Chapter 6

Other Applications

This dissertation began with the thesis that geometry-limited diffusion is a reliable and accurate means of characterizing structure in images. The geometry-limited diffusion framework grew out of an attempt to modify isotropic scale space in order to reflect the local homogeneity of image structure. I have demonstrated the effectiveness of these principles in several different ways. While Chaptors 2 and 3 presented the basic philosophy and mathematics, Chapt. 4 explained a method for using the differential structure of images as a means of segmentation and feature extraction, and it gave some examples of first- and second-order diffusion processes. Chapter 5 examined a particular case, zero-order diffusion, in much more detail in order to provide quantitative evidence for the thesis. The purpose of this chapter is to step back and reexamine how the general principles could be extended to apply to different descriptions of local image structure. My assertion is that interesting segmentations can be formed on the basis of homogeneity of local image structure, and variable-conductance diffusion offers a means of combining information at different scales in order to produce reliable measurements of this structure. This chapter explores two additional applications of this general framework, albeit in the early stages of their development.

The first application assumes an alternative representation of local image structure. Local image structure is represented as a set of Fourier descriptors rather than differential geometry of the intensity surface. The result is a further generalization of geometry-limited diffusion which produces a system of reaction-diffusion equations which are tuned to specific frequencies. Some early results show that technique is capable of distinguishing patches on the basis of texture.

The second application shows how the techniques of supervised pixel classification can be used to construct dissimilarity measurements. This approach assumes nothing about the feature space but in the spirit of statistical pattern recognition assumes that pixels fall into classes that are described by multivariate probability distributions. This work is in its early stages, and there are no experimental results.

6.1 Spectra-limited diffusion and texture

6.1.1 Texture

The following definition is a relatively simple working definition of texture, which is not meant to capture all of the richness and variety of textures as they are perceived in the real world, but instead helps to establish a set of basic principles.

Definition 3 A *texture* is a patch in an image that has a periodic or "nearly" periodic luminance profile.

An image with such a patch is shown in Fig. 6.1.

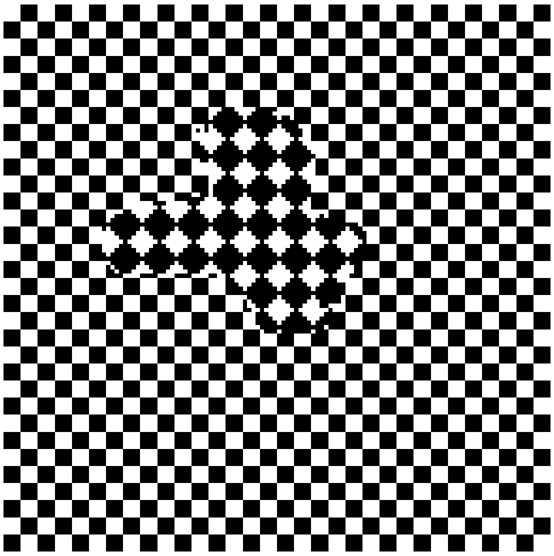
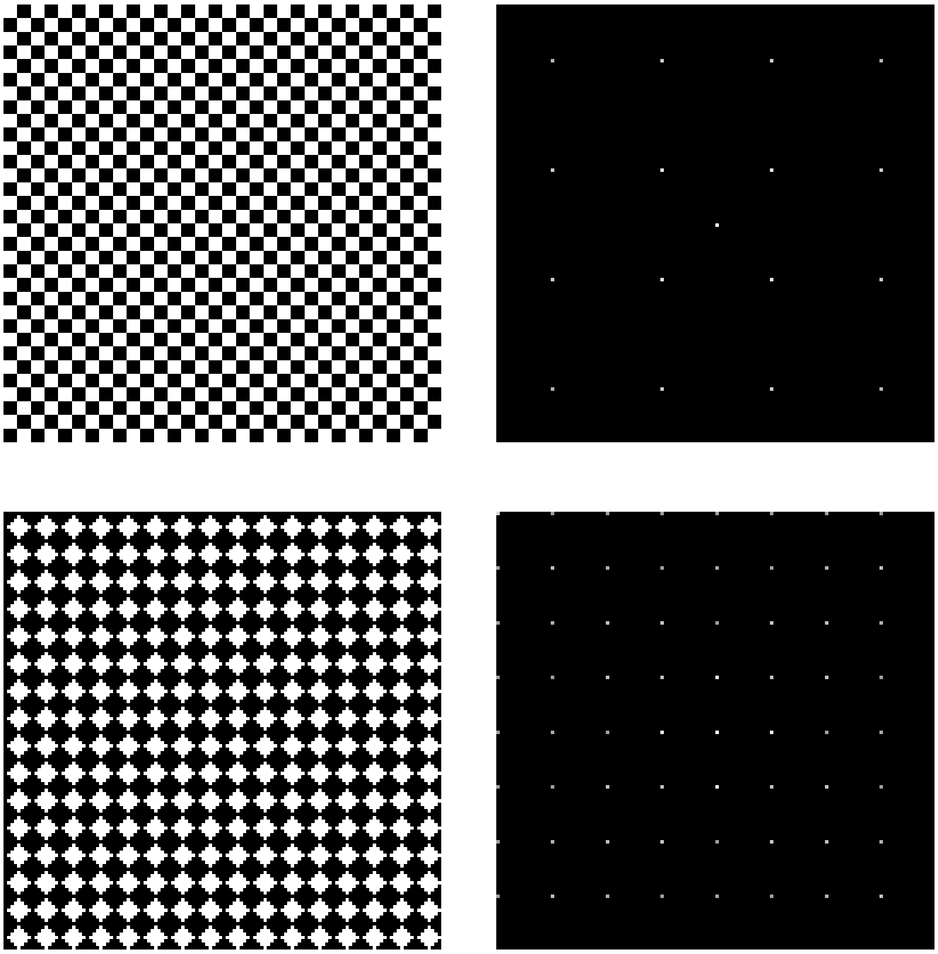


