[13] Viewpoint-Invariant Reconstruction of Visible Surfaces

Andrew Blake

Department of Computer Science
University of Edinburgh, Edinburgh EH9 3JZ, Scotland,UK

An earlier and shorter version of this paper was published in the *Proceedings of the National Conference on Artificial Intelligence*, 1984, 23-26, published by William Kaufman.

Abstract

Stereoscopic vision delivers a sparse map of the range to various "matched" points or contours, in the field of view. This paper addresses the problem of explicitly reconstructing a smooth surface that interpolates those points and contours. It is argued that any scheme for surface reconstruction should be viewpoint-invariant; otherwise the reconstructed surface would "wobble" as the viewpoint changes.

Progress has been made towards obtaining viewpoint-invariant reconstruction. A scheme has been implemented in 2-D and found to be relatively invariant to changes of viewpoint. Some remaining theoretical problems are outlined.

1 Introduction

In this paper we consider aspects of the task of generating geometrical information from stereo vision, extending the conclusions of a previous paper [4]. The aim is to derive as rich a geometric description as possible of the visible surfaces of the scene - a "viewer-centred representation of the visible surfaces" [12]. Principally this is to consist of information about surface discontinuities and surface orientation and curvature. Ideally it would be desirable to label discontinuities, and generate smooth surfaces between them, all in a single process. Some preliminary work has been done towards achieving this [3] but here we restrict discussion to reconstruction of smooth surfaces.

Grimson [9] discusses the task of interpolating smooth surfaces inside a known contour (obtained from stereo e.g. [13], [11], [9], [2]). He shows how surface interpolation can be done by minimising a suitably defined surface energy, the "quadratic variation". The interpolating surface that results is *biharmonic* and under most conditions is defined uniquely. Terzopoulos [18] derives, via finite elements, a method of computing a discrete representation of the surface; the computation uses relaxation which is widely favoured for minimisation problems in computer vision [22], largely because of its inherent parallelism. Both Grimson and Terzopoulos suggest that the surface computed represents the configuration of a thin plate under constraint or load.

In this paper we first point out that the faithfulness of the computation to the physical thin plate holds only under stringent assumptions - assumptions that do not apply for the intended use in representing visible surfaces. It is argued that physical thin plates do not anyway have the right properties for surface interpolation - it is not desirable to try and model one. Secondly, the effect of biharmonic interpolation is investigated in its own right. We show that it lacks 3-D viewpoint-invariance and demonstrate, with 2-D examples, that this results in an appreciable "wobble" of the reconstructed surface as the viewpoint is varied.

An alternative method of surface reconstruction is proposed, minimising a surface energy that is invariant to changes of viewpoint. However there is a problem of non-uniqueness: there may be more than one minimising surface - a surface that *locally* minimises the surface energy. As viewpoint changes, the surface delivered by an optimisation process could "flip" from one local minimum to another, resulting in loss of viewpoint-invariance. The best compromise seems to be the mixed membrane/plate. The membrane on its own is stable and viewpoint-invariant, with unique minimum energy solutions, but the reconstructed surface is not smooth. The plate, on the other hand, is smooth but not viewpoint-invariant. A mixture of the two energies gives a controlled trade-off between smoothness and viewpoint-invariance.

2 The thin plate

Accurate mathematical modelling of a thin plate is fraught with difficulties and, in general, generates a somewhat intractable, non-linear problem. Under certain assumptions however the energy density on the plate can be approximated by a quadratic expression; minimising the total energy in that case is equivalent to solving a linear partial differential equation with linear boundary conditions. The partial differential equation determines the displacement $u(x, y)$ of the plate, in the z-direction

*Current address: Dept. Engineering Science, Parks Rd., Oxford OX1 3PJ.

