University of Pennsylvania

ScholarlyCommons

Technical Reports (CIS)

Department of Computer & Information Science

August 1988

On the Deformation of Image Intensity and Zero-Crossing Contours Under Motion

Jian Wu
Boston University

K. Wohn
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/cis_reports

On the Deformation of Image Intensity and Zero-Crossing Contours Under Motion

Abstract

Image intensity and edge are two major sources of information for estimating the motion in the image plane. The 2-D motion obtained by analyzing the deformation of intensity and/or edges is used to recover the 3-D motion and structure. In this paper we show that the motion defined by the image intensity differs from the motion revealed by the (zerocrossing) edge. Understanding of this discrepancy is important since most of the 3-D motion recovery algorithms reported so far require accurate 2-D motion as their input.

We begin the discussion by assuming the *invariance of intensity*, that the evolution of image intensity manifests the underlying transformation of the image due solely to the motion of objects. We then raise the question if the zerocrossing of the Laplacian operating on the image intensity is invariant too. The change of perspective view due to relative motion results the zerocrossing not being preserved as the image evolves, thereby deteriorating the accuracy of the 2-D motion obtained from the zerocrossing contour. We derive how much the zero-crossing contour deviates from its "correct" position due to motion. The result may be used to determine the reliability of the zerocrossing contours for the purpose of the motion estimation.

Comments

ON THE DEFORMATION OF
IMAGE INTENSITY AND
ZERO-CROSSING CONTOURS
UNDER MOTION

Jian Wu and K. Wohn

MS-CIS-88-69
GRASP LAB 154

Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104

September 1988

On the Deformation of Image Intensity and Zero-Crossing

Contours Under Motion

Jian Wu

Laboratory for Sensory Robotics

Boston University

K. Wohn

Department of Computer and Information Science

University of Pennsyvania

Abstract

Image intensity and edge are two major sources of information for estimating the motion in the image plane. The 2-D motion obtained by analyzing the deformation of intensity and/or edges is used to recover the 3-D motion and structure. In this paper we show that the motion defined by the image intensity differs from the motion revealed by the (zerocrossing) edge. Understanding of this discrepancy is important since most of the 3-D motion recovery algorithms reported so far require accuarate 2-D motion as their input.

We begin the discussion by assuming *the invariance of intensity*, that the evolution of image intensity manifests the underlying transformation of the image due solely to the motion of objects. We then raise the question if the zerocrossing of the Laplacian operating on the image intensity is invariant too. The change of perspective view due to relative motion results the zerocrossing not being preserved as the image evolves, thereby deteriorating the accuracy of the 2-D motion obtained from the zerocrossing contour. We derive how much the zero-crossing contour deviates from its "correct" position due to motion. The result may be used to determine the reliability of the zero-crossing contours for the purpose of the motion estimation.

1   Introduction

The Analysis on the time-varying imagery often involves the problem of estimating the 2-D motion in the image plane. Many existing algorithms that attempt to recover the 3-D structure and motion assume the availability of accurate estimate for the 2-D motion in the form of the disparity vector [11] or the image flow field [12], although some algorithms bypass this intermediate step and compute the 3-D motion directly from the image intensity [10] or from contours [8]. Still, the problem of image motion estimation is an important one in many other applications such as object tracking and motion compensation for the efficient image transmission [7].

In computing the 2-D motion, two popular appoaches worth mentioning are the intensity-based method and the edge-based method. Intensity-based methods analyze the raw or filtered intensity across two or more image frames. While the image motion may be estimated in the frequency domain [5], from spatio-temporal derivatives [6] or by the intensity correlation [1], all the approaches rely on the same assumption (called *the convected-invariance of intensity*) that the intensity of physical point does not change in time, in the Lagrangian sense.

Edge-based methods compare the edge profile obtained by applying the edge detector to the image intensity. The correspondence between edge contours is established at every edge point [4] or at the entire contour [13], after imposing additional constraints to overcome the *Aperture Problem*. One must notice that any edge-based method inherently assumes the convected-invariance of edge, namely, edges are preserved in time while undergoing deformation so that the deformation of edges manifests the underlying deformation of image due solely to the 3-D motion.

In this paper we show that these assumptions on invariance of intensity and edge contour do not go along together; if intensity is invariant, contour is not; if contour is invariant, the convected invariance of intensity cannot hold. This suggests that the motion estimates obtained from two different approaches do not agree with each other. However, our analysis will show that such disagreement is relatively small when the edge strength is large. This effect becomes visible

when the edge is weak, and motion estimates obtained from the weak edge tend to be erroneous. This argument agrees with the heuristic control strategy often adopted in many motion estimation algorithms, in which weak edges are to be discarded below certain threshold. Care must be given to the motion estimation scheme which utilizes multi-resolution representation, since the smoothing increases the dynamic range of motion vector while decreasing the edge strength.

We are interested specifically in contours generated by the Marr-Hildreth edge operator [9] because of its popularity and some desirable properties discovered by several researchers [14, 17, 2]. Image is smoothed by the Gaussian G_σ, convolved with the Laplacian ∇, then the zerocrossing is traced. Throughout the rest of paper contour refers to the zerocrossing of $\nabla^2 G_\sigma * I$, unless otherwise stated.

Figure 1 shows a typical procedure for the edge-based method. $f(\mathbf{x})$ is the original gray level image of a scene and $Mf(\mathbf{x})$ is the gray level image of the same scene after motion, where M is the 2-D motion induced by the 3-D rigid body motion of the scene relative to the camera, which needs to be recovered. G and E are smoothing and edge detection operators, respectively.

One can readily see at the bottom of the figure that the procedure seeks for M', the motion between contour, instead of M. M' is defined as $M'(EGf(\mathbf{x})) = EGMf(\mathbf{x})$. The necessary and sufficient condition for $M' = M$ is that $EGM = MEG$. Only under this condition the deformation of contour will represent the underlying motion of image itself. But the condition may not always be satisfied. For example, Figures 3.a and 3.b show two consecutive image frames which contain the projection of a curved surface with a two-dimensional sinusoidal pattern undergoing a small but non-trivial 3-D motion relative to the viewer. The relative motion between two images is depicted by the vector field of Figure 3.e. Zerocrossing contours obtained from these images are shown in Figures 3.c and 3.d. If the zerocrossing contours evolve in the same way as the image intensity does, the contours in Figure 3.c will be deformed as the dashed curve in Figure 3.f according to the transformation shown in Figure 3.e. The solid curves depict the actual zerocrossings of Figure 3.d for comparison. We observe that

- The predicted contour does not coincide with the actual contour. *i.e.*, $EGM \neq MEG$.

- On large part of the contours, they do coincide to each other (or the deviation is small enough), so they can be used as features for the purpose of motion estimation.

- Large deviations occur at those portions of contour where variations of the intensity are small, *i.e.* at weak edges.