Figure 6.1: Textured patch on a textured background.

Several properties of such textures are worth noting. First, these patches are "extendable" in the sense that the periodic pattern could be repeated in order to create larger patches with the "same" texture. Second, the Fourier spectra of these textures have a great deal of structure when analyzed on regularly sized patches. Figure 6.2 shows the extended patches of the textures of Fig. 6.1 and the respective power spectra of those patches. Because of the periodic nature of these textures there are relatively few frequencies that have high energy. These power spectra suggest that there is some efficient encoding, based on the high-energy frequencies, that could be used to distinguish these textures. The difficulty is that textures are rarely analyzed in distinct rectilinear patches. The problem of segmenting images on the basis of texture typically involves distinguishing textures along irregular and often ambiguous boundaries. The spectral analyses of images that contain even a very few different textures with irregular boundaries are quite complex and typically lack the obvious structure of single texture spectra (e.g. see Fig. 6.3). The global nature of the Fourier transform makes it inadequate for segmenting textures.

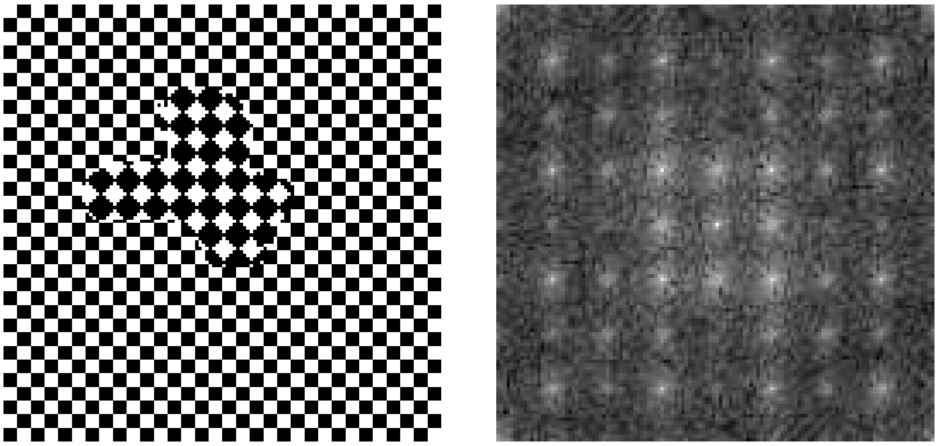
A great deal of work in both natural and artificial vision has focused on the importance of filters that are local in both frequency and space. Gabor (1946) shows that filters with optimal bandwidth in frequency and space are characterized in the spatial domain as sinusoids modulated by Gaussian envelopes. Indeed, the presence of frequency "channels" in natural vision has been argued on the basis of psychophysical data (Graham 1981, Graham 1979) In artificial vision such filters have be used to classify textures (Lee, Mumford and Yuille 1992, Julesz and Caelli 1979, Coggins and Jain 1985). Coggins and Jain (1985) compute at each pixel in a digital image an array (feature vector) of measurements from filters of various frequencies and orientations. They apply techniques of statistical pattern recognition in order to classify pixels based on their proximity in the feature space. The difficulty with this kind of "local" spectral analysis is the inverse relationship between the extent of filters in the space and frequency domains. Any attempt to make a precise measurement of a particular frequency component at a particular position an in image results in an uncertainty about the exact location of the structures that give rise to the measurement (Wechsler 1990, 48-85). The extreme case is the Fourier transform itself which is precise in frequency but global in space.

