A Unified Approach to Boundary Perception: Edges, Textures, and Illusory Contours

B. S. Manjunath, *Member, IEEE* and Rama Chellappa, *Fellow, IEEE*

Abstract—**This paper presents a unified approach to boundary perception. The model consists of a multistage system which extracts and groups salient features in the image at different spatial scales (or frequencies). In the first stage, a Gabor wavelet decomposition provides a representation of the image which is orientation selective and has optimal localization properties in space and frequency. This decomposition is useful in detecting significant features such as step and line edges at different scales and orientations in the image. Following the wavelet transformation, local competitive interactions are introduced which help in reducing the effects of noise and changes in illumination. Interscale interactions help in localizing the line ends and corners, and play a crucial role in boundary perception. The final stage groups similar features, aiding in boundary completion. This approach is consistent with some of the known neurophysiological observations regarding biological visual information processing, as the different stages can be identified with processing by simple, complex, and hypercomplex cells in the visual cortex of mammals. Experimental results are provided to indicate the performance of this model in detecting boundaries (both real and illusory) in real and synthetic images.**

I. Introduction

We suggest a simple biologically motivated approach to detecting image boundaries. Biological vision systems, especially those of mammals and in particular human's, are extremely adept at processing the vast amount of intensity data projecting from the three-dimensional external world on to the two-dimensional retina. Recent research in psychophysics and neurophysiology has begun to shed light on some of the basic mechanisms that are used in interpreting this information. The initial stages of this visual processing are very important in this respect as they detect and group various types of salient features, and transform the intensity information to a more suitable representation convenient for further processing. These stages are responsible for preliminary processing of stereo, texture, and motion, which further aid in performing one of the fundamental tasks in image understanding, namely boundary perception and scene segmentation.

Manuscript received January 21, 1991; revised February 26, 1992. This work was supported in part by the Defense Advanced Research Projects (ARPA Order 6989) and the U.S. Army Topographic Engineering Center under Contract DACA 76-89-C-0019. This work was carried out by the authors at the Signal and Image Processing Institute, University of Southern California.

B. S. Manjunath is with the Department of Electrical and Computer Engineering, Center for Information Processing Research, University of California, Santa Barbara, CA 93106.

R. Chellappa is with the Department of Electrical Engineering, Center for Automation Research, University of Maryland, College Park, MD 20742.

IEEE Log Number 9200937.

In the three-dimensional world the objects are separated from the background (as well as other objects) by depth discontinuities, which usually manifest as intensity discontinuities in two-dimensional images. Intensity changes also result from occlusion of objects, sharp changes in surface orientation, changes in reflectance properties, or illumination. As these intensity changes are a rich source of information, detecting them is an important problem both in computer vision as well as in human vision. Among the most commonly used edge detection algorithms are the zero crossings of the Laplacian of the Gaussian [1] and Canny's edge detector [2]. Textures form another important class of natural scenes and like intensity edges provide useful information regarding shape and motion. In vision, many models ranging from stochastic to structural ones have been used in analyzing textures. Random field models have been especially successful in classifying and segmenting scenes consisting of several natural textures, and parallel relaxation algorithms have been developed for this purpose [3]. Computational models for human texture perception have also been extensively studied [4]–[7]. Though intensity edges and textures are fundamental to image understanding, only recently some work has been done in integrating the detection of these features. In [8], a composite model is proposed for detecting both the intensity as well as texture edges. A random field model is proposed for a general boundary detection scheme in [9], where the problem of segmentation is formulated as an optimization process and relaxation algorithms are used to obtain the segmentation.

In addition to intensity edges and textures, human vision system can perceive object boundaries where none physically exist, giving rise to what are generally referred to as illusory contours. This perception is a consequence of the mechanisms involved in interpreting incomplete information such as those due to occlusion, which is very common in the real world. The mechanisms themselves are not well understood and surprisingly not much attention has been given to this problem in computer vision research. The problem of understanding the perception of such contours is complicated because it is difficult to separate the role of high level (or contextual) knowledge from the low level mechanisms which actually complete the boundary. Our discussion in this paper regarding such contours, hence, is limited to very simple examples such as the ones induced by line terminations (see Fig. 10).

A schematic diagram of our model is shown in Fig. 1. The input image is first processed through a bank of orientation selective bandpass filters at various spatial frequencies. Our choice of Gabor functions to model these filters has been

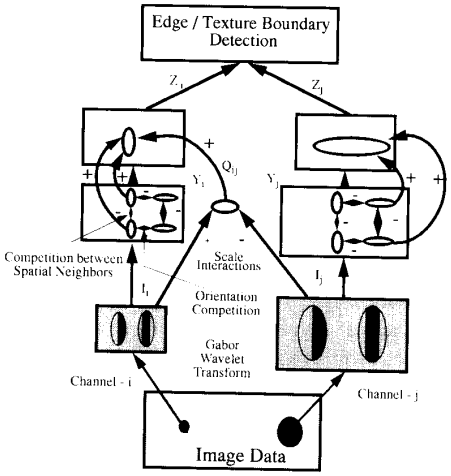

Fig. 1. Schematic diagram of the model. The input image is first processed through a wavelet transform based on Gabor functions. In the next stage local competitive interactions are introduced in each of the frequency channels. Interscale interactions help in localizing line ends. In the final stage outputs from like oriented cells are grouped to complete boundaries. Edges are located at the local maxima in Z and texture boundaries correspond to local maxima in the gradient field of Z.

mainly due to mathematical convenience and their important theoretical properties concerning localization in space and frequency. Gabor functions are modulated Gaussians having an even symmetric real part and an odd symmetric imaginary part. They have been used in many vision applications such as optical flow computations [10], image coding [11], [12], pattern recognition [13], and texture analysis [14]. The convolution of the image with these filters yields a representation which is localized in space as well as in frequency. The filter parameters determine the exact nature of this representation. A special class of this decomposition is the wavelet transformation where the filter profiles are all self-similar. Wavelets are families of basis functions obtained through dilations and translations of a *basic wavelet* and such a decomposition provides a compact data structure for representing information. In our case the basic wavelet is a Gabor function and we refer to this decomposition as the Gabor wavelet transformation (in [12] the term Gabor pyramid is used instead). A Gabor wavelet decomposition can be interpreted as extracting salient features in the image at different scales and orientations and the local maxima in their *energy* (see Section IV-A) correspond to the intensity edges in the image.

Following the wavelet decomposition we introduce local feature interactions. Three distinct types of interactions are considered: competition between spatial neighbors in each orientation channel, competition between orientations at each spatial location, and interscale interactions. These interactions are shown in Fig. 1. One extreme form of this type of interactions is the *winner-take-all* case where the dominant feature suppresses all the others, and this has been used in [7], [15]. Interscale interactions are used in localizing line ends and play an important role in boundary detection. The second stage

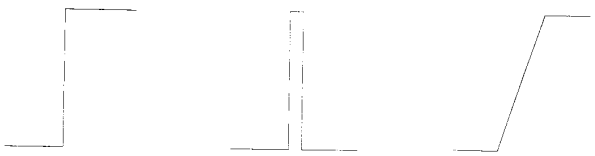

Fig. 2. (a) Step edge (b) bar or line edge (c) ramp.

of interactions groups similar features in the neighborhood. This cooperative processing helps in the boundary completion process. The receptive fields of the cells in this stage have the same orientation selectivity as their inputs and have a larger receptive field, and the filter profiles are modeled by oriented Gaussians. From a neurophysiological perspective the Gabor wavelet decomposition can thus be identified with processing by simple cells and local interactions with those of complex and hypercomplex cells.

The final step in the model involves identifying the boundaries in the image. Let the output after the grouping stage be denoted by Z_i, where i corresponds to the ith frequency channel. Features such as intensity discontinuities and illusory contours can now be located at the local maxima in Z_i and textural boundaries correspond to the local maxima in the gradient of Z_i. Experimental results on several images are presented to illustrate the performance of this model in detecting these features.

The organization of this paper is as follows: Section II discusses some related work on edge detection, preattentive segmentation and illusory contour perception. Section III gives a brief introduction to wavelets and Gabor functions. Section IV describes the different stages in our model which includes Gabor wavelet transformation to extract features and local feature interactions for segmentation and grouping. A brief analysis of the Gabor wavelets in edge detection is also given. Experimental results in detecting edges and texture boundaries in a variety of images are provided in Section V.

