Feature Detection Algorithm Based on a Visual System Model

ELI PELI

An algorithm for the detection of visually relevant luminance features is presented. The algorithm is motivated and directed by current models of the visual system. The algorithm detects edges (sharp luminance transitions) and narrow bars (luminance cusps) and marks them with the proper polarity. The image is first bandpass filtered with oriented filters at a number of scales an octave apart. The suprathreshold image contrast details at each scale are then identified and are compared across scales to find locations in which the signal polarity (sign) is identical at all scales, representing a minimal level of phase congruence across scales. These locations maintain the polarity of the bandpass-filtered image. The result is a polarity-preserving features map representing the edges with pairs of light and dark lines or curves on corresponding sides of the contour. Similarly, bar features are detected and represented with single curves of the proper polarity. The algorithm is implemented without free (fitted) parameters. All parameters are directly derived from visual models and from measurements on human observers. The algorithm is shown to be robust with respect to variations in filter parameters and requires no use of quadrature filters or Hilbert transforms. The possible utility of such an algorithm within the visual system and in computer vision applications is discussed.

Keywords*—*Biological systems, edge detection, image matching, image processing, machine vision.

I. Introduction

Imitating the successful adaptation or strategies of the human brain in the design of engineering devices is a long-standing and attractive concept. If one understood how the human brain performs its many successful functions, the same approaches could be implemented in mechanical devices needed for similar tasks and improve their performance [1]–[7]. A number of such applications were presented in a book on visual models for target detection [8]. The most recent incarnation of this approach is in the field of neural networks. The neural-networks concept pushes this idea further, suggesting that if we build our machines from building blocks similar to those composing the human neural system, we do not even have to understand how the brain solves a particular problem. The neural-network system will solve the problem in a similar way when presented with it [9].

More often, however, the transfer of concepts was actually reversed. The technology of the day was implemented in new models of the brain and of the visual system. There are various possible reasons for this direction of the process. We understand our technology (including neural-networks technology) better than we understand our brain, and our view of the brain and the visual system as complex structures leads us to form simplifications using the more familiar and simpler technological units. It is also possible that technology has advanced faster in the last two decades than has our understanding of the visual system or the brain.

The idea of transfer from human vision performance to technological solution is so attractive that in many cases a human-vision-based motivation was suggested for processes and algorithms that did not necessarily need it. For example, Oppenheim *et al.* [10] applied it to homomorphic filtering and Frei [11] applied it to histogram hyperbolization. While the proposed algorithms, or some steps in them, were similar to properties of the visual system, the development of these algorithms was essentially independent from the visual system models of the day and probably was not derived from them.

This paper describes the development of a feature (edge and narrow bar) detection algorithm that is motivated by contemporary vision models as an example of the utility of the transfer of our understanding of visual function to algorithmic design. An edge is a luminance transition from a light to dark area (or vice versa) in the image. A bar is frequently defined as a strip of light (or dark) luminance on a dark (or light) background [12], [13]. Such a bar, if it is wide enough, may be viewed as composed of two edges and will be detected as such by our algorithms and many others. Here, a bar, line, or a cusp is defined as a very narrow (in terms of the algorithm) local increment or decrement in luminance from the background (such as a dark thin line drawn on a light background). As shown by Burr *et al.* [12], some features may be considered to be a combination of an edge and a bar at the same location.

Unlike previous approaches that derived only from the general structure of the visual system, the current algorithm

Manuscript received December 18, 2000; revised August 3, 2001. This work was supported in part by National Instutes of Health under Grant EY05957 and Grant EY12890 and in part by the National Aeronautics and Space Administration under Grant NCC-2-1039.

The author is with the Schepens Eye Research Institute, Harvard Medical School, Boston, MA 02114 USA (e-mail: eli@vision.eri.harvard.edu).

Publisher Item Identifier S 0018-9219(02)00728-4.

implements parameters derived directly from visual system measurements. Like many previous visual system motivated approaches to edge detection, the algorithm developed here employs multiscale processing and a nonlinear thresholding step. However, unlike all other approaches, this algorithm uses the nonlinear thresholding both within and across the scales to compute the features. In addition, the visual threshold as a way of controlling noise is applied here in distinction from other algorithms. All of this is done within visual-model-based approaches and parameters. As such, this algorithm may represent a new family of algorithms that are directly derived from the multiscale and threshold nature of visual processing neural response.

Approaches to the development of edge or feature detection algorithms presented in the literature are too numerous to permit proper review here. Heath *et al.* [14], [15] reviewed and compared a few of these algorithms and found that for each algorithm, performance differed with the image used, was significantly affected by choice of parameters, and the image and parameter selection interacted significantly as well. Some algorithms in the literature were based on vision models or developed at least in part as a model of visual feature detection. Only a few such algorithms are addressed here. The few algorithms reviewed below were selected only as illustrations of the properties and approaches that are either similar to or different from the properties of the algorithm described here. This review is far from being exhaustive or complete even with regards to this smaller class.

Perhaps the best known visual edge detection algorithm is the one proposed by Marr and Hildreth [16]. In the Marr and Hildreth algorithm, the image is first bandpass filtered at various scales. The bandpass-filtered image at the ith scale is

$$F_i'(x,y) = \nabla^2 G_i(x,y)^* I(x,y) \tag{1}$$

where G is a two-dimensional (2-D) Gaussian, I is the image, and ∇^2 is the Laplacian. The operator $\nabla^2 G$ is similar to the isotropic bandpass filter used in our algorithm. Gaussians of different widths were used to obtain filtering at different scales. The zero-crossing locations were then found at each scale by thresholding the bandpass-filtered image at zero. If at least two zero-crossing locations from independent but adjacent scales coincided, they were said to represent a feature in the image. The features found through the coincidence test needed to be further analyzed to determine if they represented edges or bars and of what polarity. The algorithm for computing these coincidences was deemed to be too complex to be described in the original paper [16] or Marr's book [17]. The complexity in determining the coincidence is a result of the shift or drift of the zero-crossing location across scales, as pointed out in [18].

A computational algorithm detecting zero crossings is applicable without direct relation to a visual model. However, no structure in the visual system is known to detect zero crossings. In fact retinal processing, which does apply multiscale bandpass filtering, separates the positive and negative aspects of the bandpass-filtered image using unipolar responding cells. The ON-center cells carry the information about the positive output of the filtering operation and the OFF-center cells carry only the information about the negative output of the bandpass filtering.

Marr and Hildreth [16] did provide a possible biological basis for the detection of zero crossings in the retinal image. The presence of zero crossings can be detected between contiguous ON- and OFF-center cells if the response is combined with an AND operation. Cell pairs at different scales must be associated with every zero-crossing location to permit local edge detection. They indicated that they were aware of the complexity of such a network of pairs of cells or lines of cells needed to account for all the possible edge lengths and orientations at every retinal location.

The new algorithm described here is most similar to the Marr and Hildreth algorithm in the sense that both apply hard threshold functions both within and across scales. While the Marr and Hildreth [16] algorithm works, as demonstrated in their publications, the detection of zero crossings is highly sensitive to noise. At low signal-to-noise ratio (SNR) the algorithm produces spurious zero-crossings. At higher SNR, the position of the zero crossings may be affected by the noise. This sensitivity to noise might be one of the reasons the visual system evolved to use nonzero-crossing cells such as the ON- and OFF-center cells. A good engineering design would suggest that it is best to operate as far from the zero crossing as possible. The algorithm presented here does so.

Morrone and Burr [13] have shown that edges and bars (cusps) in the signal $g(x)$ are located where the energy function $D(x)$ has peaks or local maxima. The energy function is

$$D(x) = g^2(x) + h^2(x) \tag{2}$$

where $h(x)$ is the Hilbert transform of $g(x)$, which is derived from the original image $I(x)$. They also showed that the peaks in the energy function occur where one finds a maximal phase congruence of the Fourier components of the signal. Therefore, they suggested that feature detection in the visual system might be accomplished by computation of the energy function and detection of its local peaks. While the mathematical foundation of this approach is sound and the energy function concept was adapted by others including the author [19], the evidence for the sum of squares energy computation in the visual system is less directly motivated by known physiology [20]. Furthermore, Georgeson and Freeman [21] directly tested the hypothesis that the peaks in the energy function determine the position of image features indicated by observers in one-dimensional (1-D) compound gratings patterns and were able to reject it. The algorithm described here was developed to explore the possibility of determining the phase congruency at visual features using a different approach that is not based on the computation of the energy function.

Watt and Morgan [22] developed the MIRAGE theory for encoding 1-D luminance primitives. In this vision model, the 1-D image is also processed into multiple scales by filtering (convolving) with a second derivative of a Gaussian of various scales. At each scale, the response of the filtered

image is separated into the ON and OFF responses by rectification. The processing up to this stage is identical to that of the basic algorithm proposed here except that the MIRAGE uses the rectification to separate the image into the two categories, but maintains the gray levels at each of the two components. The ON and OFF responses in the MIRAGE are then averaged separately across all scales. Watt and Morgan [22] concluded that the centroid of the averaged response was the best measure of edge position in a noisy system. They indicated that the exact center frequency used for the various filters was not crucial for the proper detection of edge position. There are a number of limitations to this approach that relate to the comparison with the algorithm described below. Following rectification to separate the ON response from the whole signal, the noise is also rectified. Even if one initially assumes a zero mean noise, the rectified noise has a nonzero average and that must be subtracted from the averaged signal. It is not clear how one can determine the level of rectified noise except by guessing (for example, Hess and Dakin [23] reported using ± 1 standard deviation of the gray levels in the image as that threshold). The algorithm also required a threshold mass parameter and it is not clear where one finds that threshold and what its basis is in current vision models. The process described so far was used only to determine the location of primitives or features. However, the determination of their nature (bar versus edge) and their polarity required further application of complex rules. The physiological or psychophysical basis for these rules was not clear. Finally, the MIRAGE was offered only for processing of 1-D signals and the extension to two dimensions may not be trivial (although Hess and Dakin [23] have reported using a 2-D adaptation of the algorithm, but gave no details).

