Image Object Description Without Explicit Edge-Finding

TR91-049

November, 1991

Stephen M. Pizer
James M. Coggins
Daniel S. Fritsch
Bryan S. Morse

Medical Image Display Group
Department of Computer Science
Department of Radiation Oncology
The University of North Carolina
Chapel Hill, NC 27599-3175

The research was partially supported by NIH grant number #P01 CA47982 and NASA contract #NAS5-30428.
Submitted to the Second European Conference on Computer Vision (ECCV), 1992.

Image Object Description Without Explicit Edge-Finding

STEPHEN M. PIZER, JAMES M. COGGINS, DANIEL S. FRITSCH, BRYAN S. MORSE

Medical Image Display Research Group
University of North Carolina
Chapel Hill, NC USA

Correspondence address before 4 January 1992:
Sitterson Hall
University of North Carolina
Chapel Hill, NC 27599-3175 USA
Telephone: 1-919-962-1785
Fax: 1-919-962-1799
email: pizer@cs.unc.edu

Correspondence address as of 4 January 1992:
Buys Ballot Laboratorium
Rijksuniversiteit Utrecht
Princetonplein 5
Postbus 80.300
3508 TA Utrecht
Netherlands
Telephone: 31-30-53 39 85
Fax: 31-30-52 26 65
email: steve@cv.ruu.nl

Abstract

The various tasks of computer vision dealing with objects, such as recognition, registration, and measurement, have typically required the intermediate step of finding an object edge, or equivalently the list of pixels in the object. This paper proposes a means for characterizing object structure and shape that avoids the need to find an explicit edge but rather operates directly from the image intensity distribution in the object and its background, using operators that do indeed respond to "edgeness". The means involves a generalization of medial axis descriptions from objects defined by characteristic functions to those described by intensity distributions. The generalized axis is called the *multiscale medial axis* because it is defined as a branching curve in scale space. The result is stable to calculate and can be used to subdivide an image object into subobjects and detail subshapes as well as to characterize the shape properties of the objects, subobjects, and detail subshapes.

Edges

The dominant train of thinking in object recognition, registration, and measurement has been that grouping is based on the local detection and tracking of edges. These edges or the regions enclosed by them are first found. Then various measurements are made on the result, such as edge curvatures, medial axes, or moments of the object, and the final recognition, registration, and measurement are based on these. The difficulty of this approach is two-fold. First, from the point of view of physics, for an object in an image there exists no edge locus without a tolerance since the object can exist only via imaging and visual (here computer visual) measurements which have an associated spatial scale, and thus spatial tolerance [Koenderink, 1990b]. Second, the design of methods of detection of object boundary regions, even with tolerance, has been tried by innumerable scientists with limited general success, probably due to the fact that it is hard to build global properties into the edge-finding process.

In addition to the problems of finding edge loci, it is hard to see how to use the local measurements that determine edges to get at the global properties that have to do with finding an object, such as the relation that between opposite points on two sides of an object, called *involutes* (see figure 1 for examples).

Medial Properties

The above difficulties are ameliorated with an encoding scheme responding to the opposite object edges simultaneously, sensing the object region

