

Projective Alignment with Regions

Ronen Basri, *Member, IEEE*, and David W. Jacobs, *Member, IEEE*

Abstract—We have recently proposed an approach to recognition that uses regions to determine the pose of objects while allowing for partial occlusion of the regions. Regions introduce an attractive alternative to existing global and local approaches, since, unlike global features, they can handle occlusion and segmentation errors, and unlike local features they are not as sensitive to sensor errors, and they are easier to match. The region-based approach also uses image information directly, without the construction of intermediate representations, such as algebraic descriptions, which may be difficult to reliably compute. In this paper, we further analyze properties of the method for planar objects undergoing projective transformations. In particular, we prove that three visible regions are sufficient to determine the transformation uniquely and that for a large class of objects, two regions are insufficient for this purpose. However, we show that when several regions are available, the pose of the object can generally be recovered even when some or all regions are significantly occluded. Our analysis is based on investigating the flow patterns of points under projective transformations in the presence of fixed points.

Index Terms—Object recognition, pose estimation with regions.

1 INTRODUCTION

ESTIMATING the pose of objects is important both for manipulating the objects and as a step in their recognition. In this paper, we analyze the behavior of a region-based method of determining the pose of planar objects under projective transformations. Planar object recognition is potentially useful not only because some objects are flat, but also because some polyhedral objects can be positioned using markings on their flat surfaces, and because mobile robots may localize their position relative to flat indoor surfaces. By allowing projective transformations to represent pose, we handle a superset of true perspective distortions and allow an object model to be acquired with an uncalibrated camera of unknown pose and to be matched to an image taken with an uncalibrated camera.

A basic question in pose estimation is how to represent objects in order to estimate their pose accurately and efficiently. Recently, we proposed a method for pose estimation that uses regions ([1], [11]). This method expands the repertoire of possible representational primitives, using direct representation of region information. It does not require an exact localization of features. It handles objects with smooth curved boundaries, but does not require an algebraic description of such objects. Moreover, the method can recognize planar objects when the image regions are occluded, without needing to derive information about the extent of this occlusion. Also, region information can be seamlessly combined with local features, when they are available. Finally, this method computes the pose of objects efficiently using linear programming.

In [1], we introduced the method and explored its properties for planar objects undergoing similarity and affine

transformations. We also showed that the method could be applied to the projective case and stated some preliminary results about this case. In this paper, we more thoroughly consider the case of aligning an image to a model of a planar object undergoing a projective transformation. We focus on the most basic problem of determining when there is enough region information to determine object pose. We prove that three regions generally determine a solution uniquely. In addition, we show that for a large class of objects two regions are insufficient for this purpose. This complements our previous results showing that two regions determine a unique solution under affine transformations. At the same time, we show that when several regions are available the pose of the object can generally be recovered even when all regions are significantly occluded.

We should compare these results to corresponding ones for other pose determination methods. First, a projective transformation relating a model with an image can be determined using four or more corresponding simple local features, such as points or lines (such methods are discussed, e.g., in [9]). Our algorithm addresses the problem of pose determination when such local features are difficult to find or match. For example, in smooth convex shapes, there may not be local features that can be reliably computed and matched for arbitrary perspective views.

Second, one can determine the pose of two regions using their second moments or algebraic descriptions of conics (e.g., [8], [12]). We show that our method generally requires more than two regions. However, methods based on moments or algebraic descriptions require regions that have no occlusion. Our method correctly uses information derived from partially occluded image regions without any prior knowledge of which regions are occluded or the locus of occlusions. We show theoretically that our method produces correct results when several regions are matched and at least two are unoccluded, while others are arbitrarily occluded. We demonstrate experimentally that our method can correctly determine pose even when all regions are partially occluded. By handling arbitrarily occluded, arbitrarily shaped regions, our method handles some problems that cannot be solved by other existing methods.

• R. Basri is with the Department of Computer Science, The Weizmann Institute of Science, Rehovot, 76100, Israel.
E-mail: ronen.basri@weizmann.ac.il

• D. Jacobs NEC Research Institute, 4 Independence Way, Princeton, NJ 08540. E-mail: dwj@research.nj.nec.com.

Manuscript received 18 Aug. 1999; revised 20 Nov. 2000; accepted 29 Nov. 2000.

Recommended for acceptance by S. Dickinson.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 110460.

The paper is divided as follows: In Section 2, we review the method of recognition with regions. In Section 3, we derive necessary and sufficient conditions for unoccluded regions to produce a unique projective transformations and show that, in general, three regions produce a unique solution. Results a bit weaker than those shown in this section had been stated without proof in [1]. In Section 3.2, we show that two unoccluded regions generally do not produce a unique solution and, in Section 3.3, we explore the contribution of partially occluded regions in providing a unique solution.

2 RECOGNITION WITH REGIONS

In this section, we briefly review the scheme for recognition using region correspondences. A more detailed description can be found in [1]. We restrict our discussion to planar objects undergoing projective transformations.

In our scheme, an object is modeled as a set of convex regions (denoted by $V_1, \dots, V_k \subset \mathcal{R}^2$) and an image of the object is given by a corresponding set of regions (denoted by $R_1, \dots, R_k \subset \mathcal{R}^2$). Our task is to determine the transformation $\bar{T} \in \mathcal{T}$ that maps every model region V_i to its corresponding image region R_i ($1 \leq i \leq k$), where \mathcal{T} denotes the set of projective transformations. Basri and Jacobs [1], [11] discuss issues involved in applying the method to nonconvex regions, using their convex hulls, but we will not discuss that here.

Determining a transformation that perfectly maps a set of model regions to their corresponding image regions is generally a nonconvex optimization problem since we can have multiple, locally optimal solutions. Moreover, such a transformation would not allow for partial occlusion of the image regions. For these reasons, we divide the constraints that finding a perfect match would place on the transformation into two sets, one of which is unaffected by occlusion.

Forward constraints: Every model point $\bar{p} \in V_i$ should project inside the region R_i (that is, $\bar{T}V_i \subseteq R_i$).

Backward constraints: Every image point $\bar{q} \in R_i$ is the projection of some model point $\bar{p} \in V_i$ (that is, $\bar{T}V_i \supseteq R_i$, or equivalently, $V_i \supseteq \bar{T}^{-1}R_i$).

