Extracting the Affine Transformation from Texture Moments

Jun Sato and Roberto Cipolla

CUED/F-INFENG/TR 167

February, 1994

A condensed version of this paper will be presented at ECCV'94

Department of Engineering
University of Cambridge
Trumpington Street
Cambridge CB2 1PZ
England

Email: js2@eng.cam.ac.uk

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

Structure from motion (or stereo) and shape from texture can sometimes be conveniently analysed in a two stage framework. The first stage involves the estimation of image velocities (disparities) in structure from motion (stereo) or *texture gradients* in shape from texture [11]. The second stage involves their interpretation to infer the viewer motion and/or the distance and shape of the visible surfaces.

In structure from motion relative motion between the viewer and scene induces distortion in image detail and apparent shape. In small neighbourhoods and for smooth surfaces this distortion – the image velocity field or disparity field – can be conveniently described by an image translation and a four parameter *affine transformation* [19, 18]. In shape from texture the distortion in a given direction of an image of a surface with a repeated texture pattern – texture gradients – can also be modelled by affine transformations [14, 23].

The estimation of an affine transformation is often an integral part in recovering the image velocity field and *distortion map*. Better still the affine transformation can be decomposed into four components (translation, change in scale (divergence), image rotation (curl) and shear (deformation)) [19] which are related to the scene structure and viewer's motion in a simple geometrically intuitive way.

Many methods have been proposed in the literature to extract the affine transformations from an image sequence (structure from motion) or between different parts of a single image (shape from texture). These can be broadly divided into methods which require the correspondence of features or tokens and methods which exploit the temporal coherence in an image sequence and avoid the correspondence problem.

The simplest method is based on the accurate extraction of points or corner features and their correspondences [16]. The image motion of a minimum of three points (provided they are not colinear) is sufficient to completely define the affine transformation. Unfortunately many corner finders produce poorly localised features which are often temporally unstable [9, 17, 25, 12, 5] although clusters of corners can be used [27]. The technology for edge detection is more advanced than isolated point detection. Reliable, accurate edge detectors which localise surface markings to sub-pixel accuracy [7] can be used to recover the normal (vernier) component of velocity at edges. A minimum of six normal velocities can then be used to compute the affine transformation [24, 28]. Both corner and edge detection and tracking required finding correspondences over the image sequence. This becomes a non-trivial problem when the image motion is large or in densely textured images. Cipolla and Blake [8] presented a novel method to recover the affine transformation of a deforming closed contour from the integral of simple functions of the normal image velocities around image contours. This integration provided some immunity to image measurement noise. This is equivalent to measuring the temporal changes in the area of a closed contour. 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 extraction and explicit correspondence of tokens or features. For small visual motions or distortions a common method is to estimate the affine transform from spatiotemporal gradients of image intensity from the motion constraint equation [22, 1, 3]. The amount of visual motion allowed (especially rotation) is limited by the smoothing scale factor. For larger image motions brute force search techniques (again not requiring correspondence) have been used [13].

For estimating the texture distortion map Malik and Rosenholtz [23] and others [2, 29] 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. A more common approach exploits the *second moment* statistics of image edge orientations. Under the assumptions of directional isotropy [31] in the real texture it is possible to estimate the surface orientation from the second moment matrix of image element orientations [15, 6, 4]. Modifications of the second moment matrix which also exploit image intensity gradients have also been used [21, 10]. It is, of course, impossible to recover the affine transformation (four independent parameters) uniquely from the second moment matrix (which is symmetric and positive semidefinite). (This is equivalent to the "aperture problem in the large" [30] when trying to distinguish the rotation (curl) and pure shear components (deformation) in the distortion of an ellipse under small viewer motions.) In many existing schemes restrictions on the class of texture – isotropy or homogeneity – or on the stereo geometry [10] allowed an incomplete solution. As Malik and Rosenholtz [23] all four affine parameters are needed to completely specify surface position and 3D shape for a general repetitive texture. All four parameters of the affine transformation are also required for an arbitrary stereo configuration or in structure from motion.

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. 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), deformation and scale (divergence) components of an affine transformation. 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. 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.

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.

$$A = SU \tag{1}$$

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:

$$A = S(s)R(\theta)D(\alpha, \mu) \tag{2}$$

where the isotropic scale, S, is specified by a scale parameter, s, and the rotation, R, is specified by an angle, θ:

$$S(s) \quad = \quad \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \tag{3}$$

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (4)$$

The deformation, D, is specified by an axis of deformation, μ, and a magnitude of deformation, α, and can be described using a rotation, $R(\mu)$, and symmetric matrix, $M(\alpha)$, whose eigenvalues are α and $\frac{1}{\alpha}$.

$$\begin{aligned} D(\alpha,\mu) &= R(\mu)M(\alpha)R^T(\mu) \\ &= \begin{bmatrix} \cos\mu & -\sin\mu \\ \sin\mu & \cos\mu \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & \frac{1}{\alpha} \end{bmatrix} \begin{bmatrix} \cos\mu & \sin\mu \\ -\sin\mu & \cos\mu \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\alpha}\sin^2\mu + \alpha\cos^2\mu & \left(\alpha - \frac{1}{\alpha}\right)\sin\mu\cos\mu \\ \left(\alpha - \frac{1}{\alpha}\right)\sin\mu\cos\mu & \alpha\sin^2\mu + \frac{1}{\alpha}\cos^2\mu \end{bmatrix} \quad (5) \end{aligned}$$

Deformation is equivalent to a pure shear which preserves area, i.e. an expansion by a factor, α, in the direction, μ, 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 the affine transformation on the orientation of image detail.

Consider an element of texture represented by an unit vector, $\mathbf{v}$, with orientation, φ.

$$\mathbf{v} = [\cos\varphi, \sin\varphi]$$

The affine transformation, A, transforms the vector, $\mathbf{v}$, into $\mathbf{v}'$ and changes its orientation by $\Delta\varphi$, so that the transformed vector, $\mathbf{v}'$, has new orientation $\varphi' = \varphi + \Delta\varphi$.
The change in orientation, $\Delta\varphi$, caused by the affine transformation can be computed from the vector product ($\wedge$) of $\mathbf{v}$ and $\mathbf{v}'$ (note that in general the transformed vector $\mathbf{v}'$ will no longer be a unit vector).

$$\sin(\Delta\varphi) = \frac{|\mathbf{v} \wedge \mathbf{v}'|}{|\mathbf{v}||\mathbf{v}'|} \quad (6)$$

