AN INVESTIGATION INTO SEMI-AUTOMATED 3D CITY MODELLING

Nikolaos Kokkas, MSc, BSc

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

May 2008

ABSTRACT

Creating three dimensional digital representations of urban areas, also known as 3D city modelling, is essential in many applications, such as urban planning, radio frequency signal propagation, flight simulation and vehicle navigation, which are of increasing importance in modern society urban centres.

The main aim of the thesis is the development of a semi-automated, innovative workflow for creating 3D city models using aerial photographs and LiDAR data collected from various airborne sensors. The complexity of this aim necessitates the development of an efficient and reliable way to progress from manually intensive operations to an increased level of automation. The proposed methodology exploits the combination of different datasets, also known as data fusion, to achieve reliable results in different study areas. Data fusion techniques are used to combine linear features, extracted from aerial photographs, with either LiDAR data or any other source available including Very Dense Digital Surface Models (VDDSMs).

The research proposes a method which employs a semi automated technique for 3D city modelling by fusing LiDAR if available or VDDSMs with 3D linear features extracted from stereo pairs of photographs. The building detection and the generation of the building footprint is performed with the use of a plane fitting algorithm on the LiDAR or VDDSMs using conditions based on the slope of the roofs and the minimum size of the buildings. The initial building footprint is subsequently generalized using a simplification algorithm that enhances the orthogonality between the individual linear segments within a defined tolerance. The final refinement of the building outline is performed for each linear segment using the filtered stereo matched points with a least squares estimation.

The digital reconstruction of the roof shapes is performed by implementing a least squares-plane fitting algorithm on the classified VDDSMs, which is restricted by the building outlines, the minimum size of the planes and the maximum height tolerance between adjacent 3D

points. Subsequently neighbouring planes are merged using Boolean operations for generation of solid features. The results indicate very detailed building models. Various roof details such as dormers and chimneys are successfully reconstructed in most cases.

LIST OF PUBLISHED PAPERS

- **N. Kokkas., M. J. Smith., 2008**. A new workflow for semi-automated 3D city modelling. *Submitted to Photogrammetric Engineering and Remote Sensing.*
- **A. Taha., N. Kokkas., C. Hancock., G. Roberts., X. Meng., 2008**. A GIS approach to GNSS simulation in Urban Canyons. *European Navigation Conference (ENC-GNSS 08), France, Toulouse Space Show, April 23 – 25.*
- **M. J. Smith., N. Kokkas., A. M. Hamruni., D. Critchley., A. Jamieson., 2008**. Investigation into the orientation of oblique and vertical digital images, *EuroCOW 2008: International Calibration and Orientation Workshop, Spain, Castelldefels, Jan 30 – Feb 1.*
- **Miliaresis G., Kokkas N. 2007**. Segmentation & object based classification for the extraction of building class from LIDAR DEMs. *Computers & Geosciences, 33, 1076-1087. [*http://dx.doi.org/10.1016/j.cageo.2006.11.012*]*
- **M. J. Smith, N. Kokkas, K. S. Qtaishat., 2007.** Investigation into Self-Calibration methods for the Vexcel UltraCam D Digital Aerial Camera. *ISPRS Hannover Workshop, High Resolution Earth Imaging for Geospatial Information, 29 May - 1 June,*
- **N. Kokkas, M. J. Smith., 2007.** Automated 3D city modelling and the importance of quality assurance techniques. *ISPRS Hannover Workshop, High Resolution Earth Imaging for Geospatial Information, 29 May - 1 June, 2007*
- **Miliaresis G., Kokkas N., 2006.** Geomorphometric Segmentation Applied to the city modelling Problem. *International Symposium on Terrain Analysis and Digital Terrain Modelling, Nanjing, China, November 23-25*
- **Kokkas N., Dowman I. J., 2006**. Fusion of Airborne optical and LIDAR data for automated building reconstruction. *American Society for Photogrammetry & Remote Sensing, Reno Nevada, May 1-7.*

- **Kokkas N., 2005.** City modelling and building reconstruction with Socet Set v.5.2. BAE Systems, *2005 GXP Regional User Conference, Cambridge, UK, September 19-21.*
- **Kokkas N., Miliaresis G., 2004.** Geomorphometric Mapping of Grand Canyon from the 1-degree USGS DEM. *Int. Archives of Photogrammetry, Remote Sensing & GIS - Vol. XXXV, (XXth ISPRS Congress, 12-23 July, 2004 Istanbul, Turkey Commission IV), pp. 840-845*
- **Miliaresis G., Kokkas N., 2004.** Segmentation and terrain modelling of extra-terrestrial chasmata. *Journal of SPATIAL SCIENCES, Vol. 49, No. 2, pages 89-99.*
- **Miliaresis G., Kokkas N., 2003.** The geomophometric signature of Valles Marineris from M.O.L.A. DEM. *American Society for Photogrammetry & Remote Sensing, Anchorage, Alaska, May, 5-9, pp. 122-130.*

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude and appreciation to my academic supervisor, Associate Professor Martin J. Smith and to my internal examiner, Professor Alan Dodson for their continuous support and guidance throughout this project. Without their constructive criticism, the past 3 years would not have been so productive and successful.

I would also like to express my gratitude to Leica Geosystems Ltd, NPA Group Ltd, the Ordnance Survey, and Blom Aerofilms Ltd for providing the data for the Switzerland, London and Nottingham study areas respectively. Without their generous offer and continuous support this research would not have been so successful.

Finally but more importantly I would like to thank my family and my partner Ana, for their understanding and encouragement throughout my research. Their support was a major contribution that motivated me toward my personal goals.

LIST OF ABBREVIATIONS

DSM – Digital Surface Model
DTM – Digital Terrain Model
DEM – Digital Elevation Model
LiDAR – Light Detection And Ranging
VDDSM – Very Dense Digital Surface Model
IGN – Institut Geographique National (French National Mapping Agency)
OS – Ordnance Survey (British National Mapping Agency)
RGB – Red, Green, Blue
CCD – Charged Coupled Device
GIS – Geographic Information System
IMU – Inertial Measurement Unit
INS – Inertial navigation System
GPS – Global Positioning System
TIN – Triangulated Irregular Network
CAD - Computer Aided Design
VRML – Virtual Reality Modelling Language
LOD – Level of Detail
GCP – Ground Control Point
GSD - Ground Sample Distance
NDVI - Normalized Vegetation Index

Table of Contents

ABSTRACT **2**
LIST OF PUBLISHED PAPERS **3**
ACKNOWLEDGMENTS **4**
LIST OF ABBREVIATIONS **5**

1. INTRODUCTION **21**
1.1. Project introduction 21
1.2. Aims and objectives 24
1.3. Novelty and contribution to knowledge 30
1.4. Outline of the report 31

2. APPLICATIONS AND PREVIOUS RESEARCH EFFORTS **34**
2.1. Applications for 3D city models 34
2.1.1. Virtual Tourism 34
2.1.2. Noise pollution studies 37
2.1.3. Environmental Monitoring & Disaster management 40
2.1.4. Urban planning and Engineering 43
2.1.5. Simulation applications 47
2.1.6. Property management and cultural heritage 51
2.1.7. Navigation with portable navigation devices 55
2.1.8. Security and defence 58
2.2. Real time visualization of detail 3D city models 61
2.3. Towards a unified representation of 3D city models 63
2.4. Building reconstruction from single data sources 67
2.4.1. Building reconstruction from airborne optical data 67
2.4.2. Building reconstruction from LiDAR point clouds 73
2.5. Building reconstruction from fusion of different datasets 79
2.6. Commercial systems for semi-automated building reconstruction 88
2.7. Strengths and weaknesses of previous research efforts 90
2.8. Adopting previous methods in the proposed workflow for semi-automated 3D city modelling 92

3. AIRBORNE DATA COLLECTION SYSTEMS 94
3.1. Airborne Digital Sensor ADS40 94
3.1.1. Design principles 95
3.1.2. ADS40 Sensor Calibration 97
3.1.3. Post-Processing workflow of ADS40 data 99
3.1.4. Geometric accuracy and overall quality 100
3.2. Airborne frame film camera, Leica RC30 103
3.2.1. RC30 main Components 104
3.3. UltraCam D digital sensor 105
3.3.1. Design Concepts & Calibration 106
3.3.2. Accuracy assessment & performance 109
3.4. Airborne Laser Scanner - ALS50 113
3.4.1. Operating and design principles 114
3.4.2. Calibration and overall performance 116
3.5. Airborne Laser Terrain Mapper - Optech ALTM 3033 118
3.5.1. Operating and design principles 118
3.6. Error sources for airborne laser scanners 120
3.7. The significance of the selected data collection systems in 3D city modelling 125

4. STUDY AREAS AND AVAILABLE DATASETS 128
4.1. Study area (Heerbrugg, Switzerland) 128
4.1.1. Data Provided 129
4.1.1.1. ADS40 digital strip imagery 129
4.1.1.2. Airborne LiDAR data 131
4.1.1.3. Co-ordinate system for the study area 134
4.2. Study area (London – Bloomsbury) 135
4.2.1. Data Provided 136
4.2.1.1. Leica RC30 aerial images 137
4.2.1.2. ALTM 3033 LiDAR data 138
4.3. Study area (The University of Nottingham, University Park campus) 140
4.3.1. Data Provided 142
4.3.1.1. UltraCamD images 142

5. PRE-PROCESSING AND DATA PREPARATION 144
5.1. Switzerland data – Pre-processing and AT of the ADS40 imagery 144

5.2. London - Aerial Triangulation of RC30 scanned aerial photographs 148
5.2.1. Interior orientation..... 148
5.2.2. Automatic tie point extraction and GCP measurement 150
5.2.3. Aerial Triangulation results 153
5.3. The University of Nottingham campus - Aerial Triangulation of UltraCamD imagery 156
5.3.1. Interior orientation..... 156
5.3.2. Automatic tie point extraction and GCP measurement 158
5.3.3. Aerial Triangulation results 159

6. PROPOSED WORKFLOW FOR AUTOMATIC 3D CITY MODELLING AND ANALYSIS OF THE RESULTS 164

6.1. Overview of the proposed method 164
6.2. Extracting features from optical data 167
6.2.1. Selecting appropriate stereo pair 169
6.2.2. Applying edge detection algorithm 171
6.2.3. Optimised stereo matching 174
6.2.4. Discussion for the stage of semi-automated feature extraction from optical data..... 180
6.3. TIN Generalization and Very Dense DSM generation..... 182
6.4. Building detection and classification of LiDAR data and VDDSMs..... 189
6.4.1. Classification and building detection using high density LiDAR data..... 190
6.4.1.1. Delineating tree canopies from high density LiDAR 193
6.4.1.2. Classifying ground surface (DTM) from high density LiDAR data..... 195
6.4.1.3. Classifying low vegetation and background features from high density LiDAR data 199
6.4.1.4. Generating building hypothesis from high density LiDAR 200
6.4.1.5. Generating DSM from the initial building classification..... 202
6.4.1.6. Reclassification and raster to vector conversion of the DSM 204
6.4.1.7. Filtering and merging unclassified LIDAR points to the building class..... 205
6.4.1.8. Discussion for the stage of building extraction from high density LiDAR data 205
6.4.2. Classification and building detection using course LiDAR or VDDSM..... 207
6.4.2.1. Seed sells and histograms analysis 208
6.4.2.2. Region growing segmentation..... 211
6.4.2.3. Parametric representation and final building classification 219
6.5. Data fusion for optimizing building footprint and roof details 227
6.5.1. Data fusion process for High Density LiDAR 228
6.5.1.1. Extracting initial building footprint from High Density LiDAR..... 228
6.5.1.2. Spatial cleaning and generating topology for the 2D building footprints..... 231

6.5.1.3. Generalization and simplification of the building outline 232
6.5.1.4. Filtering stereo matched points with simplified building outline 234
6.5.1.5. Adjustment and refinement of the building outline 236
6.5.2. Data fusion for VDDSMs 238
6.5.2.1. Deriving the initial building outline from the raster building class 238
6.5.2.2. Fitting 3D Linear segments and TIN refinement to improve roof geometric detail ... 244
6.5.2.3. Converting refined TIN to a dense 3D point cloud 250
6.6. Building Reconstruction for generating final polyhedral models 252
6.6.1. Combining building footprint with 3D point cloud 253
6.6.2. Final 3D Building reconstruction 254
6.6.3. Merging adjacent roof planes utilizing Boolean functions 257
6.6.4. Discussion for the stage of building reconstruction 260

7. PROPOSED WORKFLOW FOR AUTOMATIC QUALITY ASSURANCE AND TEXTURE MAPPING 262
7.1. Semi-automated Quality Assurance workflow 262
7.1.1. Qualitative evaluation of the reconstructed building models 265
7.1.2. Semi-automated Quantitative evaluation of the reconstructed building 272
7.1.2.1. Planimetric accuracy assessment of the building footprint 272
7.1.2.2. Semi-automated vertical accuracy assessment of reconstructed building models 279
7.2. Automatic texturing using vertical airborne images 288
7.2.1. Automatic texture mapping and visualization for study area in Switzerland 291
7.2.2. Automatic texture mapping and visualization for the study area in London 292
7.2.3. Automatic texture mapping and visualization for the study area in Nottingham 294
8. CONCLUSIONS AND FURTHER PROSPECTS 300
8.1. Addressing the aims and objectives 302
8.2. Further Prospects 308
REFERENCES 309
APPENDIX I 326

List of figures

Figure 1 LOD definition for city models, adapted from Bildstein, (2005) 22

Figure 2 3D City model of Las Vegas, Nevada, USA. Adapted from Google Earth ™. Copyright Google Corporation 36

Figure 3 3D City Model of Las Vegas, Nevada, USA. Adapted from Microsoft Virtual Earth ™. Copyrights Microsoft Corporation 36

Figure 4. Photorealistic model of HK (left), 3D noise grid model projected on the City Model (right). Adapted from (Wing, 2006) 38

Figure 5 Coloured 3D noise grid model with respect to building heights (left), noise propagation from traffic represented with blue colour in Paris (right). Adapted from Butler (2004) 39

Figure 6 Visualization of flood analysis in Heerbrugg, Switzerland (test site 1). Semi-automatic building models from designed algorithms. Flooding levels are 2m, 5m, 7m and 10m for top left, top right, bottom left and bottom right pictures respectively. 42

Figure 7 Visualization of flood analysis in London, UK (test site 2). Semi-automatic building models from designed algorithms. Flooding levels are 1.5m, 1.8m and 2m for top left, top right and bottom pictures respectively. 42

Figure 8 Visualising landscape differences in an urban planning scenario. Models used from the London – Bloomsbury test site with and without trees. 45

Figure 9 The three views of (Verbree et al., 1999) applicable at different stages in urban planning. 45

Figure 10 An interactive model of the Tottenham Court Road region in London. The building highlighted with red colour can be moved from one location to another (adapted from Hudson-Smith and Evans, 2001) 46

Figure 11 Components of a virtual database for training and simulation purposes (adapted from Bildstein, 2005) 47

Figure 12 Example of LOD2, 3D city models for flight simulations. (Adapted from PC video game Flight Simulator 2004 ™, Copyright Microsoft Corporation) 48

Figure 13 Use of LOD2 & 3 for helicopter flight simulations. (Adapted from PC video game Flight Simulator 2004 ™, Copyright Microsoft Corporation) 49

Figure 14 An example of a driving simulator using LOD2 city models (Adapted from Bildstein, 2005) 51

Figure 15 General representation of geometry and topology for a 3D city model. Adapted from (Zlatonova, 1999). 52

Figure 16 3D building models in Manchester area linked to a relational database in VirtualGIS™ 53

Figure 17 3D city model of the ancient city of Augusta Raurica- Switzerland with modern orthoimages and height model (adapted from Salathé, 2001) 54

Figure 18 A sketch based navigation system using LOD2 3D city models. Look around the area (left), sketching a navigation command to walk along the street (right). (Adapted from Jurgen et al., 2005)..... 57

Figure 19 Use of LOD2 city models for military flight simulations (adapted from CAE incorporation 2008)..... 58

Figure 20 Use of highly detailed LOD4 building models in modern warfare PC video games (adapted from Call of Duty 4 – Modern Warfare™, Copyrights reserved by Activision Incorporated)..... 60

Figure 21 The logical data structure of a hybrid visualization system proposed by Gruen and Wang (1999) 62

Figure 22 Volumetric clouds using particles physics OpenGL engine (Adapted from TerrainView™, Viewtec)..... 63

Figure 23 A conceptual description for unified 3D city models (adapted from Qing et al., 2005)..... 64

Figure 24 VRML representation of a triangle mesh (Rossignac et al., 2001) 65

Figure 25 Models for building roof structures. a) gable, b) hip and c) flat (Willuhn and Van Gool, 2005)..... 70

Figure 26 Results from the building reconstruction process (Willuhn and Van Gool, 2005). 70

Figure 27 Diagram indicating an overview of the method by Taillandier and Deriche, (2004). 71

Figure 28 Coherency of the adjacent facets (Taillandier and Deriche, 2004). 72

Figure 29 Reconstructed buildings projected on an orthophoto based on the method proposed by (Taillandier and Deriche, 2004). 72

Figure 30 Final result from the reconstructed buildings overlaid on a DTM of the proposed method by NASA/ICREST Project report, (2001)..... 76

Figure 31 Gable roof building and the parameters for building modelling calculated from the invariant moments (Maas, 2004)..... 77

Figure 32 Final results illustrating the reconstructed buildings including the dormers (Maas, 2004)..... 78

Figure 33 Partitioned building outlines overlaid 82

Figure 34 Reconstructed building as resulted from the first method (Vosselman and Dijkman, 2001)..... 83

Figure 35 workflow of the method proposed by (Rottensteiner and Jansa, 2002). 84

Figure 36 Roof polygons back-projected for the model fitting step (Rottensteiner and Jansa, 2002)..... 85

Figure 37 Workflow of the proposed method (Chen et al., 2004). 86

Figure 38 Illustration for the ‘‘CyberCity Modeler’’. A saddleback roof is obtained by measurement of the points 1–6 in succession, in a single layer. For a hipped roof, eaves and

gable points are assigned different layers, for complex structures there are even more layers. From Grun and Wang (1998) 88

Figure 39 Illustration for inJECT. The operator chooses a primitive, which is then fitted to the images using manual measurement as well as automated measurement procedures..... 90

Figure 40 The major components of the ADS40 (ADS40 flyer, 2004) 94

Figure 41. Scanning principle of the three-line digital sensor (Sandau et al. 2000). 95

Figure 42 The trichroid beam splitter, adapted from ADS40 brochure, 2002. Copyrights reserved by Hexagon Corp. 96

Figure 43 Main components of the integrated IMU/GPS system in the ADS40, adapted from Sandau et al. 2000. 97

Figure 44 MTF along the flight direction (a) and across track direction (b) for the nadir looking view (Schuster and Braunecker, 2000) 98

Figure 45 Distortion in pixels estimated from the engineering models of DLR and LH systems for the nadir line (Schuster and Braunecker, 2000). 99

Figure 46 Ground processing workflow of the ADS40 dataset (Tempelmann, et al. 2000). 100

Figure 47. The Leica RC-30 aerial frame camera with navigation sight (adapted from Leica RC30 brochure, 2007, Copyrights reserved by Hexagon Corp.)..... 103

Figure 48. The UltraCamD sensor depicting the panchromatic and colour cones (Copyrights reserved by Microsoft Corporation)..... 105

Figure 49. Schematic of the 9 sub-images making up the panchromatic image with one image highlighted (Adapted from Smith et al. 2007) 106

Figure 50 The block of images taken at 1500m and distribution of control points (left), residuals of tie and control points from the different Aerial Triangulations using only the GCPs (right). (adapted from Smith et al., 2005) 109

Figure 51. Mean image residuals in 24x24 sub areas. On the left are results of AT without calibration model (low flown). The right diagram shows results of AT with IESSG calibration model (low flown) (adapted from Smith et al. 2007) 110

Figure 52 ALS50 system components representing the scanner assembly (lower left), equipment rack (right) and laptop control computer, adapted from ALS50 brochure, 2004, copyright reserved by Hexagon Corp. 114

Figure 53 Diagram illustrating the operating principles of the ALS50 (ALS40 brochure, 2002) 114

Figure 54 Estimated vertical and planimetric accuracy as a function of the field of view and operating altitude (ALS50 product description, 2003). 117

Figure 55 The latest ALTM sensor including the navigation system, onboard storage and sensor head. (Adapted from Optech, ALTM brochure 2008. Copyrights reserved by Optech Inc.) 118

Figure 56 The integrated ALTM sensor with the Applanix DSS camera (adapted from ALTM brochure, 2008)..... 119

Figure 57 Geometry of a zero offset. The actual laser beam (dotted line) and calculated direction (solid line) (Katzenbeisser, 2003). 122

Figure 58 Geometry for scaling factor. The actual laser beam (dotted line) and its calculated direction (solid line) for an offset (Katzenbeisser, 2003). 122

Figure 59 Mounting of IMU and deflection device on a carrying plate (Katzenbeisser, 2003) 123

Figure 60 Study area representing Leica's headquarter facilities and the surrounding region at Heerbrugg, Switzerland. (adapted from ©search.ch / Endoxon AG, TeleAtlas) 128

Figure 61 Panchromatic band from the strip imagery acquired with a forward look angle of 28° from the ADS40. Study area is highlighted in the red rectangle. 129

Figure 62 Study area represented by the forward looking panchromatic band. Subset of the entire scene with reference to figure 61. 130

Figure 63 Color coded, shaded relief image of the LIDAR point cloud obtained from the first flight path (Line ID: 050114_141115) 131

Figure 64 Color coded, shaded relief image of the LiDAR point cloud obtained from the second flight path (Line ID: 050114_141444) 132

Figure 65 Location of the GCPs (blue points) used for the vertical accuracy assessment.. 133

Figure 66. Overview of the second study area located in London-Bloomsbury. Adapted from Google Earth ™. Copyright Google Corporation 135

Figure 67. Example of complex building roof geometry present in the second study area.. 136

Figure 68. The available triplet of images for the 2nd study area over London-Bloomsbury. (Crown Copyright) 137

Figure 69 Subset of the original aerial photographs over London – Bloomsbury indicating the 'grain effect' (Crown Copyright) 138

Figure 70 The LiDAR data represented as a color coded raster surface and the elevation frequency histogram on the right using orthometric heights based on the OSGB36 vertical datum 139

Figure 71 3D perspective scene from the LiDAR data representing the overall density of the surface and highlighting the limitations for adequate representation of roof details. 140

Figure 72. Satellite view of the third study area located in The Nottingham University Park campus 141

Figure 73. The block of UltraCam images with the distribution of GCPs 142

Figure 74. Subset of an UltraCam D image over Nottingham test site representing one of the chemical engineering buildings of increased roof complexity 143

Figure 75 ADS40 data workflow for the production of the orientation data imagery. 144

Figure 76 ADS40 data workflow for the production of Level 1 imagery 145

Figure 77 ADS40 data workflow for the aerial triangulation 146

Figure 78 Location of the fiducial marks in the frame 148

Figure 79 Distributions of the GCPs 150

Figure 80 Distributions of the tie points in the stereo model 152

Figure 81 Ground residuals for control and check points from Trial 1 (scale = 10cm). 154

Figure 82 Ground residuals for control and check points from Trial 2 (scale = 10cm) 155

Figure 83 Camera co-ordinate system for UltraCam D as defined by the camera calibration report. 157

Figure 84 Distribution of measured control (Δ) and ckeck points (O) 158

Figure 85 Distribution of extracted tie points 158

Figure 86 Ground residuals for control and check points from Trial 1 (scale ----- = 5cm) 159

Figure 87 Ground residuals for control and check points from Trial 2 (scale ----- = 5cm) 160

Figure 88 Ground residuals for control and check points from Trial 3 (scale ----- = 5cm) 161

Figure 89 Ground residuals for control and check points from Trial 4 (scale ----- = 5cm) 162

Figure 90 Flow chart diagram indicating the proposed workflow for semi-automated 3D city modelling 165

Figure 91 Subset of overall workflow showing the process of feature extraction. Colours indicate the software used for each operation 167

Figure 92 Results from the implementation of the Sobel edge detector in the three study areas 173

Figure 93 Comparison of the optimised stereo matching process with and without the use of edges for the three study areas. 178

Figure 94 An example of inadequate extraction of stereo matched points for adjusting the building outlines over the Switzerland test site 180

Figure 95 Subset of the overall workflow indicating the proposed processing steps when LiDAR data is not available or dense enough. 182

Figure 96 Merging of LiDAR with stereo matched points in the London test site, all superimposed on the orthoimage. 183

Figure 97 Initial TIN (left) and the generalized TIN (right) from the second and third study areas. 184

Figure 98 Left, a TIN with one vertex represented by a white circle. Middle, the polygon that is created when the specific vertex is excluded. Right, a Delaunay triangulation of the polygon 185

Figure 99 The generated VDDSM for the second and third study areas from the generalized TINs 187

Figure 100 The calculated slope and aspect surfaces from the VDDSMs of the second and third study areas 188

Figure 101 Subset from the overall workflow indicating the building detection and classification of either LiDAR or VDDSMs 189

Figure 102 Diagram indicating the overall workflow of the building extraction process from high density LiDAR 191

Figure 103 Perspective scene of the combined LiDAR data, visualized as a colour coded shaded relief map. 192

Figure 104 Geometric shape of the generic tree models used to scan over the entire point cloud 193

Figure 105 Results from the tree detection using the two tree models over the LiDAR point cloud. Top view of the study area (left), TIN model with tree points superimposed (right) .. 194

Figure 106 Example of undetected individual trees resent in the study area 195

Figure 107 Diagram illustrating the iterative selection of new points at the ground surface 196

Figure 108 Resultant ground surface points (orange points) over the 1st study area, classified using the iterative selection algorithm 197

Figure 109 Classified points (red) representing the ground surface, overlaid on the shaded relief map of the LIDAR data and unclassified ground regions 198

Figure 110 Classified features and low vegetation with height in the range of 0-2.5m, superimposed on the TIN model for the study area 199

Figure 111 Remaining unclassified tree crowns and points representing building roof tops. 200

Figure 112 Top view of the initial building detection for the 1st study area, laser points representing buildings superimposed on the colour coded shaded relief map. 201

Figure 113 Perspective view of the initial building detection for the 1st study area, superimposed building points on the colour coded, shaded relief map 202

Figure 114 Remaining unclassified points (white features) representing individual trees and roof details overlaid on a TIN model for the study area 203

Figure 115 Digital Surface Model of the initial detected buildings and surrounding features in the background 203

Figure 116 Polygon layer of building hypothesis, produced form the raster to vector conversion process 204

Figure 117 Selected unclassified points located within the building hypothesis 205

Figure 118 Incorrectly classified building points (highlighted in blue ellipses), superimposed on a rectified true colour composite 206

Figure 119 The elevation frequency histogram for London and Nottingham study areas. Elevation is in range 12– 215 m and 18-86m respectively, the lighter a point is the greatest its elevation. 208

Figure 120 The difference DSM (VDDSM – Median DSM) for the second and third study areas shown in (a). Enlarged portions of selected buildings are shown in (c) and (d) with differences less than 1m depicted as white while differences greater than 9m and 3m are depicted as black for London and Nottingham respectively 210

Figure 121 Slope image for London and Nottingham on the left side (The higher the slope value the darkest a cell is). Frequency histogram of slope in the interval 1-90^{o} on the right side. 212

Figure 122 The resultant foreground and background object class from the implementation of the region growing segmentation workflow in second and third study areas..... 214

Figure 123 The resultant grouped and labelled foreground and background object class from the connecting-component labelling process in the second and third study areas..... 215

Figure 124 The resultant grouped and labelled foreground and background objects after the implementation of the object size filtering where background cells within enclosed foreground objects were merged to the foreground class..... 217

Figure 125 The merged image that contains both foreground and background objects after the implementation of size filtering and the components labelling algorithm. Different brightness values indicate different ID for each object..... 218

Figure 126 Visualization of the objects assigned to each of the 5 cluster (2,3,5,7,10) after the implementation of the K-means classification algorithm in London study area. White colour represents the objects of each cluster 224

Figure 127 Visualization of the objects assigned to each of the 5 cluster (2,3,5,6,7) after the implementation of the K-means classification algorithm in Nottingham study area. White colour represents the objects of each cluster..... 225

Figure 128. Diagram indicating the steps of the Data Fusion process 227

Figure 129 Perspective scene of reconstructed roof planes (red boundary), with building points superimposed over the ground TIN. 230

Figure 130 Extruded roof planes on the ground surface (grey lines) representing the vertical facades of the buildings 230

Figure 131 Resulting building outline from the procedure of topology generation and spatial cleaning in Switzerland test site 231

Figure 132 Simplified building outline (green boundary) compared to the initial building footprint (blue boundary) superimposed on a TIN model of the Switzerland study area 233

Figure 133 Oversimplified building outline in the London test site, as a function of the increased linear tolerance in relation to the size of the building footprint 233

Figure 134 Generated buffer zones of width 25cm at an increment of 25cm around the linear segments constituting the building outline..... 235

Figure 135 Filtered conjugate points derived from the stereo matching algorithm in LPS (red points) based on incremental buffer regions around the simplified building outline (green polyline) 236

Figure 136 Individual linear segments in Switzerland (blue lines) after the least squares adjustment using the stereo matched points (red)..... 237

Figure 137 Initial simplified building outline (green boundary) versus adjusted building footprint, superimposed on a TIN model 237

Figure 138 The raster polygons as created from the VDDSM classification process for London and Nottingham. Note each raster building object has been assigned a unique DN value 239

Figure 139 The resulted vector building polygons from the raster to vector conversion in London 240

Figure 140 The resulted vector building polygons from the raster to vector conversion in Nottingham 240

Figure 141 The result from the generalization of the building polygons (left) and the comparison with the initial building polygon prior to generalization (right) 241

Figure 142 An example of the filtered stereo matched points for adjusting the building outlines for London and Nottingham study areas 242

Figure 143 The adjusted building footprints (right) compared to the initial footprint (left) for London and Nottingham using the filtered stereo matched points 243

Figure 144 The initial 3D breaklines of the internal roof structures formed when the slope change is greater than 20 degrees. 245

Figure 145 The adjusted 3D breaklines forming closed polylines of roof details in London and Nottingham study areas 247

Figure 146 The optimized TIN surface of the building structures for London and Nottingham, for illustration purposed the TIN of the ground surface is also represented. 249

Figure 147 The dense 3D point cloud of the building structures created from the optimized TIN surface 250

Figure 148 Diagram indicating the overall workflow for the building reconstruction process 252

Figure 149 Projected building footprints at the ground level 253

Figure 150 Visualization of the reconstructed planes (hidden lines are excluded) for the three study areas. Successful reconstruction of small roof details (ventilation equipment, dormers etc.) 255

Figure 151 Perspective views for the three study areas (part of the study areas) with the reconstructed roofs and vertical building facades-hidden lines excluded 256

Figure 152 Resulting solid building models (rendered models- smooth shading) from the implementation of Constructive Solid Geometry at the planar facets. 257

Figure 153 Resulting solid building models (rendered models- smooth shading) from the implementation of Constructive Solid Geometry at the planar facets. 258

Figure 154 Perspective scene of the final building reconstruction for the entire study area in Switzerland. Rendered scene with global illumination and phong shading 258

Figure 155 Perspective scene of the final building reconstruction for the entire study area in Nottingham. Rendered scene with global illumination and phong shading 259

Figure 156 Perspective scene of the final building reconstruction for the entire study area in London. Rendered scene with global illumination and phong shading 259

Figure 157 Missing roof planes from 3D building model 260

Figure 158 Roof details in the 3D models 260

Figure 159 Deficiencies introduced as small intrusions during the plane merging function. 260

Figure 160 Examples from the semi-automatically reconstructed building models in London and Nottingham 261

Figure 161 Stereo plotted buildings for the three study areas. Rendered with a smooth shading effect and global illumination parameters 264

Figure 162 Qualitative comparison between the automatically reconstructed buildings and the reference stereo plotted 3D models for Switzerland 267

Figure 163 Qualitative comparison between the automatically reconstructed buildings and the reference stereo plotted 3D models for London 269

Figure 164 Qualitative comparison between the automatically reconstructed buildings and the reference stereo plotted 3D models for Nottingham 270

Figure 165 Example of improved planimetric accuracy at the adjusted building footprint. .. 277

Figure 166 Example of a raster DSM created from the individual roof planes 279

Figure 167 Difference images for buildings 10 & 15, indicating the vertical errors of the two models in Switzerland 281

Figure 168 Difference DSM of two buildings in the London study area representing vertical deviation between the semi-automatic and reference roof planes 284

Figure 169 Difference DSM of four buildings in the Nottingham study area representing vertical deviation between the semi-automatic and reference roof planes 287

Figure 170 Perspective view of the textured building models for the study area superimposed on the ortho-rectified true color composite and draped over a DTM 291

Figure 171 Detailed perspective view of the textured building models, superimposed on the orthorectified true color composite which was draped over a DSM 292

Figure 172 Perspective views of the automatically textured mapped building models in London 294

Figure 173 The automatically textured buildings models using the available Ultracam D airborne images with 80% overlap. 299

List of Tables

Table 1 Block configuration of the initial test sites (Fricker, 2001). 101

Table 2 Characteristics and accuracy results from the test flight in the area of Sion (Fricker, 2001) 101

Table 3 Characteristics of the block configuration for the Waldkirch test site. (Alhamlan et al. 2004). 102

Table 4 RMS for the forward/backward scene combination (Alhamlan et al. 2004). 102

Table 5. Tie and CGP analysis for the ATs using GCPs/GPS and IMU (adapted from Smith et al., 2005) 110

Table 6 Summary of high flight results 111

Table 7 Summary of low flight results 111

Table 8. Edge sharpness values present the width of an edge in pixels when considering distinct transitions from dark to bright areas. The analysis method is by Blonski et al. (2002) and is described by Perko (2005). 112

Table 9. Comparison of stereo matching results. “Fwhh” is the “full width at half height” of the histogram of errors. The columns with “Nr. of Points” presents the successful matches found in each image segment. Adapted from Perko (2005) 113

Table 10 Flight parameters during the acquisition of the LiDAR data over Heerbrugg-Szitzerland 133

Table 11 Results from the vertical evaluation of the LIDAR point cloud 133

Table 12 Parameters of the co-ordinate system used for the study area 134

Table 13 Aerial Triangulation results from the ADS40 imagery 148

Table 14 The interior orientation results from the automatic measuring process 149

Table 15 Report from the auto tie point extraction 152

Table 16 Image and ground residuals from AT results without self calibration – Trial 1 154

Table 17 Image and ground residuals from AT results Jacobsen’s self calibration model – Trial 2 155

Table 18 Image and ground residuals from AT results. No self calibration model – Trial 1. 159

Table 19 Image and ground residuals from AT results. Jacobsen self calibration model – Trial 2 160

Table 20 Image and ground residuals from AT results. No self calibration model – Trial 3. 161

Table 21 Image and ground residuals from AT results. Bauer’s simple self calibration model – Trial 4 162

Table 22 The results of the quality statistics from the implementation of the stereo matching process over the three test sites 180

Table 23 Attribute correlation matrix for the second and third area 220

Table 24 Final cluster (standardized) centroids and the number of objects in each cluster for London (first table) and Nottingham (second table) 222

Table 25 London study area - Distances between cluster centroids (divergence) and cluster compactness for clusters 2,3,5,7,10 223

Table 26 London study area - Distances between cluster centroids (divergence) and cluster compactness for clusters 2,3,5,6,7 223

Table 27 Results from the planimetric evaluation of the adjusted building outlines, derived from the combined LiDAR point cloud in Switzerland 273

Table 28 Results from the planimetric evaluation of the initial building outlines, derived from the combined LiDAR point cloud in Switzerland 273

Table 29 Results from the planimetric evaluation of the adjusted building outlines derived from the single path LIDAR point cloud in Switzerland 274

Table 30 Results from the planimetric evaluation of the initial building outlines derived from the single path LIDAR point cloud 274

Table 31 results from the planimetric evaluation using building detection metrics for the adjusted building outlines in London 275

Table 32 results from the planimetric evaluation using building detection metrics for the initial building outlines in London 275

Table 33 results from the planimetric evaluation using building detection metrics for the adjusted building outlines in Nottingham 276

Table 34 results from the planimetric evaluation using building detection metrics for the initial building outlines in Nottingham. 276

Table 35 Vertical accuracy characteristics of the semi-automatic building models derived from the combined LiDAR data in Switzerland. 281

Table 36 Vertical accuracy characteristics of the semi-automatic building models derived from the single path LiDAR data in Switzerland. 282

Table 37 Vertical accuracy characteristics of the semi-automatic building models derived from the study area in London 284

Table 38 Vertical accuracy characteristics of the semi-automatic building models derived from the study area in Nottingham 286

Table 39 Time requirements for unsupervised processing and manually adjusting parameters from the implementation of the proposed workflow in the test sites. 304

Table 40 Summary from the vertical accuracy assessment for all three test sites 306

1. INTRODUCTION

1.1. Project introduction

More than 50% of the world population lives in urban/suburban areas, so detailed and up-to-date building information is of great importance to every resident, government agencies, and many private companies. Government agencies and private companies spend millions of dollars each year obtaining aerial photographs and other forms of remotely sensed data. The large amount of airborne data, acquired every year, require automated solutions for post-processing in order to create 3D city models in a consistent and cost effective way. One of the key components of any 3D City Model is a three dimensional representation of the buildings present in the scene, the process of creating these 3D building models that includes a geometric three dimensional representation of roof and facades details is often referred to as building reconstruction. Building reconstruction is of primary importance in many varied applications, ranging from urban planning and telecommunication network propagation studies to application for next generation vehicle navigation. Currently the production of detailed 3D building models of urban areas is very expensive and time consuming and automated workflows are required in the industry.

Unfortunately, manual reconstruction from photogrammetric techniques is time consuming and not a cost effective solution. It is estimated that an operator will have to commit several days or even weeks of stereo plotting for the production of every square kilometre of 3D city models. These time requirements justify the significant cost associated with these datasets and the main reason behind the lack of widespread use in different industries. It is evident that automated techniques and tools for data acquisition from remotely sensed imagery are needed and is the focus of many current research efforts. Automatic building reconstruction from multiple aerial images is however a difficult problem due to the presence of shadows, occluded regions introduced by vegetation and the perspective geometry of the scene. There are also issues regarding accuracy and consistency of the reconstructed building

models that need to be addressed. Most automatic approaches have to account for the complexity of roof structures and thus supply a general approach that can model most of the buildings present on a scene. At the same time simplifications of the roof shapes is needed for providing a fast and reliable solution. Currently most approaches are not designed to reconstruct roof details for dormer windows, chimneys and small building recesses, except where very high density LiDAR data are available. Therefore the goal of an automatic solution is to provide a generic 3D building reconstruction function that can be efficient in a variety of different situations. It should present reliable results with data collected from different sensors and adapt to projects with different specifications regarding the geometric accuracy and the desired level of detail. It is important at this point to introduce the concept of Level Of Detail (LOD) in the context of 3D city models which has been defined in Bildstein, 2005. The definition of these categories was based on the evaluation of several applications related primarily to augmented reality and simulation but it can be adaptable to a variety of different applications. This thesis is concerned with the automation of 3D city modelling for the LOD1 and LOD2 models as illustrated in figure 1.

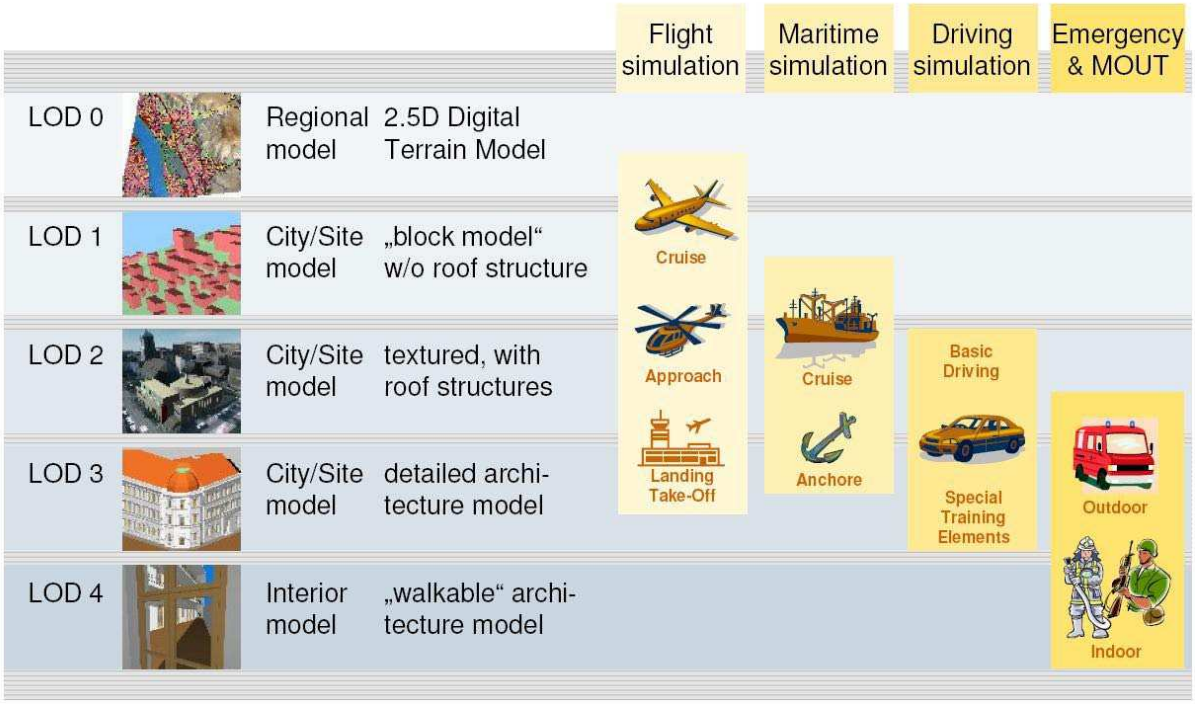


Figure 1 LOD definition for city models, adapted from Bildstein, (2005)

Methods for building detection and reconstruction can be applied to LiDAR point clouds. Typically existing methods either directly derive the surface parameters in a parameter space by clustering the point cloud or segment a point cloud based on criteria like proximity of points or similarity of locally estimated surfaces (Vosselman et al., 2004; Heuel et al., 2000). Kokkas and Dowman (2006) introduced a method that employs a semi-automated technique for generating the building reconstruction by fusing LiDAR data with stereo matched points extracted from the air photograph stereo model. The roof reconstruction is achieved by implementing a least squares plane fitting algorithm on the LiDAR point cloud and subsequently neighbouring planes are merged using Boolean operations for the generation of solid features. To complement this modelling various quantitative techniques have been developed in order to automate the interpretation of terrain features from DEMs using geomorphometric parameters in an attempt to characterize the landscape (Miliaresis and Kokkas, 2004). Most approaches in order to overcome certain difficulties introduce some kind of external knowledge, either as models of buildings (with gable, hip or flat roof shapes) which reduce the generality of the approach, or as constraints on primitives extracted from aerial images.

Data fusion techniques between different data sources can be an effective solution for automated building reconstruction since they combine the advantages of very good vertical accuracy, usually provided from LiDAR data, and the precise location of the building outlines and roof details, derived from aerial images. Building reconstruction using LiDAR, aerial and satellite imagery was introduced by Chen et al. (2004). A method proposed by Rottensteiner and Jansa, (2002) combines LiDAR with aerial photographs and performs the polyhedral modelling using Voronoi diagrams. Sohn and Dowman (2006) merges LiDAR with high resolution satellite data, and introduced a method for Binary Space Partitioning (BSP) while a combination of calibrated aerial images and DSM, produced from LiDAR datasets, was introduced by Institut Geographique National (French National Mapping Agency, IGN) for several test sites (Kaartinen et al., 2005). Although the effectiveness of data fusion between LiDAR and aerial photographs has been illustrated the potential high cost associated with

LiDAR data collection and, in addition, aerial imagery is a limiting factor for the commercial implementation of these approaches. Hence one aim of this research is to exploit the synergy between mapping products extracted using only high resolution airborne imagery.

Due to varied complexity of different urban scenes, automatic or semi-automatic workflows for building reconstruction should be accompanied by robust quality assurance tools. So far most accuracy assessment studies have been focused on assessing 3D city models using methods for assessing either the planimetric accuracy of the building footprints or selectively analysing roof shapes. This research investigates the implementation of a robust semi-automated approach for a comprehensive planimetric and vertical accuracy assessment proposed by Kokkas and Smith (2007). The methods described in this report were designed for the lower and average LOD city models up to LOD2.

1.2. Aims and objectives

The main aim of the project is the development of a semi-automated, innovative workflow for building reconstruction and 3D city modelling. The complexity of this aim dictates a staged development from manual reconstruction to an increasing level of automation. The main objective set from the beginning was to evaluate what level of automation is achievable using different workflows and datasets in order to perform 3D reconstruction and texture mapping and produce models that would satisfy a varied range of applications.

The overall scope of the project is nevertheless limited by the data sources available, obtainable resolution, accuracy and overall detail. Due to timeframe constraints it was decided that this project would be restricted to use only data collected from airborne platforms. This decision was also made to satisfy the requirements of the fourth objective (see below) concerning cost efficiency. By setting the above limitation it was immediately evident that the solution would not be able to provide models detailed enough for

applications requiring LOD3 & 4 models (see figure 1), due to the use of airborne data. LOD3 & 4 models require geometric detail for building facades and interiors currently not achievable using only airborne data. This decision would ultimately set the scope of the project with consideration to the overall research timescale and the foundations for feasible milestones. Nevertheless there are multiple airborne data sources that either individually or using data fusion techniques could provide the desirable LOD models. All these limitations led to the identification of realistic objectives based on the datasets used and the overall timeframe as outlined below.

Project objectives:

1. **The final solution would need to have a degree of automation exceeding current industry standards and previous research efforts**
2. **The solution should present comparable accuracy and LOD characteristics with 3D models extracted from manual stereo plotting procedures and a method to perform accuracy assessment should be proposed.**
3. **It should be a flexible and reliable solution that can be accommodated in different situations with different dataset characteristics (resolution, quality and number of data sources), adaptable to different scenes with different building architecture.**
4. **Cost efficiency and availability of data sources should also be considered.**

Objective 1: The final solution would need to have a degree of automation exceeding current industry standards and previous research efforts. The first objective is very broad and further description was needed to define exactly the desired level of automation. Automated or semi-automated methods are sometime difficult to quantify and this project would define it as a threshold level related to the time required to process the data and the time spent by a single operator performing manual stereo plotting. For one square kilometre of LOD2, 3D city models (geometry only) a single operator using state-of-the-art stereo plotting facilities needs at least 40 working hours (this timeframe was verified internally by

manual 3D plotting performed at three test sites by the experienced researcher). A semi-automated algorithm would be able to reduce the time requirements by at least 50% but still require the operator to perform several functions and setting parameters during these 20 hours. In contrast an automated algorithm could be named as such if it could reduce the overall processing time by at least 50% and maintain the time spent by the operator for setting parameters within 1%. Essentially for one square kilometre of 3D city models a workflow would qualify as automatic if the operator was spending no more than 30 minutes setting parameters and altering variables and the unsupervised processing of the data was complete in less than 20 hours.

Objective 2: The solution should present comparable accuracy and LOD characteristics with 3D models extracted from manual stereo plotting procedures and a method to perform accuracy assessment should be proposed. This objective was defined to set the milestones and the scope of the accuracy assessment methodology and the level of detail this solution should satisfy, although the achievable accuracy is mainly a function of the datasets used as well as the technique employed to extract features. In order for the solution to compete with manual methods it would have to achieve comparable results. Consider that high resolution state-of-the-art digital sensors, flying at a typical flying height (1000m) can achieve centimetre resolution and features extracted from these stereo pairs would have an estimated accuracy of 10-20cm on the ground. Features and shapes extracted from LiDAR data have been shown to have absolute Z accuracy in the order of 5-10 cm (Maas, 2002). Based on the above it is safe to assume that manual photogrammetric techniques employing airborne datasets can create 3D city models with an absolute vertical and planimetric accuracy of 20-25cm. Apart from the accuracy characteristics another important factor is the completeness of the model at the LOD level. Due to the limitations of airborne datasets even manual photogrammetric techniques are restricted to LOD 2 models. That is models with roof details and textures, but with no geometric detail for the facades or interiors. Considering the above characteristics for accuracy and LOD and in order for the

solution to be adaptable in the industry and not purely stay a research prototype, the project adopted the objective that the derived models would be quantitatively assessed. The overall absolute errors (standard deviation) should not be worse than 50cm (for height and plan). The derived models would also need to have a comparable level of detail with the manual models, hence the solution should be able to generate LOD 2 city models, including the necessary geometric detail and textures. At this point it's useful to consider that for most applications and projects requiring 3D city models the accuracy specifications are rarely defined, instead LOD is the predominant factor with the geometric quality in most cases unjustifiably overlooked.

Objective 3: It should be a flexible and reliable solution that can be accommodated in different situations with different dataset characteristics (resolution, quality and number of data sources), adaptable to different scenes with different building architecture. In order to meet the requirements of the third objective three distinct test sites were chosen with different architecture and varied roof details. The first site is situated in Switzerland, Heerbrugg and represents Leica Geosystems headquarters and the surrounding area. The test site has a mixture of large industrial buildings and small cottage style houses. The second test site was in London, Bloomsbury, a challenging site due to the varied details of the building structures. The third test site was situated at the University Park of The University of Nottingham consisting of a mixture of different building types.

Objective 4: Cost efficiency and availability of data sources should also be considered. The fourth objective is an assessment of the cost efficiency of the proposed methodology. This project is primarily concerned with exploiting the synergy between LiDAR and aerial photography. Data Fusion techniques for building reconstruction provide, in general, more accurate results since they combine the advantages of very good vertical accuracy characteristics, usually obtained from LiDAR data, with the precise location of the building outlines, derived either from aerial images or ground plans. A quantitative evaluation of extracted building footprints presented in Kaartinen et al. (2005) indicates the higher

geometric accuracy achieved with data fusion approaches. Considering the promising results introduced using data fusion of different datasets for building reconstruction, the project adopts a method that combines aerial imagery with LiDAR data for the first two test sites. For the Switzerland test site, airborne imagery from the ADS40 sensor was available. This sensor represents the state of the art in airborne pushbroom sensors. LiDAR data were also collected using the ALS50. The selection of the acquisition systems was related to the fact that they can be easily integrated and mounted in a single airborne platform for simultaneous data collection. Furthermore digital imagery acquired with the ADS40 is well suited for building reconstruction purposes mainly because of the three line pushbroom acquisition geometry that minimises occluded regions behind buildings when successive strips of data are available and due to the direction of the relief displacement. The three look angles (backward, nadir and forward) can effectively minimize occluded areas and provide multiple stereo pairs of imagery for feature extraction. Apart from the minimization of the occluded regions ADS40 can acquire multispectral information by capturing one band in the near infrared which can be very useful for vegetation discrimination using a classification method. The effectiveness of the roof reconstruction process and the level of extracted roof details is affected by the density of the LiDAR point cloud. The selection of the ALS50 was based on this characteristic as it can acquire data with a maximum pulse rate of 83 kHz, which was considered one of the highest pulse rates in the industry when the data were chosen in 2006. Today's LiDAR sensors can achieve up to 120kHz pulse rates. For this specific project the LiDAR point cloud has a density of approximately 5 points/m^2 by combining two crossed flight paths. For the London – Bloomsbury test site scanned aerial photographs acquired from a Wild RC30 aerial camera were available. Post-processed LiDAR data with 1m density were also obtained from the UK Environment Agency using an Airborne Laser Thematic Mapper 3033 (ALTM 3033) manufactured by Optech. These datasets were deliberately chosen to contrast with the high quality data available in the first test site. So that a study of the impact of an alternative, or more widely used (traditional) datasets, could be made.

So the main objective for using the first two test sites was the evaluation of fusing conjugate points extracted from the aerial images with the LiDAR point cloud for automated building reconstruction. Emphasis is placed on the individual processing steps and the entire workflow is carefully designed so that it addresses the issues of reliability, robustness and efficiency of time. These three parameters are crucial in the design of an automated procedure as well as the geometric accuracy of the reconstructed buildings.

What this workflow didn't address at that point though is cost-efficiency. The issue with LiDAR data is that they are considerably more expensive to capture, compared to aerial photographs. Consider that the cost to purchase off-the-shelf aerial photographs per sq.km, as of late 2008, is approximately £100 but for LiDAR is double that to £200. This price difference equates to the higher cost of operating the aircraft at lower altitudes and extended period of time to cover the necessary area since LiDAR sensors have smaller coverage compared to aerial photos. Therefore considering the cost factor of purchasing LiDAR data, the project for the third test site adopts a modified workflow to utilise high resolution airborne imagery acquired by the UltraCam D sensor. This dataset represents the state of the art digital sensors based on frame CCD architecture and the acquisition of the UltraCamD imagery is less expensive than purchasing LiDAR data.

It is evident that a lot of effort was put towards identifying adequate data sets and test sites to satisfy most of the project objectives. The final objective of the project was the design of a robust quality assessment workflow and assessing the effectiveness of vertical aerial photographs for automatic texture mapping.

1.3. Novelty and contribution to knowledge

This section describes the main innovative processes and algorithms introduced in the proposed workflow for semi-automated 3D city modelling. Overall the novelty of the proposed workflow is that it is synthesized from a number of existing and new optimised individual processes that have been combined in such a way to present a unified end-to-end workflow for complete 3D city modelling. The initial working prototype of this workflow has been submitted to the UK Intellectual Property office for consideration of a patent. The patent evaluation process is nearly complete and since no objections were raised it is highly probable that a patent will be granted within 2009 (Kokkas, 2007).

Evaluating the novelty of the individual processes comprising the proposed workflow, the following assertions can be made.

- **Edge extraction for image enhancement prior to stereo matching –** This process was devised in order to enhance the contrast between the linear features visible on the aerial images and the background objects. Although edge operators have been widely used as documented, the successful enhancement of the linear features prior to the stereo matching process is indicating that this innovative process could be used on a regular basis for enhancing the quality and density of the stereo matched point cloud.
- **Building detection process –** Although there have been several attempts to identify building models from optically derived data, this research presents an innovative generic process for identifying the building class from LiDAR data or Very Dense DSM (VDDSM) created solely from airborne imagery. More specifically the author makes use of an adaptive k-means classification algorithm in combination with geomorphometric region growing and parametric object representation to enhance the reliability of detecting building footprints.

- **Data fusion for optimising building footprint and 3D breaklines –** One of the most innovative stages introduced is a new method for fusing the LiDAR data with high density stereo matched points from aerial photographs. The novelty in this stage specifically is a new approach to creating a TIN from the merged data and generalising it in order to remove any blunders and registration issues between the two datasets. TIN generalization takes place using generic GIS functions that can be adapted in many systems available in the market. This new data fusion method was performed in order to improve the accuracy of the building footprint as well as the accuracy of the reconstructed 3D breaklines that are used subsequently to enhance the Level Of Detail of the reconstructed roof shapes.
- **Semi-automatic quality assurance –** Another important aspect introduced in this research is the proposed workflow for accuracy assessment. Apart from the qualitative visual assessment and the use of building detection metrics, which are well known in the industry, the report introduces a new innovative process for a complete quantitative vertical accuracy assessment. The semi-automatic vertical accuracy assessment is based on the generation of TIN models for each building that are subsequently subtracted from reference building models. The novelty of the proposed method for quality assurance is also represented by its easy adaptation to existing open source GIS environments since it's using standard GIS techniques to achieve a comprehensive quantitative evaluation.

1.4. Outline of the report

This section summarises the contents of the eight chapters included in this report. The first chapter provides an introduction to the challenges confronted for 3D city modelling and automatic solutions. The aims and objective are also stated in the first chapter.

The second chapter consists of an introduction to the applications requiring 3D City models, including their requirements and LOD for each case. It also provides a comprehensive literature review containing information about manual techniques and automatic or semi-automatic workflows using one or multiple datasets. This chapter presents information regarding alternative methods and data sources for building reconstruction including, terrestrial methods, Synthetic Aperture Radar and shape from shadow techniques. It also places the work presented here in the context of the work of others, presenting previous research efforts that were adopted for further refinement in the proposed workflow.

The third chapter gives details of the principles and design concepts of the sensors used to acquire the datasets for this research. Specifically the chapter contains information regarding the ADS40 pushbroom sensor, UltraCam D, Leica RC30 frame camera, the ALS50 LiDAR sensor, and the ALTM 3033.

The fourth chapter provides information on the three test sites selected with information on building architecture and the variations of each test site that contributes to the overall objective of a generic and adaptable solution.

Chapter five describes all the pre-processing stages necessary in order to orient the available imagery and perform the aerial triangulation for the different test sites. The pre-processing stage includes all the operations necessary to perform accurate aerial triangulations using self-calibration and additional parameters to obtain the most accurate results. Although this stage is not part of the designed automated workflow, this process is nevertheless essential for the accuracy required in the final 3D models.

Chapter six describes in detail the workflow and algorithms designed for automatic 3D city modelling. It provides initially an overview of the proposed method and then details the extraction of features from optical data, classification of LiDAR data for building detection, the generation and classification of very dense digital surface models, data fusion and building reconstruction algorithms. The proposed algorithms and discussion is given for each

stage. The sixth chapter also presents detailed results and critical analysis from the implementation of the designed workflow in each of the three test sites. For each stage a critical analysis and a discussion of advantages and drawbacks for the specific areas is presented.

The seventh chapter presents detailed results from the algorithm designed for semi-automatic accuracy evaluation of the automatic 3D city models. It includes a qualitative or visual comparison of the automatic 3D city models with reference models (manually created) for each of the three test sites. The 3D models from the test sites are quantitatively assessed with results for their planimetric and vertical accuracy. This chapter includes a method for automatic texture mapping using vertical airborne images. The final chapter (chapter eight) describes the conclusions and further prospects for future research in this field.

2. APPLICATIONS AND PREVIOUS RESEARCH EFFORTS

2.1. Applications for 3D city models

Compared to the application of 2D spatial data, the application of 3D city models is a communicative process that relies on the efficiency of presenting a specific meaning that can be transferred from creators and organizers of information to users and receivers of the same information (Zhu & Lin, 2004).

Many practical applications show that data content definitions of 3D city models are based on the direction of the particular application. A report by Batty et al. (2000; 2002) defined twelve different categories of applications, including emergency response, urban planning, telecommunication, architecture design, facilities management, etc.

The description of 3D city models for various applications often has distinctive characteristics. For instance, the geo-visualization applications stress the display of the real-world realistically and authentically, whereas analysis applications, such as flood modelling or signal propagation, aim at the representation of the real-world in an abstractive and compact way (Qing et al., 2005). Traditional categories of 3D city model applications usually depend on the function of application, but the classification is random and cannot be unified (Batty et al. 2002). Different groups of professionals belonging to different disciplines have great difference in understanding the details, and the application type is changing ceaselessly. A series of example applications will now be discussed to highlight these issues.

2.1.1. Virtual Tourism

The tourism industry has seen recently the immergence of ‘virtual tourism’ as an alternative to interactive exploration of distant places. The tourism industry was essentially about providing people with experiences. The tourism industry found it difficult to deliver on its

promised 'dreamlike' experiences, it has had to provide not only alternative, but also substitute and surrogate experiences and destinations. The industry has even encouraged their development, for example by building theme parks modelled on foreign destinations (Williams, 1995)

In recent years several web-based portals offering interactive 3D virtual environments have been introduced. Tourists planning their holidays can benefit from web-based interactive visualization tools with 3D city models. It can give prospective tourists a preview of their holiday area or the hotel they want to visit. They can also assess the best ways to explore and travel around the area. This is all thanks to the availability of spatial and descriptive information, available through a virtual 3D model of the holiday destination via the internet. A typical web-based tourism information system could combine a 2D map window with a 3D scene. Points of Interest of various categories can be displayed in both 2D and 3D.

From a marketing perspective, Virtual Reality (VR) has the potential to revolutionize the promotion and selling of tourism. Tour operators and travel agents using this method would have the ability to offer potential tourists a simulated experience of their planned trip. Unlike brochures and videos which are passive tools, VR offers the ability to offer an interactive experience (Hobson, 1995).

Nevertheless there are several issues with the virtual tourism concept. These applications tend to require higher Level of Detail (LOD) models (see section 1.1), usually LOD3 or even LOD4 which represent interior details of shopping malls, hotels and other tourist attraction through an immersive photorealistic environment. Occasionally LOD2 models are used for larger areas. Recent examples of online platforms that have the potential to be used for virtual tourism include Google Earth™ and Microsoft Live Local™ shown in figures 2 and 3 respectively.

Figure 2 3D City model of Las Vegas, Nevada, USA. Adapted from Google Earth ™. Copyright Google Corporation

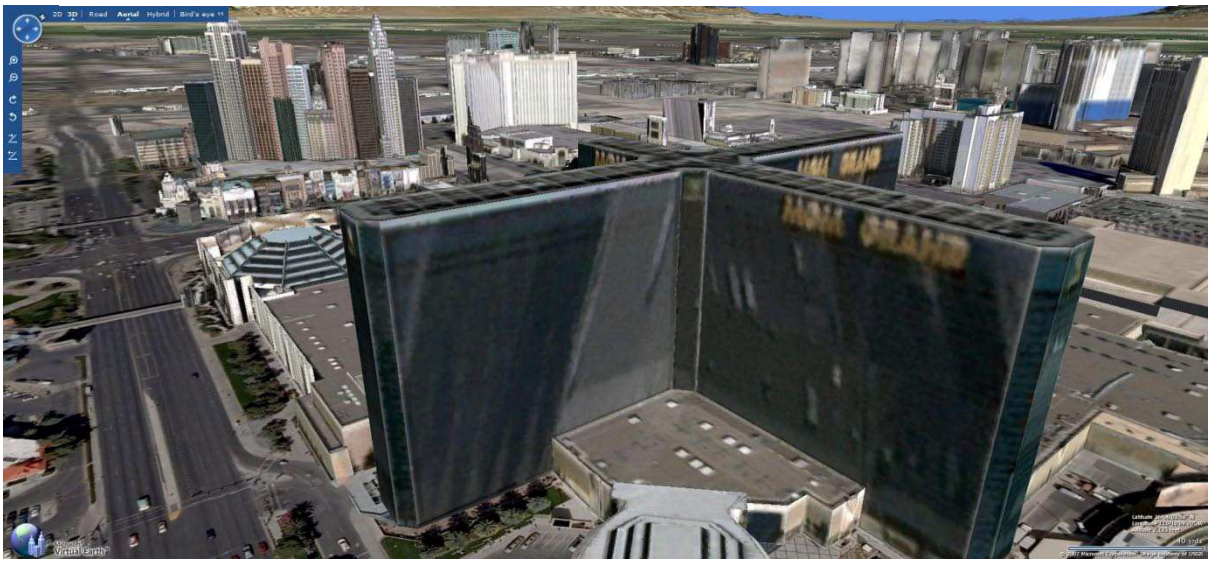


Figure 3 3D City Model of Las Vegas, Nevada, USA. Adapted from Microsoft Virtual Earth ™. Copyrights Microsoft Corporation

Although these two platforms may not yet have the necessary level of detail to satisfy a fully immersive requirement for virtual tourism they could be effectively used as the basis for introducing and updating the virtual world with more detailed models.

There is an apparent difference in the way the real world can be represented in a real time 3D rendering environment and this is illustrated when comparing figures 2 and 3. Figure 2 presents detailed 3D building models with roof details including some facade architectural details without though using any texture from images. Through a visual examination of the buildings present in figure 2, it is easily observed that these building models have been manually modelled, with respect to their exterior detail, using geometric primitives and solid colours. This method of building modelling, although very expensive and time consuming, provides very efficient models for real time rendering with small file sizes. The introduction of real time shading effects, using new generation graphics cards, assists also in the photorealism of these models. Despite the shading effects, these types of models tend to have a repetitive pattern and in most cases look artificial. In contrast figure 3 presents a different modelling method which incorporates texture from aerial photographs, draped automatically on the models, to give information about the building facade. Although this method is less expensive and more time efficient the quality of the textured facades is not always suited for virtual tourism when the observer expects a fully immersive realistic environment to explore. Although the introduction of terrestrial images can provide a high quality result consideration should be made about the efficiency of such models for real time rendering in a large scale.

Issues of real time visualization are presented in section 2.2, describing recent advancements and proposed solution for large scale applications.

2.1.2. Noise pollution studies

Noise pollution of urban areas is one of serious factors that the local agencies and state authorities have to consider in decision making processes. Large cities have very often severe problems with ubiquitous traffic noise from multiple sources. Noise pollution in large urban areas is regarded as a growing problem. Present studies have shown that more than 20% of the world population lives under unacceptable noise levels and nearly 60% of the

European population is exposed to high noise levels during the day (Silvia, 2003). In Europe, it is estimated that more than 90 million people suffer from unacceptable noise levels (European Commission News, 2005).

There have been several recent efforts to introduce 3D city models for noise studies. In Hong Kong photorealistic 3D models were used to build 3D noise models (Wing, 2006). GIS data including buildings, roads, terrain, podiums and barriers were used for the Calculation of Road Traffic Noise (CRTN). Wing (2006) calculated a 3D noise grid model taking into consideration the traffic composition, flow, speed, noise propagation, etc. All the 3m by 3m noise grids were located vertically one metre from building façades, and their colours are categorized according to the calculated noise levels, indicating the traffic noise distribution on the façade of buildings. The 3D grid noise model was projected on the photorealistic 3D model of Hong Kong for visualization purposes as illustrated in figure 4.

Figure 4. Photorealistic model of HK (left), 3D noise grid model projected on the City Model (right). Adapted from (Wing, 2006)

A 3D noise model was also implemented in Paris, (Butler, 2004) using aerial photographs to extract the building models. The most accurate way to map noise would be to place microphones every few metres throughout a city but that would take years and be expensive. Instead, noise analysts resort to 'virtual microphones', each of which is a point in a computer model that reports what the sound level would be at a certain place under given

circumstances. The 3D noise map (figure 5) introduced by Butler (2004) added microphones up the sides of buildings at 3-metre intervals vertically and every 10 metres horizontally. In total, the 3D representation of Paris contains 26 million virtual microphones. Each microphone bases the sound level it reports on computer models of how sound from nearby noise-makers should behave. It's impossible to model all noise sources, such as neighbours arguing and pneumatic drills. So analysts are first tackling the traffic noise, which accounts for some 90% of urban noise pollution.

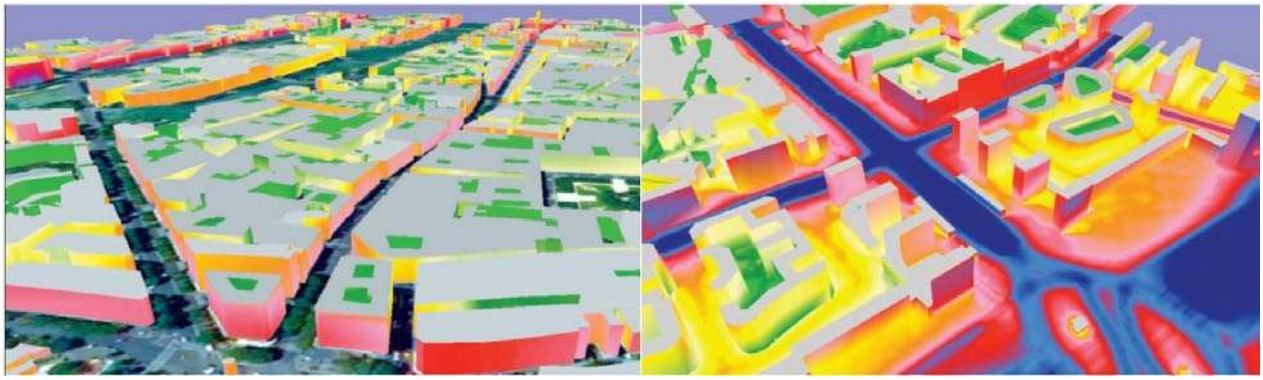


Figure 5 Coloured 3D noise grid model with respect to building heights (left), noise propagation from traffic represented with blue colour in Paris (right). Adapted from Butler (2004)

The requirements for noise simulations have been specified by the European Commission which has passed legislation that requires the member states to create noise maps for all major cities, roads and railways and airports by 2007. On 13th January 2006, 'European Commission Working Group Assessment of Exposure to Noise (WG-AEN)' has published a position paper on guidelines for strategic noise mapping (European Commission Working Group, 2006). In this report, although there are no specific requirements set for 3D city models, it indicated the need for the development of new methods for converting 2D models into 3D models using a variety of different techniques such as interpretation of height attribute information, laser scan data, use of textual height information, etc. The report indicates that the height of buildings can have a significant effect on the propagation of noise particularly in built up areas. Simplification of building heights is often carried out to reduce computation time when carrying out noise mapping so roof details are not always

desirable. Nevertheless the accuracy of the average building height is critical for the calculated accuracy of the noise level map which is estimated at 1.5dB (European Commission Working Group, 2006).

In addition, the report specifies that when computation methods are used for the purpose of strategic noise mapping in relation to noise exposure in and near buildings, the assessment points must be at the most exposed façade. For this purpose, the most exposed façade will be the external wall facing onto and nearest to the specific noise source. The requirement for architectural details for the building facades indicates the need for LOD3 building models.

2.1.3. Environmental Monitoring & Disaster management

With the global climatic changes currently becoming apparent, a protection against natural disasters like floods is necessary in both developed and developing countries. Globally, natural disasters kill an estimated one million people and leave millions more homeless each decade. Floods killed more than ten thousand people, displaced 20 million and caused 82 billion US dollars worth of damage in the year 2005 alone (Brakenridge et al., 2006). Historically, flooding has superseded all other disasters in terms of frequency of deaths, property destruction, and land degradation (Douben and Ratnayake, 2006)

In a continuously changing and unpredictable environment the quality and availability of state-of-the-art & up-to-date 3D geographic information is of vital importance in flood modelling and natural disaster monitoring. 3D city models in combination with detailed Digital Terrain Models (DTMs) serve as a basis for numerical flood inundation simulations. An urban area flooding simulation is even more resource demanding and the conventional one-dimensional (1D) hydrodynamic models are considered to be inadequate for simulating wave propagation over complex terrains, such as urban areas (*Bates and De Roo*, 2000). In contrast Yu and Lane (2006) developed a simplified 2D diffusion-wave based urban fluvial flood model, but admitted that a more sophisticated representation of the inundation process is required to effectively simulate the flooding process in an urban terrain.

Using detailed 3D city models, it is possible not only to predict whether parts of a city will be affected by the flood or not but also how severe the impact on the buildings will be. These simulations help to plan efficient flood protection measures. The 3D city models are usually integrated on the DTM of the study area and form a Digital Surface Model (DSM). The DSM is essential for the estimation of the flood volumes on the surface areas. In addition, the results are usually in the form of a flood inundation map that is based on water levels from the model simulation in conjunction with the DSM.

Thus the quality of the model depends on the quality and accuracy of the DSM and the 3D City models that represent most of the man-made features. As shown by Mark et al. (2004) the interval of the spot heights should be small enough to account for features such as road curbs and the desirable accuracy of the DSM should be in the range of 10-40 cm. The most important requirements for 3D City models used in flood simulations is the need to represent the elevation at the bottom and curb level of the road systems as well as modelling of the building façades and architectural details that might affect the flood model. For this application 3D building models rarely have roof details as these don't tend to influence the flood analysis. An illustration of a flood simulation for visualizations purposes is demonstrated in figures 6 and 7. The figures depict several flooding scenarios using the semi-automatically constructed building models presented in this report in two test sites, in Heerbrugg, Switzerland and London, UK.

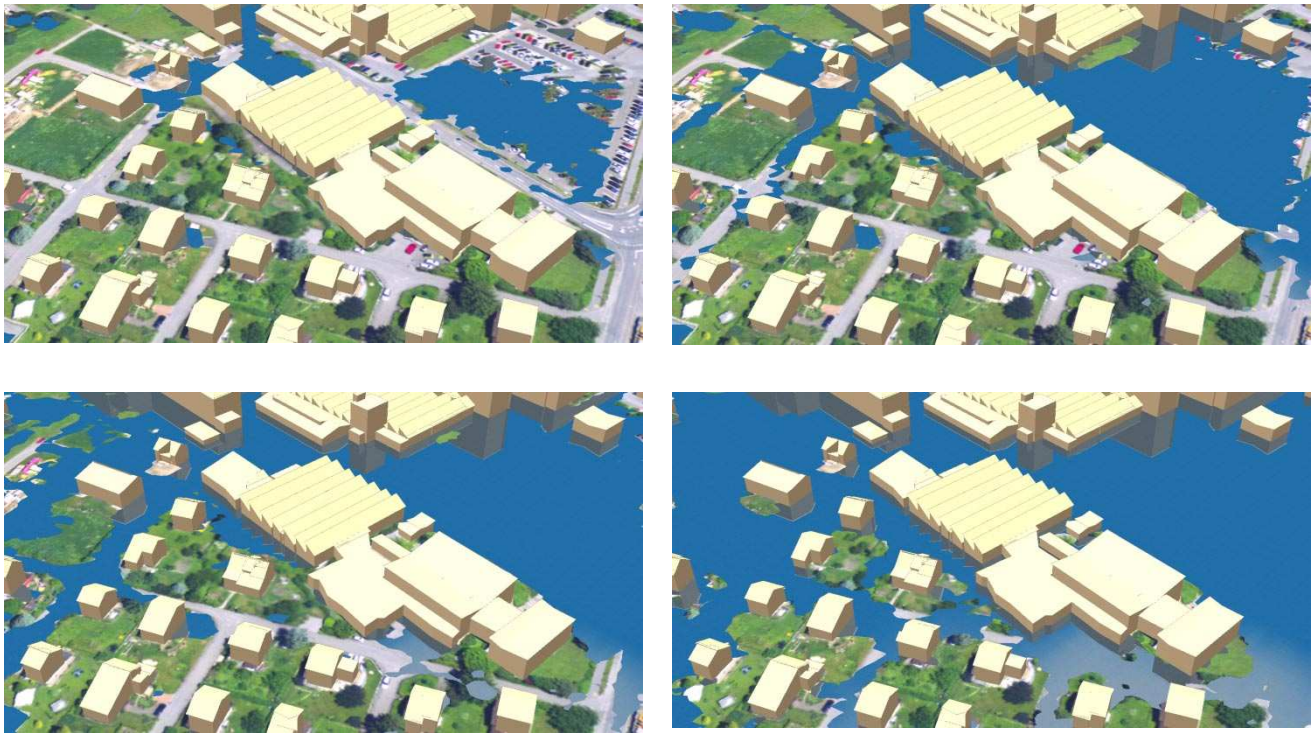


Figure 6 Visualization of flood analysis in Heerbrugg, Switzerland (test site 1). Semi-automatic building models from designed algorithms. Flooding levels are 2m, 5m, 7m and 10m for top left, top right, bottom left and bottom right pictures respectively.

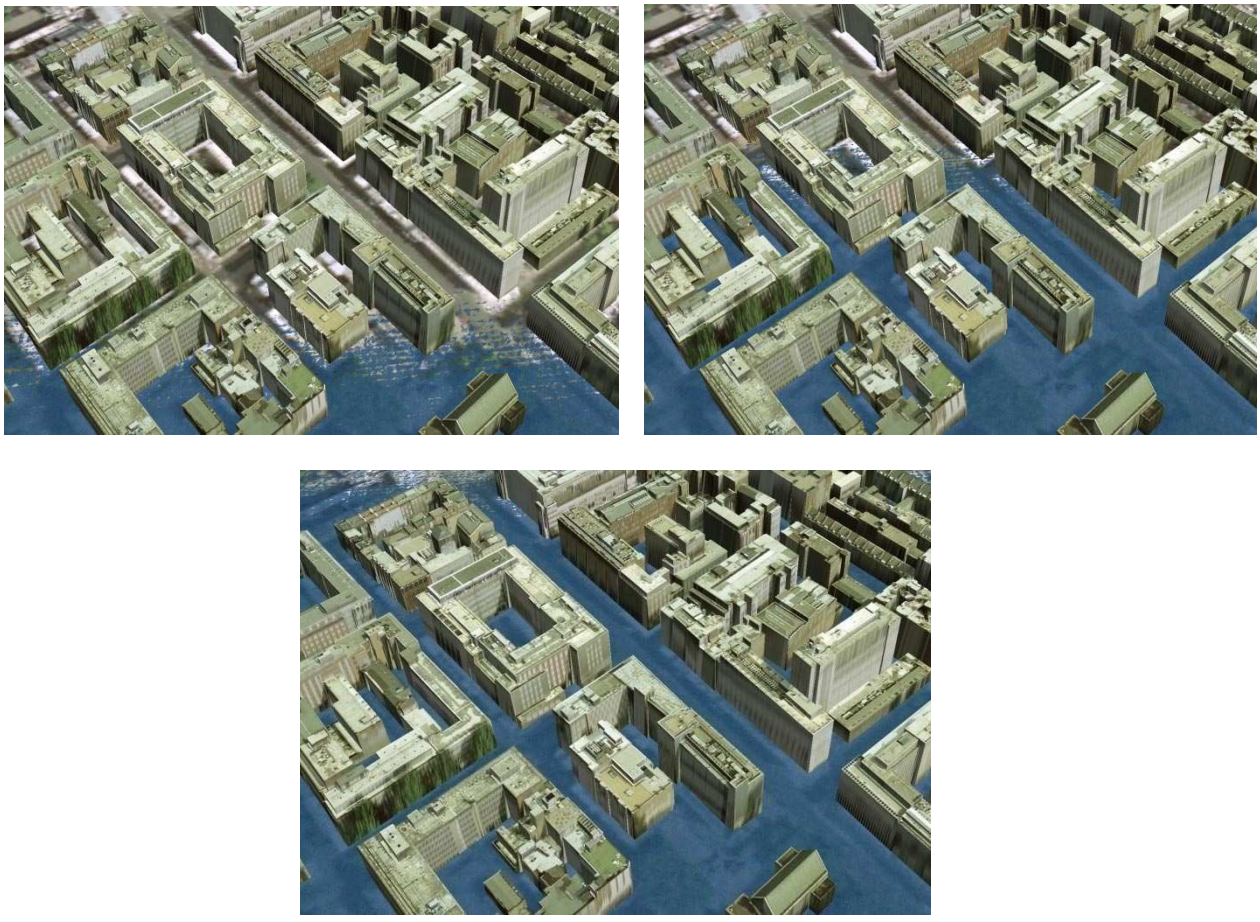


Figure 7 Visualization of flood analysis in London, UK (test site 2). Semi-automatic building models from designed algorithms. Flooding levels are 1.5m, 1.8m and 2m for top left, top right and bottom pictures respectively.

Geospatial data and especially highly detailed 3D city models can be very effective for fire risk analysis in urban areas. In a pre-fire analysis scenario, detailed 3D city models can be used to assess roof structures for emergency response operations, the number of floors at risk, as well as determine proximities with nearby buildings for risk assessment. In situations like these where the most appropriate decision is required, 3D city models of the surrounding area can be invaluable for efficient emergency response and planning in post-fire response.

2.1.4. Urban planning and Engineering

Geographical Information Systems (GIS) have, over the last three decades, revolutionised the way that spatial data is generated stored, analysed and disseminated. This information helps us to manage what we know, by making it easy to organise and store, access and retrieve, manipulate and synthesize and apply to the solution of problems (Longley et al., 2001). With the development of GIS technology and its widespread use and accessibility to home users, it is inevitable that its role in planning analysis and the property market will increase. Nevertheless the use of GIS as a 2D method for representing buildings and urban areas is fundamentally flawed since it is ignoring the complexity of the urban landscape, which is more efficiently defined by the third dimension.

city models offer city planners and engineers the opportunity to visualize complex urban environments to assist in the planning and decision making process. It's an efficient tool to visualize complex topographic structures and architectural details as well as assess future urban development projects.

city models can be an excellent tool for planners to assist them in the decision making process for projects before involving implementation costs. Common projects include:

- Traffic engineering. Allows engineers to perform pollution, noise and transport analysis
- Evaluate new infrastructure proposals and assess ecological concepts
- Fusing and visualizing urban development visions and simulating different planning options

- Shadow analysis
- City regeneration projects and visualization throughout the various stages of completion
- Decision making in competitions for architectural and engineering projects

Urban planning is certainly one of the most important fields of application for 3D city models. Local governments find themselves often confronted with complex decision processes about larger restructuring projects of old town areas and investment projects (e.g. a new shopping centre, commercial area, industrial site). Using a 3D city model in a state-of-the-art realtime visualization solution, all parties involved in the decision-making process can fly interactively through the virtual 3D city model. Since real time visualization systems support a number of CAD-formats, participating architects may place their drafts of planned projects in the 3D city model and present it to the public. When a new road construction is planned, an interactive 3D visualization may help to demonstrate the possible alternatives (variant planning). The impact on the natural landscape becomes most clear in the 3D virtual representation of the situation (Figure 8). Besides the terrain, image data and 3D city models, it is also possible to show various 3D objects like bridges.

Figure 8 Visualising landscape differences in an urban planning scenario. Models used from the London – Bloomsbury test site with and without trees.

3D visualization can be applied at different stages in the process of urban planning, and three different visual representations have been suggested by (Verbree et al., 1999). The visual representation contain three views: the Plan View (PV), an ordinary 2D colour map for initial orientation, the Model View (MV), a simple 3D map for professional volume analysis and the World View (WV), a detailed photorealistic map in 3D for public presentations (figure 9).

Figure 9 The three views of (Verbree et al., 1999) applicable at different stages in urban planning. a) PV, b) MV, and c) WV

The requirements of the 3D City models for urban planning applications vary considerably. With respect to geometry and level of detail there have been models ranging from LOD1 up to LOD3. One of the most typical requirements is the need for additional attribute information. In urban planning scenarios the building models are treated as separate entities.

Information can be tagged to individual objects and this can be coupled back to a related database. The individual objects can be made to be non-fixed if the user wishes and hence rapidly create moveable chunks of a city (Hudson-Smith and Evans, 2001). This way, different scenarios can be viewed, for example what a region would look like if one building were moved out and another one took its place (figure 10).

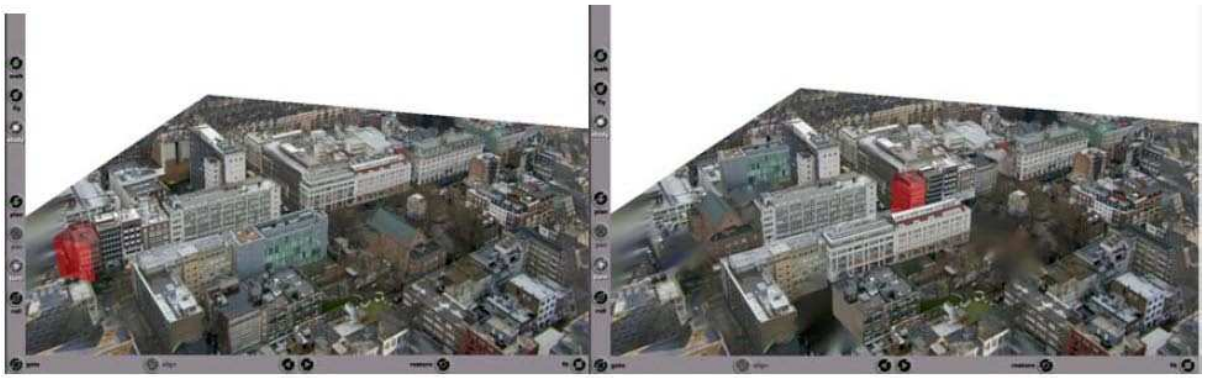


Figure 10 An interactive model of the Tottenham Court Road region in London. The building highlighted with red colour can be moved from one location to another (adapted from Hudson-Smith and Evans, 2001)

Another typical process is shadow analysis which in many countries is one of the prerequisites for planning permission considerations. Shadow analysis evaluates the relationships between the heights of the existing buildings and that of proposed buildings that could affect lighting in parts of a project area and could possibly affect the suitability, aesthetics, and functionality of some structures for their current uses.

Another potential application of 3D city models is the ventilation analysis. With the emergence of more and more skyscrapers, air circulation and ventilation problems occur. Negative pressure behind the windward side of a tall building is good for the internal ventilation, but for outside it could produce an uncomfortable micro climate in cold weather. The taller the building, the more complex the airflow is. How to compute the airflow for a building and control its change accurately is a big challenge for the architect and the urban designer. The use of 3D city models can assist in the construction of the analytical models to simulate the airflow activity.

2.1.5. Simulation applications

The use of 3D City models for training and simulation can be very effective. Extreme situations can be practised under varied weather conditions and at alternating daytimes without danger for the trainee and its environment. Training simulators usually consist of a replica of the original device, for instance the cockpit of an aircraft, the bridge of a ship or the interior of a car. To achieve a realistic environment, virtual scenery is projected on screens placed around the cabin simulator. The content of the scenery is defined within a visual database (figure 11).

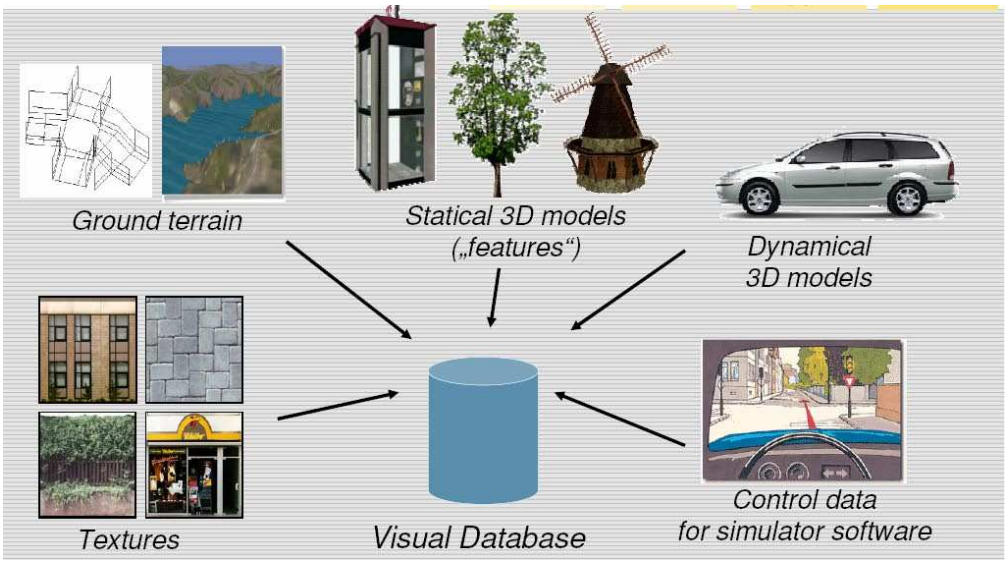


Figure 11 Components of a virtual database for training and simulation purposes (adapted from Bildstein, 2005)

The visual database contains all geometry and image data of the training area which the image generator needs for rendering the scenery. The components of these databases for a simulation environment typically contain ground terrain, static and dynamic 3D objects with animation effects and photo realistic 3D city models. The dynamic models in the virtual database are special 3D models which are equipped with movable parts, like wheels and rotors, or even switchable parts, such as traffic lights.

Flight simulations have become increasingly popular. Civil airlines use simulators to teach new pilots, to simulate emergency situations and to adapt experienced pilots to new or modified aircrafts. For flight simulations the requirements for 3D city models in state of the art flight simulation environments are LOD2 buildings for large scale areas and LOD3 models for airports and other landmarks of interest. Generic textures for the building facades are extensively used for fly over areas since they give the illusion of photorealistic representation when viewed from the flying altitude (figure 12).

Figure 12 Example of LOD2, 3D city models for flight simulations. (Adapted from PC video game Flight Simulator 2004 ™, Copyright Microsoft Corporation)

Airport buildings in flight simulations are used primarily for recognition and orientation purposes. The shape of the building should be correct in terms of collision detection when the trainee drifts from the specified route. Usage of 3D city models is also very beneficial for helicopter simulators since the trainee has the ability to hover and observe the surrounding environment and fly in lower altitude. Helicopter missions in urban environments comprise of

touch-down and take-off manoeuvres at different types of helicopter airfields. In this respect 3D city models enable helicopter training in real world urban locations within extensive and highly detailed scenarios (figure 13).

Figure 13 Use of LOD2 & 3 for helicopter flight simulations. (Adapted from PC video game Flight Simulator 2004 ™, Copyright Microsoft Corporation)

During the last few years, there has been growing interest in driving simulators because they have proved themselves to be a highly useful tool for training and instructing drivers, particularly professionals (bus, truck, emergency vehicle drivers, etc.). There are several driving simulations around the world in different research centres that are used for general driver training as well as for specific training needs like police patrol training or for education of public transport drivers.

In order to describe individual vehicle dynamics, a wide range of models has been put forward. These vehicle dynamics models range from simple models that use basic logic and consists of a few operations having high performance up to more complex car-following models that provide more realistic modelling of driver and vehicular behaviour (Bham and Benekohal, 2004).

Research related to the development and application of a driving simulator is not new. Driving simulators began to appear in primitive forms in the 1970s. With the advent of computer technologies, Daimler-Benz launched a high fidelity driving simulator in the 1980s (Drosdol and Panik, 1985), which created wide interest throughout the world. Since then, many automotive makers and research institutions worldwide have developed their own simulators. There are various levels of complexity of driving simulators ranging from a simple static base simulator to the most advanced simulators that are capable of simulating the dynamic motion and scenes of an actual vehicle.

A typical virtual environment for vehicle simulation is shown in figure 14. The typical vehicle simulation environment consists of the geometry objects, lights, sound sources, tasks, behaviours, collision detection, and actions in the systems (driving actions in this case). All the objects or nodes are structured in a hierarchical way, in a tree structure. The base node is the so-called ground. The ground node includes a three-dimensional mesh representing the complete terrain geometry. The second types of nodes are the vehicles. These are movable nodes, changing position in each simulation step.

Driving simulators databases contain an accurate representation of the roads, the sidewalks, traffic lights and signals. The 3D city models typically used for these applications should have a high level of detail for the building facades so the use of LOD2 or LOD3 building models is advisable. One of the main requirements of the driving simulations is the high refresh rate (60hz) as proposed by Bildstein (2005). In order to achieve this, the 3D city models should be very efficiently created for real time rendering.

Figure 14 An example of a driving simulator using LOD2 city models (Adapted from Bildstein, 2005)

2.1.6. Property management and cultural heritage

Property management and the real estate industry was one of the early adopters of online platforms such as Google Earth that helped the industry manage and demonstrate future buildings to potential clients. For the real estate industry the two critical aspects of any 3D visualization environment is the use of thematic information and the ability to perform basic spatial analysis within a GIS environment.

The generic idea of GIS is to incorporate geometric and semantic information in one system and to support analysis in both domains. A 3D GIS should satisfy various spatial operations. These operations can be summarized as follows: (Zlatonova, 1999).

- Access to semantic properties of one type object,
- Access to both semantic and location information,

- Operations which create object-pairs. (e.g. buildings in a given parcel that have one owner.)
- Operations to analyse semantics of object pairs from one or more types.
- Operations which create a new type of object from existing objects.
- Retrieval operations. (e.g. what is the current information about a particular building.)
- Query operations retrieve data which satisfies some given conditions.
- Retrieval and query of semantic data.
- Integrated analysis of spatial and semantic data, (classification, measurement, overlay operations.)

For the above operations, firstly the geometric and the thematic characteristics of objects and their spatial relationships should be integrated in a database. Topology between the building entities in a 3D City model as well as attributes are stored in relational databases (figure 15).

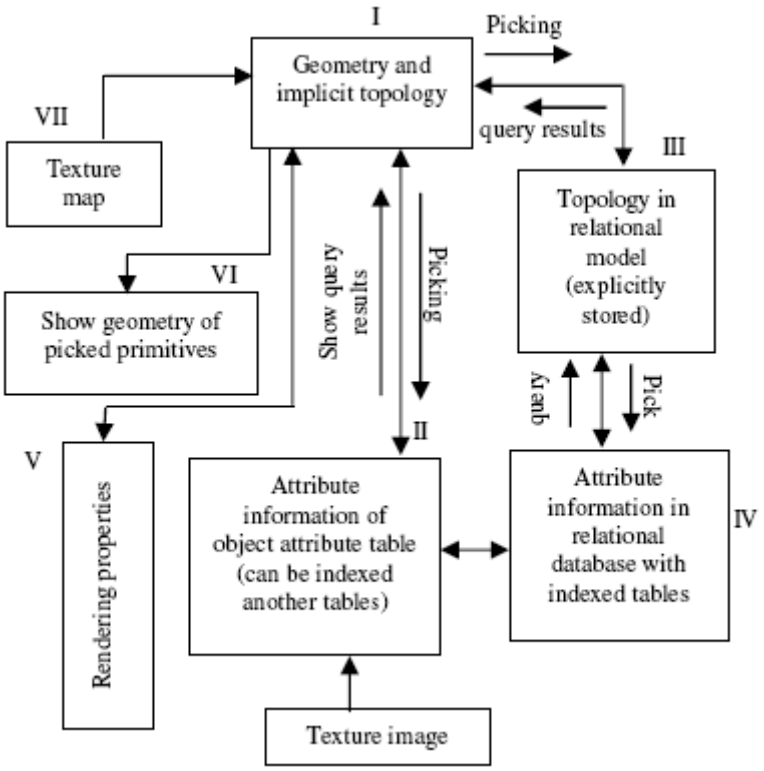


Figure 15 General representation of geometry and topology for a 3D city model. Adapted from (Zlatonova, 1999).

With the representation shown in Figure 15, it is possible to manage the visualization tasks and database tasks in a flexible manner (independent and depending to each other). With the above representation even if the relational tables are not created, the operations which are shown in the boxes with the numbers V, VI and VII are ready and effectively can be performed in the 3D visualization pipeline. This means, one can query the geometry of the 3D building's geometric features such as point coordinates, ID of any cell (face, edge, surface etc.) by clicking on it with the mouse in the rendering window (figure 16).

Figure 16 3D building models in Manchester area linked to a relational database in VirtualGIS™.

Figure 16 is an example from the addition of a relational database linked to 3D building models in Manchester. The relational database stores attribute information containing the number of floors, address, owner, house price and other valuable information for each building entity. The database query and data manipulation is handled by the database management system implemented in VirtualGIS™. Highlighting buildings that conforms to

user generated SQL queries provide a valuable tool for the property management and real estate industry. For property management applications the accuracy requirements of the 3D city models range from LOD2 to LOD3 models.

Highly detailed 3D City models have been also used to visualize cultural heritage projects (figure 17). Cultural monuments often have an inestimable value. But often this is only recognized once the cultural monuments are endangered or already destroyed.

