

3D RECONSTRUCTION FROM VERY HIGH RESOLUTION SATELLITE STEREO AND ITS APPLICATION TO OBJECT IDENTIFICATION

J. R. Kim , J.-P. Muller

Dept. of Geomatic Engineering, University College London, Gower Street, London, WC1E 6BT UK
jkim@ge.ucl.ac.uk, jpmuller@ge.ucl.ac.uk

Commission IV, IC WG II/IV

KEY WORDS: High resolution satellite image, DEM, DSM, Lidar, automated Building and tree detection

ABSTRACT:

New high-resolution satellite images such as those from IKONOS provide potentially useful information for the identification of individual surface objects such as trees and buildings compared with previous satellite image sources (e.g. SPOT, Landsat). However, the detailed structure of buildings and trees is still ambiguous for the application of existing machine vision algorithms, mainly due to the resolution limit, currently about 1 metre, the existence of strong shadows due to the large oblique angles and the local time of overpass. Other data sets such as Lidar, aerial photographs, GIS information and pre-existing DEMs, can be used to complement these IKONOS image drawbacks. In particular, owing to the current paucity of IKONOS stereo-pairs, it may be possible to utilise Lidar data as an alternative source of 3D information, where available. This hypothesis has been tested with a 1m pan sharpened multispectral image of East London using a geometrically corrected IKONOS image and coincident 2m Lidar data using a new algorithm which identifies regions of interest (ROIs) for landscape objects (such as trees & buildings). This new algorithm was developed using a Bald Earth Digital Terrain Model (DTM) derived from Lidar and NDVI derived from IKONOS. As a result, excellent building outlines and individual tree extents were detected. However, results shown here applied to an IKONOS stereo-derived DEM indicate that the existing quality of such DEMs is not yet good enough for the detection of small objects.. A new processing system to refine IKONOS DEM quality is currently being developed and higher resolution satellite images (such as Quick-Bird) have recently appeared leading to the expectation that high resolution satellite based GIS construction may be realised in the near future.

1. INTRODUCTION

IKONOS 1m resolution is superior to all previous civilian satellite data. Unfortunately, although 11-bit images are recorded, the image quality and resolution appear to make them unsuitable for the direct application of machine vision algorithms and the large size (11 x 11km) makes the search space too large for object detection.

One alternative approach is to use segmentation of DSMs (Digital Surface Models) in a focusing strategy to reduce the search space into more manageable discrete units. There are several very efficient algorithms available for Lidar and photogrammetrically derived DSMs. We have assessed stereo IKONOS DEM quality from commercially available software for this purpose, such as SocetSet, as well as our own in-house stereo-matching system (Kim & Muller, 1999) but the output DSM quality appears to be insufficient for the automated delineation of individual buildings in urban areas apart from very large structures so that such focusing strategies don't work. Also the positioning error of IKONOS Geo product is too large to register correctly multi-source data-sets. Two possible solutions are here tested. The first is to exploit secondary information such as Lidar and the second is to update the positioning accuracy and DEM quality of IKONOS stereo. Currently very good results have been obtained using the first approach. In the case of IKONOS stereo DEM updating, this work is still in progress and only preliminary results are shown here.

1.1 Data set description

Two data sets have been utilised to date:

1) East London area comprising a 12 km by 12 km IKONOS Precision 1m Pan sharpened image, 2m resolution Lidar which is provided by Infoterra Limited, 4,600 kinematic GPS (7 tracks) (kGPS) measurements and Landline® plus GIS vector data for some part of the area.

2) Barton Bandish test area in Eastern England comprising a pair of IKONOS stereo panchromatic 1m and 4m multispectral images covering an 11 x 11km area, 40,000 kGPS measurements and Landline® plus data. Field survey work was performed in the test areas (Barton Bendish, East London) by collecting k-GPS and static GPS point measurements as well as using a digital camera to obtain "ground truth".

1.2 Previous research work

The first significant research work concerning focusing strategies was published by Baillard and Maitre (1999). They proposed as their focusing strategy segmenting a scene into two different kinds of components: "at ground" and "above-ground" objects. The set of 3-D points is segmented into homogeneous altitude regions by a classical region growing and merging algorithm. In coarse DEM areas, they suggested graph based interpolation and the application of this is used as part of our gridding component development. (see later)

Haala (1997) employed a DSM segmentation technique based on local surface normals. The normal direction from a Lidar DEM was approximated using the derivatives of a local bivariate polynomial fit. The direction of the unit normal vector of a possible roof plane emerging from the ground plan is then

perpendicular to the segment of it. So within a building outline, the correct extent of heights can be determined and segmentation carried out.

