'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

Automatic aerial triangulation: results of the OEEPE-ISPRS test and current developments

CHRISTIAN HEIPKE, Hannover

ABSTRACT

The European Organisation for Experimental Photogrammetric Research (OEEPE) and the International Society for Photogrammetry and Remote Sensing (ISPRS) have carried out a test on the performance of tie point extraction in automatic aerial triangulation (AAT). The aims of the test were to investigate the geometrical block stability, the accuracy of the tie points and the derived orientation parameters, and the limitations of existing commercial and experimental software systems. In order to separate the essentially new aspect of digital processing, namely automation, from conventional issues of aerial triangulation, control information was not assessed, and the test blocks to be processed had an arbitrary block datum.

The Chair for Photogrammetry and Remote Sensing, Technische Universität München acted as pilot centre for the test. In early 1997 various small blocks of different scene content were distributed to interested participants. Their task was to generate tie points in an automatic way. The results of 21 participants, including all major software vendors of AAT and users of their systems, have been analysed and are presented in this report. If a large number of tie points per image has been extracted, the blocks were found to be mostly stable. Under good conditions (open, flat terrain) an accuracy for the tie points of up to 2.2 µm corresponding to 0.11 pixel could be reached, while under less favourable conditions, the result was 4-9 µm or 0.2-0.3 pixel. These figures were found to be very similar for the different systems. In mountainous and forested areas, some systems failed to produce acceptable results. Reliable self control is a feature missing in all systems as of 1997. Also, it seems that considerable experience is required to properly run the systems.

Besides the test results this paper also discusses recent improvements of AAT and the current state-of-the-art. While relatively large cost savings have been realised in practice using AAT instead of analytical aerial triangulation, human supervision and intervention will remain necessary. An integration of AAT with the direct measurement of the parameters of exterior orientation by means of GPS and INS is seen as the most powerful way for obtaining highly accurate and reliable image orientation in future.

1. INTRODUCTION

Automatic aerial triangulation (AAT) has been an increasingly interesting topic of research and development in digital photogrammetry for a number of years (see Schenk 1997 for an excellent review of the subject). The two tasks of measuring the image coordinates of tie points and of computing the orientation parameters, which were well separated in analytical photogrammetry, are more and more being merged into a single process, carried out in a hierarchical fashion using image pyramids. In future there will most probably be an option to also include the generation of digital terrain models (DTM) into this process. At the same time a shift of focus concerning the results of aerial triangulation can be observed. While in earlier times point densification was the primary goal, currently the orientation parameters themselves are of growing importance. There are two reasons for this shift of focus: First, the automatically determined tie points are not really useful for point densification, since in general they do not fulfil the requirements set out in the point selection phase of analytical aerial triangulation. Second, the orientation parameters themselves are increasingly used directly for subsequent tasks such as orthoprojection or vector data capture.

Over the last few years various AAT software systems with different degree of automation have been developed and have become commercially available, either as stand-alone packages or as part of a Digital Photogrammetric Workstation. These systems have been introduced into practice, and users have started to report on obtained results. At the same time a number of questions remained open, from the theoretical side (how to best select image primitives suitable for point transfer, multi image matching vs. matching only two images at a time, area based vs. feature based matching, the influence of local image texture etc.) as well as from the practical side (how accurate do initial

values of exterior orientation have to be, what is the most suitable pixel size to use, how many tie points should be available per image, which degree of automation can be reached and what does it depend on, what is the effect of image compression, how to implement an efficient procedure for quality control etc.).
In 1996 the European Organisation for Experimental Photogrammetric Research (OEEPE) and the International Society for Photogrammetry and Remote Sensing (ISPRS) launched a common test on the performance of tie point extraction in automatic aerial triangulation (Heipke, Eder 1996) in order to approach some of the open questions and to allow a comprehensive comparison between the available systems. „Tie point extraction“ is meant to include the selection, transfer and image coordinate measurement of block tie points. The test was primarily aimed at the commercial software development and the user community of AAT systems. Detailed results of this test have been published in Heipke, Eder (1999). In this paper a synthesis of these results is given. It should be noted that the reported results refer to the AAT software available in 1997. In the meantime, and perhaps partly as a reaction to the test results, improvements of the commercial systems could be observed. These improvements are discussed together with some thoughts about the future of automatic aerial triangulation at the end of this paper.

2. THE OEEPE-ISPRS TEST

2.1. Test goals

The goals of the test were developed in preliminary discussions together with potential test participants. It was decided to investigate

- the geometrical stability of the resulting block,
- the accuracy of the image coordinates of the tie points, and
- the limitations of existing commercial and experimental software systems.

Throughout the test, tie point extraction was considered to be a totally **autonomous** process, to be carried out without any user interaction. In particular, any interaction during the tie point generation process, as well as manual editing or completion of the automatically obtained results in order to improve the measurement precision, to eliminate blunders and/or to introduce new measurements in areas where the automatic process failed to determine tie points, was not allowed within the test. Only automatic blunder detection and elimination within a robust adjustment was permitted. In this way the essentially new aspect of digital imagery, namely automation, could be investigated separately from the issues which basically remain constant in the transition from analytical to digital photogrammetry (control information, block configuration, accuracy propagation, etc.).

2.2. Test organisation and test data sets

The Chair for Photogrammetry and Remote Sensing, Technische Universität München (TUM) acted as pilot centre for the test. In early 1997 various small blocks of different scene content (see table 1) were distributed to interested participants. As an example the data set Montserrat is depicted in figure 1. Guidelines for the selection of the test data were the need for a representative test data set covering different standard applications in photogrammetry, for small blocks/strips resulting in manageable data volumes, and the use of photogrammetric images and scanners only. The first point inspired the use of different scene contents, topography, cameras, scales, film material, and overlap configurations. As far as image scales were concerned, preference was given to larger scales, because in these cases, potential matching problems due to occlusions and relief displacement are more pronounced.

