PERGAMON

Vision Research 38 (1998) 2323–2333

Vision Research

Medial-point description of shape: a representation for action coding and its psychophysical correlates

Ilona Kovács *, Ákos Fehér, Bela Julesz

Laboratory of Vision Research, Rutgers University, Busch Campus-Psychology Bldg., Piscataway, NJ 08854, USA

Received 3 January 1997; received in revised form 11 September 1997

Abstract

We describe a region-based shape representation that might be particularly useful from a biological perspective because it promotes the localization of objects, and object parts relative to each other. The proposed medial-point representation is similar to medial-axis type representations, but it is more compact. The medial points are those points along the medial axis that are equidistant from the longest segments of the boundary, therefore they represent the largest amount of edge information. The main advantage is that the original image can be reduced to a small number of points. We also provide psychophysical correlates of the representation for shapes with increasing complexity. Using a reverse mapping technique, we find that variations of contrast sensitivity within figures are defined by the shape of the bounding contour, and the peaks in the sensitivity maps correspond to the medial points of the proposed representation.

Keywords: Shape; Medial-axis; Perceptual organization; Spatial interaction; Contrast sensitivity

1. Introduction

Capturing time is the essence of the recognition of events in the world. What would be a suitable representation to capture time for the perception of the complex movements of living things? The criticality of the representation has been nicely illustrated in the history of photography. As Fig. 1a shows, it is not easy to depict the relationships between various body parts in time and space within a single representation such as a static photographic plate. A clear expression of this problem, and an original solution to it was provided by E.-J. Marey, a French medical doctor of the last century: ''In this method of photographic analysis the two elements of movement, time and space, cannot both be estimated in a perfect manner. Knowledge of the positions the body occupies in space presumes that complete and distinct images are processed; yet to have such images, a relatively long temporal interval must be had between two successive photographs. But if it is the notion of time one desires to bring to perfection, the only way of doing so is to greatly augment the frequency of images, and this forces each of them to be reduced to lines.'' [1]. To prepare for his chronophotographs (an example is shown in Fig. 1c), he dressed his subject in a black costume, and marked the joints with shiny buttons connected with metal bands (Fig. 1b). By selecting what he considered the most informative points and lines, he was able to read the successive postures of the body on his plates, and follow the important trajectories of motion. The relevance of this solution for human vision was later confirmed by G. Johansson's work on the perception of biological motion in the point-light walker displays [2].

Beyond historical examples, the purpose of this paper is to suggest a concise representation of visual shape, that can be used to capture complex movements of complex shapes. The representation belongs to the class of skeletal representations introduced by Blum [3–5]. With our procedure described below, we would like to replace the skeleton-type descriptions by an even more compact, pruned shape representation which is, at the same time, suitable for implementation in a parallel neural network.

We use our simulations that transform images into the suggested type of representation to predict psychophysical data. Originally, our psychophysical experi-

* Corresponding author. Fax: +1 732 4456715; e-mail: http://zeus.rutgers.edu.

0042-6989/98/$19.00
PII: S0042-6989(97)00321-0

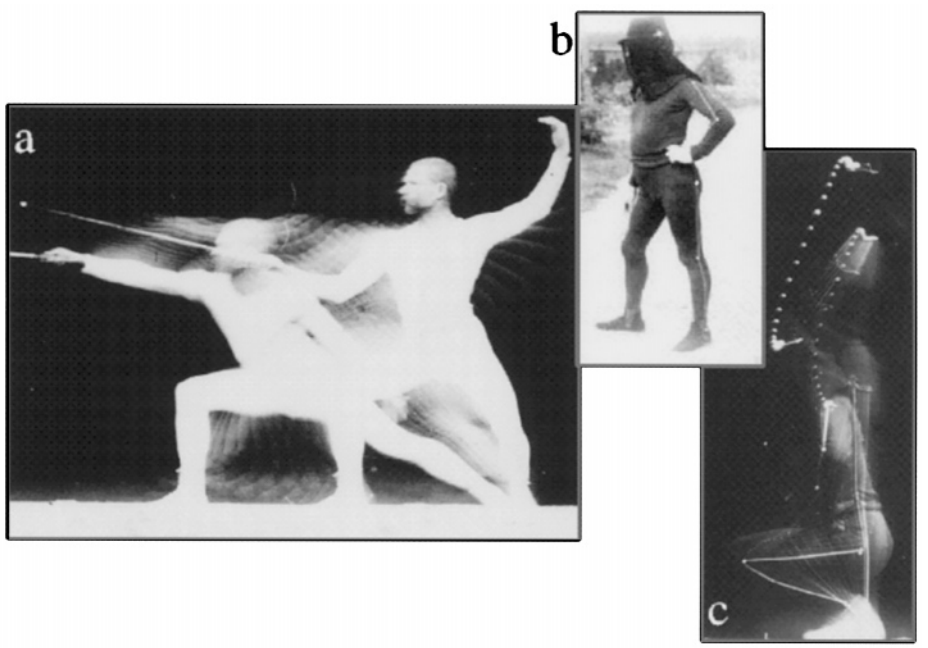

Fig. 1. Photographs taken by E.-J. Marey [1]. (a) Movement of a fencer. (b) A subject dressed in black preparation for geometric chronophotography. (c) Jump in place with bent knees. (The figure is a composition of pictures from Marta Braun's book: Capturing Time, 1992 [40] ; with the permission of Chicago University Press).

ments were planned to find correlates of perceptual organization at the level where spatial interactions occur among early cortical filters. The architecture of pairwise spatial interactions has been described by Polat and Sagi [6,7]. The activity pattern of a large number of interacting filters has been described by Kovács and Julesz [8–10] and Kovács [11]. These latter studies provided evidence that the interactions are sensitive for shape properties, and can mediate a medial-point-type representation.

In the following, we elaborate the idea of the medial-point representation in Section 2, and in Section 3 we provide further psychophysical correlates to support the biological relevance of the representation.

2. The medial-point representation

Our purpose is to provide a natural description of visual shape, which is: (a) *local* (although it should reflect global shape properties), therefore promotes fine localization of objects; (b) *compact*, therefore promotes the comparison between internal representations and the viewed image; and (c) can be easily applied at *different spatial scales*, therefore allowing for the selection of the desired spatial detail, and proper segmentation of parts. The selection of these aspects is based on our pragmatic view that emphasizes the usefulness of the representation in actual perceptual tasks.

