3. CORE MODEL OF HUMAN VISION

The visual model forms the very basis of the image quality approach pursued in this research. For this image quality measure to produce consistent predictions of human image assessment, it must implement mechanisms hypothesized to exist in the human visual system. The requirements of the model might be specifically enumerated in the following way. First, the model ought to provide the functional basis for the performance of the task. In the case of some detection task, for example, it is necessary to use or adapt the output of the model to indicate, perhaps in a graded fashion, whether the structure of interest is present or absent. Likewise, for an estimation task, the model must provide a mechanism or result as the basis for determining a measure, or estimate, such as the area of an anatomical object. Second, and more importantly, there must be some reason to believe that the manner in which the model carries out the task is consistent with, and thus predictive of, the way in which humans perform the task. It is possible to conceive of a method for computing a measurement that is unrelated to the manner in which the human visual system would perform the task. That kind of a method would appear to have little hope of providing the basis for an image quality measure that would be correlated with human performance across variations in image characteristics.

The core model of human shape perception,[1,2,3] developed by the cooperative efforts of several departments at this university, has demonstrated promise for the computation of some of the basic shape tasks involved in medical image perception. It is the visual model that will be adopted for this research, and this chapter describes it in detail. It should become clear that the model provides an attractive mechanism for shape estimation. Furthermore, the chapter attempts to develop an expectation that the model will be predictive of human performance. Of course, the experiments related in the third part of this dissertation ultimately render a verdict in the comparison of the model with human performance for the two tasks studied there.

3.1 Core Theory

Traditional shape theories possess mechanisms for computing object contours followed by, for example, a "filling-in" process.[4] These object boundaries are difficult to extract from greyscale images, and the process of deriving a shape representation from a list of boundary points can be cumbersome. The core model on the other hand capitalizes on the inherent "two-sidedness" of objects and captures, more succinctly and directly than a boundary description, an object's important properties, namely the position of its middle and corresponding widths. The model's multiscale properties allow invariant recognition of an object at any spatial size and orientation and enable a description of overall shape that is relatively independent of more localized boundary fluctuations. Width information, the basis for many estimation tasks in medical image viewing and in particular those in this research, is directly accessible from the core description. The core model is thus an effective computational approach for representing shape. More importantly for utilizing it as a model in determining image quality in the same way that the human might, the core model seems to offer a reasonable description for how the human visual system segregates and represents shapes. The following sections discuss core formation as a visual process. Later sections in this chapter offer some evidence for the model’s validity and describe its computer implementation.

3.1.1 Edge measurements

Higher level form descriptions in the human visual system are undoubtedly generated from fundamental measurements about changes, or boundaries, in perceptual quantities such as luminance, texture, and velocity. The prevalence of cells in the visual cortex responsive to luminance boundaries was first demonstrated by the pioneering work of Hubel and Wiesel.[5] We know that those cells, and others that have since then been shown to

---

[1] S.M. Pizer, C.A. Burbeck, J. Coggins, D. Fritsch, B.S. Morse, “Object Shape Before Boundary Shape: Scale-space Medial Axes,” Journal of Mathematical Imaging and Vision 4 (1993): 303-131.

[2] S.M. Pizer, C.A. Burbeck, D.S. Fritsch, B.S. Morse, A. Liu, S. Murthy, D.T. Puff, “Human Perception and Computer Image Analysis of Objects in Images,” Proceedings of the Conference of the Australia Pattern Recognition Society (DICTA) 1 (1993): 19-26.

[3] C.A. Burbeck, S.M. Pizer, “Object Representation by Cores: Identifying and Representing Primitive Spatial Regions,” Vision Research, 1995, in press.

[4] S. Grossberg, E. Mingolla, “Neural Dynamics of Perceptual Grouping: Textures, Boundaries, and Emergent Segmentations,” Perception and Psychophysics 38 (1985): 141-171.

[5] A.H. Hubel, T.N. Wiesel, “Receptive Fields and Functional Architecture of Monkey Striate Cortex,” Journal of Physiology 195 (1968): 215-243.

respond to other types of boundaries such as texture,[6] are characterized by their preferential firing in response to edges of a particular orientation. Furthermore, that rate of firing is directly proportional to how well the stimulus matches the cell's characteristic preference. Because the cells indicate, in a graded fashion, the extent to which a stimulus is a boundary consistent with their preference, they have been referred to as "boundariness" detectors.

The region in the visual field over which the cell collects information is termed its "receptive field," and the size of that region is defined as the *scale* of its measurement. Any visual measurement is made at a particular scale or aperture. A typical visual scene is composed of much information, and the human observer views only a portion of it at once. The observer's attention shifts among objects of all sizes, and the visual system utilizes in its perception of those objects receptive fields of different scales as needed. Large structural features are primarily the domain of large receptive fields. As detailed portions of the scene are attended to, the receptive fields recruited are appropriately smaller. The set of measurements about boundaries, or other quantities such as medialness (described next) derived from boundariness, form a *scale space*:[7] measurement values are represented as a function of spatial position in the visual scene for the range of measurement scales employed.