The algorithm described below uses many of the visual properties of the algorithms described above. Unlike these previous approaches, it stays explicitly within a vision model and applies it directly using a different and novel scheme of combining the multiscale bandpassed signals that were uniformly used by the previous algorithms reviewed above. Additionally, the approach is completely image independent and no parameters are fitted in the process.

Evaluations and assessments of the edge detection performance presented here are based on the author's observations that hopefully are in agreements with the reader's. While this method is not very satisfying, it is the method most commonly used for evaluation of edge detection in real scene images [15]. Other methods of evaluation of edge detection algorithm performance are essentially a derivative of the same approach—using either observer specified edges as ground truth or averaged judgment of multiple observers about the quality of the algorithms performance [14].

II. Visual Feature Detection Algorithm

A. Basic Algorithm

Image processing in the visual system is carried out in multiple scales. The image is broken into various representations at multiple scales starting with the neural processing at the retina. At the output of the retinal ganglion cells, the image is represented by isotropic bandpass-filtered versions at different scales [24]. At higher cortical levels, the representation is multiscale bandpass filtered, but is also orientation selective [25], [26]. The basic feature detection algorithm presented here applies that structure directly. The image is filtered through a multiscale set of visual channels such as bandpass filters. The filters are one-octave wide in the frequency domain and separated by one octave at their center frequencies. In the most basic algorithm, the filters are isotropic without orientation sensitivity, although the orientation selective option is presented later as well. The specific filters used are not crucial, as will be shown below. The filters used here were log-cosine filters [27] applied in the spatial frequency domain (f_x, f_y). The center frequency of the ith-order filter used to generate the ith-scale bandpass-filtered image was 2^i cycles/picture, $i \geq 1$. The ith-order filter was

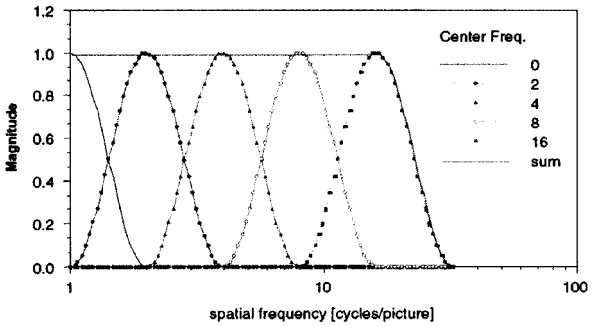


Fig. 1. Set of log-cosine filters (one-octave wide at half maximum and one-octave apart in center frequency) used to decompose the image into bandpass-filtered versions. Identical filters are added at higher frequencies for larger images. The actual filters were rotational bodies of the 1-D filters shown here. Note that the filters are symmetric on a logarithmic scale, as visual channels are, and they sum to unity, indicating that the decomposition is complete.

$$C_i(f_x, f_y) = \begin{cases} 0.5\left[1 + \cos\left(\pi \log_2 r - \pi i\right)\right], 2^{i-1} \leq r \leq 2^{i+1} \\ 0, \text{elsewhere} \end{cases} \quad (3)$$

where r is the radial spatial frequency $r = \sqrt{f_x^2 + f_y^2}$. These filters have a number of convenient properties that make them appropriate for digital implementation as well as for representation of the visual system decomposition. They have equal bandwidth and symmetrical shape on a log frequency axis and they form a complete representation of the decomposed image that can be reconstructed from them by simple addition (Fig. 1).

The isotropic bandpass filters were applied in the Fourier domain and the bandpass-filtered images were converted back to the space domain (without rescaling). Note that since the filters only affect the amplitude of the signal their phase can be considered to be zero or in cosine phase. The output images of the set of filters $F_i(x, y)$ were then binarized, creating bipolar representations $B_i(x, y)$ in which the white pixels might represent the locations where the ON-center

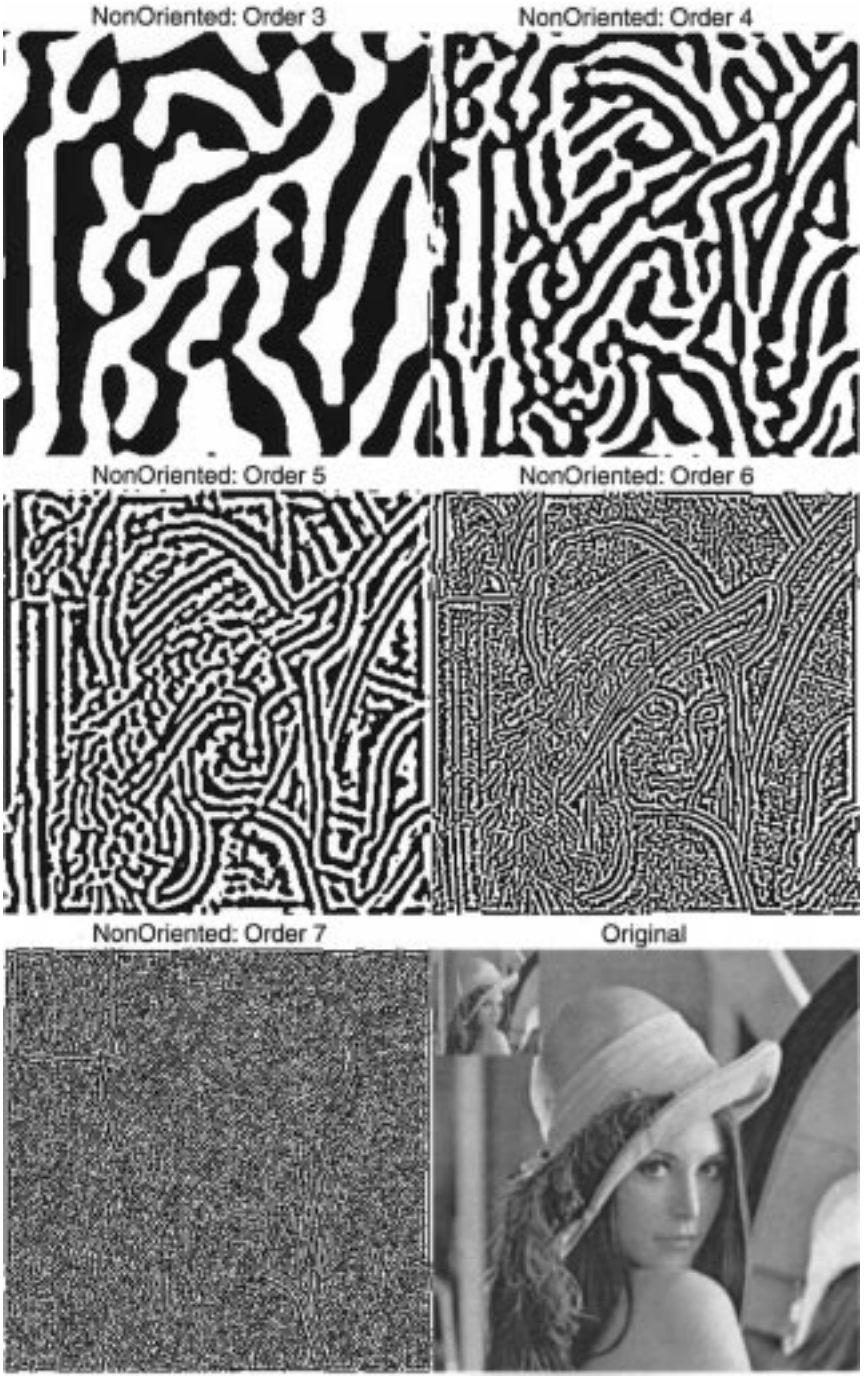


Fig. 2. Bandpass-filtered versions of the Lena image following binarization. Shown are scales with center frequency of 8, 16, 32, 64, and 128 cycles/image, corresponding to 1, 2, 4, 8, and 16 cycles/degree for an 8° image span. Original Lena image is at the lowest right.

cells respond and the black pixels the locations where the OFF-center cells respond

$$B_i(x,y) = \begin{cases} +1, & \text{if } F_i(x,y) \geq 0 \\ -1, & \text{if } F_i(x,y) < 0 \end{cases}. \tag{4}$$

Four to five such filters were used in agreement with models of the visual system [28]. Note that the bandpass-filtered images $F_i(x,y)$ (Fig. 2) are similar, but not identical to the bandpass-filtered images $F_i'(x,y)$ of (1). The small differences between the filters are not material to the functionality of the algorithm, as will be seen later. The application of a threshold hard-decision rule early in the system is not common in edge detection algorithms (except for Marr and Hildreth's [16] algorithm); usually the hard decision rule is kept until as late as possible in the process. However, in the visual system, the analog signal is thresholded and converted to impulses quite early and then repeated in the chain of events. Note also that application of contrast gain control or normalization, as have been

Fig. 3. Trilevel feature map derived from the images in Fig. 2 using the congruence rule. Each white pixel represents a location where the corresponding pixels in all scales of Fig. 2 were white. A black pixel represents a location with corresponding pixels in all scales in Fig. 2 that were black. Gray pixels are from locations in which the polarity was switched at least once across scales. Note that a polar pair of pixels of the edge polarity represents edge features. Bar features are represented by single pixels of the corresponding polarity.

