

The University of Nottingham  
Nottingham Geospatial Institute

# **Digital Images for Road Surface Monitoring**

Sarhat M. Adam

Supervisors

Dr. Martin Smith

Dr. Nikolaos Kokkas

Thesis submitted to the University of Nottingham for the degree of Doctor of  
Philosophy

May 2015

This Page is Intentionally left Blank

## DEDICATION

*“This thesis is dedicated to the soul of 1100 Peshmergas, who lost their life fighting ISIS and those who continue fighting them to make Iraqi-Kurdistan a better place to live in.”*

*Source: <http://rudaw.net/NewsDetails.aspx?pageid=112615>*

*Sarhat Mustafa Adam  
15/03/2015*

## ABSTRACT

Road networks are regarded as probably the most-important infrastructure in modern day travel. Regular assessment of their condition is necessary for implementing proper maintenance and minimizing the cost. For example, early crack detection and maintenance has proved to be an effective technique of prolonging the age of roads and maintaining safe travel conditions. Although, detection of the cracks has been the aim of a number of researches in recent years, many challenges still exist. For example, crack detection in rough texture surfaces needs more attention and investigation, as rough texture can make crack detection difficult in a digital image. In addition, the pavement surface texture can change rapidly within a few metres which affects the ride quality, skid resistance, and road safety. Existing traditional surface texture measurement techniques such as the Sand Patch Test (SPT) tend to be time consuming and of variable quality. Whereas, modern methods which include Mobile Laser Scanning Sensors (MLSS) can provide accurate results but are often regarded as too expensive.

In this study an alternative method for measuring pavement surface texture will be investigated, with the aim of testing the potential of digital images for road surface monitoring and pavement evaluation. This method is based on the rapid advances in the field of image processing and image-based 3D modelling. This research is aimed to investigate the possibility of identifying road surface distresses due to cracks and potholes, as well as characterising surface texture depth. This project is split into two parts, the first part is to investigate or examine the use of digital images in video mode for identifying the cracks and potholes. In the second part, the project will concentrate on using digital still images for characterising the road surface texture in order to extract relevant safety parameters such as texture depth.

This research showed that it is possible to measure the texture depth from digital images using different cameras with comparable results to SPT. A good accuracy and high correlation with SPT were achieved even with smartphone cameras. It has also been shown that it is possible to assemble a system with cost effective tools such as camera and hand-held GPS. Then, through algorithm development, it was possible to detect important road damages such as cracks and potholes with good accuracy when compared with measured trust data.



## ACKNOWLEDGEMENTS

First of all, I would like to thank God, most Gracious, and most Merciful, for making all this possible.

Secondly, I would like to thank the KRG (Kurdistan Regional Government) for financing my study, as without their help this would have not been possible. Many thanks for the University of Nottingham and NGI for their help and support.

Thirdly, I would like to thank both of my supervisors (Martin and Nick) and all other staff and friends in NGI.

Fourthly, I would like to thank Farhan, a PhD student from the Department of Civil Engineering for providing fractured concrete samples to test the author's developed method of texture depth measurement on those samples.

It is also worth to mention, that a paper titled "An Approach to Produce a GIS Database for Road Surface Monitoring" has been published as the outcome of this PhD study. Another paper titled "Flexural characteristic of rubberized cement-stabilized crushed aggregate from pavement structure" is in the progression for publication with two other papers targeted to ISPRS titled "An approach for Road surface depth measurement using image based 3D modelling" and "A low cost system for pothole and crack detection".

Finally, special thanks go to my wife and my family for being patient and supportive during my study time.

# TABLE OF CONTENTS

DEDICATION	i
ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	x
LIST OF TABLES	xviii
LIST OF ABBREVIATIONS	xx
<b>CHAPTER ONE: INTRODUCTION</b>	<b>1</b>
1.1 INTRODUCTION	1
1.2 AIMS AND OBJECTIVES	3
1.2.1 <i>Aims</i>	3
1.2.2 <i>Objectives</i>	3
1.3 METHODOLOGY	4
1.4 CONTRIBUTION TO KNOWLEDGE	5
1.5 THESIS OUTLINE	6
<b>CHAPTER TWO: BACKGROUND AND THEORY</b>	<b>8</b>
2.1 INTRODUCTION	8
2.2 LITERATURE REVIEW	8
2.3 ROAD PAVEMENT DESIGN, DISTRESSES AND MONITORING	14
2.3.1 <i>Flexible pavement design</i>	15
2.3.2 <i>Modes of failure in flexible pavement</i>	16
2.3.3 <i>Distresses on the flexible pavement</i>	16
2.3.3.1 Cracking distresses	18
2.3.3.2 Patching and Potholes	19
2.3.4 <i>Flexible pavement evaluation and data collection</i>	20
2.3.4.1 Non-destructive testing	20
2.3.4.2 Destructive testing	21
2.3.4.3 Visual condition surveys	21
2.3.4.4 Automated condition survey	23
2.3.4.4.1 SCANNER road condition indicator	24
2.3.5 <i>Road surface texture roughness</i>	25
2.3.5.1 Road surface texture and skid resistance	26
2.3.5.2 Methods of characterising roughness of pavement texture	27

2.3.5.2.1	Micro texture measurement techniques	27
2.3.5.2.2	Measurement of frictional forces	28
2.3.5.2.3	Macro texture measurement techniques	28
2.4	2D IMAGE ANALYSIS	30
2.4.1	<i>Image Segmentation</i>	30
2.4.2	<i>Image Filtering</i>	33
2.4.3	<i>Image de-blurring</i>	35
2.5	POINT CLOUD GENERATION TECHNIQUES	36
2.5.1	<i>Total stations</i>	37
2.5.2	<i>Global navigation satellite systems (GNSS)</i>	38
2.5.3	<i>Laser Scanners</i>	38
2.5.4	<i>Using Structure from Motion (SfM) in Computer Vision (CV)</i>	41
2.5.5	<i>Using a Photogrammetric-based approach</i>	44
2.5.5.1	Image Acquisition	48
2.5.5.1.1	Platforms	48
2.5.5.1.2	Camera	49
2.5.6	<i>Comparison between photogrammetry and Computer vision (CV) based 3D surface modelling</i>	51
2.5.6.1	Camera Calibration	52
2.5.6.2	Feature Matching	54
2.5.6.3	Outlier Rejection in Feature matching	60
2.6	SUMMARY	61
	<b>CHAPTER THREE: NEW APPROACH FOR TEXTURE DEPTH MEASUREMENT</b>	<b>63</b>
3.1	INTRODUCTION	64
3.2	NEW APPROACH-DESCRIPTION	65
3.3	AIMS AND OBJECTIVES OF THE TRIALS	68
3.4	METHODOLOGY OF THE TRIALS	68
3.5	DATA SETS AND TEST SITES	69
3.6	MATERIAL AND MODIFIED METHODS	72
3.6.1	<i>Cameras used for the trials</i>	72
3.6.2	<i>Sand patch test apparatus</i>	73
3.6.3	<i>Total station for absolute orientation</i>	74
3.6.4	<i>Sand patch test modification</i>	75
3.6.5	<i>Texture depth calculation criteria</i>	76
3.6.6	<i>Sand patch test area calculation</i>	78
3.7	PROCEDURES	78
3.7.1	<i>For testing image quality or image sharpness (objective one)</i>	78
3.7.1.1	Preparation	79

3.7.1.2	Image data capturing	79
3.7.1.3	Image data analysis	80
3.7.2	<i>Preparation for all other trials (all objectives)</i>	83
3.7.2.1	Image data capturing	84
3.7.2.1.1	Stereo pair image data collection	84
3.7.2.1.2	Multiple image data collection	84
3.7.2.2	Image processing and 3D model preparation	85
3.7.2.2.1	3D model reconstruction	85
3.7.2.2.2	3D model transformation	87
3.7.2.2.3	3D model sample preparation or cropping	91
3.7.2.2.4	3D model sample depth calculation	92
3.7.2.2.5	Sand patch surface reconstruction	95
3.7.2.3	Data analysis approaches	95
3.7.2.3.1	Average texture depth calculation	96
3.7.2.3.2	Comparison of 3D points or model (surface-to-surface or points-to-points)	96
3.7.2.3.3	AGISoft summary report	98
3.8	RESULTS AND ANALYSIS	99
3.8.1	<i>Trial 1: Investigate the effect of lenses and cameras on image quality or image sharpens</i>	99
3.8.2	<i>Trial 2: Investigate the accuracy of Stereo and multiple images</i>	102
3.8.2.1	Effect of intersection angle (IA) or base to height ratio and object to camera distance on Stereo images	102
3.8.2.1.1	Intersection angle (IA) effect	102
3.8.2.1.2	The effect of the distance from camera to object	115
3.8.2.1.3	Outcomes from intersection angle (IA) and camera to object distance investigations	127
3.8.2.2	Investigation of depth accuracy from multiple images modelling	128
3.8.3	<i>Trial 3: Comparison between depth calculated from 3D models and Sand Patch Test</i>	132
3.8.3.1	Depth from Sand patch test as Mean Texture Depth (MTD)	133
3.8.3.2	Depth from digital images (stereo pairs)	134
3.8.3.3	Depth from digital images (Multiple Images)	137
3.8.4	<i>Trial 4: Testing other materials for feature extraction, matching and RMSE</i>	141
3.8.5	<i>Trial 5: Investigating the method on concrete texture</i>	144
3.8.6	<i>Trial 6: Using the proposed method of depth calculation as a dependable measuring technique for fractured concrete</i>	153
3.9	SUMMARY AND CONCLUSIONS	156
<b>CHAPTER FOUR: A DEVELOPED METHOD OF CRACK DETECTION ON ROUGH TEXTURE</b>		<b>159</b>
4.1	INTRODUCTION	159

4.2	DESIGN OF A CRACK DETECTION METHOD	160
4.3	AIMS AND OBJECTIVES OF TRIALS	162
4.4	METHODOLOGY OF TRIALS	162
4.5	DATA SETS AND TEST SITE DESCRIPTION	163
4.6	PROBLEM DEFINITION AND TRIAL WITH CRACK DETECTION	167
4.7	MATERIALS AND METHODS	170
4.8	SEMI-AUTOMATIC PROCEDURES FOR CRACK DETECTION	172
4.8.1	<i>Image preparation</i>	173
4.8.1.1	Frame extraction with motion estimation	173
4.8.1.2	Frame Geo-tagging	175
4.8.2	<i>Exclusion of problem free images based on statistics</i>	177
4.8.3	<i>Image pre-processing</i>	180
4.8.3.1	Shadow detection and removal	180
4.8.3.2	Morphological erosion and image filtration	182
4.8.3.3	Modelling the roughness of the texture using 2D image analysis	184
4.8.4	<i>Image processing or segmentation (edge detection)</i>	185
4.8.5	<i>Image post-processing</i>	186
4.9	RESULTS AND ANALYSIS	187
4.9.1	<i>The Author's Developed Crack detection algorithm results evaluation</i>	187
4.9.1.1	Motion estimation for frame extraction	187
4.9.1.2	Results of frames Geotagging	190
4.9.1.3	Statistical testing for problems identification	194
4.9.1.4	Testing road roughness modelling using images	198
4.9.2	<i>Testing various sources of images</i>	202
4.9.3	<i>Testing the author's developed crack detection algorithm on video images captured from a mobile van</i>	209
4.10	FROM SEMI TOWARDS FULL AUTOMATION OF CRACK DETECTION	215
4.11	SUMMARY AND CONCLUSIONS	219
	<b>CHAPTER FIVE: POTHOLE DETECTION FROM VIDEO IMAGES</b>	<b>221</b>
5.1	INTRODUCTION	221
5.2	DESIGN OF A POTHOLE DETECTION METHOD	222
5.3	AIMS AND OBJECTIVES OF TRIALS	224
5.4	METHODOLOGY OF TRIALS	224
5.5	DATA SETS AND TEST SITES	225
5.6	PROCEDURES	226
5.6.1	<i>Methodology of testing matching algorithms (Trial 1)</i>	226
5.6.2	<i>Pothole and bump detection (trial 2, 3 &amp; 4)</i>	229
5.6.2.1	Point clouds generation	230

5.6.2.2	Algorithm designed for surface anomalies detection	233
5.7	RESULTS AND ANALYSIS	234
5.7.1	<i>Trial 1: Comparison of various matching algorithms</i>	234
5.7.1.1	Implementing the test without outlier detection on both SIFT and SURF	235
5.7.1.2	Implementing the test with outlier detection on both SIFT and SURF	237
5.7.2	<i>Trial 2: Comparison between open source and commercial software for 3D point cloud generation and a bump detection</i>	240
5.7.3	<i>Trial 3: Investigating the use of 3D models for small object detection</i>	243
5.7.4	<i>Trial 4: Investigating the use of video images for pothole detection</i>	246
5.7.4.1	Using only two images	246
5.7.4.2	Using multiple images	255
5.8	FROM SEMI TOWARDS FULL AUTOMATION OF POTHOLE DETECTION	256
5.9	SUMMARY AND CONCLUSIONS	259
<b>CHAPTER SIX: CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK</b>		<b>261</b>
6.1	INVESTIGATING HIGHWAY MAINTENANCE PROBLEMS AND HIGHLIGHTING LIMITATIONS OF MONITORING TECHNIQUES	261
6.2	EXAMINE THE POTENTIAL OF DIGITAL IMAGES FROM A VARIETY OF CAMERAS FOR TEXTURE DEPTH MEASUREMENT	263
6.3	INVESTIGATE THE POTENTIAL OF DETECTING THE CRACKS ON THE ROAD SURFACE FROM DIGITAL SLR EXTRACTED VIDEO FRAMES	264
6.4	INVESTIGATE THE POTENTIAL OF DETECTING POTHOLES ON THE ROAD SURFACE USING THE EXTRACTED FRAMES FROM VIDEO DSLR IMAGES	266
6.5	RECOMMENDATIONS FOR FUTURE WORK	266
<b>REFERENCES</b>		<b>269</b>
<b>APPENDICES</b>		<b>280</b>
APPENDIX A: TRANSFORMATION PARAMETERS FROM CC ALIGNING TOOL		280
APPENDIX B: SAMPLES OF DATA SET IMAGES COLLECTED USING NIKON D100		281
APPENDIX C: DIFFERENT SIZE OF STRUCTURING ELEMENTS EFFECT		282
APPENDIX D: SAMPLES FOR DIFFERENT EDGE OPERATIONS TESTING		283
APPENDIX E: MINIMUM AVERAGE COST PATH INTENSITY		284
APPENDIX F: FRAME GEO-TAGGING RESULTS		285
F01:MVI-4513		285
F02:MVI-4514-A		286
F03:MVI-4517		288
F04:MVI-4522		289
F05:MVI-4523-A		290

<i>F05:MVI-4523-B</i>	291
<i>F06:MVI-4539</i>	292
<i>F07:MVI-4540</i>	293
<i>F08: MVI-4541</i>	294
<i>F09: MVI-4542</i>	295
APPENDIX G: RESULTS OF TRIANGULATION	296
<i>G1: Estimated accuracy results of Australis</i>	296
<i>G2: Estimated accuracy results of PMS</i>	297
<i>G3: Estimated accuracy results of LPS (pixels)</i>	297
APPENDIX H: RESULTS OF ACCURACIES FOR BUMP TEST	298
<i>H1: LPS triangulation results</i>	298
<i>H2: GCPs accuracy and quality from Photomodeler</i>	298
<i>H3: Output report from AGISoft StereoScan</i>	299
APPENDIX I: CANON 5D CAMERA CALIBRATION REPORT	300

## LIST OF FIGURES

Figure 2.1: Komatsu Survey Vehicle (Wang, 2000) .....	9
Figure 2.2: Flexible and rigid pavement load distribution (OSU, 2012) .....	14
Figure 2.3: Life cycle of flexible pavement and the cost of maintenance (Alex, 2011) .....	15
Figure 2.4: Typical cross section of a flexible pavement (Mathew, 2009) .....	16
Figure 2.5: Type of distresses on flexible pavement (Miller and Bellinger, 2003) .....	18
Figure 2.6: Distresses types on the asphalt road pavement, images adapted from (Asphalt Institute, 2009; Pavement Interactive, 2012) .....	19
Figure 2.7: High severity patch and pothole (Miller and Bellinger, 2003) .....	20
Figure 2.8: Individual present serviceability rating form (Pavement Interactive, 2012) .....	22
Figure 2.9: Samples of parameters and threshold used in RCI .....	24
Figure 2.10: Relationship between texture and characteristics of the pavement surface (World Road Association, 2012) .....	26
Figure 2.11: Macro and Micro level properties of aggregates .....	26
Figure 2.12: Tyre/Pavement contact area in rough and smooth texture (Nutz and Hoffmann, 2012) .....	27
Figure 2.13: Convolution masks of Sobel edge operation .....	32
Figure 2.14: Diagram explain the K-mean clustering process .....	32
Figure 2.15: Demonstration of the effect of the neutral density filter (Wiki, 2014) .....	33
Figure 2.16: Effect of mean filtration, original image, 3x3 filtering, 5x5 filtering .....	33
Figure 2.17: Example shows the effect of median filtering on salts and pepper noise (original image left and median filtered results right) .....	34
Figure 2.18: An example illustrating the median filtering with 3x3 neighbourhood (USF, 2014) .....	34
Figure 2.19: PSFs examples, original image, out of focus, and linear motion blur (ibid) .....	35
Figure 2.20: Principle of laser triangulation (left), (Bitelli et al., 2012) and sample of laser sensor (A), (Micro-Epsilon, 2014) .....	40
Figure 2.21: Circular Texture Meter (left) and FaroArm-Platinum (right) (Applied Pavement Technology, 2008; Faro, 2014) .....	41
Figure 2.22: Merging images constantly using SfM (Michot, 2010) .....	42
Figure 2.23: Reconstruction of web photos (Google & Flickr) into 3D models (Snively, 2010) .....	43
Figure 2.24: Point cloud generation procedures (Jin-Tsong et al., 2012) .....	43
Figure 2.25: Overall processing of Photosynth point clouds using possible outlier removal tool .....	44
Figure 2.26: Coplanarity condition and intersection for relative orientation (Ghosh, 2005) .....	45
Figure 2.27: Screenshot of point clouds generated from UltraMap 3.0, (UltraCAM, 2011) ..	47



Figure 2.28:Diagram showing the relationship between Field of View and focal length .....	50
Figure 2.29: Exposure Triangle (CiC, 2012) .....	50
Figure 2.30: Rotation comparison between SIFT and SURF at different angles, repeatability or frequency vs rotation (above) and the used samples (below) (Juan and Gwun, 2009) .....	55
Figure 2.31: Extracted features and their scale (SURF) .....	57
Figure 2.32: Possible Images that SIFT & SURF can Work with(Juan and Gwun, 2009) .....	59
Figure 2.33: SURF only (left), SURF with RANSAC (Right) .....	60
Figure 3.1: The workflow diagram for depth calculation from digital images.....	67
Figure 3.2: Testing lens quality of different cameras .....	69
Figure 3.3: Test study area located near Jubilee Campus.....	70
Figure 3.4: Used samples for texture depth evaluation .....	70
Figure 3.5: Various texture or material types .....	71
Figure 3.6: Concrete samples, rough texture (left) and smooth texture (right).....	71
Figure 3.7: Nine concrete fractured sample from Civil Department .....	72
Figure 3.8: Sand Patch Test Apparatus .....	74
Figure 3.9: Comparison between SMTD and MPD in depth calculation (ibid) .....	76
Figure 3.10: Sample shows the texture depth measurement using different approaches ...	77
Figure 3.11: Three different methods of sand patch area calculation .....	78
Figure 3.12: Printed Sine pattern charts for testing Image quality .....	79
Figure 3.13: ImageJ tool for scaling and plotting profiles from sine pattern chart .....	80
Figure 3.14: Exported profile (above) and calculated SFR vs lp/mm (below).....	81
Figure 3.15: CFtool for making a best fit curve for SFR data .....	81
Figure 3.16: Sample of SFR response of different cameras at F stop 4.0 .....	82
Figure 3.17: Marker details and measurement procedures .....	83
Figure 3.18: Methodology of image collection for AGISoft StereoScan .....	84
Figure 3.19: Multiple image data collection for 3D surface modelling using AGISoft PhotoScan .....	85
Figure 3.20: 3D reconstructed sample & output report (AGISoft StereoScan) .....	86
Figure 3.21: Textured 3D model & output report description (AGISoft PhotoScan).....	87
Figure 3.22: 3D model sample before and after adding material library (AGISoft StereoScan) .....	88
Figure 3.23: Matlab subroutine for markers extraction .....	88
Figure 3.24: Outcome from Matlab subroutine extraction tool.....	89
Figure 3.25: Cluster Selection tool for extracting real markers .....	89
Figure 3.26: K-mean Clustering for grouping and finding out the centroid .....	90

Figure 3.27: 3D transformation from relative to real world system .....	90
Figure 3.28: 3D model sample cropping and sand patch (Canon 5D 50mm).....	91
Figure 3.29: Plane orientation of 3D model sample .....	92
Figure 3.30: Sample of Matlab fitted plane to surface tool .....	93
Figure 3.31: Author's Developed Algorithm for calculating depth of texture .....	93
Figure 3.32: Cross section on sample of rough texture illustrate the fitted plane, upper and lower residuals .....	94
Figure 3.33: Multiple image data collection for reconstructing sand patch surface .....	95
Figure 3.34: Distance computations between reference and model surfaces .....	97
Figure 3.35: Typical result of a distance computation process with colour scale display parameters .....	97
Figure 3.36: Histograms of distance differences.....	98
Figure 3.37: Spatial Frequency Response for various cameras.....	101
Figure 3.38: Base to height ratio and intersection angle .....	102
Figure 3.39: Designed test to read IA every 10° with fixed range.....	103
Figure 3.40: Average Feature Extraction Vs Intersection Angle for both rough and smooth texture, Smartphones (above) and DSLR Cameras (below) .....	105
Figure 3.41: Ten strongest features at 70 intersection angle, smooth texture (above) and rough texture (below) .....	106
Figure 3.42: Relationship between intersection angle and number of matches for both smooth and rough textures (Smartphones above and DSLR Cameras below) .....	108
Figure 3.43: change in the target size and shape due to intersection angle .....	109
Figure 3.44: RMSe vs Intersection angle (Smooth texture above and rough texture below).....	110
Figure 3.45: Non complete 3D reconstruction due to increase of IA (50, 60 and 70°), left to right.....	111
Figure 3.46: Share and non-share features vs. different angle of intersection in rough texture.....	111
Figure 3.47: Camera optical axis vs. aggregate vertical angle.....	112
Figure 3.48: Texture depth vs intersection angle, smooth texture (above) and rough texture (below).....	113
Figure 3.49: Samples of rough textures at different IA, (10-60°) left to right.....	115
Figure 3.50: Approach to find optimal camera range at constant intersection angle .....	116
Figure 3.51: Feature point Vs camera range, smartphone cameras (above) and DSLR cameras (below).....	117
Figure 3.52: Ten strongest features extracted from 20 cm (above) and 110 cm (below) using Samsung Note camera .....	118
Figure 3.53: Strongest extracted feature at 20 cm Samsung Wave Camera (above) and Canon 5D (below), smooth (left), rough (right) .....	120

Figure 3.54: Matches vs camera to object distance for smooth texture (above) and rough (below) .....	121
Figure 3.55: RMSe vs camera to object distance (pixels unit), smooth above, rough below.....	122
Figure 3.56: RMSe vs camera to object distance (mm units), smooth texture above, rough texture below.....	123
Figure 3.57 Texture depth vs camera's range (Smooth texture above and Rough texture below) .....	125
Figure 3.58: 3D samples of rough texture using colour as a height ramp .....	126
Figure 3.59: 3D samples prepared using multiple images smooth texture (above) and rough texture (below).....	128
Figure 3.60: Cloud to Cloud distances and their statistical outcome between Nikon D200 and Canon 5D-50mm for the smooth texture (colour visualization on top) .....	129
Figure 3.61: Cloud to cloud comparison of all samples with reference to C50, Smooth texture (above) and rough texture (below) .....	131
Figure 3.62: Texture depth variation on the used samples .....	132
Figure 3.63: Tests vs error from three different method of depth calculation .....	134
Figure 3.64: Errors between depth from Stereo images and average of sand patch texture depth.....	135
Figure 3.65: Correlation between depth from Stereo and Sand .....	136
Figure 3.66: Errors between depth from multiple images and sand patch .....	138
Figure 3.67: Correlation between depth from multiple images and sand .....	139
Figure 3.68: comparison of focal lengths between both stereo and multiple images .....	140
Figure 3.69: Various materials vs reconstruction parameters.....	142
Figure 3.70: Matching on Grass (Wave left and D200 right) .....	143
Figure 3.71: 3D reconstructed surface for various materials Wave (above) and D200 (below) .....	143
Figure 3.72: 3D samples from 5 cameras, smooth texture (above) and rough texture (below) .....	144
Figure 3.73: Colour visualisation of differences on 3D models, smooth (above) and rough (below) .....	145
Figure 3.74: Colour visualization of differences on concrete models, smooth (above) and rough (below).....	146
Figure 3.75: Horizontal profiles through centre of 10cm models, smooth texture (above) and rough (below).....	147
Figure 3.76: Differences between all samples with reference to C50, smooth texture (above) and rough (below), RMSE ( $\mu\text{m}$ ) .....	148
Figure 3.77: Sand Patch recreated surface (rough concrete texture) .....	150

Figure 3.78: Sample of cloud-to-cloud distance calculation (C50 with reference to sand smooth texture concrete) .....	150
Figure 3.79: Visual height ramp of the smooth texture with the camera to object distance, Nikon D200 camera (above) and Samsung Wave (below) .....	152
Figure 3.80: Fractured concrete .....	153
sample with markers .....	153
Figure 3.81: The reconstructed concrete fractured samples .....	154
Figure 3.82: visualization of differences of the two sets .....	155
Figure 3.83: Histograms of differences between two sets with Gaussian distribution fit results .....	155
Figure 4.1: Workflow diagram for developed method of crack detection .....	161
Figure 4.2: Path of collected data set for crack detection algorithm development.....	163
Figure 4.3: Samples of crack images from variety of sources .....	164
Figure 4.4: System for image data collection on mobile van .....	165
Figure 4.5: the path of all collected data from Radford area on mobile van.....	166
Figure 4.6: Some available texture in the road pavement.....	167
Figure 4.7: Intensity variations of crack pixel and .....	168
Its neighbours.....	168
Figure 4.8: Intensity profiles and curve fitting through image.....	168
Figure 4.9: False positives (identifying non-crack pixel as a crack) .....	169
x= pixel location along profile, y=intensities .....	169
Figure 4.10: Results from thresholding method based on profile intensities original image (left) and results (right) .....	169
Figure 4.11: Concept of crack detection algorithm.....	172
Figure 4.12: Frame extraction and motion estimation flow chart .....	173
Figure 4.13: Sample of motion estimation through transformation matrix calculation.....	174
Figure 4.14: The system concept of recording time of starting camera .....	175
Figure 4.15: Extracted frames from video file (MVI_4514.MOV) .....	175
Figure 4.16: sample of GPS track and interpolation points.....	176
Figure 4.17: Histogram and fitted Gaussian curve for sample of images .....	178
Figure 4.18: Flowchart describing the methodology of Gaussian best fit and image histograms.....	179
Figure 4.19: Flow chart shows the use of K-mean clustering for shadow detection .....	180
Figure 4.20: Results from Shadow removal algorithm and post processed shadow core mask .....	181
Figure 4.21: Original image (left), eroded and reconstructed image (right).....	182

Figure 4.22: Percentage of detected edges (cracks and non-cracks) vs size of the mask for a sample of image showed in figure 4.21 .....	183
Figure 4.23: Designed filters for different texture roughness .....	183
Figure 4.24: Texture modelling based on small object repetition .....	184
Figure 4.25: Sample of disconnected objects and the minimum path .....	186
Figure 4.26: Motion estimation between samples of frames (1494 to 1501) .....	187
Figure 4.27: Motion estimation between samples of frames (1501 to 1507) .....	188
Figure 4.28: Frames sharing area at sharp turn locations. ....	189
Figure 4.29: Road junction points used to check the accuracy of Geotagging .....	190
Figure 4.30: MVI-4512 video turn points frame and geo-tagging differences .....	191
Figure 4.31: Drift of GPS path due to the start of the GPS receiver .....	193
Figure 4.32: Sample of images for testing statistically the problem identification .....	195
Figure 4.33: Sample at range 130cm cropped to be same size of 20cm (Test01) .....	198
Figure 4.34: Results of object repetition vs measured depth at different ranges.....	199
Figure 4.35: Samples of linear correlation between objects repetition and texture depth	200
Figure 4.36: Sample of texture (test04 left and test06 right) .....	201
Figure 4.37: Collection of images for evaluation of crack detection algorithm .....	202
Figure 4.38: Samples of images after pre-processing and processing stages with applying the five filtrations (very rough left to very smooth right), green boxes are correct filters .....	203
Figure 4.39: Image processing using only edge operation without pre-processing stage ..	204
Figure 4.40: Crack detection in shadow existence samples .....	205
Figure 4.41: Final results of crack detection on samples extracted from published papers	206
Figure 4.42: Final results of crack detection for LRIS imaging samples .....	207
Figure 4.43: Final crack detection results on Google downloaded images .....	207
Figure 4.44: Crack detection results for rough and moderate texture image .....	208
Figure 4.45: Relationship between Shutter speed and Traffic speed (2m height) .....	210
Figure 4.46: Relationship between Shutter speed and Traffic speed (1m height) .....	210
Figure: 4.47 Video extracted frames with crack detection results .....	211
Figure 4.47: video extracted frames with crack detection results, from left to right .....	211
Figure 4.48: Undetectable cracks due to the width of crack (original image left and crack detection results on right, detectable are in green and undetectable on red) .....	212
Figure 4.49: process of identifying the shadow and removing its effect on crack detection algorithm.....	213
Figure 4.50: Manhole cover false positives removed using value intensities .....	213

Figure 4.51: Sample of extracted frames (shutter 1/1000 at about 35 km/hr) original images (red box), processing results (blue box) and result from four views merged in CAD tool (green box).....	214
Figure 4.52: Main key stages in crack detection (toward fully automation) .....	218
Figure 5.1: Workflow diagram for developed method of pothole detection .....	223
Figure 5.2: Trolley for 100% overlap multiple image acquisition.....	225
Figure 5.3: Sample of images acquired at 100% lateral overlap using the trolley .....	226
Figure 5.4: GCPS distribution and scale bar used for EO calculation .....	227
Figure 5.5: New project in LPS showing IOPs and EOPs, first pair .....	228
Figure 5.6: Tool for importing tie points from SIFT and SURF into LPS .....	229
Figure 5.7: Microsoft Photosynth desktop tool for image uploading .....	230
Figure 5.8: Photomodeler calibration sheet .....	231
Figure 5.9: Types of automation coded targets in Photomodeler .....	232
Figure 5.10: Procedures for detecting the Potholes or bumps on the road surface .....	233
Figure 5.11: First pair of tested images above and second below.....	234
236	
Figure 5.12: Aerial Triangulation Results, three algorithms, 1 <sup>st</sup> pair (above) & 2 <sup>nd</sup> (below)	236
Figure 5.13: Effect of Outlier Detection Algorithms on SIFT, 1 <sup>st</sup> pair (above) & 2 <sup>nd</sup> (below)	237
Figure 5.14: Effect of Outlier Detection Algorithms on SURF, 1 <sup>st</sup> pair (above) & 2 <sup>nd</sup> (below)	238
Figure 5.15: SIFT and SURF with Outlier removed vs LPS, 1 <sup>st</sup> pair (above) & 2 <sup>nd</sup> (below)....	239
Figure 5.16: Four shots showing the bump on the street.....	240
Figure 5.17: Visual analysis of the achieved 3D points clouds results .....	241
Figure 5.18: Fitted Surface (above) and residuals (down) .....	242
Figure 5.19: Bump detected using exclusion criteria at different sigma's.....	242
Figure 5.20: Location of Cameras and used images for analysis.....	243
Figure 5.21: Visual analysis of the achieved 3D points clouds results .....	244
Figure 5.22: 3 sigma, 2 sigma, and 1 sigma exclusion criteria from (top to bottom) .....	245
Figure 5.23: Extracted frames from video files with pothole existence .....	246
Figure 5.24: samples used for testing the effect of base to height ratio on the results.....	247
Figure 5.25: 3D models generated from processing 1 <sup>st</sup> frame with the next following frames (1 <sup>st</sup> with 2 <sup>nd</sup> on top and 1 <sup>st</sup> with 8 <sup>th</sup> on bottom) .....	248
Figure 5.26: Height ramp colour visualization on 3D models frames (1 <sup>st</sup> with 3 <sup>rd</sup> on top and 1 <sup>st</sup> with 8 <sup>th</sup> on bottom).....	249
Figure 5.27: Height ramp colour visualization on 3D models frames (1 <sup>st</sup> with 2 <sup>nd</sup> on top and 1 <sup>st</sup> with 8 <sup>th</sup> on bottom).....	250
Figure 5.28: Matching and RMSe vs base line or translation between two frames .....	251

Figure 5.29: Colour visualization of ramp height on 3D models generated using two frames only .....	253
Figure 5.30: 3 sigma criteria for pothole detection on all samples using only two frames. ....	254
Figure 5.31: Height ramp visualization on 3D models using multiple images .....	255
Figure 5.32: potholes detected at 3 sigmas away from mean plane.....	256
Figure 5.33: Main key stages in pothole detection (toward fully automation).....	258

## LIST OF TABLES

Table 2.1: Example of automated systems for pavement evaluation .....	23
Table 2.2: Comparison between LIDAR system and UltrCAM Eagle, (UltraCAM, 2011) .....	48
Table 2.3: Image resolution and theoretical visible crack width(Wang and Smadi, 2011) ....	51
Table 2.4: Camera Calibration in Photogrammetry and CV (Remondino et al., 2012) .....	53
Table 2.5: SIFT and SURF general performance levels (Juan and Gwun, 2009) .....	55
Table 2.6: SIFT and SURF workflow (Lowe, 2007; Bay et al., 2008) .....	56
Table 3.1: Cameras specifications and calibration results .....	73
Table 3.2 : Texture depth vs volume of sand (ASTM, 2004) .....	75
Table 3.3 : Texture depth vs volume of sand (modified) .....	75
Table 3.4: Sample of results from texture depth calculation algorithm and two techniques (SMTD and MPD) (mm) .....	94
Table 3.5: Equivalent values between B/H ratio and intersection angle .....	102
Table 3.6: Difference between the required intersection angle at field and calculated from Australis, each camera smooth (above) and rough (below) .....	104
Table 3.7: Standard deviation of different cameras at different intersection angles calculated using mean sand patch depth as true value .....	113
Table 3.8: Ground sample distance vs camera to object distance of variety of cameras....	119
Table 3.9: Standard deviation of different cameras at different ranges calculated using mean sand patch depth as true value .....	125
Table 3.10: Depth results and error of different cameras ( $\mu\text{m}$ ).....	128
Table 3.11: cloud to cloud distance calculation results and statistics .....	130
Table 3.12: Depth calculated from three different methods ( $\mu\text{m}$ ) .....	133
Table 3.13: cloud-to-cloud distance comparison results .....	147
Table 3.14: RMSE of 4GCPs from transformation between local camera and total station coordinates.....	149
Table 3.15: Depth from 3D samples vs depth from sand patch for concrete samples ( $\mu\text{m}$ )	149
Table 3.16: cloud-to-cloud distance calculation between sand recreated surfaces and 3D samples.....	151
Table 3.17: RMSE of GCPs out from transformation process in CC .....	154
Table 4.1: Video files detail for the data collected.....	166
Table 4.2: Sample of GPS positions (WGS84) transformed to OSGB36 .....	176
Table 4.3: sample of time processing for a one frame of image.....	177
Table 4.4: Results of Crack and non-Crack Edges.....	185
Table 4.5: Number of included and excluded frames using motion estimation .....	189
Table 4.6: Accuracy of GPS frames at turning points .....	192



Table 4.7: Results of statistical testing of comparing data and fitted curve .....	196
Table 4.8: Results of problem identification for video files .....	197
Table 5.1: Average Camera Calibration parameters .....	227
Table 5.2: EO parameters using different software .....	228
Table 5.3: Mean Result of EOP using three programs .....	235

## LIST OF ABBREVIATIONS

2D	.....	2-Dimensional
3D	.....	3-Dimensional
AASHO	.....	American Association of State Highway and Transportation Officials
AP	.....	Additional parameters
ARAN	.....	Automatic Road Analyzer (Fugro Roadware's)
ASTM	.....	American Society for Testing and Materials
AT	.....	Aerial Triangulation
BPT	.....	British Pendulum Tester
C28	.....	Canon 5D, 28mm Canon lens
C28S	.....	Canon 5D, 28mm Sigma lens
C50	.....	Canon 5D, 50mm Canon lens
CAA	.....	Civil Aviation Authority
CC	.....	Cloud Compare
CCD	.....	Charged Couple Device
CCPs	.....	Check Control Points
CMOS	.....	Complementary Metal–Oxide Semiconductor
CSIRO	.....	Commonwealth Scientific & Industrial Research Organisation
CT Meter	.....	Circular Texture Meter
CV	.....	Computer Vision
D100	.....	Nikon D100, 28mm Nikon lens
D200	.....	Nikon D200, 28mm Nikon lens
D300	.....	Nikon D300, 28mm Nikon lens
DAA	.....	Developed Algorithm by Author
DG	.....	Differential Gradient
DGPS	.....	Differential Global Positioning System
DSLR	.....	Digital Single-Lens Reflex
DTM	.....	Digital Terrain Model
EDM	.....	Electronic Distance Measurement
EOPs	.....	Exterior Orientation Parameters
ETC	.....	Earth Technology Corporation
FoV	.....	Field of View
GCPs	.....	Ground Control Points
GPS	.....	Global Positioning System and
GNSS	.....	Global Navigation Satellite System
GoF	.....	Goodness of Fit
GPR	.....	Ground Penetrating Radar
GSD	.....	Ground Sample Distance
HARRIS	.....	Highways Agency Road Research Information System
HMA	.....	Hot Mixture Asphalt
IA	.....	Intersection Angle
IMU	.....	Inertial Measurement Unit
IOPs	.....	Interior Orientation Parameters

## List of Abbreviations

---

ISO	.....	International Standards Organization
KRG	.....	Kurdistan Regional Government
LAR	.....	Least Absolute Residuals
LCD	.....	Liquid Crystal Display
LIDAR	.....	Laser Imaging Detection and Ranging
LMeds	.....	Least Median of Squares
LoG	.....	Laplace of Gaussian
LOWESS	.....	Locally Weighted Scatterplot Smoothing
LP/mm	.....	Line Pairs per Millimetres
LPS	.....	Leica Photogrammetric Suite
LRIS	.....	Laser Road Imaging systems
LTPP	.....	Long-Term Pavement Performance
MLSS	.....	Mobile Laser Scanning Sensors
MP	.....	MegaPixels
MPD	.....	Mean Profile Depth
MSAC	.....	M-estimator SAmple Consensus
MTD	.....	Mean Texture Depth
MTF	.....	Modulation Transfer Function
NGB	.....	Nottingham Geospatial Building
NGI	.....	Nottingham Geospatial Institute
Note	.....	Samsung Galaxy Note GT-N7000
OBJ	.....	OBJect (Wavefront)
OSGB36	.....	Ordnance Survey Great Britain 1936
OR	.....	Objects Repetition
PAVUE	.....	Automated Pavement Road Analyzer
PCES	.....	Pavement Condition Evaluation Services
PCI	.....	Pavement Condition Index
PhD	.....	Doctor of Philosophy
PMS	.....	Photomodeler Scanner
PLY	.....	PoLYgon File Format
PSF	.....	Point Spread Function
PSI	.....	Present Serviceability Index
PSR	.....	Present Serviceability Rating
PSV	.....	Polished Stone Value
RANSAC	.....	RANdom SAmple Consensus
RCI	.....	Road Condition Index
RGB	.....	Red, Green, and Blue
RMSE	.....	Root Mean Squared error
SCANNER	.....	Surface Condition Assessment for the National Network of Roads
SCRIM	.....	Sideway force Coefficient Routine Investigation Machine
SfM	.....	Structure from Motion
SFR	.....	Spatial Frequency Response
SIFT	.....	Scale Invariant Feature Transform
SMTD	.....	Sensor Measured Texture Depth

## List of Abbreviations

SNR	.....	Signal-to-Noise Ratio
SPT	.....	Sand Patch Test
SSE	.....	Sum of Squares due to Error
SURF	.....	Speed Up Robust Feature
TM	.....	Template Matching
TRACS	.....	TRAffic Speed Condition Surveys
TRL	.....	Transport Research Laboratory
UAV	.....	Unmanned Aerial Vehicle
UKPMS	.....	United Kingdom Pavement Management System
US	.....	United State
USACE	.....	United States Army Corps of Engineers
VRML	.....	Virtual Reality Model Language
Wave	.....	Samsung Galaxy Wave GT-S8500
WRA	.....	World Road Association

## **CHAPTER ONE: INTRODUCTION**

### **1.1 Introduction**

The road network is a vital part of the infrastructure for modern day living and subject to distresses after construction due to various factors such as aging, environment conditions such as temperature and frost action, loading and normal wear. The evaluation of road quality is an important task in many areas where harsh weather conditions are present such as in mountainous areas. The accurate assessment of road surface distress data is required as an essential input to any decision making process concerning the road surface management policy. This would help reduce the costs of maintenance and stop the distresses development.

According to WarrantyDirect (2012) , road surface anomalies are estimated to cause many mechanical failures in UK's roads and motorists face an estimated cost of £320 million every year. Furthermore, the United States (US) association of state highway and transportation officials reported that only half of major US roads are in good condition because of less effective monitoring. Also, there is damage claims raised against highway bodies and insurance companies worldwide. For example, the state of Michigan in 2005 faced more than 7,500 cases for pothole-related damage claims. This is in addition to more than 500,000 pothole-related claims annually received by insurance companies in the same state. Therefore, periodic road assessments and collection of up-to-date information about the road surface condition is considered to be the most effective way to conserve high road standards at minimum possible cost (Furness et al., 2007; Eriksson et al., 2008).

There are various methods and techniques adopted and available for measuring and evaluating the distress of the road surface. All the inspection that was available before 1980 was achieved manually creating long processing surveys. The importance of monitoring was addressed by many road authorities that led to develop effective methods of road health monitoring. From 1980 and onwards, the monitoring could be automated with the advancement in technology. Techniques for detecting the condition and the performance of roads are continuously being developed over time with the advent of new methods and improvements to the established equipment. Significant progress was made in the field and new approaches were proposed, many of which were designed based on slow speed Ground Penetrating Radar (GPR), or on the laser scanning detection system. Other approaches included the use of mobile sensors or accelerometers for detecting and

reporting road surface anomalies based on integrated sensor equipped vehicles (Payne, 1992; Forest and Utsi, 2004; Eriksson et al., 2008).

From all the developed methods, digital images remain a dependable source and have been widely used in road surface monitoring. Advances in digital image technology created a chance to overcome some of the problems related to labour-intensive processes. The use of image techniques seems to be more promising and has advantages over other invasive methods such as laser based systems or Ground Penetrating Radar (GPR) approach. The acquisition technique using imagery is more cost effective, easier, more dense (each millimetre), and more precise in measuring the defect (Sylvie and Jean-Marc, 2011).

There are many image acquisition techniques available nowadays. By comparing with other techniques such as satellite imagery or traditional aerial photography, the use of images from a consumer camera on a different platform such as a UAV (Unmanned Aerial Vehicle) or terrestrial based mobile system can be cost effective. On the one hand, satellite image is not preferred due to the cost and due to limited spatial resolution of the image. On the other hand, traditional aerial imagery could be an alternative choice, but the limited manoeuvrability of the platform to acquire the qualitative image data and the associated high costs are possible limitations. In contrast, UAVs and other mobile based systems are extremely flexible, low cost in collecting image data, faster and safer. Moreover, UAVs and mobile based systems are able to operate quite close to the scene and acquire images with few centimetre resolutions which provide details sufficient for identification and extraction of most distresses parameters, namely cracks and potholes (Zhang, 2008a).

Although there has been substantial progress achieved in road monitoring techniques during recent years, there are still many challenges remaining in automatically generating other types of distress data, such as potholes or rutting. Despite the fact that other distress types can also contribute considerably to the pavement surface condition, the primary interest of all previous researchers was focused on the pavement surface cracking only without considering the texture roughness. Various types of distresses will be considered in this report through using multiple images and producing 3D surface models.

Not all the distresses developed on road surfaces are cracks; another important parameter related to safety of the road is skid resistance. It is related to how rough the

texture is and how much friction is required to keep the car stopping safely while moving at high speeds. Skid resistance is related to the surface texture depth in high speed and in wet condition. Although there is a large amount of research work involved in skid resistance, there is less work dealing with the surface texture measurement using digital images.

This project aims to investigate the potential of using only a digital SLR camera and smartphone cameras as a tool for evaluating some problems in road surfaces. Roughness measurement, for example, is an important parameter which will be investigated using texture analysis techniques and 3D models as an easy alternative to the traditional sand patch test. Cracks are also considered as the most-frequent distresses that if detected and maintained can save some of the expected cost. Another important problem related to weather conditions is potholes. Using only digital images for road surface monitoring is challenging but is also important for some developing countries that cannot afford the high cost associated with special designed commercially available systems, with estimated costs up to \$100k (Pavemetrics, 2012; Engineering, 2013; Fugro Roadware, 2013).

## **1.2 Aims and Objectives**

### **1.2.1 Aims**

The aims of this research are to investigate the potential of using digital images for roughness (texture depth measurement) and detecting important road damages such as cracks and potholes using hand held and mobile vehicle devices.

### **1.2.2 Objectives**

To meet the above aims, the following objectives have been set:

1. Investigate highway maintenance problems by identifying the issues and present challenges and highlighting limitations in existing highway surveying maintenance and monitoring techniques.
2. Examine the potential of digital images from a variety of cameras, smartphone cameras and Digital Single-Lens Reflex (DSLR), for roughness or texture depth measurement on concrete and asphalt pavement surfaces through:
  - Use of different cameras with different lenses and comparison of the results.

- Investigating the effect of intersection angle and the distance of the camera from the surface on the results.
  - Comparing the results of texture depth measured from both automated 3D models and the traditional Sand Patch Test (SPT).
  - Use of the proposed methodology on other materials such as concrete texture which is not part of a road surface.
3. Investigate the potential of detecting the cracks on the road surface from digital SLR extracted video frames through:
- Algorithm development of crack detection based on the edge operator detection.
  - Time management of algorithm by applying exclusion criteria or decision making on the extracted frames based on statistical testing and Gaussian fit to exclude free of problem images or frames.
  - Examining the algorithm on different image sources (Web, Extracted from published papers, low resolution, and shadow existence) and with various texture roughness.
  - Examining and evaluating the algorithm for crack detection on video extracted images from mobile van.
4. Investigate the potential of detecting potholes on road surfaces using the extracted frames from video DSLR images through:
- Investigating different sources of 3D modelling and point cloud generation using open source and commercial software.
  - Investigating the best available digital matching algorithms.
  - Investigating, in terms of accuracy the open source and commercial 3D modelling software.
  - Testing the mobile vehicle based imagery system for potholes detection.

## 1.3 Methodology

In order to satisfy the above objectives, the following methodology was considered to be necessary for performing the research:



1. The first step in the methodology is to understand the pavement health, remaining life time, distress types, and what are the possible causes that could likely affect the pavement. This can be done through extensive reviewing of the distress types and classifying them according to the severity and frequency. To explore both the available methods and analysis techniques that perform the monitoring and maintenance of the highways, this can be done through literature reviewing and evaluating the previous methods.
2. To investigate the potential of small format images,
  - a- A set of practical trials were designed for testing the existing photogrammetric software and open source tools (examples are Leica Photogrammetric Suit (LPS), Photomodeler, Photosynth, Bundler, and AGISoft).
  - b- Evaluating the existing technique for measuring roughness and comparing it with texture analysis technique.
  - c- Developing an algorithm to detect and measure crack automatically.
  - d- Developing an algorithm to detect and measure Potholes. This includes a comparison of various available techniques for point cloud generation.
3. To test the proposed design system potential, a number of surveys were carried out using real world data and statistically evaluating and analysing the data for surface defect detection with manually measured values.

## **1.4 Contribution to knowledge**

The contribution to knowledge is to investigate the possibility of using only one camera system for the road surface anomalies evaluation. The system will be used to detect cracks, potholes, and surface roughness or skid resistance. A novel approach for measuring the roughness of the texture is another important parameter that this project will consider. This project will aim to overcome some of the existing limitations of road surface monitoring; mainly, the detection of cracks in the rough texture that is often accompanied with lots of mis-interpretation due to the raising of aggregates to the surface. Moreover, the project will investigate the use of off-the-shelf cameras for the production of 3D surfaces from point clouds for surface anomaly detection and roughness measurements.

An extensive review on the literature has not identified any previous study or research that dealt with automatic roughness measurement or depth calculation using smartphone cameras and digital SLR cameras. The rapid advancement in 3D automation

methods has opened the door for many applications and research questioning. Research into 3D method automation for roughness measurement has revealed no critical comparisons between the automated 3D surfaces of texture and the sand patch test SPT. The feasibility of the research is the fact that Smartphones with cameras are now widely available. A user only needs to capture images from different directions of the surface texture. The images can then be sent in real time to the office and a 3D automated program can be run to work out the texture depth while the user is still in the field. In order to get the scale of the model, specially designed targets were prepared and placed on the surface, with the distance measured between the different targets.

The cost of available commercial systems is not affordable by many developing countries and local road authorities. Local authorities should have a monitoring process that is effective, easy to use, and in place across the road network to meet their obligations. Therefore, building a cost effective and reasonably accurate system will be a useful alternative. The alternative system will include on-board, a Digital SLR camera and hand-held global positioning system (GPS) device. An algorithm was developed to deal with the huge number of video frames. The extracted frames underwent different processes in order to effectively detect cracks and potholes. A semi-automated solution was then achieved in identifying both cracks and postholes with manual intervention. However, a fully-automated solution is possible through linking the different algorithms in one main program interface.

## **1.5 Thesis Outline**

In this section, the outline of the thesis is displayed in such a way as to guide the reader for the contents of each chapter.

Chapter 1 starts with an introduction to the subject through briefly describing the available techniques with their limitations, and the proposed solutions. The aims and objectives of the project are demonstrated after that. Project methodologies are described later. It is then followed by the contribution to the knowledge in order to show the main contribution of the author. Finally, the chapter is finalized by thesis outline.

Chapter 2 consists of an introduction to the chapter and then the literature review is presented to provide critical review on the previous methods of road surface monitoring and the developed techniques. It then provides a comprehensive review about road pavement designs, distresses, and monitoring methods. A brief introduction to 2D

image analysis is then presented and followed by point cloud generation techniques. Finally, a summary of the chapter contents is included.

Chapter 3 concentrates on the developed methodology for roughness measurement or texture depth measurement using digital images that begins with an introduction to the chapter, followed by data sets and test site description. The material and methods required for conducting the tests are described. It then provides the procedures for performing the required trials. Results and analysis of this chapter are detailed before finalizing the chapter with a summary.

Chapter 4 includes the developed algorithm and a new methodology for crack detection that starts first with the introduction to the chapter, followed by describing the used data sets and test site location. After that, a brief description to the problem definition of cracks and a trial of detection is presented. Then the chapter methodology or procedures follows. The results and analysis of the chapter are then presented. Finally, the summary is coming at the end of the chapter to give a brief description of the chapter outcomes.

Chapter 5 focuses on the detection of potholes from video extracted digital frames and starts with an introduction to the chapter and then procedures, followed by results and analysis. Finally, the conclusion or summary of the chapter is highlighted.

Chapter 6 concludes the project and recommends areas of possibility for further research study.

## **CHAPTER TWO: BACKGROUND AND THEORY**

### **2.1 Introduction**

This chapter presents a historical review of the theory of road surface monitoring and the available manual and automated methods. Moreover, it shows how each country uses different systems. It also presents the challenges existing in road surface monitoring, and the theory behind 2D image analysis and 3D based image modelling, and new advances in the techniques used for pavement measurement.

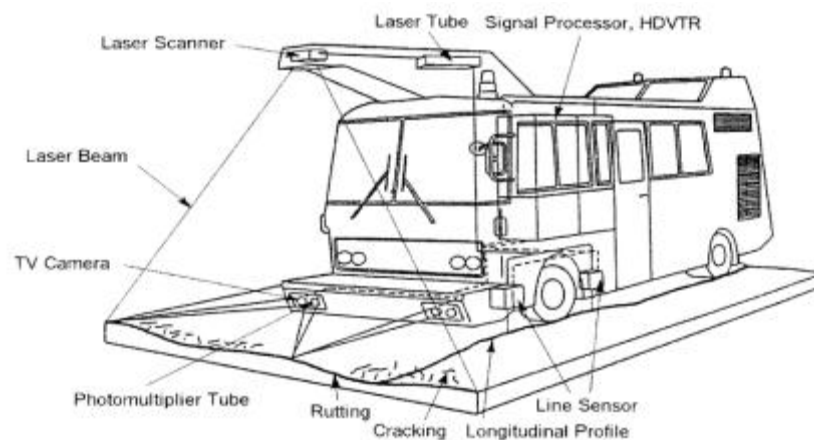
### **2.2 Literature Review**

In transportation management systems, road pavement condition data are considered to be an essential part of road maintenance. Early detection of a crack on the road surface helps the maintenance to be performed before the crack distress develops into a more serious problem, such as a pothole. The detection and measurement of pavement cracking provides valuable information on the road network condition and can help to reduce the maintenance costs. Worldwide, several damage claims are annually raised against local authorities and insurance companies regarding potholes and other types of pavement distresses (Eriksson et al., 2008).

Over the last decades, significant progress has been made in using a variety of techniques for assessing the pavement road surfaces. For efficient collection of pavement condition data, different approaches have been proposed, and various automated systems developed worldwide since the 1980s. Previous approaches to pavement condition involved a labour-intensive, time consuming, and risky process of data collection (Zhang and Elaksher, 2012).

Several systems have been designed based on different techniques such as slow speed Ground Penetrating Radar (GPR) (Forest and Utsi, 2004; Eriksson et al., 2008) or based on a laser scanning detection system (Payne, 1992). The data collection using a mobile vehicle with GPR remains a costly and disruptive method which will affect the frequency of the survey. It is also limited in terms of covering the road surface; as most of these instruments are only covering one lane. On the other hand, digital image processing is more preferable than those systems regarding the cost, coverage and frequency of the survey (Zhang, 2008b).

Image processing technology plays a vital part in analysing and evaluating the pavement surface cracking. Several systems have been designed since 30 years ago, in an attempt to automate the process of recognition and classification of pavement surface distresses in an efficient and effective approach using engineer usable indices. Starting from the late 1980s, the first automated pavement distress survey system was produced by the Japanese consortium Komatsu (figure 2.1) comprising a survey vehicle with a data-processing system on board to measure simultaneously; cracking, rutting, and longitudinal profile. The system works only at night to control the lighting condition, and a maximum resolution of 2,048x 2,048 pixels is obtained at the speed of 10 km/h (Wang and Smadi, 2011).



**Figure 2.1: Komatsu Survey Vehicle (Wang, 2000)**

From the late 1980s to the early 1990s, Pavement Condition Evaluation Services (PCES), a research unit supported by Earth Technology Corporation (ETC), created the first automated system to manage line scan cameras at 512-pixel resolution to collect pavement data. As image capturing and processing technologies associated with the system were not mature enough at that time, no significant progress was made by PCES (Wang and Smadi, 2011).

Several other commercial companies and scientific research have offered solutions for monitoring road surface condition such as Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Fugro Roadware crack detection systems. CSIRO is the Australian Commonwealth Scientific and Industrial Research Organization that developed the first fully automated road crack detection system (RoadCrack) using a line scan camera. The system was capable of identifying any crack wider than 1 mm at highway traffic speed and served Australian highways for several years. Fugro Roadware's Automatic Road Analyzer (ARAN) platform is most-broad system

employed in United States of America (USA) which utilises area scan cameras. The system has configurations with the option of 2 mm resolution or 1mm ground sampling distance, with strobe or infrared lighting, respectively. The collected image data are rated visually (visual inspection) or through the Roadware automated crack detection software, WiseCrax (Gavilán et al., 2011).

In Europe, especially in the Netherlands and Finland, the Automated Pavement Distress Analyzer (PAVUE) system has been used. In the UK, it is worth mentioning that the first shift toward automation was started in 2000, on the trunk road network, with the defining of TRAFFIC Speed Condition Surveys (TRACS) and SCANNER (Surface Condition Assessment for the National Network of Roads). A SCANNER survey vehicle is used to assess the condition of all motorways, and TRACS survey vehicle is used to assess the condition of Highways Agency owned motorways and trunk 'A' roads (the Strategic Road Network). The research was carried out on both systems in order to test the full capability of the two systems to be run on all road networks without disruption to traffic (Hawker, 2003; TRACS, 2006).

The UK Transport Research Laboratory (TRL) which is the result of ten years of extensive research, has introduced the Highways Agency Road Research Information System (HARRIS). The main objective of HARRIS was to use the state-of-the-art technology for the assessment of road conditions at traffic highway speed. This system has been used to survey national and local roads using line scan cameras at 3.5 m survey width. With the rapid development to the methods and equipment, HARRIS1 and HARRIS2 were also defined. The HARRIS2 system has reached a full capability of assessing most of the distresses occurring on the street, such as cracks, other deformations, and roughness measurement or texture depth measurement, achieved by using the laser sensors (Transport Research Laboratory, 2014).

Digital imaging is not the only solution for distresses identification. The GIE LaserVISION is an example of a laser-based system that utilizes four sensors of lasers and provides 3D measurements. The system has a low resolution 3 mm by 110 mm footprint, so it was limited to measure the transverse cracking (GIE Technology, 2009). Another promising developed survey system is to use the car as a sensor for detecting and reporting the surface conditions of roads (Eriksson et al., 2008). The system is called patrol pothole and uses the inherent mobility of a number of participating vehicles which could gather the data from vibration (accelerometers) and GPS sensors. The collected data was processed to evaluate road surface conditions. As the name of the system

suggests, it is only used for detection of the potholes. Moreover, the approach requires data training that seems to be a limitation in addition to utilizing a network of the cars.

In accordance with the system development, there was also a pace in developing the algorithms of crack detection. In most automated systems, different algorithms have been developed and adapted. An investigation was undertaken by the TRL to assess five commercial systems that are used for crack identification including Fugro, TRACS, Waylink, HARRIS1 and HARRIS2. Typical non-crack features (false-positives) or misinterpretations, such as road markings, manhole covers, road edges, patches, and joints were all considered major problems and were reported more times than the reference data by all systems. The suggestion was to better understand the crack pixels and to develop more robust crack detection algorithms (Furness et al., 2007).

Extracting distresses, such as cracks on the road pavement, in images is rather complicated and in most of the research, only a single image was used. The threshold method used by Acosta et al. (1992) was the first popular and simple method for crack detection but contained lots of false-positive cracking. Neural network-based methods have also been proposed by (Kaseko and Ritchie, 1993) to improve the problems or reduce the false detection that usually occurred with thresholding methods. Filtering methods were also produced which are based on wavelet decomposition (Subirats et al., 2006) or partial differential equations (Augereau et al., 2001). Finally, Shimamura et al. (2009) used a pattern recognition method for crack geometry to extract the quantitative data necessary for evaluating the extent and severity of cracking.

There are some difficulties of using one image, such as illumination changes due to sunlight, objects on the road, artefacts due to the dynamic acquisition, and potential false identification of cracks such as shadows. Chambon (2010) proposed a solution by using multiple images from different points of view in order to overcome some of the problems associated with false crack candidates. This proposed methodology led to an increase in the performances of detecting the false results due to the use of redundant images, but did not totally detect all the false-positive candidates.

A high-quality pavement image is not possible with the influence of sunlight and shadows or illumination changes. As a result, laser illumination based technology became available in late 2005, after two decades of struggle in acquiring high-quality pavement images without the influence of sunlight and shadows. The system was called Laser Road Imaging Systems (LRIS), which allows image acquisition without the

influence of sunlight and shadows. The advantage is that the system works at day or night, but the requirement of road pavement surface dryness and the cost of the system are limitations, and the working speed limits (0-100Km/h) are also considered as possible limitations, i.e for motorways but not residential roads. Although, a mobile platform with LRIS technology can capture a detailed image quality, it is still expensive, not efficient and time- consuming. Under consistent laser illumination, the images of LRIS are free from shadows. However, a significant amount of background noise caused by high contrast was observed from laser images. Therefore, exploiting other systems or solutions was considered essential (Chambon, 2010; Wang and Smadi, 2011).

Further research studies have looked at different sources and techniques such as satellite imaging and hyper-spectral data. Research has been carried out using remote sensing satellite to find a correlation between spectral reflectance and physical characteristics of road surfaces, such as rutting and cracking. This study showed that it was possible to describe general pavement age and surface distresses such as ravelling. However, other significant pavement quality parameters, such as rutting and cracking, were difficult to detect due to the limited spatial resolution of satellite sensors of  $>4\text{m}$  (Zhang, 2008b).

In another study, the roads were sensed using sub-meter (50cm) hyper-spectral remote sensing data. Their eventual goal was to explore relationships between remotely sensed parameters (image ratios and spatial variance) and road condition parameters such as Pavement Condition Index (PCI). Although the results of this research were promising, it proved to show high levels of uncertainty in evaluating older roads and was only perfect for roads in good conditions. This study suggested the need for further research to develop a more effective strategy of mapping the road condition through exploring other hyper-spectral mapping techniques including the short-wave infrared and small absorption features (Herold et al., 2004).

UAV based remote sensing system has also been proposed by Zhang, (2008a) for unpaved road surface image acquisition and road condition parameters, with Initial test results flew over a test area showing that the UAV images are suitable for extracting many of the parameters required for monitoring the condition of unpaved roads, and the development of image processing algorithms to detect and extract road defects from the UAV collected imagery. Various road distresses were detected from 2D imagery using the developed algorithm. Photogrammetric techniques were also used for analysing the inherent 3D geometry information in images and the fusion of 2D or 3D information to



derive road condition parameters. The current 2D image analysis uses image features (such as colour and edges), pattern recognition and image classification techniques.

The crack on the road surface is not the only defect that most of the automated systems are used to measure. Potholes for example and texture depth measurements are two additional important parameters that require attention. 2D image analysis is not sufficient to detect the pothole, although many researchers have attempted to do so (Koch and Brilakis, 2011; Koch and Brilakis, 2012). A texture difference between the pothole and the background road is the main key behind the detection of potholes using 2D image analysis. Therefore, the need for the third dimension proved to be necessary when the texture difference was not available. 3D laser scanners have also been utilized for pothole identification and detection. Many researchers starting from 2005 have used laser scanners for pothole measurement through reconstructing 3D point clouds (Liu et al., 2005; Perng et al., 2013). Other studies have used mobile sensors or vibration sensors for pothole detection (Eriksson et al., 2008; Mednis et al., 2011). The cost of building a 3D laser and the preliminary results of vibration sensors are the two main limitations of both existing pothole recognition and related assessment methods.

The texture depth measurement is another important road surface characteristic that requires attention. There is a highly significant relationship between the skid resistance in wet conditions and high speed with the road surface texture depth (Viner and Britain, 2006). Texture depth measured using the historical approach which was through the use of a volumetric technique (DoP:Transport & Infrastructure, 2012). Limitations of this technique are the lack of access to all voids in the road surface, lack of accuracy in permeable asphalt surfaces and the fact that it can only be used for dry surfaces. Recent advances in technology have allowed the use of the laser sensor to be an alternative for classical volumetric approaches (Bitelli et al., 2012; Jahanshahi et al., 2012; Sengoz et al., 2012). In addition to the cost of the laser sensors, the lack of information from only one profile is considered to be possible limitations.

This extensive study on literature has revealed that there is a wide range of automated systems that offer the monitoring of the road surface. However, the cost of these systems is a major limitation which is not affordable by many road authorities especially in developing countries. Therefore, the necessity for a cost effective alternative is critical. It has been noticed from literature that three characteristics of road defects, i.e. cracks, potholes and texture depth measurement, need further studies. Although, intensive studies were carried out on the crack detection, still some drawbacks

can be noticed. These drawbacks, such as shadow removal and roughness of texture need to be overcome. Roughness measurement or texture depth measurement is another area where a gap has been found. Based on author knowledge, no study has used images based on 3D modelling for texture depth measurement and pothole detection, so an alternative of using only digital images for texture depth measurement will be investigated. Finally, the reconstruction of 3D models from a digital image extracted from video files will be evaluated. The next sections will cover the road surface, sensors, platform, and the new processing techniques in more details.

## 2.3 Road pavement design, distresses and monitoring

A pavement can be defined as a horizontal structure supported by local available natural material with the purpose of transferring or distributing the applied traffic and other loads safely and reliably to the supporting soil. Acceptable riding quality and adequate skid resistance are more important characteristics that the pavement has to provide in addition to favourable light reflection and low noise pollution. A pavement can be classified by construction type into four different groups; unpaved, asphalt, concrete, and elemental. Unpaved road is either soil or gravel that may be sealed or unsealed. Elemental is a pavement with concrete blocks or bricks paver and may be natural stones. Asphalt is flexible in nature and called a flexible pavement due to the nature of distributing the loads beneath it. While concrete is a rigid structure, and the distribution of loads differs from the flexible pavement, figure 2.2 shows the differences between flexible and rigid pavement with regard to load distribution (Pearson, 2012).

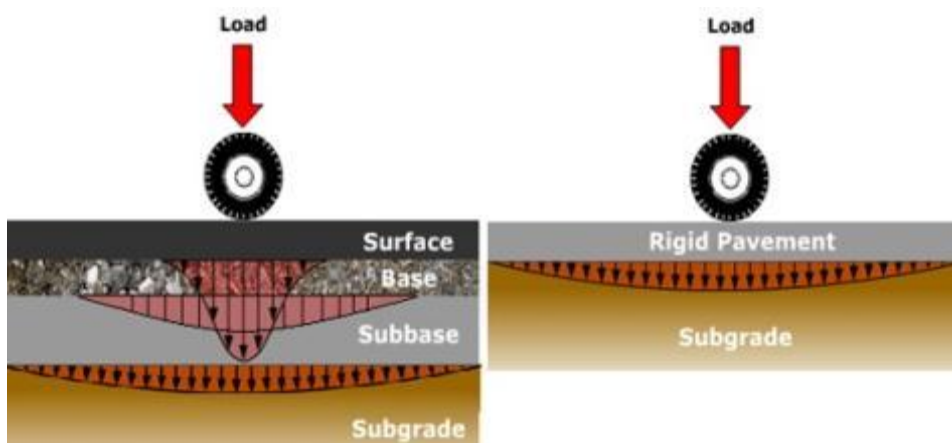


Figure 2.2: Flexible and rigid pavement load distribution (OSU, 2012)

Typical flexible pavement structure consists of several layers of materials that can provide elastic nature to the load distribution. Due to its nature of load distribution,

flexible pavement roads suffer from distresses more than rigid pavement roads. However, flexible pavement is more economic and easier to deal with in case of damage, besides that it is recyclable. Policy and economy of most highway agencies can decide the selection of pavement type. Maintenance or rehabilitation generally requires every 10 to 15 years for flexible pavements due to its life cycle, figure 2.3. On the other hand, rigid pavements can usually serve longer than 20 and up to 40 years with slight or no maintenance or rehabilitation (OSU, 2012).

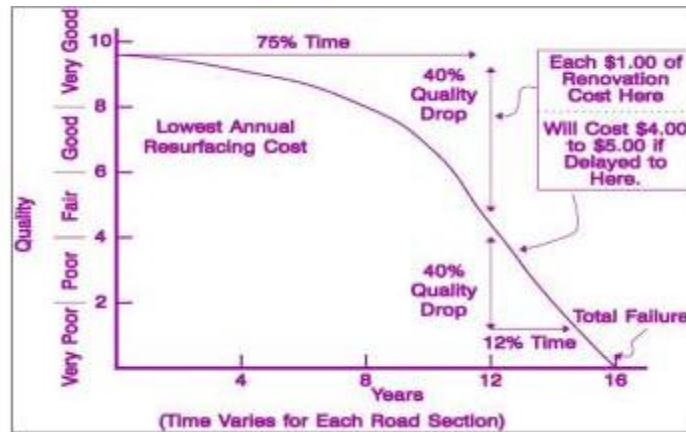


Figure 2.3: Life cycle of flexible pavement and the cost of maintenance (Alex, 2011)

### 2.3.1 Flexible pavement design

The wheel load acting on the pavement is distributed to a wider area of soil, and the pressure decreases with the increase in depth of the layer, as can be noticed from figure 2.2. The stress distribution characteristic becomes better when the flexible pavement has many layers. Therefore, the concept of a layered system is applied in the designing of the flexible pavement, and the top layer has to be of the best quality to withstand maximum compressive stress in addition to wear and tear. The top layer of flexible pavements is constructed using bituminous materials. This layer can be in the form of surface treatments such as that used in low volume traffic. Another type of top layer is asphalt concrete surface which can be generally used on high volume traffic. The top layer of the flexible pavement has to be designed in such a way in order to allow for water drainage from the surface. In the lower layers of the flexible pavement, the low-quality material can be used due to the fact that it experiences smaller amount of stresses. Figure 2.4 shows a typical cross-section of a flexible pavement that is commonly built, for more information about the used materials in each layer, refer to (Mathew, 2009).

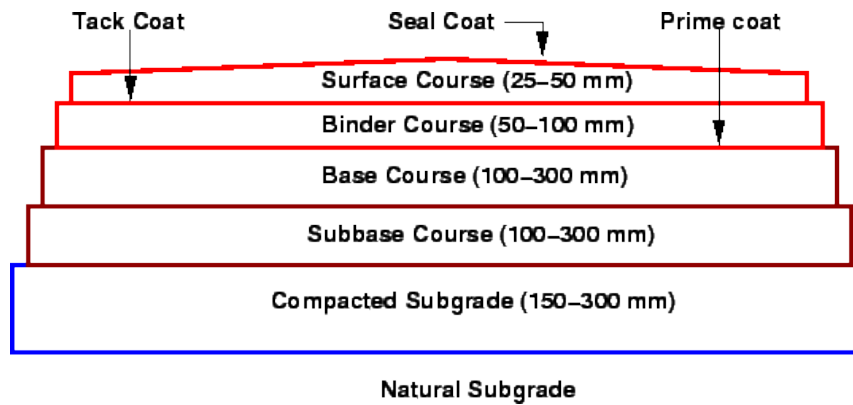


Figure 2.4: Typical cross section of a flexible pavement (Mathew, 2009)

### 2.3.2 Modes of failure in flexible pavement

According to Pearson (2012), failure in flexible pavement can be due to one of the following reasons:

- 1- Inadequate bearing capacity of the formation, this happens when the load exceeds the formation design capacity.
- 2- Failure due to frost damage, this happens when the water or moisture gets into the pavement either due to its permeability or due to lack of maintenance.
- 3- Failure of the constituent material due to fatigue, the pavement by its nature is elastic and returns to its regular shape as the loads are removed. However, due to the existence of some plastic strains, the deformation is unrecovered and stays permanently.
- 4- Failure of the constituent material due to environmental exposure. This is due to the effect of sunlight and oxygen as well as the effects of moisture in non-freezing conditions.
- 5- Inadequate quality of structure that commonly take place when the construction is carried out in a time of year that severely affect the quality and longevity of flexible pavement. Also this may be due to the lack of experience and skills of artisans to deliver a good quality pavement in all weather conditions. The failure in the road surface that takes many shapes and severities are presented in the next section.

### 2.3.3 Distresses on the flexible pavement

Road asphalt surface is subjected to distresses due to loading, environmental conditions and normal wear. Asphalt pavement distress represents a significant economic

and engineering concern. This distress is estimated to cause damage costing up to \$10 billion each year in the United States alone. Earlier reports in 1998 estimated that this figure was even more with the Federal Highway Administration stating that road surface maintenance costs peaked to \$17 billion each year in the United States. One crucial step to overcome and manage this problem is by accurately assessing the pavement condition and its change over time. The effective assessment of the road surface is applied by categorizing and verifying cracks or distresses into different classes (Cheng and Miyojim, 1998; Rababaah et al., 2005).

The classification and quantification of the type, severity, and the amount of surface distress is the primary method for assessing the condition of road surface pavements. The Long-Term Pavement Performance (LTPP) program within a 20-year life time has been carried out by highway agencies in the Department of Transportation in the United States and in 15 other countries. They have collected data on pavement condition, climate, and traffic volumes and loads from more than 1,000 pavement test sections. This program focuses on three main types of road surfaces; asphalt concrete-surface (or flexible pavement), joined Portland cement concrete (or rigid), and continuously reinforced Portland cement concrete (or rigid as well).

Many types of distresses have been identified, and their characteristics were catalogued for the flexible pavement by LTPP. The focus of this project will be on the flexible pavement since this type of pavement forms a large part of transportation infrastructure and is considered to be suffering more from distresses due to its nature of loading distribution. Distresses on flexible pavement can be divided into five generic types as can be seen from figure 2.5. Only two defect types of cracks and potholes are described in this thesis, for more information about other defects refer to (Miller and Bellinger, 2003; Rababaah et al., 2005).

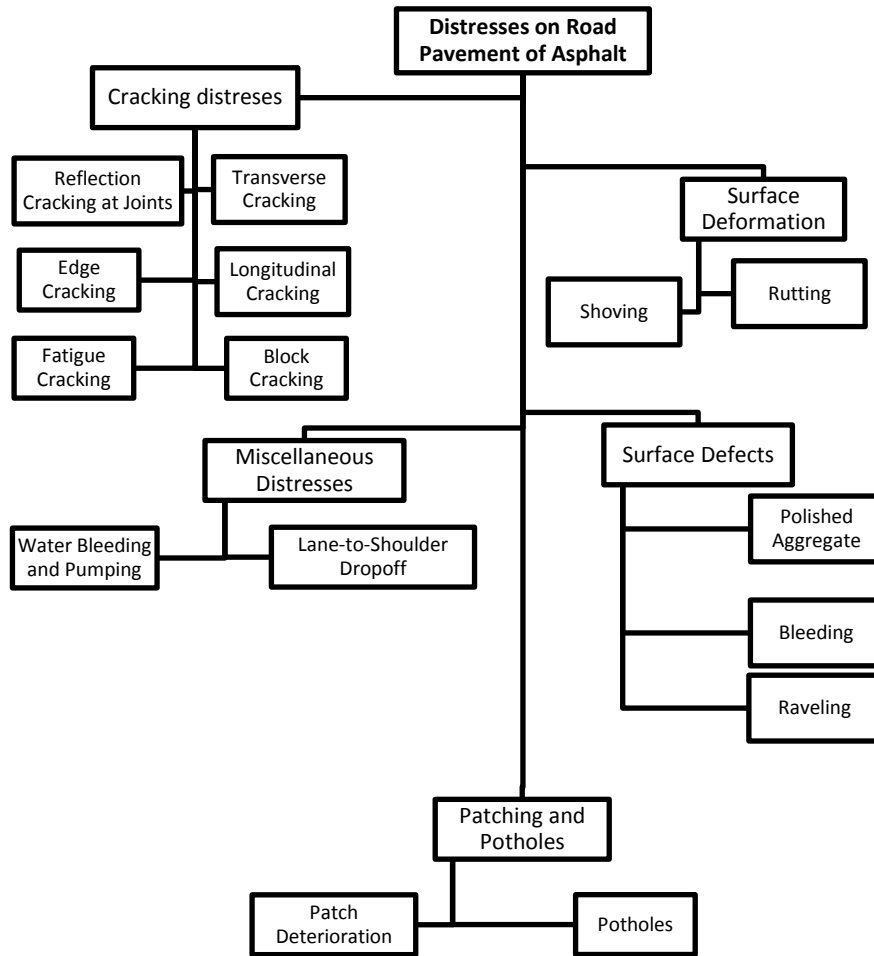


Figure 2.5: Type of distresses on flexible pavement (Miller and Bellinger, 2003)

### 2.3.3.1 Cracking distresses

Cracks are fissures resulting from incomplete or complete fractures in the flexible pavement surface. Cracking distresses on the road pavement surface can occur in a variety of patterns, ranging from isolated single cracks to a connected pattern extending over the entire or part of the pavement surface. The types, quantities, and severities of cracks distresses in the asphalt pavement can also vary depending on the age of the crack and the load of traffic. The LTPP has introduced six types of cracks that are frequently found in flexible pavement, namely fatigue cracking, block cracking, edge cracking, longitudinal cracking, reflection cracking at joint and transverse cracking. According to severity, each type is classified into three severity levels of low, moderate, and high. These various types of cracks are listed in the figure 2.6.

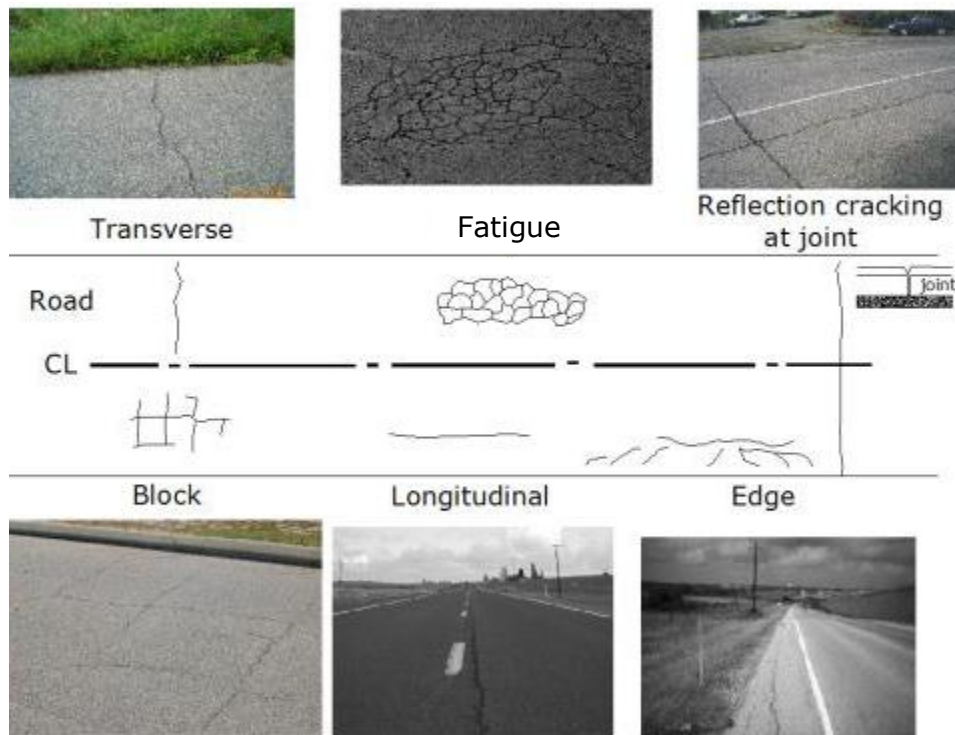


Figure 2.6: Distresses types on the asphalt road pavement, images adapted from (Asphalt Institute, 2009; Pavement Interactive, 2012)

### 2.3.3.2 Patching and Potholes

Patches are parts of the pavement surface that were subjected to removal and replacement or adding extra material to the original surface after construction. Usually, this part of the surface is greater than  $0.1 \text{ m}^2$  and can take three severity levels. Low level has a depth of less than 6 mm; moderate level has a depth of between 6-12 mm, and high level has depth of greater than (12 mm). Commonly, patches are the parts of the surface that has been previously treated; either during potholes maintenance or as part of the work conducted on the road. Potholes, on the other hand, are defects or holes in the surface that take the form of small bowl shape depressions of various sizes in the pavement surface that penetrate to the lowest surface through the asphalt layer. Also, a pothole has three levels of severity; low has depth of less than 25 mm, moderate has depth range between 25-50 mm, and high has depth of more than 50mm, figure 2.7 shows a high severity sample of patch and pothole (Miller and Bellinger, 2003).

There are four main causes of potholes occurrence according to Eaton et al. (1981):

- A. Insufficient thickness of the road to support the traffic load.
- B. Poor drainage that usually affects the thin pavement more than the thick.

- C. Failure at trench and casting utilities, examples are (manholes and drainage of water facilities).
- D. Various types of pavement defects and cracks that were left un-maintained or unsealed for a long time which led to crack development due to water penetration to the lower layers.



Figure 2.7: High severity patch and pothole (Miller and Bellinger, 2003)

### 2.3.4 Flexible pavement evaluation and data collection

To insure that the money spent in maintaining pavements is used effectively, different types of information are required in order to maximise the likely contribution to the existing asset. It is required to assess the functional and structural conditions of the road network or a section of the road either for purposes of routine monitoring or for planned corrective actions. The primary concern of the functional process is to evaluate or measure the ride quality, skid resistance and characteristic of surface texture. On the other hand, layer thickness, structural capacity of the pavement (as measured by deflection) and material properties are concerned with the structural condition of the pavement (Lenz, 2011). Four main methods are available to measure the functional and structural conditions of the flexible pavement, which are described in the next sections.

#### 2.3.4.1 *Non-destructive testing*

Non-destructive testing can be defined as a group of analysis techniques to assess the characteristics of the flexible pavement without causing damage. The collected data in the field are usually objective in nature using non-destructive testing but often subjectivity appears in the analysis techniques. Usually these types of tests are applied on the existing pavement structure without requiring maintenance work to bring back the pavement to its pre-testing state. Therefore, they are preferable due to their nature of minimizing the disruption to traffic and for determining the locations where selective material sampling should be conducted to evaluate the need for further “destructive” testing (Lenz, 2011).



Non-destructive testing is used to examine both the functional and structural properties of the pavement. Skid resistance and roughness of the texture are two examples of functional properties that can be measured through non-destructive tests. There are ranges of non-destructive tests that can be applied on the pavement in order to evaluate the health of the pavement. Examples of these tests are; falling weight deflectometer, seismic pavement analyzer, dynamic cone penetrometer, ground penetrating radar, nuclear density gauge and non-nuclear electrical gauges, for more information about each of these tests and their uses refer to (Pavement Interactive, 2012).

#### ***2.3.4.2 Destructive testing***

As the name suggests, destructive testing includes several methods (coring, boring, and trenching) where a part of road pavement is taken away in order to provide more detailed data about the pavement that is not possible to obtain through non-destructive testing. Such detailed data include, determining the laboratory mechanical properties, physical, and chemical properties. It can also provide a visual inspection of pavement layers through coring and trenching, refer to (Lenz, 2011) for more information.

#### ***2.3.4.3 Visual condition surveys***

Usually, visual condition surveys serve as a qualitative indicator of pavement overall condition which can include both functional and structural pavement condition aspects. Specific equipment is utilized to quantify characteristics of the pavement structure for both structural and functional tests. Pavement surface inspection is carried out visually by two or more evaluators. Both evaluators arrive at the same assessment location for testing the condition of the road section or travel along the highway for identifying the distress areas, stop the vehicle when one is detected, measure some parameters and record them on a specially designed sheet. This type of survey is highly subjective for many aspects of pavement evaluation and inconsistencies in distress detail are recorded over the space and across evaluations (Cheng and Miyojim, 1998; Lenz, 2011).

Different rating conditions were designed over the past period, starting from the 1950s; the first rating procedure was the American Association of State Highway and Transportation Officials (AASHTO) road test designed in Illinois. The test aim was to study the performance of known thickness of pavement structure under moving loads of known magnitude and frequency. The test was designed to study both rigid pavement

concrete and flexible pavement asphalt. The information obtained from these tests was useful to advance the knowledge in pavement structural design, load equivalencies, pavement performance, climate and environmental factors, and much more. The present serviceability rating (PSR) was the first rating designed by AASHO road test which is based on individual observation. The crew of observers usually ride around the test roadways and then will rate their ride quality using the measurable scale shown in figure 2.8 (Pavement Interactive, 2012).

Acceptable?		5	Very Good
Yes	<input type="checkbox"/>	4	Good
No	<input type="checkbox"/>	3	Fair
Undecided	<input type="checkbox"/>	2	Poor
		1	Very Poor
		0	Very Poor

Section Identification	Rating
Rater	Vehicle
Date	Time

**Figure 2.8: Individual present serviceability rating form (Pavement Interactive, 2012)**

PSR was soon replaced by the present serviceability index (PSI) due to necessity of a transition to a non-panel based system since PSR rating is not practical for large-scale pavement networks. To develop PSI equations, the real PSR measurement was correlated with various pavement measurements parameters such as cracks, slope variance, and profile, refer to Highway Research Board (1972) for more information about the developed equations and the correlation with actual measurements. Many other rating systems were developed over time in different countries. In the UK, the United Kingdom Pavement Management System (UKPMS) has not changed since 1974 (TRL, 2012). In the United States, the PSI was replaced by a more advanced and easy to use numerical rating system called pavement condition index (PCI) which was developed by the United States Army Corps of Engineers (USACE), for more information on this numerical rating system refer to ASTM D 6433 – 07 (2007).

#### 2.3.4.4 Automated condition survey

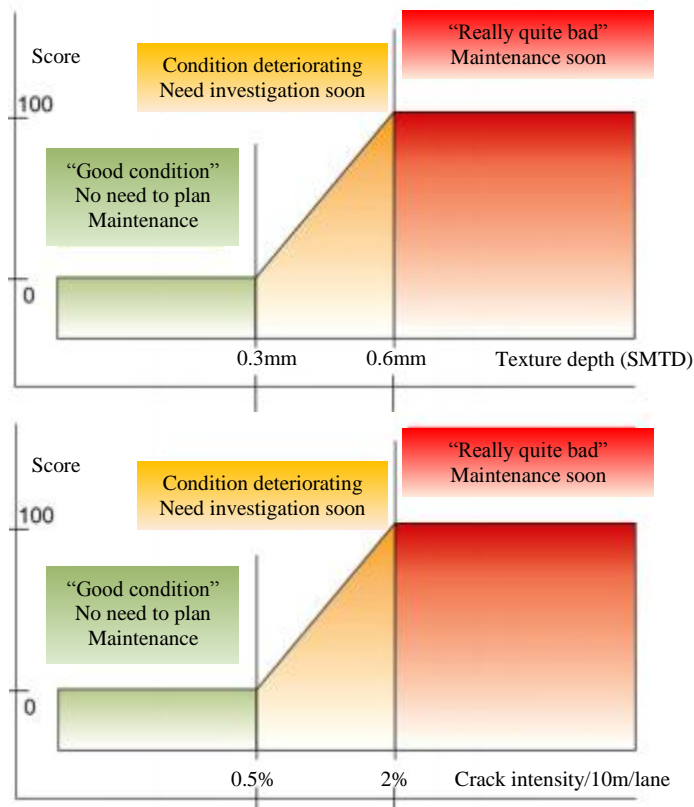
In order to minimize the difficulties of manual data collection, a range of approaches and concepts for the automated pavement distress detection have evolved, with several technologies being developed. An example of these techniques, the pavement image data collection is the most-common approach. Survey vehicles are usually adapted which are capable of collecting and storing pavement surface images. Most of these survey vehicles are mounted with video camera or photographic camera that can record pavement surface images, with the complimentary use of laser-illumination-based technology which became available in late 2005 as LRIS. Different commercial and research based systems were introduced in each country. The list of systems presented in table 2.1 shows some of the available automated systems and their capabilities. Please refer to the following references for more details about each of these systems (Wang and Elliott, 1999; Viner and Britain, 2006; CSIRO, 2010; Wang and Smadi, 2011; Pavemetrics, 2012).

Table 2.1: Example of automated systems for pavement evaluation

System	Year	Country	Speed	Features	Distress types
<b>Komatsu</b>	1980	Japan	10 km/h	Working at night only	Cracking, rutting, longitude profiles
<b>PCES</b>	1990	USA	traffic	1 <sup>st</sup> to use line scan camera	Cracking
<b>RoadWare</b>	1995	Canada	Up to 80 km/h	Two cameras sync. with a strobe illumination system	cracking and other distresses
<b>TRACS</b>	2000	UK	traffic	video, laser sensors	All distresses, texture measurement
<b>SCANNER</b>	2001	UK	traffic	Use Road Condition Indicator (RCI)	All distresses, texture measurement
<b>HARRIS</b>	2002	UK	traffic	line scan cameras and advanced lighting system over a 3.5m	All distresses, texture measurement
<b>CSIRO</b>	2008	Australia	Up to 105 km/h	Detect cracks as fine as one millimetre in width	Cracking only
<b>Pavemetrics</b>	2009	Canada	Up to 100 km/h	Laser Road Imaging System	Most of distresses

### 2.3.4.4.1 SCANNER road condition indicator

SCANNER (Surface Condition Assessment for the National Network of Roads) is a survey system that provides a consistent measuring method of the road surface condition, and has been developed by the UK Roads Board. It is usually employed for road carriageways throughout the United Kingdom. SCANNER uses an automated machine in order to collect a range of road condition parameters that includes ride quality, rut depth, intensity of cracking, texture depth and edge condition. The SCANNER uses the collected road condition parameters to quantify the road distresses in terms of an indicator called Road Condition Indicator (RCI). RCI has been developed through a process of research and development, and testing and refinement by the UK Roads Board. The main aim of RCI is to identify the length of the road or section of a network that needs urgent inspection or soon planned maintenance. To obtain RCI, each of the collected road parameters is scored between two thresholds; a lower threshold which is equal or below usually do not require considering any maintenance. Whereas upper threshold limit requires a soon planned maintenance. The full process of calculating the RCI from each of the used parameters (Cracks, Texture depth, Ride quality and etc.) are available in Department of Transport, (2011). Samples of the threshold used in RCI are in figure 2.9.



**Figure 2.9: Samples of parameters and threshold used in RCI**  
Threshold extracted from the provided tables in (Department of Transport, 2011)

### 2.3.5 Road surface texture roughness

Roughness can be defined as a deficiency or lack of smoothness in the transverse profile of the road surface that potentially causes poor ride quality. Usually, roughness can be produced by exposing the surface of the road (or even subsurface) to distresses or by poor construction quality. An example of poor construction is the compaction process of sub base and the top of pavement layers as well as the grade control and poor bonding of the surface Hot Mixture Asphalt (HMA) layer to the underlying layers. The cause of roughness might be also attributed to shrinkage or swelling activity in the subgrade or soils fill (Lenz, 2011).

Roughness measurement of texture is important for many applications and commercial purposes as well as in the industry. In road surface monitoring, there is a highly significant relationship between depth of texture and the skid resistance which is linked to the accident rate. The tyre / asphalt interface, as stated by the World Road Association (WRA) in World Road Association (2012), should be better understood and evaluated. Interaction between tyre and road surface affects not only the grip but also other important properties as well; such as rutting, cracking, ravelling, noise and long term durability, figure 2.10. In the concrete industry as well, the roughness of the texture plays an important role in when overlaying a new layer or lining with the required amount of painting. In addition, concrete surface roughness can play a crucial role in the bonding between the two layers (Leising, 2010).

Texture detail at different scales is essential. At the micro level, the texture details are related to the aggregate surface properties. However, at the macro level, the textured details are related to the actual asphalt surface, figure 2.11. At the micro level, the characteristics of the aggregate are defined. The microscopic properties of the surface function of the aggregate gives the basic level of friction achievable between the tyre and the road and thus provides adequate skidding resistance. The properties that can be measured at micro level correspond to wavelengths less than 0.5 mm and peak-to-peak amplitudes of the profile ranging between 1 micron and 0.2 mm. Current laser scanner technology cannot provide direct measurement of these properties at such scale. The only available tools for measuring the roughness at micro level is through the use of the Portable Skid Resistance tester in the field and the Accelerated Polishing machine in the laboratory (McQuaid et al., 2010; Airy, 2012).

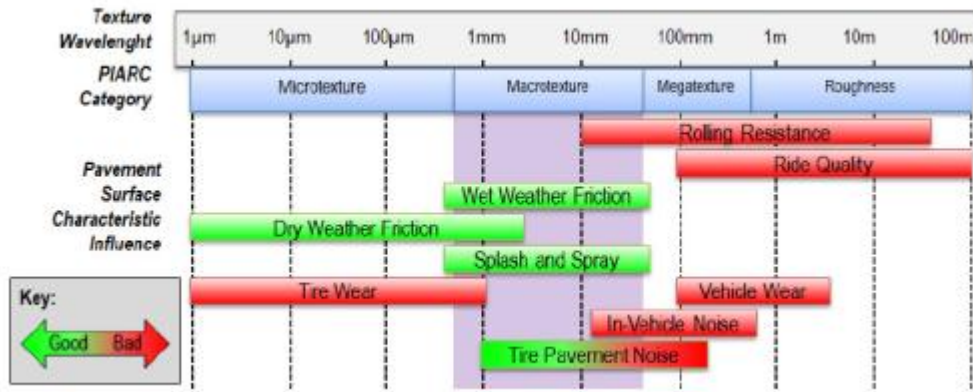


Figure 2.10: Relationship between texture and characteristics of the pavement surface  
(World Road Association, 2012)

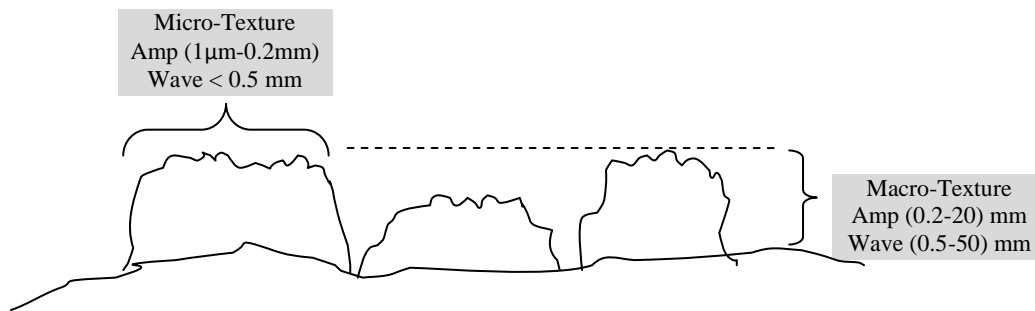


Figure 2.11: Macro and Micro level properties of aggregates

### 2.3.5.1 Road surface texture and skid resistance

Transmission of vertical, transversal and longitudinal forces mainly depends on the contact area between the road surface and the tyre. The contact area between the tyre and the surface texture available for both horizontal and vertical force transmission is insufficient due to flattening of the tyre surface in many cases. The limited contact area depends mainly on the tyre geometry, the load on the wheel and tyre pressure, and increases with lowering the tyre pressure. This contact area also decreases when a medium, such as water, separates the tyre from contacting with the road surface leading to a decrease in transmittable force and therefore skid resistance. The tyre/ pavement contact area also minimizes when increasing the texture depth or the roughness of the texture and vice versa as can be seen in figure 2.12. Therefore, the skid resistance is dependent on the texture depth in addition to the surface texture characteristic which is more closely related in wet weather condition and at high speed (Viner and Britain, 2006; Nutz and Hoffmann, 2012).



Figure 2.12: Tyre/Pavement contact area in rough and smooth texture (Nutz and Hoffmann, 2012)

### 2.3.5.2 Methods of characterising roughness of pavement texture

In order to establish if a pavement surface texture has sufficient skid resistance, two parameters need to be evaluated and provided; they are micro-texture and macro-texture, that both were described previously. Micro-texture provides the frictional forces required to bring a moving vehicle to stop at low speeds (less than 50 km/h). Macro-texture, on the other hand, plays a critical role in wet conditions and at high speeds, refer to (Fwa, 2006) for more details. These two important parameters can be measured using different equipment as explained in the next two sections.

#### 2.3.5.2.1 Micro texture measurement techniques

As previously explained, the micro-texture exists on the microscopic scale that has peak-to-peak amplitude in the range of 0.001 to 0.2mm and a wavelength of less than 0.5mm. These properties are not feasible to be measured due to the microscopic nature of the aggregate. However, the primary interest of the pavement surface is the polishing vulnerability of a particular stone under the scrubbing action of tires at the tyre /pavement interaction area. Therefore, it is important to evaluate the so-called Polished Stone Value (PSV) which can be done in the field using the British Pendulum Tester (BPT). Some other tools do exist but are limited for use in the laboratory, an example of these tools are polishing machine and the Small-Wheel Circular Track Wheel (Fwa, 2006) .

### 2.3.5.2.2 Measurement of frictional forces

The measurement of the resistance in the interaction area between the tyre and the pavement in the case of a car moving or braking is dependent on the friction coefficient. The forces resisting the motion (tyre/pavement contact area's frictional forces) can be defined as the friction coefficient when divided by vertical forces. Increases in the friction coefficient can lead to increases in the mobilized force at the interface area to resist the vehicle movement and hence the vehicle can more rapidly be stopped or slowed (Fwa, 2006).

Usually, the friction coefficient is not constant and depends on many factors that include tyre components properties, tread pattern and depth of the tyre, the micro and macro surface pavement textures, and the existence of water surface or other lubricants. The presence of water film in wet condition between the tyre and pavement surface significantly reduces the friction coefficient. Moreover, the frictional resistance is also affected by the percent of slippage occurring between the tyre and pavement surface which can be define by the flowing equation (Fwa, 2006).

$$S = \frac{v - v_r}{v} 100\% \quad 2-1$$

Where:

S=the percent slip

$v$  =velocity of the test vehicle

$v_r$ = rotational velocity of the test tire

The frictional measuring devices can be categorized into four types. The tools that are usually used to measure the skid resistance on the pavement surface are; locked wheel devices, fixed slip devices, variable slip devices, and sideways-force devices such as Sideway force Coefficient Routine Investigation Machine (SCRIM). For more details about each of these tools and their uses, please refer to (Fwa, 2006).

### 2.3.5.2.3 Macro texture measurement techniques

Two different techniques are now available to measure the macro texture of the pavement or texture depth. The first method is called Sand Patch Test (SPT) which uses a known volume of sand or glass beads spread evenly over the pavement surface to form a circle. The used sand or glass beads will normally fill the surface voids. The diameter of the circle is measured in four different directions using a ruler, and the value is averaged which is then used to calculate the Mean Texture Depth (MTD).



$$MTD = \frac{v}{area} \quad 2-2$$

Where:

MTD	The mean texture depth
$v$	Volume of sand
area	Area of circular patch

Many engineers have defined the texture of the road surface through the use of the SPT method, which has a traditional history. Through utilizing the mathematical formula listed in equation (2-2), the depth of road surface texture can be determined. The depth obtained from SPT then can be used for the process of the pavement evaluation. It is also worth to mention that this process is dependent to a large extent upon the operator. The limitation of SPT however, is the limited access all the voids in the surface and it has an uncertainty of measurements of about  $\pm 230 \mu\text{m}$  for a confidence interval of 95% (Rasmussen et al., 2011). In addition, it is a slow and labour intensive process. As well as having safety issues when being undertaken on an operating road. All of mentioned limitations have been addressed through the use of automated processes utilizing laser scanner techniques.

In a998, Nippo Sangyo Co., Ltd., Japan has introduced a new laser-based device, the Circular Texture Meter (CT Meter). A circular profile of 284 mm in diameter or 892 mm in circumference is measured using the CT Meter laser sensors. The Mean Profile Depth (MPD) is then calculated from the measured profile (China and James, 2011). In addition to the cost and the difficulty of mobilization due to the weight, the disruption of the traffic is another limitation of CT Meter. So, the high-speed laser technique is currently being used for macro-texture measurement by many transportation agencies.

The high-speed laser technique is another method that is used for texture depth measurement or macro-texture characterization. The method is used through utilizing the laser texture measurement devices which work by measuring the distance between the sensor and the road surface. As the sensor moves along the road, changes in distances due to the surface texture are recorded at short intervals, possibly 1mm. A profile of detailed surface texture is built up which requires further processing to interpret a valuable measuring index. Two techniques can be utilized for surface texture profile data processing which is Sensor Measured Texture Depth (SMTD) or Mean Profile Depth (MPD), for more information about both techniques refer to (Viner and Britain, 2006).

There are some limitations associated with high-speed laser and other laser profile measurement techniques such as lack of information particularly when trying to understand what is happening, and the data provided by these techniques is limited. Therefore, a precise 3D modelling surface of texture shaping is more important than the profiles in order to better understand surface characteristics. A new-developed technique is proposed in this project through digital imaging and 3D modelling which will be discussed in the next chapter. The following sections will concentrate on the image processing techniques and other processing techniques normally utilized for road surface monitoring.

## **2.4 2D Image analysis**

Image analysis can be defined as a process of extracting meaningful information from images and mostly from digital images by automatic or semi-automatic methods. Image analysis can be divided into many tasks and processes such as edge detection, shape finding, object counting or measuring properties of an object through morphological tools. Image analysis algorithms are commonly comprised of edge detection, shape recognition, image segmentation based on colour, and thresholding. These image processing techniques can be combined with region analysis functions to obtain statistical details from images to support human analysis with extra qualitative and quantitative data (Matlab, 2014a).

### **2.4.1 Image Segmentation**

Image segmentation in Computer Vision (CV) is a process of segmenting or partitioning the digital images into multiple parts or segments (sets of pixels). The aim of segmentation is to simplify or change the presentation of a digital image into more meaningful objects that can be easily analysed and dealt with. Normally, the segmentation process is utilized to locate objects and boundaries such as lines, curves and objects of interest within images. Pixels with the same characteristics are assigned with the same label and isolated from other pixels. The simplest process in image segmentation is the thresholding of image pixels (Roy, 2012).

In thresholding, the grey level pixels are converted to a binary image by turning all pixels below a certain threshold (or level) to zero and above that level to one, see equation (2-3). The threshold level may be designed globally for the whole digital image or locally to overcome the global threshold limitation due to the changes in illumination across the scene that cause some parts to be darker or brighter than others. The major

problem with thresholding methods is that they only consider the pixel intensity values as border limits, and do not consider any other relationship between pixels such as the location of a pixel that might lead to erroneous segmentation. In thresholding the image is separated into two regions; dark and light which in consequence means a binary image is created (Morse, 2000).

$$B(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq Th \\ 0 & \text{otherwise} \end{cases} \quad 2-3$$

Where:

$B(x, y)$	Converted binary pixel
$f(x, y)$	Original grey level pixel
$Th$	Global or local threshold

Another process that segmentation includes is edge detection that provides more rigorous results than the thresholding method for image segmentation. Edge detection is the process in which the pixel in grey scale image undergoes a set of mathematical operations that aim at identifying sharp changes in brightness or identifying discontinuities in a digital image. Edges are typically occurring on the boundary between two different regions in a digital image. Edge detection was developed 30 years ago and passed through many stages. Over this period, two main methods have been apparent: the first approach is template matching (TM), and the second one is differential gradient (DG). Both approaches aim to search for the intensity gradient magnitude ( $g$ ) for the sufficiently large values which can be reliably treated as an indicator of edge presence (Roy, 2012).

The main differences between both methods (DG and TM) lie in the way that both methods estimate the gradient  $g$  value (the pixels with minimum and maximum intensity values) and determine local edge orientation. Both operators utilize suitable convolution masks to estimate the local intensity gradient. The TM approach usually utilizes up to 12 convolution masks that are capable of calculating local gradient components in various directions which can be approximated by taking the maximum response of all components. However, the DG operator type uses only two convolution masks, one in the  $x$  direction and another in the  $y$  direction. An example of the DG operator is Sobel edge detection that uses a pair of  $3 \times 3$  convolution masks; see figure 2.13, one estimating the gradient in the  $x$ -direction and the other estimating the gradient in the  $y$ -direction. Then, the approximate absolute gradient magnitude (the strength of the edge) at each pixel can be found as shown in equation (2-4). The direction of the gradient is then easily

calculated once the gradients in the x and y directions are known, see equation (2-5). For different edge operations, their processing stages and the utilized convulsion masks, refer to (Roy, 2012).

