3D building reconstruction from high-resolution Ikonos stereo-imagery

3D building reconstruction from high-resolution *Ikonos* stereo imagery

C.S. Fraser
Department of Geomatics, University of Melbourne, Melbourne, Australia

E. Baltsavias & A. Gruen
Institute of Geodesy & Photogrammetry, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland

ABSTRACT: An investigation of the application of 1m *Ikonos* satellite imagery to 3D building reconstruction is reported. The focus of the evaluation has been upon geopositioning accuracy, radiometric quality and attributes of the imagery that support building feature extraction. Following an initial review of characteristics of the *Ikonos* system, the multi-image dataset employed is described, along with the Melbourne *Ikonos* testfield. Radiometric and image preprocessing aspects are discussed and the results of 2D and 3D metric accuracy tests are presented. A quantitative and qualitative assessment of the use of stereo *Ikonos* imagery for generating building models is then provided, using the campus of the University of Melbourne as an evaluation site. The results of this assessment are summarised, these highlighting both the submetre metric accuracy potential of *Ikonos* imagery and difficulties to be anticipated in building reconstruction when a comprehensive and detailed modelling is required.

1 INTRODUCTION

Although it has now been close to 18 months since *Ikonos* imagery first became commercially available, there have been only a few accounts presented of photogrammetric analysis of this new image data. A number of factors have contributed to the current paucity of research outcomes related to the geometric and radiometric quality of 1m panchromatic, 4m multispectral and 1m 'pan-sharpened' *Ikonos* imagery. These include significant early delays in image delivery, restrictions regarding the provision of stereo imagery and ongoing improvements to the image preprocessing phase. Also, it is fair to say that teething problems were encountered with what is a commercially new technology being made available to the user community under a distinctly new business model.

Given the present shortage of information on the photogrammetric performance of the *Ikonos* system, it has been necessary to examine essentially three salient aspects when evaluating the use of 1m stereo imagery for 3D city modelling. These comprise the geometric accuracy of geopositioning from stereo and multi-image coverage; the radiometric quality, with an emphasis on characteristics to support automatic feature extraction (e.g. noise content, edge quality and contrast); and attributes of the imagery for the special application of building extraction and visual reconstruction. In the present paper, we examine these aspects with the aid of three-fold *Ikonos* coverage of a precisely surveyed testfield covering the city of Melbourne. We first describe this testfield dataset and then examine radiometric and image preprocessing aspects. This is followed by an evaluation of the metric potential of *Ikonos* mono, stereo and three-image coverage, along with an account of experimental application of stereo imagery to building reconstruction, including topology building and visualisation.

Prior to proceeding further, however, it is useful to give a brief account of the design parameters of the *Ikonos* imaging system and to clarify specifically which image products are being employed and what general level of metric precision is sought. As regards details of the sen-

sor and mission parameters, these are summarised in Table 1. Visitors to Space Imaging's (SI's) web site (www.spaceimaging.com) will note that there are basically five product options for panchromatic and pan-sharpened 1m *Ikonos* imagery: *Geo*, *Reference*, *Pro*, *Precision* and *Precision Plus*. Except for the *Geo* product, all are ortho-rectified, with ground control being required for *Precision* ortho-imagery. The absolute geopositioning accuracies (RMS, 1-sigma) associated with these categories of imagery are 24m, 12m, 5m, 2m and 1m, respectively.

Table 1. Sensor and mission parameters for the Ikonos-2 satellite.

Focal length of steerable linear array	10 m
Orbit altitude and period	680 km & 98 minutes
Nominal image scale	1:68,000
Array length & pixel size	13,800 & 12 μm PAN
Field of view	0.93^0
Footprint in panchromatic (nadir)	0.82m, resampled to 1m
Footprint in multispectral (nadir)	3.3m, resampled to 4m
Number of bits	11
Swath width at nadir	11 km
Nominal scene size	11 x 11 km
Revisit and repeat-cycle times	1-3 & 14 days (max)
Stereo imaging: along-track, with cross-track possible	

Depending on the intended function of a digital city model, metric accuracy expectations will vary, though if the very sizeable market of modelling for mobile communications is to be accommodated, accuracies at the metre level are generally sought. The clear implication from the product offerings listed above, however, is that users of *Ikonos* imagery who seek metre-level accuracy will have to resort to acquisition of expensive *Precision* imagery. We have pursued an alternative option, with the aim of yielding high accuracy results, to metre or sub-metre level, using the least expensive *Ikonos* data, namely *Geo* images, coupled with alternative computational schemes for mono-, stereo- and multi-image positioning. A few accounts of *Geo* imagery being employed for ortho-image production to metre-level accuracies have been recently published (e.g. Kersten et al. 2000, Gruen 2001, Davis & Wang 2001), though the only results to date that the authors are aware of involving 3D geopositioning from stereo triangulation have emanated from SI (Dial 2000). The reported stereo results were, perhaps not surprisingly, both encouraging and consistent with *Ikonos* product specifications. One can imagine a number of disincentives for SI to promote the use of *Geo* imagery for higher accuracy applications.

2 INPUT DATA

2.1 *Image data*

The imagery comprised a stereopair of epipolar-resampled *Geo* PAN images, and a nadir-looking scene of PAN and multispectral imagery. The latter were also combined to a pan-sharpened image. As indicated in Table 2, the sensor and sun elevation angles for the stereopair (imaged in winter) were less than optimal. Apart from the right stereo image, the azimuths of sensor and sun differed considerably, leading to strong shadows in non-occluded areas.

Table 2. Acquisition parameters of the original 1m *Ikonos* PAN images.

	Left Stereo	Right Stereo	Nadir
Date, Time (local)	16/7/2000, 09:53	16/7/2000, 09:53	23/3/2000, 09:58
Sensor azimuth (deg)	136.7	71.9	143.0
Sensor elevation (deg)	61.4	60.7	83.4
Sun azimuth (deg)	38.2	38.3	50.0
Sun elevation (deg)	21.1	21.0	38.0

The overall geometry was close to that of 3-line imagery, with the base/height ratio of the stereopair being B/H = 1.2. Supplied with the stereo imagery were coefficients of the rational functions (RFCs) which provide a mechanism for object-to-image space transformation and 3D point determination. There were no RFCs for the near-nadir image, since the option of obtaining these coefficients for mono *Geo* imagery was not available at the time the data was acquired.

