Photogrammetric Record, 16(92): 271–291 (October 1998)

ASSESSMENT OF DEM QUALITY FOR CHARACTERIZING SURFACE ROUGHNESS USING CLOSE RANGE DIGITAL PHOTOGRAMMETRY

By J. B. Butler, S. N. Lane
University of Cambridge

and J. H. Chandler
Loughborough University

Abstract

This paper presents a procedure for assessing the quality of a digital elevation model (DEM) which has been applied to the output of a normalized cross correlation based stereomatching algorithm. Using semimetric photography of natural gravel river bed surfaces acquired in the field, digital photogrammetry was used to extract DEMs automatically for use in characterizing surface roughness properties.

The procedure for assessing DEM quality involves examination of (i) ortho-images, to provide a qualitative check on stereomatching performance; (ii) DEM collection statistics which quantify the percentage of correctly matched pixels as a function of those interpolated; and (iii) height differences between check points, measured using independent field survey, and corresponding DEM points. The concepts of precision, accuracy and reliability are defined in the context of DEM quality assessment and methods are outlined which can be used to assess these variables. The assessment is conducted for two adjacent stereopairs with similar characteristics, considering the effects of both DEM collection parameters and different lens models upon DEM quality.

Results show that digital photogrammetry, in conjunction with independent field survey, can be used successfully for extracting high resolution, small scale DEMs from natural gravel surfaces. Components (i) and (ii) of the quality assessment suggest the need to optimize DEM collection parameters, although the effects of not using a properly specified lens model were minimal at this scale. Method (iii) showed that increasing stereomatching success does not necessarily lead to more accurately estimated DEM points. However, the use of method (iii) remained difficult because of the scale of the photogrammetric application being used; check point positioning within the photogrammetric co-ordinate system was only possible to ± 10 *mm which, for a gravel bed surface, was associated with elevation variance of a similar, sometimes greater, magnitude. The next stage of this research will require the use of higher quality check data, possibly from laser profiling.*

Key words: DEM accuracy, gravel bed river, surface roughness, close range photogrammetry

INTRODUCTION

THE COLLECTION of topographic data for small terrain areas has traditionally been accomplished using field surveying methods, such as levelling, angular heighting, stadia tacheometry and electronic distance measurement (EDM). Although these methods are regarded as flexible and technically uncomplicated (Kirby, 1991), close range photogrammetry is seen to have distinct advantages over conventional survey methods mainly in terms of accuracy, speed of application and spatial coverage.

Thus far, most applications of close range photogrammetry (Kirby, 1991; Lane *et al.*, 1994) have made use of manual data collection methods, albeit using automated stereoplotters. With appropriate research design, these methods can provide data of appropriate quality and density but still require time consuming laboratory data collection. Developments in automated stereomatching methods clearly provide much potential for automating this process and preliminary close range applications are encouraging (Pyle *et al.*, 1997). However, these methods need careful assessment if they are to be applied to rough surfaces, where small changes in camera position and lighting intensity and direction may result in major changes in image quality, image perspective and, possibly, the quality of the DEMs that are acquired.

This paper develops a procedure for assessing DEM quality and applies it to DEMs of rough river gravel surfaces generated using close range digital photogrammetry. Recent research has suggested that better specification of surface topography is critical for the understanding of both the effects of roughness upon flow processes (Lane, 1998) and of roughness elements upon sediment entrainment and transport (Hassan and Reid, 1990; Clifford *et al.*, 1992). The work presented here forms part of a wider research programme which is concerned with developing close range digital photogrammetric methods to measure and quantify morphological change in gravel bed rivers and coupling this to the development of numerical modelling methods. In this paper, a field based application of the close range photogrammetric technique is used to acquire accurate bed topographic information from a natural gravel bed river. This application is to be supplemented at a later stage by a smaller scale laboratory based study. This will involve the use of similar equipment and methodologies to acquire high quality DEMs of a simulated gravel river bed. By simultaneously acquiring high resolution velocity information throughout the flow, attempts will be made to link bed roughness characteristics to turbulent flow processes which operate above and around roughness elements. Roughness statistics will be derived from the acquired DEMs to provide boundary condition information for the parameterization of two and three dimensional numerical models. The overall quality of the small scale DEMs is paramount as this will affect (i) the numerical values of the derived roughness statistics; and (ii) the predictions of the computational fluid dynamics code which is to be coupled with these DEMs.

This investigation is concerned with the quality of these small scale DEMs. In the field based application presented here, scanned images of exposed river gravel surfaces were stereomatched to acquire high resolution DEMs. These DEMs were assessed using qualitative interpretation of ortho-images, quantitative analysis of DEM collection statistics and quantitative comparison with independently acquired check points. The latter were used (i) to assess the effects of DEM collection parameters on DEM quality; (ii) to determine optimum collection parameters; and (iii) to assess the effects of different lens models on surface quality.

THE CONCEPTS OF PRECISION, ACCURACY AND RELIABILITY IN RELATION TO DEM QUALITY ASSESSMENT

The overall "quality" of the elevation data produced using stereomatching algorithms is of fundamental importance in this research, because this will directly affect any roughness statistics calculated from the DEMs. The quality of a DEM is a function of the accuracy, reliability and precision of the survey/photogrammetric measurements and the block bundle adjustment itself. In order to construct a thorough, systematic data quality assessment procedure, it is necessary to identify the different types of potential errors and their sources (Table I) and to quantify the probability of there being errors of a specified size and type.

The *precision* of survey and photogrammetric measurements is controlled by *random errors*, which are inherent in the measurement process. Repeated measurements of the same element yield different measured values and these inconsistencies cannot be removed by refining the functional model or by applying corrections. By regarding survey and photogrammetric measurements as random variables, a *stochastic model* can be used to assess random effects and obtain estimates of precision (Cooper and Cross, 1988). Local measures of precision require the variance of individual parameters to be computed from the covariance matrix of the parameter. Global measures of precision, such as the a posteriori variance of unit weight (the variance factor), can be used to quantify the effects of random errors in a complete set of co-ordinates or other derived quantities. The precision of automated image measurement and stereomatching is positively related to (i) the number of pixels that are associated with a target (Chandler and Padfield, 1996); and (ii) the signal-to-noise ratio (SNR) of the imagery (Vision International, 1995). Higher image SNRs also lead to fewer gross errors (false fixes) in the stereomatched DEMs.

Blunders or mistakes which occur during survey measurement, photogrammetric measurement or stereomatching may be referred to as *gross errors*; these can be considered to determine the *reliability* of the DEM (Table I). Cooper and Cross (1988) define reliability as “a measure of the ease with which outliers may be detected”. In their assessment of the quality of survey data sets, Cooper and Cross (1988) also make the distinction between *internal reliability* and *external reliability*. Internal reliability is seen as the size of the marginally detectable gross error in a measurement whereas external reliability is a measure of the effect of this error on the parameters (for example, co-ordinates) or on data computed from them. The tau factor (τ_i) can be used to quantify the internal reliability of a particular survey measurement and this is simply the ratio of the standard error of a measurement to that of the corresponding correction (Cooper and Cross, 1988). A method for evaluating the internal reliability of stereomatched DEMs based on (semi-) independent estimates of surface elevation is presented below. External reliability is also an important concept in this research because this will affect the measures of roughness derived from the stereomatched DEMs. In the context of experimental tests on DEM accuracy, Li (1991) has constructed a complementary definition of reliability, which is seen as “the degree of correctness to which the DEM accuracy figures (i.e. RMSE, standard deviation, mean) have been estimated”.

In the context of digital photogrammetry, Table I outlines some physical effects which result in the functional model (for example, the collinearity equations) being an inaccurate description of the relationship between derived survey quantities (for example, photocontrol co-ordinates) and the actual co-ordinates of points in the object space. In terms of DEM quality, these physical effects can result in *systematic errors* in the control point data and in the image measurements used to define the photo-co-ordinate system. Systematic errors determine the *accuracy* with which photogrammetric image measurements (and hence stereomatched DEM points) can be used to estimate the co-ordinates of a target in the object space. DEM accuracy is traditionally evaluated and summarized by computing the mean and standard deviation/r.m.s.e. of height discrepancies between the DEM and independently acquired check point elevations (Torlegård *et al.*, 1986).

