Chapter 4

The Geometry of Scale Space

4.1 Introduction

The necessity of using a multiscale analysis of images has clearly been established in the literature. The introduction of a continuous scale space can be found in Koenderink (1984), Witkin (1983), and Yuille & Poggio (1986). The fundamental constraint on a continuous scale space is that it be *causal*; that is, no spurious detail should be generated with increasing scale. Additional constraints involving linearity and symmetry lead to the fact that the Gaussian kernel is the unique scale space filter. A detailed investigation of scale space, including its natural differential operators and differential invariants is found in ter Haar Romeny, Florak, Koenderink & Viergever (1991). The issues of discretization of the operators are found in Lindeberg (1990).

The essential foundations of a scale space are that an image $I(\vec{x})$ is a physical observable with an inner scale σ_0, determined by the resolution of the sampling device, and an outer scale σ_1, limited by the field of view. A front–end vision system which allows for superposition of input stimuli (linearity), which samples and preprocesses its input in a symmetric way (rotational and translational invariance) and which has no preferred scale of measurement (scale invariance), can be modeled by the diffusion equation $B_\sigma(\vec{x}, \sigma) = \sigma \nabla^2 B(\vec{x}, \sigma)$ for $\vec{x} \in \mathbb{R}^n$ and $\sigma \in [\sigma_0, \sigma_1]$, with initial conditions $B(\vec{x}, \sigma_0) = I(\vec{x})$ for $\vec{x} \in \mathbb{R}^n$. The information derived at scale σ is $B(\vec{x}, \sigma) = K(\vec{x}, \sigma) \oplus I(\vec{x})$, which is the convolution of a radially symmetric Gaussian kernel of standard deviation σ with the input data. Researchers have investigated nonlinear diffusion processes as models of a front–end vision system (Grossberg 1984, Perona & Malik 1987, Whitaker 1993), where the assumption of linearity is not made so that object interactions within the image can be accounted for.

A key idea in ter Haar Romeny et al. (1991) is the invoking of dimensional analysis: A function relating physical observables must be independent of dimensional units. A set of *natural spatial coordinates* is proposed, namely $\vec{y} = \vec{x}/\sigma$. The natural distance between two points $\vec{x}_1$ and $\vec{x}_2$ at scale σ is $\|\vec{x}_1 - \vec{x}_2\|/\sigma$. Moreover, a *natural scale parameter* is proposed, namely $\tau = \ln(\sigma/\epsilon)$, where ϵ is a hidden scale whose units are dimension of length and which is image–dependent. The main consequence of their development is that differentiation at some selected scale σ can be made a well–posed operation if kernels constructed as derivatives of a Gaussian at the same scale are used. At a fixed scale σ, first–order dimensionless derivatives are σB_{x_i} for each spatial component x_i. Second–order dimensionless derivatives are $\sigma^2 B_{x_i x_j}$. In general, the dimensionless spatial derivatives are obtained from the usual partial derivatives by multiplying by the appropriate power of scale.

I propose a definition for scale space that is similar, but fundamentally different, from that described in ter Haar Romeny et al. (1991). It is desired that the front–end vision system show rotational invariance, translational invariance, and zoom invariance.[1] The change of variables to obtain the natural spatial coordinates $\vec{y} = \vec{x}/\sigma$ preserves translational invariance *if scale σ is assumed to be a constant.* It does not preserve translational invariance *through varying scale*; that is, the natural coordinates place an unnatural emphasis on the spatial origin $\vec{0}$. I want to develop a definition for scale space which has all the desired invariances for all scales, not just for a fixed scale.

To obtain the desired invariances, I assume that a measured *spatial difference* is meaningful only in the context of the scale at which it is measured. Similarly, when making multiscale measurements, a measured *scale difference* is meaningful only in the context of the scale at which it is measured. These assumptions suggest specifying differential forms as the measurement tools. I propose that the dimensionless 1–forms to be used for scale space

[1] I use the term *zoom invariance* rather than *scale invariance*, which is used in (ter Haar Romeny et al. 1991). The term refers to invariance with respect to changes in the scale (units) of measurement, which is equivalent to a uniform magnification in the space and scale variables. I seek to avoid confusion between the scale parameter and scale as units of measurement.

measurements are

$$\frac{d\vec{x}}{\sigma} \quad \text{and} \quad \frac{d\sigma}{\sigma}.$$

In contrast, the differential forms induced by the change of variables $\vec{y} = \vec{x}/\sigma$ and $\tau = \ln(\sigma/\epsilon)$ proposed in ter Haar Romeny et al. (1991) are

$$d\vec{y} = \frac{d\vec{x}}{\sigma} - \frac{\vec{x}}{\sigma}\frac{d\sigma}{\sigma} \quad \text{and} \quad d\tau = \frac{d\sigma}{\sigma}.$$

Note that for a *fixed scale* σ_0, the induced forms for natural coordinates are $d\vec{y} = d\vec{x}/\sigma_0$, which agree with my proposed forms. But for non–constant scale, the spatial forms are fundamentally different. In fact, one major consequence of using my forms is that the geometry of scale space is non–Euclidean, whereas the geometry of scale space using $d\vec{y}$ and $d\tau$ is still Euclidean.

This chapter provides the mathematical formalism of scale space as a geometric entity. A concise coverage of the mathematics used in this paper can be found in Kay (1988). In particular the reference covers tensor calculus, Riemannian geometry, Christoffel symbols, and covariant derivatives. Section 4.2 gives the definition for the metric tensor of scale space and shows how differentiation must be defined. Scale space is shown to be Riemannian with constant negative curvature. The isometries of the space verify that scale space has the desired invariance under rotation, translation, and zoom. Section 4.3 describes how to compute the gradient and Hessian of a real–valued function defined on scale space. The theory of curves, geodesics, distance, integration, and curvature of surfaces is discussed. Section 4.4 gives an extension of the definitions of *ridges* and *valleys* for Euclidean space, found in Chapter 2, to ones for scale space. Section 4.5 introduces a more general metric for scale space which depends on the image data itself. I show how the selection of the metric automatically determines which anisotropic diffusion process must be used to generate multiscale data. The conductance term and the density term which occur in the more general model for heat transfer show up as parameters in the metric. Finally, Section 4.6 is a discussion of the applicability of the ideas to construction of multiscale medial axes and surfaces (cores of objects). In particular, I discuss briefly the application of the ideas to image registration.

4.2 The Structure of Scale Space

Let scale space be denoted by $\mathcal{S} = \mathbb{R}^n \times (0, \infty)$ with typical element denoted by $\vec{\xi} = (\vec{x}, \sigma)$. The vector $\vec{x} \in \mathbb{R}^n$ represents the spatial information of the point and $\sigma > 0$ represents the scale information of the point. For indexing purposes, we have $\xi_i = x_i$ for $1 \le i \le n$ and $\xi_{n+1} = \sigma$.

4.2.1 Metric Tensor

The measuring tool used for component ξ_i is the 1–form $d\xi_i/\sigma$. The motivation is that a measured (spatial or scale) difference $d\xi_i$ is meaningful only in the context of the scale at which it is measured. That is, we need only be concerned with the relative measurement $d\xi_i/\sigma$. Note that the 1–forms are dimensionless quantities. The metric of the space is therefore determined by

$$ds^2 = \sum_{i=1}^{n} \lambda_i^2 \frac{dx_i^2}{\sigma^2} + \lambda^2 \frac{d\sigma^2}{\sigma^2} = \lambda^2 \left(\sum_{i=1}^{n} \rho_i^2 \frac{dx_i^2}{\sigma^2} + \frac{d\sigma^2}{\sigma^2} \right),$$

where the parameters $\lambda_i > 0$ may be selected to account for nonuniform scaling in spatial coordinates. For example, different units of measurements may be used for different coordinate directions, or the aspect ratio of an imaging device may not be 1. The parameter $\lambda = \lambda_{n+1} > 0$ may be selected to account for the fact that units of space and units of scale are not necessarily equally weighted. The second form above, where $\rho_i = \lambda_i/\lambda$, will be useful in the later developments. Define the $n \times n$ matrix $P = \mathrm{diag}(\rho_1, \ldots, \rho_n)$. The $(n+1) \times (n+1)$ metric tensor (as a 2×2 block diagonal matrix) is

$$G = [g_{ij}] = \frac{\lambda^2}{\sigma^2} \mathrm{diag}(P^2, 1).$$

Given two vectors $\vec{V}_i = (\vec{W}_i, \gamma_i) \in \mathbb{R}^n \times \mathbb{R}$, $i = 1, 2$, with initial point at $(\vec{x}, \sigma) \in \mathcal{S}$, their dot product with respect to the metric G is defined as

$$\vec{V}_1 \odot \vec{V}_2 = \vec{V}_1^{\,t} G \vec{V}_2 = \frac{\lambda^2}{\sigma^2} \left((P\vec{W}_1) \cdot (P\vec{W}_2) + \gamma_1 \gamma_2 \right),$$

where the single dot symbol on the right–hand side of the definition represents regular Euclidean dot product. Note that, unlike the regular Euclidean dot product, the scale space dot product does depend on the initial point of the vectors. The length of a vector $\vec{V} = (\vec{W}, \gamma)$ with initial point at $(\vec{x}, \sigma) \in \mathcal{S}$ is defined by

$$\|\vec{V}\| := \sqrt{\vec{V} \odot \vec{V}} = \frac{\lambda}{\sigma}\sqrt{|P\vec{W}|^2 + \gamma^2},$$

where the single bars represent regular Euclidean length. The angle between the two vectors is determined by

$$\cos\theta = \frac{\vec{V}_1 \odot \vec{V}_2}{\|\vec{V}_1\|\,\|\vec{V}_2\|}, \quad \theta \in [0, \pi].$$

The angle is well–defined since the right–hand side can be shown to be no larger than 1 in magnitude (*i.e.* the Cauchy–Schwartz inequality holds in this space). The condition for orthogonality of vectors is $\vec{V}_1 \odot \vec{V}_2 = 0$.

4.2.2 Christoffel Symbols and Covariant Derivatives

My analysis involves tensor quantities which need to be differentiated. In general, the partial derivative of a tensor is not necessarily a tensor. Tensor differentiation requires the use of some nontensorial objects called Christoffel symbols.

Let $\vec{e}_k \in \mathbb{R}^{n+1}$ denote the $(n+1) \times 1$ unit length vector whose components are all 0 except for the k^{th} component which is 1. The Christoffel symbols of the second–kind are defined by

$$\Gamma^k_{ij} = \frac{1}{2}\sum_{\ell=1}^{n+1} g^{k\ell}\left(\frac{\partial g_{j\ell}}{\partial x_i} + \frac{\partial g_{i\ell}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_\ell}\right)$$

where the g_{ij} are the components of G and the g^{ij} are the components of the inverse matrix G^{-1}. For my metric G, define the matrix $\Gamma^k = [\Gamma^k_{ij}]$; then

$$\Gamma^k = -\frac{1}{\sigma}\left\{\begin{array}{ll} \vec{e}_k\vec{e}^{\,t}_{n+1} + \vec{e}_{n+1}\vec{e}^{\,t}_k, & 1 \le k \le n \\ \vec{e}_{n+1}\vec{e}^{\,t}_{n+1} - \sum_{\ell=1}^{n}\rho_\ell^2\vec{e}_\ell\vec{e}^{\,t}_\ell, & k = n+1 \end{array}\right\}.$$

For example, if $n = 1$,

$$\Gamma^1 = -\frac{1}{\sigma}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \Gamma^2 = -\frac{1}{\sigma}\begin{bmatrix} -\rho_1^2 & 0 \\ 0 & 1 \end{bmatrix}$$

and if $n = 2$,

$$\Gamma^1 = -\frac{1}{\sigma}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \Gamma^2 = -\frac{1}{\sigma}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \Gamma^3 = -\frac{1}{\sigma}\begin{bmatrix} -\rho_1^2 & 0 & 0 \\ 0 & -\rho_2^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Note that each Γ^k is a symmetric matrix.