Individually, each set of constraints produces a convex set of feasible transformations. Thus, each set of constraints can be solved efficiently by computing a linear discriminant function.

We proceed by defining the backward constraints under a projective transformation (see Fig. 1). The forward constraints can be defined in the same way since planar objects models and images are interchangeable. Solving for the transformation using the backward constraints alone is particularly useful in the case of occlusion. Image regions that are partly occluded lie inside the corresponding model regions (after the model and the image are brought into alignment), but the inclusion may be strict due to the occlusion.

We solve the backward constraints by formulating constraints on \bar{T}^{-1} , the inverse of the transformation we seek. For notational convenience, we let $T = \bar{T}^{-1}$. Let $\bar{q} = (u, v, 1)^T \in V$ be a model point written in homogeneous coordinates. Below, we denote a line of the form $Ax + By + C = 0$ as a 3-tuple, $(A, B, C)^T$ (so the line equation is expressed as a dot product $(A, B, C)(x, y, 1)^T = 0$). Then, since V is convex, every line $\bar{l} = (A, B, C)^T$ which does not intersect the interior of V satisfies

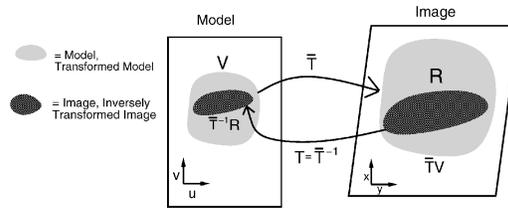


Fig. 1. The backward constraints require that a feasible transformation satisfies $\bar{T}V \supseteq R_i$. Equivalently, $V \supseteq \bar{T}^{-1}(R_i) \equiv T(R_i)$.

$$\bar{l}^T \bar{q} \geq 0. \quad (1)$$

Let $\bar{p} = (x, y, 1)^T \in R$ be an image point. The backward constraints imply that the appropriate transformation T maps \bar{p} inside V . This implies that there exists some point $\bar{q} \in V$ such that

$$\bar{q} = \alpha T \bar{p}, \quad (2)$$

where T is 3×3 nonsingular and α is a scalar factor. When T is restricted to be affine, its third row is given by $(0, 0, 1)$ and $\alpha = 1$ (and, so, $\bar{q} = T \bar{p}$). In the general projective case, α can be an arbitrary nonzero value. Furthermore, in actual images since the object is constrained to appear in front of the camera, α must be positive. Combining (1) and (2), we obtain that

$$\alpha \bar{l}^T T \bar{p} \geq 0, \quad (3)$$

with $\alpha > 0$. Since α is positive, it can be eliminated from the equation, yielding

$$\bar{l}^T T \bar{p} \geq 0. \quad (4)$$

This equation is linear in the unknown transformation parameters, which are the components of T .

The backward problem introduces such a constraint for every pair of a point in the image regions and a tangent line to the model regions. In the forward problem, the model and image change roles. The number of constraints for a curved object is therefore infinite. For such objects, we obtain a finite system of inequalities by sampling the set of constraints. For polygonal regions, the number of independent constraints is finite. These constraints are defined by the vertices of the image regions and the sides of the model regions and the rest of the constraints are redundant.

The one-way problem, therefore, can be expressed as follows: Given a finite set of constraints

$$\bar{l}_i^T T \bar{p}_i \geq 0, \quad i = 1, \dots, n, \quad (5)$$

we seek a matrix of parameters T that is consistent with the constraints.

Solving the one-way problem (5) involves finding a linear discriminant function. One method of finding a linear discriminant is by using linear programming. To generate a linear program a linear objective function should be specified. A common way of defining such a linear program is by introducing an additional unknown, λ , in the following way:

$$\max \lambda \quad \text{s.t.} \quad \bar{l}_i^T T \bar{p}_i \geq \lambda, \quad i = 1, \dots, n. \quad (6)$$

A solution to (5) exists if and only if a solution to (6) with $\lambda \geq 0$ exists. (Note that other objective functions, e.g., the

perceptron function, can be used for recovering T , see, e.g., [7] for a discussion of solutions to the linear discriminant functions problem.)

To give a rough sense of the actual runtime of our algorithm, we note that convex curves can be accurately approximated using a relatively small number of line segments (while it is possible to prove bounds on this, we just note that, in practice, a dozen or so line segments will usually provide an excellent approximation to a convex region in an image). The noise introduced by this approximation can be reasonably made smaller than the noise generally present in region segmentation. If we match k polygonal regions, each with m sides, we will have km^2 constraints. Because the linear program has a constant number of variables (nine in the projective case), it is possible to solve it in time that is linear in the number of constraints (Seidel [15] gives an algorithm that runs in expected time that is $O(d!n)$, with n constraints and d variables). More practically, Simplex is found to solve such problems in a time roughly proportional to dn^2 (Strang [16]). Matlab's linear programming code solves a problem with five, 16-sided polygons, in about one-half a second on a laptop with a 400 MHz, Pentium II processor.

2.1 Uniqueness

Whenever an error-free image contains an instance of the model, perhaps with some occlusion, the backward constraints (5) are satisfied by the correct alignment transformation. However, in certain cases, other transformations may also satisfy the backward constraints. An algorithm for finding a linear discriminant function is likely in such cases to return one of these incorrect transformations, typically by allowing the image regions to contract inside the corresponding model regions. Therefore, it is important to determine when the backward constraints specify a unique solution.

We define the uniqueness problem as follows: Let M be a model composed of k convex regions V_1, \dots, V_k . Let $T \in \mathcal{T}$ be a projective transformation and let R_1, \dots, R_k be the corresponding image regions, so that $TR_i = V_i$ ($1 \leq i \leq k$). Consider the backward constraints between the model and image. Does there exist another transformation $T' \neq T \in \mathcal{T}$ that satisfies these backward constraints? We focus first on the case where all k image regions are unoccluded. However, our method derives only valid constraints from arbitrarily occluded regions. So, if k unoccluded regions produce a unique result, so do any set of regions with k of them unoccluded.

