ELSEVIER

ISPRS Journal of Photogrammetry & Remote Sensing 53 (1998) 217–234

PHOTOGRAMMETRY
& REMOTE SENSING

Geometric and radiometric evaluation of the DSW300 roll film scanner [1]

Emmanuel P. Baltsavias [a,*], Silvio Haering [b], Thomas Kersten [b], Alex Dam [c]

[a] *Institute of Geodesy and Photogrammetry, ETH-Hoenggerberg, 8093 Zurich, Switzerland*
[b] *Swissphoto Vermessung AG, Dorfstraße 53, 8105 Regensdorf, Switzerland*
[c] *LH Systems, LLC, 10965 Via Frontera, San Diego, CA 92127, USA*

Received 5 January 1998; accepted 8 April 1998

Abstract

Geometric and radiometric investigations performed with two LH Systems DSW300 scanners are presented. Their performance evaluation was carried out using good quality test patterns and accurate processing methods. The geometric tests include global and local geometric errors, misregistration between colour channels, geometric repeatability and determination of the geometric resolution. Efforts were made to separate the contribution of various error sources (especially mechanical positioning, vibrations and lens distortion) on the total error. The radiometric tests include investigations of noise, linearity, dynamic range, spectral variation of noise, and artifacts. After a brief description of the scanner, details on the above investigations, analysis and results are presented. Regarding the geometric accuracy, the RMS was 1.3–1.9 μm and the mean maximum absolute error 4.5–8 μm. The errors are bounded, i.e. on the average the 3σ (99.7%) values are 3 RMS, and the maximum absolute error 3.7 RMS. The co-registration accuracy of colour channels was about 1 μm. The short and medium term repeatability was very high. With a linear Look Up Table (LUT) the radiometric noise level is 1 and 1–1.5 grey values for 25 and 12.5 μm scan pixel size, respectively. The dynamic range is 2 D (D ... density) with a very good linear response up to this value. One of the major remaining radiometric problems is dust. In both geometric and radiometric tests, no significant differences between R, G, B and B/W scans has been observed. These results show that the geometric and radiometric quality of the DSW300 has been very much improved as compared to the DSW200 and also other scanner models. This test was part of a long and fruitful cooperation between the manufacturer, a major user, and an academic institution and shows that honest and critical behaviour, as well as thorough understanding of the problems and a desire to search for solutions, can lead to significant improvements to the benefit of all.

Keywords: roll film scanner; scanner test; CCD; geometric evaluation; radiometric evaluation; colour misregistration

* Corresponding author. Fax: +41 (1) 633-1101; E-mail: manos@geod.ethz.ch

[1] Slightly modified version of a paper presented at the ISPRS Commission I Symposium, 25–27 February 1998, Bangalore, India.

0924-2716/98/$ – see front matter
PII S0924-2716(98)00007-0

1. Introduction

Photogrammetric film scanners are, and in the near future will be, ever more used for producing digital data, especially from aerial images. Since every subsequent processing step builds upon the scanned imagery, the analysis of the scanner accuracy and performance is of fundamental importance. Unfortunately, many users take for granted that all photogrammetric scanners perform well. However, experience with several scanners has shown that many problems of a geometric and radiometric nature may occur.

An overview of photogrammetric scanners is given in Baltsavias (1998). General scanner test procedures are given in Gruen and Slater (1983), Baltsavias (1994), and Seywald (1997). Testing of specific scanners includes geometric and radiometric evaluation of the RM-1 (Bethel, 1994, 1995; Bolte et al., 1996; Jakobsen and Gaffga, 1998), DSW200 (Miller and Dam, 1994; Baltsavias et al., 1997), and Zeiss SCAI (Baltsavias and Kaeser, 1998), image noise and sensitivity analysis of PS1, VX3000 and RM-1 (Koelbl and Bach, 1996), tests of the IDS scanner (very similar to Zeiss/Integraph PS1) (Roos, 1993), tests on the geometric accuracy, MTF and noise level of VX3000 (Leberl et al., 1992; Seywald et al., 1994; Seywald, 1996) and on the geometric accuracy of PS1 (although the model was not explicitly named) (Seywald, 1996).

The authors have cooperated on testing DSW200 scanners for over 2 years. After publication of a critical paper on different problems, particularly geometric ones, of the DSW200 (see Baltsavias et al., 1997), there was cooperation with LH Systems (LHS is a new joint venture between Leica and Helava) in San Diego with the aim of making improvements to the DSW300 and testing it using the same procedures as in the aforementioned paper. Many improvements were implemented by LHS and the tests with two DSW300 scanners were carried out successfully by three of the authors in November 1997 at LHS in San Diego. This paper presents the results of these tests. It must be noted that the processing of the data was done at ETH and Swissphoto and the results were verified by independent analysis at LHS.

A short description of the DSW300 scanner is given in Dam and Walker (1996). DSW300 is similar to DSW200, but apart from allowing roll film scanning, it has a more robust stage because of the added weight of the roll film support and film media, a more precise servo-mechanism, thicker platen, and slightly different electronics for encoders and motors to sense and control the film roll position. Mechanical positioning is achieved by two stages, with the y-stage being the secondary one. The geometric accuracy specification is 2 μm per axis. The sensor (a Kodak Megaplus with 2029 × 2044 pixels) and the optics are stable and lie below the moving scanner stage. An image larger than the sensor dimensions is scanned as a mosaic consisting of several tiles, each with user definable dimensions from 960 × 960 to 1984 × 1984 pixels in increments of 64. An overlap region of four pixel width between tiles is used to equalise radiometrically neighbouring tiles and an optional linear feathering can be performed across the borders of the tiles to smooth out any remaining radiometric differences. For colour scanning, each tile is scanned sequentially in R, G, B with the use of a rotating filter which is positioned before the liquid pipe optic and away from the stage to reduce the danger of vibrations. A uniform, diffuse illumination is produced by using a xenon lamp and a sphere diffusor. The base scan pixel size (4–20 μm) is set at the factory and for both scanners tested was 12.5 μm. Larger pixel sizes (25, 50, ... μm) can be achieved by local averaging (2 × 2, 4 × 4, ...) of the grey values using software. The radiometric accuracy (noise level) of the scanner according to coarse manufacturer specifications should be about 1–2 grey values and the maximum density 3 D. The user can specify a Look Up Table (LUT) for mapping the output 8-bit grey values from the 10-bit input. The illumination source and the electronics are positioned away from the stage and the sensor to avoid heating. The scanning throughput depends on the host computer (currently a Sun Ultra 30) and the output image format. It is currently about 2 MB/s (12.5 μm, colour) for an Ultra 30 host (used in these tests) while the maximum scan speed is 100 mm/s. Its price is about 150,000/125,000 US$ with/without roll film option.

The scanner software performs two geometric (see Miller and Dam, 1994, for a brief description) and two radiometric calibrations. The first geometric calibration (stage calibration) is performed by measuring a reference grid plate of 13 × 13 crosses with 2 cm spacing. The crosses are measured auto-

matically by cross-correlation, and after computation of an affine transformation between stage and grid reference coordinates, corrections (offsets) to the scanner stage at predefined grid positions are computed. These corrections are applied on-line in each scan. Note that the grid covers an area of 240 × 240 mm, while the possible scan area is 265 × 265 mm. In stage positions outside the calibration grid, the corrections are extrapolated and saved in a calibration file covering 15 × 15 grid nodes. The second geometric calibration (geometric sensor calibration) computes the relation between the pixel and the stage coordinate system (two scales and two shears). This is achieved by moving one grid cross at the centre of the grid plate such that a 5 × 5 grid is created, and then an affine transformation between pixel and stage coordinates is computed. The scales and shears of this transformation are used at every tile position in order to relate all local pixel coordinate systems to the global stage coordinate system. The scanner manufacturer generally suggests performing a new geometric calibration every 2 weeks.

The radiometric calibration includes an equalisation of the grey values for low and high illuminations, where grey level non-uniformities are caused mainly by differences in the CCD sensor element responses, and to a much lesser degree by illumination and glass plate non-uniformities, and vignetting. For this calibration the scanner stage glass plate is scanned at two positions for each channel and algorithms try to detect differences due to spatially varying noise (mainly dust, but also scratches, threads, etc.) and exclude these from the computation of the corrections. Using the grey values at the two illuminations, an offset and gain correction factor are computed for each sensor element. If varying dust is not correctly detected, wrong radiometric corrections are applied and 'electronic' dust is created. Stationary dust, e.g. on the lens, is corrected for. The corrections are computed in 16-bit and added to the raw 10-bit data. Finally, an equalisation of the colour response (colour balance) can be performed, using the histograms of the colour channels.

