DIGITAL ELEVATION MODEL PRODUCTION BY STEREO-MATCHING SPOT IMAGE-PAIRS: A COMPARISON OF ALGORITHMS.

T Day
J-P Muller

Department of Photogrammetry and Surveying
University College London
Gower Street
London WC1E 6BT, U.K.

JANET: tday@uk.ac.ucl.cs

Three published stereo-matching algorithms have been implemented and tested on SPOT images of an area for which we have an accurate digital elevation model (DEM) available with 30m spacing. We present the results of comparison of stereo-matcher output with the DEM, and examine the errors arising and their causes.

1 INTRODUCTION

Alvey MMI-137 is concerned with producing accurate DEMs from sources such level 1A ("raw") SPOT image data within a couple of hours for a whole 6000 by 6000 pixel image pair using a transputer-based MIMD machine [1]. This paper presents an empirical comparison of the performance of three stereo-matching algorithms when applied to SPOT image data.

Previous empirical assessments of quality of automated stereo-matcher output have suffered from noise-correlated data [2 3], limited area data [4] or a-priori data of dubious quality [5 6 7].

Sub-pixel accuracy stereo-matching has already been demonstrated by many authors using quality assessment techniques based on the use of isolated random check-points (see, for example, [8]). However, the potential of stereo-matchers to produce millions of point pairs implies that reliability and blunder rate (and the ease with which blunders can be filtered out) are also important factors.

2 TEST AREA

2.1 Images

Our test data consists of three SPOT images (scene number 50-252) of Aix-en-Provence in the South of France, provided as part of the SPOT-PEPS programme. Extracts from these are shown in Figure 2. While the vertical image is free of atmospheric effects, the left image (angle of incidence $-17.5°$) is affected by haze which acts as multiplicative noise and lowers the contrast range of features. The right image ($+22.6°$) contains several completely opaque clouds, but is clear away from these. All images are affected by horizontal and vertical striping originating in the pushbroom sensor and uncorrectable by SPOT-Image for these images.

2.2 Independent Reference Data

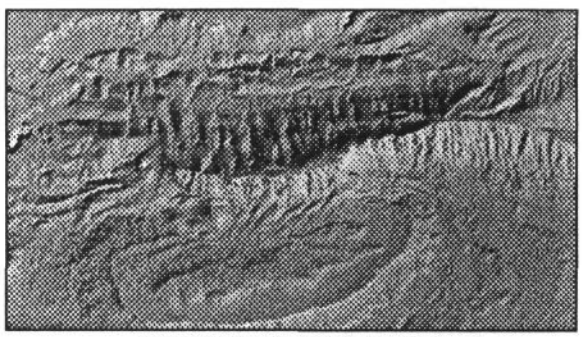
We also have a $12.42km$ by $6.9km$ DEM, with $30m$ spacing, of a region (Montagne Sainte Victoire) within the area covered by the images. The range of elevations is $191.71m$ to $1010.99m$. This was produced within the department by manual photogrammetric measurement of spot heights from aerial photographs with much higher resolution than SPOT. The DEM is unusual in that the operator measured the top of the tree canopy, where it was present, rather than attempting to measure the underlying ground level; this makes it more directly comparable with stereo-matcher output. By repeated measurement of several blocks, the accuracy of the DEM (i.e the standard deviation from the unknown ground-truth) was estimated as $1.3m$. Comparison with a lower resolution DEM of the area, supplied by IGN, indicated no systematic offset.

3 STEREO-MATCHERS TESTED

All the stereo-matchers used are adaptations of existing published algorithms, and were implemented by project collaborators (PMF and Otto & Chau on conventional hardware by UCL's Department of Computer Science [9 10] and Barnard and Thompson on transputers by the Royal Signals and Radar Establishment at Malvern [11 1]).