(the viewer direction), that interpolates a set of matched points. These matched points are assumed to be available as the output of stereopsis. With an approximate representation of the plate in a discrete (sampled) space, using finite differences or finite elements, the linear differential equation becomes a set of simultaneous linear equations. These can be solved by relaxation. The assumptions necessary to approximate the surface energy by quadratic variation are analysed by Landau and Lifshitz [10] and we enumerate them:

1. The plate is thin compared with its extent.
2. The displacements of the plate from its equilibrium position $z = 0$ are substantially in the z-direction; transverse displacement is negligible.
3. The normal to the plate is everywhere approximately in the z-direction.
4. The deflection of the plate is everywhere small compared with its extent.
5. The deflection of the plate is everywhere small compared with its thickness

Assumption 1 is acceptable - indeed intuitively it is preferable to use a thin plate that yields willingly to the pull of the stereo-matched points. Assumption 2 may also be acceptable if the pull on the plate from each matched point is normal to the plate. The remaining assumptions 3-5 are the ones which prove to be stumbling blocks for reconstruction of visible surfaces.

Assumption 3 is clearly unacceptable: any scene (for example, a room with walls, floor, table-tops etc.) is liable to contain surfaces at many widely differing orientations. By no means will they all be in or near the frontal plane (i.e. normal to the z-direction), though it seems that human vision may have a certain preference for surfaces in the frontal plane (Marr, 1982). In particular, surfaces to which the z-axis is almost tangential are of considerable interest: it is important to be able to distinguish, in a region of large disparity gradient, between such a slanted surface and a discontinuity of range (caused by occlusion).

Assumption 4 and the even stronger assumption 5 are again unacceptably restrictive. In fact assumption 5 can be removed at the cost of introducing non-linearity that makes the problem considerably harder; the non-linear formulation takes into account the stretching energy of the plate as well as it bending energy. It is this energy that represents the unwillingness of a flat plate to conform to the surface of a sphere rather than to, say, a cylindrical or other developable surface. Even without assumption 5, assumption 4 on its own is still too strong because it requires the scene to be relatively flat - to have an overall variation in depth that is small compared with its extent in the xy plane. This is clearly inapplicable in general.

One conclusion from the foregoing review of assumptions is that that faithfulness of visible surface reconstruction to a physical thin plate model is undesirable. This is because of the stretching energy discriminating against spherical surfaces, which is not generally appropriate in surface reconstruction. In fact, happily enough, we saw that quadratic variation is *not* an accurate description of the surface energy of a thin plate precisely because it omits stretching energy, so biharmonic interpolation does *not* exhibit this discrimination.

We now declare ourselves free from any obligation to adhere to a physical thin plate model and will explore the geometrical properties of biharmonic interpolation.

3 Biharmonic interpolation

We now examine biharmonic interpolation in its own right. A variety of forms of such interpolation are possible and the one preferred by Grimson [9] is to construct that surface $z = f(x, y)$ that (uniquely) minimises the quadratic variation

$$F = \int A \, dx \, dy \quad \text{where} \quad A = f_{xx}^2 + f_{yy}^2, \tag{1}$$

subject to the constraints that $f(x, y)$ passes through the stereo-matched points[1]. Landau and Lifshitz show [10] that the solution to this minimisation satisfies the biharmonic equation

$$\nabla^2 \nabla^2 f = 0, \tag{2}$$

under certain boundary conditions. For instance when the edges of the surface are fixed (constrained, for example, by stereo-matched points) the condition is that

$$f \text{ is fixed and } \partial^2 f / \partial n^2 = 0, \tag{3}$$

where $\partial / \partial n$ denotes differentiation along the normal to the boundary. Consider the effect on a simple shape such as a piece of the curved wall of a cylinder, assuming that the surface is fixed on the piece's boundary. It is easy to show that a cylindrical surface defined by

$$f(x, y) = \sqrt{a^2 - x^2} \tag{4}$$

does not satisfy $\nabla^2 \nabla^2 f = 0$, so we cannot expect the surface to be interpolated exactly. Grimson [9] demonstrates this: his interpolation of such a boundary conforms to the cylindrical surface near the boundary ends but sags somewhat in the middle.

To return to the definition in (1), a serious objection to using quadratic variation to define surface energy is that it is not invariant under change of 3D coordinate frame. As (Brady and Horn, 1983) point out, it is isotropic in

[1] An alternative formulation attaches the surface $f(x, y)$ to matched points by springs, allowing some deviation of the surface from the points.

