Statistical Analysis of a Stereo Matching Algorithm

N. A. Thacker* and P. Courtney.
AI Vision Research Unit, University of Sheffield
Sheffield, UK
(* now at the Department of Electrical and Electronic Engineering, University of Sheffield, Sheffield, UK)

Abstract

This paper discusses the problems of image processing algorithm design and comparison and suggests that a suitable approach may be to model algorithms. We introduce the corner matching algorithm which we have used to provide reliable data for 3D computation modules [5][6]. The development of a simple model of the matching process permits the understanding of the influence of various parameters in the matching algorithm. This model also allows optimisation of the algorithm using data distributions obtained from representative scenes.

1 Introduction

The vision literature contains many algorithms for image processing involving feature extraction and matching. Often these algorithms take the form of heuristic solutions which attempt to exploit natural properties of generic images. Such algorithms are rarely perfect and peformance on real images is often degraded compared to simulated images due to the presence of noise. Clearly some algorithms must be better than others at doing a particular task, but how can we determine which? The conventional method for algorithm evaluation, demonstration and comparison seems to be to show the results on a set of "standard" images. This is useful to show that the algorithm has been successfully implemented and will work on real data [10]. However, the validity of the assumptions and heuristics underlying the algorithm can rarely be seriously tested in this way. There are several reasons for this: algorithm performance is often determined by the specific images for which it has been developed. Secondly, it is nearly always impossible to obtain an absolute measure of performance (often one needs to be defined and this can be a matter of subjective choice). Finally, correct implementation and use of other people's algorithms is often difficult [1], due to a lack of information about control parameters.

It has been suggested [8] that many implementations of vision algorithms lack a stability analysis (see for example [9]). Fundamentally, the only way to evaluate the quality of output data from an algorithm is on the basis of how well the data is suited to a particular application [3]. However, we need to get away from the dependence of the evaluation of the algorithm on a specific source of input data. One alternative involves developing a model of the effects of an algorithm on input data distributions. This achieves two things, firstly

BMVC 1992 doi:10.5244/C.6.33

the statistical properties of the input data are specified and secondly the validity of any heuristics are made explicit so that any data independent behaviour can be identified. Also, by developing such a model the effects of algorithm modification may be directly assessed. Different algorithms can then be compared, either directly on the basis of their models, or on results predicted by their models for specific data distributions. Thus algorithm modeling can be fundamentally useful in understanding, optimising and comparing image processing algorithms. This paper has thus adopted this approach to describe the performance of a corner matching algorithm. Although this algorithm is relatively simple, we believe that the same basic analysis strategy could be applied to any other algorithm.

2 Feature Matching

The robust matching of any image features obtained from a pair of grey level images involves the use of a limited set of heuristics.
(a) local image similarity (eg image correlation).
(b) restricted search strategies (eg epi-polars in the case of stereo).
(c) disparity gradient (or smoothness) constraints.
(d) one to one matching.
(e) reliability.

The relative merits of any matching algorithm will be determined by the extent to which these heuristics are utilised.

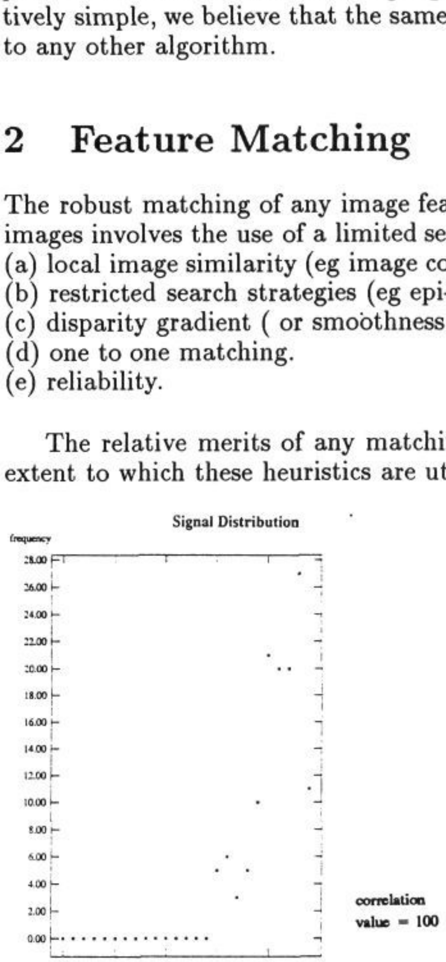


Fig 1(a): Cross correlation distribution for known correct matches.