Figure 1: Reference DEM

Intensity range image



Lambertian shaded nadir view

3.1 Barnard and Thompson

This is an interest operator based matcher [12], using the Moravec operator [13] as a feature extractor. Since the Moravec operator does not locate features to sub-pixel accuracies, the algorithm is unable to resolve elevations in steps smaller than the height change due to one pixel disparity (e.g approximately $30m$ for the left and vertical image pair used below). However, the matched points obtained can be used as initial seed points for the Otto & Chau algorithm.

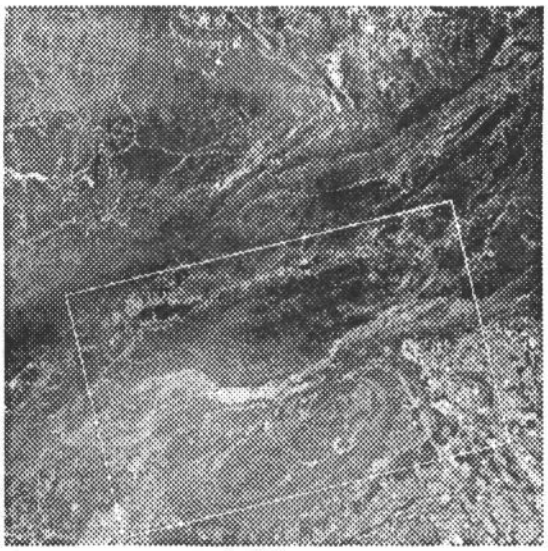
3.2 PMF

This is an edge-based matcher developed at Sheffield University as part of Alvey IKBS-025 (3D surface representations and 3D model based vision from stereo) [14]. It operates only along scanlines and therefore requires epipolar images as input. However, SPOT images are generally rotated at different angles as a function of scanline; we currently use an affine transformation to warp to near-epipolar. In practice it appears we cannot resample to true epipolar without iterative adjustment [15].

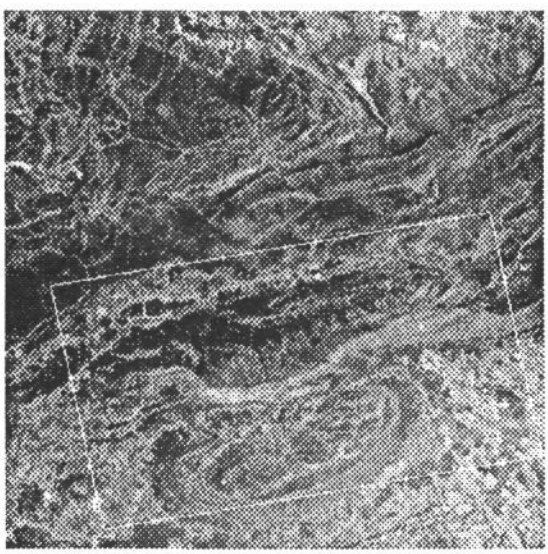
Software to produce epipolar images *when supplied with a DEM* has recently been written [16], but the PMF algorithm has not yet been applied to these due to difficulties associated with the afore-

AVC 1988 doi:10.5244/C.2.18

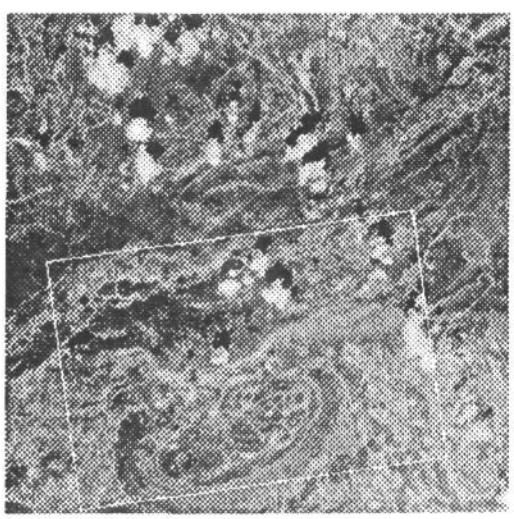
Figure 2: Extracts from left, vertical and right SPOT images showing the location of the 30m DEM. Images copyright CNES.