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

Project name	**Echallens**	**Montserrat**	**OSU**	**Kapellen**	**München**
Scene content	open, partly forest	forest, partly built-up	built-up, partly trees	settlement, partly open	city centre
Scene topography	flat	hilly	flat, buildings	flat	buildings
Image scale	1 : 5 000	1 : 15 000	1 : 4 000	1 : 4 000	1 : 2 000
Camera	Wild RC 10	Zeiss RMK TOP	Wild RC 10	Zeiss RMK A	Zeiss RMK A
Focal length [mm]	150	150	150	150	300
Flight datum	September 1982	May 1995	September 1995	April 1992	May 1975
Film material	black and white	black and white	FIR	black and white	colour
Number of images	3 x 3	3 x 3	3 x 3	2 x 3	3
Overlap	l = 60 %, q = 30 %	L = 60 %, q = 30 %	l = 60 %, q = 60 %	l = 60 %, q = 60 %	l = 60 %
Scanner used	LH DSW 200	Zeiss PS1	LH DSW 200	Wehrli RM1	Zeiss PS1
Pixel size for test	20 µm	30 µm	25 µm	24 µm	30 µm
Scanned material	negative, original	negative, original	positive, original	negative, original	positive, original
Scanned channel	pan	pan	red (= infrared)	pan	red
Scan datum	January 1996	November 1996	October 1995	June 1996	December 1996
Source	EPFL, Lausanne	ICC, Barcelona	The Ohio State University / TU München	Hanover University	Technische Universität (TU) München

Table 1: Description of the test data sets.

The second point led to the selection of blocks with 3 x 2 and 3 x 3 images, strips with 3 images and pixel sizes of 20-30 µm. While operational problems cannot be detected with such small blocks, the geometrical block stability and the accuracy of the tie points can be assessed. As for the third point, only first generation film products were scanned and all employed scanners are especially designed for photogrammetric applications.

The task of the participants was to automatically generate tie points without human intervention using an AAT software available to them, given the digital imagery together with auxiliary information. Wherever possible a common set of free parameters for the individual programs was to be used. After announcing the test 39 interested groups requested the test data. 21 participants (4 major commercial AAT software providers, 5 national/regional mapping organisations, 4 private companies, 3 research institutes employing commercial products, and 5 research institutes who had developed their own AAT software; see table 2) returned results. Four groups can be distinguished, namely users of the commercial systems HATS from LH Systems (de Venecia et a. 1996; 7 users), Match AT from Inpho (Ackermann, Krzystek 1997; 5 users), and Phodis AT from Carl Zeiss (Tang et al. 1997; 4 users), and the five participants having developed their own software (FGI - Honkavaara, Hogholen 1996; IPI - Wang 1996; TUM - Brand, Heipke 1998; DIIAR - Forlani et al. 1998; OUAT - Paszotta 1998).

