[1] PMF: A Stereo Correspondence Algorithm Using a Disparity Gradient Limit

Stephen B Pollard, John E W Mayhew and John P Frisby

AI Vision Research Unit,
University of Sheffield, Sheffield S10 2TN, UK

Reprinted, with permission of Pion Ltd, from *Perception*, 1985, **14**, 449-470.

ABSTRACT

The advantages of solving the stereo correspondence problem by imposing a limit on the magnitude of allowable disparity gradients are examined. It is shown how the imposition of such a limit can provide a suitable balance between the twin requirements of disambiguating power and the ability to deal with a wide range of surfaces. Next, the design of a very simple stereo algorithm called PMF is described. In conjunction with certain other constraints used in many other stereo algorithms, PMF employs a limit on allowable disparity gradients of 1, a value that coincides with that reported for human stereoscopic vision. The excellent performance of PMF is illustrated on a series of natural and artificial stereograms. Finally, the differences between the theoretical justification for the use of disparity gradients for solving the stereo correspondence problems presented in the paper and others that exist in the stereo algorithm literature are discussed.

1 INTRODUCTION

Useful depth information about a scene can readily be recovered from the disparities that arise between a pair of images taken from different viewing positions, but measuring these disparities requires a solution to the difficult 'stereo correspondence problem'. That is, some means must be found for resolving the considerable ambiguities that typically arise in trying to match corresponding elements in the two images. In the present paper we address this problem from the computational viewpoint (Marr 1982), in that we regard our goal as that of finding appropriate constraints for solving an information-processing task. However, the analysis we offer has been to a considerable extent stimulated by psychophysical findings about human binocular vision (in particular Burt and Julesz 1980).

Constraints for solving the stereo correspondence problem must be derived from properties of surfaces in the world being viewed and/or the image formation process. For example, in the theory presented by Marr and Poggio (1976, 1979) the principal assumption is that correct matches will arise from the projections of points lying on surfaces that are smooth almost everywhere, and that this will not be the case for incorrect matches. A surface continuity assumption was also exploited by Mayhew and Frisby (1981), but they used it to justify a figural continuity matching rule. The advantage of the latter scheme was that it did not assume surface smoothness beyond the scale at which the image primitives were themselves described. The constraint that provides the disambiguating power in the theory presented here is that the disparity gradients that exist between correct matches will be small almost everywhere. The analysis we present in section 2 shows that, given camera geometry approximating the arrangement of human eyes, a disparity gradient limit of about 1 will almost always be satisfied between the correct matches arising from a large class of the surfaces that form our visual world, whereas this is not true of incorrect matches. The value of 1 is not in itself critical - all that is required is a limit that balances satisfactorily the competing requirements of disambiguating power and ability to deal with as wide a range of surfaces as possible (section 2). However, we have chosen 1 because this seems to be roughly the limit found for the human visual system (Burt and Julesz 1980).

In subsequent sections we describe the design and performance of an algorithm called PMF that is based upon this constraint, and several others that have been identified previously[1]. Finally we briefly discuss how the theory underlying PMF relates to other computational theories of how to solve the stereo correspondence problem.

2 THE DISPARITY GRADIENT LIMIT

Consider a simple stereogram made up of two dots in each field (figure 1). Burt and Julesz (1980) have provided evidence that if both dots are to be binocularly fused simultaneously by the human visual system, then the ratio of the disparity difference between the dots to their cyclopean separation must not exceed a limit of about 1. This ratio is known as the disparity gradient and the existence of a disparity gradient limit implies that even small absolute disparity differences will not produce fusion if the spatial separation between dots is also small. This is why the notion of a disparity gradient limit represents a substantial departure from the classical concept of Panum's fusional area, which is expressed in terms of a limit not on gradients but on absolute disparities.

The concept of a disparity gradient limit provides an elegant and parsimonious description of various psychophysical phenomena (Burt and Julesz 1980); see also Tyler (1973) for an estimation of the disparity gradient limit for horizontal sinusoidal variations in depth. However, its functional significance has yet to be established. The possibility we have investigated is that the human visual system introduces a disparity gradient limit in order to solve the stereo correspondence problem. The reason for entertaining this idea is, in brief, that for most naturally occurring scene surfaces, including quite jagged ones, the disparity gradients between the correct matches of image primitives (deriving from texture markings, surface edges, etc in the scene) are usually less than 1, whereas this is seldom the case for incorrect matches formed from the same image elements.

In order to understand why correct matches generally lie within a disparity gradient of 1 it is helpful to begin by visualising this limit as imposing within disparity space a cone-shaped 'forbidden zone' for fusions around any given match [the zone is cone-shaped because the limit is isotropic (Burt and Julesz 1980)]. This cone is shown in figure 2 which also illustrates

1 See also Pollard, Mayhew and Frisby (1990), paper [11] in this book, for further details and a description of refinements added to PMF since the present paper was published.

how it implies a further cone restricting the nature of surfaces in the viewed world that can satisfy the disparity gradient limit. Because the gradient of the world cone is approximately that of the disparity cone scaled by the ratio of the viewing distance, z, to the interocular separation, I (Appendix 1), the world cone does not impose a very severe constraint on the nature of surfaces that can satisfy it.

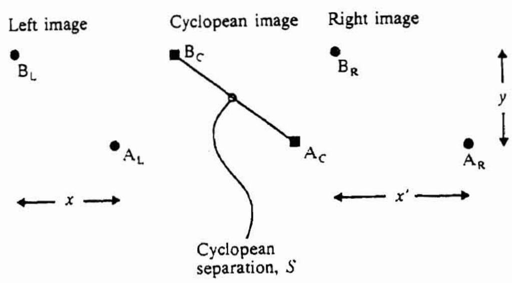

Figure 1 Definition of disparity gradient. Consider a stereogram comprising the left and right halves shown in the figure. When A_L is matched to A_R and similarly for B, the disparity gradient between them, as defined by Burt and Julesz (1980), is the difference in the disparity divided by their cyclopean separation, S. The latter is give by the distance between the midpoints of the two pairs of dots (located at A_C and B_C respectively). Hence:

$$S = \{[{}^1/_2(x + x')]^2 + y^2\}^{1/2}.$$

As the changes in disparity between the two matches is x' - x, the cyclopean disparity gradient, Γ_D, between A_C and B_C is given by:

$$\Gamma_D = |x' - x|[{}^1/_4(x + x')^2 + y^2]^{1/2}.$$

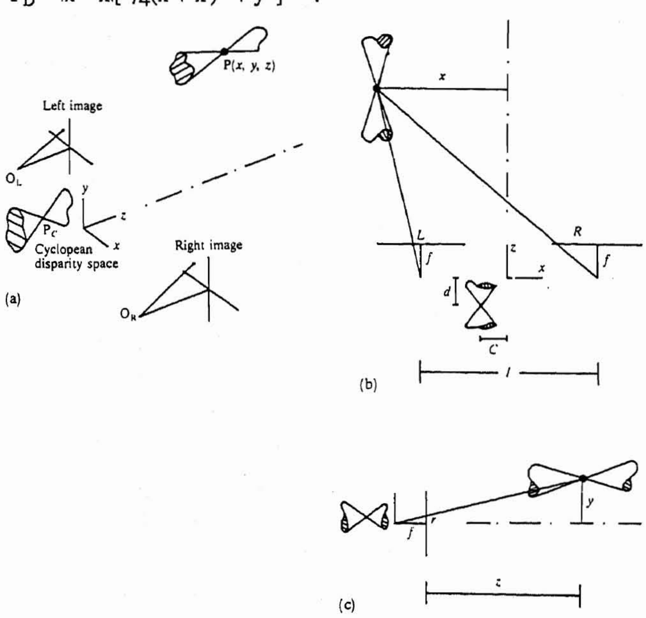

Figure 2 Forbidden cones shown schematically in cyclopean disparity space and viewed world space. Consider the imaging geometry shown in (a), with top and side elevations of the same shown in (b) and (c) to aid the presentation. The principle axes of the left and right imaging devices are parallel and have a interocular separation I . The left and right image planes are shown in front of the respective optical centres O_L and O_R for the pictorial simplicity. A world viewed space is defined with respect to a coordinate frame located midway between O_L and O_R with axes arranged as in the figure. The location of point P with respect to this frame is given by (x, y, z). The projections of P into the left and right image planes are given by (L, r) and (R, r) respectively. A cyclopean disparity space has also been constructed with its origin coincident with that of the world viewed space. In this space the point P_C corresponding to P in the viewed world space is located at (C, r, d), where C is the mean horizontal component of the projection into the left and right images given by ${}^1/_2(L+R)$, and d is the disparity defined as d = R - L.