II. Review of Previous Work

A. Energy Features and Edge Detection

Almost all techniques for edge detection in computer vision literature have been developed for detecting step edges. Consequently their performance is poor on other edges such as lines (or bars) and ramps (Fig. 2). As mentioned earlier, the popular techniques include locating edges at the zero crossings of the Laplacian of the Gaussian convolved with the image [1], or at the local maxima of the outputs of convolution with directionally selective odd-symmetric filters [2]. The limitations of these methods are well known [16]–[18]. Some important observations are that feature detection/localization by any type of linear filtering operation is not adequate, and in particular zero crossings of the result of applying any linear operator to the image do not capture all significant features in the image. Secondly, there is a need for directionally selective

quadrature filter pairs. The outputs of these filters cannot be analyzed separately (except when either the step edges alone or the line edges alone are present). In general no linear filtering operation will be able to detect and localize composite edges accurately [17], [18]. As an alternative, Morrone and Burr [17] suggested the use of energy measures in edge detection. They show that locations in an image where the Fourier components have zero phase difference constitute perceptually significant features (such as the different types of edges mentioned above), and these could be detected at the local maxima in an appropriate energy measure. This energy model for edge detection is also used to explain perceptions of Mach bands[1] and several other visual illusions [19]. An analysis on the performance of energy features in composite edge detection is given in [20], where the filters are derived from the Gaussian function and its derivatives. Note that energy is one of the measures that can be used in combining the outputs of the even and odd symmetric filters and a discussion of other means of combining the information and their relative merits can be found in [18].

B. Preattentive Segmentation

Preattentive segmentation refers to the ability of humans to perceive textures without any sustained attention. Among the basic primitives that strongly influence the perception of textural scenes are color, brightness, orientation, and size [4]. In addition, distribution of features also play an important role. For example, Fig. 9(a) shows randomly oriented L and + texture. The two regions are easily discriminable, although there is no statistical difference in either orientation or scale of the micropatterns. What they differ is in the distribution of higher order features such as corners and intersections. Central to solving this problem are the issues of what features need to be computed and what kind of processing of these features is required for texture discrimination.

Some of the early work in this field can be attributed to Julesz [21] for his theory of textons as basic textural elements. Spatial filtering approach has been used by many researchers for detecting texture boundaries not clearly explained by the texton theory [5]. Recently an elegant computational model for preattentive texture discrimination is proposed by Malik and Perona [7]. Their three stage model involves convolution with even symmetric filters followed by half wave rectification, local inhibition, and texture boundary detection using odd symmetric filters. They present convincing arguments for the necessity of each of these stages, and provide quantitative measures regarding the performance of their algorithm to establish consistency between their results and human perception of textures.

C. Boundary Contour System

We begin with a brief review of some terminology. For simplicity, the various types of cells in the early processing stages in the visual cortex are grouped into three broad functional classes, simple, complex, and hypercomplex cells [22]. Simple cell receptive fields are sensitive to bars (lines) and step edges and their orientations, and can be modeled by even-symmetric (line detectors) and odd-symmetric (step edge detectors) filters. In addition the cells are also sensitive to the direction of contrast. Complex cells respond to more complex patterns such as textures, and unlike simple cells, do not contain any phase information, are less sensitive to precise location but are tuned to respond to different specific orientations and direction of movement. These cells are usually modeled by summing the outputs of a group of simple cells of similar orientations. Hypercomplex cells in the cortex exhibit end-inhibition, in that they respond to small lines and edges, and their response decreases as the length increases [22]. These cells appear to play an important role in localizing line-ends and texture boundaries, and both simple and complex cells with this end-stopping behavior are known to exist. A related concept, the end-cut mechanism, is introduced by Grossberg and Mingolla [15] where it is hypothesized that all perceived line ends are illusory and activation of orthogonal orientations at the end of lines are due to local competitive interactions in the early stages of visual processing.

Grossberg and Mingolla's Boundary Contour System (BCS) [15] is one of the first attempts to model the early processing stages in the visual cortex. The BCS processes the intensity data and performs preattentive segmentation of the scene. The first stage of BCS consists of oriented contrast filters at various scales and orientations and extracts the contrast information from the scene. The outputs of the filters are then fed to a two stage competitive network whose main goal is to generate end-cuts. Subsequent long range cooperative interactions and a positive feedback to the competitive stage help in boundary completion. The boundary detection takes place independently in different spatial channels. A detailed description of this model and its performance in texture grouping and in detecting illusory contours such as the Kanisza's square can be found in [15].

The BCS model provides a very general framework in which many of the current models for detecting boundaries can be included, though the details might differ. The model itself as detailed in [15] is quite complicated and computationally expensive to simulate on any real image examples. The BCS model does not account for even symmetric mechanisms which are useful in detecting lines and in texture discrimination [7]. Another problem with the BCS model is its end-inhibition mechanism, described as follows.

In the BCS model it is hypothesized that all line ends are illusory. In order to detect line ends, a two stage competitive network is proposed. Lateral inhibition between similar orientations in neighboring positions and between different orientations at the same spatial location result in the generation of end-cuts. At the line ends, the simple and complex cell outputs are not strong and compensatory mechanisms are required for detecting these line ends. However, it is not clear if these competitive interactions by themselves are sufficient for end-cut generation, and if not, what additional mechanisms are needed. Further, as noted in [23], the two stage model

[1] Mach bands are bright and dark lines that are perceived close to the transition regions of a blurred edge.

hypothesizes that all cells involved in boundary completion exhibit end-inhibition, which is a debatable issue. We propose an alternate model for this using scale interactions in Section IV-C, and demonstrate the usefulness of end-inhibition in texture discrimination and illusory contour perception.

III. Multiscale Representation and Wavelets

The multiscale approach provides an elegant hierarchical framework for image analysis. The features of interest in an image are generally present in various sizes. An efficient way to analyze such features is to have a multiscale decomposition of the image. Laplacian pyramid [24] is one of the early schemes developed for such applications, though initially it was proposed for compact image coding applications. Multiscale representation also helps in parallel processing as the different channels can now be analyzed independently (at least initially). Multiscale approach has also been used in robust detection of step edges [1] and Witkin [25] describes how information at different levels can be related. The presence of parallel visual pathways consisting of cells with varying receptive field sizes and orientations are indicative of a multiscale feature extraction in biological systems as well. However, the role of interactions that exist between different scales in these systems is not well understood. In Section IV-C we suggest a simple model which uses such interactions in detecting line ends and corners, and its possible biological significance.

There has been a growing interest in the use of wavelets for multiscale representation of the image data [26]. Wavelets are families of basis functions generated by dilations and translations of a *basic wavelet*. The wavelet transform is thus a decomposition of the function (image intensity) in terms of these basis functions. One of the objectives of such a transformation is to provide a simultaneous description of the data in frequency and spatial domains.

Let us first consider the one-dimensional case. Let $g(x)$ be a wavelet, $x \in \mathbf{R}$ ($\mathbf{R}$ denotes the set of real numbers). Then the family of basis functions corresponding to $g(x)$ can be generated by translations $(g(x-s))$ and dilations $(g(\alpha x))$, where s and α are the translation and scale parameters, respectively. Let this family be denoted by $(g(\alpha(x-s))), (\alpha, s) \in \mathbf{R}^2$. The wavelet transform of a function $f(x)$ (assuming that $f(x)$ is square integrable) is defined by

$$W_f(\alpha, s) = \int_{-\infty}^{\infty} f(x) g^*(\alpha(x-s)) dx \tag{1}$$

where the (*) indicates complex conjugate. Wavelets can be discretized by a suitable sampling of the parameters α and s. For example we can write the scale parameter as α^j where $j \in \mathbf{Z}$, $\mathbf{Z}$ being the set of integers. This results in a class of discrete wavelets represented by $g(\alpha^j x - n), (j, n) \in \mathbf{Z^2}$. A function $f(x)$ can then be expanded in terms of the basis functions $g(\cdot)$ as

$$f(x) = \sum_{i,j} c_{ij} g(\alpha^j x - i).$$

The Laplacian pyramid [24] mentioned earlier is a wavelet decomposition based on the Difference of Gaussian (DOG) wavelet and has found many applications in image processing [27]. Orthogonal wavelets are a special family of discrete wavelets corresponding to $\alpha = 2$, where the basis functions are mutually orthogonal, i.e., $\int g(x) g(2^j x - k) dx = 0$ for $((j,k) \in \mathbf{Z^2})$. A discussion on orthogonal wavelets and their applications to image processing can be found in [26]. An important feature of orthogonal wavelets is that the information at different resolutions is uncorrelated. Orthogonality, in general, is a strong condition, and is difficult to achieve if arbitrary orientation selectivity is desired. Further, it is harder to give a frequency domain interpretation of the features so extracted by the decomposition. In the following we consider a transformation based on nonorthogonal Gabor basis functions and discuss its usefulness in image processing applications.