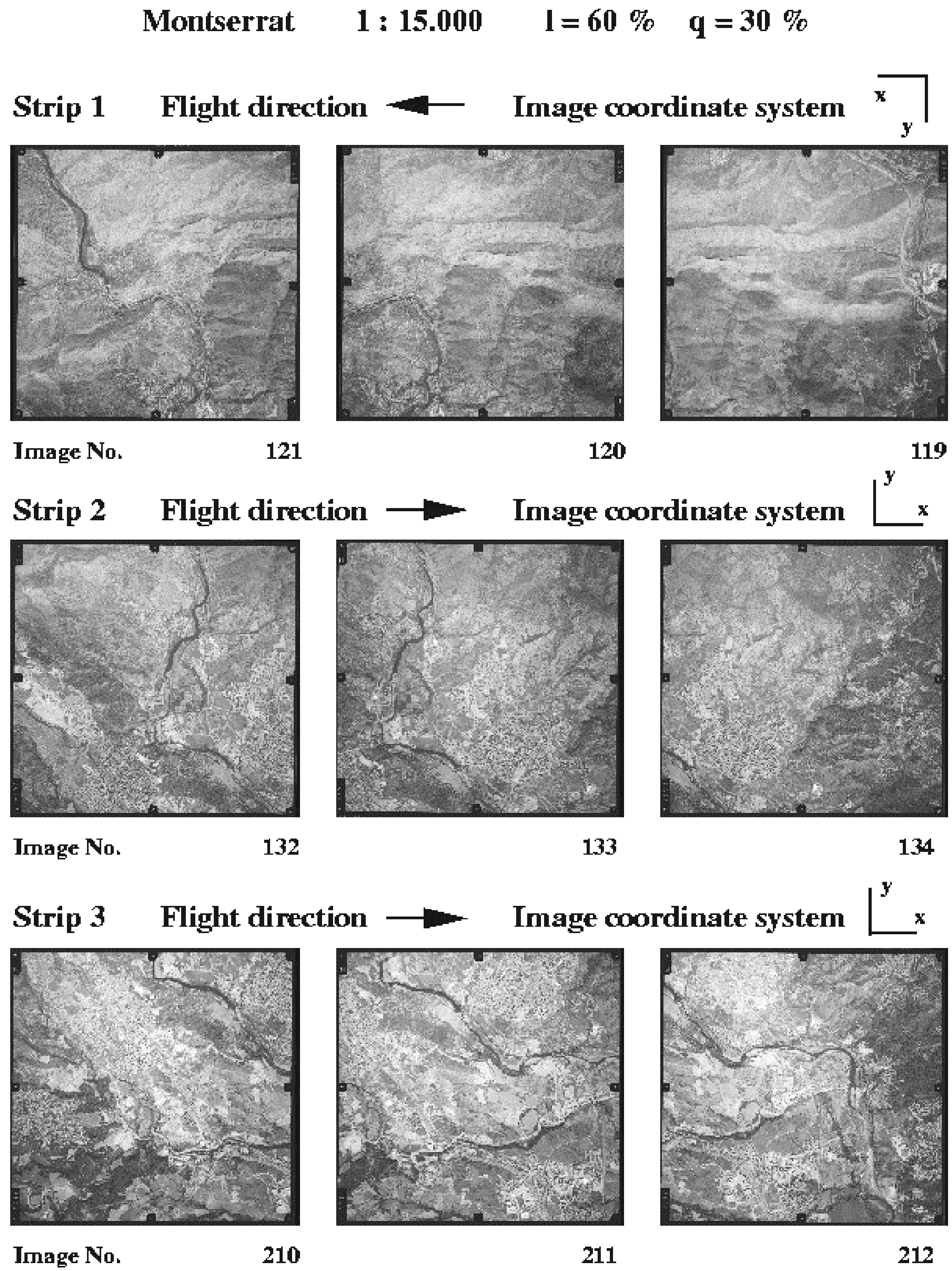

Figure 1: Imagery for test data set Montserrat.

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

full name and abbreviation of participant		**software and version no.**	**Echallens**	**Kapellen**	**Montserrat**	**München**	**OSU**
LH Systems, San Diego	LHS	HATS, 3.2.1.1	X	X	X	X	
Bundesamt für Geod. u. Kart., Frankfurt/M.	BKG	HATS, 3.1.1.2	X	X	X	X	
Institute for Photogrammetry, EPFL Lausanne	EPFL	HATS, 3.1.3k	X		X		X
National Land Survey of Finland, Helsinki	NLS-SF	HATS, 3.2.1.2			X		
National Land Survey of Sweden, Gävle	NLS-WE	HATS, 4.0.8	X	X	X		X
School of Geomatics, UNSW, Sydney	UNSW	HATS, 3.2.1	X		X	X	X
Swissphoto, Regensdorf	SWPH	HATS, 3.2.1.2 *	X	X	X	X	X
Inpho GmbH, Stuttgart	Inpho	Match AT, 2.1.0	X	X	X	X	X
Intergraph, Huntsville	I-graph	Match AT, 2.1.1	X	X	X	X	X
Compagnia Generale Ripreseaeree, Parma	CGR	Match AT, 2.1.1			X	X	X
Hansa Luftbild, Münster	HL	Match AT, 2.1.1			X		
Photogrammetrie GmbH, München	Ph GmbH	Match AT, 2.1.1			X	X	
Carl Zeiss, Oberkochen	CZ	Phodis AT, 2.0.1	X	X	X	X	X
Bayerisches Landesvermessungsamt, München	B-LVA	Phodis AT, 2.0.0	X	X	X	X	X
General Command of Mapping, Ankara	GCM	Phodis AT, 2.0.0		X	X	X	X
Landesvermessung + Geobasisdaten, Hannover	LGN	Phodis AT, 2.0.0	X		X		
Finnish Geodetic Institute, Masala	FGI	research system	X		X		
Institute of Photogrammetry and Eng. Surveys, Hannover	IPI	research system		X		X	X
Chair for Photogram. & Rem. Sensing, TU München	TUM	research system	X	X	X	X	X
Dip. Ing. e Idraul. Amb. e del Rilev., Politec. di Milano	DIIAR	research system	X		X	X	X
Chair of Ph & RS, Olsztyn Univ. of Agricul. a. Techn.	OUAT	research system			X		

Table 2: List of test participants (*: SWPH combined HATS with customised software).

2.3. Employed software systems

Neither the commercial products nor the developments of the research institutes as they stood in 1997 will be presented in detail in this report. However, some aspects shall be mentioned and have been collected in table 3. Further information is available in the given references. The participant SWPH has customised HATS for his own purposes. The resulting system while still being a HATS system has some unique features and was therefore entered in a separate row in table 3 called HATS*.

The first column of table 3 contains the names of the software systems. The next three columns deal with the **matching methods** employed. As can be seen all but the IPI development use a feature based matching scheme with points as matching entities. In most cases points are selected using the Förstner-Operator (Förstner 1991). IPI´s solution also uses relational descriptions of structures extracted from the imagery. Matching refinement in order to increase the geometric accuracy is

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

mostly based on least squares matching which is known to be the most accurate method. However, HATS and HATS* use subpixel cross correlation, OUAT uses simple cross correlation, and in the version for the test no refinement procedure is integrated into the TUM development. During matching refinement only Inpho uses all available overlapping images simultaneously. All other systems rely on matching image pairs and generating multi ray tuples in a separate step. It should be noted that the exact matching algorithms are not always published in the literature.

In order to solve the problem of obtaining **initial values** for the unknown orientation parameters all approaches are implemented in a hierarchical fashion based on image pyramids. As an option Match AT can also use an existing DTM as input which is claimed to be helpful especially in mountainous terrain. HATS, HATS* and FGI search for conjugate points only in predefined areas. Often areas around the “von Gruber positions” are used. Match AT also starts in these areas, but they are automatically shifted away from the initial position if no adequate matching results are obtained. Therefore, the X in the appropriate position in table 3 appears in brackets. On the other hand Phodis AT, and the TUM, IPI and DIIAR approaches try to match points in the whole images, at least in the upper pyramid levels. Some participants (e.g. DIIAR) have found that their system is very sensitive to the quality of the initial values of exterior orientation and have therefore changed the provided values manually prior to running their AAT software.