We are going to investigate the difference between the measured contour, EGM and the predicted one, MEG where E is the *zero-crossing* edge operator, G is the *Gaussian* convolution and M represents the image transformation such that $Mf(\mathbf{x}) = f(T^{-1}\mathbf{x})$, where T is the *affine* transformation. The analysis is based on the assumption that the intensity of a physical point does not change due to motion.

The issue on the deviation of zerocrossing edges in a single image has been studied by several researchers. Yuille and Poggio considered the deviation in edge location as a function of the Gaussian scale factor [17]. Clark discussed the same problem in the context of stereo vision, treating the deviation of the edge location as the result of a changing Gaussian scale factor [3]. In our case, as we shall see, the deviation of the zerocrossing edge caused by motion is the result of both the changing Gaussian scale factor and the shape of the edge detection operator used.

The reason that the motion causes the edge deviation is because the smoothing operation and the edge detection are both neighborhood operation. Since the image transformation induced by the 3-D motion is non-Euclidean transformation in general, the shape and the area of a neighborhood are not preserved under the transformation. As an operator with the same kernel is applied to the original and transformed images, and the transformed image actually covers different physical area, it produces different results. In Section 2 we will investigate separately two effects of the motion which cause the zerocrossing edge to deviate from its "correct" position. In Section 3 and 4 we will combine two effects to estimate the total deviation of the edge position. We will also give the estimate of the change in the edge strength. These estimates could serve as a criterion

to determine whether a contour segment should or should not be used in the 2-D motion recovery procedure.

2 The Motion-Compensated Edge Operator and Gaussian Scale Matrix

Suppose $f(\mathbf{x})$ and $\tilde{f}(\tilde{\mathbf{x}})$ are the gray level images of a scene before and after a motion, respectively. We denote $\tilde{f}(\tilde{\mathbf{x}}) \equiv Mf(\mathbf{x})$, where M is the image transformation operator corresponding to the motion from $f(\mathbf{x})$ to $\tilde{f}(\tilde{\mathbf{x}})$. The explicit relation between $f(\mathbf{x})$ and $\tilde{f}(\tilde{\mathbf{x}})$ is determined by many factors - illumination, reflectance function of the object surface, the 3-D shape of object, the relative motion between the scene and the camera, etc. However, if the motion is small, we may assume that the intensity of a physical point does not change during the motion. The image transformation is describe by a pure geometrical operation which is completely determined by the 3-D motion and the structure of scene, *i.e.*, $Mf(\mathbf{x}) = f(T^{-1}\mathbf{x})$, where T is the geometrical transformation induced by the 3-D motion such that $\tilde{\mathbf{x}} = T\mathbf{x}$. For example, Waxman and Wohn has shown that the second-order deformation is a good approximation for the motion of smooth surfaces [13]. We, however, assume that T is approximated as the affine transformation in a small neighborhood since, although the second-order model may provide the better approximation to the real motion, second-order terms are relatively small within a local neighborhood. Thus,

$$
\begin{cases}
Mf(\mathbf{x}) &= f(T^{-1}\mathbf{x}), \\
T\mathbf{x} &= A\mathbf{x} + b,
\end{cases}
\tag{1}
$$

where A is a 2 by 2 matrix which describes the linear deformation of image, and b is a 2 by 1 column matrix which specifies the average translation. As for the usefulness of this assumption, we have reported an algorithm that recovers the 3-D motion and structure from the affine parameters at two consecutive frames [15]. Assuming the inverse transformation of T exists, we see that $M^{-1}\tilde{f}(\mathbf{x}) = f(T^{-1}\mathbf{x})$ holds.

2.1 The Motion-compensated Edge Operator

We investigate the effect of motion on the edge position. As mentioned earlier, we consider the zerocrossing of the Laplacian of gray image. The edge strength is given as the magnitude of the gradient of Laplacian. Let $Z(\mathbf{x})$ be the Laplacian image of $f(\mathbf{x})$. Applying the *zero-crossing edge operator* E to a gray level picture $f(\mathbf{x})$ we get the edge image:

$$\begin{cases} Ef(\mathbf{x}) & = \begin{cases} |Z_{\mathbf{x}}(\mathbf{x})|; & Z(\mathbf{x}) = 0 \\ \\ 0; & \text{otherwise,} \end{cases} \\ \\ \\ Z(\mathbf{x}) & = tr\left[\frac{\partial^2 f}{(\partial \mathbf{x})^2}\right]_{(\mathbf{x})}, \end{cases} \tag{2}$$

where $Z_{\mathbf{x}}$ is the derivative of Z with respect to $\mathbf{x}$. The contour (edge) Γ is defined as

$$\Gamma = \{\mathbf{x} \mid Z(\mathbf{x}) = 0, \ |Z_{\mathbf{x}}(\mathbf{x})| \neq 0\}.$$

Likewise, applying E to the second image $Mf(\mathbf{x})$ we obtain

$$\begin{cases} EMf(\mathbf{x}) & = \begin{cases} |\tilde{Z}_{\mathbf{x}}(\mathbf{x})|; & \tilde{Z}(\mathbf{x}) = 0 \\ \\ 0; & \text{otherwise,} \end{cases} \\ \\ \\ \tilde{Z}(\mathbf{x}) & = tr\left[\frac{\partial^2 \tilde{f}}{(\partial \mathbf{x})^2}\right]_{(\mathbf{x})} = tr\left[A^t \frac{\partial^2 f}{(\partial \mathbf{x})^2} A\right]_{(T^{-1}\mathbf{x})}. \end{cases} \tag{3}$$

The contour obtained by applying the edge operator E to $Mf(\mathbf{x})$ is

$$\tilde{\Gamma} = \left\{\mathbf{x} \mid \tilde{Z}(\mathbf{x}) = 0, \ |\tilde{Z}_{\mathbf{x}}(\mathbf{x})| \neq 0\right\}.$$

We would like to see if contours are invariant with respect to motion, namely, all the points on contour Γ have their corresponding points on contour $\tilde{\Gamma}$. This property may be called "convected invariance of zerocrossings" referring to the similar property of intensity. One necessary condition for the convected invariance of zerocrossings is that $\tilde{\Gamma}$, when it is brought back to the initial frame

via the inverse transformation of T, must coincide with Γ.

$$\begin{cases} M^{-1}EMf(\mathbf{x}) &= \begin{cases} |Z'_{\mathbf{x}}(\mathbf{x})|; & Z'(\mathbf{x}) = 0 \\ \\ 0; & \text{otherwise}, \end{cases} \\ \\ \\ Z'(\mathbf{x}) &= \tilde{Z}(T\mathbf{x}) = tr\left[A^t \frac{\partial^2 f}{(\partial \mathbf{x})^2} A\right]_{(\mathbf{x})} \end{cases} \qquad (4)$$