Fig. 4. Features map derived from the images in Fig. 2 using the same congruence rule as in Fig. 3, but dropping the lowest scale image (order three). When compared to the features map in Fig. 3, this demonstrates a higher rate of feature detection, but also a significant increase of false alarms or noise pixels.

applied to visual models [27], has no effect here, as it will not change the binarization result. The use of binarization here does not imply that contrast is completely lost for other nonedge representations.

As mentioned above, Morrone and Burr [13] have shown that the phase of the image decomposition components in various scales is congruent at locations where relevant visual features (edges and bars) occur in an image. In order to find such features, one could look for phase congruency across scales in the bandpass-filtered series of images. Marr and Hildreth's [16] coincidence of zero crossing across adjacent scales is an extreme example of such phase congruency requiring a congruency at a single phase angle and further requiring it to be at the $\pi/2$ phase exactly. Binary phase matching or the congruency of signal polarity (or sign) might be considered the lowest level of phase congruency. At this other extreme, the congruency is accepted if just the sign is equal across scales. To find locations with this minimal level of phase congruency, the binary images from all scales were compared using a XOR-like operation (called congruence) across all scales. Each pixel that had the same polarity in all the scales maintained that polarity, while pixels with mixed polarity across scales were deselected (presented as gray pixels in Fig. 3). This congruence operation results in a trilevel map $E(x, y)$, representing all image pixels that had the same polarity across all scales tested, and only these pixels as features of that polarity. In biological terms, this operation might be similar to higher level cells responding only if all the cells (in all scales) of a given polarity (ON or OFF) at a specific retinal/field location are active. The congruence function is defined as

$$E(x,y) = \begin{cases} +1, & \text{if } B_i(x,y) = +1, \forall i \\ 0, & \text{otherwise} \\ -1, & \text{if } B_i(x,y) = -1, \forall i \end{cases}. \qquad (5)$$

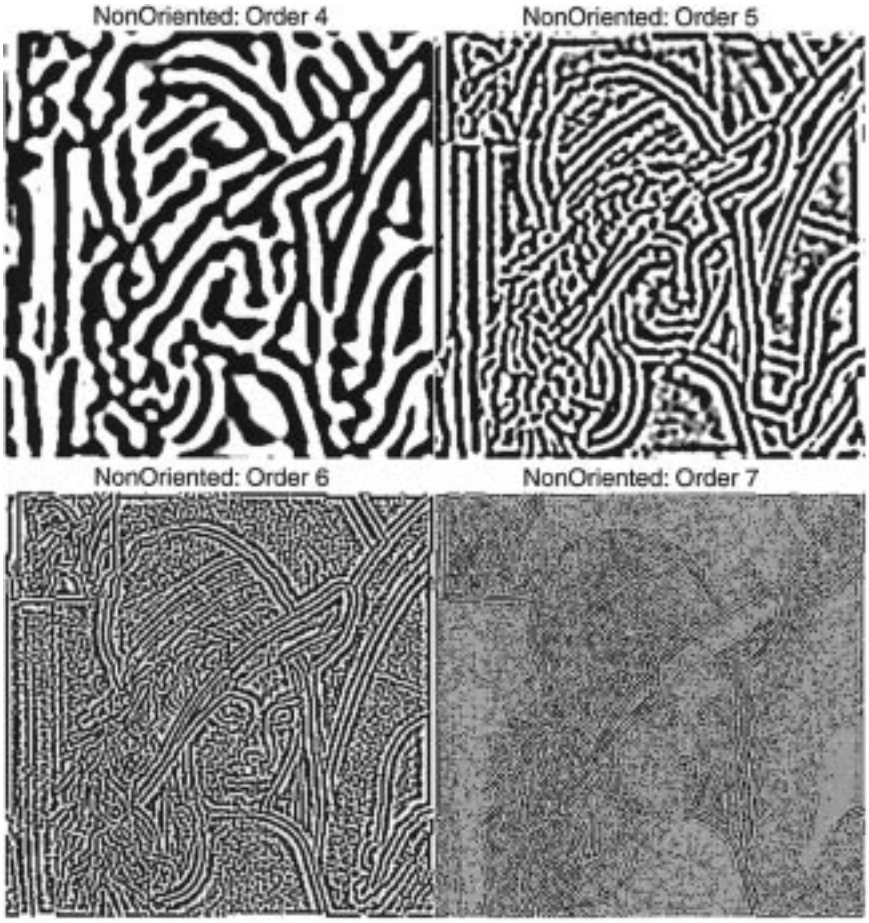
The simple and minimally restrictive nature of the phase congruence rule used in the algorithm described here makes computational implementation simple and the visual system implementation plausible without limitation on the number of scales to be tested. On the other hand, the restricted congruency of the zero crossings used in the Marr and Hildreth algorithm resulted in such a complex application (even when applied across two scales only) that the full algorithm was never described. The plausibility of the physiological implementation of these rules was not discussed either.

The features image presented in Fig. 3 represents a successful detection of relevant visual edges and narrow bars. The algorithm results are unique in that it represents edge features with pairs of black and white pixels placed at the edge with the black pixel on the dark side of the edge and the white on the bright side. Similarly, the narrow bar features (cusps) are detected and represented with the corresponding polarity without any additional processing.

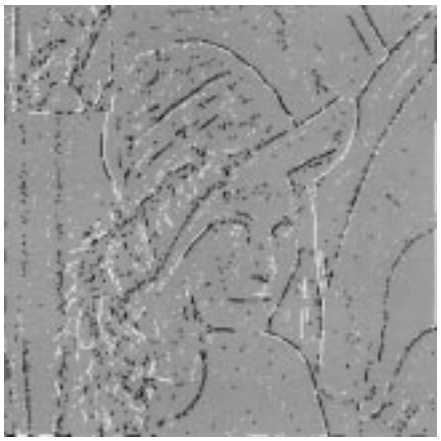
While the results of the basic algorithm illustrated in Fig. 3 using the five isotropic filters are quite satisfactory, it is clear that the algorithm did not detect many features easily detected by human observers. Improvement of the detection rate in this algorithm can be achieved by removing one of the scales, reducing the probability of a polarity change across scales. As shown in Fig. 4, the removal of the lowest scale indeed resulted in the expected improvement of features detection, however, with a simultaneous increase in noise as expected.

B. Using Contrast Detection Threshold to Control Noise

To address the noise, the algorithm applies contrast detection thresholds. This is a direct replication of the function of the visual system as described in most current models. Visual cells or psychophysical channels, ON and OFF center, are known to respond vigorously when the signal exceeds a certain level. This characteristic of the visual system is frequently measured as a contrast detection threshold and is used to represent the visual system's response by using the contrast sensitivity function $\mathrm{CSF}(r)$, which is the inverse of



(a)



(b)

Fig. 5. (a) Bandpass-filtered versions of the Lena image following trilevel processing using, at each scale, the corresponding threshold derived from contrast sensitivity measurements. Shown are scales with center frequency of 16, 32, 64, and 128 cycles/image, corresponding to 2, 4, 8, and 16 cycles/degree for an 8° image span. Notice the subthreshold portion at the highest frequency image due to the high threshold. Note, however, that gray segments appear in all orders and their proportion increases with increased spatial frequency. Comparison of the highest frequency band to the corresponding one in Fig. 2 (lower left) clearly illustrates the noise reduction effect of the threshold application. (b) Features map derived from the application of the congruency rule to the images in Fig. 5(a). This map maintains the feature detection of Fig. 4, but has reduced noise due to the effect of the visual threshold application.

the detection threshold, as a function of spatial frequency [29]. The aforementioned basic algorithm (4) was modified to include the contrast threshold or the $\mathrm{CSF}(r)$. Following bandpass filtering, the images at each scale $F_i(x, y)$ were transferred first into a trilevel thresholded image $T_i(x, y)$ instead of the binarization applied in the basic algorithm [see

Fig. 5(a)]. At each scale i, the corresponding threshold absolute value Th_i was determined and applies as follows:

$$T_i(x,y) = \begin{cases} +1, & \text{if } F_i(x,y) \geq +Th_i \\ 0, & \text{if } -Th_i < F_i(x,y) < +Th_i \\ -1, & \text{if } F_i(x,y) \leq -Th_i \end{cases} . \quad (6)$$

The threshold value Th_i is derived directly from CSF measurements in the literature [29]. These measurement results are typically given as a function of retinal spatial frequencies in cycles/degree. The image processing in the algorithm is performed in terms of object spatial frequencies in cycles/image. In order to determine the correspondence between these two domains, the size of the image in degrees must be set, depending on the specific application and situations being depicted. In the examples provided here, the images were presumed to span 8° of visual angle at the observers eyes. Thus, filters centered on 8 cycles/image correspond to 1 cycle/degree and the corresponding threshold was applied. Unless specified otherwise, the images presented here have a resolution of 256×256 pixels.

The ideal hard-threshold description applied here ignores the noisy nature of the visual system, where some cells actually discharge some spikes spontaneously at all times even in total darkness and the observers' psychophysical response is characterized by a psychometric function and not by a hard threshold. The hard threshold was used in the algorithm to simplify computations. This approach was justified by experimentation with the basic underlying vision model that showed that the use of statistical threshold provided minor improvements in controlling the occasional false contour artifacts caused by the hard threshold when computing simulations [30]. The false contours were completely eliminated also when other variabilities such as CSF changes with retinal eccentricity were included in such simulations [31], [32].