From the definition of disparity gradient given in figure 1 it follows that the illustrated cone-shaped 'forbidden zone' will exist within the cyclopean disparity space around the point P_C. Other points that lie between the forbidden cone will violate the disparity gradient limit with respect to P_C. A second cone is shown in the viewed world space to illustrate that the forbidden cone in disparity space imposes an approximately similarly-shaped restriction on the allowable depth relationship between points lying on the surfaces that constitute the visual world. It is shown in Appendix 1 that the gradient of this second cone is approximately related to that of the first by z/I. Hence at any reasonable viewing distance the constraint imposed upon the structure of the world by a disparity gradient limit of 1 is not very severe. Note that the geometry in figure 2 is only schematic. Moreover, the described approximate relationship between world and disparity gradients holds only if z is large with respect to x, y, and I (see Appendix 1).

For example, planar surfaces with maximal slopes of up to 74° will be tolerated at a viewing distance of 6 interocular units, rising to 84° for 10 interocular units. Hence, at any reasonable viewing distance only a small proportion of planar slopes will have disparity gradients upon them that violate the limit of 1.

Turning to nonplanar surfaces, the disparity gradient limit will in general be satisfied on many types of surfaces provided they do not recede too rapidly from the viewer. Furthermore, surfaces satisfying the requirements need not be C1 smooth (ie they need not be everywhere differentiable with continuous derivative): mathematically, the disparity and world cones can be described as determining Lipschitz functions of order 1 with $K = 1$ and $K = z/I$ respectively (Appendix 2).

Intuitively, one can say that $K = 1$ in disparity space will permit world surfaces to be 'jagged but not too jagged'. The amount of allowable 'jaggedness' will depend upon the relationship of interocular distance to viewing distance. As will be illustrated later, such a rule provides considerable disambiguating power in terms of solving the stereo correspondence problem while not being overly restrictive about allowable scene surfaces. Furthermore, it is not necessary to assume that the disparity gradient limit need be satisfied between all correct matches in order that disambiguation can be achieved. Hence the Lipschitz condition provides only a conservative estimate of the class of surfaces that satisfy our constraint.

So far it has been demonstrated that the correct matches derived from the elements arising from most planar and many jagged surfaces will lie within a disparity gradient limit of 1.0 (given always small interocular separation with respect to viewing distance). But, as implied above, disambiguation needs also to rely upon there being a low probability that this limit will be satisfied by incorrect matches by chance. The convergent-to-divergent disparity range allowed between a pair of matches in order that they satisfy a disparity gradient limit of 1.0 is exactly twice their cyclopean separation. From this it follows that the probability that two points from one image can find incorrect matches that satisfy the disparity gradient limit by chance is almost directly proportional to their cyclopean image proximity [Appendix 3]. Hence at close proximities it is very unlikely that a pair of matches will 'accidentally' satisfy the disparity gradient limit. Thus a distinguishing feature that allows the identification of correct matches from the pool of 'possibles' is that the disparity gradients between the former almost always lie within a limit of 1.0, whereas between the latter this is not the case, especially at close image proximities.

Burt and Julesz (1980) pointed out that one way of viewing a disparity gradient is as a measure of figural disparity between a pair of matches because its size reflects both figural dimensions on which such a 'dipole' can differ - orientation and length. Similarly, it is possible to characterise the degree of satisfaction of the disparity gradient limit between matches in a fused binocular structure as a measure of the figural similarity of the projections of that structure into the two stereo halves. Hence an alternative description of stereo projections that will satisfy the disparity gradient limit is that they conform to a 'constraint of figural similarity', and simple observation of stereograms of most natural scenes shows that they satisfy this requirement almost everywhere. Hence it is possible to summarise this section by saying that, for a wide range of surface structures, because of the similarity of the vantage points of the two eyes the projected figural structures that constitute the left and right images will in general be sufficiently similar to allow corresponding points to be identified by using the strategy of finding the best figural matches available.

3 ADDITIONAL CONSTRAINTS UTILISED BY PMF

As well as the disparity gradient constraint discussed in the last section, PMF also exploits two other constraints: (i) the *epipolar constraint*, a limitation on the possible locations of matching points; and (ii) the *uniqueness* constraint, a limitation on the number of matches allowed for a single image entity. Both these constraints are used in a variety of other stereo algorithms.

3.1 The epipolar constraint

For the class of stereo imaging geometries with which we are concerned, ie those in which the optical axes of the two imaging devices lie in the same plane, all matching primitives appear on left/right pairs of (straight) epipolar lines (Baker 1982; see figure 3). Thus points along one member of the epipolar pair can only match with points situated along the other member, and vice versa. In the special case where the principal axes are in fact parallel, all epipolar pairs will be horizontal and matching points will be found on corresponding rasters.

As with many other algorithms PMF uses this constraint to restrict the search for possible matches to one dimension. In current versions of PMF it is assumed that the camera geometry is in fact parallel or at least approximately so. In the natural scene stereograms considered below the inter-camera separation was at least one tenth the distance to the fixation point. Given that our images are 128 x 128 pixels square and cover a visual angle of at most 10 deg, a vertical disparity of no more than a half a pixel will exist at the very corners of the image. Hence for the work reported here seeking matches along corresponding rasters is an adequate approximation to the correct epipolar geometry.

In PMF all potential matches within the (large) disparity range of +/- 30 pixels (corresponding to a Panum's fusional area of up to 5 deg) that satisfy a matching criterion for left/right

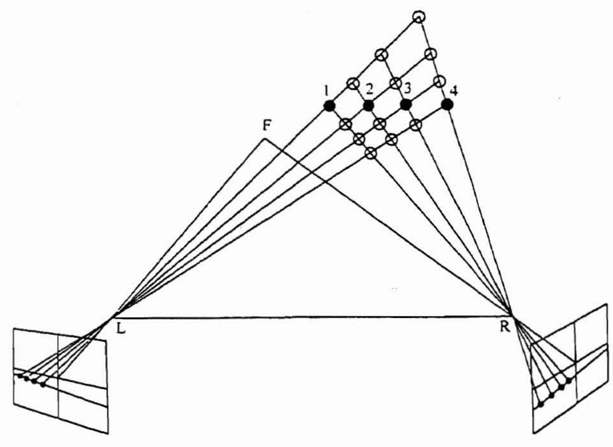

***Figure 3** Epipolar Geometry and the stereo correspondence problem. In the imaging geometries with which we are concerned the optical axes of the left and right imaging devices (LF and RF respectively) lie in a single plane and thus intersect at the fixation point F (possibly at infinity). Consider the plane, 1LR, formed by a single point in that space, labelled 1, and the optical centres of the two imaging devices, L and R. The plane 1LR is then said to intersect the two image plates along a pair of epipolar lines. Furthermore, as a result of the fact that all points along 1L and all the points along 1R are (by definition) limited to lie in the plane 1LR, all possible ghost matches associated with the point 1 must also lie in that plane and therefore project to the same pair of epipolar lines. The same is also true for all and only all the other points that lie in that same plane, for example points 2, 3, and 4 in the figure. In short, matching points can only be found along a corresponding pair of epipolar lines. The possible ghost matches that exist in this case are shown in the figure as open circles, with closed circles portraying the physically existing points that gave rise to them.*

image primitives are selected for subsequent disambiguation. The choice of matching criterion depends upon the nature of the adopted primitives, a decision which is in turn domain-dependent. PMF makes no specific requirements about what the nature of these should be, except that they should reflect scene entities. For artificial stereograms we simply use the points defined when the stimuli are being created (see figures 4-6, 9 and 10). For natural images we have found it convenient to use edge-like primitives given by zero crossings identified after the application of a single high-frequency Marr-Hildreth operator (Marr and Hildreth 1980; see figures 7 and 8). For the latter, potential matches have been restricted to those that can be formed between zero crossings of the same contrast sign and roughly similar orientation. In some domains it may also be possible to determine a measure of the 'goodness' of a match (eg degree of figural similarity of its constituent primitives) that can be made use of in the disambiguation procedure by weighting the relative importance of potential matches.

3.2 The uniqueness constraint

Owing to the fact that only rarely will a feature project into one member of a stereo pair such that it neatly masks an identical feature farther away which nevertheless remains visible in the other stereo half, it follows that matches of image primitives extracted from the two images should be unique, ie each image primitive will participate in just one match. Only in the unlikely case just mentioned, which violates the general position assumption with regard to viewpoint, would it be legitimate for a single primitive from one image to be matched with two primitives in the other.

Notice that the uniqueness constraint is itself implied by a disparity gradient limit of 1 (Burt and Julesz 1980). If a single point identified in one image is allowed to match with a pair of points identified in the other (cf Panum's limiting case), then the disparity gradient between the two matches is 2.

