Exact and Approximate Solutions of the Perspective-Three-Point Problem

Daniel DeMenthon and Larry S. Davis

Abstract—**Model-based pose estimation techniques that match image and model triangles require large numbers of matching operations in real-world applications. We show that by using approximations to perspective, 2–D lookup tables can be built for each of the triangles of the models. An approximation called "weak perspective" has been applied previously to this problem; we consider two other perspective approximations: paraperspective and orthoperspective. These approximations produce lower errors for off-center image features than weak perspective.**

Index Terms—**Inverse perspective of triangles, lookup tables, orthoperspective, paraperspective, perspective-n-point problem, pose computation, robot vision, scaled orthographic projection, 3-D object recognition, weak perspective.**

I. Introduction

Model-based pose estimation of 3-D objects from a single view has been a major focus of research in robot vision. One of the techniques commonly used consists of locating "interest points" [9] 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. Knowing that a given subset of points of an object is projected on the image into a given subset of points determines constraints on the object location in space. The validity of a match is estimated either by verifying that the object at the found location does project onto the given image [7] or by checking whether many other matches determine a similar location.

Fischler and Bolles [4] coined the term "Perspective-n-point problem" for the problem of finding the position and orientation of an object from the images of n points at known locations on the object. Their paper provides solutions and useful insights to this problem (see also [6] for a review of solutions and a solution for four points). The question of how many points should be taken as subsets in an object pose or recognition system has generated some interest. Many researchers have considered three point solutions [10], [12], [9], [7] because it is the smallest subset that yields a finite number of object poses, generally two poses (four poses in some image configurations). The perspective-three-point problem, which is also called the triangle pose problem [10], has been solved in many different ways. A review of the major direct solutions for three points under exact perspective is provided in [5]. Yet another direct solution is described in this paper with parameters that are convenient for comparison with the proposed approximate methods. The drawback to solving the triangle pose problem with exact perspective is that it is slow, requiring quite a few floating-point operations. The speed of computation of the triangle pose is of importance in the overall performance of an object pose or recognition system since it is performed for many or all possible combinations of triples of both image feature points and model interest points. The problem is significant even for present massively parallel machines because in real-world problems, the number of combinations is so large that each processing element would have to serially process several model-image triplet combinations.

Some researchers have proposed simpler computational solutions that use an approximation of perspective—the scaled orthographic perspective (which is also called weak perspective [13])—which leads to simpler computations of the triangle pose. For example, object recognition systems such as the ORA system of Huttenlocher [7] and the geometric hashing system of Lamdan *et al.* [9] have relied on the computation of the pose of triangles by weak perspective approximations. Other object recognition systems have also used weak perspective approximation on nontriangular combinations of features [14], [15]. However, the authors do not discuss the errors introduced by the approximation in comparison with an exact perspective approach. This is one of the topics of this paper. We also show that weak perspective is only one possible approximation for simplifying the problem and is not necessarily the best. Two other approximations (paraperspective [1], [11] and a new approximation that we call *orthoperspective*) are found to yield similar expressions in the proposed framework. Orthoperspective is a perspective approximation that yields simple expressions in this type of problem and may also be useful in other computer vision applications. It is *a local scaled orthographic projection using a plane normal to one of the rays*. It is, in fact, a transformation that is equivalent to a virtual camera rotation recentering the image in the image plane [8], followed by a weak perspective, followed by the inverse virtual camera rotation. Of the three approximations, weak perspective is the least attractive with respect to pose errors because its performance deteriorates for locations far from the image center. Paraperspective and orthoperspective do not have this drawback. Finally, we show that the angular terms of the approximated pose of a given triangle can be expressed simply as functions of two parameters of the image triangle and, therefore, can be precomputed in a 2-D lookup table, resulting in a very fast pose estimation algorithm.

Manuscript received August 8, 1990; revised January 8, 1992. This work was supported by the Defense Advanced Research Projects Agency and the U.S. Army Engineer Topographic Laboratories (ARPA Order No. 6350) under Contract DACA76-88-C-0008. Recommended for acceptance by Associate Editor C. Brown.

The authors are with the Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, MD 20742-3411.

IEEE Log Number 9203751.

II. Exact and Approximate Perspective: The Two-Point Problem

We consider a simple "perspective-two-point problem" (perspective of line segments) to introduce three perspective approximations: weak perspective, paraperspective, and orthoperspective. We express the constraint imposed on the position of a line segment of known length D_1 with end points M_0 and M_1 when the images m_0 and m_1 of the end points are given. The position of the segment could be expressed by the two parameters R_0 and R_1, which are distances of the end points M_0 and M_1 from the projection center along the lines of sight of the images m_0 and m_1. An alternative choice, which leads to simpler expressions and to insights on perspective approximations, is the pair (R_0, θ_1), where θ_1 is the angle that the segment M_0M_1 makes with the line of sight Om_0 (Fig. 1). Other analyses benefiting from the use of these parameters can be found in [8]. The angle between the lines of sight Om_0 and Om_1 is called γ_1.

