

# Extracting the Affine Transformation from Texture Moments

Jun Sato and Roberto Cipolla

Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, England

**Abstract.** In this paper we propose a novel, efficient and geometrically intuitive method to compute the four components of an affine transformation from the change in simple statistics of images of texture. In particular we show how the changes in first, second and third moments of edge orientation and changes in density are directly related to the rotation (curl), scale (divergence) and deformation components of an affine transformation. A simple implementation is described which does not require point, edge or contour correspondences to be established. It is tested on a wide range of repetitive and non-repetitive visual textures which are neither isotropic nor homogeneous. As a demonstration of the power of this technique the estimated affine transforms are used as the first stage in shape from texture and structure from motion applications.

## 1 Introduction

The estimation of an affine transformation is often an integral part in structure from motion (or stereo) and shape from texture. In structure from motion relative motion between the viewer and scene induces distortion in image, and in small neighbourhoods this distortion can be described by an image translation and a four parameter *affine transformation* [6]. In shape from texture the distortion in an image of a surface with a repeated texture pattern can also be modelled by affine transformations [4, 9].

Many methods have been proposed to extract the affine transformations. The simplest method is based on the accurate extraction of points or lines and their correspondences. This requirement of correspondences becomes a non-trivial problem in densely textured images. Cipolla and Blake [3] presented a novel method to recover the affine transformation from image contours. Although this method did not require point or line correspondences the extraction and tracking of closed contours is also not always possible in richly textured images.

A large number of techniques have been developed which avoid the explicit correspondence of features. For small visual motions or distortions a common method is to estimate the affine transform from spatiotemporal gradients of image intensity [1]. The amount of visual motion allowed is limited by the smoothing scale factor. For estimating the *texture distortion* map Malik and Rosenholtz [9] have attempted to solve for the affine transformation in the fourier domain although this involves the choice of a suitable window and is computationally expensive. Under the assumptions of directional isotropy [10] it is

possible to estimate the surface orientation from the second moment matrix of image element orientations [5, 2]. Modifications of the second moment matrix which also exploit image intensity gradients have also been used [8]. However it is impossible to recover the affine transformation (four independent parameters) uniquely from the second moment matrix.

In this paper we propose a novel, efficient and geometrically intuitive method to compute the four components of an affine transformation from the change in simple statistics of the images of texture. A simple implementation is described which does not require correspondences to be established. It is tested on a wide range of repetitive and non-repetitive visual textures which are neither isotropic nor homogeneous. The estimated affine transform is also used in shape from texture and structure from motion applications.

## 2 Theoretical Framework

### 2.1 Decomposition of the Affine Transformation

Generally, an affine transformation,  $A$ , can be described by the product of an isotropic scale,  $S$ , and matrix,  $U$ , whose determinant is equal to one. Furthermore, the matrix,  $U$ , can be decomposed into a symmetric matrix,  $D$ , which we will call the geometric deformation and an asymmetric 2D rotation matrix,  $R$ . An affine matrix can thus be described with these three fundamental transformations:

$$A = S(s)R(\theta)D(\alpha, \mu) \quad (1)$$

where the isotropic scale,  $S$ , is specified by a scale parameter,  $s$ , and the rotation,  $R$ , is specified by an angle,  $\theta$ . The deformation,  $D$ , is specified by an axis of deformation,  $\mu$ , and a magnitude of deformation,  $\alpha$ , and can be described using a rotation,  $R(\mu)$ , and symmetric matrix,  $M(\alpha)$ , whose eigenvalues are  $\alpha$  and  $\frac{1}{\alpha}$ .

$$\begin{aligned} D(\alpha, \mu) &= R(\mu)M(\alpha)R^T(\mu) \\ &= \begin{bmatrix} \cos \mu & -\sin \mu \\ \sin \mu & \cos \mu \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & \frac{1}{\alpha} \end{bmatrix} \begin{bmatrix} \cos \mu & \sin \mu \\ -\sin \mu & \cos \mu \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\alpha} \sin^2 \mu + \alpha \cos^2 \mu & (\alpha - \frac{1}{\alpha}) \sin \mu \cos \mu \\ (\alpha - \frac{1}{\alpha}) \sin \mu \cos \mu & \alpha \sin^2 \mu + \frac{1}{\alpha} \cos^2 \mu \end{bmatrix} \end{aligned} \quad (2)$$

Deformation is equivalent to a pure shear which preserves area, i.e. an expansion by a factor,  $\alpha$ , in the direction,  $\mu$ , with a contraction by the same amount in a perpendicular direction.

