RADIOMETRIC AND GEOMETRIC EVALUATION OF IKONOS GEO IMAGES AND THEIR USE FOR 3D BUILDING MODELLING

Emmanuel Baltsavias, Maria Pateraki, Li Zhang Institute of Geodesy and Photogrammetry, ETH-Hoenggerberg, CH-8093, Zurich, Switzerland (manos, maria, zhangl)@geod.baug.ethz.ch

KEY WORDS: high-resolution satellite imagery, IKONOS, radiometric analysis, image preprocessing, geometric evaluation, orthoimages, building reconstruction, accuracy analysis

ABSTRACT

Investigations on the radiometric and geometric characteristics of IKONOS Geo satellite imagery and its use for orthoimage generation and 3D building reconstruction are reported. The paper first starts with an analysis of the radiometric quality of IKONOS Geo images of varying preprocessing and type, focussing on noise, edge quality and definition, and various artefacts. A noise estimation method is presented and results from various IKONOS Geo images, showing an intensity-dependent noise of 1.5-10.5 grey levels, which is high considering the effectively 8-9 bit data. Methods will be presented that reduce noise, enhance edges and contrast and optimally reduce 11-bit to 8-bit, without leading to loss of small, but still important, image details. Next, the topic of the geometric potential of IKONOS Geo images using the Rational Polynomial Coefficients (RPCs) and a possible improvement will be discussed. The geometric analysis makes use of stereo images and testfield data provided by the University of Melbourne with 28 well defined and distributed GCPs with 1-2 dm accuracy in object and image space. It will be shown that the RPCs have moderate absolute accuracy (large systematic bias, conforming to the Geo product accuracy specifications of 25 m RMS) but sub-metre relative accuracy. With the help of 4-8 accurate and well-distributed GCPs and a simple translation (bias removal), 0.4-0.5 m and 0.6-0.8 m absolute accuracies in planimetry and height can be achieved with the full set of 60 RPCs per image, or reduced RPCs omitting the higher-order terms (down to 22 coefficients). Even when using just one GCP to remove the bias, the accuracy is at worst ca. 2 m. In a related paper (Fraser et al., 2001b), it is shown that other simple sensor models (including affine projection) can produce similar or better accuracy than the RPCs with similar GCP requirements. This is shown also in this work, but with a different 3D spatial resection, using relief-corrected affine transformation and down to just 3 GCPs. Results from generation of orthoimages from three projects using own methods will be shown, including quantitative analysis, and showing how with simple and fast methods the 1-2 m accuracy of the much more expensive IKONOS Precision and Precision Plus can be achieved. With 1-2 dm accurate GCPs and depending on DTM accuracy and sensor elevation, sub-metre planimetric accuracy can be achieved. Buildings of the University of Melbourne Campus were measured manually and stereoscopically. After measuring a point cloud, 3D points were automatically structured and visualised with the system CC-Modeler, including texture mapping. The point accuracy was evaluated using GPS measurements of building corners and was found to be in the meter range. The completeness of the data and their limits regarding modelling roof details were evaluated: (a) qualitatively and (b) quantitatively through a comparison to a 3D model that was generated from aerial images. The results of this assessment are summarised, these highlighting both the high metric potential of IKONOS and difficulties to be anticipated in building reconstruction when the modelling should be complete and detailed.

1. INTRODUCTION

Although IKONOS imagery has been commercially available since early 2000, the use of this imagery and especially the scientific investigations on its potential use in various applications has been restricted due to various reasons, the main ones relating to the closed policy of Space Imaging (SI). Some publications from SI scientists report on the radiometric and geometric characteristics of the sensor (Gerlach, 200), the use of rational polynomial functions as a substitute sensor model (Grodecki, 2001), and the 3D mapping accuracy

that can be achieved with this imagery using stereo triangulation (Dial, 2000). Independent investigations on the geopositioning accuracy of IKONOS using 2D transformations and full 3D analysis are reported by Hanley and Fraser (2001) and Fraser et al. (2001a, 2001b), respectively. Li et al. (2000) report 3D point positioning accuracy studies using simulated data. They mention that 4-7 GCPs suffice, but their results are too pessimistic when using GCPs (1.5 - 3 m RMS) and too optimistic when not using any GCPs at all. Generation of accurate orthoimages from IKONOS Geo imagery is reported by Kersten et al. (2000), Toutin and Cheng (2000) and Davis and Wang (2001) with achieved accuracies in the 1-2 m range. Regarding use of high-resolution spaceborne imagery for object detection, recognition and reconstruction, the first investigations using IKONOS imagery appeared only recently, while some research was performed earlier using simulated data or empirical analysis. Ridley et al. (1997) evaluated the potential of 1m-resolution spaceborne imagery for the needs of UK's Ordnance Survey using simulated data. Regarding buildings only 72.9% and 85.6% could be interpreted correctly for monoscopic and stereoscopic evaluation, respectively. Sohn and Dowman (2001) present investigations on building extraction, however the results presented were few and dealing only with large detached buildings without a comprehensive analysis of accuracy, completeness and modelling of details. Dial et al. (2001) present first investigations on automated road extraction, focussing rather on wide suburban roads of expanding US cities. Hofmann (2001) tries to perform a 2D detection of buildings and roads using spectral information, a laser scanner derived DSM, context and form with the commercial package eCognition.

In between, some commercial systems (Erdas Imagine, LHS Socet Set, Z/I Imaging ImageStation, PCI Geomatics OrthoEngine) support to one or the other extent IKONOS imagery for import, stereo viewing and processing, orthoimage generation etc. by using the rational polynomial coefficients (RPCs) provided by SI with some image products (stereo images, Geo Ortho Kit), estimating RPCs from ground control or using alternative (bundle adjustment, DLT) and partly proprietary mathematical models. While estimating RPCs from GCPs is both expensive (high number of GCPs required) and suboptimal, use of bundle adjustment due to the undisclosed IKONOS sensor model is impossible. DLT, as shown in Fraser et al. (2001a) could be used with GCPs, while proprietary sensor models (PCI) are of unknown quality, while raw IKONOS data which would be most appropriate for use with a strict sensor model are not available.

Given the present shortage of information on the photogrammetric performance of the IKONOS system, it has been necessary to examine various salient aspects when evaluating the use of 1m Geo imagery. These comprise the geometric accuracy of geopositioning from mono and stereo image coverage; the radiometric quality, with an emphasis on characteristics to support automatic feature extraction (e.g. noise content, edge quality and contrast); orthoimage generation; 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, as well as additional images in the radiometric analysis and orthoimage generation.

Details on IKONOS, including sensor and platform parameters and product description, can be found at Space Imaging (2001). Here, it is only reminded that the field of view is 0.93^o and the available products include *Geo*, *Reference*, *Pro*, *Precision* and *Precision Plus* with absolute geopositioning accuracies (RMS, 1-sigma)^{*} of 25m, 11.8m, 4.8m, 1.9m and 0.9m, respectively. In June, SI introduced the Geo Ortho Kit which provides together with the Geo imagery an Image Geometry Model (IGM) allowing users to generate their own accurate orthoimages, using DTMs and GCPs. However, this product has a substantially higher cost than the normal Geo products (Ortho Kit is currently offered only by SI USA at a 57-78% surcharge and SI Eurasia at a 280% surcharge, while the cheapest Ortho Kit product outside N. America is 62 and 98 USD/km² for the two SI companies respectively). In July 2001, SI published new prices and partly new products, e.g. stereo imagery is offered for the reference and precision products only and this not by all SI regional partners. These investigations refer to the Geo product which is a geometrically corrected product

^{*} It should be kept in mind that when no GCPs are used (as with Geo) these errors refer to a, generally unknown to the image user, reference height plane employed in the image generation and it also does not include displacements due to relief. Thus, checking the accuracy of Geo products, using GCPs, often leads to much higher errors (see Table 9) that the ones specified here.