In some systems a sophisticated automatic **blunder elimination** scheme is integrated. For instance, in the TUM development every step of the algorithm is immediately followed by a verification step. Such a design allows for the early detection and elimination of blunders. While HATS comes with interactive blunder elimination, HATS* is tuned to automatic elimination of gross errors.

For the FGI system blunder detection is performed during a robust bundle adjustment loosely coupled with the matching software (thus the brackets in table 3). Match AT and TUM compute **integrated robust bundle adjustments** at each level of the pyramid in order to improve the initial values for the unknowns from one pyramid level to the next and to eliminate additional blunders. For this step Match AT and FGI need a minimum number of 3 ground control points. HATS and HATS* include a bundle adjustment with a somewhat reduced functionality.

Also the **degree of automation** is different for the different systems. Some systems are designed as autonomous systems without any operator control (such as Match AT and Phodis AT), other approaches (such as HATS) are more flexible and usually call upon the operator in order to manually measure additional points or eliminate blunders. It should be noted that this possibility was not to be used by the test participants (see section 2.1). As evidenced by SWPH HATS can also be tuned into a fully autonomous system.

Finally, most systems have a list of free parameters, sometimes collected in a **parameter file**, which can be used to tune the results. The effect of these parameters, however, is not always clearly documented. While most participants used a standard parameter set for all test images, some did optimise the values in order to achieve better results.

Given these numerous differences in the approaches it is impossible within this test to link a certain result to a particular design feature. What makes the situation more complicated is the fact that different participants used different versions of the same software (see table 2). Nevertheless, the obtained results show some distinct trends, see chapter 3.

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

	matching entities	matching method for accuracy refinement	no. of images during matching refinement	image pyramids	DTM as input (optional)	use of von Gruber positions	automatic blunder elimination	integration of block adjustment programme	need for ground control points	autonomous system design	parameter file
HATS	points	subpixel cross correlation	2	X	-	X	-	(X)	-	(X)	X
HATS*	points	subpixel cross correlation	2	X	-	X	X	(X)	-	X	X
Match AT	points	least squares matching	all over-lapping	X	X	(X)	X	X	X	X	X
Phodis AT	points	least squares matching	2	X	-	-	-	-	-	X	-
FGI	points	least squares matching	2	X	-	X	(X)	(X)	X	X	X
IPI	points, structures	least squares matching	2	X	-	-	-	-	-	X	X
TUM	points	-	2	X	-	-	X	X	-	X	X
DIIAR	points	least squares matching	2	X	-	-	-	-	-	X	X
OUAT	points	cross correlation	2	X	-	-	X	-	-	X	X

Table 3: Comparison of the different systems used in the test
(HATS* is a customisation of HATS developed and used exclusively by SWPH).

2.4. Analysis procedure

The image coordinates of the conjugate points as received from the participants were analysed at the pilot centre using robust bundle adjustment and independent interactive reference measurements. The actual analysis procedure consisted of two different steps. In the first step for each set of image coordinates a robust bundle adjustment was carried out. Blunder detection and elimination was performed similar to the suggestions by Klein, Förstner (1984). Image coordinates not representing blunders were assumed to be uncorrelated and of an accuracy of $\sigma_{o,a\ priori}$ = 1/3 of a pixel. 1/3 of a pixel is a rather conservative estimate of the accuracy generally attributed to matching of two images. However, in the case of multiple overlapping images this value seems to be rather appropriate and was therefore selected. The influence of $\sigma_{o,a\ priori}$ onto the results was further investigated for selected cases (see Heipke, Eder 1999).
The block datum was fixed by introducing the minimum of seven orientation parameters (six parameters of one image and one base line) as constant values. Thus, it could be ensured that the resulting block would not be influenced by ground control information. Rather, the potential of the purely automatic tie point extraction could be assessed. For each bundle adjustment run the average number of tie points per image, the number and percentage of eliminated blunders, and the number of multi ray points were collected in a log file, and plots depicting the distribution of tie points connecting different images and different strips were generated. These results were used in order to obtain a first

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

impression of the quality of the received sets of conjugate points. Additional results of the robust bundle adjustment consisted in the adjusted exterior orientation parameters for each image, and the standard deviation σ_o of the image coordinates.
A second analysis step was carried out for each set of image coordinates in order to independently assess the accuracy of the obtained orientation parameters. Using interactively measured image coordinates of reference points as observations and the exterior orientation parameters obtained in the robust bundle adjustment of the first analysis step as constant values, over-determined least squares forward intersections were computed. Among other things, this computation resulted in a value for the accuracy of the image coordinates termed σ_{FI} for "forward intersection". σ_{FI} can be considered as a measure of quality for the orientation parameters determined from the results of the participants. Besides the numerical value for σ_{FI} plots showing the individual residuals of the least squares forward intersection across the whole block were also generated. Besides this comparison in image space also an object space comparison was carried out: the coordinates (X_{FI}, Y_{FI}, Z_{FI}) obtained in the over-determined least squares forward intersection were compared to the reference coordinates (X_{ref}, Y_{ref}, Z_{ref}), and the root mean square differences between the two coordinate sets were determined. These values were termed RMS(X), RMS(Y), and RMS(Z), respectively.
It is clear that neither σ_{FI} nor the RMS values fully describe the accuracy of the AAT performed by the participants. The reason is twofold: first, any effects connected to ground control were deliberately excluded from the analysis, and second the reference observations and the test results were obtained from the same images, because no independent reference measurements of adequate accuracy (one order of magnitude better than the test results, say) were available. It should also be noted that the mentioned numerical quality measures σ_{FI}, RMS(X), RMS(Y) and RMS(Z) constitute average measures for the complete block. As such they are not useful in detecting local block deformations. Within the test these local effects were investigated graphically using the mentioned plots. Nevertheless the presented analysis allows for an interesting and valuable assessment of the results of the participants as will be explained in the next chapter.