Basri and Jacobs [1] proved a basic lemma that the uniqueness of a one-way matching problem depends on the model alone. If the model is nondegenerate, a unique solution will be obtained when the model is matched to any of its images, while if the model is degenerate, multiple solutions will exist whenever the model is matched to any image of the object. The lemma states the following claim: The solution to a one-way matching problem under a group of transformations (similarity, affine, or projective) is unique if and only if there exists no transformation of that group (other than the identity) which projects the model regions entirely inside themselves. The transformation that produced an image does not affect whether it leads to a unique pose. Thus, we can modify the problem as follows: Let M be a model composed of k convex regions R_1, \dots, R_k .

Does there exist a transformation $T \neq I$ (where I denotes the identity transformation) such that $TR_i \subseteq R_i$ for all $1 \leq i \leq k$?

Another property that is used below (and was used already in [1]) is that transformations that project the model regions entirely inside themselves introduce fixed points inside the model regions. That is, if $T(R_i) \subseteq R_i$, there exists $\vec{p} \in R_i$ such that $T(\vec{p}) = \vec{p}$. The presence of fixed points, and their particular nature, will play a key role in our proofs.

3 PROJECTIVE TRANSFORMATIONS

In [1], we showed that two distinct regions generally determine the pose of a planar object undergoing similarity or affine transformations. In this section, we extend the uniqueness results to the projective case. This depends on a further analysis of the fixed points, which is given in Section 3.1. Previous researchers ([4], [17]) have analyzed the fixed points of optical flow fields as dynamical systems, obtaining characterizations of the possible local properties of fixed points. We extend that work by presenting a taxonomy of the possible combinations of fixed points that can occur in a planar projective flow field. In [4], it was pointed out that these fixed points are the eigenvectors of a transformation matrix and we use that fact as the basis of our analysis. This analysis is relevant to our uniqueness questions because we also show that any projective transformation mapping a region inside itself must have a sink point inside that region. Our analysis shows that either there is just one sink point in planar projective flow or there is a line on which all points are sink points. This implies that, if three regions have no single line intersecting them all, they must lead to a unique transformation. It also allows us to derive necessary and sufficient conditions for collinear regions to produce nonunique solutions.

We use these conditions in Section 3.2, however, to show that every pair of distinct triangles or ellipses is degenerate. Finally, in Section 3.3, we explore, given three regions, to what extent one of the regions can be occluded and still maintain uniqueness.

3.1 Conditions for Uniqueness

We begin with a general discussion of the fixed points that may occur in planar projective flow. First, we note that a planar projective transformation can be written as a 3×3 matrix. Points in the projective plane can be written in homogeneous coordinates as 3D vectors, (x, y, w) . The image coordinates of this point are then $(x/w, y/w)$. So, if $p = \lambda q$ for some nonzero scalar λ , then p and q represent the same point in the plane. p is a fixed point if $Tp = \lambda p$. Therefore, if p is a fixed point of T , p is an eigenvector of T , with λ being the associated eigenvalue.

Next, we consider a basic property of the fixed points that belong to regions obeying the one-way constraints under projective transformations. As stated above, to determine whether a model gives rise to unique solutions, without loss of generality, we may assume that the identity transformation maps the model to the image and ask whether $T \neq I$ exists such that $p \in V_i \Rightarrow Tp \in V_i$. Note that if T obeys this constraint, it must also be the case that $T(Tp) \in V_i$, and, in general, that $T^n p \in V_i$, for all n . Therefore, we are especially interested in characterizing the limit points of $T^n p$.

TABLE 1
Topologically Distinct Cases Involving Fixed Points

# eigenvectors associated with largest eigenvalue	# linearly independent eigenvectors	# real eigenvalues	Max. # regions obeying constraints
0	1	1	1
1	1,2 or 3	1,2 or 3	1
2	2	1	1
2	3	2	2
3	3	1	All, $T = I$

There are several possible cases that can occur in the limit as n goes to infinity for $T^n p$. First, $T^n p$ can fail to converge (e.g., by forming a cycle) when T has complex eigenvalues. T can have either zero or two complex eigenvalues. When it has two complex eigenvalues, it can have only one fixed point (only one eigenvector associated with a real eigenvalue). When $T^n p$ converges to a point, this point will be an eigenvector of T . In general, this will be an eigenvector associated with T 's largest eigenvalue. This is well-known and is the basis of the power method of computing eigenvalues. $T^n p$ will only converge to a different eigenvector in special cases. If p is an eigenvector to start with, then obviously $T^n p$ always remains at p . Or, if p is a linear combination of two eigenvectors associated with nonmaximal eigenvalues of T , then $T^n p$ will remain a linear combination of these two eigenvectors and will generically converge to the one that is associated with the second largest eigenvalue.

Geometrically, this means that when $T^n p$ converges to a point, there may be a line in the image such that for points on this line $T^n p$ does not converge to a fixed point associated with the largest eigenvalue. However, all other points in the plane do converge to such fixed points. We will now analyze topologically distinct cases involving these fixed points. Table 1 summarizes the results of this analysis.

First, consider the case in which T has three linearly independent eigenvectors associated with its largest eigenvalue. This case occurs only when T has a single real eigenvalue of multiplicity three. In this case, all linear combinations of these eigenvectors are also eigenvectors, implying that all points in the plane are fixed points and $T = I$. Therefore, T cannot be distinct from I and have three linearly independent eigenvectors associated with one eigenvalue.

Second, we consider the case in which there is no eigenvector associated with a real eigenvalue of greatest magnitude. This can only occur when a complex eigenvalue is the eigenvalue of greatest magnitude. In this case, there are two complex eigenvalues and one real one. So, there is only one fixed point of the transformation and there can only be one region that obeys the one-way constraints.

Third, consider the case in which there is one eigenvector associated with a real eigenvalue of greatest magnitude. In this case, for any point p that is not a linear combination of the other two eigenvectors, $T^n p$ converges to this *sink* eigenvector. Any region of nonzero area that obeys the backward constraints must contain this sink point, since for any point in the region, $T^n p$ is also in this region. Therefore, only one region can obey the backward constraints.