²

that has been rectified to a pre-specified ellipsoid and map projection, re-sampled by cubic convolution to 1m pixels and in the case of a stereopair in addition epipolarly resampled.

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 1, 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.

	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

Tał	ole	 Acquisit 	on parameters	of IKONOS	PAN images.
-----	-----	------------------------------	---------------	-----------	-------------

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 the RPCs which provide a mechanism for object-to-image space transformation and 3D point determination. There were no RPCs 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

The testfield covered a 7x7 km area over central Melbourne with 32 GCPs (mostly road roundabouts) with 1-2 dm accuarcy in object and image space. The position of the roundabout centres was estimated by using 6 or more GPS measurements along their perimeter and a least squares based ellipse fit. The image coordinates were determined similarly, using manual measurements. 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 artefacts such as cars. Thus, 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. In spite of significant disturbances, and size and shape variations of roundabouts, the matching performed correctly in most cases (in 22% of the points matching failed, and in the rest a few points had slight errors due to disturbances at the roundabout perimeter; in case of failed matching, the coordinates used came from the ellipse fit). A point, that is important also in 1D matching for DSM generation, is the fact that with both pixel coordinate measurement methods the y-pixel coordinates of the stereo pair, although usually very close to few tenths of a pixel, in some cases had differences above 0.5 and up to ca. 1 pixel. This could be due to image measurement errors but also imprecisions in the epipolar resampling.

A second essential component of the Melbourne testfield comprised information on buildings. Building height data for 19 image-identifiable and readily accessible roof corners was provided through precise GPS surveys at the University of Melbourne campus, to 10cm accuracy. This data formed the metric standard by which the 3D triangulation for building extraction would be assessed. Image coordinate observations 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. 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, 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, at least in most cases. 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 difference 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.

A fact that is not widely known is that the IKONOS linear CCDs employ Time Delay and Integration (TDI) technology, i.e. the line consists of several lines (also called stages) that accumulate the signal received from one scene object by all lines, in applications where either illumination is very low or the dwelling (integration) time of an object is very short due to high object or sensor speed. Due to the high satellite speed over ground of 6.79 km/s and the small pixel footprint, the integration time of each line must be kept small (typically, IKONOS acquires for the PAN 6000 lines/s resulting in an integration time for each line of the TDI of 0.166 ms, and a pixel footprint in flight direction of ca. 1.13 m). This integration time is too short to achieve a sufficiently strong signal and high SNR and to reproduce the high dynamic range of natural scenes. Thus, IKONOS uses TDI technology in 5 different modes, accumulating each time the signal from a different number of TDI lines, up to 32 (Gerlach, 2000), while the TDI mode remains constant during the acquisition of one particular scene (to solve the same problem, the EROS A high-resolution satellites use a so-called nimble imaging technology, whereby the satellite bends backwards at an almost constant predetermined angular speed to image the same scene object with one single line over a longer time period; the EROS B series satellites will employ TDI technology as IKONOS). Typically, IKONOS images with 16 TDI stages, resulting in a total integration time of 2.7 ms. The accumulation of the signal with TDI technology leads to a smoothing of the signal, especially in the flight direction, since the TDI lines can not exactly image the same scene surface and thus a signal mixing occurs (see Fig. 7 in Gerlach (2000)). SI uses a so-called Modulation Transfer Function Compensation (MTFC) to unsmooth the signal. This leads visually to a sharper image, but at the same time to lower contrast and artefacts (ringing, overshoot) that can be seen especially along strong edges parallel to the CCD direction (see again Fig. 7 in Gerlach (2000)). This leads to a worse edge definition for metric purposes, especially by automated procedures. A second preprocessing performed by SI is the Dynamic Range Adjustment (DRA). This, as MTFC, aims at a visual improvement by stretching the grey values to cover more uniformly the available 11-bit, but this destroys the absolute radiometric accuracy of the images (thus, it is not applied to multispectral data) plus it leads to combination (mixing) of grey values that are not frequently occupied.

All Melbourne images were preprocessed by SI with MTFC but no DRA. Although the images were 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 8-bit intensity range observed with other spaceborne linear array CCD sensors such as SPOT and MOMS. The peak of the histo-

gram 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.

To increase the reliability of the results and test images with varying spectral content and different preprocessing and imaging conditions, several additional images apart from the Melbourne images were used. The overall noise characteristics of the images were analysed in both homogeneous (sea and lake surfaces) and nonhomogeneous areas (e.g. whole image, excluding large homogeneous surfaces). 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 for CCD-imagers is not additive but intensity-dependent. Both areas are selected manually and as large as possible. A small window is moved within the area with a freely defined spacing and the standard deviation, as an indication of noise, is calculated. Small windows are justified since homogeneous areas (e.g. water) often show low frequency grey value variations, which would lead to higher standard deviation if it were computed from the whole area. For nonhomogeneous areas, a small window is imperative in order to get small homogeneous areas between edges. For homogeneous areas, the standard deviations are sorted and the N% smallest ones are used to calculate a mean standard deviation, which indicates the noise. Typical values for N are 80-95. With nonhomogeneous areas, the grey level range is divided in bins, and the standard deviations are assigned to a bin according to the mean grey value of each window. In each bin, the standard deviations are sorted, and the noise is estimated as the mean of the N% smallest standard deviations. N is chosen small, e.g. 5, using the reasonable assumption that at least some windows will be homogeneous, even in highly textured images. The noise is estimated for a bin only if the N% sample number is sufficiently large (e.g. > 100). For both areas, care should be taken with homogeneous areas that are saturated (at one of the histogram ends), as this will lead to too low estimated noise. This noise estimation method is quite general and could be applied to any type of image, without the need of on-platform calibration devices or special targets in the image.

The results using four PAN images, two MSI and one PAN-MSI and homogeneous areas led to a noise estimate (mean standard deviation) of 4.5-5, 2 and 4.5-5.5 grey levels respectively. All images were preprocessed by SI using MTFC, while one PAN was preprocessed also with DRA, and PAN-MSI was also preprocessed by an unknown projective multispectral algorithm. The results showed a very high consistency, i.e. noise was similar for all PAN and MSI images, in spite of the different image acquisition characteristics. MSI images exhibit lower noise, probably due to larger pixel spacing (48 µm) and electron well capacity of the multispectral line CCDs. As expected, the PAN-MSI images have higher noise, due to the higher noise of PAN which is injected in the sharpened image, and maybe the specific image sharpening algorithm used by SI. The three spectral channels (R, G, B) had similar noise for the MSI images, while for the PAN-MSI (NIR, R, G) the G channel had lower noise by 1 grey value. Using nonhomogeneous regions, the noise relation to the spectral image type remained generally the same, but the noise was intensity dependent being lower for the first bins (at 1.5 grey levels), and increased for higher intensities up to ca. 10.5, which is more than suggested by SI investigations (4 grey values in all 5 channels without MTFC, see Gerlach, 2000). In the multispectral channels, the noise slightly increased from B through to NIR. The Lucern/PAN filled more bins due to the DRA. Although the noise extent was not judged to be of crucial significance for the specific task of building reconstruction, if one considers the fact that the 11-bit data represent actually only 8 to 9-bit, the noise is high, while its negative influence on automated processing becomes even more pronounced, if an often necessary contrast enhancement is performed (see Fig. 2, top right).

