# Towards Qualitative Vision: Motion Parallax

**Andrew Blake, Roberto Cipolla & Andrew Zisserman**

Department of Engineering Science *
University of Oxford
Oxford, OX1 3PJ, U K

*We propose the use of motion parallax - the relative image motion of nearby image points - as a robust geometric cue for the computation of relative depth and surface curvature on specular surfaces and at extremal boundaries.*

## PARALLAX IS A ROBUST CUE

A robot vehicle moving under visual guidance needs to compute approximate geometry of obstacles in its environment. It is unreasonable to assume that egomotion is known to the sort of precision that is available for a camera mounted on a high quality robot arm. Generally a nominal estimate for egomotion is available. One possibility is to refine this estimate using optic flow data (Harris, 1987). Alternatively, the problem can be turned on its head: what geometric information remains stable under perturbation of assumed egomotion? This question has been addressed by Koenderink and van Doorn (1977), Nelson and Aloimonos (1988) and Verri et al. (1989), in the case of continuous motion fields and, in the domain of stereoscopic vision, by Weinshall (1990). Part of the answer, we claim, lies in the use of motion parallax as a geometric cue. Motion parallax, which is a relative measure of the positions of two points, can be very much more robust as a cue than the absolute position of a single point. This is true for computation of relative depth, curvature on specular surfaces and curvature on extremal boundaries.

*The authors acknowledge the support of SERC, IBM UK Scientific Centre and Esprit BRA 3274 (FIRST).

## RELATIVE DEPTH

Longuet-Higgins and Prazdny (1980) showed that whereas the structure of an image motion field is generally confounded by viewer rotation, motion parallax does not suffer that problem. Using the convention of Maybank (1985) that the image is projected onto a unit sphere through its centre, a point $\mathbf{r}$ on a visible surface projects to a vector $\mathbf{Q}(t)$ on the image sphere:

$$\mathbf{r} = \mathbf{v}(t) + \lambda(t)R(t)\mathbf{Q}(t), \qquad (1)$$

where, at time $t$, $\mathbf{v}(t)$ is the position of the viewer, $\lambda(t)$ is the distance along the ray from the viewer to the point $\mathbf{r}$ and $R(t)$ is a rotation operator describing the orientation of the camera frame relative to the world frame. The motion is assumed rigid, involving translational motion $\mathbf{U}$ and rotational motion $\mathbf{\Omega}$ defined by:

$$\mathbf{U} = \dot{\mathbf{v}} \text{ and } (\mathbf{\Omega}\times) = \dot{R}$$

where $\dot{}$ denotes differentiation with respect to time and $\times$ denotes vector product. The image motion of the projection of $\mathbf{Q}$ of the point $\mathbf{r}$ is then

$$\dot{\mathbf{Q}} = \frac{1}{\lambda}(\mathbf{U} \times \mathbf{Q}) \times \mathbf{Q} + \Omega \times \mathbf{Q}. \qquad (2)$$

The rotational contribution is apparent, and hard to extricate, although numerous solutions to that problem have, of course, been devised.

One simple solution involves parallax. A second point $\mathbf{Q}^*$ is introduced which, instantaneously, coincides with the first, so that $\mathbf{Q}(0) = \mathbf{Q}^*(0)$. Then the relative displacement of the two image points is $\boldsymbol{\Delta} = \mathbf{Q} - \mathbf{Q}^*$ and the *parallax* is its temporal derivative $\dot{\boldsymbol{\Delta}}$. It is straightforward to show that

$$\dot{\boldsymbol{\Delta}} = \left(\frac{1}{\lambda} - \frac{1}{\lambda^*}\right)(\mathbf{U} \times \mathbf{Q}) \times \mathbf{Q}, \qquad (3)$$

which is independent of rotation. This means that *relative inverse depth* $1/\lambda - 1/\lambda^*$ can be computed from parallax more *robustly* than absolute depth can be recovered from image motion. It is, theoretically, entirely immune to errors in estimated viewer rotation. In practice, of course, there is a residual sensitivity owing to the fact that $\mathbf{Q}, \mathbf{Q}^*$ will not coincide exactly at the instant of measurement.