TABLE I. Types and sources of error in conventional survey and close range digital photogrammetry.

Sources/causes of error	*Methods of assessment*	*Methods of correction*
RANDOM ERROR (PRECISION)		
(i) **W.r.t. survey measurements**		
Variation in the measurement process.	Repeat measurements of the same element. Covariance matrix used to compute local and global measures of precision.	Cannot be removed.
(ii) **W.r.t. digital photogrammetry**		
Low SNR.	Precision estimate calculated during stereomatching from the SNR and the negative curvature of the correlation curve.	Improve image quality.
GROSS ERROR (RELIABILITY)		
(i) **W.r.t. survey measurements**		
Incorrect measuring/recording procedures.	Assessment of individual measurement residuals obtained from least squares estimation.	Selective removal of gross errors.
Internal reliability is the marginally detectable gross error. *External reliability* is the effect of undetected error on derived co-ordinates.	Calculation of the tau factor (τ_i) and computation of the upper bound on the gross error that can be detected in the ith measurement with a given probability.	
(ii) **W.r.t. digital photogrammetry**		
Internal reliability. Incorrect fixes during stereomatching due to: poor image contrast/quality; poor triangulation; non-optimal DEM collection parameters; errors in ground point measurement and identification; effects of viewing geometry.	Analysis of DEM collection results and comparison of DEM points with independent elevation data to identify individual blunders. Overlapping DEMs (redundant data) can also be used to provide (semi-) independent estimates of surface elevation.	Blunder editing. Improve image contrast and quality; improve triangulation by reviewing functional models; optimize DEM collection parameter set for specific applications; re-measure GCPs.
External reliability. Effect of false fixes on DEM derived data (for example, roughness parameters).	Analysis of the effect that a marginally detectable error in the ith measurement has on a particular parameter.	
SYSTEMATIC ERROR (ACCURACY)		
(i) **W.r.t. survey measurements**		
Atmospheric and other physical effects. Instrument errors and incorrect pointing.	Extension of functional models to include systematic errors. Systematic errors which can be successfully estimated from the survey measurements can then be modelled.	Cannot be fully eliminated but review of functional models where necessary can be used to mitigate against their effects.
(ii) **W.r.t. digital photogrammetry**		
Insufficiently convergent imagery.	Ensure convergence angle of 20° to 40° when acquiring imagery.	
Over-simplification of the measurement process (for example, ignoring the effects of lens distortion when these should be considered). Physical effects include camera lens distortions, atmospheric refraction, film unflatness.	Independent accuracy assessment and assessment of the significance of additional (for example, interior orientation) parameters within the functional models.	

Modified from Cooper and Cross (1988).

METHODOLOGY

Image Data Acquisition and Automatic DEM Collection

Fieldwork was undertaken on a reach of the River Affric, Scotland (Ordnance Survey National Grid reference NH 128204), to test the field potential of close range photogrammetry for acquiring high resolution DEMs of natural, water-worked gravels. The Affric is a typical upland, coarse-grained, clearwater river within which a number of large gravel bars become exposed at low flow. Although not reported here, research has also been carried out using two media (through water) photogrammetry to acquire DEMs in submerged zones, and this application requires an accessible clearwater reach.

For the surface roughness characterization of sub-aerial river gravels, a large partially exposed gravel bar was selected for study which exhibited a pronounced downstream transition in its grainsize composition and hence surface topography. Three test sites (A, B and C) (Fig. 1(a)) were located on the bar, visually identified as having different surface roughness characteristics and distinct morphological features. These test sites were chosen such that the spatial variability of microtopography over the surface of the bar could also be examined.

Sites A, B and C were photographed using two semimetric, 80 mm Hasselblad ELX 500 cameras. These can be pinned at a desired focal setting and are fitted with 25 cross réseau plates. The cameras were gantry-mounted to produce a camera-to-object distance of approximately 2·2 m and a base:distance ratio of around 1 : 6. Each site was recorded using a small block of eight stereopairs, arranged in two parallel strips of four stereopairs each (Fig. 1(b)). The cameras were separated by 350 mm to allow approximately 60 per cent forward overlap and strips were aligned to produce approximately 20 per cent sidelap (Fig. 1(b)). The two exposure stations at the end of each strip were offset by 100 mm to acquire photographs of the same area from two slightly different positions. This produced stereopairs with approximately 90 per cent forward overlap and hence the potential to assess the reliability of the stereomatching algorithm and the effects of camera position and viewing geometry on the derived surface elevation values.

The results presented in this paper relate to DEMs extracted from two contrasting stereopairs taken from site B, within which the spatial variability of surface topography was found to be greatest (possibly due to the existence of a transverse cut-off channel which passes through the site) and for which a large number of checkpoints were available. Photocontrol targets were glued onto stones ensuring that at least four targets were visible in the overlap area of each stereopair, as well as in the overlaps between adjacent stereopairs. A Leica T1610 co-axial total station was used to establish the three dimensional co-ordinates of these targets and best estimates of their object space positions were obtained by a least squares adjustment of the survey data.

Stereomatching was performed using the Ortho*MAX* module of Erdas Imagine software mounted on a Sun workstation; this module carries out interior orientation, ground point measurement and, ultimately, automated DEM extraction using the Vision International stereomatching algorithm. The matching procedure is based on the normalized cross-correlation of orthorectified image patches, obviating the need for epipolar resampled stereo-imagery.

The Vision International algorithm relies upon small scale variations in brightness and contrast in each image. The bracketing of aperture stops and shutter speeds was thus used in an attempt to obtain optimum exposures with maximum contrast. The negatives were then scanned at 20 μm (Fig. 2) with 256 grey levels using the Helava DSW 100 scanning workstation at City University, London. Each pixel represented an area of approximately $0{\cdot}6 \times 0{\cdot}6$ mm in the object space.

