INVERSE PERSPECTIVE OF A TRIANGLE: NEW EXACT AND APPROXIMATE SOLUTIONS

Daniel DeMenthon and Larry S. Davis
Computer Vision Laboratory
Center for Automation Research
University of Maryland, College Park, MD 20742 USA

1 Introduction

One of the techniques in model-based pose estimation of 3D objects consists of locating "interest points" on models of the objects, detecting these points in the image, and matching subsets of these image points against subsets of the interest points of the models. Valid matches determine a similar object pose, which can be found by clustering techniques.

Solving for the position and orientation of an object knowing the images of n points at known locations on the object is called the n-point perspective problem (see [5] for a review and a four points solution). The three-point problem, also called the triangle pose problem [9], has been solved in various ways. A review of the major direct solutions for three points under exact perspective is provided in [4]. Another direct solution is described in this paper. Computation speed is important when many or all possible combinations of triples of image feature points and triples of model interest points are considered. Direct methods require quite a few floating point operations. Some researchers have proposed faster methods based on scaled orthographic projection [8,6,10]. We introduce an alternative approximation which we call orthoperspective, a *local* scaled orthographic projection using a plane normal to one of the rays, which causes smaller errors for off-center images. The angular terms of the pose of a given triangle depend then only on two parameters of the image triangle, and these can be precomputed in a two dimensional lookup table resulting in a very fast pose estimation algorithm.

2 Preliminaries

Figure 1 shows a triangle $M_1M_0M_2$ of known dimensions. Its sides M_0M_1 and M_0M_2 have known lengths D_1 and D_2 and the angle $M_1M_0M_2$ at vertex M_0, which we call α, is also known. This triangle projects on the image plane in $m_1m_0m_2$. The projection on the left drawing of the figure is an exact perspective projection. The projection on the right is an approximation that we call *orthoperspective* (Section 4). With both projections, the position of the triangle in space is completely determined by the location of its image, by the range R_0 of the vertex M_0 along the line of sight Om_0, and by the angles θ_1 and θ_2 that the sides M_0M_1 and M_0M_2 make with the line of sight Om_0. Once the three unknowns R_0, θ_1 and θ_2 are determined, the triangle vertices are given by
$\overrightarrow{OM_0} = R_0\overrightarrow{e_0}$, $\overrightarrow{OM_1} = [R_0 + D_1(\cos\theta_1 - \frac{\sin\theta_1}{\tan\gamma_1})]\overrightarrow{e_0} + D_1\frac{\sin\theta_1}{\sin\gamma_1}\overrightarrow{e_1}$, $\overrightarrow{OM_2} = [R_0 + D_2(\cos\theta_2 - \frac{\sin\theta_2}{\tan\gamma_2})]\overrightarrow{e_0} + D_2\frac{\sin\theta_2}{\sin\gamma_2}\overrightarrow{e_2}$
where $\overrightarrow{e_0}$, $\overrightarrow{e_1}$ and $\overrightarrow{e_2}$ are the known unit vectors along the lines of sight Om_0, Om_1 and Om_2, and γ_1 and γ_2 are the angles between $\overrightarrow{e_0}$ and $\overrightarrow{e_1}$ and $\overrightarrow{e_0}$ and $\overrightarrow{e_2}$ respectively.

3 Triangle Pose Solutions for Exact Perspective

In the exact perspective projection (Figure 1, left), the law of sines for triangle OM_0M_1 yields

$$\frac{\sin(\theta_1 - \gamma_1)}{R_0/D_1} = \sin\gamma_1 \tag{1}$$

A similar equations is written for θ_2. Dividing the two equations eliminates R_0

$$\frac{\sin(\theta_1 - \gamma_1)}{\sin(\theta_2 - \gamma_2)} = K, \text{ with } K = \frac{\sin\gamma_1/D_1}{\sin\gamma_2/D_2} \tag{2}$$