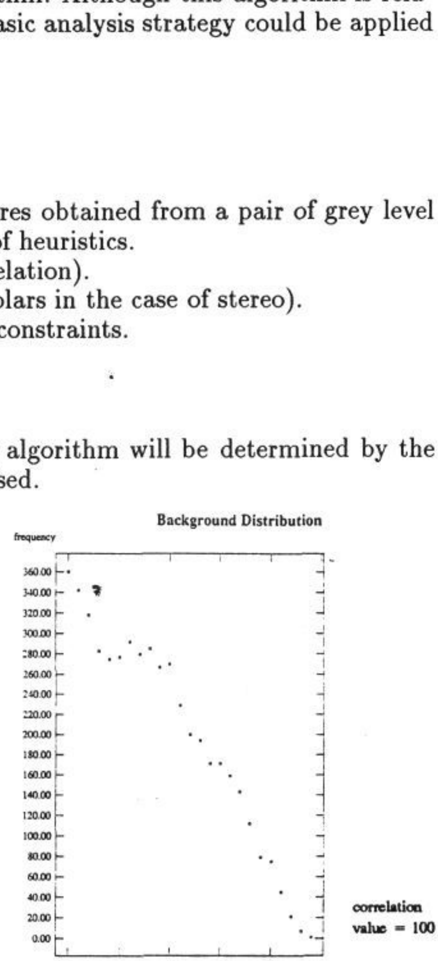


Fig 1(b): Cross correlation distribution for incorrect matches.

We use the corner detection algorithm of Harris and Stephens [2]. Our corner matching algorithm makes use of all but (c) as corners are generally too sparse to formulate a sensible constraint (but see [7]). Such information may be available from edge based matching algorithms but is not considered further here.

The basic matching algorithm we use for corners is as follows: (a) Construct a list of possible matches on the basis of limited search. This involves choosing a valid search area A. (b) Order the list according to a cross correlation measure c. (c) Select good matches on the basis of: (i) a threshold ρ on the minimum acceptable cross correlation $c < \rho$. (ii) a threshold ω on the ratio of absolute corner strengths $\frac{(c_1-c_2)^2}{(c_1+c_2)^2} < \omega$. (iii) reliability of the best candidate match c_m compared to the next best match c_n on the basis of a uniqueness parameter δ : eg: $c_m - c_n < \delta$. (iv) the same best candidate match must be obtained when matching from image 1 to image 2 and image 2 to image 1 (this enforces one to one matching). We have previously given reports on the performance of this algorithm for stereo/temporal matching for use in ego-motion determination and camera calibration [5],[6], but what we are aiming for here is a more systematic model of the effects of the parameters which control the matching process. We can consider these rules and control parameters as a prune to select a valid set of candidate matches followed by selection on the basis of image cross correlation. If the prune results in completely unambiguous assignment then the result of the matching process is already determined. If however, there are still several candidates for matching then the success or failure of the algorithm is determined by the extent to which the image correlation measure separates correct from incorrect matches. We start by justifying our image correlation measure as the best measure of its sort for choosing appropriate matches. By modeling the distribution of this measure for correct and incorrect matches we are able to assess the effects of the algorithms matching parameters.

2.1 What should we use as our match strength measure?

Given that corners are defined as the peak of an auto-correlation function it makes sense to use cross-correlation. There are many ways to construct the correlation function but we will assume that the function should be radially symmetric so we choose a function of a similar form to the corner detection definition.

$$c = \int_{-\infty}^{\infty} A^{-2} w_{uv} I_{uv} I'_{uv} \partial u \partial v$$

where I_{uv} is the image, $w_u v$ is a gaussian weighting function and with

$$A = \int_{-\infty}^{\infty} w_{uv} I_{uv}^2 \partial u \partial v \int_{-\infty}^{\infty} w_{uv} {I'}_{uv}^2 \partial u \partial v$$

We have hand selected a set of correct matches from several images of different objects including simple widget like objects, complicated machine castings, plastic childrens toys and cassette and lightbulb boxes. For this the cameras were configured at a typical verge angle of 0.1 radians and data at a distance of 5-10 inter-occular separations from the camera. The distribution of the correlation measure for these correct matches is shown in figure 1 (a). This distribution is generated by the differences in local image formation between the two images due to lighting, sensor differences and surface orientation. By computing the cross-correlation for incorrect matches we can get an idea of the shape of the background underneath the correlation signal when using this measure for matching (figure 1 (b)).

