Surface Approximation from Industrial SEM Images*

A.J.Lacey, N.A.Thacker and R.B.Yates
Dept. of Electronic and Electrical Engineering
University of Sheffield
Mappin St., Sheffield, S1 3JD
United Kingdom
email: a.lacey@sheffield.ac.uk

Abstract

We are currently involved in an industrial project to recover depth information from stereo image pairs retrieved using a scanning electron microscope, (SEM). Feature based approaches to stereo provide accurate disparity estimations, however the quantity of estimates recovered is small (typically 1-2% of the image). If a continuous approximation to the surface is to be reconstructed, as requested by potential customers, more data has to be recovered.

Our approach involves using the disparity estimates from a feature based stereo algorithm to constrain a function fitting process. Assuming the image may be represented by an iterated facet model, the algorithm attempts to fit piecewise polynomials between the feature disparity estimates, which describe the mapping of grey-levels from the left to right image along epi-polars. The problems of illumination variation between the left and right images have been addressed using a modification to rank-order filtering which we call 'soft' ranking. The fitted functions are then used to calculate intermediate disparities.

1 Introduction

The limited application of machine vision algorithms to industrial problems is something the EPSRC Industrial Machine Vision (IMV) initiative is attempting to address. We are involved in an IMV project to develop and transfer our stereo depth estimation algorithms for industrial collaborators, who are interested in surface approximations from SEM images. Our collaborators have the following requirements of the project;

- An estimate of disparity at every pixel location,
- Disparity estimates with an accuracy of 0.01 of a pixel,

*This work is supported by the IMV EPSRC grant No. GR/K45272.

BMVC 1996 doi:10.5244/C.10.28

- Minimum of user intervention, preferably a single button press.

One of the most direct ways of recovering 3-dimensional information from images is by stereopsis. Previously, we have reported upon an algorithm to perform stereo depth estimation using stretch correlation [4]. This algorithm is a cross-correlation process which compares information rich areas of the two images, while modelling the warping effects encountered in difficult stereo problems, effectively relaxing the front-o-parallel constraint. This algorithm was development of the PMF [7] feature based stereo algorithm, embodying the constraints of PMF in an computationally efficient area based algorithm.

However, given the typical quantities of features present in images, stretch correlation, (and any other feature based stereo algorithm), provides relatively sparse estimates of depth. For many applications, such as surface approximation, it is desirable that continuous depth data is presented. With such demands attempts have been made to augment the depth estimates from stereo matching approaches with results from other algorithms such as surface interpolation [8] and binocular shape-from-shading [1].

1.1 Multi-Stage Algorithm

Our starting point for this work has been the auto-calibration [5] and stretch correlation [4] algorithms. The requirements of this project have enabled us to further the stretch correlation algorithm and integrate it as part of complete approach to surface approximation. An outline of this multi-stage algorithm is given in figure 1.

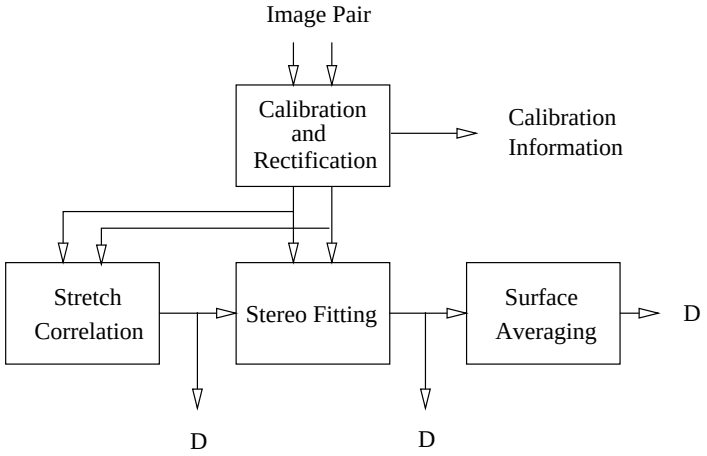


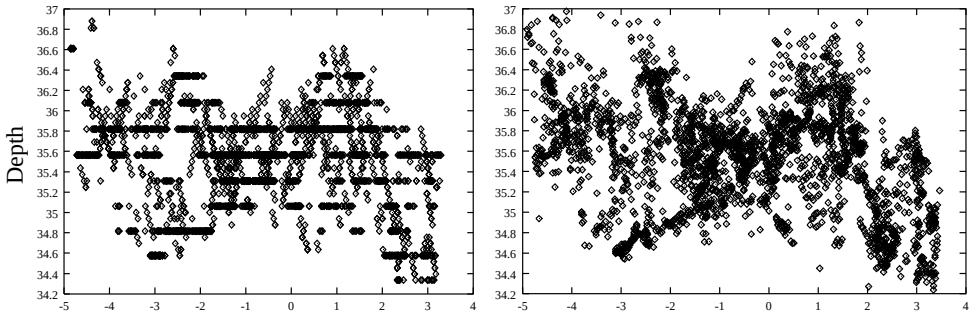
Figure 1: Overview of Algorithmic Stages. D represents the disparity image at each stage, moving left to right.

