Chapter 4

Geometry-limited diffusion

The goal is a geometry-based image segmentation, an image segmentation which relies on geometric descriptions of local image structure. A geometric description is a set of measurements at every point in the image space (be it a discrete or continuous space) that *encodes* the local shape of the intensity surface. The multiscale diffusion process forms regions of homogeneous intensity in a somewhat robust manner but does not account for higher-order properties. This chapter combines the concepts of anisotropic diffusion and local shape in order to define a process that forms patches which are locally homogeneous in higher-order information.

4.1 Multi-valued diffusion

4.1.1 Multi-valued images

A multi-valued image is a smooth mapping from the image domain to a m-dimensional range,

$$\vec{F}(\vec{x}) : I \mapsto D,$$

where $D \subset \mathbb{R}^m$ is also called the feature space. Each point in the domain is characterized by its position in the feature space, and this position is recorded at each point in I as a finite array of measurements, $\vec{F}(\vec{x}) = F_1(\vec{x}), \ldots, F_m(\vec{x})$. A multi-valued image is thus a *set* or a *stack* of single-valued images that share the same image space. Note that the set of numbers used to record the position of a pixel in the range is not indicative of the shape or nature of that space. More specifically, Euclidean distance is not necessarily the appropriate measure for comparing positions in this feature space.

Perhaps the most apparent examples of multi-valued images are the data that result from imaging devices that measure several physical quantities at the same location in space. Multi-echo magnetic resonance imaging data (MRI), multispectral LANDSAT data, and the red-green-blue (RGB) channels from a color camera are examples. One could treat each physical measurement in these data as a separate image and perform edge-affected diffusion process separately for each measurement. However, if the goal of the edge-affected diffusion process is to capture dissimilarities between pixels across a number of these measurements, then the regularization process should reflect this.

4.1.2 Diffusion with multiple features

The diffusion process introduces a time or evolution parameter $t \in T \subseteq \mathbb{R}^+$ into the function $\vec{F}$, so that

$$\vec{F}(\vec{x}, t) : I \times T \mapsto D,$$

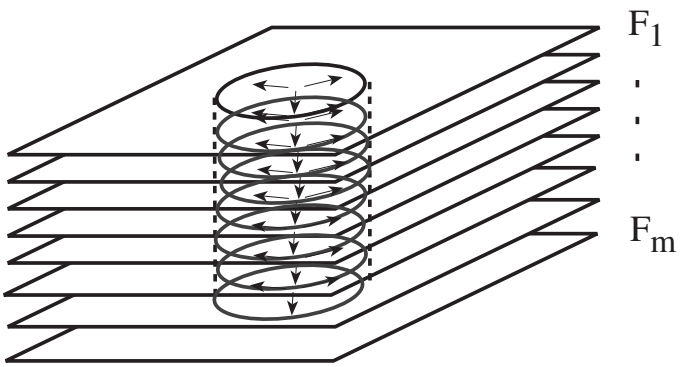

Figure 4.1: Multi-valued diffusion is a system of diffusion processes that are coupled through the conductance function.

and there is a multi-valued function at each point in time or each level of processing.

The multi-valued diffusion equation is

$$\nabla \cdot g(\mathcal{D}_s \vec{F}) \nabla \vec{F} = \frac{\partial \vec{F}}{\partial t} \quad , \tag{4.1}$$

where $\mathcal{D}$ is the *dissimilarity* operator, a generalized form of the gradient magnitude which is defined in the following section. The composition, $\mathcal{D}_s \vec{F} : I \mapsto \mathbb{R}^+$, assigns a degree of dissimilarity to every point in the image space. The s indicates the isotropic scale used to make the dissimilarity measurements so that $\mathcal{D}_s \vec{F} \stackrel{\text{def}}{=} \mathcal{D}(G(\vec{x}, s) \otimes \vec{F}(\vec{x}, t))$. The scale s decreases over time as in the multiscale diffusion equation. The derivative $\nabla \vec{F}$ takes the form of a matrix,

$$\nabla \vec{F} = \begin{matrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{matrix} ,$$

which may be called the generalized Jacobian matrix of $\vec{F}$. The conductance, $g : \mathbb{R} \mapsto \mathbb{R}^+$, is a scalar function, and the operator "$\nabla \cdot$" is a vector that is applied to the matrix $g(\mathcal{D}_s \vec{F}) \nabla \vec{F}$ using the usual convention of matrix multiplication.

$$\nabla \cdot g \nabla \vec{F} = \begin{pmatrix} \frac{\partial}{\partial x_1} & \cdots & \frac{\partial}{\partial x_n} \end{pmatrix} \cdot g \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial F_1}{\partial x_n} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}$$

Equation (4.1) is a system of separate single-valued diffusion processes, evolving simultaneously and sharing a common conductance modulating term.

$$\begin{aligned} \nabla \cdot g(\mathcal{D}_s \vec{F}) \nabla F_1 &= \frac{\partial F_1}{\partial t} \\ &\vdots \\ \nabla \cdot g(\mathcal{D}_s \vec{F}) \nabla F_m &= \frac{\partial F_m}{\partial t} \end{aligned}$$

The boundaries are not defined on any one image but are shared among (and possibly dependent on) all of the images in the system as in Fig. 4.1.

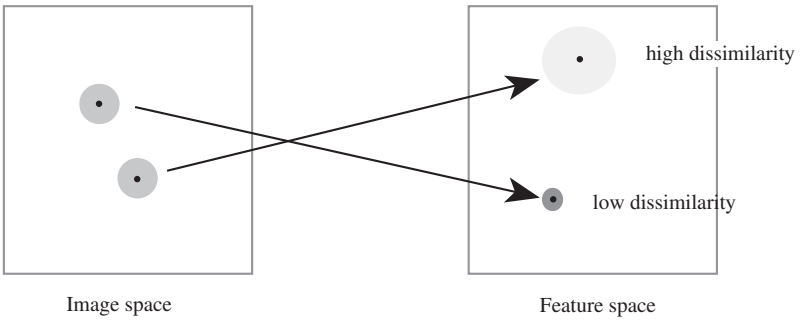

Figure 4.2: Every neighborhood in the image space maps onto a neighborhood in the feature space. The "size" of the neighborhood in the feature space indicates dissimilarity.

4.1.3 Diffusion in a feature space

The behavior of the system (4.1) is clearly dependent on the choice of the dissimilarity operator $\mathcal{D}$. In the single feature case, $m = 1$, the gradient magnitude proves to be a useful measure, that is

$$\mathcal{D}f \stackrel{\text{def}}{=} (\nabla f \cdot \nabla f)^{\frac{1}{2}} .$$

For higher dimensions the dissimilarity is evaluated on the basis distances in D, the feature space.

Applications of statistical pattern recognition to image segmentation are available in the literature (Duda and Hart 1973, Jain and Dubes 1988, Coggins and Jain 1985). Such approaches apply multivariate statistics in order to find clusters of pixels in these feature spaces. Boundaries between clusters can be defined in a way that optimizes certain metrics, such as the Hotelling trace. Pixels are classified on the basis of their distance to groups of pixels that are nearby in the feature space. Effective use of such techniques requires an appropriate measure of distance. Typically, linear transformations are performed on the feature space in order to produce meaningful clusters. These transformations can be determined by statistical measures made on the available data.

The dissimilarity operator is constructed to capture the manner in which neighborhoods in the image space (n-dimensional) map into the feature space. If $\vec{F}$ is a well behaved function (Lipschitz continuous for example), then for a point $x_0 \in I$ and a neighborhood $\mathcal{N}^{\epsilon}_{x_0}$, of size ϵ, there is a point $y_0 = \vec{F}(x_0) \in D$ and a corresponding neighborhood which contains the mapping of $\mathcal{N}^{\epsilon}_{x_0}$: $\mathcal{N}_{y_0} \supseteq \vec{F}(\mathcal{N}^{\epsilon}_{x_0})$. The dissimilarity at x_0 is a measure of the relative *size* of these neighborhoods. If $\mathcal{N}_{y_0}$ is relatively small for a given ϵ, this indicates that the neighborhood of x_0 has low dissimilarity as shown in Fig. 4.2. In the limit as $\epsilon \to 0$, the relative size of these neighborhoods is described by the first-order derivatives of $\vec{F}$. The matrix of first-order derivatives is the Jacobian matrix, $J = \nabla \vec{F}$.

I propose a dissimilarity that is the Euclidean (root sum of squares) norm of the Jacobian, also called the Frobenius norm. If J is the Jacobian of $\vec{F}$, with elements J_{ij}, then the dissimilarity is

$$\mathcal{D}\vec{F} \stackrel{\text{def}}{=} \|J\| = \left(\sum_{j=1}^{n} \sum_{i=1}^{m} {J_{ij}}^2 \right)^{\frac{1}{2}} . \tag{4.2}$$

I choose this norm for several reasons. First, this matrix norm is induced from the Euclidean vector norm so that in the case of $m = 1$, this expression is the gradient magnitude, as in anisotropic diffusion. Second, the square of this norm (as it often appears in the conductance functions (Nordstrom 1990)) is differentiable. Finally, when this norm is applied to geometric objects (tensors), it produces a scalar which is invariant to orthogonal coordinate transformations (see Sect. 4.5.1).

This approach has several advantages over other statistical pattern recognition approaches. First, because the feature space is used in conjunction with the diffusion equation, the process combines information about the image space with position in the feature space. As the process, evolves pixels that are nearby in image space are drawn together in the feature space, except in cases where the dissimilarity is large enough to reduce the conductance. The result is a clustering that is sensitive to the spatial cohesiveness of the feature space. Several authors (Gerig, Kubler, Kikinis and Jolesz 1991, Yoo and Coggins 1993) have shown experimentally that a pair of coupled nonuniform diffusion equations applied to dual-echo MRI can improve the distinction between clusters of pixels in the two-dimensional feature space. Multi-valued diffusion as described here constructs a generalized framework for such systems of coupled diffusion equations.

The second advantage is that the dissimilarity depends only on the differential structure of $\vec{F}$ and therefore does not rely on the global structure of the feature space. The dependence on the local structure of the feature space allows for a great deal of flexibility for transformations within that space.

The dissimilarity measure can be generalized to allow for coordinate transformations that are *local* in the feature space. Include an arbitrary (differentiable) transformation on $\vec{F}$ that is used in the process of computing dissimilarity. This is consistent with the idea that the relative importance of the features $F_1, \ldots, F_m$ might not be reflected in their relative magnitudes. Allow a transformation to a new set of features $\vec{G}(\vec{w})$ where $G : \mathbb{R}^m \mapsto \mathbb{R}^k$. Then the dissimilarity of $\vec{F}$ which accounts for this mapping $\vec{G}$ is

$$\mathcal{D}\vec{F} \stackrel{\text{def}}{=} \|\nabla \vec{G}(\vec{F})\| = \|\frac{\partial}{\partial \vec{w}} \vec{G}(\vec{F}) \nabla \vec{F}\| = \|\phi(\vec{F}) \nabla \vec{F}\| \tag{4.3}$$

where ϕ is an $m \times k$ matrix that indicates a local linear transformation in the feature space. The derivative with respect to feature-space coordinates takes the form of a vector:

$$\frac{\partial}{\partial \vec{w}} \stackrel{\text{def}}{=} \frac{\partial}{\partial w_1} \cdot \frac{\partial}{\partial w_m},$$

and I reserve the use of ∇ for derivatives in the image space. If one considers coordinate transformations (rotations and rescaling of axis) in feature space to represent changes in the relative importance of features, then the local coordinate transformation $\phi(\vec{y})$, for $\vec{y} \in \mathbb{R}^m$, allows the relative importance of various features to vary depending on the position in the feature space.

The practical implications of this local transformation can be imagined by considering a multi-valued data set gathered by some imaging device. Suppose that the relative importance of each feature depends not only on its own value but the values of the other features. Such behavior could be accounted for by choosing the proper $\phi(\vec{y})$. At this point it is still not evident how one would go about choosing $\vec{F}$ given set of image data. It is conceivable that $\phi(\vec{y})$ should depend on the statistical properties (local and global) of the feature space, the nature of the imaging device, and the task at hand.

For feature spaces that consist of geometric features, this metric is essential for comparing incommensurate quantities associated with axes of these spaces. It is also necessary in order to construct dissimilarity measures that are invariant to one's initial choice of spatial coordinates.

4.2 Describing local geometry

The local shape of the intensity surface at every point in the image is represented through a finite set of scalar values. These values constitute a set of geometric descriptors and an associated feature space. There are a variety of possibilities for such representations, but for this work I choose sets of derivatives of the intensity surface measured in Cartesian coordinates.

