International Archives of Photogrammetry & Remote Sensing, 32(2):136-139

TESTS ON AUTOMATIC DEM GENERATION IN A DIGITAL PHOTOGRAMMETRIC WORKSTATION

G. E. Karras, N. Mavrogenneas, D. Mavrommati, N. Tsikonis
Department of Surveying, National Technical University of Athens, GR–15780 Athens, Greece
e-mail: gkarras@central.ntua.gr

Commission II, Working Group 8

KEY WORDS: Matching, DEM, Accuracy

ABSTRACT

The massive introduction of modern digital photogrammetric workstations into routine mapping activities has obviously changed the tools and context of photogrammetric practice. Notably, DEM generation by image matching has become a standard softcopy photogrammetric technique. Inevitably, however, new questions have also arisen, often perplexing users, regarding geometric accuracy and success rates or expediency of user-defined matching strategies and suitable tuning of numerous parameters. In this contribution, tests on automatic DEM generation with the SoftPlotter system of Vision International are reported, employing several possibly "representative" DEM patches. The matching results were assessed against manual measurement (the precision of which was estimated as better than 0.1‰H). A basic finding was that, consistently, 20% of raw DEM nodes had elevations significantly different from those measured manually. In fact, 60% of nodes were correctly found directly by correlation (the remaining 20 % were interpolated elevations), which is an indication for the actual "intrinsic" success rate of matching encountered here.

1. INTRODUCTION

Digital photogrammetry is now being rapidly accepted as a standard tool in mapping, cartographic and GIS applications. In contrast to many close-range projects where own software is extensively employed, all orientation and measuring operations generally take place within powerful, commercially available, digital photogrammetric workstations (DPWs). The massive turn towards using DPWs has occurred within a short period of time; internationally, their growing number is now surpassing that of analytical plotters. In Greece, for instance, and chiefly thanks to the ongoing National Cadastre Project, two dozens of DPWs (exceeding in number the APs ever used in this country) have been installed in private firms in the last few years and are currently involved in routine mapping work.

Admittedly, DPW has revolutionised daily photogrammetric practice and multiplied possibilities, above all by introducing automation via image matching techniques. Yet, it is a complicated 'multi-parametric' system addressing the market and, often functioning as a "black box", raises – not only to the common user, one might add – as many questions as the answers it provides. When viewed within the framework of their somewhat sudden massive introduction into mapping practice, this may well account for the requests of users/customers for data, mainly regarding geometric accuracy of DPWs, referred to in literature (Walker, 1995). Data are now not all that scarce but, explicably (differences among systems, multiplicity of tested parameters, varying evaluation approaches), they convey a rather contradictory picture of DPW performance.

Automatic DEM generation, in particular, now a standard technique in softcopy photogrammetry based on sophisticated software and equipped with efficient editing tools, is an outstanding feature of DPWs; yet success rates seem to vary strongly, mostly with terrain and image characteristics, making interactive editing a key factor for satisfactory products. Thus, a striking range of data on geometric accuracy of raw DEMs have been reported, ranging from below 0.1‰H to above 1‰H (Miller & de Venecia, 1992; Abanmy et al., 1995; Krzystek & Ackermann, 1995; Scarpace & Saleh, 1996; Gasior, 1996). In this context, the importance of interactive editing has been stressed (Boniface, 1994). Generally, acceptable results are expected only at rather small scales and over open terrain; in other cases, editing of match-results requires efforts almost invalidating the assumed advantages (Leberl, 1994). Besides, it is possible to change one to two dozens of (sometimes "obscure") parameters of automatic DEM generation; this, of course, allows numerous variations in matching strategies which, in turn, drastically affect final results (Smith & Smith, 1996).

In this context (which, of course, is quite understandable in view of the complex nature of the problem but still no less confusing for many users), the tests carried out here rely on two assumptions. First, that – since users will not easily investigate the effects of the numerous parameters and trust user-defined correlation strategies – the default values as suggested by manuals for each case would be adopted at this stage. On the other hand, most tests reported in recent literature rely on comparisons with elevation data from analytical plotters; however, few users are expected to have direct access to APs.

Therefore, second, manual DEM measurement within the digital system itself was to serve as reference. Generally, DPW precision (: repeatability) of measured spot heights has been shown to verge on accuracy (±0.3 pixel for both has been reported by Walker, 1995). This has not been contradicted here using elevation data from a Wild BC3. Hence, all automatically generated DEMs were examined node by node and edited. The rms differences of 'edited'-'unedited' DEMs (differences from manual measurement by the operator) were regarded as describing inaccuracy attributed to both manual and automatic collection.