The application of contrast normalization [26] in this CSF-based thresholded version of the algorithm does have an effect on the result, particularly in very dark and very bright areas in the image. The effect, however, was found to be small with all the images shown here as well as with a group of 125 other images and, therefore, for the sake of computation simplicity, that aspect of the visual model was not included in further testing.

The trilevel threshold images were then evaluated across scales using the same congruence operation used in the basic algorithm to derive the trilevel feature image $E'(x,y)$

$$E'(x,y) = \begin{cases} +1, & T_i(x,y) = +1, \forall i \\ 0, & \text{otherwise} \\ -1, & T_i(x,y) = -1, \forall i \end{cases} . \quad (7)$$

The result of this processing shown in Fig. 5(b) demonstrates a successful removal of much of the noise seen in Fig. 4 while maintaining an excellent level of feature detection. The CSF-based thresholding applied here has its maximal effect at the high frequency scale(s). The visual system's high threshold at high frequencies resulted in preservation of only strong signal locations at these scales. It is assumed that a strong signal is more likely at a real edge or bar locations than at locations of mere noise. There are various sources for noise to be considered in the feature images. Some real luminance features or texture may be deemed by the observers as irrelevant for image or object recognition task. Our algorithm and any other similar luminance-based algorithm will detect such features in the same way that it detects relevant features. With such a "low level" algorithm there is no way to address such feature "noise." There is also noise in the images resulting from the imaging process and computational noise (ringing) due to the processing. These noise sources are expected to be of low magnitude in most images and thus will be effectively controlled by the CSF-based thresholding. In particular, such noise is excessive at the higher frequency channels as seen in Fig. 2 and is effectively reduced by the CSF-based thresholding, as seen in Fig. 5(a).

Note that the algorithm as implemented here has no free parameters that were fitted to the images or adjusted in any way. The image size is set arbitrarily, but as illustrated by the effective performance of the algorithm on the 1/16th size inset in the upper left corner of the images, the algorithm is effective across a wide range of sizes without adjustment. The threshold parameters used were derived from human visual measurement and were directly applied in the algorithm; no fitting or scaling was necessary.

To gain an appreciation of the performance of the visual algorithm, the result was compared with the performance of another feature detection algorithm [33]. This algorithm used polynomial transforms to detect edges by finding the local maxima in the gradient magnitude at multiple scales. An edge found in the smallest scale was maintained if it was also found in larger scales. This computationally intensive algorithm claimed no physiological basis. A best performance result for a 512×512 Lena image was provided by Martens and is shown in Fig. 6(a). It is clear that the algorithm has high sensitivity, but also high false alarm rate. The falsely detected edges are not only single noise pixels, but they also created false connections between real edges not seen in the original. When compared with the parameter free application of our algorithm, [see Fig. 6(b) and (c)], it is apparent that our algorithm is essentially free from this type of false connection artifacts. Note that the performance comparison is not made with respect to detection of signal (edge) and its differentiation from noise, as this performance may be equal or similar. The difference seen between the algorithms refers to the type of systematic errors represented by the false alarms produced by the other algorithm. In addition, note that our algorithm provided not just the edge position, but also the polarity of the features and a distinction between edge and bar features [see Fig. 6(b)].

C. Orientation-Selective Filters

The algorithm described so far applied isotropic filters with no orientation selectivity. While such isotropic processing takes place in the retina, it is established that in the

(a)

(b)

(c)

Fig. 6. Comparison of the feature detection of the polynomial transform algorithm of [33] (a) with the performance of our algorithm and (b) with nonoriented filters. (c) To facilitate comparison, this image provides all the features detected in (b), presented in single polarity. A higher resolution (512 × 512) was used here to match the resolution of the image provided by Martens.

visual cortex the processing cells or channels are orientation selective. It was, therefore, important to determine how the addition of the orientation selectivity to the algorithm would affect its performance.

To obtain orientation selectivity, the spatial frequency filters $C_i(f_x, f_y)$ in (3) were multiplied by oriented filters. An

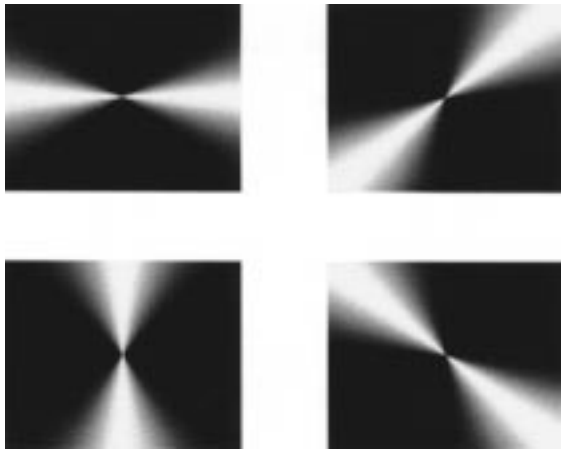


Fig. 7. Oriented filters used in the case of four orientations. These filters multiply the filters of Fig. 1 to generate the set of filters used in the orientation selective version of the algorithm. Note that the DC is in the center of these filters.

even number N of raised cosine filterswas generated [1]. The filter for the jth orientation was centered at $\theta = N\pi j$ and

$$O_j(\theta) = \begin{cases} 0.5\left[1+\cos(N\theta+\pi j)\right], & \frac{\pi}{N}(j-1) \le \theta \le \frac{\pi}{N}(j+1) \\ 0, & \text{elsewhere} \end{cases} \tag{8}$$

where $\theta = \tan^{-1}(f_y/f_x)$. For example, two orientations would consist of a horizontal and vertical direction, while four orientations are depicted in Fig. 7. As is the case with the isotropic filters, independent of the number of orientations, the total oriented filter response sums to unity over the 2-D spatial frequency domain. Thus, the bandpass-oriented filter $K_{i,j}$ that was applied at a scale i and orientation j was

$$K_{i,j}(f_x, f_y) = K_{i,j}(r, \theta) = C_i(r) \cdot O_j(\theta). \tag{9}$$

The algorithm described above was applied to each of the oriented bandpass-filtered images. Features detected in any of the orientations were represented in the final image. The results of processing two images with two, four, eight, and 16 orientations are compared with the nonoriented processing in Figs. 8 and 9. As seen in these figures, the addition of orientation selectivity increases feature sensitivity and improves performance, at least up to four orientations. At the level of eight orientations, some artifacts are apparent and many more are apparent at 16. From these results, it appears that the optimal performance is obtained between four to eight orientations. This finding is rewarding in view of the fact that psychophysical studies place the number of oriented channels in the visual system at about six to 12 [34]; [35]. Using a partial overlap as implemented here, the eight-orientation case corresponds to orientation bandwidth of 45°. De Valois *et al.* [36] found that the median orientation bandwidth for all cortical cells they measured was 41.9°.

D. Robustness to Variations in Filter Parameters

The algorithm described so far is strongly motivated by current vision models. However, it is clear that the filters

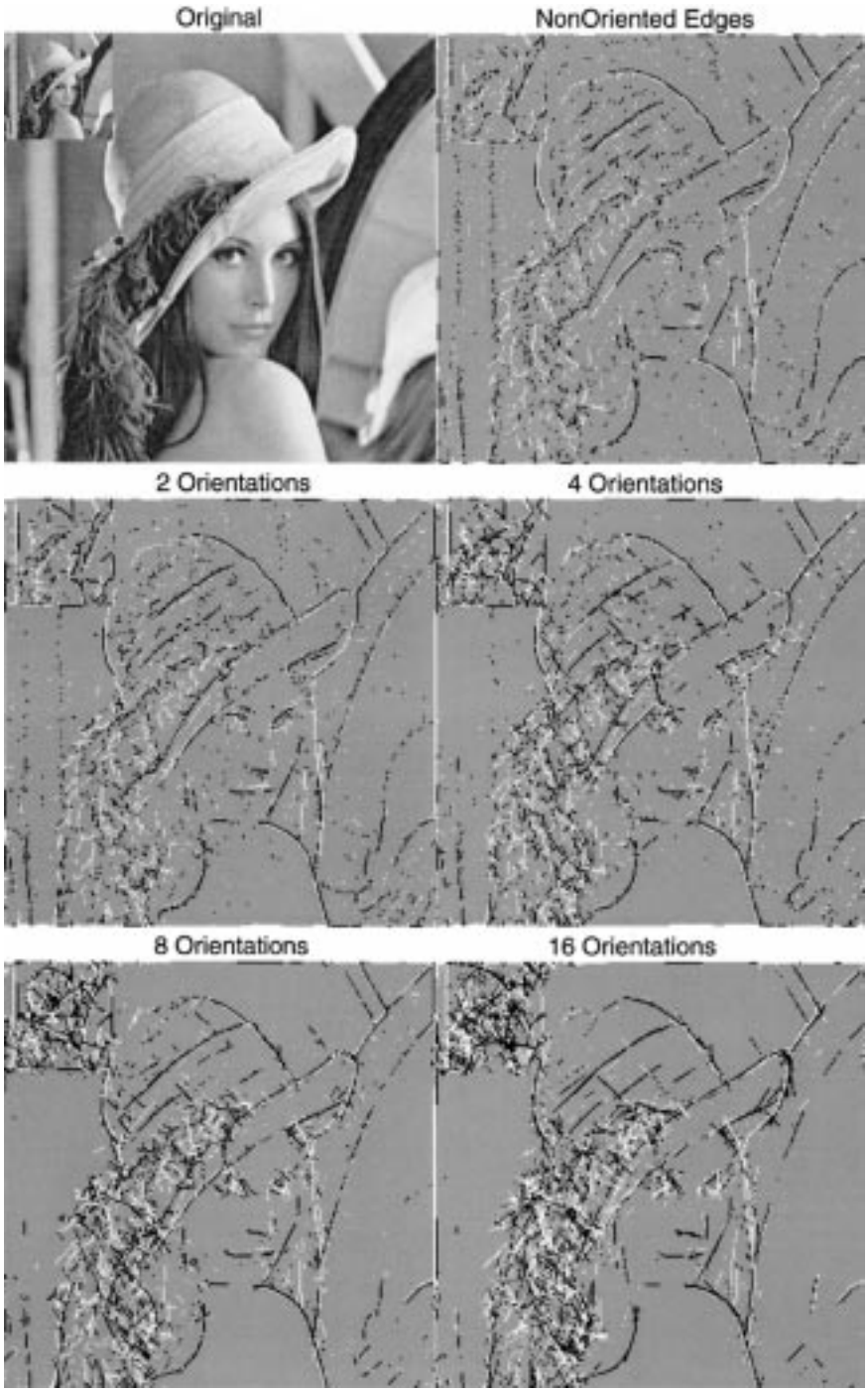


Fig. 8. Comparison of the performance of the nonoriented algorithm to the performance of the orientation selective algorithm at different number of orientations. Note that detection performance increases with an increased number of orientations. At a very high number of orientations (16), distortion artifacts reduce the algorithm performance. Note in particular the improvement and then decline for the small inset image.