2D - invariant under rotation of axes in the xy plane. However, under a change of coordinate frame in which the z-axis also moves, the quadratic variation proves not to be invariant.

Is it altogether obvious that 3D invariance is required? Certainly the situation is not entirely isotropic in that the visible surface is single valued in z - any line perpendicular to the image plane intersects the visible surface only once - the z-direction is special. On the other hand it is also desirable that the interpolated surface should be capable of remaining the same over a wide range of viewpoints. Specifically, given a scene and a set of viewpoints over which occlusion relationships in the scene do not alter, so that the points matched by stereo do not change, the reconstructed surface should remain the same over all those positions. Such a situation is by no means a special case and is easy to generate: imagine, for example, looking down the axis of a "beehive" (fig 1). There is

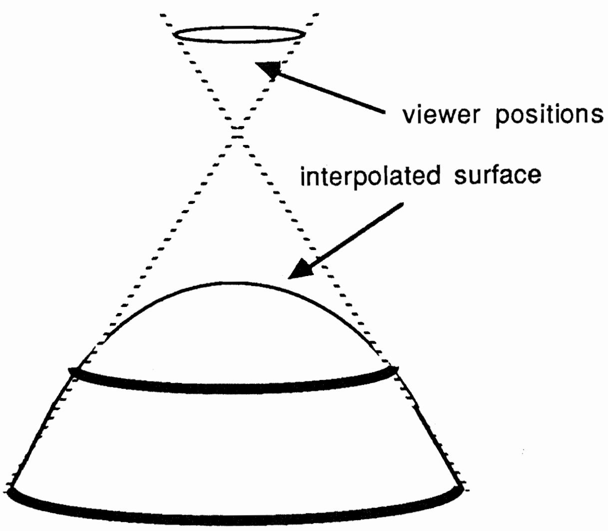

Figure 1: **Viewpoint-invariant surface reconstruction.** *The two solid rings can be interpolated by a "beehive"-shaped surface. Within the cone of viewer directions shown, the lower ring is not obscured by the upper one. It is argued that, for viewer directions lying in that cone, the interpolated surface should remain static (invariant) in 3-D.*

no change of occlusion over a range of viewer directions that lie inside a certain cone. The reconstructed surfaces of both beehive and table should remain static in 3D as viewpoint varies within that cone. The point is that, over such a set of viewpoints, the available information about the surface does not change; neither then should there be any change in the estimate of its shape[2]. Without invariance, a moving viewer would perceive a wobbling surface.

[2] Since the short version of this paper [4] appeared there has been some logomachy in the literature over the term "viewpoint-invariance". It has apparently been misunderstood by some as referring to invariance over *all* viewpoints, who proposed instead a cacology that was allegedly more accurate. What our term refers to, of course, is invariance over *some set* of viewpoints.

To demonstrate the wobble effect, surface interpolation using quadratic variation has been simulated in 2-D (fig 2) over a range of viewpoints. In the 2-D case, biharmonic interpolation simply fits a piecewise cubic polynomial to set of points. There is continuity of second derivative at those points and the second derivative is zero at the end-points. In other words, interpolation in 2-D reduces simply to fitting cubic splines [6]. As expected, the wobble effect is strong when boundary conditions are such that the reconstructed surface is forced to be far from planar.

4 A viewpoint-invariant surface energy

4.1 Deriving the energy expression

In order to obtain the desired invariance to viewpoint while still constraining the surface to be single valued along the direction of projection, the interpolation problem can be reformulated as follows: first surface energy is defined for an arbitrary 3D surface, defined by

$$g(x, y, z) = 0 \tag{5}$$

then the single value constraint is applied, that g must have the form

$$g(x, y, z) = f(x, y) - z \tag{6}$$

In this way we can generate a new energy expression to replace (1) that does have 3D invariance, because it is defined in terms of surface properties. The energy is:

$$F = \int E dS \text{ where } E = \kappa_1^2 + \kappa_2^2. \tag{7}$$

and where κ_1, κ_2 are principal curvatures and dS is the area of an infinitessimal surface element. This can be expressed quite straightforwardly, as a non-linear function of the derivatives [7]. It is not the *only* possible invariant energy but is consistent with the old expression (1) when $f_x = f_y = 0$ - the normal to the surface lies everywhere along the viewer direction. It is approximately consistent if the surface normal is everywhere close to the viewing direction. This is simply assumption 3, for the thin plate approximation, appearing again. Indeed, this consistency property leads to a proof that (1) is not in general an invariant expression: for a given surface element dS, we know that