4 THE PMF ALGORITHM

We have developed a stereo correspondence algorithm called PMF based on the foregoing considerations. It has two distinct stages of processing.

First, the matching strength of each potential match is computed from the sum of contributions received from all potential matches in its neighbourhood that satisfy the disparity gradient of 1 with respect to it. We find a neighbourhood defined by a circle of radius 7 pixels satisfactory for all textures so far considered. Because the probability of a neighbouring match falling within the limit by chance increases (almost linearly) with its distance away from the match under consideration (section 2 and Appendix 3), the contribution of a match is weighted inversely by its distance away. Uniqueness is exploited at this stage by requiring that at most one match associated with a single primitive in one or other image makes a contribution to the matching strength.

Second, on the completion of this procedure, correct matches are chosen on the basis of matching-strength scores by using a form of discrete relaxation (Rosenfeld et al 1976). At the first iteration any matches which have the highest matching strength for both of the two image primitives that formed them are immediately chosen as 'correct', ie matches are selected whose primitives have no higher matching-strength scores with any other matches they can form. Then, in accordance with the uniqueness constraint, all other matches associated with the two primitives that formed each chosen match are eliminated from further consideration. This allows further matches, that were not previously either accepted or eliminated, to be selected as correct because they now have the highest strengths for both constituent primitives. Usually only four or five iterations are needed to propagate the uniqueness constraint in this way to the point at which the disambiguation process of PMF is complete and satisfactory (see section 6).

It is important to emphasise the fact that PMF is only interested in the quantity of within-disparity-gradient-limit support that exists for a particular match. Hence the extent to which the disparity gradient limit is offended in the neighbourhood of a candidate match does not directly affect the selection procedure of PMF. This design feature is perfectly in line with the justification given in section 2 for seeking within-disparity-gradient-limit support as the disparity gradient limit need not necessarily be satisfied everywhere.

We have also thought it sensible not to weight contributions by the actual magnitude of their disparity gradients, but instead to treat all within-disparity-gradient-limit contributions homogeneously. This makes sense for the stereo correspondence problem, as there seems to be no reason to penalise any perturbations that lie within the range that is to be expected in the stereo projections of interest. However, many other functions using disparity gradients could be employed. In one interesting example suggested by Prazdny (1985), the strength of the support that flows between a pair of matches is scaled in a gaussian fashion with respect to the size of the disparity gradient between them.

A further point to notice in this implementation of PMF is that within-disparity-gradient-limit support is sought independently from all possible matches in the neighbourhood of the match under consideration. This means that it is possible that two or more matches that give within-disparity-gradient-limit support might not themselves share a within-limit disparity gradient. This design feature has been dictated by considerations of computational efficiency. Further research is in hand with the object of examining the desirability of insisting on within-neighbourhood support consistency.

Although the use of a disparity gradient limit in PMF was stimulated by observations of the human visual system, various details of its design were shaped by more practical constraints introduced by the need to achieve reasonable efficiency and robustness on [near]-state-of-the-art computer machinery. The speed criterion is met by the intrinsically parallel nature of the structure of PMF: each matching strength could, on appropriate computer architecture, be computed independently. Extensive examination of its performance on various artificial and natural stereo images bears out satisfaction of the robustness requirement. These two factors suggest that PMF might prove valuable for industrial application in the short term.

5 HORIZONTAL SECTIONS

Our current implementation does not address fully the special difficulties posed by horizontal edge segments, a characteristic shared with all other stereo algorithms we know, many of which simply ignore horizontal sections altogether. Such segments have the dual problems of being difficult to locate (their locations can easily migrate onto an adjacent raster/epipolar line) and of being intrinsically ambiguous (points along a horizontal edge segment in one image can potentially match all points along the corresponding edge segment in the other image). The first of these problems will be met in future implementations of PMF by the inclusion of a small two-dimensional search window for near-horizontal sections. As for the second, it is clear that for long segments it will be difficult or impossible to identify matches correctly using only information based upon disparity gradients and hence some later stage of interpolation seems mandatory. Note, however, that horizontal segments can still provide figural similarity information for other matches if not always for themselves, which means that they should not be excluded from consideration by the early stages of PMF.

6 PERFORMANCE EXAMPLES

No common basis exists in the literature for assessing the performance of stereo algorithms in terms of either the images/scenes to be used or measures of performance for them. This deficiency represents a severe problem in evaluating different algorithms. Here we report some examples of the quantitative performance of PMF on a series of artificial stereograms and we show some qualitative results for a variety of natural scene stereograms which are not amenable to quantitative assessment at the present time.

For the artificial images, where the points to be matched can be specified accurately, we report the percentages of points (i) matched correctly, (ii) matched incorrectly, and (iii) left unmatched. We provide these measures for a range of surfaces chosen to sample the hardest problems typically found in many natural scenes, such as steep slopes, shears, and transparencies.

In the case of natural images, there are difficulties in providing quantitative measures because there is no simple definition of what constitutes a point to be matched. In most tests of stereo algorithms on natural scenes the points used are the outputs of an edge operator of some kind, which poses problems about whether it is the operator or the stereo algorithm which is deficient. In common with most other papers in the field, therefore, we simply present illustrations of PMF's treatment of natural images without attempting any quantitative assessment, using for this purpose primitives provided by a high-spatial-frequency Marr-Hildreth edge operator.

For the representative cross-section of performance examples presented in this section (both natural and artificial), PMF was run with the set of parameter values for allowable matching range, neighbourhood support zone, etc., described earlier. Alterations in any or all of these produced little appreciable change in the performance of PMF as long as extremes were avoided. Furthermore, the performance of PMF was largely unaffected by the addition of even quite large quantities of noise (eg the addition of more than 30% extra unmatched points to either or both halves of the artificial stereograms).

6.1 Artificial stereograms

The primitives of the pair of images that formed the artificial stereograms (figures 4-6) were positioned at the intersections of a 128 x 128 pixels square grid. Each grid point had the same probability (0.1) of carrying an image primitive, subject to the same number of points occurring on each line (n = 13). The ambiguity problem posed by these stereograms (as measured by the density of possible matches) was generally greater than that observed in the natural images (figures 7 and 8). This was partly because the densities of the primitives themselves were greater, and partly because the primitives were all identical so that it was not possible to restrict the set of initial matches by requiring left/right primitives for a match to be similar in contrast and orientation. For the various examples of artificial images shown below, each primitive had on average six possible matches in the other image.

6.1.1 Sloping surfaces As shown in section 2, stereo images of sloping surfaces will satisfy a disparity gradient limit of 1.0 provided the slope they depict is not too great (and assuming camera geometry typical of human stereo vision). Furthermore, even if the slope in one direction violates the disparity gradient limit it need not necessarily be violated in all other directions, and hence it is still possible that PMF may be successful. This needs to be borne in mind when considering the fact that PMF often does better than the limit would seem on the face of it to allow.

The performance of PMF for slopes was examined on surfaces of the kind shown in figure 4a, which depicts a surface whose vertical cross-section is a triangular wave. In this case the stereogram has a peak-to-trough separation of 20 pixels and a peak-to-trough disparity difference of 10 pixels; hence the disparity gradient in the vertical direction is approximately 0.5 and in the horizontal direction zero. Measures of the performance of PMF for such surfaces were obtained for peak-to-trough disparities ranging from 4 pixels to 36 pixels with 2-pixel intervals (corresponding to a range in vertical component of disparity gradient of 0.2 to 1.8). The result of running PMF on 4a is given in 4b with intensity used to code disparity (with crossed-eye viewing darker points are closer). Over 99% of the disparity values recovered from this stereogram are correct with the few errors being almost equally divided between incorrectly matched and unmatched points. The percentage of points matched correctly is plotted in figure 4c against the size of the vertical component of the disparity gradient present over the whole range of such images.

As can be seen from the graph, the performance of PMF is excellent for disparity gradients up to about 1.0, always being able to identify an excess of 98% of the true matches. Beyond this imposed limit, performance tends to fall off with only about 50% of correct matches being obtained for a disparity gradient of 1.8. As explained above, the fact that PMF copes at all with disparity gradients larger than the limit of 1 is a result of the fact that the disparity gradients are not as large in other directions.

6.1.2 Depth discontinuities and lacy surfaces The depth profile of the stereogram in figure 5a consists of a vertical square wave. Hence the surface it depicts includes a number of depth discontinuities. Figures 5b and 5c show that PMF copes quite well at the corresponding disparity shears, although it can be seen that it fails to match a small number of the points correctly. Similar results were also obtained for vertically oriented square waves, with the exception that some points close to the disparity shears are obscured in one or other image and thus remain correctly unmatched.