A similar set of difficulties is confronted in Chapters 3 and 4 when dealing with the problem of segmenting images into patches on the basis of intensity levels or luminance (as opposed to texture). In the presence of noise and ambiguous region boundaries, decisions about patch boundaries can



Textured patches Power spectrum

Figure 6.2: Textures extended to a rectilinear grid show a great deal of regularity in their respective power spectra.



Textured patches Power spectrum

Figure 6.3: Irregular textured patches show the difficulty of using the Fourier transform as a means of segmenting textured patches. The regularly spaced bright areas indicate the presence of energy, but do not offer any immediate way of separating the effects of the two different textures or retrieving local information.

become difficult. For this reason higher-order measurements of digital images are often combined with low pass transforms in order to provide stable measurements. However, low-pass filters such as the Gaussian can have adverse effects on the characterization of objects or patches whose shapes depend on high frequency information. Nonuniform or anisotropic diffusion has been proposed in earlier chapters as an alternative to stationary filtering with Gaussian kernels.

Gaussian blurring (diffusion) and nonuniform diffusion are inadequate when trying to distinguish patches on the basis of high frequency information, particularly when such patches are similar in their low frequency characteristics. This is often the case with textures. In this section I develop a type of blurring which acts as a band-pass filter and a diffusion process which implements such blurring in a local, iterative fashion. I then introduce a conductance term analogous to the variable conductance term, $g(\mathcal{D}\vec{F})$, which controls geometry-limited diffusion. This conductance term allows characteristic wave forms to flow up to, but not across, textural boundaries.

6.1.2 Blurring the envelopes of functions

The difficulty with low-pass blurring is that it can undermine the periodic nature of textures and thereby destroy the information that distinguishes different textural regions. As an alternative, consider the representation of a luminance function (one dimension) as a periodic function, with a characteristic frequency ω, multiplied by an envelope that describes the magnitude. In one dimension

$$f(x) = \varphi(x, \omega)\psi(x, \omega), \tag{6.1}$$

where $f : \mathbb{R} \mapsto \mathbb{R}$ is the luminance profile and $\psi : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{C}$ is a periodic function, independent of $f(x)$, that has a characteristic frequency ω. The function $\varphi : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{C}$ modulates the amplitude and phase of the periodic function $\psi(x, \omega)$ and describes the local presence of the characteristic frequency. Given a nonzero $\psi(x, \omega)$ and a luminance profile $f(x)$, solve for $\varphi(x, \omega)$:

$$\varphi(x, \omega) = \frac{f(x)}{\psi(x, \omega)}.$$

The function φ is the *envelope* of f for a frequency ω. If the characteristic wave form is defined for a continuous range of frequencies, there is an envelope $\varphi(x,\omega) = f(x)/\psi(x,\omega)$ for each frequency.