If $f(x, y)$ is the original image ($n = 2$), the local surface at each point in the image can be

represented by the local Taylor series expansion:

$$h(x', y', x, y) = f\,|_{x,y} + f_{x'}\,|_{x,y}\,x' + f_{y'}\,|_{x,y}\,y' + \frac{1}{2} f_{x'x'}\,|_{x,y}\,x'^2 + \cdots \tag{4.4}$$

The x' and y' are local coordinates that exist at each choice of x and y. The function, $g(x', y', x, y)$, is the same function at every x and y, except that it is expressed in local coordinates. Thus $g(0, 0, x, y) = f(x, y)$ as can be seen in Eq. (4.4). In practice, the series can be truncated in order to produce a local approximation to the surface and to produce a finite number of terms. Truncated versions of this function are no longer the same at each point in the image space—at each point there is a distinct local approximation to the surface structure. The coefficients of the truncated expansion can be considered a multi-valued function of the original image space.

$$\vec{h}(x, y) = (f\,|_{x,y}, f_{x'}\,|_{x,y}, f_{y'}\,|_{x,y}, \frac{1}{2} f_{x'x'}\,|_{x,y}, \ldots)$$

If the local coordinates x' and y' are aligned with the image coordinates, then $\vec{F}(x, y)$ consists of partial derivatives (to within a constant factor) of $f(x, y)$.

$$\vec{F}(x, y) = (f(x, y), f_x(x, y), f_y(x, y), f_{xx}(x, y), \ldots)$$

This expression for local shape has several useful properties. First, derivatives of the intensity surface can be measured directly from a digital image using linear filters as described in Sect. 2.5.1. Second, in analyzing the behavior of this system it is useful to be able to rely on the linearity of these operators. Finally, the use of complete sets of derivatives (all of the derivatives of a particular order treated as a tensor) provides means of constructing processes which are invariant to orthogonal group transformations. A set of derivatives on the original image create a multi-valued function $\vec{F} : I \mapsto D$, where $D \subset \mathbb{R}^m$ is the feature space.

4.3 Diffusion in the jet space

The differential measurements that comprise the Taylor series description of local surface shape form feature vectors that become the initial conditions of a multi-valued diffusion equation. In the feature space of geometric features, each point corresponds to a different shape, and the metric ϕ quantifies these differences in shape; it is precisely the *shape metric* as discussed in Sect. 2.4.2. The strategy is to choose ϕ so that it creates boundaries that are of interest and choose the scale, $s(t)$, in the diffusion equation so that unwanted fluctuations in the feature measurements (noise) are eliminated.

As the anisotropic scale increases, measurements are taken from the resulting set of features with progressively smaller Gaussian kernels, and decisions about presence of specific geometric properties can be made on the basis of those measurements. This strategy is depicted in Fig. 4.3. Information about $(N + 1)$th-order derivatives can be obtained in three steps. First, make differential measurements of the image up to Nth-order using derivative-of-Gaussian filters of a scale that does not compromise interesting image structures. Second, smooth these derivatives in a way that breaks these "images" into patches that start out regions of local homogeneity and become nearly piecewise constant as the process evolves. Third, study the boundaries of these patches by looking at the derivatives of the features and make decisions about the presence of higher-order features which include $(N + 1)$th-order derivatives.

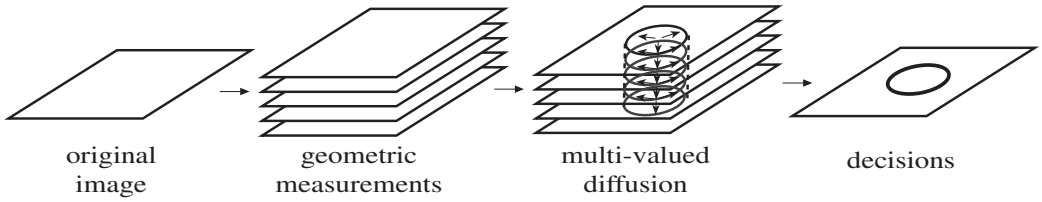

Figure 4.3: Geometry-limited diffusion in which features are geometric measurements made on a single valued image.

4.4 First-order geometry and creases

4.4.1 Creases

This approach has already been studied in the zeroth-order case (anisotropic diffusion based on gradient magnitude), and a logical next step is to include first-order derivatives as features. For a two-dimensional image space the resulting feature space is three-dimensional. The remainder of this section investigates the use of first-order geometric features for the purpose of finding "creases" or "ridges".

Pizer and others (Pizer, Gauch, Cullip and Fredericksen 1990, Lifshitz and Pizer 1990) have shown that creases and the corresponding flank regions are very powerful for forming hierarchical segmentations of medical images. Although the notion of creases has no one precise definition (Eberly, Gardner, Morse, Pizer and Scharlach 1993, Rosin et al. 1991, Koenderink and van Doorn 1993), there is an intuition that one can test via a simple experiment. Imagine an image as an intensity surface, and suppose that the intensity axis is aligned with a gravitation field so that the direction of increasing intensity is "up". Place a drop of water on the surface. Creases are places where a drop of water tends to split and run in two or more different directions. Likewise, one could turn the surface over and repeat the experiment. The results in the first case are "ridges" and in the second case "valleys". If the surface is continuous, then every place on the image, except local extrema, has exactly one downhill direction. However, the second order structure (a larger drop of water) describes, among other things, the *degree of variation* in gradient directions in the neighborhood of a point. Creases are places that have a large disparity of gradient directions in an immediate neighborhood. If one characterizes flanks as contiguous regions that have the same (or nearly the same) uphill direction, then creases are the places where flank regions meet.

This definition of a crease is independent of image intensity; the height of the intensity surface does not enter into this definition. Therefore, it is appropriate and convenient for this discussion to exclude the zeroth-order term as one of the features. Including it would have no effect on the results. For a two-dimensional image space encode the local image geometry as two features F^x and F^y, the pair of first partial derivatives of intensity. The raised indices denote the fact that these features are undergoing nonlinear diffusion and are no longer the derivatives of the initial image but have the image derivatives as their initial conditions;

$$F_1(x,y,t) = F^x(x,y,t), \text{ and } F^x(x,y,0) = \frac{\partial f(x,y)}{\partial x},$$

while

$$F_2(x,y,t) = F^y(x,y,t), \text{ and } F^y(x,y,0) = \frac{\partial f(x,y)}{\partial y}.$$

Thus, the raised indices are merely labels, and these labels indicate the individual features associated with a vector of such features. Subscript indices, however, are used to indicate differentiation in the direction of the axis associated with that index. For instance, $f_x(x,y) \stackrel{\text{def}}{=} \frac{\partial f(x,y)}{\partial x}$ and

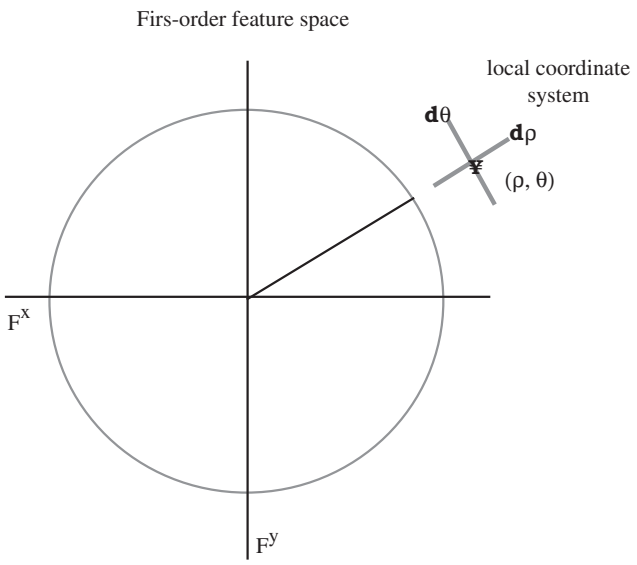

Figure 4.4: First-order feature space as represented in polar coordinates. Every point in this feature space has a local coordinate system that is aligned with the dθ and dρ directions.

$F^x(x,y,0) \stackrel{\text{def}}{=} f_x(x,y)$. These two measurements comprise a two-dimensional feature space. Effectively, this choice of features models each local surface patch as a plane of a particular orientation.

The features, F^x and F^y, are derivatives of $f(x,y)$ only at $t=0$. For $t>0$ it might be impossible to construct $f(x,y,t)$ such that $(\partial f(x,y,t)/\partial x) = F^x(x,y,t)$ and $(\partial f(x,y,t)/\partial y) = F^y(x,y,t)$. Derivatives do not commute with the nonlinear diffusion operator, so, in general, the relationship $F^x_y = F^y_x$ does not hold, and there is no such $f(x,y,t)$. However, the reconstruction of a diffused luminance function, $f(x,y,t)$ is not the goal of this type of processing—the goal is to characterize changes in first-order information—so the lack of $f(x,y,t)$ is of little consequence. This reasoning brings up a very important question about the *geometry* that results from a nonlinear diffusion of first derivatives. Does such geometry make sense? Are differential invariants constructed from $F^x(x,y,0)$ and $F^y(x,y,0)$ invariant for $t>0$? These questions will be addressed in Sect. 4.5.1.

4.4.2 A shape metric for creases

Choose the coordinate transformation ϕ, which determines the way in which distances are measured (locally) in the feature space, specifically to capture the notion of creases. The appropriate choice of ϕ is best understood by considering the feature space; $F^x,\ F^y$; as it is described by polar coordinates (Fig. 4.4). Each position $(x,\ y)$ in the image has a position, $(v,\ w) = \vec{F}(x,y)$, in this feature space. The polar coordinates of each point in this feature space have a geometric interpretation. The magnitude of the gradient vector ($|\nabla f|$) at the point $(x,\ y)$ is $\rho = (v^2+w^2)^{\frac{1}{2}}$, and $\theta = \arctan(w/v)$ is the direction of image gradient. Now define a dissimilarity operator, ϕ, which captures changes only in the direction of the vector, $\vec{F}$. I express ϕ as a rotation R followed by a scaling S. The matrix R is a rotation of the local feature space coordinates into a coordinate system which is aligned with dρ and dθ in the polar coordinates. The matrix S controls the scaling and thereby the relative effects of each of these properties. This allows for control of the amount of influence that the gradient magnitude and gradient direction have on the dissimilarity operator. The notion of creases does not depend on the strength of the gradient magnitude, only on direction. Therefore, differences between pixels in the radial direction does not contribute to the dissimilarity. This has the effect of locally collapsing the feature space onto a circle with its center at the origin. The perpendicular

component, $\mathrm{d}\theta$, is scaled by the inverse of the radius so that small changes in the in the $\mathrm{d}\theta$ direction capture changes in the angle. Use the local metric

$$\begin{aligned}\phi &= R^{-1}SR \\ &= \begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix}\begin{bmatrix}\frac{1}{\rho} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{bmatrix} \\ &= \begin{bmatrix}\frac{\cos^2(\theta)}{\rho} & \frac{-\sin(\theta)\cos(\theta)}{\rho}\\ \frac{-\sin(\theta)\cos(\theta)}{\rho} & \frac{\sin^2(\theta)}{\rho}\end{bmatrix} = \frac{1}{((F^x)^2+(F^y)^2)^{\frac{3}{2}}}\begin{bmatrix}(F^x)^2 & F^xF^y\\ -F^xF^y & (F^y)^2\end{bmatrix}\end{aligned} \tag{4.5}$$

When this transformation is used to compute dissimilarity $\mathcal{D}\vec{F} = \|\phi(\vec{F})\nabla\vec{F}\|$, it has the same effect (in continuous images, but necessarily in discrete approximations—see Appendix A) as computing the Jacobian on a pair of intermediate features $\vec{G} = G_1, G_2$ which are normalized with respect to the Euclidean length of the vector $\vec{F}$:

$$G_1 = \frac{F_1}{|\vec{F}|} = \frac{F^x}{(F^{x2}+F^{y2})^{\frac{1}{2}}}, \qquad G_2 = \frac{F_2}{|\vec{F}|} = \frac{F^y}{(F^{x2}+F^{y2})^{\frac{1}{2}}} \text{ and} \tag{4.6}$$

$$\mathcal{D}\vec{F} \quad \stackrel{\text{def}}{=} \quad \|\nabla\vec{G}\|$$

This normalization makes explicit the fact that this distance measure is invariant to any monotonic intensity transformation although the pair of features, F^x and F^y, are clearly not.

4.4.3 Results

Experiments (Whitaker 1993) show this first-order geometry-limited diffusion process forms sets of homogeneous regions that have similar gradient directions. These regions correspond to "flanks" or "hillsides" in the intensity surface of the original image.

The test image in Fig. 4.5a (hereafter the *blob* image) is created by drawing a white figure on a black background and then blurring the result. Uniformly distributed noise is added to the image so that the range of the noise at each pixel is half the overall intensity of the foreground (50%), where the image is 256 × 256 pixels in size. Figures 4.5b–d show the initial values of the features, F^x and F^y, and the dissimilarity measure at start of the diffusion process. Because the dissimilarity measure normalizes features with respect to the gradient magnitude, areas of the image that were initially flat are susceptible to noise. Figures 4.6 shows the features and the dissimilarity measure after processing. The areas of high dissimilarity indicate boundaries between flank regions. The pixels within flank regions have been "regularized" so that they have virtually the same gradient values.

The boundaries of these regions appear to correspond well with the intensity ridges in the original image. Because of the tendency of this process to produce piecewise constant solutions, the boundaries are distinct and allow for easy detection with one of the techniques already mentioned in Sect. 2.4.2. The results of this process appear to be robust with respect to noise. The use of scale in measuring dissimilarity controls the size of these regions or patches without distorting the shapes of the boundaries between patches.

