[2] Disparity Gradient, Lipschitz Continuity, and Computing Binocular Correspondences

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

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

Reprinted, with permission of MIT Press, from Faugeras, O D and Giralt, G (Eds) *Robotics Research: The Third International Symposium*, 1986, 19-26.

Abstract

A theoretical formulation of the stereo correspondence problem and an algorithm for its solution are described. One way of characterising the stereo matching problem is to find a one-to-one locally continuous mapping between the two eyes' subject to the epipolar geometry. This abstract formulation unfortunately provides no algorithmic *mechanism* with which to obtain the solution. This has long been recognised and researchers in the field have been concerned to identify and exploit heuristic continuity and smoothness constraints to resolve the stereo ambiguity problem (for a review see Mayhew 1983). One such constraint, motivated by psychophysical observation (Burt and Julsez 1980), is the disparity gradient constraint which was first exploited heuristically in the PMF stereo algorithm (Pollard et al 1985). The disparity gradient constraint enforces Lipschitz continuity on the mappings between the eyes' views, on the surfaces in the scene and on the depth map. The Lipschitz constant (corresponding to the limiting disparity gradient) provides a free parameter that can be exploited algorithmically. Accordingly we also examine how varying the value of the disparity gradient limit effects both the disambiguating power of the PMF algorithm, and restricts the range of surface orientations that can be fused (full details can be found in Pollard *et al* 1985; Pollard 1985).

1. Introduction

Suppose we have a stereo pair of views of n points in space. The only purely geometrical constraint on matching between the two views is that matches should lie on corresponding epipolar lines. In principle this would be enough. For generic views of generic point sets no two points will lie on the same epipolar and so every point has a unique match. In practice we do not deal with generic point sets; quantisation error also forces points into the same epipolar *raster*. Further matching constraints are thus required for uniqueness.

Objects in the world are usually bounded by continuous opaque surfaces. If we assume that the observed feature points lie on such a surface so as to be simultaneously visible to both eye's we arrive at the *ordering constraint* (see: Baker and Binford 1981; Burt and Julesz 1981; Baker 1982; Mayhew 1983; Yuille and Poggio 1984; Ohta and Kanade 1985): points on the same epipolar line are in the same order in both eyes' views. Once again the matching problem is solved in principle. Points are matched in left-to-right order along epipolars, starting with the leftmost in each view. Again this is not sufficient in practice; the two eyes' views will not cover the same area, so the leftmost points will not both be available; also extra unmatchable noise points will be present. A further constraint is required. One possibility is to propagate information between rasters exploiting figural continuity (Mayhew and Frisby 1980, 1981) since continuous edges in the scene will project to continuous curves in the image (see also: Baker 1982; Grimson 1983, Ohta and Kanade 1985). Once we have matched one point on such a curve, we can follow it continuously across rasters (note that a practical definition of continuity would be required here).

Another between-rasters constraint that has proved useful is based on the concept of *disparity gradient limit* (derived from psychophysical observations: Burt and Julesz 1980a, 1980b; Tyler 1973, 1974, 1975). It is the basis of a successful feature based stereo matching algorithm (the PMF algorithm: Pollard *et. al.* 1984, 1985; Pollard 1985). Disparity gradient can be thought of as a simple measure of continuity, and the disparity gradient limit encapsulates previous stereo constraints in a strong form. For example, enforcing a disparity gradient limit of $DG<2$ requires that the observed surface be Lipschitz continuous (this is defined in the Appendix) and not self-occluding; a recent result by Trivedi & Lloyd (1985) shows that it also imposes continuity on the mapping between the two eyes' views. This can be considered as an algorithmic description of the continuous and opaque nature of most surfaces in the world. Some simple transparent self-occluding surfaces (*e.g.* a pair of transparent planes at different depths), though not satisfying the disparity gradient limit globally, can be built up from surfaces patches which do satisfy this limit locally (this is similar to the concept of a *cohesiveness* discussed by Prazdny 1985). The set of such surfaces forms a wider domain in which disparity gradient limit is still a useful tool for disambiguation, and in fact the PMF algorithm can cope with such *lace curtain* stereograms (see Pollard *et al* 1985).

2. Disparity Gradients and Lipschitz Continuity

In this section we will show that an isotropic disparity gradient limit is only one member of a whole family of measures of continuity which impose scene-to-view and view-to-view Lipschitz continuity. In the process we obtain a simplified proof of the result of Trivedi & Lloyd (1985) on view-to-view continuity. In §3 the PMF stereo algorithm (Pollard *et al* 1985) which exploits the disparity gradient constraint will be briefly described.