The envelope characterizes each point in the image over a range of characteristic frequencies. Note several things about the envelope of a function. First, it is generally complex valued. The pair of values for each frequency at every point encodes local spectral information, both magnitude and phase. Second, integrating $\varphi(x,\omega)$ over the entire domain produces a transform of $f(x)$. If the characteristic waveforms comprise a basis, then the integral

$$\int_I \varphi(x,\omega)\mathrm{d}x$$

is a complete representation of $f(x)$ over the image space $x \in I$. Third, if $\psi(x,\omega)$ is a sinusoid $e^{\imath\omega x}$, then the combination of blurring $\varphi(x,\omega)$ with Gaussian kernels and multiplying by $\psi(x,\omega)$ is precisely the same as a convolution of $f(x)$ with a Gaussian band-pass filter centered at frequency ω. Define a function $F(x,\omega,s)$

$$\begin{aligned} F(x,\omega,s) &= \psi(x,\omega)(G(s,x) \otimes \varphi(x,\omega)) \\ &= e^{\imath\omega x}\left(G(s,x) \otimes (f(x)e^{-\imath\omega x})\right) = (G(s,x)e^{-\imath\omega x}) \otimes f(x)) \end{aligned} \tag{6.2}$$

The resulting filter $Z(x,\omega,s) = G(s,x)e^{\imath\omega x}$ has both real and imaginary parts which form a pair of Gabor blobs (sine and cosine respectively). The magnitude of this kernel is a Gaussian $G(x,s)$, where the scale parameter s defines the standard deviation of the Gaussian and the effective width of the Gabor filter. For larger values of s these filters are wider in space and narrower in the frequency domain. The luminance function $f(x)$ can be described by measurements made with Gabor blobs over a range of scales and frequencies. Thus $F : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+ \mapsto \mathbb{C}$ is a frequency-tuned scale space as described by Florack [6].

A degree of spatial coherence for the envelope $\varphi(x,\omega)$ is achieved by applying some low-pass filter. The more smoothing that is done to $\varphi(x,\omega)$ the narrower the bandwidth of the corresponding band pass filter. Consider Gaussian blurring as a process described by the heat equation,

$$\frac{\partial\varphi(x,\omega)}{\partial t} = \frac{\partial^2\varphi(x,\omega)}{\partial x^2} \tag{6.3}$$

and let the characteristic wave form be a sinusoid, $\psi(x,\omega) = e^{\imath\omega x}$. There is an equivalent process which modifies $f(x)$ with respect to a characteristic frequency ω in order to produce a smoother envelope for that frequency:

$$\frac{\partial F(x,\omega,t)}{\partial t} = \frac{\partial^2 F(x,\omega,t)}{\partial x^2} - 2\imath\omega\frac{\partial f(x)}{\partial x} - \omega^2 F(x,\omega,t), \tag{6.4}$$

where $F(x,\omega,0) = f(x)$. This is a diffusion equation with fundamental solutions of the form

$$G(x,\omega,s) = \frac{1}{\sqrt{2\pi}s}e^{\frac{x^2}{2s^2}-\imath\omega x}. \tag{6.5}$$

These are, once again, Gabor blobs. This process describes the "flow" of these Gabor blobs as they spread out in space. As the process evolves, it produces a progressively more narrow bandpass filter. The steady state solution of this equation (given the appropriate boundary conditions) is a sine and cosine wave pair.

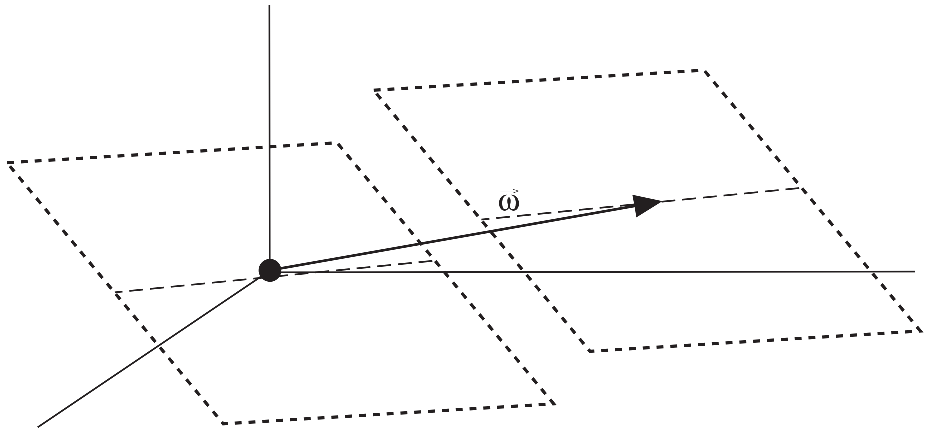


Figure 6.4: The operator $\nabla_{\vec{\omega}}$ is equivalent to an ordinary gradient ∇ that is shifted in the Fourier domain by the vector $\vec{\omega}$