2.2 *Melbourne* Ikonos *testfield*

Covering an area of 7 x 7km over central Melbourne, the testfield comprises an array of GPS-surveyed ground control points (GCPs) of which 32 are road roundabouts. As mentioned, feature extraction accuracies to metre level were sought from the stereo *Ikonos* data and thus there was a need to measure image-identifiable ground points to sub-pixel accuracy within the imagery, as well as to sub-metre precision on the ground. Road roundabouts were selected as control points since they constituted an elliptical target which was usually well contrasted against its background (the surrounding road) within the image, and which was typically 8 to 25 pixels in diameter. Within the imagery, it was the centroid of the roundabout which was measured and so the GPS survey needed also to determine the centre point coordinates.

As described in Hanley & Fraser (2001), the centroids of each of the 32 roundabouts were determined by measuring six or more edge points around the circumference of the feature, both in the image and on the ground, and then employing a best-fitting ellipse, computed by least-squares, to determine the ellipse centres (see Fig. 1). By this means, it was possible to achieve accuracies of image and object point determination, in 2D and 3D space, respectively, to better than 0.2 pixels. Figure 1 shows one of the roundabouts as recorded in the nadir-looking *Ikonos* image, along with a best-fitting ellipse to edge points of the same roundabout.

To complement the image mensuration via ellipse fitting, least-squares template matching was also employed within the stereo imagery, with special care being taken to alleviate problems such as target occlusions from shadowing and the presence of artifacts such as cars (e.g. Fig. 1). The matching utilised gradient images instead of grey values and the observational weights were determined from the template gradients, i.e. only pixels along the circular template edge were used in matching, thus making the method insensitive to contrast differences and other disturbances. Three different template sizes were used to accommodate the greatly varying roundabout sizes and in most cases an affine or conformal transformation was used.

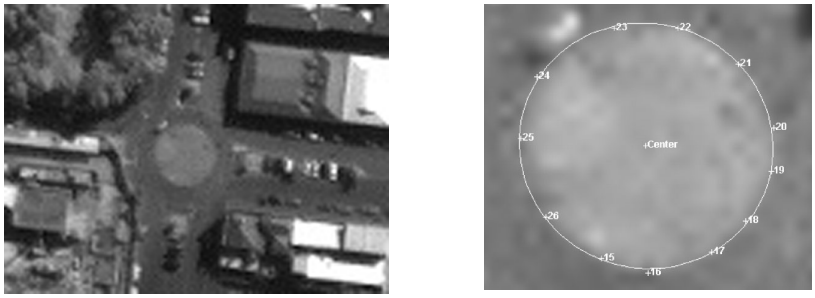

Figure 1. Left *Ikonos* image of roundabout and right recorded edge points and best-fitting ellipse.

In spite of significant disturbances, and size and shape variations of roundabouts, the matching performed correctly in almost all cases. Since the same template was matched simultaneously using the same parameters in all images, the image measurements were mutually quite consistent. This means that parallax errors were reduced and height estimation accuracy could be expected to surpass that of the ellipse fitting method. The images used for measurement were initially preprocessed for noise reduction and edge enhancement, as described in the following section. Two examples of modified least squares matching are indicated in Figure 2. In each, the top row shows the template and the two stereo images with the starting (light) and final (dark) position after a conformal transformation. The bottom row shows the matched gradients. The

examples illustrate the robustness regarding disturbances along the roundabout edge and outside, and in the presence of contrast variations.

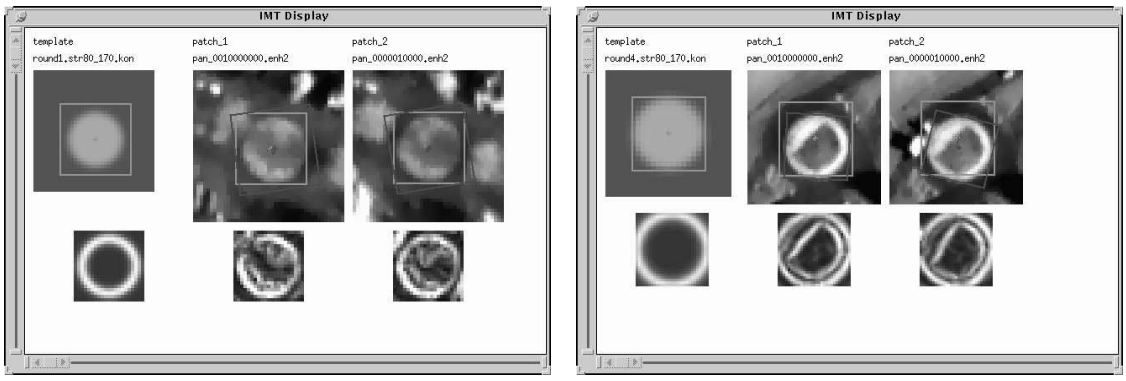

Figure 2. Two examples of modified least squares matching.

A second essential component of the Melbourne testfield in the context of city modelling comprised information on buildings. Due to both time constraints and the unavailability of accurate height data for central city high-rise buildings, many over 30 stories, it was decided that the campus of the University of Melbourne would serve as the test site for the building reconstruction and visualisation phase of the project. The campus comprises a few dozen major buildings which vary in design characteristics, size, shape, roof superstructure and height, though most are between four and 10 stories tall. Building height data for 19 image-identifiable and readily accessible roof corners was provided through precise GPS surveys, to 10cm accuracy. This data formed the metric standard by which the 3D triangulation for building extraction would be assessed. Image coordinate observations to these feature points were carried out by manual recording, both in stereo and monoscopically (for the nadir image), nominally to 0.5-1 pixel precision. However, the definition of corner points was often weak, leading to the possibility of significant errors. Least-squares matching was not appropriate for the recording of such corner features due to both occlusions and the geometric differences between the images which could not be modelled through affine transformation.

An existing stereopair of 1:15,000 scale wide-angle colour aerial photography of the campus was employed in the evaluation of the building extraction potential of *Ikonos* imagery. This imagery had previously been used to create a reasonably detailed 3D model of the campus, which could be compared to that derived from *Ikonos* data in terms of recoverable feature content.