A. Gabor Functions and Wavelets

Gabor functions are Gaussians modulated by complex sinusoids. In its general form, the two-dimensional Gabor function and its Fourier transform can be written as [28]

$$g(x, y; u_0, v_0) = \exp(-[x^2/2\sigma_x^2 + y^2/2\sigma_y^2] + 2\pi i[u_0 x + v_0 y]) \tag{2}$$

$$G(u, v) = \exp\left(-2\pi^2(\sigma_x^2(u-u_0)^2 + \sigma_y^2(v-v_0)^2)\right) \tag{3}$$

σ_x and σ_y define the widths of the Gaussian in the spatial domain and (u_0, v_0) is the frequency of the complex sinusoid. A well known property of these functions is that they achieve the minimum possible joint resolution in space and frequency domains [28]. A signal such as a delta function which is concentrated at a point in space has no frequency localization. Likewise, a function concentrated in frequency has no spatial localization. A good measure of localization in the two domains is the product of the bandwidths in space and frequency. The effective bandwidth of a signal is defined as the square root of the variance of the energy of the signal. Let δx and δy be the effective widths of the signal in the horizontal and vertical directions in space respectively and $\delta u, \delta v$ denote the corresponding widths in frequency. Then the following inequalities (also called the uncertainty relations) hold: a) $\delta x \delta u \geq 1/4\pi$ and b) $\delta y \delta v \geq 1/4\pi$. Gabor family of functions are unique in attaining the minimum possible value of this joint uncertainty. This localization property has received considerable attention among vision researchers and has led to many applications [11]–[14].

The Gabor functions form a complete but nonorthogonal basis set and any given function $f(x, y)$ can be expanded in terms of these basis functions. Such an expansion provides a localized frequency description and has been used in image compression [11] and texture analysis [14]. Local frequency analysis, however, is not suitable for feature representation as it requires a fixed window width in space and consequently the frequency bandwidth is constant on a linear scale. However, in order to optimally detect and localize features at various scales, filters with varying support rather than a fixed one

are required. This would suggest a transformation similar to wavelet decomposition rather than a local Fourier transform. We now consider such a wavelet transform where the *basic wavelet* is a Gabor function of the form

$$g_\lambda(x,y,\theta) = e^{-(\lambda^2 x'^2 + y'^2)+i\pi x'} \quad (4)$$
$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta$$

where λ is the spatial aspect ratio and θ is the preferred orientation. To simplify the notation, we drop the subscript λ and unless otherwise stated assume that $\lambda = 1$. For practical applications, discretization of the parameters is necessary. The discretized parameters must cover the entire frequency spectrum of interest. Let the orientation range $[0,\pi]$ be discretized into N intervals and the scale parameter α be sampled exponentially as α^j. This results in the wavelet family

$$(g(\alpha^j(x-x_0, y-y_0), \theta_k)), \alpha \in \mathbf{R}, j = \{0,-1,-2,..\} \quad (5)$$

where $\theta_k = k\pi/N$. The Gabor wavelet transform is then defined by

$$W_j(x,y,\theta) = \int f(x_1,y_1) g^*(\alpha^j(x_1-x, y_1-y), \theta) dx_1 dy_1. \quad (6)$$

At each resolution in the representation hierarchy these wavelets localize the information content in both frequency and spatial domains simultaneously. Any desired orientation selectivity can be obtained by controlling the orientation parameter θ. The Gabor wavelet decomposition also has an important physical interpretation of the type of features detected and this is further discussed in Section IV-A.

IV. Stages In Boundary Detection

We now discuss the various processing stages in our model shown in Fig. 1. We begin with a brief analysis of the Gabor wavelets as edge detectors.

A. Line and Edge Detectors

The wavelet decomposition using Gabor functions has an important physical interpretation. The complex Gabor function has an even-symmetric (cosine) real part and an odd-symmetric (sine) imaginary part, which respond maximally to *line edges (or bars)* and *step* edges (of appropriate sizes and orientations), respectively, in the image. This wavelet decomposition can be viewed as obtaining a primal sketch of the raw intensity data by detecting perceptually significant features at different scales. These features can be detected at the local maxima in their *energy* [17]. If R_i and I_i represent the response from the even and odd symmetric feature detectors at a position i, then the local energy E_i at i is given by $E_i = \sqrt{R_i^2 + I_i^2}$.

1) Performance Analysis: In the following we assume one-dimensional functions for simplicity and give an analysis of the signal-to-noise ratio (SNR) and localization properties analogous to the one given in [2]. The SNR is defined as the ratio of the signal power output to the noise power at the true location of the edge. Localization (L) gives a measure of the performance of the detector in accurately localizing the edge in the presence of noise and is defined as the inverse of the square root of the variance in this deviation. In [2], the product of SNR and localization is maximized in deriving an appropriate filter for detecting step edges. Notice that in general these two criteria contradict, as better localization implies poorer SNR.

For one-dimensional line and step edges it is shown in Appendix A that

$$\text{SNR(step)} = 0.2774/\sqrt{\alpha}, \text{L(step)} \approx 1.93\sqrt{\alpha} \quad (7)$$

$$\text{SNR(line)} = 0.7511\sqrt{\alpha}, \text{L(line)} \approx 0.33\alpha\sqrt{\alpha}. \quad (8)$$

In general (except for the special case of line edge), the SNR improves with increasing filter width whereas localization deteriorates. The SNR (for step edges) is poor compared to the first derivative of Gaussian used in [2], but its overall performance in the presence of composite edges is better. Note that both the Marr–Hildreth [1] and Canny [2] operators will fail to detect line edges at their true locations. In fact it is easy to see that at the true locations of line edges the SNR is zero for these operators.

B. Local Spatial Interactions

Following feature extraction using Gabor wavelets, we now consider local competitive and cooperative processing of these features. Competitive interactions help in noise suppression, and in reducing the effects of illumination.

These interactions are modeled by nonlinear lateral inhibition between features. Two types of such interactions are considered. The first type includes competition between different spatial neighbors within each orientation and scale. Fig. 1 shows the various interactions in two frequency channels in the system. For simplicity the transfer function $g(x)$ of all feature detectors is assumed to be the same. The following notation is used in explaining the interactions: The output of a cell at position $s = (x,y)$ in the ith spatial frequency channel with a preferred orientation θ is denoted by $Y_i(s,\theta)$, with $I_i(s,\theta)$ being the excitatory input to that cell from the previous processing stage. For example $I_i(s,\theta)$ could be the energy in the filter output corresponding to feature (s,θ) in the ith frequency channel. For convenience we will drop the subscript i indicating the frequency channel whenever there is no ambiguity. Let N_s be the local spatial neighborhood of s. The competitive dynamics is represented by:

$$\dot{X}(s,\theta) = -a_{s,\theta}X(s,\theta) + I(s,\theta) - \sum_{s'\in N_s} b_{s,s'}Y(s',\theta) - \sum_{\theta'\neq\theta} c_{\theta,\theta'}Y(s,\theta') \quad (9)$$

$$Y(s,\theta) = g(X(s,\theta)) \quad (10)$$

where (a, b, c) are positive constants. In our experiments we have used a sigmoid nonlinearity of the form $g(x) = 1/(1 + \exp(-\beta x))$. The dynamics of (9) can be visualized as follows: At each location within a single frequency channel, the corresponding cell receives an excitatory input from a similarly oriented feature detector (of the same spatial frequency). Further it also receives inhibitory signals from the neighboring cells within the same channel. We assume that all these interactions are symmetric ($b_{s,s'} = b_{s',s}$ and $c_{\theta,\theta'} = c_{\theta',\theta}$). The competitive dynamics of the above system can be shown to be stable. The Lyapunov function for the system [29], [30] can be written as

$$E(Y) = \frac{1}{2}\sum_{s,s'} b_{s,s'} Y(s,\theta)Y(s',\theta) + \frac{1}{2}\sum_{\theta,\theta'} c_{\theta,\theta'} Y(s,\theta)Y(s,\theta') + \sum_{s,\theta}\int_0^{Y(s,\theta)} \left(a_{s,\theta} g^{-1}(y) - I(s,\theta)\right) dy. \tag{11}$$

Under the assumptions that the interactive synapses are symmetric and that $g(\cdot)$ is monotone nondecreasing, the time derivative of E is negative and the system represented by (9) always converges.