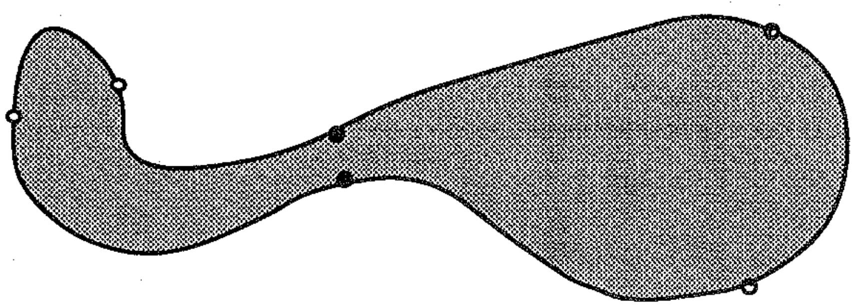

Figure 1: Involutes: visually related opposite points on an object

rather than its separate edges. Beginning with Blum [1967], many in the field of computer vision have been attracted by a scheme of this type in which an object is represented in terms of a medial axis or skeleton running down the center of the object, together with a width value at each point on the medial axis. Leyton [1984, 1987] has suggested that the long known fact that corners and other object boundary locations of locally maximal curvature are perceptually important is related to the correspondence of these locations to endpoints of these central axes. It has also been noted [von der Heydt, 1984] that subjective edge perceptions derive especially strongly from extensions of edges from corners, and this idea has been generalized for computer vision by Heitger & Rosenthaler [1991].

The Blum medial axis is formally defined as the locus of centers of maximal disks in the object (see figure 2). As a result every axis point corresponds to two (or occasionally more) object boundary points where the maximal disk tangentially touches the boundary. These two boundary points appear to correspond to each other in a way consistent with the visual percept. The medial axis carries with it (in the radii of the disks) straightforward access to the angle of the object boundary at each of these two boundary points relative to the axis direction at the corresponding axis point. Moreover, the curvature of the axis and of the boundary pair relative to the axis is also straightforwardly accessible. The maximal disks at axis endpoints select the visually important vertices of protrusions via the locations at which their disks touch the boundary, and the axis branch points correspond to indentations, i.e., branching, in the object itself.

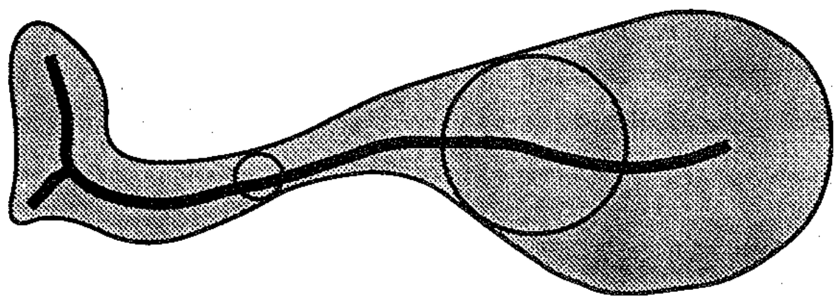

Figure 2: The medial axis for an object

Blum also suggested a more general form of the medial axis: the locus of the centers of all disks that are tangent to the object boundary over two or more connected boundary segments. This *global* form of the axis includes sections in which the tangent disks are external to the object; these sections select indentations into the object or equivalently protrusions in the object's background. Global axis sections for which the disks overlap the object and its background select symmetries of larger width than the object, for example, the longer symmetry of a rectangle.

Multiscale Geometry Detectors

Many investigators have suggested that grouping into objects must be based on measurements in scale space, i.e., by sets of detectors that sense a regional rather than curvilinear (e.g., edge) property, with each detector sensing the same property but at different spatial scales. Among the detector kernels suggested have been derivatives of Gaussians [Koenderink, 1990a], differences of Gaussians [Crowley, 1984], Gabor functions [Daugman, 1980; Watson, 1987], Wigner operators [Wechsler, 1990], and wavelets [Mallat, 1989, 1991]. The most persuasive case for how to choose the form of receptive fields, by ter Haar Romeny et al [1991], is that any visual system, including computer vision systems, that must be invariant to translation, rotation, and size change must have multiscale receptive fields which are solutions to a diffusion equation, e.g., linear combinations of derivatives of a Gaussian. These receptive fields or combinations of them can be thought of as measuring geometrical properties such as "edgeness", "cornerness", and "t-junctionness", in many cases with an orientation. The Laplacian of the Gaussian has certainly been a popular choice [Marr, 1982].