Left image



Vertical image



Right image

mentioned striping effects.

3.3 Otto & Chau

This stereo matcher is based on the Gruen adaptive least-squares correlator (ALSC) [17], claimed to be of extremely high accuracy (approximately 0.05 pixels, based on figures reported for aerial photography [8]). The correlator can only refine an initial estimate of the disparity at a point, and has a limited pull-in range (on the order of two pixels). However, it also produces shaping information which is used in the Otto & Chau algorithm to estimate the disparity locally and so the matcher can sheet-grow out from a few initial seed-points [18]. These could be obtained by manual measurement, or from some other stereo matcher.

4 QUALITY ASSESSMENT

Since we are comparing our stereo-matcher output for SPOT with (assumed accurate) manual measurements of aerial photography, it must be stressed that the results given below will also include errors due to the camera model (which transforms SPOT image co-ordinates to and from ground co-ordinates). We are therefore testing the accuracy of our stereo-matching system as a whole, and not just the stereo-matching component.

4.1 Disparity-space analysis

To evaluate stereo-matcher output in image co-ordinates we first transform the DEM to a 'digital disparity model' (DDM) using an appropriate camera model [19]. The DDM takes the form of a list of corresponding points in each image (thus there is a different DDM for each image pair). Given a stereo-matcher output point in one image of the pair (usually the closest to nadir, since this contains the most even distribution) we compare the stereo-matcher derived disparity vector with the disparity predicted from the manual measurements.

4.2 Analysis in ground co-ordinates

Here we consider the quality of stereo-matcher output transformed to ground co-ordinates through the camera model. We examine both the error statistics of the raw transformed points, and of a regular gridded DEM interpolated from them.

In the first case we simply compare the elevation of a stereo matched point with the nearest reference point, provided one exists within a specified distance. Obviously we would like to make this distance small to minimise the additional error due to the variation of the terrain away from the manually measured reference points (we can estimate this error from the terrain's variogram [20]), but this will in turn reduce the number of points we are evaluating.

A problem with previous methods is that they only assess the quality of the points the stereo-matcher chose to output. If the points are unevenly distributed, a gridded DEM produced from those points could be of significantly lower quality than indicated by the previous method. So we also interpolate the irregularly spaced stereo-matcher output to the same grid as the 30m DEM, using Laserscan's 'Matrix' package; the quality of the stereo-matcher derived DEM is then assessed by comparing coinciding points.

5 RESULTS

To compare the performance of each stereo-matcher implemented we have applied each to the images in Figure 2.

5.1 Barnard and Thompson

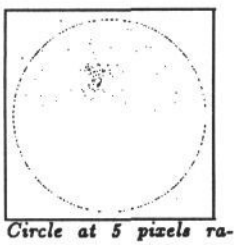
Barnard and Thompson produces relatively sparse output; 115 points from 240 by 240 pixel extracts taken from the left and vertical images, compared with an expected 560 from filtered PMF and 1800 from 5 pixel spaced Otto & Chau. We are currently unable to test this algorithm on larger images due to it being implemented on transputers virtual storage. Since it does not produce sub-pixel disparities or sufficient density for DEM generation we intend to use it only as a source of seed-points for the sheet-growing Otto & Chau stereo-matcher (and so we give no ground co-ordinate results).

From Table 1 it can be seen that the majority of the points are accurate to within 2 pixels and suitable for use as seed-points for the Otto & Chau algorithm (see below). The Chau & Otto algorithm itself filters bad seed-points and corrects those within the pull-in range.

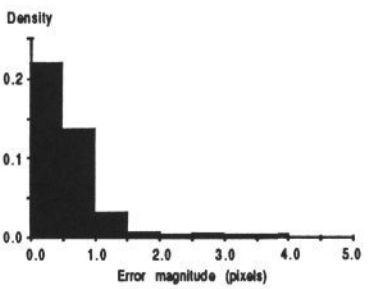
Table 1: Barnard and Thompson disparity errors (produced by comparison with nearest rederence DDM point within *two* pixels in vertical image).