The specific form of the dynamics such as the one in (9) is not very critical, as long as there is some form of local inhibition to suppress weak responses. For example, one can use the inhibition scheme proposed by Malik and Perona in [7], or the one suggested in Grossberg and Mingolla's BCS [15]. In [15] the orientation competition is separated from the local spatial competition between neighbors. An advantage of our method as well as that of BCS is that since the dynamics is expressed in terms of differential equations, it is amenable for analog implementations. One of the important differences between our model and that of BCS is in the generation of end-inhibition, discussed as follows.

C. Local Scale Interactions

We now suggest a simple mechanism to model the end-inhibition property of hypercomplex cells. For this the hypercomplex cell receptive field must have inhibitory end zones along the preferred orientation. Such a profile can be generated either by modifying the profile of the simple cell itself or through interscale interactions, discussed below. The fact that both simple and complex cells often exhibit this end-stopping behavior further suggests that both these mechanisms are utilized in the visual cortex. A schematic diagram of the model which utilizes interscale interactions is shown in Fig. 3. If $Q_{ij}(x, y, \theta)$ denotes the output of the cell C at position (x, y) receiving inputs from two frequency channels i and j ($\alpha^i > \alpha^j$) with preferred orientation θ, then

$$Q_{ij}(x, y, \theta) = g(||W_i(x, y, \theta) - \gamma W_j(x, y, \theta)||) \tag{12}$$

where $\gamma = \alpha^{-2(i-j)}$ is the normalizing factor. Fig. 4 shows a typical receptive field profile of such an end-inhibited

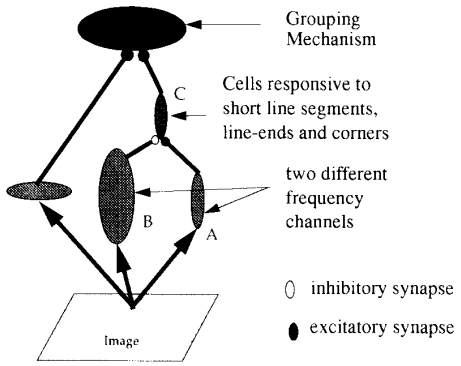

Fig. 3. Interscale interactions: Cells with larger receptive profiles (B) inhibit those with shorter receptive fields (C), which also receive excitatory inputs from similar sized cells (A). Due to these interactions cell C exhibits end-inhibition, and in turn cooperates with orthogonal orientations in grouping the edges.

cell, corresponding to an even symmetric receptive field, and without taking into account the sigmoid nonlinearity $g(\cdot)$. The parameter values used are $\alpha = \sqrt{2}, i = -2, j = -5$. In Fig. 3, unit C represents the hypercomplex cell, which receives an excitatory input from unit A and an inhibitory input from unit B. All three units A, B, and C have the same orientation preference and unit B has a larger receptive field profile compared to A. Unit C thus responds to only line ends and short line segments and its response decreases as the output of B increases for larger line segments. The logic behind this is simple. At line ends, cells with shorter receptive fields will have a stronger response than those with larger fields, and consequently will be able to excite the hypercomplex cells. At other points along the line, both small and large receptive field cells are equally excited and in the process the response of the hypercomplex cells is inhibited. It appears that such scale interactions to generate end inhibition do exist in the visual cortex. Bolz and Gilbert [31] observe that connections between layers 6 and 4 in the cat striate cortex play a role in generating end inhibition. The cells in layer 4 are of hypercomplex type exhibiting end inhibition. Layer 6 cells have large receptive fields and require long bars (or lines) to activate them. In addition, cells in both layers show orientation selectivity. Inactivating layer 6 cells resulted in the loss of end-inhibition property of layer 4 cells, while preserving other properties such as orientation selectivity. Thus, in the absence of layer 6 activity, cells in layer 4 could be excited by short bars and their response did not decrease as the bar lengths increased, suggesting that layer 6 cells have inhibitory effect on the cells of layer 4.

The model suggested in Fig. 3 is one of the ways of generating end-inhibition (and probably the most simple one) through scale interactions. The original idea of using such interactions dates back to the early work of Hubel and Wiesel [32]. A similar model is also suggested in [33], where the role of end-inhibition in detecting curvature is also discussed. It has generally been suggested that hypercomplex cells help

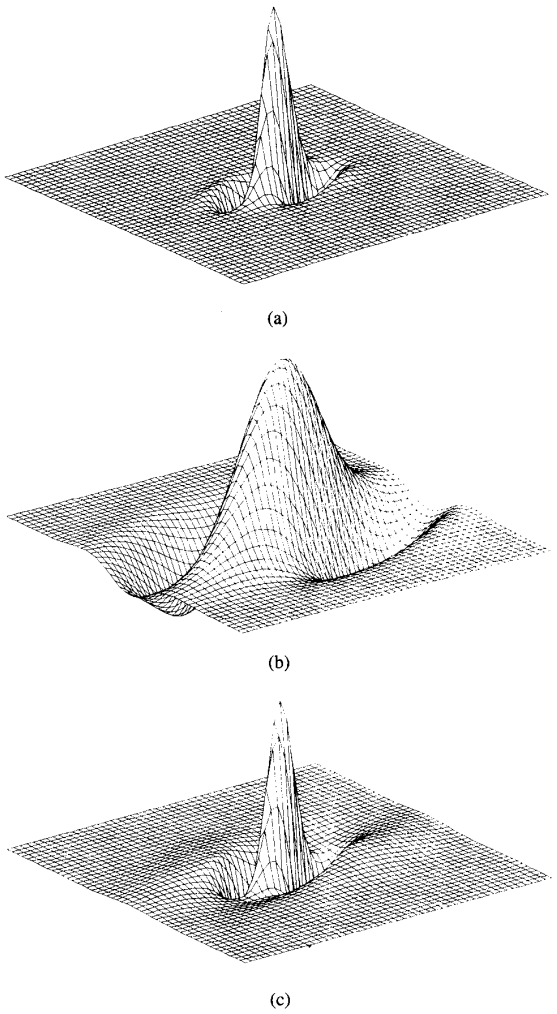

Fig. 4. Even symmetric cell's receptive fields (a) for scale $\alpha^i = 1/2$, (b) for $\alpha^i = 1/2\sqrt{2}$, and (c) Profile generated due to interactions between the above two fields.

in localizing texture boundaries. We provide here for the first time a demonstration of their usefulness in detecting texture boundaries (see example 4 in Section V). von der Heydt and Peterhans [23], [34] were the first to clearly demonstrate that these cells play a major role in the perception of illusory contours, and some of their observations are used in our model for boundary detection. Before getting into the details of their role in the perception of illusory contours, it would be interesting to consider again the type of features that these cells represent. The inhibitory end-zones of these cells help in making them respond to local curvature changes in the input intensity data (see also [33]). That these cells respond to line-ends is nothing but one extreme example. These cells thus form the first stage in extracting meaningful shape information. Appropriate grouping of these cells help in detecting texture boundaries, shape recognition or perception of illusory contours, depending on the context.

D. Grouping and Boundary Detection

The final stage involves grouping similar orientations. The grouping process receives inputs both from the competitive stage (9) and from the end-detectors (hypercomplex cells) described in Section IV-C. Note that the orientation of the activating end-detector is orthogonal to the actual orientation of the grouping process. This incorporates the observation made in [23], [34] that hypercomplex cells are responsible for detecting illusory contours. Abrupt line endings signal an occluding boundary almost orthogonal to the edge orientation, and this is represented by these end-inhibited cells providing input to the grouping process nearly orthogonal in their orientation preference. If $Z_i(s,\theta)$ represents the output of this process, then

$$Z_i(s,\theta) = g\left(\int d_i(s-s',\theta)(Y_i(s',\theta) + Q_{ij}(s',\theta')ds'\right) \tag{13}$$

$d_i(s,\theta)$ represents the receptive field of $Z_i(s,\theta)$ and in our experiments we have used

$$d(s=(x,y),\theta) = \exp(-(2\sigma^2)^{-1}[\lambda^2(x\cos\theta + y\sin\theta)^2 + (-x\sin\theta + y\cos\theta)^2]) \tag{14}$$

where θ is the preferred orientation, θ' is the corresponding orthogonal direction, and λ is the aspect ratio of the Gaussian. The Z cells thus integrate the information from similar oriented cells within each frequency channel and from hypercomplex cells of appropriate orientation, and thus help in grouping the features and in boundary completion. Since the various frequency channels are sampled, the effective standard deviation of the Gaussian is σ/α^i, where α^i is the scale parameter for channel i.