2.1. Properties of the Disparity Gradient Limit Constraint

Let L and R be the left and right views of a given scene (they could be regions in the image planes, or finite sets of feature points). If points $p \in L$ and $p' \in R$ have been matched we write $p \to p'$ (this match considered as an object in itself will be denoted in later sections by $M_{pp'}$). The matching may be many-to-many i.e. many points on the left could be matched to the same point on the right and *vice versa*. Suppose we have a pair of matches $p \to p'$ and $q \to q'$. Their disparity gradient is defined to be

$$DG = \frac{\textit{difference in disparities}}{\textit{cyclopean separation}}$$

We also use the notation $K=DG/2$ and when we wish to refer explicitly to the pair of matches we will write $DG(M_{pp'}, M_{qq'})$ for DG. Since the cyclopean image points are at $(p+p')/2$ and $(q+q')/2$ and the associated disparity vectors are $(p'-p)$ and $(q'-q)$

$$\frac{DG}{2} = K = \frac{\|(p'-q') - (p-q)\|}{\|(p'-q') + (p-q)\|}$$

where $\|\cdot\|$ denotes the vector norm.

Proposition 1: Suppose the matching $L \to R$ satisfies the DG-limit constraint with $DG<2$ ($K<1$), that is, for all pairs of matches $p \to p'$, $q \to q'$ we have the inequality

$$\|(p'-q') - (p-q)\| \le K \, \|(p'-q') + (p-q)\|$$

Then the matching $L \to R$ is one-to-one.

Proof

Suppose $p \to p'$ and $q \to p'$. Substituting into the disparity gradient constraint gives

$$\|p - q\| \le K \, \|p - q\|$$

and since $K<1$ this is only possible if $p=q$. Similarly if $p \to p'$ and $p \to q'$ then $p'=q'$.

In the case when the matching is one-to-one, if we restrict L and R to contain only matched points, then we can define a one-to-one and onto map $f{:}L \to R$ by $f(p)=p'$ if p matches with p'. In this case we can prove the Trivedi & Lloyd (1985) result for the map f.

Proposition 2: Under the assumptions of Proposition 1 the map f is known to exist. This map and inverse f^{-1}. are then Lipschitz of order one with Lipschitz constant

$$\frac{(1+K)}{(1-K)}$$

If L is an open set in the image plane (rather than just a finite set of points) then the map f is a homeomorphism from L to R.

Proof

Using the inequalities

$$\|a\| - \|b\| \le \|a - b\| \qquad \|a + b\| \le \|a\| + \|b\|$$

the disparity gradient constraint gives

$$\|p'-q'\| - \|p-q\| \le K\|p'-q'\| + K\|p-q\|$$

and since $K<1$ we can simplify to get the Lipschitz constraint

$$\|f(p) - f(q)\| = \|p' - q'\| \le \frac{1+K}{1-K}\|p - q\|$$

The symmetry of the disparity gradient constraint immediately gives the same result for the inverse mapping f^{-1}. Continuity of f and its inverse follow automatically when L is an open set, and then $f{:}L \to R$ must be a homeomorphism (see appendix for details).

Neither of the above Propositions uses any stereo geometry. However to relate the results to properties of the scene we need to consider this geometry. To simplify the working we will prove the next proposition only within the framework of small-angle stereo geometry. Let (x,y,z) be a point in space, and p, p' its two views. If d is the distance to a fixation point on the forward direction of the cyclopean eye and I is the interocular distance then the disparity vector is

$$p' - p = \begin{bmatrix} \delta \\ 0 \end{bmatrix}$$

and we have the relationships

$$z = d + \frac{d^2}{I}\delta \qquad \begin{bmatrix} x \\ y \end{bmatrix} = d\,\frac{(p+p')}{2}$$

valid for small interocular distance and for points close to the fixation point. The left-right matching process generates four maps of interest. The two view-to-scene maps $p \to (x,y,z)$ and $p' \to (x,y,z)$, the cyclopean map $(p+p')/2 \to (x,y,z)$ and the depth map $(x,y) \to z$. All these can be shown to be Lipschitz with appropriate constants. We will state a result only for the depth map.

Proposition 3: Suppose the matching f satisfies the disparity gradient limit, then the map $(x,y) \to z$ is Lipschitz order one with constant $DG \cdot d/I$. In particular, if the scene is a surface, it is continuous.