Finally, we consider the case in which the largest eigenvalue, λ , is real and has two linearly independent

eigenvectors, p_1 and p_2 . In this case, from elementary linear algebra, we know that since $Tp_1 = \lambda p_1$ and $Tp_2 = \lambda p_2$, $T(ap_1 + bp_2) = \lambda(ap_1 + bp_2)$, and $ap_1 + bp_2$ is a fixed point of T , for any choice of a and b . That is, all points on the line formed by p_1 and p_2 are fixed points and we say this line is *pointwise fixed*. For any p , if p is not the third eigenvector of T , $T^n p$ converges to some point on this line. Therefore, every region that is not a point and obeys the backward constraints must contain a point on this line and we have proven:

Theorem 1. Let $V_1, V_2, V_3 \subseteq \mathcal{R}^2$ be three regions, each containing more than a single point such that there exists no straight line passing through all three regions. Then, the solution to the one-way matching problem with these regions as a model under a projective transformation is unique.

We have also shown that if two or more regions do obey the one-way constraints, then there must be a transformation T mapping the regions inside themselves such that T has exactly two linearly independent eigenvectors associated with its largest real eigenvalue. We now describe the form T must have more precisely.

First, we show that if T obeys the one-way constraints on two distinct regions, T must have a third eigenvector. Suppose this were not the case. Then, T must have only one eigenvalue with two associated eigenvectors. Without loss of generality (WLOG), we may assume that the pointwise fixed line is the x axis in which case we have:

$$T = \begin{pmatrix} 1 & b & 0 \\ 0 & 1 & 0 \\ 0 & d & 1 \end{pmatrix}$$

(since we can scale the matrix without changing the projective transformation it produces, we may assume WLOG that $\lambda = 1$). Note that such a T has only two eigenvectors, unless $T = I$. Then, we have:

$$T^n = \begin{pmatrix} 1 & nb & 0 \\ 0 & 1 & 0 \\ 0 & nd & 1 \end{pmatrix}.$$

For $p = (x, y, 1)$, $T^n p = (x + nby, y, ndy + 1)$. In the limit as n goes to infinity, $T^n p$ therefore converges to the point $(\frac{x}{d}, 0)$. Therefore, in this case, any region with nonzero area that obeys the backward constraints must contain the point $(\frac{x}{d}, 0)$. This cannot be the case for two distinct regions. This proves that T must have a third eigenvector.

In summary, T can obey the backward constraints with two or more distinct regions of nonzero area if and only if T has three eigenvectors, two of which are associated with the largest eigenvalue. These two eigenvectors imply that T has a pointwise fixed line, l . If we call the third eigenvector q , it follows that all lines through q are mapped by T onto themselves since T maps lines to lines and all lines through q intersect l also and, so, have two fixed points. These are called *fixed lines*. If WLOG, we choose our coordinate system so that l is the x axis and $q = (0, 1, 1)$, then:

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - a & 0 \\ 0 & a & 1 \end{pmatrix}$$

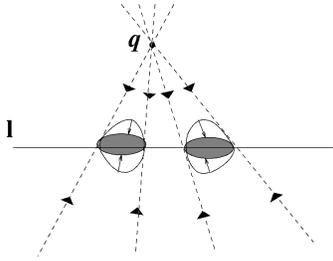


Fig. 2. Two model regions lead to nonunique projective transformations when a line l , exists such that the tangents at all intersection points meet at a single point q . In this case, the regions can contract toward l in the directions emanating from q .

for $0 < a < 1$. T maps all points that are not fixed away from q and towards l , as shown in Fig. 2.

We can derive one further useful condition if T obeys the one-way constraints with respect to a region R_i . Consider any point p that is on the boundary of R_i , so that the line through p and q is tangent to R_i . Then, p must also lie on l . Otherwise, Tp also lies on the line tangent to R_i , but $Tp \neq p$, and so Tp is not in R_i , violating the one-way constraints. (It is also possible that a tangent line to R_i contains more than one point, if part of the boundary of R_i is a straight line. In this case, similar reasoning shows that at least one point on R_i on this tangent line must also lie on l). Alternately, we can say that, if we draw lines through q tangent to all the regions, then the tangent points on the boundaries of these regions must all be collinear. This is also depicted in Fig. 2. It is easy to see that this is not only a necessary but also a sufficient condition for T to obey the one-way constraints.

We have therefore shown (see Fig. 2):

Theorem 2. Let $V_1, V_2 \dots V_k \subseteq \mathcal{R}^2$ be distinct regions with nonzero areas. Then, the solution to the one-way matching problem with these regions as a model under a projective transformation is nonunique if and only if there exists a line l through the regions and a point q outside l such that contracting V_i in directions emanating from q toward l (denoted by $T_{l,q}$) implies

$$T_{l,q}(V_i) \subset V_i, \quad i = 1, 2 \dots k.$$

This theorem is the natural generalization of the two region case under affine transformations. In that case, a degeneracy occurs when the tangent lines are parallel (i.e., intersect at a point at infinity). In the projective case, a degeneracy occurs when the tangent lines intersect at any point in the plane. Note also that the contracting transformation, which keeps a line pointwise fixed and has an additional fixed point, is known as a *projective homology*. Homologies were used recently to detect repeated structures and coincident sets of parallel lines in images [13], [14].

Using this result, we can show that generically, even when three smooth regions are intersected by a single line segment, they will give rise to a unique solution. This is because there are two degrees of freedom in placing a line that intersects all three regions. However, the condition that all six points where this line intersects the regions will have tangents that intersect at a single point provides four constraints on the location of the line. Simple equation



Fig. 3. Two views of a drawing by Kandinsky.

counting shows that generically, we will not be able to satisfy these constraints.

Figs. 3 and 4 demonstrate this result using real images. We have manually extracted regions from the images in Fig. 3 and used these regions to recover the projective transformation relating the two images. Fig. 4 shows the result of aligning the two images using three regions. As can be seen, three regions led to a fairly good alignment, indicating that the projective transformation was correctly recovered.

3.2 Nonuniqueness of Two Regions

Similar reasoning shows that two regions will not give rise to a unique solution when one can use three degrees of freedom to satisfy three constraints. The resulting equations are nonlinear so we are not guaranteed that a solution will exist, but it is plausible that often two regions provide insufficient constraint to determine the solution uniquely. To test this we examine two types of shapes, triangles and ellipses. Below, we show that both objects composed of two triangles and objects composed of two ellipses give rise to nonunique solutions, supporting the intuition that for a large class of objects two regions do not introduce sufficiently many constraints to uniquely determine a solution.