To summarize, this approach consists of three distinct steps a) feature detection using Gabor wavelets, b) local interactions between features, and c) scale interactions to generate end-inhibition. The output $Z(\cdot)$ from different frequency channels is now used to detect edges and texture boundaries.

1) Intensity edges and illusory contours: In Section IV-A the usefulness of energy detectors in localizing image features was discussed. In detecting the intensity edges in the image we used the energy features as input to the competitive stage. Thus the input to a cell in the competitive stage at a position (x,y) in the ith frequency channel is given by

$$I_i(x,y,\theta) = ||W_i(x,y,\theta)|| \tag{15}$$

where $W(\cdot)$ is as in (6), and $i = \{0,-1,-2,-3,\ldots\}$ and $\theta = k\pi/N, k = \{0,1,\ldots,N-1\}$, N is the number of discrete orientations. The edges are located at the local maxima of the $Z(\cdot)$ field in (13). These energy features are also used in our experiments to detect the line-ends through scale interactions. The perceptual boundaries for the examples in Fig. 10 are marked at the local maxima of the $Z(\cdot)$ field.

Texture boundaries: The information extracted by the wavelets can be used in several ways to detect textures, though the results reported here are obtained using the energy measure. Texture boundaries are located at the local maxima of the gradient of the Z field. Scale interactions also play an important role in texture boundary detection as is evident from

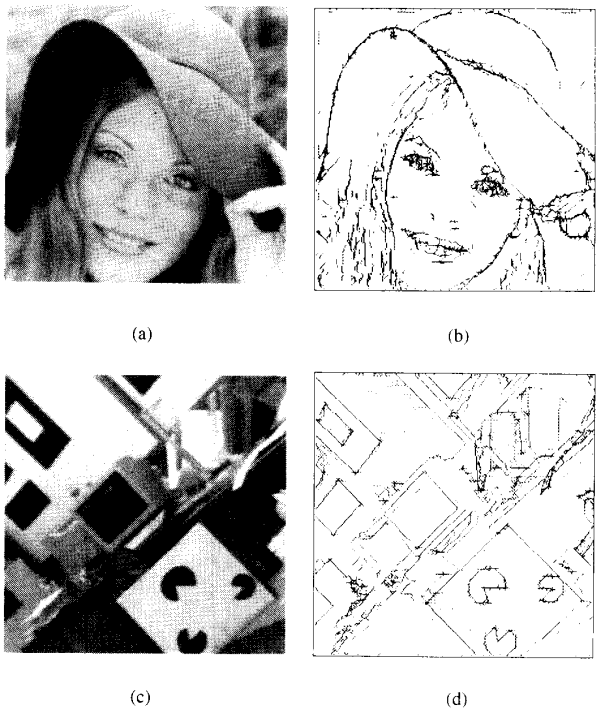

Fig. 5. (a) and (c) show two 256×256 images and the corresponding edges detected are shown in (b) and (d). In (b) the edges are from two channels $\alpha^i = \{1/\sqrt{2}, 1/2\}$ and in (d) $\alpha^i = 1/\sqrt{2}$. For both examples $\sigma = 1$. (Fig. 5(c) is taken from a University of Massachusetts image sequence.)

the example in Fig. 9 where the two regions differ only in the distribution of intersections and corners.

V. Experimental Results

The performance of the model is illustrated on several images. The following parameter values were used in our experiments described here: $\beta = 4.0$ in the transfer function $g(\cdot)$. The strengths of the inhibitory synapses in (9) are $b_{s,s'} = 1/||N_s||$ and $c = 1/N$, where $||N_s||$ is the cardinality of the neighborhood set and N is the number of discrete orientations used. Unless otherwise stated, $N = 4$ and N_s consists of the four nearest neighbors of s. The aspect ratio of the Gaussians in both the Gabor wavelets (4) and in the receptive field of Z cells (14) is set to 0.5. If more than one channel is mentioned then the result shown is a superposition of the boundaries detected in the individual channels. Since the various frequency channels are sampled, in order to bring them to the original image size the outputs of the grouping stage are first convolved with appropriate size Gaussian smoothing filters, and then the boundaries are detected.

Regarding the implementation of the dynamics of competition, we used a simple gradient descent on the corresponding energy function (11) instead of solving the set of differential equations. The equilibrium points in general for these two methods will be different, but gradient descent on E in (11) will be much faster (typically it takes less than 50 iterations to converge on a 256 × 256 image).

Example 1 (intensity edges): Fig. 5 shows two examples of edge detection using the energy measures. Fig. 5(a) and (c) show the original 256 × 256 images. The edges shown in Fig. 5(b) are detected in channels $\alpha^i = \{1/\sqrt{2}, 1/2\}$ and in (d) they correspond to the channel $\alpha^i = 1/\sqrt{2}$. In both cases σ is set to 1.

Example 2 (natural textures): Fig. 6 shows the boundaries detected in an image consisting of four textures, grass, water, wood, and raffia. The wood texture is present at two regions at different orientations. The parameter values used are $\alpha^i = \{1/2, 1/2\sqrt{2}, 1/4\}$ and $\sigma = 5.0$.

Example 3 (LT-T texture): Fig. 7 shows the results on a synthetic texture which is often used in psychophysical experiments. The boundary between L and Ts is not easily perceived where as that between straight and oriented Ts clearly stands out. This boundary can be easily detected in almost all frequency channels, and the parameters values used in this example are the same as in the previous example.

Example 4: Fig. 8(a) shows a texture consisting of micropatterns which differ in their sign. This particular texture is generated by adding to a constant intensity background (intensity value 120 on a 0–255 scale) patterns formed of bright (intensity 200) and dark (intensity 40) regions. The two regions in such textures can not be distinguished using energy measures. This was one of the motivations for using halfwave rectification in [7], and is based on the assumption that the filters are exactly zero mean and no nonlinear transformations of the intensity prior to filtering. The cosine component of the Gabor wavelet filters used in here is not exactly zero mean (though very close to zero). The nonlinearities following filtering enhance the differences at the boundaries as is illustrated by the boundary detected (Fig. 8(b)). Slight bias in the patterns toward one of the grey levels as in Fig. 8(c) (which has a background level of 150, brighter region at 200 and the darker region at 80) significantly influences the strength of the boundary (Fig. 8(d)).

Example 5 (L-Plus texture): This example illustrates the importance of end-inhibition in texture boundary detection. Fig. 9 shows another of commonly used texture consisting of randomly oriented Ls and +s. Unlike the previous example, orientation information can not be used for segmentation. The line segments forming Ls and +s have the same length (7 pixels). The two regions differ in the distribution of corners, line-ends, and intersections. As we discussed in Section IV-C, scale interactions play an important role in detecting these features. None of the scales by themselves contain enough information to segment the two regions, but using these interscale interactions the boundary between the Ls and +s can be detected (Fig. 9(b)). The boundary shown is for the case of using the interactions between scales corresponding to $\{1/2, 1/4\}$ with a $\sigma = 16$. In this context it is interesting to note the observation in [5] that the L+ texture can be discriminated by simple linear filtering followed by rectification, where they used size tuned center-surround filters. These filters are the simplest case of filters having inhibitory end-zones and as such respond to *blobs* of certain size. Hence in a sense, they are sensitive to the distribution of line ends and corners. Filters which have such inhibitory end zones include Laplacian of the Gaussians and the DOG's, and have been used in texture discrimination of L+ patterns in [7], [35].