2. Description of test procedures and test patterns

Our geometric and radiometric investigations were performed with two DSW300 scanners, after first performing all the necessary scanner calibration procedures with an accuracy of less than 1.5 μm for the geometric calibrations. The scanners were located at LHS in rooms without temperature and humidity control and more dust than the 'clinical' room environments at Swissphoto where previous tests have been conducted. One was a demo scanner (called DS) in an office, the other in the factory (called FS), on a concrete floor and without a cover, so more problems due to vibrations and flare light would be expected. Both scanners used a Kodak Megaplus 4.2i model and the firmware revision 3. The host computer was a Sun Ultra 1, 167 MHz with 256 MB RAM. Both scanners were equipped with an about 18 kg roll film transport system but no roll film (an additional 7 kg) was mounted on it.

For the tests three glass plates were used. One high precision réseau glass plate came from Rollei, which has been produced by Heidenhain, with a 2 mm grid spacing (116 × 116 crosses), 200 μm cross length, 15 μm line width, and an accuracy of the reference cross positions better than 1 μm. The calibration glass plate of the scanner, with a 2 cm grid spacing (13 × 13 lines), 25 μm line width, and an accuracy of better than 2 μm, is called the DSW300 plate in this paper. The third plate was more planar than the second one (10 μm maximum out of plane deviation over the whole area), had a 1 cm grid spacing (23 × 23 lines), 20 μm wide lines, and an accuracy of better than 1 μm, and will be referred to as the planar plate in this paper. The Rollei and DSW300 plates were used exclusively for testing the geometric accuracy and performing the stage calibration of the scanner, respectively. The use of the planar plate served two purposes. Firstly, it was employed to check the influence of cross density on the accuracy results, i.e. comparison to the Rollei plate, and secondly, to check whether with a denser and more planar plate than the DSW300 one, better calibration results and thus higher geometric accuracy could be obtained. To determine the scanner resolution, a standard USAF resolution pattern on glass produced by Heidenhain was used. The radiometric performance was mainly checked by scanning a calibrated Kodak grey level wedge on film (21 densities with density step of approximately 0.15 D; density range 0.055 D–3.205 D). The densities were determined by repeated measurements (4 to 15) using a

Gretag D200 microdensitometer with a resolution of 0.01 D.

All test patterns were scanned with DS, while for FS only the Rollei grid plate was scanned. All scans were with a 12.5 μm pixel size, if not otherwise mentioned. After performing a calibration with the DSW300 plate (called calib2), the Rollei plate was scanned with DS three times in colour to check the short term geometric repeatability as well as misregistration between the colour channels. This was repeated after 1 day to check the medium term repeatability of the scanner (using calib2 again). Between these two tests the Rollei plate was scanned once in colour but this time after performing a calibration with the planar plate (calib3). Finally, the Rollei plate was also scanned once in B/W and colour using the FS scanner. In addition, with DS the planar plate was scanned twice in colour, once with calibration using the DSW300 plate (calib2) and once using calib3. The resolution pattern was scanned three times, the second and third time by shifting the scan area by half a pixel in x and y, respectively, in order to account for an unknown arbitrary phase shift between sensor elements and lines of the resolution pattern, which can influence the results for high line frequencies. The grey level wedge was scanned in B/W and colour, with 12.5 μm (with linear and logarithmic LUT) and 25 μm (only linear LUT) pixel size to check differences between colour channels and B/W scans, the effect of the LUT, and the effect of pixel size on the radiometric performance. Use of a logarithmic LUT results in taking the logarithm of the 10-bit input values and then scaling them to the range [0, 255]. The wedge was masked with a black carton to avoid stray light.

The pixel coordinates of the grid crosses were measured by fully automatic Least Squares Template Matching (LSTM). This algorithm is described in Gruen (1985) and details can be found in Baltsavias (1991). The software implementation of the algorithm that was used employs on-the-fly generation of the templates and is described in Kersten and Haering (1997). An option of the algorithm that reduces the influence of dust and other noise on the cross measurement was used. The accuracy of LSTM, as indicated by the standard deviations of the translations, for these targets was 0.02–0.03 pixels. Matching results with bad quality criteria (low cross-correlation coefficient, etc.) were automatically excluded from any further analysis. In addition, the matching results of all crosses with large errors were interactively controlled. However, smaller errors (5–6 μm) due to dust have remained in the data set. For some of the grid plates, the crosses were also measured by the cross-correlation algorithm of the scanner calibration software. Both results were very similar, so here only the results from LSTM will be reported.

The geometric tests performed include:

(1) *Global geometric tests*. For this purpose an affine transformation between the pixel and the reference coordinates of all crosses was computed with three versions of control points (all crosses, 8 and 4, the latter two versions simulating the fiducial marks used in the interior orientation of aerial images). The use of multiple plates permits an analysis of the influence of pattern density on the ability to detect errors reliably.

(2) *Misregistration errors between the channels*. Such errors were checked by comparing pairwise the pixel coordinates of each channel (R–G, R–B, G–B).

(3) *Local geometric tests (only for DS)*. For this purpose an affine transformation between the pixel and the reference coordinates of the crosses of each individual image tile was computed. The errors and the affine parameters of each individual tile were compared to each other. Errors influencing the whole tile (mechanical positioning, vibrations) are absorbed by the translations of the affine transformation, so the local tile errors reflect primarily errors due to the optical components, especially lens distortion.

(4) *Repeatability (only for DS)*. It was checked by comparing the results between different scan dates using the same scanner and geometric calibration.

(5) *Stability, robustness*. It was checked by comparing: (a) the results of the same plate but using different calibrations; (b) the results between Rollei and planar plates; and (c) the results between the two scanners.

(6) *Geometric resolution*. It was determined by visual inspection of the scanned resolution pattern, i.e. the smallest line group that was discernible was detected, whereby it was required that the contrast between lines is homogeneous along the whole line length.

In the first three tests listed above, efforts were made to separate the contribution of various error sources (especially mechanical positioning, vibrations and lens distortion) to the total error.

The radiometric tests include:

(1) *Estimation of the noise level, linearity and dynamic range*. This was done by determining the mean and standard deviation for each density of the grey level wedge. In previous tests it has been noticed that the grey level wedges of our film, especially for the high densities, are not homogeneous, i.e. they are lighter towards the borders. There is also a very small decrease in the grey values across each density rectangle as one goes from low to darker densities. To avoid the influence of such inhomogeneities on the computed grey level statistics, only the central region of each wedge was used (the same region for all wedges and test scans, independently of the scan pixel size). In addition, in previous tests when scanning with small pixel size a corn pattern was sometimes visible. To reduce the effect of such dark corn and also of dust, etc., grey values that are outside a specified range are excluded from the computation of the statistics. The range is computed for each grey wedge as (mean $\pm 3 \times$ standard deviation), whereby the minimum and maximum allowable range is 4 and 20 grey values, respectively. The minimum range is used to avoid excluding too many pixels in high density wedges with small standard deviations due to saturation. The linearity was checked by plotting the logarithm of the mean grey value of each wedge against the respective calibrated density (when using a logarithmic LUT the grey values were first transformed to the original 10-bit values entering the LUT, before taking the logarithm). These points should ideally lie along a line and be equidistant. The dynamic range is determined as follows. Firstly, the minimum unsaturated density is selected. Then, the maximum detectable density i is determined using the following conditions: (a) $M_{i+1} + \mathrm{SD}_{i+1} + \mathrm{SD}_i < M_i < M_{i-1} - \mathrm{SD}_{i-1} - \mathrm{SD}_i$, with M and SD the mean and standard deviation of the wedges and i increasing with increasing density (i.e. the distance of the mean grey value of a detectable density from the mean values of its two neighbouring densities must be at least equal to the sum of the SD of the detectable density and the SD of each of its neighbours); (b) $\mathrm{SD}_i > 0.1$ (to avoid cases when other conditions, especially condition (a), are fulfilled but the signal is in reality saturated and therefore has a very small SD); (c) nint $(M_i) \neq$ nint (M_j), with j any other density except i (i.e. since grey values are integer the mean grey value value of a detectable density must differ from the mean grey value of all other densities).

(2) *Artifacts*. Some of the above-mentioned scanned patterns were very strongly contrast-enhanced by Wallis filtering (Baltsavias, 1991). This permits the visual detection of various possible artifacts such as radiometric differences between neighbouring tiles, 'electronic' dust, etc. However, the quantification of radiometric errors is always performed using the original images.

3. Evaluation of geometric performance

3.1. Global geometric accuracy and repeatability

The results of this evaluation are shown in Table 1 and some examples are illustrated in Fig. 1a and Fig. 2a. The transformation with eight control points (CP) was left out of the table to make it more readable. Generally, they were slightly better than the results with four CP (14% and 5% lower RMS in x and y, respectively). When using few CP, the transformation results depend heavily on the CP quality, so a higher redundancy (eight instead of four points) is positive.