A. Exact Perspective

In the exact perspective projection (Fig. 1), M_1 is on the line of sight Om_1. In triangle OM_0M_1, the angle at the vertex O is γ_1, and the angle at vertex M_1 is $\theta_1 - \gamma_1$. The law of sines for this triangle

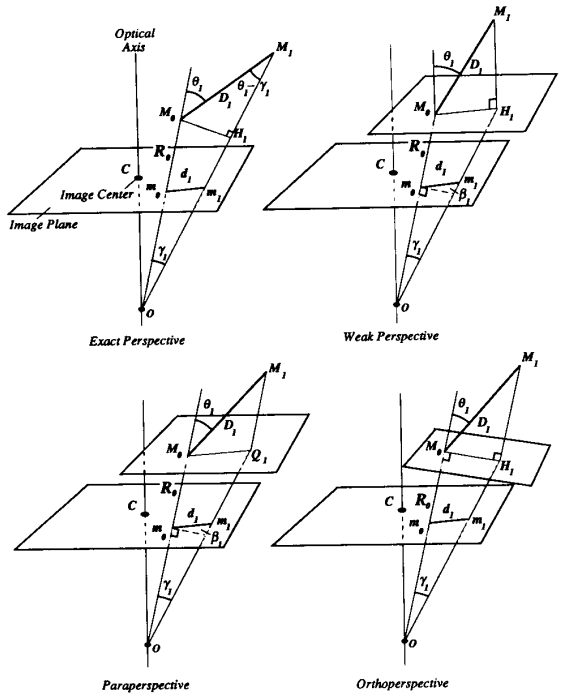

Fig. 1. Exact perspective and perspective approximations for a pair of points.

yields the following relation:

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

Of course, this relation does not define the position of the pair of points M_0 and M_1 uniquely. One more constraint is necessary, for example, an angular constraint with another pair (Section V), or the requirement that the segment be horizontal.

B. Weak Perspective

In the weak perspective approximation, a plane Π parallel to the image plane is drawn through one of the points of the pair, say, M_0 (Fig. 1). Point M_1 is projected to H_1 by orthographic projection onto this plane, and m_1 is the perspective image of H_1. The image segment m_0m_1 is the image of M_0H_1, and since these two segments are parallel, the image m_0m_1 is just a scaled version of the orthographic projection M_0H_1 of M_0M_1; hence, this approximation is sometimes called a "scaled orthographic projection."

In this configuration, the angle θ_1 is defined as the angle between M_0M_1 and the optical axis OC. We also define the angle β_1 between m_0m_1 and a plane perpendicular to Om_0. This angle is calculated from the dot product of Om_0 and m_0m_1. Applying the law of sines in triangles OM_0H_1 and $M_0H_1M_1$ yields

$$\frac{\sin\theta_1}{R_0/D_1} = \frac{\sin\gamma_1}{\cos(\gamma_1 + \beta_1)}. \tag{2}$$

C. Paraperspective

In this approximation, as in weak perspective, a plane Π through point M_0 parallel to the image plane is considered. However, point M_1 projects onto this plane to Q_1 in an *oblique* projection parallel to the line of sight Om_0 such that m_1 is the perspective image of Q_1. Applying the law of sines in triangles OM_0Q_1 and $M_0Q_1M_1$ yields

$$\frac{\sin\theta_1}{R_0/D_1} = \frac{\sin\gamma_1\cos\beta_1}{\cos(\gamma_1 + \beta_1)}. \tag{3}$$

D. Orthoperspective

In this approximation, a plane Π through point M_0 perpendicular to the line of sight Om_0 is considered. Point M_1 projects onto this plane at H_1 using an orthogonal projection such that m_1 is the perspective image of H_1. Considering triangles OM_0H_1 and $M_0H_1M_1$ yields

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

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, the image plane becomes parallel to the plane Π on which we performed the orthogonal projection. This type of camera rotation is called a *standard rotation* by Kanatani [8]. After this camera rotation, orthoperspective is simply a scaled orthographic projection. Therefore, orthoperspective is equivalent to 1) a virtual rotation of the camera around the center of projection to make a chosen line of sight (Om_0 in our line segment example) coincide with the optical axis, 2) a scaled orthographic projection of the recentered image, and 3) the inverse camera rotation that brings the camera back to its original position. The camera rotations remove the dependency on the offset of the projection in the image plane. For this reason, orthoperspective is a better approximation to perspective than weak perspective for image elements that are not centered in the image plane.