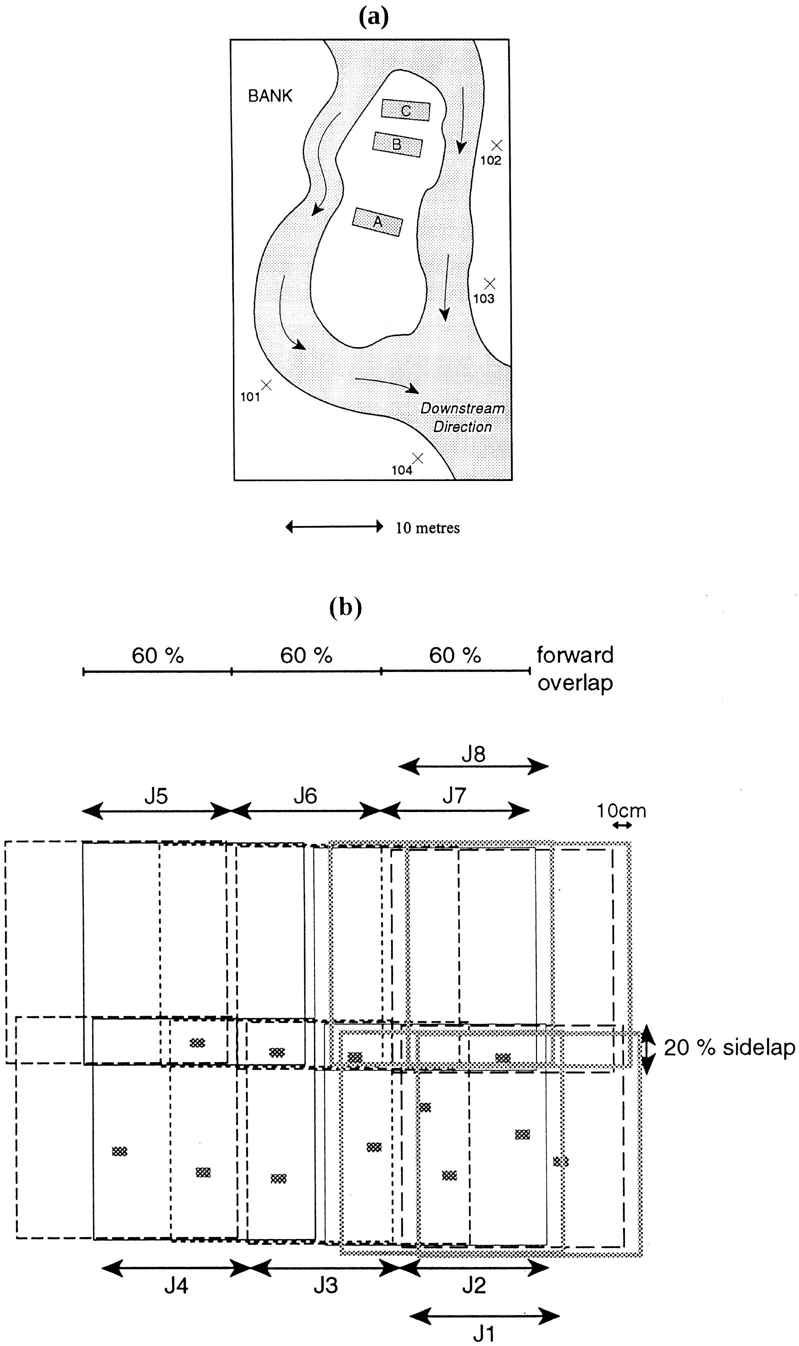

FIG. 1. (a) Schematic map of the study area; (b) location of stereopairs J1 to J8 acquired from site B.

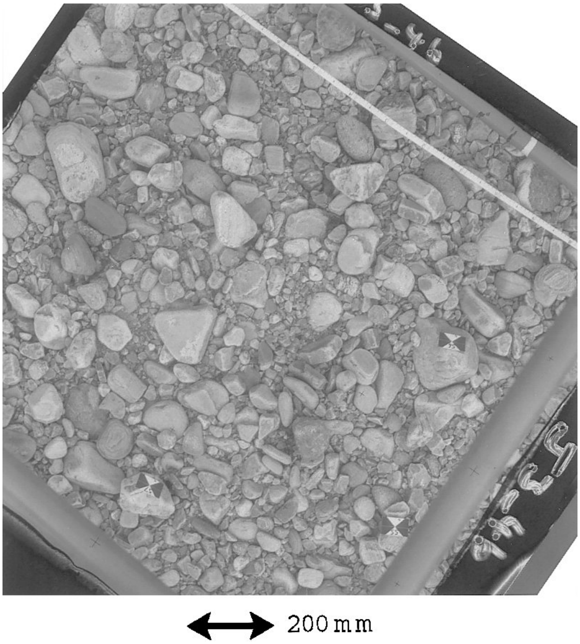

FIG. 2. Example of the scanned imagery used for stereomatching.

Radial lens distortion was modelled using a polynomial and the camera principal distances at the time of exposure were estimated to an acceptable level of precision using a self calibrating bundle adjustment which included additional significant lens parameters (Chandler *et al.*, 1989). The output from this bundle adjustment and also the remaining interior orientation parameters, photocontrol co-ordinates and ground point measurements were then submitted to the Ortho*MAX* block bundle adjustment. DEMs were generated from the block once an acceptable photogrammetric solution had been achieved.

Variation of DEM Collection Parameters

DEMs were initially collected using the default DEM collection parameters. These parameters control various aspects of the correlation process and may be altered by the user (Table II). The determination of optimized parameters is explained subsequently.

For successful correlations, a user-defined set of signal-to-noise ranges is used to "rank" the matched points into three categories: "good"; "fair"; and "poor". The thresholds for each category are defined on the basis of the minimum precision (MP) parameter value. For example, if MP is set to 0·5 pixels, this defines the upper bound of the "poor" category, such that the estimated precision of matches falling into this category cannot exceed 0·5 pixels. Although increasing MP has the effect of increasing the number of "successfully" matched points, many of these matched points will be inaccurate. For this analysis, the default MP value of 0·5 was used throughout, which sets thresholds (in pixels) for the "good" and "fair" categories at 0·17 and 0·33

TABLE II. Ortho*MAX* DEM collection (strategy) parameters.

Collection parameter	*Description of parameter*	*Default*	*Optimized*
Minimum threshold	Minimum acceptable correlation coefficient used to accept a point	0·6	0·6
Noise threshold	Minimum acceptable correlation coefficient used to consider a point	0·4	0·4
Maximum parallax	Maximum search range (in pixels of x parallax) around a point	5	5
Minimum template size	The smallest (initial) template size (in pixels) used by the area correlator	7	9
Maximum template size	The largest (final) template size (in pixels) used by the area correlator	9	11
Minimum precision	Minimum acceptable estimated precision (in pixels) of a point passing the minimum threshold test	0·5	0·5
Rejection factor	Smoothing factor for rejecting spikes and pits during post-processing	1·5	2
Skip factor	(Pixel) spacing at which DEM points are collected for all but the end RRDS	2	2
Edge factor	Controls non-unique correlations which lead to false fixes	2·5	3
Start RRDS	Starting image resolution to be used	4	4
End RRDS	Final (highest) image resolution	0	0
y parallax allowance	Allows the correlator to move around in cases for which the triangulation results are poor	0	0
Resampling	Use bilinear interpolation (versus nearest neighbour) during orthorectification of patches	Bilinear	Nearest neighbour
Post processing	To govern whether post-processing (including blunder editing) is performed after each RRDS	Yes	Yes
Ignore interpolated	(TIN collections only) used to ignore interpolated points when outputting elevations	—	—
Band selection	(Multiband images only) to select which band is to be used as the source for correlation	—	—

Source: Vision International (1995).

respectively. Those pixels which cannot be matched with a precision less than or equal to the value of MP are interpolated, using an *n*th order polynomial based on the elevations and radial distances of neighbouring points. Thus, a point that is "good" has the highest probability of being correctly matched.

The DEM collection parameters were then systematically varied, using a range of pre-specified values, and the DEMs re-collected. This repetition was carried out to assess the performance of the stereomatching algorithm when using small format imagery containing complex small scale roughness features and to examine the effects that extreme parameter values might have upon the DEM collection results. Because Ortho*MAX* has been designed for use with large format aerial photography and satellite imagery, the ranges over which the parameters were tested exceeded the manufacturer's suggested ranges.

In order to reduce processing time, three small area subsets of the full area DEMs were subjected to systematic parameter change. Two of the areas had different surface textures and the third area included a percentage of peripheral pixels so that stereomatching success in peripheral areas could be analysed. The results of these collections were then used to define an "optimum" parameter set for each stereopair which (i) maximized the number of "good" matches; and (ii) minimized the number of interpolated points. Due to the nature of this close range application, which aims to conserve small scale roughness features of the terrain, it was felt that the maximization of good matches was appropriate. The smoothing effect of interpolation would be more appropriate for modelling more homogeneous terrain areas, such as parcels of arable land found in large scale aerial photography. Once optimized parameter sets had been defined, the full area DEM was re-collected using these parameters. Small differences in image texture, quality and contrast between the sub-areas resulted in only slightly different optimum parameter sets for each stereopair and so a single optimum parameter set was finally selected for the full area re-collection.

Use of Alternative Lens Models

Uncertainties exist in the functional models used by Ortho*MAX* to relate the camera interior orientation parameters, estimated photocontrol co-ordinates and consequent object space point estimates. The augmentation of these functional models to include the output from the self calibrating bundle adjustment (that is, estimates of the principal distances) was done in an attempt to reduce any systematic errors associated with the use of semimetric photography and hence improve the accuracy of the DEM.

