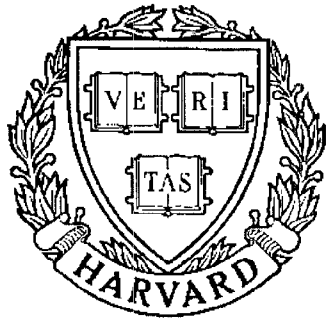
TECHNICAL RESEARCH REPORT

SYSTEMS RESEARCH CENTER



Supported by the National Science Foundation Engineering Research Center Program (NSFD CD 8803012), Industry and the University

A Perceptually Motivated Three-Component Image Model

by X. Ran and N. Farvardin

TR 92-32

A Perceptually Motivated Three-Component Image Model †

Xiaonong Ran and Nariman Farvardin
Electrical Engineering Department
and
Systems Research Center
University of Maryland
College Park, Maryland 20742

Abstract

In this paper, results of psychovisual studies of the human visual system are discussed and interpreted in a mathematical framework. The formation of the perception is described by appropriate minimization problems and the edge information is found to be of primary importance in visual perception. Having introduced the concept of edge strength, it is demonstrated that strong edges are of higher perceptual importance than weaker edges (textures). We have also found that smooth areas of an image influence our perception together with the edge information, and that this influence can be mathematically described via a minimization problem. Based on this study, we have proposed to decompose the image into three components: (i) primary, (ii) smooth and (iii) texture, which contain, respectively, the strong edges, the background and the textures. An algorithm is developed to generate the three-component image model and an example is provided in which the resulting three components demonstrate the specific properties as expected. Finally, it is shown that the primary component provides a superior representation of the strong edge information as compared with the Laplacian-Gaussian Operator scheme which is a popular edge extraction method.

†This work was supported in part by National Science Foundation grants NSFD MIP-86-57311 and NSFD CDR-85-00108.

I Introduction

In most image compression systems, the primary objective is that of a high fidelity reconstruction of the original image with as small a number of bits as possible. Naturally, the fidelity criterion used for the design of the compression system plays a significant role on the system performance. A commonly used fidelity criterion is the mean squared-error (MSE) distortion. The main attributes of the MSE are (i) its mathematical tractability and (ii) the fact that small values of MSE correspond to perceptually high quality reconstructed images. The latter fact is important since human eye is commonly the final judge of the reconstructed image. With the MSE and its variants as fidelity criterion, various image coding techniques such as adaptive discrete cosine transform (ADCT) coding, various forms of vector quantization (VQ) and subband coding (SBC) have been studied [1]-[4]. While all these techniques have led to relatively high quality reconstructed images at bit rates about 1.0 bit/pixel and above, they usually produce specific types of visible distortion (blockiness, blurred edges, jagged edges, etc.) at lower bit rates.

Over the past decade, a new class of image coding schemes, generally referred to as *second-generation techniques*, has been developed [5]-[8]. These methods, which have exhibited improvements over the MSE-based schemes, attempt to describe the image in terms of more nearly physical entities such as regions or contours for a more compact representation of the image and hence a higher compression ratio [6]; this is in contrast with the waveform representation of the image in the MSE-based schemes.

In this paper, in an approach not unlike that of second-generation image coding techniques, we have developed a model for real-world imagery based on those features of the image signal that are of distinct significance to human perception. Specifically, we have formalized some previous psychovisual studies [9] [10] to characterize the features of the image signal that are responsible for the formation of the perception in the human visual system (HVS). In particular, some interesting evidences on the binocular nature of the human vision in [9] are used to formulate the formation of

the perception as a minimization of the intensity variation energy; this has led to the notion of "strong edges" which apparently plays a significant role in the perception. Additionally, we have used the psychovisual observations in [10] to mathematically formulate the interaction between the strong edges and the areas of smooth intensity variations. To characterize the strong edges, we have introduced the concept of the "stressed image" and defined the strong edges as the high curvature energy pixels of the stressed image. This stressed image, which has an interesting analogy to the stable configuration of a mechanical structure, is generated by a space-variant low-pass filtering of the original image. The above formalism has led to a three-component image model consisting of (i) the strong edge component, (ii) the smooth intensity variation component and (iii) the texture component.

The three-component model developed here is quite general and might prove useful in various image processing algorithms. However, it was primarily developed for an image coding situation where the different features of the image signal can be classified (according to the role they play in the formation of human perception), extracted and treated separately for subsequent encoding. This idea is similar to the sketch-based image coding scheme of [7] [8] which consists of two components: one is similar to the strong edge component of the three-component model developed here, and the other is simply the residual between the original image and an image obtained merely from the strong edge information. In the sketch-based scheme, the strong edges are extracted by the Laplacian-Gaussian Operator (LGO) followed by a gradient operator. This edge extraction scheme suffers from certain drawbacks which will be discussed in detail. In addition, a careful comparison between the LGO-based method and the edge extraction scheme associated with the three-component model will be presented. The applications of the three-component image model in defining a perceptual distortion criterion for specific image coding schemes are presented in a sequel paper [11].

This paper is organized as follows. In Section II, certain psychovisual evidences of the human perception are presented and discussed. The evidences lead to the characterization of the strong edge information in Section III, followed by an algorithm for

the extraction of strong edges along with the development of the three-component image model in Section IV. In Section V, an example of the three-component image model is provided and certain comparisons against the LGO-based edge extraction scheme are presented. Summary and conclusions are provided in Section VI.

II Observations on Psychovisual Aspects of Human Visual System

In the next three subsections, we describe some observations on psychovisual aspects of the human visual system. The objective is to extract and discriminate different properties of image signals which are of significance to our perception. Interpretations in the context of image coding are also provided, and will form the motivation and basis for the work described later. In the first two subsections, we use the binocular nature of human visual system to explain the relationship between the image signals and our perception.

A. Edge Information of Image Signals

As described in [9], in natural binocular vision, when two views are presented with two forms which are different (in the sense that they do not admit of being combined into the image of a single object), the images of both forms will generally be seen at the same time superposed on one another in the field of view. Usually, in some locations of the field of view, one image dominates the other, and vice versa in other parts of the field. For the situation where broad black-and-white figures are displayed to both views, the general rule is that the dominating image along an edge and in its vicinity will be the one which owns the edge. As an example, let us consider Fig. 1, in which we have two bars, one vertical and one horizontal. When the vertical bar is seen by the left eye and the horizontal bar by the right eye, without devoting exclusive attention to any of the two, the total effect will be an image similar to Fig. 2. As shown in Fig. 2, the perception will be a cross which is black over the

center square; the background appears white. The four arms of the cross are perfectly black at their ends and almost entirely white near the center square, with transitions in between.

Based on this phenomenon, we may conclude that the collection of the individual pixel intensity values without any interaction between them (the most primitive property of the image) is not what produces the visual perception. For otherwise, the image in the field of view should be formed at each pixel with an intensity produced only by the two corresponding pixel intensities of the two images according to a certain law, and thus the perception should be a uniform combination of the two pictures. Obviously, the image shown in Fig. 2 cannot be constructed by a *uniform* combination of the two images in Fig. 1.

Based on this observation, we conjecture that here it is the edge property (or information), to be defined next, of the image signal that arouses our visual perception. For the images in Fig. 1, the edge information is described by (i) the *locations* of variations of intensity values and (ii) the related *intensity values* at these locations. For example, in Fig. 1 (a), the locations of the intensity variations are along the border of the vertical black bar, and the related intensity values are those intensity values immediately inside and outside the border. The above actually describes what we generally refer to as an *edge*. This edge basically has two related intensity values, zero (255) corresponding to black (white), representing the intensity variation. We call the contour formed by the lower intensity value the *lower brim* of the edge (or lower edge brim), and the contour formed by the higher intensity value the *upper brim* of the edge (or upper edge brim). For example, in Fig. 1 (a), the rectangular contour of the lower edge brim is just inside the border, while the one for the upper edge brim is just outside. The locations and intensity values of the lower and upper edge brims completely define the corresponding edge, and, therefore, we may define an edge through the concept of edge brims which will be characterized mathematically later in this paper.

We now go back to our conjecture that the edge information is responsible for the formation of the visual perception. Remarkably, this conjecture is justified in this

case by actually producing Fig. 2 *only* from the knowledge of lower and upper edge brims of the two images in Fig. 1 by minimizing the variation in intensity values as described below.

Let a generic digital image of size $M \times M$ be denoted by an array of real numbers $\{x_{i,j}\}$, $i,j = 0,1,\ldots,M-1$, where $x_{i,j}$ is the intensity value of the pixel (i,j) at the ith row and the jth column. Note that an image $\{x_{i,j}\}$, $i,j = 0,1,\ldots,M-1$, can be defined as a set $\mathcal{X}$ of triples: $\mathcal{X} \equiv \{(i,j,x_{i,j}),\, i,j = 0,1,\ldots,M-1\} \subset \mathbb{R}^3$. Then the lower and upper edge brims of an image $\mathcal{X}$ are a subset of $\mathcal{X}$. Let the subsets of lower and upper edge brims of the images (a) and (b) in Fig. 1 be denoted by $\mathcal{B}_a$ and $\mathcal{B}_b$, respectively. We combine the information in $\mathcal{B}_a$ and $\mathcal{B}_b$ by forming a set $\mathcal{B}$ from $\mathcal{B}_a \cup \mathcal{B}_b$ in the following way. We first define two projection functions, $f_1 : \mathbb{R}^3 \to \mathbb{R}^2$ where $f_1((i,j,x)) = (i,j)$, and f_2: $\mathbb{R}^3 \to \mathbb{R}$ where $f_2((i,j,x)) = x$. We denote by $f_k(\mathcal{A})$ the image of a set $\mathcal{A}$ under f_k, $k = 1,2$. For a set $\mathcal{A} \subset \mathbb{R}^3$, we also define $\mathcal{A}^{(i,j)} = \{s : s \in \mathcal{A} \text{ and } f_1(s) = (i,j)\}$. Then the set $\mathcal{B}$ is defined as,

$$\mathcal{B} = \{(i,j,\overline{x}_{i,j}) :\ (i,j) \in f_1(\mathcal{B}_a \cup \mathcal{B}_b)\}, \tag{II.1}$$

where $\overline{x}_{i,j}$ is the average value of the elements of the set $f_2\left((\mathcal{B}_a \cup \mathcal{B}_b)^{(i,j)}\right)$. In other words, the set $\mathcal{B}$ is a "linear" combination of $\mathcal{B}_a$ and $\mathcal{B}_b$ in the above sense.

Before describing how the image shown in Fig. 2, denoted by $\mathcal{X}_c$, can be obtained from the edge information contained in $\mathcal{B}$, we make the following definition. The *variation energy* of intensity values of an image $\mathcal{X}$ is defined as:

$$V_{\mathcal{X}} = \sum_{i=0}^{M-2}\sum_{j=0}^{M-2}[(x_{i,j} - x_{i,j+1})^2 + (x_{i,j} - x_{i+1,j})^2]. \tag{II.2}$$

As indicated above, our objective is to generate an image $\mathcal{X}_c$, as shown in Fig. 2, solely from the information contained in $\mathcal{B}$. In other words, we seek to find an image $\mathcal{X}_c$ having minimum information while containing the information in $\mathcal{B}$. We call this concept the *minimum information principle* [1]. We quantify the information content of an image $\mathcal{X}$ by its variation energy $V_{\mathcal{X}}$, and define $\mathcal{X}_c = \{(i,j,x^c_{i,j})\}$ to be the

[1]The term "minimum information principle" is motivated by a concept, referred to as "*no news is good news*," introduced in [12].

solution of the minimization problem:

$$\min_{\{x_{i,j}\}} V_{\mathcal{X}}, \quad \text{subject to } \mathcal{X} \cap \mathcal{B} = \mathcal{B}, \tag{II.3}$$

where $\mathcal{X} = \{(i, j, x_{i,j})\}$. For example, an image with uniform intensity values has zero variation energy, and thus contains the smallest amount of information by the above definition. Because of the quadratic nature of the objective function $V_{\mathcal{X}}$, we may write it in matrix notation: $\mathbf{x}^T L_V \mathbf{x} + \mathbf{x}^T H \mathbf{x}_{\mathcal{B}} + \mathbf{x}_{\mathcal{B}}^T D \mathbf{x}_{\mathcal{B}}$, where $\mathbf{x}$ is a vector in $\mathbb{R}^{M^2-|\mathcal{B}|}$ containing elements $x_{i,j}$, $(i,j) \notin f_1(\mathcal{B})$ in a certain order, $\mathbf{x}_{\mathcal{B}}$ is a vector in $\mathbb{R}^{|\mathcal{B}|}$ with elements $x_{i,j}$, $(i,j) \in f_1(\mathcal{B})$, the matrix L_V is a non-negative definite (positive definite when $\mathcal{B} \neq \emptyset$) matrix, and the superscript ^{T} indicates vector transpose. Then, when $\mathcal{B} \neq \emptyset$, the unique solution of problem (II.3) is given by the vector $\mathbf{x}_c \equiv -L_V^{-1} H \mathbf{x}_{\mathcal{B}}/2$. For the given example, the resulting image $\mathcal{X}_c$ is shown in Fig. 2; this image, obtained from our minimization problem, is very similar to the corresponding illustration shown in Fig. 73 of [9].

Therefore, in an image information system for which the human visual system is the final receiver, the necessary and sufficient information to be transmitted and/or stored for the images in Fig. 1 is the edge information.

To summarize, we explain the phenomenon in Fig. 2 with a conjecture that the visual perception is derived from the edge information of the image. Furthermore, the perception can be illustrated by an image which is the solution of a minimization problem similar to (II.3). In this sense, we may take the mechanism governed by this minimization problem as the one which forms our perception.

We now question whether all edges are of the same importance in forming our perception. In the next subsection, we describe a binocular phenomenon in an effort to answer this question in the context of image coding.

B. Strong Edges and Textures

Similar to Fig. 1, shown in Fig. 3 are two images, (a) and (b), separately presented to the two eyes. In Fig. 3, image (a), examined by the left eye, is a black cross, while image (b) for the right eye is a network of slanted black lines over a white

background. Without any special attention to either image, the usual binocular perception would be similar to the image shown in Fig. 4. That is, the image of the cross prevails along its edges; it is only at the center of the cross and away from its edges that the pattern of the network texture may be visible [9].

In view of this phenomenon, the two images in Fig. 3 do not have edge information of the same importance, for otherwise in the vicinity of the edge of the cross the network texture should not be invisible. Thus, in this case, we cannot combine the edge information "linearly" as in (II.1). We speculate that an edge has a *strength* property associated with it. Edges of relatively high (low) strength have stronger (weaker) influence on our perception, and may be called *stronger (weaker) edges.* Weaker edges commonly correspond to *textures* in the usual sense, and thus are referred to as textures. As compared with textures, we may simply refer to stronger edges as *strong edges.*

Apparently, edges with larger intensity variations and shorter widths (distance between the corresponding edge brims) are relatively more significant to our perception, and are generally called sharp edges. However, edges do not individually influence our perception; rather, they interact with other neighboring edges. For example, a thin black line in the network of Fig. 3 (b) actually consists of two edges of the same intensity variation and width as for the edge of the cross, but it has less perceptual significance when compared with the edge of the cross. Based on this observation, we conjecture that every edge has an original strength which is proportional to (i) the intensity variation between its two edge brims and (ii) its width. Neighboring edges interact with each other in an inhibitive way. The more closely two edges stand, the more severely their strengths are weakened. When two edges of equal original strength stand side by side, as in the case of a thin black line on a white background, their strengths vanish. The resulting strength of an edge after it interacts with the neighboring edges will be called the *strength* of the edge. Thus stronger edges are those edges of higher intensity variation and shorter width which are relatively isolated, while weaker edges are those of lower intensity variation and longer width, being relatively crowded with other edges.