When is the expression for orthoperspective a good approximation for the exact perspective expression? Equation (1) for exact perspective can be rewritten as

$$\frac{\sin\theta_1}{R_0/D_1} = \frac{\tan\gamma_1}{1 - \tan\gamma_1/\tan\theta_1}.$$

This equation approaches the orthoperspective equation when the term $\tan\gamma_1/\tan\theta_1$ becomes small compared with 1. This occurs when the world line segment is nearly perpendicular to its lines of sight ($\tan\theta_1$ large) and/or when its size is small with respect to its distance from the camera ($\tan\gamma_1$ small).

Orthoperspective and paraperspective yield very similar results. One way to understand this is to rewrite paraperspective equation (3) as

$$\frac{\sin\theta_1}{R_0/D_1} = \frac{\tan\gamma_1}{(1 - \tan\gamma_1\tan\beta_1)}.$$

In the denominator, the term $\tan\gamma_1\tan\beta_1$ is small compared with 1 since it is the product of two terms that are both generally small. The expression is then close to

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

which is the expression obtained for orthoperspective.

III. The Perspective-Three-Point Problem

The simple expressions developed in Section II for the perspective-two-point problem are now combined to solve the perspective-three-point problem. The points m_0, m_1, and m_2 are the known images of world points M_0, M_1, and M_2 with known relative geometry but unknown positions along their lines of sight (Fig. 2). The line segment m_0m_1 in the image plane is the image of segment M_0M_1 of length D_1, and m_0m_2 is the image of M_0M_2 of length D_2. An additional constraint is provided by the knowledge of the angle $M_1M_0M_2$ at vertex M_0, which we call α. The triangle pose in space is completely determined by the location of the image points m_0, m_1, and m_2

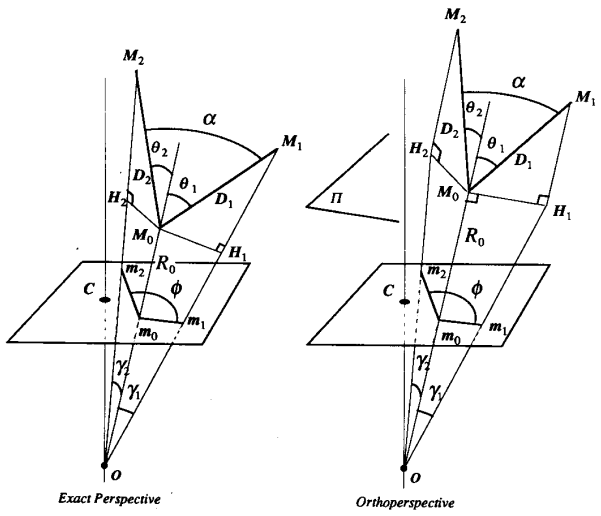

Fig. 2. Exact perspective and orthoperspective for a triplet of points.

(which is known from the image) by the range R_0 of the vertex M_0 along the line of sight Om_0 (which is to be computed from the image/world triangle pair) and by the angles θ_1 and θ_2 (which are also to be computed) that the sides M_0M_1 and M_0M_2 make with the line of sight Om_0 (with the optical axis direction in the case of weak perspective). Once these unknowns are determined, the 3-D positions of the three vertices of the model triangle can be computed.

The known relation between range and angle of a world segment, given its image, can be applied both to M_0M_1 and M_0M_2. For example, in exact perspective (see (1))

$$\frac{\sin(\theta_1-\gamma_1)}{R_0/D_1}=\sin\gamma_1,\quad \frac{\sin(\theta_2-\gamma_2)}{R_0/D_2}=\sin\gamma_2.$$

A solution for the angles is found first. Dividing the two equations eliminates R_0 and yields

$$\frac{\sin(\theta_1-\gamma_1)}{\sin(\theta_2-\gamma_2)}=K \tag{5}$$

where K is a known quantity depending only on the observables γ_1, γ_2, D_1, and D_2:

$$K=\frac{\sin\gamma_1/D_1}{\sin\gamma_2/D_2}.$$

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

Similarly, for all three approximations, the same procedure yields

$$\frac{\sin\theta_1}{\sin\theta_2}=K \tag{6}$$

where K is a known quantity depending only on the observables. For weak perspective, paraperspective, and orthoperspective, the expressions for K are, respectively

$$K=\frac{\sin\gamma_1/D_1}{\sin\gamma_2/D_2}\frac{\cos(\gamma_2+\beta_2)}{\cos(\gamma_1+\beta_1)},$$

$$K=\frac{\sin\gamma_1/D_1}{\sin\gamma_2/D_2}\frac{\cos\beta_1\cos(\gamma_2+\beta_2)}{\cos\beta_2\cos(\gamma_1+\beta_1)},$$

$$K=\frac{\tan\gamma_1/D_1}{\tan\gamma_2/D_2}.$$

A second relation between θ_1 and θ_2 is obtained by expressing α, which is the angle at vertex M_0 of the world triangle, in terms of θ_1 and θ_2.