The results using all crosses as control points are examined first. The differences in accuracy between R, G, B is negligible. The results of a B/W scan with the FS scanner (not listed here) were also very similar to the R, G, B scans. With the DS, the results in the x-direction are clearly worse than in y; with the FS, the results in the x-direction are very slightly better. In all tests we previously performed with the DSW200, the results in y (secondary stage) were consistently worse. The DSW300, however, has a new stage and servos, and the accuracy in the two directions is more balanced. Accuracy in x may be poorer because x-positioning comes after y-positioning and, due to the high scan speed, it might not have fully converged. The FS, although operating under bad factory conditions, was slightly more accurate than the DS with respect to RMS and maximum absolute errors, especially in x. The short

Table 1
Statistics of differences between pixel and reference coordinates after an affine transformation for the Rollei plate with the DS scanner (if not otherwise mentioned) and using the DSW300 plate (if not otherwise mentioned) for calibration (units in μm)

Scan version, No. of control/check points	Statistics [a]	Red channel			Green channel			Blue channel		
		mean	min.	max.	mean	min.	max.	mean	min.	max.
First 3 scans, calib2, 13,444/0	RMS x	1.8	1.8	1.8	1.9	1.8	1.9	1.9	1.9	1.9
	RMS y	1.3	1.3	1.3	1.3	1.3	1.4	1.5	1.5	1.6
	max. abs. x	7.0	6.7	7.3	7.9	7.3	8.4	8.0	7.8	8.2
	max. abs. y	4.5	4.2	5.2	5.5	5.2	5.7	5.6	5.4	5.7
First 3 scans, calib2, 4/13,440	RMS x	2.4	2.1	2.5	2.8	2.6	3.1	2.8	2.4	3.1
	RMS y	1.5	1.5	1.5	1.5	1.4	1.6	1.6	1.6	1.7
	mean x	1.4	1.0	1.6	2.0	1.6	2.3	2.0	1.4	2.4
	mean y	0.7	0.4	0.8	0.5	0.4	0.7	0.5	0.4	0.7
	max. abs. x	9.6	8.7	10.7	10.3	10.2	10.3	8.7	7.5	9.7
	max. abs. y	6.6	6.3	7.0	6.9	6.2	7.2	7.9	7.7	8.2
Second 3 scans, calib2, 13,448/0	RMS x	1.7	1.7	1.8	1.8	1.7	1.8	1.8	1.8	1.8
	RMS y	1.5	1.5	1.5	1.4	1.4	1.5	1.6	1.6	1.6
	max. abs. x	6.0	5.9	6.1	7.2	6.8	7.4	7.8	7.0	8.2
	max. abs. y	6.3	6.2	6.5	5.4	5.1	5.5	5.5	5.5	5.6
Second 3 scans, calib2, 4/13,444	RMS x	2.5	2.4	2.5	2.9	2.8	3.0	2.9	2.7	3.1
	RMS y	1.7	1.6	1.9	1.7	1.7	1.7	1.7	1.6	1.7
	mean x	1.3	1.2	1.4	2.0	1.9	2.3	2.2	1.9	2.4
	mean y	0.7	0.4	1.0	0.8	0.6	0.9	0.5	0.3	0.5
	max. abs. x	7.9	7.8	8.0	8.8	8.7	9.0	8.4	8.1	8.6
	max. abs. y	10.1	9.5	10.9	8.9	8.2	9.7	8.8	8.4	8.8
One scan, calibration with planar plate, 13,451/0	RMS x	1.8			1.8			1.6		
	RMS y	1.4			1.3			1.3		
	max. abs. x	6.2			7.0			5.7		
	max. abs. y	5.0			4.6			5.1		
One scan, FS scanner, 13,452/0	RMS x	1.3			1.3			1.4		
	RMS y	1.4			1.4			1.5		
	max. abs. x	4.8			5.5			5.6		
	max. abs. y	6.5			4.8			5.4		

[a] When all points are used as control points, the mean values are zero and thus listed in this and all subsequent tables.

term repeatability, as indicated by the difference between minimum and maximum values of the first three or the second three scans (see Table 1), is very good. The same applies to the medium term repeatability, when comparing the results of the first three scans to those of the second three scans. The calibration with the planar plate gave an improvement only in the blue channel as compared to the results using the DSW300 plate for calibration. However, its use resulted in a much more homogeneous and smooth error distribution (compare Fig. 1a and Fig. 2a). Summarising, the RMS in x/y are (1.6–1.9)/(1.3–1.6) for the DS and (1.3–1.4)/(1.4–1.5) for the FS. The mean maximum absolute errors in x/y were (6–8)/(4.5–6.3) for DS and (4.8–5.6)/(4.8–6.5) for FS. Table 1 does not include the accuracy results when using DS with the planar plate, and for calibration both the DSW300 and the planar plate. These results were even better, particularly in RMS y and the maximum absolute errors (in y especially this error was between 3.0 and 4.2 μm!). It might be that the Rollei plate can better detect local large errors due to its higher density. A final remark on maximum absolute errors. Most of them are only local and in these tests were often caused by wrong measurements due to dust that still remained in the data set. For that rea-

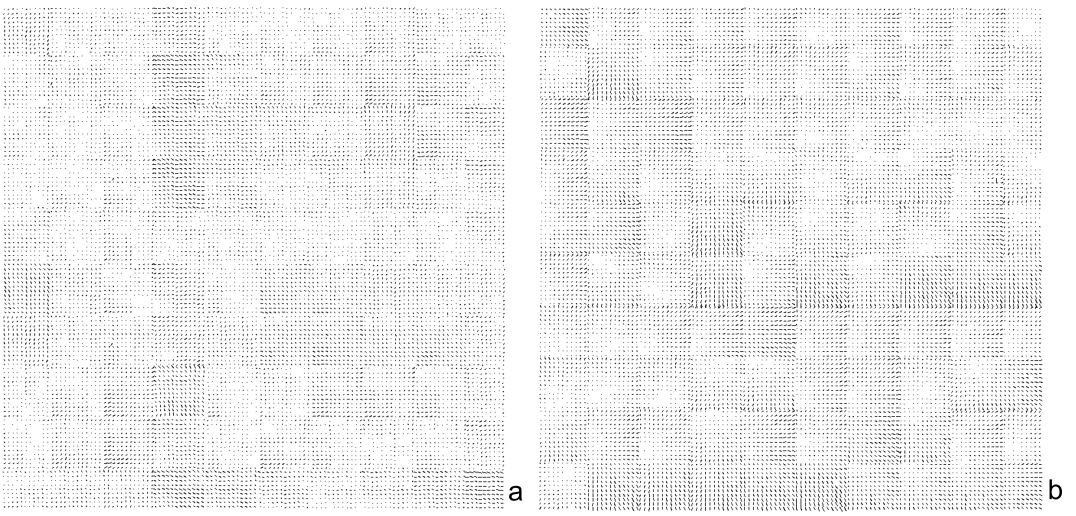

Fig. 1. Rollei plate, DS scanner, calibration with the DSW300 plate (calib2) (vectors enlarged by a relative factor of 3 left, and 5 right). (a) Residuals of global affine transformation from pixel to reference coordinates using all grid crosses as control points (red channel). The tile structure of the image is only slightly visible as compared to previous results with the DSW200 (Baltsavias et al., 1997; see also Fig. 2b). The errors are almost constant within each tile. Note the systematically larger errors in the fourth tile column. These were similar for all channels and all first and second three scans, so this is an indication that they are caused by errors in the stage calibration. (b) Colour misregistration errors, i.e. pixel coordinates of red minus blue channel. The errors are similar within each tile, and vary from tile to tile, the main reason being vibrations between the spectral scans.

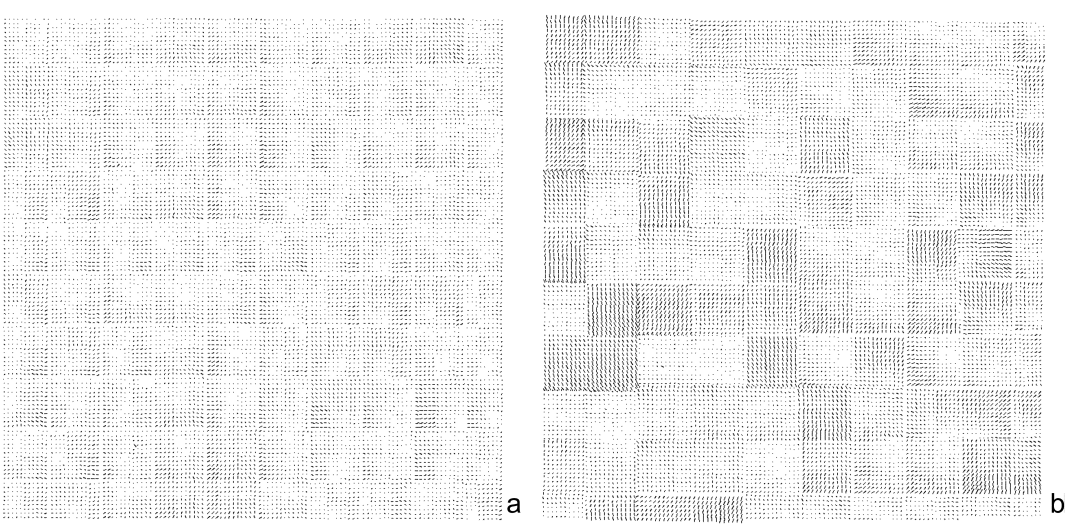

Fig. 2. Rollei plate (vectors enlarged by a relative factor of 3). (a) DS scanner, calibration with the planar plate (calib3). Residuals of global affine transformation from pixel to reference coordinates using all grid crosses as control points (red channel). Although the statistical values of the errors shown in Fig. 1a and 2a do not differ a lot, the residuals of Fig. 2a are smoother and more uniform. (b) One of the results using the same test procedures with a DSW200 scanner (Baltsavias et al., 1997). This is one of the best results achieved with the DSW200. Compare to (a) and Fig. 1a.