A suitable class of descriptors is provided by the skeletal representation of shape, that was introduced by Blum [3–5] as the result of the grassfire transformation. The grassfire transformation has several modern versions, also known as the medial-axis transformation or MAT [12]. Among the advantages of the MAT are that it is compact and it captures important properties, such as symmetry and complexity, in a translation- and rotation-invariant manner. It is, however, very sensitive to perturbations along the boundary of the shape [13]. We claim that this is not necessarily a drawback, because it is this kind of sensitivity that provides an opportunity to build up a good multiresolution representation. The real flaw from our point of view is that although the skeletal representation reduces the amount of information to be stored about the shape by keeping only the symmetric axes, it is still not local enough to promote fine localization. In other words, defining the position of a line in space (such as that of a skeletal line) is still difficult as compared to defining the position of a point.

Our procedure is intended to advance the MAT-type descriptions by increasing locality and compactness, and by easing the computational costs of a multi-scale representation. The goal is to find the most informative points along the skeleton (medial axis) of a shape. This goal is achieved by computing the D_ϵ-function, which is illustrated in Fig. 2. The D_ϵ-function is based on an equidistance metric. The D_ϵ value of an internal point represents the degree to which this point can be considered as the center of the local boundary segment around it. D_ϵ is defined for each internal point by the percentage of the boundary points which are equidis-

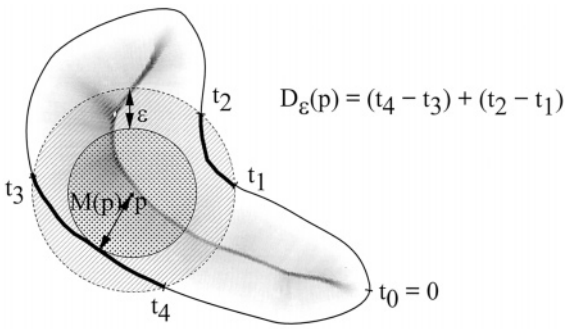

Fig. 2. The D_ϵ-function within a simple shape. The D_ϵ-function integral can be obtained as the sum of the curve segments falling inside the ϵ neighborhood (thick boundary segments within the grey ring) of the $M(p)$ radius circle (dark-grey area) around p. Dark shading corresponds to increasing values of D_ϵ, and the white spot denotes the maximum of the function.

tant from the internal point within a tolerance of ϵ. See the mathematical formulation of the D_ϵ-function in the Appendix A.

Fig. 2 presents the D_ϵ-function as a grey-level map of the D_ϵ values within an arbitrary shape. Dark shading corresponds to increasing values of D_ϵ, and the white spot denotes the maximum of the function. With this transformation we obtained a non-uniform skeleton with a peak value. A peak is an important point, it is equidistant from the longest segment of the boundary, therefore it *represents the largest amount of edge information*. The number of peaks within an object will be defined by the shape of the object and by the thresholding operation that finds the local maxima of the D_ϵ-function. We call *the local maxima of the D_ϵ-function medial points*. The classical MAT does not provide medial points, because the points along the skeleton are not differentiated from each other in terms of their value. The points along the skeleton of a curved object (Fig. 3a) or along a worm with cusps (Fig. 3b) are equivalent in the classical case, while the value of D_ϵ is higher where the smoothly curving boundary approximates circular arcs.

The medial-point representation (MPT) obtained with the D_ϵ-function is sensitive to details along the boundary, just as the classical skeletal representation. That sensitivity helps us to obtain a multiscale MPT. In fact, multiple scales are inherently included in the transformation, and no preprocessing of the boundary is required (although the boundary has to be extracted before the suggested computation is applied). The value of parameter ϵ (Fig. 2) defines the resolution of the D-function. The larger the value of ϵ is, the fewer details will be represented, as it is illustrated in Fig. 4.

We have been successfully using the D_ϵ-function as a model to predict psychophysical data [8–10]. Here we will extend these studies, and directly test the power of the D-function to predict human contrast sensitivity maps of different shapes. We include shapes of different elongations, a shape that has a curved skeleton, and a shape that has a branching skeleton. If our model is able to provide good predictions for shapes with increasing complexity, that should indicate the biological relevance of the sparse skeletal description.

3. Psychophysical correlates

Our original purpose with the psychophysical experiments was to see whether there are correlates of perceptual organization and explicit shape representation at the earliest levels of cortical processing. Based on the correlation between psychophysical evidence on the properties of spatial filters and physiological data on receptive field properties in V1 [14,15], we presume that the first cortical representation of the visual image is mediated by linear spatial filters, such as orientation-

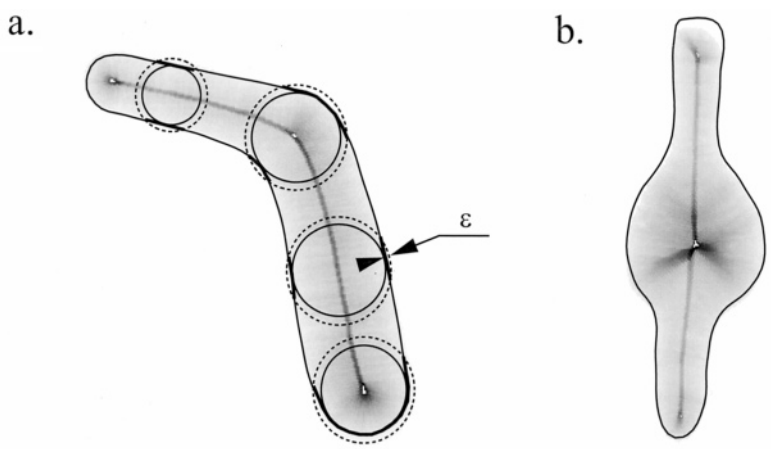

Fig. 3. (a) The D_ϵ-function within a curved object. The maxima of the function (white spots) occur on those locations where the largest portion of the boundary is equidistant from an internal point (compare the length of the thick lines within the rings defined by ϵ at different locations within the figure). (b) The thick part of the 'worm' is represented by a high peak at its center.