implemented in the visual system are unlikely to represent such a uniform and accurate set as the one implemented here. The selection of the one-octave bandwidth and the one-octave separation between filter center frequencies was inspired by the biological data and was very convenient for computations. However, it is known that the bandwidth varies between channels [35]. It is also the case that the contrast sensitivity function, which represents the envelope of all the channels, varies between observers [37] and, thus, both center frequencies and channel bandwidth probably vary across observers. It was, therefore, of interest to examine, for the sake of the vision model and for the sake of the implemented algorithm, the effect of variation in these parameters on the performance of the algorithm. The impact of varying both the center frequency and the bandwidth was evaluated separately.

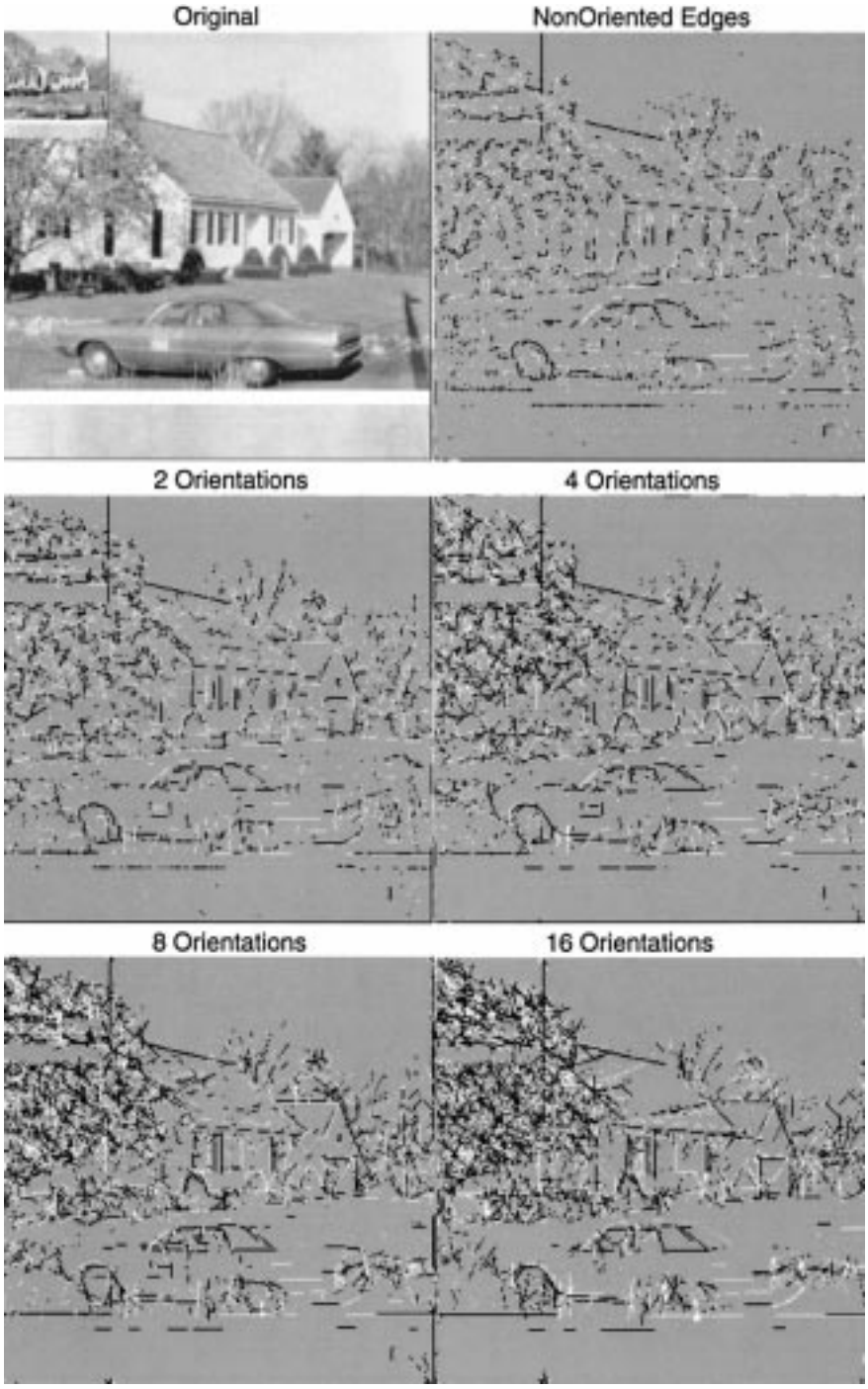


Fig. 9. Comparison of the performance of the nonoriented algorithm to the performance of the orientation selective algorithm at different number of orientations for the car image. Best result appears to be with four orientations in this case.

To test the effect of center frequencies, the filters were modified by shifting each filter's center frequency independently from the others. A uniform random distribution was used to shift each filter's center frequency and the processing was otherwise unchanged. The two sets of filters used are shown in Fig. 10. In this figure, the filters shown were shifted from their nominal frequency by as much as $\pm 25\%$ of the frequency range to the adjacent filters. The results of applying such randomly shifted filters to both the isotropic filters and the oriented (two orientations) versions were compared (Fig. 11) with the results obtained with the original deterministic set of filters. As can be seen, the features images derived with the randomly varied center frequencies filters are not identical, but are very similar to those derived using the deterministic filters. Most of the visually relevant feature pixels were detected in both cases and the main differences are in the noise pixels. Although there are minor differences between a few feature pixels detected with both sets of filters, the algorithm is clearly robust to such a large variation in the filters.

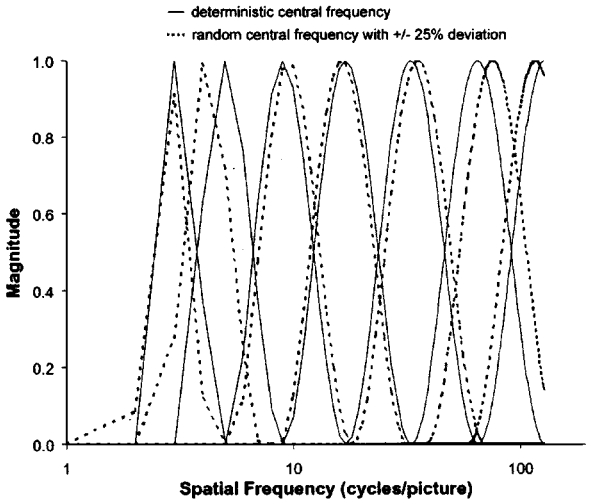


Fig. 10. Deterministic (solid lines) and randomly (dashed lines) shifted filter banks used to create the images in Fig. 11. Filters used were 2-D rotational bodies versions of the 1-D filters shown here. Only four of the filters—orders four to seven—were used. Deterministic filters are one-octave wide and one-octave apart, centered on spatial frequencies that are powers of two in terms of cycles/picture. Randomly shifted filters are the same filters shifted randomly in center frequencies either to the right or the left. As a result of that shift, they also vary slightly in bandwidth.

To determine the level of correspondence between the feature maps derived with the randomly shifted center frequencies, ten such images were computed. There were almost 56 000 feature points in all ten images, but the correspondence was such that only 16 000 locations (27%) were occupied all together (73% of correspondence). Of these feature points, 16% were in identical locations in all ten images (38% if a correspondence across at least five images were determined). The feature images created only by these corresponding points clearly represented the visually relevant details while the complementary—noncorresponding points created a noisy representation of the image. A visual assessment of the robustness to changes in the center frequency parameters can be derived also from the consistency of the images shown in Figs. 12 and 13 with the relevant visual features in the images.

To evaluate the impact of bandwidth variation on the algorithm, a wide set of filters was generated implementing a bandwidth variations in steps of 0.1 octaves between 0.2 and 2.0 octaves for each one of the center frequencies. The program enabled us to select the bandwidth for each of the four filters independently and observe the changes in the feature map generated. Experimenting with this program clearly illustrated that a significant variation in bandwidth resulted in fairly minimal changes in the feature detection, similar in magnitude to those illustrated in Fig. 11. Even with extreme variations in the available range, the algorithm continued to perform reasonably well in detecting and marking the various important visual features and their polarities.