Let $\vec{T} : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ be a vector field, say $\vec{T}(\vec{x}) = [T_i(\vec{x})]$. As a vector field defined on Euclidean space, the *derivative of* $\vec{T}$ is the $(n+1) \times (n+1)$ matrix of partial derivatives given by $d\vec{T}/d\vec{\xi} = [\partial T_i/\partial \xi_j]$, where i is the row index and j is the column index. In a space whose curvature is not identically zero, the matrix of partial derivatives is not a tensor. The derivative of a tensor should again be a tensor. Intuitively, the measured differences are in a curved space, so the the usual derivative measurements need to be "corrected" to be consistent with the curvature of the space. *Covariant differentiation* is the correct generalization of ordinary differentiation. The application of covariant differentiation to vectors requires knowing whether the vector is *covariant* or *contravariant*. A covariant vector is analogous to the gradient of a function, which is normal to level surfaces. A contravariant vector is analogous to tangent vectors of the level surfaces. Let $\vec{U} = [U_i] = (\vec{W}, \gamma)$ be a covariant vector defined on scale space. The *covariant derivative of* $\vec{U}$ is the second–order tensor defined by

$$\frac{\mathrm{d\!l}\vec{U}}{\mathrm{d\!l}\vec{\xi}} = \left[\frac{\partial U_i}{\partial \xi_j} - \sum_{k=1}^{n+1} \Gamma_{ij}^k U_k\right] = \frac{d\vec{U}}{d\vec{\xi}} + \frac{1}{\sigma}\left[\begin{array}{c|c} -\gamma P^2 & \vec{W} \\ \hline \vec{W}^{\,t} & \gamma \end{array}\right].$$

Let $\vec{V} = [V^i] = (\vec{W}, \gamma)$ be a contravariant vector. The *covariant derivative of* $\vec{V}$ is the second–order tensor defined by

$$\frac{\mathrm{d\!l}\vec{V}}{\mathrm{d\!l}\vec{\xi}} = \left[\frac{\partial V^i}{\partial \xi_j} + \sum_{k=1}^{n+1} \Gamma_{jk}^i V^k\right] = \frac{d\vec{V}}{d\vec{\xi}} + \frac{1}{\sigma}\left[\begin{array}{c|c} -\gamma I & -\vec{W} \\ \hline P^2\vec{W}^{\,t} & -\gamma \end{array}\right].$$

In either case, the covariant derivative definitions include not only the usual partial derivatives but also correction terms dependent on the Christoffel symbols. A contravariant vector $\vec{V}$ can be converted to a covariant vector by $\vec{U} = G\vec{V}$. The two covariant derivative tensors are related by

$$\frac{\mathrm{d\!l}\vec{U}}{\mathrm{d\!l}\vec{\xi}} = \frac{\mathrm{d\!l}(G\vec{V})}{\mathrm{d\!l}\vec{\xi}} = G\frac{\mathrm{d\!l}\vec{V}}{\mathrm{d\!l}\vec{\xi}}\ ;$$

that is, the metric tensor is treated as a constant by covariant differentiation.

4.2.3 Riemann Tensor and Curvature

The implication of using the 1–forms $d\xi_i/\sigma$ for relative measurements is that the geometry of scale space is Riemannian. The Riemann tensor of the second kind is defined by

$$R^i_{jk\ell} = \frac{\partial \Gamma^i_{j\ell}}{\partial x_k} - \frac{\partial \Gamma^i_{jk}}{\partial x_\ell} + \sum_{m=1}^{n+1} \Gamma^m_{j\ell}\Gamma^i_{mk} - \sum_{m=1}^{n+1} \Gamma^m_{jk}\Gamma^i_{m\ell}.$$

The Riemann curvature relative to the metric G is defined for each pair of contravariant vectors $\vec{U} = [U^i]$ and $\vec{V} = [V^i]$ as

$$K(\vec{x}; \vec{U}, \vec{V}) = \frac{\sum_{i,j,k,\ell} R_{ijk\ell} U^i V^j U^k V^\ell}{\sum_{i,j,k,\ell} G_{ijk\ell} U^i V^j U^k V^\ell}$$

where $R_{ijk\ell} = \sum_{m=1}^{n+1} g_{im} R^m_{jk\ell}$ and $G_{ijk\ell} = g_{ik}g_{j\ell} - g_{i\ell}g_{jk}$. It can be shown that the only nonzero independent components are

$$R_{ijij} = -\sigma^{-4}\left(\frac{\lambda_i\lambda_j}{\lambda_{n+1}}\right)^2 \quad \text{and} \quad G_{ijij} = \sigma^{-4}(\lambda_i\lambda_j)^2,$$

where $1 \leq i < j \leq n+1$. The other components are either 0 or are determined by the values of the terms mentioned above. Consequently, the Riemann curvature of the space is identically a constant, $K = -1/\lambda^2$. Surprisingly enough, the curvature of scale space depends on the weight λ of the 1–form $d\sigma/\sigma$, but not on the spatial weights λ_i. When $\lambda_i = \lambda = 1$ for all i, this particular Riemannian geometry is called *hyperbolic geometry* and has been studied extensively in the differential geometry literature. A basic description of hyperbolic geometry is found in Fenchel (1989). An advanced treatise on the subject is Benedetti & Petronio (1987).

4.2.4 Isometries

When analyzing an image using multiscale techniques, the measurements should be invariant with respect to certain transformations. In particular, they should be invariant under spatial rotations, spatial translations, and spatial reflections. Object shape information produced by an algorithm should not depend on the orientation of the object being measured. Invariance with respect to *zoom* is also required. That is, if scale measurements are made for a given object, the scale measurements for a magnified version of the object should be the magnified measurements for the original object.

An *isometry* for scale space is a function $\vec{\psi} : S \to S$ whose differential preserves the dot product of vectors. Specifically, if $\vec{\psi} = (\psi_1, \ldots, \psi_{n+1})$ and $d\vec{\psi}/d\vec{\xi}$ is the matrix whose $(i,j)^{\text{th}}$ entry is $\partial\psi_i/\partial\xi_j$, then $\vec{\psi}$ is an isometry iff

$$\left(\frac{d\vec{\psi}}{d\vec{\xi}}\,\vec{V}_1\right) \odot \left(\frac{d\vec{\psi}}{d\vec{\xi}}\,\vec{V}_1\right) = \vec{V}_1 \odot \vec{V}_2,$$

where the vectors $\vec{V}_i$ are positioned at $\vec{\xi}$ and the vectors $(d\vec{\psi}/d\vec{\xi})\vec{V}_i$ are positioned at $\vec{\psi}(\vec{\xi})$. Scale space has the following isometries which provide the invariances described previously. It has an additional isometry which is mentioned for completeness.

$$\begin{array}{rcll}
\vec{\psi}(\vec{x},\sigma) & = & (\vec{x}+\vec{a},\sigma), & \text{translation by constant vector } \vec{a} \\
\vec{\psi}(\vec{x},\sigma) & = & (P^{-1}RP\vec{x},\sigma), & R \text{ is a rotation or reflection matrix} \\
\vec{\psi}(\vec{x},\sigma) & = & (\mu\vec{x},\mu\sigma), & \text{for any } \mu > 0 \text{ (zoom with magnification } \mu) \\
\vec{\psi}(\vec{x},\sigma) & = & \frac{(\vec{x},\sigma)}{\lambda^2(|P\vec{x}|^2+\sigma^2)}, & \text{inversion with respect to a hyperellipsoid}
\end{array}$$

Intuitively, the first three isometries indicate that if measurements involving angles or lengths are made at a given point, the measurements will be the same if you translate or rotate (assuming $P = I$) the spatial coordinates or if you change the units of measurement by zooming both both space and scale. A more detailed discussion of the isometries is found in Thorpe (1985).

4.3 Measurements in Scale Space

In this section I derive some basic formulas that are needed in multiscale algorithms. Since the space is Riemannian, the measurements must take into account the curvature of the space.

4.3.1 Gradient and Hessian

Let $f : S \to \mathbb{R}$ be a twice–differentiable function. Let f_{ξ_i} denote the partial derivative of f with respect to ξ_i and let $f_{\xi_i\xi_j}$ denote the second partial derivative of f with respect to ξ_i

and ξ_j. In Euclidean space the gradient of f is the vector $\nabla f = [f_{\xi_i}]$, and the Hessian of f is the matrix $\mathrm{H}f = [f_{\xi_i \xi_j}]$. The differential of f is given by

$$df = \underbrace{\sum_{i=1}^{n+1} f_{\xi_i} d\xi_i}_{\text{Euclidean}} = \underbrace{\sum_{i=1}^{n+1} \left(\frac{\sigma f_{\xi_i}}{\lambda_i}\right) \frac{\lambda_i d\xi_i}{\sigma}}_{\text{scale space}}.$$

Therefore, the *scale space gradient of* f is defined to be

$$\widehat{\nabla} f := \sqrt{G^{-t}} \nabla f = \left(\frac{\sigma f_{x_1}}{\lambda_1}, \ldots, \frac{\sigma f_{x_n}}{\lambda_n}, \frac{\sigma f_\sigma}{\lambda}\right).$$

The natural derivative operator is $(\sigma/\lambda_i)\partial/\partial\xi_i$, as indicated in Lindeberg (1993*b*) and ter Haar Romeny et al. (1991). The relationship of scale space gradient and Euclidean gradient is reminiscent of the equation obtained when making a change of variables $\vec{\zeta} = M\vec{\xi}$ and $g(\vec{\zeta}) = f(\vec{\xi})$, where $\nabla g = M^{-t}\nabla f$. (The notation M^{-t} is a concise representation of $(M^{-1})^t$.) Taking second derivatives yields $\mathrm{H}g = M^{-t}\frac{d\nabla f}{d\vec{\xi}} M^{-1} = M^{-t}\mathrm{H}f M^{-1}$. The similarity carries over to defining the *scale space Hessian of* f, but use covariant differentiation of the covariant vector ∇f must be used, denoted by $\mathrm{d\!l}$:

$$\begin{aligned}
\widehat{H} f &= \sqrt{G^{-t}}\, \frac{\mathrm{d\!l}\nabla F}{\mathrm{d\!l}\vec{\xi}}\, \sqrt{G^{-1}} \\
&= \frac{1}{\lambda^2}\mathrm{diag}(P^{-1}, 1) \left(\sigma^2 \mathrm{H} f + \sigma \left[\begin{array}{c|c} -f_\sigma P^2 & \nabla_{\vec{x}} f \\ \hline \nabla_{\vec{x}} f^{\,t} & f_\sigma \end{array}\right]\right) \mathrm{diag}(P^{-1}, 1),
\end{aligned}$$

where $\mathrm{H}f$ is the usual Hessian matrix of f in Euclidean space and where $\nabla f = (\nabla_{\vec{x}} f, f_\sigma)$. In Lindeberg (1993*b*) and ter Haar Romeny et al. (1991) it is suggested that the natural second derivatives are just $\sigma^2 f_{\xi_i \xi_j}$. But in this Riemannian setting, the correction terms must be included. The derivatives $\sigma^2 f_{\xi_i \xi_j}$ are natural only if the scale is fixed and the differential calculations are made on the manifold $\sigma = \sigma_0$, which is embedded in scale space.