Figure 17 3D city model of the ancient city of Augusta Raurica- Switzerland with modern orthoimages and height model (adapted from Salathé, 2001)

The basis for the model shown in figure 17 was an existing wooden 3D model at a scale of 1:500 which had been established around 1960. The wooden model was based on an abundance of archaeological evidence and had thus gained a certain level of acceptance (Salathé, 2001).

In archaeological and cultural heritage projects object semantics are typically just as important as the actual geometry. Thus, it is a key requirement to assign thematic information to entire objects and to individual geometric elements. This also makes it possible to select, analyse or edit the geometry and the appearance of objects based on semantic criteria.

Cultural heritage 3D models present the most challenging task from an automation perspective. This is because cultural heritage objects often have a very irregular complex geometry. Thus, a good digital reconstruction requires a very detailed 3D model with a lot of geometry elements. So there are two main requirements to a 3D GIS. The first one is a support for the acquisition and handling of large amounts of complex and non-planar 3D geometry. The second one is the visualisation of these objects which consists of a lot of geometry elements. Typical 3D models suitable for these applications include LOD2, 3 and LOD4 models which have interior spaces also modelled.

2.1.7. Navigation with portable navigation devices

The last few years the car navigation and location based services industries have experienced significant growth. Competition in the car navigation industry forces companies to innovate and add unique features in their car navigation devices to improve the experience for the driver and passengers. One of the main features that most companies have focused on the introduction of 3D city models for car navigation systems. This presents a number of technical and safety issues that will need to be overcome before such systems are fully implemented in the mass market.

Mobile applications of virtual 3D city models represent a major and complex research challenge due to limited bandwidth and graphics capabilities, restricted interaction capabilities, data standardizations and distribution techniques as well as and digital rights

issues. One of the biggest issues in implementing 3D city models in mobile units is the 3D rendering engine.

In 3D computer graphics, numerous rendering techniques are available to cope with complex virtual environments, including discrete and continuous multi-resolution geometry and texture representations, view-frustum culling, occlusion culling, imposter techniques, and scene graph optimizations (Akenine-Möller and Haines, 2002). Virtual 3D city model visualizations require an efficient management of large-scale texture data, for example, for aerial photography and building facades (Buchholz and Döllner, 2005), and level-of-detail management for large heterogeneous 3D object collections (Davis et al., 1999) and 3D terrain surfaces (Döllner et al., 2000). Although these rendering techniques enable real-time rendering of complex 3D scenes, they generally cannot be transferred directly on mobile devices due to limited computational resources and power.

One principal approach to efficient mobile 3D rendering was proposed by Royan et al. (2003) that describe a client-server architecture for mobile 3D virtual city visualizations based on a progressive and hierarchical representation, whereby the progressive and hierarchical representation is based on a tree data structure that holds the merging and simplifications of buildings. Another solution proposed by Cheng et al. (2004), makes use of a progressive, compressed transmission of image sequences. They investigate a client-server approach for visualizing complex 3D models on thin clients (PDAs and mobile phones) by applying real-time MPEG-4 streaming to compress, transmit, and visualize rendered image sequences. To accelerate the MPEG-4 encoding process, Cheng et al. (2004) developed an online algorithm to calculate the block motion vectors using 3D information without having to employ an expensive search. Moreover, this computation is performed on the Graphics Processing Unit (GPU), which can be performed in parallel with the video encoding. The MPEG block was used as a matching technique to further search for the blocks whose motion vectors cannot be directly determined (Cheng et al., 2004)

Other initiatives for the use of 3D city models in mobile devices include sketch based navigation techniques. Jurgen et al. (2005) presented a solution for accessing virtual 3D city

models on mobile devices with the user controlling the navigation within the virtual 3D city model by navigation command sketches drawn directly on the rendering window of the mobile client (Figure 18). The sketches are sent to the server, which reprojects the sketches onto the 3D scene correlating the sketches to scene objects, interprets these sketches in terms of navigation commands, and sends the resulting video-encoded image stream to the mobile client. That is, the mobile client enables users to specify and retrieve step-by-step created video sequences that correspond to their navigation intentions (Jurgen et al., 2005)

Figure 18 A sketch based navigation system using LOD2 3D city models. Look around the area (left), sketching a navigation command to walk along the street (right). (Adapted from Jurgen et al., 2005)

Due to the limited system resources and low-end 3D graphics accelerators 3D city models for portable navigation devices are restricted to LOD2 models without any roof details and in most cases using generic texture libraries.

2.1.8. Security and defence

3D city models have played an important role for military training and homeland security ranging from battle simulations to visibility analysis in police enforcement scenarios. Military agencies around the world have a long history of investing in modelling and simulation to support objectives such as training and analysis programs and have indirectly supported the development of many of the fundamental computer graphics and networking technologies that underlie both military and entertainment applications of modelling and simulation.

Figure 19 Use of LOD2 city models for military flight simulations (adapted from CAE incorporation 2008)

Military agencies use modelling and simulation for a variety of purposes, such as to train individual soldiers, conduct joint training operations, develop doctrine and tactics, formulate operational plans, assess war-fighting situations, evaluate new or upgraded systems, and analyze alternative force structures. As a result of this, defence models and simulations range in size and scope from components of large weapons systems through system-level and engagement-level simulations, to simulations of missions and battles, and theatre-level campaigns.

For military simulations there are four main critical points according to Smith (2006) which are tools for creating the 3D city models, the interface, networking technologies and computer generated forces and autonomous agents.

Automated tools for creating 3D city models are very important for the military industry. Computer-based tools are needed to efficiently generate LOD2 and LOD3 building models with objects that can interact with participants, like doors that can open or destructible buildings. The interfaces provide the portal through which participants interact with a system. They include displays, entry devices such as keyboards or touch-sensitive screens, VR systems, and a host of other input/output devices that link the participant to the simulator. The increase in the richness of the participant's ability to interact with the synthetic environment and other people and agents similarly integrated there, is especially important as large-scale simulations are constructed. Networking technologies enable large numbers of participants to join in a simulation regardless of their physical locations. The network must be able to accommodate the volume of messages between and among participants in a timely fashion with a minimum amount of delay or latency. Computer-generated forces and autonomous agents control the actions of elements not directly under the control of a human participant in a simulation. They can be adversaries or companions and can represent individual players or aggregated forces. Computer-generated forces are critical in any simulation intended to be used by an individual participant or in large networked simulations in which it may not always be possible to ensure enough players to control all the necessary entities.

In addition, the video game industry has demonstrated the potential use of 3D video games for military simulation. The military has been one of the first and most avid adopters of game technologies. These games originated from military roots in the 1990’s and contain many similarities with the training devices that are used to train soldiers (Smith, 2007).

The gaming industry demonstrates the capability to create highly detailed LOD4, 3D city models that have created new areas of co-operation for military and homeland security agencies. The emergence and application of these technologies has been visible for a number of years. The National Research Council Committee on Modelling and Simulation described the convergence of entertainment and defence simulation in their 1997 report (National Research Council, 1997).

Figure 20 Use of highly detailed LOD4 building models in modern warfare PC video games (adapted from Call of Duty 4 – Modern Warfare™, Copyrights reserved by Activision Incorporated)

For most military applications realism and interaction for the 3D city models is valued more highly compared to the accuracy of the models, hence the close synergy with the video games industry. Interaction and the introduction of physics engines (rendering algorithms and hardware that take into the physical properties of materials and objects in order to create particles from destructible objects and the effects of various interactions with the user) for realistic destructible environments are also very important. Typically 3D city models for

land simulations consist of LOD3 and LOD4 building models with interactive features such as doors, windows, etc.

2.2. Real time visualization of detail 3D city models

As highlighted above several applications would benefit from large scale detailed 3D city models deployed either through the internet or locally in the users platform. These applications in most cases require powerful rendering engines to facilitate the interactive visualization of not only the 3D City models but other geospatial data critical for the photorealism of the product including high resolution Digital Terrain Models, high resolution true-orthophotos and other vector data.

During the past few years a number of different technologies have been introduced to handle these datasets such as adaptive triangulation and paging of digital terrain models. LOD management of texture data and volumetric rendered clouds have helped generate more realistic 3D scenes (Beck, 2003).

Even though over the past 5 years 3D graphics systems have achieved stunning performance gains there are still two remaining constraints in real time rendering of large 3D city models that is; the database and texture size constraints. In order to cope with these constraints efficient system architectures are needed.

One of the first requirements is a real time rendering engine that should maintain a 30Hz frame rate at a reasonable flying speed while paging the database from disk in real time. LOD switching is also an effective practice during which geometric objects are displayed with the appropriate LOD depending on the distance from the viewer. LOD switching of discrete objects, for example buildings and vehicles, is relatively simple since the entire object can be replaced at once. LOD-switching of terrain data is more problematic though because a DTM needs to have a sustained continuation for efficient visualization. This effect is normally achieved by dividing the DTM in smaller tiles or rectangular blocks. Each tile can have

multiple LOD layers and can switch between these independently. Tiling schemes are categorized as the ‘replacement LODs’ where all LOD representations for a tile are the same size and the ’Quadtree LODs’ are where the current LOD tile is quartered at the next LOD level into four smaller tiles. In order for the grid of tiles to appear as a continuous surface, each tile must match its neighbour along their shared edge. This is typically accomplished by forcing the geometry of the shared edge to remain constant throughout all LOD representations of a tile (Beck, 2003).

Objects such as trees, buildings walls and roof textures can also be represented with different LOD depending on the distance from the viewer. For this purpose image pyramids are employed for the building textures. An efficient data hybrid data structure has been proposed by Gruen and Wang (1999) that combines the geometry of 3D models, images and attribute information. The image data can be attached to the surface object, body object and DTMs while the thematic attributes are built in a separate data table. According to Gruen and Wang (1999) the 3D models are grouped into four different geometric object types: Point objects, Line objects, Surface objects and Body objects. The geometric data set contains information regarding the position, shape, size, structure definition and image index as illustrated in figure 21.

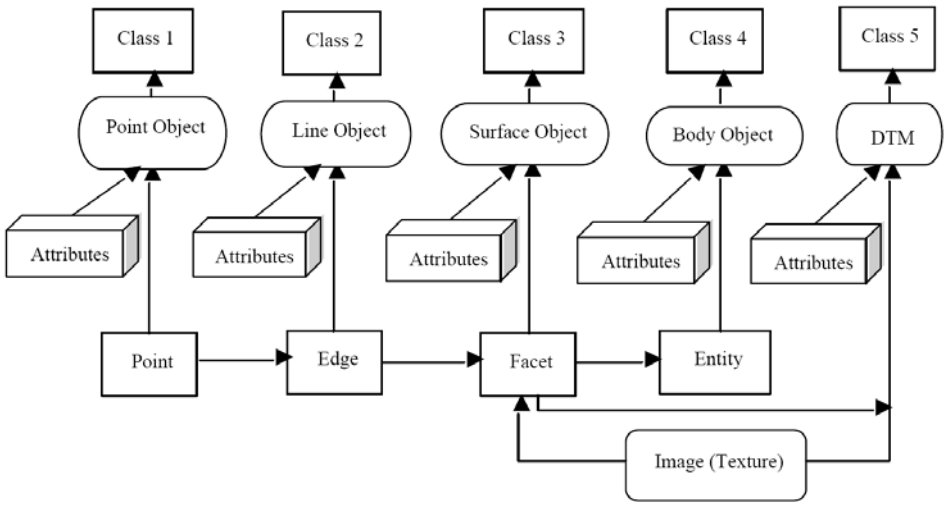


Figure 21 The logical data structure of a hybrid visualization system proposed by Gruen and Wang (1999)

In the above data structure the topological relationships between geometric elements are implicitly defined by the data structure.

Realistic environmental effects and cloud dispersion (figure 22) is another important component that improves the immersion in a virtual scene. Shading effects are also incorporated in many applications to improve photorealism. For a precise shading model two components should be considered. The first one is a multiple scattering due to particles in clouds and the other factor to be considered is sky light (Dobashi et al., 2000).

Figure 22 Volumetric clouds using particles physics OpenGL engine (Adapted from TerrainView™, Viewtec)

2.3. Towards a unified representation of 3D city models

During the last few years there has been an effort to propose a unified representation of 3D city models. Qing et al. (2005) proposed an interesting conceptual model which is classified based on Feature entities and Fields. According to the proposed model the Fields is an

association between a geometric description of entities and a set of attributes in OpenGIS specification. Compared with natural objects, the urban man-made objects are relatively regular in the terms of shape, distribution, and structure. This makes it possible for data model to abstract and represent urban spatial objects with some primary geometric elements. As shown in Figure 23, the point object, curve object, surface object, volume object, group object, LOD object, DEM and Texture object are then employed as the eight basic objects to describe urban spatial entities.

In addition, Qing et al. (2005) propose a unified model for the spatial object property characteristics based on geometry properties, material properties, textural properties, resolution, thematic information, model recognition, integrity and consistency.

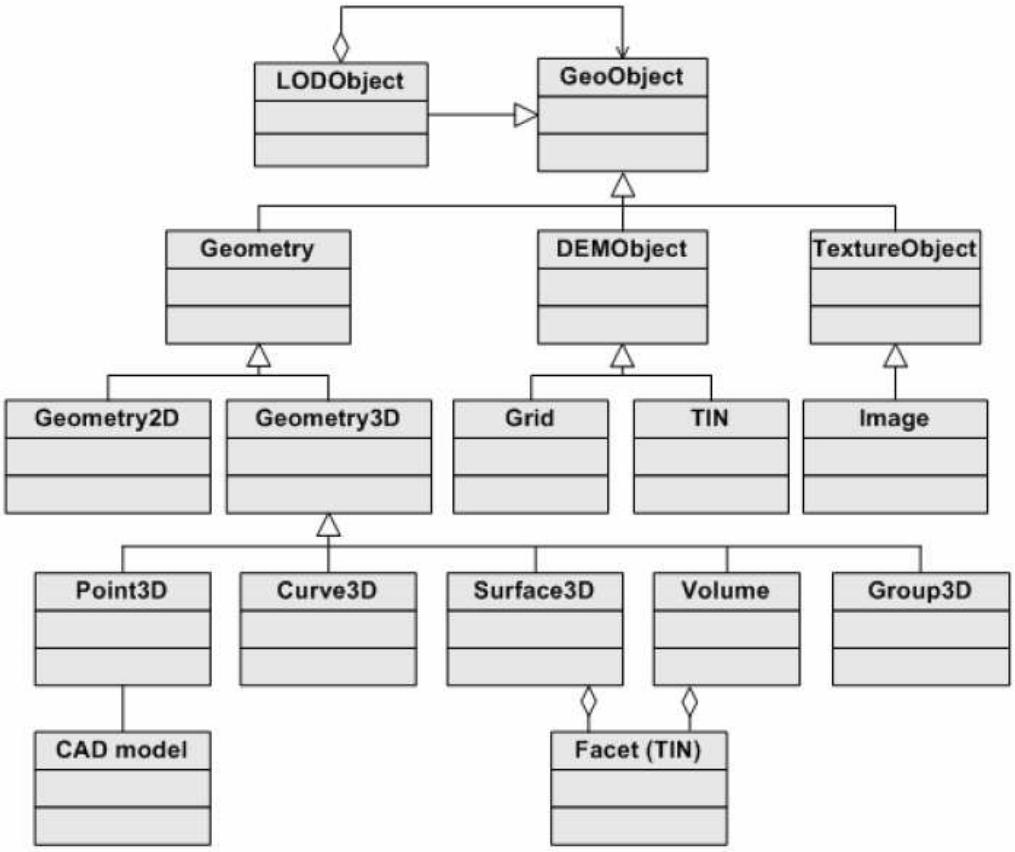


Figure 23 A conceptual description for unified 3D city models (adapted from Qing et al., 2005)

Unified protocols for exchanging 3D city models are becoming increasingly important. These should be designed for efficient manipulation, minimising file sizes for real time rendering. Currently triangular meshes have become very popular for exchanging and viewing 3D data

sets. This trend is reinforced by the wide spread use of 3D graphic libraries (openGL, VRML) and other 3D data exchange file formats and hardware 3D engines optimised for rendering triangles. A triangle mesh is usually stored as a list of coordinates, called vertex list and a list of triangles that reference to these coordinates, called connectivity (figure 24).

A complex three-dimensional model with n vertices has about 2n triangles. 18 bytes per triangle are needed to store a vertex list and geometry plus additional cost for texture mapping and colour information.

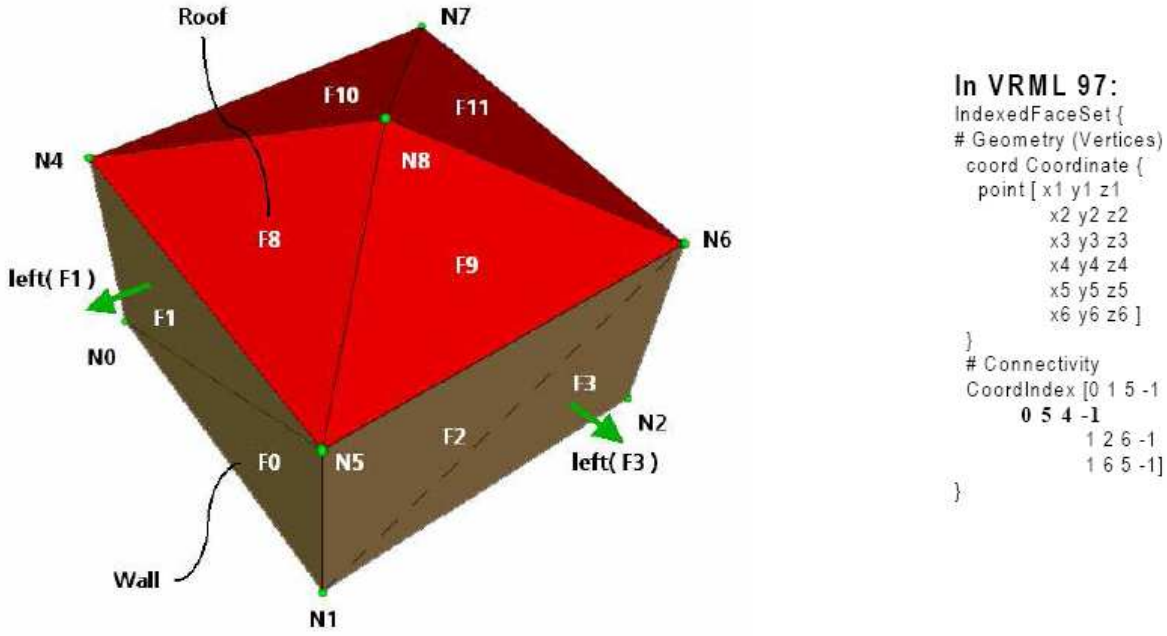


Figure 24 VRML representation of a triangle mesh (Rossignac et al., 2001)

With increasing file sizes of 3D City models the importance of efficient compression techniques has become apparent for storing and handling the information. Although generic file compression algorithms can compress up to 60-70% of the file size, algorithms specialised on 3D meshes achieve a much better compression rate of about 95%.

CityGML is a common information model for representing 3D urban objects. It defines classes and relations for the most relevant topographic objects in cities and regional models with respect to their geometric, topological, semantic and appearance properties. "City" is broadly defined to include not just built structures, but also elevation, vegetation, water

bodies, “sidewalk furniture” and more. Included are generalization hierarchies between thematic classes, aggregations, relationships between objects and spatial properties. These thematic information types go beyond graphic exchange formats and allow users to employ virtual 3D city models for sophisticated analysis tasks in different application domains. CityGML is an open data model and XML-based format for storing and exchanging virtual 3D city models. It is implemented as an application schema of GML3, the extensible international standard for spatial data exchange developed within the Open Geospatial Consortium (OGC) and ISO TC211. GML3, used with other OGC standards - mainly the OpenGIS Web Feature Service (WFS) Specification - provides a framework for exchange of simple and complex 3D models. However, WFS and GML3 only establish syntactic interoperability. A GML3 document needs to be structured by the definition of an application schema that is tailored to a specific application domain. In this case, the application domain is 3D city modeling and the GML3 application schema (or profile) is CityGML.

CityGML takes advantage of other open standards and its development has proceeded in careful cooperation with other groups. For example, graphic rendering of data encoded in CityGML can be accomplished using standardized computer graphics data formats like VRML, GeoVRML, X3D or Universal 3D (U3D). To ensure European acceptance, the developers of CityGML have coordinated with EuroSDR, a spatial data research organization consisting of delegates from geographic information production organizations and research centers from 18 member states in Europe, together with participants from industry and the commercial sector.

2.4. Building reconstruction from single data sources

Building reconstruction is of primary importance in many applications as highlighted in the previous section. Unfortunately, manual reconstruction of 3D building models is very time consuming and therefore several research studies have been presented that deal with semi-automatic building reconstruction, either from aerial images or from LiDAR datasets. Nevertheless the practical implementation for operational use remains limited. The main reason is the lack of reliability and the need of manual intervention which then only introduces a marginal gain in productivity. For instance, certain parts of the buildings will be hidden in optical data due to the scene perspective or occlusions and the feature extraction may not yield all the necessary roof planes for a reliable reconstruction. In contrast building models derived from LiDAR data lack the high planimetric accuracy of the building footprint and the ability to reconstruct small roof details. A summary of different methods for building reconstruction is now introduced.

2.4.1. Building reconstruction from airborne optical data

Automatic building reconstruction from aerial images is a difficult problem due to occlusions, introduced mainly from vegetation, scene perspective and the complexity of the roof structures. Early approaches tried to use a single image only (Huertas, 1988; Lin, 1998). Since the inference of 3D information from one image is very difficult and there are still some ambiguities in the detected buildings that can be only resolved by feature matching in multiple images, the application of the single-image approach is very limited. In this context, most of the recent work in this area has focused on the multiple-view analysis. In the context of building reconstruction from multiple aerial images the proposed methods can be categorized as data-based or model-based. In the data-based methods (Vozikis and Jansa, 2008; Heuel et al., 2000; Ameri and Fritsch, 2000; Scholze et al., 2002) the building reconstruction is performed without any assumption for the structure of the roof and

therefore no restrictions are introduced in the method. On the contrary the model-based methods (Katartzis and Sahli, 2008; Dong-Min et al., 2008; Willuhn and Van Gool, 2005) use some models of buildings to restrict the set of possible shapes. This external knowledge enables the user to overcome the lack of detection due to occlusions.

In the data-based methods it is very common to extract linear primitives from the imagery to assist in the reconstruction process. In Vozikis and Jansa (2008) buildings are extracted from airborne and spaceborne imagery by applying Hough Transformation procedures. The workflow in Vozikis and Jansa (2008) consists of 4 steps, namely the pre-processing of the data, nDSM (normalized Digital Surface Model) creation, building localization and building extraction. They concluded that the Hough Transfrmation is a very powerful tool for increasing the degree of automation during building extraction, while it is very robust against noisy data. Additionally the level of detail of the extracted buildings can easily be adjusted, but dark shadows in the images can make the algorithm produce erroneous results. In Scholze et al. (2002) 3D segments are extracted, edges are extracted in Heuel and al. (2000) and in Ameri and Fritsch (2000) planar patches are used to solve the lack of generality that is inherent in these strategies, by introducing heuristic rules (Fischer et al., 1998). Despite the promising results this approach is still limited to simple forms and thus cannot handle all the shapes available in urban or suburban areas as a function of the limited number of models. Increasing the library of models would result in an increased complexity and a less robust method (Taillandier, F., Deriche, R., 2004).

Most methods for building reconstruction divide the task using a two-phased approach whereby a detection phase is initially applied extracting the locations of single buildings and then a reconstruction phase is implemented. In Katartzis and Sahli (2008) a building reconstruction method from aerial images is employed using a stochastic image interpretation model, which combines both 2-D and 3-D contextual information of the imaged scene. Building rooftop hypotheses are extracted using a contour-based grouping hierarchy that emanates from the principles of perceptual organization. Katartzis and Sahli (2008) use

a Markov random field model to describe the dependencies between all available hypotheses with regard to a globally consistent interpretation. The hypothesis verification step is treated as a stochastic optimization process that operates on the whole grouping hierarchy to find the globally optimal configuration for the locally interacting grouping hypotheses, providing also an estimate of the height of each extracted rooftop.

In Dong-Min et al. (2008) a method for building reconstruction is presented using the concept of disparity maps. In this approach the epipolar images are generated from the aerial images by epipolar resampling process. The disparity map is obtained between the epipolar pairs by stereo matching using area-based matching with non-parametric technique. From the disparity map, a DEM is generated as a 3D terrain model. The building location information extracted from the disparity map is used to remove the unnecessary line segments extracted in the low level process. After 2D lines are generated, perceptual grouping is applied to the filtered line segments in order to obtain the structural relationship features such as parallel line segment pairs and U-shapes. These are used to generate rooftop hypotheses. Among the generated hypothesis, the candidate rooftop is selected by searching close cycles in the undirected graph. An undirected graph is a graph in which the nodes are connected by undirected arcs. An undirected arc is an edge that has no arrow; both ends of an undirected arc are equivalent (Thomas et al., 2001). By using undirected graph, hypothesis selection becomes a simple graph search for close cycles. This significantly improves the performance of the system over the traditional hypothesis selection methods. Finally, the 3D buildings are retrieved by using 3D triangulation for each line segment of detected rooftops (Dong-Min et al., 2008).

Apart from the above fully automated methods, semi- automated building detection approaches have been introduced with reliable results as in the case of Willuhn and Van Gool, (2005). This model based method combines a semi-automatic building detection, requiring minimal manual interaction, with a fully-automatic building reconstruction. The efficient semi-automatic detection is based on the concept of a building row. A building row as defined in Willuhn and Van Gool, (2005) is a number of buildings that are aligned along a

street. The detection starts with the user identifying a building row and the system calculating the exact orientation, approximate width and height range of the building row, as well as the start and end parameter of each building. Furthermore a pixel-based segmentation is used. The gray scale values of the pixels inside the strips on both sides along this line are analyzed by using a similar approach to region growing, starting with segments of high homogeneity, and extending them in both directions to account for the transition at the border of the segments. Then the homogeneous segments are determined based on homogeneity criteria and then the segments are merged together. Subsequently the merged segments are treated as seed points for a region growing algorithm that expands the segmented regions. Additional cues are introduced, such as 2D contours, to assist in the segmentation hypothesis. In the final step of the building detection the building hypothesis is evaluated and individual detected buildings belonging in the same roof structure are merged together.

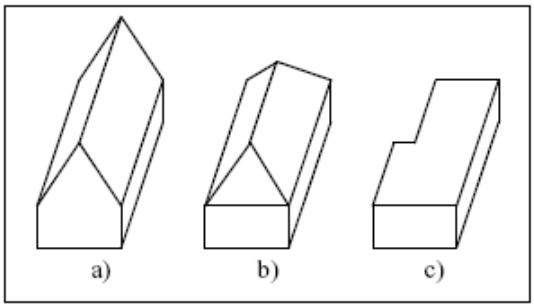


Figure 26 Models for building roof structures. a) gable, b) hip and c) flat (Willuhn and Van Gool, 2005).

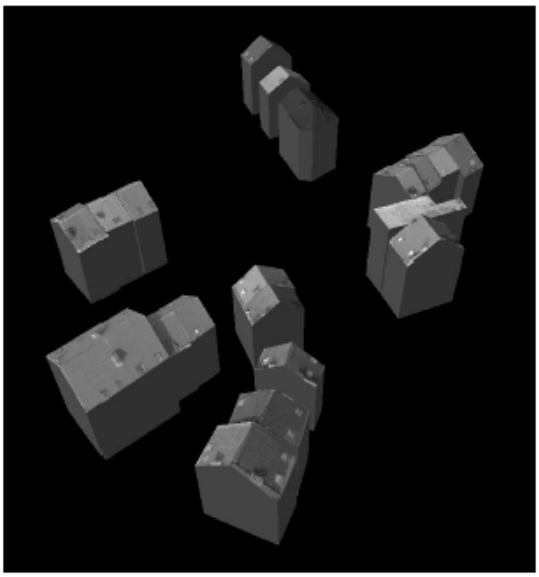


Figure 25 Results from the building reconstruction process (Willuhn and Van Gool, 2005).

The building reconstruction in the model-based methods uses some kind of conceptual models. The conceptual model takes into account the various roof forms, such as flat, gabled, or hip form as illustrated in figure 25. This model, however, requires a high quality of data in order to ensure that the reconstructed object is indeed the roof that is captured in the image (Willuhn and Van Gool, 2005). For the building reconstruction process the building hypotheses are used together with straight line segments and planar surfaces in 3D. The automatic roof reconstruction produced high quality results for 20 out of the 25 buildings in the scenes. A sample of

the reconstructed buildings is illustrated in figure 26. Nevertheless problems still exist in shadowed areas and in occluded regions, that resulted in insufficient quality of the extracted 3D line segments and planar surfaces. This method demonstrates the advantages of separating the task into two phases. The first phase illustrates the efficiency of the segmentation to produce regions containing single buildings and the quality of an automatic reconstruction that can be achieved by focusing on individual buildings. The combination of a region-based and a contour-based segmentation in 2D and fusion of the results with 3D data, extracted from the aerial images, has been shown effective in keeping the user interaction at a minimum (Willuhn and Van Gool, 2005). In contrast Taillandier and Deriche, (2004) have presented a data-based approach which uses a novel system for automatic building reconstruction from multiple aerial images. This approach uses a generic concept that treats the buildings as polyhedral models.

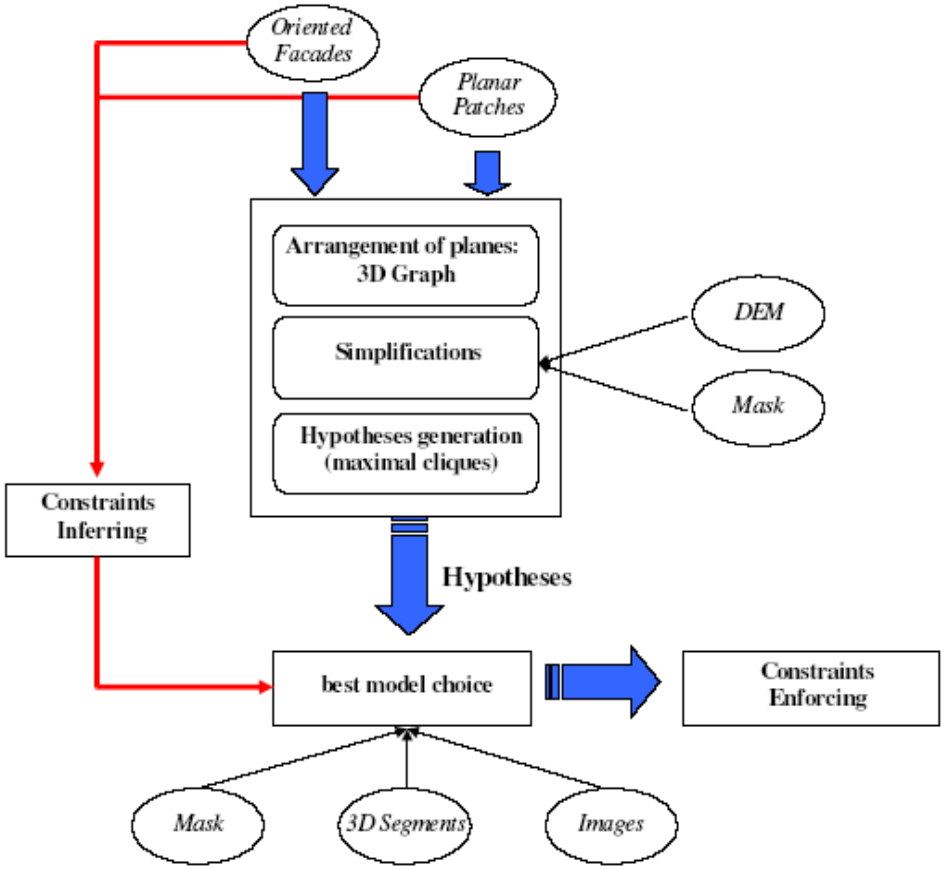


Figure 27 Diagram indicating an overview of the method by Taillandier and Deriche, (2004).

The method can be subdivided into three main steps that include the primitive detection, hypotheses extraction and choice of the best representation, together with geometric refinement (figure 27). In the primitive detection step, planar patches are extracted from a produced Digital Elevation Model (DEM) and subsequently are parametrically described based on their orientation (Taillandier and Deriche, 2004).

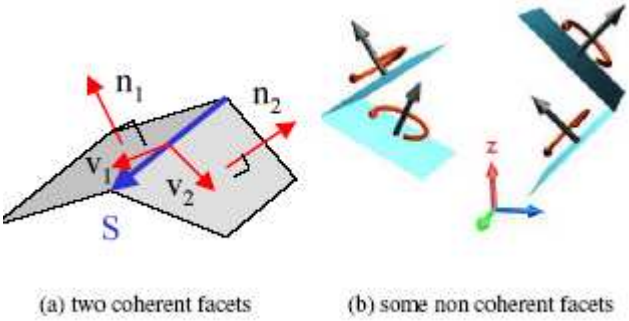


Figure 28 Coherency of the adjacent facets (Taillandier and Deriche, 2004).

From the initial hypotheses, a graph of hypothesis is generated that estimates the coherency of the adjacent facets. The coherency of the adjacent facets is determined as a function of the triple scalar product, as indicated in figure 28. In general two facets are coherent if the normal direction remains consistent when switching from one facet to the other one.

Among these hypotheses the choice is done through a Bayesian formulation (Taillandier and Deriche, 2004).

Furthermore, additional constraints are introduced, such as parallelism, edges or orthogonality constraints that are derived from any prior architectural knowledge of the scene. These constraints can enhance the overall procedure of selecting the appropriate building model, especially when ambiguities remain in the choice of the model.

Figure 29 Reconstructed buildings projected on an orthophoto based on the method proposed by (Taillandier and Deriche, 2004).

This aproach compared to model based methods, uses a very generic modelling of buildings without assuming any prior information or restraining the model in predefined shapes. The use of additional constraints is also an important element, but complex roof structures still require a more brute force search method for reliable results (Taillandier and Deriche, 2004).

2.4.2. Building reconstruction from LiDAR point clouds

The main objective of the building reconstruction solely using point clouds is to extract surfaces from the dataset. In general these methods can be divided into two categories. The first includes the methods that directly derive the surface parameters in a parameter space by clustering the point cloud, which can be a very effective and robust approach when planes or other simple shapes are extracted (Karsli and Kahya, 2008; Rottensteiner, 2003). The second category includes methods that segment a point cloud based on criteria like proximity of points or similarity of locally estimated surfaces (Vosselman et al., 2004).

Several approaches have been presented for building extraction from LiDAR data. Maas and Vosselman (1999) extracted parameters of standard gable roof type using invariant moment analysis. The method was based on intersection of planes fitted into a TIN model and had the ability to determine even more complex buildings. Merging of TIN meshes was used by Gorte, (2002) in order to compose the surfaces of the polyhedral building models. In this method the initial planar surfaces are created by the TIN mesh and then adjacent planes patches are merged if their plane equations are similar. The merging process is based on a similarity measure that is computed for each pair of neighbouring surfaces. The merging process continues until there are no more similar adjacent surfaces. Additional cues were used by Wang, (1998) that implemented a Laplacian of Gaussian edge detector to extract edges from a DSM produced from LiDAR data. Moment analysis was used to describe edge properties while shape and morphological parameters were used to classify the buildings edges from other features.

One of the most frequent methods for plane extraction, used for polyhedral modelling, is the 3D Hough transform. The 3D Hough transform is an extension of the (2D) Hough transform used for the extraction of line segments in imagery. The 3D Hough transform uses the following formula that describes every non-vertical plane.

$$Z = s_x X + s_y Y + d \tag{2.1}$$

From the above equation the (s_x) and (s_y) represent the slope of the plane along the X and Y axis respectively and *d* is the height of the plane at the origin. These three plane parameters define the parameter space. Every point (*sx*, *sy*, *d*) in this parameter space corresponds to a plane in the object space. The detection of planar surfaces in a point cloud can be performed by mapping all these object points to planes in the parameter space (Vosselman et al., 2004).

In Karsli and Kahya (2008) a sequential Hough transform process is introduced for detection of buildings in point clouds. The algorithm consists of two sequential steps of low dimensional Hough transforms. The first step, called Orientation Estimation, uses the Gaussian sphere of the input data and performs a 2D Hough Transform for finding strong hypotheses for the direction of building axis. The second step of Position and Radius Estimation consists of a 3D Hough transform for estimating building position and radius. This sequential breakdown reduces the space and time complexity while retaining the advantages of robustness against outliers and multiple instances.

Rabbani and Van den Heuvel (2005) presented an efficient 3D Hough transform for detection of cylinders in point clouds. Although 3D Hough transform can be used for automatic detection of cylinders, the required Hough space has a prohibitively high time and space complexity for most practical applications. They addressed this problem and presented a sequential 3D Hough transform for automatic detection of cylinders in point clouds.

Tarsha-Kurdi et al., (2007) focused on the 3D Hough transform and the RANSAC (Random Sample Consensus) algorithm. They made an analytic comparison of both algorithms, in terms of processing time and sensitivity to 3D point cloud characteristics. Despite the limitation encountered in both methods, RANSAC algorithm is more efficient than the Hough transform. The main advantage of RANSAC algorithm is rapidity and the percentage of successful detected roof planes. On the other hand, the 3D Hough transform is very sensitive to the segmentation parameters values.

One of the major problems related with the building extraction process from LiDAR point clouds is the discrimination between buildings and vegetation. One solution proposed by Brunn and Weinder, (1998) was based on differential geometry via Bayesian networks in which edge information was used to extract vegetation areas and building roof structures.

Another method introduced in NASA/ICREST Project report, (2001) makes use of watershed analysis on a DSM by extracting ridgelines that represent edges on the roof structure. The proposed approach for extracting building outlines consists of three processes, which include the generation of a DSM, detecting building outlines and simplifying the building footprints. Initially a Digital Surface Model (DSM) is generated from LiDAR data using a nearest neighbour interpolation. The nearest neighbour interpolation method was chosen because it will preserve the sharp difference between buildings and their surrounding ground.

The building detection is based on general knowledge about buildings that utilize height and shape characteristics to discriminate buildings from other objects. A height threshold can remove objects with lower height, such as cars, and a size threshold will remove some smaller objects, such as single trees. The main problem of filtering larger tree canopies which cannot be addressed using height or size criteria can be resolved based on the roughness of the surface, measured by differential geometric criteria. The main assumption at this stage is that trees present more irregular shapes which can be modelled and detected from the shape description, given by the differential calculus.

In the following step the building outlines are transformed into vector format using raster-to-vector conversion and simplified using an orthogonal simplification algorithm to reduce details in the outlines, while maintaining their essential shape and size (NASA/ICREST Project report, 2001).

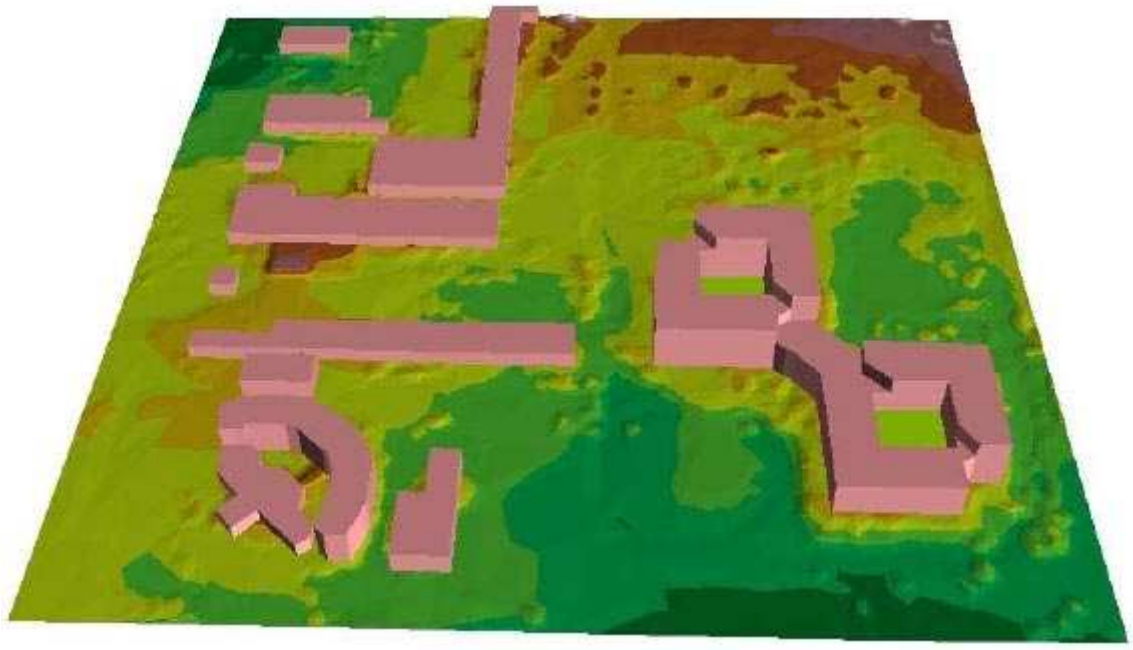


Figure 30 Final result from the reconstructed buildings overlaid on a DTM of the proposed method by NASA/ICREST Project report, (2001)

The 3D building reconstruction process in this method is based on the three common building roof models: flat, gabled, and hipped. The building roof type is determined from a watershed analysis that is conducted to extract ridgelines on the building roofs. Additionally, the height of each flat building can be determined by the average height of pixels within each building boundary.

With the watershed analysis the accumulation and the direction of flow can be calculated. These two parameters can be used to define the watershed boundaries. Cells with a flow accumulation of zero are local topographic heights and can be used to identify ridges. After extracting the ridgelines from the building roofs, hipped and gabled buildings can be distinguished based on shape characteristics (NASA/ICREST Project report, 2001).

The evaluation of the extracted buildings indicated in figure 30, is based on the completeness and correctness criteria. The completeness represents the percentage of the extracted buildings compared to the total number of buildings located in the study area. Correctness indicates the percentage of correct extraction. The comparison that was conducted indicated that 93.7% of the 79 buildings present in the scene were extracted while the correctness approached 90.9%.

Building reconstruction utilizing invariant moments has been proposed by Maas, (2004), applied to LiDAR data for the determination of roof parameters for simple building types. One of the main advantages of the specific method is that it can be applied directly on the irregularly distributed LiDAR point cloud, thus avoiding effects caused by an interpolation to a regular grid. The invariant moments are used to calculate basic building parameters. The ratios from the 1st order invariant moments derived from the segmented point cloud can estimate the length, width and height of a building as illustrated in figure 31. Assuming a standard gable roof as represented in figure 31, the height of the building and the inclination of the roof can be derived from 2nd order moments. In addition, more complex roof structures can be modeled using higher order moments.

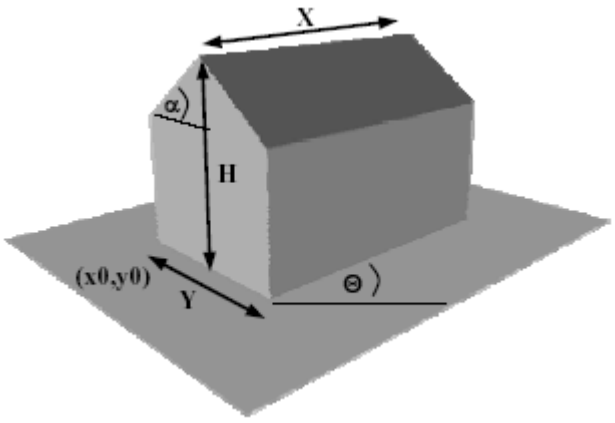


Figure 31 Gable roof building and the parameters for building modelling calculated from the invariant moments (Maas, 2004).

The method initially is segmenting the dataset using height texture measures, followed by morphological filtering and connected component labelling for building detection. From the segmented regions the building parameters are derived as closed solutions from ratios of 0th, 1st and 2nd order invariant moments. After the determination of building model parameters, a goodness of fit can be determined by projecting the model into the point cloud and computing residuals for every

data point. This allows for a rejection of the computed house model in case of bad fit and for the detection and elimination of outliers in the data points, or for a refinement of the segmentation (Maas, 2004). By analyzing differences between point cloud and building model, systematic deviations can be detected from the assumed model and thus dormers on roofs can be modelled.

Alternative roof types may then be determined by the comparison of sets of invariant moments of detected houses with existing house models, stored in a database. The large number of roof types, including varied inclination values, would create a very large database, resulting in decreased reliability of the solution, because of the noise sensitivity of the higher order moments. The precision for the building parameters, is reported by Mass (2004) in the order of 0.1-0.2m for the building dimensions and 1-2 degrees for building orientation and roof inclination.

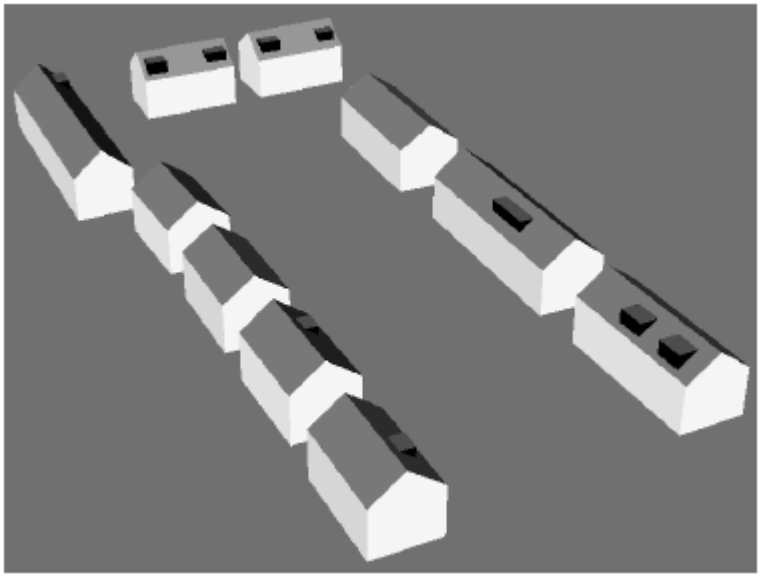


Figure 32 Final results illustrating the reconstructed buildings including the dormers (Maas, 2004).

Figure 32 represents the reconstructed buildings and the successful modelling of the dormers present on the roofs. In addition, this approach presents a time effective method with the computation time per building in the order of 0.8 seconds that was measured at an HP 9000 workstation. The accuracy assessment conducted in the proposed method consisted of model fitting analysis that yielded an RMS of 10cm between point clouds and models of buildings (Maas, 2004).

Nevertheless we should note that the success of this method, in modelling even the dormers, is related with the high density of the point cloud which was approximately 5 points per m^2 as reported in Mass (2004). This characteristic also influences the accuracy of the

extracted building outlines and it was the main reason for the low RMS error estimated in the accuracy assessment. Advantages of the technique are the fact that the parameters of standard gable roof type buildings can be formulated as a closed solution as well as the avoidance of the time consuming interpolation procedure (Maas, 2004).

2.5. Building reconstruction from fusion of different datasets

Various research efforts (Wang, 2008; Cheng et al., 2008; Kaartinen et al., 2005; Vosselman and Dijkman, 2001) confirm that fusion techniques of different datasets for building reconstruction yield superior results. The LiDAR datasets can be very effective for deriving building heights or extracting planar roof faces but it is hard to obtain a detailed and geometric precise boundary using only LiDAR point clouds considering its low spatial resolution. To eliminate noise effects and get building boundaries with precise geometric position, some researchers used the minimum description length (MDL) method to regularize the ragged building boundaries (Weidner and Forstner, 1995; Suveg and Vosselman, 2004). Zhang et al., (2006) used Douglas–Peucker algorithm (Douglas and Peucker, 1973) to remove noise in a footprint, then adjusted the building footprints based on estimated dominant directions. Sampath and Shan (2007) performed building boundary tracing by modifying a convex hull formation algorithm (Preparata and Hong, 1977), then implemented boundary regularization by a hierarchical least squares solution with perpendicularity constraints. However, regularization quality is also dependent on the point density of Lidar data; and limitation of Lidar data resolution and errors in filtering processes may cause obvious offset and artefacts in the final regularized building boundary (Sampath and Shan, 2007).

On the other hand, aerial images provide better results for length determination and deriving the building outline. Although many research efforts have been proposed for building

reconstruction from aerial imagery, few methods have been proposed for integrating 2D map data for building reconstruction.

In Wang (2008), a new approach of *Model-based Building Reconstruction* (MBBR) from topographic maps and LiDAR data called *Floating Models* is proposed. Floating models are a series of pre-defined primitive models which are floating in space. A model's size is adjustable by shape parameters, while its location and rotation is controlled by pose parameters. A building is reconstructed by adjusting these model parameters so the wire-frame model adequately fits into the building's outlines among the topographic maps, LiDAR data, aerial photos and DEM. This model-based reconstruction provides good constraints to the shape of the model in contrary to the data-based approach. In Wang (2008) the model parameters are re-arranged into two groups: plane and height parameters. The plane parameters are determined by fitting the top or bottom boundary of the model to the topographic maps. The height parameters are decided by fitting the top surface of the model to the LiDAR data and interpolating the datum point's height from DEM. The proposed reconstructing procedure is semi-automated. First, the operator chooses an appropriate model and approximately fit to the building's outlines on the topographic map. Second, the computer computes the optimal fit between the model and the topographic map based on an *ad hoc* least-squares model fitting algorithm. Third, the computer computes the roof or ridge height from the LiDAR points within the roof's boundary. Finally, the model parameters and standard deviations are provided, and the wire-frame model is superimposed on all overlapped aerial photos for the operator to check the result. The operator can make any necessary modification by adjusting the corresponding model parameter.

A method proposed by Jibrini et al. (2000) extracted the position and orientation of roof planes from aerial imagery, which was constrained from existing ground plans. The main problem at this stage was that the building height was not known from the ground plans and thus a back projection of the building outlines in the imagery would not be accurate. Hence in order to estimate the building main height, an affine transformation between the image

edges and the edges of the ground plan was introduced. Apart from the affine transformation further alignment is performed, between the ground plans and the edges present on the aerial images, by extending the roof faces beyond the vertical walls and thus taking into account the situations where inconsistencies exist between the two datasets. Research on the fusion of map information and aerial imagery for building reconstruction is focused on model-based reconstruction. Segments of a ground plan usually contain few extracted points and edges for the fitting of shape primitives per segment. Hence, fitting is only performed on aggregated segments. A more data driven segmentation of ground plans will be required in order to be able to deal with more complex buildings (Vosselman, 2002).

A combination of aerial images and DSMs, produced from LiDAR datasets, was introduced by Institut Geographique National for several test sites (Kaartinen et al., 2005). The aerial images were used for the semi-automatic production of pseudo-cadastral maps, that described accurately the building outlines, as well as for the automatic production of true-orthophoto. The building reconstruction was either fully automated or semi-automated, depending on the test site, using a model driven approach (prismatic models) by integrating the LiDAR DSM restricted by the pseudo cadastral maps. The overall workflow included a quality control step by comparing 3D building models with the existing DSM.

Apart from the integration of aerial images with 2D ground plans, considerable research has been made for 3D building model generation by combining extracted roof faces from LiDAR point clouds with ground plans. A method proposed by Vosselman and Dijkman, (2001) utilizes a 3D Hough transformation for plane extraction combined with existing ground plans. The ground plans can accurately estimate the outline of the roof face and subsequently yield the precise location of the vertical walls. In addition this method uses the ground plans for locating roof face edges of buildings.

The method introduces two strategies for combining the two datasets. The first strategy extracts planes from the LiDAR point cloud and segments the ground plan for restricting the plane extraction. Then the initial ground plan segmentation is refined until every segment

corresponds to only one planar face. The second strategy is more generic having the ability to reconstruct more complex roof structures. The second strategy refines the initial models on the bases of fitting models to point clouds that did not correspond to the initial models (Vosselman and Dijkman, 2001).

As mentioned previously, the planar facets are extracted with a 3D Hough transform but spurious planes are generally introduced. In order to minimize the possibility of a region having many different faces, the proposed method splits the dataset into smaller partitioned regions. The partitioned regions are obtained by segmenting the ground plan and extending the edges of the building outline until intersected, as illustrated in figure 33. In addition, the planar facets are further refined by region growing the facets and merging them over the bounds of the segment, introduced from the partitioned ground plans and result in one planar point cloud for every roof plane. For each planar point cloud the final plane facet is then obtained by a least squares plane fitting process.

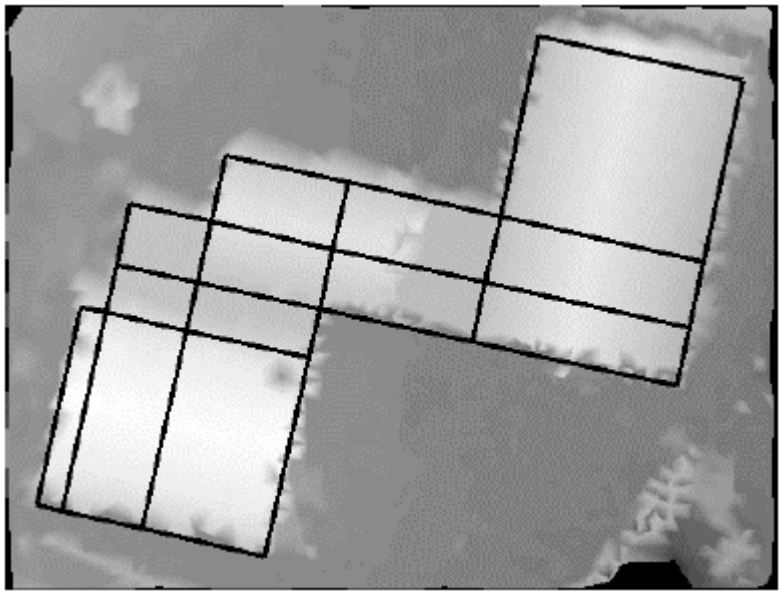


Figure 33 Partitioned building outlines overlaid on a DSM (Vosselman and Dijkman, 2001).

At a subsequent step, the 3D models for each segment are created by combining the planimetric bounds of the ground plan segment with the detected plane. The building reconstruction in most cases requires the detection of 'height jump' edges, especially when complex structures need to be modelled and this stage presents considerable difficulties. In most cases several assumptions need to be integrated. In the method proposed by

Vosselman and Dijkman, (2001) it is assumed that the height jump edge is parallel to one of the edges of the ground plan segment, for simplifying the process.

The detected height jump edges together with intersection lines are two crucial features for the 3D building reconstruction. This is because the process of creating the 3D building models is based on splitting and merging the ground plan segments until there is a one-to-one relationship between the segments and the roof faces. Therefore, the criteria for splitting the segments is based on the presence of intersection lines of height jump edges. The limitation of this strategy is that it can fail to further refine the initial ground plan segmentation in cases of insufficient evidence for the presence of intersection lines or height jump edges.

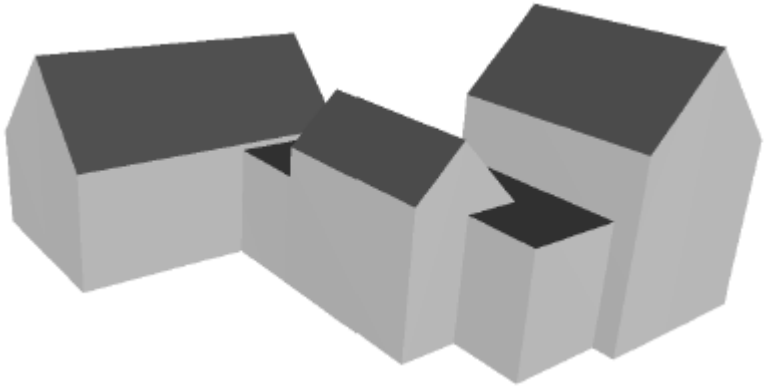


Figure 34 Reconstructed building as resulted from the first method (Vosselman and Dijkman, 2001)

The second strategy can reconstruct more details of the buildings and demonstrated the capability to reconstruct even chimneys, using LiDAR point cloud with density in the order of 5-6 points/m^2. The second strategy may lead to better results even with fewer points, but in some cases it leads to small details that are incorrect.

Various research efforts have tried to combine LiDAR point clouds with aerial images. These methods usually consist of three main steps. Initially, hypothesis and building delineation is performed, either from an interpolated DSM or directly from the LiDAR point cloud. Then polyhedral building models are created for the delineated building regions by a bottom up procedure using both the DSM and the aerial images. Finally the polyhedral building models are verified. A method proposed by Rottensteiner and Jansa, (2002) combines LiDAR with aerial photographs and performs the polyhedral modelling using Voronoi diagrams

(Aurenhammer, 1991). The method follows the three steps described previously and verifies the results by projecting the building models in the aerial images to improve the accuracy of their geometric parameters.

The method initially begins with the generation of the building hypothesis as indicated in figure 35, illustrating the overall workflow. The building hypothesis includes the discrimination between buildings and bare earth as well as the segmentation of the vegetation from the buildings. The method applies initially a hierarchical classification procedure of the LiDAR points with height thresholds to the normalized DSM. The detection of the buildings and the initial building mask is derived by thresholding the differences between the DSM and the DTM. As a result, regions of interest for the geometric building reconstruction are obtained by performing a connected component analysis on the aerial images (Rottensteiner and Jansa, 2002). The building detection from DSM is based on the method of linear prediction but additional processing is required for filtering the vegetation. A possible solution is the use of textural classifiers or Morphological opening filters (Rottensteiner and Jansa, 2002). This method uses terrain roughness, which is similar with the textural classifiers used for optical data, for filtering any remaining groups of trees.

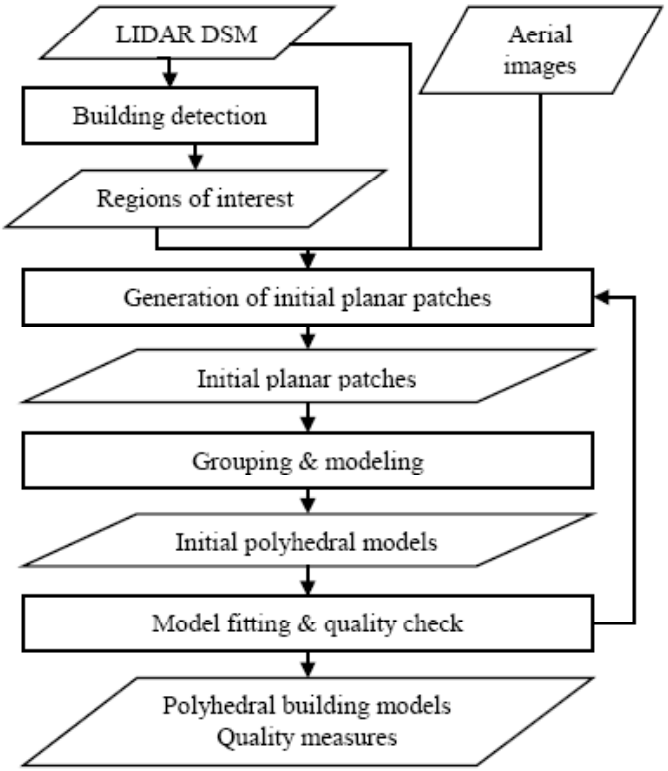


Figure 35 workflow of the method proposed by (Rottensteiner and Jansa, 2002).

After the building detection, the polyhedral models are created based on the initial building hypotheses for detecting planar patches in object space. At this stage the geometric parameters and the initial border polygons for each planar facet are determined. The detection of planar facets is based on a "segment label image" that is defined in the object space. The segmented pixel image is a binary image that contains pixels classified as being homogenious in relation with their surrounding pixels that have similar components to the normal vectors. The initial planar segments are obtained by an iterative connected component analysis process, restricted by 3D line segments that are extracted from the aerial images and a colour-based segmentation.

In a subsequent step the polyhedral models are generated by grouping neighbouring planar segments, utilizing a Voronoi diagram based on a distance transformation of the segment label image. From the Voronoi diagram the neighbourhood relations of the planar segments are derived and the borders of the Voronoi regions can be extracted as the first estimates for the polygons of the planar segments (Rottensteiner and Jansa, 2002).

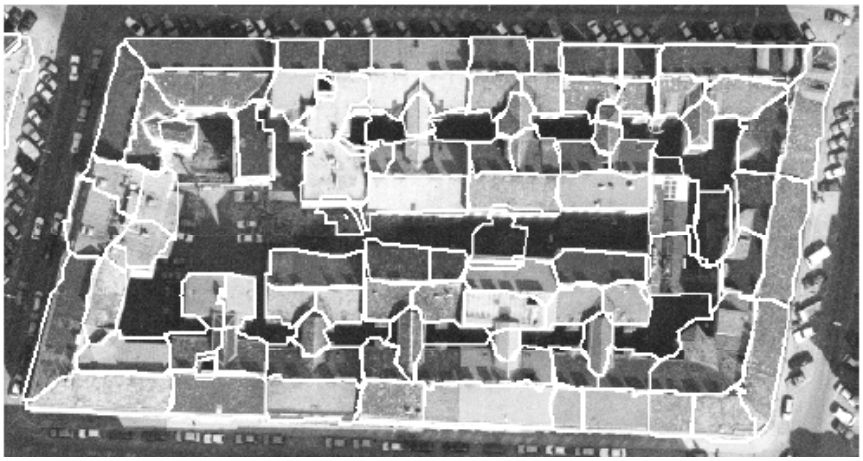


Figure 36 Roof polygons back-projected for the model fitting step (Rottensteiner and Jansa, 2002).

Once the polyhedral model is completed it can be back-projected to the image where its edges can be matched with image edges. This step is crucial for assessing and improving the geometric accuracy of the building models, especially with respect to the building outlines (Rottensteiner and Jansa, 2002).

Building reconstruction using more than two data sources has also been introduced by Chen et al. (2004). This approach combines LiDAR data, aerial imagery and satellite images obtained from the Quickbird sensor. The proposed method is basically a two step procedure that initially detects building regions and then reconstructs the building models. In the stage of building detection, a region-based segmentation and knowledge-based classification are integrated, as indicated in figure 37.

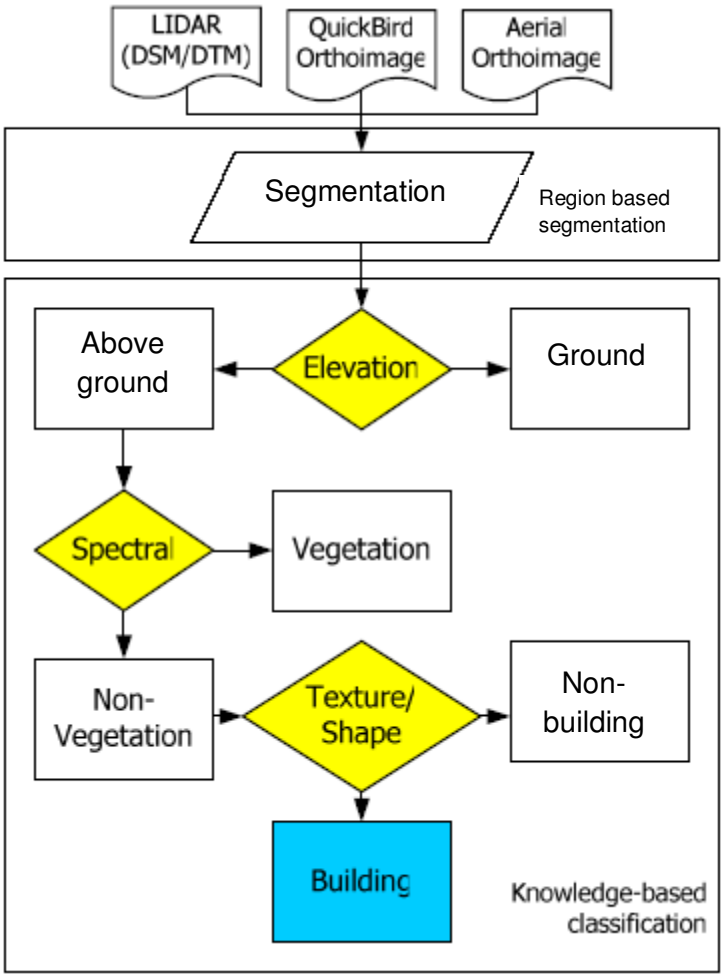


Figure 37 Workflow of the proposed method (Chen et al., 2004).

During the segmentation procedure the LiDAR point cloud are interpolated to create a DSM. The segmentation initially uses a region growing technique, in the produced DSM, to merge pixels with similar attributes. Furthermore a bare earth surface is obtained by applying

elevation threshold for a rough discrimination of objects above ground and on the ground. The above ground surface includes buildings and vegetation that are higher than the elevation threshold. For discriminating between vegetation and buildings, spectral information is used from the multispectral images acquired from QuickBird sensor, together with shape, texture, and elevation information in a knowledge based classification procedure in order to detect the building regions.

The near infrared band provides useful spectral information for estimating the Normalized Vegetation Index (NDVI), which is included as an additional layer in the knowledge-based classification.

Once the building regions are extracted, a TIN-based region growing is performed, to generate 3D planes from the building regions. In addition, the position of the vertical walls is determined by the edges extracted from optical imagery. The method uses approximate location of the edges, constituting the vertical walls extracted from the LiDAR point cloud and subsequently back-projected on the aerial images. The approximate edges are treated as initial values for the Hough transform that is used for extracting accurate straight lines in the image space. The final three dimensional edges are generated, given the image coordinates and the height information, from 3D planes together with the exterior orientation parameters.

After the straight line extraction, the reconstructed vertical walls are combined with the reconstructed rooftops, generated by a 'Split-Merge-Shape' method for building reconstruction. The Split and Merge steps are the two main procedures for topology reconstruction from non-related roof-edges. The Shape step uses the available roof-edge height information to define an appropriate rooftop (Chen et al., 2004). The Split-Merge-Shape method provides high reliability and flexibility, even in cases where the 3D building lines are broken.

2.6. Commercial systems for semi-automated building reconstruction

CyberCity Modeler

Following the semi-automatic approach ‘‘TOBAGO’’ (Grun and Dan, 1997), ‘‘CyberCity Modeler’’ was proposed in 1998 (Grun and Wang, 1998). It is now being commercially offered by CyberCity AG. This modeler is based on conventional, manual measurement of corresponding image points. It uses two basic steps. In the first step (see Fig. 38, left), a so-called ‘‘weakly structured’’ point cloud is obtained by manual stereo measurement. It consists of all gable and roof points, possibly with additional points covering roof structures such as dormer windows. The point cloud is structured by two measures. First, the point measurement has to take place in a certain order, and second, points are assigned certain codes which will guide the second, automatic step. This can be obtained by placing the points into corresponding layers, a functionality which is usually available in existing photogrammetric or CAD software. The second step is a fully automatic processing which uses a relaxation approach to derive the topology of the roof (see Fig. 38, right).

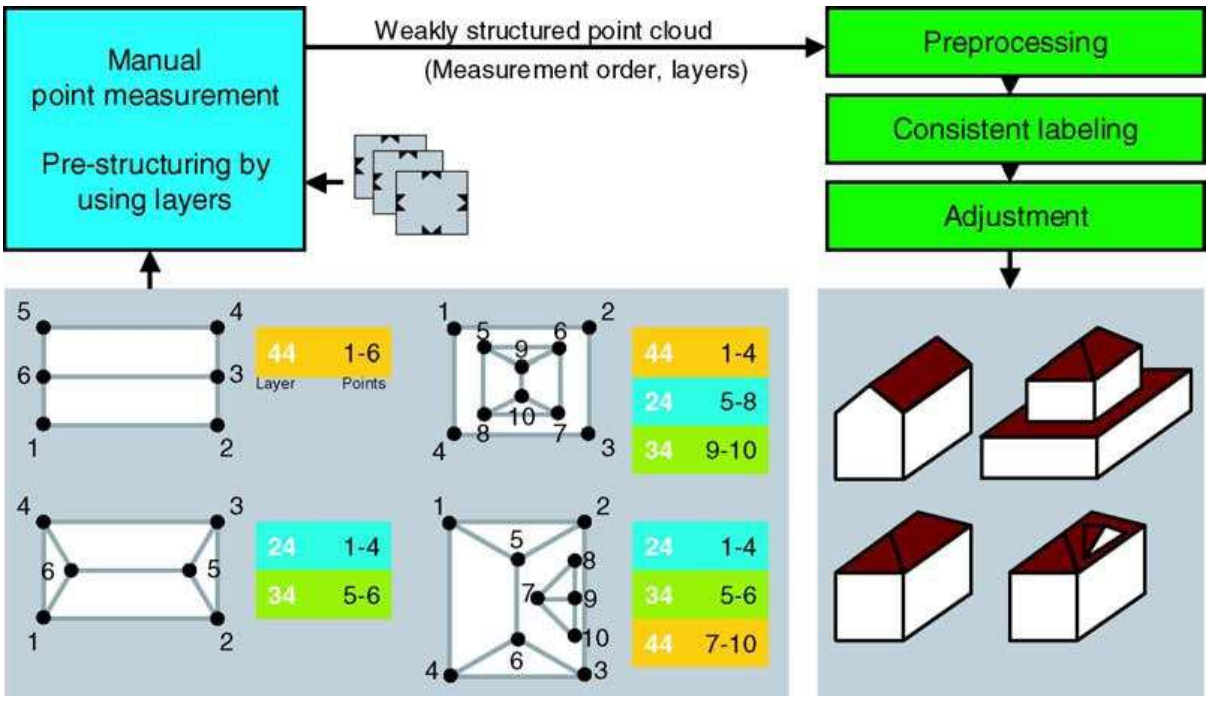


Figure 38 Illustration for the ‘‘CyberCity Modeler’’. A saddleback roof is obtained by measurement of the points 1–6 in succession, in a single layer. For a hipped roof, eaves and gable points are assigned different layers, for complex structures there are even more layers. From Grun and Wang (1998)

Then, using the known correspondences between planes and points, the point positions can be corrected using a least squares adjustment. The result has to be checked, since in some cases, the relaxation might have recovered the wrong roof topology. However, in Grun and Wang (1998) success rates of 95% are given for the automatic structuring step. As compared to a strictly manual measurement, CyberCity Modeler only automates the topology generation step. The point measurement itself is not accelerated. Since the modelling is based on points and planar patches only, the approach is able to reconstruct general polyhedral roof surfaces, as compared to methods using a limited set of building primitives.

inJECT

This system has been developed at the University of Bonn over a long period. It has evolved from earlier approaches termed ‘‘Hase’’, ‘‘Hase+’’, and ‘‘ObEx’’ (Gulch et al., 1998, 1999). It is now being commercially offered under the name ‘‘inJECT’’ by inpho GmbH. Buildings are modelled using a fixed number of parametric primitives like flat-, desk-, saddleback-, hipped-roof, etc. (see Fig. 39). The selection of the appropriate primitive is carried out manually by an operator. Then, the wireframe model is overlaid in two images and the operator can adapt the parameters accordingly. A guided mode is available where the operator is asked to measure certain points only. For example, for a saddleback roof, the operator has to measure only the two gable points in one image—after that, the system tries to find the corresponding points in the other image as well as the remaining parameters automatically. After the automatic step, the operator can assess the result and can correct it by manual measurements, if necessary. For the automatic methods, a success rate of 50–90% is given in Gulch et al. (1999). Complex buildings are subdivided by the operator into building parts which are modelled independently. The parts are then later on merged by an external CAD software to obtain the final buildings. Using this constructive solid geometry (CSG) approach the number of required measurements by the operator is reduced considerably. It is even further reduced by the provided automatic matching and snapping procedures. On the other

hand, as with any system which uses CSG as a modelling principle, some more complex buildings might be quite hard to model using the fixed set of primitives.

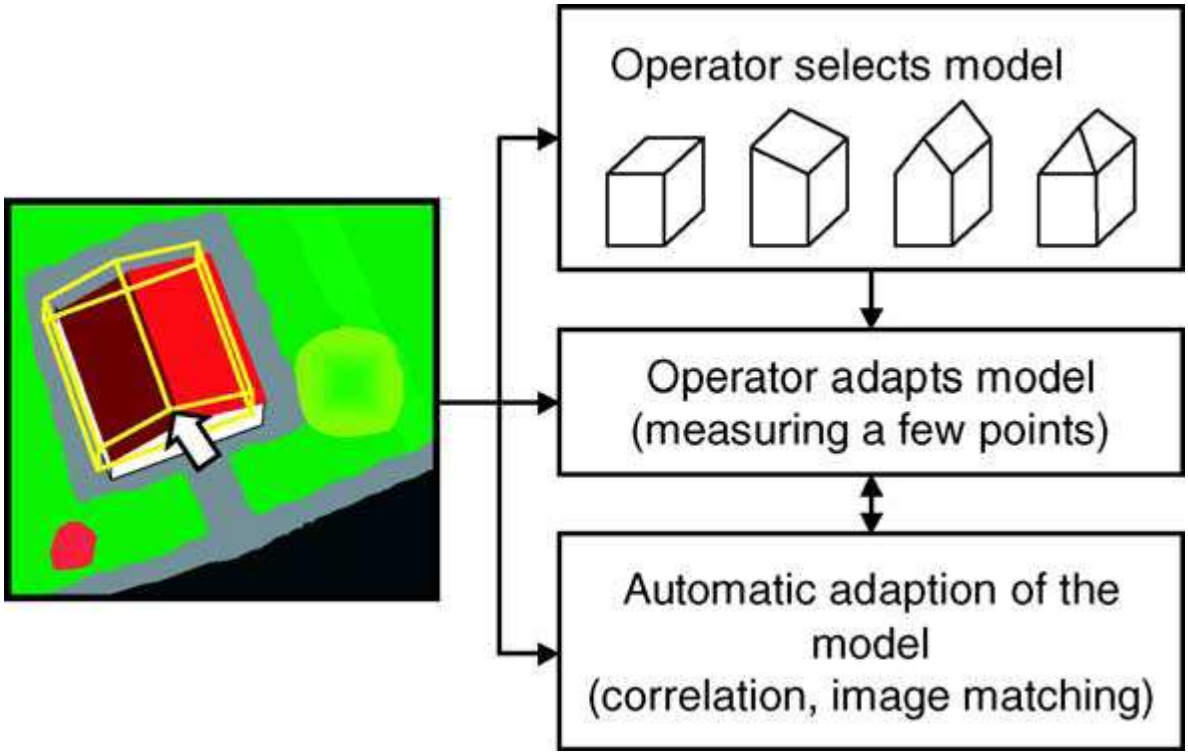


Figure 39 Illustration for inJECT. The operator chooses a primitive, which is then fitted to the images using manual measurement as well as automated measurement procedures.

2.7. Strengths and weaknesses of previous research efforts

In summary most of the research efforts introduced so far for automatic building reconstruction from aerial images have several reliability issues and problems due to occlusions, introduced mainly from vegetation, scene perspective and the complexity of the roof structures. That is especially true for the data-based methods as introduced before (Wang, 2008; Cheng et al., 2008; Heuel et al., 2000; Ameri and Fritsch, 2000; Scholze et al., 2002) because the building reconstruction is performed without any assumption for the structure of the roof and in most cases artefacts and errors are introduced. One of the main

strengths of all the data-based methods using aerial photography is that there are no restrictions introduced by using a set of finite roof models and these methods can potentially reconstruct a wide variety of roof styles. On the contrary the model-based methods (Willuhn and Van Gool, 2005) use some models of buildings to restrict the set of possible shapes. This external knowledge enables the user to overcome the lack of detection due to occlusions but also restricts the solution to the pre-defined roof styles.

Previous research efforts for automatic building reconstruction from LiDAR data have one characteristic weakness. That is the lack of adequate planimetric accuracy for the building footprints and loss of roof detail when only coarse LiDAR point clouds are available. Nevertheless when high density LiDAR is used with methods that directly derive the surface parameters in a parameter space by clustering the point cloud, the solutions are very effective and robust when planes or other simple shapes are extracted (Vosselman, 1999; Gorte, 2002). Mass, 2004 introduced a very efficient and fast method for building reconstruction utilizing invariant moments on LiDAR point clouds.

Fusion techniques for building reconstruction provide more accurate results, since they combine the advantages of very good vertical accuracy, usually provided from LiDAR data, and the precise location of the building outlines derived either from aerial images or ground plans. A quantitative evaluation presented in Kaartinen et al. (2005) indicates the higher geometric accuracy achieved by data fusion approaches. The evaluation indicates that in the assessment of all test sites, the RMS error of photogrammetric methods for extracting building outlines ranged from 14 to 36 cm with a standard deviation of 7.2 cm. The corresponding values for aerial images assisting laser scanning ranged from 20 to 76 cm with std of 18.5 cm. Laser scanning-based building outline errors ranged from 20 to 150 cm with a std of 33.2 cm (Kaartinen et al. 2005).

This thesis presents an effort to improve the weaknesses of the data-based methods by improving the reliability of the building detection and 3D reconstruction stage. This is achieved using data-fusion techniques between stereo matched points and LiDAR point

clouds to improve the building footprint and 3D roof breaklines. The building detection is performed using an adaptive k-means classification algorithm in combination with geomorphometric region growing and parametric object representation to enhance the reliability of detecting building footprints. The strength of the proposed workflow is that it doesn't use any building primitives for the 3D building reconstruction as opposed to the model based methods but maintains a high level of reliability, and level of detail. Another weakness of the previous research efforts discussed above is the issue of adaptability. Most methods are suitable for specific datasets and sites with specific building architecture. This thesis presents a unified, end-to-end workflow that can use data-fusion techniques or Very Dence DSMs created solely from airborne imagery to detect buildings and reconstruct the roof structures

2.8. Adopting previous methods in the proposed workflow for semi-automated 3D city modelling

From the previous section it is evident that the efficiency of different approaches is strongly related to the data at hand and each method presents some advantages and weaknesses. The proposed method presented in this thesis has four major steps; the feature extraction from the optical data, the building detection, the adjustment of the building outline using data fusion techniques and the building reconstruction.

The method for feature extraction follows the general workflow of the method proposed in Chen et al. (2004). According to this method the approximate location of the edges, constituting the vertical walls, is extracted initially from the LiDAR point cloud and subsequently is refined based on the features derived from the optical data. The main difference in the proposed method of this thesis is that the features are conjugate points from stereo pairs of aerial photographs, instead of edges used in Chen et al. (2004). Another

difference is that the adjustment of the building outline is performed using a least square fitting solution and not the Hough transform used in Chen et al. (2004).

The classification of vegetation in the study area is performed using two generic tree models that scan the LiDAR data. Therefore, the proposed method for classifying vegetation does not follow any of the previously mentioned methods. Further details for this procedure and the reasons for choosing this approach are described in section 5.3. Nevertheless the classification of low features, above the ground, is performed by adopting the method used in Chen et al. (2004). That is, after the classification of the ground points, low features that are not related to buildings are filtered using a range of relative heights above the ground surface.

The procedure of building detection and generating the building hypothesis is performed with the method proposed in Rottensteiner and Jansa, (2002), whereby the generation of the building hypotheses is based on the detection of planar patches in the LiDAR point cloud.

The roof reconstruction process utilizes the first approach introduced by Vosselman and Dijkman, (2001). According to this approach, the adjusted building boundaries can accurately estimate the outline of the roof face and subsequently yield the precise location of the vertical walls. The first strategy introduced in Vosselman and Dijkman, (2001), extracts planes from the LiDAR point cloud and segments the ground plan for restricting the plane extraction. A similar approach is used in the proposed method with the adjusted building outline restricting the plane fitting of the roof structures and therefore producing an accurate result for the building models.

3. AIRBORNE DATA COLLECTION SYSTEMS

This chapter describes in detail the characteristics of the data collections systems used to capture aerial photography and LiDAR data from the three test sites, Switzerland, London and Nottingham. The chapter describes the ADS40 digital camera, the ALS50 LiDAR system, the Leica RC30 frame camera, the ALTM 3033 LiDAR system and Vexcell UltraCam D digital camera.

3.1. Airborne Digital Sensor ADS40

The ADS40 (figure 40) was used to capture the airborne imagery available for the Swiss, (Heerbrugg) test site. It was introduced at the XIXth ISPRS Congress in Amsterdam in July 2000. It was developed jointly with LH systems and Deutsches Zentrum für Luft-und Raumfahrt (German Aerospace Centre – DLR) utilizing forward, nadir and backward-looking linear arrays on the focal plane to acquire panchromatic and multispectral imagery. It presents one of the first attempts to design a digital airborne sensor that could achieve similar spatial resolution and coverage compared to a film based aerial camera.

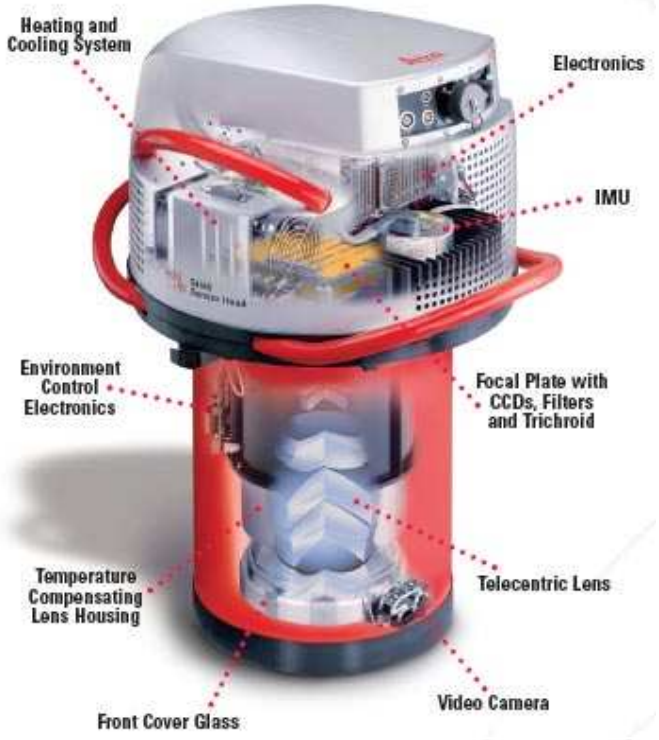


Figure 40 The major components of the ADS40 (ADS40 flyer, 2004)

The development of the ADS40 was split into four stages to ensure the quality of the digital sensor: The functional model, the engineering model, the prototype model and finally the series model, which included the final ADS40 product (Eckardt et al., 2000).

One of the major advantages of the

three line airborne digital sensor compared to area array sensors is the acquisition of strip imagery which produces a seamless scene that provides continuous processing in all photogrammetric procedures. Therefore it considerably increases the coverage and reduces the time requirements during the tedious process of mosaicking numerous individual images (Fricker, 2001).

3.1.1. Design principles

The basis of the design is the three-line scanning principle that results in three different views, forward from the aircraft, vertically down and looking backward (figure 41). The imagery from each scan line provides information about the objects on the ground from the different viewing angles assembled into three strips.

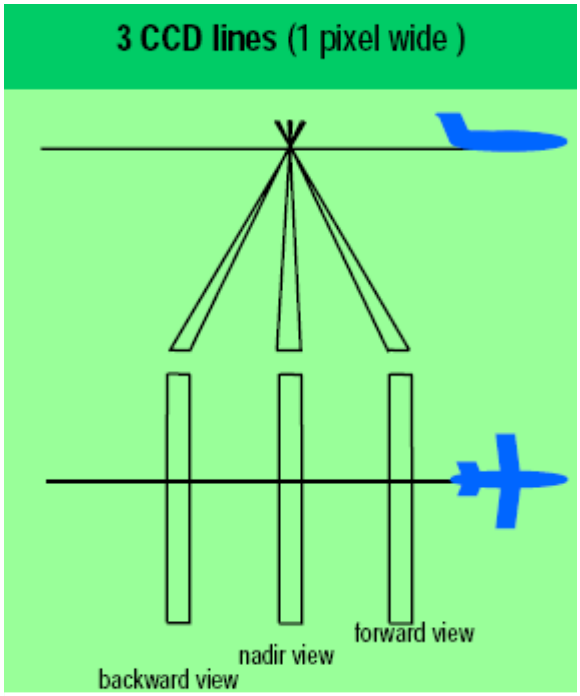


Figure 41. Scanning principle of the three-line digital sensor (Sandau et al. 2000).

The panchromatic linear sensors consist of pairs of 12,000 pixel arrays, one member of the pair staggered laterally from the other by half a pixel (3,25µm). The result is that every portion of the ground surface is imaged three times, far superior to the 60% triple coverage of the conventional film aerial photography. Additionally on the focal plane are four further linear arrays of 12,000-pixels acquiring imagery, through interference filters, in the red, green, blue, and near infrared portions of the spectrum (Fricker, 2001).

The colour linear CCD arrays are optically superimposed during the flight. This is accomplished with trichroitic mirrors as illustrated in figure 42, which divide a single light ray into its three colour components without significant energy loss. The advantage of this approach is that the RGB bands are co-registered without significant post-processing, thus resulting in a good quality colour composite. The near infrared sensor lines are 12000 pixels long and slightly offset from the RGB triplet. The precise position of each pixel is known after the calibration process (Tempelmann et al. 2000).

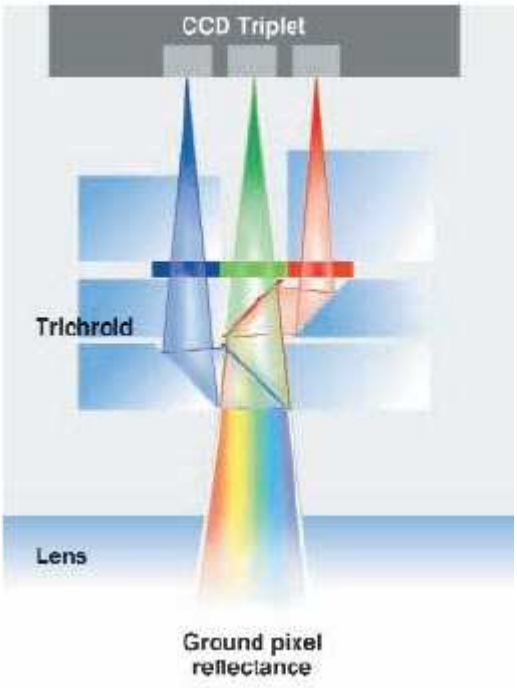


Figure 42 The trichroid beam splitter, adapted from ADS40 brochure, 2002. Copyrights reserved by Hexagon Corp.

The performance of the linear CCD arrays and the filters is maximized through the use of a telecentric lens. The incoming light rays are impinged on the focal plane at right angles regardless of their angle of incidence at the front nodal point (Fricker, 2001).

The attitude and positioning of the ADS40 is provided by a Position and Orientation System (POS) from Applanix. It consists of an inertial measurement unit (IMU), located internally in the camera head and therefore eliminating the problem of relative motion between the camera and the IMU, that may be present if the IMU were mounted externally. The data generated by the POS are stored as part of the mission data in the mass memory system and can be retrieved for post-processing after the flight. For flight guidance and navigation purposes the real-time data of the POS are provided to the navigation module (Sandau et al. 2000).

The integrated GPS/Position and Orientation System (POS) is used during the post processing stage for direct georeferencing and the production of Level 0 images. The main

key features of the integrated GPS/POS system include the reduction of the time requirements for data processing and price/performance ratio of the system.

The main components of the GPS/POS system are illustrated in figure 43. It consists of the IMU mounted on the focal plane head that tracks the six degrees of freedom providing high frequency information for the velocity, attitude and relative position.

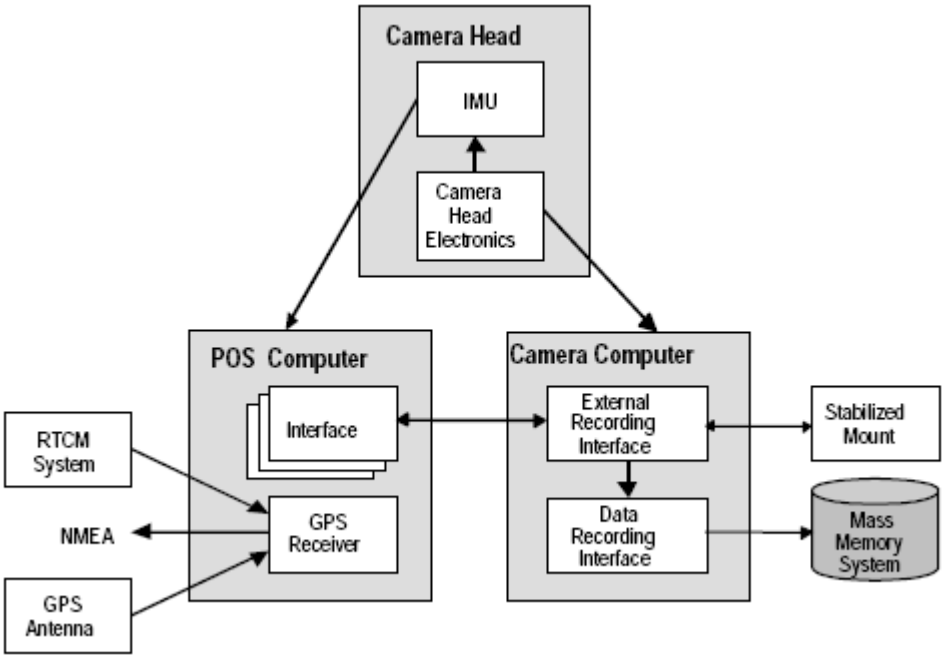


Figure 43 Main components of the integrated IMU/GPS system in the ADS40, adapted from Sandau et al. 2000.

3.1.2. ADS40 Sensor Calibration

The geometric calibration of a digital sensor includes the determination of two basic parameters.

• Accurate measurement of the position of the CCD detectors and determination of the image co-ordinates.

• Determination of the modulation transfer function (MTF), see below, along-truck and across flight directions.

The determination of the position of the light sensitive elements is made by illuminating the pixels with a collimating instrument. The direction of the illuminated pixel to the collimator axis is determined by the angle along the line direction and the angle perpendicular to the line direction. These two angles can be exchanged in the case of horizontal mounting of the camera (Schuster and Braunecker, 2000)

The MTF is determined for both directions along the flight path and across track by estimating initially the point spread function (PSF). The PSF characterizes the optical performance of the digital camera system, which is mainly determined by the pixel size and the optics. With the estimation of the PSF the MTF can then be calculated by a Fast Fourier Transform (FFT) along an arbitrary number of points on the linear CCD. Figure 44 indicates the MTF curves for the along and across track. From early calibration tests it is evident that the resolution always seems to be better in the flight direction rather than in the across track direction.

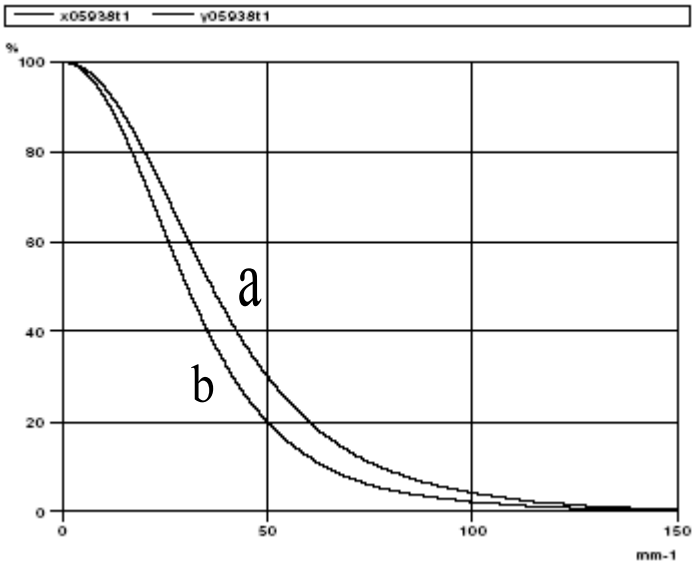


Figure 44 MTF along the flight direction (a) and across track direction (b) for the nadir looking view (Schuster and Braunecker, 2000)

Another stage of the geometric calibration is the measurement of the distortions present along the linear CCD sensor. The distortion curves along the nadir looking CCD sensor are given in figure 45, determined from two calibration centres that indicate distortions of less than one pixel over the whole line of 12,000 pixels. The measurements performed from the two calibration centres for the geometrical calibration process include two polar angular co-ordinates, assigned to each measured pixel.

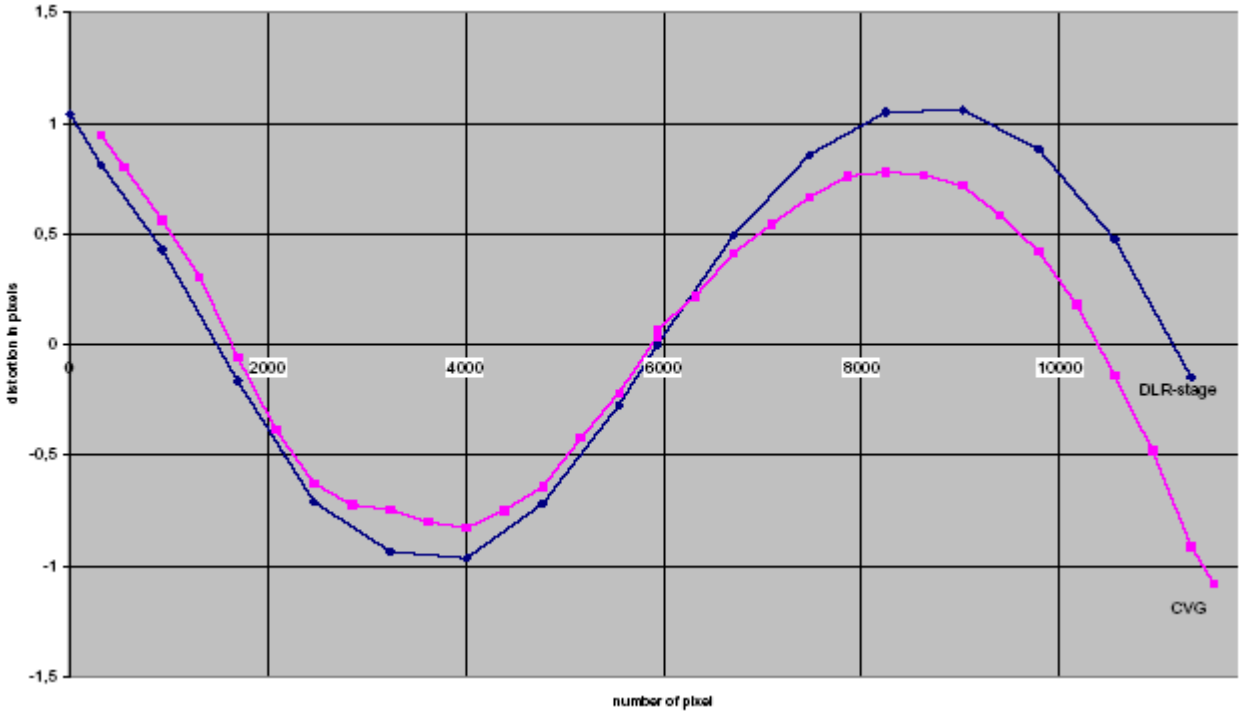


Figure 45 Distortion in pixels estimated from the engineering models of DLR and LH systems for the nadir line (Schuster and Braunecker, 2000).