3 RADIOMETRIC ASPECTS AND IMAGE PREPROCESSING

3.1 *Radiometric quality analysis*

Prior to discussing radiometric features of the *Ikonos* imagery, a subject which is discussed further in Baltsavias et al. (2001), it is worthy of note that operational aspects of the image acquisition are likely to have a more profound effect on the homogeneity, or non-homogeneity, of image quality than specific radiometric characteristics of the sensor system. For example, large changes in image quality and suitability for automated feature extraction are associated with variations in the sensor view angle, the sun angle and shadowing, the seasons, atmospheric conditions, and whether the scene is recorded in mono or stereo. These influences are well known, but it needs to be appreciated that with the exception of the last aspect, they are largely beyond the control of the image user. There is limited opportunity for the user to dictate specific imaging dates, times and weather conditions. Of the three images employed, the near-nadir image was superior in terms of both contrast and visual resolution. This can be explained by the higher sensor and sun elevation, though it is uncertain if this fact alone accounted for the modest dif-

ference in image quality, or whether differences were in addition due to changes in atmospheric conditions or aspects of the epipolar resampling of the stereo images.

All images were preprocessed by SI with Modulation Transfer Function Compensation (MTFC) but no Dynamic Range Adjustment (DRA). Although the images were delivered as 11-bit data, the number of grey values having a substantial frequency was much less than 2048. For the left, right and nadir PAN images, grey values with a frequency of more than 0.01% were covering only the ranges of 44-343, 56-423 and 61-589, respectively. The corresponding effective grey value ranges of 299, 367 and 528 (37, 46 and 66 grey values for a linear stretch to 8-bit) are quite similar to the effective intensity range observed with other spaceborne linear array CCD sensors such as SPOT and MOMS. The peak of the histogram is typically towards the darker values (at 62, 73 and 82 for left, right and nadir, respectively) with the right part of the histogram decreasing smoothly and slowly towards the higher grey values.

The overall noise characteristics of the images were analysed in both homogeneous (sea surface) and nonhomogeneous areas (e.g. city) by a procedure explained in Baltsavias et al. (2001). The use of nonhomogeneous areas in image noise evaluation is justified as large homogeneous areas do not always exist, plus this allows an analysis of the noise variation as a function of intensity, as noise is not additive but intensity-dependent. This analysis revealed the presence of somewhat more noise (2-14 grey values, intensity-dependent) than is suggested by SI investigations (4 grey values, see Gerlach 2000), though its extent was not judged to be of crucial significance for the specific task of building reconstruction.

Geo images have been found to exhibit artifacts in addition to their noise content, of which a number remain unexplained. Many are visible only in homogeneous areas, especially after contrast enhancement, and some appear in areas of rich texture. Although the artifacts often lead to small grey level variations (e.g. 2-4 DN in 11-bit images), after contrast enhancement, which is often required for visual interpretation and measurement or automated computer processing, they can lead to erroneous high-contrast texture patterns. Apart from chess pattern noise and dark or bright stripes in the flight direction, which are typical of linear array CCD spaceborne sensors, the following concerns regarding artifacts have been identified: some striping in the stereo images in the flight direction, an apparent MTFC effect similar to the unsharp masking used to sharpen edges, a staircase variation of the grey values within homogeneous areas of the epipolar-resampled stereo images, jumps in grey value (visible in homogeneous areas) across the flight direction, the ghosting of moving objects, and some striping normal to the flight direction.

These radiometric concerns are not related solely to the sensor imaging parameters but also to the subsequent image processing methods and to image compression. The compression to 2.6 bits leads to some artifacts, which are more visible in homogeneous areas. Some additional problems include shadows and image saturation. The shadow areas in the Melbourne imagery did not have a significant signal variation, and thus in spite of a strong contrast enhancement, feature details in shadow areas were often hardly visible. Saturation occurred routinely, with bright vertical walls, especially those approximately in the middle of the illumination-to-sensor angle. This led again to a loss of detail and poor definition or disappearance of edges when two saturated walls were intersecting.

Shortcomings in edge definition, partly inherent in 1m imagery, and problems with variability in image quality, often lead to instances where buildings and trees, for example, of nominally sufficient size are almost impossible to identify. Figure 3 illustrates the large differences in object definition quality that can occur between two *Geo* images. The lefthand image is 1m PAN of Lucerne, Switzerland, while that on the right, of Nisyros in Greece, is from the red channel (corresponding to NIR) of a pan-sharpened image. Note also the artifact (bright, nearly vertical line) at the centre of the lefthand image.

3.2 *Image preprocessing*

In order to reduce the effects of the above mentioned radiometric problems and optimise the images for subsequent computer processing (the measurement of GCPs and buildings, and in other *Ikonos* projects orthoimage generation and mapping), various preprocessing methods were developed, implemented and tested, the full details of which are given in Baltsavias et al. (2001). Figures 4 and 5 illustrate the impact of the adopted image processing methods. The first con-

sisted of an anisotropic Gaussian low-pass filtering to reduce noise and stripes in the flight direction (which for non-stereo images are almost vertical), coupled with application of a Wallis filter for contrast enhancement (Fig. 4 middle). The filter anisotropy was aimed at filtering more in the line direction in order to reduce vertical striping.

In a second preprocessing step, after the low-pass filtering, an unbiased anisotropic diffusion was used to further reduce noise and sharpen edges (Fig. 5). The improvement of the Wallis filter in shadow regions was not as much as previously achieved for aerial and other satellite imagery, possibly due to a low signal variation in these regions. Both these approaches were applied to 8-bit data that were generated by a linear compression of the original 11-bit imagery.