Rieger and Lawton (1985) have implemented a scheme using motion parallax for robust estimation of relative depth and direction of translation from real image sequences. They also shown how the results degrade with increasing separation of the 2 points, $\mathbf{Q}, \mathbf{Q}^*$.

## CURVATURE OF SPECULAR SURFACES

It has been shown (Blake and Brelstaff, 1988) that parallax of a specularity – relative motion of a highlight and nearby surface feature – is a robust cue for surface curvature. It is more robust than stereo-based curvature estimates because spatial derivatives need not be computed in the specular case. It appears that human vision is capable of using this cue (Blake and Bülthoff, 1990). An earlier qualitative model dealt only with creation and annihilation of specularities at parabolic lines (Koenderink and van Doorn, 1980). This is structurally a very robust cue but somewhat sparse. Specular parallax cues are more common, being stable with respect to viewpoint. Surface curvature in the form of the *Weingarten map* $W$ (Thorpe 1979) is constrained by observed parallax $\dot{\mathbf{\Delta}}$:

$$2\lambda^2 W \mathcal{P}_1 \dot{\mathbf{\Delta}} = -\mathcal{P}_2 \mathbf{U}. \quad (4)$$

Here $\mathcal{P}_1$ and $\mathcal{P}_2$ are dimensionless projection operators involving combinations of projections into image and surface tangent planes. A derivation is presented in (Blake et al,1988,1990). The details are not important, but suffice it to say that $\mathcal{P}_1$ and $\mathcal{P}_2$ reduce to identity operators for the special case of fronto-parallel viewing. Note that this equation is correct in the limit that the source and viewer are distant from the surface. Thus the singular effects of the crossing of the caustic by the viewer are neglected. However the full model is easily obtainable (Blake and Brelstaff, 1988).

Clearly the constraint is independent of, and hence robust to rotational motion $\mathbf{\Omega}$. If instead of parallax, the *absolute* image motion of the specularity were substituted the result would be sensitive to rotation and, even worse, sensitivity to translational motion $\mathbf{U}$ would be greatly enhanced. However, even with parallax there is a remaining, though linear and hence well-conditioned, dependence on viewer motion $\mathbf{U}$. There is also an implied (by $\mathcal{P}_1, \mathcal{P}_2$) dependence on estimated source position. This too can be eliminated by using a still more qualitative measure than parallax $\dot{\mathbf{\Delta}}$, namely the sign of its epipolar component (Bolles et al. 1987). This sign is sufficient for convex/concave discrimination, is independent of light source position and the magnitude of of translational velocity $\mathbf{U}$, depending only on the direction of $\mathbf{U}$ to define epipolar lines on the image sphere (Zisserman et al. 1989). Human observers are able to use this cue for disambiguation of reversible figures (Blake and Bülthoff 1990).

## SURFACE CURVATURE ON EXTREMAL BOUNDARIES

Building on earlier work (Marr 1977, Barrow and Tenenbaum 1978, Koenderink 1984, Giblin and Weiss 1987) we have developed a general model of the inference of surface curvature from the deformation of apparent contours (Blake and Cipolla 1989). As a viewer moves past a curved surface, apparent contours, the projections of extremal boundaries, move non-rigidly on the image sphere. Their motion and deformation can be used to characterise local surface shape including curvature, along the extremal boundaries. It can also be used, of course, to discriminate between fixed surface features and extremal boundares.

The characterisation turns out to be robust only if a parallax-based measure is used. Intuitively it is relatively difficult to judge, moving around a smooth, featureless object, how large the normal curvature across the extremal boundary is. However, this judgment is much easier to make for objects which have at least a few surface features. Under small viewer-motions, features are "sucked" over the extremal boundary, at a rate which depends on surface curvature. Our theoretical findings exactly reflect the intuition that the "sucking" effect is a reliable indicator of normal curvature, regardless of the exact details of the viewer's motion.