$$\cos\alpha=\sin\theta_1\sin\theta_2\cos\phi+\cos\theta_1\cos\theta_2 \tag{7}$$

where ϕ is the angle between the image segments m_0m_1 and m_0m_2 in weak perspective and the angle between the planes of sight of M_0M_1 and M_0M_2 in paraperspective and orthoperspective. This equation was also derived in [8].

Therefore, solving for the angles θ_1 and θ_2 of the triangle sides M_0M_1 and M_0M_2 requires solving a system of two equations (an expression of the form of (7) and either (5) for exact perspective or (6) for the three approximations) with different expressions for the foreshortening ratio K according to the type of approximation and with a different interpretation of θ_1, θ_2, and ϕ in the case of weak perspective. Finding an equation in θ_1 only is more complex for the exact perspective than for the approximations since in the exact case, the trigonometric functions involve $\theta_1-\gamma_1$ in one of the equations and θ_1 in the other. The elimination process is given in the next section.

IV. Triangle Pose Solutions for Exact Perspective

Equations (5) and (7) for exact perspective can be rewritten

$$\sin t_1/\sin t_2=K.$$

$$p(c_1\sin t_1+s_1\cos t_1)(c_2\sin t_2+s_2\cos t_2)$$
$$+(c_1\cos t_1-s_1\sin t_1)(c_2\cos t_2-s_2\sin t_2)=q$$

with new unknowns

$$t_1=\theta_1-\gamma_1,\quad t_2=\theta_2-\gamma_2$$

and coefficients

$$p=\cos\phi,\ q=\cos\alpha,\ c_1=\cos\gamma_1,$$
$$s_1=\sin\gamma_1,\ c_2=\cos\gamma_2,\ s_2=\sin\gamma_2.$$

The system is solved by transforming $\cos^2 t_2$ into $(1-\sin^2 t_2)$ in the second equation by isolating $\cos t_2$ on the left-hand side, squaring both sides, eliminating the unknown t_2 in the second equation by replacing $\sin t_2$ by $\sin t_1/K$, obtaining an expression in $\sin t_1$ only by isolating $\cos t_1$ on the left-hand side, and, finally, squaring both sides again. This yields a fourth-degree equation in $\sin^2 t_1$:

$$u_4\sin^8 t_1+u_3\sin^6 t_1+u_2\sin^4 t_1+u_1\sin^2 t_1+u_0=0 \tag{8}$$

with coefficients

$$u_0=-r_2^2,\ u_1=q_1^2-2q_2r_2,\ u_2=-q_1^2-q_2^2-2(p_1q_1+p_2r_2),$$
$$u_3=p_1^2-2p_1q_1-2p_2q_2,\ u_4=-p_1^2-p_2^2.$$

These coefficients are themselves defined with respect to the following coefficients:

$$p_1=-2(a_1b_1+a_2b_2),\ q1=2(a_1b_1K^2-qb_2K),$$
$$p_2=a_1^2+a_2^2-b_1^2-b_2^2,$$

$$q_2=K^2(-a_1^2+b_1^2)+b_1^2K^2+2a_2qK,\ r_2=K^2(-b_1^2+q^2)$$

with

$$a_1=-s_1c_2+pc_1s_2,\ b_1=c_1c_2+ps_1s_2,$$
$$a_2=-s_1s_2-pc_1c_2,\ b_2=c_1s_2-ps_1c_2.$$

Equation (8) is a fourth-degree equation in $\sin^2 t_1$. We explored its solutions in only a slice of the parameter space; the parameter γ_1 was fixed by the constraint $\tan\gamma_1=0.1$. In this case, the equation yields only two real positive solutions. The unknown t_1 is $\theta_1-\gamma_1$, which