Both methods are unsuitable for our use, due to the lack of sufficiently good quality and/or sufficiently high enough resolution DEMs.

External data such as maps and digital map data stored in a GIS, can also be used to define a ROI as in Roux (1997), Brenner and Haala (1998). However, such additional data is not always available.

One simple method to detect a building or a tree is a morphological opening filter. By applying a “top hat” filter to a DEM, building area height points can be removed. This method is usually applied to low resolution imagery since around very large buildings or clusters of buildings, the filter often fails to derive the correct ground DEM.

We used a modified version of this filter for the initial stage to define the ROI.

According to our end user specifications, bald surface DTMs are also required. There is some research about bald surface DEM construction (often referred to as a Digital Terrain Model or DTM) from a DEM. Usually such methods are based on a hierarchical fitting scheme such as the one described by Kraus & Pfierfer (1998).

Another example is the method described by Tao et al (2001) using Lidar and very high resolution airborne SAR. The authors initially fitted some ground seed points starting at low resolution. “Above ground” height points were detected by their height difference and the SAR amplitude information, and applied to a next level fitting stage with a low weighting value. Lu et al (1998) developed a method, which is very similar to Tao’s method but used image classification results instead of SAR amplitude.

1.3 Fundamental Approach

A DSM segmentation based approach cannot produce correct landscape object detection with low density height points and/or poor quality heights, because the gridding error in object boundaries prevents correct building definition.

Therefore, the identification of a correct building and tree extent in our method is dependent on the accuracy to which we can retrieve a Bald Earth DEM. By splitting all height points into “above-ground” and “at ground” components using this information, a new gridding scheme with only “above-ground” height points has been developed to obtain relatively clear building and tree boundaries. Moreover, the “above-ground” points can also be separated by an NDVI mask so that the trees, which are near to a building boundary don’t impinge on the building ROI quality. Through a hierarchical scheme, the quality of the ROI can be updated, because the slope is smoother at coarser resolution. A tension factor based gridding method has been employed to take this effect into account.

1.4 Structure of this Paper

In section 2, the overall processing scheme and some outline details of the algorithmic method are described for object detection including bald surface extraction, gridding of lidar points and the extraction of a normalised DEM. Assessments of

this scheme were made with the two input data-sets, viz. 1) IKONOS and multispectral pair, 2) Lidar and IKONOS pan sharpened multispectral image combination. In both cases, a positioning accuracy check was performed and an additional processing step to update the IKONOS Geo stereo based product’s positioning quality was applied. This step was crucial to update the registration accuracy to compensate IKONOS stereo derived DEM quality with other data sources through, for example, the pan-sharpened multi-spectral images.

The quality of the ROIs for building and tree areas was assessed by comparing them against the Landline® plus digital map vector data.

2. ALGORITHM DESCRIPTION

The overall approach is shown schematically in Figure 1. To save CPU time and construct a more accurate object boundary, the “above-ground” and “ground” points are separately gridded. The most crucial factor in this algorithm is the selection of “ground seed points”. If the seed points are correctly selected, the constructed bald surface should be accurate enough to define a sensible ROI.

2.1 Bald Earth seed points selection

The simplest method for this step is to use a local minimum operator (Kilian et al., 1996). However, the local minimum based Bald Earth reconstruction is very sensitive to noise and the reconstructed surface based on these seeds points is usually lower than the actual bald Earth surface. Therefore, we developed a more robust technique based on the local height distribution. At the start, every height point is divided by the local minimum value to remove the effect of any slope bias. Using a moving window, whose size is determined by an initial estimate of maximum building size, the local height distribution was analysed.

If the moving window size is large enough, a height distribution should consist of two Gaussian functions, so that the mean value of a lower distribution can be used for reliable seed points. To fit two normal distributions, we use Kittler and Illingworth (1986)’s criterion.

$$J(t) = 1 + 2.0(P_1(t) \log \sigma_1(t) + P_2(t) \log \sigma_2(t)) - 2.0(P_1(t) \log P_1(t) + P_2(t) \log P_2(t))$$

(where $P(t)$ is the sum of Probability Density Functions (PDFs) and σ is the standard deviation)

A threshold value for two normal distributions is then minimized using the above function. After finding this threshold value, two mean values can then be calculated.