Although fully metric photography has traditionally been recommended for photogrammetric measurement, the use of semimetric imagery has been shown to provide a cost effective alternative which allows point co-ordinates to be estimated with high levels of precision and accuracy (Chandler and Padfield, 1996). In order to assess the effects of small, systematic errors associated with the use of semimetric photography on overall DEM quality, the second phase of the analysis involved the use of alternative lens models. Camera parameters (principal distance and radial distortion model) were altered and the Ortho*MAX* bundle adjustment repeated. The DEM was then re-collected using the optimized DEM collection parameters and a principal distance value (i) estimated using the self calibrating bundle adjustment; and (ii) set at a previously calibrated value. By comparing the DEMs from (i) and (ii) with the check point elevations, any reduction in DEM accuracy resulting from the incursion of additional systematic error would be observed.

As Table III shows, *Lens Model 1* was obtained from the separate self calibration of the left and right Hasselblad cameras at close range. The use of *Lens Model 2* represents the degradation of the lens model with calibration data from the left camera only used to define both the left and right cameras and the principal distance set at a previously calibrated value. *Lens Model 1* is thus the more realistic because (i) the left and right cameras are calibrated separately; and (ii) calibration data for *Lens Model 1* were obtained from a self calibrating bundle adjustment based on photogrammetric measurements made from the close range imagery itself.

Assessment of the Precision of Survey and Photogrammetric Measurements

The cofactor matrix produced as a by-product of the least squares adjustment of survey observations provided a measure of the precision with which the co-ordinates of ground control points could be estimated. During ground point measurement, the automatic matching of corresponding tie and control points in the left and right images provided estimates of the precision (in pixel units) with which individual points would be stereomatched in the image. This information gave an appreciation of the level of precision which could be expected in DEM elevation.

A standard least squares bundle adjustment is also used within Ortho*MAX* to compute the parameters of exterior orientation. Before each DEM collection, analysis of the “triangulation results” was used to assess estimated parameters and their associated precisions. In this case, the a posteriori variance of unit weight (or variance factor) is a measure of the conformity of frame parameter and point residuals and their associated a priori standard deviations. This provides a global measure of the

TABLE III. Alternative lens models used in the least squares bundle adjustment.

Lens Model	*Used to define*	X_0 *(μm)*	Y_0 *(μm)*	*C (mm)*	K_1	K_2	K_3
Lens Model 1	Right camera	− 3·6	335·7	90·5795	7·358e-9	1·301e-11	-3·245e-21
Lens Model 1	Left camera	185·3	251·7	91·6721	− 3·911e-9	9·261e-12	4·681e-22
Lens Model 2	Left/right camera	185·3	251·7	81·0209	2·727e-10	1·150e-11	7·671e-22

X_0 and Y_0 are the co-ordinates of the principal point in the fiducial co-ordinate system; C is the principal distance of the taking camera; and K_1, K_2, and K_3 are radial distortion parameters.

precision of the photogrammetric solution. Attempts were made to neither under constrain nor over constrain the photogrammetric solution so that this value approximated to unity, signifying that parameter residuals are consistent with realistically specified estimated precisions.

Assessment of DEM Reliability

As Table I shows, gross errors which occur in the photogrammetric measurement of control points and tie points in the image space and object space are identified by analysing the residuals at the block bundle adjustment stage. The variation of the collection parameters was done in an attempt to improve the internal reliability of the automatic DEM collection by reducing the probability of incorrect fixes. However, because "false fixes" can occur for "good" points, independent accuracy assessment was also required to (i) identify and edit stereomatching failures; and (ii) examine the effect that varying DEM collection parameters has upon DEM accuracy.

The DEMs collected from the overlapping stereopairs, which were acquired by offsetting the exposure station by 100 mm, were used to provide (semi-) independent estimates of elevation. By comparing field check point elevations with corresponding points in both DEMs, the internal reliability of stereomatching was evaluated by analysing the height discrepancies between the two DEMs.

Check Point Acquisition for DEM Accuracy Assessment

Independent ground based measurements were used to assess DEM accuracy. The field work involved using a level and staff to acquire a set of check points from each test site. A steel band was laid across the gravel surface to mark out a profile which was approximately 2 m in length. In order to locate the profile in the object space, measurements were made to the start and end points of the tape using the total station. A series of elevations adjacent to the tape was then obtained using the level and staff. The spacing of the height measurements along the profile was varied to reflect the roughness of the terrain such that more points were collected if the tape happened to rest over clasts with greater elevation variability.

RESULTS

Automatic DEM Collections

Fig. 3 shows a shaded relief visualization of a portion of the extracted DEM collected at 3 mm resolution. This portion of the DEM was produced from the stereomatching of stereopair J4 (see Table IV). The lighter areas represent greater elevation values.

The DEM contains most of the dominant small scale terrain features present in the imagery and larger clasts can easily be identified. The stereomatching algorithm

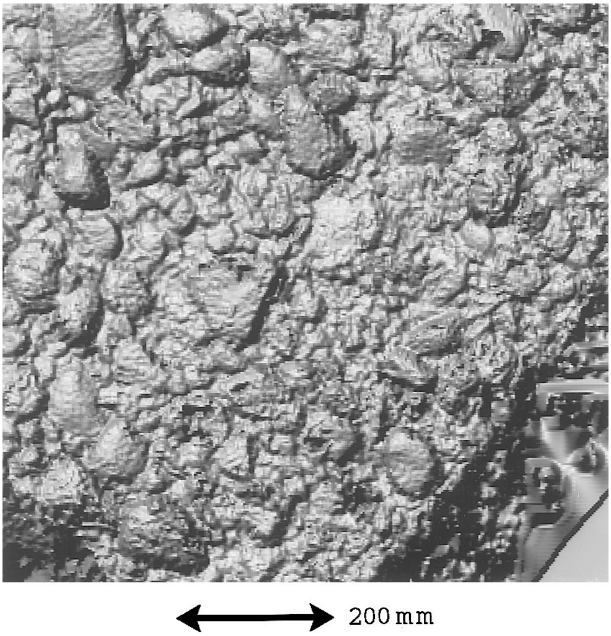

FIG. 3. Planimetric view of a portion of the full area DEM. The region shown here corresponds to the overlap area of stereopair J4. This DEM was collected using the most realistic lens model (*Lens Model 1*) and optimized collection parameters.

TABLE IV. Elevation statistics for the DEMs collected using default parameters.

Sub-area DEM	*Minimum (m)*	*Maximum (m)*	*Mean (m)*	*Standard deviation (m)*	*Range (m)*
J2	98·629	99·169	99·084	0·024	0·540
J4	94·947	99·145	98·962	0·170	4·198
J2(sub)	99·043	99·147	99·086	0·015	0·104
J4(sub)	98·925	99·062	98·990	0·021	0·137

has failed in the darker regions located toward the bottom left and bottom right of Fig. 3, which contain abnormally low elevation values relative to their surroundings. It is possible that excessive obliquity due to perspective geometry may be responsible for this failure.

Table IV shows elevation statistics for stereopairs J2 and J4 which were used in the production of the full area DEM. Elevations are recorded in metres above an arbitrary datum. Inspection of Table IV reveals the abnormally low elevation values which are found in the regions of matching failure. Because the total height difference found on the field surveyed profile was approximately 110 mm, the minimum elevation value collected from J4 is obviously erroneous. Table IV also includes elevation statistics for the small area, subsetted DEMs also collected from stereopairs J2 and J4 (J2(sub) and J4(sub)). These DEMs were found not to contain regions of significant matching failure and their elevation ranges are seen to compare more realistically to that observed in the field. The *F*-statistic produced from a one way analysis of variance test revealed that the standard deviation of elevation values in

TABLE V. Effect of parameter optimization on collection results for DEM J4.

