Image Registration Using Multi-Scale 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 second circular moments of edge orientation are directly related to the rotation (curl), scale (divergence) and deformation components of an affine transformation, and how these components can be computed from multi-scale texture moments. A simple implementation is described which does not require point, edge or contour correspondences to be established. It is tested on repetitive and non-repetitive visual textures which are neither isotropic nor homogeneous. The theoretical accuracy and the noise sensitivity of this method are compared with other linear moment and circular moment methods.

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. In small neighbourhoods, this distortion can be described by an image translation and a four parameter *affine transformation* [7]. 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 do not require 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

BMVC 1994 doi:10.5244/C.8.9

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 limited to repetitive textures. Under the assumptions of directional isotropy [11] it is possible to estimate the surface orientation from the second moment matrix of image element orientations [2, 5]. 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, although all four parameters of the affine transformation are required for an arbitrary stereo configuration or in structure from motion.

Recently, a novel method [10] was proposed to compute all four parameters of the affine transformation from simple linear moments. Although this method succeeds in deriving the affine transformation of texture images which are neither isotropic nor homogeneous, it suffers from the aliasing problem, that is distribution of orientation is not continuous at 0 and π radians, and the theory breaks down at these points. Although this problem can be avoided by using circular moments, second circular moments are not sufficient to compute all the four parameters of affine transformation. One solution is to use higher moments but these are sensitive to noise.

In this paper we propose a novel, efficient and geometrically intuitive method to compute the four components of an affine transformation only from the change in second circular moments of the images of texture using multi-scale-space representations. This method does not require any correspondence of the image feature to be established, and does not suffer from the aliasing problem. The theoretical accuracy and noise sensitivity of this method are compared with other linear and circular moment methods. A simple implementation is described, and is tested on repetitive and non-repetitive visual textures which are neither isotropic nor homogeneous.

2 Theoretical Framework

In this section, we formalize the method of computing the parameters of an affine transformation from the moments of multi-scale representation. To do this, we first show that the affine transformation can be computed using differential invariants of the image velocity field. Next, we describe how the change in orientation of the image detail is related to the parameters of an affine transformation. This observation is then used to derive the relationship between these parameters and the moments of the orientation of image detail. Unfortunately, this relationship provides only two equations, but there are four unknown parameters of the affine transformation. To compute the affine parameters uniquely we combine this relationship with multi-scale-space representation, and formalize the method to compute all four parameters of an affine transformation from the moments of the orientation of image detail in closed form.

2.1 Differential Invariants of the Image Velocity Field

Generally, an affine transformation, A, which represents the image distortion, can be described by the 2×2 identity matrix, I, and a 2×2 differential component, Q, as follows:

$$A = I + Q \tag{1}$$

If the image distortion is small, the differential component, Q, of the affine transformation which describes the image distortion can be approximated by the first order partial derivatives of the image velocity field or disparity field. Then, the matrix Q can be described using the first order differential invariants of the image velocity field, i.e. curl, divergence and deformation [3]:

$$Q = \frac{\mathrm{div}\vec{\mathbf{v}}}{2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{\mathrm{curl}\vec{\mathbf{v}}}{2}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} + \frac{\mathrm{def}\vec{\mathbf{v}}}{2}\begin{bmatrix}\cos 2\mu & \sin 2\mu\\ \sin 2\mu & -\cos 2\mu\end{bmatrix} \tag{2}$$

where μ is the orientation of the axis of maximum expansion. $\mathrm{curl}\vec{\mathbf{v}}$, $\mathrm{div}\vec{\mathbf{v}}$ and $\mathrm{def}\vec{\mathbf{v}}$ are the curl, divergence and deformation components of the image velocity field, $\vec{\mathbf{v}} = [u, v]$, and are defined by:

$$\begin{aligned} \mathrm{div}\vec{\mathbf{v}} &= u_x + v_y & \qquad \mathrm{def}\vec{\mathbf{v}}\cos 2\mu &= u_x - v_y \\ \mathrm{curl}\vec{\mathbf{v}} &= -u_y + v_x & \qquad \mathrm{def}\vec{\mathbf{v}}\sin 2\mu &= u_y + v_x \end{aligned}$$

where, u_x, u_y, v_x and v_y are the first order partial derivatives of the image velocity field, $\vec{\mathbf{v}}$.

2.2 Changes in Image Orientation and the Affine Transformation

We now investigate the effect of these first order differential invariants on the orientation of image detail.