The Multiscale Medial Model

Collectively the above ideas have led us to the development of a new model for visual grouping and description of object shape. This model appears reasonable not only for computer vision but also at the neural level as a model of human visual processing. It produces a group of global-form medial axes by multiscale, regional, two-edge-engaging geometric measurements. It is based on a set of measurements $R(\mathbf{x},s)$ in scale space (location ($\mathbf{x}$) × scale (s)) that have a particularly strong response relative to nearby positions and scales when the measurement has a strong contribution by two opposing boundary regions at a distance s from $\mathbf{x}$. That is, R measures "medialness" in the sense that points which are medial between two boundaries and have a scale corresponding to the distance between the medial point and the boundary give strong responses. An example of $R(\mathbf{x},s)$ from the literature is the difference of Gaussians normalized by its absolute value integral [Crowley, 1984], where s is the standard deviation of the larger Gaussian. Other response functions R that we find promising will be given in the next section.

The notion of having a strong response relative to nearby positions and scales is formalized to mean a sort of ridge in scale space (see figure 3), as follows. A medial point $\mathbf{x}$ should have two properties. First, a slightly larger or slightly smaller scale should give a smaller response there, and second, this response should form a ridge in image space. To be more precise,

1) $R(\mathbf{x},s)$ must be a relative maximum with respect to s for that fixed $\mathbf{x}$. Let M be the set of ($\mathbf{x}$, s) such that $R(\mathbf{x},s)|_{\mathbf{x}}$ is a such a relative maximum with respect to s. Partition M into its connnected subsets, M_i, i = 1,2, ... In each M_i there exists a connected region of image points $\mathbf{x}$ not necessarily covering the whole image space, and there exists at most one scale s associated with any such position $\mathbf{x}$.

2) For each M_i, project $R(\mathbf{x},s)$ for $(\mathbf{x},s) \in M_i$ onto $\mathbf{x}$ to form the image or subimage $R^{max}{}_i(\mathbf{x})$: $R^{max}{}_i(\mathbf{x}) = R(\mathbf{x},s)$ for $(\mathbf{x},s) \in M_i$. Then $(\mathbf{x},s)$ is in the multiscale medial axis if $\mathbf{x}$ is a ridge point in any such portion of $R^{max}{}_i(\mathbf{x})$ for any i.

Among the many non-equivalent ridge definitions in literature, we use the definition that a ridge point of a function $f(\mathbf{x})$ is a place where a level curve of f has maximal curvature, i.e, a place where the orientation of the gradient of f changes maximally along the direction perpendicular to the gradient.

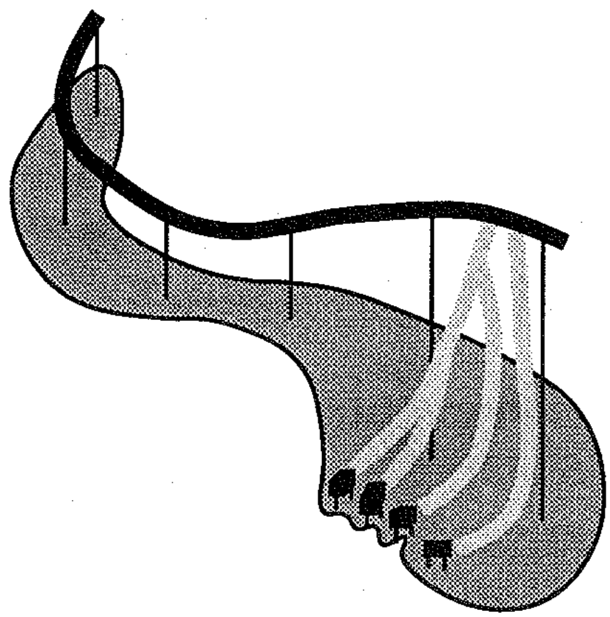

Figure 3: Scale space medial axis traces for an object. Dotted traces are less strong than solid traces.