Geo images have been found to exhibit artefacts in addition to their noise content, of which a number remain unexplained. Many are visible only in homogeneous areas, especially after contrast enhancement. Although the artefacts often lead to small grey level variations (e.g. 2-4 grey levels 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 artefacts have been identified: some striping in the stereo images in the flight direction (Fig. 1 e, width of stripes ca. 630 pixels); signal saturation due to strong reflectance and spilling of this bright signal due to read-out to neighbouring lines in flight direction (Fig. 1 f); an apparent MTFC effect similar to the unsharp masking used to sharpen edges (Fig. 1 f, note black contour at the bottom of the white saturated area); a staircase variation of the grey values within homogeneous areas of the epipolar-resampled stereo im-

ages (Fig. 1 a shows the right stereo image, Fig. 1 b the left one; note also the white dotted line on water, which is physically nonexisting); jumps in grey value across the flight direction (visible in homogeneous areas, see Fig. 1 c); the ghosting of moving objects leaving colour patterns behind, visible especially for blue, due to the 0.5 s time difference between acquisition of PAN and MSI and partly deficiencies of the image sharpening method (Fig. 1 g and h show a stereo Pan-Sharpened pair; note the blue car in 2 h which is however not visible in 2 g, while the corresponding MSI imagery of 2 h is, at the position of the blue car, only very slightly blueish); physically nonexisting bright lines, neither along or across flight direction; and some striping normal to the flight direction (horizontal bright stripes in Fig. 1 d). 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 artefacts, 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 with surface normal approximately in the middle of the illumination-tosensor 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.

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 first consisted 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. 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. 2). 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.

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.

Next, a new version of the Wallis filter (Baltsavias, 1991), which estimates automatically some of the filter parameters, is applied, and finally a reduction to 8-bit imagery by histogram equalisation is performed (Figs. 5 and 6). We used histogram equalisation, as it is optimal for general-purpose computer processing, since it preserves more grey values that are more frequently occurring. 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 first transforming to 8-bit and then preprocessing, even though the 11-bit imagery was effectively only 8 to 9 bits. The third preprocessing approach produced sharper edges than the first two ones (see Fig. 2). The histogram equalisation may lead to too strong bright and dark regions for visual interpretation, and thus for this purpose it could be replaced either by a reduction of Gaussian type or an optimisation of a selected target grey level range. A noise estimation of the Melbourne stereo images after the fuzzy noise reduction method (third approach) led to a noise decrease by factor three, while after Wallis the noise increased. However, the images had still slightly lower noise than the original ones, but with better image quality. After reduction to 8-bit the noise increased, with only few bins showing more than 1-3 grey level noise, as encountered in scanned aerial imagery. The noise had a Gaussian form, being lower at the two histogram ends due to compression of grey levels at these histogram positions by the histogram equalisation.



Fig. 1. Artefacts (see explanations in text).



Fig. 2. 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 and Sensor Orientation Models

The last 15 years various mathematical models have been formulated to derive 3D information and generate orthoimages from spaceborne line CCD sensors, especially SPOT, IRS-1 C/D and MOMS. These models have varying complexity, rigour and accuracy. Strict sensor models have been developed using the known sensor information and modified collinearity equations, in some cases including parameters for modelling errors in the interior orientation or in-flight calibration, or incorporating orbital information and orbital constraints. Several empirical tests have shown that with such models, reasonable B/H ratio and good quality GCPs in both object and image space 3D accuracies in the order of 0.25 - 0.3 pixels can be achieved for well-defined points. It would be reasonable to except that high-resolution imagery as IKONOS would exhibit similar accuracy characteristics. However, for IKONOS strict models can not be used, as both the sensor model and ephemeris data are proprietary information of SI, while raw data are also not commercially available but only one that have been geometrically processed by unknown transformations. Thus, alternative models come into play.

One possibility is the use of RPCs provided with some IKONOS products, like stereo and Geo Ortho Kit (Grodecki, 2001; Yang, 2000). Since the RPCs for Geo products are estimated by using only position and attitude data of the sensor and no GCPs, it is expected that the accuracy in both image and object space that can be achieved by RPCs will not be very high and would correspond to the accuracy specifications of this

product). As it will be shown below, this has been verified by tests, while it has been also shown that the positional errors are very systematic due to a large bias and can be reduced to sub-metre errors in the simplest case by just a translation, which can be computed using some GCPs. Other models that do not need sensor and orbit information have been developed and used for spaceborne linear CCDs. RPCs can be estimated by using GCPs and not as an approximation of a strict sensor model as in the case of the RPCs provided by SI. Other nonrational polynomial models (Kratky, 1989; Papapanagiotu and Hatzopoulos, 2000) can be estimated similarly. However, estimating such polynomial coefficients using GCPs needs many GCPs that cover the whole planimetric and height range, which is difficult to impossible and costly, and also leads to extrapolation errors and possible undulations between GCPs, especially with high degree polynomials. DLT has been used by El-Manadili and Novak (1996) and Savopol and Armenakis (1998) with SPOT and IRS-1C images respectively, while Wang (1999) expanded the DLT by adding corrections for selfcalibration, and Yang (2001) used it in piece-wise functions. Hattori et al. (2000) present 4 models termed 1D perspective, 1D affine, parallel perspective and 2D affine models based on extensive work of Okamoto (see e.g. Okamoto et al., 1999). For larger FOVs (e.g. SPOT), the affine projection introduces errors, which can be eliminated by a pretransformation of the images from central perspective to affine projection using terrain height information. Hattori et al. (2000) present results with SPOT images which show that both affine models perform better than the perspective ones. Palà and Pons (1995) used similarly a 2D affine model, incorporating terrain relief corrections in image space, however, only for generation of orthoimages, not 3D positioning.

All above models refer mostly to original raw images. However, Geo is rectified to a so-called inflated ellipsoid and map projected without terrain relief corrections. To relate object information to such imagery, the object information could be first corrected by relief displacements using a flat Earth or curved Earth model. Then, a simple affine transformation can relate object and image spaces (see Section 4.3), without the need to apply any image transformation (from central perspective to affine or correction for terrain relief). In fact, since the relative accuracy of Geo images is in the sub-metre level, the major correction needed are shifts between the two spaces. The two scales and shears are needed to model smaller deviations between the two spaces. Using various Geo images with GCPs in different projection systems and affine transformation between the two spaces, showed that the two scales deviate from 1 in the 4th or 5th decimal only, while the rotation is in the order of 1-2 deg (except of course for stereo images which are rotated), and the nonorthogonality of the axes is in the order of 0.003-0.03 deg. This affine transformation can be used not only for object to image transformation, but also 3D point positioning using two or more images, using essentially 6 parameters per image, and not 8, as sensor elevation and azimuth are known (see Eq. (8)). Simple linear models like the DLT and affine projection perform well with IKONOS images that cover a small area, but as the area increases, unmodelled higher-order terms are expected to lead to an accuracy deterioration. However, this also depends on the quality of the unknown methods used by SI to generate Geo images, i.e. to what extend non-linear terms can be modelled in generation of large Geo images.

4.2 Use of RPCs for Object to Image Transformation and 3D Point Positioning

In the rational function model, image pixel coordinates (px, py) are expressed as the ratios of polynomials of object coordinates (X, Y, Z), which in the case of the SI RPCs correspond to latitude, longitude and height. In order to improve the numerical stability of the operations, the two image and three object coordinates are each offset and scaled to fit the range from -1.0 to 1.0. For an image, the ratios of polynomials have the following form:

nx	$=\frac{fl(X_n,Y_n,Z_n)}{fl(X_n,Y_n,Z_n)}$	
P. n	$f2(X_n, Y_n, Z_n)$	(1)
D V	$-\frac{f3(X_n, Y_n, Z_n)}{f3(X_n, Y_n, Z_n)}$	
РУ _п	$-\frac{1}{\mathrm{f4}(\mathrm{X}_{\mathrm{n}},\mathrm{Y}_{\mathrm{n}},\mathrm{Z}_{\mathrm{n}})}$	

with px_n and py_n normalised pixel coordinates and $X_n Y_n Z_n$ normalised object coordinates.