The good level of performance achieved for depth discontinuities illustrates the fact that PMF is able to make use of the within-disparity-gradient-limit structure that lies to either side of the shear and at the same time ignore the fact that the limit is exceeded across it. The desirability of this characteristic is clear-cut (see earlier remarks in section 4 regarding the objective of not penalising perturbations lying within the range to be expected in the stereo projections of interest). The underlying reason why PMF achieves this goal is that it takes advantage of such within-disparity-gradient-limit support as exists; it does not impose a cost if neighbouring primitives fall outside the limit (cf Prazdny 1985).

The same property is apparent in the way in which PMF is able to deal with the lace surface portrayed in figure 6. Here the within-limit neighbourhoods that facilitate disambiguation are actually superimposed on top of each other. Whilst it is clear that the performance of PMF is considerably worse for this lace surface, in that only two thirds of the points in each depth plane are matched correctly, the results are qualitatively in keeping with human performance (comparable quantitative data not being available for human vision)[2].

6.2 Natural images

Figures 7 and 8 portray the stages of stereo processing for two very different natural image pairs, each of which is 128 x 128 pixels square. Zero crossings serving as edge-like primitives were extracted from each image with the use of a single high-frequency (w = 4 pixels) Marr-Hildreth operator (Marr and Hildreth 1980). In the figures the resulting edge primitives are portrayed with a grey level proportional to their contrast strengths (darker primitives have a greater absolute contrast value). The orientation and the contrast polarity associated with a primitive were used to restrict the initial set of matches.

---

[2] This issue is considered in more detail in Pollard and Frisby (1990) Transparency and the Uniqueness Constraint in Human and Computer Stereo Vision, *Nature* (in press).

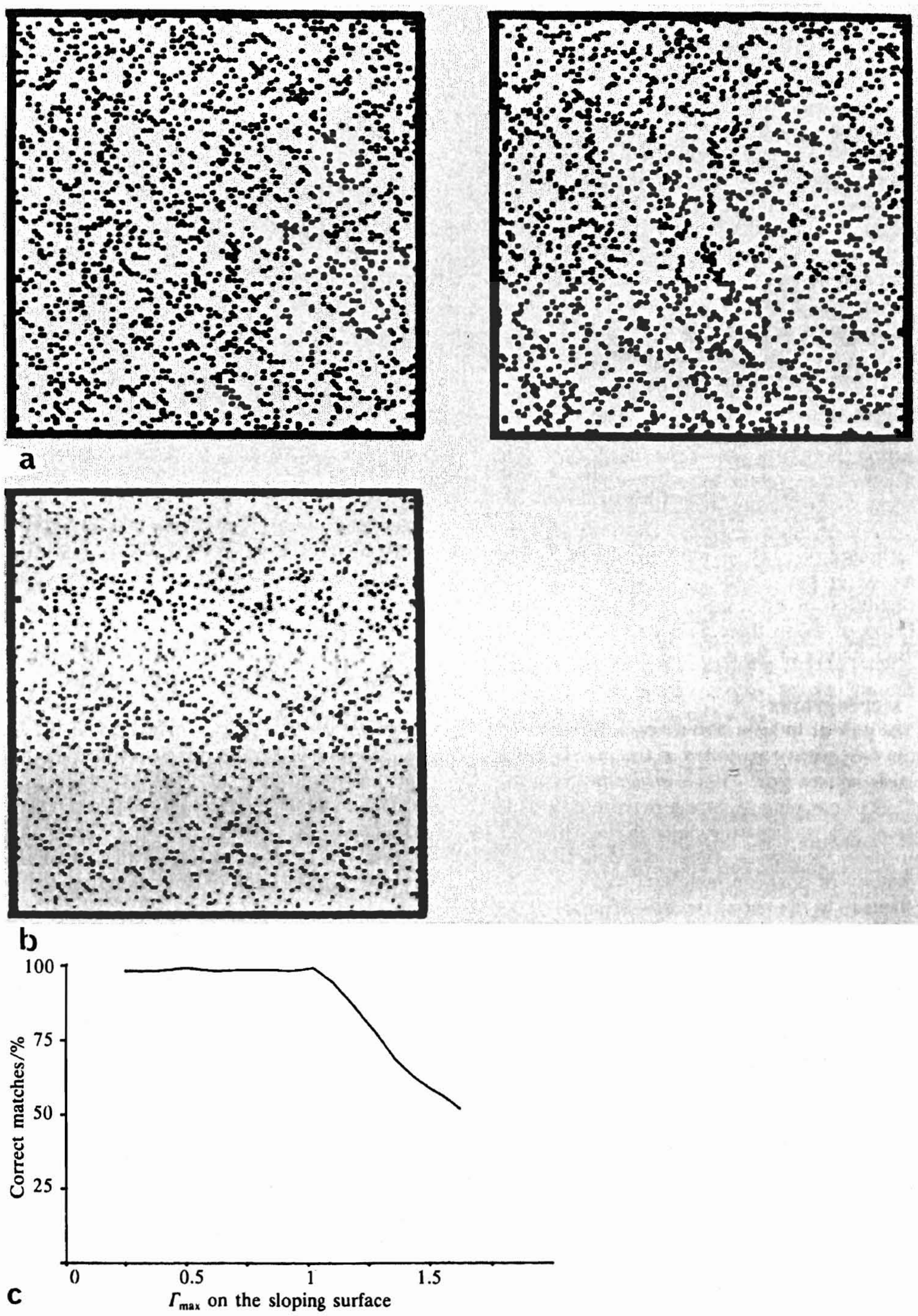

Figure 4 *Random-dot stereograms used to test the capacity of PMF to deal with sloping faces. (a) An example of a stereogram portraying a triangular waveform oriented horizontally, with peak-to-trough disparity amplitude equal to 10 pixels (each stereo half comprised of 128 x 128 pixels, of which about 10% are shown as black dots) and with maximum disparity gradient,* Γ_{max}, *on each sloping face equal to 0.5. (b) The result of applying PMF to (a), 99% of correct matches were found by PMF and these are shown with their disparities coded by intensity [darker points shown as the closest when (a) is viewed with cross-eye fusion]. (c) The correctly located matches recovered by PMF for triangular waveforms with varying* $\Gamma_{ma c}$. *See text for further details.*