Figure 4.7 shows the MRI image from Fig. 1.1, the initial values of the first-order features, and the initial value of the dissimilarity measure. Figure 4.8a–b shows the pair of first-order feature measures after processing at a scale of 3.0 (measured in units of pixel separation). The features F^x and F^y form homogeneous patches that seem to correspond to flank-sides in the original image. Increasing the scale of the process increases the size of the patches that remain after the process is complete. Figure 4.8c shows that the dissimilarity has formed distinct boundaries for these patches. Figure 4.8d shows another second-order measure that reflects the sign of the "curvature" associated

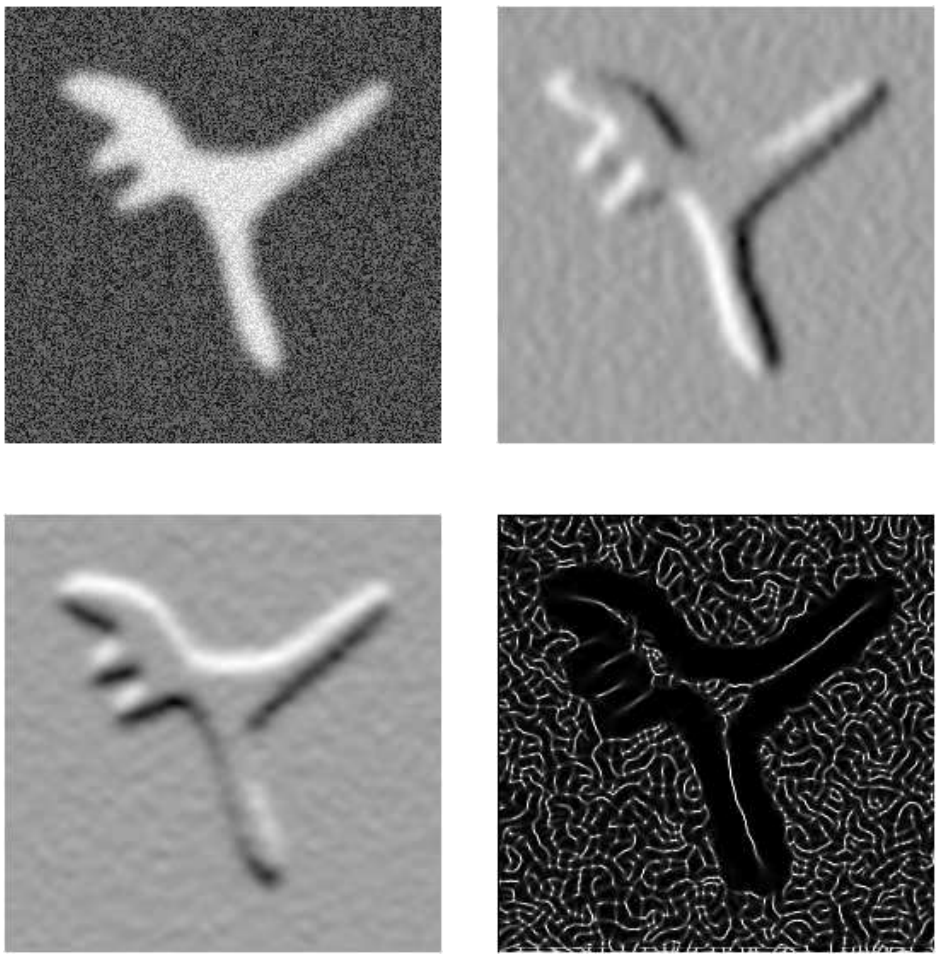

Figure 4.5: From the top left to right: (a) A noisy blob is created by blurring a white on black figure and then adding uniformly distributed random noise. Shown here are the initial values of the first-order features (b) F^x and (c) F^y and (d) the dissimilarity measure on these features.

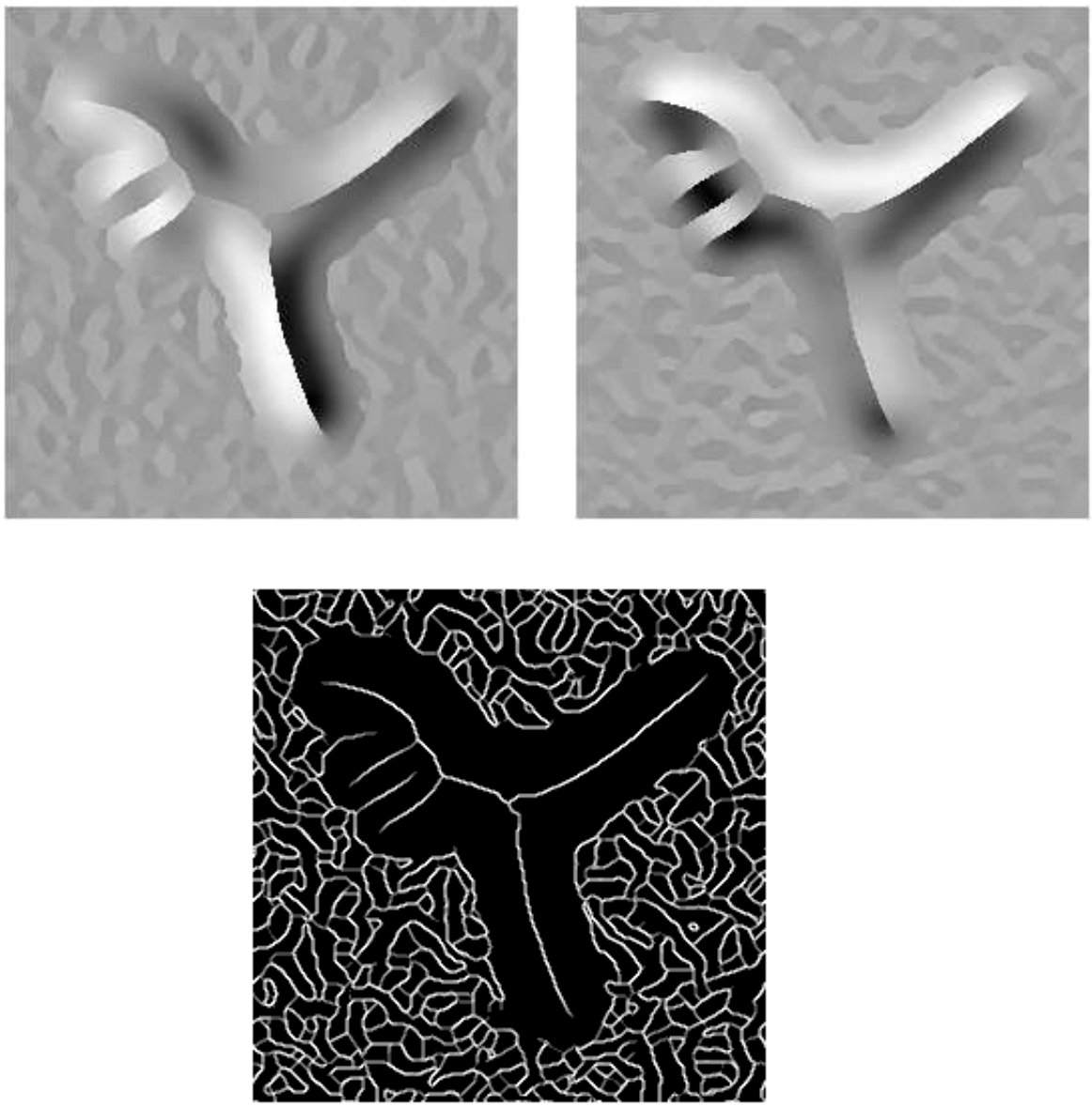

Figure 4.6: From the top left to right: The result values of the first-order features (a) F^x and (b) F^y after the geometry-limited diffusion process and (c) the dissimilarity measure $\mathcal{D}\vec{F}$ on these features.

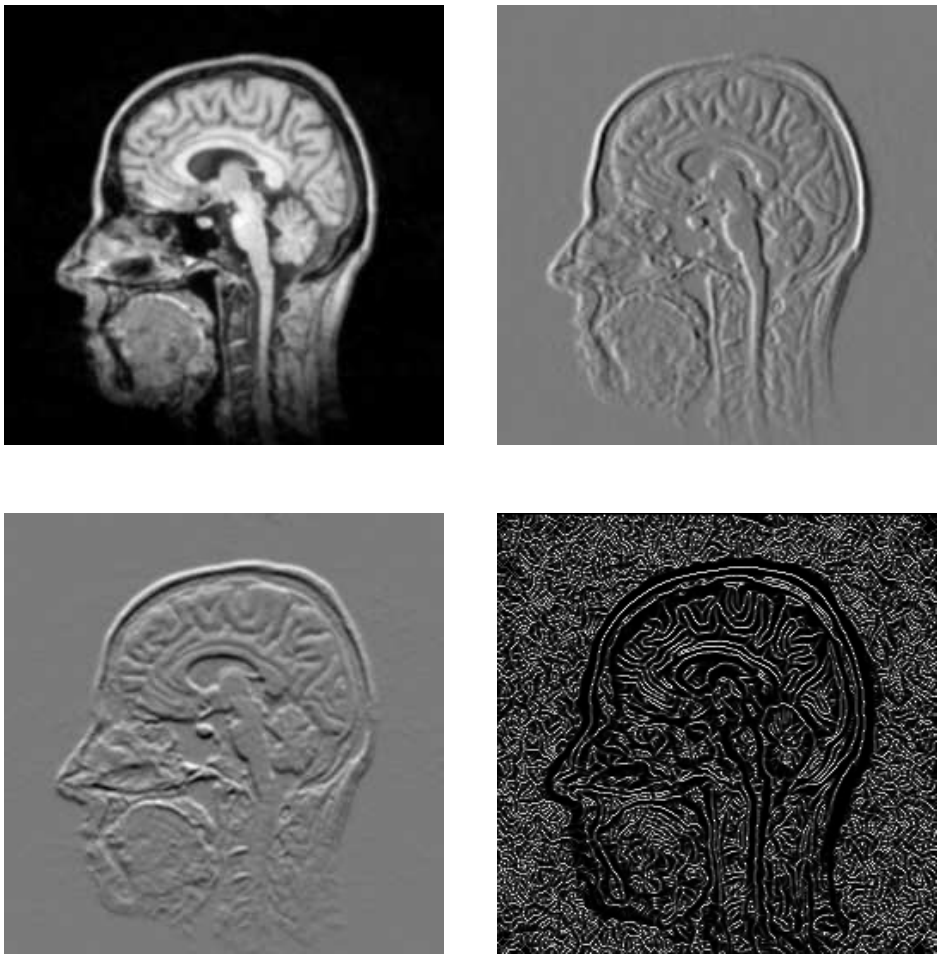

Figure 4.7: From the top left to right: (a) A slice of an MRI image of a head, (b)–(c) the first partial derivatives in the x and y directions, and the d) dissimilarity measure computed in the initial data.

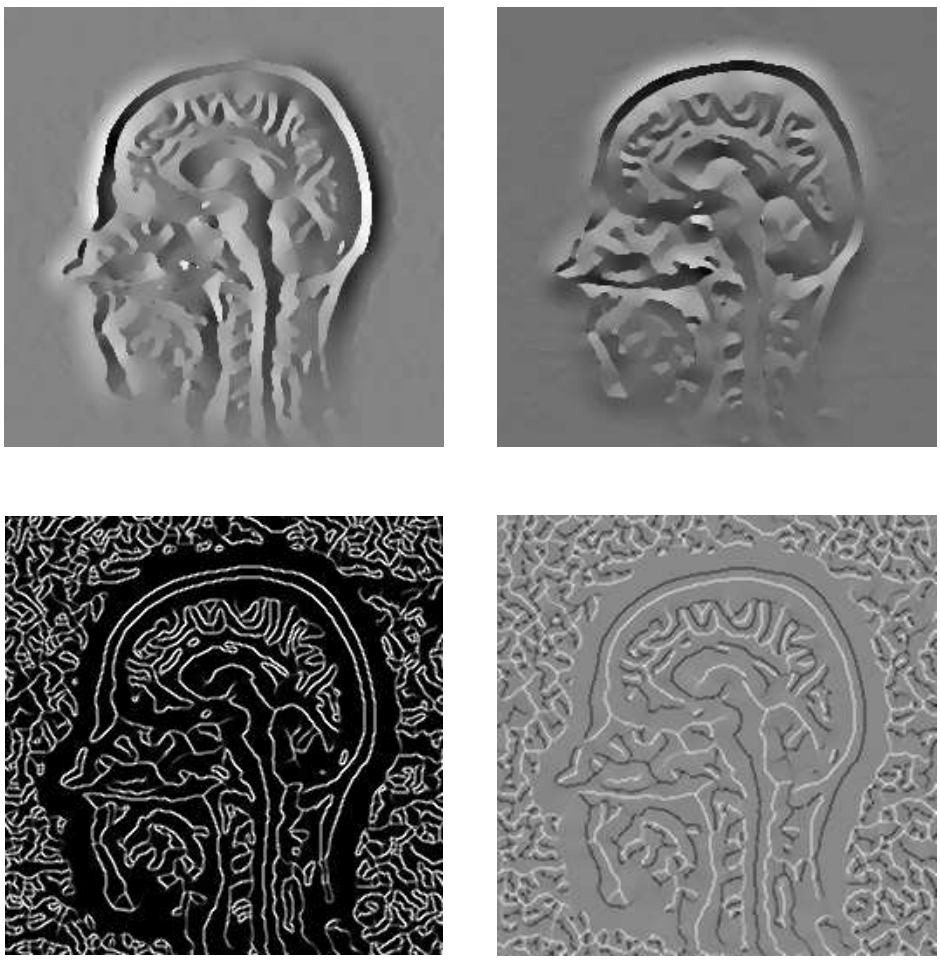

Figure 4.8: From the top left to right: (a)–(b) The first partial derivatives in the x and y directions after processing at a scale of 3.0, (c) the dissimilarity measure after processing, and (d) the curvature measure from (4.7).

with each boundary. This is analogous to the isophote curvature of a scalar function $f(\vec{x},t)$ except that it is constructed from *diffused* derivatives. For G_1 and G_2 defined above, this measure is

$$\frac{\partial G_1}{\partial x} + \frac{\partial G_2}{\partial y}, \tag{4.7}$$

and the sign of this measure distinguishes "ridges" (dark) from "valleys" (light). The first-order diffusion is essentially a piecewise linear approximation to the original image which ignores the height differences between patches (the "dc" component). The sign of the curvature at boundaries indicates the manner in which these linear patches meet. If the boundary of two patches "sticks up", the curvature is negative, and the boundary is a ridge. If the two patches slope down toward the place where they meet, the boundary "sticks down", the curvature is positive, and the boundary is a valley. The fact that $\vec{G}$ is normalized means that the steepness of the patches does not enter into the value of (4.7); this measure captures only the isophote curvature.