3. TEST RESULTS AND DISCUSSION

3.1. Blunder elimination, multi ray points, and point distribution

In this section the results of the analysis of the geometric stability of the blocks are reported. Montserrat turned out to be the most difficult data set. The scene is rather mountainous and contains forest, especially in the mountainous area in the upper part of the scene between the first and the second strip, leading to unfavourable conditions for image matching.
Therefore, the Montserrat results are presented here, see table 4 for the numerical values. Concentrating on this table a number of observations can be made:

- The average number of correct tie points per image (this is the number of tie points after blunder elimination) and the total number of multi ray points in object space differ considerably between the participants and systems. Some of the participants using HATS delivered rather few points. On the other hand Phodis AT and two research systems (TUM and FGI) extracted between 330 and 495 points per image and between 1005 and 1969 object points.
- It can be seen (and comes at no surprise) that within AAT a robust adjustment is absolutely necessary. In the systems which do not include an internal blunder elimination scheme up to 24% of the measurements were eliminated. The actual number of detected blunders differs according to the number of extracted tie points.
- A closer look at the number of rays per object point reveals that only some MatchAT users, TUM and OUAT obtained a large number of multi ray points. Expressed in relative figures for some participants, LHS obtained 60 % 2 ray points (125 out of 209) and 7 % 5 + 6 ray points (8+6 out of

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

209), the figures for Inpho are 46 % and 8 %, and those for CZ 63 % and 2 %. When interpreting these percentages one has to keep in mind that given the nominal overlap configuration of l=60 % and q=30 %, about 67 % of the block is depicted in two images, 12 % in three, 17 % in four and 4 % in six images.

The plots depicting the tie point distribution (not shown here due to space constraints) clearly convey the general philosophy of the different approaches: HATS determines a moderate number of conjugate points well distributed across the images. To some extent HATS emulates the results of interactive measurements. Match AT generates considerably more conjugate points and point clusters in the areas of multiple overlap areas. Phodis AT creates a very large number of conjugate points. However, the major part consists of 2 ray and 3 ray points.
The results for the other data sets confirm these findings reported here. The average number of tie points per image is rather high, a reliable blunder detection mechanism is absolutely necessary, and the number of multi ray points varies considerably between the systems.

Participant	System	av. no. of correct tie pts. per image	elim. blunders		no. of multi ray points in object space					
			no.	%	total	2 ray pts.	3 ray pts.	4 ray pts.	5 ray pts.	6 ray pts.
LHS	HATS	62	43	7	209	125	44	26	8	6
BKG		18	10	6	66	48	9	5	3	1
EPFL		49	60	12	168	103	36	19	6	4
NLS-SF		17	21	12	60	36	16	6	1	1
NLS-SWE		22	23	10	81	56	16	6	2	1
UNSW		10	18	17	32	17	8	4	2	1
SWPH	HATS*	69	0	0	243	165	40	26	2	10
Inpho	Match AT	184	0	0	574	265	182	82	13	32
I-graph		148	0	0	508	286	154	49	11	8
CGR		160	0	0	550	334	138	54	2	22
HL		102	0	0	337	182	89	51	5	10
Ph GmbH		98	0	0	352	227	76	41	5	3
CZ	Phodis AT	358	371	10	1315	824	413	56	15	7
B-LVA		330	373	12	1245	841	335	58	7	4
GCM		495	573	11	1969	1523	384	51	7	4
LGN		349	429	12	1307	849	396	52	6	4
FGI	research systems	395	0	0	1506	1112	286	74	26	8
TUM		325	0	0	1005	473	285	148	58	41
DIIAR		123	354	24	524	475	39	9	1	0
OUAT		147	0	0	493	285	122	62	0	24

Table 4: Results for the test data set Montserrat, blunder elimination and multi ray points.

3.2. Accuracy analysis

Again, only the results for Montserrat are presented. In table 5 the values σ_0 from the robust bundle adjustment (both in pixels and in µm), σ_{FI} (in µm) and the RMS values in object space (in cm) can be found. A look at the individual figures reveals some interesting findings:

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

Participant	System	σ_o		σ_{FI}	RMS values		
		[pel]	**[μm]**	**[μm]**	**X [cm]**	**Y [cm]**	**Z [cm]**
LHS	HATS	0.19	5.8	4.9	8.9	8.5	12.5
BKG		0.10	3.1	9.7	10.6	7.8	35.8
EPFL		0.20	6.0	13.4	17.4	14.8	42.2
NLS-SF		0.22	6.5	12.2	15.5	17.5	40.4
NLS-SWE		0.25	7.4	18.3	50.0	41.8	42.9
UNSW		0.14	4.3	17.6	28.6	32.5	66.9
SWPH	HATS*	0.21	6.4	5.4	7.8	8.6	55.7
Inpho	Match AT	0.11	3.3	11.4	13.9	10.1	17.9
I-graph		0.20	6.0	7.2	14.3	10.0	30.4
CGR		0.14	4.3	5.9	6.5	7.3	15.0
HL		0.15	4.6	10.6	17.7	11.9	16.2
Ph GmbH		0.17	5.2	6.3	9.4	10.0	50.6
CZ	Phodis AT	0.22	6.7	6.4	19.6	13.7	14.1
B-LVA		0.21	6.2	5.0	7.7	8.8	16.7
GCM		0.19	5.7	5.2	9.5	9.9	12.5
LGN		0.20	5.9	4.4	5.1	5.3	15.6
FGI	research systems	0.18	5.4	5.5	8.0	5.6	32.5
TUM		0.32	9.6	4.5	7.0	6.1	9.2
DIIAR		0.25	7.4	20.1	28.3	23.2	65.5
OUAT		0.25	7.4	13.6	24.0	15.9	29.9

Table 5: Results for the test data set Montserrat, accuracy figures.