3.1.2 Medialness

The issue that remains in developing a model for how humans perceive form is to hypothesize how boundary measurements are combined and related to represent an object. The hypothesis that is the basis of the core model is that edges opposite each other in the object, involutes, are perceptually linked. This section presents the theory as though an explicit linkage mechanism might occur to combine individual boundary responses. At the end of the section, recent evidence is revealed for the existence of cells in the visual cortex sensing *directly* the presence of object middles.

Weber's law[8] for separation states that the accuracy with which the distance between two locations may be determined is inversely proportional to the distance between them: the accuracy for estimating the distance between two bars decreases as the bars are moved farther apart. A perceptual mechanism that would explain this observation has been proposed. Burbeck and Hadden[9] measured separation discrimination for two parallel bars while varying the position of a third, "background" bar placed outside of, and also parallel to, the two stimulus bars. They found that accuracy for bar separation diminished when the background bar was placed a distance from the stimulus bars that was less than the separation between the stimulus bars. The distance at which the background bar was able to disturb separation judgment provides an indication of the area over which information is integrated in judging the separation, so the finding that the distance to the background bar was proportional to the separation distance suggests that the integration area scales with the separation distance. In other words, the size of the boundariness detectors recruited for the judgment is proportional to the distance being judged.

This finding has led to a hypothesis about how boundaries are linked: opposite edges of an object are linked with a resolution proportional to their separation.[10] Small scale measurements would only be utilized by the visual system in the presence of small objects, for example. Thus small scale detectors are linked over small distances, and larger scale detectors are connected over larger distances (Figure 3.1).

The theory for the core model is that boundariness measurements are made in some region of the visual scene at a range of scales. A boundariness detector contributes, with a strength proportional to the extent to which the local information in the scene is consistent with its scale and orientation preferences, to the formation of medialness. Detectors contribute at a distance proportional to their scale, and in a direction prescribed by their orientation (Figure 3.2). Furthermore, the detectors contribute in proportion to the magnitude of the boundary. Such relationships might exist in the visual cortex as excitatory synapses between boundary detectors and hypothetical medialness cells.

According to the model, the boundariness contributions are accumulated at each scale. It will be the case that where two boundaries bilaterally contribute to medialness, the accumulated response will be locally particularly high (Figure 3.3). These accumulated contributions generate a scale-space of "medialness," a mapping of the extent to which each position is a middle of an object of the size proportional to that position's scale value (Figure 3.4). Thus, positions in the medialness scale space with high values are more likely to be object middles.

---

[6] H.C. Northdurft, J.L. Gallant, D.C. VanEssen, "Neural Responses to Texture Borders in Macaque Area V1," Society of Neurosciences Abstract 18 (1992): 1274.,

[7] T. Lindeberg, Scale-Space Theory in Computer Vision, Kluwer International Series in Engineering and Computer Science (Dordrecht, the Netherlands: Kluwer Academic Publishers, 1994).

[8] H.K. Wolfe, "On the Estimation of the Middle of Lines," American Journal of Psychology 34 (1923): 313-358.

[9] C.A. Burbeck, S. Hadden, "Scaled Position Integration Areas: Accounting for Weber's Law for Separation," Journal of the Optical Society of America A 10 (1993): 5-15.

[10] C.A. Burbeck, S.M. Pizer, B.S. Morse, D. Ariely, G.S. Zauberman, J. Rolland, "Linking Object Boundaries at Scale: A Common Mechanism for Size and Shape Judgments," accepted to Vision Research, 1995

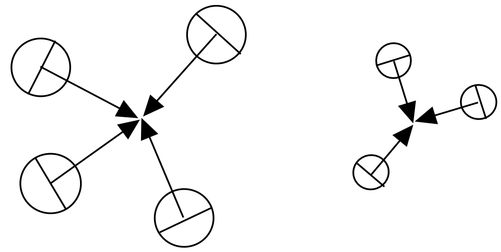


Figure 3.1. Boundariness detector linkage. Boundariness detectors are linked at a distance proportional to their scale. The scale of the detector is indicated here by the radius of the circle, and linkage occurs perpendicular to the preferred orientation of the detector at a distance proportional to that scale.

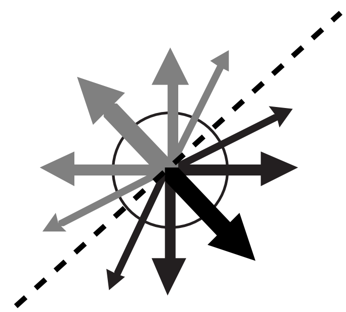


Figure 3.2. Boundariness detector responses. At a boundary (dashed line), detectors contribute to medialness formation at a distance proportional to their scale (indicated by the radius of the circle) and in a graded manner (indicated by the width of the arrows) related to their orientation preference. The detector may contribute in both perpendicular directions, or alternatively may be selective for object polarity and only vote for medialness of, for instance, light objects on a dark background.