Consider an element of texture represented by an unit vector, $\mathbf{v}$, with orientation, φ. The affine transformation, A, transforms the vector, $\mathbf{v}$, into $\mathbf{v}'$ with orientation φ'. Therefore, the x and y component of the transformed vector, $\mathbf{v}'$, is described as follows:

$$\begin{aligned} \begin{bmatrix} L'(\varphi)\cos\varphi' \\ L'(\varphi)\sin\varphi' \end{bmatrix} &= A\mathbf{v} \\ &= \begin{bmatrix} 1+s+d_1 & d_2-c \\ d_2+c & 1+s-d_1 \end{bmatrix}\begin{bmatrix}\cos\varphi \\ \sin\varphi\end{bmatrix} \end{aligned} \tag{3}$$

where $L'(\varphi)$ is the length of $\mathbf{v}'$ (note that the length of the transformed vector depends on the original orientation, φ, and is no longer a unit vector), and for simplification we changed the expression of the differential invariants as follows:

$$\begin{aligned} s &= \tfrac{1}{2}\mathrm{div}\vec{\mathbf{v}} & \qquad d_1 &= \tfrac{1}{2}\mathrm{def}\vec{\mathbf{v}}\cos 2\mu \\ c &= \tfrac{1}{2}\mathrm{curl}\vec{\mathbf{v}} & \qquad d_2 &= \tfrac{1}{2}\mathrm{def}\vec{\mathbf{v}}\sin 2\mu \end{aligned}$$

We now compute $\cos 2\varphi'$ and $\sin 2\varphi'$, because we will use second circular moments

in the next stage. (Note that first circular moments are always zero, because we reflect each orientation, φ, of the image detail to $\varphi + \pi$ before computing the moments to avoid an aliasing problem.) From (3), $\cos 2\varphi'$ and $\sin 2\varphi'$ are computed by:

$$\begin{aligned} \cos 2\varphi' &= \cos^2 \varphi' - \sin^2 \varphi' \\ &= \frac{1}{L'(\varphi)^2} (k_{11} \cos 2\varphi + k_{12} \sin 2\varphi + k_{13}) \end{aligned} \tag{4}$$

$$\begin{aligned} \sin 2\varphi' &= 2 \sin \varphi' \cos \varphi' \\ &= \frac{1}{L'(\varphi)^2} (k_{21} \sin 2\varphi + k_{22} \cos 2\varphi + k_{23}) \end{aligned} \tag{5}$$

where:

$$\begin{aligned} k_{11} &= 1 + d_1^2 - d_2^2 + s^2 + 2s - c^2 & \qquad k_{21} &= 1 - d_1^2 + d_2^2 + s^2 + 2s - c^2 \\ k_{12} &= 2(d_1 d_2 - c - cs) & \qquad k_{22} &= 2(d_1 d_2 + c + cs) \\ k_{13} &= 2(d_1 + d_1 s - d_2 c) & \qquad k_{23} &= 2(d_2 + d_2 s + d_1 c) \end{aligned}$$

If the image distortion is small, that is $s \ll 1$, $c \ll 1$, $d_1 \ll 1$ and $d_2 \ll 1$, the second order products of these differential invariants can be neglected. Then, (4) and (5) are approximated to first order by:

$$\cos 2\varphi' \simeq \frac{1}{L'(\varphi)^2} (\cos 2\varphi + 2s \cos 2\varphi - 2c \sin 2\varphi + 2d_1) \tag{6}$$

$$\sin 2\varphi' \simeq \frac{1}{L'(\varphi)^2} (\sin 2\varphi + 2s \sin 2\varphi + 2c \cos 2\varphi + 2d_2) \tag{7}$$

2.3 Texture Moments under Affine Transformation

In this section, we formalize the relationship between the circular moments of the orientation of the texture and the four components of the affine transformation. This allows us to compute an affine transformation without any correspondence of spatial image features. 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. The unit texture elements whose orientations lie in the small interval $(\varphi, \varphi + d\varphi)$ move to the small interval $(\varphi', \varphi' + d\varphi')$ by the affine transformation, and the length of each element changes to $L'(\varphi)$. Then, the number of unit elements, $g(\varphi')$ which lie in the interval $(\varphi', \varphi' + d\varphi')$ is described as follows:

$$g(\varphi')d\varphi' = L'(\varphi) f(\varphi) d\varphi$$

Because $f'(\varphi')$ and $f(\varphi)$ are defined per unit area, the transformed distribution $f'(\varphi')$ is described as follows:

$$\begin{aligned} f'(\varphi')d\varphi' &= \frac{T}{T'} g(\varphi') d\varphi' \\ &= \frac{T}{T'} L'(\varphi) f(\varphi) d\varphi \end{aligned} \tag{8}$$

where T and T' are the total length of the texture elements in the original and distorted area. Second circular moments of the distorted texture, $I_{\sin 2\varphi'}$ and $I_{\cos 2\varphi'}$ can therefore be described as follows by definition and (8):

$$I_{\sin 2\varphi'} = \int_0^{2\pi} \sin 2\varphi' f'(\varphi') d\varphi' = \frac{T}{T'} \int_0^{2\pi} \sin 2\varphi' L'(\varphi) f(\varphi) d\varphi \quad (9)$$

$$I_{\cos 2\varphi'} = \int_0^{2\pi} \cos 2\varphi' f'(\varphi') d\varphi' = \frac{T}{T'} \int_0^{2\pi} \cos 2\varphi' L'(\varphi) f(\varphi) d\varphi \quad (10)$$

Substituting (6) and (7) into (9) and (10):

$$I_{\sin 2\varphi'} = \frac{T}{T'} \int_0^{2\pi} (\sin 2\varphi + 2s \sin 2\varphi + 2c \cos 2\varphi + 2d_2) \frac{1}{L'(\varphi)} f(\varphi) d\varphi \quad (11)$$

$$I_{\cos 2\varphi'} = \frac{T}{T'} \int_0^{2\pi} (\cos 2\varphi + 2s \cos 2\varphi - 2c \sin 2\varphi + 2d_1) \frac{1}{L'(\varphi)} f(\varphi) d\varphi \quad (12)$$

Because $L'(\varphi) \simeq 1$ under small image distortion, $1/L'(\varphi)$ of (11) and (12) can be approximated to first order by:

$$\frac{1}{L'(\varphi)} \simeq 1 - (L'(\varphi) - 1) \quad (13)$$

where $L'(\varphi)$ is computed from (3) by:

$$L'(\varphi) \simeq 1 + d_1 \cos 2\varphi + d_2 \sin 2\varphi + s \quad (14)$$

Substituting (13) and (14) into (11) and (12), and approximating to first order by neglecting the second order products of the differential invariants, second circular moments of the distorted texture are described by second and fourth circular moments of the original texture as follows:

$$I_{\sin 2\varphi'} \simeq \frac{T}{T'} (I_{\sin 2\varphi} + s I_{\sin 2\varphi} + 2c I_{\cos 2\varphi} - \frac{d_1}{2} I_{\sin 4\varphi} + \frac{d_2}{2} (3 + I_{\cos 4\varphi})) \quad (15)$$

$$I_{\cos 2\varphi'} \simeq \frac{T}{T'} (I_{\cos 2\varphi} + s I_{\cos 2\varphi} - 2c I_{\sin 2\varphi} + \frac{d_1}{2} (3 - I_{\cos 4\varphi}) - \frac{d_2}{2} I_{\sin 4\varphi}) \quad (16)$$

where $I_{\sin 2\varphi}$, $I_{\cos 2\varphi}$, $I_{\sin 4\varphi}$ and $I_{\cos 4\varphi}$ denote second and fourth circular moments of the original texture respectively. Note that the moments of distorted texture are described by simple linear combination of the moments of original texture.

2.4 Multi-Scale Texture Moments

In the previous section, we derived the relationship between the moments of the texture and the components of the affine transformation. We have only two equations, and there are four unknown parameters in these equations, i.e. four components of the affine transformation. Due to this lack of constraints, we cannot compute the affine components uniquely from the moments of the texture. In this section, we propose an efficient method to compute four components of the affine transformation reliably, using the moments of scale-space representations of the image.

As shown in the literature [6, 12], we can observe the different image structures in one image using scale-space representation: a certain scale extracts a certain structure in the image and a different scale extracts a different structure. Image features extracted by different scales in an image distorted by an affine transformation belong to physically different image structures, but are distorted by the same affine transformation. This means we can automatically obtain moments which are derived from different image features but are affected by the same affine transformation by choosing different scale-space representations. Combining two or more scale-space representations with derived equations of moments and the affine transformation, we can compute all four components of the affine transformation directly and reliably without any correspondence.