In general $Z'(\mathbf{x})$ (defined in Equation (4)) does not agree with $Z(\mathbf{x})$ (defined in Equation (2)). In fact one can show that

$$\Gamma' = \left\{\mathbf{x} \mid Z'(\mathbf{x}) = k\{Z(\mathbf{x}) + F(\mathbf{x})\} = 0, \; |Z'_{\mathbf{x}}(\mathbf{x})| \neq 0\right\},$$

where k and F are determined by the transformation T. The zerocrossing - being the root of the above equation - is affected by the presence of $F(\mathbf{x})$, whereas k is a scalar constant which does not change the root. At this point, we could proceed to estimate the deviation of edge position by studying the difference between Γ and Γ'. Instead, we postpone the analysis on the zerocrossing and move our focus to another source of edge deviation due to the Gaussian smoothing, since in practice edges are extracted from the Gaussian-convolved image, rather than the original one. The exact formulars for k and F will be given later as we combine two sources of deviation together. At this moment, we simply observe that Γ is identical to Γ' iff $F(\mathbf{x}) = 0$. This result may be paraphrased as; *The image motion M and the edge detector E do not commute, i.e., $EM \neq ME$.* It is interesting to ask if there exist an edge detector commutable with any affine transformation T. If so, we can apply this operator to a pair of intensity images to yield the matching pair of invariant contours. Unfortunately, it is unlikely that such operator exists since characteristics of the intensity profile (such as inflection points, maxima, minima) which are related to the edge profile change as the image undergoes the (affine) transformation. At best, we can conceive an edge detector which changes itself according to the image transformation so that edges extracted from one frame are preserved in other frames.

Definition 1 *Motion-compensated Edge Operator*

We call the edge operator $E' = M^{-1}EM$ the motion-compensated edge operator for the motion M. We also define the motion-compensated edge function as the corresponding functional form $Z'(\cdot)$.

From this definition, it is easy to show that $ME' = EM$. The net result of applying E' followed by the transformation M is identical to the edge profile obtained from the transformed image. In practice, the operator E' is of little use since the motion M must be known prior to the edge detection. However, it will serve as a key concept later in Section 3 as we analyze the total deviation due to the edge detection and the Gaussian smoothing together.

2.2 Motion-compensated Gaussian Scale Matrix

In order to study the effect of the Gaussian smoothing on edge position we define the *generalized Gaussian convolution* of a gray level image as

$$G_\Sigma f(\mathbf{x}) = \frac{1}{2\pi\sqrt{|\det\Sigma|}} \int e^{-\frac{1}{2}(\mathbf{x}-\xi)^t\Sigma^{-1}(\mathbf{x}-\xi)} f(\xi)d\xi \tag{5}$$

where Σ is a 2 by 2 matrix which we call a *Gaussian scale matrix*. In case $\Sigma = \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, we call Σ an isotropic Gaussian scale matrix and σ the scale factor.

The convolution of the transformed image $Mf(\mathbf{x})$ with the Gaussian kernel defined by the same scale matrix is

$$
\begin{aligned}
G_\Sigma M f(\mathbf{x}) &= \frac{1}{2\pi\sqrt{|\det\Sigma|}} \int e^{-\frac{1}{2}(\mathbf{x}-\xi)^t\Sigma^{-1}(\mathbf{x}-\xi)} f(T^{-1}\xi)d\xi \\
&\overset{\xi=T^{-1}\zeta}{=} \frac{|\det A|}{2\pi\sqrt{|\det\Sigma|}} \int e^{-\frac{1}{2}(\mathbf{x}-T^{-1}\zeta)^t\Sigma^{-1}(\mathbf{x}-T^{-1}\zeta)} f(\zeta)d\zeta \\
&= \frac{1}{2\pi\sqrt{|\det(A\Sigma A^{-1})|}} \int e^{-\frac{1}{2}(T\mathbf{x}-\zeta)^t(A\Sigma A^t)^{-1}(T\mathbf{x}-\zeta)} f(\zeta)d\zeta \\
&= (G_{A\Sigma A^t} f)(T^{-1}\mathbf{x})
\end{aligned}
\tag{6}
$$

Again, we bring $G_\Sigma M f(\mathbf{x})$ back into the original frame:

$$M^{-1}G_\Sigma M f(\mathbf{x}) = G_{A\Sigma A^t} f(\mathbf{x}). \tag{7}$$

We can see that $M^{-1}G_{\Sigma}M$ still defines the Gaussian convolution but with a different scale matrix, and that the Gaussian-smoothed image is not invariant with respect to motion since $G_{A\Sigma A^t} \neq G_{\Sigma}$ in general. Thus we have

Proposition 1 *Motion-compensated Gaussian Scale Matrix*

If G_{Σ} is a Gaussian convolution and M is defined as in Equation (1), then $G_{\Sigma'} = M^{-1}G_{\Sigma}M$ is still a Gaussian convolution with a different scale matrix $\Sigma' = A\Sigma A^t$. We call Σ' the motion-compensated Gaussian scale matrix for the motion.

Notice that Σ' is not necessarily isotropic even if Σ is.

3 The Deviation of the Edge

We are now ready to consider the total deviation for the zerocrossing contour of Gaussian-smoothed image. The approach is similar to that of Section 2.1 except that we are now dealing with the Gaussian convolved image. The presence of the Gaussian introduces additional complexity to our derivation. The new quantities defined in the previous section enables us to simplify the derivation quite elegantly. By using Definition 1 and Proposition 1 we have

$$M^{-1}EG_{\Sigma}M = (M^{-1}EM)(M^{-1}G_{\Sigma}M) = E'G_{\Sigma'} \tag{8}$$

Thus, the edge image obtained by smoothing a transformed image with the Gaussian mask and then applying the edge operator can also be obtained by smoothing the original image with the Gausssian mask determined by the motion-compensated Gaussian scale matrix, applying the motion-compensated edge operator, and then transforming the resulting image ($EG_{\Sigma}M = ME'G_{\Sigma'}$).