To compute the normalised coordinates, the following equations are used:

$$px_{n} = \frac{px - px_{0}}{px_{s}}, \quad py_{n} = \frac{py - py_{0}}{py_{s}}$$

$$X_{n} = \frac{X - X_{0}}{X_{s}}, \quad Y_{n} = \frac{Y - Y_{0}}{Y_{s}}, \quad Z_{n} = \frac{Z - Z_{0}}{Z_{s}}$$
(2)

with px_{0} , py_{0} and px_{s} , py_{s} offset and scale values for the two image coordinates respectively, and similarly X_{0} , Y_{0} , Z_{0} and X_{s} , Y_{s} , Z_{s} offset and scale values for the object coordinates.

In the RPCs, the maximum power of each object coordinate and the total power of all object coordinates are limited to 3. In such a case and following the SI definition of coefficient sequence, each polynomial has the following form (for convenience, the subscripts are omitted), which leads to the 80 RPCs per IKONOS image:

$$f(X,Y,Z) = c_1 + c_2 Y + c_3 X + c_4 Z + c_5 XY + c_6 YZ + c_7 XZ + c_8 Y^2 + c_9 X^2 + c_{10} Z^2 + c_{11} ZYX$$

$$c_{12} Y^3 + c_{13} X^2 Y + c_{14} Z^2 Y + c_{15} Y^2 X + c_{16} X^3 + c_{17} XZ^2 + c_{18} Y^2 Z + c_{19} X^2 Z + c_{20} Z^3$$
(3)

With the SI RPCs, the two polynomials used in denominator are identical, while the first coefficient of the denominator is 1, to avoid the case when the denominator becomes close to zero, leading thus to practically 59 parameters. Using the RPCs, the distortions caused by the optical projection can generally be represented by the ratios of first-order terms. Thus, these terms are used by the space intersection procedure below to derive the initial object coordinate values.

The RPCs can be directly used for the transformation from object to pixel coordinates. However, the transformation from pixel to object coordinates is the inverse procedure and needs an iterative calculation due to the non-linearity of the RPCs:

 Calculation of initial values for ground coordinates. As mentioned before, the distortions caused by the optical projection can be represented by the ratios of first-order terms in the RPCs. Omitting the higherorder RPCs, the functions which transform the object coordinates into pixel coordinates can be expressed as follows:

$$px^{i} = \frac{(a_{1} + a_{2}Y + a_{3}X + a_{4}Z)^{i}}{(b_{1} + b_{2}Y + b_{3}X + b_{4}Z)^{i}}; \quad py^{i} = \frac{(c_{1} + c_{2}Y + c_{3}X + c_{4}Z)^{i}}{(d_{1} + d_{2}Y + d_{3}X + d_{4}Z)^{i}}$$
(4)

where *i* is the index of the images (*i* = 2), and $(a_k, b_k, c_k, d_k)_{k=1,\dots,4}$ are first-order RPCs. After making some changes of these equations, we can get the following equation groups:

$$((a_{1} - b_{1}px) + (a_{2} - b_{2}px)Y^{0} + (a_{3} - b_{3}px)X^{0} + (a_{4} - b_{4}px)Z^{0})^{i} = 0$$

$$((c_{1} - d_{1}py) + (c_{2} - d_{2}py)Y^{0} + (c_{3} - d_{3}py)X^{0} + (c_{4} - d_{4}py)Z^{0})^{i} = 0$$
(5)

Obviously, given two or more images and their first-order RPCs, the initial object coordinates (X^0, Y^0, Z^0) can be derived from all above equations or just 3, two involving px and one py.

2) Derivation of the final object coordinates. By performing a Taylor expansion, the observation equations can be formulated as

$$px^{i} = F(X, Y, Z)^{i} = \frac{f1(X^{0}, Y^{0}, Z^{0})}{f2(X^{0}, Y^{0}, Z^{0})} + \frac{\partial F^{0}}{\partial X}?X + \frac{\partial F^{0}}{\partial Y}?Y + \frac{\partial F^{0}}{\partial Z}?Z$$

$$py^{i} = G(X, Y, Z)^{i} = \frac{f3(X^{0}, Y^{0}, Z^{0})}{f4(X^{0}, Y^{0}, Z^{0})} + \frac{\partial G^{0}}{\partial X}?X + \frac{\partial G^{0}}{\partial Y}?Y + \frac{\partial G^{0}}{\partial Z}?Z$$
(6)

and the final object coordinates can be derived by least square adjustment.

Out of the 32 testfield points, two were not available and another two seemed to be erroneous. Using the pixel coordinates of the remaining 28 GCPs, the corresponding pixel coordinates in left and right IKONOS images were computed using the RPCs and compared to the known pixel coordinates of the two datasets from ellipse fitting and least squares template matching (see Table 2). The differences had a very large systematic component. However, the relative accuracy as shown by the standard deviations is in the subpixel range, so after removing the bias by the use of some GCPs, e.g. 6, the achieved RMS become similar to the standard deviations listed in Table 2. The results with least squares matching are slightly worse than those of the ellipse fitting dataset, since occlusions and especially shadows at the roundabout perimeter influence matching, while with ellipse fitting the perimeter points were selected manually. However, with least squares matching the errors in left and right image are similar (i.e. the parallax errors are less), since for both images the template image and matching parameters were identical, and the shadow disturbances very similar. This consistency leads to better height accuracy as shown in Table 3, although the planimetry is still worse.

The values of the higher order RPCs, especially in the denominator, are very small. Their effect in the resulting pixel coordinates, even after multiplication with scale, is not significant even for the worst case, when the normalised object coordinates have an absolute value of 1. This was observed in various sets of available RPCs. To investigate how many RPCs are really needed, we repeated the computations of Table 2, combining the first 1 and 4 coefficients in the denominator (1 corresponds to nonrational polynomials and 4 include the 1st order terms), with the first 4, 9, 10 and 20 coefficients of the numerator (4, 10 and 20 correspond to 1st, 2nd and 3rd order polynomials, while 9 is like 10 without the quadratic term for the height, which in various RPCs was consistently smaller). Leaving out some coefficients, could change the value of the remaining ones, in presence of correlations between the parameters, which are very probable in the overparametrised RPCs. The correct procedure would be to recompute the RPCs with reduced coefficients, as an approximation to the strict sensor model, but clearly this could not be done. The results for the 20/4, 10/4 and 9/4 RPCs in the numerator/denominator give very similar results to the 20/20 version, which is what we expected from the numerical values of the left-out coefficients, i.e. 9 parameters in the numerator and 4 in the denominator or 21 parameters in total seem to suffice. This could be useful to speed up operations, if RPCs are to be used. The worse relative accuracy of the remaining versions (3-10 pixels standard deviation), combined with the better results of the 4/1 version (ca. 1 pixel), are in our opinion due to correlations between the RPCs and not an indication that more coefficients are needed. Indeed, in this paper and in Fraser et al. (2001b) it is shown that an affine model, computed from 4-6 GCPs gives equal or better accuracy than the full set of RPCs.

		Ellipse	fitting		Least squares matching					
	Left Image		Righ	t Image	Left	ft Image Right Ima				
	х	X V		у	х	у	х	У		
Mean	28.94	-16.07	28.00	-16.52	28.94	-15.91	27.93	-16.14		
Stand. Dev.	0.40 0.35		0.56	0.46	0.51	0.62	0.65	0.62		

Table 2. Differences (in pixels) between known pixel coordinates and ones computed via RPCs.

