LBL-32546
UC-405
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

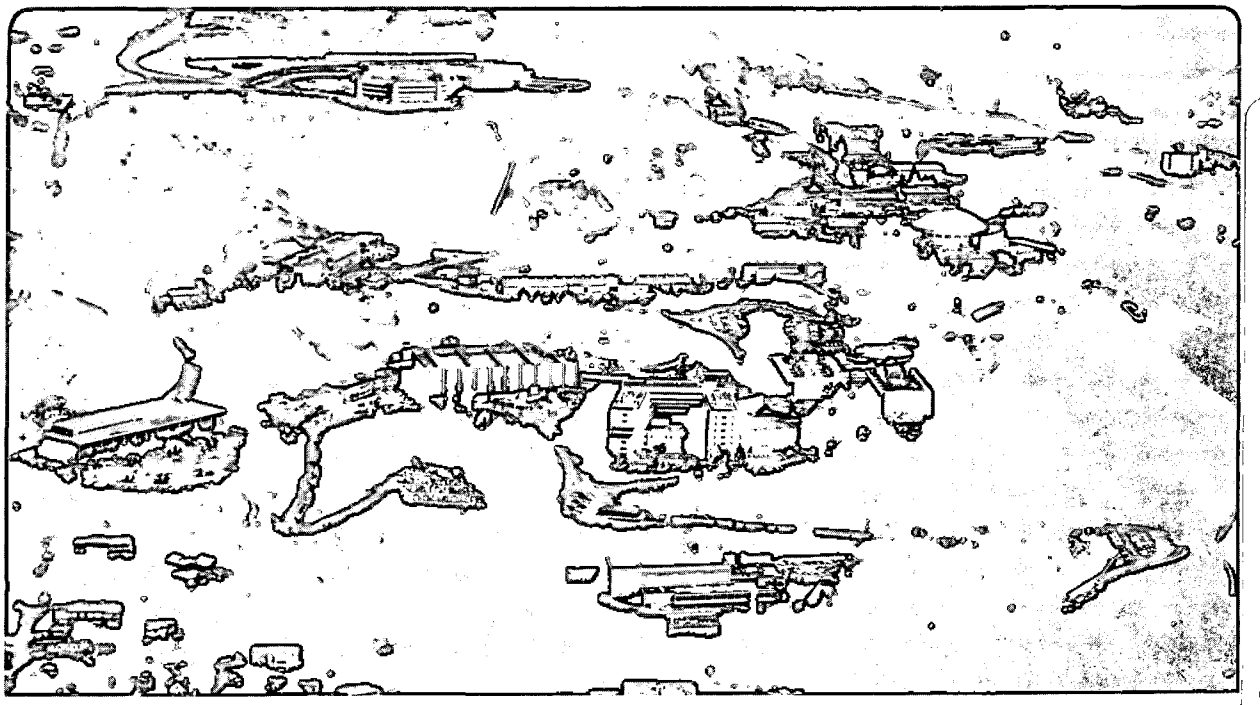
Engineering Division

Submitted to IEEE Engineering in Medicine and Biology

The Representation of Medical Images by Visual Response Functions

J. Llacer, B.M. ter Haar Romeny, L.M.J. Florack,
and M.A. Viergever

July 1992



REFERENCE COPY
Does Not
Circulate
Bldg. 50 Library.
LBL-32546
Copy 1

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-32546
UC-405

The Representation of Medical Images by Visual Response Functions

Jorge Llacer

Engineering Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

and

Bart M. ter Haar Romeny, Luc M.J. Florack, and Max A. Viergever

University Hospital Utrecht
3584CX Utrecht, The Netherlands

July 1992

This work was supported in part by the Director, Office of Energy Research, Office of Helath and Environmental Research, Physical and Technological Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

The representation of medical images by Visual Response Functions

Jorge Llacer
Engineering Division, Lawrence Berkeley Laboratory
Berkeley, CA 94720, USA

Bart M. ter Haar Romeny, Luc M.J. Florack
and Max A. Viergever,
University Hospital Utrecht,
3584CX Utrecht, The Netherlands

Medical Imaging, beyond the presentation of X-ray images in photographic plates, has received an important boost in the last two decades with the development of new technologies and their application to medical diagnostics: X-ray Computer Assisted Tomography, Ultrasound Imaging, Positron Emission Tomography, Single Photon Emission Tomography and Magnetic Resonance Imaging, just to mention those modalities in wide spread use in the clinic. Research is continuing in both the instruments and in the algorithms used to extract useful information from them and new generations of "scanners" are coming to the market with regularity. The final objective of those instruments and algorithms is, in general, the presentation of an image to a physician which is useful in establishing a diagnostic. There is also a substantial amount of research on how best to present those images and on how to extract features from the images to assist the physician's task. Ultimately, one could conceive of an automated image scanning machine that only alerts a physician when some characteristics of the image indicate a certain possible abnormality. The latter is still a remote goal, but substantial inroads are being made into "segmenting" medical images into its component parts in a meaningful way. A current example would be the extraction of all the contours of a particular organ from a set of Magnetic Resonance Imaging (MRI) contiguous image "slices" and then presenting that organ as a separate entity in a three-dimensional visualization screen, which currently can be done more or less automatically at a number of medical institutions for specific situations.

Segmentation is one of the areas of image analysis that requires some way of representing image features mathematically. Two more research areas which have a similar requirement are: Finding a specific feature or structure in an image by computer and the description of anatomical "prior" knowledge to aid in the reconstruction of tomographic data. So, the question is how does one represent image features mathematically? This article does not attempt to answer that question in a general way, but will focus, instead, on the use of a natural representation in terms of Visual Response Functions (VRF's), based on the pioneering work on multi-scale vision receptive fields of J.J. Koenderink, at the University of Utrecht. With the exception of the work being carried out by S.M. Pizer, J.M. Coggins and associates at the University of North Carolina, Chapel Hill, we feel that the medical imaging engineering community in the U.S. and abroad has not been sufficiently exposed to the use of

Koenderink's ideas on image analysis. This article is intended to begin providing a remedy for that and it describes work that some of us have published elsewhere, as well as new work that can be useful for the representation of "prior" information in image reconstruction.

SCALE SPACE AND VISION RECEPTIVE FIELDS

The concept of scale and the origin of the operators needed to represent images in a multi-scale environment have been described in substantial detail in [1, 2] using concepts of differential geometry. The following is a practical summary of the ideas presented there. The concept of scale corresponds directly with the intuitive idea of the meaning of that word: Some things have a large scale, i.e., their structure is gross and their frequency spectrum contains mostly low frequencies. Other things have a small scale, i.e., their structure is fine ahd their frequency spectrum contains high frequencies. When observing an image, however, the smallest and largest scale are well defined and provide a frame of reference for the meaning of scale in that image: the smallest scale, the *inner scale*, has a lower bound determined by the size of the picture elements (grain size in a photographic film, element size in a CCD detector, etc.). In the absence of degrading factors that can introduce correlation between elements, that lower bound becomes the inner scale. The largest scale, the *outer scale*, has an upper bound determined by the size of the image itself. In practice, that upper bound is usually the outer scale. The image can only contain features with scales between the inner and outer scales.

An image can be observed at different scales: one can focus one's attention to some small details in a corner, for example, or look at the shape of the large objects that contain those details. We need an operator (or set of operators) that allow an observer (human or machine) to move in "scale space" and extract features that appear in the image at different scales. A *scale* parameter s is defined as a scalar quantity corresponding to the scale at which an image is being observed and *scale space* is defined as a family of images generated in a continuous manner from a given image L by means of convolution with a suitable spatial kernel [3,4,5,6] of size determined by the 1-dimensional parameter s.

The optimal linear kernel $G(\vec{r},s)$ for front-end vision blurring can be found by considering some fundamental requirements:

1 - Linearity: allowing for superposition of input stimuli.

2 - Spatial shift invariance: follows from the assumption that there should be no preferred location within the field of view. The operator that transform the image when we increase scale has to be a *convolution* with some suitable spatial kernel.