- The standard deviation σ_o of the tie point coordinates generally lies between 0.1 and 0.2 pixels or 3 and 6 μm. This result has been obtained although the expectation for the accuracy of the image coordinates as expressed in $\sigma_{o,\ a\ priori}$ were set to only 1/3 pixel (see above). Some systems yielded larger values for σ_o. At least for TUM this result was to be foreseen, since the version used for the test relies uniquely on feature based matching without a matching refinement stage (see again table 3).
- While the σ_o column seems to suggest correct results for all participants an inspection of σ_{FI} reveals the opposite. Some systems obtained a high accuracy in the order of 0.2 pixel and a good agreement between σ_o and σ_{FI}. Thus, the exterior orientation parameters computed in the robust bundle adjustment are confirmed. In many cases, however, σ_{FI} is significantly larger than σ_o. This disagreement demonstrates that in AAT the σ_o value from the robust adjustment alone cannot be considered as an indicator for the quality of the aerial triangulation results. The reason is that in contrast to analytical photogrammetry in AAT an appropriate point distribution in each image and proper connections between the images and strips are not necessarily ensured. Blocks generated from rather few tie points (BKG, NLS-SF, NLS-SWE, UNSW) or from an overwhelming number of 2-ray points (DIIAR) were found to be severely deformed.
- Discrepancies between σ_o and σ_{FI} also exist in other cases (EPFL, Inpho, HL, OUAT). In order to further analyse these results all the residuals of the forward intersection were plotted. It was found that for a number of participants points in the overlapping area between the first and the second image strip the residuals in the flight direction are unacceptably large. Apparently, most matching algorithms had major difficulties in the mountainous and forested area connecting the first and the second strip.
- When looking at the RMS values the size of the deformations is quantified. Only for the blocks of 6 participants (LHS, CGR, B-LVA, GCM, LGN and TUM) out of 20 the RMS values are sufficiently small to consider the block free of deformations. In all other cases partly severe

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

deformations were found. RMS(Z) is more sensitive to distortions than σ_{FI}, note for instance the results of SWPH, I-graph or Ph GmbH.

Given these results the block was split up into two subblocks, one containing only the first strip, the other one containing strips no. 2 and 3, and the complete analysis was repeated with both subblocks. For NLS-SWE and DIIAR the first strip contained too few tie points, and thus no results were obtained. For many participants the results improved and fulfilled the expectations, however, in cases where very few tie points per image were used only the strips 2 and 3 still yielded deformed blocks. The block Montserrat has shown the potential but also the limitations of the available AAT systems. For the correct blocks an accuracy of about 4.5 to 6 µm in σ_{FI} corresponding to approximately 0.15 to 0.2 pixel was reached, but this was only the case for very few participants. It is interesting to note that success and failure occurred with one and the same system, and a demonstrated failure was not signalised by the systems. The results for the other data sets were found to be better than those for Montserrat. In case enough tie points per image were generated a correct solution was obtained. The accuracy of the tie points in image space was in the same range as for Montserrat.

4. CONCLUSIONS OF THE TEST

Compared to the test goals (see chapter 2) and taking into account the results of all participants and all data sets the following conclusions can be drawn (it should be emphasised again, that point extraction is considered to be a totally **autonomous** process within AAT):

- A good **geometric block stability** can be guaranteed, if and only if a sufficiently large number of tie points (say 100 to 300 per image) is extracted. The reason is that local matching procedures, as they are employed in the tested systems in order to achieve an acceptable level of accuracy, are subject to blunders, and these blunders can only be reliably eliminated if their percentage is relatively small. If too few points are extracted the resulting block can be heavily deformed. Within the test this problem occurred mainly for results generated with HATS. As mentioned before, HATS calls upon the operator if points are missing or need to be remeasured, but this feature was deliberately not used in the test.
- Especially in larger blocks the geometric stability also depends on the number and distribution of the available ground control points (GCP) and/or the quality of the direct measurements for the orientation parameters from GPS and/or INS. Such information can lead to a somewhat reduced number of necessary tie points per image. As mentioned before, however, no such effects were investigated within the test.
- The high redundancy in the adjustment leads to a smaller theoretical standard deviation and an improved reliability for the exterior orientation parameters as compared to analytical photogrammetry. These parameters, of course, must be regarded as the prime result of AAT.
- While the significance of a large number of multi ray points is not as high as in analytical photogrammetry neglecting this aspect too much can also lead to severe block deformations. In the test all commercial systems generated enough multi ray points, but it seems safe to predict that more emphasis should be concentrated on this point.
- Under favourable conditions (open and flat terrain, good texture as given in the data set Echallens) the **accuracy of the tie point coordinates** as expressed by σ_o can reach 0.15-0.2 pixels or 3-4 µm using only natural tie points if least squares matching is employed for coordinate refinement. In one of the projects Match AT has even achieved 0.11 pixel or 2.2 µm. In analytical photogrammetry a comparable accuracy has only been achieved using signalised points.
- Taking all test results into account a realistic value for σ_o lies in the range of 0.2-0.3 pixels or 4 - 9 µm (again with only natural tie points and least squares matching), at least when the images

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

were scanned with a pixel size of 20-30 µm. The values are rather similar across the different systems. Since most systems use least square matching in the final coordinate measurement this result seems plausible. In this test the effect of pixel size was not separately investigated. Experience and the literature (e.g. Ackermann, Krzystek 1997) suggest that pixel sizes smaller than about 20 µm will not increase the accuracy of the tie points accordingly.