We now use this qualitative notion of edge strength to explain the phenomenon of Fig. 4. Similar to Subsection II.A, we denote the subsets of lower and upper edge brims of images (a) and (b) in Fig. 3 by $\mathcal{B}_a$ and $\mathcal{B}_b$, respectively. We combine the information in $\mathcal{B}_a$ and $\mathcal{B}_b$ based on the concept of edge strength in the following way. When stronger edges meet weaker ones, the stronger ones dominate. This domination of the stronger edge will be extended to its neighborhood and the influence will diminish gradually as a function of the distance from the stronger edge. After a certain distance from the stronger edge, weaker edges will begin to show their effects which will eventually dominate at distances far enough from the stronger edge.

Our experiments indicate that this phenomenon for Figs. 3 and 4 can be explained as follows. In the close neighborhood of the edges represented by $\mathcal{B}_a$, the image for the perception is formed by solving a problem similar to (II.3) with $\mathcal{B}_a$ as the fixed-point set; we call this image $\mathcal{X}_A$. For distances sufficiently far from the edge in $\mathcal{B}_a$, where the influence of the strong edge diminishes, the image of the perception is formed by solving a problem like (II.3) with $\mathcal{B}_b$ as the fixed-point set; we call this image $\mathcal{X}_B$. In between these two extreme cases, the influence of the strong edge in $\mathcal{B}_a$ decreases, and the perception is formed by linearly combining $\mathcal{X}_A$ and $\mathcal{X}_B$ $(\lambda\mathcal{X}_A + (1-\lambda)\mathcal{X}_B)$ where λ decreases quadratically from 1 to 0 as the distance from the strong edge gets large. Our results indicate that this linear combination with the chosen weighting factor results in an image which agrees with our perception as well as the description given in [9, pp. 496-497]. Again as in Subsection II.A, we tested our explanation by actually constructing Fig. 4 according to the process described above.

We conclude that stronger edges are of higher importance to our perception. Therefore, to design an image information transmission system, one should treat edges in accordance with their strengths to achieve high efficiency. This concept is especially important for image transmission systems at low bit rates.

C. *Smooth Areas of Image Signals*

We now investigate the influence of smooth areas of an image on our perception. In the smooth areas of an image, intensity values change slowly. For example, in Fig. 1 (a), the areas inside and outside of the black bar are smooth areas where intensity values are actually constant. The smooth areas of constant intensity values are of little interest here, since we have indicated in Subsection II.*A* that they have no influence on the perception. Instead we will consider the case of smooth areas with non-constant intensity values.

Let us examine the patterns in Fig. 5 which are also considered in [10] [12]. In Fig. 5 (a) we have two concentric disks with their intensity values along the diameter as illustrated in Fig. 5 (b). While the perception of the image in Fig. 5 (a) varies from person to person, generally, one feels that it is a small dark disk in the middle of a larger and brighter one. In the extreme case, the inner disk is perceived as having constant intensity values. However, as shown in Fig. 5 (b), in the regions near the center of the disks and the border of the bigger disk, the intensity values are, in fact, identical.

We now describe this phenomenon quantitatively. We denote the set of lower and upper edge brims by $\mathcal{B}$ representing the two circular edges of the inner and outer disks in Fig. 5 (a). We denote the image in Fig. 5 (a) by $\mathcal{X}$ and the image of our perception by $\mathcal{X}_p = \{(i, j, x^p_{i,j})\}$. Then we conjecture that $\mathcal{X}_p$ is the solution of the following minimization problem:

$$\min_{\{y_{i,j}\}} \lambda \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} (y_{i,j} - x_{i,j})^2 + V_{\mathcal{Y}}, \tag{II.4}$$

subject to $\mathcal{Y} \cap \mathcal{B} = \mathcal{B}$, where $\lambda \geq 0$ is a weighting factor on the squared errors between $\mathcal{X}$ and $\mathcal{Y}$ in the objective function, $\mathcal{Y} = \{(i, j, y_{i,j})\}$, and $\mathcal{B} \neq \emptyset$. Notice that when $\lambda = 0$, $\mathcal{X}_p$ would be an image of two disks with constant intensity values for the inner disk and a linear transition of intensity values on the outer disk, which corresponds to the extreme case mentioned before. When λ gets larger, $\mathcal{X}_p$ would be closer to $\mathcal{X}$ in the Euclidean sense, and would have larger variation energy than the case with $\lambda = 0$.

We summarize the psychovisual results described in the above three subsections by stating that basically the edge information in an image is responsible for our perception, stronger edges are of higher importance to our perception and the smooth areas influence our perception together with the edge information.

In the next two sections, we will introduce a three-component image model based on these observations.

III Characterization of Strong Edge Information and Generation of the Stressed Image

As pointed out in Section II, for our visual perception, the strong edge information of an image plays a more important role as compared with texture and smooth areas. Therefore, in designing image coding systems, strong edges need special attention. To do this, we may first characterize and then extract this strong edge information so that it can be treated separately. In what follows, we will consider the characterization of strong edge information which leads to the concept of a stressed image.

A. Characterization of Strong Edge Information

In this subsection, we first provide a mathematical description of edges, strong or weak, and then develop a scheme for the discrimination of strong edges.

In what follows, we explain briefly a traditional way of describing edges [7] [8] [13]. For an image $\mathcal{X} = \{(i, j, x_{i,j}),\ i, j = 0, 1, \ldots, M-1\}$, we define the directional variations $v^r_{i,j}$ and $v^c_{i,j}$ at pixel (i, j),

$$v^r_{i,j} = x_{i,j} - x_{i,j+1}, \tag{III.1}$$

for $i = 0, 1, \ldots, M-1$, and $j = 0, 1, \ldots, M-2$,

$$v^c_{i,j} = x_{i,j} - x_{i+1,j}, \tag{III.2}$$

for $i = 0, 1, \ldots, M-2$, and $j = 0, 1, \ldots, M-1$. We define the locations of edges to

be those points where $v^r_{i,j}$ and/or $v^c_{i,j}$ assume large absolute values [2]. More precisely, we say that there is an edge point at $(i, j+0.5)$ if $|v^r_{i,j}| > T_v$, and an edge point at $(i+0.5, j)$ if $|v^c_{i,j}| > T_v$, for a positive threshold T_v. We then describe the corresponding edge at location $(i, j+0.5)$ or $(i+0.5, j)$ by the intensity values of neighboring pixels, $x_{i,j}$ and $x_{i,j+1}$, or, $x_{i,j}$ and $x_{i+1,j}$, respectively.

This scheme works well for edges which are one-pixel wide, but not for edges of multi-pixel width as illustrated by the following example. In Fig. 6 (a), we have an image with an edge which has an intensity jump from 20 to 220, and a variable-width: one pixel wide at the top and ten pixels wide at the bottom. The intensity values associated with the 256th scanline (center row) in Fig. 6 (a) is shown in Fig. 7 (a). On this scanline, the edge is 6 pixels wide: from pixel $(256, 255)$ to $(256, 261)$, where the intensity values change linearly from 20 to 220. Notice that $|x_{256,j} - x_{256,j+1}| = (220 - 20)/6$ for $255 \leq j \leq 260$. Therefore, for $T_v \geq 200/6$, no point on this scanline will be identified as an edge point. On the other hand, for $T_v < 200/6$, the points $(256, j + 0.5)$, $255 \leq j \leq 260$ will all be taken as edge points. While a complete description of the edge on the scanline is provided by these six detected edge points, this description is redundant, since the two intensity values at locations $(256, 255)$ and $(256, 261)$ completely define the edge here. These two pixels, $(256, 255, 20)$ and $(256, 261, 220)$, are referred to as edge brim points in Section II.

We now develop a new way of describing edges in the following. We define the second-order directional variations $D^r_{i,j}$ and $D^c_{i,j}$ for $\mathcal{X}$ at pixel (i, j),

$$D^r_{i,j} = x_{i,j-1} - 2x_{i,j} + x_{i,j+1}, \tag{III.3}$$

for $i = 0, 1, \ldots, M-1$ and $j = 1, \ldots, M-2$,

$$D^c_{i,j} = x_{i-1,j} - 2x_{i,j} + x_{i+1,j}, \tag{III.4}$$

for $i = 1, \ldots, M-2$ and $j = 0, 1, \ldots, M-1$. We then use these second-order directional variations to describe the *edge brim points* and then the corresponding

[2]The actual schemes introduced in [7] [8] [13] are further extensions of the concept here; they are combined with techniques of Gaussian filtering and gradient operator, and will be briefly described in Section V.

edges. More specifically, we define the *pixel row-curvature energy* $C^r_{i,j}$ and the *pixel column-curvature energy* $C^c_{i,j}$ by

$$C^r_{i,j} \equiv \begin{cases} (D^r_{i,j})^2 & \text{for } i = 0, 1, \ldots, M-1 \text{ and } j = 1, \ldots, M-2, \\ 0 & \text{otherwise,} \end{cases} \tag{III.5}$$

and,

$$C^c_{i,j} \equiv \begin{cases} (D^c_{i,j})^2 & \text{for } i = 1, \ldots, M-2 \text{ and } j = 0, 1, \ldots, M-1, \\ 0 & \text{otherwise,} \end{cases} \tag{III.6}$$

respectively, and will refer to them collectively as *pixel curvature energies.* We then define the set $\mathcal{B}_T(\mathcal{X})$ of edge brim points of $\mathcal{X}$ as follows,

$$\mathcal{B}_T(\mathcal{X}) \equiv \left\{(i, j, x_{i,j}) : \; C^r_{i,j} > T \text{ or } C^c_{i,j} > T\right\}, \tag{III.7}$$

where $T \geq 0$. We apply this characterization of edge brim points to the image in Fig. 6 (a) with a threshold $T = 32$; the detected edge brim points are indicated by bright points in Fig. 6 (b). We superimpose the image (b) on image (a), and show the result in image (c) where dark points are used for the upper edge brim points. These detected edge brim points are exactly at the places where the edge starts and ends. To demonstrate this, we repeat Fig. 7 (a) along with two vertical lines indicating the edge brim points in Fig. 7 (b). For the case of Fig. 6, the set $\mathcal{B}_T(\mathcal{X})$ completely defines the edge; in fact, we can recover the image in Fig. 6 (a) merely from the information in $\mathcal{B}_T(\mathcal{X})$ by solving the minimization problem (II.3).

We note that definition (III.7) gives a set of edge brim points regardless of whether they correspond to strong edges. To circumvent this problem, we introduce the concept of a *stressed image*: $\mathcal{X}^s = \{(i, j, x^s_{i,j})\}$ associated with the image $\mathcal{X}$. We first state here the properties which the stressed image $\mathcal{X}^s$ is required to possess, and then, in the next subsection, develop a scheme to generate the stressed image. We require that, at strong edges, the stressed image $\mathcal{X}^s$ closely approximate the original image $\mathcal{X}$, i.e., the squared-errors, $(x_{i,j} - x^s_{i,j})^2$, are small at these pixels; the pixel curvature energies of $\mathcal{X}^s$ of these pixels have no additional constraint other than that imposed by the squared-errors constraint. In other areas such as smooth and

texture areas, we require the pixel curvature energies of $\mathcal{X}^s$ to be small while placing a loose constraint on the squared-error $(x_{i,j} - x^s_{i,j})^2$. Thus $\mathcal{B}_T(\mathcal{X}^s)$, as compared with $\mathcal{B}_T(\mathcal{X})$, will only contain the edge brim points corresponding to strong edges because large pixel curvature energies occur only at the strong edges. Note that this stressed image is generally smooth, except at strong edges where it usually assumes large pixel curvature energies, and thus is "stressed". Note also that in the edge extraction schemes introduced in [7] [8] [13], the textures are suppressed by uniformly smoothing the image. While here we also smooth the image to suppress the textures, we do the smoothing *non-uniformly.* The term "stressed image" is used to indicate this non-uniformity of smoothing.

B. *Generating the Stressed Image*

In this subsection, we consider the generation of a stressed image $\mathcal{X}^s$ from an original image $\mathcal{X}$. Since the main property of the stressed image $\mathcal{X}^s$ is described in terms of (i) the squared-errors $(x_{i,j} - x^s_{i,j})^2$ and (ii) the pixel curvature energies of $\mathcal{X}^s$, we consider the following quantity at pixel (i,j), which is a combination of the above two quantities,

$$E_{i,j}(\lambda^1_{i,j}, \lambda^2_{i,j}, \lambda^3_{i,j}) = \lambda^1_{i,j}(x_{i,j} - x^s_{i,j})^2 + \lambda^2_{i,j} C^r_{i,j} + \lambda^3_{i,j} C^c_{i,j}, \tag{III.8}$$

where parameters $\lambda^1_{i,j}, \lambda^2_{i,j}, \lambda^3_{i,j}$ are three non-negative real numbers. We define the summation of these $E_{i,j}$ as

$$E(\mathcal{X}^s, \mathcal{X}, \Lambda) \equiv \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} E_{i,j}(\lambda^1_{i,j}, \lambda^2_{i,j}, \lambda^3_{i,j}), \tag{III.9}$$

where Λ represents the collection of the parameters $\lambda^1_{i,j}, \lambda^2_{i,j}$ and $\lambda^3_{i,j}$.

We then consider the following minimization problem for a given parameter set Λ

$$\min_{\{y_{i,j}\}} E(\mathcal{Y}, \mathcal{X}, \Lambda), \tag{III.10}$$

where $\mathcal{Y} \equiv \{(i,j,y_{i,j})\}$. Note that the objective function $E(\mathcal{Y}, \mathcal{X}, \Lambda)$ is convex as a function of $\mathcal{Y}$, since each $E_{i,j}(\lambda^1_{i,j}, \lambda^2_{i,j}, \lambda^3_{i,j})$ is a convex function of $\{y_{i,j}\}$ [14].