We now turn to showing that the solution for the projective one-way matching problem with a model consisting of two distinct triangles is always nonunique. To show that such a model is degenerate, we have to find a line l through the two triangles and a point q outside the two triangles such that contracting the two triangles toward l in directions emanating from q will contract the triangles inside themselves. Given two distinct triangles notice that one can construct a line that enters each triangle at a vertex and exits the triangle at an edge or vice versa (see, e.g., Fig. 5 in which the line l enters the triangles at \vec{p}_1 and \vec{p}_4 and exits the triangles at the edges connecting \vec{p}_2, \vec{p}_3 and \vec{p}_5, \vec{p}_6). Denote this line by l ; l will be the fixed line. Now, continue the two edges from which l exits until they intersect (which if they are parallel will be at infinity). Denote the point of

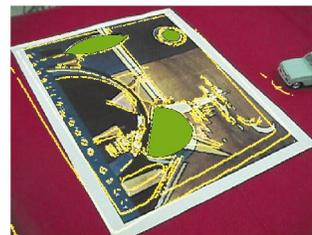


Fig. 4. Aligning the two views with the one-way constraints using three regions. The figure shows an overlay of the two images. The green regions represent the regions used for alignment. These regions are the sun, the tree top, and the tilted dome.

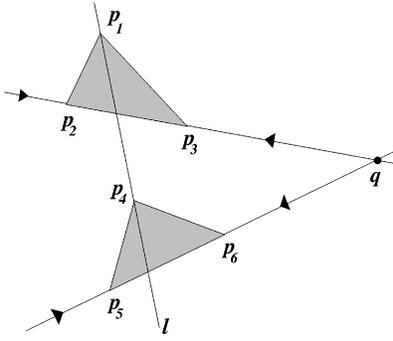


Fig. 5. Two model regions lead to nonunique projective transformations when a line l , exists such that the tangents at all intersection points meet at a single point q . In this case, the regions can contract toward l in the directions emanating from q .

intersection by \bar{q} . It can be readily verified that contracting the triangles toward l in directions emanating from \bar{q} will keep the triangles inside their original regions. Consequently, under projective transformation, any model consisting of two distinct triangles is degenerate.

Next, we show that the solution for the projective one-way matching problem with a model consisting of two distinct ellipses is always nonunique. An ellipse in the projective plane is defined by a quadratic form $\bar{x}^T A \bar{x} = 0$, where A is a 3×3 symmetric matrix with negative determinant. Given an ellipse (or a general nondegenerate conic), two points \bar{x} and \bar{y} are called *conjugate* with respect to the conic if they satisfy the bilinear equation $\bar{x}^T A \bar{y} = 0$. For a given point \bar{x} , the set of conjugate points to \bar{x} forms a straight line. This line, l , called the *polar line* of \bar{x} , is given simply by $\bar{x}^T A \bar{y} = 0$ since the quadratic form is linear in \bar{y} . Similarly, \bar{x} is called the *pole* of l .

Here are a few examples of conjugate points and polarities. A point is *self-conjugate* if and only if it lies on the conic. The polar line of a point on the conic is the tangent to the conic through that point. The polar line of a point \bar{u} outside a conic is given by the following construction: Connect \bar{u} to the conic by two tangent lines. Each tangent meets the conic at a single point. The polar line of \bar{u} is the line connecting the two intersections of the tangents with the conic.

The following lemma establishes that when two ellipses are contracted inside themselves, the fixed line is, in fact, the polar line of the fixed point, \bar{u} , from which contraction is emerging. This polarity relationship is used in Lemma 1 to relate the uniqueness of two ellipses with the generalized eigenvalue problem.

Lemma 1. *The solution to the projective one-way problem for two distinct ellipses $\bar{x}^T A \bar{x} \leq 0$ and $\bar{x}^T B \bar{x} \leq 0$ is nonunique if and only if the matrix $A^{-1}B$ has a positive eigenvalue.*

Proof. According to Theorem 2, two regions V_1 and V_2 give rise to a nonunique solution if and only if there exists a line l through the regions and a point \bar{u} outside the regions such that contracting the regions in directions emerging from \bar{u} toward l will map V_1 and V_2 to inside themselves. Notice that, for smooth bounded regions, the directions of contraction at the points of intersection of l with the boundaries of V_1 and V_2 must be tangential to the boundaries or else the regions will not contract within themselves. Consequently, suppose that now V_1

and V_2 are two ellipses, $\bar{x}^T A \bar{x} \leq 0$ and $\bar{x}^T B \bar{x} \leq 0$, respectively, then l must be the polar line of \bar{u} with respect to both A and B .

Since l is the polar line of \bar{u} with respect to both A and B , it is given by both $\bar{x}^T A \bar{u} = 0$ and $\bar{x}^T B \bar{u} = 0$. These two equations represent the same line if and only if they are scaled versions of one another, that is

$$\lambda A \bar{u} = B \bar{u} \quad (7)$$

for some $\lambda \neq 0$. λ and \bar{u} , respectively, are called the generalized eigenvalue and eigenvector of B with respect to A . Alternatively, if we multiply the last equation by A^{-1} we obtain

$$A^{-1} B \bar{u} = \lambda \bar{u}. \quad (8)$$

We have proven so far that contraction is possible if and only if there exists a point \bar{u} outside the two ellipses which is an eigenvector of $A^{-1}B$.