-1	0	+1
-2	0	+2
-1	0	+1

Gx

+1	+2	+1
0	0	0
-1	-2	-1

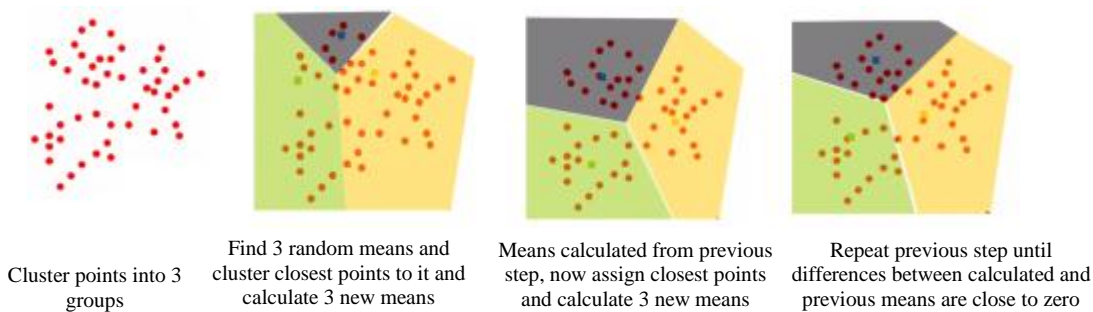
Gy

**Figure 2.13: Convolution masks of Sobel edge operation**

$$|G| = |Gx| + |Gy| \quad 2-4$$

$$\theta = \tan^{-1}\left(\frac{Gy}{Gx}\right) \quad 2-5$$

Segmentation can also be done based on colour which is called colour based segmentation. A good example of colour based segmentation is K-mean clustering. K-means is a least-squares partitioning method that aims at dividing or partitioning n observations in to k clusters in which each observation belongs to the cluster with the nearest mean. Clustering based on K-mean is done first through calculating the mean of each cluster and then assigning each observation to the nearest cluster by calculating the distance between each observation and the mean of each cluster. The previous two steps iterate until the sum of squared errors within each group cannot be minimized anymore. The created groups are geometrically as compact as possible around their respective means, see figure 2.14.



**Figure 2.14 :Diagram explain the K-mean clustering process**

The K-mean clustering is more rigorous than other previous segmentation tools for some applications because it includes, in addition to the intensity values of pixels, the geometry as well for image segmentation (UTAH, 2002; Lugade, 2011). After image segmentation, image filtration, is another important operation in image analysis, and will be reviewed in the next section.

## 2.4.2 Image Filtering

Image filtering is the process that deals with random variations in intensity, variations in illumination, and poor contrast (that may be associated with an imaging system). Image filtering is an image enhancement tool that acts as a noise removal through windowing operations that achieve a useful outcome, see figure 2.15 for the clear effect of image filtration (USF, 2014).



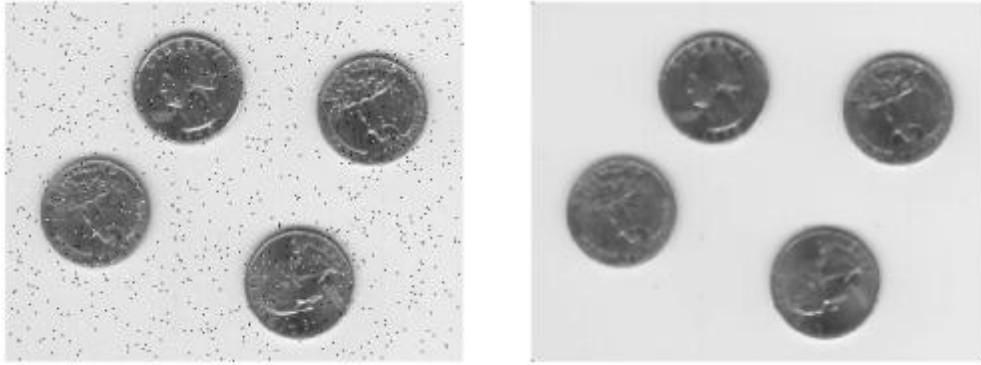
**Figure 2.15: Demonstration of the effect of the neutral density filter (Wiki, 2014)**

There are many approaches of image filtration such as median, local mean, mode, low-pass, range, and Gaussian filter. Each of these filters has a special effect on image enhancement and pixel manipulation. For example, in the case of mean filtration, a window size such as 3x3 neighbourhood is selected for pixel manipulation. The mean filter is a simple linear filter which is implemented by local averaging operation in which the value of the central pixel is replaced by the mean of all neighbouring pixels. Figure 2.16 shows the effect of the mean filter using different size windowing (Matlab, 2014b).

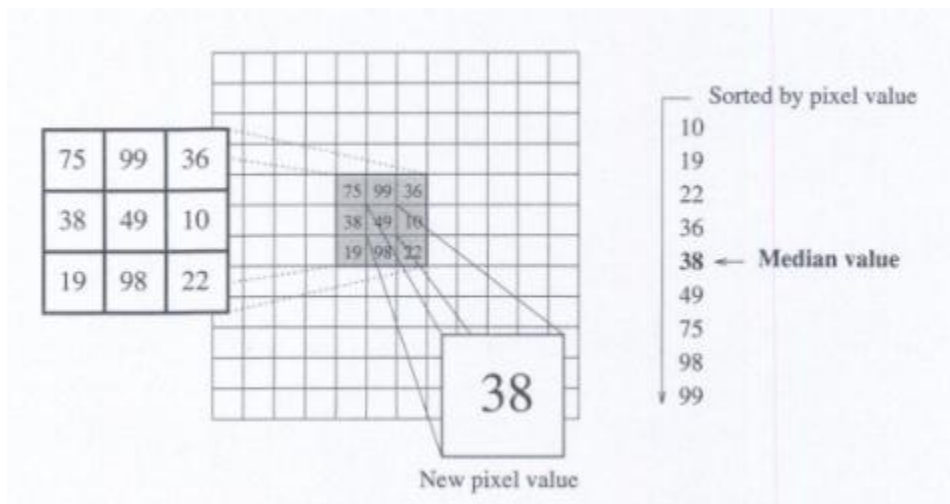


**Figure 2.16: Effect of mean filtration, original image, 3x3 filtering, 5x5 filtering**

Each filter was designed for a special effect, for example, the ‘salt and pepper’ noise in digital image can be removed using median filtering as can be noticed from figure 2.17. As the name suggest, the median filtering works by selecting the median value in the (i-by-j) neighbourhood around the input pixel in the digital image as shown in figure 2.18.



**Figure 2.17: Example shows the effect of median filtering on salts and pepper noise (original image left and median filtered results right)**



**Figure 2.18: An example illustrating the median filtering with 3x3 neighbourhood (USF, 2014)**

Image filtering can also be used for smoothing the image and reducing the effect of image sharpness. The task is especially important for some applications such as reducing the effect of sensitivity of edge detection. In this project, different filters with various smoothing effects were designed to reduce the effect of the texture roughness, please refer to section 4.8.3.2. It is important to note that the image filtering can be also used for image de-blurring when the image is blurred during the process of acquisition. More information is available in the following section.

### 2.4.3 Image de-blurring

Image blurring or degradation can be caused by many factors such as movement during the image acquisition process, by capturing images with longer exposure time, when the camera optics is out of focus, or in light-limited situation. The captured image quality in some cases might degrade in various ways which urge the demand for image restoration, or restoration of image quality. Image restoration or image de-blurring is a process of restoring the original image by modelling the degradation or applying the inverse process which can only be done through the prior knowledge of the amount of degradation. Degradation or blurring can occur in two different ways; shift variant and shift invariant. The main differences between the two are the ways blurring affects on different parts of the image. The blurring can be approximately modelled as shown in the equation (2-6) (Vollmerhausen and Driggers, 2000).

$$g(x, y) = H[f(x, y)] + n(x, y) \quad 2-6$$

where

$g(x, y)$	is the blurred pixel in the digital image.
$H$	the distortion operator or point spread function
$f(x, y)$	pixel of original or undistorted image or restored quality pixel
$n(x, y)$	additive noise, introduced during image acquisition.

In ideal image acquisition, any point from image space is captured in a digital image and projected to the image-plane as one point. However, due to many reasons this might not be the case and the object point spreads out over several points in the image plane. The spreading takes place when the object scene is blurred by a number of factors. The spreading of a point over the image plane is called Point Spread Function (PSF), examples shown in figure 2.19 (ibid).



Figure 2.19: PSFs examples, original image, out of focus, and linear motion blur (ibid)

There are many algorithms that deal with image restoration or image de-blurring. An example of these algorithms is de-blurring an image using a blind deconvolution algorithm. In this technique, the image is restored using an initial PSF value. Many different methods were attempted to restore the original image without knowing the PSF, but most of them had very little success. Methods for estimating PSF are usually known as blind deconvolution because deconvolution is performed without knowledge of blurring function. As can be seen from equation (2-6), in addition to  $H$  or PSF there are two additional unknowns (original pixel and additive noise). In most blind deconvolution approaches, they assume the noise to be Gaussian (Levin et al., 2009). Using a blind image de-convolution approach is the approximate solution and is important when no prior information is available about PSF. To further improve the accuracy, the initial PSF for example can be roughly measured using inexpensive gyroscopes and accelerometers to measure a camera's acceleration and angular velocity during an exposure (Joshi et al., 2010). There are some available free tools such as SmartDeblur that can be used for image de-blurring based on blind de-convolution (Yuzhikov, 2014). For all other restoration and de-blurring algorithms, please refer to (Matlab, 2014a).

## 2.5 Point cloud generation techniques

The interest in 3D models is growing due to the rapid development of information technology, the availability of cheap high-quality digital cameras, the advancement in the survey instruments, and the computational power of personal computers. 3D surface models include Digital Terrain Model (DTM) generation from the filtered digital surface models or point cloud generation. Point clouds are groups of data points in any known coordinate system where each point has three-dimensional coordinates. These points are typically distinct by X, Y, and Z coordinates. A triangulated mesh and contour lines computation can be carried out on a point cloud to generate a DTM, or vice versa (computation can be carried out to produce a point cloud from a triangulated mesh).

Many methods are available for the collection of point clouds, the first and the more traditional one is through the use of a total station, but many other techniques are available such as 3D modelling using images, laser scanning techniques, Global Navigation satellite systems (GNSS), and satellite radar interferometry. Total station and GNSS can produce accurate and comprehensive models through 'walk over' surveying techniques. However, dense surface modelling requires a substantial increase in the number of data points, which substantially delays the process of surveying. Laser scanner devices are able to generate highly accurate, high-resolution models using short



wavelength light pulses, but also remain expensive, large, heavy, and with low precision for texture measurement. So, digital elevation models from photogrammetric techniques remain the most cost effective, rapid, and reliable method of obtaining more dense models than the other methods (Peng et al., 2009; Hudzietz and Saripalli, 2011; Neitzel and Klonowski, 2011). This section will cover the currently available techniques for point cloud generation.

### **2.5.1 Total stations**

A total station is an efficient and accurate tool of measuring angles and distances. In the field, this instrument is capable of distance measurement through utilizing Electronic Distance Measurement (EDM) and angle measurement through utilizing an electronic theodolite. In addition, it has the ability to perform computations in the field and produce a 3D (X, Y, and Z) coordinates of points through utilizing microprocessor on-board. It works in most environments and both outdoors and indoors. Usually the collected points can make clouds, and the density of the point distribution is mainly based on the specifications required for particular purposes. This tool is capable of providing features in 3D space and making a DTM of the site. It is also very efficient at obtaining accurate target point locations and establishing datum points in a topographical survey (Leica, 2006b).

There is a wide range of total stations that have different operation ranges and precisions which can seriously affect the price of a total station. The range of distance measurement can reach up to 3000 m, and an average cost of total station can range between £3,000 to £15,000 depending on the utilized features such as precision, measurement range, reflectorless characteristics, and data roaming and advance data communications. It can also be integrated with GNSS to provide positions with centimetre accuracy which can save time by avoiding the use of control points. Usually a total station survey is a two person operation, but recent advances in technology have allowed it to be used by one operator as in the case of the Leica icon robot 50 robotic total station (Leica Geosystem, 2014b).

The accuracy of a total station is dependent on many sources of errors including blunders that can be avoided through careful handling. Collimation errors or line of sight errors (Horizontal and Vertical) are the main source of instrumental error which can be eliminated by observing on two face measurements. Other sources of errors that are related to atmospheric effects can be reduced, but not fully eliminated, as it depends on the temperature and water content over the distance range. The accuracy of distance

measurement of a total station is normally in the linear form of  $\pm (\alpha \text{ mm} + \beta \text{ ppm})$ . The constant ( $\alpha$ ) is length independent and is the sum of instrument internal errors that are beyond the control of the user.  $\beta$  is the range dependent term which is based on parts per million, for example 1mm for every 1 km. Normal specifications of a total station vary from  $\pm (2\text{mm} + 1\text{ppm})$  to  $\pm (5\text{mm} + 5\text{ppm})$ , and angular accuracy can range between 1" to 10" (Uren and Price, 2010).

### **2.5.2 Global navigation satellite systems (GNSS)**

Global Positioning System (GPS) is an example of a GNSS and is a range based navigation system that offers operation in 24 hours and in all weather conditions. GPS was designed to work fully in a constellation of 24 satellites distributed in 6-orbital planes in inclination of about  $55^\circ$  with the equator. The orbit of the GPS satellites is approximately 20200 km above the surface of the earth. The position of satellites in orbit is calculated through distribution of ground control station all over the world. The position (X, Y and Z) in addition to the clock offset of the user unit can be measured using triangulation to a minimum of four satellites (Kaplan and Hegarty, 2006).

GPS positioning is based on measuring the transmitted time of radio wave signals emitted from satellites. For a receiver to compute its stand-alone position, it must be in view of at least four satellites. This method is widely used in navigation applications and has sub-metre accuracy, less than 15 m at 95% confidence level in the simplest pseudo range operation (Garmin, 2004). However, with the current constellation and advancements in technology, the accuracy of positioning can be further improved to reach centimetres or millimetres level by either, relative to a known position by double differencing, or in precise point positioning (PPP) mode. The accuracy of GPS positioning can be affected by many parameters such as the surrounding environment, multipath, number of satellites, satellite geometry, distance from the reference station, ionosphere/ tropospheric condition in addition to the quality of GPS receivers. For more information about each of GPS utilized techniques and the accuracy limits of each method, please refers to (Kaplan and Hegarty, 2006).

### **2.5.3 Laser Scanners**

The use of total stations and GPS for point cloud generation is not sufficiently dense and for some application, such as a road surface's texture depth measurement may not be practical. Therefore, other instruments and techniques should be investigated. In laser scanning, a laser beam is emitted and projected onto rotating mirrors whose

horizontal and vertical angles are recorded. The range that the emitted signal is projected and reflected from the object is calculated using either the time of flight or phase based measurement. For more information about the principle of laser scanning, please refer to Pfeifer and Briesse (2007).

Laser scanning can be classified in terms of method /concept of distance measurement into three categories; pulse round trip measurement (time of flight), phase based measurement, and a laser-based triangulation or laser profile. Most of the available terrestrial laser scanners utilize the time of flight or phase based measurement. The range, data rate, precision, and scanning speed are the main differences between the two techniques. Time of flight tends to be more suitable for longer ranges than the phase based methods but with low data rate productivity and scanning speed. Usually, the time of flight data tends to be noisier than the phase measurement. Although the resolution of point clouds obtained from the two techniques is very low, reaching 1mm or less depending on the range (Callieri et al., 2009), it is not preferred for road surface monitoring and texture depth measurement due to the low precision of the laser scanning measurements being in the range of  $\pm(0.4-10)$  mm (Leica Geosystem, 2014a) depending on the distance measurement technique, colour of the reflected surface, the available illumination, and many other parameters.

As previously mentioned, there is another concept that laser systems work on which is called laser-based triangulation or laser triangulation sensors. These types have been used for application that requires the closer range and for small scenes that require a high precision. The triangulation laser technology may reach precision of  $\pm 0.020$ mm (Faro, 2014). The basic idea of laser triangulation is the use of an active stereoscopic technique that is based on the principle of forward intersection for topographic creation. It is capable of determining the position of an object in space through projecting it to a Charge Couple Device (CCD), as can be seen from figure 2.20 (Bitelli et al., 2012).

As figure 2.20 shows, the laser emitter produces a beam of energy which is projected to the mirror and emitted from the instrument at angle ( $\alpha$ ). The laser hits the object's surface at the point (A) and is reflected back with a magnitude depending on the surface characteristic. The reflected laser signal is received by CCD/ Complementary Metal–Oxide–Semiconductor CMOS sensor at known distance (b) from the emitter. The angle ( $\beta$ ) which is unknown can be calculated using the trigonometric formulas listed in equation (2-7). With the knowledge of focal length (c) and location (Px, Py) of point (A)

on CCD, the object space point ( $X_A$ ,  $Y_A$ , and  $Z_A$ ) can be calculated from equation (2-8).

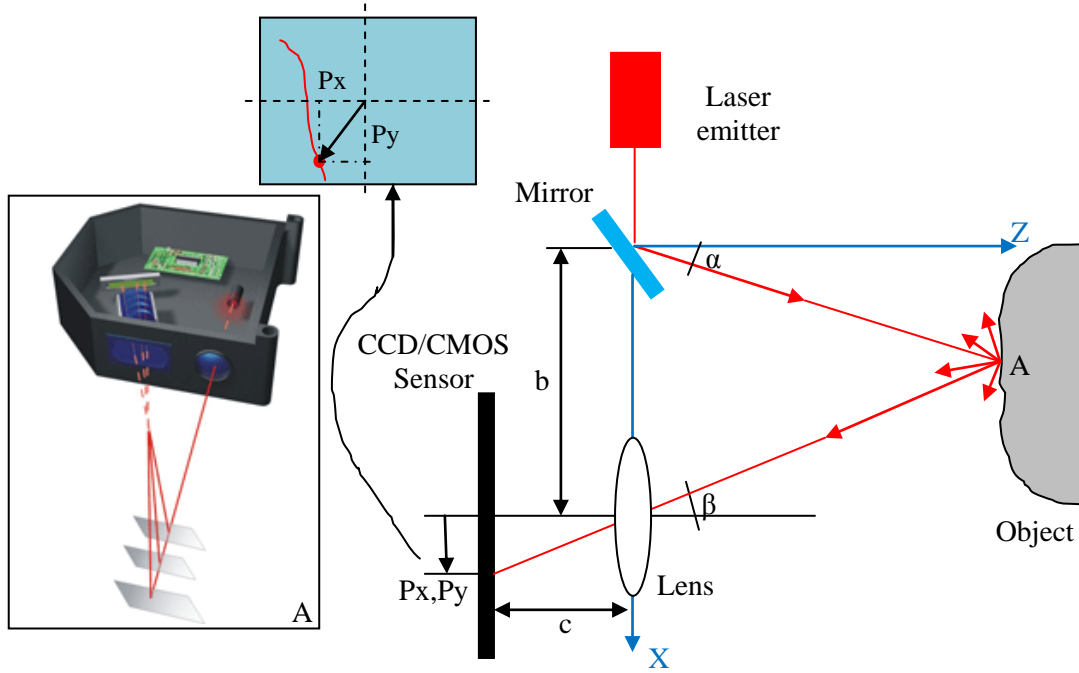


Figure 2.20: Principle of laser triangulation (left), (Bitelli et al., 2012) and sample of laser sensor (A), (Micro-Epsilon, 2014)

$$\tan\beta = \frac{P_x}{c}, \quad \tan\gamma = \frac{P_y}{c} \quad 2-7$$

$$X_A = \frac{b}{1 + \frac{\tan\beta}{\tan\alpha}}, \quad Y_A = \frac{b}{\frac{\tan\alpha}{\tan\gamma} + \frac{\tan\beta}{\tan\gamma}}, \quad Z_A = \frac{b}{\tan\alpha + \tan\beta} \quad 2-8$$

Although the laser profile has been extensively used in the industry sector, their uses in the real world are limited due to the difficulty of controlling its axial movement. There is also another limitation of using the laser triangulation which is the limited range of object measurement in some of the low-cost sensors which can only reach up to 25mm, and the cost of the sensors rises with the increases of the measurement range and precision (Micro-Epsilon, 2014). The texture depth measurement of the road surface is not possible with the use of laser sensors such as those shown in figure 2.20 due to the mentioned limitations. Therefore, the use of other more expensive alternatives is the solution such as FaroArm laser or Circular Texture Meter (CT Meter) as figure 2.21 shows. For more information about these two instruments and their technical specifications refer to (Applied Pavement Technology, 2008; Faro, 2014). Due to the

limitations of the laser scanner and laser triangulation, other alternative method should be investigated which will be discussed in the next sections.



**Figure 2.21: Circular Texture Meter (left) and FaroArm-Platinum (right)**  
(Applied Pavement Technology, 2008; Faro, 2014)

#### **2.5.4 Using Structure from Motion (SfM) in Computer Vision (CV)**

In computer vision (CV), Structure from motion (SfM) denotes to the process of finding out the three-dimensional structure of an object by analysing local motion signals over time. In CV, SfM refers to the general phenomenon by which humans can produce 3D structure from the projected 2D motion field of a moving object. A very closely related technique to SfM is stereo vision. The latter is the most-distinguished process in the projective geometry technique applications in the CV field. In contrast to general motion, stereo vision assumes that there are only two shots of the scene (Jin-Tsong et al., 2012).

SfM works by including consecutive scenes one at a time when the camera is moving, and each view is registered by using the Fundamental (F)-matrix that is based on the epipolar constrained resection. The F-matrix describes the epipolar geometry in the form of 3x3 matrix of rank two that relates a pair of image points and defines the rotation, translation and scale changes, equation (2-9). A partial reconstruction is extended by computing the positions of all 3D points that are observable in two or more images using a triangulation technique. The process is planned in a way to track the extracted points of interest by constantly matching the adjacent images (Fisher, 1997) as shown in figure 2.22.

$$[x', y', 1] \cdot \begin{bmatrix} F \end{bmatrix} \cdot \begin{Bmatrix} x \\ y \\ 1 \end{Bmatrix} = 0 \quad 2-9$$

$x', y'$  : the correspondence points or matching points of first image.

$x, y$  : the correspondence points or matching points of second image.

$F$  : are 3x3 matrix elements

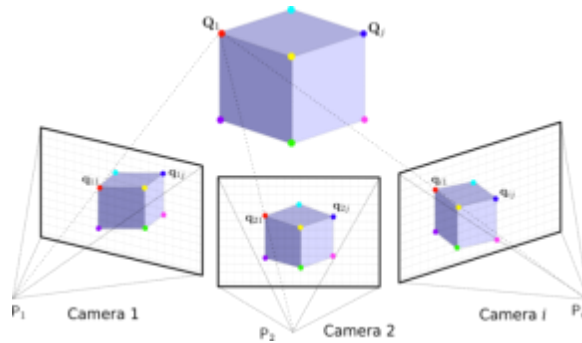
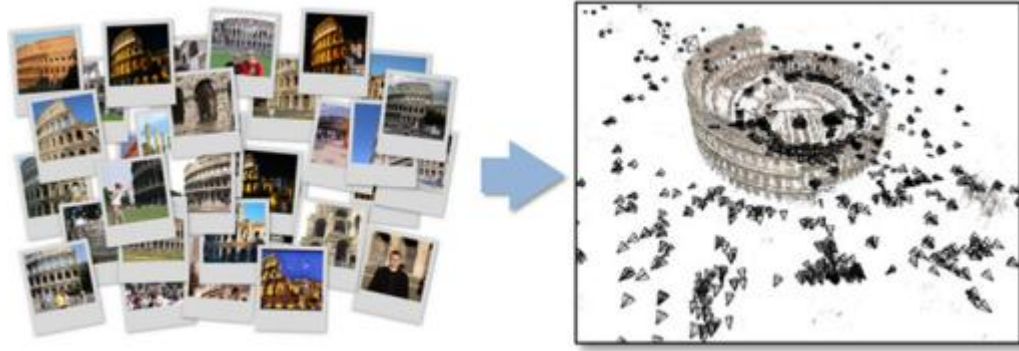


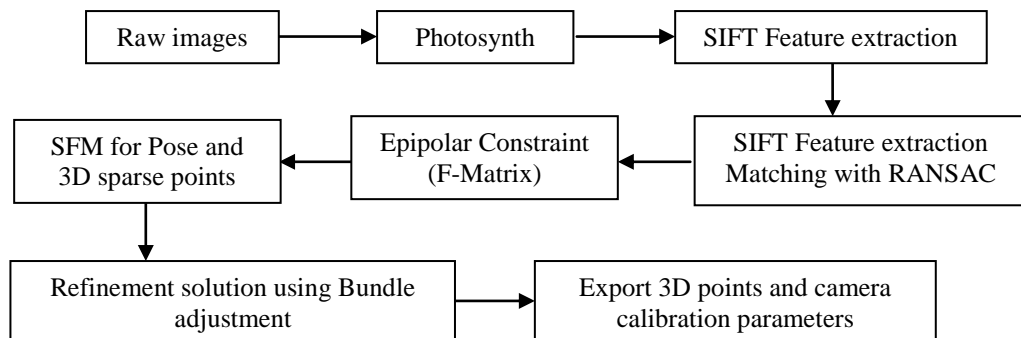
Figure 2.22: Merging images constantly using SfM (Michot, 2010)

The SfM technique is used extensively for researching in the CV community. Although used in CV, however it is still similar to the stereo vision that uses parallax between two images to create 3D models (Hudzietz and Saripalli, 2011). The first successful demonstration of the SfM technique applied to a real-world image was by Snavely (2010), who used a 3D reconstruction model from SfM by applying it to a set of Google and Flickr tourist images. The images were typically captured from hundreds of different cameras, scales, resolutions, illumination, and different times of day or seasons, see figure 2.23. SfM has been explored enormously in recent years, and has led to the development of various products and web services that reconstruct the 3D scene or models from unordered or different set of images (Snavely, 2010; VisualSFM, 2011; Photosynth, 2012; VISICS, 2012).



**Figure 2.23: Reconstruction of web photos (Google & Flickr) into 3D models (Snavely, 2010)**

Photosynth is one example that uses the SfM technique. It is a free internet processing tool run by Microsoft Live Labs and the University of Washington. This tool generates point clouds and the pose of a camera automatically from a series of unordered images. The two main processes of Photosynth include uploading the images to the service and then exporting the point cloud and the camera calibration parameters by using SynthExport. The procedures required for generating point cloud are shown in figure 2.24.



**Figure 2.24: Point cloud generation procedures (Jin-Tsong et al., 2012)**

Because the generated point clouds always contain outliers, the exported point clouds are recommended to be post-processed using a tool such as Geomagic studio. There are two main processing steps that are recommended to be applied to an exported point cloud. In the first step, the point clusters that are disconnected from the main body of the point cloud are removed. In the second step, the points that are not apart from the point cloud and further away with one standard error (66.667%) in z direction are removed. The final processed point clouds can then be used as a dependable source of DTM generation but with varying accuracy. This variation in accuracy might depend on many parameters that will be explained later in this project. Figure 2.25 shows the

possible procedures required for processing of the exported point clouds from Photosynth web service.

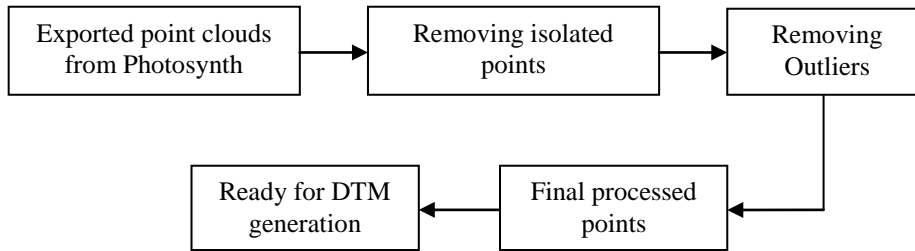


Figure 2.25: Overall processing of Photosynth point clouds using possible outlier removal tool

### 2.5.5 Using a Photogrammetric-based approach

The photogrammetric approach for the production of point clouds is not yet fully automated in most currently available commercial tools. The 3D point is determined based on defining interior and exterior orientation parameters and these parameters are then used with perspective geometry to produce object space point coordinates (X, Y, and Z) that appear in the overlapping areas between two or more images (Leica Geosystems., 2003).

Photogrammetry over many decades dealt with the precise 3D reconstruction of objects from images. Suitable commercial packages (Australis, Canoma, ImageModeler, iWitness, PhotoGenesis, PhotoModeler, and ShapeCapture) are available for close range photogrammetric solutions. However, these tools still require precise calibration and orientation procedures, and most are based on manual or semi-automated measurements. Recently, a new feature was added to Photomodeler scanner called SmartMatch. The SmartMatch feature can help in generating 3D (XYZ) coordinates of points automatically through searching out natural feature points within images and then matching them. The automatic matching process can eliminate the use of coded targets for the production of the 3D models (Remondino and El-Hakim, 2006; PhotoModeler®, 2009).

The operator needs to define some parameters in order to extract point clouds from photogrammetric software. A minimum constraint is required in order to produce 3D point clouds based on the photogrammetric approach. If the matching points or correspondence points are known, the relative orientation can be built between two images using the coplanarity condition. In the coplanarity condition (see figure 2.26), the



two perspective centres, image corresponding points, and any object space points all lie in a common plane. Theoretically, the  $\vec{b}$ ,  $\vec{R}_1$  and  $\vec{R}_2$  vectors should all be equal to zero, see equations (2-10) and (2-11). Each object point has one coplanarity equation and is independent of the object space coordinate system. At least five image points are required for solving the relative orientation between two images, for more information about the coplanarity condition refer to (Ghosh, 2005).

$$\vec{b} \cdot \vec{R}_1 \times \vec{R}_2 = 0 = F \quad 2-10$$

$$F = \begin{bmatrix} b_x & b_y & b_z \\ X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{bmatrix} \quad 2-11$$

Where

$\vec{b}$

Is the vector of baseline between perspective points ( $O_1$  and  $O_2$ )

$\vec{R}_1$

Is the vector between ( $O_1$  and  $P$ )

$\vec{R}_2$

Is the vector between ( $O_2$  and  $P$ )

$b_x, b_y, b_z$

Are projection of baseline on X, Y, and Z coordinates

$X_1, Y_1, Z_1$

Image point  $p_1$  with reference to image 1 coordinate system.  $Z_1$ =focal length

$X_2, Y_2, Z_2$

Image point  $p_2$  with reference to image 2 coordinate system.  $Z_2$ =focal length

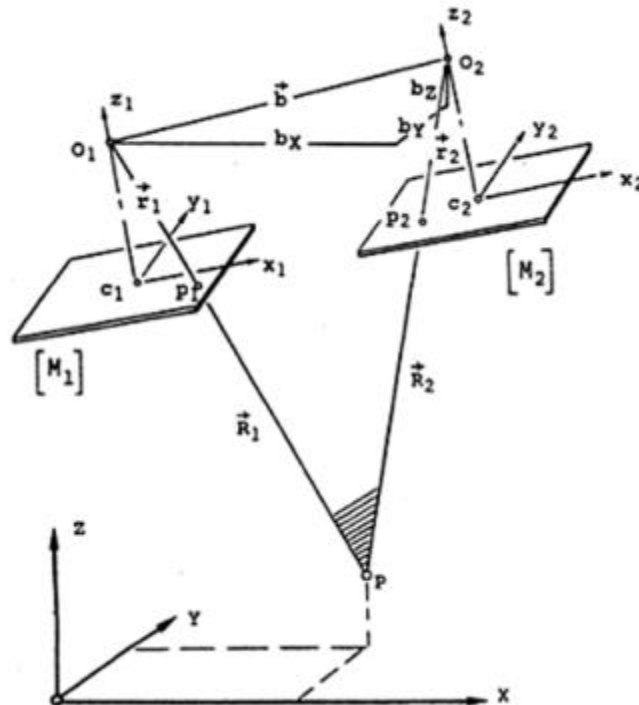


Figure 2.26: Coplanarity condition and intersection for relative orientation (Ghosh, 2005)

In Leica Photogrammetric Suite (LPS) for example, based on coplanarity condition equation and without knowledge of the exterior orientation parameters of the images, the operator needs to measure at least five tie points to solve for 5 unknowns ( $\omega$ ,  $\phi$ ,  $k$ ,  $B_y$ , and  $B_z$ ). This is required in order to obtain an approximate relative orientation before automatic tie point generation can be started. In addition to the minimum requirement for automatic tie point generation, Ground Control Points (GCPs) are also required for space point resection (Leica Geosystems., 2003). However, in the case of Australis and Photomodeler, there is no need for GCPs, and only tie points are required for the production of 3D point clouds, as a local coordinate system is adopted inside each of these tools.

In photogrammetry, several techniques are available for the production of 3D surface models, with each having unique strength and weakness characteristics. Stereo vision is one of the techniques used for 3D generation modelling. Stereo vision uses the parallax between two images to generate a 3D model (much like human vision), and requires placing two cameras at a known distance apart. However, due to the difficulty of placing two cameras with a required distance apart, overlapping images using only one camera (the camera moves to produce parallax) for production of 3D models is widely utilized in photogrammetry. The process finds the similar points (common points) between two or more images. The 3D coordinates of ground points (object space) can then be computed from image matched points by using the principle of photogrammetry. The most widely used approach is known as Collinearity equations, which relates the projection centre or perspective points, the ground coordinates of an object and the image coordinates through a straight line. The collinearity equations for a metric camera are given in equations (2-12) and (2-13). Non metric or digital consumer cameras do not normally fulfil the collinearity equations listed on (2-12) and (2-13). These two mathematical model equations should be extended to include principal offset points ( $x_p$ ,  $y_p$ ) and can be further extended to include lens distortion parameters.

$$x = f \frac{[r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)]}{[r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)]} \quad 2-12$$

$$y = f \frac{[r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)]}{[r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)]} \quad 2-13$$

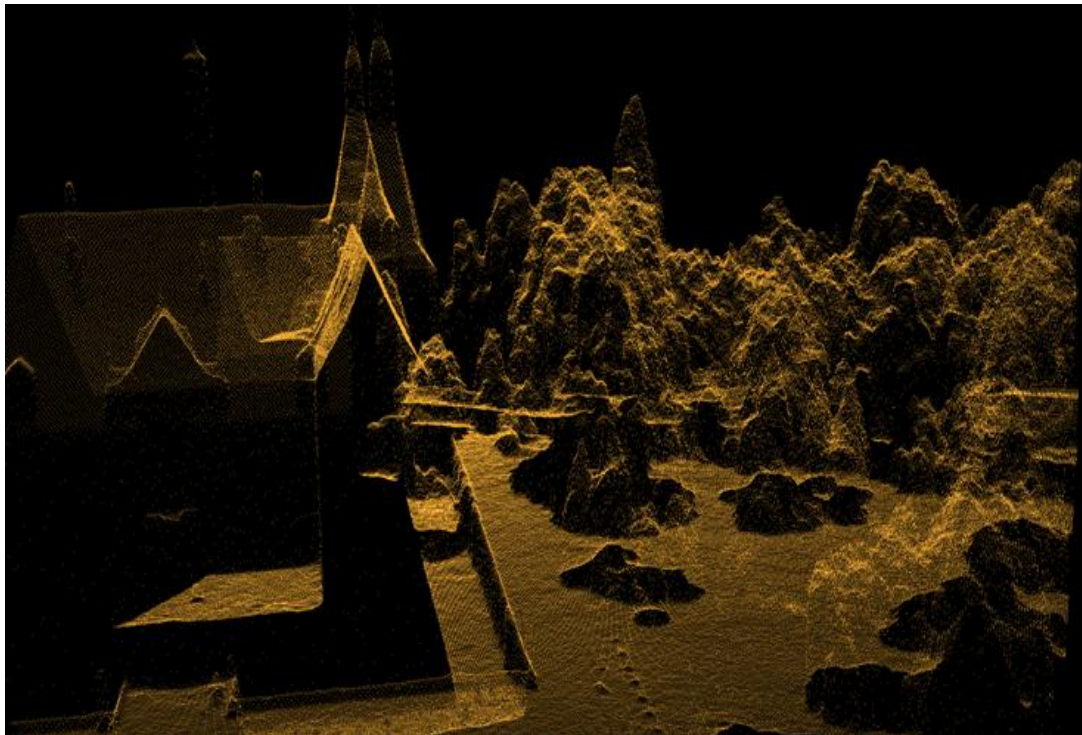
Where  $x, y$  = image coordinates

$X, Y, Z$  = ground or object coordinates

$X_0, Y_0, Z_0$  = projection centres.

And  $r_{11}, r_{21}, \dots, r_{33}$  are the elements of rotation matrix (Smith, 2009).

Due to advances in technology, a photogrammetric one click solution for 3D reconstruction is available today from Vexcel by Microsoft. This web service is designed for the rapid production of digital surface and terrain models with high-density point clouds (see figure 2.27). This software is specially designed for the UltraCAM digital aerial camera systems. According to the web-seminar from the developer of UltraCAM, the accuracy of the point cloud (X, Y and Z) for 15 cm resolution or Ground Sample Distance (GSD) is 10cm and 10-15cm for position and vertical accuracies, respectively. A comparison was carried out by the developer between a Laser Imaging Detection and Ranging (LIDAR) system and UltraCAM Eagle, using UltraMap for the production of point clouds. The results were highly promising for UltraCAM Eagle; the results of the comparison are listed on table 2.2. Gatewing cloud solution is another web-based service from Trimble. This web service has the ability to create orthophotos and digital terrain models form UAV imagery. It builds a DTM through automatically processing the images based on the users' requirements (UltraCAM, 2011; Questexmedia, 2012). The advances in technology have opened the possibilities for new photogrammetric acquisition techniques which will be explained in the next section.



**Figure 2.27: Screenshot of point clouds generated from UltraMap 3.0, (UltraCAM, 2011)**

Table 2.2: Comparison between LIDAR system and UltraCAM Eagle, (UltraCAM, 2011)

<b>LIDAR System</b>	<b>UltraCAM Eagle</b>
<b>Effective pulse: 400 khz</b>	Image GSD: 15 cm
<b>Strip width 600m</b>	Strip width 3000m
<b>20% overlap</b>	60% sidelap
<b>Used strip 480m</b>	Used strip 1200m
<b>Point density 10 pt/sqm</b>	Point density 300 pt/sqm
<b>Project 120 km<sup>2</sup>, Length 40 km, Width 30 km, Speed 120 kts</b>	Project 120 km <sup>2</sup> Length 40 km, Width 30 km, Speed 120 Kts
<b>63 flight Lines, 11 min per line, 3 min/turn</b>	25 flight lines, 9 min per line, and 3 min / turn
<b>882 mins, 15 hrs, 3 days</b>	300 mins, 5 hrs, 1 day

### ***2.5.5.1 Image Acquisition***

In all photogrammetric projects, image acquisition is the most influential and expensive step. The primary aim for many photogrammetric researchers and practitioners was to develop flexible and efficient techniques to obtain high-resolution imagery with the minimum possible cost. The acquisition technique consists of a platform and a camera, which are described in more details in the following sub-sections.

#### **2.5.5.1.1 Platforms**

There are different photogrammetric measurement platforms available for image acquisition such as satellites, traditional aircraft, mobile vans, and UAV platforms. Remote sensing satellites have temporal and spatial resolution limitations. Traditional aerial platforms could be used to obtain high image quality, but the cost of using large and medium format cameras is a constraint which cannot be afforded for many applications. Mobile vans can be a reliable source but are also limited to close range applications. UAVs combine both aerial and terrestrial photogrammetry which can open various new applications in the close range domain, and also introduces low-cost alternatives to the classical manned aerial photogrammetry. Without endangering human life, UAVs can be used in high risk situation or harsh environment.

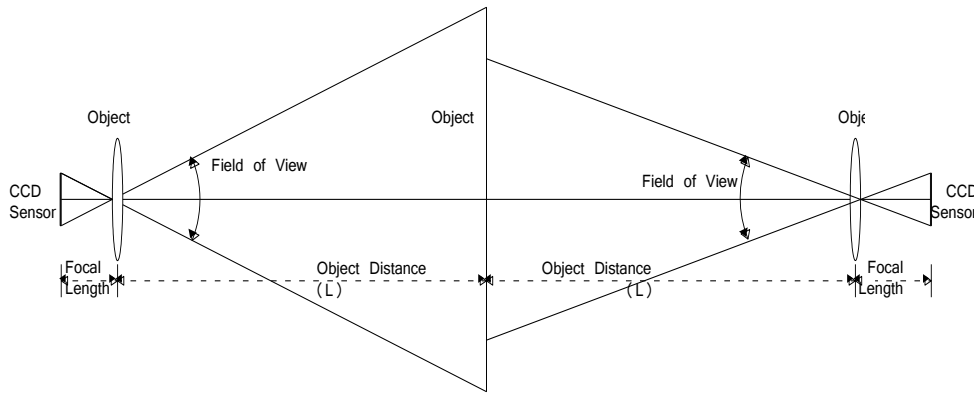
There are several reasons that made a UAV an extremely useful source of platform for image acquisition in many situations. The capability of image and video acquisition in real-time, rapid response in a number of critical situations, ability to collect image data at a lower cost, in addition to the high-resolution imagery acquisition, are all considered to be significant advantages. However, UAVs, especially low-cost types and

those limited to the carrying payload, have low accuracy of defining their position & orientation. UAVs are normally not equipped with air traffic communication and collision avoidance systems. Another limitation might also be the small coverage footprint on the ground and the necessity of the presence of at least two persons for system manoeuvres and transportation. It can also be a real threat to the people on the ground in case of malfunction (Zhang, 2008b; Nagai et al., 2009; Agüera et al., 2011; Briese et al., 2011; Haala et al., 2011; Neitzel and Klonowski, 2011; Remondino et al., 2011; Thamm, 2011).

In the UK, according to the current regulations, UAVs must only be used in segregated airspace, which restricts them to fly below 400 feet and in a few designated areas for testing purposes only. It was anticipated that by the end of 2012 UAVs will be technically capable of flying in unsegregated airspace. This was due to the primary interest of ASTRAEA, which is a combination of leading aerospace companies supported by government technology strategy to permit UAVs to operate anywhere in the UK (ESS, 2011). However, the final regulations require obtaining permission from the Civil Aviation Authority (CAA) before commencing a flight within a congested area or in proximity to people or property (Civil Aviation Authority, 2013). Due to the limitations of UAVs, mobile vans remain a very dependable source for image acquisition for close range applications, particularly for road surface monitoring.

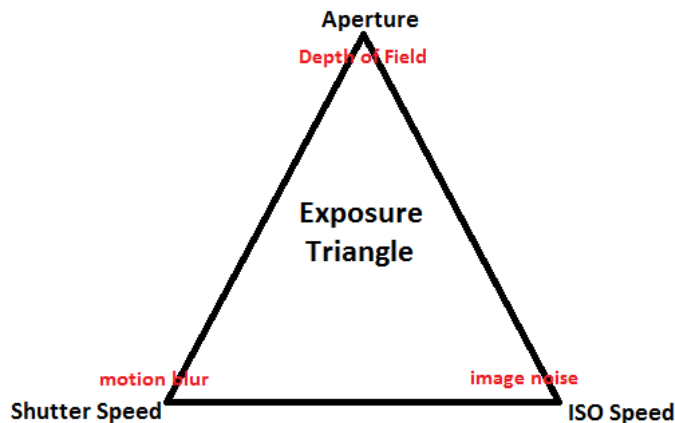
#### **2.5.5.1.2 Camera**

In road maintenance and crack detection, the camera has to be high resolution in order to be able to detect as thinner crack as possible. Some camera parameters can help to achieve higher resolution and produce excellent image acquisition. These parameters include Field of View (FoV), radiometry, exposure settings, and pixel size on the object (Ground Sample Distance (GSD)). The camera FoV defines how many scenes can be covered with the camera and this is a function of the focal length and the size of CCD or CMOS. The larger format sensor and smaller focal length have a larger FoV and vice versa. Different cameras with various camera focal lengths will be investigated in this project. Platforms with on-board sensors far away from their objects normally can view a larger area but cannot provide considerable details. The relationship between format size, focal length and the field of view is shown in figure 2.28.



**Figure 2.28: Diagram showing the relationship between Field of View and focal length**

The second parameter that can affect the quality of the image is radiometry. Radiometry is a critical parameter due to the fact that image matching process is dependent to a high extent on image radiometric quality. The major radiation components affecting the object and entering the sensor of a DSLR camera in a traditional aerial, low-altitude UAV, and mobile van are: the direct sun illumination, skylight, and illumination reflected from surrounding objects in the area. The radiation component dominating from objects is reliant on the state of the atmosphere and shadowing conditions. The amount of light reaching the CCD sensor or detector with consumer cameras is controlled by exposure time, f-stop, and International Standards Organization ISO sensitivity settings. Fundamental quality indicators of the camera are sensitivity to lighting, stability of the sensor, and the sensor effect by Signal-to-Noise Ratio (SNR) (Rosnell T et al., 2011). The exposure process includes the amount of light entering the CCD. Three elements are involved in exposure which makes a triangle called the exposure triangle, see figure 2.29. The high quality image acquisition is a balance or a wise selection of these three elements.



**Figure 2.29: Exposure Triangle (CiC, 2012)**

The image resolution or GSD is another important parameter in the camera, especially for road surface monitoring, with the larger the GSD, the more details can be captured by the camera see table 2.3 (Wang and Smadi, 2011). Advances in technology have allowed for the advent of new instruments with powerful specifications. Recently, Nokia mobile phone has revealed its new state-of-the-art Nokia 808 purview phone with 41 megapixel (MP) sensors. This new technology could be considered revolutionary, especially for micro drones because of their limited payload, using this new-light camera with a powerful resolution will further broaden their applications (BBC, 2012).

Table 2.3: Image resolution and theoretical visible crack width(Wang and Smadi, 2011)

<b>Traverse Resolution (pixel)/4m lane</b>	<b>1300</b>	<b>2048</b>	<b>4096</b>
<b>Visible Crack Width (mm)</b>	3	2	1

### 2.5.6 Comparison between photogrammetry and Computer vision (CV) based 3D surface modelling

In recent years, different solutions were available for automated processing of images for the derivation of 3D information or 3D models. Examples of automated 3D models generation are the open source Microsoft Photosynth, Bundler, Photofly, and ARC3D. These solutions are based on the SfM derived from CV. An investigation of the performance of these public sources is necessary in order to determine their reliability. Some data set images were used for making a comparison between both photogrammetric and vision approach for 3D surface modelling in this project. The procedures that both communities share are; calibration, image matching, and inverse triangulation with bundle adjustment for 3D point cloud generation. All of these processes were investigated separately in more detail in order to conclude the limitations of each method in creating the 3D surfaces. A comparison of different available tools from both communities was evaluated in this project. Both commercial photogrammetric software such as LPS, Photomodeler, Bundler, AGISoft StereoScan and open source tool (such as Microsoft Photosynth) were utilized for analysis evaluation in the next chapters.

There are some problems associated with SfM or CV's 3D reconstruction programs and solutions. Although, they are easy to operate and can provide a fascinating 3D point cloud out of unordered images automatically, these solutions do not have photogrammetric mathematical rigour but are appropriate for visualisation purposes only. In contrast, for achieving high accuracy in addition to photorealism, user interaction is considered essential in order to provide scale, e.g. by measuring GCPs or introducing scale by the use of a scale bar. Using both communities, i.e. the best of the

both photogrammetry and computer vision, it should be possible to have tools, fully automated and precise, allowing 3D modelling of the complex environment from digital images (Pierrot Deseilligny and Clery, 2011). The next section will cover the main differences between photogrammetry and computer vision techniques in some of most important parameters, such as camera calibration and image matching.

#### **2.5.6.1 Camera Calibration**

Accurate camera calibration is considered as a necessary prerequisite for the extraction of precise and reliable 3D metric information from images. The topic of camera calibration has been widely researched in photogrammetry (Abraham and Hau, 1997; Clive S, 1997; Zhang, 2000; Läbe and Förstner, 2004; Remondino and Fraser, 2006). The camera calibration is used to determine precise and accurate values, called interior orientation parameters. If the principal distance, principal point offset and lens distortion parameters are known, the camera is considered to be calibrated. In general, camera calibration methods may be classified into three primary classes; laboratory methods, field methods and stellar methods (Wolf and Dewitt, 2000).

The number of existing low-cost digital consumer cameras has significantly increased over the last decade. The stability of consumer camera calibration parameters is crucial and needs to be considered due to the excessive use of camera features such as zoom or auto focus. The change of calibration parameters over time was investigated by Läbe and Förstner (2004). In their study, different camera types, which cover the spectrum of consumer cameras, were used. The calibration of cameras was repeated over a period of four months from three to seven times, and no systematic temporal increase or decrease of calibration parameters was observed. The study concluded that consumer cameras, under certain limited accuracy requirements, can be used for photogrammetric purposes but have to be calibrated periodically if camera features such as focusing were used (Läbe and Förstner, 2004).

In computer vision and many other applications, the calibration is only applied to recover the focal length as other parameters are neglected or assumed known. In photogrammetry, all calibration parameters are employed for precise photogrammetric measurements. The focus of treatment can be so different in each of these two communities regarding calibration approaches. In photogrammetry, the calibration requirement might be designed to sustain a successive object space measurement demanding say 1:20,000 of the accuracy. On the other hand, the prerequisite for a



structure from motion applications may require positioning object points to the accuracy of only 5% of the camera to object distance (Remondino and Fraser, 2006).

The photogrammetric community add more emphasis on the accuracy and reliability of the camera calibration than the CV community. Although the camera calibration is a necessary step in 3D modelling in order to extract metric information from 2D images, the automation of the 3D modelling is considered more important in CV. In photogrammetry, however, accurate camera calibration is an essential requirement for the extraction of precise and reliable 3D metrics without putting an emphasis on the automation. Much work has been done in the photogrammetric community and recently, full automation using target-less features has become available (Barazzetti et al., 2011). Camera calibration can be classified roughly into two categories based on the used techniques of photogrammetric calibration and self-calibration (Zhang, 2000; Remondino et al., 2012), table 2.4 shows the main differences between both communities for camera calibration.

Table 2.4: Camera Calibration in Photogrammetry and CV (Remondino et al., 2012)

item	Photogrammetry	Computer Vision
<b>Geometry</b>	Perspective Geometry	Projective geometry
<b>Parameters</b>	Interior Orientation (IO): principal point (2 parameters), principal distance (1), radial lens distortion (2-3), tangential distortion (2), affinity (1) and shearing (1).	Interior Orientation (IO): are limited to the principal point (2), principal distance (1) and radial lens distortion (1-2) coefficients
<b>Calibration method</b>	Utilize all methods	Self- calibration only
<b>Automation</b>	Manual and Automatic using coded or target-less(recently)	automated using SfM without target

A camera model based on perspective projection considers the Interior Orientation parameters (IOPs) to be stable (at least for a given focal length setting) and that all departures from Collinearity, linear and non-linear, can be accommodated. Generally, the collinearity equation requires a minimum of five correspondence points within a multi-image network. Moreover, approximations of parameter values for the least-squares bundle adjustment in which the calibration parameters are recovered are necessary due to its non-linear nature (Remondino and Fraser, 2006). On the other hand, the projective geometry is more supporting rather than Euclidean scene reconstruction with projective camera model. The Essential matrix and fundamental matrix models are characterized by such a model and can accommodate some variables, such as unknown focal lengths. This model requires a minimum of 6 - 8 correspondence points to enable a linear solution that is invariable although quite unstable. In such models, non-linear image distortion parameters, such as lens distortion, are not easily dealt with (Remondino and Fraser,

2006). Usually in CV, the self-calibration approach is utilized for defining the camera calibration parameters. However, not all parameters are considered important for the CV community (Pierrot Deseilligny and Clery, 2011). In some approaches like Photosynth, only focal length and two radial distortions (K1 and K2) are worked out.

In the photogrammetric approach, camera calibration is achieved by observing a calibration target with known geometric 3D spaces to a high precision. Examples of these high precision geometric observers are a flat sheet in case of PhotoModeler Scanner (PMS), LCD chessboard in case of AGISoft Lens, and the frame based coded target calibration in case of Australis. AGISoft lens is a free tool that allows for automatic calibration of cameras through utilizing LCD screen target as a calibration sheet. It provides a full assessment of camera calibration matrix, comprising coefficients of non-linear distortion by means of brown's model (Agisoft, 2015).

An example of CV camera calibration is Photosynth, which is easy to use and can produce an attractive 3D point clouds out of unordered images. However, the service is incomplete in terms of photogrammetric rigour in the formulation of equations, which may lead to unacceptable precision for some applications, for example, large blocks of images (Remondino et al., 2012). Photosynth allows each image to have its focal length; this is traditionally the cause of imprecision. The proposed calibration model by Photosynth is relatively simple, as it is limited to three coefficients: focal length and radial distortion (K1 and K2). Some precise aero-triangulation of a large set of images acquired by consumer cameras from UAVs requires the possibility of more sophisticated distortion rather than the focal length and two radial distortions. For example, the de-centric distortion and sometimes arbitrary polynomial distortion is necessary for these types of cameras (Pierrot Deseilligny and Clery, 2011).

### **2.5.6.2 Feature Matching**

Digital image matching can be divided into three general categories that are feature-based, area based and symbolic or hybrid method. Image matching on area based methods is straightforward and achieved through numerically comparing the digital numbers in small sub-arrays of each image. On the other hand, feature-based methods are quite complicated and involve the extraction of features, which include edges at different scales (Wolf and Dewitt, 2000). In recent years, in the photogrammetric field, interest in feature-based matching, which is widely used in computer vision, is increasing due to the availability of new acquisition techniques. The techniques that do not observe normal acquisition conditions are images acquired by UAVs, mobile mapping

technology, and oblique photogrammetric cameras. Examples of those image matching methods are the two operators Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF) or its newer version OpenSURF (Lingua et al., 2009).

In the literature, a comparison between SIFT and SURF, the previous version of OpenSURF, was carried out by Juan and Gwun, (2009). The study used KNN (K-Nearest Neighbour) for finding the matches and RANdom SAMple Consensus (RANSAC) to reject inconsistent matches of these two methods. The repeatability measurement was used to evaluate the performance of detection and was applied to compare the two methods for scale changes, rotation, blur, illumination changes and affine transformations. The overall conclusion of the study that was SIFT is slower than SURF and is not good at illumination changes. SURF is also not good in illumination changes, however, in time of processing SURF outperforms SIFT, see table 2.5. While SIFT is invariant to rotation, scale changes and affine transformations, SURF is not stable in rotation. The performance of both tests for rotation changes is shown in figure 2.30. Overall process of both algorithms SIFTS and SURF is summarized in table 2.6.

Table 2.5: SIFT and SURF general performance levels (Juan and Gwun, 2009)

Method	Time	Scale	Rotation	Blur	Illumination	Affine
SIFT	common	best	best	best	common	good
SURF	best	good	common	good	best	good

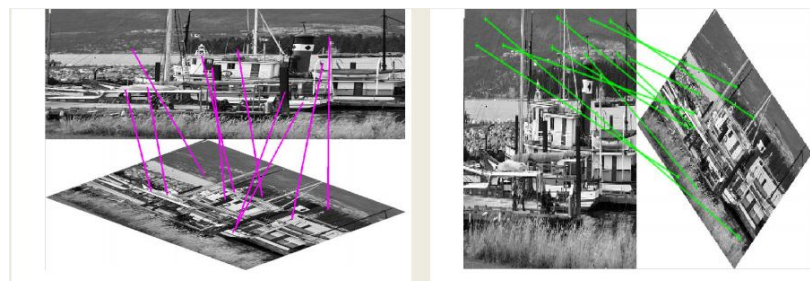
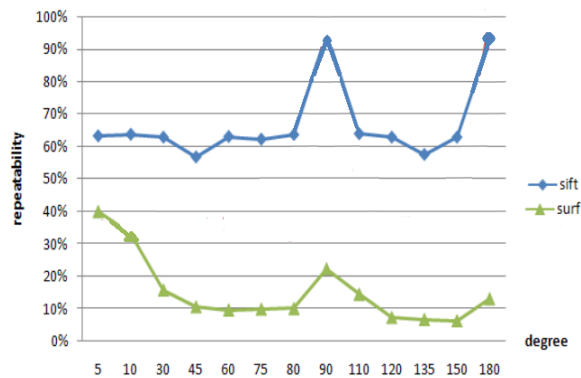
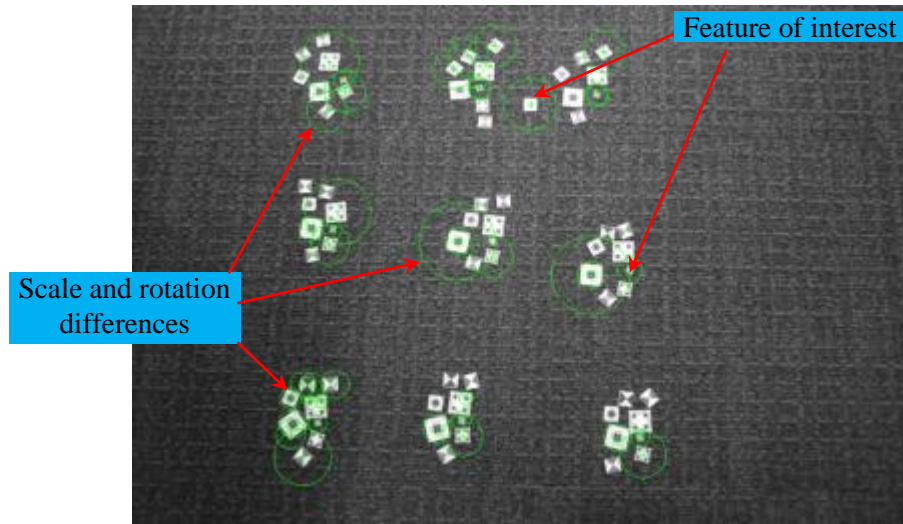


Figure 2.30: Rotation comparison between SIFT and SURF at different angles, repeatability or frequency vs rotation (above) and the used samples (below) (Juan and Gwun, 2009)

Table 2.6: SIFT and SURF workflow (Lowe, 2007; Bay et al., 2008)

SIFT	Detection	scale-space, extrema detection	Uses Difference of Gaussian (DOG) instead of Gaussian to improve the computation speed. This step is to detect the best candidate key-point. Laplacian: for region with rapid intensity change (edge detection). First been smoothed with something approximating (Gaussian) reduce its sensitivity to noise. Single grey level image as input and produces another grey level image as output.	
		Key-point localization	Too many candidates, discarding of low contrast key-points and eliminating edge responses. Interpolation of nearby data is used to accurately determine its position.	
	Description	Orientation assignment	In this step, each key-point is assigned one or more orientations based on local image gradient directions. Describe the key-point with respect to this orientation. 36 orientation bins (representing 10° increments) is computed. Largest bin is chosen as the key-point's orientation.	
		Key-point descriptor	The area around the Key-point is divided into 4 x 4 sub-regions, Build an orientation histogram with 8 bins for each sub-region. This results in a vector with 128 dimensions (4 x 4 x 8).	
SURF	Detection	scale-space, extrema detection	Fast Hessian Detector	The Hessian matrix uses to classify local maximum of pixel intensities. Calculate the determinant of the Hessian for each pixel in the image and use the value to find interest points.
			Constructing the Scale-Space is a continuous function which can be used to find extrema across all possible scales. Different idea of scale space applied here. The SURF approach leaves the original image unchanged and varies only the filter size. Due to using integral images, filters of any size can be applied at exactly the same speed.	
	Description	Key-point localization	Accurate Interest Point Localisation It done through three stages. The responses are thresholded by comparing each pixel in the scale-space with 26 neighbours, and interpolating the nearby data to find the location in both space and scale to sub-pixel accuracy.	
		Orientation assignment	Haar wavelet responses of size 4 σ are calculated for set of pixels within a radius of 6σ of the detected point. σ is scale where interest point found. The dominant orientation is selected by rotating a circle segment covering an angle of 60° around the origin. Here the integral images is used with Haar wavelet to limit the computation steps to only 6	
		Key-point descriptor	Construct a square window around the interest point with 20σ and is oriented along the direction found on the previous step. As in SIFT, SURF uses 4x4 sub-regions, Within each of these Haar wavelets of size 2 σ are calculated for 25 regularly distributed sample points. Each subregion contributes four values to the descriptor vector leading to an overall vector of length 4 x 4x 4 = 64. If use 128 as in SIFT will be more discriminatory.	
	Much of the performance increase in SURF can be attributed to the use of an intermediate image representation known as the “Integral Image”.			

Dedicated area-based or feature-based algorithms with the extraction of points or regions of interest are used to identify the image correspondences on the overlapped area. Image matching is slightly different between the two communities. In CV, there are more interest in feature-based extraction using algorithms such as (SIFT and SURF) which provides highly distinctive features invariant to image scaling and rotations, figure 2.31 shows a possible extracted features and their scale in the image using SURF.



**Figure 2.31: Extracted features and their scale (SURF)**

By comparing the descriptors (64 or 128) for SURF and SIFT, respectively, the corresponding points can be found with an exhaustive analysis of all possible image combinations. Two different methods are available for comparing the descriptors, the first is rigorous but slow procedures (e.g. quadratic matching), and while the second is fast but approximate methods (e.g. kd-tree search) can be used. Therefore, it is normal to have lots of outlier in case of SURF and SIFT since they utilize the second approach.

In photogrammetry however, the area-based matching approach is still adoptable through numerically comparing the digital numbers in small sub array of each image or area based matching. It is not invariant to large rotation and scale changes and also needs initial epipolar geometry to narrow down the search area and to increase the accuracy of matching. The most used method is cross-correlation (C) see equation (2-14) and the Least Square Matching techniques which have been implemented widely in commercial photogrammetric software. Many researches have demonstrated their high reliability as well as accuracy, which can reach up to sub-pixel level (Lingua et al., 2009). In the cross-correlation method, the correlation coefficient of the grey value is computed

between the template window from the left image and the search window in the right as follow (Wolf and Dewitt, 2000).

$$c = \frac{\sum_{i=1}^m \sum_{j=1}^n [(A_{ij} - \bar{A})(B_{ij} - \bar{B})]}{\sqrt{[\sum_{i=1}^m \sum_{j=1}^n (A_{ij} - \bar{A})^2][\sum_{i=1}^m \sum_{j=1}^n (B_{ij} - \bar{B})^2]}} \quad 2-14$$

Where

- C : the correlation coefficient.
- m,n : number of rows and columns respectively in the template window.
- A<sub>ij</sub>, B<sub>ij</sub> : are the digital number from right and left image at row i, column j.
- $\bar{A}$ ,  $\bar{B}$  : are the average of all digital number in left and right image respectively.

Least squares' matching is closely related to the correlation method. Least squares estimate parameters that best fit a search window to a reference window. Because it is an iterative process, the least squares can generate a high positional accuracy to about 0.1 pixels. Cross correlation requires a prior less accurate position than the least square matching but usually limited in precision to about one pixel. Least square on other hand, requires more accurate prior positional accuracy to about 2 pixels but can achieve high precision. Therefore, in most practical cases, cross correlation is first used for prior position estimation and then followed by least square for achieving high accuracy (Leica, 2006b). The least square uses the equation listed on (2-15) for estimating the position of search window to about 0.1 pixel.

$$\begin{aligned} g_2(x_2, y_2) &= h_0 + h_1 g_1(x_1, y_1) \\ x_2 &= a_0 + a_1 x_1 + a_2 y_1 \\ y_2 &= b_0 + b_1 x_1 + b_2 y_1 \end{aligned} \quad 2-15$$

Where

- $x_1, y_1$  : The pixel coordinate in the reference window
- $x_2, y_2$  : The pixel coordinate in the search window
- $g_1(x_1, y_1)$  : The grey value of pixel  $(x_1, y_1)$
- $g_2(x_2, y_2)$  : The grey value of pixel  $(x_2, y_2)$
- $h_0, h_1$  : Linear grey value transformation parameters
- $a_0, a_1, a_2$  : Affine geometric transformation parameters
- $b_0, b_1, b_2$  : Affine geometric transformation parameters



Traditional feature extraction and matching techniques used in photogrammetric applications are generally inadequate for many new applications such as UAVs and low-quality image acquired using mobile platforms. These algorithms fail to produce reliable results under extreme geometrical conditions of the new applications. Photogrammetric detection and description algorithm represented by cross-correlation is not invariant to scale changes, rotation, blur, illumination changes and affine transformations (Lingua et al., 2009). Therefore, the needs for another alternative is vital in order to deal with the new acquisition techniques.

Figure 2.32 shows possible scenes for which computer vision matching algorithms can work perfectly while photogrammetric ones could probably fail to do so. A and H are the affine transformed images, B and C are the scale changed images, D are the rotation images, E and F are the blurred images, G are the illumination changed images (Juan and Gwun, 2009). For a comparison between the three different methods (SIFT, SURF, and cross-correlation with least square refinement), trials were designed for this project based on some available data sets to examine the accuracy of each method in terms of total RMSE, for more information refer to section (5.7.1).



Figure 2.32: Possible Images that SIFT & SURF can Work with(Juan and Gwun, 2009)

### 2.5.6.3 Outlier Rejection in Feature matching

There are many outliers in the matching solutions of both SIFT and SURF, since both uses the less rigorous matching classifiers, especially SURF. In order to remove possible outliers, robust estimators LMeds (Least Median of Squares), RANSAC (RANdom SAMple Consensus), and MSAC (M-estimator SAMple Consensus) are normally employed to validate the eppipolar constraint between pair of image. In the case of RANSAC, a model such as geometric transformation is built using a random sample of subsets (7 matches) through normalized 8 point algorithm. It then determines the inliers that are consistent with the model and outliers that are not. Finally, it is repeated to a predefined number of times or until the set of inliers is large enough, or by defining the distance threshold. Figure 2.33 shows the original SURF matching results and after applying RANSAC as the outlier rejection. It can be visually concluded from the results presented in figure 2.33 that the accuracy of the matching is massively improved after applying the RANSAC outlier removal algorithm. This result was achieved using 5000 iterations and with projection error (distance threshold) equal to or less than 0.01.

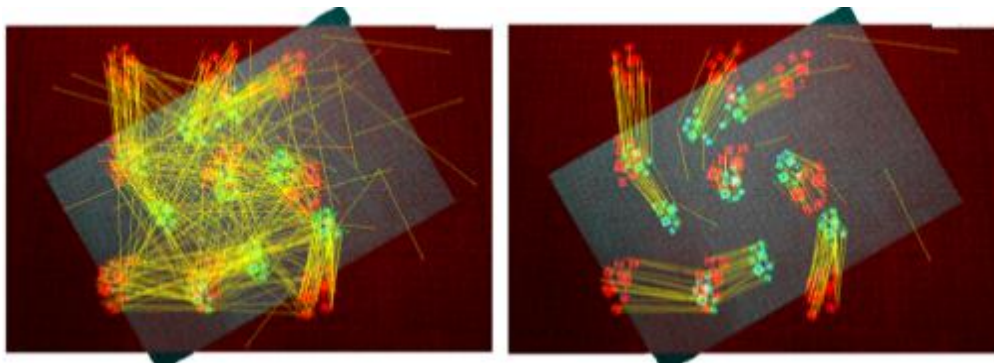


Figure 2.33: SURF only (left), SURF with RANSAC (Right)



## 2.6 Summary

In this chapter, the literature review revealed that various automated systems are already developed for road surface monitoring. These automated systems were developed to overcome the previous approaches to assessing pavement condition which involved labour-intensive, time consuming, and risky processing of data collection. The review also showed that automated systems are different from one country to the other, but with all sharing one objective which is accurate data collection and monitoring. It has been discussed that the cost of these automated systems is not affordable by many developing countries and local road authorities. Therefore, the need for other cheap alternative methods should be investigated in this project, see (chapter 4 and 5).

An extensive review on the literature has identified that previous techniques (sand patch testing (SPT) and Laser sensors) for characterising road surface texture have some limitations. Texture characterising is an important factor for assessing the ride quality, noise control and skid resistance. It has been proved that all previous techniques have limitations such as cost of the technique, limitations in providing satisfactory information, and limitation in the scanning range used to measure the surface. The most widely used classic approach for road surface texture depth measurement is SPT. It has been showed that this approach has a limitation in providing accuracy which is in the range of  $\pm 0.227$  mm of 95% confidence interval and is also a labour intensive process as well as having safety issues, in addition to the difficulty of using it in wet weather and windy conditions. The rapid advancement in 3D automation methods from digital images has opened the door for many applications that require precision such as roughness measurement. Research into 3D automation method for roughness measurement has revealed that no critical comparison between the automated 3D surfaces of texture and SPT are yet available. Therefore chapter 3 of this project deals with investigating this new approach for texture depth measurement and surface characterizing using digital images.

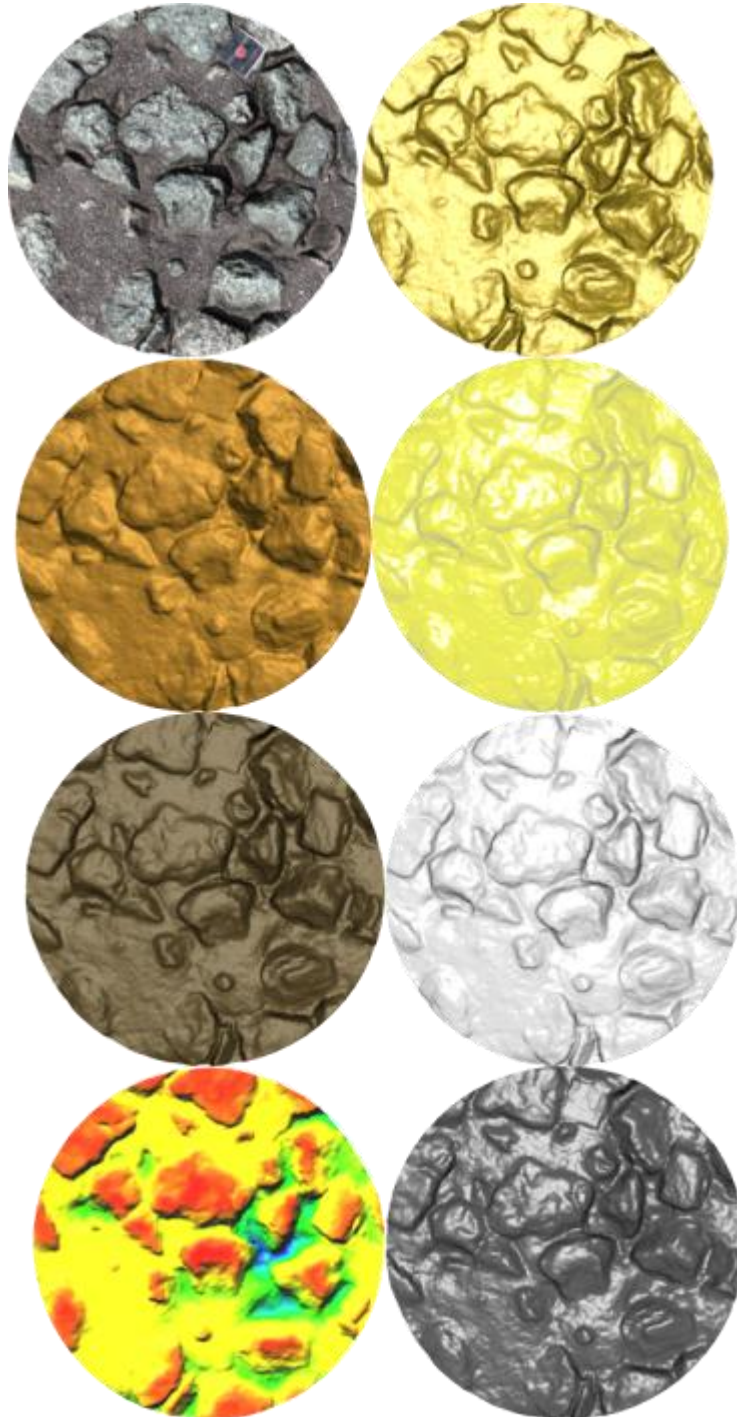
It has also been showed that although many other monitoring techniques are available, the images remain the only dependable and affordable source. Images can be used successfully for crack and other road surface defect detection. The literature review revealed that although detection of the cracks has been the aim behind many researches in recent years, many problems still need to be solved. For example, crack detection in rough texture surfaces needs more attention and investigation, as rough texture can easily hide crack pixels, and the processed image may contain lots of false positives due to

raised up aggregates. The crack detection on rough texture is investigated further in chapter 4.

Review on literature also showed that keeping roads safe is a challenging process for many road authorities. Damages on the road surface such as potholes can be a serious problem for many local government and road agencies. Claims are rise annually against road authorities and insurance companies for the potholes related problems. Millions of pounds are spent annually in the UK and Wales for compensation of motorist and for repairing the potholes. Therefore, periodic road assessments and collection of up-to-date information about the road surface condition seems to be the most effective way to conserve high road standards at a minimum price.

Many automated methods such as laser sensors, mobile network sensors, the use of accelerometers, and single digital image processing were developed, and reporting of potholes through a dedicated website is still available even in modern countries such as UK. Each of these developed methods have their advantage and disadvantages, for example, the pothole reporting method is mainly based on the users' involvement. On other hand, 2D image-based methods are easy to use; however, challenges remain in the methods of processing, especially when texture differences do not exist between potholes and the background. To overcome the above problems, an alternative 3D surface modelling has been proposed which will be discussed in chapter 5 of this project.

## CHAPTER THREE: NEW APPROACH FOR TEXTURE DEPTH MEASUREMENT



Prepared using render tool in MeshLab

### 3.1 Introduction

For pavement design and management, the measurement of pavement surface texture depth plays an important role. The pavement surface texture's depth measurement determines many important parameters in road management, such as ride quality and pavement durability as well as skid resistance at high speed and in wet conditions. When the road surface is wet, the risk of accidents is increased as the breaking distance becomes longer. Therefore, a mass of accidents in wet weather conditions can be an indicator of friction deficiency. The skid resistance of a road pavement is an important road safety factor which needs further attention (Kennedy et al., 1990; Viner and Britain, 2006).

Several methods are available for measuring the texture depth using different techniques. The sand patch test (SPT) is the most widely used and historic approach that determines mean texture depth (MTD) of pavement surfaces by applying and spreading a known volume of sand to cover a circular shape on the road pavement surface. Mean texture depth is calculated by measuring the area of the sand and dividing a known volume of sand by the area [ASTM E965]. More recent advances in technology have allowed for the use of laser sensors to calculate the depth through texture profiling. As an example, in the UK texture depth measurements made by laser systems are currently reported as Sensor Measured Texture Depth (SMTD) which is treated as an alternative to SPT (Viner and Britain, 2006). A few reported papers have used photogrammetric techniques for measurement of texture depth but these are limited to laboratory use only. Other papers work used the integration of laser measurement with photogrammetry for obtaining results (Gendy and Shalaby, 2007; McQuaid et al., 2010).

In this chapter, a methodology has been developed to determine texture depth measurement to be an alternative method to the SPT. The new developed method is based on reconstructing the 3D texture of the road surface using digital images. Figure 2.1 describes in details the workflow for the proposed methodology. This chapter will investigate the use of digital images for texture depth measurement and compare it with SPT. Smartphone cameras and DSLR cameras were used for the 3D modelling generated from digital images. A comparison is carried out based on real data collected from different road surfaces with different texture depths and levels of deteriorations. The results presented in this chapter show the promising outcomes of an alternative method to SPT using only digital images.

## 3.2 New approach-description

Extensive literature review showed that SPT tends to be time consuming and not always applicable, especially in wet and windy conditions. The review also showed that laser sensor system tend to be expensive and limited to a two-dimensional array that lacks some information (Bitelli et al., 2012; Jahanshahi et al., 2012a; Sengoz et al., 2012). Recent advances in technology have allowed for the introduction of a new imaging technique and a processing methodology based on a 3D automated method.

In this approach, a camera will be used for capturing the texture of the road surface. Different cameras will be utilized in this project to examine the effect of pixel size, focal length, and CCD sensor size. It is well-known in photogrammetry that cameras in close range applications can achieve ‘mm accuracy’ if well calibrated. However, the challenges remain in automatic processing of the calibration and 3D model generation. New smartphone device technologies are available which have a very compact camera (CCD and focal length) which add additional challenges to close range photogrammetric applications. 3D model reconstruction communities have appreciated these limitations that led to development of new automated methods of camera calibration and 3D surface model reconstruction. These new automated methods are the core of processing in this project.

Figure 3.1 shows a workflow diagram that describes in detail the proposed new method for surface texture depth measurement as follows:

1. The camera is used to capture the surface texture for measuring the depth. Five cameras, 2 of DSLR with 3 lenses and 2 smartphone camera were utilized for the overall investigation in this chapter.
2. Before image data capturing, specially designed markers were placed on the object to relate the image coordinate system to the real coordinate system. This was achieved through using a reflectorless total station.
3. Two scenarios were investigated to generate surface models or texture recreation.
  - a. Using stereo vision (AGISoft StereoScan) to reconstruct 3D surfaces from only two images, investigations were carried out to find out the optimum camera to object distance and intersection angle.

- b. Using bundle adjustment through multiple images (AGISoft PhotoScan) to reconstruct 3D surfaces.
4. Transformation (absolute orientation) is performed to transform the 3D reconstructed surface from the image coordinate system to the real world coordinate system. A minimum of 5 GCPs were utilized through the Cloud Compare (CC) free tool.
5. After transformation, a 10 cm circular sample is cropped from the 3D surface using CC. This was useful in comparison (SPT with photogrammetric) to be compatible and dependable.
6. The circular 10 cm sample data (point cloud) is fitted with a mean plane in Matlab to measure the average depth of texture using the residuals that have been calculated between data and fitted plane.
7. Finally, different analysis approaches are carried out to assess the author's developed approach through:
  - a. Comparing the results of depth worked out from digital images using different cameras and materials (asphalt and concrete) with depth measured from SPT to report the differences, calculate RMSE from differences, and find out correlation degree using Matlab.
  - b. Comparing the generated 3D surfaces from different cameras with the 3D reference data (generated from Canon 5D-50mm) to measure the reliability of the 3D automated tool (PhotoScan) as well as to measure the precision of each used camera using CC.

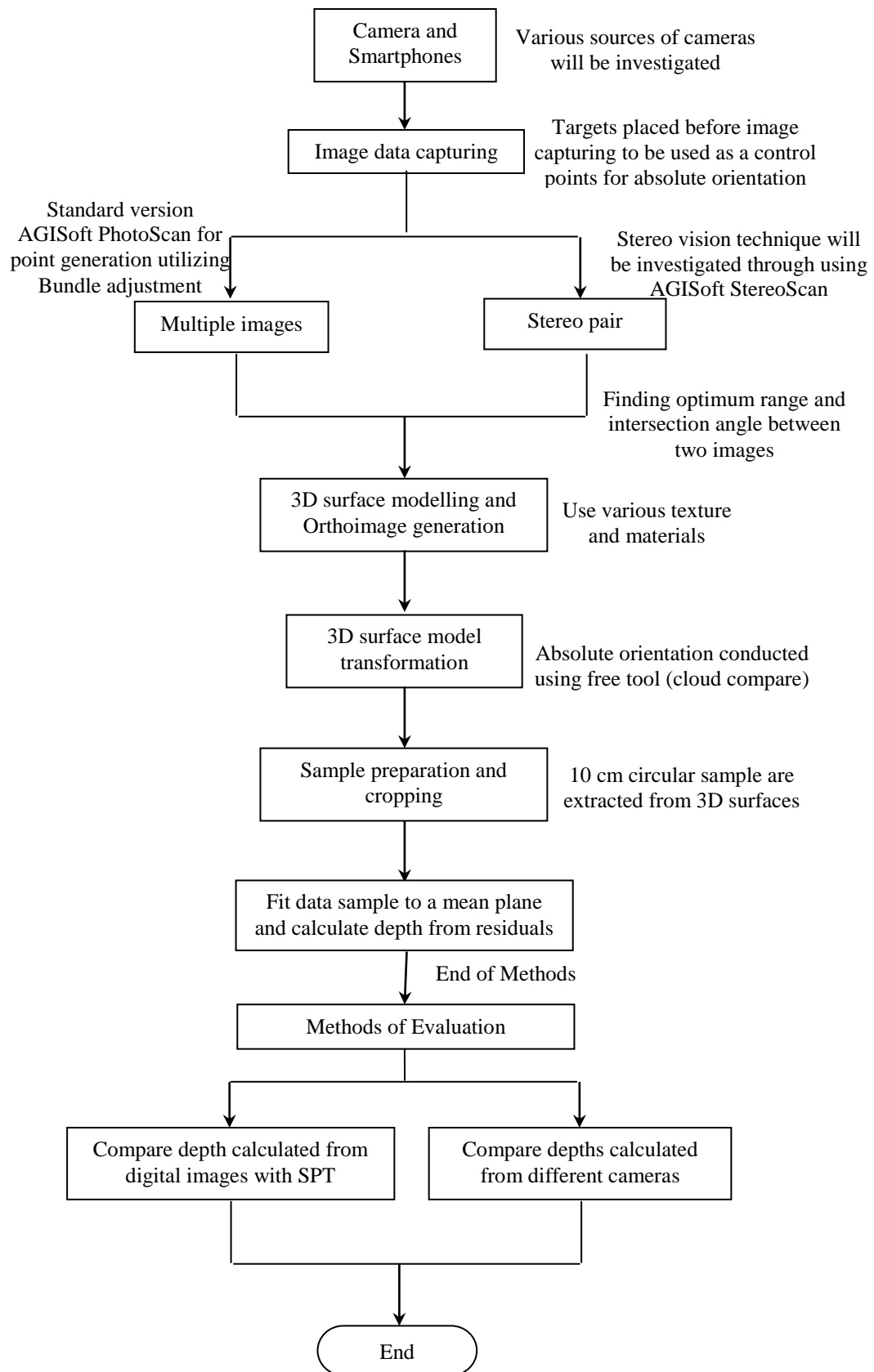


Figure 3.1: The workflow diagram for depth calculation from digital images

### 3.3 Aims and objectives of the trials

The aim of the trials was to assess the potential of the new developed approach, achieved through the following objectives:

- 1- To investigate image quality or image sharpness through testing different lenses and cameras (Trial 1).
- 2- To investigate the accuracy of both stereo images and multiple images for the production of 3D models (Trial 2).
- 3- To carry out an investigation of texture depth calculation from digital images and compare their results with SPT using two tools AGISoft StereoScan and PhotoScan (Trial 3).
- 4- To examine different texture or material types (gravel rough, gravel smooth, soil, concrete tiles, grass and asphalt pavement both ,smooth and rough) for feature extraction and matching points production (Trial 4).
- 5- To investigate the new approach of depth calculation from digital images on materials different to asphalt, such as concrete samples (Trial 5).
- 6- To test the author's developed method as a dependable measurement technique for use by other PhD studies in the Civil Engineering department (Trial 6).

### 3.4 Methodology of the trials

To fulfil the above objectives, the following methodology was adopted:

- i. **Objective 1** : A lab image was prepared using the standard Koren 2003 lens test chart printed sheet (Koren, 2004), see data sets on figure 3.2.
- ii. **Objective 2**: Two control samples, namely test 09 and test 10 rough and the smooth texture were used for evaluation, figure 3.4.
- iii. **Objective 3**: Ten samples of various texture depths were utilized, see data sets on figure 3.4.
- iv. **Objective 4**: Different materials were utilized through the use of AGISoft StereoScan on a pair of images, see figure 3.5.
- v. **Objective 5**: Two samples of concrete material were used in order to validate the proposed methodology of depth calculation on this important material, figure 3.6.
- vi. **Objective 6**: Nine samples of fractured concrete structure were used, see figure 3.7.



### 3.5 Data sets and test sites

A test location was prepared in the Visualisation and Digital Image Processing lab. An image sheet was printed and placed on the wall with the lighting condition as recommended by Imatest (2013), see figure 3.2. A variety of DSLR cameras available at the Nottingham Geospatial Institute (NGI) including Canon 5D-50mm (C50), Canon 5D-28mm (C28), Canon 5D-Sigma-28mm (C28S), Nikon-D100 (D100), Nikon-D200 (D200) and Nikon-D300 (D300) were used for this trial, utilizing different aperture modes. In order to avoid camera shaking, a tripod and remote controller or timer were employed; more details about the trial procedures are available in section 3.7.1.

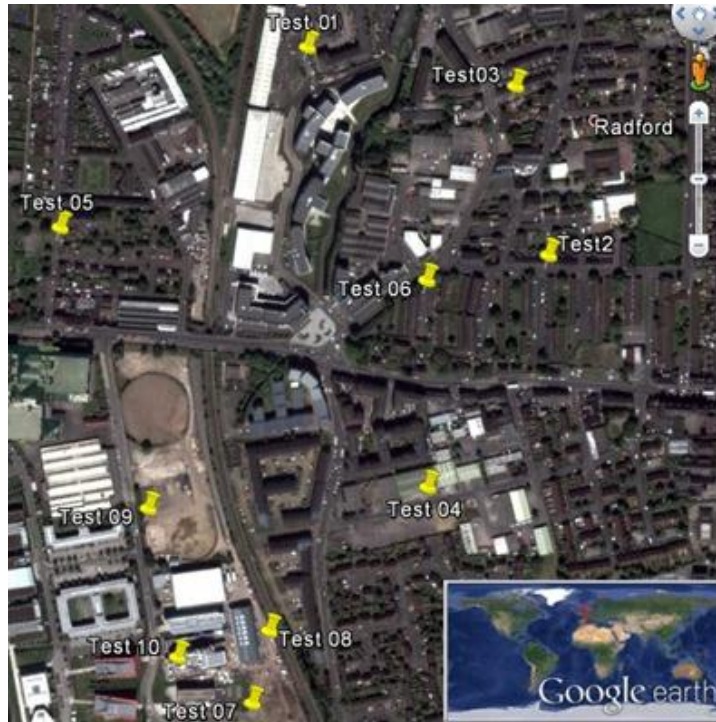


**Figure 3.2: Testing lens quality of different cameras**

In order to study the depth calculated from the digital images, the Jubilee campus road network and the area around it was utilized. The Jubilee campus road network is newly built, and the surface texture of the road was found to be less than 1mm. Therefore, the area around the Jubilee campus has been used in order to understand the variation on texture depth. The selection of the test areas and the degree of roughness were based on visual inspection by eye. Ten road test sites were chosen, see figure 3.3. In each position, a minimum of 5 markers as Ground Control Points (GCPs) were placed on the surface to act as fixed points and measured to determine their 3D coordinates using a reflectorless total station.

Having test sites with variations in texture depth is important as this relates to the real world situation. The test samples used in this examination ranged from very smooth to very rough texture, and the range was found to be between 0.5mm to 4mm. The samples used for this investigation consisted of the macro level properties of actual asphalt with different wavelength and peak to peak amplitude. Moreover, the shape and size of aggregate textures varied from sample to sample, figure 3.4. Among these ten

samples, two samples, test 09 and test 10, were used as control samples for examining the effect of intersection angle and range on some output parameters of the free tool StereoScan.



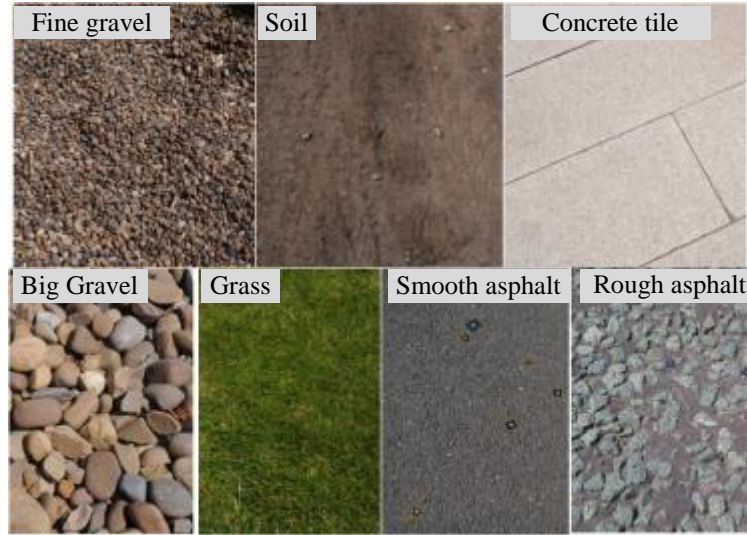
**Figure 3.3: Test study area located near Jubilee Campus**



**Figure 3.4: Used samples for texture depth evaluation**

To carry out investigation on different texture or material types and to compare their results with each other, seven texture types were used, figure 3.5. The location of these texture samples were all around the Nottingham Geospatial Building (NGB). The time of the day was an afternoon and with full sunny conditions. The data collection for

image capturing was conducted within one hour in order to minimize the variation in lighting conditions for these different textures.



**Figure 3.5: Various texture or material types**

In addition to asphalt surface texture, two concrete surface tests were also carried out in this investigation which consisted of both very rough and smooth texture, see figure 3.6. The location of these samples was the University Park Campus in Nottingham. One of these samples, the rough concrete texture is a lab sample. The other sample is a real world sample. The aim of testing the concrete surface is to examine the methodology of depth calculation on a different texture rather than asphalt. Five cameras with different resolution and focal lengths were used to examine the potential methodology of texture depth measurement from digital images.



**Figure 3.6: Concrete samples, rough texture (left) and smooth texture (right)**

Finally, data consisting of nine samples of fractured concrete structure, see figure 3.7, were measured. Again, four markers were used as GCPs for conducting absolute orientation.

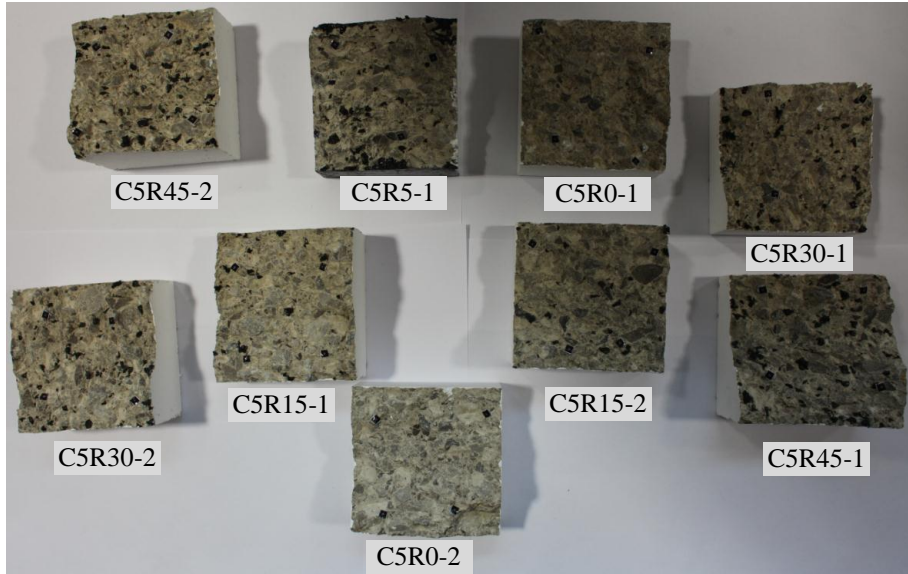


Figure 3.7: Nine concrete fractured sample from Civil Department

## 3.6 Material and modified methods

### 3.6.1 Cameras used for the trials

It has been reported that the camera in close camera to object distance measurement can provide better accuracy than a total station if it has been well calibrated (Luhmann et al., 2006; Luhmann, 2010). However, the challenge is in the automation process of calibrating the camera and generating the 3D model. Moreover, the use of smartphone digital cameras which are characterized by a very compact charge-coupled device (CCD) sensor and small focal length can add more challenges to the process of automatic 3D modelling. These challenges are related to the low image resolution and relatively inadequate lens sharpness. Both of these factors can affect the feature extraction and matching processes, as explained later in this chapter. Small focal length and CCD size can limit the image coverage, so more images need to be captured, leading to more processing time, or the camera to object distance needs to be increased, which subsequently affect the ground sample distance (GSD).

For this project, several cameras (DSLR and smartphone cameras) were employed in order to generate 3D models on an automated basis without any intervention. Cameras with various focal lengths and CCD sizes, from full frame to the extremely compact size that are used in today's smartphone market were used, in order to assess the accuracy and reliability of digital images for the purpose of texture depth measurement. Table 3.1



shows the calibration details and manufacturers specifications of each camera used. The calibration of DSLR cameras was conducted using the Australis AutoCal tool, while the smartphone camera calibrations were conducted using flat sheet in Photomodeler, as the calibration of smartphones cameras failed in Australis AutoCal. This failure in calibrating smartphone cameras using Australis AutoCAL might be attributed to the size of the calibration object (the utilized frame with coded targets in the visualization lab), which was too big in size for smartphone camera calibration. Overall Root Mean Squares errors (RMSE) were 0.129, 0.113, 0.145, 0.628, and 0.638 pixels for Canon 5D-50mm (C50), Canon 5D-28mm (C28), Nikon-D200 (D200), Samsung Galaxy Note (Note), and Samsung Wave (Wave) respectively.

Table 3.1: Cameras specifications and calibration results

Camera	Manufacturer Specifications			Australis <sup>1</sup> and Photomodeler <sup>2</sup>		
	Focal Length (mm)	Sensor W*H (mm)	Sensor Pixel Size(mm)	Calibrated Focal length (mm)	Calibrated Sensor Size W*H (mm)	Calibrated Sensor pixel Size (mm)
<b>C50<sup>1</sup></b>	50	36 *24	0.0064	53.7276	36.48*24.33	0.006497
<b>C28<sup>1</sup></b>	28	36 *24	0.0064	29.2048	36.49*24.33	0.006498
<b>D200<sup>1</sup></b>	28	23.6*15.8	0.00605	29.5330	23.99*16.06	0.006197
<b>Samsung Wave<sup>2</sup></b>	4	3.7*2.8	0.00144	3.939206	3.765*2.823	0.00147
<b>Samsung Galaxy Note<sup>2</sup></b>	4	4.4*3.3	0.00134	3.923945	4.419*3.310	0.00135

### 3.6.2 Sand patch test apparatus

The SPT for measuring texture depth is required in order to check the accuracy of 3D models through measuring the texture depth through a 3D model. The sand patch test apparatus is a tool, figure 3.8, used for measuring pavement surface texture to determine the average macrotexture depth using a volumetric patch technique. It comprises the following tools and materials in addition to the sand that has specific characteristics, refer to (Matest, 2010) for more details:

- Spreader disc with handle and rubber coated surface.
- Wind shield.
- Soft and wire brushes.
- Screw-adjusted dividers.
- 300 mm graduated rule.
- Metallic cylinder for sphere's volume measurement.
- Two glass pycnometers with metallic screw top and pouring hole.
- Three graduated cylinders 10, 25 and 50 ml cap.
- Knee-guard.
- Carrying case.



Figure 3.8: Sand Patch Test Apparatus

### 3.6.3 Total station for absolute orientation

A total station is an effective and reliable tool of measuring angles and distances and for Ground Control Points (GCPs) collection. In order to transform the camera coordinate system (3D models) into the real world, the Leica smart station (TCR1201) total station was employed. Red colour targets with a black background that will be explained later were placed on the road surface texture, and the reflectorless mode was used for measuring the centre of the target. A local coordinate system was aimed for each particular test that wants to relate between the 3D model and SPT. Therefore, no attempt was made to reference the targets into the national grid system. The reflectorless mode was utilized in order to gain as accurate a result as possible within a small area of the road texture surface.

Many parameters affect the accuracy of coordinating GCPs in the reflectorless mode, such as inclined angle, type of texture and colour of the reflecting surface. The mean square errors in measuring the slope distance, based on the distance and the colour of the used surface, was carried out by Beshr and Abo Elnaga (2011). They concluded that increasing the distance between the total station and the target leads to increase in the errors in measuring slope distance for all colours. The distance range was up to 60 m, and the errors were in the range of 0.05-1.2 mm for different colours. Accuracy of the red marker which has been used in this project, is followed by both white and yellow markers. Based on the work done by Beshr and Abo Elnaga (2011), mean square errors of less than 0.2mm for the red marker based on a small range of less than 2m, and an inclination of less than  $45^\circ$  is expected for the tests carried out in this project.

### 3.6.4 Sand patch test modification

According to the ASTM E965 Standard, a volume of known sand (typically 50 cm<sup>3</sup>) for texture depth ranged between 0.3-2 mm has to be poured into the road surface. The sand is spread out in order to make a circular shape, and a minimum diameter of 170 mm is preferred. The quantity of the sand used for the texture depth measurement should cover the patch to keep the mentioned diameter, table 3.2 shows the recommended volume of sand for the range of texture depths (ASTM, 2004) . According to British Standard, BS598, Part 3, 1985, a known volume recommended to be used is 25 cm<sup>3</sup>(HIGHWAYS DEPARTMENT, 1989). However, according to the modified British Standard BS598-105:1990, the recommended volume of sand is consistent with the American standard and it is 50 cm<sup>3</sup>.

Table 3.2: Texture depth vs volume of sand (ASTM, 2004)

Surface Texture Depth (mm)	Volume of Sand (cm <sup>3</sup> )
0.3 to~ 2.0	50
~2.0 to~4.0	100
~4.0 to~6.0	150
~6.0 to~8.0	200
~8.0 to~10.0	250

If the known sample of 50 cm<sup>3</sup> is used for the depth ranging from 0.3 to2.0 mm, the diameter of the patch will range from 46 cm to 17.8 cm for 0.3 and 2.0 mm depth respectively. To keep the comparison between 3D models and the SPT as compatible as possible, the volume of sand was modified in this project depending on the roughness of the texture in order to keep the diameter as close as possible to a 10 cm circular sample. The relationship of modified volume of sand and the texture depth are listed in table 3.3. Using the modified volume of sand showed that the sample diameter for most tests was in the range of 10-13cm, considered to be very close to the design sample of 10 cm and to have a little effect on the accuracy of the comparison.

Table 3.3: Texture depth vs volume of sand (modified)

Surface Texture Depth (mm)	Volume of Sand (cm <sup>3</sup> )
0.3 to~ 0.5	5
~0.5 to~1.0	7
~1.0 to~1.5	10
~1.5 to~2.0	15
~2.0 to~2.5	18
~2.5 to~3.0	22
~3.0 to~3.5	26
~3.5 to~4.0	30
~4.0 to~4.5	35
~4.5 to~5.0	40

### 3.6.5 Texture depth calculation criteria

This section aims to provide the theory behind depth calculation of the texture sample and the available methods for depth calculation. In the UK, laser systems provide texture depth measurement using Sensor Measured Texture Depth (SMTD), which are measured every 10m in the nearside wheel-track, the centre of the lane, or offside wheel-track, depending on the highway authorities' requirements. In Europe, however, the Mean Profile Depth (MPD) is an alternative method for texture depth measurement (Viner and Britain, 2006).

The laser sensors are fixed on the mobile platform. As the platform moves along the road, changes in the distance are measured between the laser sensors and the road surface. In order to record the small wavelength and amplitude in the texture, the sampling interval is typically 1mm. The main differences between both SMTD and MPD are in the way that the depth or height of the texture is calculated, figure 3.9. In the SMTD method, the root mean square (RMS) is measured both above and below the mean level of the texture, see equation (3-1). Whereas in MPD process, the highest peaks above the mean level are essentially estimated as the height of the texture. Mean level (ML) being the level positioned in such a way to minimize the residuals (ibid).

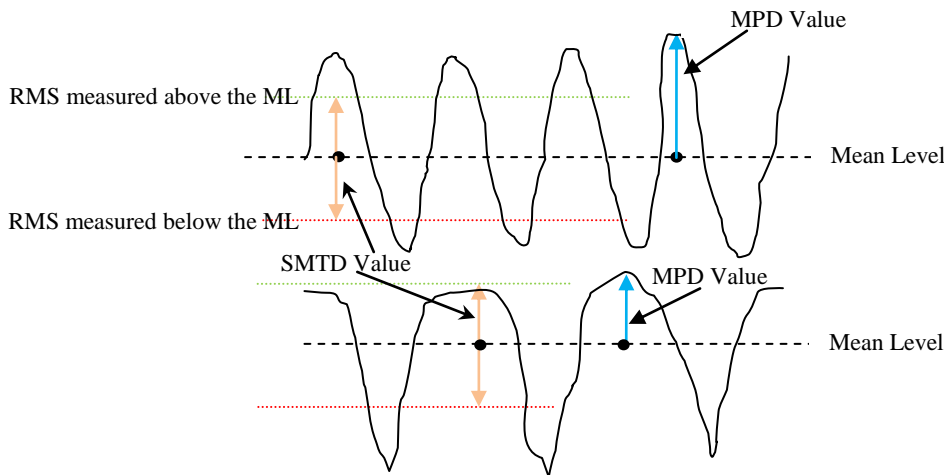


Figure 3.9: Comparison between SMTD and MPD in depth calculation (ibid)

$$\text{RMS} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_4^2 + x_n^2}{N}} \quad 3-1$$

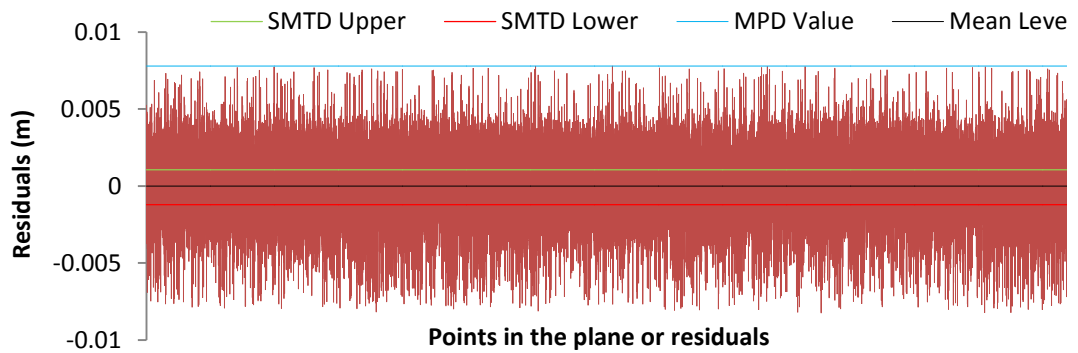
Where

$$x_1^2 + x_2^2 + x_3^2 + \dots + x_4^2 + x_n^2 = \text{observations or residuals from ML}$$



As can be noticed from figure 3.9, the calculated value of texture depth is different between both methods. The relationship between both methods calculated by Viner and Britain, (2006) for a sample of 42 km road was found to be approximately linear, see equation (3-2). The value of depth calculated from volumetric sand patch is the mean depth which is closely related to SMTD value, but it differs depending on the texture shape and size. In this project, the aim is to find out the most accurate depth that is closely related to the depth calculated from the sand patch test. For achieving this aim, the average texture depth was found to be more closely related to the sand patch depth and was used in this project, see section 3.7.2.2.4.

$$MPD = 1.42 \times SMTD^{0.840} \quad 3-2$$



**Figure 3.10: Sample shows the texture depth measurement using different approaches**

In figure 3.10, a sample of rough concrete texture, discussed further in section 3.8.5, is tested for the texture depth calculation using SMTD, MPD and the author's proposed approach. The whole procedure of depth calculation using the author approach includes the 3D model generation, best fitted plane and residual calculation, and can be found on section 3.7.2.2. The mean depth using sand is measured for this sample and found to be 5221 $\mu$ m using the average diameter of patch measured by ruler, see table 3.15. The black line in figure 3.10 is the mean level or the best fitted plane, the green and red line are the root mean square from the mean level or fitted plane considered to be the SMTD value. The SMTD value was found to be 2657 $\mu$ m comparing with 5221 $\mu$ m in the case of sand patch test depth; the difference is -2564 $\mu$ m. On the other hand, the value of MPS was found to be 7781  $\mu$ m; a difference of +2560  $\mu$ m if compared with the sand patch test depth. For depth measured using the author's proposed approach, please refer to section 3.8.5. The errors from both methods are clearly not acceptable considering that SPT can provide texture depth in the range of  $\pm 230$   $\mu$ m for a confidence interval of 95% (Rasmussen et al., 2011).