Let us consider the effects of each component of the affine transformation on orientation. Let the deformation, D, change $\mathbf{v}$ into $\mathbf{v}''$ such that:

$$\mathbf{v}'' = D(\alpha,\mu)\mathbf{v} \quad (7)$$

Then this vector, $\mathbf{v}''$, is transformed by S and R to $\mathbf{v}'$ such that:

$$\mathbf{v}' = S(s)R(\theta)\mathbf{v}'' \quad (8)$$

Substituting (8) into (6):

$$\begin{aligned} \sin(\Delta\varphi) &= \frac{(\mathbf{v}\cdot\mathbf{v}'')}{|\mathbf{v}||\mathbf{v}''|}\sin\theta + \frac{|\mathbf{v}\wedge\mathbf{v}''|}{|\mathbf{v}||\mathbf{v}''|}\cos\theta \\ &= \cos(\Delta\varphi_d)\sin\theta + \sin(\Delta\varphi_d)\cos\theta \\ &= \sin(\theta + \Delta\varphi_d) \quad (9) \end{aligned}$$

where $(\cdot)$ means inner product. Equation (9) trivially shows us that the change in orientation, $\Delta\varphi$, can be written as the sum of two components: one due to rotation, θ, and one due to deformation, $\Delta\varphi_d$.

$$\Delta\varphi = \theta + \Delta\varphi_d \quad (10)$$

Note that the isotropic scale, S, does not affect orientation, while the change in orientation due to the deformation term (described below) depends on the initial orientation, φ, the axis of deformation, μ, and magnitude of deformation, α.

$$\begin{aligned}
\sin(\Delta\varphi_d) &= \frac{|\mathbf{v} \wedge \mathbf{v}''|}{|\mathbf{v}||\mathbf{v}''|} \\
&= \frac{\frac{1}{2}\left(\frac{1}{\alpha} - \alpha\right)\sin 2(\varphi - \mu)}{\sqrt{\frac{1}{\alpha^2}\sin^2(\varphi-\mu) + \alpha^2\cos^2(\varphi - \mu)}} \qquad (11)
\end{aligned}$$

If the magnitude of deformation is small, the change in orientation can be described as follows from (10) and(11).

$$\Delta\varphi \simeq \theta + \frac{1}{2}\left(\frac{1}{\alpha} - \alpha\right)\sin 2(\varphi - \mu) \qquad (12)$$

If we consider the change in the relative orientation between two oriented elements, $\Delta\varphi_1 - \Delta\varphi_2$:

$$\Delta(\varphi_1 - \varphi_2) \simeq \left(\frac{1}{\alpha} - \alpha\right)\sin(\varphi_1 - \varphi_2)\cos(\varphi_1 + \varphi_2 - 2\mu) \qquad (13)$$

We notice that rotation, $R(\theta)$, changes the orientation of all the elements equally, and does not affect relative orientation between two elements. Relative orientation (13) are only affected by deformation. Koenderink and Van Doorn [20] derived a similar approximate equation for small displacements.

2.3 Change in Area and Density

We will now consider the effect on the area and density of a texture under an affine transformation.

By definition, area will only be affected by the scale, $S(s)$. Consider an area, a, described by two vectors, $\mathbf{v_1}, \mathbf{v_2}$, which is transformed into area, a', composed by transformed vectors, $\mathbf{v'_1}, \mathbf{v'_2}$, by the affine transformation. Then, the transformed area, a', computed by vector product of $\mathbf{v'_1}$ and $\mathbf{v'_2}$ can be described using the original area, a, as follows:

$$\begin{aligned}
a' &= |\mathbf{v'_1} \wedge \mathbf{v'_2}| && (14)\\
&= s^2\left(\sin^2\mu + \cos^2\mu\right)^2(\cos\varphi_1\sin\varphi_2 - \sin\varphi_1\cos\varphi_2) && (15)\\
&= s^2|\mathbf{v_1} \wedge \mathbf{v_2}| && (16)\\
&= s^2 a && (17)
\end{aligned}$$

where, φ_1, φ_2 are the orientation of the vector $\mathbf{v_1}$ and $\mathbf{v_2}$ respectively. Then, the scaling factor, s, is:

$$s^2 = \frac{a'}{a} \qquad (18)$$

For homogeneous texture, the density of texture elements will be scaled inversely by the same amount:

$$s^2 = \frac{\rho}{\rho'} \qquad (19)$$

where, ρ and ρ' are the densities of texture elements in the original and transformed image respectively.

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 (i.e. corners, edges or closed contours) 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 kind of texture 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 (12), 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 $f(\varphi)$ (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 $f(\varphi)$ (the second moment). Furthermore, because there is a term μ 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 $f(\varphi)$ (the third moment) will be affected. Fig.1 shows the changes in distribution due to the deformation, and we can find out the effect of the deformation mentioned above. 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. As above scale, S, affects the area of texture elements and their density, leaving orientations unaffected.

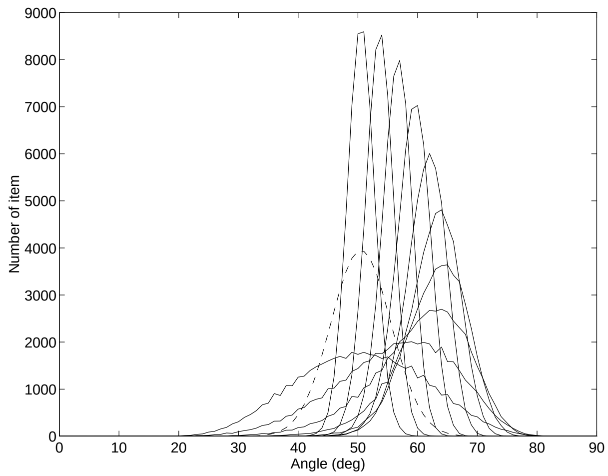


Figure 1: Distribution of the orientation of the original and deformed texture elements. Original symmetric distribution (dashed line) is changed to asymmetric distribution (solid line) by deformation whose axis is not 0 or 90 degree. The axes of deformation are 0, 10, 20, ..., 90 (deg) from the lowest distribution respectively.

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. Better still this can be done without making explicit point, edge or contour correspondences. The derivation follows.

From (12), an element with orientation φ is transformed into φ' such that:

$$\begin{aligned} \varphi' &= \varphi + \theta + \lambda \sin 2(\varphi - \mu) \\ &= \varphi + \theta + \lambda(\sin 2\varphi \cos 2\mu - \cos 2\varphi \sin 2\mu) \end{aligned} \tag{20}$$

where, λ is related to the magnitude term of the deformation by:

$$\lambda = \frac{1}{2}\left(\frac{1}{\alpha} - \alpha\right) \tag{21}$$

The change, ΔI_φ, in the first moment of $f(\varphi)$, in terms of the rotation, θ, axis, μ, and magnitude, λ, of the deformation is given by summing equation (20) for all elements.

$$\Delta I_\varphi = \lambda(I_{\sin 2\varphi}\cos 2\mu - I_{\cos 2\varphi}\sin 2\mu) + \theta \tag{22}$$

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 (Appendix A). They are given by:

$$\Delta I_{\varphi\varphi} = 2\lambda(I_{\varphi \sin 2\varphi}\cos 2\mu - I_{\varphi\cos 2\varphi}\sin 2\mu) \tag{23}$$

$$\Delta I_{\varphi\varphi\varphi} = 3\lambda(I_{\varphi^2 \sin 2\varphi}\cos 2\mu - I_{\varphi^2\cos 2\varphi}\sin 2\mu) \tag{24}$$

where $I_{\varphi \sin 2\varphi}$ and $I_{\varphi \cos 2\varphi}$ are the covariances between φ and $\sin 2\varphi$, and φ and $\cos 2\varphi$ respectively. $I_{\varphi^2 \sin 2\varphi}$ and $I_{\varphi^2 \cos 2\varphi}$ are third moments. Note that the first moment is related to the deformation and the rotation. The second and third moments are related to the deformation only, and the change in scale does not affect any of the moments of orientation.