4.3.2 Curves

Consider curves $\vec{\xi}(t) = (\vec{x}(t), \sigma(t))$ in scale space. The speed of a particle traveling along the curve is $ds/dt = \|\vec{\xi}'(t)\|$, the length of the tangent vector $\vec{\xi}'(t)$. If $t \in [t_0, t_1]$, the arc length of the curve is $\int_{t_0}^{t_1} ds/dt$. A curve is parametrized by arc length if $ds/dt \equiv 1$. To obtain a unit length tangent vector $\vec{T}(t)$, make the usual adjustment $\vec{T}(t) = \vec{\xi}'(t)/\|\vec{\xi}'(t)\|$.

If the curve is parameterized by arc length s, the unit length tangent vector is just $\vec{T}(s) = \vec{\xi}'(s)$. In Euclidean geometry, the s– derivative of $\vec{T}(s)$ is taken to obtain a vector which is normal to the curve. In Riemannian geometry the ordinary derivative is not necessarily a tensor. The analog is to take the *absolute derivative* along the curve, which does yield a tensor quantity. For a contravariant vector $\vec{V}(\vec{\xi})$ defined on a curve $\vec{\xi}(t)$, its absolute derivative is defined by

$$\frac{\mathbb{d}\vec{V}}{\mathbb{d}t} = \frac{\mathbb{d}\vec{V}}{\mathbb{d}\vec{\xi}}\frac{d\vec{\xi}}{dt}.$$

For a curve parametrized by arc length s, the unit principal normal is the contravariant vector $\vec{N}(s)$ defined by $\mathbb{d}\vec{T}/\mathbb{d}s = \kappa(s)\vec{N}$, where $\kappa(s) = \|\mathbb{d}\vec{T}/\mathbb{d}s\| \geq 0$ is the curvature of the curve. The explicit components of the absolute derivative of the tangent vector are

$$\begin{aligned}\left(\frac{\mathbb{d}T}{\mathbb{d}s}\right)_i &= \frac{d^2\xi_i}{ds^2} + \sum_{j=1}^{n+1}\sum_{k=1}^{n+1}\Gamma^i_{jk}\frac{d\xi_j}{ds}\frac{d\xi_k}{ds} \\ &= \left\{\begin{array}{ll} \frac{d^2x_i}{ds^2} - \frac{2}{\sigma}\frac{dx_i}{ds}\frac{d\sigma}{ds}, & 1 \leq i \leq n \\ \frac{d^2\sigma}{ds^2} - \frac{1}{\sigma}\left[\left(\frac{d\sigma}{ds}\right)^2 - \sum_{j=1}^{n}\left(\rho_j\frac{dx_j}{ds}\right)^2\right], & i = n+1 \end{array}\right\}.\end{aligned}$$

4.3.3 Geodesics

A curve is a *geodesic* if its curvature is identically zero, which means $(\mathbb{d}\vec{T}/\mathbb{d}s)_i = 0$ for all i. The geodesics are then solutions to the system of ordinary differential equations

$$\frac{d^2\vec{x}}{ds^2} = \frac{2}{\sigma}\frac{d\sigma}{ds}\frac{d\vec{x}}{ds} \text{ and } \frac{d^2\sigma}{ds^2} = \frac{1}{\sigma}\left[\left(\frac{d\sigma}{ds}\right)^2 - \left|P\frac{d\vec{x}}{ds}\right|^2\right].$$

These can be solved in closed form to obtain either

$$(\vec{x}(s), \sigma(s)) = (\vec{c}, r\exp(s/\lambda)) \tag{4.1}$$

for some constants $\vec{c}$ and $r > 0$, in which case the geodesic is a line in the direction of the scale axis, or

$$(\vec{x}(s), \sigma(s)) = (\vec{c} + r\tanh(s/\lambda)P^{-1}\vec{u}, r\,\mathrm{sech}(s/\lambda)) \tag{4.2}$$

for constants $\vec{c}, \vec{u} \in \mathbb{R}^n$ with $|\vec{u}| = 1$, and $r > 0$, in which case the geodesics are curves on half–ellipses of the form $|P(\vec{x} - \vec{c})|^2 + \sigma^2 = r^2$ with center on the hyperplane $\sigma = 0$.

Geodesics act as paths of minimum distance between points in scale space; see the next subsection. They also can be used to illustrate why scale space derivatives (of first

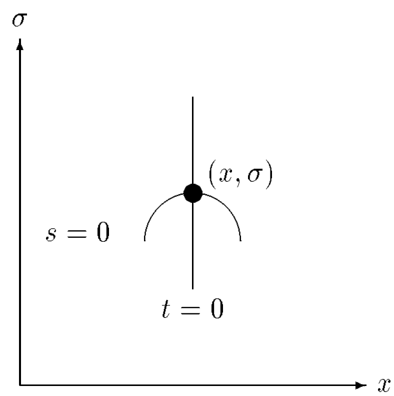

Figure 4.1: Geodesic coordinate axes

and second order) are the natural derivatives to compute in this space. Define $\vec{\xi}(s,t) = (\vec{x} + \sigma e^{t/\lambda} \tanh(s/\lambda)\vec{e}_k, \sigma e^{t/\lambda} \operatorname{sech}(s/\lambda))$. The curve $\vec{\xi}(s,0)$ is a half–ellipse geodesic whose north pole is at $(\vec{x}, \sigma)$, and the curve $\vec{\xi}(0,t)$ is a straight–line geodesic. The point $(\vec{x}, \sigma)$ can be thought of as the origin of a coordinate system (corresponding to $(s,t) = (0,0)$). Figure 4.1 shows typical (local) coordinate axes given by the curves when $s = 0$ and $t = 0$. Define $\phi(s,t) = f(\vec{\xi}(s,t))$. Some computations will show that $\phi_s(0,0) = \sigma f_{x_k}/\lambda_k$, $\phi_t(0,0) = \sigma f_\sigma/\lambda$, $\phi_{ss}(0,0) = (\sigma^2 f_{x_k x_k} - \rho_k^2 \sigma f_\sigma)/\lambda_k^2$, $\phi_{st}(0,0) = (\sigma^2 f_{x_k \sigma} + \sigma f_{x_k})/(\lambda_k \lambda)$, and $\phi_{tt}(0,0) = (\sigma^2 f_{\sigma\sigma} + \sigma f_\sigma)/\lambda^2$, where the derivatives of f are all evaluated at $(\vec{x}, \sigma)$. These relationships show that at the origin of the geodesic coordinates (s,t) the components of the first– and second–order covariant derivatives for f are the same as the first– and second–order partial derivatives of the function $\phi(s,t)$ at the origin. The analogy does not hold for third and higher order derivatives of ϕ since covariant differentiation of those orders is not necessarily commutative because of the effects of the curvature of the space.

4.3.4 Distance Between Points

Given two points $(\vec{x}_k, \sigma_k)$ for $k = 1, 2$, the distance between them is measured along the unique geodesic path connecting the points. Let $(\vec{x}(s), \sigma(s))$ be a parameterization by scale–space arc length of a geodesic such that $(\vec{x}_k, \sigma_k) = (\vec{x}(s_k), \sigma(s_k))$ for some s_k, $k = 1, 2$. Without loss of generality assume that $\sigma_1 \leq \sigma_2$ and $s_1 \leq s_2$. The distance between the two

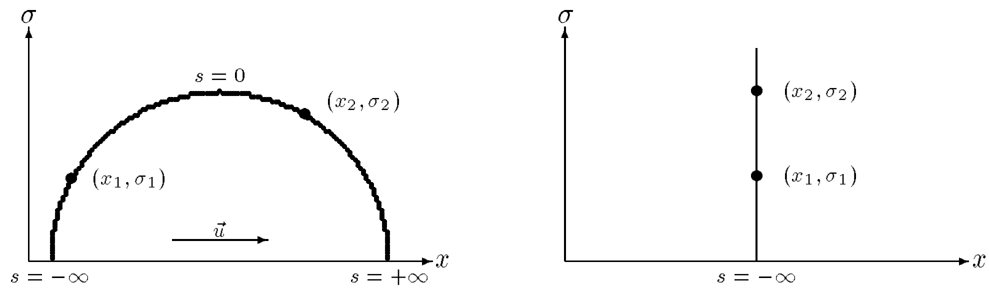

Figure 4.2: Geodesics as shortest paths

points is $s_2 - s_1$.

If the geodesic is a line in the σ direction, then $\vec{x}_1 = \vec{x}_2$ and the distance between the two points is

$$\operatorname{dist}((\vec{x}_1, \sigma_1), (\vec{x}_2, \sigma_2)) = \lambda \ln\left(\frac{\sigma_1}{\sigma_2}\right).$$

Otherwise, the geodesic connecting the two points is a half–ellipse, and the distance between the two points is derived as follows. Let $\vec{u} = P(\vec{x}_2 - \vec{x}_1)/L$ where $L = |P(\vec{x}_2 - \vec{x}_1)|$. The choice of $\vec{u}$ and the ordering of scales $\sigma_1 \le \sigma_2$ implies that $s_1 < 0$ and $s_2 > s_1$ (see Figure 4.2). From the geodesic equations, $\operatorname{sech}(s_k/\lambda) = \sigma_k/r$. Using the identity $\tanh^2(z) + \operatorname{sech}^2(z) \equiv 1$, $\tanh(s_1/\lambda) = -[1 - \operatorname{sech}^2(s_1/\lambda)]^{1/2} = -[1 - (\sigma_1/r)^2]^{1/2}$, where the minus sign is a result of $s_1 < 0$. Also, $P(\vec{x}_2 - \vec{x}_1) = r[\tanh(s_2/\lambda) - \tanh(s_1/\lambda)]\vec{u}$, which leads to $\tanh(s_2/\lambda) = L/r - \sqrt{1 - (\sigma_1/r)^2}$. Finally, it can be shown that

$$\rho = \frac{1}{r} = \frac{2L}{\sqrt{(\sigma_1^2 - \sigma_2^2)^2 + L^2[L^2 + 2(\sigma_1^2 + \sigma_2^2)]}},$$

where r is the quantity that appears in the parameterization in equation (4.2). The distance between the two points is

$$\begin{aligned} \operatorname{dist}((\vec{x}_1, \sigma_1), (\vec{x}_2, \sigma_2)) &= \lambda\left(\ln\left(\frac{\operatorname{sech}(s_2/\lambda)}{1 - \tanh(s_2/\lambda)}\right) - \ln\left(\frac{\operatorname{sech}(s_1/\lambda)}{1 - \tanh(s_1/\lambda)}\right)\right) \\ &= \lambda \ln\left(\frac{\sigma_2}{\sigma_1} \frac{1 + \sqrt{1 - (\rho\sigma_1)^2}}{1 + \sqrt{1 - (\rho\sigma_1)^2 - \rho L}}\right). \end{aligned}$$

Note that this distance formula also applies in the case $\vec{x}_1 = \vec{x}_2$ because $L = 0$. Figure 4.2 illustrates two geodesic curves and the relative position of the points on them.