### 2.2 Relationship between Changes in Image Orientation and the Affine Transformation

We now investigate the effect of these components of an affine transformation on the orientation of image detail. Consider an element of texture represented by an

unit vector,  $\mathbf{v}$ , with orientation,  $\varphi$ . The affine transformation,  $A$ , transforms the vector,  $\mathbf{v}$ , into  $\mathbf{v}'$  and changes its orientation by  $\Delta\varphi$ . The change in orientation,  $\Delta\varphi$ , can be written as the sum of two components: one due to rotation,  $\theta$ , which changes the orientation of all elements equally and one due to deformation,  $\Delta\varphi_d$ .

$$\Delta\varphi = \theta + \Delta\varphi_d \quad (3)$$

$\Delta\varphi_d$ , can be computed from the vector product ( $\wedge$ ) of  $\mathbf{v}$  and the deformed vector,  $\mathbf{v}'' = D(\alpha, \mu)\mathbf{v}$  as follows:

$$\begin{aligned} \sin(\Delta\varphi_d) &= \frac{|\mathbf{v} \wedge \mathbf{v}''|}{|\mathbf{v}||\mathbf{v}''|} \\ &= \frac{\frac{1}{2} \left( \frac{1}{\alpha} - \alpha \right) \sin 2(\varphi - \mu)}{\sqrt{\frac{1}{\alpha^2} \sin^2(\varphi - \mu) + \alpha^2 \cos^2(\varphi - \mu)}} \end{aligned} \quad (4)$$

Note that the isotropic scale,  $S$ , does not affect orientation, while the change in orientation due to the deformation term depends on the initial orientation,  $\varphi$ , the axis of deformation,  $\mu$ , and magnitude of deformation,  $\alpha$ .

If the magnitude of deformation is small, the change in orientation can be described as follows from (3) and(4).

$$\Delta\varphi \simeq \theta + \frac{1}{2} \left( \frac{1}{\alpha} - \alpha \right) \sin 2(\varphi - \mu) \quad (5)$$

Koenderink and Van Doorn [7] derived a similar approximate equation for small displacements.

### 2.3 Texture Moments under Affine Transformation

In this section, we propose a novel method to calculate the four parameters of the affine transformation reliably without any correspondence of spatial image features using moments of the orientation and density of the texture. In previous work on shape from texture, the texture was often assumed either to be spatially homogeneous or isotropic in orientation, though such textures are limited in the real world. Here, we consider any visual pattern in the real world as a texture, and consider the change in the statistics of the visual texture under an affine transformation.

Consider the texture to have oriented elements with distribution,  $f(\varphi)$ , which will be changed to  $f'(\varphi)$  by an affine transformation. From (5), the rotation term,  $R$ , changes the orientations of the texture elements equally. This means that rotation is related to a shift in the mean value of  $f(\varphi)$  (i.e. the first moment of  $f(\varphi)$ ), and does not affect higher moments. The deformation term, on the other hand, depends on the original orientation of the element and hence affects the variance of  $f(\varphi)$  (i.e. the second moment). Furthermore because there is a term  $\mu$  in the deformation term, we can infer that the changes in the distribution of orientation will not generally be symmetric about the mean of the orientations, (except for the case when  $\mu = 0$ ) and hence the skewness of  $f(\varphi)$  (i.e. the third

moment) will be affected. Thus changes in first moment of orientation are related to the rotation,  $R$ , and the deformation,  $D$ . Changes to the second and third moments are only affected by  $D$ . Scale,  $S$ , affects the area of texture elements and their density, leaving orientations unaffected.