3 - Isotropy: there should not be a preferred direction.

4 - Scale invariance: There should be no preferred scale.

The above conditions lead necessarily to the Gaussian kernel

$$G(s) = \frac{1}{\left(2\pi s^2\right)^{D/2}} e^{-\frac{\vec{r}\cdot\vec{r}}{2s^2}}, \tag{1}$$

where D is the dimensionality of the image space.

One very important characteristic of the Gaussian kernel is that it is the only kernel that leads to monotonic destruction of detail under repeated blurring. Gaussian blurring is a causal process, no *spurious resolution* is created, i.e., no new details will appear at a given scale s that did not exist at smaller scales.

The Gaussian kernel (1) is the Green's function of the isotropic diffusion equation. In two dimensions that is

$$\frac{\partial L}{\partial s} = \frac{\partial^2 L}{\partial^2 x} + \frac{\partial^2 L}{\partial^2 y} \tag{2}$$

and the process of going up in scale space is easy to visualize if we consider L to be a heat distribution at time t = 0 in a metal plate, for example, and with the passage of time (analogous to increasing s) the heat distribution becomes more and more uniform until the temperature is constant everywhere in the plate.

The above front-end vision formulation is very elegant but suffers from at least one drawback when compared to the way human vision appears to operate. When a human looks at a large scale object in an image, for example, the edges of that large object may appear perfectly sharp while the convolution of an image with a Gaussian kernel of large s will result in blurred edges. In order to remedy this problem, work is being carried out in several research groups based on the concept of an anisotropic diffusion equation, first described by Perona and Malik [7], to replace (2).

The Gaussian kernel (1) is not the only solution to the diffusion equation (2). Partial derivatives of (1) are also solutions and the methodology of image analysis in scale space has been extended to include partial derivatives of (1) of different orders as useful kernels in image analysis. A scaled differential operator in two dimensions can be described as

$$G_{n,m}(x,y,s) = \frac{\partial^{n+m} G(x,y,s)}{\partial x^n \partial y^m}. \tag{3}$$

The effects of the Gaussian kernel (1) and the differential operators (3) operating on an image show a remarkable resemblance to the processing occurring in human vision in the connections between the retina and the first layer of ganglion cells behind it. *Receptive fields* are circular areas of different sizes in the retina whose rods and cones are connected to a single ganglion cell. The latter responds to light intensity or to derivatives of light intensity up to perhaps 4th order at a scale corresponding to the circle size. Hubel [8] has written a highly recommended book on the subject. Regarding the operation of differentiation of an image by convolution with a differential kernel (3), it should be pointed out that the spatial integration effect of the convolution operation more than compensates for the noise increase inherent in the differentiation of noisy images, so that the outcome is well behaved even to high orders[9]. Because of the similarity of (1) and (3) to the functions governing front-end human vision, we shall call those kernels Visual Response Functions (VRF's).

INVARIANT VISUAL RESPONSE FUNCTIONS

In extracting information about a specific feature from an image, we do not want, in general, the results of the operation to depend on the orientation of the feature. We could say, for example, that the "edgeness" or "cornerness" of a specific element in an image should not change if the element is rotated. Bearing in mind that in image analysis we need to use a coordinate system, we will be interested in combinations of VRF's that are invariant under rotation in the chosen coordinates. An infinite number of invariant VRF's can be generated with the tools of tensor calculus and the derivation of a number of interesting VRF's up to third order is shown in detail in [1]. It is also shown in that publication that there are five irreducible invariants of second order from which all the other invariants of the same order can be formed. For the purposes of this paper, we shall examine several invariant VRF's (irreducible or not) whose application to image representation and analysis is evident. Before doing that, however, we have to establish a number of concepts which will become particularly clear when discussing the specific examples. We will restrict the discussion to 2-dimensional images, although a generalization to three dimensions is straight forward.

PRELIMINARY CONCEPTS

The image L: We will consider the image L as being given by a surface above the (x,y) plane whose height at each point corresponds to the intensity at that point. In this way, we can talk about mountains, valleys, etc.

Isophotes: This are lines of equal intensity, corresponding to the familiar lines of equal elevation in topographical maps.

Summation convention: We use the Einstein convention, which implies summation over all repeated indices. For example, $L_i L_{ij} = \sum_i L_i L_{ij}$.

Notation in Cartesian coordinates: Subscripts following L imply the operation of convolving the image L with the Gaussian derivative kernels of (3); in two dimensions, i and j can each take the values x and y. For example, $L_{ii} = \frac{\partial^2 L}{\partial^2 x} + \frac{\partial^2 L}{\partial^2 y}$.

The (v,w) gauge coordinates: VRF's are local in nature, i.e., they gather information over an image region with size determined by the scale s. It is often convenient to use a coordinate frame (v,w) which is a rotated version of the (x,y) frame, with a rotation angle that depends on the image properties at the point P in whose neighbourhood we are interested. Particularly useful is the frame in which the w-axis is along the gradient of the image at point P, i.e., along the direction of steepest ascent. Then, the v-axis is tangent to an isophote at point P.

Homogeneity of degree n: As used in this work, the term indicates behaviour of an invariant VRF under multiplication of image values by a constant. For λ = constant, we say that a VRF given by $f(L)$ is homogeneous of degree n in L if

$$f(\lambda L) = \lambda^n f(L). \tag{4}$$

For example, if $L \rightarrow \lambda L$, then $L_i L_i \rightarrow \lambda^2 L_i L_i$. The invariant VRF $L_i L_i$ is homogeneous of degree 2. Homogeneity describes how the result of applying a particular invariant VRF wil be affected by a uniform change in illumination, for example.

General intensity transformations: Florack et al, in [10], describe in detail the behaviour of a specific form of the five second order irreducible invariant VRF's that allows a clear visualization of what are the effects of operating on an image which suffers an intensity transformation of a more general nature than the simple multiplication by a constant described above. The particular transformations considered are:

1) The intensities L are transformed into a new set

$$\tilde{L} = \tau(L) \quad \text{with } \frac{d\tau}{dL} > 0. \tag{5}$$

An example of this transformation would be when the illumination on an image is increased non-uniformly, with the new intensity at a point P depending only on the original image intensity at the same P. The condition on the derivative insures that the change in illumination is reversible, i.e., going back from $\tilde{L}$ to L is single valued. For this case, very interesting results are obtained for first and second order VRF's:

$$\tilde{L}_i = L_i \tau' \tag{6}$$

$$\tilde{L}_{ij} = L_{ij} + L_i L_j \tau'' \tag{7}$$

where the primes imply derivatives with respect to L. We will make use of those results below.

2) The intensities L are transformed into a new set

$$\hat{L} = a + bL, \tag{8}$$

i.e., the original intensities are multiplied by a constant and a "pedestal" is added. We note that this case is included in case 1). As we shall see below, however, an image that undergoes the less general transformation (8) can be analyzed by one additional VRF with characteristics that are easy to interpret.

Zero crossings: The zero crossings of some VRF's play a significant role in image representation. We examine here what is the general effect of intensity transformations on the location of zero crossings of a VRF by looking at one example. Let's consider the Gaussian Laplacian operator L_{ii} whose zero crossings are often used as an edge finder (more on this below). We start with (7),

$$\tilde{L}_{ii} = L_{ii}\tau' + L_i^2\tau'' \tag{9}$$

and consider first the case in which the intensities are transformed by (8). Then, replacing into (9) we get

$$\tilde{L}_{ii} = bL_{ii} \tag{10}$$

and the locus of points where $L_{ii} = 0$ will also be the locus of points with $\tilde{L}_{ii} = 0$.