EXAMPLE 4.1: Let $\lambda_1 = \lambda_2 = 1$. Define $a = 1/\sqrt{2}$ and let $(x(t), \sigma(t)) = (t, a)$ for $|t| \le a$. The scale space arc length between $(-a, a)$ and (a, a) along the specified constant scale arc is

$$\ell = \int_{-a}^{a} \frac{\sqrt{\dot{x}^2(t) + \dot{\sigma}^2(t)}}{\sigma(t)}\, dt = \int_{-a}^{a} \frac{1}{a} dt = 2.$$

Along the geodesic circular arc $(x(s), \sigma(s)) = (\tanh(s), \operatorname{sech}(s))$ between $(x_1, \sigma_1) = (-a, a)$ and $(x_2, \sigma_2) = (a, a)$, the scale space distance is

$$\delta = \left| \ln\left(\frac{a}{a} \frac{1 + \sqrt{1 - a^2} - 2a}{1 + \sqrt{1 - a^2}} \right) \right| = 2\ln(\sqrt{2} + 1) \doteq 1.763 < \ell.$$

If the same lengths were computed in Euclidean space, the circular arc would have had a larger length than the straight line segment. The reason the circular arc has smaller length than the straight line segment in scale space is that the spatial domain is in effect "less dense" at larger scales. The distance between spatial locations x_1 and x_2 at scale $\sigma = 2$ is half the distance between the same locations at scale $\sigma = 1$. Initially traversing the circular arc through increasing scale will accumulate a total distance which is smaller than that accumulated by traversing the line segment at a given scale. The north pole of the geodesic represents the point at which further increasing the scale of the path is no longer cost effective (as measured by total distance traveled). $\square$

4.3.5 Volume Integrals, Hyperspheres

Integration in scale space must also take into account the relative 1–forms that define the space. If $f(\vec{x}, \sigma)$ is a real–valued function defined on a region V, the *scale space integral of f over V* is given by

$$\int_V f(\vec{x}, \sigma) \prod_{i=1}^{n+1} \frac{\lambda_i d\xi_i}{\sigma}.$$

For example, let $\lambda_i = 1$ for all i, and consider the scale space hypersphere of radius R centered at $(\vec{x}_1, \sigma_1)$. The set consists of all points $(\vec{x}, \sigma)$ which are R units of scale space distance from the central point. The defining equation is

$$\left| \ln\left(\frac{\operatorname{sech}(s)}{1 - \tanh(s)} \frac{1 - \tanh(s_1)}{\operatorname{sech}(s_1)} \right) \right| = R,$$

where s and s_1 are the arc length parameter values for the points $(\vec{x}, \sigma)$ and $(\vec{x}_1, \sigma_1)$, respectively, along the unique geodesic containing the points. Some algebraic computation will show that this equation is equivalent to

$$|\vec{x} - \vec{x}_1|^2 + [\sigma - \sigma_1 \cosh(R)]^2 = [\sigma_1 \sinh(R)]^2,$$

which is the equation of a Euclidean hypersphere centered at $(\vec{x}_1, \sigma_1 \cosh(R))$ and whose radius is $\sigma_1 \sinh(R)$. Note that the center of the scale space hypersphere is not the same as the center of the Euclidean hypersphere. Let S_{n+1} be the set of points $(\vec{x}, \sigma)$ satifsying $|\vec{x} - \vec{x}_1|^2 + [\sigma - \sigma_1 \cosh(R)]^2 \le [\sigma_1 \sinh(R)]^2$. In scale space the volume of the hypersphere is

$$V_{\mathrm{SC}}^{(n+1)}(R) = \int_{S_{n+1}} \frac{d\vec{x}}{\sigma^n} \frac{d\sigma}{\sigma} = \frac{2\pi^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} \int_0^R \sinh^n(r)\, dr.$$

Note that the volume is independent of the center of the hypersphere, even though the hypersphere appears to be "larger" as the scale component of the center is increased. In comparison, in Euclidean space the volume of the hypersphere is

$$V_{\mathrm{eu}}^{(n+1)}(R) = \int_{S_{n+1}} d\vec{x}\, d\sigma = \frac{2\pi^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} \int_0^R r^n\, dr.$$

For the special cases $n = 1, 2, 3$, the scale space volumes are $2\pi[\cosh(R) - 1]$, $\pi[\sinh(2R) - 2R]$, and $2\pi^2[\cosh^3(R) - 3\cosh(R) + 2]/3$, respectively. For the same special cases, the Euclidean volumes are πR^2, $4\pi R^3/3$, and $2\pi^2 R^4$, respectively.

4.3.6 Curvature of Surfaces

Given an n–dimensional surface embedded in $(n+1)$–dimensional Euclidean space, *principal curvatures* and *principal directions* can be constructed at each point on the surface. I derive the analogous quantities for an n–dimensional surface embedded in $(n+1)$–dimensional Riemannian space. For a self–contained discussion, I give the constructions for Euclidean space first.

Graphs in Euclidean Space

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, say $\sigma = f(\vec{x})$, whose graph lives in Euclidean space $\mathbb{R}^n \times \mathbb{R}$. If $(\vec{x}(s), \sigma(s))$ is a curve on the graph which is parametrized by arc length, then $\sigma(s) = f(\vec{x}(s))$, $\sigma' = \vec{x}' \cdot \nabla f$ and $\sigma'' = \vec{x}'' \cdot \nabla f + \vec{x}' \cdot (\mathrm{H}f)\vec{x}'$. Unit tangent vectors to the curve are $\vec{T}(s) = (\vec{x}'(s), \sigma'(s))$. A unit normal to the graph of f is $\vec{N}(s) = (-\nabla f, 1)/\ell_N$, where $\ell_N^2 = 1 + |\nabla f|^2$. For an arbitrary vector $\vec{V} \in \mathbb{R}^n$, the unit tangents to the graph are $\vec{T} = (\vec{V}, \vec{V} \cdot \nabla f)/\ell_T$, where $\ell_T^2 = \vec{V}^{\,t}(I + \nabla f \nabla f^{\,t})\vec{V}$. The curvature of the curve in the normal section determined by $\vec{V}$ and $\vec{N}$ is given by

$$\kappa = \vec{N} \cdot \frac{d\vec{T}}{ds} = \frac{\vec{V}^{\,t}\left[\frac{1}{\ell_N}\mathrm{H}(f)\right]\vec{V}}{\vec{V}^{\,t}(I + \nabla f \nabla f^{\,t})\vec{V}}.$$

The principal curvatures κ and principal directions $\vec{V}$ are the eigenvalues and eigenvectors solving the general eigensystem

$$\left(\frac{\mathrm{H}(f)}{\sqrt{1+|\nabla f|^2}} - \kappa\left(I + \nabla f \nabla f^{\,t}\right)\right)\vec{V} = \vec{0}.$$

The matrices representing the first and second fundamental forms for the graph of f are $I + \nabla f \nabla f^{\,t}$ and $\mathrm{H}(f)/\sqrt{1+|\nabla f|^2}$, respectively. The shape operator is the matrix

$$W = \left(I + \nabla f \nabla f^{\,t}\right)^{-1} \frac{\mathrm{H}(f)}{\sqrt{1+|\nabla f|^2}},$$

so the principal curvatures and principal directions are also solutions to the regular eigensystem $W\vec{V} = \kappa\vec{V}$.

Graphs in Scale Space

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, say $\sigma = f(\vec{x})$, whose graph lives in scale space $\mathbb{R}^n \times (0, \infty)$ with metric $G = (\lambda/\sigma)^2 \mathrm{diag}(P^2, 1)$. As in the Euclidean case, a curve $(\vec{x}(s), \sigma(s))$ on the graph and which is parameterized by arc length satisfies $\sigma(s) = f(\vec{x}(s))$, $\sigma' = \vec{x}' \cdot \nabla f$ and $\sigma'' = \vec{x}'' \cdot \nabla f + \vec{x}' \cdot (\mathrm{H}f)\vec{x}'$. The unit tangent vectors are still $\vec{T}(s) = (\vec{x}'(s), \sigma'(s))$, but unit normals to the graph of f are now $\vec{N}(s) = (-P^{-2}\nabla f, 1)/\ell_N$, where $\ell_N^2 = (\lambda/\sigma)^2(1 + |P^{-1}\nabla f|^2)$. Select a constant vector $\vec{V} \in \mathbb{R}^n$ such that the unit tangents are $\vec{T} = (\vec{V}, \vec{V} \cdot \nabla f)/\ell_T$, where $\ell_T^2 = (\lambda/\sigma)^2 \vec{V}^{\,t}(P^2 + \nabla f \nabla f^{\,t})\vec{V}$. The curvature of the curve in the normal section determined by $\vec{V}$ and $\vec{N}$ is given by

$$\kappa = \vec{N} \odot \frac{\mathbb{d}\vec{T}}{\mathbb{d}s} = \frac{\vec{V}^{\,t}\left[\frac{1}{\ell_N}\mathrm{H}\left(\frac{|P\vec{x}|^2 + f^2}{2}\right)\right]\vec{V}}{\vec{V}^{\,t}(P^2 + \nabla f \nabla f^{\,t})\vec{V}}.$$

The principal curvatures κ and principal directions $\vec{V}$ are the eigenvalues and eigenvectors solving the general eigensystem

$$\left(\frac{\mathrm{H}\left(\frac{|P\vec{x}|^2+f^2}{2}\right)}{\sqrt{1+|P^{-1}\nabla f|^2}} - \kappa(P^2 + \nabla f \nabla f^{\,t})\right)\vec{V} = \vec{0}.$$

The matrices representing the first and second fundamental forms for the graph of f, as a surface embedded in scale space, are $P^2 + \nabla f \nabla f^{\,t}$ and $\mathrm{H}((|P\vec{x}|^2 + f^2)/2)/\sqrt{1 + |P^{-1}\nabla f|^2}$, respectively. The shape operator is the matrix

$$W = \left(P^2 + \nabla f \nabla f^{\,t}\right)^{-1} \frac{\mathrm{H}\left(\frac{|P\vec{x}|^2 + f^2}{2}\right)}{\sqrt{1 + |P^{-1}\nabla f|^2}},$$

so the principal curvatures and principal directions are also solutions to the regular eigensystem $W\vec{V} = \kappa\vec{V}$.

Implicitly Defined Surfaces in Euclidean Space

Let $F : \mathbb{R}^n \times (0,\infty) \to \mathbb{R}$ be a function whose level surfaces implicitly defined by $F(\vec{\xi}) = c$ live in Euclidean space $\mathbb{R}^n \times (0,\infty)$. If $\vec{\xi}(s)$ is a curve parametrized by arc length and lives on the level surface defined by $F \equiv 0$, then $F(\vec{\xi}(s)) \equiv 0$. The unit tangent vectors are $\vec{T}(s) = \vec{\xi}'(s)$. Unit scale space length normals are given by $\vec{N}(s) = \hat{\nabla} F(\vec{\xi}(s))/|\hat{\nabla} F(\vec{\xi}(s)|$. Assume that the curve is such that $\vec{T}$ and $\vec{N}$ determine a normal section. The curvature of the curve is

$$\kappa = \vec{N} \cdot \frac{d\vec{T}}{ds} = -\vec{T} \cdot \frac{d\vec{N}}{ds} = -\vec{T} \cdot \frac{d\vec{N}}{d\vec{\xi}}\,\vec{T} = -\frac{\vec{T}^{\,t}\frac{d\vec{N}}{d\vec{\xi}}\vec{T}}{\vec{T}^{\,t}\vec{T}},$$

where the identities $\vec{T} \cdot \vec{N} \equiv 0$ and $\vec{T} \cdot \vec{T} \equiv 1$ were used. The principal curvatures κ and principal directions $\vec{T}$ are the eigenvalues and tangential eigenvectors solving

$$-\frac{d\vec{N}}{d\vec{\xi}}\,\vec{T} = \kappa\vec{T}.$$

The matrix $-d\vec{N}/d\vec{\xi}$ represents the shape operator as an operation applied to the tangent spaces in the ambient $(n+1)$–dimensional space, as compared to the usual representation as an operator on the n–dimensional tangent spaces.