We have control over two aspects of the nature of this correlation measure the first is the range of the correlation. This clearly should be large compared to the localisation accuracy (0.3 pixels) but not so large that the correlation computation is costly or that we demand image similarity on a scale which is unrealistic. We use a range of 3 pixels.

Secondly we are free to choose the form for our cross-correlation measure by any non-linear rescaling of the image values:

$$I_u v := I_u v^n, I'_u v := I'_u v^n$$

We find that the best signal to noise ratio is given when $n = 1$ (figure 2). This is presumably because the corner detection auto correlation is also defined on the original image ($n = 1$) and this is therefore the correlation measure that we us in the matching algorithm.

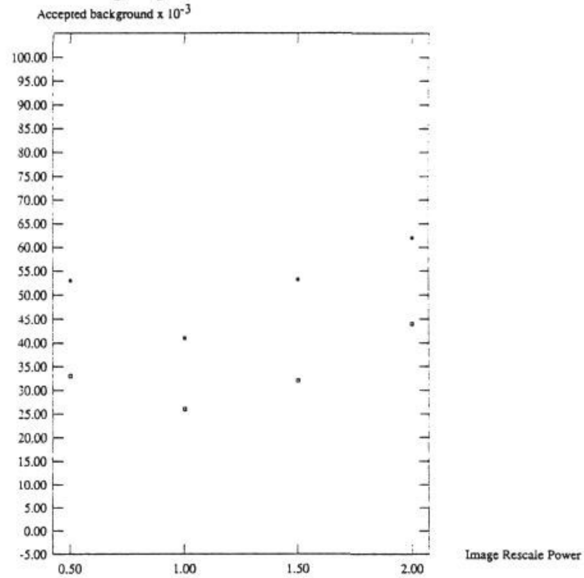


Figure 2: Incorrect cross correlations above 90% and 95% signal cut.

3 Getting an Incorrect Match

The possible forms of mismatch and signal rejection are determined by the reliability of corner detection process. An inefficient corner detector and occlusion will generate cases where some corners do not have a detected partner and can only be matched incorrectly. We can analyse the conditions under which we will get a mismatch and reject a correct match by considering each corner feature and its available match candidates in turn. We will thus show how the probabilities of accepting noise and rejecting signal can be controlled by the parameters ρ, ω, A and δ in the matching algorithm. In the following analysis we assume that the cross correlation distributions for correct and incorrect matches are independant of the detection process.

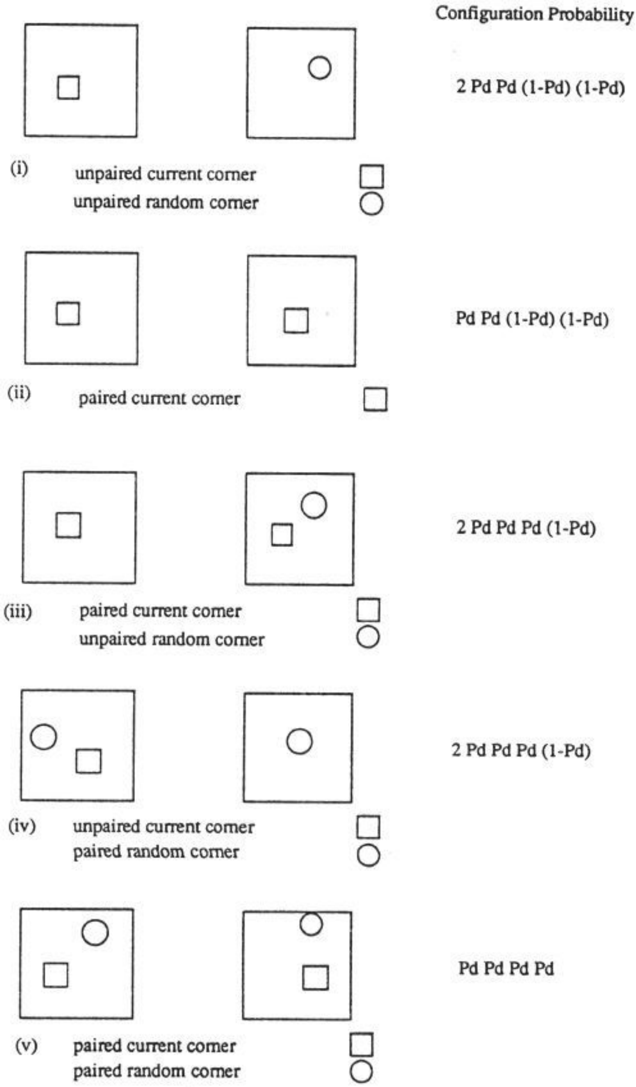


Figure 3: Basic Detection Configurations.