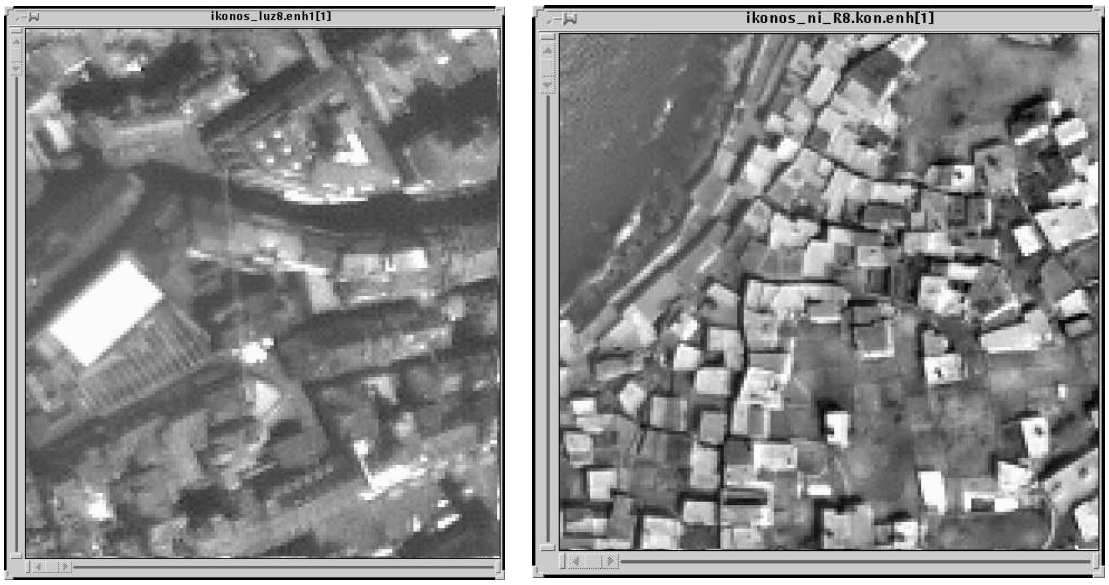

Figure 3. Large differences in object definition quality between *Geo* images of Lucerne, Switzerland and Nisyros, Greece.

Later, a third, improved approach was adopted. All preprocessing was applied to the 11-bit data. First, two adaptive local filters were developed. They reduced noise, while sharpening edges and preserving even fine detail such as one-pixel wide lines, corners and line end-points. Optionally, salt-and-pepper noise can be eliminated to a certain extent. The effect of the two local filters is generally quite similar, although they use a different number and size of masks, and one employs a fuzzy method. They require as input an estimate of the noise, which may be known or estimated by the methods mentioned above.

Figure 4. Original image (left), after first preprocessing (middle), and after final preprocessing (right).

Next, a new version of the Wallis filter, which estimates automatically some of the filter parameters, is applied, and finally a reduction to 8-bit imagery by histogram equalisation is performed (Figs 4 & 5). The histogram equalisation was iterative with the aim being to occupy all 8 bits with similar frequencies. Processing in 11-bit led to slightly better 8-bit images than the

first two preprocessing methods, even though the 11-bit imagery was effectively only 8 to 9 bits. The histogram equalisation, which is optimal for general computer processing but may lead to strong bright and dark regions for visual interpretation, could be replaced by an alternative reduction of either Gaussian type or through an optimisation of a selected target grey level range, for example.

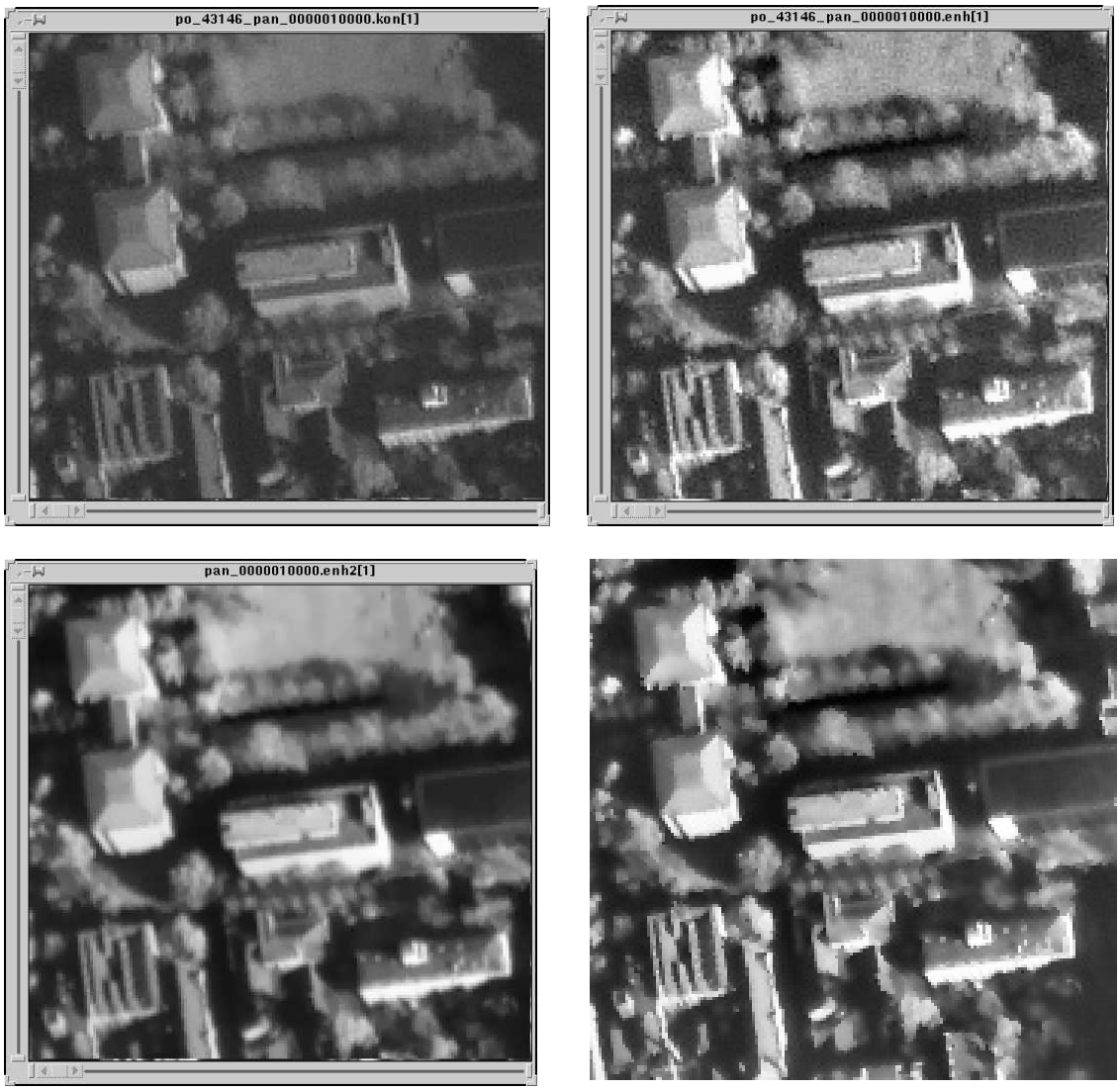

Figure 5. Top: original image (left) and after Wallis filtering to indicate noise content (right). Bottom: preprocessing by the second (left) and third (right) methods.