3.1.3. Post-Processing workflow of ADS40 data

The mass memory MM40 is removed from the aircraft and connected to a docking station on the ground, where the ground processing system obtains the data through SCSI interface. The images and metadata are downloaded which can take several hours since the capacity of the mass memory is 0.5 TB.

After archiving, the GPS/IMU data are processed with the GPS base station data using the Applanix PosProc software. The orientation data are then used together with the camera calibration parameters to create Level 1 rectified images, which can be used for photogrammetric processing in digital photogrammetric workstation. Further products and level 2 rectified orthophotos can also be created (Figure 46).

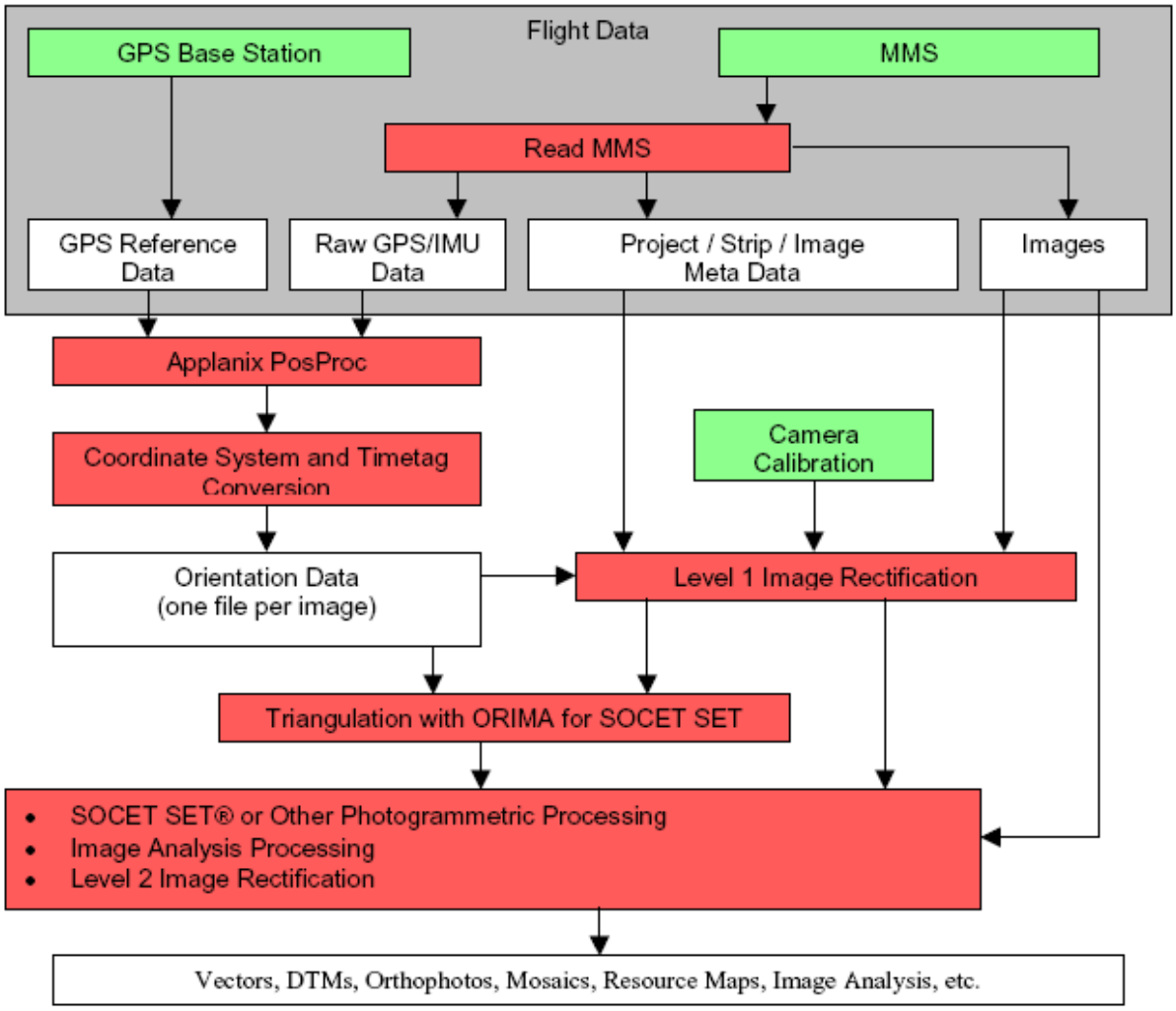


Figure 46 Ground processing workflow of the ADS40 dataset (Tempelmann, et al. 2000).

3.1.4. Geometric accuracy and overall quality

The first series of test flights for ADS40 were conducted in Switzerland above Waldkirch and Sion. The aerial triangulation for these blocks of images was performed with the ORIMA software package (Fricker, 2001). Table 1 summarises the characteristics of the early test flights.

Location	Waldkirch	Sion	Chofu	Tokyo	Tsukuba	Parma	Pavia	Adria	Verona
strips flown	8	8	4	2	4	8	3	3	3
shape	block	block	block	single line	block	single lines	block	block	block
cross-strips	yes (3)	yes (3)	yes (2)	-	yes (2)			no	no
flying heigth (above ground)	2000 m	1750 m	1950 m	1950 / 2750 m	2000 m	1500 / 1800 / 2000 / 3000 m	1600 / 2300 / 5750 m	1550 m	1500m
general direction	E-W	120°	E-W	E-W	N-S	120°	E-W	160°	120°
GSD [cm]	20 cm	17 cm	20 cm	20 & 15/30 cm	20 cm	15 & 20cm	15/30/50cm	15cm	15cm
Number of CCD's	7	3	7	10	7	7	7	7	7
Total project size	10.1 GB	5.96 GB	73.6 GB		126 GB	85.2 GB	64.5 GB	111 GB	

Table 1 Block configuration of the initial test sites (Fricker, 2001).

From the post-processing of the early test flights it was concluded that a robust bundle adjustment, few ground control points and automatically extracted tie points were still necessary in order to produce a parallax free stereo model, suitable for accurate feature extraction. This was evident since the Level 1 rectified images using the GPS/IMU data can produce good stereo viewable images but in most cases they maintain a small amount of parallax and therefore, hinder the ability to comfortably view the stereo model and extract features for large scale applications.

Strips	4
Cross strips	2
Control points	18
Orientation fixes	216
Tie points	5420
Bundle adjustment options	Self-calibration; automatic blunder elimination
σ_0 (standard error of unit weight)	5.3 µm (0.8 pixel)
RMS for ground control points	<5 cm

Table 2 Characteristics and accuracy results from the test flight in the area of Sion (Fricker, 2001)

One of the early triangulation tests was conducted over the area of Sion in the southwest of Switzerland. The block consisted of four parallel strips with two strips at right angles. A total of 18 check points were used for assessing the geometric accuracy of the aerial triangulation. No preparation was made to signalize points or determine ground control points in advance. The bundle adjustment included self calibration parameters and automatic gross errors detection, yielding a standard error of unit weight equal with 5.3µm

and a RMS of 5cm, compared to the ground control points. Table 2 summarizes the block configuration characteristics and the accuracy of the results.

Evaluation of the ADS40 was additionally conducted over the area of Waldkirch in Switzerland described in Alhamlan et al. (2004). The characteristics of the block configuration are indicated in table 3.

ADS40 imagery	Six strips with GPS/IMU data
Coverage	8 x 8 km
Location	Waldkirch, Switzerland
No. of GCPs	30
Focal length	62.5 mm
Flying height	2000 m
Pixel size	6.5µm
GSD	200 mm

Table 3 Characteristics of the block configuration for the Waldkirch test site. (Alhamlan et al. 2004).

The evaluation was based on 18 check points, measured by GPS, used for evaluating the accuracy of the different geometric configurations, created from the multiple look angles.

The quality of the georectified imagery can be improved with an aerial triangulation. The four parallel strips consisting the block of rectified imagery were triangulated and the accuracy was evaluated based on 30 check points. Table 4 summarizes the RMS errors of the check points for the forward and backward scenes. Similar results were obtained for the other combinations of look angles.

From the specific test site it was concluded that despite the use of GPS/IMU data, the aerial triangulation of ADS40 images requires a few GCPs in order to provide good precision and the required accuracy for photogrammetric applications.

Config.	GCPs	RMS X (m)	RMS Y (m)	RMS Z (m)
1	0	0.342	0.229	1.039
2	4	0.172	0.212	0.342
3	9	0.160	0.154	0.252
4	12	0.275	0.245	0.337

Table 4 RMS for the forward/backward scene combination (Alhamlan et al. 2004).

3.2. Airborne frame film camera, Leica RC30

The Leica RC30 was used to capture the pair of aerial photographs of the second test site in London, Bloomsbury area.

Figure 47. The Leica RC-30 aerial frame camera with navigation sight (adapted from Leica RC30 brochure, 2007, Copyrights reserved by Hexagon Corp.)

The Leica RC30 aerial film camera is the continuation of a series that started with Wild's first aerial camera in the 1920s. Since then there have been several new features added and the latest model is enhanced with the PAV30 gyro-stabilized mount and other sub-systems such as the NSF3 navigation sight. The camera can also be supplemented with airborne GPS and IMU systems. The Leica RC series of aerial camera systems has been characterised by the robust design and their overall quality and long term reliability and has made them a popular choice for many mapping companies and national mapping agencies.

The latest model, the RC30, offers two interchangeable lenses, minimal distortion and reaches lens/film resolutions well over 100 lp/mm. Apertures up to f/4 and shutter speeds from 1/100 to 1/1000 second maximise applicability. Forward motion compensation

produces sharp photos at low light levels and altitudes. One of the advantages of the RC30 is that additional film cassettes are inexpensive, so prolonged missions are feasible. Other big advantages include the PEM-F automatic exposure control and the flexible data annotation on each photograph. Navigation sights, viewfinders and a wide range of filters complete this solution.

The RC30 can be combined with the PAV30 Gyro-stabilized Camera Mount and the Aerial Survey Control Tool (ASCOT) for a complete aerial survey system.

3.2.1. RC30 main Components

The PTW30 drive unit consists of the Forward Motion Compensation mechanism, the PEM automatic exposure control, the EDI film data recording and a RS232 interface.

The EDI records several internal and external data such as the aircraft voltage, shutter and aperture settings, mission data from aircraft as well as navigation data. The FMC mechanism can provide a maximum correction of 640µm image motion with both manual and automatic modes. The PEM automatic exposure control consists of a control board with an embedded microprocessor that controls the shutter and aperture settings, the film speed, exposure correction etc. The microprocessor is linked to a sensor that records the illumination conditions with a spectral sensitivity of around 700nm and a spectral range from 400nm to 1000nm. The film cassette compartment consists of a pair of PKA4 film cassettes with a width of 240mm and three different film lengths ranging from 120m up to 219m. Additional film cassettes can be used during the flight.

Two interchangeable lens cones are provided with the RC30. The 15/4 UAG-S lens cone and the 30/4 NAT-S lens cone. The first lens cone has a wide angle lens with 90^{o} field of view and a calibrated focal length of 15cm. The later lens cone has a normal angle lens with a 55^{o} field of view and a 30cm calibrated focal length.

The Leica PAV30 Gyro-Stabilized Camera Mount compensates for angular aircraft movements due to turbulence, including pitch, roll and drift. As a result, survey flights are more efficient and result in sharper images, as well as reduced stress on the flight crew. The PAV30 is designed to carry a variety of camera payloads.

3.3. UltraCam D digital sensor

The UltraCam D digital sensor was used to capture the block of aerial images of the third test site, covering The University of Nottingham University Park campus. The UltraCam D digital sensor was introduced in 2003 by Vexcel Corporation which since 2006 has been owned by Microsoft Corp.

Figure 48. The UltraCamD sensor depicting the panchromatic and colour cones (Copyrights reserved by Microsoft Corporation)

3.3.1. Design Concepts & Calibration

The sensor unit of Vexcel's UltraCamD consists of eight independent cameras, so-called cones. Four of them create the large format panchromatic image at a size of 11500 by 7500 pixels. The other set of four cones is responsible for the multi spectral channels of the UltraCamD, i.e. red, green blue and near infrared (figure 48) (Smith et al., 2005, Kruck, 2006, Gruber and Ladstädler, 2006). The main sensor specifications are given below

• Panchromatic, RGB and Near Infrared imagery captured on a single pass

• 11500 pixels perpendicular to the flight direction

• 7500 pixels along the flight direction

• Focal length = 101.400mm

• CCD array sensor size =103.5 x 67.5mm

• CCD pixel size = 9 µm

The 4 lens cones in a line through the centre of the cone cluster are used to capture the panchromatic image which is made up of 9 overlapping sub-images to create a composite image as shown schematically in figure 49. The sub-images have been given a letter to show which images were captured by the same lens cone.

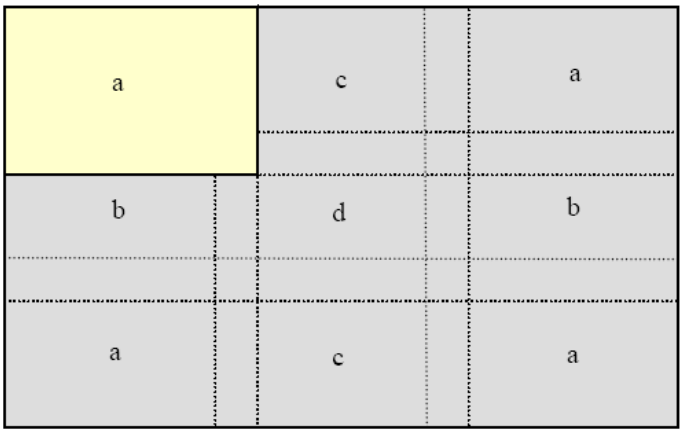


Figure 49. Schematic of the 9 sub-images making up the panchromatic image with one image highlighted (Adapted from Smith et al. 2007)

The panchromatic part of the camera combines a set of 9 medium format CCD sensors into a large format panchromatic image. The multispectral channels are supported by four additional CCD sensors. Each of these 13 CCD sensors is the front end of a separate imaging module. It consists of the sensor, the sensor electronics, a high end analog/digital converter (ADC) and the IEEE 1394 data transfer unit. The raw image data is transferred via the IEEE 1394 interface to a separate storage module of the Storage and computing Unit of UltraCamD. The computing unit consists of the UltraMap Server with a tape robot, with 200,000 uncompressed images available near line via the Internet or an Intranet. Included are RAID disks and multiple CPUs, plus considerable software for cataloging, archiving, project management and processing of UltraCam images.The camera offers a frame rate of more than 1 frame per second, exploiting the benefit of its parallel system architecture (Leberl et al. 2003).

The UltraCam produces for each collected image an independent set of 5 spectral bands: pan, blue, green, red and near-infrared. As an object moves through the camera's field-of-view, it gets repeatedly imaged and a set of new spectral observations gets collected. The UltraCamD produces the pan-channel at a geometric resolution exactly 3 times greater than the resolution of the color bands. The visualization therefore employs a process of "pan-sharpening" at the ratio 1:3, to take the best from each channel (Leberl et al. 2005).

The 'new' multi cone digital camera systems are geometrically complex systems. The image used for photogrammetric analysis is made up of a number of images produced by a cluster of camera cones and various groups of CCD arrays. This produces a resultant image which is not just based on traditional single lens/focal plane camera geometries but is dependent on the joining of images from multiple lens (different perspectives), groups of focal planes and the matching of overlapping image areas (Smith et al. 2007)

This joining of images from the multiple lenses is of critical importance and is determined through a calibration procedure. The same calibration procedure estimates also the

additional interior orientation parameters. The basis for the calibration is a set of images from a well defined and precisely surveyed control point field (Kropfl et al. 2004).

The mathematical background of the calibration procedure is the co-linearity equation as it is known from traditional analogue cameras and is used to determine the parameters of the pinhole camera plus lens distortions. In addition though, a third set of parameters is involved in the UltraCamD geometric calibration.

The parameters derived through the calibration procedure build up the full geometric description of each cone of the UltraCamD. The internal linear transformation between cones is determined by a stitching process, based on highly redundant tie point matching results. Calibration parameters and internal transformation parameters are then used to produce the distortion-free UltraCamD output image.

More specifically the entire adjustment procedure is a fourfold process and includes the following steps:

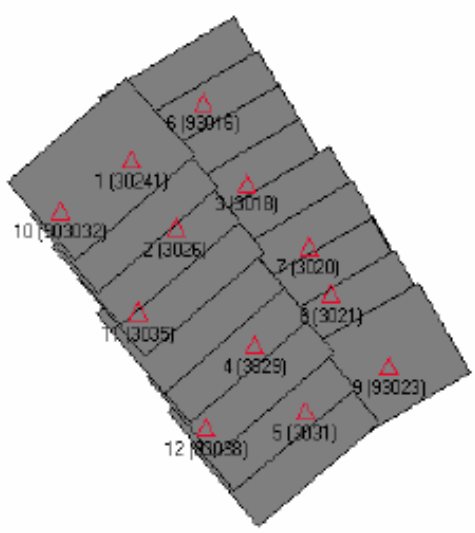
- Calculation of the initial solution of camera parameters including parameters of CCD position, principal distance and principal point coordinates and lens distortion.
- Transformation of image coordinates (the measurements) to clear CCD position parameters. This step needs several iterations in order to avoid any eccentricity of the radial distortion parameters of the lens cone.
- Description of remaining distortion on the CCD defined by a look up table.
- Estimation of transform parameters between cones in order to guide the post-processing (stitching) of the large format panchromatic image and the registration of the multispectral channels to the panchromatic image (pansharpening) (Kropfl et al. 2004).

3.3.2. Accuracy assessment & performance

There have been several attempts to quantify and evaluate the performance of the UltraCamD digital camera. Smith et al., 2005 were one of the first groups to critically evaluate the sensor using data from the UK, specifically over Milton Keynes. The data collected from that particular study included a block of 18 images taken from a nominal flying height of 1500m with a Ground Sample Distance (GSD) of 0.13m and a second block of 30 images taken from a nominal flying height of 760m and a 7cm GSD. Both blocks had a 60% overlap. Because the data were captured initially for commercial purposes the amount and distribution of ground control points (GCPs) was not ideal and the GPS base station used was at Northampton some 30km from the test flight. That was a long baseline for high quality kinematic GPS (Smith et al., 2005). The tests were divided into two categories.

- Aerial triangulation based only on ground control.
- Aerial triangulation based on ground control and inflight GPS and IMU measurements.

The results from the aerial triangulations using only ground control points are given in figure 50.



No of Strip /GCP	Tie points RMSE (m) of standard deviations			Ground control points RMSE (m) of residuals		
	X	Y	Z	X	Y	Z
Two /12	0.076	0.067	0.271	0.145	0.077	0.029
Two /4	0.063	0.058	0.271	0.071	0.067	0.004
One* /7	0.082	0.068	0.260	0.048	0.049	0.019
One* /4	0.083	0.071	0.313	0.036	0.040	0.020

Figure 50 The block of images taken at 1500m and distribution of control points (left), residuals of tie and control points from the different Aerial Triangulations using only the GCPs (right). (adapted from Smith et al., 2005)

The results are very good when considering the flying height and a 60% overlap instead of 80% overlap that would give a base to height ratio closer to one. Tie point RMSE values are reasonably consistent where as there is some variation in the RMSE of the residual for the GCPs. When the number of control points is small the influence of an individual point becomes more significant. As can be seen there is a small RMSE in Z for the two strips of 12 GCP solution and the Z RMSE for the tie points standard deviations is probably showing the effects of the relatively small airbase (base to height ratio, 0.27) (Smith et al., 2005).

It is interesting to note that in that study the results from the aerial triangulations including in-flight GPS and IMU didn't present any particular improvement over the original GCP only solutions (Table 5).This might be the expected effect from the long baseline that was used to compute the in-flight GPS values (30km). With a shorter baseline that would normally be used an improved solution might be expected. In addition, it does show the strong solution produced by the imagery and ground control.

No of Strip /GCP	Tie points RMSE (m) of standard deviations			Ground control points RMSE (m) of residuals		
	X	Y	Z	X	Y	Z
Two /12	0.124	0.107	0.388	0.150	0.114	0.044
Two /4	0.125	0.109	0.357	0.107	0.110	0.045
Two /0	0.121	0.112	0.276			

Table 5. Tie and CGP analysis for the ATs using GCPs/GPS and IMU (adapted from Smith et al., 2005)

Smith et al. 2007 presented a new methodology for self calibrating the UltraCamD sensor. The project involved datasets from a EuroSDR test site in Fredrikstad-Norway with two

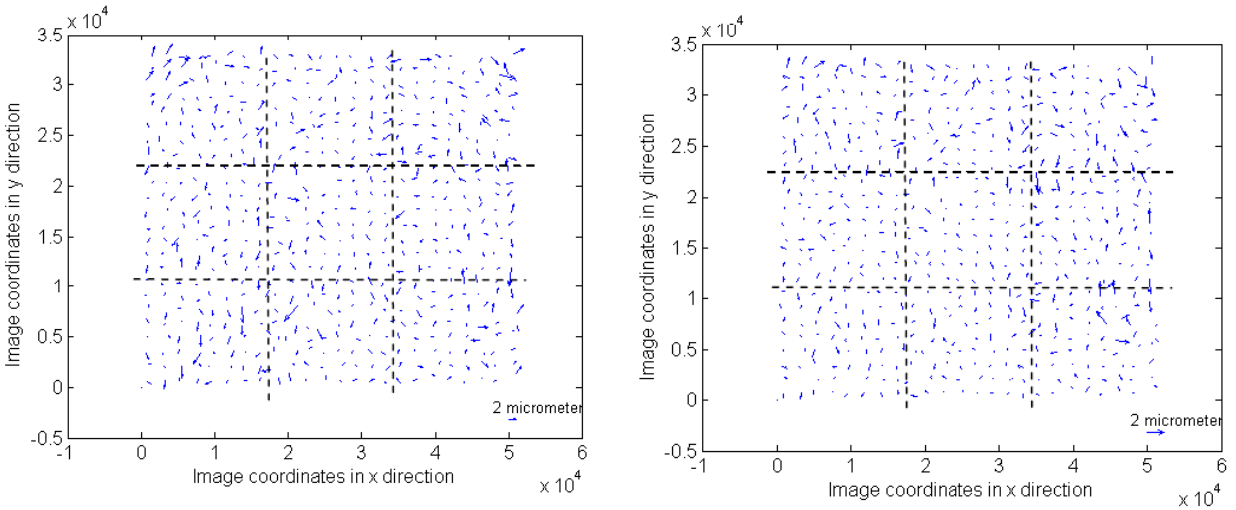


Figure 51. Mean image residuals in 24x24 sub areas. On the left are results of AT without calibration model (low flown). The right diagram shows results of AT with IESSG calibration model (low flown) (adapted from Smith et al. 2007)

blocks of images obtained from 3800m and 1900m altitude respectively.

By identifying and quantifying the systematic residuals in the image space a new calibration method was proposed that re-computes the bundle adjustment based on the analysis of the systematic residual patterns. Only very small systematic patterns could be visually identified in small areas of the image (figure 51).

The IESSG calibration approach for the low flight has been particularly beneficial in improving the RMSE in Z and reducing image residuals. However, the method was relatively less successful at improving the high flown results. (Smith et al. 2007). The results from the processed blocks (table 6 & 7) of images indicate that the UltraCamD camera achieves under certain condition sub pixel accuracy horizontally with slightly worse results for the vertical residuals.

Self Calibration method	Ground control points RMSE (m) of residuals			Ground check points RMSE (m) of residuals			RMSE (μm) of Image residuals	
	X	Y	Z	X	Y	Z*	x	y
No	0.048	0.026	0.031	0.108	0.102	0.278	1.69	1.82
Yes	0.042	0.024	0.020	0.120	0.104	0.248	1.59	1.73
IESSG	0.038	0.022	0.018	0.129	0.098	0.280	1.53	1.62

Table 6 Summary of high flight results
(* removing a dominant residual and the values drop to around 0.1m)

Self Calibration method	Ground control points RMSE (m) of residuals			Ground check points RMSE (m) of residuals			RMSE (μm)of Image residuals	
	X	Y	Z	X	Y	Z	x	y
No	0.054	0.034	0.042	0.042	0.038	0.186*	1.32	1.31
Yes	0.052	0.037	0.033	0.031	0.032	0.093	1.24	1.20
IESSG	0.055	0.038	0.028	0.037	0.037	0.038	1.06	1.00

Table 7 Summary of low flight results
(* includes a dominant residual)

Another noteworthy capability of the UltraCam D camera that makes it suitable for urban modelling is the image sharpness. “Image sharpness” is usually being judged subjectively by the naked eye. However, it can also be measured and assessed quantitatively. Perko (2005) has used the edge response function to obtain edge sharpness values for scanned

film imagery as well as UltraCam D images. Generally the findings are very favourable for the digitally sensed data. Table 8 provides a summary of some of Perko's results with µ denoting microns.

Film System	GSD	Scanned Film	UltraCam
		µ	µ
RMK-TOP, scanned @ 20µm	8.0 cm	1.88	1.10
LMK, scanned @ 14µm	12.5 cm	2.01	1.65
RC20, scanned @ 20µm	12.5 cm	1.55	1.30

Table 8. Edge sharpness values present the width of an edge in pixels when considering distinct transitions from dark to bright areas. The analysis method is by Blonski et al. (2002) and is described by Perko (2005).

Some comparisons were made with imagery flown simultaneously with a film and a digital camera, some show comparisons when the data were collected over the same terrain but at different times. In all cases investigated by Perko, the UltraCamD images were superior in edge sharpness to the scanned film data, with the differences encompassing a factor of ~1.7. The UltraCam D has demonstrated the ability to capture images with as small pixel size as 3cm (Leberl et al. 2005). In order to achieve this ground resolution the aircraft should fly at an altitude of 330 meters above the ground. At such a large scale sufficient forward motion compensation (FMC) and a high frame rate is essential. To obtain stereo overlaps at a pixel size of 3 cm requires a rapid succession of image triggers. At a typical velocity of 75 m per second, a 70% forward overlap with 3 cm pixels results in covering each ground point 3 times. To achieve this coverage the camera needs to be triggered every 0.96 seconds.

Another benefit of using the UltraCamD digital camera for 3D Urban Modelling is the good radiometric quality that enables dense stereo matching. As will be shown later, the proposed workflow for semi-automatic 3D city modelling uses stereo matching techniques to substitute for LiDAR data were these are not available (third study area). The good quality result of the

stereo matching process is related to the higher dynamic range compared to film diapositive and also the benefit of having 'no-grain'. The improved density of match points when applying the same acceptance criterion for the match points is shown in table 9.

Film Camera	GSD	Scanned Film		UltraCam	
		Fwhh (Pixel)	Nr of Points	Fwhh (Pixel)	Nr of Points
RMK-Top, @ 20 µm	8.0 cm	0.52	18% (1838)	0.17	73% (7292)
LMK 2000, @ 14 µm	12.5 cm	0.25	35% (3466)	0.16	70% (6964)
RC20, @20 µm	12.5 cm	0.31	50% (2516)	0.11	63% (3152)

Table 9. Comparison of stereo matching results. "Fwhh" is the "full width at half height" of the histogram of errors. The columns with "Nr. of Points" presents the successful matches found in each image segment. Adapted from Perko (2005)

Digital source data produce more match points than scanned film. We can also study the match accuracy by a method using epipolar relationships between image pairs. Table 9 provides root mean square match errors found from scanned film and from UltraCamD image pairs. It is evident from table 9 that UltraCamD provides on average 100% more matched points with half the RMS compared to film based cameras.

3.4. Airborne Laser Scanner - ALS50

Elevation data from LiDAR systems is an alternative method for capturing ground elevations, preferably in featureless terrain, for the production of digital surface models under conditions that are challenging for aerial photogrammetry. The Airborne Laser Scanner ALS50, from Leica Geosystems is designed for the acquisition of elevation information with the ability to also record intensity information of the reflected signal.

Figure 52 ALS50 system components representing the scanner assembly (lower left), equipment rack (right) and laptop control computer, adapted from ALS50 brochure, 2004, copyright reserved by Hexagon Corp.

The ALS50 is a compact laser-based system designed to capture elevation data using laser range and return signal intensity measurements, which are recorded in-flight along with position and attitude data derived from airborne GPS and inertial subsystems. Further analysis of the operating principles is given in the following section.

3.4.1. Operating and design principles

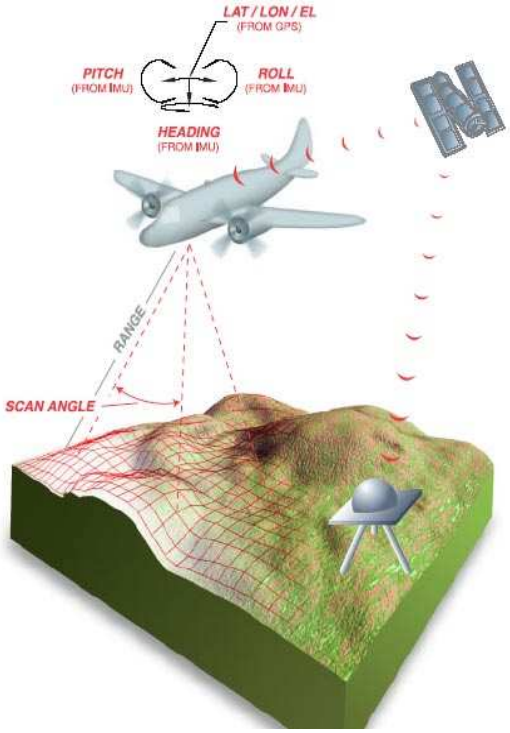


Figure 53 Diagram illustrating the operating principles of the ALS50 (ALS40 brochure, 2002)

The computation of the elevation for each laser pulse can be determined when the location and attitude of the aircraft, the distance-to-ground and the scan angle are accurately known. The operating principle of the airborne laser scanners is illustrated in figure 53.

The main characteristics of the ALS50 include:

• 83 kHz maximum pulse rate and three-return range detection system (1st, 2nd, 3rd and last)

• Automatic adaptive roll compensation

• Up to 75° field of view

The scanner assembly consists of several components for generating, aiming and receiving the laser pulse. The laser pulses are generated initially by a diode-pumped laser transmitter that includes a beam expander/collimator and output optics that focus the laser output to the scanner mechanism. The scanner mechanism consists of a galvanometer that controls the movement of a scan mirror. The scan mirror is controlling the direction of the transmitted pulse, with the direction measured by an accurate optical encoder. The high speed scan mirror is optimized for fields-of-view up to 75° at maximum operating altitudes of 4000 m (ALS50 product description, 2003). The scanner assembly includes also a receiver that collects a sample of the reflected laser pulse and an IMU, mounted in the scanner housing.

The equipment rack contains several components, responsible for raw data measurement and data recording. One of the main components is the System Controller that measures and stores the slant range distance for each laser pulse and records the scan angle along with GPS timing information. In addition the equipment rack includes the aircraft's Position and Orientation System (POS). As in the case of the ADS40 the POS system is responsible for recording the information provided by the GPS receiver and the IMU.

The laptop control computer provides a graphical user interface for system setup, operation and monitoring. In addition it includes the AeroPlan software for in-flight management and for improving the mission planning monitoring process. Furthermore, it contains post-processing software for processing the raw data from the POS system and subsequently calculating the elevation points and their intensity values. The default data output is a compact binary format in WGS84 coordinates (ALS50 product description, 2003).

Another useful feature of the ALS50 is the multi-return intensity digitization capability of up to 3 return reflections from each outbound laser pulse**.** Thus, the sensor has the ability to approximately represent the shape of the waveform by providing 3 amplitudes and 3 ranges, without degrading the maximum pulse rate. With this feature multiple reflected pulses from different levels of a forest canopy are recorded. Multiple reflected laser pulses from a forest

canopy can assist in filtering the data and creating digital terrain models by identifying the tip and bottom of the forest canopy. This can be achieved, since the intensity image can reveal the presence of differing surface materials, by detecting the higher or lower reflectivity of a surface and ultimately distinguish between different surfaces. The raw intensity is initially digitized at the 8-bit level and subsequently an automatic gain control function effectively increases the dynamic range to12bit.

While in operation the ALS50 provides sinusoidal scan patterns of the LiDAR points in a plane nominally orthogonal to the longitudinal axis of the aircraft, with an Instantaneous Field of View equal with 0.33 mrad.

3.4.2. Calibration and overall performance

One of the most common methods for calibrating an ALS sensor is also the least rigorous, and it requires extensive manual measurements. In this method an experienced operator compares profiles of overlapping strips and manually adjusts the misalignment angles until the strips appear to visually fit (Morin and El-Sheimy, 2002). Since this method consists of a visual adjustment, it presents a qualitative method without providing any statistical measure on the quality of the calibration. Additionally this method is time consuming and labour intensive.

According to initial reports, the system produces data after post-processing with a standard deviation in plan of 0.13-0.61m and a vertical accuracy of 0.14-0.36m, depending on the flying height (ALS50 product description, 2003). Figure 54 illustrates the relationship of the vertical and planimetric accuracy for an atmospheric visibility of 23.5 km and flying heights up to 4000m. From figure 54 it is evident that a FOV of 75 degrees (dashed lines) have a

more significant negative impact in the planimetric and vertical accuracy. The planimetric accuracy is represented with the blue lines, while the orange lines represent the vertical accuracy.

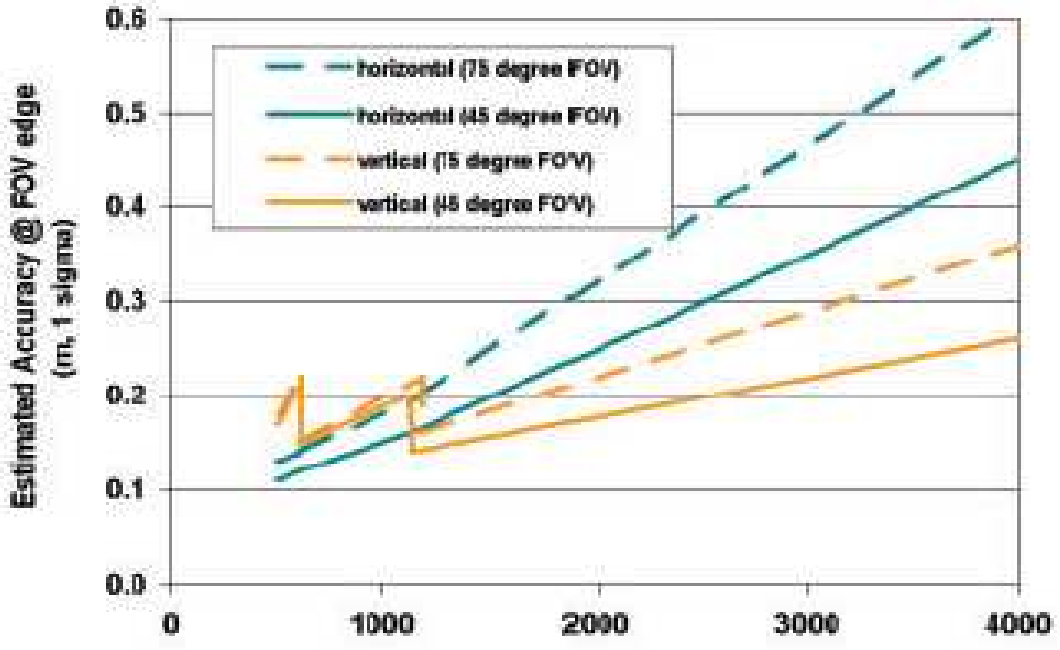


Figure 54 Estimated vertical and planimetric accuracy as a function of the field of view and operating altitude (ALS50 product description, 2003).

The performance of the LiDAR sensor is determined primarily by three factors, which include the maximum and minimum slant range limits, the scan rate and the pulse rate. The maximum slant range for the standard scan mirror configuration is approximately 3800m, whilst the minimum slant range is 500m.

The scan rate is user-selectable, ranging from 0 to 70 Hz in 0.1 Hz increments, but it should be noted that the maximum scan rate degrades as a function of increasing FOV. The maximum pulse rate is determined by the maximum slant-range and the desired number of intensity responses, that the system is recording. These two factors are user-defined and affect the operational pulse rate. For example, a system set to collect data at maximum slant range of 1100m with 2 responses of the reflected intensity stored, would be capable of

operating at a pulse rate of 47 kHz, but a further increase in the intensity responses would decrease the operational pulse rate.

3.5. Airborne Laser Terrain Mapper - Optech ALTM 3033

The ALTM 3033 was designed and distributed by Optech Inc. and was used to collect the LiDAR data for the second test site over London – Bloomsbury area. The sensor, operated by the UK Environment Agency, has been used to cover most of the rural areas and floodplains in the UK. Figure 55 shows the ALTM sensor with the different components.

Figure 55 The latest ALTM sensor including the navigation system, onboard storage and sensor head. (Adapted from Optech, ALTM brochure 2008. Copyrights reserved by Optech Inc.)

3.5.1. Operating and design principles

The ALTM sensor was introduced in 2001 as a replacement for the ALTM 1225. It is a 33 kHz laser scanner usually flown at 3km altitude. It consists of 5 main units, the pulsed, solid state laser scanner, range finder, the IMU, GPS and the onboard storage rack. It can capture up to 4 pulse returns using a multi-pulse operation, with intensity and waveform capture

provided the return pulses are separated in time. The waveform digitizer is another important feature that enables advanced analysis to be carried out. The waveform digitizer is responsible for capturing the whole waveform, capable of recording up to 70,000 waveforms per second. The advantage of the full waveform digitization is that additional details and features can be detected even within the same pulse footprint. In addition waveform information can be valuable for automatic classification.

The ALTM sensor has been already integrated with a variety of different sensors including

- Rollei AIC, 22 megapixel digital camera
- Applanix DSS 322, 22.2 megapixel digital camera(figure 56)
- Vexcel UltraCam digital camera
- Intergraph DMC digital camera
- Hyperspectral – Itres CASI 150 digital sensor

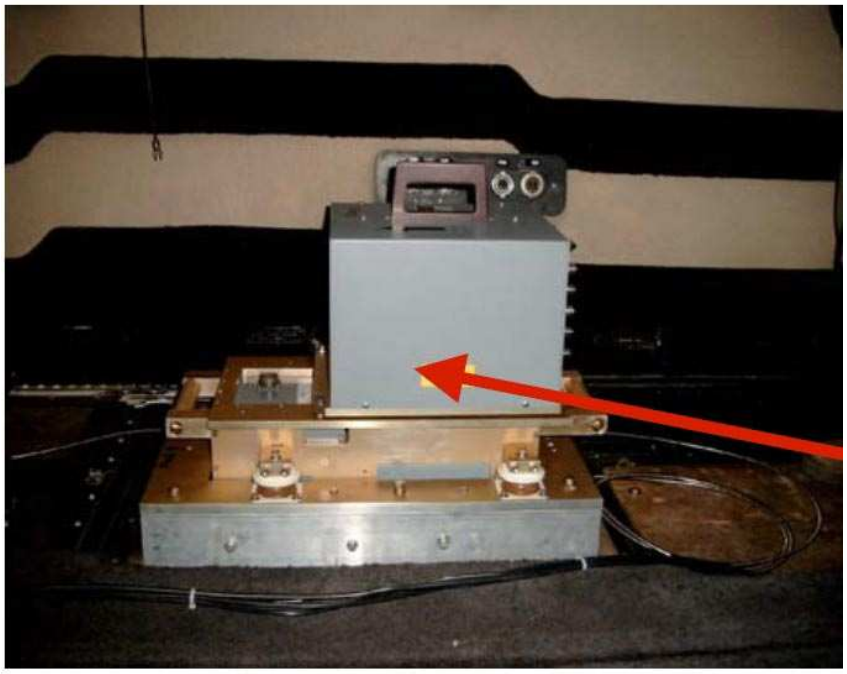


Figure 56 The integrated ALTM sensor with the Applanix DSS camera (adapted from ALTM brochure, 2008)

The ALTM LiDAR sensor has an operating altitude that ranges from 80 – 3,500 m. The vertical accuracy as stated by the manufacturer (ALTM brochure, 2008) is less than 15cm at

1.2km altitude and less than 35cm at 3km. It provides up to 4 range measurements for each pulse with a 12-bit dynamic range and a 1cm vertical discrimination. The scan angle is variable ranging from 0 – 25° with a 1° increment and achieves a swath width of approximately 3000m from a 3.5km altitude.

The ALTM sensor is integrated with an Applanix POSAV 510 position orientation system, including an internal 12 channel, dual frequency 2Hz GPS receiver. The sensor comes with onboard storage facilities that consists several removable hard drives that can record a minimum of 7 hours continuous data.

3.6. Error sources for airborne laser scanners

There are many sources of error that affect the quality of the laser scanning data. The resulting errors are related to the laser ranging computation, the scanning system, topography, the atmosphere, positioning and navigation systems, and system integration factors (Alharthy et al. 2004). The three parameters that are directly related with the airborne laser scanner that could introduce errors are the following.

- sensor position
- distance to reflecting object
- viewing direction to reflecting object

The position of the sensor in relation to the GPS antenna can be determine ideally with an accuracy of 0.05m and is directly related to the quality of the GPS data (Katzenbeisser, 2003). Any disruption of the spacecraft signals can incorporate errors in the determination of the position. Therefore, in most cases Differential GPS is used, acquiring data at a time interval of 1Hz, using a dual frequency receiver that records code and phase information for post processing. Another major requirement is a good distribution of the GPS satellite constellation and the reference GPS station positioned within 25km away from the rover

receiver onboard the aircraft. Ultimately the precision of the Differential GPS is a function of the accuracy in the solution of the ambiguities and from the various positioning errors.

Apart from the position of the sensor, the accuracy of the slant range distance directly affects the quality of the calculated elevation values. The distance is measured with a time of flight method by counting the number of cycles (*n*) of an oscillator operating at a frequency *f* and then converted into distance using the following formula.

$$t = \frac{n}{f} + \Delta t \quad \text{or} \quad s = \frac{n}{2f} \cdot c_a + \Delta s \qquad \text{t =(the time of flight), s = (slant range distance)} \tag{3.1}$$

Where *f* is the frequency of the oscillator, Δt and Δs compensate for delays and optical paths within the sensor and c_a is the speed of light within the atmosphere. From the above equation the frequency of the oscillator f can introduce significant errors even with minor deviations from the nominal frequency.

Another important parameter is the speed of light c_a within the atmosphere, because it should take into account the deviations in density (or pressure), humidity and temperature. The difference in pressure is critical, especially when the data are acquired at varied altitude.

The accuracy of the viewing direction to the reflecting object is also affecting the quality of the estimated elevation values. The viewing direction or more precisely the beam direction is a function of the sensor orientation, which is given by the IMU system, and the beam deflection that describes the direction of the laser beam with respect to the laser generating device.

Since the beam direction is directly related to the sensor orientation, the quality of the IMU data are influencing the overall accuracy. The IMU devices include gyros and accelerometers that are not free of errors and thus, simple integration will lead to a drift of the results (Katzenbeisser, 2003). The attainable accuracy in determining the orientation is greatly improved by integrating DGPS position, movement direction and speed derived

information. With operational IMU devices used in airborne sensors the roll and pitch can be determined with accuracies in the range of 0.004 deg to 0.02 deg.

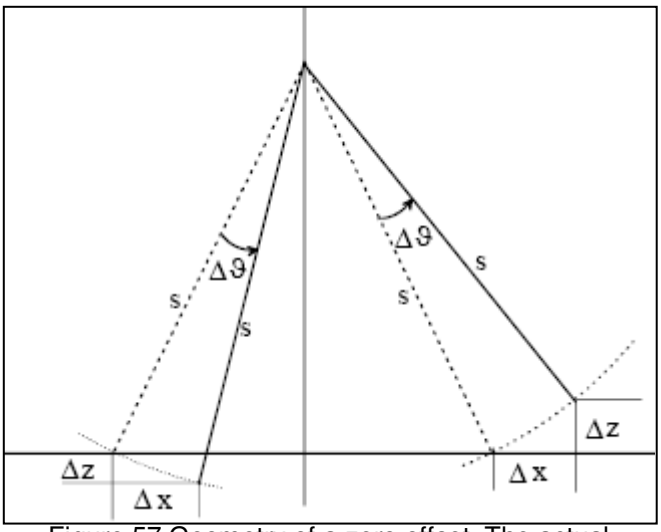


Figure 57 Geometry of a zero offset. The actual laser beam (dotted line) and calculated direction (solid line) (Katzenbeisser, 2003).

In the case of the ALS50 the beam deflection is achieved by an oscillating mirror. The equation used for calculating the instantaneous look angle is given by

$$\vartheta(t) = \Theta/2 \cdot \sin(\omega t) \qquad (3.2)$$

Where (Θ/2) is the maximum angular position. The first error that might be introduced at this stage is known as the zero-offset (Δ□), possibly related to mechanical miss-alignments of mirror and encoder, or a zero-shift within the Analog/Digital converter.

A second error presented at this stage is a scaling factor, introduced by a false gain-control within the A/D converter or by the encoder itself. Therefore, the instantaneous angle of view now becomes

$$\vartheta_D(t) = \vartheta(t) \cdot (1+\varepsilon) \qquad (3.3)$$

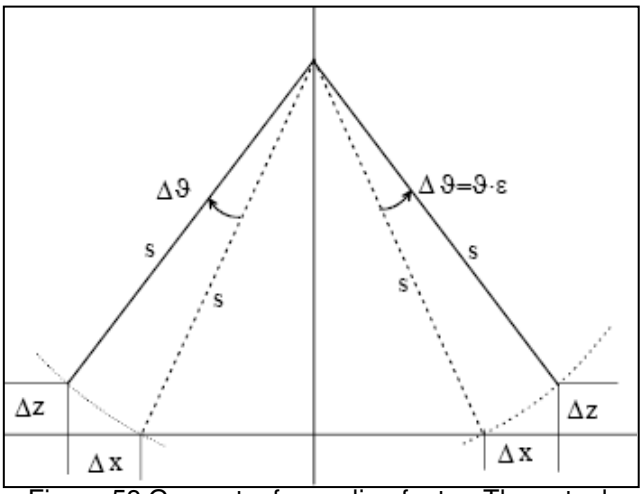


Figure 58 Geometry for scaling factor. The actual laser beam (dotted line) and its calculated direction (solid line) for an offset (Katzenbeisser, 2003).

Figure 58 shows the geometry of the beams in the case of a false scaling in the viewing angle, which is depicted here as a positive error. The false scaling factor results in a scaling of the swath width which can be either wider or narrower and results in miscalculated

elevation values. These errors can only be corrected by calibration, when they are systematic in nature, or in other words if they are introduced because of a mechanical movement. A possible solution for a fibre scanner can be given, in which the individual beam direction is defined by the number of the fibres in use. As these fibres are tightly coupled the deflection cannot vary over time.

As it was mentioned previously, the beam direction is a function of the beam deflection angle with respect to the aircraft orientation. For this reason the position of the IMU must be precisely known relative to the beam deflection device in order for the beam direction to be determined. Hence an alignment between the IMU and the beam deflection device is established by an initial calibration in the airborne laser scanners, as illustrated in figure 59. Nevertheless the distance between the two devices may become unstable, as indicated by the red dashed line, because of bending forces during mounting the sensor into the aircraft or from thermal effects that result in the deformation of the carrying plate. A thorough description of the deformation effect on the carrying plate is described in Katzenbeisser, (2003) whereby a carrying plate of distance d = 300mm and a deformation a = 1mm are assumed, which result in a 3m displacement on the ground from a measuring distance of 1000m.

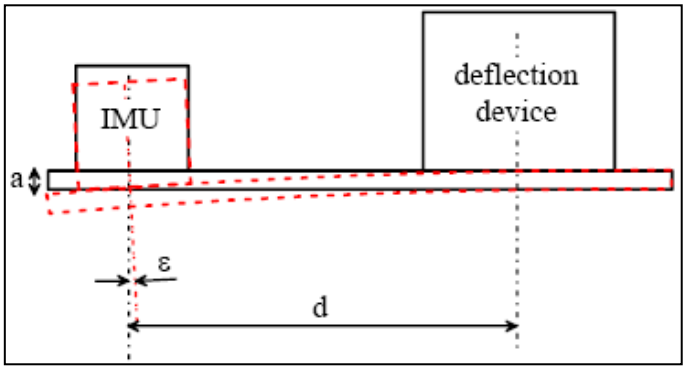


Figure 59 Mounting of IMU and deflection device on a carrying plate (Katzenbeisser, 2003)

The main problem introduced when calibrating the sensor, for correcting the errors described so far, is the requirement that the individual measurements from the different elements of the sensor are taken at the same instant of time. The time requirement is critical for associating the individual errors together and applying the necessary corrections. The precise time requirements between two devices are even more difficult, if we consider that they operate at different frequencies. For example, establishing the time instances between the IMU

(operating at 200Hz) and the deflection device (20-80Hz), can introduce errors especially when bumpy conditions are present during flight.

Apart from the errors introduced from the mechanical and electronic components of the airborne laser scanners, topography and surface features can also introduce erroneous effects in the LiDAR point cloud. The large footprint of the laser beam on the ground is a limiting factor in the cases of steep slopes or rough terrain, which can introduce erroneous effects and elevation errors. These effects are related to the operating altitude and the beam direction, combined with the slope angle, but most important the flight direction with relation to the slope. These errors will be more severe when the flight direction is parallel to the slope contour, but is minimized when the flight direction is parallel to the direction of the gradient. Tree canopies and other features introduce systematic errors, since they don't represent the actual terrain surface.

In addition errors can be introduced when the laser pulse hits the side of a vertical object, which yields a misleading profile. Furthermore gaps in the data might be introduced if the return pulse cannot be detected due to the weak response. Specular reflections may also produce regions of missing data, as well as occluded regions for high features. In most cases these gaps can be corrected by crossed flight paths. These errors are not correctable in a rigorous sense, but some of their values can be minimized with careful mission planning and operation (Alharthy et al. 2004).

Daveport et al. 2004 performed an extensive accuracy assessment of the Airborne Laser Terrain Mapper. The study was conducted to analyze the fine temporal detail and elevation accuracy. Daveport et al. 2004 analyzed 37 min of laser altimetry data acquired over a very flat 120 m 120 m area, and compared it with 408 ground-acquired differential GPS measurements of the surface. They concluded that on average a point measured using the laser altimetry system is displaced 36 cm from the mean in each dimension, in 78% of cases is 40 cm or closer, and no more than 70 cm displacement.

In comparing heights between swathes, if the area is being measured in two swathes, the difference in height due to instrument error is on average 6.3–6.4 cm, seemingly due to the DGPS-induced drift (Daveport et al. 2004). The maximum possible systematic difference between swathes depends on GPS optimization. In addition to the systematic swathe–swathe error, there is a random error contribution to each measured point, due to the measurement of the laser return time and INS/DGPS system random noise. This imparts a normal distribution noise with a standard deviation of 4.1 cm, or 3.5 cm for points closer than 50 ms. This could be eliminated on large-scale data by collecting points into, say 3 m 3 m cells, and using the mean of the ten or so points that would fall within each cell using typical acquisition parameters

3.7. The significance of the selected data collection systems in 3D city modelling

This section summarizes the main reasons for the selection of the previously mentioned data collection system and their advantages in automating building reconstruction.

The ADS40 represents one of the next generation digital sensors and is well suited for building reconstruction, mainly because of the three line pushbroom acquisition geometry and the form of the relief displacement that minimizes occluded regions. The minimization of occluded regions in the optical data is critical, especially in the proposed workflow, because the extracted stereo matched points are subsequently used for adjusting the building outline. Therefore, having an adequate number of conjugate points, representing most of the building façades, will immediately affect the effectiveness of the proposed method. This is the key point of using data from the ADS40 since the pushbroom scanning mechanism results in one dimensional relief displacement instead of two dimensional, as in the case of frame cameras. The three look angles (backward, nadir and forward) can provide multiple stereo pairs of imagery for feature extraction with different base to height ratios, but most

importantly it can provide alternative views in the case some features are not observable from a given stereo pair. In addition, ADS40 can acquire multispectral information that can be very useful for vegetation discrimination using classification methods. The ability to obtain multispectral information is a major advantage over conventional aerial mapping cameras and can be used to filter the LiDAR point cloud before the process of building reconstruction.

During the building reconstruction process, and depending upon the complexity of the buildings, the extraction of roof details is affected by the density of the LiDAR point cloud. The ALS50 has the advantage of using a maximum pulse rate of 83 kHz, which depending on the flight altitude, can yield very dense point clouds. Since the proposed method uses the initial building outline, derived from LiDAR data, to filter the stereo matched points, the planimetric accuracy of the initial building outline is of major concern. Therefore with a planimetric accuracy in the range of 0.13-0.61m and a vertical accuracy of 0.14-0.36m, the ALS50 can produce good approximations for the initial building outlines and furthermore, accurate elevations for the reconstructed roof planes.

For the second test site in London Bloomsbury instead of using state of the art LiDAR and digital imagery, it was thought that the designed workflow should be tested against more conventional or traditional datasets. Hence the selection of one of the most widely used film based cameras (RC30). The use of the RC30 would introduce all the disadvantages of a scanned film dispositive which was especially important in evaluating the deterioration of the feature extraction quality.

The LiDAR device selected was used extensively by the Environment Agency to capture data across the UK so LiDAR data from that ALTM sensor have been utilised by several public and commercial organisation. In contrast though to the high density data collected from the ALS50, the ALTM data are down sampled and distributed with a 1m density. Hence presenting a particular challenge for the designed algorithms, especially in an urban environment like London with complicated building structures.

The data selected for the third test site reflect the willingness to 'breakthrough' the dependence on LiDAR data. Avoiding the use of LiDAR in the designed workflow would introduce additional advantages, considering the limited availability and cost of LiDAR data. Therefore the decision to adapt the workflow in cases where LiDAR data are not available demanded the use of state of the art digital sensors capable of substituting the LiDAR data. The selection of the UltraCam D camera was based on that reason since it brings a number of qualities for 3D urban modelling, including:

- Digital sensor based on a frame geometry able to collect images with as high ground resolution as 3cm.
- Good quality stereo matching, compared to frame cameras, resulting in dense point clouds, potentially able to substitute LiDAR. This quality is due to the higher dynamic range compared to film dispositive and also the benefit of having 'no-grain'.
- High image sharpness, compared to film photography, that improves the efficiency of edge operators and subsequently the feature extraction process.

It should be noted that although the UltraCam D imagery has a very good radiometric quality and therefore it is expected to perform well in the stereo matching process, recent research in evaluating digital sensors (Jacobsen, 2008) indicates that the DMC sensor has an even better radiometric quality. Unfortunately DMC imagery could not be obtained for any of the test sites.

4. STUDY AREAS AND AVAILABLE DATASETS

This chapter describes the three different test sites and characteristics concerning urban modelling as well as introducing the available datasets collected from the sensors described in the previous chapter.

4.1. Study area (Heerbrugg, Switzerland)

The study area is situated 5 miles west of Heerbrugg-Switzerland (figure 60), representing Leica's headquarter facilities as well as the surrounding region.

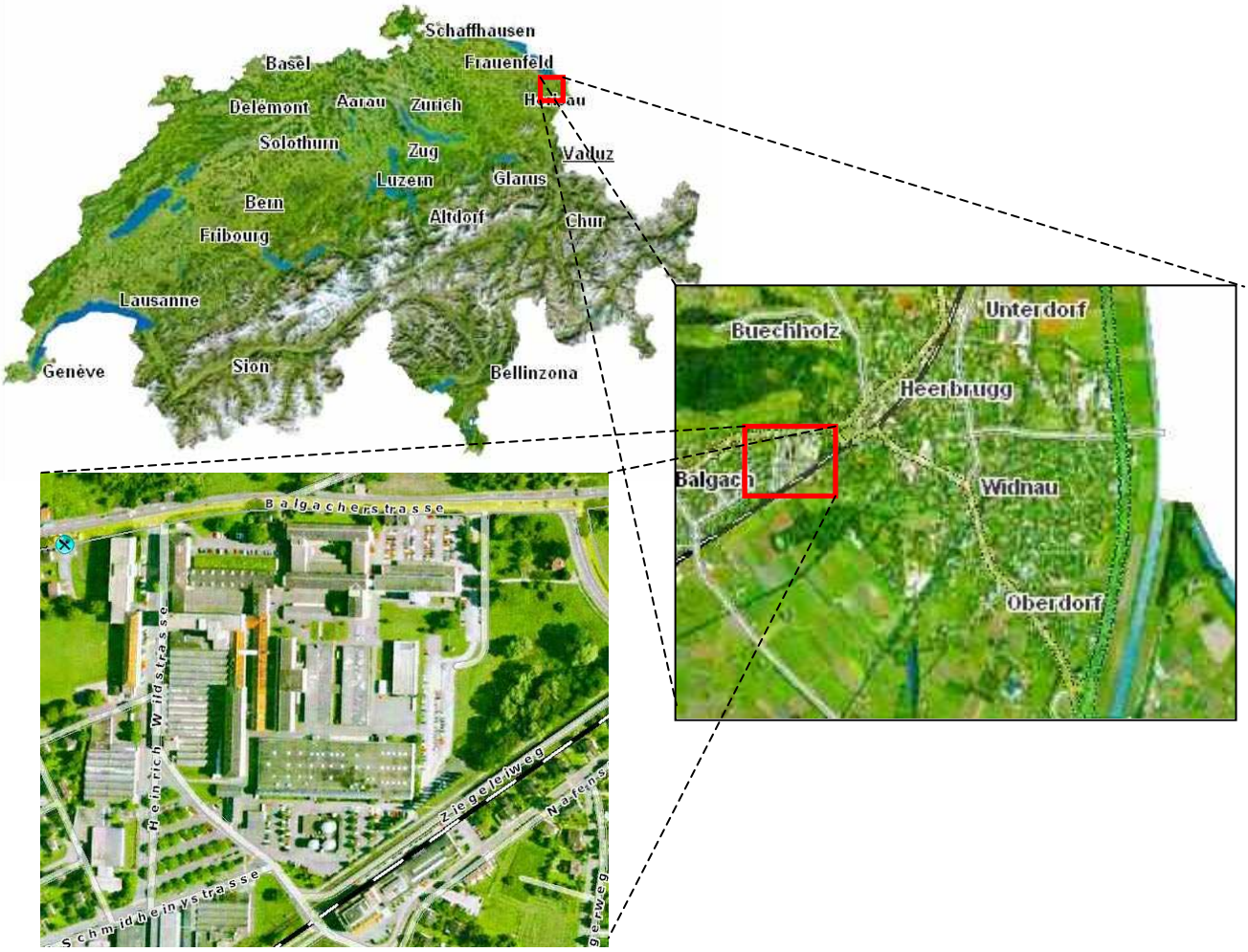


Figure 60 Study area representing Leica's headquarter facilities and the surrounding region at Heerbrugg, Switzerland. (adapted from ©search.ch / Endoxon AG, TeleAtlas)

The study area situated approximately 5 miles west of the centre of Heerbrugg is mostly large industrial buildings with complex roof structures which are part of Leica's facilities. These buildings provide an excellent opportunity to test the efficiency of the proposed method for building reconstruction on large industrial buildings. In addition, the study region represents two semi-urban regions, west and southeast of the main facilities, with smaller houses having a gable or hip type roofs. These semi-urban regions are very useful for evaluating the influence of the density of the LiDAR point cloud in the building reconstruction process. Various tree canopies of different shape and height can be identified in the study area, which introduce the necessity for designing a robust and reliable method for vegetation filtering from the LiDAR data.

4.1.1. Data Provided

The data and the supporting material were provided by Leica Geosystems, Heerbrugg for the needs of the specific project. The available dataset included, an oriented stereo-model acquired with Leica's Airborne Digital Sensor (ADS40) and LiDAR data for the study region obtained with Leica's Airborne Laser Scanner (ALS50).

4.1.1.1. ADS40 digital strip imagery

The airborne optical data included three bands of strip imagery acquired from the same flight

Figure 61 Panchromatic band from the strip imagery acquired with a forward look angle of 28^{o} from the ADS40. Study area is highlighted in the red rectangle.

path. Specifically the data consisted of a green band, obtained from the nadir looking view, one panchromatic band acquired from the 14° backward looking view and another panchromatic band from the 28° forward view. Figure 61 illustrates the last band.

The provided data included the raw Level 0 support files and the precision Level 1 support and image files, for the given bands, that formed high resolution stereo viewable images. The Level 0 imagery provides the initial raw data calculated by the ground post-processor and is required in order to recover the instantaneous position and attitude associated with any pixel in the Level 1 image and subsequently used for the aerial triangulation.

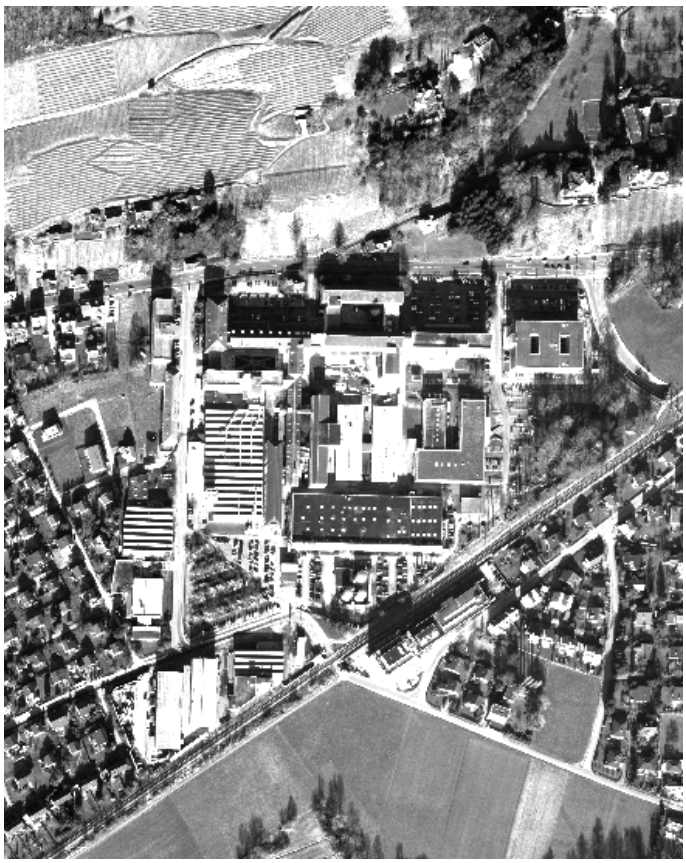


Figure 62 Study area represented by the forward looking panchromatic band. Subset of the entire scene with reference to figure 61.

The green nadir looking view and forward panchromatic bands have an overlap of 84% and provided full coverage over the study region indicated in figure 62. In contrast the overlap between the green and backward looking panchromatic band was 92%, without providing any coverage over the study region. Therefore, only the first stereo model was used for the subsequent processing steps. The accompanying data included the calibration and orientation files for the sensor, which are essential for establishing the relationship from the Level 0 to Level 1 images.

The data were acquired from a flight altitude of approximately 2430m, yielding images with photo scale equal to 1:38900. The entire scene represented in figure 64 is covering an area of 2558 hectares, bounded by co-ordinates X= 544526.50m, Y= 5244511.84m for the upper left corner and X= 547797.64m, Y=5252845.91m for the lower right corner. The co-ordinate

system and projection parameters for the study area are based on the UTM system which will be described in section 4.1.1.3 in detail. The sensor had a lens with a focal length of 64.7mm. Each detector on the CCD has a size of 6.5 μm, which resulted in a Ground Sample Distance (GSD) equal to 0.25m. The digital images had approximately 33690 columns and 12520 rows. Note that the amount of columns and rows was slightly altered between the three images. In addition the number of rows exceeds the nominal number of CCD detectors due to the rectification process which is evident in figure 61 from the irregular outline and black regions along the edges of the image.

4.1.1.2. Airborne LiDAR data

The LiDAR data were acquired with the ALS50 and consist of two separate point clouds acquired from different flight paths, which intersected at approximately right angles over the study area. The LiDAR point cloud acquired from the first and second flight paths cover an area of approximately 803170 m^2, having a density of 2-3 points per square metre. Figures 63 & 64 illustrate a colour coded, shaded relief map of the two flight paths. The visualizations were generated by triangulating the raw point clouds and subsequently applying ray tracing methods for creating the shading effects.

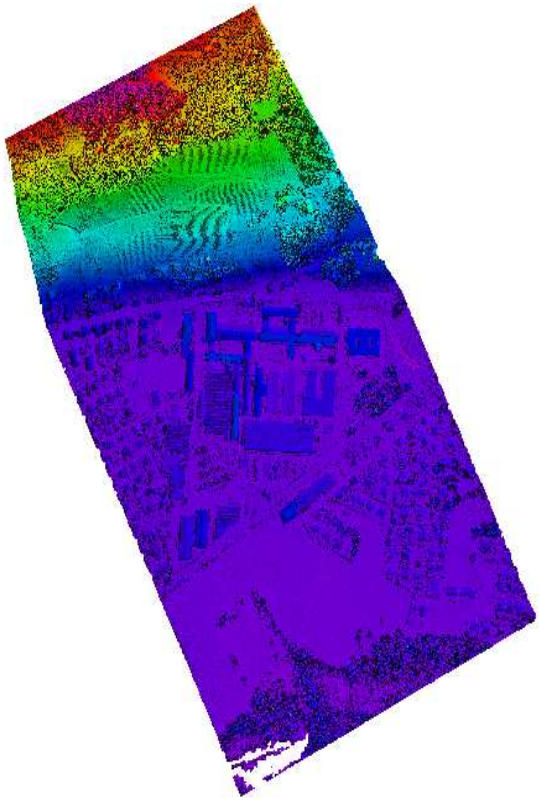


Figure 63 Color coded, shaded relief image of the LIDAR point cloud obtained from the first flight path (Line ID: 050114_141115)

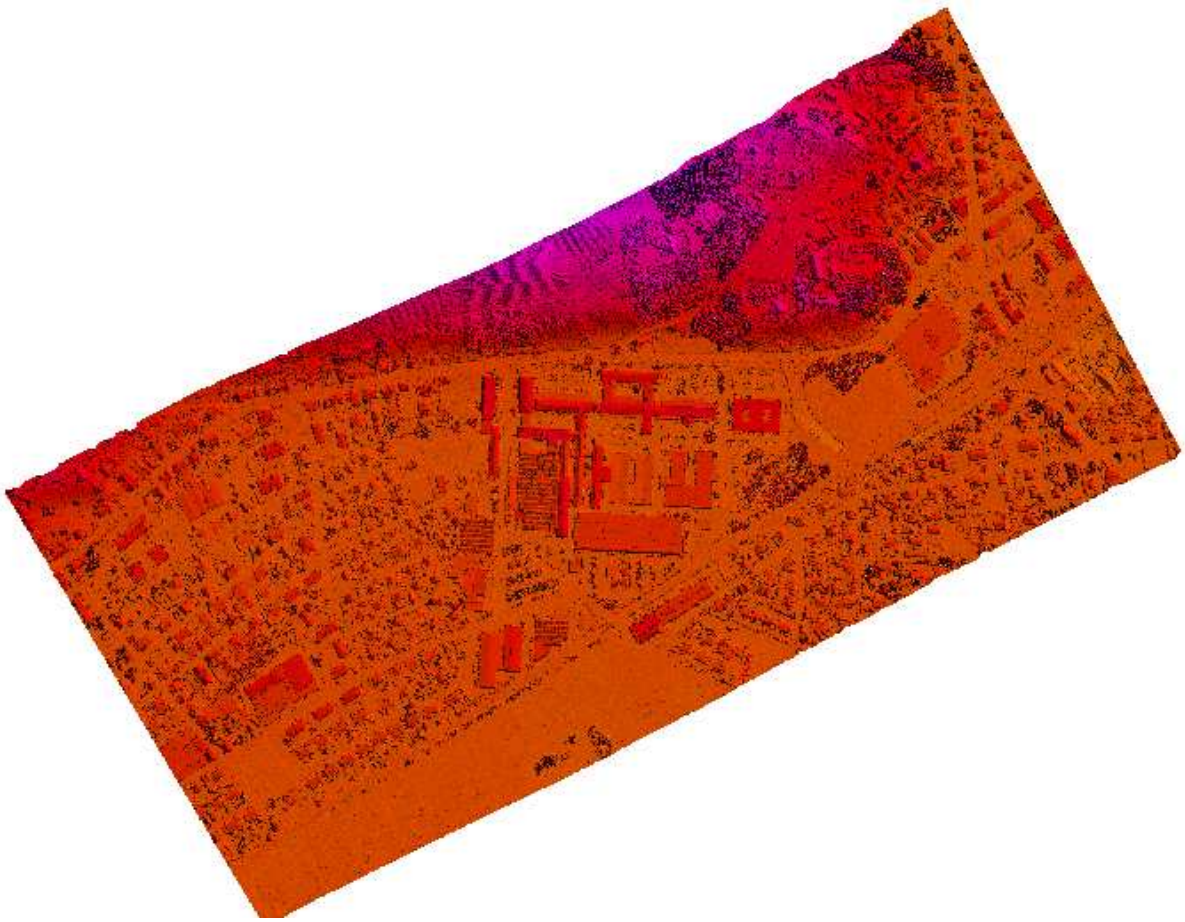


Figure 64 Color coded, shaded relief image of the LiDAR point cloud obtained from the second flight path (Line ID: 050114_141444)

The first point cloud (line ID: 050114_141115) was obtained by the sensor from an operating altitude of 1207m (770m mean AGL), using a Field of View (FoV) equal to 45.798°. The aircraft was flying at a speed of 81 knots, while the oscillating mirror was operating at a scan rate set at 34.1 Hz and the laser transmitter operating at a pulse rate of 84.9 kHz.

During the second flight path (line ID: 050114_141444) the operating altitude was almost identical at 1211m, while the FoV was slightly decreased to 45.07°. The scan rate of the oscillating mirror and the pulse rate remained the same as in the previous flight path. In both cases the sensor was adjusted to detect and store the first and second signal return of the laser pulse. A total of 4.8 million points were collected from both flight paths with a density of two to three points per square meter. Table 10 summarizes the most important parameters from the specific flight mission.

Flight Path ID	050114_141115	050114_141444
Sensor	ALS50	ALS50
Actual FOV	45.798°	45.07°
Scan Rate	34.1 Hz	34.1 Hz
Pulse rate	84.9 kHz	84.9 kHz
Altitude	1207m	1211m
Heading of flight path	151.217° (Clockwise from North)	238.718 ° (Clockwise from North)
Speed	81 knots	85.6 knots
Swath width	628m	631

Table 10 Flight parameters during the acquisition of the LiDAR data over Heerbrugg- Szitzerland

A vertical accuracy assessment was also conducted by Leica Geosystems, SAT Group, using 34 GCPs. The GCPs are located in the study area which corresponds to the overlapping regions between the two LiDAR datasets. The location of the GCPs used for the vertical evaluation is depicted in figure 65.

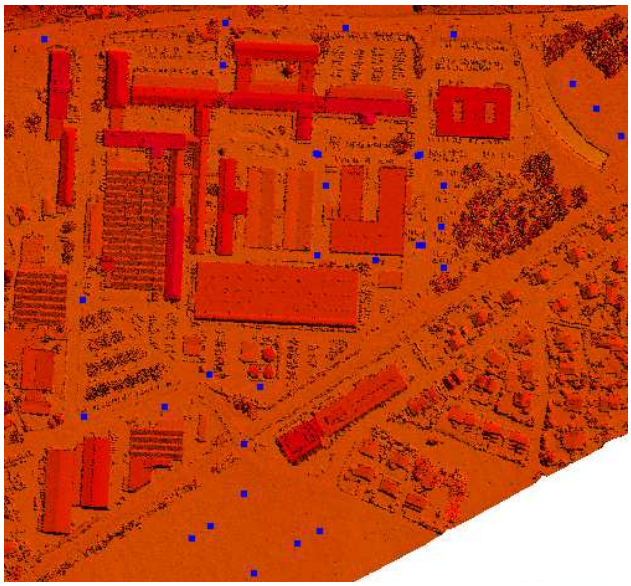


Figure 65 Location of the GCPs (blue points) used for the vertical accuracy assessment

Vertical accuracy assessment (34 GCP	
Average dz	+0.004m
Minimum dz	-0.130m
Maximum dz	+0.160m
Average magnitude	0.064m
Root mean square	0.075m
Std deviation	0.077m

Table 11 Results from the vertical evaluation of the LIDAR point cloud

The assessment indicates a good vertical

accuracy with the standard deviation in the order of 7.7 cm (Table 11), while the maximum vertical difference observed was in the order of 16cm.

4.1.1.3. Co-ordinate system for the study area

The co-ordinate system used for the specific project was based on the WGS84 system with UTM projection. The specific co-ordinate system was defined mainly because of the integrated GPS system onboard the aircraft that directly measures on this specific datum and was subsequently used for the entire post-processing of the optical and LiDAR data. In addition GCPs used for direct georeferencing and the aerial triangulation are usually measured with geodetic GPS antennas in the same co-ordinate system. Therefore instead of applying geometric transformations between different datums and spheroids it is more convenient and reliable to sustain this co-ordinate system for the entire processing stage.

The study area is located at the north zone 32 of the UTM system. This zone covers regions from longitudes 6^{o} – 12^{o} east of the prime meridian. The characteristics of the co-ordinate system for the specific project are given in table 12.

Co-ordinate system parameters	
Projection system	Universal Transverse Mercator
Zone	32, north of the equator
Vertical spheroid	World Geodetic System 1984
Datum	World Geodetic System 1984

Table 12 Parameters of the co-ordinate system used for the study area

4.2. Study area (London – Bloomsbury)

Bloomsbury area is located in central London (figure 66) encompassing very popular landmarks such as the British Library and the Kings Cross, St. Pancras train stations.

Figure 66. Overview of the second study area located in London-Bloomsbury. Adapted from Google Earth ™. Copyright Google Corporation

The area presents one of the most challenging environments when it comes to urban modelling due to the complexity of building structures and the varied architectural design. Building designs range from modern block skyscrapers and industrial buildings to gothic style architecture. Figure 67 illustrates the complexity of some of the building sites with pipes and ventilation equipment covering most of the roof space. It was evident that in many cases the density of the LiDAR data would not be able to sufficiently represent structures of less than 1m^2 and generalization of the roof structures would have to take place. The introduction of vegetation and individual trees between buildings will potentially introduce significant problems during the tree segmentation and building hypothesis stage.

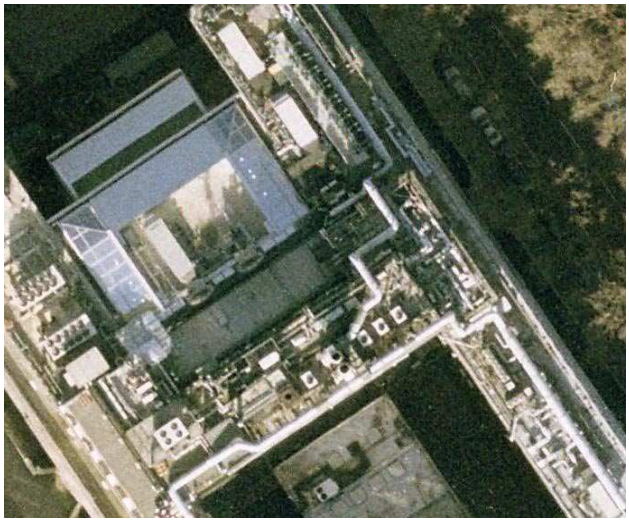


Figure 67. Example of complex building roof geometry present in the second study area

4.2.1. Data Provided

The data for this test site were provided by Ordnance Survey and the Environment agency for the purpose of evaluating the designed workflow for semi-automated building reconstruction. This test site has a particular significance since it will enable the performance evaluation of the proposed workflow under very difficult circumstances. That is, trying to model a very complicated urban environment with more widely accessible or traditional data sources that are primarily used by mapping professionals.

4.2.1.1. Leica RC30 aerial images

This test site consists of a single strip of three consecutive aerial photographs acquired by the Ordnance Survey in 2000 using a Leica RC30 frame camera with the 30/4 NAT-S lens cone having a calibrated focal length of 304.791mm. A camera calibration certificate was also provided but no in-flight GPS or IMU data were collected during the aerial survey. The aerial photographs (figure 68) are only a subset from an aerial survey with a 60% overlap. The images were acquired from an altitude of 1800m, yielding an image scale of 1:6,000. The film negative was subsequently scanned at 13.6μm creating true-colour imagery with 16861x16861 pixels and a GSD of 8cm. The true colour images contain only three bands (R,G,B) with an 8bit dynamic range and a file size of approximately 350MB each. The triplet of images was collected during a leaf-on season over central London covering an area of 2sq. km

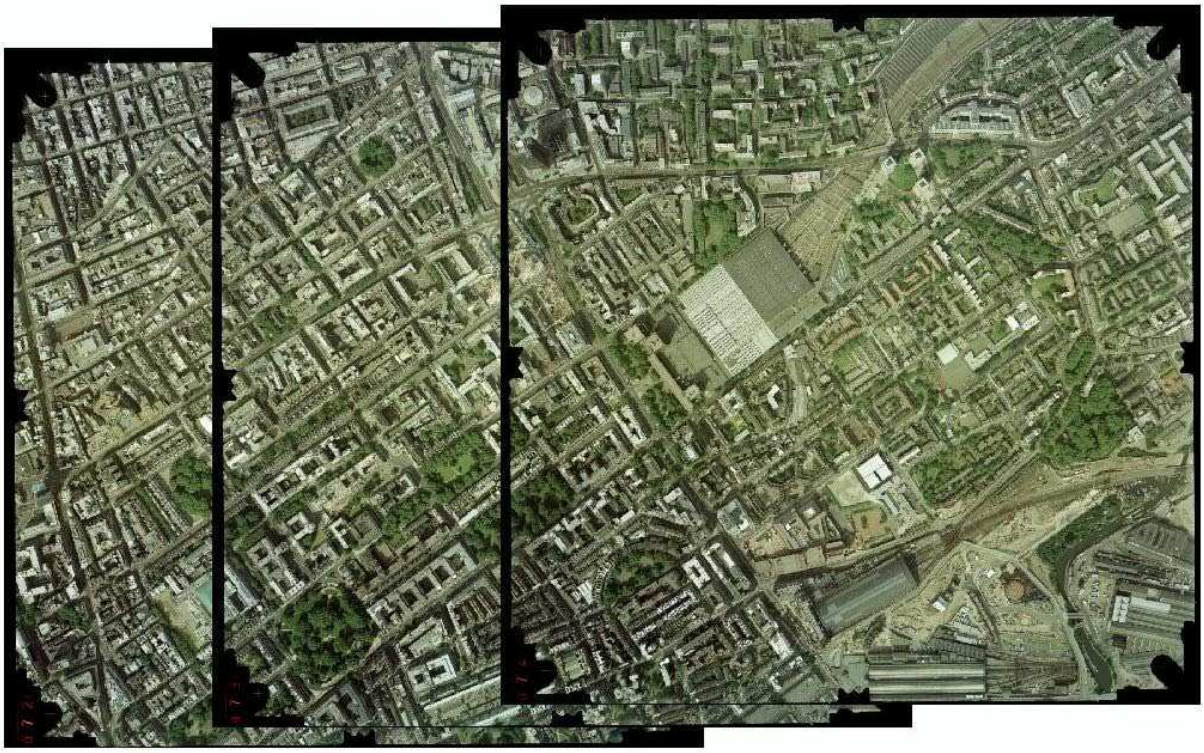


Figure 68. The available triplet of images for the 2nd study area over London-Bloomsbury. (Crown Copyright)

The scanned images don't appear to have any scratches or any other artefacts. Nevertheless there seems to be a few overexposed areas and small variations in the overall brightness are evident. In addition despite the high ground resolution the feature identification is affected by the 'grain effect' as illustrated in figure 69 and is expected to have an adverse effect during the feature extraction.

Figure 69 Subset of the original aerial photographs over London – Bloomsbury indicating the 'grain effect' (Crown Copyright)

4.2.1.2. ALTM 3033 LiDAR data

The LiDAR data used for the second study area were acquired by the Environment Agency using their Airborne Laser Terrain Mapper - Optech ALTM 3033 in 2003 during the leaf-on season. Due to the three year difference there have been some temporal changes between the aerial photography and the LiDAR data. The LiDAR point cloud represents Bloomsbury area in central London bounded by co-ordinates (lower left: 529174, 181500, upper right:

530199, 182525) in the British National Grid (OSGB36). The point cloud has been post-processed by the UK Environment Agency and was delivered with a 1m by 1m sampling resolution covering an area of approximately 2 km^2 (1700 rows and 1100 columns consist the derived regular grid DEM). The post processed raster surface resulting from the raw point cloud is illustrated in figure 70.

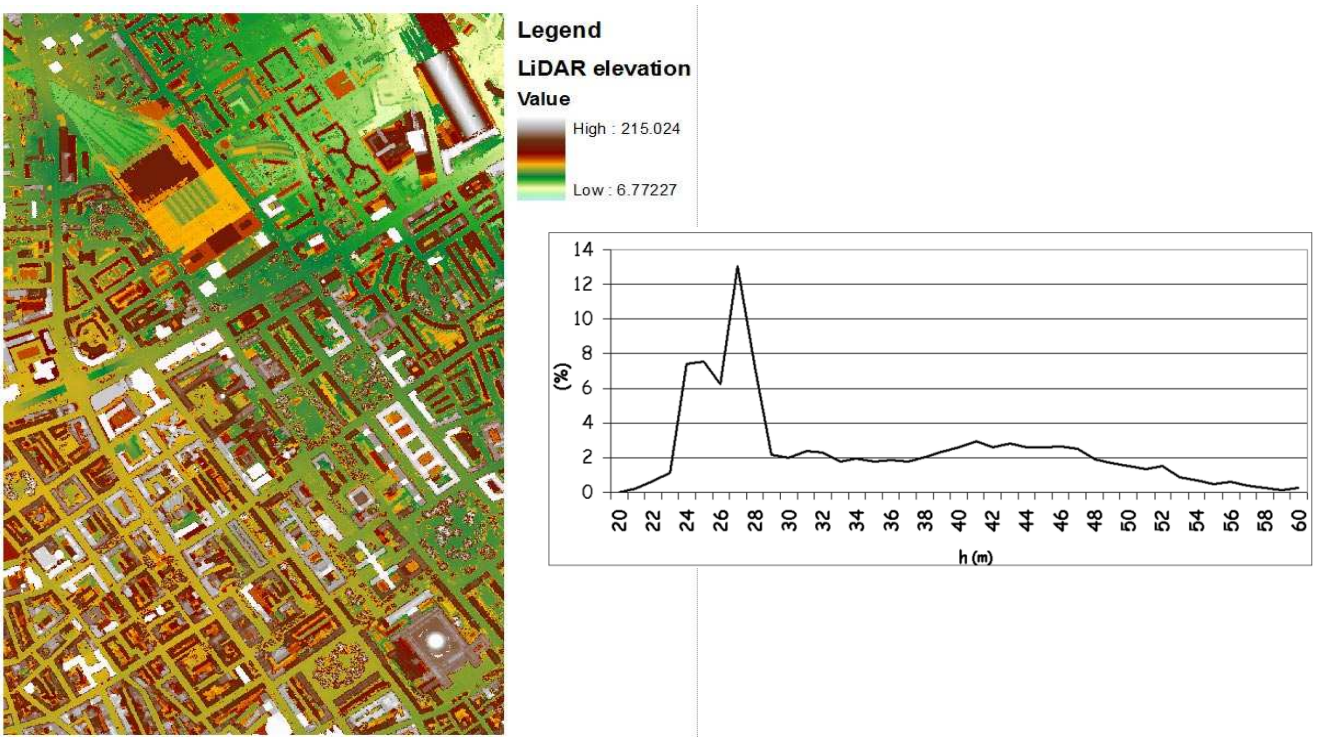


Figure 70 The LiDAR data represented as a color coded raster surface and the elevation frequency histogram on the right using orthometric heights based on the OSGB36 vertical datum

As illustrated in figure 70 the elevation frequency histogram indicates that the majority of DEM cells (98.94%) present elevation in the range 20 to 60 m. It is also important to note that the post processed LiDAR data from the UK Environment Agency don't have any additional return pulses; instead they represent only the first pulse return, thus introducing additional difficulties for effective tree segmentation. Although this means the proposed workflow will need to take this into account, this situation represents the majority of distributed data. Hence a more generic workflow able to address this issue can have a wider implementation in the research and commercial community.

Upon further investigation of the LiDAR data over London – Bloomsbury it's evident that the post processing and down sampling of the raw point cloud to a regular 1x1m raster surface will impact the level of detail when it comes to building detection and 3D reconstruction of roof details. Note, that in the specific study area there are buildings with very complex roof details as illustrated in figure 67. Figure 71 presents a 3D perspective view of a subset area from the LiDAR surface.

Figure 71 3D perspective scene from the LiDAR data representing the overall density of the surface and highlighting the limitations for adequate representation of roof details.

Upon further visual inspection it is evident that the LiDAR dataset by itself will be unable in most cases to provide the necessary information to model roof details and instead will probably result in very generalised roof models. Thus additional roof features will need to be extracted from the aerial photography as mentioned before.

4.3. Study area (The University of Nottingham, University Park campus)

The third and final study area selected for further implementing and optimizing the semi-automated workflow was the University Park, the main campus for The University of

Nottingham. University Park campus is located just 3.2km west of Nottingham's city centre with a total area of 133 hectares. The study area consists of various building structures with different architecture ranging from residential buildings with flat and gable type roofs to industrial buildings with complicated roof details. The topography of the terrain is characterized by rolling hills with dense vegetation and tree clusters present in many locations (figure 72).

Figure 72. Satellite view of the third study area located in The Nottingham University Park campus

The third study area presents a less difficult environment to model since most of the building roofs are less complex than the second study area over London - Bloomsbury. One of the most complicated buildings is one of the chemical engineering buildings (figure 74) due to the ventilation equipment and pipes present on the roof. Because this area is not as complex as the previous, it was decided that this area should be the basis for evaluating an optimized or adapted workflow for 3D modelling without the use of costly LiDAR data. Note that the

proposed workflow for the first two study areas is based on the synergy or data fusion of LiDAR with aerial photographs. In contrast for the third study area a revised workflow is presented for semi-automatic 3D modelling using only aerial photographs and effectively substituting the LiDAR data with Very Dense DSMs (VDDSM). The decision to exclude LiDAR was based on the fact the there is still limited availability for urban areas and due to the increased cost for data supply. It was also an excellent opportunity to test whether the high resolution UltraCamD images could substitute LiDAR data.

Hence the third study area presents a significant deviation from the initial workflow and highlights the adaptability of the proposed method in alternative scenarios with limited data sources. In the subsequent sections this study area will be simply referred to us 'Nottingham test site' for convenience.

4.3.1. Data Provided

A block of high resolution aerial imagery from the UltraCam D sensor was collected and distributed by BlomAerofilms Ltd. for research purposes.

4.3.1.1. UltraCamD images

A block of 85 UltraCamD images was provided by Blom Aerofilms Limited captured in November 2006 over the University Park campus. The block of images consists of 4 strips flown west to east and reverse, with 20% lateral overlap and 60% forward overlap. Flight altitude was 500m that led to a 6cm GSD. The composite true-color images have 11500x7500 pixels.

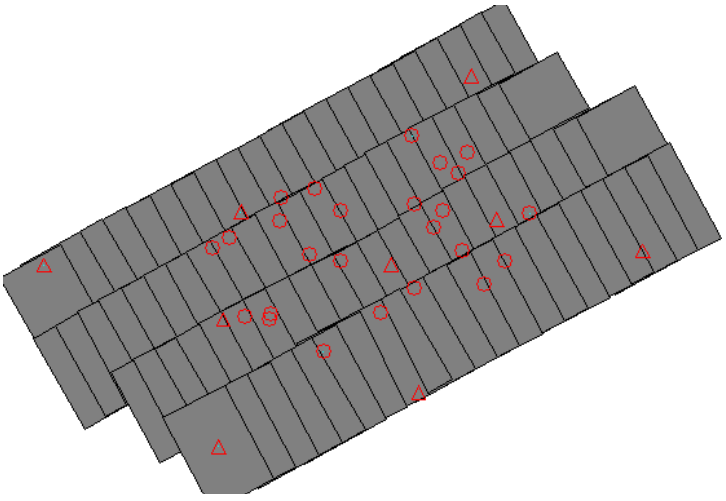


Figure 73. The block of UltraCam images with the distribution of GCPs

Each pixel has 9μm physical dimensions while the nominal focal length was equal to 101.4mm. In-flight GPS and IMU data were also available.

A total number of 37 coordinated ground points were available with a good distribution in the block of images (figure 73). These points were collected using static GPS observation of at least 20 minutes in duration and an estimated accuracy of 3cm which was used as the standard deviation of the ground control points in the aerial triangulations.

One of the main advantages of using the UltraCam D images for this study area is the impressive radiometric quality and ground resolution as indicated in section 3.8 that has the potential to create very dense DSMs to substitute the absence of LiDAR data. The overall quality and ground resolution of the UltraCam images is indicated in figure 74.

Figure 74. Subset of an UltraCam D image over Nottingham test site representing one of the chemical engineering buildings of increased roof complexity

5. PRE-PROCESSING AND DATA PREPARATION

This chapter describes in detail the pre-processing analysis implemented in the three different study areas in order to prepare the datasets before utilising the proposed workflow. The chapter contains information and critical analysis of the results obtained from the aerial triangulations of the airborne images. It is important to note that trying to automate the AT process was beyond the scope of this research. Instead conventional workflows and software packages were used during this stage in order to define the 'benchmark' or the optimum results for each dataset and study area.

5.1. Switzerland data – Pre-processing and AT of the ADS40 imagery

The first step of the post-processing includes the determination of the position and orientation of each scan line in the imagery based on the GPS and IMU data. At this stage optional GPS base station data are also included, to calculate the smoothed best estimate of trajectory.

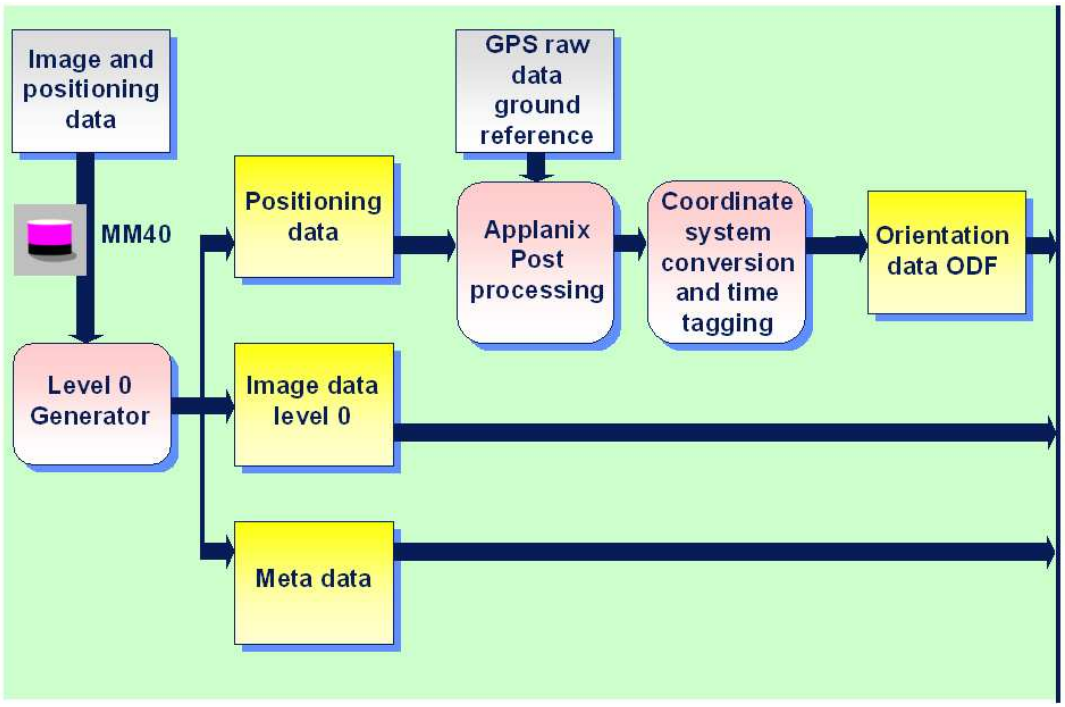


Figure 75 ADS40 data workflow for the production of the orientation data imagery.

Since different lines may be using different sampling frequencies, the data are processed into separate files for all sensor lines. Thus, at the end of this processing step, there is one data file for each sensor line, containing the position and attitude for each line of the image. Figure 75 indicates the workflow for the generation of the orientation data required for the rectification of the Level 0 imagery. Note that because this operation requires access to the raw positioning data and the use of the proprietary software Leica GPro the processing of the orientation data ODF and generation of Level 0 imagery was performed by Leica Geosystems.

After post-processing the GPS/IMU data, the position and attitude files are created which are represented as the orientation data ODF. These parameters together with the camera calibration and a simplified interior orientation of the camera are used to rectify the images. The rectification consists of correcting the distortions present in the Level 0 images by projecting the data to a ground plane at a user-specified elevation. This allows the correction of the aircraft motion and results in the Level 1 imagery that can be used in various software packages for further processing. Figure 76 indicates the workflow for the production of Level 1 images.