The rotation, θ, the axis of deformation, μ, and the magnitude of deformation, α, can be computed from:

$$\mu = \frac{1}{2}\tan^{-1}\left(\frac{M_1}{M_2}\right) \tag{25}$$

$$\alpha = \frac{1}{2M_3}\left(\sqrt{M_1^2 + M_2^2 + 4M_3^2} - \sqrt{M_1^2 + M_2^2}\right) \tag{26}$$

$$\theta = \Delta I_\varphi - \frac{1}{2M_3}\left(I_{\sin 2\varphi}M_2 - I_{\cos 2\varphi}M_1\right) \tag{27}$$

where:

$$M_1 = 3\Delta I_{\varphi\varphi}I_{\varphi^2 \sin 2\varphi} - 2\Delta I_{\varphi\varphi\varphi}I_{\varphi \sin 2\varphi} \tag{28}$$

$$M_2 = 3\Delta I_{\varphi\varphi}I_{\varphi^2 \cos 2\varphi} - 2\Delta I_{\varphi\varphi\varphi}I_{\varphi \cos 2\varphi} \tag{29}$$

$$M_3 = 3\left(I_{\varphi^2 \cos 2\varphi}I_{\varphi \sin 2\varphi} - I_{\varphi^2 \sin 2\varphi}I_{\varphi \cos 2\varphi}\right) \tag{30}$$

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 using (18) and (19). Having computed the rotation, θ, the axis of deformation, μ, the magnitude of deformation, α, 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. It 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. 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 [21], this assumption is no longer required.

For efficiency we have chosen to use a combination of linear and circular moments. In fact linear moments suffer from an aliasing problem, i.e. linear moments have singular points at angles of 0 and π. This is avoided by using circular moments, generating a symmetric distribution by representing each orientation twice (φ, and $\varphi + \pi$) or shifting the origin of the distribution.

4 Implementation

4.1 Detection of Local Direction

The proposed method first detects the edge points and their orientations [7] (Samples should in principle be random though this was not implemented for simplification). The accuracy of the orientation estimates is important and a typical corner finder — in fact the Noble's corner detector [26] — was used to reject high curvature points and corners. This is because the orientations at these points cannot be accurately determined due to lack of local image support [12].

4.2 Calculation of The Affine Transformation

Once the orientations of edges in both the original and transformed images have been estimated the directional moments of equation (22) - (24) can be computed. Equation (25) - (27) allow the estimation of the rotation and deformation components of the affine transformation.

The change in the scale can be obtained from the ratio of the numerical density of the texture or the ratio of the area of the subjects [8] between the original image and the deformation image. In this preliminary implementation we have used a simple estimation method using the edge density. If the deformation is small and the sampling is random, the ratio of the number of the edge points is equal to the ratio of the numerical density of the texture. Unfortunately this requires spatial homogeneity. An alternative method to compute the scale is presented by Lindeberg and Garding [21].

5 Experiments

5.1 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.2 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 (left most images in Fig.2) 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 column of images in Fig.2). The two ellipses in Fig.2 show that the estimated orientations are qualitatively good even with non-uniform textures. Table 3.1 compares the accuracy of this method quantitatively for each sample image with the known fiducial orientation. 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 problem of linear moments.