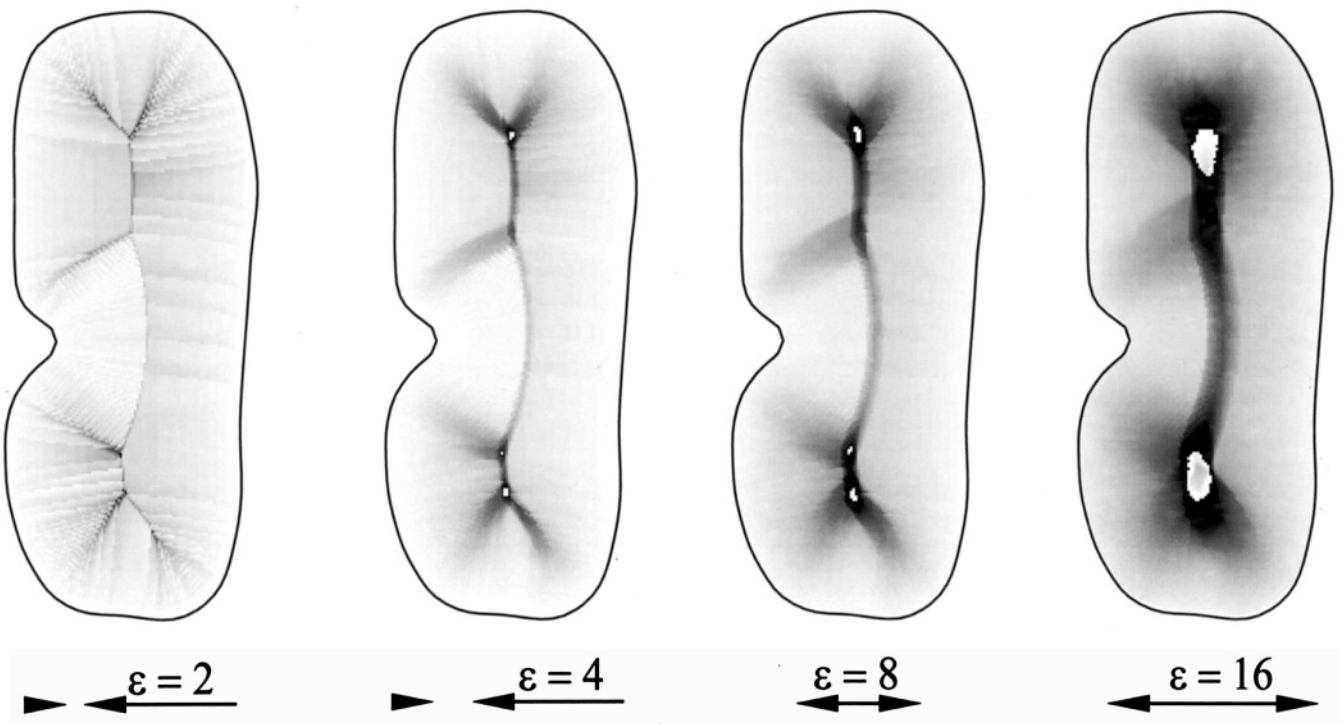

Fig. 4. The D_ϵ-function at different spatial scales as defined by the value of ϵ (ϵ is indicated by the arrows).

specific Gabor filters [16] of different sizes. Since these filters are spatially localized, and respond to input coming from a limited retinal region, spatially extended image patterns involve interactions between them. The architecture and dynamics of lateral interactions between oriented spatial filters have been studied in a contrast masking paradigm by Polat and Sagi [6,7], where local contrast sensitivity is measured for a foveal Gabor target in the presence of two masking Gabor patches. Three fundamental interaction types have emerged from these lateral masking studies: spatially short-range, non-linear, suppressive interactions; spatially longer range, 'side-way' orientation selective facilitatory interactions (parallel arrangement of target and maskers); and particularly strong, orientation-selective facilitatory interactions, extending along the preferred orientation of the filters to large spatial distances (colinear arrangement of target and maskers). Grouping and figure-ground segregation might be subserved by the longer-range facilitatory interactions [6].

Subsequent studies have shown that the pattern of spatial interactions depends as much on past experience as on the actual stimulation, thus the functional architecture of these interactions is not fixed, it is dynamic and context dependent [17–19]. We have also demonstrated that the known types of orientation specific pairwise interactions might not be enough to explain even such a simple event as the perceptual integration of closed contours [8]. The description of the neural activity pattern in response to more complex stimulus arrangements is required.

Simultaneous activity of 'many' interacting neural elements can be revealed by tracking the activity of several units simultaneously in search of their higher-order correlations, such as in electrophysiological cross-correlation and multiunit studies. An alternative is to estimate how the activity of one unit is affected in the context of the activity of other units. We use the latter approach in the reverse mapping technique, where the activity of one unit is measured as a function of the gradually changing context. In order to examine the global interaction pattern of a large number of local spatial filters, we introduced the psychophysical reverse mapping technique [8–11]. Psychophysically measured local contrast sensitivity reflects the local activity of single units, and the context of the interaction pattern is being manipulated by changing the stimulus pattern.

3.1. Method: reverse mapping

We measured local differential contrast thresholds of human observers for a single Gabor target that was surrounded by a background of randomly oriented and positioned Gabor patches and by a contour, forming different shapes. The only cue that defined the path of the contour was spatial continuity across gaps, as it is shown in Fig. 5. That ensured that the perceptual integration of the figure is mediated by orientation specific interactions among a limited set of spatial filters. Contrast sensitivity maps of different shapes were obtained by measuring thresholds at different