To assess the level of correspondence across changes in filter bandwidth, ten feature images were created. In computing these images, each filter bandwidth was randomly varied between 0.5 and 1.5 octaves. The ten images yielded about 65 000 feature points in total, but only 16 000 locations were occupied (75% correspondence). Of these feature points, 19% were at the same locations in all ten images (45% of the points for correspondence across at least 5 images). In all cases, the corresponding points-generated feature map gave an easily recognizable image and the complementary noncorresponding points were presenting a noisy hardly recognizable image.

E. Effects of Variable Filter Sets

It is interesting to consider the consequence of having variable filters in the visual system. The filters applied by the visual system vary not only across observers, but also within one observer from place to place in the visual field. As both the voluntary larger scanning saccades and the involuntary micro eye movements scan a scene or an image, the same image feature is presented to different sets of filters. As the visual system integrates the perception derived from these separate glimpses, it is reasonable to assume that if it performs feature detection, and these feature maps are integrated as well, possibly by averaging them (though other combination rules are possible). To appreciate the effect of such processing, Figs. 12 and 13 illustrate the outcome. In these figures, ten trilevel feature images generated from randomly shifted filters, as shown in Fig. 11, were averaged. In this case, the filters were shifted by as much as $\pm 50\%$ of the range, so that two adjacent filters could be shifted to the same frequency. The result is a high-quality feature map in which the averaging filtered much of the noise in the separate maps out while the visually relevant features are maintained. These images may be considered a measure of the correlation between the feature images obtained with the various filters and the consistency of the features is an indication of the high level of consistency of the output despite the large changes in filters' center frequencies. In addition, the resulting averaged feature maps appear to have a higher resolution than the single feature maps (compare Fig. 12 with Fig. 11, for example). The averaging process of randomly generated feature maps permits, in effect, the development of feature maps with subpixel accuracy. This approach may be applied in image-processing applications and can provide feature detection with position accuracy at subpixel levels.

III. Discussion

A. Applications of Feature Detection in Vision

The families of image-processing algorithms that were most closely associated with a vision model justification or support were image enhancement and edge detection. The selection of these two categories for such support is intriguing as there is no direct proof or demonstration that the visual system is particularly good at either image enhancement or in generating edge maps or that it actually performs either of these tasks (although the phenomenon of Mach bands may be considered an indication of both functions). It is, therefore, interesting to consider the possible utility of edge detection in the visual system.

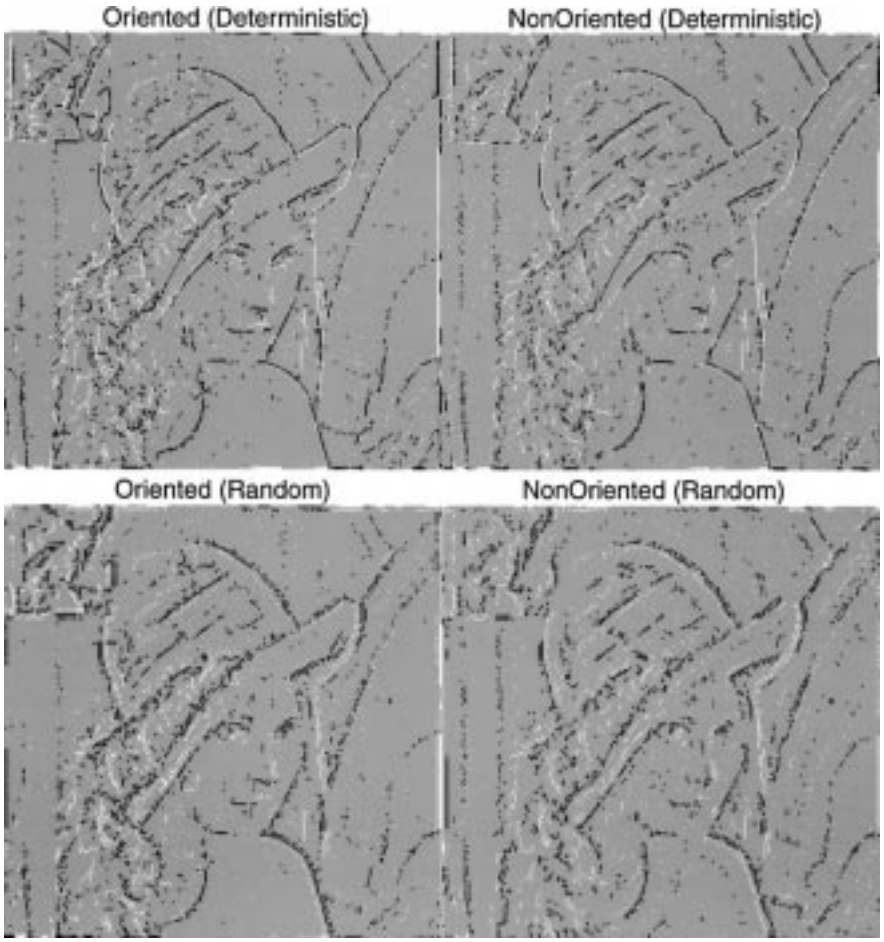
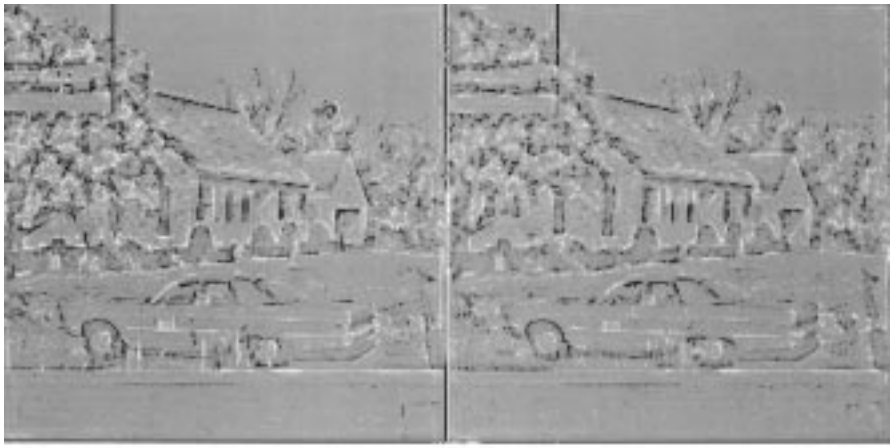


Fig. 11. Comparison of the algorithm feature detection output with the deterministic filters and the randomly shifted filters as shown in Fig. 10. Both isotropic and orientation selective (two orientations) are presented. As can be seen, the impact of the large variations in the filters on the algorithm performance is minimal in both cases.

(a) (b)

Fig. 12. Result of averaging ten separate features maps of the Lena image. Each features map is generated with a set of filters with center frequency randomly shifted as shown in Figs. 10 and 11. Note the increased resolution of the processed images apparent most clearly at the small minified inset image. (a) Two orientations. (b) Isotropic filters.

The application of contour and edge detection in computerized image processing is a long-standing practice with clear incentives related to the needs of segmentation of images [38]. Many other applications of edge detection in image processing have been reported [39]–[42]. The role or even existence of such capability in the visual system is less clear. Some visual system models use edge detection in a similar fashion as a precursor to image segmentation, in particular



(a) (b)

Fig. 13. Result of averaging ten separate features maps of the Car image. Same processing was applied as in Fig. 12. In this image, the benefit of two-orientation processing is more apparent. (a) Two orientations. (b) Isotropic filters.

with relation to modeling of "filling in" [43], [44] that requires an edge detection segmentation step. Many vision researchers seem to feel intuitively that the visual system performs an edge detection task and many studies evaluated object recognition from edges [45]. However, we are not aware of any direct indication that edge or bar detection is actually carried out in the visual system (although see [46] and [12]). It would, nevertheless, be interesting to consider what would be the possible utility for such function in the visual system and, in particular, what would the features detected by the algorithm presented here be especially useful for.

A frequently used image-processing function is the registration of images [47], [48]. Images are frequently taken in a sequence and need to be registered and aligned either to provide tiling of a larger image or to register images taken in different spectral bands [49]. The visual system might be considered to perform this task all the time and as frequently as five times per second. The nonuniform structure of the visual system provides for a wide field of view at low resolution and a very narrow central zone of high resolution. The high-resolution fovea is directed at objects of interest, detected by the periphery, using saccadic eye movements. Once the eyes move, the new retinal image has to be brought into registration with the previous retinal image to provide for a stable perception [50], though there are opposing views of this function that imply no requirement for such registration [51]. Image registration in the visual system is required also in the alignment of the binocular images into single vision and to provide stereo depth [52]. It is not known how the visual system performs this registration.

We would propose that the bipolar features derived from the current algorithm may be particularly useful for image registration and thus may serve such a purpose in the visual system. Typically, image registration algorithms use some similarity measure computed between two image segments that are shifted in relation to each other and the registration is defined as the shift position providing the maximum similarity [53], [54]. The performance of the algorithm is directly related to the sharpness of the maximum peak in the similarity function. Peli *et al.* [55] have shown that for a given algorithm, the performance improves if the similarity measure is computed from the original images, but only at the locations of image features (in one of the two images). This may be particularly useful if one image is at a lower resolution. In any case, if the similarity measure is to be computed for the edge images themselves it could provide an even sharper peak. The registration of the bipolar edge features detected by our algorithm will result in a particularly sharp similarity function, since a small deviation from correct registration will bring opposite polarity edge components together and will strongly reduce the similarity measure with small registration error. It is, therefore, plausible that if the visual system detects features such as edges and bars and if it applies an algorithm as described here, the result would lend itself well to the registration of images when needed in the visual system. Of course, the same advantage is available for use in computer image registration algorithms for various applications.