Consider two different scale-space representations, whose scales are t_1 and t_2, to have different texture moments defined by (9), (10), that is $I_{\sin 2\varphi_1}$ and $I_{\cos 2\varphi_1}$ for t_1 and $I_{\sin 2\varphi_2}$ and $I_{\cos 2\varphi_2}$ for t_2, and these moments are changed to $I_{\sin 2\varphi'_1}$, $I_{\cos 2\varphi'_1}$, $I_{\sin 2\varphi'_2}$ and $I_{\cos 2\varphi'_2}$ by the affine transformation. Then, from (15) and (16), the relationship between these texture moments and the components of the affine transformation can be described as follows:

$$\begin{bmatrix} \frac{T'_1}{T_1} I_{\cos 2\varphi'_1} - I_{\cos 2\varphi_1} \\ \frac{T'_1}{T_1} I_{\sin 2\varphi'_1} - I_{\sin 2\varphi_1} \\ \frac{T'_2}{T_2} I_{\cos 2\varphi'_2} - I_{\cos 2\varphi_2} \\ \frac{T'_2}{T_2} I_{\sin 2\varphi'_2} - I_{\sin 2\varphi_2} \end{bmatrix} = \begin{bmatrix} I_{\cos 2\varphi_1} & -2I_{\sin 2\varphi_1} & \frac{1}{2}(3 - I_{\cos 4\varphi_1}) & -\frac{1}{2} I_{\sin 4\varphi_1} \\ I_{\sin 2\varphi_1} & 2I_{\cos 2\varphi_1} & -\frac{1}{2} I_{\sin 4\varphi_1} & \frac{1}{2}(3 + I_{\cos 4\varphi_1}) \\ I_{\cos 2\varphi_2} & -2I_{\sin 2\varphi_2} & \frac{1}{2}(3 - I_{\cos 4\varphi_2}) & -\frac{1}{2} I_{\sin 4\varphi_2} \\ I_{\sin 2\varphi_2} & 2I_{\cos 2\varphi_2} & -\frac{1}{2} I_{\sin 4\varphi_2} & \frac{1}{2}(3 + I_{\cos 4\varphi_2}) \end{bmatrix} \begin{bmatrix} s \\ c \\ d_1 \\ d_2 \end{bmatrix}$$

where T_1 and T_2 are the total length of the original texture elements in the scale t_1 and t_2, and T'_1 and T'_2 are those of the distorted texture elements. We can compute four differential components, s, c, d_1, d_2, of the affine transformation as a solution to this matrix equation (For two different scales, the matrix will not in general be singular). The absolute components of the affine transformation are computed using the derived differential components and (1). This method requires minimal information to compute an affine transformation using two different scales. We can also use more than two scales and raise the accuracy and robustness of this method.

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 will only be meaningful if the two patches are projections of "world" textures with similar properties. This therefore requires that corresponding areas of interest are identified. (4) This method does not suffer from an aliasing problem which occurs when linear moments are used.

3 Theoretical Accuracy and Noise Sensitivity

In this section we compare the systematic error and sensitivity to noise of this method with those of the linear moments method [10] and another circular moment method, Kanatani's stereological method [5]. For comparison purpose, in the following we chose to use the affine transformation (arising from changing the

orientation of a textured surface viewed under *weak* perspective) to compute the surface orientation (i.e. slant and tilt).

These three methods use approximations to solve the problem in closed form, and these approximations cause systematic errors in the estimated orientation of the surface. We investigate these errors using 100 texture elements whose orientations are randomly sampled from a Gaussian distribution with a mean of 0 degrees and a standard deviation of 30 degrees. Fig.1 (a) and (b) show the systematic error in the slant and tilt angles of the proposed method, linear moments method and Kanatani's stereological method. The proposed method exhibits the best accuracy in slant angle and good accuracy in tilt angles, independent of the slant and tilt angles, although the accuracy of Kanatani's stereological method degrades rapidly with slant and tilt angles.

Next we compare the noise sensitivities of these three methods. Random Gaussian noise with a standard deviation of 1.0 degree was added to the orientation data of the texture elements. The errors in slant and tilt angle caused by the random Gaussian noise are shown in Fig.1 (c) and (d). Although all methods are sensitive to noise in the case of small slant, the slant estimated by the proposed method is less affected by noise than that of the other methods. As shown in (d), all methods have similar noise sensitivity with respect to the tilt angle. Although the accuracy and noise sensitivity change in terms of the type of texture, the proposed method is more accurate and less sensitive to noise in most of the cases. It also has the advantage that it does not suffer from the aliasing problem of using linear moments and can be used to compute the rotation component and the change in scale as well.

