facturers should improve their products, stability and quality control, and most important users should thoroughly evaluate and test scanners before buying and also during their use.

ACKNOWLEDGEMENTS

REFERENCES

1. E. P. Baltsavias, and R. Bill, "Scanners - a survey of current technologies and future needs", *International Archives of Photogrammetry and Remote Sensing*, Vol. 30, Part 1, pp. 130-143, 1994.

2. E. P. Baltsavias, "Multiphoto geometrically constrained matching", Ph. D. dissertation, *Institute of Geodesy and Photogrammetry, ETH Zurich, Mitteilungen No. 49*, 221 p., 1991.

3. E. P. Baltsavias, "Test and calibration procedures for image scanners", *International Archives of Photogrammetry and Remote Sensing*, Vol. 30, Part 1, pp. 163-170, 1994.

4. E. P. Baltsavias, and B. Waegli, "Quality analysis and calibration of DTP scanners", *International Archives of Photogrammetry and Remote Sensing*, Vol. 31, Part B1, pp. 13-19, 1996.

5. J. S. Bethel, "Geometric alignment and calibration of a photogrammetric image scanner", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 50, No. 2, pp. 37-42, 1995.

6. A. Dam, and A. S. Walker, "Recent developments in digital photogrammetric systems from Leica-Helava", *International Archives of Photogrammetry and Remote Sensing*, Vol. 31, Part B2, pp. 66-71, 1996.

7. A. Gruen, "Adaptive least squares correlation: a powerful image matching technique", *South African Journal of Photogrammetry, Remote Sensing and Cartography,* Vol. 14, No. 3, pp. 175-187, 1985.

8. U. Bolte, K. Jakobsen, and H. Wehrmann, "Geometric and radiometric analysis of a photogrammetric image scanner", *International Archives of Photogrammetry and Remote Sensing*, Vol. 31, Part B1, pp. 72-77, 1996.

9. Th. Kersten, and S. Haering, "Automatic interior orientation of digital aerial images", *Photogrammetric Eng. and Remote Sensing*, Vol. 63, No. 8, 1997 (to be published).

10. O. Koelbl, and U. Bach, "Tone reproduction of photographic scanners", *Photogrammetric Eng. and Remote Sensing*, Vol. 62, No. 6, pp. 687-694, 1996.

11. F. Leberl, M. Best, and D. Meyer, "Photogrammetric scanning with a square array CCD camera", *International Archives of Photogrammetry and Remote Sensing*, Vol. 29, Part 2, pp. 358-363, 1992.

12. S. Miller, and A. Dam, "Standards for image scanners used in digital photogrammetry", *International Archives of Photogrammetry and Remote Sensing*, Vol. 30, Part 2, pp. 134-136, 1994.

13. R. Seywald, F. Leberl, and W. Kellerer, "Requirements of a system to analyze film scanners", *International Archives of Photogrammetry and Remote Sensing*, Vol. 30, Part 1, pp. 144-149, 1994.

14. R. Seywald, "On the automated assessment of geometric scanner accuracy", *International Archives of Photogrammetry and Remote Sensing*, Vol. 31, Part B1, pp. 182-186, 1996.

Examination of the histograms showed that for small grey values certain values occur more frequently (see Figure 7a) and Table 8). This occurred in all observed histograms for the same grey values but only for the 12.5 μm scans. In previous tests with negative films, we had observed that certain high grey values (high, since the LUT was inverted to generate a positive digital image) occurred more frequently (but in that case every fourth grey value). We believe that this problem relates to the conversion of 10- to 8-bit. The behaviour in the current tests could be explained by the application of a logarithmic function during the 10- to 8-bit conversion to stretch the dark grey values and a subsequent grey level interpolation when assembling the tiles to an image. This effect is not visible for scans with pixel size larger than 12.5 μm, maybe due to the pixel averaging. Note that the histograms were generated by two own programmes. The scanner creates automatically for each image a histogram file. But this file was totally wrong for the 25 μm scan pixel size of the grey scale wedge, and has only some correct values for the 12.5 μm scans. In some histogram files we have even observed a number of pixels that exceeded by far the actual one.

Table 8: Part of a histogram. The grey value frequencies are divided by 1000. Some values occur more often than others. Latter are shown in bold. The same pattern occurs in all observed histograms of images scanned with 12.5 μm.

Gr. Val.	#	Gr. Val.	#	Gr. Val.	#	Gr. Val.	#	Gr. Val.	#	Gr. Val.	#	Gr. Val.	#
0	21817	10	900	20	290	30	94	40	72	50	47		
1	20	**11**	30	**21**	30	**31**	24	**41**	36	51	54		
2	20	**12**	37	**22**	33	**32**	22	42	83	52	64		
3	48	**13**	45	23	217	33	59	**43**	36	**53**	12		
4	98	14	589	**24**	34	**34**	14	44	82	54	77		
5	1726	**15**	29	**25**	35	35	45	45	65	55	78		
6	17	**16**	31	26	168	**36**	17	**46**	22	56	70		
7	25	17	401	**27**	38	37	47	47	53	57	60		
8	43	**18**	25	28	136	38	54	48	46	**58**	13		
9	66	**19**	31	**29**	33	**39**	26	**49**	11	59	52		

5. SUMMARY AND CONCLUSIONS

The major problems of DSW200 refer to the geometry. The major error sources are vibrations, mechanical positioning accuracy and the related scanner calibrations (stage and sensor) and the random line shift (electronic error) for some 4.2i Kodak models, while the errors due to the optical system are smaller and stable. Vibrations and electronic errors are random, and even the mechanical positioning errors vary within a few hours. There are significant error variations between different DSW200 scanners. The errors are always larger in y, and increase as we go from blue to green to red channel. The differences between colour channels (ignoring the random line shift) can amount up to 6 μm. The biggest problem are the maximum errors that affect whole tiles. These tile shifts have serious effects on subsequent operations with aerial images like interior orientation, aerial triangulation, DTM, orthoimage generation etc. A decrease of the errors could be achieved by a more stable scanner stage and slower scan speed in order to achieve a good repeatability and lack of vibrations. In addition, a grid plate with denser crosses (e.g. 5 mm) could be used for calibration. This would allow a better modelling of the positioning errors and a determination of the sensor calibration (pixel size) for each individual tile. In addition, a dense grid plate allows a better understanding of the distribution and nature of errors.

Regarding the radiometric errors, the noise level for 12.5 μm scans was ca. 3 grey levels for this test, but in most older tests was rather 2 grey levels. The dynamic range lies between 0.1-0.2D and 1.6-1.9D (incl. older tests) for 12.5 μm scans. The difference in dynamic range and noise level between colour channels is small. The use of logarithmic LUTs is problematic and should be avoided. The blemished pixels are very few and do not constitute a problem. Dust is a more serious problem and can lead to "electronic" dust, if the radiometric equalisation procedure of the sensor elements is not improved. The spikes in the dark region of the histograms should be avoided.

As it has been shown, scanners can have various geometric and radiometric errors that are beyond the manufacturer specifications. Results with other photogrammetric scanners, which can not be published yet, have shown that geometric and/ or radiometric errors occur with all scanners, and with some of them geometric or radiometric errors are larger than those of the DSW200. Researchers need to increase their activities on scanner evaluation, calibration and better design, manu-

Table 7: Statistics of the pairwise differences between the four scans of the green channel (in grey values).

|	1 - 2	1 - 3	1 - 4	2 - 3	2 - 4	3- 4
Mean	1.11	1.12	1.12	1.11	1.11	1.12
Stand. Dev.	0.94	0.95	0.95	0.94	0.94	0.95
RMS	1.45	1.47	1.47	1.46	1.46	1.46
Max	20	97	20	84	14	92

4.3. Artifacts

To detect blemishes we used the scans of the scanner stage plate. For each tile of each colour channel we defined a range of acceptable grey values as (mean ± sqrt(mean)). All grey values outside this range should be due to blemishes or dust. To separate dust from dark blemished pixels one could check whether they occur at the same pixel positions for spatially varying tiles. However, the amount of pixels outside the acceptable range was so small (1-3 for dark, 2 - 18 for light "blemishes") that we did not follow this approach. The scanner performs as mentioned a radiometric equalisation of the sensor element output that can correct the blemished pixels. However, dust can lead to wrong corrections values, causing "electronic" dust (see Figure 9a).