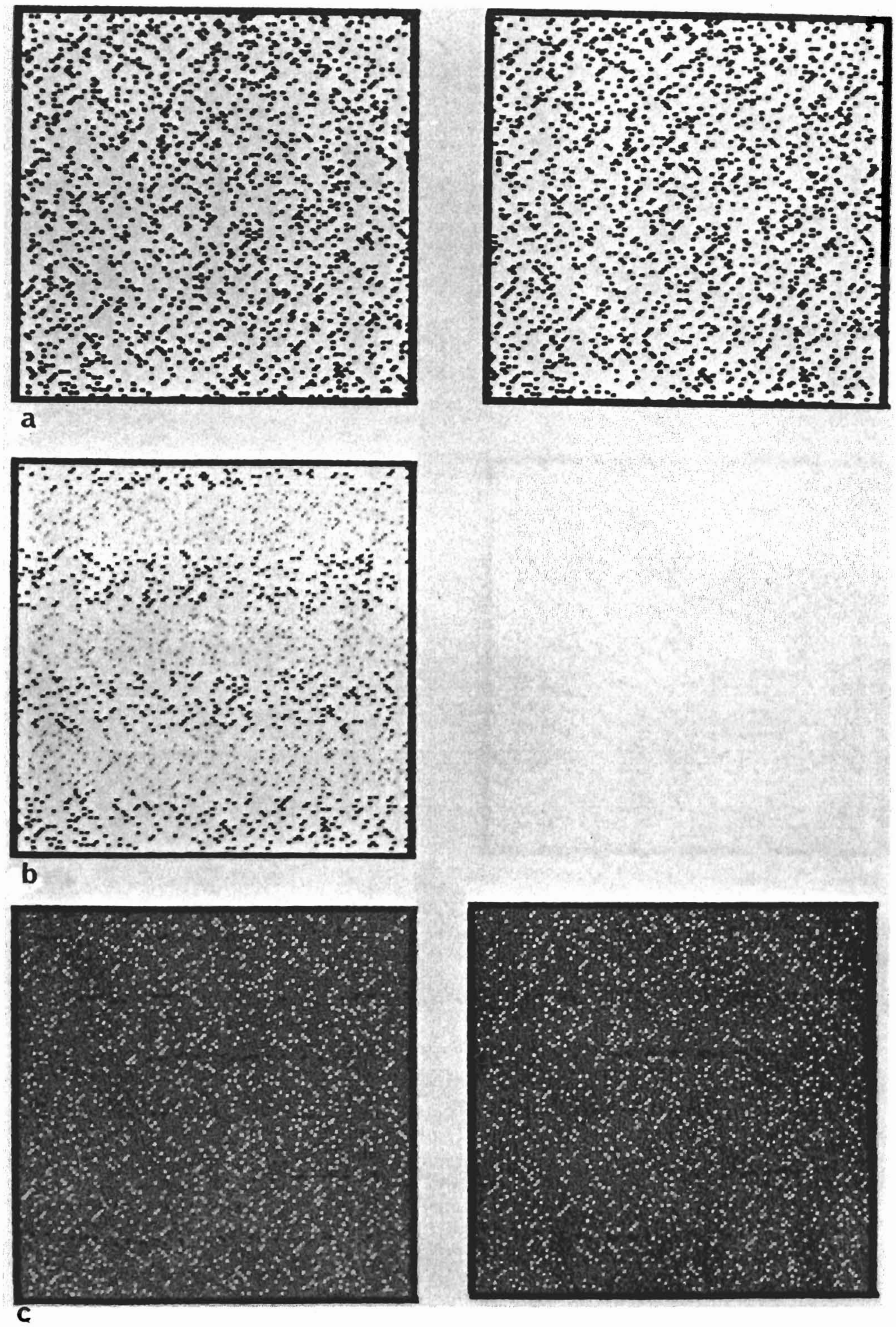

Figure 5 *Random dot stereogram used to test the capacity of PMF to deal with depth shears. (a) Stereogram portraying a square wave orientated horizontally and with peak-to-trough amplitude of 5 pixels. (b) Over 95% of points are matched correctly, with 4.5 matched incorrectly and a further 0.5% being left incorrectly unmatched. (c) A reconstruction of the stereogram displayed in (a) after processing with PMF. Dots matched correctly are shown white (hence the mid-grey back-ground), and the remainder are shown black. It can be seen that the performance of PMF is in general good with a few incorrect or missing matches confined to some parts of the borders of the shears. See text for further details.*

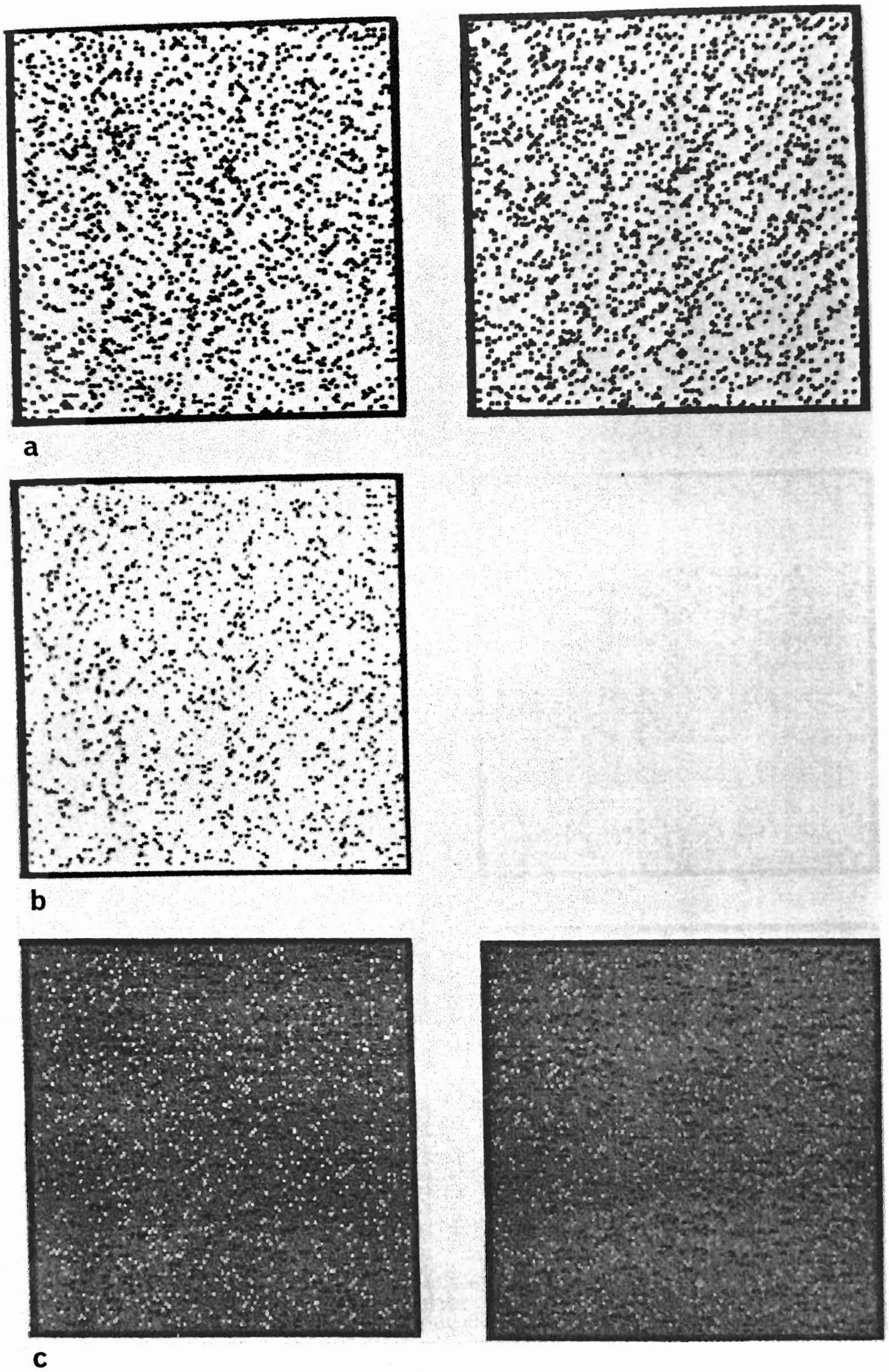

Figure 6 *Random-dot stereogram used to test the capacity of PMF to deal with superimposed transparent surfaces. (a) Stereogram portraying two superimposed planar surfaces with a disparity difference of 5 pixels. (b) The correct matches (64%) located by PMF for this stereogram. Of the remainder, 35.5% were matched incorrectly and 0.5% were left incorrectly unmatched. (c) A reconstruction of the stereogram displayed in (a) after processing with PMF. Dots matched correctly are shown white and the remainder are shown black. See text for further details.*