	Collection percentages by status	
Status categories	*Default*	*Optimized*
Good	19·1	42·8
Fair	30·8	21·7
Poor	22·2	13·3
Interpolated	26·9	20·9
Off-Image	0	1·3

The DEM was collected using the most realistic lens model.

J4(sub) was significantly greater than that of J2(sub). This result would suggest that the former DEM is probably statistically “rougher” than the latter. However, this inference cannot necessarily be drawn for the gravel surfaces themselves, because the spread of elevation values is likely to be influenced by gross DEM errors such as anomalous spikes and false fixes. “Blunder editing” is thus required prior to the derivation of roughness statistics from the extracted DEMs.

Effects of DEM Collection Parameters on DEM Collection Results

Table II shows the default and optimized DEM collection parameter sets, the latter of which was defined by choosing the value for each parameter which maximized the percentage of matched points versus the percentage of interpolated points.

Variation of individual collection parameters revealed that the percentage of successful matches could be increased by enlarging the maximum and minimum template sizes used for cross-correlation. Contrary to expectations, variation of the minimum threshold (MT) and noise threshold (NT) parameters could not be used to increase the percentage of “good” matches without reducing the proportion of matching versus interpolation. The default values of these parameters were thus maintained. Increases in the rejection factor (RF) and edge factor (EF) were also found to improve matching results, as did the selection of a nearest neighbour resampling strategy rather than the default bilinear method.

Table V shows the change in the collection percentages by status which resulted from collecting the DEM using (i) the default parameter set; and (ii) the optimized parameter set. The most realistic lens model was used in both cases. These results are for the DEM extracted from stereopair J4. The use of the optimized collection parameter set is seen to improve the overall quality of the DEM in terms of the significant increase in the percentage of “good” matches, and the reduction in the percentage of “fair”, “poor” and “interpolated” points which are observed.

Effects of Alternative Lens Models on DEM Collection Results

Table VI shows that the degradation of the lens model does not appear to significantly effect the collection percentages by status.

Qualitative Assessment of DEM Quality

By removing the effects of image perspective and relief, the production of ortho-images from the stereomatched DEMs and the source imagery provides a means of qualitatively assessing DEM quality (Pyle *et al.*, 1997). Poorly matched

TABLE VI. Effect of lens model degradation on collection results for DEM J4.

	Collection percentages by status	
Status categories	*Realistic lens model*	*Degraded lens model*
Good	42·8	42·7
Fair	21·7	21·2
Poor	13·3	13·1
Interpolated	20·9	21·7
Off-Image	1·3	1·2

In both cases, the DEM was collected using optimized parameters.

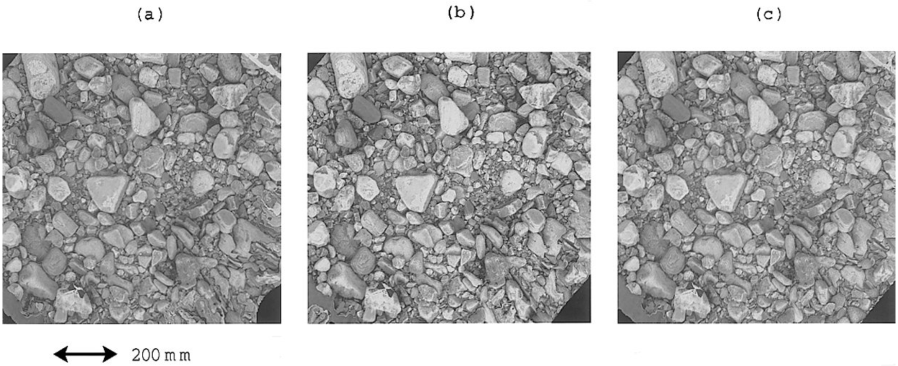

FIG. 4. Visual identification of DEM quality using ortho-images produced from DEMs collected (a) using default collection parameters and most realistic lens model; (b) using optimized collection parameters and most realistic lens model; and (c) using optimized collection parameters and degraded lens model.

areas would be associated with over distortion of the ortho-images (Fig. 4) as compared with the source image (Fig. 2). Fig. 4 suggests that the DEM gives a reasonable representation of microtopography for most of the area but also shows areas where stereomatching has clearly failed (Fig. 4(a)). The straightness of the tape measure in the top right hand corner confirms successful matching, although localized warping and deformation may again be the result of obliquity or radial lens distortion.

Qualitative examination of the ortho-images produced from the DEM which was collected using optimized parameters (Fig. 4(b)) suggests some improvement in quality with a reduction in number and extent of deformations. Contrary to expectations, the ortho-image produced using the DEM collected using optimized parameters and the degraded lens model (Fig. 4(c)) records some improvement in matching success as compared with Fig. 4(b).

Although the use of the optimized DEM collection parameter set resulted in (i) an ortho-image which showed an improvement in the overall quality and successfully matched area of the DEM (Fig. 4(b)); and (ii) an increase in the number of "good" matches, comparison of the re-collected elevations with the check point heights is necessary to establish whether or not these effects are accompanied by an overall improvement in DEM accuracy.

Quantitative Assessment of DEM Quality

(i) Estimation of precision. The precision of the control survey and the image matching process can be used to provide quantitative measures of resultant DEM

quality. Although very high levels of precision can be achieved using image matching techniques, in close range photogrammetric applications the absolute precision of the DEM is constrained by the relative precision of the control survey. The least squares bundle adjustment of the survey data suggested that photocontrol target co-ordinates could be estimated to within ± 1 mm in X and Y and $\pm 0{\cdot}5$ mm in Z. Because image measurement using Ortho*MAX* showed that the precision of image matching was likely to be around 0·1 pixel (equivalent to 0·06 mm in the object space), the precision of the control survey served as the primary constraint on the precision of the final DEM.

(ii) Estimation of DEM internal reliability. Fig. 5 shows the results of comparing corresponding points in the two DEMs extracted from overlapping stereopairs J1 and J2. The deviation of points away from the line of equality indicates differences in elevation between the DEMs at each field check point location. Because these stereopairs cover the same area but are taken from different camera positions, these differences can be used to analyse the effects of camera position on stereomatching and hence the internal reliability of the DEMs.

The average value of the unsigned differences between corresponding DEM points is not significantly different from zero at the 95 per cent level. Because there is, on average, no significant difference in the elevation values between the two DEMs, this sample of points suggests that the effects of camera position are negligible. Points shown to be outliers would, however, be removed from the DEM before further processing to ensure the maximization of external reliability.

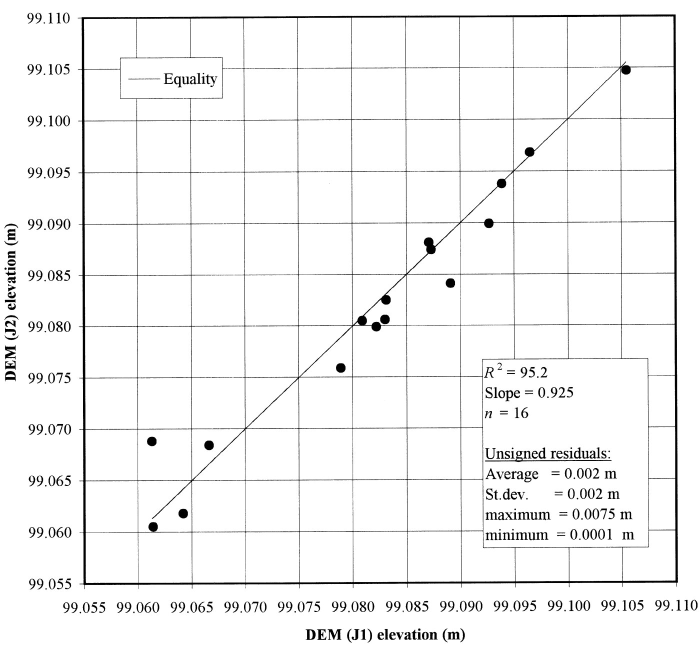

FIG. 5. Analysis of the reliability of automated DEM extraction.