B&T disparity errors	
Number of points	115
Line error mean μ_l	-1.90
Line error S.D σ_l	0.48
Sample error mean μ_s	-0.25
Sample error S.D σ_s	3.42
Maximum deviation from (μ_l, μ_s)	14.61

Figure 3: B&T output: disparity error distribution.



Circle at 5 pixels radius, centre at zero error



Error magnitude relative to (μ_l, μ_s)

5.2 PMF

The striping in the SPOT images currently presents severe difficulties for edge-detector based matching, and we have yet to find a combination of edge-detector and matcher parameters which will ignore these artifacts while still accurately locating other genuine image features. We have overcome this to some extent by resampling images to lower resolutions; the results given here for PMF were produced from a left and vertical pair with pixels approximately 24.3m across (c.f the original 10m). In these images, one pixel disparity corresponds to approximately 70m elevation.

After stereo-matching, the near-epipolar image co-ordinates were converted back to original SPOT image co-ordinates. The disparity-space results given below are for SPOT image co-ordinates. Error vectors were produced for points within one pixel of a refererence DDM point in the vertical image; statistics of these are given below. It is currently unknown how much of a quality increase we can expect from PMF applied to full resolution SPOT images if the aforementioned noise problems can be overcome.

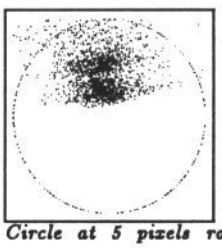
These results were obtained from the left and vertical images by supplying PMF with disparity limits (-13 to 1 pixel) and a disparity gradient limit based on an assumed maximum terrain gradient of 45°. The Marr-Hildreth feature extractor [21] [22] was used. Output was post-processed with a noise filter.

After conversion to ground co-ordinates, and comparison of points with reference points within 10m, the error statistics in Table 4 were obtained. By interpolation to the same grid as the reference DEM, the errors in Table 5 were obtained.

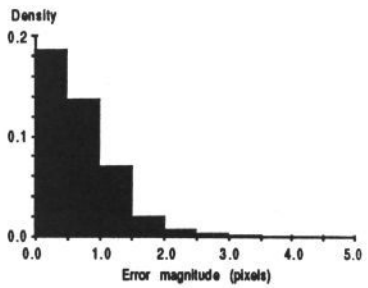
Table 2: PMF disparity errors

Filtered PMF disparity errors	
Number of points	2950
Line error mean μ_l	-2.20
Line error S.D σ_l	0.97
Sample error mean μ_s	-0.46
Sample error S.D σ_s	1.56
Maximum deviation from (μ_l, μ_s)	24.19

Figure 4: PMF output: disparity error distribution.

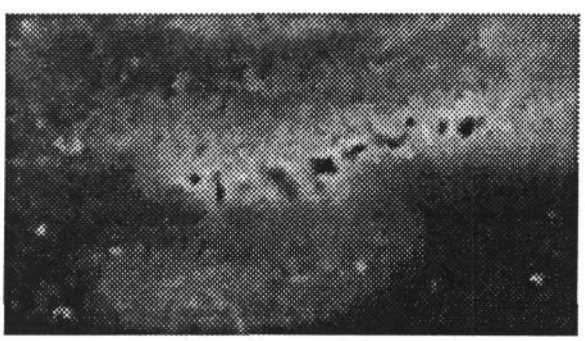


Circle at 5 pixels radius, centre at zero error

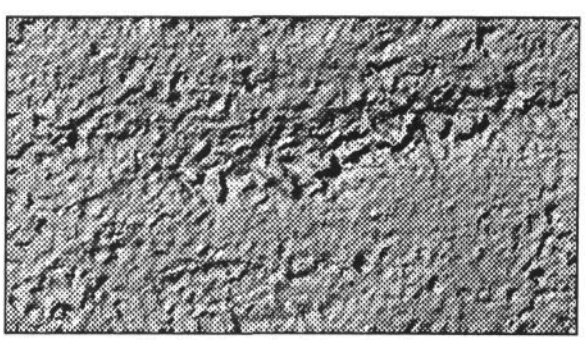


Error magnitude relative to (μ_l, μ_s)

Figure 5: PMF derived DEM



Intensity range image.