Using the known pixel coordinates of the GCPs measured by ellipse fitting and least square template matching, 3D object coordinates of these points can be computed by the above described space intersection algorithm. Compared with the GPS measured coordinates after transformation of both coordinate sets to UTM, the RMS errors in X, Y and Z were 8.0, -31.3, 1.5 m respectively and the results again show a very large systematic trend (see Table 3). Then, six well-distributed points were selected as control points and other points served as check points. By using the six control points, two correction methods, only removing a positional bias and 3D similarity transformation between the RPC-computed and the known object coordinates, were performed, leading to sub-metre 3D accuracies (Table 3). The results are not sensitive to the selection of the six points as long as they are fairly well distributed. Fig. 3 shows the deviation of the GCP differences from the mean difference (bias). Although this test dataset is limited to be conclusive, it seems that in planimetry, where the bias is higher, the deviations vary within the image and in some cases they are locally systematic. Thus, the use of well-distributed GCPs for bias removal is suggested. The difference between the two corrections (translation and 3D similarity) is small, with the 3D similarity transformation giving slightly better results, especially in Z. Another version that was tried, was the subtraction of the bias in the pixel space, before performing 3D spatial intersection, and the results were similar to the ones with bias subtraction in the object space. Table 4 shows the maximum deviation from the mean bias. This shows that if only one GCP is used to remove the bias, the maximum additional error which will be added to the submetre accuracy after bias removal is in the order of 1-2 m.



Fig. 3. Plot of the deviations from the mean bias computed from all 28 GCPs for planimetry (left) and height (right).

In a related paper, Fraser et al. (2001b) present results of 2D transformations using similarity, affine and projective models and 3D positioning using DLT and affine models with 2 and 3 image configurations. Their results show that with all these simple models sub-metre accuracy can be achieved, and that the affine model performs the best with 0.3 - 0.4 m planimetric accuracy (RMS) for 2D transformations and 0.25-0.45 m and 0.45-0.70 m planimetric and height accuracy using 2-3 images, 4-8 GCPs and bundle adjustment for 3D spatial resection. This accuracy is higher than the one achieved using the full set of RPCs.

Processing	Least Squares Matching					Ellipse Fitting						
Method	RMS (m)		n)	Ι	Max. (m)		RMS (m)			Max. (m)		
	Х	Y	Ζ	Х	Y	Ζ	Х	Y	Z	Х	Y	Z
Use of RPCs	8.00	31.37	1.55	8.96	-32.40	-2.18	8.28	31.50	1.78	9.74	-32.40	-3.52
Position Bias Removal	0.65	0.50	0.68	-1.89	1.08	-1.80	0.49	0.43	0.79	1.48	1.21	-1.92
3D similarity	0.65	0.44	0.58	2.07	-1.01	1.10	0.48	0.37	0.74	-1.18	-1.13	-1.82

Table 3. Errors of 3D point determination using RPCs.

Least Squares Matching					Ellipse Fitting						
Mean Bias Max. Dev. from Bias			Mean Bias Max. Dev. from Bias				m Bias				
Х	Y	Z	Х	Y	Z	Х	Y	Ζ	Х	Y	Ζ
7.98	7.98 -31.37 -1.42 -2.02 -1.09 -2.09					8.26	-31.49	-1.59	1.48	1.22	-1.92

Table 4. Mean bias computed from all 28 GCPs and maximum deviation of GCP errors from mean bias (in m).

4.3 Use of Relief-Corrected Affine Transform for Object to Image Transformation and 3D Point Positioning

The GCPs in UTM were projected on a reference height of 0m, and then several transformations (affine, bilinear, quadratic and biquadratic) from object to image space were computed with all GCPs. The results (sigma-0 after the least squares adjustment) were practically identical for all transformations. The results for the ellipse fit and matching datasets for the left/right/nadir images were 0.41/0.46/0.36 and 0.60/0.60/0.45 pixels respectively. The nadir image gave better results probably due to better image quality and more accurate relief correction, since height errors influence the relief displacement by factor 5 less than for the stereo images due to the higher sensor elevation. The residuals showed that some image measurements, especially with matching were poor, with more than 1 pixel errors (thus, also the results of the matching dataset are less accurate). The affine transformation was repeated with variable number of GCPs (3, 6, 8, all and "all, cleaned" without few pixel measurements with large errors). The results were very similar to the ones published in Fraser et al. (2001b) and the results from the matching dataset were worse in almost all cases, by ca. 0.1-0.2 pixels. However, it was interesting to note than 3 GCPs gave similar or better results than 6 or 8 (see Table 5). This is a confirmation that the most important factor is the point quality not their number. The 3 GCPs for the stereo and nadir images were points 4, 8, 16 and 9, 26, 29 respectively (see Fig. 1 in Fraser et al. (2001b)), with good but not optimal distribution. The results with 3 GCPs were worse than the ones with all GCPs only by 13-14% and 22-24% for the stereo and nadir images respectively. The conclusion is that even 3 highly accurate GCPs with fair distribution suffice for subpixel accurate transformation from object to image space.

As a next step, we performed 3D spatial resection using the following method. The above affine transformations which were computed for each image independently provided six parameters (see Eq. (8) below). The parameters a₄ and b₄ were computed as explained at Eq. (8). These 8 parameters were used with the same programme for 3D resection using RPCs (see Section 4.2), without bias removal naturally. The results using the ellipse fit dataset are listed in Table 6. Although the affine parameters from the "all, cleaned" version were used, all points were used in the computation of the error statistics. It is very interesting to note that the results with 3 GCPs are very similar to the "all, cleaned" version. In all cases, planimetric accuracy is in the 0.3-0.5 m range and height accuracy in the 0.6-0.8 m range. The results are very similar to the ones in Fraser et al. (2001b), however some significant differences exist. Fraser et al. compute 8 affine parameters (as in Eq. (8)) using GCPs and bundle adjustment. We estimate only 6 parameters (excluding a_4 and b_4 , which are functions of known quantities and other parameters) for each image separately and using relief-corrected planimetric coordinates of GCPs. Thus, we need minimum only 3 GCPs and no image pretransformation from central perspective to affine projection (Hattori et al., 2000) is needed. Later, is time consuming, needs a DTM, interior orientation and approximate exterior orientation and may introduce image errors. A use of bundle adjustment in our estimation of the 6 affine parameters could be favourably used but has not been implemented. Summarising, with our approach 3 accurate GCPs can be used for the computation of 6 affine parameters from relief-corrected object space to image space without any knowledge about the sensor and its orbit, but knowing the sensor elevation and azimuth and assuming they are constant within one scene. These parameters can be used for any transformation between the original object space and image space, including orthoimage generation (see Section 5), as well as 3D spatial resection for point positioning and automatic DSM generation, including its use to constrain matching along a line or perform an epipolar resampling.

Image	Coordinate	# control	# check	x-RMS	x-Mean	x-Max.	y-RMS	y-Mean	y-Max.
	dataset	points	points		with Sign	Abs.		with Sign	Abs.
Left	Ellipse fit	3	25	0.45	-0.06	1.16	0.43	0.06	0.91
	Matching	3	25	0.51	0.07	1.17	0.75	0.21	1.96
Right	Ellipse fit	3	25	0.47	-0.03	1.46	0.53	0.09	1.27
	Matching	3	25	0.54	-0.01	1.47	0.75	0.19	1.93
Nadir	Ellipse fit	3	25	0.39	0.01	1.31	0.44	-0.18	0.95
	Matching	3	25	0.37	0.17	0.93	0.67	0.16	1.95

Table 5. Error statistics (in pixels) for affine transformation from object to image space for various images and pixel coordinate datasets. The statistics refer to the check points only.