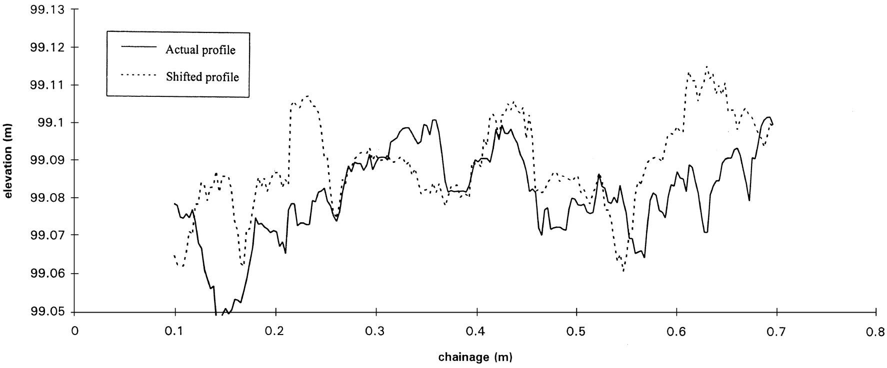

FIG. 6. DEM profiles showing the effect of shifting the location of the transect by 10 mm.

(iii) Estimation of the accuracy of restitution and automatic DEM derivation. The a posteriori variance of unit weight associated with the Ortho*MAX* block bundle adjustment was used to evaluate the effects of using alternative lens models on the global accuracy of the photogrammetric solution. A variance factor of 0·989 resulted from the use of the most realistic *Lens Model 1*. Degradation of the lens model produced a value of 1·053 which is not significantly different from the previous value at the 95 per cent significance level. This suggests that lens model degradation was not accompanied by a reduction in the global accuracy of restitution.

As Fig. 6 shows, it was found that a 10 mm deviation of the profile taken across the stereomatched DEMs from the actual location of the profile on the ground would lead to significant error in the results of the comparison. In order to compensate for check point positioning errors, the elevation range within a 10 mm search radius of the nearest DEM point was used for comparison with each check point.

Fig. 7 presents the results of the comparison between check point elevations obtained from the field profile and the nearest corresponding points in the DEM. In order to compensate for possible shifts in the positioning of the field profile, the "error bands" at each point reflect the range of elevations found within a 10 mm radius of the estimated location of the profile. The percentage of "hits" (that is, points having error bands which cross the line of perfect correspondence) can thus be used to assess quantitatively the overall DEM accuracy. The ortho-images were used at this stage of the analysis to identify check points which were located within areas of the DEM where the stereomatching algorithm had failed. These check points were then discarded from the remainder of the analysis.

Inspection of Fig. 7 shows the slight improvement in the accuracy of individual DEM points which is achieved by optimizing collection parameters, in terms of increasing the proportion of "hits". As Fig. 7(a) shows, the majority (74 per cent) of error bands cross the line of perfect correspondence and a significant number of the DEM points themselves fall very close to this line. A range of error band widths is present, the largest band widths being frequently associated with points which lie on the boundaries of individual clasts. Points which do not cross the line of perfect correspondence ("misses") may be due to false fixes or poorly interpolated points.

Fig. 7(b) shows a very similar pattern of correspondence between check point

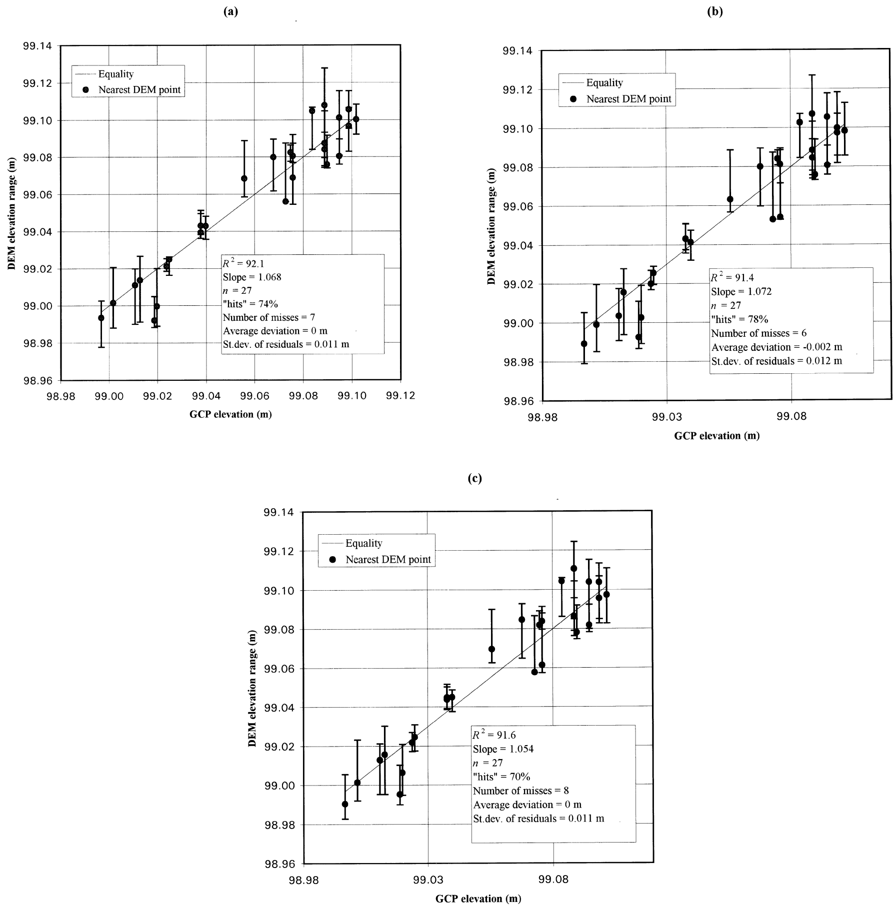

FIG. 7. Elevation plots showing ground control point (GCP) elevation versus the nearest corresponding DEM point for (a) default collection parameters; (b) optimized collection parameters; and (c) optimized collection parameters and the degraded lens model. Error bands show the range of elevations which is found within a 10 mm radius of the estimated location of each check point.

and DEM elevations. Although the positions and widths of individual error bands have not changed significantly, the optimization process resulted in one additional hit. A number of DEM points have been more accurately estimated, whereas others are either unchanged or have migrated slightly further away from the line of perfect correspondence. Use of the degraded lens model (Fig. 7(c)) resulted in a reduced proportion of hits and an increased divergence of a number of DEM points further from the line of equality, indicating that the accuracy of a number of points was reduced. The standard deviation of height residuals does not, however, suggest that the degradation of the lens model has significantly affected the overall accuracy of the DEM.