Implicitly Defined Surfaces in Scale Space

Let $F : \mathbb{R}^n \times (0,\infty) \to \mathbb{R}$ be a function whose level surfaces implicitly defined by $F(\vec{\xi}) = c$ live in scale space $\mathbb{R}^n \times (0,\infty)$ with metric $G = (\lambda/\sigma)^2 \mathrm{diag}(P^2, 1)$. As in the Euclidean case, a curve $\vec{\xi}(s)$ parametrized by arc length and living on the level surface defined by $F \equiv 0$ satisfies $F(\vec{\xi}(s)) \equiv 0$. The unit tangent vectors are still $\vec{T}(s) = \vec{\xi}'(s)$. Unit normals are now given by $\vec{N}(s) = \nabla F(\vec{\xi}(s))/\|\nabla F(\vec{\xi}(s)\|$, where the length calculation is with respect to the metric. Assume that the curve is such that $\vec{T}$ and $\vec{N}$ determine a normal section. The curvature of the curve is

$$\kappa = \vec{N} \odot \frac{\mathrm{d}\!\mathrm{l}\vec{T}}{\mathrm{d}\!\mathrm{l}s} = -\vec{T} \odot \frac{\mathrm{d}\!\mathrm{l}\vec{N}}{\mathrm{d}\!\mathrm{l}s} = -\vec{T} \odot \frac{\mathrm{d}\!\mathrm{l}\vec{N}}{\mathrm{d}\!\mathrm{l}\vec{\xi}}\,\vec{T} = -\frac{\vec{T}^{\,t}\,\mathrm{d}\!\mathrm{l}\vec{N}/\mathrm{d}\!\mathrm{l}\,\vec{\xi}\vec{T}}{\vec{T}^{\,t}\vec{T}},$$

where the identities $\vec{T} \odot \vec{N} \equiv 0$ and $\vec{T} \odot \vec{T} \equiv 1$ were used. The principal curvatures κ and principal directions $\vec{T}$ are the eigenvalues and tangential eigenvectors solving

$$-\frac{\mathrm{d}\!\mathrm{l}\vec{N}}{\mathrm{d}\!\mathrm{l}\vec{\xi}}\,\vec{T} = \kappa\vec{T}.$$

Note that the development for scale space is identical to that for Euclidean space, except that differentiation of tangents and normals is replaced by covariant differentiation.

4.4 Ridges in Scale Space

Representation of object shape by medial structures is an important aspect of image analysis. For 2–dimensional binary images, the Blum medial axis (Blum & Nagel 1978) encodes the shape information in the form of centers and radii of maximal disks contained in the objects. For 3–dimensional binary images, medial surfaces encode shape information in the same way. Construction of medial structures for objects in gray scale images has also been considered by researchers. For example, in many types of medical images, the intensity tends to be bright at pixels which are centrally located in objects and tends to be dark near boundaries of objects. The bright centers of objects show up as ridges on the graph of intensity. The projection of the ridges onto the image plane may be treated as medial structures.

In Chapter 2, various definitions for ridges of n–dimensional images are given, but they were formulated for graphs or hypersurfaces which are embedded in Euclidean space. As such, the ridge structures obtained from any one of the definitions do not provide information about object size. The absence of a scale parameter in the process does not allow measurement of global information such as object width. The definitions are, however, suitable for $(n+0.5)$–dimensional applications (for example, see Chapter 3, (Lindeberg 1993a)) where a sequence of increasingly Gaussian–blurred images is built from the initial image. The ridge structures on the graph of each blurred image give information about object widths where the widths are proportional to the scale (standard deviation of the Gaussian) of measurement, but no measurements through scale are made.

A true $(n+1)$–dimensional analysis (n spatial variables and 1 scale variable) should use the scale space geometry I have described here. A general discussion of cores (formerly called multiscale medial axes) is found in Pizer, Burbeck, Coggins, Fritsch & Morse (1992). Specific approaches to constructing cores from gray scale images using scale space ideas can be found in Eberly, Fritsch & Kurak (1992), Fritsch (1993), and Morse, Pizer & Liu (1993). The main idea in each of these papers is to apply a filter to the image to build a function $M(\vec{x}, \sigma)$ which measures the "medialness" of a position $\vec{x}$ relative to object boundaries (if any) located at a "distance" $\sigma > 0$ away from the position and using boundariness measurements relative to the aperture size σ. In Eberly et al. (1992) and Fritsch (1993), functions of the form $\sigma = f(\vec{x})$ are constructed by requiring at each position $\vec{x}$ that $\Theta(M)(\vec{x}, \sigma) = 0$ for some appropriate differential operator Θ. Cores are then derived as ridges on the graph of f as a surface in scale space. In Morse et al. (1993), the medialness measurements are iteratively refined using a Hough–like transformation to obtain a function $F(\vec{x}, \sigma)$. Medial structures are derived directly as ridges of the function values for F. In all cases, the extraction of medial structures was based on geometric methods applied to scale space with only the Euclidean metric.

I extend the (Euclidean) ridge definitions in Chapter 2 to ones that apply in scale space $\mathcal{S}$ with metric G. The height definition and level definition are suitable for real–valued functions $F(\vec{x}, \sigma)$. Each definition depends only on the metric associated with the domain space of the function, not on the metric associated with the space in which the graph of F lives. Thus, for functions $\sigma = f(\vec{x})$, the height and level definitions are exactly those given in Chapter 2. The principal direction definition is suitable for functions $\sigma = f(\vec{x})$ where the graph of f is contained in (Riemannian) scale space.

4.4.1 Height Ridges

I briefly review the concepts of *relative extreme points* and *height ridges and valleys* for functions $f : \mathbb{R}^n \to \mathbb{R}$, where $\mathbb{R}^n$ has the usual Euclidean metric. The extension of the ideas is then given for functions $f : \mathcal{S} \to \mathbb{R}$ where $\mathcal{S}$ has metric G.

Euclidean Space

Let $\vec{V}_1, \ldots, \vec{V}_d$ be a set of constant linearly independent vectors in $\mathbb{R}^n$, where $1 \leq d \leq n$. Let V be the $n \times d$ matrix whose columns are the given vectors. The function has a *relative minimum (maximum) of type $n-d$* at $\vec{x}$ if $V^t \nabla f(\vec{x}) = 0$ and $V^t \mathrm{H}(f(\vec{x}))V$ is positive (negative) definite. Such points $\vec{x}$ are called *relative extreme points of type $n-d$ for f with respect to V*. When $d = n$, the classification is the usual one for extreme points. Generally, the solution sets of $V^t \nabla f = 0$ are $(n-d)$–dimensional manifolds since there are d equations in n unknowns; hence the use of "type $n-d$" in the definition.

The condition $V^t \nabla f = 0$ is a first derivative test. The test for definiteness of $V^t(\mathrm{H}f)V$ is a second derivative test. Both tests can be phrased in terms of directional derivatives computed on a special hypersurface which is a subset of the domain of f. Without loss of generality, let the $\vec{V}_k$ be orthonormal vectors. Let $\vec{s} \in \mathbb{R}^d$ and define $\phi(\vec{s}) = f(\vec{x} + V\vec{s})$; then

$$\phi(\vec{0}) = f(\vec{x}), \quad \nabla\phi(\vec{0}) = V^t \nabla f(\vec{x}) \quad \text{and} \quad \mathrm{H}\phi(\vec{0}) = V^t \mathrm{H} f(\vec{x}) V.$$

A point $\vec{x}$ is a relative extreme point of type $n-d$ for f with respect to V if and only if $\vec{0}$ is an extreme point (in the usual sense) for the function $\phi(\vec{s})$.

It is important to note that the first–order directional derivatives are independent of the choice of parameterization of the domain of f, but the second–order directional derivatives are dependent on that choice. Generally, let $\vec{y}(\vec{s})$ be a parameterized hypersurface contained in the domain of f, $\vec{y} : \mathbb{R}^d \to \mathbb{R}^n$, such that $\vec{y}(\vec{0}) = \vec{x}$ and $d\vec{y}(\vec{0})/d\vec{s} = V$. Define $\phi(\vec{s}) = f(\vec{y}(\vec{s}))$; then

$$\phi(\vec{0}) = f(\vec{x}) \quad \text{and} \quad \nabla\phi(\vec{0}) = V^t \nabla f(\vec{x}),$$

but

$$\mathrm{H}\phi(\vec{0}) = V^t \mathrm{H} f(\vec{x}) V + \frac{d^2\vec{y}(\vec{0})}{d\vec{s}^2} \nabla f(\vec{x}),$$

where $d^2\vec{y}(\vec{0})/d\vec{s}^2$ is a triply–indexed quantity $(d \times d \times n)$ and its product with the $n \times 1$ vector ∇f is a contraction on the index with common range n. For the special case $\vec{y}(\vec{s}) = \vec{x} + V\vec{s}$, the only term on the right–hand side is the Hessian of f (as seen earlier). But in general, the second term on the right is not identically zero, which means that the second derivative calculations for ϕ depend on more than just the vector directions V at $\vec{s} = \vec{0}$.

The height definition for ridges and valleys in Euclidean space is described now. Let κ_i and $\vec{V}_i$, $1 \le i \le n$, be the eigenvalues and eigenvectors for the matrix $-\mathrm{H}f$. Assume that the eigenvalues are ordered as $\kappa_1 \ge \cdots \ge \kappa_n$. Since $-\mathrm{H}f$ is symmetric, the eigenvectors are orthonormal.

- A point $\vec{x}$ is a *ridge point of type $n-d$* if $\kappa_d(\vec{x}) > 0$ and $\vec{x}$ is a relative maximum point of type $n-d$ for f with respect to $V = [\vec{V}_1|\cdots|\vec{V}_d]$. Since $-V^t(\mathrm{H}f)V = \mathrm{diag}\{\kappa_1|\vec{V}_1|^2, \ldots, \kappa_d|\vec{V}_d|^2\}$ and the eigenvalues are ordered, the test for a ridge point reduces to $V^t\nabla f(\vec{x}) = 0$ and $\kappa_d(\vec{x}) > 0$. A ridge point $\vec{x}$ is a *strong ridge point* if $\kappa_d(\vec{x}) > |\kappa_n(\vec{x})|$; otherwise, it is a *weak ridge point.*

- A point $\vec{x}$ is a *valley point of type $n-d$* if $\kappa_{n-d+1}(\vec{x}) < 0$ and $\vec{x}$ is a relative maximum point of type $n-d$ for f with respect to $V = [\vec{V}_{n-d+1}|\cdots|\vec{V}_n]$. Since $-V^t(\mathrm{H}f)V = \mathrm{diag}\{\kappa_{n-d+1}|\vec{V}_{n-d+1}|^2, \ldots, \kappa_n|\vec{V}_n|^2\}$ and the eigenvalues are ordered, the test for a valley point reduces to $V^t\nabla f(\vec{x}) = 0$ and $\kappa_{n-d+1}(\vec{x}) < 0$. A valley point $\vec{x}$ is a *strong valley point* if $\kappa_{n-d+1}(\vec{x}) > |\kappa_1(\vec{x})|$; otherwise, it is a *weak valley point.*

Scale Space

The generalization of the relative extreme point definitions to scale space is technically somewhat more complicated, but intuitively, the idea of relative extreme points for a function $f : \mathcal{S} \to \mathbb{R}$ is clear. In Euclidean space, I restricted f to a hypersurface $\vec{y}(\vec{s}) = \vec{x} + V\vec{s}$, say $\phi(\vec{s}) = f(\vec{x} + V\vec{t})$, and computed extreme points for ϕ. The selected hyperspace is flat in the sense that $d^2\vec{y}(\vec{0})/d\vec{s}^2 = 0$. Moreover, since $|\vec{V}_i| = 1$, as you walk in that direction along the hypersurface, the induced parametrization of the path is one of arc length.