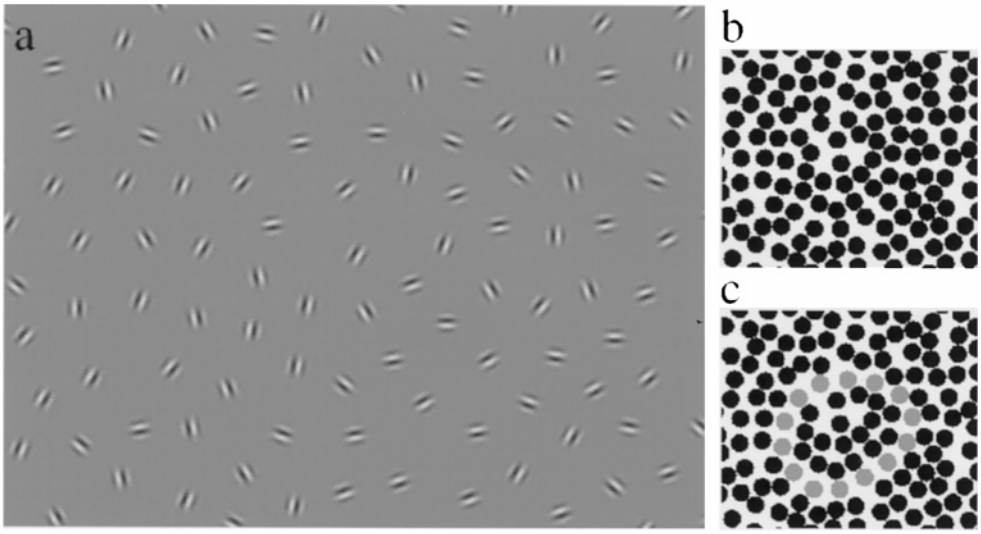

Fig. 5. Example of the stimulus context that we use in the reverse mapping technique. (a) A closed contour embedded in noise (these types of contours were introduced earlier by Field et al. [27] and Kovács and Julesz [8] in contour detection experiments). (b) The Gabor elements were replaced by non-oriented elements (and the size of the image was reduced). Once orientation is removed, the contour becomes invisible. (c) For reference, we added luminance contrast to the invisible contour in (b). (The actual ratio of element spacing in the noise background and spacing along the contour was 0.83. As long as this ratio remains smaller than 1, the contour remains invisible in (b).)

relative target-contour distances, as illustrated in Fig. 6. In other words, the shapes were moved around the target, and in separate experiments, contrast thresholds were estimated for a large number of locations within and outside of the shapes. Contrast thresholds on each of these locations were compared to the baseline threshold that was estimated in the absence of the contour. The difference between baseline and actual threshold was plotted, resulting in an activity surface (or sensitivity map) for each shape with valleys and peaks of sensitivity change (Figs. 7–9).

Stimuli were presented as grey-level modulation on a Mitsubishi color monitor, controlled by a Silicon Graphics Iris Indigo R 4000 machine. The display had a mean luminance of 32 cd/m^2. The patterns were viewed from a distance of 90 cm, and each pattern subtended 16° × 16° of visual angle. As described earlier [9], stimuli were composed of Gabor patches of the same parameters, where the wavelength (λ) of the Gabors was 0.18°, and the Gaussian envelop size was equal to λ. Signal amplitude was 16% of the mean luminance for all the Gabor patches, except for the target Gabor patch that had a varying contrast.

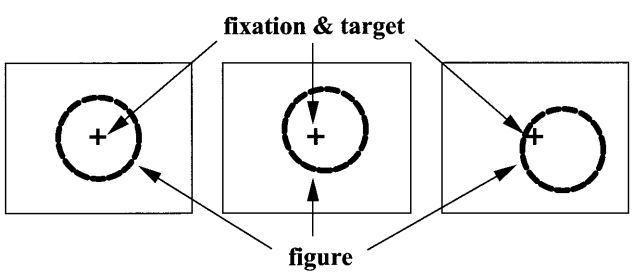

Fig. 6. Relative target-contour positions in different blocks of the reverse mapping experiment. Differential contrast thresholds were measured for a Gabor target (its position is marked by the +) in the presence of a closed contour and random noise (only the contour is shown here schematically; an example of the contour embedded in noise is illustrated in Fig. 5.)

Two sequential frames of one trial (following a central fixation mark) both contained a random background of Gabor patches, an embedded contour and a target. The two frames were equivalent except for the contrast of the target. The target was a centrally positioned Gabor patch with a contrast increment from the background in one of the two frames, and its orientation was parallel to the closest segment of the embedded contour. The orientation of the target (and the contour) was randomized across trials. The background contained 2500 randomly positioned and oriented Gabor patches, and the contours were composed of 16 Gabor patches. We applied regular circle shaped, ellipse shaped (with varying aspect ratios, ranging between 1.2 and 1.8), cardioid shaped and triangle shaped contours.

Contrast thresholds for the target were measured as a function of target-contour distance in a two-alternative temporal forced-choice procedure. One trial consisted of two successive presentations of 170 msec stimulus frames (interframe interval was 500 msec), with the target presented either in the first or second frame. Contrast thresholds were estimated by a staircase procedure [20], 4–8 times for each relative target-contour distance. Contour-target distance was kept constant during one block. To determine the effect of the circular contour on target sensitivity as a function of contour-target separation, we used at least eight different