Proof

By substitution the positions and depths of a pair of points satisfy

$$\frac{|z_2 - z_1|}{\left\| \begin{bmatrix} x_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\|} = \frac{2d}{I}\,\frac{|\delta_2 - \delta_1|}{\|(p_2-p_1) + (p'_2-p'_1)\|} \le \frac{d}{I}\cdot DG$$

which is the Lipschitz condition. It can be considered as a relationship between disparity gradients in the image and world gradients. Given a point $\mathbf{x}_1$ we call the set of points $\mathbf{x}_2$ such that the associated pair of matches do not satisfy the disparity gradient limit the excluded region of $\mathbf{x}_1$. If $\mathbf{x}_1$ is the fixation point $(d,0,0)$ the inequality above shows that the excluded region is the set of points interior to the right circular cone

$$(z-d)^2 = \frac{d}{I}\cdot DG\,(x^2 + y^2)$$

Because this cone is rotationally symmetric about the z-axis we say that the disparity gradient limit constraint is isotropic. The condition $DG<2$ is necessary to ensure that the surface of this cone is simultaneously visible to both both eye's. This is the opacity constraint. However the opacity constraint in itself does not impose any between rasters continuity, so we would expect the isotropy of the $DG<2$ constraint to be *sufficient but not necessary* for the truth of the propositions above. In the next section we shall see that this is in fact the case.

Another point is worth noting briefly here. Suppose we observe a cone with axis through our *cyclopean* eye, on which the generators formed by intersection with a vertical plane are drawn. Then points on these generators will have different left-to-right orders in each eye's view. Thus the geometry of binocular vision does not impose any off epipolar ordering constraint as has sometimes been assumed (Yuille and Poggio 1984).

2.2. Generalised Disparity Gradient Measures

The proofs of Propositions 1 and 2 did not use any properties other than the fact that the distance function $\|\cdot\|$ was a norm (see Appendix). The results would thus be true for any other norm defined in the image planes. The proof of Proposition 3 did require that distance along epipolar lines be the usual distance measure (since disparity in this direction is related directly to depth). Thus if we can find a norm $\|\cdot\|$ with the property

$$\left\| \begin{pmatrix} x \\ 0 \end{pmatrix} \right\| = |x|$$

then, using the disparity gradient definition of the last section with this new norm, the proofs of all the propositions will be unchanged. There are many such norms. We mention here three families:

$$\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|_\infty = \max \left\{ |x|,\, a|y| \right\}$$

$$\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|_2 = \sqrt{x^2 + a^2 y^2}$$

$$\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|_1 = \|x\| + a\|y\|$$

(proofs of the norm property can be found in any analysis textbook). The constant a allows us to vary the *between rasters* strength of the constraint. This is easiest to see if we consider the boundary of the excluded region of the fixation point. It is a cone whose cross-section is, for the norm $\|\cdot\|_\infty$ a flattened diamond, for $\|\cdot\|_1$ a flattened rectangle, and for $\|\cdot\|_2$ an ellipse. As $a \to \infty$ the cross-section of the excluded region goes to zero. In the limit $a = \infty$ none of the above are norms, and the Propositions become false. When $a=1$ norm $\|\cdot\|_2$ gives the usual disparity gradient measure. The norm can thus be varied to give excluded regions of varying shape and cross-sectional area. This will affect the exact nature of the *between rasters* constraint being imposed.

2.3. Conclusions

The disparity gradient limit is a constraint on scene *jaggedness* embracing simultaneously the ideas of opacity, scene continuity, and continuity between views. It is a sufficient but not necessary condition for these properties to hold. The concept of *continuity* employed in the preceding sections is that obtained by imposing a bound on the Lipschitz constant. It is argued in the Appendix that this is the most useful approach in practical situations.

There are whole families of stronger and weaker sufficient conditions of the same type, of which the usual disparity gradient limit is the only isotropic example. This isotropy is not required by the geometry of binocular viewing or to impose continuity. The human visual system may prefer the isotropic condition however, because if an object can be viewed stereoscopically in one position, then a rotation about the line of sight does not alter this.

Perhaps more importantly we have a free parameter, the disparity gradient limit DG, which can be varied over the range $0<DG<2$ while still retaining the continuity of the view-to-view and scene-to-view maps. Taking a value $DG \approx 0$ gives a very strong constraint, allowing only nearly fronto-parallel surfaces (this has been used locally as the basis of a stereo correspondence algorithm, Marr & Poggio 1976). At the other extreme a value $DG \approx 2$ gives the weakest isotropic Lipschitz constraint which is consistent with non-self-occlusion of the underlying surfaces. We are free to choose the value of DG between these two limits, and in the following sections we will show how an intermediate value of DG supplies sufficient disambiguating power while imposing a (statistically) weak constraint on the observable surfaces in the world.