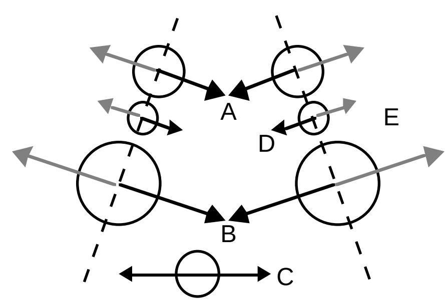


Figure 3.3. Medialness accumulation. At positions where boundariness detectors have voted bilaterally (A , B), the summation of the contributions will be large. Conversely, at other positions medialness may be correspondingly smaller because there was little evidence for an edge (C), or because the scale is such that the contributions fail to combine (D) or the detectors vote perpendicular to the edge and "outside" the object (E).

Recent evidence suggests that there may exist cortical cells that respond preferentially to object middles. Lee, *et.al.*,[11] measured spatio-temporal response profiles for cells in the primate striate cortex. Stimuli were strips and squares distinguished from their backgrounds by texture difference alone. Of primary interest for this discussion was the finding of strong response sensitivities when the middles of these stimuli were positioned at the cells' receptive field centers. These results are preliminary evidence for the possibility that the human visual system might also sense directly, without additional neural linkage, the existence of object middles. In any case the multiscale principles and resulting medialness distribution suggested by the core model remain as described

[11]T.S. Lee, D. Mumford, P.H. Schiller, "Neural Correlates of Boundary and Medial Axis Representations in Primate Striate Cortex," <u>Investigative Ophthalmology and Visual Science</u>, Annual Meeting of the Association for Research in Vision and Ophthalmology 36, no. 4 (March 15, 1995): abstract 2205-172.

throughout this chapter. In fact, it is a direct mechanism that is adopted for the core model implementation (Section 3.3) in this research.

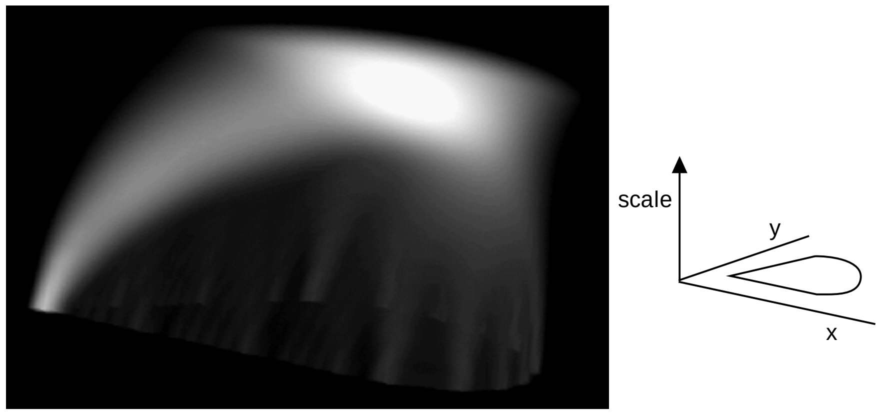


Figure 3.4. Medialness scale space for a teardrop object. This volume rendering depicts medialness for a "teardrop" as a function of two dimensions of position together with scale (see axes at right). Medialness is high at positions corresponding to the middle of the object and at scales proportional to the width of the object.

3.1.3 The Core

The core is a one-dimensional track through the regions of highest medialness. Mathematically, the core may be defined as a *ridge*, an extension of the definition of a local maximum to multiple dimensions.[12] In two dimensions, the ridge is analogous to the crest of a mountain range; in three dimensions (such as the two spatial dimensions and one scale dimension of scale space), a ridge may be thought of as the heart, or "core," of a cloud (a more formal definition and implementation of ridges is provided in Section 3.3). The ridge of medialness that is the core is thus a continuous curve through scale space. Its position reflects the middle of the object while its height is proportional to the width of the object (Figure 3.5).

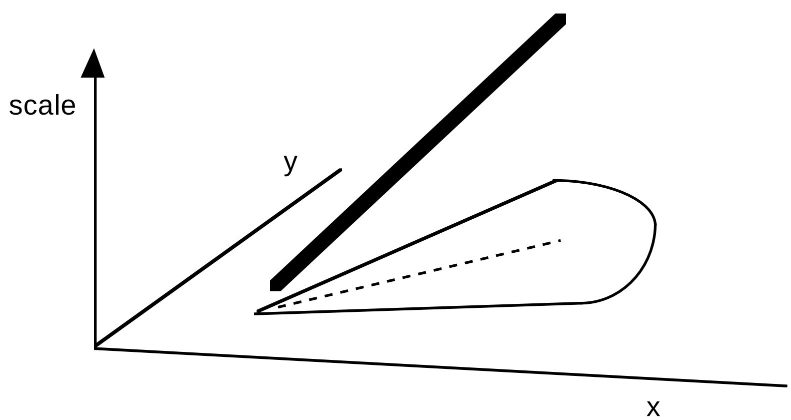


Figure 3.5. Core for teardrop object. In theory, the core for a "teardrop" increases in scale in proportion to its increasing width. The projection (dashed line) of the core lies at the middle of the object.