### 3.6.6 Sand patch test area calculation

Three different methods were used to calculate the area of the sand patch, the first method is recommended by ASTM E965 which is based on taking the average measurement of 4-5 rulers. The ruler is used to measure the diameter of the patch on the road surface from 4-5 different directions in order to calculate the area of the sand coverage, figure 3.11. The second and third methods have been adopted in order to get more accurate results. These methods are utilizing digital images in order to calculate the area of the patch on the road surface. The second method is through utilizing the CAD tool to calculate the accurate area through taking a perpendicular image of the patch. This is followed by exporting the image into the AUTOCAD program in order to calculate the area of the patch. In the third method, the Photomodeler programme is utilized in order to measure the area through recreating the model using 5-6 images, figure 3.11. For results and differences among three methods, refer to section 3.8.3.1. The depth of the texture is calculated by dividing the known volume of the sand by the area, see section 2.3.5.2.3.

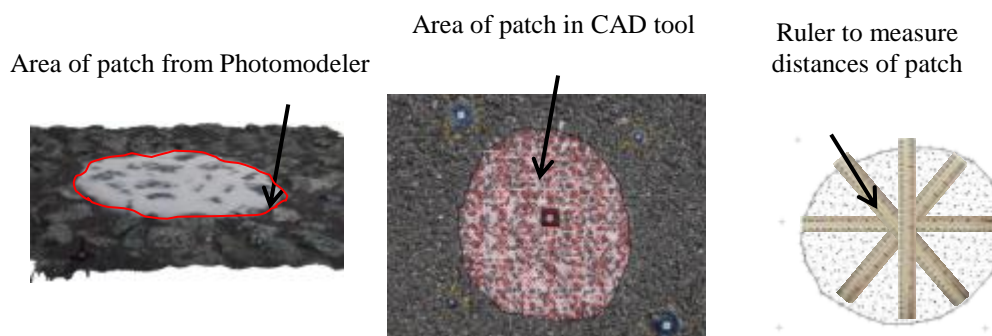


Figure 3.11: Three different methods of sand patch area calculation

## 3.7 Procedures

### 3.7.1 For testing image quality or image sharpness (objective one)

This section provides details about the methodology used in order to perform the image quality test. In order to select the camera that produces the sharpest images, a range of available DSLR cameras Canon 5D-50mm (C50), Canon 5D-28mm (C28), Nikon-D100 (D100), Nikon-D200 (D200) and Nikon-D300 (D300)) were considered. Although many internet sources can provide some excellent lens test results (Photozone, 2013), carrying out these tests in the NGB gave direct control of the trials. The test followed the steps developed by Koren (2004) which describe the procedure in details.

### 3.7.1.1 Preparation

A number of sine and bar pattern sheet were printed on a paper using the correct chart as recommended for laser jet printers. In addition, the ISO 12233 chart was added in case it is required, figure 3.12. The sine and bar pattern used for this procedure was a 5mm chart magnified 50 times as recommended by the developer. The sheet was firmly fixed on the wall and placed horizontally using a spirit level. Modulation Transfer Function (MTF) or Spatial Frequency Response (SFR) of 50% contrast was aimed for comparing the sharpness of different lenses which is located in the middle of the sine or bar chart, see figure 3.12.

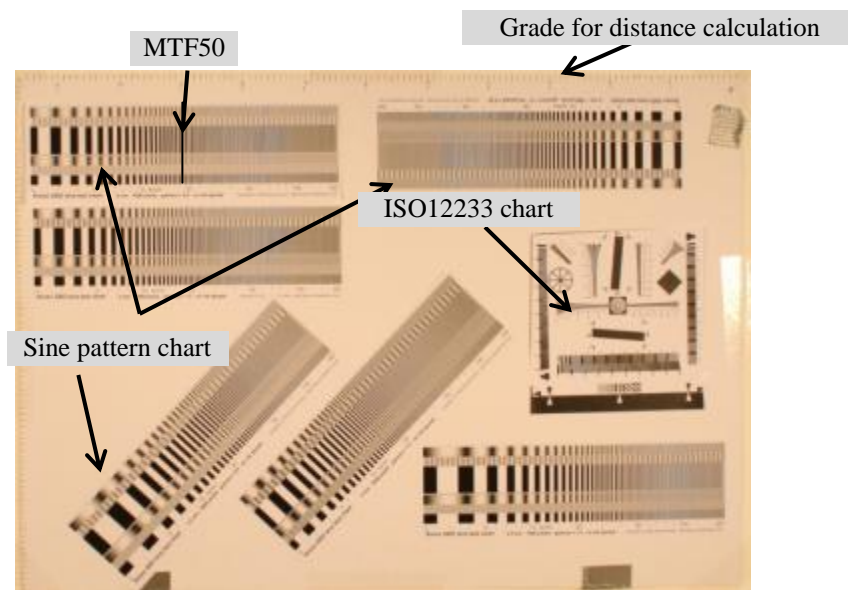


Figure 3.12: Printed Sine pattern charts for testing Image quality

The lighting condition and environment were used as recommended by Imatest (2013) in order for the light to be evenly distributed and free of glare illumination. Two sources of flood lights with 500 watt was used and placed with the degree of intersection as shown in figure 3.2 in section 3.5.

### 3.7.1.2 Image data capturing

Five DSLR cameras Canon 5D-50mm (C50), Canon 5D-28mm (C28), Nikon-D100 (D100), Nikon-D200 (D200) and Nikon-D300 (D300) were used for the image sharpness examination. In order to avoid any vibration in image capture, a tripod and timer or remote controller were employed. A spirit level was also utilized to place the camera as horizontal as possible. All camera settings were set to ISO 500 option and the aperture

was set to priority in order to test the lens sharpness at different apertures (2.8, 4.0, 7.1, 10 and 13). The RAW camera formats (NEF and CR2) of an image were recorded in order to avoid loss in compression and to preserve the full size of images.

### 3.7.1.3 Image data analysis

There are many different commercial tools that can be used for testing a lens quality with different capabilities to carry out all or some of the tests (sharpness, distortion, contrast, flare, uniformity of illumination, colour rendition and bokeh (out of focus behaviour) and many others). Depending on the software capability, the prices start from \$100 for conducting only the sharpness test (Avangate, 2013) to \$2,200 which carry out all the tests required to measure the quality of lenses (Imatest, 2013). SFRCal is a free-open-source Matlab code which can be used for testing the sharpness of the image (Koren, 2004). It can be used in accordance with ImageJ (2013) the free-open-source tool that can be used to calculate the profile from images. These two free tools were utilized to measure the quality of the lenses in term of sharpness only and as follows:

- 1- The ImageJ tool was used to prepare the profiles from the sine pattern by first scaling the image and then exporting the profiles, figure 3.13.

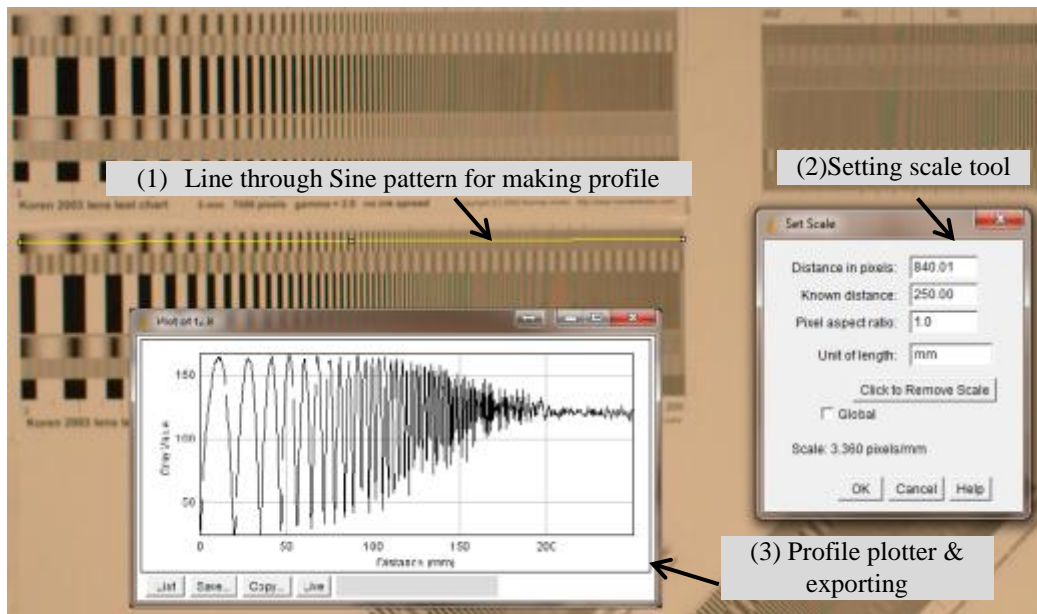


Figure 3.13: ImageJ tool for scaling and plotting profiles from sine pattern chart

- 2- Calculating the Spatial Frequency Response (SFR) graphs for the exported profiles by using the free Matlab tool SFRCal, figure 3.14.

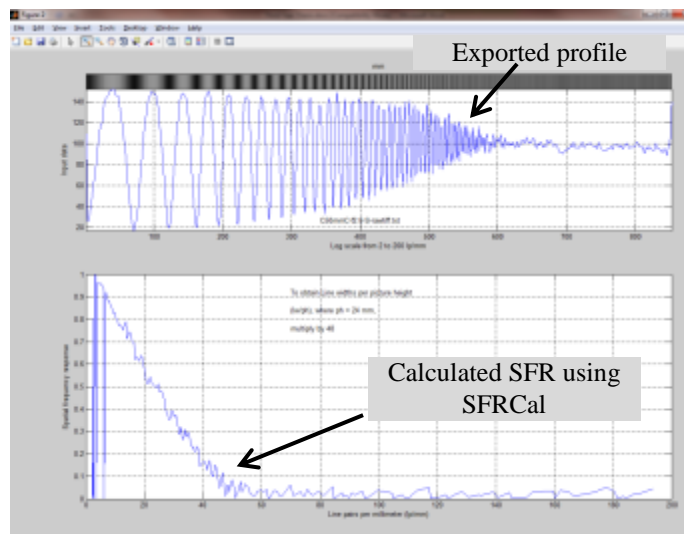


Figure 3.14: Exported profile (above) and calculated SFR vs lp/mm (below)

- 3- At this stage, the graphs of SFR are not interpretable enough due to the noise or outliers in the solution, see figure 3.15. In order to represent these graphs, the CFtool from Matlab was utilized to fit the data into the best fit curve using the 5<sup>th</sup> degree polynomial with bisquare robust option to exclude outliers as shown in figure 3.15. Only 2-100 Line Pairs per millimetre (lp/mm) were used in order to minimize the curve smoothing due to the noisy data over the 100 lp/mm.

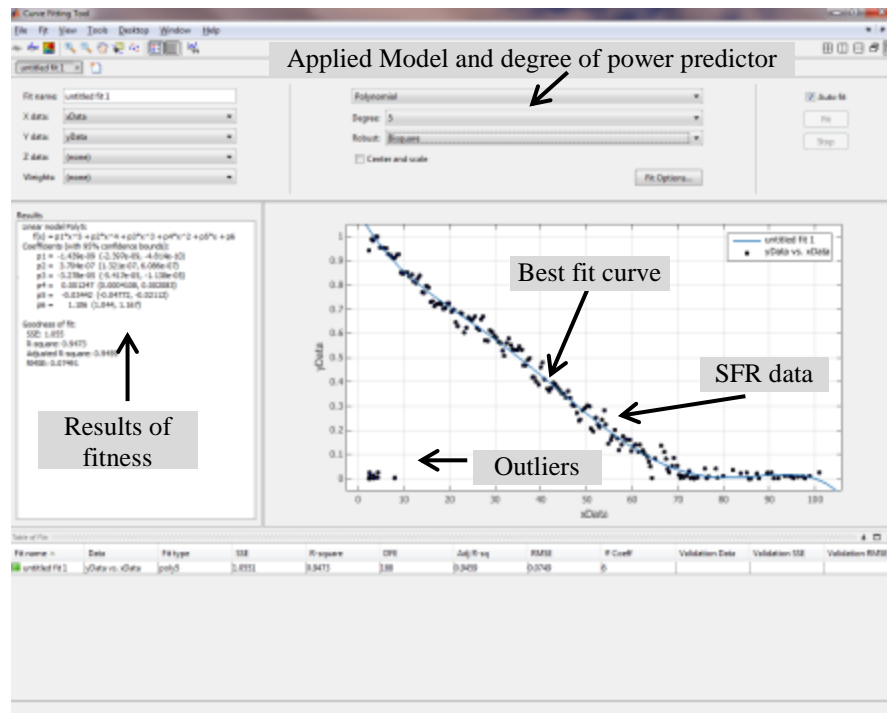


Figure 3.15: CFtool for making a best fit curve for SFR data

- 4- Finally, the output from the third process is the curve that is describing the SFR of the lens in lp/mm which can be used to test the quality of the lenses, figure 3.16, more results and details are given in section 3.8.1.

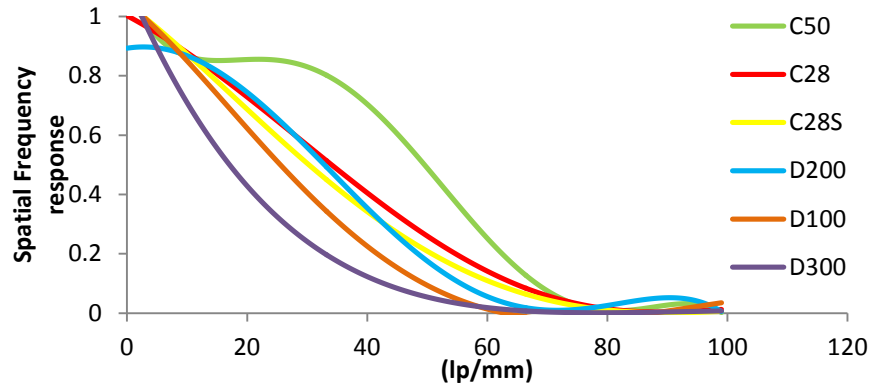


Figure 3.16: Sample of SFR response of different cameras at F stop 4.0

### 3.7.2 Preparation for all other trials (all objectives)

In this section, the overall methodology is described in detail of how the data was collected, processed, analysed and presented. The first step in the procedure is the data collection step that involves the methodology used for image capturing and SPT. For each sample test, the first step included cleaning the sample of road surface from any debris that exists on the surface that may affect the results. After that, a minimum of five specially designed markers were placed on the surface of the road. Finally, image capturing followed along with; processing and analysis, as covered in the next sections. A marker consisting of a red circle dot of (0.2cm diameter) covered in a black background of 1x1cm was used, see figure 3.17. The red dot colour was chosen for two reasons. First, it is expected that the mean square errors of less than 0.2mm will be achieved using the total station in reflectorless mode, see section 3.6.3. Second, it can be extracted and isolated from the road surface texture more easily, see section 3.7.2.2.2. The black colour background was used to separate the red dot from the road surface texture. These markers were used as Ground Control Points (GCPs) and Control Check Points (CCPs) and were measured using reflectorless total station, figure 3.17.

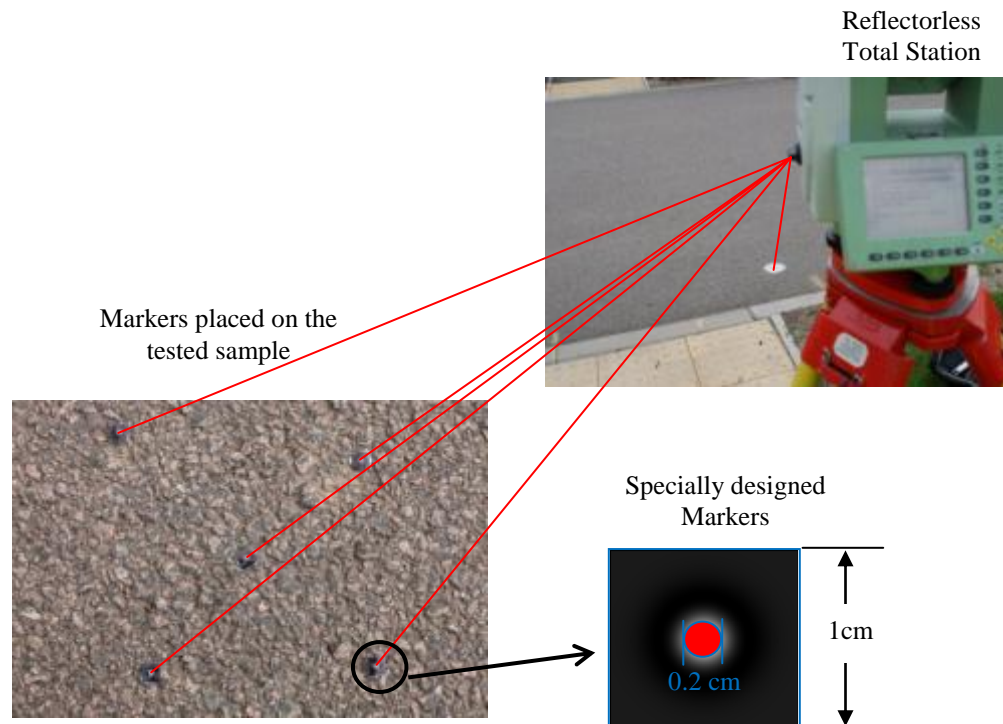


Figure 3.17: Marker details and measurement procedures

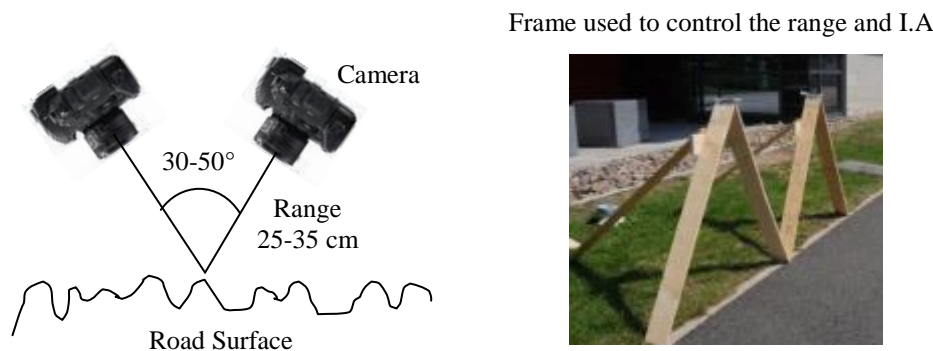


### 3.7.2.1 Image data capturing

After placing the targets onto the road surface and collecting GCPs using a reflectorless total station, the next step is image data collection for comparing 3D models with SPT; for this two methods were considered.

#### 3.7.2.1.1 Stereo pair image data collection

A stereo pair of images were captured with a convergence angle and with a suitable camera to object distance (this will be discussed further in section 3.8.2.1) for use in the free AGISoft StereoScan processing software. The results of texture depth measurement from different trials showed that the best possible intersection angle (IA) was found to be between  $30^{\circ}$  -  $50^{\circ}$  and the distance from the camera to the surface was found to be as close to the surface as the camera will allow whilst remaining in focus, see section 3.8.2.1. Depending on the focal length of the camera and the CCD size, the camera to object distance can be between 25-35 cm, figure 3.18.

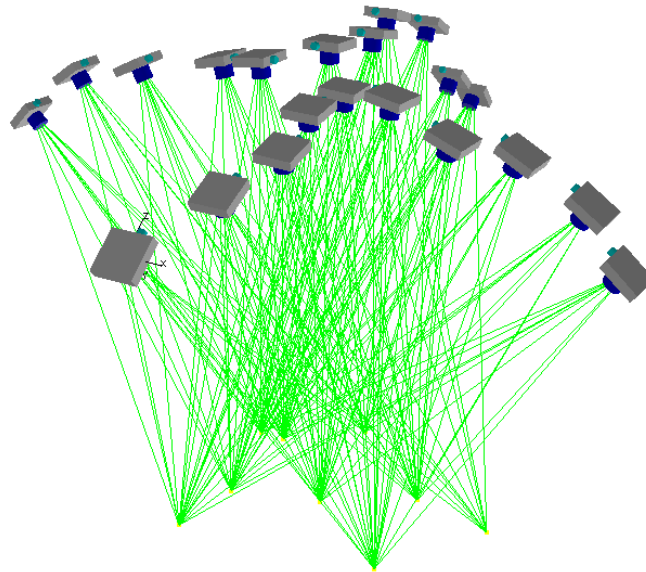


**Figure 3.18: Methodology of image collection for AGISoft StereoScan**

#### 3.7.2.1.2 Multiple image data collection

With the multi-image approach, 12 – 22 images were collected in two perpendicular rows as shown in figure 3.19. The distance from camera to the object is again as close to the surface as the camera can remain in focus, i.e. 25-35cm with different angles of convergence. AGISoft StereoScan tool can only allow for two images processing, therefore, the standard commercial version (AGISoft PhotoScan) was used for multi-image processing.





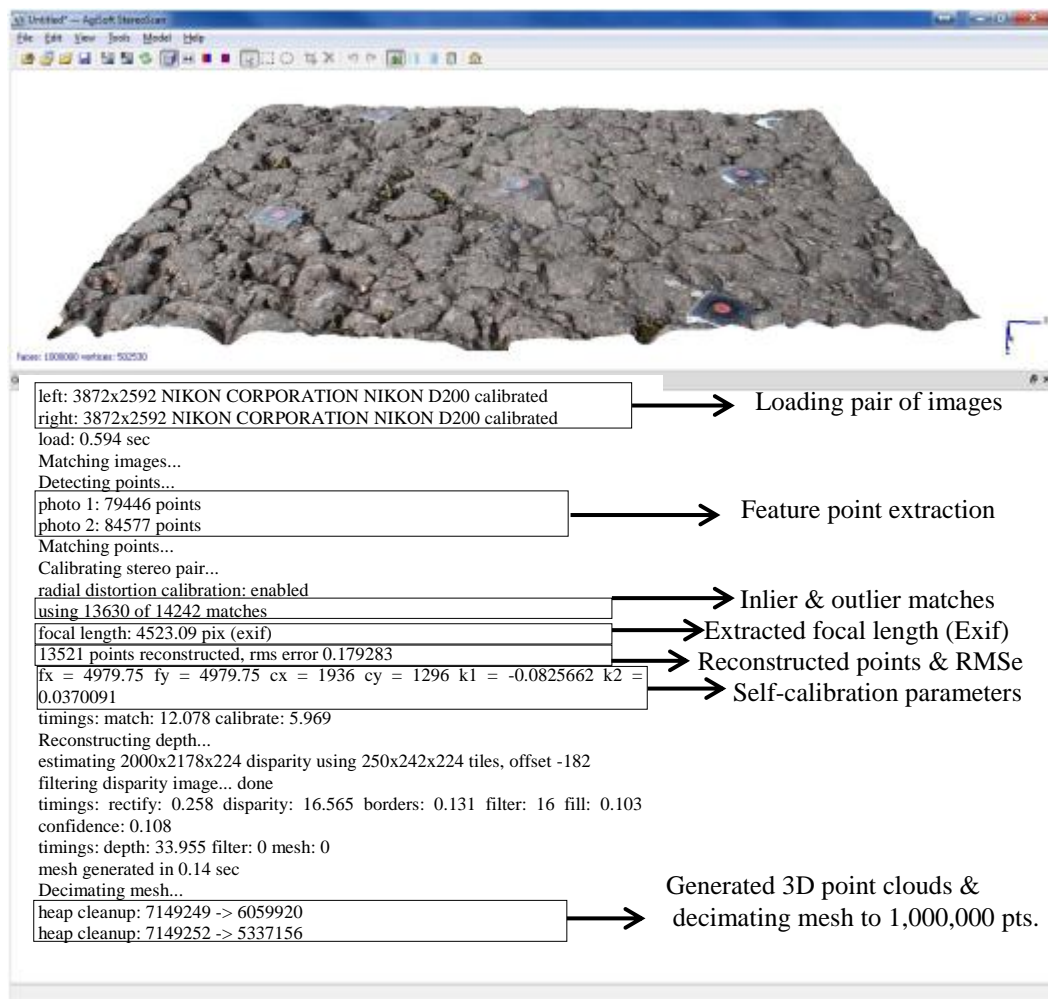
**Figure 3.19: Multiple image data collection for 3D surface modelling using AGISoft PhotoScan**

In addition to the comparison between the sand patch test and 3D modelling from digital images, other investigations were carried out in order to find out the optimum camera to object distance and intersection angle (IA) and are discussed in more details in the results section 3.8.2.1. Furthermore, the best camera settings (Aperture, ISO sensitivity and Shutter speed) are all examined in more detail in the next sections. However, the settings of the camera were found to have no real influences on the 3D geometry.

### ***3.7.2.2 Image processing and 3D model preparation***

#### **3.7.2.2.1 3D model reconstruction**

In this stage, the collected images are processed using two techniques (pair of images and multiple images). AGISoft StereoScan is a graphical interface tool used to process individual pair of images in order to produce 3D models in a very short time, with full automation without manual intervention. Depending on the resolution of the used pair of images, the full process of image feature extracting, matching, calibrating and 3D model reconstructing takes less than a minute. A report of all procedures is listed on the output panel as part of the software, figure 3.20. For more information about this software and its capabilities refer to Agisoft (2015). Due to the commercial nature of the software, little information is available about utilized algorithms and only the basics of photogrammetry are provided with the software documentation.



**Figure 3.20: 3D reconstructed sample & output report (AGISoft StereoScan)**

On the other hand, AGISoft PhotoScan is a commercial version of the software that has some additional capabilities for processing more than two images. In addition to multiple images processing, PhotoScan has the capability of transforming the local coordinates system of images into the real world coordinate system through the use of GCPs. Besides the 3D model automation, the AGISoft PhotoScan is capable of camera calibration and georeferenced orthophotos generation. Depending on the number and size of images and the settings used; the process of aligning, calibrating, and 3D reconstructing may take several hours. Reporting of all procedures in the output panel is another feature that this software offers for the users, figure 3.21. For more information about this software and its capabilities refer to Agisoft (2015). The next step is to transform the relative coordinate system of the 3D reconstructed model into the real world coordinate system using the provided GCPs data collected from designed target through the total station surveying.

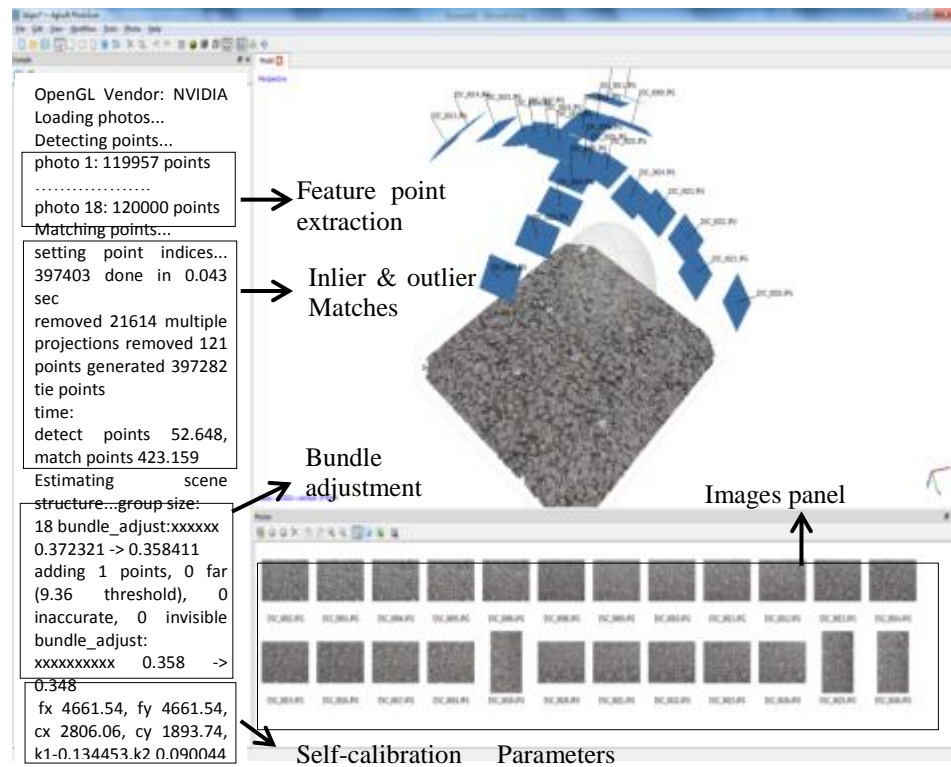


Figure 3.21: Textured 3D model & output report description (AGISoft PhotoScan)

### 3.7.2.2.2 3D model transformation

Both of the available tools (the free StereoScan and the commercial standard version of PhotoScan) do not support the coordinate transformation or interacting with GCPs. Therefore, the need for alternative 3D transformation software was necessary. Cloud Compare (CC) is an open source project that allows users to edit and process point clouds. It has many capabilities of editing, viewing and analysing point clouds. It has many other capabilities of visualizing the 3D model based on colorization (the user chooses the display colour of the 3D points), high ramp (colour visualization based on Z value) and the scalar field (gradient, Gaussian filter, and many other useful process), as well as normal to plane calculation. It also has abilities to calculate roughness, curvature and density of the point clouds in addition to the cloud-to-cloud comparisons or cloud to mesh or TIN (Triangulated Irregular Network) comparison. For more information on this open source project refer to CloudCompare (2013).

After point cloud generation from StereoScan and PhotoScan, the points are exported to the standard OBJ format. StereoScan and PhotoScan are both capable of exporting the 3D point clouds into different formats, such as OBJ (Wavefront) OBJ

models, Virtual Reality Modelling Language (VRML) models, and PoLYgon file format (PLY) models. However, StereoScan does not support the texture (coloured vertex or point with RGB) to be exported with 3D surface model on all the mentioned formats, figure 3.22. Therefore additional header information (mtllib 3d.mtl ) should be added to the OBJ file to inform the file about the material library file (3d.mtl) which contains the following information:

Characters added to header	Meaning/explanation
<b>newmtl Textured</b>	Start a definition of a new material
<b>Ka 0.6 0.6 0.6</b>	ambient color (r,g,b)
<b>Kd 0.6 0.6 0.6</b>	diffuse color (r,g,b)
<b>Ks 0.9 0.9 0.9</b>	specular color (r,g,b)
<b>d 1.0</b>	the transparency of the material 1.0 (not transparent at all)
<b>Ns 0.0</b>	shininess of the material
<b>illum 2</b>	Define the illumination model: <b>illum</b> = 1 a flat material with no specular highlights, <b>illum</b> = 2 denotes the presence of specular highlights
<b>map_Kd 3d.jpg</b>	name of file containing a texture map, which should just be an ASCII of RGB values

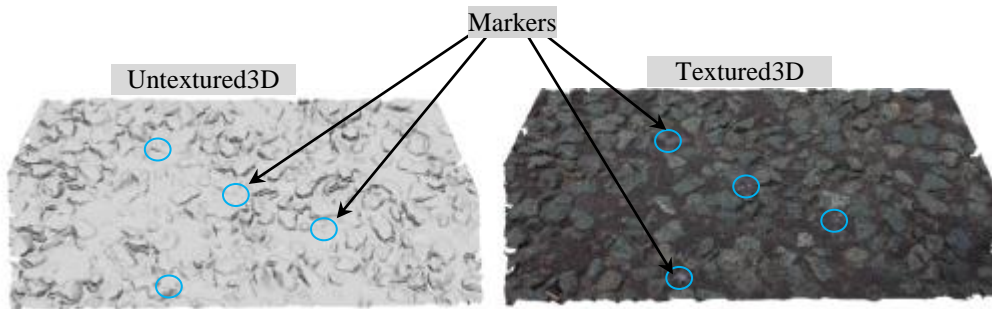


Figure 3.22: 3D model sample before and after adding material library (AGISoft StereoScan)

Standard PhotoScan version does not have this technical problem, and the exported 3D model already has the texture assigned with it. The textured 3D model plays an important role in identifying the position of the markers that will be used for 3D transformation in CC. A Matlab subroutine was designed in order to detect automatically and measure markers based on the intensities or Red, Green, and Blue (RGB) variation, figure 3.23.

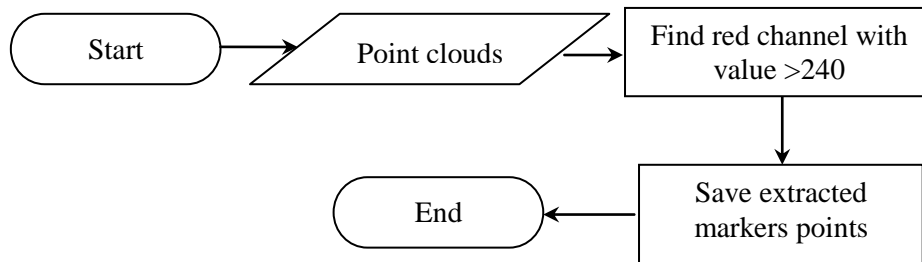


Figure 3.23: Matlab subroutine for markers extraction

The outcome of this process is not only the markers, but some identified objects (outliers) that have equal intensities to the extracted markers, figure 3.24. Therefore, the MeshLab (SourceForge, 2011) cluster extraction tool is used in order to separate the real markers from other outlier objects, see figure 3.25. The next step is to locate the centroid of these markers using Matlab, to be used as the centres of GCPs for the 3D transformation.

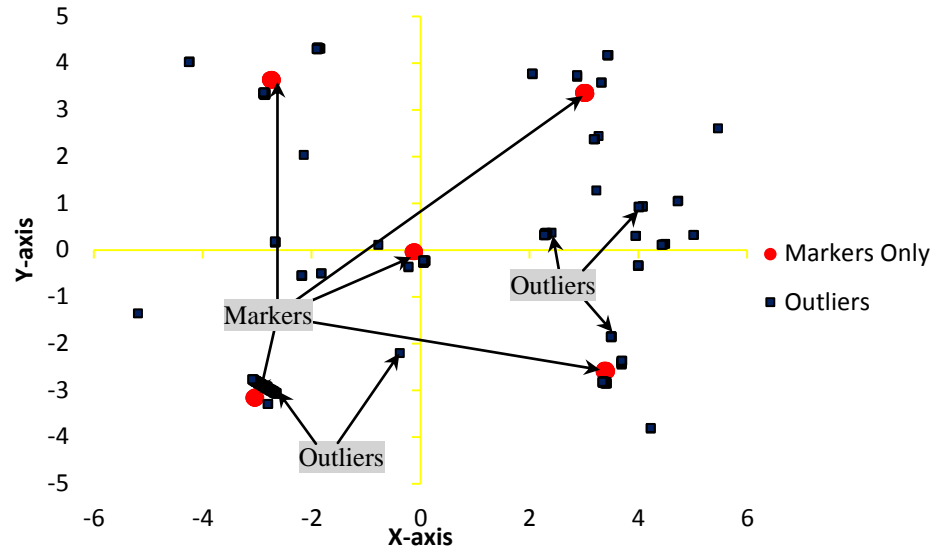


Figure 3.24: Outcome from Matlab subroutine extraction tool

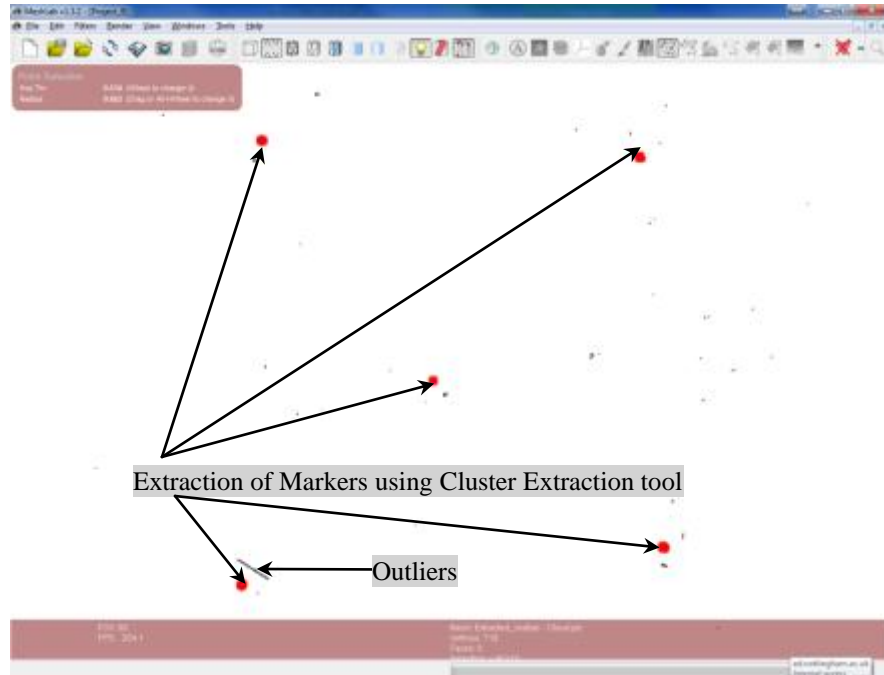


Figure 3.25: Cluster Selection tool for extracting real markers

K-mean clustering (see section 2.4.1) is used in Matlab to cluster the markers (extracted from the previous stage) into 5 groups (number of GCPs) and find out the centroid (coordinate for local image system) of each cluster as shown in figure 3.26. These centroids are considered to be the centre of the marker cluster points and will be used for transformation (absolute orientation).

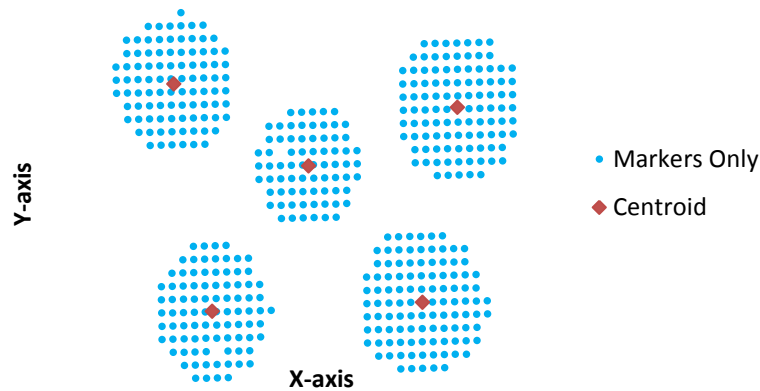


Figure 3.26: K-mean Clustering for grouping and finding out the centroid

At this stage, the centres of the markers are defined, and the real world coordinates of these markers collected via reflectorless total station instrument are also available. The next step is to use CC to convert from the 3D model imagery coordinate system into the real world coordinate system. Both the RMSE and transformation parameters of this 3D transformation process are provided by CC tool in the output report, see figure 3.27.

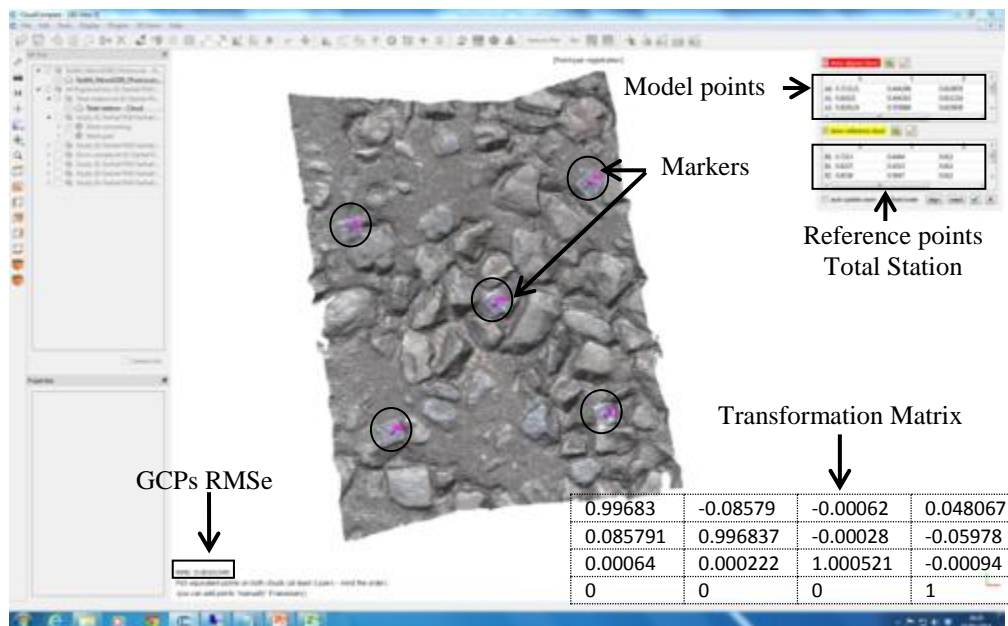


Figure 3.27: 3D transformation from relative to real world system



### 3.7.2.2.3 3D model sample preparation or cropping

The sample designed for the comparison with SPT was chose to be a circular 10 cm diameter. The reason of choosing this sample is because the distance from camera to the object was found to be as close as possible with the camera not out of focus. This range was found to be between 25-35cm depending on the focal length of the used camera, see section 3.8.2.1.2. The 3D modelling footprint of the wider lens camera used in this experiment was found to be 15 cm by 12cm, see figure 3.28. To control the samples and to make the comparison as compatible as possible, the 10 cm diameter was found to be very suitable for processing, analysis, and comparison.

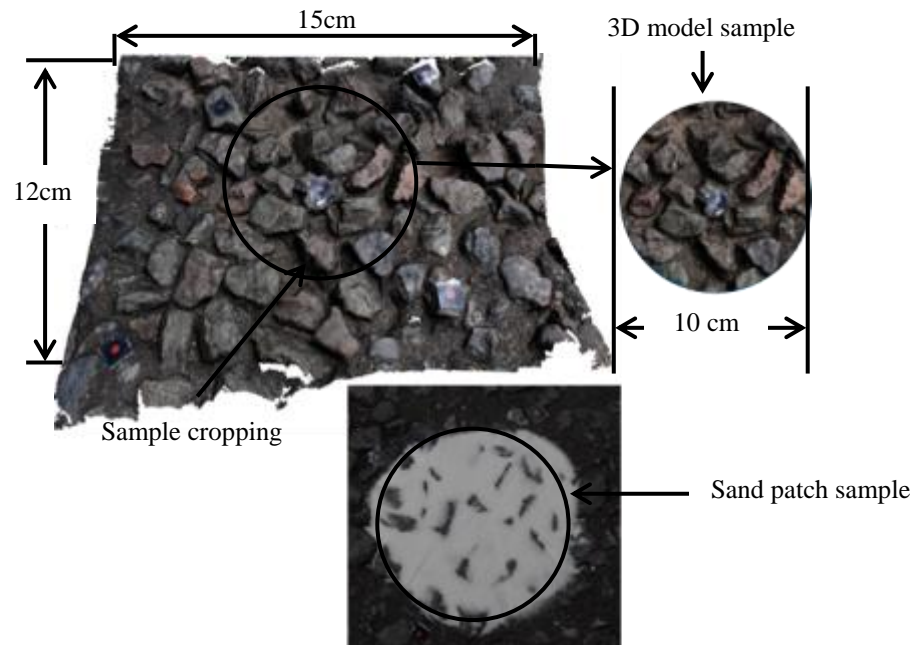


Figure 3.28: 3D model sample cropping and sand patch (Canon 5D 50mm)

The final stage in sample processing is to make a sample's horizontal or vertical axis normal toward  $Z^+$  (this being defined by the real world coordinates from the total station). It is true that the 3D transformation will make the 3D sample  $Z$  axis toward normal  $Z^+$  but due to many reasons the  $Z$  axis of the sample might not be totally aligned with  $Z^+$ ; one cause being due to the tilt of the pavement where the sample was taken. In order to make the sample truly horizontal, the CC plane orientation tool was utilized, figure 3.29.

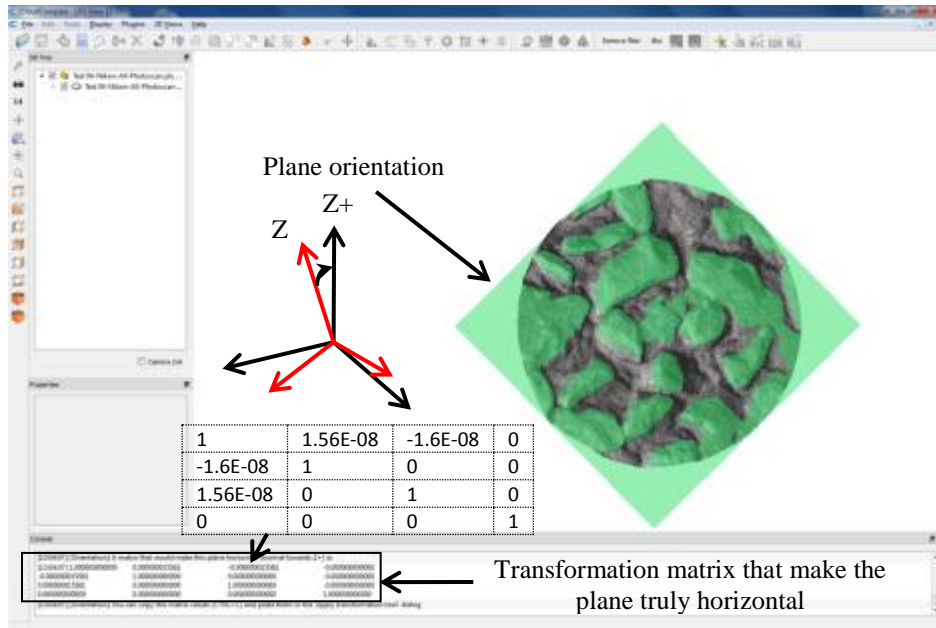


Figure 3.29: Plane orientation of 3D model sample

#### 3.7.2.2.4 3D model sample depth calculation

After the sample preparation stage, the sample is further processed in order to calculate the depth of the texture. This step involves calculating the depth of the texture based on fitting a mean plane to the point cloud. Most of the available depth calculation techniques (refer to section 3.6.5) are based on the surface profile (2D) measurement and do not include the volume of the texture, which is required if a comparison is to be made with SPT. As previously discussed, techniques based on 2D surface profile represented by SMTD and MPD are not accurate and showed lots of discrepancy when compared with the texture depth calculated by SPT. Therefore, the necessity for an algorithm development to deal with volume rather than surface profile was important.

The author's developed algorithm first fits a plane to the cloud of points defining the surface. Then the residuals from the plane to the point cloud surface are determined. The fitted plane is normal polynomial with xy axes set to first degree, positioned such that the sum of the upper and lower residuals is equal, figure 3.30. The residuals are divided into two parts; the upper residual parts, the positive residuals, that are located above the mean fitted plane and the lower residual parts, the negative residuals, which are located below the mean fitted plane. The sum of the absolute mean upper and mean lower residuals is considered to give the depth of the texture that is more closely related to the mean depth of the sand patch. Figure 3.31 show the flowchart of the overall process of the algorithm using Matlab program.



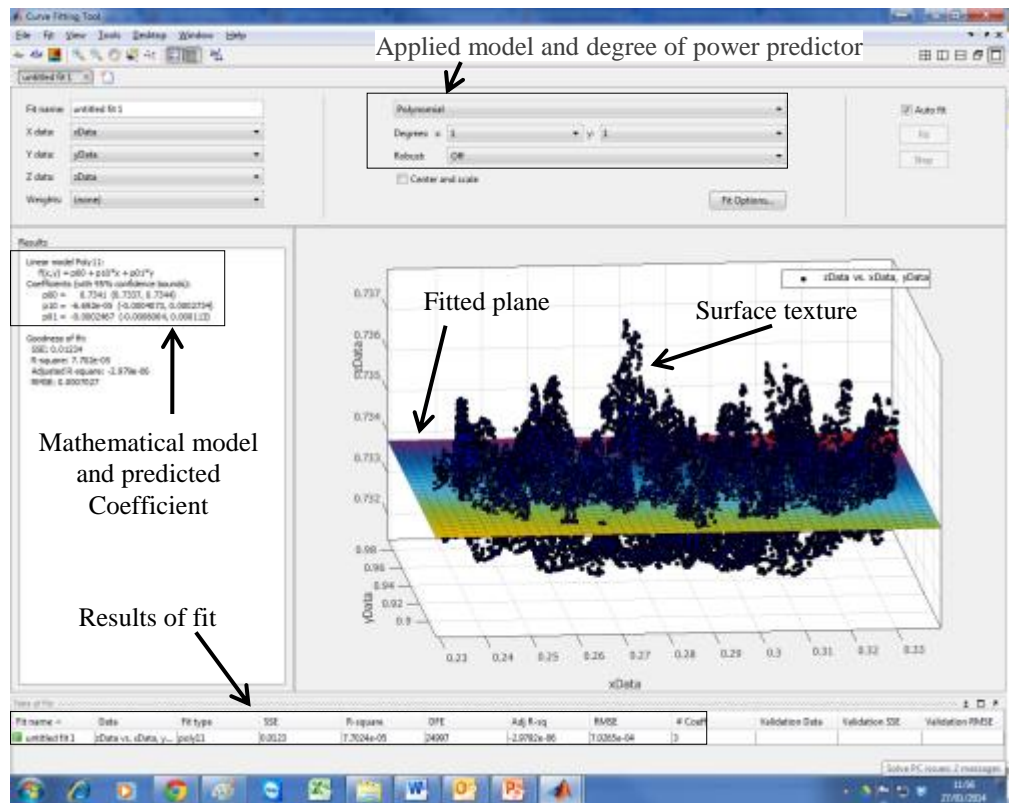


Figure 3.30: Sample of Matlab fitted plane to surface tool

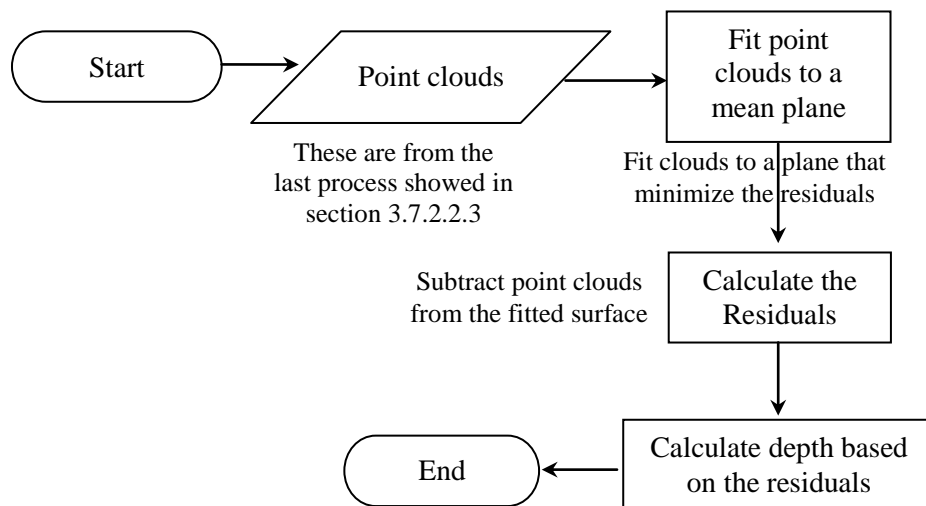
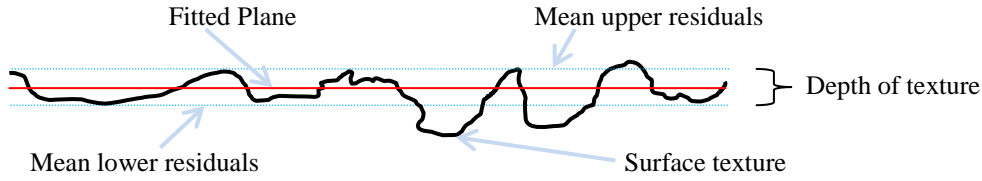


Figure 3.31: Author's Developed Algorithm for calculating depth of texture

To demonstrate an example of a 3D model sample depth calculation, test 10 asphalt rough sample (figure 3.4) was used. The cross section or profile sample was extracted from the centre of test10 asphalt circular sample, see figure 3.32. The profile cross section sample shown in this figure is only for the centre of the sample.



**Figure 3.32: Cross section on sample of rough texture illustrate the fitted plane, upper and lower residuals**

The fitted plane is shown in red colour; the texture surface is in black and the blue lines are the location of the mean upper and lower residuals. The fitted plane is positioned such that the upper and lower residuals are equal, see table 3.4. The depth of the texture is the sum of the two absolute mean residuals (upper and lower).

The algorithm calculates the upper and lower residuals and from those upper and lower residuals a mean is calculated, equation (3-3). The mean lower residual was -1.032 mm and the mean upper residual was +1.194 mm for this sample used. The depth of the texture is the sum of the absolute mean upper and lower residuals which is 2.226 mm. Results for the other two methods are also presented and as can be noticed from table 3.4, the more closely related depth to the SPT depth is the average texture depth with about 0.2 mm differences. In the next sections, more results of comparison between SPT and average depth will be discussed.

$$\text{Average depth} = \left\{ \left| \frac{\sum_1^{n_1} -r}{n_1} \right| + \left( \frac{\sum_1^{n_2} +r}{n_2} \right) \right\} \quad 3-3$$

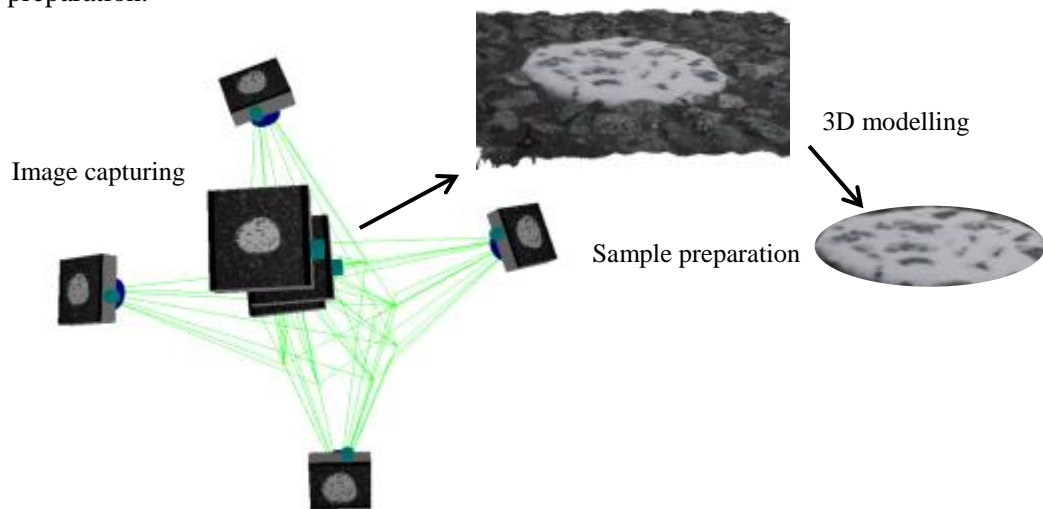
Where *average depth* : depth of the texture  
 $n_1$ : Number of points or responses of lower residuals  
 $n_2$ : Number of points or responses of upper residuals  
 $+r, -r$  : Is upper and lower residuals response

Table 3.4: Sample of results from texture depth calculation algorithm and two techniques (SMTD and MPD) (mm)

Sum(-)	Sum(+)	Mean(-)	Mean(+)	Average Depth	MPD	(SMTD)	SPT depth
- 479.671	479.672	-1.032	+1.194	2.22	3.12	1.30	1.96

### 3.7.2.2.5 Sand patch surface reconstruction

Another method to test the accuracy of 3D models from images will be investigated through using the sand surface as a reference rather than using a mean fitted plane, which has been previously discussed. In order to validate the depth calculated from the 3D models, the surface of sand was recreated as it was in the field using the PhotoScan software. Multiple images were collected for the sand surface, after spreading out sand on the surface, and processed as previously discussed. Figure 3.33 shows a sample of recreated sand surface with the process of collection the images and sample preparation.



**Figure 3.33: Multiple image data collection for reconstructing sand patch surface**

The recreated sand surface is used as a reference surface to subtract from the texture model surface. Then the absolute differences between the two surfaces (reference and model) are calculated. The absolute average distance is the average of all perpendicular distances between the two surfaces (TIN) at user defined intervals, or it might be point-to-point distance comparison, more details are in section 3.7.2.3.2. The average absolute distances should be very close to the sand patch test depth if the reconstructed sand patch and the road surface texture are accurate which will be investigated in the next sections. The SPT depth method can be also evaluated using the reconstructed sand surface which according to the specification should be a flat surface, and that has also been investigated.

### 3.7.2.3 Data analysis approaches

In order to carry out analysis, different approaches were considered; the statistical testing of the results was one option. Average texture depth is compared statistically with

SPT depth to find out the degree of correlation. Another option included was comparing the results from different cameras with each other through surface-to-surface comparison. Also, the mean and standard deviation of each camera for different texture samples were worked out. Both precision and accuracy were studied, as explained in the following sections. It is also possible to evaluate the reliability of the 3D automated software through the use of comparisons between different samples of 3D models.

#### **3.7.2.3.1 Average texture depth calculation**

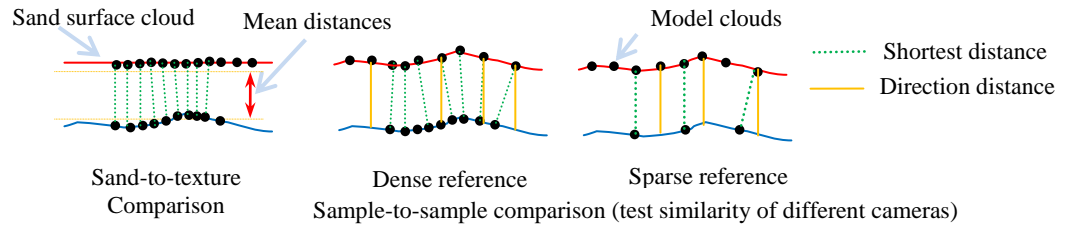
The results from the depth calculation Matlab algorithm for different cameras and samples were used to examine the accuracy of the texture depth calculated from generated point clouds, see section 3.7.2.2.4. These results were used to carry out a comparison between calculated depth from this method and SPT to find the regression or correlation between depth from digital images and SPT, see section 3.8.3.

#### **3.7.2.3.2 Comparison of 3D points or model (surface-to-surface or points-to-points)**

This comparison is usually conducted to evaluate the similarity between two sets of 3D models or point clouds. For example, to compare between 3D models generated from different cameras or between laser and image clouds. One sample is set as a reference and the other one as a model; the model is subtracted from the reference. This comparison was used in this project in two different ways. In the first approach, the degree of similarity between point clouds generated using various cameras was tested (all cameras with reference to Canon 5D-50mm (C50)). In the second approach, the depth worked out between the sand reconstructed surface and the texture model.

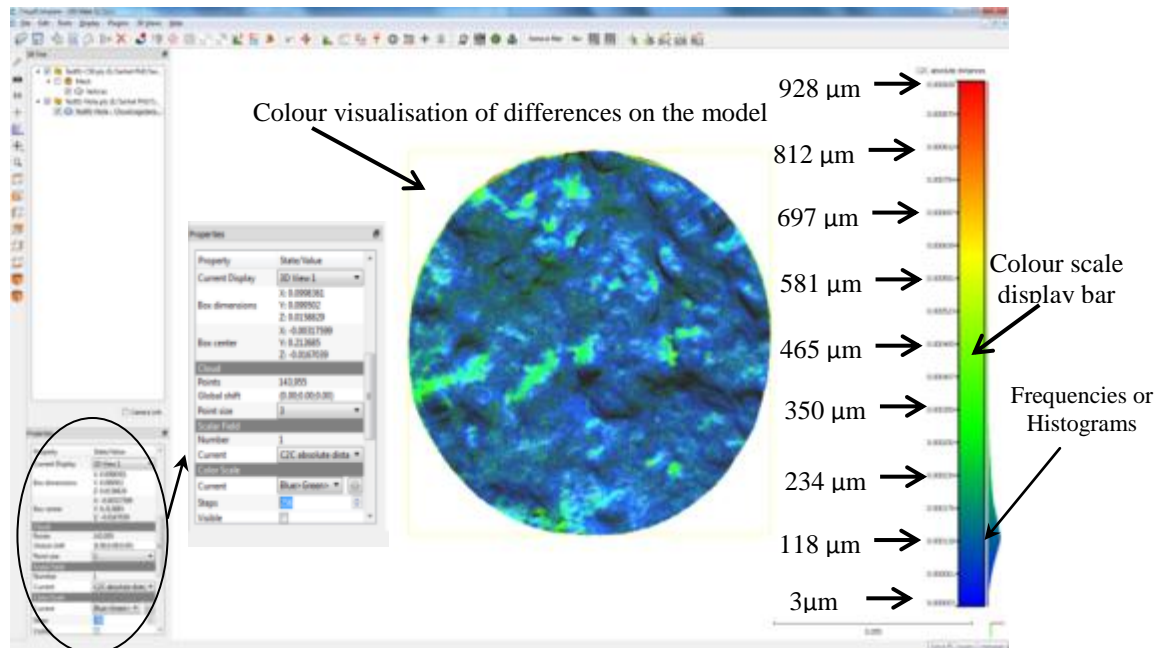
Most 3D point cloud and mesh processing commercial programs such as Geomagic Studio, Pointools, and VRMesh Studio (Geomagic, 2010; Bentley, 2014; VirtualGrid, 2014) have the ability of calculating and visualizing the differences between a set of models. They also have the capability of reporting the relative accuracy as well as the colour visualization of differences on the model. However, CC also has the same capability as the commercial tools, and it is a free to use software. Therefore, CC is used for comparing between surfaces or point clouds. Two approaches are available for comparison, the shortest distance or defined direction such as toward Z. By default, CC program utilizes the shortest distance for comparison. For each point on the model, CC searches for the closest point in the reference cloud or mesh and calculates their

(Euclidean) distance, see figure 3.34. Therefore, it is recommended for the reference cloud to be as dense as possible in order to increase the reliability of difference calculation by reducing interpolation (CloudCompare, 2013).



**Figure 3.34: Distance computations between reference and model surfaces**

All point clouds generated from AGISoft StereoScan and PhotoScan were dense enough, 1528 to 1910 points/cm<sup>2</sup> or 65 to 52  $\mu\text{m}$  resolution. Therefore, any local modelling options such as least square best fitting plane, or a 2D Delaunay triangulation, or a quadratic height function were not used for further intensifying the reference clouds. The result from comparing the surfaces is a colour based visualization (figure 3.35) ranging from blue as being minimum distances to red treated as maximum distances. Two surfaces for rough concrete texture from two different cameras, i.e. (Canon 5D-50mm (C50) and Samsung Galaxy Note (Note)), see section 3.8.5, were used as an example in figure 3.35 for explanation.



**Figure 3.35: Typical result of a distance computation process with colour scale display parameters**

The calculated distances from comparing the two surfaces (reference and model) are stored as histograms for further analysis. From these histograms, many important parameters such as mean distances, standard deviation and the percentage of cut off point's limit can be calculated, figure 3.36. Mean distances is considered to be texture depth or 3D deviation for both (sand-to-sample and sample-to-sample), respectively. Standard deviation is considered to be uncertainty in the mean distances of texture depth or 3D deviation. Cut off point at 1mm is a limit to show the degree of similarity between 3D models (sample-to-sample) from different cameras compared with the Canon 5D-50mm (C50).

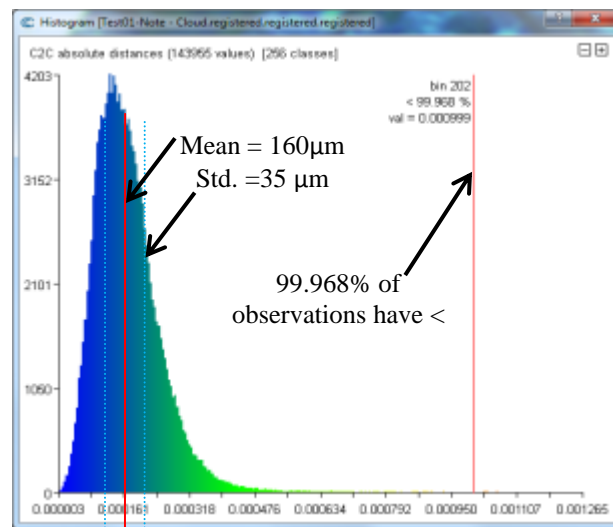


Figure 3.36: Histograms of distance differences

### 3.7.2.3.3 AGISoft summary report

AGISoft StereoScan provides a summary report of the 3D reconstruction process; this summary report includes a lot of important information on the outcomes from the analysis procedures of reconstructing the 3D models, refer to section 3.7.2.2.1 for more details. The outputs from this reports are feature points number on each image, inlier and outliers of matching points, 3D reconstructed points, RMSE of the process represented in image unit, and camera calibration results represented as a focal length and radial distortion. These output result parameters were used as a tool to measure or examine the accuracy and reliability of the 3D modelling automation solutions. These parameters were used in investigating the angle of intersection and camera to object distance of cameras on both smooth and rough texture of asphalt surface. The relationship of each of these parameters, for example, features, matches, and RMSE, were compared against the intersection angle (IA) and camera to object distance of cameras, refer to section 3.8.2 for more details.

## 3.8 Results and Analysis

### 3.8.1 Trial 1: Investigate the effect of lenses and cameras on image quality or image sharpens

In order to select the best available lenses with DSLR cameras Canon 5D-50mm (C50), Canon 5D-28mm (C28), Canon 5D-Sigma-28mm (C28S), Nikon-D100 (D100), Nikon-D200 (D200) and Nikon-D300 (D300) from the NGI; it was decided to test the available DSLR cameras for the important parameter which is the image sharpness. Smartphone cameras were not presented in the results due to a failure in processing them using the free SFRCal tool. Although, many internet sources can provide some excellent lens test results (Photozone, 2013), carrying out tests on lenses is important for many reason. For example, lenses may be different from the same product, and also to avoid using the old lenses which have been affected by shaking or wearing (loss elements) that may affect the accuracy. The tests followed the procedures developed by Koren (2004).

The procedures of conducting this test are explained in procedure and methodology, section 3.7.1. Four cameras (Canon 5D Mark II, Nikon-D300, Nikon-D200, and Nikon-D100) with a mixture of six lenses (Canon28mm, Canon-50mm, Sigma-28mm for Canon-5D, Nikon-28mm for D300, Nikon-28mm for D200 , and Nikon-28m for D100) were used in this test. The test was conducted utilizing five aperture modes (F2.8, F4.0, F7.1, F10, and F13). Only up to of F13 was used because in the field further aperture settings will result in blurring of the picture as the shutter speed will reduce. Instead, the ISO sensitivity has to be increased to substitute for low light coming to the camera that will lead to noisy image.

The results of the five aperture modes, figure 3.37, show that the spatial frequency response increases with minimising the aperture. The increase is due to the fact that the depth of field increases with decreasing the aperture. The trend for almost all cameras and in all aperture modes shows that the Canon 5D with a 50 mm lens (C50) outperforms all other lenses. This outperformance might be attributed to the fact the 50 mm lens was made with special care (more expensive) or might be due to its long focal length. However, it is worth to mention that equivalent distances with relation to the focal length were considered when the test was carried out. Using the Canon 5D with the other three lenses (C28, C28S), and the Nikon D200 recorded similar performance. Lastly, the D300 gave the worst results which might be due to the wear process or excessive use.

Therefore, only 2 DSLR cameras (Canon-5D and Nikon-D200) with 3 lenses (Canon-28mm, Canon-50mm, and Nikon-28mm for D200) were used in addition to both smartphone cameras Samsung Galaxy Note (Note) and Samsung Wave (Wave) in the remaining tests of this chapter.



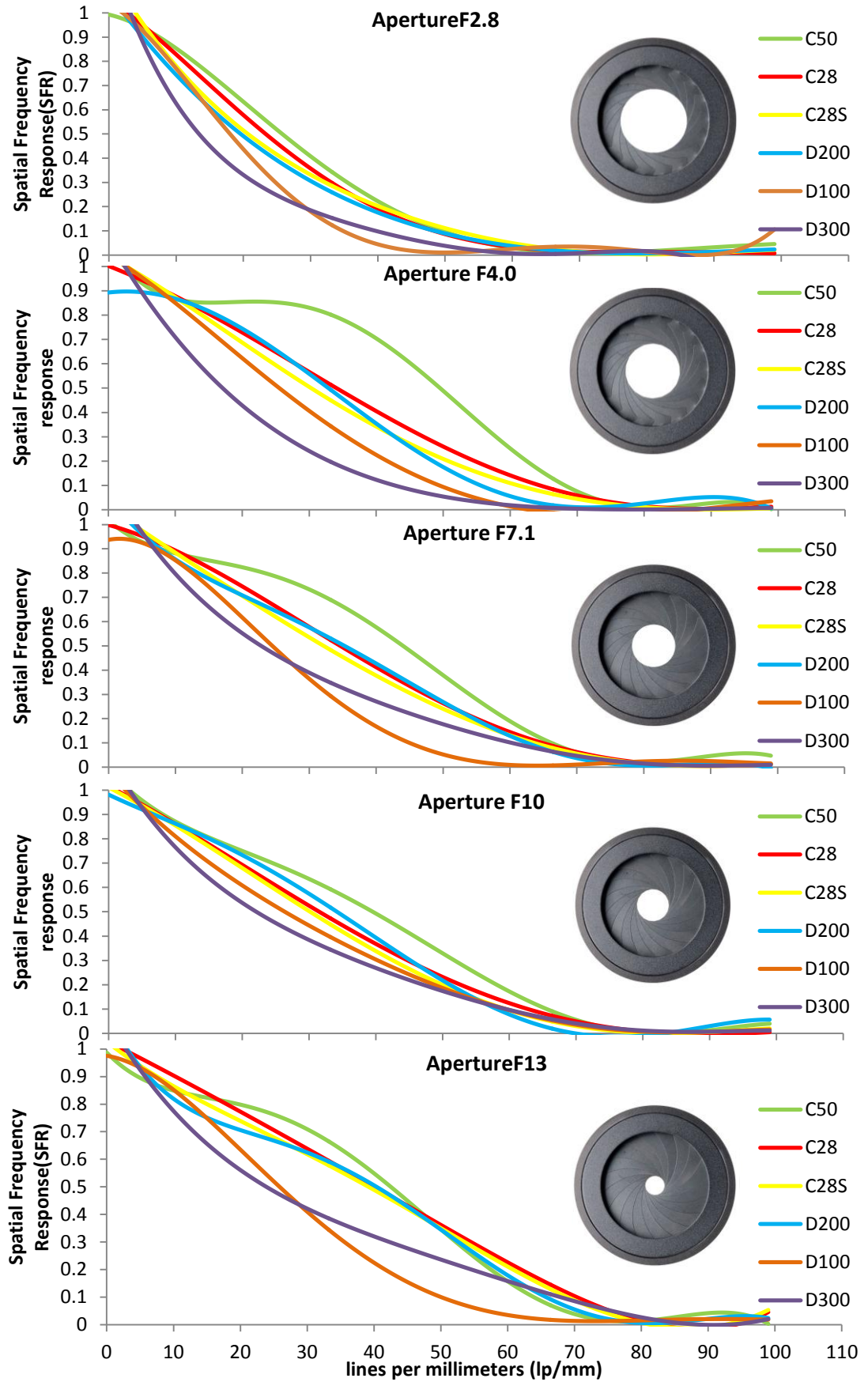


Figure 3.37: Spatial Frequency Response for various cameras

### 3.8.2 Trial 2: Investigate the accuracy of Stereo and multiple images

#### 3.8.2.1 Effect of intersection angle (IA) or base to height ratio and object to camera distance on Stereo images

The purpose of this section is to investigate the effect of the camera range, or the distance from the camera to the object, and intersection angle between the stereo pair using free software AGISoft StereoScan. To achieve this aim, a series of tests were carried out using the five cameras mentioned in section 3.6.1, i.e. 3 DSLAR cameras and 2 smartphone cameras. Two controls sample rough and smooth texture (test09 & test10) shown previously on figure 3.4 were used for this trail. The surface texture in both test 09 and test 10 is asphalt pavement. The depth of both textures was measured using the SPT and used as a reference for 3D model accuracy.

##### 3.8.2.1.1 Intersection angle (IA) effect

Intersection angle (parallactic angle) can be defined as the angle between the two lines of sight from the camera position to the object; it also has a relation with the base to the height B/H ratio, figure 3.38. By increasing the base or decreasing the height, the intersection angle increases and vice versa. Table 3.5 shows the equivalent values in degree of intersection to the B/H ratio calculated using the CAD tool.

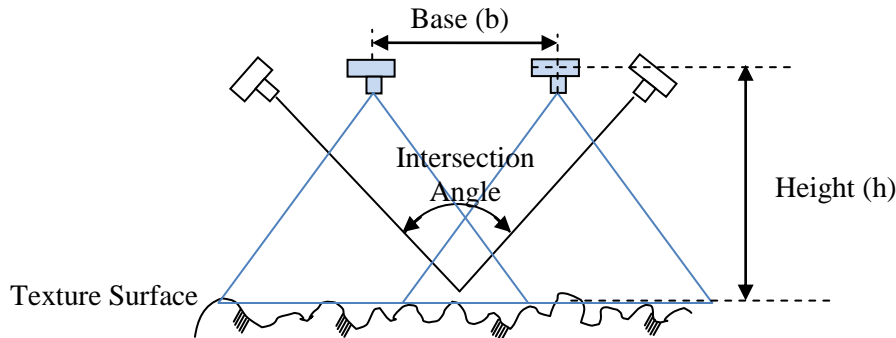


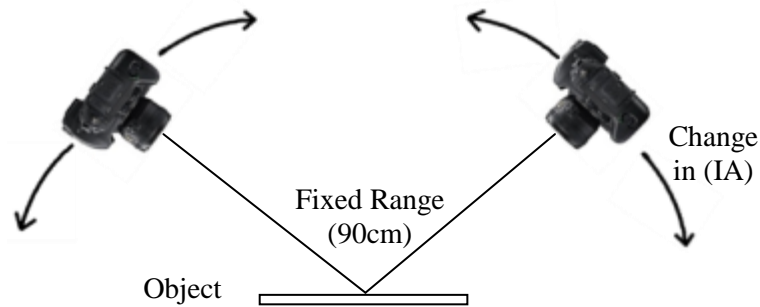
Figure 3.38: Base to height ratio and intersection angle

Table 3.5: Equivalent values between B/H ratio and intersection angle

B/H	Intersection Angle~ °	B/H	Intersection Angle~ °
1.9/1.0=1.9	87°	1.0/1.0= 1.0	53°
1.8/1.0=1.8	84°	0.9/1.0= 0.9	48°
1.7/1.0=1.7	81°	0.8/1.0=0.8	44°
1.6/1.0=1.6	77°	0.7/1.0=0.7	39°
1.5/1.0=1.5	74°	0.6/1.0=0.6	33°
1.4/1.0=1.4	70°	0.5/1.0=0.5	28°
1.3/1.0=1.3	66°	0.4/1.0=0.4	23°
1.2/1.0=1.2	62°	0.3/1.0=0.3	17°
1.1/1.0=1.1	58°	0.2/1.0=0.2	11°

It is well known in the basics of photogrammetry that the accuracy and reliability increase when the intersection angle increases. According to Hasegawa et al (2000), height accuracy decreases as the B/H ratio decreases. However, height accuracy also decreases as the B/H ratio increases if B/H ratio becomes more than 1.0. So, they concluded that a B/H ratio ranging from 0.5 to 0.9 is the best value for automatic DEM generation from a stereo pair. B/H ratio ranging from 0.5 to 0.9 is equivalent to  $\sim 48$  to  $\sim 28$  degrees, see table 3.5. Geodetic Systems (2013) recommend using a convergent image with the angle of intersection between  $60 - 120^\circ$ . This is not just to increase the base length but could also be to increase the area of stereo coverage.

However, it is also well known that the increase of intersection angle can lead to a decrease in the digital matching points between overlapped images. As discussed latter, this is due to the fact that images become different to each other at higher IA. Therefore, in this investigation, the aim is to find the optimum angle of intersection that can be used with a stereo pair using AGISoft StereoScan to generate the textured 3D models from individual stereo pairs. In order to test the two samples (test09 and test10) with a different angle of intersection, the test was designed to read every 10 degrees from 10 to 90 with the distance of object to the camera fixed to 90 cm for all tests. This means that the distance of object to the camera remains the same while the intersection angle changes, figure 3.39.



**Figure 3.39: Designed test to read IA every  $10^\circ$  with fixed range**

The lowest Intersection Angle (IA) is  $10^\circ$  and it is then gradually increased every  $10^\circ$  until it reaches  $90^\circ$  that was the final possible IA that allows for automatic matching between the two overlapped images. Each camera has a total of 18 samples for both rough and smooth textures (test09 and test10). The five cameras mentioned before were utilized in this investigation, so a total of 90 samples were subjected to processing and analyses. To test the designed intersection angle, Australis tool was used for each pair of

images to work out the real intersection angle. The intersection angle calculated from Australis was found to be very close to the intersection angle acquired in the field for most samples and found to have less than 2.5 degree difference, except for some cases. Therefore, these differences were ignored and the overall calculation and analysis was based on the designed intersection angle, see table 3.6.

Table 3.6: Difference between the required intersection angle at field and calculated from Australis, each camera smooth (above) and rough (below)

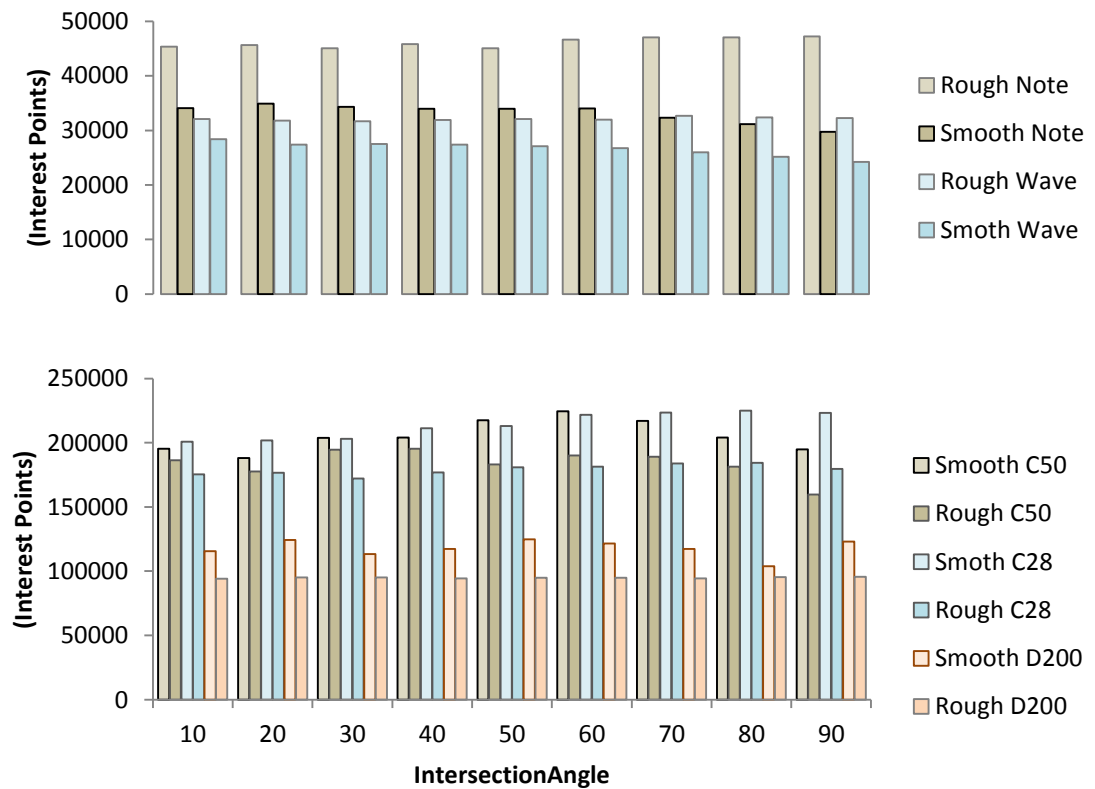
Field I. A. degree		10	20	30	40	50	60	70	80	90	
Cameras	Australis I.A. degree	Canon 5D 50 mm	10.5	20.8	29.4	41.2	52.2	60.6	71.5	81.8	92.5
			11.1	20.5	31.2	40.3	51.5	59.1	68.1	82.3	90
		Nikon D200	11	21.6	30.8	42.4	52.4	61.2	73	81.6	92
			7.0	18.8	29.8	41.0	48.0	62.6	69.5	81.8	89.5
		Samsung Wave	11.5	22.4	31.2	43.6	51.6	61.8	71	81.4	93.5
			10.1	21.5	28.4	40.5	50.7	62.2	70.7	82.0	90.3
		Samsung Note	12	22.2	31.6	44.8	51.8	62.4	72.8	81.2	91.5
			9.9	21.5	29.0	41.3	50.3	62.6	70.5	81.8	90.7
		Canon 5D 28 mm	10.2	20.8	30.8	41.8	50.9	61.2	70.4	81.8	91.3
			13.7	21.4	28.4	37	48.8	57.8	69.7	79.8	90.4

The objectives that will be presented in the following sections are to test the number of features or interest point's extraction, and tie points or matching point's generation at different angles of intersection on both rough and smooth textures. In addition, another objective is to investigate the effect of intersection angle on RMSE of the image points with re-projected solution. Another important parameter which requires to be examined at different intersection angles is the level of details that both cameras from a different view can share. Lastly, 3D model accuracy in terms of texture depth will be tested for both surfaces with utilizing different cameras compared with the reference depth from the sand patch test. The first parameter that needs addressing is the feature extraction, which is explained in the next section.

#### *A) Number of feature points*

The first step in digital image matching is a feature or interest point's extraction on both pair of images. Different techniques are available in order to extract important and significant features that can be visible on both images. The quality, but not the number, of features has influence on the matching process. A robust feature, in terms of scale and orientation, plays a significant role in finding out the relative orientation between the stereo pair. In this investigation, the number of features and their quality are studied using the rough and smooth texture tests at different intersection angles.

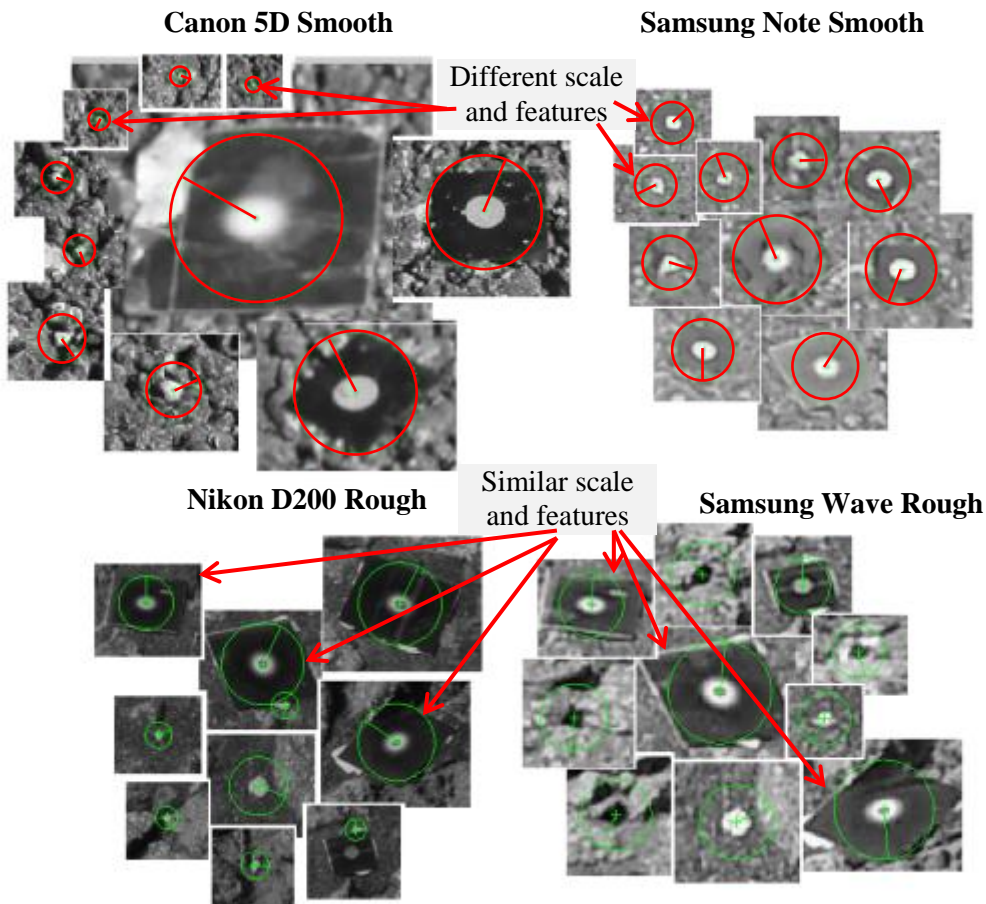
The feature point's extraction on both smartphones cameras, i.e. Samsung Note (Note) and Samsung Galaxy Wave (Wave), was different from DSLR cameras, the Nikon D200 (D200), the Canon 5D-50mm (C50) and Canon 5D-28mm (C28). On smooth texture, both DSLR cameras provided more feature points than on the rough texture. In contrast, smartphone cameras provided more feature points on rough than smooth texture, figure 3.40 shows the trend of average feature points from both images against the intersection angle on both surfaces (rough and smooth).



**Figure 3.40: Average Feature Extraction Vs Intersection Angle for both rough and smooth texture, Smartphones (above) and DSLR Cameras (below)**

It can be concluded from figure 3.40 that the number of features is dependent on the texture type and image resolution. It was found that the number of features was not affected by intersection angle within the camera to object distance tested. The features at 12 MP camera of Nikon D200 and 21.1 MP of Canon 5D are considerably different from those with 8 MP in the Samsung Note and 5 MP in the Samsung Galaxy Wave. In order to visualize the differences between the generated features at different camera resolutions, ten strongest features from both smooth and rough textures were extracted

using the SURF feature descriptor. For the smooth texture, the Canon 5D with 21 MP and the Samsung Galaxy Wave with 5 MP were used. For the rough texture, the Nikon D200 of 12 MP and the Samsung Note of 8 MP were used, see figure 3.41. Whereas, the ten strongest features on the smooth texture were entirely different in terms of scale, orientation and the feature itself, the repetition of the same feature scale and orientation on rough texture showed an apparent similarity, which clearly indicates the effect of resolution, texture type or feature extraction robustness.



**Figure 3.41: Ten strongest features at 70 intersection angle, smooth texture (above) and rough texture (below)**

It can be concluded from figure 3.41 that the size of features has a relation with the size of the pixel on the ground, with relatively bigger feature sizes on both smartphone cameras, i.e. the Samsung Note and Samsung Galaxy Wave when compared with the DSLR cameras, i.e. (the Canon 5D-50mm and Nikon D200).

### *B) Matching points or automatic tie points*

After the feature extraction on each image, the next step is to match keypoint descriptors between the pair of images. This helps to find the same features on both images. There are many methods and techniques for finding the descriptor that both images share, one example being kd-tree (Muja and Lowe, 2009). The results of different cameras showed that the automatic matching is decreased dramatically between the two images with the increase of the intersection angle and the relationship had an exponential trend line for all camera samples with a minimum regression of 95%, figure 3.42.

Difference in the number of matches between the two surfaces can be noticed on figure 3.42. In the previous section, both smartphone cameras (Samsung Note and Samsung Galaxy Wave) provided more feature points on rough than smooth texture, whereas, DSLR cameras, Canon 5D-50mm (C50), Canon 5D-28mm (C28) and Nikon D200 (D200), provided more feature points on smooth texture, see figure 3.40. Interestingly, the same pattern was exactly noticed on the matching points; the DSLR cameras provided more matches on smooth than on the rough texture which is opposite to smartphone cameras.

Another important conclusion is the degree of similarity between the data sets in the case of rough texture as can be noticed from figure 3.42. The t-test was used to find the probability of similarity between the mean of the smooth group and rough group data. Rough group data from both smartphone cameras registered probability of 0.9516; while that of the smooth group data from both smartphone cameras registered 0.8422. Average t-test probabilities from three DSLR camera sets were as follows; rough =0.8885 and Smooth =0.8012. These probability values support the conclusions drawn for the feature extraction which shows that the feature scale and orientation are almost same for the rough texture despite the camera resolution differences, see figure 3.41.

The results from figure 3.42 prove that the rough texture can provide matches and produces 3D surfaces even in higher intersection angle greater than 70°; for example around 90° that was not possible in the smooth texture. The maximum possible degree for 3D reconstruction in smooth texture was found to be 70°, despite the fact that the settings on the cameras remained constant and tests were carried out on the same day. These results revealed that the matching is not only dependent on the intersection angle, but also can be affected by texture type which can play an important role on the matching process. Different texture types can provide different features and subsequently affect the matching process using the same camera.



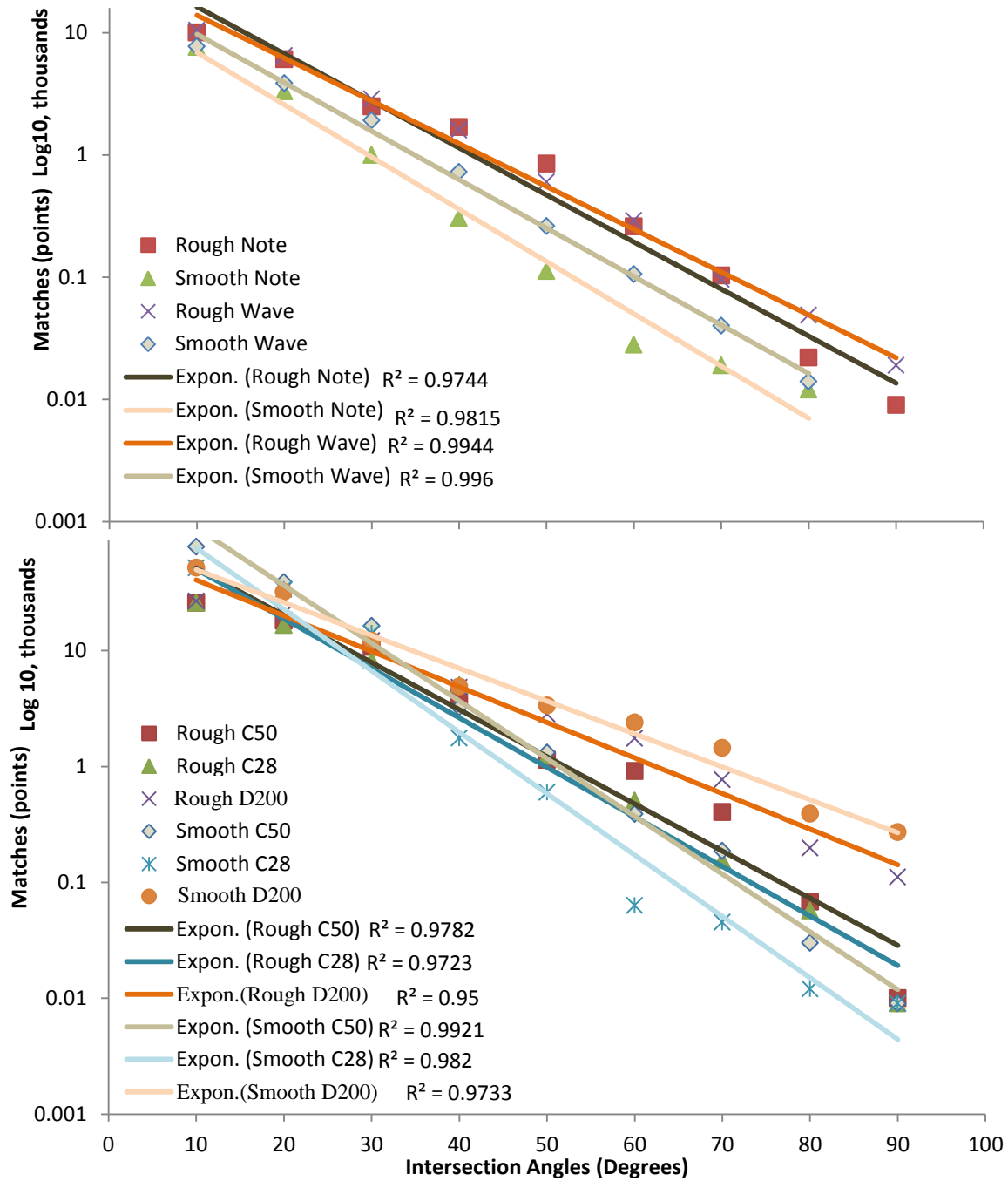


Figure 3.42: Relationship between intersection angle and number of matches for both smooth and rough textures (Smartphones above and DSLR Cameras below)

AGISoft and most automated 3D models, such as ARC3D, PhotoSynth, Photo truisim, Bundler, ATiPE, and AGISoft employ scale invariant feature transform algorithms such as SIFT and SURF for describing the point of interest (Barazzetti et al., 2010; Photosynth, 2012; VISICS, 2012; Agisoft, 2015). Therefore, it is usual to have such a relationship that is shown on figure 3.42, because as the angle of intersection



increases, the affine transformation between the two images increases which results in less matching points. For more information about the affine transformation effect on matching points refer to Juan and Gwun (2009). The feature of interest's shape and size are changing with increasing intersection angle resulting in a less accurate descriptor. Figure 3.43 shows a sample of a square target cropped from the Canon 5D, 50mm images at different intersection angles. It can be noticed from figure 3.43 that the size and shape of target changes as the intersection angle increases. This normally can happen with all features or keypoints at higher intersection angle.



**Figure 3.43: change in the target size and shape due to intersection angle**

Results from this test were found to be compatible with the work carried out by Hasegawa et al. (2000) which was based on cross-correlation matching for aerial photography. Their results confirmed that the view angle differences between the left and the right images become larger as the B/H ratio increases. The highest possible B/H ratio for successfully producing matches was found to be 1.4 which is equivalent to  $70^\circ$ , see table 3.5. However, they proved that the correlation coefficient becomes smaller as undulation of terrain increases and that was not the case in this project. The same conclusions as stated in Hasegawa et al. (2000) can be applied to DSLR cameras but not to smartphone cameras. The results on this project showed that the smooth texture provides more points than rough texture in case of DSLR cameras but not smartphone cameras. The possible reasons might be due to the relative small size of pixel sensor size in smartphone cameras, which is about 1/6 of DSLR cameras.

### *C) RMSe of image point residuals from reconstruction process for different IA*

RMSE is an indication of how large or small the residuals are from the solution, while in turn it can relate to how good the derived parameters are used for estimation. RMSE for all cameras (DSLR and smartphone cameras) and for both surfaces rough and smooth was found to increase with the increases of intersection angle, figure 3.44. The increases in RMSE might be attributed to the decrease in the number of matching points as the angle of intersection increases. The high number of matching points allows the

least square solution to pick up the best available matches with low residuals that are not possible in case of limited available matching points in higher intersection angles. Therefore, it is necessary for any statistical process that a reasonable sample size is available (matching points) for all tests in order to get comparable results.

The linear relationship is the best trend describing the RMSE data at different intersection angles for all DSLR cameras (C50, D200) and with a minimum regression of about 96% except for Nikon D200 (D200) in case of the smooth texture which is about 80%. RMSE of rough texture data is smaller and more consistent than the RMSE of the smooth texture data. In addition, the minimum  $R^2$  (measure of consistency) value recorded for the rough texture is 96% that indicate the good consistency between the RMSE data at various intersection angles. In contrast, RMSE for both smartphone cameras (Note and Wave) are inconsistent, and the  $R^2$  values tend to be lower than that of the DSLR cameras. The  $R^2$  value might be a good indication of the best camera that has a systematic trend relation, as in the case of Canon 5D-50mm (C50) and Canon 5D-28mm (C28) that have both  $R^2$  value of more than 97%.

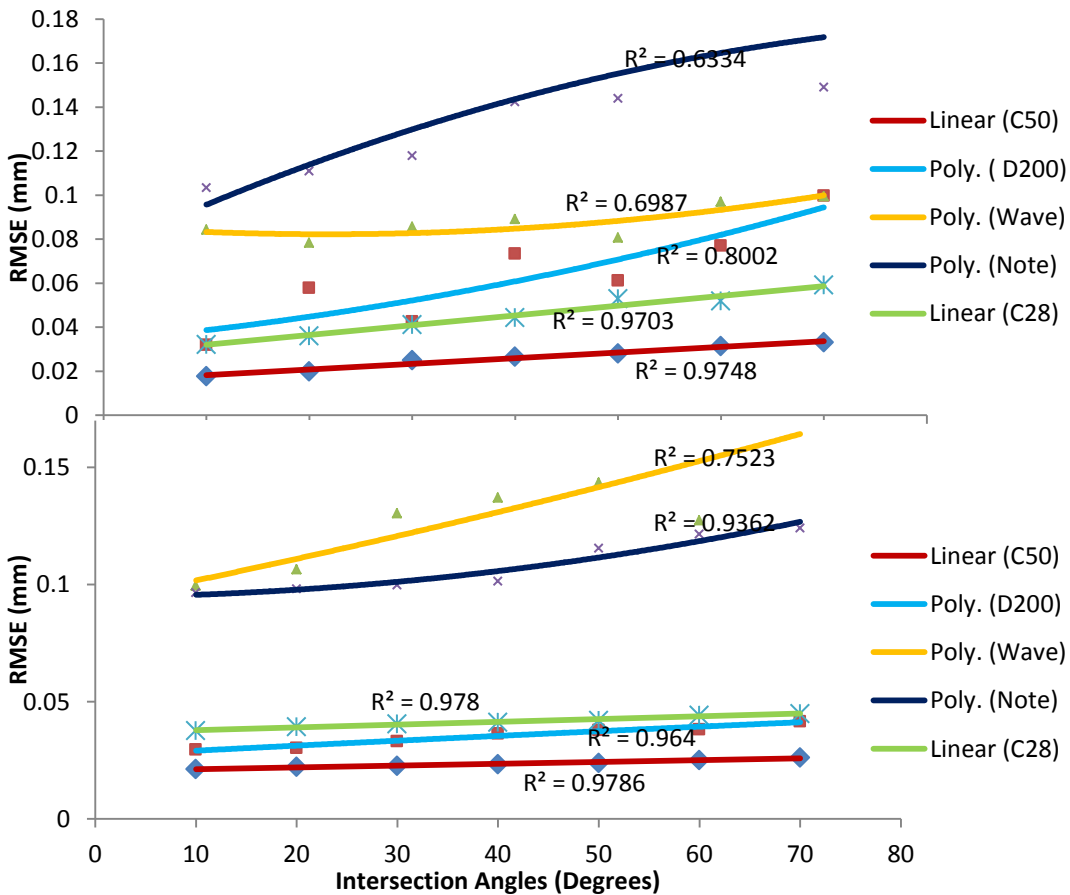
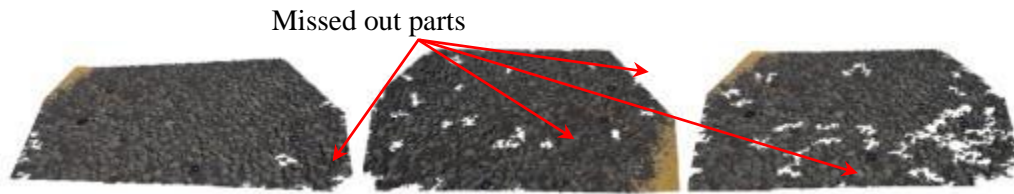


Figure 3.44: RMSe vs Intersection angle (Smooth texture above and rough texture below)

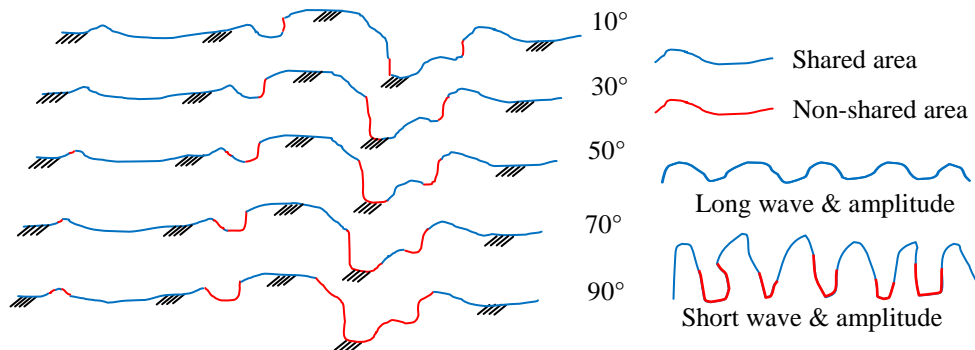
#### D) 3D model level of details vs. intersection angle

The increase in intersection angle in the rough texture leads to missing out of some part of the reconstructed surface, figure 3.45. When the angle of intersection becomes more than  $50^\circ$ , the 3D reconstructed model experiences some missed out parts on the rough texture due to non-availability of matching points on those parts (dead ground). The cameras from two different views observe different feature points due to raised up aggregate which results in hiding some of the details that are not shared between the two views.



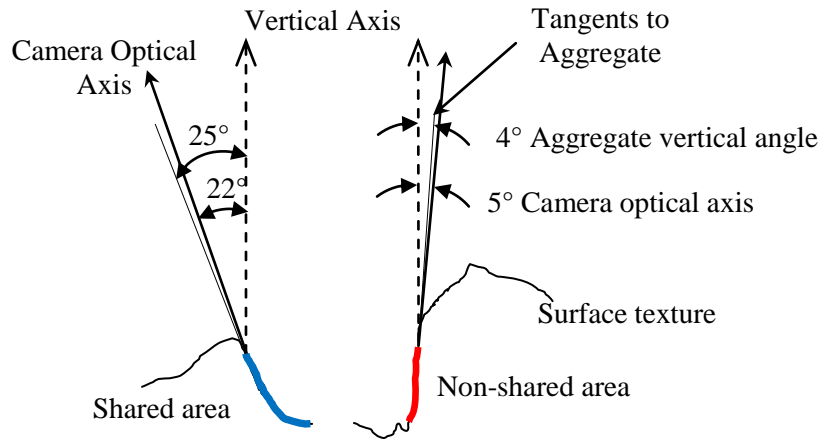
**Figure 3.45: Non complete 3D reconstruction due to increase of IA ( $50^\circ$ ,  $60^\circ$  and  $70^\circ$ ), left to right**

Figure 3.46 is made using CAD, and it shows how increase in intersection angle can hide some sharing features between the two views. The angle of intersection in this case is starting from  $10^\circ$ ,  $30^\circ$ ,  $50^\circ$  and up to  $90^\circ$ , and the cross section is made by using a plane passing horizontally through the centre of the rough texture (test10). The Z axis scale is exaggerated for demonstration purposes. The blue colour shows the shared area that both cameras can see from two different views and the red colour shows the areas that either only one camera can see or none of them. The details of data that can be shared between the two views are not only intersection angle dependant but also wavelength and peak-to-peak amplitude of the texture dependant (see section 2.3.5 for more details about these two characteristics), as figure 3.46 shows.