Therefore, $\{y^*_{i,j}\}$ is a solution of this minimization problem if and only if it is a solution of the following system of linear equations,

$$\nabla_{\{y_{i,j}\}} E(\mathcal{Y}, \mathcal{X}, \Lambda) \equiv 0, \tag{III.11}$$

where $\nabla_{\{y_{i,j}\}} E(\mathcal{Y}, \mathcal{X}, \Lambda)$ denotes the gradient of $E(\mathcal{Y}, \mathcal{X}, \Lambda)$ with respect to $\{y_{i,j}\}$. We may write the objective function in problem (III.10) in matrix notation:

$$[y_{i,j}]^T L[y_{i,j}] - 2[\lambda^1_{i,j} x_{i,j}]^T [y_{i,j}] + [\lambda^1_{i,j} x_{i,j}]^T [x_{i,j}], \tag{III.12}$$

where vectors $[y_{i,j}] \in \mathbb{R}^{M^2}$ and $[\lambda^1_{i,j} x_{i,j}] \in \mathbb{R}^{M^2}$ contain elements $y_{i,j}$ and $\lambda^1_{i,j} x_{i,j}$, respectively, in a certain order (one-to-one function) r: $\{0, 1, \ldots . M-1\} \times \{0, 1, \ldots, M-1\} \rightarrow \{0, 1, \ldots, M^2 - 1\}$, i.e., the $(r(i_0, j_0))$th element in the vectors $[y_{i,j}]$ and $[\lambda^1_{i,j} x_{i,j}]$ are y_{i_0,j_0} and $\lambda^1_{i_0,j_0} x_{i_0,j_0}$, respectively. The matrix $L = [\ell_{u,v}]$ is a non-negative definite matrix, where $\ell_{u,v}$ is the coefficient of the term $y_{r^{-1}(u)} y_{r^{-1}(u)}$ when $u = v$, and half of the coefficient of the term $y_{r^{-1}(u)} y_{r^{-1}(v)}$ when $u \neq v$, in $E(\mathcal{Y}, \mathcal{X}, \Lambda)$. With this matrix notation, (III.11) simplifies to the following:

$$L[y_{i,j}] = [\lambda^1_{i,j} x_{i,j}]. \tag{III.13}$$

We now show the existence and uniqueness of the solution for the minimization problem (III.10) when $\lambda^1_{i,j} > 0$ for all (i, j).

Lemma III.1 *The matrix L is positive definite if $\lambda^1_{i,j} > 0$ for all (i, j).*

Proof: Setting $x_{i,j} = 0$ for all (i, j) in the objective function (III.12), we have

$$E(\mathcal{Y}, \mathcal{Z}, \Lambda) = [y_{i,j}]^T L[y_{i,j}], \tag{III.14}$$

where $\mathcal{Z} = \{(i, j, 0)\}$. While

$$E(\mathcal{Y}, \mathcal{Z}, \Lambda) = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \lambda^1_{i,j} (y_{i,j})^2 + \lambda^2_{i,j} C^r_{i,j} + \lambda^3_{i,j} C^c_{i,j}, \tag{III.15}$$

we have that $[y_{i,j}]^T L[y_{i,j}] = 0$ if and only if $y_{i,j} = 0$ for all (i, j), when $\lambda^1_{i,j} > 0$ for all (i, j) (under the assumption that $\lambda^2_{i,j}, \lambda^3_{i,j}$ are non-negative). QED.

From now on, we assume that the parameter set Λ is such that all $\lambda^1_{i,j}$ are greater than zero, and therefore we will talk about *the* solution of problem (III.10), which

is $L^{-1}[\lambda^1_{i,j}x_{i,j}]$. The problem (III.10) is solved iteratively using the Gauss-Seidel iteration and with the Multi-Grid technique to speed up the convergence. The Gauss-Seidel iteration and the Multi-Grid technique are well known [15]. Nevertheless, the details of the implementation of these two techniques are included in [16].

We now consider the influences of the choices of $\lambda^1_{i,j}$, $\lambda^2_{i,j}$ and $\lambda^3_{i,j}$ on the solution, $\mathcal{Y}^* = \{(i,j,y^*_{i,j})\}$, of problem (III.10). A close examination of (III.8) indicates that if $\lambda^1_{i,j}$ is large compared with $\lambda^2_{i,j}$ and $\lambda^3_{i,j}$, $(x_{i,j}-y^*_{i,j})^2$ would be small, and the constraint on the pixel curvature energies of $\mathcal{Y}^*$ would be loose; on the other hand, if $\lambda^2_{i,j}$ and $\lambda^3_{i,j}$ are large as compared to $\lambda^1_{i,j}$, the pixel curvature energies of $\mathcal{Y}^*$ would be small, and the constraint imposed on the squared-error $(x_{i,j} - y^*_{i,j})^2$ would be relaxed. To make this observation more precise, we consider the relationship between the original image $\mathcal{X}$ and the solution $\mathcal{Y}^*$ governed by the minimization problem (III.10) as a filtering operation with input $\mathcal{X}$ and output $\mathcal{Y}^*$. In Appendix A, we investigate the frequency response of this filter (in its continuous form for simplicity) and show that it is a low-pass filter with cutoff frequencies at (i,j) controlled by the parameters $\lambda^1_{i,j}$, $\lambda^2_{i,j}$ and $\lambda^3_{i,j}$, and thus is space-variant. The relationship between the cutoff frequencies and the parameters given in Appendix A can be described as follows: Larger values of $\lambda^1_{i,j}$ give higher cutoff frequencies in both directions, while larger values of $\lambda^2_{i,j}$ and $\lambda^3_{i,j}$ lead to lower cutoff frequencies in row-direction and column-direction, respectively. It is important to note that it is not the absolute values of $\lambda^1_{i,j}$, $\lambda^2_{i,j}$, and $\lambda^3_{i,j}$ but rather the ratios $\lambda^2_{i,j}/\lambda^1_{i,j}$ and $\lambda^3_{i,j}/\lambda^1_{i,j}$ which influence the solution $\mathcal{Y}^*$.

The minimization problem (III.10) actually has an interesting interpretation in that the solution $\mathcal{Y}^*$ can be thought of as the stable configuration of a mechanical structure, a part of which is depicted in Fig. 8. In this structure, at each pixel location (i,j) we have a vertical spring with both ends fixed on a floor (height = 0) and a ceiling (height = 255), and a cylinder fixed on it at height $x_{i,j}$. This cylinder is constrained to move only in the vertical direction. For each row and column we have a flexible bar of the shape shown in Fig. 8 with a slot in the middle. The cylinders associated with each (i,j) are fitted inside the corresponding column bars, and the

column bars, in turn, are fitted inside row bars. Shown in Fig. 8 is the structure at pixel $(M-1,0)$. We now suppose that this mechanical structure assumes a configuration described by the heights $y_{i,j}$ of the cylinders. Then the potential energy of this configuration $\{y_{i,j}\}$ is approximated by $E(\mathcal{Y},\mathcal{X},\Lambda)$ defined in (III.9), where $\lambda^1_{i,j}(x_{i,j}-y_{i,j})^2$, $\lambda^2_{i,j}C^r_{i,j}$ and $\lambda^3_{i,j}C^c_{i,j}$ approximate the potential energies of the spring, the row bar and the column bar at pixel (i,j), respectively. The parameters $\lambda^1_{i,j}$, $\lambda^2_{i,j}$ and $\lambda^3_{i,j}$ control the rigidness of the spring, the row bar, and the column bar at pixel (i,j) [17].

To visualize the formation of the stable configuration for this mechanical structure, we note that if all the bars are taken out, the cylinders on the vertical springs for each pixel (i,j) will assume their stable positions at $x_{i,j}$, i.e., this configuration represents the original image $\mathcal{X}$. After we slide in all the row and column bars whose rigidness (or flexibility) are determined by the parameters, $\{\lambda^2_{i,j}\}$ and $\{\lambda^3_{i,j}\}$, the configuration of the structure will change, and will reach the final stable configuration $\{y^*_{i,j}\}$ which has the minimum potential energy, i.e., $\mathcal{Y}^*$ is the solution of problem (III.10). Due to this analogy to the mechanical system, we refer to (III.10) as an *Energy Minimization Model* (EMM) problem.

To obtain the stressed image $\mathcal{X}^s$ from the original image $\mathcal{X}$, it suffices to solve the EMM problem with a proper parameter set Λ, namely, small ratios $\lambda^2_{i,j}/\lambda^1_{i,j}$ and $\lambda^3_{i,j}/\lambda^1_{i,j}$ at locations of strong edges and larger ratios at other places. However, we can not specify the parameter set Λ a priori, since this requires the knowledge of the locations of strong edges - the very purpose of generating the stressed image.

We now proceed to explain qualitatively an approach for solving this problem with the help of the above mechanical structure. We assume that in the process of forming the stressed image $\mathcal{X}^s$, the bars and springs first have uniform rigidness, and the stable configuration reached by the structure is a uniformly smoothed version, denoted by $\mathcal{Y}^1 \equiv \{(i,j,y^1_{i,j})\}$, of the original image $\mathcal{X}$. Then we allow the bars to have more flexibility at locations corresponding to large pixel curvature energies (where the bars are most severely bent), and let the structure stabilize to a new configuration, denoted by $\mathcal{Y}^2 \equiv \{(i,j,y^2_{i,j})\}$. Then again we measure the pixel cur-

vature energies of this new configuration $\mathcal{Y}^2$, and allow higher flexibility of the bars at places of large pixel curvature energies. We continue the process by repeating the above procedure. Notice that the textures are suppressed in the first configuration $\mathcal{Y}^1$, and will continue to be suppressed in later configurations, while the strong edges which correspond to locations of large pixel curvature energies will be gradually approximated in the later configurations since we change the flexibilities of the bars to accentuate the strong edges.

To express the above idea mathematically, we start with a uniform parameter set Λ, i.e., parameter $\lambda^k_{i,j} = \lambda^k$ for all (i,j), $k = 1,2,3$, and solve problem (III.10) to get $\mathcal{Y}^1$. We then update the parameter set Λ by changing the ratios, $\lambda^2_{i,j}/\lambda^1_{i,j}$ and $\lambda^3_{i,j}/\lambda^1_{i,j}$, according to $C^r_{i,j}$ and $C^c_{i,j}$ of $\mathcal{Y}^1$. The updating strategy should be such that large values of $C^r_{i,j}$ and $C^c_{i,j}$ give rise to small ratios, $\lambda^2_{i,j}/\lambda^1_{i,j}$ and $\lambda^3_{i,j}/\lambda^1_{i,j}$, respectively. In the algorithm proposed here, we set $\lambda^2_{i,j}/\lambda^1_{i,j}$ and $\lambda^3_{i,j}/\lambda^1_{i,j}$ to be inversely proportional to $C^r_{i,j}$ and $C^c_{i,j}$, i.e.,

$$\lambda^2_{i,j}/\lambda^1_{i,j} = \beta/C^r_{i,j} \text{ and } \lambda^3_{i,j}/\lambda^1_{i,j} = \beta/C^c_{i,j} \tag{III.16}$$

where β is a constant. Then we solve problem (III.10) again with the new parameter set to obtain $\mathcal{Y}^2$. We repeat the above iteration until the relative variation of the objective function in (III.10) for the two consecutive iterations is smaller than a given threshold. The final solution $\mathcal{Y}^K$, where K is the total number of iterations, will be called a stressed image (denoted by $\mathcal{X}^s$ before) of the image $\mathcal{X}$.

IV Extraction of Strong Edge Information and a Three-Component Image Model

As indicated before, the strong edge information can be extracted by identifying pixels of large pixel curvature energies in a stressed image $\mathcal{X}^s$, and is represented by $\mathcal{B}_T(\mathcal{X}^s)$ in (III.7). While generating $\mathcal{B}_T(\mathcal{X}^s)$ based on (III.7) is straightforward, the choice of the threshold T is crucial and fine-tuning of T is computationally extensive. To circumvent these difficulties, further selection of the pixels in $\mathcal{B}_T(\mathcal{X}^s)$ is made by

identifying pixels of large *local maximum* pixel curvature energies as the resulting edge brim points. More specifically, the definition of edge brim points in (III.7) is modified as follows:

$$
\begin{aligned}
(i,j,x^s_{i,j}) \;& \in \mathcal{B}_T(\mathcal{X}^s) \\
\text{if} \quad & D^r_{i,j-1}D^r_{i,j} \le 0,\ D^r_{i,j}D^r_{i,j+1} \le 0,\ \text{and } C^r_{i,j} > T; \\
\text{or if} \quad & D^r_{i,j-1}D^r_{i,j} > 0,\ D^r_{i,j}D^r_{i,j+1} \le 0,\ \text{and } C^r_{i,j} > \max\{C^r_{i,j-1}, T\}; \\
\text{or if} \quad & D^r_{i,j-1}D^r_{i,j} \le 0,\ D^r_{i,j}D^r_{i,j+1} > 0,\ \text{and } C^r_{i,j} > \max\{C^r_{i,j+1}, T\}; \\
\text{or if} \quad & D^r_{i,j-1}D^r_{i,j} > 0,\ D^r_{i,j}D^r_{i,j+1} > 0,\ \text{and } C^r_{i,j} > \max\{C^r_{i,j-1}, C^r_{i,j+1}, T\}; \\
\text{or if} \quad & D^c_{i-1,j}D^c_{i,j} \le 0,\ D^c_{i,j+1}D^c_{i,j} \le 0,\ \text{and } C^c_{i,j} > T; \\
\text{or if} \quad & D^c_{i-1,j}D^c_{i,j} > 0,\ D^c_{i,j+1}D^c_{i,j} \le 0,\ \text{and } C^c_{i,j} > \max\{C^c_{i-1,j}, T\}; \\
\text{or if} \quad & D^c_{i-1,j}D^c_{i,j} \le 0,\ D^c_{i,j}D^c_{i+1,j} > 0,\ \text{and } C^c_{i,j} > \max\{C^c_{i+1,j}, T\}; \\
\text{or if} \quad & D^c_{i-1,j}D^c_{i,j} > 0,\ D^c_{i,j}D^c_{i+1,j} > 0,\ \text{and } C^c_{i,j} > \max\{C^c_{i-1,j}, C^c_{i+1,j}, T\}.
\end{aligned}
\tag{IV.1}
$$

Therefore, the set $\mathcal{B}_T(\mathcal{X}^s)$ contains pixels not only of high pixel curvature energies, but also of curvature energies larger than those of the neighboring pixels with the same sign of the second-order variations.

To complete the process of the extraction of the strong edge information, we will consider the generation of the contours of edge brims from $\mathcal{B}_T(\mathcal{X}^s)$ in the next subsection. The formation of these contours is also important for the coding of strong edge information [11] [16] [18] [19].

A. Generation of Edge Brim Contours

We now consider the problem of generating contours from the set $\mathcal{B}_T(\mathcal{X}^s)$ described in (IV.1). We define a contour as a sequence of triples: $\overline{b} \equiv \{(i^k, j^k, x^s_{i^k,j^k})\}$, $k = 0, 1, \ldots, m-1$, such that

$$
|i^{k-1} - i^k| \le 1,\ |j^{k-1} - j^k| \le 1, \tag{IV.2}
$$

and $(i^{k-1}, j^{k-1}) \neq (i^k, j^k)$, for $k = 1, \ldots, m-1$, and the *maximum-variation* of intensity values

$$\sigma_{\overline{b}} \equiv \max_{k=0,1,\ldots,m-1} |x^s_{i^k,j^k} - \overline{x}| \leq T_c, \tag{IV.3}$$

where $T_c \geq 0$, $\overline{x}$ is the average of the pixel intensity values on $\overline{b}$, and m is called the *length* of the contour. This definition is different from previous contour definitions [20] - [22] in condition (IV.3); condition (IV.3) is introduced to prevent a mixing of the upper and lower edge brims into one contour. In the following, we will describe the details of the contour extraction procedure used here.