A figure may be defined as a shape represented by a single core. Thus an object may actually be composed of multiple figures and therefore be described by multiple cores. The cores describing protrusions, indentations, and subobjects may all be related in a hierarchy to describe object shape. It is the case that fluctuations at an object boundary are represented by small-scale cores of their own, while the core describing the global shape of the object may not reflect the presence of the smaller boundary fluctuations. The core is therefore

[12] D. Eberly, R.B. Gardner, B.S. Morse, S.M. Pizer, C. Scharlach, "Ridges for Image Analysis," Journal of Mathematical Imaging and Vision 4 (1994): 351-371.

a particularly concise representation of figure or object shape. There is also evidence to suggest that it is a reasonable model for the way that humans segregate and encode shape.

3.2 Psychophysical Verification

Experiments have been conducted[13] to test the essential hypothesis of the model, that boundariness detectors are linked at a distance proportional to their scale. Several related studies were conducted that employed objects with sinusoidally-modulated edges to measure the perception of figure middles. The figures attempt to expose the perceptual relationships between the scale of the judgments made at the boundary of the a figure and the perception of its middle. The stimuli (Figure 3.6) possessed two types of edge modulations. “In-phase” modulations created objects of constant width and modulating middles (Figures 3.6a, 3.6c), while “out-of-phase” modulations created objects with a straight middle and modulating width (Figures 3.6b, 3.6d).

A B C D

Figure 3.6. Edge-modulated stimuli. These figures were used in an experiment testing the core model hypothesis of boundary linkage at a distance proportional to the scale of the boundary measurements. The wider stimuli in Figures C and D have proportionately larger edge modulation amplitudes. Figures A and C have constant width and modulating middle position. Alternatively, Figures B and D have straight middles and modulating width.

The hypothesis of the core model would predict that figures with edge variations that are small relative to the width of the figure will have little influence on the perception of its middle. The model posits that the figure middle is determined by a linkage of boundariness measurements at a scale proportional to the width of the figure. Larger scale measurements would be relatively insensitive to the smaller edge variations. For example, Figure 3.6c might, under this hypothesis, appear straight even though in-phase edges appear to cause the middle of the corresponding smaller width figure (Figure 3.6a) to “wiggle.” On the other hand, if the scale of boundariness detectors is selected regardless of figure width (that is, the basic assumptions of the core model are erroneous), the influence of edge modulations on the perception of the figure middle should be essentially constant as figure width is varied. In the experiment, the observer’s perception of the middle of the figure was inferred by asking him/her to say whether a probe dot, that was over many trials positioned at different horizontal and vertical locations within the figure, was to the left or right of the figure middle. Other important details of the experiment are presented in the reference.

The results were in fact largely consistent with the core model hypothesis. The perceived modulation of the figure middle was approximately inversely proportional to figure width, suggesting that figure boundaries are indeed linked at a scale related to figure width. Core computations performed with a computer implementation of the model support the perceptual judgments; calculated cores exhibited less central modulation for larger width figures.

An interesting finding of the experiments was that the proportionality factor for the relationship between boundary integration area and figure size was not constant. While the perceived figure middle modulated with an inverse relationship to figure width as hypothesized by the core model, the effect was not related by a single factor that would imply zoom invariance. Specifically, the experimenters measured perceived central modulation for two stimuli, one a zoomed version of the other. The edge amplitude and modulation were scaled by a factor of two for the wider figure. There were significant differences in the correspondence of the data for the two figures; simply shifting the data by a factor of two failed to superimpose the results.

Burbeck and Pizer then suggested that there might exist a common mechanism for size discrimination and shape perception. That is, the perceptual mechanism that produces the property for size that an increase in bisection thresholds with increased distance might at the same time underlie the shape results (a perceived middle modulation increase with increased figure width). If so, some scaling factor other than perfect zoom invariance

[13] C.A. Burbeck, et.al., “Linking Object Boundaries at Scale.”

could be discovered to account for both results. The authors tested whether the scaling factor for bisection thresholds, calculated as the ratio of the wider to narrower thresholds, could predict the perceived middle modulation results. Remarkably, this ratio was found to be exactly that needed to predict the effect of figure width on shape perception.

Further psychophysical evaluation of the core model is of course necessary and in fact underway. The preparation for the experiments in this research preceded the demonstration of imperfect zoom invariance in humans. Furthermore, the nature of this non-proportionality is still not completely understood. Therefore, the model tested here embodies the perfect zoom invariance premise. For the most part, the initial experimental results and inferences indicate the potential viability of the core model in predicting human visual performance for multiscale shape judgments. It was for that same reason that there was an interest in putting the model to work in an image quality investigation. The model might be expected to produce results that correlate with human performance as image properties that influence the apparent shape of anatomical objects are manipulated.

3.3 Implementation

This section describes the implementation of core-based image analysis. The core computations consist of a multitude of algorithms, such as finding minima of multidimensional functions, differentiation to third or higher orders, eigenvector analysis, and differential equation solving, to name a few. It is beyond the scope of this document to relate the details of these computational recipes, so the reader should consult the references if desired. What is important to understand from this description is the overall operation of the model and the decisions that were made to produce cores in a manner hopefully consistent with the way that humans would carry out the task.