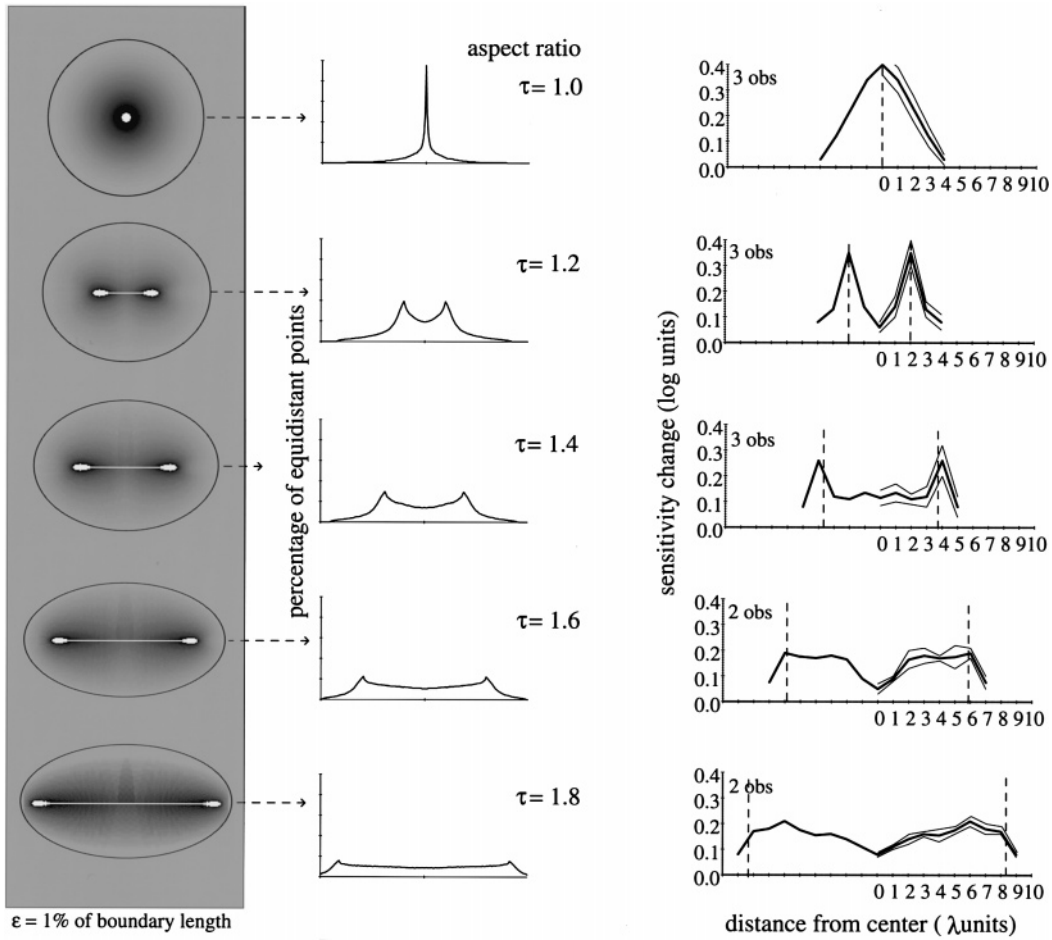

Fig. 7. Simulation and psychophysical results for the medial-point representation. Left column: 2-D grey-level maps showing the values of D_ϵ inside a set of ellipses. Dark shading corresponds to increasing values of D_ϵ, and the white spots denote the maxima of the D_ϵ-function. ϵ was 1% of the length of the contours. Middle column: 1D sections along the main axes of the ellipses. The percentage of equidistant points (D_ϵ) is expressed on the ordinate, and the abscissa represents locations along the axes. (The D_ϵ-function was plotted by the DAVID visualization software, Fluid Sciences Inc. 1992.) Right column: Contrast sensitivity change as a function of target-contour distance along the main axes of the ellipses. Dashed lines correspond to the peaks predicted by the D_ϵ-function. Thin lines show the standard error. (Results for aspect ratio 1.0, 1.2 and 1.4 are replotted from Kovács and Julesz [9].)

measuring sites with one target period (λ) resolution along one or two sections across the figures. We computed the relative change in sensitivity as compared to the sensitivity for the probe in the absence of the contour.

One of the authors (I. Kovács), and a total of seven naive observers participated in the experiments.

3.2. Results

We reported earlier [8,9] that there are two separate enhancement regions within the contrast sensitivity map of a regular circle. The first region is related to the boundary and the second is related to the region bounded by the contour. At short spatial distances from the contour, sensitivity change is similar to detection threshold variations in the paradigm of lateral masking. Sensitivity is reduced when target and contour overlap; and a facilitatory effect occurs symmetrically on both sides of the contour with peaks at 2λ distance from the perimeter (for comparison, see Fig. 6, Ref. [8], and Fig. 2, Ref. [6]). These short-range effects can be interpreted by within-filter interactions. The second enhancement region is far from the perimeter, and can only be observed within the contour. Here we only report sensitivity changes corresponding to the second, interior-specific enhancement region.

The right column in Fig. 7 depicts data for a set of ellipses with aspect ratios ranging between 1.0 and 1.8. Corresponding 2D and 1D simulation results are in the left, and middle columns, respectively. Dark shading corresponds to increasing values of D_ϵ, and the white spots denote the maxima of the function. For clarity, we also provide 1D cross-sections, where the percentage of equidistant points is on the vertical axes. The data graphs in the right column plot contrast sensitivity

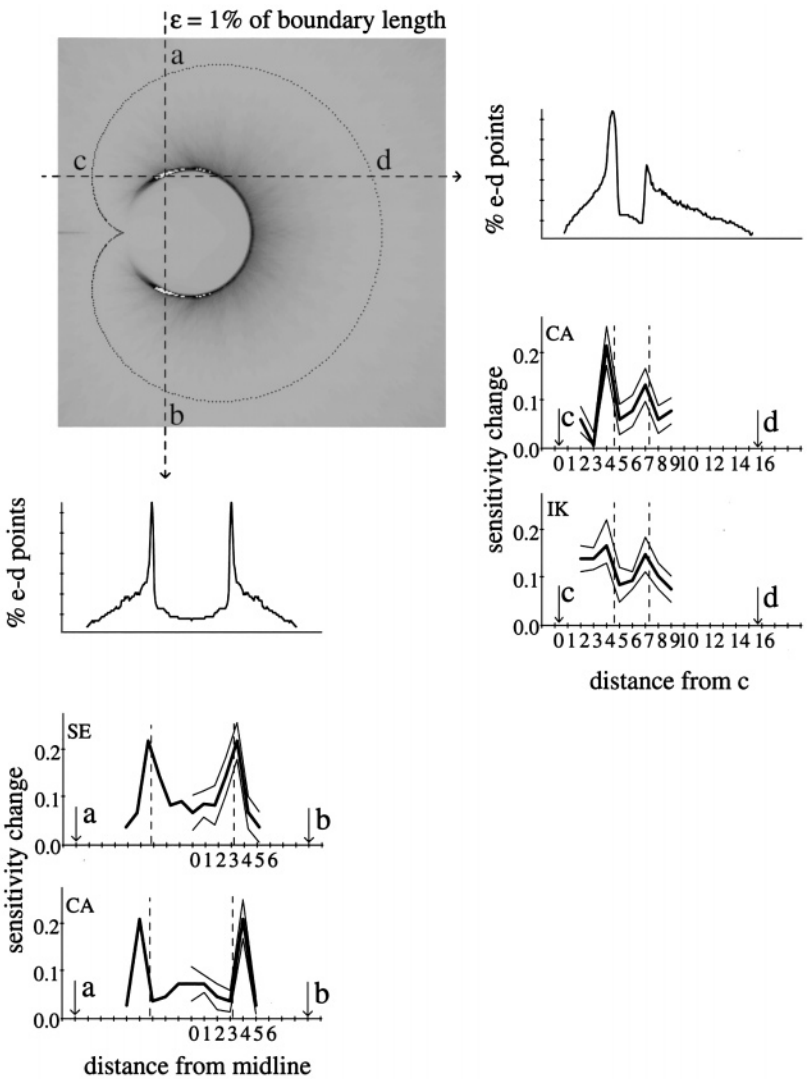

Fig. 8. Same as in Fig. 7 for a cardioid shape. 1D cross-sections were taken along the *a–b* (see graphs below grey-level map) and *c–d* (see graphs to the right from the grey-level map) directions. *e–d* on the ordinate of the 1D plot of the simulation results refers to 'equidistant.' (Results of observer SE are replotted from Kovács, [11].)