Now,

If $\mu_2(t) - \mu_1(t) <$ the estimated object height, then $\mu_1(t)$ is the mean value height of a larger object surface or it is assumed to be flat earth so that the window size then needs to be extended and the estimation repeated.

If $\mu_2(t) - \mu_1(t) >$ the estimated object height, then $\mu_1(t)$ is selected as a seed point within the window

Applying this method, suitable seed points can then be selected.

2.2 Gridding of Bald earth

From the initial seed points, an approximate Bald Earth DTM can be constructed by grid-point interpolating the seed points selected in the previous step. However, the interpolation interval needs to be too large enough to avoid interpolating noise. Hence, we use Smith & Wessel (1990)'s method to interpolate bald earth seed points. It's one kind of fitting method for the solution of the following function:

$$(1-T) \nabla^2 Z - TVZ = 0$$

(where Z height points, T : Tension factor between 0.0 to 1.0)

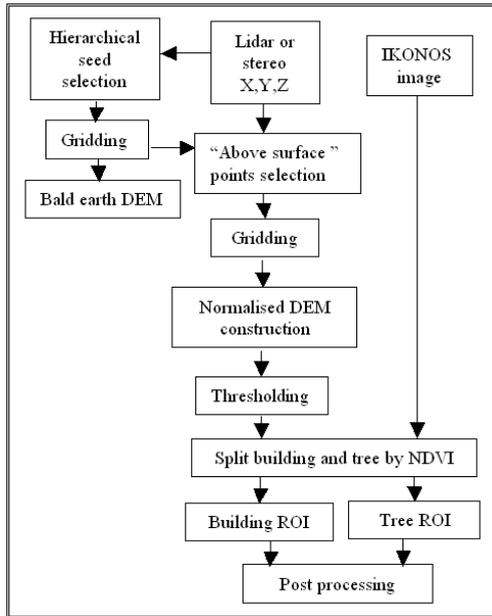


Figure 1 ROI definition scheme from height and NDVI

Using this method, a higher tension factor will produce a higher curvature surface. To construct a smoother surface at a low hierarchical stage, a low tension factor is required. In the first stage, the lowest tension factor and largest window size are used to save CPU time. From experience, they do not appear to seriously affect the quality of the derived bald earth surface. This implies that the accuracy of the derived Bald Earth is relatively free from user defined parameters in our suggested processing method if and only if, the seed point selection is appropriate and not subject to gross errors.

Then the height difference between the bald surface and Lidar measurements can be derived and the height points which are higher than some critical value are excluded. This critical value is calculated from the assumption that most natural Earth slope values are always lower than 25 degrees. These points, which are derived from this step will consist of the next level Bald Earth DEM seeds points and these are then interpolated using a higher tension factor and finer interpolation interval. The iteration continues until a termination criterion is reached.

2.3 Normalised DEM construction and ROI definition

Comparing the constructed Bald Earth DTM and height values, "above-ground" points are then selected. After "filling" gaps with a reasonable search area, all points are re-gridded using the highest tension factor (≈ 1.0). This step is useful for preserving

edges. Then by overlapping with the IKONOS NDVI, these normalised DEM patches can be divided into tree and building areas using any thresholding algorithm, because the NDVI distribution within "above-ground" areas have clearly two Gaussian distributions. We expect a better result by isolated point removal, such as morphological pruning, because if trees are removed close to a nearby building, the building boundary line will be distorted.

2.4 Correspondence with IKONOS image features

Building roofs in images are sometimes not very well matched with the corresponding DEM patches in the case of very tall buildings because of local height relief (Figure 2).

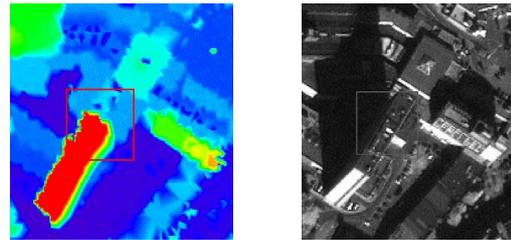


Figure 2. Planimetric shift due to local height relief

If an ortho-image is created using a Bald Earth DTM, this shift can be found without detailed camera information (Baltsavias et al., 2001) using the relationship

$$\Delta X_i = -\Delta Z \sin(a) / \tan(e), \Delta Y_i = -\Delta Z \cos(a) / \tan(e)$$

(where ΔZ : normalised height points-Bald earth height
a : sensor azimuth
e : sensor elevation)

Now, by merging a building area's normalised heights and roof features, the correct building structure can then be reconstructed.