Lambertian shaded nadir view.

5.3 Otto & Chau

The results below were obtained from the vertical and right images by supplying automatically produced seed points generated by an interest-operator based matcher developed for this purpose by the Department of Computer Science. A sheet was then grown outwards from these points at five pixel intervals, using a 21-by-21 pixel (in the vertical image) correlation patch. Blunder detection was by parametric constraint [23], with post-processing using geometric considerations.

Comparison of the areas in which no points were matched in Figure 7 with the vertical image in Figure 2 demonstates the algorithms ability to sheet-grow around features (in this case clouds) which cannot be matched.

By comparing the lambertian-shaded views in Figure 1 and 7 it can be seen that we have failed to detect much of the fine detail present. By re-processing the data obtained with a smaller patch size (9-by-9 pixels in this case) we can correct this to some extent (see Figure 8); however, in places this introduces large blunders (note changes in maximum and minimum error in Table 4).

6 ERROR SOURCES

Errors are introduced at several stages within the image-pair to DEM process:

- Errors from the stereo-matcher itself. This includes both incorrect matches of completely different features, and small sub-pixel errors when features are correctly matched, but not perfectly, due to fuzzy edges.
- Errors from the camera-model itself. Our model is known to have deficiencies since it only approximates a polynomial distortion for the effect of attitude variations. Such errors manifest themselves as systematic shifts when we compare stereo-matcher output with the reference DDM e.g the non-zero mean disparity errors for Barnard & Thompson and Otto & Chau output.

  On conversion to ground co-ordinates non-zero mean elevation-errors may be seen on comparison with the reference DEM, but in the examples shown here the disparity error is primarily in the line direction and has little effect on elevation. There may also be a planimetric shift of the stereo-matcher derived DEM; this will increase the elevation-error variance by an amount equal to the terrain's variogram at that distance. These shifts are not constant across the images, so it is currently unknown what proportion of the deviation from the mean offset will be due to camera-model introduced error.

Other considerations are:

- The area based nature of the Gruen least-squares correlator implies that the elevations produced are some weighted average of those occuring in the correlation patch, whereas the manually measured points they are being compared with are effectively spot heights. Other than noting the effect of reducing the patch size (see also [24]), this has not been investigated.
- Errors from interpolating stereo-matcher output to generate a gridded DEM.
- In the quality assessment process, errors are introduced by attempting to compare stereo-matcher output with reference points some distance away. Currently we do not compare with an interpolated surface. For the results presented here (comparison vs. a point within $10m$ / 1 pixel), we estimate (from the terrain's variogram) that approximately $10m^2$ of the elevation error variance and 0.06pixels^2 of disparity error variance is due to this effect.

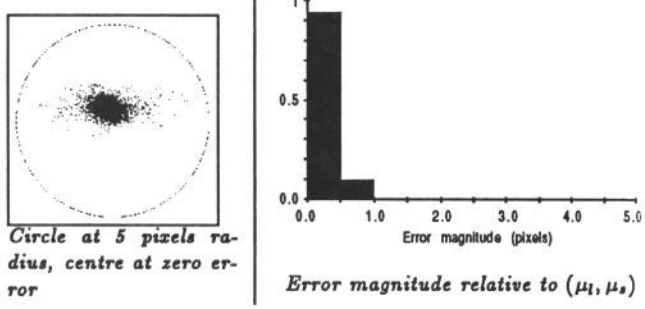
7 CONCLUSIONS

The ability of Barnard and Thompson, and Otto & Chau, to handle non-epipolar image pairs means that high accuracies for

Table 3: Otto & Chau (patch size 21) disparity errors.