Case (a): The first matching case we consider is that of finding that the cross-correlation for the best candidate match x_m to the corner under consideration ("current") is incorrect in the case where neither candidate has their pair detected in the other image (Figure 3 (i)). The probability P_m^a for the best candidate match to the current corner being one of n_u unpaired random corners can be computed as a product of several factors. These are: the probability of getting the configuration shown in figure 3 (i) and the probability $P_{I(\rho)}$ that a random match drawn from the cross correlation for incorrect matches P_N will have a value greater than ρ:

$$P_m^a = 2n_u P_d^2(1 - P_d)^2 P_I(\rho)$$

where P_d is the probability of finding a corner again given that it has already been detected in one image and

$$P_I(x) = \int_x^1 P_N(a)da$$

Notice that this source of matching error is very sensitive to the reliability of corner detection P_d, and the only way to remove such matches is to increase the minimum accepted cross-correlation ρ.

Case (b): The next case we consider is that of obtaining an incorrect match with one of n_u unpaired random corners when the correct match to one of the corners was also present (Figure 3 (iii) & (iv)). This is slightly more complicated than the previous case because the existence of the correct match in the matching list may still prevent this getting accepted as a match due to the uniqueness parameter δ. The probability of this happening P_m^b is given by:

$$P_m^b = 4n_u P_d^3(1 - P_d)P_n(\delta, \rho)$$

where $P_n(\delta, \rho)$ is the probability that an incorrect match can be chosen even when the correct match is present in the match list above a value of ρ. Given that P_N and P_S (the cross correlation distribution for correct matches) are uncorrelated this is given by:

$$P_n(\delta, \rho) = \int_\rho^{1-\delta} P_S(x) \int_x^{1-\delta} P_N(a - \delta)dadx$$

We can see from this that as the uniqueness factor increases the probability of keeping a noisy match of this type is reduced.

Case (c): Finally we consider the case where the current corner is paired but has been matched incorrectly with one of $2n_p$ paired random corners (figure 3 iii). We may wish to write the probability for the acceptance rate for mismatches P_m^c as:

$$P_m^c = 2n_p P_d^4 P_n(\delta, \rho)^2$$

This equation assumes that the two probabilities for mismatch $P_n(\delta, \rho)$ are uncorrelated which is unrealistic, as the two corners must be figurally similar if they are to have mismatched in one matching direction. Thus it is better to write this as:

$$P_m^c = 2n_p P_d^4 P_n(\delta, \rho) P_k(\delta, \rho)$$

where $P_k(\delta, \rho)$ is the probability that the cross correlation value for the complementary pair of the original random match will also be bigger than the correlation value for the correct match.

4 Rejected Signal

We now consider ways in which corner pairs are rejected by the matching process.

Case (a): The first case we consider is when the current corner has been detected in both images (ie paired) and a random matching feature has not been detected in either image (Figure 3 (ii). The probability of rejecting this match is given by:

$$P_r^a = P_d^2(1 - P_d)^2 P_J(\rho)$$

where

$$P_J(\rho) = \int_0^\rho P_S(x)dx$$

Case (b): When the current match is paired and there is a random unpaired match present in either image (Figure 3 (iii)). The probability of rejecting a correct match due to the presence of n_u unpaired random corners P_r is given by:

$$P_r^b = 2P_d^2(1 - P_d)^2(P_J(\rho) + n_u P_l(\delta, \rho))$$

where $P_l(\delta, \rho)$ is the probability of rejecting a correct match due to the proximity of a random corner.