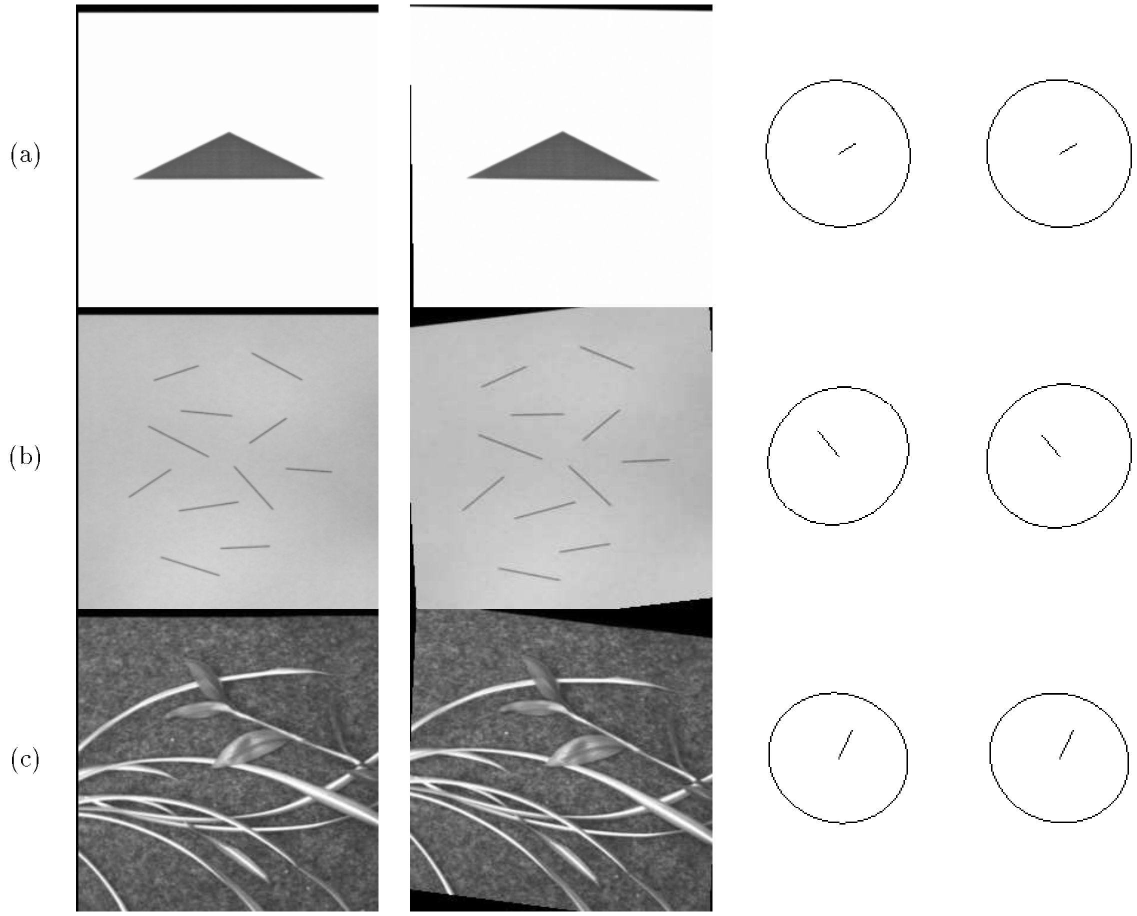
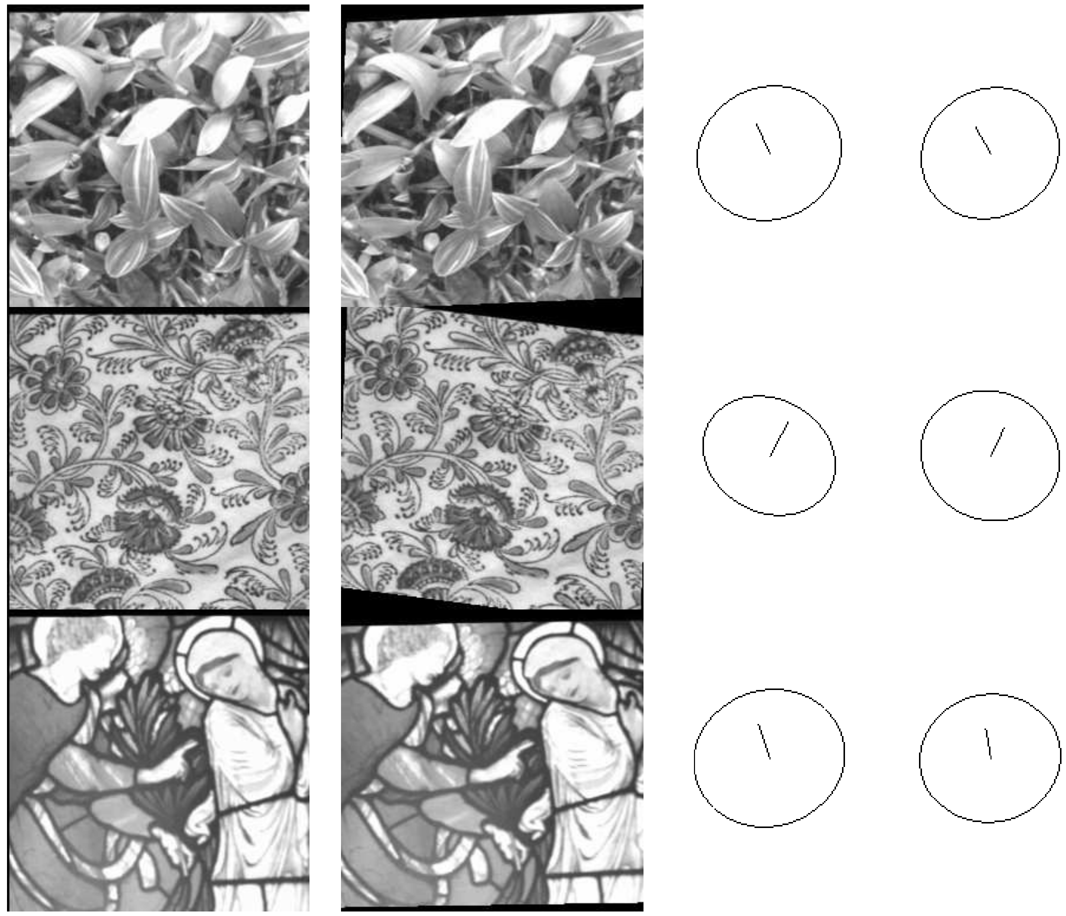


Figure 2: Results of preliminary experiments. Examples of the images distorted by arbitrary affine transformations were processed by our affine transform from texture moments algorithm. Images in the first and second column are fronto-parallel and transformed images respectively after changing the position and orientation of the plane viewed under *weak* perspective. The estimated orientations (left ellipses) and true orientations (right 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.

(d)

(e)

(f)



(See previous page for the caption.)

Table 3.1 Accuracy of the surface parameters, scale, s, rotation, θ, tilt, τ, and slant, σ.

	True				Estimated			
Images	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

5.2 Exploiting the Affine Transformation

A second method for testing the accuracy of this method is to apply the results to real applications which can exploit this visually derived information. We consider two examples: a simple application of shape from texture [23] and then an example in which the divergence, curl and deformation components are used in qualitative visual navigation [8].

5.2.1 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 [23].

As a simple example consider the affine transform relating the distortion between a patch which is fronto-parallel and another which has an orientation specified by tilt, τ, and slant, σ. The orientation can be recovered directly from the deformation component of the affine transform:

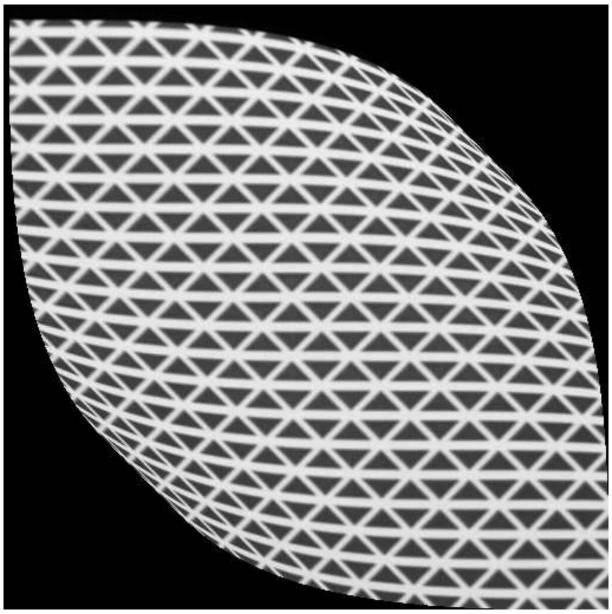
$$\tau = \mu + \frac{\pi}{2} \tag{31}$$

$$\sigma = \cos^{-1}\left(\frac{1}{\alpha^2}\right) \tag{32}$$

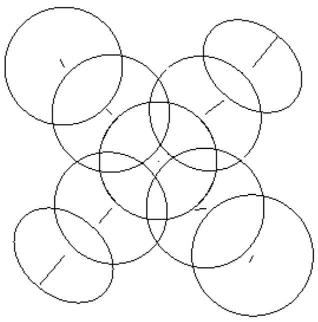
A similar relationship can be derived between two image patches with arbitrary orientations [10].

Fig.3 shows the result of using affine transforms by the method presented in this paper to recover the shape of a curved surface. Fig.3 (b) shows the estimated surface orientation of (a). The proposed method derives qualitatively good results, though there are some errors in the estimated orientations. These errors arise from:

1. the difference of the sampling areas;
2. errors caused in the sampling of orientation;
3. small deformation approximation used in the proposed method.