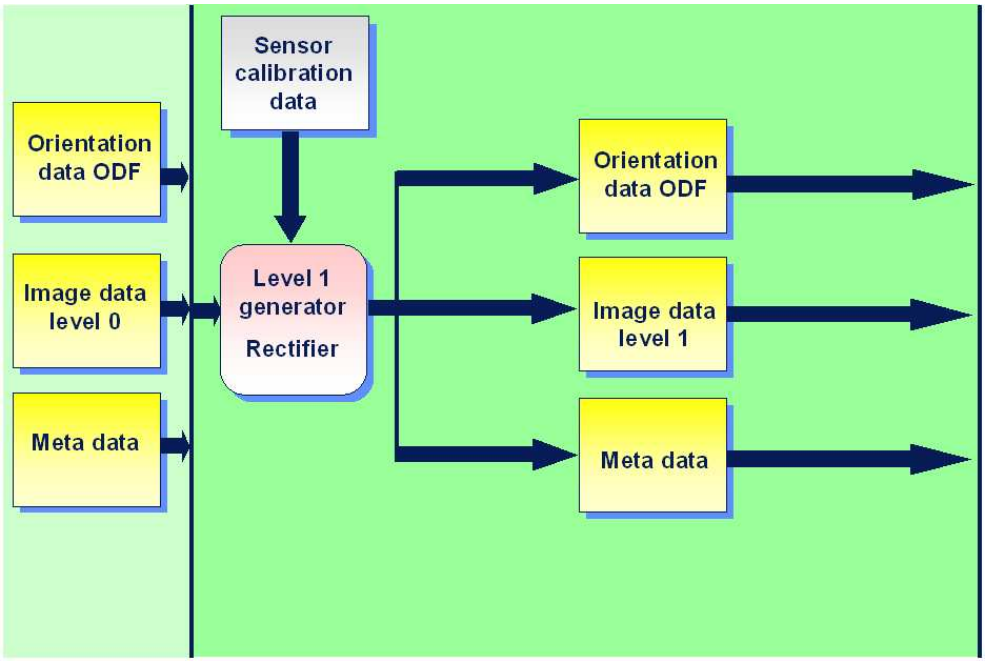


Figure 76 ADS40 data workflow for the production of Level 1 imagery

During the generation of the Level 1 imagery, the continuous strip of data is divided in user defined blocks for displaying the data. However the imagery is still treated as continuous strips by the ground processing system. Additionally during the rectification process, a look up table is created which provides a fast method for applying a transformation between the raw Level 0 images and the Level 1 images. This information is necessary, in order to recover the instantaneous position and attitude associated with any pixel in the Level 1 image and subsequently used for an aerial triangulation (Tempelmann, et al. 2000).

Apart from the rectified images the ground processing system can produce the “precision Level 1” imagery that results in a high resolution stereo viewable imagery. It uses an existing DEM and combines both lines of the staggered pairs and the geometric model of the sensor.

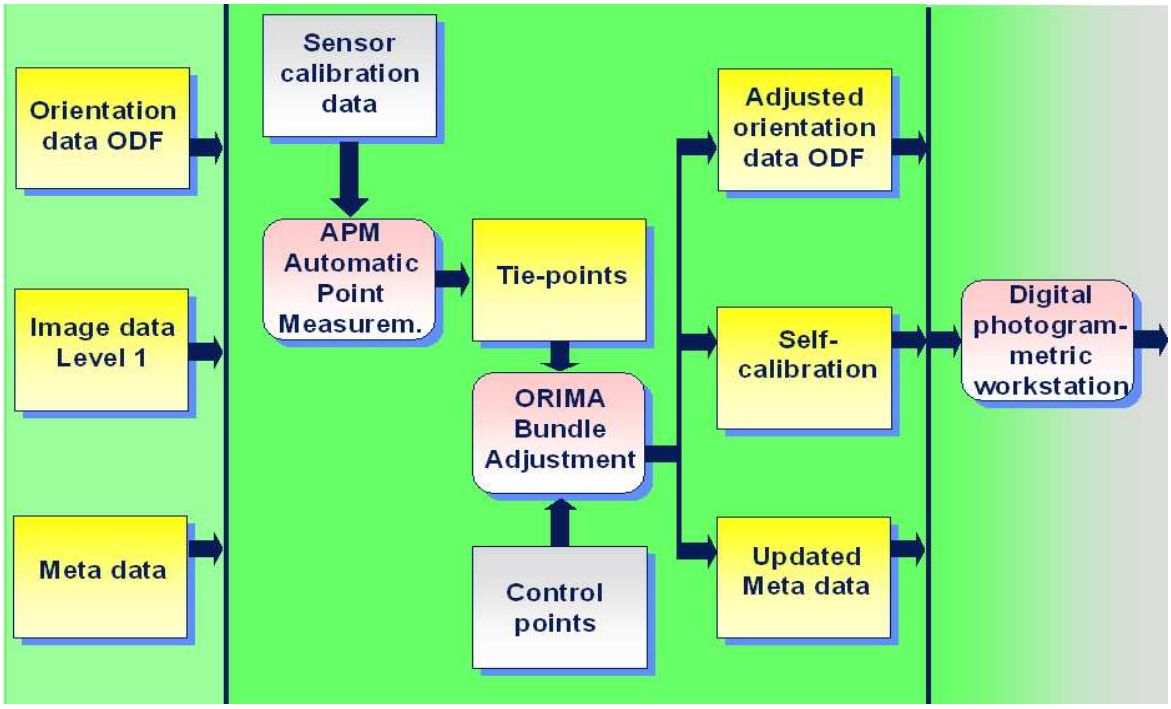


Figure 77 ADS40 data workflow for the aerial triangulation

Further processing of the ADS40 Level 1 imagery was applied by performing an aerial triangulation. This stage is implemented in order to precisely determine the exterior orientation by combining the short-term, high frequency precision of the IMU system with the

high global accuracy of GPS. Because of the integrated GPS/IMU, the requirements for ground control points are greatly reduced. Figure 77 indicates the workflow for the aerial triangulation using SOCET SET®.

The aerial triangulation is performed using a combined bundle adjustment that compensates systematic effects, which include primarily the misalignment between IMU and camera axes and the datum differences between IMU/GPS. The combined bundle adjustment uses the generalized collinearity equations to determine the position and orientation for the sensor lines at a certain interval, which are called orientation fixes. The orientation and position of the sensor lines between two orientation fixes are determined by interpolation using the IMU/GPS measurements. The generalized collinearity equations are given below.

$$\begin{aligned} x_i &= F(X_i, Y_i, Z_i, X_k, Y_k, Z_k, \omega_k, \varphi_k, \kappa_k, X_{k+1}, Y_{k+1}, Z_{k+1}, \omega_{k+1}, \varphi_{k+1}, \kappa_{k+1}) \\ y_i &= G(X_i, Y_i, Z_i, X_k, Y_k, Z_k, \omega_k, \varphi_k, \kappa_k, X_{k+1}, Y_{k+1}, Z_{k+1}, \omega_{k+1}, \varphi_{k+1}, \kappa_{k+1}) \end{aligned} \tag{5.1}$$

Where the X_k,Y_k,Z_k and X_{k+1},Y_{k+1},Z_{k+1} define the position of the two GPS fixes and ω_k, φ_k, k_k, ω_{k+1}, φ_{k+1}, k_{k+1} define the orientation angles given by the IMU. The triangulation process involves automatic measurement of tie points and interactive measurement of control points using the APM module from SOCET SET®. The observations are image coordinates, and the position/attitude values computed by the IMU/GPS post processing software. The observation equations used to compensate datum differences are:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{GPS} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \lambda \cdot \mathbf{R} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{PC} \tag{5.2}$$

The left-hand side indicates the co-ordinates calculated from the integrated IMU/GPS system, derived from the post-processing software, which has corrected the measurements from the offset between the antenna and the projection centre. The right-hand side indicated the position of the projection centre as calculated from the bundle adjustment. In between are the 7 parameters of the 3D similarity transformation that are treated as unknowns in the bundle adjustment (Tempelmann, et al. 2000).

The results from the aerial triangulation performed on the ADS40 image strips is given in table 13. The aerial triangulation was based on the 34 GCPs provided from Leica Geosystems Ltd with an estimated precision of 5cm. Due to the non-optimal distribution (all gathered around Leica's facilities) and due to the relatively small image scale (25cm GSD) the results obtained are not indicative of the achievable performance of the ADS40 sensor compared to results from other researchers as shown in section 3.1.4.

Total image unit weight RMSE = 7.8µm	Control Point RMSE (no pts)	Check Point RMSE (no pts)
Ground X m	0.188 (10)	0.205 (24)
Ground Y m	0.162 (10)	0.192 (24)
Ground Z m	0.215 (10)	0.271 (24)
Image x µm	4.6 (30)	3.8 (72)
Image y µm	3.5 (30)	3.9 (72)

Table 13 Aerial Triangulation results from the ADS40 imagery

5.2. London - Aerial Triangulation of RC30 scanned aerial photographs

In contrast to the ADS40 data the RC30 aerial photographs were processed using traditional workflows in Leica Photogrammetry Suite (LPS), by performing the interior orientation, measuring GCPs and automatic tie points and carrying out the aerial triangulation. In addition in order to identify the optimum solution several AT results are presented using different self calibration models.

5.2.1. Interior orientation

The process of interior orientation reconstructs the bundle of light rays passing through the perspective centre as they existed during the image acquisition. The first stage is to transform the pixel co-ordinate system of the digital images in the camera's co-ordinate system defined by the axes of the fiducial marks and the position of the principal point. The specific frame camera has eight fiducial marks and figure 78 provides an overview of their position in the frame.

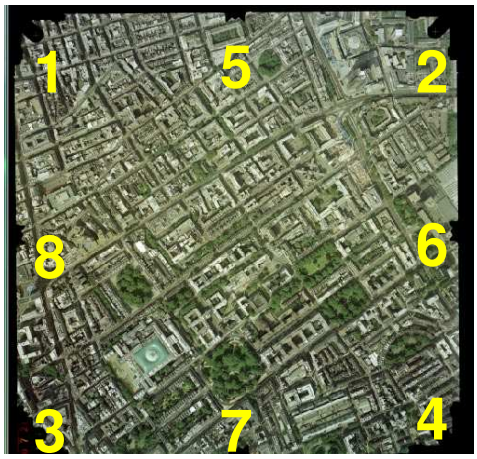


Figure 78 Location of the fiducial marks in the frame

The measurement of the fiducial marks using a semi-automatic process, whereby the first two points are measured manually and the rest are automatically measured using a cross-correlation matching algorithm using a primitive target matched against the fiducial mark. The cross-correlation matching uses a cross correlation coefficient as the main variable. An important key parameter in the internal orientation is the selection of the most appropriate mathematical model for the two dimensional transformation. Most softcopy photogrammetric system have either a 2D similarity transformation (four unknown parameters) or an affine transformation to correct distortions caused by non orthogonality and different scale distortions along the two axes. In order to account for any film distortions the affine transformation is selected in this project. The results from the interior orientation are summarized in table 14.

Image 1 (residuals in pixels)			
RMSE: 0.47 pixels or 6.7µm			
Point	ResidualX	ResidualY	Coeff
1	0.15	-0.18	
2	0.13	0.29	
3	-0.31	-0.43	0.92
4	0.51	-0.39	0.96
5	-0.04	-0.15	0.99
6	-0.62	-0.09	0.92
7	0.04	0.67	0.91
8	0.12	0.33	0.96

Image 2 (residuals in pixels)			
RMSE: 0.48 pixels or 6.8µm			
Point	ResidualX	ResidualY	Coeff
1	0.26	-0.56	
2	-0.01	0.40	
3	-0.28	-0.08	0.92
4	0.35	-0.56	0.96
5	0.09	0.10	0.98
6	-0.50	-0.15	0.97
7	0.25	0.55	0.93
8	-0.18	0.29	0.96

Image 3 (residuals in pixels)			
RMSE: 0.40 pixels or 5.6µm			
Point	ResidualX	ResidualY	Coeff
1	0.23	-0.27	
2	-0.02	0.17	
3	-0.10	-0.34	
4	0.30	-0.31	0.97
5	0.18	0.04	0.99
6	-0.46	-0.15	0.96
7	0.19	0.55	0.94
8	-0.32	0.29	0.97

Table 14 The interior orientation results from the automatic measuring process

Investigating the results from the interior orientation we can easily identify the high cross correlation coefficient used in all matches. This indicates significant reliability and matching precision. The total RMS of 6μm for the transformation is typical of what is expected from relatively old scanned photographs where scanning and film distortions might be present. The interior orientation was also performed using manual observations but in most cases the results were not improved and even got worse.

5.2.2. Automatic tie point extraction and GCP measurement

A total of 11 GCPs were available for this study area. The ground control point observations were performed using static observations using dual frequency GPS receivers and in many cases the duration was extended well beyond 30min in order to account for errors caused by multipath and limited open horizon due to tall buildings. The estimated precision from the differentially post processed GPS observations was 10cm.

The observation of the GCPs on the aerial photographs in many cases was hindered by the poor radiometric quality of the scanned images. In addition the points did not have an optimal distribution (figure 79) since the GCPs were initially measured for a previous pair of photographs that had less coverage of the study area.

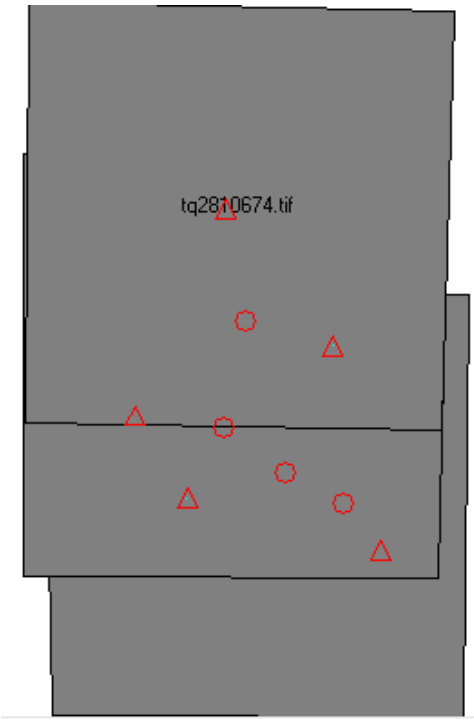


Figure 79 Distributions of the GCPs

The automatic tie point measurement in LPS is an area based matching algorithm, utilizing a least squares correlation workflow.

Least squares correlation uses the least squares estimation to derive parameters that best fit a search window to a reference window. It accounts for both gray scale and geometric differences, making it especially useful when ground features on one image look somewhat different on the other image (differences which occur when the surface terrain is quite steep or when the viewing angles are quite different).

Least squares correlation is iterative. The parameters calculated during the initial pass are used in the calculation of the second pass and so on, until an optimum solution is determined. Least squares matching can result in high positional accuracy (about 0.2 pixels). However, it is sensitive to initial approximations. The initial coordinates for the search window prior to correlation must be accurate to about two pixels or better. Thus the process requires either a good distribution of existing tie points measured by the operator or previously measured GCPs to determine the Exterior Orientation values of each image.

In most softcopy photogrammetric systems the change in gray values between two correlation windows is represented as a linear relationship given below.

$$g2(c2, r2) = h0 + h1g1(c1, r1) \tag{5.3}$$

Also assume that the change in the window's geometry is represented by an affine transformation.

$$\begin{aligned} c_2 &= a_0 + a_1 c_1 + a_2 r_1 \\ r_2 &= b_0 + b_1 c_1 + b_2 r_1 \end{aligned} \tag{5.4}$$

In the above equations,

c1,r1 = the pixel coordinate in the reference window

c2,r2 = the pixel coordinate in the search window

g1(c1,r1) = the gray value of pixel (c1,r1)

g2(c2,r2) = the gray value of pixel (c2,r2)

h0, h1 = linear gray value transformation parameters

a0, a1, a2 = affine geometric transformation parameters

b0, b1, b2 = affine geometric transformation parameters

Based on the above assumptions, the error equation for each pixel is derived, as shown in the following equation:

$$v = (a1+a2c1+a3r1)gc+(b1+b2c1+b3r1)gr - h1 - h2g1(c1, r1)+\Delta g \quad (5.5)$$

$$\text{with } \Delta g = g2(c2, r2) - g1(c1, r1) \quad (5.6)$$

The values gc and gr are the vertical and horizontal gradients for pixels g2(c2,r2) based on the brightness difference with surrounding pixels. The results from the implementation of the auto tie point extraction are given in table 15 with the distribution represented in figure 80.

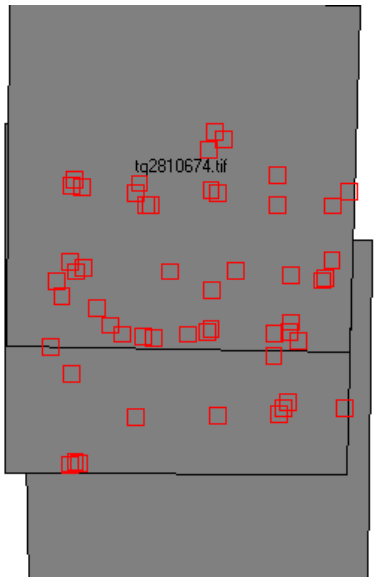


Figure 80 Distributions of the tie points in the stereo model

Img ID	Num of Intended Pts	Num of Pts Found	Num of Patterns	Point Success Rate	Pattern Success Rate
1	50	33	25	66.00%	46.00%
2	50	52	25	100.00%	76.00%
3	50	36	25	72.00%	52.00%
Average Point Success Rate: 79.33%					
Average Pattern Success Rate: 58.00%					
Total unique tie points found: 53					

Table 15 Report from the auto tie point extraction

In total 53 unique points were extracted with a satisfactory distribution since most of the overlapping area was covered by tie points. At this point we have to note that the auto tie point extraction process is much more reliable when adequate number of GCPs or tie points are measured beforehand. This is because by having measured GCPs or tie point before hand the exterior orientation and therefore the relationship of the images can be established and effectively minimise any mismatches of conjugate points. A visual inspection was also carried to identify and gross errors and mismatched points without excluding any tie points.

5.2.3. Aerial Triangulation results

The AT in LPS employs a least squares adjustment to estimate the unknown parameters, including the exterior orientation parameters, co-ordinates of tie points and any additional self-calibration parameters we may have. The least squares adjustment is achieved by minimizing and distributing data error through the network of observations. Data errors are attributed to the inaccuracy associated with the input GCP coordinates, measured tie point and GCP image positions, camera information, and systematic errors. The least squares approach requires iterative processing until a solution is attained. A solution is obtained when the residuals, or errors, associated with the input data are minimized.

A simplified version of the least squares condition can be broken down into a formula that includes a weight matrix *P,* as follows

$$V = A X - L \tag{5.7}$$

Where:

V = the column matrix containing the image coordinate residuals

A = the matrix containing the partial derivatives with respect to the unknown parameters, including exterior orientation, interior orientation, XYZ tie point, and GCP coordinates

X = the column matrix containing the corrections to the unknown parameters

L = the column matrix containing the input observations (i.e., image coordinates and GCP coordinates)

The *A* matrix is formed by differentiating the functional model, which is based on collinearity equations, with respect to the unknown parameters such as exterior orientation, etc. The *L* matrix is formed by subtracting the initial results obtained from the functional model with newly estimated results determined from a new iteration of processing. The *X* matrix contains the corrections to the unknown exterior orientation parameters. The *X* matrix is calculated in the following manner.

$$X = (A^t P A)^{-1} A^t P L \tag{5.8}$$

Where:

X = the column matrix containing the corrections to the unknown parameters t

A = the matrix containing the partial derivatives with respect to the unknown parameters

t = the matrix transposed

P = the column matrix containing the weights of the observations

L = the column matrix containing the observations

The first set of results for the London datasets were computed without any additional self calibration model. In addition 5 points were chosen as ground control points with 10cm precision, 5 points were selected as check points while one GCP was excluded due to high residuals (distribution of GCPs and check points is given in figure 81). The interior orientation parameters were kept fixed and the initial Exterior Orientation (EO) values were left floating. Note that no initial values were provided since no in-flight GPS/IMU data were available, instead the EO values were set as 'unknown'. The triangulation converged due to the amount of GCPs available to initialise the EO parameters. Results from the first trial AT trial are summarised in figure 81 and table 16.

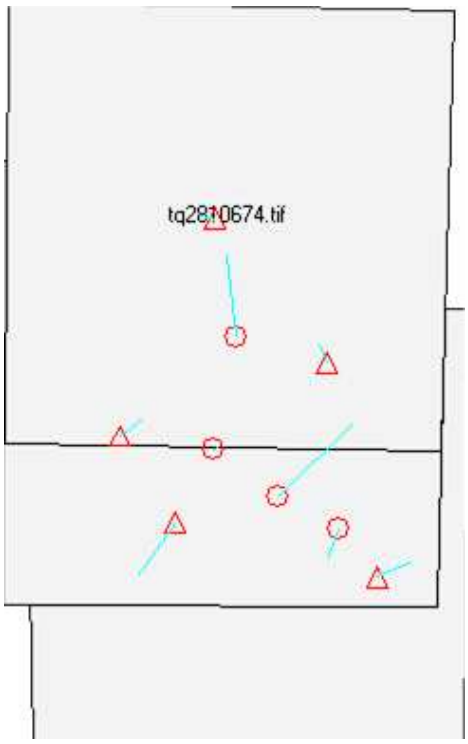


Figure 81 Ground residuals for control and check points from Trial 1 (scale —— = 10cm). (Δ = control points, O = check points)

1ST Aerial Triangulation trial

Total image unit weight RMSE = 6.2µm	Control Point RMSE (no pts)	Check Point RMSE (no pts)
Ground X m	0.069 (5)	0.103 (4)
Ground Y m	0.071 (5)	0.150 (4)
Ground Z m	0.064 (5)	0.171 (4)
Image x µm	2.5 (12)	1.8 (86)
Image y µm	2.8 (12)	4.9 (86)

Table 16 Image and ground residuals from AT results without self calibration – Trial 1

From the results shown in table 16 it is evident that even without any self calibration parameters and no in-flight GPS the results are satisfactory given the poor radiometric quality and the ground resolution (8cm). Most of the ground residuals of the GCPs are of subpixel accuracy while the check points exhibit larger residuals as expected. There is no apparent systematic pattern in the image residuals as represented in figure 85 but this is primarily related to the limited number of GCPs. No significant image residuals are present exempt from very few exemptions in some of the check points. No mismatches or gross errors were observed in the tie points while the overall image residuals are within the range of 0.4-0.5 pixels.

In addition different self calibration models were also defined in an effort to optimise the AT results. However the use of self calibration models introduced even larger ground residuals in the check points. A typical example of this is given in figure 82 and table 17 using the Jacobsen model with 4 additional parameters.

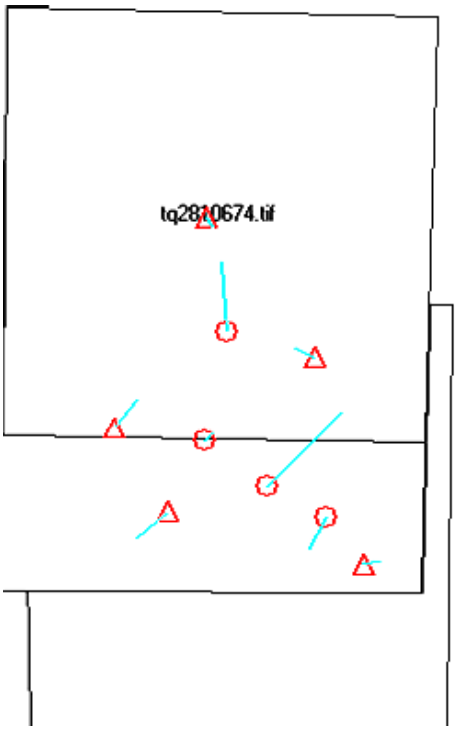


Figure 82 Ground residuals for control and check points from Trial 2
(scale —— = 10cm)
(Δ = control points, O = check points)

2nd Aerial Triangulation trial

Total image unit weight RMSE = 5.6µm	Control Point RMSE (no pts)	Check Point RMSE (no pts)
Ground X m	0.056 (5)	0.106 (4)
Ground Y m	0.050 (5)	0.152 (4)
Ground Z m	0.031 (5)	0.241 (4)
Image x µm	1.8 (12)	1.8 (86)
Image y µm	2.7 (12)	4.9 (86)

Table 17 Image and ground residuals from AT results Jacobsen's self calibration model – Trial 2

In table 17 we can clearly notice the introduction of significant vertical ground residuals in the check points. Clearly the self calibration model in an effort to distribute the image residuals is introducing large vertical ground residuals. There is also a noticeable forcing of the solution on the GCPs, particularly in the vertical direction, even though the block was left floating. There are only minor differences as to the direction and pattern of the residuals as illustrated in figure 82.

In conclusion better results might have been obtained from self calibration models if more tie points were extracted but the first trial was considered satisfactory for the objectives and requirements of this project.

5.3. The University of Nottingham campus - Aerial Triangulation of UltraCamD imagery

This section describes in detail the pre-processing of the block of 85 UltraCamD images consisting of the interior orientation, ground control and tie point measurement as well as several aerial triangulation results using different parameters.

5.3.1. Interior orientation

The UltraCam D imagery was delivered as a single image tile with 11500x7500 pixels and as such the interior orientation can be defined by specifying the calibrated focal length, the position of the principal point, lens distortion parameters and the physical dimensions of the CCD. Since digital sensors do not have any fiducials marks the physical dimensions of a single CCD detector will suffice for the reconstruction of the camera geometry. The block of images was also accompanied by an up-to-date camera calibration report, performed one year before the image acquisition.

The radiometric calibration is based on a series of 60 flat field images for each aperture size and sensor. The flat field is illuminated by two normal light lamps with known spectral illumination curves. These images are used to calculate the specific sensitivity of each pixel to compensate local as well as global variations in sensitivity. Sensitivity tables are calculated for each sensor and aperture setting, and applied during post processing from level 0 to level 1. Outlier Pixels that do not have a linear behaviour as described in the CCD specifications are marked as defective during the calibration procedure. These pixels are not used or only partially used during post processing and the information is restored by interpolation between the neighbourhood pixels surrounding the defective pixels. Certain pixels that are named Qmax pixels due to the fact that they can only store and transfer charge up to a certain maximum amount are detected in an additional calibration step. These pixels are treated differently during post processing, since their behaviour can affect not only single pixel values but whole columns.

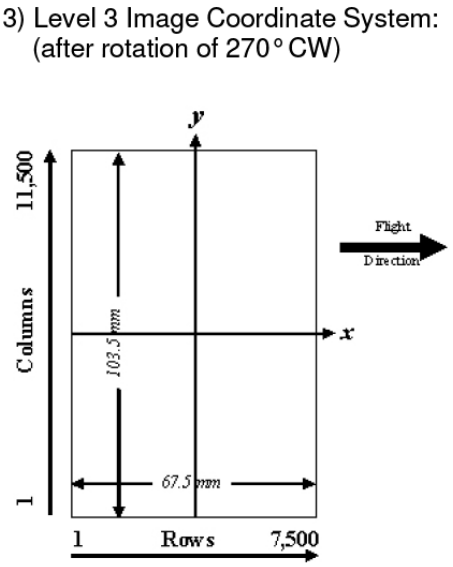


Figure 83 Camera co-ordinate system for UltraCam D as defined by the camera calibration report

The orientation of the camera co-ordinate system is also defined in the calibration report in relation to the flight direction as indicated in figure 83. As there are no measurements made by the operator to establish the relationship between the pixel co-ordinate system and the photo-co-ordinate system there are no errors reported for the transformation. This does not mean the interior orientation is error free. Lens distortions (radial and tangential) could still be present for example or deformation on the CCD sensor due to thermal variations.

5.3.2. Automatic tie point extraction and GCP measurement

A total number of 37 coordinated ground points were available. These points were collected using static GPS with an estimated accuracy of 5cm which was used as the standard deviation of the ground control points in the triangulations. The observation for the GCP in UltraCamD images was performed manually. Most of the GCP were easily identifiable due to the good radiometric quality of the images. The tie points for the UltraCamD block were automatically extracted using a cross-correlation area based matching technique available in Leica Photogrammetric Suite (LPS) as described in section 5.2.2. Blunders and mismatched points were identified manually by the operator based on the image residuals and were excluded in an iterative process after rerunning the AT. A total of 1000 tie points were extracted while 100 tie points were manually excluded after analyzing the image residuals.

From the 37 points available in the area 5 points were excluded from all the Aerial Triangulation results due to high residuals, 8 were selected as control points and 24 check points for the first two AT trials while only one control point was selected for the third and fourth AT trials. The distribution of the tie points and control/check points are shown in figures 84 and 85 respectively.

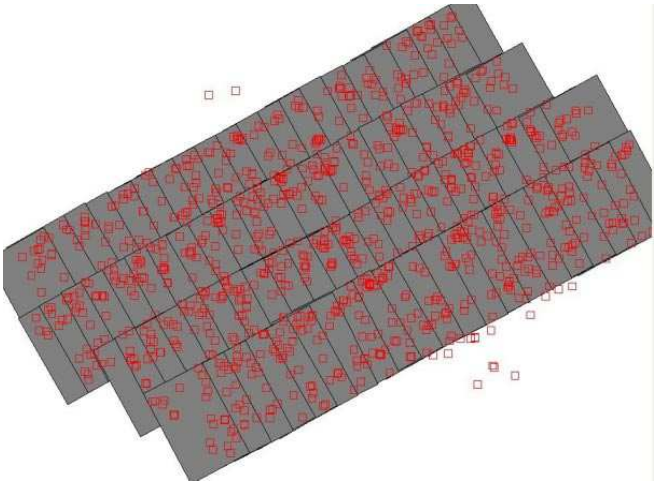


Figure 85 Distribution of extracted tie points

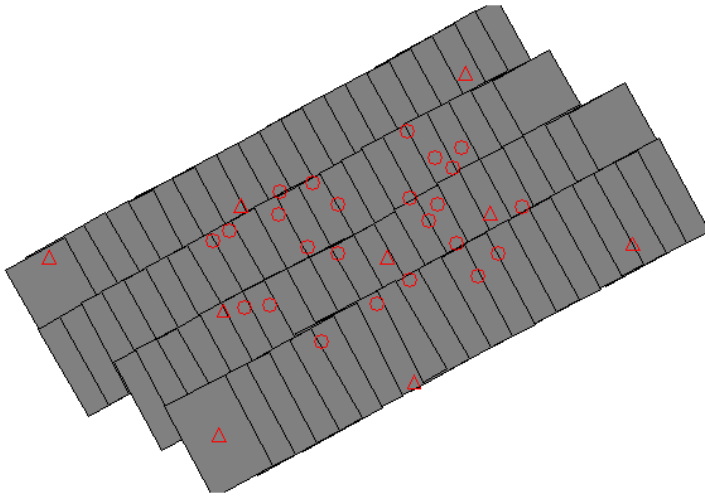


Figure 84 Distribution of measured control (Δ) and ckeck points (O)

5.3.3. Aerial Triangulation results

The aerial triangulation trials consist of four different solutions using a combination of different parameters. In the first two trials in-flight GPS and IMU data are used only as initial values without any weighting so that effectively the block is not constrained during the least squares iterations. The difference between the first and second trial is the introduction of a self calibration model. In addition the third and fourth trials use the in-flight GPS/IMU with appropriate weighting and only one control point in an effort to evaluate the direct georeferencing performance of the provided imagery. Again the third and fourth AT trials are distinguished by the utilization of self calibration models.

During the aerial triangulation computations 5 points were rejected due to large image or ground control residuals. The results from the best self-calibration method are presented.

1st AT trial: The first Aerial Triangulation trial was conducted using the in-flight GPS/IMU only as initial values. The block was left floating. Using 8 control and 24 check points. Control points had 5cm standard deviation for the weighting. No additional self calibration model was used and the interior orientation parameters were kept fixed. Figure 86 and table 18 illustrate the distribution and magnitude of ground residuals respectively.

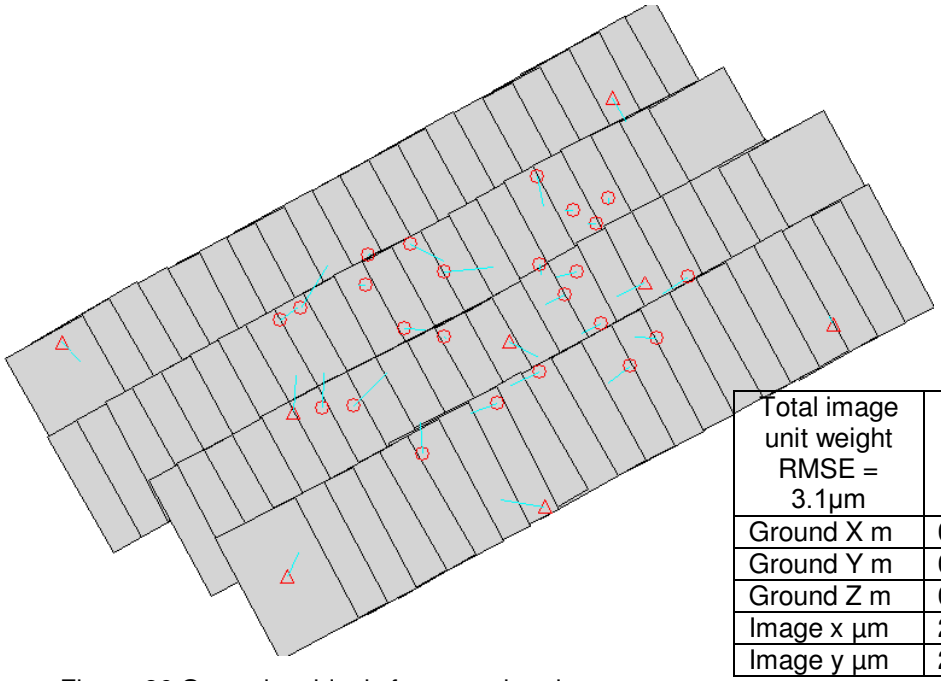


Figure 86 Ground residuals for control and check points from Trial 1
(scale —— = 5cm)
(Δ = control points, O = check points)

Total image unit weight RMSE = 3.1μm	Control Point RMSE (no pts)	Check Point RMSE (no pts)
Ground X m	0.064 (8)	0.058 (24)
Ground Y m	0.057 (8)	0.049 (24)
Ground Z m	0.057 (8)	0.148 (24)
Image x μm	2.7 (21)	3.0 (86)
Image y μm	2.8 (21)	2.5 (86)

Table 18 Image and ground residuals from AT results. No self calibration model – Trial 1

2nd AT trial: The second Aerial Triangulation trial was conducted using the in-flight GPS/IMU only as initial values. The block was left floating, using 8 control and 24 check points. Control points had 5cm standard deviation for the weighting. The best self calibration model was the Jacobsen model with 4 additional parameters. Figure 87 and table 19 illustrate the distribution and magnitude of ground residuals respectively.

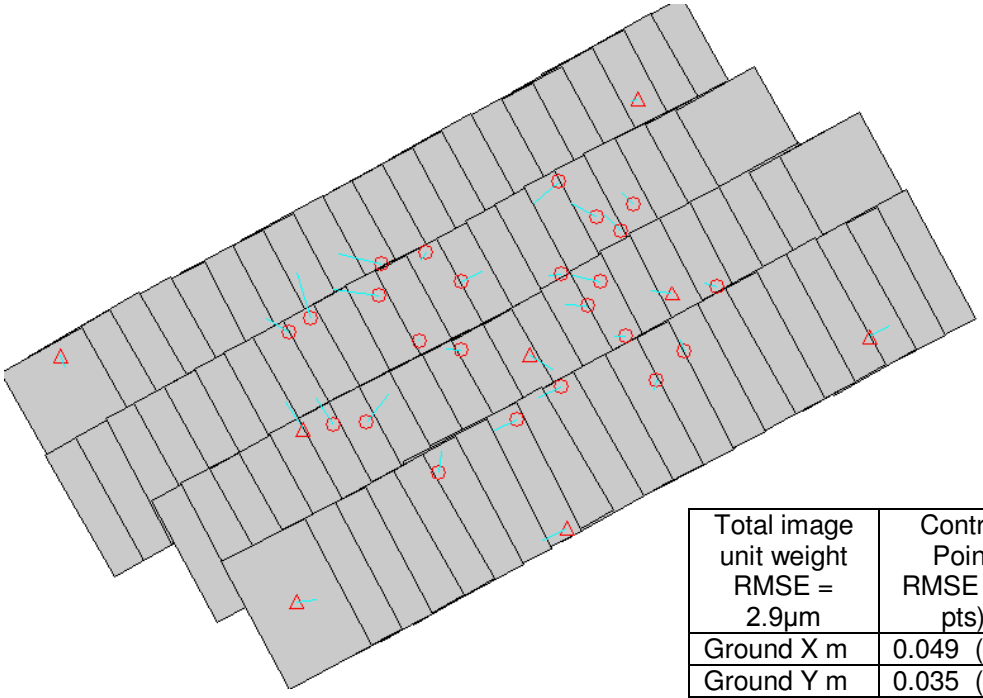


Figure 87 Ground residuals for control and check points from Trial 2
(scale ——= 5cm)
(Δ = control points, O = check points)

Total image unit weight RMSE = 2.9μm	Control Point RMSE (no pts)	Check Point RMSE (no pts)
Ground X m	0.049 (8)	0.055 (24)
Ground Y m	0.035 (8)	0.042 (24)
Ground Z m	0.025 (8)	0.106 (24)
Image x μm	2.7 (21)	2.9 (86)
Image y μm	1.8 (21)	2.5 (86)

Table 19 Image and ground residuals from AT results. Jacobsen self calibration model – Trial 2

3rd AT trial: The third Aerial Triangulation trial was conducted using the in-flight GPS/IMU with an appropriate standard deviation as weighting that reflects the expected precision of the GPS/IMU performance and minimizes the ground and image residuals. Thus in order to evaluate the direct georeferencing of the provided data only 1 control point was used, located in the centre of the block, while 31 check points were used. In a direct geoferencing scenario the weighting of the EO values is critical to the performance of the aerial triangulation and in many cases adjustment needs to be made in the weighting so that optimal results are achieved. The reader should note that the weighting should reflect the

expected precision of the EO values. In this project after a trial an error the weighting for the Xo,Yo,Zo, was set to 0.3m and for the ω,φ,κ to 0.05°. The selected standard deviation values minimised the residuals as well as reflecting the expected precision of the raw EO values. Any increase or decrease from the selected weighting was increasing ground residuals. Note that the positional accuracy of the EO values was less than expected due to the quality of the GPS data and reference stations used during the GPS post-processing. For this trial no self calibration model was used. Figure 88 and table 20 illustrate the distribution and magnitude of ground residuals respectively.

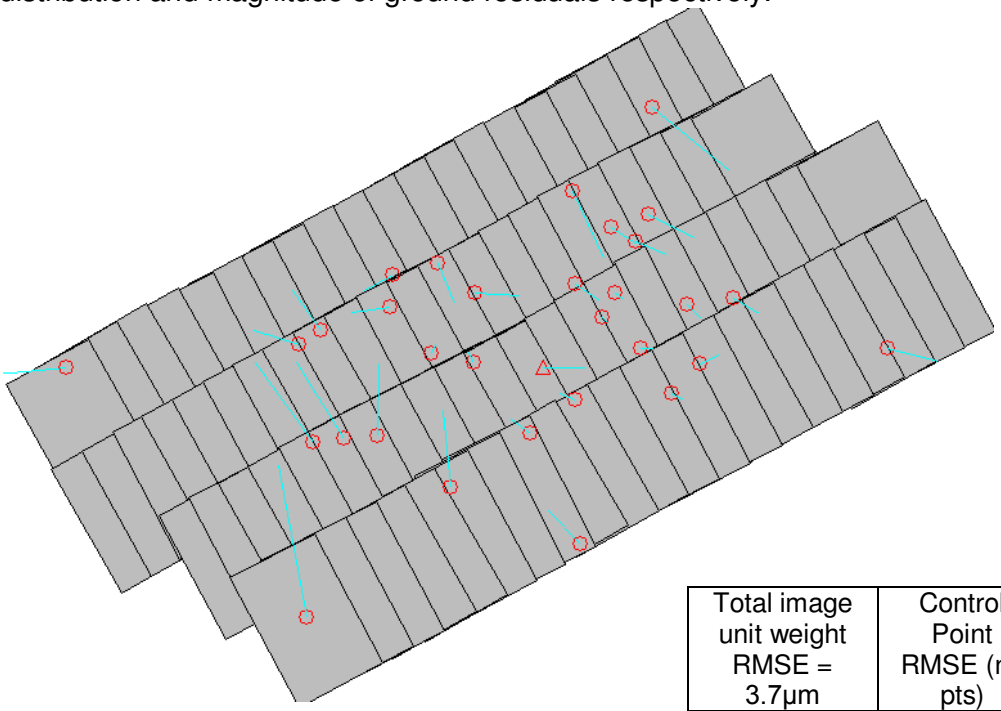


Figure 88 Ground residuals for control and check points from Trial 3
(scale —— = 5cm)
(Δ = control points, O = check points)

Total image unit weight RMSE = 3.7μm	Control Point RMSE (no pts)	Check Point RMSE (no pts)
Ground X m	0.112 (1)	0.088 (31)
Ground Y m	0.002 (1)	0.116 (31)
Ground Z m	0.048 (1)	0.135 (31)
Image x μm	3.0 (2)	2.8 (105)
Image y μm	2.6 (2)	2.5 (105)

Table 20 Image and ground residuals from AT results. No self calibration model – Trial 3

4th AT trial: The fourth and last Aerial Triangulation trial was conducted using the in-flight GPS/IMU with an appropriate standard deviation as explained previously. Only 1 control point was used, located in the centre of the block, while 31 check points were used. The weighting of the raw EO values for the Xo,Yo,Zo, was set to 0.3m and for the ω,φ,κ to 0.05°.

The selected standard deviation values minimised the residuals as well as reflecting the expected precision of the raw EO values. Any increase or decrease from the selected weighting was increasing the ground residuals. For this trial Bauer's simple self-calibration model with 3 additional parameters presented the best results. Figure 89 and table 21 illustrate the distribution and magnitude of ground residuals respectively.

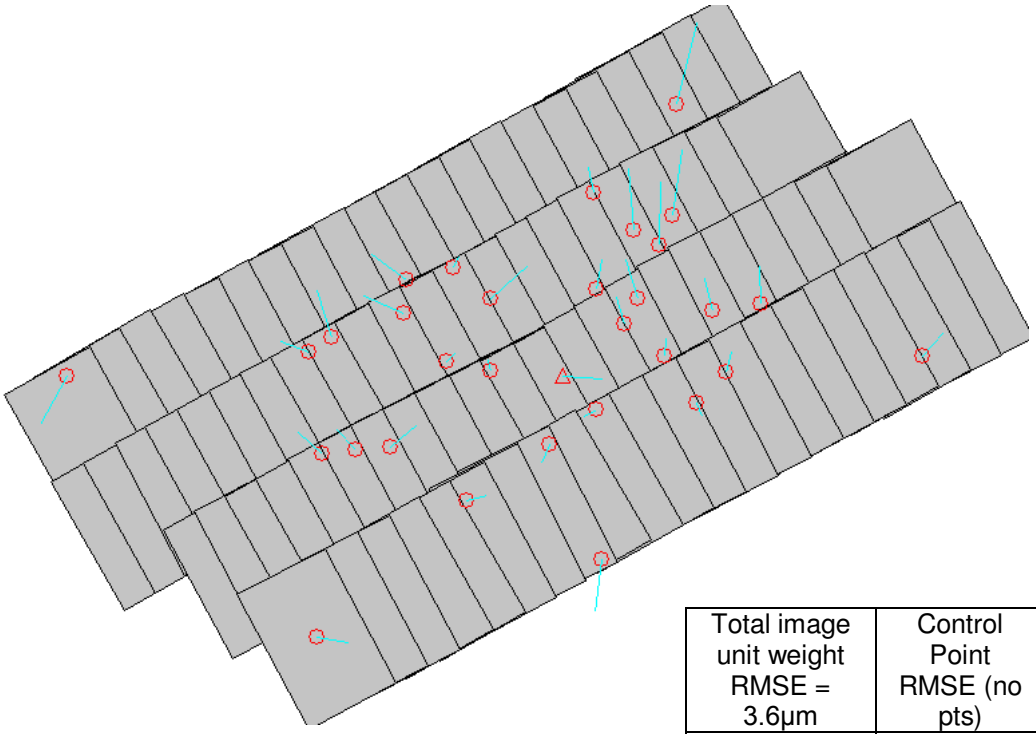


Figure 89 Ground residuals for control and check points from Trial 4
(scale ——= 5cm)
(Δ = control points, O = check points)

Total image unit weight RMSE = 3.6µm	Control Point RMSE (no pts)	Check Point RMSE (no pts)
Ground X m	0.105 (1)	0.046 (31)
Ground Y m	0.009 (1)	0.087 (31)
Ground Z m	0.009 (1)	0.116 (31)
Image x µm	3.2 (2)	2.8 (105)
Image y µm	1.1 (2)	2.5 (105)

Table 21 Image and ground residuals from AT results. Bauer's simple self calibration model – Trial 4

From the trials presented above it is clear that the optimum result with the lowest ground residuals is the second trial using Jacobsen's self-calibration model and 8 ground control points. Nevertheless it is important to note the impressive performance of the direct geo-referencing when it comes to in-flight GPS/IMU. The fourth trial with the use of Bauer's simple self calibration model presents satisfactory results even with the use of only one GCP. However we could expect even better performance for the in-flight GPS data. From a

comparison between the raw EO values and the calculated EO values after the AT there were significant positional deviation of the order of 0.5m. The 0.3m standard deviation chosen for the in-flight position reflects the non optimal GPS performance. The in-flight inertial system performed much better with the standard deviation in the order of 0.05 degrees. In both cases the use of self calibration parameters seems to improve primarily the vertical ground residuals and the overall image residuals as well. Investigating the distribution of the ground residuals in the first two cases there is no noticeable pattern with some higher residuals concentrated on the left from the block's centre. In contrast the aerial triangulation results using the in-flight GPS/IMU have noticeable patterns. In the third trial the residuals seem to have two predominant directions while in the fourth trial the influence of the self calibration model in re-distributing the residuals is evident and the patterns noticed before less obvious.

For the remaining tasks of this project and the implementation of the design algorithms the result from the second trial are adopted. With a planimetric accuracy, as given by the check points, in the sub-pixel range and 10cm vertical ground residuals, the results are considered satisfactory for the requirements of the project.

6. PROPOSED WORKFLOW FOR AUTOMATIC 3D CITY MODELLING AND ANALYSIS OF THE RESULTS

This chapter provides a detailed description of the proposed semi-automatic workflow for 3D city modelling and analyzes the results from the implementation of the algorithms in the three study areas. It is divided in four main sections each describing the algorithms and analysing the results created for extracting features from optical data, classifying LiDAR or Very Dense Digital Surface Models (VDDSM), data fusion for optimising building footprint and 3D polyhedral modelling. The design workflow and algorithms are implemented in such a way as to decrease the level of user interaction and it has been refined so that it produces acceptable results in all the three test sites and as such is reliable in different areas with various data sources available. The individual stages were designed to address the issues introduced in the building reconstruction process with the potential view to be fully automated.

6.1. Overview of the proposed method

The proposed workflow (see figure 90) for semi-automated 3D city modelling can be summarised as follows. The dashed outlines in figure 90 indicate that the enclosed processes are described in a particular section.

- Perform aerial triangulation (AT) on the block of images. Extracting edges from images using Laplace or Sobel operators. Extracted edges used for optimized stereo matching.
- If LiDAR data are available linear segments are extracted using geomorphometric parameters and used to refine the building footprints. If LiDAR data are not available stereo matched points are used to refine footprints and roof details by fitting 3D linear segments. The linear segments are used to create an optimized TIN which is converted to a VDDSM.

- Classification is performed either on the LiDAR data or on the VDDSM to extract the ground surface, tree clusters and perform the building detection. Finally 3D city modelling is performed using least squares plane fitting on classified building points.

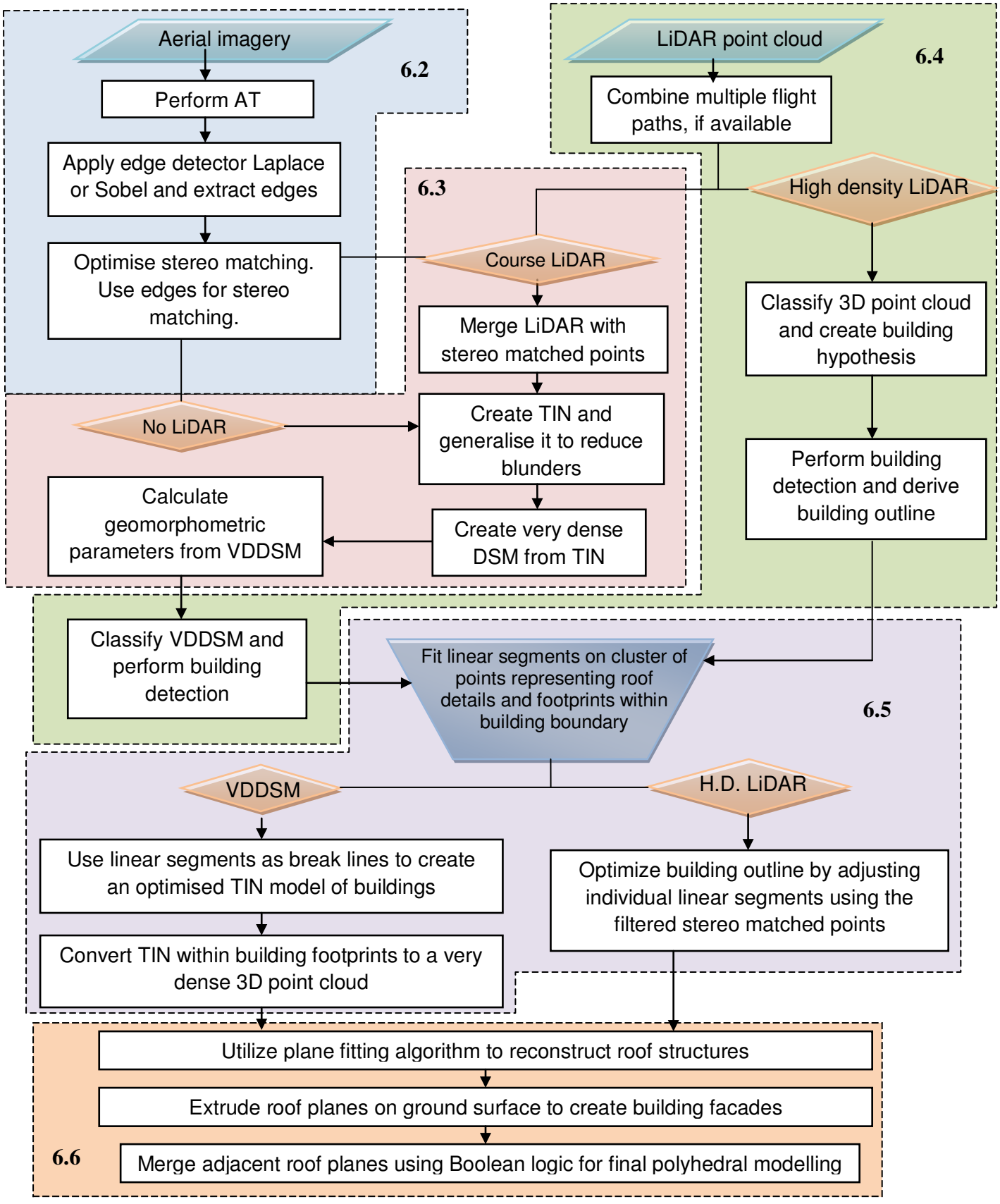


Figure 90 Flow chart diagram indicating the proposed workflow for semi-automated 3D city modelling

Although the method utilizes the provided datasets, most of the processing steps were designed in order to be efficient in a variety of different situations. That is, the proposed method is not limited in using for example only sensors that collect multispectral optical data, but instead provide reliable results even when scanned aerial photographs are used. In fact the provided data were limited to panchromatic bands and true colour composites. Hence, the method had to utilize more generic approaches for solving certain aspects of the processing stages.

Nevertheless the crucial aspect of this method is the use of data fusion techniques between two different datasets, optical and LiDAR data in this case, in order to enhance the overall result. The improvement from the proposed method is apparent in both the generation of the building hypothesis and the building reconstruction phases. In addition in order to improve the scope of the workflow and deal with the limited availability of LiDAR data (note that the third test site didn't include any LiDAR data) there is an alternative process for creating very dense digital surface models (VDDSMs) to substitute LiDAR data.

The method for feature extraction follows the general workflow of the method proposed in Chen et al. (2004). According to this method the approximate location of the edges, constituting the vertical walls and additional roof features, are extracted initially from the LiDAR point cloud (if available) and subsequently are refined based on the linear features derived from the optical data. The information from the linear features will not only refine and improve the accuracy of the building outlines but provide information for smaller roof details that are not modelled correctly in the case of using only low density LiDAR data. A major difference of the proposed method is that the adjustment of the building outlines and related features will be performed using a least square adjustment and not the Hough transform used in Chen et al. (2004). The reason for not implementing the Hough transform is because it's a brute-force method which can be very complex and computationally demanding. Thus it has two main disadvantages, large memory requirements and slowness.

From the diagram depicted in figure 90, it's evident that the creation of building hypothesis (building detection) is applied on a LiDAR point cloud or VDDSM for classifying the data. This stage is briefly mentioned in the diagram but in fact is one of the most complicated stages of the entire project. This stage requires a certain level of user interaction and therefore the entire procedure of building detection is only semi-automated.

6.2. Extracting features from optical data

The section describes the process of extracting conjugate points from the available stereo pair of images as shown in figure 91. The results from the aerial triangulations were presented in chapter 5.

There are three main steps for extracting the desired features from the optical data:

- Select the most appropriate stereo pair, if multiple overlapping images exist, for the study area.
- Apply edge detector and add the extracted features to the radiometric values of the initial image.
- Optimize stereo matching algorithm for extracting conjugate points in urban areas.

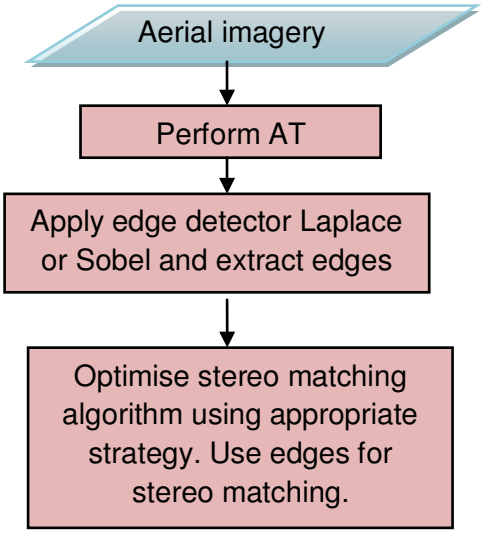


Figure 91 Subset of overall workflow showing the process of feature extraction. Colours indicate the software used for each operation

At this point it is important to summarise some of the previous research efforts in stereo matching from airborne and satellite imagery to set the background for the work that was carried out in this section. Most stereo matching techniques can be categorised as either area-based methods or feature-based methods.

Area-based methods, sometimes called correlation-like methods or template matching (Fonseca and Manjunath, 1996) merge the feature detection step with the matching part.

These methods deal with the images without attempting to detect salient objects. Windows of predefined size or even entire images are used for the correspondence estimation during the second registration step, (Althof et al., 1997). The limitations of the area-based methods originate in their basic idea. Firstly, the rectangular window, which is most often used, suits the registration of images which locally differ only by a translation. If images are deformed by more complex transformations, this type of the window is not able to cover the same parts of the scene in the reference and sensed images (the rectangle can be transformed to some other shape). Several authors proposed to use circular shape of the window for mutually rotated images. However, the comparability of such simple-shaped windows is violated too if more complicated geometric deformations (similarity, perspective transforms, etc.) are present between images.

Classical area-based methods like cross-correlation (CC) exploit image intensities for matching directly, without any structural analysis. Consequently, they are sensitive to intensity changes, introduced for instance by noise, varying illumination, and/or by using different sensor types.

In feature-based matching techniques spatial features are usually detected in both images and the aim is to find the pair-wise correspondence between them using their spatial relations or various descriptors of features.

Goshtasby and Stockman (1985) described the registration based on the graph matching algorithm. He was evaluating the number of features in the sensed image that, after the particular transformation, fall within a given range next to the features in the reference image. The transformation parameters with the highest score were then set as a valid estimate. Clustering technique, presented by Stockman et al. 1982, tries to match points connected by abstract edges or line segments. The assumed geometrical model is the similarity transform. For every pair of spatial features from both the reference and sensed images, the parameters of the transformation which maps the points on each other are computed and represented as a point in the space of transform parameters. The parameters of transformations that closely map the highest number of features tend to form a cluster,

while mismatches fill the parameter space randomly. The cluster is detected and its centroid is assumed to represent the most probable vector of matching parameters. Mapping function parameters are thus found simultaneously with the feature correspondence. Local errors do not influence globally the registration process. The clustering technique was implemented, for example, in (Chang et al., 1997). Borgefors (1988) proposed an improved version, where better measure of correspondence, the sequential distance transform, together with the root mean square average was applied. The algorithm employs also the pyramidal speed-up. Even though this overview does not intend to cover 3D registration methods, here the well-known Iterative Closest Point (ICP) algorithm, introduced by Besl and McKay (1992) is mentioned, because it represents a key approach for registering 3D shapes (including free-form curves and surfaces).

Feature-based matching methods are typically applied when the local structural information is more significant than the information carried by the image intensities. They allow the registration of images with completely different nature (like aerial photograph and map) and can handle complex between-image distortions. The common drawback of the feature-based methods is that the respective features might be hard to detect and/or unstable in time. The crucial point of all feature-based matching methods is to have discriminative and robust feature descriptors that are invariant to all assumed differences between the images.

This section presents an effort to combine the advantages of both techniques mentioned above by detecting edges on the aerial photographs to enhance the geometric shapes of the buildings and then performing a cross correlation, area-based matching.

6.2.1. Selecting appropriate stereo pair

The selection of the most appropriate stereo model may seem a relatively unimportant step but in the case of airborne digital sensors such as the ADS40, it becomes an important consideration. The reason is that the sensor collects 4 bands of data in three look angles resulting in different combinations of stereo pairs. One of the parameters that must be taken

into account is the base to height ratio in the vertical accuracy of the extracted conjugate points. Therefore, considering only the base to height ratio as a factor, a straightforward solution would be to use the 28° forward and 14° backward looking panchromatic bands. At this stage, there are two major issues that should also be taken into consideration. That is, the possibility of the study area being located outside the overlapping region of the two look angles and also the occlusions introduced from the relief displacement. The latter issue is the most critical, especially in this specific project, because the stereo matched points are subsequently used for adjusting the building outline and therefore, having conjugate points representing the planimetric position of all the building façades is more important than having the highest vertical accuracy possible.

The location of the study area should not be neglected, when designing the method, since rarely an automatic building reconstruction process will take place in the entire scene, because it is very computationally intensive. Therefore it is evident that a check mechanism should be introduced in the algorithm to determine the location of the study area on the imagery. The proposed method incorporates the generalized collinearity equation for back-projecting a polygon enclosing the study area at the image space and checking if it is within the extend of the image (equation 5.1).

This check mechanism does not require any substantial user interaction since in most cases the polygon, enclosing the study region, is previously defined from the project specifications.

From the discussion so far it is evident that the optimum selection is to have a stereo pair of the nadir and 14° backward looking panchromatic bands in order to have a small convergence angle and thus decreased occluded regions. The second option is a stereo pair between the nadir and forward looking images. This step applies only to the Swiss study area with the availability of the ADS40 data. For this study area the second option was selected since there was no overlap between the nadir and backward looking images.

6.2.2. Applying edge detection algorithm

The selected edge detector implemented for extracting edge features was the Sobel edge detection operator. In principle this operator uses the first derivative of a continuous function that is subsequently approximated by the following equation (Duda and Hart, 1972).

$$\frac{\Delta f(x,y)}{\Delta x} = \frac{f(x+1,y) - f(x,y)}{(x+1) - x} = f(x+1,y) - f(x,y) \tag{6.1}$$

$$\frac{\Delta f(x,y)}{\Delta y} = \frac{f(x,y+1) - f(x,y)}{(y+1) - y} = f(x,y+1) - f(x,y) \tag{6.2}$$

Where f(x,y) represents a continuous function for the pixel values.

The above equation leads to the concept of the image gradient for a discrete digital representation. Nevertheless in order to be more computationally efficient, this algorithm uses a kernel window of size 3x3 to scan the entire digital image. The defined window has the following form:

A	B	C
D	E	F
G	H	I

With the utilization of the above kernel window the updated value for the central pixel is calculated based on the following equation. The value (S) represents the gradient of the central pixel in relation to the surrounding pixel values.

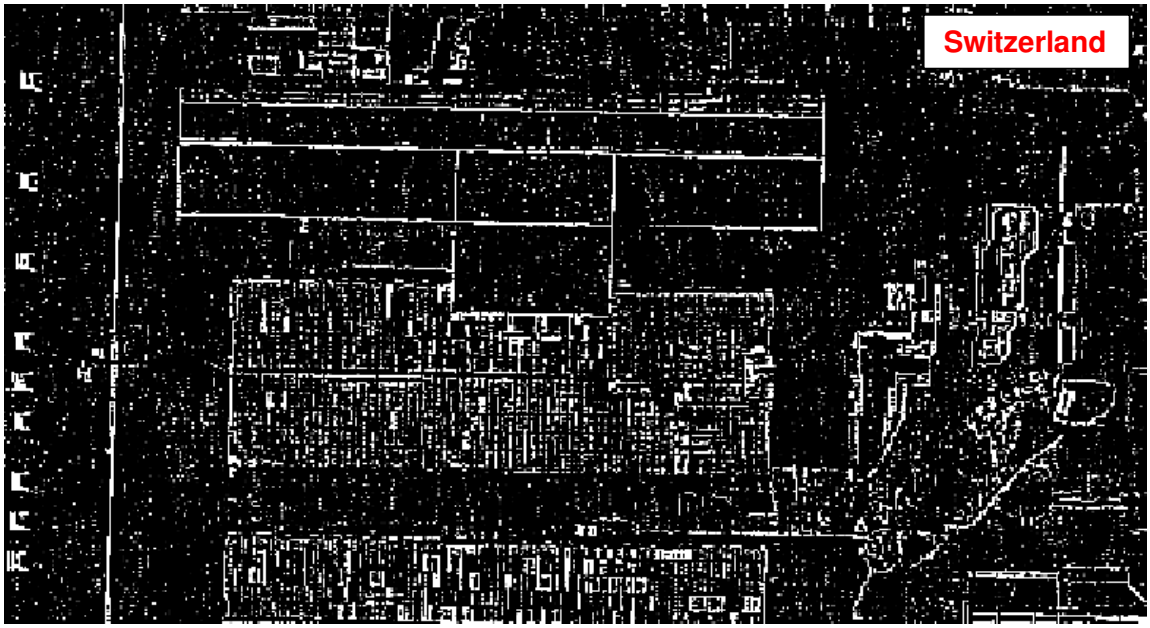
$$S = \sqrt{X^2 + Y^2} \tag{6.3}$$

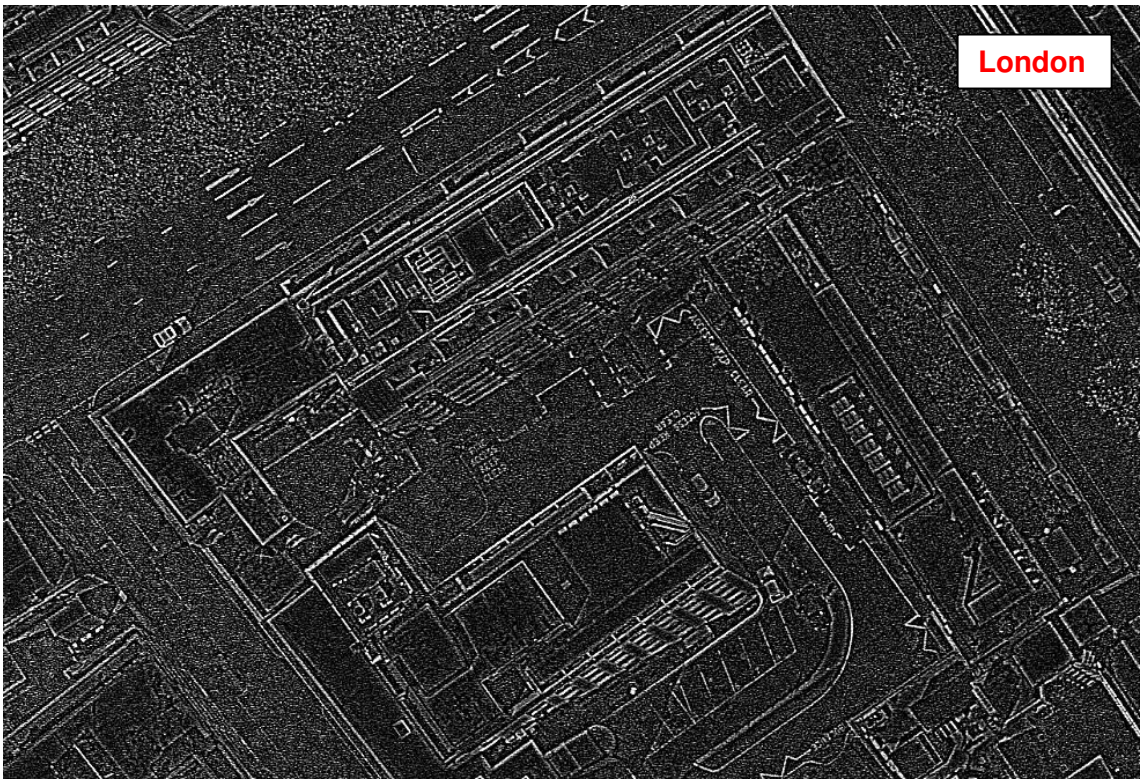
Where

$$X = (C + 2F + I) - (A + 2D + G) \tag{6.4}$$

$$Y = (A + 2B + C) - (G + 2H + I) \tag{6.5}$$

Although there is a variety of edge detection operators, Sobel presents one of the most widely used operators. The extracted edges from the implementation of the Sobel operator with a 3x3 kernel size in the three study areas are illustrated in figure 92.





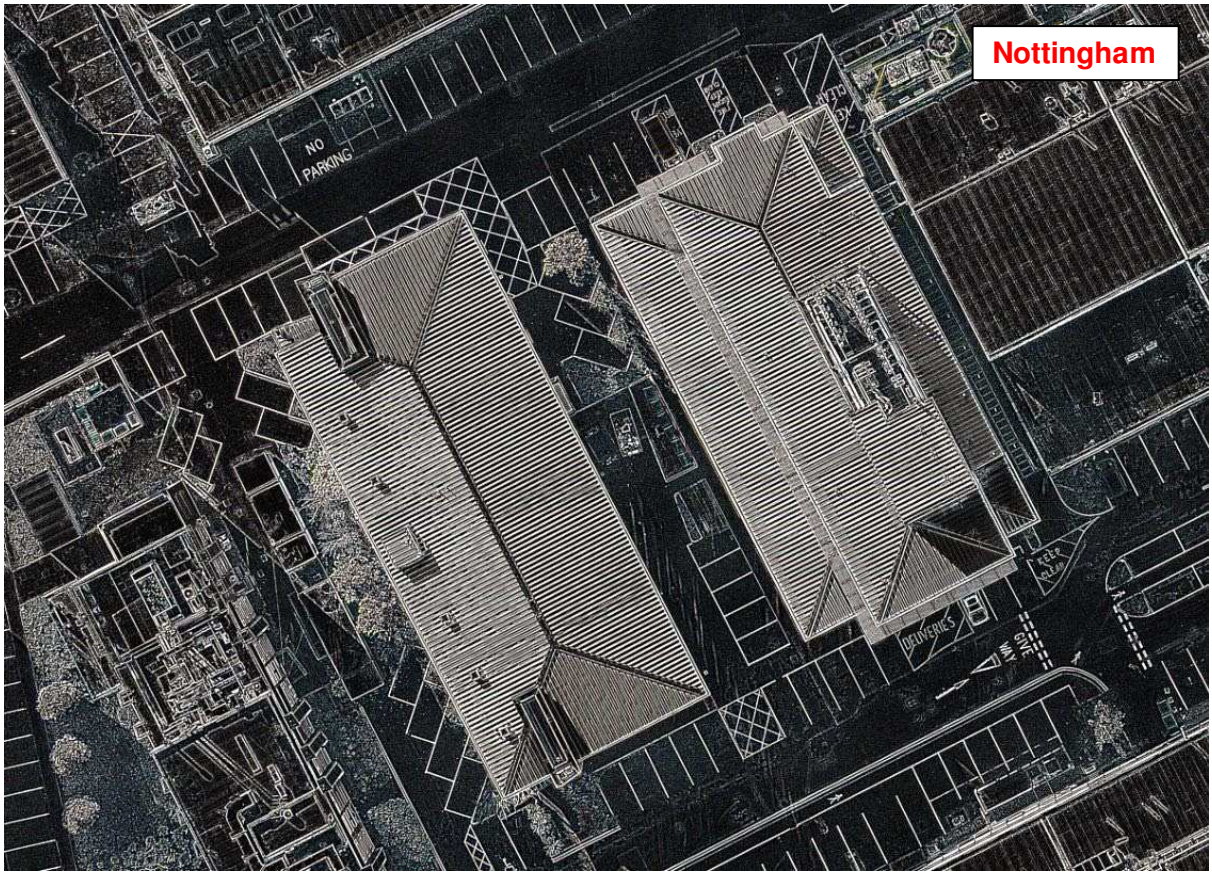


Figure 92 Results from the implementation of the Sobel edge detector in the three study areas

At this point it should be noted, that this research effort was focused on implementing a reliable edge detector operator that can enhance the image for assisting the stereo matching process. Hence the motivation for applying an edge detector is to increase the probability of the stereo matching algorithm producing as many conjugate points as possible along the enhanced linear features.

Upon further investigation of the results shown in figure 92, it is evident that the best results from the edge detection are achieved in the third study area, using the UltraCam D images over Nottingham. The edges from the third study area are sharply defined and there is little noise or any major artefacts. The results obtain from the third test site are related to the good radiometric quality of the specific images in combination with the high ground resolution. A visual comparison of the edges obtained in the three study areas indicates that the worst results are obtained over London with significant noise and background artefacts

present in the scene which was due to the poor radiometric quality of the 8bit scanned photos, the excessive scattering that took place during the image acquisition and the low ground resolution. The result from the Swiss study area were not as impressive as expected, although there are several well defined edges it seems that the low resolution of the ADS40 images for the Switzerland area deteriorated the effectiveness of the Sobel operator.

Subsequently the extracted edges are merged with the initial aerial imagery in order to enhance the discrimination of the linear features present on the optical data. This process has no impact in the geometric properties of the image since it influences only the radiometric values of the pixels coinciding with the detected edges

6.2.3. Optimised stereo matching

For the specific project the stereo matching was performed using the automatic terrain extraction tool available in Leica Photogrammetry Suite (LPS). The stereo matching algorithm implemented in LPS calculates the cross correlation coefficient between the template window and the search window, in order to identify and match conjugate points according to the following equation.

$$\rho = \frac{\sum_{i,j}[g_1(c_1,r_1)-\bar{g}_1][g_2(c_2,r_2)-\bar{g}_2]}{\sqrt{\sum_{i,j}[g_1(c_1,r_1)-\bar{g}_1]^2\sum_{i,j}[g_2(c_2,r_2)-\bar{g}_2]^2}} \quad (6.6)$$

with

$$\bar{g}_1 = \frac{1}{n}\sum_{i,j} g_1(c_1,r_1) \qquad \bar{g}_2 = \frac{1}{n}\sum_{i,j} g_2(c_2,r_2) \quad (6.7)$$

From the above equation

ρ = the correlation coefficient

g(c,r) = the DN value of the pixel (c,r)

c_1,r_1 = the pixel coordinates on the left image

n = the total number of pixels in the window

i,j = pixel index into the correlation window

Based on the cross correlation formulation the proposed method should optimize 3 basic parameters in order to provide a reliable solution in urban areas. These parameters include the size of the search window, the size of the correlation (template) window and the correlation coefficient limit.

The search window defines the search area along the X and Y direction, where X direction is equivalent to the epipolar lines (in aerial photos) in order to locate the conjugate points (in the right image) of the previously extracted interest points. The search length in the X direction is directly related to the amount of relief displacement present on the scene and thus, in urban cases a value of **20 pixels** is adequate to account for large relief displacements caused by tall buildings. The search length in the Y direction is related to the geometric configuration of the stereo pair. In a normal case of a stereo pair, consisting of aerial images, the epipolar lines pass over the same scan line between the two images. Deviations from the normal case or inaccurate results from the aerial triangulation result in an increase in the search length along the Y direction. These deviations from the normal case are also increased in the case of the pushbroom sensors. The proposed method utilizes a **search length of 5 pixels** in the Y direction, which provides adequate results over urban areas.

The correlation or template window defines the size of the area to be matched in the left image. This area defines the density of the points of interest and subsequently the density of the stereo matched points. For regions containing a high degree of topographic relief this window should be small to increase the density of the DSM. The proposed method implements a **correlation window of size 3x3** in order to extract as many conjugate points as possible for the urban area.

The correlation coefficient limit defines the correlation threshold used to determine whether or not two points are considered possible matches. This parameter should be balanced

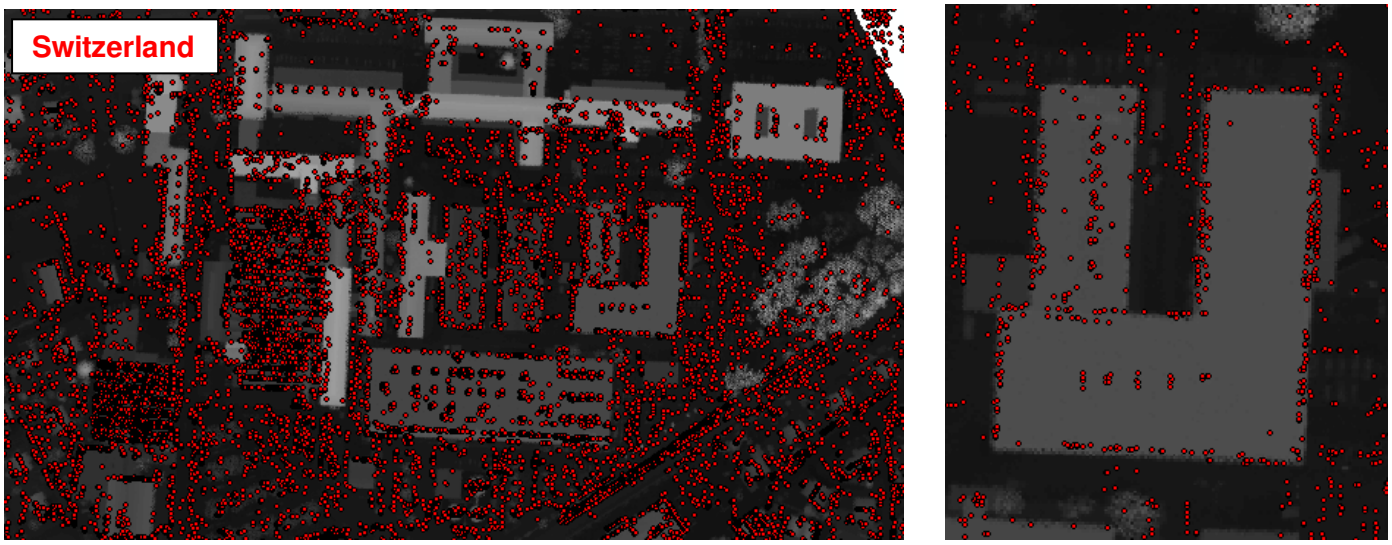
between the desired reliability and density of the extracted points and requires a certain level of user interaction. The proposed method used the value of 0.80 with acceptable results in all three study areas and could be treated as a default value. At this point it should be noted that for the specific method there is more emphasis given to the density of points rather than the reliability since any outliers introduced in this stage will be identified and removed in the step of data fusion, before adjusting the building outlines and roof polygons. Having an adequate amount of conjugate points, describing most of the building footprints and roof details, is an important consideration.

In summary, the values proposed for the parameters of the automatic stereo matching are; the search window values are 20 and 5 pixels respectively along the X and Y direction; the selected correlation window size is 3x3 and the correlation coefficient limit is 0.80. In addition, the stereo matching algorithm was applied only inside the polygons representing the study areas in order to reduce the computation requirements. The results from the stereo matching are the conjugate points in a 3D shapefile format and therefore no interpolation or DEM filtering was applied in the extracted features.

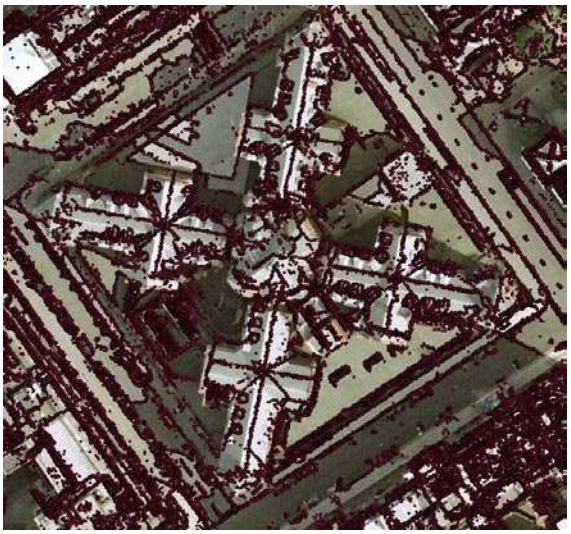
The stereo matching algorithm can additionally estimate the quality of the extracted points which provides a rough indication of the overall performance. This quality statistic is in the form of percentage per category, computed from external information, such as GCPs available for the study area or the tie points calculated during the aerial triangulation. The results from the automated quality system use the following categories:

Excellent % (1-0.85):
Good % (0.85-0.70):
Fair % (0.70-0.5):
Isolated % (less than 0.3):
Suspicious % (0.5-0.3):

The number in parenthesis is the cross correlation coefficient range calculated for the extracted tie points and it's related to the standard deviation compared to the reference data (tie or GCPs). These quality statistics provide the opportunity for automating the stereo matching algorithm and minimizing any user interaction when considering the best cross correlation coefficient to use. Further details of the iterative adjustment of the cross correlation coefficient are given in the proposed algorithm described in Appendix I. Figure 93 presents the results from the implementation of the optimized stereo matching strategy in the three study areas and a comparison against stereo matching results without the use of edges .



London – stereo matching with edges



London – stereo matching with edges

London – stereo matching without edges

Nottingham – stereo matching with edges

Nottingham – stereo matching without edges

Figure 93 Comparison of the optimised stereo matching process with and without the use of edges for the three study areas.

From figure 93 it is evident that the majority of the building outlines are adequately described by a number of extracted stereo matched points. It is also clear that the introduction of the edge enhancement stage prior to stereo matching managed to increase the density of the points representing roof details. This is evident in all the three study areas with few exceptions. As in the edge extraction process the Nottingham test site presents the best

results which are indicative from the high density of the extracted points that represent adequately most of the building outlines and roof details. In many cases the stereo matched points in the Nottingham test site represent the outline of small roof details such as pipes, chimneys and small ventilation equipment. The efficiency of the stereo matching algorithm over the Nottingham test site is a combination of the good radiometric quality and very small ground resolution of the UltraCam D images. The stereo matched points over London also present better results than expected, considering the lower resolution and poor radiometric quality from the film scanning. In most cases roof outlines are adequately represented by a number of stereo matched points but the level of detail for the inner roof structures is limited to larger structures, larger than chimneys and small dormers. The London test site presents an additional challenge, that is in many cases the roof texture and material resembles the texture of the pavements and road tarmac that makes the differentiation of breaklines and, as an extension, the extraction of stereo points very difficult. This is evident in some cases when certain facades of smaller gable type buildings are not well represented. In general the results are considered acceptable and could be subsequently used for the next stages of adjusting the initial building footprints. The stereo matching algorithm can additionally estimate the quality of the extracted points which provides a rough indication of the overall performance and as described in the next section they are important for the quality check loop of the optimized stereo matching workflow. The stereo matching quality output is given below (table 22) and it highlights the general observations that were made before about the efficiency of the stereo matching process over the test sites.

Switzerland	London	Nottingham
Excellent (1-0.85): 57.5 %	Excellent (1-0.85): 60.3 %	Excellent (1-0.85): 65.9 %
Good (0.85-0.70): 29.5 %	Good (0.85-0.70): 28.4%	Good (0.85-0.70): 30.2%
Fair (0.70-0.5): 0%	Fair (0.70-0.5): 0%	Fair (0.70-0.5): 0%
Isolated : 0 %	Isolated : 0%	Isolated: 0%
Suspicious : 13 %	Suspicious : 11.3 %	Suspicious : 3.9 %

Table 22 The results of the quality statistics from the implementation of the stereo matching process over the three test sites

The implementation of the quality check loop introduced in the proposed algorithm is shown in the next section.

6.2.4. Discussion for the stage of semi-automated feature extraction from optical data

Based on the description of the proposed method for feature extraction in section 6.2.3., a generic algorithm can be formed as shown in Appendix I (algorithm 1). The algorithm also indicates which software was used and whether any customization of third party software has taken place.

The proposed method for feature extraction produces an overall acceptable result since most of the linear features were represented by at least five conjugate points. Having at least five points would ensure in most cases that a linear segment can be fitted subsequently. In addition the method is relatively straightforward and easy to implement on digital airborne images or photographs from frame cameras. Furthermore, the method has the potential for being fully automated since in most operations user interaction is not

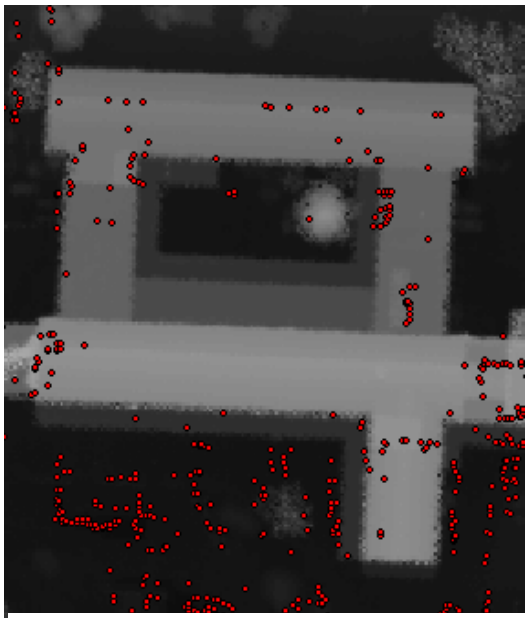


Figure 94 An example of inadequate extraction of stereo matched points for adjusting the building outlines over the Switzerland test site

required. Even at the tedious process of optimizing the parameters of the stereo matcher, the incorporation of the quality statistics for checking the results and accordingly adjusting the parameters can minimize the need of user interaction.

Nevertheless this step introduces a number of disadvantages because of the assumptions inherent in the process. One of the critical characteristics is the assumption that the density of the points is more important than the vertical accuracy, in order to have an adequate number of points representing as many linear features as possible. Based on this assumption the method accepts the selection of image pairs with small convergence angles in case of ADS40 data (nadir and 14^{o} backward look angle) and also a relatively small cross correlation coefficient for all the data sources available. This assumption introduces another problem, that the reduced vertical accuracy of the extracted points will also affect the planimetric position of the features and thus decrease the accuracy of the building outline adjustment. However blunders and large errors in the stereo matching will, in most cases, be identified and removed from the quality statistics check loop as shown in algorithm 1 which acts as the decision making tool for assessing the quality of the conjugate points.

Finally despite the overall acceptable results of the specific step, there are still few building outlines without a significant number of extracted points (figure 94). This problem is introduced due to the lack of radiometric differences present in some building outlines, as a result of shadows, noise in the CCD and the angle of view with respect to the direction of the illumination (BRDF).

6.3. TIN Generalization and Very Dense DSM generation

As can be seen in the subset flowchart in figure 95 the process described in this section is triggered only when there is low density LiDAR or no LiDAR data available. At this point it should be clarified that the proposed workflow will classify a LiDAR dataset as ''coarse'' if after the density analysis process the raw point cloud has less the 3 points per square meter. In contrast LiDAR data with more the 3 pt / sq.m are classified as high density. This distinction is critical for the reliability of the workflow and the threshold is based on the minimum number of points required to fit a plane on the point cloud so that it is possible to reconstruct roof details as small as $1m^2$. Hence the smallest geometric detail this workflow is aiming to reconstruct is $1m^2$. This threshold is also reinforced by the results from the implementation of the proposed workflow in the study areas where the LiDAR data had less than 3 points per square meter density.

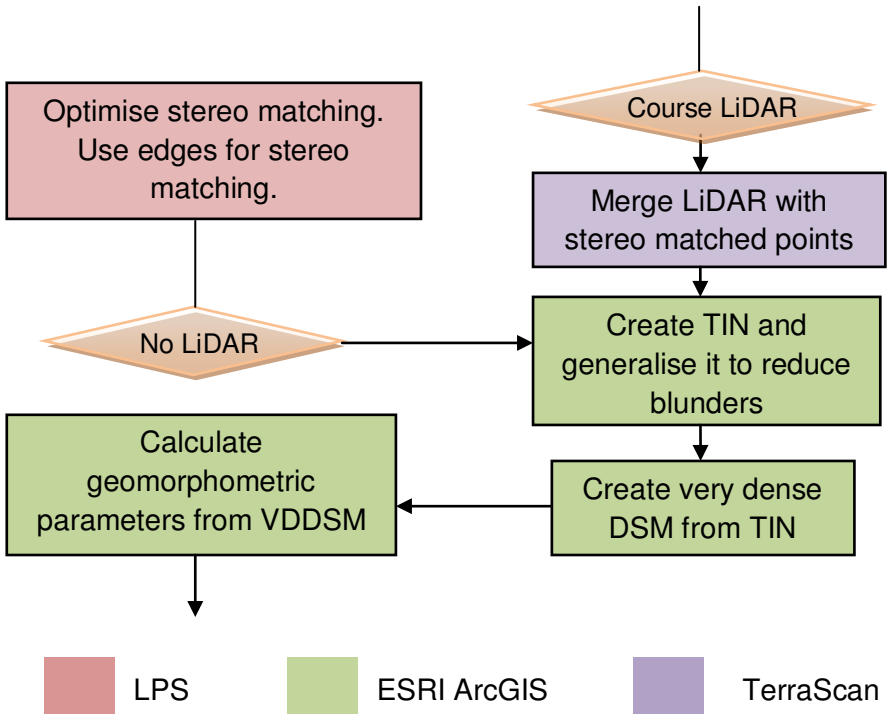


Figure 95 Subset of the overall workflow indicating the proposed processing steps when LiDAR data is not available or dense enough.

Once the stereo matched points are extracted as described in the previous section the algorithm creates a density map to evaluate the density of the LiDAR data (if available). If the density is less than 3 points per square meter then the LiDAR data are merged together with the stereo matched points. Merging of the stereo matched points with the LiDAR took place in the second study areas in London since the LiDAR data only had 1 point per square meter density (figure 96).

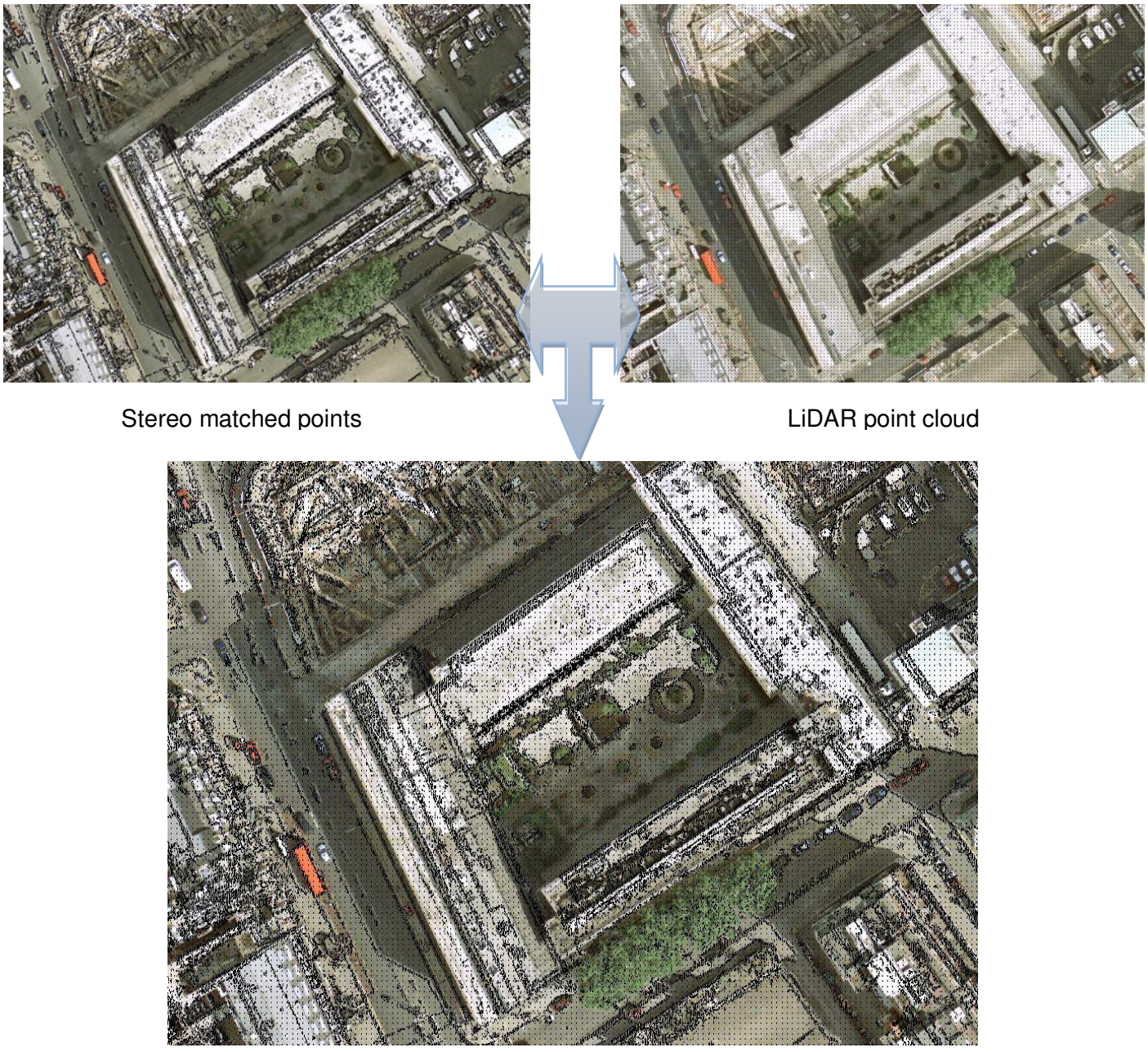


Figure 96 Merging of LiDAR with stereo matched points in the London test site, all superimposed on the orthoimage.

The merged point cloud is subsequently used to create a TIN surface using the Delaunay Triangulation. In the absence of LiDAR data as was the case in the third study area in Nottingham the TIN surface is created using only the stereo matched points. In order to minimize the number of blunders and small spikes appearing due to noise and low radiometric quality the TIN surface is generalized to reduce the number of vertices present but without filtering breaklines of the roof details (Figure 97).

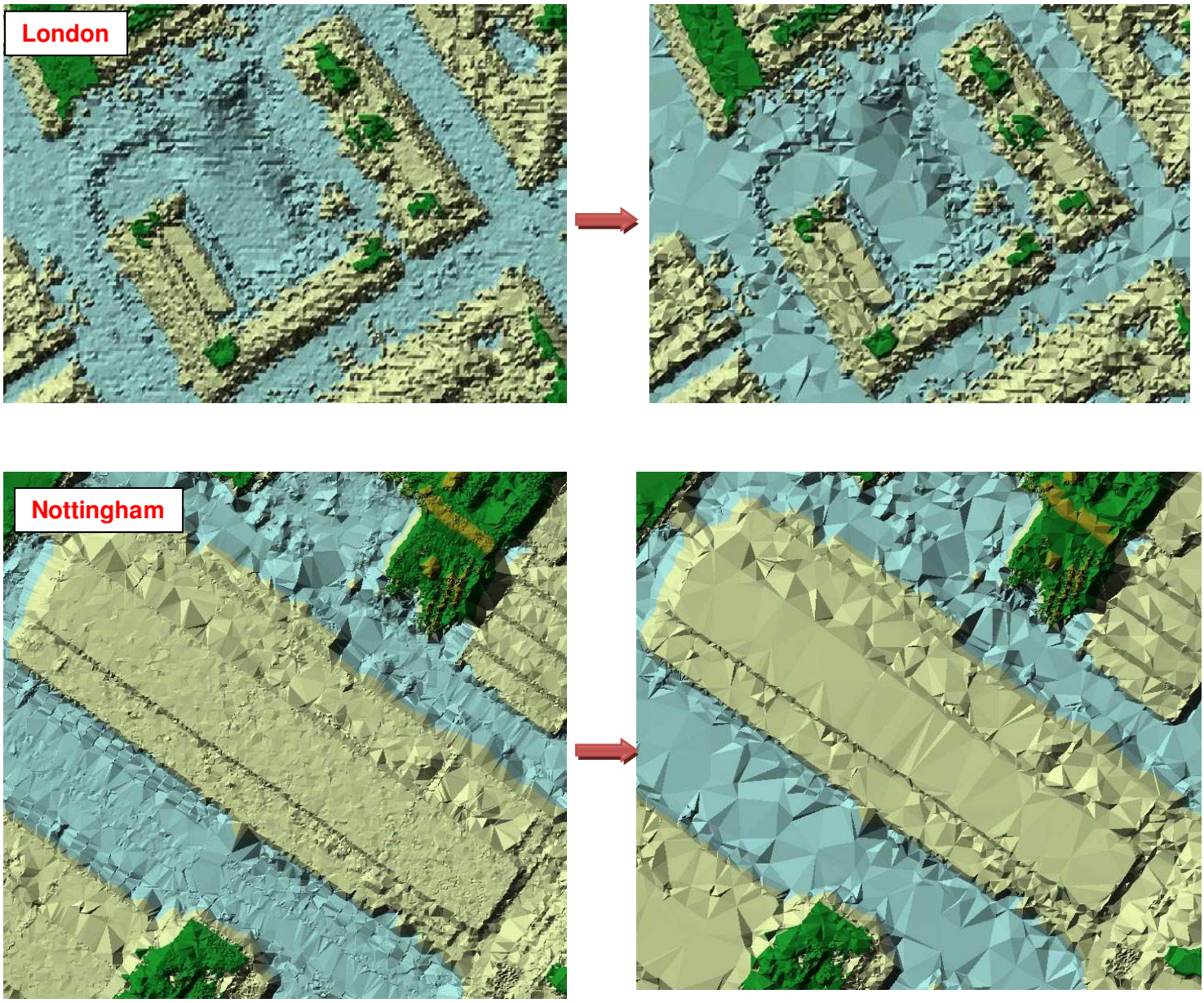


Figure 97 Initial TIN (left) and the generalized TIN (right) from the second and third study areas.