Filtered PMF disparity errors	
Number of points	9916
Line error mean μ_l	-0.68
Line error S.D σ_l	0.28
Sample error mean μ_s	-0.27
Sample error S.D σ_s	0.43
Maximum deviation from (μ_l, μ_s)	6.17

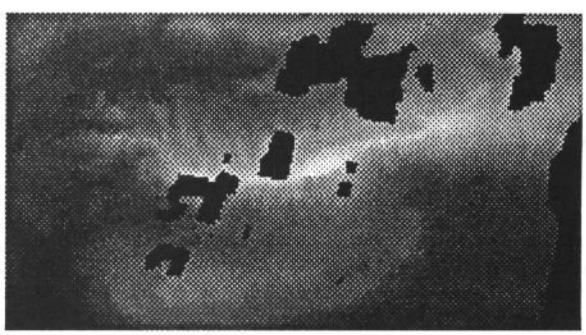
Figure 6: Otto & Chau (patch size 21) output: disparity error distribution.



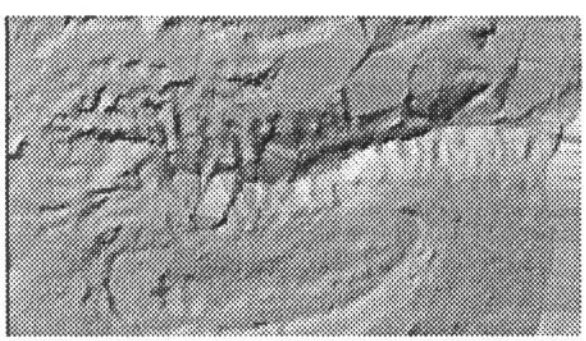
Circle at 5 pixels radius, centre at zero error

Error magnitude relative to (μ_l, μ_s)

Figure 7: Otto & Chau derived DEM (patch size 21 pixels)

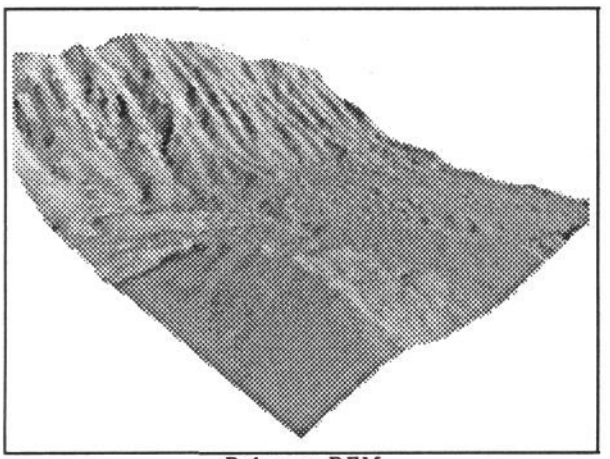


Intensity range image. No data in black areas.

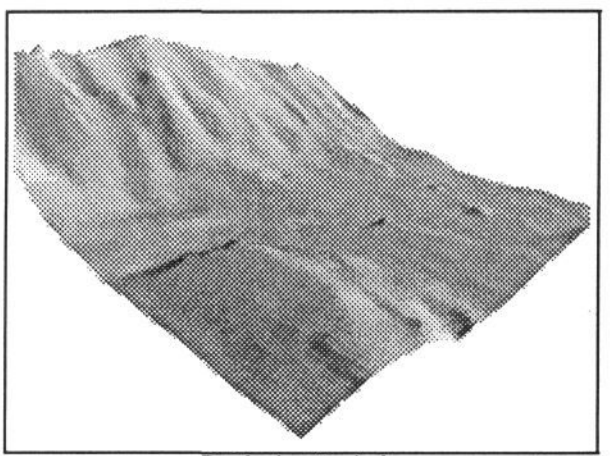


Lambertian shaded nadir view.

