Mapping Image Properties into Shape Constraints: Skewed Symmetry, Affine-Transformable Patterns, and the Shape-from-Texture Paradigm

John R. Kender and Takeo Kanade
Computer Science Department
Carnegie-Mellon University
Pittsburgh, Pa. 15213

1. Introduction

Certain image properties, such as parallelisms, symmetries, and repeated patterns, provide cues for perceiving the 3-D shape from a 2-D picture. This paper demonstrates how we can map those image properties into 3-D shape constraints by associating appropriate assumptions with them and by using appropriate computational and representational tools.

We begin with the exploration of how one specific image property, "skewed symmetry", can be defined and formulated to serve as a cue to the determination of surface orientations. Then we will discuss the issue from two new, broader viewpoints. One is the class of Affine-transformable patterns. It has various interesting properties, and includes skewed symmetry as a special case. The other is the computational paradigm of shape-from-texture. Skewed symmetry is derived in a second, independent way, as an instance of the application of the paradigm.

This paper further claims that the ideas and techniques presented here are applicable to many other properties, under the general framework of the shape-from-texture paradigm, with the underlying meta-heuristic of non-accidental image properties.

2. Skewed Symmetry

In this section we assume the standard orthographic projections from scene to image, and a knowledge of the gradient space (see [4]).

Symmetry in a 2-D picture has an axis for which the opposite sides are reflective; in other words, the symmetrical properties are found along the transverse lines perpendicular to the symmetry axis. The concept *skewed symmetry* is introduced by Kanade [1] by relaxing this condition a little. It means a class of 2-D shapes in which the symmetry is found along lines not necessarily perpendicular to the axis, but at a fixed angle to it. Formally, such shapes can be defined as 2-D Affine transforms of real symmetries. Figures 2-1(a)(b) show a few key examples.

Stevens [5] presents a good body of psychological experiments which suggests that human observers can perceive surface orientations from figures with this property. This is probably because such qualitative symmetry in the image is often due to real symmetry in the scene. Thus let us associate the following assumption with this image property:

> "A skewed symmetry depicts a real symmetry viewed from some unknown view angle."

Note that the converse of this assumption is always true under orthographic projection.

We can transform this assumption into constraints in the gradient space. As shown in Figure 2-1, a skewed symmetry defines two directions: let us call them the skewed-symmetry axis and the skewed-transverse axis, and denote their directional angles in the picture by α and β, respectively (Figure 2-1(c)). Let $G = (p,q)$ be the gradient of the plane which includes the skewed symmetry. We will show that

$$p'^2\cos^2\left(\frac{\alpha-\beta}{2}\right) - q'^2\sin^2\left(\frac{\alpha-\beta}{2}\right) = -\cos(\alpha-\beta) \qquad (1)$$

where
$$p' = p\cos\lambda + q\sin\lambda$$
$$q' = -p\sin\lambda + q\cos\lambda$$
$$\lambda = (\alpha+\beta)/2.$$

Thus, the (p,q)'s are on a hyperbola. That is, the skewed symmetry defined by α and β in the picture can be a projection of a real symmetry if and only if the gradient is on this hyperbola. The skewed symmetry thus imposes a one-dimensional family of constraints on the underlying surface orientation (p,q).

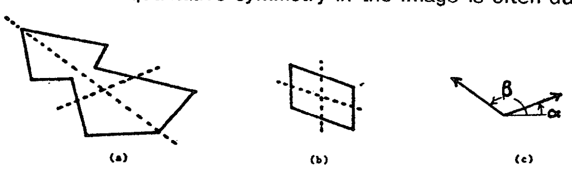

Figure 2-1: Skewed symmetry

3. Affine-Transformable Patterns

In texture analysis we often consider small patterns (texel = texture element) whose repetition constitutes "texture". Suppose we have a pair of texel patterns in which one is a 2-D Affine transform of the other; we call them a pair of *Affine-transformable* patterns. Let us assume that

> "A pair of Affine-transformable patterns in the picture are projection of similar patterns in the 3-D space (i.e., they can be overlapped by scale change, rotation, and translation)".

Note that, as in the case of skewed symmetry, the converse of this assumption is always true under orthographic projections. The above assumption can be schematized by Figure 3-1.