**Figure 3.46: Share and non-share features vs. different angle of intersection in rough texture**

It can be concluded from figure 3.46 that the vertical aggregate angle (the angle that is enclosed between the tangent to the aggregate surface and the vertical axis) should be more than the angle that the optical axis of the camera makes with the vertical axis in order for the area behind aggregate to be seen. Figure 3.47 shows some parts of the aggregate from the rough texture sample presented on figure 3.46.



**Figure 3.47: Camera optical axis vs. aggregate vertical angle**

#### *E) 3D reconstructed model accuracy vs. intersection angle*

The generated 3D model is registered, cropped and oriented to Z+ (see sections 3.7.2.2.1, 3.7.2.2.2, and 3.7.2.2.3 for more details about these procedures). The depth of the texture within the 10 cm sample is then compared against the intersection angle for both control areas, smooth and rough, see figure 3.48. Depths at different intersection angles were subtracted from the SPT depth that was treated as a mean value for standard deviation calculation. The depth calculated from SPT was found to be 670  $\mu\text{m}$  and 1970  $\mu\text{m}$  for both smooth and rough texture, respectively. The differences between SPT depth and depths at different intersection angles for each camera is squared and summed and divided by the number of observations in order to calculate the standard deviation, see table 3.7. Standard deviation is a good indication of the camera and software reliability for providing consistent data with the mean value (SPT depth).

The standard deviations listed in table 3.7 show that the best result of depth calculation was for the Canon 5D-50mm (C50) followed by the Canon 5D-28mm (C28), the Nikon D200 (D200) and the last two smartphone cameras (Note and Wave) and for both textures (smooth and rough). The relative high standard deviation of some cameras

is due to the distance of the camera to object (90cm) which will be explained later in section 3.8.2.1.2.

Table 3.7: Standard deviation of different cameras at different intersection angles calculated using mean sand patch depth as true value

Cameras	Standard deviation	
	Smooth Texture ( $\mu\text{m}$ )	Rough Texture ( $\mu\text{m}$ )
Canon 5D-50mm	203	225
Canon 5D-28mm	337	426
Nikon D200	444	656
Samsung Galaxy Wave	590	1400
Samsung Note	490	1250

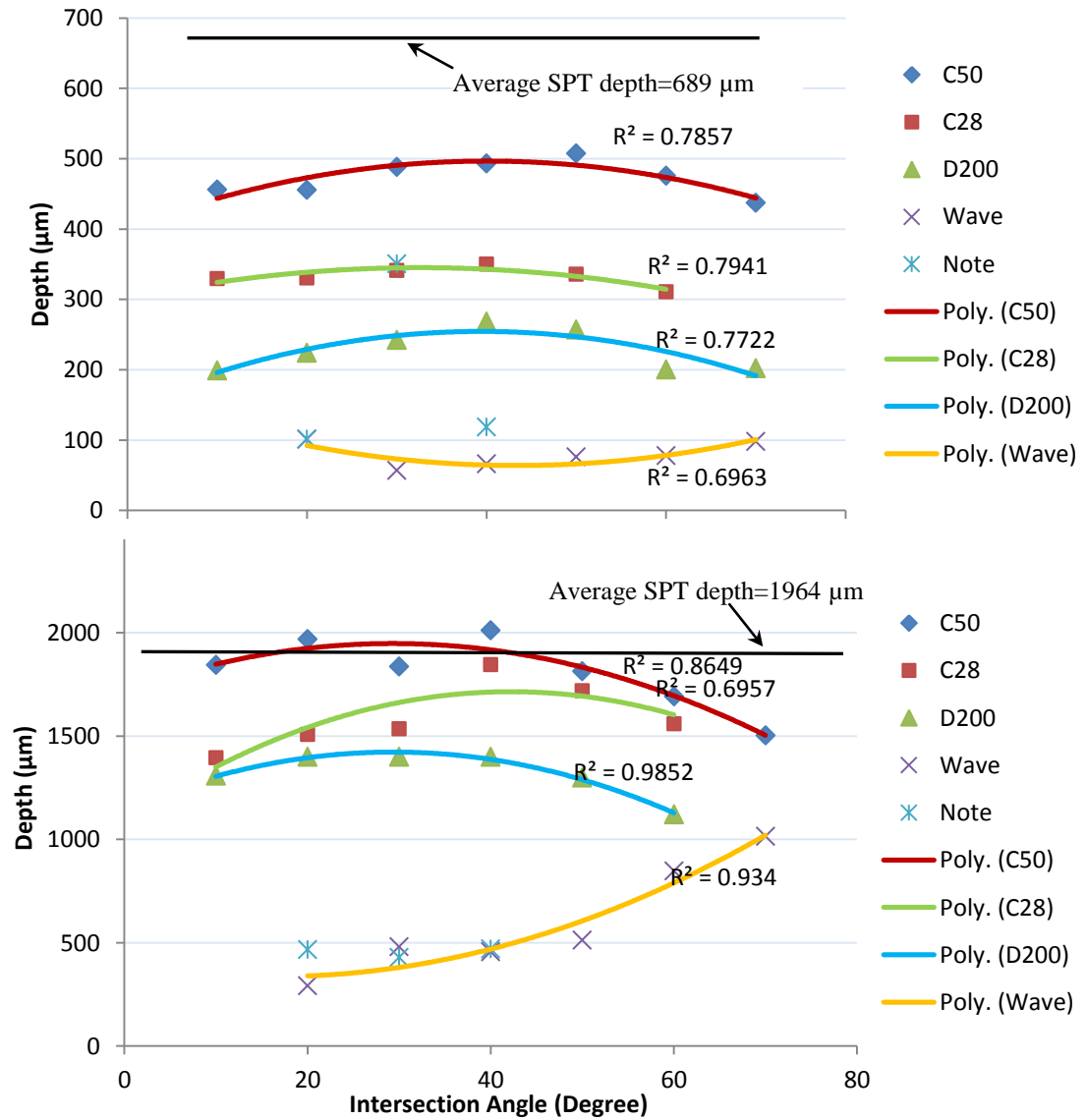


Figure 3.48: Texture depth vs intersection angle, smooth texture (above) and rough texture (below)

As can be noticed from figure 3.48 and the standard deviations listed in table 3.7, the Canon 5D-50mm (C50) recorded the best results. Many reasons are behind this. First of all is the camera geometry in the case of Canon 5D for both 28 mm (C28) and 50 mm (C50) which utilizes the full frame CCD size (36mm) and the lens quality as well. The focal length plays an important role in providing the best results as in case of 50mm. Both Canon 5D (C28) and Nikon D200 (D200) share the 28mm focal length property, however, the results of Canon-5D are more accurate than Nikon D200 (D200), which could be attributed to CCD size or the better quality of the lens in case of Canon 5D (see section 3.6.1 for details and specification about each used camera).

It is worth to mention that both smartphone cameras have provided less accurate results which could be attributed to their low focal length (4mm) and small CCD size. Regardless to the accuracy differences of each camera, DSLR cameras are showing that the best results of intersection angle is between 30-50°. This range is a balance between retaining the positional and depth accuracy with a successful matching and feature point generation. Therefore, the intersection angle of 45° was decided to be used for testing the distance of the camera to object, as explained in more detail in the next sections. Figure 3.49 shows samples of rough textures with visual coded colours that describe the effect of intersection angles on geometry. The mid samples 30-50° represent the balance between sustaining accuracy and better reconstruction parameters such as the number of matching and reconstruction points that helps in 3D automation.

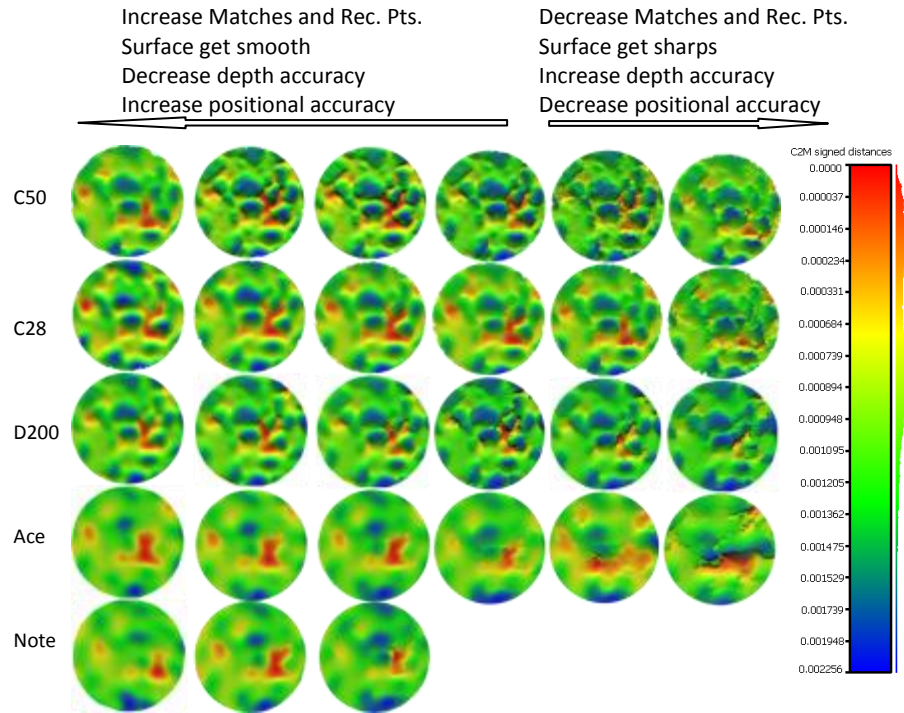


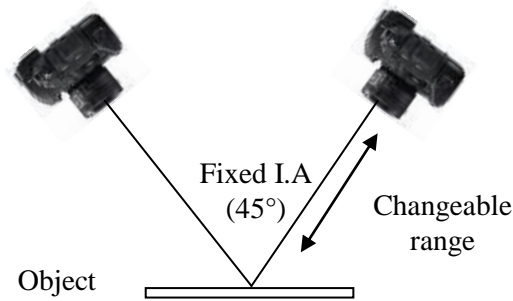
Figure 3.49: Samples of rough textures at different IA, (10-60°) left to right

### 3.8.2.1.2 The effect of the distance from camera to object

The two control textures (rough and smooth, test10 and test09, respectively for control samples shown in figure 3.4) were again examined to test for the effect of camera to object range on the accuracy of the 3D generated models. Again the five cameras presented on section 3.6.1 were used for investigating the distance of the camera to object and their effect on accuracy and results. In theory, the camera to object distance, or flying height in the case of aerial photography, has an influence on the ground sample distance (GSD) and hence on the accuracy. The closer the camera to the object, the smaller the GSD and the better the accuracy of positioning (Abdullah, 2010; Olivier Kung et al., 2011; Strecha, 2011) .

In this study, the camera to object distance of the camera is variable from the minimum focusing distance and up to 120cm on a 10cm interval basis. The minimum focusing of camera to object distance for different cameras was found to be as follows (Canon 5D-50mm (C50)=35 cm, Canon 5D-28mm C(28)=20.8cm, Nikon D200 (D200)=20.5cm, Samsung Wave (Wave)= 19,5cm and Samsung Galaxy Note (Note)=21cm). While the camera to object distance is changing, the intersection angle remains the same for all the tests and is fixed at 45° (figure 3.50), the best possible

intersection angle concluded from section 3.8.2.1.1. The camera to object distance of all cameras was started from 20cm up to 110cm except for the Canon 5D-50mm (C50) that started at 35cm and up to 125cm. Each camera has a total of 10 samples for each texture and so 20 samples for both rough and smooth surface textures. As mentioned before, four cameras with five lenses were utilized, making a total of 100 samples. All samples were generated and prepared using the methodology mentioned in previous sections. The aim of this test is to study the effect of the camera to object distance of the camera on each of feature, matches, RMSE or reprojection error, and the accuracy of 3D generated model using StereoScan software on a pair of images.



**Figure 3.50: Approach to find optimal camera range at constant intersection angle**



### A) Number of feature points vs camera to object distance

A relationship was drawn between the feature points or points of interest and the range for each of the mentioned cameras and for both textures (rough and smooth). The relationship found was different from the rough to smooth surface and from one camera to the other. Both smartphone cameras showed that the number of feature points decrease with the increase of camera to object distance for the smooth texture, figure 3.51. The decrease in the number of feature points in both smartphone cameras indicated that the size of extracted feature increases when the distance increases between the camera and the object.

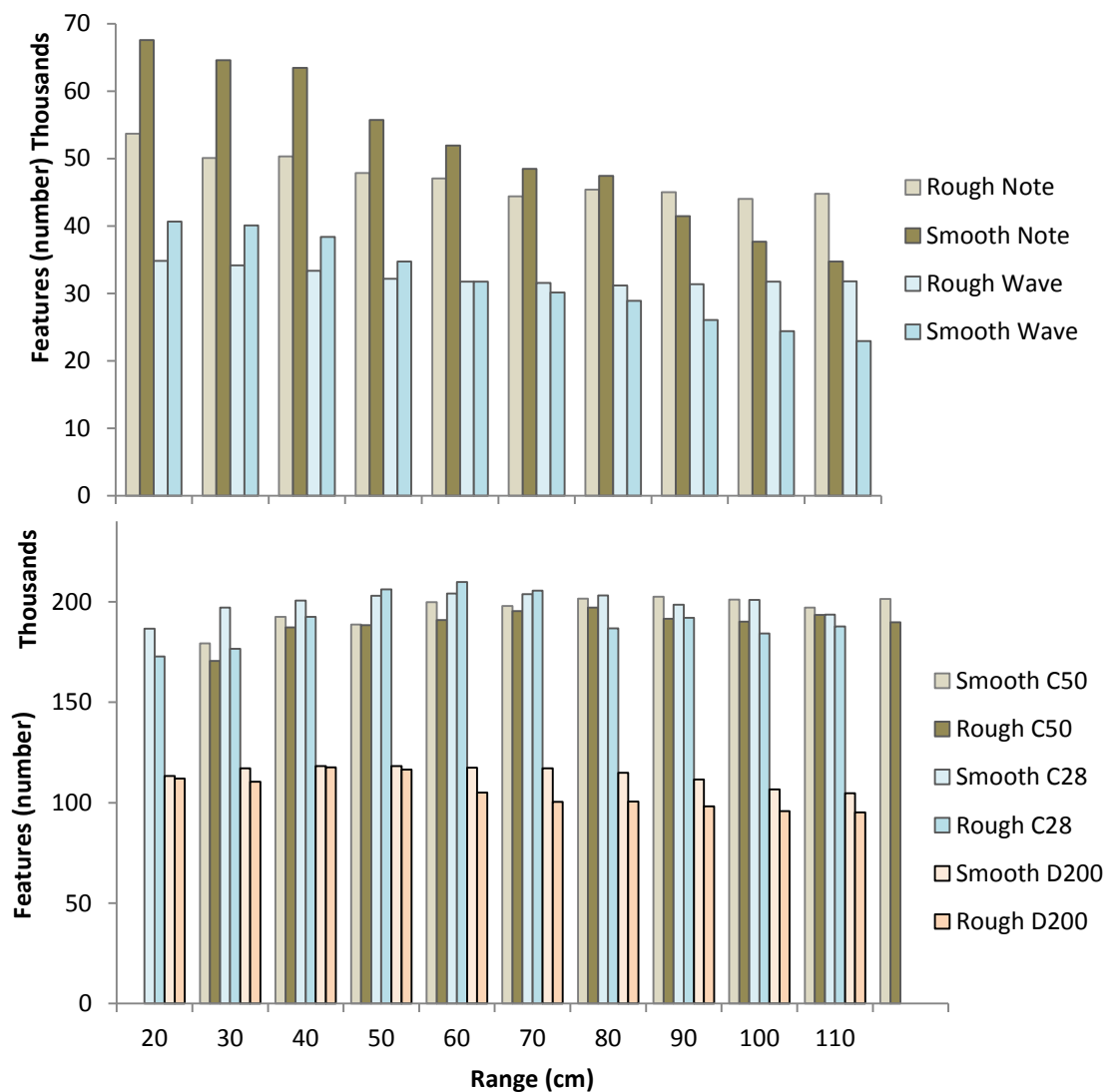
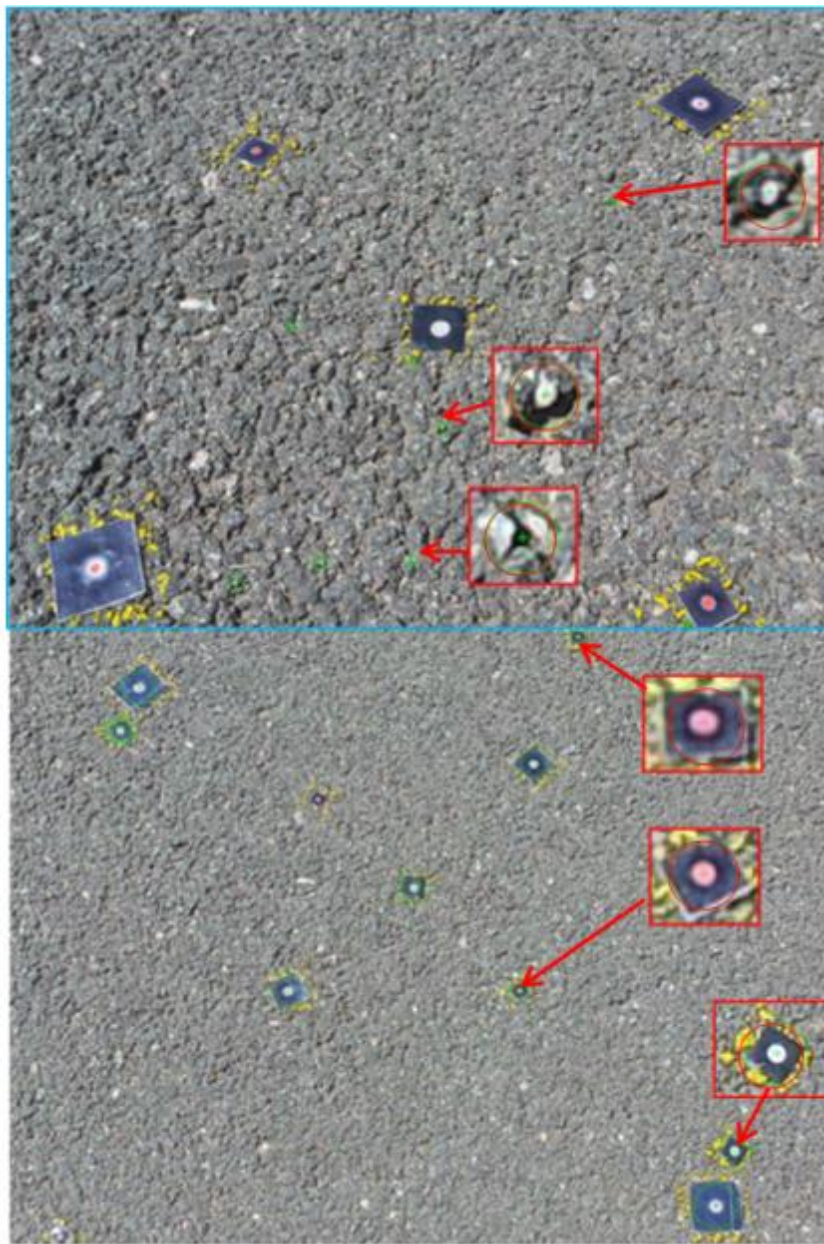


Figure 3.51: Feature point Vs camera range, smartphone cameras (above) and DSLR cameras (below)

Although the image is covering a greater area at 110 cm range than at 20 cm range (about 18 times bigger), the features extracted at 20cm are much smaller in size than the features extracted at 110cm as shown in figure 3.52. The size of the smaller feature at 20 cm is a 10<sup>th</sup> of the size of the feature at 110 cm and more features can be extracted at the shorter camera to object distance as can be seen in figure 3.51 for both smartphone cameras. However, with rough texture and in the case of smartphone cameras shows a steady decrease as the camera to object distance of the camera increases, see figure 3.51.



**Figure 3.52: Ten strongest features extracted from 20 cm (above) and 110 cm (below) using Samsung Note camera**

The trend of DSLR cameras with regards to feature extraction and camera to object distance of the camera is totally different from smartphone cameras for both textures. While the camera to object distance increases, the features are also increased for all DSLR cameras, i.e. Canon 5D-50mm(50), Canon 5D-28mm (28) mm and Nikon D200 (D200), up to some limits then it starts to decrease, see figure 3.51. However, the only difference between both DSLR cameras is the limit of the peak range. With the peak of the Nikon D200 (D200) camera around 60 cm while for Canon 5D-28mm (C28) and Canon 5D-50mm (C28) it is around 60cm and 95 cm, respectively. In order to precisely relate the peak's camera to object distance to the GSD, a table of GSD for each of the camera calculated against the camera to object distance, see table 3.8.

Table 3.8: Ground sample distance vs camera to object distance of variety of cameras

Cameras	C50	C28	D200	Wave	Note
Camera to object distance (cm)	Focal length(mm)				
	50	28	28	4	4
	Sensor Pixel Size (mm)				
	0.0064	0.0064	0.0061	0.0015	0.0014
GSD (mm)					
20	0.026	0.046	0.044	0.075	0.070
30	0.038	0.069	0.065	0.113	0.105
40	0.051	0.091	0.087	0.150	0.140
50	0.064	0.114	0.109	0.188	0.175
55	0.070	0.125	0.120	0.206	0.192
60	0.077	0.137	0.131	0.225	0.210
70	0.090	0.160	0.153	0.263	0.245
80	0.102	0.183	0.174	0.300	0.280
90	0.115	0.206	0.196	0.338	0.315
95	0.121	0.217	0.207	0.356	0.332
100	0.128	0.229	0.218	0.375	0.350
110	0.141	0.251	0.240	0.413	0.385
120	0.154	0.274	0.261	0.450	0.420

As can be noticed from table 3.8, the GSD at 95 cm for the Canon 5D-50mm (C50) is about 0.12 mm. This value is the peak feature for Canon 5D-50mm (C50), and it is located between 50 and 60 cm for Canon 5D-28mm (C28) and Nikon D200 (D200), refer to figure 3.51. On the other hand, the peak value for both smartphone cameras is located 'between' 20 to 40 cm. In theory, the relationship for both smartphone cameras should be close to the DSLR cameras and the values should start increasing until it reaches the peak around 30-40 cm. However, considering the camera calibration, the real camera to object distance of the camera from the ground, the quality of the lenses, the exposure settings differences and the oblique image acquisition, the relationship drawn in practice will be different from those in theory.

The peak of the extracted features also shows slight differences from the smooth to the rough texture. In order to investigate this fact, the strongest feature from both the rough and the smooth textures were extracted, figure 3.53. The extracted strongest feature is not in term of size only but it is in terms of how the feature is identical and robust to scale and rotation changes from all other features. As can be noticed from the extracted strongest features, the size of the feature is smaller in the smooth than in rough texture.

To summarize, there is no fixed relationship that can be drawn between the features and the camera to object distance of the camera. The relationship is changing depending on many factors, for example the texture type, the feature distribution, and the size of features can all play an important role in this relationship. It is also important to bear in mind that there is a link between feature size and camera resolution. Different resolution cameras can extract different features at the same level due to the GSD size on the object. It can be concluded that feature extraction is dependent on many factors, such as camera to object distance of camera, focal length, camera's resolution and texture type.

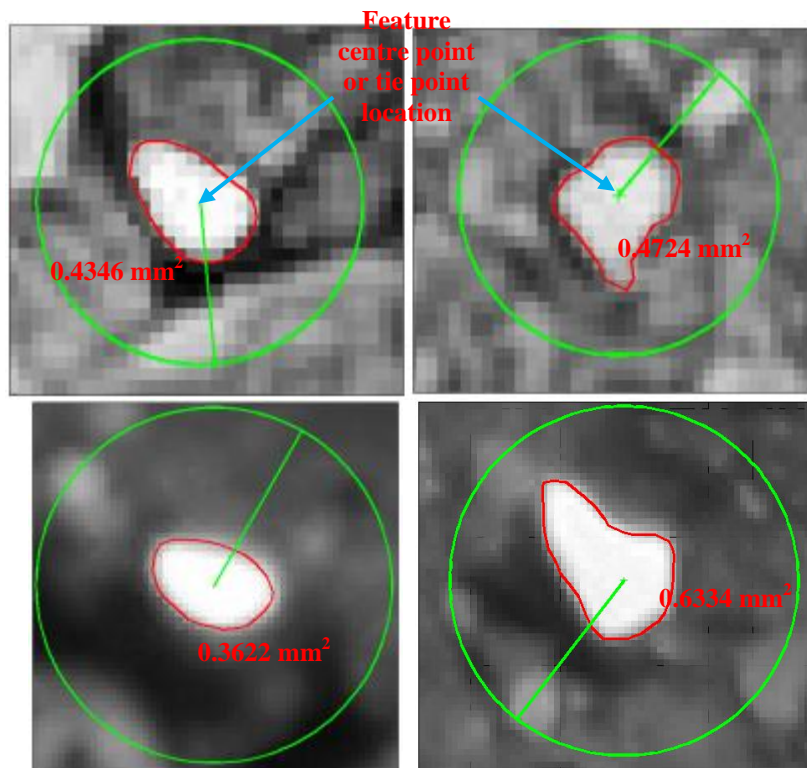


Figure 3.53: Strongest extracted feature at 20 cm Samsung Wave Camera (above) and Canon 5D (below), smooth (left), rough (right)

*B) Matching points or automatic tie points vs camera to object distance*

Not surprisingly, the relationship between the distance of the camera to object and the number of matching points is reversed. It shows that the matches decline with increasing the distance between the camera and the object for both smooth and rough textures, figure 3.54. These results confirm the theory that as the camera moves away from an object the GSD gets bigger, therefore, less matching points are observed (Abdullah, 2010; Olivier Kung et al., 2011; Strecha, 2011).

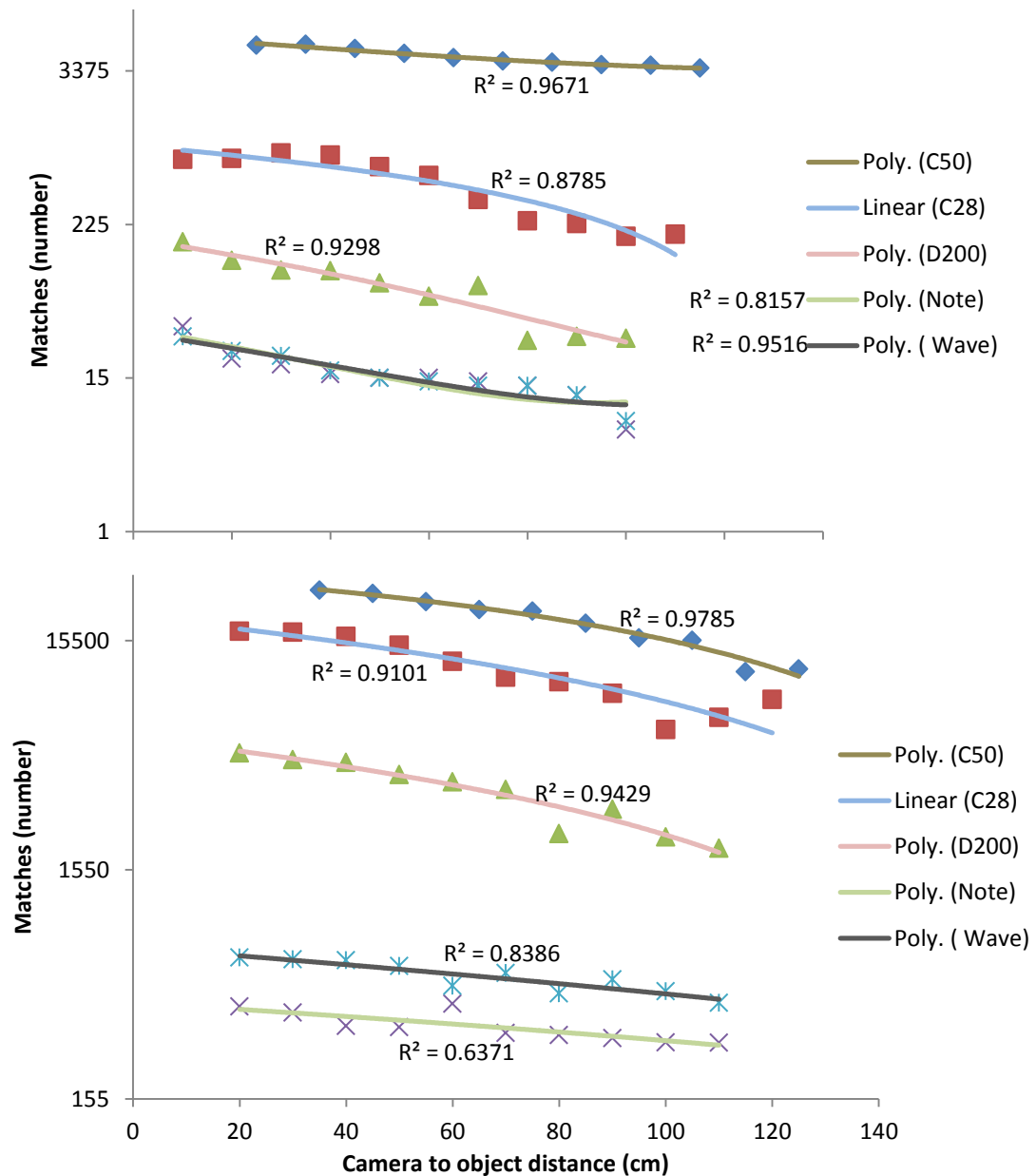


Figure 3.54: Matches vs camera to object distance for smooth texture (above) and rough (below)

Both surfaces, smooth and rough, showed that the number of matching points decreases with the increase in the distance of the camera from object. However, there is a difference in the number of matches for the two surfaces. Smooth texture provides very few matches compared with the rough texture. It is also worth mentioning that both tests were carried out on different days but with the same lighting condition (sunny days). It is clear that the texture type has a great influence on matching points as can be seen from figure 3.54.

### C) RMSE of image points vs camera to object distance

On first inspection, the RMSE of image points is not really influenced by the distance of the camera to object. The only conclusion is the apparent effect of the camera resolution on the RMSE or reprojection values. The Canon 5D-28mm (C28) provided better results if compared with other cameras, see figure 3.55. It is important to bear in mind that values in figure 3.55 are the outcome from AGISoft report and in units of pixels which need to be converted to mm to reflect the correct measurement units.

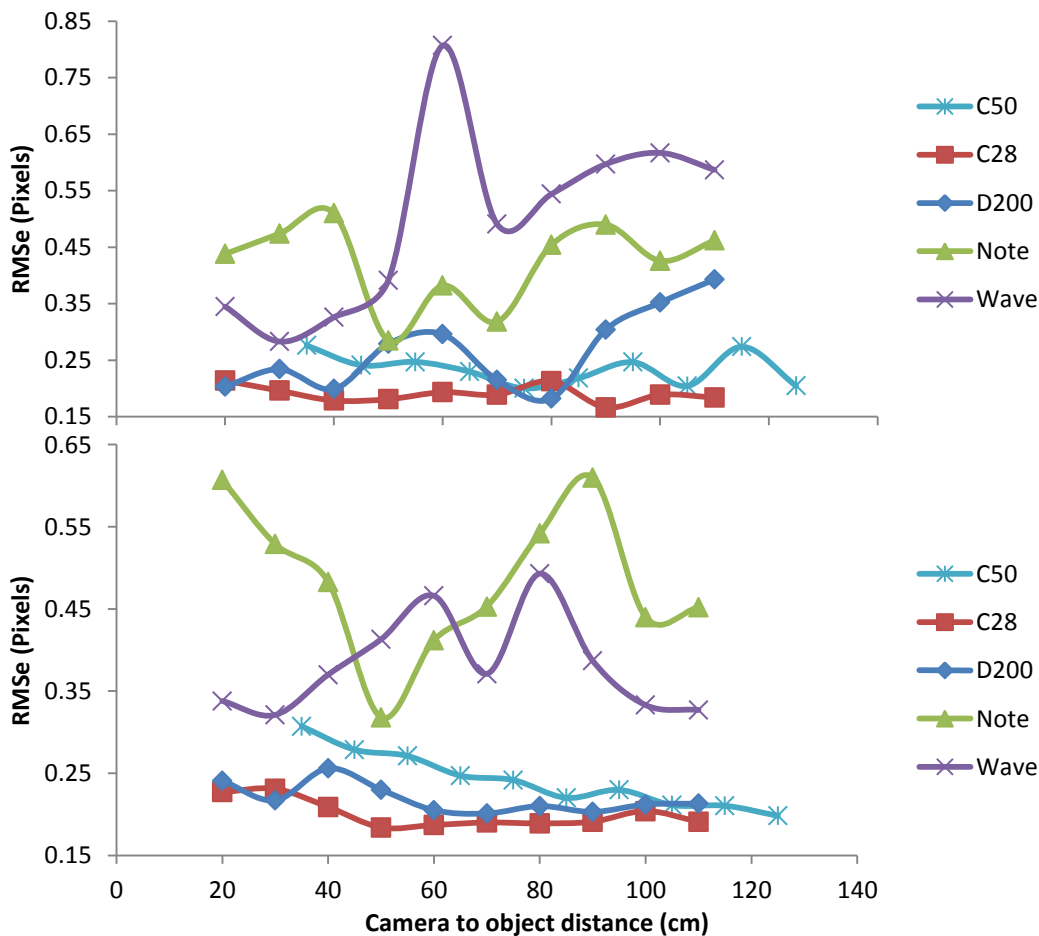


Figure 3.55: RMSe vs camera to object distance (pixels unit), smooth above, rough below



After converting units from pixels into mm, the relationship showed that the RMSE is increased with the increase of the camera range, see figure 3.56. As usual, the best accurate results are again for the Canon 5D-50mm (C50), followed by the Canon 5D-28mm (C28), Nikon D200 (D200) and at last the two smartphone camera of Samsung Galaxy Note (Note) and Wave (Wave). It is worth to mention that these RMSE values are from reprojection of object points into the 2D image.

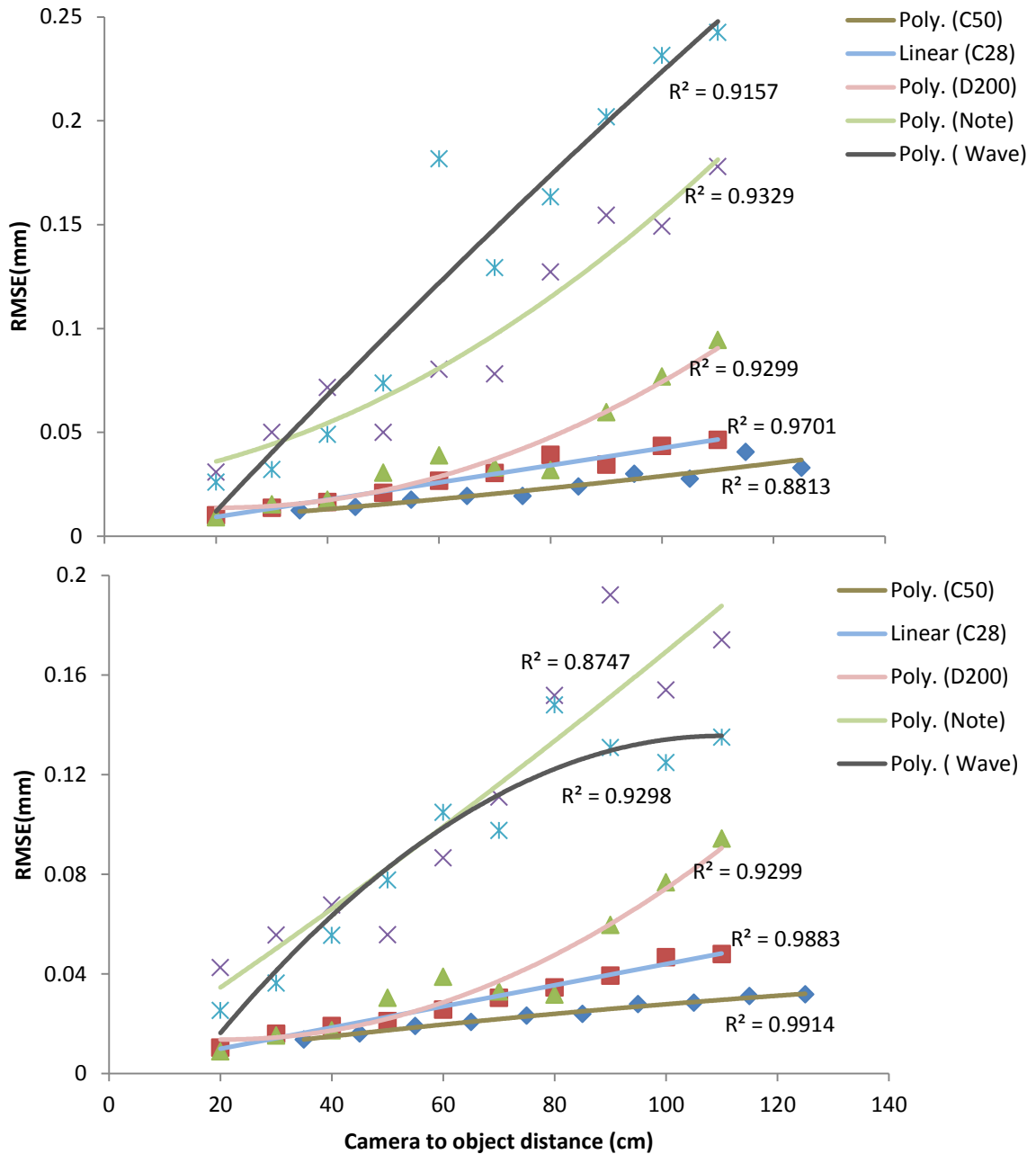


Figure 3.56: RMSe vs camera to object distance (mm units), smooth texture above, rough texture below

*D) 3D reconstructed model accuracy vs. camera to object distance*

As it has been previously discussed, the 3D models of both surface textures (rough and smooth) were prepared as 10 cm circular samples considering different camera to object distances at fixed intersection angle. The depth for each of samples was calculated using the developed methodology for texture depth measurement explained in section 3.7.2.2.4. The depths measured from 3D samples were all plotted against the camera to object distance, and the results present in figure 3.57.

As it can be seen from figure 3.57, the trend of all cameras is showing that the depth is getting smaller as the distance of the camera to object increases. It is also proved that the rough texture is better in terms of 3D model generation as it was possible to create 3D models at different camera to object distances for both smartphone cameras, which was not the case for smooth texture. In the smooth texture, it was only possible to create the 3D models for both smartphone cameras when the camera was very close to the object (at 20cm). For all other ranges of the smooth texture in smartphone cameras, the free program AGISoft StereoScan was unable to produce 3D models. The failure in 3D reconstruction for the smartphone cameras in the case of smooth texture could be attributed to the texture types and feature robustness which subsequently affect the matching process as showed previously.

The standard deviation of all cameras for both textures was calculated with reference to the mean depth from SPT, see table 3.9. The mean depth of SPT from both surfaces was found to be 670  $\mu\text{m}$  and 1970  $\mu\text{m}$  for both smooth and rough textures, respectively. As the table shows, the standard deviation results revealed that the depth calculation using 3D models became better when different camera to object distances were considered at a fixed intersection angle of 45°, see table 3.7. Standard deviation from the rough texture showed much better improvement than the smooth texture, especially in the case of the Canon 5D-50mm (C50) that recorded  $\pm 93 \mu\text{m}$  only.

Visual inspection of the rough texture samples in figure 3.58 reveals that Canon 5D-50mm (C50) provides enough detail at different camera to object distances starting from 35cm on upper left and up to 135 cm on the lower right. As it can be noticed from figure 3.58 the selected area within the 3D model is showing a bump in its lower zone (circled). This bump starts to disappear on all the samples depending on the distance of the camera to object, except for the Canon 5D-50mm (C50) where it remained the same.



This is because the details disappear as GSD gets bigger as the camera moves away from the object.

Table 3.9: Standard deviation of different cameras at different ranges calculated using mean sand patch depth as true value

Cameras	Standard deviation	
	Smooth Texture ( $\mu\text{m}$ )	Rough Texture ( $\mu\text{m}$ )
Canon 5D-50mm	200	93
Canon 5D-28mm	291	338
Nikon D200	281	347
Samsung Galaxy Wave	303	785
Samsung Note	350	740

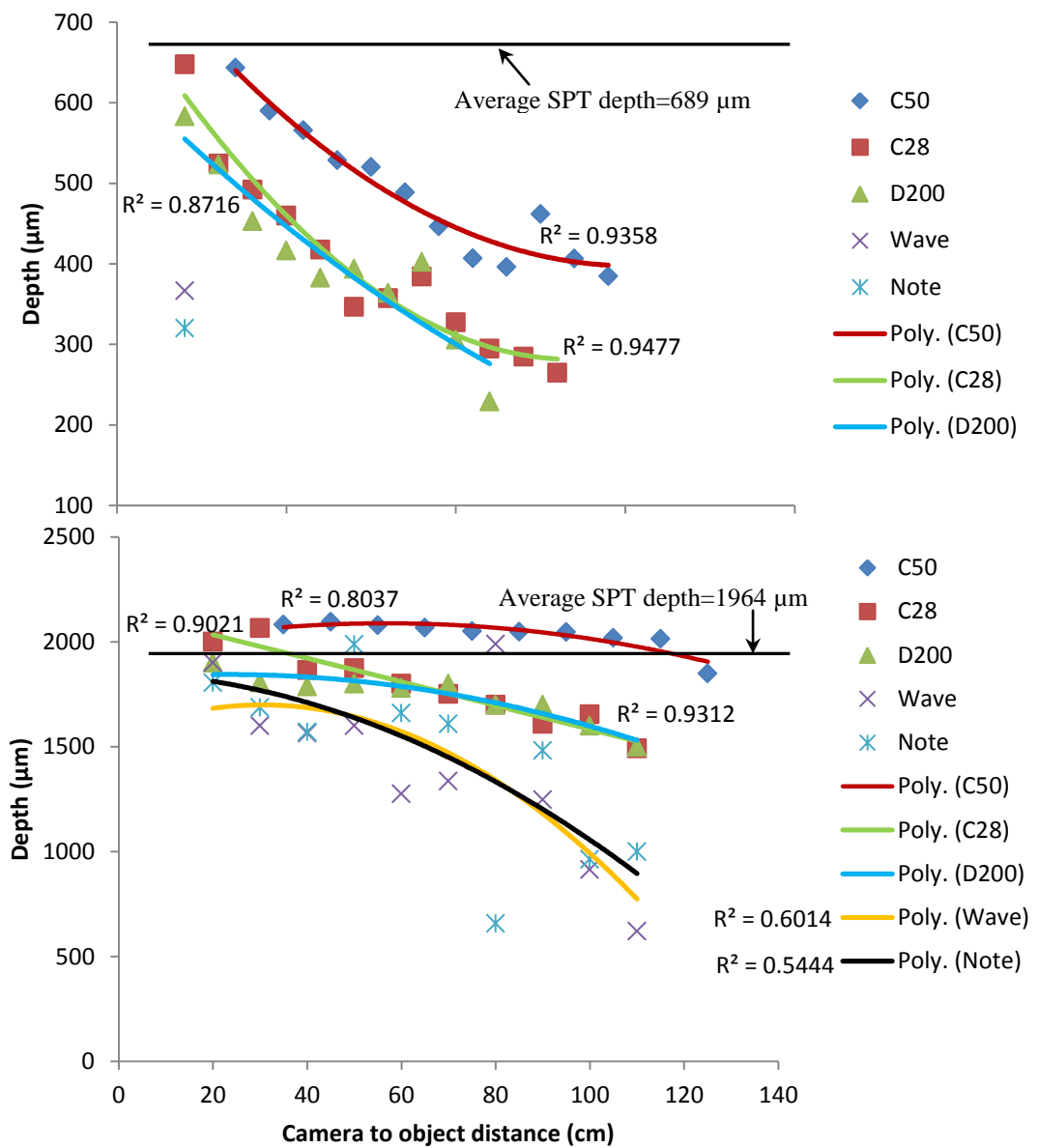


Figure 3.57: Texture depth vs camera's range (Smooth texture above and Rough texture below)

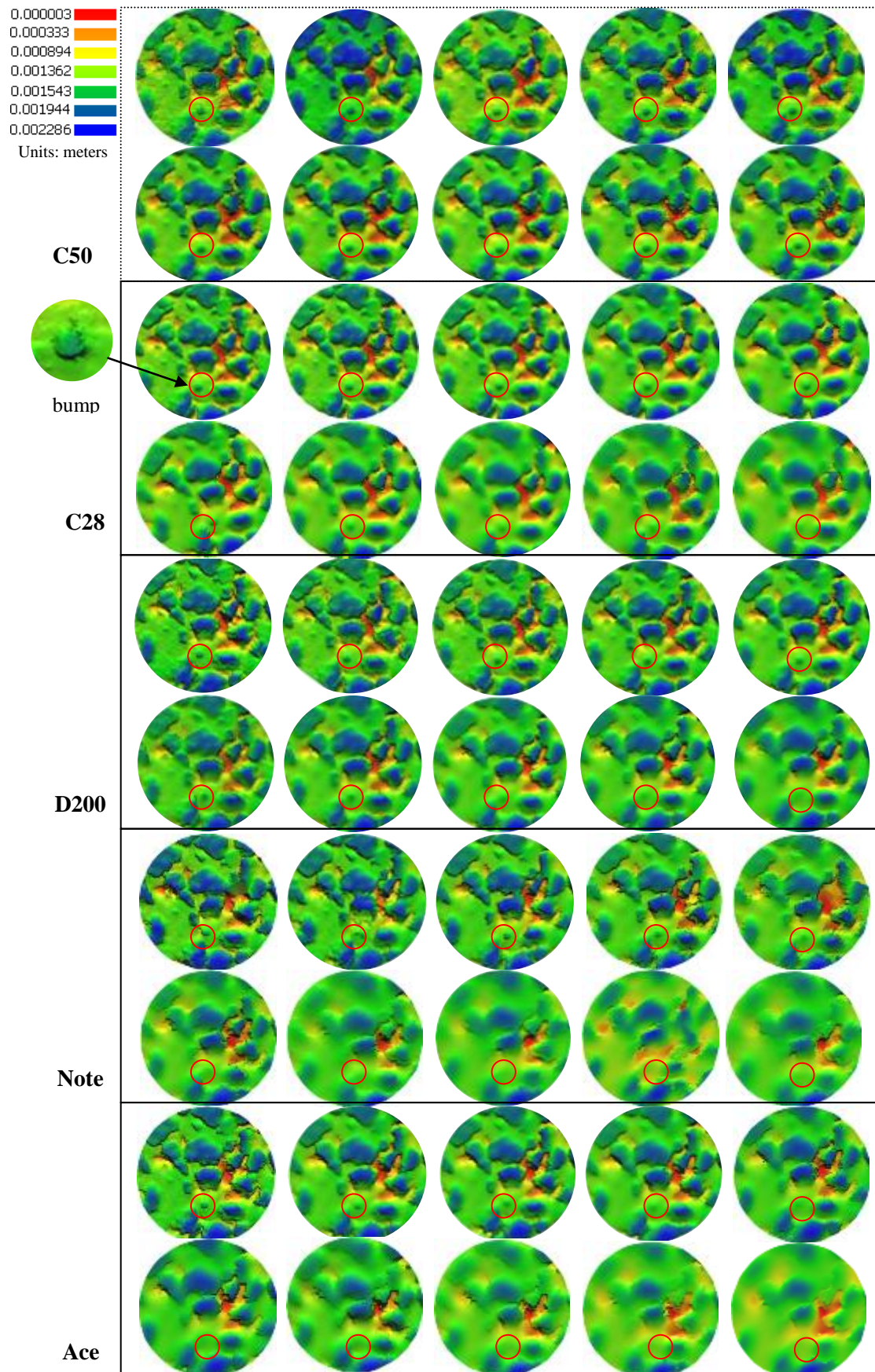


Figure 3.58: 3D samples of rough texture using colour as a height ramp

### **3.8.2.1.3 Outcomes from intersection angle (IA) and camera to object distance investigations**

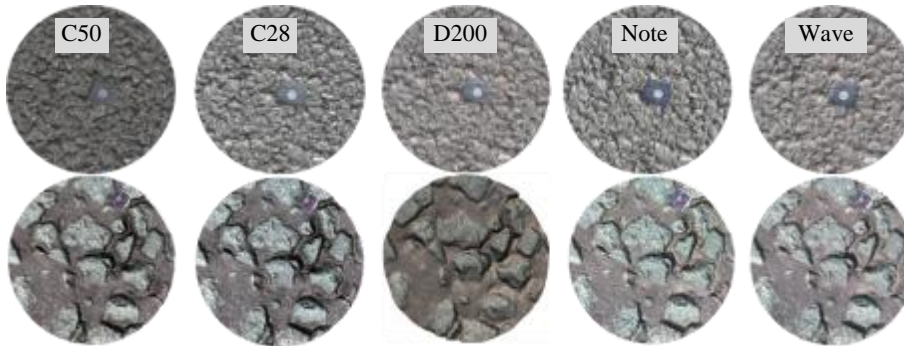
A series of investigations were carried out in order to examine both intersection angle and the effect of the camera to object distance on a number of outcome parameters by using the free 3D automation tool (AGISoft). Investigations of both the camera to object distance and the intersection angles proved that the feature points are not affected by the intersection angle at all for both smooth and rough textures. Nevertheless, it was concluded that the distance of the camera to object has a considerable influences on the feature point extraction. The later was also shown to be affected by the texture type. The distance of the camera to object can affect the size of the features recorded, and thus plays an important role in determining the number of extracted feature points. The results also showed that the affine transformation does not affect the feature extraction techniques adopted by SIFT and SURF; however, the feature robustness was found to be decreased with an increase in the intersection angle.

Matching points are affected by both intersection angle and distance of the camera to object. The tests proved that by increasing the intersection angles, the matching points were reduced dramatically. Such results could be best described by the exponential relationship. On the other hand, an opposite trend of a linear relationship was found between the distance of the camera to object and the number of matching points. Camera to object distance results revealed that there is a steady decrease in the number of matching points with increased range.

Texture depth measurement showed that the intersection angles have significant influences on the measured depth. The best depth results were only possible with intersection angles between 30-50°. In addition, both smartphone cameras showed poor results even with intersection angle ranged between 30-50°. This could be attributed to their inefficient lenses, small CCD size or due to high distance of the camera to object (at 90cm). These results were found to be improved by fixing the intersection angle around 45° as trials from the distance of the camera to object proved. However, the tests showed that in order to reliably get better results, the camera is preferred to be as close as possible to the surface, refer to figure 3.57. Therefore, in order to use only two images for surface texture shaping, the camera to object distance of 20-35cm and intersection angles of between 30-50° are recommended to be used. Otherwise, an alternative method of using more than two images needs to be investigated, such as using a commercial version of AGISoft (PhotoScan) as discussed in the next section.

### 3.8.2.2 Investigation of depth accuracy from multiple images modelling

The commercial version of AGISoft (PhotoScan) has the capability of producing 3D models from multiple images on a fully automated basis. See section 3.7.2.1.2 to know how data are collected for multiple images. The five cameras (DSLR and smartphone cameras) mentioned in section 3.6.1 namely (Canon 5D-50mm (C50), Canon 5D-28mm (C28), Nikon D200 (D200), Samsung Galaxy Note (Note), and Samsung Wave (Wave) were used again for both rough and smooth surface textures. A total of ten samples were processed for both surfaces and the methodology of sample preparation was applied as explained in section 3.7.2.2, see figure 3.59 for prepared 3D samples.



**Figure 3.59: 3D samples prepared using multiple images smooth texture (above) and rough texture (below)**

The depths of all of these samples were calculated using the author's developed methodology for texture depth measurement explained in section 3.7.2.2.4 and section 3.7.2.3.1. The results of depth calculation showed a significant improvement especially for rough texture results. However, for smooth texture, the biggest error was reported for the Samsung Wave smartphone camera; see table 3.10. It is worth to mention that these results were compared with MTD calculated from the SPT; SPT might not necessarily to be accurate, but the aim of this study was to find a correlation between SPT and depth calculated from 3D models.

Table 3.10: Depth results and error of different cameras ( $\mu\text{m}$ )

Cameras	Smooth Texture	Rough Texture	Error Smooth	Error Rough
<b>C50</b>	607	2063	-62	93
<b>C28</b>	577	2044	-92	74
<b>D200</b>	535	2008	-134	38
<b>Note</b>	512	2005	-157	35
<b>Wave</b>	456	1940	-213	-29



In order to further investigate the degree of similarity between these surfaces with reference to Canon 5D-50mm (C50), the CC cloud-to-cloud distance tools were utilized. Canon 5D-50mm (C50) was selected as a reference for the other cameras after proving its improvement over other cameras during the investigation for angle of intersection and the distance of the camera to object; see section 3.8.2.1. For each camera, a comparison was carried out with reference to Canon 5D-50mm (C50). For each sample and each camera, analysis parameters such as colour visualization, histograms, mean and standard deviation of differences can be worked out. Figure 3.60 shows an example of a comparison of the Nikon D200 (D200) camera with reference to the Canon 5D-50mm (C50) for the smooth texture.

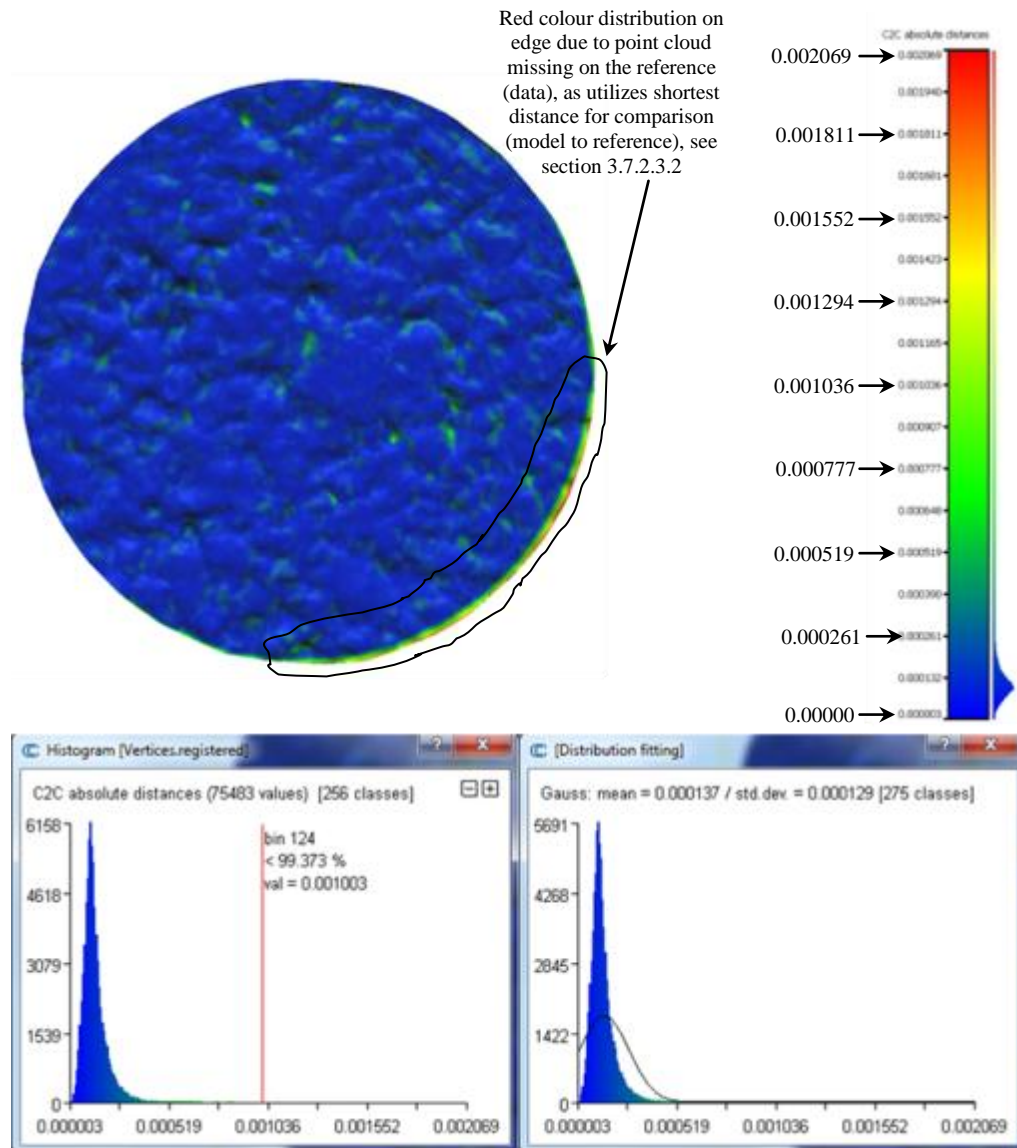


Figure 3.60: Cloud to Cloud distances and their statistical outcome between Nikon D200 and Canon 5D-50mm for the smooth texture (colour visualization on top)

As figure 3.60 shows, the colour visualization of differences represents the distribution of the distance frequencies. The blue colour on the bottom of the colour bar represents differences of 0.003mm, and the red colour on the top of the bar represents differences of about 2 mm. The histogram distribution of the differences on the lower left side of figure 3.60 shows the cut off red line that shows the percentage of included differences of less than 1mm. As can be noticed from figure 3.60, 99.373% of differences are less than 1mm. Through Gaussian curve fitting curve to the differences, the mean and standard deviation are also shown on the lower right of the figure 3.60 and both values are 137 $\mu$ m and 129  $\mu$ m, respectively. For all other samples, comparisons were carried out using cloud-to-cloud distance calculation with reference to the Canon 5D-50mm (C50) and all related information are extracted and presented in table 3.11. The visualizations of differences in the models are shown in figure 3.61.

Table 3.11: cloud to cloud distance calculation results and statistics

	Samples	% < 1mm	Mean ( $\mu$ m)	$\pm$ SD( $\mu$ m)
<b>Smooth</b>	<b>C28 to C50</b>	100	93	38
	<b>D200 to C50</b>	99.373	137	129
	<b>Note to C50</b>	100	160	92
	<b>Wave to C50</b>	99.907	162	102
<b>Rough</b>	<b>C28 to C50</b>	100	119	64
	<b>D200 to C50</b>	99.929	111	81
	<b>Note to C50</b>	99.98	126	82
	<b>Wave to C50</b>	99.586	140	119

As table 3.11 shows, the results of the mean differences for both smooth and rough textures are within the camera to object distance of about 90 to 160  $\mu$ m. These results revealed that the author's methodology of texture depth measurement from multiple images is reliable, and the differences between the reference surface (C50) and other surfaces are acceptable. A visual inspection of the colour coded differences presented in figure 3.61, shows the consistency of the surface matches for both DSLR and smartphone cameras. Smartphone cameras show some of red colour distributions on both texture types which indicate the mismatch on those places. Moreover, the edge differences on some of the samples may affect the visual appearance which is due to missing points in the reference data. Finally, this test proved that it is possible to measure the texture depth using low quality smartphone cameras such as Samsung Wave and Note. The next step is to apply these two techniques (stereo and multiple) image based 3D modelling and depth calculation in a wider range of texture variation which is discussed in the next section.

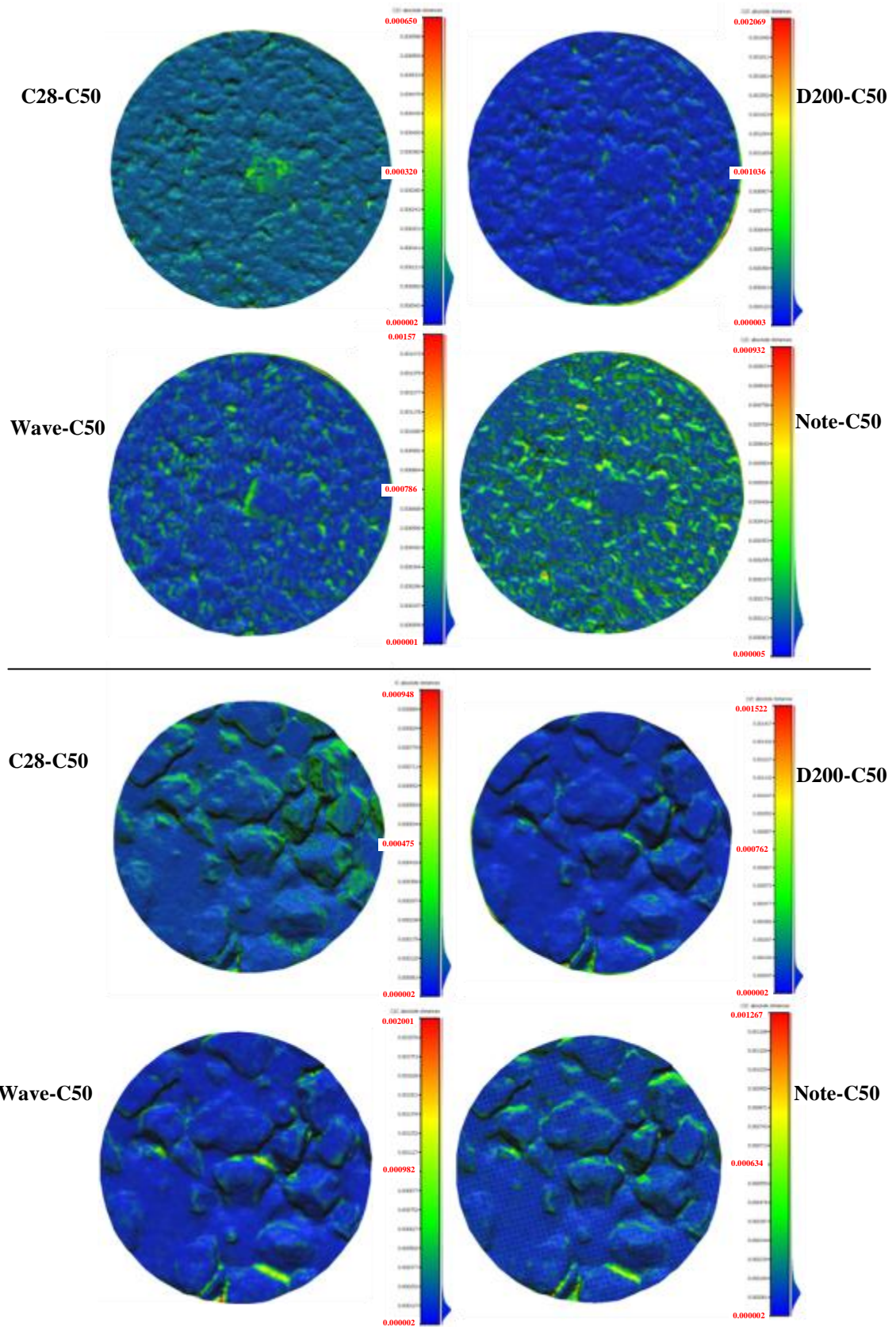


Figure 3.61: Cloud to cloud comparison of all samples with reference to C50, Smooth texture (above) and rough texture (below)

### 3.8.3 Trial 3: Comparison between depth calculated from 3D models and Sand Patch Test

This test is designed to compare the 3D surfaces generated using both AGISoft StereoScan from two images and AGISoft PhotoScan from multiple images with reference to the 3D surface generated using the Canon 5D-50mm (C50). Another aim of this examination is to compare the depth calculated from 3D samples with the MTD calculated from the SPT. The methodology of depth calculation of the 3D samples is described in section 3.7.2.2.4 and section 3.7.2.3.1 in more details. The digital imagery data used for this examination is described in section 3.5 which contains 10 samples collected from different areas around the Jubilee campus, see figure 3.3.

This test aimed at investigating the degree of correlation between the depth calculated from both StereoScan and PhotoScan and the depth calculated from SPT, using five different cameras, see section 3.6.1 for the used camera details. Therefore, it was necessary to include a range of various possible texture depths for the comparison. The texture depths calculated from SPT for each particular test were plotted against each sample, and the second order relation was found with about 98% regression which indicated a good variation on the tested samples, figure 3.62.

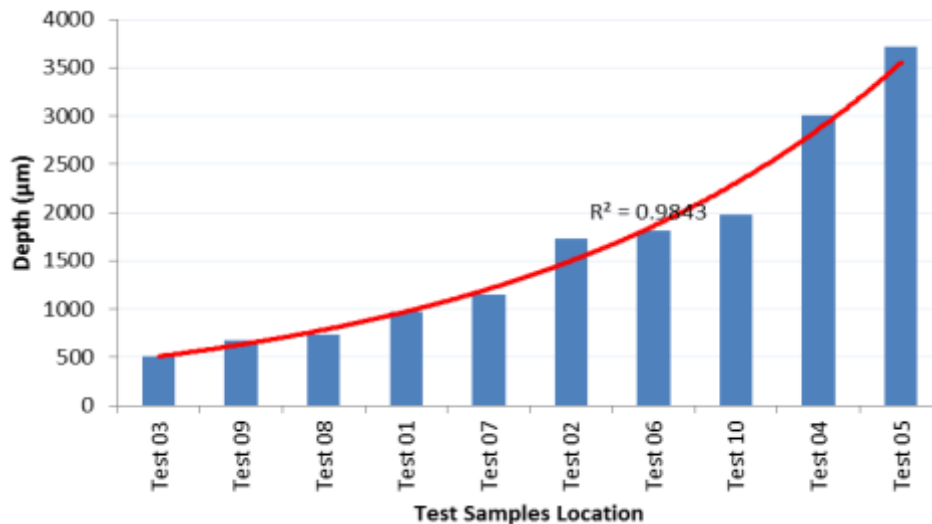


Figure 3.62: Texture depth variation on the used samples

Two methods were investigated using only a pair of images through the use of free tool AGISoft StereoScan and using multiple images through the use of the standard version of PhotoScan which will be explained in more details in the next sections.



### 3.8.3.1 Depth from Sand patch test as Mean Texture Depth (MTD)

Mean Texture Depth (MTD) was calculated using a known volume of sand by dividing it by the area of the patch. The procedures usually followed for calculating the area of the patch is through the averaging of 4 to 5 ruler dimensions of the diameter of the patch. However, two other methods were designed in order to validate further the area of the patch through using CAD tool and Photomodeler, see section 3.6.6. The results of the depth calculated from these three methods for all the tests are listed in table 3.12.

Table 3.12: Depth calculated from three different methods ( $\mu\text{m}$ )

Tests	Ruler Depth	CAD depth	Photomodeler Depth	Average (Depth)	Error (Ruler)	Error (CAD)	Error (Photomodeler)
Test01	978	956	945	960	-18	4	15
Test02	1729	1697	1706	1711	-18	14	5
Test03	511	497	501	503	-8	6	2
Test04	3205	3268	3241	3238	-33	30	3
Test05	3724	3601	3622	3649	-75	48	27
Test06	2160*	1821	1802	1812	348	9	-10
Test07	1052	1075	1066	1064	12	-11	-2
Test08	581	478	493	517	-64	39	24
Test09	673	692	703	689	16	-3	-14
Test10	1974	1954	1963	1964	-10	10	1

The errors of the three methods were calculated from the average depth plotted against the tests and presented on figure 3.63. As can be noticed from figure 3.63, the errors from using CAD and Photomodeler showed better compatibility than the ruler measurement, see section 3.6.6 for more details about CAD, Photomodeler and Ruler area calculation and depth measurement. The differences or errors from the average depth for each method are less than 75  $\mu\text{m}$  except for test06. The differences on test06 showed that the depth from CAD and Photomodeler are compatible. However, the depth from the ruler showed a significant error that is much different from both other depths, and is found to be about 350  $\mu\text{m}$ . It can be concluded that a blunder error from the ruler might be due to taking a wrong reading. This can occur even if the average readings are considered and can seriously affect the depth as in the case of test06 that was excluded from calculations. For this reason, the depth measurement for sample 06 using ruler was excluded for the average depth measurement.

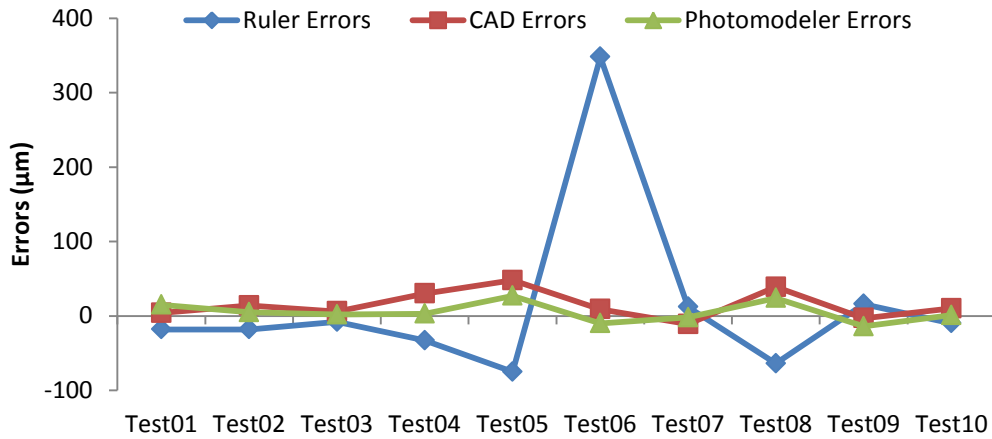
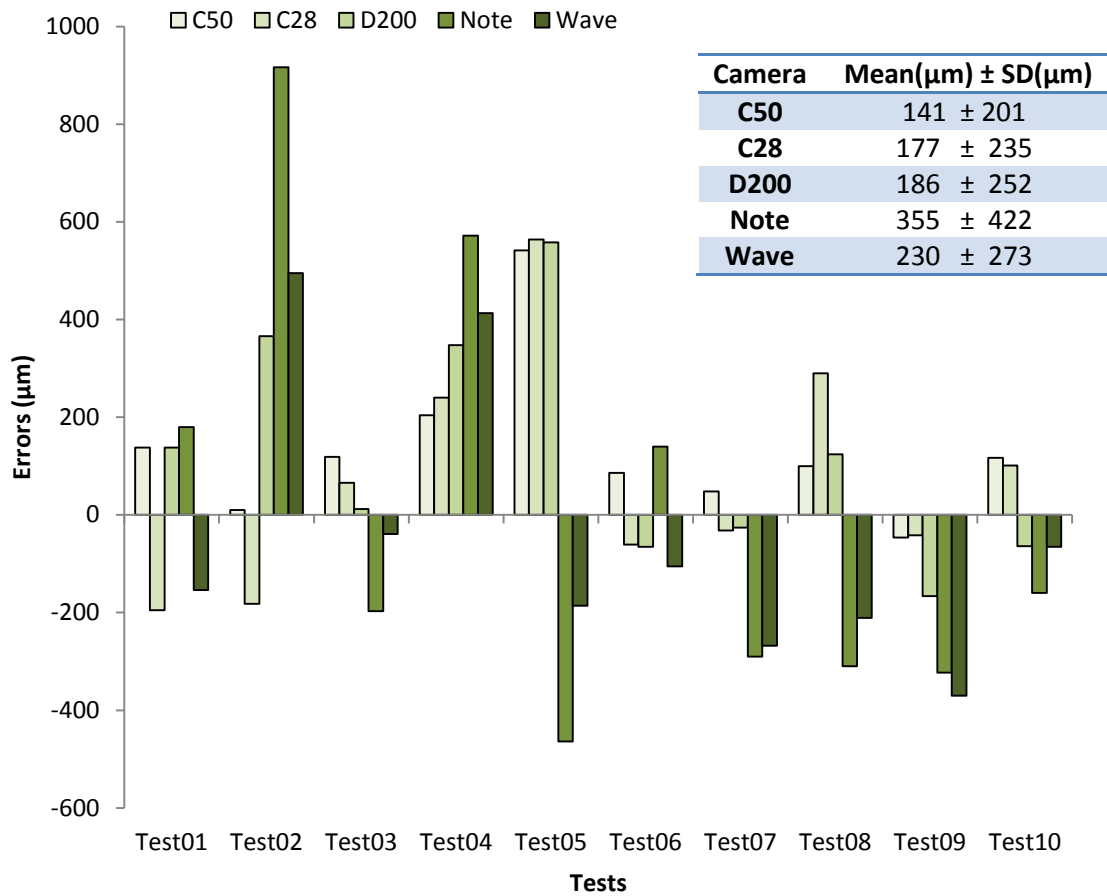


Figure 3.63: Tests vs error from three different method of depth calculation

### 3.8.3.2 Depth from digital images (stereo pairs)

As found from the camera to object distance and intersection angle investigation on section 3.8.2.1, the best degree of intersection is between 30 to 50 ° for both smooth and rough textures. The camera to object distance is as close to the object as the camera can be without being out of focus (20-35 cm). Therefore, for every sample (test01 to test10), two images were collected at different angles of intersection ranging from (30 to 50°) depending on the degree of roughness. The procedures of data collection and 3D model reconstruction are described in section 3.7.2. The differences between depth from digital images for various tests and the average texture depth from three SPT methods (CAD, Photomodeler, and Ruler) were calculated and are presented on figure 3.64. The mean and the standard deviation from differences for each camera were calculated and included on figure 3.64.



**Figure 3.64: Errors between depth from Stereo images and average of sand patch texture depth**

Not surprisingly, as can be seen from figure 3.64 and the mean with standard deviation, the best camera result was obtained for the Canon 5D-50mm (C50). The better result of the Canon 5D-50mm (C50) is due to the fact that it has a longer focal length with full CCD frame which provides more reliable results if compared with other cameras. The next two DSLR cameras, i.e. the Canon 5D-28mm (C28) and Nikon D200 (D200), come after the Canon 5D-50mm (C50) with slightly worth results. However, the results from two smartphone cameras showed that the Samsung Wave (Wave) with about 5.0 megapixels or lower resolution has recorded better accuracy when compared with the Samsung Galaxy Note (Note) that has 8 megapixels. Those adverse results from Note might be attributed to the smaller sensor pixel size that the Samsung Note utilizes, see section 3.6.1 for details of used cameras. The next step was to find a degree of correlation between depths from 3D models and the average depth of the sand patch tests from three different methods for each particular camera. A simple linear correlation is applied as both depths showed a high degree of compatibility as can be seen from figure 3.65.

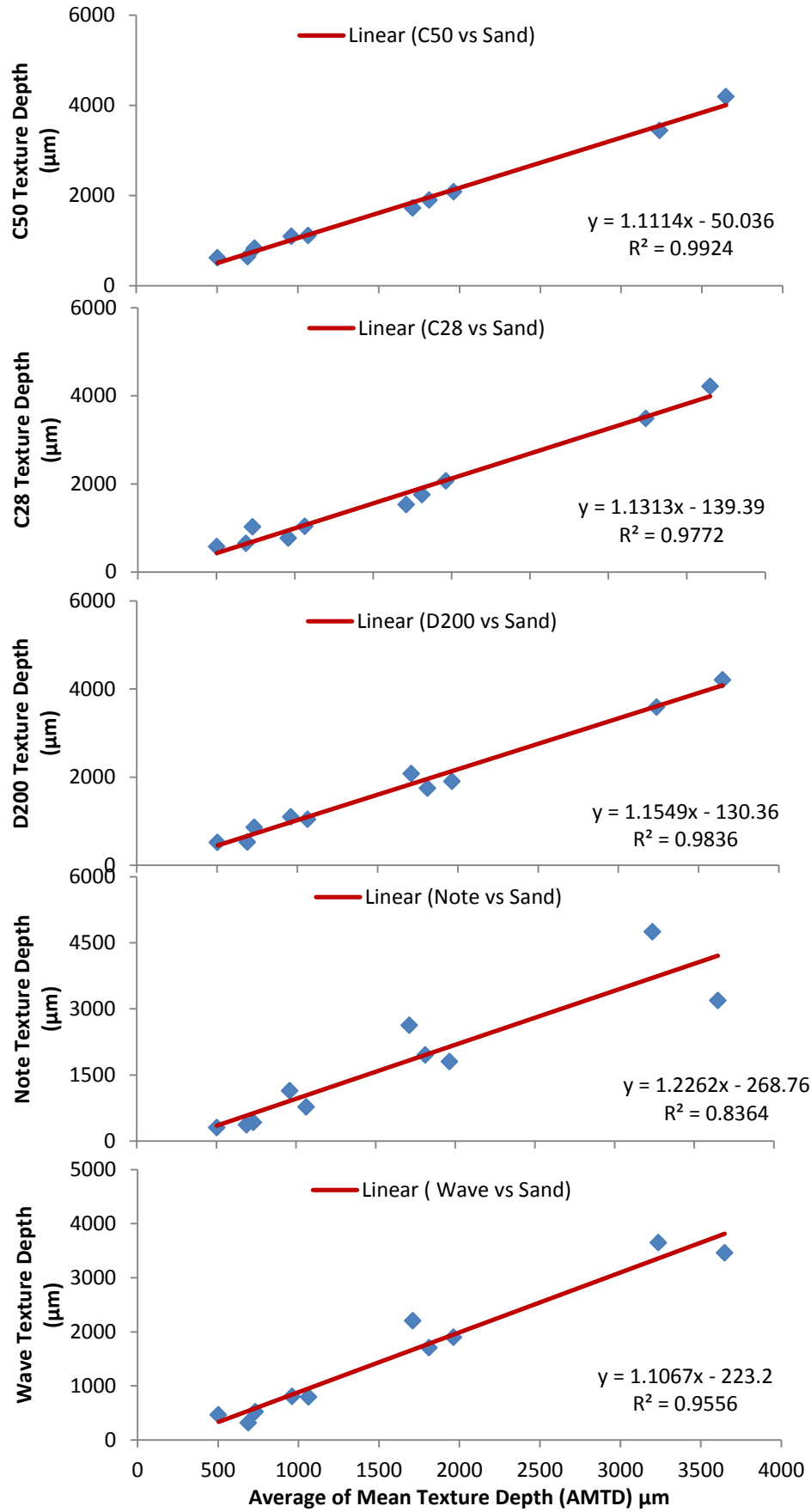


Figure 3.65: Correlation between depth from Stereo and Sand

As figure 3.65 shows, highly significant correlations are found between depth from each of the used cameras and the depth from the sand patch test. The minimum correlation is about 0.83 which is for the Samsung Galaxy Note (Note) smartphone camera. The highest correlation is documented for Canon 5D-50mm (C50) DSLR camera with about 0.99. The slope from the equation for all cameras is near to 1.0 which indicates that both depths are increasing with the same rate. The other value at the end of each of the equation is describing the shift or the average error between the two data sets for the used camera. The shift is found to be about 50.0 ( $\mu\text{m}$ ) in the case of Canon 5D-50mm (C50) and the worst case was found to be about 270.0 ( $\mu\text{m}$ ) for the Samsung Note (Note).

The correlation shown in figure 3.65 revealed that there is a strong relationship between the depths calculated from a pair of stereo images using a free tool AGISoft StereoScan and the depth from the sand patch test that represented the mean texture depth. It is worth to mention that the correlation is made between 10 samples and that might not be enough especially due to the lack of information on some texture depths as can be noticed from the figure 3.62. Clearly in future, it would be better to add more samples to this correlation in order to cover as wide as possible samples, including the depths of 1100 to 1600  $\mu\text{m}$  and 2000 to 3000  $\mu\text{m}$ .

### ***3.8.3.3 Depth from digital images (Multiple Images)***

As previously discussed, the use of multiple images will be considered using the standard version of AGISoft PhotoScan. The process of image collection, 3D modelling, processing and preparing the 10 cm sample for comparison is described in sections 3.7.2.1.2, 3.7.2.2.1, 3.7.2.2.2, 3.7.2.2.3, and 3.7.2.2.4. The depth was calculated from the 3D 10 cm samples for each camera using the author's developed methodology. The depth from each camera and for each test is subtracted from the average depth of the sand patch test listed in table 3.12. The differences (errors) are plotted against each test for each of the camera and presented as figure 3.66. The errors are then averaged, and the standard deviation was calculated for each of the used cameras and included on figure 3.66. As can be seen from figure 3.66 and attached results, all cameras showed better improvement on the average errors and standard deviation when compared with the results from Stereo images that are listed in figure 3.64. The mean error for all cameras was less than 100  $\mu\text{m}$  except for the Samsung Wave, which has mean errors of 114  $\mu\text{m}$ . However, an error of 114  $\mu\text{m}$  is still acceptable in texture depth measurement as no other

techniques can provide this in the field; e.g. it is well known that 95% of SMTD is within  $\pm 0.25\text{mm}$  which is due to limited laser spot size of 0.25 mm diameter (Viner and Britain, 2006; Halloway, 2011).

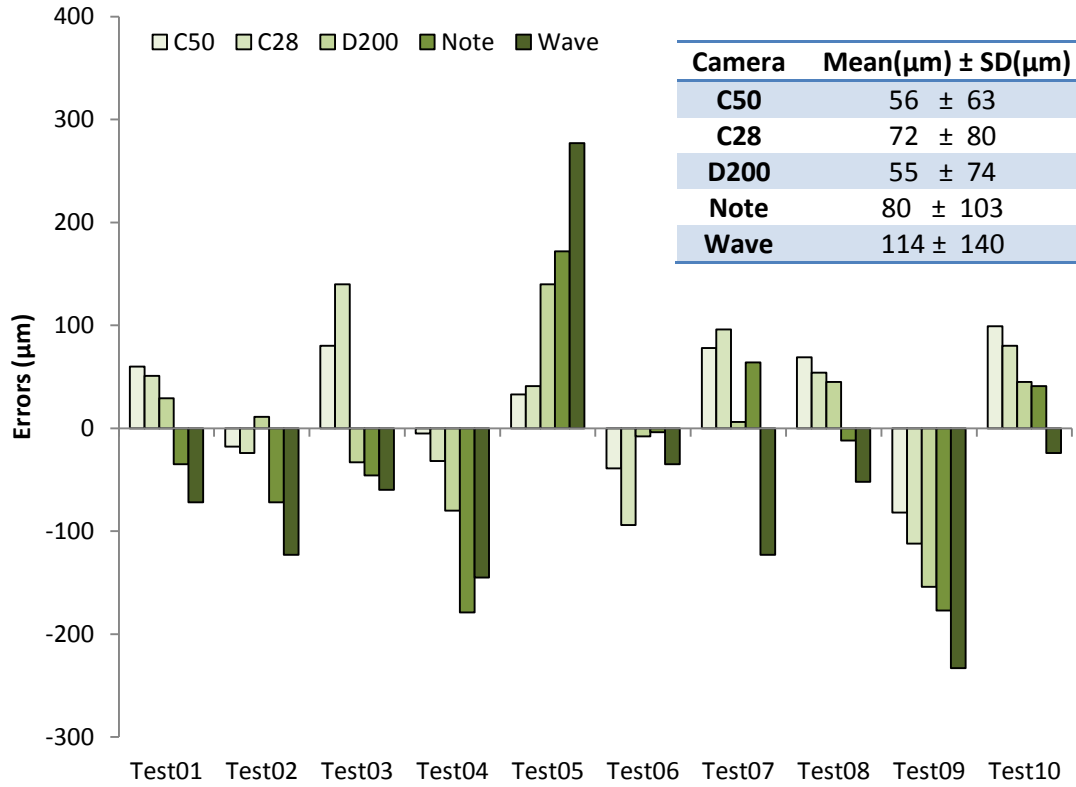


Figure 3.66: Errors between depth from multiple images and sand patch

The next step in the analysis was to find the correlation between mean texture depth from the sand patch test and the depth calculated from 3D models using multiple images. The correlation results from both methods (depth from multiple images and sand patch test) showed that both methods are in a high correlation with a percentage of as high as 0.99 and for all cameras, figure 3.67. Much improvement has occurred for smartphone cameras using multiple images as can be seen from both correlations (figure 3.65 and figure 3.67) which might be attributed to better geometry, full coverage of the texture, and better camera self-calibration parameters, see figure 3.68 in case of multiple images. As can be seen from figure 3.68, the extracted focal length from multiple images is stable over different tests, whereas focal length in the case of stereo fluctuated. The high correlation from multiple images is a good indication that the author's developed methodology of texture depth measurement from multiple images can be used reliably for measuring the surface texture depth of asphalt surfaces.

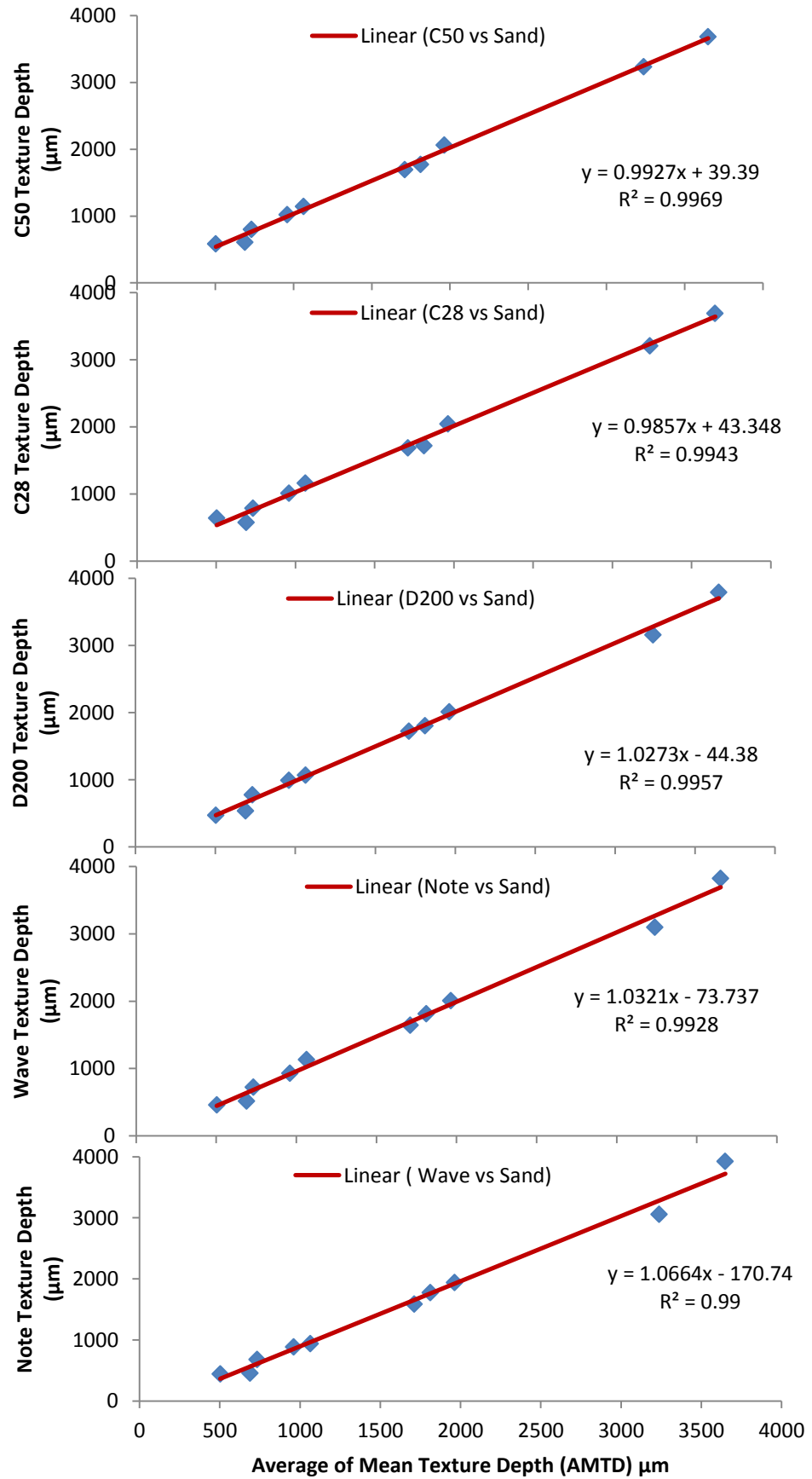


Figure 3.67: Correlation between depth from multiple images and sand

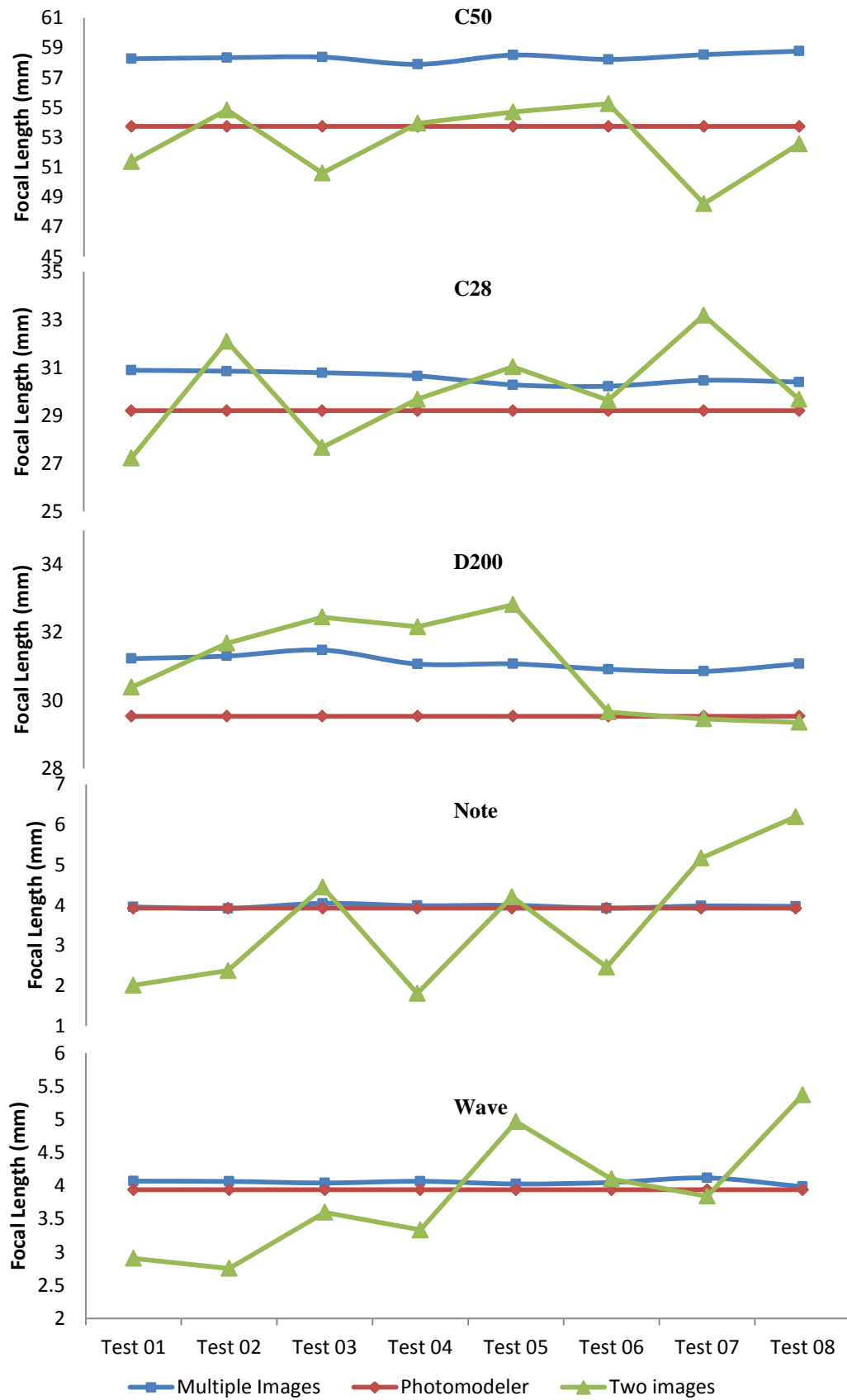


Figure 3.68: comparison of focal lengths between both stereo and multiple images



### 3.8.4 Trial 4: Testing other materials for feature extraction, matching and RMSE

Before testing on concrete, it was considered worth testing the author's developed methodology on various types of materials for feature extraction, matching, and RMSE using stereo images. The aim of this section is to examine some different types of materials (big gravel, small gravel, soil, concrete tiles, grass, and asphalt pavement both smooth and rough), figure 3.5 of section 3.5. The tests were carried out on the Jubilee campus using two different cameras the Nikon D200 (D200) and the Samsung Wave (Wave), mentioned in section 3.6.1. In order to get comparable results from different textures and cameras, the tests were carried out on the same day and with the same time span (only within one hour). In addition, the camera settings remained the same for all tests. Moreover, the intersection angle and the camera to object distance remained constant at  $45^\circ$  and 90 cm, respectively.

The results of AGISoft StereoScan were used to make a comparison between these various materials. The comparison was in terms of the number of extracted features, matches and RMSE. The aim of this analysis is to find the best robust feature among these different textures which can affect the number of matches and reconstructed points and hence the RMSE value. The results of each of the stated parameters are presented in figure 3.69. The reconstruction results are not showing here because it did not show any differences from matching numbers, all correct matches are reconstructed.

As figure 3.69 shows, the feature extraction in all materials was reasonably good. The least feature extracted in all materials was for big gravel texture with about 68 thousand and 22 thousand for both Nikon D200 (D200) and the Samsung Wave (Wave), respectively. The matching relationship shows that the number of matches is not dependent on the feature number as demonstrated in previous sections (3.8.2.1.1-A, 3.8.2.1.2-A). Although grass could provide as many features as concrete tiles, it produced very little matches compared with all other materials. The difference in the matching number between grass and concrete tiles might be due to the fact that grass features are repetitive or similar, but this might not be robust. The matching process might contain lots of outliers which were removed, figure 3.70. The grass can be classified as low textured imagery which is similar to forest textured imagery. Digital image matching often fails in low textured imagery materials such as in forest, desert and mountain areas as reported by Martin (2004).

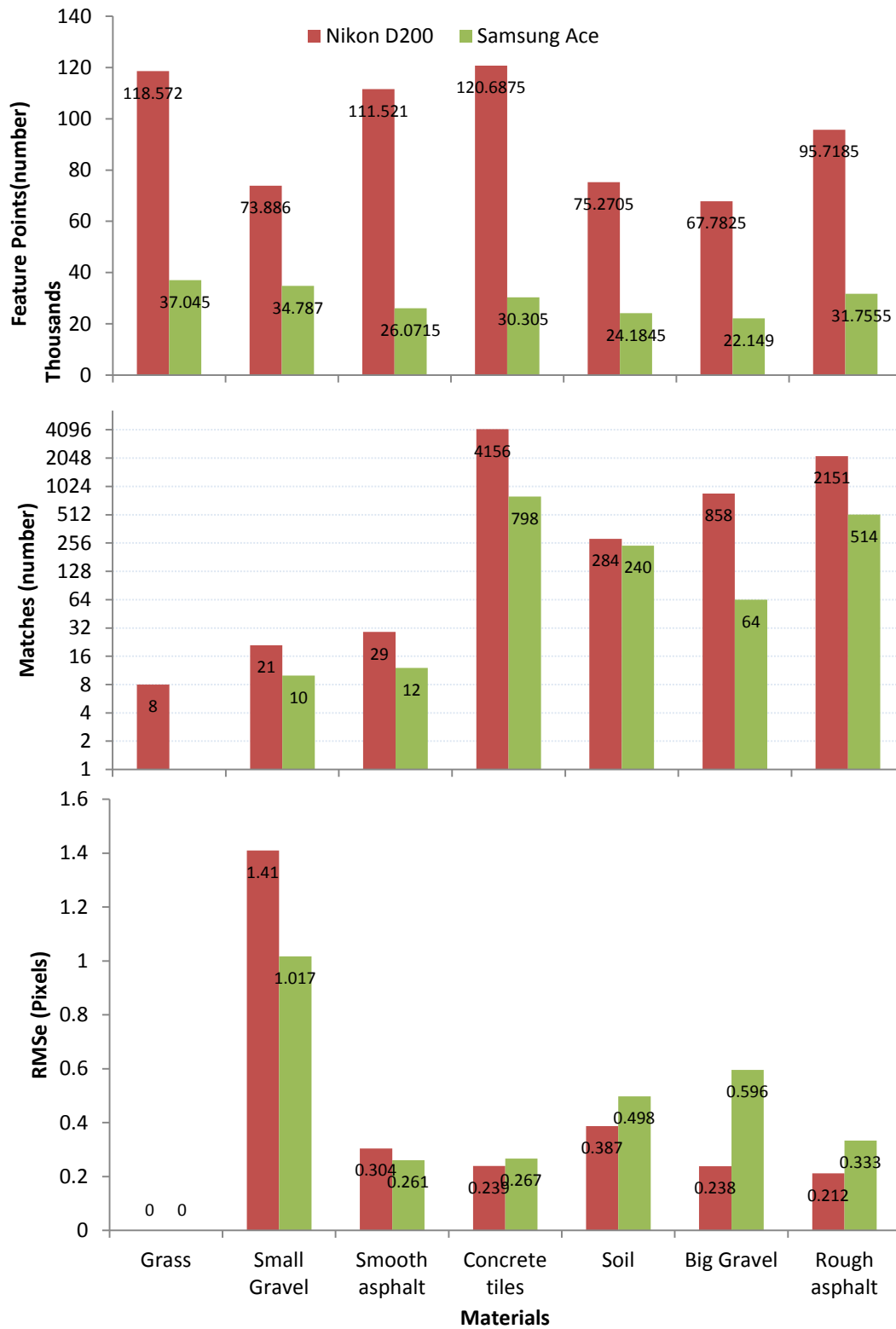


Figure 3.69: Various materials vs reconstruction parameters



**Figure 3.70: Matching on Grass (Wave left and D200 right)**

RMSE values are found to be high for small gravel due to fewer number of matching points. RMSE was found to be depending on the number of matches and found to be increased as the matches decreased, see section 3.8.2.1.1-C. However, points of matching on smooth asphalt are also as few as small gravel but little RMSE could be documented. The differences in RMSE between smooth asphalt and small gravel might only be due to the fact that the matching points in case of smooth asphalt are more accurate than small gravel. It can be concluded from this investigation that probably both the concrete tiles and rough asphalt textures provided robust features that led to production of reasonable matching points with small RMSE values which substantially helped in successful 3D modelling. The soil texture came as the third material in term of robust features and successful matching and 3D modelling on an automated basis. The smooth asphalt texture was found to rank fourth and came after soil texture as it was not possible to build a 3D model in the case of a low resolution camera such as Wave, see figure 3.71.



**Figure 3.71: 3D reconstructed surface for various materials Wave (above) and D200 (below)**

### 3.8.5 Trial 5: Investigating the method on concrete texture

It was shown in previous sections that it is possible to use digital images for the measurement of roughness or texture depth on asphalt surfaces. The author's developed methodology was successfully used for measuring texture depth with a highly significant correlation to sand patch test results. In this section, the aim will be to investigate the author's proposed methodology of texture depth measurement on concrete texture. This is of particular interest when adhesions or screeds are to be added to the concrete face or surface. The depth from using multiple images will be utilized on two control surfaces of concrete, one considered to be smooth and the other to be rough, see section 3.5 and figure 3.6 for more details about both surfaces. The five cameras (three DSLR and two smartphone cameras) were used when testing the concrete surface texture; see section 3.6.1 for details about used cameras.

Ten samples of the circular 10cm 3D model prepared using the methodology described in sections 3.7.2.1.2, 3.7.2.2.1, 3.7.2.2.2, 3.7.2.2.3, and 3.7.2.2.4; figure shows samples of prepared 3D models. The height ramp colour visualization on the models showed no real differences between the samples, see figure 3.73. This visualization is a good indication of the correct geometry of each sample and can also provide the degree of similarity of all cameras. Visual inspection of the colour coded ramps show no real differences between the 3D models. As can be seen from figure 3.73, the details from each camera appear exactly the same if compared with each other. For example, the green and red colour coded zones localized on the centre and edge of the rough texture are repeated for all models without any changes.

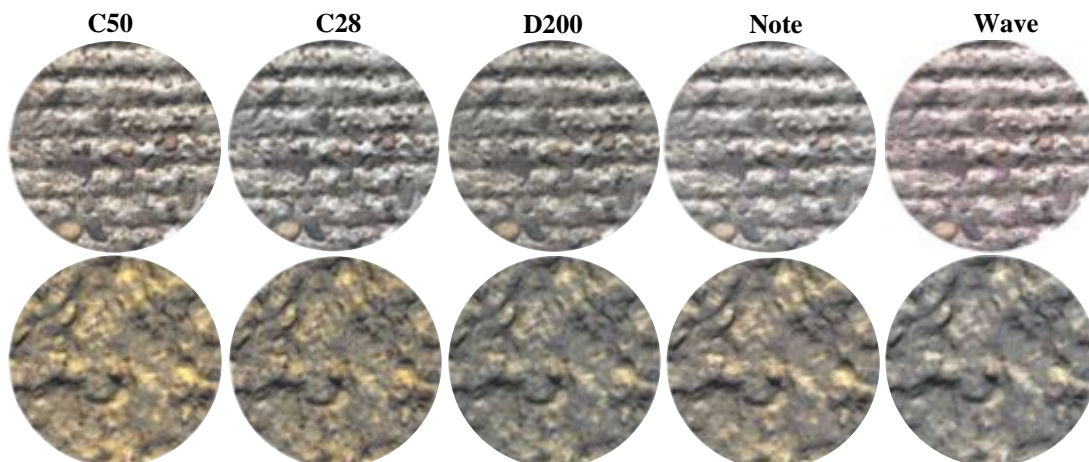
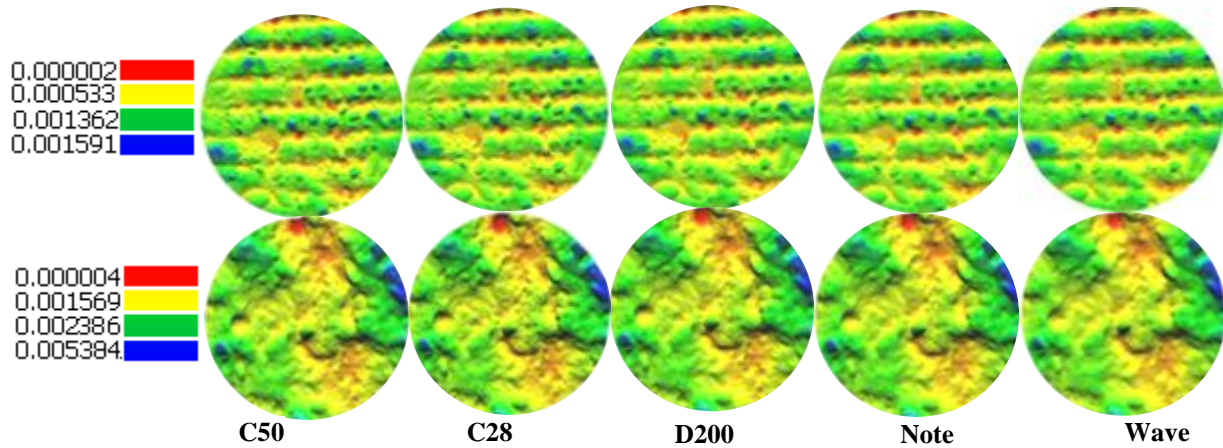


Figure 3.72: 3D samples from 5 cameras, smooth texture (above) and rough texture (below)



**Figure 3.73: Colour visualisation of differences on 3D models, smooth (above) and rough (below)**

Carrying out further analysis is useful in order to test for the degree of similarity between each sample with reference to the Canon 5D-50mm (C50), using the comparison of cloud-to-cloud distances in CC tool. The colour visualisation of the differences on the models were made for all samples with reference to the Canon 5D-50mm (C50) and presented as figure 3.74. As can be seen, the visual differences of the comparison between each model and the Canon 5D-50mm (C50) as reference show that all the samples are compatible with the Canon 5D-50mm (C50) for both smooth and rough texture. The blue and green colour zones visualized on the models are showing differences that are on average less than 100  $\mu\text{m}$ . The upper limits of the differences for most cases are less than 1mm except for Samsung Wave in the rough texture which is about 1.84 mm.