4 Experiments

We present experimental results which show that this method does not need any assumptions like directional isotropy or spatial homogeneity to estimate the four components of the affine transformation. To demonstrate the accuracy of the extracted affine transform, we have again chosen to assume that the original images are of textures on a fronto-parallel plane and we use the affine transformation to estimate the new orientation of the plane and the scale assuming it is viewed under weak perspective. Fig.2 and Fig.3 shows the results from this method tested on an artificial texture and stained glass images. The center and right ellipses in Fig.2 and Fig.3 show the calibrated and estimated orientations and the changes in scale using shape distortion from the original circle shown in the left, and are qualitatively good even with non-uniform textures. Table 5 compares the accuracy of this method quantitatively for each sample image with the known fiducial orientation. The error seen in the case of the stained glass image is mainly caused by the difference of the sampling points between the original image and the distorted image, that is the difference of the area of interest between the two images.

5 Conclusion

In this paper we have proposed a novel method to compute the four components of an affine transformation from the changes in circular moments of edge orientation. A method to combine the moments of the texture image and the scale-space representation is described. This method does not require any point, edge or contour correspondences to be established, and is simple and efficient. The estimated affine transformation is accurate enough to be useful. However the problem of selecting the area of interest and choice of smoothing scale remain to be investigated.

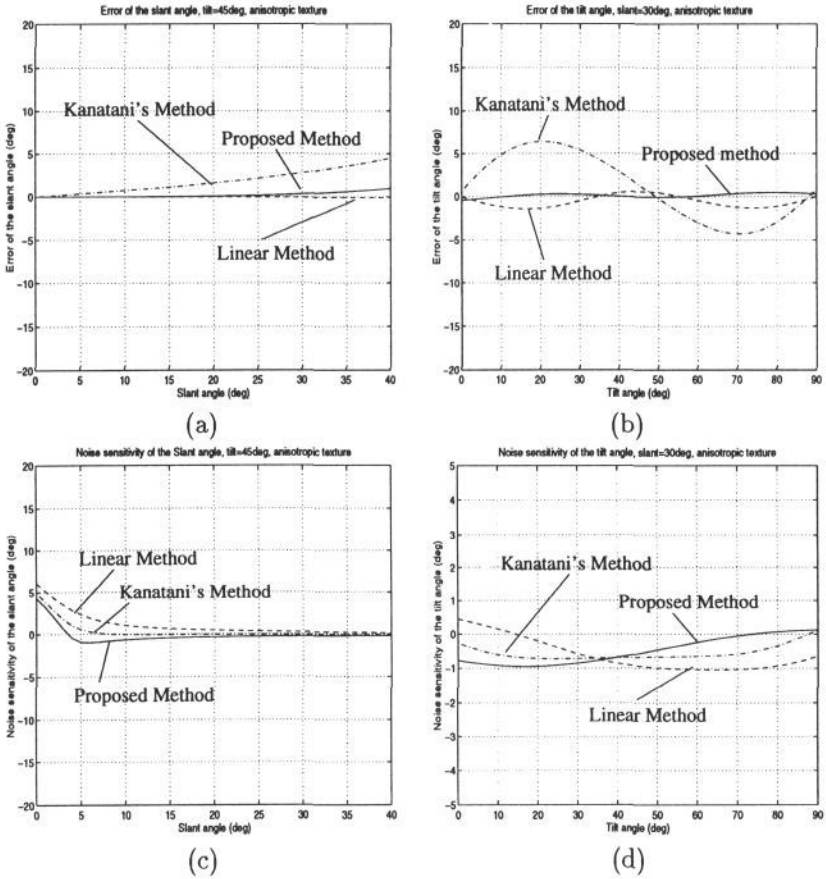

Figure 1: Results of systematic error and noise sensitivity analysis. (a) and (b) show the systematic errors in the estimation of the slant and tilt angles of the surface. The solid line, dashed line and dash-dot line show the error of the proposed method, linear moment method [10] and Kanatani's stereological method [5] respectively. (c) and (d) show errors in slant and tilt angle of the surface caused by additive Gaussian noise with standard deviation of 1.0 degree.