The statistical relationship between surveyed and corresponding DEM elevations was assessed by carrying out a regression analysis. The strength of this relationship for both the default and optimized DEMs is indicated by very high R^2 values which

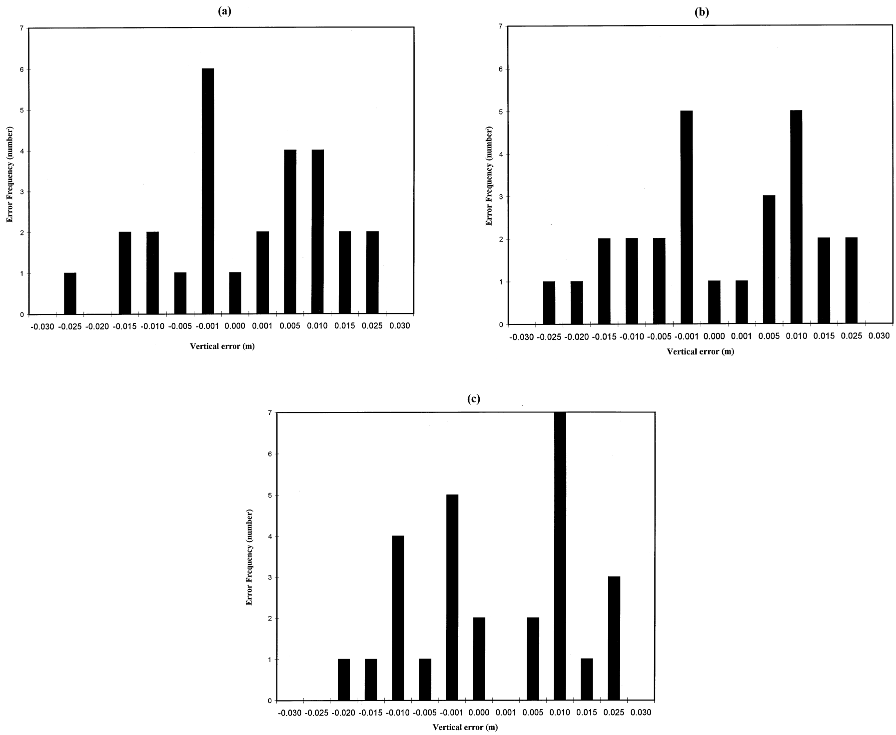

FIG. 8. Histograms of DEM errors: (a) default DEM; (b) optimized DEM; (c) optimized DEM using degraded lens model.

represent the proportion of the variance in DEM elevation that can be attributed to its linear regression on check point elevation. Further statistical analysis revealed that the slope of the regression line was not significantly different from 1. However, nearly identical values for the standard deviation of height residuals (Fig. 6) suggests that collection parameter optimization has not been accompanied by an overall improvement in DEM accuracy. This conclusion is also supported by the R^2 terms which do not suggest any significant improvements after optimization.

Histogram plots reveal that the majority (approximately 75 per cent) of height residuals are between ± 10 mm (Figs. 8 (a), (b) and (c)) and, of these, around 30 per cent are less than ± 1 mm in size. Collection parameter optimization has little effect on this pattern, although lens model degradation is accompanied by an increase in the number of larger residuals especially in the ($\pm$) 5 mm to 10 mm class.

DISCUSSION

The scanned imagery of rough gravel surfaces used in this study contained excellent fine texture which, in general, facilitated successful stereomatching and automated DEM production using Ortho*MAX*. The qualitative examination of orthoimages provided an effective means for identifying regions of stereomatching failure, although the precise causes of this failure remain unclear, other than the obvious importance of edge effects, which seem to be particularly dependent upon certain DEM collection parameters. Further research is required to assess if optimization may be generalized for a particular (in this case gravel) surface, or whether optimization is needed for each stereopair analysed.

The occurrence of false fixes which may be classified as “good” undermines the extent to which DEM collection statistics can be used to evaluate DEM quality and makes independent accuracy assessment critical. Sensitivity analysis is required to quantify the effects of individual collection parameters on DEM accuracy and the derivation of an alternative optimized set of parameters based on this strategy will be carried out in subsequent stages of the research. An automatic method for detecting and correcting false fixes is also required to improve the quality of the DEMs, which will ensure that roughness statistics can be calculated from them with high levels of external reliability.

The quality of the DEM was quantified on the basis of the precision, accuracy and reliability of the derived survey quantities and digital photogrammetric measurements. Despite the uncertainties inherent in the measurement process, it was possible to determine the spatial co-ordinates of photo-control targets with sub-millimetre precision. The precision of automatic image matching was 0·06 mm in the object space and the internal reliability analysis of the overlapping DEMs showed that the effects of camera position on stereomatching were negligible.

Although the quantitative tests of DEM accuracy suggested that the use of a degraded lens model was accompanied by a slight decrease in the accuracy of individual DEM points, this was not supported by the qualitative interpretation of the ortho-images. The effect of not accounting for minor systematic inaccuracies during camera self calibration may explain this anomaly and further quantitative methods are required to analyse the impact of a change in the functional model on the quality of the DEM.

In this application, uncertainties in the positioning of check points in the field made independent accuracy assessment difficult. Further, the number, density and quality of profile-aligned check points may not have been sufficient to pick up the smaller variations in DEM values which occurred as a result of altering collection parameters and degrading the lens model. In addition, the placement of the surveyed profile through the central portions of the imagery may also have failed to pick up changes in elevation occurring towards the DEM edges, which the ortho-images suggest seem to be more dependent upon optimization. The analysis of height differences between field check points and corresponding DEM points suggested a DEM accuracy of approximately ± 10 mm, which would be expected from the planimetric accuracy of the field check points.

An understanding of the relationships which exist between viewing angle, lighting direction and intensity and the resultant image quality/content is also required to enable the selection of collection parameters which are more likely to represent the terrain better. The smoothing effect which results from enlarging the template size and the effect of interpolation may also lead to the propagation of errors during automated DEM collection and the next stages of this research will endeavour to quantify these errors using refined independent accuracy assessment procedures.

Despite the need for a greater number of higher quality check points, the analysis showed that an improvement in stereomatching success, as measured by an increased number of “good” matches, is not necessarily accompanied by an improvement in the overall accuracy of the DEM. This is most likely due to the distribution of check points, which rarely covered peripheral areas that were significantly improved by optimization. A more thorough assessment of the effects of DEM collection parameters, lens models, image quality/contrast, shadowing and grid resolution on DEM accuracy will require the use of a dense grid of check points which can be reliably positioned and co-ordinated with the DEM. Hence, it is anticipated that subsequent phases of the research will involve the use of a high resolution laser profiling device, in a laboratory flume. Because the device can be used to acquire point elevations at 0·5 mm spacing over small (about 300×300 mm) areas, a matrix of elevations would be provided which could be compared to photogrammetrically derived elevations, in a common datum.

CONCLUSIONS

The aim of this research has been to develop and apply a non-contact method for acquiring high quality topographic information from complex, natural gravel surfaces for use in the roughness parameterization of hydraulic models. Although the results presented here are site specific, in that they are based on scanned photography acquired from a single exposed gravel bar, the close range photogrammetric system can equally be applied to any complex topographic surface for which sufficient ground truth information is available. Developments in two media (through water) photogrammetry are also being used to enable the comparison of subaqueous and subaerial gravel bed sedimentologies in natural and simulated channels.

The close range photogrammetric system used here comprises specific “off the shelf” hardware, software and semimetric cameras, to acquire and process small format imagery of complex gravel surfaces. Small format, semimetric scanned photography has been shown to be a viable, low cost primary data source for precise and accurate sub-pixel image measurement using digital photogrammetry (Chandler and Padfield, 1996). However, image quality (and hence subsequent stereomatching success) is dictated by a wide variety of factors such as viewing conditions, image contrast, contribution of graininess in the image, as well as the photographic resolving power of specific camera/lens configurations (Thomson, 1988). The quality of DEMs extracted from imagery acquired using different cameras under variable lighting conditions is thus likely to be variable and this research illustrates how, with proprietary software, close attention should be given to default parameter values.

The procedure for assessing the quality of the derived DEMs, based on ortho-images and height differences between check points and corresponding DEM points, could be applied to aerial photography and stereoscopic satellite imagery. When attempting to assess the overall quality of DEMs acquired from imagery at a variety of spatial scales, comparable methods for assessing precision, accuracy and reliability must be found for each application. By using conventional survey to acquire check points, precision estimates for the control survey can easily be obtained using least squares adjustment procedures. For this application, it was felt that there are fewer uncertainties in the functional models used to relate survey measurements and consequent co-ordinate estimates than is the case for the digital photogrammetry. Further research is thus required to test the comparability of results obtained from using alternative hardware and software configurations to acquire DEMs from specific stereopairs.

ACKNOWLEDGEMENTS

J. B. Butler is in receipt of NERC studentship GT4/95/55/F and this research was supported by the Royal Society. R. Littleworth and S. Robson*, City University, London, provided access to and help with image scanning which is gratefully acknowledged.