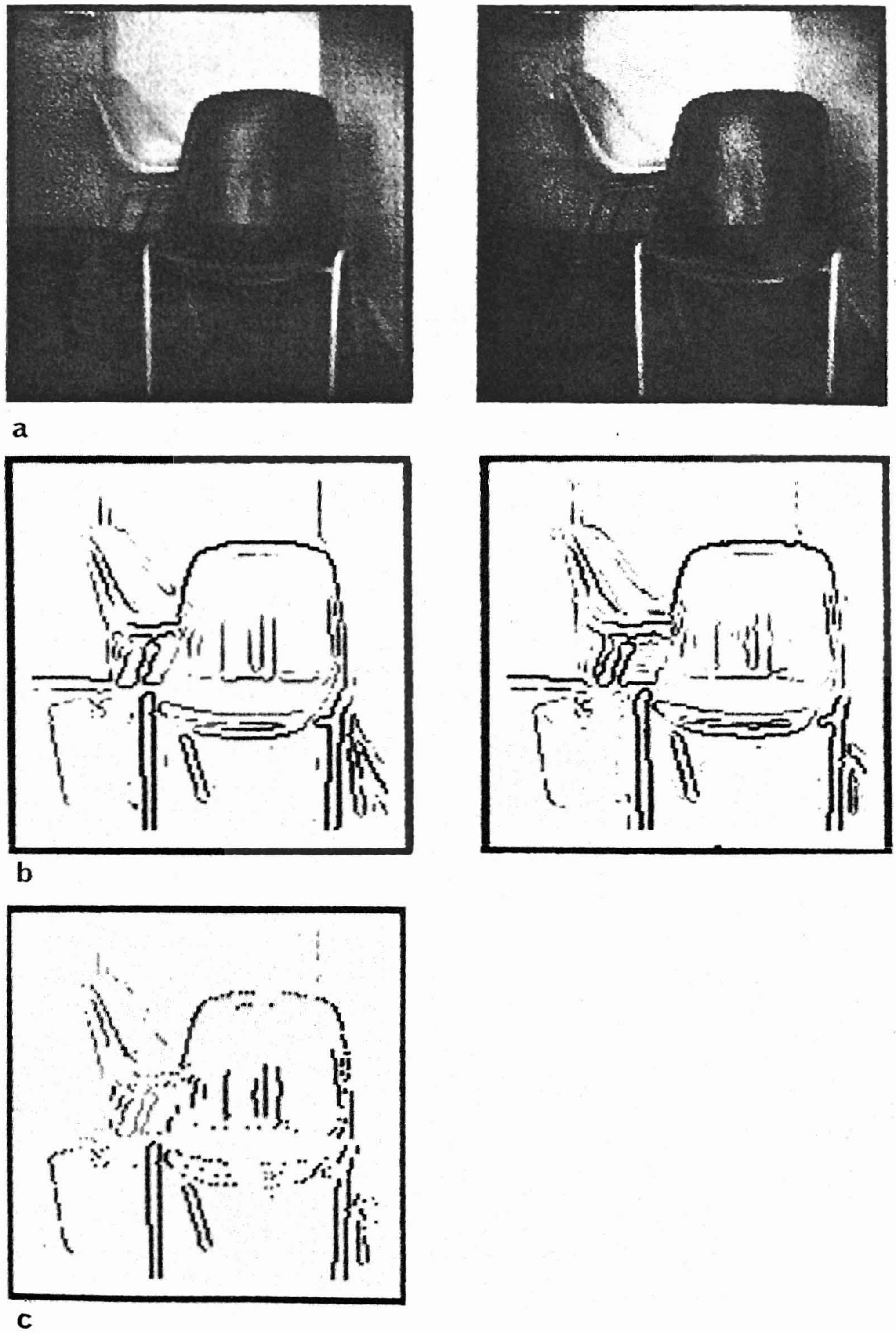

Figure 7 The performance of PMF on a stereo pair comprised of natural images of an office scene. (a) Stereo pair arranged suitably for cross-eyed fusion. (b) Edge-like primitives extracted with the use of a Marr-Hildreth operator. (c) Matches found by PMF, with intensity used to code relative disparities (dark = near, faint = far). See text for details.

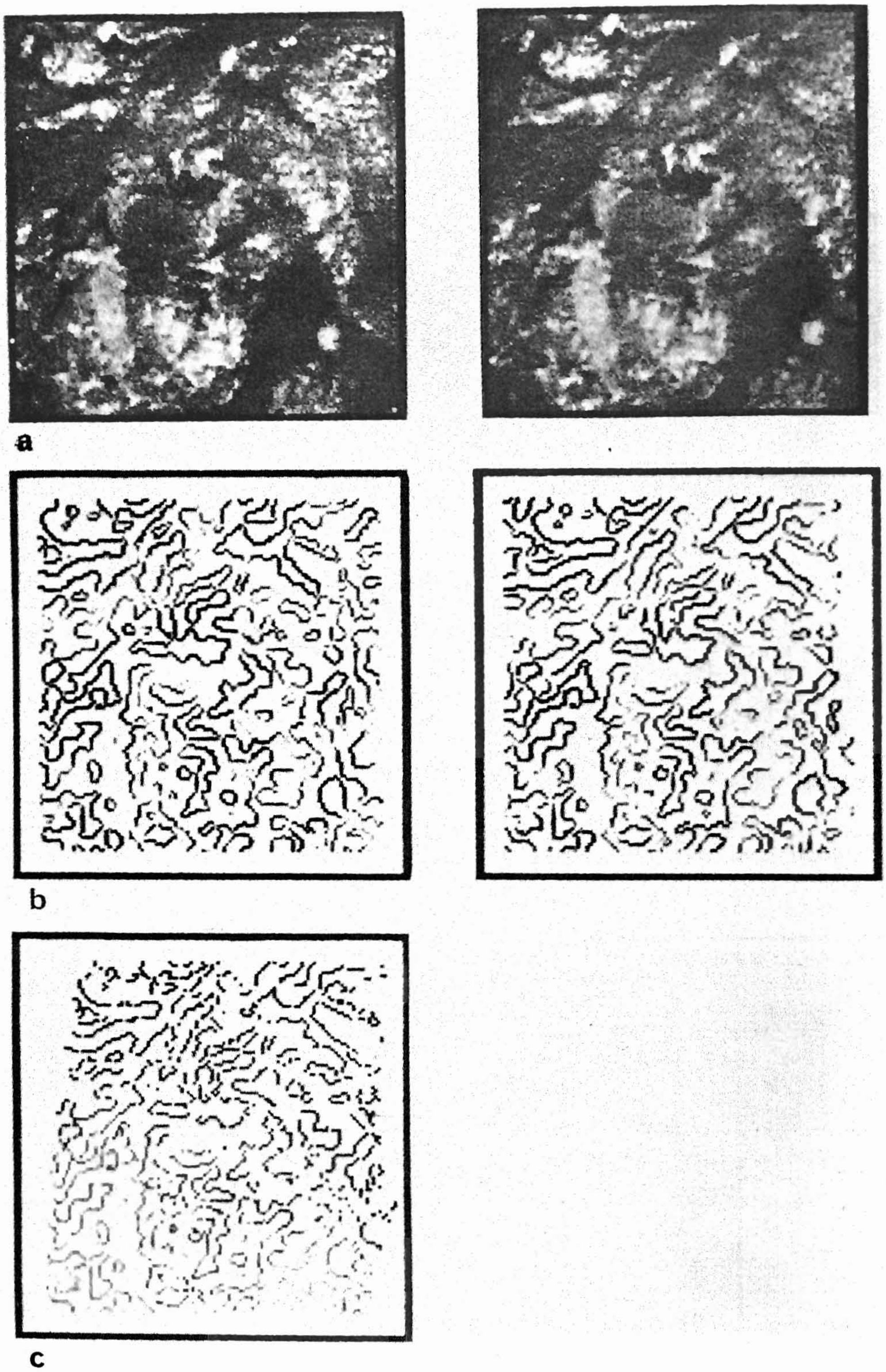

Figure 8 A further example of the performance of PMF on natural images, this time a rocky terrain viewed from above. Details as for figure 7.

The disparity gradient limit was again used to determine the allowable differences in orientation: a matched edge segment was not allowed between left and right zero crossings whose orientations were so different as to imply that the disparity gradient existing along a binocular line formed from those orientations would exceed the disparity gradient limit.

The 'goodness' of each match was given by the contrast strength of the weakest of its two constituent edges rather than as a function of the similarity of their orientation and contrast values. A goodness measure of the latter kind was thought to be inappropriate, given that changes in orientation and contrast do occur between the stereo projections with which we are concerned. Using a simple contrast measure has the advantage that it gives preference to the more reliable data, given that strength of an edge (its absolute contrast value) is directly related to its reliability (whether or not it is likely to be noise).

The disparity information (approximating for present purposes to relative depth) recovered from the matches identified as correct by PMF is displayed in the form of a cyclopean edge image in which disparity is encoded by edge intensity (darker points are closer to the imaging device). These points that remained unmatched by PMF are not displayed in these images.

Figure 7 is of a simple office scene, the edges recovered from which are quite sparse in comparison to those encountered in the artificial stereograms above. Nevertheless, despite the fairly small neighbourhood window exploited in PMF, it is able to cope quite easily with this scene. Figure 8 involves a natural scene that is more like those encountered in the artificial stereograms as it has a dense distribution of image primitives. It portrays a rocky terrain viewed from above, presenting therefore a similar task to that solved by human operators in the field of photogrammetry from aerial photographs. For both of these stereograms it can be seen that for the most part the resulting disparity output is at least qualitatively correct, the main exception being the horizontal sections located in the office scene (see remarks in section 5).

The accuracy to which depths can be measured in natural scenes is directly dependent upon the accuracy to which edges can be located in their images. In the example above, edges are only located to the nearest pixel, but it would be easy to locate the majority of strong edges to a tenth or even a hundredth of a pixel. The use of more accurately located image primitives would not interfere with the robustness of PMF.

7 HOW PMF RELATES TO OTHER THEORIES OF STEREOPSIS

As we have already stated, the theoretical justification for PMF is the fact that, for the stereo projections for which it is designed, the majority of disparity gradients that exist between correct matches will be less than 1. In this section we discuss how this constraint relates to others that have been considered in the literature on stereo algorithms. We do not, however, give a full review of the field as the theoretical aspects of many stereo algorithms have little in common with PMF.