(a)



(b)

Figure 3: 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.

5.2.2 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 are given by [18, 8]:

$$\begin{aligned} \mathrm{curl}\mathbf{v} &= -2\mathbf{\Omega}\cdot\mathbf{Q} + |\mathbf{F}\wedge\mathbf{A}| & (33)\\ \mathrm{div}\mathbf{v} &= \frac{2\mathbf{U}\cdot\mathbf{Q}}{\lambda} + \mathbf{F}\cdot\mathbf{A} & (34)\\ \mathrm{def}\mathbf{v} &= |\mathbf{F}||\mathbf{A}| & (35)\end{aligned}$$

where, curl $\mathbf{v}$, div $\mathbf{v}$ and def $\mathbf{v}$ are the curl (rotation), divergence (scale) and deformation components of the image velocity field respectively. $\mathbf{U}$ is the translational velocity, $\mathbf{\Omega}$ is the rotational velocity, $\mathbf{Q}$ is the viewing direction, $\mathbf{A}$ is the component of the translational velocity parallel to the image plane scaled by distance and $\mathbf{F}$ represents the surface orientation as a depth gradient. The axis of deformation, μ, bisects $\mathbf{A}$ and $\mathbf{F}$,

$$\mu = \frac{\angle\mathbf{A} + \angle\mathbf{F}}{2} \tag{36}$$

The geometric significance of these equations is easily seen with a few examples. For example a translation towards the surface patch leads to a uniform expansion in the image encoded as a change in the scale. This determines the distance to the object which due to the speed–scale ambiguity is more conveniently expressed as a time to contact, t_c:

$$t_c = \frac{\lambda}{\mathbf{U}\cdot\mathbf{Q}} \tag{37}$$

A translational motion perpendicular to the visual direction and parallel to the horizontal axis of the image plane results in image deformation with a magnitude determined by the slant of the surface and an axis determine by the tilt. In this case the axis of deformation bisects the tilt angle and direction of translation in the image. Changes in scale and rotation will also occur.

Fig.4 and 5 shows the images taken from the moving observer, motion and optical direction are parallel in Fig.4 and perpendicular in 5. Table 3.2 shows the results of estimation of tilt angle of the surface and time to contact computed from Fig.4 and 5. Good accuracy of estimated divergence and magnitude of deformation exploit the good result of time to contact. The systematic error shown in the result of tilt angle was caused by the computation error in the axis of deformation which is related to the error in rotation.

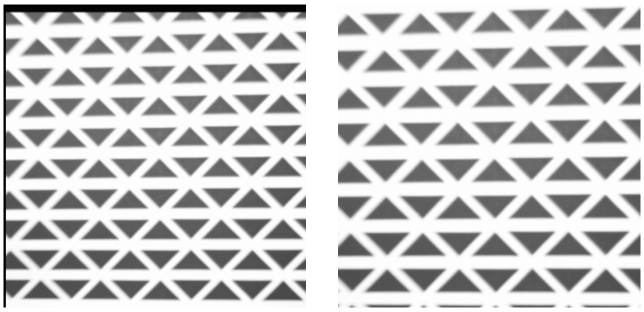


Figure 4: Two images used in the experiments for visual navigation.
Two images are taken from a moving observer. The observer moves towards the object along the optical axis. The time to contact, estimated from the change in scale, is shown in table 3.2.

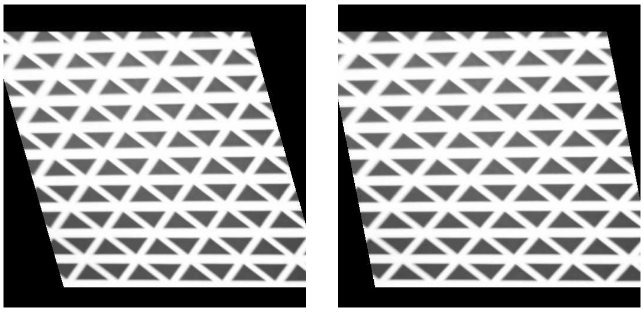


Figure 5: Two images used in the experiments for visual navigation.
Two images are taken from a moving observer. The observer moves in the direction perpendicular to the optical axis from left to right. The axis of deformation recovered from the texture moments can be used to compute the tilt angle of the surface (shown in table 3.2).

Table 3.2 Tilt angle and time to contact. Scale, s, rotation, θ, axis and magnitude of the deformation, μ, α, recovered from the texture moments were used to compute the time to contact, t_c, in (a) and tilt angle of the surface, τ, in (b). Mean values and variances were estimated by changing the area of interest in each image.

Images		s	$\theta(°)$	α	$\mu(°)$	$\tau(°)$	t_c
	True	1.20	0.0	1.0	-	-	10.0
(a) **parallel**	Estimated	1.19 ± 0.07	0.1 ± 0.1	1.01 ± 0.01	-	-	10.4 ± 3.2
perpen-	True	1.04	3.2	1.07	28	56	∞
(b) **dicular**	Estimated	1.03 ± 0.01	1.3 ± 0.1	1.07 ± 0.001	33 ± 2	66 ± 4	882 ± 496

5.3 Theoretical Comparison with Other Methods

In this section we compare the theoretical error and sensitivity to noise of this method with those of the second moment method proposed by Kanatani [15].

5.3.1 Review of Kanatani's Method

It has long been known as the Buffon transform from a stereological point of view, that the probability function, $f(\theta)$, of the orientation of the image features, θ, can be described using the expected number of intersections, $N(\varphi)$, of sampling lines whose orientation is φ with image features as follows:

$$N(\varphi) = \int_0^{2\pi} |\sin(\varphi - \theta)| f(\theta) d\theta \tag{38}$$

Kanatani showed the inverse formula of this Buffon transform using Fourier series.

$$\begin{aligned}
f(\theta) &= \frac{\frac{1}{4}C}{2\pi}\left[1 - \sum_{n=2}^{\infty}(n^2-1)(A_n \cos n\theta + B_n \sin n\theta)\right] && (39)\\
C &= \int_0^{2\pi} N(\varphi) d\varphi && (40)\\
A_n &= \frac{2}{C}\int_0^{2\pi} N(\varphi)\cos n\varphi d\varphi && (41)\\
B_n &= \frac{2}{C}\int_0^{2\pi} N(\varphi)\sin n\varphi d\varphi && (42)
\end{aligned}$$