$$P_l(\delta, \rho) = \int_\rho^1 P_S(x) \int_x^{1+\delta} P_N(a + \delta)dadx$$

Case (c): The final case for consideration is when the current match is paired and there is a random paired match (Figure 3 (v)). The rejection rate for good corner matches P_r^c is given by

$$P_r^c = P_d^4(P_J(\rho) + 2n_p P_l(\delta, \rho) - n_p^2 P_l(\delta, \rho)^2)$$

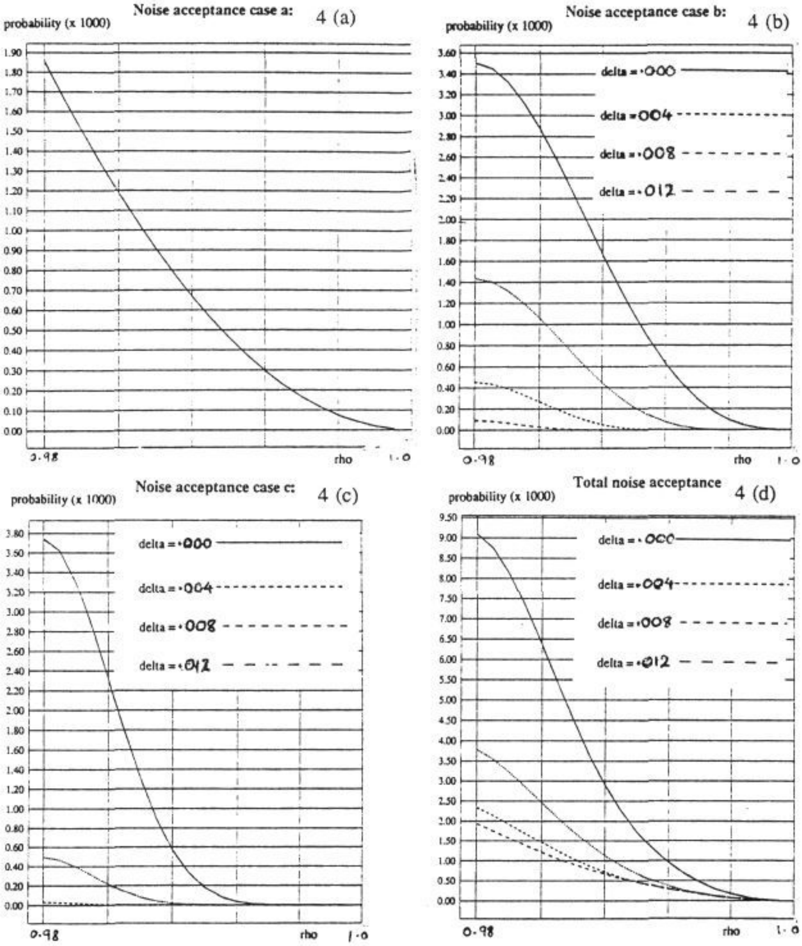
5 Algorithm Analysis and Conclusions

Although the above probabilities can be computed directly from examples of the cross correlation histograms for signal and background there are advantages to modelling the data as a functional form. With an analytic model of the matching process we can compute directly the effects of varying parameters in the matching algorithm and can thus minimise the number of background matches obtained, for various amounts of signal, using standard numerical minimisation methods.

The cross-correlation distributions $P_N(x)$, $P_S(x)$ and $P_k(\delta, \rho)$ can be approximated by triangular distributions. The remaining unknown parameters are the detection efficiency P_d and the numbers of paired and unpaired random corners n_p and n_u. In some ways these values are closely related as the ratio $n_p : n_u$ has a maximum value of $P_d : 1 - P_d$. In an application where the full image contains several hundred corners and the search regions are of the order of a few percent of the image we estimate these values as;

$$P_d = 0.85, n_u = 0.75, n_p = 4.25$$

With these values we can now compute typical signal rejection and noise acceptance curves for the matching algorithm as a function of the matching parameters ρ and δ (Figures 4 and 5).