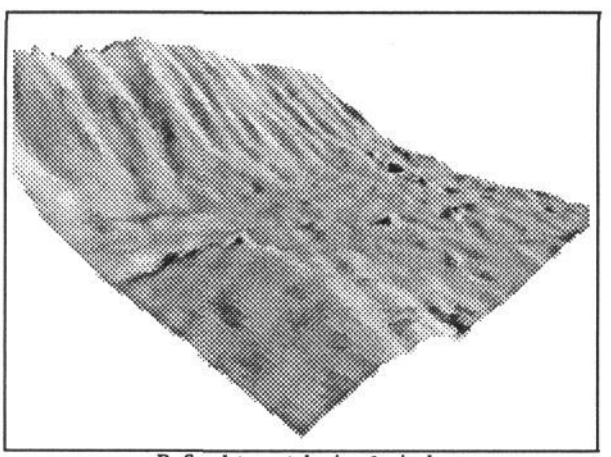
Figure 8: Comparison of 3km by 3km extracts from Otto & Chau output DEMs and manually measured DEM. Height exageration x1.5.



Reference DEM



Patch size 21 pixels.



Refined to patch size 9 pixels.

Table 4: Stereo matcher output elevation errors

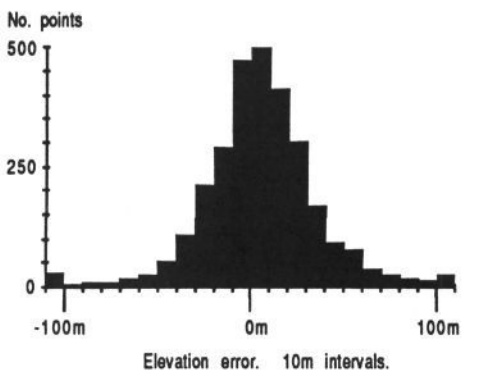
	PMF	*Otto & Chau*	
		21 pixel patch	*9 pixel patc*
Points matched within DEM area	8380	28053	2675
Points compared with reference point	2903	9794	930
Mean	3.80m	−0.09m	−0.07$_7$
S.D	45.19m	11.24m	12.71$_7$
R.M.S	45.35m	11.24m	12.71$_7$
Max.	548.70m	126.42m	547.81$_7$
Min.	−669.77m	−115.05m	−347.94$_7$
$\mid error - \mu \mid > 3\sigma$	1.24%	1.54%	3.03%

Table 5: Error statistics of stereo-matcher derived DEMs (95865 points)

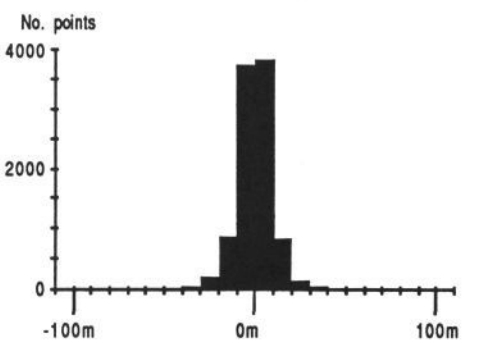
	PMF	*Otto & Chau*	
		21 pixel patch	*9 pixel patch*
Mean	−2.26m	−3.49m	−3.13m
S.D	63.26m	25.14m	27.19m
R.M.S	63.30m	25.39m	27.37m
Max.	549.14m	138.01m	500.03m
Min.	−684.86m	−268.56m	−330.58m
$\mid error - \mu \mid > 3\sigma$	2.20%	2.67%	3.03%

Figure 9: Histograms of stereo-matcher output elevation errors

PMF



Otto & Chau



geocoding/resampling to epipolar are not required (see [15] on the difficulties of creating epipolar image pairs from SPOT); they also seem to be relatively unaffected by image striping. Further, Otto & Chau has the ability to create arbitrarily dense disparity maps, and to sheet-grow around features such as clouds.

Future stereo-matcher development work is likely to concentrate on the Otto & Chau algorithm, with B&T or some similar feature-point matching algorithm to create initial seed points. However, refinements of the camera model will also be of vital importance in obtaining increased accuracies.

References

1 J-P Muller, K A Collins, G P Otto, and J B G Roberts. Stereo matching using transputer arrays. In *Proceedings of the XVIth International Congress of ISPRS, Kyoto, Japan, IAPRS 27-A3*, 1988.