2. TEST IMAGES AND TERRAIN

Choice of 'typical' images and terrain is crucial if claims to 'representative' results are to be realistic. Taking this into account, care has been taken to select models possibly characteristic of several routine jobs. A total of five DEM patches from three good quality aerial 23/153 stereopairs (M_1–M_3) were used. Additionally, a non-conventional low-altitude stereopair (L) of a prehistoric excavation, taken with a Wild P31 camera from a helicopter, was studied. Table 1 gives the essential information.

Table 1
Images and DEMs Used

	M_1	M_{21}, M_{22}	M_{31}, M_{32}	L
Scale Factor	10.000	7.000	6.500	300
H (m)	1.550	1.090	1.020	30
B:H	0.53	0.53	0.53	0.24
Pixel Size (ìm)	25	15	25	42
Spacing (m)	10	10	10	0.3
Nodes	1190	441, 930	1517, 1296	3286

Chosen terrains, also regarded as typical for many applications, may be grossly described as follows:

M_1 slope: 35 % undulating, open (Fig.2)
M_{21} slope: 20 % rolling; sparse bushes and trees (Fig.3)
M_{22} slope: 3 % flat; trees; sparse buildings (Fig.4)
M_{31} slope: 10 % chiefly flat; small dried up river; sparse buildings and trees
M_{32} slope: 2 % flat; sparse buildings and bushes
L flat; $\Delta H < 1$ m; breaklines; shadows (Fig.5).

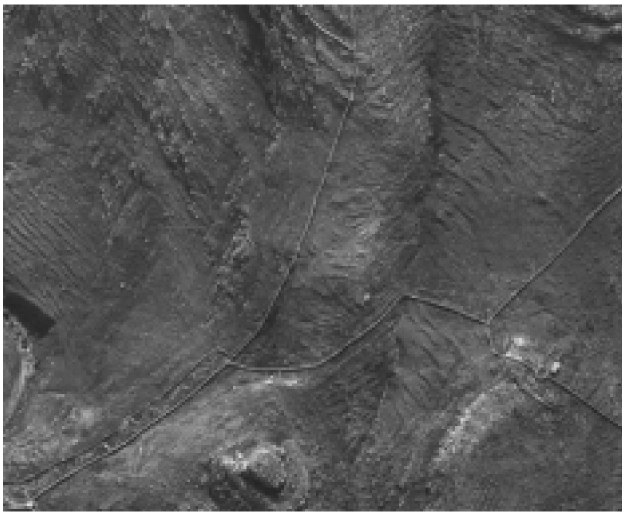

Figure 2. Area of model M_1

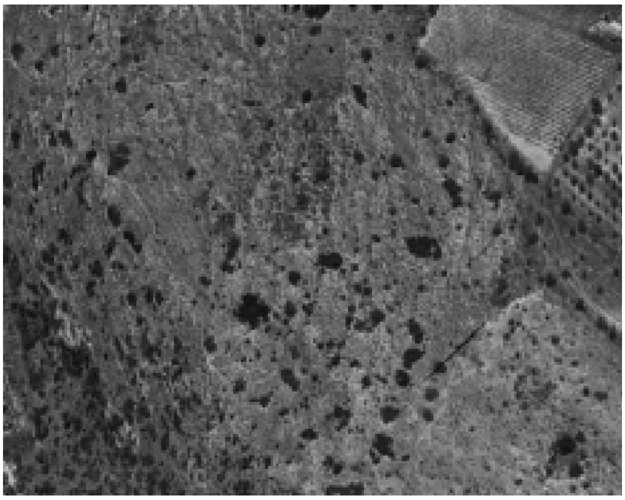

Figure 3. Area of model M_{21}

Figure 4. Area of model M_{22}

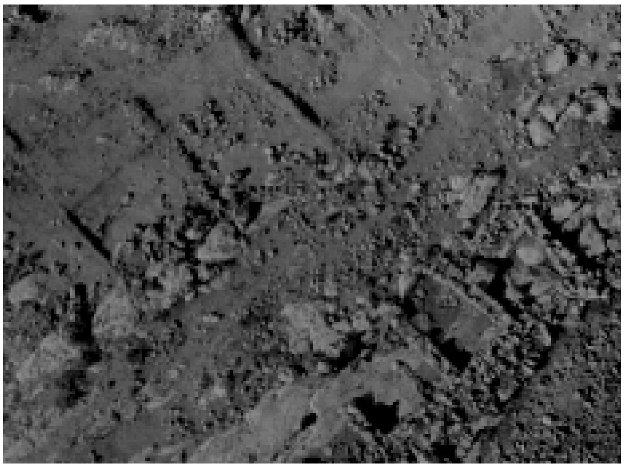

Figure 5. Area of model L

All tests were performed with the SoftPlotter™ system of Vision International which employs proprietary digital correlation techniques in a matching algorithm incorporating a hierarchical approach. For reasons given above, default values suggested by the User's Manual (1997) were used leaving user-modifiable strategies to future investigations.