As before, parallax is defined as $\dot{\boldsymbol{\Delta}}$ where $\boldsymbol{\Delta} = \mathbf{Q} - \mathbf{Q}^*$, and now $\mathbf{Q}$ is a point on the apparent contour whereas $\mathbf{Q}^*$ is the image of a reference feature on the surface. It is assumed that the surface reference point lies close (in 3D space) to the extremal boundary. The normal curvature $\kappa$ of the surface *along the line of sight* can be computed from parallax as follows:

$$\frac{1}{\kappa} = \frac{\lambda^3}{(\mathbf{U}\cdot\mathbf{n})^2}\ddot{\boldsymbol{\Delta}}, \tag{5}$$

where $\mathbf{n}$ is the normal (in the tangent plane to the image sphere) to the apparent contour. Normal curvature $\kappa$, together with curvature of the apparent contour, is in fact sufficient to compute full surface curvature at a point (Blake and Cipolla 1989). Again the estimate of $\kappa$ is robust in the sense that it is independent of rotational velocity $\Omega$. Better than that, curvature estimates based on absolute image motion and acceleration is also sensitive to linear and rotational *acceleration*, something which is eliminated when the "rate of parallax" $\ddot{\boldsymbol{\Delta}}$ is used as above. The sign of $\kappa$ determines the "sidedness" of the extremal boundary – on which side of the image contour lies the curved surface. This qualitative sidedness cue depends only on knowledge of the *direction* of translational motion, to establish the epipolar lines, as before. Finally, ratios of normal curvatures are completely independent of *any* assumption about viewer motion! Terms depending on absolute depth and translational velocity are cancelled out in equation (5). As before, independence degrades as the assumption that the reference point is close to the extremal boundary is relaxed. The degradation can be theoretically predicted and is analysed in (Cipolla and Blake 90).

Results in figures 3,4,5 illustrate that the theoretical robustness of rate-of-parallax are borne out in practice. They show the sensitivity of the estimated radius of curvature, ($R = 1/\kappa$) at a point on an extremal boundary (figure 2) computed from known viewer motion when an error is introduced in the camera positions and orientations. Estimates of curvature not using parallax require precise knowledge of viewer motion (1 part in 1000). This sensitivity is reduced by an order of magnitude if parallax measurements are used instead. It is further reduced by another order of magnitude for ratios of parallax measurements. The residual sensitivity to rotation and velocity is, as predicted, due the finite separation of the point on the extremal boundary and the surface marking.

## CONCLUSIONS

We believe that the use of motion parallax in the recovery of surface geometry from surface texture, specularities and apparent contours represents a significant step in the development of practical techniques for robust, qualitative 3D vision.

We are currently working on the realtime implementation of algorithms using motion parallax for use in the active exploration of the 3D geometry of visible surfaces for navigation. Preliminary results involving the tracking of image contours and specularities; discrimination between fixed and extremal features; and the recovery of strips of surfaces in the vicinity of an extremal boundary or specularity are described in (Cipolla and Blake, 1990) and (Zisserman et al, 1990).

Figure 1. Sample of monocular image sequence of motion of specularities across the curved surface of a Japanese cup with viewer motion. 4 images of the sequence (magnified window) showing the relative motion (parallax) of a point specularity (**Q**) and fixed surface marking (**Q***) are shown in Figure 1b. Parallax measurements can be used to determine surface curvature and normal along the path followed by the specularity as the viewer moves. A more qualitative measure is the sign of the epipolar component of the parallax measurement. With viewer motion the specularity moves in opposite directions for concave and convex surfaces.

Figure 2. Sample of monocular image sequence showing the image motion of apparent contours with viewer motion. 4 images of the sequence (magnified window) showing the relative motion between the apparent contour (projection of the extremal boundary) and a nearby surface marking (shown as a cross) are shown in Figure 2b. The relative image motion as the feature moves away from the extremal boundary can be used for the robust estimation of surface curvature.