- **Limitations of existing systems** showed up in the Montserrat example which contains mountainous and forested terrain. Some participants failed to produce correct and accurate results. The strip connection seems to be the weak point.
- Failure to produce an acceptable result is not indicated by the systems (with the partial exception of HATS, see above), because internal self control is not sufficiently accounted for. Elements of self control are the individual matching results, the distribution of the tie points and the number of multi ray points within the block, the measurement accuracy, and the covariance matrix of the unknowns. As was shown in a number of cases the σ_0 of the block adjustment is by itself not a valid indicator of errors or deformations within the block. The adjustment theory developed for analytical photogrammetry including measures for reliability and blunder detection and elimination seems to be the proper starting point for the necessary improvements.
- Due to the large amount of required observations (see above) the self control mechanism should be automatic.
- A minimum requirement for assessing the quality of the results is a graphical output similar to the plots produced during the test analysis. In larger blocks one should be able to roam through the whole block and zoom in and out in such graphical representations of the results, possibly even with the images as back drops.
- It is interesting that both success and failure occurred partly within one and the same system. This suggests that an extensive amount of experience in handling the software is necessary in order to appropriately tune any available free parameters. Taking also the results into account which due to gross errors are not contained in the presented tables this experience seems to be especially necessary for using Match AT and for HATS. If the number of free parameters cannot be significantly reduced additional effort should be focused on training of the AAT operators.

Obviously, not all topics related to a complete system analysis were investigated within this test. For instance, issues related to an economical use (e.g. the time and cost needed for preparation, computation, and post processing) have not been considered. Furthermore, the behaviour of AAT systems for larger and non-regular blocks, and the influence of control information were not investigated. From the results obtained, it can be concluded that the AAT systems of 1997 after only a few years of market presence, showed a remarkable level of performance. A number of details, however, need further refinement.

5. DEVELOPMENT IN AAT SINCE 1997

As mentioned before the test results were obtained with software versions available in 1997. In the meantime a number of new developments have taken place. Many of these developments were motivated by customer feedback (Kersten, O`Sullivan 1996; Hartfield 1997; Kersten et al. 1998; Köhler 1998; Käser et al. 1999; Kersten 1999; Masala 1999; Urset, Maalen-Johansen 1999), some of them were perhaps also a reaction on the results of the OEEPE-ISPRS test. Many of the improvements were exhibited during the OEEPE workshop “Automation in Digital Photogrammetric Production” held in Paris in June 1999:

- LHS has integrated a robust bundle adjustment into their solution, which is now called APM for automated point measurement. In APM there is the possibility to import a DTM as initial

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

information for the terrain height. LHS also offers a variety of graphical output to assess the results of the automatic process.
- A new matching strategy was adopted for Match AT in order to reduce the correlation of the observations within a point cluster and to increase the number of multi ray points.
- Phodis AT now creates more multi ray points, especially between image strips.

Also two new commercial systems, one by ERDAS called OrthoBase and another one by VirtuoZo were on display in Paris. OrthoBase is a commercial version of the IPI development (Wang 1996) and has been benchmarked with some of the OEEPE-ISPRS data sets. The results reported by ERDAS are comparable to those determined in the test described in this paper. The VirtuoZo system seems to be similar in design to HATS, further details were not available. According to both companies these systems are primarily designed for applications within GIS and engineering projects, rather then for the high-end photogrammetric market.
What then is the current status of AAT? After some preliminary enthusiasm shortly after commercial systems became available a more realistic evaluation is now possible based on test results and the cited user reports. Progress over the last years has been fast and steady, and considerable cost savings have been reported by various users when comparing analytical to automatic aerial triangulation. Three main issues have turned out to be decisive for further improvement:

(1) In AAT systems tie points are selected using local interest operators only, the criteria used in analytical photogrammetry (a point should lie on a flat and stable area and should be well visible in all overlapping images) are ignored. Consequently, points sometimes lie on moving cars, or in the middle of lakes and more often on the top of trees leading to many incorrect results. Also, points often lie on shadows edges. Shadow points within a strip are of course not critical because these images are acquired immediately one after the other. For strip connections the situation is often worse, because especially for long strips a considerable amount of time can go by before the plane comes back to the same area, and in the meantime the sun and thus the shadow has moved. In DTM matching it was realised some time ago when buildings and vegetation needed to be removed from the digital surface model in order to obtain a bare Earth DTM that despite the fact that a DTM is a purely geometric product a limited interpretation of the imagery was necessary. In AAT the same necessity arises when it comes to a proper selection of suitable and well distributed tie points.
(2) Large height differences within image blocks, imagery of different scale, flown on different dates, and/or with noticeable azimuth differences are a major challenge to existing software systems, and in most cases lead to incorrect results.
(3) Internal self control is not sufficiently accounted for. Elements of self control are the individual matching results, the distribution of the tie points and the number of multi ray points within the block, the measurement accuracy, and the covariance matrix of the unknowns. As was shown in a number of cases the σ_0 of the block adjustment is by itself not a valid indicator of errors or deformations within the block. The adjustment theory developed for analytical photogrammetry including measures for reliability and blunder detection and elimination seems to be the proper starting point.