3. OBJECT DETECTION IN IKONOS STEREO PAIRS

3.1 Positional and radiometric properties' check

Initially, the positioning accuracy of IKONOS was assessed by using GCPs from GPS measurements and Landline® plus data. In this case about 10 metres of planimetric shift was observed (Figure 3).



Figure 3 Error of ortho-image positioning error by IKONOS RPC (red: original, white: corrected)

Table 1 shows the positional error in the ortho-image measured using 24 static GPS measurements before and after updating the RPCs.

	X direction (metre)		Y direction (metre)		X,Y RMS (metre)	
	A*	B*	A	B	A	B
Average	9.733	0.002	-3.366	0.002	10.749	0.002
Std dev	2.449	3.352	3.264	3.791	2.614	5.061

Table 1 24 GPS points' positional error in the IKONOS ortho image which is produced by the initial (A) and updated (B) RPC parameters

These positional errors will be in WGS84 ellipsoid co-ordinates and are a serious problem for multi-data registration, which is one solution for object detection.

Another problem of IKONOS images are there radiometric quality (Figure 4). Edges are not clearly detected in IKONOS images. That's one reason why the correct ROI definition using 3D information is necessary.

3.2 Positional accuracy updating for DEM construction

Using a simple RPC analysis, we updated positional accuracy to within 3-4 m (Figure 5), so that a better registration between data sets is possible. However, not from Table 1 that although all the bias is removed, there is still some 3m of random error which we believe to be due to higher order RPC parameters not currently taken into account

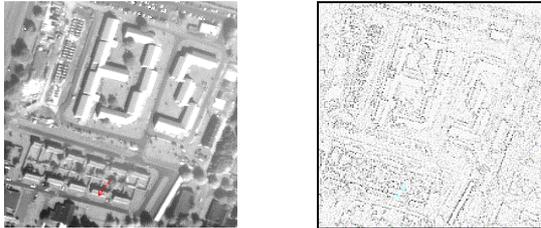


Figure 4 Edge detection in IKONOS image

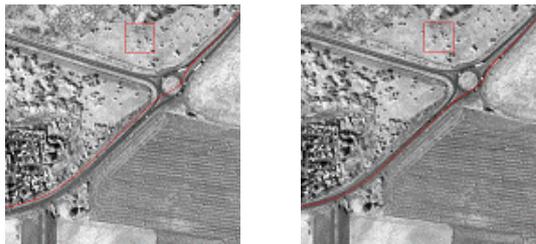


Figure 5 Positioning error removing (left : before updating, right : after updating)

The DEM (Figure 6) and ortho-images (Figure 7) were constructed with this updated RPC and our in-house P-Gotcha matcher (Kim & Muller, 1999). The 3 metres of vertical error which was previously reported by (Muller et al., 2001) is removed, so that a more accurate Bald Earth DTM construction is now possible.

3.3 ROI definition result

We have tested our ROI method using the previously described method. Because the target area is very flat, the region growing based method, and the multi-channel classification (4 pan-sharpened multispectral + DSM) was applied to the flat area

(Muller et al., 2001) with the height distribution based method. The Bald Earth construction result (Figure 8) appears reasonable but the ROI boundary is not correctly defined, because of the delineation problem with the building boundaries. The solution of this problem might be an improved image matcher, which can treat features and can address problems caused by high order RPC errors.