We show below how these simple geometrically intuitive relations can be used to directly recover the parameters of the affine transformation from changes in the density and orientation statistics of image textures.

From (5), an element with orientation,  $\varphi$ , is transformed into  $\varphi'$  such that:

$$\begin{aligned}\varphi' &= \varphi + \theta + \lambda \sin 2(\varphi - \mu) \\ &= \varphi + \theta + \lambda(\sin 2\varphi \cos 2\mu - \cos 2\varphi \sin 2\mu)\end{aligned}\quad (6)$$

where,  $\lambda$  is related to the magnitude term of the deformation by:

$$\lambda = \frac{1}{2} \left( \frac{1}{\alpha} - \alpha \right) \quad (7)$$

The change in the first moment of  $f(\varphi)$ ,  $\Delta I_\varphi$ , in terms of the rotation,  $\theta$ , axis,  $\mu$ , and magnitude,  $\lambda$ , of the deformation is given by summing equation (6) for all elements.

$$\Delta I_\varphi = \lambda(I_{\sin 2\varphi} \cos 2\mu - I_{\cos 2\varphi} \sin 2\mu) + \theta \quad (8)$$

where,  $I_{\sin 2\varphi}$  and  $I_{\cos 2\varphi}$  are the mean values of  $\sin 2\varphi$  and  $\cos 2\varphi$  respectively. If we assume that the deformation is small, that is  $\lambda \ll 1$ , then we can derive the relationships between the changes in second and third moments,  $\Delta I_{\varphi\varphi}$ ,  $\Delta I_{\varphi\varphi\varphi}$ , and the rotation and deformation in closed form.

$$\Delta I_{\varphi\varphi} = 2\lambda(I_\varphi \sin 2\varphi \cos 2\mu - I_\varphi \cos 2\varphi \sin 2\mu) \quad (9)$$

$$\Delta I_{\varphi\varphi\varphi} = 3\lambda(I_{\varphi^2 \sin 2\varphi} \cos 2\mu - I_{\varphi^2 \cos 2\varphi} \sin 2\mu) \quad (10)$$

where  $I_\varphi \sin 2\varphi$  and  $I_\varphi \cos 2\varphi$  are the covariances between  $\varphi$  and  $\sin 2\varphi$ , and  $\varphi$  and  $\cos 2\varphi$  respectively.  $I_{\varphi^2 \sin 2\varphi}$  and  $I_{\varphi^2 \cos 2\varphi}$  are third moments.

The rotation,  $\theta$ , the axis of deformation,  $\mu$ , and the magnitude of deformation,  $\alpha$ , can be computed from:

$$\mu = \frac{1}{2} \tan^{-1} \left( \frac{M_1}{M_2} \right) \quad (11)$$

$$\alpha = \frac{1}{2M_3} \left( \sqrt{M_1^2 + M_2^2 + 4M_3^2} - \sqrt{M_1^2 + M_2^2} \right) \quad (12)$$

$$\theta = \Delta I_\varphi - \frac{1}{2M_3} (I_{\sin 2\varphi} M_2 - I_{\cos 2\varphi} M_1) \quad (13)$$

where:

$$M_1 = 3\Delta I_{\varphi\varphi} I_{\varphi^2 \sin 2\varphi} - 2\Delta I_{\varphi\varphi\varphi} I_\varphi \sin 2\varphi \quad (14)$$

$$M_2 = 3\Delta I_{\varphi\varphi} I_{\varphi^2 \cos 2\varphi} - 2\Delta I_{\varphi\varphi\varphi} I_\varphi \cos 2\varphi \quad (15)$$

$$M_3 = 3(I_{\varphi^2 \cos 2\varphi} I_\varphi \sin 2\varphi - I_{\varphi^2 \sin 2\varphi} I_\varphi \cos 2\varphi) \quad (16)$$

The special case where  $M_3 = 0$  does not occur in practice. The change in scale,  $s$ , of the affine transformation can be obtained from the first moment of density or area of the texture elements [6, 3]. Having computed rotation,  $\theta$ , the axis of deformation,  $\mu$ , the magnitude of deformation,  $\alpha$ , and the change in scale,  $s$ , we have recovered all four independent parameters of the affine transformation.