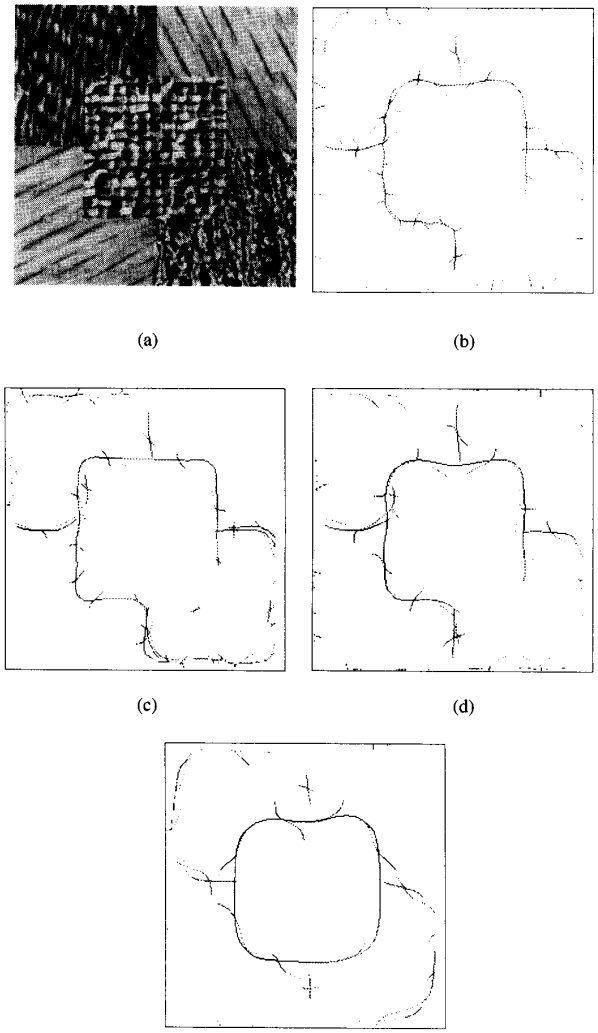

Fig. 6. (a) Image consisting of four natural textures, water, wood (in two regions at different orientations), raffia and grass. (b) Texture boundary detected using the scales $\alpha^i = \{1/2, 1/2\sqrt{2}, 1/4\}$ and $\sigma = 5$ pixels. (c), (d), and (e) Show the texture boundaries detected in each of these individual frequency channels separately. The result in (b) is obtained by superimposing the boundaries in (c)–(e), filtering using a Gaussian filter (to smoothly combine the boundaries) and thresholding. The filter used has a standard deviation of 4 pixels.

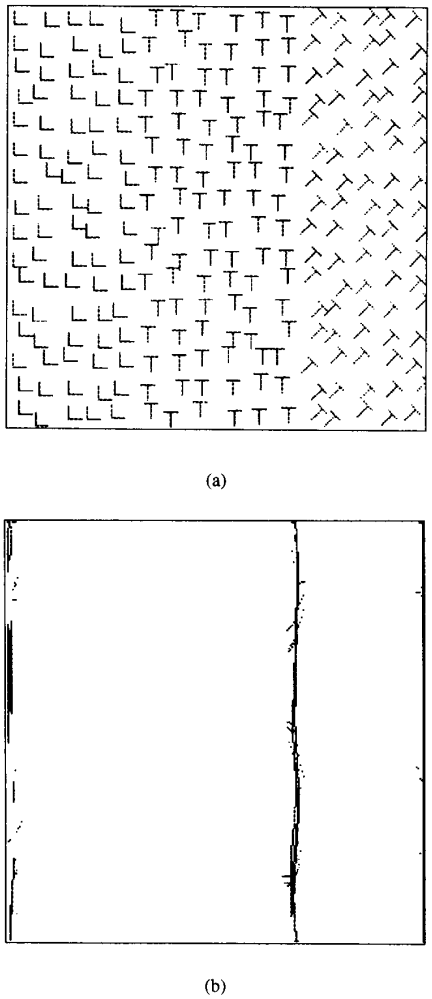

Fig. 7. Texture consisting of three regions, L, T and tilted-Ts. The boundary between L and T can not be easily detected. However the orientation difference between the two T regions is enough to discriminate between the two regions in almost all frequency channels. The boundary shown in (b) corresponds to the combined output from channels $\alpha^i = \{1/2, 1/2\sqrt{2}, 1/4\}$and $\sigma = 5$ pixels.

Example 6 (Illusory contours): The usefulness of scale interactions in detecting line endings and their subsequent grouping to detect illusory contours is illustrated in Fig. 10. For the line (Fig. 10(d)) and sine wave (Fig. 10(e)) contours the results shown are for $\alpha^i = \{1/2, 1/4\}$, $\sigma = 8$. For the circle (Fig. 10(f)) $\alpha^i = \{1/\sqrt{2}, 1/2\}$ and $\sigma = 2$.

VI. Conclusions

In this paper we have developed a common framework for detecting perceptually significant features such as edges, textures, and illusory contours. We have suggested a simple model based on detecting oriented features at different spatial scales and on local interactions between features. Interaction between frequency channels is used in generating end-inhibition which plays an important role in boundary perception. Several examples are provided to illustrate the performance of this approach in detecting different types of boundaries. We are also considering possible extensions of this model to include segmentation based on shading information in the image.

Appendix A

Here we derive the SNR and the localization error for the two cases (step and line edges) in one-dimensional. Let the

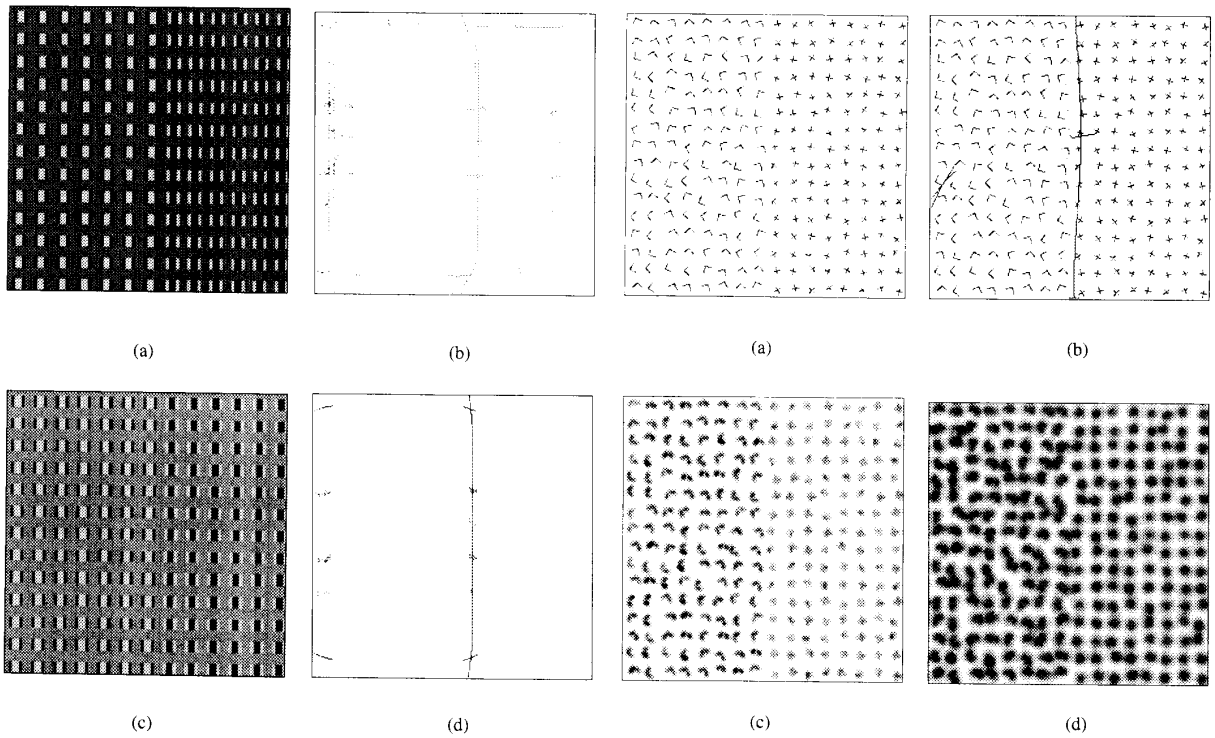

Fig. 8. (a) Primitives in this texture are zero mean (i.e., mean equals the background intensity level) patterns, with the intensity levels of the background, brighter and darker regions, respectively at 120, 200, and 40 (on a 0–255 scale), (b) boundary detected using $\alpha^i = 1/2\sqrt{2}$. $\sigma = 5$ pixels. Even a slight offset in the mean of the patterns can result insignificant increase in the strength of the boundary. In (c) the intensity levels are adjusted to be nonzero mean (at 150, 80, and 200, respectively for the background, darker and brighter regions, a net difference of 10 intensity levels between the background and the patterns) and the boundary detected is shown in (d).