Next, we show that \bar{u} lies outside the two ellipses if and only if $\lambda > 0$. Multiplying the two sides of (7) by \bar{u}^T from the left we obtain that

$$\lambda \bar{u}^T A \bar{u} = \bar{u}^T B \bar{u}. \quad (9)$$

Suppose that \bar{u} lies outside the two ellipses, then $\bar{u}^T A \bar{u} > 0$ and $\bar{u}^T B \bar{u} > 0$, and, consequently, $\lambda > 0$. Conversely, suppose $\lambda > 0$, then $\bar{u}^T A \bar{u}$ and $\bar{u}^T B \bar{u}$ share the same sign. If $\bar{u}^T A \bar{u} < 0$ and $\bar{u}^T B \bar{u} < 0$, then \bar{u} must be contained in both ellipses, but this is impossible because the two ellipses are distinct. Therefore, both $\bar{u}^T A \bar{u}$ and $\bar{u}^T B \bar{u}$ must be positive and \bar{u} lies outside the two regions. \square

Using Lemma 1, we can now show:

Theorem 3. *The solution to the one-way projective problem with a model consisting of two ellipses is always nonunique.*

Proof. According to Lemma 1, the solution to the one-way projective problem with a model consisting of two distinct ellipses $\bar{x}^T A \bar{x} \leq 0$ and $\bar{x}^T B \bar{x} \leq 0$ is nonunique if and only if the matrix $A^{-1}B$ has a positive eigenvalue. Note that both $\det(A)$ and $\det(B)$ are negative. Therefore, $\det(A^{-1}B)$ must be positive. This implies that $A^{-1}B$ has a positive eigenvalue. To see this, note that the eigenvalues of C are the roots of the characteristic polynomial of C which is a third order polynomial. Every third order polynomial has either three real roots or two complex conjugate roots and one real root. In case all three roots are real, they cannot all be negative since their product must be equal to $\det(C)$ which is positive. In case only one of the three roots is real, this root still must be positive because the two other roots are conjugate and so their product is positive. Consequently, $A^{-1}B$ must have at least one positive eigenvalue and, so, the solution to the one-way projective problem for any two ellipses is nonunique. \square

An example of two ellipses and the fixed point and line that they produce is shown in Fig. 6. One eigenvector produces the fixed point; the other two eigenvectors, which correspond to negative eigenvalues, lie on the fixed line and are located inside the two ellipses. Note that it is known that, given two pairs of ellipses related by a projective transformation, the transformation that aligns the two pairs of ellipses exactly is determined uniquely. Theorem 3 shows

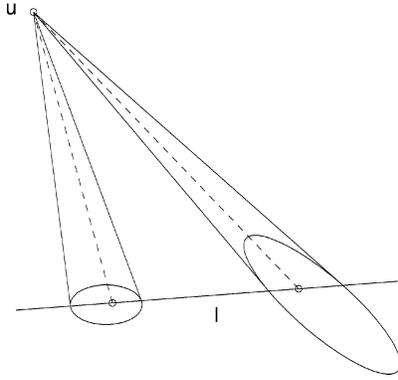


Fig. 6. Contraction directions with two ellipses. The generalized eigenvalues of the ellipses are 13.14, -31.10, and -0.04. The eigenvector corresponding to the only positive eigenvalue determines the fixed point (denoted by u) and its polar line with respect to either of the ellipses is the pointwise fixed line (denoted by l). The two eigenvectors that correspond to the negative eigenvalues determine two points that lie on the fixed line and are located inside the two ellipses. (The points are denoted by circles and their respective polar lines are denoted by the dashed lines.)

that this transformation is not determined uniquely by the backward (or forward) constraints alone. However, by considering the backward constraints, we can align more than two convex shapes even in the presence of unknown occlusion.

Fig. 7 illustrates these results on a real image. Here, two regions are used to determine projective pose, producing an incorrect pose with significant contraction.

3.3 Uniqueness in the Presence of Occlusion

The implication of Section 3.2 is that, for a large class of objects, two regions are insufficient to determine a projective transformation uniquely. In this section, we examine analytically the case of a model composed of three regions when one of the regions appears partly occluded. We show below that it is possible to occlude significant portions of one of the regions and still obtain a unique solution. We then describe experiments that indicate that a unique solution is often obtained with zero or one unoccluded regions.

In the analysis below, we assume that the regions are ellipses and that no straight line can traverse all three regions. Furthermore, we can assume WLOG that the occluded region is a circle. This will require preprocessing the model by applying to it an affine transformation, but such a transformation does not effect whether a unique solution exists and it does not change relative areas. According to Theorem 2, a nonunique solution is obtained only if it is possible to contract the two ellipses inside themselves toward a pointwise-fixed line in directions emanating from a point outside the two ellipses. As shown in Section 3.2, the number of different contractions can be either one of three according to the number of positive eigenvalues.

Consider the case in which the two ellipses introduce a single contracting transformation. (It can be readily shown that it is generally impossible to have more than one contracting transformation.) Suppose that the fixed point is at infinity (in which case, the contracting transformation is in fact affine). In this case, we obtain the situation illustrated in Fig. 8. The contracting transformation due to the two unoccluded ellipses is not feasible because the boundaries

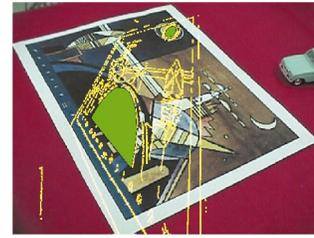


Fig. 7. Aligning the two views with the one-way constraints using two regions. The figure shows an overlay of the two images. The green regions (the sun and the large dome) represent the regions used for alignment.

of the third region that face the fixed line will prevent the third region from contracting. Only when this side is occluded entirely will the contracting transformation become feasible and the solution cease to be unique.

To get a sense of the likelihood that occlusions may destroy uniqueness, we now restrict ourselves to the case that the occluded region is a circle of unit radius. (The following results do not extend to the case that the occluded region is an ellipse since in that case the likelihood that occlusion destroys uniqueness will vary with the orientation and the aspect ratio of the ellipse.) Let occlusion be determined by a half-plane passing through two randomly selected points on the boundary of the circle. (It can be readily verified that a uniform selection of pairs of points on the boundary of the circle yields a uniform selection of half-planes.) To construct a destructive occlusion, we need to select the two points in the upper half of the circle and, in addition, select the side of the plane that includes the center of the circle. The chance of this happening (assuming uniform distributions) is 0.125. Using this model, we can also estimate the expected area of occlusion that can be applied to the third region without destroying uniqueness. Next, we show that under this model we can on average occlude up to 93.5 percent of the area of the third region and still maintain uniqueness.