Let Γ be the edges of the original image;

$$\begin{cases} \Gamma &= \{\mathbf{x} \mid Z(\mathbf{x}, \Sigma) = 0\} \\ Z(\mathbf{x}, \Sigma) &= tr\left[\frac{\partial^2 G_{\Sigma}I}{(\partial \mathbf{x})^2}\right]_{(\mathbf{x})} \end{cases} \tag{9}$$

In order to estimate the deviation, we bring the edges of the transformed image back to the original image frame. Let us call such edge Γ'. (See Figure 2). Using the results we obtained from the previous sections and combining the effects of E' and Σ' together, we have

$$
\begin{cases}
\Gamma' & = & \{\mathbf{x} \mid Z'(\mathbf{x}, \Sigma') = 0\} \\
Z'(\mathbf{x}, \Sigma') & = & k[Z(\mathbf{x}, \Sigma') + F(\mathbf{x}, \Sigma')],
\end{cases}
\tag{10}
$$

where k and $F(\cdot)$ are determined by Σ and transformation T (see Equation (1)). If we denote

$$
A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}
\quad \text{and} \quad
\left[\frac{\partial^2 (G_\Sigma f)}{(\partial \mathbf{x})^2} \right]_{(\mathbf{x}, \Sigma)} = \begin{bmatrix} J_{xx} & J_{xy} \\ J_{xy} & J_{yy} \end{bmatrix}_{(\mathbf{x}, \Sigma)}
$$

then one can show that

$$
\begin{cases}
k & = & \frac{1}{2}\left[a_1^2 + a_2^2 + a_3^2 + a_4^2\right] \\
F(\mathbf{x}, \Sigma) & = & \frac{1}{k}\left[(a_1^2 + a_2^2 - a_3^2 - a_4^2)(J_{xx} - J_{yy}) + 2(a_1 a_3 + a_2 a_4)J_{xy}\right]_{(\mathbf{x}, \Sigma)}
\end{cases}
\tag{11}
$$

Notice that a_i's are determined solely by the motion whereas J_{ij}'s are functions of image intensity and the Gaussian scale. For a given motion, if $\Gamma = \Gamma'$ for a particular image then we say the zero-crossing contours of the image are invariant under the motion; if it holds for all images then we say that the zero-crossing edge operator is invariant under the motion. It is clear that the zero-crossing edge operator is invariant if $F(\cdot) = 0$ and $\Sigma' = \Sigma$.

Proposition 2

The zero-crossing edge operator is invariant under the (2-D) rotation and the (2-D) translation if an isotropic Gaussian scale matrix is used.

Proof: From Equation (11) it is trivial to show that $F(\cdot) = 0$ and $\Sigma' = \Sigma$. So $Z(\mathbf{x}) = 0$ and $Z'(\mathbf{x}) = 0$ define the same contour. ∎

In general $\Gamma \neq \Gamma'$. The size of the deviation between the contours gives us a measure of the variation of the zero-crossing contour under the motion. (See Figure 2). We use the *normal distance* as the measure of the deviation between two contours:

Definition 2 *(Deviation)*

For every point $\mathbf{x}_0$ *on* Γ, *let* ℓ *be the line perpendicular to the tangent of* Γ *at* $\mathbf{x}_0$ *and let* $\mathbf{x}'$ *be the point where* ℓ *intersects* Γ'. *We define the deviation at* $\mathbf{x}_0$ *to be the distance between* $\mathbf{x}_0$ *and* $\mathbf{x}'$.

4   Estimation of the Deviation

Using Proposition 2 we can simplify Equation (1) by dropping the 2-D translational and rotational parts of the transformation, since any 2 × 2 matrix A^{-1} can be decomposed as $A^{-1} = R(\frac{\alpha}{2})DR(-\frac{\alpha}{2})R(\theta)$, where the R's are rotational matrices and $D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$ is a diagonal matrix. Let $\mu = (d_1 + d_2)/2$ and $\nu = (d_1 - d_2)/2$; dropping $R(\theta)$, Equation (1) becomes

$$\begin{cases} Mf(\mathbf{x}) &= f(A^{-1}\mathbf{x}) \\ A^{-1} &= \mu I + \nu S_\alpha \end{cases} \tag{12}$$

where

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \qquad S_\alpha = \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix}$$

We can see that the transformation represented by A^{-1} is basically a dilation with μ as the average scale factor, and ν is the relative difference of the scale factors in the two principle directions, these two directions are obtained by rotating x and y axes by $\alpha/2$.

4.1   Estimation of Deviation when the Motion is Small

In case the motion is small, we can write

$$\mu = 1 + \delta \quad \text{and} \quad |\nu|, |\delta| \ll 1 \tag{13}$$

Suppose the original image f and the transformed image Mf are both convolved with the mask determined by the Gaussian scale matrix $\Sigma_0 = \sigma^2 I$. Then we will have Γ and Γ' defined by

$$\left\{ \begin{array}{rcl} \Gamma & = & \{\mathbf{x} \mid Z(\mathbf{x}_0, \Sigma_0) = 0\} \\ \Gamma' & = & \{\mathbf{xI} \mid k[Z(\mathbf{x}, \Sigma') + F(\mathbf{x}, \Sigma')] = 0\} \end{array} \right. \tag{14}$$

where

$$\left\{ \begin{array}{rcl} k & = & \mu^2 + \nu^2 \\ F(\mathbf{x}, \Sigma) & = & \frac{\mu\nu}{\mu^2 + \nu^2}[\cos\alpha\,(J_{xx} - J_{yy}) + \sin\alpha\,J_{xy})]_{(\mathbf{x}, \Sigma)} \\ \Sigma' & = & k(\Sigma_0 + \Delta\Sigma) \\ \textbf{where} & & \Delta\Sigma \equiv \frac{\mu\nu}{\mu^2 + \nu^2}(\Sigma_0 S_\alpha + S_\alpha \Sigma_0). \end{array} \right. \tag{15}$$

Now let $\mathbf{x}_0$ be a point on Γ and $\mathbf{x}' = \mathbf{x}_0 + \Delta\mathbf{x}$ be the corresponding point on Γ'. Expanding $Z(\mathbf{x}', \Sigma')$ at $(\mathbf{x}_0, \Sigma_0)$:

$$Z(\mathbf{x}', \Sigma') = Z(\mathbf{x}_0, \Sigma_0) + Z_\mathbf{x}(\mathbf{x}_0, \Sigma_0)\Delta\mathbf{x} + Z_\Sigma(\mathbf{x}_0, \Sigma_0)\Delta\Sigma + h.o.t.,$$

where Z_Σ is a row vector of partial derivatives of Z with respect to the elements of Σ and $\Delta\Sigma$ is written as a column vector. Substituting the above into Equation (14) and omitting the higher order terms we get

$$Z_\mathbf{x}(\mathbf{x}_0, \Sigma_0)\Delta\mathbf{x} \approx F(\mathbf{x}', \Sigma') - Z_\Sigma(\mathbf{x}_0, \Sigma_0)\Delta\Sigma \tag{16}$$

On the other hand, by the definition of the deviation (see Figure 2), the vector $\Delta\mathbf{x}$ is perpendicular to the tangent vector $Z_\mathbf{x}(\mathbf{x}_0, \Sigma_0)$, so

$$Z_\mathbf{x}(\mathbf{x}_0, \Sigma_0) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \Delta\mathbf{x} = 0. \tag{17}$$

We can solve for $\Delta\mathbf{x}$ from Equations (16) and (17):

$$\Delta\mathbf{x} \approx \frac{F(\mathbf{x}', \Sigma') - Z_\Sigma(\mathbf{x}_0, \Sigma_0)\Delta\Sigma}{|Z_\mathbf{x}(\mathbf{x}_0, \Sigma_0)|^2} Z_\mathbf{x}(\mathbf{x}_0, \Sigma_0)$$

Thus we get the estimate of the deviation:

$$|\Delta\mathbf{x}| \approx \frac{|F(') - Z_\Sigma(0)\Delta\Sigma|}{|Z_\mathbf{x}(0)|} \tag{18}$$

11

where $F(')$ is evaluated at $(\mathbf{x}', \Sigma')$ and $Z_{\mathbf{x}}(0)$ and $Z_{\Sigma}(0)$ are evaluated at $(\mathbf{x}_0, \Sigma_0)$.