From figure 97 it is evident that the superior quality and performance of the stereo matching workflow on the UltraCam D images has produced a stereo matched point cloud with few artefacts, noise or even spikes. After the generalization, the TIN in the Nottingham area is in most cases free from any small artefacts on the roof structures while preserving most of the roof breaklines. In contrast the TIN over London is the result of the union between the stereo matched points and the 1m LiDAR data and is easily identifiable that there is significant noise and several artefacts in the stereo matched points. Once the initial TIN is generalized most of the excessive noise is reduced significantly but the overall quality of the generalized TIN is still affected in some cases by the remaining artefacts. As a result these artefacts are expected to have an adverse effect during the extraction of roof breaklines.

The TIN generalization process follows a drop heuristic method as introduced by Lee (1989). The process takes a TIN as an input, and iteratively rejects one vertex at a time to create a TIN with less vertices. It can also be applied to a grid of points as input if we regard it to be a triangulated regular grid. If a vertex is excluded, the incident edges are also rejected and a polygon appears in the subdivision. To get back to a triangulation, the polygon entity is triangulated using the Delaunay triangulation. This will make certain that if the algorithm starts with a Delaunay triangulation, then after each iteration we'll still end up using a Delaunay triangulation.

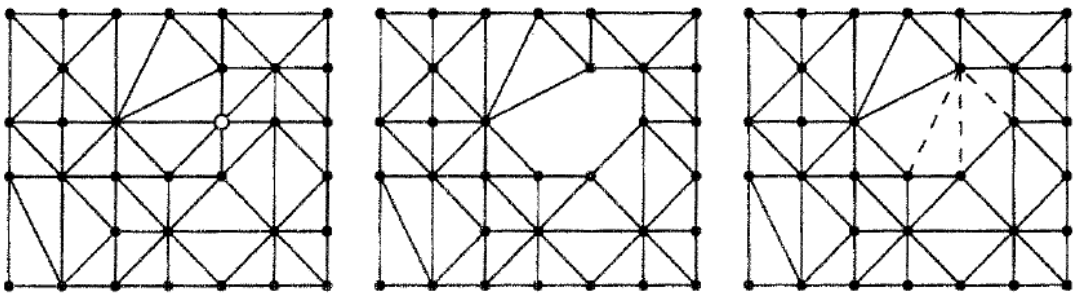


Figure 98 Left, a TIN with one vertex represented by a white circle. Middle, the polygon that is created when the specific vertex is excluded. Right, a Delaunay triangulation of the polygon

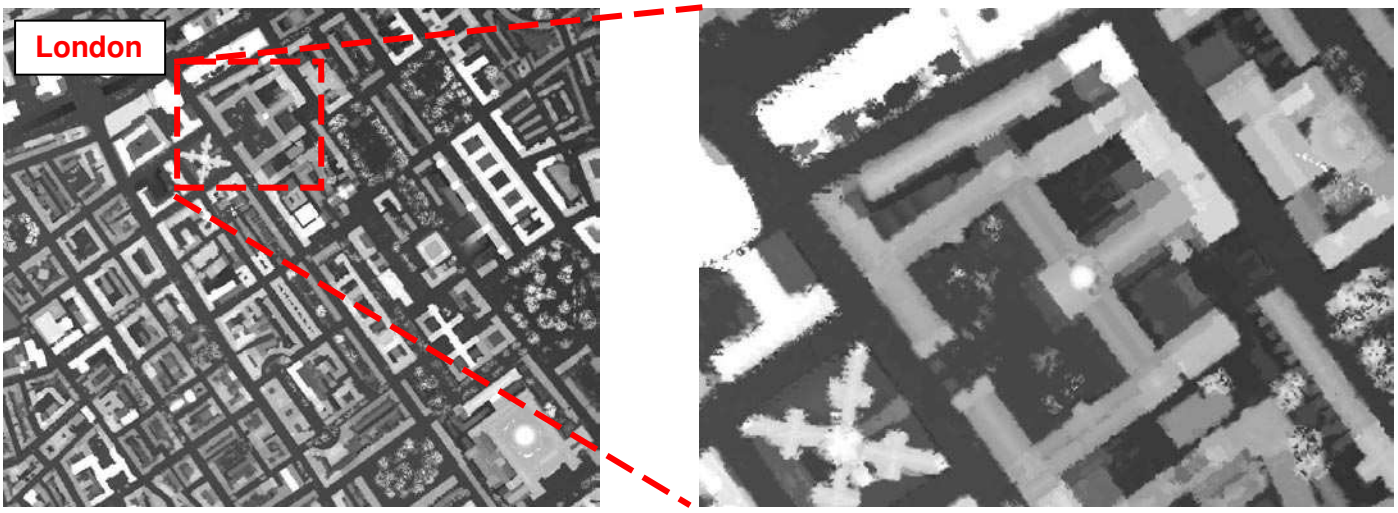
To identify which vertex should be excluded, each vertex is temporarily isolated and the appearing polygonal shape is triangulated (Figure 98). Then the process calculates the

vertical distance between the isolated point and the new, generalised TIN. The excluded vertex is located within one of the new triangles in the polygon.

This vertical distance represents essentially the error introduced by the deletion. Once we know the error that would be introduced, we insert the excluded vertex back to the TIN and temporarily isolate another vertex. After we have done so for a number of vertices, the algorithm identifies the one for which the vertical distance (or the computed error) is smallest and permanently excludes it. The process is repeated until the created error is more than the predefined allowed error by the user.

For the semiautomatic identification of the error threshold when trying to generalize a TIN surface, the proposed workflow utilizes the vertical ground residuals of the check points as calculated from the AT process. This provides generalized TINs with most of the mismatches and blunder points successfully filtered as illustrated in figure 97.

The generalized TIN is then converted to a very dense digital surface model (VDDSM) using an inverse distance weighted, quadratic polynomial interpolation algorithm as introduced by Franke and Nielson (1980). The VDDSMs have a sampling size of 0.20m and will be used subsequently for performing the building hypothesis and building detection (figure 99).



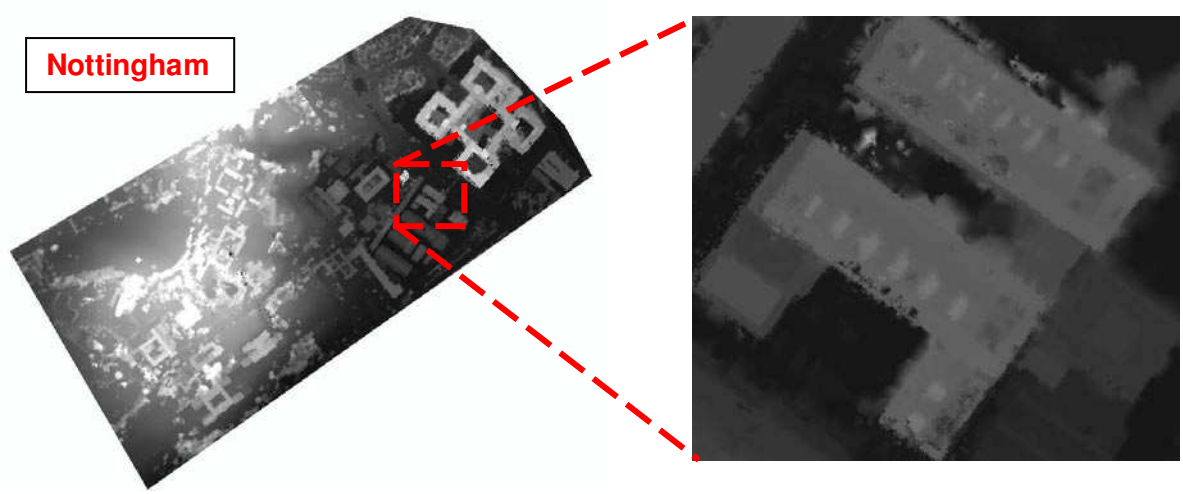


Figure 99 The generated VDDSM for the second and third study areas from the generalized TINs

Once the VDDSM is created the proposed workflow calculates two basic geomorphometric parameters, slope and aspect (slope pointing direction). The two parameters are calculated for all the grid cells in the VDDSM and will be used during the stage of extracting 3D linear primitives of roof structures. If the partial derivatives of elevation (H) along the east (x) and the north (y) direction are known then slope and aspect (slope pointing direction) are computed from Burrough, 1987.

$$slope = \sqrt{\left(\frac{dH}{dx}\right)^2 + \left(\frac{dH}{dy}\right)^2} \tag{6.8}$$

$$Aspect = arc\tan\left(\frac{-dH/dy}{dH/dx}\right) \tag{6.9}$$

Figure 100 illustrates the calculated slope and aspect from the VDDSMs for the two study areas.

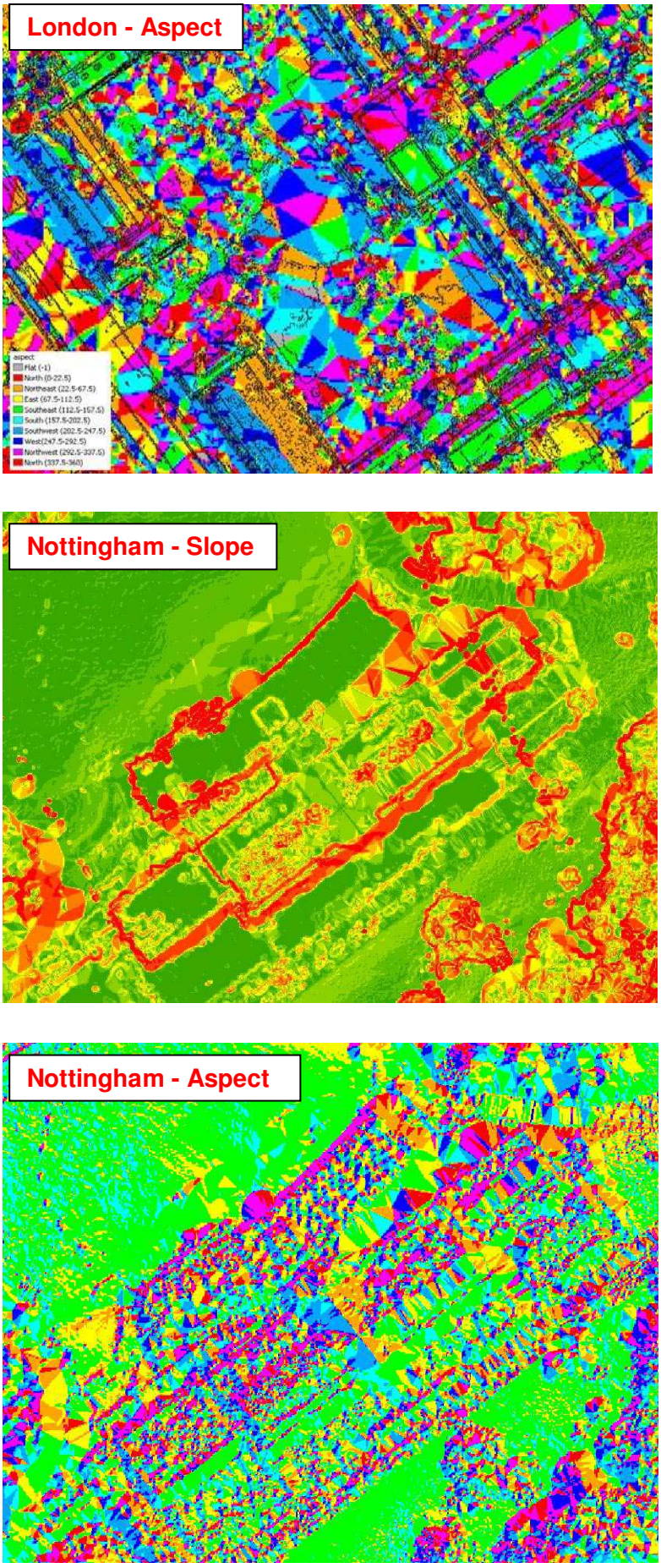


Figure 100 The calculated slope and aspect surfaces from the VDDSMs of the second and third study areas

The slope and aspect is critical for the reconstructions of linear features representing roof shapes because the stereo matched points that form these linear segments will be grouped in clusters of points with the same aspect and slope. The generic algorithm of the TIN generalization and VDDSM creation as described before is given in Appendix I (Algorithm 2).

6.4. Building detection and classification of LiDAR data and VDDSMs

This section provides a detailed description of the proposed workflow for semi-automatic building detection. Even though these steps seem trivial in the overall workflow (figure 101) this is the most complicated stage of the entire process and certain degree of user interaction is still required to optimize certain parameters.

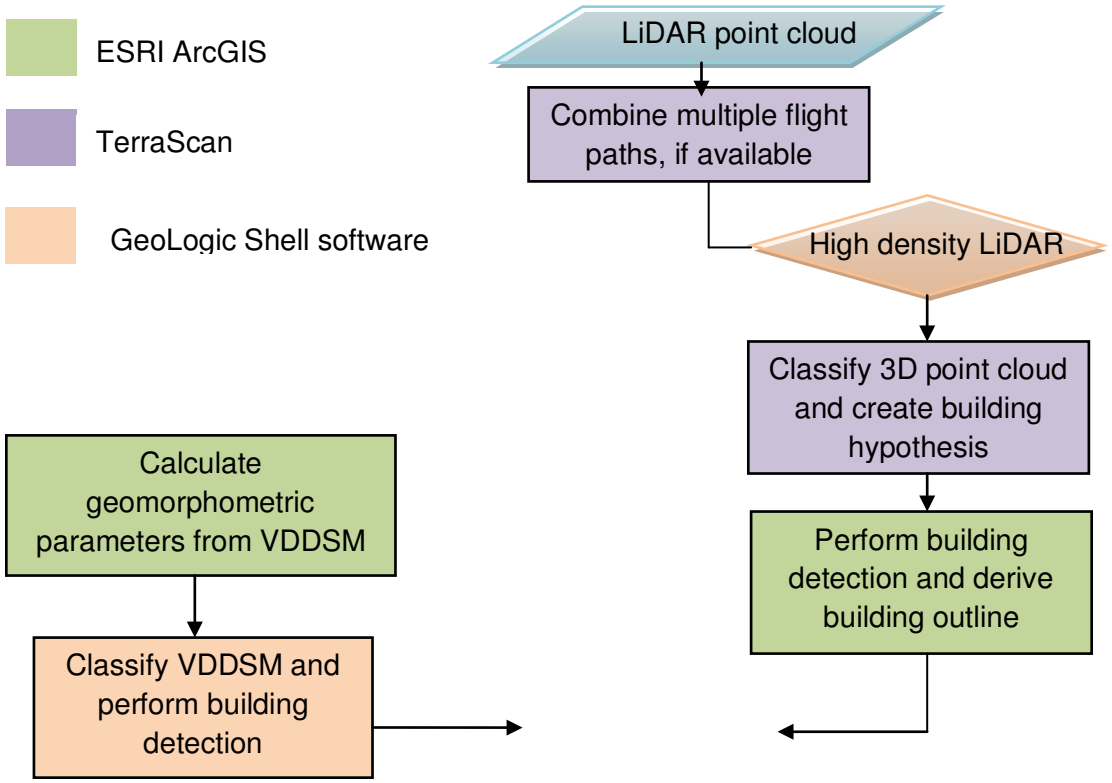


Figure 101 Subset from the overall workflow indicating the building detection and classification of either LiDAR or VDDSMs

As shown in figure 101 the process of building detection follows two alternative routes. As discussed in the previous section it depends whether high density LiDAR are available or

not. The differentiation is based on the density of the point cloud, exactly as described in section 6.3. Due to the nature of the two different input datasets (VDDSM versus high density LiDAR) the proposed workflow utilizes alternative classification techniques in order to optimize the building detection process.

There are two main subsections that describe the proposed building detection method when a VDDSM is available and an alternative method to use when high density LiDAR is available. The final extracted information in both cases are building outlines or building polygons. As shown in figure 101 the process starts initially by merging all the available LiDAR point clouds since multiple flight paths could have been used.

6.4.1. Classification and building detection using high density LiDAR data

This section provides a description of the workflow, designed to address the issues of classifying the high density LiDAR data and detecting building entities. From the study areas and the data available, only the first study area was classified as having high density LiDAR and therefore this workflow was implemented only in the Switzerland study area. The diagram in figure 102 provides a summary of the individual processes implemented to detect buildings in the Switzerland study area.

The proposed method does not employ any information derived from the multispectral imagery, not even for delineating tree canopies. Many researchers have proposed tree detection undertaken using multispectral information and calculating Normalized Difference Vegetation Index (NDVI), (Chen et al. 2004). A reason for not applying classification techniques using the near infrared band collected from few digital sensors is that in this case the method would not be suitable for use with traditional film aerial photographs. Hence at this point a textural classifier would be more appropriate since it can be applied with LiDAR and VDDSM. Despite the potential use of a textural classifier there is a major disadvantage. The methods of textural classifiers and multispectral classification are extremely dependent

upon seasonal conditions, in other words if the tree canopies are in a leaf-on or leaf-off condition. In a leaf-off condition it is obvious that both methods will fail to delineate tree canopies, since there are neither textural differences nor high radiometric responses in the near infrared bands.

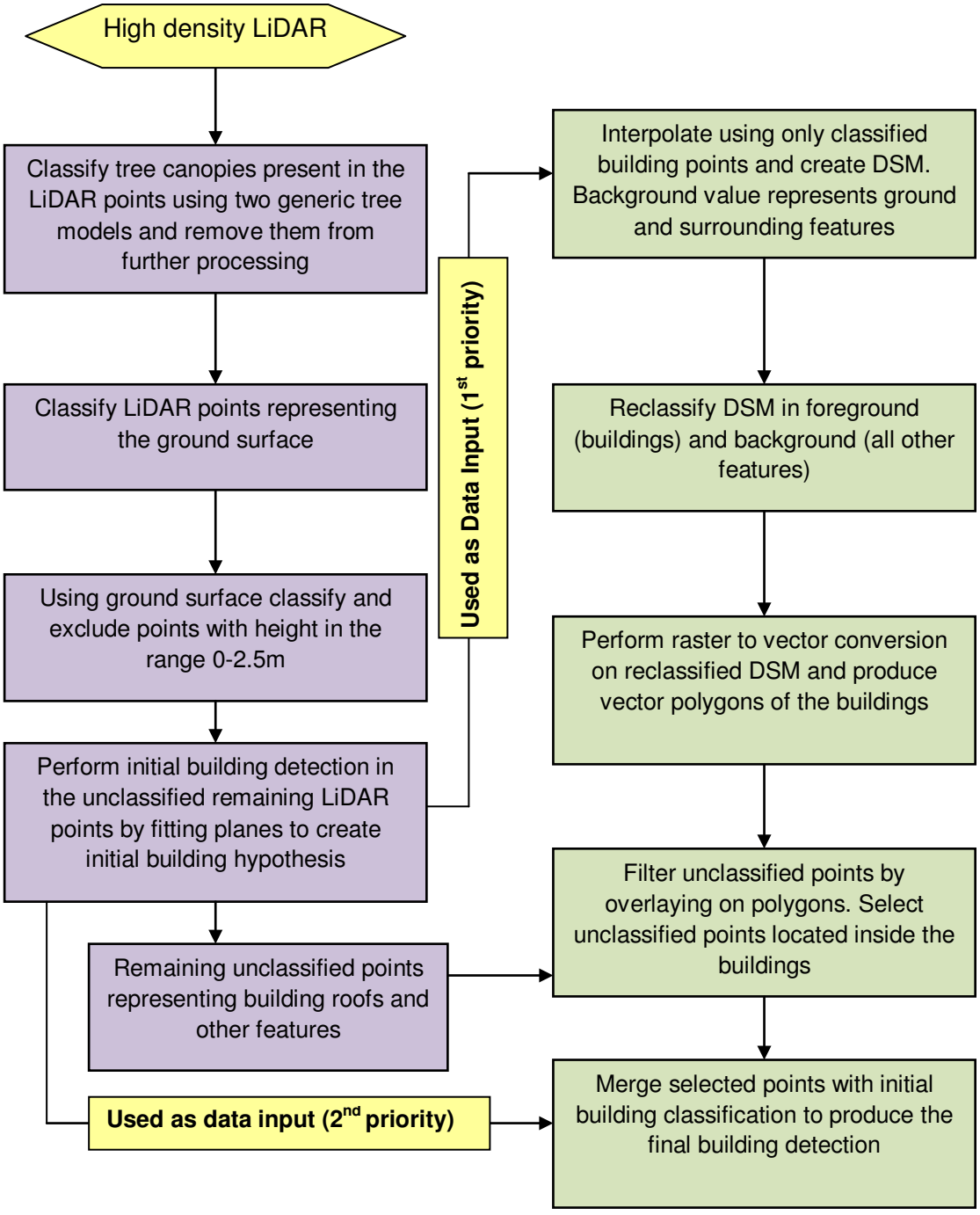


Figure 102 Diagram indicating the overall workflow of the building extraction process from high density LiDAR

The disadvantages of the textural and multispectral classifier mentioned before are also evident when the optical and LiDAR data have been acquired in different seasons. This was the case for the data provided in the specific project for the first two tests sites. For example, for the first test site the ADS40 imagery was collected with the trees in a leaf-off condition. Therefore the proposed method is designed, taking into account the above considerations to present reliable results even if the tree canopies in the LiDAR data are in a leaf-on or leaf-off condition.

As described above before the implementation of the classification workflow any LiDAR data acquired from different flight paths are merged together. The first study area in Switzerland had two LiDAR point clouds available which are merged together. The overlapping region between the two flight paths is located over the study area. This procedure is performed in order to increase the density of the point cloud from 3 points per sq.m to 5 points per sq.m. The advantage of having a denser point cloud is the ability to reconstruct even smaller roof details, during the plane fitting procedure, which results a detailed polyhedral model. The entire procedure of the proposed method was repeated using only a single flight path, in an attempt to quantitatively assess the influence of the decreased density in the building reconstruction stage. The resulting building models from the single path LiDAR are given in chapter 7. The resulted surface model from the combination of the two LiDAR point clouds is illustrated in figure 103, where a colour-coded, shaded relief image is indicated.

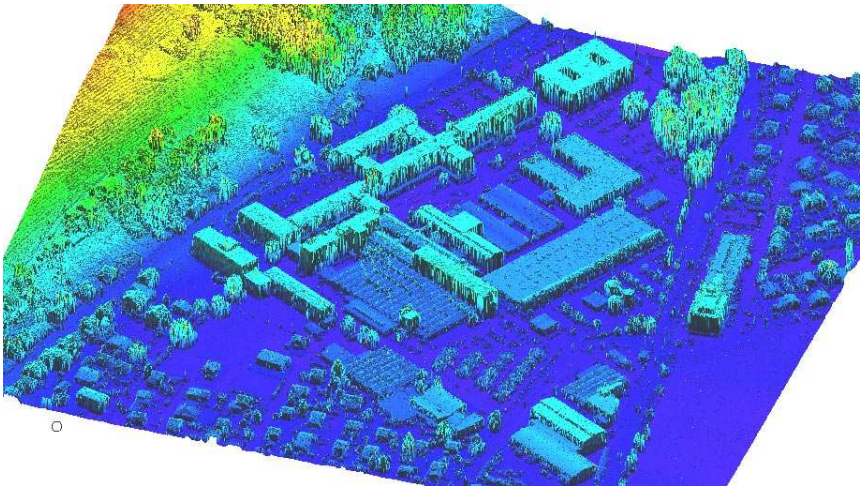


Figure 103 Perspective scene of the combined LiDAR data, visualized as a colour coded shaded relief map.

6.4.1.1. Delineating tree canopies from high density LiDAR

This stage is implemented in order to apply an initial delineation of the tree canopies. The primary purpose is to perform a rough detection of large trees without necessarily detecting all types of vegetation, present on the study area. For the tree detection, two generic tree models are utilized in order to scan the entire point cloud. The algorithm uses the tree models as geometric primitives to identify clusters of points that match the shape of the primitives through a correlation process. This process was carried out using the TerraScan software package. The geometric shape of the tree models used in the proposed workflow is indicated in figure 104.

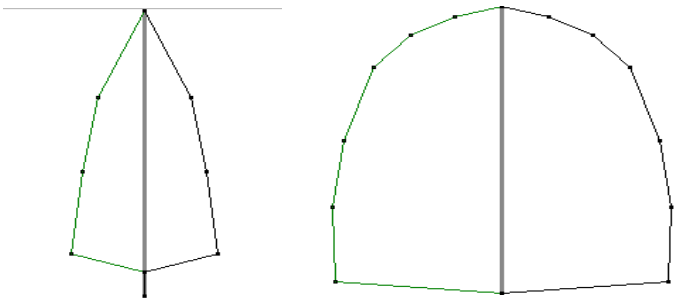


Figure 104 Geometric shape of the generic tree models used to scan over the entire point cloud

There are three parameters associated with the specific procedure that should be optimized in order to achieve acceptable results. These parameters specify the shape of the tree models and include the minimum and maximum height as well as the width variation in percentage. The width variation factor determines the width of the tree model as a function of the height. Although one might consider that this stage requires pre-existing knowledge of the tree heights in the scene, it's not a prerequisite since the minimum and maximum range can be defined as broad as possible, in order to encompass most of the potential sizes present at the scene. The algorithm essentially scans the point cloud using successive increments between the defined tree height ranges, to match clusters of points that resemble the geometric shapes shown in figure 104. As noted before, for each height the width is adjusted accordingly as a percentage of the search height. Hence the increase of the height

range has a negative impact only on the computational requirements, since the search range is increased. Once the scan has been completed, using the generic tree models, the algorithm labels the clusters of points that match the models to the high vegetation category (figure 105).

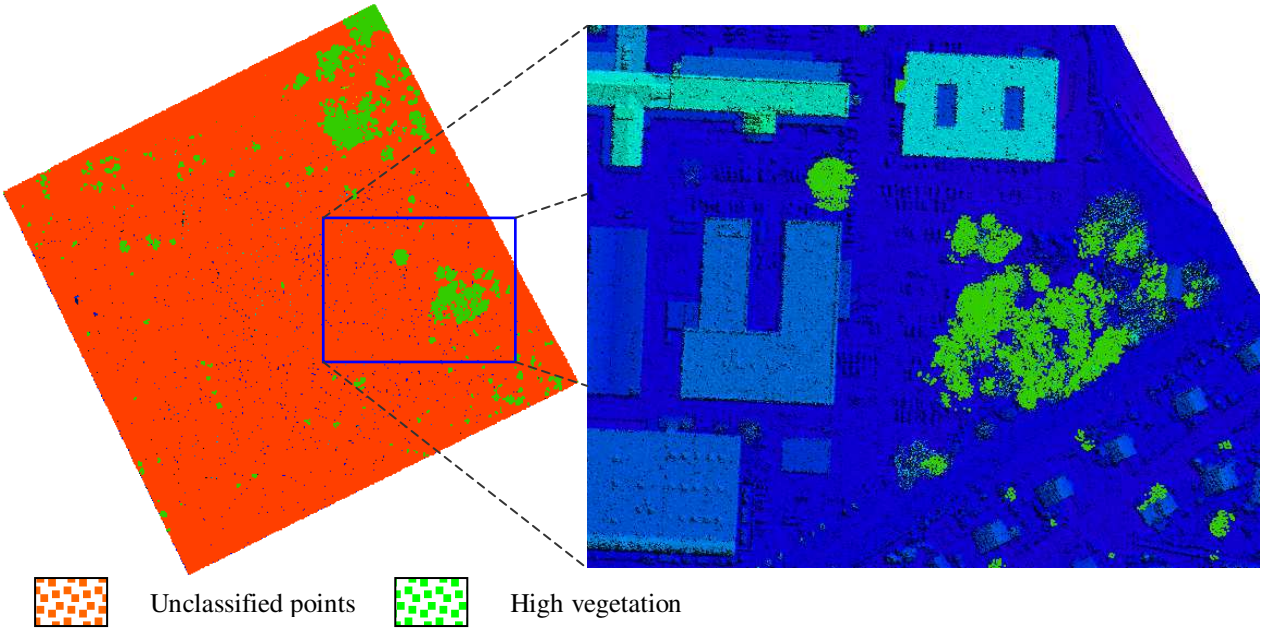


Figure 105 Results from the tree detection using the two tree models over the LiDAR point cloud. Top view of the study area (left), TIN model with tree points superimposed (right)

The proposed workflow utilizes a minimum tree height of 2.5m since features lower than 2.5m will be filtered in the subsequent steps. The maximum height for both tree models was selected equal to 40m in order to encompass a broad range of tree heights. The width percentage was defined equal to 30%. These two tree models with the previous parameters can yield satisfactory results in many different situations and can be treated as default parameters in the design of the algorithm. The results illustrated in figure 105 provide an indication of the effectiveness of the method, with most of the tree canopies successfully detected in the study area.

Nevertheless, there are occasions were the two tree models are inadequate to detect every individual tree, present in the study area (figure 106). This is a function of the limited number of tree models used and the non optimization of the width percentage parameter. The width

percentage is a critical factor, but selecting the appropriate percentage is a tedious process that is very difficult to be automated. Instead the method suggests the use of 30% as a general guideline.

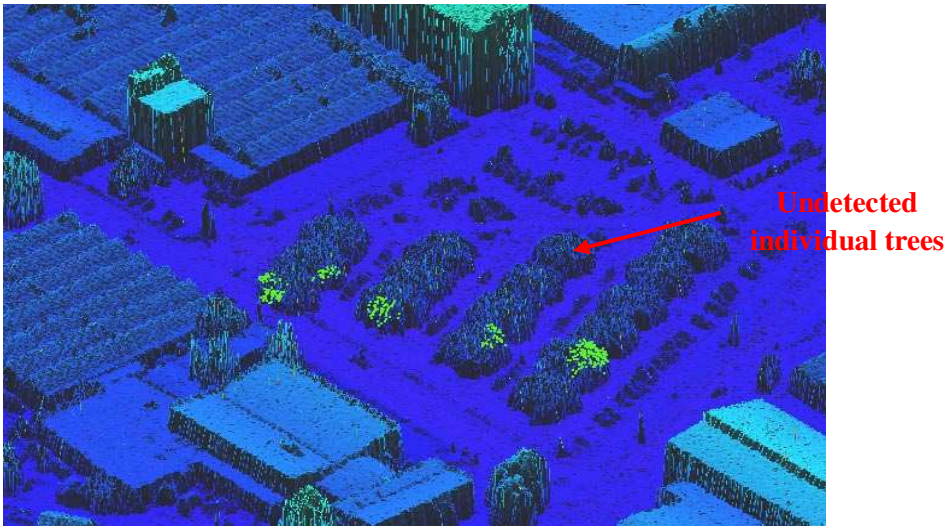


Figure 106 Example of undetected individual trees resent in the study area

Despite the existence of undetected individual trees the method successfully filters tree canopies which are the crucial aspect of this stage. The main reason is that even if individual trees remain in the LiDAR point cloud, they can be filtered when applying a minimum plane size criterion during the building detection process. If tree canopies are not detected they can introduce significant problems during the building detection.

6.4.1.2. Classifying ground surface (DTM) from high density LiDAR data

This stage of classifying LiDAR points belonging to the ground surface is an important step that is required before the building detection process, as well as for classifying points with a specified relative height. The classified ground surface is also crucial in assigning elevation to the projected building outlines during the building reconstruction process.

The proposed method employs TerraScan's algorithm that detects ground points by iteratively building a triangulated surface model. There are four parameters that must be

optimized during this procedure, which include the maximum building size, the maximum terrain angle, maximum iteration angle and distance values (see below). The maximum building size controls the number of initial points selected for the generation of the initial TIN model. For the proposed workflow a maximum building size of 300m was selected as the most appropriate value. According to this value the algorithm will assume that any area of size 300x300m will have at least one ground point and will select the point with the lowest elevation within this area. The other factor that affects the initial selection of ground points from the initial TIN is the maximum terrain angle. The maximum terrain angle is used to restrict the selection of initial points, if the slope between them exceeds the defined value. The proposed workflow utilizes a maximum terrain angle of 60°.

After the generation of the initial TIN, the algorithm iteratively adds new points to the existing TIN model. The iterative selection of new ground points uses as conditions the maximum iteration angle and maximum iteration distance (TerraScan 2004). The **iteration angle** is the maximum angle between a point, its projection on triangle plane and closest triangle node. **Iteration distance** prevents abrupt vertical changes when the triangles of the TIN are large (figure 107). These two parameters prevent buildings from being classified as ground surface.

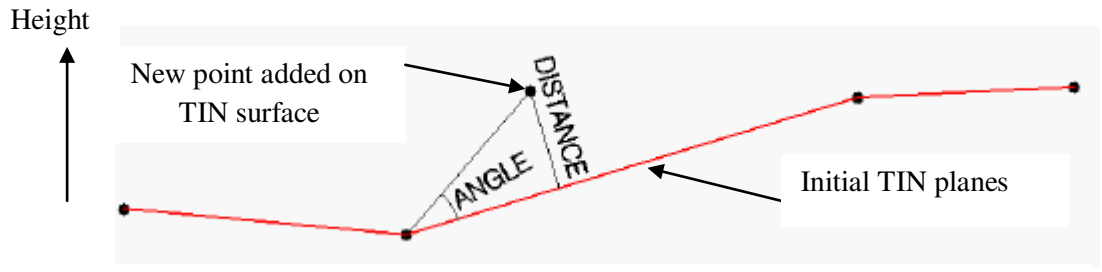


Figure 107 Diagram illustrating the iterative selection of new points at the ground surface

The proposed method utilized a maximum iteration angle equal to 5° and a maximum iteration distance of 1m which were determined after a trial and error process. The resultant ground surface from the filtered points is illustrated in figure 108.

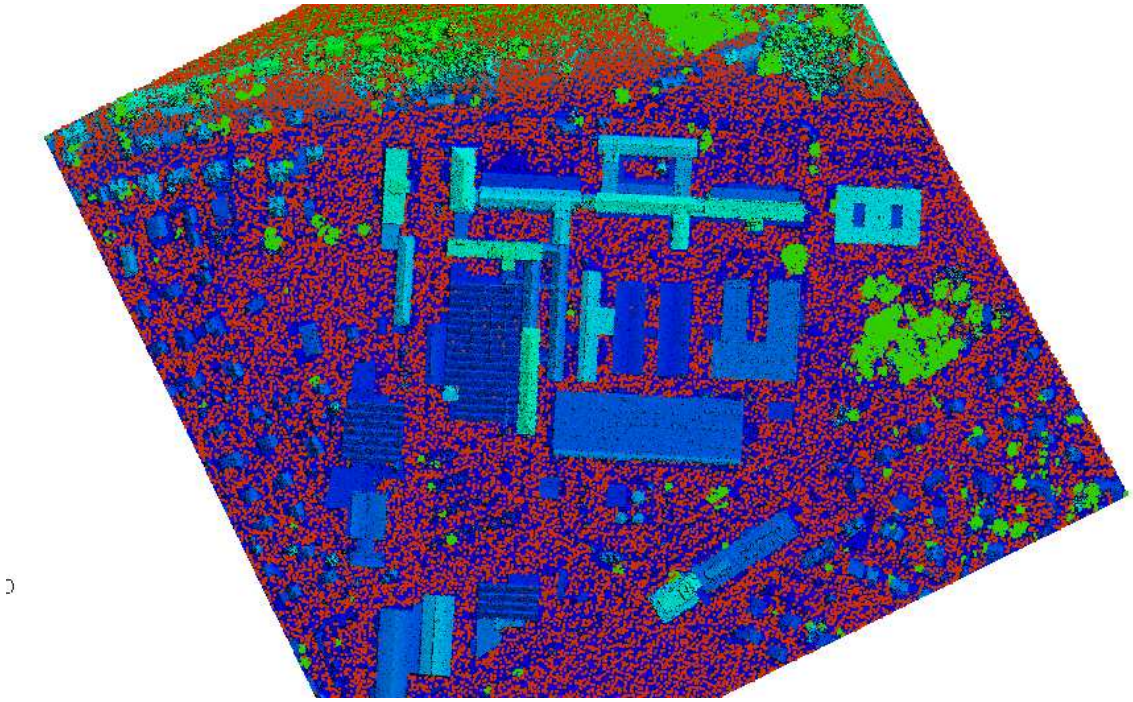


Figure 108 Resultant ground surface points (orange points) over the 1st study area, classified using the iterative selection algorithm

Classifying the ground surface requires a certain amount of user interaction that can be minimized if assumptions are integrated in the algorithm, as described below. The proposed algorithm for estimating the maximum building size employs the value of 300m as a rule of thumb. The advantage using a large value is that it's suitable for smaller as well as for industrial buildings because the algorithm can populate the TIN with the iterative approach as depicted in figure 97. In contrast the maximum terrain angle is more crucial and at the same time very difficult to automatically estimate. The difficulty arises because the value must be balanced in such a way, to take into account the topography present in the scene but at the same time excluding buildings from further calculation. In most cases this requires the user to have prior knowledge of the topography and the technical characteristics of the project (flightpath, density, altitude of sensor) or as an alternative to experiment with the data and optimize the value. In an attempt to simplify the problem the proposed algorithm assumes that the steepness of the slope, present at the building facades, is only related with the density of the LiDAR point cloud. This assumption is oversimplified because in fact the slope of the building facades are related with the altitude of the sensor, the look angle with

relation to the building height and other variables, but it serves the purpose of approaching a semi-automated solution.

Therefore the value of 60^{o} as the maximum terrain angle determined is treated as being representative for any LiDAR point cloud with density of 4 points/m^2. Hence point clouds with coarser or higher density are adjusted accordingly with respect to a linear relationship. The density of the point cloud can also be determined automatically by calculating a "**density map**". The density map is a popular function among GIS packages that use a kernel window of size 1 by 1m to scan the point cloud and create a raster image of the density. Each pixel of the raster image represents the number of enclosed points. The overall density is then calculated from the average value of the floating point values, in each pixel.

The **maximum iteration angle and iteration distance** can be treated as a rule of thumb, in the proposed algorithm, since the values are very small and will avoid detecting buildings in almost every situation unrelated to the density of the LiDAR point cloud, but at the same time populate the initial TIN model. Because the selected values are relatively small there are still remaining unclassified ground surfaces (figure 109).

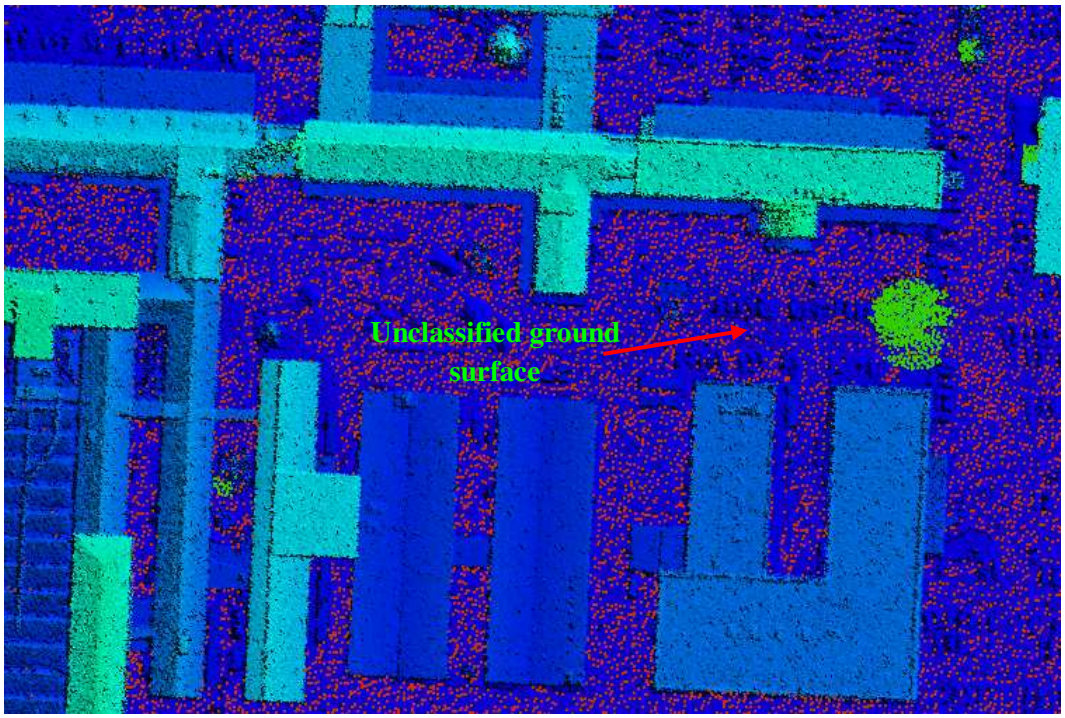


Figure 109 Classified points (red) representing the ground surface, overlaid on the shaded relief map of the LIDAR data and unclassified ground regions

Despite the presence of unclassified points as ground surface, this is not a major concern since this will be solved in the following stage. Figure 109 illustrates the efficiency of the method since the proposed process was able to distinguish cars and other small features from the ground surface.

6.4.1.3. Classifying low vegetation and background features from high density LiDAR data

Points not related with the building entities, such as parked cars and low vegetation, can be effectively removed by applying a height range filter above the ground surface. As a representative value the range of 0-2.5m was selected that can be used for most applications. This step will also classify the points representing the ground surface that weren't classified from the previous stage and exclude them from further calculation. This process uses the initial ground model to create a temporary TIN model and then compare the unclassified points to estimate the height from the initial TIN model. Features and low vegetation with height equal or less than 2.5 meters are filtered out. This classification step was also implemented with the use of TerraScan software package. The results from the above procedure are illustrated in figure 110.

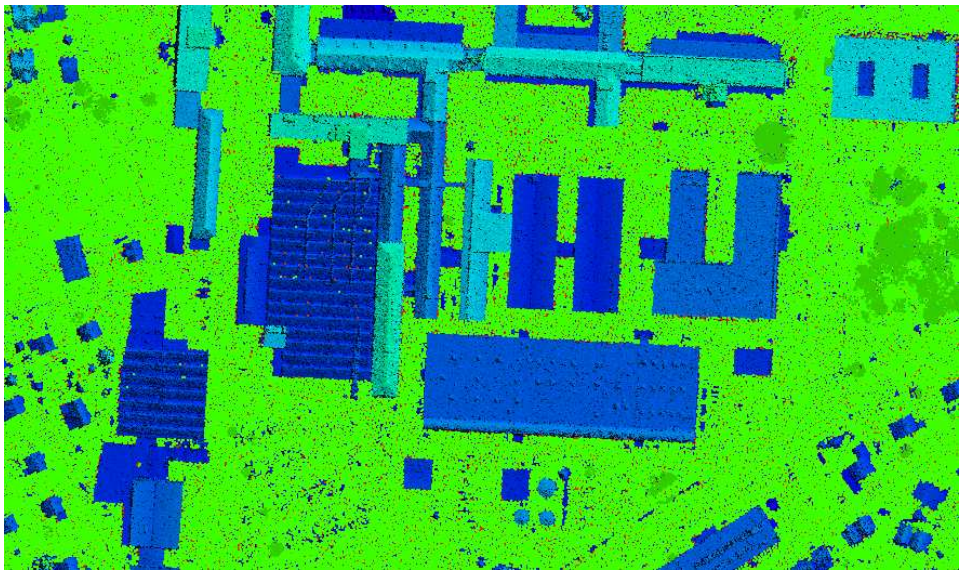


Figure 110 Classified features and low vegetation with height in the range of 0-2.5m, superimposed on the TIN model for the study area

6.4.1.4. Generating building hypothesis from high density LiDAR

With the completion of the previous stage, ideally the only remaining points would correspond to the buildings present on the scene. Nevertheless, as discussed previously, the stage of tree detection in most cases will not be able to filter all the individual trees. As a consequence the unclassified points at this stage include remaining individual tree crowns with height greater than 2.5m as indicated in figure 111.

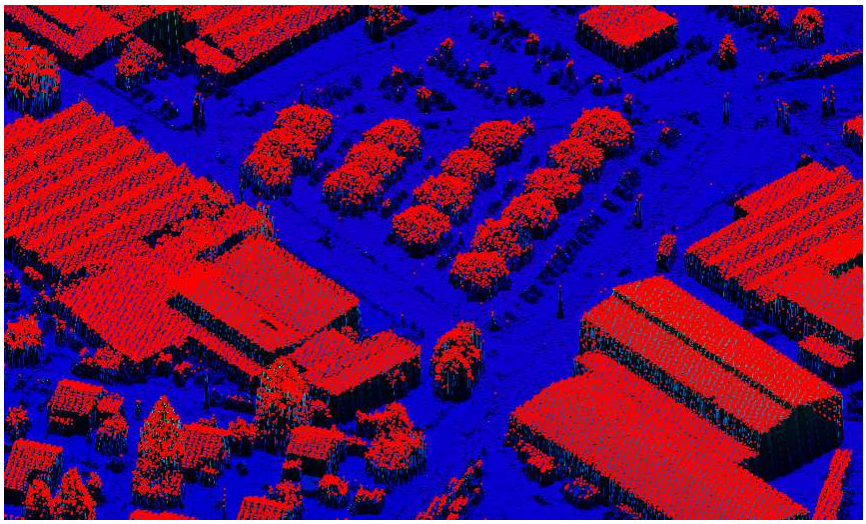


Figure 111 Remaining unclassified tree crowns and points representing building roof tops.

The building classification algorithm is based on a plane fitting method that will be further described in the building reconstruction section. The plane fitting algorithm is applied inside the void areas introduced from the classified ground points. At this point it should be noted that before the procedure of building detection, the proposed method merges the ground points with the low vegetation (height up to 2.5m) points in order to minimize void regions not related to building entities. Several conditions are introduced at this stage to optimize the performance of the algorithm. The **minimum building** parameter is used in order to avoid performing the plane fitting in holes, with areas smaller than the one defined for the minimum building criterion. This value should be large enough to bypass individual remaining trees but at the same time, taking into account smaller cottage style houses that might be present in the scene.

The proposed method implements a minimum building size of 40m^2 with acceptable results that can be treated as the default value for many different situations since it is small enough to detect a variety of building sizes. This value assumes that the tree segmentation step has successfully filtered tree canopies from the study area and only individual trees remain in the unclassified LiDAR points. Another criterion used, during the selection of void areas on the ground, is the **maximum building size**. Although this condition is used for avoiding extremely large gaps in the ground surface, in the proposed method every large void area is related to a building entity. Therefore the proposed method utilized a maximum building size of 30000m^2.

During the plane fitting, on the unclassified points, there are four conditions that must be addressed. The **minimum detail parameter** defines the minimum size of the desired plane. This parameter can be automatically determined with respect to the density of the LiDAR point cloud. Considering that a plane requires at least three LiDAR points, the minimum detail can be defined as three times the average spacing calculated from the density map previously.

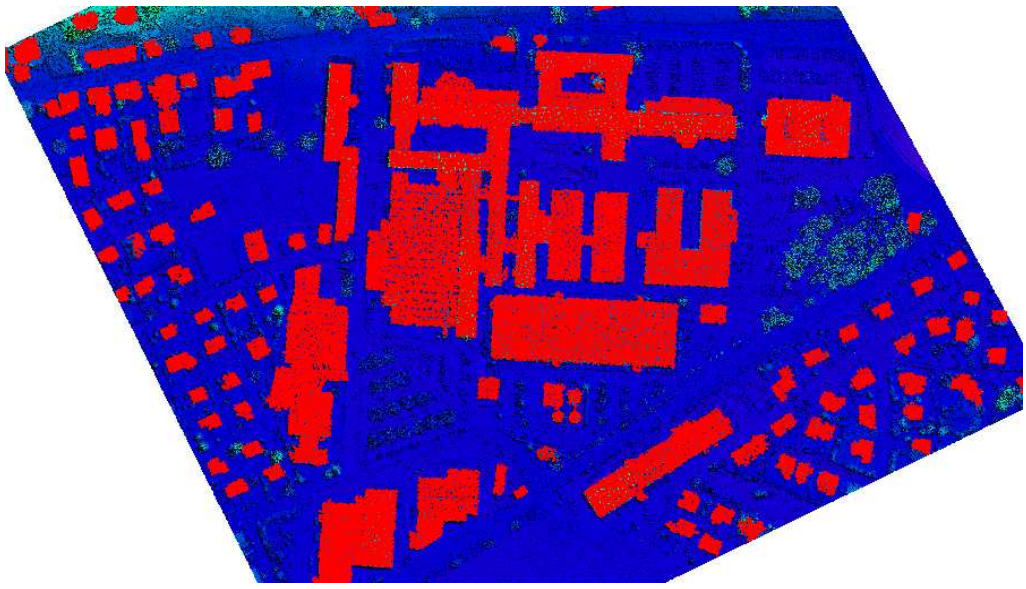


Figure 112 Top view of the initial building detection for the 1st study area, laser points representing buildings superimposed on the colour coded shaded relief

Another important parameter is the **maximum roof angle**. This parameter is useful for discriminating between gable or hip type roofs and remaining tree crowns with an area

greater than 40m^2. The valid assumption at this point is that in most cases tree crowns will introduce steeper plane angles (depending on the tree type) than the roof planes. The proposed method uses a maximum roof angle of 60^o, since small cottage style houses exist in the study areas. The **elevation tolerance** is utilized as a restriction for the least squares plane fitting process. That is, only clusters of points within the specific height tolerance will be used for estimating the least squares location of the individual planes. For the specific project the elevation tolerance can be defined as equal to the minimum vertical discrimination of the LiDAR system as defined by the sensor manufacturer. In the absence of this information the elevation tolerance can be safely assumed to be equal with the horizontal spacing between the LiDAR points. From the above procedure, an initial building classification is performed as indicated in figures 112 and 113.

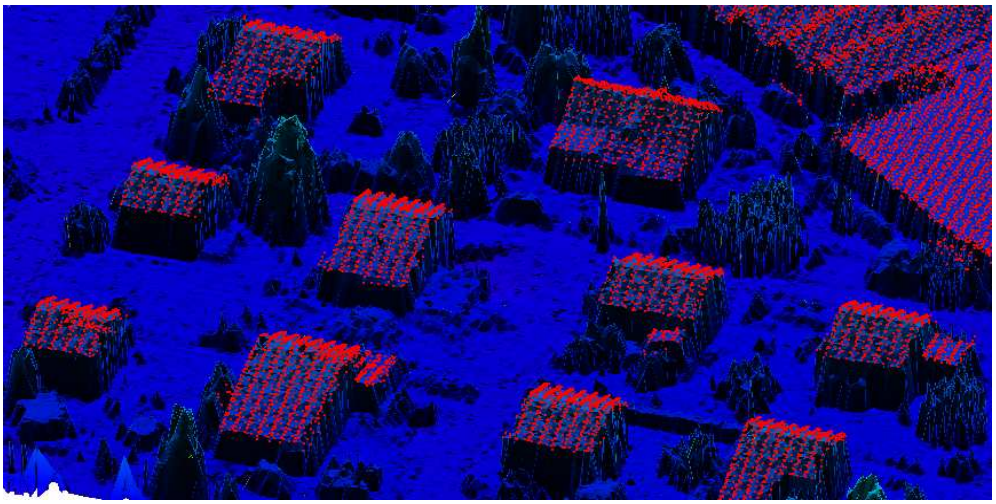


Figure 113 Perspective view of the initial building detection for the 1st study area, superimposed building points on the colour coded, shaded relief map

6.4.1.5. Generating DSM from the initial building classification

From the initial classification obtained with the previously mentioned procedures, vegetation, ground surface, features with height up to 2.5m and buildings should be successfully segmented. The main issue introduced at this stage are the LiDAR points representing roof details with area smaller than 40m^2. Recall that for classifying buildings a restriction was used that excluded points inside void regions with area smaller than 40m^2. The restriction

was used so that the process does not take into account small remaining trees that have not been previously classified. Therefore, there are still few remaining unclassified LiDAR points representing small roof details (figure 114).

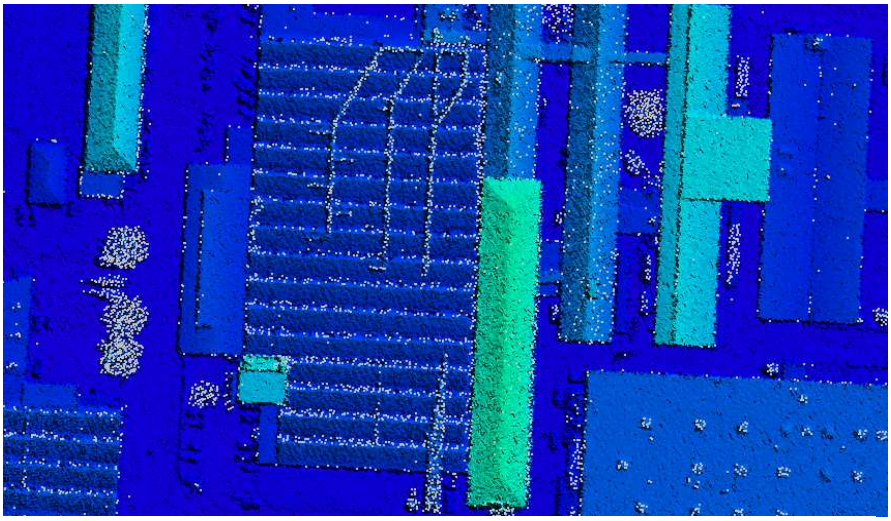


Figure 114 Remaining unclassified points (white features) representing individual trees and roof details overlaid on a TIN model for the study area

In order to assign the unclassified building points, a building vector hypothesis should be created to filter the desired points. The first stage toward the creation of the vector building hypothesis is to generate a raster DSM using only the detected building class. The selected spacing can be equal to the LiDAR point spacing. The resultant DSM represents only the buildings of the study area, with every other feature assigned in the background (figure 115).

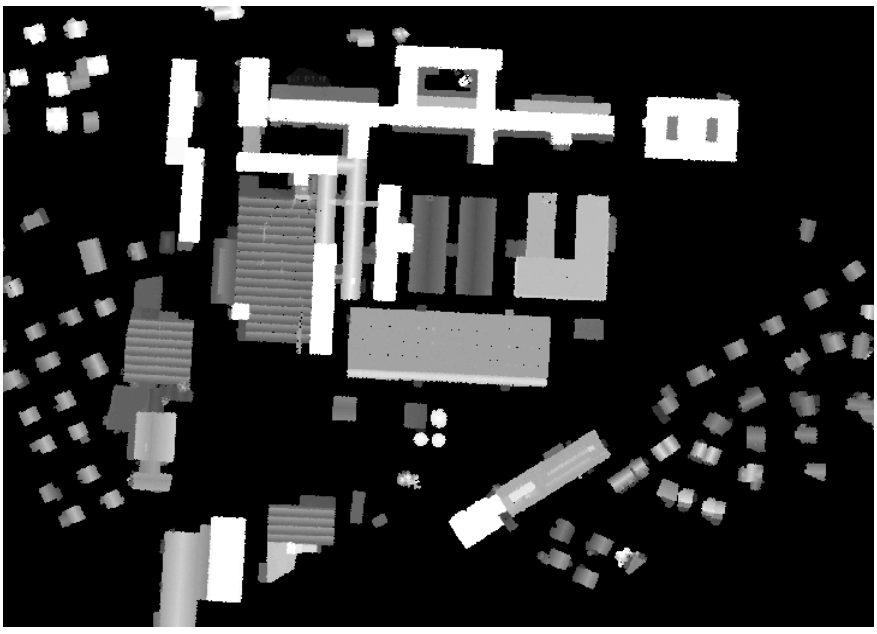


Figure 115 Digital Surface Model of the initial detected buildings and surrounding features in the background

6.4.1.6. Reclassification and raster to vector conversion of the DSM

Before the raster to vector conversion, the raster DSM has to be reclassified into a binary image. During the reclassification, pixels representing building entities are assigned to the foreground (value 1), while the surrounding features remain in the background (value 0).

The raster to vector conversion was performed with the ArcScan module in ArcGIS, using a smoothing weight of 2, without any gap closure parameters and a void area size closure of 3 pixels. The resultant polygon layer from the raster to vector conversion is depicted in figure 116.

The polygon layer is subsequently used to filter the unclassified LiDAR points, as described in the following section.

Figure 116 Polygon layer of building hypothesis, produced form the raster to vector conversion process.

6.4.1.7. Filtering and merging unclassified LIDAR points to the building class

The filtering is performed in a GIS environment by overlaying the unclassified LiDAR point cloud on the building polygon hypothesis. Only the LiDAR points located within the polygon layer are selected (figure 117) and subsequently merged with the initial building detection. This is the final step of the proposed method for the building detection procedure from high density LiDAR data.

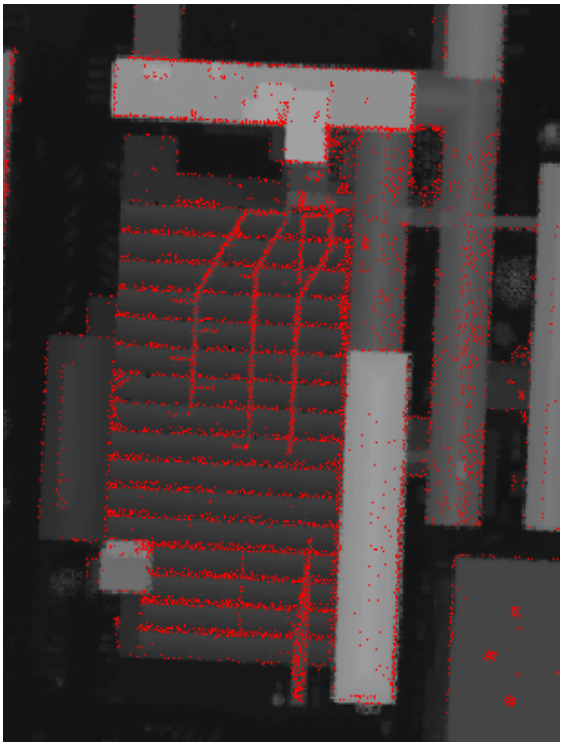


Figure 117 Selected unclassified points located within the building hypothesis

6.4.1.8. Discussion for the stage of building extraction from high density LiDAR data

From the proposed method for building detection using a high density point cloud a generic algorithm can be formed as shown in Appendix I (Algorithm 3).

The proposed algorithm includes an additional stage of splitting the entire LiDAR point cloud into subsets, if the study area is larger than 1km^2. This stage takes place before the building classification and it should be incorporated in the algorithm in order to avoid any RAM overloading on the workstation. This problem is introduced, because during the building classification process the fitted planes are stored temporarily and in general the entire process of building detection is very computationally intensive. Therefore by splitting the

study area and running the procedure separately can effectively reduce the computation requirements. The proposed method for building detection described in this section produces very promising results with nearly all buildings in the study area successfully detected, unrelated to their size and roof type. The main advantage is the generic approach of the algorithm that can be implemented in a variety of situations. In addition, the method is not restricted by the condition of the tree canopies (leaf-off or leaf-on condition) and despite that the tree delineation step (first step) is not able to detect all the individual trees, the method can effectively disregard them during the building classification stage.

Figure 118 Incorrectly classified building points (highlighted in blue ellipses), superimposed on a rectified true colour composite

Nevertheless the proposed method for building detection has a major disadvantage. It requires a certain level of user interaction for some crucial parameters which are very difficult to define automatically. Even though the algorithm suggests some default values, their effectiveness in varied situations is questionable and further research is required. Therefore the proposed stage for building detection presents a semi-automated solution rather than a fully automated one. Another difficulty is introduced when individual trees, not segmented in the initial step, are bordering buildings. In these circumstances the minimum

area parameter specified during the building classification process will not exclude the tree points, since the void area is including both buildings and vegetation. An example of this problem is illustrated in figure 118.

6.4.2. Classification and building detection using course LiDAR or VDDSM

As mentioned before the building detection workflow presented previously is very efficient and reliable when high density LiDAR data are available. If no LiDAR data are available an alternative building detection process is utilized on the VDDSM. This alternative workflow is based on geomorphometric analysis of raster surfaces rather than using the raw 3D points as in the workflow introduced in section 6.4.1. The reason for this is that the previous workflow introduced is very effective when high quality, very dense LiDAR point clouds are available but it is unreliable when course LiDAR or stereo matched points are used, because of the lower density and increased number of outliers. The workflow presented in this section is even more flexible when dealing with low density surfaces and maintains a satisfactory reliability at the expense of few artefacts and errors introduced in the building footprints. The workflow described in the chapter is also implemented when LiDAR data with a density of less than 3 points per square meter is available.

This alternative building detection algorithm introduces a new method for extraction of the building class from LiDAR DEMs on the basis of the geomorphometric segmentation principles. Thus, an object partition framework of the LiDAR DEM will be defined (Milliaresis and Kokkas, 2007). Objects will be represented on the basis of geomorphometric parameters combined by object classification which is expected to allow the extraction and mapping of the building class. The workflow starts by initially analysing the hypsometric characteristics of the study area. Then seeds cells and region growing criteria are defined. Finally, region growing segmentation is performed and an object partition framework is

defined. Connected components labelling identifies background and foreground objects that are parametrically represented on the basis of elevation and slope and then classified.

6.4.2.1. Seed sells and histograms analysis

The first step before any analysis can be carried out is for the operator to quantify the elevation frequencies present in the study area. The elevation histograms for the two study areas in London and Switzerland (figure 119) can be useful during the interactive statistical analysis at a later stage.

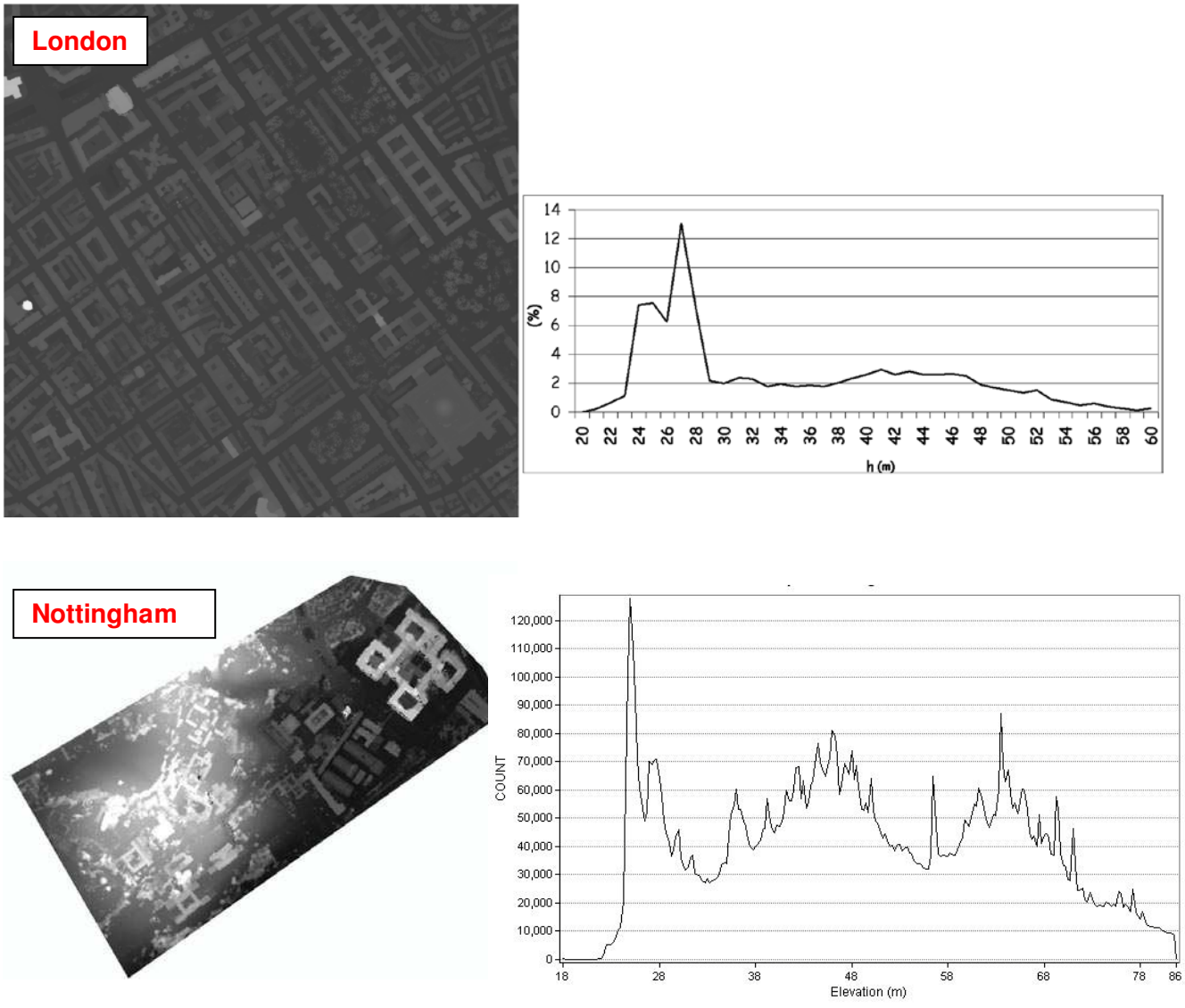
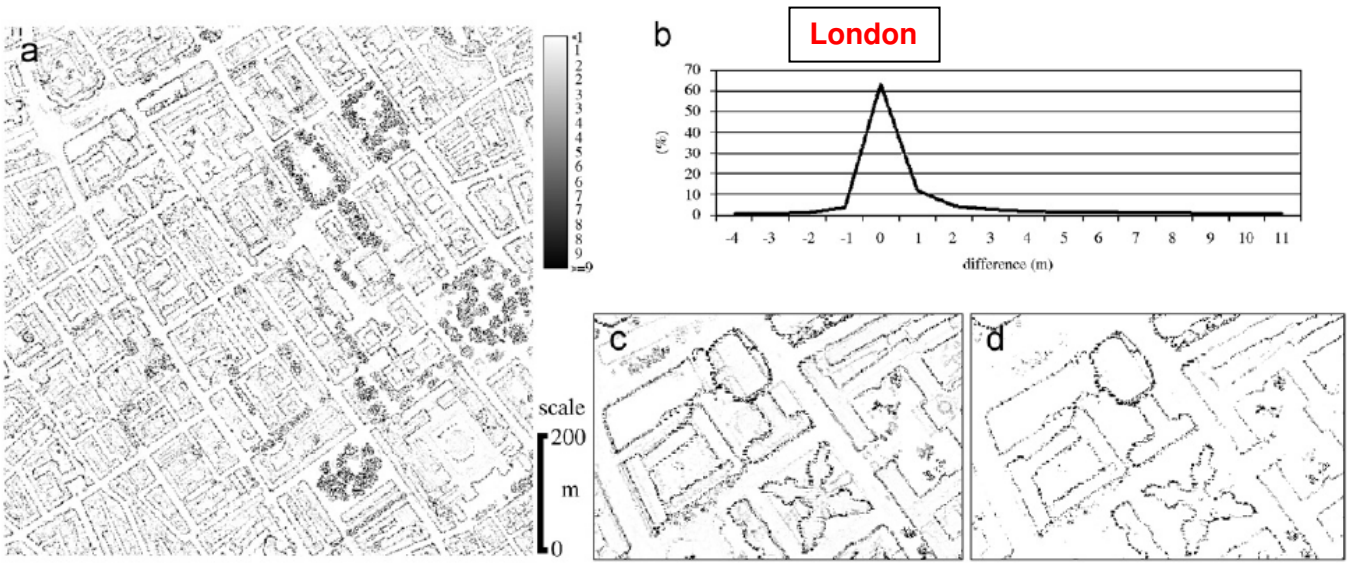


Figure 119 The elevation frequency histogram for London and Nottingham study areas. Elevation is in range 12– 215 m and 18-86m respectively, the lighter a point is the greatest its elevation.

The elevation frequency histogram (figure 119) indicates that the majority of the VDDSM cells (98.9 %) present elevation in the range 20-60m and 18-86m for London and Nottingham study areas respectively.

The next step in the building detection workflow is to define reliable seed cells along building edges prior to the region growing segmentation. In order to achieve this, a median filter is applied for seed pixels to be defined. The seed points are computed in a limited size kernel window. More specifically, the kernel size is 5x5 grid cells; note that each grid cell is related to the size of the LiDAR DSM sampling size. Usually the sampling size for the VDDSM or LiDAR DSM is no greater than 1-2m. Due to the very small kernel size, it is very difficult for the seed points to be affected by the regional topographic surface. Additionally, the regional topographic surface in all the study areas is a gently sloping terrain. Even if the regional topographic surface is more complicated (either 2nd or 3rd order polynomial surface), the elevation differences computed within a 25-50m^2 area would be residual elevation anomalies and not regional elevation anomalies especially in an urban (city) environment. Median filtering removes very high and very low elevation values within the selected kernel size while at the same time preserves edges (Mather, 1987). Once the Median filtered DSM is calculated, it is spatially subtracted from the original VDDSM or LiDAR DSM and forms a Difference DSM as depicted in figure 120.



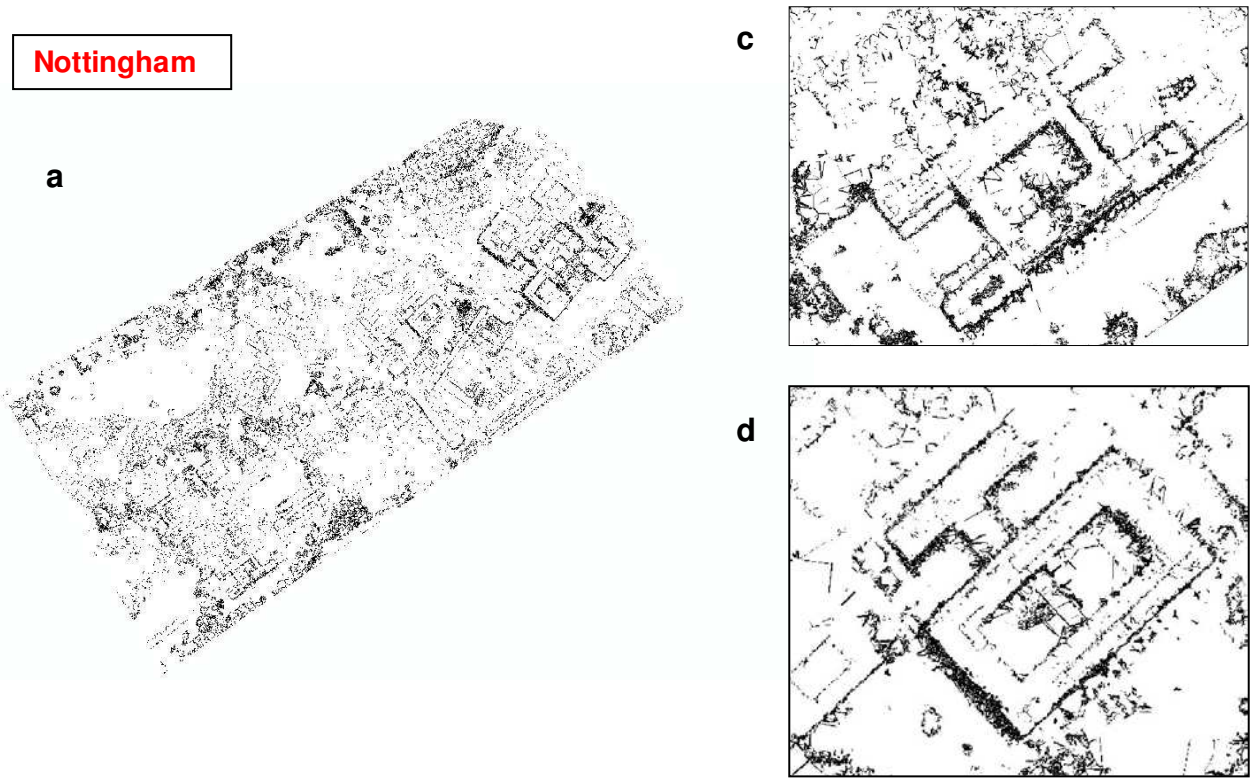


Figure 120 The difference DSM (VDDSM – Median DSM) for the second and third study areas shown in (a). Enlarged portions of selected buildings are shown in (c) and (d) with differences less than 1m depicted as white while differences greater than 9m and 3m are depicted as black for London and Nottingham respectively

The Difference DSM (figure 120) revealed the cells of the LiDAR DSM or VDDSM where severe median filtering occurs. Once the difference DSM is calculated a histogram of the frequency of the elevation differences is generated that indicates the concentration of the elevation differences. The operator can then visually identify if the distribution is positively skewed (greater frequencies occur for positive elevation differences with respect to the corresponding negative elevation differences). Note that in figure 120 elevation differences less than 1m are depicted in both study areas as white while for the London area, differences greater than 9m are depicted black. For the Nottingham area differences greater than 3m are depicted black due to the shorter buildings present in the scene. The frequency of the elevation differences indicates that the majority of the DSM cells (97.1%) present differences in the range of (-4,11) for London and (-10,25) for Nottingham.

Positive differences are interpreted to be associated to building edges. As can be seen by figures 120 (c) and (d), the greater the elevation difference, the more the detected cells stand along the borders of the interpreted city buildings. Under a trial and error procedure though experimentation and visual interpretation of the histogram and difference DSM, it was found that positive elevation differences in the range of 5-9m are more associated to the building edges in London. For the Nottingham study area the elevation difference range of 3 -12m is more associated to the building edges. On the contrary differences greater than 9m in London and 12m in Nottingham were interpreted to be associated to the internal building/roof structure (chimneys, dormers etc). Thus seed cells were selected as those that are located within the pixels with elevation differences in the range of 5-9m for London and 3-12m in Nottingham (Miliaresis and Kokkas, 2007).

6.4.2.2. Region growing segmentation

Once the seed cells have been selected a region growing segmentation can be applied. Region growing segmentation is an effective method to form the initial building cells and it is therefore employed in the proposed workflow. Although it can be very efficient, one of the main disadvantages is that it requires a certain level of user interaction to optimise the region growing criteria. Initially the region growing criteria should be defined (Miliaresis and Kokkas, 2007) on the basis of slope and elevation. If the partial derivatives of elevation (H) along the east (x) and the north (y) direction are known then slope and aspect (slope pointing direction) are computed as given in equations 6.8 and 6.9.

In order to estimate the partial derivates of elevation over a 3x3 kernel, Evans (1980) fitted a 6-parameter quadratic equation while Zevenbergen and Thorne (1987) used a 9-parameter quadratic equation. Evans (1980) method results in a polynomial surface which will not necessarily pass through the 9 grid cells included in the 3 by 3 kernel, while the opposite is true for Zevenbergen and Thorne (1987). A review and a comparison of these methods (Skidmore, 1989) proved that Evans (1980) method is more precise, since Zevenbergen and Thorne (1987) method is affected more by random noise and errors evident in DEMs. The slope image as well as the frequency histogram for the second and third study area is given in figure 121.

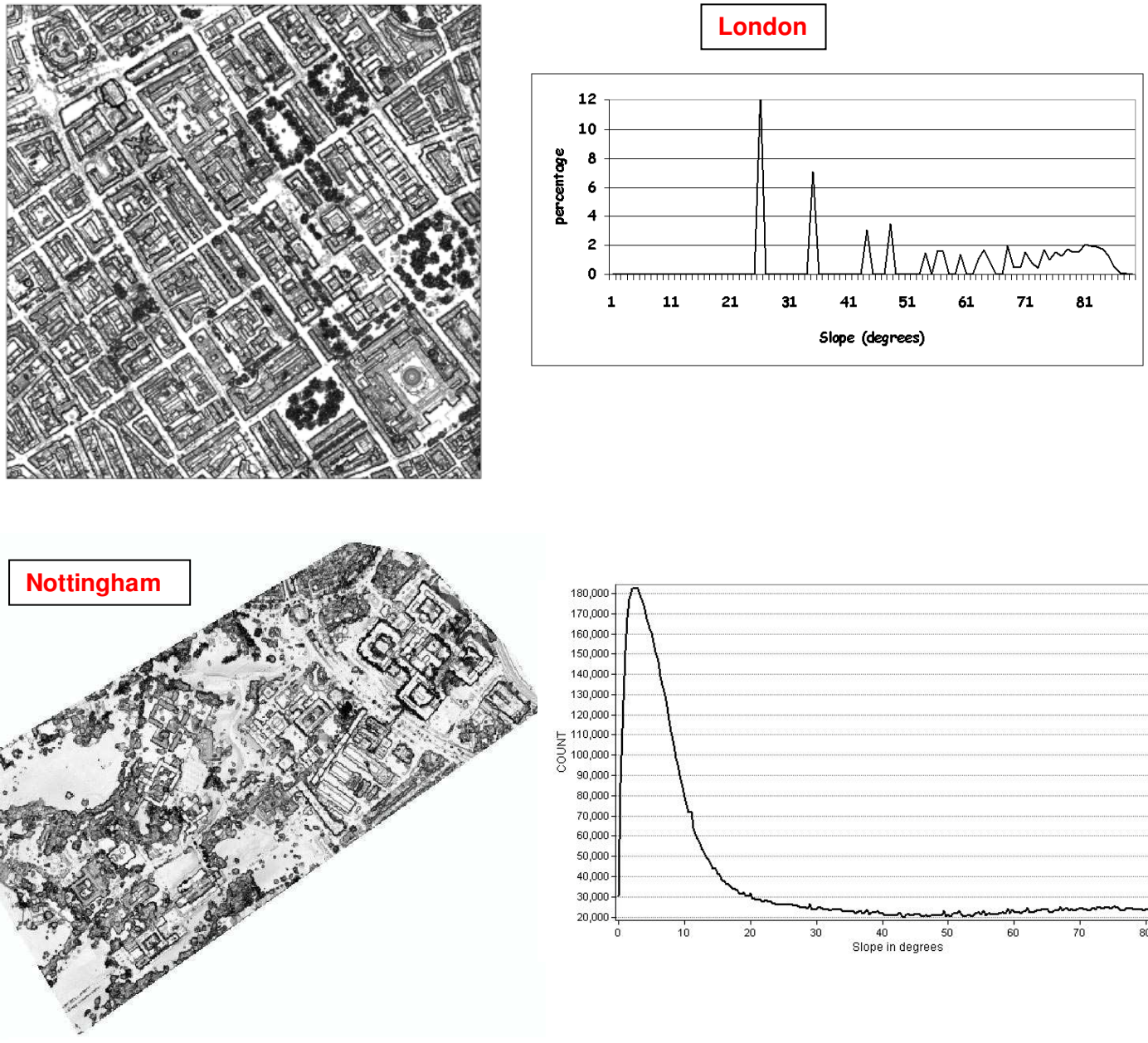


Figure 121 Slope image for London and Nottingham on the left side (The higher the slope value the darkest a cell is). Frequency histogram of slope in the interval 1-90° on the right side.

The frequency histogram of the slope (figure 121) for the two study areas in London and Nottingham are completely different than the slope frequency distributions normally observed in geomorphologic applications (the slope frequency decreases gradually for increasing slope values) where digital terrain models (DTMs) are used (Miliaresis and Kokkas, 2004; Miliaresis et al., 2005). So in DTMs (with spacing 30–1000 m), the ground elevation prevails and thus the slope frequency distribution is the outcome of geomorphologic (erosion, deposition) and geologic (tectonism) processes (Miliaresis, 2006). On the contrary in urban LiDAR DSMs or VDDSMs, the building structure prevails and due to the high DSM resolution (usually less than 1-2m), the slope frequency distribution observed is the outcome of the interaction of the building surface (abrupt changes of elevation, high and abrupt changes of slope).

The interpretation and study of both a slope image and the slope frequency histogram is crucial for the operator to determine the slope cut-off values. Through a trial and error procedure the slope cut-off values for the region growing criteria were defined in the interval of 40^0-90^0. These slope cut-off values were used both in London and Nottingham test sites.

Once the slope cut-off values are determined an iterative region growing segmentation algorithm is implemented. The seed cell is determined from the initial set of building cells. For each iteration of the region growing algorithm, if a cell of the DSM has a slope within the cut off angles is an 8-connected neighbour to the current set of cells, it is flagged as a new building cell and the current set of building cells are updated. Note that the 8-connected neighbours are the 8 pixels surrounding the central pixel in a 3x3 kernel. The segmentation stops if no more cells are added during the current iteration. For the London test site a total of 448,865 cells were identified after 29 iterations while for the Nottingham test site 269,734 cells were identified after 25 iterations (figure 122).

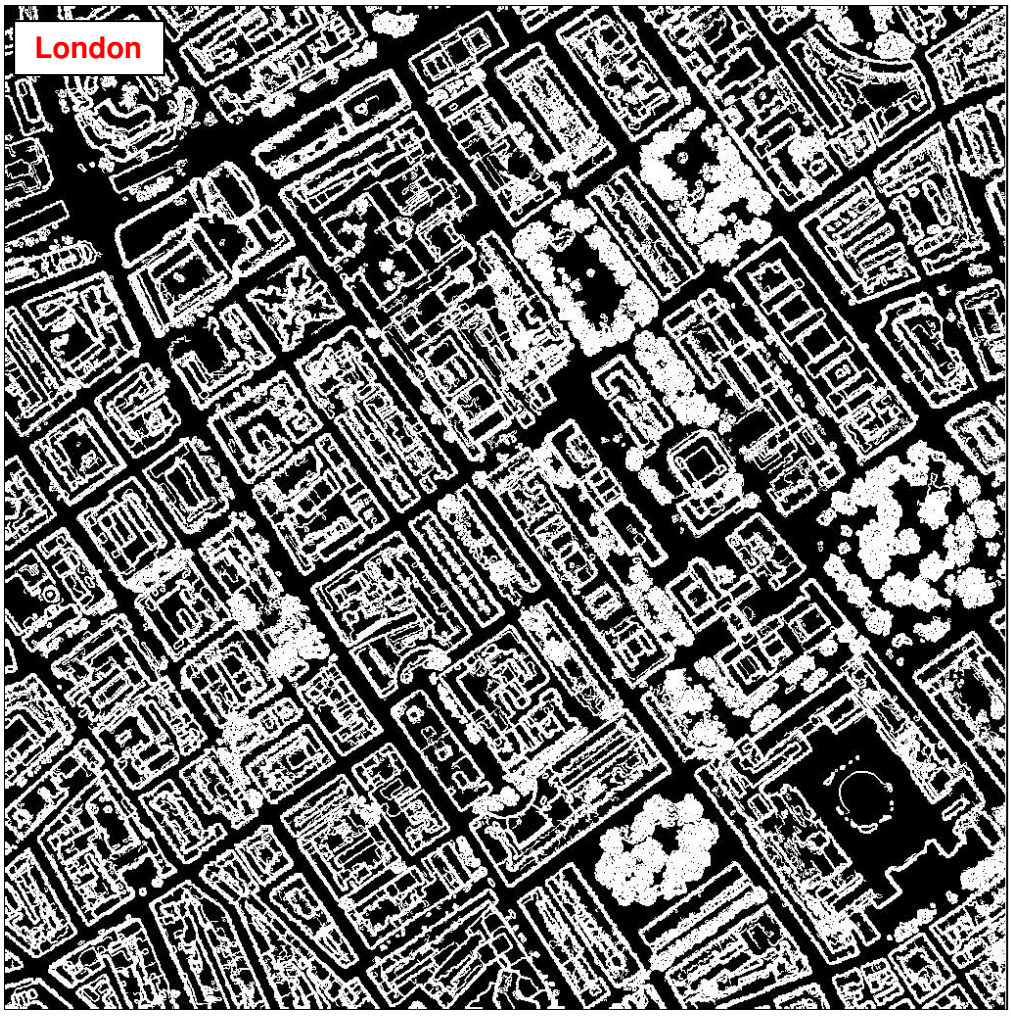


Figure 122 The resultant foreground and background object class from the implementation of the region growing segmentation workflow in second and third study areas

Once the region segmentation is complete a connected component-labelling algorithm (Pitas, 1993) is applied and regions formed by the 8-connected background cells (black pixels) are grouped and labelled (figure 123).

Figure 123 The resultant grouped and labelled foreground and background object class from the connecting-component labelling process in the second and third study areas

At this stage many objects (most of them very small in size) are evident in the segmented images in both study areas (figure 123). The very small-sized objects correspond to residual objects evident either on the roofs or in the non-building class. These residual objects are classified as such when their area size is less than 40m^2 which is the minimum building size we have adopted for the building detection workflow. Small in size foreground (white) and background (black) objects are eliminated by applying object size filtering. More specifically, foreground objects with size less than 40m^2 cells were initially merged to the background class. For the London study area 6633 cells were merged to the background and so the foreground cell class was formed by 441,828 cells. For the Nottingham study area 8532 cells were merged to the background and so the foreground cell class was formed by 245,982 cells.

Once the foreground cells are merged to the background class the reverse process takes place. That is background objects (formed by black cells) that are located within foreground closed objects are merged in the foreground class (figure 124).

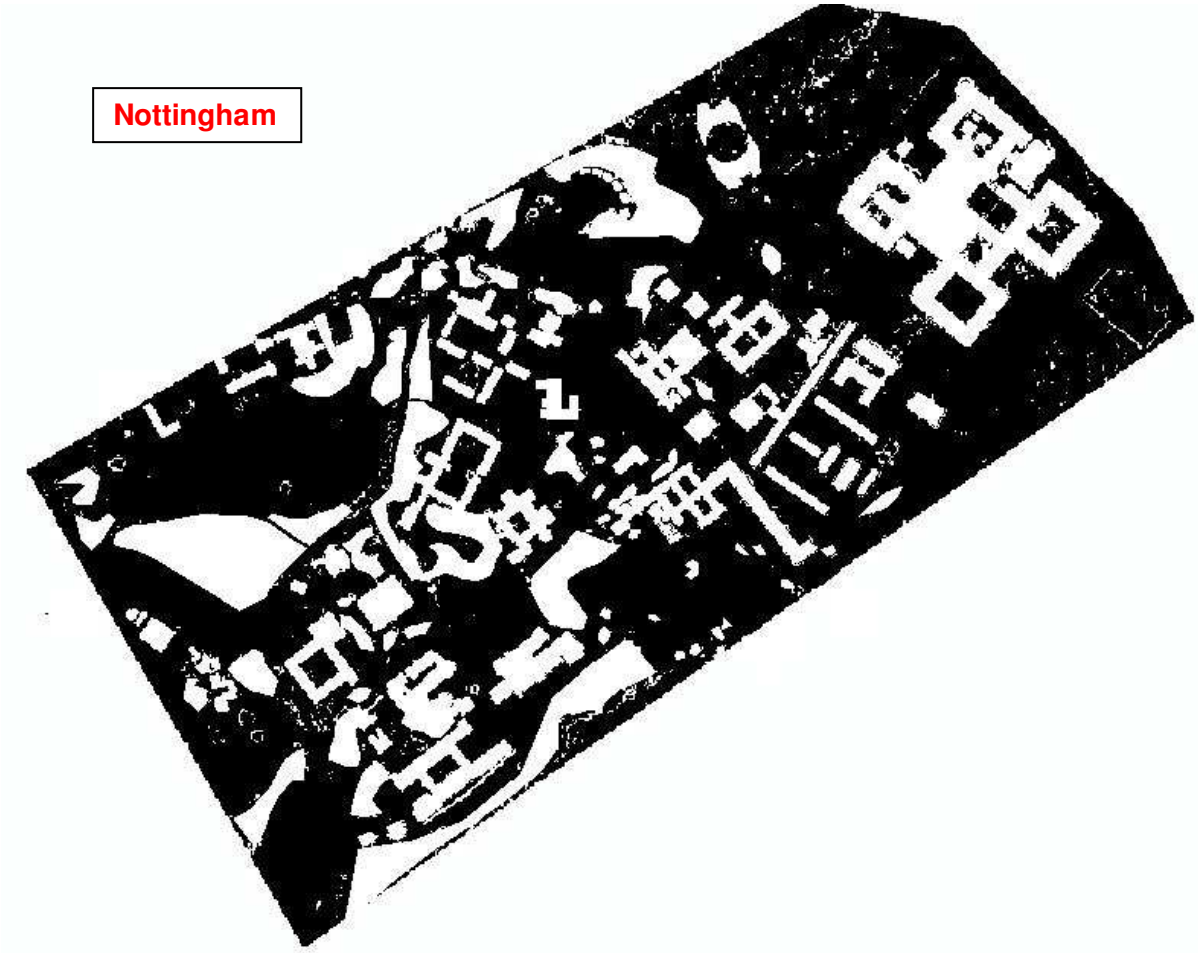


Figure 124 The resultant grouped and labelled foreground and background objects after the implementation of the object size filtering where background cells within enclosed foreground objects were merged to the foreground class.

For the London test site a total of 114,009 of the background cells (black cells) were merged to the foreground object class while 134,551 background cells were merged to the foreground class in Nottingham. Figure 124 shows the resultant foreground and background objects in Nottingham. Any of the foreground or background objects with an area of less than 40m^2 were also filtered out. Finally the remaining foreground and background objects were labelled by the connected components algorithm and merged into a single image. Thus each of the background and foreground objects has a unique ID (figure 125).

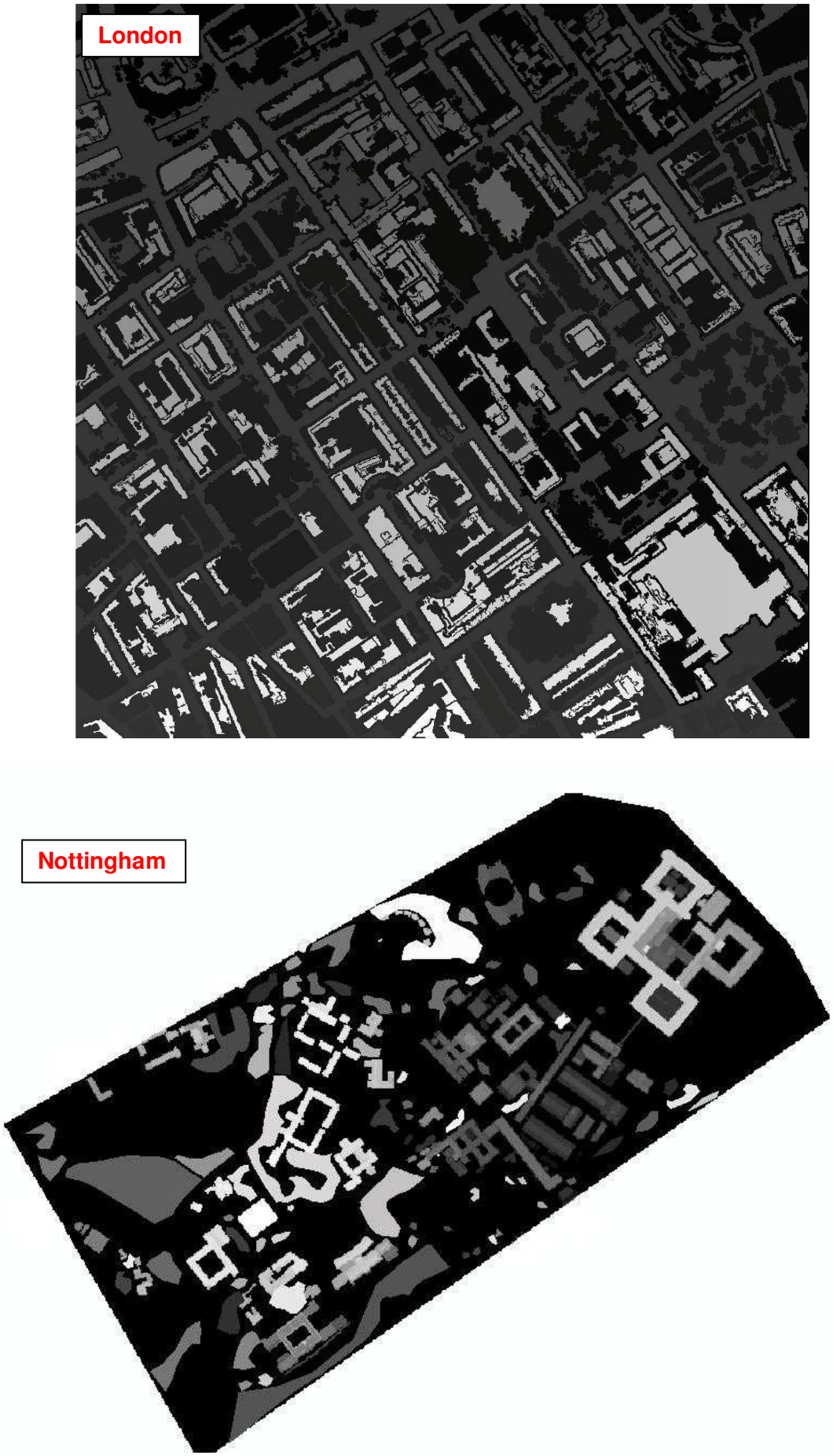


Figure 125 The merged image that contains both foreground and background objects after the implementation of size filtering and the components labelling algorithm. Different brightness values indicate different ID for each object

From figure 125 it is evident that in both study areas most of the artefacts representing smaller features in figure 124 have been successfully filtered out and there are three predominant object classes remaining in the segmented images. That is the building class, non building class and vegetation. There have been several instances were problems occur when vegetation and buildings are sharing common borders or overlapping and this issue will be addressed in the following section.

The region growing segmentation, connected components labelling, size filtering and the subsequent object parametric representation were performed with GeoLogic Shell software which is freely available for download from the website of the International Association for Mathematical Geology (GeoLogic Shell software, 2008).

6.4.2.3. Parametric representation and final building classification

The objects from the implementation of the connected components process as described in the previous section are interpreted to belong either to the building class or to the non-building class (roads, vegetated areas, components of a building, etc.). The main problem at this stage is that objects that are adjacent or ones enclosing others could both belong to the building class. In order to identify which additional clusters should be merged in the building class each object is parametrically represented by a set of geomorphometric attributes and then an unsupervised classification defines clusters of objects that belong either to the building or the non-building class.

For the parametric representation of the objects resulting from the connected components algorithm (figure 125), four attributes are used.

- H, mean elevation that was computed as the average elevation of the cells that belong to an object's region.
- R, roughness, the standard deviation of elevation of the cells forming the object (a measure of the vertical variability of object's height).

- S, object's mean slope.
- SR, standard deviation of slope (a measure of the variability of object's slope).

From the implementation of the unsupervised classification based on the above parameters the proposed workflow calculates the cross correlation matrix of the attributes based on the cluster of objects. The attribute correlation matrix for the London and Nottingham study areas is given in table 23.

Attribute correlation matrix (London)

	H	R	S	SR
H	1			
R	0.28	1		
S	-0.12	0.56	1	
SR	-0.03	0.63	0.72	1

Attribute correlation matrix (Nottingham)

	H	R	S	SR
H	1			
R	0.32	1		
S	-0.08	0.67	1	
SR	-0.15	0.51	0.79	1

Table 23 Attribute correlation matrix for the second and third area

It is observed that in both study areas a partial correlation exists between S and SR, as well as R and SR. In other words, if the slope is high then for most of the objects standard deviation of slope and elevation should be high, but this is not always true.