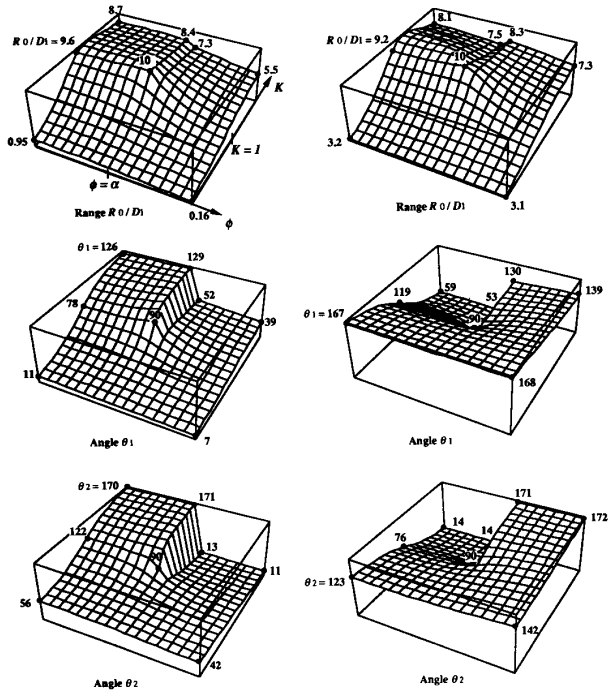

Fig. 3. Exact perspective solutions for triangle pose for model triangle with $D_1/D_2 = 0.5, \alpha = 45°$ and $\tan\gamma_1 = 0.1$ in image. Left column: First set of solutions $(R_0/D_1, \theta_1, \theta_2)$. Right column: Second set of solutions.

is the angle at vertex M_1 in the triangle OM_0M_1, and its value is between zero and π. Thus, the solutions for $\sin t_1$ must be positive, and two solutions for $\sin t_1$ are found. Each of these $\sin t_1$ solutions yields a single $\sin t_2$ solution from (5) and a single solution for the range R_0 of point M_0 from (1). One thus finds two complete solutions $(R_0, \sin t_1, \sin t_2)$. However, each set leads to four possible triangles corresponding to (R_0, t_1, t_2), $R_0, t_1, \pi - t_2)$, $(R_0, \pi - t_1, t_2)$, and $(R_0, \pi - t_1, \pi - t_2)$ for a total of eight triangle solutions for the two sets; extra solutions were introduced by the successive squarings, and the angle at M_0 is not necessarily α for these triangles. Choosing among the eight triangles requires determining the corresponding angles θ_1 and θ_2 $(\theta_1 = t_1 + \gamma_1, \theta_2 = t_2 + \gamma_2)$ and ensuring that they satisfy (7), which specifies that the angle at M_0 must be α. From these eight triangles, only one of them in each set was found to have an angle equal to the required angle α at its vertex M_0. Therefore, a total of two triangle poses was found to be compatible with (7) in the parameter space slice corresponding to $\tan\gamma_1 = 0.1$. Other constraints for γ_1 or a wider range for the other parameters could have allowed a total of four triangle poses; Fischler and Bolles [4] described specific geometric configurations in which four triangles of the same shape project to a single image triangle; more recently, Wolfe *et al.* [9] showed, with a simple graphic method, that while most of the parameter space yields only two triangle pose solutions, there are small islands that yield four triangle solutions.

Representative pose surfaces are presented graphically in Fig. 3. The left column of Fig. 3 shows the three diagrams of one set of solutions $(R_0, \theta_1, \theta_2)$, and the right column shows the second set. However, this separation into two sets is artificial. There is actually only one double-valued surface of solutions obtained by combining the surfaces from each set. The step discontinuity of the surface of one set matches the step discontinuity of the surface of the other set so that

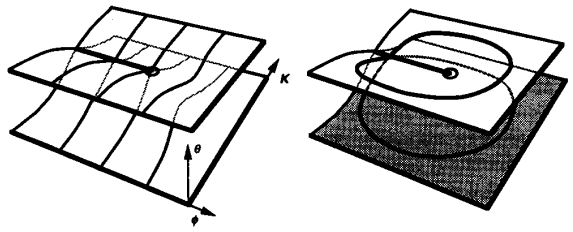

Fig. 4. Left: Shape of solution surface for angles θ_1 and θ_2 Right: Some closed cycles of image deformation correspond to the triangle moving from one pose solution to the other. The motion of the triangle is represented by a trajectory on the solution surface that loops from one ply of the surface to the other ply.

the resulting surface does not possess any discontinuity. For example, the general shape of the surface representing θ_1 is the double-valued surface shown in Fig. 4. The surface for R_0 also combines two layers that cross each other but are closer together.