Histograms of these visual differences were also prepared with the mean and standard deviation extracted from each comparison and listed in table 3.13. As this table reveals, with the cut off limit set to 1mm almost all samples have more than 99 % of differences falling within 1mm. Mean and standard deviation of both smooth and rough textures showed that the results are within 140-180  $\mu\text{m}$  and 65-90 $\mu\text{m}$  for the mean and standard deviation respectively. The worst result is documented for the Samsung Wave and for the rough texture which is 234  $\mu\text{m}$  and 147  $\mu\text{m}$  for mean and standard deviation, respectively. These results showed that the author's developed methodology is working for concrete texture, and not much different from the asphalt pavement texture results listed in table 3.11. Another approach of analysing the similarity of the models will be considered through drawing a profile of differences and calculating mean and standard deviation.



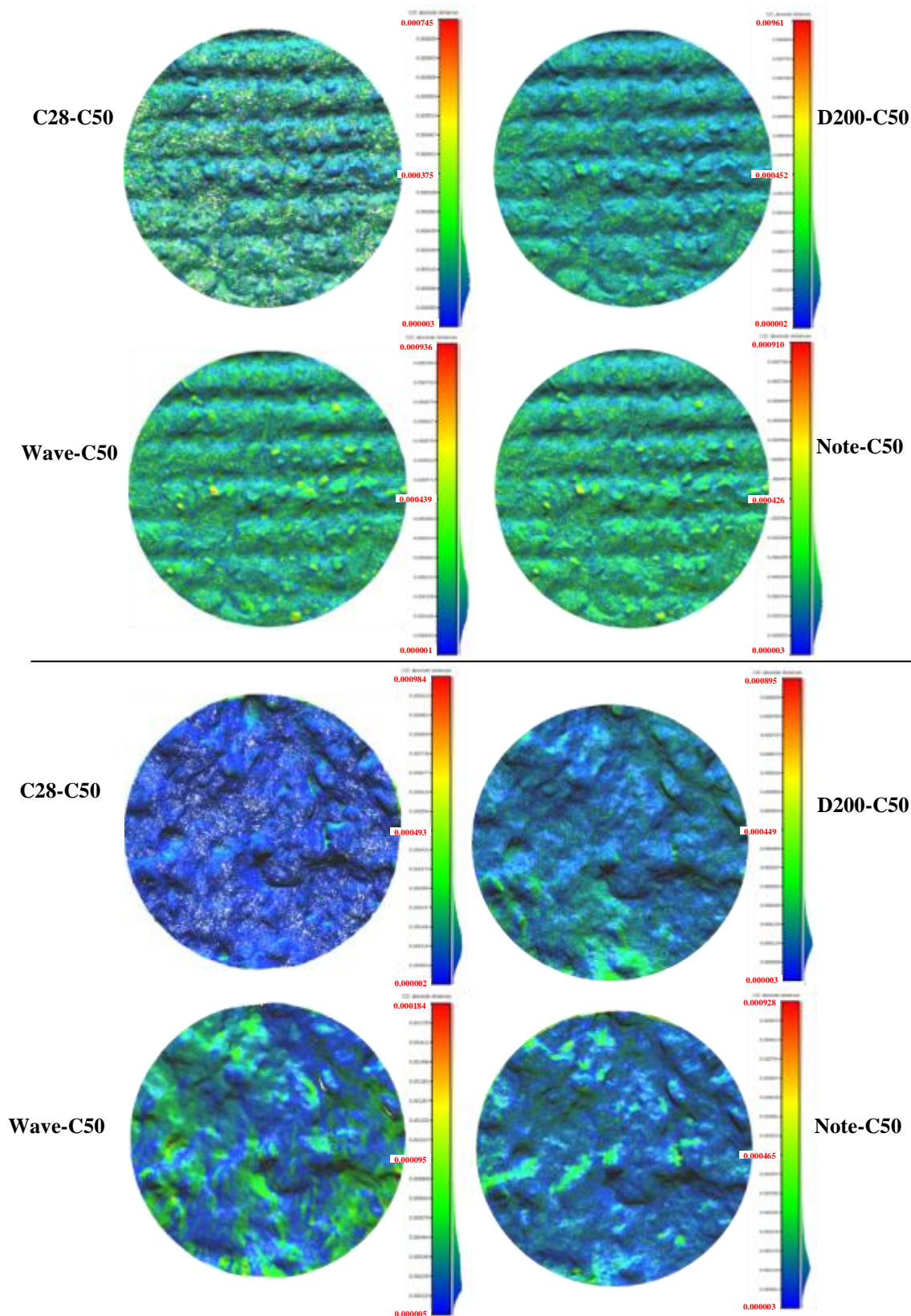


Figure 3.74: Colour visualization of differences on concrete models, smooth (above) and rough (below)

Table 3.13: cloud-to-cloud distance comparison results

	Samples	% < 1mm	Mean	SD
Smooth	C28 to C50	100	140	66
	D200 to C50	100	158	76
	Note to C50	100	176	79
	Wave to C50	100	179	82
Rough	C28 to C50	99.855	150	91
	D200 to C50	99.998	139	64
	Note to C50	99.968	161	81
	Wave to C50	99.541	234	147

As before, the profiles were drawn horizontally in the centre of both smooth and rough texture samples, figure 3.75. Then, differences of each sample profile with reference to the Canon 5D-50mm (C50) profile were reported at every 1mm of range, see figure 3.76. The RMSE of each sample is calculated from the differences and included in figure 3.76.

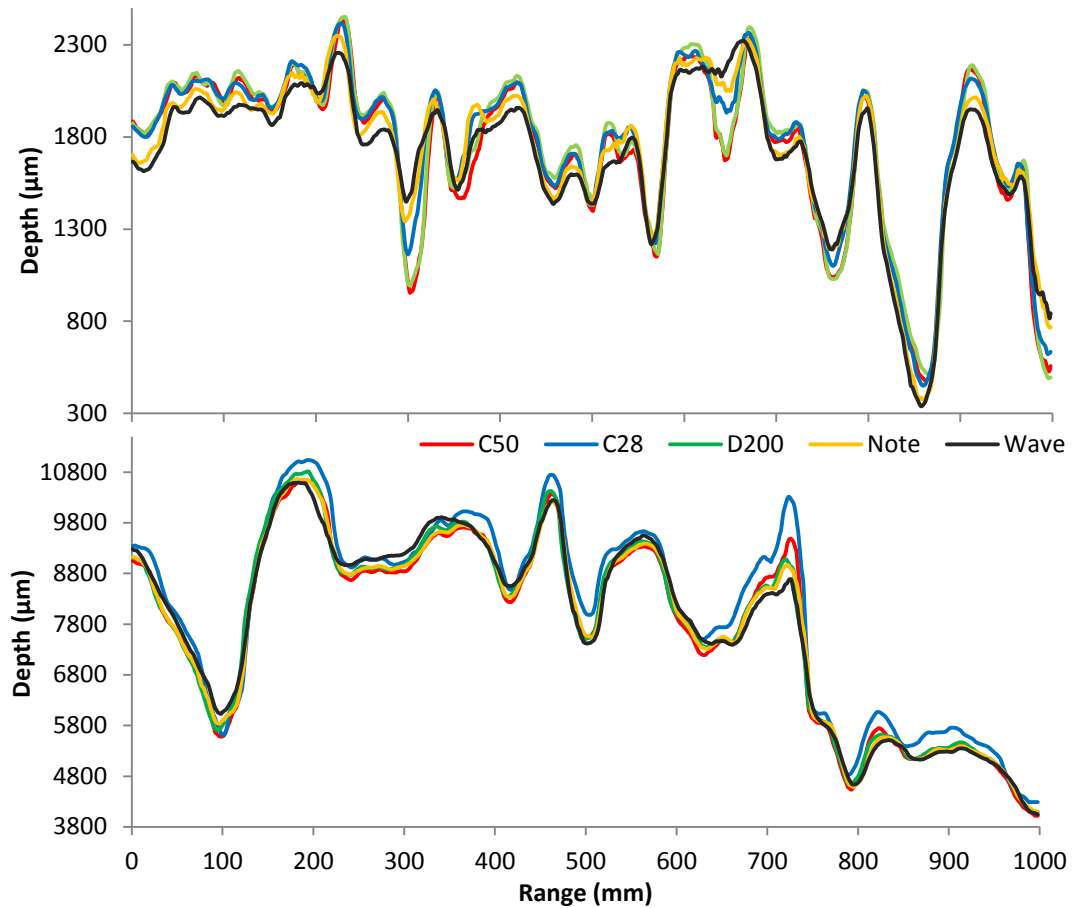
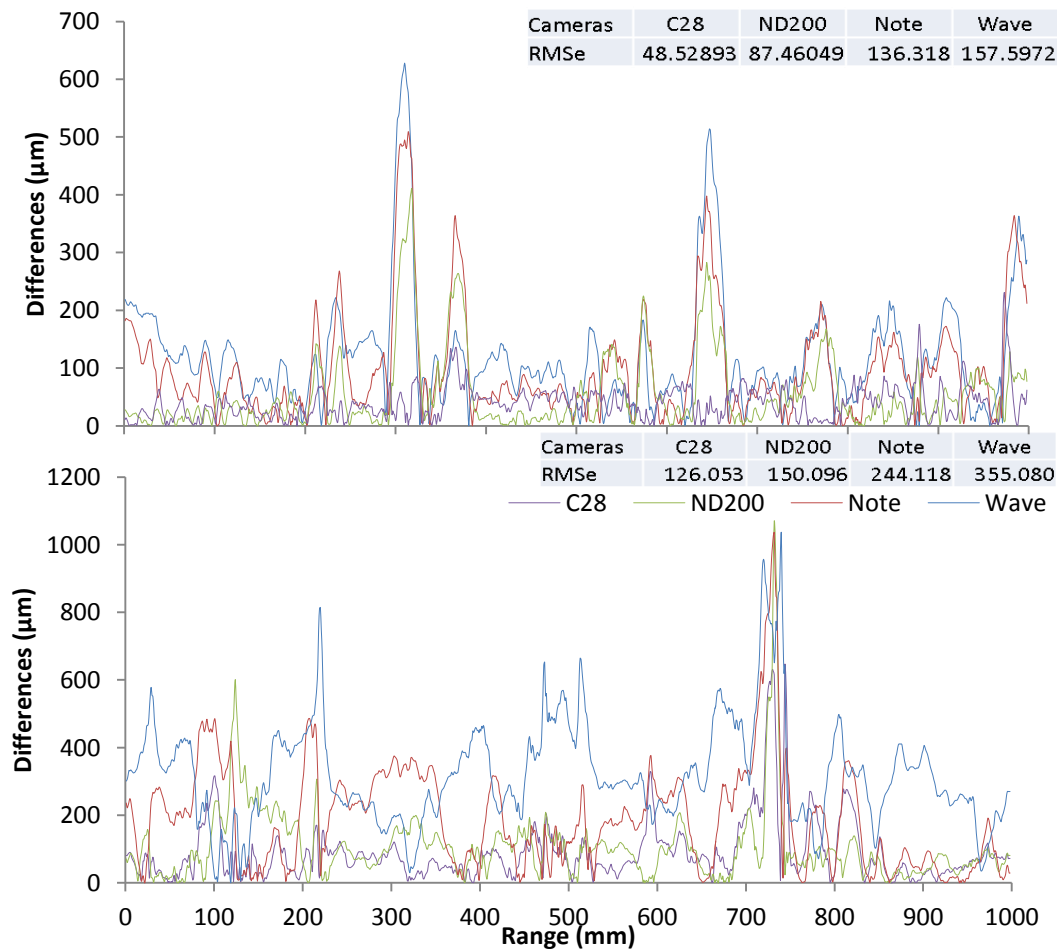


Figure 3.75: Horizontal profiles through centre of 10cm models, smooth texture (above) and rough (below)



**Figure 3.76: Differences between all samples with reference to C50, smooth texture (above) and rough (below), RMSE ( $\mu\text{m}$ )**

As figure 3.76 and the RMSE values show, the differences at some intervals are too high. For example, in the case of rough texture at 740mm range, the differences reached about 1mm. These differences might be due to inadequate quality registration of the surfaces with GCPs collected using reflectorless total station. However, RMSE of registration of four GCPs from Cloud Compare results showed that the highest value was 691  $\mu\text{m}$ , recorded for the Samsung Galaxy Wave (Wave) in the rough texture, table 3.14. The results of RMSE listed in this table conclude that the errors are not only from registration but also from other sources, such as a mismatch of matching points, lens quality, focal length, and camera resolution as it is apparent from the RMSE results in figure 3.76, in the case of rough texture. Occlusion might also be another source of errors as the differences around ranges of 305mm and 650mm record the highest values in the smooth texture, figure 3.76.



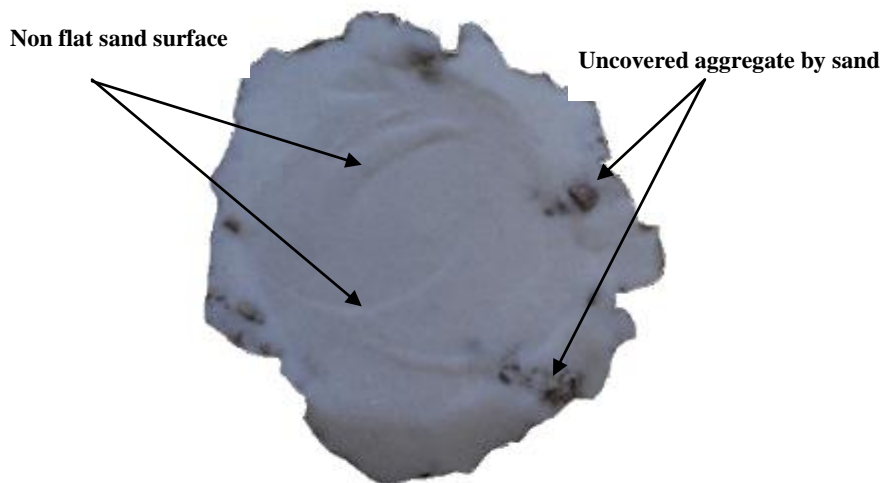
Table 3.14: RMSE of 4GCPs from transformation between local camera and total station coordinates

	Samples	RMSE( $\mu\text{m}$ ) from 4 GCPs
<b>Smooth</b>	<b>C50</b>	159
	<b>C28</b>	164
	<b>D200</b>	222
	<b>Note</b>	291
	<b>Wave</b>	401
<b>Rough</b>	<b>C50</b>	295
	<b>C28</b>	298
	<b>D200</b>	281
	<b>Note</b>	410
	<b>Wave</b>	691

The texture depth from 10 cm circular samples was worked out, using the author's developed methodology described in section 3.7.2.2.4. The differences between the sand patch test mean depth from three methods, and the depth from 3D models are apparent, especially in the case of rough texture. These differences are because errors from the sand patch test increase as the roughness of texture increases, due to the difficulty of spreading out the sand in a perfect circular shape and due to the slope that that sand makes at edges. In addition, some aggregate might not be covered with the sand, as can be seen from figure 3.77, where the sand surface is not flat, and some parts of the aggregate are uncovered by sand; in addition, the patch was not in a circular shape which can affect the sand depth results obtained using the other methods, particularly the ruler method. The recreated 3D sand patch test surface showed that the surface has a depth of 810  $\mu\text{m}$  in the rough texture and 326  $\mu\text{m}$  in the smooth texture. The reason for making the comparison with SPT is that is an industry standard, although the results from SPT might not necessarily be better than the photogrammetric solutions.

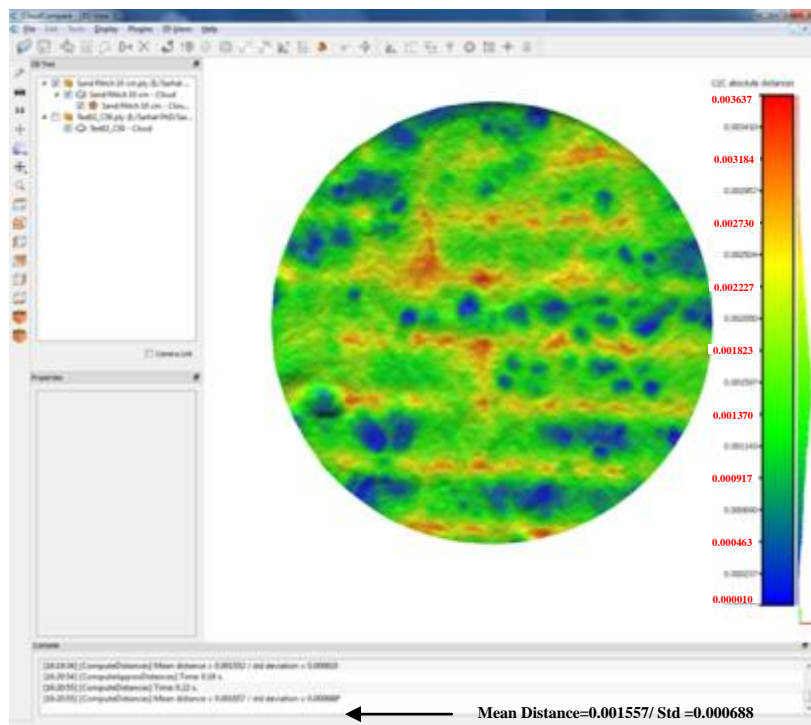
Table 3.15: Depth from 3D samples vs depth from sand patch for concrete samples ( $\mu\text{m}$ )

	Samples	Depth from 3D sample	Depth from Sand Patch		
			Ruler	AutoCAD	Photomodeler
<b>Smooth</b>	<b>C50</b>	1505	1683	1570	1585
	<b>C28</b>	1502			
	<b>D200</b>	1492			
	<b>Note</b>	1478			
	<b>Wave</b>	1444			
<b>Rough</b>	<b>C50</b>	4155	5221	5211	5181
	<b>C28</b>	4055			
	<b>D200</b>	4105			
	<b>Note</b>	4125			
	<b>Wave</b>	4006			



**Figure 3.77: Sand Patch recreated surface (rough concrete texture)**

Differences between the depths from 3D models and depth from the sand patch test are apparent. Therefore, it was decided to use cloud-to-cloud distance calculation between the 3D samples and the sand patch test recreated surface as explained in section 3.7.2.2.5. Figure is an example of the cloud-to-cloud distances between the Canon 5D-50mm (C50) and sand reconstructed surfaces for the smooth texture. The recreated sand surface is subtracted from the real texture samples and the mean distances with standard deviation are reported using cloud compare software, the results from all samples are listed in table 3.16.



**Figure 3.78: Sample of cloud-to-cloud distance calculation (C50 with reference to sand smooth texture concrete)**

Table 3.16: cloud-to-cloud distance calculation between sand recreated surfaces and 3D samples

	Samples	Mean distances( $\mu\text{m}$ )	Std. deviation( $\mu\text{m}$ )
<b>Smooth</b>	<b>C50</b>	1557	688
	<b>C28</b>	1513	612
	<b>D200</b>	1563	608
	<b>Note</b>	1569	600
	<b>Wave</b>	1543	584
<b>Rough</b>	<b>C50</b>	4283	2029
	<b>C28</b>	4151	2028
	<b>D200</b>	4185	2028
	<b>Note</b>	4029	2042
	<b>Wave</b>	3999	2029

The results in table 3.16 showed that the mean distances calculated between the 10 cm 3D samples and the recreated 3D surface of the sand is closely related to the depth calculation results. The differences from all cameras are within the range of about 6-289  $\mu\text{m}$  for both smooth and rough surfaces. These differences reveal the fact that the 3D models are accurate and reliable using multiple images. Although the depth calculated from 3D samples using the author's developed methodology described in section 3.7.2.2.4 showed significant discrepancy with the sand patch test as illustrated in table 3.15, the variation is not due the miscalculation of the depth using the methodology explained in section 3.7.2.2.4, but it is probably due to non-flat surface of the sand as the results in table 3.16 proved.

Multiple images based 3D surface modelling seems be work very good when the camera to object distance is close to the surface. In all of the above cases, the camera was close to the object, and the distance ranged between 20-35cm. Farther ranges were considered in this test in order to examine the effect of the camera to object distance on the texture depth. This was achieved by using only the smooth concrete texture for investigation. For this purpose, two cameras Nikon D200 and Wave were utilized with two ranges and used to capture photos at 60- 70cm and 90-110cm. In addition to the close distance of camera to object of 20-35cm which was already tested, four other samples were prepared making a total of six. The methodology of preparing circular 10cm samples is explained in sections 3.7.2.1.2, 3.7.2.2.1, 3.7.2.2.2, 3.7.2.2.3, and 3.7.2.2.4. The depth calculated from the 10 cm circular samples showed that the camera to object distance has the same effect as previously proved in section 3.8.2.1.2 for the stereo process. The increase in the camera to object distance decreases the depth of the

texture as visual inspection at figure 3.79 shows, with details starting to disappear as the camera moves away from the object and subsequently the texture depth get smoother. The stated depths are those calculated using the author's developed methodology from multiple images illustrated in previous section (3.7.2).

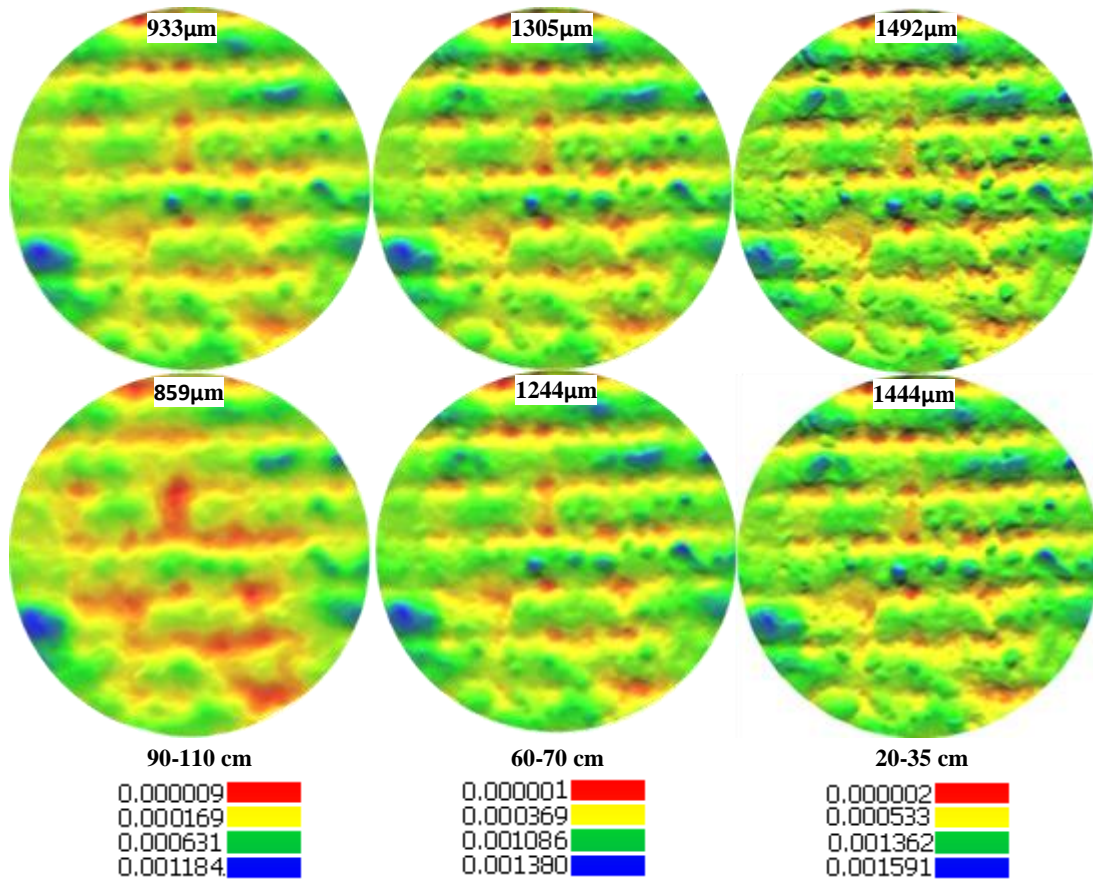
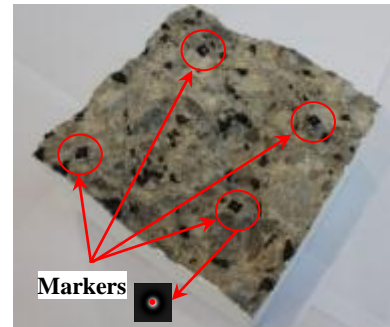


Figure 3.79: Visual height ramp of the smooth texture with the camera to object distance, Nikon D200 camera (above) and Samsung Wave (below)

### 3.8.6 Trial 6: Using the proposed method of depth calculation as a dependable measuring technique for fractured concrete

The author's proposed methodology of texture depth measurement and characterizing was also used to study the 3D surface modelling of fractured concrete samples. This was to quantify the fractured surfaces so they could be described in terms of fracture phenomenon of cementitious materials. One of the important parameters in the fractal analysis concept is the fractal dimension which can be calculated by cube counting method based in 3D analysis. The fractured surface should be first geometrically characterized by constructing a 3D surface prior to subsequent analysis. For more information about the analysis and characterization techniques of impact-fractured concrete surfaces refers to (Ahmed, 2012; Erdem et al., 2012; Erdem and Blankson, 2013; Santos and Júlio, 2013; Werner et al., 2013).

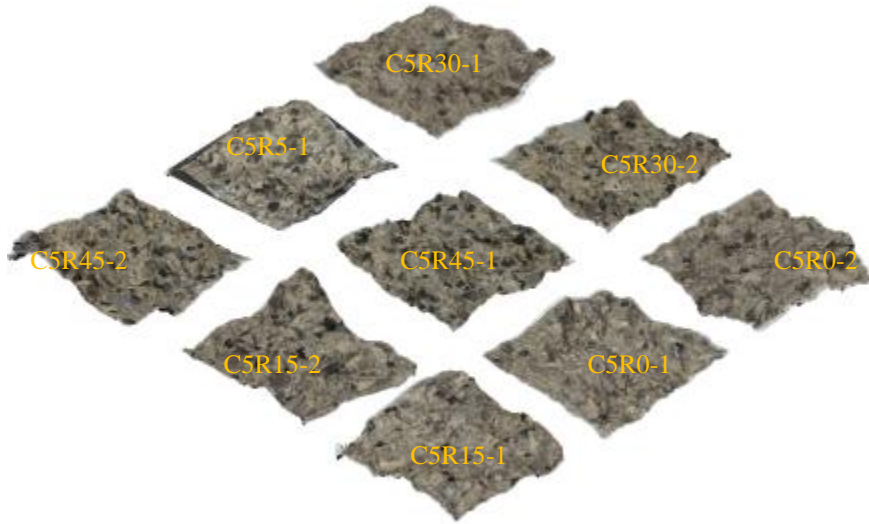
To minimize the processing time and to help the extraction of the fractured surface, all sides from each sample were painted with white paint and placed on a white background and captured using Canon 5D-28mm (C28). Using the white paint on the sample sides and background will help the trimming process as it will be easier to extract the fractured surface from the white background. As it was previously discussed, nine concrete fractured samples were subjected to processing in order to reconstruct the 3D surfaces based on the digital image and photogrammetry. The procedures of reconstructing the 3D surfaces from the digital image based on photogrammetry were described in sections 3.7.2.1.2, and 3.7.2.2. Four markers with the size of (0.5x0.5 cm) were placed on each of the used samples and surveyed using a reflectorless total station (refer to section 3.6.3, for accuracy of total station) in order to be used as GCPs for absolute orientation process; figure 3.80.



**Figure 3.80: Fractured concrete sample with markers**

The nine samples were reconstructed, see figure 3.81 and transformed to absolute orientation utilizing the aligning tool in CC which was described in detail in section 3.7.2.2.2. Cloud compare provides details of the transformation process such as transformation matrix, RMSE of the transformation and the recovered scale. The RMSE values of transforming each sample from the image coordinate system to the local

coordinate system were recorded. The RMSE which can be treated as GCPs accuracy was worked out using four GCPs for each sample. Nine observations are usually utilized for 3D rigid transformation and in this processing, 12 observations were used in each sample. RMSE values for each sample are listed in table 3.17. For more information about the recovered scale and transformation matrix refer to appendix A.



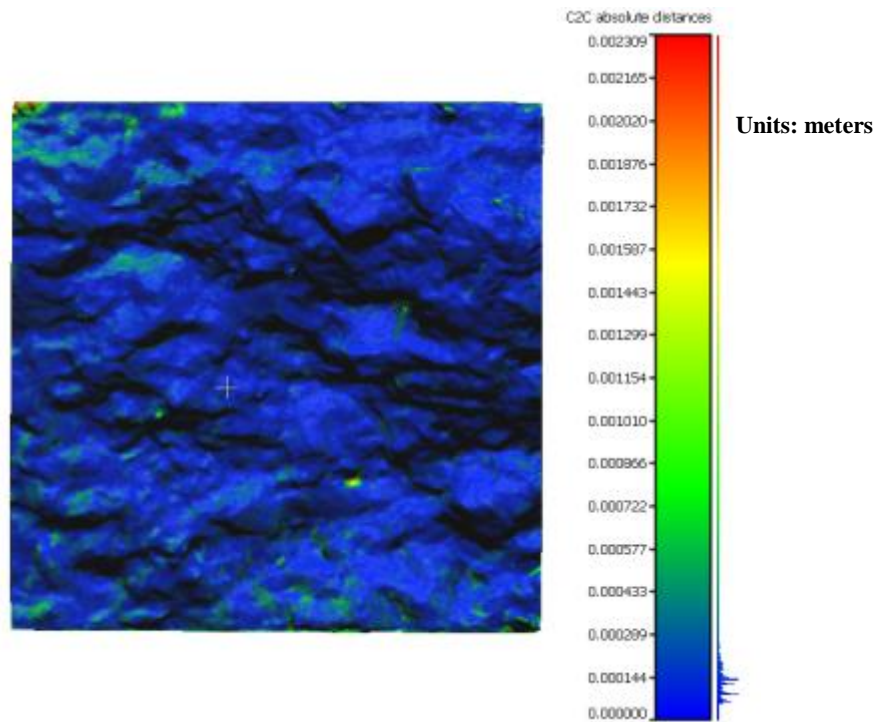
**Figure 3.81: The reconstructed concrete fractured samples**

Table 3.17: RMSE of GCPs out from transformation process in CC

Sample IDs(C5--)	R0-1	R0-2	R5-1	R15-1	R15-2	R30-1	R30-2	R45-1	R45-2
RMSE (mm)	0.54	0.20	0.61	0.68	0.44	0.84	0.82	0.88	0.73

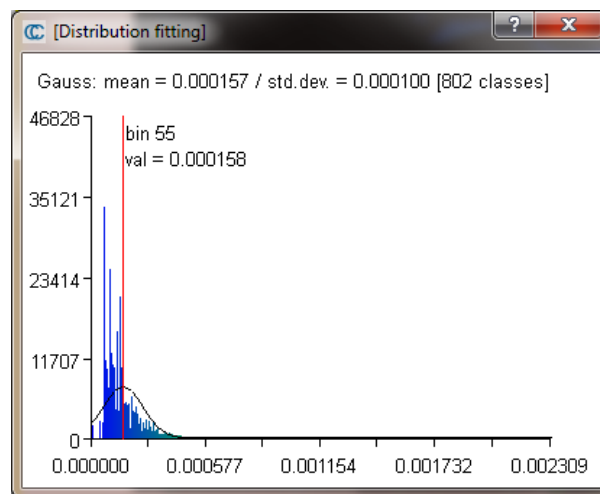
As table 3.17 shows, the RMSE values for all samples fall within less than 1.0 mm; the biggest RMSE was recorded for C5R45-1 as 0.88 mm. These RMSE values pointed to the presence of better consistency between the local coordinate system and the recreated surfaces represented in GCPs. To further analyse the accuracy or the precision of the reconstructed surfaces, the C5R5-1 sample was subjected to second recreation using the same camera C50 but with a different set of images. Two sets of images were collected for the C5R5-1 sample, each set consist of 17 images. However, the distribution and angle of acquisition of images was different from one set to the other. The reconstruction processes of 3D surfaces followed the procedures mentioned in the previous examples, and detailed in sections 3.7.2.1.2, 3.7.2.2.1, and 3.7.2.2.2. The cloud-to-cloud distance calculation was then carried out on these two sets and the results of visualizing the differences on the model with colour coded bar scale are presented in figure 3.82.





**Figure 3.82: visualization of differences of the two sets**

As figure 3.82 shows, the overall differences between the two sets are in the range of less than 200  $\mu\text{m}$ . In order to precisely work out the mean of the difference, the histogram of the differences is drawn with Gaussian fit, see figure 3.83. As can be seen from this figure, the mean and standard deviation differences are  $157 \pm 100 \mu\text{m}$  concluding that the author's proposed methodology can be used for the geometric characterization of a fractured concrete sample.



**Figure 3.83: Histograms of differences between two sets with Gaussian distribution fit results**

### 3.9 Summary and Conclusions

The aim of this chapter was to examine the use of digital images for texture depth measurement of various materials, particularly, asphalt and concrete textures. A number of experiments were carried out in order to investigate the use of digital images for texture depth measurement. A mixture of DSLR cameras and smartphone cameras were investigated and used in all tests conducted in this chapter.

Testing DSLR cameras for lens quality was an objective of this chapter and many of the available cameras in NGB along with their lenses were tested. The results of this study showed that the image sharpness is mainly due to the lenses and it has been proved that some old lenses produced inadequate image sharpness, which could be due to excessive use or due to any source of accident that may degrade the quality of the lens, as in the case of Nikon D300 which showed unsatisfactory results when compared with other cameras. Based on this test, two DSLR cameras (Canon 5D, Nikon D200) with three lenses (Canon 50mm, Canon 28mm, and Nikon 28mm) were selected to be used in addition to the two smartphone cameras, for the trials in this chapter.

Another objective of this chapter was to qualify the effect of the intersection angle and distance of the camera to object on many reconstruction parameters such as feature extraction, matching of correspondence points, RMSE of the least square solution, and the accuracy of the generated 3D surfaces. The aim was to find the optimum intersection angle and distance of the camera to object when only two images used, through utilizing the free tool AGISoft StereoScan. Two controls of asphalt texture, rough and smooth were employed for performing these trials. The findings from this examination revealed that the best intersection angle is between 30 to 50°. Testing the distance of the camera to object also showed that the depth of the texture minimized as the camera distance from the surface increased. The recommended distance of the camera to object was found to be as close as the camera can be while not being out of focus, that was found to be between 20 and 35cm.

Results from AGISoft StereoScan with the use of two images showed some discrepancies. The reliability of the two images was proved to be low. This was found to be true especially for both smartphone cameras, i.e. Samsung Galaxy Note and Wave that reported a standard deviation of about 1250  $\mu\text{m}$  and 1400 $\mu\text{m}$ , respectively when different intersection angles were used at a 90cm camera to object distance, while



standard deviations of about 740 $\mu$ m and 785 $\mu$ m, respectively were found when different distances of the camera to object were used with an intersection angle of 45°. The results of DSLR cameras were recorded better accuracy with 225 $\mu$ m, 426 $\mu$ m, and 656 $\mu$ m for the Canon 5D-50mm, Canon 5D-28mm, and Nikon D200, respectively for different intersection angles at a 90cm camera to object distance. The results became even better when different distances of camera to object were used with an intersection angle of 45°, with 93 $\mu$ m, 338 $\mu$ m, and 347 $\mu$ m for the Canon 5D-50mm, Canon 5D-28mm, and Nikon D200, respectively. Due to the large errors in the results, especially for smartphone cameras, the proposed methodology of texture depth measurement was performed using multiple images. The results from testing two control textures, rough and smooth showed that the calculated texture depth was significantly improved to the largest error of about 210  $\mu$ m recorded for the Samsung Wave and for smooth texture.

The use of stereo and multiple images were considered for greater range of asphalt texture through including ten test areas. Different analysis was performed to validate the author's developed methodology for texture depth measurement from images through using five cameras. The differences between depths from two images using the optimum camera to object distance and intersection angle revealed that the mean errors were 141 $\mu$ m, 177 $\mu$ m, 186 $\mu$ m, 355 $\mu$ m, and 230 $\mu$ m for the Canon 5D-50mm, Canon 5D-28mm, Nikon D200, Samsung Galaxy Note, and Samsung Wave, respectively.. On the other hand, differences between depths from multiple images with the depth from SPT showed that the mean errors became better, with 56 $\mu$ m, 72 $\mu$ m, 55 $\mu$ m, 80 $\mu$ m, and 114 $\mu$ m for the Canon 5D-50mm, Canon 5D-28mm, Nikon D200, Samsung Galaxy Note, and Samsung Wave, respectively. The experiments of measuring the correlation between depths from images and sand patch test were also considered. The outcome from these trials showed a highly significant correlation, with a minimum value of about 0.80 reported for the Samsung Galaxy Note and for stereo process.

The testing of other materials rather than asphalt pavement was also included in this chapter. Many materials such as concrete tile, gravel, soil, and grass were examined for reconstruction parameters output from AGISoft StereoScan. The findings from this test proved that the concrete tile material outperforms other materials in terms of providing a sufficient number of features and low RMSE with highly significant matching points. The rough asphalt pavement texture also showed an outperformance as it comes after the concrete tile as a second best material for surface modelling on an

automated basis. Visual inspection of the generated 3D surfaces of various materials revealed that almost all the materials were successfully generated automatically except for both gravel sizes and grass which might be due to low textured imagery, see section 3.8.4.

Two control concrete textures, rough and smooth were tested for texture depth measurement using multiple images for surface generation. Colour visualization of differences on concrete models revealed good consistency of different cameras with reference to the Canon 5D-50mm (C50) data. Cloud-to-cloud distance calculations showed that the largest error reported was for the Samsung Wave with the mean error of 234  $\mu\text{m}$ . Mean errors of other cameras with reference to the Canon 5D-50mm (C50) were 150 $\mu\text{m}$ , 139 $\mu\text{m}$ , and 161 $\mu\text{m}$  for the Canon 5D-28mm, Nikon D200, and Samsung Galaxy Note, respectively. Differences between depth from digital images and SPT were also calculated and found to be 295 $\mu\text{m}$ , 298 $\mu\text{m}$ , 281 $\mu\text{m}$ , 410 $\mu\text{m}$ , and 691 $\mu\text{m}$  for the Canon 5D-50mm, Canon 5D-28mm, Nikon D200, Samsung Galaxy Note, and Samsung Wave, respectively. Due to the apparent differences between depth from the sand patch test and digital images especially for smartphone cameras, different analysis approaches were considered for this test using the recreated surface of sand as a subtraction layer. It was concluded that the apparent differences between the sand patch test and digital images were due to errors in the sand patch test method of depth calculation, see section 3.8.5 and table 3.16.

The methodology of reconstruction of a 3D surface was utilized for characterization of fractured concrete. Nine fractured concrete samples were reconstructed. Four GCPs were utilized in each sample and collected using total station via reflectorless mode. RMSE of transformation (absolute orientation) values revealed that the reconstructed surfaces were accurate and in the range of less than 1mm. The precision was also considered and resulted in utilizing one of the used samples through reconstructing two repeated measurement. Trial of comparing the respective measurement suggested that the precision of the characterized concrete sample via Cloud-to-cloud distance calculation is in the range of  $(157 \pm 100)$   $\mu\text{m}$ , see section 3.8.6.

Finally, it can be concluded that the author's developed methodology of texture depth measurement from digital images showed promising results, especially for multiple images as an alternative method for SPT with which a highly significant correlation was achieved.

## **CHAPTER FOUR: A DEVELOPED METHOD OF CRACK DETECTION ON ROUGH TEXTURE**

### **4.1 Introduction**

Cracks on the road surface after construction are due to loading, environmental condition, ordinary wear and structure aging. It has been proved that early detection of cracks and maintenance can prolong the age of the road and minimize the expected cost. There has been a lot of research into crack detection on the road surface but very limited work has been conducted when it occurs on rough texture. Different techniques have been developed for crack detection. Some of them use image thresholding with image filtration to remove the noise, others use morphological tools while some use Neural networks (Chambon et al., 2009). In this project, a combination of methods of image thresholding, image filtering and morphological tools were used in order to detect cracks on the road surface efficiently.

The novelty of this method is focused on the automatic detection of cracks on the rough texture which is usually accompanied by lots of mis-interpretations due to raised aggregates. Most previous researchers have used only one image filtration technique for smoothing images (Subirats et al., 2006; Chambon et al., 2009). This is considered to be less useful because various texture roughnesses are available in the real world. The purpose of this chapter is to investigate crack detection using digital images. The main aim is to develop an algorithm that can deal with different image sizes, sources and lighting conditions. To fulfil this aim, many trials were conducted in order to assess the algorithm on different image sources (web, extracted from published papers, low resolution, and shadow existence) and with various texture roughnesses.

The development of the algorithm was carried out in this project because extensive literature reviewing revealed that:

- No particular study concentrates on the detection of cracks on the rough texture.
- Most of image filtration designed used only one filter for all types of textures.
- No open source online or commercially developed algorithms for crack detection were available that could be used for this investigation.

This chapter starts with describing the developed method for crack detection, followed by data set collection and test site location with different experiments for

identifying and locating cracks on the road surface. The problem of crack identification will be reviewed, and a trial of crack identification will be presented by using a thresholding method. Then the used materials are presented, followed by the automation procedures of the design concept. Finally, the results and analysis of different experiments will be discussed before presenting the summary and conclusions.

## **4.2 Design of a crack detection method**

Over decades, many methods have been developed to deal with crack detection on asphalt road pavements. Some of the developed methods are based on radar penetration or laser scanning techniques, while almost all of the established methods were based on images for identifying and quantifying the crack on the road surface. This was because it was showed that images are cost effective and more competitive than other systems. For images based methods, many of the algorithms were developed based on different techniques. Most of these techniques used 2D image analysis for segmenting or isolating the crack pixels from the rest of the image background.

In this project, a method was developed based on a digital image as a source of acquisition. The Canon 5D camera was utilized to capture the data of the road using video mode while on a moving van. The camera was fixed (2 m) on a mobile van over the road surface. As the van was moving, the video frames were recording the scene below the camera. A hand-held GPS and data logger were assembled with the camera in order for each particular frame to be geolocated. The processing stage started with the extraction of the video frames which was done through the use of feature based matching algorithms in order to include the required frames only. Later, the extracted frames were geotagged using the GPS log to assign the position and GPS time.

After the frame extraction and geotagging steps, the frames were subjected to the process of shadow detection and removal. Then the pre-processing stage started with image filtration and morphological operations to smooth the image, strengthen the crack signal and remove the noises. Different image filters were applied based on roughness modelling using 2D image analysis. Next, image segmentation based on edge detection was started to identify possible cracks, followed by a post-processing stage to connect crack candidates using a minimum cost path (locate the minimum intensity path that connects between two crack objects) and remove mis- interpretations based on crack

candidate statistics. This process is described in the workflow diagram presented in figure 4.1.

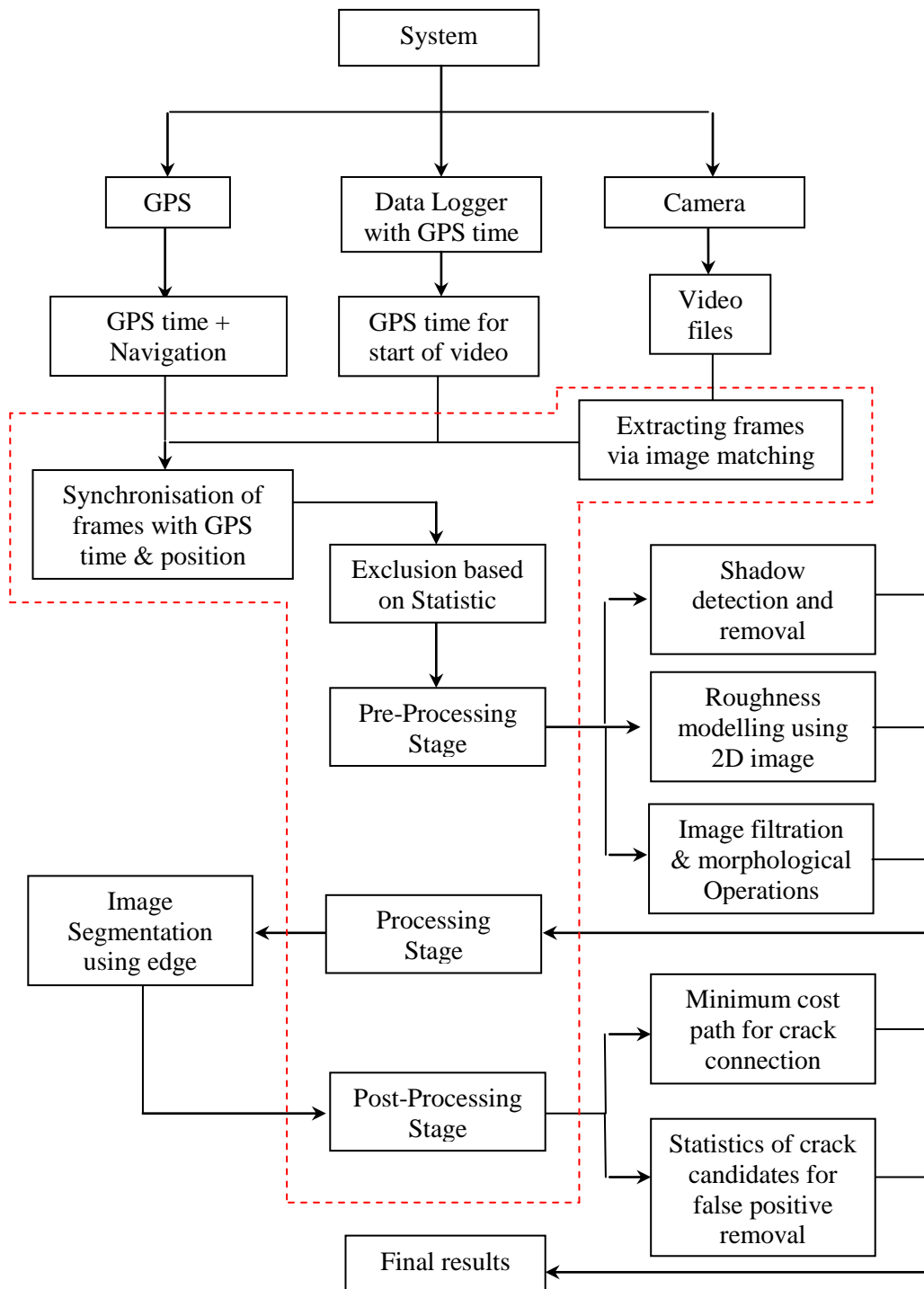


Figure 4.1: Workflow diagram for developed method of crack detection

--- Box containing the developed algorithm processes

### 4.3 Aims and Objectives of trials

The aim of this chapter is to design and assess a crack detection algorithm and to test each processing stage, which will be achieved through the following objectives:

1. Investigating video frame extraction using the invariant feature matching algorithm (SURF).
2. Investigating the use of hand-held GPS for defining the location of the extracted frames.
3. Investigating the use of statistics for identifying the frames where problems exist (potholes, manholes, shadow, cracks and other unusual objects of road surface).
4. Investigating the use of a 2D analysis method for identifying the roughness of the texture on image.
5. Investigating the crack detection algorithms for different sources of images.
6. Investigating the crack detection algorithm on video images captured from a mobile van.

### 4.4 Methodology of trials

To fulfil the above objectives, following methodology was considered:

1. Objective 1: Video files listed in table 4.1 were used for testing the developed methods of frame extraction.
2. Objective 2: Extracted frames from the junction points from the conducted survey path were used, see figure 4.5.
3. All frames extracted from objective 1 were used for analysis and statistical testing.
4. Ten samples of texture with known texture depth were used, see figure 4.4.
5. Different source of images were used for testing the crack detection rate, see figure 4.3.
6. Some samples of extracted frames from video files with different problems were utilized, see figure 4.47.

## 4.5 Data sets and test site description

The research investigated the possibility of detecting cracks on close range images through developing an algorithm based on thresholding, image filtration, and morphological operations. For this purpose, the first data set was acquired using a Nikon D100 camera in a mobile car with average traffic speed. The vehicle moved over different street areas in Nottingham, starting from the Nottingham Geospatial Building to the Northgate Primary School, figure 4.2. The camera was held obliquely by hand and at about 1.2m height over the road surface. Camera exposure settings were fixed, with the aperture priority used to maintain a sharply focused image of the road surface, and the ISO sensitivity was set to 200. The shutter speed changes accordingly to compensate for both aperture and ISO sensitivity fixed values; a sample of images is given in appendix B. The collected images were visually checked for the presence of cracks to be used as trial images. This set of images was used for developing the algorithm, testing for the crack detection as well as for the texture analysis or roughness measurement.

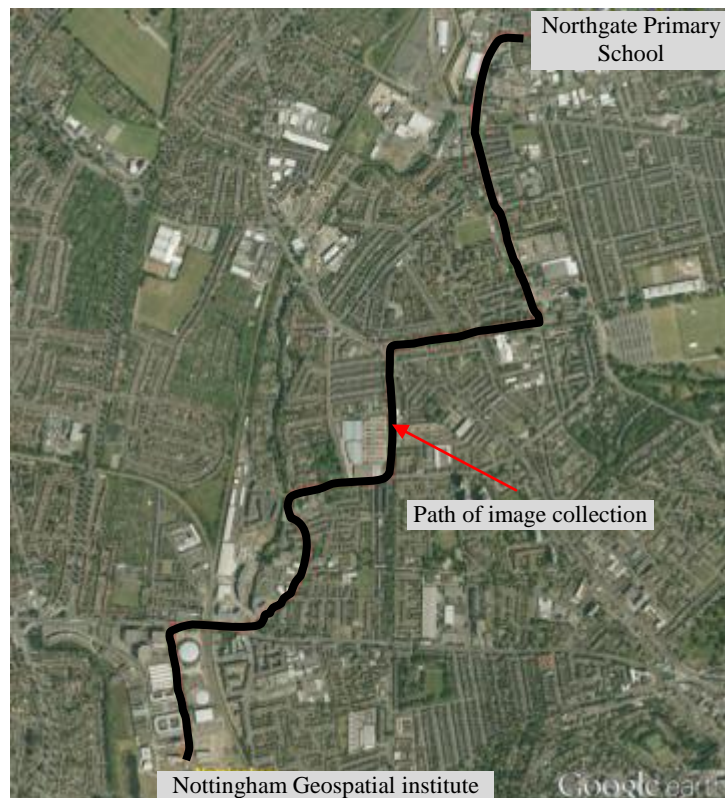
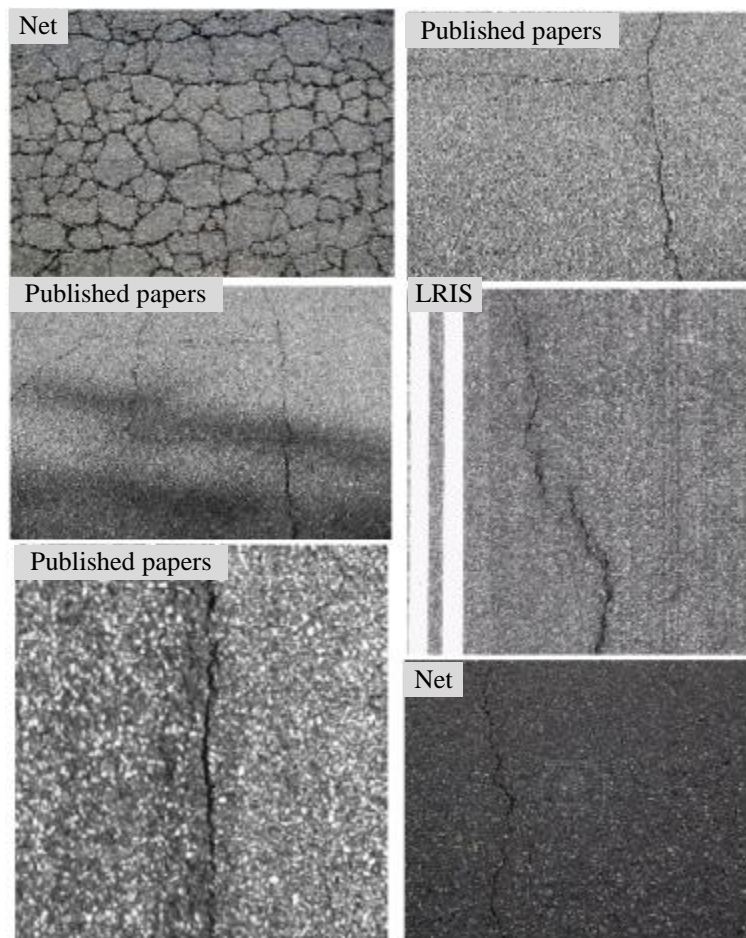


Figure 4.2: Path of collected data set for crack detection algorithm development

In order to investigate the potential of the author's developed crack detection algorithm, different sources of images (web net, extracted from published papers, low resolution, shadow existence, Laser Road Imaging System (LRIS) and collected from the real world) were examined. There are different sources of images available online which have different types of crack extension and levels of severity and with different resolutions. More information regarding these images is available in the results and analysis (section 4.9). Images of cracks were also extracted from some published papers using the cut option and the low resolution available on the presented figures. Image pyramid reduction techniques were applied on some images in order to test the algorithm on the low-quality resolution images. Figure 4.3 shows some samples of the images extracted from published articles, downloaded from the net or from LRIS (Sy et al., 2008; Pavemetrics, 2012; Zou et al., 2012).



**Figure 4.3: Samples of crack images from variety of sources**



Another set of data was collected for investigating the author's developed crack detection algorithm using video mode. This set was obtained using a designed system which consisted of a Canon-5D camera with hand-held GPS mounted on the top of the NGI's mobile van. The camera was mounted with its optical axis perpendicular to the road surface at a height of about 2m, figure 4.4. The camera used on this system was a Canon 5D, which was fixed on the end of a wooden frame. Along with the camera, different sensors (GPS receivers and data logger) were also assembled on the mobile van.

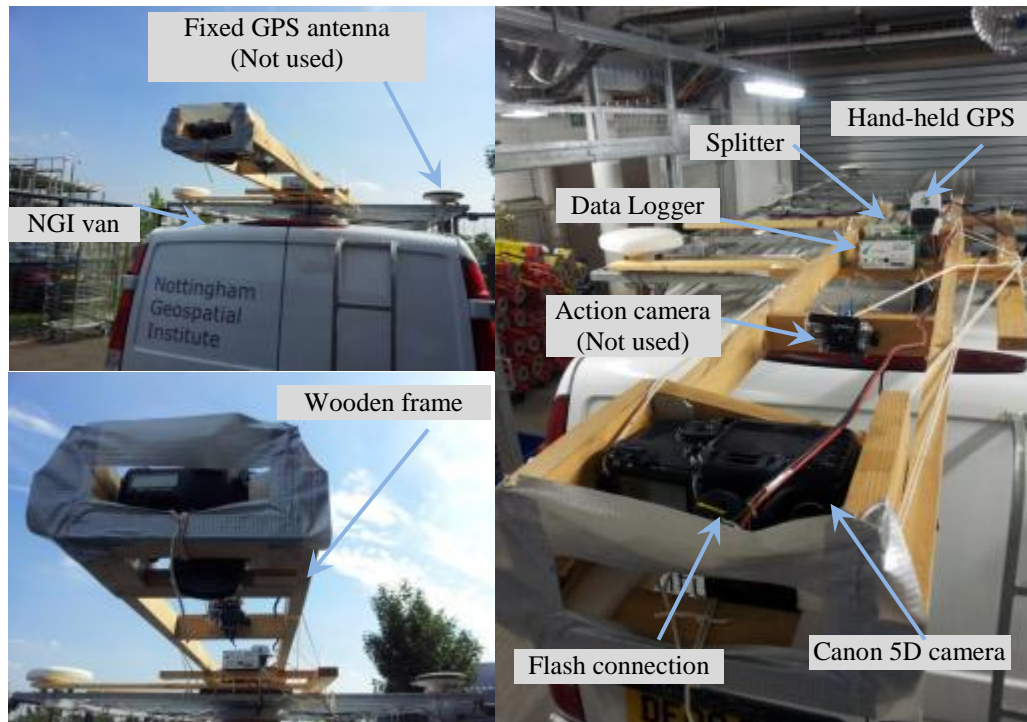


Figure 4.4: System for image data collection on mobile van

The data collected from this system was used for three purposes. First, to test the possibility of acquiring quality images without blurring effects under different traffic speeds, see section 4.9.3. Second, for assessing the author's developed crack detection algorithm on the extracted video frames by comparing it with manually measured values, see section 4.9.3. Finally, to examine the video extracted frames for the detection and extraction of potholes using 3D based image modelling techniques on a fully automated basis, see section 5.7.4. More details about the collected data and the proposed system are described in the next sections. The extracted frames were tested for both 2D image analysis and 3D image-based modelling. A method was developed to deal with frame extraction and handling of extensive data which will be explained in the following sections.

Adequate images were obtained from different roads at various traffic speeds using the video mode of the Canon 5D camera. A total length of about (27 km) of data was gathered for different roads and during different times of day (morning, noon and afternoon) under day lighting conditions. Some of these images were overlapped for the same road surface, see figure 4.5 for the path of collected data and refer to table 4.1 for details about individual video files and the total number of frames.



Figure 4.5: the path of all collected data from Radford area on mobile van

Table 4.1: Video files detail for the data collected

Video Files	Time (min.sec)	Length (Km)	Total number of frames
MVI_4512	8.26	2.6	15280
MVI_4513	1.52	0.5	3360
MVI_4514	13.17	4.16	23910
MVI_4517	4.13	2.20	8760
MVI_4522	10.27	1.94	19783
MVI_4523	14.30	4.1	27029
MVI_4539	11.00	3.6	19800
MVI_4540	10.08	3.2	20700
MVI_4541,42,43	17.52	5.3	32160

## 4.6 Problem definition and trial with crack detection

A crack is considered difficult in the context of road surface evaluation by using image techniques because it is weakly represented (about 1.5% of the whole image) and weakly contrasted (easily hidden by road surface texture). Many methods have shown limitations because the detected results contained lots of mis-interpretations due to variations in the road texture (Sylvie and Jean-Marc, 2011). Rough texture can also cause slight variations in the average intensity of local pixel regions even if the image is without cracks; figure 4.6 shows a sample of texture variations and the effect of texture on blending the crack pixels.



**Figure 4.6: Some available texture in the road pavement**

Conventionally, methods of crack detection can be divided into four major groups based on using Thresholding, Morphology, Neural Networks or Transformation techniques (Chambon et al., 2009). The general approach developed in this research for crack detection was based on combining segmentation using edge operations with image filtration and morphology methods. For the same sample used, the morphology operation was applied to the RGB image in order to enhance the image, reduce noise and remove unwanted objects from the image, such as lanes or marks on the road; whereas, the thresholding method was used to identify the crack pixel candidates. The threshold method is a simple and effective method; however, because the aim was to use automatic detection of cracks, automatic thresholding was required.

A trial was carried out to detect cracks based on intensity variation through local thresholding definition.

Crack detection is mainly based on intensity variations, because the crack pixels are darker than their neighbours, figure 4.7 shows a sample of crack pixels. The segmentation of a crack pixel from the background texture is considered as a challenging process due to the variation of illumination around the image. The variation in intensities between crack pixels and neighbouring pixels can help the segmentation process by setting threshold values which can be chosen automatically or manually.

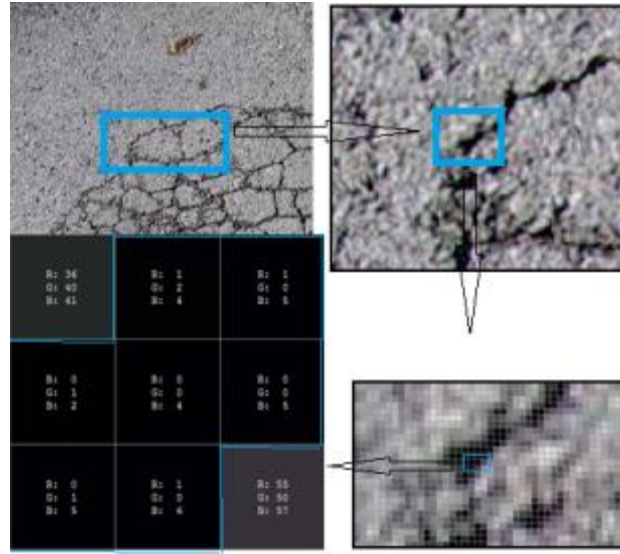


Figure 4.7: Intensity variations of crack pixel and its neighbours

However, the limitation of the threshold method is the variability of the intensity values due to the change of illumination around the image, see figure 4.8.

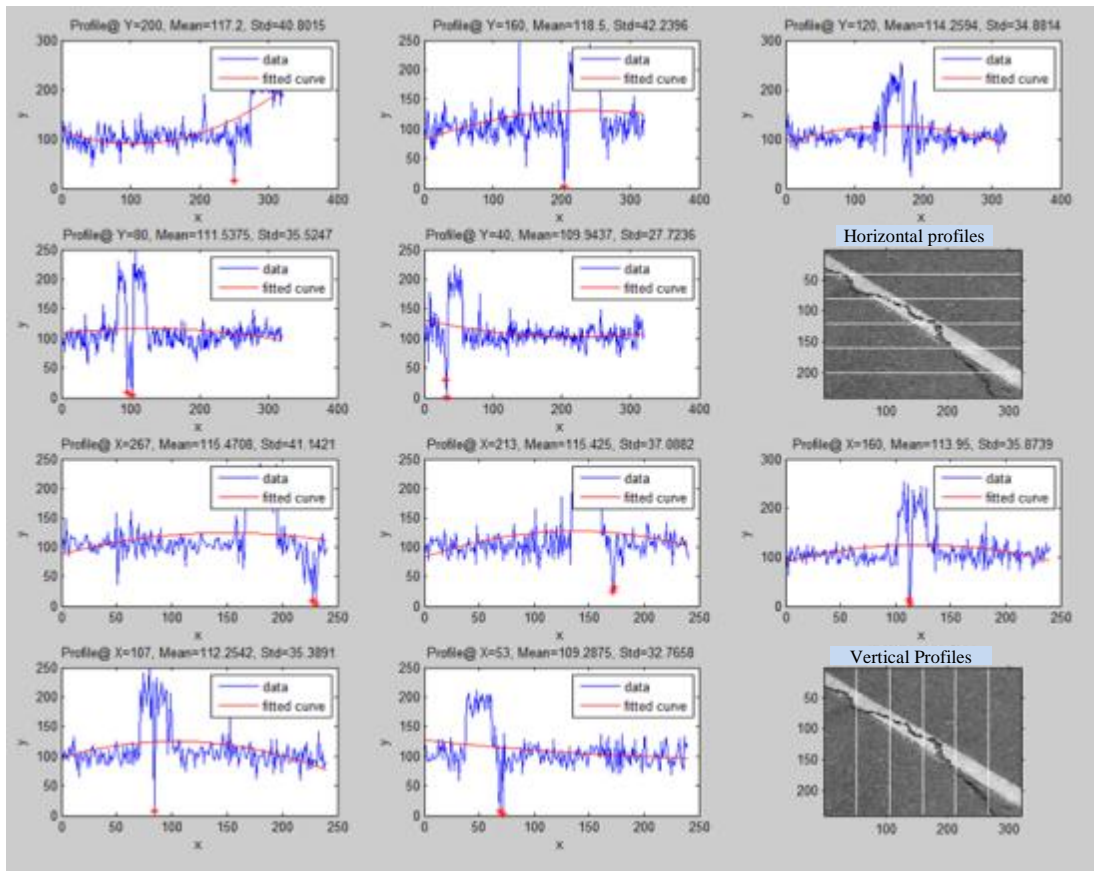




Figure 4.8: Intensity profiles and curve fitting through image

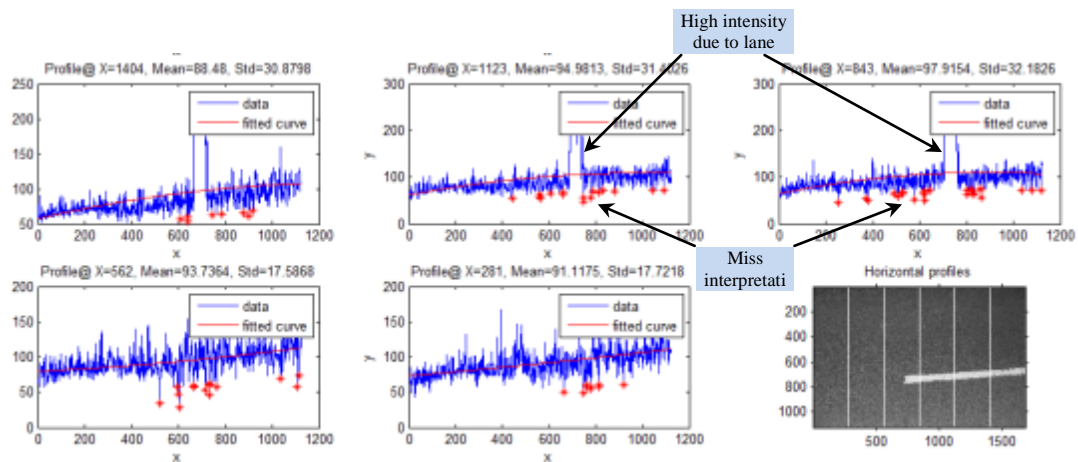
• Crack candidates,  Profile intensities,  fitted curve

x= pixel location along profile, y= intensities

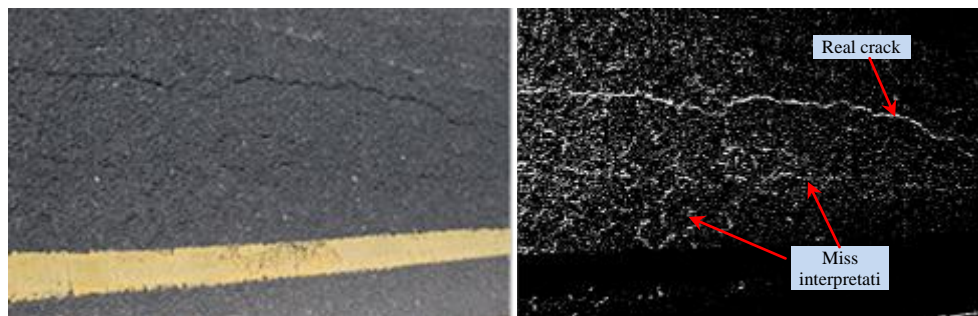


Based on these profiles and the fitted curve, the mean and standard deviation of intensities can be calculated. The crack candidate pixels can be identified along the profile based on the mean and standard deviation of intensities. The deep dark pixel values were considered to be crack pixels if they had 3- sigma values away from the mean value. From these profiles, the threshold limits used for segmenting the crack pixels from the whole image can be calculated.

This process will identify crack seeds but not the whole crack pixels. Moreover, the fitted curve and the crack detection idea based on the 3-sigma criteria might in some cases detect non-crack seeds, or mis- interpretations. As shown in figure 4.9, although there are no cracks along the profiles, this process identified cracks based on the 3-sigma criteria, however these can be avoided through removing the effect of bright objects. The results from this method showed many mis- interpretations after testing the sample as shown in figure 4.10.



**Figure 4.9: False positives (identifying non-crack pixel as a crack)**  
x= pixel location along profile, y=intensities



**Figure 4.10: Results from thresholding method based on profile intensities**  
original image (left) and results (right)

## 4.7 Materials and methods

The camera had a primary contribution in this study through the use of a Canon 5D and Nikon D100 (for details about Canon 5D refer to section 3.6.1). The Canon 5D was used in this project through using the video mode for the collection of road surface imagery under day lighting conditions. The Canon 5D was chosen because it can offer a 1080\*1920 pixels image resolution with about 0.018  $\mu\text{m}$  sensor pixel size in video mode. Over 2m height and with a 28mm focal length, the pixel size (GSD) on the ground will be about 1.3 mm. Tests of the author's developed crack detection algorithm proved that detection of crack was successful in most cases if the crack width is more than 3 pixels(see section 4.9.2 and 4.9.3). Therefore, it is expected that the Canon 5D camera can detect cracks with about 4 mm width. The Ground Sample Distance (GSD) in the recent LRIS system is in the range of 0.5-1mm for the 8000 and 4000 pixel options, respectively (Pavemetrics, 2012). Therefore, the pixel resolution in the case of the Canon 5D camera is sufficient if compared with the LRIS system.

Other cameras were also utilized for still image collection. For example, the Nikon D100 was used to collect some imagery data for algorithm development purposes. The Nikon D100 has a 28 mm focal length with 23.7\*15.5 mm CMOS frame size and 0.0079  $\mu\text{m}$  effective sensor pixel size, with a total of 6.1 effective megapixels.

Another sensor, a hand-held GPS receiver (GPS 76) was also used with the camera to measure roughly the location of the video frame when the camera was moving. Theoretically, it is estimated that the accuracy of hand-held GPS (GPS 76) is always less than 15m with RMS 95% for standalone and less than 5m with RMS 95% for Differential GPS (DGPS). The cost of this hand-held GPS receiver on the day of production was about £200 (Yeazel and Birch, 2002). For more information regarding the GPS 76 and the accuracy of the positioning refer to Garmin (2004). A precise time measurement tool data logger, Precise Time Data Logger (PTDL1), was assembled with the camera in order to calculate the start time of when the camera was triggered with regards to GPS time. The data logger was used for all video files for the purpose of synchronisation between camera time and GPS time. For more information about the

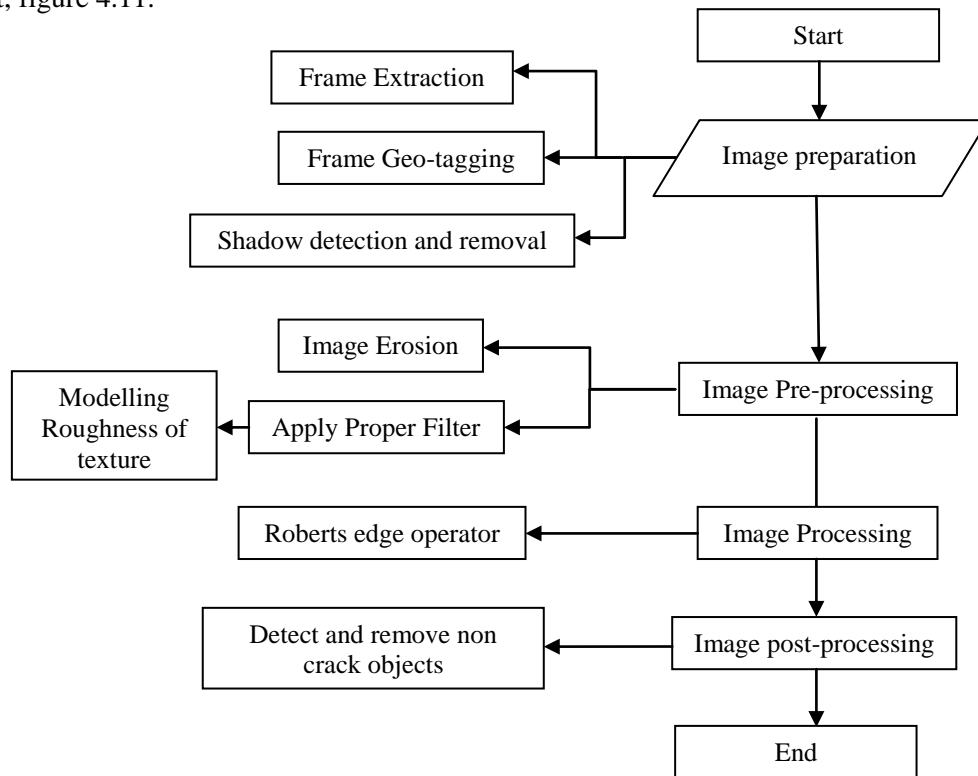
used data logger PTDL1, please refer to the manual at Barnsdale K.P. and Andreotti (2008).

The author's developed crack detection method used in this project was based on detecting the crack automatically and involves:-

- 1- Image noise reduction through applying a designed filter based on measuring the degree of roughness of the texture.
- 2- Segmentation based on Roberts edge detection for identifying crack objects.
- 3- Crack object identification based on geometry.
- 4- Crack objects connection based on a designed method of minimum path identification.
- 5- Morphological tool for pre-processing the image and post-processing the results in order to remove or mitigate the outliers (more details are discussed in the following sections).

## 4.8 Semi-automatic procedures for crack detection

This chapter presents the work or procedures followed to identify cracks on the road surface. The methodology described here is to show the algorithm details in each step progressively. The algorithm called crack detection was developed using the Matlab program. The goal of the algorithm was to successfully identify crack candidate pixels and non-crack pixels within the image frame. To achieve this aim, a number of operations were designed to detect the crack with the minimum amount of noise or misinterpretations. The concept of the algorithm can be summarized in the following flow chart, figure 4.11.



**Figure 4.11: Concept of crack detection algorithm**

Procedures shown in figure 4.11 were developed using a variety of available or developed tools. For example, at the image preparation stage, the video frame extraction methods are available in Matlab toolbox. However, the process in Matlab was not enough to remove unwanted frames that overlapped when the car was fully stopped. Therefore, a methodology was developed to remove unwanted frames that significantly reduced the time of processing. It is also worth to mention that at each of these stages, a mixture of thresholding, filtration, and morphological operations were used which will be described in more details in the following sections. The first stage in algorithm development was image preparation that will be discussed in the next section.



### 4.8.1 Image preparation

The objective is to detect a crack on the video extracted frames obtained using the Canon 5D camera which records 30 frames per second with full HD resolution (1080\*1920 pixels). Plenty of unwanted data was recorded especially when the car was fully stopped in a traffic light control zone or when the car was in a traffic jam. The image preparation stage involved two principle processes; the first one was frame extraction with motion estimation. A motion estimation routine was developed to remove unwanted images of the fully overlapped areas or with a range that is less than the required distance, which will be explained later. In the second operation, the extracted frame was geotagged which involved assigning GPS position to the extracted frames. These two operations are described in the following sections.

#### 4.8.1.1 Frame extraction with motion estimation

In this section, a method of extracting the video frames and removing frames that shared areas with other frames is presented. As mentioned before, the Canon 5D camera in video mode records 30 frames per second. With the car travelling on an average traffic speed of about 30 mph which is equivalent to 13.41m/s, 1 km of data will provide about 2237 frames, which is an enormous amount of data and could be more if the car was fully stopped at any time or was travelling at a slower speed. In order to manage this vast number of frames, a subroutine algorithm was developed based on the feature matching algorithm (SURF). SURF was used to estimate the motion between a pair of images. The first frame is compared with the next frames until finding the required shift distance which shared about 20%, or the shift distance of more than or equal to 800 pixels, figure 4.12 shows the procedures or flowchart of the subroutine.

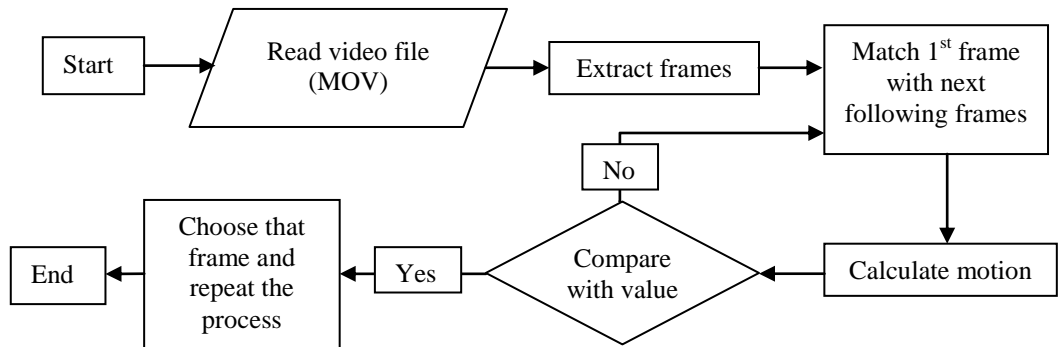
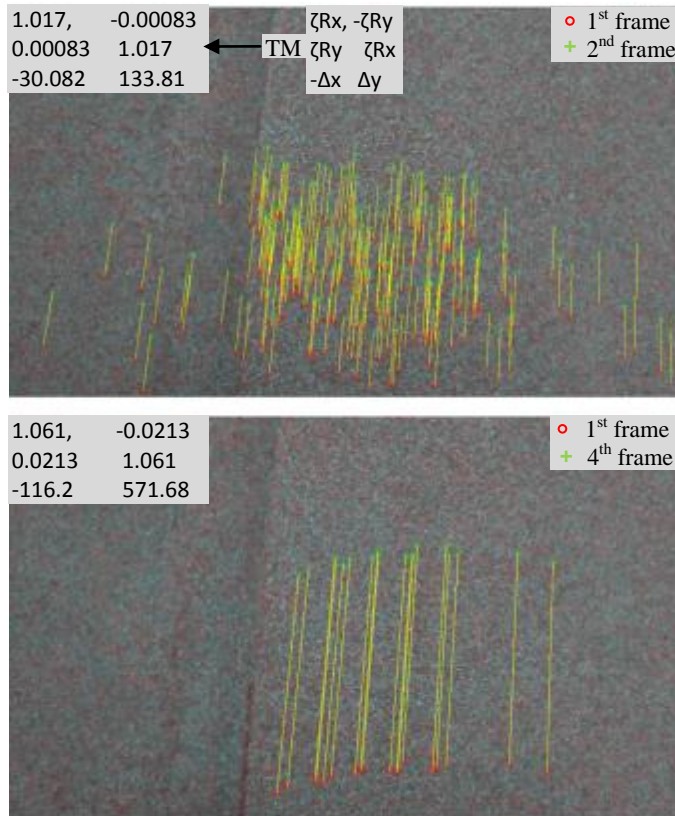


Figure 4.12: Frame extraction and motion estimation flow chart

After extracting frames from video files, the first frame was compared with the following frames by matching the two images and then calculating the transformation matrix between successive pair of images. In the sample shown in figure 4.13, the matching and 2D transformation matrix (TM) results are shown comparing the first with second and fourth frames successively.

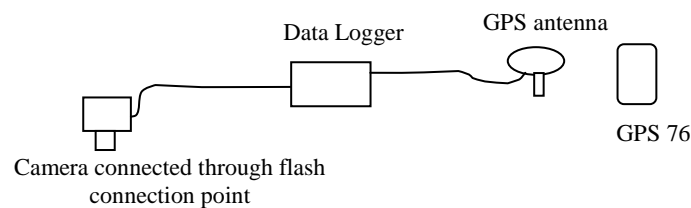


**Figure 4.13: Sample of motion estimation through transformation matrix calculation**

The transformation matrix is computed using the inlier matches calculated through a RANSAC outlier removal algorithm. The model parameter is estimated from seven matches and then applied on the rest of matches. The residuals calculated from each model and the small registered amount through required iterations was chosen. As can be noticed from figure 4.13, the transformation matrix provides information such as (rotation, scale and translation) between two frames. From this information, the proper image frame was selected, and all other frames in between were removed. More details about the results of this method and their effects on frame reduction percentage are available in section 4.9.1.1.

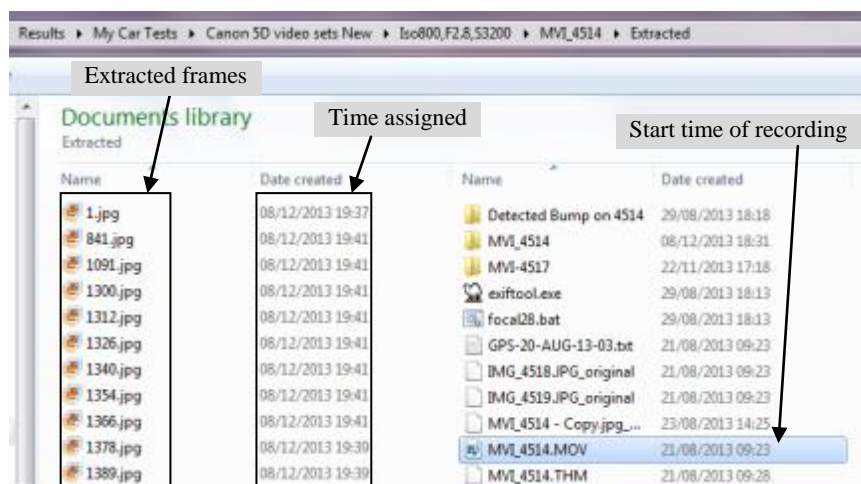
### 4.8.1.2 Frame Geo-tagging

At this stage, the GPS log is used to position each extracted frame after the motion estimation phase. As mentioned before, the GPS receiver (GPS 76) was mounted with the Canon 5D camera to register the position every second. The time of starting the GPS log and the camera shooting was recorded through the PTDL1 data logger which uses GPS time to record all actions occurring at the same time. When pressing the recording mode on the camera, it sends a pulse through the flash connection point to the data logger and the GPS time is attached with this pulse. This time is then synchronised with the time from GPS log recorded through the GPS 76 receiver, figure 4.14.



**Figure 4.14: The system concept of recording time of starting camera**

The Canon 5D camera, and most cameras, record only the beginning of the video file time. When the video file was extracted to frames, the timestamp of the frame was changed to the time when the frame was created using a Matlab program (it uses the clock time of the computer to assign time to the frames). This time needs to be corrected using the start time of the recording video file which was synchronized with the GPS time, and adding it to the frame number which is attached to the file name as can be noticed in figure 4.15.



**Figure 4.15: Extracted frames from video file (MVI\_4514.MOV)**

For assigning the correct timestamp to the extracted frames, the ExifTool (Harvey, 2013) was installed on the system so that it can be recalled inside the Matlab program. The Exiftool has the capability of assigning a timestamp to the image frame. However, the frame rate is 30 per seconds, and this time fraction cannot be assigned to the Exif header, which required another solution. To overcome the problem, track of GPS time and position required interpolation.

The GPX file format was converted into a tab delimited file format (txt) in order to be dealt with more easily. All the latitudes and longitudes were transformed into Eastings and Northings Ordnance Survey Great Britain 1936 (OSGB36) system using Grid InQuest coordinate transformation tool (Ordnance Survey, 2014), see a sample in table 4.2. The time of frames is more frequent than the epoch of the GPS. One second GPS time is equal to 30 frames of images. Therefore, in order to assign the correct positions to the proper frames, two GPS positions with one second-time span were interpolated to 30 positions in between, figure 4.16.

Table 4.2: Sample of GPS positions (WGS84) transformed to OSGB36

Index	Time (Hourse)	Position (Lat, Lon)	Easting	Northing
120	15:44:43	N52 57.116 W1 10.997	454969.1	339704.6
121	15:44:45	N52 57.109 W1 10.997	454969.2	339691.6
122	15:44:47	N52 57.108 W1 11.003	454962.6	339689.7
123	15:44:49	N52 57.106 W1 11.010	454954.8	339685.9
124	15:44:51	N52 57.107 W1 11.020	454943.5	339687.6
125	15:44:53	N52 57.105 W1 11.028	454934.6	339683.8
126	15:44:55	N52 57.103 W1 11.035	454926.8	339680
127	15:44:56	N52 57.102 W1 11.037	454924.6	339678.1
128	15:44:58	N52 57.099 W1 11.044	454916.8	339672.5
129	15:44:59	N52 57.098 W1 11.045	454915.7	339670.6
130	15:45:00	N52 57.097 W1 11.046	454914.6	339668.7

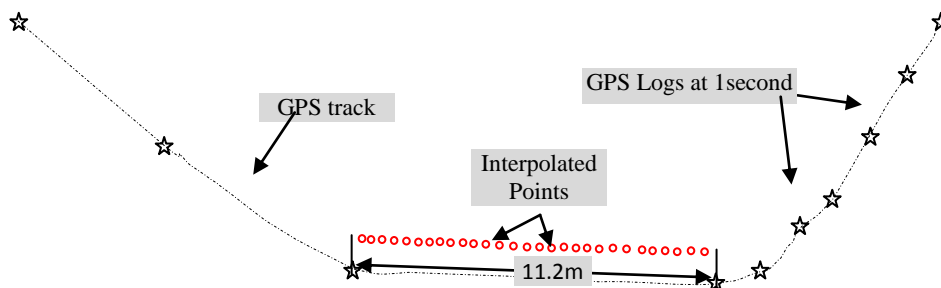


Figure 4.16: sample of GPS track and interpolation points

### 4.8.2 Exclusion of problem free images based on statistics

The processing time of the author's developed crack detection algorithm is mainly dependent on the number of candidate objects. The time required for pre-processing and post-processing is within less than five seconds for almost all image types. However, the time required for the post-processing stage is 300 times longer than the two other stages, table 4.3 shows the time required to process only one frame of the image.

Table 4.3: sample of time processing for a one frame of image

Process	Time required(sec)
Pre-Processing	2.5
Processing	1.53
Post Processing	1510

It is obvious from table 4.3 that the post-processing stage is very time-consuming and this needs to be carefully considered. This time was only required for processing one frame, and it was mainly due to the use of the iteration process or for-loops in Matlab. The candidate objects for post-processing were 2316, which means that the program had to iterate 2316 loops. However, it is only necessary to run the program for the frames that contain problems or distresses (lane painting, cracks, shadows, and potholes).

The basic method for this was to examine histogram distributions. It was found that clear road pavement images will follow the Gaussian pattern, figure 4.17. The residuals from the data and the fitted curve were calculated, and the RMSE from the residuals was designed as a limit to include or exclude the frame in the next processes. The algorithm starts by converting the true colour image RGB to the grayscale intensity image, eliminating hue and saturation information while retaining luminance. The next step is to calculate the histogram for the intensity image or grayscale image by defining the number of bins and saving a plot of the histogram. Fit Gaussian is then applied to the histograms utilizing the try/catch statement to cause the program not to encounter any error if the Gaussian fit is not working (in case data are not normally distributed). The residuals are finally calculated, and the RMSE from the residuals calculated, which helps for decision making, see figure 4.18.

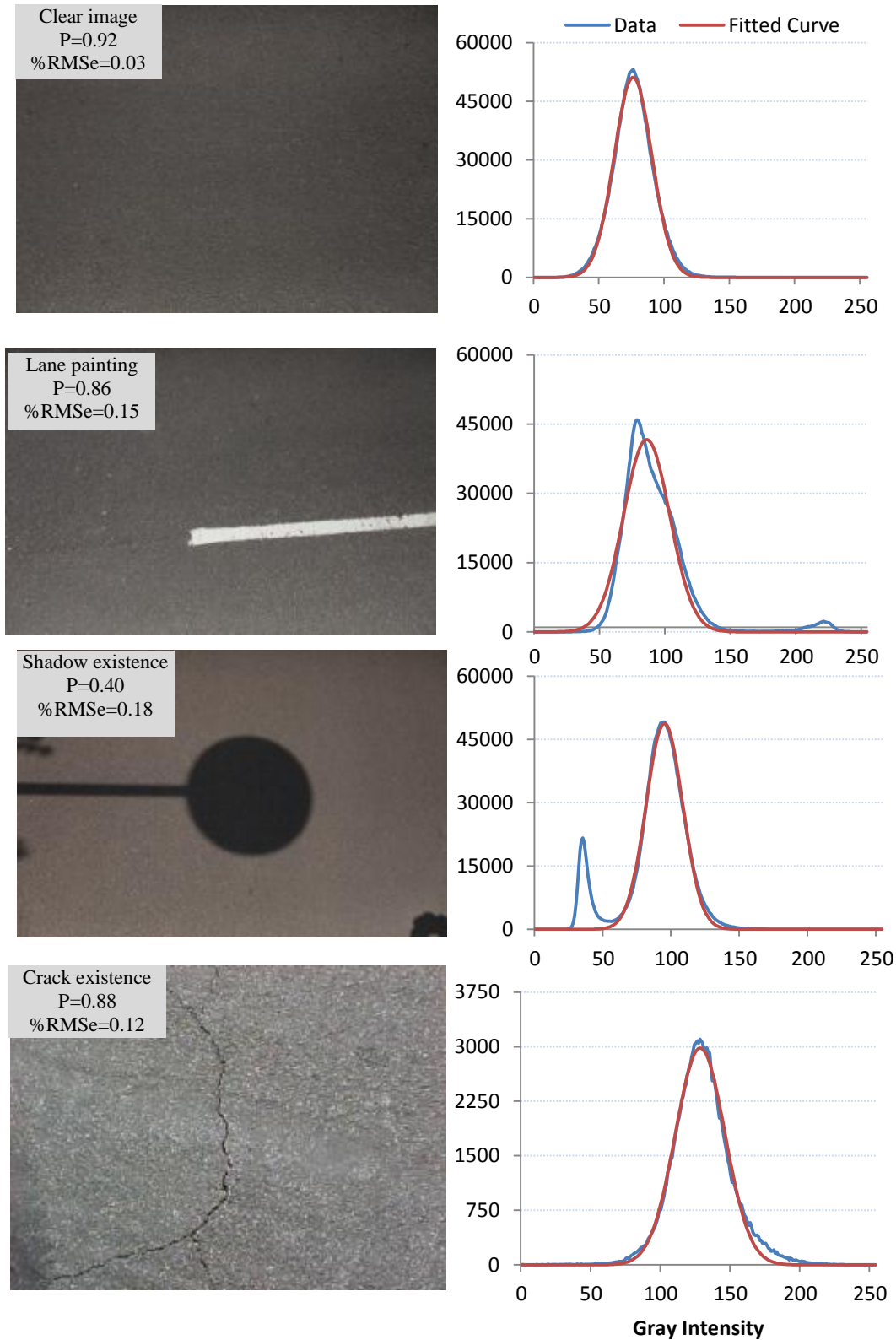
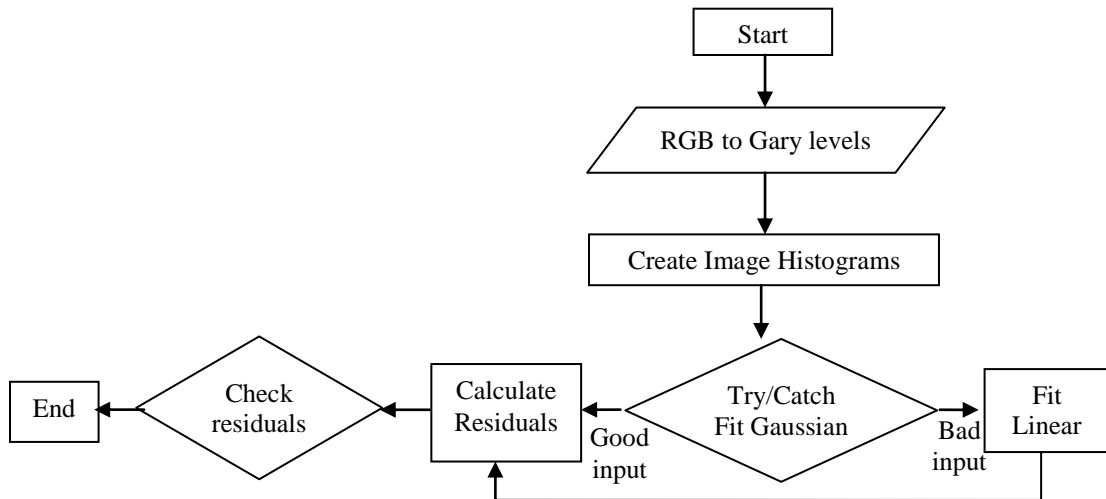


Figure 4.17: Histogram and fitted Gaussian curve for sample of images



**Figure 4.18: Flowchart describing the methodology of Gaussian best fit and image histograms**

In addition to RMSE from residuals, the statistical t-testing was applied as well in order to compare the similarity of the two data sets (data and fitted curve). T-test is usually used to determine if two sets of data are significantly different from each other. T-test is commonly applied when the two sets of data follow a Normal distribution. The null hypothesis is that the two sets of data are from equal means. The probability (p)-value calculated from the outputs of the T-test to measure the strength of null hypothesis. A p-value of one indicates that both sets of data are in identical (accept null hypothesis with 100% confidence interval), whereas a p-value of zero indicates that the sets of data are entirely different (reject null hypothesis with 100% confidence interval) (GraphPad, 2014; Matlab, 2014d).

Figure 4.17 shows a sample of images that were either rated as free of problems, or clears, or containing different sources of problems, such as shadow, lane painting and crack existence. As can be noticed from figure 4.17, the p-value output from the t-test calculated between the two sets of data (data and fitted curve) and RMSE from residuals are changing from one sample to the other according to the source of the problem. The hypothesis that a p-value of less than 0.90 and an RMSE value of greater than 0.1% indicated an existing problem was applied. The RMSE values from data and fitted curve were calculated as pixel counts from the residual calculation. In order to eliminate the size of image effect on the RMSE value, the residuals in pixel counts number was divided by the total count of pixels and then converted into a percentage. This method was evaluated based on testing the methodology on various types of images, and the results are presented on section 4.9.1.3.

### 4.8.3 Image pre-processing

After frames were extracted and geotagged using interpolated GPS position logs and statistically evaluated for problem existence, the next step was image pre-processing. This process involved a number of operations in order to remove shadows, enhance the image, mitigate noise, and remove unwanted objects, was will be explained in the following sections.

#### 4.8.3.1 Shadow detection and removal

The first step in the pre-processing stage is to examine the video frame or still image for shadow existence. This step was important because the author utilized natural lighting for image collection, as this was considered to be a cost effective process compared with LRIS. The process of shadow detection and removal was applied through two main steps. In the first step, the image or frame was subjected to a process of examination for shadow existence that could be achieved via K-mean clustering algorithm. This process was conducted using the K-mean clustering algorithm with pre-defined detection groups of ten colour space objects. Then, the correct number of clusters is defined using Silhouette tool, for more information refer to (Matlab, 2014c). The procedures of all steps are described in the flow chart presented as figure 4.19.

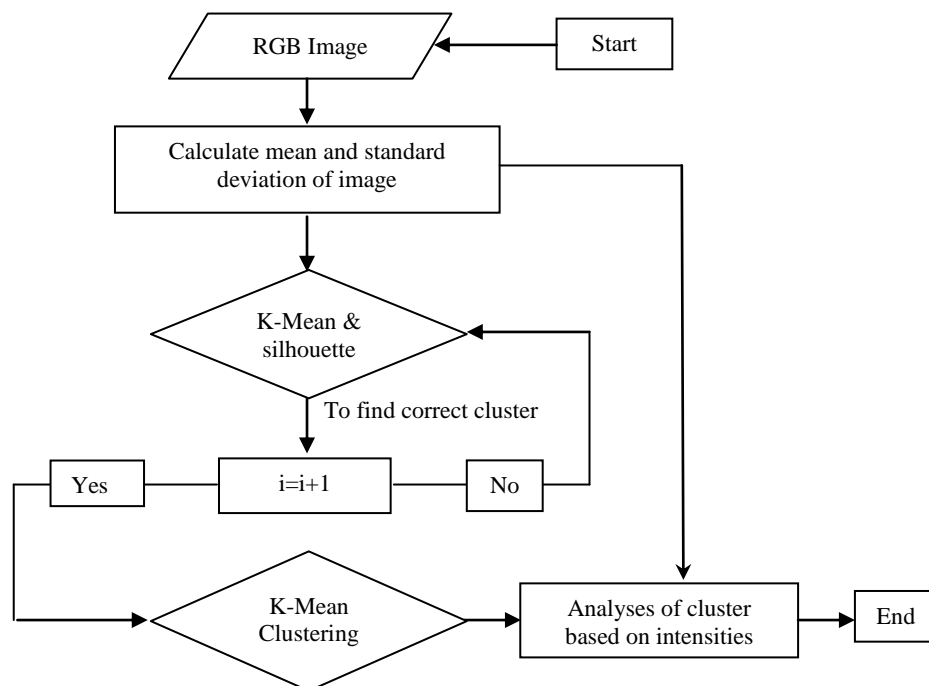


Figure 4.19: Flow chart shows the use of K-mean clustering for shadow detection



After the frame was tested for shadow existence, the next step was to remove the effect of the shadow in order for the segmentation process to be as correct as possible. For shadow removal, the work by Blajovici et al. (2011) was adapted. The central idea of their work was based on statistical analysis of the intensities of images converted to colour space objects. This algorithm has been successfully tested on many applications such as segmentation, object detection, stereo, and tracking, and the authors provide lots of examples at <http://www.inf.u-szeged.hu/projectdirs/ssip2011/teamF/>. A binary shadow core mask is prepared as part of algorithm procedures which could be used as a reference for mis-interpretations removal.

The outputs from the shadow removal algorithm are free shadow image and shadow core mask which were both used as input for the author's developed crack detection algorithm, see figure 4.20. The free shadow image is directly used as input for the algorithm, and the whole process of crack detection is running on this image. The shadow core mask represents the output from statistics of intensities. It is modified before it can be used successfully for removal of outliers in the final stage of crack detection. The shadow core mask can then undergo two further processes in order to be successfully used for mis-interpretations removal. The first step in preparing the mask removal includes filling the holes. In the second step, small objects such as crack pixels which treated as a shadow mask are removed, as these removed crack objects have equal intensities to the shadow but have much smaller sizes.

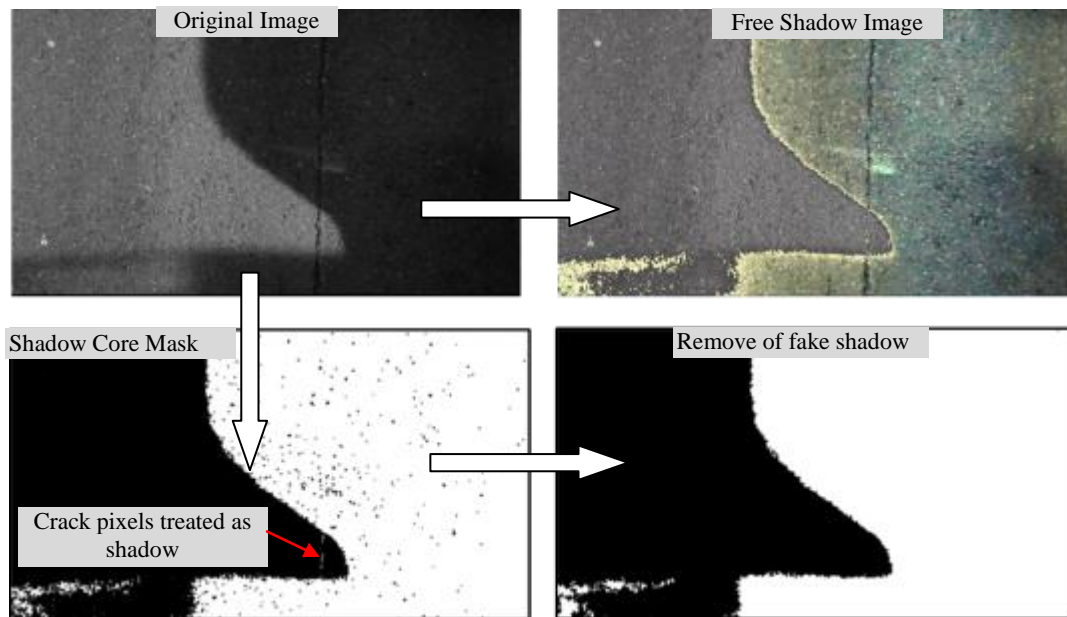


Figure 4.20: Results from Shadow removal algorithm and post processed shadow core mask

#### 4.8.3.2 Morphological erosion and image filtration

It was shown in the previous section how a frame or image can be examined for shadow existence and removal. In this stage, the image undergoes two operations in order to enhance the image and strengthen the crack signal. In the first operation, the grey scale image is applied with the structure element in order to remove paints or markers on the road surface. This procedure is called morphological erosion and it is used to remove the effect of bright pixels, such as the painting on the street. The mark on the street is considered as a foreground pixel because the intensity of these markers is always bright. The structuring element used with this process is highly dependent on the size of the mark. A disk shape with the size of  $(0.12\%)^2$  pixels from the image size was found to be appropriate for reducing the effect of the marking with different sizes. Refer to appendix C for more details about the effect of using different structuring element size. After erosion, the image is reconstructed using the original grey scale image as a mask and the result of these two steps is shown in figure 4.21.

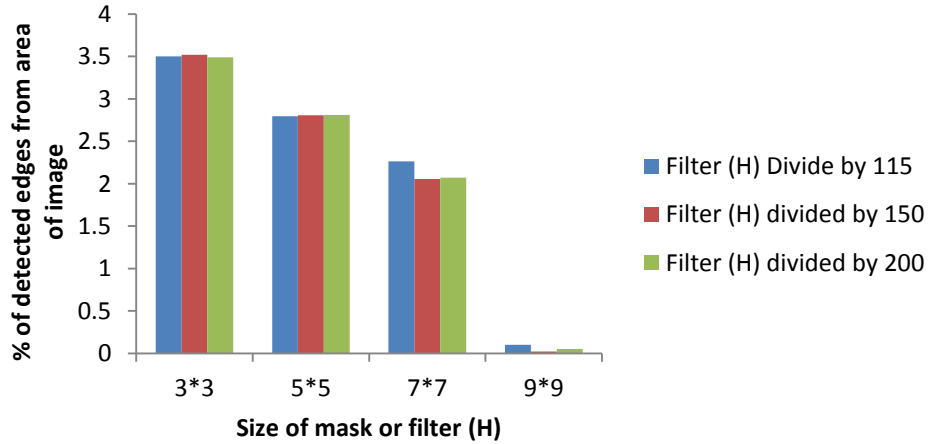


**Figure 4.21: Original image (left), eroded and reconstructed image (right)**

After reconstruction of the eroded image, a further pre-processing operation, which is image filtrating, was applied in order to smooth the image. The purpose of filtration is to enhance and remove any noise in the original image before trying to locate and detect the edges. For this purpose, the image filtering can be computed using a well-studied simple multidimensional linear mask or filter (H). From different trials, the size of a  $7 \times 7$  filter mask was found to be in a midpoint between both sensitive and non-sensitive edge percentage, figure 4.22.

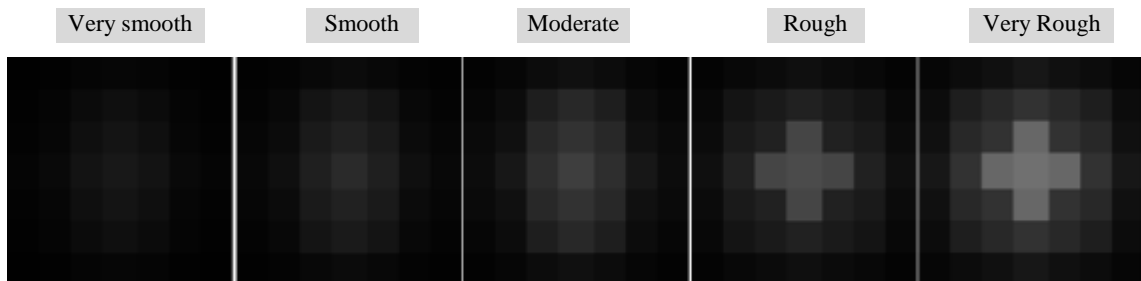
The mask of filter (H) is usually much smaller than the actual image size. As a result, the mask window is shifted over the image, amending a square of pixels at each pixel location. Therefore, the larger the size of the mask or filter, the lesser is the

detector's sensitivity to noise, (figure 4.22 shows the effect of size and smoothing on edge detection percentage for the image shown in figure 4.21). As can be seen from figure 4.22, the edge rate calculated from the total area dramatically decreases with the size of the mask.