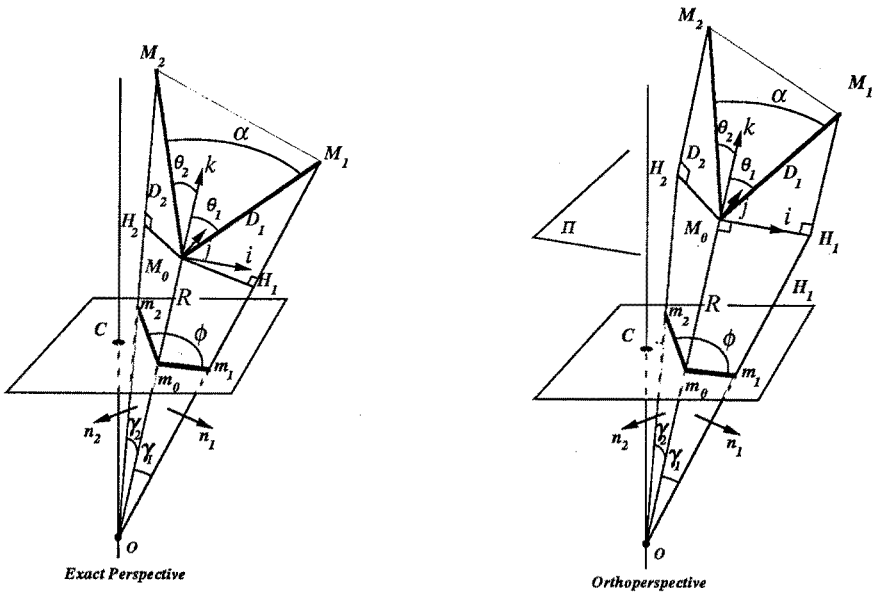

Figure 1: Exact perspective and orthoperspective for a triangle

K is a combination of known parameters. We call K the *foreshortening ratio*, because it is a ratio of foreshortening measures for the two sides of the triangle.

A second relation between θ_1 and θ_2 is obtained by expressing the cosine of α, the angle at vertex M_0 of the world triangle, as a dot product [7]:

$$\cos\alpha = \sin\theta_1 \sin\theta_2 \cos\phi + \cos\theta_1 \cos\theta_2 \tag{3}$$

where ϕ is the angle between the plane containing O, m_0 and m_1 and the plane containing O, m_0 and m_2. This equation also applies to the orthoperspective approximation (Figure 1, right).

Eliminating θ_2 yields a fourth degree equation in $\sin^2(\theta_1 - \gamma_1)$. The coefficients of the equation and details are given in [3]. A total of two triangle poses compatible with the original equations is generally found. Representative pose surfaces are presented graphically in Figure 2. The top row of Figure 2 shows the three diagrams of one set of solutions (R, θ_1, θ_2), the bottom row the second set. However, this separation into two sets is artificial. The step discontinuity of the surfaces of one set matches the step discontinuity of the surfaces of the other set so that the resulting surface does not possess any discontinuity. For example, the general shape of the surface giving θ_1 is the double valued surface shown in Figure 3; the surface for θ_2 is similar. The surface for R_0 also combines two layers which cross each other, but are very close together.

A startling consequence is that if the image of a triangle deforms smoothly (corresponding to a smooth 3-D motion of the 3D triangle) so that the final image is identical to the initial image, then the corresponding world triangle has not necessarily returned to its original position. It would require a second cycle of the image deformation sequence to bring back the world triangle to its initial position (Figure 3, right). This occurs when the trajectory corresponding to the triangle poses on the solution surface describes a closed cycle around the point $(K = 1, \phi = \alpha)$.

4 Orthoperspective

Figure 1 (right) describes the orthoperspective approximation. A plane Π through point M_0 perpendicular to the line of sight Om_0 is considered. Points M_1 and M_2 project onto this plane at H_1 H_2 using an orthogonal projection; the image points m_0, m_1 and m_2 are the perspective images of M_0, H_1 and H_2.