3. SPOT HEIGHT MEASUREMENT

Initially, the repeatability of spot height measurement was estimated. The purpose was to assess the accuracy of manual versus automatic (terrain following cursor mode: TFC) spot height measurement, and obtain measures for evaluating automatic DEM results. In all 4 stereomodels 16 points – not detail points but rather elevations on flat surfaces – were measured 10 times, in turn among many more to exclude "automatic repetition" by the operator; points were selected at different heights, which each time implies changes in relative position of epipolar images. Care was taken to choose both 'good' and 'bad' points. Table 6 presents the standard deviations (σ_H) of height measurements and their translation to parallax errors (σ_p) through scale and base-to-distance ratio. A further indication is given in the last two rows which show the percentage of automatically measured points falling outside two confidence intervals of the manual measurements.

For manual measurement, a high precision $\sigma_{HO} < 0.1‰H$ was obtained for the three aerial models, with parallax repeatability σ_p in the range of 0.3 – 0.4 pixels. For model L results are not directly comparable due to its less favourable base-to-distance ratio but parallax precision was of the same (or slightly smaller) magnitude.

Table 6 Spot Height Measurements (O: operator; A: automatic)				
	M_1	M_2	M_3	L
σ_{HO} (cm)	14	9	9	0.9
$ó_{HA}$ (cm)	22	11	20	2.2
σ_{HO}:H	0.09‰	0.08‰	0.09‰	0.31‰
σ_{HA}:H	0.14‰	0.10‰	0.19‰	0.75‰
σ_{pO} (pixels)	0.30	0.42	0.29	0.18
σ_{pA} (pixels)	0.45	0.50	0.64	0.42
outside $H_O \pm 2\sigma_{HO}$	37%	23%	37%	25%
outside $H_O \pm 3\sigma_{HO}$	10%	2%	22%	12%

For elevations collected automatically with cursor kept on the terrain by digital correlation, a satisfactory TFC precision of $\sigma_{HA} < 0.2‰H$ was obtained (σ_p between 0.4 – 0.7 pixels). Matching precision thus ranged from equal to half that of manual measurements while hinting at a possible effect of scanning resolution (cf. M_2 – M_3).

As far as the actual deviations in elevation are concerned and regarding operator measurements as more accurate, 30% of individual automatic measurements were found to fall outside the range $H_O \pm 2\sigma_{HO}$, indicating a trend towards significantly different elevation measurements in the automatic mode. Evidently, L is not directly comparable mainly due to its different base-to-distance ratio. Here again, however, the operator measurements appeared as being twice as precise.

4. EVALUATION OF AUTOMATIC DEM GENERATION

4.1. Comparison with Data from an Analytical Plotter

As mentioned above, for users not having access to APs a means may be required for checking their own results. This could succeed as a comparison with careful point-to-point height measurements with stereoviewing. Here, elevation data from a Wild BC3 were available only for M_2. Unfortunately, these were not spot heights but directly digitised contours to which a DEM was later interpolated. Ignoring interpolation error, the estimated DEM accuracy should be close to 0.25‰H (Kraus, 1994). For models M_{21}, M_{22}, the comparison of this reference DEM with the automatically derived DEMs before and after editing gave the results of Table 7.

Table 7 ÄH Differences: Comparison with Analytical Plotter				
	Range (m)	Mean ± σ (m)	rms	
Reference – Undited DEMs				
M_{21}	-4.18 / 2.76	-0.31 ± 0.62	0.70 m	0.64‰H
M_{22}	-2.11 / 6.86	-0.15 ± 0.42	0.45 m	0.41‰H
Reference – Edited DEMs				
M_{21}	-0.84 / 0.95	-0.16 ± 0.23	0.28 m	0.25‰H
M_{22}	-1.25 / 0.90	-0.06 ± 0.25	0.26 m	0.24‰H

The edited DEMs may be viewed as practically equivalent to the reference DEM, conforming to the assumption that repeatability is close to accuracy. This is a further indication that edited DEMs may indeed be used as reference.