7.1 The constraint of surface continuity

As outlined in the introduction, Marr and Poggio (1976) chose to base their theory of stereo disambiguation upon the observation that "matter is cohesive, it is separated into [reasonably large] objects", from which they derived their constraint of local surface smoothness. Whatever they might want to imply in detail by the notion of 'surface smoothness', it is clear that the way a disparity gradient limit is imposed in PMF explicitly allows many jagged (ie 'non-smooth') surfaces. An illustration of the 'degree of jaggedness' allowed by a disparity gradient limit of 1 is given by the random-dot stereogram portrayed in figure 9. At no point do any disparity gradients in this stereogram exceed 1 and yet the perceived surface is not obviously well described as smooth. This surface satisfies the Lipschitz condition discussed above and, as can be seen from the figure, the performance of PMF on it is extremely good. An example of the performance of PMF on a stereogram which depicts a surface that does not satisfy the Lipschitz condition is given in figure 10. The performance for this stereogram is rather better than may be expected for a surface that is so far from smooth, providing a further illustration that the Lipschitz condition is a rather conservative estimate of the class of surfaces suitable for PMF.

This suggestion that the use of a disparity gradient by PMF is not usefully characterised as imposing a surface smoothness constraint is further supported by the observation that there exist perfectly smooth surfaces that violate the disparity gradient limit of PMF. Even a planar surface can provide many violations of this limit if it recedes sufficiently rapidly. This is therefore another reason for distinguishing between Marr and Poggio's surface smoothness constraint and that developed here in terms of surfaces needing (conservatively) to meet a Lipschitz criterion.

7.2 The ordering constraint

It has been pointed out that stereo projection almost always preserves the order of primitives extracted from the two images along matching epipolar lines (Baker 1982; Mayhew 1983). The underlying reason for this is that it is geometrically impossible for points arising from the same opaque surface to be differently ordered in the two images (different orderings can come about only for scenes comprised of surfaces of small extent such as isolated small blobs or thin wires). Hence order has been exploited explicitly to constrain the selection of matches in a number of stereo algorithms (eg Baker and Binford 1981; Arnold 1982; Ohta and Kanade 1983).

It has been frequently noted that order reversals correspond to disparity gradients of magnitude greater than 2 (eg Burt and Julesz 1980; Mayhew 1983). Hence a limit of 1 will prevent matches that violate the ordering constraint from mutually supporting each other and to that extent a limit of 1 can be said to be a conservative implementation of the ordering constraint. However, we think there are some difficulties with that notion, particularly if it is used to justify the design of PMF.

First, the ordering constraint and its associated disparity gradient limit of 2 are concerned only with events along epipolar lines. The disparity gradient limit between epipolars can be in principle infinite even for opaque surfaces. PMF uses its limit isotropically, and justification for this cannot logically be sought just in terms of the physical limits that are possible along epipolars.

Second, we have found that a limit of 2 is not sufficiently restrictive for disambiguating purposes. Hence some other justification for a lower limit is required, even if attention is directed only to disparity gradients between points on the epipolars themselves. It may be helpful to point out in this context that a limit of 1 prevents the separation of a pair of points along an epipolar line in one image from being greater than three times the separation of a corresponding pair in the

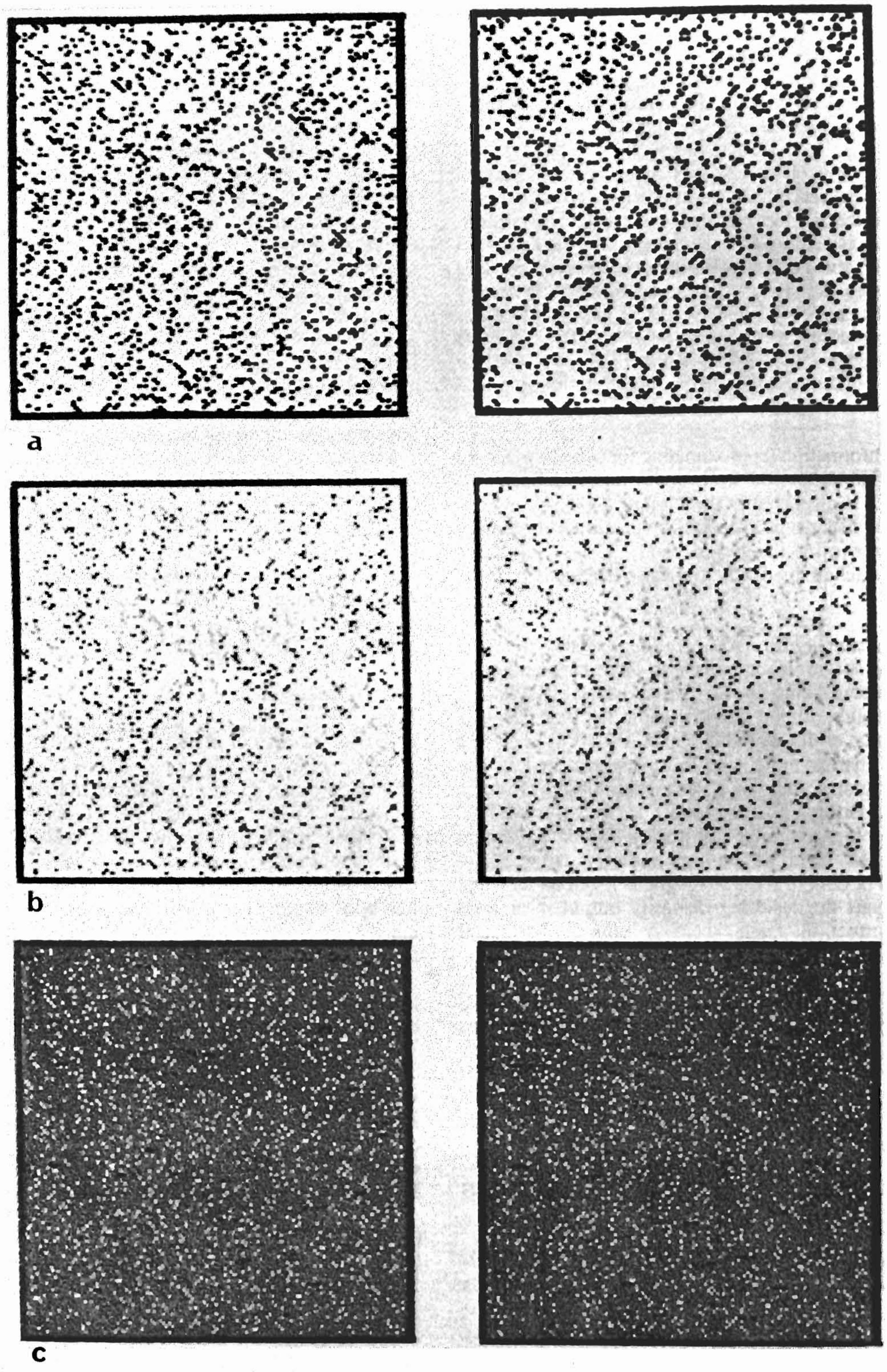

Figure 9 *Random-dot stereogram used to test the capacity of PMF to deal with jagged surfaces. (a) A stereogram in which the disparities were randomly selected from a range of 6 pixels except that they everywhere had to satisfy a disparity gradient limit of 1. (b) Dots matched by PMF are shown on the right-hand side with intensity used to code relative disparity (again, dark = near, faint = far, for the cross-eyed fusion of the original stereogram). For comparison the correct relative disparities are shown on the left-hand side with the same intensity code being used. The similarity between these figures brings out the fact that PMF found the vast majority of correct matches (93%). To help further in this regard, the positions of the dots in these two images have been adjusted such that when the two are fused as though there were a stereo pair, the correct matches are seen to lie in a plane, with the very few incorrect matches seen as dots lying outside the plane occupied by the majority. (c) A reconstruction of the stereogram displayed in (a) after its processing by PMF. Dots matched correctly are shown white and the remainder are shown black. See text for further details.*