In scale space, I restrict f to a "flat" hypersurface $\vec{y}(\vec{s})$ and compute extreme points for the restricted function $\phi(\vec{s})$. The concept of flatness depends on the metric for scale space; the flat hypersurfaces will lie on hyperellipsoids with center on the $\sigma = 0$ hyperplane. The Hessian that occurs in the second derivative test for ϕ will be the scale space Hessian.

More specifically, for $1 \le k \le d$ let $\vec{V}_k = (\vec{W}_k, \gamma_k) \in \mathbb{R}^{n+1}$ be orthonormal vectors: $\vec{V}_i \odot \vec{V}_j = \delta_{ij}$. Define the flat hypersurface $\vec{\xi} : \mathbb{R}^d \to \mathcal{S}$ by

$$\vec{\xi}(\vec{s}) = \left(\vec{x} + \sum_{i=1}^{d} \alpha_i(\vec{s})\vec{W}_i, \sigma\beta(\vec{s})\right),$$

where $|\mathrm{diag}(P,1)[\vec{\xi}(\vec{s})-(\vec{c},0)]| = r$ for some constant $\vec{c} \in \mathbb{R}^n$ and constant $r > 0$, and where $\alpha_i(\vec{0}) = 0$ and $\beta(\vec{0}) = 1$. Thus, $\vec{\xi}(\vec{0}) = (\vec{x},\sigma)$. Also, I require that $\partial\vec{\xi}(\vec{0})/\partial s_i = \vec{V}_i$, a unit length vector with respect to the scale space metric. Consequently, $\partial\alpha_i(\vec{0})/\partial s_j = \delta_{ij}$ and $\partial\beta(\vec{0})/\partial s_i = \gamma_i/\sigma$. Differentiating $|\mathrm{diag}(P,1)[\vec{\xi}-(\vec{c},0)]| = r$ with respect to s_k and s_ℓ, and then evaluating at $\vec{0}$, yields

$$\sigma\left(\frac{\partial^2\beta(\vec{0})}{\partial s_k \partial s_\ell} + \frac{\delta_{k\ell}}{\lambda^2}\right) = \sum_{i=1}^{d} \gamma_i \frac{\partial^2 \alpha_i(\vec{0})}{\partial s_k \partial s_\ell}. \tag{4.3}$$

Define $\phi(\vec{s}) = f(\vec{\xi}(\vec{s}))$. Taking a derivative and evaluating at $\vec{0}$ yields

$$\frac{\partial\phi(\vec{0})}{\partial s_k} = \vec{V}_k \cdot \nabla f(\vec{x},\sigma) = \vec{T}_k \cdot \widehat{\nabla} f(\vec{x},\sigma),$$

where $\vec{T}_k = \sqrt{G}\vec{V}_k$ and $\widehat{\nabla} f = \sqrt{G^{-t}}\nabla f$. Taking a second derivative and evaluating at $\vec{0}$ yields

$$\frac{\partial^2\phi(\vec{0})}{\partial s_k \partial s_\ell} = \vec{V}_k^{\,t}\left(\mathrm{H}f(\vec{x},\sigma)\right)\vec{V}_\ell + \frac{\partial^2\vec{\xi}(\vec{0})}{\partial s_k \partial s_\ell} \cdot \nabla f(\vec{x},\sigma).$$

The second term on the right–hand side of the previous equation needs to be evaluated. Let $(\vec{x},\sigma)$ be a point such that $\vec{V}_i \cdot \nabla f(\vec{x},\sigma) = 0$ for all $i = 1,\ldots,d$. Let $\vec{N}_j = (\vec{U}_j,\eta_j)$, $1 \le j \le n+1-d$, be vectors such that $\vec{V}_i \cdot \vec{N}_j = 0$ for all valid i and j, and $\vec{N}_i \cdot \vec{N}_j = \delta_{ij}$. The gradient of f therefore can be expressed in this basis as

$$\nabla f(\vec{x},\sigma) = \sum_{i=1}^{n-d+1} \mu_i \vec{N}_i.$$

Also $\vec{V}_i \cdot \vec{N}_j = 0$ implies that $\vec{W}_i \cdot \vec{U}_j = -\gamma_i\eta_j$. At $\vec{s} = \vec{0}$ for each j,

$$\begin{aligned}
\tfrac{\partial^2\vec{\xi}(\vec{0})}{\partial s_k \partial s_\ell} \cdot \vec{N}_j &= \textstyle\sum_{i=1}^{d} \tfrac{\partial^2\alpha_i(\vec{0})}{\partial s_k \partial s_\ell}\vec{W}_i \cdot \vec{U}_j + \sigma\tfrac{\partial^2\beta(\vec{0})}{\partial s_k \partial s_\ell}\eta_j \\
&= \eta_j\left(\sigma\tfrac{\partial^2\beta(\vec{0})}{\partial s_k \partial s_\ell} - \textstyle\sum_{i=1}^{d}\gamma_i\tfrac{\partial^2\alpha_i(\vec{0})}{\partial s_k \partial s_\ell}\right) \\
&= -\tfrac{\sigma}{\lambda^2}\eta_j\delta_{k\ell},
\end{aligned}$$

where I have used equation (4.3). Therefore,

$$\frac{\partial^2\vec{\xi}(\vec{0})}{\partial s_k \partial s_\ell} \cdot \nabla f = \left(-\frac{\sigma}{\lambda^2}\sum_{j=1}^{n-d+1}\mu_j\eta_j\right)\delta_{k\ell} = -\frac{\sigma}{\lambda^2} f_\sigma \delta_{k\ell}.$$

Replacing this in the equation for the second derivatives of ϕ yields the following formula:

$$\frac{\partial^2\phi(\vec{0})}{\partial s_k \partial s_\ell} = \vec{V}_k^{\,t}\left(\mathrm{H}f(\vec{x},\sigma)\right)\vec{V}_\ell - \frac{\sigma}{\lambda^2} f_\sigma(\vec{x},\sigma)\delta_{k\ell} = \vec{T}_k^{\,t}\widehat{H}f(\vec{x},\sigma)\vec{T}_\ell.$$

The last equality follows by a direct application of the definitions for $\vec{T}_k$ and the scale space Hessian matrix.

The definitions for relative extreme points in scale space is given by the following. Let $\vec{V}_1, \ldots, \vec{V}_d$ be a set of constant (Euclidean) orthonormal vectors in $\mathbb{R}^n \times \mathbb{R}$ where $1 \leq d \leq n+1$. Let V be the $(n+1) \times d$ matrix whose columns are the given vectors. The function $f : \mathcal{S} \to \mathbb{R}$ has a *relative minimum (maximum) of type* $n+1-d$ at $\vec{\xi}$ if $V^{\,t}\widehat{\nabla} f(\vec{\xi}) = 0$ and $V^{\,t}\widehat{H}(f(\vec{\xi}))V$ is positive (negative) definite. This definition is identical in structure to that of the Euclidean one, except that the Euclidean gradient and Hessian are replaced by the scale space gradient and Hessian.

The generalization of the height ridge/valley definitions to scale space is straightforward now that I have defined relative extreme points for functions $f : \mathcal{S} \to \mathbb{R}$. Let κ_i and $\vec{V}_i$, $1 \leq i \leq n+1$, be the eigenvalues and eigenvectors for the matrix $-\widehat{H}f$. Assume that the eigenvalues are ordered as $\kappa_1 \geq \cdots \geq \kappa_{n+1}$.

- A point $\vec{\xi}$ is a *ridge point of type* $n+1-d$ if $\kappa_d(\vec{\xi}) > 0$ and $\vec{\xi}$ is a relative maximum point of type $n+1-d$ for f with respect to $V = [\vec{V}_1 \cdots \vec{V}_d]$. The test for a ridge point is equivalent to $V^t\widehat{\nabla} f(\vec{\xi}) = 0$ and $\kappa_d(\vec{\xi}) > 0$.

- A point $\vec{\xi}$ is a *valley point of type* $n+1-d$ if $\kappa_{n-d+2}(\vec{\xi}) < 0$ and $\vec{\xi}$ is a relative maximum point of type $n+1-d$ for f with respect to $V = [\vec{V}_{n-d+2} \cdots \vec{V}_{n+1}]$. The test for a valley point is equivalent to $V^t\widehat{\nabla} f(\vec{\xi}) = 0$ and $\kappa_{n-d+2}(\vec{\xi}) < 0$.

The additional classifications as strong or weak points still hold.

4.4.2 Principal Direction Ridges

I briefly review the principal direction definition for ridges and valleys on the graph of functions $f : \mathbb{R}^n \to (0, \infty)$ and on level surfaces of functions $F : \mathbb{R}^n \times (0, \infty) \to \mathbb{R}$, where $\mathbb{R}^n \times (0, \infty)$ has the Euclidean metric. The extension of the ideas is then given for graphs of functions $f : \mathbb{R}^n \to (0, \infty)$ and for level surfaces of functions $F : \mathbb{R}^n \times (0, \infty) \to \mathbb{R}$ where $\mathbb{R}^n \times (0, \infty)$ is scale space with metric G.

Euclidean Space

Consider functions $f : \mathbb{R}^n \to (0, \infty)$, say $\sigma = f(\vec{x})$. The matrices representing the first and second fundamental forms are $A = I + \nabla f \nabla f^t$ and $B = \mathrm{H}(f)/\sqrt{1 + |\nabla f|^2}$, respectively. The shape operator is $W = A^{-1}B$. The principal curvatures κ_i and principal directions $\vec{V}_i$, $1 \leq i \leq n$, are the eigenvalues and eigenvectors of W. Assume that the eigenvalues are ordered as $\kappa_1 \geq \cdots \geq \kappa_n$, and assume that d is such that $1 \leq d \leq n$. The principal direction definition for ridges and valleys is formulated so that such points correspond to local extrema of the principal curvatures as the surface is traversed along integral curves of the principal directions.