Cluster analysis is a multivariate procedure, which is commonly used for regional classification. It is based on some measurement of distance among objects (for example Euclidean distance), which is calculated in a c-dimensional space, where c represents the number of attributes used in the clustering process (Mather, 1987). The centroid method is employed that requires a priori definition of the number of clusters. The main problem in cluster analysis has to do with the different value range of the different attributes (for example elevation and slope) included in the distance calculation. In order to solve this problem, the data are standardized (transformed to a new dataset in which the corresponding attribute mean value is zero and the corresponding attribute std. dev. is 1) before the exploratory (K-means) clustering algorithm is applied (Miliaresis and Argialas, 2002).

Thus, the data values are transformed by subtracting the corresponding attribute mean value and dividing the outcome by the corresponding attribute standard deviation (Clark and Hosking, 1986). Ten clusters of objects are considered, giving satisfactory results for both study areas and are implemented as a default value. Each object is assigned to a cluster on the basis of its Euclidean distance from the cluster centroid. The cluster gravity centres are computed on the basis of the standardised centroid co-ordinates as defined in table 24 which also represents the total number of objects defined in each of the 10 clusters.

Cluster	standardised centroid co-ordinates				Number of
ID	H	R	S	SR	objects
1	3.23	8.50	-0.46	2.24	1
2	-1.16	-0.34	-0.58	-0.37	81
3	0.12	-0.35	-0.37	-0.50	119
4	4.29	9.47	1.91	2.37	1
5	-0.32	0.93	1.63	1.97	56
6	0.46	2.56	2.27	1.34	2
7	-0.61	0.86	2.33	0.06	19
8	4.90	-0.13	-0.39	-0.02	2
9	6.35	8.78	-0.18	2.24	1
10	0.98	-0.35	-0.56	-0.42	85

Cluster	standardised centroid co-ordinates				Number of
ID	H	R	S	SR	objects
1	4.52	9.12	-0.23	-1.78	2
2	4.67	-0.11	1.01	0.56	37
3	5.01	0.13	0.98	-0.42	158
4	6.54	3.32	-0.91	3.08	-
5	1.02	-1.22	-1.54	-2.01	34
6	4.98	3.12	5.02	1.34	29
7	5.12	4.54	6.54	1.09	32
8	-3.90	-0.15	-1.23	0.77	1
9	10.21	7.12	0.44	-4.89	1
10	-0.23	1.12	-0.05	-1.67	3

Table 24 Final cluster (standardized) centroids and the number of objects in each cluster for London (first table) and Nottingham (second table)

The similarity among cluster centroids can be assessed by the operator by the divergence of the clusters as well as the cluster compactness. Divergence indicates the distance between cluster centroids. The cluster compactness is computed on the basis of

$$\text{Compactness} = 100 \times \frac{mean - st.dev.}{Mean} \tag{6.10}$$

Where mean and std. dev. are the mean distance and the corresponding standard deviation of the objects from the gravity centre of the cluster. The similarity assessment is performed only for the clusters with a significant number of objects, in other words classes 1,4,6,8,9 were excluded from further investigation in London and classes 1,4,8,9,10 were also excluded in Nottingham. Tables 25 and 26 indicate the distances between the cluster centroids (divergence) and the cluster compactness for London and Nottingham respectively.

Distance between cluster						Statistics for the objects distance from The corresponding cluster centroid			
Cluster	2	3	5	7	10	Cluster	Mean	St.dev	Compact.
2	0					2	0.57	0.26	54.4 %
3	**1.32**	0				3	0.44	0.21	52.3 %
5	3.58	3.47	0			5	0.65	0.34	47.7 %
7	3.24	3.12	**2.05**	0		7	0.63	0.38	39.7 %
10	**2.15**	**0.88**	3.74	3.56	0	10	0.61	0.38	37.7 %

Table 25 London study area - Distances between cluster centroids (divergence) and cluster compactness for clusters 2,3,5,7,10

Distance between cluster						Statistics for the objects distance from The corresponding cluster centroid			
Cluster	2	3	5	6	7	Cluster	Mean	St.dev	Compact.
2	0					2	0.60	0.25	58.3 %
3	**0.98**	0				3	0.64	0.28	56.2 %
5	3.79	4.21	0			5	0.72	0.38	47.2 %
6	4.01	3.78	**1.34**	0		6	0.68	0.42	38.2 %
7	3.35	5.01	**0.94**	**1.46**	0	7	0.75	0.45	40.0 %

Table 26 London study area - Distances between cluster centroids (divergence) and cluster compactness for clusters 2,3,5,6,7

By analysing tables 25 and 26, based on the distance between clusters, it is concluded that the clusters can be grouped in two categories on the basis of similarity/divergence. For the London study area clusters 2, 3 and 10 form the first category and clusters 5 and 7 form the second category. For the Nottingham study area the first category consists of clusters 5,6 and 7 and the second category consists of the clusters 2 and 3.

In the first category the clusters are differentiated on the basis of mean elevation while the second category is formed by objects with almost equal mean elevation. In addition tables 36 and 37 indicate that clusters 7,10 in London and clusters 6,7 in Nottingham are less compact, meaning that they could be divided into possible to sub-clusters.

At this stage it should be pointed that the VDDSM on Nottingham and the enhanced LiDAR DSM in London are still quite noisy while at the same time residual elevation objects are observed over the roofs. These objects correspond either to artificial objects induced by

noise effects or to small elevated features that are difficult to recognise and assign to a certain feature type.

Through a visual interpretation of the classified objects (figure 126 and figure 127) to the selected cluster and with the assistance of cluster statistics the operator can identify which clusters belong to the non-building, vegetation and building classes.

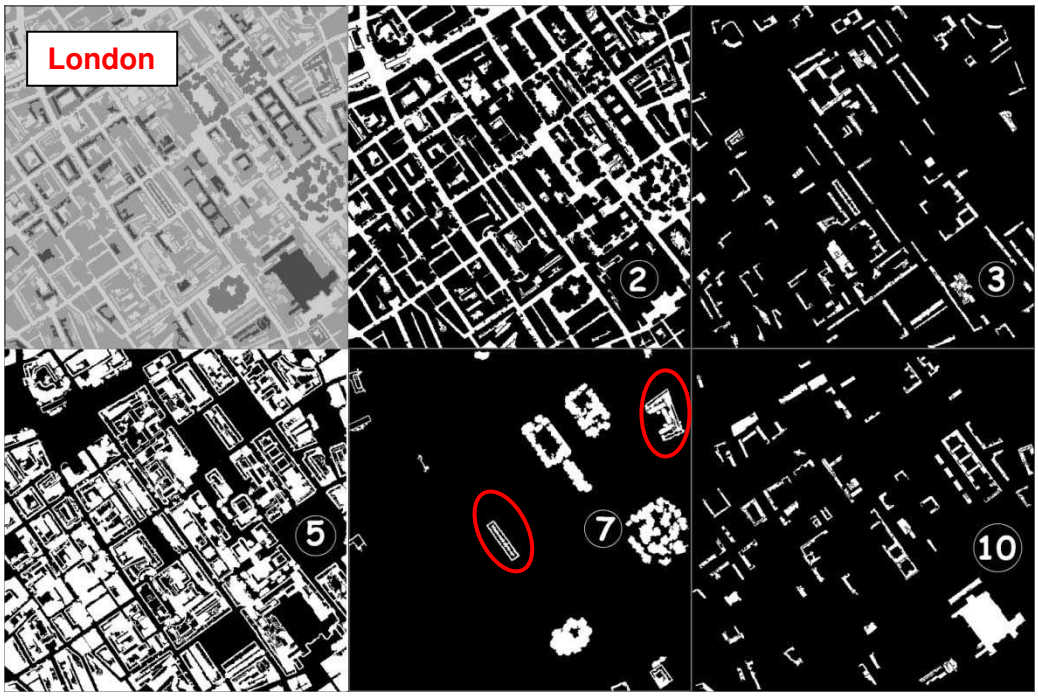


Figure 126 Visualization of the objects assigned to each of the 5 cluster (2,3,5,7,10) after the implementation of the K-means classification algorithm in London study area. White colour represents the objects of each cluster

Figure 126 indicates that in the London study area the non-building class is formed by cluster 2 while the urban vegetation class corresponds to cluster 7. It is interesting to note though that the classification workflow presented in this section is not without problems as in some cases building roofs are incorrectly assigned to the tree cluster (cluster number 7, circled in red) as highlighted in figure 126. This is due to the similar roof slope variability with gable type roofs and a similar mean elevation. The interpretation of the cluster centroids and the statistics presented in tables 25 and 24 it is evident that the non-building class is formed by the less elevated, mostly flat, less variable with respect to slope and elevation

objects. The urban vegetation is formed by objects with the highest slope and with significant slope variability.

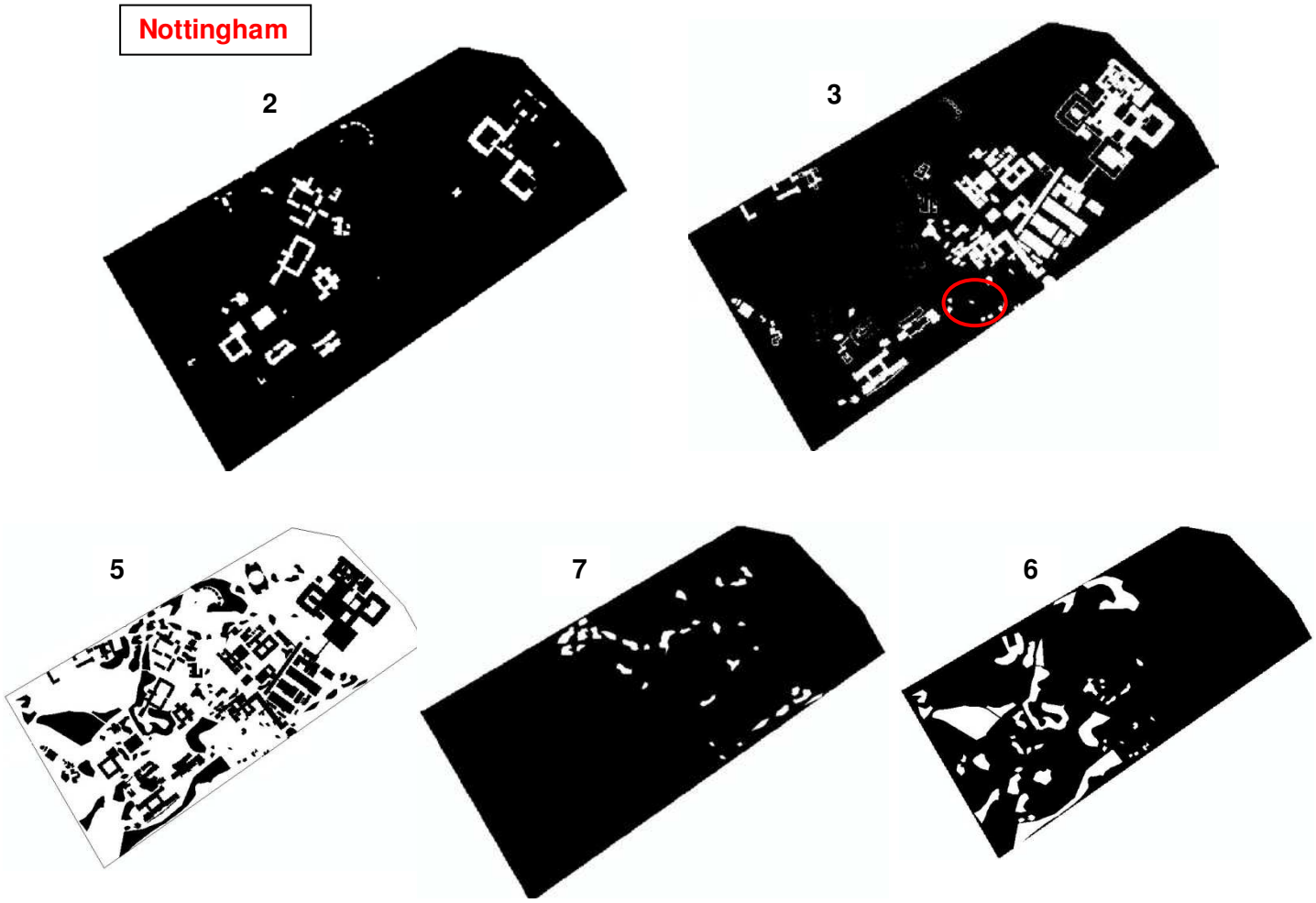


Figure 127 Visualization of the objects assigned to each of the 5 cluster (2,3,5,6,7) after the implementation of the K-means classification algorithm in Nottingham study area. White colour represents the objects of each cluster

Thus the vegetation class is usually formed by low mean elevation (but higher than the elevation of the non-building class), while slope is maximised while the standard deviation of slope and elevation is high. The last remarks are functions of tree canopy properties and thus different vegetation types might present different parametric representation. For the Nottingham area the non building class is formed by the objects in cluster 5 while the vegetation category is a combination of the objects in cluster 6 and 7.

The building class in London is formed mainly by the objects of cluster 5 while in Nottingham is formed by cluster 3. For the objects of the building class, slope variability of maximised

while slope is among the maximum ones. The interpretation for this is due to the building edges that are in contact with the non-building class. The clusters 10 and 3 in London correspond primarily to internal portions of building and in Nottingham, cluster 2 represents some buildings and building rooftops. That is why the roughness, slope and slope variability is among the minimum while mean elevation is maximised. At this stage it should be pointed that another difficulty was introduced in Nottingham where the building class (cluster 3) had incorrectly classified some tree canopies due to the similarity in slope variability and mean height. These misidentifications were manually removed from the classified objects.

As a conclusion of the workflow presented in this section, the geomorphometric region growing segmentation combined by median filtering for identifying seed cells, connected components labelling, size filtering and object labelling, object parametric representation on the basis of slope and elevation attributes and classification has proved capable of delineating the building class from both study areas. The interpretation of cluster centroids allowed the identification of building sub-classes with different geomorphometric characteristics that are associated with different building portions.

Nevertheless, the proposed method for building class detection has a major disadvantage. It requires a certain level of user interaction for some crucial parameters, which are very difficult to define automatically in varied situations and further research is required for full automation. In addition, although most of the building footprints were successfully extracted there were several instances of building roofs assigned to the tree clusters and small tree clusters assigned to the building class that required a small level of manual intervention to rectify.

6.5. Data fusion for optimizing building footprint and roof details

This section describes the proposed method for the initial generation of the building outlines and subsequently the data fusion step, with the stereo matched points, for refining and adjusting the initial building outlines and roof details. The diagram shown in figure 128 indicates the steps of the proposed workflow that will be covered in the section.

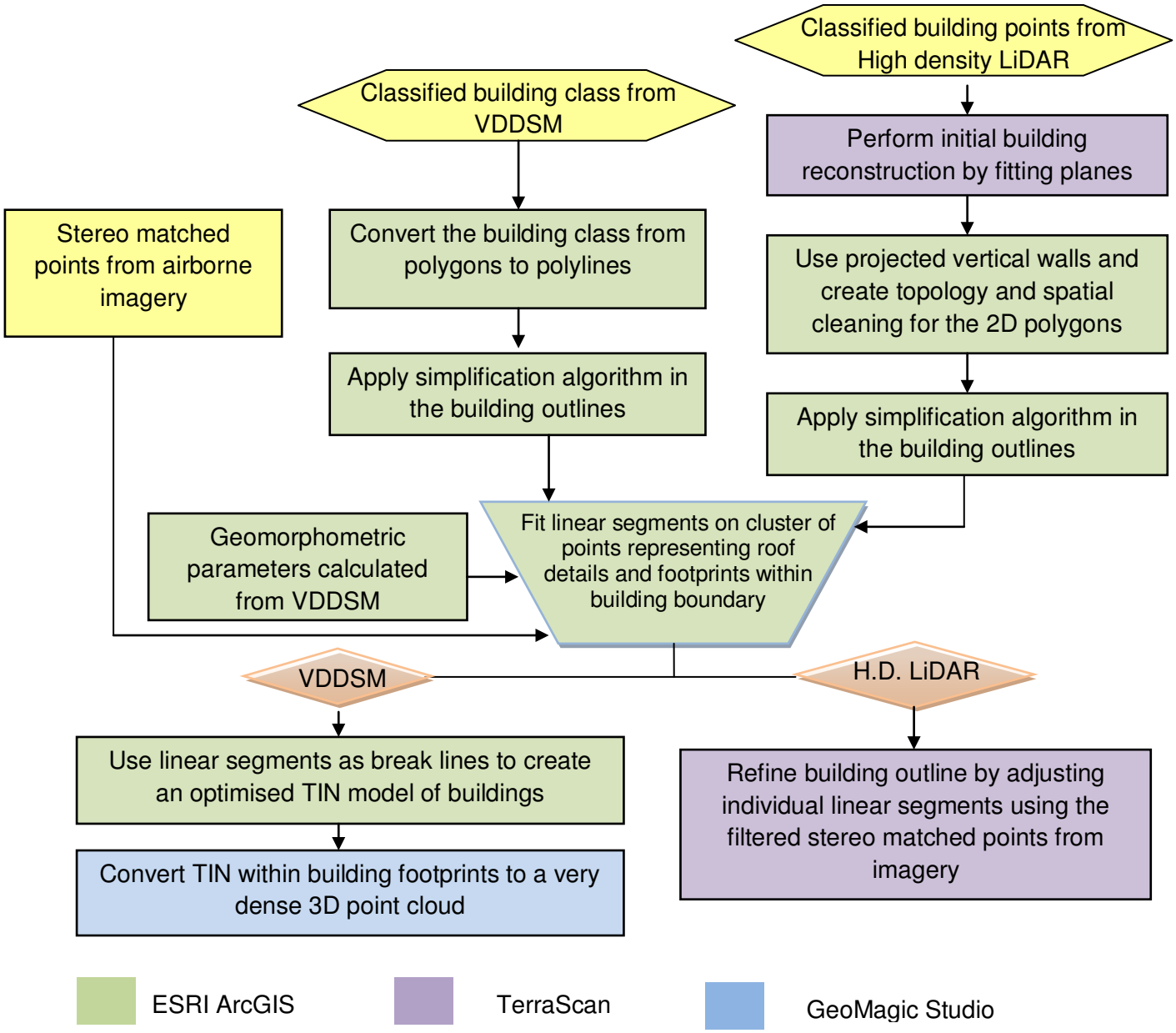


Figure 128. Diagram indicating the steps of the Data Fusion process

Before the introduction of the data fusion method it is useful at this point to recap some of the previous steps and the processing that has taken place since the entire workflow can be complicated at times. The data fusion method as shown in figure 128 can have two main inputs at a time, which are the stereo matched points and the classified buildings. The process of extracting the stereo matched points from the available aerial photos was described in section 6.2.3. The classified buildings can come from two sources, either from high density LiDAR or from a VDDSM. Note that a VDDSM is created either by combining low density LiDAR with a DSM extracted from the aerial photographs or just using the DSM if no LiDAR data exist. The VDDSM generation process was described in section 6.3. In the case where a VDDSM is used for the data fusion process the additional input of the calculated slope and aspect parameters are used. The calculation of the slope of aspect from the VDDSM was also described in section 6.3.

The classification of buildings was divided in two processes, one for high density LiDAR as explained in section 6.4.1 and another method for the VDDSM, described in section 6.4.2.

6.5.1. Data fusion process for High Density LiDAR

Since the data fusion process depends on whether we have high density LiDAR or low density/no LiDAR this section will describe the data fusion process separately for high density LiDAR and for the VDDSMs.

6.5.1.1. Extracting initial building footprint from High Density LiDAR

From the workflow described in section 6.4.1 the high density LiDAR data have been classified and the points representing buildings have been isolated. Nevertheless there is still the need to create the building polygons and subsequently derive the building outlines from the classified 3D points. For this reason a least median of squares (LMS) plane fitting

algorithm is implemented on the classified building points. Parameters such as minimum plane size and height tolerance between adjacent roofs are automatically determined by the density of the point cloud.

The least medians of squares (LMS) is a robust regression method that estimates the parameters of a plane by minimizing the median of the absolute residuals, these are defined as the difference between the measurement (y_i) and estimation ($f(X_i)$) : for the i th sample $r_i = f(X_i) - y_i$. That is, we search the parameters β that minimizes the median of the residuals:

$$\underset{\beta}{argmin} = \underset{ri}{median} |f_\beta(X_\iota) - y_i| \quad (6.11)$$

The above equation can be solved using the following random sampling algorithm (Fleishman et al. 2005); k points are selected at random from a total of N points, and a plane is fitted to the points. Then the median $r_{i,\beta}$ of the residuals of the remaining N −k points is computed. The process is repeated T times to generate T candidate planes. The plane with minimal $r_{i,\beta}$ is selected as the final plane. If g is the probability of selecting a single good sample at random from the sample set, then the number of iterations that are required to have a probability of success of P can be computed by $P = 1 - (1 - g^k)^T$. A small value of k does not use the entire available sample to fit a plane, while a larger value of k requires more iterations.

Based on the classified building points, the LMS algorithm iteratively fits planes on the LiDAR points. The basic difference with the building detection algorithm is that the procedure is restricted only to the classified points without having the need to differentiate between building and vegetation. Furthermore, instead of the planes being stored temporarily at the building reconstruction stage, the planes are visualized at the user interface dialog window. The complexity and the time requirements for reconstruction is a function of the specified minimum size of the plane that the algorithm will try to detect, the size of the roofs and the density of the LiDAR point cloud.

An additional parameter is the possible merging between planes with a specified vertical separation. This option is useful for applying a certain level of generalization in the reconstructed roofs and potentially avoids multiple superimposed planes, representing the same roof surface. The proposed method was implemented in the Switzerland study area since the LiDAR data provided are classified as high density. The algorithm was implemented using a merging function between planes with a vertical difference up to 25cm. An example of a reconstructed roof from that study area is shown in figure 129.

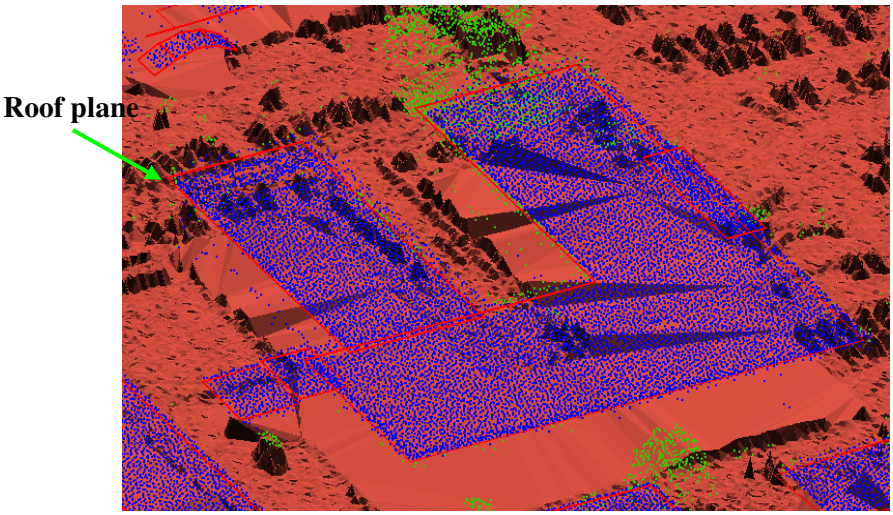


Figure 129 Perspective scene of reconstructed roof planes (red boundary), with building points superimposed over the ground TIN.

The next step is to extrude the roof planes on the ground surface and create the vertical facades of the buildings (figure 130).

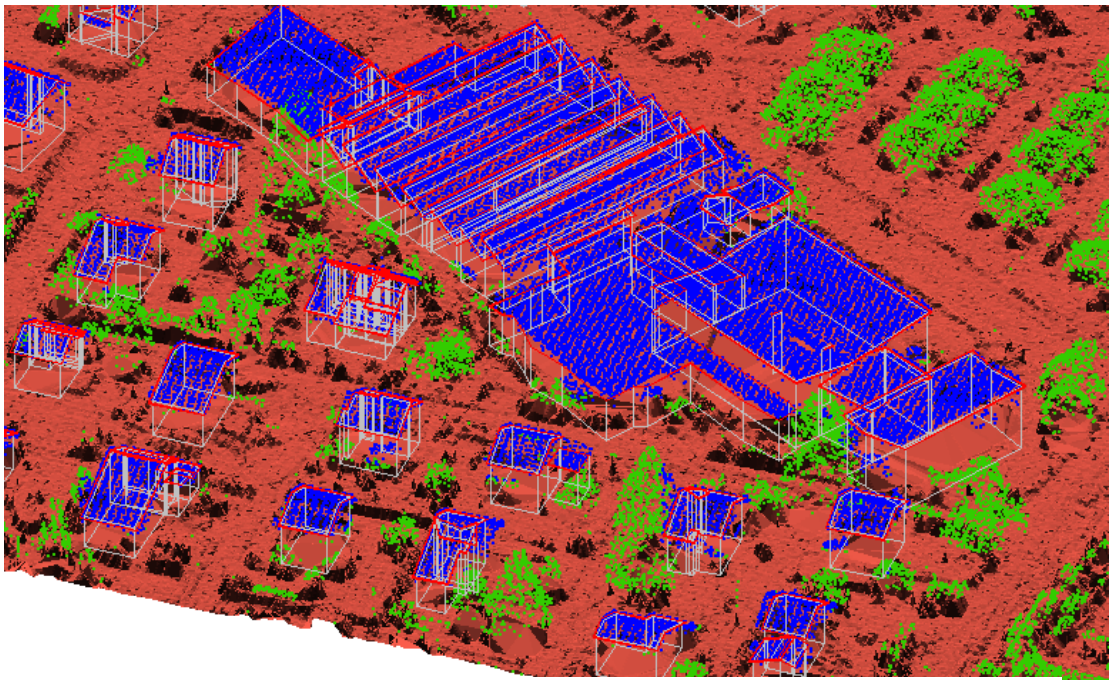
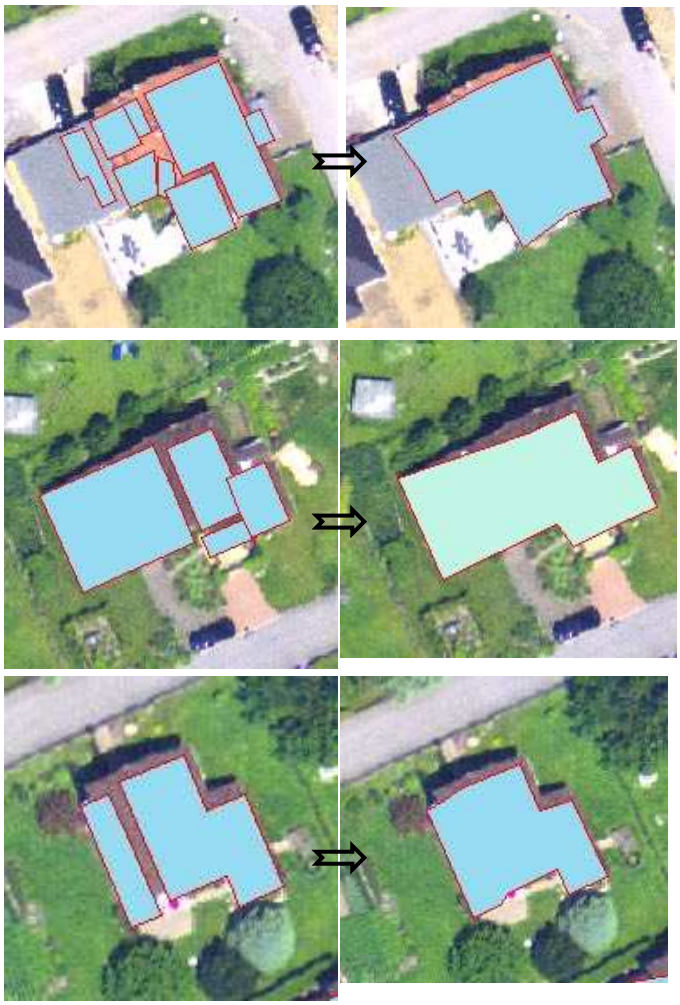


Figure 130 Extruded roof planes on the ground surface (grey lines) representing the vertical facades of the buildings

6.5.1.2. Spatial cleaning and generating topology for the 2D building footprints

After the initial 3D building reconstruction, the vertical facades of the buildings are transformed in two dimensional polygons for subsequent use in a GIS environment.

The individual polygons introduced from the building reconstruction are not consistent since adjacent planes are not automatically merged (left side of figure 131). In order to overcome this problem and create a single polygon for each building, spatial cleaning is utilized. The spatial cleaning process is the fundamental function for topology generation. Topology organizes the spatial relationships between features in a set of feature classes, using specific topological rules that will constrain different feature's topological relationships. Once the participating feature classes have been added to the topology and the rules defined, the topology is validated. Two basic topology rules are used during the spatial cleaning. The first rule requires that the interior of polygons in the feature class don't overlap. The polygons can share edges or vertices.



Extruded planes before topological correction	Building outline after spatial cleaning

Figure 131 Resulting building outline from the procedure of topology generation and spatial cleaning in Switzerland test site

This rule is used when an area cannot belong to two or more polygons. The second rule requires that polygons not have voids within themselves or between adjacent polygons. Polygons can share edges, vertices, or interior areas.

There are two parameters that are used to validate the second rule, which include the dangle length and fuzzy tolerance.

The dangle length removes dangling arcs that are shorter than the specified dangle tolerance. The dangling arc is an arc having the same polygon on both its left and right sides and having at least one node that does not connect to any other arc. It often occurs where a polygon does not close properly (undershoot) or where arcs don't connect properly. The fuzzy tolerance defines small distances used to resolve inexact intersection locations. It defines the resolution of a coverage resulting from the spatial clean operation.

This value of the dangle length should be optimized so that roof planes representing the same building are merged together but at the same time avoid merging neighboring buildings. This is especially evident in densely structured areas and a certain level of user interaction is required for specifying these parameters. For the Switzerland study area the proposed method utilized a dangle length tolerance and fuzzy tolerance equal to 2m with satisfactory results.

6.5.1.3. Generalization and simplification of the building outline

This important stage is implemented so that the number of vertices describing the building footprint are decreased and therefore simplifying the building outline. The main assumption at this stage is that all the buildings in the scene are described by orthogonal boundaries and therefore the simplification is preserving and enhancing the orthogonality between the linear segments of the building footprint.

There are two parameters which regulate the simplification algorithm, which include the linear tolerance and the minimum size of the polygon. In general, straight lines will be enhanced so that all linear near 90 degrees angles become exactly 90 degree. Based on the given tolerance, isolated small intrusions will be either filled up or widened. Isolated small extrusions will be filtered out. Any building or group of connected buildings with a total area smaller than the minimum area will be excluded from the result (figure 132).

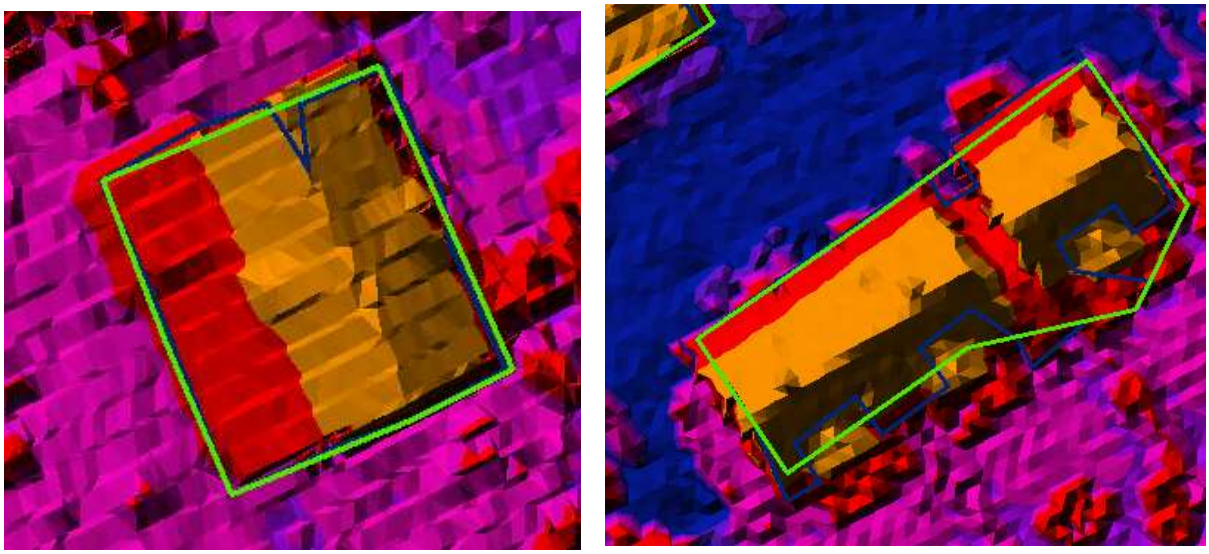


Figure 133 Simplified building outline (green boundary) compared to the initial building footprint (blue boundary) superimposed on a TIN model of the Switzerland study area

The proposed method utilizes a linear tolerance of 3m that presents acceptable results for all three study areas. The minimum area was specified equal to 10m^2. Despite the efficiency of

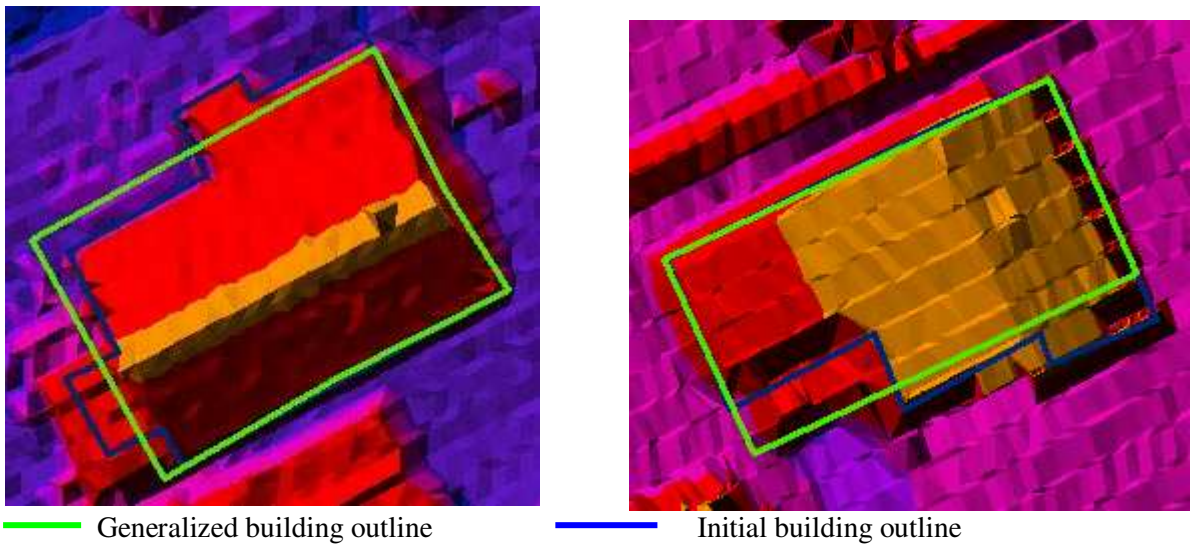


Figure 132 Oversimplified building outline in the London test site, as a function of the increased linear tolerance in relation to the size of the building footprint

the generalization algorithm, to enhance the orthogonality as indicated in figure 132, there is a potential problem of oversimplifying the footprint. An example of oversimplified building outlines over the London study area is depicted in figure 133.

The oversimplification is directly related to the specified linear tolerance and therefore, the results should be evaluated from the operator in order to avoid oversimplifications in the majority of the buildings.

6.5.1.4. Filtering stereo matched points with simplified building outline

Apart from few occasions, where the building boundary is oversimplified from the simplification process, in most cases it could be considered as a good approximation of the optimum building outline and it is subsequently used to filter the stereo matched points.

Initially this stage requires the polygon layer to be converted into a polyline feature class where it is subsequently used to create buffer regions around the generalized footprint. The size of the buffer regions is directly related to the expected planimetric accuracy of the generalized footprint. The buffer distance will also define the maximum possible planimetric correction that can be applied in the simplified footprint. Therefore, this parameter is related to a number of different variables, like the laser footprint size on the ground, the density of points, the building height and the position and direction of the flight path. Even if the above characteristics where known, it could still be difficult to automatically determine a reliable value for the buffer size.

Instead, the proposed method utilizes a search algorithm to determine the buffer size for each individual linear segment. The method initially breaks the polylines into the constituent linear segments. Each linear segment is then treated independently and buffer zones with width of 25cm are created at an increment of 25cm around the linear segment (figure 134). The assumption at this point is that the generalized building footprint will have a similar direction to the actual footprint and therefore the two building outlines will be nearly parallel. For each buffer zone of 25cm width, the algorithm counts the number of stereo matched

points that lie within the zone. If more the 5 points are located inside the zone, the process is completed and the buffer size is defined for the specific linear segment. If less the 5 points are present, the algorithm will create the next buffer zone between the range of 25-50cm from the position of the linear segment and the counting will be repeated.

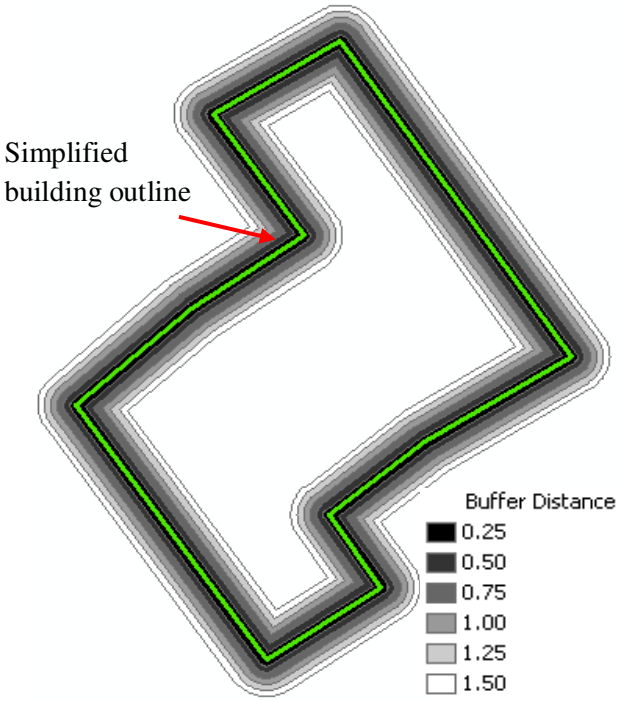


Figure 134 Generated buffer zones of width 25cm at an increment of 25cm around the linear segments constituting the building outline.

Because the stereo matching process may not yield conjugate points, for every segment of the building outline, the proposed algorithm employs a maximum search range that is acting as a termination criterion. This termination function is useful in order to minimize the possibility of adjusting the simplified footprint using stereo matched points not related with the actual building outline. The maximum search range can be estimated approximately as 2-3 times the point spacing of the LiDAR point cloud, as calculated from the density map. For the Switzerland test site the maximum search range was defined equal to 1.5m, which defines the maximum possible planimetric correction as well. Figure 135 depicts the filtered stereo matched points selected with the procedure described above.

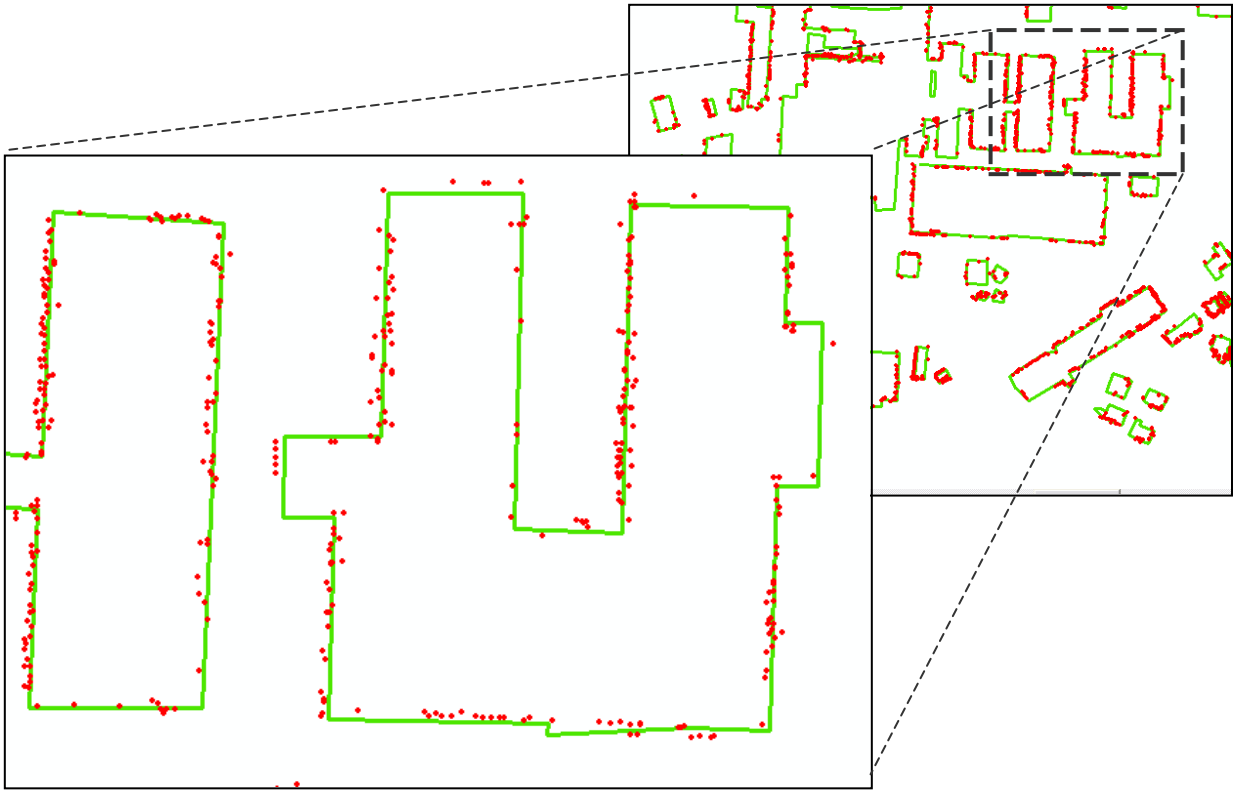


Figure 135 Filtered conjugate points derived from the stereo matching algorithm in LPS (red points) based on incremental buffer regions around the simplified building outline (green polyline)

6.5.1.5. Adjustment and refinement of the building outline

The final adjustment of the building footprint is performed using the filtered stereo matched points. The footprint adjustment uses the individual linear segments that are refined when the sum of the squared distances of the stereo matched points from the line is minimized. This algorithm is implemented in TerraScan and essentially minimizes the perpendicular distance between the points and the line. Using the linear segments instead of the entire outline for the adjustment has the advantage of avoiding distortions from the adjustment and therefore preserving the orthogonality between the lines obtained from the previous stage.

The least squares adjustment is restricted in taking into account only filtered points within the buffer distance. The results from the adjusted linear elements for the Switzerland test site are illustrated in figure 136. The refinement of the building footprints was deemed necessary

for improving the planimetric accuracy, something that was proved successful from the results of the quantitative accuracy assessment, presented in chapter 7.

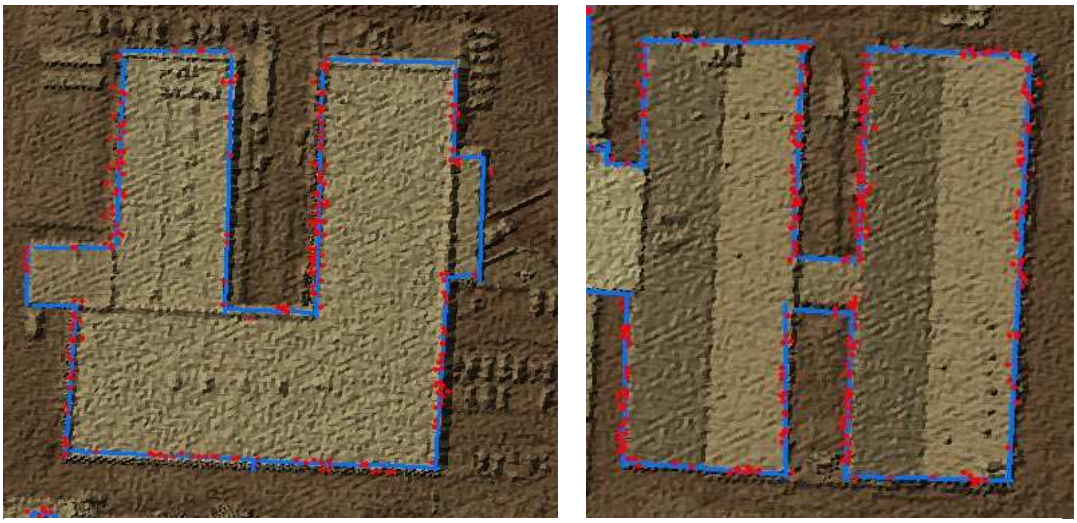


Figure 136 Individual linear segments in Switzerland (blue lines) after the least squares adjustment using the stereo matched points (red)

After the least squares adjustment, each individual linear segment is extended in both directions until it is intersected with a neighbouring line segment. This operation is implemented for creating topologically correct closed polylines representing the building footprint. Figure 137 indicates the differences between the adjusted outlines and the simplified footprint for the Switzerland study area.

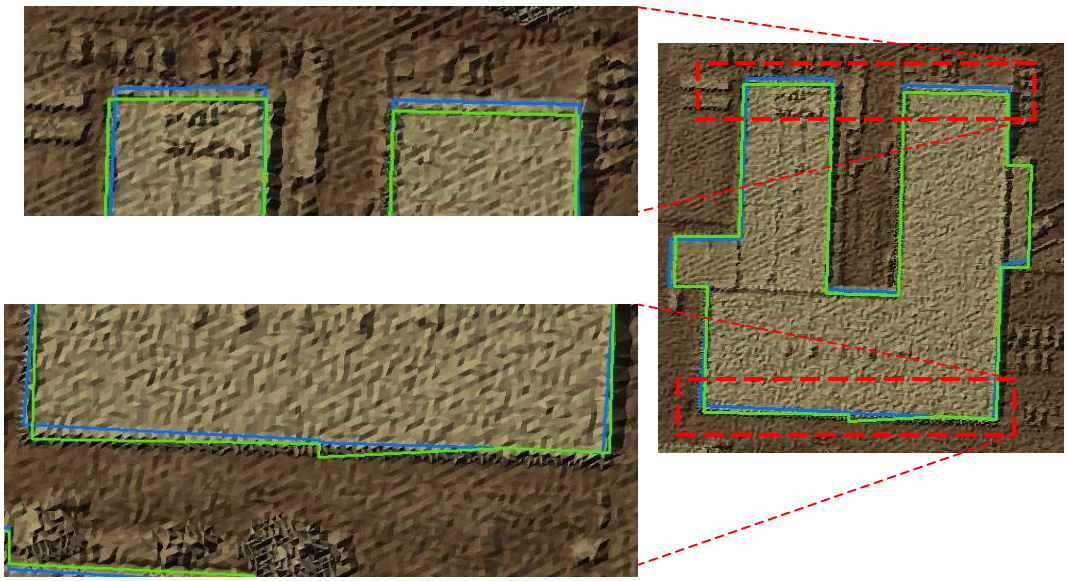


Figure 137 Initial simplified building outline (green boundary) versus adjusted building footprint, superimposed on a TIN model

6.5.2. Data fusion for VDDSMs

When a VDDSM has been classified as described in section 6.4.2, the building class is a raster representation of all the building in the study area. Because a VDDSM is characterised by not only larger planimetric errors of the building footprint but also a lower detail of the building roofs there is the need to enhance the roof details using the stereo matched points. Thus the main difference in the data fusion technique for VDDSMs is that the stereo matched points are not only used to refine the planimetric accuracy of the building footprint but enhance the level of detail of the roof geometry within the classified buildings. The data fusion process that takes place when a VDDSM is present contains the following main steps.

- Deriving the initial building outline from the raster building class
- Generalization and simplification of the building outline *
- Filtering stereo matched points with simplified building outline *
- Adjustment and refinement of the building outline *
- Fitting 3D Linear segments and TIN refinement to improve roof geometric detail
- Converting refined TIN to a dense 3D point cloud

The steps marked with an asterisk (*) follow the same process for the high density LiDAR as described in section 6.5.1.3, 6.5.1.4, and 6.5.1.5 respectively and as such they won't be described in detail.

6.5.2.1. Deriving the initial building outline from the raster building class

From the classification of the VDDSM, as described in section 6.4.2, the building class has been formed as a collection of raster grid cells having the same DN values to represent the building polygons (figures 138).

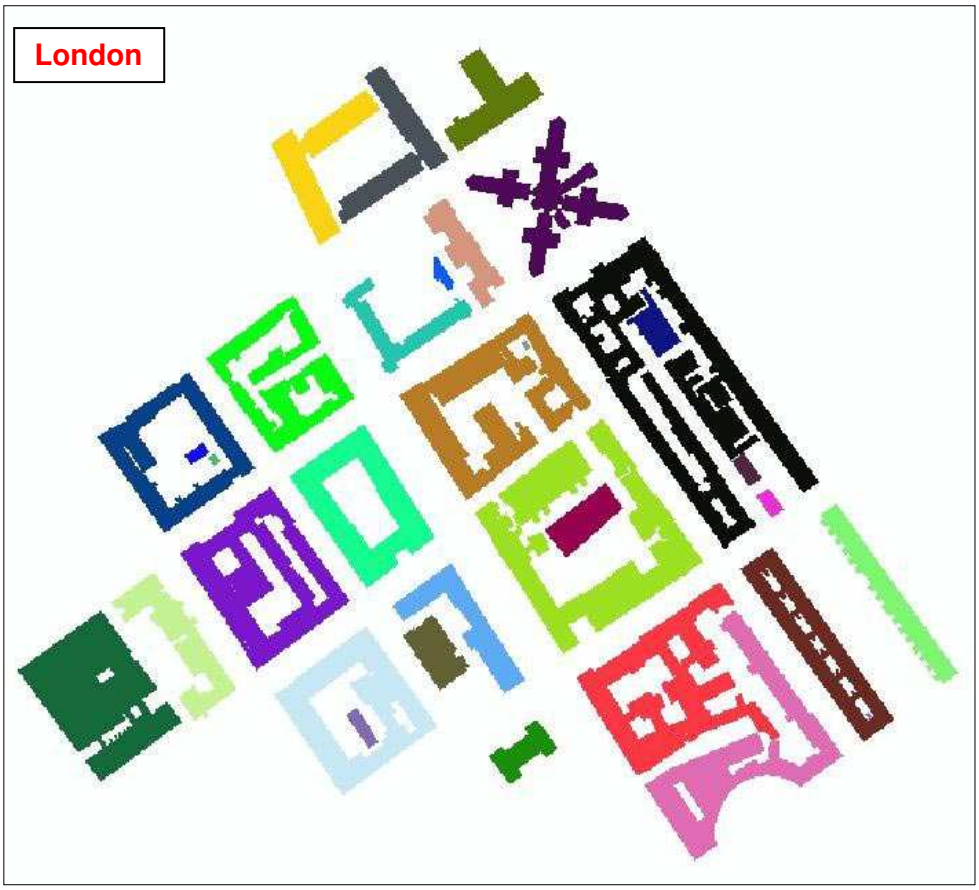


Figure 138 The raster polygons as created from the VDDSM classification process for London and Nottingham. Note each raster building object has been assigned a unique DN value

At this stage the only process that needs to take place is a simple raster to vector conversion so that the raster building class is converted to vector building polygons. The results from the raster to vector conversion are illustrated in figures 139 and 140.

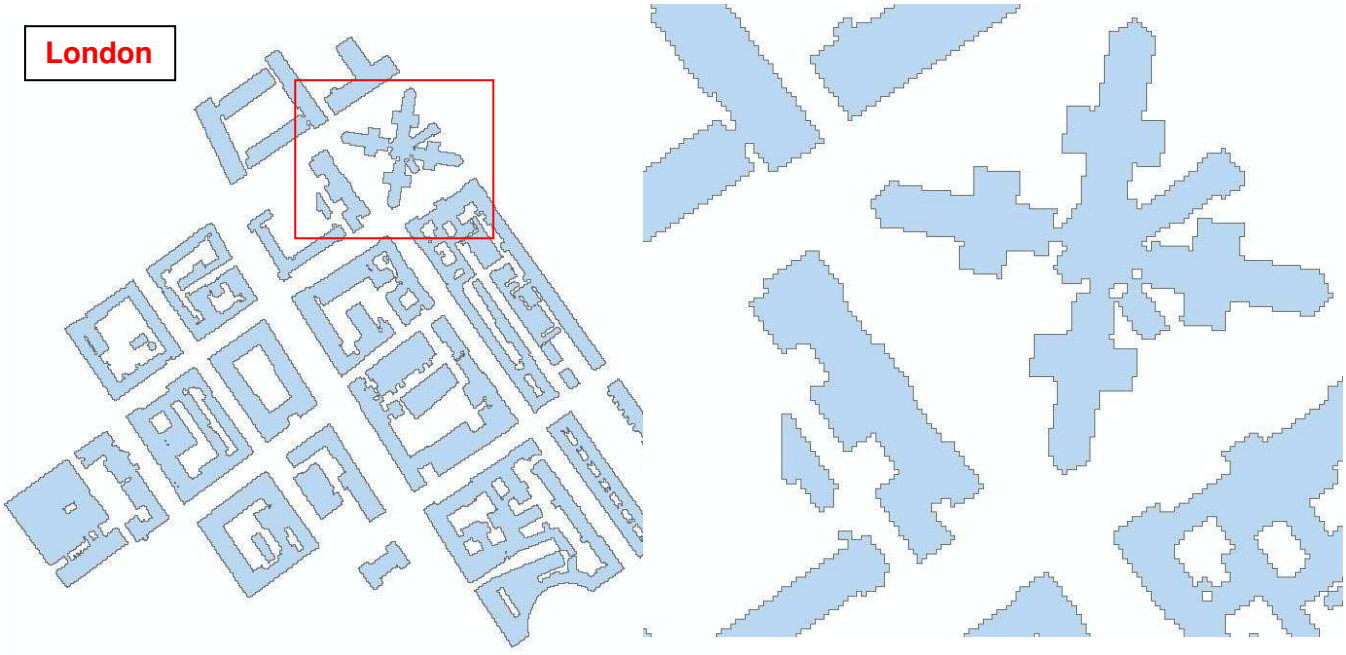


Figure 139 The resulted vector building polygons from the raster to vector conversion in London

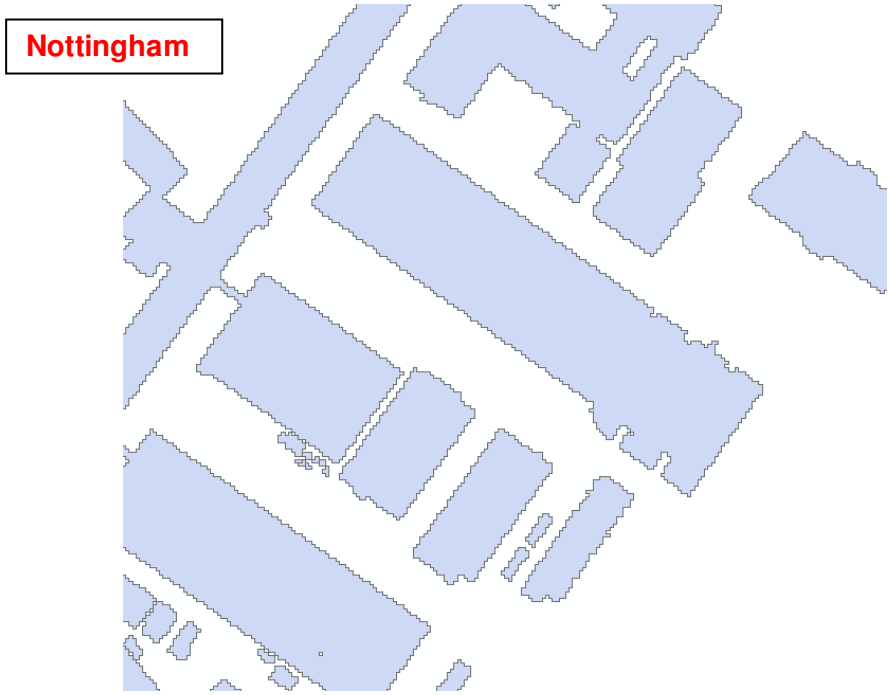


Figure 140 The resulted vector building polygons from the raster to vector conversion in Nottingham

As can be seen from figures 139 and 140 in all study areas the raster to vector conversion creates non-orthogonal/fuzzy outlines and there is a need to implement the generalisation algorithm as described in section 6.5.1.3. The results from the generalisation algorithm are given in figure 141.

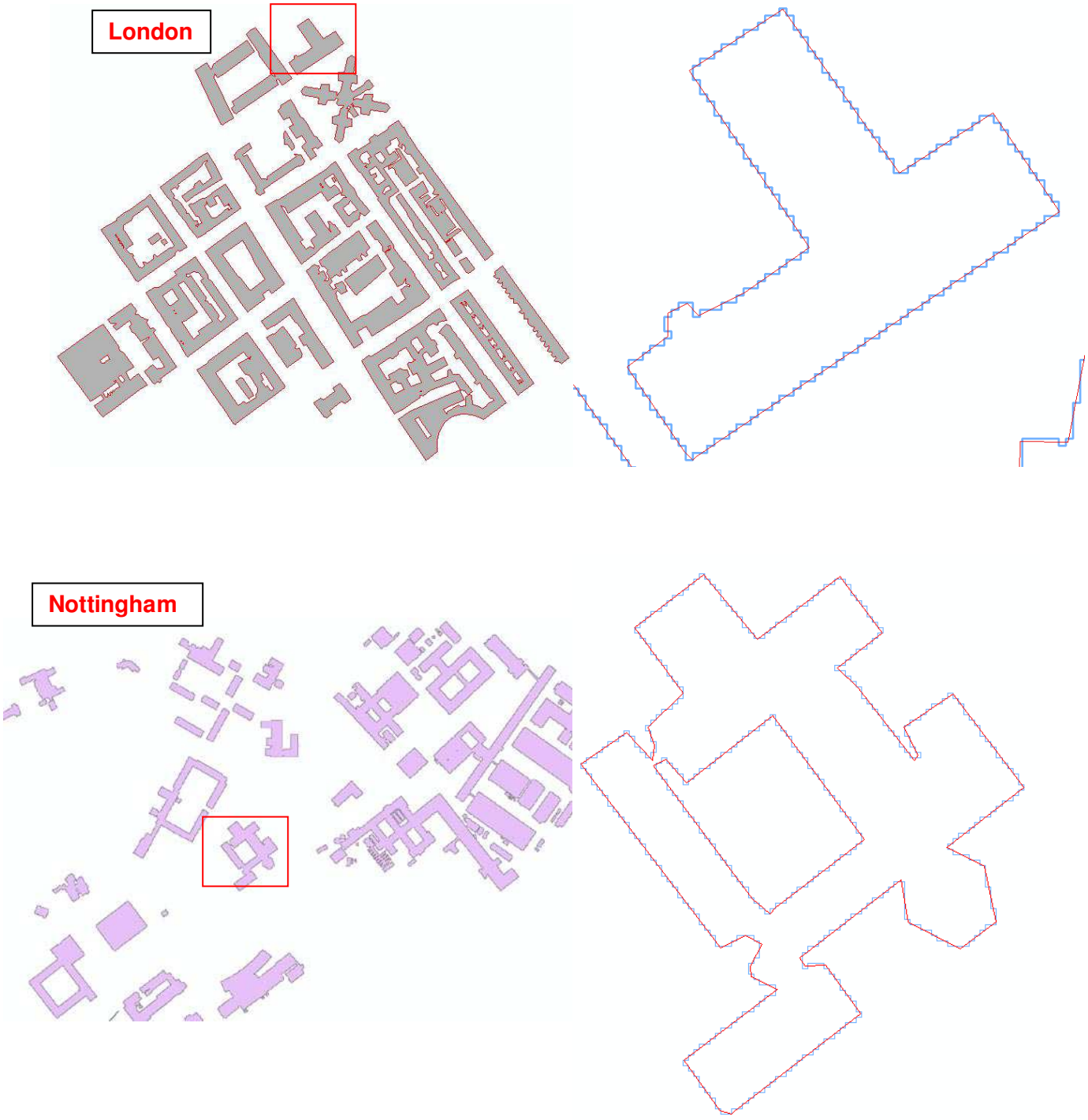


Figure 141 The result from the generalization of the building polygons (left) and the comparison with the initial building polygon prior to generalization (right)

Once the generalised building polygons are created the building outlines can be used to filter the stereo matched points just as described in section 6.5.1.4. The filtered stereo matched points will be used subsequently for the adjustment of the individual linear segments of the generalized building footprints. The filtered stereo matched points are illustrated in figure 142.

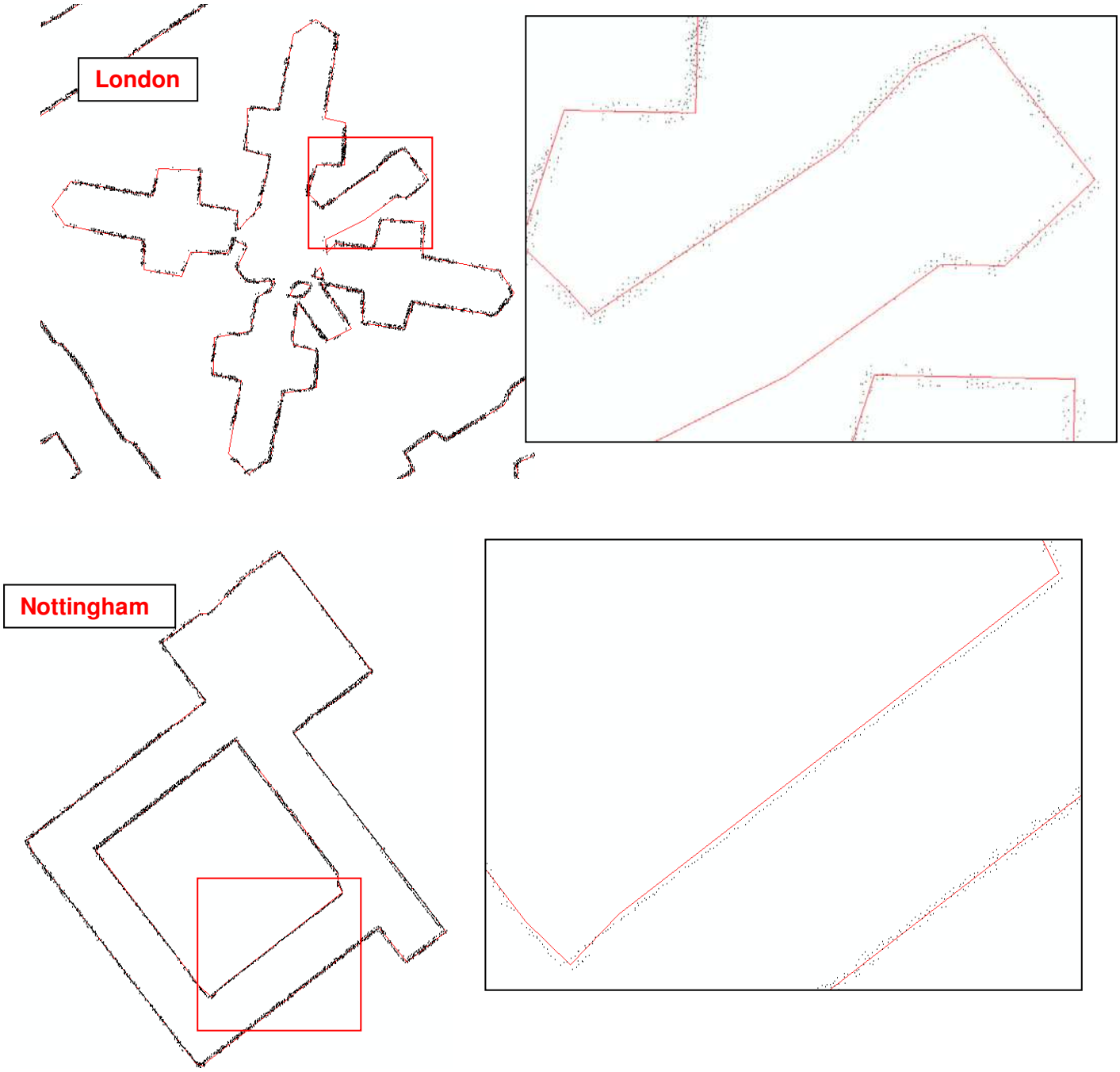


Figure 142 An example of the filtered stereo matched points for adjusting the building outlines for London and Nottingham study areas

Once the stereo matched points have been successfully filtered by the adaptive algorithm, as described in section 6.5.1.4, the workflow proceeds with the implementation of the least square adjustment for the individual linear segments. At this point it should be noted that the planimetric adjustments in most cases are very small (quantitative results given in chapter 7) since the initial stereo matched points were previously merged with the initial LiDAR or VDDSM (figure 143).

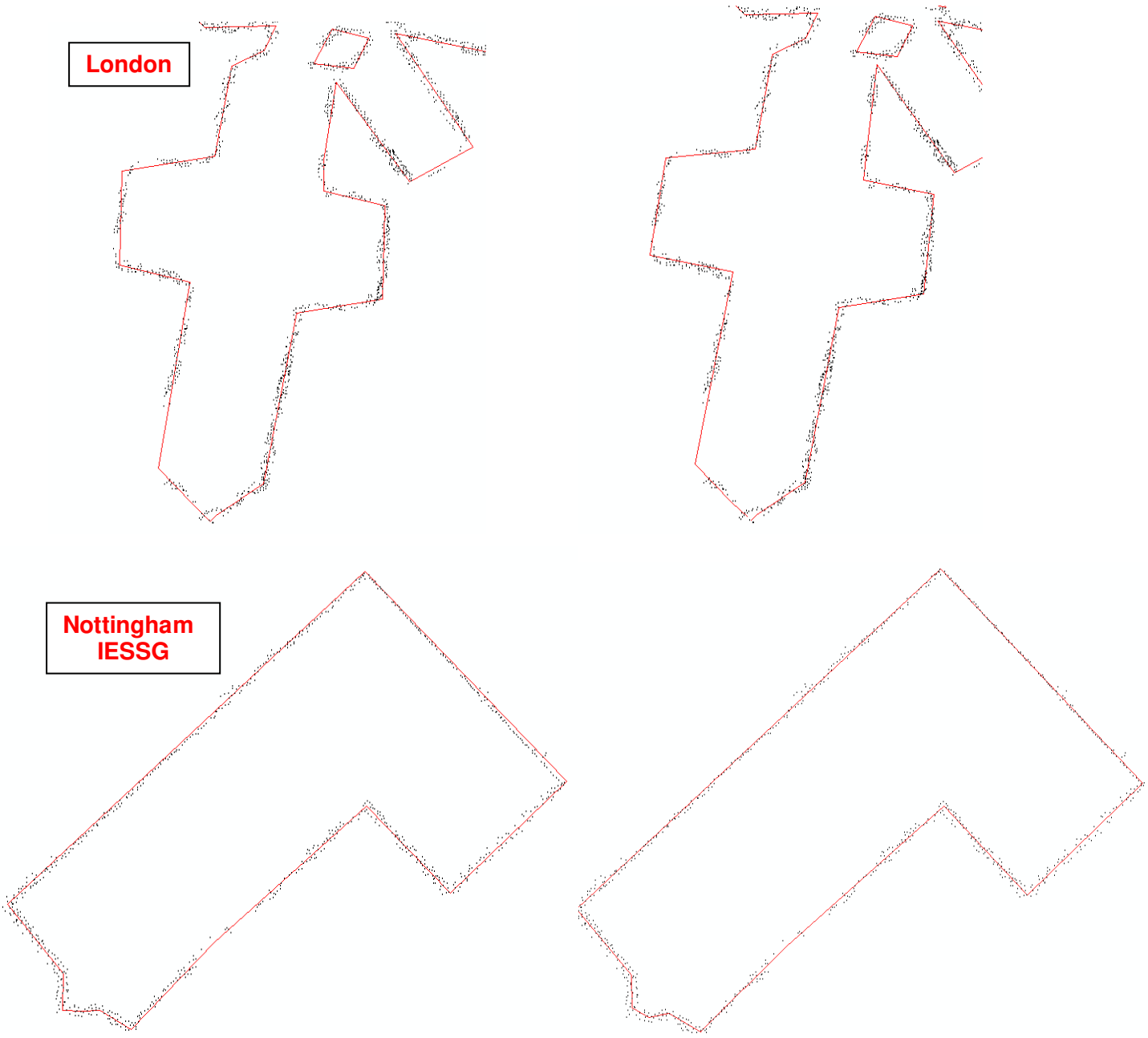
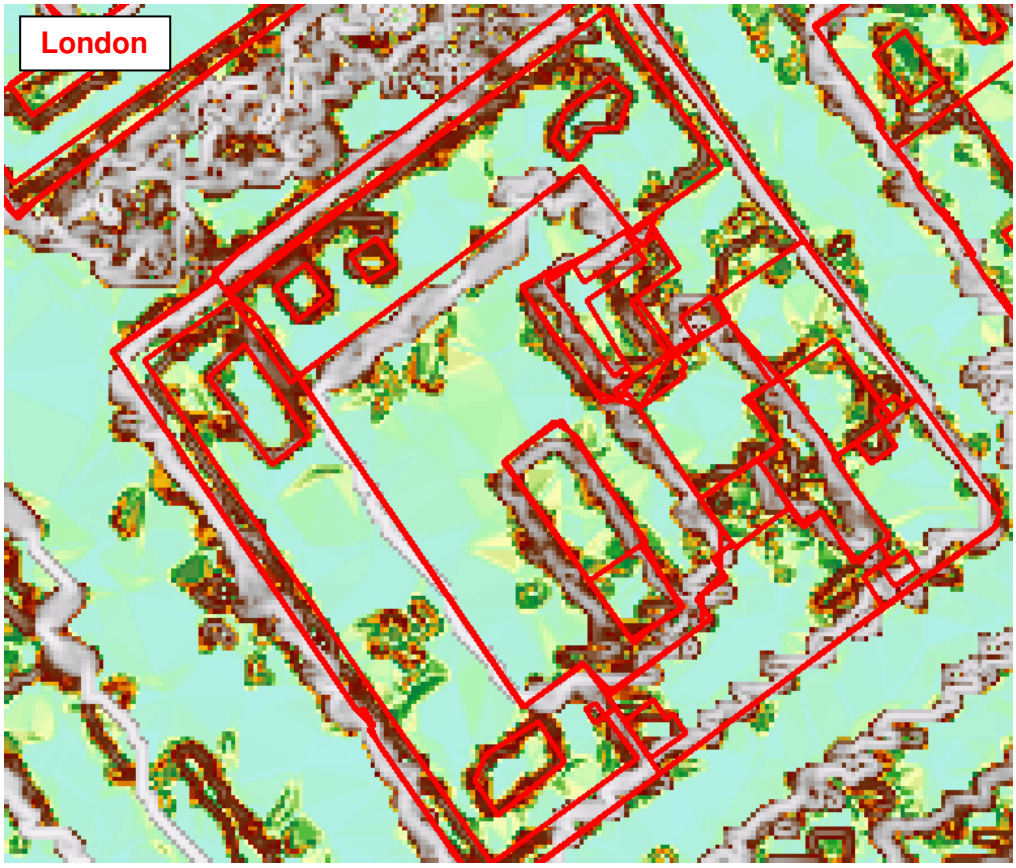


Figure 143 The adjusted building footprints (right) compared to the initial footprint (left) for London and Nottingham using the filtered stereo matched points

6.5.2.2. Fitting 3D Linear segments and TIN refinement to improve roof geometric detail

Within the building footprint the process employs two geomorphometric parameters, the slope and aspect from the VDDSM to define the clusters of points from the stereo matched points. If the partial derivatives of elevation (H) along the east (x) and the north (y) direction are known then slope and aspect (slope pointing direction) are computed from Burrough, 1987 as shown in equations 6.8 and 6.9.

With the slope calculated the workflow applies a slope threshold that segments the breaklines where the slope change between the two sides exceeds 20^0. Note that the slope change threshold is applied only inside the building footprints and the extracted breaklines are also generalized to remove any excessive vertices. From this operation the initial 3D breaklines of roof details are formed (figure 144).



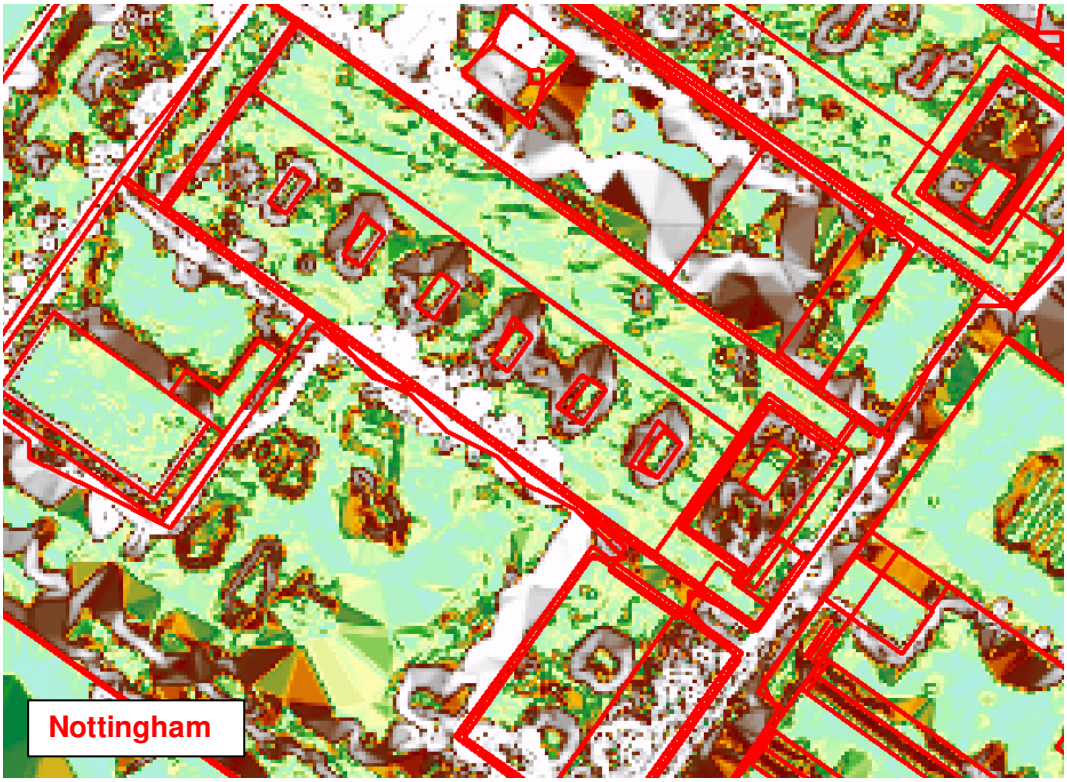


Figure 144 The initial 3D breaklines of the internal roof structures formed when the slope change is greater than 20 degrees.

At this points it should be noted that the efficiency to extract initial 3D breaklines depends on the quality of the original data (LiDAR, VDDSMs and stereo matched points etc). In both study areas there are several instances were smaller breaklines are not successfully extracted based on the slope threshold criteria simply because there is not enough detail on the roof structure. These initial 3D breaklines are nevertheless derived from the VDDSM and may not represent an accurate depiction of the roof geometry. Thus they need to be refined using the stereo matched points extracted from the aerial photography.

At this stage the process employs the iterative buffering technique described in section 6.5.4.1 but applied in the three dimensional space. In more detail the initial 3D breaklines are used as a starting point for the generation of the first 3D buffer zone (creating essentially a cylinder of 25cm diameter). If more than 5 stereo matched points lie within the first buffer the operation stops, otherwise it continuous until the variable is satisfied. At this stage though it is important to note that once the stereo matched points are selected within the buffered there is an additional parameter, that is the calculated aspect. Note that the aspect is

calculated for each pixel of the VDDSM. Therefore the selected stereo matched points will need to have the same aspect. This ensures that the selected stereo matched points belong to the same linear segment. As it has been discussed the calculated aspect is categorized in 8 orientations. The final outcome of this operation is filtered stereo matched points that form cluster of points with the same aspect that are used to subsequently fit the final 3D linear segments representing the roof details. These clusters of points represent in most cases roof outlines as well as several roof details such as chimney outline and dormers.

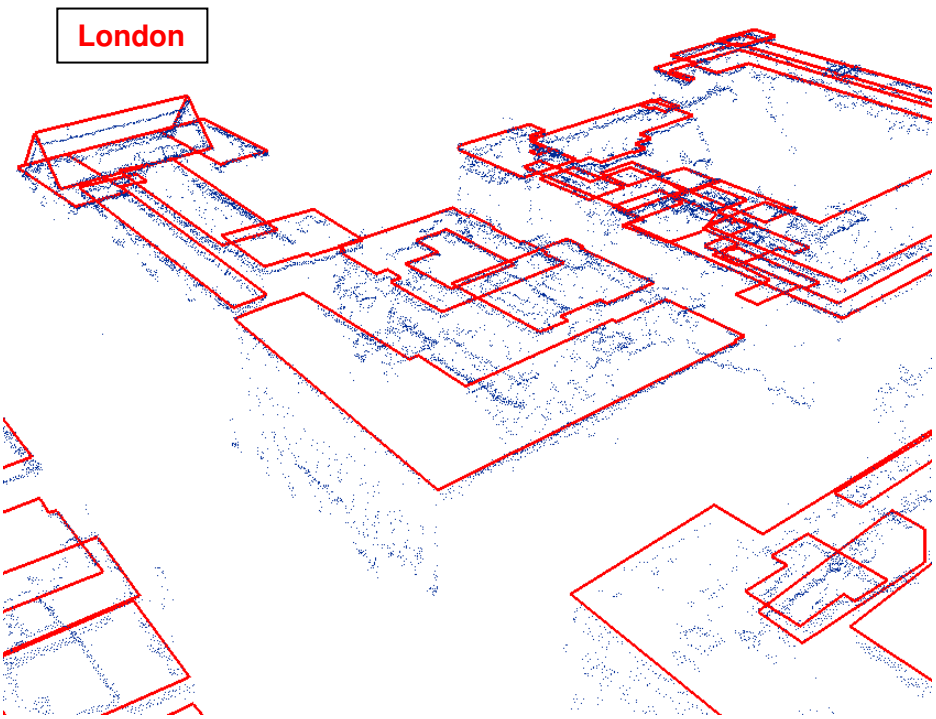
3D Linear segments are then fitted when the sum of the squared distances of the 3D points from the line is minimized. This process essentially minimizes the perpendicular distance between the points and the line. The parametric equation for a 3D line is:

$$Xp = X0 + Vx*t$$

$$Yp = Y0 + Vy*t \qquad (6.12)$$

$$Zp = Z0 + Vz*t$$

Where (X0,Y0,Z0) is some point on the line and <Vx,Vy,Vz> is a vector defining the direction of the line and *t* is the parameter whose value scales the vector to define points away from the line. The 3D linear segments are extended and snapped to adjacent line nodes based on proximity criteria to create closed 3D polylines of roof details although this does not ensure that all breaklines will form closed polylines (figure 145).



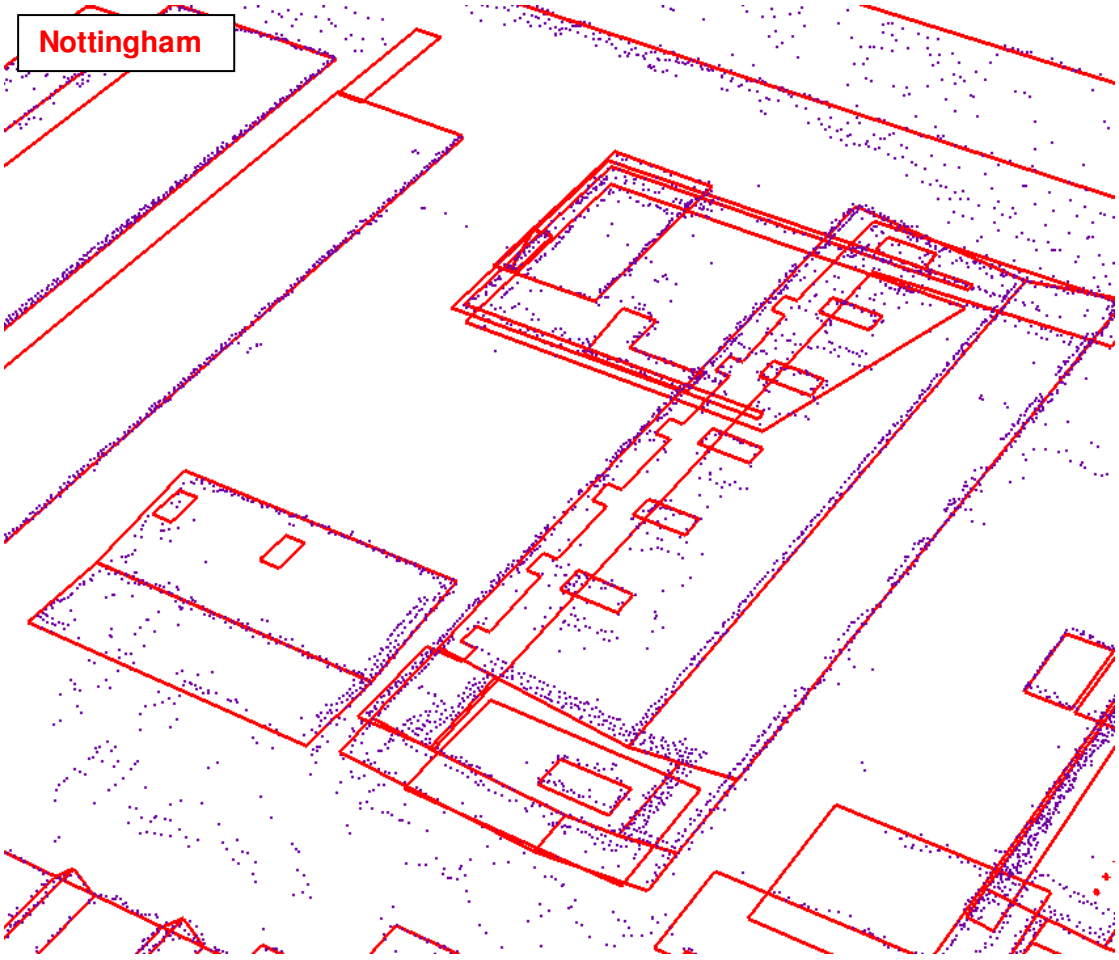


Figure 145 The adjusted 3D breaklines forming closed polylines of roof details in London and Nottingham study areas

In order to avoid potential issues from linear primitives that do not form closed 3D shapes the workflow utilizes a new iterative process for creating a detail TIN model before the final building reconstruction.

In a TIN surface every node is joined with its nearest neighbours by edges to form triangles, which satisfy the Delaunay criterion. The Delaunay criterion ensures that no vertex lies within the interior of any of the triangles in the network. If the Delaunay criterion is satisfied everywhere on the TIN, the minimum interior angle of all triangles is maximized. The result is that long, thin triangles are avoided as much as possible. There is a wide variety of algorithms available to create a Delaunay triangulation from a point cloud. The algorithm described by Watson, (1981) has been one of the most robust methods.

There are certain challenges when trying to generate a TIN building model using detailed 3D linear features. The first issue is that the vertical building facades cannot be represented as such in a TIN surface and the second issue is that the footprints of small roof details exist on the ground level rather than on the main roof top. The first issue can be resolved by expanding the main footprint by a small distance using buffer regions. To tackle the second issue the individual small roof polylines will have to be categorized in a hierarchical way depending on their height level. The proposed workflow for detail TIN generation is described below.

- Employ Delaunay triangulation using the extracted points representing the ground surface to create a bare-earth TIN surface.
- Categorize in different height levels the roof polygons depending on their size and height level.
- Begin iterative TIN generation

    For the first roof level **do**

        Project vertically the 3D polygon outlines on the ground TIN

        Get ground elevation for each node of polygons

        Create final footprint by creating a buffer of 5cm

        Add 3D polygons and footprints on ground TIN and create **Level 1 TIN**

    end

    For each subsequent roof level repeat loop using the TIN generated from the previous roof level.

end

From the implementation of the above workflow in the second and third study area an optimised TIN surface of the building structures is generated as illustrated in figure 146.

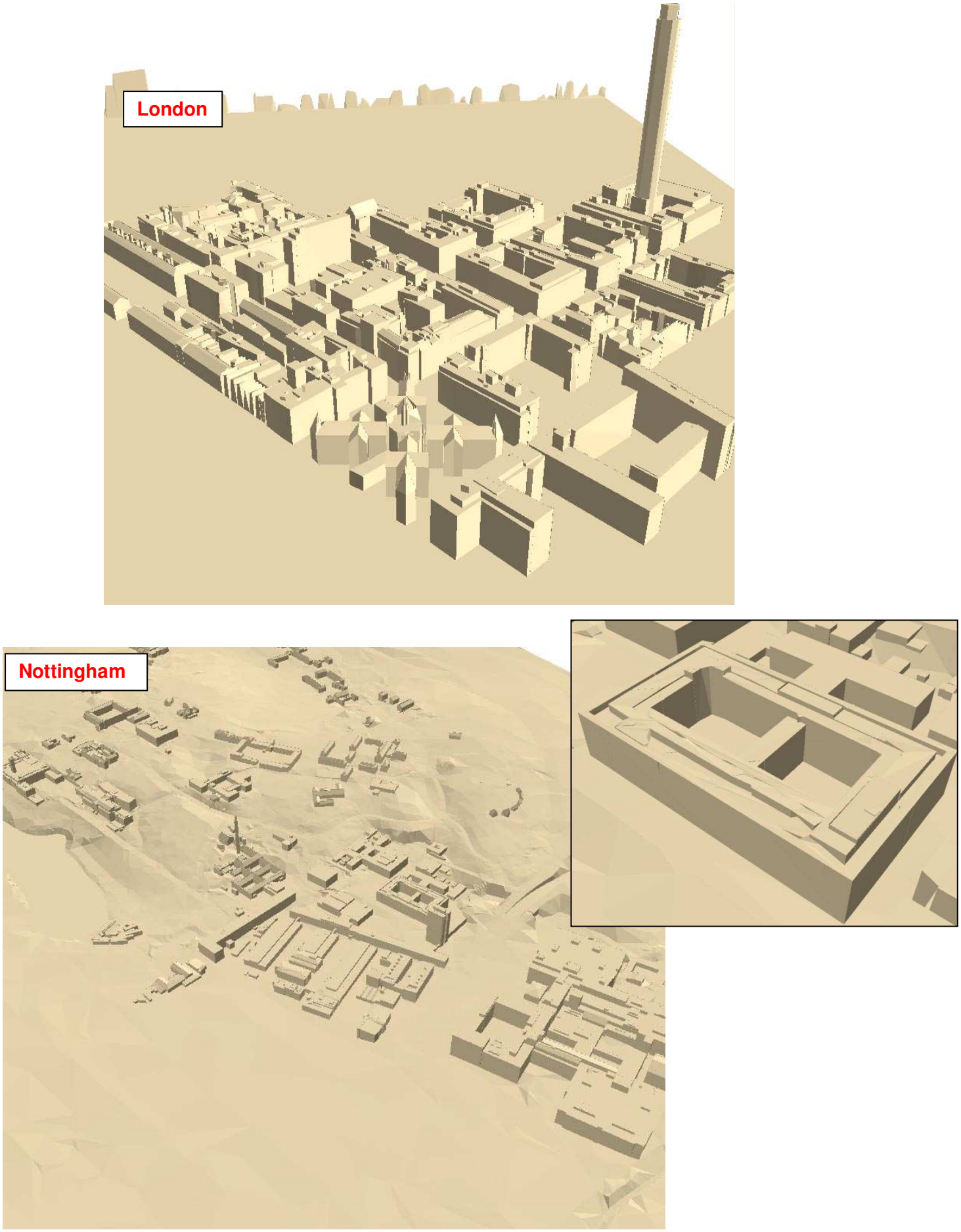


Figure 146 The optimized TIN surface of the building structures for London and Nottingham, for illustration purposed the TIN of the ground surface is also represented.

6.5.2.3. Converting refined TIN to a dense 3D point cloud

Once the refined TIN is generated it needs to be converted to a dense point cloud. The need for 3D points is related to the data input format during the final building reconstruction. The plane fitting algorithm employed uses 3D points.

The conversion of the refined TIN to a 3D point cloud is a trivial process that takes place within a GIS environment using basic GIS functionality. The first step is to create an artificial grid of points with a desired spacing. In our case a 20cm spacing is adequate for most areas and will provide enough points to reconstruct even small roof details. Once the artificial grid of 2D points is generated every point is projected orthogonally upwards until it intersects the TIN surface. At that intersection the elevation value is calculated and assigned to each point. The final output of this operation is a dense 3D point cloud for the second and third study area as illustrated in figure 147.

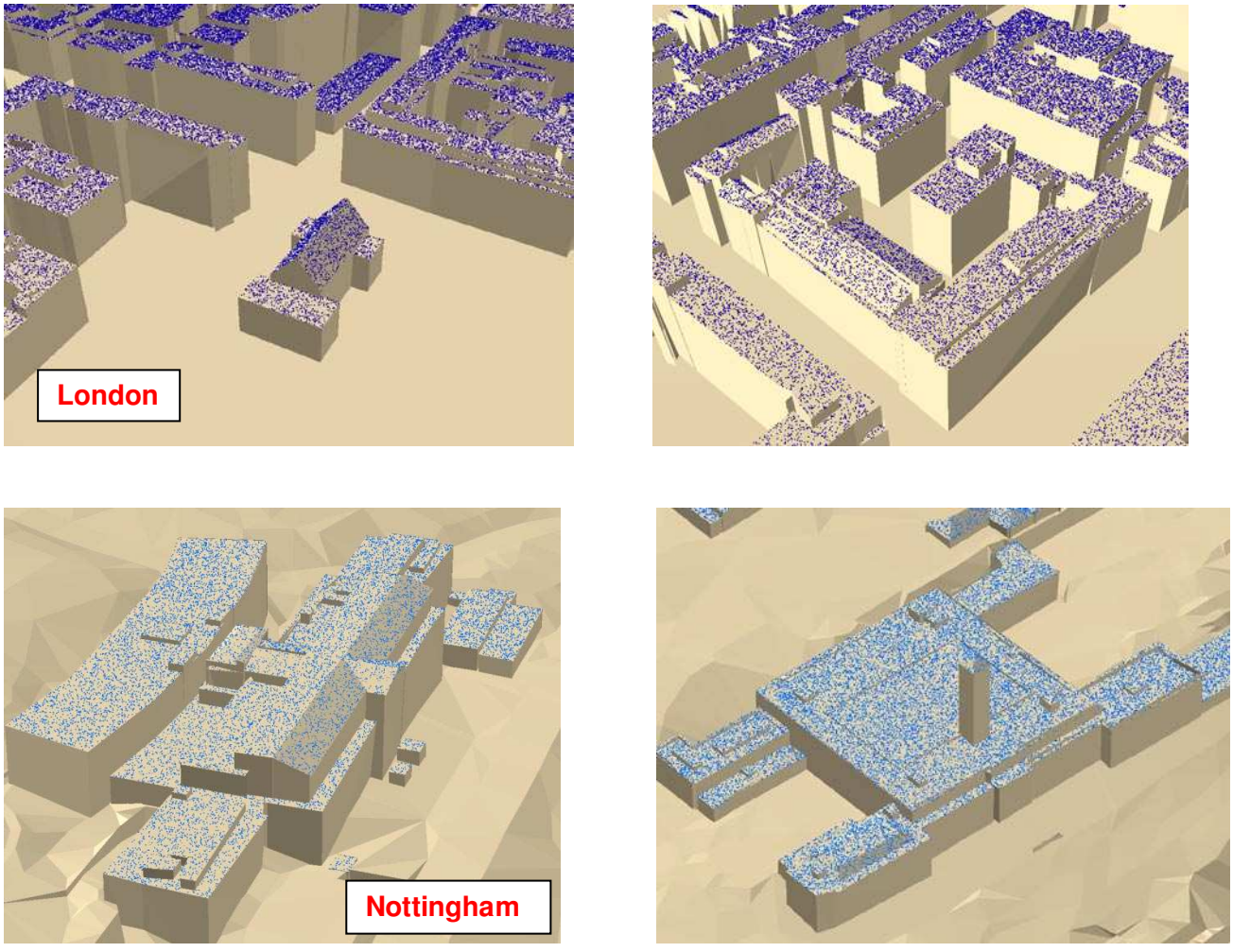


Figure 147 The dense 3D point cloud of the building structures created from the optimized TIN surface

6.5.3. Discussion for the stage of data fusion

From the proposed method for data fusion a generic algorithm can be written as shown in Appendix I (Algorithm 4). The proposed method for generating and adjusting the building outlines introduces promising results. In most cases the method successfully improved the planimetric accuracy (results from the evaluation are given in chapter 7) of the building footprint by incorporating fusion techniques with the stereo matched points. The proposed method is fairly robust by minimizing user interaction, since most of the critical parameters can be automatically defined. The proposed algorithm employs automated estimation for two of the most critical parameters which include the linear tolerance, during the simplification of the footprint, and the robust estimation of the buffer size for filtering the stereo matched points. The robust estimation for the linear tolerance of the footprint simplification initially obtains the lengths of the linear segments between each node. The lengths can be retrieved automatically from the geodatabase that has been created from the generated topology. Then the algorithm estimates the average length of the linear segments. Based on the average length, the algorithm in the next step, selects only the small linear segments with length below the average value. Based on the assumption that these lines will contain the unnecessary intrusions, the algorithm estimates the linear tolerance as the average length value of the selected short linear elements. Nevertheless, the efficiency of the simplification step is arguable and further research is needed to assess the effectiveness of simplifying the footprints. One of the main assumptions in the creation of the buffer zones and the entire concept for estimating the maximum search size is that both data have been geo-rectified with the same level of horizontal accuracy. In other words the accuracy in the determination of the exterior orientation parameters for all the data involved is nearly identical. If the two datasets have varied absolute horizontal accuracy, then the proposed method will not provide reliable results.

The reconstruction of the 3D breaklines in most cases was deemed to provide acceptable results as several roof details are enhanced in the optimized TIN surface. Nevertheless there

are still limitations to the amount of extracted 3D breaklines from a VDDSM that is primarily related to the quality and density of the LiDAR DSM, the efficiency of the stereo matcher and the overall quality of the VDDSM. There were several instances were initial 3D breaklines of roof structures were not successfully extracted and that was especially evident in the London study area.

