RECONSTRUCTING A VISIBLE SURFACE

A. Blake
Computer Science Department,
University of Edinburgh,
King's Buildings,
Mayfield Rd,
Edinburgh,
Scotland.

ABSTRACT

We address the problem of reconstructing the visible surface in stereoscopic vision. We point out the need for viewpoint invariance in the reconstruction scheme and demonstrate the undesirable "wobble" effect that can occur when such invariance is lacking. The design of an invariant scheme is discussed.

I INTRODUCTION

In this paper we consider aspects of the task of generating geometrical information from stereo vision. 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" (Marr, 1982). 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 (Blake, 1983) but here we restrict discussion to reconstruction of smooth surfaces.

Grimson (Grimson, 1982) discusses the task of interpolating smooth surfaces inside a known contour (obtained from stereo e.g. (Mayhew and Frisby, 1981), (Marr, 79), (Grimson, 1982), (Baker, 1981)). 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 (Terzopoulos, 1983) 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 (Ullman, 1979), 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 that does have the requisite viewpoint-invariance.

II 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 $f(x,y)$ of the plate, in the z-direction (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 in (Landau and Lifshitz, 1959) 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.

III 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 (Grimson,1982) is to construct that surface $z=f(x,y)$ that (uniquely) minimises the quadratic variation

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

subject to the constraints that $f(x,y)$ passes through the stereo-matched points.* In (Landau and Lifshitz, 1959), the solution to this minimisation is given by the biharmonic equation

$$\Delta^2 f = 0, \text{ where} \tag{2}$$

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

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

($\partial/\partial n$ 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 $\Delta^2 f = 0$, so we cannot expect the surface to be interpolated exactly. Grimson (Grimson, 1982) 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 2D - invariant under rotation of axes in the x-y 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

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

- 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". There is no change of occlusion over a range of viewer directions that lie inside a certain cone. We want the reconstructed surfaces of both beehive and table to remain static in 3D as viewing position changes. 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.

Without invariance, a moving viewer would perceive a wobbling surface. To demonstrate the wobble effect, surface interpolation using quadratic variation has been simulated in 2-D (fig 1) over a range of viewpoints. In the 2-D case, biharmonic

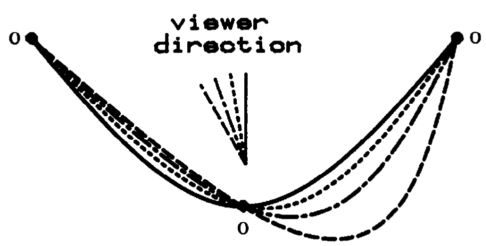

Figure 1: Biharmonic interpolation scheme.
Here is an example of the interpolation scheme operating in 2-D rather than 3-D. The curve interpolates 3 points (marked by circles). As the viewer direction varies from 0 to 30 degrees there is marked movement of the interpolating curve. Clearly the scheme is far from invariant to change of viewpoint.

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. As expected, the wobble effect is strong when boundary conditions are such that the reconstructed surface is forced to be far from planar.

IV A VIEWPOINT-INVARIANT SURFACE ENERGY

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 interpolation 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:

$$F = \int_S E \, dS \text{ where } E = k_1^2 + k_2^2$$

and where k_1, k_2 are principal curvatures and dS is the area of an infinitessimal surface element. This can be expressed in cartesian coordinates using the "Weingarten map" (Thorpe, 1979) to yield:

$$F = \int_S (p + 2q + r)(1 + u_x^2 + u_y^2)^{-1/2} \, dx \, dy \tag{7}$$

$$\text{where } p = u_{xx}^2 (1+u_x^2)^{-2}$$

$$q = u_{xy}^2 (1+u_x^2)^{-1} (1+u_y^2)^{-1} \tag{8}$$

$$r = u_{yy}^2 (1+u_y^2)^{-2}$$

This 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
- A dx dy = E dS in one coordinate frame but not in certain others

therefore A dx dy cannot be invariant under change of coordinate frame.

The original energy (1) has a unique minimum (Grimson, 1982) but with the new energy (7) the situation is more complicated. To understand this we will consider, for simplicity, a 2-D form of (7):

$$F = \int_{x=x_1}^{x=x_2} E(f_x, f_{xx}) ds \tag{9}$$

$$\text{where } E(t,u) = u^2(1+t^2)^{-3} \quad \text{and}$$

$$ds = w(f_x)dx, \text{ where } w(t)=(1+t^2)^{1/2}.$$

A standard result from the calculus of variations (Akhiezer, 1962) states certain sufficient conditions for a minimum of F to exist, one of which is that:

there exist $a>0$, $p>1$, b s.t. for all t,u

$$E(t,u)w(t) >= a|u|^p + b.$$

This condition is not satisfied by (9) because the term in t 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 t is bounded below.

There remains a uniqueness problem: $w(t)E(t,u)$ fails to satisfy a certain sufficient condition for uniqueness (Troutman, 1983): it is not convex. This too can be remedied by replacing E in (9) by E+P, where P is a positive constant, representing the energy of a flexible rod under a stretching load. Now, for t-values in a certain range $|t|<=T$ (T depends on P and may be made arbitrarily large by choosing a large enough P), $w(t)(E(t,u)+P)$ becomes convex. Thus the energy functional (7) is convex in f_x, f_{xx} provided that, for all x in the appropriate interval,

$$|f_{xx}|<=T. \tag{10}$$

The consequence is that any admissible function f for which the functional F (9) is stationary uniquely minimises F. This suggests that, in a discrete version of the problem suitable for computation, optimisation by gradient descent (using relaxation) could succeed in finding the surface f(x) that has minimum energy. If, for this f(x), there is equality in condition (10) then viewpoint invariance is lost. But provided T is chosen sufficiently large this will occur only for reconstructed curves of extreme slope and/or curvature. The case of extreme slope, for example, occurs at extremal boundaries - for which changing viewpoint affects occlusion - in which case reconstruction is not expected to be viewpoint invariant.

V CONCLUSION

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 possible alternative reconstruction scheme uses an energy that is a function of surface curvature and area.

This method is viewpoint invariant and certainly possesses the necessary existence and uniqueness properties, in the 2D case. It remains to extend these results to 3D and to develop and test a discrete computation to perform the reconstruction.

ACKNOWLEDGEMENT

The author is grateful to Professor J. Ball and to Dr J. Mayhew and Professor J. Frisby for valuable discussion. He is indebted to the University of Edinburgh for the provision of facilities.

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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