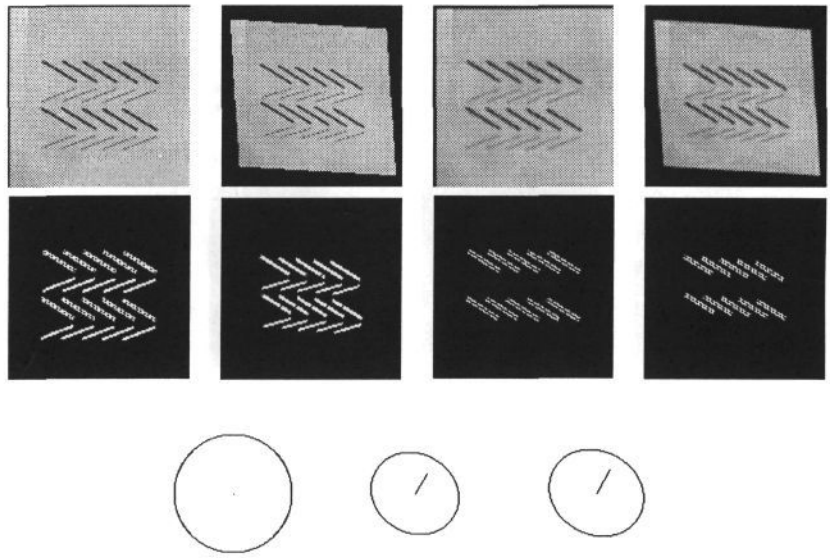

Figure 2: Results of preliminary experiments (artificial texture image). Examples of the images distorted by arbitrary affine transformations were processed by our affine transform from texture moments algorithm. Images in the top row are the original and distorted images of scale (variance) $t = 0$, and the original and distorted images of scale $t = 15$. Images in the second row show the edges detected from the images in the top row. These images show that different scale-space representations have different image structures. The center and right ellipses in the third row show the real and estimated orientation and the change in scale using normal vectors and oriented circles whose size and shape correspond to the scale change and distortion from the original fronto-parallel circle shown in the left.

Table 1: Accuracy of the surface parameters, change in distance, r, rotation, θ, tilt, τ, and slant, σ which are computed from the differential invariants, s, c, d_1, d_2.

Images		r	$\theta(^\circ)$	$\tau(^\circ)$	$\sigma(^\circ)$
(a) **artificial texture**	True	1.10	0.0	60.0	30.0
	Estimated	1.12	0.3	61.1	31.9
(b) **stained glass**	True	1.00	0.0	120.0	30.0
	Estimated	0.97	0.6	123.5	32.4

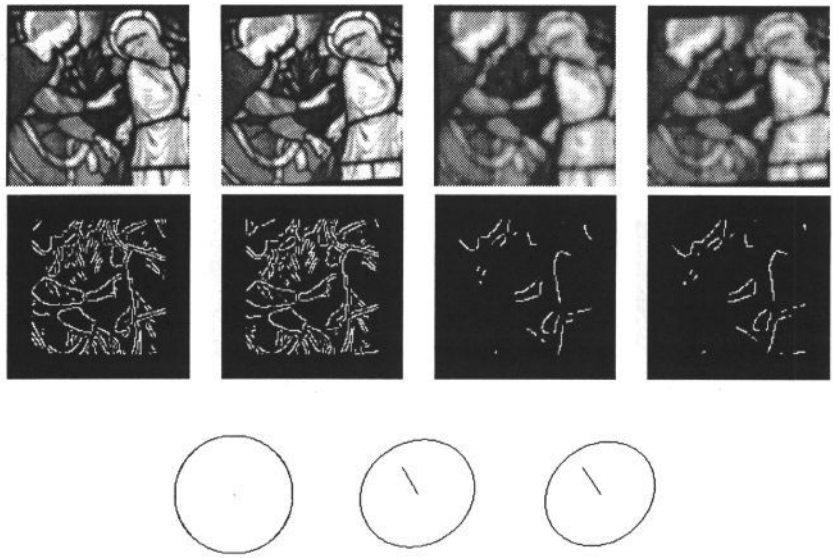

Figure 3: Results of stained glass image. See Fig.2 for the caption.

References

[1] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International 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. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[7] J.J. Koenderink. Optic flow. *Vision Research*, 26(1):161–180, 1986.

[8] T. Lindeberg and J. Garding. Shape from texture from a multi-scale perspective. In *Proc. 4th International Conference 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. In *Proc. Conference on Computer Vision and Pattern Recognition*, pages 267–273, 1993.

[10] J. Sato and R. Cipolla. Extracting the affine transformation from texture moments. In Jan-Olof Eklundh, editor, *Proc. 3rd European Conference on Computer Vision*, volume 2, pages 165–172. Springer–Verlag, 1994.

[11] A.P. Witkin. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17:17–45, 1981.

[12] A.P. Witkin. Space-scale filtering. In *Proc. International Joint Conference on Artificial Intelligence*, pages 1019–1022, 1983.