The grey level variations across the tile borders in Figure 9b) are within the noise level of the sensor. However, the saw tooth pattern of the radiometric feathering across the tile borders is peculiar. The interference pattern in Figure 9c) occurs at the corners of the grid plate. When dark horizontal signal exists, it is "echoed" in an inverse fashion in regions of light signal along the same image lines. This does not happen in vertical direction. The change of grey values across this echo is just 1.5 grey values. To visually detect artifacts we strongly enhance the contrast by Wallis filtering. However, the quantification of radiometric errors always occurs in the original images.
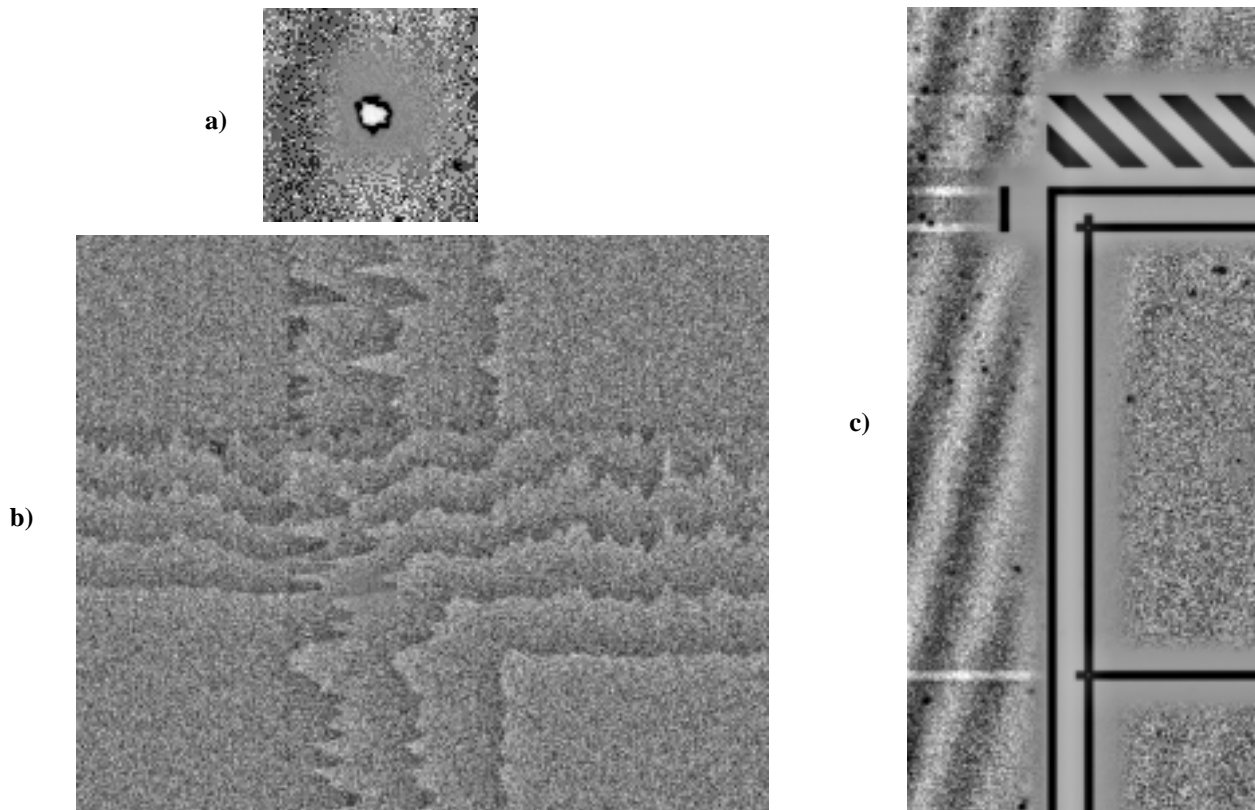
Figure 9. Artifacts: a) "electronic" dust ; b) radiometric feathering at the borders of four image tiles ; c) ETH plate: interference patterns and inverse "echoes" at the prolongation of black horizontal lines. All images are strongly enhanced by Wallis filtering.
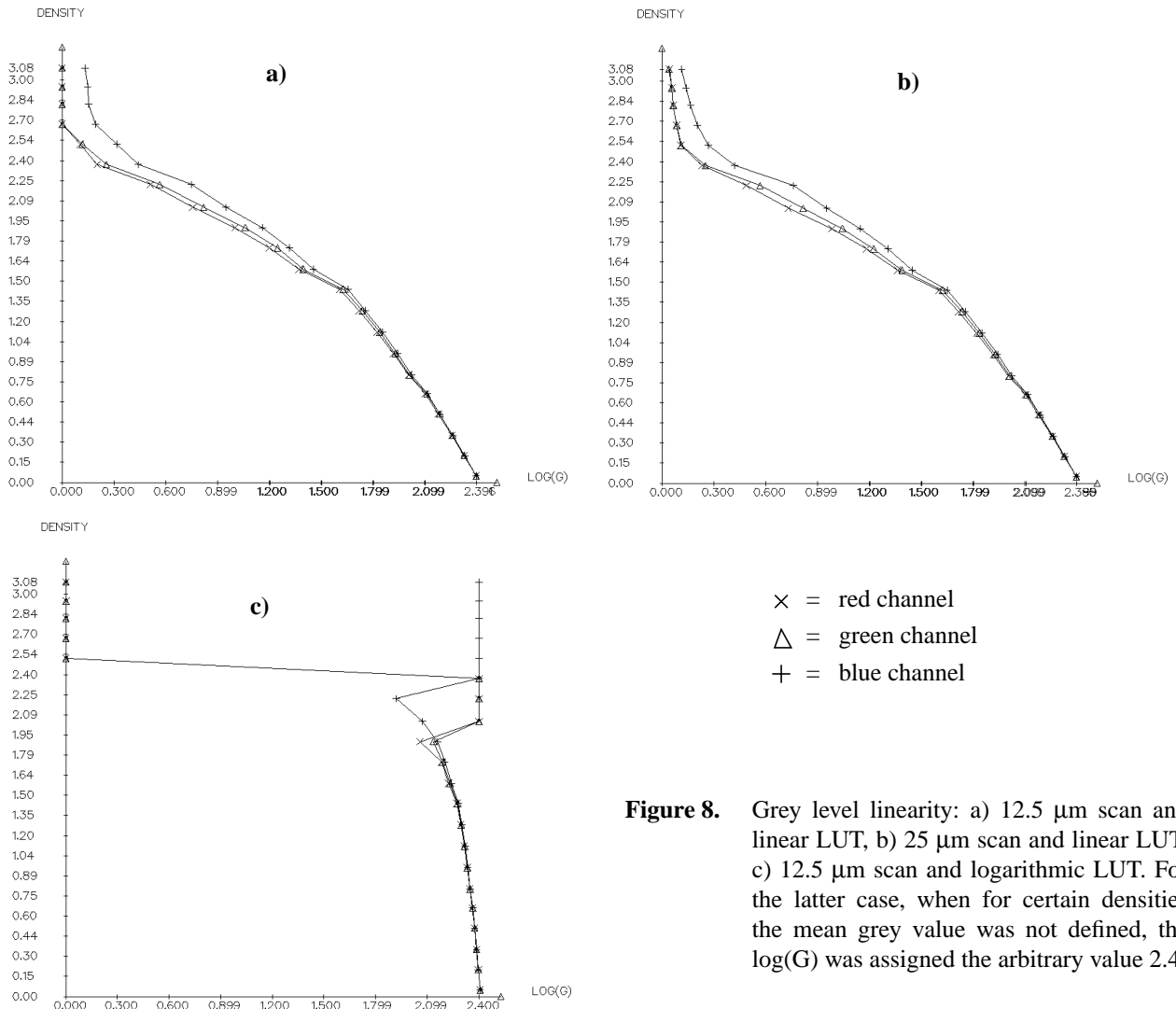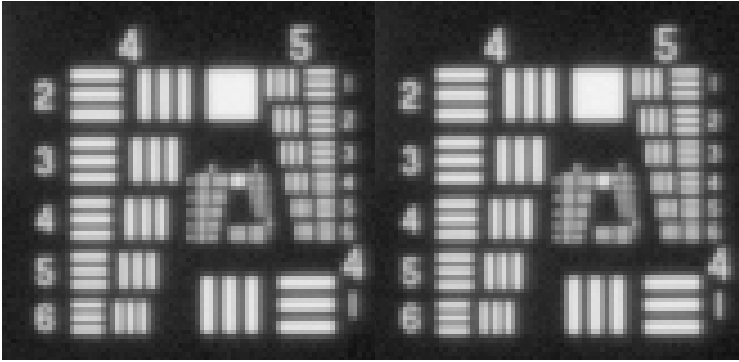
Figure 8. Grey level linearity: a) 12.5 μm scan and linear LUT, b) 25 μm scan and linear LUT, c) 12.5 μm scan and logarithmic LUT. For the latter case, when for certain densities the mean grey value was not defined, the log(G) was assigned the arbitrary value 2.4.