3. The PMF stereo algorithm

The basis for disambiguation in the PMF stereo algorithm (Pollard *et al* 1984, 1985; Pollard 1985) is provided by the local satisfaction of a moderate disparity gradient limit. Initial matching strengths are computed for each potential match (denoted $MS_{pp'}$ for match $M_{pp'}$) from the sum of within gradient limit support they receive from other potential matches in their immediate neighbourhood. The ubiquitous uniqueness constraint (Marr and Poggio 1976, 1979) is then exploited in order to resolve ambiguity on the basis of these matching strengths.

3.1. Computing matching strengths

Consider the points in one image, the left say. Each point, for example point p, identified in that image can take part in a number of matches (eg $M_{pp'}$, $M_{pq'}$ etc). For convenience a circular neighbourhood is defined in the left image about p. Only those matches associated with points that lie within this neighbourhood are allowed to contribute support to each of the possible matches for p, and only then if the disparity gradient between them is less than a moderate predetermined limit. Furthermore, in accordance with the uniqueness constraint only a single match for each primitive in the neighbourhood is allowed to make a contribution.

Consider a single pair of points in one image. The range of disparity difference over which the gradient limit is satisfied between them increases linearly (but not isotropically) with their physical separation. Accordingly it was decided that support should be scaled inversely with relative proximity. The advantages of alternative scaling factors are discussed in Pollard (1985).

The matching strength can be expressed more formally as

$$MS_{pp'} = \sum_{i \in N(p)} \max_{all\ j'} \frac{Cij' \times DG(M_{pp'}, M_{ij'})}{S(p, i)}$$

where $N(p)$ is the set of points in the neighbourhood of p. $DG(M_{pp'}, M_{ij'})$ is a function of the disparity gradient that exists between match $M_{pp'}$ and $M_{ij'}$, it has value one if the gradient is less than the chosen limit and zero otherwise. $S(p, i)$ is the magnitude of the separation between points p and i. $C_{ij'}$ reflects the goodness of the match between primitives i and j. In those situations where there is more than one match for a single primitive that satisfies the gradient limit, only the stronger contributes to the matching strength.

3.2. Resolving ambiguity

Uniqueness is enforced via a simple discrete iterative winner take all procedure. At each iteration those matches having the highest matching strength for both of the two image primitives forming them are immediately chosen as *correct*, ie matches that are maximal with respect to both lines of sight are selected. Subsequently, alternative matches associated with the two primitives that form each selected match can be eliminated from further consideration. This allows further matches, not previously either accepted or eliminated, to be selected as correct provided that they now have the highest strengths for both constituent primitives. In practice convergence of this procedure usually occurs after only 4 or 5 iterations, with the overwhelming majority of matches being identified at the first iteration.

3.3. Support

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. The extent to which the disparity gradient limit is violated in the neighbourhood of a candidate match does not directly effect PMF's selection procedure. Alternatively, it would be possible to reduce the strength of a match in accordance with the number of points over the neighbourhood that did not possess matches that satisfied the gradient limit. However it cannot be assumed that a moderate disparity gradient limit will be satisfied everywhere. The disparity gradients that exist across depth shears, for example, will generally be large. Fortunately the satisfaction of the gradient limit that exists to either side of the discontinuity is generally sufficient to resolve ambiguity in their vicinity.

3.4. Consistency

A further point to note is that within-disparity gradient support is sought independently from all possible matches in the neighbourhood of the match under consideration. Hence it is possible that two or more matches that give within-disparity gradient limit support might not themselves share a within-limit disparity gradient. An alternative approach would be to require that all matches that give support be mutually consistent, that is the disparity gradient limit must be satisfied amongst them. However the computational overheads involved in recovering the best such score are prohibitive even for quite small neighbourhoods. Hence this requirement is relaxed in the design of PMF for reasons of computational efficiency. The effect of this simplification is illustrated below.

4. Statistics of Projection

In practice the selection of a suitable limiting disparity gradient in PMF is doubly constrained. On the one hand the chosen limit should provide sufficient generality, ie it should admit as wide a range of surfaces as possible. And on the other it should provide sufficient disambiguating power to resolve any ambiguity that may be present. In this section we shall examine the statistics of stereo projection in order to show that the restriction on the set of possible surfaces resulting from the imposition of a moderate disparity gradient limit is not severe.