4.4.4 The diffused Hessian

The second-order information derived from first-order patches provides a number of interesting measurements. Consider the *diffused Hessian*,

$$H^{(1)} = \begin{bmatrix} F^x{}_x & F^x{}_y \\ F^y{}_x & F^y{}_y \end{bmatrix}, \tag{4.8}$$

where the superscript (1) indicates that this measurement results from a first-order diffusion process. Unlike the conventional Hessian, this matrix is not necessarily symmetric. The trace of this matrix, $\mathrm{Tr}[H^{(1)}] = F^x{}_x + F^y{}_y$, is a kind of Laplacian, and $\mathrm{Tr}[H^{(1)}] = 2K^{(1)}$, where $K^{(1)}$ might be called the *mean curvature*. In the context of this first-order diffusion process it describes the curvature at patch boundaries. Unlike the isophote curvature it includes information about the steepness of the patches. This measure gives sign information, which indicates whether creases are ridges or valleys, and its magnitude reflects the angle (in two dimensions) at which two linear patches meet. The Euclidean norm of this matrix $\|H^{(1)}\|$ is sometimes called *deviation from flatness*, and it indicates the overall change in first-order information; magnitude and direction. These two measures, the mean curvature and the Euclidean norm of H, can be used to sort creases on the basis of their "sharpness" and thereby exclude, if desired, creases that result from noise in otherwise very flat regions. Fig. 4.9 shows these two measures computed with the features of Fig. 4.8. Some creases respond more strongly than others based on the magnitude of the gradients within patches. Unlike the dissimilarity these measurements are sensitive to monotonic intensity transformations.

4.4.5 Combining information of zeroth and first-orders—corners

In the examples of Sect. 4.4.3, intensity is not used as one of the features. There may exist a three-dimensional feature space, a corresponding multi-valued function $\vec{F} = F^x, F^y, F$, and an associated shape metric $\phi(F^x, F^y, F)$ which together would allow a segmentation to produce visually interesting regions and boundaries. This section explores an alternative method for combining zeroth and first-order information for the purpose of detecting corners in images.

Consider the visual feature called a "corner". The difficulty of a finding corners using geometry-limited diffusion is that corners in two dimensions are not contours or lines, but points. Therefore, they cannot form boundaries between geometric patches. If one were to define a dissimilarity measure to capture corners, features in the surrounding space would most likely flow around these points and undermine the local changes in geometry which indicate a corner. An alternate strategy is to compute

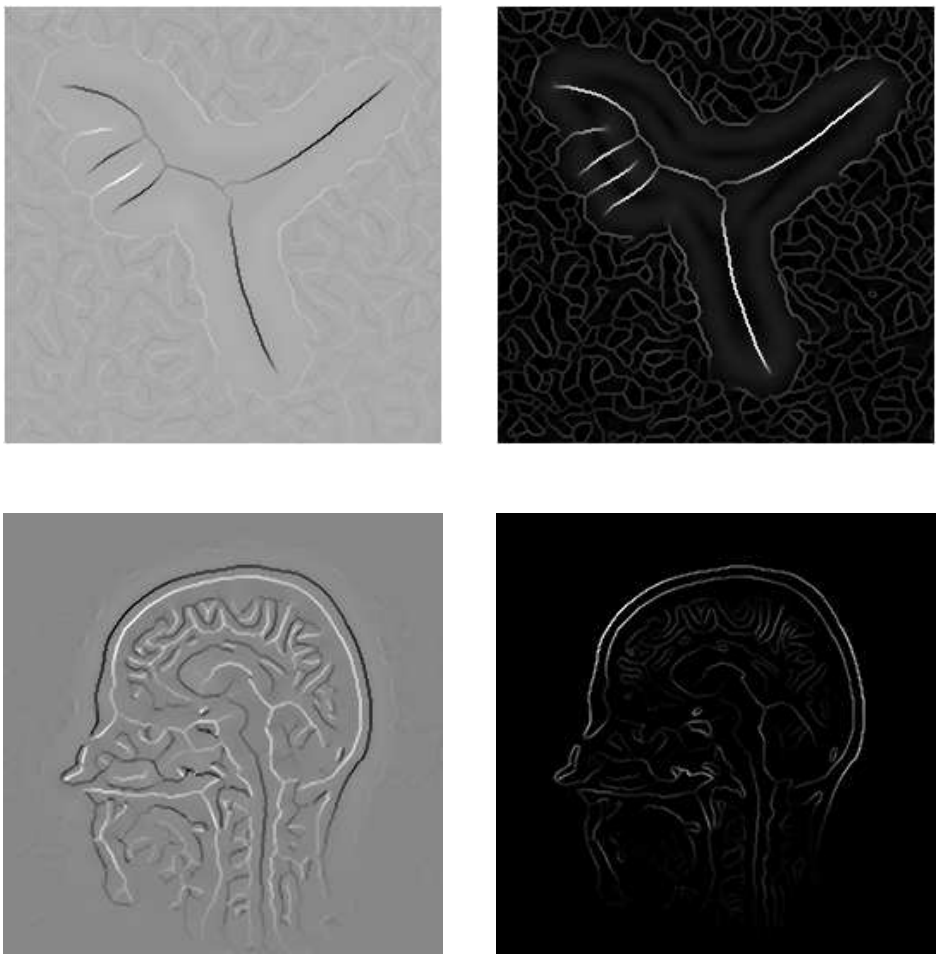

Figure 4.9: The trace (left) and norm (right) of the diffused Hessian for the *blob* (top) and *head* (bottom) images.

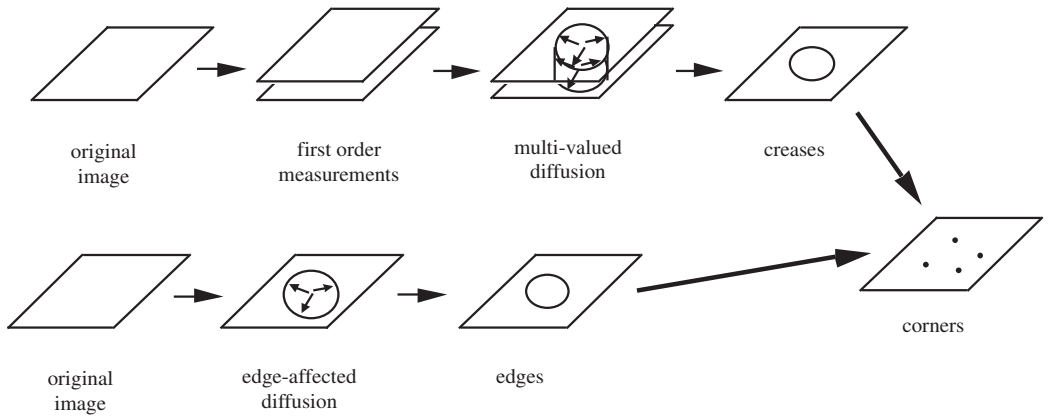

Figure 4.10: The strategy for finding corners is to isolate the zeroth-order and first order patches separately and then combine the results.

and process zeroth and first-order feature spaces separately and then combine the information that results from each of these processes. A reasonable description of a corner is a place on the edge of an object where the boundary turns very sharply. In order for a boundary to turn suddenly, the gradients along that boundary must undergo a large turn in a small neighborhood. This is precisely the definition used previously for creases. *A corner is a place in the image that is both a crease and an edge.* This suggests that one could obtain corners by combining the results of the crease calculations with edges discerned from zeroth-order diffusion. Because a corner must have both a high gradient (derivative of zeroth-order quantity) and a large variation in gradient direction (derivative of a first-order quantity), it is natural to multiply the measurements which result from the zero and first-order diffusion processes. To this end, Blom (1992) proposes a corner detector of the form $f_{vv}{f_w}^2$ which has some very useful properties, including affine invariance. By analogy a strategy for locating corners is to produce zero and first-order segmentations and then combines the results of the separate processes using the product of zeroth- and first-order dissimilarity measures (depicted in Fig. 4.10).

The sample image in Fig. 4.11a is a white hexagon on a black background. Uncorrelated random noise is added to the image. The value of $\|H^{(1)}\|$ which results from the first-order diffusion processing and the value of $|\nabla F|$ which results from the zeroth-order processing are shown in Figs. 4.11b–c respectively. The image of Fig. 4.11d is the result of multiplying the pixel values of the images in Figs. 4.11b and c. The corners that result from a simple threshold of Fig. 4.11d are accurate to within 3 pixels.

4.5 Invariance

4.5.1 Geometric invariance

Any image processing algorithm should produce results that are independent of the coordinate system that is used to describe the image space. In particular, the results should not be influenced by translations and rotations of the original image. The encoding of local geometry (Sect. 4.3) in terms of coefficients of Taylor series forces the choice of a coordinate system. The absolute positions of points in the feature space depend on this choice. This is because the axes of the feature space are associated with sets of directional derivatives and the directions of these derivatives are aligned with the coordinate axis in the image space as described in Sect. 4.2. This presents a serious question. How can one construct a system which produces invariant results but which relies on a feature space

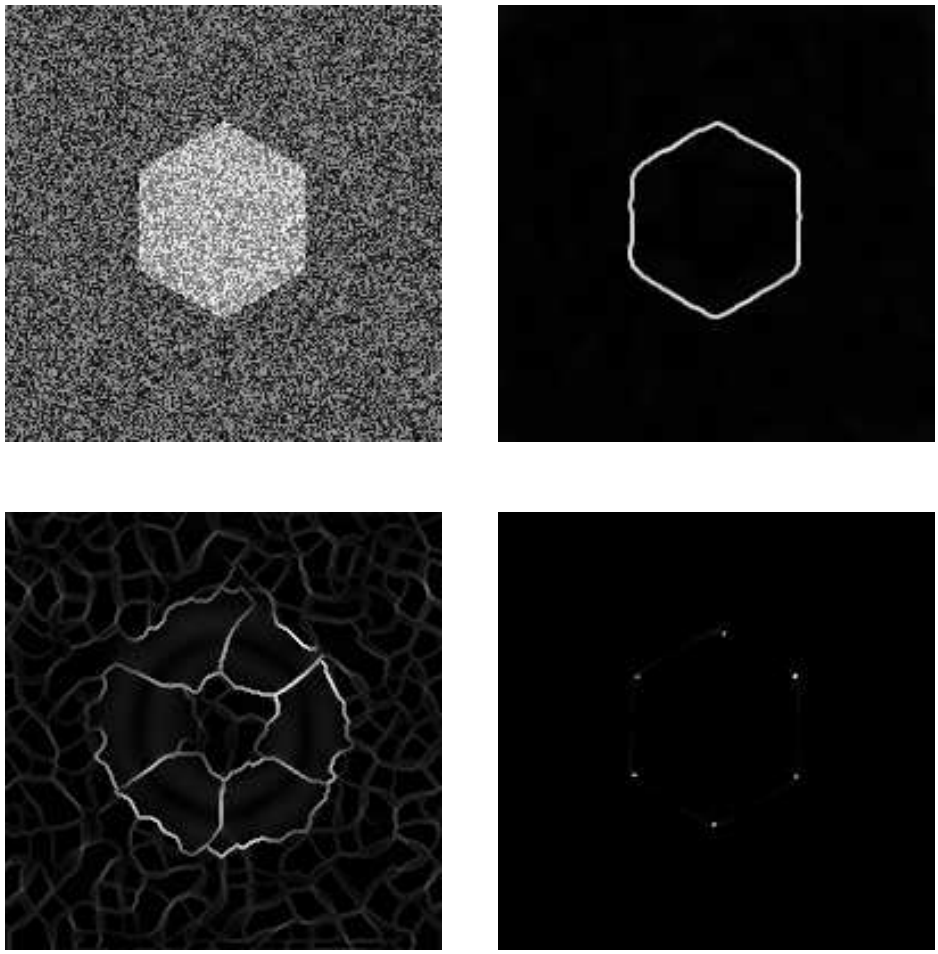

Figure 4.11: Clockwise from the upper left: (a) A test image composed of white hexagon on a black background with additive random noise (pixelwise, uniformly distributed, 50%). (b) The gradient squared after zeroth-order diffusion. (c) The creases that result from first-order diffusion. (d) The result of multiplying the edge and crease measures in order to detect corners.

that is intimately tied to our choice of coordinate systems?