**Figure 4.22: Percentage of detected edges (cracks and non-cracks) vs size of the mask for a sample of image showed in figure**

Due to the availability of various types of roughness of textures in the real world, five filters with the same size window (7\*7) were designed, from very smooth to very rough. It can be noticed from figure 4.23 that the sensitivity increases with the increasing roughness of texture. The manipulation of pixels for a very smooth texture is slight but increases when increasing the roughness of the texture. After defining different filters, it was necessary to test the image for roughness. This was important to decide which filter is to be used for each tested image.



**Figure 4.23: Designed filters for different texture roughness**

#### 4.8.3.3 Modelling the roughness of the texture using 2D image analysis

The image filtration process is mainly based on the type of texture and roughness degree as explained in the previous section. A method was developed for modelling the roughness of the texture. This method, which can be named as Objects Repetition (OR) included subjecting the original image to different operations in order to calculate the number of all objects, see figure 4.24. This process can be summarized as follows:

- a- Range filtering operation on the original image. Range filtering operation is carried out on a 3-by-3 neighbourhood around the corresponding pixel in the input image in order to work out the range (maximum value - minimum value) as the output pixel. The range filtering is usually used to identify raised up aggregates and texture features in the image texture samples.
- b- Edge operation on the range filtered image. The Canny edge process is utilised in order to locate the edges of possible objects on the texture such as small features and raised up aggregates.
- c- Morphological operation for object geometry definition. This action includes closing the edges and isolating the closed objects for area calculation.
- d- Calculating the number of total objects with various sizes using the prop region tool used extensively in Matlab.

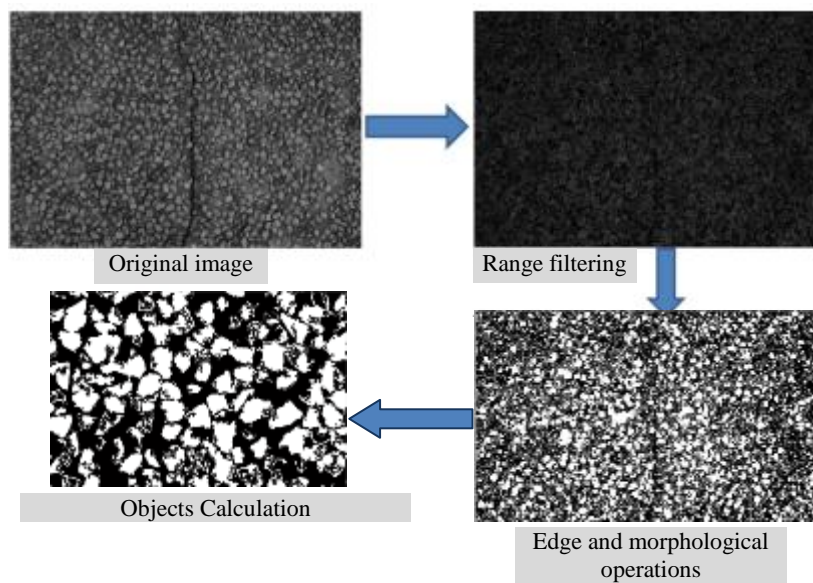


Figure 4.24: Texture modelling based on small object repetition

#### 4.8.4 Image processing or segmentation (edge detection)

After smoothing the image and eliminating the noise using the designed filter (depending on the texture type), the next process was to identify and locate sharp discontinuities in images or calculate significant changes in image intensities. Edges are typically occurring on the boundary between two different regions such as a crack object and the background texture. For the purpose of edge detection, there are some available algorithms such as Canny, Roberts, Laplacian of Gaussian (LoG), Sobel, and Prewitt edge detection (MathWorks, 2014).

In order to test the best edge detection operator for the road pavement surface, trials were carried out on different images, table 4.4. The optimal detector must minimize the probability of false-positives (detecting non-crack edges caused by noise), as well as that of false-negatives (missing real cracks). The results were all carried out using the default threshold values without any changes, which is useful for automating the process without human intervention.

Table 4.4: Results of Crack and non-Crack Edges

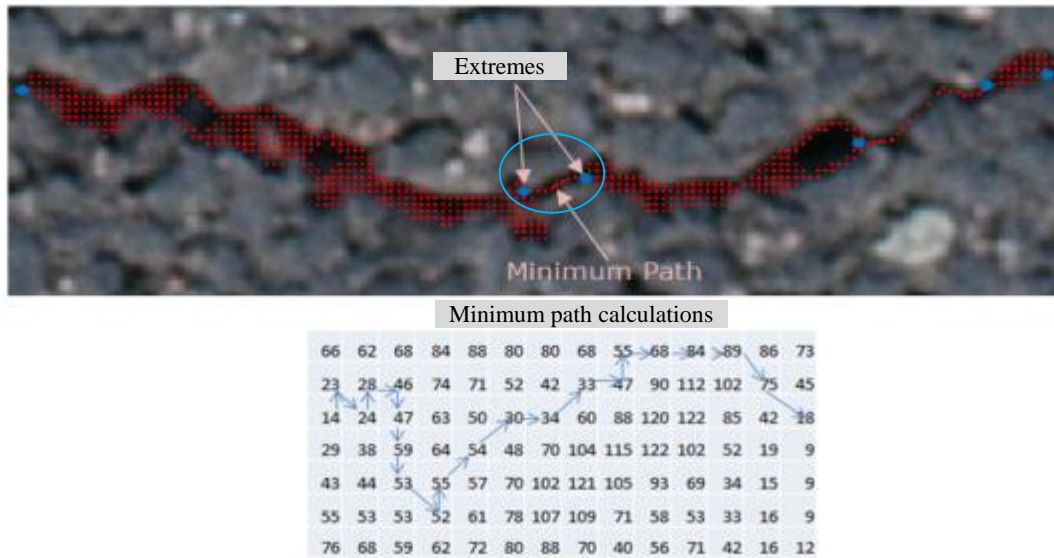
Operator	False-Positive %	False-Negative %	Crack Detected
Canny	7.887	0.0334	0.2652
LoG	8.0788	0.0356	0.263
Roberts	6.5632	0.0211	0.2775
Prewitt	6.6749	0.0234	0.2752
Sobel	5.8304	0.0225	0.2761

Each of these algorithms was tested with different image types and different texture sizes. The results shown in table 4.4 are the mean results of the different images presented in appendix D. It can be noticed that Sobel, Prewitt and Roberts detected about the same number of crack features, whereas, Canny and LoG detected slightly less crack features. LoG is more sensitive to noise as it differentiates twice, and that is apparent from the slightly higher percentage of false positives. The lowest false-positive rate was for Roberts, and the second lowest was for Sobel. However, the false-negative percentage of Sobel is more than that for Roberts. Therefore, the Roberts edge operator was selected and decided to be used in this project. As long as the crack candidate was detected, the next stage is to connect the crack objects and remove any non-crack objects, which was achieved by developing a Matlab subroutine as explained in the next section.

### 4.8.5 Image post-processing

The post-processing consists of connecting disconnected crack objects, and false-positive crack eliminating process and real crack measurement. For the connection stage, the required processes are as follows:

- 1- The end points or extremes of crack and non-crack objects are defined.
- 2- The minimum average cost path intensity between the nearest points is calculated and then compared with the mean value of the two connecting objects, figure 4.25, (for details about minimum average cost path intensity, refer to appendix E). If the average cost path intensity was equal to or less than the average intensity value of the two connected areas the connection path was treated as a single crack, otherwise it was treated as disconnected cracks.



**Figure 4.25: Sample of disconnected objects and the minimum path**

In the false-positive crack eliminating stage, some statistic characteristics such as length, area, numbers of pixels, orientation were calculated for every connective area. Connective areas with fewer pixels were treated as fake-crack objects and were eliminated. After that, the crack objects underwent the operation of removal of crack objects due to shadow and other unusual things such as manhole cover objects; the latter having intensity of more than crack intensity. In shadow false removal, the output shadow mask from shadow removal algorithm (which is described in section 4.8.3.1) was used. The final result is obtained with only real cracks left and a small percentage of false-positives. More details about the results of false-positive and real cracks are available in the results and analysis section 4.9.



## 4.9 Results and analysis

This section shows the results from the author's developed crack detection algorithms and their operations used in each step. The sections are outlined as follows: first start with presenting the results from each algorithm step then followed by the results of crack detection from various image sources and finally presenting the results from video extracted frames.

### 4.9.1 The Author's Developed Crack detection algorithm results evaluation

#### 4.9.1.1 Motion estimation for frame extraction

As previously explained, the aim of a motion estimation algorithm is to extract the frames that share about 20% with the next successive frame, which is achieved by defining the distance shift in pixels. The objective was to reduce the number of frames dramatically in order to save processing time. The developed method was described in section 4.8.1.1. In the current section, the results from a set of video frames will be presented.

As discussed previously in section 4.3, the Canon 5D camera was mounted on the top of NGI's van for the purpose of road surface data collection. Sufficient data were collected for the road surface using the video mode. The total numbers of frames were presented on table 4.1. The aim of motion estimation was to reduce the total number of frames by only including the frames that share the required area with successive frames which achieved as can be seen from figure 4.26.

		Images	Scale	Rotation	Matrix	
					$\zeta Rx, -\zeta Ry$	$\zeta Ry, \zeta Rx$
					$-\Delta x$	$\Delta y$
1494	1498	1494-1494	1	0.00003	1.000016	-6.02E-07
1495	1499	1494-1495	0.9913	-0.16850	-6.01E-07	1.000016
1496	1500	1494-1496	0.9824	-0.09650	-0.01272	-0.01175
1497	1501	1494-1497	0.9726	-0.66520	1.008813	-2.97E-03
		1494-1495	0.9913	-0.16850	2.97E-05	1.008813
		1494-1496	0.9824	-0.09650	-21.5845	106.4121
		1494-1497	0.9726	-0.66520	1.017887	-1.71E-03
		1494-1498	0.9659	-0.33040	1.71E-05	1.017887
		1494-1499	0.9496	-1.86800	-37.1605	210.3687
		1494-1500	0.9464	-0.12430	1.028073	-1.19E-02
		1494-1501	0.9507	-3.79680	1.19E-02	1.028073
		1495-1496	0.9824	-0.09650	-59.9349	323.7255
		1495-1497	0.9726	-0.66520	1.035292	-5.97E-03
		1495-1498	0.9659	-0.33040	5.97E-03	1.035292
		1495-1499	0.9496	-1.86800	-73.9111	421.2205
		1495-1500	0.9464	-0.12430	1.052526	-3.43E-02
		1495-1501	0.9507	-3.79680	3.43E-02	1.052526
		1496-1497	0.9726	-0.66520	-108.076	555.6115
		1496-1498	0.9659	-0.33040	1.056581	-2.29E-03
		1496-1499	0.9496	-1.86800	2.29E-03	1.056581
		1496-1500	0.9464	-0.12430	-111.298	630.2368
		1496-1501	0.9507	-3.79680	1.049519	-6.96E-02
		1497-1498	0.9659	-0.33040	6.96E-02	1.049519
		1497-1499	0.9496	-1.86800	-117.075	807.8513

Figure 4.26: Motion estimation between samples of frames (1494 to 1501)

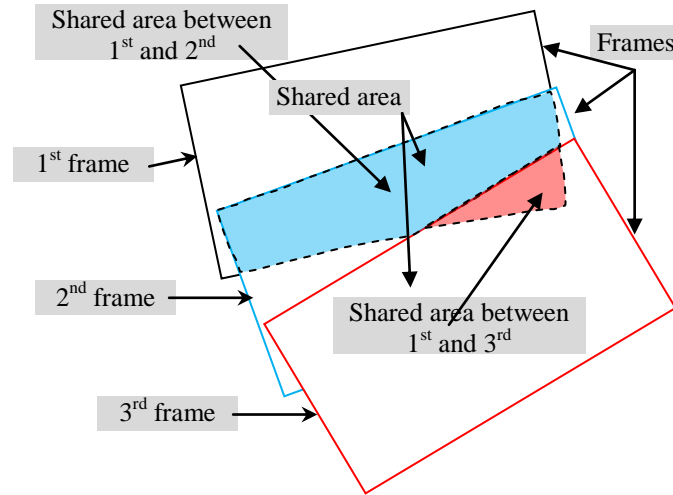
Figure 4.26 shows a sample of frames for the video file named (MVI\_4517). As can be seen, the results of each frame with the next successive frames are calculated using the transformation matrix. From this matrix, the scale and rotation between two frames can be recovered. The distance of the shift that matches with the criteria is between frames 1494 and 1501. Therefore, all frames in between are excluded as they are effectively repeating the same scene. After a frame was selected, it will again be compared with the next following frames to include the required frame and exclude others and so on, figure 4.27.

1501	1505	Images	Scale	Rotation	Matrix
1502	1506	1501-1501	0.9884	-0.18300	1.011755 -3.23E-03 3.23E-03 1.011755 -21.3059 104.3049
1503	1507	1501-1502	0.9665	-0.02000	1.034645 -3.61E-04 3.61E-04 1.034645 -38.2351 202.5105
1504	1508	1501-1503	0.9496	-0.15780	1.05307 -2.90E-03 2.90E-03 1.05307 -57.8672 308.6316
		1501-1504	0.9322	-0.08820	1.072776 -1.65E-03 1.65E-03 1.072776 -78.051 415.5262
		1501-1505	0.9167	-0.35260	1.09087 -6.71E-03 6.71E-03 1.09087 -101.033 524.3494
		1501-1506	0.9442	-0.31050	1.059092 -5.74E-03 5.74E-03 1.059092 -112.917 635.707
		1501-1507	0.95231	-0.42520	1.054093 -4.54E-03 4.54E-03 1.054093 142.9031 973.856

**Figure 4.27: Motion estimation between samples of frames (1501 to 1507)**

The video file MVI\_4517 was processed and the number of frames are reduced dramatically from a total of 8760 frames to only 1271 frames (the data reduces by about 85 %). The frames extracted and excluded were from the starting of the path; traffic lights control zone, priorities for other cars, and at the end of the road when the car had fully stopped. For example on the start of video file MVI\_4517, a total of 840 frames were excluded because the car had not yet moved. Another example was when the car was fully stopped in a traffic control zone, wherein a total of 1050 frames was excluded. All included frames from the used video file were visually checked to see if any errors occurred and only two out of 1271 frames were found to be erroneously included that due to wrong matching or mismatching. During the processing of different video files, some problems occurred, when the algorithm was not working efficiently. For example, when there was a sharp turn of the van (and the camera), the recovered scale is equal to zero because frames are sharing limited area, figure 4.28. To overcome this problem, the distance shift for this case was decreased to 500 pixels instead of 800 pixels.





**Figure 4.28: Frames sharing area at sharp turn locations.**

The results from processing of all the video files showed the success of the developed methodology for motion estimation, based on SURF matching algorithm. All the results were checked visually to identify any erroneously included frames (which were not found). The percentage of the included extracted frames differed from one video file to the other. The percentage difference was dependent on whether the car had fully stopped due to traffic. Table 4.5 shows the results of excluded and included frames from each video file used in this study.

**Table 4.5: Number of included and excluded frames using motion estimation**

Video Files	Total number of frames	Included frames	Excluded frames	Percentage Excluded
MVI_4512	15280	3612	11668	76%
MVI_4513	3360	697	2663	79%
MVI_4514	23910	4585	19325	80%
MVI_4517	8760	1271	7489	85%
MVI_4522	19783	3198	16585	83%
MVI_4523	27029	4794	22235	82%
MVI_4539	19800	3211	16589	83%
MVI_4540	20700	3912	15740	85%
MVI_4541	9000	1526	7474	83%
MVI_4542	14340	1799	12541	87%
MVI_4543	9031	1109	7922	87%

As can be noticed from table 4.5, the percentage of excluded frames in some cases is 87%, as the stopping time is bigger than other cases. After repeated frames were excluded, the next stage was to present geotag results of the image frames by using the GPS log path in order to relate the position (Latitude and Longitude) to each particular frame.

#### 4.9.1.2 Results of frames Geotagging

The methodology of geotagging the GPS position to single frame images can be performed using available ‘free tools ‘ such as Geosetter, ExifTool, and PhotoME (Duttke, 2009; Schmidt, 2011; Harvey, 2013). However, geotagging the extracted video frames is not available in these tagging tools due to the fact that they are working based on time synchronisation between the GPS log and the time the image was captured. All ExifTool editor tools only read the seconds but not the fraction of the seconds required for videos. Therefore, a method was designed and described in section 4.8.1.2 in order to assign the position of the GPS log to the extracted frames. This has been tested on all video files mentioned in section 4.5, table 4.1. All road junction points were analysed and presented, as it was difficult to present all the data frames, see figure 4.29. The images of junction points were all projected into Google Earth, and the position of image centre was compared to the image centre of the junction point and the differences were reported. Figure 4.30 shows all junction points projected to Google Earth with the difference in metres for video file MVI\_4512. For all other video file results and their graphical presentation, please refer to appendix F. All the errors from junction points were reported and presented in table 4.6.

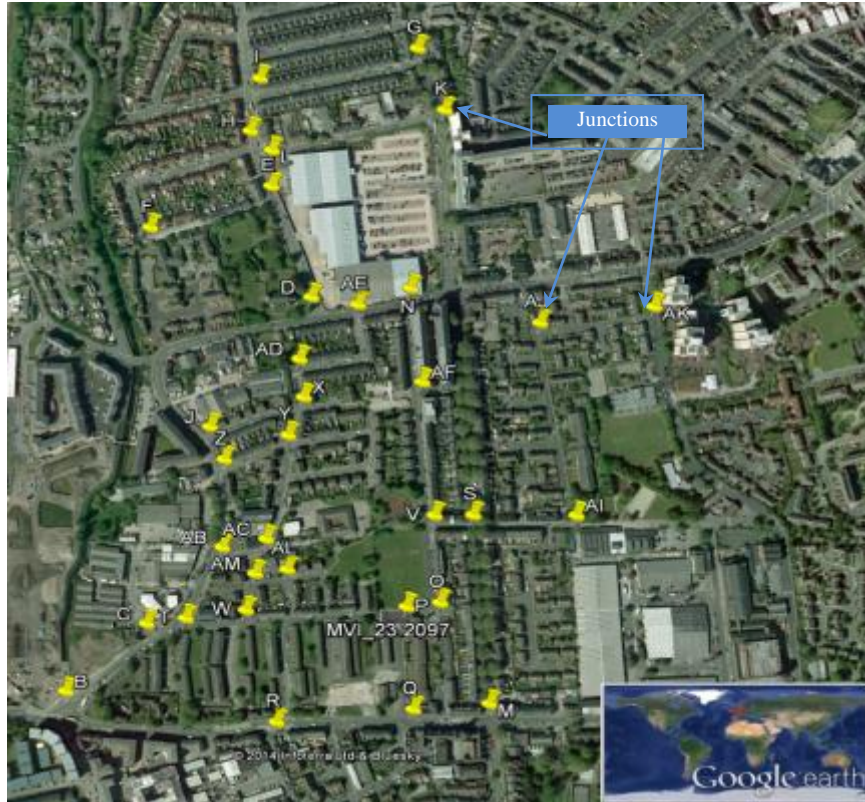


Figure 4.29: Road junction points used to check the accuracy of Geotagging

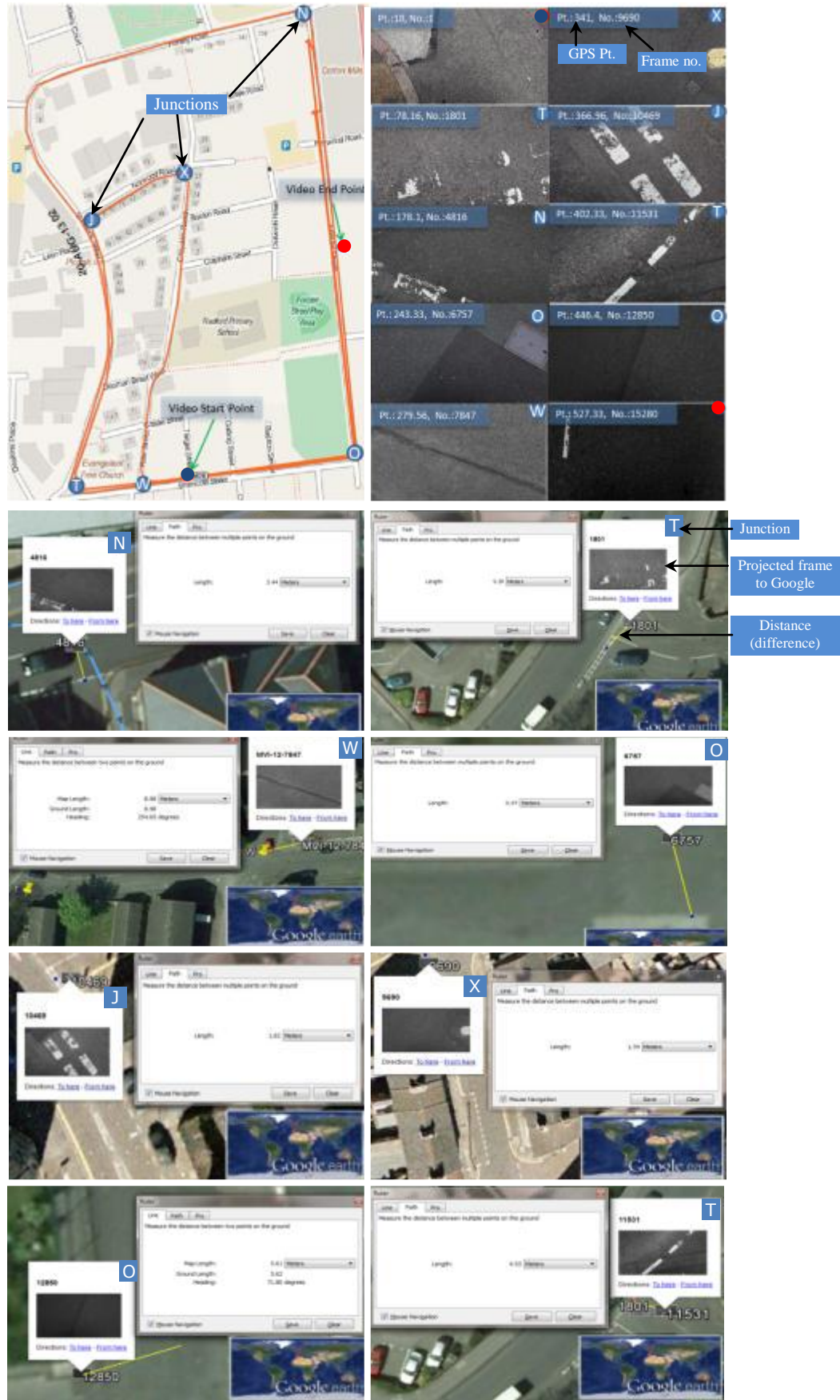


Figure 4.30: MVI-4512 video turn points frame and geo-tagging differences

Table 4.6: Accuracy of GPS frames at turning points

Turn Points	Video Files (Difference of frame position with Google Earth in meters)											
	MVI_4512		MVI_4514		MVI_4522		MVI_4539		MVI_4541		MVI_4543	
	MVI_4513		MVI_4517		MVI_4523		MVI_4540		MVI_4542			
	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In
A	----	----	9.92	10.20	----	----	----	----	----	----	----	----
B	----	----	----	12.34	----	4.38	----	----	----	----	----	----
C	----	----	----	----	3.85	3.75	----	----	----	----	----	----
D	----	----	----	----	----	24.29	3.08	4.79	5.78	3.48	----	----
E	----	----	----	----	----	----	----	----	----	1.79	----	----
F	----	----	----	----	----	----	----	----	----	2.40	----	----
G	----	----	----	----	----	----	----	----	4.87	----	----	----
H	----	----	----	----	----	----	----	----	5.77	----	----	----
I	----	----	----	----	----	----	----	----	----	2.18	----	----
J	----	----	----	----	----	----	----	----	----	15.32	----	----
K	----	----	----	----	----	----	----	----	----	4.21	----	----
L	----	----	----	----	----	----	----	----	2.65	----	----	----
M	----	----	----	5.66	----	----	----	----	----	----	----	----
N	3.44	----	----	3.56	8.14	5.22	4.39	3.46	----	----	----	----
O	3.37	5.61	4.74	----	----	9.12	8.09	----	----	----	11.31	2.82
	7.71	----	3.27									
P	----	----	----	5.13	1.83	----	----	----	----	----	----	----
			4.58									
Q	----	----	4.97	----	3.11	----	----	----	----	----	----	----
			4.13									
R	----	----	----	----	1.16	1.93	----	----	----	----	----	----
S	----	----	----	11.56	----	----	----	----	----	----	----	----
T	5.30	4.53	----	3.37	----	----	----	----	----	----	----	----
U	----	----	----	11.03	----	----	----	----	----	----	----	----
V	----	----	4.95	----	----	----	2.87	8.63	----	----	----	----
W	----	----	5.72	----	7.83	----	----	----	----	----	----	----
X	----	----	----	1.86	----	----	----	----	4.00	----	----	----
Y	1.34	----	----	10.57	4.62	----	----	----	----	----	----	----
Z	1.62	----	6.96	----	----	5.04	----	----	----	----	----	----
AB	----	----	----	----	----	7.00	----	----	22.15	----	----	----
AC	----	----	----	----	----	2.21	----	----	25.83	----	----	----
AD	----	----	----	----	----	5.46	----	----	----	----	----	----
AE	----	----	----	----	3.09	----	----	----	----	----	----	----
AF	----	----	----	----	----	8.56	----	----	----	----	----	----
AG	----	----	----	----	3.58	----	----	----	----	----	----	----
AH	----	----	----	----	9.75	----	----	----	----	----	----	----
AI	----	----	----	----	----	----	5.16	4.47	----	----	----	----
AJ	----	----	----	----	----	----	2.79	2.54	----	----	----	----
AK	----	----	----	----	----	----	2.98	2.77	----	----	----	----
AL	----	----	----	----	----	----	----	----	21.20	----	----	----
AM	----	----	----	----	----	----	----	----	19.61	----	----	----

Out: from Local to Main Street (Double lines)

In: from Main to Local (one line)

Table 4.6 shows that the largest error recorded was about 26 m for junction point AC for video file MVI\_4541. In total, 72 out of 78 positions were less than the claimed positional error of 15 m, and the RMSE percentage was about 92% which is 3% percent less than the claimed RMSE percentage. These differences might be due to the fact that the author was not aiming at measuring the absolute error of position rather than depending on Google Earth. Although the claimed accuracy of hand-held GPS standalone single-frequency code positioning is less than 15 m; however, errors of as big as 26 m were reported which might be attributed to one of the following reasons:

- The urban environment which can obstruct some satellites and full geometry cannot be seen.
- Multipath errors
- The errors with reference to Google map which might also contribute some sources of errors.
- The shift between the synchronized times of GPS and camera.
- For start points only, when the GPS receiver has not found enough satellites yet, which can be overcome by waiting as long as possible to allow the GPS to log all available satellites. Figure 4.31 shows the drift of the GPS path due to the starting of the GPS receiver.

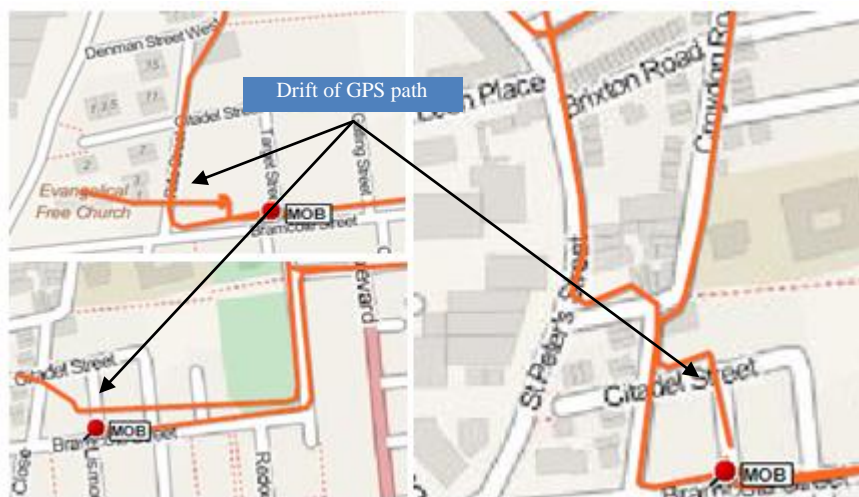


Figure 4.31: Drift of GPS path due to the start of the GPS receiver



#### ***4.9.1.3 Statistical testing for problems identification***

After the frames were geotagged in the previous step, all frames were subjected to a further process in order to identify whether they have any distresses or objects (such as paints on road, cracks, potholes, and manhole covers) in the image. This operation aimed to roughly identify the existence of problems through the use of histogram curve fitting for residuals calculation (Gaussian fitted curve). The method includes the statistical testing for identification of the problems (see section 4.8.2 for more details). The idea of this methodology was based on identifying the texture differences between any sources of problems and the image background.

This method was tested and applied on a group of images (extracted from published papers, downloaded from the web, taken with Nikon D100 camera, from the LRIS system) as shown in figure 4.32 and the results of statistical testing presented on table 4.7. As can be noticed from the table, the p-value of most images was equal to or less than 0.9, and %RMSE was greater than 0.1% which indicated that the image has a problem existence and will be passed to the next stage for further analysis. Hence, considering that only 6 images were selected erroneously, the correct rate of problem identification was found to be about 87% for this group of images. Out of 45 images tested, the resulted identification showed that 6 were selected erroneously as problem free, 6 selected correctly as problem free, 33 were selected with existing problems, while none were erroneously selected with problem existence.

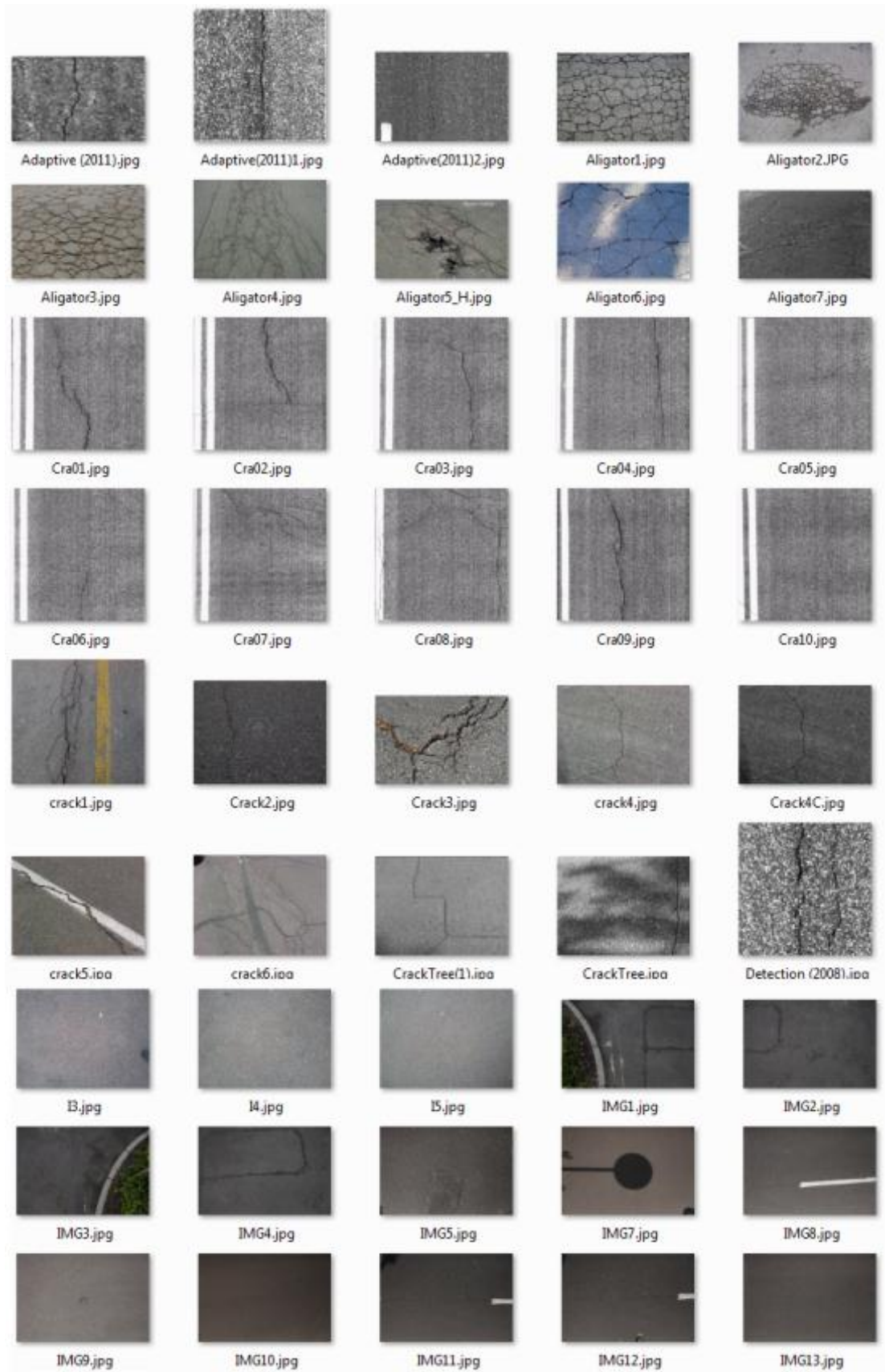


Figure 4.32: Sample of images for testing statistically the problem identification

Table 4.7: Results of statistical testing of comparing data and fitted curve

ImageID	% RMSE	RMSE	R <sup>2</sup>	P-value
Adaptive-2011	0.036397	126.7836	0.99611	0.85536
Adaptive(2011)1	0.13221	725.9363	0.85162	0.36695
Adaptive(2011)2	0.092415	856.759	0.97061	0.57252
Aligator1	0.07381	706.8004	0.96412	0.43955
Aligator2	0.10147	1948.283	0.92805	0.31417
Aligator3	0.14838	169.1512	0.93748	0.40179
Aligator4	0.12382	62.2133	0.97505	0.87198
Aligator5_H	0.082767	211.4274	0.97929	0.52315
Aligator6	0.15731	253.9392	0.93635	0.16733
Aligator7	0.095692	1202.608	0.96273	0.39646
Cra01	0.52097	2700.712	0.26836	0.0152
Cra02	0.5229	2710.694	0.23343	0.00125
Cra03	0.41117	2131.494	0.4121	0.24549
Cra04	0.32791	1699.871	0.53886	0.27893
Cra05	0.28537	1481.423	0.6115	0.29455
Cra06	0.29923	1553.341	0.58983	0.32615
Cra07	0.33115	1719.089	0.53113	0.33001
Cra08	0.1915	994.0931	0.81	0.39893
Cra09	0.22158	1150.286	0.77373	0.3924
Cra10	0.54431	2825.596	-0.07836	0.0024
Crack2	0.11984	150.0408	0.96358	0.51287
Crack3	0.069422	1037.161	0.96453	0.56024
Crack4C	0.053929	74.8436	0.99346	0.76868
CrackTree(1)	0.15083	190.9381	0.97388	0.64024
CrackTree	0.14046	848.8861	0.87094	0.81846
Detection-2008	0.10889	559.8046	0.86161	0.26949
I3	0.063072	3100.109	0.99228	0.90658
I4	0.049009	2408.904	0.9968	0.96481
I5	0.1146	5141.094	0.98557	0.9742
IMG1	0.18009	3407.844	0.92353	0.1669
IMG10	0.029733	562.625	0.99857	0.97755
IMG11	0.10826	2048.522	0.97516	0.94857
IMG12	0.086479	1636.399	0.98297	0.97998
IMG13	0.033703	637.7458	0.99807	0.9198
IMG2	0.034275	882.7292	0.99838	0.94354
IMG3	0.16913	4355.956	0.93387	0.29671
IMG4	0.041826	791.4465	0.99777	0.90116
IMG5	0.023359	442.0196	0.9986	0.97498
IMG7	0.18685	3535.707	0.93095	0.40946
IMG8	0.12331	2333.343	0.96695	0.86645
IMG9	0.065359	1236.755	0.99068	0.8759
crack1	0.020142	68.0405	0.9989	0.98259
crack4	0.051228	71.0948	0.99379	0.76825
crack5	0.09213	70.7555	0.98136	0.33878
crack6	0.15062	116.1585	0.97697	0.67507



As stated previously, from testing the above group of images, it is clear that the method has a success rate of about 87% in identifying the existing problem images. So, it was applied on all the included the frames from motion estimation for frame exclusion stage, refer to table 4.5. The results of this operation (Gaussian fitted curve with statistical t-testing) proved that it was possible to identify frames with existing problems with success rate of about 73% to 87%, see table 4.8. This method was ‘not so successful’ in very dark intensities, very bright intensities, ravelling texture, with blended cracks on the rough texture or when the crack was very thin. It was also found that most of potholes that do not show texture differences from the background were not identified as existing problems by this method. However, the method of problem identification was significantly useful in reducing the processing time and proved to be time saving by more than 73%.

Table 4.8: Results of problem identification for video files

<b>Video File ID</b>		<b>True</b>	<b>False</b>	<b>~ % of correct identification</b>
MVI_4512	Problem Free	1228	270	78%
	Problem Exist	1589	573	
MVI_4513	Problem Free	184	23	85%
	Problem Exist	411	81	
MVI_4514	Problem Free	1185	356	80%
	Problem Exist	863	183	
MVI_4517	Problem Free	656	145	78%
	Problem Exist	368	138	
MVI_4522	Problem Free	1379	323	84%
	Problem Exist	1310	187	
MVI_4523	Problem Free	3735	687	82%
	Problem Exist	1198	173	
MVI_4539	Problem Free	1736	404	83%
	Problem Exist	926	154	
MVI_4540	Problem Free	2878	502	84%
	Problem Exist	430	94	
MVI_4541	Problem Free	1109	179	87%
	Problem Exist	426	63	
MVI_4542	Problem Free	1151	258	83%
	Problem Exist	538	71	
MVI_4543	Problem Free	54	21	73%
	Problem Exist	763	271	

#### 4.9.1.4 Testing road roughness modelling using images

As previously discussed in the procedures section, the principle of crack detection is based on image filtering. Therefore, it was necessary to design a proper filter based on the texture roughness of the tested image, see section 4.8.3.3. The method of object repetition was examined in a number of known texture depth images previously used in chapter 3 of this project. Ten texture samples from different areas around the Jubilee campus were investigated for roughness measurement using the author's developed method. From these, 120 images at different ranges were obtained, which could be divided into five groups based on the degree of roughness (very smooth to very rough), (see section 3.5 and figure 3.4 for a sample of these images) and were used later to measure the success rate of the proposed method of object repetition.

The object repetition method was investigated using the images collected for roughness measurement (Test01 to Test10), see section 3.5 and figure 4.4. For each test, twelve images at different ranges (20 to 130 cm) were obtained with the camera perpendicular to the road surface texture. The texture depth was calculated previously for this set of tests, see section 3.8. The camera used for this data collection was the Nikon D200 with its optical axis perpendicular to the texture. Images at various ranges (20-130 cm) were captured and cropped to be of a 20 cm range, see figure 4.33. Cropping images of different ranges makes the investigation more reliable and comparable, as there are more objects observed as the area increases which will make the results less dependable.



Figure 4.33: Sample at range 130cm cropped to be same size of 20cm (Test01)

The method of object repetition was run on the total of 120 images (test01 to test10), and the results are presented on figure 4.34. The figure shows the relationship between the degree of roughness or texture depth and the number of repeated objects. The maximum number of repeated objects was recorded for the smoothest texture or the minimum depth test (test03, 511  $\mu\text{m}$ ). Also, the minimum number of objects was recorded for roughest texture or the maximum texture depth (test05, 3724  $\mu\text{m}$ ), see section 3.8.3.1 for more details about the calculated texture depth.

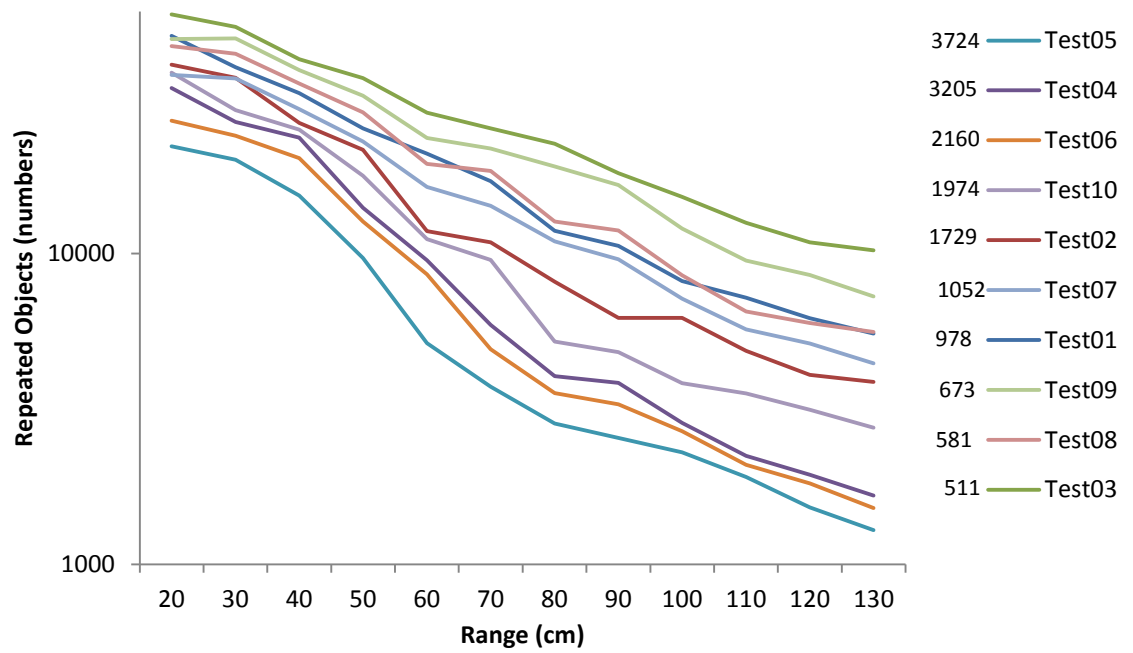


Figure 4.34: Results of object repetition vs measured depth at different ranges

As figure 4.34 shows, the object repetition method proved that it is dependent on the distance or range of the camera. All samples showed that object repetition decreases as the camera range increases. The decrease in the object repetition is due to the fact that the pixel size on the ground decreases as the camera becomes further from the surface texture. It is worth to mention that object repetition for all texture samples at different ranges was calibrated to the correct pixel size on the ground. For example, using the Nikon D200 over 20cm can produce a GSD of about 0.044mm, refer to table 4.8. So, any value of object repetition at a different range for all samples was corrected to be compatible with the sample on the ground, in order for the investigation and comparison of different textures to be more reliable.

As figure 4.34 shows, there is some degree of correlation between object repetition and texture depth. However, some tests (test09 and test06) showed that the results from object repetition are not correlated with texture depth. To find out how much correlation existed between both methods a linear correlation was drawn for all tests between texture depths and objects repetition at various ranges (20-130) cm and some samples are presented in figure 4.35.

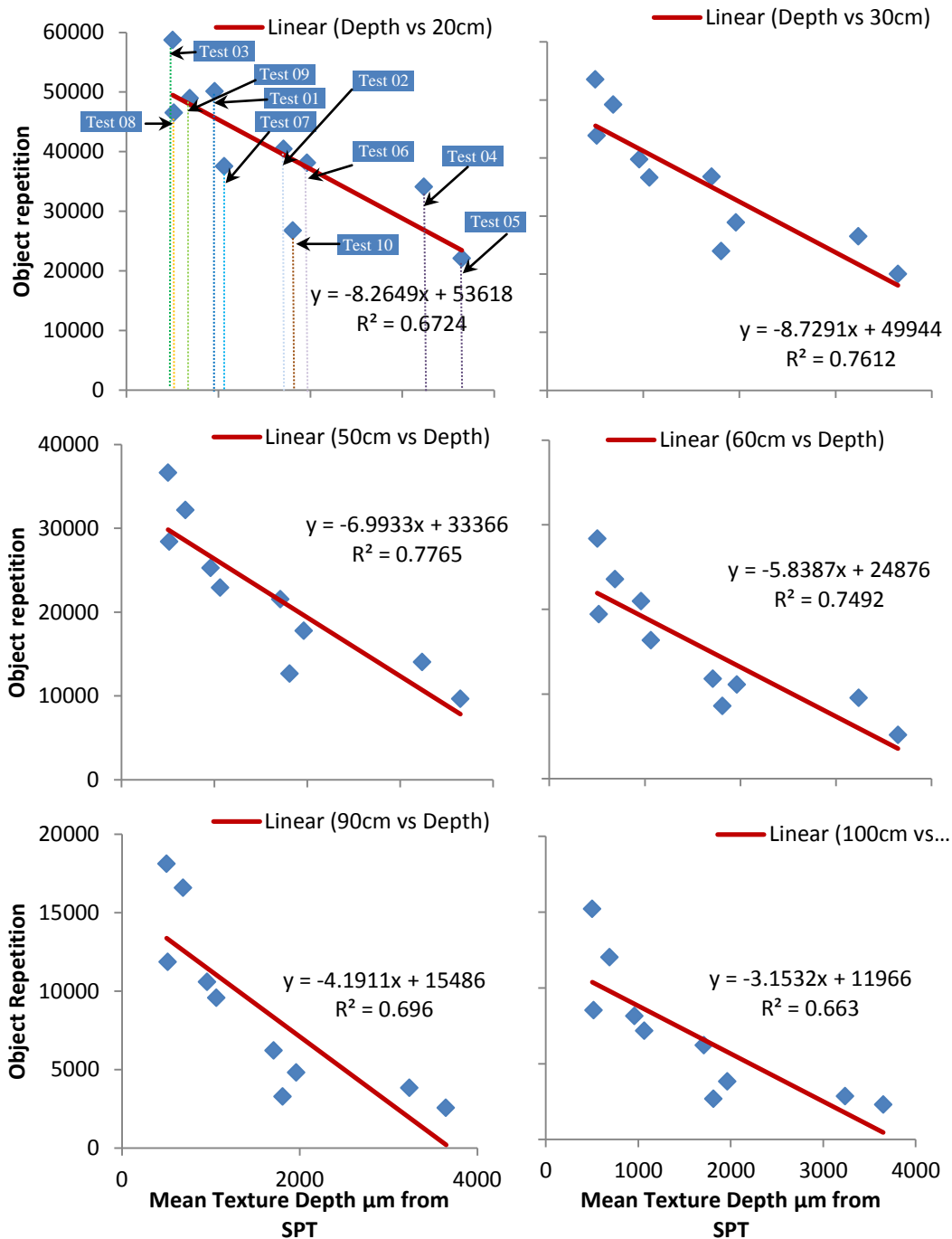


Figure 4.35: Samples of linear correlation between objects repetition and texture depth

As shown in figure 4.35, the minimum correlation was found to be 0.66 which was recorded for the range 100cm while the maximum correlation was found to be 0.77 which was recorded for the range 50cm. The average correlation from different ranges and samples was found to be about 0.67. The average correlation showed that there is some evidence of a relationship between texture depth and object repetition. As an example, figure 4.34 shows that by using SPT method for calculated depth, texture roughness was highest for test05, then 04 followed by 06. However, by using the repeated object method, the order of texture roughness changed into test 05, 06 then 04 (highest to lowest). This exaggeration of roughness on test 06 might be due to the nature of the feature and aggregate distribution as can be noticed from figure 4.36.

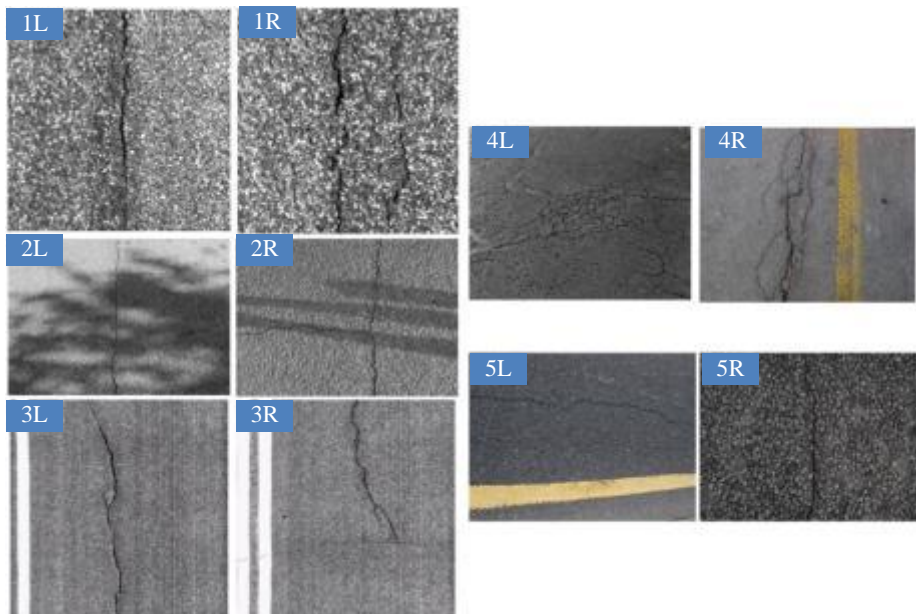


**Figure 4.36: Sample of texture (test04 left and test06 right)**

Although, texture depth on test 04 was more than that on test 06, the repetition of objects on test 04 were more than test 06 as can be noticed from figure 4.34. The raised of aggregates on the surface occupy more space of image in case of test 06 than that of test04. It is true that no significant correlation was found between the object repetition and texture depth, but object repetition gives a good indication of how smooth or rough the texture is in order to apply the proper filter. The rough texture needs more manipulation than the smooth texture in order to minimize the effect of the raised aggregate edges.

## 4.9.2 Testing various sources of images

In order to investigate the success rate of the author's developed crack detection algorithm, different sources of images (web net, extracted from published papers, low resolution, shadow existence, LRIS and collected from the real world) were examined. The aim of this section is to provide the outcome of testing the algorithm on each of the particular samples. As previously discussed in section 4.3, images were prepared from different sources. A total of ten images were used for investigating the success rate of the proposed algorithm, figure 4.37 shows the images with different lighting condition, roughness, texture's type and crack extent being available in this collection.



**Figure 4.37: Collection of images for evaluation of crack detection algorithm**

For figure 4.37, image sets 1(1L and 1R) and 02 (2L and 2R) were extracted from published papers (Sy et al., 2008; Zou et al., 2012), respectively. Both sets were extracted with low resolution and shadow problem existence. The first set was considered to be a high texture which can easily hide cracks and the second one is where shadows exist in the images. The third set (3L and 3R) of images were extracted with low resolution from a laser imaging systems available at the website of Pavemetrics (2012). The real pixel resolution was 4000 but reduced to 720 pixels due to extraction from video file. The fourth set (4L and 4R) was downloaded from Google images (Wolf, 2014) and contained two images, one has marker existence and the other has acquired obliquely. The last set of images (5L and 5R) was collected from a real road surface data and captured using the Nikon D100 camera.

All samples were processed using the pre-processing and processing stages mentioned in section 4.8. In order to examine the effects of filtering on the texture, all five filtering techniques presented in figure 4.23 were applied; see figure 4.38 for a sample of pre-processing and processing stage results. As figure 4.38 shows, the effect of correct filtering is apparent, as bounded by green boxes. This figure shows how crack pixel candidates and false-positives change with the filtering technique. As an example, the results on the 3R image showed that the correct filter can even detect the thinner crack which is not possible when using different filters. Other examples are in the case of 2L and 3R where nothing was detected when using very rough texture filtration. However, detection of cracks failed in shadow cases even if the correct filter was chosen, as in case 2L.

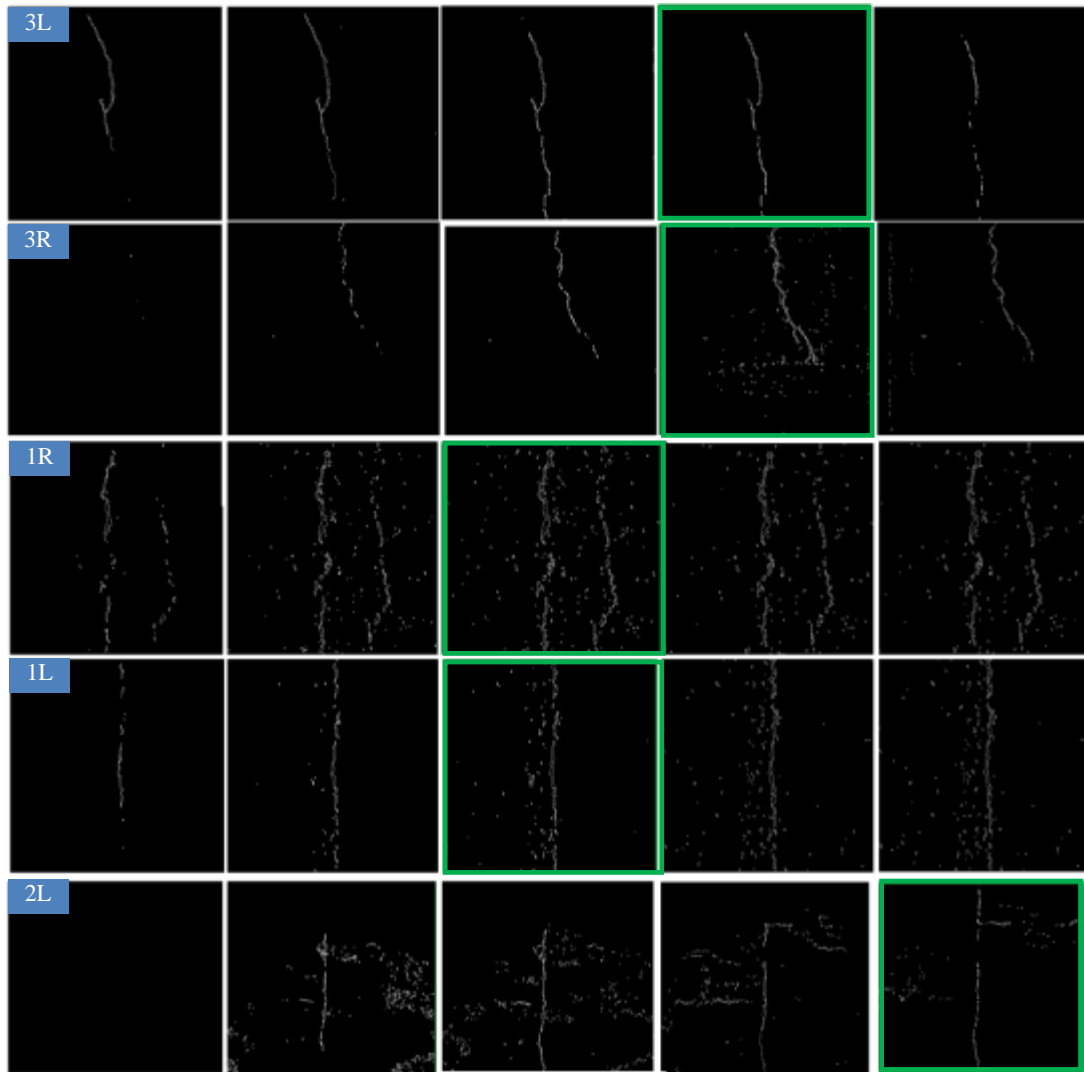
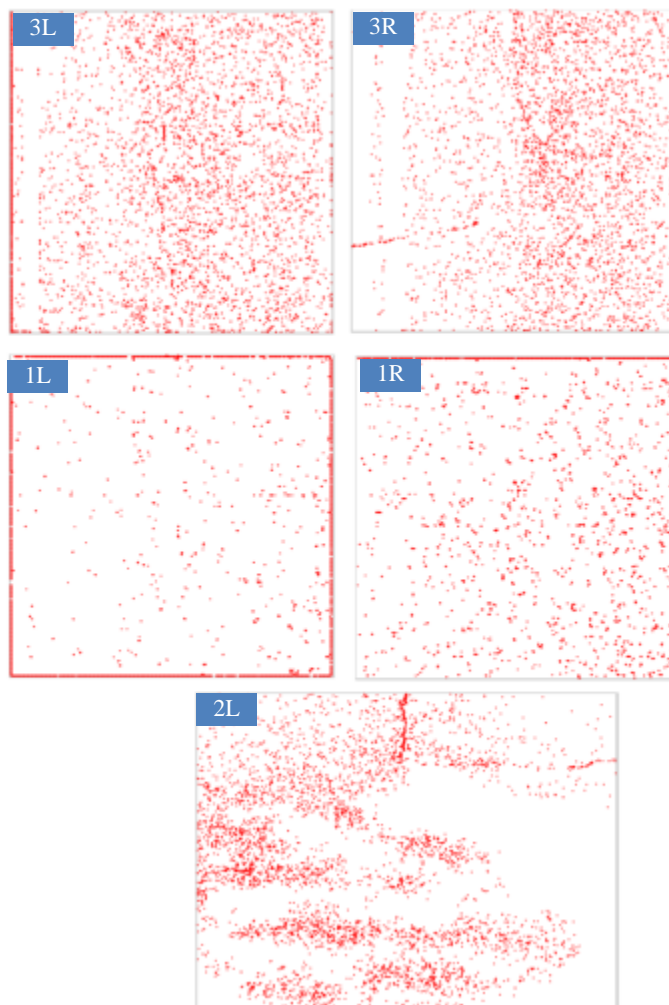


Figure 4.38: Samples of images after pre-processing and processing stages with applying the five filtrations (very rough left to very smooth right), green boxes are correct filters



In the previous paragraph, it was shown how proper filtration can improve the crack detection and reduce noise. The improvement is not only due to image filtration but due to other image enhancement processes applied in the author's developed crack detection algorithm before running the edge operation (which was discussed in section 4.8.3). All samples shown in figure 4.38 were processed again but this time without the pre-processing stage in order to see the effect of image enhancement and pre-the processing stage. Results from this trial showed that the pre-processing stage is very important in identifying and strengthening the crack signal and reducing the noise as can be seen from figure 4.39 by comparing it with figure 4.38. As figure 4.39 shows, no signs of cracks are visible on the images and all the samples are noisy. These results prove the importance of applying a pre-processing step on the images before running the edge operation, which leads to much improvement in identifying cracks and reducing the noise.



**Figure 4.39: Image processing using only edge operation without pre-processing stage**



After the pre-processing and processing stages, the final post-processing stages, it is required to connect crack objects utilizing the developed subroutine and removing the noise (refer to section 4.8.5 for more information). The success rate of the author's developed crack detection algorithm is different from sample to sample depending on many parameters such as degree of roughness, texture type, image resolution, lighting condition and shadow existence. For example, it was not possible to detect the thin cracks on shadow areas due to difficulties of connecting the crack objects using the minimum path method. The path between the crack objects in most shadow cases is inaccessible as sample shown in figure 4.40. The success rate and rate of false-positive crack objects were calculated for the samples shown in figure 4.40. The success rate was calculated by comparing with the real crack measured from CAD tool as a reference and from the algorithm as a data value and presented as a percentage. The success and false-positive rates were found to be 90.4% and 15.4% for the left image and 65.4% and 0% for the right image, respectively.

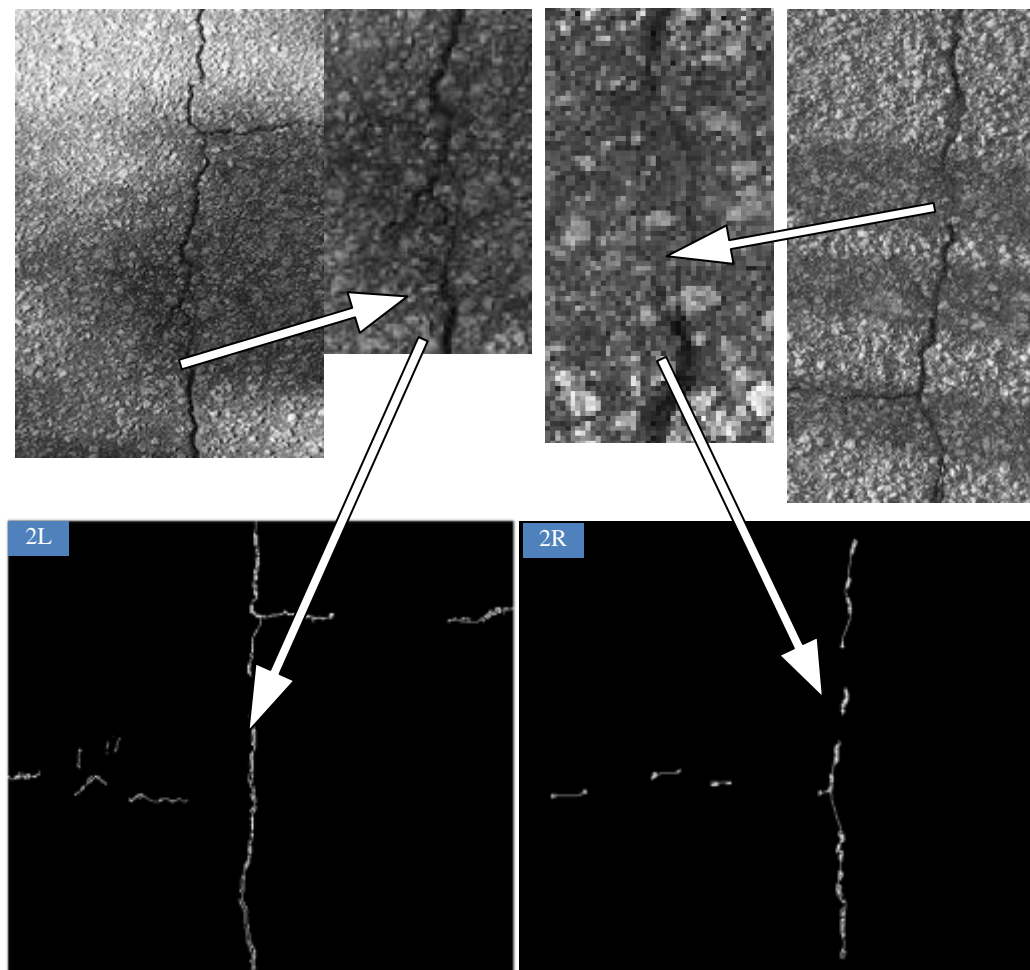
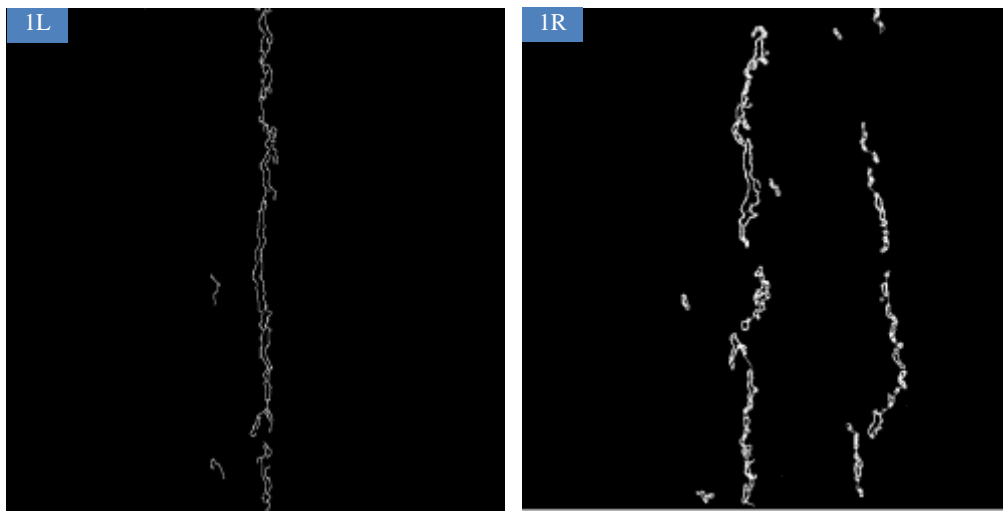


Figure 4.40: Crack detection in shadow existence samples

Although the 1L and 1R images were extracted from published papers with low resolution, the success rate of both samples was high, as can be noticed from figure 4.41. The success rate of crack detection was found to be 98.2% for the left image and 96.5% for the right image, and the false-positive rate was found to be 11.6% and 7.7% for left and right images, respectively. This high-success rate can be attributed to the crack features, like width and objects with low intensity. Thus, such crack objects can still be clearly visible despite the fact that the type of texture can easily hide the crack object. In addition, the low resolution imagery can add more detection challenges as it minimizes crack visibility.



**Figure 4.41: Final results of crack detection on samples extracted from published papers**

LRIS imagery samples were also examined for the success rate of the author's developed crack detection algorithm. It is well known that LRIS can minimize the effects caused by variations in lighting conditions and shadows cast by roadside objects. Another important feature of LRIS is the increase in crack visibility through the formation of shadows in thin cracks by the incident illumination angle of the laser (Pavemetrics, 2012). The success rate of LRIS images 3L and 3R was found to be 88.4% and 58.7 % for left and right images, respectively, with a false-positive rate of 0% and 3.1% for left and right images, respectively, see figure 4.42. As the figure shows, the success rate of the right image sample is low which is due to the fact that the crack was very thin, i.e. the author's developed crack detection algorithm is not capable of detecting cracks that are less than three pixels width, as was the case in the 3R sample.

With regards to set 4 of data, downloaded from Goggle images, the final post-processing results showed that the correct crack rate was around 62% for the left image

and 86% for the right image, figure 4.43. Although the left image was downloaded from the net with the resolution and more than three pixels width of crack all over the image, the dark texture of the background blended the crack which reduced the chance of detection to about 62%. The false-positive rate for both images was found to be 0%.

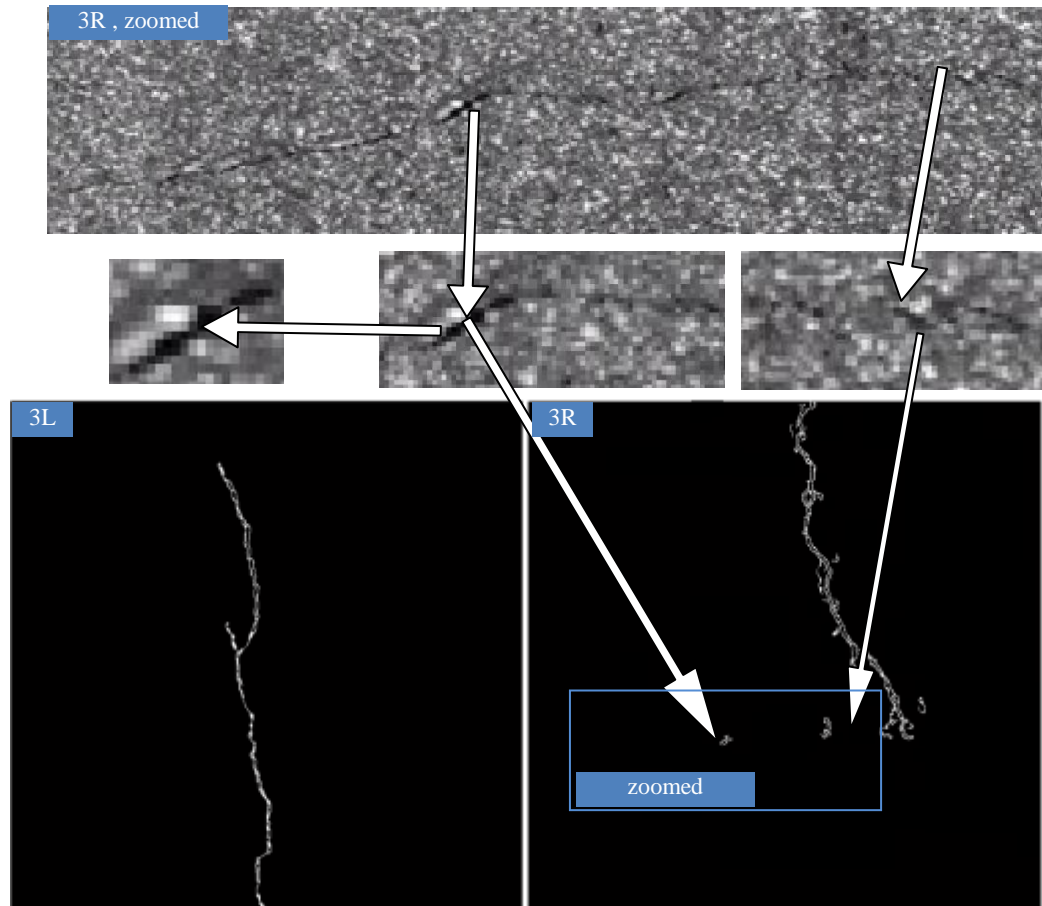


Figure 4.42: Final results of crack detection for LRIS imaging samples

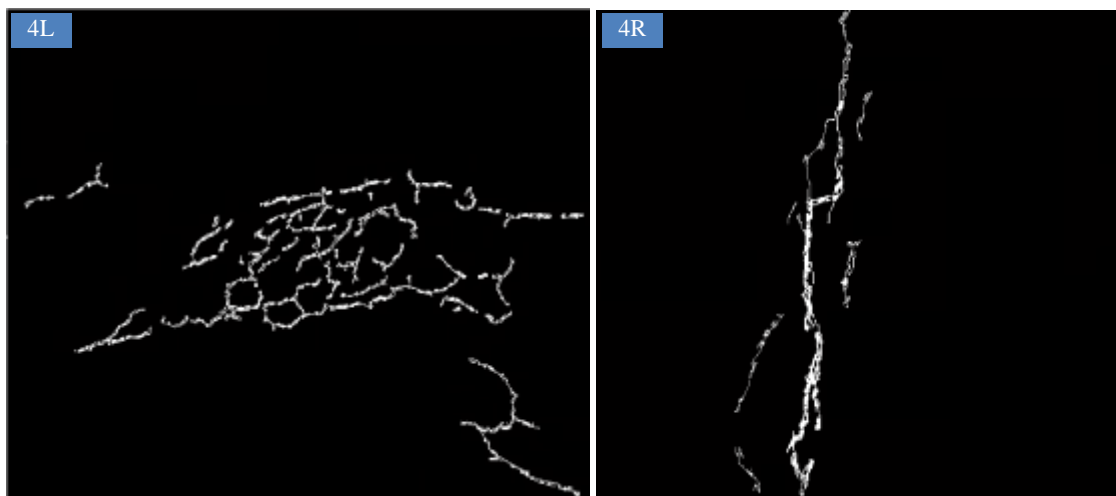
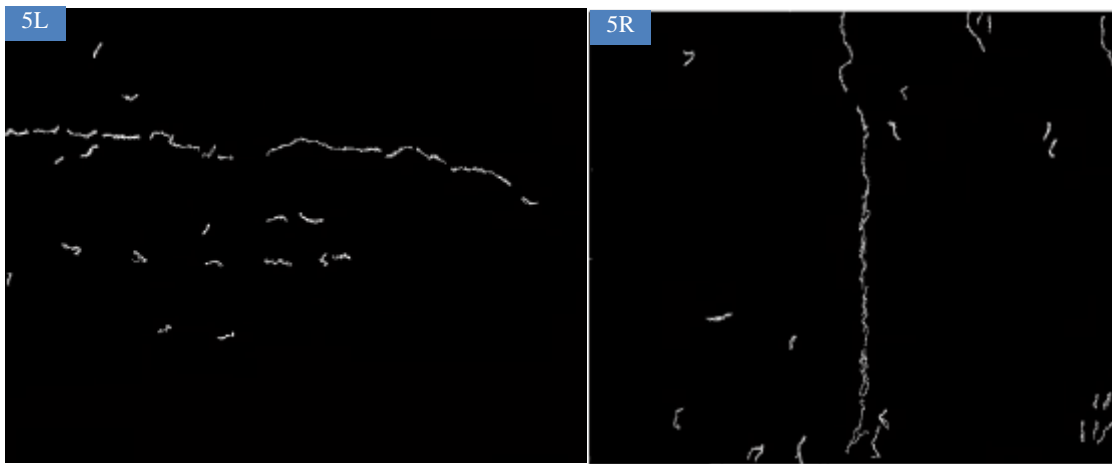


Figure 4.43: Final crack detection results on Google downloaded images

The final 5L and 5R of images were captured via a Nikon D100 camera and both images were captured using natural lighting condition for a real world pavement. The author's developed crack detection algorithm was run on this set of images and the final results are presented on figure 4.44. The results from this set were different from all previous sets. It showed that the false-positive rate was higher than all other previous samples with about 48% for the left image and 55% for the right image. This high rate of false-positives in both images might be attributed to the raised aggregate in the surface texture which casted a shadow on the image. The correct detection rate on both images was found to be 92.7% and 98.9% for left and right images, respectively.



**Figure 4.44: Crack detection results for rough and moderate texture image**

The false and correct detection rates were different from one sample to the other. The detection of cracks and other false-positive objects depends on many parameters as the above examples showed. Degree of roughness, texture type, illumination condition, crack width, image resolution and crack intensity can all play an important role in the detection process. The average crack detection and false-positive rates for all samples were calculated and found to be 83.72 % and 14.08 %, respectively. These two average values can be used as limits for the author's developed crack detection algorithm which is based on morphological tools to pre-process images and threshold limits using edge operation. The author's developed crack detection algorithm also needs to be validated on video extracted frames which will be examined in the following section.

### **4.9.3 Testing the author's developed crack detection algorithm on video images captured from a mobile van**

The aim of this test was to investigate the possibility of using video images to analyse and detect cracks through a fully automated mode using the author's developed crack detection algorithm. To achieve this aim, a Canon 5D camera was used for data collection along different paths which were shown in section 4.3 and figure 4.5. The camera was held vertically at about 2m height over the road surface through utilizing a designed wooden frame as discussed previously. The camera was in video mode and in shutter speed priority to avoid blurring.

The car was moved with traffic speed, and the shutter was changed to test the possibility of blurring occurrence. According to Fricker and Pacey (2005), blurring does not occur if the pixel size on the ground is bigger than the translation due to the car movement or forward motion. They concluded that in order to overcome the effects of blurring, the time of exposure has to be shorter (or equal to the pixel size), the focal plane has to be moved mechanically, or a time delay integration solution has to be applied. Due to the difficulty of using the last two methods, the first one (exposure time) was utilized in acquiring video imagery with the fact of accepting the radiometric loss. If the shift and direction of blurring are well known, the image can be de-blurred utilizing different available methods, refer to Matlab (2014a) documentation for more information about the utilized algorithms and methods.

A relationship was derived between the shutter speed and the traffic speed based on the pixel size and the height from the ground. At a 2 m height above the road surface, the pixel size on the ground is 0.458 mm which is more than the shift due to forward motion. Therefore, no blurring will occur even if the used shutter speed is 1/500, as can be noticed from figure 4.45. Visual inspection of video files at different traffic speeds and utilizing different shutter speeds showed that no blurring occurred. However, when using a camera at a 1 m height, the pixel size on the ground becomes 0.229 mm, which is less than the shift due to forward motion, and in some occasions with lower shutter speeds blurring will occur according to the relationship shown in figure 4.46.

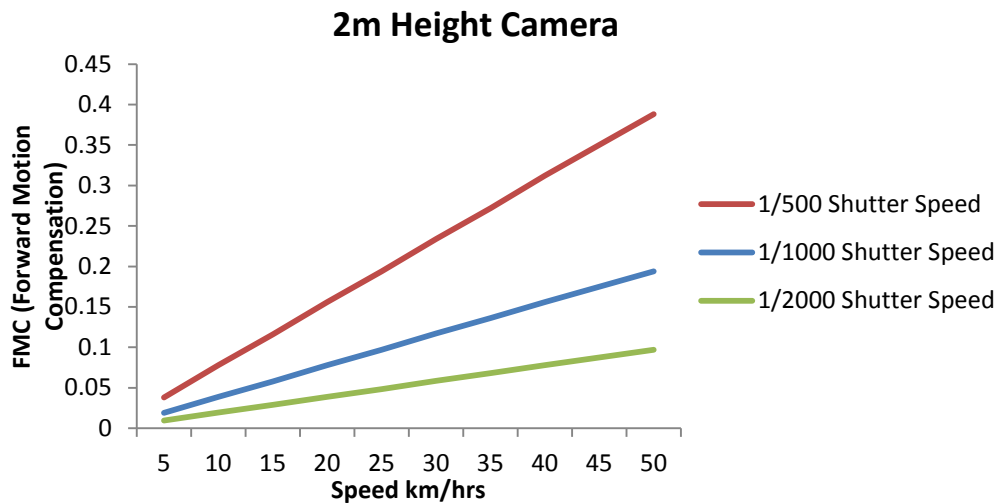


Figure 4.45: Relationship between Shutter speed and Traffic speed (2m height)

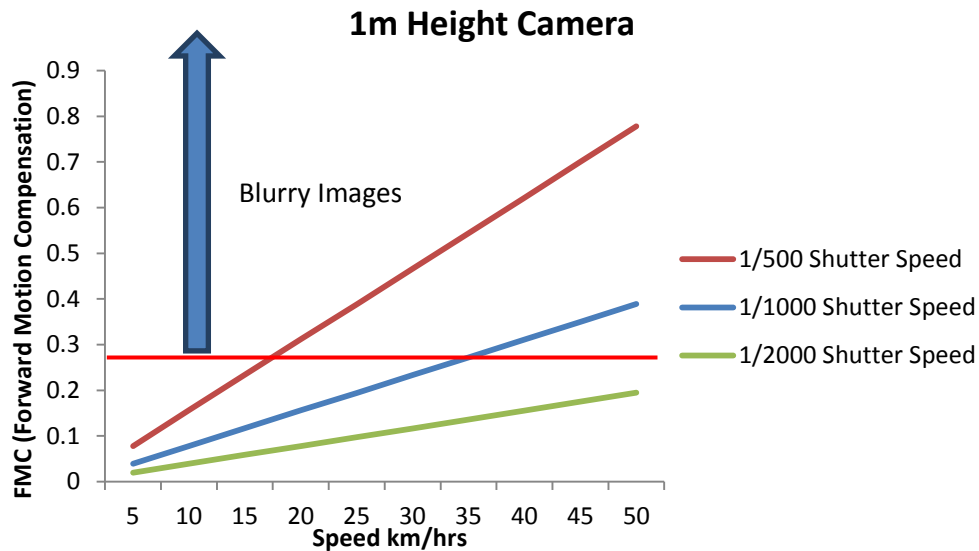


Figure 4.46: Relationship between Shutter speed and Traffic speed (1m height)

Quality images with high radiometric dynamism contrast are not possible to acquire using the video mode at high speed. The lack of radiometry in the video images at high speed is probably due to the reduction of the exposure trying to avoid blurring. However, this can counterbalance the lack of incoming light that enters to CCD leading to radiometric loss. The aim from this section was to evaluate the author's developed crack detection algorithm on video extracted frames. Video extracted frames were checked visually for the presence of cracks, and some examples were used for evaluation of the algorithm. The used examples are presented in figure 4.47.



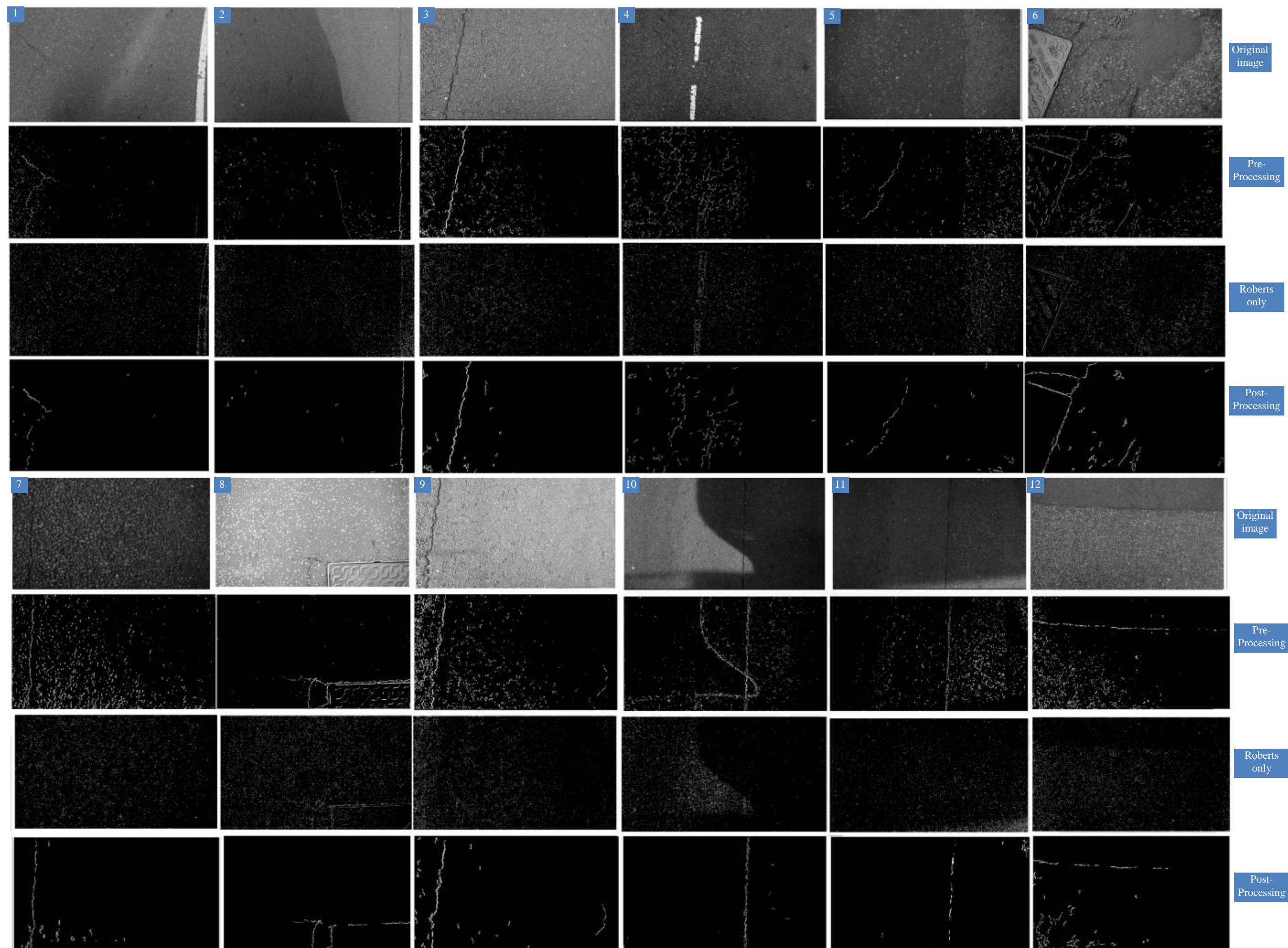


Figure 4.47: Video extracted frames with crack detection results

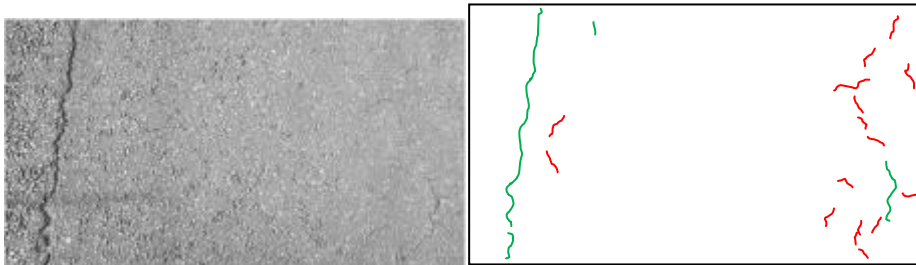
This Page is Intentionally left Blank



This Page is Intentionally left Blank

Figure 4.47 includes a wide range of real world problem images that contain shadows, texture differences, different crack widths, illumination changes, manhole existence and paint existence. The author's developed crack detection algorithm was tested on these samples of the extracted frames, and the results from each separate step (Roberts edge detection only, pre-processing, and post-processing) are shown in figure 4.47. As the figure shows, the pre-processing stage can greatly reduce the marker or paint effects and maximizes the effect of crack objects. The results from all samples were noisy and without any trace of crack objects; however, when applying the Roberts edge operation tool, the improvements can be attributed to the utilized image analysis tools and image filtrations which were described in section 4.8.3.

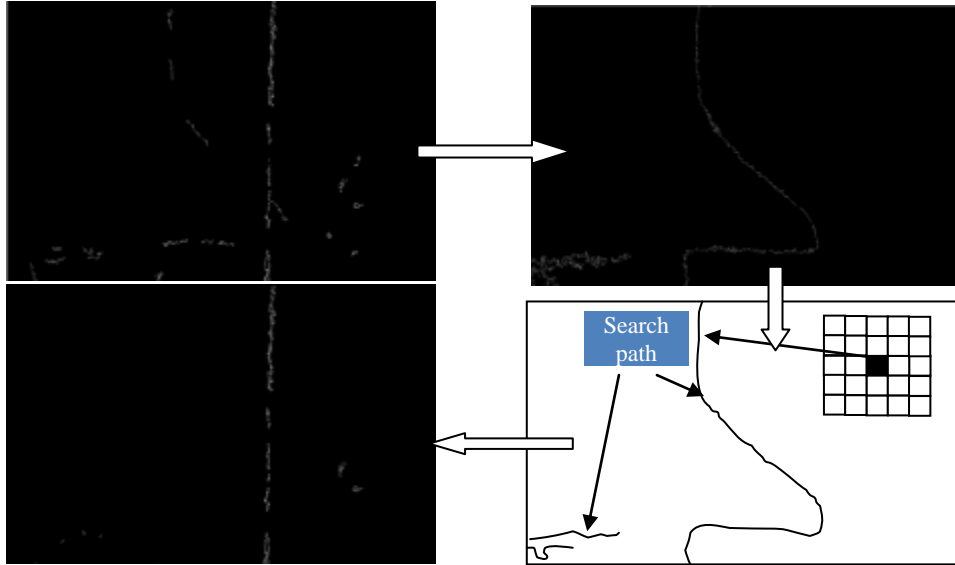
The correct and false-positive crack rates were calculated for each sample using a CAD tool. As it was previously stated in section 4.9.2, the correct and false crack rates were found to be different from one sample to the other, and depend on many parameters such as crack width, crack intensity values, shadow existence, texture type, and the degree roughness. The minimum correct rate was recorded for sample 9 which was found to be 44.5% followed by sample 12 which was found to be 48.6%. The lower correct rate in the sample 9 was recorded due to the existence of thin cracks which were not detected by the algorithm as figure shows.



**Figure 4.48: Undetectable cracks due to the width of crack (original image left and crack detection results on right, detectable are in green and undetectable on red)**

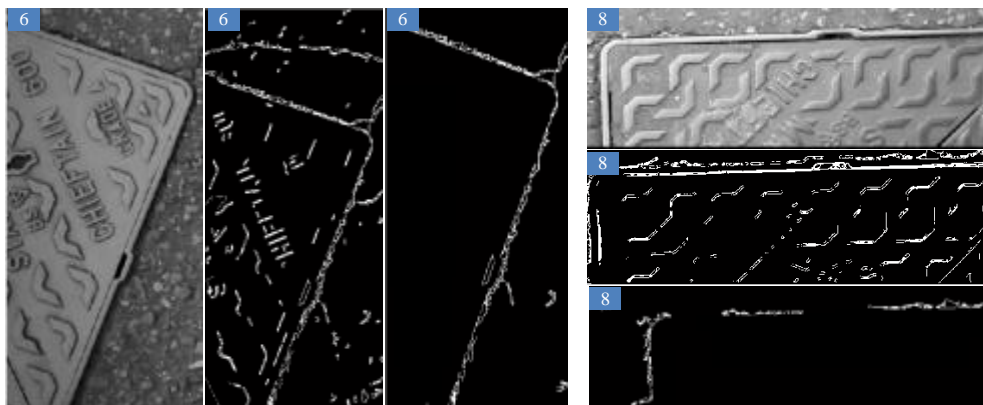
For any shadow detected image, the shadow removal algorithm mentioned in section 4.8.3.1 was utilized. The outputs from the shadow algorithm were used as input for the crack detection algorithm. All shadow existence images were subjected to the operation of shadow removal, and the image free of shadow was used for crack detection. Due to the differences between the lit and the shadow area, the crack detection algorithm treats the connection area as crack candidates. In the post-processing stage, the shadow core mask was used to differentiate the false-positive from the border area (between the lit and shadow area). The shadow mask is undergoing the process of edge detection which will be used as a search area

for false-positive detection and removal in the crack detection results. A square disk shape with  $5 \times 5$  pixels will be used for searching any false-positive cracks and will be removed as figure 4.49 shows.



**Figure 4.49: process of identifying the shadow and removing its effect on crack detection algorithm**

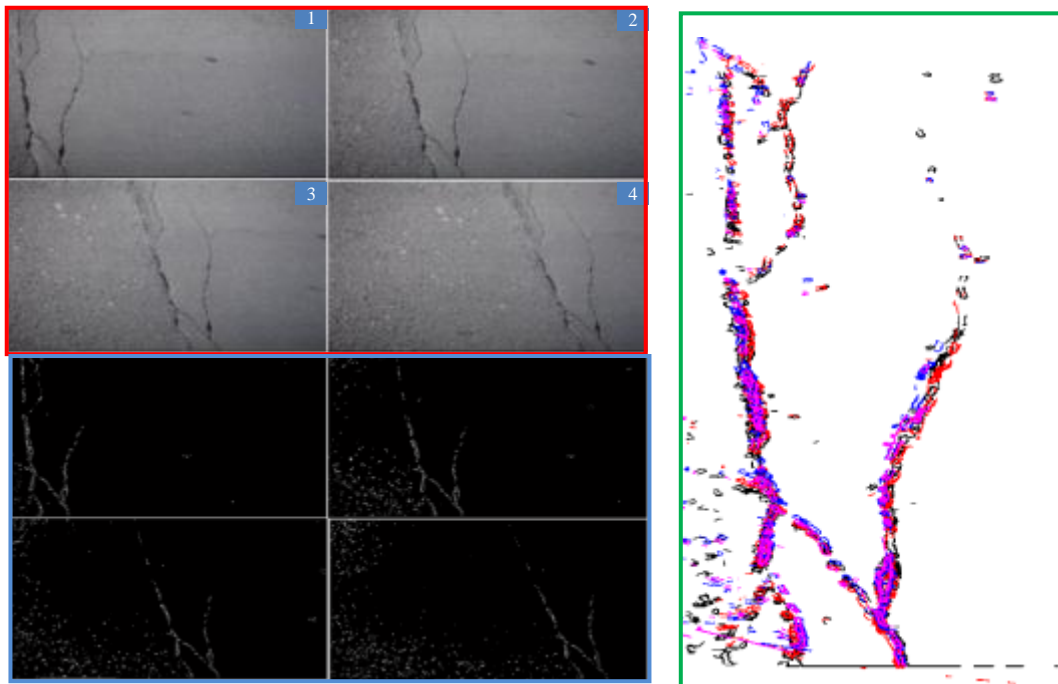
The maximum false positive rate was recorded for sample 6 which treated the manhole existence on the road as crack pixels, followed by sample 12 that applied a smooth instead of a rough filter. The false-positive rate was found to be 75.15%, and 65.9% for samples 6 and 12, respectively. The false-positives were effectively reduced using some statistical properties of crack objects such as straightness of crack candidate and shadow detection algorithms to remove the effects of the shadow. The manhole cover effect was reduced through testing the intensity value as it is clear from figure 4.50. However, crack pixels from sample 6 were merged with manhole pixels which had not been removed by intensity values checking and were treated as false-positives.



**Figure 4.50: Manhole cover false positives removed using value intensities**

For all 12 samples of HD video frames tested, the average false-positive and success rates were calculated and found to be 38.9% and 68.15 %, respectively. The success rate of crack detection for video digital images was dropped if compared with the other image sources in previous test, which was about 83.72%, see section 4.9.2. The decrease in the success rate of the author's developed crack detection algorithm can be attributed to the quality of digital image in case of video file. As mentioned before, the quality of image is dropping with video filming in order to avoid blurring, which can reduce the radiometry of the image. The other reason for this low success rate can also be attributed to utilizing real world images with various sources of problems such as shadow and manhole existence as well as different degrees of roughness.

Finally, the tests showed that the HD video resolution is adequate for the purpose of crack identification. The success of the author's developed crack detection algorithm might be further improved by utilizing different overlapped frames which can be used to merge the outcome from each frame. Utilizing different frames with different views might improve the detection rate of cracks. Figure 4.51 shows an example of crack from four different views. The four results were merged in a CAD tool and the results showed the possibility of increasing the crack detection rate by carefully calibrating images and applying appropriate transformation between frames which could be the basis of future research.



**Figure 4.51: Sample of extracted frames (shutter 1/1000 at about 35 km/hr) original images (red box), processing results (blue box) and result from four views merged in CAD tool (green box)**

## 4.10 From Semi towards full automation of crack detection

The main aim of chapter 4 was to build an automated system for crack detection at a reasonable price to be affordable by many local authorities, especially in developing countries. The overall cost of the system is in the range of £3000-4000. A Canon 5D camera with 28mm focal length was utilized for video image capturing of the road surface using a mobile van. The camera was chosen because enough GSD was possible over 2m, and also the user can control the exposure setting in video mode. It has been shown through different trials that the maximum possible speed that the system can reach without affecting the image quality (blurring) is 30 mph. 16GB of memory was used for data capturing which offers a continuous 4 hours of image storing at HD quality. The user can download the captured image directly to a PC utilizing the image acquisition toolbox from Matlab. This option can allow for performing image analysis on live-feed of images.

It is worth mentioning that after data collection, the processing stage was started by analysing the images for detecting cracks. This process involved the use of an algorithm to deal with different analysis approaches. In crack detection, the semi-automatic approach was achieved and full automation would be under focus in near future. For example in crack detection, different procedures were developed for each stage of pre-processing, processing and post processing. Images with a crack existence were visually checked and the output from each process was used as an input for other. Clearly, it is possible to link all of these procedures through one main program interface to make the process fully automated.

The main key stages in crack detection can be summarised as below, see figure 4.52:

- 1- Image preparation phase, which involved two main stages
  - a. Frame extraction, where images were extracted from video files using the author's developed motion estimation algorithm which is based on feature matching.
    - i. **Input** for this process: video files collected via a mobile van (a **link** is required if real time processing in the field, utilizing the image acquisition toolbox for downloading the captured image directly to a PC).
    - ii. **Output** from this process: frames extracted with non-repeated scenes.

- b. Frame Geotagging, where a GPS log is used to assign location to each extracted frame from the first stage.
    - i. **Input:** Extracted frames from previous stage (a **link** is required to use 1.a.ii data).
    - ii. **Output:** Geotagged frames (position assigned to each frame).
- 2- Reduction of image percentage: this process involved the use of statistics to identify images with problem existence such as manholes, cracks, potholes, paint, etc.
  - a. **Input** for this process: images from the first stage (a **link** is required to use 1.b.ii data).
  - b. **Output** from this process: images with problem existence.
- 3- Shadow detection and removal, where images were tested for shadow existence and removal.
  - a. **Input** to this process: images from the second stage (a **link** is required to use 2.b data)
  - b. **Output** from this process:
    - i. Images with shadows removed.
    - ii. A searchable shadow path (border between lit and shadow) for post processing stage.
- 4- Image pre-processing stage, where images underwent a process of enhancement, noise minimization, minimizing the effect of marking paint and strengthening the crack signal.
  - a. **Input** for this process, images from third stage (a **link** is required to use 3.b.i data)
  - b. **Output** from this process, enhanced images with noise minimizing and strengthening crack signal.
- 5- Image processing stage which involved two procedures
  - a. Image filtering, where images underwent a process of filtering based on image roughness measurement
    - i. **Input** for this step were images from fourth stage (a **link** is available for using 4.b data)
    - ii. **Output** from this process was filtered image.
  - b. Roberts edge operation, where image segmentation was used to detect crack candidates

- i. **Input** for this step were filtered images (a **link** is available for using 5.a.i data)
  - ii. **Output** from this process was crack candidates
- 6- . Image post-processing stage, which involved the process of false crack detection and removal through the author's developed crack connection algorithm (minimum path intensity) and removal of non-crack pixels.
  - a. **Input** for this process were
    - i. crack candidates from fifth step (a **link** is available for using 5.b.ii)
    - ii. Searchable shadow path to remove false crack due to shadow effect (a **link** is required to use 3.b.ii data).
  - b. **Output** from this process: detected crack pixels and the remaining non-crack pixels.

The term link refers to a subroutine that need to be designed in Matlab linking each step with a manual interface were required. It is possible in the future work to design a main program that can do all the procedures fully automatically without human intervention. It is worth mentioning that all codes and subroutines were designed using Matlab programme.

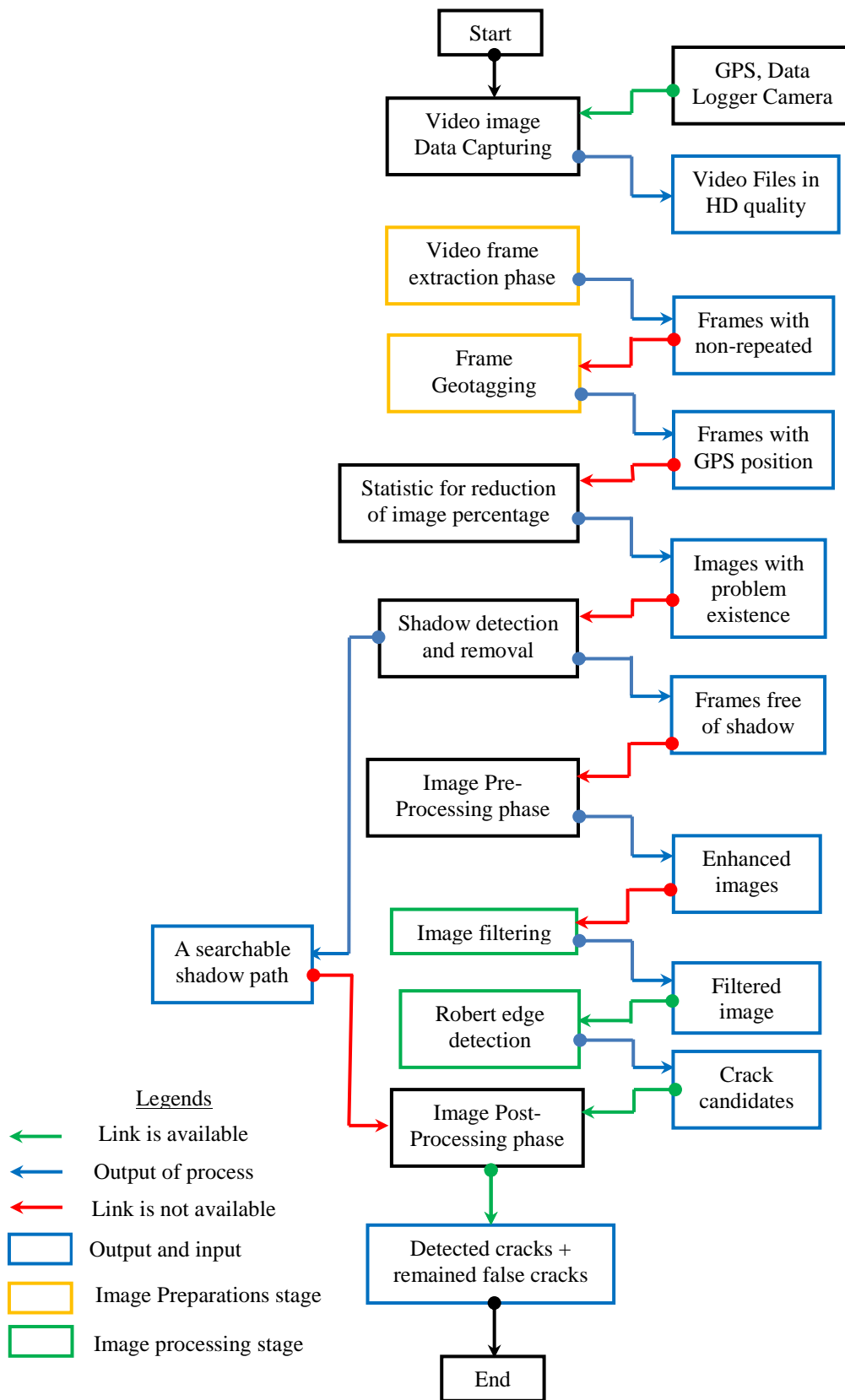


Figure 4.52: Main key stages in crack detection (toward fully automation)



## 4.11 Summary and Conclusions

This chapter started with the definition of the crack detection problem and showed that the classic approach of crack identification through thresholding was not sufficient with rough texture. The thresholding method is not efficient due to the availability of quite a large amount of false-positives, even if the local weighted limits are utilized through making profiles. The necessity for developing an algorithm to deal with cracks in the rough texture was critical due to the non-availability of any open source or commercial software to deal with crack analysis, especially for the rough texture.

It was also shown that it is possible to assemble low-cost tools of camera and cheap sensors for monitoring and detecting the cracks on the road surface. The results proved that by developing an algorithm, the cracks pixels could be detected and isolated from other image pixels. The author's developed crack detection algorithm was developed using thresholding and morphological operation tools and through image filtration. The algorithm was developed based on the roughness measurement in order to smooth images depending on the degree of roughness. The algorithm for crack detection also considered the large amount of data and a subroutine was designed to deal with this challenge.

Data handling was an important part of the algorithm development in this chapter. Due to the large amount of data and processing time for each particular frame, the author decided to handle this amount of data by only including the frames that share a certain percent (about 20%) of the area between successive frames. This aim was achieved by developing an algorithm named as motion estimation. It was shown that the motion estimation can estimate the translation between two frames accurately through utilizing the SURF matching algorithm. Further time management was achieved through statistical testing.

Due to the time consuming nature of running the algorithm, statistical testing was used to minimize the amount of frames and to include only the frames that have a real problem such as shadow, manhole, paint, cracks, and pothole existence. Statistical testing showed that is possible to identify the source of the problem through Gaussian normal distribution curve fitting to histograms. The results of this operation (Gaussian fitted curve with statistical t-testing) proved that it was possible to roughly identify the video frames with problem existence, with an average success of about 81%.

This chapter has also showed that video frame geotagging is possible through accurately calculating the starting time of the video camera trigger with regard to GPS time. Experiments showed that it was possible to locate the video frame 92% of the time with RMSE of less than 15m, which is 3% percent less than the claimed RMSE percentage. It is worth to mention that errors of as large as 26m were recorded in the current project but only 8% of the data had an error of less than 26 m and more than 15m.

A method was developed to deal with roughness measurement using only a single image. The method called object repetition was developed using a mixture of image and texture analysis tools. The method was tested on a number of predefined texture depths which included a total of ten samples. Trials showed that the method of object repetition is dependent on GSD. The farther away the camera was, the smoother the look of the texture was. Tests proved that there was no significant correlation found between object repetition and texture depth. However, object repetition could be correlated with a better prediction of the roughness of the texture. As the roughness of the texture increased, the more manipulation of pixels was required in order to smooth the image.

Tests of examining the success rate of the developed algorithm on different sources of images revealed that the success rate of crack detection reached up to 98.9% in one example of the image, but the lower rate was found to be about 58%. The average success rate of crack detection was found to be about 83% for all image samples. The reason behind recording different success rates was discussed for each particular sample and found to be dependent on many parameters, such as illumination changes, shadow existence, texture type, crack width and intensity limits. It was also shown that the author's developed crack detection algorithm can sometimes erroneously identify non-crack pixels and treat them as crack candidates, with an average false-positive for all samples of about 14%.