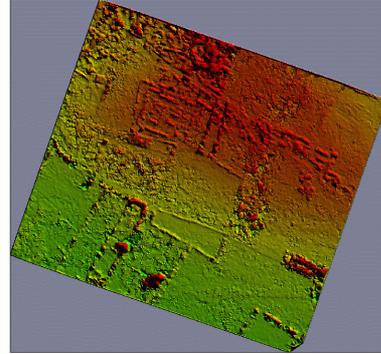


Figure 6 DEM construction result in BB test area



Figure 7 1 metre resolution IKONOS pan-sharpened multispectral ortho-image

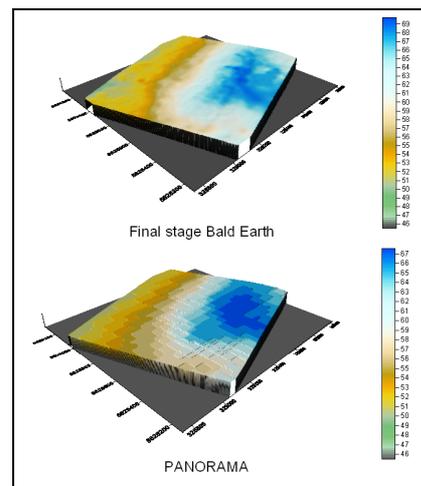


Figure 8 Bald Earth construction result compared to the OS® PANORAMA® DTM

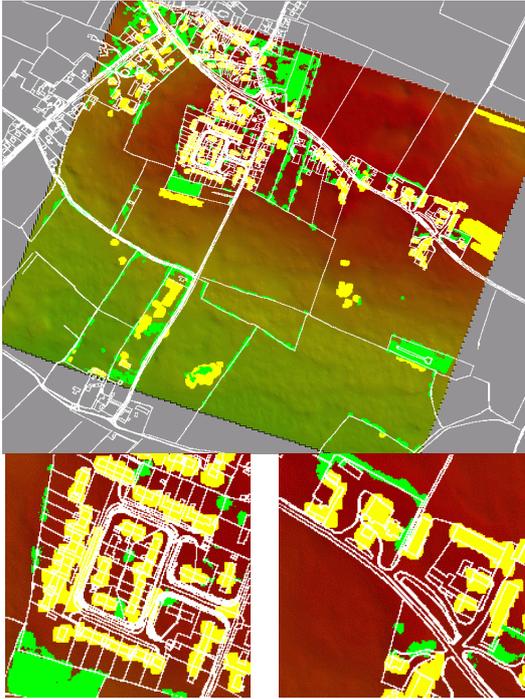


Figure 9. ROI definition and visual assessment with Landline@ plus

4. OBJECT DETECTION IN LIDAR & IKONOS PRECISION PRODUCT

4.1 Positioning

The East London (Precision) image was orthorectified with a 2-metre horizontal and 3-metre vertical accuracy using ground control points. In this case, no RPC exists. However, the horizontal and vertical accuracy using k-GPS points, was checked (Figure 10). As observed, the horizontal accuracy is within 2 metres as per the Space Imaging specification. Also the edge detection quality was assessed compared to Landline@ plus.



Figure 10. Example of Positioning and radiometric accuracy of IKONOS Precision by k-GPS tracks

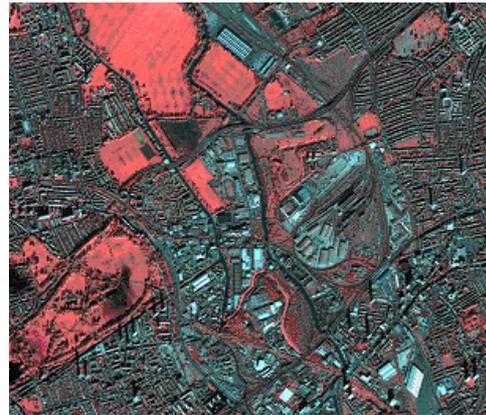
4.2 Bald Earth construction

The maximum estimated building size and target object area height were set as 100 metres and 4 metres respectively. So the moving window size and threshold value are set following these values. Figure 11 (a)-(d) shows the initial input data and results. This figure shows that the Bald Earth construction appears to be

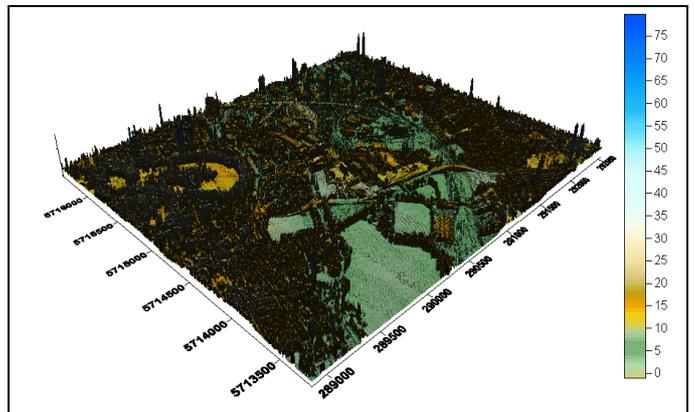
more accurate so that in the normalised DEM, the detailed roof structures over the buildings are now well observed.