6.6. Building Reconstruction for generating final polyhedral models

This section describes the procedure of the building reconstruction for the creation of the final polyhedral building models. This is the last part of the proposed workflow and it follows the same process independent of whether High Density LiDAR or VDDSM were initially available.

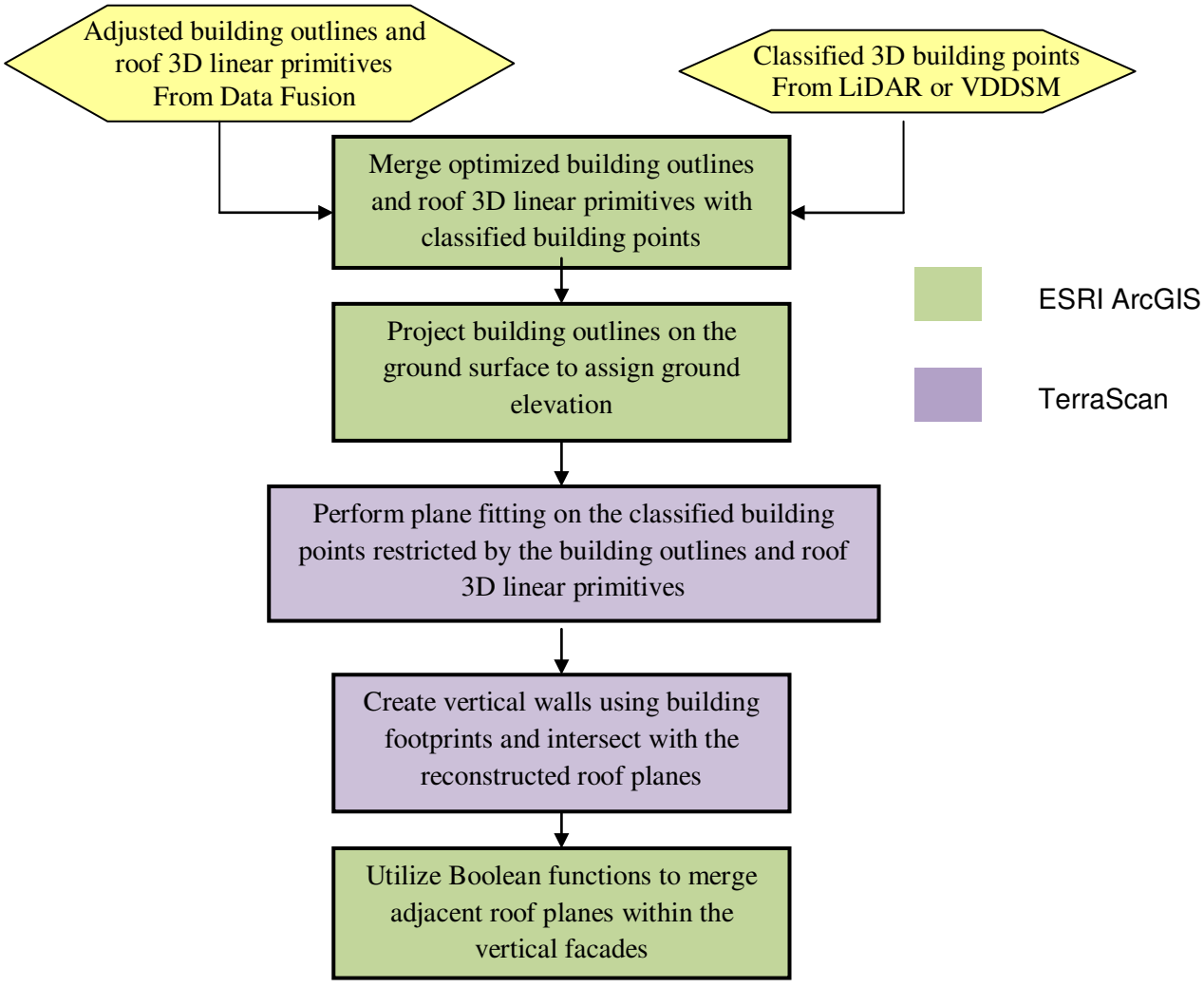


Figure 148 Diagram indicating the overall workflow for the building reconstruction process

Note that as described in the previous section the VDDSM has been converted into a dense 3D point cloud with enhanced point density for the roof details. Hence the 3D point cloud from the VDDSM should ideally be comparable to a high density LiDAR point cloud. The building reconstruction is based on the least medians of squares (LMS) algorithm introduced in section 6.5.1.1. for fitting planes on high density 3D point clouds, while Boolean logic is implemented for merging adjacent planes to complete the polyhedral models. The diagram in figure 148 depicts the overall workflow for building reconstruction.

6.6.1. Combining building footprint with 3D point cloud

This stage is implemented in order to merge the two dimensional building outlines and the 3D breaklines of the roof details with the classified 3D point cloud. Combining the building outlines and roof breaklines with the classified building points will restrict the algorithm during the roof plane reconstruction and provide adequate information for the planimetric position of the vertical building facades.

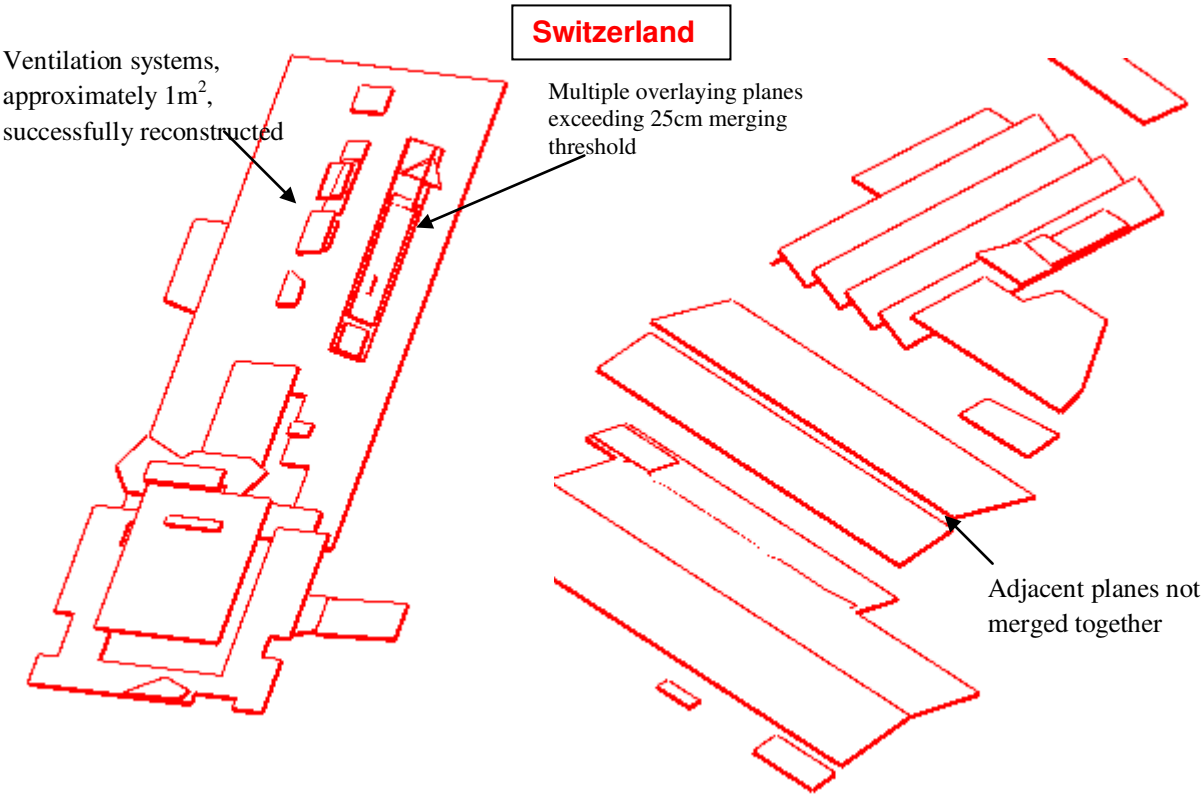
The building footprints are orthographically projected on the ground surface (using the classified ground points) in order to assign elevation values at the vertex of the polylines (figure 149). The workflow automatically defines any intermediate vertices of the building footprint so that the resulting polylines follow the topography. The vertex density depends on the density of the 3D point cloud.

Figure 149 Projected building footprints at the ground level

6.6.2. Final 3D Building reconstruction

The building reconstruction process is based on the least medians of squares (LMS) algorithm introduced in section 6.5.1.1. At this stage the starting locations are not the void regions introduced in the classified ground points, but the building footprints. Furthermore the roof planes are restricted to be within the boundaries specified by the building outlines and the rood 3D breaklines.

This stage consists of the plane reconstruction process and the generation of the vertical walls from the building outline. For the roof reconstruction the same parameters as previously are used, which include the minimum plane size equal to 1m^2, the elevation tolerance equal to 20cm and the minimum vertical difference between adjacent planes equal to 25cm. Figure 150 illustrates a selection of reconstructed roof planes for the three study areas. The triangles for London and Nottingham were formed on the roof shapes because the models were converted to the VRML format for efficient rendering. The VRML format automatically stores 3D facets as triangles.



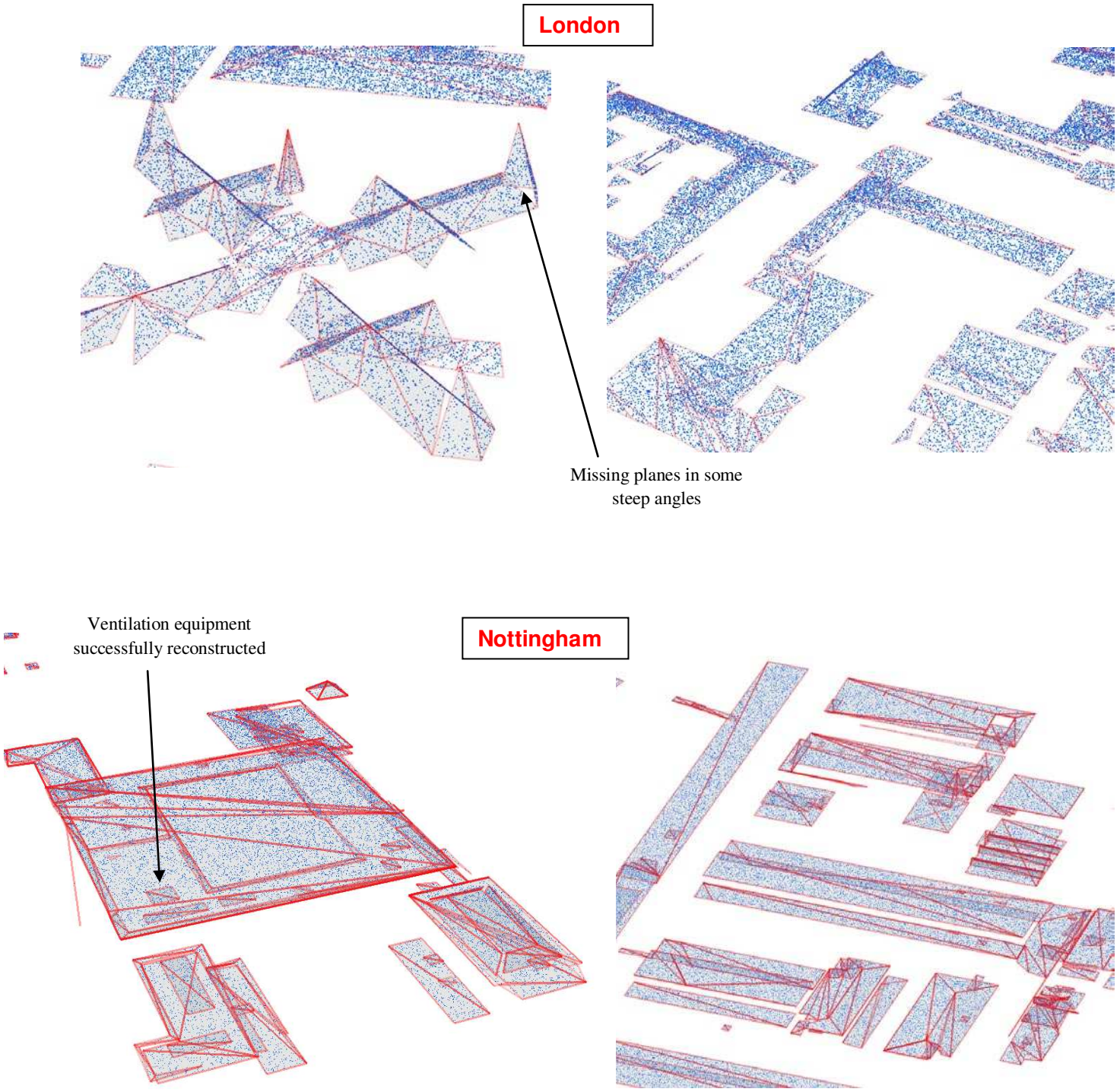


Figure 150 Visualization of the reconstructed planes (hidden lines are excluded) for the three study areas. Successful reconstruction of small roof details (ventilation equipment, dormers etc.)

After the successful roof reconstruction, the vertical building facades are generated by extruding the planes on the projected building footprint as depicted in figure 151.

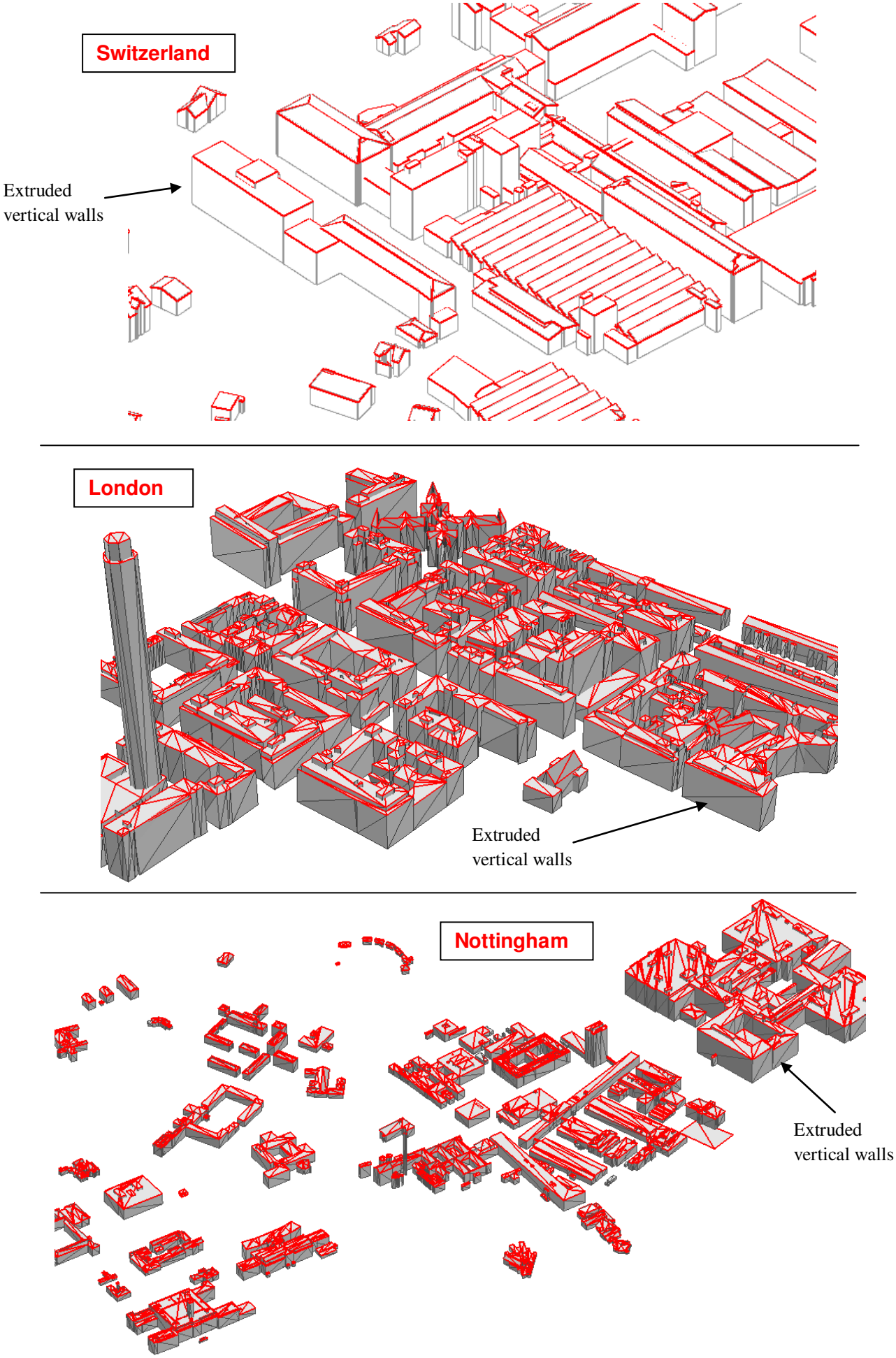


Figure 151 Perspective views for the three study areas (part of the study areas) with the reconstructed roofs and vertical building facades-hidden lines excluded

6.6.3. Merging adjacent roof planes utilizing Boolean functions

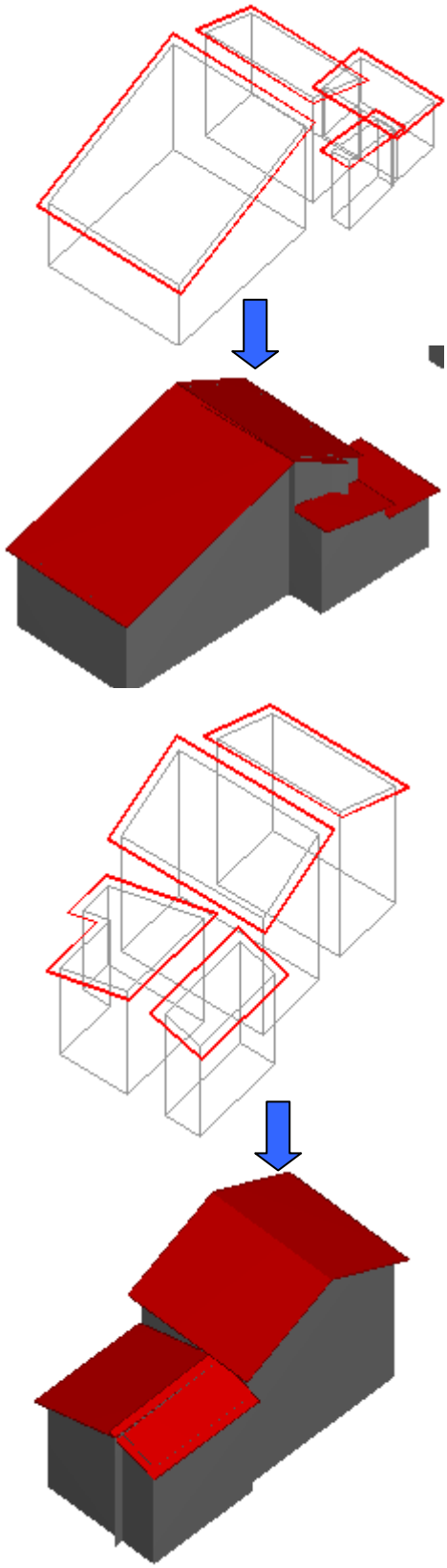


Figure 152 Resulting solid building models (rendered models- smooth shading) from the implementation of Constructive Solid Geometry at the planar facets.

The final stage in the creation of polyhedral building models is the implementation of Boolean functions for merging adjacent roof planes of the same building. The Boolean merging function utilized is essentially converting the "boundary representation" (individual planar facets) of the buildings into a solid feature. In order for this conversion to be performed the planar facets must be transformed into volume primitives and then merged together using the Constructive Solid Geometry (CSG).

The CSG modelling is used widely in computer aided design (CAD) systems, since the modelling is much more intuitive and the primitives can be parameterized. In addition CSG enables the association of the primitives with other, additional information and the determination of volumetric primitive parameters is quite robust.

The proposed method utilizes the CSG for each separate building entity, as defined from the optimized building footprint and therefore there isn't any possibility of merging planes that do not belong to the same building. Within each building boundary adjacent planes are extended and intersected if they are located within 2m from each other. This buffer region is calculated similarly as the "linear tolerance" during the spatial cleaning, as two times the minimum plane size specified

for the plane fitting. Figures 152 and 153 illustrate some examples of the implementation of the CSG in the study areas.

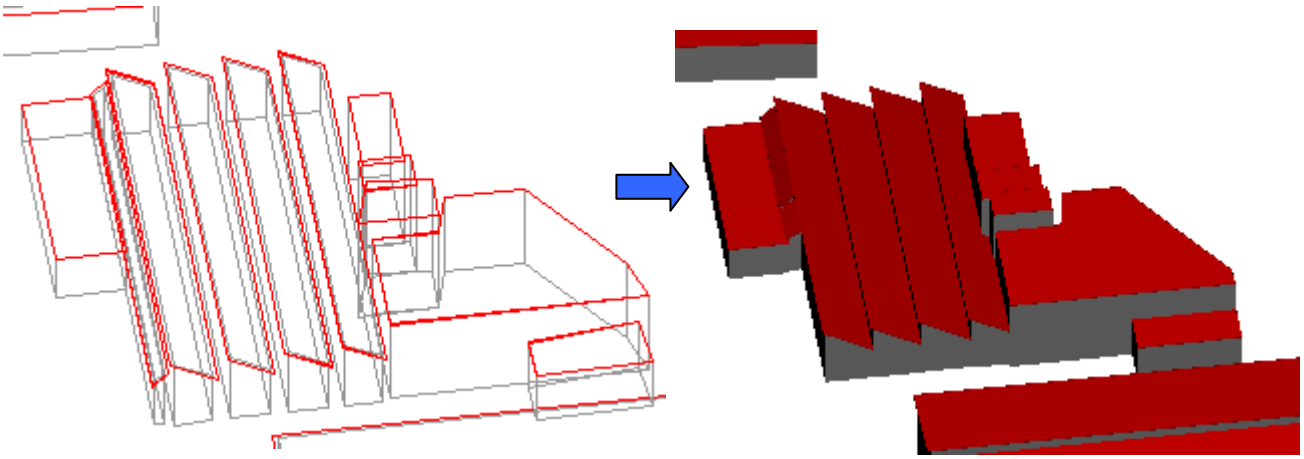


Figure 153 Resulting solid building models (rendered models- smooth shading) from the implementation of Constructive Solid Geometry at the planar facets.

With the implementation of Boolean logic, the building reconstruction process is completed. The final results of the reconstructed building models for Switzerland, London and Nottingham are depicted in figures 154, 155 and 156 respectively.

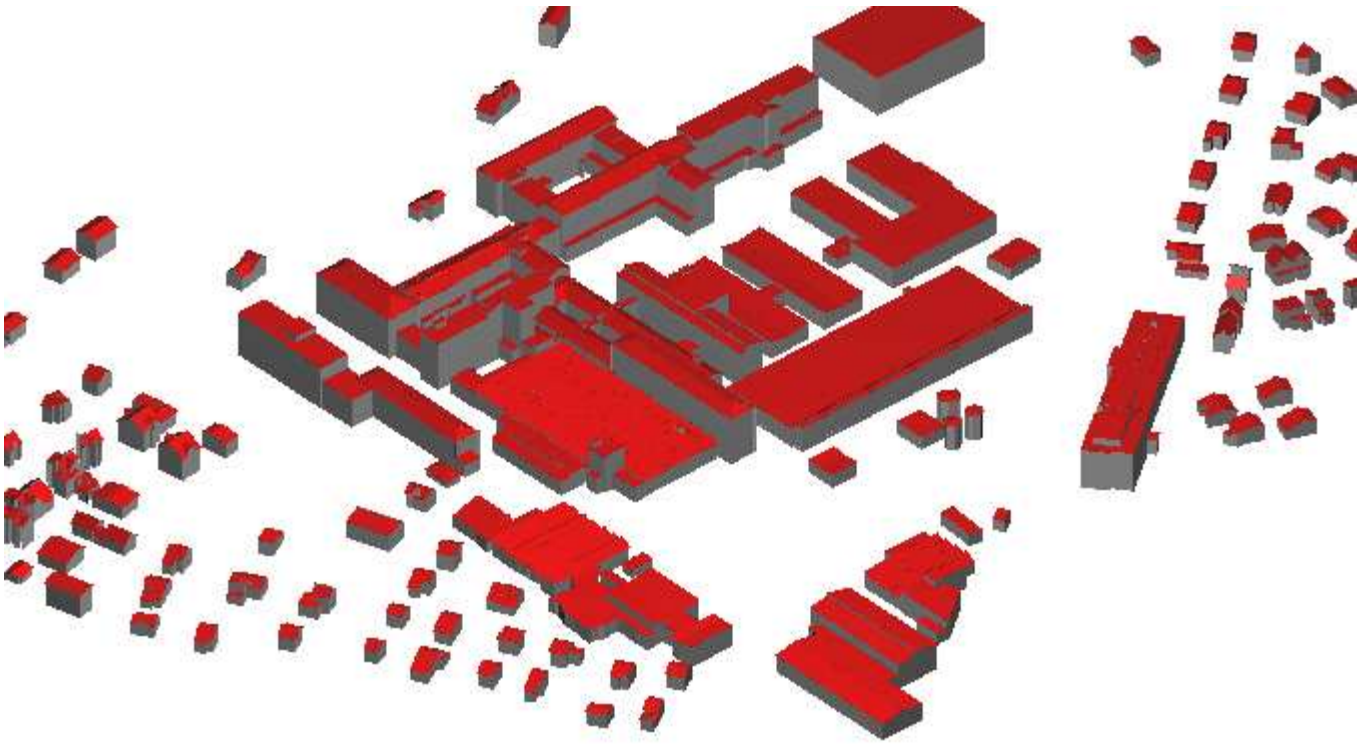


Figure 154 Perspective scene of the final building reconstruction for the entire study area in Switzerland. Rendered scene with global illumination and phong shading

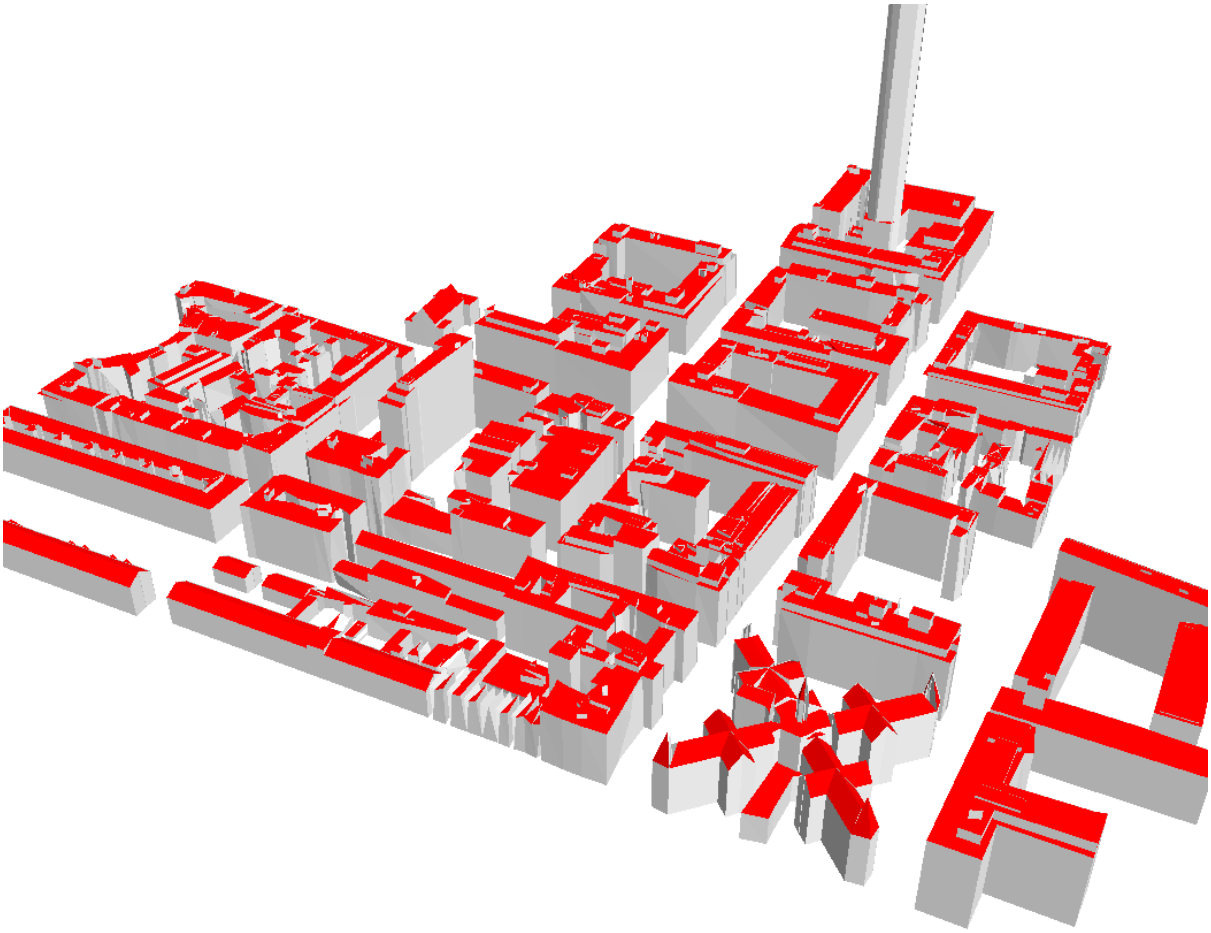


Figure 156 Perspective scene of the final building reconstruction for the entire study area in London. Rendered scene with global illumination and phong shading

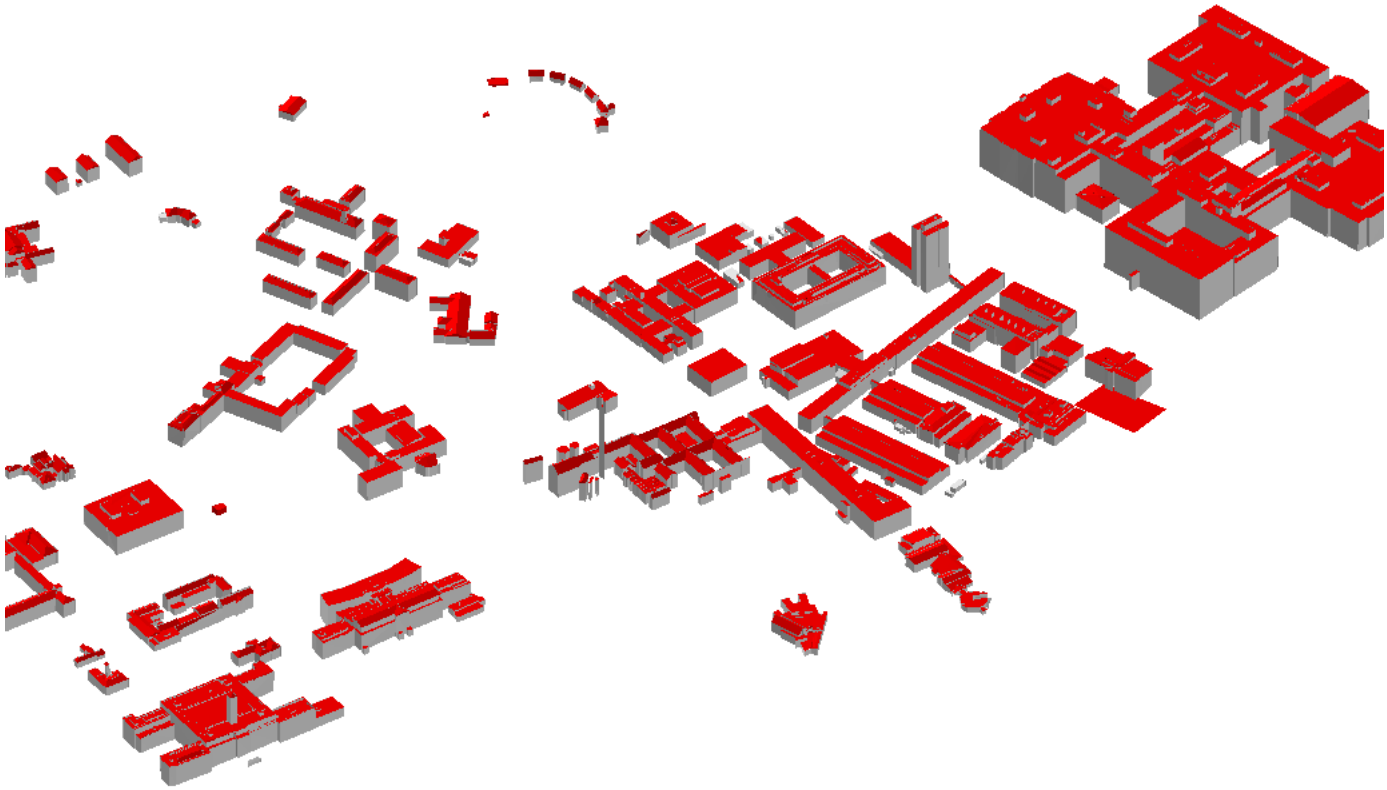


Figure 155 Perspective scene of the final building reconstruction for the entire study area in Nottingham. Rendered scene with global illumination and phong shading

6.6.4. Discussion for the stage of building reconstruction

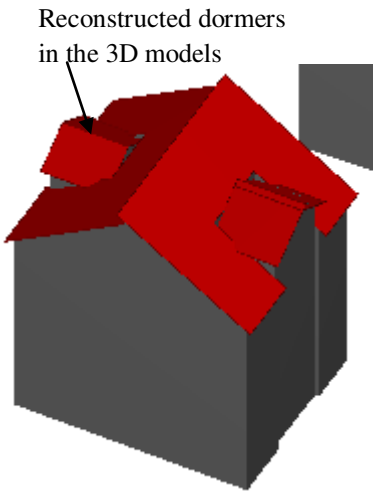


Figure 158 Roof details in the 3D models

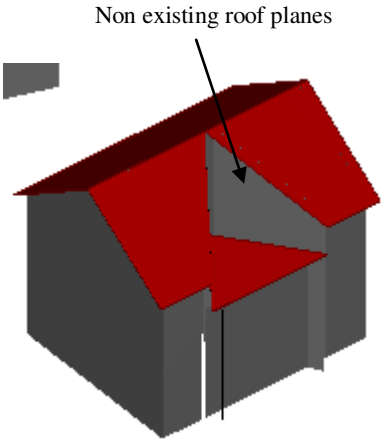


Figure 157 Missing roof planes from 3D building model

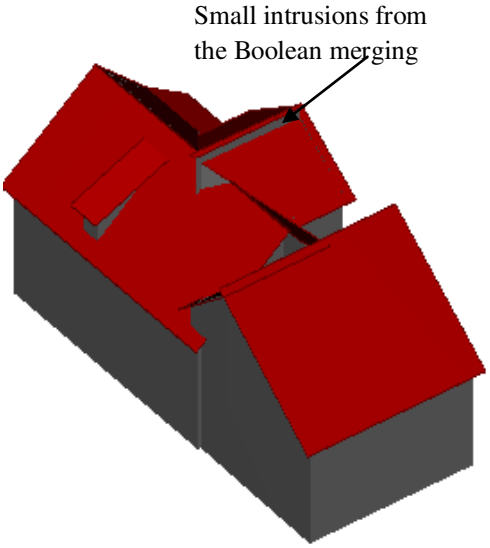


Figure 159 Deficiencies introduced as small intrusions during the plane merging function

From the previously described method for building reconstruction, a generic algorithm is formulated as shown in Appendix I (Algorithm 5). The method seems to be very reliable, with most building models visually correct. In addition, the amount of reconstructed roof details is impressive since ventilation equipment, dormers and chimneys were obtained in many cases (figure 157). The level of detail at the 3D building models is related to the density of the point cloud or the VDDSM the quality of the extracted stereo matched points and the overall radiometric quality of the aerial photographs. In addition the vertical facades of the buildings are created from the refined building footprint and therefore improving the overall planimetric accuracy of the solid models.

Nevertheless, despite the high density of the LiDAR point cloud or the stereo matched features, there are few situations where small roof planes are not reconstructed from the algorithm (figure 158). These situations are a function of the decreased point density on the specific roof planes (less than 4 points) or lack of adequate stereo points to recreate the 3D breaklines, due to occlusions or from the arrangement of the plane (orientation) with respect to the look angle of the sensor. One of the crucial steps is the implementation of Boolean functions for implementing the Constructive Solid Geometry. In most

cases the process yields reliable results with the majority of the adjacent roof planes merged together. There are a few occasions where the snapping function between neighbouring planes is not very efficient, as small intrusions are introduced in the 3D models (figure 159). The following figure shows some additional examples of the semi-automatically extracted building models in London and Nottingham.

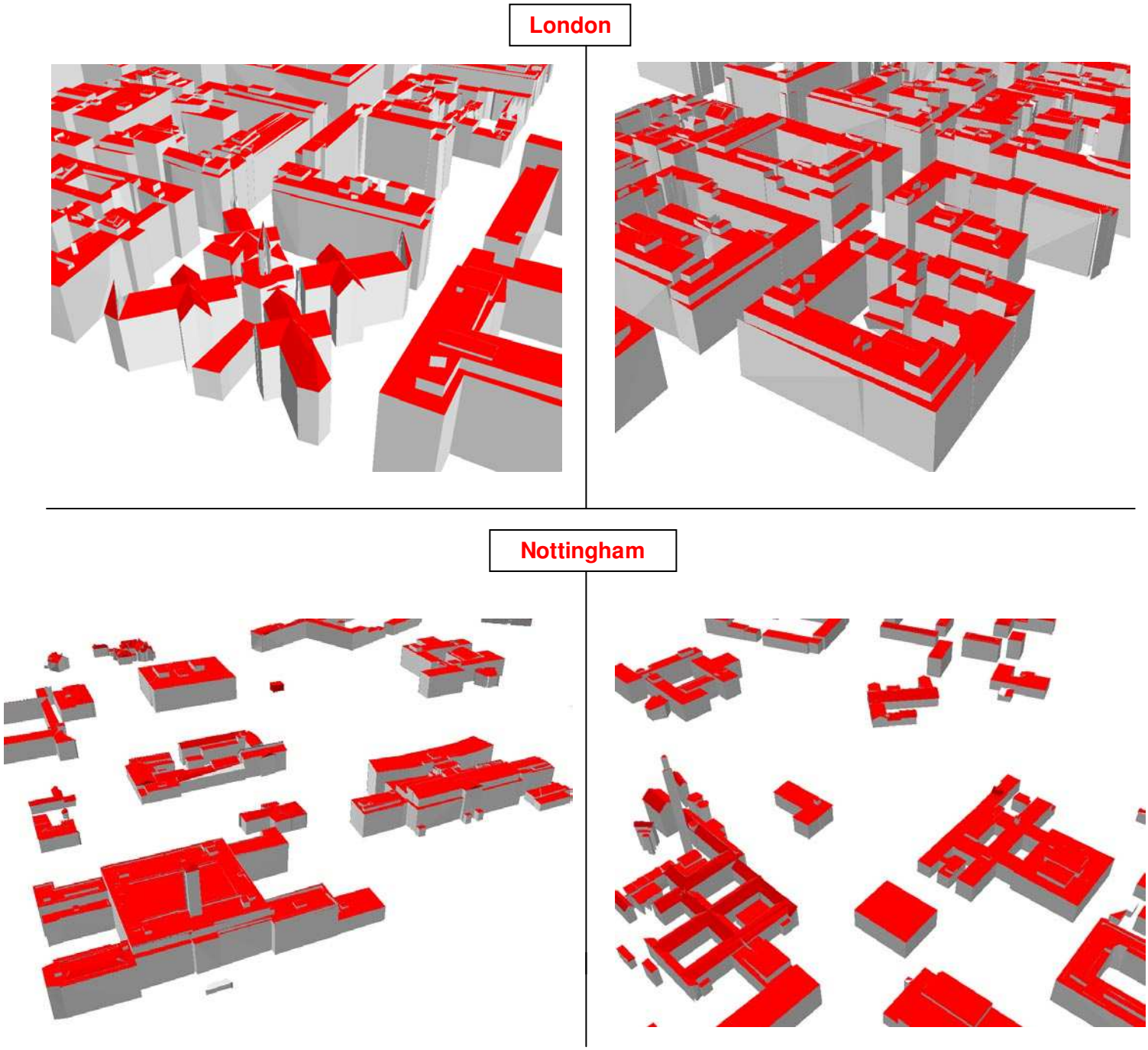


Figure 160 Examples from the semi-automatically reconstructed building models in London and Nottingham

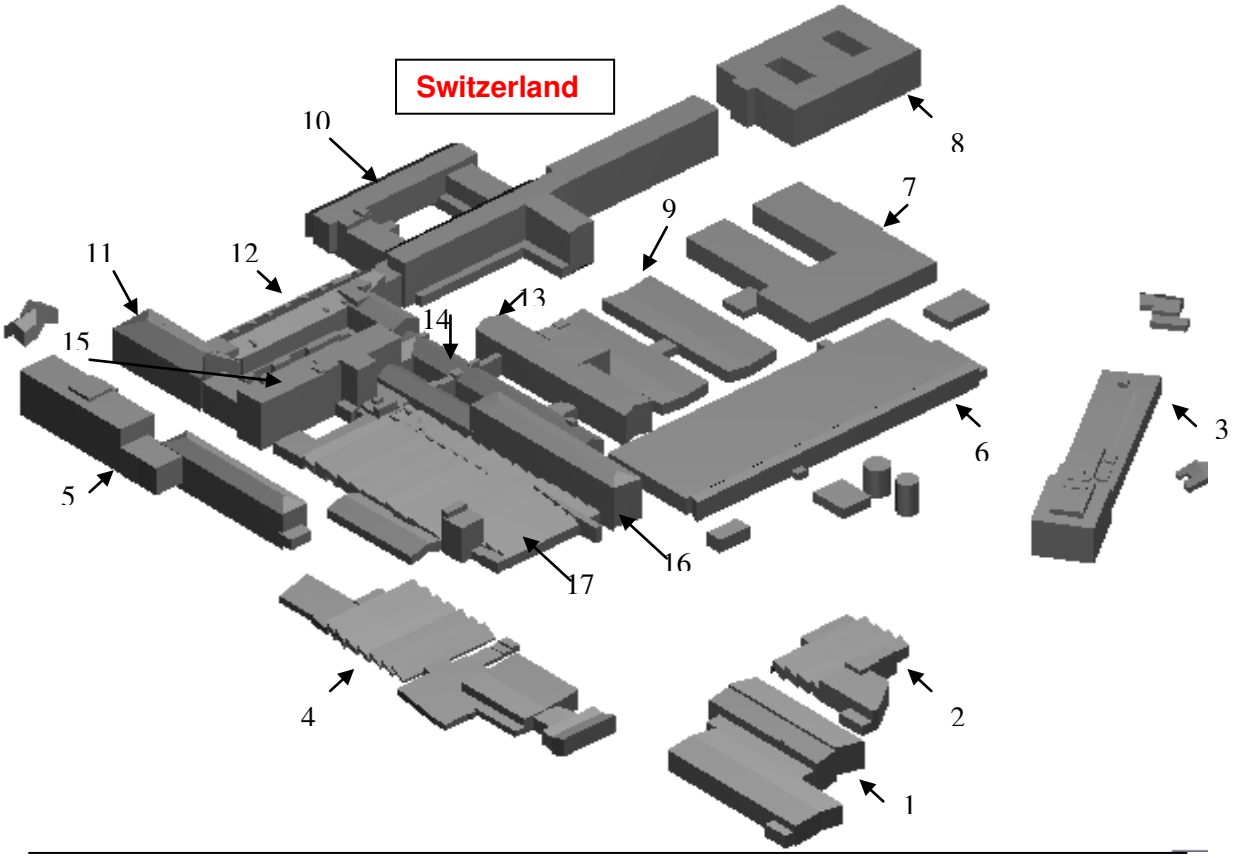
7. PROPOSED WORKFLOW FOR AUTOMATIC QUALITY ASSURANCE AND TEXTURE MAPPING

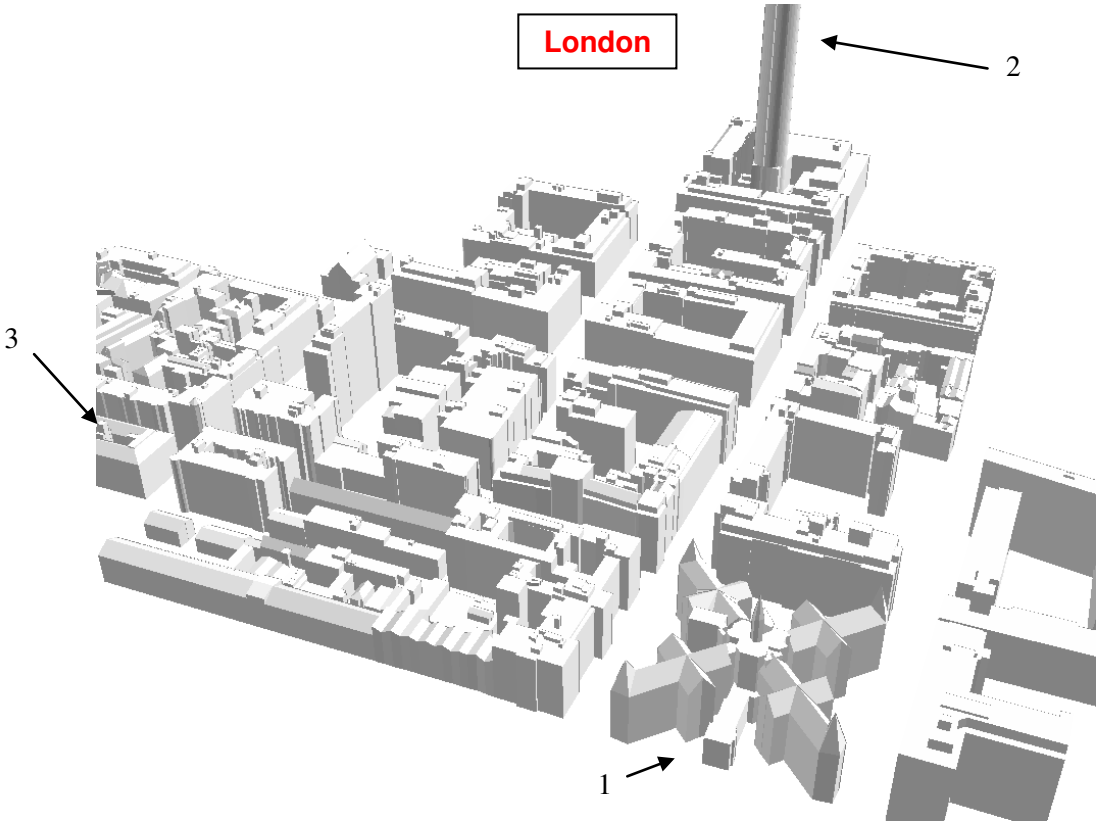
This chapter provides a description of two important aspects of a typical 3D city modelling project. That is the quality assurance and accuracy assessment stage and the automatic texture mapping of the reconstructed models. Due to varied complexity of different urban scenes, automatic or semi-automatic workflows for building reconstruction should be accompanied by robust quality assurance tools. The existence of these tools is critical to the adoption of automated algorithms in the industry to ensure the final products meet client requirements and project specifications. So far most accuracy assessment studies have been focused on assessing 3D city models using abstract methods by assessing either the planimetric accuracy of the building footprints or visually comparing roof shapes. This research presents a robust semi-automated approach for a comprehensive planimetric and vertical accuracy assessment of a selection of the reconstructed 3D building models. It should be noted that the design and reliability of automatic quality assurance methods is a function of the Level of Detail (LOD) of the city models. The methods described in this report were designed for the lower and mid LOD city models up to LOD2 (Kokkas and Smith, 2007).

7.1. Semi-automated Quality Assurance workflow

In order to perform any kind of quality assessment there is the need to collect some reference data for the three study areas. The reference data should have higher accuracies when it comes to the planimetric and vertical accuracy of the reconstructed buildings and also have superior roof detail. The reference data for evaluating the planimetric and vertical accuracy of the semi-automatically reconstructed building were obtained from manual stereo plotting using the stereo pairs of aerial photographs. In an industry environment where large datasets need be checked the proposed workflow can be utilised for a sample of buildings in

the area. The manual plotting of the 3D building models was performed by Leica Stereo Analyst and was a tedious process since the highest amount of roof detail was required for a meaningful comparison. In addition, certain roof details in the Switzerland study area were difficult to extract due to shadow effects or low contrast of the ADS40 imagery. The stereo plotted buildings obtained for the three study areas are depicted in figure 161.





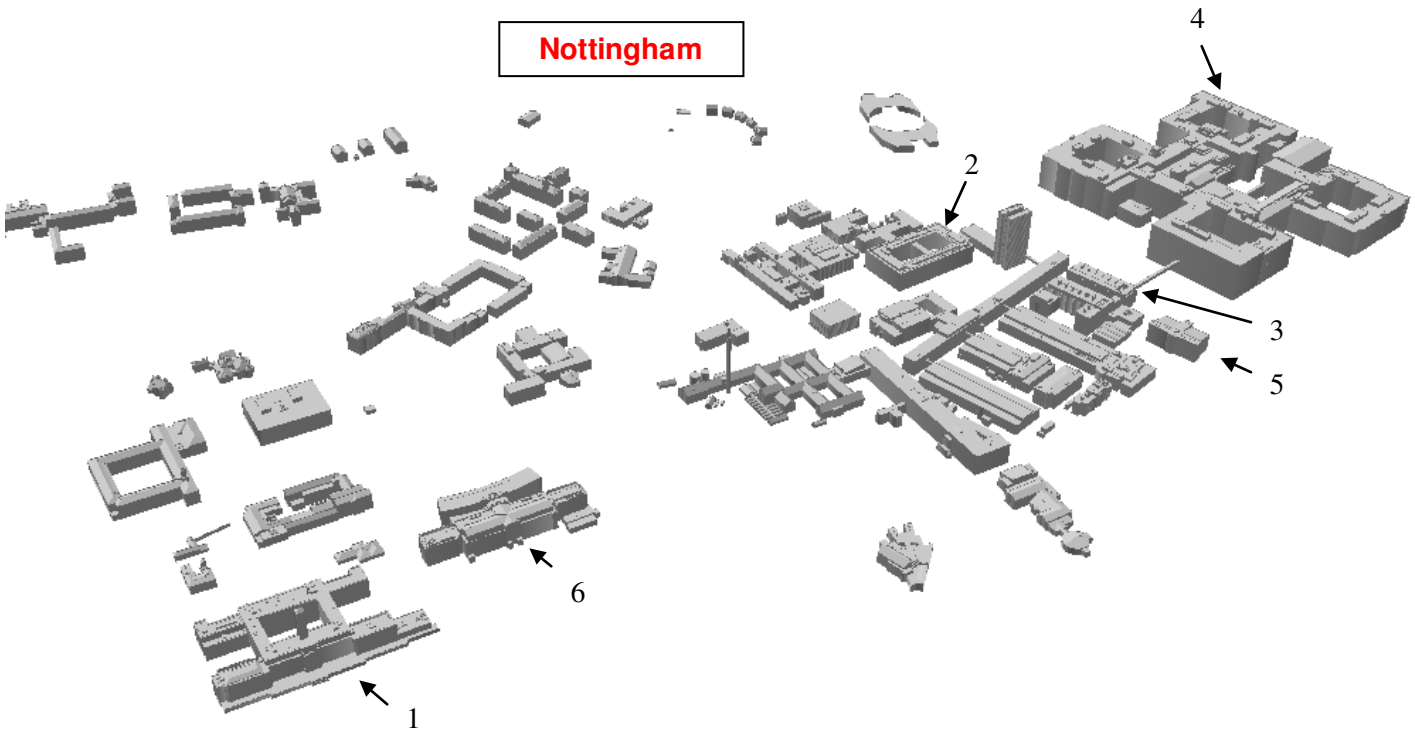


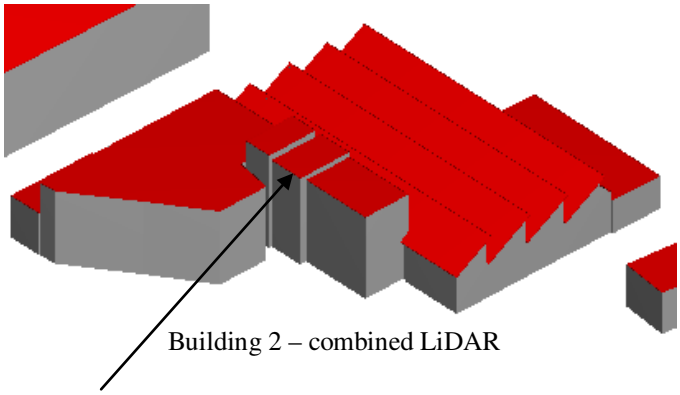
Figure 161 Stereo plotted buildings for the three study areas. Rendered with a smooth shading effect and global illumination parameters

The accuracy assessment is divided in the quantitative and qualitative evaluation. The qualitative evaluation includes a visual comparison between the reconstructed buildings and the reference stereo plotted buildings. This comparison will provide a useful indication of the overall quality and therefore, only a selection of the most interesting structures will be illustrated in this section.

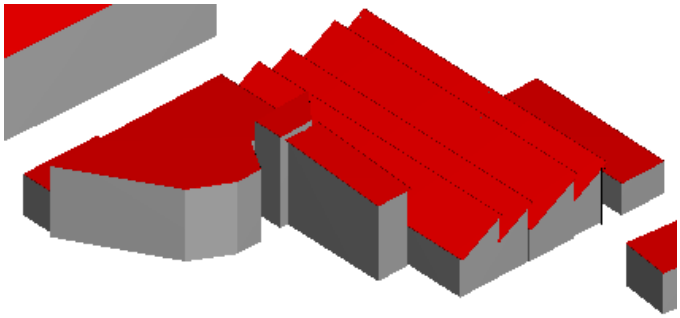
The quantitative assessment performed on the reconstructed building models evaluated the planimetric and vertical accuracy. The planimetric evaluation is conducted by employing the widely used building detection metrics suggested by Shufelt, (1999). For the vertical accuracy assessment a robust method is proposed for evaluating each building separately, in order to provide statistical parameters for the vertical accuracy.

7.1.1. Qualitative evaluation of the reconstructed building models

The qualitative evaluation is performed for a selection of the total reconstructed buildings. The numbering system used for identifying each building is depicted in figure 161. At this point it should be mentioned that for the first study area in Switzerland the entire process of the building reconstruction workflow as described in chapter 6 was applied once on the combined LiDAR data (density of 5pt/ sq.m) and once on the single LiDAR data (density of 3 pt/ sq.m). This will allow us to evaluate the efficiency of the proposed workflow under different LiDAR densities. Note that in both cases the data for the Switzerland area were classified as High Density LiDAR and thus followed exactly the same workflow as described in the previous chapter. The qualitative comparison for the reconstructed building in Switzerland is given in figure 162.



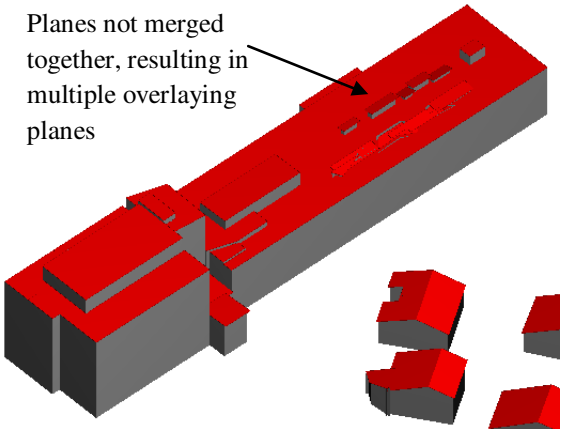
Building 2 – combined LiDAR



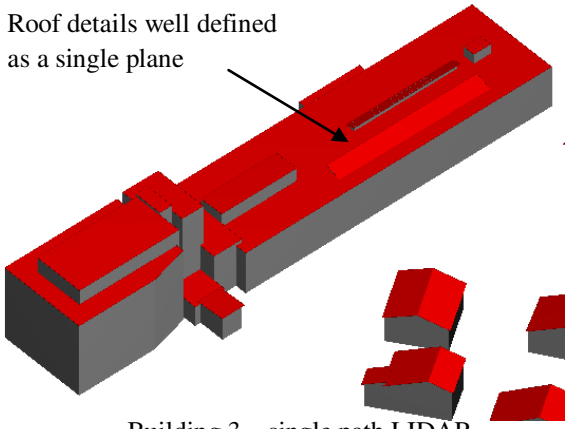
Building 2 – single path LiDAR

Roof details not detected in the single path LiDAR

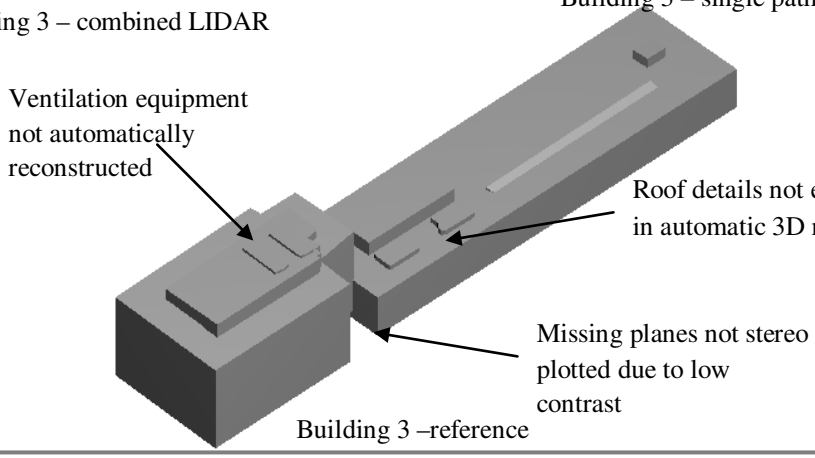
Building 2 – reference



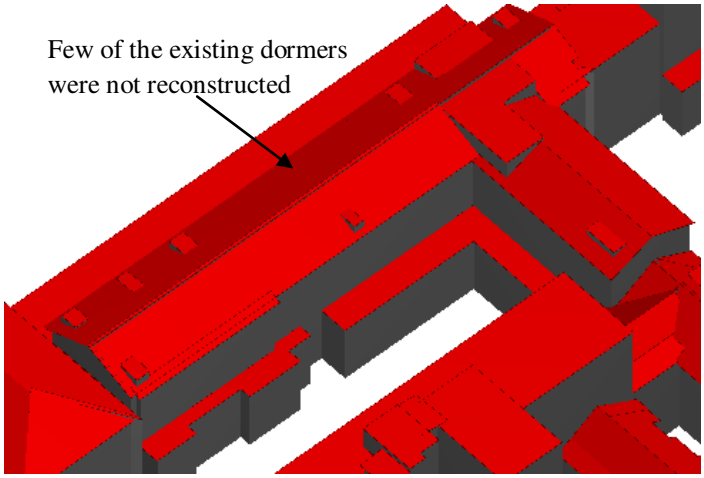
Building 3 – combined LIDAR



Building 3 – single path LIDAR

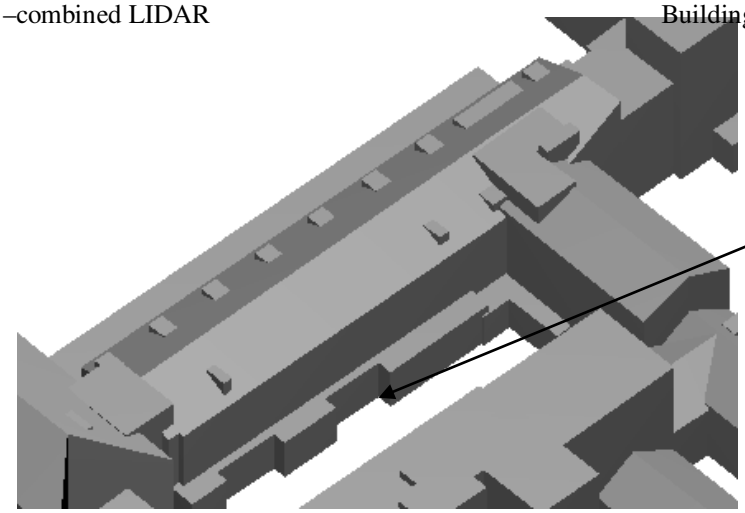


Building 3 –reference



Building 12 –combined LIDAR

Building 12–single path LIDAR



Building 12 –reference

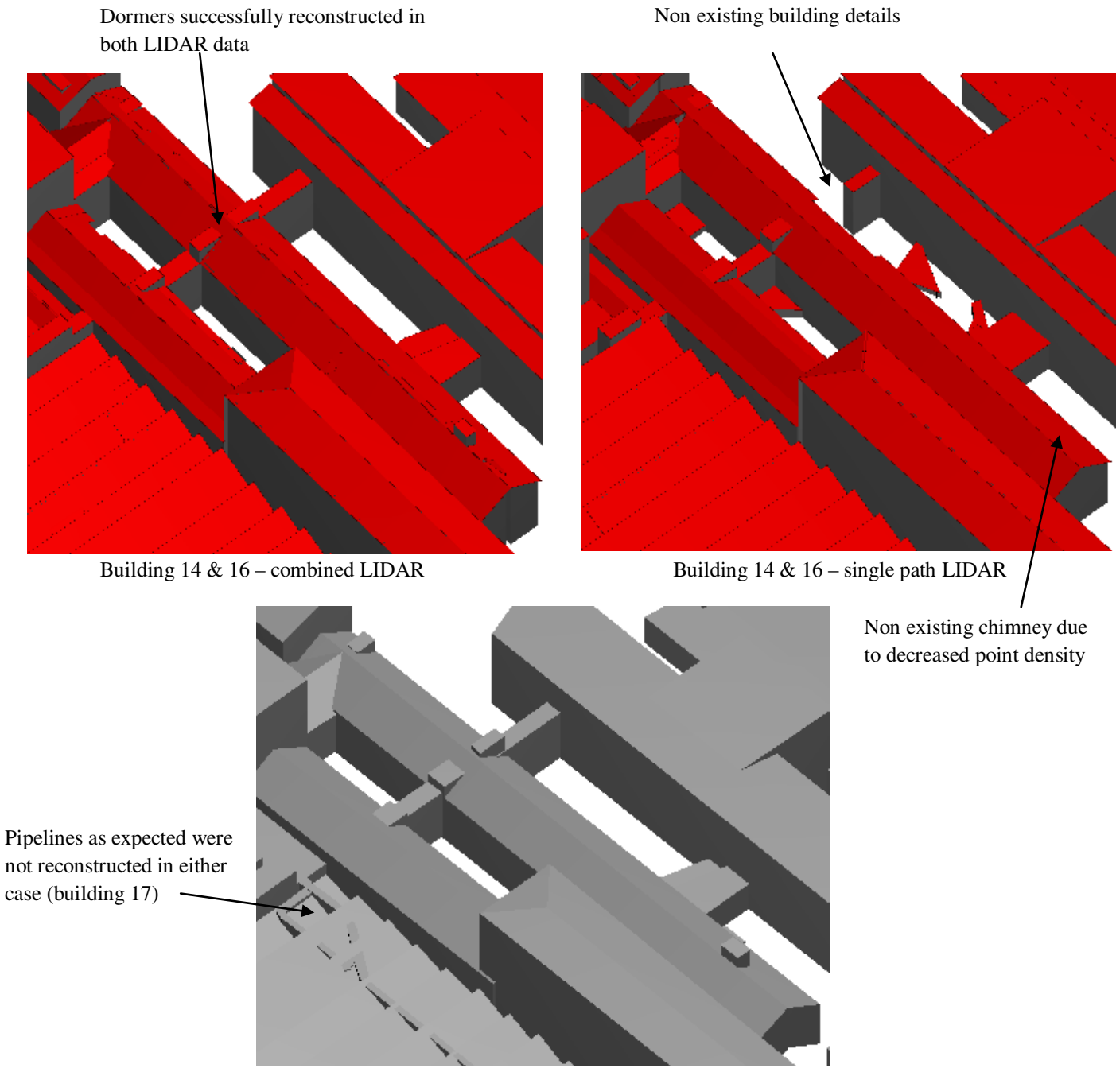
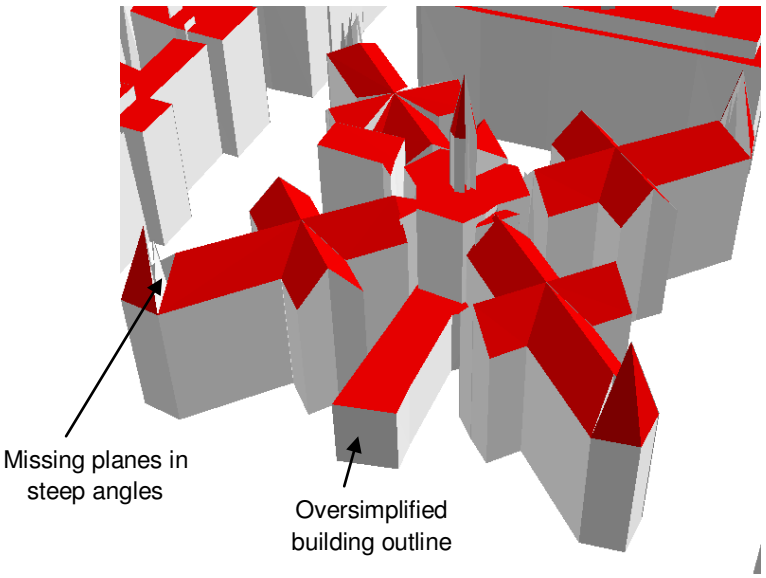
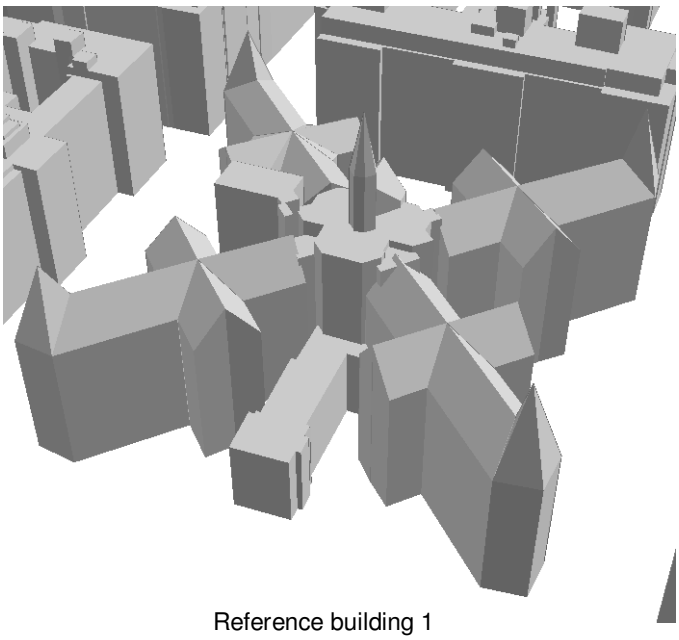


Figure 162 Qualitative comparison between the automatically reconstructed buildings and the reference stereo plotted 3D models for Switzerland

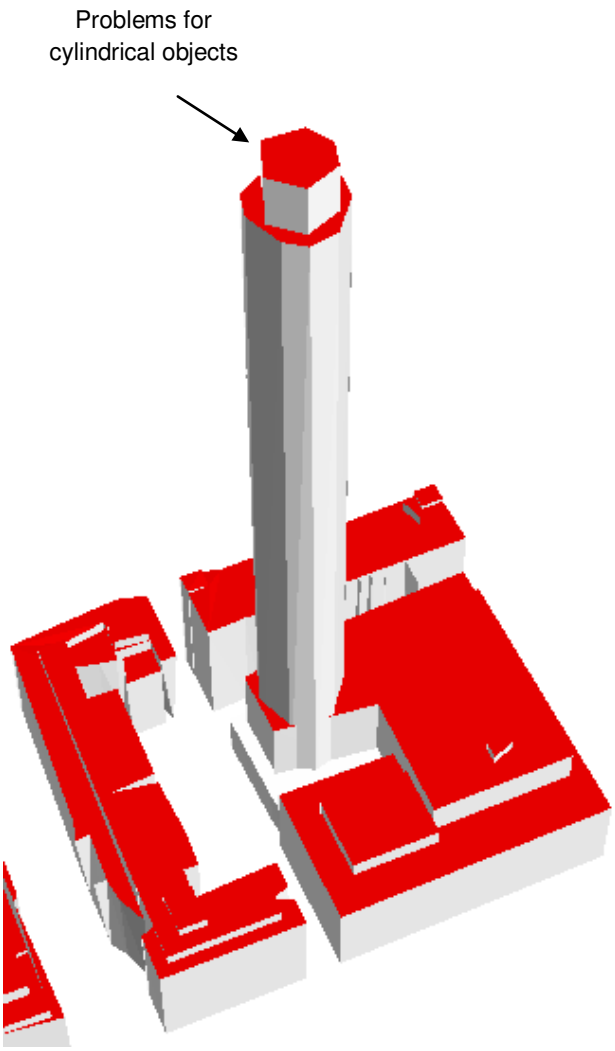
The qualitative comparison for the reconstructed building models in London is given in figure 163.



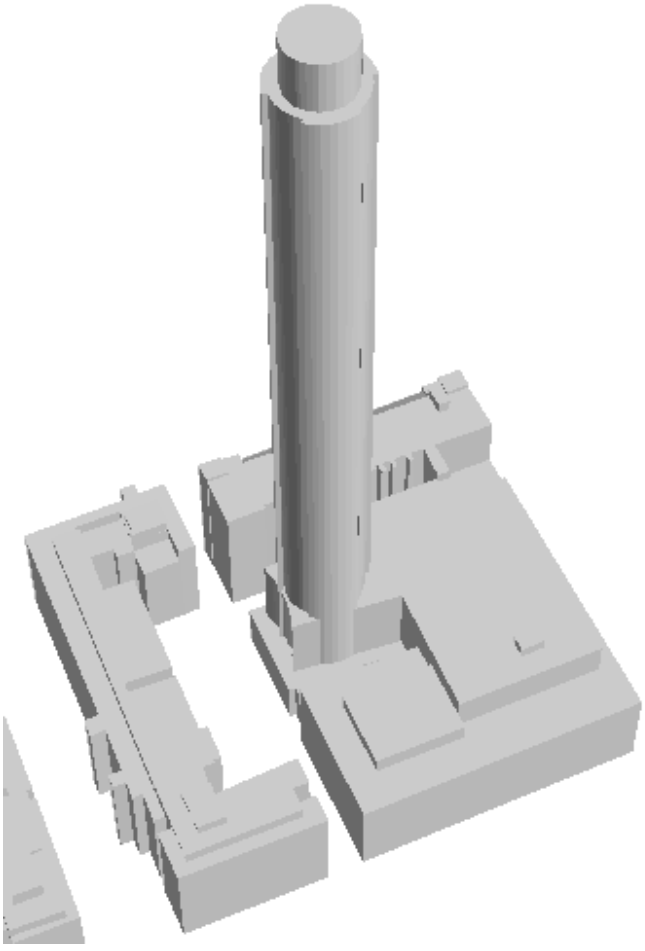
Building 1



Reference building 1



Building 2



Reference building 2

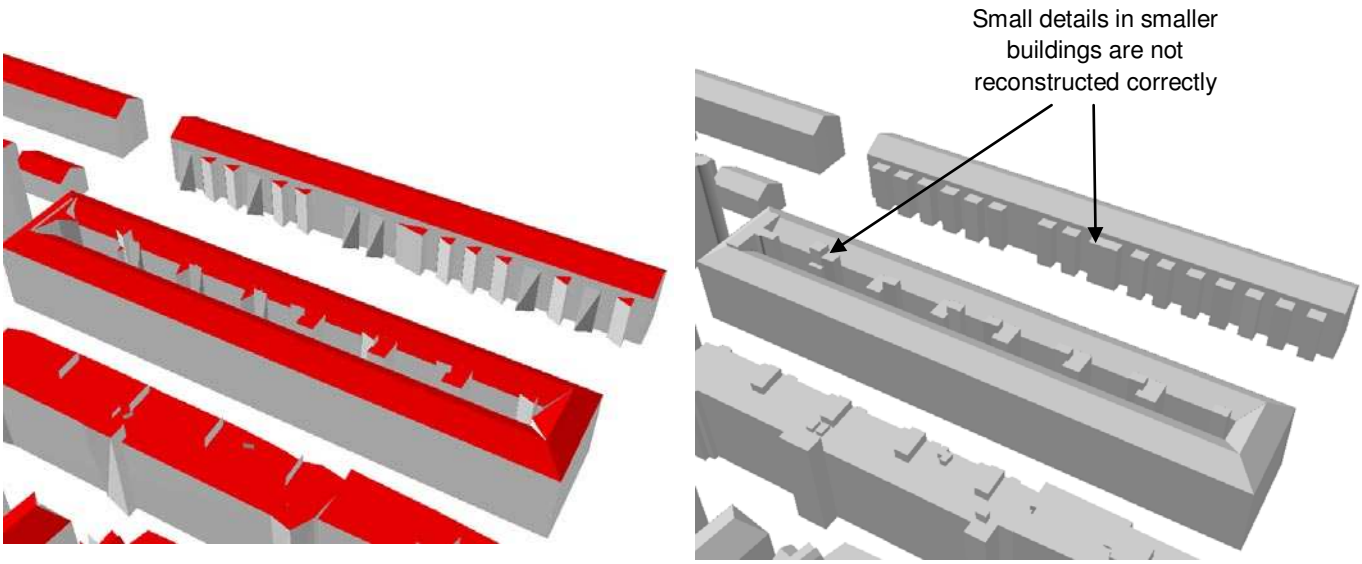
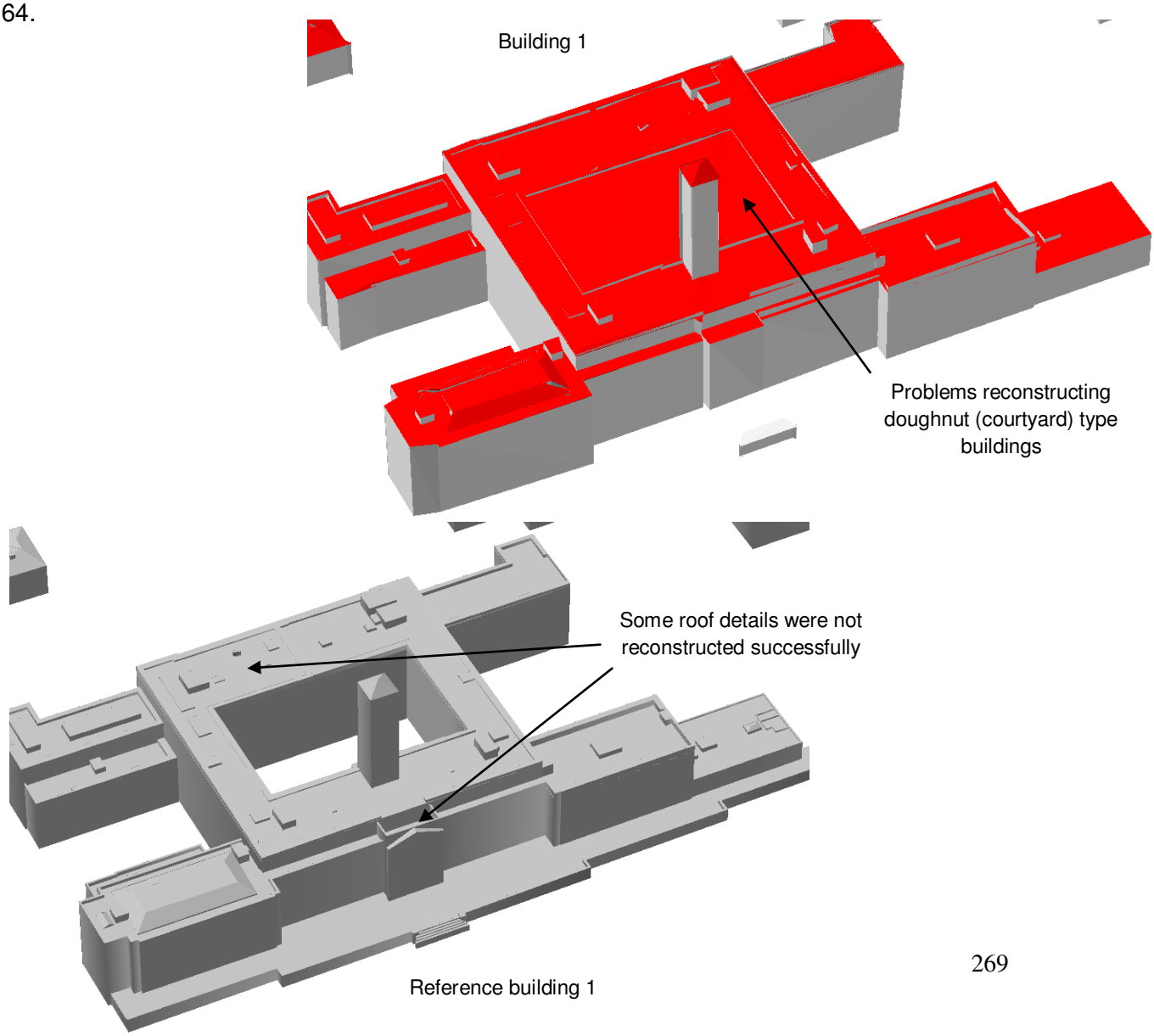


Figure 163 Qualitative comparison between the automatically reconstructed buildings and the reference stereo plotted 3D models for London

The qualitative assessment for the reconstructed buildings in Nottingham is given in figure 164.



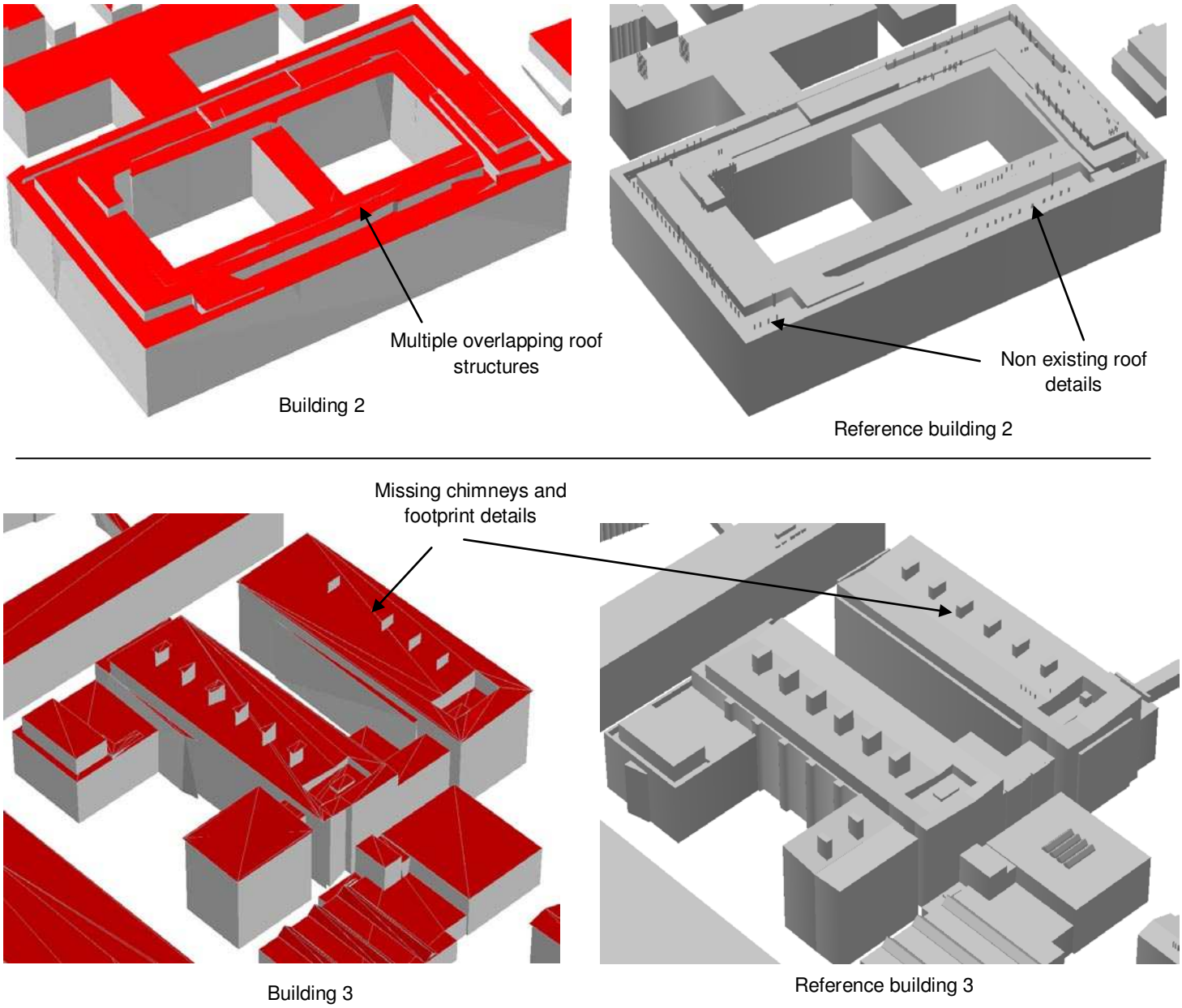


Figure 164 Qualitative comparison between the automatically reconstructed buildings and the reference stereo plotted 3D models for Nottingham

Figures 162, 163 and 164 illustrated only a selection of the reconstructed buildings that showed significant differences between them. The majority of the buildings in the study area presented minor differences that weren't able to be distinguished visually. Nevertheless, the qualitative assessment of the reconstructed buildings depicted before indicates the impressive overall quality of the 3D models.

The level of detail for the 3D models in Switzerland, derived from the combined LiDAR point cloud (density 5 points/m^2), is comparable with the level of detail acquired from the manual stereo plotting. In contrast the 3D models derived from the single path LiDAR data, although they don't present significant differences, in many cases present an incomplete building model since dormers and chimneys are not modelled. At this point special reference to building 3 in the Switzerland study area has to be made. This is one of the few situations where the 3D model from the single path LiDAR data, presented visually a better result. This is evident from the way that the roof details were modelled. The 3D model from the combined LiDAR has multiple superimposed planes that exceed the 25cm merging threshold and therefore the continuous feature is introduced as several individual planes. In contrast, as the merging threshold is increased in the single LiDAR data to 35cm (due to the decreased density), the features are properly modelled as a single plane.

The qualitative building comparison for the London study area indicates several cases where 3D planes of steep angled roof features are not reconstructed properly due to the insufficient quality of the 3D breaklines and the lower point density from the stereo matched points. In general the London study area presents the most artefacts in the semi-automatically reconstructed building models and the overall lower quality and level of detail is evident throughout the entire study area.

In contrast the Nottingham study area is reconstructed to a higher level of detail with a better overall quality since there are fewer artefacts introduced. Nevertheless there are certain cases where the plane fitting algorithm is performing very poorly and this is especially evident in courtyard (doughnut) type buildings.

7.1.2. Semi-automated Quantitative evaluation of the reconstructed building

The quantitative assessment was performed for 17 buildings depicted in figure 161 (Switzerland study area) and the majority of all the other buildings present in London and Nottingham. The semi-automated quantitative evaluation was performed by evaluating the planimetric accuracy of the building footprints and the vertical accuracy of the roof planes. At this point it should be noted that the expected accuracy of the reference models used in this comparison is equivalent to the RMS of the check points calculated from the Aerial Triangulations of each test site.

7.1.2.1. Planimetric accuracy assessment of the building footprint

The planimetric evaluation was performed using the building detection metrics suggested by Shufelt, (1999). The evaluation was conducted using as a reference source the stereo plotted building outlines from the stereo pairs of the three study areas. The comparison was performed twice, once using the optimized building outline and the second time using the initial building outlines. Recall that the initial building outlines were generated from the LiDAR point cloud or VDDSM whilst the optimized footprint was adjusted using the filtered stereo matched points. The accuracy assessment therefore, attempts to investigate the level of improvement in the horizontal accuracy of the adjusted footprints. The building detection metrics adopted for the evaluation can be defined as:

Building Detection Percentage = 100*TP / (TP+FN)

Branching Factor = FP / TP (7.1)

Quality Percentage = 100*TP / (TP + FP + FN)

The branching Factor or building delineation errors are caused when the building outlines are not properly extracted by the automated method. *TP* (True Positive) is a building classified by both datasets, *TN* (True Negative) is a non-building object classified by both datasets, *FP* (False Positive) is a building defined only by the automated building

reconstruction process, and *FN* (False Negative) is a building defined only by the reference stereo plotted buildings. The building detection metrics were obtained with basic GIS functions (overlying, clipping), using the building footprints as polygon entities. Consequently the detection metrics are expressed in square meters and not as number of pixels.

Tables 27, 28, 29 and 30 present the results from the evaluation between the reconstructed building footprints and the reference building outlines for the 1st study area in Switzerland.

Building detection metrics for adjusted building outlines in Switzerland Derived from the combined LiDAR data			
entire study region	308537 m^2	**building detection percentage**	96.6127
true positive	47181.34 m^2		
false positive	4055.17 m^2	**Branching Factor**	0.08595
false negative	1654.21 m^2		
true negative	255646.28 m^2	**Quality Percentage**	89.2053

Table 27 Results from the planimetric evaluation of the adjusted building outlines, derived from the combined LiDAR point cloud in Switzerland

Building detection metrics for initial building outlines in Switzerland derived from the combined LiDAR data			
entire study region	308537 m^2	**building detection percentage**	96.3032
true positive	47035.07 m^2		
false positive	4096.58 m^2	**Branching Factor**	0.0871
false negative	1805.56 m^2		
true negative	255599.79 m^2	**Quality Percentage**	88.8507

Table 28 Results from the planimetric evaluation of the initial building outlines, derived from the combined LiDAR point cloud in Switzerland

Building detection metrics for adjusted building outlines in Switzerland derived from the single path LiDAR data			
entire study region	308537 m^2	**building detection percentage**	95.2698
true positive	46525.03 m^2		
false positive	3017.16 m^2	**Branching Factor**	0.06485
false negative	2310 m^2		
true negative	255646.28 m^2	**Quality Percentage**	89.7263

Table 29 Results from the planimetric evaluation of the adjusted building outlines derived from the single path LIDAR point cloud in Switzerland

Building detection metrics for initial building outlines in Switzerland derived from the single path LIDAR data			
entire study region	308537 m^2	**building detection percentage**	94.9095
true positive	46349.58 m^2		
false positive	3476.05 m^2	**Branching Factor**	0.075
false negative	2485.97 m^2		
true negative	255646.28 m^2	**Quality Percentage**	88.6029

Table 30 Results from the planimetric evaluation of the initial building outlines derived from the single path LIDAR point cloud

Tables 31 and 32 provide the building detection metrics for the adjusted and initial building footprints respectively for the study area of London.

Building detection metrics for adjusted building outlines in London			
entire study region	105189 m^2	**building detection percentage**	85.35
true positive	39538.01 m^2		
false positive	11834.33 m^2	**Branching Factor**	0.30
false negative	6784.10 m^2		
true negative	47032.56 m^2	**Quality Percentage**	67.98

Table 31 results from the planimetric evaluation using building detection metrics for the adjusted building outlines in London

Building detection metrics for the initial building outlines in London			
entire study region	105189 m^2	**building detection percentage**	85.11
true positive	39281.75 m^2		
false positive	12201.93 m^2	**Branching Factor**	0.31
false negative	6871.30 m^2		
true negative	46834.02 m^2	**Quality Percentage**	67.31

Table 32 results from the planimetric evaluation using building detection metrics for the initial building outlines in London

Tables 33 and 34 present the calculated building detection metrics for the adjusted and initial building footprints respectively, for the study area in Nottingham

Building detection metrics for adjusted building outlines in Nottingham			
entire study region	941671 m^2	**building detection percentage**	89.33
true positive	151632.78 m^2		
false positive	14052.43 m^2	**Branching Factor**	0.092
false negative	18093.86 m^2		
true negative	757891.93 m^2	**Quality Percentage**	82.51

Table 33 results from the planimetric evaluation using building detection metrics for the adjusted building outlines in Nottingham

Building detection metrics for the initial building outlines in Nottingham			
entire study region	941671 m^2	**building detection percentage**	88.87
true positive	151255.29 m^2		
false positive	15931.22 m^2	**Branching Factor**	0.105
false negative	18945.45 m^2		
true negative	755539.04 m^2	**Quality Percentage**	81.26

Table 34 results from the planimetric evaluation using building detection metrics for the initial building outlines in Nottingham

The accuracy assessment depicted in the previous tables presents promising results as the building detection and quality percentage are significantly high in all the study areas with the worst results obtained for the detected buildings in London. The building detection percentage was increased for the adjusted building outlines as the area of the buildings defined only by the reference stereo plotted buildings (false negative) was decreased in most study areas. It's worth noting that in most cases the building areas, defined only by the

automated procedure (false positive) are very similar with the exception of the Nottingham study area where the false positive between the initial and adjusted building outlines presents more significant differences.

The Building Delineation errors (branching factor) are caused when the building outlines are not properly extracted by the automated method. These errors are estimated with respect to the size of the building footprint, defined by both datasets (true positive). These errors are related to the inherent planimetric accuracy of the input data and subsequently with the point density of the LiDAR point cloud and the overall quality of the stereo matched points. The building delineation errors for the Switzerland study area in both building footprints were minimized to approximately 0.08 indicating that the high density LiDAR point clouds are usually adequate in extracting a good approximation of the building footprint without the presence of significant linear artefacts. The same case is presented for the building footprints of the Nottingham study area indicating that the VDDSM created from high quality stereo matched points present satisfactory results of the building footprint. In contrast the London study area presents the worst branching factor (0.30) indicating the presence of significant artefacts for the footprint geometric shape due to the low density of the initial LiDAR point cloud and the overall degraded quality of the stereo matched points.

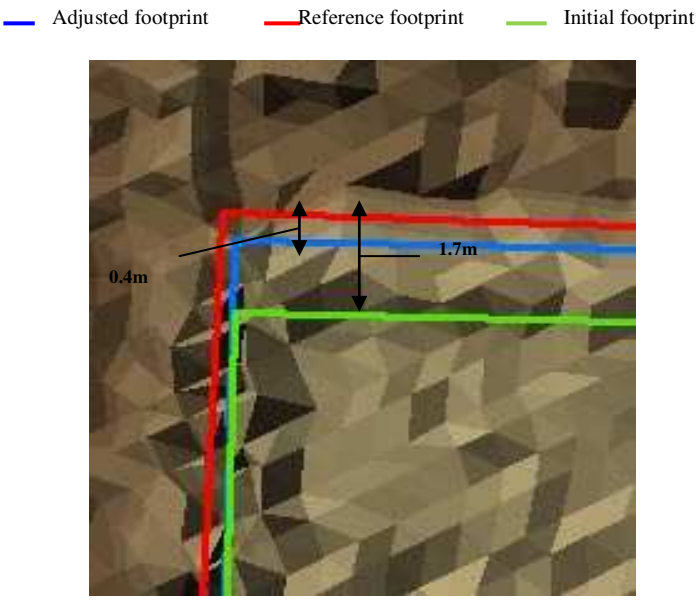


Figure 165 Example of improved planimetric accuracy at the adjusted building footprint.

The overall quality of the extracted of the building outlines is given by the quality percentage, which in all cases is indicating a good overall quality (greater than 65%). Nevertheless, the difference in the quality between the adjusted building outline and the initial footprints is not

as large as was expected. Although, there is an improvement in the statistical parameters for the adjusted footprint, it is evident that the difference is relatively small. The main reasons for this situation are the procedure of simplifying the building outline and the high density of the LiDAR or the VDDSM. The high density of the combined LiDAR point cloud in Switzerland is improving the accuracy of the extracted initial building footprints and therefore, small horizontal corrections are applied in the adjusted footprints. A similar situation is also observed for the second and third study areas with the optimised building outlines presenting better results by small margins compared to the initial building outlines. As a result it is assumed that the implementation of the simplification algorithm for generalizing the adjusted building outlines may have a negative impact since small building areas might be oversimplified.

At this point it's interesting to note that the planimetric accuracy from the single path LiDAR point cloud is very similar with that obtained from the combined LiDAR data in Switzerland. Although the density of the LiDAR data seems to have a significant negative impact on the level of reconstructed roof details, the building outlines on the contrary are accurately extracted in both cases. Switzerland produced the best results during the building detection stage with high overall quality (89%), very small branching factor and high building detection percentage (95%). These impressive results for the Switzerland area where obtained due to the high quality of density of the LiDAR point cloud, coupled with the effectiveness of the proposed workflow for LiDAR classification and building detection. The building detection for the Nottingham study area was also very impressive, with a lower building detection percentage due to the presence of misclassified objects of the building class to the tree class, but with very small branching factor indicating that the shape of the objects classified correctly as buildings was representing the correct shape of the footprints. The impressive results for the Nottingham study area is evidence of the effectiveness of the high quality airborne imagery that produced a stereo matched point cloud of superior quality. In contrast the results for the London area, although deemed acceptable, were far worse compared to

the other two study regions. The relatively poor results obtained for London were a function of the low quality of the initial LiDAR, poor quality of the airborne imagery that produced a low density stereo matched point cloud and complex building geometry. It is concluded that further research is needed for alternative workflows when it comes to building detection from low quality datasets to enhance the final result of the building outlines.

7.1.2.2. Semi-automated vertical accuracy assessment of reconstructed building models

The proposed method for a robust estimation of the vertical accuracy includes the generation of raster Digital Surface Models for each building. With the use of DSMs, functions such as overlaying and subtraction with reference DSMs can provide a robust estimation of the vertical accuracy using statistical parameters. The procedure of creating the DSMs from the 3D building models consists of two basic steps. Initially each separate roof plane of the building model is treated independently to create a TIN plane. Each TIN plane can accurately represent each roof plane, since the planes are treated independently. The next step includes the conversion of the TIN features into a raster representation. The raster representation has a small spacing to handle even the small boundary details, while the elevation information is stored as a floating point format to preserve the precision of the interpolated elevation values. The TIN to raster conversion is performed for each roof plane independently.