We also can estimate the edge strength of Γ' at $\mathbf{x}'$:

$$
\begin{aligned}
|\tilde{Z}_{\mathbf{x}}(T\mathbf{x}', \Sigma_0)| &= |A^{-T} Z_{\mathbf{x}}(\mathbf{x}', \Sigma')| \\
&= |A^{-T} Z'_{\mathbf{x}}(\mathbf{x}', \Sigma')| \\
&\approx k^{3/2} |Z_{\mathbf{x}}(0) + Z_{\mathbf{xx}}(0)\Delta\mathbf{x} + Z_{\mathbf{x}\Sigma}(0)\Delta\Sigma + F(')|
\end{aligned}
\tag{19}
$$

The ratio of the edge strength is

$$
\begin{aligned}
\frac{|\tilde{Z}_{\mathbf{x}}(T\mathbf{x}', \Sigma_0)|}{|Z_{\mathbf{x}}(\mathbf{x}_0, \Sigma_0)|} &\approx k^{3/2} \frac{|Z_{\mathbf{x}}(0) + Z_{\mathbf{xx}}(0)\Delta\mathbf{x} + Z_{\mathbf{x}\tau}(0)\Delta\Sigma + F(')|}{|Z_{\mathbf{x}}(0)|} \\
&\leq k^{3/2} \left| 1 + \frac{|Z_{\mathbf{xx}}(0)\Delta\mathbf{x} + Z_{\mathbf{x}\tau}(0)\Delta\Sigma + F(')|}{|Z_{\mathbf{x}}(0)|} \right| .
\end{aligned}
\tag{20}
$$

4.2 Special Cases

4.2.1 One-dimensional case:

Let $T^{-1}x = \mu x$; where $\mu = 1 + \delta; |\delta| \ll 1$. The 1-D Gaussian scale matrix is just a scalar σ^2, so we can reparameterize $Z(\mathbf{x}, \Sigma)$ as $Z(x, \sigma)$. The relation between the edge functions is

$$
Z'(x, \sigma') = \mu^2 Z(x, \mu\sigma).
\tag{21}
$$

Thus the deviation is given by:

$$
|\Delta x| \approx |\frac{Z_\sigma}{Z_x} \delta\sigma_0|,
\tag{22}
$$

It is caused purely by the change of the effective Gaussian mask size, and is proportional to the real mask size σ_0 used to convolve the images. We can compare Equations (21) and (22) to Equations (4) and (19) of Clark [3] — setting $\mu = (1 + \beta_1)$, $\delta = \beta_1$, and $Z_x/Z_\sigma = m$ — where the equations are for the edge deviation in the case of stereo vision. Our result can also be compared to that of Yuille [17], which gives the deviation of edge location with respect to the change in the Gaussian scale factor.

The ratio of the edge strength in this case is

$$
\left| \frac{\tilde{Z}_x(Tx_0)}{Z_x(x_0)} \right| \approx a^{-3} \left| 1 + \frac{Z_{xx}(0)\Delta x - Z_{x\sigma}(0)\delta\sigma_0}{Z_x(0)} \right|
\tag{23}
$$

4.2.2 $\nu = 0$ case :

In this case $A = \mu I$, and the transformation is an isotropic dilation. If the original Gaussian scale matrix Σ is isotropic, i.e., $\Sigma_0 = \sigma_0^2 I$, then Σ' is still an isotropic matrix. So we can parametrize the scale matrix by a scale σ. Equation (15) becomes

$$\begin{cases} k &=& \mu^{-2} \\[2mm] F(\cdot) &=& 0 \\[2mm] \sigma' &=& \mu\sigma_0 = (1+\delta)\sigma_0. \end{cases}$$

The deviation estimate is

$$|\Delta \mathbf{x}| \approx \frac{|Z_\sigma \delta \sigma_0|}{|Z_\mathbf{x}|} \tag{24}$$

and the ratio of the edge strength is

$$\left| \frac{\tilde{Z}_\mathbf{x}(T\mathbf{x}_0)}{Z_\mathbf{x}(\mathbf{x}_0)} \right| \approx k^3 \left| \frac{|Z_\mathbf{x}(0) + Z_\mathbf{xx}(0)\Delta\mathbf{x} - Z_{\mathbf{x}\sigma}(0)\delta\sigma_0|}{|Z_\mathbf{x}(0)|} \right|. \tag{25}$$

All results are the same as in the 1-D case except that x is a vector.

4.2.3 $\alpha = 0$ case :

In this case, $A = uI + vS_0 = \begin{bmatrix} \mu+\nu & 0 \\ 0 & \mu-\nu \end{bmatrix}$, the transformation is a dilation but with different scale factors along the x and y axes. For an original Gaussian scale matrix $\Sigma_0 = \sigma_0^2 I$, Σ' is a diagonal matrix, so we can use a vector $\epsilon = \begin{bmatrix} \sqrt{\Sigma_{11}} \\ \sqrt{\Sigma_{22}} \end{bmatrix}$ **to parameterize the Gaussian scale matrix.** Thus Equation (15) becomes

$$\begin{cases} k &=& \mu^2 + \nu^2 \\[2mm] F(\cdot) &=& \frac{\mu\nu}{\mu^2+\nu^2}(J_{xx} - J_{yy}) \\[2mm] \epsilon' &=& \begin{bmatrix} \mu+\nu \\ \mu-\nu \end{bmatrix}\sigma_0. \end{cases}$$

The deviation estimate is

$$|\Delta\mathbf{x}| \approx \frac{|F(') - Z_\epsilon(0)\Delta\epsilon|}{|Z_\mathbf{x}(0)|} \tag{26}$$

where $\Delta\epsilon = \begin{bmatrix} \delta + \nu \\ \delta - \nu \end{bmatrix} \sigma_0.$

5 Experiment and Discussion

5.1 Experiment on Synthetic Images

To check the soundness of our theory, we conducted an experiment on sythetically generated images (Figures 4). Figures 4.a and 4.b are two 256 by 256 images which contain the projection of a planar surface with a two-dimensional sinusoidal pattern undergoing a non-trivial 3-D motion relative to the viewer. The 2-D image transformation induced by this 3-D motion is an affine transformation with:

$$A = \begin{bmatrix} 1.5 & 0.5 \\ -0.5 & 0.9 \end{bmatrix} ; \quad \text{and} \quad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} . \tag{27}$$

A Gaussian convolution with $\sigma = 2$ is applied to these images and followed by the zerocrossing edge detection. The results is shown in Figure 4.c — edge obtained from the original image and 4.d — edge obtained from the transformed image. To compare the results, we apply the affine transformation (Equation (27)) to the countour in Figure 4.c to get Figure 5.a. Figure 5.b is the superimposition of Figure 5.a and Figure 4.d. We can see the deviations on part of the countour.