The properties of the proposed method are: (1) It does not require correspondence of individual image features. (2) This allows much greater interframe motions than spatio-temporal techniques. (3) The method relies on the comparison of statistics of the image patches. This therefore requires that corresponding areas of interest are identified. (4) The recovery of scale from the texture density assumes that the texture is homogeneous. If, instead, we can determine the changes in area of the texture elements [8], this assumption is no longer required.

### 3 Preliminary Results

In this section, we will present several results which show that this method does not need any assumptions like directional isotropy or spatial homogeneity to estimate the rotation and the deformation, though we need homogeneity of the texture to compute the scale component properly. Fig.1 shows the results from this method tested on a wide range of images. To demonstrate the accuracy of the extracted affine transform we have chosen to assume that the original images (upper most images in Fig.1) are of textures on a fronto-parallel plane and we use the affine transformation to estimate the new orientation of the plane assuming it is viewed under *weak* perspective (second row of images in Fig.1). The two ellipses in Fig.1 show that the estimated orientations are qualitatively good even with non-uniform textures. Table 1 compares the accuracy of this method quantitatively. The accuracy is seen to degrade when the texture in the image does not have preferred orientations. This was caused by filtering of orientation to avoid aliasing.

A second method for testing the accuracy of this method is to exploit the results in real applications.

**Shape from Texture:** For a repetitive texture on a curved surface the texture distortion in different directions is well modelled by an affine transform and the scale and deformation components of this affine transform can be used to recover the relative orientations and positions of the surface patches [9]. Fig.2 shows the result of using affine transforms by the method presented in this paper to recover the shape of a curved surface. The proposed method derives qualitatively good results, though there are some errors in the estimated orientations. These errors arise from (1) the difference in the sampling areas, (2) errors caused in the sampling of orientation, (3) the small deformation approximation used in the proposed method.

**Qualitative Visual Navigation:** In the next application, we will show how a moving observer can determine the object surface orientation and time to contact from the affine transformation estimated from texture moments. The relations between the motion parameters and the surface position and orientation

were presented in [3, 6]. A translation along the optical axis towards the surface patch leads to a uniform expansion in the image. This determines the distance to the object which is conveniently expressed as a time to contact. A horizontal translation perpendicular to the visual direction results in image deformation with a magnitude determined by the slant and an axis determined by the tilt of the surface. In this case the axis of deformation bisects the tilt and direction of translation in the image. Table 2 shows the results of estimation of tilt angle of the surface and time to contact computed from Fig.3.