- EdS is invariant with respect to change of coordinate frame
- $Adxdy = E$dS in one coordinate frame but not in certain others

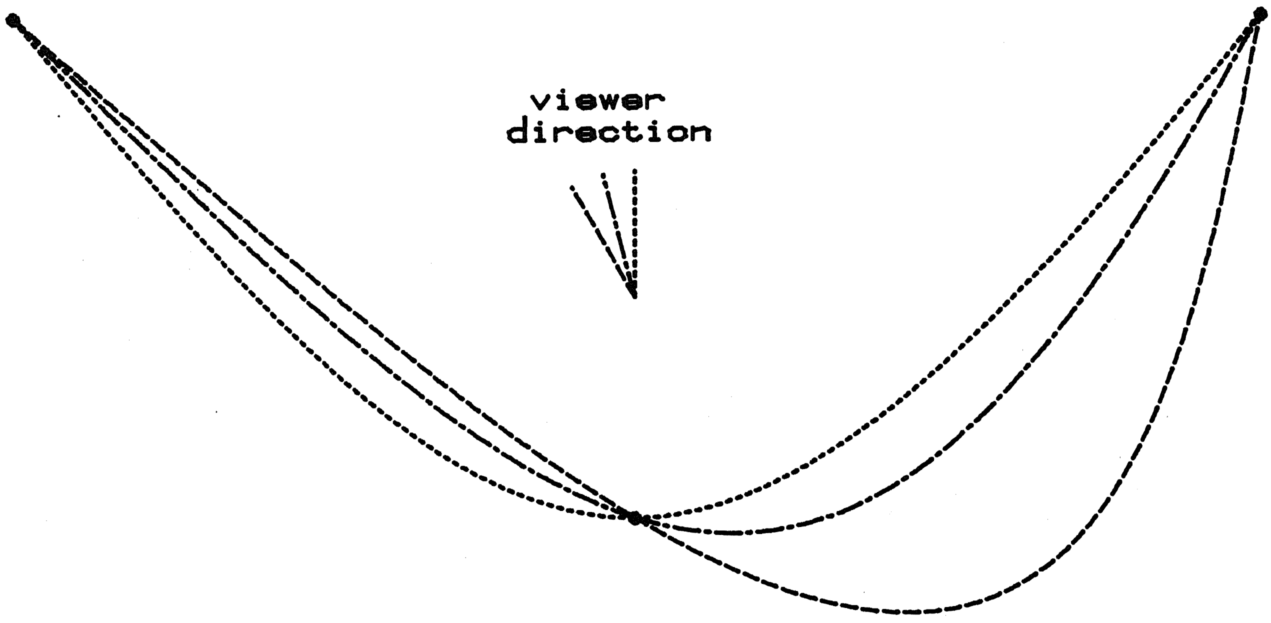

Figure 2: **Biharmonic interpolation scheme.** *Here is an example of the interpolation scheme operating in 2-D rather than 3-D. The curve interpolates 3 points (marked by circles). As the viewer direction varies from 0 to 30 degrees there is marked movement of the interpolating curve. Clearly the scheme is far from invariant to change of viewpoint.*

therefore $Adxdy$ cannot be invariant under change of coordinate frame.