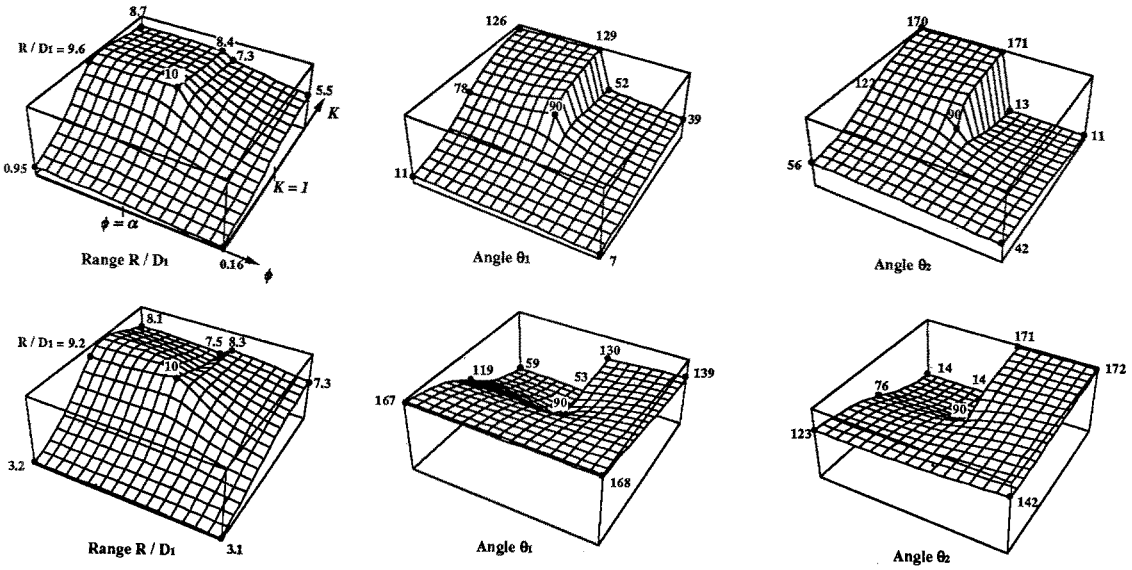

Figure 2: Solutions for triangle pose using exact perspective for $\tan\gamma_1 = 0.1$, $D_1/D_2 = 0.5$ and $\alpha = \pi/4$. Range of K from $1/4$ to 4. Range of $\tan(\phi/2)$ from one quarter to four times $\tan(\alpha/2)$.

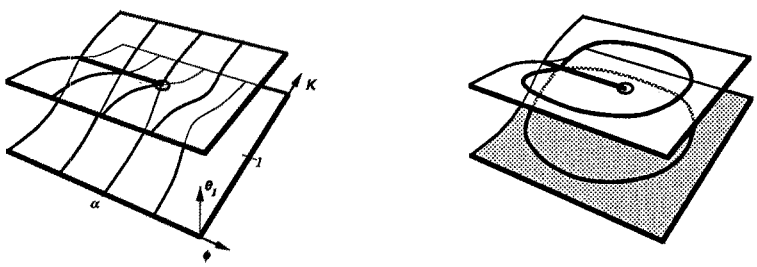

Figure 3: Solution surface for angle θ_1 or θ_2. Two cycles may be necessary in the image for the world triangle to return smoothly to its original position

Notice that if we rotate the camera around the center of projection O to bring the optical axis to the line of sight Om_0 (a *standard rotation*[7]), the image plane becomes parallel to the plane Π on which we performed the orthogonal projection. After this camera rotation, orthoperspective is simply a scaled orthographic projection. The camera rotation removes the dependence of the construction on the offset of the world object from the optical axis. For this reason orthoperspective is a better approximation to perspective than classical scaled orthographic projection for image elements which are not centered in the image plane [1,3].