4.2. Scanner stage glass plate

The exposure time was set such that the glass plate had an average grey value of 159. The variation of grey level statistics between the four scans for each colour channel was irrelevant (maximum 0.2 grey values for the mean and 0.1 for the standard deviation). The variation between the colour channels for all four scans was also very small (maximum 0.9 for the mean and 0.4 for the standard deviations). For each scan and each channel the variation between the four tiles was maximum 0.6 for the mean and 0.2 for the standard deviation. The minimum values were 119-127 and the maximum 255. These large deviations from the mean are due to dust (real or "electronic", see section 4.3) and blemishes. The standard deviations (noise level) for the red, green and blue channels was 1.3, 1.6 and 1.5 grey values respectively. These values are ca. 1.5 grey values less than the values in Table 6 for similar mean grey value (0.51D). Statistics of the pair-wise differences between the four scans (including all tiles) were also computed. Table 7 shows the results for the green channel. The RMS includes the sensor noise and its temporal variation as well as illumination variations. The maximum differences occur when there is new dust between the scans or when there is dust (real or "electronic") and positioning errors between the scans. The large maximum values in the differences involving the third scan are due to blemishes (occur at same pixel coordinates of all tiles) with value 255. Inexplicably, these blemishes, do not appear in the other three scans.

Table 6: Radiometric test with grey scale wedge. Mean and standard deviation of grey values. Maximum density that can be detected shown in bold.

| Density	Red channel 12.5 μm linear LUT Mean	St.D.	Red channel 25 μm linear LUT Mean	St.D.	Red channel 12.5 μm log. LUT Mean	St.D.	Green channel 12.5 μm linear LUT Mean	St.D.	Green channel 25 μm linear LUT Mean	St.D.	Green channel 12.5 μm log. LUT Mean	St.D.	Blue channel 12.5 μm linear LUT Mean	St.D.	Blue channel 25 μm linear LUT Mean	St.D.	Blue channel 12.5 μm log. LUT Mean	St.D.
0.05	248.2	1.1	247.5	0.9	254.1	0.3	245.3	1.4	244.7	1.1	254.0	0.2	245.4	1.4	244.4	1.1	254.0	0.2
0.2	211.8	2.1	211.1	1.6	248.5	0.5	209.6	2.5	208.9	1.8	248.1	0.5	210.8	2.6	209.9	1.9	248.3	0.6
0.35	179.9	2.4	179.2	1.8	242.5	0.6	179.4	2.8	178.8	2.0	242.4	0.7	180.0	3.1	179.2	2.1	242.6	0.7
0.51	150.3	2.7	149.7	2.0	236.0	0.7	151.0	3.1	150.4	2.2	236.2	0.8	151.8	3.3	151.0	2.3	236.3	0.9
0.66	125.9	2.7	125.3	2.1	229.4	0.8	127.4	3.2	126.8	2.3	229.8	0.9	128.6	3.4	127.8	2.4	230.2	1.0
0.8	100.5	3.2	99.8	2.6	221.5	1.4	101.0	3.5	100.6	2.6	221.8	1.5	103.6	3.7	103.0	2.9	222.6	1.5
0.96	82.0	2.9	81.4	2.2	213.5	1.4	83.1	3.2	82.7	2.3	214.1	1.4	85.9	3.4	85.3	2.5	215.3	1.5
1.12	65.7	3.1	65.1	2.3	205.4	1.7	67.3	3.3	66.9	2.4	206.4	1.7	70.0	3.5	69.4	2.6	207.8	1.8
1.28	51.4	3.1	50.8	2.1	196.1	2.2	53.2	3.3	52.8	2.3	197.7	2.3	55.7	3.4	55.1	2.3	199.2	2.3
1.44	39.4	3.3	39.0	2.3	186.2	3.0	41.0	3.6	40.6	2.4	187.8	3.1	43.8	3.6	43.2	2.5	190.3	3.0
1.59	**22.5**	**3.9**	22.0	2.6	166.9	4.9	**23.8**	**3.9**	23.3	2.5	169.5	4.2	**27.3**	**3.9**	26.8	2.7	173.6	4.5
1.75	14.9	4.3	14.2	2.6	151.9	4.3	16.5	3.9	15.7	2.6	153.1	4.4	19.5	4.6	19.2	2.8	158.4	5.7
1.9	9.1	5.0	**8.6**	**2.8**	113.3	4.1	10.4	5.1	**10.0**	**2.8**	135.7	1.9	13.4	5.2	**13.0**	**3.1**	143.4	6.7
2.05	4.7	4.6	4.4	2.5	-	-	5.6	4.7	5.5	2.7	-	-	7.8	5.5	8.0	3.1	116.9	0.9
2.22	2.2	3.4	2.1	1.8	-	-	2.7	3.6	2.7	2.0	-	-	4.6	4.6	4.7	2.7	82.3	1.1
2.37	0.6	1.6	0.7	0.9	-	-	0.8	1.9	0.8	1.0	-	-	1.8	2.9	1.6	1.8	-	-
2.52	0.3	1.1	0.3	0.6	0.0	0.0	0.3	1.2	0.3	0.6	0.0	0.0	1.1	2.3	0.9	1.3	-	-
2.67	0.0	0.0	0.2	0.5	0.0	0.0	0.0	0.2	0.2	0.5	0.0	0.0	0.6	1.3	0.6	1.1	-	-
2.82	0.0	0.0	0.2	0.4	0.0	0.0	0.0	0.1	0.2	0.4	0.0	0.0	0.4	1.0	0.5	0.9	-	-
2.95	0.0	0.0	0.1	0.4	0.0	0.0	0.0	0.0	0.1	0.4	0.0	0.0	0.4	1.0	0.4	0.8	-	-
3.09	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.4	1.0	0.3	0.7	-	-
Mean St. D.[1]		2.7		1.7		1.7		2.9		1.8		1.8		3.1		2.1		2.2
Mean St. D.[1] (0.51-1.44D)		3.0		2.2		1.6		3.3		2.4		1.7		3.5		2.5		1.7

Figure 6. The USAF resolution pattern with the smallest detectable line group 5/3 in the first scan (left) and 5/2 in the second scan (right).