son the 3σ value was also computed, i.e. the errors were sorted and the one which is greater than 99.7% of the others was found. For 116×116 measurements with the Rollei plate this is the 40th value, and for 23×23 measurements with the planar plate it is the 2nd value. The ratio σ value to RMS varied between

2.6 and 3.4 with 3 being the average ratio. This shows that the maximum errors as expressed by the 3σ values do not vary a lot and are bounded to about 3 RMS. The errors higher than 3σ are very few and local. The ratio maximum absolute error to 3σ was also computed and this varies between 1 and 1.5 with 1.24 being the average value, showing that use of the maximum absolute error is in most cases sensitive to very few local errors and thus pessimistic.

The error patterns as shown in Fig. 1a and Fig. 2a are smooth with the exception of the fourth column in Fig. 1a. This systematic effect is probably due to stage calibration errors and can be improved with a better calibration plate (see also Section 3.4). The improvement in comparison to results with DSW200 (see Baltsavias et al., 1997; Fig. 2b) is very significant.

The results using only four control points were, as expected, worse but in accordance with the above statements. The RMS in x/y are (2.1–3.1)/(1.4–1.9) for the DS and (1.3–1.6)/(1.5–1.6) for the FS. The mean maximum absolute errors in x/y were (7.9–10.3)/(6.6–10.1) for DS and (5.1–6.8)/(5.6–7.3) for FS. A problem is the high mean x-value for DS showing a systematic bias, larger than half the RMS. When using the planar plate for calibration (results not listed in Table 1), the mean x-value drops below 1 μm, still another indication that a more accurate and denser calibration plate can improve the results.

3.2. Misregistration between colour channels

The results are summarised in Table 2 and one example is given in Fig. 1b. Generally the differences between the R, G, B channels are decreasing from (R minus B) to (G minus B) and (R minus G). The latter was a bit unexpected, since the sequence of scanning is R, G, B and vibrations should cause larger errors in the R channel. The mean x- and y-values are less than 1 μm, indicating that no systematic differences exist. There is no significant difference between the x- and y-directions with DS while with FS the x-errors are significantly larger. The results of FS in x are generally worse, which was expected, since vibrations cause more problems for this scanner as it stands on a concrete floor and the scanning movement is mainly in x. The RMS values in x for FS even exceed its RMS global accuracy values (1.3–1.4 μm; see Table 1). The repeatability (check difference between minimum and maximum values) is very good. Summarising, the RMS in x/y are (0.7–1.3)/(0.8–1.4) for the DS and (1.5–1.7)/(0.8–1.3) for the FS, while the maximum absolute errors vary between 2.5 and 5 μm.

3.3. Local geometric accuracy and repeatability

The affine transformation between pixel and reference coordinates was computed for each tile excluding the border tiles which had fewer grid crosses (64 tiles were used with 142 to 144 crosses each).

Table 2
Statistics of pairwise differences between pixel coordinates of the colour channels for the Rollei plate (in μm)

Scan version/No. of comparison points	Statistics	Red–Green channel			Red–Blue channel			Green–Blue channel		
		mean	min.	max.	mean	min.	max.	mean	min.	max.
DS scanner, average of 6 scans/13,441	RMS x	0.8	0.7	0.8	1.2	1.2	1.3	1.2	1.0	1.3
	RMS y	1.1	0.9	1.1	1.3	1.2	1.4	0.9	0.8	1.1
	mean x	−0.1	−0.1	0.0	−0.9	−1.0	−0.8	−0.8	−0.9	−0.7
	mean y	−0.1	−0.2	0.0	0.3	0.1	0.4	0.3	0.3	0.4
	max. abs. x	2.5	2.3	2.7	4.1	3.8	4.6	3.5	3.3	3.7
	max. abs. y	3.1	2.8	3.2	4.5	4.2	5.4	4.0	3.2	4.8
FS scanner, one scan/13,450	RMS x	1.7			1.5			1.5		
	RMS y	1.2			1.3			0.8		
	mean x	1.0			0.4			−0.6		
	mean y	−0.7			−0.7			0.0		
	max. abs. x	3.9			4.4			4.8		
	max. abs. y	3.1			3.7			2.8		

Table 3
Statistics of differences between pixel and reference coordinates after an affine transformation for each individual image tile (in μm)

Scan version/No. of control points per tile	Statistics	Red channel			Green channel			Blue channel		
		mean	min.	max.	mean	min.	max.	mean	min.	max.
First 3 scans/142–144	RMS x	0.8	0.7	1.0	0.9	0.7	1.0	0.9	0.7	1.0
	RMS y	0.7	0.6	0.9	0.7	0.6	0.9	0.8	0.6	0.9
	max. abs. x	2.7	1.9	6.0	2.7	1.9	5.7	2.7	1.9	4.1
	max. abs. y	2.1	1.5	5.7	2.1	1.6	5.7	2.2	1.5	6.0
Second 3 scans/142–144	RMS x	0.8	0.7	0.9	0.9	0.7	1.0	0.9	0.7	1.0
	RMS y	0.7	0.6	0.9	0.7	0.6	0.9	0.7	0.7	0.9
	max. abs. x	2.6	1.8	3.6	2.6	1.9	3.9	2.7	2.0	4.0
	max. abs. y	2.0	1.5	4.7	2.0	1.5	3.7	2.0	1.5	3.0

The statistics are computed from all tiles of all three scans of each colour channel. In all scans the Rollei plate, DS scanner and calibration with the DSW300 plate were used.

Table 4
Statistics of parameters of affine transformation from pixel to reference coordinates for each individual image tile (in μm)

Scan version/No. of control points per tile	Parameters	Red channel			Green channel			Blue channel		
		mean	min.	max.	mean	min.	max.	mean	min.	max.
First 3 scans/										
142–144	x-shift	1810.7	1798.5	1819.5	1810.5	1801.4	1820.7	1810.9	1797.5	1824.2
	y-shift	630.2	626.9	635.0	630.1	623.3	638.8	630.8	620.2	643.7
	x-scale	12.5021	12.5016	12.5026	12.5016	12.5011	12.5020	12.5009	12.5004	12.5014
	y-scale	−12.5039	−12.5045	−12.5035	−12.5033	−12.5040	−12.5029	−12.5027	−12.5034	−12.5023
	x-shear	0.0176	0.0172	0.0181	0.0176	0.0172	0.0181	0.0176	0.0172	0.0181
	y-shear	0.0178	0.0174	0.0183	0.0178	0.0174	0.0183	0.0178	0.0174	0.0184
	$\sigma 0$	0.8	0.7	0.9	0.8	0.7	0.9	0.8	0.7	1.0
Second 3 scans/										
142–144	x-shift	1916.7	1899.8	1928.4	1916.6	1902.8	1924.6	1917.0	1907.7	1927.6
	y-shift	735.9	728.9	743.2	735.8	731.6	740.5	736.7	731.0	745.4
	x-scale	12.5026	12.5021	12.5030	12.5021	12.5016	12.5025	12.5014	12.5009	12.5018
	y-scale	−12.5043	−12.5050	−12.5038	−12.5038	−12.5045	−12.5033	−12.5032	−12.5038	−12.5027
	x-shear	0.0285	0.0281	0.0290	0.0285	0.0281	0.0290	0.0285	0.0281	0.0290
	y-shear	0.0287	0.0282	0.0291	0.0287	0.0283	0.0291	0.0287	0.0282	0.02920
	$\sigma 0$	0.8	0.7	0.8	0.8	0.7	0.9	0.8	0.7	0.9

The statistics are computed from all tiles of all three scans of each colour channel. In all scans the Rollei plate, DS scanner and calibration with the DSW300 plate were used.