The original energy (1) has a unique minimum [9] but for the new energy the situation is more complicated. To understand this we will consider, for simplicity, a 2-D form of the new energy:

$$F = \int_{x=x_1}^{x=x_2} E(f_x, f_{xx})\,\mathrm{d}s \qquad (8)$$

where

$$E(t,u) = u^2(1+t^2)^{-3}$$

and the arc length $\mathrm{d}s = w(f_x)\,\mathrm{d}x$ with

$$w(t) = \sqrt{1+t^2}.$$

A standard result from the calculus of variations [1] states a certain sufficient set of conditions for a minimum of F to exist, one of which, in the case of (8), is that:

$$\exists a > 0, p > 1, \text{ s.t. } \forall t, u, E(t,u)w(t) \geq a|u|^p + b. \qquad (9)$$

This condition is not satisfied as $E(t,u)$ becomes arbitrarily small for large enough t. This problem can be circumvented by restricting f to a family of functions whose normal is nowhere perpendicular to the line of sight - say at most 85^o away. Now the term in u is bounded below. However, the boundary of this set of functions (defined by a condition $|u| < U$ for some U) is coordinate-frame dependent. So a local minimum f of F is guaranteed to be invariant (to small changes of viewpoint) if it lies in the interior of this set. If it lies on the boundary it may be viewpoint-dependent.

Note that, by the Morse-lemma [14], an f which locally minimises F in one coordinate frame (u,l) also minimises F in another frame (u',l'), provided that the change of frame is a "diffeomorphism". This property is needed for viewpoint-invariance of f. The change of frame is not diffeomorphic if, somewhere on the surface, the gradient becomes unbounded in the new frame. This is as expected: the surface acquires an extremal contour in the new frame and some of the previously reconstructed surface is lost to view. Clearly there is no viewpoint-invariance in this situation.

There remains a more serious problem: that of uniqueness. The integrand of (8) fails to satisfy a certain sufficient condition for uniqueness [21] because it is not convex. It can easily be shown that its Hessian matrix with respect to u,l is not positive definite. Therefore the integrand $E(u,l)w(u)$ is not convex in u,l. In the absence of convexity there may be more than one extremal function (non-uniqueness). Practical methods of finding local minima (such as the descent method used in the discrete computation in the next section) are coordinate frame dependent. Both the initial state (initial estimate of f) and the path taken from that state depend on the coordinate frame. This would not matter if there were a unique minimum; in each coordinate frame the descent method would reach that minimum, albeit via different paths. However if there are several local minima, the final state may flip from one to another as the coordinate frame is varied. This would cause viewpoint-invariance to be lost.

A certain modification of the energy in (8) can be shown to make it's integrand convex, at least asymptotically. The modified energy is

$$F = \int (E + K) w \mathrm{d}x \tag{10}$$

where K is a positive constant. The additional term Kw adds a component of energy that is simply proportional to the length of the curve $f(x)$ between endpoints. In three dimensions this is simply the energy of an elastic membrane — energy is proportional to surface area. If K is very large, so that the membrane term dominates the energy F, the effect, in 2D, is simply to link the interpolated points by straight lines. Moreover the integrand is convex in the limit of large K. That is, the term $Kw(u)$ in (10) is convex in u because its second derivative

$$K w_{uu} = K \left(1 + u^2\right)^{-3/2} \tag{11}$$

is positive everywhere. The convexity result for the membrane holds also in 3D.

The membrane limit may be inappropriate to visible surface reconstruction because gradient discontinuities are introduced at interpolated points. (A drawn bow-string has a V-shaped kink at the archer's finger). An *intermediate* value of K produces a compromise between invariant interpolation and avoiding high curvature at interpolated points. Examples are shown in the next section. Note that the mixed membrane/plate is only approximately invariant. To make it fully invariant, it would be necessary to find a *convex, viewpoint-invariant set in u, l space over which the integrand is a convex function.* This is shown, in the appendix, to be impossible.

The modified energy (10) can easily be extended to the full 3-D case simply by adding a positive constant K to the integrand, as before. In 3-D, for large K, the surface behaves as a membrane having minimal area (and creases). An intermediate K achieves a compromise, as in 2-D.