The success rate of the author's developed crack detection algorithm was also tested using video extracted frames of real world data. It could be proved that the success rate declined due to low radiometric and contrast qualities of the video images. The low quality of the video frame could be attributed to the camera exposure settings at traffic speed. A relation was drawn from theory between the speed of the vehicle and the forward motion for various exposure times and this showed that it changes with the height of the camera from the ground. Testing twelve extracted frames with various problem existences showed that the average success rate of the algorithm dropped from about 83% to about 70%, while the average false-positive increased from 14% to about 39% when using video extracted frames.

## CHAPTER FIVE: POTHOLE DETECTION FROM VIDEO IMAGES

### 5.1 Introduction

A pothole is a serious type of failure in an asphalt pavement caused by many factors. According to Paige et al. (2011), water is the primary recognized source of pothole development. However, many other factors can seriously affect pothole development such as crack growth, poor road design over certain subgrade, and insufficient pavement thickness to support traffic. Also, other non-structural causes exist such as falling rocks, spillage of diesel or any chemical, and even animal hooves in hot weather.

Keeping roads safe is a challenging process for many road authorities. Many pothole claims against road authorities and insurance companies have risen annually. In the year 2012, in the UK and Wales, in addition to the cost of repairing the potholes motorists were paid more than £22 million in compensation for pothole damage (Massey, 2013). The pothole problem is worldwide, for example, the state of Michigan in 2005 faced more than 7,500 cases for pothole-related damage claims. This is in addition to more than 500,000 pothole-related claims annually received by insurance companies. Therefore, periodic road assessments and collection of up-to-date information about the road surface condition seems to be the most effective way to conserve high road standards at a minimum price (Furness et al., 2007; Eriksson et al., 2008).

Many methods have been developed for the identification of potholes using different techniques such as laser sensors, mobile network sensors, the use of accelerometers, and through digital images (Eriksson et al., 2008; Koch and Brilakis, 2011; Mednis et al., 2011; Jahanshahi et al., 2012; Koch and Brilakis, 2012). Each method has their pros and cons, for example, 2D image-based methods are easy to use; however, challenges remains in the methods of processing, especially when the texture differences do not exist between potholes and the background. Existing methods such as high-maintenance range sensors tend to be expensive. Acceleration data on the other hand, are only providing preliminary and rough condition surveys. Thus, it is essential to provide an alternative method that is both accurate and easy to use.

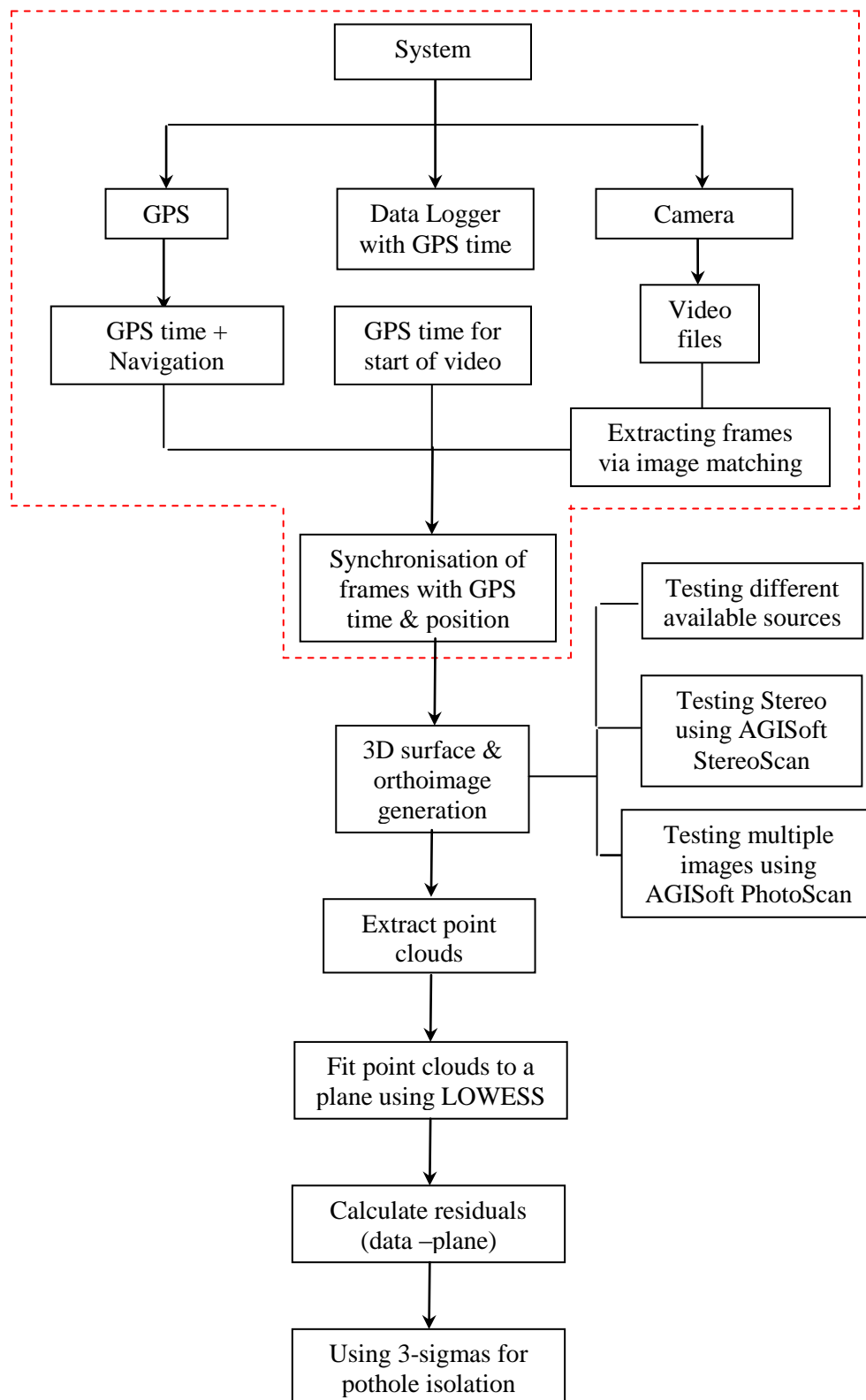
In this chapter, a method was developed based on using digital images for pothole detection through 3D based image modelling. Potholes or any deviated objects on the road surface will be investigated using an automated method through the use of video

imaging using a Canon 5D camera. This chapter will first describe the pothole detection method followed by a presentation of the data sets and test site. Then, procedures will be described, followed by the results and analysis, which will first evaluate available open source matching algorithms by comparing them with the traditional cross-correlation. Another objective included assessing the automated open source 3D models such as (AGISoft, Bundler, and Photosynth) and commercial photogrammetric programs in terms of accuracy and potentiality of detecting potholes and bumps. Lastly, the use of extracted digital frames from video files for pothole identification will be examined and followed by assessing the system for possibility of achieving full automation of pothole detection.

## **5.2 Design of a pothole detection method**

Many existing road monitoring systems use laser scanning based techniques for pothole identification. However, the manual reporting of potholes is still widely used in most countries across the world, even in the UK, where potholes can be reported in specific websites related to local agencies (GOV.UK, 2014). This method depends mainly on the users' contribution and attention. Other monitoring techniques such as digital images, mobile sensing systems, and networks of accelerometer sensors are also available. The limitations of these methods were previously discussed in the introduction of this thesis.

In this study, a new approach was developed to deal with pothole detection. The method relies on digital images extracted from video files for creating 3D surfaces based on photogrammetry. The preference of 3D models over existing 2D image analysis is due to the ability of the first to identify the potholes even without differences between the texture of pothole and the surrounded area, as trials have proved. This method showed that potholes as small as 2.5 cm depth can be recognized. The method relies on developing a cost effective system that consists of a camera (with the capability of video recording in HD quality such as the Canon 5D), cheap GPS and data logger. The system was discussed previously in chapter 4 of this project. The purpose of the system is to provide image frames with GPS position and time. The extracted frames can then be used as an input for the processes of 3D model generation. Figure 5.1 shows the overall process of the used system and the developed method.



**Figure 5.1: Workflow diagram for developed method of pothole detection**

----- Box contain the system that described in chapter four

### 5.3 Aims and Objectives of trials

The aim of this chapter is to assess the use of extracted video frames for pothole identification which will be achieved through the following objectives:

1. Trial 1: Investigating the potential of new feature matching algorithms such as SIFT and SURF adapted in many available automated 3D models by comparing them with a traditional cross-correlation algorithm.
2. Trial 2: Investigating open source and commercial software for 3D point cloud generation and bump detection.
3. Trial 3: Investigating the possibility of detecting a small hole using open source and commercial software
4. Trial 4: Investigating the use of video extracted images from a mobile system for pothole detection

### 5.4 Methodology of trials

In order to fulfil the above mentioned objectives, the following methodologies were carried out:

1. **Objective 1:** To investigate the potential of new feature matching algorithms, two sets of totally overlapping images were prepared by firmly fixing two Nikon D200 cameras on a trolley, see figure 5.2.
2. **Objective 2:** To investigate open source and commercial software for 3D point cloud generation and bump detection, a bump on the road surface was used for the evaluation using four images from a Canon 5D camera.
3. **Objective 3:** To investigate the possibility of detecting a small hole using open source and commercial software, a hole with approximately  $(5 \text{ cm}^3)$  in the asphalt road surface was utilized and captured using three images from a Nikon D200 camera.
4. **Objective 4:** To investigate the use of video extracted images from a mobile system for pothole detection, a special system was designed which consists of a camera, GPS and data logger assembled in a mobile van. Nine samples of pothole existence images were extracted from different videos with various shapes and sizes that will be used for the purpose.

## 5.5 Data sets and test sites

Two sources of data were used for this project; the first source was still image collection for assessing different tools of 3D model generation, while the second is the use of video images to test the proposed methodology of pothole detection. The first sets of data which are still images were divided into the following sources:

1. Two Nikon D200 cameras were fixed on a trolley, figure 5.2. The digital cameras were tilted from the nadir by about 12 degrees in the y-axis direction in order to have a 100% lateral overlap. The height over the road surface was approximately 1.2 m. The overlapped digital images from this set will be used for making a comparison between the matching algorithms namely cross-correlation, SIFT, and SURF.



**Figure 5.2: Trolley for 100% overlap multiple image acquisition**

2. A Canon 5D-28mm camera was used for collecting a set of still images for a bump on the road surface, for more details about the data and the results of this refer to section 5.7.2. The purpose of this trial was to evaluate, in terms of accuracy, point cloud density and processing time of 3D automated open source and commercial photogrammetric software.
3. The last source of still digital images was used to test the potential of open source automated 3D modelling tools and available photogrammetric applications for identification of a pothole as small as 5cm<sup>3</sup>. Three images were used with a sufficient amount of overlap. The camera used for image capturing was Nikon D200; more details are available in section 5.7.3.

The second set of data is the image frames extracted from video files. A special system was designed which consisted of a camera, GPS and data logger, all assembled in a mobile van. This data set was discussed previously in section 4.3. The aim of using this

data set was to evaluate the video extracted frames for pothole detection using automated 3D surfaces.

## 5.6 Procedures

### 5.6.1 Methodology of testing matching algorithms (Trial 1)

As described in section 5.5, a trolley was used to acquire a pair of images with 100% overlap. Two Nikon D200 cameras were used to capture images simultaneously for the road surface texture. The two cameras were assembled with a data logger and distribution box in order to control the timing of taking pictures at 1 second epochs. The exposure setting on both cameras was set to manual setting with aperture priority of F8.0, ISO sensitivity of 500, and shutter speed at 1/2000 seconds in order to guarantee that the quality of images acquired was not affected by blurring. Thirty-two images were obtained from each of the used cameras for a road near to NGBon the Jubilee campus; figure 5.3 shows samples of collected images. Only two pairs of the 32 images were used for making the comparison between different matching algorithms. The trolley was utilized in this trial in order to accurately determine the Exterior Orientation Parameters (EOPs) of both cameras that will be needed for conducting the rest of the procedures in the Leica Photogrammetric Suit (LPS).

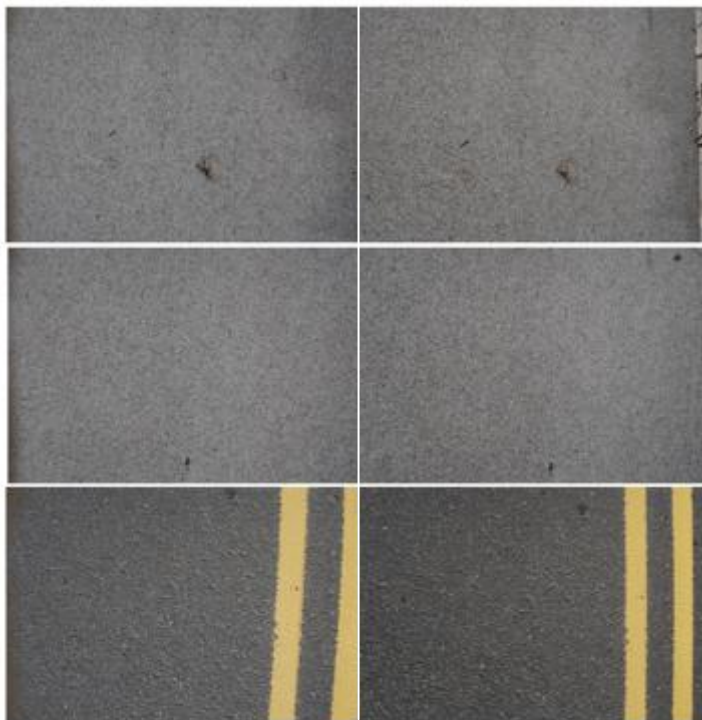


Figure 5.3: Sample of images acquired at 100% lateral overlap using the trolley



The comparison of different matching algorithms was carried out using LPS through utilizing the aerial triangulation tool. The first step in making the comparison was to accurately recover the position and orientation of the two cameras as well as the camera calibration parameters. There are different programs capable of camera calibration, such as Australis, PMS, and AGISoft Lens. Each program has its unique calibration method and model. Australis, PMS, and AGISoft Lens were all used to model the calibration parameters of the two Nikon D200 cameras, and the average results of these three programs are shown in table 5.1. Only the average of the two cameras was used because LPS only allows for calibration parameters of a single camera to be entered. Differences are apparent between the three programs which might be due to the fact that each program utilizes a different network of targets, calibration models or methods. Flat sheet and LCD screen were used in the case of Photomodeler and AGISoft Lens, respectively which do not account for differences in z direction (limited to 2D). Whereas, 3D target frames were used in the case of Australis that consider 3D object observation. The 2D targets are geometrically weak which might lead to instabilities in the least-squares estimation (Remondino and Fraser, 2006). For this reason, the calibration parameters of Australis were used in section 5.7.1.1.

Table 5.1: Average Camera Calibration parameters

Program	F(mm)	xp(mm)	yp(mm)	K1	K2	K3
<b>Australis</b>	28.7065	-0.0704	0.0233	-8.8E-05	4.15E-06	-2E-08
<b>PMS</b>	29.533	-0.0367	0.0057	1.4E-04	-1.5E-07	0
<b>AGISoft</b>	29.77	-0.1	-0.15	6.88E-04	8.66E-06	3.38E-07

The next step was to calculate accurately the relative orientation between the two cameras that are fixed in the trolley. For this purpose, a calibrated bar and several ground control points (GCPs) were used, figure 5.4. Three different photogrammetric programs were used for working out the EOPs, table 5.2, see appendix G for accuracy results of these three programs. The differences between the three programs used for location of the two cameras were less than  $\pm 0.015\text{m}$  indicating accurate location recovery. However, for the rotation values, the largest error recorded was for Omega which is about  $0.22^\circ$ . The mean value from the three programs was used for testing the matching algorithms inside LPS.



Figure 5.4: GCPs distribution and scale bar used for EO calculation

Table 5.2: EO parameters using different software

software	X(m)	Y(m)	Z (m)	Omega	Phi	Kapa
<b>Australis*</b>	454896.406	339696.554	32.232	-12.415	4.159	-115.894
	454896.141	339695.962	32.231	14.449	-6.746	-112.992
<b>LPS**</b>	454896.406	339696.548	32.247	-12.021	4.123	-115.914
	454896.138	339695.96	32.242	14.423	-6.888	-112.932
<b>PMS</b>	454896.387	339696.545	32.254	-12.437	4.282	-115.934
	454896.114	339695.937	32.253	14.699	-6.811	-112.99
<b>Differences</b>	$\pm 0.011$	$\pm 0.004$	$\pm 0.011$	$\pm 0.234$	$\pm 0.083$	$\pm 0.020$
	$\pm 0.015$	$\pm 0.014$	$\pm 0.011$	$\pm 0.153$	$\pm 0.071$	$\pm 0.034$

\*: transformed from Azimuth, Elevation and roll to Omega, Phi, Kapa.

\*\* : GCPs: 11, (0.02m ) same weighted values, IOP : fixed. No AP and Blunder check.

At this stage, the EOPs and IOPs were defined from utilizing different programs as explained previously. The next step was to generate matching points using three algorithms (SIFT, SURF and cross-correlation). Open source SIFT and SURF algorithms for research purposes were used in the Matlab program (Lowe, 2007; Bay et al., 2008) to generate matching points. Cross-correlation is the implemented algorithm in LPS, which was used to generate points using automatic tie point's generation. Two separate projects for first and second pairs in LPS were created using the values of EOPs and IOPs worked out in the previous steps, figure 5.5.

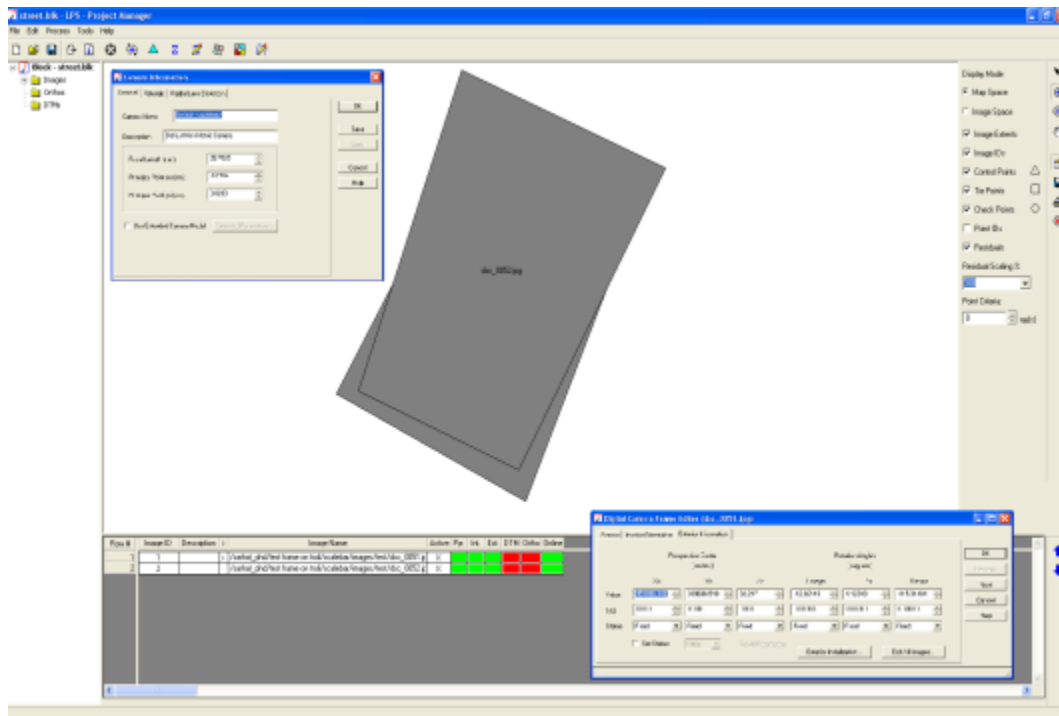
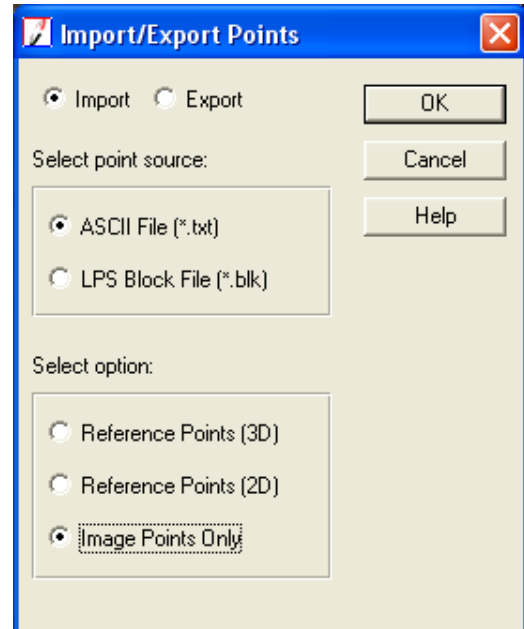


Figure 5.5: New project in LPS showing IOPs and EOPs, first pair

Points generated from both SIFT and SURF were imported into the LPS project through point measurement, import/export points tool as shown in figure 5.6 in addition to the points that were already produced using cross-correlation. In LPS, the point measurement tool provided control on the tie points to be active or de-active through the active column tab. From the point measurement tool, followed the aerial triangulation (AT) process for every ten points until reaches 150 points. For example, the LPS points' generated using cross-correlation needed to be processed 15 times for each pair. A total of 270 trials were processed for all matching points using the three matching algorithms and for the two pairs of images, refer to section 5.7.1 for more details.



**Figure 5.6: Tool for importing tie points from SIFT and SURF into LPS**

### 5.6.2 Pothole and bump detection (trial 2, 3 & 4)

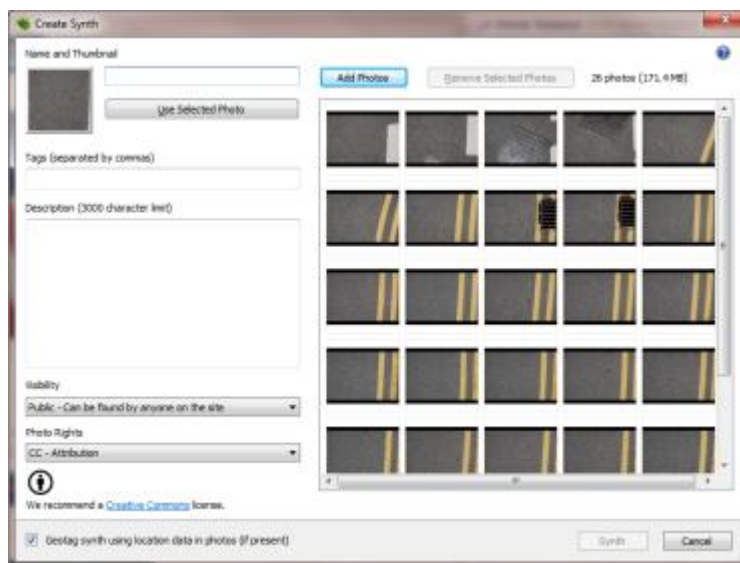
Potholes and bumps are considered to be highly visible defects on road pavement surfaces. Severe weather conditions are regarded as the main cause of pothole development. However, there are many other causes that can lead to pothole development. One of these causes is water penetration through cracks to the sub-base, and thus expanding the cracks from the base to the surface and eventually leading to a pothole. On the other hand, bumps in asphalt pavements are objects that are formed due to many reasons. A common example of bump formation in developing countries, such as Iraq, is when concrete is spilled onto the road surface due to poor transportation.

It is possible to identify both potholes and bumps on the road surface using 2D image analysis but with a limited level of accuracy. Due to this limited accuracy, using 2D image analysis is not sufficient, and to accurately locate potholes and bumps, the 3D surface model generation was required. In order to create a 3D surface of the pavement, the image has to overlap with a sufficient amount in order to reconstruct the scene using the stereo-matching technique, based either on photogrammetry or computer vision

system, or a merger between the two techniques. The methodology used in this project for reconstructing 3D models from different sources is described in the next section

### 5.6.2.1 Point clouds generation

Open source web services such as Microsoft Photosynth are designed to offer the user a free service to turn images into 3D point clouds (Photosynth, 2012). This web service has been used in this project to investigate the generated point clouds for the purpose of bump and pothole detection. Each user after registration with this web service is provided with 25 GB of free space in order to upload images and to convert them into 3D point clouds and panoramic scenes. The first step in using this service is to upload the images into the service through Photosynth desktop tool shown in figure 5.7. After uploading images into the service, the processing may take minutes to several hours depending on the size and number of images. After this processing stage is finished, a link to the processed images is prepared and through the provided link, the point clouds and camera calibration can be exported using the SynthExport tool. Different formats are offered for point cloud exportation such as ply, vrml, obj, and x3d.



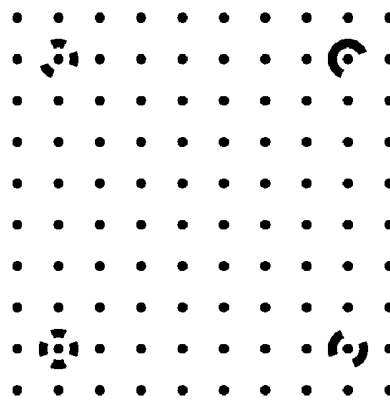
**Figure 5.7: Microsoft Photosynth desktop tool for image uploading**

Another free tool such as AGISoft StereoScan was also utilized in this project for the aim of detecting potholes and bumps. For more information about this free tool, refer to section 3.7.2.2.1. Bundler 3D is another free, open source tool that was investigated in this project which takes a set of images, image features, and image matches as input. As a result, it produces a 3D reconstruction of camera and (sparse) scene geometry as

output. For more information about Bundler 3D and the procedures of 3D reconstruction from images, refer to the documentation of Snavely (2010).

Two commercial tools, Photomodeler and LPS were also used for producing point clouds for investigating pothole and bump detection throughout this chapter. Photomodeler® (2009) was developed by Eos Systems Inc., Vancouver, Canada, and can be used for the production of 3D models from plane images including all other processes such as calibration and image matching. Camera calibration is an essential part of Photomodeler, and all photogrammetric tools, due to the capacity of the calibration function to improve the accuracy. The camera calibration parameters in Photomodeler usually include the focal length, lens distortion, format aspect ratio, and the principal point.

The camera calibration in Photomodeler (PhotoModeler®, 2009) is straightforward and fully automated; calibration photos can be acquired as recommended by the system developer, based on a flat pattern. The flat sheet can be positioned onto the floor, and three images can be collected from each side of the sheet's four sides. Figure 5.8 shows the flat sheet calibration from Photomodeler that is usually utilized for the calibration of the camera. As recommended by the developer, the focal length has to be first put on auto focus then fixed. With the focus fixed, the images were captured in the standard way. The camera position has to be close to  $45^\circ$  from the horizontal and vertical axes and the camera was rotated at each position.



**Figure 5.8: Photomodeler calibration sheet**

The calibration results can be further improved by considering two procedures; firstly, taking the calibration images at a range similar to the camera / object working

range in the project. The second procedure includes utilizing a field of 3D targets instead of a flat plane. It is expected that self-calibration bundle adjustment utilizing a field of 3D targets might increase the accuracy. This is due to the fact that they account for the changes in camera zooming and focusing. The camera also changes with temperature or by changing the camera settings. The field calibration in photomodeler was not carried out in this project due to unavailability of 3D coded targets (PhotoModeler®, 2009).

Image matching in Photomodeler can be performed using coded targets to speed up the process of referencing and matching the photos. Coded targets are used in Photomodeler to automate the process of 3D point extraction and/or the project setup (initial marking, referencing and orientation of photos). There are different shapes and types of coded targets available in Photomodeler which can be printed to a printer or to a postscript file; figure 5.9 shows a sample of coded targets.



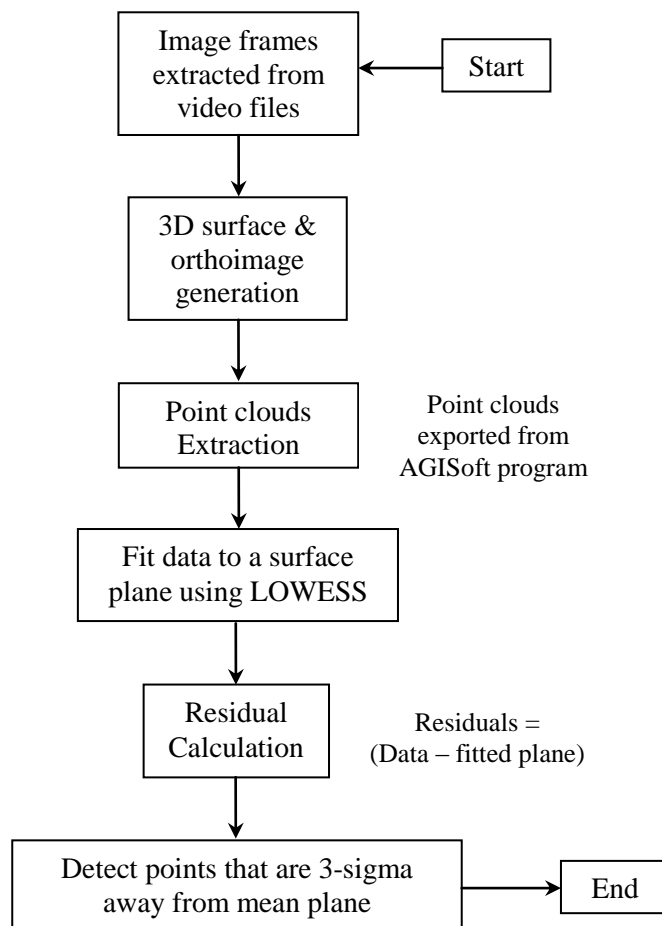
**Figure 5.9: Types of automation coded targets in Photomodeler**

In the absence of coded targets, the user needs to mark points in the processing of the photographs and then cross-reference them in different photographs that were taken from different directions. However, the latest available image matching techniques can eliminate the use of coded targets. In the most recent version of Photomodeler, there is a new feature available called SmartMatch. It searches out natural feature points in images as a replacement for coded targets in certain scenes, and then matches them between photos and generates 3D coordinates automatically.

After image matching and running the bundle block adjustment, the next step is image idealization or re-mapping (pixel by pixel). Image idealization will remove any lens distortion, non-centred principal point and any non-square pixels. The next and final process in Photomodeler is to produce point clouds via running the 3D dens surface at the required sampling distance. On the other hand, turning digital images into usable 3D data or point clouds for extracted features in LPS core can be found in Leica (2006a).

### 5.6.2.2 Algorithm designed for surface anomalies detection

In the previous steps, it was shown how point clouds can be produced using a mixture of open sources, free tools, and commercial photogrammetric software. The generated 3D point clouds will be used as input for detecting potholes or any deviations from the street surface, such as bumps. An algorithm was developed as part of this project to fit the xyz data (point clouds) onto a surface plane using a Locally Weighted Scatterplot Smoothing (LOWESS) fitted function. After the point clouds were fitted with the best plane that minimizes the residuals, the residuals in z-direction from the fitted surface plane are calculated using the curve fitting tool in Matlab. Any residuals greater than 3-sigma (standard deviation) were then considered to be a part of a pothole or bump. Image frames with the pothole detected can then be isolated for further analysis by the user; figure 5.10 shows the overall procedure.



**Figure 5.10: Procedures for detecting the Potholes or bumps on the road surface**

## 5.7 Results and Analysis

A number of trials were designed in order to evaluate the proposed objectives. The first test was designed to evaluate the matching algorithm that is used in commercial photogrammetric software and the available open source software. Another point considered for testing was how likely it is to improve the results of the feature interest matching algorithms. The second test was designed to compare between computer vision and photogrammetric software in terms of visually inspecting the resultant 3D object coordinates, processing time and the resultant point cloud density. The potential of the available point cloud generation tools for detecting small objects was investigated in the following section in the third trial. The possibility of using video images for the purpose of analysing and detection of potholes through 3D point cloud generation with acceptable accuracy will then be presented last.

### 5.7.1 Trial 1: Comparison of various matching algorithms

The aim of this experiment was to make a comparison between a traditional feature matching algorithm such as cross-correlation that is used in LPS and the feature interest approach represented by SIFTS and SURF. Most of the common automated 3D systems are employing SIFT or SURF in their digital matching processes. An example of an automated system that utilizes SIFT in the matching between images is open source web services Photosynth. The comparison will be carried out using two pairs of collected images as shown in figure 5.11 through employing a trolley mounted with two Nikon D200 cameras, as explained in section 5.5. The comparison will be in terms of the resultant total image unit-weight RMSE from the aerial triangulation process in LPS core. Two different scenarios were considered, the first was with no attempt to remove outliers and the second was with outlier rejection algorithms.

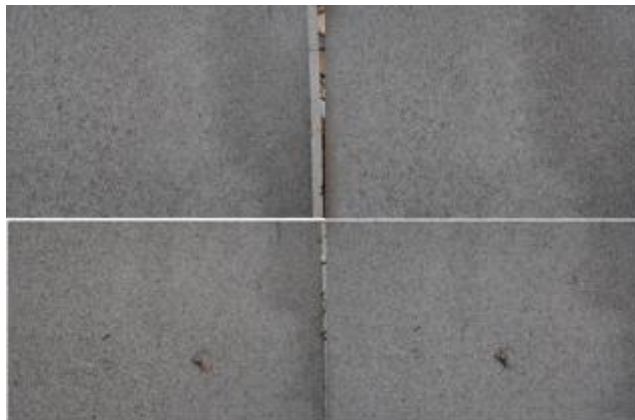


Figure 05.11: First pair of tested images above and second below



### 5.7.1.1 Implementing the test without outlier detection on both SIFT and SURF

A new project and two separate projects were created in LPS, with the first and second pair of images (Left and Right) shown in figure 5.11. As LPS does not allow for different camera parameters, the IOPs were set to the known and fixed option using the average camera calibration parameters from Australis (see table 5.1). The EOPs are also fixed using the average EOPs from the three programs as shown in table 5.3. Using EOPs as fixed values will allow the errors to propagate to the tie or matching points and give a good indication of the investigated matching point accuracy.

Table 5.3: Mean Result of EOP using three programs

ID	X (m)	Y(m)	Z(m)	Omega	Phi	Kapa
Left	454896.406	339697.325	32.25	-12.02	4.123	-115.91
Right	454896.138	339696.412	32.24	14.423	-6.89	-112.93

In LPS, since the EOPs are set to fixed values, tie points can be extracted and generated automatically using cross-correlation algorithms (Leica, 2006b). Alternatively, the two open source matching algorithms based on feature extraction (SIFT and SURF) were also used to generate the tie points using the Matlab programme. The three algorithms produced about 150 points that were sufficient for making the comparison. Both algorithms, SIFT and SURF, generated matching points using the original implemented algorithm without any aim to exclude outliers.

It is worth mentioning, that the generation of tie or matching points in LPS was conducted using the calibration parameters (IOPs) calculated from Australis. While, for both SURF and SIFT, the matching point's generation were conducted on non-calibrated images. The location of matching points has systematic errors (lens distortion) which subsequently can affect the accuracy of both SURF and SIFT matching results. However, the improvements were possible in both SURF and SIFT matching results through using outlier rejection algorithms as explained later.

The results of these three algorithms were tested every ten tie points and drawn against the total image unit-weight RMSE values that result from aerial triangulation, figure 5.12. The results from the three algorithms showed that as the number of tie point increased the total image unit-weight RMSE decreased. These results are as expected as long as the observations are free from outliers. The results from LPS were found to be the best as LPS has a filter for detecting and removing the mismatched points in the automatic tie point generation process. The trend for LPS is a steady decrease with no jump in the total image unit-weight RMSE. However, both matching algorithms SIFT and SURF showed that

outliers were feeding into the bundle solution as points were added to the aerial triangulation process, as can be noticed from figure 5.12.

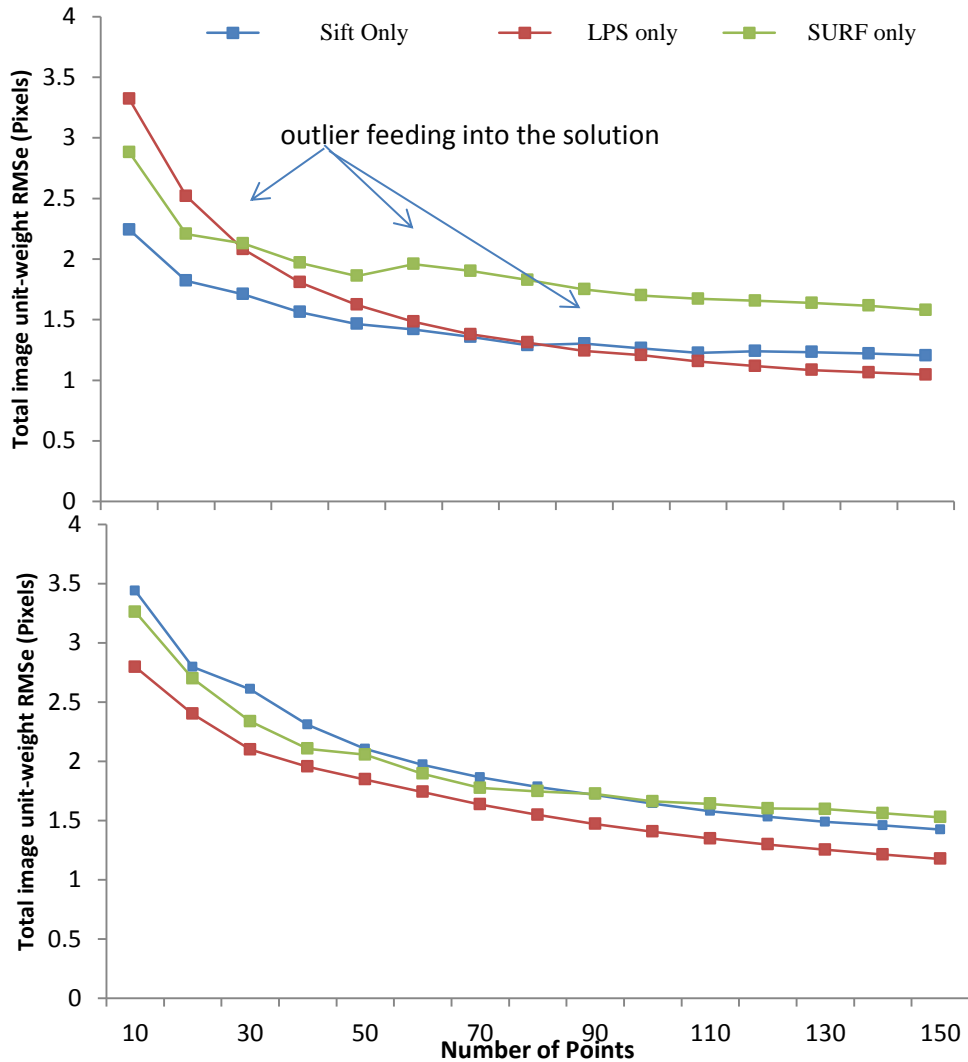


Figure 5.12: Aerial Triangulation Results, three algorithms, 1<sup>st</sup> pair (above) & 2<sup>nd</sup> (below)

It can be concluded from the two samples that the total image unit-weight RMSE is decreased with increasing the number of tie points. However, the incorrect matches that are feeding to the solution can increase the RMSE. LPS can provide tie points with outstanding accuracy estimated from a robust outlier detection technique using automated gross error checking or a bundler checking model. The measured image coordinates were analysed to determine the erroneous input data contributing to the overall instability of output aerial triangulation results and thus were removed in the next least squares iteration (Leica, 2006b). Whereas, an outlier rejection algorithm needs to be used to detect the incorrect matches from both SIFT and SURF results which will be covered in the next section.

### 5.7.1.2 Implementing the test with outlier detection on both SIFT and SURF

The generated points from both algorithms (SIFT and SURF) were tested for the effect of outlier detection algorithms. Different outlier detection algorithms are available; examples are LMeds (Least Median of Squares), RANSAC (RANdom SAMple Consensus), and MSAC (M-estimator SAMple Consensus). The results from different outlier detection algorithms showed that the total image unit-weight RMSE value decreased for both sets of images (first and second pair) and for both matching algorithms (SIFT and SURF) as the number of the points is increased, see figure 5.13 and figure 5.14.

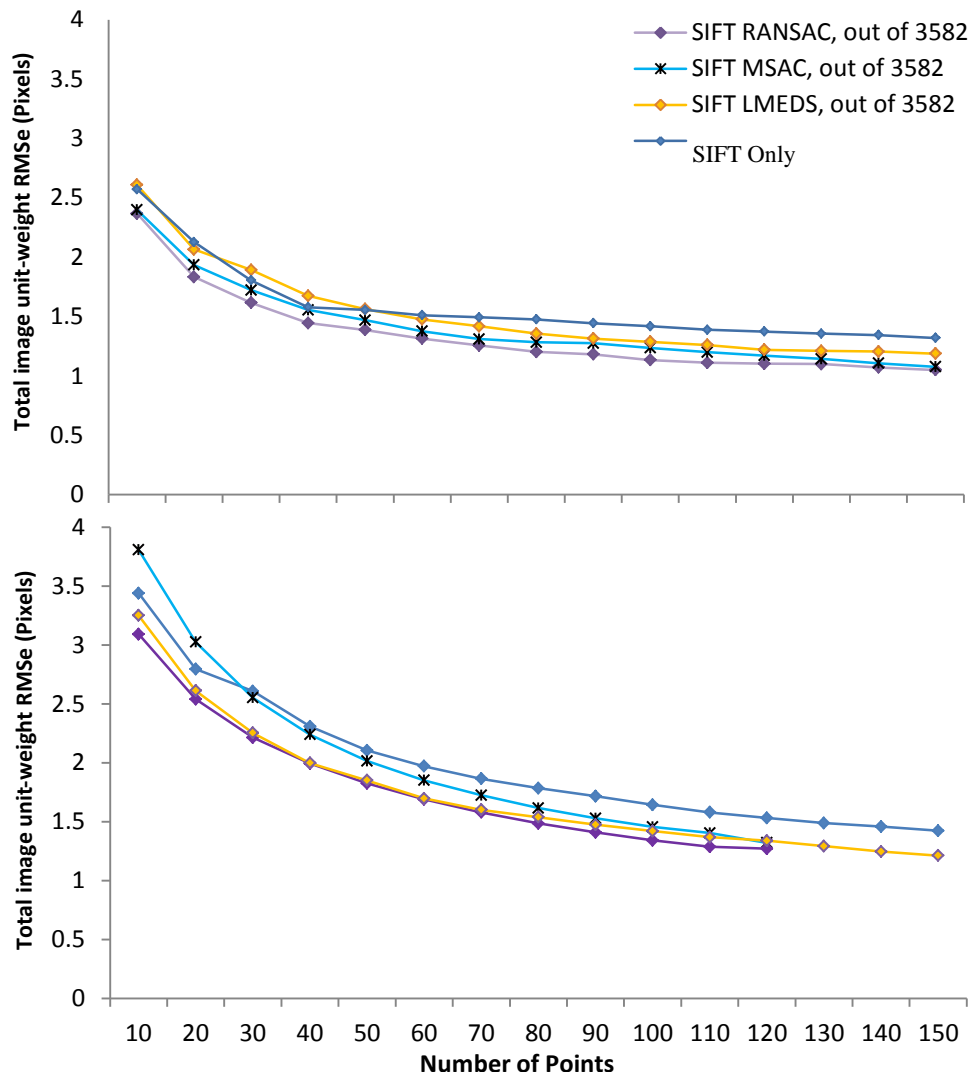


Figure 5.13: Effect of Outlier Detection Algorithms on SIFT, 1<sup>st</sup> pair (above) & 2<sup>nd</sup> (below)

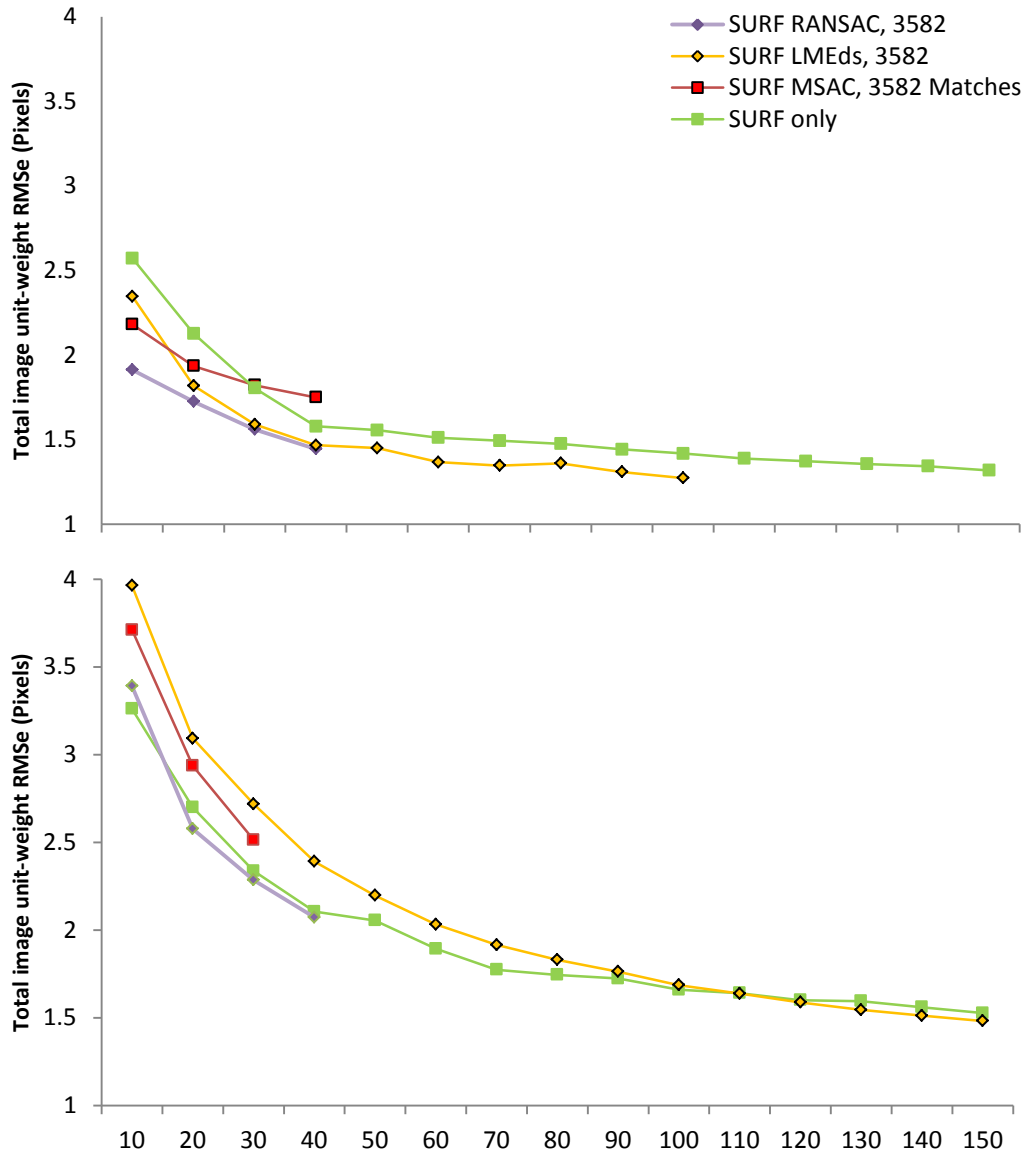


Figure 5.14: Effect of Outlier Detection Algorithms on SURF, 1<sup>st</sup> pair (above)  
& 2<sup>nd</sup> (below)

In the SIFT case, the best results were obtained using the RANSAC outlier detection and removal algorithm. In the SURF case, however, the RANSAC is working fine as well but the number of inlier or correct matches is always less than the SIFT case, which might be due to the fact that SURF contains more outliers than SIFT. In this case, LMEds was found to be better for working with SURF as can be noticed from figure 5.14. The final result of outlier rejection algorithms showed that the best accuracy achieved, especially for SIFT is with RANSAC. The result of total image unit-weight RMSE for SIFT with RANSAC nearly achieved the same accuracy as in LPS core, see figure 5.15.

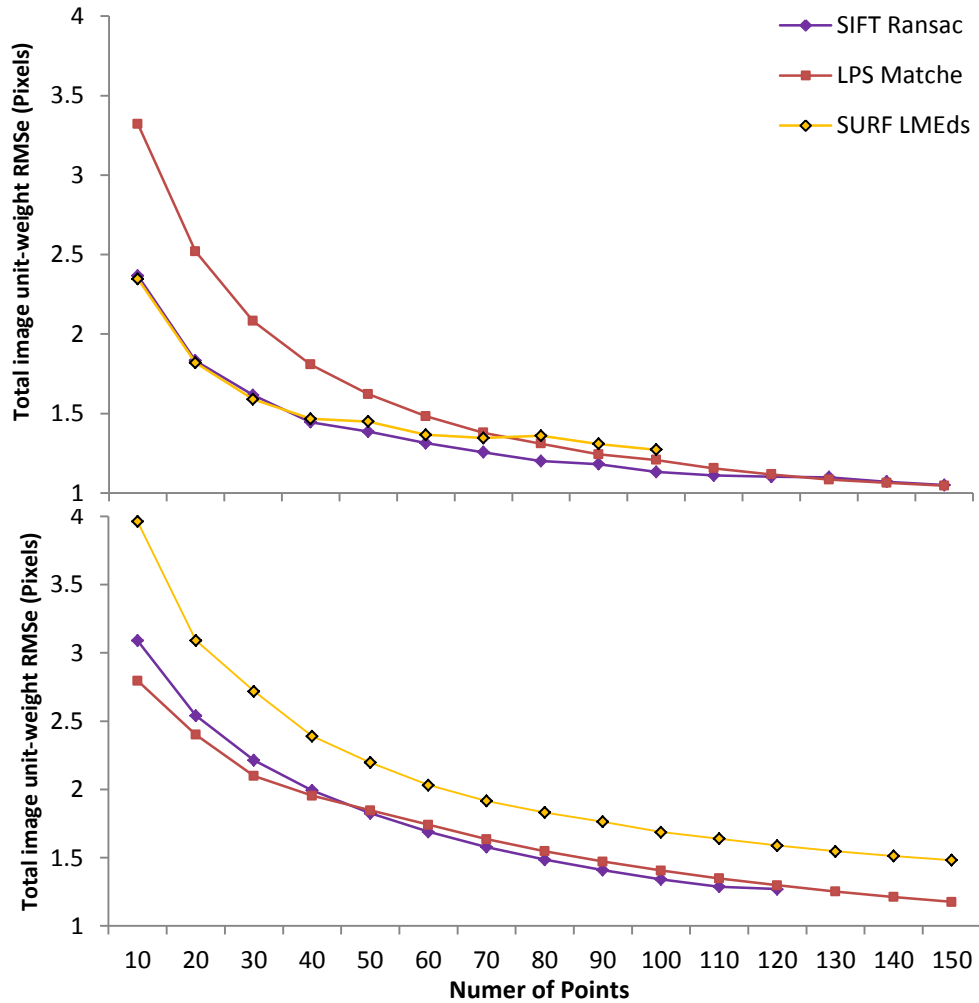


Figure 5.15: SIFT and SURF with Outlier removed vs LPS, 1<sup>st</sup> pair (above) & 2<sup>nd</sup> (below)

These results revealed that SIFT outperformed SURF in providing accurate matching points, that are as accurate as the tie points produced from LPS. The lowest total image unit-weight RMSE is about one pixel. There are many reasons that could be behind this comparable total image unit-weight RMSE. First of all is the IOPs that are set to the average and fixed. The second cause might be due to the number of used images (two) in the bundle solution. Another factor might be constraints used in the bundle solution, both IOPs and EOPs fixed with no standard deviation; the errors from both EOPs and IOPs might propagate into the total image unit weight RMSE. It is worth to mention that in the case of SIFT and SURF; no prior information was required of camera pose and the points are generated on a fully automated basis. However, in the case of LPS, minimum constraints were required which are six tie points or three GCPs, or direct geo-referencing with known EOPs values. The next sections were designed to compare between different programs for anomalies detection using the previously-investigated algorithms for digital image matching.

### 5.7.2 Trial 2: Comparison between open source and commercial software for 3D point cloud generation and a bump detection

The aim of this trial is to compare between the available open source and commercial photogrammetric software in terms of inspecting visually the resultant 3D object coordinates, processing time and the resultant point cloud density. The test site showed a bump on a surface of the road; the bump is excess asphalt which forms a hard solid object on the road surface. The size of the bump is 0.3227m, 0.5066m, 0.0600m for width, length, and maximum height, respectively. Four images were collected using the Canon 5D camera from different views, figure 5.16. The camera positions were close to about 45° from the horizontal and vertical axes. The average pixel size on the ground was about 3.5\*3.5 mm.






The four images were processed using different commercial photogrammetric software and open source software, such as Bundler, Photosynth, LPS and AGISoft StereoScan. Refer to section 5.6.2.1 for the procedures of point clouds generation from each of the used softwares. Appendix H shows the processing accuracies and output results for the used commercial software and AGISoft StereoScan.



**Figure 5.16: Four shots showing the bump on the street**

The results from the different set of softwares showed that the best available option for automatic point cloud generation was through using AGISoft StereoScan, which produce more than enough points for geometrically recognizing a pothole or a bump in the street, see figure 5.17 for visual inspection. In terms of point cloud

generation density, once more AGISoft StereoScan could provide a highly dense number of points in a very short time, within less than 20 seconds. LPS on the other hand, showed lots of outliers which are not forming a part of the bumps as can be seen from figure 5.17. Here it should be noted that LPS is designed for aerial applications, although the manufacturer claims that it can also be used in close range applications. However, the test showed that it is not accurate for point cloud generation in close range applications, despite an adequate number of GCPs (11), accurate calibration, and adequate triangulation results, see appendix H and Appendix I. The main reason of LPS giving low quality results might be due to the use of cross-correlation for matching between images that cannot handle the big rotation and scale changes as in this case.

	<b>LPS</b> Oriented Images: 4/4 # Point Clouds: 22028 Elapsed Time : 15 mins
	<b>PhotoModeler</b> Oriented Images: 4/4 # Point Clouds: 5357 Elapsed Time: 3 hrs
	<b>PhotoSynth</b> Oriented Images: 4/4 # Point Clouds: 5455 Elapsed Time: 3 mins
	<b>Bundler</b> Oriented Images: 4/4 # Point Clouds: 1320 Elapsed Time: 5 mins
	<b>AGISoft StereoScan</b> Oriented Images: 2/2 # Point Cloud: 26344 Elapsed Time: 14.5 sec

**Figure 5.17: Visual analysis of the achieved 3D points clouds results**

Resultant point clouds from Photomodeler, Photosynth and AGISoft StereoScan, can be used for identifying the bump using an algorithm developed as part of this project and explained in section 5.6.2.2. The results showed that it is possible to identify such an anomaly on the road surface using the automatically generated point clouds from Photosynth and AGISoft StereoScan or semi-automated point clouds from Photomodeler, figure 5.18 and figure 5.19. AGISoft StereoScan can precisely detect the bump and isolate it from other parts of the point cloud at the three different sigmas away from the fitted curve. The next section is to investigate the methodology on a smaller object than the one examined in this trial.



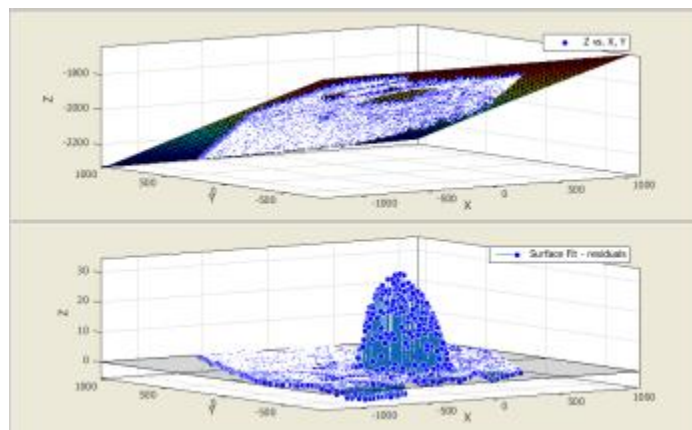


Figure 5.18: Fitted Surface (above) and residuals (down)

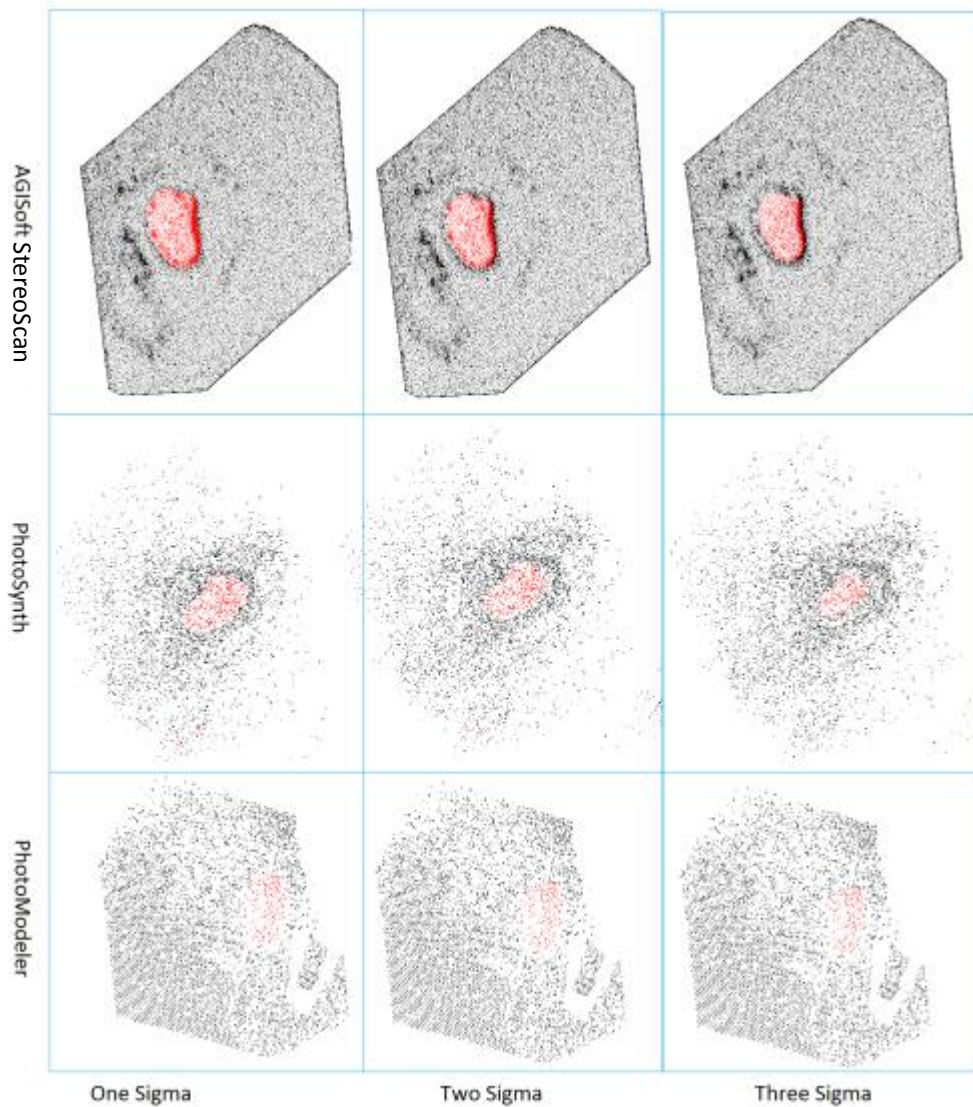


Figure 5.19: Bump detected using exclusion criteria at different sigma's



### 5.7.3 Trial 3: Investigating the use of 3D models for small object detection

The aim of this test was to investigate whether it will be possible to detect a small hole of approximately  $5 \text{ cm}^3$  size and with a depth of about 1.5 cm. As previously mentioned, the aim again was to make a comparison between the 3D point clouds generated from different software packages. Three images were collected using Nikon D200 with the camera plane perpendicular to the road surface for the middle image and low oblique for other two images, figure 5.20. The base to height ratio for successive pairs was ranging from 0.40 to 0.60.

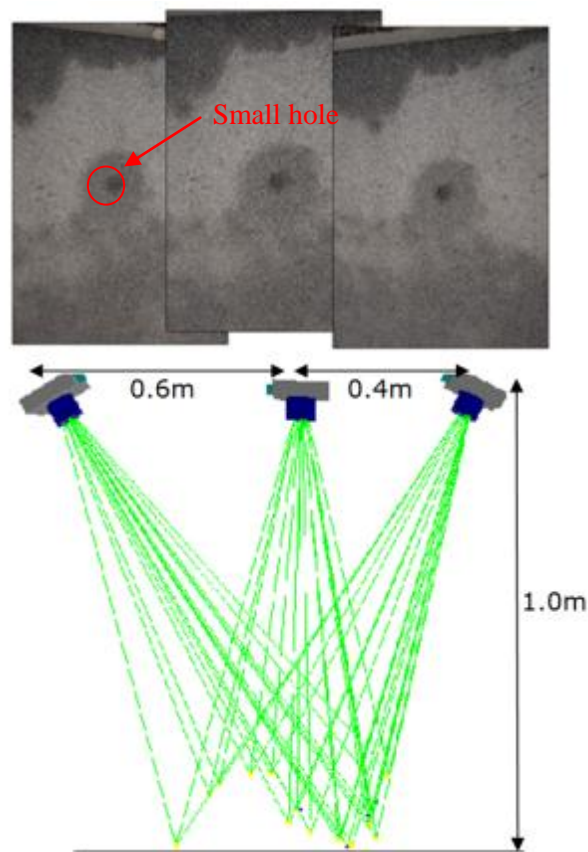


Figure 05.20: Location of Cameras and used images for analysis

The results from this test showed that Photosynth and Bundler produced less point clouds when compared with other softwares. LPS and Photomodeler produces the highest density of point clouds, but contained lots of outliers, which might be due to one of the following reasons:

- 1- Bad exterior orientation parameters (both).

- 2- Mismatch of correspondence points (both).
- 3- Not being designed for close range applications, especially in this case where the camera was perpendicular to the ground, and assumed as aerial photographs (LPS).
- 4- The base to height ratio was less than what is required in LPS; the accuracy increases with the increase of the base to depth ratio (Remondino and El-Hakim, 2006).

The best results achieved were through using AGISoft StereoScan, despite the fact that only two images have been used for the production of 3D points. The produced points were not as dense as in Photomodeler and LPS, but were more accurate and contain no outliers, even the hole could be clearly visible, figure 5.21. LPS and Photomodeler point clouds cannot be edited and outliers cannot be removed and excluded from further investigation. The remaining point clouds from Photosynth, Bundler, and AGISoft StereoScan were investigated using the designed algorithm for surface anomalies detection, see section 5.6.2.2. The point clouds from these three tools were exported to Matlab for detecting the small hole on the road using one, two, and three standard deviation or sigma exclusion criteria.

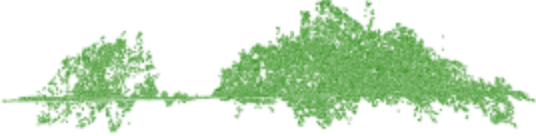




	<b>LPS</b> Oriented Images: 3/3 # Point Clouds: 53526 Time Elapsed: 35 mins
	<b>Photosynth</b> Oriented Images: 3/3 # Point Clouds: 6963 Time Elapsed: 3 mins
	<b>Agisoft StereoScan</b> Oriented Images: 2/2 # Point Clouds: 25477 Time Elapsed: 13sec
	<b>PhotoModler</b> Oriented Images: 3/3 # Point Clouds: 415568 Time Elapsed: 3.5 hrs
	<b>Bundler</b> Oriented Images: 3/3 # point Clouds: 11078 Time Elapsed: 6 mins

Figure 5.21: Visual analysis of the achieved 3D points clouds results

The results from the three exported point clouds revealed that neither Photosynth nor Bundler can be used to detect such a small size of pothole on the road surface. The reason is that both softwares are designed to generate sparse point clouds. However, the results from AGISoft StereoScan showed that this size of the hole can be detected using all exclusion criteria but with some amount of errors, especially on the edge of the point clusters, as can be noticed from figure 5.22. Therefore, it can be concluded that the algorithm can detect such a size of deformation on the road surface using the merging of technology from photogrammetry and computer vision represented by AGISoft StereoScan software, and through only two images. In the next section, the methodology of pothole detection will be tested using AGISoft StereoScan on some real video images collected from different road asphalt surfaces around the Jubilee campus, see section 3.5 for details about mobile van that was used for data collection, and for a list of video files see table 4.1 and for the video paths see figure 4.5.

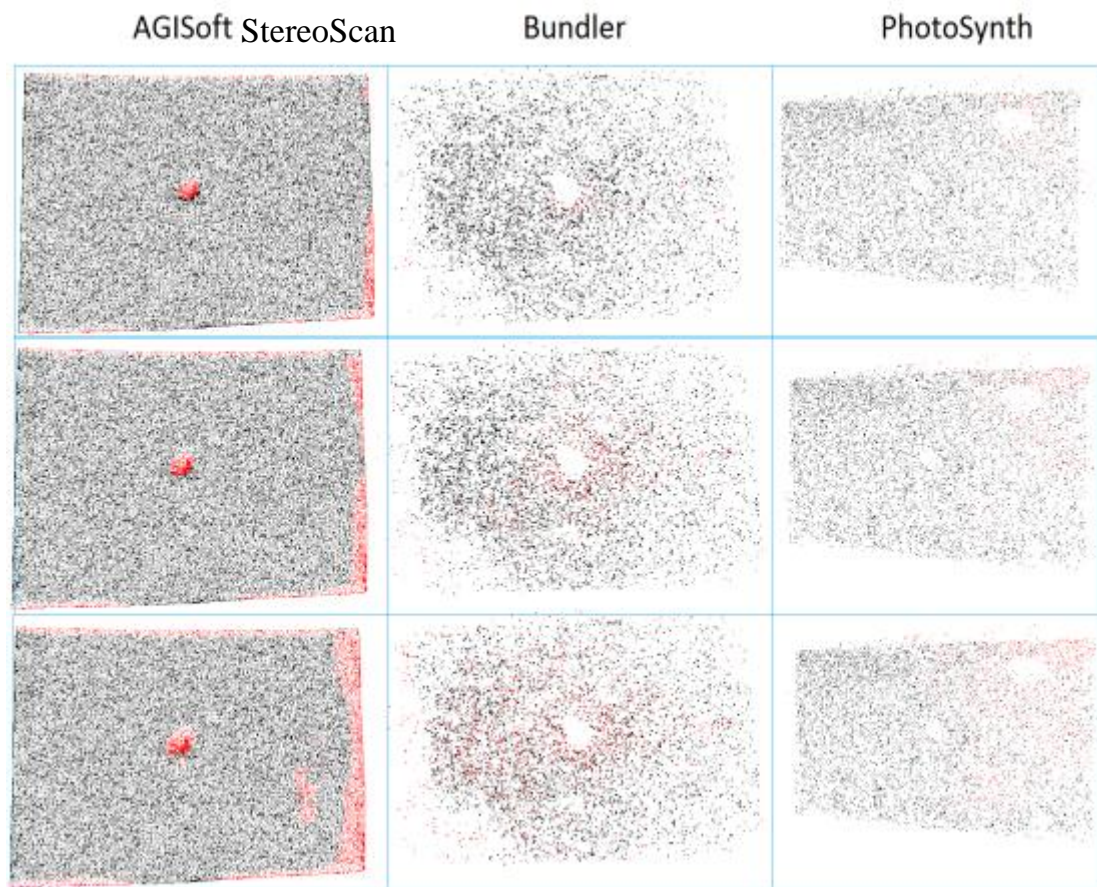


Figure 5.22: 3 sigma, 2 sigma, and 1 sigma exclusion criteria from (top to bottom)

### 5.7.4 Trial 4: Investigating the use of video images for pothole detection

In previous sections, it was shown that a bump or a hole as small as 5 cm<sup>3</sup> can be detected using AGISoft StereoScan software. In this section, the AGISoft StereoScan software will be tested for digital images extracted from video files. The data used for this experiment were the video files listed in table 4.1. The procedure of video files collection, handling and image extraction were all extensively discussed in section 4.8.1. Due to the huge number of extracted images from video files, the extracted video frames with existing potholes were visually checked for further investigations. Nine samples of pothole existence images were extracted from the different videos, with various shapes and sizes that were used for the purpose of testing the 3D modelling for anomalies detection, as shown in figure 5.23.

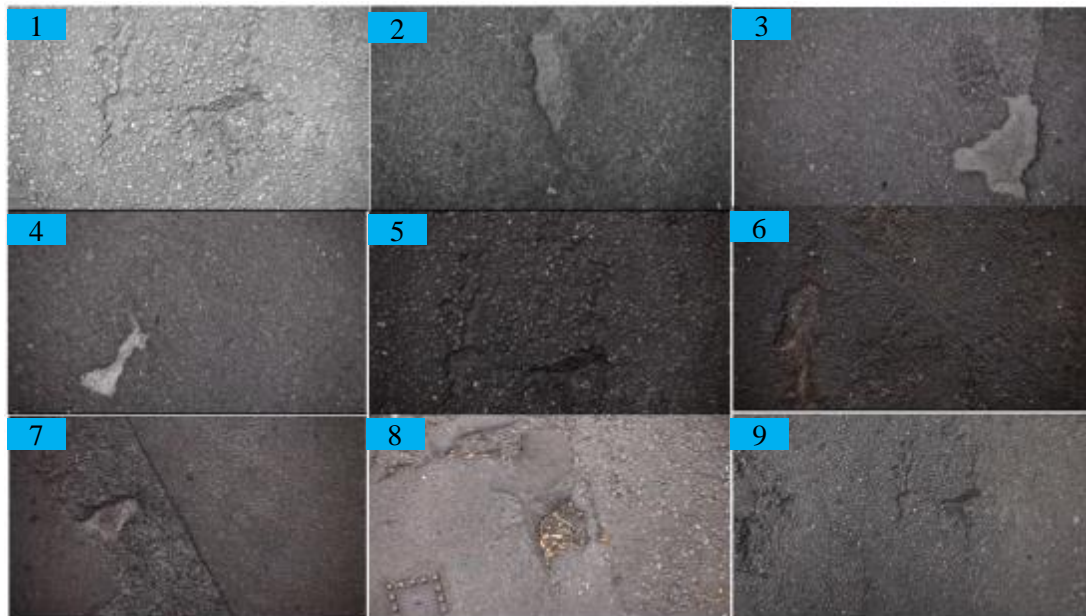


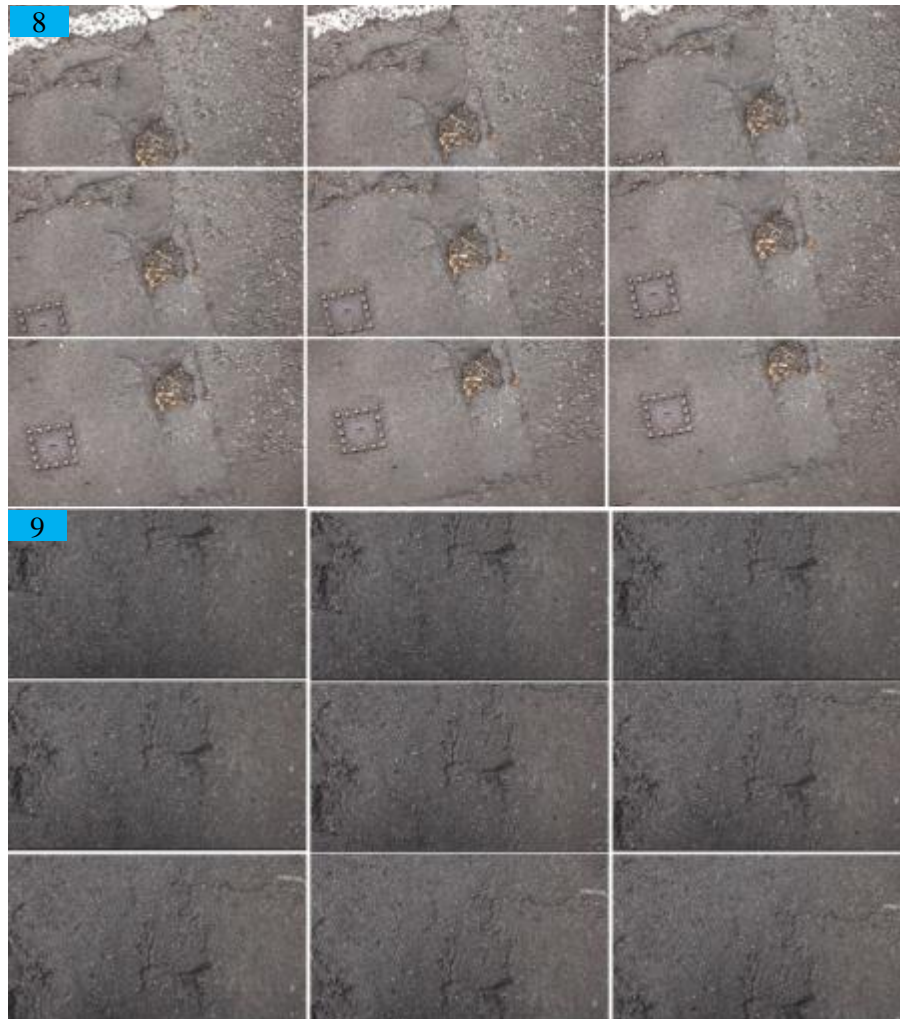
Figure 5.23: Extracted frames from video files with pothole existence

#### 5.7.4.1 Using only two images

The process of reconstructing 3D models from two images using AGISoft StereoScan is described in section 3.7.2.2.1. Two samples (8 and 9) from the samples shown in figure 5.23 were used for testing the effect of intersection angle or the base to height b/h ratio effect on the accuracy of the 3D reconstructed surface. All frames that share the pothole were used for testing the intersection angles, as can be seen from figure 5.24. The results from the two groups showed that as the baseline between the two image frames increased, the geometry of the 3D models accuracy was improved. The results

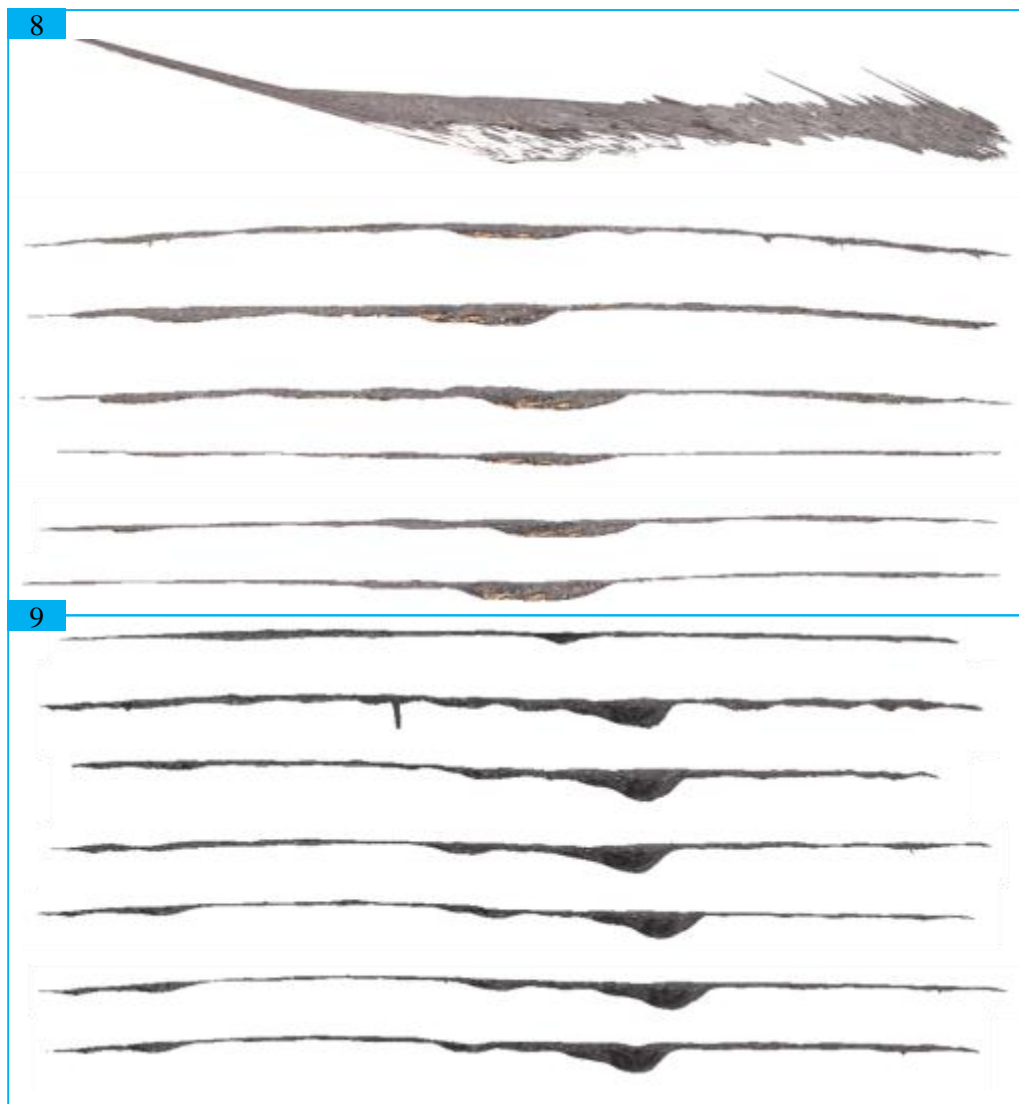


showed that it is possible to detect the pothole when the baseline between the two is about 60 cm or more.



**Figure 5.24: samples used for testing the effect of base to height ratio on the results**

A visual inspection of the results showed that there is a change in the geometry of the 3D models as the baseline increased. As figure 5.25 shows, the 3D models with low baseline between frames are showing that the geometry is not correct. These results revealed that the best baseline to be used when utilizing two images in AGISoft StereoScan is more than 60 cm. All results from these 3D models (prepared with different baselines) were processed in the Cloud Compare (CC) to build models with the colour visualization of ramp height differences, figure 5.26 and figure 5.27. As can be noticed from both figures, the details of the potholes become clearer as the translation between frames become bigger. This result further confirms that the geometry of 3D models can become better with the increasing baseline between the pair of images.



**Figure 5.25: 3D models generated from processing 1<sup>st</sup> frame with the next following frames (1<sup>st</sup> with 2<sup>nd</sup> on top and 1<sup>st</sup> with 8<sup>th</sup> on bottom)**

It is important to bear in mind that enough translation between the frames is necessary in order to reconstruct dependable and accurate 3D models. However, figure 5.26 and figure 5.27 show that the reconstructed area between the two frames was reduced as the translation increased. Matching points as well decreased with the increase in translation or shift between the two images which led to an increase in the RMSE value as can be noticed from figure 5.28. Therefore, it is recommended to use frames that provide enough translation that make a balance between accuracy and provide enough matching points with adequate reconstructed area, which is possible using the first extracted frame with the fifth or the sixth which means that the distance or shift range between 50 to 70 cm.

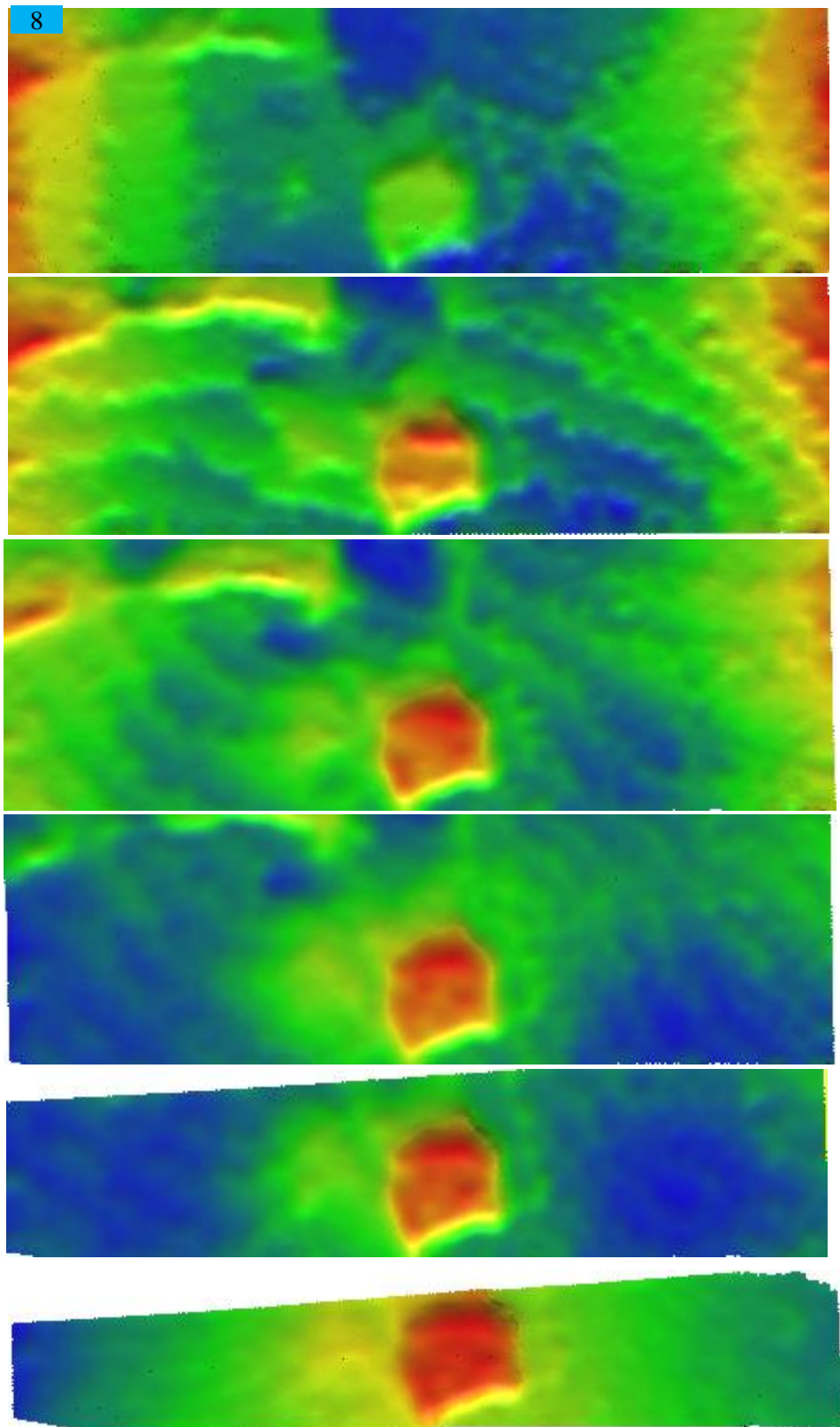


Figure 5.26: Height ramp colour visualization on 3D models frames (1<sup>st</sup> with 3<sup>rd</sup> on top and 1<sup>st</sup> with 8<sup>th</sup> on bottom)

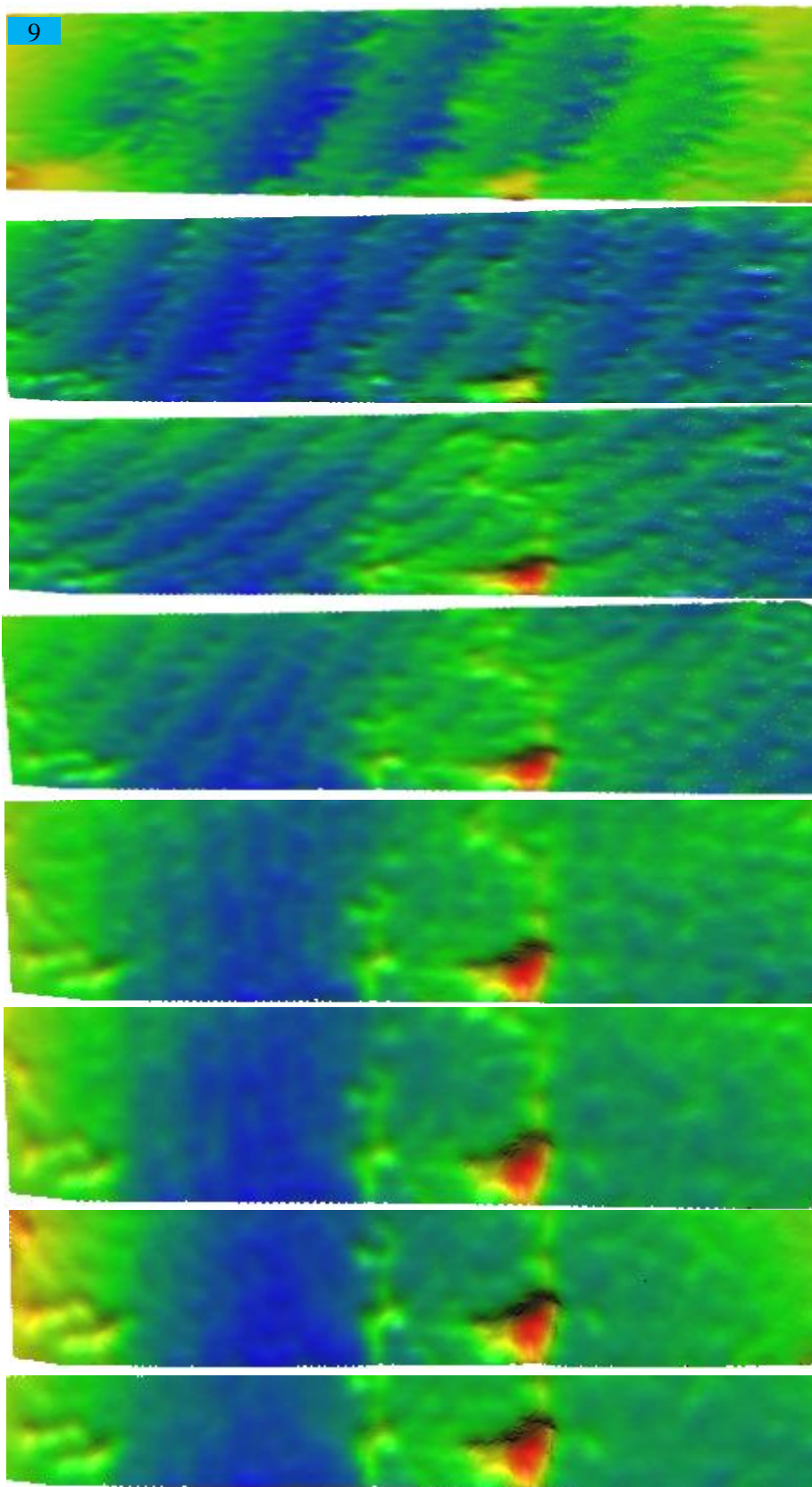


Figure 5.27: Height ramp colour visualization on 3D models frames (1<sup>st</sup> with 2<sup>nd</sup> on top and 1<sup>st</sup> with 8<sup>th</sup> on bottom)



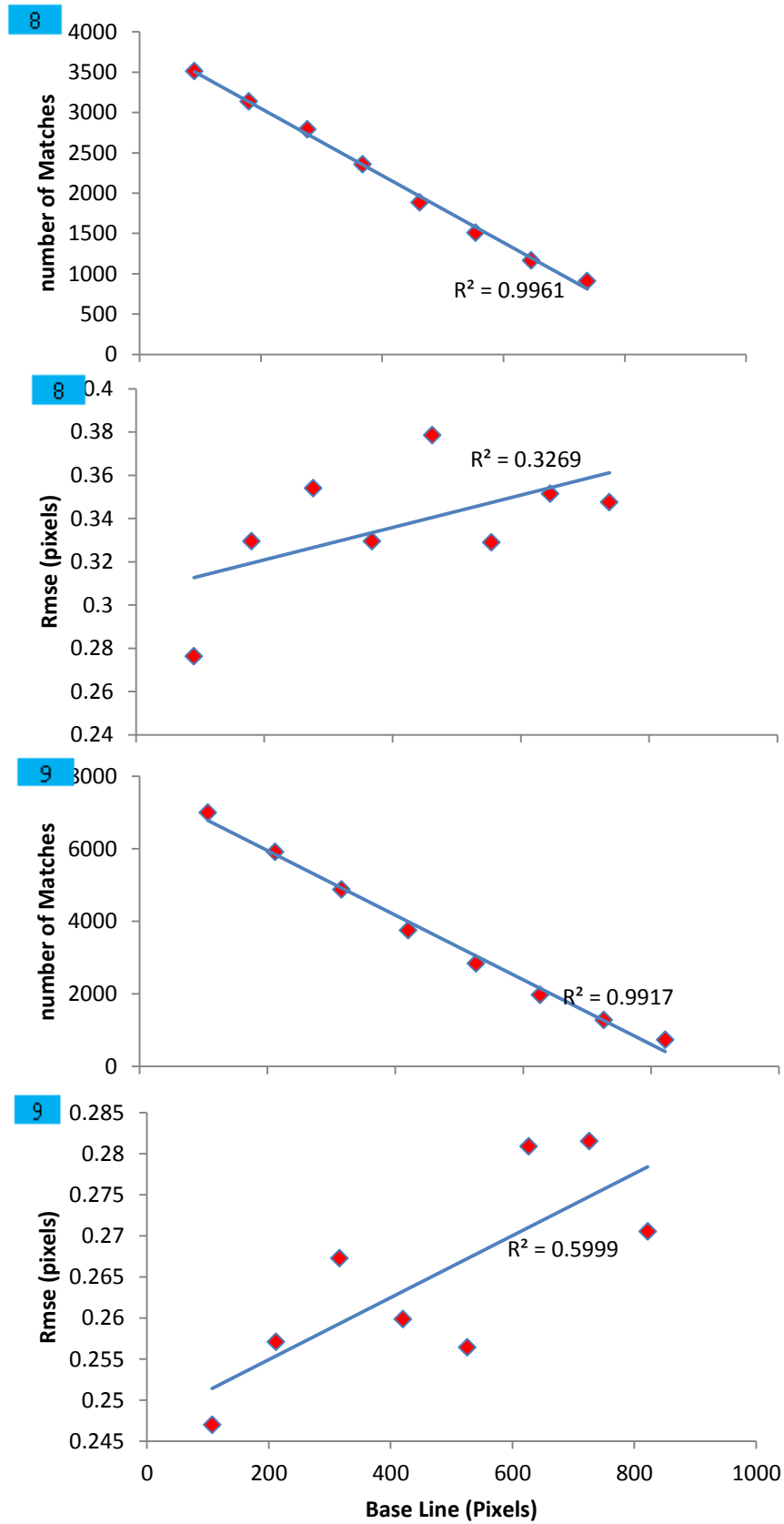


Figure 5.28: Matching and RMSe vs base line or translation between two frames

As trials showed, the methodology of detecting potholes on the road surface using a generated 3D surface from only on pair of images was successful. Therefore, all nine samples shown in figure 5.23 were used for generating 3D models based on two frames that provided sufficient translation. The translation between frames has to be between 50 to 70 cm which can be calculated using the motion estimation algorithm that was developed as part of this project; for more information about this algorithm and its results refer to sections 4.8.1.1 and 4.9.1.1.

Colour visualization of height or depth on the 3D samples was prepared and showed the possibility of detecting the potholes on all the tested samples. As figure 5.29 shows, the red colour is an indication of the pothole's existence which can be detected using the developed methodology of pothole detection. The fitted LOWESS planes to 3D models were all done using a Matlab developed algorithm mentioned on section 5.6.2.2. The results from the fitted plane and the calculated residuals at 3-sigmas with cluster checking showed that each sample has at least one cluster of object that is treated as a pothole, figure 5.30. These results revealed that the methodology of pothole detection using 3D automated modelling from two images and the fitted LOWESS plane can be used successfully. However, it is possible that these results can be further improved by using multiple images rather than just a stereo pair as will be explained in the next section.

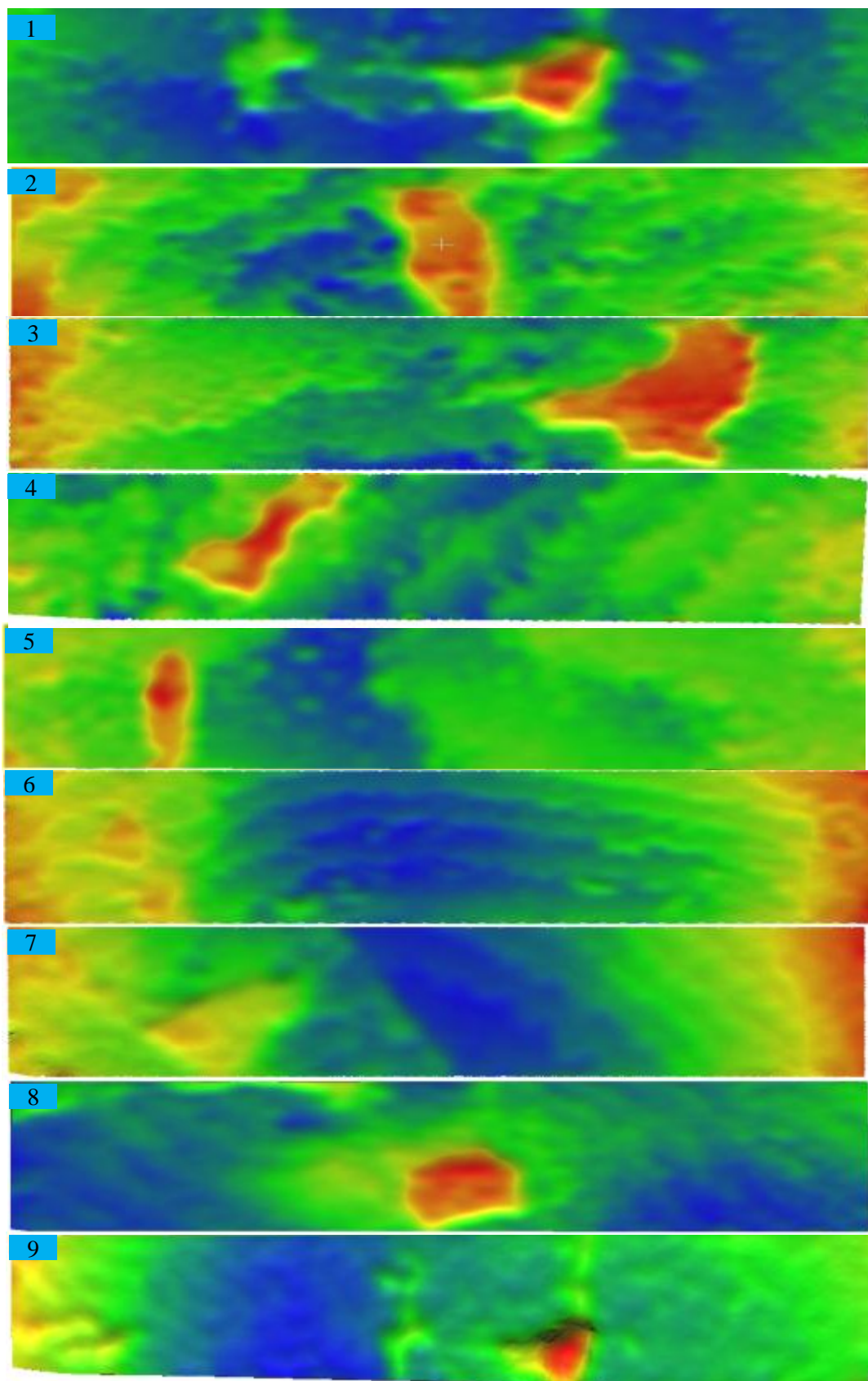


Figure 5.29: Colour visualization of ramp height on 3D models generated using two frames only

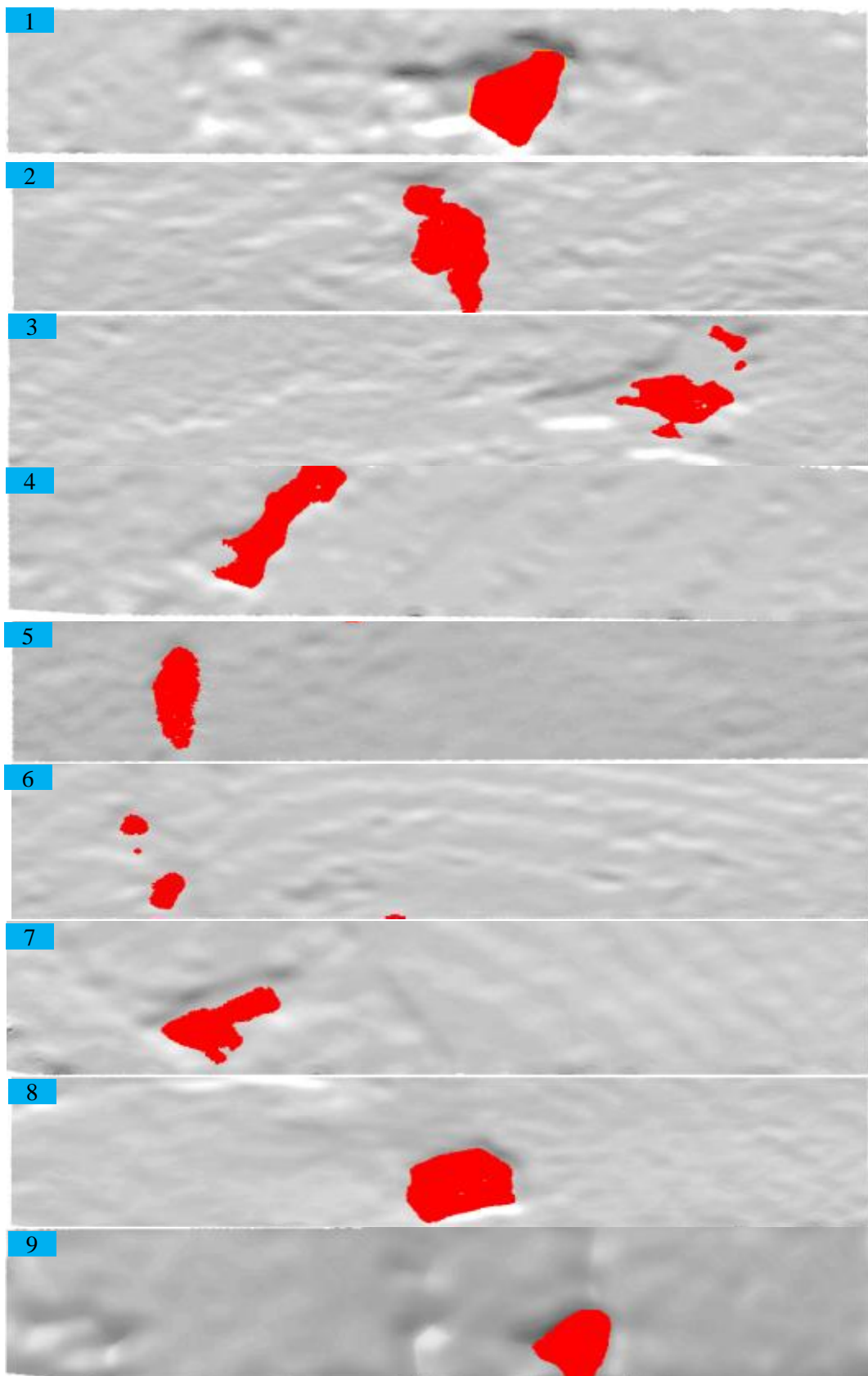


Figure 5.30: 3 sigma criteria for pothole detection on all samples using only two frames

#### 5.7.4.2 Using multiple images

In the previous section, the possibility of detecting the potholes using only two images was shown. In this section, the aim is to examine the potential of including all of the frames that share the pothole instead of just two frames. For achieving this purpose, the commercial standard version of AGISoft PhotoScan was utilized for two samples (8 and 9). The procedures of reconstructing 3D models from multiple images using AGISoft PhotoScan is described in section 3.7.2.2.1. Two samples (8 and 9) showed in figure 5.24 were again tested for the effect of multiple images, examining the accuracy of 3D modelling and possibility of pothole detection.

The results of 3D models from both samples were processed in Cloud Compare (CC) for visualization of height or depth on 3D models as can be noticed from figure 5.31. The results of height ramp visualization on the models showed much improvement when compared with the two images results. More geometrical details can be noticed in the case of multiple images that were not possible in the case of stereo pair reconstruction. However, the aim was to detect the pothole on the road surface, which was successfully done in case of two images. Testing the methodology of pothole detection described in section 5.6.2.2 on these two models showed the possibility of detecting the potholes in both samples at three sigmas exclusion criteria as can be noticed from figure 5.32.

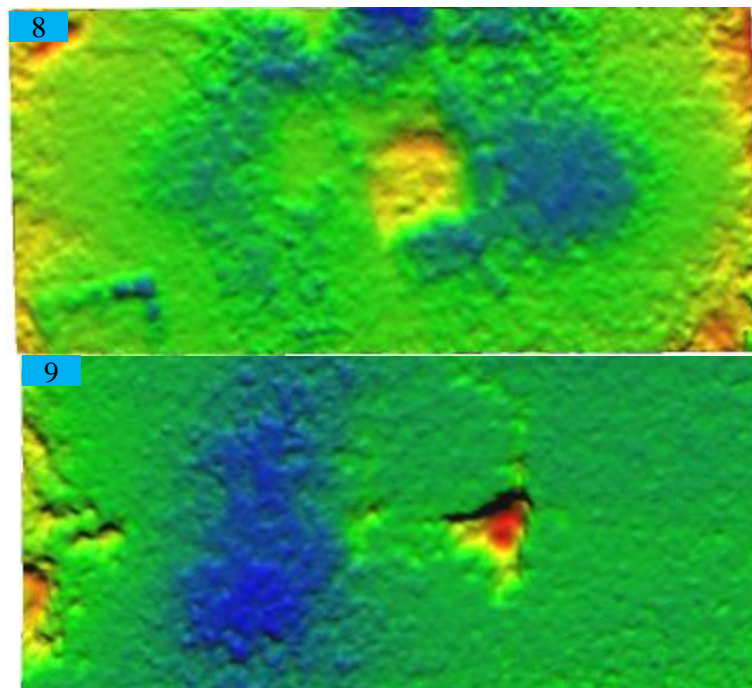


Figure 5.31: Height ramp visualization on 3D models using multiple images



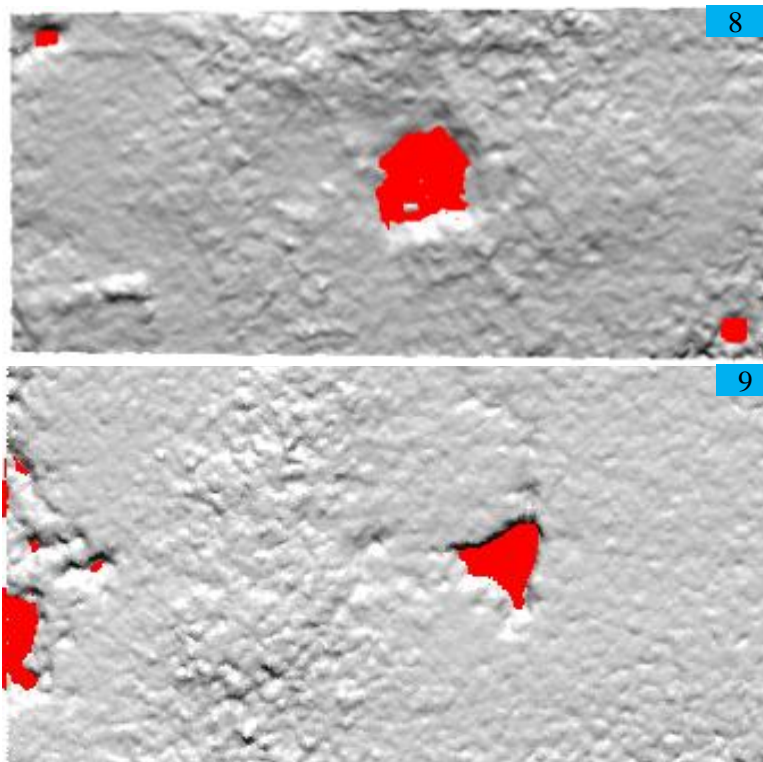


Figure 5.32: potholes detected at 3 sigmas away from mean plane

## 5.8 From semi towards full automation of pothole detection

As previously discussed in section 4.10, the main aim of the thesis was to build a cost effective system to automatically detect road damages such as cracks and potholes. The system was useful in collecting road images at HD quality covering a lane as previously mentioned. An algorithm was developed by author to extract images from video files; details are available on chapter 4. Due to limited time, the aim of full automation was not achieved for both crack and pothole detection. However, as discussed in section 4.10, it is possible to make the detection process fully automatic for both crack and pothole through linking different code and algorithms.