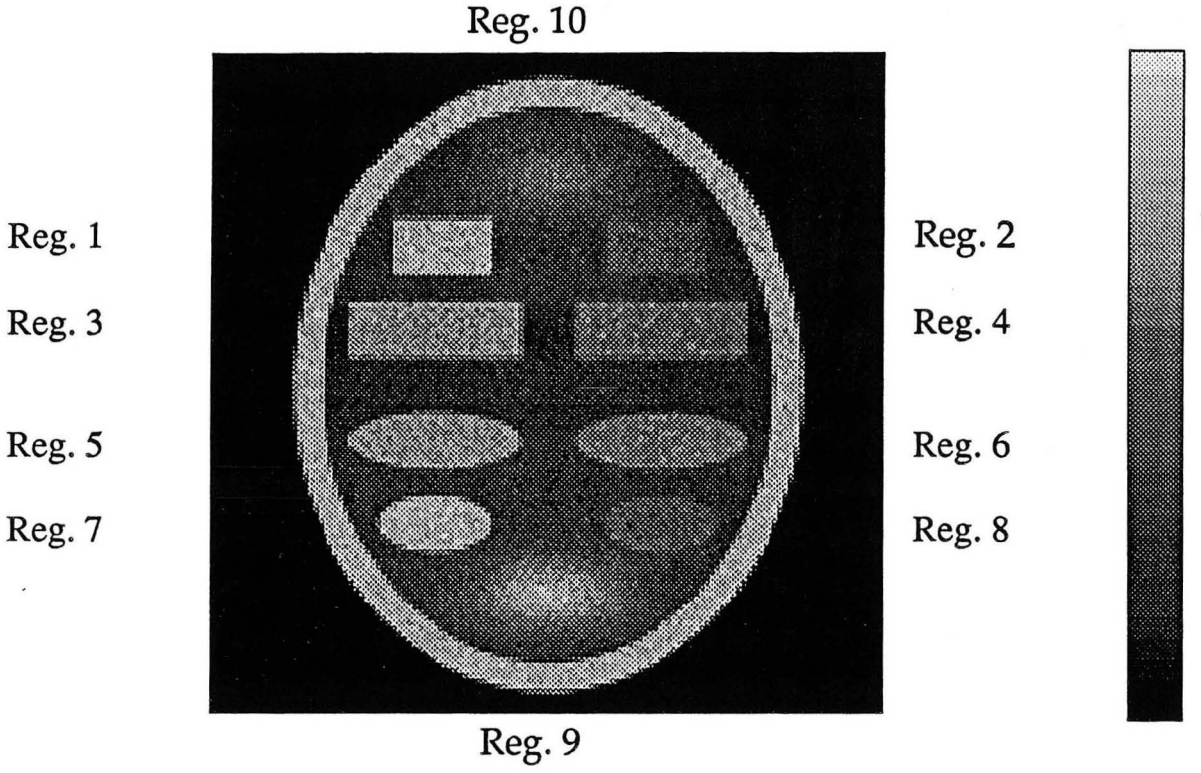
For the more general transformation (5) the result is

$$\tilde{L}_{ii} = L_i^2\tau'' \tag{11}$$

at the locus of $L_{ii} = 0$. Thus, the zero crossings of L_{ii} may shift significantly under the more general intensity transformation, but the effect will depend on the characteristics of the image, the scale s and the form of the transformation. Since VRF's are local operators, the general transformation (5) can be approximated by a transformation of the type given by (8) in a region surrounding the point of interest at a desirable scale, so that the zero crossings of the Gaussian Laplacian are useful in defining some local features of an image, even if the illumination changes substantially.

SOME INTERESTING VRF'S

In order to study the behavior of VRF's when operating on an image, we have devised a very simple phantom, shown in Fig. 1. It consists of a series of rectangles and ellipses of different intensities placed in a background of 20% intensity, surrounded by an elliptical ring of



Relative intensities:

Reg. 1 : 100%	Reg. 2 : 40%
Reg. 3 : 80%	Reg. 4 : 60%
Reg. 5 : 80%	Reg. 6 : 60%
Reg. 7 : 100%	Reg. 8 : 40%
Reg. 9 : max 90%	Reg. 10: max 45%

Fig. 1 - Simple reference phantom to be used to demonstrate the effect of different Visual Response Functions.

100% intensity. Two Gaussian regions are also included. The relative intensities of the different elements are shown in the figure. The image contains Poisson noise corresponding to a total of 10^6 photons.

The VRF's that we shall examine here are members of the second order irreducible invariants, in two different representations: in cartesian coordinates and in (v,w) gauge coordinates. A third representation in gauge coordinates which is most appropriate to study intensity transformations will also be considered. Each of these three representations is complete in the sense that any member of one representation can be generated in terms of the complete set of members of any of the other representations. In the listing of each of the representations, the invariants that we shall examine in this paper are in bold characters.

1) In cartesian coordinates, $J \equiv \left\{ \boldsymbol{L}, \boldsymbol{L_i L_i}, \boldsymbol{L_i L_{ij} L_i}, \boldsymbol{L_{ii}}, \boldsymbol{L_{ij} L_{ji}} \right\}$.

a) The image L: Observing the phantom of Fig. 1 at different scales corresponds to convolutions of the image with Gaussian kernels with different parameters s. The results are of a familiar nature but will be presented along with the description of other VRF's to help in understanding their effects.

b) The Gaussian Laplacian L_{ii}: The effect of this kernel is to extract a function that is closely related to edges. In particular, the zero crossings of this kernel have been used for edge detection [11]. We note, however, that transforming the operator to gauge coordinates gives an interesting insight on what the zero crossings actually yield.

$$L_{ii} = L_{vv} + L_{ww} = L_{ww} + \frac{L_{vv}}{L_w} L_w \tag{12}$$

The invariant $-\frac{L_{vv}}{L_w}$ yields the *isophote curvature*, to be discussed further below. If we define an edge as the locus of points of inflection defined by the vanishing of the second derivative of the image along its gradient direction, i.e., $L_{ww} = 0$, than the zeros of L_{ii} will define that edge approximately only if the isophote curvature is very small at the point of interest. For that reason, the L_{ii} edge finding filter often fails as an edge finder at small indentations[1]. Figure 2 shows the image L, L_{ii} and the zero crossings of L_{ii} superimposed on L at two different scales, $s = 1.65$ pixels (left column) and $s = 4.48$ pixels (right column), representing a small scale and a medium scale for the elements of the phantom.

A number of observations can be made from Fig. 2. First, we see that the zero crossings of L_{ii} behave quite well as edge finders in spite of the sharp edges of the rectangular sections and that the shape of the edges found is basically independent of the intensity of the elements, to the extent that the element intensity is not affected by statistical noise. In fact, the intensity transformation in going from Region 1 to 2, for example (labels on Fig. 1), is of the type