In the right triangle OM_0H_1 of Figure 1 (right)

$$\frac{\sin\theta_1}{R_0/D_1} = \tan\gamma_1 \tag{4}$$

Writing a similar equation for θ_2 and dividing the two equations eliminates R_0, yielding

$$\frac{\sin\theta_1}{\sin\theta_2} = K, \text{ with } K = \frac{\tan\gamma_1/D_1}{\tan\gamma_2/D_2} \tag{5}$$

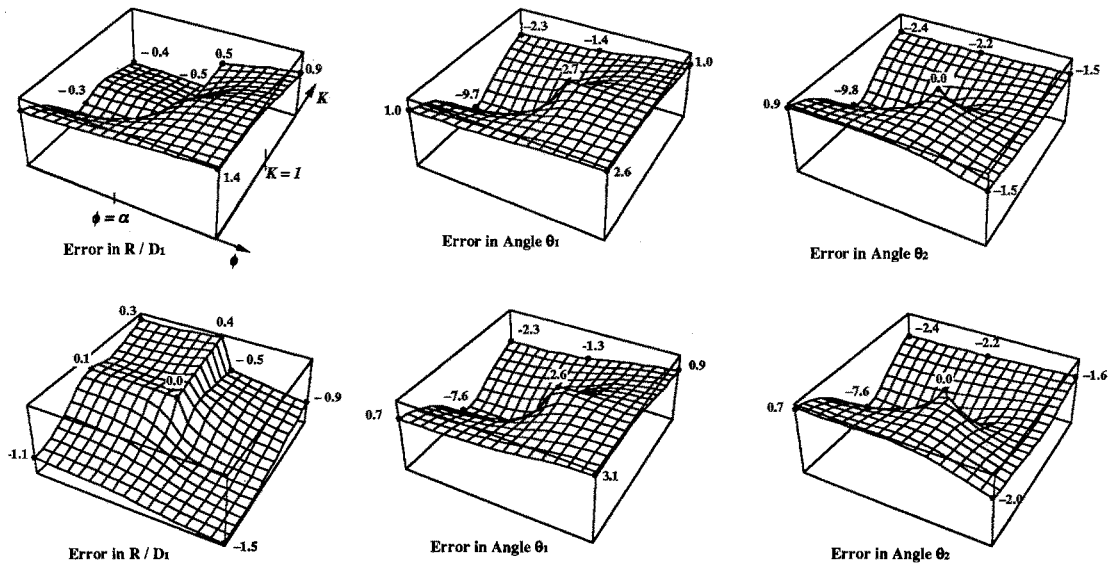

Figure 4: Errors for triangle pose using orthoperspective

Eliminating θ_2 between Equation 5 and Equation 3 yields a second degree equation in $\sin^2\theta_1$

$$\sin^2\phi \sin^4\theta_1 - (K^2 - 2K\cos\alpha\cos\phi + 1)\sin^2\theta_1 + K^2\sin^2\alpha = 0 \quad (6)$$

Equation 6 always has two positive solutions, but only the smaller solution has magnitude less than 1 and can thus be equated to the square of the sine of an angle. The single solution for $\sin\theta_1$ results in two complementary solutions for θ_1 corresponding to mirror image directions of M_0M_1. The corresponding value for $\sin\theta_2$ obtained from Equation 5 also yields two complementary solutions for θ_2. Equation 3 has to be used to discard two of the four combinations of θ_1 and θ_2. Finally, the distance R_0 of the vertex M_0 from the center of projection is computed using Equation 4. This yields a single solution for R_0. Thus the two solution triangles share a common vertex M_0, whereas in exact perspective the two solution triangles have slightly distinct M_0 vertices. Therefore a single-valued surface is obtained for R_0/D_1. Otherwise the diagrams for the approximation are almost identical to those obtained for exact perspective (Figure 2) and are omitted for lack of space. They can be found in [3].