4.3 ROI definition by thresholding and NDVI splitting

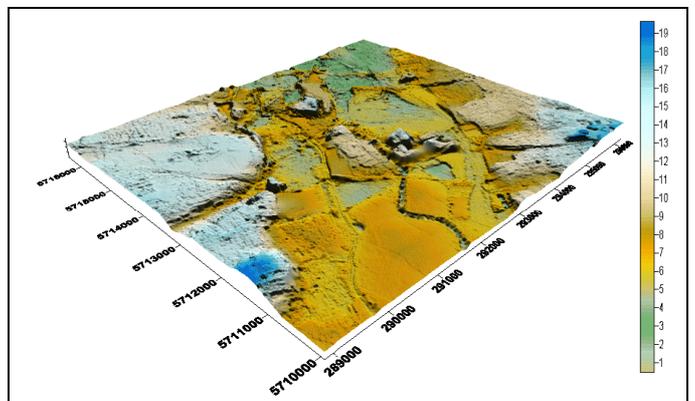
The tree and building ROI segmentation is produced using optimal thresholding and shows good agreements with the IKONOS image (Figure 12). If the normalised DEM is accurate enough, in other words, global slope effects are completely removed, a higher threshold value should give a clearer building boundary. We used a 4 metres threshold value. Because after segmenting trees around buildings, the boundary is distorted, these ROIs must be rectified by a morphological filter.



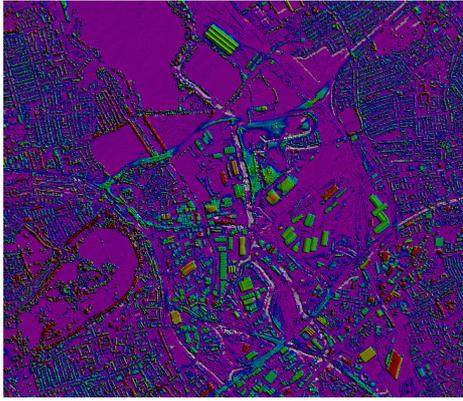
(a) IKONOS Precision image in test area (R:IR, G:Green, B:Red)



(b) Original Lidar DEM



(c) Final stage Bald Earth



(d) Re-constructed normalised DEM

Figure 11 Input Lidar data, final Bald Earth DTM and normalised DEM

4.4 ROI assessment

Landline® plus data was used to do a visual assessment of the building ROI. Even for very small buildings, it appears to detect buildings very well (Figure 13). The outlines of buildings can now be directly input to building reconstruction

4.5 Connection to the IKONOS image

The average normalised heights of each building shape can be extracted and applied to find out the correct roof location in the optical image. Now building structure reconstruction is feasible with both DEM and optical images.

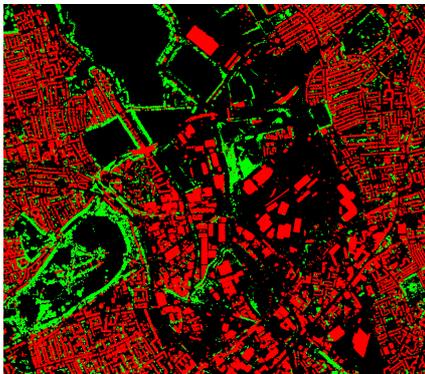


Figure 12 ROI definition (Red : building, Green : tree)

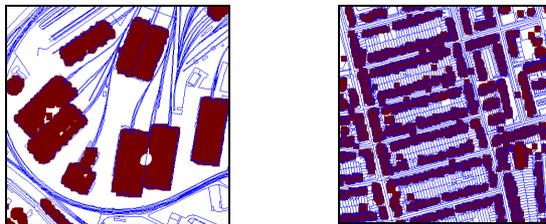


Figure 13 Assessment of ROI by Landline® plus

5. CONCLUSIONS

Because of the 1m resolution limit, the existence of strong shadows, the radiometric characteristics and the large search space, object detection in IKONOS image is very difficult to