## 4. EVALUATION OF RADIOMETRIC PERFORMANCE

4.1. Noise, linearity and dynamic range for each colour channel

The results from the grey scale wedge are shown in Table 6. Use of a logarithmic LUT leads to various problems. The useful density range 0.2D - 1.9D is compressed to about half of the grey level range. In addition, some grey levels occur more frequently resulting in a saw tooth histogram (see Figure 7b)). The "-" in the table represent wedges where no grey value within the maximum allowable range (mean ± 10) occurs. In this case, the grey values were 0 or took values in the range 80 to ca. 225. The same occurred with the highest densities that have a mean of 0, but in this case 0 was the only value included in the allowable range. For the linear LUT we can note the following. The noise is small for 0.05D due to partial saturation. For low and middle densities the noise is higher than for the high densities, since latter are to a large extent saturated. The maximum detectable density is 1.59D for 12.5 µm and 1.9D for 25 µm. The mean values are similar for 12.5 and 25 µm scan pixel size, but latter has, as expected, lower noise (mean noise smaller by one grey value). The red channel has slightly larger grey value range and saturation and lower noise than the green channel, and this than the blue one. These results are similar to all previous tests (however they were with B/W scans) with two exceptions. In this test, firstly the noise is higher by 1 - 1.5 grey values. A part of this additional noise is due to the fact that the grey scale was scanned in two long horizontal stripes of image tiles because of a large top border, while previously only in one. Thus, all wedges included at least two image tiles. The tiles have slightly different grey values especially at the borders and this increases the standard deviation of the grey values. Secondly, the noise for the density range 1.59 - 2.22D is much higher than in old tests and higher than the one for the low densities, which previously always had the highest noise. This is also observed in the linearity plots in Figure 8a) and b). After 1.44D there is a change in the line direction and the points continue at approximately equal distances until 2.22D.
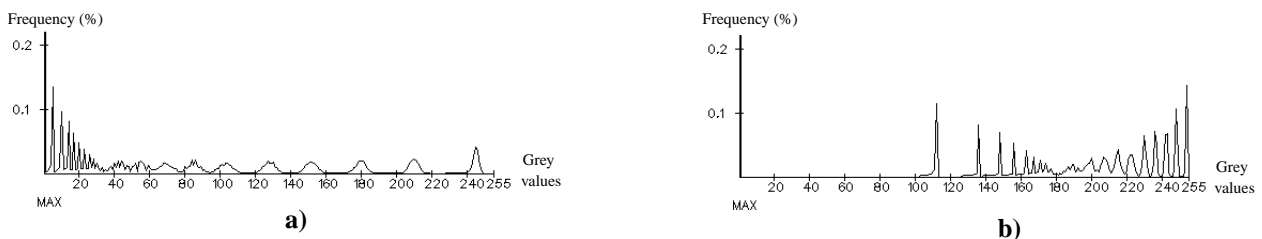
Figure 7. a) histogram of grey scale for green channel, 12.5 µm, linear LUT ; b) same as a) but for logarithmic LUT.

12

pixel and reference coordinate systems are) the translations gradually increase/decrease. The scanner determines in the sensor calibration the scales just from one central cross. But as the results of Table 5 show, the scales vary a lot from tile to tile. Using the wrong scales from the sensor calibration leads to overlap (or gaps) between neighbouring tiles. This is clearly shown in Figure 4b) where a cross at the border of two tiles is imaged twice. The overlap of the two tiles is ca. one pixel. We also controlled the scales and shears of the global transformation which fit very well to the respective values of the sensor calibration, and compared them to those of the local transformations. The shears differ on the average by 0.001 and maximum 0.002 for both scanners. Larger differences occur with the scales. The x/y scales differ for Polaris 0.0014/ 0.0006 on the average, and maximum 0.0032/0.0027. But for Antares the x/y scale differences are 0.0064/0.0072 on the average, and maximum 0.0083/0.0090! For a distance from the origin of 10,000 pixels as at the borders of the grid plates, the average scale difference would lead to a shift of 64 and 72 μm in x and y respectively. Although Polaris has a larger variation of scales in the local transformations, these scale values fit much better to the global and sensor calibration values. These considerations fit well with the behaviour of the residuals in Figures 1a) and 2a).
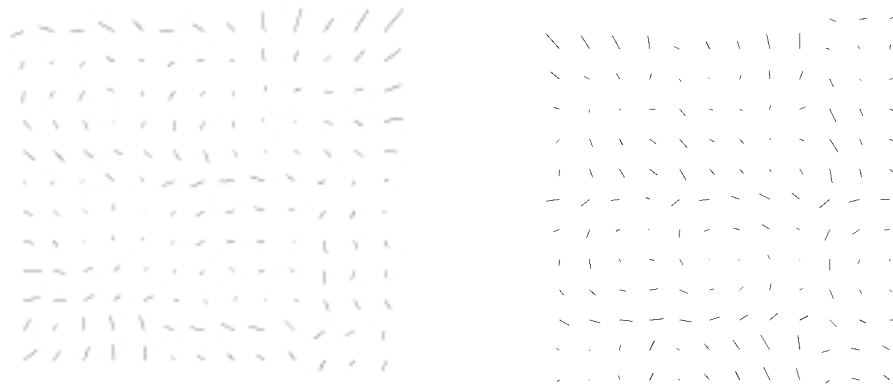
Figure 5. Residuals of a local (referring to one tile) affine transformation from pixel to reference coordinates. Left Antares, right Polaris. In both plots vectors are enlarged by 1200. The residuals are locally systematic and have similar magnitude and pattern for both scanners.

3.4. A short note on the stage calibration

On Polaris we performed two stage calibrations before and after scanning the grid plates with a time interval of a few hours. The corrections of the stage calibration were generally larger towards the borders (especially the top and bottom one) and reached values up to 10 μm (excluding the border rows and columns of the 15 x 15 correction grid which are extrapolated). The differences between the border and middle rows of the 13 x 13 grid of the two calibrations were up to 4.5 μm. This poses questions on the repeatability of the local positioning errors and the possibility to correct them by an offline calibration.

3.5. Geometric resolution

For the first scan the smallest line group that could be sufficiently detected had a line width of 12 μm for both vertical and horizontal direction. However, the line groups corresponding to 18 and 16 μm had a worse definition, although they were also detectable. For the second scan the smallest line group that could be sufficiently detected had a line width of 14 μm. Vertical lines had a slightly worse definition that horizontal ones. Again line groups corresponding to 18 and 16 μm had a worse definition. In both scans the edges of long lines were well-defined (not noisy).

Table 5: Statistics of parameters of affine transformation from pixel to reference coordinates for each individual image tile (in μm). The statistics are computed from all tiles of all four scans of each colour channel.

Scanner	Affine parameters	Red channel			Green channel			Blue channel		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Antares	# of points/tile	152.4	143.0	169.0	152.4	142.0	169.0	152.3	143.0	169.0
	x-shift	-1152.1	-1199.0	-1109.2	-1150.8	-1195.1	-1111.1	-1150.3	-1189.0	-1115.0
	y-shift	1397.7	1335.2	1466.4	1397.9	1338.4	1463.1	1398.4	1342.7	1459.0
	x-scale	12.4964	12.4956	12.4972	12.4959	12.4952	12.4967	12.4952	12.4945	12.4960
	y-scale	-12.4972	-12.4979	-12.4966	-12.4967	-12.4974	-12.4961	-12.4961	-12.4968	-12.4955
	x-shear	3.3e-2	3.3e-2	3.4e-2	3.3e-2	3.3e-2	3.4e-2	3.3e-2	3.3e-2	3.4e-2
	y-shear	3.3e-2	3.2e-2	3.3e-2	3.3e-2	3.2e-2	3.3e-2	3.3e-2	3.2e-2	3.3e-2
	sigma 0	0.81	0.69	0.98	0.83	0.72	0.99	0.84	0.73	1.01
Polaris	# of points/tile	152.1	144.0	169.0	152.0	144.0	169.0	152.0	142.0	169.0
	x-shift	-99.4	-121.2	-82.6	-99.2	-119.0	-84.4	-99.1	-118.5	-85.3
	y-shift	2625.8	2611.1	2641.1	2625.8	2610.6	2642.2	2626.0	2608.7	2641.7
	x-scale	12.5004	12.4980	12.5022	12.4999	12.4975	12.5021	12.4995	12.4971	12.5014
	y-scale	-12.4996	-12.5017	-12.4973	-12.4992	-12.5014	-12.4968	-12.4988	-12.5008	-12.4965
	x-shear	8.9e-2	8.8e-2	8.9e-2	8.9e-2	8.8e-2	8.9e-2	8.9e-2	8.8e-2	8.9e-2
	y-shear	8.9e-2	8.8e-2	8.9e-2	8.8e-2	8.8e-2	8.9e-2	8.8e-2	8.8e-2	8.9e-2
	sigma 0	0.63	0.54	0.71	0.65	0.54	0.76	0.65	0.54	0.74