4 METRIC QUALITY

4.1 *Accuracy potential*

The recovery of 3D cartographic information from satellite line scanner imagery has been the subject of photogrammetric investigation for the last decade and a half. Mathematical models have been formulated to support triangulation of cross-track *SPOT* and *IRS-1C/D* imagery, as well as *MOMS-02* 3-line imagery which is an along-track configuration with similarities to the *Ikonos* coverage of the Melbourne testfield. Fully rigorous mathematical models for orientation and triangulation, which have in turn implied provision of sensor calibration data and, to a degree, prior information on the satellite orbit and sensor attitude data, have been a prerequisite for these developments. In summary, it could be said that under ideal conditions of high-quality im-

age mensuration and ground control/checkpoints, coupled with favourable imaging geometry (e.g. B/H > 0.8) and provision of sensor calibration data, ground point determination to 0.3 pixel accuracy is possible (Ebner et al. 1996) with medium-resolution stereo satellite imagery. However, accuracies of between 0.5 and 2 pixels are more commonly encountered in practical tests.

Through a simple extrapolation of these findings to 1m-resolution imagery, we might anticipate that under similar operational constraints sub-metre 3D geopositioning accuracy should be obtainable from stereo *Ikonos* panchromatic and pan-sharpened image data. Consider, for example, a stereo *Ikonos* configuration similar to that covering the Melbourne testfield. Under the assumption of an image measurement accuracy of 0.4 pixel (σ_{xy} = 5μm), the ground point triangulation precision to be anticipated for the stereo configuration is σ_{XY} = 0.3m (planimetry) and σ_Z = 0.7m (height). If this stereo geometry is extended to three along-track images (addition of a nadir-looking image), again as we have for the Melbourne testfield, the triangulation precision in Z remains unchanged (see also Ebner et al. 1992), whereas the planimetric precision is improved to σ_{XY} = 0.25m, or ¼ of the ground sample distance. In order to apply familiar collinearity-based models in this triangulation process, a camera model comprising sensor calibration data must be available. Moreover, for the stereo image restitution, prior knowledge of the satellite exterior orientation is required, though this requirement can be relaxed in the 3-image triangulation.

4.2 *Sensor orientation models*

At the present time, however, the option of applying rigorous, collinearity-based image orientation and triangulation models to *Ikonos* data is effectively denied to the end-user of this imagery, since the necessary camera model (calibration data) and precise ephemeris data for the satellite are withheld as proprietary information. Alternative orientation/triangulation models are required, with the Rational Function Model being currently supported (Grodecki 2001, Yang 2000). The 80 rational polynomial coefficients (RFCs) per image for the object-to-image (latitude, longitude, height to image line, sample coordinates) transformation function can be expected to yield absolute accuracies in subsequent 3D object space positioning which are consistent with the specifications for the particular class of *Ikonos* product, i.e. *Geo*, *Precision*, etc. RFCs for the *Geo* product, which are expected to produce an RMS positioning accuracy of about 25m, are derived solely from satellite ephemeris and attitude data, whereas those for *Precision* products are computed with the additional aid of ground control. As will be demonstrated, there is considerable scope for enhancing the absolute accuracy of XYZ data obtained via RFCs through post transformation of the object point coordinates to control points. This transformation may in the simplest case be a coordinate translation or it may comprise a translation and rotation.

Beyond rational functions, there are further models displaying a favourable level of sensor independence, which can be employed in order to overcome the imposed absolute accuracy limitations of RFCs. Potential approaches which take account to varying degrees of the mixed projective/parallel projection of *Ikonos* include the Direct Linear Transformation (DLT), either in extended form (Wang 1999) or as piece-wise functions (Yang 2001), and affine projection (Okamoto et al. 1999, Fraser 2000). The former is fundamentally a projective model, whereas the latter is better suited to *Ikonos'* parallel along-track projection. The very small field of view of the sensor of only 0.93^0 leads to the overall projection approaching skew parallel, especially over smaller sub-scene areas. As it happens, both these models may be thought of as linear rational functions, though in the case of the affine model a perspective-to-affine transformation of the image is first warranted in stereo scenes containing moderate to high relief variation (Okamoto et al. 1999). The object-to-image transformation functions for the extended DLT and affine models are given, respectively, as:

$$x = (L_1X + L_2Y + L_3Z + L_4) / (L_9X + L_{10}Y + L_{11}Z + 1) + L_{12}\,x\,y \quad (1)$$

$$y = (L_5X + L_6Y + L_7Z + L_8) / (L_9X + L_{10}Y + L_{11}Z + 1)$$

and

$$x = A_1X + A_2Y + A_3Z + A_4 \quad (2)$$

$$y = A_5X + A_6Y + A_7Z + A_8$$

where x,y are image space coordinates in flight and CCD direction respectively and X,Y,Z are Cartesian object space coordinates. Additional parameters can be appended to these expressions to account for non-linearities when projection or geographic coordinates are employed. As linear approximations to higher-order rational functions, the DLT and affine models can be expected to show limitations, especially as the area covered by the *Ikonos* scene becomes large. As will be discussed, however, in the 7 x 7km Melbourne testfield both models yielded sub-metre 3D positioning accuracy when applied with as few as six GCPs. For both the 2- and 3-image orientation/triangulation with the DLT and affine models, all parameters where determined simultaneously. Further details are provided in Fraser et al. (2001).

4.3 *Planimetric positioning*

In many respects a prerequisite to the application of approaches such as affine projection and the DLT is that the imagery itself has to be of high metric integrity. It should be recalled that the *Geo* product is "a geometrically corrected product that has been rectified to a pre-specified ellipsoid and map projection" (Space Imaging 2001). Moreover, the imagery is re-sampled by cubic convolution to 1m pixels and in the case of a stereopair it is epipolar resampled. As a means of both ascertaining the planimetric positioning accuracy of 1m *Ikonos* imagery and of evaluating 'sensor linearity', a number of 2D transformations from image to object space were performed. The aim was to examine the accuracy in XY coordinates resulting from mapping the image data to height-corrected 'planes of control' using various ground GCP configurations.

The results of this 2D accuracy analysis, which are reported in Hanley & Fraser (2001), were very encouraging. For three configurations of six GCPs and 20-25 independent checkpoints, RMS positioning accuracies of 0.3-0.5m were obtained. When a best-fit mapping was carried out using all 32 GCPs, the XY discrepancies displayed RMS values ranging from 0.27m to 0.38m, i.e. from about 0.3 to 0.4 pixels. In the absence of any camera model information, it was felt that this straightforward analysis gave a strong indication of the metric integrity of *Ikonos Geo* imagery, which augured well for the use of linear models in 3D geopositioning. Moreover, the tests demonstrated that in the presence of good quality control and DTM data (or moderately flat terrain) *Ikonos* imagery readily yields XY positioning accuracies to 0.5 pixels and better.