The results are shown in Tables 3 and 4 and two tile examples in Fig. 3. Errors due to mechanical positioning and vibrations are absorbed by the translations of the affine transformation. Thus, the errors in Table 3 and Fig. 3 represent mostly optical errors, especially lens distortion. For each scanner and each tile the errors are very similar for all scans and all colour channels. There are some changes in the error distribution from tile to tile (see Fig. 3) but their magnitude remains the same. Since lens distortion should not change, the only explanation is that this variation is due to spatial variation of the new thicker stage platen (such variations were not observed with the DSW200). The RMS errors are 0.6–1 μm, and the maximum errors are on the average 2–2.7 μm and can reach up to 6 μm. The largest errors (e.g. 5–6 μm) are due to matching errors because of the dust (see Fig. 3b). There is no significant dif-

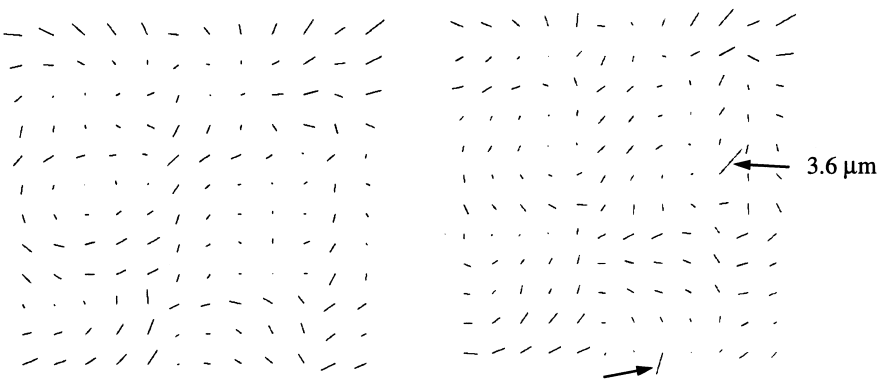

Fig. 3. Residuals of a local tile affine transformation from pixel to reference coordinates. Both plots show the residuals of two neighbouring tiles. The residual distribution changes slightly from tile to tile due to variations of the new thicker stage platen. The arrows show measurement errors by LSTM due to dust.

ference between the colour channels, and results in the y-direction are slightly better than in x. The fit between the first and the second three scans is excellent with the exception of the maximum values for the maximum absolute errors which are less in the second three scans due to continuous cleaning of the plate and thus less dust.

More interesting are the affine parameters of the individual tile transformations in Table 4. The translations of the affine transformation give the position of the origin of the pixel coordinate system ($x = 0$, $y = 0$) with respect to the origin of the reference coordinate system of the known crosses. For all tiles global coordinate systems referring to the whole plate are used. Both are at the centre of the plate and the pixel coordinates reach values between about -9200 to 9200. The translations vary in x/y by (21–27)/(8–24) μm for the first and (20–29)/(9–14) μm for the second scans and this might create the impression that the tiles are not accurately mosaicked. However, the shift variations are only partly due to shifts of individuals tiles (positioning errors, vibrations). These are mainly caused by variations in scale.

The x-, y-scales vary (maximum–minimum value) by about 0.001 for each channel and first or second scans. This shows the variability of scales from tile to tile when scanning one plate, whereby the differences of the mean scales between red and green, and green and blue channels, and between first and second scans are constant and about 0.0005–0.0006 (for an imaging scale factor of about 1.4 and a camera constant of 12 cm this scale variation can be caused by a vertical shift of the scanner stage of just 7–8 μm!). Thus, a scale variation of 0.001 would lead in the worst case (pixel coordinate = 9200) to a shift of 9.2 μm in a positive or negative direction (in Baltsavias et al., 1997, sect. 3.3 it was not clearly stated that this error — in that case up to 72 μm — does not occur in this magnitude, because the scales are multiplied in the worst case by half the tile size; see below). In addition, the scale range is not centred with respect to the scales of the geometric sensor calibration used to scan the plate and this causes an additional translation error. The variation of the translations is less in y than in x, exactly because the scale range in y is better centred over the y-scale of the sensor calibration. The calibration values for the x- and y-scale were 12.5022 and -12.504, respectively. Note that these values fit very well to the scales determined from the transformation of all crosses to their reference values using all crosses as control (the difference is 0.0002 and 0.0000 for the x- and y-scale).

The scanner software uses the scales and shears of the sensor calibration to transform from pixel to the stage coordinate system, so the scales are multiplied by maximum 992 (half the maximum image tile size), and the error due to scale variations is much smaller. For all scans (see Table 4) the x-/y-scales differ from the nominal scales by (-0.0018 to 0.0008)/(-0.0017 to 0.001) and this corresponds to maximum errors of (-1.8 to 0.8)/(-1.7 to 1) μm. Regarding the shears the following can be observed.

The values do not practically change from channel to channel and the difference between x and y is very small. Their variation for each channel and scan is as for the scales about 0.001. Their difference to the x-, y-nominal values corresponds to maximum errors of (−0.8 to 0.1)/(−0.6 to 0.4) μm for the first three scans and for the second ones are similar. Given the very high accuracy of the scanner as shown by the results of Table 1, these errors, particularly those due to scale variation, are not negligible. The manufacturer plans to improve the sensor calibration by using for its performance not just the central cross but a 3×3 grid of positions over the whole scan format.

3.4. A short note on the stage and sensor calibration

Two stage calibrations were performed on DS with the DSW300 plate, with a time interval of approximately 20 h. The corrections of the stage calibration were generally larger towards the borders (especially the right and bottom ones) and reached values up to 11–12 μm at the border rows and columns of the 15×15 correction grid which are extrapolated. The differences in the border rows/columns between the two calibrations were about 2–3 μm, with a maximum value of 4.5 μm. In the non-extrapolated 13×13 grid nodes, the corrections were less than 8–9 μm and the differences between the two calibrations were generally up to 2 μm. Also a calibration with the planar plate was performed. From its 23×23 crosses, 29×29 grid corrections were derived, i.e. the three border rows/columns on each side were extrapolated. However, the extrapolation errors, instead of increasing in comparison to the DSW300 plate due to the larger extrapolation distance, were smaller. The reason is that the corrections are much smoother, i.e. neighbouring corrections differ by 1.5 μm at the most, while for the DSW300 plate these were up to 4–5 μm. So, the planar denser grid plate results in smoother corrections (compare also Fig. 1a and Fig. 2a) and due to this also smaller extrapolation errors. In addition, the smaller grid spacing results in smaller errors when interpolating corrections, and the effect of measurement errors during the calibration is propagated in a smaller area. Thus, the manufacturer will consider employing a 25×25 grid plate with 1 cm grid spacing for the calibrations. In addition the line width will be increased to 4–5 pixels, which will make automatic measurement more accurate and robust with respect to dust and other noise.

We also checked the variation of the geometric sensor calibration within 1 day. The shears and the shifts of the affine transformation changed slightly but the scales changed up to 0.002. According to the manufacturer, scale changes above 0.001 (as well as corresponding changes in the correction values of the stage calibration) are mainly due to errors in the estimation of the affine transformation in the sensor calibration. These errors should be reduced by a new calibration procedure that will use more measurements distributed over a 3×3 grid covering the whole scan format, instead of using just the central grid cross as is currently the case. Scale variations below 0.001 are probably due to changes in the environmental conditions (temperature, etc.) and do not constitute a problem as long as the calibrations, when verified with the scanner software, give small residuals.

3.5. Geometric resolution

Taking all three scans into account, the smallest line group that could be sufficiently detected had a line width of 12 μm (line group 5/3) for both the vertical and horizontal directions (see Fig. 4). However, depending on the shift between pattern and sensor element, the line groups corresponding to 16 μm (group 5/1) and 14 μm (group 5/2) had a worse definition, and were partly undetectable. Vertical lines had a slightly worse definition than horizontal ones. In all scans the edges of long lines were well-defined (not noisy).

4. Evaluation of radiometric performance

4.1. Noise, linearity and dynamic range

The results from the grey scale wedge are shown in Table 5. The results of B/W scans are not shown here due to lack of space, but are very similar to those given in Table 5. The histograms for linear and logarithmic LUT are shown in Fig. 5. The numbers of samples for the 25 and 12.5 μm scans were 13,400 and 53,680, respectively. The numbers in

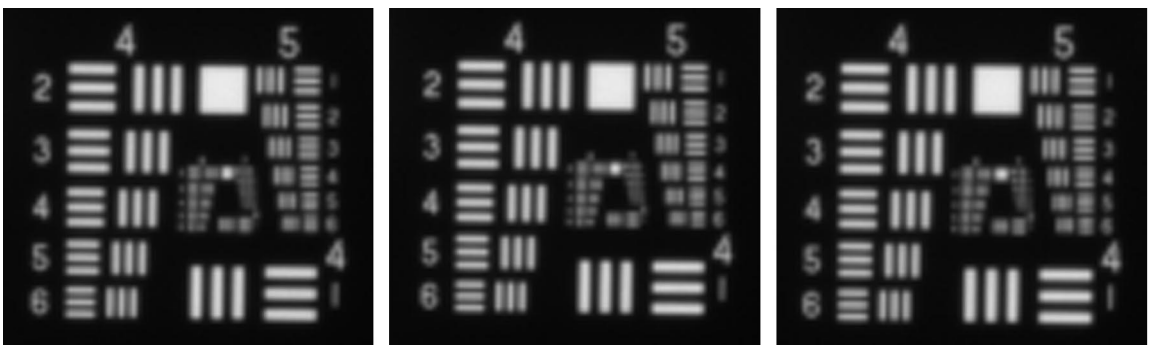

Fig. 4. The USAF resolution pattern scanned at position x/y (a), $x/y + 0.5$ pixel (b), $x + 0.5$ pixel$/y + 0.5$ pixel (c). The smallest detectable line groups in x-, y-direction (vertical–horizontal lines) are: 5/1–5/2 (a), 5/1–5/3 (b), and 5/3–5/3 (c). The x/y directions here represent the y/x directions of the original pixel coordinates.