*Now at University College London.

REFERENCES

CHANDLER, J. H. and PADFIELD, C. J., 1996. Automated digital photogrammetry on a shoestring. *Photogrammetric Record*, 15(88): 545–559.

CHANDLER, J. H., COOPER, M. A. R. and ROBSON, S., 1989. Analytical aspects of small format surveys using oblique aerial photographs. *Journal of Photographic Science*, 37(7): 235–240.

CLIFFORD, N. J., ROBERT, A. and RICHARDS, K. S., 1992. Estimation of flow resistance in gravel-bedded rivers: a physical explanation of the multiplier of roughness length. *Earth Surface Processes and Landforms*, 17(2): 111–126.

COOPER, M. A. R. and CROSS, P.A., 1988. Statistical concepts and their application in photogrammetry and surveying. *Photogrammetric Record*, 12(71): 637–663.

HASSAN, M. A. and REID, I., 1990. The influence of microform bed roughness elements on flow and sediment transport in gravel bed rivers. *Earth Surface Processes and Landforms*, 15(8): 739–750.

KIRBY, R. P., 1991. Measurement of surface roughness in desert terrain by close range photogrammetry. *Photogrammetric Record*, 13(78): 855–875.

LANE, S. N., 1998. Hydraulic modelling in hydrology and geomorphology. A review of high resolution approaches. *Hydrological Processes*, 11: 1131–1150.

LANE, S. N., CHANDLER, J. H. and RICHARDS, K. S., 1994. Developments in monitoring and modelling small scale river bed topography. *Earth Surface Processes and Landforms*, 19(4): 349–368.

LI, Z., 1991. Effects of check points on the reliability of DTM accuracy estimates obtained from experimental tests. *Photogrammetric Engineering & Remote Sensing*, 57(10): 1333–1340.

PYLE, C. J., RICHARDS, K. S. and CHANDLER, J. H., 1997. Digital photogrammetric monitoring of river bank erosion. *Photogrammetric Record*, 15(89): 753–764.

THOMSON, G. H., 1988. Evaluating image quality of small format aerial photography systems. *Ibid.*, 12(71): 595–603.

TORLEGÅRD, K., ÖSTMAN, A. and LINDGREN, R., 1986. A comparative test of photogrammetrically sampled digital elevation models. *Photogrammetria*, 41(1): 1–16.

VISION INTERNATIONAL, 1995. *IMAGINE Orthomax user's guide*. Unpublished document. 121 pages.

Résumé

On présente dans cet article une méthode d'evaluation de la qualité d'un modèle numérique des altitudes (MNA) que l'on a appliquée en sortie d'un algorithme d'appariement stéréoscopique basé sur une corrélation croisée normalisée. On a utilisé des photographies semi-métriques des surfaces naturelles du lit d'une rivière recouvertes de gravier, prises sur le terrain, desquelles on a tiré automatiquement des MNA par photogrammétrie numérique pour déterminer les caractéristiques de rugosité de surface.

Pour évaluer la qualité des MNA, on utilise une méthode comportant trois phases d'examen: (i) examen des ortho-images pour avoir un contrôle qualitatif des résultats de l'appariement; (ii) examen des statistiques de saisie du MNA pour quantifier le pourcentage de pixels correctement appariés par rapport à ceux qui sont interpolés; et (iii) examen des écarts aux points de vérification, entre les altitudes déterminés sur le terrain par un levé indépendant et celles correspondantes du MNA. Dans le cadre de cette évaluation de la qualité du MNA, on définit les notions de précision, d'exactitude et de fiabilité et l'on donne un aperçu des méthodes utilisables pour déterminer ces paramètres. On a mené l'évaluation sur deux couples adjacents de caractéristiques semblables, en mettant en évidence à la fois les effets des paramètres de saisie du MNA et ceux de diverses sortes d'objectifs sur la qualité du MNA.

Les résultats montrent que l'on peut utiliser avec succès la photogrammétrie numérique, associée à un levé indépendant sur le terrain, pour extraire des surfaces naturelles de graviers, des MNA à petite échelle et à haute résolution. Les phases (i) et (ii) précédentes suggèrent qu'il faudrait optimiser les paramètres de saisie du MNA, même si l'on peut constater l'influence minime à cette échelle d'un type d'objectif aux caractéristiques non convenablement déterminées. La phase (iii) précédente *a montré qu'en accroissant la réussite de l'appariement stéréoscopique on n'aboutissait pas obligatoirement à une meilleure exactitude sur les points du MNA. En fait l'examen lors de la phase (iii) est resté délicat à cause de l'échelle utilisée dans cette application photogrammétrique; on n'a pu obtenir qu'une précision de ±10 mm sur les déterminations photogrammétriques des coordonnées des points de vérification, ce qui, pour la surface du lit en graviers, correspondait à peu près à l'ordre de grandeur de la variance des altitudes, si ce n'est pas davantage. On utilisera, dans la poursuite de cette recherche, des données de vérification d'une qualité supérieure, probablement à l'aide d'un profilage au laser.*

Zusammenfassung

Im Artikel wird ein Verfahren zur Qualitätsbewertung eines digitalen Höhenmodells (DHM) dargestellt, das auf die Ergebnisse eines Stereokorrelationsalgorithmus auf der Grundlage normalisierter Kreuzkorrelation angewandt wurde. Unter Verwendung semimetrischer Photos natürlicher Kiesbettoberflächen von Flüssen im Gelände wurde die Digitalphotogrammetrie zur automatischen Erzeugung von DHM angewandt, um die Oberflächenrauheit zu beschreiben.

Das Verfahren zur Bewertung der DHM-Qualität umfaßt die Prüfung 1. von Orthobildern, um einen qualitativen Test zur Leistungsfähigkeit der Stereokorrelation zu erhalten, 2. der Statistik der DHM-Gewinnung, die den Prozentsatz korrekt korrelierter Pixel als Funktion der interpolierten quantifiziert und 3. der Höhendifferenzen zwischen Kontrollpunkten, die aus unabhängigen Geländemessungen und entsprechenden DHM-Punkten ermittelt wurden. Die Begriffe Präzision, Genauigkeit und Zuverlässigkeit werden im Kontext der DHM-Qualitätsbewertung definiert, und es werden Methoden skizziert, die zur Bewertung dieser Variablen verwendet werden können. Die Analyse wird für 2 benachbarte Bildpaare mit ähnlichen Charakteristika, durchgeführt, wobei sowohl der Einfluß der DHM-Erfassungsparameter als auch unterschiedlicher Objektivmodelle auf die DHM-Güte betrachtet werden.

Die Ergebnisse zeigen, daß die Digitalphotogrammetrie in Verbindung mit unabhängigen Geländemessungen erfolgreich zur Erzeugung kleinmaßstäbiger DHM mit hoher Auflösung von natürlichen Kiesoberflächen genutzt werden kann. Die Komponenten 1. und 2. der Qualitätsbewertung zeigen die Notwendigkeit, die Erfassungsparameter für DHM zu optimieren, obwohl die Einflüsse von nicht passend gewählten Objektivmodellen bei diesem Maßstab minimal waren. Die Methode 3. zeigte, daß ein erhöhter Stereokorrelationserfolg nicht zwangsläufig zu genauer geschätzten DHM-Punkten führt. Jedoch bleibt die Nutzung des Verfahrens 3. wegen des Maßstabs der photogrammetrischen Anwendung schwierig. Die Positionierung von Kontrollpunkten war innerhalb des photogrammetrischen Koordinatensystems nur auf ±10 mm möglich, was für eine Kiesbettoberfläche mit einer Höhenvarianz verbunden war, die ähnlich aber manchmal auch größer war. Die nächste Stufe dieser Forschungen erfordert die Verwendung von besseren Kontrolldaten, möglicherweise aus einer Profilbestimmung mittels Lasern.