4.4 *3D positioning of roundabouts*

4.4.1 *Rational functions*

As a first stop in analysing the triangulation accuracy of *Ikonos* stereo images, 3D coordinates for the GPS-surveyed roundabouts were computed using the supplied RFCs, for the cases of both manually measured image coordinates (via ellipse fitting) and least-squares template matching. The resulting RMS discrepancy in XYZ ground points coordinates was close to 20m in each image of the stereopair, which was consistent with *Geo* specifications. The presence of a positioning bias was clearly apparent, however, and a simple coordinate translation using a few control points reduced the RMS discrepancy at remaining checkpoints to between 1.5 and 2m. Use of six GCPs, coupled with 3D similarity transformation, reduced the RMS checkpoint discrepancies further, to about 0.5m in planimetry and height. In regard to a comparison of results obtained from least-squares matching and ellipse fitting, there was no significant distinction in accuracy, though the image coordinates obtained from matching yielded more consistent results in height determination.

Although the evaluation of geopositioning via RFCs is ongoing, results obtained to date within the Melbourne *Ikonos* testfield confirm that high relative accuracies are attainable with *Geo* RFCs in spite of their rather modest absolute accuracy specification. Provision of only one or two GCPs facilitated removal of positioning biases to yield 1-2m level 3D positioning, whereas sub-pixel accuracies were achieved with additional GCPs.

4.4.2 *DLT and affine models*

With the DLT and affine orientation/triangulation models it is a straightforward matter to extend the analysis of stereo *Ikonos* imagery to an evaluation of 3-image configurations. A series of 'bundle adjustments' were computed using both the extended DLT (Eq. 1) and the affine model

(Eq. 2) with and without additional parameters. A full account of the results is beyond the scope of the present paper and readers are referred to Fraser et al. (2001) for further details. The results can, however, be summarised as follows: For control configurations of 4 to 8 GCPs and 20-25 checkpoints, the affine model produced RMS geopositioning accuracies of 0.35-0.5m in XY and 0.5-0.8m in Z, for both stereo and 3-image geometries. For the same GCP sets, the DLT yielded slightly lower triangulation precision, by around 0.1 pixel, and it was found to exhibit stability problems for certain GCP configurations. The fact that first-order models can readily yield 0.5 pixel positioning accuracy is very encouraging for both metre-level city modelling applications and for GIS mapping for local government planning and management, which requires nominal 1-sigma precision of 1.5-2m (e.g. Davis & Wang 2001).

4.5 *Accuracy of building extraction*

Whereas the foregoing 3D ground point determination was centred upon 'high-quality' targets, accurate positioning of building features, generally corners and edges, involves not only metric factors but issues of image resolution and feature identification. The approach adopted in the reported investigation to ascertaining the metric quality of building extraction in stereo *Ikonos* imagery again involved independent checks of photogrammetrically triangulated distinct points against their GPS-surveyed positions. As mentioned, 19 image-identifiable roof corners were precisely surveyed to serve as accuracy checkpoints in the building extraction phase.

Within the stereo triangulation, the RFCs produced RMS accuracies of 0.7m in planimetry and 0.9m in height after removal of the bias in object space using the known GCP coordinates. Corresponding accuracy estimates resulting from the 19 checkpoints for the affine and DLT models with 6 GCPs were 0.7m and 0.6m, and 1.0m and 0.8m, respectively, for planimetry and height. Triangulation of the 3-fold image coverage produced results which were not significantly different than those for the stereo restitution. Whereas in practice it may be unlikely that a 3-image coverage would be employed for building extraction, provision of the near-nadir image can prove very useful for subsequent ortho-image generation of built-up urban areas.

5 EVALUATION OF BUILDING EXTRACTION FROM *IKONOS*

In order to provide a qualitative and quantitative assessment of the potential of stereo *Ikonos* imagery for the generation of building models, the University of Melbourne campus was measured manually in stereo with both an in-house developed software tool for the *Ikonos* stereo images, and using an analytical plotter for the 1:15,000 colour aerial imagery. The resulting plots of extracted building features are shown in Figures 6 & 7. The manual measurements of roof corners and points of detail were topologically structured automatically using the software package CC-Modeler and also visualised (Fig. 7).

This process revealed that many building points could neither be identified nor subsequently measured in the *Ikonos* images. Moreover in a number of cases, buildings could only be coarsely generalised (Fig. 8). Measurement and interpretation in stereo is a considerable advantage, as is the use of colour which was unfortunately not available in the *Ikonos* stereopair. Other influencial factors are shadows, occlusions, edge definition (related also to noise and artifacts), saturation of bright surfaces, sun and sensor elevation and azimuth, and atmospheric conditions. The 1m resolution of *Ikonos* also leads to certain interpretation restrictions, as is illustrated in Figure 9. This investigation was aimed in part at quantifying the extent of such limitations.

A comparison of the two models in Figure 6, one from aerial photography and the other from *Ikonos* 1m imagery, revealed the following regarding the *Ikonos* stereo feature extraction: about 15% of the building area as measured in aerial images could not be modelled; a number of both small and large buildings could not be identified and measured; some new buildings could however be reconstructed, even if small; and buildings could often be only generalised with a simplified roof structure and variations to their form and size.

Figure 6. Buildings of University of Melbourne campus reconstructed from 1:15,000 aerial images (left) and from stereo *Ikonos* imagery (right). To simplify visualisation, first points and first lines have been omitted in the left figure.