3.3.1 General Principles

As the previous sections have outlined, the core analysis consists of medialness formation followed by ridge tracking to form a connected set of scale space positions that is the core. Much of the medialness theory originates with the dissertation work of Morse,[14,15] while the theory of ridge analysis was proposed by Eberly.[16,17] Fritsch,[18] as well as the author, developed the implementations utilized in this research.

There are several important properties of the receptors, or operators, utilized in core analysis. The medialness operators are implemented as derivatives, or combinations of derivatives, of a Gaussian. In general, first derivative operators will, when convolved with the image intensity distribution, provide a large output at an intensity boundary when oriented properly with respect to that boundary. Other structures, such as object outlines or "bars," may be captured with higher-order derivatives. There is some physiological evidence that receptive fields are indeed shaped like combinations of Gaussians: Young has shown the shape of measured receptive fields in the primate cortex to be approximated by the sum of a Gaussian and its second derivatives.[19] There are additional considerations for the use of Gaussian-weighted measurements in a computer vision application.[20,21] To recognize an object regardless of its position, orientation, or size in a scene, an object representation scheme must operate with measurements that are *invariant* to such transformations. Also, the scale space of measurements must be generated in a causal manner: spurious detail may not arise in the representation acquired with larger scale measurements that did not exist in that made with smaller scales. The Gaussian operator, a solution to the more general diffusion equation, satisfies these invariance and causality requirements. Lastly, it should be noted that for all derivatives, the scale of the measurement, the analog of the receptive field size, is dictated by the standard deviation of the Gaussian.

---

[14] B.S. Morse, S.M. Pizer, A. Liu, "Multiscale Medial Analysis of Medical Images," Information Processing in Medical Imaging, ed. H.H. Barrett, A.F. Gmitro, Lecture Notes in Computer Science 687 (Berlin: Springer-Verlag , 1993): 112-131.

[15] B.S. Morse, "Computation of Object Cores from Grey-Level Images," Ph.D. dissertation, (University of North Carolina-Chapel Hill, 1994).

[16] D. Eberly, "Geometric Methods for Analysis of Ridges in n-Dimensional Images," Ph.D. dissertation, (University of North Carolina-Chapel Hill, 1994).

[17] D. Eberly, S.M. Pizer, Ridges in Image Analysis, Computational Imaging and Vision Series, (the Netherlands: Kluwer Academic Publishers, to appear 1995).

[18] D.S. Fritsch, D. Eberly, S.M. Pizer, M.J. McAuliffe, "Stimulated Cores and Their Applications in Medical Imaging," submitted to Information Processing in Medical Imaging '95 .

[19] R.A. Young, "The Gaussian Derivative Model for Spatial Vision: I. Retinal Mechanisms," Spatial Vision 2 (1987): 273-293.

[20] J.J. Koenderink, A.J. van Doorn, "Representation of Local Geometry in the Visual System," Biological Cybernetics 55 (1987): 367-375.

[21] J.J. Koenderink, A.J. van Doorn, "Receptive Field Families," Biological Cybernetics 63 (1990): 291-298.

Because of the different natures of the two tasks examined in this research, the medialness for the model was implemented somewhat differently in each case. Specifically, medialness for the estimates about blood vessels was determined with a fairly standard operator, the Laplacian. For the core estimates of treatment field distances, medialness was formed from horizontally-constrained, oriented boundariness contributions. The two tasks are described extensively in Chapters 4 and 5, but some justification is provided in Sections 3.3.2 and 3.3.3 for the model implementation decisions.

Cores are computed as one-dimensional ridges in the medialness scale space.† A ridge is simply an extension of the concept of a local maximum in multiple dimensions. In general, a point in an n-dimensional function is a local maximum if it 1) is a critical point, and thus the first derivative, with respect to all directions, of the function is at that point *zero*, and 2) lies at a convexity; that is, the second derivative, with respect to all directions, of the function at that point is *negative*. For a function of two dimensions, ridges correspond to the common terrestrial notion of the crest of a mountain (Figure 3.7).

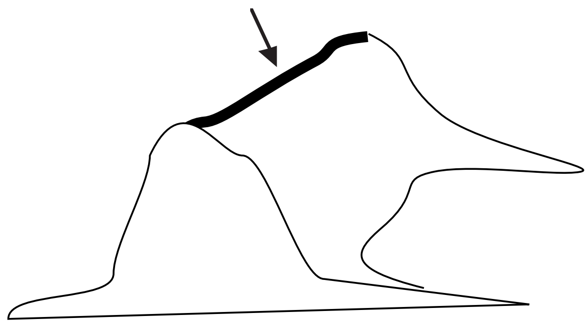


Figure 3.7. Ridge in two dimensions. A ridge in two dimensions corresponds to the crest of a mountain, or those positions where the elevation is a maximum as you hike in the direction of maximum change in steepness.