Let us first consider the so-called *contour tracing* problem [20] - [22], which ignores condition (IV.3). This problem is solved by a local search algorithm. To describe this algorithm, consider Fig. 9 where a generic pixel (i, j) is circled, and its neighboring pixels are labeled by indices $0, 1, \ldots, 7$. Starting from a pixel $(i^0, j^0) \in f_1(\mathcal{B}_T(\mathcal{X}^s))$, the tracing algorithm searches among its neighboring pixels (starting from 0 in the clockwise direction) for one in $\mathcal{B}_T(\mathcal{X}^s)$. As soon as the algorithm finds a pixel $(i^1, j^1) \in f_1(\mathcal{B}_T(\mathcal{X}^s))$ in these neighboring pixels, it moves the search center to (i^1, j^1), and repeats the search again until the search fails. The resulting sequence of pixels, $(i^0, j^0, x^s_{i^0,j^0})$, $(i^1, j^1, x^s_{i^1,j^1}), \ldots$, defines the contour extracted.

Suppose that we have a contour $(i^0, j^0, x^s_{i^0,j^0})$, $(i^1, j^1, x^s_{i^1,j^1})$, $\ldots$, $(i^{m-1}, j^{m-1}, x^s_{i^{m-1},j^{m-1}})$ satisfying the geometric condition (IV.2). We now impose the constraint (IV.3) on this contour. While the whole contour may not satisfy this constraint for a given threshold T_c, short segments of the contour are likely to satisfy this condition. Thus this added constraint essentially breaks the contour into segments, which leads to a problem similar to the *straight-line-fitting* problem in [23], with the exception that the straight-line-segments in [23] are replaced by constant-intensity curve-segments here.

The contour extraction algorithm actually implemented is a combination of the above techniques and works as follows. Starting from a pixel $(i^0, j^0, x^s_{i^0,j^0}) \in \mathcal{B}_T(\mathcal{X}^s)$, in its eight neighboring pixels we search for pixels in $\mathcal{B}_T(\mathcal{X}^s)$, among which we choose a pixel that gives the minimum maximum-variation; if this maximum-variation is also lower than T_c, we take this pixel as the next pixel $(i^1, j^1, x^s_{i^1,j^1})$ on the contour. Then

we move the search center to $(i^1, j^1, x^s_{i^1,j^1})$, delete $(i^0, j^0, x^s_{i^0,j^0})$ from $\mathcal{B}_T(\mathcal{X}^s)$, and repeat the above procedure again. The process stops when we cannot locate a pixel in $\mathcal{B}_T(\mathcal{X}^s)$ having maximum-variation lower than T_c among the neighboring pixels of the current search center. This is the basic form of the algorithm. In the following, we discuss two other considerations in the actual algorithm implemented, namely, *back-tracing* and *prediction.*

Note that when we take an arbitrary pixel in $\mathcal{B}_T(\mathcal{X}^s)$ to begin a contour tracing, we might start in the middle of an actual contour, and would only get a portion of this contour using our basic algorithm. We circumvent this problem by introducing a flag such that when in the first search step we have more than one neighboring pixel in $\mathcal{B}_T(\mathcal{X}^s)$ giving maximum-variations lower than T_c, this flag is set to "1", and otherwise, it is set to "0". When this flag is "1", we first trace to one end of the contour, and then trace back to the other end of the contour.

Due to the iterative nature of the algorithm for generating the stressed image, it is possible that we lose a few pixels on some edge brim contours for a given threshold T_c. Some of the missing pixels may be recovered by a predictive process based the information of the generated contours. The details of the prediction is included in Appendix B.

Having determined all contours with the above algorithm, we eliminate those contours of small length, namely, those with length smaller than a certain threshold T_ℓ. Since strong edges are represented by two contours, namely, lower and upper edge brims, we may use this pairing property to further reduce spurious contours. More precisely, let us define a d-neighborhood of a pixel $(i, j, x^s_{i,j})$ as follows,

$$N(d, (i^k, j^k, x^s_{i^k,j^k})) \equiv \{(i, j, .) : \max\{|i - i^k|, |j - j^k|\} \leq d\}, \tag{IV.4}$$

where $d \geq 0$. We say that a pixel in $\mathcal{B}_T(\mathcal{X}^s)$ is paired in distance d if and only if the d-neighborhood of this pixel contains pixels on another contour of $\mathcal{B}_T(\mathcal{X}^s)$; and that a contour in $\mathcal{B}_T(\mathcal{X}^s)$ is paired in distance d if and only if it has at least T_ℓ pixels being paired in distance d. Having checked all contours in $\mathcal{B}_T(\mathcal{X}^s)$, we eliminate those contours not paired in distance d.

We will denote the final collection of the pixels on the remaining contours by $\mathcal{B}_T(\mathcal{X}^s)$ again, and we will refer to this set as the collection of extracted strong edge brim contours.

B. *A Three-Component Image Model*

We now proceed to generate an image $\mathcal{P} = \{(i, j, p_{i,j})\}$, $i, j = 0, 1, ..., M-1$, solely from $\mathcal{B}_T(\mathcal{X}^s)$, such that the difference image $\mathcal{X} \ominus \mathcal{P} \equiv \{(i, j, x_{i,j} - p_{i,j})\}$ has no strong edges represented by $\mathcal{B}_T(\mathcal{X}^s)$. We generate the image $\mathcal{P}$ based on the minimum information principle introduced in Section II (ref. eq. (II.3)). We will call the image $\mathcal{P}$ the *primary image* (or *primary component*) of the original image $\mathcal{X}$, since the strong edge information contained in image $\mathcal{P}$ is most important to human visual system.

Notice that, the stressed image $\mathcal{X}^s$, from which the set $\mathcal{B}_T(\mathcal{X}^s)$ is derived, can be considered as the output of a space-variant filter for the input $\mathcal{X}$. This filter is basically low-pass except at locations of strong edges. Thus the stressed image $\mathcal{X}^s$ contains the low-frequency component, or smooth component, and the strong edge information of the image $\mathcal{X}$. Therefore, the difference image $\mathcal{X} \ominus \mathcal{X}^s$ consists of the high-frequency component, but without the strong edge information. We define $\mathcal{T} = \mathcal{X} \ominus \mathcal{X}^s = \{(i, j, t_{i,j})\} = \{(i, j, x_{i,j} - x^s_{i,j})\}$ and $\mathcal{S} = \mathcal{X}^s \ominus \mathcal{P} = \{(i, j, s_{i,j})\} = \{(i, j, x^s_{i,j} - p_{i,j})\}$ corresponding to the texture and the smooth component, respectively. Thus, the image $\mathcal{X}$ can be expressed as

$$\mathcal{X} \equiv \mathcal{T} \oplus \mathcal{S} \oplus \mathcal{P} \equiv \{(i, j, t_{i,j} + s_{i,j} + p_{i,j})\}. \tag{IV.5}$$

In the sequel, this model will be referred to as the *three-component image model.*

In the next section, we present an experiment for this three-component image model and a comparison with the LGO.

V Experimental Results and Comparisons

Consider the test image shown in Fig. 10. This 256×256 image has a circular strong edge in the middle and textures or smooth areas in the rest of the image. A scanline (row 128) of the test image is plotted in Fig. 11. Note that around the center of the scanline we have a convex curve with the shape of a sine wave from 0 to π, then two strong edges, and up-down waveforms with linearly increasing frequency and amplitudes.

To get a picture of the edge location of the edge, we draw a grid in Fig. 12 with the intersections of vertical and horizontal lines indicating the pixel locations; the upper and lower edge brims are indicated by small square-dots (Fig. 12 is only the center portion of the image). Two typical edge segments are shown in Fig. 13. Note that the width of the edge is not limited to one pixel in contrast with the restriction in [7] [8] (see Subsection V.*B*).

A. An Example of the Three-Component Image Model

The stressed image is generated with the scheme developed in Section III. For simplicity, we assumed that $\lambda^2_{i,j} = \lambda^3_{i,j} \equiv \lambda_{i,j}$, for all (i,j). Thus $\lambda_{i,j}/\lambda^1_{i,j}$ is the only independent parameter. Therefore, instead of updating two parameters $\lambda^2_{i,j}$, $\lambda^3_{i,j}$, we may simply update $\lambda^1_{i,j}$ by the formula: $\lambda^1_{i,j} = \gamma(C^r_{i,j} + C^c_{i,j})$, ($\gamma$ is a constant), which is similar reciprocally to (III.16) [16].

The stressed image is shown in Fig. 14 (a), which is clearly stressed only at the strong edge of the test image. The computer time used on a SPARCstation-1 is about 1400 seconds. Edge brim points and the corresponding contours are extracted by the scheme introduced in Section IV, with $T = 512$, $T_c = 32$, $T_\ell = 16$ and $d = 3$. The resulting two contours are shown in Fig. 14 (b). An enlarged picture of these contours is shown in Fig. 15. These two contours correspond to the lower and upper edge brims. In Fig. 16, we superimpose Fig. 15 and Fig. 12 to demonstrate the effectiveness of our scheme in locating the strong edge.

To better understand the behavior of the stressed image, we have plotted scanline

128 for the test image (solid curve) and for the stressed image (dotted curve) in Fig. 17 (a), and with vertical lines indicating the locations of edge brim contours on this scanline in Fig. 17 (b). Note that the scanline associated with the stressed image is smooth except at the strong edge locations where it matches the edge, and that the locations of edge brim contours coincide with the upper and lower edge brims of the strong edge.

The primary image associated with the extracted strong edge is shown in Fig. 18. The other two components, namely, the smooth and texture components, are derived as in Subsection IV.*B*. Notice that the smooth and texture components are differences of two images, and thus may have negative intensity values. We have added constants to these two components in order to bring their intensity values in the displayable range, namely, $[0, 255]$. The resulting images of these two components are shown in Fig. 19. Note that the primary image contains the strong edge information, the smooth component provides the background slow intensity variations and the texture component contains all the textures.

Note that there are some traces of brighter pixels at around the location of the circular edge in the texture image. The reason for this phenomenon is basically the orientation anisotropy of the EMM problem since there are only two directions for the bars and the discrete structure of the problem (ref. Section III). This orientation anisotropy can be most clearly seen for edge segments with an orientation slightly tilted away from the vertical or the horizontal directions. However, since the primary image gives a perceptually good copy of the edge, these traces of brighter pixels can be treated as textures.

In what follows, we compare the performance of the strong edge extraction of the three-component image model with that of the LGO scheme.

B. The Laplacian-Gaussian Operator

The LGO scheme refers to an edge extraction scheme which is introduced by Marr and Hildreth [13]. In the following, we provide a brief description of the LGO and discuss some of its applications in the context of image coding [7], [8]. Then some

limitations of the LGO are illustrated by an example.

For simplicity, let us consider a real-valued function $x(t)$ of a real variable t. We define the edge locations of $x(t)$ by the zero-crossing locations of the second derivative of $x(t)$. To suppress weak edges, a Gaussian filter is used to smoothen $x(t)$ before second derivative zero-crossings are located. More precisely, a 1-D Gaussian kernel $g_\sigma(t)$ of standard deviation σ is convolved with $x(t)$:

$$y_\sigma(t) = x(t) \otimes g_\sigma(t) = \int_{-\infty}^{\infty} x(u) \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(t-u)^2}{2\sigma^2})\, du, \qquad \text{(V.1)}$$

and the points at which $y''_\sigma(t) = 0$ are defined as the edge locations. Note that the standard deviation σ of the Gaussian kernel controls the strengths of the extracted edges, i.e., larger values of σ correspond to higher edge strengths [13].

In the 2-D case, we consider a function $x(t,s)$ of two variables t and s. The edges with strengths above a threshold are defined as the points where the Laplacian of $y_\sigma(t,s)$, $\nabla^2 y_\sigma(t,s) \equiv \partial^2 y_\sigma/\partial t^2 + \partial^2 y_\sigma/\partial s^2$, vanishs, where $y_\sigma(t,s)$ is the convolution of $x(t,s)$ and a 2-D Gaussian kernel: $G_\sigma(r,\theta) = \exp(-r^2/(2\sigma^2))/(2\pi\sigma^2)$, where (r,θ) denotes the polar coordinates, with $r \in [0,\infty)$, $\theta \in [0,2\pi)$. Note that

$$\nabla^2 y_\sigma(t,s) = \nabla^2(x(t,s) \otimes G_\sigma(r,\theta)) = x(t,s) \otimes (\nabla^2 G_\sigma(r,\theta)), \qquad \text{(V.2)}$$

and

$$\nabla^2 G_\sigma(r,\theta) = \frac{\partial^2 G}{\partial r^2} + \frac{1}{r}\frac{\partial G}{\partial r} = \frac{1}{\pi\sigma^4}(\frac{r^2}{2\sigma^2} - 1)\exp(-\frac{r^2}{2\sigma^2}). \qquad \text{(V.3)}$$

To apply the LGO to digital images, the 2-D kernel $\nabla^2 G_\sigma(r,\theta)$ is discretized and convolved with the image $\{x_{i,j}\}$. We denote the result of this convolution by $\{y_{i,j}\}$ and locate its zero-crossings in the horizontal and vertical directions. For example, we say that there is a zero-crossing at $(i, j+0.5)$ if $y_{i,j}y_{i,j+1} \leq 0$. As in [7], we apply a gradient operator across each zero-crossing; if the absolute value of the output of the gradient operator is above a certain threshold, the zero-crossing is taken as an edge location. The resulting edge locations are transmitted along with the intensity values of their neighboring pixels for a complete description of the corresponding edges.

As indicated above, the standard deviation σ of the Gaussian filter controls the strengths of the extracted edges. Additionally, σ influences another important property of the resulting zero-crossings, namely, the accuracy of their locations with respect to the actual edges. Smaller values of σ give better accuracy, and larger values of σ result in poor accuracy [24]. We illustrate this situation with the following example.

Consider the 1-D signal $x(i)$ shown by the solid curve in Fig. 20 with an edge at $i = 179.5$. More specifically, $x(i) = 105$ for $i < 180$, $x(180) = -105$ and $x(i)$ linearly increases afterward. The dashed curve in Fig. 20 is a normalized version of $-y''_\sigma(i)$ (ref. eq. (V.1)) for $\sigma = 3$. Notice that $-y''_\sigma(i)$ has an isolated zero-crossing at $i = 179.5$. Thus, in this case the LGO locates the edge accurately; the intensity values of the left and right pixels of this zero-crossing are $x(179) = 105$ and $x(180) = -105$. These intensity values plus the location of the zero-crossing provide a complete description of the edge. Fig. 21 shows $-y''_\sigma(i)$ for $\sigma = 11$. In this case, the zero-crossing point is $i = 177.5$ which is clearly an inaccurate result.

Thus, on the one hand, small values of σ are needed to ensure accurate detection of the edge location and, on the other hand, large values of σ should be used for detection of "strong" edges. These requirements are conflicting and constitute a drawback of the LGO edge-detection scheme. In addition, the LGO scheme implicitly assumes the edge is *one* sample wide. This is in contrast with real-world situations (especially for images) where edges are typically several pixels wide.

In what follows we proceed to use the LGO scheme to detect the strong edge of the test image presented in Fig. 10.

C. *An Example of Locating Strong Edges with the LGO*

We now present the results of the LGO-based edge detection for the test image with five different values of σ namely, $1, 2, 3, 4$ and 5 pixel-widths. In each case, the threshold for the output of the gradient operator is determined empirically as the largest possible to keep the zero-crossings for the strong edge and to suppress, as much as possible, the zero-crossings corresponding to the textures. The results

are summarized in Fig. 22 in the form of images where a bright point at pixel (i, j) indicates a zero-crossing at $(i, j + 0.5)$ and/or $(i + 0.5, j)$. The computer times, in seconds (s), used on a SPARCstation-1 to do the convolution (ref. (V.8)) are also included.