Once the stereo images have been calibrated they are rectified in order to align their rasters, taking care not to degrade the information content of the images. The stretch correlation algorithm is then used to recover sparse disparity estimates, denoted D_{sc}. These estimates are passed to the next stage which attempts to fit nth order polynomials to the left image data, between pairs of points in D_{sc} on the

same epi-polar, in order to predict data in the right image. The functions are then used to recover intermediate disparities estimates at fixed points along the function relative to the left image. Having already identified surface discontinuities using the stretch correlation algorithm, it seems appropriate to fit smooth functions between these estimates. The final stage is used to 'close' the surface by averaging local areas of the image and is essentially the algorithm described by Grimson [8].

2 Stretch Correlation

The stretch correlation algorithm is an area based solution to the stereo problem, matching discrete blocks in the left image to blocks in the right. However, focusing the matching process on information rich areas of the image (those containing non-horizontal edges) improves the robustness of the approach. Further, by warping blocks (stretching or shearing the right image blocks) rotation effects between views may be modelled, improving the disparity accuracy.



(a) pixel disparity (b) sub-pixel disparity

Figure 2: (a) Original estimates of disparity, (b) disparity estimates recovered by post-matching Canny edges.

Due to compromises built into the algorithm for the purposes of hardware implementation, the stretch correlation algorithm constrains the estimation of edge location to the nearest pixel and assumes that the stereo disparities lie on matched disparity planes. The effects of these compromises are clearly demonstrated by graph (a) of figure 2 which shows a horizontal cross-section through some disparity estimates recovered with this approach. The planes of depth, shown as horizontal quantisation banding, are caused by the disparity quantisation. The skewed linear correlations, shown as diagonal correlations in the graph show how the disparity estimates are constrained to lie on fitted planes.

It is a relative straight-forward matter to modify the algorithm in order to remove these constraints. In the original algorithm edge enhancement was performed by thresholding a horizontally differenced image, providing pixel accurate edge locations. This data is used both to direct the block matching process and also to calculate pixel disparities from the matched block disparities. By replacing the edge enhancement algorithm with Canny edge detection the edges may be

located with much greater accuracy, (approximately 0.1 pixel). Matching these edges using the results from the stretch correlation algorithm improves the accuracy of the disparity estimates and removes the planarity constraint imposed by pixelation.

The plot of figure 2(b) shows disparity estimates recovered using matched Canny edges, extracted from the same image pair as figure 2(a). With this modification the stretch correlation algorithm provides accurate disparity estimates gained from matched Canny edges, while using a computationally efficient correlation based framework.

3 Piecewise Surface Approximation

The estimates of disparity recovered using the stretch correlation algorithm are known to be reliable and well located, these points have been denoted D_{sc}. It is the distinct nature of the features used which enables this algorithm to reliably match regions in the left and right images. Attempting to perform a similar process at all points in the images would result in high numbers of incorrect matches, leading to unreliable estimates of disparity.

If it is assumed that the surface structure between points in D_{sc} is smooth, then it follows that the disparity of points between those in D_{sc} will also vary smoothly. Indeed given the data being analysed by our collaborators, (crystal structures, see figure 6), a facet model of near planar surfaces is a justifiable hypothesis. By approximating the surface between these estimates using smooth functions it should possible to recover disparity estimates at points along the function.

Piecewise polynomials are fitted between points in D_{sc} using a least-squares metric. The algorithm estimates the function parameters which most accurately map data in the left image to data in the right. The images have already been rectified so that the left and right rasters align. Horizontally consecutive pairs of points in D_{sc} are used to constrain the ends of the fit, however the fitting process stops short of these points as the behaviour of grey-levels at such discontinuities is unreliable.

3.1 Rank order filtering

The process of matching grey-levels in the left and right images is inhibited by the effective variation in image illumination caused by the change in viewpoint. Consequently, we would not expect to be able to simply match image grey-levels directly using a least-squares approach, except under very limited circumstances, i.e. photogrammetric invariance. However, some authors [3] have applied *ranking* schemes to stereo images in order to regain much of this property.