The answer lies in the dissimilarity measure, which defines distance in the feature space. Positions in the feature space need not be invariant to orthogonal transformations in order to obtain invariant results. It *is* essential, however, that the *relative positions* of points and the distances between those points be invariant. To ensure this invariance, express the dissimilarity as a geometric invariant of the intensity surface of the original image. Such a dissimilarity is invariant at the start of the process because it is a geometric invariant of the intensity surface. As the process progresses, however, the terms of the invariant change so that they are no longer the derivatives of the intensity surface. Does dissimilarity remain invariant (under the orthogonal group) as the nonlinear, nonuniform diffusion progresses?

In this Section I will prove a theorem which states that the behavior of the geometry limited diffusion process is invariant to transformations of the original coordinate system. I am primarily concerned with *polynomial invariants* as described by ter Haar Romeny et al. (1991a), where polynomials of image derivatives are expressed using the Einstein summation convention of summing expressions like indices over the basis. For example, the square of the gradient magnitude of a function, $f : \mathbb{R}^2 \mapsto \mathbb{R}$, is expressed as

$$f_i f_i \stackrel{\text{def}}{=} \sum_{i \in x,y} f_i f_i = {f_x}^2 + {f_y}^2 \quad , \tag{4.9}$$

while the Laplacian is

$$f_{ii} \stackrel{\text{def}}{=} \sum_{i \in x,y} f_{ii} = f_{xx} + f_{yy} \quad . \tag{4.10}$$

Any expression of this form with the property that all indices are paired (no free indices) is a scalar which is invariant to orthogonal coordinate transformations (rotations and translations) and is thereby independent of one's choice of coordinates.

In geometry-limited diffusion we are concerned not only with invariant expressions of derivatives of scalar functions, but also with *diffused* derivatives. Diffused derivatives are functions that evolve over time according to the multi-valued diffusion equation and have image derivatives as their initial conditions. Recall that the raised indices indicate the initial conditions associated with a single value of a multi-valued function. These functions satisfy the geometry-limited diffusion equation. For instance, if $f(\vec{x})$ is an image (scalar function), and $\vec{F}(\vec{x},t) = F^x(\vec{x},t)\ F^y(\vec{x},t)$,

$$\nabla \cdot g(\mathcal{D}\vec{F}) F^x(\vec{x},t) \text{ where } F^x(\vec{x},0) = f_x(\vec{x}),$$

and

$$\nabla \cdot g(\mathcal{D}\vec{F}) F^y(\vec{x},t) \text{ where } F^y(\vec{x},0) = f_y(\vec{x}).$$

Diffused polynomial invariants are expressions that have *paired upper and lower case indices.* Thus, $F^i F^i$ is one instance of a diffused polynomial invariant, while F^i_i is another. These expressions are functions of both time and space. They mimic the polynomial invariants described by the Einstein summation convention, but we cannot know a priori that their behavior is invariant—their invariant properties remain to be proven. At time $t = 0$, however, the invariant properties of diffused polynomial invariants are well understood; they are polynomial invariants. That is, at $t = 0$ the values of *diffused polynomial invariants* are invariant under orthogonal group transformations.

I show invariance by considering a transformed system which is expressed in terms of a new set of coordinates (Olver 1986). That is,

$$\vec{F}(\vec{x},t) \mapsto \underline{\vec{F}}(\underline{\vec{x}},\underline{t}).$$

Invariance is proven by showing that the geometry-limited diffusion equation is satisfied by the new function $\underline{\vec{F}}$ in terms of the new coordinates $\underline{x}$ and $\underline{t}$. That is,

$$\nabla \cdot g(\mathcal{D}_s \vec{F}) \nabla \vec{F} - \frac{\partial \vec{F}}{\partial t} = 0 \implies \underline{\nabla} \cdot g(\underline{\mathcal{D}}_s \underline{\vec{F}}) \underline{\nabla} \underline{\vec{F}} - \frac{\partial \underline{\vec{F}}}{\partial \underline{t}} = 0$$

For this work I consider only rotations in the image space, where

$$\underline{\vec{x}} = A\vec{x}, \;\; \underline{t} = t,$$

and the elements $\vec{F}$ transform as the derivatives that they represent.

As a notational convenience I use the Einstein summation convention to represent the various tensor operations. Thus, $\vec{x} \stackrel{\text{def}}{=} x_j$, and $\underline{x}_i = a_{ij} x_j$. Let $B = A^{-1}$, so that $x_i = b_{ij} \underline{x}_j$ Derivatives are expressed in a similar fashion. Using the chain rule,

$$\frac{\partial}{\partial \underline{x}_i} = \frac{\partial}{\partial x_j} \frac{\partial x_j}{\partial \underline{x}_i} = b_{ji} \frac{\partial}{\partial x_j},$$

and

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial \underline{x}_j} \frac{\partial \underline{x}_j}{\partial x_i} = a_{ji} \frac{\partial}{\partial \underline{x}_j}.$$

Note that $a_{ij} b_{jk} = \delta_{ik}$, where

$$\delta_{ik} = \begin{cases} 1 & \text{for } i = k \\ 0 & \text{for } i \neq k \end{cases}$$

is the Kronecker delta (ter Haar Romeny et al. 1991a). Because A is a rotation matrix, $a_{ij} = b_{ji}$ (in conventional matrix notation $A = B^T$) and $a_{ij} a_{ik} = \delta_{jk}$. Also note that δ_{jk} behaves as the identity matrix, so $\delta_{jk} = \delta_{kj}$ and for any tensor c_{jm}, $\delta_{kj} c_{jm} = c_{km}$. The Einstein summation convention has the advantage (over convential matrix notation) of not relying on the ordering of the parts of an expression, e.g., $a_{ij} b_{jk} = b_{jk} a_{ij}$.

Before confronting the invariance of the general geometry-limited diffusion framework, I show that the first-order diffusion process from Sect. 4.4 is invariant under orthogonal transformations. This proof of invariance in first-order case will motivate a theorem and a proof of the general case of geometry-limited diffusion of higher-order features. For first-order diffusion, $\vec{F} = F^i$. I use the raised indices in order to be consistent with earlier notation [1]. I denote all diffused features and functions of those features with raised indices. Despite the raises indices, F^i acts like any other 1-tensor. We would like the new function $\underline{F}^i$ to transform in the same way as the gradient of the original scalar image f. That is,

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial \underline{x}_j} \frac{\partial \underline{x}_j}{\partial x_i} = a_{ji} \frac{\partial f}{\partial \underline{x}_j}$$

gives the relationship $F^i = a_{ji} \underline{F}^j$. In addition we need to consider derivatives of F^i:

$$\frac{\partial F^i}{\partial x_j} = a_{ki} \frac{\partial \underline{F}^k}{\partial x_j} = a_{ki} a_{lj} \frac{\partial \underline{F}^k}{\partial \underline{x}_l}.$$

Derivatives with respect to t require no transformations, because none of the spatial transformations depend on time.

Invariance is established in two steps. The first step is to show that the dissimilarity is invariant,

[1] The reader should not be confused with the conventional use of raised indices in tensor notation, which are routinely used in the context of general affine transformations in order to preserve the ordering of tensors.

and the second step is to show that the entire diffusion equation is invariant. Recall that for the first-order diffusion the dissimilarity is

$$\begin{aligned}
\mathcal{D}\vec{F} \stackrel{\text{def}}{=} \|\frac{\partial}{\partial x_i}\frac{F^j}{(F^k F^k)^{\frac{1}{2}}}\| &= \|a_{li}\frac{\partial}{\partial \underline{x}_l}\frac{a_{jq}\underline{F}^q}{(a_{km}\underline{F}^m a_{kn}\underline{F}^n)^{\frac{1}{2}}}\| \\
&= \|a_{jq}a_{li}\frac{\partial}{\partial \underline{x}_l}\frac{\underline{F}^q}{(\delta_{mn}\underline{F}^m\underline{F}^n)^{\frac{1}{2}}}\| \\
&= \|a_{jo}a_{li}\frac{\partial}{\partial \underline{x}_l}\frac{\underline{F}^j}{(\underline{F}^k\underline{F}^k)^{\frac{1}{2}}}\| \\
&= \|\frac{\partial}{\partial \underline{x}_i}\frac{\underline{F}^j}{(\underline{F}^k\underline{F}^k)^{\frac{1}{2}}}\| \\
&= \underline{\mathcal{D}\vec{F}},
\end{aligned} \tag{4.11}$$

where the next to last step is because the Euclidean matrix norm is unchanged by rotations.

Now show invariance for the diffusion equation by using the tensor notation in lieu of the gradient operators:

$$\begin{aligned}
0 &= \frac{\partial}{\partial x_j}\left(g(\mathcal{D}\vec{F})\frac{\partial F^i}{\partial x_j}\right) - \frac{\partial F^i}{\partial t} \\
&= a_{lj}\frac{\partial}{\partial \underline{x}_l}\left(g(\underline{\mathcal{D}\vec{F}})a_{ki}a_{mj}\frac{\partial \underline{F}^k}{\partial \underline{x}_m}\right) - a_{qi}\frac{\partial \underline{F}^q}{\partial t} \\
&= a_{lj}a_{ki}a_{mj}\frac{\partial}{\partial \underline{x}_l}\left(g(\underline{\mathcal{D}\vec{F}})\frac{\partial \underline{F}^k}{\partial \underline{x}_m}\right) - a_{qi}\frac{\partial \underline{F}^q}{\partial t} \\
&= b_{in}a_{ki}a_{lj}a_{mj}\frac{\partial}{\partial \underline{x}_l}\left(g(\underline{\mathcal{D}\vec{F}})\frac{\partial \underline{F}^k}{\partial \underline{x}_m}\right) - b_{in}a_{qi}\frac{\partial \underline{F}^q}{\partial t} \quad \begin{array}{l}\text{(can multiply by}\\ b_{in}\text{ because}\\ \text{quantity is zero)}\end{array} \\
&= \delta_{kn}\delta_{lm}\frac{\partial}{\partial \underline{x}_l}\left(g(\underline{\mathcal{D}\vec{F}})\frac{\partial \underline{F}^k}{\partial \underline{x}_m}\right) - \delta_{qn}\frac{\partial \underline{F}^q}{\partial t} \\
&= \frac{\partial}{\partial \underline{x}_m}\left(g(\underline{\mathcal{D}\vec{F}})\frac{\partial \underline{F}^n}{\partial \underline{x}_m}\right) - \frac{\partial \underline{F}^n}{\partial t},
\end{aligned}$$

which is of the required form, a first-order, geometry-limited diffusion. Thus, the first-order diffusion is invariant to orthogonal group transformations (in the image space).

This result is important because it states that solutions to the first-order equation transform in the same manner as the initial conditions provided that the transformations are orthogonal and do not explicitly depend on space or time. Thus, the solutions to the first-order diffusion equation behave like a 1-tensor. This means that the dissimilarity measure, as well as any other function of polynomial invariants, remains insensitive throughout the process to the original choice of coordinates. For example, the gradient magnitude $(F^iF^i)^{\frac{1}{2}} = ((F^y)^2 + (F^x)^2)^{\frac{1}{2}}$ and the Laplacian (twice the mean curvature) $F^i{}_i = F^x{}_x + F^y{}_y$ are both invariants. As another example consider the Hessian define by diffused first-order derivatives $H^{(1)}$ (Sect. 4.4.4). This tensor can be expressed in a new coordinate system, $\underline{x}\underline{y}$, that is rotated by an angle θ. In two dimensions

$$\begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix},$$

were $a = \cos\theta$ and $b = \sin\theta$. The Hessian is

$$\begin{aligned} H^{(1)} &= \begin{bmatrix} \underline{F^x}_{\underline{x}} & \underline{F^x}_{\underline{y}} \\ \underline{F^y}_{\underline{x}} & \underline{F^y}_{\underline{y}} \end{bmatrix} \\ &= \left(\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial x} \end{bmatrix} \right) \left(\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} F^x \\ F^y \end{bmatrix} \right) \\ &= \begin{bmatrix} a^2F^x_x + ab(F^x_y + F^y_x) + b^2F^y_y & -abF^x_x + a^2F^x_y - b^2F^y_x + abF^y_y \\ -abF^x_x - b^2F^x_y + a^2F^y_x + abF^y_y & b^2F^x_x - ab(F^x_y + F^y_x) + a^2F^y_y \end{bmatrix} \end{aligned}$$

The trace of this matrix is

$$\begin{aligned} a^2F^x_x &+ ab(F^x_y + F^y_x) + b^2F^y_y + b^2F^x_x - ab(F^x_y + F^y_x) + a^2F^y_y \\ &= F^x_x + F^y_x, \end{aligned}$$

which confirms the invariant behavior of the *diffused* mean curvature.