We can thus approximate the total number of incorrect matches (Figure 4 (d)) by

$$P_m^T = P_m^a + P_m^b + P_m^c$$

and the total rejection rate for paired corners (Figure 5 (d)) as

$$P_r^T = P_r^a + P_r^b + P_r^c$$

In specific applications where the detected corners have correlated properties the probability distributions for cross correlations of signal and background may be significantly different. In these cases the probabilities for mismatch and signal rejection would also be different. However, we can still draw some qualitative conclusions about the generic case of corner matching which must be true regardless of the signal and background distributions. These are:

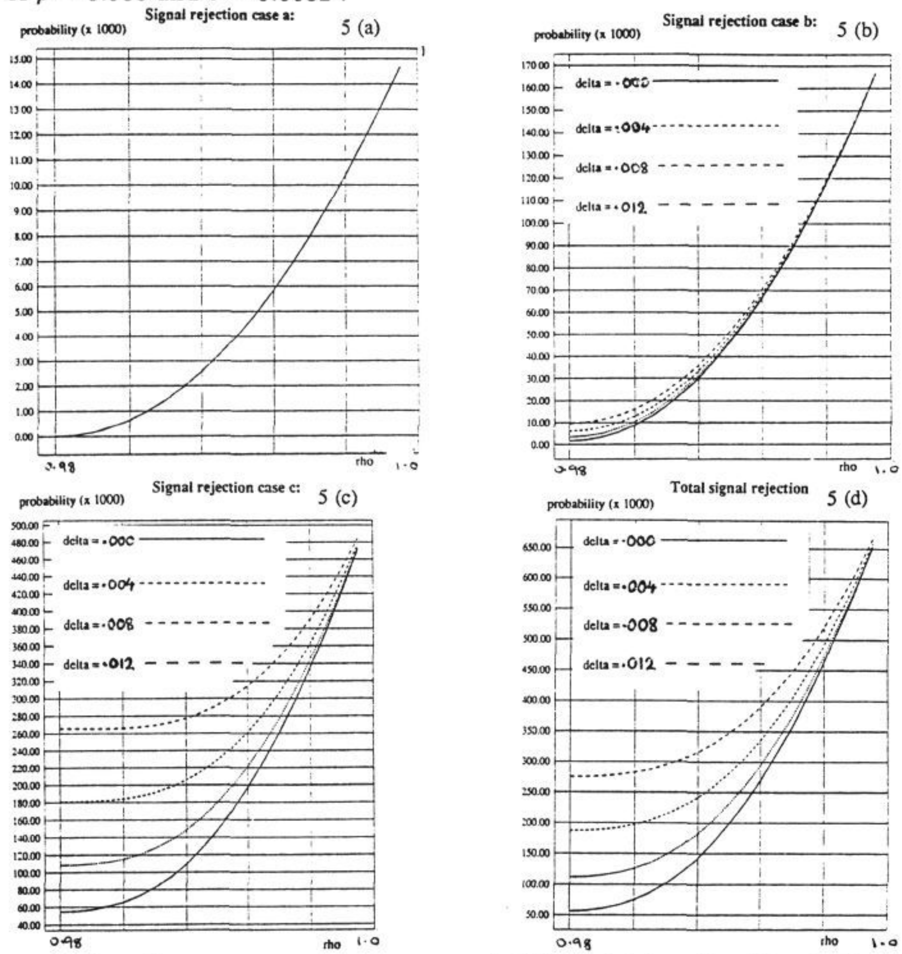
1) All terms in P_m^T are proportional to the mean number of candidate matches, thus we would expect the total number of mismatches to vary proportionately with the search area A.

2) We expect type (a) mismatches to be a very small fraction of the total number of mismatches. The only way to remove these is to increase the minimum required cross correlation value ρ.

3) We expect type (b) and (c) mismatches to be of roughly equal importance and both are reduced considerably by use of the uniqueness parameter δ at the cost of only marginal reduction in the overal number of matches.

4) There is no improvement obtained by increasing δ beyond a value of $1-\rho$ as at this point all mismatches of type (b) have already been rejected.

5) There is no set of parameters which give an optimal signal to noise ratio, this value keeps on rising with increasing ρ. There are however optimal values of ρ and δ corresponding to the minimum noise obtainable for a required proportion of signal. For example using the above model for the data the minimum noise obtainable at a signal level of 60% is 0.2% at parameter values of $\rho = 0.985$ and $\delta = 0.0032$.



6) Even in very severe cases we expect this matching algorithm to have a

signal to noise ratio in excess of 100:1.

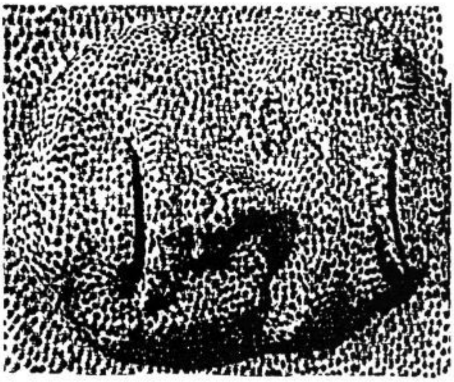


Figure 6 (a) Left Image.