change as a function of target-contour distance along the main axes of the ellipses. Results are averaged across observers. Zero on the horizontal axis indicates the center of the shape, and the distance of the probe from the center is expressed in Gabor wavelength (λ) units. Data were obtained only along one radial direction, and the same data were reflected along the other direction. The vertical dashed lines correspond to the maxima of the medial axis computations (peaks in the left panels), in other words, they show the predicted peak locations. For aspect ratios of 1.0, 1.2 and 1.4 there is an extremely good fit between simulation and experimental results. The predicted and the measured peaks overlap within measurement sampling error (1λ). For more elongated ellipses, the amplitude of the sensitivity change also has to be considered, not only the location of the peaks. The 1D model results show that increased aspect ratio (with a constant perimeter) results in both a larger separation between the peaks, and in lower peak amplitude. The more elongated the shape is, the fewer boundary points will be equidistant from the main points of the medial axis. In fact, the high peaks become replaced by long medial axes for very elongated shapes. Because of that, it is more difficult to match the predicted and measured peak locations for aspect ratios of 1.6 and 1.8. However, the amplitude changes do match the predicted tendencies. We conclude that for shapes with a single and straight medial axis the measured contrast sensitivity changes can be well predicted by the simulation results of the medial-point representation.

Sensitivity change data for a cardioid shape are shown for two observers along two cross-sections in the lower two panels of both the left and right columns in

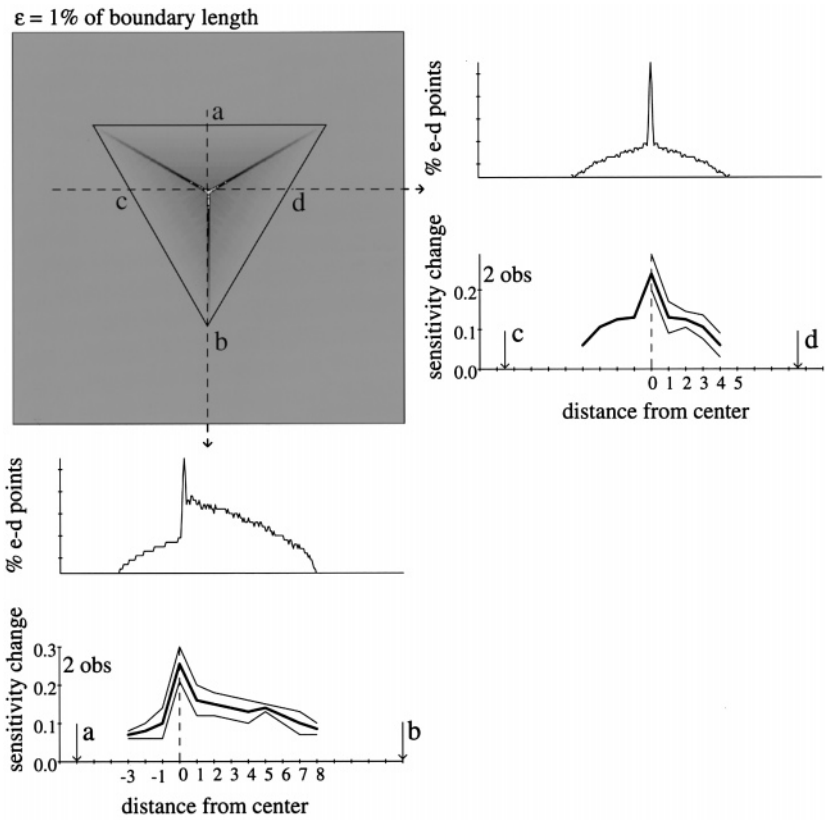

Fig. 9. Same as in Fig. 7 for an equilateral triangle. 1D cross-sections were taken along the $a–b$ (see graphs below grey-level map) and $c–d$ (see graphs to the right from the grey-level map) directions. $e–d$ on the ordinate of the 1D plot of the simulation results refers to 'equidistant'.

Fig. 8. The 2D plot of the simulation results is in the upper left panel, and the 1D plots corresponding to $a–b$, and to $c–d$ directions are shown next to the 2D map in the top panels of the left and right columns. Data are plotted as in Fig. 7, but the results are shown separately for each observer. Zero on the horizontal axis indicates the half distance between a and b along the $a–b$ direction, and c along the $c–d$ direction. Data were obtained only along one direction, and the same data were reflected along the other direction in the $a–b$ cross-section. The fit between simulation and experimental results is good both in terms of peak locations, and in terms of the amplitude of the sensitivity changes. The high peaks are well predicted by the simulation along the $a–b$ section, and along the $c–d$ section we have evidence for the presence of secondary peaks within a shape with a curved medial axis.

Sensitivity change data for an equilateral triangle shape are shown for two observers along two cross-sections in the lower panels of both the left and right columns in Fig. 9. Simulation and experimental data are plotted as in Figs. 7 and 8. Results are averaged across observers. Zero on the horizontal axes indicates the symmetric center of the triangle. Data were obtained only along one direction, and the same data were reflected along the other direction in the $c–d$ cross-section. Along the $a–b$ section we find an exact match between the predicted and measured peak locations (sampling error was reduced by placing a sampling point exactly at the predicted peak location), and the peak is very clearly evidenced along the $c–d$ direction as well. We conclude that the measured contrast sensitivity changes can be well predicted by the simulation results for a complex shape with a branching medial axis.