5 Discrete computation in 2-D

Interpolation in 2-D, using the mixed membrane/plate of the previous section, has been implemented on a computer, using a parallel, iterative method. First the energy is expressed in a discrete form, by a finite element approximation [17]. Trial functions are represented as quadratic piecewise polynomials. A quadratic spline basis [6] allows the piecewise polynomials to be represented as vectors, each of whose components affects the energy function only locally. Hence when one of these components is adjusted only a local computation is necessary to update the total energy. This situation is typical of optimisation by parallel relaxation [22]. Computation is further simplified by approximating the energy (8) within each polynomial piece. The gradient u, in a polynomial piece, is approximated by its average value over that piece. A simple application of the patch test [17] shows that this is allowable.

The discrete scheme has been applied to the problem of fig 2, for which interpolation with quadratic variation was shown to be viewpoint-dependent. In practice, for an appropriate choice of K, the interpolated curve is very nearly static over viewpoints in the range $\pm 30^o$ (fig 3), without excessively high curvatures.

6 Conclusion

We have shown that:

1. Biharmonic interpolation does not accurately model a thin plate and, in any case, a thin plate model would be inappropriate for use in surface interpolation.

2. Biharmonic interpolation of the visible surface is not viewpoint-invariant and that, in specific 2-D cases, this lack of invariance certainly causes significant surface wobble.

3. A proposed alternative reconstruction scheme uses an energy that is a function of surface curvature and area — the mixed membrane/plate. In 2-D simulation the scheme appears to be relatively invariant to change of viewpoint. However the theoretical basis for invariance is still incomplete, because of problems demonstrating uniqueness. A possible line of investigation to try and resolve this would examine the Euler equation of the energy integrand. This would define the extrema. It might reveal, for instance, that for large values of the parameter K, any extremal surface must approximate to the surface of minimal area. In that case, reconstruction is, at least approximately, invariant to change of viewpoint.

4. If the visual task does not require smoothly interpolated surfaces then a computational membrane can be used which *is* viewpoint-invariant.

5. Before attempting to proceed to a full 3-D implementation, it is worth questioning whether it is anyway appropriate to perform full, explicit reconstruction of a surface as a range-map. An alternative would be to represent the surface in terms of prototype (e.g. quadric) surface patches [8], [15]. Such a representation could be computed from a range map; but a more direct route would be to perform surface reconstruction using the surface patches themselves, thus eliminating the need for the range map as an intermediate representation. This direct route might be