6.1.3 Variable-conductance diffusion

Equation 6.4 describes a means of characterizing local frequency information. It specifies a scale-space for *each* characteristic frequency ω and provides a means of "smoothing" the function with respect to a particular frequency. The difficulty with this smoothing, as with Gaussian blurring, is the inherent trade-off between the size of the fundamental solutions in the spatial and Fourier domains. This trade-off has the same mathematical form as the uncertainty principle which dictates the fundamental limitations in quantum mechanical observations. I avoid this trade-off by employing the principles of nonuniform or anisotropic diffusion. The strategy is to introduce a variable-conductance term that limits the spread of the fundamental solution. Notice that Eq. (6.4) is the result of applying two first-order operators.

$$\begin{aligned} \frac{\partial F(x,\omega,t)}{\partial t} &= \frac{\partial^2 F(x,\omega,t)}{\partial x^2} - 2\imath\omega\frac{\partial F(x,\omega,t)}{\partial x} - \omega^2 F(x,\omega,t) \\ &= \left(\frac{\partial}{\partial x} - \imath\omega\right)\cdot\left(\frac{\partial}{\partial x} - \imath\omega\right) F(x,\omega,t) \end{aligned} \tag{6.6}$$

Equation 6.6 has the same form, $D \cdot DF = \partial F/\partial t$, as the heat equation, where D is some linear differential operator. This similarity in form suggests a family of differential operators, analogous to the gradient operator, each sensitive to its associated characteristic frequency, ω. This family of operators $(\partial/\partial x - \imath\omega)$ operators has a somewhat intuitive interpretation when analyzed in the Fourier domain. Consider the Fourier representation of a directional derivative in a two-dimensional image space – it is a plane passing through the origin. The gradient operator is a set of planes (two planes if $I \subset \mathbb{R}^2$). Linear combinations of these planes produce a single plane – a directional derivative. The frequency-tuned derivative operators have the same interpretation except that they have zero response not at the origin, but at the characteristic frequency ω. The operator $(\partial/\partial x - \imath\omega)$ is a gradient that has undergone a translation (by an amount ω) in the Fourier domain as shown in Fig. 6.4.

The expressions (6.1-6.6) extend easily to image spaces of arbitrary dimension, and for this I use the vector notation to represent the space and frequency variables. Assuming an n-dimensional image, then $\vec{x}, \vec{\omega} \in \mathbb{R}^n$. Denote this "shifted" operator as $\nabla_{\vec{\omega}}$, and the corresponding Laplacian as $\Delta_{\vec{\omega}}$.

$$\nabla_{\vec{\omega}} = (\nabla - \imath\vec{\omega}) = e^{\imath\vec{\omega}\cdot\vec{x}}\nabla e^{-\imath\vec{\omega}\cdot\vec{x}}$$

$$\Delta_{\vec{\omega}} = (\nabla - \imath\vec{\omega}) \cdot (\nabla - \imath\vec{\omega}) = e^{\imath\vec{\omega}\cdot\vec{x}} \nabla \cdot \nabla e^{-\imath\vec{\omega}\cdot\vec{x}}$$

The gradient ∇ is the set of first-order derivatives in the form of a vector.

$$\nabla \equiv \frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_n}$$

Notice that the conventional gradient and Laplacian are a special cases of the operators $\nabla_{\vec{\omega}}$ and $\Delta_{\vec{\omega}}$ with $\vec{\omega} = 0$. Thus the diffusion Eq. (6.6) becomes

$$\frac{\partial F}{\partial t} = \nabla_{\vec{\omega}} \cdot \nabla_{\vec{\omega}} F \tag{6.7}$$