4.2. Comparison of Edited and Unedited DEMs

All individual nodes of the six DEM patches were checked stereoscopically and corrected carefully with the editing tools of SoftPlotter™. Thus, edited DEMs actually reflect what an operator would measure manually. These results are gathered in the following Table 8.

Table 8 Edited Nodes (%) of Automatically Generated DEMs					
M_1	M_{21}	M_{22}	M_{31}	M_{32}	L
29.6	30.4	32.5	27.9	33.1	29.6

The results show a consistency for all terrain types, with about 30% of nodes needing editing. In literature, varying estimates are reported based on different matching software and terrain. These range from 3% (Mikhail, 1992) to 20% (Saleh & Scarpace, 1994). Yet estimates may not be directly comparable to each other as “need for editing” is probabaly not understood in a uniform way. Here, for instance, some corrections included in Table 8 fell within the corresponding precisions $\pm\sigma_{HO}$ of manual measurement. In this sense, further checks were made with the three DEM patches M_1, M_{21}, M_{22}. It was assumed that a node needed no editing if this would result in a correction within $\pm 2.6\sigma_{HO}$ (stereoviewing precision; cf. Table 6). The results are given in Table 9.

Table 9 Accuracy of Automatically Generated DEMs (editing of significant differences in elevation)			
	M_1	M_{21}	M_{22}
Edited Nodes (%)	20.3	19.7	20.2
ΔH Range (m)	–6.10/7.78	–5.13/2.30	–6.67/2.64
Mean $\Delta H \pm \sigma_{ÄH}$ (m)	0.26±1.12	–0.15±0.64	0.08±0.41
$rms_{\Delta H}$ (m)	1.15	0.66	0.42
$rms_{\Delta H}$:H (‰)	0.74	0.60	0.39

Compared to manual measurement, the automatic DEMs consistently showed (regardless of terrain) significant deviations in 20% of nodes. In this sense, 80% is regarded here as an estimate for success rate.

Finally, it is interesting to identify the percentage of nodes causing these large overall discrepancies in elevation. A check in this direction was to retain only the nodes giving $rms_{\Delta H} = 0.14‰H$, thus defining automatic DEMs with the same accuracy 0.1‰H as manual collection. The results are given in the following Table 10.

Table 10 Accuracy of Automatically Generated DEMs (keeping nodes giving acceptable overall accuracy)			
	M_1	M_{21}	M_{22}
Used Nodes (%)	88.1	90.2	91.4
ΔH Range (m)	–0.81/0.93	–0.65/0.58	–0.43/0.54
Mean $\Delta H \pm ó_{ÄH}$ (m)	0.02±0.21	–0.04±0.15	0.05±0.14
$rms_{\Delta H}$ (m)	0.21	0.16	0.15
$rms_{\Delta H}$:H (‰)	0.14	0.14	0.14

Thus, an omission of 10% of nodes has led to acceptable results with the mean differences being close to zero and the maximal individual elevation differences in the range of $3.6 – 4.40 \times \sigma_{HO}$. All rms values have fallen drastically to 18% – 36% of those in Table 9. (In this context, Mikhail, 1992, has reported that editing in 3% of nodes improved accuracy up to 40%.)

4.3. Evaluation of Matching

Of course, the results reported above refer to raw DEMs, namely nodes either matched or (in case matching failed) interpolated to neighbouring ones. The statistical analysis of generated points provided by the system allows further assessments of success rate. In Table 11 information on two aerial DEM patches and model L are given.

Table 11
Evaluation of Matching (% of Nodes)

	M_{31}	M_{32}	L
Matched /*of which corrected*	67/*15*	70/*19*	88/*27*
"Good" /*of which corrected*	26/ *6*	23/ *8*	33/*20*
"Fair" /*of which corrected*	37/*17*	43/*18*	32/*34*
"Poor" /*of which corrected*	37/*21*	34/*26*	35/27
Interpolated /*of which corrected*	33/*53*	30/*66*	12/*49*
Successful Matching	57	57	64

Thus, data provided by the system reveal that about 70% of elevations in the aerial DEMs were directly estimated by matching; of these, 25% were internally rated "good", 40% "fair", 35% "poor". This indeed comes close to the statement of the User's Manual (1997) that, in general, a 75+% of nodes are expected to be collected automatically, with 20–30% of these rated as "good". Only 7% of the "good" points needed editing in the aerial models, 17% of the "fair" and 23% of the "poor" ones. The exclusion of all points where correlation either failed (31%) or its product had to be edited (12%) lets 57% be an estimate of the "intrinsic" success rate of matching in the aerial models tested. On the other hand, the low altitude DEM exhibited a much higher percentage of matched points followed, however, by a larger proportion of editing, thus leading to a similar overall success rate estimate of 64%.