Proposition 3 relates gradients in the scene to disparity gradients in the image. Expressing the world gradient as $\tan s$ and rearranging gives

$$DG = \frac{I}{d} \cdot \tan s$$

Notice how the magnitude of the disparity gradient is scaled with respect to the viewing parameters I and d. Generally, for viewing systems approximating that of the human visual system, d is large in comparison to I and the vast majority of disparity gradients in the image will be small. Following Arnold and Binford (1980), also reported by Kass (1984), it is possible to derive a probability density function for disparity gradient based upon the assumption that surface orientation is uniformly distributed over the gaussian sphere (details in Pollard 1985). That is

$$Pdf(DG) = \frac{\cos(\tan^{-1}(\frac{DG}{a}))a}{a^2 + DG^2}$$

where

$$a = \frac{I}{d}$$

for DG in the range 0 to infinity.

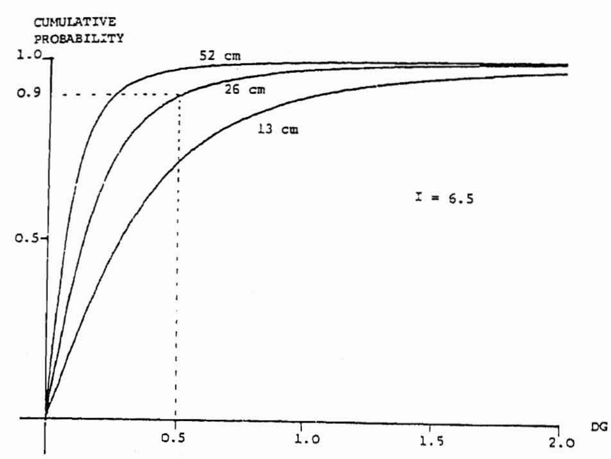


Figure 1

In figure 1 the cumulative density function is plotted for a range of d values representative of the distances for which binocular stereo is considered to be an important depth cue in human vision. The inter-ocular separation I is assumed to be 6.5cm. Inspection of figure 1 clearly shows that at any reasonable viewing distance the majority of disparity gradients will be small. For example, less than 10% of world surfaces viewed at more than 26cm will present with disparity gradient in excess of 0.5. Hence one can argue that to enforce a disparity gradient well below the *theoretical* limit (of 2) imposes negligible restrictions on the *worlds* that can be fused by the stereo algorithm.

5. Disambiguation

So far it has been demonstrated that the binocular projections of most planar and many jagged surfaces will present small disparity gradients almost everywhere (given always that interocular separation is small with respect to viewing distance). Consequently, assuming some cohesiveness of the world, the disparity gradients that exist between the correct matches of the image features that describe the structure of the viewed surfaces will almost always be within a moderate limit. But, as implied above, disambiguation also needs to rely upon there being a low probability for this limit to be satisfied between incorrect matches (also called ghosts for convenience) of the same image features.

This section presents the results of a simple computational experiment designed to illustrate the effect on disambiguating power of varying the magnitude of the limiting disparity gradient. For this purpose it is convenient to consider the matching problem associated with patterns of random point features (dots generated to 32 bit floating point resolution) of a uniform density. Whilst this restriction is not entirely satisfactory, the ambiguity problem associated with dot patterns makes them typical of some of the most problematic textures that exist in natural imagery.

5.1. Strength Ratios

For our purposes disambiguating power is defined as the extent to which it is possible to distinguish those matches that are correct from their associated ghosts. The ratio of incorrect to correct match strength provides a useful metric in this regard. If this ratio is generally small the resulting disambiguation power will be sufficient to resolve the vast majority of correct matches even in the presence of large quantities of noise and/or close to disparity discontinuities. Experimental results are presented in the form of a frequency distribution compiled as a result of a large number of independent trials. At each trial:

(i) a single dot is chosen from a random dot pattern

(ii) the strength of its correct match is computed by matching the dot with the same dot in an identical dot pattern

(iii) the strength of an incorrect match is computed by matching the dot with a dot in an independent dot pattern

(iv) their ratio is computed

(v) this single contribution is added to the distribution

The magnitude of the strength of a correct match, given by (ii), is equivalent for all situations in which the gradient limit is satisfied amongst the correct matches, ie where the word projects as a suitable Lipschitz disparity surface. This follows from the fact that in such situations the matching strength of PMF is only dependent upon the projection into the left image. For noiseless data, the matching strength for a correct match (CS) will receive a single contribution from each dot in the neighbourhood.

$$CS = \sum_{i \in N} \frac{1}{S(i)}$$

where N is the set of dots in the neighbourhood and $S(i)$ is the physical separation of dot i.