As shown in Fig. 22, the result for $\sigma = 5$ pixel-widths detects the strong edge only, while all the other results contain zero-crossings corresponding to the textures, as well. The enlarged picture of the zero-crossing points for $\sigma = 5$ pixel-widths is shown in Fig. 23 (compared with the actual edge location in Fig. 24), where the positions of the zero-crossings are one half pixel-width to the right or below the square-dots (ref. Subsection V.*B*). Thus, the set of the edge brim points consists of the pixels on two circles. One circle is shown in Fig. 23, and the other one (not shown) is one pixel away to the right and below of the first one. The primary image of these edge brim points is generated and shown in Fig. 25. Notice that simply there is no edge in Fig. 25, as compared with Fig. 18 obtained using the three-component image model.

Therefore, the representation of the edge in the test image generated with the LGO gives location error, and, more seriously, fails to provide the intensity variation of the edge, while the strong edge extraction scheme associated with the three-component image model provides a superior edge extraction performance.

VI Summary and Conclusions

We have studied several psychovisual evidences of human visual system and attempted to describe them mathematically. In particular, the strong edges and the areas of smooth intensity variations are carefully studied and their effect as well as their interaction in formation of the perception are described in terms of simple minimization problems. An algorithm, based on a space-variant low-pass filtering operation, for the generation of the stressed image is developed from which the strong edges can be easily extracted. This has led to introducing a three-component image model based on the (i) strong edges, (ii) areas of smooth intensity variations and

(iii) textures - the three components which apparently play different roles in the formation of the perception.

In a specific comparison, we have shown the superiority of the strong edge extraction scheme developed here over the LGO-based edge extraction method. The three-component model was motivated by the need for subjective-based criteria for the design of low bit rate image coding systems (such as adaptive DCT and subband coding) in which the three components of our model are treated based on their relative perceptual importance. The preliminary results [11] [18] [19] are quite encouraging. In addition, we are currently working on applying the three-component model in the complex domain for representing planar curves - a problem which has applications in computer vision [25].

Appendix A EMM – a Space-Variant Filter

In this Appendix, we show that the EMM problem can be described as a space-variant filtering process. The input-output relation of this filter is determined by the minimization problem (III.10), and will be discussed in the following.

For simplicity, we consider the continuous form of (III.10). Let $x(u,v)$ be a continuous function of two variables u, v defined on a simply-connected closed domain [26] $G \subset \mathbb{R}^2$ with a piecewise smooth boundary C. We want to solve the following minimization problem,

$$\min_y J_x(y) \equiv \int\int_G \{\lambda_1(u,v)[x(u,v)-y(u,v)]^2 + \lambda_2(u,v)[y_{uu}(u,v)]^2 + \lambda_3(u,v)[y_{vv}(u,v)]^2\}\, dudv, \qquad \text{(A.1)}$$

where $\lambda_i(u,v) \geq 0$, $i = 1,2,3$, are functions of u, v on G having continuous second-order partial derivatives, $y(u,v)$ is an admissible function defined on G having continuous fourth-order partial derivatives with y, $\partial y/\partial u$, $\partial y/\partial v$ vanishing on the boundary C and $y_{uu}(u,v)$ and $y_{vv}(u,v)$ are second-order partial derivatives of y with respect to u and v, respectively. In (A.1), $J_x(y)$ is used to denote the objective function. All admissible functions form a linear space, denoted by $\mathcal{D}(G)$ endowed by a norm $||y||$

defined as

$$||y|| \equiv \max_{(u,v)\in G} |y(u,v)| + \max_{(u,v)\in G} |y_{uu}(u,v)| + \max_{(u,v)\in G} |y_{vv}(u,v)|. \tag{A.2}$$

Thus $\mathcal{D}(G)$ becomes a normed linear space [26].

We say J_x has a minimum at $y^* \in \mathcal{D}(G)$ if and only if $J(y) - J(y^*) \geq 0$ for all $y \in \mathcal{D}(G)$. Then, the necessary and sufficient condition for y^* to be a minimum of J_x is that [16]

$$\int\int_G \{\lambda_1(y^* - x)h + \lambda_2 y^*_{uu} h_{uu} + \lambda_3 y^*_{vv} h_{vv}\}\, dudv = 0, \tag{A.3}$$

for all $h \in \mathcal{D}(G)$. By the Green's theorem [27], (A.3) reduces to [16]

$$\lambda_1(u,v)y^*(u,v) + (\lambda_2 y^*_{uu})_{uu}(u,v) + (\lambda_3 y^*_{vv})_{vv}(u,v) = \lambda_1(u,v)x(u,v), \tag{A.4}$$

for $(u,v) \in G$. Equation (A.4) is known as an *Euler's equation* [26].

Now we show that at all $(u,v) \in G$ satisfying $\lambda_1(u,v) > 0$, the solution of the Euler's equation is unique. Let y_1 and y_2 be two solutions of the Euler's equation, then $y_1 - y_2$ is a solution of the Euler's equation with $x \equiv 0$, and thus $y_1 - y_2$ is a minimum point of J_0. Then $y_1(u,v) - y_2(u,v) = 0$, for every (u,v) such that $\lambda_1(u,v) > 0$, for otherwise $J_0(\epsilon(y_1 - y_2)) < J_0(y_1 - y_2)$ for any $0 < \epsilon < 1$.

Now we investigate the frequency domain behavior of the input-output relation governed by the above Euler's equation. Let (u_1, v_1) be an interior point in G, and A be an open neighborhood containing the point (u_1, v_1). To further simplify the problem, we assume that the parameter functions, $\lambda_1 > 0$, λ_2, λ_3, are constants on A (certainly when A is a small neighborhood, this is a reasonable assumption). Then the above Euler's equation becomes

$$\lambda_1(u_1,v_1)y^*(u,v) + \lambda_2(u_1,v_1)y^*_{uuuu}(u,v) + \lambda_3(u_1,v_1)y^*_{vvvv}(u,v) = \lambda_1(u_1,v_1)x(u,v) \tag{A.5}$$

on A. Equation (A.5) is a linear partial differential equation which describes the input-output relationship of a linear system with input $x(u,v)$, output $y^*(u,v)$ and frequency response $H(w_u, w_v)$ given by

$$H(w_u, w_v) = \frac{1}{1 + \frac{\lambda_2(u_1,v_1)}{\lambda_1(u_1,v_1)}\omega_u^4 + \frac{\lambda_3(u_1,v_1)}{\lambda_1(u_1,v_1)}\omega_v^4} \tag{A.6}$$

which clearly describes a 2-D low-pass filtering operation. The cutoff frequencies of $H(w_u, w_v)$ are controlled by the ratios $\lambda_2(u_1, v_1)/\lambda_1(u_1, v_1)$ and $\lambda_3(u_1, v_1)/\lambda_1(u_1, v_1)$. Larger values of $\lambda_1(u_1, v_1)$ give rise to higher cutoff frequencies in u and v, while larger values of $\lambda_2(u_1, v_1)$ and $\lambda_3(u_1, v_1)$ lead to lower cutoff frequencies in u and v directions, respectively.

Note that we have assumed that all admissible functions and their first-order partial derivatives vanish on the boundary C, while in the original case (III.10) we do not have any such constraint. This difficulty can be avoided by solving a slightly modified problem in which the boundary constraints are removed. The details of this are omitted here but can be found in [16].

Appendix B Contour Prediction

The first-order prediction works in the following way. After the basic algorithm (with back-tracing) stops, we locate an additional pixel in the neighborhood of the end-pixel of the traced contour in the direction determined by the last two pixels. More precisely, for the traced contour, $(i^0, j^0, x^s_{i^0,j^0})$, $(i^1, j^1, x^s_{i^1,j^1})$, ..., $(i^{\ell-1}, j^{\ell-1}, x^s_{i^{\ell-1},j^{\ell-1}})$, the additional pixel, $(i^\ell, j^\ell, x^s_{i^\ell,j^\ell})$, is defined as:

$$i^\ell = i^{\ell-1} + (i^{\ell-1} - i^{\ell-2}); \;\; j^\ell = j^{\ell-1} + (j^{\ell-1} - j^{\ell-2}). \tag{B.1}$$

If this new contour, $(i^1, j^1, x^s_{i^1,j^1})$, ..., $(i^\ell, j^\ell, x^s_{i^\ell,j^\ell})$, satisfies condition (IV.3), and by the above basic search procedure we can locate in the neighborhood of $(i^\ell, j^\ell, x^s_{i^\ell,j^\ell})$ a pixel other than $(i^{\ell-1}, j^{\ell-1}, x^s_{i^{\ell-1},j^{\ell-1}})$ in $\mathcal{B}_T(\mathcal{X}^s)$, we move the search center to this new pixel in $\mathcal{B}_T(\mathcal{X}^s)$; otherwise, the contour terminates at $(i^{\ell-1}, j^{\ell-1}, x^s_{i^{\ell-1},j^{\ell-1}})$. The first-order prediction is illustrated in Fig. 26, where the end pixel and its previous pixel are indicated by a circle and a cross, respectively, and the predicted pixel is indicated by a square box.

If the first-order prediction fails, a second-order prediction can be used. Using the same notation as before, we define pixels $(i^\ell, j^\ell, x^s_{i^\ell,j^\ell})$ and $(i^{\ell+1}, j^{\ell+1}, x^s_{i^{\ell+1},j^{\ell+1}})$

such that

$$i^{\ell} = i^{\ell-1} + (i^{\ell-2} - i^{\ell-3}); \;\; j^{\ell} = j^{\ell-1} + (j^{\ell-2} - j^{\ell-3});$$
$$i^{\ell+1} = i^{\ell} + (i^{\ell-1} - i^{\ell-2}); \;\; j^{\ell+1} = j^{\ell} + (j^{\ell-1} - j^{\ell-2}). \tag{B.2}$$

If this extended contour does not satisfy condition (IV.3), we stop and end the contour at pixel $(i^{\ell-1}, j^{\ell-1}, x^s_{i^{\ell-1},j^{\ell-1}})$; otherwise, we continue as follows. If pixel $(i^{\ell+1}, j^{\ell+1}, x^s_{i^{\ell+1},j^{\ell+1}})$ is in $\mathcal{B}_T(\mathcal{X}^s)$, we move the search center to this pixel and continue; if not, we perform the basic search step at $(i^{\ell+1}, j^{\ell+1}, x^s_{i^{\ell+1},j^{\ell+1}})$, and if we successfully locate a pixel in $\mathcal{B}_T(\mathcal{X}^s)$, we then move the search center to this new pixel and continue. We illustrate the second-order prediction in Fig. 27 with a similar legend as in Fig. 26. This procedure can be extended to higher order predictions in a similar manner.

References

[1] W. H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images ," *IEEE Trans. Commun.*, vol. COM-25, pp. 1285-1292, Nov. 1977.

[2] N. M. Nasrabadi and R. A. King, "Image coding using vector quantization: a review," *IEEE Trans. Commun.*, vol. COM-36, pp. 957-971, Aug. 1988.

[3] J. W. Woods and S. D. O'Neil, "Subband Coding of images," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-34, pp. 1105-1115, Nov. 1986.

[4] W. A. Pearlman, "Adaptive cosine transform image coding with constant block distortion," *IEEE Trans. Commun.*, vol. COM-38, pp. 698-703, May. 1990.

[5] M. Kunt, A. Ikonomopoulos, and M. Kocher, "Second-generation image-coding techniques," *Proceedings of the IEEE,* vol. 73, No. 4, pp. 549-574, April 1985.

[6] M. Kunt, M. Benard, and R. Leonardi, "Recent Results in High-Compression Image Coding," *IEEE Trans. Circuit and Systems*, vol. CAS-34, No. 11, pp. 1306-1336, Nov. 1987.

[7] S. Carlesson, "Sketched based coding of grey level images," *Signal Processing*, Vol. 15, No. 1, pp. 57–83, July 1988.

[8] S. Carlesson, C. Reillo, and L. H. Zetterberg, "Sketched based representation of grey value and motion information," in *From Pixels to Features* by J. C. Simon (ed.) Elsevier Science Publishers B.V. (North-Holland), 1989.

[9] H. Helmholtz, *Treatise on Physiological Optics*, edited by J. Southall, vol. III, *The Perceptions of Vision*, the Optical Society of America, George Bonta Publishing Company, Menasha, Wisconsin, 1925.

[10] T. Cornsweet, *Visual Perception*, Academic Press, New York and London, 1970.

[11] X. Ran, "Low Bit Rate Image Coding Using a Three-Component Image Model for Human Visual Perception," in preparation.

[12] W. Grimson, "Surface consistency constraints in vision," *Computer Vision, Graphics, and Image Processing* 24, pp. 28-51, 1983.

[13] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond.* B 207, pp. 187–217, 1980.

[14] D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, Menlo Park, California, etc., 1984.

[15] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1985.

[16] X. Ran, "A Three-Component Image Model for Human Visual Perception and Its Application in Image Coding and Processing," Ph.D. dissertation, University of Maryland, College Park, MD.

[17] P. Chou and N. Pagano, *Elasticity, Tensor, Dyadic, and Engineering Approaches*, D. Van Nostrand Company, Inc., Princeton, New Jersdy, Toronto, London, 1967.

[18] X. Ran and N. Farvardin,"A Three-Component Image Model for Coding with a Subjective Criterion," presented in the Seventh Workshop on Multidimensional Signal Processing, Lake Placid, NY, Sept. 1991. p.7.2.

[19] X. Ran and N. Farvardin,"Adaptive DCT Image Coding Based on a Three-Component Image Model," accepted to be presented in 1992 ICASSP, San Francisco, CA, March 1992.

[20] T. S. Huang, "Coding of Two-Tone Images," *IEEE Trans. on Communications*, Vol. COM-25, No. 11, pp. 1406-1424, November, 1977.

[21] D. N. Graham, "Image Transmission by Two-Dimensional Contour Coding," *Proceedings of IEEE*, Vol. 55, No. 3, pp. 336-346, March, 1967.

[22] H. Freeman, "On the Encoding of Arbitrary Geometric Configuration." *IRE Trans. Electron. Comput.*, EC-10, pp. 260-268, June, 1961.

[23] J. K. Yan and D. J. Sakrison, "Encoding of Images Based on a Two-Component Source Model," *IEEE Trans. on Communications*, Vol. COM-25, No. 11, pp. 1315-1322, November, 1977.

[24] V. S. Nalwa and T. O. Binford, "On Detecting Edges," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. PAMI-8, No. 6, pp. 699–714, Nov. 1986.

[25] H. Asada and M. Brady, "The Curvature Primal Sketch," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. PAMI-8, No. 1, pp. 2–14, Jan. 1986.

[26] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, translated and edited by R.A. Silverman, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963.

[27] W. Kaplan, *Advanced Calculus*, Addison-Wesley Publishing Company, Inc., MA., 1952.

[28] B. O'Neill, *Semi-Riemannian Geometry, with applications to relativity*, Academic Press, New York, London, etc., 1983.

[29] D. M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York and London, 1971.

[30] O. Axelsson, *Finite Element Solution of Boundary Value Problems : Theory and Computation*, Academic Press, Orlando, Fla., 1984.

(a) (b)

Figure 1: (a) Vertical bar for the left eye; (b) horizontal bar for the right eye.