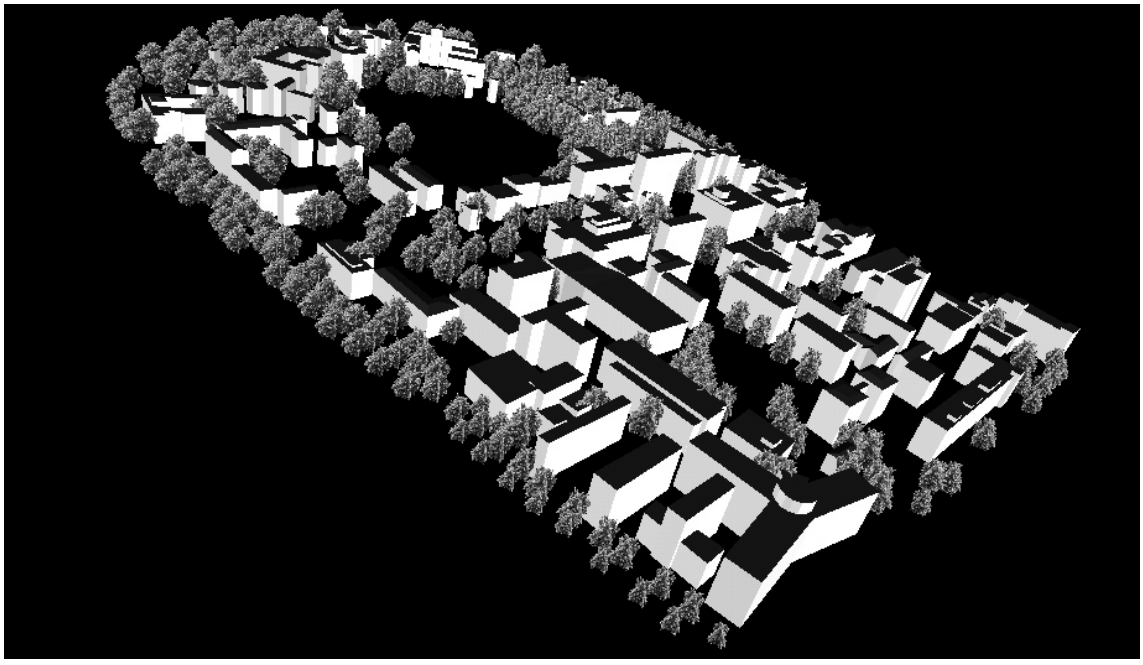

Figure 7. Visualisation of University of Melbourne campus with buildings and trees extracted from stereo *Ikonos* images.

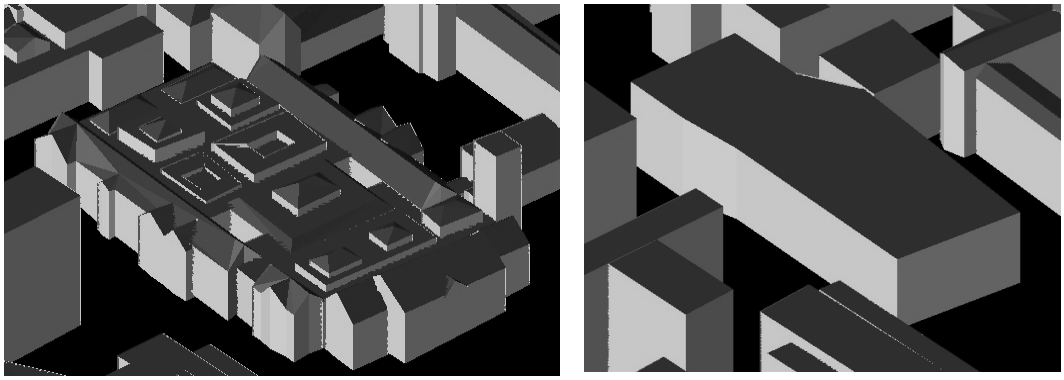

Figure 8. Building with complicated roof structure as measured from aerial images (left) and *Ikonos* images (right).