For a one-dimensional ridge of the three-dimensional medialness function, the previously-described local maxima criteria must only be evaluated along two specially-selected directions. The theory for ridges specifies that the test for the existence of a local maximum of a three-dimensional function f at position $\vec{x}$ is performed in local neighborhood about $\vec{x}$ that is spanned by linearly-independent vectors $\vec{v}_k(\vec{x})$, $1 \le k \le 2$. Of primary interest for ridges of the medialness function, $\mathrm{M}(\vec{x},\sigma)$, is the determination of local maxima with respect to scale. Where the local maximum criteria are satisfied in the scale direction ($\mathrm{M}_\sigma(\vec{x},\sigma)=0$ and $\mathrm{M}_{\sigma\sigma}(\vec{x},\sigma)<0$), a collection of n-dimensional surfaces is in turn defined. Optimal-scale ridges for two-dimensional images then are ridges on the collection of surfaces $\overline{\sigma}(\vec{x})$. A one-dimensional ridge satisfies $\vec{v} \bullet D\mathrm{M}=0$ and $\vec{v}^t D^2 \mathrm{M} \vec{v} < 0$, where $\vec{v}$ is an eigenvector of the Hessian, $D^2\mathrm{M}\left(\vec{x},\overline{\sigma}(\vec{x})\right)$, corresponding to the most negative eigenvalue.

In the particular algorithm used for this application, ridges are computed on a two-dimensional manifold that is obtained by projecting medialness at local scale maxima. While discontinuities in this projected medialness surface may occur that force the termination of any ridge tracking that would be attempted across such positions, in a well-defined medialness function in the region of a single figure such occurrences are infrequent. Local maxima of the projected medialness function are evaluated along the direction of the maximal (negative) second derivative. A core point then is one for which the first derivative of the function, in the direction of the Hessian eigenvector corresponding to the most negative eigenvalue, is zero.

The derivatives computed in the ridge determinations should reflect the principles of scale space. Derivatives, or differences, must be normalized by the scale of the measurement so that they can be meaningful or comparable across scale. For example, a difference of one millimeter may be inconsequential in the comparison of the size of two planets, but that same difference between two molecules would be substantial. Thus, the ridge computations in scale space are performed with scaled derivatives: differences are normalized (multiplied) by the scale of the measurement.

† While Eberly has defined d-dimensional ridges for n-dimensional functions, it is the one-dimensional cores of three-dimensional medialness that will in particular be described here.

Computations are performed with exponential sampling of scale,[22] based upon an initial scale and a base for the exponent ($\sigma_i = \sigma_0 \mathrm{f}^i$), to create a three-dimensional (for a two-dimensional image) scale space of measurements. Consistent with Weber's Law, the exponential metric generates sampling intervals that correspond to perceptually equal increments in the acquisition of spatial information.

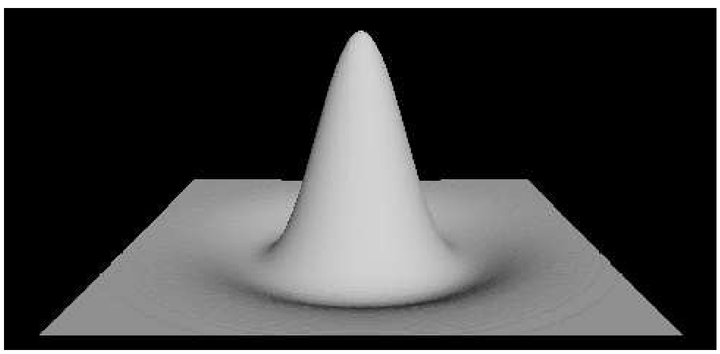
Just as the human visual system focuses its attention at any moment on objects of roughly some size at some position, cores may be calculated by a similar attentional mechanism. Given an initial guess about the scale and spatial position of a point on a core, a system of ordinary differential equations may be solved that specifies the directional flow towards an actual ridge. The solution provides the initial core point. Alternatively, the solution may determine a local minimum, in which case a different guess must be made.

From this initial core point, the ridge may be traversed in two directions, again via the solution of a closed formula ordinary differential equation for the ridge direction. The vector that is the ridge direction may be derived from the surface normal vectors at the initial starting point. Successive points generated during the ridge traversal process are tested against the ridge definition to determine termination.[†]

3.3.2 Medialness for Stenosis Estimation

The objects in question in the angiography task were tall, thin blood vessels. A great deal of preliminary experimentation went into choosing a best medialness operator. Stenosis estimates were computed on simulated angiograms with a range of blur and noise conditions. Several operators[23] were evaluated on the basis of their 1) reliability in producing legitimate stenosis estimates without premature termination and 2) potential consistency with the visual system's behavior for stimuli such as the vessels.