Having confirmed that the first-order diffusion process produces invariant results, I wish to consider the general geometry-limited diffusion framework. The invariant behavior of a system will depend on the form of the dissimilarity. In transforming the dissimilarity for the first-order case, the term inside of the Euclidean norm (Eq. (4.11)) had the form of a 2-tensor. After transformation we were left with a 2-tensor that had precisely the same form of the original tensor with some additional rotation terms "out front". That is,

$$\mathcal{D}\vec{F} = \|T_{ij}\| = \|a_{jm}a_{il}\underline{T}_{ij}\| = \|\underline{T}_{ij}\| = \mathcal{D}\vec{F}.$$

A particularly useful property of tensors is that they are transformed to new coordinate systems by repeated applications of a rotation tensors. In general I will say that a quantity is a *tensor* if, when changing coordinate systems, it transforms as

$$T_{i_1 \ldots i_p} = a_{j_1 i_1} \ldots a_{j_p i_p} \underline{T}_{j_1 \ldots j_p}. \tag{4.12}$$

This gives rise to the following theorem.

Theorem 1 *Given a geometry-limited diffusion system as described in Sect. 4.3 with*

1. *a set of features, $\vec{F} = F_1, ..., F_m$, with initial values that are complete sets of derivatives of a smooth image and*
2. *a dissimilarity measure, $\mathcal{D}\vec{F}$, that has the form*

$$\mathcal{D}\vec{F} = \|\nabla\vec{G}(\vec{F})\|,$$

where $\vec{G}$ is a set of intermediate features that form a tensor.

then solutions are invariant to orthogonal group transformations for all $t \geq 0$.

Proof of Theorem 1 The proof proceeds in three parts. First, establish some basic properties of higher-order tensors (i.e. not just vectors, which are 1-tensors). Second, show that the dissimilarity is invariant. Finally, show that the diffusion equation is satisfied in the new coordinates.

Consider a geometry-limited diffusion system which incorporates features that represent (have as their initial conditions) image derivatives up to order N. The set of features $\vec{F} = F_1, ...F_m$ can be represented as set of tensors which correspond the order of the derivatives that they represent. The zero-order tensor is simply F, while the first-order tensor is F^i, as described earlier. The Nth-order

tensor is $F^{i_1 \ldots i_N}$. The initial conditions of each tensor can be expressed using the same notation:

$$F^{i_1 \ldots i_N} \mid_{t=0} = \frac{\partial^N f}{\partial x_{i_1} \ldots x_{i_N}},$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is the image. In order to have geometric invariance, each of these tensors associated with the features must transform in the same way as the initial conditions. The transformations associated with higher-order tensors have a distinct pattern. Consider the derivatives of f:

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= a_{ji} \frac{\partial f}{\partial \underline{x}_j}, \qquad (4.13) \\ \frac{\partial^2 f}{\partial x_{i_1} x_{i_2}} &= a_{j_1 i_1} a_{j_2 i_2} \frac{\partial f}{\partial \underline{x}_{j_1} \underline{x}_{j_1}}, \\ &\vdots \\ \frac{\partial^N f}{\partial x_{i_1} \ldots x_{i_N}} &= a_{j_1 i_1} \ldots a_{j_N i_N} \frac{\partial f}{\partial \underline{x}_{j_1} \ldots \underline{x}_{j_N}}. \qquad (4.14) \end{aligned}$$

These equations prescribe the transformations of the higher order features, which are treated as sets of tensors:

$$\begin{aligned} F^i &= a_{ji} \underline{F}^j, \qquad (4.15) \\ F^{i_1 i_2} &= a_{j_1 i_1} a_{j_2 i_2} \underline{F}^{i_1 i_2}, \\ &\vdots \\ F^{i_1 \ldots i_N} &= a_{j_1 i_1} \ldots a_{j_N i_N} \underline{F}^{i_1 \ldots i_N}. \end{aligned}$$

Now let us establish the invariance of the dissimilarity. Recall that the dissimilarity can be expressed in terms of the gradient of a set of intermediate features, i.e.,

$$\mathcal{D}\vec{F} \stackrel{\text{def}}{=} \|\nabla \vec{G}(\vec{F})\|.$$

Invariance requires that

$$\mathcal{D}\vec{F} \stackrel{\text{def}}{=} \|\nabla \vec{G}(\vec{F})\| = \|\underline{\nabla} \vec{G}(\underline{\vec{F}})\| = \underline{\mathcal{D}\vec{F}}.$$

If $\vec{G}$ is a M-tensor, as we have assumed, the dissimilarity can be expressed in tensor notation:

$$G^{j_1 \ldots j_M}(F, F^{i_{11}}, F^{i_{21} i_{22}}, \ldots, F^{i_{N1} \ldots i_{NN}}),$$

where I have used i_{mn} to represent the nth index of the mth tensor. If $G^{j_1 \ldots j_M}$ transforms as a tensor, then

$$\begin{aligned} & G^{j_1 \ldots j_M}(F, F^{i_{11}}, F^{i_{21} i_{22}}, \ldots, F^{i_{N1} \ldots i_{NN}}) \\ &= G^{j_1 \ldots j_M}(\underline{F}, a_{j_{11} i_{11}} \underline{F}^{j_{11}}, a_{j_{21} i_{21}} a_{j_{22} i_{22}} \underline{F}^{i_{21} i_{22}}, \ldots, a_{j_{N1} i_{N1}} \ldots a_{j_{NN} i_{NN}} \underline{F}^{i_{N1} \ldots i_{NN}}) \\ &= a_{j_1 i_1} \ldots a_{j_M i_M} G^{j_1 \ldots j_M}(\underline{F}, \underline{F}^{i_{11}}, \underline{F}^{i_1 i_2} \ldots \underline{F}^{i_1 \ldots i_N}) \end{aligned}$$

That is, if $\vec{G}(\vec{F})$ is a tensor $T^{i_1 \ldots i_p}$, the new coordinates yield another tensor, $T^{i_1 \ldots i_p} = a_{j_1 i_1} \ldots a_{j_p i_p} \underline{T}^{j_1 \ldots j_p}$. The gradient of this tensor generates another tensor with yet another rota-

tion in front:

$$\frac{\partial}{\partial x_{i_{p+1}}} T^{i_1 \ldots i_p} = a_{j_1 i_1} \ldots a_{j_p i_p} \frac{\partial}{\partial x_{i_{p+1}}} \underline{T}^{i_1 \ldots i_p} = a_{j_1 j_1} \ldots a_{j_p i_p} a_{j_{p+1} i_{p+1}} \frac{\partial}{\partial \underline{x}_{j_{p+1}}} \underline{T}^{j_1 \ldots j_p}.$$

Now we need only to show that the Euclidean norm of a tensor does not depend on the orthogonal transformations which sit in front. The Euclidean norm of a tensor takes the form $\|T^{i_1 \ldots i_p}\| = (T^{i_1 \ldots i_p} T^{i_1 \ldots i_p})^{\frac{1}{2}}$. Thus,

$$\begin{aligned}
\|T^{i_1 \ldots i_p}\| &= \|a_{j_1 i_1} \ldots a_{j_p i_p} \underline{T}^{j_1 \ldots j_p}\| \\
&= \left(a_{l_1 i_1} \ldots a_{l_p i_p} \underline{T}^{l_1 \ldots l_p} a_{m_1 i_1} \ldots a_{m_p i_p} \underline{T}^{m_1 \ldots m_p} \right)^{\frac{1}{2}} \\
&= \left(a_{l_1 i_1} a_{m_1 i_1} a_{l_2 i_2} a_{m_2 i_2} \ldots a_{l_p i_p} \ldots a_{m_p i_p} \underline{T}^{l_1 \ldots l_p} \underline{T}^{m_1 \ldots m_p} \right)^{\frac{1}{2}} \\
&= \left(\delta_{l_1 m_1} \ldots \delta_{l_p m_p} \underline{T}^{l_1 \ldots l_p} \underline{T}^{m_1 \ldots m_p} \right)^{\frac{1}{2}} \\
&= \left(\underline{T}^{l_1 \ldots l_p} \underline{T}^{l_1 \ldots l_p} \right)^{\frac{1}{2}} \\
&= \|\underline{T}^{l_1 \ldots l_p}\|
\end{aligned}$$

Thus, if the intermediate features associated with the dissimilarity of a diffusion process behave like tensors, the dissimilarity is invariant to orthogonal group transformations. For the first-order case above, I showed explicitly that $F^i/|F^k|$ is a tensor. In general, any expression which can be constructed using combination of tensor products and scalar invariants (such as the gradient magnitude $|F_k|$) is a tensor and will therefore form an acceptable (invariant) set of intermediate features in the geometry-limited diffusion equation.

Now let us show that the geometry-limited diffusion equation is invariant. Once again, treat the features $\vec{F}$ in terms of the tensors which correspond to the derivatives that they represent. If the pth tensor (recall that there are a total of $N+1$ including the zero-tensor) is $F^{i_1 \ldots i_p}$, then

$$\begin{aligned}
0 &= \frac{\partial}{\partial x_{i_{p+1}}} \left(g(\mathcal{D}\vec{F}) \frac{\partial}{\partial x_{i_{p+1}}} F^{i_1 \ldots i_p} \right) \\
&\quad - \frac{\partial F^{i_1 \ldots i_p}}{\partial t} \\
&= \frac{\partial \underline{x}_{l_{p+1}}}{\partial x_{i_{p+1}}} \frac{\partial}{\partial \underline{x}_{l_{p+1}}} \left(g(\mathcal{D}\vec{\underline{F}}) a_{j_1 i_1} \ldots a_{j_{p+1} i_{p+1}} \frac{\partial}{\partial \underline{x}_{j_{p+1}}} \underline{F}^{j_1 \ldots j_p} \frac{\partial \underline{x}_{j_{p+1}}}{\partial x_{j_{p+1}}} \right) \\
&\quad - a_{k_1 i_1} \ldots a_{k_p i_p} \frac{\partial \underline{F}^{k_1 \ldots k_p}}{\partial t} \\
&= a_{l_{p+1} i_{p+1}} \frac{\partial}{\partial \underline{x}_{l_{p+1}}} \left(g(\mathcal{D}\vec{\underline{F}}) a_{j_1 i_1} \ldots a_{j_{p+1} i_{p+1}} \frac{\partial}{\partial \underline{x}_{j_{p+1}}} \underline{F}^{j_1 \ldots j_p} a_{j_{p+1} i_{p+1}} \right) \\
&\quad - a_{k_1 i_1} \ldots a_{k_p i_p} \frac{\partial \underline{F}^{k_1 \ldots k_p}}{\partial t} \\
&= a_{j_1 i_1} \ldots a_{j_p i_p} a_{l_{p+1} i_{p+1}} a_{j_{p+1} i_{p+1}} \frac{\partial}{\partial \underline{x}_{l_{p+1}}} \left(g(\mathcal{D}\vec{\underline{F}}) \frac{\partial}{\partial \underline{x}_{j_{p+1}}} \underline{F}^{j_1 \ldots j_p} \right) \\
&\quad - a_{k_1 i_1} \ldots a_{k_p i_p} \frac{\partial \underline{F}^{k_1 \ldots k_p}}{\partial t}
\end{aligned}$$

$$
\begin{aligned}
&= b_{i_1 q_1} \ldots b_{i_p q_p} a_{j_1 i_1} \ldots a_{j_p i_p} \delta_{l_{p+1} j_{p+1}} \frac{\partial}{\partial \underline{x}_{l_{p+1}}} \left(g(\underline{\mathcal{D}\vec{F}}) \frac{\partial}{\partial \underline{x}_{j_{p+1}}} \underline{F}^{j_1 \ldots j_p} \right) \\
&\quad - b_{i_1 q_1} \ldots b_{i_p q_p} a_{k_1 i_1} \ldots a_{k_p i_p} \frac{\partial \underline{F}^{k_1 \ldots k_p}}{\partial t} \\
&= \delta_{q_1 j_1} \ldots \delta_{q_{p+1} j_{p+1}} \frac{\partial}{\partial \underline{x}_{j_{p+1}}} \left(g(\underline{\mathcal{D}\vec{F}}) \frac{\partial}{\partial \underline{x}_{j_{p+1}}} \underline{F}^{j_1 \ldots j_p} \right) \\
&\quad - \delta_{q_1 k_1} \ldots \delta_{q_p k_p} \frac{\partial \underline{F}^{k_1 \ldots k_p}}{\partial t} \\
&= \frac{\partial}{\partial \underline{x}_j} \left(g(\underline{\mathcal{D}\vec{F}}) \frac{\partial}{\partial \underline{x}_j} \underline{F}^{q_1 \ldots q_p} \right) - \frac{\partial \underline{F}^{q_1 \ldots q_p}}{\partial t}
\end{aligned}
$$

which is the geometry-limited diffusion equation in the new coordinates. □

This proof suggests that geometry-lmited diffusion can be viewed as a diffusion process that operates on *tensors*. The features associated with a geometry-limited diffusion process can be treated in groups according to the order of derivative that they represent, and these groups, or tensors, transform in the same manner as the initial conditions. Because of this property, a large number of invariants can be created by treating the labels associated with diffused features (raised indices) as the actual derivatives, and these invariants will be invariant to spatial coordinate transformations throughout time. Theorem 1 in combination with the Einstein notation convention allows one to construct differential invariants by placing indices interchangeably in both the raised and lowered positions, where the difference in position indicates whether the derivatives are taken before or after the diffusion process.