Such a scale-space ridge is a possibly branching trace in scale space (*x,y, scale*). The *x,y* positions of these ridges form a medial axis for an object, and their scales specify its width at each axis point. Just as with the Blum medial axis, width (scale) angles and curvatures (boundary orientation and curvature relative to the axis) are straightforwardly available. Also, excitatory connections along the ridge and inhibitory connections across the ridge should produce subjective edges in the appropriate way. Note that this operation applies to grey scale objects with fuzzy edges as well as those with sharp edges.

As seen in figure 3, the unbranching component of the axis at the largest scale describes the gross orientation and width properties of the object. It establishes the boundary of the object only to a tolerance proportional to the width of the object. The branches into smaller scales correspond to smaller boundary detail (see figure 3) or objects within the main object (see figure 4), either with boundary tolerance proportional to their widths. Yet tighter tolerance on the boundaries can be obtained from smaller scale operators responding to single boundaries within the boundary regions associated with the medial ridge.

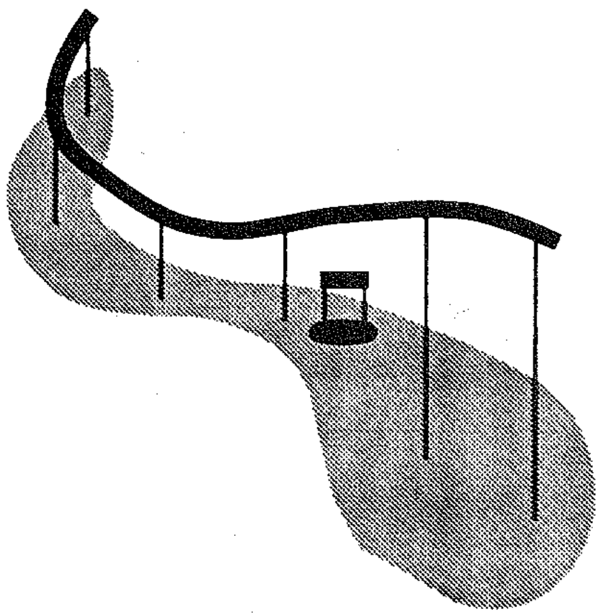

Figure 4: Scale space medial axis traces for an object within an object. Dotted traces are less strong than solid traces.

Response Functions

Medial response functions can be thought of as produced via an axis-centered operator or an edge-centered operator. An axis-centered operator is centered at a point that responds to edgeness at some range of distances from it. An edge-centered operator measures edgeness of some orientation at a given scale and contributes to a medial response at that scale but at a distance from the measurement point proportional to the scale of measurement and in a direction perpendicular to the orientation at which the edgeness was measured (see figure 5). An example of the first approach is that based on the normalized Laplacian of a Gaussian (Crowley [1984] uses a similar normalization on a difference of Gaussians). The second approach has the flavor of a Hough transform -- each point is voting for medial points in scale space.