The medialness operator that was chosen, a second derivative, or Laplacian, of a Gaussian kernel (Figure 3.8), is a natural object detector: it reflects the extent to which opposing edges of an object are simultaneously engaged. Medialness at a position and scale then is the convolution of the intensity distribution with the sum of the second partial derivatives of the Gaussian. Medialness for a vessel with a constriction appears as shown in Figure 3.9.



$$\mathrm{M}(x,y,\sigma) = -\sigma^2 \left(\frac{\partial^2 \mathrm{G}}{\partial x^2} + \frac{\partial^2 \mathrm{G}}{\partial y^2}\right) \otimes \mathrm{L}(x,y), \text{ where } \mathrm{G}(x,y,\sigma) = \frac{1}{2\pi\sigma^2}\, \mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}$$

Figure 3.8. Vessel medialness operator. Medialness at position x,y and scale σ is the result of the convolution of a kernel like the one shown with the luminance distribution, L. The filter, $-\sigma^2 \left(\frac{\partial^2 \mathrm{G}}{\partial x^2} + \frac{\partial^2 \mathrm{G}}{\partial y^2}\right)$, is the scale-normalized Laplacian of a Gaussian, G, with scale σ. When the scale of this kernel is such that the kernel's zero crossings co-align with adjacent object boundaries, medialness at the position of the kernel center is maximum.

---

[22] B.M. ter Haar Romeny, L. Florack, "A Multiscale Geometric Model of Human Vision," B. Hendee, P.N.T. Wells, eds., Perception of Visual Information (Berlin: Springer-Verlag, 1991).

[†] Theory for the calculation of cores from a medialness distribution, including the many numerical issues involved with eigensystem and differential equation solution, are documented thoroughly by Eberly in the aforementioned references.

[23] S.M. Pizer, D. Eberly, D.S. Fritsch, "Objects, Boundaries, and Medical Image Analysis," (University of North Carolina Department of Computer Science Technical Report, TR94-069, 1994).

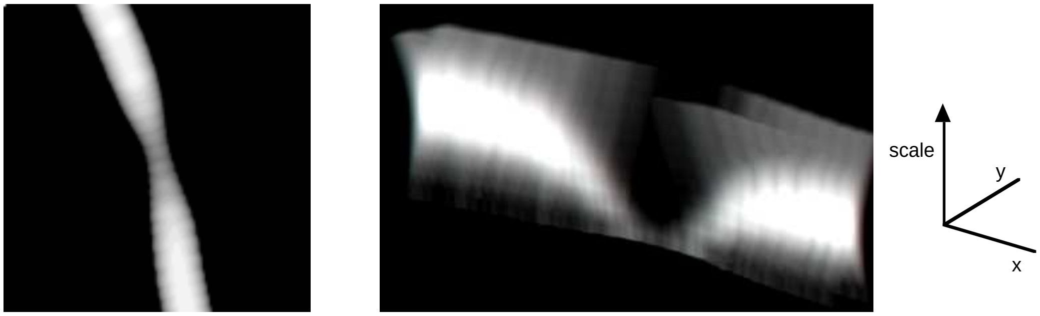


Figure 3.9. Vessel medialness. Volume rendering (middle) of the medialness scale space for a simulated vessel (left). In the rendering, intensities, reflecting the extent to which each position is an object middle, are plotted as a function of the three axes (right): two image dimensions and geometrically increasing scale.

The details of precisely how stenosis estimates were determined from core calculations are described in Section 4.4.

3.3.3 Medialness for Treatment Field Distance Estimation

Because of some psychophysical issues present in the human observer experiment (see Section 6.4), observers were instructed to make a distance judgment between two vertically-oriented edges at a vertical position specified by a marker. Normally the core's distance estimates would be along an orientation that is perpendicular to the edges that contributed to core formation. In an attempt to produce estimates in a manner that was potentially more similar to the human's viewing and decision process, medialness construction was constrained to contributions from only horizontally-neighboring boundariness operators.[24] To accomplish this, spatial derivatives were computed with respect to the horizontal, or x, direction. Thus an expression for the output of a luminance boundariness operator applied at position x,y is simply the value of the normalized Gaussian-weighted derivative of scale σ :

$$\mathrm{B}(x, y, \sigma) = \sigma \mathrm{L}_x^\sigma \qquad 3.1$$

Under the core model hypothesis, the presence of an edge measured at some scale is consistent with the presence of an object middle at some distance away that is proportional to that scale. Thus a proportionality constant, k, relates the scale with the true object radius via $r = k\sigma$. The value at some position in the medialness scale space is the accumulation of the boundariness responses from the two adjacent positions along the horizontal a distance r away:

$$\mathrm{M}(x, y, \sigma) = \mathrm{B}(x + r, y, \sigma) + \mathrm{B}(x - r, y, \sigma) \qquad 3.2$$

It was also the case in the distance estimation experiments described later that the images were always oriented with the anterior anatomy at the right of the image. This meant that the human observers knew and expected that the orientations of the two edges between which the distance estimate was made were always the same. The edge formed by the dark interior of the treatment field and the edge at the posterior of the vertebral body generated an intensity profile such as that in Figure 3.10.

This additional *a priori* information, clearly available to the human, was built into the model. Boundary operators were tailored to only contribute to medialness when the edge orientation was appropriate (Figure 3.11).

---

[24] C.A. Burbeck, et.al., "Linking Object Boundaries at Scale."