A startling consequence of the shape of the solution surface is that the image of a triangle may be seen to deform in a closed cycle and return to its initial shape, whereas the corresponding world triangle has not actually returned to its original position but has gone to its other pose solution instead. With a second identical cycle of image deformation, the world triangle returns back to its initial position at the beginning of the first cycle (see the right side of Fig. 4). This occurs when the image deformations correspond to a closed cycle in the parameter plane (K, α) around the point $(K = 1, \phi = \alpha)$; the world triangle position is then represented by points along trajectories on the surfaces of the diagrams for $(R_0, \theta_1, \theta_2)$, such as the trajectory shown on the right side of Fig. 4. In this figure, the trajectory will follow the ramp connecting one level of the surface to the other and end up at a different level at the end of the first image cycle. With one more image cycle, the trajectory will follow the second ramp and complete its cycle.

V. Solutions with Approximate Perspectives

With the approximate perspectives, we must solve the system of equations

$$\frac{\sin\theta_1}{\sin\theta_2} = K \tag{9}$$

and

$$\sin\theta_1 \sin\theta_2 \cos\phi + \cos\theta_1 \cos\theta_2 = \cos\alpha. \tag{10}$$

In contrast with Section III, by isolating the term $\cos\theta_1 \cos\theta_2$ on the right-hand side of (10) and squaring the equation, one transforms both $\cos\theta_1$ and $\cos\theta_2$ into the corresponding sines in a single squaring. Then, the terms in $\sin\theta_2$ are eliminated using (9). This yields a second-degree equation in $\sin^2\theta_1$. We define $X_1 = \sin^2\theta_1$ and find

$$\sin^2\phi\, X_1^2 - (K^2 - 2K\cos\alpha\cos\phi + 1)X_1 + K^2\sin^2\alpha = 0. \tag{11}$$

Equation (11) always has two positive solutions, but only the smaller solution has magnitude less than 1 and can thus be equated to the sine of an angle. Thus, the solution of this quadratic always yields a single solution for $\sin\theta_1$

$$\sin\theta_1 = \left[\frac{-B - \sqrt{\Delta}}{2\sin^2\phi}\right]^{1/2} \tag{12}$$

with

$$B = -(K^2 - 2K\cos\alpha\cos\phi + 1), \quad \Delta = B^2 - 4K^2\sin^2\alpha\sin^2\phi.$$

This single sine solution results in two complementary solutions for θ_1. They correspond to mirror image directions of M_0M_1 with respect to the plane normal to Om_0 through M_0. The corresponding value for $\sin\theta_2$ obtained from (9) also yields two complementary solutions for θ_2. Note that we cannot combine the two solutions for θ_1 and the two solutions for θ_2 arbitrarily. Two of these combinations yield a positive number for $\cos\theta_1\cos\theta_2$, whereas the other two combinations result in a negative product. One of these pairs is unacceptable and was introduced when (7) was squared. This equation can be written

$$\cos\theta_1\cos\theta_2 = \cos\alpha - \sin\theta_1\sin\theta_2\cos\phi$$

and unacceptable pairs produce opposite signs for the left- and right-hand sides of the equation. Finally, the distance R_0 of the vertex M_0 from the center of projection is computed using (4). Note that only a single value of R_0 is recovered. Thus, the two solution triangles share a common vertex M_0, whereas in exact perspective, the two solution triangles have distinct M_0 vertices.

The solution surfaces are almost identical to those shown in Fig. 3. They can be found in [2]. The *angular* solutions do not involve γ_1, whereas the range R_0 does depend on γ_1.

VI. Comparison of Triangle Pose Obtained by Exact and Approximate Perspective

We compare the results provided by the exact perspective and the approximations for the image and triangle geometry used for Fig. 4. The results are plotted in Fig. 5. Since Fig. 4 was plotted using the expression of K for orthoperspective, the resulting error diagrams are to be interpreted either as the errors for orthoperspective independent of the position of the triangle in the image or as the errors for all approximations under the condition that the image vertex m_0 be at the image center. The following characteristics of the errors can be observed:

1) The error surfaces obtained by combining the solutions of the left and right column of Fig. 5 are smooth two-layer surfaces.
2) 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).
3) The largest angular errors occur along the line ($K = 1, \phi < \alpha$) and reach almost 10° in these valleys of the surfaces. At the edges of the diagrams, the angular errors are around 2 or 3°.

VII. Lookup Tables

The main advantage of using approximate perspective resides in the fact 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), whereas the range R_0 requires the use of γ_1 in (4) once the angles are found. For each triangle of features of an object, a 2-D 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, according to expressions that are specific to the type of approximation chosen (Section III), and the angles are read from the table. Then, the range R_0 is calculated using the parameter γ_1 computed from the image and using (4).

We implemented this technique on a 16K Connection Machine without floating-point processors. The pose estimate of a single polyhedron with 40 triangles with a smooth background is found in around 1 s [3]. This involves comparing all pairs of image and model triangles and clustering the resulting set of pose estimates.