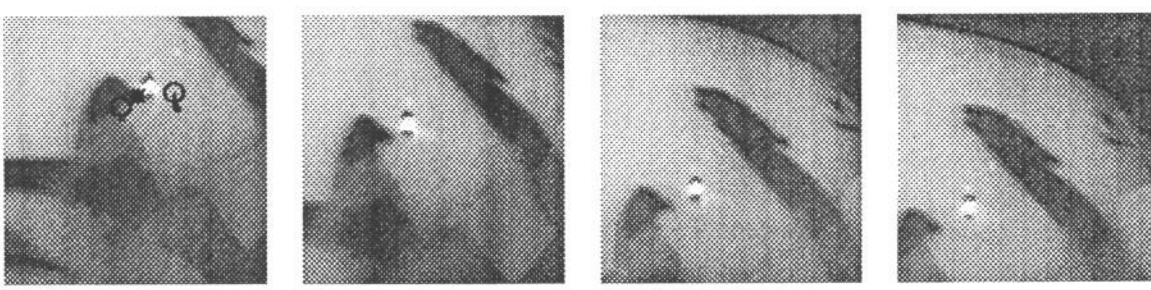


Figure 1b. Relative motion of specularity and nearby surface marking

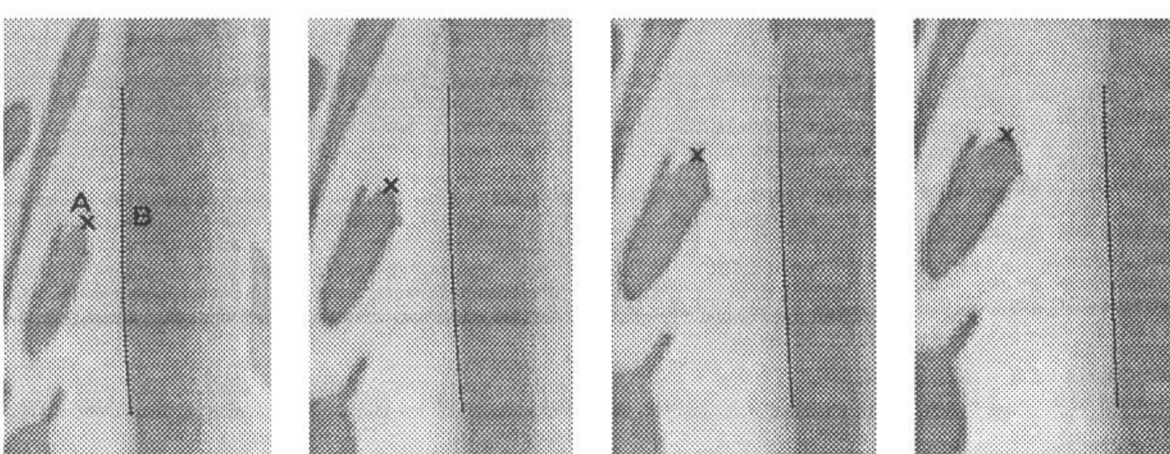


Figure 2b. Relative motion of an apparent contour

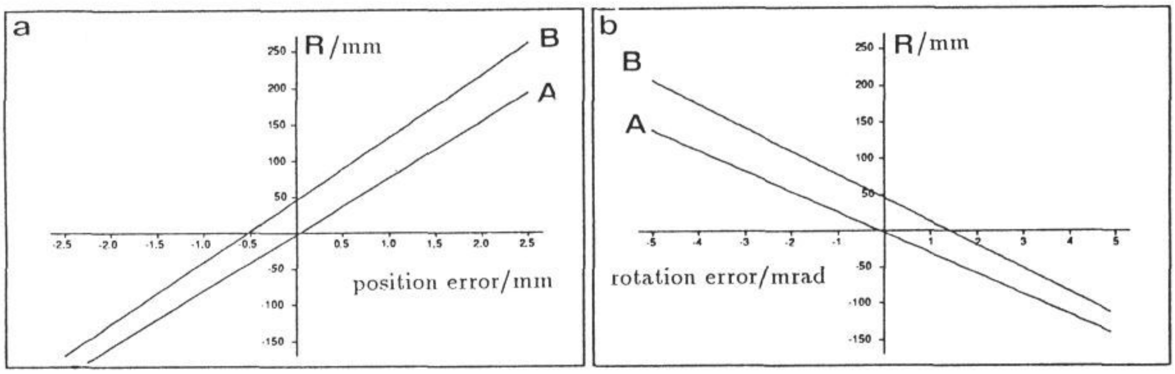


Figure 3. Sensitivity of curvature estimated from absolute measurements to errors in motion.
The radius of curvature (mm) for both a point on a surface marking (A) and a point on an extremal boundary (B) is plotted against error in the estimate of position (a) and orientation (b) of the camera for view 2. The estimation is very sensitive and a perturbation of 1mm in position produces an error of 190% in the estimated radius of curvature for the point on the extremal boundary. A perturbation of 1mrad in rotation about an axis defined by the epipolar plane produces an error of 70%.

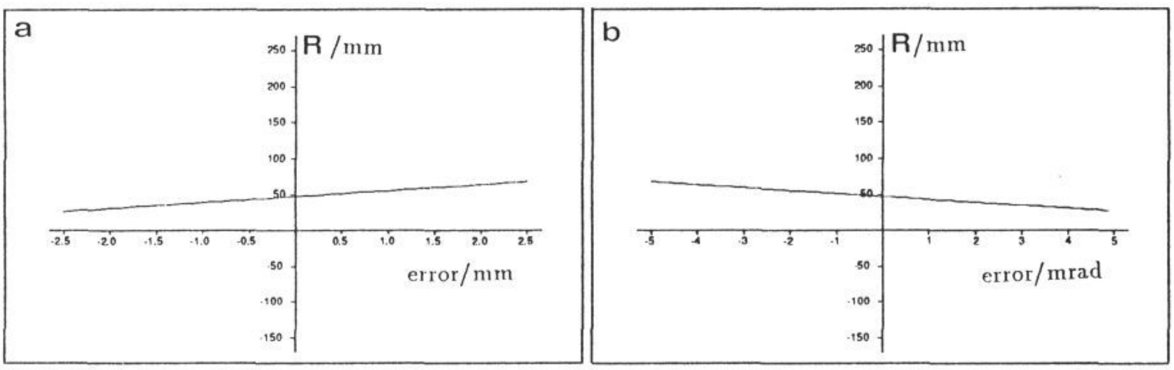