4. Conclusions

In our search for a natural and effective shape description we found a representation that might be employed to simplify the classical skeleton-type representations. The suggested representation describes shapes by a small number of medial points that are important in terms of representing maximal portions of the boundary. The representation is local, compact, and nicely works on multiple spatial scales. These properties make it attractive from the point of view of real perceptual tasks to be solved in biological systems. With respect to neural computation, an important aspect of the medial-point representation is that it reduces the redundancy of the image, and small cell assemblies can

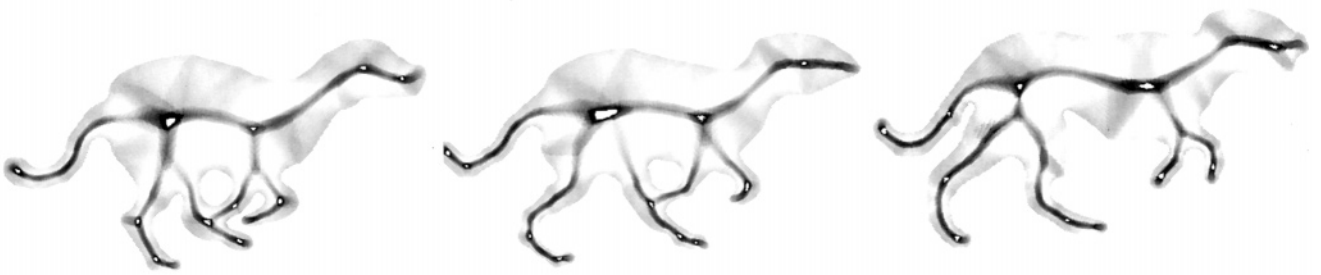

Fig. 10. The D_ϵ-function for sequential frames of the movement of an animal. The maxima of the function are good candidates as primitives for biological motion computations.

carry information about large, extended objects. The medial points can be followed easily between frames of, for instance, a biological motion sequence (Fig. 10). It also seems to be extremely straightforward to implement this transformation in a neural network with spreading activation, where the equidistance metric can be reinterpreted in terms of temporal synchrony of neural events [21].

The suggested representation integrates the advantages of curvature extrema computations (an edge-based representation, exemplified by Attneave [22] and Hoffman and Richards [23]) and medial-axis computations. Curvature extrema may provide the most important *points* along the boundary, however, all the drawbacks of an edge-based representation have to be encountered, and the spatial organization of the object components is not made explicit. The medial-axis is a region-based representation, however, it is not as local as the point-based curvature extrema computation. The medial-point representation has the benefits of both representations: it defines local points, and it is region-based. Notice the other advantage of our simple procedure: symmetry and curvature can be taken into account at the same time (see the discussion on the duality of symmetry and curvature by Leyton [24]).

The medial-point representation bears similarities with the annular symmetry representation of Kelly and Levine [25], and with the 'core' representation proposed by Burbeck and Pizer [26]. The machine vision system described by Kelly and Levine [25] produces a symmetry categorization on the internal points of a figure, discriminating global center-symmetric points of blob-like shapes, and axial-symmetric points of limb-like shapes. As a result, a rough classification of objects and object parts is obtained. In spite of the similarity in using annular operators, our local circularity measure provides a detailed description of shapes, not only a rough classification as the Kelly and Levine model. In the core representation of Burbeck and Pizer [26], the skeleton is weighed according to the width of the object, and the result is a fuzzy axis at multiple spatial scales. From a machine vision point of view, the Burbeck and Pizer model has a great advantage: it works on grey-level images, not requiring segmentation of object contours. Our process, however, requires the definition of object contours first. The biological plausibility of our process is supported by psychophysical results regarding the detection of gapped contours embedded in noise [8,27–29], that seem to indicate that edge definition should precede higher-level shape related computations. Another difference is that the Burbeck and Pizer model assumes separate medialness detectors that would be activated by boundary detectors of the same scale, but all possible orientations. That implies that neural correlates of this kind of connectivity should be found at a very early level (since the core computation precedes any contour definition). However, the existing neural connectivity at the level of the primary visual cortex may only support contour integration processes by linking neurons of similar orientation tuning [30]. Finally, although the result of our computation and the core computation may look similar at an intermediate spatial scale, the medial points are made explicit only in the former case.

In our search for psychophysical correlates of perceptual organization we found very specific changes in contrast sensitivity according to the context of different visual shapes. The reverse mapping technique provided a good tool to see these sensitivity changes within different shapes in detail. We found that contrast sensitivity is highly affected by shape in this paradigm. The peaks in the sensitivity map are extremely well predicted by the D-function that we use to create the medial-point representation. Our results provide correlates of the representation for shapes with increasing complexity.

We interpret our psychophysical results within the framework of early spatial interactions [6,7,11,31,32], and suggest that 'inside' and 'outside' regions, or 'figure' and 'ground' are distinguished at the level where the measured variations in contrast sensitivity occur (probably in the primary visual cortex). Observations, analogous to our results, were reported in the primary visual cortex of the macaque monkey, where a strong asymmetry was found in single cells' responses depending on whether the receptive field was positioned

inside or outside of a figure [33–36]. These results were obtained with the reverse mapping technique applied in neurophysiology, where texture, motion, color or disparity defined figures were flashed at different spatial locations relative to the V1 cell's classical receptive field position. Neural correlates of the shape specific sensitivity maps were recently found in the modulation profiles of single-cell activity in the primate striate cortex [35,37–39], also employing the reversed mapping technique. These neuronal correlates suggest an early cortical locus for shape-specific interactions as well.

To conclude, the most provocative possibility is that an explicit, and sparse skeletal representation of shape is generated early in visual processing. The fit between our simulation results and psychophysical data supports the potential relevance of the medial-point description. It should be an interesting question whether spreading neural activity within a layer of V1, or intricate feedforward-feedback connections are better candidates for implementing such a description.

Acknowledgements

We thank Thomas V. Papathomas for his useful comments on the manuscript. Preparation of this paper was supported by a grant from the J.S. McDonnell Foundation 9560.

Appendix A. Mathematical formulation of the D_ϵ-function

Let the bounding contour B of an object be characterized in a parametric form by a curve $b(t):[0,1]\rightarrow\mathscr{R}^2$ where $b(t)=[x(t),\ y(t)]$ is a vector in 2D space. For every internal point $p=[x_p,\ y_p]$, let $M(p)$ be the smallest distance between p and the bounding contour. Obviously, (Eq. (A1))

$$M(p)=\min_{0\le t\le 1}|p-b(t)| \tag{A1}$$

By definition, $M(p)$ is the radius of the largest inscribed circle $C_{\max}$ centered at p (see solid circle in Fig. 2).

The value of D_ϵ is defined for each internal point as the percentage of boundary-contour points whose distance from p is within tolerance ϵ of $M(p)$. We call these points 'ϵ-equidistant' from p, and show that portion of B by thicker lines in Fig. 2. Thus D_ϵ is the arclength of the thick portion of B and is given by

$$D_\epsilon(p)=\int_{|p-b(t)|\le M(p)+\epsilon} dt \tag{A2}$$

Obviously, $0\le D_\epsilon(p)\le 1$

When the object boundary cannot be described as a parametric curve, only as a subset in 2D space, the generalized form of the above definition is the following:

Let $B'\subset\mathscr{R}^2$ be the set of points of the bounding contour (it can be a finite set of points). For any internal point p, let $M(p)=min_{b\in B'}|p-b|$, as before and let

$$T=\int_{b\in B'} db$$

be the total measure of the set. The D-function in this case is given by (Eq. (A3))

$$D_\epsilon(p)=\frac{1}{T}\int_{|p-b|\le M(p)+\epsilon} db \tag{A3}$$

Again, as in Eq. (A2) the value of D_ϵ is between 0 and 1.