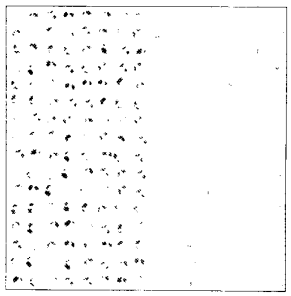

Fig. 9. Texture consisting of randomly oriented L and +. The line segments of the primitives are 7 pixels wide and the image is 256 ×256 pixels. The two regions differ in the distribution of line-ends, intersections and corners. The boundary shown in (b) (superimposed on the original texture) is detected using the output of the scale interactions with $\sigma = 16$. The scales used in this example are $\alpha^i = \{1/2.\, 1/4\}$, and (c) and (d) show the result of convolution and (e) shows the output after the interactions.

true location of the edge $e(x)$ be at the origin $x = 0$ and due to the presence of noise $\eta(x)$ the observed maximum in the energy $E(x)$ is at x_0. Let $y(x) = f(x) + n(x)$ be the response of the complex filter $g(x)$ due to the noisy input $e(x) + \eta(x)$, where

$$g(x) = g_r(x) + \imath g_i(x) = \exp(-\alpha^2 x^2/2 + \imath\pi\alpha x)$$

and $f(x)$ and $n(x)$ denote the signal and noise terms respectively at the output. The output noise energy for a white noise input is

$$\int_{-\infty}^{\infty} g(x)g^*(x)dx = \int_{-\infty}^{\infty} \exp(-\alpha^2 x^2)dx = \sqrt{\pi}/\alpha. \quad (16)$$

If the edge feature $e(x)$ is centered at the origin, the SNR at the true location of the edge is given by the ratio of the signal power to the noise power at the origin:

$$\text{SNR} = \frac{|f(0)|}{\sqrt{\sqrt{\pi}/\alpha}}. \quad (17)$$

The edges are located at the local maxima of the energy $E(x) = y(x)y^*(x)$ and $E'(x_0) = 0$. Further, assuming that the noise power is small compared to the signal power, we can approximate the energy by neglecting the terms containing $n^2(x)$, $E(x) \approx ff^*(x) + 2(f_r(x)n_r(x) + f_i(x)n_i(x))$, where $f_r(x) = g_r * e(x)$, $f_i(x) = g_i * e(x)$. Similarly n_r and n_i denote the real and complex parts of the output noise signal. Now,

$$E'(x_0) \approx (ff^*)'(x_0) + 2(f_r n_r + f_i n_i)'(x_0) = 0. \quad (18)$$

Expanding the first term in a Taylor's series around the origin,

$$(ff^*)'(x_0) = (ff^*)'(0) + x_0(ff^*)''(0) + \mathcal{O}(x_0^2). \quad (19)$$

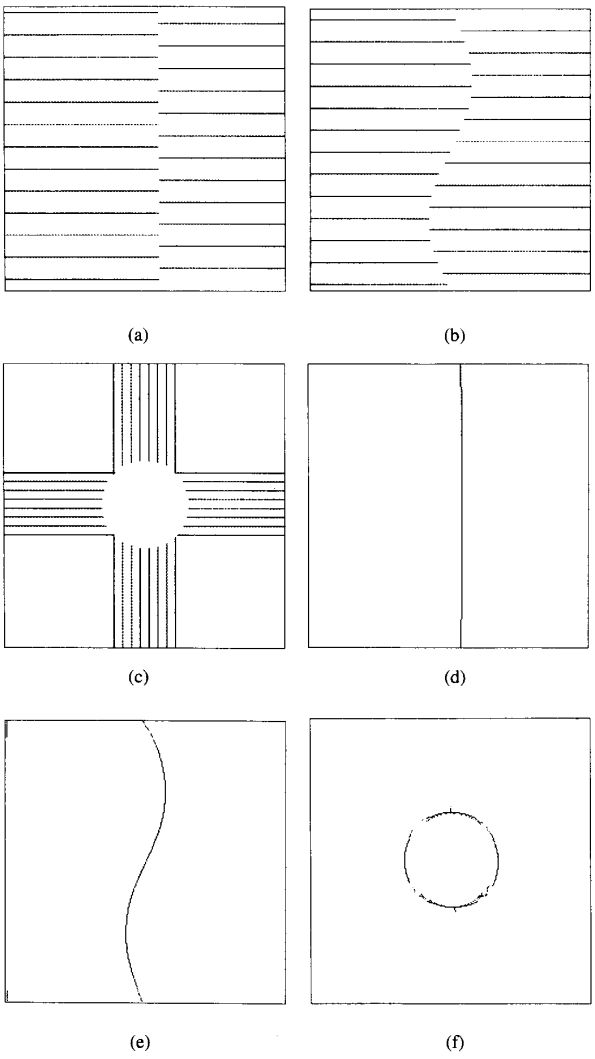

Fig. 10. Some examples of illusory contours formed by line terminations ((a), (b), and (c)) and the detected contours (d) and (e) correspond to the interaction between scales $\alpha^i = \{1/2, 1/4\}$ and $\sigma = 8$. In (f) the scales are $\alpha^i = \{1/\sqrt{2}, 1/2\}$ and $\sigma = 2$.

Noting that $(ff^*)'(0) = 0$ and substituting (19) in (18) and ignoring higher order terms, we get

$$x_0(ff^*)''(0) + 2(f_r n_r + f_i n_i)'(x_0) = 0. \tag{20}$$

As in [2], we use the inverse of the variance of x_0, $(\mathcal{E}(x_0^2))^{-1}$ as a measure of localization, and from (20) we get,

$$\mathrm{L} = [\mathcal{E}(x_0^2)]^{-\frac{1}{2}} \approx \frac{|[(ff^*)''(0)]|}{2\sqrt{\mathcal{E}[(f_r n_r + f_i n_i)'(x_0)]^2}}. \tag{21}$$

We now evaluate the SNR and localization for the cases of line and step edges. For a line edge, $e(x) = \delta(x)$

$$\mathrm{SNR(line)} = |f(0)|/\sqrt{\sqrt{\pi}/\alpha} = 0.7551\sqrt{\alpha} \tag{22}$$

$$\mathcal{E}[(f_r n_r + f_i n_i)'(x_0)]^2 \approx \mathcal{E}[(f_r n_r)'(x_0)]^2 \approx f_r^2(x_0)\mathcal{E}[(n_r)'(x_0)]^2.$$

Further, it can be shown that

$$\mathcal{E}[(n_r)'(x_0)]^2 = \mathcal{E}[(n_i)'(x_0)]^2 \approx 9.19\alpha$$

$$f_r(x_0) = (g_r * \delta)(x_0) = \exp(-\alpha^2 x_0^2/2)\cos(\pi\alpha x_0).$$

To evaluate the numerator in (21),

$$(ff^*)'' = f''f^* + f(f^*)'' + 2f'(f^*)' \tag{23}$$

and for the line edge, $f(x) = \delta(x) * g(x) = g(x)$, $f'(0) = g'(0) = \imath\pi\alpha$ and $f''(0) = g''(0) = -\alpha^2(1 + \pi^2)$. Substituting these values in (23) $(ff^*)''(0) = 2\alpha^2$, and finally from (21)

$$\mathrm{L(line)} = \frac{2\alpha^2}{2\exp(-\alpha^2 x_0^2/2)\cos(\pi\alpha x_0)\sqrt{9.19\alpha}} \approx 0.33\alpha\sqrt{\alpha}. \tag{24}$$

For the step edge we have $e(x) = \int_{-\infty}^{x} \delta(x')dx'$, $f(x) = \int_{-\infty}^{x} g(x')dx'$, $|f(0)| = |\int_{-\infty}^{0} g(x')dx'| \approx 0.3694/\alpha$, and

$$\mathrm{SNR(step)} = |f(0)|/\sqrt{\sqrt{\pi}/\alpha} \approx 0.2774/\sqrt{\alpha}. \tag{25}$$

Further we have $f'(x) = g(x)$, $f''(x) = g'(x)$ and from (23)

$$(ff^*)''(0) = g'(0)\int_{-\infty}^{0} g^*(x)dx + (g^*)'(0)\int_{-\infty}^{0} g(x)dx + 2g(0)g^*(0) = 4.32$$

and

$$\mathcal{E}[(f_r n_r + f_i n_i)'(x_0)]^2 \approx \mathcal{E}[(f_i n_i)'(x_0)]^2 \approx f_i^2(x_0)\mathcal{E}[(n_i)'(x_0)]^2 \approx 9.19\alpha(0.3693/\alpha)^2.$$

Substituting all these values in (21)

$$\mathrm{L(step)} = 1.93\sqrt{\alpha}. \tag{26}$$

ACKNOWLEDGMENT

The authors would like to thank Prof. von der Malsburg for his suggestions. They also thank Dr. S. Chandrashekhar and Dr. Q. Zheng for some useful discussions.