Figure 4. Sensitivity of *differential curvature*
The difference in radii of curvature between a point on the extremal boundary and the nearby surface marking is plotted against error in the position (a) and orientation (b) of the camera for view 2. The sensitivity is reduced by an order of magnitude to 17% per mm error and 8% per mrad error respectively.

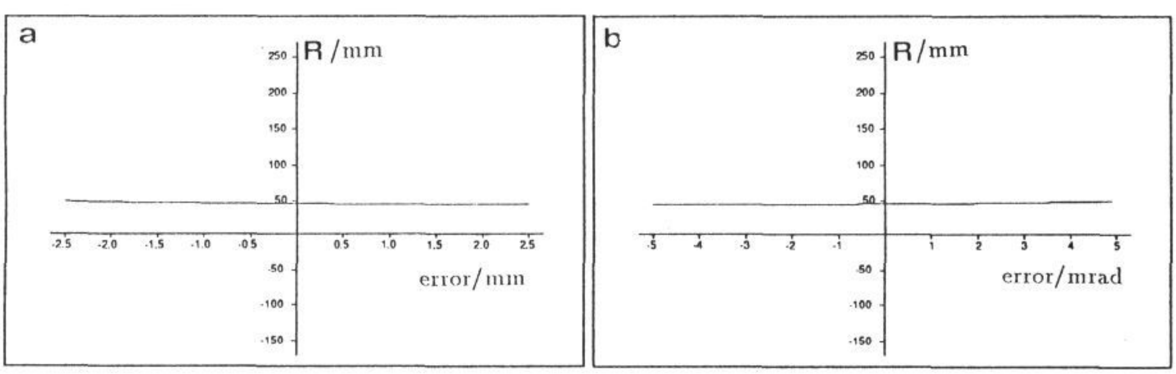


Figure 5. Sensitivity of ratio of *differential curvatures*
The ratio of *differential curvatures* measurements made between 2 points on an extremal boundary and the same nearby surface marking is plotted against error in the position (a) and orientation (b) of the camera for view 2. The sensitivity is further reduced by an order of magnitude to 1.5% error for a 1mm error in position and 1.1% error for 1mrad error in rotation. The vertical axes are scaled by the actual curvature for comparision with figures 3 and 4.