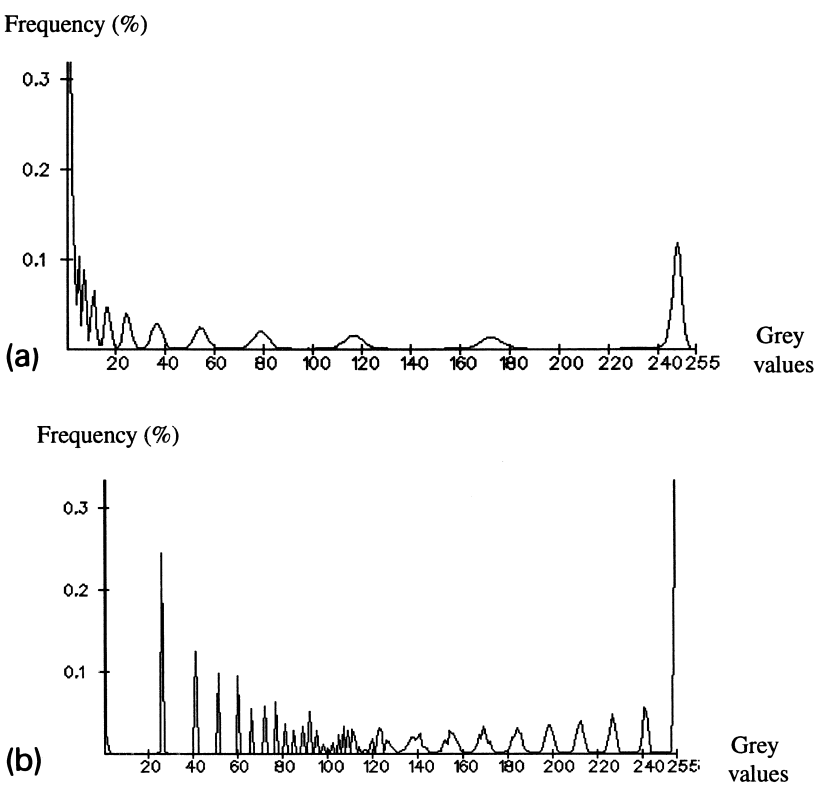

Fig. 5. (a) Histogram of grey scale for B/W scan, 12.5 μm, linear LUT. (b) Same as (a) but for logarithmic LUT. Frequencies scaled by a factor 4.

bold show the statistics for the maximum detectable density according to the criteria listed in Section 2. For the logarithmic LUT two values are listed for the following reason. As explained in Section 2, a test is applied to exclude blunders like dust, corn, etc., from the computation of the statistics. This test worked well for the linear LUT, where less than 1% of the samples were rejected and this only for the low densities. With the logarithmic LUT, however, the densities more than 2 D were strongly stretched, resulting in a grey level range for each wedge of 50 to 90 grey values, much more than

Table 5
Radiometric test with grey scale wedge. Mean and standard deviation of grey values. Maximum density that can be detected shown in bold. See explanation in text for values in bold italics

Density	Red channel						Green channel						Blue channel					
	12.5 μm linear LUT		25 μm linear LUT		12.5 μm log. LUT		12.5 μm linear LUT		25 μm linear LUT		12.5 μm log. LUT		12.5 μm linear LUT		25 μm linear LUT		12.5 μm log. LUT	
	mean	SD	mean	SD	mean [a]	SD [a]	mean	SD	mean	SD	mean	SD [a]	mean	SD	mean	SD	mean	SD [a]
0.055	242.6	1.8	244.9	1.3	254.3/254.3	0.5/0.5	240.8	2.1	240.1	1.5	253.6	0.5	240.2	2.4	240.4	1.7	253.6	0.5
0.214	168.2	3.3	170.0	2.6	240.9/240.9	0.8/0.8	168.3	4.0	167.6	2.9	240.3	1.0	167.8	4.3	167.8	3.1	240.3	1.1
0.375	113.3	3.2	114.4	2.5	226.3/226.3	1.1/1.1	114.8	3.8	114.2	2.8	226.1	1.3	114.7	4.2	114.7	3.0	226.3	1.5
0.530	76.4	2.6	76.6	2.1	211.5/211.5	1.4/1.3	78.1	3.3	77.5	2.3	211.9	1.6	78.3	3.7	77.7	2.5	212.0	1.8
0.690	52.6	2.1	52.7	1.6	197.7/197.7	1.5/1.5	54.4	2.6	53.8	1.9	198.5	1.8	54.6	3.0	54.0	2.0	198.7	2.0
0.840	35.4	1.6	35.3	1.3	183.2/183.2	1.8/1.8	37.1	2.1	36.6	1.5	184.3	2.1	37.3	2.4	36.8	1.6	184.6	2.4
0.997	23.6	1.3	23.4	1.0	167.9/167.9	2.0/2.0	24.9	1.6	24.5	1.1	169.6	2.4	25.0	1.8	24.6	1.2	169.8	2.7
1.160	15.9	1.0	15.7	0.8	153.3/153.3	2.2/2.2	17.0	1.2	16.5	0.9	155.2	2.6	17.0	1.4	16.6	0.9	155.4	2.9
1.320	10.4	0.7	10.1	0.6	137.2/137.2	2.6/2.6	11.2	0.9	10.8	0.7	139.8	3.0	11.2	1.0	10.8	0.7	139.9	3.3
1.485	6.9	0.6	6.6	0.5	122.0/122.0	2.8/2.8	7.5	0.7	7.1	0.5	124.9	3.3	7.5	0.8	7.1	0.6	124.5	3.7
1.660	4.7	0.5	4.3	0.5	107.2/107.2	3.3/3.3	5.1	0.6	4.7	0.5	110.1	3.6	5.1	0.6	4.7	0.5	109.9	4.0
1.830	3.0	0.3	2.8	0.4	89.4/ 89.4	4.1/4.0	**3.2**	**0.4**	2.9	0.3	92.7	4.1	3.2	0.5	2.9	0.3	92.8	4.3
2.000	**2.0**	**0.2**	**1.9**	**0.3**	74.0/ 74.2	5.0/4.3	2.1	0.4	**2.0**	**0.2**	78.0	4.0	**2.1**	**0.4**	**1.9**	**0.2**	77.5	4.1
2.160	1.2	0.4	1.0	0.1	***59.9***/ 59.5	***6.3***/5.0	1.5	0.5	1.1	0.3	***64.4***	***4.5***	1.5	0.5	1.1	0.3	***64.0***	***4.5***
2.305	1.0	0.1	0.9	0.2	42.4/ **45.7**	11.1/**5.0**	1.0	0.2	1.0	0.2	**46.8**	**4.9**	1.0	0.2	1.0	0.2	**46.5**	**5.0**
2.440	0.8	0.4	0.5	0.5	26.9/ 26.0	14.7/0.0	0.9	0.3	0.7	0.5	32.9	7.5	0.9	0.4	0.6	0.5	26.0	0.0
2.575	0.5	0.5	0.1	0.3	15.9/ 26.0	15.0/0.0	0.7	0.5	0.3	0.4	26.0	0.0	0.6	0.5	0.2	0.4	26.0	0.0
2.728	0.2	0.4	0.0	0.0	5.2/ 0.0	10.7/0.2	0.3	0.5	0.0	0.1	0.0	0.2	0.2	0.4	0.0	0.1	0.0	0.2
2.890	0.1	0.3	0.0	0.0	1.9/ 0.0	6.9/0.0	0.1	0.3	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
3.050	0.0	0.2	0.0	0.0	1.0/ 0.0	5.1/0.0	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
3.205	0.0	0.2	0.0	0.0	0.7/ 0.0	4.2/0.1	0.0	0.2	0.0	0.0	0.1	0.2	0.0	0.2	0.0	0.0	0.1	0.3
Mean SD [b]		1.0		0.9		3.6/2.2		1.3		0.9		4.5/2.7		1.4		1.0		4.9/2.5
Mean SD [b] (0.53–1.485 D)		1.4		1.1		2.0/2.0		1.8		1.3		2.4/2.4		2.0		1.4		2.7/2.7

[a] Values without/with the blunder test (see explanation in text). If only one value is given, it is with the blunder test.
[b] Excluding standard deviations that are 0.

the maximum allowable range for this test, which was 20. Thus, many grey values (up to 50%) were incorrectly excluded from the test, resulting in lower standard deviations, even zero ones when only 1 grey value was included in the allowable range. Thus, the maximum detectable density without this test was also estimated and this is indicated in bold italics. As an example, the statistics without and with this test are given for the red channel. The remaining channels and the B/W scan had similar values. However, the noise level (standard deviation) is given without/with blunder test for all scans.