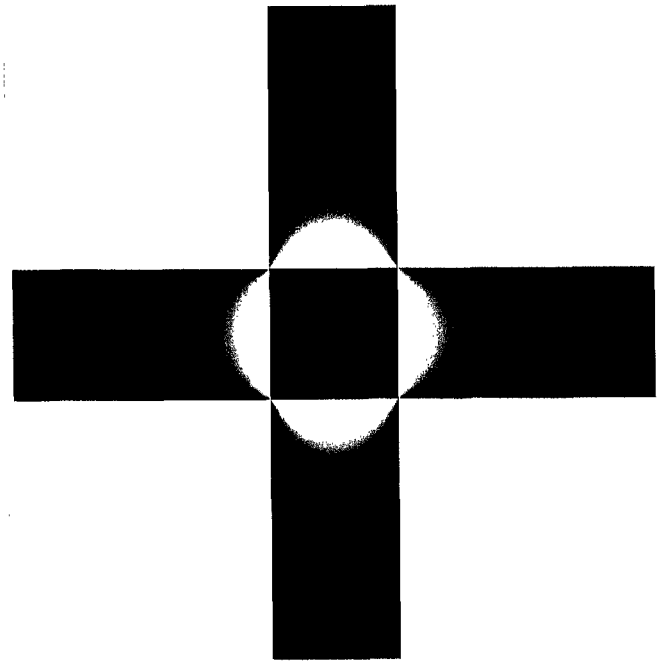
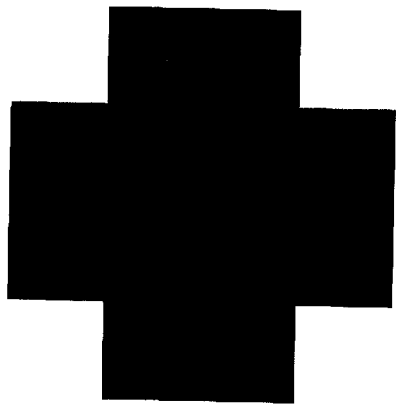


Figure 2: Binocular perception of the two images in Fig. 1.



(a) (b)

Figure 3: (a) A cross; (b) a network of lines.

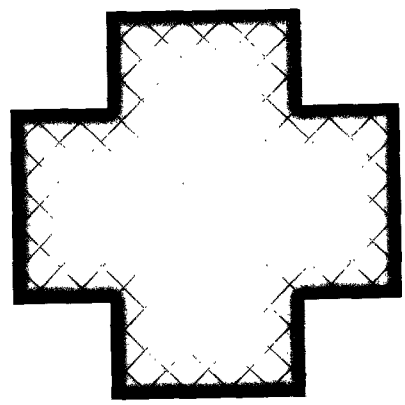
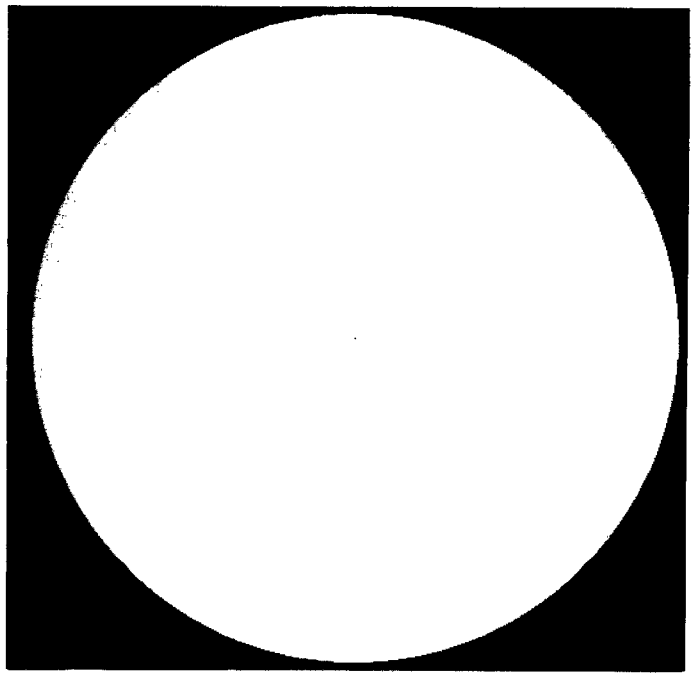
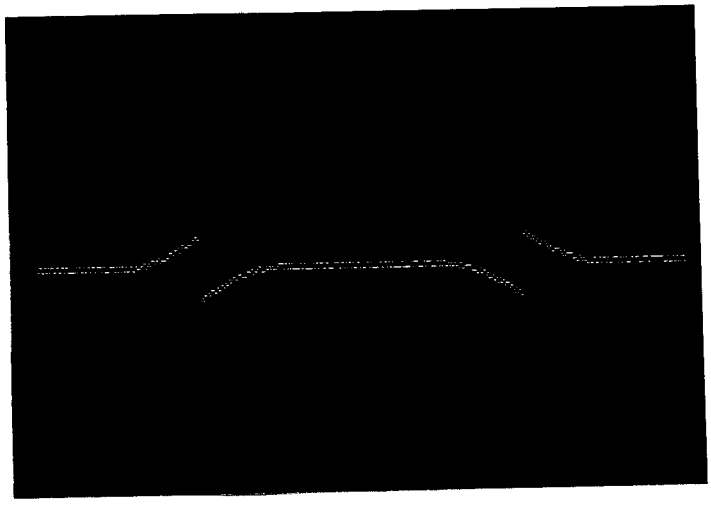


Figure 4: Binocular perception of the two images in Fig. 3.



(a)



(b)

Figure 5: (a) Two concentric disk image; (b) intensity values along the diameter of the disks.

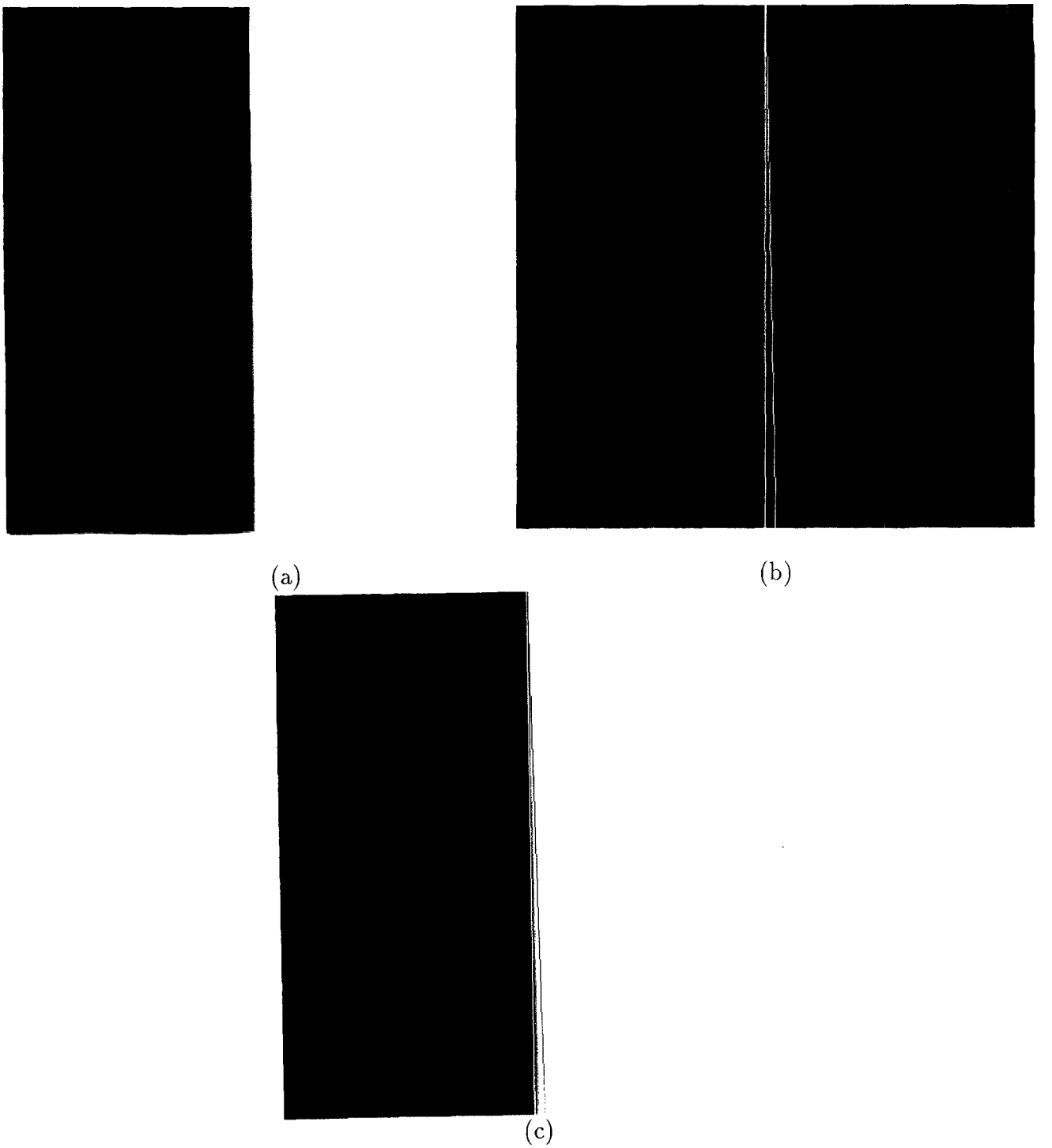
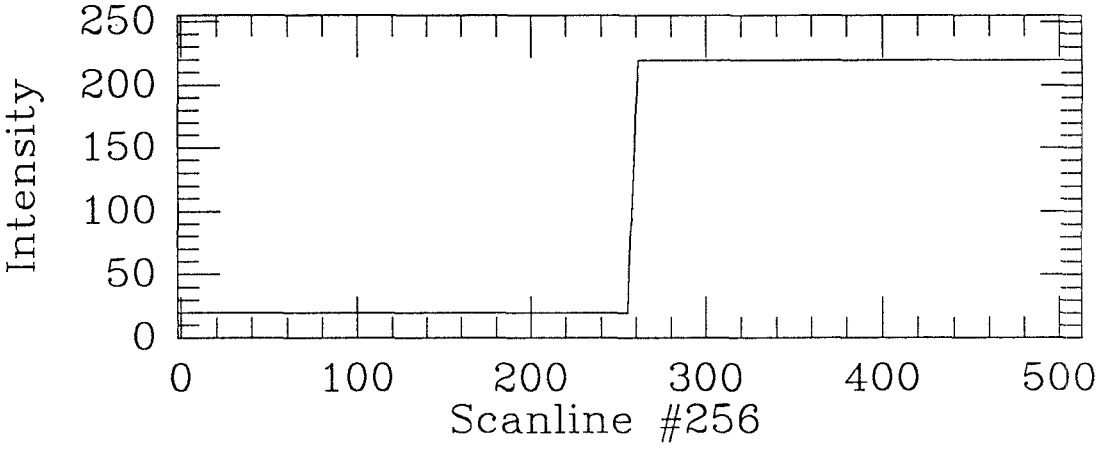
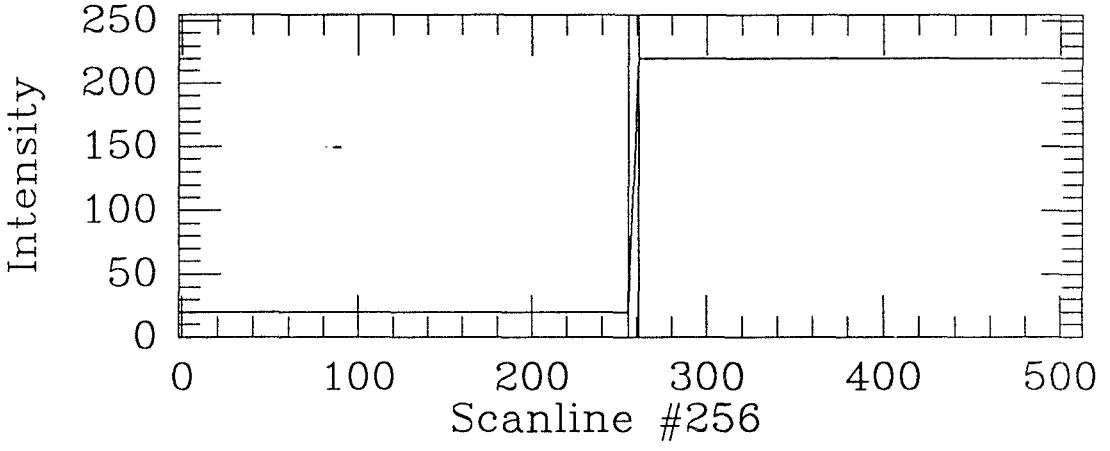


Figure 6: (a) An edge of variant width; (b) lower and upper edge brims; (c) image (a) superposed with image (b).



(a)



(b)

Figure 7: (a) Scanline #256 of image (a) in Fig. 6; (b) with indications of edge brims.

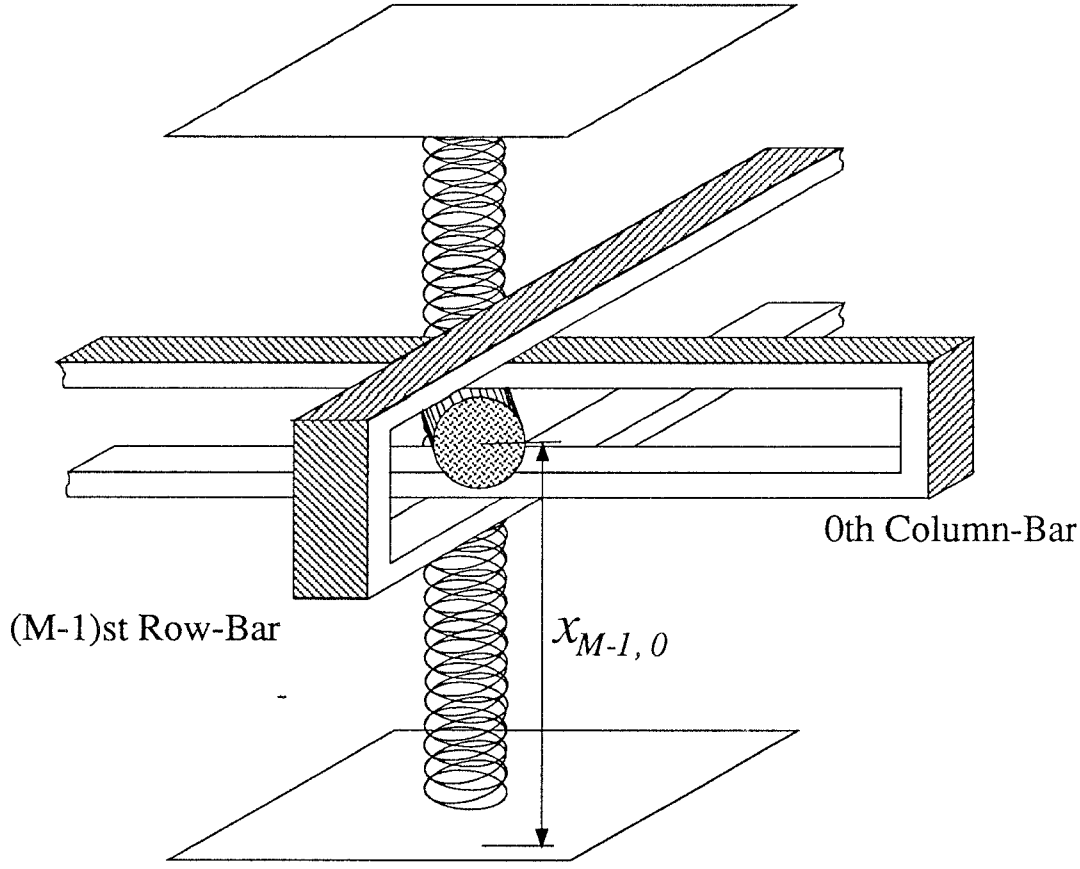


Figure 8: Mechanical structure of EMM at pixel $(M-1,0)$.