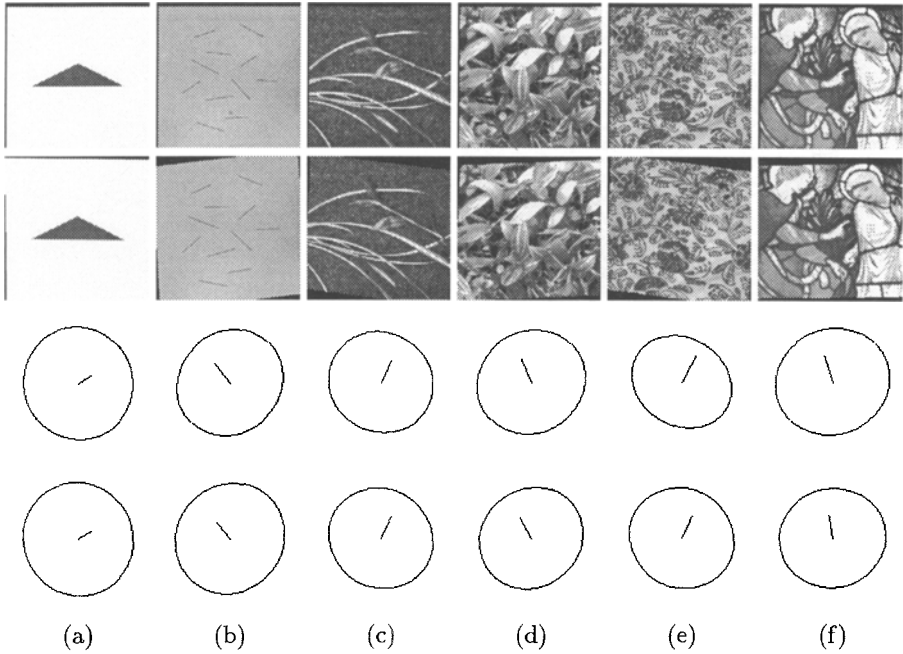
## 4 Conclusions

In this paper we have proposed a novel method to compute the four components of an affine transformation from the changes in moments of edge orientation and density. This method does not require point, edge or contour correspondences to be established, though the problem of selecting the area of interest still remains. It is extremely simple and efficient and the four parameters are linked to changes in orientation and texture density in a geometrically intuitive way.

Preliminary results have been presented and tested in simple applications exploiting the derived affine transformation. The estimated affine transformation has been of useful accuracy. However the sensitivity and error analysis as well as a quantitative comparison with other methods remain to be carried out.

## References

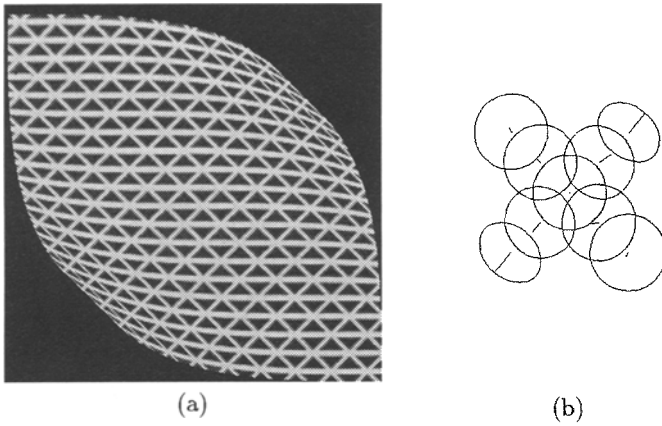
1. P. Anandan. A computational framework and an algorithm for the measurement of visual motion. In *Int. Journal of Computer Vision*, pages 283–310, 1989.
2. A. Blake and C. Marinos. Shape from texture: estimation, isotropy and moments. *Artificial Intelligence*, 45:323–380, 1990.
3. R. Cipolla and A. Blake. Surface orientation and time to contact from image divergence and deformation. In G. Sandini, editor, *Proc. 2nd European Conference on Computer Vision*, pages 187–202. Springer-Verlag, 1992.
4. T. Kanade and J.R. Kender. Mapping image properties into shape constraints: Skewed symmetry, affine-transformable patterns, and the shape-from-texture paradigm. In J.Beck et al, editor, *Human and Machine Vision*, pages 237–257. Academic Press, NY, 1983.
5. K. Kanatani. Detection of surface orientation and motion from texture by a stereological technique. *Artificial Intelligence*, 23:213–237, 1984.
6. J.J. Koenderink. Optic flow. *Vision Research*, 26(1):161–179, 1986.
7. J.J. Koenderink and A.J. Van Doorn. Geometry of binocular vision and a model for stereopsis. *Biological Cybernetics*, 21:29–35, 1976.
8. T. Lindeberg and J. Garding. Shape from texture from a multi-scale perspective. *Proc. 4th Int. Conf. on Computer Vision*, pages 683–691, 1993.
9. J. Malik and R. Rosenholtz. A differential method for computing local shape-from-texture for planar and curved surfaces. *Proc. Conf. Computer Vision and Pattern Recognition*, pages 267–273, 1993.
10. A.P. Witkin. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17:17–45, 1981.