The noise level is 1–2 and 2–3/4–5 (with/without blunder test) grey values for the linear and logarithmic LUT, respectively. For the densities 0.5–1.5 D, the linear LUT results in just 0.6 grey values less noise than the logarithmic one. The noise is increasing in the sequence: red, B/W, green, blue scan. For the linear LUT, the lowest density is partly saturated, and the noise is decreasing from low to high densities, while for the highest densities it gets very small due to saturation. For the logarithmic LUT, the noise increases up to about 2.5 D and then decreases again due to saturation. Actually when the standard deviation stops to decrease/increase (for linear/logarithmic LUT), this is an indication that one is close to the maximum detectable density. Scanning with 25 μm leads to a 15%–30% noise reduction. The dynamic range, according to the conditions of Section 2, is about 2 D for the linear LUT and 2.16 D/2.31 D (without/with blunder test) for the logarithmic LUT. It is slightly higher for red, while green and blue are similar. This does not mean that the signal beyond the maximum detectable density is useless. There is some information there, with the exception of the last 3–4 densities that are saturated. For example, visual inspection allows separability up to 2.6 D for both LUTs (for linear LUT using a gamma 3). Also all mean grey values were nicely decreasing even for the highest densities. The plots in Fig. 6 are in accordance with the above-stated maximum detectable density and show a good linear behaviour up to this density. They also show a very good fit between the three spectral and the B/W scan.

The mean values are very similar for all spectral channels (for both LUT versions), indicating a very good colour balance. They are also similar between the 12.5 and 25 μm scans. Use of a logarithmic LUT leads to a larger dynamic range, while visually the grey values representing the densities are almost equally spaced and thus more homogeneous. On the other hand, the noise in the higher densities is increasing a lot, the strong stretching of the dark grey values (the highest densities were each covering a range of 50–90 grey values) leads to a mixing of grey values from different densities (e.g. see grey level range 90–110 in Fig. 5b), and the most common range 0–1.5 D is occupying only half of the grey level range. Thus, a logarithmic LUT should be used with caution.

4.2. Artifacts and other radiometric problems

To detect artifacts visually, the contrast was strongly enhanced using Wallis filtering. However, the quantification of radiometric errors always occurs in the original images. Fig. 7 shows different artifacts. 'Electronic' dust is usually bright, since it is normally the dark dust that is sometimes not detected as such and then wrong bright corrections are applied. The two correction factors (gain and offset) are wrong, whereby the gain factor is usually (in the case of dark dust) larger than 1. An example is shown in Fig. 7c,d. The white dot causes a grey level error of up to 14 grey values when scanning the scanner stage glass plate, but with a gain factor larger than 1 the error would increase with increasing background grey value. The grey level variations across the tile borders in Fig. 7e are small, generally 1 and up to 4 grey levels, and are not visible in the original images. The sawtooth pattern of the radiometric feathering across the tile borders is done on purpose to make remaining radiometric differences along the seamline less easily visible. The number of sawtooth patterns along each seamline gives directly the difference in grey levels, e.g. 2 in Fig. 7e. Concluding, by far the major problem is dust, natural and 'electronic', whereby the latter can be reduced, if better algorithms are employed by the scanner software. Smaller problems, such as differences between tiles and artifacts can have an influence, if the image contrast is strongly enhanced, as this is sometimes done in order to generate more texture for matching in DTM generation and automatic point measurement in aerial triangulation.

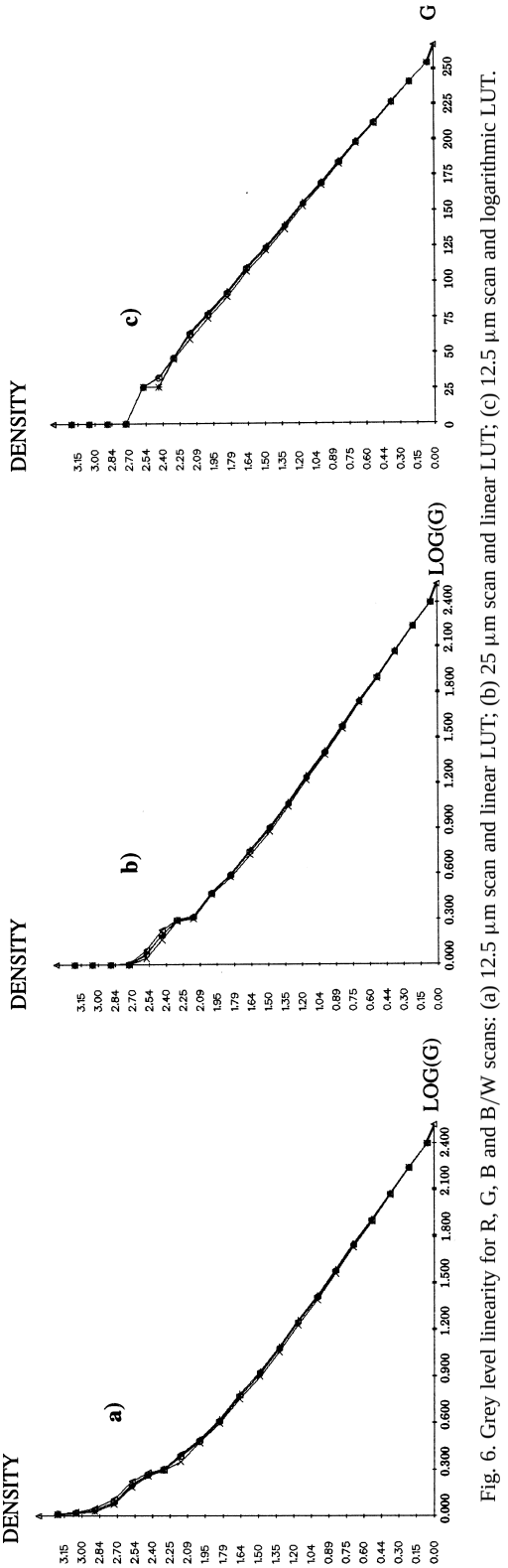

Fig. 6. Grey level linearity for R, G, B and B/W scans: (a) 12.5 μm scan and linear LUT; (b) 25 μm scan and linear LUT; (c) 12.5 μm scan and logarithmic LUT.

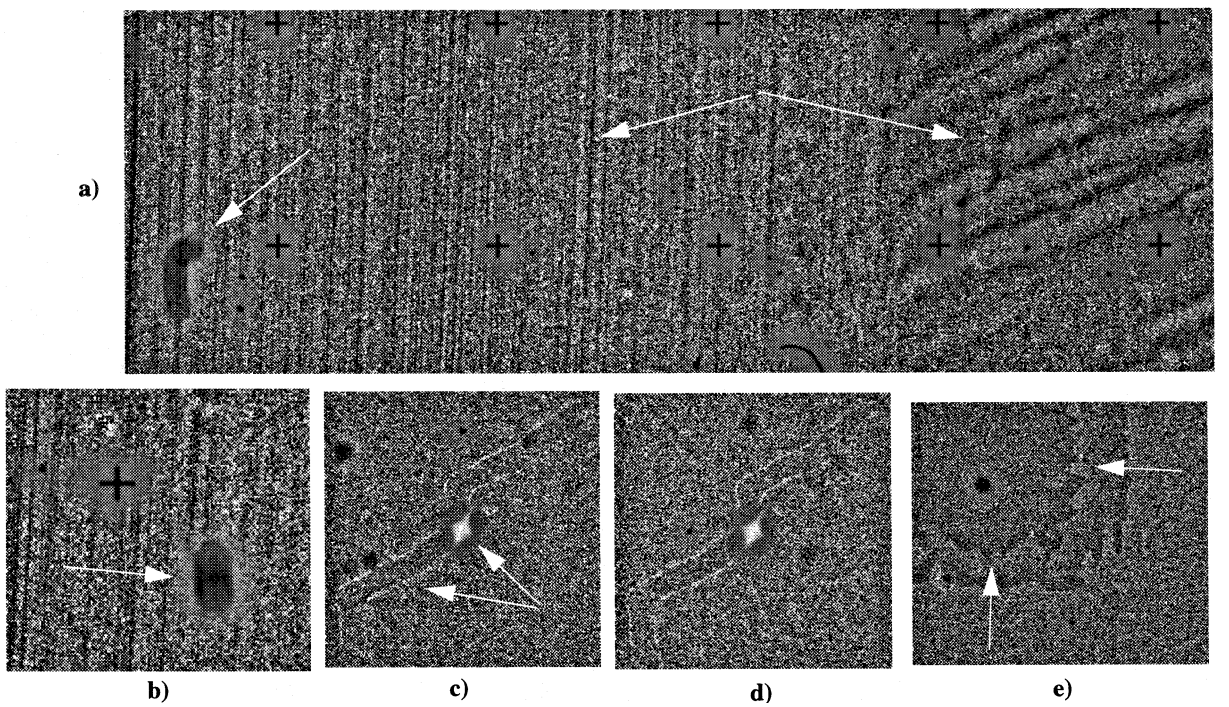

Fig. 7. Artifacts. (a) Unfocused dust, vertical stripes and strange linear structures in the left and right image part, respectively. These structures either exist in the Rollei grid plate or are due to interface effects between the Rollei and the scanner stage glass plate. (b) Black circles imaged twice. Lots of similar unfocused single dots, due to dust, can be observed over the whole format. They cause a grey value deviation of up to 20 grey values when scanning the scanner stage glass plate. (c) 'Electronic' dust. The white dot and scratches were observed in all image tiles of various scans at the same image tile position. (d) The same sensor position at a neighbouring tile. (e) Radiometric feathering between image tiles.