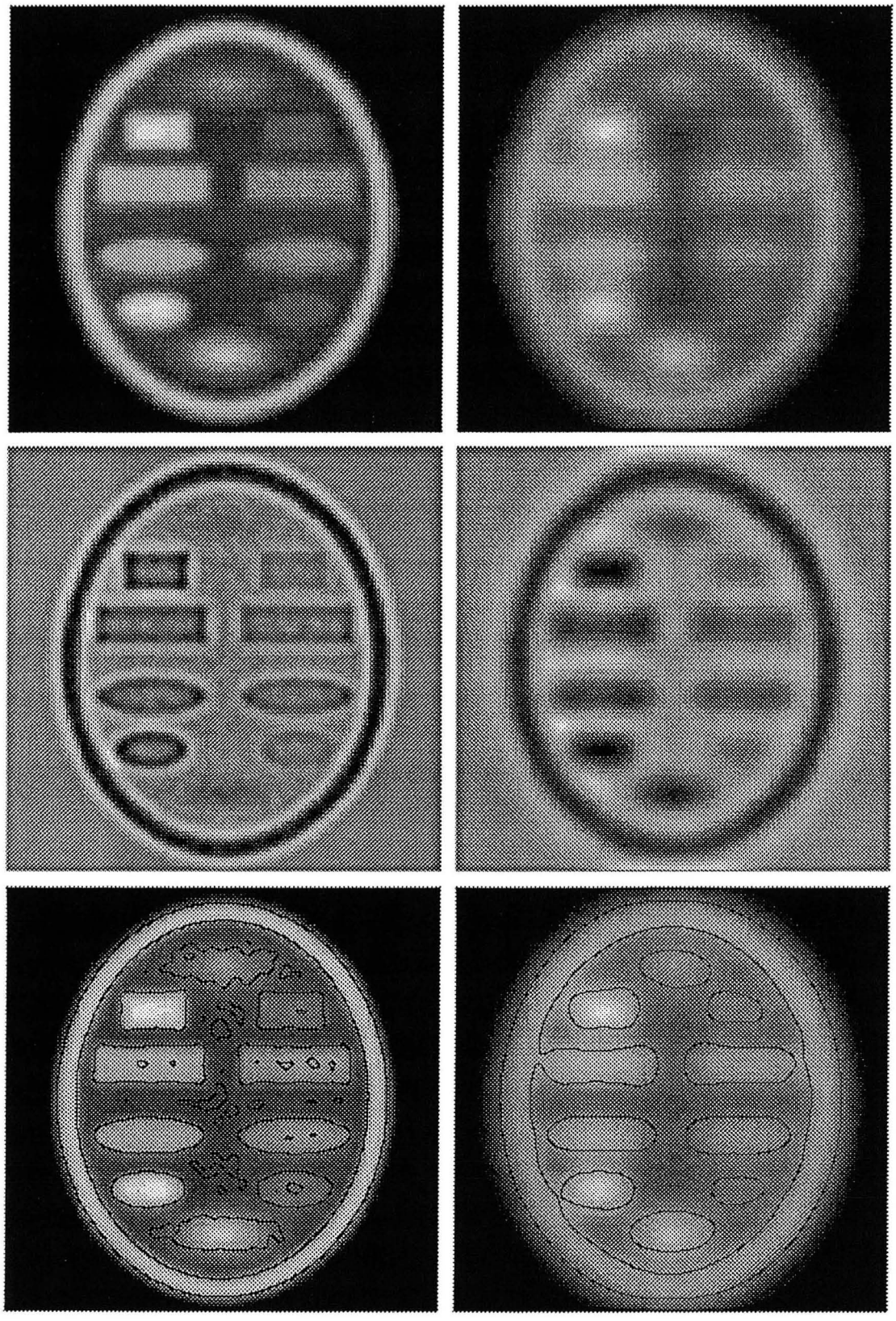


Fig. 2 - Top: the image L seen at two different scales, σ=1.65 and 4.48 pixels. Middle: The Laplacian operator applied to the reference phantom, at the same two scales. The zero level for these two images is the background grey. Bottom: The zero crossings of the Laplacian operator superimposed on the the top images.

described by (8) and we have already seen that, in that case, zero crossings are independent of intensity. We also note that, at small scales, the zero crossings of L_{ii} may bring out features due exclusively to noise and that edges in cases of elements that do not have abrupt bounds (like the Gaussian elements in Fig. 1) may be strongly distorted in the presence of noise. At higher scales, though, the effects of noise become less obtrusive.

At large scales, the L and L_{ii} operators yield results that appear to be too removed from the original image to be of much use. However, Pizer [12] is working on a description of shape based on medial axes, obtained from the loci of maxima of L_{ii} as a function of s which makes use of all scales. Research on segmentation of medical images is now underway [13] that can be based on following the maxima of L down scale space, for example, starting from an adequately large scale.

The complete range of scales is also being considered as a way to represent prior information in the iterative reconstruction of tomographic images. In the case when only the geometrical properties (but not the intensities inside specific features) of the "prior" image are to be represented, the zero crossings of L_{ii} (among other VRF's) can be used to convey feature shape information [14]. Figure 3 shows a cut through scale space of the zero crossings of L_{ii} for the phantom of Fig. 1. As s increases, the closed loci of $L_{ii} = 0$ at fixed s become closed surfaces in the 3-D scale space. An iterative Bayesian reconstruction of noisy tomographic data using relative noise-free prior information given in the form of Fig. 3, for example, can proceed by penalizing a reconstruction whose $L_{ii} = 0$ loci in some range of scales are sufficiently different from the prior values. At the early iteration steps, when the image estimates do not have much detail, the range of interest may be from large values of s to some intermediate values. As the iterative process continues, the penalizing can be extended or shifted to lower values of s.

c) The *deviation from flatness* $L_{ij}L_{ji}$: This invariant yields an image whose brightness is proportional to the deviation from flatness in the original image. It is interesting to remark that a flat plane that is inclined with respect to the vertical is found to be flat by this VRF. In practice, it may be more useful to consider the form $\sqrt{L_{ij}L_{ji}}$, which is homogeneous of order 1/2, making it less sensitive to illumination than in the original form. Figure 4 shows the (square root of) deviation from flatness of the image L at scales $s = 1.65$ pixels (left) and 4.48 pixels (right). The tops of the features are indicated as being flat at the smaller scale, but not so at 4.48 pixels, as expected. The inclined planes at the edges of the features are found to be reasonably flat in both cases, more so in the smaller scale.

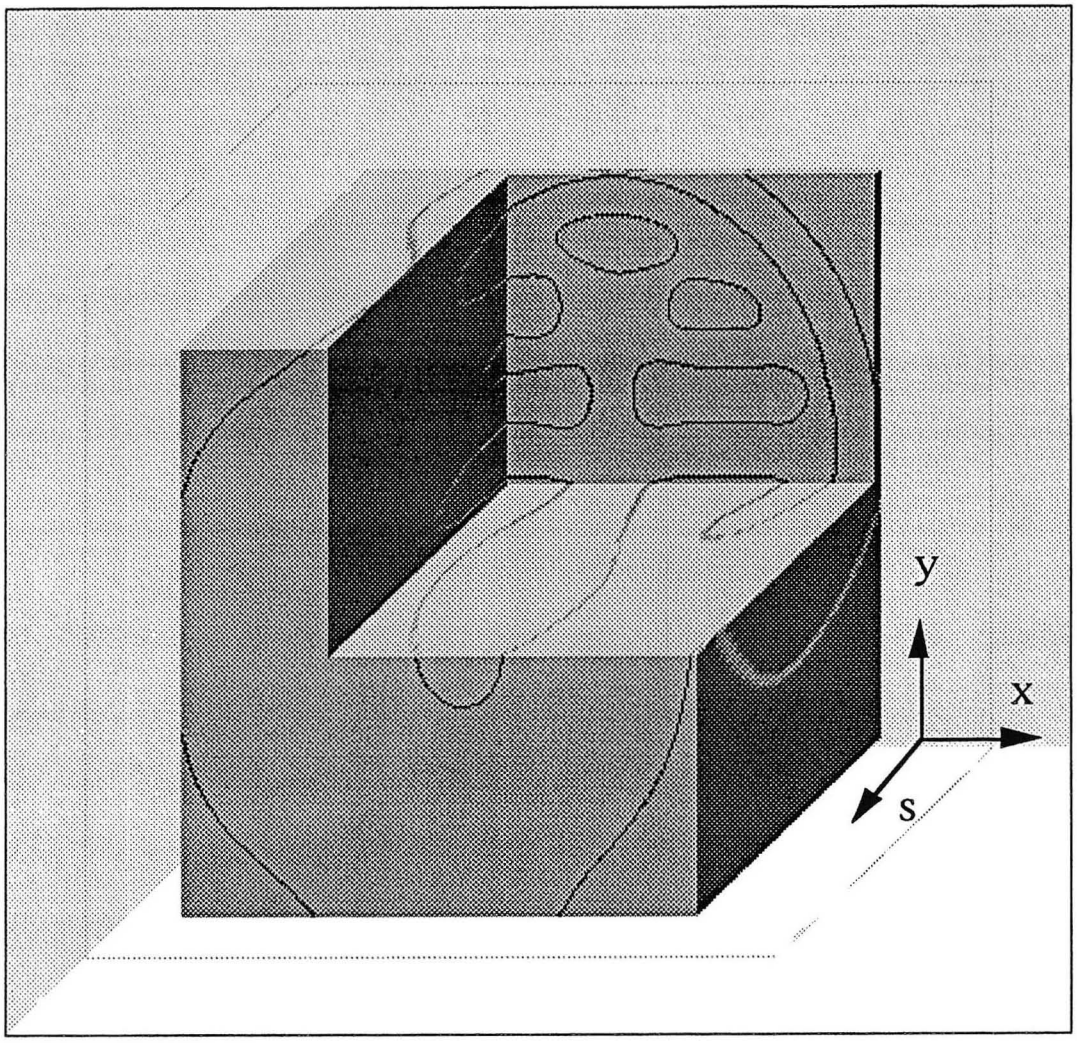


Fig. 3 - Behavior of the zero crossings of the Laplacian operator as the scale parameter *s* increases.