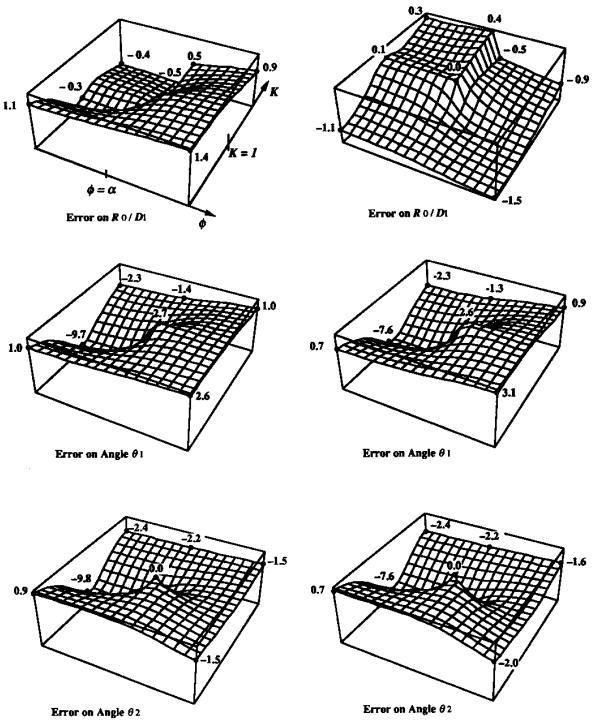

Fig. 5. Errors for triangle pose using orthoperspective for model triangle with $D_1/D_2 = 0.5, \alpha = 45°$, and $\tan\gamma_1 = 0.1$ in image. Left column: Errors for first set of solutions $(R_0/D_1, \theta_1, \theta_2)$. Right column: Errors for second set of solutions.

The error diagrams shown in Fig. 5 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 Fig. 5) can be flagged, and the pairs of model and image triangles corresponding to these table cells can be disregarded.

VIII. Conclusions

We have compared several approximate methods for the perspective-three-point problem: orthoperspective, paraperspective, and weak perspective. Orthoperspective is equivalent to a combination of a camera rotation that recenters the triangle and a weak perspective. It yields results similar to paraperspective and produces approximation errors that are lower than those with weak perspective for off-centered triangles. We have shown that with these approximate methods, 2-D lookup tables can be used to reduce the number of runtime floating-point operations needed to compute pose estimates.

References

[1] J. Aloimonos and M. Swain, "Paraperspective projection: Between orthography and perspective," Cent. for Automat. Res. CAR-TR-320, Univ. of Maryland, College Park, May 1987.
[2] D. DeMenthon and L. S. Davis, "New exact and approximate solutions of the three-point perspective problem," Tech. Rep. CAR-TR-471, Cent. Automat. Res., Univ. of Maryland, College Park, Oct. 1989.
[3] L. S. Davis *et al.*, "RAMBO—vision and planning on the connection machine," in *Proc. 1989 DARPA Image Understanding Workshop*, pp. 631–639.
[4] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and

automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, June 1981.
[5] R. M. Haralick, "The three point perspective pose estimation problem," Internal Note, Dept. of Elect. Eng., FT-10, Univ. of Washington.
[6] R. Horaud, B. Conio, O. Leboulleux, and B. Lacolle, "An analytical solution for the perspective four-point problem," *Comput. Vision Graphics Image Processing*, vol. 47, pp. 33–44, 1989.
[7] D. Huttenlocher and S. Ullman, "Recognizing solid objects by alignment," in *Proc. 1988 DARPA Image Understanding Workshop*, pp. 1114–1122.
[8] K. Kanatani, "Constraints on length and angle," *Comput. Vision Graphics Image Processing*, vol. 41, pp. 28–42, 1988.
[9] Y. Lamdan, J. T. Schwartz, and H. J. Wolfson, "On recognition of 3-D objects from 2-D images," in *Proc. 1988 IEEE Int. Conf. Robotics Automat.*, pp. 1407–1413.
[10] S. Linnainmaa, D. Harwood, and L. S. Davis, "Pose determination of a three-dimensional object using triangle pairs," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 10, pp. 634–647, 1988.
[11] Y. Ohta, K. Maenobu, and T. Sakai, "Obtaining surface orientation of plane from texels under perspective projection," in *Proc. IJCAI*, 1981, pp. 746–751.
[12] K. Pehkonen, D. Harwood, and L. S. Davis, "Parallel calculation of 3-D pose of a known object in a single view," *Patt. Recogn. Lett.*, vol. 12, pp. 353–361, June 1991.
[13] L. G. Roberts, "Machine perception of three dimensional solids," in *Opt. Elec.-Optical Inform. Processing* (J. T. Tippet *et al.*, Eds.). Cambridge, MA: MIT Press, 1965.
[14] D. W. Thompson and J. L. Mundy, "Three-dimensional model matching from an unconstrained viewpoint," in *Proc. IEEE Robotics Automat.*, 1987, pp. 208–220.
[15] ——, "Model-directed object recognition on the connection machine," in *Proc. DARPA Image Understanding Workshop*, 1987, pp. 98–106.
[16] W. J. Wolfe, D. Mathis, C. W. Sklair, and M. Magee, "The perspective view of three points," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, pp. 66–73, 1991.