The form of (6.7) allows one to see more clearly the means of incorporating variable conductance into this frequency-sensitive diffusion process. Recall the analogy described in Sect. 2.5.2 where the diffusion equation describes the "flow" of a substance for which F is a density function. The first derivative $\nabla_{\vec{\omega}} F$ is the difference in concentration of the substance and thereby the flow of the quantity F from one location to another (infinitesimally close). The Laplacian $\nabla_{\vec{\omega}} \cdot \nabla_{\vec{\omega}} F(\vec{x})$ is the net flow into a location $\vec{x}$. Modulate the flow of F with a conductance term as follows:

$$\frac{\partial F}{\partial t} = \nabla_{\vec{\omega}} \cdot g(\vec{x}) \nabla_{\vec{\omega}} F. \tag{6.8}$$

Experiments in one dimension show that this equation behaves as expected. Solutions to (6.8) take the form of waves that "flow" but are impeded by regions of low conductivity. Figure 6.5 shows the evolution of a single pulse in a nonconducting "box". The pulse gradually spreads to fill the box forming a pair of waves whose combined magnitude is virtually constant.

Equation (6.8) provides the framework for a diffusion equation that is sensitive to frequency and also incorporates a conductance term which is sensitive to local structure.

$$\frac{\partial F}{\partial t} = \nabla_{\vec{\omega}} \cdot g(\mathcal{D}F) \nabla_{\vec{\omega}} F \tag{6.9}$$

The dissimilarity $\mathcal{D}F$ depends on spatial derivatives of $F(\vec{x}, \vec{\omega}, t)$. The gradient measurement for the conductance term must incorporate some finite isotropic scale. The derivatives in the conductance term are measured with a Gaussian neighborhood function. This isotropic scale must reflect the same frequency sensitivity as the diffusion process, and gradient measurements should be conducted with kernels composed of Gabor blobs as in Eq. (6.5). This scale term slowly decreases as the process progresses in order to provide a large to small scale process that "narrows in" on relevant boundaries.

Experiments in one dimension, with a single nonzero frequency, confirm the desired behavior of (6.9). This process produces distinct regions with sinusoidal waves of nearly constant amplitude. The regions and the amplitude of the waves within regions reflect the presence of the characteristic frequency in the original data.

There is no need to implement this mixed-order partial differential equation as it is shown above, because the operator $\nabla_{\vec{\omega}}$ corresponds to a frequency shift followed by the application of the gradient operator ∇. To implement this alternative strategy, first shift the initial conditions (the luminance profile) in the frequency domain and then use the operator ∇ instead of $\nabla_{\vec{\omega}}$. Shifting the initial conditions and performing a variable-conductance diffusion is the same as blurring the envelope $\varphi(\vec{x}, \vec{\omega}, t)$. Assuming that the dissimilarity in (6.9) is some function $\mathcal{D}F \equiv ||\phi(F)\nabla_{\vec{\omega}} F||$,