The countour in Figure 5.c is obtained from the transformed image (Figure 4.b) by locating the zero-crossings with the motion-compensated Gaussing scale mask and motion-compensated edge operator as discribed in the Section 2. Again we superimpose Figure 5.c on top of Figure 5.a (shown as Figure 5.d). We can see that with motion-compensated operators the countour obtained is much closer to the "correct" position.

5.2 Discussion

We showed that the invariance of intensity and the invariance of zerocrossing contours conflict with each other. The image motion revealed by the intensity change does not agree with the

motion revealed by the contour evolution in general. Assuming that intensity is invariant we derived the formular for the size of deviation in localizing contours. The deviation consists of two parts, one contributed by the change of the motion-compensated Gaussian mask, and the other by the change of the motion-compensated edge operator. Both changes reflect the fact that image transformation induced by the 3-D motion are non-Euclidean and does not preserve shape and area of a neighborhood. Therefore, any edge detector which does not change itself adaptively to the underlying motion possesses this undesirable property, although the exact size of deviation may differ from what we have derived in this paper. A special case of changing the Gaussian mask is that of changing the mask size but keeping the mask isotropic; this effect has been dicussed in detail by Yuille and Poggio [17].

Both the deviation of the edge location and the change of the edge strength are related to the transformation parameters, the Gaussian mask parameters and the the intensity derivatives. In general the transformation parameters are unknown, but if we have an estimate of these parameters, (as in the iterative procedure [16], we can use the parameters obtained in the previous step as the estimate of this step), then we can estimate the deviation of the contour by using Equation (18). If this deviation exceeds the maximum deviation allowed in the 2-D motion computation procedure, one can simply discard the corresponding segment of the contour from the matching process. We can also see that the deviation is inversely proportional to $|Z_{\mathbf{x}}(0)|$, the strength of the edge at this point. So, as we would expect, a strong edge is much more reliable than a weak one. In particular, $F(\cdot)$ and $\Delta\Sigma$ both contain a factor of $\frac{\mu\nu}{\mu^2+\nu^2} \ll 1$, so if $|Z_{\mathbf{x}}(0)|$ is not too small the deviation will not be too large. This is why strong zero-crossing contours can be used as features for the purpose of recovering the 2-D motion.

References

[1] P. Anandan, "A Unified Perspective on Computational Techniques for the Measurement of

Visual Motion", Proc. *1'st Intl. Conf. Computer Vision*, 219-230, June 1987.

[2] J. Babaud, A. P. Witkin, M. Baudin and R. Duda, "Uniqness of the Gaussian Kernel for Scale-Space Filtering", *IEEE Trans. PAMI-8*, 26-33, January 1986.

[3] J. J. Clark and P. D. Lawrence, "A Theoretical Basis for Diffrequency Stereo", *Computer Vision, Graphics, and Image Processing* **35** 1-19, 1986.

[4] E. C. Hildreth, "Computation Underlying the Measurement of Visual Motion", *Artificial Intelligence* **23**, 309-355, 1984.

[5] D. Heeger, "Optical Flow from Spatiotemporal Filters", Proc. *1'st Intl. Conf. Computer Vision*, 181-190, June 1987.

[6] B. K. P. Horn and B. G. Schunck, "Determining Optical Flow", *Artificial Intelligence* **17**, 185-204, 1981.

[7] T. S. Huang (ed.), **Image Sequence Processing and Dynamic Scene Analysis**, Heidelberg: Springer-Verlag, 1983.

[8] K. Kanatani, "Coordinate Rotation Invariance of Image Characteristics for 3D Shape and Motion Recovery", Proc. *1'st Intl. Conf. Computer Vision*, 55-64, June 1987.

[9] D. Marr and E. C. Hildreth, "Theory of Edge Detection", *Proc. Royal Soc. London* **B 207**, 187-217, 1980.

[10] S. Negahdaripour and B. K. P. Horn, "Direct Passive Navigation", *IEEE Trans. PAMI-9*, 168-176, January 1987.

[11] R. Tsai, "Estimating 3-D Motion Parameters and Object Surface Structures from the Image Motion of Curved Edges", Proc. *IEEE Conf. Computer Vision and Pattern Recognition*, June 1983.

[12] A. M. Waxman and S. Ullman, "Surface Structure and 3-D Motion from Image Flow: A Kinetic Analysis", *Intl. Journal of Robotics Research* **4**, 72-94, 1985.

[13] A. M. Waxman and K. Wohn, "Contour Evolution, Neighborhood Deformation and Global Image Flow: Planar Surface in Motion", *Intl. Journal of Robotics Research* **4**, 95-108, 1985.

[14] A. P. Witkin, "Scale Space Filtering" *Intl. Joint Conf. Art. Intell.*, 1019-1022, 1983

[15] K. Wohn and J. Wu, "3-D Motion Recovery from Time-varying Optical Flows", Proc. *AAAI-86*, 670-675, September 1986.

[16] J. Wu, "Motion Estimation from Image Sequences", *Ph.D. Thesis*, Division of Applied Sciences, Harvard University, September 1987.

[17] A. L. Yuille and T. A. Poggio, "Scaling Theorems for Zero Crossings", *IEEE Trans. PAMI-8*, 15-25, January 1986.
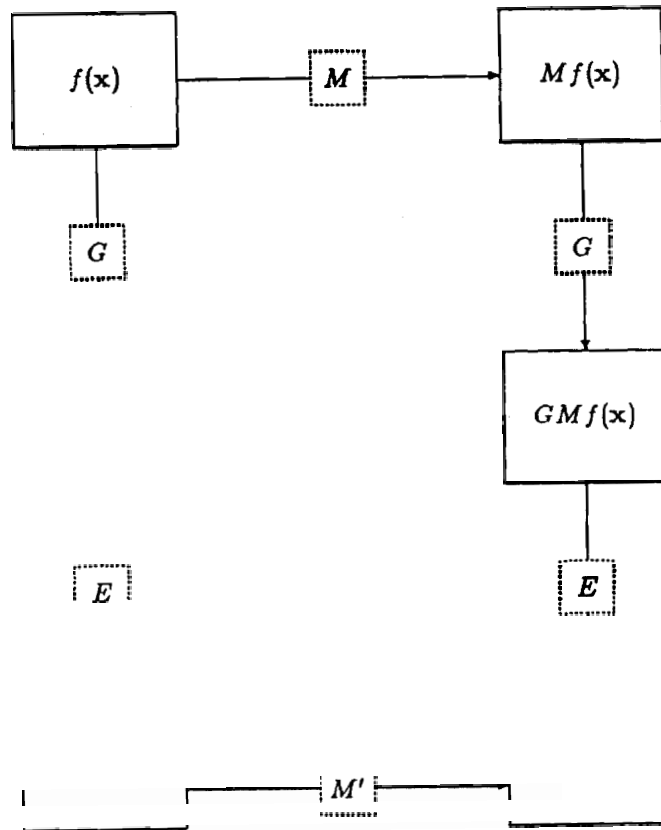
Figure 1: A typical procedure for a contour-based method.