# of images	2					3				
# of GCPs	3	6	8	All, cleaned	3	6	8	All, cleaned		
X-RMS	0.49	0.44	0.45	0.42	0.32	0.33	0.33	0.32		
Y-RMS	0.38	0.47	0.46	0.38	0.32	0.39	0.37	0.30		
Z-RMS	0.76	0.75	0.73	0.71	0.57	0.62	0.60	0.58		

Table 6. Error statistics (in m) for 3D resection using relief-corrected affine transform computed from different number of GCPs. The statistics refer to the check points only, except the "all, cleaned" versions. The number of check points is 28 minus the number of GCPs.

5. ORTHOIMAGE GENERATION

5.1 Orthoimage Generation Methods

Our group has been involved in various projects aiming at generation of accurate orthoimages with use of GCPs and DTMs starting from IKONOS Geo. Thereby, two approaches have been used to model the transformation from object to image space. The first approach is using Kratky's polynomial mapping functions (PMFs) (Kratky (1989), Baltsavias and Stallmann (1992)), which include 14-16 terms up to 4th degree. The PMFs, like the RPCs, are a very good approximation of a strict sensor model and are estimated similarly to the RPCs. They have been successfully used for orthorectification of spaceborne imagery (Baltsavias and Stallmann, 1996). In the IKONOS case, with unknown sensor model, the PMFs must be estimated using GCPs, which is suboptimal, as explained in Section 4.1. The second approach is using an affine transformation. At least 3 GCPs are needed, but we have used in our projects 4-6 to have a certain redundancy (although 3 good GCPs suffice as shown in Section 4.3). Before computing the affine transformation from object to image space, the GCPs are projected on a reference plane and the planimetric X and Y coordinates are corrected according to Eq. (7):

 $\Delta X_i = -\Delta Z_i * \sin(a) / \tan(e)$ $\Delta Y_i = -\Delta Z_i * \cos(a) / \tan(e)$ (7)

with $\Delta Z_i = Z_i - Z_o$

 Z_i = height of ground point i

 $Z_o =$ height of reference plane

a = sensor azimuth

e = sensor elevation

 ΔX_i , ΔY_i = planimetric displacement of point i due to projection onto the reference plane

whereby the azimuth and elevation are taken from the metadata file delivered with the images. Eq. (7) shows that the radial planimetric displacement depends only on elevation, while its distribution among the two axes X and Y depends on the azimuth. For scenes imaged away from the orbit footprint, the displacement would be more in the X direction and vice versa. The above equation has been used for the whole Geo scenes, as the variation of azimuth and elevation within a scene, when the satellite is not rotating, is so small, that their influence on the planimetric displacement is negligible. The choice of the reference plane is irrelevant (at least in our implementation, where the affine transformation is computed from centred coordinates), but we use for less computations (no need to compute ΔZ_i) the same reference plane as for the DTM, i.e. at height zero. The GCPs do not have to cover the whole image, i.e. their distribution is not so important, as long as the affine parameters can be determined reliably. This can be achieved, if the range of their coordinates in the 2 directions (in object or image space) covers e.g. ca. 1/3 of the area dimensions. Then, for each orthoimage pixel, knowing the height, we project it onto the same reference plane and then perform the affine transformation and grey value resampling. In our implementation, we combine the projection to the reference plane and the affine transformation to one common operation involving 8 parameters (instead of 6 for the 2D affine transformation) according to Eq. (8), and this operation is applied to the DTM nodes, which are treated as anchor points. An additional advantage of this approach is that the regular DTM structure is preserved, and thus an update philosophy can be used in the computations leading to even faster orthoimage generation.

 $\begin{array}{l} x_i = a_1 + a_2 \; X_i + a_3 \; Y_i - a_4 \; Z_i \\ y_i = b_1 + b_2 \; X_i + b_3 \; Y_i - b_4 \; Z_i \end{array}$

(8)

with x_i , y_i ...pixel coordinates

 X_i, Y_i, Z_i ...DTM coordinates without any relief displacement of the planimetric position $a_1, a_2, a_3, b_1, b_2, b_3$...six 2D affine parameters computed from the GCPs after projection to the reference plane $a_4 = [a_2 * \sin(a) + a_3 * \cos(a)] / \tan(e)$ $b_4 = [b_2 * \sin(a) + b_3 * \cos(a)] / \tan(e)$

The advantage of the second approach over the first one is the need for much less GCPs (which are difficult to get in non-urban regions) and more freedom in their choice (distribution). Furthermore, the computations are less. The accuracy of the second approach can be better since there are no extrapolation effects and possible numerical instabilities. In addition, if GCPs along a water body can be selected, the height must be known only for one GCP. For both approaches, the higher the sensor elevation, the less important the height becomes. In the extreme case of very high elevation, the orthoimage can be produced from planimetric GCPs and respective 2D transformations, without a DTM or with a much less accurate one. The planimetric influence of height errors can be estimated by Eq. (7), e.g. for an elevation of 60, 70, 80 and 85 deg, a height error of 1 m, causes 0.58, 0.36, 0.18 and 0.09 m radial displacement, respectively. Since Geo products usually have elevations larger than 60 deg, height errors always influence planimetry to a lesser extent. The relief corrected affine transformation used in orthoimage generation but also 3D positioning (see Section 4.3) assumes than elevation and azimuth are constant within one scene. As mentioned above, the variation of azimuth and elevation is minimal within a scene, when the satellite does not rotate during scene acquisition, which seems to be the common scanning mode used by Ikonos. However, Ikonos has also other scanning modes, where azimuth and especially elevation may change within a scene, e.g. consider the case when the sensor is looking forward and rotates backwards constantly (i.e. elevation change) to scan a scene backwards. Although such scanning modes seem to be seldom, are more complicated and for flat terrain and non-Lambertian object reflection cause varying imaging conditions (illumination-to-sensor angle changes), caution should be paid that for some scanning modes the constant angle model may be insufficient and should be replaced by ones using higher order (especially linear) terms in the angle variation.

5.2 Influence of DTM and GCPs on Orthoimage Accuracy

The planimetric accuracy of the orthoimage depends on the accuracy of the GCPs and the DTM. For IKO-NOS compared to other spaceborne sensors, DTM accuracy is less important due to the small FOV, while GCP accuracy becomes more important, due to the small pixel footprint. As shown in Sections 4.2 and 4.3, the planimetric potential of IKONOS Geo lies in ca. 1/3 pixel. Thus, the GCPs should be 1-2 dm accurate in both object and image space. While getting this accuracy with GPS in object space is no problem, finding image points suitable for measurement by image analysis techniques with 0.1-0.2 pixel accuracy and accessible in the scene for GPS measurement can be problematic. This is one more reason in favour of the second orthoimage generation approach. Best GCPs have good contrast, are preferably on the ground and are intersections of straight, long enough lines or centres of gravity of circular/elliptical features. The planimetric accuracy of the orthoimage can be easily estimated by the GCP accuracy and Eq. (7) using as input the DTM accuracy and the known azimuth and elevation. For example, with 1-2 dm GCP accuracy, DTM accuracy of 2 m and elevation larger than 70 deg, a sub-metre planimetric accuracy can be achieved, similar to the much more expensive Precision Plus product.

5.3 Test Results

We have produced orthoimages from IKONOS Geo in three projects. Table 7 shows some characteristics of the images and input data for the orthoimage generation used. In all projects, the aim was to generate more accurate orthoimages, starting from the cheapest product of SI. In the first project described in Kersten et al.

(2000), our first approach was used (see Zug / 1 and Zug / 2 in Table 8) and a second one from another company (Zug / 3 and Zug / 4), similar to our second one but more complicated in estimating the azimuth and elevation. Zug / 1 has check points only within the perimeter of the used control points, while Zug / 2 also outside, showing the extrapolation errors that can occur with our first approach, especially notable in the maximum absolute errors. The achieved accuracy with both methods was 1.5-2.5 m RMS, but in this case the potential of IKONOS was not fully exploited, because the GCPs, as checked with more accurate reference data made available after the project end, had unfortunately the same accuracy as the produced IKONOS orthoimages and a high bias error, close to the RMS error. The orthoimages were not accurate enough, because the DTM used in their production was not a very accurate one and most probably orientation errors causing the bias.