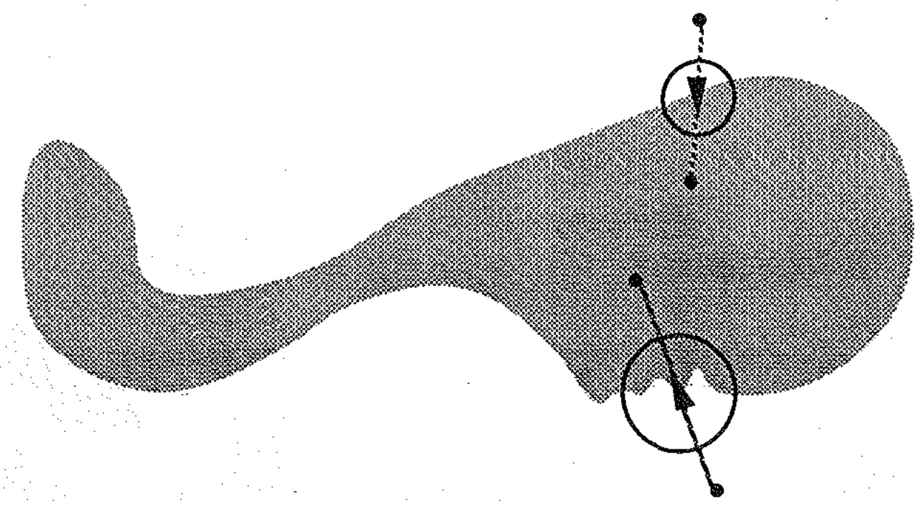

Figure 5: Two examples of the effect on the response function of an edge-centered response at an orientation. The circle indicates the scale of a directional derivative and the arrow its orientation. The heavy dots indicate the center of the region where the result is applied as votes. Note that derivatives at any point are taken in all orientations, including the non-orthogonal ones shown above.

Among the axis-centered operators that we are investigating are the Laplacian of the Gaussian at the scale in question (i.e., the trace of the Hessian at the selected scale), the maximum over directions of applying the second directional derivative of a Gaussian (i.e., the maximum eigenvalue of the Hessian at the selected scale), the sum of the the squares of the eigenvalues of the Hessian at the selected scale (sometimes called the deviation from flatness), and the magnitude of the determinant of the Hessian at the selected scale, and, per Crowley, scale-normalized versions of these such that the maximal response to a step edge for each position is independent of the distance from that edge [Fritsch, 1991]. A multiscale medial axis from a scale-normalized Laplacian at each scale is given in figure 6. Other rotation- and translation- and scale change-invariant operators can be derived as linear combinations of the receptive field sets of Koenderink [1990]. The difficulties with such operators is that they cannot analyze an object that has contrasts of different polarity at different positions along the boundary.