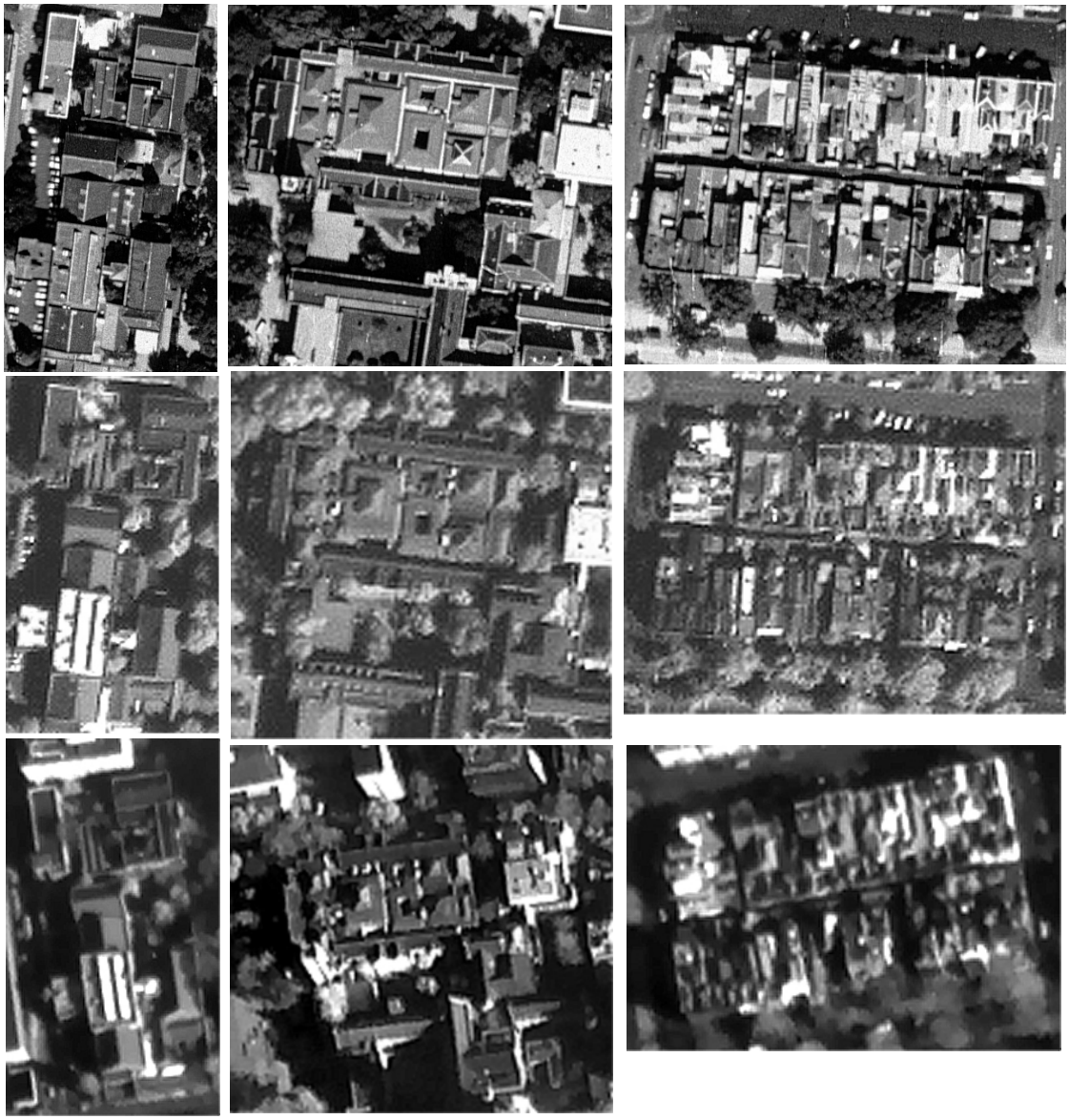

Figure 9. From top to bottom: buildings in aerial images (pixel footprint 0.3 m), nadir pan-sharpened and one of the stereo images. From left to right: buildings interpreted and measured in stereo *Ikonos* images rather well, only to a certain extent, and almost not at all.

6 CONCLUSIONS

This investigation has shown that *Ikonos Geo* stereo imagery has the potential to yield 1m geopositioning accuracies and better. These results are impressive in the context of building reconstruction, but they must be weighed against some notable shortcomings of 1m satellite imagery for building feature detection and identification. While the limitations of an image resolution of 1m are well understood, it is more the variability of image quality from scene to scene that will limit application to building reconstruction and city modelling. The influence of factors which are largely beyond the current control of the image user, such as date and time of image collection, specification of favourable sun angles and atmospheric conditions, etc. will likely generate difficulties if one is seeking to exploit the imagery to its full potential. We have witnessed in this investigation a significant difference in radiometric quality between the stereo and near-nadir *Geo* imagery, a difference that has increased the difficulty of comprehensively describing the stereo scene. Moreover, problems have been encountered not only in identifying all buildings of a certain size, but also in accurately reconstructing their form without excessive generalisation. The shortcomings in radiometric quality, though not of major concern in this investigation, do give rise to problems of accuracy and interpretability in both manual and computer mensuration. The issues of radiometric inhomogeneity both within and between *Ikonos* scenes are unlikely to be affected by the level of image product purchased and this investigation has demonstrated that very high metric performance is achievable with the least expensive product offering, namely *Geo* imagery.

ACKNOWLEGEMENTS

The computational and general research support of Harry Hanley and Takeshi Yamakawa at the University of Melbourne, and Maria Pateraki and Li Zhang at ETH Zurich in this collaborative project is gratefully acknowledged. The *Ikonos* image of Lucerne was kindly provided by the National Point of Contact, Swiss Federal Office of Topography, Wabern, while the *Ikonos* image of Nisyros was made available by the EU project Geowarn.

REFERENCES

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 September 2001* (accepted for publication) (on CD ROM).

Davis, C.H. & Wang, W. 2001. Planimetric accuracy of Ikonos 1-m panchromatic image products. *Proc. ASPRS Annual Conference, St Louis, 23-27 April 2001*, 14 p. (on CD ROM).

Dial, G. 2000. Ikonos satellite mapping accuracy. *Proc. ASPRS Annual Conference, Washington D.C., 22-26 May 2000*, 8 p. (on CD ROM).

Ebner, H., Kornus, W. & Ohlhof, T. 1992. A simulation study on point determination for the MOMS-02/D2 space project using an extended functional model. *International Archives of Photogrammetry and Remote Sensing*, vol. 29, part B4, 458-464.

Ebner, H., Ohlhof, T. & Putz, E. 1996. Orientation of MOMS-02/D2 and MOMS-2P imagery. *International Archives of Photogrammetry and Remote Sensing*, vol. 3, part B3, 158-164.

Fraser, C.S. 2000. High-resolution satellite imagery: a review of metric aspects. *International Archives of Photogrammetry and Remote Sensing*, vol. 33, part B7/1, 452-459.

Fraser, C.S., Hanley, H.B. & Yamakawa, T. 2001. Sub-metre geopositioning with *Ikonos GEO* imagery. *Proc. Joint ISPRS Workshop High Resolution Mapping from Space 2001, Hannover, 19-21 September 2001* (accepted for publication) (on CD ROM).

Gerlach, F. 2000. Characteristics of Space Imaging's One-Meter Resolution Satellite Imagery Products. *International Archives of Photogrammetry and Remote Sensing*, vol. 33, part B1, 128-135.

Grodecki, J. 2001. Ikonos stereo feature extraction - RPC approach. *Proc. ASPRS Annual Conference, St. Louis, 23-27 April 2001*, 7 p. (on CD ROM).

Gruen, A. 2001. Potential and limitations of high-resolution satellite imagery. *Keynote Address, 21st Asian Conference on Remote Sensing, Taipei, 4-8 December 2000*, 13 p. (unpublished).

Hanley, H.B. & Fraser, C.S. 2001. Geopositioning accuracy of Ikonos imagery: indications from 2D transformations. *Photogrammetric Record* 17(98): 12 p. (in press).

Kersten, T., Baltsavias, E., Schwarz, M. & Leiss, I. 2000. Ikonos-2 Carterra Geo – Erste geometrische Genauigkeitsuntersuchungen in der Schweiz mit hochaufgeloesten Satellitendaten. *Vermessung, Photogrammetrie, Kulturtechnik* (8): 490-497.

Okamoto, A., Ono, T., Akamatsu, S., Fraser, C.S., Hattori, S. & Hasegawa, H. 1999. Geometric characteristics of alternative triangulation models for satellite imagery. *Proc. ASPRS Annual Conference, Portland, Oregon, 17-21 May 1999*, 12 p. (on CD ROM).

Space Imaging 2001. Company web site. http://www.spaceimaging.com (accessed 20 June 2001).

Wang, Y. 1999. Automated triangulation of linear scanner imagery. *Proc. Joint ISPRS Workshop on Sensors and Mapping from Space, Hannover, 27-30 September 1999*, 5 p. (on CD ROM).

Yang, X. 2000. Accuracy of rational function approximation in photogrammetry. *Proc. ASPRS Annual Conference, Washington D.C., 22-26 May 2000*, 10 p. (on CD ROM).

Yang, Y. 2001. Piece-wise linear rational function approximation in digital photogrammetry. *Proc. ASPRS Annual Conference, St Louis, 23-27 April 2001*, 14 p. (on CD ROM).