2. Representation in gauge coordinates, appropriate for the investigation of behavior under intensity transformations: $H = \left\{ L, L_w, \frac{L_{vv}}{L_w}, \frac{L_{vw}}{L_w}, \frac{L_{ww}}{L_w} \right\}$.

a) The *edgeness* or *gradient magnitude* $L_w = \sqrt{L_w^2} = \sqrt{L_i L_i}$. This is equivalent to the Canny edge detector for steep edges [15, 16]. It yields the magnitude of the slope of the tangent plane at a point P. It is homogeneous of degree 1, i.e., the edgeness of a feature will be proportional to the local intensity and, in that respect, it behaves quite differently from the zero crossings of L_{ii}. Figure 5 shows the edgeness of the phantom of Fig. 1 at the same two scales indicated for Fig. 2.

b) The *isophote curvature* $-\frac{L_{vv}}{L_w}$. This invariant yields the curvature of isophote lines (lines of equal intensity) at all points in the image. This VRF is independent of transformations of the general type given by (5). Under those changes of illumination, the shapes of the isophote lines will not change, although the absolute intensity represented by a particular line will change. Figure 6 (left) shows the image L at a scale s = 3.49 pixels on which isophote lines corresponding to increments in brightness of 10% of maximum have been superimposed.

Also shown (right) is an image of $-\frac{L_{vv}}{L_w}$ with the same lines superimposed. It is evident that the isophote curvature image is bright at the points of high isophote curvature, independent of whether a particular feature is at full intensity or not. It must be pointed out that this invariant has to be used with some care: The denominator L_w is the magnitude of the gradient along the direction of maximum ascent, which goes to zero at local maxima or minima of the image. The numerator L_{vv} should also go to zero at those same points, with a result that is indeterminate. In practice, in the presence of noise, large positive and negative values appear in the results of the convolution at those points, visible in Fig. 6 as spots of maximum brightness and darkness. In using this invariant, it may be necessary to establish a "degree of confidence" factor in the determination of isophote curvature, based on the value of L_w.

c) The *flow line curvature* $-\frac{L_{vw}}{L_w}$. This VRF generates an image whose brightness coresponds to the curvature of the flow lines generated by the image gradient vector field. Those flow lines are perpendicular to the isophotes being crossed. The results for the image L at s = 3.49 are shown in Fig. 7 (left). This VRF is also invariant with respect to transformations defined by (5), so that the left and right sides of Fig. 7 (left) are symmetric, except for noise effects. The somewhat unstable behavior at points L_w = 0 is also evident in this VRF. Also in Fig. 7 (right) is the image L with the zero loci of flow line curvature

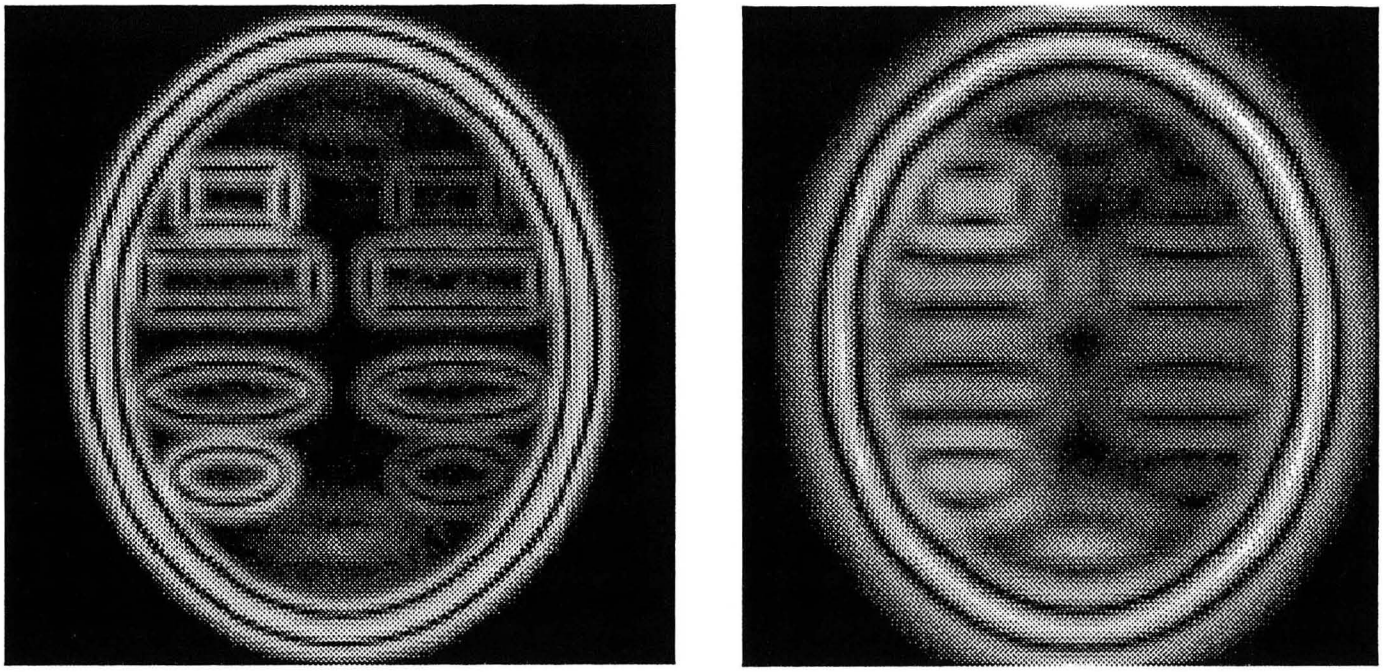


Fig. 4 - The *deviation from flatness* operator applied to the reference image at two scales, σ=1.65 and 4.48 pixels. Flat regions appear close to the zero background level.

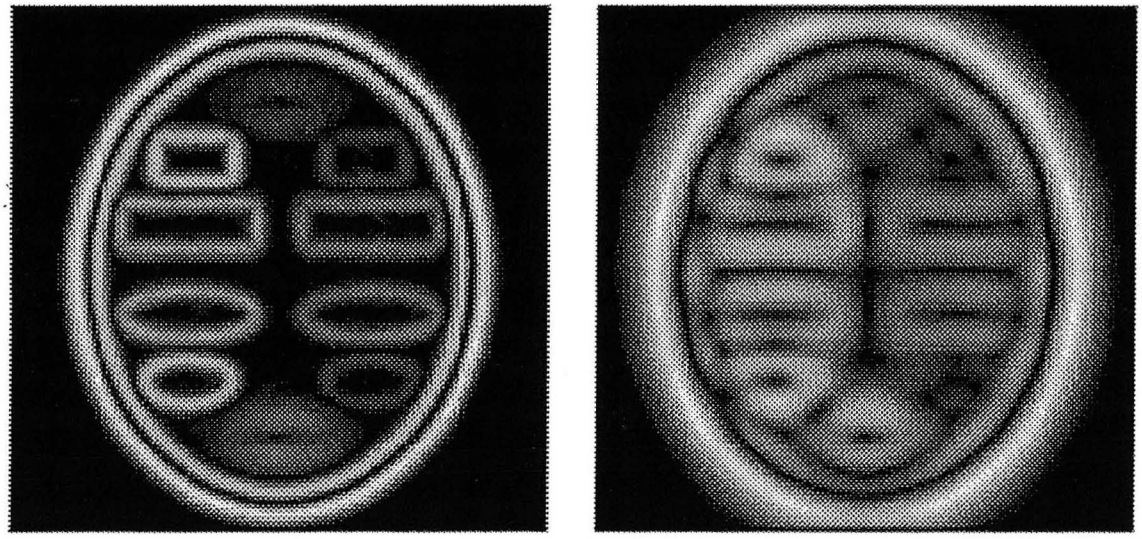


Fig. 5 - The *edgeness* operator at σ=1.65 and 4.48 pixels.