In pothole detection, the main key stages can be summarized below, see figure 5.33:

- 1- AGISoft StereoScan was used to generate 3D surfaces (using stereo process on pair of frames with enough base length)

- a. **Input** to this process were extracted video frames with pothole existence, images with pothole existence were visually checked to save the processing time, otherwise a **link** is required to use the images with pothole existence through 2D image analysis for rough identification of images or frames with potholes.
  - b. **Output** from this process is point clouds
- 2- Fit point clouds to a plane to minimize the residuals between point clouds and fitted plane.
  - a. **Input:** point clouds from first step (a **link** is required to use 1.b data).
  - b. **Output:** residuals calculated between point clouds and fitted plane.
- 3- Detection of potholes or any deviation using 3-sigmas away criteria
  - a. **Input:** residuals from second step (a **link** is available).
  - b. **Output:** cluster of points that are possibly potholes or bumps.

The term link refers to a subroutine that need to be designed in Matlab linking each step with a manual interface were required. It is possible in the future work to design a main program that can do all the procedures fully automatically without human intervention. It is worth mentioning that all algorithms were designed using Matlab except for AGISoft StereoScan. An executable version of the program is required in order to be used in the computer system which can be called in Matlab whenever needed. It is worth mentioning that the aim of this project was to design a survey tool for crack and pothole detection, the quantification of crack and potholes were not aimed in this project. However, it may be possible to quantify the detected crack and pothole to a usable measuring unit such as the severity in term of three categories (good, moderate, and bad) and assign each of them with a specific colour (green, orange, and red), respectively. The quantification could be in terms of road condition index (RCI) which is usually used in UK, more details can be found on section 2.3.4.4.1.

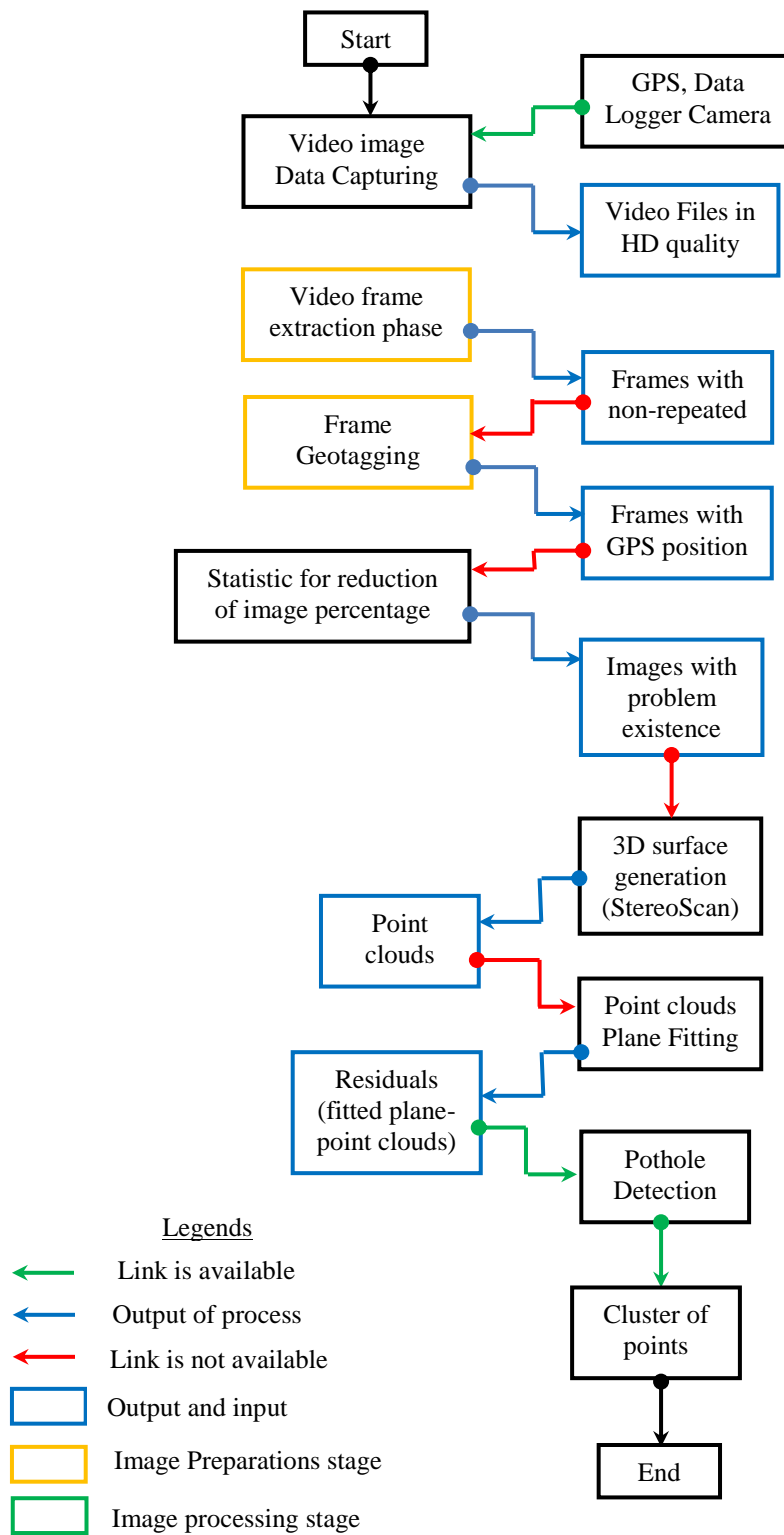


Figure 5.33: Main key stages in pothole detection (toward fully automation)



## 5.9 Summary and conclusions

Recently, different digital image matching algorithms have been developed and adapted in many 3D automated image modelling applications. Both SIFT and SURF are two matching algorithms that are invariant to scale and rotation changes. In the current study, these two digital matching algorithms were tested and compared with a cross-correlation matching algorithm that is utilized by LPS. Tests showed that both SIFT and SURF can produce enough matching points, comparable with cross-correlation matching. However, it was also proved that using SIFT and SURF for simply matching between the features of interest was not accurate enough, as it was found that lots of outliers fed into the solution as the number of tie points included in the bundle solution was increases. The accuracy of tie points using SIFT and SURF was subsequently improved through employing outlier rejection algorithms, such as RANSAC and LMedS, and it was shown that accurate matching points could be achieved that were comparable with the case of cross-correlation.

Many image-based automated 3D modelling tools and open source softwares are utilizing SIFT and SURF in matching between a pair of images. A trial was carried out in order to make a comparison between these open source and available commercial photogrammetric softwares for bump detection. Experiments proved that AGISoft StereoScan and Photosynth are sufficiently accurate in identifying a bump with about 6.0 cm height. It was demonstrated that AGISoft StereoScan tool can provide an enormous number of cloud points in less than a minute of processing time, with only using two images. LPS, with the use of plenty number of GCPs and good triangulation results, was not capable of providing accurate point clouds, which might be attributed to the cross-correlation algorithms not being able to handle big rotation changes.

Examination was carried out in this chapter to test the possibilities of detecting holes as small as 5 cm<sup>3</sup> in the road surface, using again the open source and commercial softwares. Experiments showed that it was possible to successfully detect such a small hole, but only by using AGISoft StereoScan. Other tools such as Photosynth and Photomodeler proved lacking in being able to identify such a small object. This test revealed that AGISoft StereoScan outperformed all other softwares in providing accurate and reliable results from using only two images, in a very short processing time which is less than a minute, and in a fully automated basis. This outperformance might be

attributed to the use of SIFT with RANSAC for matching between two images, knowing that the used images have angle of tilt of about  $13^\circ$ .

In the last trials, an investigation was carried out to test the method of pothole detection using frames extracted from video files. Extracted video frames were considered to be low in resolution and of low radiometric quality. Nine samples with various pothole shapes and sizes were visually identified and extracted for further analysis. Both height ramp visualization and fitted plane with residuals calculations were considered for analysis. The inspection was carried out on two samples to examine the effects of intersection angle, or the base to height ratio, on the geometry of 3D reconstructed surfaces. It was possible to prove that low baseline or distance between extracted images may affect the accuracy of the resultant 3D models. The results also revealed that AGISoft PhotoScan used to generate 3D surfaces through the processing of multiple images can further improve the detection results.

Lastly, it is worth to mention that it is possible to automatically identify any source of problems such as potholes or bumps in the road surface using extracted video frames. However, it is also worth to mention that the potholes detected, based on three sigmas criteria did not consider measuring the pothole size. In future work, it is recommended to quantify the size of the potholes to a usable engineering index.

## **CHAPTER SIX: CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK**

This chapter summarizes the research presented in the thesis, providing conclusions about the results that were obtained, and presenting the research outcomes. It also provides recommendations for possible useful future work in the field of digital image processing for road surface monitoring. As stated in the introductory chapter, the aims of this study were to investigate the potential of using digital images for roughness or texture depth measurement and for detecting significant road damages such as cracks and potholes using a mobile vehicle. A number of objectives were carried out in order to achieve the primary aims of the research. The outcomes and findings from each objective are described and summarized in the next sections.

### **6.1 Investigating highway maintenance problems and highlighting limitations of monitoring techniques**

Road surface monitoring is essential in order to keep this infrastructure in the best operational order. The previous approaches that considered evaluation of pavement conditions were found to be time consuming, labour-intensive, and hazardous (during the process of data collection). The manual process involved a team of trained technicians to walk over the road and record the defects in specifically designed forms. Assessment of pavement conditions seems to be a serious challenge in different countries all over the world that raised the demand for developing automated systems to overcome the limitations of previous approaches. Current automated systems are different from one country to the other but are still sharing one objective that is precise data collection and monitoring. Despite the advancement in technologies and the introduction of new automated systems, the walk over survey is still used in many areas around the world due to the high cost of the automated systems, the importance of a cheap alternative system was, therefore, addressed and investigated.

Among all available monitoring techniques, such as laser scanners and penetrating radar, it was concluded from the literature that digital images remain are a dependable and affordable source. Defects, such as cracks, can be successfully recognized using digital images. However, despite the fact that the detection of the cracks has been the aim behind many researches in recent years, many problems still need to be addressed. Further

consideration and exploration of crack detection in rough texture surfaces is required because rough texture can easily hide crack pixels. The processed images in the rough texture may contain lots of mis-interpretations due to raised aggregates. Digital images can also be a useful tool for identification of other damages on the road surface, such as potholes. Many authors have used a single digital image for pothole detection based on 2D image analysis processing. The success rate of pothole identification by using individual image processing is based on the variation of both textures (pothole and surrounded areas). The inclusion of the third dimension may increase the chance of detection as investigated in this research.

An extensive review on the literature also addressed some gaps in the area dealing with measurement of road surface texture depth or roughness. It has been shown that no applicable and cost effective measuring technique was available for road surface texture characterising in the field. Texture characterising is an important factor for assessing the ride quality, noise control and skid resistance. It has been revealed that all previous approaches for texture analysis have limitations such as cost, providing enough information and the range measurement of the object. An example of the most widely used traditional approach used for texture depth measurement is the sand patch test (SPT). However, this method lacks accuracy (that is in the range of  $\pm 0.227$  mm), is a labour intensive process and can only be used on dry surfaces.

The rapid advancement in 3D automation methods from digital images has opened the door for many applications that require precision, such as texture characterizing or roughness measurement and pothole identification. Research into the automation of a 3D method for roughness measurement has revealed that no critical study compared between texture depth measured using digital images and SPT are available. A number of trials were designed to test the author's developed method of texture depth measurement and texture characterization exploiting advancements in the area of 3D automated methods based on digital images

In the area of pothole detection, 3D based image modelling could be a useful alternative tool. Using 3D models can overcome the limitations of single image analysis, as single image lacks recognition and cannot provide enough information about the pothole; especially in cases where no real differences exist between pothole and non-pothole textures. Different trials were proposed for assessing the 3D automated technique of pothole detection on real data collected via a mobile van using digital images.

## **6.2 Examine the potential of digital images from a variety of cameras for texture depth measurement**

The second objective was to assess the texture depth measurement using a variety of cameras and sensors. Recent technology has allowed for 5 MP cameras to be integrated with smartphones that can possibly help in the acquisition of better images. Smartphone cameras are an essential part of every individual in the society and are almost always with user. Due to the availability of smartphone cameras, these were investigated along with other DSLR cameras for texture depth measurement. Five cameras, two DSLR cameras with three lenses and two smartphone cameras were used as a part of investigations; see section 3.6.1 for details of the used cameras. The objective was to examine texture depth measurement from digital images for various materials, particularly the asphalt and concrete textures. Experiments were carried out in order to inspect the use of digital images for texture depth measurement.

Many of the DSLR cameras with their lenses available in NGI shelves were tested for lens quality as the first objective. It was concluded that image sharpness is mainly due to the lens quality and it was also proved that some old lenses provided poor results. Inadequate image sharpness possibly being due to sensor excessive use or due to any source of accident that may degrade the quality of the lens.

Finding the optimum intersection angle and the camera to object distance through the stereo process was another objective. Two controls of asphalt texture (rough and smooth) were employed for performing these trials. The findings from this inspection revealed that the best intersection angle is between 30-50° and the distance of the camera to object was found to be as close as possible with the camera not being out of focus. Due to the low-quality results from the stereo process, a method of texture depth measurement was examined using multiple images. The reliability of the stereo method was found to be low, especially for both smartphone cameras. Therefore, the author's developed method of texture depth measurement was performed using multiple images. The results from testing two control textures (rough and smooth) showed that the calculated texture depth was significantly improved to the largest error of about 200  $\mu\text{m}$ , see section 3.8.2.2.

For further validation of the author's developed texture depth measurement method, wider depths of asphalt texture including ten test areas were examined using

both the stereo and multiple images processes. Using the recommended optimum camera to object distance and intersection angle via stereo process, mean errors as large as 355  $\mu\text{m}$  were reported. Whereas, the multiple images process exposed that the largest mean error became about 110  $\mu\text{m}$ . Trials of measuring the correlation between depths from images and the sand patch test were also considered. The outcome from these experiments showed a high significant correlation with a minimum value of about 0.80.

Another objective of this study was to examine other materials rather than asphalt pavements. Reconstruction parameters such as the number of extracted features, matches, RMSE of the least square solution and visual inspection of generated 3D surfaces were considered as examination parameters. These parameters were used for testing different materials such as a concrete tile, gravel with two sizes, soil, grass, and asphalt pavement. It was found from these trials that the process of 3D automation is mainly based on the feature characterises of the material. Repetitive features such as gravel and grass texture can lead to failure in the 3D automation process. Based on these results, it was decided to investigate the author's developed texture depth measurement method on concrete texture. The multiple image process for surface generation was used on two control samples of concrete texture (rough and smooth). Different analysis approaches were used on the generated surfaces to investigate the quality of the technique. The outcome from this trial showed that the method is successful with concrete as well as asphalt.

Finally, it can be concluded that the author's developed methodology of texture depth measurement and surface characterization from digital images is showing promising results, especially for the multiple image process. An alternative method for SPT is therefore proposed, and results of highly significant correlation were obtained.

### **6.3 Investigate the potential of detecting the cracks on the road surface from digital SLR extracted video frames**

A methodology was proposed to be an alternative for commercial non-affordable systems based on low-cost materials to collect images from the road surfaces. The approach was established to extract the desired frames from video files using image matching techniques. Subsequent steps included geotagging of the frames using a GPS log and detecting the cracks from the extracted frames using the author's developed crack detection algorithm. The method is based on roughness modelling using 2D image and texture

analysis of the extracted frame in order to apply the proper smoothing filter. Roughness measurement trials were carried out on samples with known surface texture depths in order to evaluate the process of roughness modelling. The author's developed crack detection algorithm was tested on still digital image data collected from the real world, the net source and published articles. Finally, the algorithm was used for a number of frames with different problem existences such as shadows, manholes, different illumination and lighting conditions.

Through algorithm development, it was possible to show that detection of cracks and the isolation of non-crack pixels is possible using 2D image analysis. Thresholding, morphological operation and image filtration were the tools used to develop the author's crack detection algorithm. Applying the proper image filtering was based on the roughness measurement of the image. Data handling was an important part of the algorithm in order to reduce the time of processing efficiently. Motion estimation based on image matching was used effectively to predict the distance between the frames. Further time management was achieved through statistical testing to minimize the amount of frames and to include only the frames that have a real problem such as a shadow, manhole, paint, cracks, and pothole existence.

Tests for examining the success rate of the developed algorithm on different sources of images revealed that the success rate of crack detection reached up to 98.9% on one example of the image, with the lower rate detected to be as low as about 58% on other images. The average success rate of crack detection was found to be about 83% for all tested image samples. However, it was demonstrated that the author's developed crack detection algorithm can erroneously identify non-crack pixels and treat them as crack candidates, but the average false-positive for all samples was found to be about 14%. The crack detection success rate of the developed algorithm was also tested using video extracted frames from real world data. The study proved that the success rate declined due to the low quality of radiometric and contrasts of video images. Testing twelve extracted frames with various problem existences showed that the average success rate of the algorithm dropped from about 83% to about 70% and that the average false-positive was increased from 14% to about 39% when using video not still images.

In summary, the trials showed that it is possible to assemble a series of low-cost tools of camera and cheap sensors for monitoring and detecting the cracks on the road surface. It was also possible to prove that using small format images for road surface monitoring has

some potential. For example, the cracks on the road surface can be identified using a semi-automated crack detection algorithm from still images, and even from images extracted from video-sequences with low resolution quality, albeit with a slight degradation in performance.

## **6.4 Investigate the potential of detecting potholes on the road surface using the extracted frames from video DSLR images**

The use of low-contrast and radiometry digital video images for road surface monitoring made the work with commercial photogrammetric tools more complicated. This could be attributed to the fact that these tools were designed for terrestrial or standard aerial photogrammetric applications. Trials showed that Leica Photogrammetric Suit (LPS) and Photomodeler are not working well when the camera is close to an object or when a small rotation is available. LPS, Australis, and Photomodeler require camera calibration parameters, and LPS requires at least 3 ground control points GCPs to run aerial triangulation (AT). In LPS, the setup of the block file and camera calibration requires time, and the process depends on the number of images. Without knowledge of the exterior orientation parameters (EOPs) in LPS, the operator needs to include manually at least five tie points for each image pair in order to achieve approximately the relative-orientation and thus help in automatic tie point generation.

AGISoft StereoScan and Photosynth however allow running the triangulation procedure without any manual intervention at all. The success behind that is the use of invariant feature matching such as SIFT and SURF; although outlier removal such as RANSAC was important to ensure the correct matches are in the final solution. Trials showed that using AGISoft StereoScan software and Photosynth can detect the defect and the developed mythology of pothole detection showed that a minimum depth of pothole of 1.5cm could be found using the idea of 3D based image modelling from images.

## **6.5 Recommendations for future work**

As discussed before, correlations between SPT and texture depth measurement from digital images were significantly high. The correlation was carried out using only ten samples collected from real world data while considering as different a range of



texture depths as possible. However, using only ten samples might not be enough for correlation measurement. It is recommended to include a further extensive range of texture depths, especially for depths of 1100 to 1600  $\mu\text{m}$  and 2000 to 3000  $\mu\text{m}$  to cover the whole range of depths in order for the results to be more reliable.

It was possible to show that the small and compact cameras integrated with current smartphones can provide average texture depth results with as good as 100 $\mu\text{m}$  if compared with SPT. Therefore, software needs to be developed for the smartphone cameras to measure and characterize the road surface texture on a fully automated basis without requiring intervention from the users. Targets can be designed and placed with a known distance in order to ease the process of scaling the 3D surface into a usable engineer index or units.

Currently, the pothole areas are manually identified by an operator using the exported point clouds generated from the 3D automated tools. An automated method needs to be developed in order to integrate the generated 3D surface model with image analysis data to detect, extract, and measure surface distresses. Firstly, image analysis technique should be used to identify roughly the video image frames with pothole existence, and then the successive frames with potholes should be extracted and used for point cloud generation using AGISoft StereoScan. The extraction and measurement of the defect would be then based on an automated basis without the needs for operator intervention.

It could be shown that new feature matching algorithms such as SIFT and SURF, with outlier rejection such as RANSAC, can reach the accuracy of the traditional cross-correlation commonly used in current photogrammetric softwares such as LPS. It is recommended to use these new feature matching algorithms (SIFT and SURF) instead of the traditional area based matching used in traditional photogrammetric softwares such as LPS, Australis, and Photomodeler. This might help to increase the productivity, reduce the processing time and automate the relative-orientation process. A merge between computer vision and photogrammetry is also required in order to take advantage of both.

Finally, it was shown that a semi-automatic approach was achieved in both crack and pothole detection, in which man intervention was required in some processing stages. However, it was discussed that full automation of the processes would be possible

through linking different algorithms and tools in to one main program. It is recommended in the future to develop a main interface program using Matlab to connect between the different stages of procedures. It may also possible to speed up the processing time of the crack and pothole detection programmes through translating it to C code using the C++ library.

## REFERENCES

- ABDULLAH, Q. A. 2010. *Mapping Matters* [Online]. Available: [http://www.asprs.org/a/mapping\\_matters/](http://www.asprs.org/a/mapping_matters/) [Accessed 3 March 2014].
- ABRAHAM, S. & HAU, T. 1997. Towards autonomous high precision calibration of digital cameras. *Videometrics V, Proceedings of SPIE Annual Meeting*, 3174, 82-93.
- ACOSTA, J. A.; FIGUEROA, J. L. & MULLEN, R. L. 1992. Low-cost video image processing system for evaluating pavement surface distress. *Journal of the Transportation Research Board—Transportation Research Record*, 1348, 63-72.
- AGISOFT. 2015. *Image-based 3D modeling* [Online]. Available: <http://www.agisoft.ru/products/lens/> [Accessed 18 April 2013].
- AGÜERA; CARVAJAL; PÉREZ & ABZAL ALI 2011. Measuring Sunflower Nitrogen Status from an UAV-based System and on the Ground Device. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-1/C22, 33-37.
- AHMED, H. N. 2012. *Microstructural characterization of rubber modified asphalt mixtures*. PhD thesis, University of Nottingham
- AIRY, G. 2012. Pavement Evaluation. *Course handouts, University of Nottingham*.
- ALEX. 2011. *Mooneapolis, A.D. 2030* [Online]. Available: <http://gettingaroundmpls.wordpress.com/category/equity/> [Accessed 20 Dec. 2013].
- APPLIED PAVEMENT TECHNOLOGY. 2008. *Circular Texture Meter (CT Meter)* [Online]. Available: [http://www.appliedpavement.com/techResources\\_equipLoanProg\\_ctmeter.html](http://www.appliedpavement.com/techResources_equipLoanProg_ctmeter.html) [Accessed 25 July 2014].
- ASPHALT INSTITUTE. 2009. *Asphalt Pavement Distress Summary* [Online]. Available: [http://www.asphaltinstitute.org/public/engineering/maintenance\\_rehab/asphalt-pavement-distress-summary.dot](http://www.asphaltinstitute.org/public/engineering/maintenance_rehab/asphalt-pavement-distress-summary.dot) [Accessed 23 Dec. 2013].
- ASTM 2004. Standard Test Method for Measuring Pavement Macrot texture Depth Using a Volumetric Technique. *ASTM Designation E965*. American Society for Testing and Materials.
- ASTM D 6433 – 07 2007. Standard Practice for Roads and Parking Lots Pavement Condition Index Surveys1. *ASTM International-Active Standard 2007*, PA 19428-2959.
- AUGEREAU, B.; TREMBLAIS, B.; KHOUDEIR, M. & LEGÉAY, V. 2001. A differential approach for fissures detection on road surface images. in *Proceedings of the International Conference on Quality Control by Artificial Vision*.
- AVANGATE. 2013. *Quick MTF, an image quality testing application* [Online]. Available: <http://www.quickmtf.com/> [Accessed 20 July 2013 ].
- BARAZZETTI, L.; MUSSIO, L.; REMONDINO, F. & SCAIONI, M. 2011. Targetless camera calibration. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-5/W16, 335-342.
- BARAZZETTI, L.; REMONDINO, F. & SCAIONI, M. 2010. Automation in 3D reconstruction: results on different kinds of close-range blocks. *International Archives of Photogrammetry and Remote Sensing*, 38, 55-61.

- BARNSDALE K.P. & ANDREOTTI, M. T. 2008. GRC Precise Time Data Logger PTDL1: User Manual. *Geospatial Research Centre*.
- BAY, H.; ESS, A.; TUYTELAARS, T. & VAN GOOL, L. 2008. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110, 346-359.
- BBC. 2012. *MWC 2012: Nokia reveals 41MP cameraphone* [Online]. Available: <http://www.bbc.co.uk/news/technology-17178014> [Accessed 13 Mar 2012].
- BENTLEY. 2014. *Bentley Pointools V8i, Uniquely Powerful and Intuitive Pre-processing of Point Clouds* [Online]. Available: [http://www.bentley.com/en-US/Promo/Pointools/pointools.htm?skid=CT\\_PRT\\_POINTOOLS\\_B](http://www.bentley.com/en-US/Promo/Pointools/pointools.htm?skid=CT_PRT_POINTOOLS_B) [Accessed 23 Jan. 2014].
- BESHR, A. A. A. & ABO ELNAGA, I. M. 2011. Investigating the accuracy of digital levels and reflectorless total stations for purposes of geodetic engineering. *Alexandria Engineering Journal*, 50, 399-405.
- BITELLI, G.; SIMONE, A.; GIRARDI, F. & LANTIERI, C. 2012. Laser scanning on road pavements: a new approach for characterizing surface texture. *Sensors*, 12, 9110-9128.
- BLAJOVICI, C.; KISS, P. J.; BONUS, Z. & VARGA, L. 2011. Shadow detection and removal from a single image. *19th Summer School on Image Processing*. Szeged, Hungary.
- BRIESE, C.; GLIRA, P. & PFEIFER, N. 2011. Assembling and operating a UAV. *course Handouts, Institute of Photogrammetry and Remote Sensing*. Helsinki Vienna University of Technology.
- CALLIERI, M.; CIGNONI, P.; DELLEPIANE, M. & SCOPIGNO, R. 2009. Pushing Time-of-Flight Scanners to the Limit. *Proceedings of the 10th International conference on Virtual Reality, Archaeology and Cultural Heritage*. Julians, Malta
- CHAMBON, S. 2010. Detection of road cracks with multiple images. *in Proceedings of the 5th International Conference on Computer Vision Theory and Applications*. Angers, France.
- CHAMBON, S.; SUBIRATS, P. & DUMOULIN, J. 2009. Introduction of a wavelet transform based on 2D matched filter in a Markov Random Field for fine structure extraction: Application on road crack detection. *Proceedings of SPIE, the International Society for Optical Engineering*.
- CHENG, H. D. & MIYOJIM, M. 1998. Automatic pavement distress detection system. *Information Sciences*, 108, 219-240.
- CHINA, S. & JAMES, D. E. 2011. Comparison of laser-based and sand patch measurements of pavement surface macrotexture. *Journal of Transportation Engineering*, 138, 176-181.
- CIC. 2012. *CAMERA EXPOSURE* [Online]. Available: <http://www.cambridgeincolour.com/tutorials/camera-exposure.htm> [Accessed 13 Mar 2012].
- CIVIL AVIATION AUTHORITY. 2013. *Unmanned Aircraft and Aircraft Systems* [Online]. Available: <http://www.caa.co.uk/default.aspx?catid=1995&pagetype=90> [Accessed 01 August 2014].
- CLIVE S, F. 1997. Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52, 149-159.

- CLOUDCOMPARE. 2013. *3D point cloud and mesh processing software Open Source Project* [Online]. EDF. Available: <http://www.danielgm.net/cc/> [Accessed 4 May 2014].
- CSIRO. 2010. *Automated detection of road cracks* [Online]. Australia. Available: <http://www.csiro.au/Outcomes/ICT-and-Services/People-and-businesses/RoadCrack.aspx> [Accessed 27 April 2014].
- DEPARTMENT OF TRANSPORT. 2011. *SCANNER surveys for Local Roads (User Guid and Specification)* [Online]. Available: [http://www.pcis.org.uk/iimni/UserFiles/Applications/Documents/Downloads/SCANNER%20and%20TTS/SCANNER%20Specification/SCANNER\\_Spec\\_2011\\_Volume\\_3.pdf](http://www.pcis.org.uk/iimni/UserFiles/Applications/Documents/Downloads/SCANNER%20and%20TTS/SCANNER%20Specification/SCANNER_Spec_2011_Volume_3.pdf) [Accessed 3 March 2013].
- DOP:TRANSPORT & INFRASTRUCTURE. 2012. *Determination of Average Texture Depth of a Pavement Surface Using the Sand Patch Method* [Online]. Available: [www.dpti.sa.gov.au/\\_data/assets/word\\_doc/0005/47255/TP346.DOC](http://www.dpti.sa.gov.au/_data/assets/word_doc/0005/47255/TP346.DOC) [Accessed 22 October 2013].
- DUTTKE, J. 2009. *PhotoME Digital Photo Metadata Editor* [Online]. Available: <http://www.photome.de/> [Accessed 25 Feb. 2014].
- EATON, R. A.; JOUBERT, R. H. & WRIGHT, E. A. 1981. *Pothole Primer. A Public Administrator's Guide to Understanding and Managing the Pothole Problem* [Online]. DTIC Document. Available: <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA107294> [Accessed 19 Sep. 2013].
- ENGINEERING, G. 2013. *Greenwood - profiling roads* [Online]. Denmark. Available: <http://www.greenwood.dk/road.php> [Accessed 16 Dec. 2013].
- ERDEM, S. & BLANKSON, M. A. 2013. Fractal–fracture analysis and characterization of impact-fractured surfaces in different types of concrete using digital image analysis and 3D nanomap laser profilometry. *Construction and Building Materials*, 40, 70-76.
- ERDEM, S.; DAWSON, A. R. & THOM, N. H. 2012. Impact load-induced micro-structural damage and micro-structure associated mechanical response of concrete made with different surface roughness and porosity aggregates. *Cement and Concrete Research*, 42, 291-305.
- ERIKSSON, J.; GIROD, L.; HULL, B.; NEWTON, R.; MADDEN, S. & BALAKRISHNAN, H. 2008. The pothole patrol: using a mobile sensor network for road surface monitoring. *In Proceedings of the 6th international conference on Mobile systems, applications, and services, MobiSys (2008)*. Breckenridge, CO, USA
- ESS 2011. Drones are on: UAVs for Survey Apps . *Engineering Surveying Showcase, issue two*. Stevenage, United Kingdom: PV Publications Ltd.
- FARO. 2014. *FaroArm® Platinum - Technical and Specification* [Online]. Available: [http://www2.faro.com/ap/resource/techsheets/sea/Platinum\\_Arm.aspx](http://www2.faro.com/ap/resource/techsheets/sea/Platinum_Arm.aspx) [Accessed 24 July 2014].
- FISHER, B. 1997. *The Fundamental Matrix* [Online]. Available: [http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/EPSRC\\_SSAZ/node22.html](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/EPSRC_SSAZ/node22.html) [Accessed 01 August 2014].
- FOREST, R. & UTSI, V. 2004. Non destructive crack depth measurements with ground penetrating radar. *in Proceedings of the 10th International Conference on Ground Penetrating Radar*, 799-802.

- FRICKER, P. & PACEY, R. 2005. Forward Motion Compensation (FMC) - Is It the same in the Digital Imaging World? *Photogrammetric Engineering & Remote Sensing* 1241-1242.
- FUGRO ROADWARE. 2013. *Infrastructure Asset Management Solutions* [Online]. Canada. Available: <http://www.roadware.com/> [Accessed 16 Dec. 2013].
- FURNESS, S. BARNES & A. WRIGHT. 2007. *Crack Detection on Local Roads-Phase 2* [Online]. . Available: <http://www.ukroadsliasongroup.org/en/asset-condition/road-condition-information/data-collection/scanner/scanner-research/crack-detection.cfm> [Accessed 26 July 2012].
- FWA, T. F. 2006. *The handbook of highway engineering*, CRC Press.
- GARMIN. 2004. *GPS 76 chartplotting receiver: Owner's manual & reference guide* [Online]. Available: [http://ferm.mtu.edu/resources/Manuals/GPSMAP76\\_manual.pdf](http://ferm.mtu.edu/resources/Manuals/GPSMAP76_manual.pdf) [Accessed 27 November 2013].
- GAVILÁN, M.; BALCONES, D.; MARCOS, O.; LLORCA, D. F.; SOTELO, M. A.; PARRA, I.; OCAÑA, M.; ALISEDA, P.; YARZA, P. & AMÍROLA, A. 2011. Adaptive road crack detection system by pavement classification. *Sensors*, 11, 9628-9657.
- GENDY, A. & SHALABY, A. 2007. Mean Profile Depth of Pavement Surface Macrottexture Using Photometric Stereo Techniques. *Journal of Transportation Engineering*, 133, 433-440.
- GEODETIC SYSTEMS. 2013. *The Basics of Photogrammetry* [Online]. Melbourne, USA. Available: <http://www.geodetic.com/v-stars/what-is-photogrammetry.aspx> [Accessed 30 Dec. 2013].
- GEOMAGIC 2010. Geomagic Studio 10: User manual help (Software help menu).
- GHOSH, S. K. 2005. *Fundamentals of computational photogrammetry*, Concept Publishing Company.
- GIE TECHNOLOGY. 2009. *Laservision* [Online]. Available: [http://www.gieinc.ca/main\\_en.html](http://www.gieinc.ca/main_en.html) [Accessed 05 May 2012].
- GOV.UK. 2014. *Report a pothole* [Online]. Available: <https://www.gov.uk/report-pothole/nottingham> [Accessed 27 July 2014].
- GRAPHPAD. 2014. *QuickCalcs* [Online]. Available: <http://www.graphpad.com/quickcalcs/ttest1.cfm> [Accessed 10 Dec. 2014].
- HAALA, N.; CRAMER, M.; WEIMERB, F. & TRITTLERB, M. 2011. Performance Test on UAV-based Photogrammetric Data Collection. *International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g): ISPRS*. Zurich, Switzerland.
- HALLOWAY, D. 2011. Calibration of Laser Devices used for the Measurement of Pavement Surface Texture: Accuracy of Pavement Texture Measurement *3rd International Surface Friction Conference*. Australia.
- HARVEY, P. 2013. *ExifTool (Read, Write and Edit Meta Information)* [Online]. Available: <http://www.sno.phy.queensu.ca/~phil/exiftool/> [Accessed 23 Mar 2013].
- HASEGAWA, H.; MATSUO, K.; KOARAI, M.; WATANABE, N.; MASAHARU, H. & FUKUSHIMA, Y. 2000. DEM accuracy and the base to height (B/H) ratio of stereo images. *International Archives of Photogrammetry and Remote Sensing*, 33, 356-359.
- HAWKER, J. D. K. E. A. L. G. 2003. *TRACS Type Surveys for Local Roads: Scoping Study* [Online]. Available: [http://www.pcis.org.uk/iimni/UserFiles/Applications/Documents/Downloads/SCANN ER%20and%20TTS%20Research/dft\\_roads\\_023919.pdf](http://www.pcis.org.uk/iimni/UserFiles/Applications/Documents/Downloads/SCANN ER%20and%20TTS%20Research/dft_roads_023919.pdf) [Accessed 06 March 2012].

- HEROLD, M.;ROBERTS, D. A.;SMADI, O. & NORONHA, V. 2004. Road condition mapping with hyperspectral remote sensing. *presented at the 2004 AVIRIS Workshop*. Pasadena, Calif, USA.
- HIGHWAY RESEARCH BOARD 1972. *NCHRP synthesis of Highway Practice 14: Skid resistance* Washington D.C, Highway research Board
- HIGHWAYS DEPARTMENT. 1989. *GUIDANCE NOTES ON ROAD TESTING* [Online]. Available:  
[http://www.hyd.gov.hk/en/publications\\_and\\_publicity/publications/technical\\_document/guidance\\_notes/pdf/gn009.pdf](http://www.hyd.gov.hk/en/publications_and_publicity/publications/technical_document/guidance_notes/pdf/gn009.pdf) [Accessed 4 June 2013].
- HUDZIETZ, B. P. & SARIPALLI, S. 2011. An Experimental Evaluation Of 3D Terrain Mapping With An Autonomous Helicopter. *UAV-g 2011, Conference on Unmanned Aerial Vehicle in Geomatics*. Zurich, Switzerland: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.
- IMAGEJ. 2013. *Image Processing and Analysis in Java* [Online]. Available: <http://rsbweb.nih.gov/ij/>.
- IMATEST. 2013. *The Imatest Test Lab* [Online]. Available: <http://www.imatest.com/docs/lab/> [Accessed 07 July 2013].
- JAHANSHAHI, M. R.;JAZIZADEH, F.;MASRI, S. F. & BECERIK-GERBER, B. 2012. A novel system for road surface monitoring using an inexpensive infrared laser sensor. *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, 8345.
- JIN-TSONG, H.;JIE-SHI, W. & YI-TING, T. 2012. 3D modeling and accuracy assessment- a case study of photosynth. *Geoinformatics (GEOINFORMATICS), 20th International Conference on Geoinformatics* 1-6.
- JOSHI, N.;KANG, S. B.;ZITNICK, C. L. & SZELISKI, R. 2010. Image deblurring using inertial measurement sensors. *ACM Transactions on Graphics (TOG)*, 29, 30.
- JUAN, L. & GWUN, O. 2009. A comparison of sift, pca-sift and surf. *International Journal of Image Processing (IJIP)*, 3, 143-152.
- KAPLAN, E. D. & HEGARTY, C. J. 2006. *Understanding GPS: principles and applications*. Norwood, MA: Artech House. Inc.
- KASEKO, M. S. & RITCHIE, S. G. 1993. A neural network-based methodology for pavement crack detection and classification. *Transportation Research Part C: Emerging Technologies*, 1, 275-291.
- KENNEDY, C. K.;YOUNG, A. E. & BUTLER, I. C. 1990. Measurement of skidding resistance and surface texture and the use of results in the United Kingdom. *ASTM STP*, 1031, 87-102.
- KOCH, C. & BRILAKIS, I. 2011. Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, 25, 507-515.
- KOCH, C. & BRILAKIS, I. 2012. Improving Pothole Recognition through Vision Tracking for Automated Pavement Assessment. *Advanced Engineering Informatics in Press*, 1-8.
- KOREN, N. 2004. *Understanding image sharpness part 5: Lens testing* [Online]. Available: <http://www.normankoren.com/Tutorials/MTF5.html#ImageJ> [Accessed July 2013].
- LÄBE, T. & FÖRSTNER, W. 2004. Geometric stability of low-cost digital consumer cameras. *International Archives of Photogrammetry and Remote Sensing*, 35, 528-535.



- LEICA 2006a. Leica Photogrammetry Suite, *Automatic Terrain Extraction, User's Guide* Atlanta: Leica Geosystem.
- LEICA 2006b. Leica Photogrammetry Suite, *Project Manager User's Guide* Atlanta: Leica Geosystem
- LEICA GEOSYSTEM. 2014a. *HDS Laser Scanners & SW* [Online]. Available: [http://www.leica-geosystems.co.uk/en/HDS-Laser-Scanners-SW\\_5570.htm](http://www.leica-geosystems.co.uk/en/HDS-Laser-Scanners-SW_5570.htm) [Accessed 24 July 2014].
- LEICA GEOSYSTEM. 2014b. *Leica home webpage* [Online]. Available: <http://hds.leica-geosystems.com/en/index.htm> [Accessed 23 July 2014].
- LEICA GEOSYSTEMS. 2003. IMAGINE OrthoBASE® User's Guide. *Georgia, USA*.
- LEISING, C. 2010. *Surface Roughness of Concrete with 3D Profilometry* [Online]. Available: <http://www.nanovea.com/Application%20Notes/concreteroughness.pdf> [Accessed 26 August 2012].
- LENZ, R. W. 2011. *Pavement Design Guide* [Online]. Texas Department of Transportation. Available: [http://onlinemanuals.txdot.gov/txdotmanuals/pdm/manual\\_notice.htm](http://onlinemanuals.txdot.gov/txdotmanuals/pdm/manual_notice.htm) [Accessed 15 September 2013].
- LEVIN, A.;WEISS, Y.;DURAND, F. & FREEMAN, W. T. Understanding and evaluating blind deconvolution algorithms. *Computer Vision and Pattern Recognition*, 2009. CVPR 2009. IEEE Conference on, 2009. IEEE, 1964-1971.
- LINGUA, A.;MARENCHINO, D. & NEX, F. 2009. Performance Analysis of the SIFT Operator for Automatic Feature Extraction and Matching in Photogrammetric Applications. *Sensors*, 9, 3745-3766.
- LIU, J. K.;CHANG, K. T. & CHANG, J. R. 2005. Detection of Pavement Distresses Using 3D Laser Scanning Technology. in: *Proceeding of the ASCE International Conference on Computing in Civil Engineering*. Maxico, Cancun.
- LOWE, D. 2007. *SIFT: Scale Invariant Feature Transform* [Online]. Available: [http://web.eecs.umich.edu/~silvio/teaching/EECS598/lectures/lecture10\\_1.pdf](http://web.eecs.umich.edu/~silvio/teaching/EECS598/lectures/lecture10_1.pdf) [Accessed 5 May 2013].
- LUGADE, V. 2011. *Clustering Part 2: K-means clustering* [Online]. Available: <http://matlabgeeks.com/tips-tutorials/clustering-part-2-k-means-clustering/> [Accessed 07 August 2014].
- LUHMANN, T. 2010. Close range photogrammetry for industrial applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 558-569.
- LUHMANN, T.;ROBSON, S.;KYLE, S. & HARLEY, I. 2006. *Close range photogrammetry: Principles, methods and applications*, Dunbeath, Scotland, Whittles Publishing.
- MARTIN, S. 2004. Accuracy of automated aerotriangulation and DTM generation for low textured imagery. *international Archives of Photogrammetry and Remote Sensing*, 35, 521-526.
- MASSEY, R. 2013. *Motorists paid £22MILLION in compensation for pothole damage* [Online]. Mail Online. [Accessed 21 Mar 2013].
- MATEST. 2010. *SAND PATCH EQUIPMENT* [Online]. Italy. Available: <http://www.matest.com/products/bitumen-asfalt/b099-10-sand-patch-equipment-0.aspx> [Accessed 10 Jan 2014].
- MATHEW, T. 2009. *Introduction to pavement design* [Online]. Indian Institute of Technology Bombay. Available:



- [http://www.civil.iitb.ac.in/tvm/1100\\_LnTse/401\\_LnTse/plain/plain.html](http://www.civil.iitb.ac.in/tvm/1100_LnTse/401_LnTse/plain/plain.html) [Accessed 20 Dec. 2013].
- MATWORKS. 2014. *Find edges in intensity image* [Online]. Available: <http://www.mathworks.co.uk/help/images/ref/edge.html?searchHighlight=edge> [Accessed 25 Feb. 2014].
- MATLAB. 2014a. *Deblurring* [Online]. Available: <http://www.mathworks.co.uk/help/images/image-restoration-deblurring.html> [Accessed 07 July 2014].
- MATLAB. 2014b. *Image Filtering* [Online]. Available: <http://www.mathworks.co.uk/help/images/linear-filtering.html> [Accessed 08 August 2014].
- MATLAB. 2014c. *K-Mean Clustering* [Online]. Available: <http://www.mathworks.co.uk/help/stats/k-means-clustering.html> [Accessed 14 July 2014].
- MATLAB. 2014d. *Two-sample t-test* [Online]. Available: <http://uk.mathworks.com/help/stats/ttest2.html?refresh=true#btqtvxyx-2> [Accessed 15 Jan 2014].
- MCQUAID, G.; MILLAR, P.; WOODWARD, D. & RYAN, D. 2010. Application of Close Range Photogrammetry and 3D Laser Scanning to better understand Tyre/Asphalt Interaction. *Proceedings of the 12th Annual International Conference on Sustainable Construction Materials, Pavement Engineering and Infrastructure*. Liverpool, UK.
- MEDNIS, A.; STRAZDINS, G.; ZVIEDRIS, R.; KANONIRS, G. & SELAVO, L. Real time pothole detection using Android smartphones with accelerometers. Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011 International Conference on, 27-29 June 2011 2011. 1-6.
- MICHOT, J. 2010. *Methods for Structure-from-Motion (SfM)* [Online]. Available: <http://michot.julien.free.fr/drupal/?q=content/research>.
- MICRO-EPSILON. 2014. *More precision* [Online]. Available: <http://www.micro-epsilon.co.uk/glossar/Laser-Triangulation.html> [Accessed 24 July 2014].
- MILLER, J. S. & BELLINGER, W. Y. 2003. *Distress identification manual for the long-term pavement performance program* [Online]. Washington, DC: Federal Highway Administration. Available: [http://www.fhwa.dot.gov/pavement/pub\\_details.cfm?id=91](http://www.fhwa.dot.gov/pavement/pub_details.cfm?id=91) [Accessed 21 March 2013].
- MORSE, B. 2000. *Thresholding* [Online]. Birmingham Young University Available: [http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/MORSE/threshold.pdf](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/threshold.pdf) [Accessed 10 Decmber 2013].
- MUJA, M. & LOWE, D. G. 2009. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. In: *Procceding of International Conference on Computer Vision Theory and Applications* 331–340.
- NAGAI, M.; CHEN, T.; SHIBASAKI, R.; KUMAGAI, H. & AHMED, A. 2009. UAV-Born 3D mapping System by multisensore integration. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 47, 701–708.
- NEITZEL, F. & KLONOWSKI, J. 2011. Mobile 3D Mapping With A Low-Cost UAV System. *UAV-g 2011, Conference on Unmanned Aerial Vehicle in Geomatics*. Zurich, Switzerland: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

- NUTZ, P. & HOFFMANN, M. 2012. Towards real-time skid resistance forecast. *Ratio*, 5, 04.
- OLIVIER KUNG;CHRISTOPH STRECHA;ANTOINE BEYELER;JEAN-CHRISOPHE ZUFFEREY;DARIO FLOREANO;PASCAL FUA & FRANÇOIS GERVAIX 2011. The Accuracy of Automatic Photogrammetric Techniques on Ultra-Light UAV Imagery. *Conference on Unmanned Aerial Vehicle in Geomatics*,. Zurich, Switzerland.
- ORDNANCE SURVEY. 2014. *Grid InQuest Coordinate transformation* [Online]. Ordnance Survey. Available: <http://www.ordnancesurvey.co.uk/business-and-government/help-and-support/navigation-technology/os-net/grid-inquest.html>.
- OSU. 2012. *Pavement Types* [Online]. Oregon State University. Available: [http://classes.engr.oregonstate.edu/cce/winter2012/ce492/Modules/02\\_pavement\\_type/s/02-2\\_body.htm](http://classes.engr.oregonstate.edu/cce/winter2012/ce492/Modules/02_pavement_type/s/02-2_body.htm) [Accessed 20 Dec. 2013].
- PAIGE, P.;MAHARJ, A. & KOMBA, J. 2011. *POTHOLES: Technical guide to their causes, identification and repair* [Online]. Available: [http://www.csir.co.za/pothole\\_guides/docs/Pothole\\_CSIR\\_tech\\_guide.pdf](http://www.csir.co.za/pothole_guides/docs/Pothole_CSIR_tech_guide.pdf) [Accessed 2 March 2014].
- PAVEMENT INTERACTIVE. 2012. *Pavement Evaluation* [Online]. Available: <http://www.pavementinteractive.org/category/pavement-management/pavement-evaluation/?show=all> [Accessed 8 April 2012].
- PAVEMETRICS. 2012. *Leader in vision systems for the automated inspection of transportation infrastructures*. [Online]. Canada. Available: [www.pavemetrics.com](http://www.pavemetrics.com) [Accessed 09 Dec. 2013].
- PAYNE, L. D. 1992. Automating road surface analysis. in *Proceedings of the ACM/SIGAPP Symposium on Applied Computing*. New York, USA: ACM Press.
- PEARSON, D. 2012. *Deterioration and maintenance of pavements*, London, ICE Publishing
- PENG, K.;CHEN, X.;ZHOU, D. & LIU, Y. 2009. 3D reconstruction based on SIFT and Harris feature points. *Proceedings of the IEEE International Conference on Robotics and Biomimetics*, 960-964.
- PERNG, J. W.;TAI, C. H.;KUO, C. H. & MA, L. S. 2013. 3D Environment Mapping and Pothole Detection for a Mobile Robot. *Applied Mechanics and Materials*, 431, 287-295.
- PFEIFER, N. & BRIESE, C. 2007. Laser scanning-Principles and applications. *3rd International Exhibition and Scientific Congress on Geodesy, Mapping, Geology, Geophysics*. Novosibirsk, Russia.: Citeseer.
- PHOTOMODELER®. 2009. *Eos Systems Inc* [Online]. Vancouver, Canada: EOS. Available: <http://www.photomodeler.com/index.htm> [Accessed 26 Mar 2012].
- PHOTOSYNTH. 2012. *Microsoft Photosynth* [Online]. Available: <http://photosynth.net/> [Accessed 22 March 2012].
- PHOTOZONE. 2013. *All Tests / Reviews* [Online]. Available: <http://www.photozone.de/> [Accessed 26 July 2013].
- PIERROT DESEILLIGNY, M. & CLERY, I. 2011. APERO, an open source bundle adjustment software for automatic calibration and orientation of set of images. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38.
- QUESTEXMEDIA. 2012. *Trimble Acquires UAV Mapping Company Gatewing* [Online]. Available: <http://www.gpsworld.com/gis/earth-imaging-and-remote->

- [sensing/news/trimble-acquires-uav-mapping-company-gatewing-12845](http://sensing/news/trimble-acquires-uav-mapping-company-gatewing-12845) [Accessed 7 April 2012].
- RABABAHAH, H.;VRAJITORU, D. & WOLFER, J. 2005. *Asphalt pavement crack classification: a comparison of GA, MLP, and SOM* [Online]. Computer and Information Sciences. Available: <http://www.cs.bham.ac.uk/~wbl/biblio/gecco2005lbp/papers/78-rababaah.pdf> [Accessed 25 April 2013].
- RASMUSSEN, R. O.;WIEGAND, P. D.;FICK, G. J.;HARRINGTON, D. S. & TPF, P. F. 2011. How to Reduce Tire-Pavement Noise: Better Practices for Constructing and Texturing Concrete Pavement Surfaces. *National Concrete Pavement Technology Center*.
- REMONDINO, F.;BARAZZETTI, L.;NEX, F.;SCAIONI, M. & SARAZZI, D. 2011. UAV Photogrammetry for Mapping and 3D Modelling – Current Status and Future Perspectives –. *UAV-g, Conference on Unmanned Aerial Vehicle in Geomatics*. Zurich, Switzerland: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.
- REMONDINO, F. & EL-HAKIM, S. 2006. Image-based 3D Modelling: A Review. *The Photogrammetric Record*, 21, 269-291.
- REMONDINO, F. & FRASER, C. 2006. Digital camera calibration methods: considerations and comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 266-272.
- REMONDINO, F.;PIZZO, S.;KERSTEN, T. & TROISI, S. 2012. Low-Cost and Open-Source Solutions for Automated Image Orientation – A Critical Overview. In: IOANNIDES, M., FRITSCH, D., LEISSNER, J., DAVIES, R., REMONDINO, F. & CAFFO, R. (eds.) *Progress in Cultural Heritage Preservation*. Springer Berlin Heidelberg.
- ROSNELL T;HONKAVAARA, E. & NURMINE K 2011. On Geometric Processing Of Multi-Temporal Image Data Collected By Light Uav Systems. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-1/C22, 63–68.
- ROY, D. 2012. *computer and machine vision theory algorithms practicalities*, Academic Press.
- SANTOS, P. & JÚLIO, E. N. 2013. A state-of-the-art review on roughness quantification methods for concrete surfaces. *Construction and Building Materials*, 38, 912-923.
- SCHMIDT, F. 2011. *GeoSetter* [Online]. Available: <http://www.geosetter.de/en/> [Accessed 25 Feb 2014].
- SENGOZ, B.;TOPAL, A. & TANYEL, S. 2012. Comparison of pavement surface texture determination by sand patch test and 3D laser scanning. *Civil Engineering*, 56, 73-78.
- SHIMAMURA, H.;OONUMA, K. & YASUDA, Y. 2009. Image Analysis for Automated Pavement Cracking Evaluation. *IAPR Workshop on Machine Vision Applications*, 83-86.
- SMITH, M. 2009. Photogrammetry Handouts IESSG: The University of Nottingham.
- SNAVELY, N. 2010. *Bundler: Structure from Motion (SfM) for Unordered Image Collections* [Online]. Available: <http://phototour.cs.washington.edu/bundler/> [Accessed 22 Mrch 2012].
- SOURCEFORGE. 2011. *Meshlab* [Online]. Available: <http://meshlab.sourceforge.net/> [Accessed 28 Nov. 2011].

- STRECHA, C. 2011. Automated Photogrammetric Techniques on Ultra-light UAV Imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-1, 125-130.
- SUBIRATS, P.; FABRE, O.; DUMOULIN, J.; LEGEAY, V. & BARBA, D. 2006. Automation of pavement surface crack detection with a matched filtering to define the mother wavelet function used. in *Proceedings of the 14th European Signal Processing Conference (EUSIPCO)*. Florence, Italy.
- SY, N. T.; AVILA, M.; BEGOT, S. & BARDET, J. C. Detection of defects in road surface by a vision system. Electrotechnical Conference, 2008. MELECON 2008. The 14th IEEE Mediterranean, 5-7 May 2008 2008. 847-851.
- SYLVIE, C. & JEAN-MARC, M. 2011. Automatic Road Pavement Assessment with Image Processing: Review and Comparison. *International Journal of Geophysics*, 2011, 1-20.
- THAMM, H. P. 2011. SUSI 62: a Robust and Safe Parachute UAV with Long Flight Time and Good Payload. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-1/C22, 19-24.
- TRACS, H. A. 2006. *Statement of Administrative Sources-National Road Condition Database* [Online]. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/10967/dft-statement-natinlroad-condbase.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/10967/dft-statement-natinlroad-condbase.pdf) [Accessed 15 Jan. 2012].
- TRANSPORT RESEARCH LABORATORY. 2014. *Highways Agency Road Research Information System (HARRIS)* [Online]. Available: [http://www.trl.co.uk/facilities/mobile\\_test\\_equipment/highways\\_agency\\_road\\_research\\_information\\_system.htm](http://www.trl.co.uk/facilities/mobile_test_equipment/highways_agency_road_research_information_system.htm) [Accessed 01 May 2014].
- TRL. 2012. *UKPMS* [Online]. Berkshire Available: <http://www.pcis.org.uk/index.php?p=25/42/0> [Accessed 24 Dec. 2013].
- ULTRACAM. 2011. *Web Seminar: UltraMap v3.0 Tech Preview* [Online]. Available: <http://www.microsoft.com/ultracam/en-us/webevents.aspx> [Accessed 22 Mar 2012].
- UREN, J. & PRICE, W. F. 2010. *Surveying for engineers*, Palgrave Macmillan.
- USF. 2014. *Chapter Four: Image Filtering* [Online]. Available: [http://www.cse.usf.edu/~rlk/MachineVisionBook/MachineVision.files/MachineVision\\_Chapter4.pdf](http://www.cse.usf.edu/~rlk/MachineVisionBook/MachineVision.files/MachineVision_Chapter4.pdf) [Accessed 08 August 2014].
- UTAH. 2002. *Segmentation using K-Means Algorithm* [Online]. Available: <http://www.cs.utah.edu/~arul/report/node18.html> [Accessed 07 August 2014].
- VINER, H. & BRITAIN, G. 2006. *Surface texture measurement on local roads* [Online]. TRL Limited. Available: <http://www.roadsafetyobservatory.com/Evidence/Details/10555> [Accessed 10 Feb. 2013].
- VIRTUALGRID. 2014. *VRMesh for point cloud and triangle mesh processing* [Online]. Bellevue. Available: <http://www.vrmesh.com/> [Accessed 23 Jan. 2014].
- VISICS. 2012. *ARC 3D Webservice* [Online]. Available: <http://www.arc3d.be/> [Accessed 22 Mar 2012].
- VISUALSFM. 2011. *VisualSFM : A Visual Structure from Motion System* [Online]. University of Washington at Seattle. Available: <http://www.cs.washington.edu/homes/ccwu/vsfm/> [Accessed 23 May 2012].
- VOLLMERHAUSEN, R. H. & DRIGGERS, R. G. 2000. Analysis of sampled imaging systems. *SPIE Press*, 111-138.

- WANG, K. C. P. 2000. Designs and Implementations of Automated Systems for Pavement Surface Distress Survey. *Journal of Infrastructure Systems*, 6, 24-32.
- WANG, K. C. P. & ELLIOTT, R. P. 1999. Investigation of image archiving for pavement surface distress survey. Mack-Blackwell Transportation Center, Department of Civil Engineering, University of Arkansas, Fayetteville, NC, USA.
- WANG, K. C. P. & SMADI, O. 2011. Automated Imaging Technologies for Pavement Distress Surveys. *Transportation Research E-Circular*, E-C156, 1-15.
- WARRANTYDIRECT. 2012. *Hit a pothole, found a pothole, or just fed up with potholes?* [Online]. Available: <http://potholes.co.uk/> [Accessed 12 April 2012].
- WERNER, S.; NEUMANN, I.; THIENEL, K.-C. & HEUNECKE, O. 2013. A fractal-based approach for the determination of concrete surfaces using laser scanning techniques: a comparison of two different measuring systems. *Materials and structures*, 46, 245-254.
- WIKI. 2014. *Neutral density filter* [Online]. Available: [http://en.wikipedia.org/wiki/Neutral\\_density\\_filter](http://en.wikipedia.org/wiki/Neutral_density_filter) [Accessed 07 August 2014].
- WOLF, P. R. & DEWITT, B. A. 2000. *Elements of Photogrammetry: with applications in GIS*, McGraw-Hill New York.
- WOLF, S. 2014. *WHAT TO DO WHEN YOU SEE ALLIGATOR CRACKING IN ASPHALT* [Online]. Available: <http://www.wolfpaving.com/blog/what-to-do-when-you-see-alligator-cracking-in-asphalt> [Accessed 10 Feb 2014].
- WORLD ROAD ASSOCIATION. 2012. *TC 4.2 Road Pavements* [Online]. Available: <http://www.piarc.org/en/Technical-Committees-World-Road-Association/Technical-Committee-Road-Pavements/#tab-sccbdy34d1> [Accessed 17 Dec. 2013].
- YEAZEL, J. & BIRCH, M. 2002. *Garmin GPSmap 76 -New Larger "Marine" Hand-held Model with WAAS* [Online]. Available: <http://www.gpsinformation.net/waas/g76/g-76-map.html> [Accessed 09 Dec. 2014].
- YUZHNIKOV, V. 2014. *SmartDeblur* [Online]. Available: <http://smartdeblur.net/> [Accessed 12 August 2014].
- ZHANG, C. 2008a. *Development of a UAV-based remote sensing system for unpaved road condition assessment* [Online]. Available: <http://www.asprs.org/a/publications/proceedings/portland08/0092.pdf> [Accessed 5 Feb. 2013].
- ZHANG, C. 2008b. An UAV-based photogrammetric mapping system for road condition assessment. *CiteSeer*, 627-631.
- ZHANG, C. & ELAKSHER, A. 2012. An Unmanned Aerial Vehicle-Based Imaging System for 3D Measurement of Unpaved Road Surface Distresses1. *Computer-Aided Civil and Infrastructure Engineering*, 27, 118-129.
- ZHANG, Z. 2000. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22, 1330-1334.
- ZOU, Q.; CAO, Y.; LI, Q.; MAO, Q. & WANG, S. 2012. CrackTree: Automatic crack detection from pavement images. *Pattern Recognition Letters*, 33, 227-238.



## APPENDICES

### Appendix A: Transformation parameters from CC aligning tool

<p><b>Align info</b></p> <p><b>C5R0-1</b></p> <p>Final RMS: 0.0005445</p> <p>Transformation matrix</p> <pre>-0.013 -0.011 0.005 998.791 -0.012 0.012 -0.005 1001.273 -0.001 -0.007 -0.016 100.680 0.000 0.000 0.000 1.000</pre> <p>Scale: 0.0175592 (already integrated in above matrix!)</p> <p>Refer to Console (F8) for more details</p> <p>OK</p>	<p><b>Align info</b></p> <p><b>C5R0-2</b></p> <p>Final RMS: 0.00020334</p> <p>Transformation matrix</p> <pre>-0.030 0.028 -0.016 998.808 0.032 0.026 -0.016 1001.205 -0.001 -0.022 -0.038 100.670 0.000 0.000 0.000 1.000</pre> <p>Scale: 0.0438881 (already integrated in above matrix!)</p> <p>Refer to Console (F8) for more details</p> <p>OK</p>
<p><b>Align info</b></p> <p><b>C5R15-1</b></p> <p>Final RMS: 0.000685659</p> <p>Transformation matrix</p> <pre>0.016 -0.015 0.003 998.623 -0.015 -0.015 0.006 1001.168 -0.001 -0.006 -0.021 100.690 0.000 0.000 0.000 1.000</pre> <p>Scale: 0.0221562 (already integrated in above matrix!)</p> <p>Refer to Console (F8) for more details</p> <p>OK</p>	<p><b>Align info</b></p> <p><b>C5R15-2</b></p> <p>Final RMS: 0.00044708</p> <p>Transformation matrix</p> <pre>-0.029 0.004 0.000 998.625 0.004 0.029 -0.005 1001.244 -0.001 -0.005 -0.029 100.716 0.000 0.000 0.000 1.000</pre> <p>Scale: 0.0295875 (already integrated in above matrix!)</p> <p>Refer to Console (F8) for more details</p> <p>OK</p>
<p><b>Align info</b></p> <p><b>C5R30-1</b></p> <p>Final RMS: 0.00084757</p> <p>Transformation matrix</p> <pre>-0.026 0.021 -0.002 998.610 0.021 0.025 -0.006 1001.241 -0.002 -0.006 -0.033 100.697 0.000 0.000 0.000 1.000</pre> <p>Scale: 0.0331917 (already integrated in above matrix!)</p> <p>Refer to Console (F8) for more details</p> <p>OK</p>	<p><b>Align info</b></p> <p><b>C5R30-2</b></p> <p>Final RMS: 0.000821897</p> <p>Transformation matrix</p> <pre>-0.020 -0.021 0.007 998.621 -0.022 0.019 -0.006 1001.176 0.000 -0.009 -0.028 100.689 0.000 0.000 0.000 1.000</pre> <p>Scale: 0.0293077 (already integrated in above matrix!)</p> <p>Refer to Console (F8) for more details</p> <p>OK</p>
<p><b>Align info</b></p> <p><b>C5R45-1</b></p> <p>Final RMS: 0.000884146</p> <p>Transformation matrix</p> <pre>-0.018 -0.019 0.011 998.578 -0.021 0.015 -0.008 1001.160 -0.000 -0.014 -0.024 100.678 0.000 0.000 0.000 1.000</pre> <p>Scale: 0.0277729 (already integrated in above matrix!)</p> <p>Refer to Console (F8) for more details</p> <p>OK</p>	<p><b>Align info</b></p> <p><b>C5R45-2</b></p> <p>Final RMS: 0.00073645</p> <p>Transformation matrix</p> <pre>-0.020 -0.024 0.008 998.602 -0.026 0.019 -0.007 1001.177 0.000 -0.010 -0.031 100.700 0.000 0.000 0.000 1.000</pre> <p>Scale: 0.032655 (already integrated in above matrix!)</p> <p>Refer to Console (F8) for more details</p> <p>OK</p>

## Appendix B: Samples of data set images collected using Nikon D100

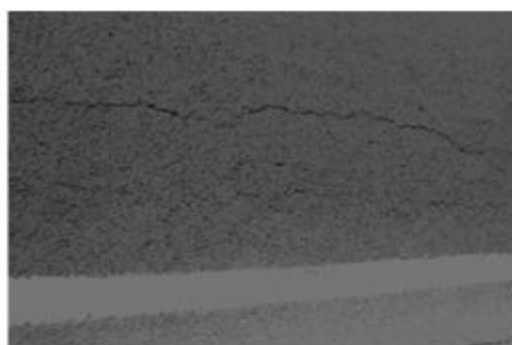


## Appendix C: different size of Structuring Elements effect

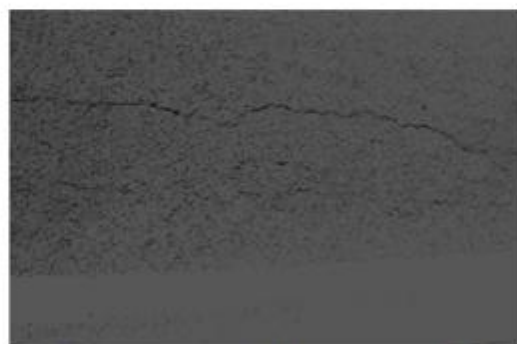
C1: Effect of increasing size of structuring elements



Structuring Element=5



Structuring Element=55



Structuring Element=65

C2: A sample of disk shape structuring element=5

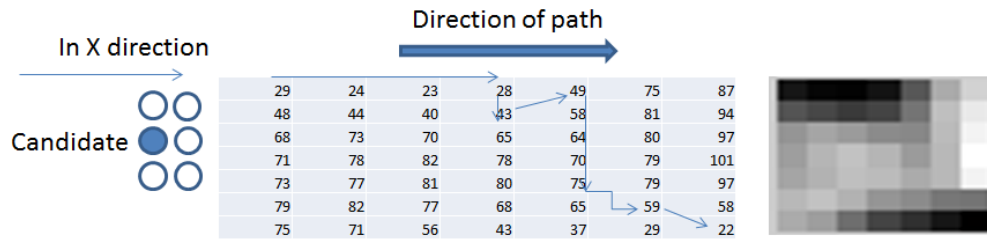
0	0	1	1	1	1	1	0	0
0	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	0
0	0	1	1	1	1	1	0	0



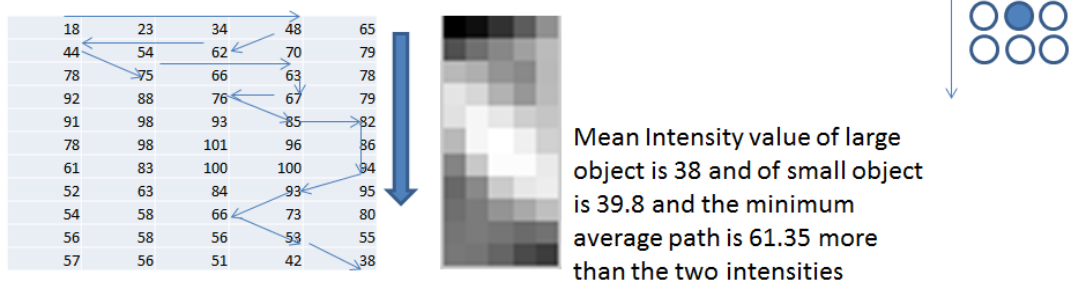
## Appendix D: Samples for different edge operations testing



## Appendix E: Minimum Average Cost Path Intensity



Mean Intensity value of large object is 38 and of small object is 41.2 and the minimum average path is 45.34 more than the two intensities

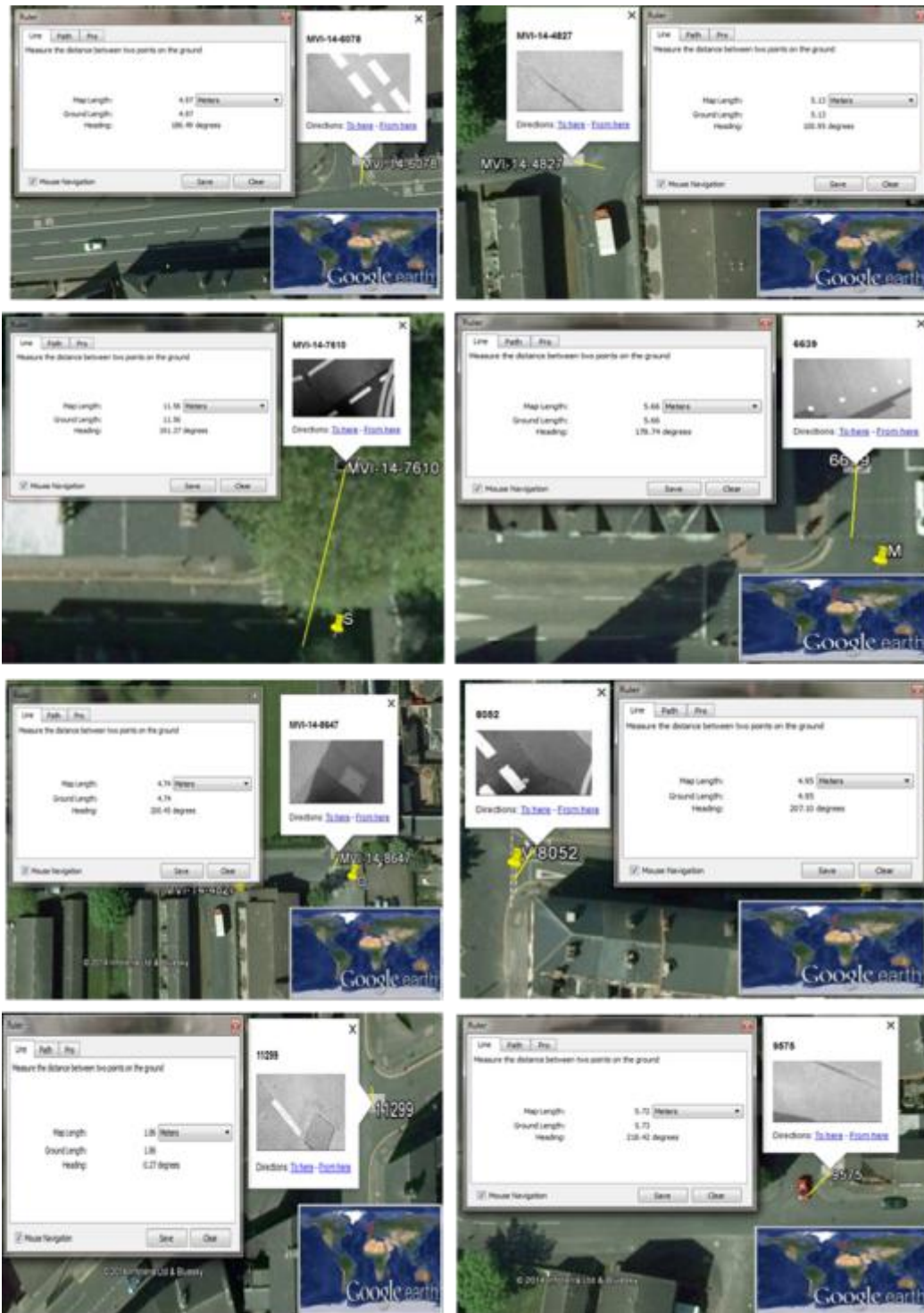


## Appendix F: Frame Geo-tagging results

### F01:MVI-4513

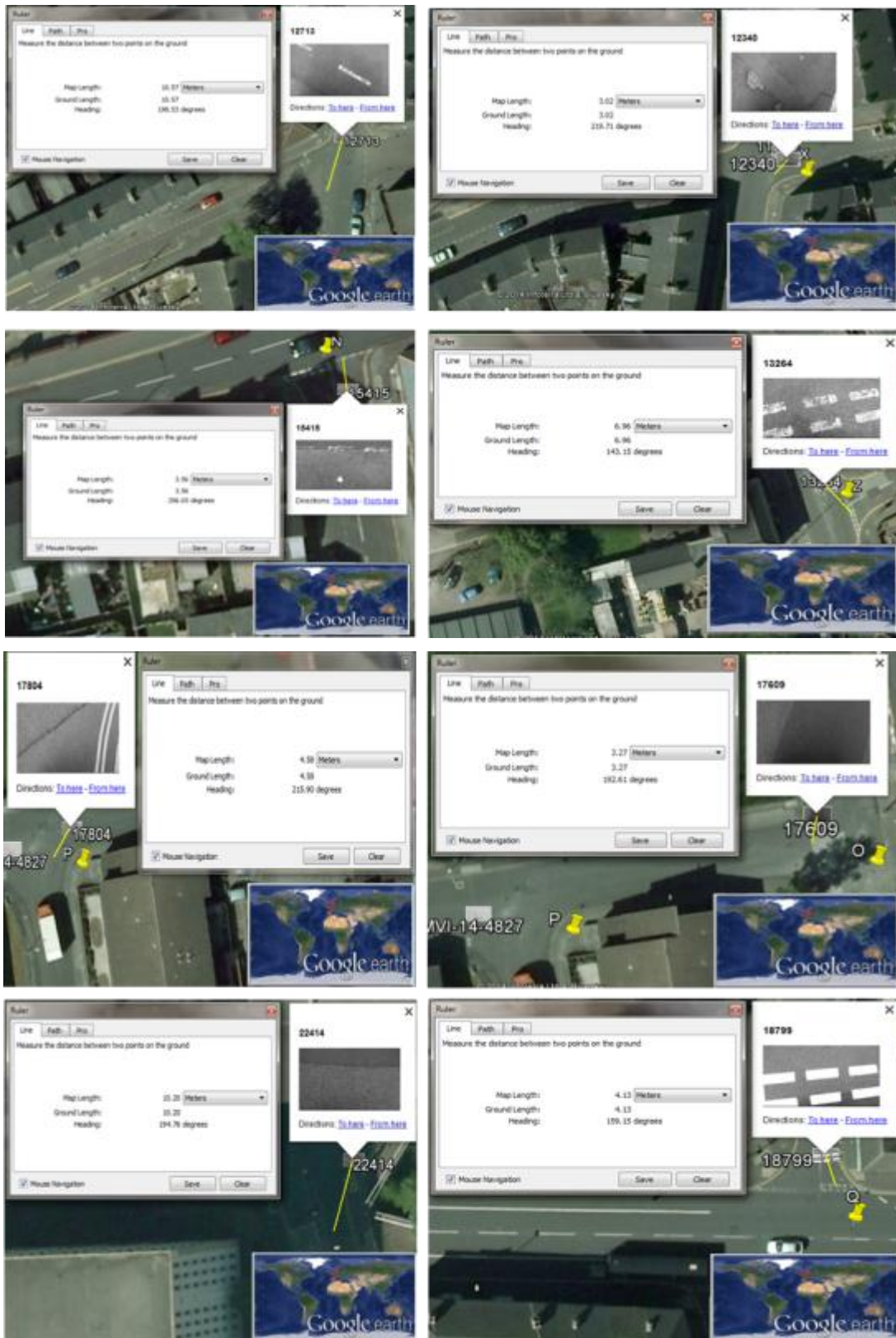


F02:MVI-4514-A

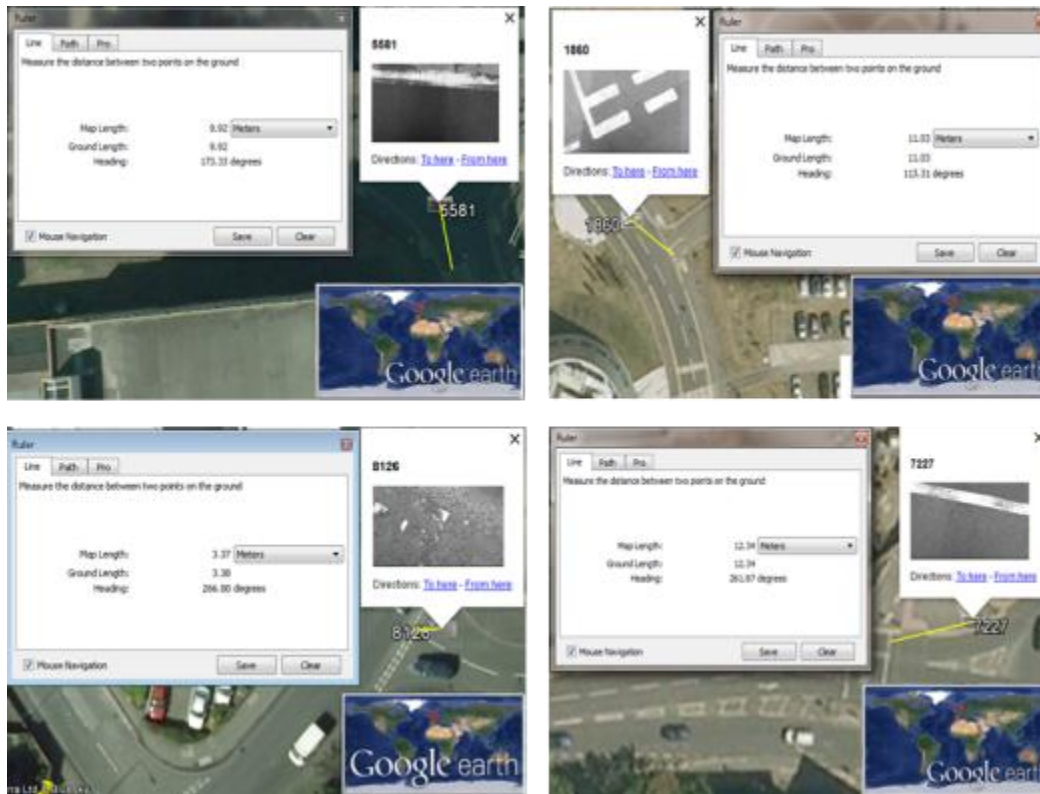




F02:MVI-4514-B



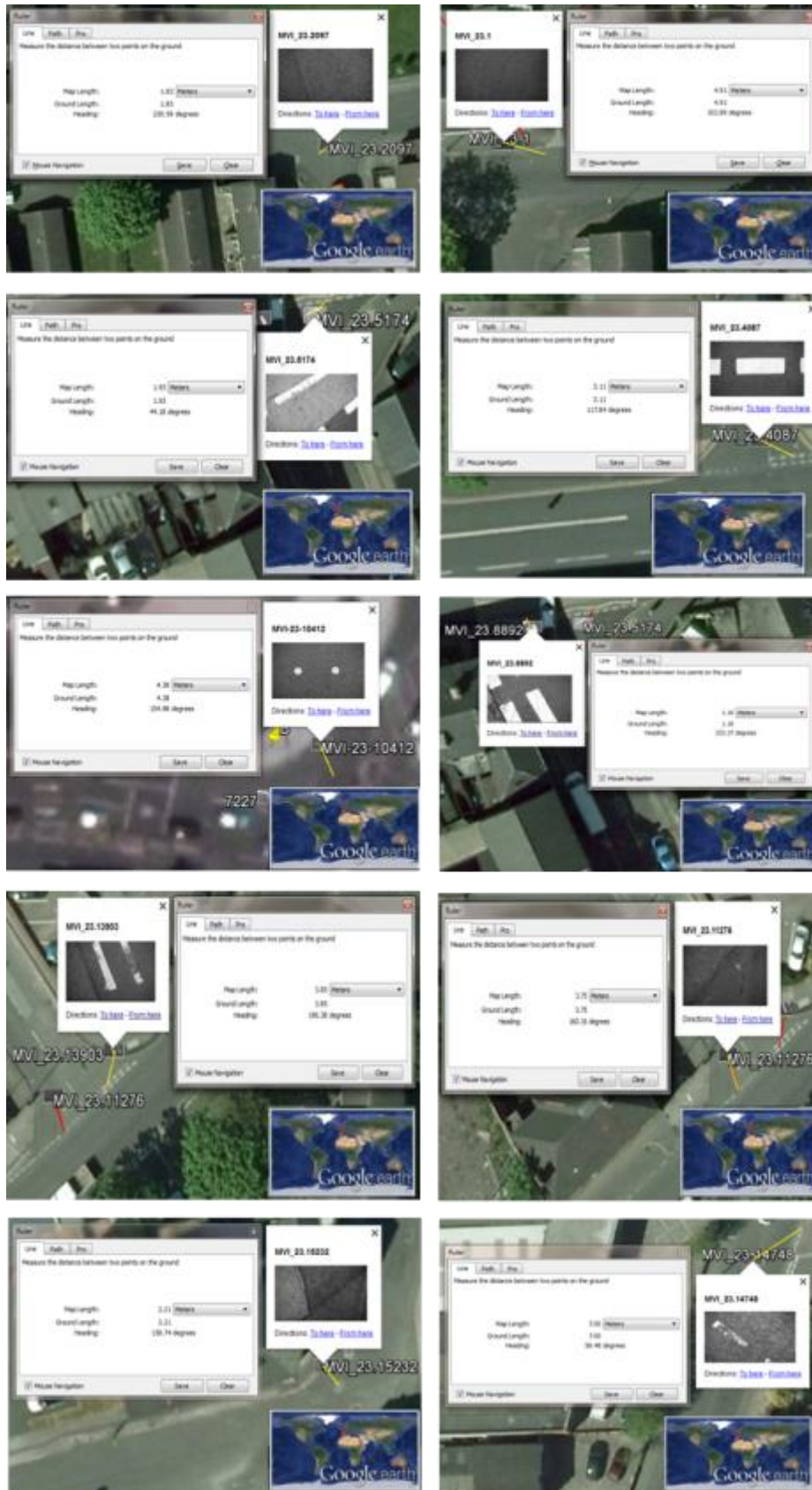
F03:MVI-4517



F04:MVI-4522

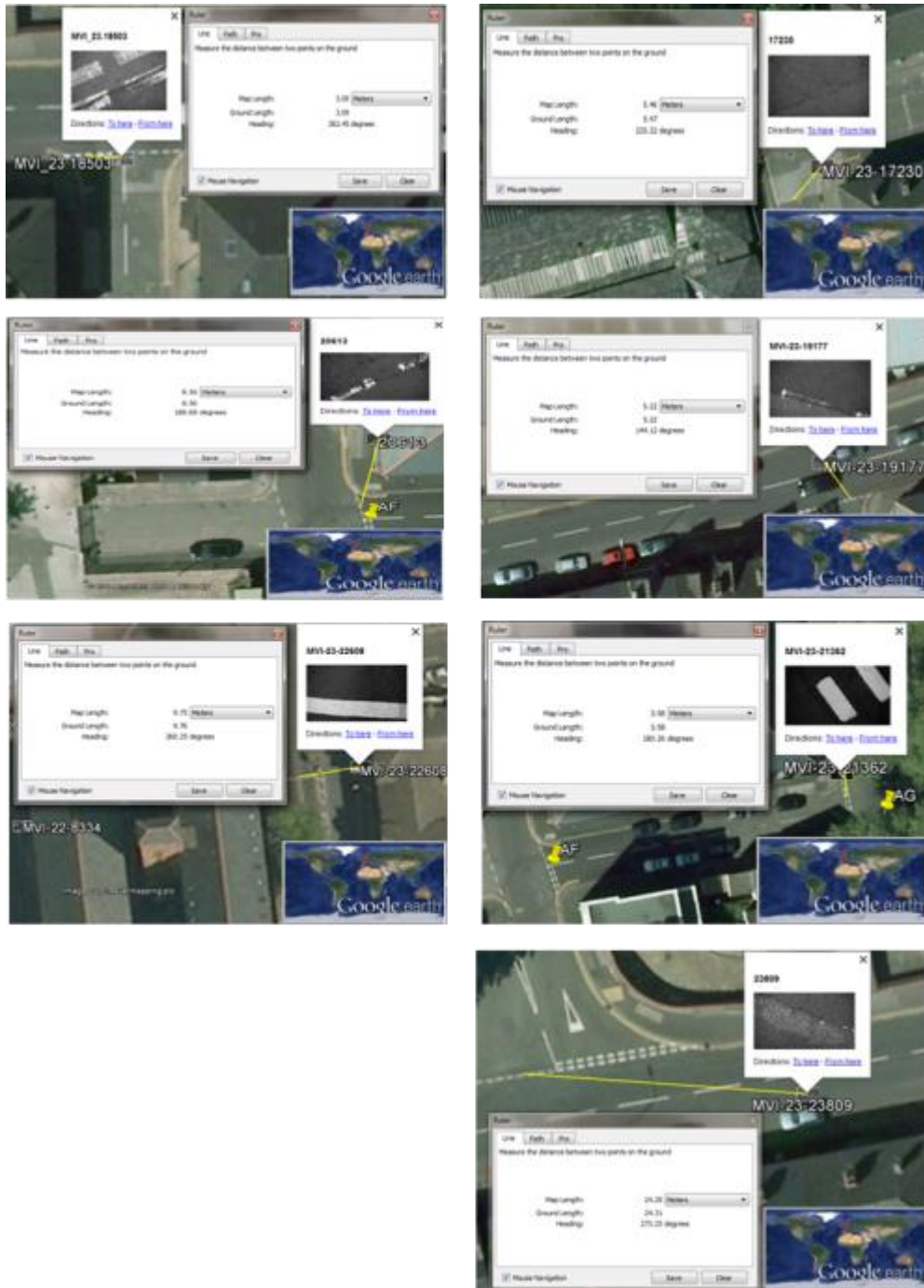


F05:MVI-4523-A





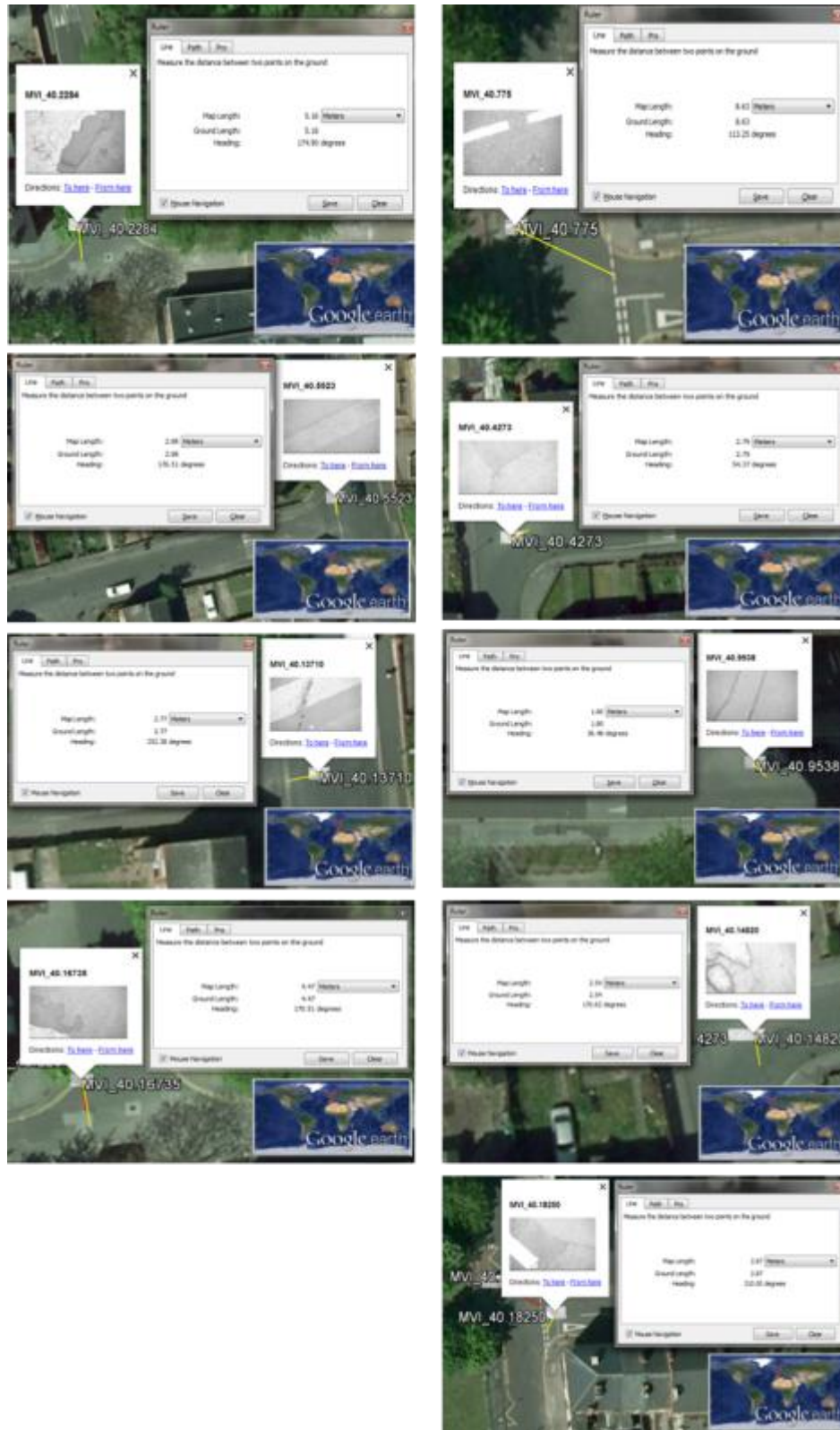
F05:MVI-4523-B



F06:MVI-4539



# F07:MVI-4540

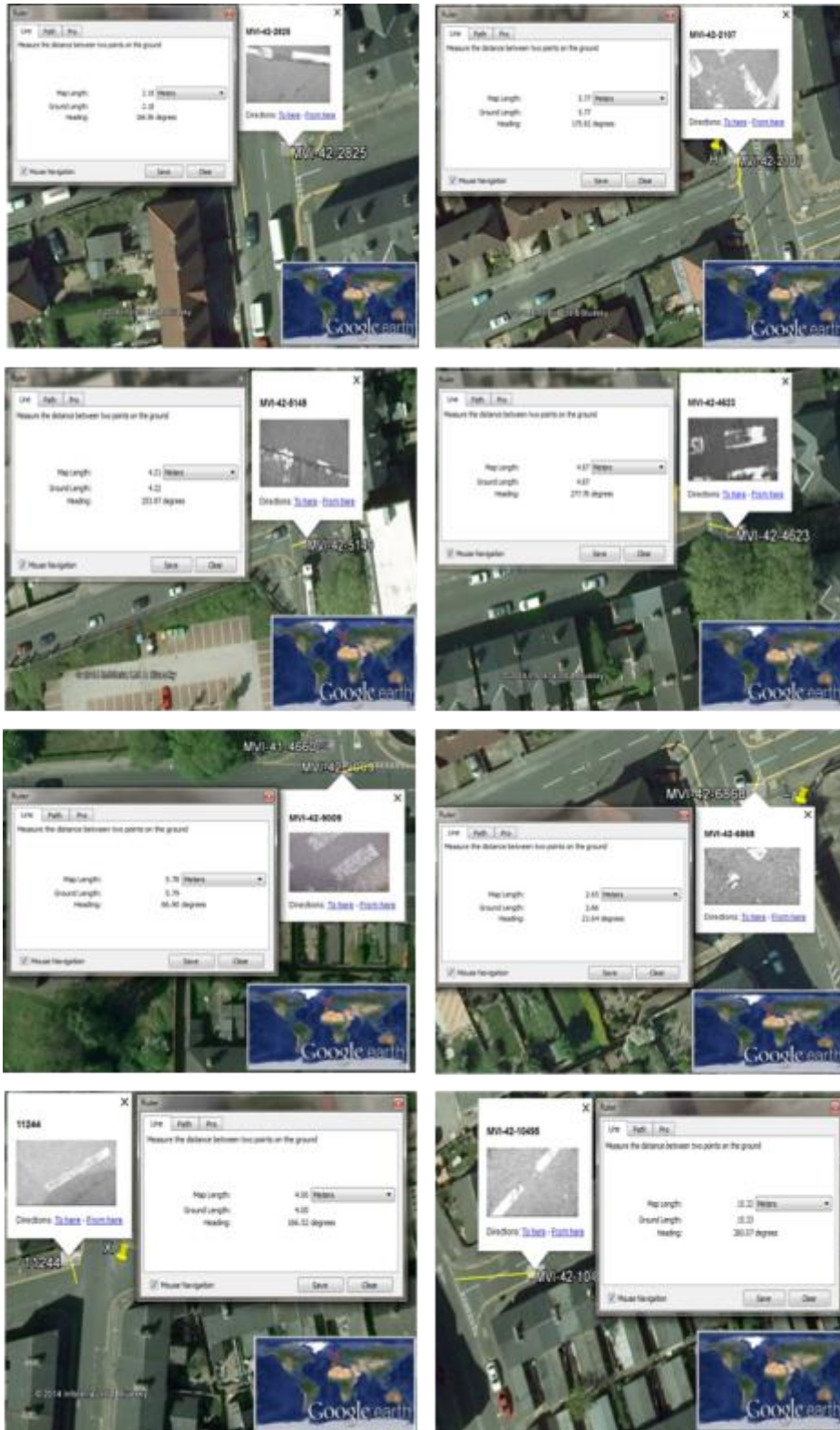




## F08: MVI-4541



# F09: MVI-4542

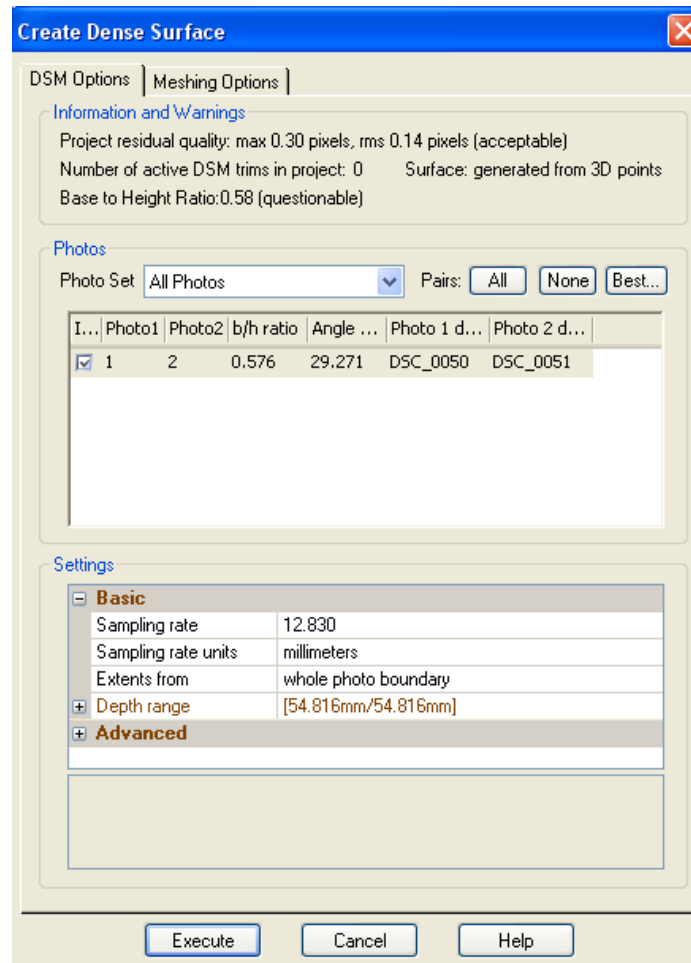


## Appendix G: Results of Triangulation

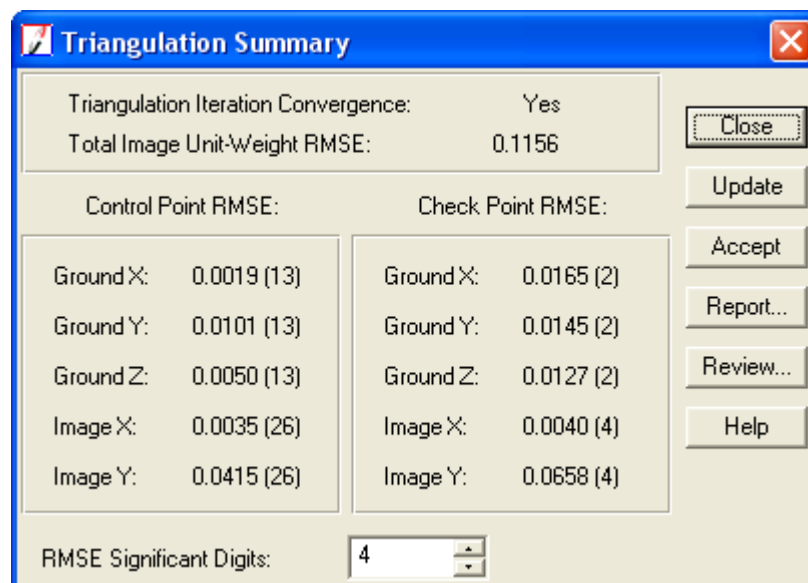
G1: Estimated accuracy results of Australis

Australis Project Status Summary	
<b>Australis Project Status Summary</b>	
Project Name:	Cal. aus
Camera Name(s):	NIKON CORPORATION NIKON D200 (unique ID: D200_right)
<b>Measurement Network</b>	
Number of images:	2
Number of referenced points:	12
Number of cameras:	1
Quality of geometry:	10.4 (poor)
Minimum number of points on an image:	14 on image DSC_0050.JPG DSC_0051.JPG
Minimum point intersection angle:	27 degs for point 8
Number of points referenced on:	
2 images only	12
3 or more images	0
4 or more images	0
6 or more images	0
<b>Measurement Accuracy Summary</b>	
Scale set?	Yes
Estimated accuracy of 3D point coordinates (RMS 1-sigma level)	
X	1.666 mm, or 1:800
Y	0.624 mm, or 1:2300
Z	0.899 mm, or 1:1600
Overall	1.063 mm, or 1:1300
Estimated accuracy of image referencing	0.04 pixels (RMS 1-sigma level)
Quality of self-calibration (if applied)	calibration not performed
<b>Output Files</b>	
3D coordinate file (text):	N/A
3D coordinate file (DXF):	N/A
<div>Print</div> <div>OK</div>	

## G2: Estimated accuracy results of PMS



## G3: Estimated accuracy results of LPS (pixels)



## Appendix H: Results of Accuracies for Bump Test

### H1: LPS triangulation results

Control Point RMSE:		Check Point RMSE:	
Ground X:	0.0117 (11)	Ground X:	0.0115 (1)
Ground Y:	0.0122 (11)	Ground Y:	0.0107 (1)
Ground Z:	0.0063 (11)	Ground Z:	0.0116 (1)
Image X:	0.3279 (44)	Image X:	0.1030 (4)
Image Y:	0.2749 (44)	Image Y:	0.4636 (4)

RMSE Significant Digits: 4

Using 11 GCPs (0.025m plan, 0.04m Z), Interior orientation parameters fixed,  
Jacobsen's simple model

### H2: GCPs accuracy and quality from Photomodeler

GCPs	X precision	Y precision	Z precision
1	0.00093	0.000924	0.003267
2	0.001282	0.001268	0.00154
3	0.001319	0.001299	0.001753
4	0.001073	0.001062	0.001166
5	0.001065	0.001056	0.001173
6	0.001029	0.001023	0.00118
7	0.000783	0.000784	0.000884
8	0.001473	0.001438	0.001908
9	0.001491	0.001474	0.001808
10	0.001221	0.00119	0.001401
11	0.001068	0.001043	0.001161
12	0.000824	0.000827	0.001022

**Project Status Report**

- Photo Coverage
  - Number of referenced points outside of the Camera's calibrated coverage: 0
- Point Marking Residuals
  - Overall RMS: 0.272 pixels
  - Maximum: 0.3364 pixels
  - Minimum: 0.208 pixels
  - Maximum RMS: 0.349 pixels
  - Minimum RMS: 0.157 pixels
- Point Tightness
  - Maximum: 0.0099 m
  - Minimum: 0.00011 m
- Point Precisions
  - Overall RMS Vector Length: 0.00022 m
  - Maximum Vector Length: 0.00227 m
  - Minimum Vector Length: 0.000525 m
  - Maximum X: 0.00077 m
  - Maximum Y: 0.00108 m
  - Maximum Z: 0.00198 m
  - Minimum X: 0.000135 m
  - Minimum Y: 0.00021 m
  - Minimum Z: 0.000457 m



### H3: Output report from AGISoft StereoScan

Output
<p>Detecting points...</p> <p>photo 1: 195613 points</p> <p>photo 2: 196584 points</p> <p>Matching points...</p> <p>Calibrating stereo pair...</p> <p>radial distortion calibration: enabled</p> <p>using 6648 of 8008 matches</p> <p>focal length: 4275.71 pix (exif)</p> <p>7552 points reconstructed, rms error 0.303558</p> <p>fx = 4300.11 fy = 4300.11 cx = 2808 cy = 1872 k1 = -0.0959398 k2 = 0.101991</p> <p>timings: match: 43.003 calibrate: 1.651</p> <p>Reconstructing depth...</p> <p>estimating 1904x2971x320 disparity using 272x248x320 tiles, offset -320</p> <p>filtering disparity image... done</p> <p>timings: rectify: 0.306 disparity: 31.485 borders: 0.144 filter: 24.999 fill: 0.125 confidence: 0.151</p> <p>timings: depth: 60.752 filter: 0 mesh: 0</p> <p>mesh generated in 0.211 sec</p> <p>Decimating mesh...</p> <p>Error: bad allocation</p> <p>Error: bad allocation</p> <p>Finished processing in 126.865 sec (exit code -1)</p> <p>Importing mesh...</p> <p>Finished processing in 0.103 sec (exit code 0)</p> <p>load: 2.522 sec</p>

## Appendix I: Canon 5D Camera calibration report

<b>Focal Length</b>	<b>Point Marking Residuals</b>
Value: 29.293731 mm	Overall RMS: 0.093 pixels
Deviation: Focal: 0.001 mm	Maximum: 0.381 pixels
<b>Xp - principal point x</b>	Point 281 on Photo 2
Value: 18.094230 mm	Minimum: 0.067 pixels
Deviation: Xp: 6.2e-004 mm	Point 63 on Photo 6
<b>Yp - principal point y</b>	Maximum RMS: 0.236 pixels
Value: 12.191040 mm	Point 281
Deviation: Yp: 7.2e-004 mm	Minimum RMS: 0.045 pixels
<b>Fw - format width</b>	Point 63
Value: 36.503023 mm	<b>Point Tightness</b>
Deviation: Fw: 1.9e-004 mm	Maximum: 0.0003 m
<b>Fh - format height</b>	Point 281
Value: 24.333200 mm	Minimum: 5.3e-005 m
<b>K1 - radial distortion 1</b>	Point 63
Value: 1.340e-004	<b>Point Precisions</b>
Deviation: K1: 2.4e-007	Overall RMS Vector Length: 2.95e-005 m
<b>K2 - radial distortion 2</b>	Maximum Vector Length: 3.74e-005 m
Value: -1.555e-007	Point 281
Deviation: K2: 7.7e-010	Minimum Vector Length: 2.87e-005 m
<b>K3 - radial distortion 3</b>	Point 73
Value: 0.000e+000	Maximum X: 1.65e-005 m
<b>P1 - decentering distortion 1</b>	Maximum Y: 1.69e-005 m
Value: -1.876e-005	Maximum Z: 2.92e-005 m
Deviation: P1: 2.6e-007	Minimum X: 1.1e-005 m
<b>P2 - decentering distortion 2</b>	Minimum Y: 1.09e-005 m
Value: 1.892e-005	Minimum Z: 2.36e-005 m
Deviation: P2: 2.6e-007	