For an incorrect match only those dots in the neighbourhood that have matches that satisfy the gradient limit by chance are allowed to contribute to the matching score.

$$IS = \sum_{i \in N} \frac{INDG(i)}{S(i)}$$

where $INDG(i)$ is a boolean that is satisfied only if there exists a match for dot i that satisfies the disparity gradient limit with respect to the match under consideration.

Figure 2 presents the normalised frequency distribution of strength ratios (IS/CS) (obtained over 1000 independent trials) for several values of limiting disparity gradient. The random dot patterns used were of uniform density with 0.1 dots per unit area [pixel] and the neighbourhood size was of 7 units radius. A vertical matching range of one unit was allowed for matching each dot in the neighbourhood. In the absence of noise the maximum possible ratio magnitude is 1. Hence good disambiguating power is characterised by a distribution of strength ratios clustered about some value much less than 1. This is the case where the disparity gradient is of moderate size, ie in the range between 0.5 and 1 (even within this range considerably better disambiguation power is achieved with the lower limit). Beyond such values the degree of deformation allowed between the two images (without affecting the matching strength) increases rapidly, and thus the disambiguation power provided falls off equally rapidly. A disparity gradient limit of 2 (approximated by a limit of 1.99), the *physical* limit along epipolar lines, provides almost no disambiguating power when used in this way; almost all incorrect matches obtain the same quantity of within-gradient support as their associated correct matches.

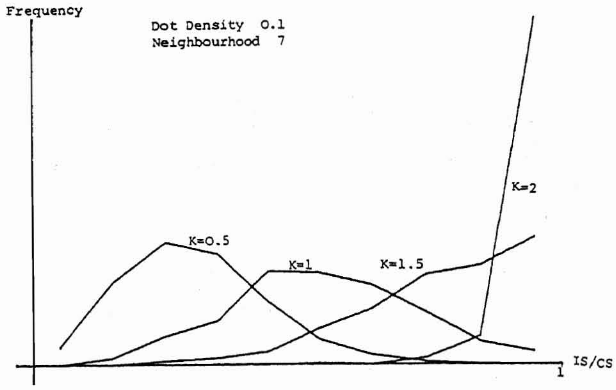


Figure 2

5.2. Mutual Consistency

It is also possible to examine the effect on the matching strength computation of PMF of the decision to relax the mutual consistency constraint. Matching strengths for correct matches remain the same as by definition correct matches defined here will all mutually satisfy the gradient limit over the extent of the neighbourhood. For incorrect matches, however, it is necessary to search the set of possible matches for each dot in the neighbourhood for the mutually consistent subset that maximises the matching strength, the initial set being limited to those matches, of each dot, that satisfy the gradient limit with respect to the match in question. Frequency distributions for mutually consistent matches are displayed in figure 3 for gradient limits 0.5, 1 and 1.5 (1.99 being too expensive to compute).

Some improvement, for each chosen value of the disparity gradient limit, is observed. However the overall trend of disambiguating power decreasing with the magnitude of the gradient limit still occurs. In practical terms the actual improvement in disambiguation power that is gained for a small disparity gradient limit, such as 0.5, is not appreciable when compared with the computational simplicity of computing matching strengths without insisting upon consistency. Enforcing mutual consistency in this way is combinatorially explosive, being closely related to the max clique problem (which is known to be NP complete; see Aho, Hopcroft and Ullman 1974).

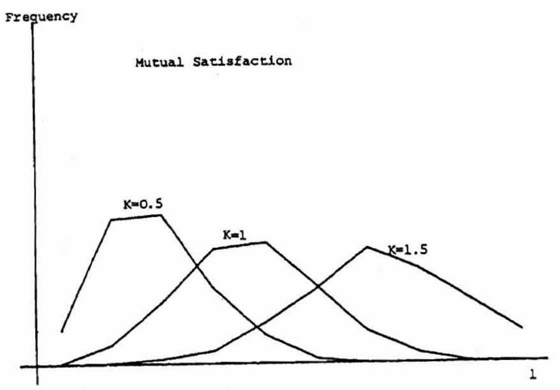
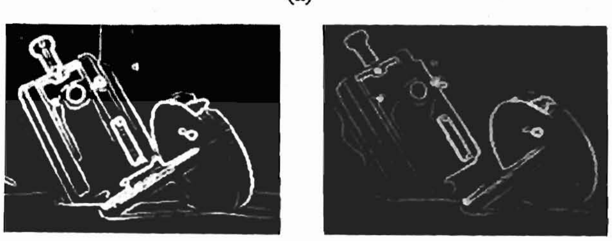