The process of image ranking is analogous to powerful ranking techniques used in statistical analysis. Comparing the rank of data sets is appropriate if the qualitative behaviour of the sets is similar but the absolute quantities are unreliable. We have chosen to investigate this approach to see if this assumption is justified. The ranking scheme adopted by O'Neil [3] replaces the grey-level of each pixel with the total number of neighbouring grey-levels it is greater than. This value

is called the rank of the pixel, and we will term this approach *hard* ranking. The weighting scheme for this is shown in figure 3(a).

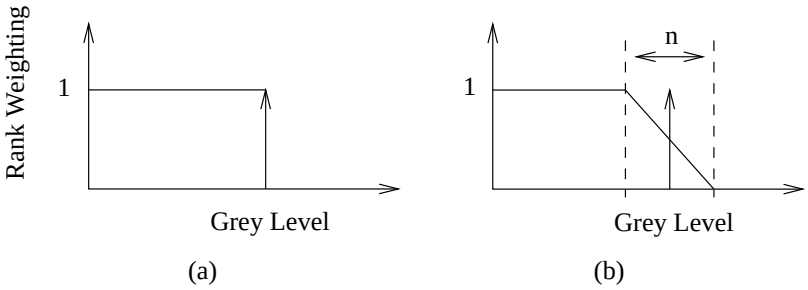


Figure 3: Weighting schemes for (a) hard rank-order filtering, (b) soft rank-order filtering

What should result is an image whose values relate to the statistical significance of the grey-levels in the original image. Pixels from locally smooth regions in the image will be replaced by mid-values, pixels at discontinuities will be replaced by extreme values.[1]. However, we have found that hard ranking is very sensitive to noise. Therefore, using a probabilistic interpretation of this ranking process as a guide, we have modified this scheme to take account of image noise, replacing the hard cutoff with a continuous ramp, see figure 3(b). The width of the ramp is specified directly by the expected error on the grey-level measurements due to quantisation. For an example of the visual effects of hard and soft ranking see figure 4.

(a) Original image region

(b) Hard ranked image region

(c) Soft ranked image region

Figure 4: Example rank order filter results on a section of a Lego brick house

In order to demonstrate the benefit of rank-order filtering, an artificial surface has been shaded using two different illumination sources. The shading has been performed in order to model variations in the left and right images due to a change in viewpoint. In both cases the shading was performed using a point Lambertian illumination model. In the first case the slant and tilt of the light source, relative to the surface were $slant = 0.5\ radians$, $tilt = 0.5\ radians$. In the second case

[1] The actual values depend upon the size of the window being used, typically 11×11

$slant = 1.0\ radian$, $tilt = 1.0\ radian$, providing an approximate 30^o rotation between the two images, figure 5 shows the resulting 2D images. In the final images, figures 5(b) and 5(c), the illuminated surfaces were both sampled and quantised in a manner typical of a CCD camera.

(a) (b) (c) (d) (e)

Figure 5: 2D views of illuminated surface

The image of figure 5(d) shows the result of subtracting (c) from (b). Figure 5(e) shows the result of rank-order filtering (c) and (b) and then performing subtraction. As may be observed, the rank ordered differenced image, figure 5(e), behaves more like a random variable than the difference of the original images, figure 5(d), which is badly affected by the central illumination artifact. Further, the rank ordering process increases the information ratio, (shown below), from approximately 1 to 10, suggesting a significant improvement in the S/N ratio. Finally the original differenced image contains an offset, something which the ranking process removes.

$$Information\ ratio\ =\ \frac{Dynamic\ range\ of\ original\ images}{Dynamic\ range\ of\ differenced\ image}$$

The variance on the ranked image, where n is the dimension of the ranking kernel in this case 11, becomes;

$$Var\left(R\left(I_i\right)\right)\ =\ R\left(I_i\right) - \frac{R\left(I_i\right)^2}{n^2} \qquad Var\left(\arcsin\left(\frac{2R\left(I_i\right)}{n^2} - 1\right)\right)\ =\ k$$

However, generally for a smooth continuous surface;

$$\frac{R\left(I_i\right)}{n^2} \approx 0.5 \quad \Rightarrow \quad Var\left(R\left(I_i\right)\right)\ =\ k$$

Obviously this limited experiment does not completely model the stereo imaging process. The consequences of projection transformations such as self occlusion, which also affect the reliability of the matching process have not been been modelled. Ranking produces data which is sensitive only to qualitative variations in the input image. Thus a ranking process generates data suitable for grey-level fitting provided that the ranking kernel is defined over equivalent sections of the data in the two images.