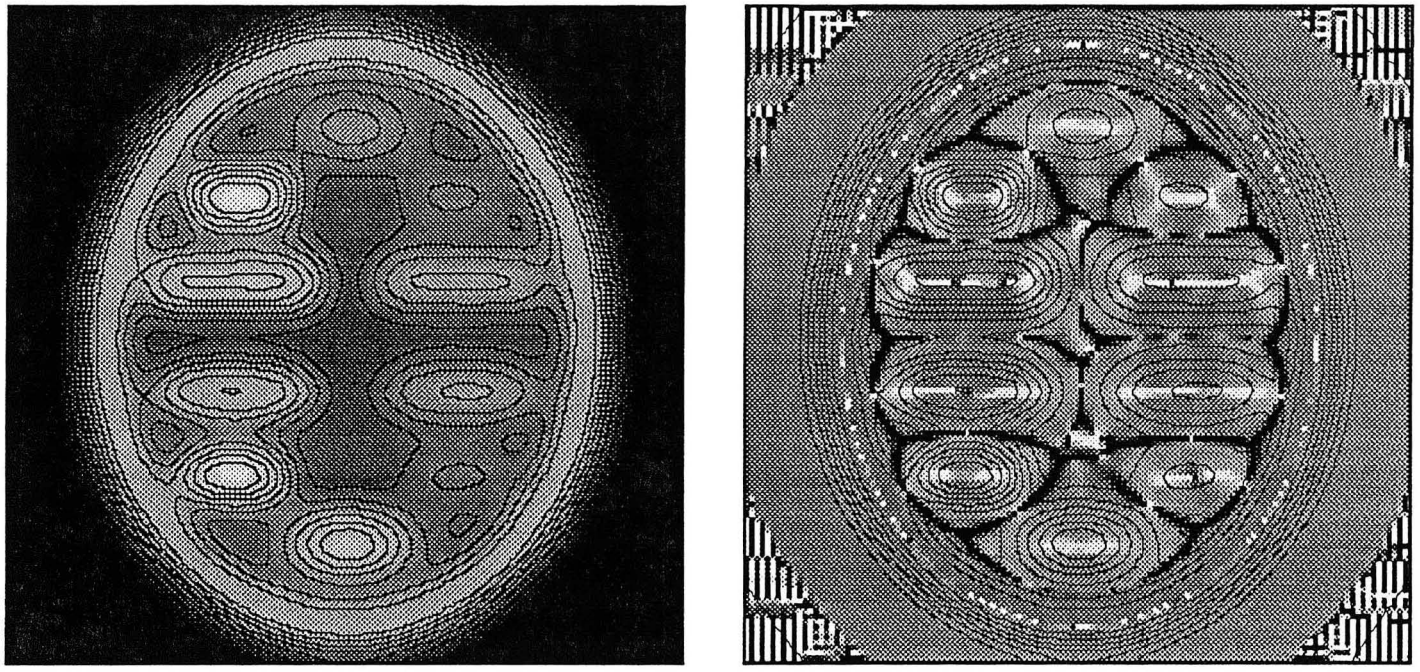


Fig. 6 - The *isophote curvature* operator at a scale of σ=3.49 pixels. The left image shows isophote (isointensity) lines superimposed on the reference phantom L at that scale. The right image shows the isophote lines superimposed on the results of applying the isophote curvature operator in the reference phantom. The image is brighter at places where the isophote lines are most curved. Middle grey is zero curvature.

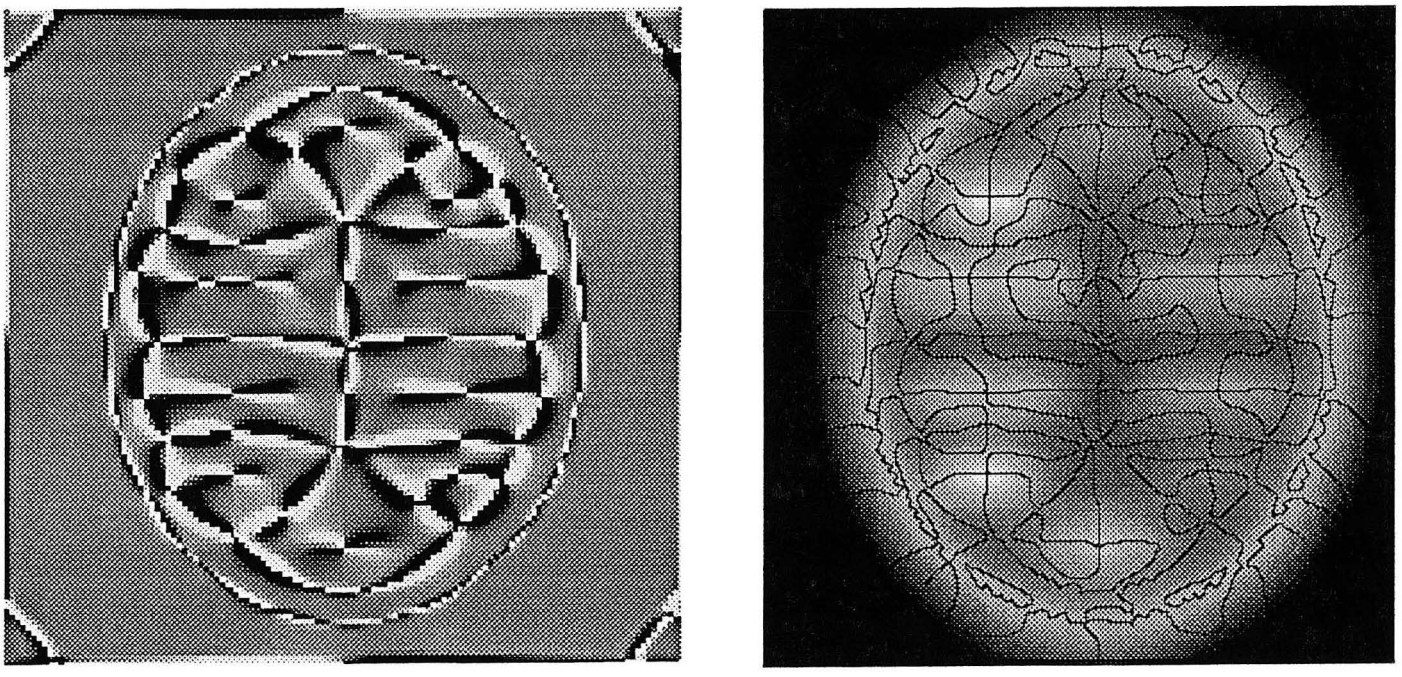


Fig. 7 - The *flow line curvature* operator at a scale σ=3.49 pixels. The left image is the result of applying the operator to the reference phantom at that scale. Zero flow line curvature is middle grey. The right image shows the zeros of the flow line curvature superimposed on the image L at that same scale.

superimposed. Those loci correspond to points in which the flow lines are straight. This is most evident in the centers of the various elliptical or rectangular features, except when disturbed by noise. Above and below, and to the right and left of the those centers, the flow line curvature alternates sign as the flow lines bend outwards from the center of the features.

d) The *relative variation of isophote density along a flow line* $\frac{L_{ww}}{L_w}$. If we follow a flowline in the ascending direction towards the center of one of the smoothed elliptical features of L, for example, the density of isophotes will start growing as we approach an edge, will reach a maximum at the edge and decrease again as we approach the top. The change in that isophote density is described by $\frac{L_{ww}}{L_w}$, with the particularity that the results are relative to the intensity of the feature and insensitive to pedestals, i.e., $\frac{L_{ww}}{L_w}$ is invariant to transformations in the intensity scale of the more restricted type described by (8). This VRF also is sensitive to points in which the denominator approaches zero, as was the case in the two previous VRF's. It is interesting to note that the loci of $\frac{L_{ww}}{L_w} = 0$, or what would be the same $L_{ww} = 0$, are another type of edge finder which may have advantages over L_{ii}, since it is not sensitive to the curvature of the isophotes. Figure 8 shows $\frac{L_{ww}}{L_w}$ (left) and L (right) at $s = 3.49$, with the loci of $\frac{L_{ww}}{L_w} = 0$ superimposed.