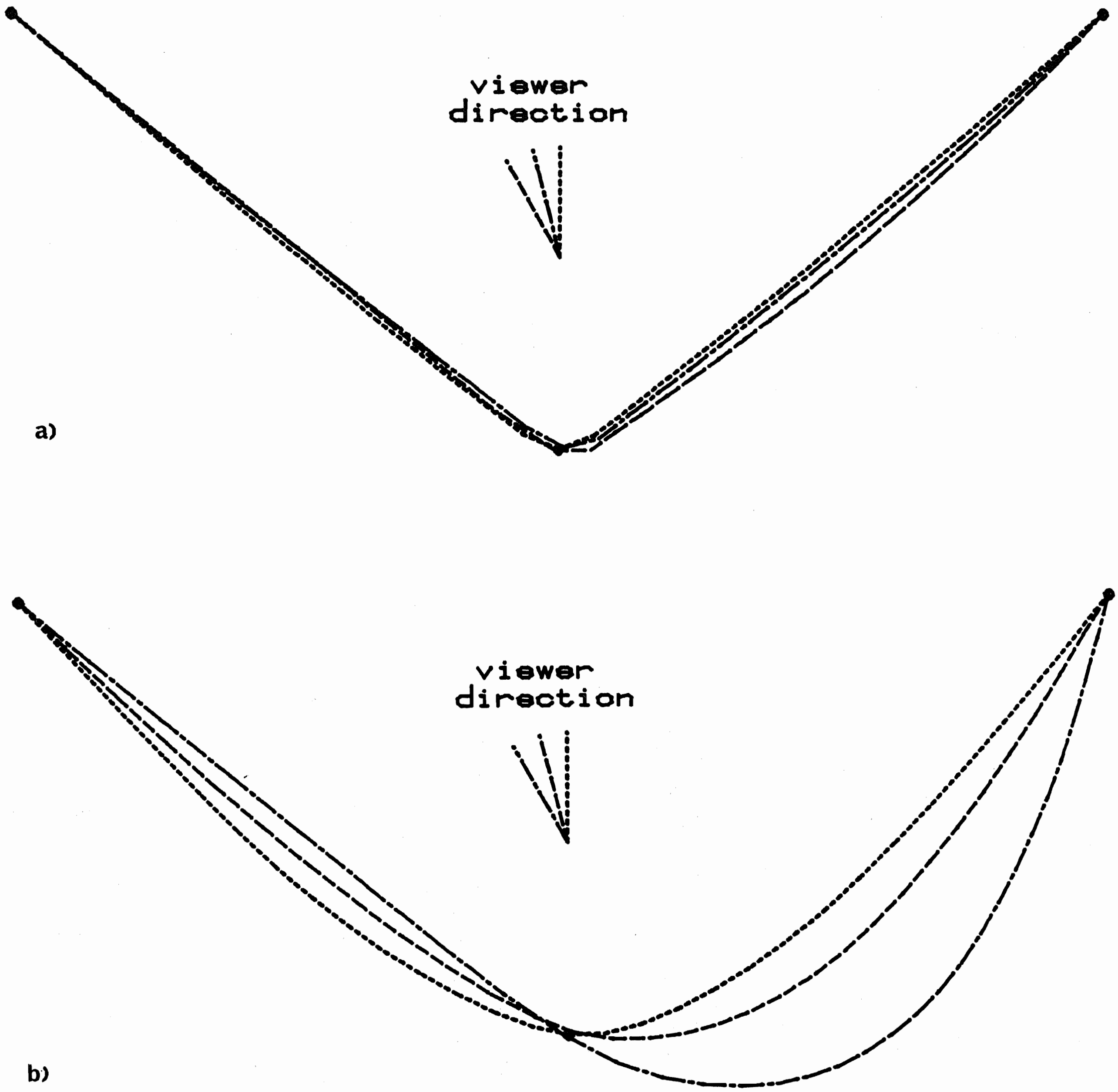

Figure 3: **Implementation of the proposed scheme.** *Compare results here with those for quadratic variation, in fig 2. The new scheme (a) is, for a modest value of the parameter* K*, fairly independent of viewpoint. As the parameter* K *is decreased there is more viewpoint dependence (b).*

attained by restricting the admissible family of functions in the finite element method to assemblies of prototype patches, and finding the member of that family with lowest energy.

6. The problem of detecting discontinuities has not been dealt with in this paper. Terzopoulos [19] suggests labelling zero-crossings of the reconstructed surface $f(x, y)$ as discontinuities. However, this method is not viewpoint-invariant (fig 5). Further work is needed here: one possibility is to incorporate a penalty for surface discontinuities into the surface energy function, an extension[3] of the method in [3].

7. The principle of 3-D invariance appears to be important in 2-1/2D sketch processes. Another potential area of application is shape-from-shading. The inferred shape of a beehive (fig 1) with a lambertian surface should be viewpoint-invariant, as with shape-from-stereo.

Acknowledgement

The author is grateful to Professor J. Ball and to Dr J. Mayhew and Professor J. Frisby for valuable discussion, to John Canny for pointing out an error in a previous version of this paper and to Andrew Zisserman for valuable comments.

References

1. Akhiezer,N.I. (1962). *The calculus of variations.* Blaisdell, New York.

2. Baker,H.H. (1981). Depth from edge and intensity based stereo. *IJCAI conf. 1981*, 583-588.

3. Blake,A. (1983). *Parallel computation in low-level vision.* Ph.D. Thesis, University of Edinburgh, Scotland.

4. Blake,A. (1984). Reconstructing a visible surface. *Proc AAAI conf. 1984*, 23-26.

5. Brady,M. and Horn,B.K.P. (1983). Rotationally symmetric operators for surface interpolation. *Comp. Vis. Graph. Image Proc.*, 22, 70-94.