Denote the points of intersection of the occluding half-plane with the circle by v_α and v_β (where α and β denote the angles that v_α and v_β , respectively, create with the x -axis, assuming the origin is set at the center of the circle; see Fig. 9). The half-plane divides the circle into two portions, the visible part, whose angular extent is $\alpha - \beta$, and the occluded part, whose angular extent is $2\pi - \alpha + \beta$. Note that to destroy uniqueness, the occluded part must include the center of the circle. Denote $\phi = \alpha - \beta$. The area of the visible part is given by

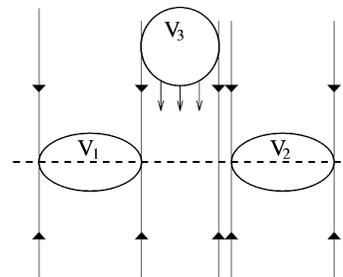


Fig. 8. The effect of contracting two ellipses (V_1 and V_2) inside themselves on a third region. This figure illustrates the case that the external fixed point is at infinity. The lower half of the circle (V_3) prevents the contraction. Consequently, to destroy uniqueness the entire lower half of the circle must be occluded.

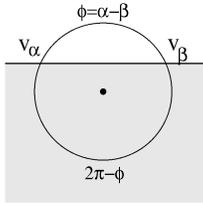


Fig. 9. A circle partly occluded by a half-plane. v_α and v_β denote the points of intersection of the half-plane with the circle. The angular extent of the visible portion is $\phi = \alpha - \beta$ and the angular extent of the occluded portion is $2\pi - \phi$.

$$0.5(\phi - \sin \phi). \quad (10)$$

The occluded part complements the visible portion to π and so its area is given by

$$\pi - 0.5(\phi - \sin \phi). \quad (11)$$

To compute the expected area of destructive occlusion, we need to integrate this measure over all choices of α and β in the upper portion of the circle. This can be achieved by varying α between 0 and π and β between 0 and α . The measure of such choices is $\pi^2/2$. Consequently, the expected area of destructive occlusion is given by

$$A = \frac{2}{\pi^2} \int_0^\pi \int_0^\alpha \left(\pi - \frac{\alpha - \beta}{2} + \frac{1}{2} \sin(\alpha - \beta) \right) d\beta d\alpha = \frac{5\pi}{6} + \frac{1}{\pi}. \quad (12)$$

And since the total area of the circle is π , we get that this area composes

$$F = \frac{A}{\pi} = \frac{5}{6} + \frac{1}{\pi^2} \approx 0.9347 \quad (13)$$

of the area of the circle. Consequently, the expected area of occlusion that can be tolerated by the scheme is about 0.9347 (93.47 percent).

This analysis also holds in most cases when the fixed line intersects the third region. In general, in such cases, the boundary of the circle includes two portions that prevent contraction so that both portions must be occluded to destroy uniqueness. The likelihood of that is the same as the likelihood of destroying uniqueness when the fixed line does not intersect the third region. An exception to that is the case that the tangent to the boundary of the region at one of the two intersection points happens to be pointing toward the external fixed point. In this case, the likelihood of destroying uniqueness increases considerably. Nonetheless, there are only measure zero models that are subject to this condition.

A similar analysis, with similar results, can be applied in case the fixed point is located at a finite position.

To provide a more detailed sense of how much information is really needed to determine pose, we have also run controlled simulations in which we have varied the number of unoccluded and partially occluded regions. First, we generated random convex hexagons by choosing six vertices on the boundary of the unit circle and then applying a randomly chosen 2×2 linear transformation with eigenvalues chosen from a uniform distribution between zero and one. These served as our model regions. We formed models by arranging the centers of the regions on grid points in a five by five rectangle (we tried varying the size of the rectangle,

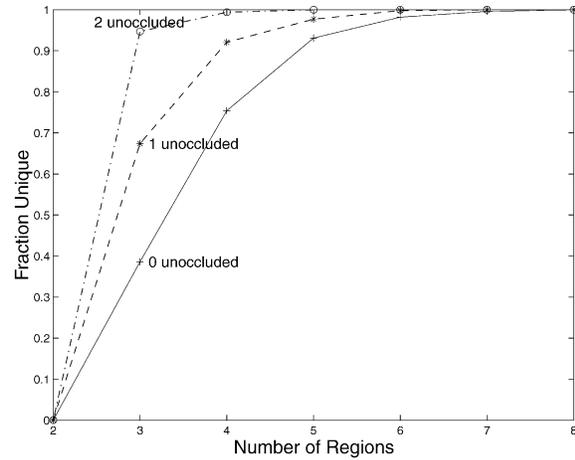


Fig. 10. Results of simulations. Three curves show cases when zero, one, or two regions are completely unoccluded. The fraction of times that models generate unique solutions are shown as a function of the total number of regions.

but this seemed to have little effect). We varied the total number of regions from two to eight, with zero, one, two, or three of these completely unoccluded and the rest randomly occluded. Each occluded region was partially covered by a half-plane. Half-planes were chosen from a uniform distribution, subject to the constraint that they occlude the chosen region. Since each half-plane and its complement are equally likely occluders, on average, half of each region's area was occluded and half of its perimeter was occluded. Randomly chosen occluded models were then checked to see whether they yield unique poses.

The results of these simulations are shown in Fig. 10. Each data point is based on 1,000 simulations. First, when three unoccluded regions were chosen, this always led to a unique solution, though this is not shown in the figure. When only two regions were used, a unique solution was never produced. These results are predicted by our theoretical results.

However, we also see that a small number of occluded regions can quickly lead to a unique solution. For example, when we use five regions that are all occluded, we obtain a unique solution more than 90 percent of the time. Roughly speaking, we see that two regions occluded on average by half contribute about as much to a unique solution as does one unoccluded region and that about two and a half regions are needed to usually produce a unique solution. In Fig. 11, we show an example of a real image in which four occluded regions are used to find a unique solution.

The results of this section demonstrate that although three unoccluded regions are needed to guarantee uniqueness, in practice our algorithm can also combine constraints from multiple partially occluded regions to produce a unique pose. In fact, even when each region is partially occluded, we can obtain correct poses by combining constraints derived from each region that are correct, even in the presence of occlusion. This is in contrast to many other methods of pose determination.