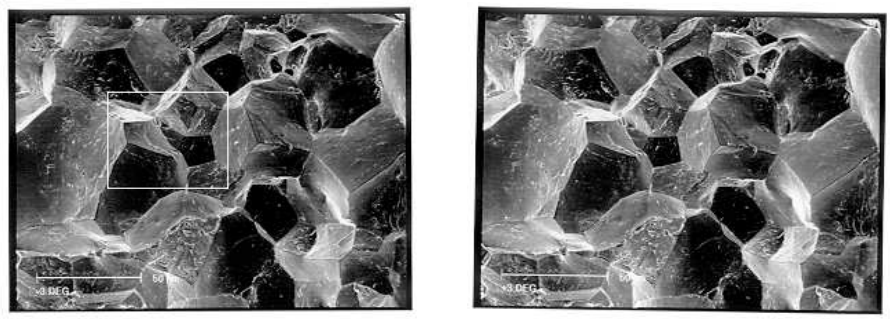
4 Surface Averaging

The function fitting process does not completely close the surface. Functions cannot be fitted unless the separation between the disparity estimates of D_{sc} is sufficiently large. It is also possible for the fit to terminate incorrectly. Therefore the final stage in the algorithm is to close the surface using a simple local averaging process.

The algorithm described by Grimson [8] attempts to fit the smoothest surface[2] across fixed points in the image. We have implemented this algorithm as a local smoothing algorithm, using a simple 3×3 Laplacian filter, which can be shown to be mathematically equivalent. This operation is only used to interpolate at points where no disparity estimate has previously been made.

5 Demonstration

As an example of what we are currently able to achieve using this technique, the images of figure 6 have been analysed. These image were provided by our collaborators and are typical of the images they wish to analyse. For the purposes of this demonstration the surface approximation has been performed within the region identified by the white box in figure 6(a).



(a) (b)

Figure 6: Typical SEM image pair, (a) is used as the left image (white rectangle shows region of interest) and (b) is used as the right

5.1 Image capture

The SEM images were captured using a back-scattered electron (BSE) beam microscope. The stereo pairs are constructed by capturing one image and then tilting the table on which the structure is laid by approximately 5^o, before capturing the second image.

[2]The surface with the smallest second derivative, (an elasticity constraint).

5.2 Camera calibration

The images in figure 6 were recovered using an unknown camera system. In order for absolute depth information to be recovered and the benefit of epi-polar geometry to be utilised, it was first necessary to calibrate the images. To do this the calibration process described by Thacker [5] was used. This modular approach to calibration allows the epi-polar error between matched image corners to be minimised, permitting the auto-calibration of arbitrary image pairs. The algorithm is currently using a physical camera model which can include terms for parameters such as focal length, the centre of the C.C.D. elements and radial distortion.

5.3 Disparity results

The plot in figure 7(b), shows the output from the stretch correlation process. The plots in figures 7(c), 7(d) and 7(e) show the results of fitting first, second and third order models, respectively to the data between the disparity estimates in figure 7(b). The image of figure 7(a) shows the region of the left image used, re-mapped to compliment the disparity plots.

6 Future Work

We have already designed, fabricated and tested a VLSI device which is capable of accelerating a large percentage of the stretch correlation algorithm [6]. We are currently developing an integrated system, incorporating this and other devices, in order to accelerate many of our vision algorithms. We then intend to explore the benefits afforded by temporal sequence stereo.

The function fitting is able to fit parameters once the order of the model has been defined. However, as is demonstrated by the plots in the figures 7(c), 7(d) and 7(e) the appropriate model order is data dependent. Because of the planar nature of the surfaces in this data the linear model seems the most appropriate, although in some areas a second order fit is justified. For a more generic fitting process the technique must be able to select the order of the model itself, using only the data. A possible solution to the problem of automatic model selection using the Bhattacharyya metric is described in the paper [2]. In future we hope to integrate this process into the function fitting algorithm and evaluate its performance.

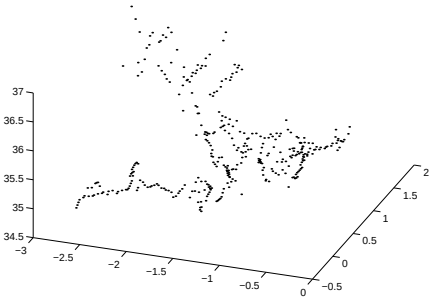
The camera calibration parameters used, do not directly relate to physical characteristics of the SEM, however a useful camera model is still achieved and is used to rectify the images. Our collaborators are providing a more suitable description of the SEM setup and this will be integrated into the calibration process.