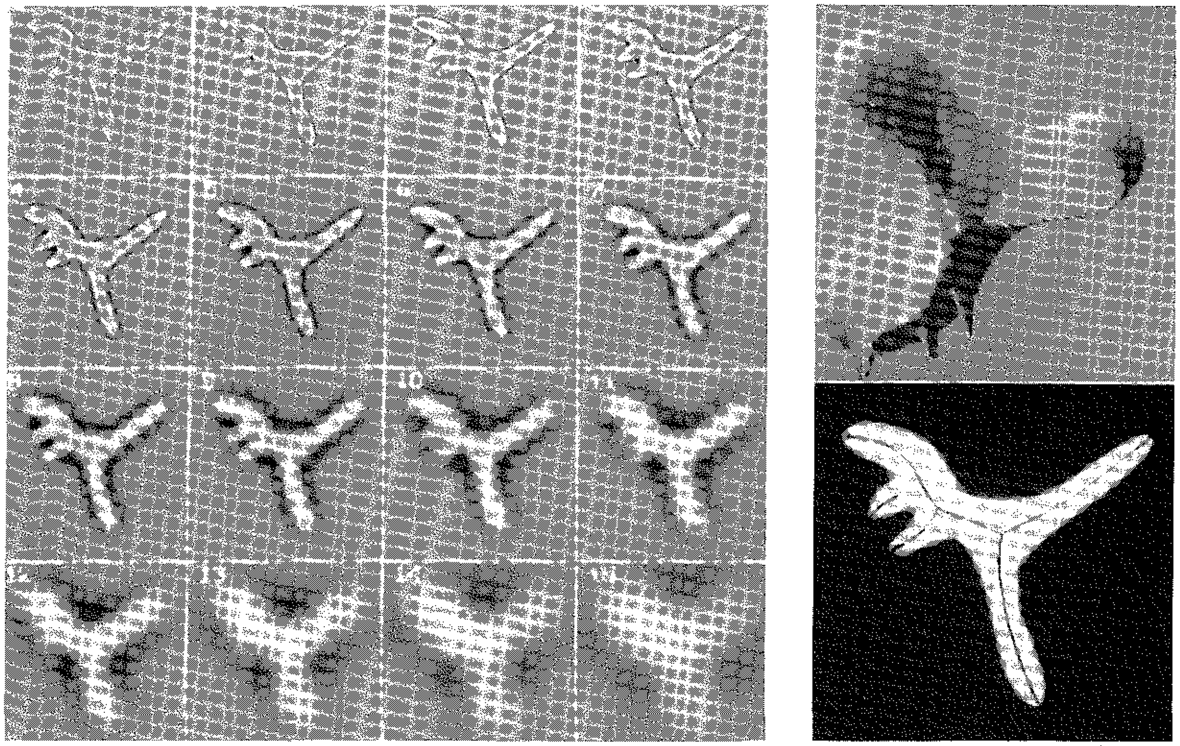

Figure 6: a) Scale-normalized Laplacian response function values, b) scale space ridge $R^{max}(\mathbf{x})$, and c) multiscale medial axis superimposed on the original image.

The Hough-like approach has been tried with the magnitude of the result of applying the first directional derivative of a Gaussian as the vote strength [Morse, 1991]. The vote, with weight given by this scaled derivative magnitude, is produced for each combination of derivative orientation and point location. This vote is fuzzily applied at a distance from that point proportional to the standard deviation of the Gaussian in both directions along the orientation of the derivative. A scale-space ridge from this approach is given in figure 7.