3. The set of 5 irreducible invariant VRF's of second order can also be represented as $I \equiv \{L, L_w^2, L_{vv}L_w^2, L_{vw}L_w^2, L_{ww}L_w^2\}$, and we shall describe one of them here.

a) The *cornerness* $L_{vv}L_w^2$: This VRF is the product of isophote curvature and the third power of the gradient strength in gauge coordinates. One can expect that both components are important in corners and, therefore, $L_{vv}L_w^2$ acts as a corner detector. This invariant is homogeneous of degree 3, i.e., it is quite sensitive to illumination. Fig. 9 shows the results of applying $-L_{vv}L_w^2$ to the image L at a scale $s = 1.5$ pixels. It appears that corners are best detected at relatively small scales.

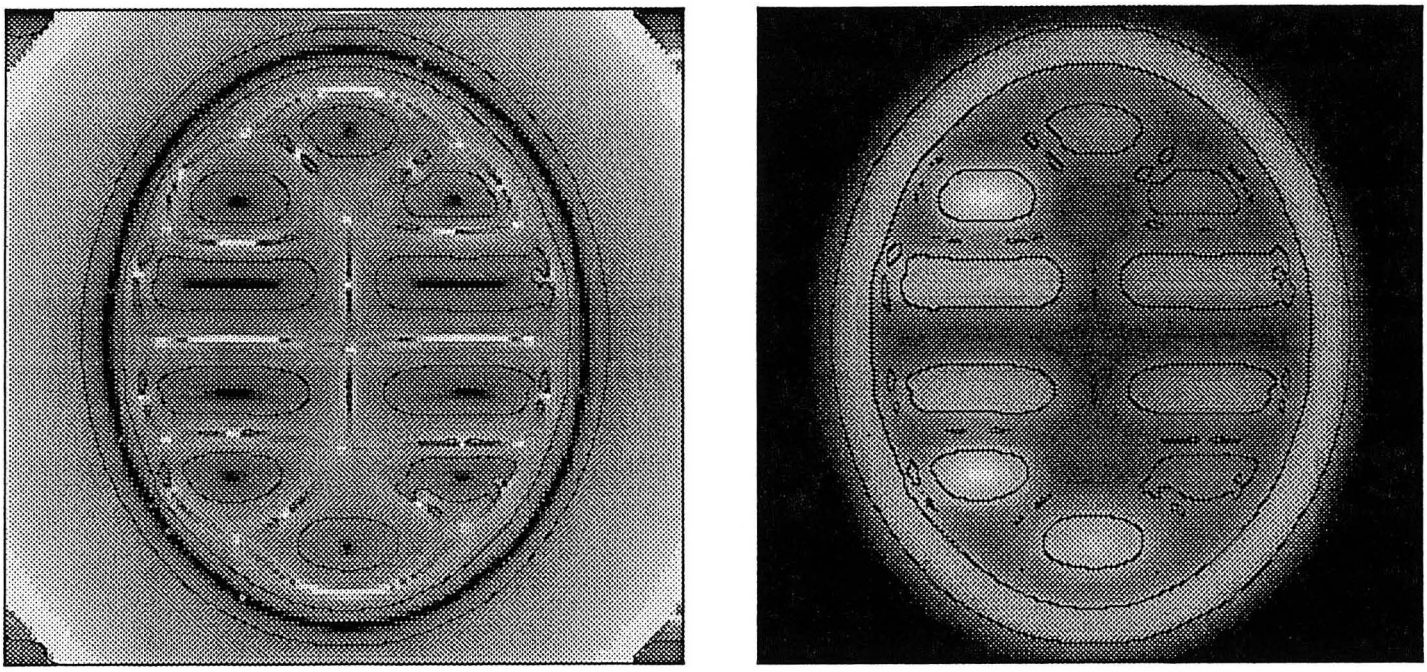


Fig. 8 - The *relative variation of isophote density along a flow line* operator at σ=3.49 pixels. The left image is the result of applying that operator to the reference phantom. Zero variation is middle grey, which has been made prominent by a black line overlay. Right image, the same overlay place on the image L at the same scale.

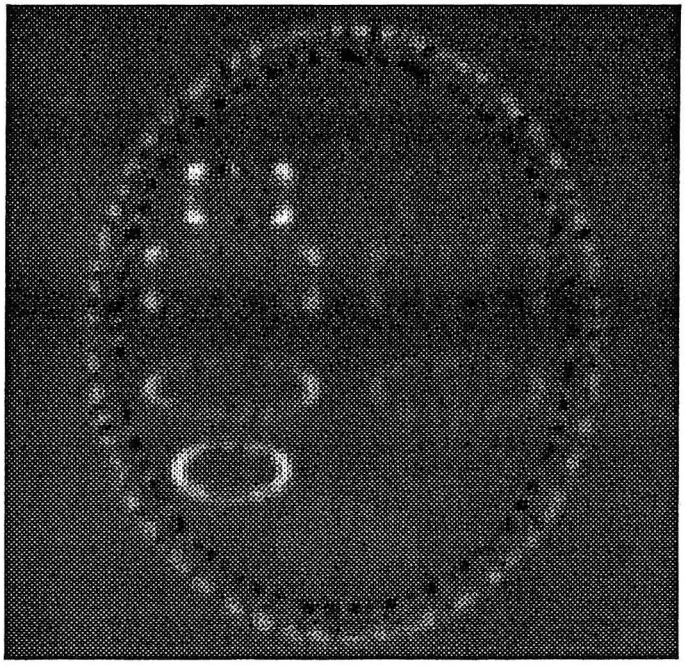


Fig. 9 - The *cornerness* operator at σ=1.5 pixels. Corners are clearly indicated by this operator. It is very sensitive to the intensity of the object.

The last VRF that we will describe here is not one of the irreducible forms and it acts as a patch classifier: the *umbilicity* $U = \frac{L_{ii}^2}{L_{ij}L_{ji}} - 1$. In order to describe this VRF, we start by discussing the concept of curvature at a point P in the surface of an image, which can be defined as the reciprocal of the radius of a circle that fits in the neighborhood of the point considered. The directions of maximum and minimum curvature (the principal curvatures) at a point are always at right angles with respect to each other. If both curvatures are positive, we have a "mountain", if both are negative, we have a "hole". If the principal curvatures are of opposite sign, we have a saddle point. Then, *Gaussian curvature* is the product of the two principal curvatures, so that regions with positive Gaussian curvature are "mountains" or "holes" (elliptical areas), and regions with negative Gaussian curvature are saddle points (hyperbolic areas). Umbilicity, then, gives the product of the Gaussian curvatures. The loci of $U = 0$ separate the elliptical from the hyperbolic areas. Mountains or holes are separated from saddle point areas by parabolic lines. In some simple cases, the parabolic lines can act as edge finders, but it is best to consider those lines as separating patches of different geometric nature.

The (-1) in the definition of U is not relevant to the nature of the VRF. This invariant is homogeneous of order 0, i.e., it is not sensitive to uniform changes in illumination. Fig. 10 (left) shows the umbilicity of L at a scale $s = 2.72$ pixels and the image L (right) at the same scale with the $U = 0$ loci superimposed.