Table 3: Statistics of pairwise differences between pixel coordinates of the colour channels for the Rollei plate (in μm)

| Scanner/ # of comparison points	Statistics	Red - Green channel (for all 4 scans)			Red - Blue channel (for all 4 scans)			Green - Blue channel (for all 4 scans)		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Antares/ 13456	RMS x	1.4	1.2	1.5	1.8	1.6	1.9	0.8	0.7	1.0
	RMS y	1.7	1.2	2.2	2.0	1.7	2.4	1.2	1.0	1.6
	mean x	1.1	0.9	1.3	1.5	1.2	1.6	0.3	0.3	0.4
	mean y	-0.5	-0.6	-0.3	-1.3	-1.3	-1.2	-0.8	-0.9	-0.7
	max abs. x	3.7	3.6	4.0	4.5	4.3	4.6	3.8	3.4	4.1
	max abs. y	11.1	4.8	14.0	11.0	6.0	13.0	6.4	4.0	12.6
Polaris/ 13456	RMS x	1.2	1.1	1.3	1.2	1.1	1.2	1.2	1.1	1.2
	RMS y	2.7	2.4	2.9	3.0	2.8	3.1	2.4	2.4	2.5
	mean x	-0.0	-0.1	0.1	-0.1	-0.2	0.0	-0.1	-0.2	0.1
	mean y	-0.2	-0.3	0.0	-0.7	-1.1	-0.4	-0.5	-0.8	-0.1
	max abs. x	5.0	4.0	6.3	5.2	4.6	6.7	4.5	3.5	6.5
	max abs. y	7.4	6.9	8.0	7.9	7.6	8.4	7.4	6.6	8.0

Table 4: Statistics of differences between pixel and reference coordinates after an affine transformation for each individual image tile (in μm). The statistics are computed from all tiles of all four scans of each colour channel.

Scanner/ # of control points per tile	Statistics	Red channel			Green channel			Blue channel		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Antares/ 144-169	RMS x	0.8	0.7	1.0	0.8	0.7	1.0	0.8	0.7	1.0
	RMS y	0.8	0.7	1.0	0.8	0.7	1.0	0.8	0.7	1.0
	max abs. x	2.4	1.7	5.0	2.5	1.8	5.0	2.5	1.8	5.3
	max abs. y	2.7	1.6	3.7	2.7	1.6	3.8	2.8	1.7	3.9
Polaris/ 144-169	RMS x	0.6	0.5	0.7	0.6	0.5	0.7	0.6	0.5	0.7
	RMS y	0.7	0.5	0.8	0.7	0.6	0.9	0.7	0.6	0.9
	max abs. x	1.6	1.1	4.7	1.7	1.1	4.7	1.7	1.1	4.7
	max abs. y	1.9	1.2	4.2	1.9	1.3	4.3	1.9	1.3	4.1

Figure 3. A broken line of the ETH plate (left). 1-D LSTM (x-shift, x-scale) with a synthetic template (right). The grey rectangle on the left is the starting value for the matching, the black one the final match position.

a)   **b)**
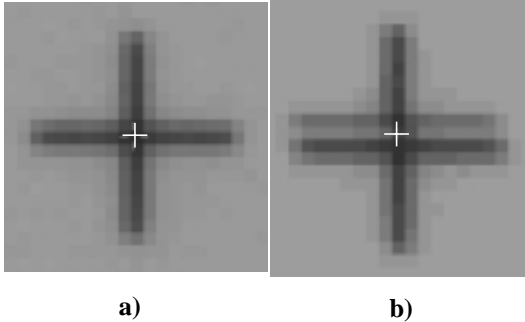
Figure 4. Matching of Rollei crosses. The result is shown by the white cross.

a) the cross exhibiting the largest residual in the global geometric accuracy test (see Table 2, Antares, red channel). The residual of 26.4 µm is obviously not due to wrong matching.

b) a cross at the horizontal border of two tiles close to the plate centre. Due to wrong scale factor (pixel size) determined in the geometric sensor calibration of the scanner, the two tiles overlap by ca. 1 pixel and the horizontal cross bar is imaged twice. The matching is approximately correct and the cross does not belong to the group with the largest residuals in the global geometric accuracy test, but the position of the tiles is still wrong.

3.2. Misregistration between colour channels

The results are summarised in Table 3 and some examples are given in Figures 1b) and 2b). Generally the differences between the channels R, G, B are larger for R-B, then R-G, and G-B. This is to be expected since the sequence of scanning is R, G, B and vibrations cause larger differences between R and B. Polaris has worse values than Antares regarding the maximum errors (excluding the random y-shift of Antares) and the RMS y, but latter has a larger bias (mean x and y values). The RMS values generally vary between 1 and 3 µm (latter being the declared scanner geometric accuracy). These errors could be due to properties of the optics (not achromatic) but also vibrations. As it will be shown in section 3.3 the optical properties are similar for all colour channels, so the greater part of the misregistration errors is due to vibrations caused by high scanning speed. In addition, for Antares the random line shift causes the largest local errors. Even without the random y-shift the maximum errors are 5 - 8 µm (Polaris). The RMS and maximum errors are worse in y for both scanners.

3.3. Local geometric accuracy and repeatability

The affine transformation between pixel and reference coordinates was computed for each tile excluding the border tiles that had less grid crosses (64 tiles were used with 144 to 169 crosses each). The results are shown in Tables 4 and 5 and an example of a tile in Figure 5. Errors due to mechanical positioning, vibrations and random y-shift (Antares) are absorbed by the translations of the affine transformation. Thus, the errors in Table 4 and Figure 5 represent mostly optical errors, especially lens distortion. The errors are very similar for all tiles, all four scans and all colour channels. Also the differences between the two scanners are in this case very small. The RMS errors are less than 1 µm, and the maximum errors are on the average 1.5 - 3 µm and can reach up to 5 µm.