B. No Need for Quadrature Filters

The oriented filters applied in the algorithm were similar to cosine-phase Gabor filters (inphase filters). A number of vision models suggested that proper representation requires quadrature pair filters [20], [56], [57]. The quadrature sine-phase filter can be computed from the inphase filter using a Hilbert transform [58]. Existence of quadrature pairs of cells in the visual system was reported as well [59], [60]. It has been generally assumed that while the inphase filter or channel is needed for the detection of bar features, the quadrature (sine-phase) filters are needed for the detection of the edges, since the response of an inphase filter is zero just at the edge or luminance transition [12], [13], [19], [56], [61], [62]. In view of all this, it might be expected that the feature detection algorithm using the inphase filters would perform well for bar features but that sine-phase filters (asymmetric mechanisms) would be needed for the detection of edge features.

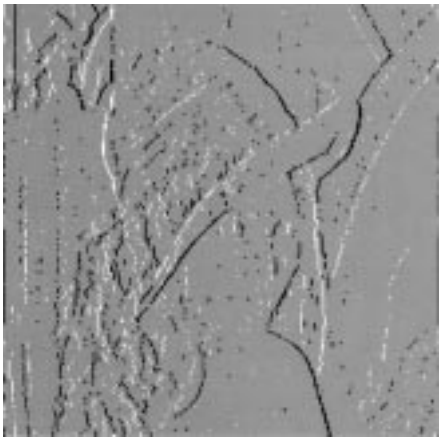


Fig. 14. Result of applying the algorithm to the quadrature filters (Hilbert transform of the oriented filters) in the vertical direction of the two-orientations case. Edges and bars are detected at about the same rate as with the corresponding inphase filters. Polarity of the detected features, however, is arbitrary and not related to the luminance transitions.

As seen above, the inphase filters detected both the bar and edge features. To determine if the use of sine-phase filters would improve edge detection, the algorithm was also applied with the use of sine-phase filters. The Hilbert transform commonly used to derive the quadrature filters is defined formally only for the 1-D domain, as it requires causality. An approximate expansion of the Hilbert transform into the 2-D domain is possible for oriented filters [57], [63], [64]. The Hilbert transform for our oriented filters design (two-orientation case) was computed by setting to zero the fast Fourier transform values of the filtered image in half of the frequency domain orthogonal to the direction of the filter. Following inverse transform to the space domain, the real part of the complex result is identical to the inphase filter and the imaginary part is the sine-phase filter output.

An example of the results obtained with the Hilbert transformed vertical filters is shown in Fig. 14. The Hilbert transformed results do not add significant detection of edges as compared with the inphase results. The main difference is that the polarity of the edge pairs obtained using the Hilbert transform is arbitrary and uniform across the image, independent of the direction of luminance change in the image. This arbitrary direction is set by the arbitrary decision regarding which of the two halves of the frequency domain is being zeroed. There is some additional noise in these images that results from the ringing caused by the sharp filtration associated with zeroing half the frequency plane.

It is clear from this example that the application of Hilbert transforms or quadrature pairs of filters is not necessary in the context of the algorithm described here. Quadrature filters are necessary within a single scale bandpass-filtered image to detect the location of edges in the original image [13], [16]. As noted by Jones and Palmer, "pairs of cells with cospatial receptive fields and quadrature relative phase angles could be used to represent stimulus position accurately regardless of the absolute phase angle of their response profiles" [65]. However, the integrated multiscale nature of our algorithm is able to derive the edges and their position from the multiscale inphase processing by identifying the phase congruence that occurs at both bar and edge locations.

C. Appearance of Images and the Features Maps

The output of the algorithm described here is a map of bipolar features consisting of pairs of bright and dark features at luminance edges and single features at bar or cusp locations. This output was derived from processing of the outputs of multiscale band pass (laterally inhibiting) filters. Davidson and Whiteside [66] have shown that the perception of image brightness near sharp contours does not conform satisfactorily to predictions of basic spatial interaction (lateral inhibition) models. They have demonstrated that the appearance was that of a bright and a dark features pair superimposed on the contours. Thus, they argued that perceived contours in images are shaped, to a large extent, just like images in which the detected bipolar features were superimposed over the original image. We are studying the effectiveness of superimposing the features maps on the original image as a method of enhancing images for visually impaired observers [62].

Davidson and Whiteside [66] were interested mostly in the apparent brightness farther away from the contour and did not address directly the appearance of these cusps, although they pointed out that most lateral interaction type models of Mach bands failed to predict them in the limit when the contour was very sharp. They concluded that, to account for the appearance of these contours, other mechanisms are needed. The edge detection mechanisms proposed here could be a model for such mechanisms responsible for the cusp perceptions they documented.

In conclusion, a visually based feature detection algorithm was developed and evaluated. It is based strictly on current vision models and makes use of visually determined parameters only. The algorithm is robust to filter parameter variations and in fact may benefit from such variations. The algorithm lends itself to computer vision applications that require such feature maps

ACKNOWLEDGMENT

The author would like to thank J. Nye and M. Butnaru, who provided programming help.

REFERENCES

[1] E. Peli, "Adaptive enhancement based on visual model," *Opt. Eng.*, vol. 26, no. 7, pp. 655–660, July 1987.
[2] A. C. Bovik, R. K. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 55–73, Jan. 1990.
[3] A. C. Bovik, "Analysis of multi channel narrow-band filters for image texture segmentation," *IEEE Trans. Signal Processing*, vol. 39, pp. 2025–2043, Sept. 1991.
[4] B. Chanda, B. B. Chaudahuri, and D. D. Magumder, "Some algorithms for image enhancement incorporating human visual response," *Pattern Recognit.*, vol. 17, no. 4, pp. 423–428, July–Aug. 1984.
[5] A. P. Ginsburg, "Visual information processing based on spatial filters constrained by biological data," Wright-Patterson AFB, OH, Aerosp. Med. Res. Lab. Rept. AMRL-TR-78-129, 1978.