DISCUSSION AND CONCLUSION

The above is a short introduction to the representation of images by Visual Response Functions. The aim has been to show how some specific characteristics or features of an image can be found by a "machine observer". Edges, flat regions, regions with saddle points, corners, etc. can be found by appropriate convolution of the image with the VRF's at a scale such that noise effects are minimized and/or the desired features are emphasized, followed by a selection operation, usually non-linear, like a thresholding. In the particular case of medical images, the zero crossings of the Gaussian Laplacian and of the Umbilicity operators are being investigated as descriptors of prior information for the following specific case: In dynamic studies of the metabolic evolution of specific pharmaceuticals labelled with radioisotopes by Positron Emission Tomography of human organs , data are obtained from a patient over a number of consecutive time periods. The reconstruction of the data set obtained by adding all the time periods is often used by the physician for the purpose of defining one or more Regions-of-Interest (ROI's). Then, the individual data at different time periods are reconstructed and examined successively in order to obtain the time dependence of the uptake or clearance of the radiopharmaceutical in the ROI's. Invariably, the reconstruction of the individual time periods is very noisy, since it contains relatively few photon detection events, resulting in distortion of organ features and high variability in the data. Could the geometrical knowledge obtained from the reconstruction of the sum data help in the reconstructions of the individual time

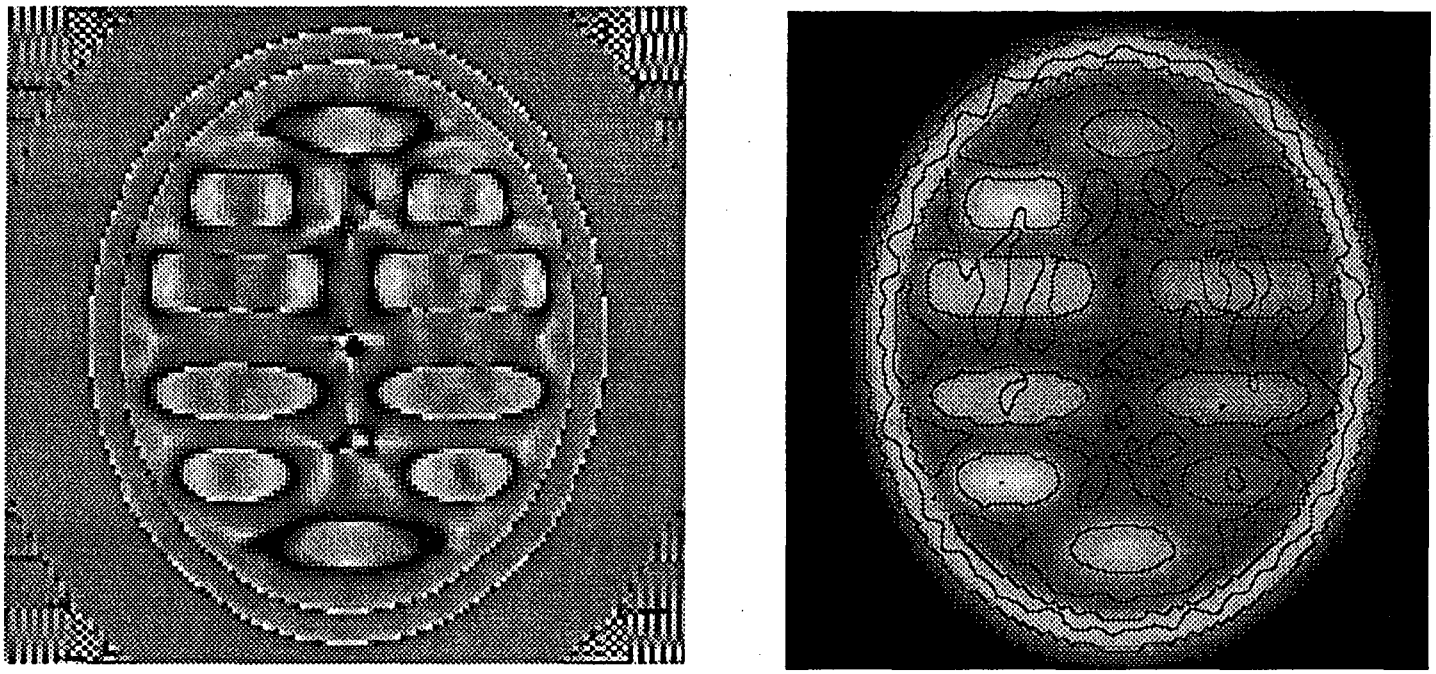


Fig. 10 - The *umbilicity* operator at σ=2.72 pixels. Left image is the result of applying the operator to the reference phantom. The right image shows the zero crossings of the umbilicity superimposed on the reference phantom L at that same scale.

period data? In that particular case, the kind of geometrical knowledge carried by the zero crossings of those two invariants, for example, can be used as prior information in a Bayesian algorithm for the reconstruction. The fact that those VRF's are homogeneous of zero order is of primary importance, because we want the radioactivity measured in a ROI in the individual time period images to be depending only on the corresponding individual tomographic data, although we want the shape information from the sum image to be carried through. First results of this approach have been described in a preliminary paper[14]. We expect the usefulness of the VRF formulation for the segmentation of medical images and for detection of specific shapes to become evident in the future.

ACKNOWDGMENTS

Discussion and help from the students and administrative personnel at the Computer Vision Group, University Hospital Utrecht, is gratefully acknowledged. The work of the first author of this paper has been supported, in part, by the Director, Office of Energy Research, Office of Health and Environmental Research, Physical and Technological Division, of the U.S. Department of Energy under Contract DE-AC03-765SF00098, and by an NFI grant from NWO, the Scientific Research Organization of the Netherlands.

REFERENCES

1. **ter Haar Romeny BM, Florack LMJ**: A multiscale geometric model of human vision. In: Perception of Visual Information, Hendee B, Wells PNT (Eds), Springer-Verlag, Berlin 1992, in press.

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

3. **Witkin AP**: Scale space filtering. *Proc. IJCAI*: 1019-1021, 1983.

4. **Koenderink JJ**: The structure of images. *Biol. Cybern.* 50: 363-370, 1984.

5. **Babaud J, Witkin A, Duda R**: Uniqueness of the gaussian kernel for scale-space filtering. *IEEE PAMI* 8, X, XXXX-XXXX, 1986.

6. **Korn A**: Toward a symbolic representation of intensity changes in images. *IEEE PAMI* 10, 5: 610-625, 1988.

7. **Perona P, Malik J**: Scale space and edge detection using anisotropic diffusion. *IEEE PAMI* 12, 7: 629-639, 1990.

8. **Hubel DH**: *Eye, Brain and Vision.*, Scientific Americal Library Series #22, W.H. Freeman and Comp., New York, 1988.

9. **Blom J, ter Haar Romeny BN, Bel A, Koenderink JJ**: Spatial derivatives and the propagation of noise in Gaussian scale space. *J. Vis. Comm. Im. Repr.*, in press.

10. **Florack LMJ, ter Haar Romeny BM, Koenderink JJ, Viergever, MA**: General intensity transformations and second order invariants. In: Scandinavian Conference on Image Analysis, Johansen P and Olsen S (Eds), Pattern Recognition Soc. of Denmark, Copenhagen, pp 346-353, 1991.

11. **Marr D, Hildreth EC:** Theory of edge detection. *Proc. R. Society London* 200: 269-294, 1980.

12. **Pizer, SM...** (need reference for paper discussing medial axis obtained from maxima of the Lii operator as a function of scale)

13. **de Graaf CN, Vincken KL, Viergever MA, Koenderink JJ, Appelman FJR, et al**: A hyperstack for the segmentation of 3D images. In: Information Processing in Medical Imaging, Ortendahl DA and Llacer J (Eds): *Progress in Clinical and Biological Research, 363* , Wiley-Liss, New York, pp 399-413, 1991.

14. **Llacer J, ter Haar Romeny BM, Viergever MA:** The use of geometric prior information in Bayesian tomographic image reconstruction: a preliminary report. To be published by SPIE, proc. of conf. on Mathematic Methods in Medical Imaging, 1992.

15. **Canny J:** A computational approach to edge detection. *IEEE PAMI* 8, 6: 679-698, 1987.

16. **De Micheli E, Caprile B, Ottonello P, Torre V**: Localization and noise in edge detection. *IEEE PAMI* 10, 11: 1106-1117, 1989.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720