The proof of Theorem 1 reveals another interesting aspect of these transformations. The transformations on the features $\vec{F}$ need not be the same as those used in the diffusion operator. Recall that the invariance of both dissimilarity and the diffusion equation relied on the *cancellation* of rotation matrices, i.e., the fact that rotation matrices which share indices collapse to delta functions. In all cases the rotations that resulted from the transformations on $\vec{F}$ were never required to interact with the rotations on the spatial coordinates $\vec{x}$. Thus, the transformations on $\vec{F}$ and $\vec{x}$ need not be the same; they could be any two rotation matrices. In the first-order diffusion, for instance, the diffusion equation is invariant to simultaneous transformations $\underline{x}^i = a_{ji} x^j$ and $\underline{F}^i = c_{ji} F^j$, where a_{ij} and c_{ij} are orthogonal transformations. This means that the features need not be computed in the same coordinate system that we "carry out" the diffusion process. The features could be first measured as derivatives in some arbitrary coordinate system and then diffused in any other coordinate system that is convenient. This result is a consequence of the fact the diffusion operator is scalar invariant, so transformations introduced into this operator have no effect on its behavior.

4.5.2 Greyscale invariance and dimensionless metrics

One might like to construct systems that are robust with respect to other kinds of transformations in addition to the orthogonal group. One of the drawbacks of anisotropic diffusion is its sensitivity to the contrast of edges in the initial image. This sensitivity requires that parameter k be set for each image on the basis of the range of intensities that are represented in that particular image. This sensitivity to the intensity range is particularly troublesome because the notion of "intensity" is ambiguous—the units used to measure luminance may differ among different images and applications. The parameter k can be set by manually for each image or automatically using some kind of heuristic such as multiplying it by the average gradient magnitude in each image (Sect. 2.7). Alternatively one could construct *dimensionless* distance measures in the feature space, and thereby compute dissimilarity measures that are independent of contrast. Such measures could allow the use of a

single k across a wider range of images.

Dimensionless measures are independent of the units in which the system is expressed and are not sensitive to linear transformations of values expressed in those units (homogeneity zero). Such dimensionless quantities play an important role in physics and are often used to capture intrinsic properties of physical systems (Cooper and West 1988). By using dimensionless measures in constructing dissimilarity, one can make shape comparisons that are not effected by any consistent rescaling of the image.

In the example of Sect. 4.4, where the goal is the detection of creases, the dissimilarity captures changes in orientation of the gradient vector. Orientation is a dimensionless quantity, which is represented as a position on the unit circle in order to aid the arithmetic. In dimensional analysis the Pi theorem (Cooper and West 1988), describes how many functionally independent dimensionless quantities can be constructed from a set of physical quantities. This theorem indicates that a system consisting of two measurements with the same dimensions encodes a single dimensionless quantity. Further constrain the system by rotational invariance (up to a constant), and orientation is unique. In general, I consider systems where the dissimilarity can be expressed in terms of the Jacobian of a set of intermediate, dimensionless quantities as in (4.16):

$$\mathcal{D}\vec{F} \stackrel{\text{def}}{=} \|\nabla \vec{G}(\vec{F})\|, \tag{4.16}$$

where $\vec{G} : \mathbb{R}^m \mapsto \mathbb{R}^k$ is dimensionless and generally $k \leq m$. The diffusion process is applied to the original features $\vec{F}$, and the intermediate features are used only for purpose of computing dissimilarity. The dissimilarity is the norm of the Jacobian of $\vec{G}$ and can, by the chain rule, always be expressed by a local transformation, $\phi(\vec{F})$ times Jacobian of $\vec{F}$ as described in Sect. 4.1.3. Typically, the intermediate features $\vec{G}$ are normalized versions of the features $\vec{F}$, and the intermediate features are expressed as the original features divided by some norm (Salden, ter Haar Romeny, Florack, Koenderink and Viergever 1991).

An important advantage of such normalized features is that they can provide a limited dynamic range for the dissimilarity measure. In the case of the first-order features, derivatives are measured on a compact subset of the original feature space—the unit circle. Because derivatives are measured at a finite scale using Gaussian kernels, these derivatives are bounded. This is important in light of the unstable "staircasing" effects that can occur in the case of very low conductance. Normalized features that are restricted to compact subspaces allow a choice of the conductance parameter k that enforces a lower limit on the conductance values. Normalizing the intermediate features by choosing the appropriate conductance parameter k can prevent staircasing for any image regardless of the range of intensities or the geometry. Experimental results confirm this analysis.

4.6 Second-order diffusion

4.6.1 Normalized Hessian

This section investigates the possibility of describing boundaries of objects in terms of the second-order features, F^{xx}, F^{xy}, and F^{yy} in Cartesian coordinates [2]. Note that $F^{xy} = F^{yx}$ because $F^{xy}(x, y, 0) = F^{yx}(x, y, 0) = f_{xy} = f_{yx}$. I wish to express boundaries in the image in terms of some dimensionless quantities of the second-order tensor (the Hessian) using the normalization approach described in Sect. 4.5.2. This approach offers the possibility of producing boundaries in a way that

[2] The x and y are not indices associated with the Einstein summation convention. I use the convention that letters from the middle of the alphabet, i, j, k, l, etc., represent indices to be used in a summation while letters from the end of the alphabet, u, v, w, x, and y, are used to represent coordinates systems. In this case the x and y are the axes associated with two-dimensional Cartesian coordinates.

does not depend on the overall contrast of objects. Using intermediate features which are normalized by the Euclidean norm of the Hessian (deviation from flatness (ter Haar Romeny et al. 1991a)),

$$\begin{aligned} \|\vec{F}\| = (F^{ij}F^{ij})^{\frac{1}{2}} = (F^{xx2} + 2F^{xy2} + F^{yy2})^{\frac{1}{2}} \\ G_1 = \frac{F^{xy}}{\|\vec{F}\|};\ G_2 = \sqrt{2}\frac{F^{xy}}{\|\vec{F}\|};\ G_3 = \frac{F^{xy}}{\|\vec{F}\|}, \end{aligned} \tag{4.17}$$

one is able to obtain boundaries that correspond to sharp variations in the "orientation" of this tensor. These intermediate features are constructed so that the dissimilarity (norm of Jacobian of $\vec{G}$) is an orthogonal group invariant. While large variations in the orientation of this tensor do correspond to certain kinds of visual boundaries, this measure does not respond well when boundaries exhibit high isophote curvature (e.g. at corners of objects).

This "breakdown" at corners results from the fact that the second order tensor includes information about second derivatives in all directions. At places of high isophote curvature the second-order tensor is dominated by curvature in the isophote direction. Changes in the "direction" of this tensor (gradient of the normalized Hessian) are also dominated by this isophote curvature. The visual feature called an edge seems to be more closely related to curvature in the gradient direction. This reasoning suggests that the effects of isophote curvature at corners can be reduced by considering the behavior of second-order properties in the gradient direction. The gradient direction is a first-order property, so boundaries between light and dark regions in images should described by both first- and second-order properties.

4.6.2 Combining first- and second-order tensors

Experiments with first-order and second-order features (a five-dimensional feature space) suggest that it is possible to construct a dissimilarity operator that captures the boundaries of objects as well as the "middles" of objects. This dissimilarity operator describes the manner in which the curvature of the intensity surface changes in the direction of the image gradient. Compute a pair of intermediate features which are the first partial derivatives of the gradient magnitude;

$$\begin{aligned} \nabla|\nabla f| &= \nabla\left((f_x^2 + f_y^2)^{\frac{1}{2}}\right) \\ &= (f_{xx}f_x + f_{xy}f_y,\ f_{xy}f_y + f_{yy}f_y)(f_x^2 + f_y^2)^{-\frac{1}{2}} \\ &= f_{ij}f_i\,(f_k f_k)^{-\frac{1}{2}}. \end{aligned} \tag{4.18}$$

Then divide this quantity by the Euclidean norm of the Hessian in order to produce a pair of dimensionless intermediate features:

$$\vec{g} = \frac{\nabla|\nabla f|}{|H|} = f_{ij}f_i\,((f_k f_k)(f_{lm}f_{lm}))^{-\frac{1}{2}}. \tag{4.19}$$

Use the raised indices to denote the first- and second-order parts of $\vec{F}$ in order to describe a pair of intermediate features $\vec{G}$ that depend on the diffused features, so $F^i = (F^x\ F^y)$ and $F^{jk} = \begin{pmatrix} F^{xx} & F^{xy} \\ F^{yx} & F^{yy} \end{pmatrix}$.

$$\vec{G} = \frac{F^l F^{lm}}{|F^i||F^{jk}|} = \left(\frac{F^x F^{xx} + F^y F^{xy}}{|F^i||F^{jk}|},\ \frac{F^x F^{xy} + F^y F^{yy}}{|F^i||F^{jk}|}\right) \tag{4.20}$$

The result of this computation is the product of the Jacobian and the Hessian divided by the norms of those two tensors. The vector described by $\vec{G}$ in (4.20) points in the direction of increasing gradient magnitude, and its orientation changes rapidly near places that are ridges or valleys of gradient magnitude as depicted in Fig. 4.12.

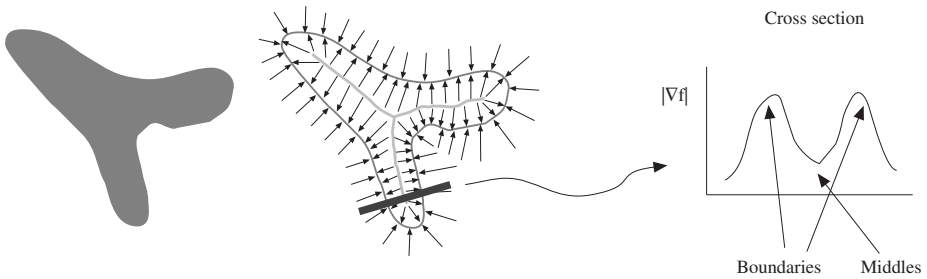

Figure 4.12: (a)–(b) The pair of intermediate features in 4.20 (when measured at finite scale) form a vector which points in the direction of edges. Sharp changes in the orientation of this vector occur near edges and at places that lie between edges, i.e. middles. (c) A cross section the gradient magnitude of the object shows that edges have influence across the entire object Ridges and valleys form at edges and middles respectively.

Figure 4.13: A *blob* image

Ridges of the gradient magnitude are places in an image that have high gradient values relative to points in a one-dimensional neighborhood that lies *transverse* to the ridge line. They correspond well with visual edges in the original image. Similarly, *valleys* of the gradient magnitude are places that have low gradient magnitudes with respect to the transversely oriented neighborhood. These are places that lie "away" from edges or are equally influenced by edges at two different places. In this respect they are the duals of edges and they represent the "middles" of objects. This analysis assumes that all measurements are made at some finite Gaussian scale, so that even for the flat object shown in Fig. 4.12 the edges have a range of influence that extends to the middle of the object.

Experiments with this system show that it provides a segmentation along edges that is not dependent on contrast or steepness but on the "sharpness" of edges, and at the same time it produces boundaries along valleys of gradient magnitude which are roughly the middles of objects. Figure 4.14a–b shows the dissimilarity of the *blob* image before and after processing at a scale of 3.0. As with the first-order case the dissimilarity forms the boundaries of coherent patches. The "flat" regions far from the object are sensitive to noise. In this case, the straight lines and right angles are indicative of artifacts introduced by the discrete Fourier transform which was used to create the the initial

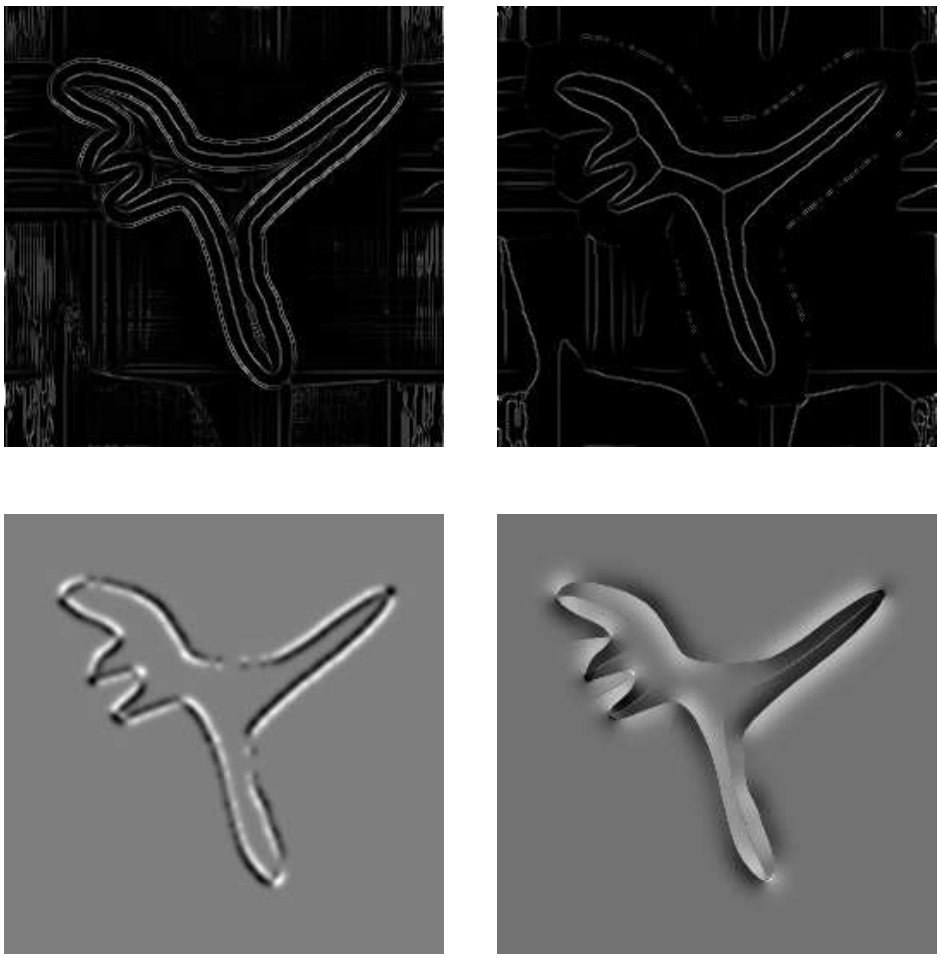

Figure 4.14: From the top left to right: (a)–(b) For the *blob* image (Fig. 4.13, the dissimilarity associated with the second-order process before and after processing at a scale of 2.0, and (c)–(d) the feature F^{xy} before and after processing.