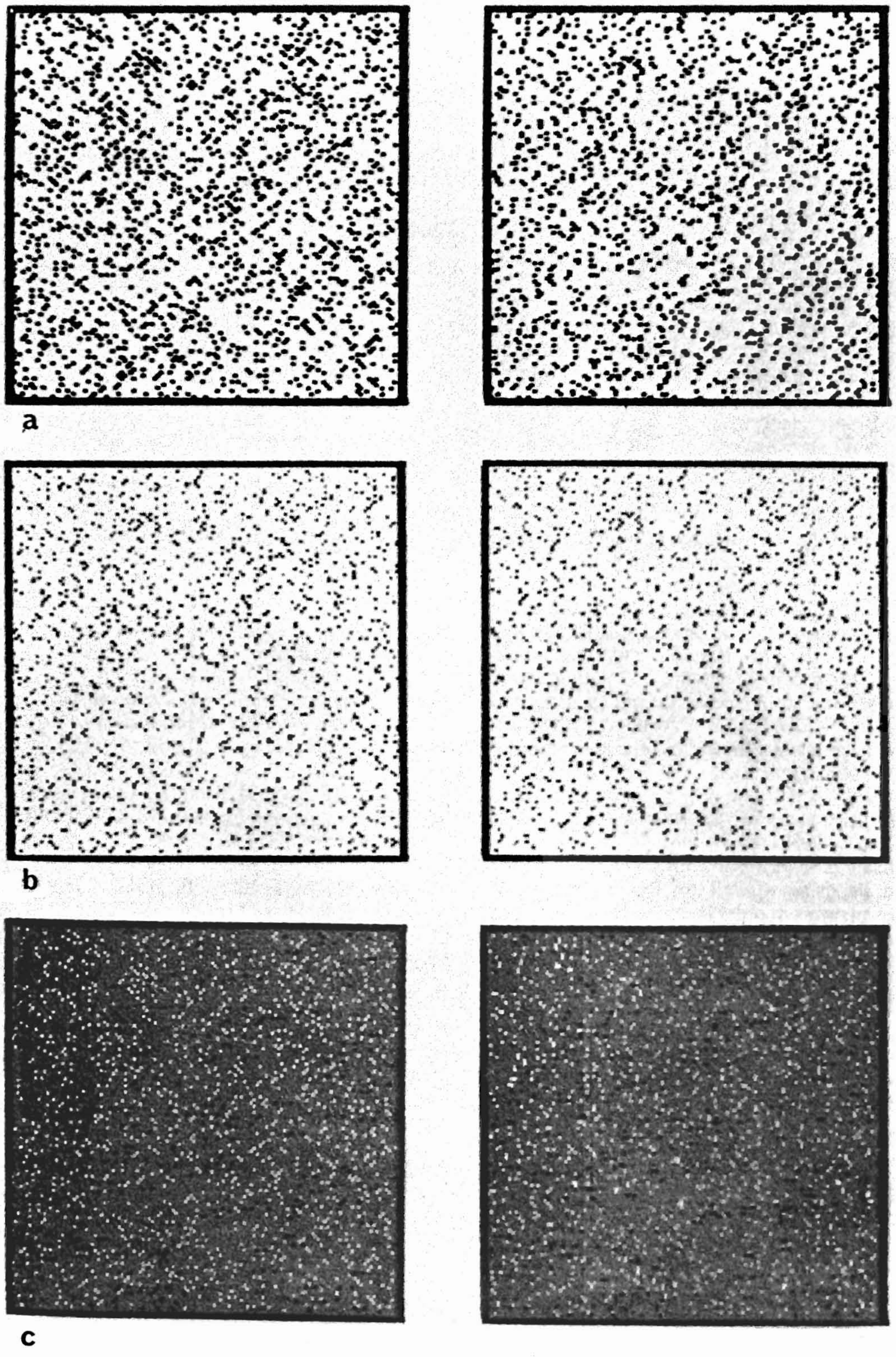

Figure 10 *A second random-dot stereogram used to test the capacity of PMF to deal with jagged surfaces. (a) A stereogram made by selecting disparities at random from a gaussian distribution of zero mean disparity and standard deviation 2 pixels with no limit on allowable disparity gradients. Hence there are many severe local variations in disparity. Despite this, PMF manages to match correctly over 79% of dots; (b) and (c) display this competence on using the method of figures 9b and 9c. The text gives further details.*

other; a gradient limit of 2 on the other hand allows the separation in one image to be infinitely larger than that in the other.

7.3 The constraint of figural continuity

Our previous stereo algorithm (STEREOEDGE: Mayhew and Frisby 1980, 1981) was based on taking advantage of the figural continuity that existed between the two images. Edge-like primitives were matched on the basis of their continuity and the similarity of their geometrical structure. That algorithm was justified in terms of the surface continuity constraint, and we continue to believe that that justification was appropriate for that algorithm. When surfaces do happen to be locally spatially continuous, then their edges and extended surface markings must be spatially continuous also: hence the rationale for the stereo combination rule of STEREOEDGE of preferring disparity matches which preserve figural continuity (see also Grimson 1984).

As we have already discussed in section 2, it is clear that a disparity gradient limit also provides a limit upon the allowable geometrical deformation that exists locally between the two images. For instance, if the disparity gradient limit were in fact zero, then the two images would have to be locally figurally identical. Imposing a limit of 1 allows some dissimilarity in the figural structures that constitute the two images. Hence, PMF can be viewed as effecting a correlation between stereo halves in which the figural distortions that characterise the differences between left and right stereo projections are not allowed to lower the correlation score. This way of considering PMF shows that it is a natural development of the STEREOEDGE algorithm. However, PMF has the advantage that its matching primitives are not restricted to extended edge segments.

8 SUMMARY

We have demonstrated the considerable disambiguating power that results from imposing a limit on the magnitude of allowable disparity gradients in order that matches are allowed to support one another. Imposing such a limit does not greatly limit the class of allowable surfaces. We have developed a simple stereo algorithm that utilises disparity gradients in a cooperative manner (Fender and Julesz 1967; Julesz 1971) to solve the stereo disambiguation problem and have illustrated its performance on a variety of natural and artificial stereograms. The rationale underlying the use of the disparity gradient limit has been contrasted with other computational treatments of the stereo correspondence problem.

Note added in proof Since the submission of this paper further results have been obtained regarding the mathematical details of the imposition of a disparity gradient limit (see Trivedi and Lloyd 1985; Pollard et al 1985).

Acknowledgements We would like to thank John Porrill and Tony Pridmore for their valuable advice and Chris Brown for his technical assistance. PMF was first invented on SERC Research Grant GR/B.43265 and was the subject of further development by S B Pollard supported on a SERC CASE studentship jointly supervised by Dr M M McCabe of GEC/Hirst Laboratories and Dr S A Lloyd. JEWM is supported by Alvey/SERC contract GR/D/1679.6.

REFERENCES

Arnold R D, (1982) *Automated Stereo Perception* PhD thesis, Stanford University, Stanford, CA, USA

Baker H H, (1982) *Depth from edge and intensity based stereo*, AI Memo 347, Stanford University, Stanford, CA, USA

Baker H H, Binford T O, (1981) Depth from edge and intensity based stereo in *Proceedings of 7th IJCA* (Los Altos, CA: William Kaufmann) 631-636

Burt P, Julesz B, (1980) Modifications of the classical notion of Panum's fusional area *Perception* **9** 671-682

Fender D, Julesz B, (1967) Extensions of Panum's fusional area in binocularly stabilised vision *Journal of the Optical Society of America* **57** 819-830

Grimson W E L, (1984) Computational experiments with a feature based stereo algorithm, AI Memo 762, MIT, Cambridge, MA, USA

Julesz B, (1971) *Foundations of Cyclopean Perception* (Chicago: University of Chicago Press)

Marr D, (1982) *Vision* (San Francisco: W H Freeman)

Marr D, Hildreth E, (1980) Theory of edge detection *Proceedings of the Royal Society of London, Series B*, **207** 187-217

Marr D, Poggio T, (1976) A cooperative computation of stereo-disparity *Science* **194** 283-287

Marr D, Poggio T, (1979) A theory of human stereo vision *Proceedings of the Royal Society of London, Series B*, **204** 301-328

Mayhew J E W, (1983) Stereopsis in *Physical and Biological Processing Images* eds O J Braddick, A C Sleigh (New York: Springer) 204-216

Mayhew J E W, Frisby J P, (1980) The computation of binocular edges *Perception* **9** 69-86

Mayhew J E W, Frisby J P, (1981) Psychophysical and computational studies towards a theory of human stereopsis *Artificial Intelligence* **17** 349-386

Ohta Y, Kanade T, (1983) Stereo by intra- and inter-scanline search using dynamic programming Technical Report CMU-CS-83-162, Carnegie Mellon University, Pittsburgh, PA, USA

Pollard S B, Porrill J, Mayhew J E W, Frisby J P, (1985) Disparity gradient, Lipschitz continuity and computing binocular correspondences in *Proceedings of the Third International Symposium of Robotics Research, Gouviex, France* (Cambridge, MA: MIT Press) in press

Prazdny K, (1985) Detection of binocular disparities *Biological Cybernetics* (in press)

Rosenfeld A, Hummel R, Zucker S W, (1976) Scene labelling by relaxation operations *IEEE Transactions on Systems, Man, and Cybernetics* **6** 420-433

Trivedi H P, Lloyd S A, (1985) The role of disparity gradient in stereo vision *Perception* in press

Tyler C W, (1973) Stereoscopic vision: cortical limitations and a disparity scaling effect *Science* **181** 276-278