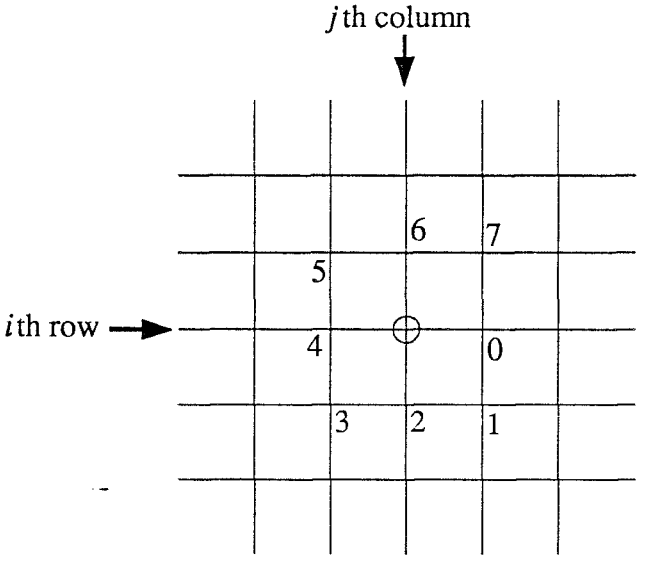


Figure 9: The eight neighboring pixels.

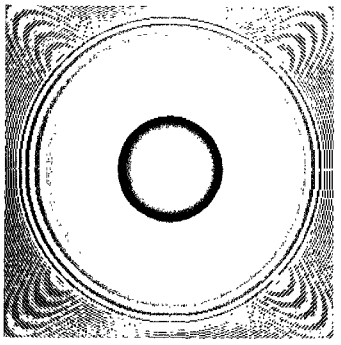


Figure 10: Test image.

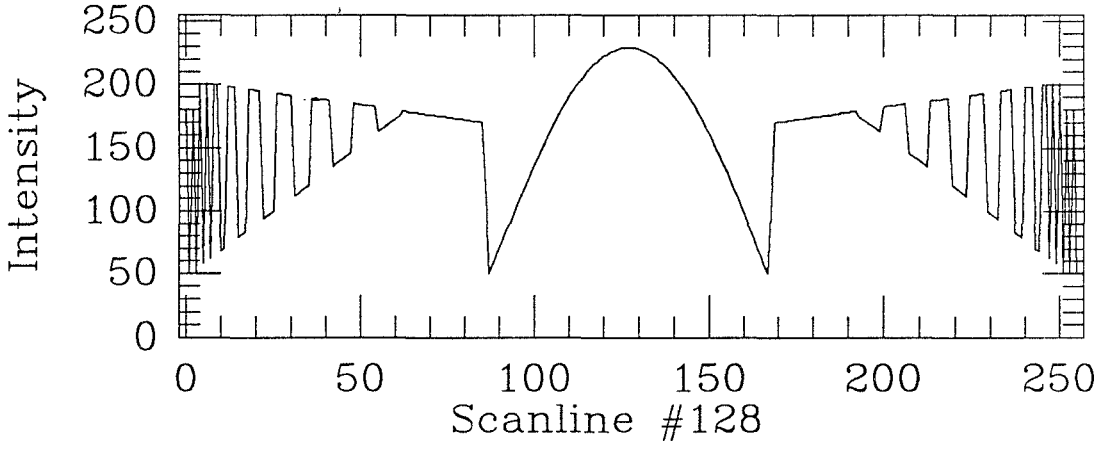


Figure 11: Scanline #128 of the test image.

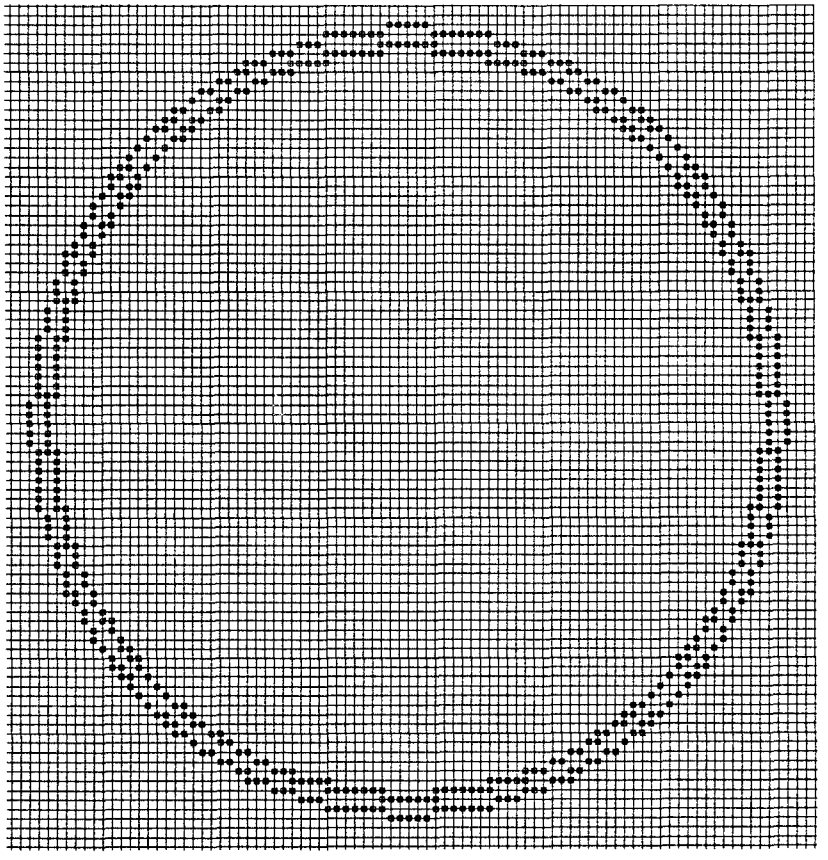


Figure 12: Enlarged picture of the edge-brim locations in the test image.

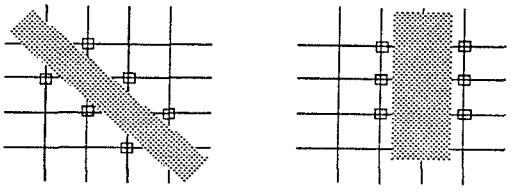
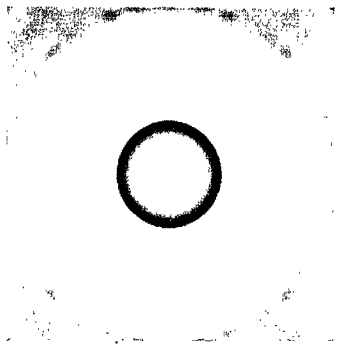
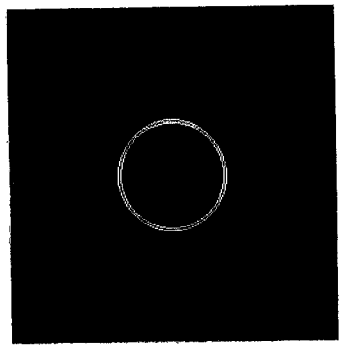


Figure 13: Typical segments of the edge in the test image.



(a) The stressed image



(b) The picture of edge brim contours

Figure 14: The stressed image and the picture of edge brim contours extracted for the test image.

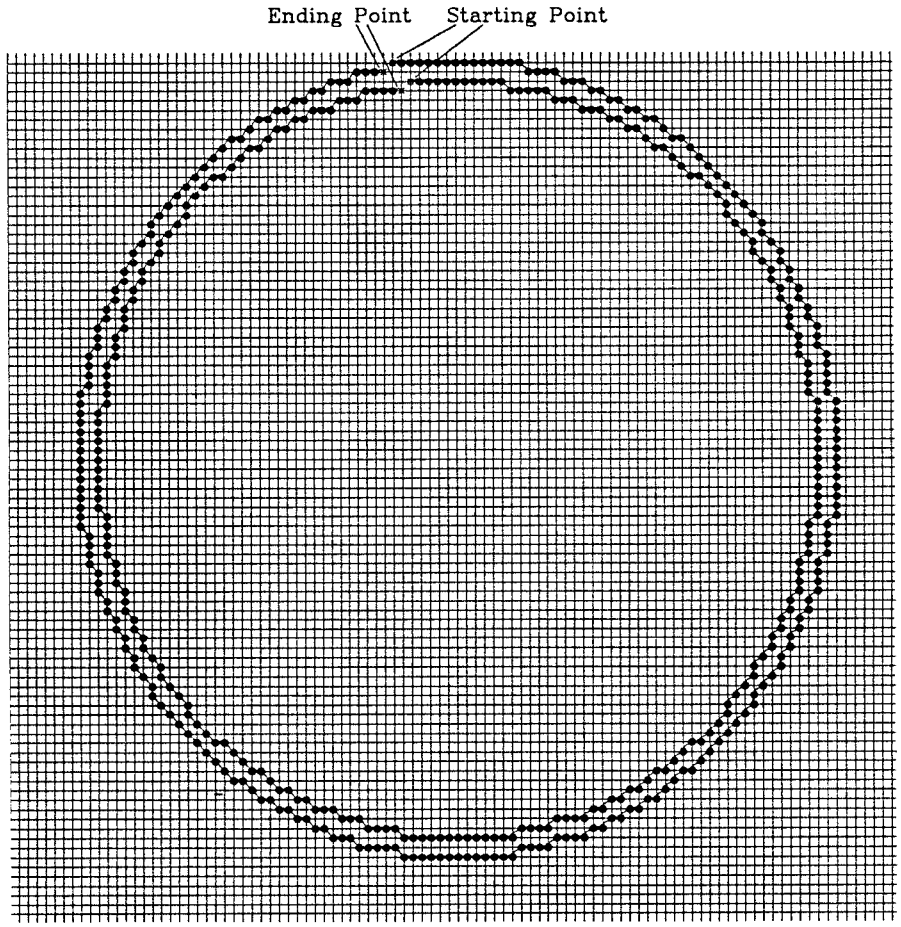


Figure 15: Enlarged picture of the contours extracted from the stressed image shown in Fig. 14 (a).

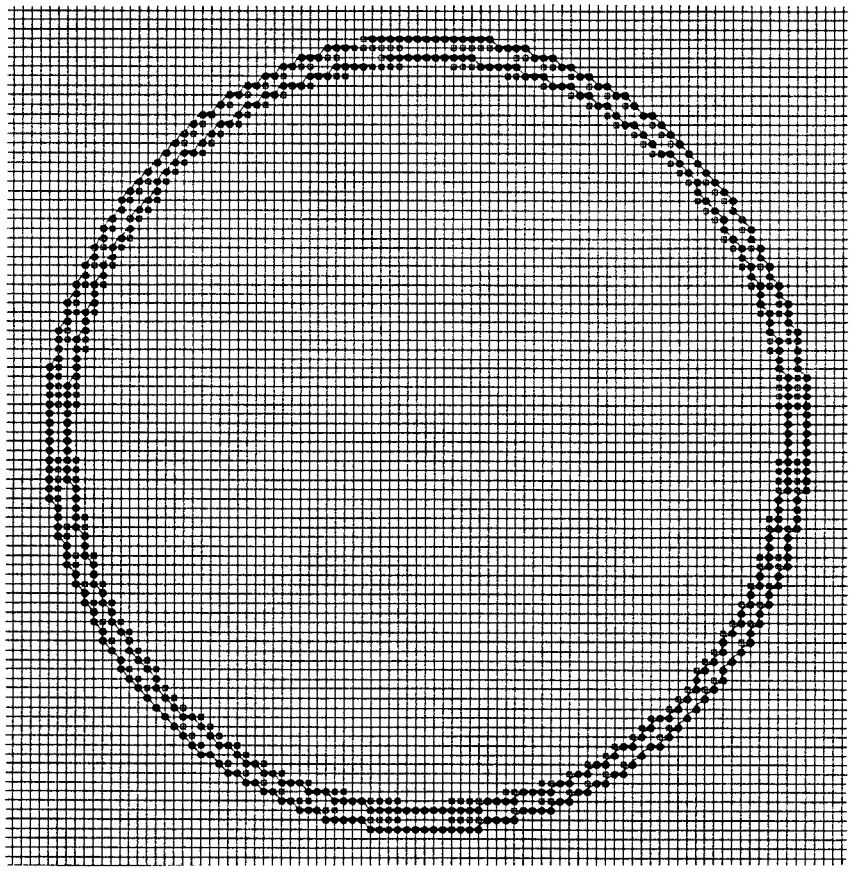
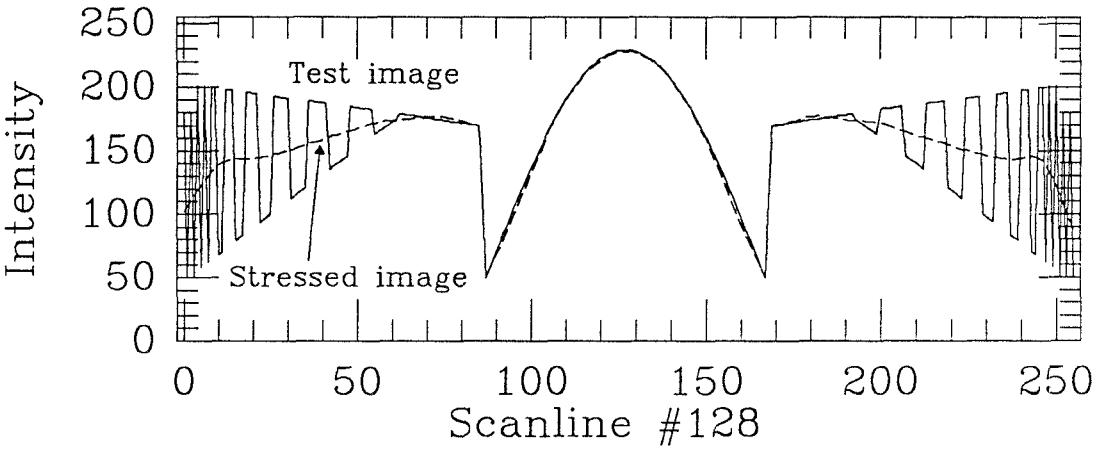
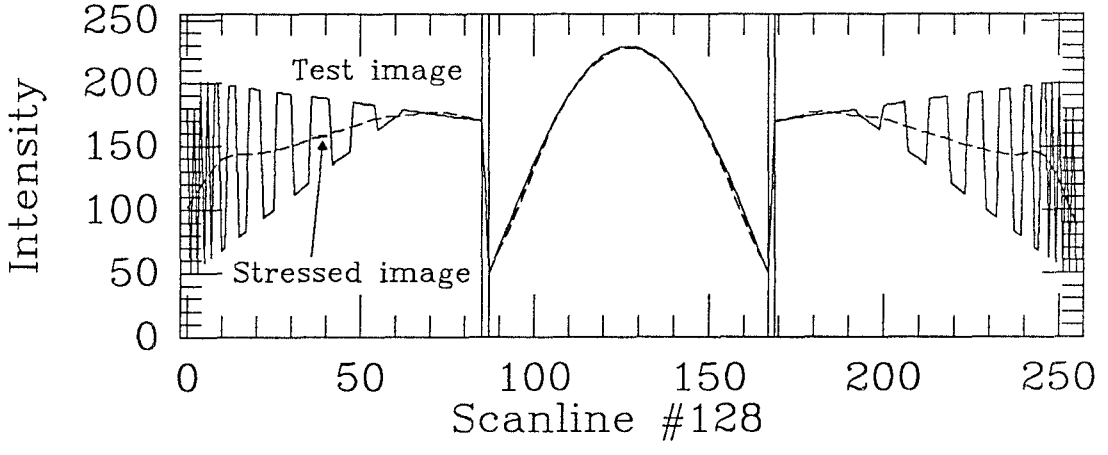


Figure 16: The superimposition of Fig. 15 and Fig. 12.