Consider two texel patterns P_1 and P_2 in the picture, and place the origins of the x-y coordinates at their centers, respectively. The transform from P_2 to P_1 can be expressed by a regular 2x2 matrix $A = (a_{ij})$. P_1 and P_2 are projections of patterns P'_1 and P'_2 which are drawn on the 3-D surfaces. We assume that P'_1 and P'_2 are small enough so that we can regard them as being drawn on small planes. Let us denote the gradients of those small planes by $G_1 = (p_1,q_1)$ and $G_2 = (p_2,q_2)$, respectively; i.e., P'_1 is drawn on a plane $-z = p_1x + q_1y$ and P'_2 on $-z = p_2x + q_2y$.

Now, our assumption amounts to saying that P'_1 is transformable from P'_2 by a scalar scale factor σ and a rotation matrix $R = \left(\begin{smallmatrix}\cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha\end{smallmatrix}\right)$. (We can omit the translation from our consideration, since for each pattern the origin of the coordinates is placed at its gravity center, which is invariant under the Affine-transform). Thinking about a pattern drawn on a small plane, $-z = px + qy$, is equivalent to viewing the pattern from directly overhead; that is, rotating the x-y-z coordinates so that the normal vector of the plane is along the new z-axis (line of sight). For this purpose we rotate the coordinates first by φ around the y-axis and then by θ around the x-axis. We have the following relations among φ, θ, p, and q:

$$\sin\varphi = p/\sqrt{p^2+1} \qquad \cos\varphi = 1/\sqrt{p^2+1} \qquad (2)$$
$$\sin\theta = q/\sqrt{p^2+q^2+1} \qquad \cos\theta = \sqrt{p^2+1}/\sqrt{p^2+q^2+1}$$

The plane which was represented as $-z = px + qy$ in the old coordinates is, of course, now represented as $-z' = 0$ in the new coordinates.

Let us denote the angles of the coordinate rotations to obtain P'_1 and P'_2 in Figure 3-1 by (φ_1, θ_1) and (φ_2, θ_2), individually. The 2-D mapping from P'_i (x'-y' plane) to P_i (x-y plane) can be conveniently represented by the following 2x2 matrix T_i which is actually a submatrix of the usual 3-D rotation matrix.

$$T_i = \begin{pmatrix} \cos\varphi & -\sin\varphi\sin\theta \\ 0 & \cos\theta \end{pmatrix}$$

Now, in order for the schematic diagram of Figure 3-1 to hold, what relationships have to be satisfied among the matrix $A = (a_{ij})$, the gradients $G_i = (p_i, q_i)$ for i = 1,2, the angles (φ_i, θ_i) for i = 1,2, the scale factor σ, and the matrix R ? We equate the two transforms that start from P'_2 to reach at P_1: one following the diagram counter-clockwise $P'_2 \to P_2 \to P_1$, the other clockwise $P'_2 \to P'_1 \to P_1$. We obtain

$$AT_2 = T_1 \sigma R.$$

By eliminating σ and α, and substituting for sines and cosines of φ_i and θ_i by (2), we have two (fairly complex) equations in terms of p_i, q_i, and the elements of A. We therefore find that the assumption of Affine-transformable patterns yields a constraint determined solely by the matrix A. The matrix is determined by the relation between P_2 and P_1 *observable* in the picture: without

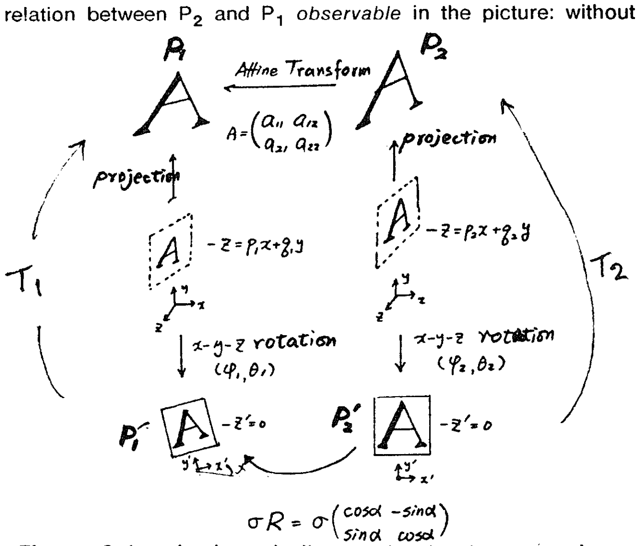

Figure 3-1: A schematic diagram showing the assumptions on Affine transformable patterns.

knowing either the original patterns (P'_1 and P'_2) or their relationships (σ and R) in the 3-D space.