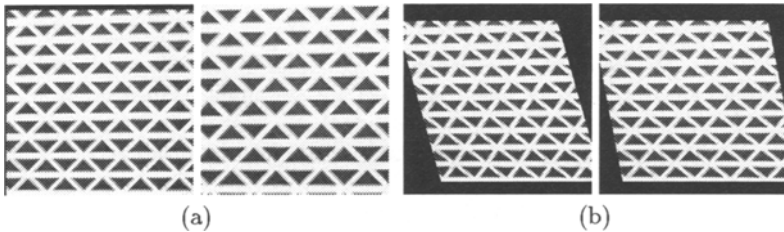
**Fig. 1.** Results of preliminary experiments. Examples of the images distorted by arbitrary affine transformations were processed by our affine transform from texture moments algorithm. Images in the first row and the second row are fronto-parallel images and transformed images respectively after changing the position and orientation of the plane viewed under *weak* perspective. The estimated orientations (upper ellipses) and true orientations (lower ellipses) of the transformed images are shown using normal vectors and oriented circles whose size and shape correspond to the scale change and distortion. Examples include (a) single triangle, (b) randomly oriented lines, (c) oriented grass, (d) leaves, (e) a cloth with texture and (f) stained glass as an example of a non-uniform texture.

**Table 1.** Accuracy of the surface parameters, scale,  $s$ , rotation,  $\theta$ , tilt,  $\tau$ , and slant,  $\sigma$ .

| Images            | True |                    |                  |                    | Estimated |                    |                  |                    |
|-------------------|------|--------------------|------------------|--------------------|-----------|--------------------|------------------|--------------------|
|                   | $s$  | $\theta(^{\circ})$ | $\tau(^{\circ})$ | $\sigma(^{\circ})$ | $s$       | $\theta(^{\circ})$ | $\tau(^{\circ})$ | $\sigma(^{\circ})$ |
| (a) triangle      | 1.0  | 0.0                | 30.0             | 15.0               | 1.0       | 0.0                | 30.0             | 15.0               |
| (b) lines         | 1.0  | -5.0               | 135.0            | 20.0               | 0.98      | -3.8               | 133.0            | 25.1               |
| (c) grass         | 0.95 | 5.0                | 60.0             | 25.0               | 0.95      | 5.1                | 60.3             | 24.8               |
| (d) leaves        | 1.0  | 0.0                | 120.0            | 25.0               | 0.99      | -2.7               | 117.0            | 25.8               |
| (e) cloth         | 0.95 | 5.0                | 60.0             | 25.0               | 0.92      | 5.4                | 56.6             | 34.0               |
| (f) stained glass | 0.95 | 0.0                | 100.0            | 25.0               | 1.02      | -2.3               | 112.3            | 26.9               |



**Fig. 2.** Shape from texture using the affine transform. Surface orientation of patches on a cylindrical object are estimated using the affine transform from texture moments. Estimated local orientations at each point are shown in (b) using oriented circles and their normal vectors.



**Fig. 3.** Two set of images are taken from a moving observer. Motion and optical direction are parallel in (a) and perpendicular in (b). The time to contact and the tilt angle of the surface recovered from the texture moments are shown in table 2.

**Table 2.** Tilt angle and time to contact. Scale,  $s$ , rotation,  $\theta$ , axis and magnitude of the deformation,  $\mu$ ,  $\alpha$ , recovered from the texture moments were used to compute the time to contact,  $t_c$ , in (a) and tilt angle of the surface,  $\tau$ , in (b). Mean values and variances were estimated by changing the area of interest in each image.

| Images       |           | $s$             | $\theta(^{\circ})$ | $\alpha$         | $\mu(^{\circ})$ | $\tau(^{\circ})$ | $t_c$          |
|--------------|-----------|-----------------|--------------------|------------------|-----------------|------------------|----------------|
| (a) parallel | True      | 1.20            | 0.0                | 1.0              | -               | -                | 10.0           |
|              | Estimated | $1.19 \pm 0.07$ | $0.1 \pm 0.1$      | $1.01 \pm 0.01$  | -               | -                | $10.4 \pm 3.2$ |
| (b) dicular  | True      | 1.04            | 3.2                | 1.07             | 28              | 56               | $\infty$       |
|              | Estimated | $1.03 \pm 0.01$ | $1.3 \pm 0.1$      | $1.07 \pm 0.001$ | $33 \pm 2$      | $66 \pm 4$       | $882 \pm 496$  |