References

[1] Marey E-J. Emploi des photographies partielles pour etudier la locomotion de l'homme et des animaux (1883). As cited in: Braun M, editor. Picturing Time. Chicago IL: Chicago University Press, 1992:83.
[2] Johansson G. Visual perception of biological motion and a model for its analysis. Perception Psychophys 1973;14:201–11.
[3] Blum HJ. A new model of global brain function. Perspect Biol Med 1967;10:381–407.
[4] Blum HJ. A transformation for extracting new descriptors of shape. In: Wathen-Dunn W, editor. Symposium on Models for the Perception of Speech and Visual Form. Cambridge, MA: MIT Press, 1967b.
[5] Blum HJ. Biological shape and visual science (part I). J Theor Biol 1973;38:205–87.
[6] Polat U, Sagi D. Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. Vis Res 1993;33:993–9.
[7] Polat U, Sagi D. The architecture of perceptual spatial interactions. Vis Res 1994;34:73–8.
[8] Kovács I, Julesz B. A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. Proc Natl Acad Sci USA 1993;90:7495–7.
[9] Kovács I, Julesz B. Perceptual sensitivity maps within globally defined visual shapes. Nature (London) 1994;370:644–6.
[10] Kovács I, Julesz B. Long-range spatial interactions of early vision. In: Julesz B, Kovács I, editors. Maturational Windows and Adult Cortical Plasticity, SFI Studies in the Sciences of Complexity, vol. XXIII. Reading, MA: Addison-Wesley, 1995:127–136.
[11] Kovács I. Gestalten of today: early processing of visual contours and surfaces. Behav Brain Res 1996;82:1–11.
[12] Ogniewicz RL. Discrete Voronoi Skeletons. Konstanz: Hartung-Gorre Verlag, 1993.
[13] Marr D. Vision. New York: W.H. Freeman, 1982.
[14] Hubel DH, Wiesel TN. Receptive fields of single neurons in the cat's striate cortex. J Physiol (Lond) 1959;148:574–91.

[15] Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol (Lond) 1962;160:106–54.
[16] Daugman JG. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J Opt Soc Am 1984;2(7):1160–9.
[17] Polat U, Sagi D. Spatial interactions in human vision: From near to far via experience dependent cascades of connections. Proc Natl Acad Sci USA 1994;91:1206–9.
[18] Polat U, Sagi D. Plasticity of spatial interactions in early vision. In: Julesz B, Kovács I, editors. Maturational Windows and Adult Cortical Plasticity, SFI Studies in the Sciences of Complexity, vol. XXIII. Reading, MA: Addison-Wesley, 1995:111–126.
[19] Sagi D, Tanne D. Perceptual learning: learning to see. Curr Opin Neurobiol 1994;4:195–9.
[20] Tolhurst D, Barfield L. Interactions between spatial frequency channels. Vis Res 1978;18:951–8.
[21] Kovács I. Points of synchrony in retinotopic maps. Hebb symposium on neurons and biological dynamics. Fields Inst Res Math Sci 1994;May:14.
[22] Attneave F. Some informational aspects of visual perception. Psychol Rev 1954;61:183–93.
[23] Hoffman DD, Richards WA. Parts of recognition. Cognition 1984;18:65–96.
[24] Leyton M. Symmetry-curvature duality. Comput Vis Graph Image Processing 1987;38:327–41.
[25] Kelly and Levine. Annual symmetry operators: a method for locating and descibing objects. IEEE International Conference on Computer Vision. IEEE Piscataway, NJ, USA, 1995: 1016–21.
[26] Burbeck CA, Pizer SM. Object representation by cores: Identifying and representing primitive spatial regions. Vis Res 1996;35(13):1917–30.
[27] Field DJ, Hayes A, Hess RF. Contour integration by the human visual system: Evidence for a local "association field". Vis Res 1993;33(2):173–93.
[28] McIlhagga WH, Mullen KT. Contour integration with colour and luminance contrast. Vis Res 1996;36:1265–79.
[29] Pettet MW, McKee SP, Grzywacz NM. Constraints on long range interactions mediating contour detection. Vis Res 1997;in press.
[30] Rockland KS, Lund JS. Widespread periodic intrinsic connections in the tree shrew visual cortex. Science 1982;214:1532–4.
[31] Polat U, Norcia AM. Neurophysiological evidence for contrast dependent long-range facilitation and suppression in the human visual cortex. Vis Res 1996;36:2099–109.
[32] Mizobe K, Kasamatsu T, Polat U, Norcia AM. Contrast and context modulation of single-cell responses elicited in cat area 17. Soc Neurosci Abstr 1996;254.18.
[33] Lamme VAF. Neuronal correlates of figure-ground segregation in primary visual cortex. Invest Ophthalmol Vis Sci 1994;35:1489.
[34] Lamme VAF. The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci 1995;15:1605–15.
[35] Lee TS, Mumford D, Schiller PH. Neuronal correlates of boundary and medial axis representations in primate visual cortex. Invest Ophthalmol Vis Sci 1995;36:477.
[36] Zipser K, Lee TS, Lamme VAF, Schiller PH. Invariance of figure-ground segregation mechanisms in V1 for depth, orientation, luminance, and chrominance cues. Invest Ophthalmol Vis Sci 1994;35:1973.
[37] Lee TS. Evidence for scale-invariant medial axis representation in primate striate cortex. Soc Neurosci Abstr 1995;21:1648.
[38] Lee TS. Neurophysiological evidence for image segmentation and medial axis computation in primate V1. In: Bower JM, Norwell MA, editors. Computation and Neural System, Proceedings of the Fourth Annual Computational Neuroscience Conference. MA: Kluwer Academic Publishing, The Netherlands, 1996:in press.
[39] Lee TS, Mumford D, Romero R, Lamme V. The role of primary visual cortex in object representation and recognition. Vis Res 1996;(present issue).
[40] Braun M. Picturing Time. Chicago, IL: Chicago University Press, 1992.