Comparing each coefficient of this inverse Buffon transform with Fourier coefficients of the distribution of orientation of the image features, he derived the next relations between slant, σ, tile angle, τ, of the surface and the second Fourier coefficients, A_2, B_2, of the inverse Buffon transform as follows:

$$\begin{aligned}
A &= (\frac{2}{3} - (B_2)^2)(A_2' - A_2) + A_2 B_2 (B_2' - B_2) && (43)\\
B &= A_2 B_2 (A_2' - A_2) + (\frac{2}{3} - (A_2)^2)(B_2' - B_2) && (44)\\
\sigma &= 2(A^2 + B^2)^{\frac{1}{4}} / |\frac{2}{3} - (A_2)^2 - (B_2)^2|^{\frac{1}{2}} && (45)\\
\tau &= \frac{1}{2}\tan^{-1}\frac{B}{A} \qquad (if A < 0) && (46)\\
& \quad \frac{1}{2}\tan^{-1}\frac{B}{A} + \frac{1}{2}\pi \qquad (if A > 0) && (47)\\
A_2 &= 2\sum_{k=0}^{m-1} N_k \cos(\frac{2\pi k}{m}) / \sum_{k=0}^{m-1} N_k && (48)\\
B_2 &= 2\sum_{k=0}^{m-1} N_k \sin(\frac{2\pi k}{m}) / \sum_{k=0}^{m-1} N_k && (49)\\
& && (50)
\end{aligned}$$

where, N_k is the number of intersections in the kth sampling of total m samplings. A_2 and B_2 are the second Fourier coefficients of the inverse Buffon transform in the original image, and A_2 and B_2 are those in the deformed image respectively.

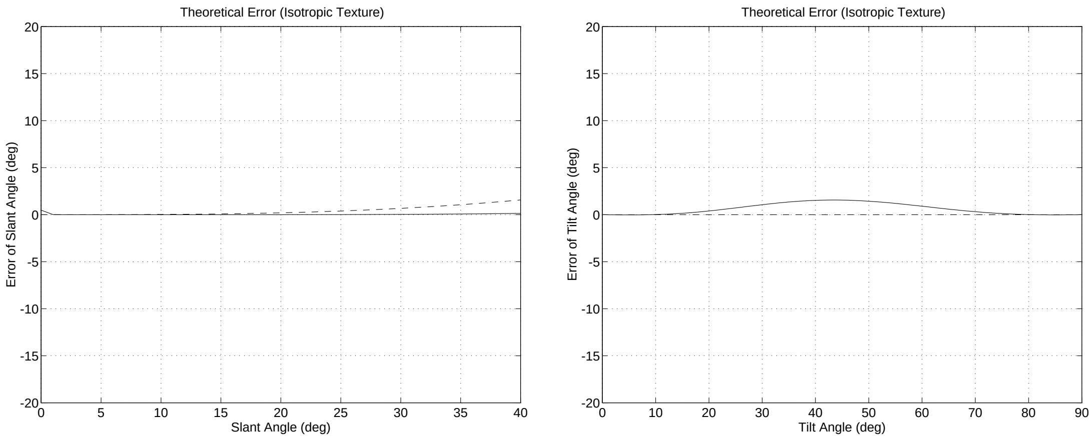
5.3.2 Comparison of the Theoretical Error

Both methods use approximations to solve the problem in closed form, and these approximations cause theoretical errors in the estimated orientation of the surface. Fig.6 and Fig.7 shows comparison of the error caused by approximations of proposed method with that of Kanatani's method.

The proposed method has good accuracy in slant and tilt angles in both isotropic and anisotropic textures, although accuracy of Kanatani's method degrades rapidly with slant and tilt angles, especially in anisotropic texture.

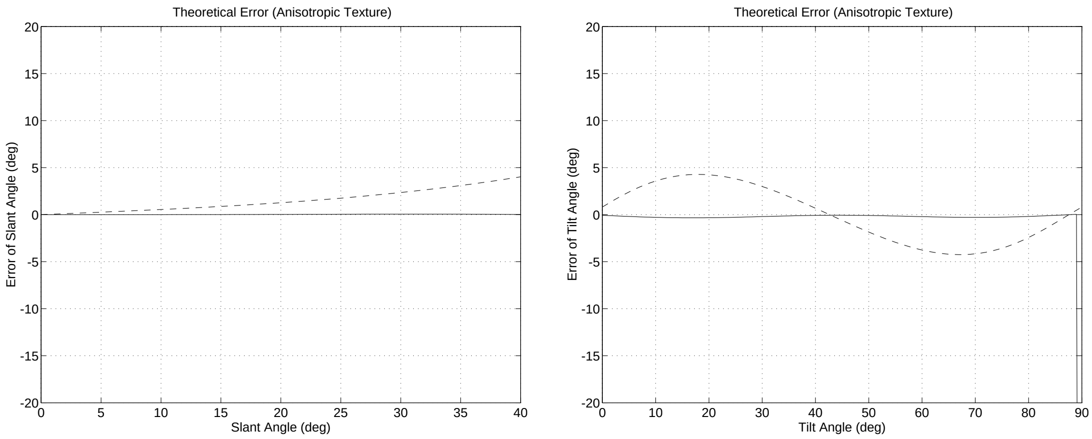
5.3.3 Comparison of the Noise Sensitivity

Noise sensitivities of these two methods are shown in Fig.8 and Fig.9. The random gaussian noises whose standard deviations are 0.5 degree are added to the orientation data. The slant estimated by the proposed method is less affected by noise than that of Kanatani's method in isotropic textures, although both methods are weak to noise in the case of small slant, especially in anisotropic textures. The tilt angle estimated by the proposed method is less affected by noise than that of Kanatani's method especially in anisotropic textures. Generally, both methods are more sensitive to noise in anisotropic textures than in isotropic textures. This is because the noise makes it difficult to distinguish the axis of deformation and rotation, if the texture pattern has less variety in orientation.



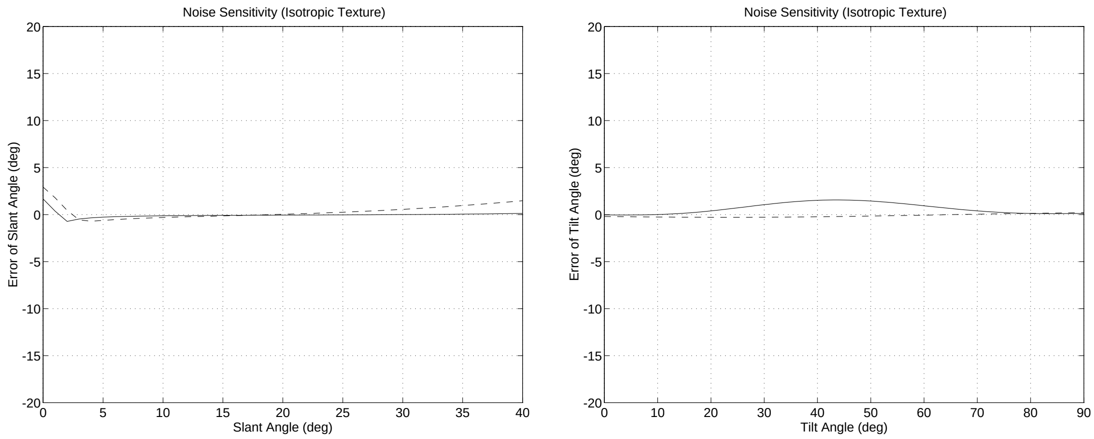
(a) (b)

Figure 6: Results of the theoretical error analysis in the case of isotropic texture. (a) and (b) show errors in slant and tilt angle of the surface respectively. The solid line shows the error of the proposed method, and dashed line shows that of Kanatani's method.