$\Gamma : EG_\Sigma f(\mathbf{x})$

$\tilde{\Gamma} : EG_\Sigma M f(\mathbf{x})$

$\Gamma' : M^{-1}EG_\Sigma M f(\mathbf{x})$
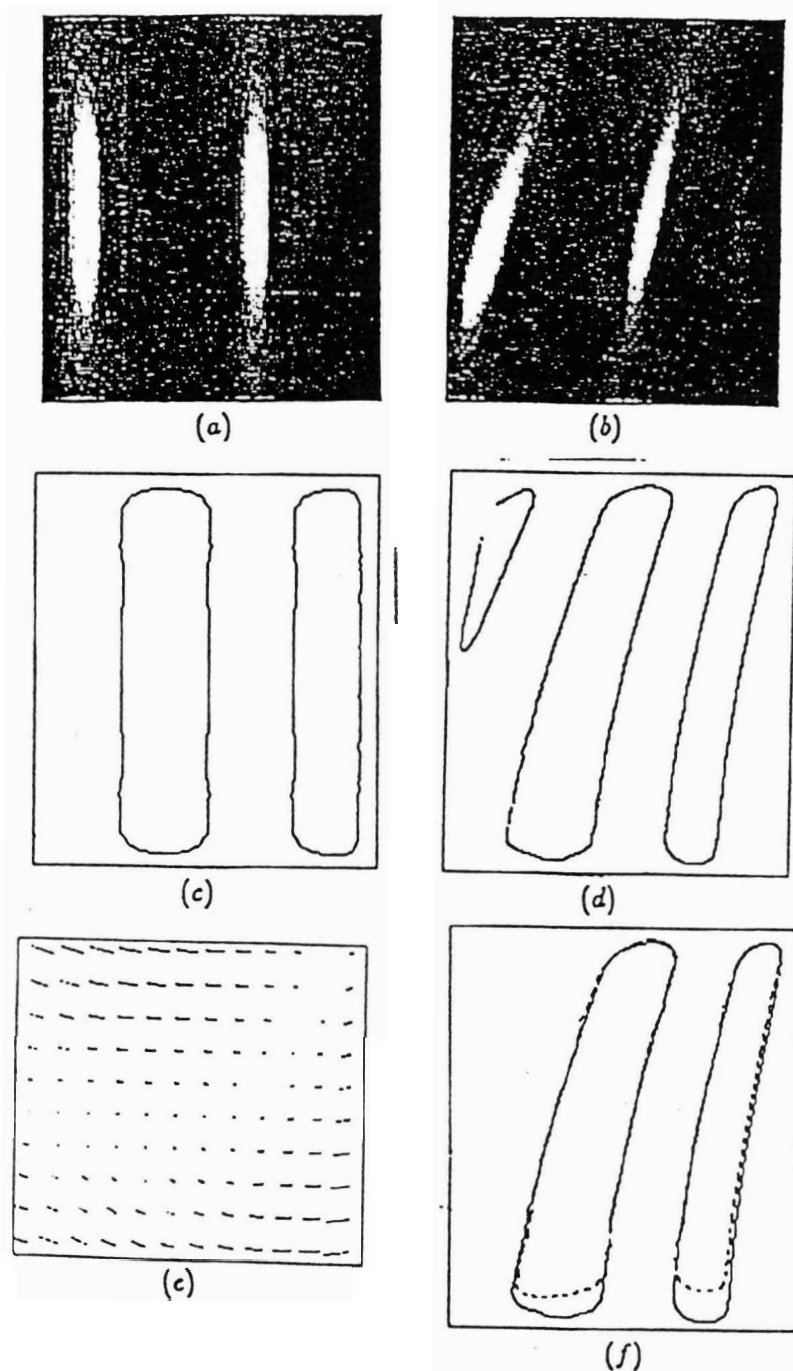
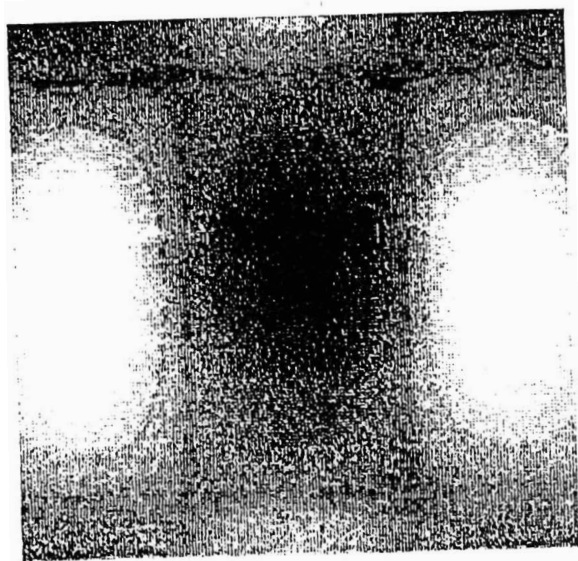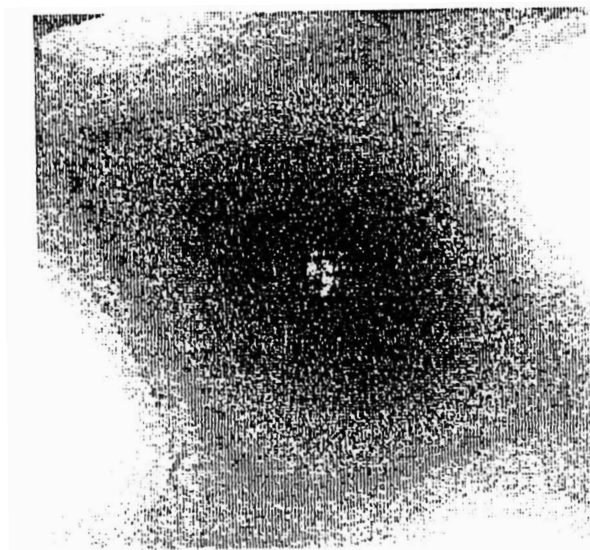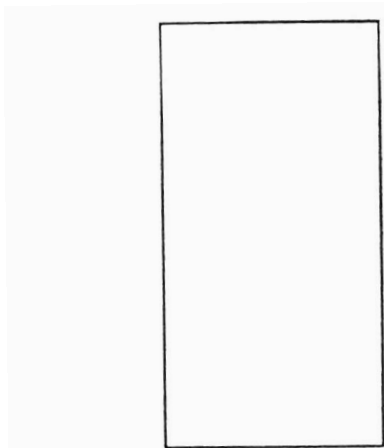Figure 2: Γ, $\tilde{\Gamma}$ and Γ'