image (Fig. 4.13). Figure 4.14c–d shows one of the features, F^{xy}, before and after processing. The richness of second-order information provides a number of interesting measurements. Boundaries are distinguished from middles on the basis of an orientation measure (Fig. 4.15a),

$$\mathcal{O} = \left(\frac{F^l F^{lm}}{|F^i| \|F^{jk}\|} \right)_m . \tag{4.21}$$

The dissimilarity does not distinguish between objects that are light or dark relative to their surroundings. After diffusion, light and dark regions can be distinguished on the basis of the sign of the mean curvature computed using the second-order features (Fig. 4.15b).

$$K^{(2)} = \frac{1}{2} F^{jj} \tag{4.22}$$

Compute the same measure on the first-order derivatives in order to distinguish ridges from valleys (Fig. 4.15c).

$$K^{(1)} = \frac{1}{2} F^j{}_j \tag{4.23}$$

Figure 4.16 shows these three measures for the processed MRI image. The dissimilarity indicates patch boundaries (Fig. 4.16b). The orientation (Fig. 4.16c) distinguishes edges (dark) from ridges (light), and the first-order curvature (Fig. 4.16d) indicates the curvature at patch boundaries and distinguishes ridges (dark) from valleys (light).

Figure 4.17 shows an image of radiograph as well as the dissimilarity $\mathcal{D}\vec{F}$, the orientation measure $\mathcal{O}$, and the curvature $K^{(1)}$ after processing at a scale of 3.5. The combination of these measures provides means of distinguishing middles from boundaries, light regions from dark regions, and ridges from valleys.

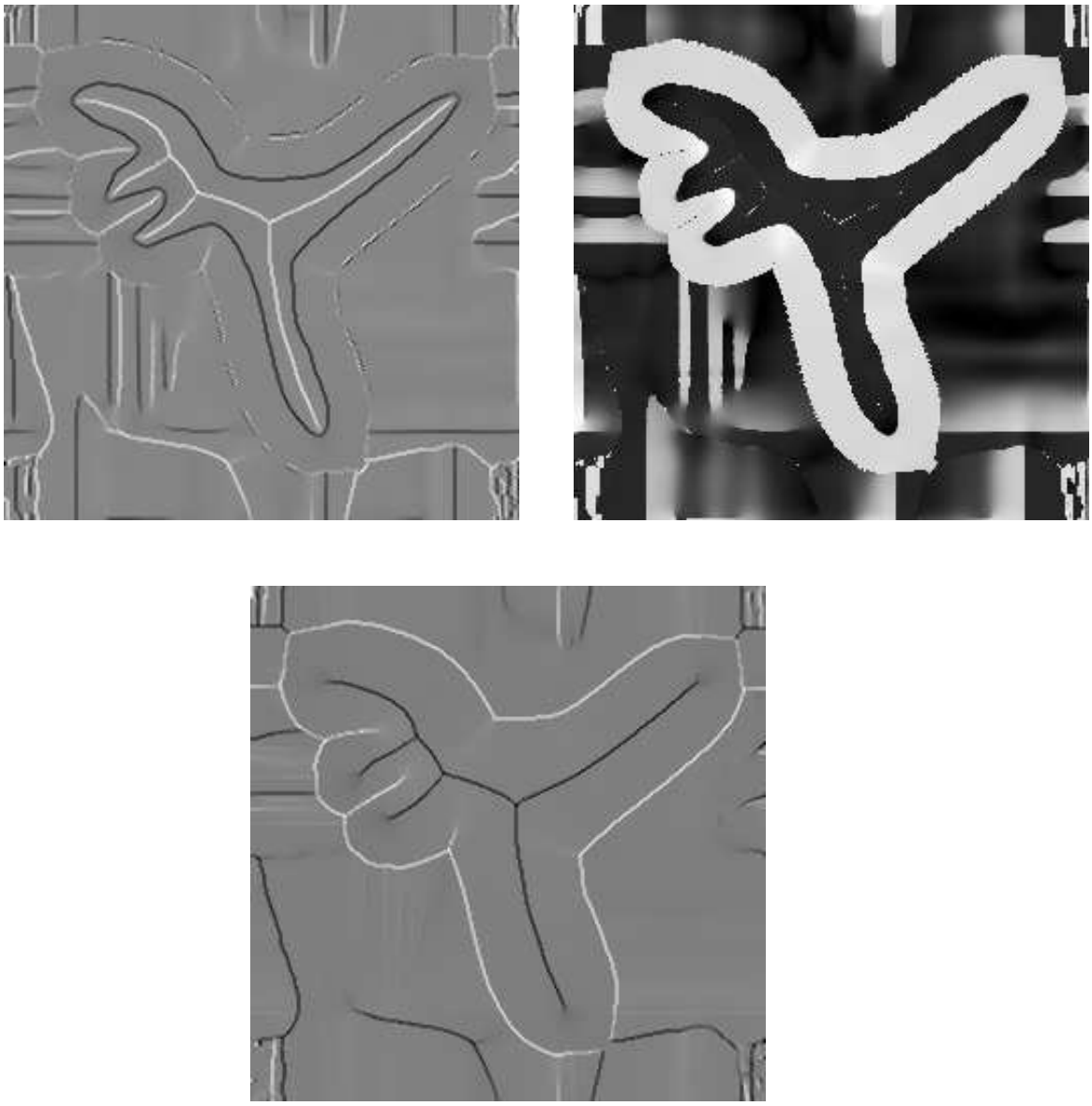

Figure 4.15: From the top left to right: Measurements made on the diffused second-order features include (a) the orientation $\mathcal{O}$ (black and white pixels denote negative and positive values respectively), (b) the second order curvature $K^{(2)}$, and (c) the first-order curvature $K^{(1)}$.

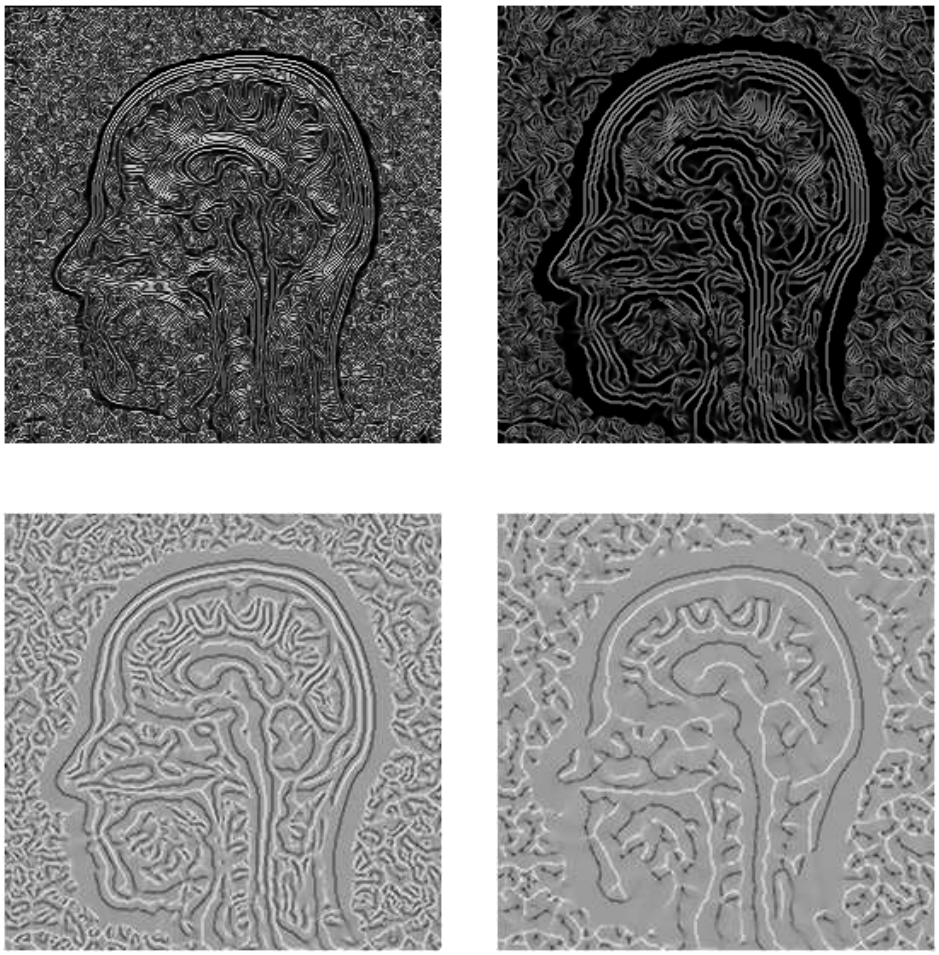

Figure 4.16: From the top left to right: First- and second-order features from the MRI head image produce (a) the dissimilarity before processing (b) the dissimilarity after processing at a scale of 3.5, (c) the orientation measure, $\mathcal{O}$, and (d) the curvature $K^{(}1)$.

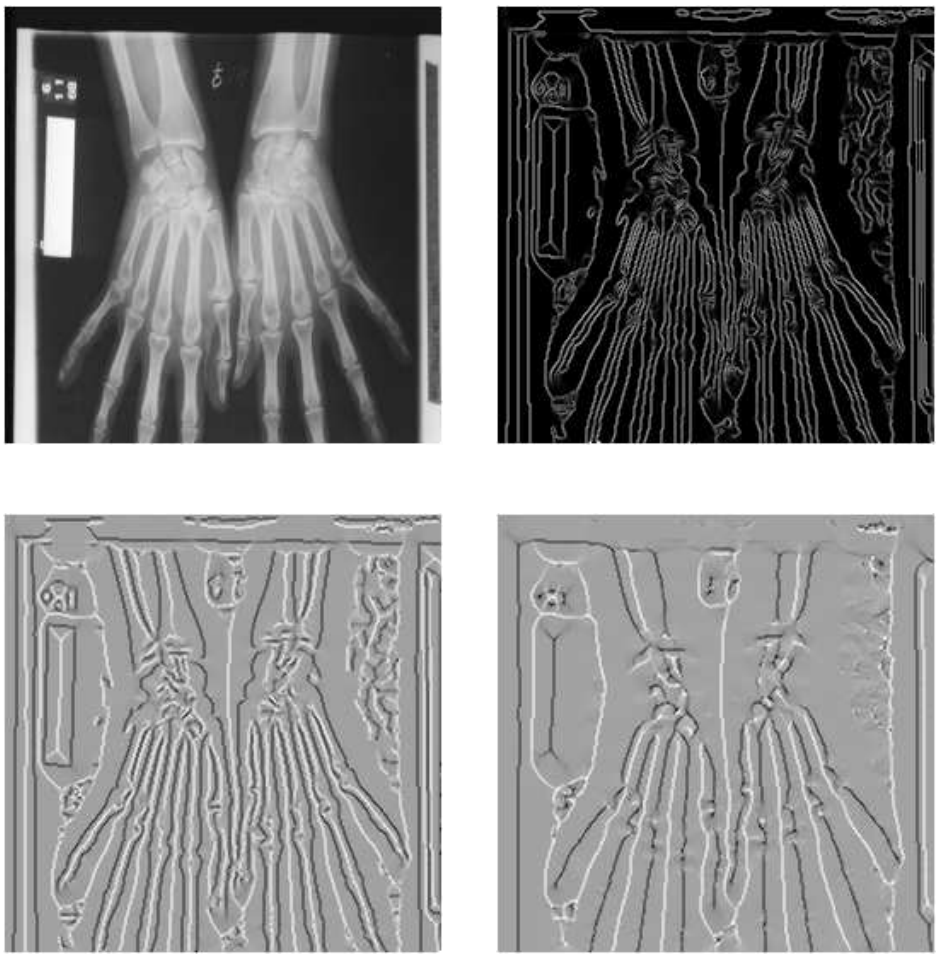

Figure 4.17: From the top left to right: (a) A radiograph of a pair of hands, (b) the dissimilarity after processing at a scale of 3.5, (c) the orientation measure, $\mathcal{O}$, and (d) the curvature $K^{(}1)$.