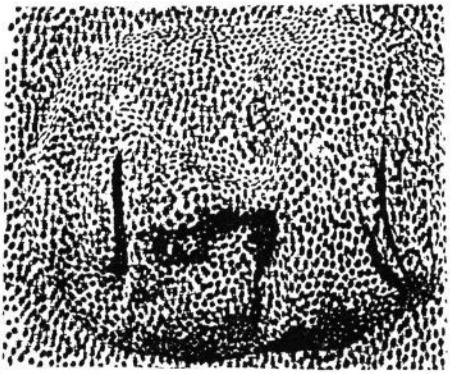


Figure 6 (b) Right Image.

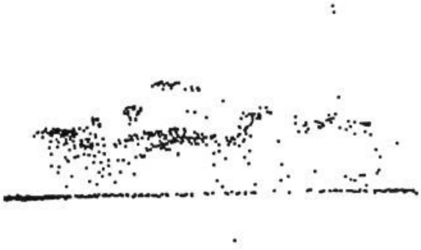


Figure 6 (c) Reprojected 3D data from stereo.

The algorithm is demonstrated here with stereo images of a highly textured head (figure 6(a) and (b)). In this case there were in excess of 1000 corners detected with an average of 36 candidate matches for each corner. For parameter values of $\rho = 0.99$ and $\delta = 0.002$ the predicted number of correct and incorrect matches of 600 and 7. As can be seem from the reprojection of the reconstructed 3D data (figure 6(c)), this is close to what is observed.

The discussion of the matching algorithm has centred on selecting a set of candidate matches and then choosing between them. Given that the cross correlation distributions can be replaced with any relevant similarity measure this work can be considered as a general model of constrained feature matching. This model puts us on a sound footing for considering potential modifications to the corner detection and matching processes.

Acknowledgements

We gratefully acknowledge the grant holders Prof. John E.W. Mayhew Dr. Paul Dean and Prof. John Frisby, the support of ESRC/MRC/SERC and the EEC, under the ESPRIT VOILA project, for the funding of this work. We would also like to thank Prof. David Hogg for letting us have this extra page for our paper.

References

[1] Day, T. and J.P.Muller, Digital Evaluation Model Production by Stereo-Matching SPOT Image Pairs: A comparison of Algorithms, Proc. 4th Alvey Vision Conference, pp.117-122, 1988.

[2] Harris, C. and M.Stephens, A Combined Corner and Edge Detector, Proc. 4th Alvey Vision Conference, pp.147-151, 1988.

[3] McLauchlan, P.F., J.E.W.Mayhew and J.P.Frisby, Stereoscopic Recovery and Description of Smooth Textured Surfaces, Proc. BMVC, pp.199-204, 1990.

[4] Moravec, H.P., Obstacle avoidance and navigation in the real world by a seeing robot rover, Ph.D Thesis, Stanford Univ., Sept. 1980.

[5] Thacker, N.A., Y.Zheng and R. Blackbourn, Using a Combined Stereo/Temporal Matcher to Determine Ego-motion, Proc. BMVC, pp.121-126, 1990.

[6] Thacker, N.A. and J.E.W.Mayhew, Optimal Combination of Stereo Camera Calibration from Arbitrary Stereo Images, Image and Vision Computing, pp.27-32, vol 9 no 1, Feb. 1991.

[7] Wang, H. and J.M.Brady, A Structure-from-Motion Algorithm for Robot Vehicle Guidance, Proc. IEE Symp. Intel. Vehic., Detroit, June 1992.

[8] Aloimonos, J., Integration of Visual Modules, Academic Press, 1989.

[9] Grimson, W.E.L., D.P.Huttenlocher and D.W.Jacobs, A Study of Affine Matching with Bounded Sensor Error, pp.291-306, ECCV, Genoa, Italy, May 1992.

[10] Ferrari, D., G.Garibotto and S.Masciangelo, Towards Experimental Computer Vision, Performance Assessment of a Trinocular Stereo System, EEC ESPRIT Day, Santa Margherita Ligure, Genoa, Italy, 18 May, 1992.