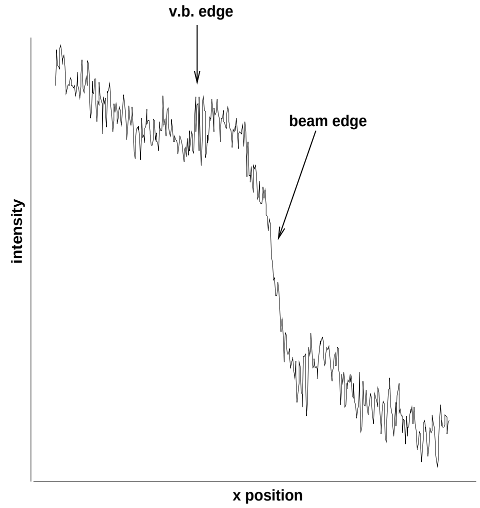
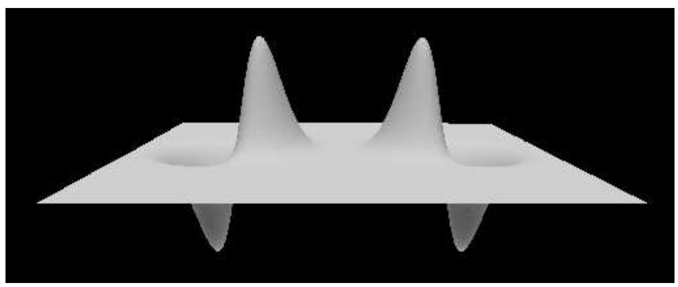


Figure 3.10. Portal image edge profile. The edges between which observers were asked to judge a distance possessed the orientations shown in this horizontal profile through a typical image. The vertebral body edge exhibited an increase in intensity with respect to the positive x direction while the treatment beam edge decreased in intensity.



$$\mathrm{M}(x,y,\sigma) = \mathrm{B}(x-r,y,\sigma) + \mathrm{B}(x+r,y,\sigma)$$

$$\text{where } \mathrm{B}(x-r,y,\sigma) = \begin{cases} -\mathrm{L}_x^\sigma, & -\mathrm{L}_x^\sigma > 0 \\ 0, & -\mathrm{L}_x^\sigma < 0 \end{cases}$$

$$\mathrm{B}(x+r,y,\sigma) = \begin{cases} \mathrm{L}_x^\sigma, & \mathrm{L}_x^\sigma > 0 \\ 0, & \mathrm{L}_x^\sigma < 0 \end{cases}$$

Figure 3.11. Field distance medialness operator. Medialness at position x,y,σ is calculated by convolution with a kernel like the one shown. When separated by a distance such that they simultaneously engage the horizontally adjacent edges, the oriented boundariness operators contribute to a maximum medialness response at the position corresponding to the center of the kernel. The scale of the boundariness operators is related to the separation between them by a proportionality constant.

Further details about distance estimates derived from the cores that employ this medialness function may be found in Section 5.4.

3.4 Summary

The core model has several properties that make it advantageous as the visual model for this image quality approach. The core provides an extremely efficient mechanism for representing greyscale figures. The width or distance judgments that form the basis for many diagnostic estimates in medical image viewing may be extracted readily from the core representation. Furthermore, the core captures shape information at the scale of the figure width. The result is that the core for a reasonably-sized figure is to a great degree independent of small fluctuations at the object boundary, small scale detail such as noise, minor edge degradations caused by blurring, or even boundary contrast. However, to the extent that the structure of the core *is* influenced by image characteristics such as blur, noise, and contrast, it is hoped that this research will show that the changes in the estimates derived from the core are consistent with variations in human performance under those same circumstances.

There is evidence to suggest that the core is constructed from fundamental visual measurements that the human visual system employs. If the core is indeed constructed from receptors that operate in the same manner as those found in the visual system, then the estimates from the core representation might be influenced in the same way as human judgment in the presence of changing physical characteristics. This argument from physiological principles, together with psychophysical evidence, prompts the hope that the core model might be predictive of human performance. It was therefore of interest to test whether this model might be applied in medical image quality investigations where it would be particularly advantageous to model the human.

While it is the case that the results of the image quality investigations in this research may have much to say about the efficacy of the core model as a theory for human vision, this dissertation attempts to develop a predictive measure of medical image quality and is not explicitly a test of any visual model. The research generated both information and questions related to core construction and variability in the presence of image conditions that affect quality, but for the most part those issues are left as feedback for others. The purpose of the dissertation was to apply, not develop, this particular model. Similarly, many estimation tasks, such as the two described in the previous chapter, involve the perception of shape, and the core model might be recruited in calculating many of them. But if the task decided upon had been of a different nature, say the detection of lung nodules, the model adopted might have been different.

Also, while the core model computations will be used here to perform several tasks, this research is not about the development of methods for computer-aided measurement and diagnosis. Task performance is used here as the basis for the functional determination of image quality. It is hoped that the model will perform in the same way that the human would, even if that performance is not entirely accurate or optimal.

The first part of this dissertation has attempted to motivate the need for a computed image quality approach that would be predictive of human performance in many task settings. This chapter described the visual model that is the foundation for the image quality approach proposed in this research. The purpose of this research is to test such a proposal, and the next part of this dissertation outlines the imaging systems and tasks that served as the bases for that evaluation.