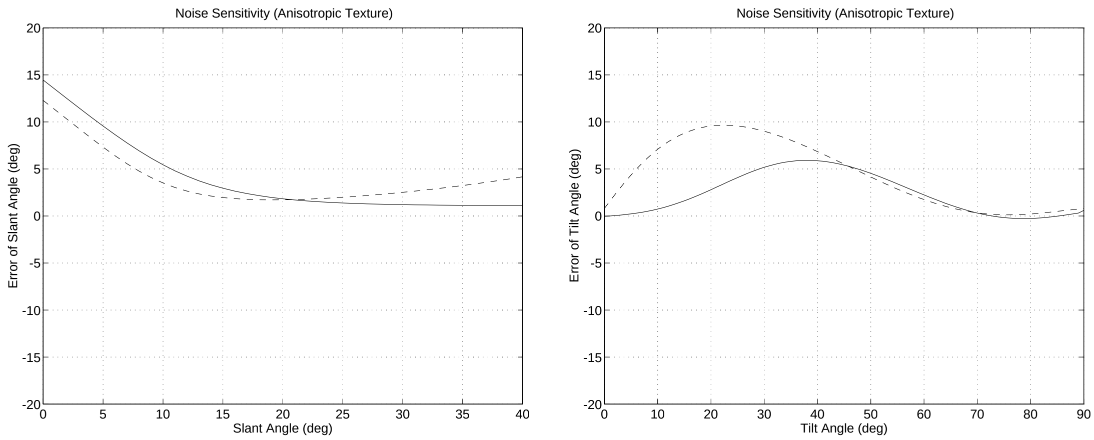
(a) (b)

Figure 7: Results of the theoretical error analysis in the case of anisotropic texture. The anisotropic texture used in this experiment has gaussian distribution whose mean value is 0 degrees and standard deviation is 20 degrees. (a) and (b) show errors in slant and tilt angle of the surface respectively. The solid line shows the error of the proposed method, and dashed line shows that of Kanatani's method.



(a) (b)

Figure 8: Results of the noise sensitivity analysis in the case of isotropic texture. (a) and (b) show errors in slant and tilt angle of the surface respectively. These errors were caused by the gaussian noise whose standard deviation is 0.5 degrees. The solid line shows the error of the proposed method, and dashed line shows that of Kanatani's method.



(a) (b)

Figure 9: Results of the noise sensitivity analysis in the case of anisotropic texture. The anisotropic orientation data used in this experiment has gaussian distribution whose mean value is 0 degrees and standard deviation is 20 degrees. (a) and (b) show errors in slant and tilt angle of the surface respectively. The solid line shows the error of the proposed method, and dashed line shows that of Kanatani's method. These errors were caused by the gaussian noise whose standard deviation is 0.5 degrees.

6 Conclusions

In this paper we proposed a novel, efficient and geometrically intuitive method to compute the four components of an affine transformation (i.e. rotation, axis of deformation, magnitude of deformation and scale) from the changes in first, second and third moments of edge orientation and changes in density. This method does not require point, edge or contour correspondences to be established. 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. The remaining problems of this method are:

- **Selection of the Area of Interest:** In this method it was assumed that the sampling area of the original image and that of the transformed image correspond to the same region of texture. In the above implementation we have simply sampled equal areas in both images. In practice the affine transform will distort the position, size and shape of the sample and thus an error in the computed affine transformation will result for textures which are not homogeneous.

  It is of course possible to make these two sampling areas equivalent by using the estimate of the affine transform to define the new region and improve the accuracy of the estimates, using iterative schemes.

- **Aliasing Problem:** Generally linear moments suffer from an aliasing problem, i.e. linear moments have singular points at angles of 0 and π. This is avoided by using circular moments, generating a symmetric distribution by representing each orientation twice (φ, and $\varphi + \pi$) or shifting the origin of the distribution.

Appendix A

As shown in equation (12), the absolute orientation of the transformed vector φ' is:

$$\begin{aligned}\varphi' &= \varphi + \theta + \lambda \sin 2(\varphi - \mu) \\ &= \varphi + \theta + \lambda(\sin 2\varphi \cos 2\mu - cos2\varphi \sin 2\mu)\end{aligned} \tag{51}$$

where,

$$\lambda = \frac{1}{2}\left(\frac{1}{\alpha} - \alpha\right) \tag{52}$$

Summing equation (51) for all elements in the sample:

$$\begin{aligned}I_{\varphi'} &= \frac{1}{N}\sum \varphi' \\ &= I_\varphi + \lambda(I_{\sin 2\varphi}\cos 2\mu - I_{\cos 2\varphi}\sin 2\mu) + \theta\end{aligned} \tag{53}$$

where, I_φ and $I_{\varphi'}$ are the mean value (first moment) of the original orientation, φ, and the deformed orientation, φ', respectively. $I_{\sin 2\varphi}$ and $I_{\cos 2\varphi}$ are the mean values of $\sin 2\varphi$ and $\cos 2\varphi$. Taking the difference, ΔI_φ, between $I_{\varphi'}$ and I_φ:

$$\Delta I_\varphi = \lambda(I_{\sin 2\varphi}\cos 2\mu - I_{\cos 2\varphi}\sin 2\mu) + \theta \tag{54}$$

The second moment, $I_{\varphi'\varphi'}$, of the deformed orientations is:

$$\begin{aligned}I_{\varphi'\varphi'} &= \frac{1}{N}\sum(\varphi' - I_{\varphi'})^2 \\ &= I_{\varphi\varphi} + 2\lambda(\cos 2\mu I_{\varphi \sin 2\varphi} - \sin 2\mu I_{\varphi\cos 2\varphi}) \\ &\quad +\lambda^2(\cos^2 2\mu I_{\sin^2 2\varphi} - 2\sin 2\mu\cos 2\mu I_{\sin 2\varphi\cos 2\varphi} + \sin^2 2\mu I_{\cos^2 2\varphi})\end{aligned} \tag{55}$$

where $I_{\varphi\varphi}$ is the variance of original orientations, $I_{\varphi\sin 2\varphi}$ and $I_{\varphi\cos 2\varphi}$ are covariances between φ and $\sin 2\varphi$, and φ and $\cos 2\varphi$ respectively, $I_{\sin^2 2\varphi}$ and $I_{\cos^2 2\varphi}$ are variances of $\sin 2\varphi$ and $\cos 2\varphi$ respectively, $I_{\sin 2\varphi\cos 2\varphi}$ is a covariance between $\sin 2\varphi$ and $\cos 2\varphi$. If the deformation is small, $\alpha \simeq 1$ and $\lambda \ll 1$. Then, to a first order approximation, the second moment of the deformed orientations is given by:

$$I_{\varphi'\varphi'} = I_{\varphi\varphi} + 2\lambda(\cos 2\mu I_{\varphi\sin 2\varphi} - \sin 2\mu I_{\varphi\cos 2\varphi}) \tag{56}$$