In summary in a production environment fully autonomous tie point extraction while feasible in some cases must in general be followed by a verification and editing stage carried out by a human operator. Therefore, software development needs to be concentrated on designing more user friendly man-machine interfaces for an efficient verification and editing of the AAT results including a stereo measurement capability for high accuracy requirements. Promises were made that photogrammetry would become easier in the digital era. However, since the job of the operator has largely shifted from routine measurements to supervising an automatic process, these promises seem to be hard to fulfil. On the contrary, efficient training seems to become a critical issue in successfully handling the related

'Photogrammetric Week '99' D. Fritsch & R. Spiller, Eds., Wichmann Verlag, Heidelberg, 1999.

software. Work is also needed to create proper quality specifications for the results of automatic aerial triangulation, especially for the parameters of exterior orientation.
Of course, further developments of AAT must be seen with regard to the improvements in the direct measurement of the image orientation parameters using GPS and INS. Recent advances in this area have given rise to the question whether and when aerial triangulation regardless of the degree of automation might become obsolete. While this topic is out of the scope of this paper (see e.g. Wewel et al. 1998; Colomina 1999 and Cramer 1999 for excellent results, discussions and comparisons) it is obvious that a combination of all three techniques (AAT, GPS and INS) will yield the most accurate and most reliable results. Details of such a combination are at present a focus of research and development.

6. ACKNOWLEDGEMENTS

The team effort at TU Munich was a prime condition for conducting this test. Not a single data set would have been distributed, not a single bundle adjustment run would have been computed, and not a single plot would have been created without the help of Konrad Eder, Rüdiger Brand and Roland Winkler. I am also grateful to the organisations and individuals having provided the test data. Finally, this test would not have been possible without the enthusiasm of the test participants. Our team was very surprised and excited by the large interest in the test and by the number of groups participating. They all put a considerable effort into processing the data sets while keeping up with a rather tight time schedule. More often than not processing had to be done in addition to the every day work load. We hope that we have fulfilled their expectations.

7. REFERENCES

Ackermann, F. and P. Krzystek (1997): Complete automation of digital aerial triangulation, Photogrammetric Record (15) 89,pp. 645-656.

Brand, R. and C. Heipke (1998): A system for automatic aerial triangulation, IntArchPhRS (32) 2, pp. 27-34.

Colomina, I. (1999): GPS, INS and aerial triangulation: what is the best way for the operational determination of photogrammetric image orientation?, Proceedings, ISPRS Conference "Automatic extraction of GIS objects from digital imagery", Munich, September 1999.

Cramer, M. (1999): Direct geocoding – is aerial triangulation obsolete? in: Photogrammetric Week ´99, D. Fritsch, R. Spiller (Eds.), Wichmann, Karlsruhe.

Förstner, W. (1991): Statistische Verfahren für die automatische Bildanalyse und ihre Bewertung bei der Objekterkennung und –vermessung, DGK C, p. 370.

Forlani, G., Pinto, L. and M. Scaioni (1998): The TRIADIGIT program for automatic aerial triangulation, IntArchPhRS (32) 2, pp. 67-73.

Hartfield, P. (1997): Higher performance with automated aerial triangulation, in: Photogrammetric Week ´97, D. Fritsch, D. Hobbie (Eds.), Wichmann, Karlsruhe, pp.109-113.

Heipke, C. and K. Eder (1996): Performance of tie point extraction in automatic aerial triangulation, description of the OEEPE-ISPRS test, Technische Universität München.

Heipke, C. and K. Eder (1999): Performance of tie point extraction in automatic aerial triangulation, OEEPE Official Publications No. 36, pp. 125-185.

Honkavaara, E. and A. Hogholen (1996): Automatic tie point measurement in aerial triangulation, IntArchPhRS (31) B3, pp. 337-342.

Käser, C., Czaka, T. and T. Kunz (1999): Digital aerotriangulation for map revision with Match AT, Proceedings OEEPE Workshop "Automation in Digital Photogrammetric Production", Paris, June 22-24, 1999.

Kersten, T. (1999): Results of digital aerial triangulation using different software packages, Proceedings OEEPE Workshop “Automation in Digital Photogrammetric Production”, Paris, June 22-24, 1999

Kersten, T., Häring, S. and W. O´Sullivan (1998): Digital aerial triangulation in alpine regions – a challenge, IntArchPhRS (32) 2, pp. 149-156.

Kersten, T. and W. O´Sullivan (1996): Experiences with the Helava Automated Triangulation System, IntArchPhRS (31) B3, pp. 591-596.

Köhler, G. (1998): Digitale Photogrammtrie – Konzept und Umsetzung beim Hessischen Landesvermessungsamt, ZfV (123) 9, pp. 301-306.

Klein, H. and W. Förstner (1984): Realisation of automatic gross error detection in the block adjustment program PAT-M43 using robust estimators, IntArchPhRS (25) 3.

Masala, B. (1999): Experiences on automatic aerial triangulation, Proceedings OEEPE Workshop “Automation in Digital Photogrammetric Production”, Paris, June 22-24, 1999

Paszotta, Z. (1998): Report on the OEEPE-ISPRS test on “Performance of tie point extraction in automatic aerial triangulation”, unpublished.

Schenk, T. (1997): Towards automatic aerial triangulation, ISPRS Journal of Photogrammetry & Remote Sensing (52) 3, pp. 110-121.

Tang, L., Braun, J. and R. Debitsch (1997): Automatic aerial triangulation – concept, realisation and results, , ISPRS Journal of Photogrammetry & Remote Sensing (52) 3, pp. 122-131.

Urset, A. and I. Maalen-Johansen (1999): Automated triangulation in nordic terrain – experiences and challenges with Match AT, Proceedings OEEPE Workshop “Automation in Digital Photogrammetric Production”, Paris, June 22-24, 1999.

de Venecia, K., Miller, S., Pacey, R. and S. Walker (1997): Experiences with a commercial package for automated aerial triangulation, ASPRS/ACSM Annual Convention, (1) pp. 548-557.

Wang, Y. (1996): Structural matching and its applications for photogrammetric automation, IntArchPhRS (31) B3, pp. 918-923.

Wewel, F., Scholten, F., Neukum, G. and J. Albertz (1998): Digitale Luftbildaufnahme mit der HRSC – ein Schritt in die Zukunft der Photogrammetrie, PFG 6, pp. 337-348.