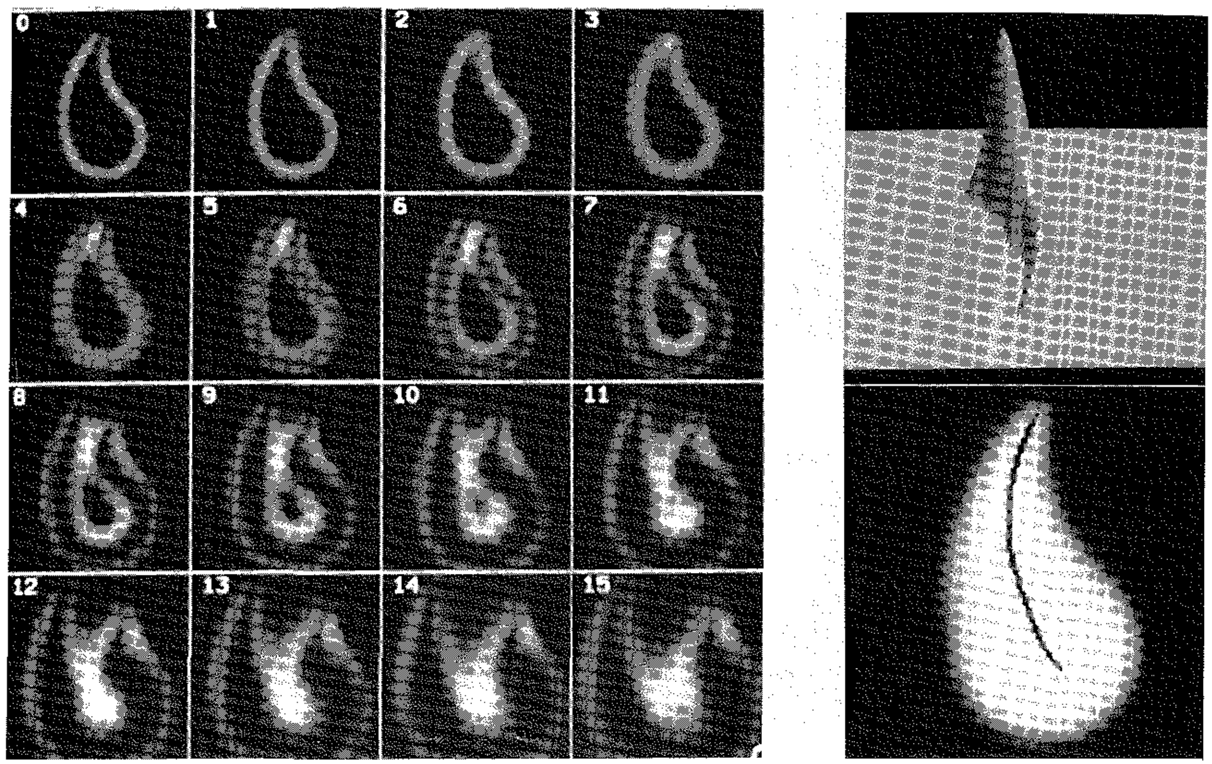

Figure 7: a) Response function values derived from magnitude of edge-centered derivative of Gaussian, b) scale space ridge $R^{max}(\mathbf{x})$, and c) multiscale medial axis superimposed on the original image.

Implementation and results

The calculation of the response function itself is simply the application of various filters, possibly followed by the calculation of a sum or product (e.g., to obtain a trace or determinant) of these results at each position. All of the response relative maxima across scale are calculated independently at each pixel by scanning the response values across scale at that pixel. We calculate the ridges in all loci of these relative maxima that are continuous in scale space. These ridges are calculated using the geometry-limited diffusion approach of Whitaker [1991], with the conductance equal to the exponential of the negative square of the gradient of intensity gradient orientation. The means of specification of continuity in scale space is still being researched.

We are applying these analysis techniques to medical images. An example is given in figure 8.

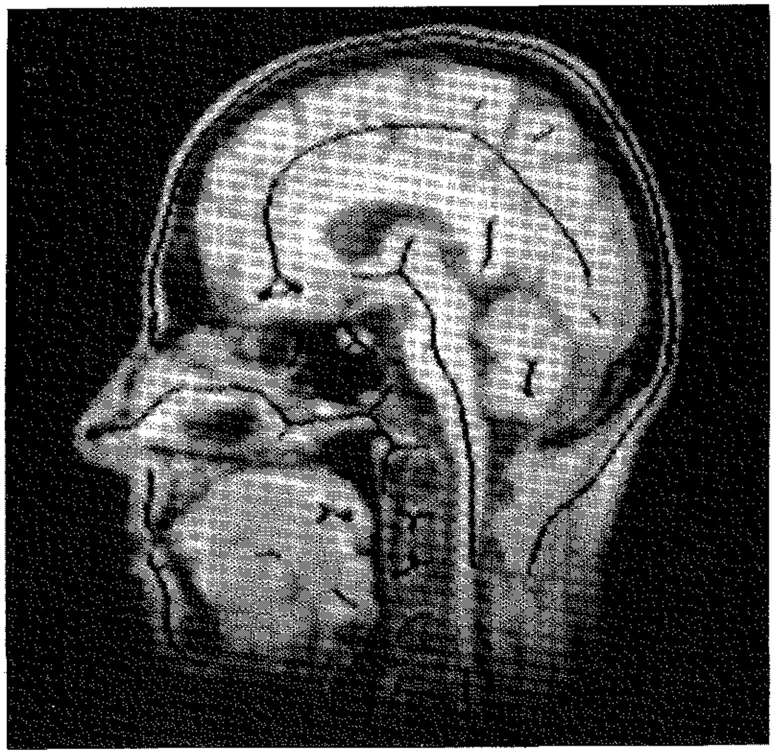

Figure 8: Multiscale medial axis superimposed on an MRI image of the head. The response function used is a scale-normalized Laplacian.

Discussion

The multiscale medial axis has many fine properties. Like Blum's global medial axis, it

1) separates object curvature from width properties, thus preserving shape measures across small changes in local orientation produced by warping or bending,

2) allows the identification of the visually important ends of protrusions and indentations, i.e., points of extremal boundary curvature,

3) allows the identification of involutes, i.e., visually opposite points on the boundary, for a range of scales of symmetry, and

4) naturally incorporates size constancy and orientation independence.