More interesting are the affine parameters of the individual tile transformations in Table 5. 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 stage coordinate system. However, these translations vary a lot. For Antares up to 45 and 69 µm, and for Polaris up to 22 and 18 µm in x and y respectively. This could be explained by shifts of individual tiles. However, the tile shifts of Antares were not so large, and if they were, they would show up in the errors of the global geometric test, which was not the case. We believe that the explanation lies in the scales. As the tiles move away from the plate centre (where the origins of the
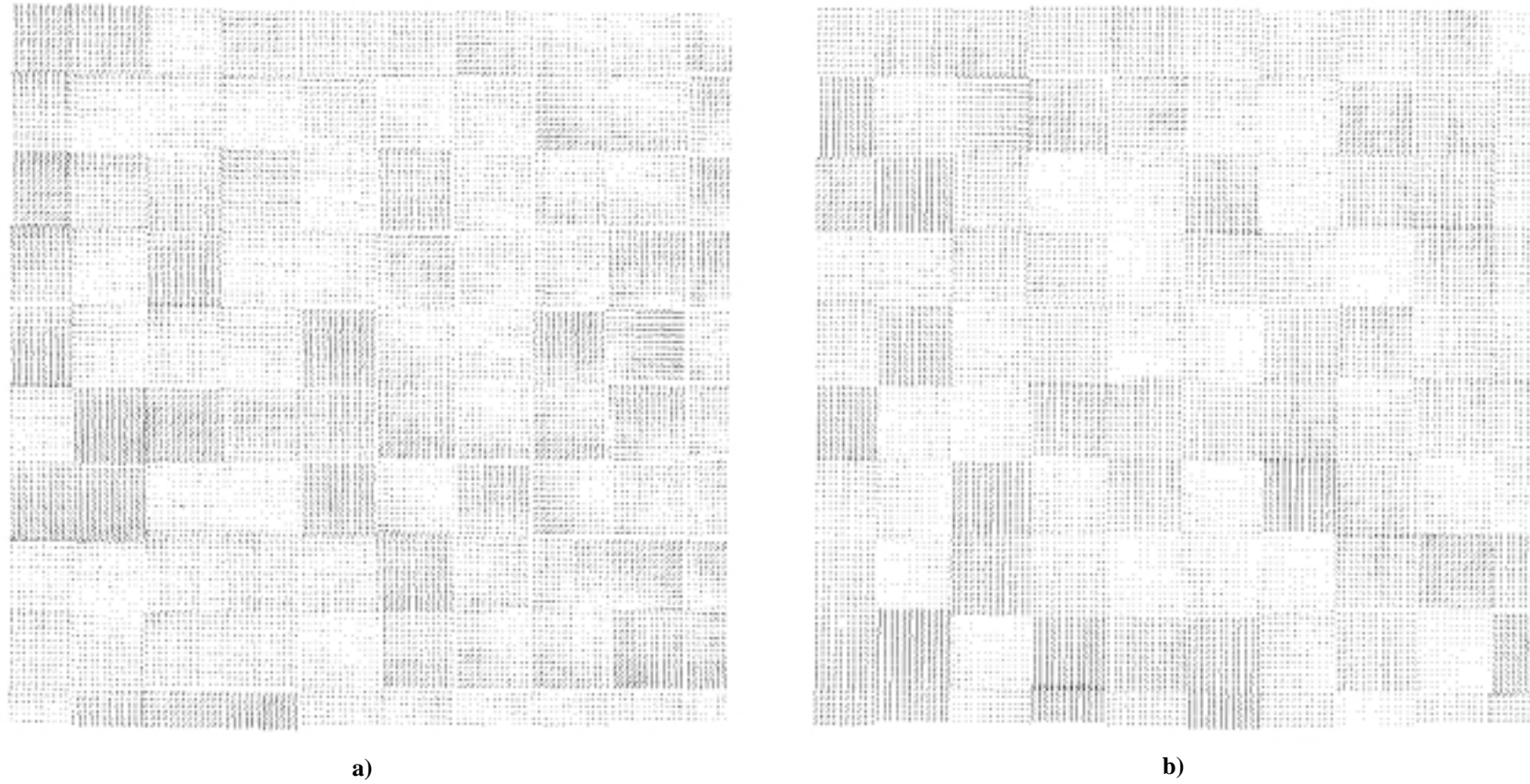
Figure 2. Polaris (vectors enlarged by 150).

a) residuals of global affine transformation from pixel to reference coordinates using all grid crosses as control points (red channel, third scan). The residuals are almost constant within each tile and much smaller that those of Antares.

b) colour misregistration errors, i.e. pixel coordinates of red minus blue channel (third scan). The errors are almost constant within each tile, larger than those of Antares, and of similar magnitude to the errors of the global affine transformation. The cause of this colour misregistration are vibrations during the sequential colour acquisition (R, G, B) at each tile position, caused by the high speed movement of the scanner stage.

a)

b)
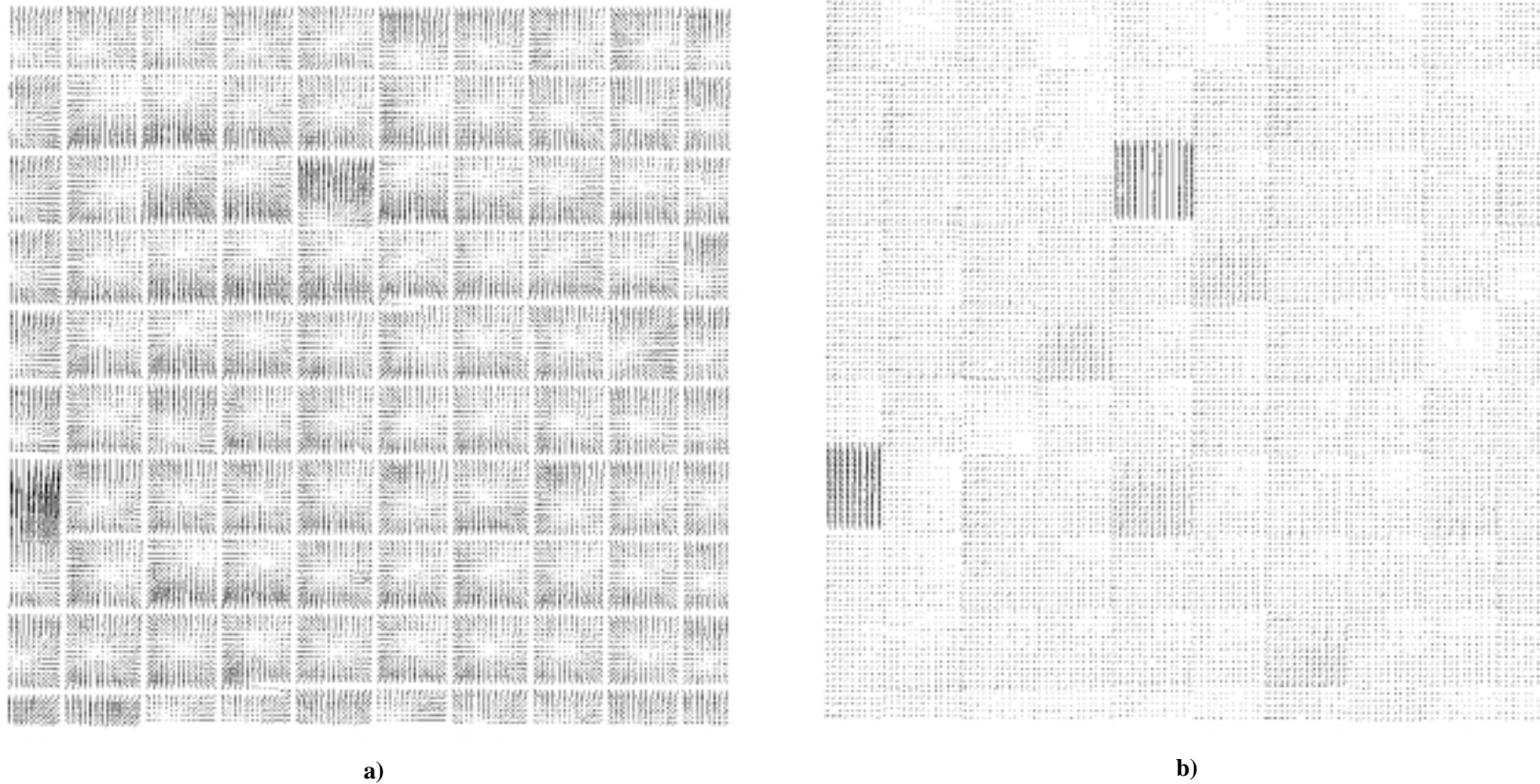
Figure 1. Antares (vectors enlarged by 150).

a) residuals of global affine transformation from pixel to reference coordinates using all grid crosses as control points (red channel, second scan). The tile structure of the image is clearly visible. The residuals within each tile are very systematic resembling errors due to lens distortion. Within each tile there is a position with small residuals, however this position is not always at the centre of each tile. The large residuals at the borders of the tiles cause shifts of each tile with respect to its neighbouring ones.

b) colour misregistration errors, i.e. pixel coordinates of red minus blue channel (second scan). The errors are almost constant within each tile. The two tiles with the large y-residuals on the right plot show the random line shift of one pixel. This error is also reflected in the larger errors of the corresponding tiles for the global affine transformation on the left.