achieve directly. We have developed a new technique to define the building and tree extents using 3D height points. It is robust and fast, if the DEM quality is good enough. Currently, the application of this method to IKONOS stereo DEMs is not completely successful, even if it is done using improved positioning information and a state-of-the-art image matcher (P-Gotcha), which was developed for geometrically distorted images. However, we showed that with 2 metre resolution Lidar data and IKONOS Precision images excellent results can be obtained mainly because of the good delineations of building outlines. Currently we are developing a new stereo DEM construction process, which uses an improved parallelised P-Gotcha (PVM Gotcha) algorithm, an unidirectional positioning error removal method and other matching data sources (colour, features). It is expected that using this scheme it will be feasible to extract GIS data from high resolution satellite stereo data including new high resolution satellite data such as the Quick-Bird satellite.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the financial assistance of the BNSC-LINK RISKMAP project especially the Project Manager at Infoterra Limited for permission to publish these results. In addition, the authors would like to thank the NASA Earth Science Enterprise project (under contract No. NAS13-98048) for supply of the stereo IKONOS imagery for Barton Bendish and Infoterra Limited for supply of the 2m lidar data over East London. Includes material which is © Space Imaging L.P and is Crown Copyright (Landline® digital vector map data and PANORAMA® DTM).

REFERENCES

- Baillard, C., Maite, H., 1999. 3D reconstruction of urban scenes from aerial stereo imagery, A focusing strategy, *Computer Vision and Image understanding*, 76(3), pp.244-258.
- Baltsavias, E., Pateraki, M. & Zhang, L., 2001. Radiometric and geometric evaluation of Ikonos Geo images and their use for 3D building modelling. *Proc. Joint ISPRS Workshop High Resolution Mapping from Space 2001*, Hannover, 19-21, CD ROM)
- Brenner, C., Haala, N., 1998, Fast production of virtual reality city models, in ISPRS Comm.IV Symposium, "GIS between visions and applications," Stuttgart, ISPRS, 32, part 4, pp.77-84.
- Halla, N., Brenner, C., 1999. Virtual city model from laser altimeter and 2d map data, *Photogrammetric Engineering & Remote Sensing*, 65(7), pp.787-795.
- Kilian, J., Halla, N., English, M., 1996, Capture and evaluation of airborne laser scanner data, *Int. Archive of Photogrammetry & Remote Sensing*, Vienna, V.XXXI, pt.B3., pp.383-388.
- Kim, T., Muller J.P., 1999, Development of a graph-based approach for building detection, *Image and Vision Computing* 17, pp. 3-14.
- Kittler, J., Illingworth, J., 1986, Minimum error thresholding *Pattern Recognition*, 19, pp. 41-47.

Kraus, K., Pfeifer N., 1998, Determination of terrain model in wooded areas with airborne laser scanner data, *ISPRS Journal of Photogrammetry & Remote Sensing*, 53, pp.193-203.

Muller J.P., Kim, J., Tong L., 2001, Automated mapping of surface roughness and landuse from simulated and spaceborne 1m data, *Automated Extraction of Man-Made Objects from Aerial and Space Images (III)*, E. Baltsavias, A. Gruen, L. Van Gool (Eds.), Balkema Publishers, Lisse, The Netherlands, pp. 369-380

Lu, Y., Kubik, K., Bennamoun, M., 1998, An accurate approach to localize house area for 3D terrain reconstruction, *Proc. of the 9th Australasian Remote sensing & Photogrammetry Conference, Sydney, Australia*,

Paskoand, M., Gruber, M., 1996, Fusion of 2D GIS data and aerial images for 3d building reconstruction, in XIX congress of ISPRS, Comm.III, *Int. Archives of Photogrammetry & Remote Sensing*, Vienna, 31, pp.257–260.

Roux, M. , 1997, Cooperative analysis of maps and aerial images for urban scene description ,in SPIE Europto : *Image and Signal Processing for Remote Sensing III*, London, 32 (17), pp.254–267.

Smith, W., Wessel, P., 1990, Gridding with continuous curvature splines in tension, *Geophysics*, 55(3), pp.293-305.

Tao, V., Wang Y., J. Mercer J., 2001, Automatic reconstruction of bald digital terrain models from digital surface models generated from airborne SAR system, The 3rd International Symposium on Mobile Mapping Technology. Cairo, Egypt, *CD-ROM*