(a)　(b)

(c)　(d)

(e)　(f)
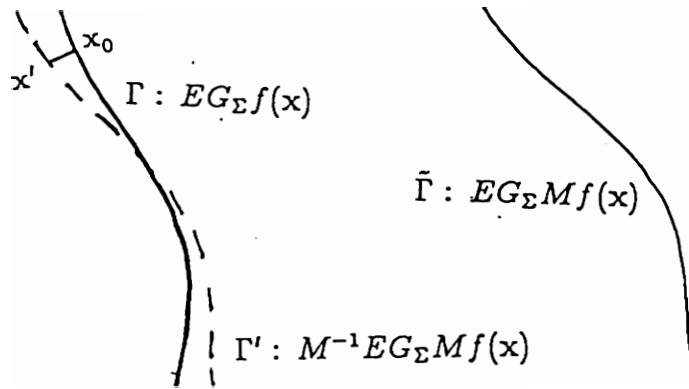
Figure 3: The variation of the zerocrossing contour.

(a)

(b)

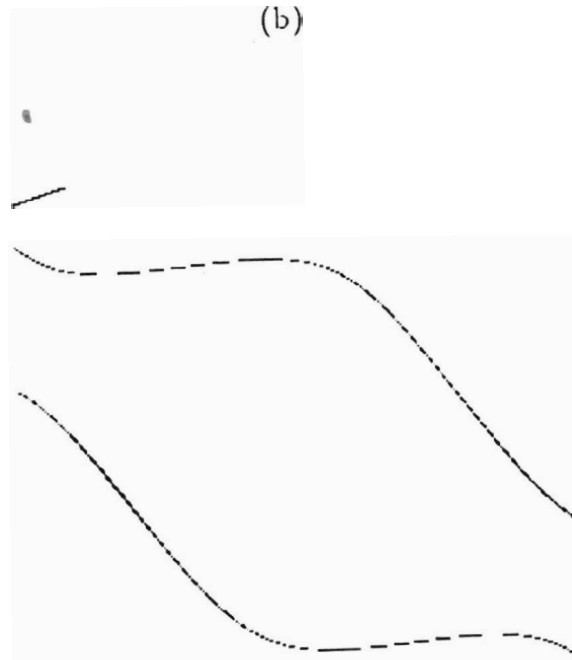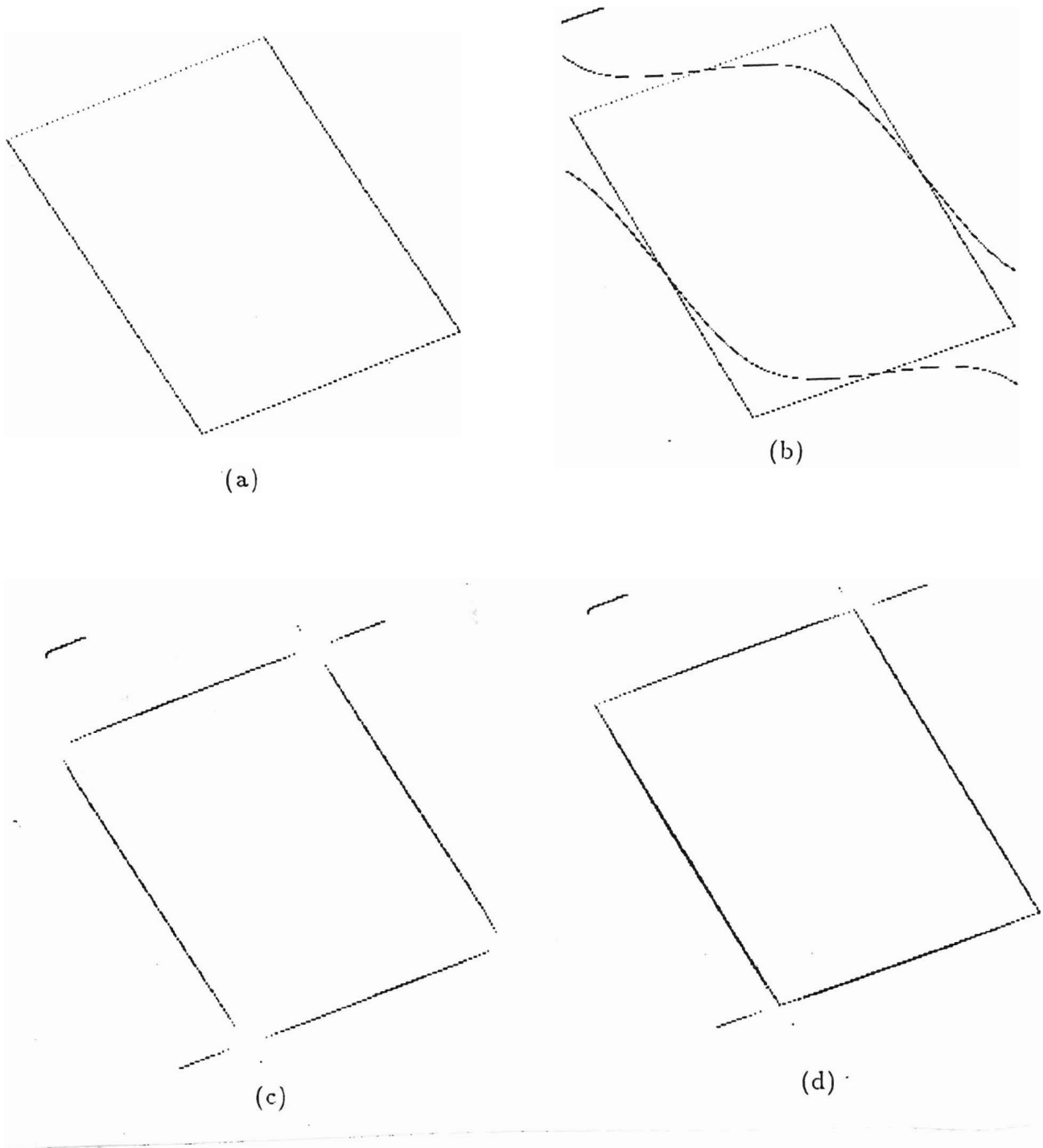(c)

(d)

Figure 4: Experiment on sythetic images.

(a)

(b)

(c)

(d)

Figure 5: Experiment on sythetic images.