Table 2: Statistics of differences between pixel and reference coordinates for the Rollei and ETH plates after an affine transformation (in μm)

Scanner, # of control/ check points	Statistics*	Rollei, Red channel (for all 4 scans)			Rollei, Green channel (for all 4 scans)			Rollei, Blue channel (for all 4 scans)			ETH, B/W scan
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	
Antares, 13456/0	RMS x	3.7	3.7	3.7	3.5	3.4	3.5	3.1	3.1	3.2	
	RMS y	5.5	5.3	5.6	5.0	5.0	5.2	4.7	4.7	4.7	
	max abs. x	10.0	9.5	10.7	9.8	9.6	10.1	9.3	8.8	9.6	
	max abs. y	20.7	13.9	26.4	15.9	13.7	21.9	13.1	12.7	13.4	
Antares, 4/13452	RMS x	3.8	3.7	3.9	3.7	3.6	3.8	3.5	3.2	3.6	
	RMS y	8.5	8.0	8.9	8.2	7.9	8.5	7.8	7.6	8.0	
	mean x	0.8	0.5	0.9	1.0	0.7	1.5	1.3	0.6	1.8	
	mean y	6.4	5.8	6.9	6.2	6.0	6.5	6.0	5.8	6.3	
	max abs. x	10.7	10.4	10.9	10.9	10.3	11.6	10.5	9.5	11.5	
	max abs. y	21.7	20.9	22.5	20.8	20.7	21.1	20.0	19.5	20.4	
Polaris, Rollei 13456/0, ETH 625/0	RMS x	1.8	1.7	1.9	1.8	1.7	1.8	1.7	1.6	1.8	2.1
	RMS y	2.7	2.6	2.9	2.3	2.2	2.5	2.4	2.4	2.5	3.1
	max abs. x	7.1	6.1	9.0	6.7	5.9	7.2	6.7	6.0	7.3	7.4
	max abs. y	8.7	7.8	10.7	8.2	6.8	8.9	9.3	7.4	10.5	10.1
Polaris, Rollei 4/13452, ETH, 4/621	RMS x	1.9	1.8	2.1	2.0	1.9	2.3	2.0	1.8	2.7	2.4
	RMS y	3.7	3.0	4.2	3.3	3.0	3.5	3.2	2.9	3.6	3.7
	mean x	0.3	0.1	0.5	0.9	0.5	1.1	0.9	0.1	2.0	0.7
	mean y	-2.1	-2.9	-1.0	-1.9	-2.5	-1.3	-1.4	-2.7	-0.5	1.5
	max abs. x	7.7	6.7	10.1	7.9	7.1	8.2	7.5	6.6	8.6	7.9
	max abs. y	11.0	8.9	13.5	9.6	9.2	10.3	11.0	10.4	11.6	11.1

* When all points are used as control points, the mean values are zero and thus not listed in this and all subsequent tables.

5

times visible. To reduce the effect of such dark corn and also of dust etc., grey values that are outside a range are excluded from the computation of the statistics. The range is computed for each grey wedge as (mean ± 3 x 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 deviation due to saturation. The linearity was checked by plotting the logarithm of the mean grey value of each wedge against the respective calibrated density. These points should ideally lie along a line and be equidistant. The dynamic range is determined as following. Firstly, the minimum unsaturated density is selected. Then, the maximum detectable density is determined such that the following relation is fulfilled: $M_i + 2\,SD_i < M_{i-1}$, with M and SD the mean and standard deviation of the wedges and i increasing with increasing density.

2. Scanner stage glass plate

Grey level statistics (mean, standard deviation, minimum, maximum) were computed for each of the four tiles of each scan and each colour channel, as well as statistics of the differences between the multiple scans of the same channel. This permits analysis of the spatial (relating to the variation from tile to tile), temporal and spectral variation of the noise. Additionally, scanning of multiple, spatially different tiles of a homogeneous surface permits the detection of blemished pixels, and their separation from other noise sources like dust which also cause local grey level differences.

3. Artifacts

The above mentioned scanned patterns were very strongly contrast enhanced by Wallis filtering [2]. This permits the visual detection of various possible artifacts like radiometric differences between neighbouring tiles, interference patterns, echoes, etc. However, the quantification of radiometric errors is always performed using the original images.

In addition, we had previously performed six other similar tests with one of those and two additional DSW200 scanners over a time period of one and a half year (four tests with the ETH and two with the Rollei plate). This allows to check the long term temporal stability of the systems and variations between different systems of the same brand. The results of the old tests will not be presented here but we will comment on their compatibility with the new test results.

3. EVALUATION OF GEOMETRIC PERFORMANCE

3.1. Global geometric accuracy and repeatability

The results of this evaluation are shown in Table 2 and some examples are illustrated in Figures 1a) and 2a). The transformation with 8 control points (CP) was left out from the table to make it more readable. Generally, they were slightly better than the results with 4 CP with the exception of Polaris in x-direction. When using few CP, the transformation results depend heavily on the CP quality, so a higher redundancy (8 instead of 4 points) is positive. The accuracy is generally better for the blue, then green, then red channel. That is because of the vibrations which are more prominent in red since it is the first colour to be scanned after moving the scanner stage to a new position to scan another tile. The accuracy in y is clearly worse than in x. The short term repeatability (comparison of four scans) is on the average good, with the exception of the maximum errors for Polaris, red channel (vibrations) and Antares (random y-shift). The RMS for Antares varies between 3 - 6 μm and 3.5 - 9 μm for all and 4 CP respectively, while for Polaris these values are 1.5 - 3, and 2 - 4 μm. A B/W scan is expected to deliver similar results as the red channel. The mean error in y reaches 7 μm for Antares and 4 CP, in all other cases it is less than 3 μm. The mean maximum errors with all CP are for Antares ca. 10 and 14 μm in x and y (excluding the random y-shift), while for Polaris are 7 and 9 μm respectively. The results of the ETH plate are compatible with those of the Rollei plate for Polaris (see bottom half of Table 2). Previous tests generally confirm the above results, although in some cases the maximum errors (excluding random y-shift) reached 15-17 μm for an affine transformation with all CP.

The error patterns as shown in Figures 1a) and 2a) are very systematic. For Antares they are quite similar for all tiles and increase towards the border of each tile due to a scale (pixel size) problem which will be explained in section 3.3. For Polaris the errors vary from tile to tile and are similar within each tile (main error sources are mechanical positioning and vibrations which influence the whole tile). In both cases the borders of the tiles do not fit together. This is shown well with the ETH grid plate. The thick grid lines appear broken at the tile borders (see Figure 3). This happens in both x- and y-direction. In previous tests we have measured the displacement between the tiles with 1-D LSM (Figure 3) and found maximum displacements between 6 and 10 μm. Thus, the large errors do not affect only a certain position but whole tiles having 4 million pixels.

hain was used. The radiometric performance was mainly checked by scanning a calibrated Kodak CAT grey level wedge on film (21 densities with density step of approximately 0.15 D; density range 0.05 D - 3.09 D). The densities were determined by repeated measurements using a Gretag D200 microdensitometer. In addition the empty scanner stage glass plate was scanned (see second radiometric test below).

All test patterns were scanned with Polaris, while for Antares only the Rollei grid plate was scanned. All scans were with 12.5 μm pixel size, if not otherwise mentioned. The Rollei plate was scanned four times in colour to check the geometric repeatability as well as misregistration between the colour channels. The resolution pattern was scanned twice, the second time by shifting the scan area by half a pixel in order to account for an unknown arbitrary phase shift between pixels and lines of the resolution pattern, which can influence the results for high line frequencies. The grey level wedge was scanned in colour, with 12.5 and 25 μm pixel size to check differences between colour channels and the effect of pixel size on the radiometric performance. For the 12.5 μm version a linear and a logarithmic LUT was used. The wedge was masked with a black carton to avoid stray light. The empty scanner glass stage was scanned four times in colour. The size of the scanned area included four full image tiles. The exposure time was set such that the resulting mean grey value was sufficiently different that 255.