2 M Ehlers. The effects of image noise on digital correlation probability. *Photogrammetric Engineering and Remote Sensing*, 51(3):357–365, 1985.

3 S H Paine. An evaluation of errors in automated digital image correlation. *Canadian Journal of Remote Sensing*, 12(2):94–102, 1986.

4 M Li. A comparative test of digital elevation model measurement using the Kern DSR-11 analytical plotter. *Fotogrammetriska Meddelanden*, 53, 1987. 13 pages.

5 P R Cooper, D E Friedman, and S A Wood. The automatic generation of digital terrain models from satellite images by stereo. *SPIE*, 660:124–135, 1986.

6 P R Cooper, D E Friedman, and S A Wood. The automatic generation of digital terrain models from satellite images by stereo. *Acta Astronautica*, 15(3):171–180, 1987.

7 R G Ley. Some aspects of height extraction from SPOT imagery. *Photogrammetric Record*, 12(72), 1988.

8 A W Gruen. High precision image matching for digital terrain model generation. *International Archives of Photogrammetry and Remote Sensing*, 26(3/1):284–296, 1986.

9 T K W Chau. *Specification and Implementation of the PMF Stereo Ranging Algorithm.* MMI-137 (UCL-CS) working paper 10, University College London, Department of Computer Science, February 1987. 7 pages.

10 G P Otto and T K W Chau. A "region-growing" algorithm for matching of terrain images. In *Proceedings of the AVC conference*, 1988.

11 K A Collins, J B G Roberts, and K J Palmer. Implementation of a feature point stereo image matching algorithm on a transputer network. In *Proc. Alvey Vision Conference*, pages 157–162, 1987.

12 S T Barnard and W B Thompson. Disparity analysis of images. *IEEE Trans. Patt. Anal. Mach. Intel.*, PAMI-2(4):333–340, 1980.

13 H P Moravec. Towards automatic visual obstacle avoidance. In *Proc. 5th International Joint Conference on Artificial Intelligence*, 1977.

14 S B Pollard, J E W Mayhew, and J P Frisby. PMF : a stereo correspondance algorithm using a disparity gradient limit. *Perception*, 14(4):449–470, 1985.

15 G P Otto. Rectification of SPOT data for stereo image matching. In *Proceedings of the XVIth International Congress of ISPRS, Kyoto, Japan, IAPRS 27-A3*, 1988.

16 M A O'Neill and I J Dowman. The generation of epipolar synthetic stereo mates for SPOT images using a DEM. In *Proceedings of the XVIth International Congress of ISPRS, Kyoto, Japan, IAPRS 27-A2*, 1988.

17 A W Gruen. Adaptive least squares correlation : a powerful image matching technique. *S. Afr. J. of Photogramm., Rem. Sens. and Cart.*, 14(3):175–187, 1982.

18 T K W Chau and G P Otto. *Design and Implementation notes for UCL-CS's Version of Gruen's Stereo Matching Algorithm.* MMI-137 (UCL-CS) report 17, University College London, Department of Computer Science, November 1987. 7 pages.

19 D J Gugan. Practical aspects of topographic mapping from SPOT imagery. *Photogrammetric Record*, 394–355, 1987.

20 J-P Muller and T Saksono. Fractal properties of terrain. *IAPRS*, 27-A3, 1988.

21 D Marr and E C Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London B.*, 207:187–217, 1980.

22 D Marr and T Poggio. A computational theory of human stereo vision. *Proceedings of the Royal Society of London B.*, 204:301–328, 1979.

23 T K W Chau. *Short note on the implementation of Parametric Constraints in Gruen.* MMI-137 (UCL-CS) working paper 24, University College London, Department of Computer Science, June 1988.

24 G P Otto and T K W Chau. Stereo matching of terrain images using a "region-growing" algorithm. In *Proceedings of the International Conference on Computer Vision, Florida*, 1988. Submitted.