## References


[1] **H.G. Barrow and J.M. Tenenbaum.** *Recovering Intrinsic Scene Characteristics from Images.* A.I Center Technical Report 157, SRI International, (1978).

[2] **A. Blake and R. Cipolla.** *Robust Estimation of Surface Curvature from Deformation of Apparent Contours.* Technical Report OUEL 1787/89, University of Oxford, 1989. Also *Proc. 1st European Conf. Computer Vision,* Springer-Verlag, (1990).

[3] **A. Blake and H. Bülthoff.** "Does the brain know the physics of specular reflection?" *Nature* 343, 165-168 (1990).

[4] **A. Blake and H. Bülthoff.** "Shape from Specularities: Computation and Psychophysics" *(to be published).*

[5] **A. Blake and G.J. Brelstaff.** "Geometry from Specularities." In *Proc. 2nd Int. Conf. Computer Vision*, 394–403 (1988).

[6] **R.C. Bolles, H.H. Baker, and D.H. Marimont.** " Epipolar-plane image analysis: an approach to determining structure." *International Journal of Computer Vision*, vol.1:7–55, (1987).

[7] **R. Cipolla and A. Blake.** "The dynamic analysis of apparent contours." *Technical Report OUEL 1828/90*, University of Oxford, (To appear in *Proc. 3rd Int. Conf. Computer Vision*, Osaka)(1990).

[8] **P. Giblin and R. Weiss.** " Reconstruction of surfaces from profiles." In *Proc. 1st Int. Conf. on Computer Vision*, 136–144, (1987).

[9] **C.G. Harris.** "Determination of ego - motion from matched points." In *3rd Alvey Vision Conference*, 189–192, (1987).

[10] **J.J. Koenderink and A.J. van Doorn,** "How an ambulant observer can construct a model of the environment from the geometric structure of the visual inflow." *Kibernetic 1977 G.*, eds. G. Hauske and E. Butendant, Oldenbourg, München (1977).

[11] **J.J. Koenderink and A.J. van Doorn,** "Photometric invariants related to solid shape." *Optica Acta,* 27, 7, 981-996 (1980).

[12] **J.J. Koenderink.** "What does the occluding contour tell us about solid shape?" *Perception*, 13:321–330, (1984).

[13] **H.C. Longuet-Higgins and K. Pradzny.** "The interpretation of a moving retinal image." *Proc. Royal Soc. London*, B208:385–397, (1980).

[14] **D. Marr.** " Analysis of occluding contour." *Proc. Royal Soc. London*, 197:441–475, (1977).

[15] **S.J. Maybank.** "The angular velocity associated with the optical flow field arising from motion through a rigid environment." *Proc. Royal Soc. London*, A401:317–326, (1985).

[16] **R.C. Nelson and J. Aloimonos,** "Using flow field divergence for obstacle avoidance: towards qualitative vision." *Proc. 2nd Int. Conf. Computer Vision*, (1988).

[17] **J.H. Rieger and D.L. Lawton.** "Processing differential image motion." *J. Optical Soc. of America*, vol A2, no. 2, 354–359 (1985).

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

[19] **A. Verri, F. Girosi and V. Torre.** "Mathematical properties of the two-dimensional motion field: from singular points to motion parameters." *JOSA* vol A6, no. 5, 698–712, (1989).

[20] **D. Weinshall.** " Qualitative depth from stereo with applications." *Computer Vision, Graphics, and Image Processing,* 49, 222-241, (1990).

[21] **A. Zisserman, P. Giblin and A. Blake.** "The information available to a moving observer from specularities." *Image and Vision Computing,* 7, 38–42, (1989).

[22] **A. Zisserman and A. Blake** "Shape from tracked specular motion." *Technical Report, University of Oxford (in preparation)*(1990)