Projection-Based Approach to Image Analysis: Pattern Recognition and Representation in the Position-Orientation Space

Hayit Greenspan, Moshe Porat, and Yehoshua Y. Zeevi

Abstract— **A method for image analysis, based on the formalism of functional analysis, is proposed. This approach, which is employed for pattern recognition in the combined position-orientation space, is motivated by a variety of neurophysiological findings that emphasize the importance of the orientation feature in visual image processing and analysis. The approach is applied to the so-called "Glass patterns," which are special cases of images characterized uniquely by their orientational feature of local correlations between pairs of corresponding dots. The approach can be of interest in the analysis of optical flow.**

Manuscript received July 18, 1990; revised June 20, 1992. This research was supported by the S. Faust Research Fund (VPR Fund 050–684), by the Ollendorff Center, and the Fund for Promotion of Research at the Technion. Recommended for acceptance by Associate Editor R. Woodham.

H. Greenspan was with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel. She is now with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

M. Porat is with the Department of Signal Processing Research, AT&T Bell Laboratories, Murray Hill, NJ, 07974. He is on sabbatical leave from the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel.

Y. Y. Zeevi is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel.

IEEE Log Number 9203760.

I. Introduction

The present study is concerned with the general problem of pattern classification and recognition with special emphasis on images that are fully characterized by their representation over the positional-orientational space. The approach is based on one of the most salient properties exhibited by cells in the primary visual cortex—their orientational selectivity. Physiological findings, which were first reported by Hubel and Wiesel [1], show that cells in the visual cortex are sensitive to specific spatial domains in the visual field of view (the so-called receptive fields) and have orientational selectivity. These physiological findings indicate that the orientational feature is extracted locally but that there also exists a global representation that can be considered as a mapping or projection onto the cortex according to orientation. Our approach deals specifically with the orientation as the only feature to be extracted locally and over which a global percept can be formulated. A similar processing scheme can be formulated over other feature domains as well as extended to multifeature domains.

Random-dot textures with local correlations between adjacent dots are uniquely characterized by their orientational feature. The Glass patterns, which were first introduced by L. Glass in 1969 [2], are such a well-studied example. The basic family of these patterns, as presented in the literature, is shown in Fig. 1(a); they are created by superimposition of two identical random-dot fields with one of them rotated, expanded, or translated relative to the other, thus inducing a structured global percept consisting of concentric circles, stars, and straight lines, respectively. In the present study, we use a variant technique of constructing the patterns from pseudorandom dot fields with homogeneous displacements [3].

Previous studies related to Glass patterns dealt with psychophysical experiments [3]–[6] and algorithmic analysis of the transition from the dot input x-y plane to dipoles, which are groupings of correlated dots in the input defined in the x-y-θ plane [3], [4], [7]–[9]. This transition is illustrated in Fig. 1(b). Regarding the algorithmic aspect, previous studies have related so far only to the local processing stage of determining the corresponding dots in the patterns. The purpose of the present work is to investigate global processing with regard to the spatial-orientational space. As such, it extends previous studies and considers the recognition of the global patterns. A theoretical approach dealing with the global effect as perceived by our visual system is presented.

Other studies have also found the position-orientation space important in image analysis. One example is the "ρ space" [10], where line-orientation discontinuities are extracted for enhancing the contrast of the most perceptually significant lines in an image and for segmenting the image into perceptually significant segments. Other related examples are referred to and discussed in [10]. Although the approach presented in this paper is motivated by human vision, it does not necessarily intend to provide an explanation of the human perception. Rather, it mimics its capabilities and motivates biological models similar to the proposed formalism.

II. Analysis of Position-Orientation Space

A. Objective

The proposed theoretical approach is based on image representation in a vector space, which is spanned by an orthogonal basis, and on an isometric operator that projects the input pattern into the $\mathbb{R}^N$ parameter space (feature space), where N is the dimensionality of