Finally, a further empirical evaluation was attempted by visually checking all edited nodes with failed or faulty correlation. About half of these could be directly associated with obstacles (vegetation, buildings); the remaining half had to be attributed to software facing locally unsuitable, albeit not always immediately identifiable, image/terrain conditions. (It is noted that in this section all edited points were considered. Thus, results are related to Table 8.)

5. DISCUSSION

It is believed that a testing of raw DEMs against operator measurement serving as reference, though demanding in manual work, is valid and may help users avoid resort to external data. It has been estimated that, with the system tested, precisions of manual spot elevation measurement $\sigma_{HO} < 0.1‰H$ can be expected (corresponding to parallax precision $\sigma_p \approx 0.3$ pixel). For automatically measured spot elevations, imprecisions maximally exceeding by a factor of 2 those of manual measurements were obtained (probably more affected by pixel size). However, about 30% of individual differences fell outside the range $\pm 2\sigma_{HO}$.

A consitent 30% of raw nodes were edited in all models tested. Under the assumption that no editing is needed for differences within $\pm 2.6\sigma_{HO}$ (stereoviewing precision), an equally consistent 20% of nodes had to be corrected. This is viewed here as a measure of raw DEM accuracy. (It must also be noted, however, that a consistent 10% of nodes were in fact held responsible for the overall deviations between unedited and edited models.)

Of all elevations, 70% were directly matched, 30% had to be interpolated; 17% of the former needed editing. Exclusion of edited and interpolated points leads to 60% being an estimate for the "intrinsic" success rate of matching. Thus, a gross assessment of the present tests might be that 80% of DEM nodes generated by the system were indeed correct (within the measuring tolerance of manual collection), consisting of 60% found directly by correlation plus a further 20% produced by interpolation (generally, of course, the correctness of interpolated points depends upon relief and DEM spacing). If image and terrain types used here are indeed somehow "typical", then the reported results may be viewed – within their obvious context – as representative.

Two final points are made. DEMs which might be viewed as "inaccurate" with this approach may well be accurate enough for the production of orthoimages. Second, user-modifiable correlation strategies, intended to account for variable image radiometry and terrain fluctuations, pose more complicated questions now under investigation.

REFERENCES

Abanmy, F., Khamees, H., Scarpace, F., Vonderohe, A., 1995. An evaluation of DEM and orthophoto generation in Orthomax. *Technical Papers ASPRS/ACSM Annual Convention & Exposition*, vol. 2, pp. 479-487.

Boniface, P.R.,1994. State-of-the-art in soft copy photogrammetry. *Mapping and Remote Sensing Tools for the 21st Century*. ASPRS Publication, pp. 205-210.

Gasior, D., 1996.Automatic derivation of a DTM with the Helava system. *Workshop on Application of Digital Photogrammetric Workstations*, OEEPE Official Publication no. 33, pp. 241-250.

Kraus, K., 1994. Photogrammetrie, Bd. 1. Dümmler, Bonn.

Krzystek, P., Ackermann, F., 1995. New investigations into the practical performance of automatic DEM generation. *Technical Papers ASPRS/ACSM Annual Convention & Exposition*, vol. 2, pp. 488-499.

Leberl, W. F., 1994. Practical issues in softcopy photogrammetric systems. *Mapping and Remote Sensing Tools for the 21st Century*. ASPRS Publications, pp. 223-230.

Mikhail, E.M., 1992.Quality of photogrammetric products from digitized frame photography. *International Archives of Photogrammetry & Remote Sensing*, vol. 29(B2), pp. 390-396.

Miller, S., de Venecia, K., 1992. Automatic elevation extraction and the digital photogrammetric workstation. *Technical Papers ASPRS/ACSM Annual Convention & Exposition*, vol. 1, pp. 572-579.

Scarpace, L. F., Saleh, A., R., 1996. Investigation of aerial triangulation and surface generation using a softcopy photogrammetric system. *International Archives of Photogrammetry & Remote Sensing*, vol. 31(B2), pp. 340-344.

Smith, G.D., Smith, J.M., 1996. Operational experiences of digital photogrammetric systems. *International Archives of Photogrammetry & Remote Sensing*, vol.31(2),pp.357-362.

SoftPlotter™, 1997. *User's Manual*, Version 1.7.

Walker, S., 1995. Analogue, analytical and digital photogrammetric workstations: practical investigations of performance. *Photogrammetric Record*, 15(85), pp. 17-25.