- The point $\vec{x}$ is a *ridge point of type* $n-d$ if $\kappa_d(\vec{x}) > 0$, $D_{\vec{V}_i}\kappa_i(\vec{x}) = 0$, and $D_{\vec{V}_i}D_{\vec{V}_i}\kappa_i(\vec{x}) < 0$ for $1 \leq i \leq d$. Additionally $\vec{x}$ is a *strong ridge point* if $\kappa_d(\vec{x}) > |\kappa_n(\vec{x})|$; otherwise it is a *weak ridge point.*

- The point $\vec{x}$ is a *valley point of type* $n-d$ if $\kappa_{n-d+1}(\vec{x}) < 0$, $D_{\vec{V}_i(x)}\kappa_i(\vec{x}) = 0$, and $D_{\vec{V}_i(x)}D_{\vec{V}_i(x)}\kappa_i(\vec{x}) > 0$ for $n - d + 1 \leq i \leq n$. Additionally $\vec{x}$ is a *strong valley point* if $\kappa_{n-d+1}(\vec{x}) > |\kappa_1(\vec{x})|$; otherwise it is a *weak valley point.*

For a vector field $\vec{V}(\vec{x})$ and real–valued function $\kappa(\vec{x})$, the indicated directional derivatives are defined as

$$D_{\vec{V}}\kappa(\vec{x}) = \vec{V}(\vec{x})^t \nabla\kappa(\vec{x}) \quad \text{and} \quad D_{\vec{V}}D_{\vec{V}}\kappa(\vec{x}) = \vec{V}^t \mathrm{H}(\kappa(\vec{x}))\vec{V} + \nabla\kappa(\vec{x})^t \, \frac{d\vec{V}}{d\vec{x}} \, \vec{V}.$$

Now consider functions $F : \mathbb{R}^n \times (0, \infty) \to \mathbb{R}$. Consider the hypersurface implicitly defined by $F(\vec{\xi}) = 0$. The principal curvatures κ_i and principal directions $\vec{V}_i$ are the eigenvalues and tangential eigenvectors of the matrix $W = -d\vec{N}/d\vec{x}$ where $\vec{N} = \nabla F/|\nabla F|$. The ridge and valley definitions are identical to the ones given above, except now the curvatures, directions, and derivatives depend on $\vec{\xi} \in \mathbb{R}^n \times (0, \infty)$ as compared to the previous definition where they depended on $\vec{x} \in \mathbb{R}^n$.

Scale Space

As pointed out in the section on scale space measurements, the principal curvatures and principal directions for graphs of $\sigma = f(\vec{x})$ and implicit surfaces defined by $F(\vec{x}, \sigma) = c$ are computed exactly as in the Euclidean setting, except that the Euclidean shape operators are replaced by their scale space counterparts, which depend on covariant differentiation.

For the graph of $\sigma = f(\vec{x})$, the matrices representing the first and second fundamental forms are $A = P^2 + \nabla f \nabla f^t$ and $B = \mathrm{H}((|P\vec{x}|^2 + f^2)/2)/\sqrt{1 + |P^{-1}\nabla f|^2}$, respectively. The shape operator is $W = A^{-1}B$. For the hypersurface defined by $F(\vec{x}, \sigma) = 0$, the shape operator is $W = -\mathrm{d\!l}\vec{N}/\mathrm{d\!l}\vec{x}$ where $\vec{N} = \widehat{\nabla} F/\|\widehat{\nabla} F\|$. The scale space principal direction definitions are identical to those for Euclidean space, except that the scale space shape operators are used in place of the Euclidean ones.

4.4.3 Level Ridges

The Euclidean level definition for ridges and valleys for a function $f : \mathbb{R}^n \to \mathbb{R}$, $n \geq 2$, involved applying the principal direction definition to finding ridges and valleys for *all* of the level surfaces $f(\vec{x}) = c$. The set of all principal direction ridges and valleys for the level surfaces makes up the level ridges and valleys.

The extension to a scale space level definition for ridges and valleys is also straightforward. In the Euclidean definition, for each level surface defined by $f(\vec{x}) = c$, the shape operator for the surface is $W = -d\vec{N}/d\vec{x}$ where $\vec{N} = \nabla f/|\nabla f|$. In scale space, the shape operator is $W = -\mathrm{d\!l}\vec{N}/\mathrm{d\!l}\vec{x}$, where $\vec{N} = \widehat{\nabla} F/\|\widehat{\nabla} F\|$.

4.4.4 Invariance Properties

The ridges constructed by the level definition for functions $f : \mathbb{R}^n \to \mathbb{R}$, where $\mathbb{R}^n$ has the Euclidean metric, are invariant under rotations, reflections, translations, and uniform magnifications in $\mathbb{R}^n$. That is, if ridges$(f(\vec{x})) \subset \mathbb{R}^n$ denotes the set of ridges for f in coordinates $\vec{x}$, and if $\vec{\psi}(\vec{x})$ represents a rotation, reflection, translation, or magnification $\vec{x} \to \mu\vec{x}$, then $\vec{\psi}($ridges$(f(\vec{x})) =$ ridges$(f(\vec{\psi}(\vec{x})))$. The level ridges are also invariant with respect to monotonic transformations of f, ridges$(g \circ f(\vec{x}))$ = ridges$(f(\vec{x}))$, where $g' > 0$. The ridges constructed by the height and principal direction definition are also invariant under rotations, reflections, and translations, but not with respect to uniform magnifications in $\mathbb{R}^n$.

The scale space level ridge definitions for functions $f : \mathcal{S} \to \mathbb{R}$ produces ridges which are invariant under rotations, reflections, and translations in the spatial components. The ridges are additionally invariant to zoom, $(\vec{x}, \sigma) \to \mu(\vec{x}, \sigma)$, for $\mu > 0$. The level ridges are invariant with respect to monotonic transformations in f, as in the Euclidean case. The scale space height and principal direction definition for functions $f : \mathbb{R}^n \to (0, \infty)$ also produces ridges which are invariant to rotations, reflection, and translations in the spatial components. As in the Euclidean case, the ridges are not invariant to uniform spatial magnifications, but they are invariant to zoom (uniform magnification in both space and scale).

4.5 Anisotropic Diffusion as a Consequence of the Metric

The purpose of this section is to show that the generation of multiscale data via anisotrophic diffusion is intimately related to the geometry of the underlying scale space. In particular, I will show that anisotropic diffusion is in a loose sense "linear" diffusion in non–Euclidean space. This is in contrast to the view that anisotropic diffusion is a nonlinear diffusion in Euclidean space. Moreover, the selection of the metric for scale space based on invariance requirements for a front–end vision system is a more natural approach to solving vision problems. Once the metric is selected, the anisotropic diffusion process for generating the multiscale data is automatically determined. I conclude with some ideas on how the metric may be useful in image analysis and in developing stable numerical algorithms when discretizing the diffusion equation.

4.5.1 Linear Diffusion in Euclidean Space

The simplest diffusion process is, of course, given by

$$u_\sigma = \nabla \cdot (\nabla u), \; x \in \mathbb{R}^n, \; \sigma > 0.$$

As a heat transfer process, the conductance is a constant ($c \equiv 1$). No preference is given for direction of transfer (rotational invariance) or for location of origin (translational invariance). The units of measurement are significant here (no scale invariance). For example, if space and scale are transformed by $(x, \sigma) \to \lambda(x, \sigma)$ and if $v(x, \sigma) = u(\lambda x, \lambda\sigma)$, then $v_\sigma = \lambda^{-1}\nabla \cdot (\nabla v)$. To retain invariance, the temperatures must also be transformed by $u \to \lambda v$.

The standard analysis of the linear diffusion equation assumes that the underlying space is Euclidean. That is, the metric of the space is determined by the form for arc length

$$ds^2 = dx \cdot dx + d\sigma^2,$$

and the total derivative is given by

$$du = \nabla u \cdot dx + u_\sigma d\sigma.$$

Any measurements in the (x, σ) space are independent of the multiscale data $u(x, \sigma)$. For example, distance between points in scale space are measured using the standard formula for Euclidean distance. Objects in the initial image $I(x)$ are analyzed essentially independently of the image intensities and of the proximity of other objects.

The standard finite difference scheme (forward difference in scale, central difference in space) in solving the linear diffusion equation uses a grid consisting of a rectangular lattice of points. In one spatial dimension, the approximations

$$u_\sigma \doteq \frac{u(x, \sigma + k) - u(x, \sigma)}{k}$$

and

$$u_{xx} \doteq \frac{u(x+h, \sigma) - 2u(x, \sigma) + u(x-h, \sigma)}{h^2}$$

lead to the difference scheme

$$u(x, \sigma + k) = u(x, \sigma) + \frac{k}{h^2}\left[u(x+h, \sigma) - 2u(x, \sigma) + u(x-h, \sigma)\right].$$

For n spatial variables, the scheme is stable if $2nk < h^2$. For a given scale, a spatial grid point is always a constant (Euclidean) distance h from any other neighbors at the same scale. For a given spatial location, a scale grid point is always a constant (Euclidean) distance k from its neighbors at adjacent scales.

4.5.2 Linear Diffusion in Non–Euclidean Space

In addition to requiring rotational and translational invariance in space, a front–end vision system might be required to have invariance with respect to units of measurement. A metric which has all three invariances is determined by the arc length form

$$ds^2 = \frac{dx \cdot dx}{\sigma^2} + \frac{d\sigma^2}{\sigma^2}.$$

Consequently scale space becomes non–Euclidean (the geometry is hyperbolic). The total derivative is given by

$$du = \sigma \nabla u \cdot \frac{dx}{\sigma} + \sigma u_\sigma \frac{d\sigma}{\sigma}.$$

Any measurements in the (x, σ) space are still independent of the multiscale data $u(x, \sigma)$, but there is interaction between the space and scale variables. Geodesics are now semicircles with centers at $\sigma = 0$ as compared to lines, which are the geodesics in Euclidean space. Objects in the initial image $I(x)$ are still analyzed independently of the image intensities and of the proximity of other objects, but the dependence on units of measurement has been removed.

The linear diffusion process corresponding to this metric is given by

$$\sigma u_\sigma = \sigma^2 \nabla \cdot (\nabla u), \; x \in \mathrm{I\!R}^n, \; \sigma > 0.$$

In Euclidean space, if the diffusion scale is t $(u_t = u_{xx})$, then the relationship to the scale in this non–Euclidean space is $t = \sigma^2/2$. However, now the derivatives are in a unitless form due to the scale invariance of the metric. Note that the left–hand side of the diffusion equation is a single application of the scale space σ derivative (as specified in the total derivative formula for du) and the right–hand side of the equation is two applications of the scale space x gradient operator: $(\sigma\partial/\partial\sigma)u = (\sigma\nabla)^2 u$.

A finite difference scheme for the one spatial dimension case may be used to solve the diffusion equation in the current setting. The extension to more spatial dimensions is apparent. The following approximations are used,

$$\sigma u_\sigma \doteq \frac{u(x, b\sigma) - u(x, \sigma)}{\ln b},$$

and

$$\sigma^2 u_{xx} \doteq \frac{u(x + h\sigma, \sigma) - 2u(x, \sigma) + u(x - h\sigma, \sigma)}{h^2},$$

where $b > 1$. The difference scheme is

$$u(x, b\sigma) = u(x, \sigma) + \frac{\ln b}{h^2} \left[u(x + h\sigma, \sigma) - 2u(x, \sigma)) + u(x - h\sigma, \sigma) \right].$$

For n spatial variables with the same pixel spacing $h > 0$, the scheme is stable as long as $2n \ln b < h^2$. Note that this finite difference scheme will get you to a larger scale more quickly than the one using the Euclidean metric, since the scale parameter increases as a geometric sequence rather than as an arithmetic sequence.

Now notice that the implied grid of points is no longer a "rectangular" grid as in the Euclidean case. The implied sampling in scale requires us to use a geometric sequence of scales. As scale increases, the implied spatial samples are sparsely placed as compared to the placement at small scale. A closer look shows that in fact the implied grid points are a "constant" distance apart, but now distance is measured with respect to the metric. For example, the distance between $(x, b\sigma)$ and (x, σ) is $\ln b$, a constant distance in this non–Euclidean space. For a given scale σ, the distance between $(x + h\sigma, \sigma)$ and (x, σ) is h units, again a constant.