Some other problems mentioned in Baltsavias et al. (1997) referring to DSW200, such as echoes and spikes in the histograms, have not been observed in these tests.

5. Summary and conclusions

With respect to the geometric accuracy, the RMS was 1.3–1.9 μm and the mean maximum absolute error 4.5–8 μm. Lens distortion contributes to this error by an RMS of 0.7–0.9 μm (a priori calibration of the lens distortion and application of appropriate corrections could remove this error and lead to even more accurate results). The errors are bounded, i.e. on the average the 3σ (99.7%) values are 3 RMS, and the maximum absolute error 3.7 RMS. The co-registration accuracy of colour channels was about 1 μm, i.e. better than the geometric accuracy as it should be. The short and medium term repeatability was very high. With a linear LUT, the radiometric noise level is 1 and 1–1.5 grey values for 25 and 12.5 μm scan pixel size, respectively, and a logarithmic LUT, 3.5–5 grey values. The dynamic range is 2 D/2.16 D for linear/logarithmic LUT with a very good linear response up to this value. One of the major remaining radiometric problems is dust. In both geometric and radiometric tests no significant differences between R, G, B and B/W scans has been observed.

The results are in all aspects, and especially with respect to geometry, much better than all previous six tests with four DSW200 scanners. The improvement in radiometry is primarily due to better software algorithms. The geometric accuracy (RMS) is better than published tests and tests that some of the authors themselves have performed with other scanners. This high geometric accuracy was achieved with both colour and B/W scans, different grid plates, different scanner calibrations, multiple scans of the same plate and two different scanners under suboptimal

environmental conditions. Therefore, our confidence that these results are objective and repeatable is high. It must be noted, however, that such results can only be achieved, if scanners operate under proper environmental and maintenance conditions, and users perform all calibrations carefully and as often as they are required.

The use of the dense Rollei plate has been proven to be an invaluable tool in detecting, separating and modelling different error sources. Although it is too expensive to be supplied with the scanner to users, such a dense plate could be used by the manufacturers for diagnostic purposes. It could also be used, together with other good quality test patterns, to test each scanner before delivery and generate a quality assurance certificate, a kind of guarantee for the customer and also a measure against which he can compare the scanner performance after installation or in periodic tests.

The scanner can be further improved. In particular, the use of a planar 25×25 grid plate with 1 cm grid spacing and thick line width can improve the geometric accuracy through a more accurate stage calibration. The geometric sensor calibration also needs improvement to better model the variation of scale within the scan format. A longer stage settling when driving the stage to a certain position can improve the geometric accuracy, especially in x. In the radiometric equalisation, the algorithm for detection of dust should be improved to avoid creation of 'electronic' dust. Investigations on the most appropriate, depending on the application, reduction of 10- to 8-bit need to be performed. Finally, a fast prescan should be made possible. This will facilitate an overview image on which to select the scan area but also possibilities of finding automatically the darkest and lightest regions in an image and the setting of appropriate scan parameters. This is very important, especially for good quality unattended roll film scanning. The authors plan to continue their cooperation and tests, especially with respect to geometric and radiometric calibration, analysis of colour accuracy, and geometric tests on scanning with loaded film rolls. Such tests have been performed by LHS and led to an RMS increase by 0.5 μm.

A final remark on this unique cooperation between manufacturers, users and academia. Some important characteristics of this cooperation are: manufacturers who do not try to block or cover criticism but listen carefully, are cooperative, knowledgeable and implement improvements; users that care about the quality of the products they buy and use, and the products they generate with scanned images; academicians who try with their in-depth research and in cooperation with the industry to contribute in performance improvements of algorithms and equipment. The very significant improvements that have been achieved with the DSW300 in scanner hardware and software is proof that such a cooperation is fruitful and leads to benefits for all.

Acknowledgements

The authors would like to thank Dr. Karsten Jakobsen, Hannover University, for kindly providing the Rollei grid plate.

References

Baltsavias, E.P., 1991. Multiphoto Geometrically Constrained Matching. Ph.D. dissertation, Institute of Geodesy and Photogrammetry, ETH Zurich, Mitteilungen 49, 221 pp.

Baltsavias, E.P., 1994. Test and calibration procedures for image scanners. Int. Arch. Photogramm. Remote Sensing 30 (1), 163–170.

Baltsavias, E.P., 1998. Photogrammetric scanners — survey, technological developments and requirements. Int. Arch. Photogramm. Remote Sensing 32 (1), 44–52.

Baltsavias, E.P., Kaeser, Chr., 1998. Evaluation and testing of the Zeiss SCAI roll film scanner. Int. Arch. Photogramm. Remote Sensing 32 (1), 67–74.

Baltsavias, E.P., Haering, S., Kersten, Th., 1997. Geometric and radiometric performance evaluation of the Leica/Helava DSW200 photogrammetric film scanner. Proc. of Videometrics V Conference, 30–31 July, San Diego. Proc. Soc. Photo-Opt. Instrum. Eng. 3174, 157–173.

Bethel, J., 1994. Calibration of a Photogrammetric Image Scanner. Technical Papers of ASPRS/ACSM Annual Convention, 25–28 April, Reno, Vol. 1, pp. 81–88.

Bethel, J.S., 1995. Geometric alignment and calibration of a photogrammetric image scanner. ISPRS J. Photogramm. Remote Sensing 50 (2), 37–42.

Bolte, U., Jakobsen, K., Wehrmann, H., 1996. Geometric and radiometric analysis of a photogrammetric image scanner. Int. Arch. Photogramm. Remote Sensing 31 (B1), 72–77.

Dam, A., Walker, A.S., 1996. Recent developments in digital photogrammetric systems from Leica–Helava. Int. Arch. Photogramm. Remote Sensing 31 (B2), 66–71.

Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique. S. Afr. J. Photogramm., Remote Sensing Cartogr. 14 (3), 175–187.

Gruen, A., Slater, P.N., 1983. A test strategy for high resolution image scanners. Dep. of Geodetic Science and Surveying, Ohio State Univ., Columbus, Rep. 350.

Jakobsen, K., Gaffga, R., 1998. Calibration of the photogrammetric image scanner Rastermaster RM1. Int. Arch. Photogramm. Remote Sensing 32 (1), 75–79.

Kersten, Th., Haering, S., 1997. Automatic interior orientation of digital aerial images. Photogramm. Eng. Remote Sensing 63 (8), 1007–1011.

Koelbl, O., Bach, U., 1996. Tone reproduction of photographic scanners. Photogramm. Eng. Remote Sensing 62 (6), 687–694.

Leberl, F., Best, M., Meyer, D., 1992. Photogrammetric scanning with a square array CCD camera. Int. Arch. Photogramm. Remote Sensing 29 (2), 358–363.

Miller, S., Dam, A., 1994. Standards for image scanners used in digital photogrammetry. Proc. of ISPRS Comm. II Symposium. Int. Arch. Photogramm. Remote Sensing 30 (2), 134–136.

Roos, M., 1993. The image digitising system (IDS). Technical Papers of ACSM/ASPRS Annual Convention, 15–18 February, New Orleans, LA, Vol. 2, pp. 301–309.

Seywald, R., 1996. On the automated assessment of geometric scanner accuracy. Int. Arch. Photogramm. Remote Sensing 31 (B1), 182–186.

Seywald, R., 1997. Automated and Interactive Procedures for Efficient and Objective Evaluation of Film Scanners. Publications of OCG (Austrian Computer Society, Wollzeile 1–3, A-1010 Vienna, Austria; ocg@ocg.or.at), Vol. 98, 160 pp.

Seywald, R., Leberl, F., Kellerer, W., 1994. Requirements of a system to analyze film scanners. Int. Arch. Photogramm. Remote Sensing 30 (1), 144–149.