6. de Boor,C. (1978). *A practical guide to splines.* Springer-Verlag, New York.

7. do Carmo,M.P. (1976). *Differential geometry of curves and surfaces.* Prentice-Hall, Englewood Cliffs, USA.

8. Faugeras,O.D. and Hebert,M. (1983). A 3-D recognition and positioning algorithm using geometrical matching between primitive surfaces. *IJCAI 83*, 996-1002.

9. Grimson,W.E.L. (1982). *From images to surfaces.* MIT Press, Cambridge, USA.

10. Landau,L.D. and Lifschitz,E.M. (1959). *Theory of elasticity.* Pergamon.

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

12. Marr,D. (1982). *Vision.* Freeman, San Francisco.

13. Mayhew,J.E.W and Frisby,J.P. (1981). Towards a computational and psychophysical theory of stereopsis. *AI Journal*, 17, 349-385.

14. Poston,T. and Stewart,I. (1978). *Catastrophe theory and its applications.* Pitman, London.

15. Potmesil,M. (1983). Generating models of solid objects by matching 3D surface segments. *IJCAI 83*, 1089-1093.

16. Roberts,A.W. and Varberg,D.E. (1976). *Convex functions.* Academic Press, New York.

17. G. Strang and G.J. Fix, *An analysis of the finite element method.* Prentice-Hall, Englewood Cliffs, USA, 1973.

18. Terzopoulos,D. (1983). The role of constraints and discontinuities in visible-surface reconstruction. *IJCAI 83*, 1073-1077.

19. Terzopoulos,D. (1984). *Multiresolution computation of visible-surface representations.* Ph.D. Thesis, M.I.T., Cambridge, USA.

20. Thorpe,J.A. (1979). *Elementary topics in differential geometry.* Springer-Verlag, New York.

21. Troutman,J.L. (1983). *Variational calculus with elementary convexity.* Springer-Verlag, New York.

22. Ullman,S. (1979). Relaxed and constrained optimisation by local processes. *Computer graphics and image processing*, 10, 115-125.

A. Convexity in the 2-D case

"Adding in" the convex component in (10) can make the entire integrand convex, but only over a subset of values of u, l. To see this we examine the Hessian of the integrand in (10) with respect to u, l. If the Hessian is strictly positive definite in some domain then the integrand is strongly convex there [16].

$$T(u, l) = (E(u, l) + K)w(u) \tag{12}$$

where $E = l^2 w(u)^{-6}$ and $w(u) = \sqrt{1 + u^2}$. Examining the hessian of T w.r.t u, l, a necessary and sufficient condition for convexity is that the hessian be positive definite [16]. Moreover a sufficient condition for strict convexity

[3] see the paper by Blake and Zisserman, in this volume.

is that the hessian be strictly positive definite. Differentiating T twice, we obtain:

$$\begin{aligned} T_{ll} &= 2w^{-5}, \\ T_{uu} &= 5(6u^2 - 10w^{-9} + Kw^{-3}), \\ T_{ul} &= -10ulw^{-7}. \end{aligned} \quad (13)$$

The eigenvalues of the Hessian are

$$\frac{1}{2}\left((T_{ll} + T_{uu}) \pm \sqrt{(T_{ll} + T_{uu})^2 - 4(T_{uu}T_{ll} - T_{ul}^2)}\right) \quad (14)$$

and, since T_{ll}, T_{uu} are positive, the smallest eigenvalue is positive iff

$$T_{ll}T_{uu} \geq T_{ul}^2. \quad (15)$$

Substituting (14) into this condition and simplifying yields the condition

$$K \geq 5\kappa^2(1 + 4u^2). \quad (16)$$

Making this inequality strict yields a condition for strict convexity.

If a viewpoint-invariant sufficient condition for convexity could be found, which also formed a convex set in u, l space, in all coordinate frames, then constrained optimisation within that set would be viewpoint-invariant. No such condition exists however. The *only* viewpoint-invariant function of u, l is κ and, from (16), κ-intervals are not convex sets in the u, l space of any viewer coordinate-frame.