4.5.3 Anisotropic Diffusion

The last two sections have a common theme. The linear diffusion equation is related to the metric assigned to the space. In both cases, the left–hand side of the diffusion is one application to u of the scale derivative which is natural to the metric. The right–hand side is two applications to u of the spatial gradient which is natural to the metric. In the Euclidean metric case, the diffusion is

$$\left(\frac{\partial}{\partial\sigma}\right) u = (\nabla)^2\, u$$

and in the non–Euclidean metric case, it is

$$\left(\sigma\frac{\partial}{\partial\sigma}\right) u = (\sigma\nabla)^2\, u.$$

More generally, anisotropic diffusion can be viewed as the multiscale process one must apply given an appropriate metric for scale space. Let $c()$ denote the conductance function for anisotropic diffusion. The conductance may be a function of space, scale, or image data (and its derivatives). Let $\rho()$ denote the density function which also appears in the more general model of heat transfer:

$$\frac{\partial u}{\partial\sigma} = \frac{1}{\rho}\nabla\cdot(c\nabla u)\,.$$

In the Perona–Malik model (Perona & Malik 1987), no density term is included, so $\rho \equiv 1$. Some candidates for the conductance are

$$c(|\nabla u|) = \exp(-|\nabla u|^2/k^2) \text{ or } c(|\nabla u|) = \left(1 + |\nabla u|^2/k^2\right)^{-1}$$

for some parameter $k > 0$. Of course this makes the diffusion process nonlinear. The idea is that "homogeneous regions" (where gradient intensity is small) have large conductance and are diffused significantly, but "edge regions" (where gradient intensity is large) have small conductance and are not diffused much. In this way the image is smoothed to preserve noticeable boundaries.

The conductance function has been viewed as a "stretching" of space. This notion is made formal by introducing a metric on scale space which contains both the conductance and density functions. Specifically, let the metric be determined by the arc length form

$$ds^2 = \left(\frac{dx \cdot dx}{c^2}\right)^2 + \left(\frac{d\sigma}{\rho c}\right)^2.$$

Both c and ρ are allowed to depend on space, scale, and image data. It makes sense that a front–end vision system will make measurements (via a metric) that are data–dependent. Object interference is a natural phenomena, so the metric should reflect this and depend on functionals of the intensity. The total derivative in this metric is therefore

$$du = c\nabla u \cdot \frac{dx}{c} + \rho c u_\sigma \frac{d\sigma}{\rho c}.$$

The natural diffusion equation for this metric has left–hand side given by one application of the scale derivative and right–hand side given by two applications of the spatial gradient, namely

$$\left(\rho c \frac{\partial}{\partial \sigma}\right) u = (c\nabla)^2 u.$$

This simplifies to the anisotropic diffusion equation mentioned earlier: $u_\sigma = \rho^{-1}\nabla \cdot (c\nabla u)$. As special cases, for the linear diffusion equation in Euclidean space, the parameters are $c \equiv 1$ and $\rho \equiv 1$. For the linear diffusion in a non–Euclidean space, the parameters are $c \equiv \sigma$ and $\rho \equiv 1$. Finally, a number of researchers have been studying

$$u_\sigma = |\nabla u| \nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right)$$

either in the context of embedded curve evolution or for edge and corner detection (Alvarez, Lions & Morel 1992, Florack 1993, Kimia, Tannenbaum & Zucker 1992). In this case, $c = 1/|\nabla u| = \rho$.

Now that I have quantified what the metric is in scale space, appropriate choices can be made for c and ρ so that the metric has the desired invariances for the given application. The construction of multiscale data is then a *consequence* of metric selection. Any object analysis in the image is performed with respect to the metric, for example, measuring distances between objects. Such distances now naturally depend on the multiscale image data, so phenomena such as object interference can be accounted for.

Moreover, the metric should be used in constructing finite difference schemes for solving the anisotropic diffusion equation. In the linear diffusion cases, the grid points were positioned at constant distances from each other, distance being measured according to the metric. Grid selection for the general anisotropic case could be done similarly. But now, the selection is adaptive depending on what the values of conductance and density are. For example, if conductance is $c(x,\sigma)$ and density is $\rho(x,\sigma)$ (no image dependence for now), then the first derivative approximations are

$$cu_x \doteq \frac{u(x+\frac{h}{2}c(x,\sigma),\sigma)-u(x-\frac{h}{2}c(x,\sigma),\sigma)}{h}$$

and

$$\rho c u_\sigma \doteq \frac{u(x,\sigma+k\rho(x,\sigma)c(x,\sigma))-u(x,\sigma)}{k}$$

for some small positive constants h and k. The second derivative spatial approximation is somewhat more complicated, but it just involves the first–order difference operator applied twice to the function. In the following equation, I drop the dependence on σ just for notational simplicity. The second derivative is approximated by

$$\begin{aligned} c(cu_x)_x \quad \doteq \quad & \frac{1}{h^2}\Big[u(x+\tfrac{h}{2}c(x)+\tfrac{h}{2}c(x+\tfrac{h}{2}c(x)))-u(x+\tfrac{h}{2}c(x)-\tfrac{h}{2}c(x+\tfrac{h}{2}c(x))) \\ & -u(x-\tfrac{h}{2}c(x)+\tfrac{h}{2}c(x-\tfrac{h}{2}c(x)))+u(x-\tfrac{h}{2}c(x)-\tfrac{h}{2}c(x-\tfrac{h}{2}c(x)))\Big] \end{aligned}$$

It can be shown that the right–hand side converges to $c(cu_x)_x$ as $h \to 0$. Notice that when $c \equiv 1$ and $\rho \equiv 1$, the finite difference method is exactly the one given earlier for $u_\sigma = u_{xx}$. When $c = \sigma$ and $\rho \equiv 1$, the scale derivative approximation is slightly different. The scales are sampled geometrically, but in place of the term $\ln b$ appearing in the approximation we have $b-1$. The replacement by $\ln b$ can be viewed as a minor refinement which has some effect on the approximation error and on the range of b which provide a stable algorithm.

If the conductance and/or density depend on image values, now the difference schemes become implicit. They may be more difficult to implement, but hopefully, like many implicit schemes, they will exhibit unconditional stability.

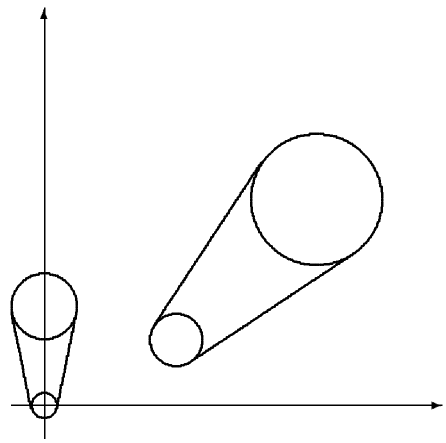

Figure 4.3: Objects to be registered

4.6 Discussion

The development of the properties of scale space is not to be viewed as simply an exercise in mathematics. The consequences of scale space measurements are far–reaching in applications to image analysis. Objects in an image impose their own "geometry" in the image. Any measurements made should depend on both position and scale of measurement. As one example, consider the problem of image registration. Let an object in a 2–dimensional image have core computed as $(\vec{x}(t), \vec{y}(t), \sigma(t)) = (0, \alpha t, (1-t)\sigma_0 + t\sigma_1)$ for $t \in [0,1]$, where $\alpha \neq 0$ and $0 < \sigma_0 < \sigma_1$. In a second image, the same object has been zoomed by a factor $\mu > 0$, rotated by an angle θ, and translated by an amount (a, b) from its original location. Its core will be computed as $(a - t\mu\alpha \sin\theta, b + t\mu\alpha\cos\theta, \mu[(1-t)\sigma_0 + t\sigma_1])$. Figure 4.3 below shows the two views of the same object.

If an attempt is made to register the two objects using chamfer matching, the results will be acceptable if the zoom factor is close to 1. In this case, the matching will provide information about how much the second object has been rotated and translated relative to the first object. But if the zoom factor is not close to 1, the matching becomes less reliable. A better registration will occur if the matching is done in scale space. In fact, the two objects match exactly. Note that the scale space arc length of the core of the first object is

$$\ell_1 = \int_0^1 \frac{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{\sigma}^2}}{\sigma(t)}\, dt = \int_0^1 \frac{\sqrt{\alpha^2 + (\sigma_1 - \sigma_0)^2}}{(1-t)\sigma_0 + t\sigma_1}\, dt$$

The scale space arc length of the core of the second object is

$$\ell_2 = \int_0^1 \frac{\sqrt{(-\mu\alpha\sin\theta)^2 + (\mu\alpha\cos\theta)^2 + [\mu(\sigma_1 - \sigma_0)]^2}}{\mu[(1-t)\sigma_0 + t\sigma_1]}\, dt = \int_0^1 \frac{\sqrt{\alpha^2 + (\sigma_1 - \sigma_0)^2}}{(1-t)\sigma_0 + t\sigma_1}\, dt = \ell_1,$$

as expected based on the invariances which were built into the definition of the metric for scale space.

One problem that was encountered in experiments with $2D$ registration was the 4–to–3 pixel sampling scheme which is used for display devices with the related aspect ratio. To make sure the true distances are calculated correctly, the values $\lambda_1 = 1$ and $\lambda_2 = 3/4$ were used in the metric G. Moreover, the scale samples were selected as a geometric sequence with base $b > 1$, so $\lambda = 1/\ln(b)$ was used. As a consequence, the scales σb^k and σb^{k+1} are one unit distance apart in scale space for any k.

More general distortions can occur in the sampling of pixels. If one knows the type of distortion *a priori*, then perhaps the same development given in this paper could be modified for more general metrics $G = \sigma^{-2}\mathrm{diag}(J^t J, \lambda^2)$ where J is an $n \times n$ matrix representing the derivative of the distortion $\vec{y} = \vec{F}(\vec{x})$. By using such a metric, the calculations made should be true to the coordinates of the original scene rather than to the distorted coordinates of the image representing the scene.

Another application currently being investigated is object–based interpolation of images (Puff, Eberly & Pizer 1994). The usual interpolation schemes simply compute a weighted average of intensities through various slices, but at a fixed spatial location. This approach treats pixels at this location, but in different slices, equally without regard to the pixel classification. If in one slice the pixel is white matter, but in the next slice it is gray matter, there is no anatomical justification for interpolating the two intensities. The interpolation should be based on relating pixels of the same classifications.

A solution to this problem is to phrase it in a scale space setting. The adjacent image slices are each segmented using multiscale methods to obtain a collection of objects and their corresponding cores. The objects are registered between pairs of slices. Interpolation of the intensities for a pair of registered objects now involves both interpolation of positions and of the intensities at those positions. The interpolation of positions is performed in a scale space setting. Pairs of points are identified with each point being in the same relative location to one object as the other point is relatively located to the paired object. The scale space distance measurements that are used essentially incorporate “width” information about the objects themselves. The geography of the objects should naturally tell how to pair up points. Once paired, the intensities between those points can be interpolated.

For objects that consist of a hierarchy of figures, each figure having its own core, an interpolation can be based on pairs of figures that are registered, but now one must take into account how the figures interrelate. The problem is being currently investigated.

The algorithms being developed in the Medical Image Processing group at UNC are successful because of their foundations in scale space geometry. I predict that many other imaging applications will succeed only if the problems are similarly formulated in a scale space setting. As indicated in the section relating anisotropic diffusion to the scale space metric, the metric selection for scale space becomes a very important aspect of any program of image analysis. Once selected, all the corresponding geometric tools such as distance, curvature, and ridges come to play in the analysis of the original image.