The Affine transform from P_2 to P_1 is more intuitively understood by how a pair of perpendicular unit-length vectors (typically along the x and y coordinate axes) are mapped into their transformed vectors. Two angles (α and β) and two lengths (τ and ρ) can characterize the transform. Components of the transformation matrix $A = (a_{ij})$ are represented by:

$$a_{11} = \tau\cos\alpha \quad a_{12} = \rho\cos\beta \qquad (3)$$
$$a_{21} = \tau\sin\alpha \quad a_{22} = \rho\sin\beta$$

Let us consider the case that α and β are known, but τ and ρ are not. Using (3), eliminate τ and ρ. Then, we obtain

$$(p_1\cos\alpha + q_1\sin\alpha)(p_1\cos\beta + q_1\sin\beta) + \cos(\alpha-\beta) = 0$$

which is exactly the same as the hyperbola (1).

4. The Shape-from-Texture Paradigm

This section derives the same skewed-symmetry constraints from a second theory, different from the Affine-transformable patterns. The shape-from-texture paradigm is briefly presented here; a fuller discussion can be found in [3].

The paradigm has two major portions. In the first, a given image textural property is "normalized" to give a general class of surface orientation constraints. In the second, the normalized values are used in conjunction with assumed scene relations to refine the constraints. Only two texels are required, and only one assumption (equality of scenic texture objects, or some other simple relation) to generate a well-behaved one-dimensional family of possible surface orientations.

The first step in the paradigm is the normalization of a given texel property. The goal is to create a normalized texture property map (NTPM), which is a representational and computational tool relating image properties to scene properties. The NTPM summarizes the many different conditions that may have occurred in the scene leading to the formation of the given texel. In general, the NTPM of a certain property is a scalar-valued function of two variables. The two input variables describe the postulated surface orientation in the scene (top-bottom and left-right slants: (p,q) when we use the gradient space). The NTPM returns the value of the property that the textural object would have had in the scene, in order for the image to have the observed textural property. As an example, the NTPM for a horizontal unit line length in the image summarizes the lengths of lines that would have been necessary in 3-D space under various orientations: at surface orientation (p,q), it would have to be $\sqrt{p^2+1}$.

More specifically, the NTPM is formed by selecting a texel and a texel property, back-projecting the texel through the known imaging geometry onto all conceivable surface orientations, and measuring the texel property there.

In the second phase of the paradigm, the NTPM is refined in the following way. Texels usually have various orientations in the image, and there are many different texel types. Each texel generates its own image-scene relationships, summarized in its NTPM. If, however, assumptions can be made to relate one texel to another, then their NTPMs can also be related; in most cases only a few scenic surface orientations can satisfy *both* texels' requirements. Some examples of the assumptions that relate texels are: both lie in the same plane, both are equal in textural measure (length, area, etc.), one is k times the other in measure, etc. Relating texels in this manner forces more stringent demands on the scene. If enough relations are invoked, the orientation of the local surface supporting two or more related texels can be very precisely determined.

What we now show is that the skewed symmetry method is a special case of the shape-from-texture paradigm; it can be derived from considerations of texel slope.

To normalize the slope of a texel, it is back-projected onto a plane with the postulated orientation. The angle that the back-projected slope makes with respect to the gradient vector of the plane is one good choice (out of many) for the normalized slope measure. Under perspective, the normalized value depends on the image position and camera focal length; under orthography it is much simpler.

Using the construction in Figure 4-1, together with several lemmas relating surfaces in perspective to their local vanishing lines, slope is normalized as follows. Assume a slope is parallel to the p axis; the image and gradient space can always be rotated into such a position. (If rotation is necessary, the resulting NTPM can be de-rotated into the original position using the standard two-by-two orthonormal matrix.) Also assume that the slope is somewhere along the line $y = y_s$, where the unit of measurement in the image is equal to one focal length. Then, the normalized slope value--the normalized texture property map -- is given by

$$[q - y_s(p^2 + q^2)] / [p\sqrt{(1 + p^2 + q^2)}].$$

This normalized value can be exploited in several ways. Most important is the result that is obtained when one has two slopes in the image that are assumed to arise from equal slopes in the scene. Under this assumptions, their normalized property maps can be equated. The resulting constraint, surprisingly, is a simple straight line in the gradient space.