In the second project, a Pan-Sharpened image of a Greek island with little man-made objects was orthorectified. Although 38 GCPs measured with GCPs were available, about half of them were not well-defined in the image. The DTM has a grid spacing of 2 m and was interpolated from 1:5,000 map contours and additionally measured breaklines. Its accuracy, estimated using the GCP points, was ca. 3.3 m RMS. Both approaches were employed with 38 and 28, and 4 GCPs respectively for the geometric transformation (Nisyros / 5, / 4, / 3 and / 2 in Table 9). Version 3 is like 2 but only with the points that were well identifiable and measurable, and shows the accuracy potential with good GCPs. Nisyros / 1 shows the accuracy of the original Geo image. The planimetric accuracy of both approaches was similar, with the second approach being slightly better. The second approach was not more accurate, as we had expected, due to the lack of accurate enough GCPs. This project is dealt with in detail in another publication (Vassilopoulou et al., 2002).

In the third project, a PAN image of Lucerne was processed with the second approach and 6 GCPs covering 1/2 to 1/3 of the image extent for the affine transformation. The GCPs included 5 GPS points, 46 points measured in triangulated aerial imagery and 17 points measured from Swissimage orthoimages (see product generation parameters below). The expected accuracy of these 3 groups was 0.5, 1 and 2-3 m, however, the accuracy of the last two groups seemed to be worse, based on a 2D affine transformation using all GCPs after reduction to a reference height plane (excluding two large blunders, 16 points had residuals more than 3m and up to 8.3m). The GCP accuracy varied a lot, with the orthoimage GCPs being insufficient and most photogrammetric GCPs equally inadequate. The majority of GCPs were not well-defined in the image, including the GPS points. Version Lucerne / 1 in Table 9 shows the accuracy of the original Geo product. Versions 2 to 5 show the accuracy using all GCPs and sequentially reducing their number, deleting the poorer ones. Again, versions 4 and 5 show that with good GCPs, the accuracy can be in meter level or below. Production of the same orthoimage with identical GCPs using the PCI OrthoEngine software which uses a self-developed sensor model for IKONOS led to X and Y RMS of 6.1 and 1.5 m respectively with very systematic residuals.

	Zug	Lucerne	Nisyros
Date, Time (local)	8/4/2000, 10:11	22/4/2000, 10:21	8/4/2000, 10:35
Sensor azimuth (deg)	266.0	256.3	134.9
Sensor elevation (deg)	85.7	67.7	73.5
Sun azimuth (deg)	151.3	153.6	136.6
Sun elevation (deg)	47.1	53.0	53.4
DTM spacing/accuracy (m)	5 / 0.4	25 / 2.5 in lowland,	2/3.3*
	(raw irregular data)	10 in Alps	
GCP accuracy (m)	1.5-2	$0.5 - 3^{**}$	ca. 0.5
GCP definition	Medium to good	Very poor to good	Poor to good
Elevation range (m)	400-990	0-700	400-2100

* Estimated from the GCPs

** See comment on real accuracy in text

Table 8. Acquisition parameters of IKONOS images used in orthoimage generation.

Project / Version	Image type / Control Points	Check Points*	Mean w	vith sign	RM	S	Maximum absolute	
			X	Y	Х	Y	X	Y
Zug / 1	PAN / 27	41	0.9	-1.3	1.5	1.6	3.8	3.2
Zug / 2	PAN / 27	69	1.1	-1.5	2.5	2.0	11.3	6.5
Zug / 3	PAN / 34	26	-1.3	-2.5	1.8	2.7	3.4	5.0
Zug / 4	MSI / 39	26	-0.8	-1.5	1.4	1.7	3.6	3.5
Lucerne / 1	PAN	66	-70.9	-15.4	134.2	30.6	501.5	118.1
Lucerne / 2	PAN / 6	65 **	-0.47	0.92	2.6	2.2	9.9	5.9
Lucerne / 3	PAN / 6	47	-0.14	0.86	2.2	2.2	6.7	5.9
Lucerne / 4	PAN / 6	21	-0.23	-0.15	0.8	0.8	1.5	1.5
Lucerne / 5	PAN / 6	14	-0.22	0.07	0.6	0.6	1.0	1.1
Nisyros / 1	PAN-MSI	38	-102.6	70.0	106.1	75.5	153.1	122.8
Nisyros / 2	PAN-MSI / 4	34	-0.7	0.2	1.7	1.0	4.4	2.3
Nisyros / 3	PAN-MSI / 4	15	-0.6	-0.1	0.9	0.6	1.5	1.4
Nisyros / 4	PAN-MSI / 28	10	-0.3	1.0	1.8	1.5	4.4	2.6
Nisyros / 5	PAN- MSI / 38	0	-0.4	0.8	1.5	1.3	3.7	2.3

* Check points in Lucerne include 6 control points ; ** One GCP was outside the produced orthoimage

Table 9. Statistics of planimetric errors (in m) of generated orthoimages (see explanations in text).

The above results are a proof that orthoimages of 1-2 m accuracy can be routinely produced from various Geo products (PAN, MSI, PAN-MSI), with varying image quality and sensor elevation, with DTM and GCPs of rather moderate accuracy, and hilly to steep terrain and without needing a strict sensor model nor RPCs. With better quality GCPs, even sub-metre accuracies can be achieved. Use of the Geo Ortho Kit product results in significantly higher costs and much more mathematical operations, while DTM and GCPs are still needed and probably also transformations from the local map coordinate system to the one supported by the IGM. The accuracy improvement of the produced orthoimages compared to the original Geo images regarding both RMS and maximum absolute errors as shown in Table 9 is enormous.

6. BUILDING EXTRACTION FROM IKONOS

6.1 Accuracy of Building Extraction

Depending on the application, accuracy requirements vary. Regarding the here treated building extraction within the wider application frame of digital city models, metric accuracy expectations again vary. To fulfil requirements for mobile communications, one of the major markets for city models, accuracies in the 1-2 m range are generally needed. This implies the need to use the expensive Precision or Precision Plus imagery (with costs outside N. America 140-180 USD/km² and 220-250 USD/km² respectively). However, as it will be shown here, 3D positioning at 1-m accuracy level or better can be achieved by using the cheapest product Geo.

Whereas the 3D ground point determination in Section 4 was centred upon 'high-quality' targets, accurate positioning of building features, generally corners and edges, involves not only metric factors but also 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 RPCs 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 orthoimage generation of built-up urban areas.

6.2 Evaluation of Building Extraction

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 Fig. 4. The manual measurements of roof corners and points of detail were topologically structured automatically using the software package CC-Modeler (Gruen and Wang, 1998) and also visualised (Fig. 5).

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 reconstructed in a coarse generalised form (Fig. 6). 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 influential factors are shadows, occlusions, edge definition (related also to noise and artefacts), saturation of bright surfaces, sun and sensor elevation and azimuth, and atmospheric conditions. The 1m resolution of IKONOS also leads to certain interpretation restrictions. This investigation was aimed in part at quantifying the extent of such limitations.

A comparison of the two models in Fig. 4, one from aerial photography and the other from IKONOS 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 be often only generalised with a simplified roof structure and variations to their form and size. It is interesting to note that this 15% fits very well with the findings of Ridley et al. (1997). However, additional tests with different IKONOS imagery are needed in order to draw safer conclusions.



Fig. 4. 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.