5 Comparison of Triangle Pose obtained by Exact and Approximate Perspective

We numerically compare the results provided by the exact perspective and the approximate perspectives for the parameter values used in Figure 2. The results are plotted in Figure 4. The numbers given on the 3D plots of Figure 4 are nondimensional errors in R_0/D_1 and angular errors in degrees. The following characteristics of the errors can be observed:

1. The largest range errors occur for small values of K, i.e. when the side M_0M_1 is much more foreshortened than the other side (M_0M_1 and M_0M_2 do not play symmetric roles in these diagrams because the size of the image of M_0M_1 is fixed since γ_1 and K are fixed).
2. The largest angular errors occur along the line ($K = 1, \phi < \alpha$) and reach almost 10 degrees in these valleys of the surfaces. At the edges of the diagrams the angular errors are only around two or three degrees. Other error diagrams are provided in [3].

6 Discussion

One advantage of using approximate perspective is that small lookup tables can be constructed. The solutions for the angles θ_1 and θ_2 depend only on two parameters, K and ϕ. For each triangle of features of an object a two dimensional table can be generated, in which the possible values for θ_1 and θ_2 are stored for a range of the observable parameters K and ϕ. From an image of the triangle, the parameters K and ϕ are calculated and the angles are read from the table. Then the range R_0 is obtained using Equation 4. On a 16K Connection Machine without floating point processors, this technique programmed in StarLisp provides a pose estimate of a single polyhedra with 40 triangles with a smooth background in around one second [2]. This involves matching all pairs of image and model triangles, and clustering the resulting set of pose estimates.

The error diagrams shown in Figure 4 are useful for increasing the accuracy of pose estimation by lookup tables. The table cells for which the approximate pose is very different from the exact pose (for example for K close to one and $\phi < \alpha$, as seen in Figure 4) can be flagged, and the pairs of model and image triangles corresponding to these table cells can be disregarded.

Acknowledgements

The support of the Defense Advanced Research Projects Agency and the U.S. Army Engineer Topographic Laboratories under Contract DACA76-88-C-0008 (DARPA Order No. 6350) is gratefully acknowledged.

References

[1] J. Aloimonos and M. Swain, "Paraperspective Projection: Between Orthography and Perspective", Center for Automation Research CAR-TR-320, University of Maryland, College Park, May 1987.

[2] L.S. Davis, D. DeMenthon, T. Bestul, D. Harwood, H.V. Srinivasan, S. Ziavras, "RAMBO—Vision and Planning on the Connection Machine", Proc. 1989 DARPA Image Understanding Workshop, 631-639.

[3] D. DeMenthon, L.S. Davis, "New Exact and Approximate Solutions of the Three-Point Perspective Problem", Center for Automation Research CAR-TR-471, University of Maryland, College Park, November 1989.

[4] R.M. Haralick, "The Three Point Perspective Pose Estimation Problem", Internal Note, Dept. of Electrical Engineering, FT-10, University of Washington.

[5] R. Horaud, B. Conio, O. Leboulleux, B. Lacolle, "An Analytical Solution for the Perspective Four-Point Problem", to appear in *Computer Vision, Graphics, and Image Processing*, Academic Press, 1989.

[6] D. Huttenlocher and S. Ullman, "Recognizing Solid Objects by Alignment", Proc. 1988 DARPA Image Understanding Workshop, 1114-1122.

[7] K. Kanatani, "Constraints on Length and Angle", *Computer Vision, Graphics and Image Processing*, 41, 1988, 28–42.

[8] Y. Lamdan, J.T. Schwartz, and H.J. Wolfson, "On Recognition of 3-D Objects from 2-D Images", Proc. 1988 IEEE Int. Conf. on Robotics and Automation, 1407-1413.

[9] S. Linnainmaa, D. Harwood, and L.S. Davis, "Pose Determination of a Three-Dimensional Object using Triangle Pairs", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 10, 1988, 634–647.

[10] D.W. Thompson and J.L. Mundy, "Model-Directed Object Recognition on the Connection Machine", Proc. DARPA Image Understanding Workshop, 1987, 98–106.