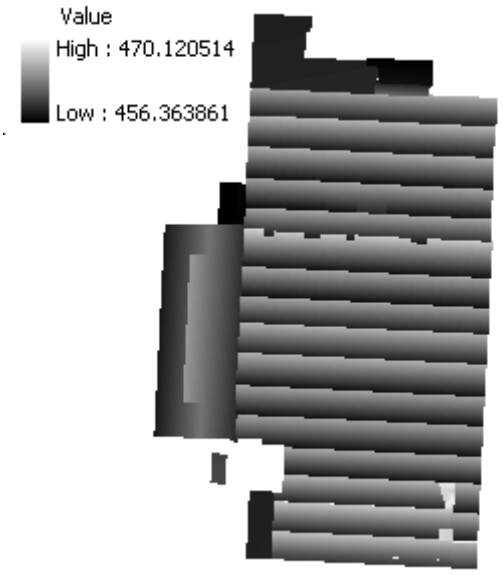


Figure 166 Example of a raster DSM created from the individual roof planes

Finally the raster planes (with spacing of 5cm), constituting a single building are merged together using Map Algebra functions. The procedure was conducted in a GIS environment with Visual Basic scripts, designed to automate the process since the buildings used in the evaluation of the three study areas had over 2000 individual roof planes. Figure 166 depicts a created DSM for one of the reconstructed buildings in Switzerland.

The procedure is repeated for the reference 3D building models and the comparison is performed by subtracting the two DSMs. The result is a difference image indicating the height differences between automatically derived roof planes and reference roof planes. The difference image is estimated only where the buildings overlap and therefore the calculation of the standard deviation does not take into account differences in the footprint of the buildings. The results from the vertical evaluation in Switzerland are summarized in table 35.

Vertical accuracy parameters for study area in Switzerland

(3D building models from the combined LiDAR data compared to the reference models)

Building number	Total pixels	Min. error (m)	Max. error (m)	Mean value(m)	Standard deviation (m)
Building 1	1171303	-0.65	1.1	0.34	0.35
Building 2	554643	-0.39	0.97	0.23	0.36
Building 3	810867	-1.27	1.56	0.21	0.53
Building 4	1549708	-0.64	2.51	0.61	0.56
Building 5	717449	-1.42	0.36	-0.59	0.48
Building 6	2317549	-0.56	0.90	-0.26	0.30
Building 7	1331752	-0.99	0.36	-0.49	0.22
Building 8	807336	-0.23	0.09	-0.08	0.07
Building 9	1120424	-0.09	0.54	0.16	0.13
Building 10	1958496	-2.45	3.63	-0.47	1.09
Building 11	312927	-0.86	0.60	0.12	0.22
Building 12	780007	-2.09	0.62	-0.84	0.39
Building 13	514909	-1.78	0.51	-0.51	0.46
Building 14	460316	-1.10	0.91	0.06	0.31
Building 15	394981	-2.23	1.23	-0.75	0.78

Building 16	605751	-1.20	1.31	0.19	0.58
Building 17	1974890	-2.94	1.23	-0.37	0.79
			Average values	-0.14m	0.45m

Table 35 Vertical accuracy characteristics of the semi-automatic building models derived from the combined LiDAR data in Switzerland.

The results from the evaluation of the automatic building models, derived from the combined LiDAR point cloud, indicate a small overall vertical shift of -14cm. The average standard deviation of the 17 buildings is in the order of 45cm, indicating the presence of significant vertical errors. This is evident also from the maximum and minimum vertical deviation, with building 10 introducing the largest vertical differences in the study area.

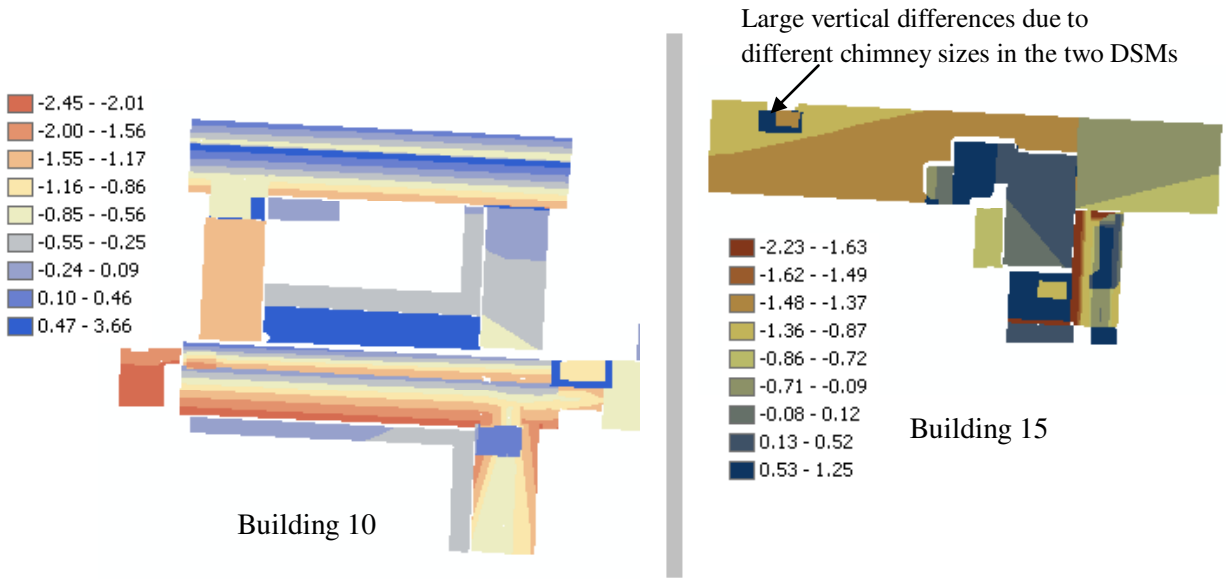


Figure 167 Difference images for buildings 10 & 15, indicating the vertical errors of the two models in Switzerland

The majority of the large vertical errors were introduced as a function of the planimetric deviation between the roof planes. An example is the vertical errors introduced when planes representing a chimney in the two DSMs have different sizes. As a result, the planimetric deviations between the DSMs have a significant impact on the estimated standard deviation

(figure 167). The second evaluation was a comparison between the reconstructed building models in Switzerland, derived from the single path LiDAR data and the reference building models. The results are summarized in table 36.

Vertical accuracy parameters for study area in Switzerland (3D building models from the single path LIDAR data compared with the reference models)					
Building number	Total pixels	Min. error (m)	Max. error (m)	Mean value(m)	Standard deviation (m)
Building 1	1175297	-0.61	1.25	0.37	0.36
Building 2	557484	-1.05	1.45	0.27	0.38
Building 3	839059	-2.38	2.70	0.19	0.74
Building 4	1520296	-1.09	2.24	0.56	0.53
Building 5	709248	-1.34	0.18	-0.57	0.48
Building 6	2090061	-0.70	0.10	-0.31	0.05
Building 7	1311480	-0.96	0.05	-0.47	0.21
Building 8	800887	-0.18	0.12	-0.03	0.06
Building 9	1136059	-0.23	0.52	0.18	0.16
Building 10	1902309	-4.09	3.52	-0.46	1.06
Building 11	297410	-0.84	3.90	0.19	0.45
Building 12	779186	-3.99	1.89	-0.88	0.47
Building 13	511057	-1.78	1.52	-0.47	0.50
Building 14	458164	-2.28	1.30	0.04	0.44
Building 15	392560	-3.79	2.78	-0.76	0.79
Building 16	603290	-1.67	1.93	0.22	0.62
Building 17	1935371	-3.29	1.50	-0.47	0.94
			Average values	-0.141m	0.48m

Table 36 Vertical accuracy characteristics of the semi-automatic building models derived from the single path LiDAR data in Switzerland.

The accuracy evaluation of the reconstructed buildings from the single path LiDAR point cloud in Switzerland indicates very similar results with the first evaluation. The overall vertical shift is in the same level as before and surprisingly the standard deviation is

introducing a small reduction. It is evident that the reduced density of the point cloud (approximately half the density of the combined LiDAR data) does not have a significant influence in the vertical accuracy. Table 37 indicates the results from the vertical evaluation of the reconstructed models in London.

Vertical accuracy parameters for study area in London (3D building models compared to the reference models)					
Building number	Total pixels	Min. error (m)	Max. error (m)	Mean value(m)	Standard deviation (m)
Building 1	134528	-0.69	1.05	0.23	0.43
Building 2	309029	-0.75	2.46	-0.11	0.22
Building 3	598132	-1.38	0.78	0.32	0.12
Building 4	733498	-1.94	1.98	0.21	0.87
Building 5	891109	-0.39	0.72	-0.34	1.12
Building 6	1090745	-0.95	0.23	-0.42	1.42
Building 7	1609639	-0.88	0.94	0.21	0.74
Building 8	529076	-2.25	0.82	0.29	0.48
Building 9	289407	-0.42	1.88	0.17	0.38
Building 10	310752	-3.45	2.05	0.09	0.62
Building 11	529104	-0.63	3.22	-0.10	1.01
Building 12	1105982	-2.63	2.07	-0.28	0.29
Building 13	407276	-1.16	2.11	-0.31	0.15
Building 14	728710	-1.94	1.03	0.02	0.41
Building 15	953072	-2.01	1.53	-0.20	0.69
Building 16	947010	-0.38	1.21	-0.31	1.05
Building 17	1682913	-0.92	0.83	0.16	1.38
Building 18	1590387	-0.74	2.83	0.19	0.44
Building 19	429690	-1.49	2.59	0.22	0.73
Building 20	1210078	-2.03	2.02	0.29	0.28
Building 21	1230784	-2.44	3.01	-0.31	0.31
Building 22	735960	-2.31	3.28	0.09	0.71
Building 23	100638	-1.55	1.05	-0.41	0.63

Building 24	297057	-1.23	1.59	0.27	0.42
Building 25	410836	-1.02	2.01	0.12	0.19
Building 26	372289	-0.39	1.99	0.36	0.72
Building 27	191067	-0.82	1.42	-0.21	0.32
Building 28	1107949	-0.92	2.03	-0.17	0.98
Building 29	1302989	-1.37	1.80	0.39	1.42
Building 30	1220989	-4.86	1.25	0.03	0.48
Building 31	1568924	-3.72	0.34	-0.08	1.38
Building 32	583807	-2.85	0.39	0.10	1.01
			Average values	0.02m	0.67m

Table 37 Vertical accuracy characteristics of the semi-automatic building models derived from the study area in London

From table 37 it is evident that in contrast to the Switzerland study area in London there is a very small vertical shift in the reconstructed building models. This indicates that there were very few systematic errors in the initial LiDAR data and GCPs used in the aerial triangulation of the aerial photographs. In contrast the standard deviation of the reconstructed building roof planes is higher compared to Switzerland that is caused by the differences in the roof plane shape and the absence of the reconstructed roof details in comparison to the reference buildings. Figure 168 depicts some examples from the vertical accuracy assessments.

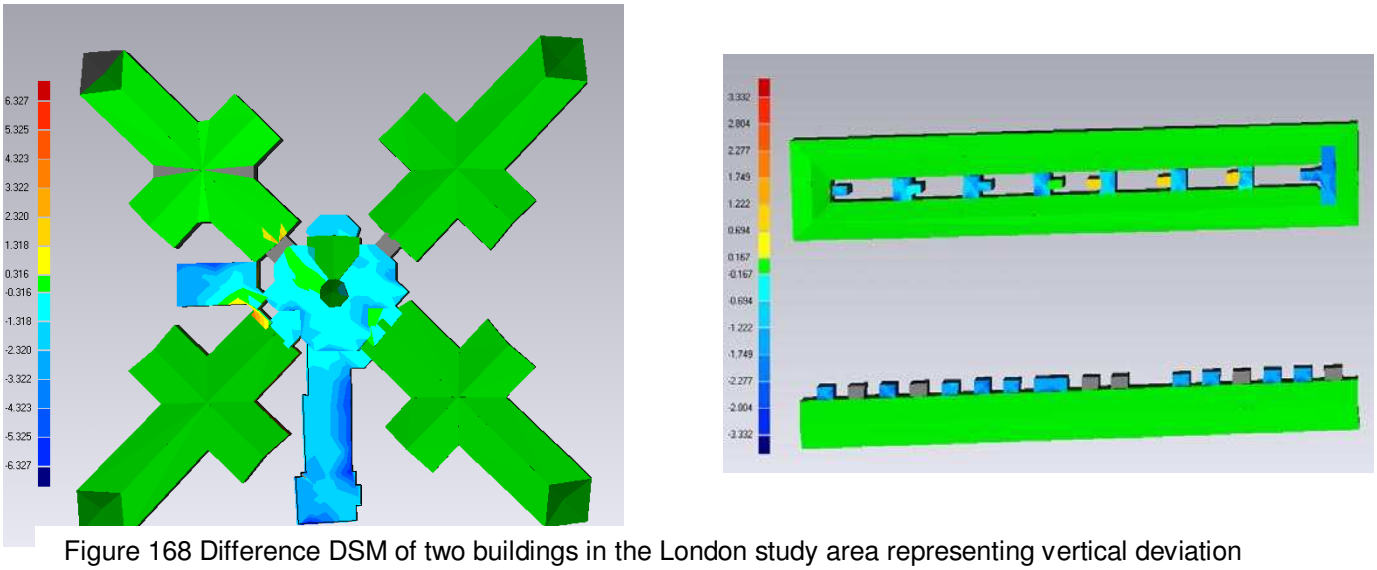


Figure 168 Difference DSM of two buildings in the London study area representing vertical deviation between the semi-automatic and reference roof planes

The following table 38 provides the summary statistics from the vertical accuracy assessment performed on the buildings in the test site of Nottingham.

Vertical accuracy parameters for study area in Nottingham (3D building models compared to the reference models)					
Building number	Total pixels	Min. error (m)	Max. error (m)	Mean value(m)	Standard deviation (m)
Building 1	736821	-1.73	2.05	-0.03	0.29
Building 2	308919	-1.24	1.49	0.04	0.17
Building 3	142368	-0.62	0.94	0.21	0.11
Building 4	857932	-0.83	0.39	0.34	0.03
Building 5	650233	-0.92	0.78	-0.23	0.08
Building 6	1123293	-0.04	1.94	-0.12	0.59
Building 7	1274329	-0.73	1.02	-0.31	0.39
Building 8	1345702	-2.64	1.75	0.23	1.05
Building 9	436875	-2.29	2.08	0.28	0.20
Building 10	328764	-1.39	1.75	0.41	0.05
Building 11	589692	-2.42	2.37	-0.23	0.10
Building 12	445703	-1.02	2.54	0.21	0.31
Building 13	879324	-0.03	2.19	0.28	0.11
Building 14	438593	-0.10	1.93	-0.16	0.02
Building 15	632489	-0.38	0.49	-0.18	0.28
Building 16	1232892	-0.92	0.71	-0.37	0.71
Building 17	1263189	-0.51	0.60	0.15	0.06
Building 18	326472	-0.29	0.31	0.22	0.19
Building 19	215381	-0.42	0.95	0.25	0.30
Building 20	879234	-0.71	0.89	-0.10	0.09
Building 21	675237	-0.62	1.49	-0.20	0.14
Building 22	237498	-0.79	1.04	0.14	0.15
Building 23	329478	-0.12	2.94	0.36	0.20
Building 24	874932	-0.28	2.39	-0.28	0.18

Building 25	675371	-0.31	2.16	0.64	0.05
Building 26	1223987	-0.49	1.68	-0.44	0.10
Building 27	328746	-0.10	0.84	0.11	0.14
Building 28	873932	-1.33	0.47	-0.31	0.27
Building 29	762824	-1.59	0.55	0.29	0.22
Building 30	643287	-2.03	1.39	0.13	0.19
Building 31	234989	-2.79	1.84	0.21	0.68
Building 32	564483	-3.62	2.04	-0.26	0.20
Building 33	218937	-0.52	1.92	-0.41	0.38
Building 34	986434	-1.84	0.39	0.25	0.22
Building 35	891274	-0.38	0.10	0.18	0.32
Building 36	274832	-2.85	1.28	0.24	0.09
Building 37	673243	-2.69	2.68	-0.21	0.13
Building 38	383429	-2.24	1.83	0.13	0.08
Building 39	782163	-3.81	2.38	-0.31	0.60
Building 40	487367	-0.39	2.59	-0.11	1.03
Building 41	541736	-0.29	1.38	0.36	0.21
Building 42	785374	-0.84	1.82	0.24	0.17
Building 43	216375	-0.92	0.49	0.28	0.23
Building 44	347826	-1.04	0.29	-0.23	0.12
Building 45	897923	-1.25	0.38	-0.34	0.27
Building 46	973864	-1.92	0.77	-0.25	0.07
Building 47	216347	-0.29	1.39	0.36	0.30
Building 48	1327864	-0.62	2.95	0.17	0.24
Building 49	1230934	-0.49	2.58	0.19	0.15
Building 50	372842	-1.92	2.93	-0.26	0.18
Building 51	912738	-2.71	0.93	0.49	0.05
			Average values	0.04m	0.25m

Table 38 Vertical accuracy characteristics of the semi-automatic building models derived from the study area in Nottingham

Table 38 presents promising results from the vertical accuracy assessment from the Nottingham study area. The small mean errors indicate the absence of significant systematic

vertical errors on the reconstructed roof planes. The standard deviation is significantly lower compared to the other two test sites and it is comparable to the vertical accuracy statistics obtained for the aerial triangulation of the aerial imagery. It is understandable that the vertical accuracy of the stereo matched points and the VDDSM created has an equivalent vertical accuracy. Thus it is safe to assume that for the Nottingham test site the plane fitting algorithm hasn't introduced any significant artefacts to affect the vertical accuracy. Figure 169 depicts some examples from the vertical accuracy assessment performed on the reconstructed buildings in Nottingham.

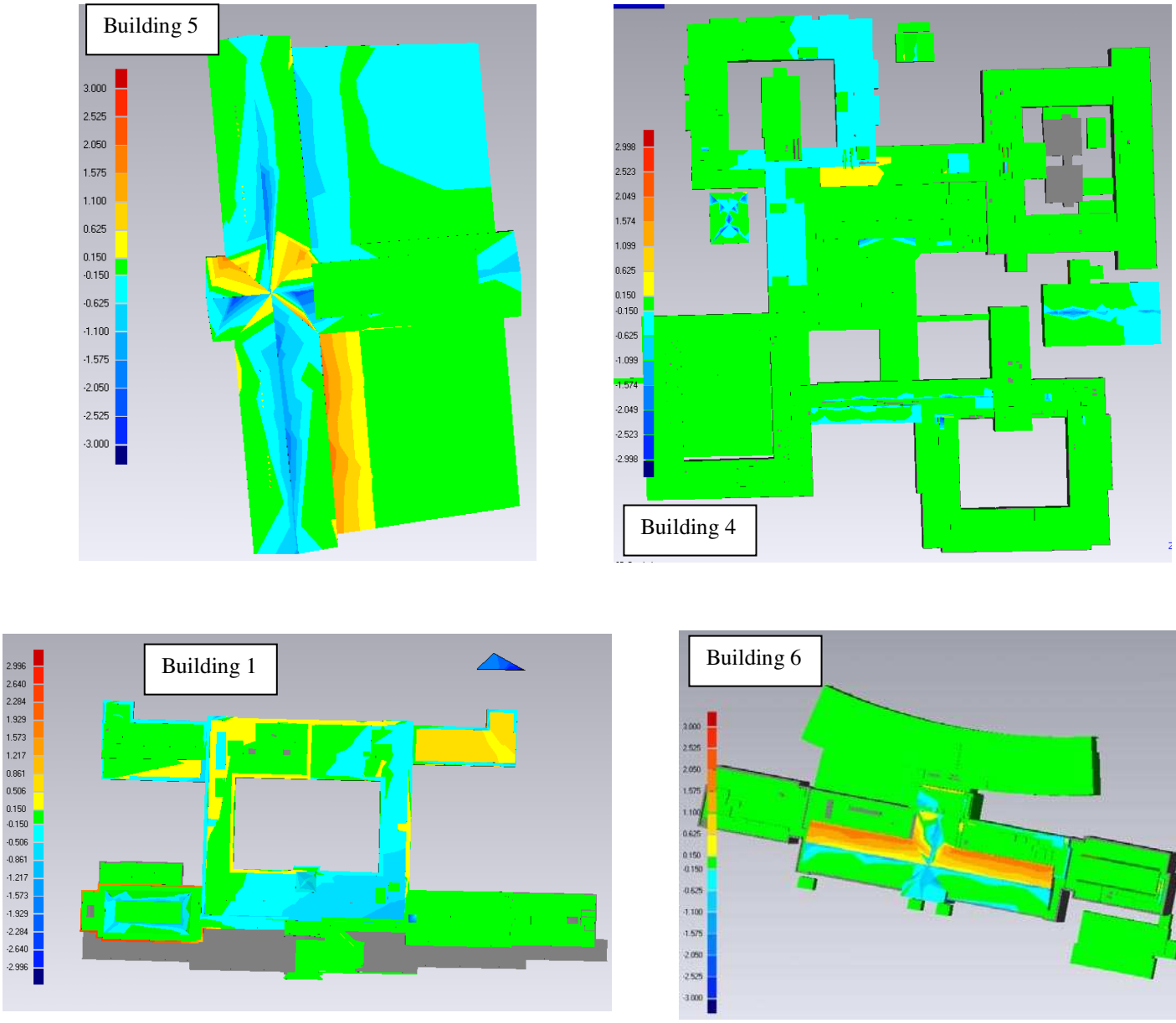


Figure 169 Difference DSM of four buildings in the Nottingham study area representing vertical deviation between the semi-automatic and reference roof planes

At this point it should be noted that the evaluation for the Switzerland study area was conducted for relatively large buildings and the influence of the decreased point density was evident only in the ability to reconstruct small roof details. Therefore, it is apparent from the evaluation that small roof structures, apart from the visual differences, don't seem to greatly affect the vertical accuracy of the building models. As a general conclusion from the evaluation it could be said that the LiDAR dataset with a density of 3 points/m^2 yield reconstructed models with very similar planimetric and vertical accuracy results, compared to the 3D building models derived from the combined LiDAR data (density of 5 points/m^2).

7.2. Automatic texturing using vertical airborne images

Automatic texture mapping of building rooftops and facades is of particular significance for an aesthetically pleasant representation of reconstructed 3D city models. There have been various techniques and data sources used to texture city models. Previous methods have tried to tackle the issue into texturing building rooftops using vertical aerial photographs, and texturing facades by utilizing ground-based images (Brenner et al. 2001).

In these cases, roofs and terrain are often texture mapped with only one single vertical aerial image while building facades not accessible from terrestrial photography are not textured (Lee et al. 2002). However, to ensure that all building facades in a city area are covered and properly textured, it is necessary to utilize multiple oblique and vertical aerial images with different viewing angles. In this case, it is necessary to cope with most parts of a model being visible in several images, under varied viewing angles, resolutions, and potentially different lighting conditions.

Despite the necessity for multiple viewing angles, oblique airborne images are not always easily accessible for all the study areas. This project presents interesting findings from the implementation of an automatic texture mapping workflow in the three study areas. Note that the first two study areas in Switzerland and London consist of vertical airborne imagery

(Switzerland test site had oblique imagery from ADS40 but only one band and thus true-color texture mapping was not properly implemented). The Nottingham test campus utilises the vertical UltraCamD airborne photography.

The texture mapping workflow described in this section was implemented in Leica Photogrammetry Suite once the final building models were generated and the aerial triangulations of the available aerial images were completed.

The first stage in the automatic texture mapping workflow is the registration of the 3D models with the aerial images. In a photogrammetric system this is easily achieved using the colinearity equations and solving for the image x,y of each polygon node. In other words each polygon is back-projected on the images using the linear form of the collinearity equations (see equation 5.1). Since the exterior and interior orientation parameters have been accurately determined from the aerial triangulations the registration between the 3D models and the airborne images is usually to a high degree of accuracy. Hence all the 3D polygonal facets of the 3D city models are projected on all the available aerial images. When the 3D polygon is projected from the object space to the image space it will either return a set of x,y image co-ordinates or will return null values in case the polygons falls outside photographs image extent.

The second stage of the workflow deals with the selection of the optimum image for texturing each polygon. This stage follows the method proposed by (Früh et al., 2004). Note that from the above process a single polygon might be projected onto multiple aerial photographs. The assumption made at this point by the workflow is that that all the images were taken under similar lighting conditions, thus difference in illumination conditions are not taken into account. This might have an adverse effect on the overall texture quality but the complexity of the workflow otherwise would have been beyond the scope of this project.

The selection of the optimal aerial image for texturing a specific 3D polygon is based on the following criteria as defined in (Früh et al., 2004).

- Resolution

Resolution is defined as the number of pixels within the projected 3D polygon at the image space. Since there are multiple images sometimes with different viewing angle the number of corresponding pixels within a projected polygon ranges. To determine the resolution R_{ij} for each image I_i and triangle T_j the workflow divides the number of pixels within each projected 2D triangle by the area of the corresponding 3D triangle.

- Occlusion

Occlusion η_{ij} is determined as a percentage of a 3D triangle T_j that is visible on an image I_i. The workflow detects the occlusions on a pixel by pixel basis by implementing a z-buffer algorithm. The z-buffer algorithm is an image space visibility determination workflow, it may also be referred to as hidden surface removal process.

- Viewing angle

The desirable aerial image for texturing a 3D polygon should have approximately the same orientation as the 3D polygon, in other words taken from a direct, perpendicular viewing position. Otherwise large distortions could be introduced. For each image I_i and triangle T_j the viewing direction $\vec{v}_{ij}$ as the vector between the image principal point and the centre point of the 3D polygon. The workflow then uses the scalar product between the viewing direction and the surface normal $\vec{n}_{ij}$ of the 3D polygon to quantify how direct the view of an aerial image is.

The selection of the most appropriate image is then given from the highest score obtained from the following equation (Früh et al., 2004).

$$\lambda_{ij} = R_{ij} \times \eta_{ij} \times \vec{v}_{ij} \times \vec{n}_{ij} \qquad (7.2)$$

Leica Geosystems has implemented the workflow described previously in Leica photogrammetry Suite and it was utilized for the automatic texture mapping of the 3D city models.

7.2.1. Automatic texture mapping and visualization for study area in Switzerland

This section describes the process of texture mapping, applied on the reconstructed building models in Switzerland for visualization purposes. Initially an ortho-rectified true colour composite was generated using the available bands of the stereo model and a DTM of the study area. The DTM was created from the filtered combined LiDAR point cloud.

Figure 170 Perspective view of the textured building models for the study area superimposed on the ortho-rectified true color composite and draped over a DTM

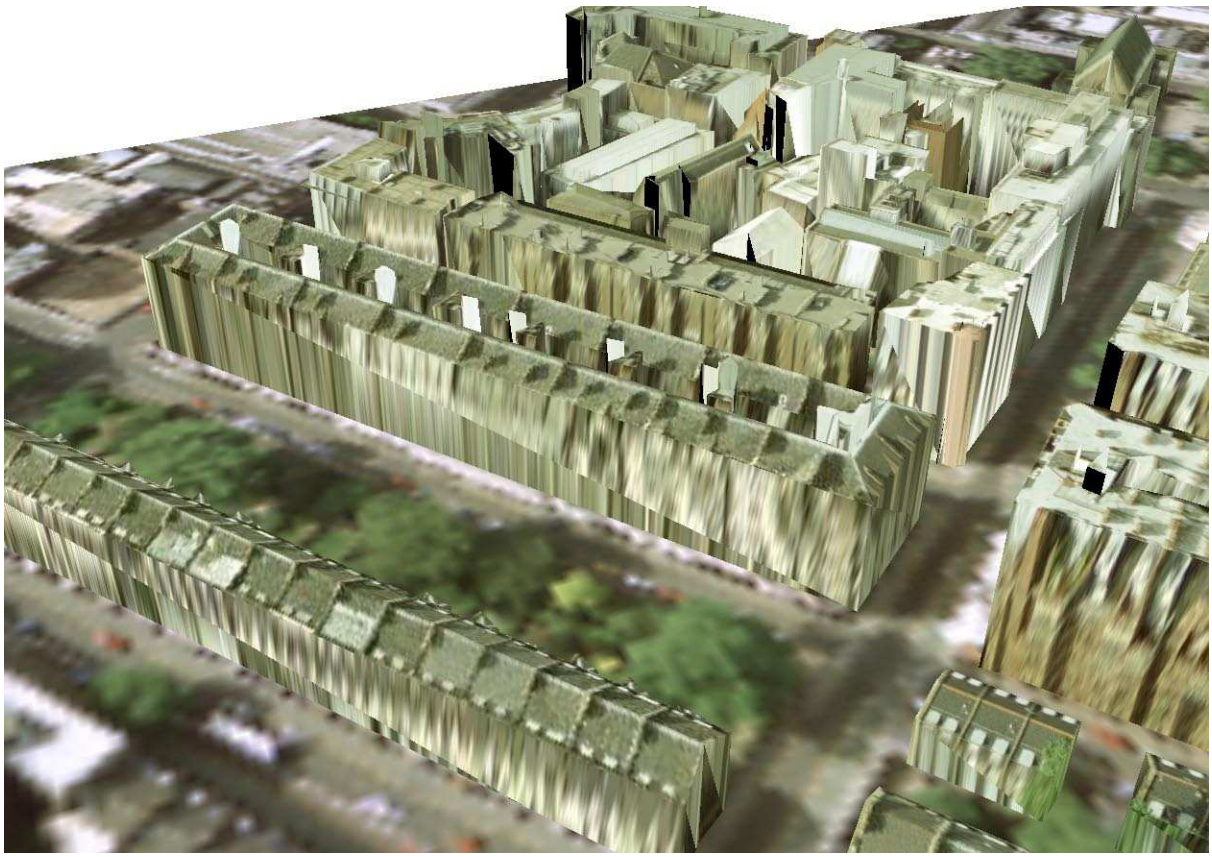
The true colour composite was subsequently used in conjunction with the reconstructed building models to obtain texture information from the image. The texture information was obtained by back-projecting only the roof planes on the true colour composite (image space) by implementing the workflow described in the section 7.2 with the ADS40 true colour composite.

Figure 171 Detailed perspective view of the textured building models, superimposed on the orthorectified true color composite which was draped over a DSM

The generic images, representing vertical building facades were used from a material library for texturing the reconstructed facades. The procedure described in this section was performed using Leica Photogrammetry Suite and the Virtual GIS module in ERDAS Imagine. Visualizations of the reconstructed buildings are depicted in figures 170 & 171.

7.2.2. Automatic texture mapping and visualization for the study area in London

The automatic texture mapping performed on the reconstructed building models in London was implemented by utilizing only the available stereo pair of aerial photographs in the area. The virtual scene consists of a DTM, created from the filtered LiDAR data, and an ortho-rectified aerial photograph which is draped on the DTM. Figure 172 illustrates some examples of the textured mapped building models.



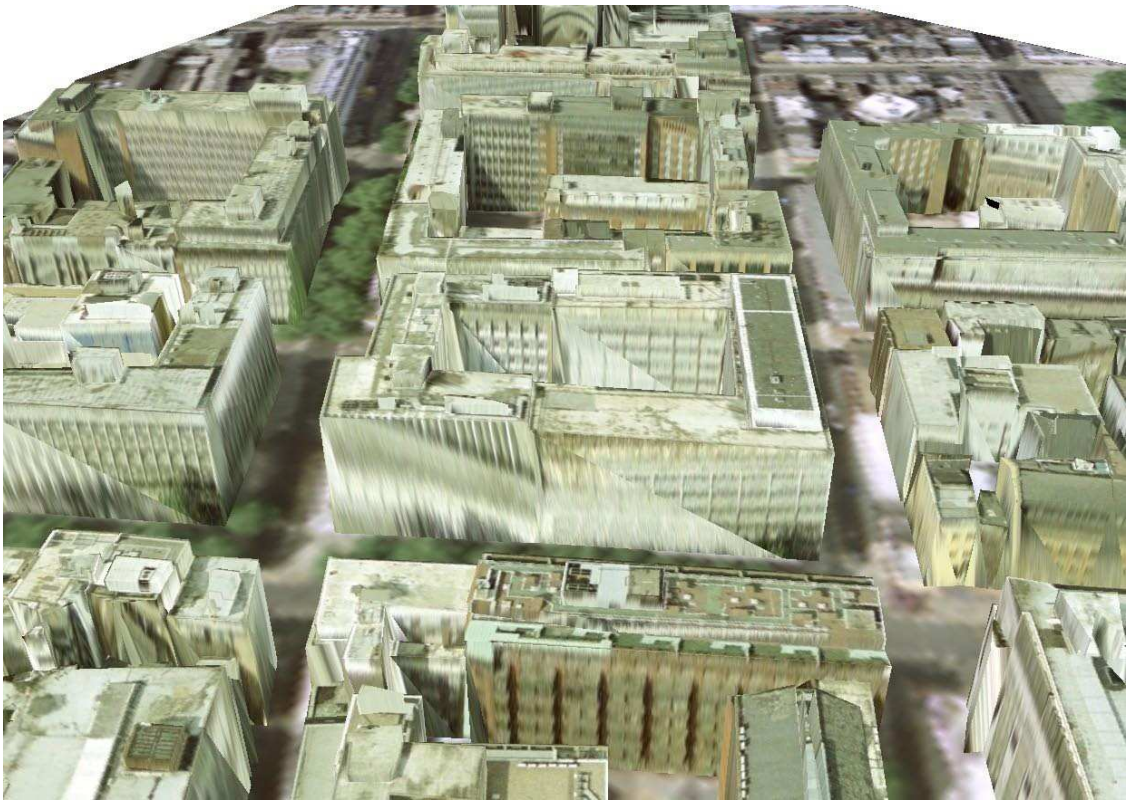


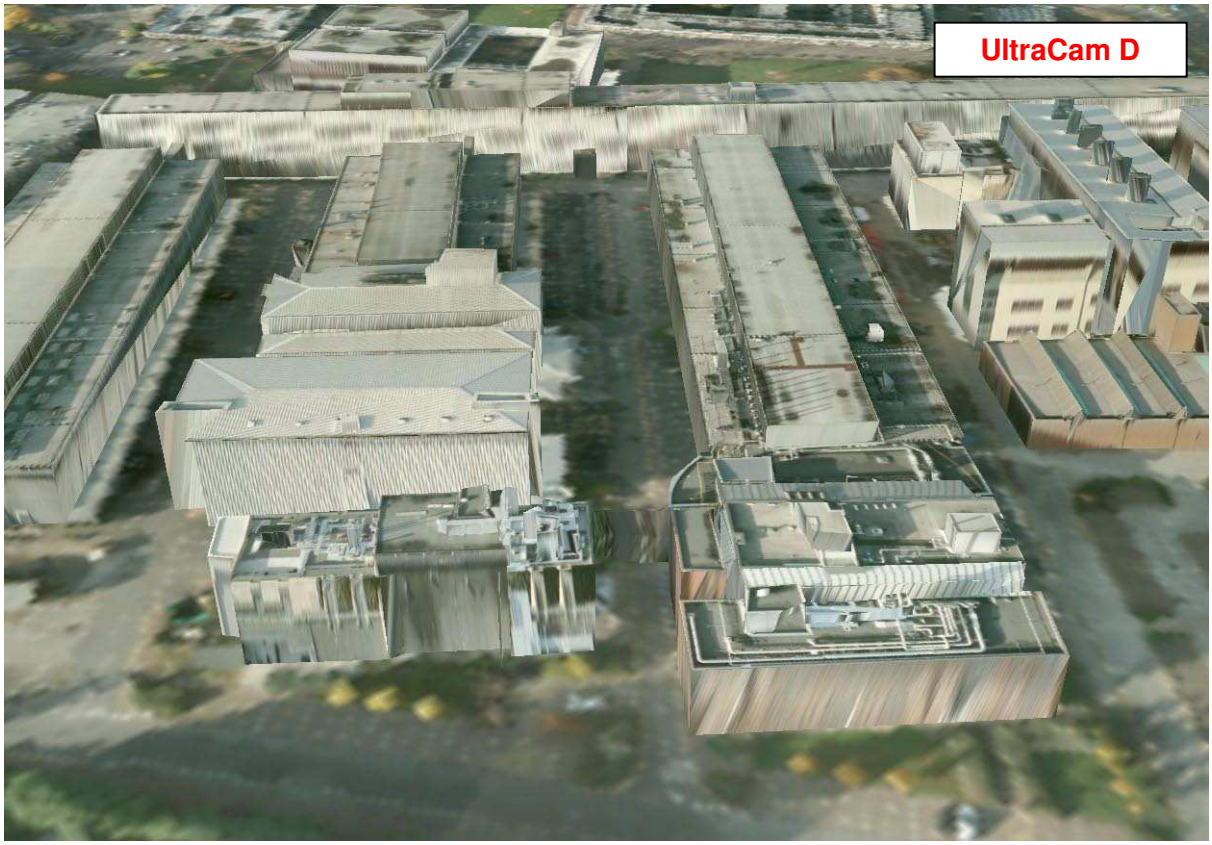
Figure 172 Perspective views of the automatically textured mapped building models in London

As can be easily seen from figure 172 the texture quality of the building models in London is considerably degraded especially for the facades. Most of the roof structures have good quality textures which are equivalent to the quality of the aerial photographs without any significant artefacts. In contrast the vertical lower resolution imagery with 60% overlap proves to be inadequate for façade texturing which in most cases is significantly distorted except for a few occasions.

7.2.3. Automatic texture mapping and visualization for the study area in Nottingham

The texture mapping process for the reconstructed building models in Nottingham presents the most interesting results compared with the other two study areas due to the availability of high resolution vertical UltraCam D imagery. The automatic texture mapping process was performed using the UltraCam D imagery with the workflow described in section 7.2. The results from the texture mapping are given for both cases in figure 173.


UltraCam D


UltraCam D


UltraCam D


UltraCam D


UltraCam D


UltraCam D


UltraCam D

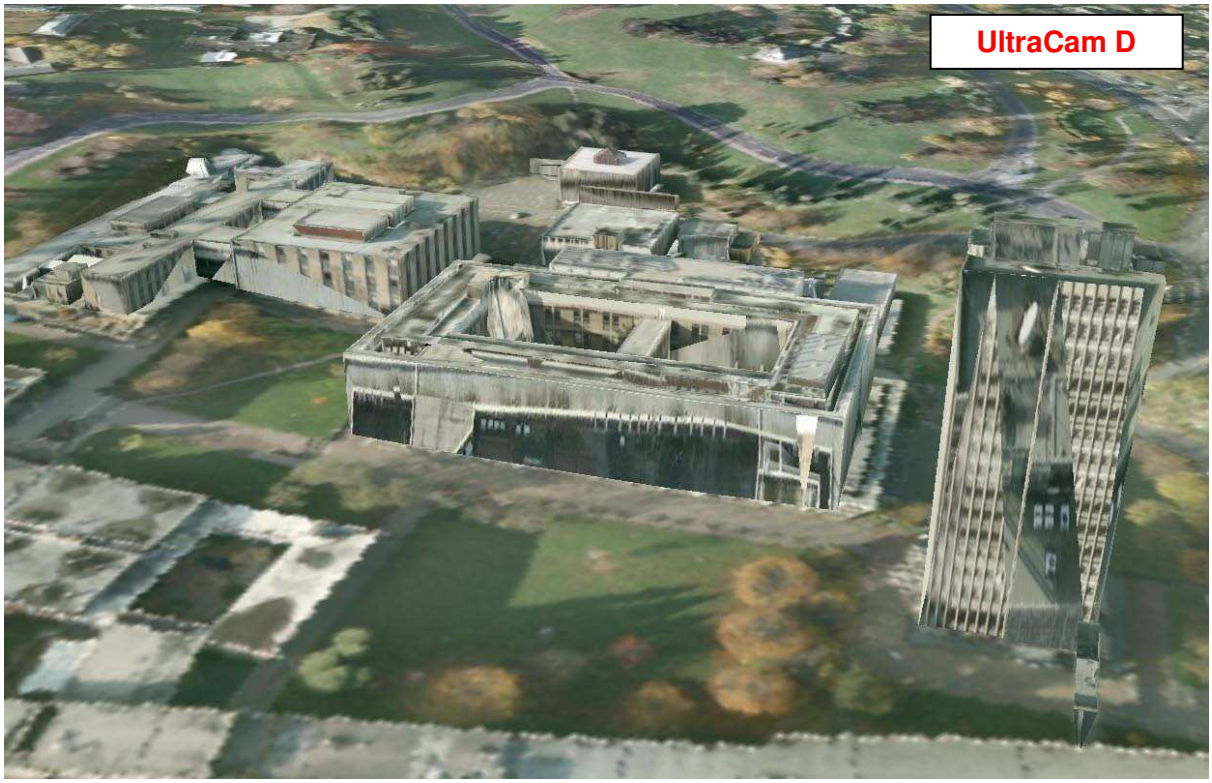

UltraCam D

Figure 173 The automatically textured buildings models using the available Ultracam D airborne images with 80% overlap.

From the images presented in figure 173 it is safe to assume that in most cases with very few exceptions the automatically textured building models using the Ultracam D images achieved superior quality for the roof textures but even with the 80% overlap the façade textures where poor at best and heavily distorted in most cases.

In addition there are several texture artefacts on the façade polygons and it is speculated that these artefacts are caused by errors or occasional bugs of the texture mapping algorithm, coupled with flipped vector normal of the 3D polygons comprising the building models. Clearly more research is required in automatic or semi-automatic texture mapping before high quality and reliable results can be produced.

8. CONCLUSIONS AND FURTHER PROSPECTS

The research introduced a method for the entire workflow of semi-automated building reconstruction, quality assurance and texture mapping using data fusing techniques to increase the accuracy and Level of Detail of the building models. The proposed method was implemented on a variety of different datasets of varied accuracy, resolution and overall quality with the aim to evaluate the effectiveness of the proposed workflow in different situations. The feature extraction process from the ADS40, film/scanned aerial photographs and UltraCam D images, all produced an overall acceptable result, with most of the linear features adequately represented. Furthermore, the implementation of the feature extraction stage was straightforward and easily adaptable to data collected from different sensors. One of the main disadvantages during the stage of feature extraction was the extraction of conjugate points instead of matching linear features directly in the model space. Extracting conjugate linear features could potentially enhance even more the stage of reconstructing the 3D breaklines, instead of trying to detect the 3D breaklines from the VDDSM at a later stage. Despite the acceptable results, there were still a number of building outlines without an adequate number of extracted points.

The stage of building detection produced very promising results since most of the buildings were successfully detected, independent of their size, roof type or when high density LiDAR were available or using VDDSMs. The main advantage is the generic approach of the algorithm that can be implemented in a variety of different situations and datasets of alternative density of resolution. The method is not restricted by the leaf condition of the tree canopies and even undetected individual trees can be filtered during the building classification stage. Nevertheless the proposed method requires a certain level of user interaction since some crucial parameters are very difficult to estimate automatically. Undetected individual trees that are bordering buildings can introduce a false building detection.

The adjustment of the extracted building outlines and 3D breaklines in most cases improved the planimetric accuracy of the building footprint and the Level of Detail of the VDDSMs by incorporating fusion techniques with the stereo matched points. The proposed method is fairly robust by minimizing user interaction since most of the critical parameters can be automatically defined. Nevertheless the accuracy improvement of the building outlines was hindered in some cases by the simplification process.

The procedure of building reconstruction presented accurate and visually impressive results. The method seems to be very reliable, having the ability to reconstruct a large number of roof details. One of the crucial steps is the use of Boolean functions for implementing the Constructive Solid Geometry. In most cases the process yields reliable results with the majority of the adjacent roof planes merged together.

The planimetric evaluation indicated in most cases a small improvement in the accuracy of the building footprint with the use of the stereo matched points for the London and Nottingham study areas. For the Switzerland area it is interesting to note that the improvement in the accuracy was more noticeable in the building outlines derived from the single path LiDAR data. In addition, from the accuracy evaluation in the Switzerland study region it was evident that the planimetric and vertical accuracy of the building models, obtained from the single path LiDAR data were very similar with those obtained from the combined LiDAR data. Although the lower density of the single path LiDAR point cloud seems to have a negative impact on the level of reconstructed roof details, the building outlines and the overall vertical accuracy is similar in both cases. It is also concluded that the vertical accuracy assessment has highlighted the direct link between the accuracy of the reconstructed roof planes and the quality of the Aerial Triangulation and the quality of the GCPs used. It was evident that the best vertical accuracy was obtained for the Nottingham study area were very precise GPS were used and accurate Aerial Triangulation results were obtained.

8.1. Addressing the aims and objectives

This section will address the specific aims and objectives that were set for this project based on the analysis of the results and the presented workflow. The main aim of this project as described in section 1.2 was the development of a semi-automated, innovative workflow for building reconstruction and 3D city modelling.

The main aim as defined above has two important phrases, semi-automated and innovative workflow. The aim of semi-automation will be address at a later stage. The aim to have some innovative processes in the overall workflow was of particular importance. Throughout the proposed workflow for 3D city modelling there were several innovative processes introduced as summarised below.

Addressing the aim of innovation and novelty:

- **Edge extraction for image enhancement prior to stereo matching –** This process was devised in order to enhance the contrast between the linear features visible on the aerial images and the background objects. Although edge operators have been widely used as documented, the successful enhancement of the linear features prior to the stereo matching process is indicating that this innovative process could be used on a regular basis for enhancing the quality and density of the stereo matched point cloud.
- **Building detection process –** Although there have been several attempts to identify building models from optically derived data, this research presented an innovative generic process for identifying the building class from LiDAR data or Very Dense DSM (VDDSM) created solely from airborne imagery.
- **Data fusion for optimising building footprint and 3D breaklines –** One of the most innovative stages introduced is a new method for fusing the LiDAR data with high density stereo matched points from aerial photographs. This new data fusion method was performed in order to improve the accuracy of the building footprint as

well as the accuracy of the reconstructed 3D breaklines that are used subsequently to enhance the Level Of Detail of the reconstructed roof shapes.

- **Semi-automatic quality assurance –** Another important aspect introduced in this research is the proposed workflow for accuracy assessment. Apart from the generic visual assessment and the use of building detection metrics, which are well known in the industry, the report introduces a new innovative process for a complete vertical accuracy assessment. The semi-automatic vertical accuracy assessment is based on the generation of TIN models for each building that are subsequently subtracted from reference building models.

Apart from the main aim of novelty and innovative processes the following objectives were also set at the beginning of the project.

Project objectives:

The project objectives as defined in section 1.2 are summarised below.

1. The final solution would need to have a degree of automation exceeding current industry standards and previous research efforts
2. The solution should present comparable accuracy and LOD characteristics with 3D models extracted from manual stereo plotting procedures and a method to perform accuracy assessment should be proposed.
3. It should be a flexible and reliable solution that can be accommodated in different situations with different dataset characteristics (resolution, quality and number of data sources), adaptable to different scenes with different building architecture.
4. Cost efficiency and availability of data sources should also be considered.

Addressing 1st objective on the level of automation:

The first objective states that the final ‘’workflow should have a degree of automation exceeding current industry standards’’. The ‘’current industry standards’’ were confirmed by

the time required to manually stereo plot the reference 3D building models used in the accuracy assessment presented in chapter 7. The manual stereo plotting was performed by the author, who is an experienced operator of Leica Stereo Analyst, and the time required for an area of one square kilometre was approximately 40 working hours. Although this timeframe fluctuated between the test sites depending on the complexity and density of the building models, it can be considered a representative timeframe for an urban area.

The time requirements for unsupervised processing and manually adjusting parameters from the implementation of the proposed workflow are given in table 39. The processing was performed using a high specification dual core workstation with 4GB or RAM and a powerful 3D graphics card.

Test Site	Area covered by models (sq.km)	Unsupervised processing (hours)	Manual intervention (hours)
Switzerland	1	12	2
London	0.5	8	3
Nottingham	2	15	2

Table 39 Time requirements for unsupervised processing and manually adjusting parameters from the implementation of the proposed workflow in the test sites.

In table 39, the column 'manual intervention' represents the time spent by the operator to adjust parameters and prepare the data for processing in the different software packages. Note that a number of packages were used for the completion of the entire workflow including ERDAS Imagine, LPS, ArcGIS, TerraScan and Geomagic Studio. The time for the 'manual intervention' is in addition to the 'unsupervised processing'. So the total time required for the production of 3D city models in Switzerland, London and Nottingham was 14, 11 and 17 hours respectively. Comparing the required time with the manual methods it is evident that all three test sites have introduced reduction on the required timeframe. Nottingham test site presented the bigger gain, by reducing the time requirements as much as 78%. The time requirements for Switzerland were reduced by 65%, while London presented the smallest gain by reducing time requirements by nearly 50%. Analysing the

results in table 50 we can conclude that the benefit of the proposed workflow in reducing the required time for the production of 3D city models is maximised as the size of the area gets bigger. It is also evident that increased complexity of building models (London site) has a negative impact on the time reduction benefits.

As stated in section 1.2 in order for the workflow to be considered semi-automatic it would need to reduce time requirements by 50% but still require the operator to perform several functions and setting parameters during this time. In contrast an automated algorithm could be named as such if it could reduce the overall processing time by at least 50% and maintain the time spent by the operator for setting parameters within 1%. Essentially for one square kilometre of 3D city models a workflow would qualify as automatic if the operator was spending no more than 30 minutes setting parameters and altering variables and the unsupervised processing of the data was complete in less than 20 hours.

Based on the above consideration the proposed workflow does not qualify as automatic because the time spent for manual intervention in all three test sites exceeds 1% of the total processing time. Since though the benefit from the time reduction was more than 50% and still require time for manual intervention the workflow can be classified as semi-automatic and has effectively met the first objective of the project.

Addressing 2nd objective on accuracy and LoD:

The second objective was set to define a milestone for the accuracy assessment methodology and the level of detail this solution should satisfy. Section 1.2 stated ‘’that in order for the solution to be adaptable in the industry and not purely stay a research prototype, the derived models should be quantitatively assessed and the overall accuracy (standard deviation) should not be worse than 50cm (for height and plan) compared to current manual methods’’. The derived models would also need to have a comparable level of detail with the manual models, hence the solution should be able to generate LOD 2 city models, including the necessary geometric detail and textures.

As discussed in chapter 7 the accuracy of the 3D city models was evaluated based on a novel workflow that combined a planimetric and vertical accuracy assessment method. The advantages of the proposed workflow for accuracy assessment is that it presented a unified workflow for a complete quantitative evaluation by using building detection metrics and a new method for evaluating vertical deviation. For the comparison reference models are used created by manual stereo plotting techniques. The use of an adequate number of stereo plotted reference models is the main limitations of this workflow since in a production environment only a sample of reference building models need to be created to reduce the overall time requirements. In a production environment it is estimated that the reference models should cover at least 10% of the area in order to produce reliable quantitative results. The results from the vertical accuracy assessment are summarised in table 40.

Test Site	Vertical accuracy (St.dev.)	Vertical accuracy (Mean value)
Switzerland	0.45m	-0.14m
London	0.67m	0.02m
Nottingham	0.25m	0.04m

Table 40 Summary from the vertical accuracy assessment for all three test sites

As it is indicated in table 40 two out of three test sites achieved the accuracy milestone set in the project objectives section. The London models fell short of achieving the accuracy requirements and it was due to a number of reasons as described in section 7.1.2.2, including the quality of the initial data as well as the complexity of the buildings. It is safe to assume that the proposed workflow creates accurate 3D city models that satisfy the project objectives when the study area does not consist of very complicated building structures. It's clear that in areas with very complicated building models there is the need for higher density LiDAR and airborne imagery of high ground resolution to satisfy these accuracy requirements.

The completeness or Level of Detail of the reconstructed building models was assessed visually by comparing them against the reference models as presented in section 7.1.1. In all

three test sites the models were fairly similar and despite some inconsistencies with very small roof details most of the reconstructed models would classify as LoD 2 models, thus satisfying the project objective set initially.

Addressing 3rd objective on adaptability and reliability:

The third objective set in section 1.2 stated that ''the solution should be flexible and reliable with different dataset characteristics (resolution, quality and number of data sources), adaptable to different scenes with different building architecture''. The results presented from implementation of the proposed workflow in three test sites are a testament of the adaptability and reliability of the solution. These three test sites presented areas with different building architecture, landscape and model complexity. In addition the datasets available presented different challenges in each case but the solution, despite few differences, was able to produce similar results with respect to the level of detail and accuracy of the models.

Addressing 4th objective on cost efficiency:

The fourth and last project objective as defined in section 1.2 stated that ''consideration should be given to the cost efficiency and availability of data sources''. The first two study areas (Switzerland and London) presented an ideal scenario where both LiDAR and aerial photography is available. These datasets presented the opportunity to utilise the benefits of data fusion techniques for obtaining superior 3D modelling results as discussed in section 1.2. For the third test site though in Nottingham, instead of obtaining LiDAR and airborne imagery it was decided that this study area should present a more cost-effectively solution. Thus the decision to constrain the available datasets to only airborne imagery was made and that decision signified a major challenge since the existing workflow would need to be changed. Therefore the proposed workflow as presented in this report was defined after seriously considering the best alternatives for a cost efficient solution able to provide reliable results even when LiDAR data are not provided.

8.2. Further Prospects

Finally, further research is needed in order to assess the effectiveness of the simplification algorithm, during the adjustment of the building outlines and the reconstruction of the 3D breaklines from VDDSMs. One aspect that requires further examination is the improvement in the automation of the building detection stage since in all cases irrespective of the datasets used, the process required significant user interaction and assessment of the statistics by the operator. Furthermore, additional research for efficient texture mapping using oblique airborne imagery and terrestrial moving platforms (such as a van with direct geo-referencing capabilities and multiple cameras) that could prove particularly useful for providing superior visual results.

REFERENCES

ADS40 brochure, 2002. *Leica Geosystems GIS & Mapping, LLC.* Part No. ADS40 cc 7/02, available online at http://gis.leica-geosystems.com/products/ads40/default.asp, accessed April 2008.

Akenine-Möller, T.,and Haines, E., 2002. Real-Time Rendering, *2nd Ed., A K Peters*

Alhamlan, S., Mills, J. P., Walker, A. S., Saks, T., 2004. *The influence of ground control points in the triangulation of leica ads40 data.* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 35 (B1): 495–500.

Alharthy, A., Bethel, J., Mikhail, E., 2004. *Analysis and accuracy assessment of airborne laserscanning system*, XXth ISPRS Congress, 12-23 July 2004 Istanbul, Turkey, Commission2.

ALS40 brochure, 2002. *Leica Geosystems GIS & Mapping, LLC. Part No. ALS40BRO cc 8/02*, available online at http://gis.leica-geosystems.com/products/als40/default.asp, accessed July 2005

ALS50 brochure, 2004. *Leica Geosystems GIS & Mapping, LLC. Part No. ALS50 brochure cc 01/04*, available online at http://gis.leica-geosystems.com/products/als50/default.asp, accessed July 2005.

ALS50 product description, 2003. *Leica Geosystems GIS & Mapping, LLC. Part No. ADS50PD cc 1/03*, available online at http://gis.leicageosystems.com/products/als50/ default.asp, accessed July 2005.

Althof, R.J., Wind, M.G.J., Dobbins, J.T., 1997. *A rapid and automatic image registration algorithm with subpixel accuracy,* IEEE Transactions on Medical Imaging 16 (1997) 308–316.

Ameri, B. and Fritsch, D., 2000. *Automatic 3D building reconstruction using plane-roof structures.* In ASPRS, Washington DC.

Aurenhammer, F., 1991. Voronoi Diagrams - A Survey of a Fundamental Geometric Data Structure. ACM Computing Surveys, 23(3):345-405, 1991

Bates, P. D., and De Roo, A. P. J., 2000. A simple raster-based model for flood inundation simulation, *Journal of Hydrology*, *236*, 54-77.

Batty, M., Chapman, D., Evans, S., 2000. Visualizing the City: Communicating Urban Design to Planners and Decision-Makers [EB/OL]. http://www.casa.ucl.ac.uk/visualcities.pdf

Batty, M., Smith, A., 2002. Virtuality and cities: Definitions, geographies, designs. In: Peter Fisher and David (Eds.), Virtual Reality in Geography, *Taylor & Francis, London, pp. 270-291.*

Beck, M., 2003. Real – Time Visualization of big 3D city models. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV-5/W10

Besl, P.J., McKay, N.D., 1992. *A method for registration of 3D shapes,* IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (1992) 239–254.

Bham, G. H., and Benekohal, R. F., 2004. A high fidelity traffic simulation model based on cellular automata and car-following concepts. Transport Research, Part C: Emerging. Technology 12(1), 1–32.

Bildstein, F., 2005. 3D city models for Simulation & Training - Requirements on Next Generation 3D city models. *ISPRS WG III/4, Next Generation city models, Bonn, 21-22 June 2005.*

Blonski, S., Pagnutti, M., Ryan, R. E. and Zanoni, V., 2002. In-flight edge response measurements for high spatial-resolution remote sensing systems. *In: W. L. Barnes (ed.), Proceedings of the SPIE: Earth Observing Systems VII, Vol. 4814, pp. 317–326.*

Borgefors, G., 1988. *Hierarchical chamfer matching: a parametric edge matching algorithm,* IEEE Transactions on Pattern Analysis and Machine Intelligence 10 (1988) 849–865.

Brakenridge, G. R., 2006. 2005 Flood Archive, edited, Dartmouth Flood Observatory, Hanover, USA.

Brenner C., Haala N., and Fritsch D., 2001. "Towards fully automated 3D city model generation", *Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images III, 2001*

Brunn, A., and U. Weidner, 1998, *Hierarchical Bayesian nets for building extraction using dense digital surface models*, ISPRS Journal of Photogrammetry & Remote Sensing, 53(5): 296-307.

Buchholz, H., and Döllner, J., 2005. "View-Dependent Rendering of Multiresolution Texture Atlases", *Proc. IEEE Visualization, IEEE CS Press*

Burrough, P.A., 1987. Principles of Geographical Information Systems for Land Resources Assessment. Oxford University Press, Oxford, 194pp

Butler, D., 2004. Noise management: Sound and vision. Nature, 427(6974): 480 - 482.

Chang, S.H., Cheng, F.H., Hsu, W.H., Wu, G.Z., 1997. *Fast algorithm for point pattern matching: Invariant to translations, rotations and scale changes,* Pattern Recognition 30 (1997) 311–320.

Chen , L. C., Teo, T. A., Shao, Y. C., Lai, Y. C., Rau J. Y., 2004. Fusion of LIDAR data and optical imagery for building modelling. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 35(B2): 586.591.*

Cheng, L., Bhushan, A., Pajarola, R., and El Zarki, M., 2004. "Real-Time 3D Graphics Streaming using MPEG-4*", BroadWise '04, San Jose, CA, USA, July 2004.*

Clark, W., Hosking, P., 1986. Statistical Methods for Geographers. *Wiley, New York, 518pp.*

Davenport, I., Holden, N., and Gurney, R., 2004. Characterising Errors in Airborne Laser Altimetry Data to extract Soil Roughness. *IEEE transactions on geoscience and remote sensing, vol. 42, no. 10, october 2004*

Davis, D., Ribarsky, W., Jiang, T. Y., Faust, N., and Ho, S., 1999. Real-Time Visualization of Scalably Large Collections of Heterogeneous Objects, *Proceedings of IEEE Visualization 1999, pp. 437-440.*

Dobashi, Y., Kaneda, K., Yamashita, H., Okita, T., Nishita, T., 2000. A simple, efficient method for realistic animation of clouds, *Proceedings of the 27th annual conference on Computer graphics and interactive techniques. Pages 19-28*

Döllner, J., Baumann, K., and Hinrichs, K., 2000. Texturing Techniques for Terrain Visualization, *Proc. IEEE Visualization, IEEE CS Press, pp. 207-234.*

Dong-Min, W., Quoc-Dat, N., Quang-Dung, N. T., Dong-Chul, P., Young-Kee J., 2008. *Building detection and reconstruction from aerial images.* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B3b. Beijing 2008.

Douben, N., and Ratnayake, R., 2006. Characteristic data on river floods and flooding; facts and figures, in *Floods, from Defence to Management. Proceedings of the 3rd International Symposium on Flood Defence, 25–27 May 2005, Nijmegen, the Netherlands*, edited by A. J. van, et al., Taylor & Francis/Balkema Publishers, Leiden, the Netherlands.

Douglas, D., Peucker, T., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, *The Canadian Cartographer 10(2), pp. 112-122.*

Drosdol, J., and Panik, F., 1985. The Daimler-Benz Driving Simulator, *SAE Paper 85-0334.*

Duda, R., and Hart, P., 1973. Pattern Classification and Scene Analysis, John Wiley and Sons, pp 271-2

Eckardt, A., Braunecker, B., Sandau, R., 2000. *Performance of the imaging system in the LH systems ADS40 Airborne Digital Sensor.* International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B1, Amsterdam, Netherlands

European Commission News, 2005. 'QCITY' and 'SILENCE' – EU projects target urban noise. http://ec.europa.eu/research/transport/news/article_3339_en.html. Access date: March 2008.

European Commission Working Group, 2006. Good Practice Guide for Strategic Noise Mapping and the Production of Associated Data on Noise Exposure. http://ec.europa.eu/environment/noise/pdf/wg_aen.pdf, Access Date: March 2008.

Evans, I., 1980. An integrated system for terrain analysis and slope mapping. Zeitschrift fuer Geomorphologie N.F. Suppl. -Bd. 36, 274–290

FEMA Report I, 2005. Providing structural elevation data – Appendix I: Pictometry accuracy analyses. *Available online at: http://www.fema.gov/pdf/nfip/alt_elevations/elevations_appi.pdf , accessed: May 2008.*

Fischer, A., Kolbe, T., Lang, F., Cremers, A., Forstner, W., Plumer, L., and Steinhage, V., 1998. *Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D.* CVIU, 72(2):163.185.

Fleishman, S., Cohen-Or, D., Silva, T. C., 2005. Robust moving Least-squares Fitting with Sharp Features. *In proceedings of 32nd SIGRAPH 2005 conference.*

Fonseca, L. M. G., Manjunath, B. S., 1996. *Registration techniques for multisensor remotely sensed imagery,* Photogrammetric Engineering and Remote Sensing 62 (1996) 1049–1056.

Förstner, W., 1986. A feature-based correspondence algorithm for image matching, *International Archives of Photogrammetry and Remote Sensing, vol.26, no.3, pp.150-166, 1986.*

Franke, R., Nielson, G. M., 1980. "Smooth interpolation of large sets of scattered data," *International Journal for Numerical Methods in Engineering 15, 1691-1704.*

Fricker, P., 2001. *ADS40 - Progress in digital aerial data collection.* in: Fritsch/Spiller (Eds.), Photogrammetric Week ´01 Herbert Wichmann Verlag, Hüthig GmbH, Heidelberg, Germany, pp. 105 - 116.

Früh, C., Sammon, R., Zakhor, A., 2004. Automated Texture Mapping of 3D city models With Oblique Aerial Imagery. *2nd International Symposium on 3D Data Processing, Visualization, and Transmission, Thessaloniki, Greece, September 2004, pp. 396-403*

Geologic Shell software ., 2008. Computers and Geosciences, software code section. Code available from server at http://207.176.140.93/documents/oldftp/VOL27/v27-10-05.zip

Gorte, B., 2002. *Segmentation of TIN-Structured Surface Models,* In: Proceedings Joint International Symposium on Geospatial Theory, Processing and Applications, on CDROM, Ottawa, Canada, 5 p.

Goshtasby, A., Stockman, G.C., 1985. *Point pattern matching using convex hull edges, IEEE Transactions on Systems,* Man and Cybernetics 15 (1985) 631–637.

Gruber M., Ladstädler R., 2006. Geometric Issues of the digital large format aerial camera UltraCamD. *International Calibration and Orientation Workshop,* EuroCOW 2006. January 2006, Castelldefels, Spain. EuroSDR Commission I and ISPRS Working Group 1/3.

Gruen, A., Wang, X., 1999. CyberCity Spatial Information System (SIS): A new concept for the management of 3D city models in the hybrid GIS. Proc. 20th Asian Conference on Remote Sensing. November, Hong Kong, pp. 121-128

Grun, A., Dan, H., 1997. *TOBAGO—a topology builder for the automated generation of building models.* In: Grun, A., Baltsavias, E., Henricsson, O. (Eds.), Automatic Extraction of Man- Made Objects from Aerial and Space Images (II). Birkhauser, Basel, pp. 149–160.

Grun, A., Wang, X., 1998. *CC-modeler: a topology generator for 3- D city models*. In: Fritsch, D., Englich, M., Sester,M. (Eds.), Int. Arch. Photogramm. Remote Sensing, vol. 32, Part 4. Stuttgart, pp. 188–196.

Gulch, E., Muller, H., Labe, T., 1999. *Integration of automatic processes into semi-automatic building extraction.* In: Int. Arch. Photogramm. Remote Sensing, vol. 32, Part 3-2W5. Munchen.

Gulch, E., Muller, H., Labe, T., Ragia, L., 1998. *On the performance of semi-automatic building extraction.* ISPRS Comm. III Symposium, Columbus, OH.

Heuel, S., Forstner, W., Lang, F., 2000. Topological and geometrical reasoning in 3D grouping for reconstructing polyhedral surfaces. International Archives of Photogrammetry and Remote Sensing 33 (B3/1), 397–404.

Hobson, J. S.., and Williams, P., 1995. Virtual reality: a new horizon for the tourism industry. *Journal of Vacation Marketing **1** 2 (1995), pp. 125–136.*

Hudson-Smith, A., Evans, S., 2001. Information rich 3D Computer Modelling of Urban Environments. CASA, *RICS Geomatics Biennial conference, 2001*

Huertas, A., Nevatia, R., 1988. *Detecting buildings in aerial images.* Computer Vision, Graphics and Image Processing, 41(2), pp. 131–152.

Jacobsen, K., 2008. Tells the number of pixels the truth? Effective resolution of large size digital frame cameras, ASPRS 2008 Annual Convention, Portland.

Jay Lee., **1989**. A drop heuristic conversion method for extracting irregular network for digital elevation models. In GIS/LIS '89 Proc., volume 1, pages 30-39. American Congress on Surveying and Mapping, November 1989.

Jensen, J. R., and D. C. Cowen, 1999, *Remote sensing of urban/suburban infrastructure and socio-economic attributes,* Photogrammetric Engineering & Remote Sensing, 65(5): 611-622.

Jibrini, H., Paparoditis, N., Deseilligny, M., P., Maître. H., 2000. *Automatic building reconstruction from very high resolution aerial stereopairs using cadastral ground plans,* Presented at XIX'th ISPRS Congress, Amsterdam, 16-23 July 2000. Available at http://www.tsi.enst.fr/~maitre/HM:ISPRS 00.ps.gz

Jürgen, D., Benjamin, H., Steffen, S., 2005. An Approach towards Semantics-Based Navigation in 3D city models on Mobile Devices, 3rd Symposium on LBS & TeleCartography, page 171-176 - November 2005

Kaartinen, H., Hyyppä, J., Gülch, E., Hyyppä, H., Matikainen, L., Vosselman, G., Hofmann, A.D., Mäder, U., Persson, Å., Söderman, U., Elmqwist, M. Ruiz, A., Dragoja, M., Flamanc, D., Maillet G., Kersten, T. , Carl, J., Hau, R. , Wild, E., Frederiksen, L., Holmgaard, J. and Vester, K., 2005. EuroSDR building extraction comparison. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Hannover Workshop, 36 (1/W3), CD-ROM .*

Karsli, F., Kahya, O., 2008. *Building extraction from laser scanning data,* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B3b. Beijing 2008.

Katartzis, A., Sahli, H., 2008*. A Stochastic Framework for the Identification of Building Rooftops Using a Single Remote Sensing Image*. In IEEE transactions on Geoscience and Remote Sensing, vol. 46, no. 1, January 2008

Katzenbeisser, R., 2003. *About the calibration of LIDAR sensors*, ISPRS Workshop "3-D Reconstruction form Airborne Laser-Scanner and InSAR data; 8 – 10 October 2003, Dresden

Kokkas, N., 2005. City modelling and building reconstruction with Socet Set v.5.2. BAE Systems, GXP Regional User Conference, Cambridge, UK, September 19–21.

Kokkas, N., 2007. Geodata fusion for automated 3D city modelling. United Kingdom patent application GB0704368.0, 07 Mar.

Kokkas, N., Dowman, I., 2006. Fusion of airborne optical and LiDAR data for automated building reconstruction. In: Proceedings of the American Society for Photogrammetry and Remote Sensing, Reno Nevada, May 1–7.

Kokkas, N., Smith, J, M., 2007. Automated 3D city modelling and the importance of quality assurance techniques. *ISPRS Hannover Workshop, High Resolution Earth Imaging for Geospatial Information, 29 May - 1 June, 2007*

Kropfl, M., Kruck, E., Gruber, M., 2004. Geometric Calibration Of The Digital Large Format Aerial Camera Ultracam. *International archives of photogrammetry remote sensing and spatial information sciences, VOL 35; PART 1, pages 42-44*

Kruck, E., 2006. Simultaneous Calibration of Digital Aerial Survey Cameras. *International Calibration and Orientation Workshop*, EuroCOW 2006. January 2006, Castelldefels, Spain. EuroSDR Commission I and ISPRS Working Group 1/3.

Leberl, F. et al. 2003. The UltraCam Large Format Aerial Digital Camera System, *Proceedings of the American Society For Photogrammetry & Remote Sensing, 5-9 May, 2003, Anchorage, Alaska*

Leberl, F., and Gruber, M., 2005. UltraCam D: Understanding some noteworthy capabilities. (Ed. D. Fritsch), *Photogrammetric Week 2005, Wichmann Verlag, Heidelberg, pp. 57-68.*

Lee S. C., Jung S. K., and Nevatia R., 2002. "Automatic pose estimation of complex 3D building models", *Workshop on Application of Computer Vision, 2002.*

Leica Photogrammetry Suite 2003. LPS Orthobase & Orthobase Pro User's Guide.Leica Geosystems, GIS & Mapping, LLC

Leica RC30 brochure, 2007. Leica Systems overview. Available online at www.**leica**-geosystems.com/common/shared/downloads/inc/downloader.asp?id=6999. Assessed: February 2008

Liang, C., Jianya, G., Xiaoling, C., Peng, H., 2008. *Building boundary extraction from high resolution imagery and LiDAR data.* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B3b. Beijing 2008.

Lin, C. A., Nevatia, R., 1998. *Building detection and description from a single intensity image.* Computer Vision and Image Understanding, 72(2), pp. 101–121.

Longley, P. A.; Goodchild, M. F; Maguire, D. J.; and Rhind, D. W., 2001. Geographic Information Systems and Science; publisher: John Wiley & Sons (August 1, 2001)

Maas, H. G. and Vosselman, G., 1999. *Two algorithms for extracting building models from raw laser altimetry data*, ISPRS Journal of Photogrammetry & Remote Sensing, Vol. 54, 1999, pp. 153–163.

Maas, H. G., 2004. *Closed solutions for the determination of parametric building models from invariant moments of Airborne Laser Scanner data.* Commemorative volume for the 60th birthday of Prof. Dr. A. Grun (IGP, ETH Zurich, 2004), available online at http://www.ty-dresden.de/fghgipf/index.htm

Maas, H., G., 2002. *Methods for Measuring Height and Planimetry Discrepancies in Airborne Laserscanner Data*, Photogrammetric Engineering & Remote Sensing Vol. 68, No. 9, September 2002, pp. 933–940.

Mark, O., Weesakul, S., Apirumanekul, C., Aroonnet, B, S., Djordjevic, S., 2004. Potential and limitations of 1D modelling of urban flooding*, Journal of Hydrology 299, pp 284–299*

Mather, P., 1987. Computer Processing of Remotely-sensed Images. Wiley, New York, 353pp.

Maune, D. (Ed.), 2001. The DEM users manual. American Society for Photogrammetry and Remote Sensing, Bethesda, 549pp

Miliaresis, G., 2006. Geomorphometric mapping of Asia minor from globe digital elevation model. Geografiska Annaler 88A, 209–221

Miliaresis, G., Argialas, D., 2002. Quantitative representation of mountain objects extracted from the GTOPO30 DEM. International Journal of Remote Sensing 23, 949–964

Miliaresis, G., Illiopoulou, P., 2004. Clustering of Zagros ranges from the Globe DEM representation. *International Journal of Applied Earth Observation and GeoInformation 5, 17–28.*

Miliaresis, G., Kokkas, N., 2004. Segmentation and terrain modelling of extra-terrestrial chasmata. Journal of SPATIAL SCIENCES, Vol. 49, No. 2, pages 89-99.

Miliaresis, G., Kokkas, N., 2007. Segmentation & object based classification for the extraction of building class from LiDAR DEMs. Computers & Geosciences, 33, 1076-1087. [http://dx.doi.org/10.1016/j.cageo.2006.11.012]

Miliaresis, G., Sabatakakis, N., Koukis, G., 2005. Terrain pattern recognition and spatial decision making for regional slope stability studies. Natural Resources Research 14, 91–100.

Morin, K., and DR. El-Sheimy, N., 2002. *Post-mission Adjustment Methods of Airborne Laser Scanning Data*, FIG XXII International Congress Washington, D.C. USA, April 19-26 2002

NASA/ICREST Project report, 2001. *Automated building extraction and reconstruction from LIDAR data,* LIDAR R&D Program for NASA/ICREST studies, Project report 09/16/01 available online at http://icrest.missouri.edu/Projects/NASA/Feature Extraction Buildings /index.htm.

National Research Council., 1997. Modelling and Simulation: Linking Entertainment and Defense. *Washington DC: National Academies Press.*

Optech, ALTM brochure 2008. ALTM 3100 brochure, available online at http://www.optech.ca/pdf/Specs/specs_altm_3100.pdf, accessed :April 2008

ORIMA 6.2 Tutorial 2005. *How to do an ADS40 triangulation,* © 1991 - 2001 Dr. L. Hinsken, LH systems, LLC.

Perko, R., 2005. Image Quality: Digital Pansharpening Versus Full Color Film. *Manuscript submitted for publication. Available online at:*

http://www.gtbi.net/export/sites/default/GTBiWeb/soporte/descargas/DigitalPansharpeningVsColorFilm.pdf

Pitas, I., 1993. Digital Image Processing Algorithms. Prentice- Hall, London, 362pp

Preparata P. F., Hong, S. J., 1977. Convex Hulls of Finite Sets of Points in Two and Three Dimensions, *Commun. ACM, vol. 20, no. 2, pp. 87–93, 1977.*

Qing Zhu., Fengchun Li., Yeting Zhang., 2005. Unified representation of three dimensional city models, *ISPRS Workshop on Service and Application of Spatial Data Infrastructure, XXXVI(4/W6), Oct.14-16, Hangzhou, China*

Rabbani, T., Van den Heuvel, F., 2005. *Efficient Hough transform for automatic detection of cylinders in point clouds.* ISPRS Proceedings. Workshop Laser scanning. Enschede, the Netherlands, September 12-14, 2005.

Rossignac, J., 2001. 3D Compression Made Simple: Edge breaker on a Corner-Table, *Invited lecture at the shape modelling international conference, Genoa, Italy, May 2001.*

Rottensteiner, F., 2003*. Automatic generation of high-quality building models from Lidar data.* IEEE Computer Graphics and Applications 23(6), pp. 42-51.

Rottensteiner, F., and Jansa, J., 2002. Automatic extraction of buildings from LIDAR data and aerial images. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences,.34(4): 295.301*

Royan, J., Bouville, C., and Gioia, P., 2003. A New Progressive and Hierarchical Representation for Network-Based Navigation in Urban Environments", *Proc. Vision Modelling Visualization (VMV 2003), pp. 299-307.*

Salathé, R., 2001. Digitales 3D-Stadtmodell Augusta Raurica. Civilian service, *FHBB Fachhochschule beider Basel, Muttenz*

Sampath, A. and Shan, J., 2007. *Building boundary tracing and regularization from airborne Lidar point clouds. photogrammetric Engineering and Remote Sensing, 73(7), pp. 805-812.*

Sandau, R., Braunecker, B., Driescher, H., Eckardt, A., Hilbert, S., Hutton, J., Kirchhofer, W., Lithopoulos, E., Reulke, R., Wicki, S., 2000. *Design Prinsiples of the LH Systems ADS40 Airborne Digital Sensor.* International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B1, Amsterdam, Netherlands, pp.258-265.

Schenk, T., 1999. *Digital Photogrammetry,* Volume I. TerraScience, pages 52-54.

Schuster, R., Braunecker, B., 2000. *Calibration of the LH systems ADS40 Airborne Digital Sensor.* International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B1, Amsterdam, Netherlands, pp.288-294

Shaw, G., Wheeler, D., 1985. Statistical Techniques in Geographical Analysis. *Wiley, Chichester, 364pp.*

Shufelt, J.A., 1999. Performance evaluation and analysis of monocular building extraction from aerial imagery. *IEEE Pattern Analysis and Machine Intelligence, 21(4):311-326.*

Silvia, R., Ricardo, H., Luis, C, J., 2003. Evaluation and prediction of noise pollution levels in urban areas of Cdiz (Spain). *Acoustical Society of America Journal: Volume 114, pp-2439-2439.*

Skidmore, A., 1989. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. International Journal of Geographical Information Systems 3, 323–334.

Smith, M. J., Kokkas, N. and Qtaishat, K. S., 2007. Investigation into Self-Calibration methods for the Vexcel UltraCam D Digital Aerial Camera. *In:* ISPRS Hannover Workshop, High Resolution Earth Imaging for Geospatial Information, 29 May - 1 June. pp. 6

Smith, M. J., N. Kokkas., A. M. Hamruni., D. Critchley., A. Jamieson., 2008. Investigation into the orientation of oblique and vertical digital images, *EuroCOW 2008: International Calibration and Orientation Workshop, Spain, Castelldefels, Jan 30 – Feb 1*

Smith, M. J., Qtaishat, K., Park, D., and Jamieson, A., 2005. Initial Results from the Vexcel UltraCam D Digital Aerial Camera. *ISPRS Hannover Workshop on High resolution*

Earth imaging for geospatial information Proceedings, Volume XXXVI Part I/W3 ISSN No. 1682-1777.

Smith, R., 2006. "Technology disruption in the simulation industry". *Journal of Defense Modelling and Simulation, 3(1), pp. 3-10.*

Smith, R., 2007. Five forces of game technology adoption in military simulation, *SimTecT Simulation Conference, Brisbane, Queensland, Australia, June 4 - 7, 2007*

Sohn, G., Dowman, I., 2006. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. PE&RS (62)*, No. 1, May 2007, pp. 43-63.*

Stockman, G., Kopstein, S., Benett, S., 1982. *Matching images to models for registration and object detection via clustering,* IEEE Transactions on Pattern Analysis and Machine Intelligence 4 (1982) 229–241.

Suveg, I. and Vosselman, G., 2004. *Reconstruction of 3D building models from aerial images and maps.* ISPRS Journal of Photogrammetry and Remote Sensing, 58(3/4), pp. 202-224.

Taillandier, F., Deriche, R., 2004. *Automatic building reconstruction from aerial images: A generic Bayesian framework,* XXth ISPRS Congress, 12-23 July 2004 Istanbul, Turkey, Commission III, WG III/4.

Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P., 2007. *Hough transform and extended RANSAC algorithm for automatic detection of 3D building roof planes from Lidar data,* IAPRS Volume XXXVI, Part 3 / W52, 2007.

Tempelmann, U., Börner, A., Chaplin, B., Hinsken, L., Mykhalevych, B., Miller, S., Recke, U., Reulke, R., Uebbing, R., 2000. *Photogrammetric software for the LH Systems ADS40 Airborne Digital Sensor.* International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B1, Amsterdam, Netherlands, pp.552-559.

TerraScan 2004. TerraScan user guide, Terrasolid Limited © 1998-2004 Arttu Soininen.

Thomas, H. Cormen., Leiserson, E. C., Rivest, L. R., Stein, C., 2001. Introduction to Algorithms, Second Edition. *MIT Press and McGraw-Hill. pp. 527–529 of section 22.1: Representations of graphs. ISBN 0-262-03293-7.*

Verbree, E., Van Maren, G., Germs, R., Jansen, F., Kraak, M. J., 1999. Interaction in virtual worlds views – Linking 3D GIS with VR, *International Journal of Geographic Information Science, 13, 4, 385-396.*

Vosselman, G., 2002. *Fusion of Laser Scanning Data, Maps, and Aerial Photographs for Building Reconstruction,* IEEE International Geoscience and Remote Sensing Symposioum IGARSS'02, Toronto, Canada, June 24-28

Vosselman, G., and Dijkman, S, T., 2001. 3D building model reconstruction from point clouds and ground plans. *International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences, 36(3/W4): 37.44*

Vosselman, G., Gorte, B.G.H., Sithole, G., Rabbani, T., 2004. *Recognizing structure in laser scanner point clouds,* International Archives of Photogrammetry, Remote Sensing and Spatial Information sciences, Vol. 46, part8/W2, Freiburg, Germany, October 4-6, pp. 33-38

Vozikis, G., Jansa, J., 2008. *Advantages and Disadvantages of the Hough Transformation in the Frame of Automated Building Extraction*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B3b. Beijing 2008

Wang, S., 2008. *Floating model for building reconstruction from topographic maps and lidar data.* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B3b. Beijing 2008.

Wang, Z., 1998. *Extracting building information from LIDAR data*. ISPRS Commission III Symposium on Object Recognition and Scene Classification from Multi-Spectral and Multi-Sensor Pixels, Columbus, Ohio

Watson, D. F., 1981. Computing the N-Dimensional Delaunay Tessellation with Apllication to Voronoi Polytopes, *The Computer Journal (24) 1981*

Weidner, U., Forstner, W., 1995. *Towards automatic building extraction from high resolution digital elevation models.* ISPRS Journal of Photogrammetry and Remote Sensing, 50(4), pp, 38-49.

Williams, P., Hobson, J. S., 1995. Virtual reality and tourism: fact or fantasy? *Tourism Management, Volume 16, Number 6, September 1995 , pp. 423-427(5)*

Willuhn, W., Van Gool, L., 2005. *Building reconstruction from aerial images using efficient semiautomatic building detection,* International Society of Photogrammetry & Remote Sensing, Hannover Workshop 2005

Wing, K., Kwong, A., 2006. Visualization of Complex Noise Environment by Virtual Reality Technologies, *Environment Protection Department (EPD), Hong Kong. http://www.science.gov.hk/paper/EPD_CWLaw.pdf, Access Date: March 2008*

Yotsumata, T., Okagawa, M., Fukuzawa, Y., Tachibana, K., Sasagawa, T., 2002. *Investigation for mapping accuracy of the Airborne Digital Sensor-ADS40.* ISPRS Comission I, Land Satellite Information IV Conference 10-15 November 2002, Denver, CO USA

Yu, D., Lane, S. N., 2006. Urban fluvial flood modelling using a two dimensional diffusion-wave treatment, part 1: mesh resolution effects, *Hydrological Processes*, *20*, 1541-1565.

Zevenbergen, L., Thorne, C., 1987. Quantitative analysis of land surface topography. Earth Surface Processes and Landforms 12, 47–5

Zhang, K., Yan, J., Chen, S. C., 2006. *Automatic construction of building footprints from airborne Lidar data.* IEEE Transaction on Geoscience and Remote Sensing, 44(9), pp. 2523-2533.

Zhu Qing, Lin Hui, 2004. CyberCity Geographic Information System. *Wuhan: Wuhan University Press, pp. 199-203*

Zlatanova, S., 1999. 3D GIS for urban development. *PhD. Thesis. http://www.gdmc.nl/zlatanova/PhDThesis/html/content.html (accessed March 2008).*

APPENDIX I

Algorithm 1 Feature extraction from optical data

Data: Multiple stereo pairs of aerial images, polygons of the AOI

Result: Stereo matched points

Begin

Select panchromatic bands for all look angles

If only vertical images are used **then**

Proceed to step 6.2.2

Else

continue from step 6.2.1

End

Erdas Imagine

6.2.1 **foreach** interest area polygon **do**

Back project on image space and calculate image co-ordinates

If image co-ordinates are within the 14^0 backward view **then**

Select 14^o backward look angle for stereo matching

else

Select 28^o forward look angle for stereo matching

end

Select nadir look angle for stereo matching

end

Erdas Imagine

6.2.2 **foreach** selected image of the stereo pair **do**

Apply Sobel edge operator with kernel size 3x3

Merge extracted edges and calculate new DN values for images

end

Erdas Imagine

6.2.3 **foreach** enhanced stereo pair **do**

Back-project polygons of interest areas on image space

Set parameters for stereo matcher

Search window size (x,y) = 20,5

Correlation window (x,y) = 3x3

Correlation coefficient limit = 0.80

Apply stereo matching on overlapping areas inside interest polygons

Calculate quality statistics

If "suspicious percentage" < 20% **then**

Accept results and export 3D points in a shapefile format

else

Increase correlation coefficient by 0.02 and repeat from 6.2.3

end

Continue loop from point 6.2.3 until "suspicious percentage" > 20%

up to maximum of 10 iterations

LPS

end

end

Algorithm 2 TIN generalization and VDDSM generation

Data: Previously stereo matched points, AT results, LiDAR data

Result: VDDSM, Geomorphometric parameters

Begin

Check if liDAR data are available

If LiDAR data exist **then**

Create density map and measure density of points (D)

If (D) < 4 pts/sq.m then merge LiDAR with stereo matched points and proceed to step 6.3.1 **else** Skip process and go to building detection

End

Else

Proceed to step 6.3.1

End

TerraScan and ArcGIS

6.3.1 Create TIN surface with a Delaunay Triangulation from merged dataset

Perform TIN generalization

Set parameters for TIN generalization

Error threshold = vertical RMS check point residuals from AT

Convert generalized TIN to VDDSM

Set parameters for VDDSM generation

Grid Sampling= 4 pixels per square meter

foreach grid cell of the VDDSM **do**

Calculate Slope in degrees

Calculate Aspect and classify pixels in 8 orientation groups

End

ArcGIS

End

Algorithm 3 Building detection from high density LiDAR data

Data: LiDAR data from multiple flight paths, polygon of the study area

Result: Classified LiDAR points representing only buildings

```
Begin
6.4.1   foreach study area polygon do
            Calculate density map from LiDAR data & average density value (D)
             If multiple LiDAR point clouds overlap then
             |      Merge LiDAR points from multiple flight paths
            else
             |      Select LiDAR point cloud where (D) is higher
            end
        end                                                          } TerraScan
        foreach selected point cloud do
6.4.1.1     Set parameters for scanning point cloud with generic tree models
                Minimum tree height = 2m
                Maximum tree height = 40m
                Width percentage = 30%
            Delineate trees and remove detected points from selected point cloud
6.4.1.2     Set parameters for classifying ground
                Maximum building size = 300m
                Maximum terrain angle (Ta) = "user defined"
                    If maximum terrain angle (Ta) = "null" then
                    |      Calculate Ta = 60° * D / 4
                    |      Set maximum value for (Ta) = 80°
                    |              If (Ta) > 80 ° then
                    |              |      assign Ta = 80°
                    |              else
                    |              |      maintain calculated (Ta)
                    |              end
                    else
                    |      Maintain user defined (Ta)
                    end                                              TerraScan
                Maximum iteration angle = 5°
                Maximum iteration distance =1m
            Detect ground surface and create ground TIN model
6.4.1.3     Classify LiDAR points up to 2.5m above ground TIN as low features
            Merge points classified as low features with the ground surface
                 If study region has area > 1km² then
                 |      Split region into four subset areas
                else
                 |      Retain existing study area
                end
        end
        foreach subset point cloud do
            Calculate average point spacing from the density map (Sm)
6.4.1.4     Set parameters for building detection
                Minimum building size = 40m²
                Maximum building size = 30000m²
                Minimum detail = 4* Sm
                Maximum roof angle = "user defined"
```

If maximum roof angle = “null” **then**
Set maximum roof angle = 60^0
else
Utilize user defined value
end
Elevation tolerance = Sm
Perform building detection and assign points to building class
End
Unite individual subsets to create a single classified point cloud
6.4.1.5 Create DSM using only the building class
6.4.1.6 Reclassify DSM into a binary image
Perform raster to vector conversion and create building polygons
6.4.1.7 Overlay unclassified laser points and select remaining building points
Merge selected points with initial building detection
end

ArcGIS

Algorithm 4 Data fusion for adjusting building outlines and 3D breaklines

Data: classified High density LiDAR (H.D.LiDAR), Stereo Matched Points, Classified building class from VDDSM

Results: Adjusted building footprint for H.D.LiDAR, optimised 3D point cloud from VDDSM

Begin

Check input

If input is H.D.LiDAR **then**

Calculate average point spacing from the density map (Sm)

6.5.1.1 **foreach** void area larger then $10m^2$ on the classified ground points **do**

Check if classified building points are present

If classified building points exist **then**

Set parameters for plane fitting algorithm

Minimum plane size (Ps) = 4*Sm

Elevation tolerance = Sm

Merge planes (increase in tolerance) = Sm + (Sm/4)

Perform plane fitting and visualize roof planes

else

Exclude void area from further processing

end

end

TerraScan

foreach roof plane **do**

Extrude plane boundary on the ground surface

Transform vertical facades into 2D polygons

6.5.1.2 **Set** topology rules for spatial cleaning

ArcGIS

Dangle length tolerance = “user defined”

Fuzzy tolerance = “used defined

If either “dangle length” or fyzzy tolerance =“null” **then**

Dangle length tolerance = 2*(Ps)

Fuzzy tolerance = 2*(Ps)

else

Utilize user defined values

end

Perform spatial cleaning and produce unified building polygons

Convert vector polygons to vector polylines

End

foreach vector polyline **do**

6.5.1.3 **Set** parameters for simplification algorithm

linear tolerance = “user defined”

minimum area = 10m^2

If linear tolerance = “null” **then**

Obtain lengths of the constituent linear segments

Calculate average length value (Al)

Select linear segments (Ls) with length< Al

linear tolerance = average of (Ls) lengths

else

Utilize user defined values

end

Create simplified building footprints

end

foreach simplified polyline **do**

6.5.1.4 Break polylines into the constituent linear segments
Overlay stereo matched points

foreach linear segment **do**

ArcGIS

Estimate buffer size

Set maximum search size (Ms) = 3* (Sm)

Generate buffer zone rings of 25cm until (Ms) is reached

Obtain number of points (Np) within buffer zone1

If (Np) within buffer zone1 (0-25cm) < 5 **then**

Repeat loop for second zone (0-50cm) until (Np) >5

else

Specify buffer zone size = 25cm

end

ArcGIS

Create buffer zones and filter stereo matched points

6.5.1.4 Apply least squares adjustment using filtered points
Extend linear segments until intersected

TerraScan

Create closed building footprint

end

end

If input is Classified building class from VDDSM **then**

Calculated Slope and Aspect raster map

6.5.2.1 Convert raster building class to vector building polygon

Convert vector building polygons to vector polylines

foreach vector polyline **do**

6.5.1.3 **Set** parameters for simplification algorithm

linear tolerance = “user defined”

minimum area = $10m^2$

If linear tolerance = “null” **then**

Obtain lengths of the constituent linear segments

Calculate average length value (Al)

Select linear segments (Ls) with length< Al

linear tolerance = average of (Ls) lengths

else

ArcGIS

Utilize user defined values

end

Create simplified building footprints

end

foreach simplified footprint **do**

6.5.1.4 Break polylines into the constituent linear segments
Overlay stereo matched points

foreach linear segment **do**

Estimate buffer size

Set maximum search size (Ms) = 3* (Sm)

Generate buffer zone rings of 25cm until (Ms) is reached

Obtain number of points (Np) within buffer zone1

If (Np) within buffer zone1 (0-25cm) < 5 **then**

Repeat loop for second zone (0-50cm) until (Np) >5

else

Specify buffer zone size = 25cm

end

Create buffer zones and filter stereo matched points

ArcGIS

6.5.1.4 Apply least squares adjustment using filtered points
Extend linear segments until intersected

TerraScan

Create closed refined building footprint

end

end

6.5.2.2 **foreach** refined building footprint **do**

Project footprint on the slope raster map

Apply threshold for slope map (SL.M) = 20^0

Identify breaklines where slope change is > SL.M

Create initial 3D breaklines

ArcGIS

end

foreach initial 3D breakline **do**

1.5.2.2. Break 3D breaklines into the constituent 3D linear segments
Overlay stereo matched points

foreach 3D linear segment **do**

Estimate 3D buffer size

Set maximum search size (Ms) = 3* (Sm)

Generate buffer zone rings of 25cm until (Ms) is reached

Obtain number of points (Np) within buffer zone1

If (Np) within buffer zone1 (0-25cm) < 5 **then**

Repeat loop for second zone (0-50cm) until (Np) >5

else

Specify buffer zone size = 25cm

end

Create buffer zones and filter stereo matched points

Foreach filtered stereo matched check point **do**

Project point on the Aspect raster map and get value

End

ArcGIS

If all filtered points have same aspect value **then**

Fit refined 3D linear segment using filtered Points

else

Reject points with different aspect value

TerraScan

Extend linear segments until intersected

end

Create closed refined 3D roof polygons

End

Employ Delaunay triangulation using the extracted points representing the ground surface to create a bare-earth TIN surface

Group 3D roof polygons in height levels depending on their size and elevation.

ArcGIS

6.5.2.2. **Begin** iterative TIN generation

For the first roof level **do**

Project vertically the 3D polygon outlines on the ground TIN

Get ground elevation for each node of polygons

Create final footprint by creating a buffer of 5cm

Add 3D polygons and footprints on ground TIN and create **Level 1 TIN**

End

For each subsequent roof level repeat loop using the TIN generated from the previous roof level.

end

ArcGIS

6.5.2.3 **Begin** generation of optimized 3D point cloud

Create grid of 2D points (input : spacing and coverage)

Specify spacing = 20cm

Specify coverage = User input or polygon of study area if available

End

Foreach 2D point **do**

Project point vertically until it's intersected with the refined TIN

Calculate elevation from the TIN triangle and store in point Database

Keep 3D points that are within the building footprint

end

end

end

Geomagic Studio

Algorithm 5 Building reconstruction from using 3D point cloud

Data: Classified LiDAR point could, classified 3D point cloud from VDDSM, adjusted building outlines, roof 3D breaklines

Result: Solid 3D building models

Begin

6.6.1 Merge building footprints and roof breaklines with 3D point cloud

Orthographically project building outlines on the ground surface

} ArcGIS

6.6.2 **foreach** projected building footprint **do**

Set parameters for plane fitting algorithm

Minimum plane size (Ps) = 4*Sm (spacing from density map)

Elevation tolerance = Sm

Merge planes (increase in tolerance) = Sm + (Sm/4)

Restrict plane fitting within roof 3D breaklines

} TerraScan

Fit planes and visualize results

Create vertical facets from footprints and intersect with roof planes

6.6.3 Convert planar facets to volumetric primitives } AutoCad

Set parameters for Constructive Solid Geometry and Boolean merging

Buffer distance = 2* (Ps)

} ArcGIS

Apply Boolean union and create solid 3D building models

end

end