Fig. 5. Visualisation of University of Melbourne campus with buildings and trees extracted from stereo IKONOS images.



Fig. 6. Building with complicated roof structure as measured from aerial images (left) and IKONOS images (right).

7. CONCLUSIONS

This investigation has shown that IKONOS Geo stereo imagery has high geometric integrity and the potential to yield sub-metre geopositioning accuracies, using either simple geometric models for sensor modelling or provided RPCs and a moderate number (4-8) of accurate GCPs. Similarly, object to image transformations and orthoimage generation can be achieved with 1-2 m accuracy, or even sub-metre one with good GCPs, with a simple and fast 2D affine transformation after relief displacement corrections (or incorporating the later in a 3D affine transform). This affine transformation gives sub-metre results with just 3 good GCPs and can be also employed in 3D spatial resection for point positioning and DSM generation. Building points can be measured with an accuracy of ca. 1m. 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 (or other) 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 of the imagery 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, restriction to 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 problems in trying to comprehensively describe the stereo scene and identify not only all buildings of a certain size, but also to accurately reconstruct their form without excessive generalisation. The shortcomings in radiometric quality, although not a major influence factor for this manual building extraction, do cause problems in both manual and computer mensuration accuracy and interpretability. Noise is relatively high, the applied MTFC destroys to a large extent the edge information, and many artefacts exist. Thus, better radiometric sensor performance is needed to fit the high geometric quality of IKONOS and a better preprocessing needs to be developed by SI, along the lines presented in this work. The issues of radiometric inhomogeneity both within and between IKONOS scenes are unlikely to be affected by the level of IKONOS product purchased and this investigation has demonstrated that very high metric performance is achievable, at least for the small area of single scenes, with the least expensive product offering, namely Geo imagery, and without the need of sensor model information, ephemeris data or RPCs.

ACKNOWLEGEMENTS

The provision of imagery and ground control information by the Department of Geomatics, University of Melbourne and the co-operation with C. Fraser, H. Hanley and T. Yamakawa are gratefully acknowledged. The IKONOS image of Lucerne was kindly provided by the National Point of Contact, Swiss Federal Office of Topography, Wabern, the images of Zug by the company Swissphoto AG, while the IKONOS image of Nisyros was made available by Prof. E. Lagios, University of Athens, Greece within the EU project Geowarn (www.geowarn.org). Gene Dial and Laurie Gibson (Space Imaging) are acknowledged for valuable discussions and suggestions and provision of information.

REFERENCES

- Baltsavias, E.P., 1991. Multiphoto geometrically constrained matching. Ph.D. dissertation, Institute of Geodesy and Photogrammetry, ETH Zurich, Report No. 49.
- Baltsavias, E.P, Stallmann, D., 1992. Metric information extraction from SPOT images and the role of polynomial mapping functions. International Archives of Photogrammetry and Remote Sensing, Vol. 29, Part B4, pp. 358-364.
- Baltsavias, E., Stallmann, D., 1996. Geometric potential of MOMS-02/D2 data for point positioning, DTM and orthoimage generation. International Archives of Photogrammetry and Remote Sensing, Vol. 31, Part B4, pp. 110-116.
- 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).
- Dial, G., Gibson, L., Poulsen, R., 2001. IKONOS satellite imagery and its use in automated road extraction. In: Automated Extraction of Man-Made Objects from Aerial and Space Images (III), E. Baltsavias, A. Gruen, L. Van Gool (Eds.), Balkema Publishers, Lisse, pp. 349-358 (in press).
- El-Madanili, Y., Novak, K., 1996. Precise rectification of SPOT imagery using the direct linear transformation model. Photogrammetric Engineering and Remote Sensing, 62(1), pp. 67-72.
- Fraser, C., Baltsavias, E., Gruen, A., 2001a. 3D building reconstruction from high-resolution IKONOS stereo imagery. In: Automated Extraction of Man-Made Objects from Aerial and Space Images (III), E. Baltsavias, A. Gruen, L. Van Gool (Eds.), Balkema Publishers, Lisse, pp. 325-338(in press).
- Fraser, C.S., Hanley, H.B., Yamakawa, T., 2001b. Sub-metre geopositioning with Ikonos GEO imagery. Proc. Joint ISPRS Workshop High Resolution Mapping from Space 2001, Hannover, 19-21 September (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, pp. 128-135.
- Gruen, A., Wang, X., 1998. CC-Modeler: a topology generator for 3-D city models. ISPRS Journal of Photogrammetry and Remote Sensing, 53(5), pp. 286-295.

- Grodecki, J., 2001. Ikonos stereo feature extraction RPC approach. Proc. ASPRS Annual Conference, St. Louis, 23-27 April 2001, 7 p. (on CD-ROM).
- Hanley, H.B., Fraser, C.S., 2001. Geopositioning accuracy of Ikonos imagery: indications from 2D transformations. Photogrammetric Record, 17(98), 12 p. (in press).
- Hattori, S., Ono, T., Fraser, C., Hasegawa, H., 2000. Orientation of high-resolution satellite images based on affine projection. International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B3/1, pp. 359-366.
- Hofmann, P., 2001. Detecting buildings and roads from IKONOS data using additional elevation information. GIS, (6), pp. 28-33.
- 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, pp. 490-497.
- Kratky, V., 1989. Rigorous photogrammetric processing of SPOT images at CCM Canada. ISPRS Journal of Photogrammetry and Remote Sensing, 44(2), pp. 53-71.
- Li, R., G. Zhou, S. Yang, G. Tuell, N.J. Schmidt and C. Fowler 2000. A Study of the Potential Attainable Geometric Accuracy of IKONOS Satellite Imagery. International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B4/2, pp. 587-595.
- 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, 12 p. (on CD-ROM).
- Palà, V., Pons, X., 1995. Incorporation of relief in polynomial-based geometric corrections. Photogrammetric Engineering and Remote Sensing, 61(7), pp. 935-944.
- Papapanagiotu, E.G., Hatzopoulos, J.N., 2000. Automatic extraction of 3D model coordinates using digital stereo images. International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B4/2, pp. 805-812.
- Ridley, H.M., Atkinson, P.M., Aplin, P., Muller, J.P., Dowman, I., 1997. Evaluating the potential of the forthcoming commercial US high-resolution satellite sensor imagery at the Ordnance Survey. Photogrammetric Engineering and Remote Sensing, 63(8), pp. 997-1005.
- Savopol, F., Armenakis, C., 1998. Modelling of the IRS-1C satellite pan imagery using the DLT approach. International Archives of Photogrammetry and Remote Sensing, Vol. 32, Part 4, pp. 511-514.
- Sohn, G., Dowman, I., 2001. Extraction of buildings from high resolution satellite data. In: Automated Extraction of Man-Made Objects from Aerial and Space Images (III), E. Baltsavias, A. Gruen, L. Van Gool (Eds.), Balkema Publishers, Lisse, pp. 339-348 (in press).
- Space Imaging, 2001. Company web site. http://www.spaceimaging.com (accessed 10 August 2001).
- Toutin, Th., Cheng, P., 2000. Demystification of IKONOS! EOM, 9(7), pp. 17-21.
- Vassilopoulou, S., Hurni, L., Dietrich, V., Baltsavias, E., Pateraki, M., Lagios, E., Parcharidis, I., 2002. Orthophoto Generation Using Ikonos Imagery and High Resolution DEM: Study Case on Volcanic Hazard Monitoring of Nisyros Island (Greece). ISPRS Journal of Photogrammetry and Remote Sensing (submitted for publication).
- Wang, Y., 1999. Automated triangulation of linear scanner imagery. Proc. Joint ISPRS Workshop on Sensors and Mapping from Space, Hannover, 27-30 September, 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, 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, 14 p. (on CD-ROM).