Then, the change in the second moment, $\Delta I_{\varphi\varphi}$, between $I_{\varphi'\varphi'}$ and $I_{\varphi\varphi}$ is,

$$\Delta I_{\varphi\varphi} = 2\lambda(I_{\varphi\sin 2\varphi}\cos 2\mu - I_{\varphi\cos 2\varphi}\sin 2\mu) \tag{57}$$

The third moment, $I_{\varphi'\varphi'\varphi'}$, of the deformed orientations is (with subscripts explained as above):

$$\begin{aligned}I_{\varphi'\varphi'\varphi'} &= \frac{1}{NI^3_{\varphi'\varphi'}}\sum(\varphi' - I_{\varphi'})^3 \\ &= I_{\varphi\varphi\varphi} + 3\lambda(\cos 2\mu I_{\varphi^2\sin 2\varphi} - \sin 2\mu I_{\varphi^2\cos 2\varphi}) \\ &\quad +3\lambda^2(\cos^2 2\mu I_{\varphi\sin^2 2\varphi} - 2\sin 2\mu\cos 2\mu I_{\varphi\sin 2\varphi\cos 2\varphi} + \sin^2 2\mu I_{\varphi\cos^2 2\varphi}) \\ &\quad +\lambda^3(\cos^3 2\mu I_{\sin^3 2\varphi} - 3\cos^2 2\mu\sin 2\mu I_{\sin^2 2\varphi\cos 2\varphi} \\ &\quad +3\sin^2 2\mu\cos 2\mu I_{\cos^2 2\varphi\sin 2\varphi} - \sin^3 2\mu I_{\cos^3 2\varphi})\end{aligned} \tag{58}$$

For small deformations:

$$I_{\varphi'\varphi'\varphi'} = I_{\varphi\varphi\varphi} + 3\lambda(\cos 2\mu I_{\varphi^2\sin 2\varphi} - \sin 2\mu I_{\varphi^2\cos 2\varphi}) \tag{59}$$

and a change in the third moment, $\Delta I_{\varphi\varphi\varphi}$, between $I_{\varphi'\varphi'\varphi'}$ and $I_{\varphi\varphi\varphi}$ is given by:

$$\Delta I_{\varphi\varphi\varphi} = 3\lambda(I_{\varphi^2\sin 2\varphi}\cos 2\mu - I_{\varphi^2\cos 2\varphi}\sin 2\mu) \tag{60}$$

References

[1] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *Int. Journal of Computer Vision*, pages 283–310, 1989.

[2] R. Bajcsy and L. Lieberman. Texture gradient as a depth cue. In *CGIP, 5*, pages 52–67, 1976.

[3] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In G. Sandini, editor, *Proc. 2nd European Conference on Computer Vision*, pages 237–252. Springer–Verlag, 1992.

[4] A. Blake and C. Marinos. Shape from texture: estimation, isotropy and moments. *Artificial Intelligence*, 45:323–380, 1990.

[5] M. Brady and H. Wang. Vision for mobile robots. *Philosophical Transactions of the Royal Society of London, Ser ies B*, 337:341–350, 1992.

[6] L.G. Brown and H. Shyvaster. Surface orientation from projective foreshortening of isotropic texture autocorrelation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-12(6):584–588, 1990.

[7] J.F. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-8:679–698, 1986.

[8] 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.

[9] L. Dreschler and H.H. Nagel. Volumetric model and 3 trajectory of a moving car derived from monocular TV-frame sequence of a street scene. In *IJCAI*, pages 692–697, 1981.

[10] J. Garding and T. Lindeberg. Direct computation of shape cues by multi-scale retinotopic processing. *Technical Report, Royal Institute of Technology, TRITA-NA-P9304*, 1993.

[11] J.J. Gibson. *The Perception of the Visual World.* Houghton Mifflin, 1950.

[12] C. Harris. Geometry from visual motion. In A. Blake and A. Yuille, editors, *Active Vision.* MIT Press, Cambridge, USA, 1992.

[13] D.G. Jones and J. Malik. Determining three-dimensional shape from orientation and spatial frequency disparities. In G. Sandini, editor, *Proc. 2nd European Conference on Computer Vision*, pages 661–669. Springer–Verlag, 1992.

[14] 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.

[15] K. Kanatani. Detection of surface orientation and motion from texture by a stereological technique. *Artificial Intelligence*, 23:213–237, 1984.

[16] K. Kanatani. Structure and motion from optical flow under orthographic projection. *Computer Vision, Graphics and Image Processing*, 35:181–199, 1986.

[17] L. Kitchen and A. Rosenfeld. Grey-level corner detection. *Pattern Recognition Letters*, 1:95–102, 1982.

[18] J.J. Koenderink. Optic flow. *Vision Research*, 26(1):161–179, 1986.

[19] J.J. Koenderink and A.J. van Doorn. Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22(9):773–791, 1975.

[20] J.J. Koenderink and A.J. van Doorn. Geometry of binocular vision and a model for stereopsis. *Biological Cybernetics*, 21:29–35, 1976.

[21] T. Lindeberg and J. Garding. Shape from texture from a multi-scale perspective. *Proc. 4th Int. Conf. on Computer Vision*, pages 683–691, 1993.

[22] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. of the 7th International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.

[23] 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.

[24] D. Murray and B. Buxton. *Experiments in the machine interpretation of visual motion.* MIT Press, Cambridge, USA, 1990.

[25] H.H. Nagel. Principles of (low level) computer vision. In J.P. Haton, editor, *Fundamentals in computer understanding: speech and vision*, pages 113–139. Cambridge University Press, 1987.

[26] J.A. Noble. Finding Corners. *Image and Vision Computing*, 6(2):121–128, May 1988.

[27] I.D. Reid and D.W. Murray. Tracking foveated corner clusters using affine structure. In *Proc. 4th Int. Conf. on Computer Vision*, pages 76–83, 1993.

[28] H.S. Sawhney and A.R. Hanson. Identification and 3d description of 'shallow' environmental structure in a sequence of images. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 179–186, 1991.

[29] B.J. Super and A.C. Bovik. Solution to shape-from-texture by wavelet-based measurement of local spectral moments. In *Proc. Conf. Computer Vision and Pattern Recognition*, 1992.

[30] A.M. Waxman and K. Wohn. Contour evolution, neighbourhood deformation and global image flow: planar surfaces in motion. *Int. Journal of Robotics Research*, 4(3):95–108, 1985.

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