Figure 3

(a)



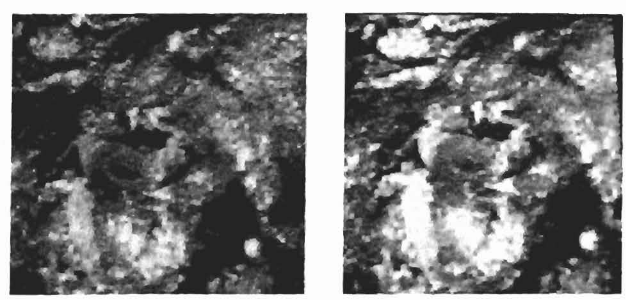
(b)

(c)
Figure 4

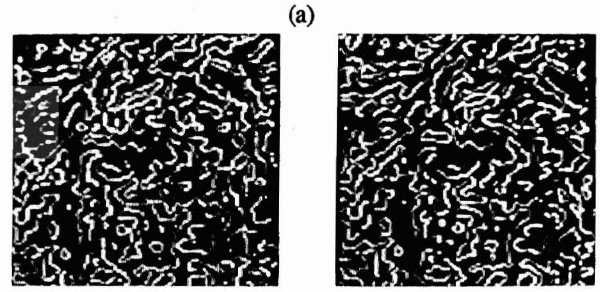
6. Examples of Performance

Two examples of the performance of the PMF algorithm on natural scenes are given here. The adopted disparity gradient limit is that reported by Burt and Julesz (1980a,1980b) for the human visual system, ie DG=1. Note that in terms of generality and disambiguating power this provides a fairly liberal constraint.

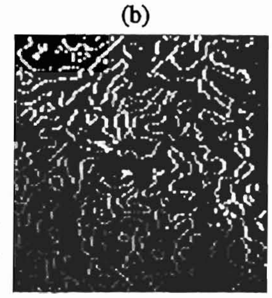
The stereogram given in figure 4(a) is representative of a simple industrial scene. Each image is 256 pixels square. Edge-like primitives, identified with a single scale (σ=2) Canny edge detector (Canny 1983) are displayed in (b), with intensity used to code edge contrast. As edge primitives are represented as pixels the constraint provided by epipolar geometry is implemented by limiting search to corresponding rasters in the two images (vertical disparities between the two images are always less than half a pixel). Potential matches are limited to a disparity range of $\pm$30 pixels and are required to be of the same contrast polarity and of a similar orientation. The actual limit on reorientation is given by the magnitude of the disparity gradient limit. Hence near vertical edge segments are allowed to reorient more so than horizontal ones. Disparity data recovered from matches selected by PMF (with a neighbourhood of radius 10 pixels) is displayed in (c) as a cyclopean image with intensity used to approximate relative depth (brighter points are closer to the imaging device). Unmatched points are not displayed.



(a)



(b)



(c)
Figure 5

A very different scene is the subject of the stereogram in figure 5(a). 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. The images are just 128 pixels square. All other parameters are the same as those used to process figure 4. The edge-like primitives in (b) are provided by a single high frequency (ω=5) Marr-Hildreth operator (Marr and Hildreth 1980). Disparity data is presented in (c) with intensity coding depth.

7. Concluding Comments

This paper has explored the theoretical and practical justifications for the use of the disparity gradient constraint in stereo matching. It has been shown that provided the gradient limit is less than 2 Lipschitz continuity is enforced between the stereo images and furthermore both the world and the cyclopean disparity map will be Lipschitz (with respect to a cyclopean coordinate frame). The selection of a suitable gradient limit is subject to pragmatic constraints. It has been shown that a moderate gradient limit (between 0.5 and 1) captures the statistics of small baseline stereo projection and provides considerable disambiguating power. The PMF stereo algorithm exploits the local satisfaction of the gradient limit in a simple and computationally attractive fashion. The strength of each match could, on a suitable computer architecture, be computed entirely in parallel and because of the considerable disambiguating power provided by the gradient limit in a single pass over the images.

References

Aho A. V., J. E. Hopcroft and J. D. Ullman (1974) *The design and analysis of computer algorithms,* Addison Wesley.