However, unlike Blum's medial axis it provides this information at a scale appropriate to the object width, does not require the unstable preliminary calculation of a boundary, and is a more stable property of the object. This stability derives from the two facts that it is tied to the center of the object and so cannot get lost like an edge boundary can, and that it incorporates the "noise averaging" inherent in Gaussian convolution, i.e. in considering objects in scale space.

The definitions and implementations of the multiscale medial axis all extend to three dimensions.

Moreover, the multiscale medial axis has many potential uses in computer vision:

1) The object-subobject relationships it defines can be computed for any image to produce a quasi-hierarchy that can be used in interactive computer systems for the fast definition of objects in images [Pizer, 1989]. These defined objects can in turn serve 3D display and object measurement.

2) The groupings defined by this approach can be used to define object inclusion likelihoods that in turn can be used to produce automatic measurements of object volumes, e.g. tumor volumes, or other object properties such as integrated metabolic function.

3) The position and scale co-ordinates along the medial ridges and the outputs of various receptive fields there can be used as a basis for matching of structures in tasks involving registration between objects in rather different images of the same anatomy, such as a simulation and a portal image in radiation oncology.

Work in all of these directions is proceeding in our laboratory.

Acknowledgements

The research reported here was partially supported by NIH grant #P01 CA47982 and NASA contract #NAS5-30428. We are indebted to Christina Burbeck for collaboration on human vision aspects of the multiscale medial axis. We are grateful to Guido Gerig for helpful suggestions. We thank Carolyn Din and Bo Strain for help in manuscript preparation and Graham Gash for systems support.

References

Blum, H, A new model of global brain function, *Perspectives in Biol. & Med.* **10**:381-407, 1967.

Crowley, JL, AC Parker, A representation for shape based on peaks and ridges in the difference of low-pass transform, *IEEE Trans PAMI* **9**(2):156-170, 1984.

Daugman, JG, Two-dimensional spectral analysis of cortical receptive field profiles, *Vis Res* **20**:847-856, 1980

Fritsch, DS, JM Coggins, SM Pizer, A multiscale medial description of greyscale image structure, Proc. Conference on Intelligent Robots and Computer Vision X: Algorithms and Techniques, SPIE, 1991

ter Haar Romeny, BM, LMJ Florack, JJ Koenderink, MA Viergever, Scale-space: its natural operators and differential invariants, *Information Processing in Medical Imaging*, ACF Colchester & DJ Hawkes, eds, Lecture Notes in Computer Science **511**:239-255, Springer-Verlag, 1991.

Heitger,F, L Rosenthaler, R von der Heydt, E Peterhans, O Kübler, Simulation of Neural Contour Mechanisms: From Simple to End-Stopped Cells, Tech. Rep. BIWI-TR-126, ETH, Zurich, Commun. Tech. Lab, Image Sci. Div., 1991.

von der Heydt, R, E Peterhans, G Baumgartner, Illusory contours and cortical neuron responses, *Science* **224**:1260-1262, 1984

Hubel AH, TN Wiesel, Receptive fields and functional architecture of monkey striate cortex, *J Physiol* **195**:215-243, 1968

Koenderink, JJ, AJ van Doorn, Receptive Field Families, *Biol Cyb* **63**(4):291-297, 1990a.

Koenderink, JJ, *Solid Shape*, MIT Press, 1990b.

Leyton, M, Perceptual organization as nested control, *Biol Cyb* **51**:141-153, 1984

Leyton, M, Symmetry-curvature duality, *Comp Vis, Graphics, Im Proc* **38**:327-341, 1987

Mallat, S, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans PAMI* **11**(7): 674-693, 1989

Mallat, S, Zero-crossings of a wavelet transform, *IEEE Trans Info Theory* **37**(4):1019-1033, 1991

Marr, D, *Vision*, Freeman, San Francisco, 1982.

Morse, BS, A Hough-like medial axis response function, TR91-044, UNC Dept. of Comp. Sci., 1991