In the final stages of the project we will be advising our collaborators on how to integrate the algorithms into their software.

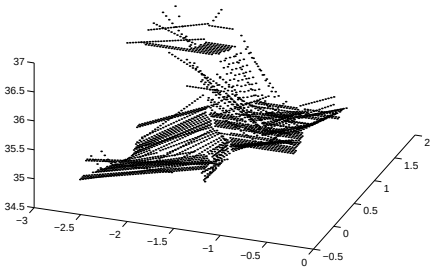
7 Conclusions

We have demonstrated how an existing stereo algorithm may be applied to a real industrial vision problem. In order to increase the density of disparity estimates, a function fitting process has been developed, integrating with the stretch correlation

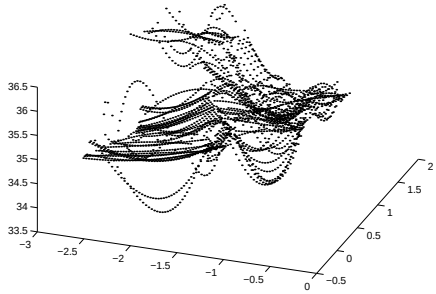
(a) ROI from left image



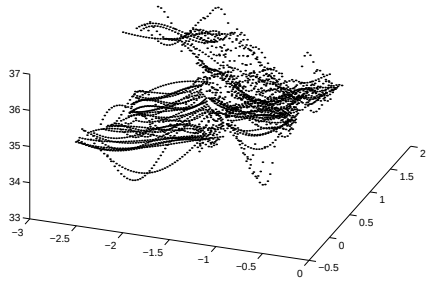
(b) Disparity estimates from stretch correlation



(c) Fitting results using first order function



(d) Fitting results using second order function



(e) Fitting results using third order function

Figure 7: Disparity estimates

algorithm. By rank order filtering the data, the fitting process is made more stable as the data attains the desired properties, i.e. approximating photogrammetric invariance with uniform variance. We would advocate such preprocessing before any correlation driven algorithms, e.g. auto-correlation based corner detection.

The user has requested accurate, dense disparity estimates achieved with the minimum of operator intervention. We are currently able to provide a low main-

tenance calibration algorithm and a process which provides dense estimates of disparity with varying degrees of accuracy. The most accurate of these measurements we now estimate to be in the region of 0.1 pixels. However, we do not believe it possible to provide highly accurate estimates of disparity at all points in the image. In fact, after discussions with our collaborators, we do not believe that such a requirement is actually necessary. Therefore we have recommended that quantitative measurements of the data are made only at feature point disparity estimates, D_{sc}, and that all other disparity estimates are used for qualitative analysis only.

8 Software Availability

In order to evaluate these algorithms for yourself, Sun Sparc binary demonstrations are available for the calibration, image rectification, stretch correlation, soft ranking and function fitting algorithms from :

`http://www.shef.ac.uk/~eee/esg/research/tina.html`

References

[1] Jones A.G. and Taylor C.J. Scale space surface recovery using binocular shading and stereo information. In *Proceedings of the BMVC*, volume 1, pages 77–86. BMVA, 1995.

[2] Thacker N.A. Prendergast D.J. and Rockett P.I. B-fitting: A statistical estimation technique with automatic parameter selection. In *Proceedings of the BMVC*, 1996.

[3] O'Neil M.A. and Denos M.I. Practical approach to the stereomatching of urban imagery. *Image and Vision Compt.*, 10(2):89–98, March 1992. 92.

[4] Lane R.A. Thacker N.A. and Seed N.L. Stretch correlation as a real-time alternative to feature based stereo matching algorithms. *Image and Vision Computing*, 12(4):203–212, 1994.

[5] Thacker N.A. and Mayhew J.E.W. Optimal combination of stereo camera calibration from arbitrary stereo images. *Image and Vision Computing*, 9(1):27–32, February 1991.

[6] Lane R.A. Thacker N.A. Seed N.L. and Ivey P.A. A stereo vision processor. In *Proceedings of the CICC*. IEEE, May 1995.

[7] Pollard S.B. *Indentifying Correspondences in Binocular Stereo*. PhD thesis, Psychology : AIVRU, University of Sheffield, 1985.

[8] Grimson W.E.L. An implementation of a computational theory of visual surface interpolation. *Computer Vision, Graphics and Image Processing*, 22:39–69, 1983.