Arnold R. D. and T. O. Binford (1980) Geometric constraints in stereo vision, *Soc. Photo-Optical Instr. Engineers,* 238, 281-292.

Baker H. H. (1982) Depth from edge and intensity based stereo, *PhD Thesis,* University of Illinois.

Baker H. H. and T. O. Binford (1981) Depth from edge and intensity based stereo, *IJCAI 7,* Vancouver, B. C., 631-636.

Burt P. and B. Julesz (1980a) A disparity gradient limit for binocular fusion, *Science* 208, 615-617.

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

Canny J. F. (1983) Finding edges and lines in images, *MIT AI Lab. Memo* 720.

Grimson W. E. L. (1983) Computational experiments with a feature base stereo algorithm. *MIT AI Lab. Memo* 762.

Kass M. H. (1984) Computing stereo correspondence, *MSc Thesis,* Dept. of Elect. Eng. and Computer Sci., MIT, Cambridge, Mass.

Marr D. and E. Hildreth, (1980) Theory of edge detection, *Proc. R. Soc. London B* 207 187-217.

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

Marr D. and T. Poggio (1979) A theory of human stereo vision, *Proc. R. Soc. London B* 204 301-328.

Mayhew J. E. W. (1983) *Stereopsis, in Physiological and Biological Processing of Images,* O. J. Braddick and A. C. Sleigh (Eds), Springer Verlag, Berlin, 204-216.

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

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

Ohta and Kanade (1985) Stereo by two-level dynamic programming, *IJCAI* 9, Los Angeles, 1120-1126.

Pollard S. B. (1985) Identifying correspondences in binocular stereo, *PhD Thesis,* Department of Psychology, University of Sheffield.

Pollard S. B., Mayhew J. E. W. and Frisby J. P. (1984) PMF: A stereo correspondence algorithm using a disparity gradient limit, *paper presented at ECVP* 7, *Perception* 13 *(abstract only).*

Pollard S. B., Mayhew J. E. W. and Frisby J. P. (1985) PMF: A stereo correspondence algorithm using a disparity gradient limit, *Perception (in press).*

Prazdny (1985) Detection of binocular disparities, *Biol. Cyb. (in press).*

Trivedi H. P. and S. A. Lloyd (1985) The role of disparity gradient in stereo vision, *Comp. Sys. Memo* 165, GEC Hirst Research Centre, Wembley, England.

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

Tyler C. W. (1974) Depth perception in disparity gratings, *Nature* 251, 140-142.

Tyler C. W (1975) Spatial organization of binocular disparity sensitivity, *Vision Research* 15, 583-590.

Yuille A. L. and T. Poggio (1984) A generalised ordering constraint for stereo correspondence, *MIT AI Memo* 777.

Appendix: Norms, Lipschitz conditions, and Continuity

A norm $\|\cdot\|$ on a vector space V is a map

$$V \rightarrow \mathbf{R} : a \rightarrow \|a\|$$

with the three properties

$$\|a\| \geq 0$$

$$\|\lambda a\| = |\lambda| \, \|a\|$$

$$\|a + b\| \leq \|a\| + \|b\|$$

for all $a,b \in V$ and $\lambda \in \mathbf{R}$. A function $f: V \rightarrow W$ is Lipschitz order α, $(0<\alpha<1)$ constant K if

$$\|f(y) - f(x)\|_W \leq K \, \|y - x\|_V^{\alpha}$$

for all $x,y \in V$. Such a Lipschitz function is always continuous since for any $\varepsilon>0$ by putting $\delta=(\varepsilon/K)^{1/\alpha}$ we have

$$\|f(y) - f(x)\|_W \leq \varepsilon$$

whenever

$$\|y - x\|_V \leq \delta$$

A function is a homeomorphism if it is one-to-one and it and its inverse are continuous, and hence if it and its inverse are Lipschitz.

If f is Lipschitz order 1 constant K then if it is also differentiable its gradient has magnitude bounded by K. This Lipschitz condition can thus be regarded as a measure of jaggedness intermediate between continuity and bounded differentiability.

In practical situations we are working with functions defined only at a finite set of points. For such functions the notions of continuity and differentiability are meaningless. We can however calculate the Lipschitz constant K and this will give us a simple measure of *continuity*. In practice continuity always means a particular choice of bound on K. For example we may have a set of image points on adjacent screen rasters. We will say that they form a continuous curve if the change in position between rasters is less than, say, 2 pixels. This is a Lipschitz condition order 1 constant 2. The mathematical definition of continuity is not useful here since any finite set of points lies on some continuous curve.