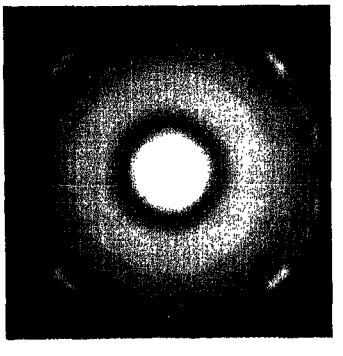
(a) Intensity values along scanline #128 for the test image and the stressed image.



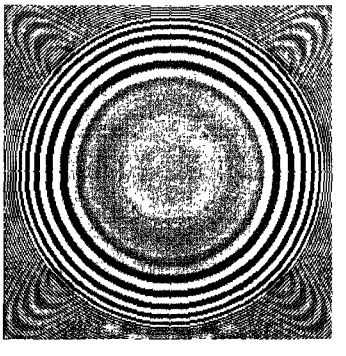
(b) Vertical lines indicating location of edge brim points on this scanline.

Figure 17: Scanline #128 of the test image and the stressed image.

Figure 18: The primary image for the test image.



(a) Smooth component



(b) Texture component

Figure 19: The smooth and texture components.

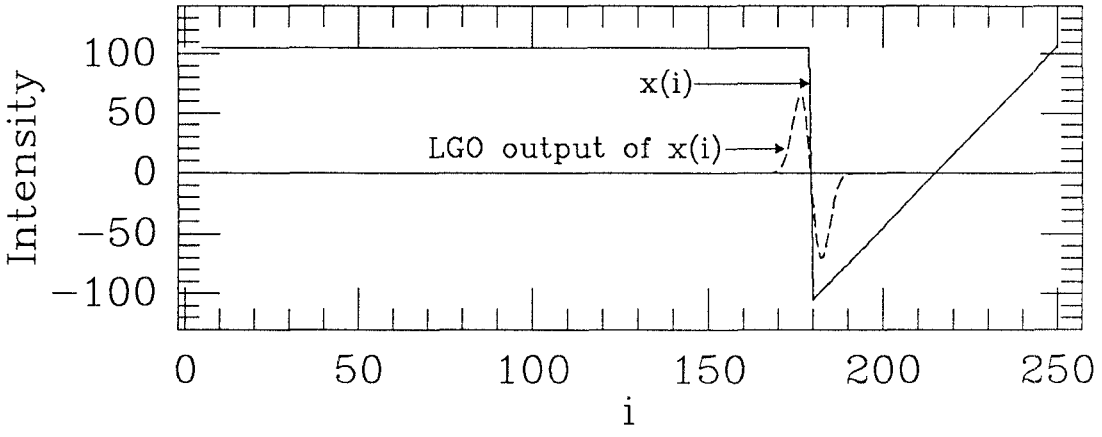


Figure 20: An example for the locational accuracy of LGO, $\sigma = 3$.

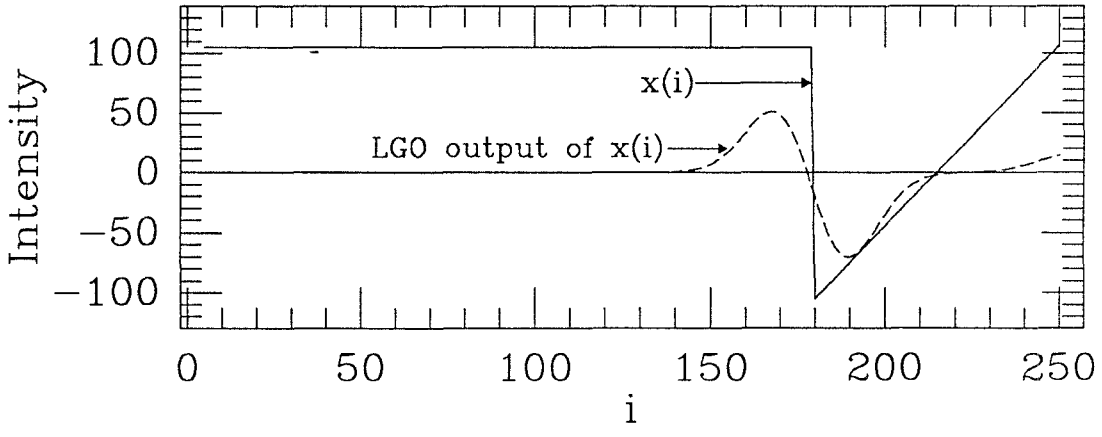
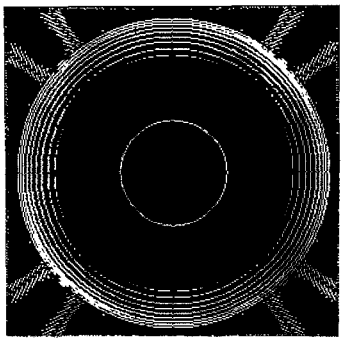
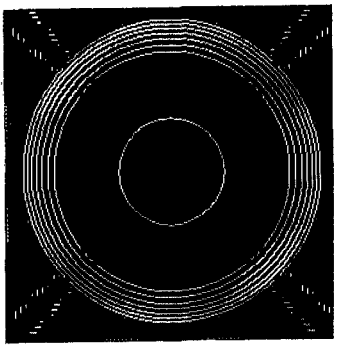


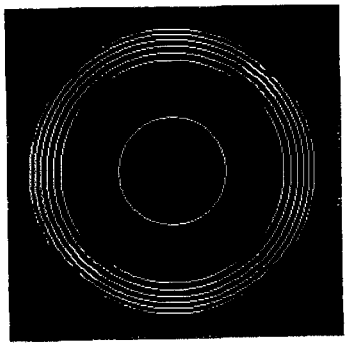
Figure 21: An example for the locational accuracy of LGO, $\sigma = 11$.



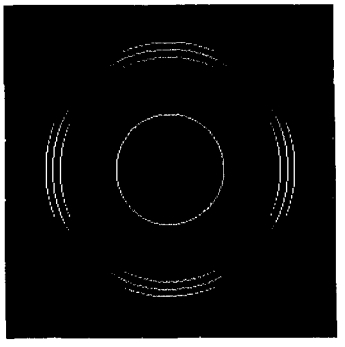
$\sigma = 1$ pixel-width (33 s)



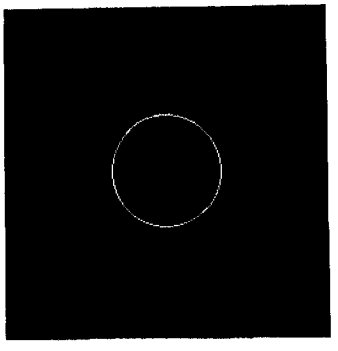
$\sigma = 2$ pixel-width (125 s)



$\sigma = 3$ pixel-width (345 s)



$\sigma = 4$ pixel-width (603 s)



$\sigma = 5$ pixel-width (1019 s)

Figure 22: Zero-crossing results with the LGO scheme for the test image.

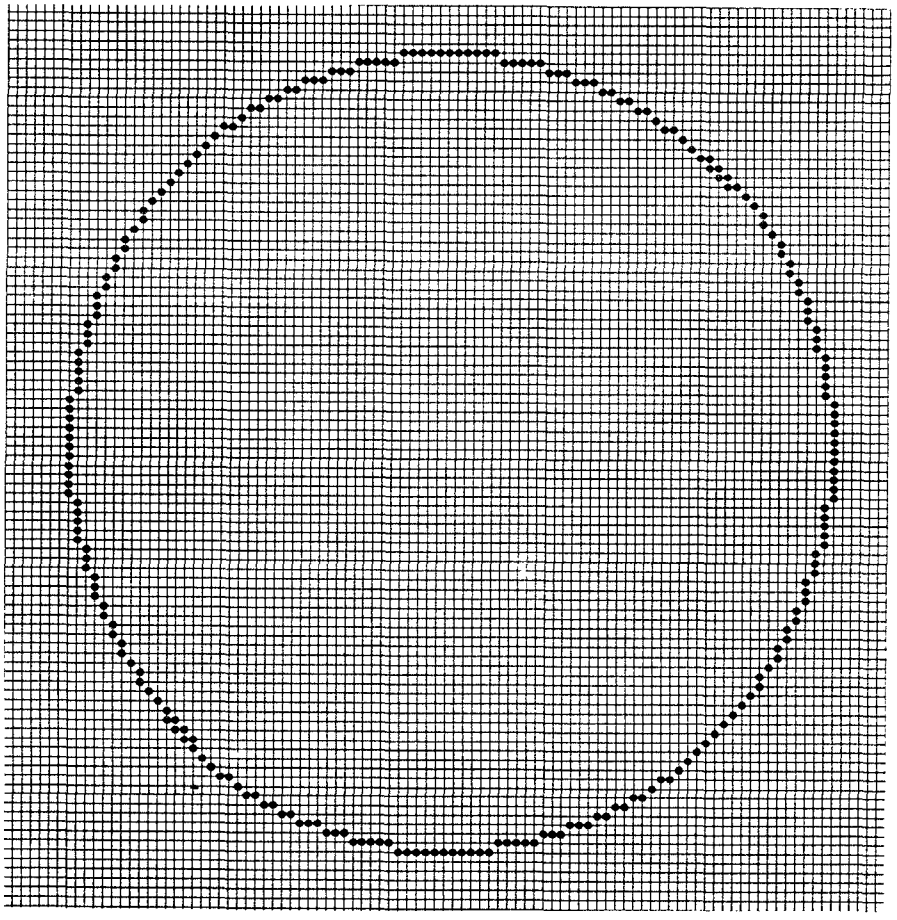


Figure 23: Enlarged picture of the zero-crossings for $\sigma = 5$ pixel-widths.

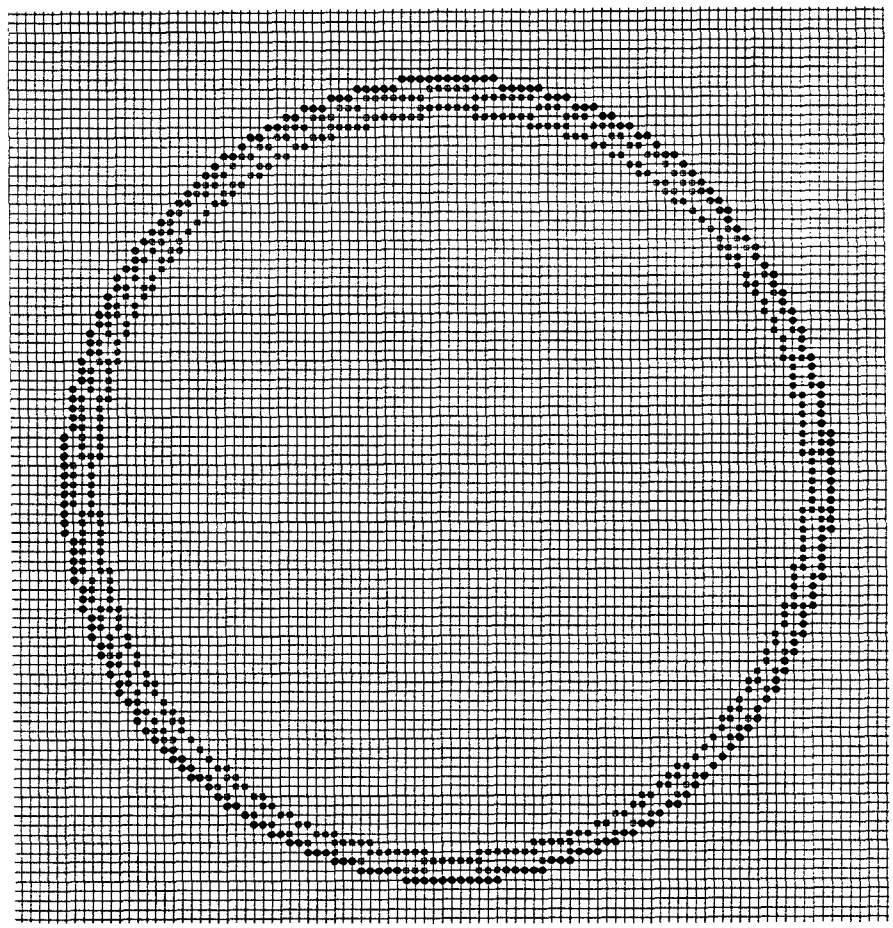


Figure 24: The superimposition of Fig. 23 and Fig. 12.

Figure 25: The primary image of the neighboring pixels of zero-crossings shown in Fig. 23.

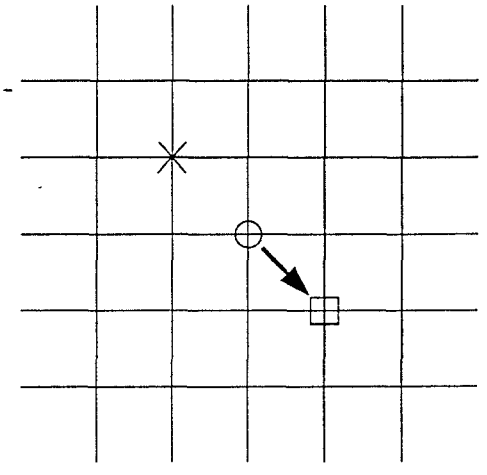


Figure 26: The first-order prediction.

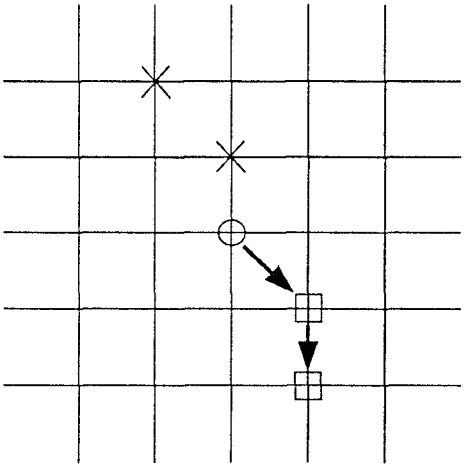


Figure 27: The second-order prediction.