4 CONCLUSION

In this paper, a method for recognition that uses regions to determine the pose of objects was analyzed. We considered the case of planar objects undergoing projective



Fig. 11. (a) A new picture of the Kandinsky with occlusion. (b) The pose that is determined using the four shaded regions (the sun, the tree top, and both domes), which are all partially occluded.

transformations and proved that the method produces unique, correct solutions in previously unexplored situations. Our results were obtained by looking at the flow patterns of points under 2D projective transformations in the presence of fixed points. We have proven that three visible regions are sufficient to determine the transformation uniquely. In addition, by considering general pairs of triangles and ellipses, we showed that two regions are often insufficient to determine the correct transformation uniquely. However, when three regions are available we showed that, depending on the model, it is possible to occlude significant portions of the area of one of the regions without destroying uniqueness. Furthermore, our simulations indicate that with a larger number of regions it is possible to occlude more of the regions and still recover the pose. In fact, when the number of regions available is at least five pose can often be determined even if all of the regions are partly occluded.

These results indicate that by using only a small number of regions it is possible to determine the pose of objects in the presence of significant occlusion. This is particularly interesting because when we use the backward constraints we do not need to specify the location of occlusion. Consequently, in many situations it is not mandatory that we identify the locations of occlusion in order to determine the pose of objects in the image. The backward constraints implicitly account for these occlusions and if sufficient information about the shape of the object appears in its nonoccluded portions then the pose of the object can be recovered by solving the backward constraints.

ACKNOWLEDGMENTS

The research of Ronen Basri was supported in part by the Israeli Ministry of Science, grant no. 6281 and by the Unites States-Israel Binational Science Foundation, grant no. 94-100. The vision group at the Weizmann Institute is supported in part by the Israeli Ministry of Science, grant no. 8504. Ronen Basri is an incumbent of Arye Dissentshik Career Development Chair at the Weizmann Institute.

REFERENCES

- [1] R. Basri and D.W. Jacobs, "Recognition Using Region Correspondences," *Int'l J. Computer Vision*, vol. 25, no. 2, pp. 141-162, 1997.
- [2] R. Basri and D.W. Jacobs, "Projective Alignment with Regions," *Int'l Conf. Computer Vision*, 1999.
- [3] L. Brand, *Differential and Difference Equations*. John Wiley, 1966.
- [4] S. Carlsson, "Information in the Geometric Structure of Retinal Flow Fields," *Proc. Int'l Conf. Computer Vision*, pp. 629-633, 1989.
- [5] J.B. Conway, *A Course in Functional Analysis*. Springer-Verlag, 1990.

- [6] H.S.M. Coxeter, *The Real Projective Plane*. Springer-Verlag, 1993.
- [7] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, John Wiley and Sons, 1973.
- [8] D. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell, "Invariant Descriptors for 3D Object Recognition and Pose," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 10, pp. 971-991, Oct. 1991.
- [9] R. Horaud, "New Methods for Matching 3D Objects with Single Perspective Views," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 9, no. 3, pp. 401-412 Mar. 1987.
- [10] M.K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans. Information Theory*, vol. 8, pp. 169-187, 1962.
- [11] D.W. Jacobs and R. Basri, "3D to 2D Recognition with Regions," *Proc. Int'l J. Computer Vision*, vol. 34, nos. 2/3, pp. 123-145, 1999.
- [12] K. Kanatani, *Geometric Computation for Machine Vision*. Oxford Univ. Press, 1993.
- [13] M. Pollefeys and L. Van Gool, "A Stratified Approach to Metric Self-Calibration," *Computer Vision and Pattern Recognition*, pp. 407-412, 1997.
- [14] F. Schaffalitzky and A. Zisserman, "Geometric Grouping of Repeated Elements within Images," *Proc. Ninth British Machine Vision Conf.*, 1998.
- [15] R. Seidel, "Linear Programming and Convex Hulls Made Easy," *Proc. Sixth Annual Symp. Computational Geometry*, pp. 211-215, 1990.
- [16] G. Strang, *Linear Algebra and Its Applications*. Harcourt, Brace, Jovanovich, 1988.
- [17] A. Verri, F. Girosi, and V. Torre, "Mathematical Properties of the Two-Dimensional Motion Field: From Singular Points to Motion Parameters," *J. Optical Soc. Am.*, vol. 6, no. 5, pp. 698-712, 1989.
- [18] I. Weiss, "Geometric Invariants and Object Recognition," *Int'l J. Computer Vision*, vol. 10, no. 3, pp. 207-231, 1993.



Ronen Basri received the BSc degree in mathematics and computer science from Tel Aviv University in 1985, where he graduated summa cum laude. He received the PhD degree in computer science from the Weizmann Institute of Science in 1990. From 1990 to 1992, he was a postdoctoral fellow at the Massachusetts Institute of Technology in the Department of Brain and Cognitive Science and the Artificial Intelligence Laboratory under the McDonnell-Pew and the Rothchild programs. Since then, he has been affiliated with the Weizmann Institute of Science in the Department of Computer Science and Applied Mathematics, where he is an associate professor, incumbent of Arye Dissentshik Career Development Chair. From 1999 to 2000, he spent a sabbatical at NEC Research Institute in Princeton, New Jersey. His research has focused on computer vision, especially in the areas of object recognition, perceptual organization, and visually guided navigation. He is member of the IEEE.



David W. Jacobs received the BA degree from Yale University in 1982. From 1985 to 1992, he attended Massachusetts Institute of Technology, where he received the MS and PhD degrees in computer science. From 1982 to 1985, he worked for Control Data Corporation on the development of database management systems and attended graduate school in computer science at New York University. Since then, he has been at NEC Research Institute, where he is a senior research scientist. In 1998, he spent a sabbatical at the Royal Institute of Technology (KTH) in Stockholm. Dr. Jacobs has also been a visiting member of the Courant Institute at New York University, where he has taught classes in computer vision and learning. His research has focused on human and computer vision, especially in the areas of object recognition and perceptual organization. He has also published articles in the areas of motion understanding, memory and learning, and computational geometry. He is an associate editor of the *IEEE Transactions of Pattern Analysis and Machine Intelligence*. He and his coauthors received honorable mention for the best paper award at CVPR 2000. He is member of the IEEE.