REFERENCES

[1] D. Marr and E. Hildreth, "Theory of edge detection," in *Proc. Royal Society of London (B)*, pp. 187–217, 1980.

[2] J. Canny, "A computational approach to edge detection," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. PAMI-8, pp. 679–698, Nov. 1986.

[3] B. S. Manjunath, T. Simchony, and R. Chellappa, "Stochastic and deterministic networks for texture segmentation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1039–1049, June 1990.

[4] J. Beck, K. Prazdny, and A. Rosenfeld, "A theory of textural segmentation," in *Human and Machine Vision*, J. Beck, B. Hope, and A. Rosenfeld, Eds. New York: Academic Press, 1983, pp. 1–38.

[5] J. R. Bergen and E. H. Adelson, "Early vision and texture perception," *Nature*, vol. 333, pp. 363–364, May 1988.

[6] B. Julesz and B. Krose, "Features and spatial filters," *Nature*, vol. 333, pp. 302–303, May 1988.

[7] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *J. Opt. Soc. Amer. A*, vol. 7, pp. 923–932, May 1990.

[8] K. B. Eom and R. Kashyap, "Composite edge detection with random field models," *IEEE Trans. Syst. Man Cybern.*, vol. 20, pp. 81–93, Jan 1990.

[9] D. Geman, S. Geman, C. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, pp. 609–628, July 1990.

[10] D. Heeger, "Optical flow from spatiotemporal filters," in *Proc. Int. Conf. Computer Vision*, London, England, June 1987, pp. 181–190.

[11] J. G. Daugman, "Relaxation neural network for nonorthogonal image transforms," in *Proc. Int. Conf. on Neural Networks*, vol. 1, San Diego, CA, June 1988, pp. 547–560.

[12] M. Porat and Y. A. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. PAMI-10, pp. 452–468, July 1988.

[13] J. Buhmann, J. Lange, and C. von der Malsburg, "Distortion invariant object recognition by matching hierarchically labeled graphs," in *Proc. Int. Joint Conf. on Neural Networks*, vol. 1, Washington DC, July 1989, pp. 155–159.

[14] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 12, pp. 55–73, Jan. 1990.

[15] S. Grossberg and E. Mingolla, "Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading," *Computer Vision, Graphics, and Image Processing*, pp. 116–165, Jan. 1987.

[16] E. D. Micheli, B. Caprile, P. Ottonello, and V. Torre, "Localization and noise in edge detection," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 11, pp. 1106–1117, Oct. 1989.

[17] M. C. Morrone and D. C. Burr, "Feature detection in human vision: a phase dependent energy model," in *Proc. Royal Society of London (B)*, vol. 235, 1988, pp. 221–245.

[18] C. Ronse, "A twofold model of edge and feature detection," preprint, Sept. 1990.

[19] D. C. Burr and M. C. Morrone, "Feature detection in biological and artificial visual systems," in *Vision: coding and efficiency*, C. Blakemore, Ed. Cambridge, MA: Cambridge University Press, 1990.

[20] P. Perona and J. Malik, "Detecting and localizing edges composed of steps, peaks and roofs," in *Proc. Int. Conf. Computer Vision '90*, Tokyo, Japan, Dec. 1990, pp. 52–57.

[21] B. Julesz, "Textons, the elements of texture perception and their interactions," *Nature*, vol. 290, pp. 91–97, Mar. 1981.

[22] D. H. Hubel and T. N. Wiesel, "Functional architecture of macaque monkey visual cortex," in *Proc. Royal Society of London (B)*, vol. 198, 1977, pp. 1–59.

[23] E. Peterhans and R. von der Heydt, "Mechanisms of contour perception in monkey visual cortex. ii. contour bridging gaps," *J. Neuroscience*, vol. 9, pp. 1749–1763, May 1989.

[24] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, pp. 532–540, Apr. 1983.

[25] A. Witkin, "Scale-space filtering," in *Int. Joint Conf. Artificial Intelligence*, Karlsruhe, Germany, 1983, pp. 1019–1021.

[26] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 11, pp. 674–693, July 1989.

[27] P. J. Burt, "The pyramid as a structure for efficient computation," in *Multiresolution image processing and analysis*, A. Rosenfeld, Ed. New York: Springer-Verlag, 1984, pp. 6–35.

[28] J. G. Daugman, "Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer. A*, vol. 2, pp. 1160–1169, 1985.

[29] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks," *IEEE Trans. Syst. Man Cybern.*, pp. 815–825, Sept. 1983.

[30] J. Hopfield and D. Tank, "Neural computation of decisions in optimization problems," *Biological Cybernetics*, vol. 52, pp. 114–152, 1985.

[31] J. Bolz and C. D. Gilbert, "Generation of end-inhibition in the visual cortex via interlaminar connections," *Nature*, vol. 320, pp. 362–365, Mar. 1986.

[32] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat," *J. Neurophysiology*, vol. 28, pp. 229–289, Mar. 1965.

[33] A. Dobbins, S. W. Zucker, and M. S. Cynader, "Endstopped neurons in the visual cortex as a substrate for calculating curvature," *Nature*, vol. 329, pp. 438–441, Oct. 1987.

[34] R. von der Heydt and E. Peterhans, "Mechanisms of contour perception in monkey visual cortex. i. lines of pattern discontinuity," *J. Neuroscience*, vol. 9, pp. 1731–1748, May 1989.

[35] H. Voorhees and T. Poggio, "Computing texture boundaries from images," *Nature*, vol. 333, pp. 364–367, May 1988.

B. S. Manjunath (S'88–M'91) received the B.E. degree (distinction) in electronics engineering from Bangalore University in 1985, the M.E. degree (distinction) in systems science and automation from the Indian Institute of Science in 1987, and the Ph.D. degree in electrical engineering from the University of Southern California in 1991.

During September 1987 to July 1991 he was a Research Assistant at the Signal and Image Processing Institute, USC, and in the summer of 1990 he worked at the I.B.M Thomas J. Watson Research Center, Yorktown Heights, NY. He joined the Department of Electrical and Computer Engineering at the University of California, Santa Barbara, as an Assistant Professor in July 1991.

Dr. Manjunath was awarded the University Gold Medal for the best graduating student in electronics engineering in 1985. He was also a recipient of the National Merit Scholarship from the Government of India during the period 1978-1985. His research interests include image processing, computer vision, neural networks, and pattern recognition.

Rama Chellappa (S'78–M'79–S'79–M'81–SM'83–F'92) was born in Tanjore, Madras, India in 1953. He received the B.S. degree (with honors) in electronics and communication engineering from the University of Madras in 1975 and the M.S. degree (with distinction) in electrical communication engineering from the Indian Institute of Science, Bangalore, in 1977. He received the M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1978 and 1981, respectively.

From 1979 to 1981, Chellappa was a faculty research assistant at the Computer Vision Laboratory, University of Maryland, College Park. From 1981 to 1991, he was a faculty member in the Department of Electrical Engineering-Systems, University of Southern California. From 1988 to 1990, he was also the Director of Signal and Image Processing Institute at USC. Presently he is a Professor in the Department of Electrical Engineering at the University of Maryland, where he is also affiliated with the Center for Automation Research, Computer Science Department and the Institute for Advanced Computer Studies. He is the coeditor of two volumes of selected papers on digital image processing and analysis published in 1985. He is a coauthor of a research monograph on *Artificial Neural Networks for Computer Vision* published by Springer-Verlag in 1992. He was an associate editor for IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING from 1987 to 1989. Currently, he is a coeditor-in-chief of *Computer Vision, Graphics, and Image Processing: Graphic Models and Image Processing* and is an associate editor for the IEEE TRANSACTIONS ON NEURAL NETWORKS and the IEEE TRANSACTIONS OF IMAGE PROCESSING. His current research interests are in signal and image processing, computer vision, and pattern recognition.

From 1969 to 1975, Dr. Chellappa received a national scholarship from the Government of India. He also received the 1975 Jawaharlal Nehru Memorial Award from the Department of Education, Government of India. In addition, he received the 1985 National Science Foundation (NSF) Presidential Young Investigator Award and the 1985 IBM Faculty Development Award. In 1990, he received the Excellence in Teaching Award from the School of Engineering at USC. He was the general chairman of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition and of the IEEE Computer Society Workshop on Artificial Intelligence for Computer Vision, both held in San Diego in June 1989. He was a program cochairman for the NSF-sponsored Workshop on Markov Random Fields, also held in San Diego in June 1989. He was a corecipient of a best paper award given at the International Conference on Pattern Analysis in 1992.