[6] A. Toet, "Multiscale image fusion," in *1992 SID International Symposium Digest of Technical Papers*. San Jose, CA: Soc. Inform. Display, 1992, pp. 471–474.
[7] T. G. Stockham, "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, pp. 828–842, July 1972.
[8] E. Peli, *Visual Models for Target Detection and Recognition*, Singapore: World Scientific, 1995, vol. 1.
[9] B. S. Manjunath and R. Chellappa, "A unified approach to boundary perception: edges, textures, and illusory contours," *IEEE Trans. Neural Networks*, vol. 4, pp. 96–108, Jan. 1993.
[10] A. V. Oppenheim, R. N. Schafer, and T. G. Stockham, "Nonlinear filtering of multiplied and convolved signals," *Proc. IEEE*, vol. 56, pp. 1264–1291, Aug. 1968.
[11] W. Frei, "Image enhancement by histogram hyperbolization," *Comput. Graphics Image Process.*, vol. 6, no. 3, pp. 286–294, 1977.
[12] D. C. Burr, M. C. Morrone, and D. Spinelli, "Evidence for edge and bar detectors in human vision," *Vis. Res.*, vol. 29, no. 4, pp. 419–431, Apr. 1989.
[13] M. C. Morrone and D. C. Burr, "Feature detection in human vision: A phase-dependent energy model," in *Proc. R. Soc. Lond. B*, Dec. 1988, pp. 221–245.
[14] M. D. Heath, S. S. T. Sarkar, T. Sanocki, and K. W. Bowyer, "A robust visual method for assessing the relative performance of edge-detection algorithms," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 1338–1359, Dec. 1997.
[15] M. Heath, S. S. T. Sarkar, and K. Bowyer, "Comparison of edge detectors: A methodolgy and initial study," *Comput. Vis. Image Understanding*, vol. 69, no. 1, pp. 38–54, Jan. 1998.
[16] D. Marr and E. Hildreth, "Theory of edge detection," in *Proc. R. Soc. Lond. B Bio. Sci.*, vol. 207, Feb. 1980, pp. 187–217.
[17] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: Freeman, 1982.
[18] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 629–639, July 1990.
[19] E. Peli, "In search of a contrast metric: matching the perceived contrast of gabor patches at different phases and bandwidths," *Vis. Res.*, vol. 37, no. 23, pp. 3217–3224, Dec. 1997.
[20] J. G. Daugman, "Quadrature-phase simple-cell pairs are appropriately described in complex analytic form," *J. Opt. Soc. Amer. A*, vol. 10, no. 2, p. 375, Feb. 1993.
[21] M. A. Georgeson and T. C. A. Freeman, "Perceived location of bars and edges in one-dimensional images: computational models and human vision," *Vis. Res.*, vol. 37, no. 1, pp. 127–142, Jan. 1997.
[22] R. J. Watt and M. J. Morgan, "A theory of the primitive spatial code in human vision," *Vis. Res.*, vol. 25, no. 11, pp. 1661–1674, Nov. 1985.
[23] R. F. Hess and S. Dakin, "Absence of contour linking in peripheral vision," *Nature*, vol. 390, no. 6660, pp. 602–604, Dec. 1997.
[24] R. W. Rodieck, "Quantitative analysis of cat retinal ganglion response to visual stimuli," *Vis. Res.*, vol. 5, no. 11–12, pp. 583–601, Nov.–Dec. 1965.
[25] L. Maffie and A. Fiorentini, "The visual cortex as a spatial frequency analyzer," *Vis. Res.*, vol. 13, no. 7, pp. 1255–1267, July 1973.
[26] A. B. Watson, "The cortex transform: Rapid computation of simulated neural images," *Comput. Vis., Graphics, Image Process.*, vol. 39, no. 3, pp. 311–327, Sept. 1987.
[27] E. Peli, "Contrast in complex images," *J. Opt. Soc. Amer. A*, vol. 7, no. 10, pp. 2032–2040, Oct. 1990.
[28] H. R. Wilson and J. R. Bergen, "A four mechanism model for threshold in spatial vision," *Vis. Res.*, vol. 19, no. 1, pp. 19–32, Jan. 1979.
[29] E. Peli, L. Arend, G. Young, and R. Goldstein, "Contrast sensitivity to patch stimuli: Effects of spatial bandwidth and temporal presentation," *Spatial Vis.*, vol. 7, no. 1, pp. 1–14, Jan.–Mar. 1993.
[30] R. Goldstein and E. Peli, "Simulations of low-vision perception of images and the CSF," *Invest. Ophthamol. Vis. Sci.*, vol. 30, no. 3, pp. 397–397, Mar. 1989.
[31] E. Peli, J. Yang, and R. Goldstein, "Image invariance with changes in size: The role of peripheral contrast thresholds," *J. Opt. Soc. Amer. A*, vol. 8, no. 11, pp. 1762–1774, Nov. 1991.
[32] E. Peli and G. A. Geri, "Discrimination of wide-field images as a test of a peripheral-vision model," *J. Opt. Soc. Amer. A*, vol. 18, no. 2, pp. 294–301, Feb. 2001.
[33] V. Kayargadde and J. B. Martens, "Estimation of edge parameters and image blur using polynomial transforms," *Graph. Models Image Process.*, vol. 56, no. 6, pp. 442–461, Nov. 1994.
[34] G. C. Phillips and H. G. Wilson, "Orientational bandwidths of spatial mechanisms measured by masking," *J. Opt. Soc. Amer. A*, vol. 1, no. 2, pp. 226–232, Feb. 1984.
[35] H. R. Wilson, D. K. McFarlane, and G. C. Phillips, "Spatial frequency tuning of orientation selective units estimated by oblique masking," *Vis. Res.*, vol. 23, no. 9, pp. 873–882, Sept. 1983.
[36] R. L. D. Valois, E. W. Yund, and N. Hepler, "The orientation and direction selectivity of cells in macaque visual cortex," *Vis. Res.*, vol. 22, no. 5, pp. 531–544, May 1982.
[37] M. A. Garcia-Perez and E. Peli, "Lack of covariation of the effects of luminance and eccentricity on contrast sensitivity," *Optom. Vis. Sci.*, vol. 76, no. 1, pp. 63–67, Jan. 1999.
[38] J. M. Gauch, "Image segmentation and analysis via multiscale gradient watershed hierarchies," *IEEE Trans. Image Processing*, vol. 8, pp. 69–79, Jan. 1999.
[39] R. Alter-Gartenberg and R. Narayanswamy, "Image coding by edge primitives," in *Proc. SPIE Visual Communications Image Processing*, vol. 1360, pt. 3, Oct. 1990, pp. 1608–1619.
[40] A. Beghdadi and A. L. Negrate, "Contrast enhancement technique based on local detection of edges," *Computer Vis., Graphics, Image Process.*, vol. 46, no. 2, pp. 162–174, May 1989.
[41] T. Chen, H. R. Wu, and Z. H. Yu, "Efficient deinterlacing algorithm using edge-based line average interpolation," *Opt. Eng.*, vol. 39, no. 8, pp. 2101–2105, Aug. 2000.
[42] P. H. Eichel, E. J. Delp, K. Koral, and A. J. Buda, "A method for a fully automatic definition of coronary arterial edges from cineograms," *IEEE Trans. Med. Imag.*, vol. 7, pp. 313–320, Dec. 1988.
[43] S. Grossberg and D. Todorovic, "Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena," *Percept. Psychophys.*, vol. 43, no. 3, pp. 241–277, Mar. 1988.
[44] S. Grossberg and E. Mingolla, "Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading," *Psychol. Rev.*, vol. 92, no. 2, pp. 173–211, Apr. 1985.
[45] I. Biederman and P. Kalocsai, "Neurocomputational bases of object and face recognition," *Philos. Trans. R. Soc. Lond. B, Bio. Sci.*, vol. 352, no. 1358, pp. 1203–1219, Aug. 1997.
[46] M. A. Georgeson, "Human vision combines oriented filters to compute edges," in *Proc. R. Soc. Lond. B*, Sept. 1992, pp. 235–245.
[47] M. Svedlow, C. D. McGillem, and P. E. Anuta, "Image registration: Similarity measure and processing method comparisons," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-14, pp. 141–150, Jan. 1978.
[48] A. Venot, J. Y. Devaux, M. Herbin, J. F. Lebruchec, L. Dubertret, Y. Raulo, and J. C. Roucayrol, "An automated system for the registration and comparison of photographic images in medicine," *IEEE Trans. Med. Imag.*, vol. 7, pp. 298–303, Dec. 1988.
[49] M. L. Nack, "Temporal registration of multispectral digital satellite images using their edge images," in *Proc. AAS/AIAA Astrodynamics Specialist Conf.*, July 1975, pp. 1–72.
[50] G. W. McConkie and C. B. Currie, "Visual stability across saccades while viewing complex pictures," *J. Exper. Psychol. Hum Percept. Perform.*, vol. 22, no. 3, pp. 563–81, June 1996.
[51] J. K. O'Regan and A. Levy-Schoen, "Integrating visual information from successive fixations: Does trans-saccadic fusion exist?," *Vis. Res.*, vol. 23, no. 8, pp. 765–769, Aug. 1983.
[52] Y. Yeshurun and E. Schwartz, "Cepstral filtering on a columnar image architechture: A fast algorithm for binocular stereo segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 759–767, July 1989.
[53] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast image registration," *IEEE Trans. Comput.*, vol. C-21, pp. 179–186, Feb. 1972.
[54] A. Goshtashby, S. H. Gage, and J. F. Bartholic, "A two stage cross correlation approach to template matching," *IEEE Tran. Pattern Anal. Mahine Intell.*, vol. PAMI-6, pp. 374–378, May 1984.
[55] E. Peli, R. Augliere, and G. T. Timberlake, "Feature-based registration of retinal images," *IEEE Trans. Med. Imag.*, vol. MI-6, pp. 272–278, Sept. 1987.
[56] C. F. Stromeyer and S. Klein, "Spatial frequency channels in human vision as asymmetric (edge) mechanisms," *Vis. Res.*, vol. 14, no. 12, pp. 1409–20, Dec. 1974.
[57] E. Peli, "Hilbert transform pairs mechanisms," *Invest. Ophthalmol. Vis. Sci.*, vol. 30, no. 3, pp. 110–110, Mar. 1989.

[58] A. Papoulis, "Fourier transforms," in *Systems and Transforms With Applications in Optics*. Melbourne, FL: Krieger, 1968, vol. 3, pp. 61–98.
[59] V. A. Pollen and S. F. Ronner, "Phase relationships between adjacent simple cells in the visual cortex," *Science*, vol. 212, no. 4501, pp. 1409–1411, June 1981.
[60] D. A. Pollen and S. F. Ronner, "Visual cortical neurons as localized spatial frequency filters," *IEEE Trans. Syst. Man Cybern.*, vol. 13, pp. 907–916, Sept.–Oct. 1983.
[61] D. J. Tolhurst, "On the possible existence of edge detector neurons in the human visual system," *Vis. Res.*, vol. 12, no. 5, pp. 797–804, May 1972.
[62] E. Peli, "Wide-band image enhancement for the visually impaired," *Invest. Ophthalmol. Vis. Sci.*, vol. 39, no. 4, p. 398, Mar. 1998.
[63] N. K. Bose and K. A. Prabhu, "Two-dimensional discrete hilbert transform and computational complexity aspects in its implementation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 356–360, Aug. 1979.
[64] N. K. Bose, "Problems in stabilization of multidimensional filters via hilbert transform," *IEEE Trans. Geosci. Electron.*, vol. GE-12, pp. 146–147, Oct. 1974.
[65] J. P. Jones and L. A. Palmer, "The two-dimensional spatial structure of simple receptive fields in cat striate cortex," *J. Neurophysiol.*, vol. 58, no. 6, pp. 1187–1211, Dec. 1987.
[66] M. Davidson and J. A. Whiteside, "Human brightness perception near sharp contours," *J. Opt. Soc. Amer.*, vol. 61, no. 4, pp. 530–536, Apr. 1971.

Eli Peli received the B.Sc. degree (*cum laude*) and the M.Sc. degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 1976 and 1978, respectively, and the O.D. degree from the New England College of Optometry, Boston, MA in 1983.

He is currently a Senior Scientist with the Schepens Eye Research Institute, Boston, MA, and is also an Associate Professor of Ophthalmology with the Havard Medical School Department of Ophthalmology and an Adjunct Professor of Optometry and Visual Sciences with the New England College of Optometry. Since 1983, he has been caring for visually impaired patients as the Director of the Vision Rehabiliation Service at the New England Medical Center Hospitals in Boston, MA. He is a Consultant to companies in the ophthalmic instrumentation area, manufacturers of head-mounted displays, and national committees advising the National Institute of Health and NASA. He has authored or coauthored more than 75 scientific papers, the book *Driving with Confidence: A Practical Guide to Driving with Low Vision* (Singapore: World Scientific, to be published), and four U.S. patents and has edited *Visual Models for Target Detection* (Singapore: World Scientific, 1995). His current research interests include image processing in relation to visual function and clinical psychophysics, image understanding, evaluation of display-vision interaction, oculomotor control, and binocular vision.

Dr. Peli is a Fellow of the American Academy of Optometry and the Optical Society of America.