The pixel coordinates of the grid crosses were measured by fully automatic Least Squares Template Matching (LSTM). This algorithm is described in [7], while details are given in [2]. The software implementation of the algorithm that was used employs on-the-fly generation of the templates and is described in [9]. An option of the algorithm that reduces the influence of dust and other noise on the cross measurement was used. The accuracy of LSTM for well-defined, high contrast, flat targets is very high and for the ETH and Rollei plates was 0.02 - 0.03 pixels, as indicated by the standard deviations of the parameters. In LSTM an affine geometric transformation was used and a two parameter radiometric correction. The patch size was 25^2 pixels. Matching results with bad quality criteria (low crosscorrelation coefficient etc.) were automatically excluded from any further analysis. In addition, the matching results of all crosses with large errors were interactively controlled (e.g. Figure 4a)).

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 (a) an analysis of the influence of pattern density on the ability to reliably detect errors and (b) an analysis of the different cross line widths (15, 187 microns) on the LSTM accuracy.

2. Misregistration errors between the channels (applied only for the Rollei plate)

   Such errors were checked by comparing pairwise the pixel coordinates of each channel (R-G, R-B, G-B).

3. Local geometric tests (applied only for the Rollei plate)

   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, random line shift) 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. Geometric resolution

   It was determined by visual inspection of the two scanned resolution patterns, 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 above first three tests efforts were made to separate the contribution of various error sources (especially mechanical positioning, random line shift and lens distortion) on 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 we have 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. To avoid 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 some-

The aim of these investigations was to test the geometric and radiometric performance of the Leica/Helava DSW 200 scanner. A short description of the scanner is given in [6] while its main characteristics are summarised in Table 1. The sensor 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 dimensions of 1984 x 1984 pixels. A feathering is performed across the borders of the tiles to smooth out 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 fiber optic and away from the stage to reduce the danger of vibrations. The base scan pixel size 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 x 2, 4 x 4, ...) of the grey values in software. The scanner manufacturer does not give an exact value for the radiometric accuracy (noise level) of the scanner, but rather vague statements indicating an accuracy of 1-2 grey values. The user can specify a LookUp Table (LUT) for mapping the output grey values. 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) and the output image format. Although the transmission of the signal from the camera is digital, a framegrabber is still needed and used.

The scanner software performs two geometric (see [11] for a coarse description) and two radiometric calibrations. The first geometric calibration (stage calibration) is performed by scanning a reference grid plate of 13 x 13 crosses with 2 cm spacing. The crosses are measured automatically by crosscorrelation, and through computation of an affine transformation between pixel and reference coordinates, corrections to the scanner stage at the grid positions are computed. These corrections are applied on-line in each scan. Note that the grid covers an area of 240^2 mm^2 (in earlier software versions only 11 x 11 crosses were used with an area of 200^2 mm^2), while the possible scan area is 265^2 mm^2. In stage positions outside the calibration grid, the corrections are extrapolated and saved in a calibration file covering 15 x 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 x 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 manufacturer makes vague statements as to how often the geometric calibrations should be repeated (ranging from once per week to once per month). The radiometric calibrations include an equalisation of the CCD sensor element response for a low and high illumination, and an equalisation of the colour response (colour balance) by using the histograms of the colour channels.

Table 1: Main features of Leica/Helava DSW200 scanner.

Mechanical movement	Sensor and dimensions (pixel)	Scanning format (mm)	Base scan pixel size (μm)	Quantis. bits (internal/ output)	Illumination	Colour passes	Density range	Geometric accuracy (μm)	Scanning throughput, speed
flatbed, moving stage	Kodak Megaplus 2029 x 2044	265^2	5-15	10/8	xenon, fiber optic, sphere diffusor	1	3D	< 3	1.2 MB/s, max. 35 mm/s

2. DESCRIPTION OF TEST PROCEDURES AND TEST PATTERNS

Our geometric and radiometric investigations were performed with two DSW200 scanners, after performing all necessary scanner calibration procedures, as suggested by the manufacturer, with an accuracy of less than 2 μm for the geometric calibrations. The scanners were located at Swissphoto Vermessung AG in a room with temperature and humidity control. The two scanners tested, named Antares and Polaris, had different Kodak Megaplus models, namely 4.2i and 4.2 respectively. According to the manufacturer scanners with early versions of the 4.2i model exhibit a shift of one pixel in the y-direction due to interface problems between the camera and the framegrabber. This shift is random, i.e. it occurs only for some tiles. The older 4.2 model and new versions of the 4.2i model do not exhibit this error.

The geometric performance was tested by scanning with 12.5 μm two different glass grid plates. A custom one produced by a Swiss company specialising in high precision optical components (IMT) with a 1 cm grid spacing and 187.5 μm line width (25 x 25 crosses), called ETH plate thereafter, and one high precision réseau glass plate from Rollei, which has been produced by Heidenhain, with a 2 mm grid spacing, 200 μm cross length and a 15 μm line width (116 x 116 crosses). The coordinates of the crosses were known with an accuracy of 2 - 3 μm for the ETH plate and better than 1 μm for the Rollei plate. To determine the scanner resolution a standard USAF resolution pattern on glass produced by Heiden-

Geometric and radiometric performance evaluation of the Leica/Helava DSW200 photogrammetric film scanner

Emmanuel P. Baltsavias[a], Silvio Haering[b], Thomas Kersten[b]

[a]Institute of Geodesy and Photogrammetry, ETH-Hoenggerberg, CH-8093 Zurich, Switzerland
[b]Swissphoto Vermessung AG, Dorfstr. 53, CH-8105 Regensdorf, Switzerland

ABSTRACT

In this paper geometric and radiometric investigations performed with two Leica/Helava DSW200 scanners will be presented. Good quality test patterns and accurate processing methods for their performance evaluation have been employed. 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, electronic errors and lens distortion) on the total error. The radiometric tests include investigations on noise, linearity, dynamic range, spatial, temporal and spectral variation of noise, and artifacts.

After a brief description of the scanner, details on the above investigations, used analysis methods and achieved results will be presented. The most important errors, especially for organisations employing the scanner in production, are non-repeatable geometric errors that are due to mechanical positioning inaccuracies and vibrations. This causes whole CCD image tiles (4 million pixels) to be systematically shifted with respect to their neighbouring tiles, with obvious negative influence on any subsequent measuring task.

Keywords: film scanner, scanner test, CCD, geometric evaluation, radiometric evaluation, colour misregistration

1. INTRODUCTION

Photogrammetric film scanners are and in the near future will be even more used for producing digital data especially from aerial images, but also from close-range and satellite-based film sensors. Since every subsequent processing step builds upon the scanned imagery, the analysis of the scanner accuracy and performance is of fundamental importance. Unfortunately, there are very few publications on this topic and most users take for granted that photogrammetric scanners perform well. However, our experiences with several scanners as well that of other researchers and users has shown that many problems of geometric and radiometric nature occur. Our investigations are not important only for film scanners. Similar scanning systems with linear or area CCDs are also employed in digital close-range cameras, and the CCD technology and its associated problems is common in most vision systems.

An overview of photogrammetric scanners is given in [1]. Related work on test procedures for evaluation of photogrammetric film scanners is reported in [3, 5, 8, 10, 11, 12, 13, 14], while a comparative test and calibration of five DTP scanners is given in [4]. Out of the 11 photogrammetric scanners that are available today, the ones that are or are expected to be used more extensively include: Leica/Helava DSW200/300, Vexcel VX 3000/4000, Wehrli RM-1, XL Vision's OrthoVision, Zeiss/Intergraph PS1, Zeiss SCAI/Intergraph TD. However, for most of them there are no published test results. Such publications include geometric and radiometric evaluation of the RM-1[8], image noise and sensitivity analysis of PS1, VX3000 and RM-1[10], and limited tests on the geometric accuracy, MTF and noise level of VX3000 [11,13,14] and on the geometric accuracy of PS1[14] (although the model was not explicitly named).