Under orthography, nearly everything simplifies. The normalized slope of a texel becomes

$$q / [p\sqrt{(1 + p^2 + q^2)}]. \tag{4}$$

It is independent of y_s; in effect, all slopes are at the focal point.

Consider Figure 2-1(a). Given the angle that the two texels form $(\beta-\alpha)$, rotate the gradient space so that the positive p axis bisects the angle. Let each half-angle be δ, so $\delta = (\beta-\alpha)/2$. Calculating the normalized value of either slope is obtained directly from the standard normalized slope formula, corrected for the displacement of $+\delta$ and $-\delta$ respectively. That is, for the slope at the positive δ orientation, instead of formula (4), we use the

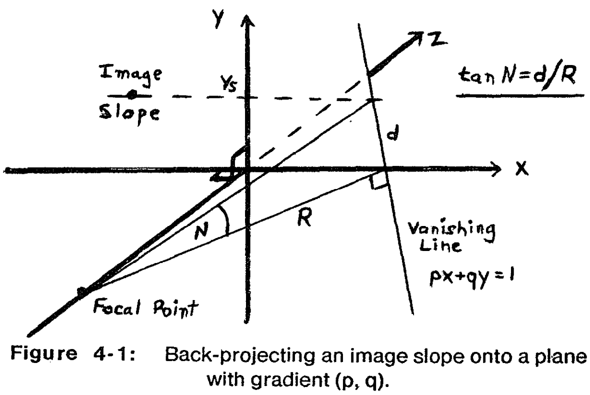

Figure 4-1: Back-projecting an image slope onto a plane with gradient (p, q).

formula under the substitution $p\cos\delta + q\sin\delta$ for p, $-p\sin\delta + q\cos\delta$ for q. We do similarly for the slope at $-\delta$.

The fact that the normalized slopes are assumed to be perpendicular in the scene allows us to set one of the normalized values equal to the negative reciprocal of the other. The resultant equation becomes

$$p^2\cos^2\delta - q^2\sin^2\delta = \sin^2\delta - \cos^2\delta = -\cos 2\delta.$$

This is exactly the hyperbola (1) with $2\delta = \beta-\alpha$.

5. Conclusion

The assumptions we used for the skewed symmetry, the Affine-transformable patterns, and texture analysis can be generalized as

> "Properties observable in the picture are not by accident, but are projections of some preferred corresponding 3-D properties."

This provides a useful meta-heuristic for exploiting image properties: we can call it the meta-heuristic of *non-accidental image properties*. It can be regarded as a generalization of general view directions, often used in the blocks world, to exclude the cases of accidental line alignments.

Instances that can fall within this meta-heuristic includes: parallel lines in the picture vs. parallel lines in the scene, texture gradients, and lines convergent to a vanishing point.

One of the most essential points of our technique is that we related certain image properties to certain 3-D space properties, and that we map the relationships into convenient representations of shape constraints. We explicitly incorporate assumptions based either on the meta-heuristic or on *apriori* knowledge of the world. The shape-from-texture paradigm provides a computational framework for our technique. In most part of our discussion we assumed orthography. Similar--though more involved and less intuitive--results can be obtained under perspective projections.

This work is further discussed in a technical report with the same title as this paper.

References

[1] Kanade, T.
Recovery of the 3-Dimensional Shape of an Object from a Single View.
Technical Report CMU-CS-79-153, Computer Science Department, Carnegie-Mellon University, Oct., 1979.

[2] Kender, J.R.
Shape from Texture.
PhD thesis, Computer Science Department, Carnegie-Mellon University, 1980.

[3] Mackworth, A.K.
Interpreting Pictures of Polyhedral Scenes.
Artificial Intelligence 4(2), 1973.

[4] Stevens, K.A.
Surface Perception from Local Analysis of Texture and Contour.
Technical Report AI-TR-512, MIT-AI, 1980.