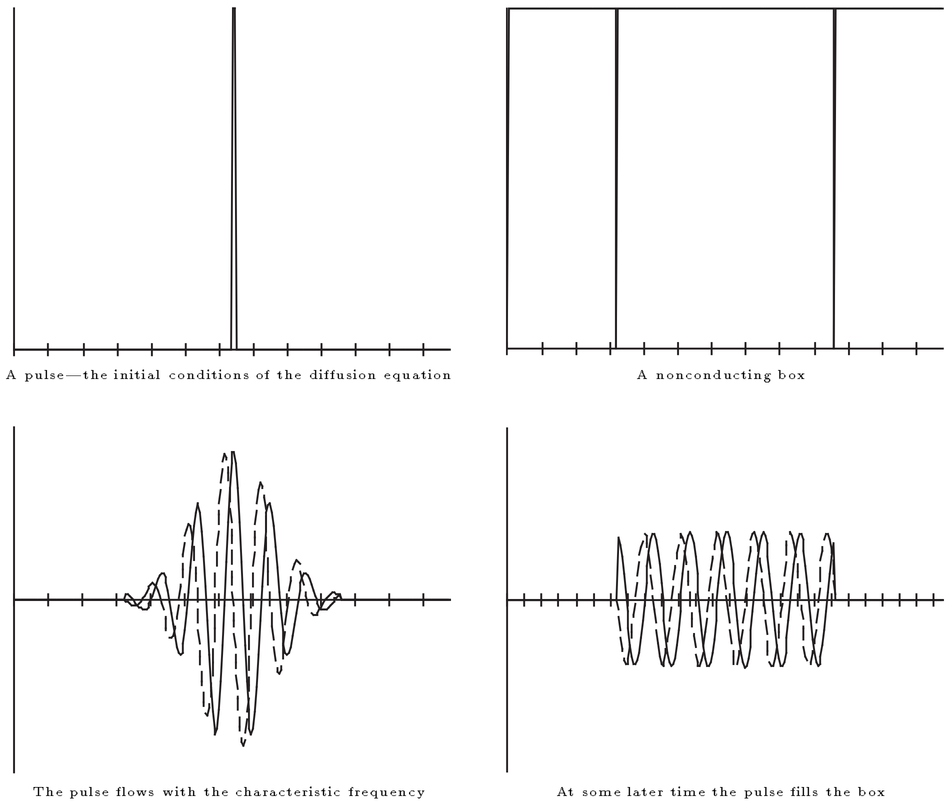


Figure 6.5: From top-left: A one-dimensional nonconducting box is constructed (a) and a single pulse (b) spreads at the characteristic frequency (c) finally filling the box (d).

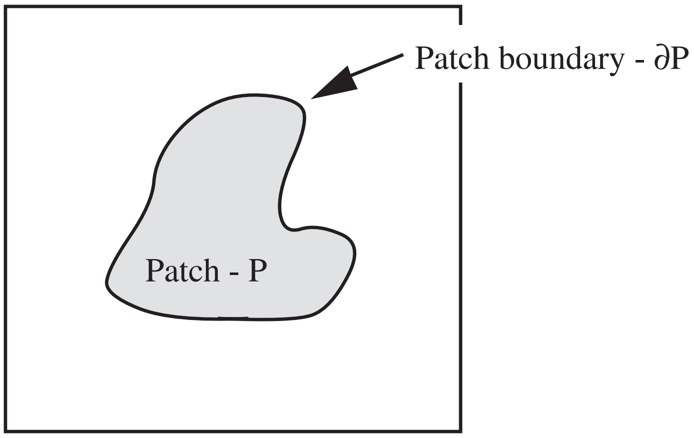


Figure 6.6: A patch with a nonconducting boundary is used to show that systems of frequency-sensitive variable-conductance diffusion equations are capable of computing a patchwise Fourier analysis.

then Eq. (6.9) becomes

$$\frac{\partial \varphi}{\partial t} = \nabla \cdot g(\|\phi(\varphi) e^{-\imath \vec{\omega} \cdot \vec{x}} \nabla \varphi\|) \nabla_{\vec{\omega}} \varphi. \tag{6.10}$$

This form is easier than (6.9) to implement on a discrete grid, because it involves fewer derivatives and enables the conductance measurements be made with a Gaussian kernel instead of the Gabor pair.

6.1.4 Patchwise Fourier decomposition

In this section I make some simplifying assumptions about the conductance function and show that Eq. (6.10) implements a Fourier decomposition within patches. Consider an image space I with luminance function $f(\vec{x})$. Assume that I contains a patch P with a boundary ∂P as shown in Fig. 6.6. For the purposes of this example, assume that the conductance function is constant over time as in (6.8), identically 1 over the interior of P, and 0 along the boundary ∂P.

Consider the steady state solutions of Eq. (6.10) within the patch P. Notice that the nonconducting border around P recasts the problem of finding the steady state of (6.10) as a boundary value problem in the irregular patch with Neuman boundary conditions (adiabatic). This insures two properties in the solutions:

1. By the divergence theorem the integral

$$\int_P F(\vec{x}, \vec{\omega}, t) \mathrm{d}\vec{x}$$

   remains constant over time.

2. The steady state solution is a boundary value problem (Laplacian) with constant boundary conditions, and solutions are constant over the patch.

Define a new function

$$f_P(\vec{x}) = \begin{cases} = f(\vec{x}) & \text{if } \vec{x} \in P \\ = 0 & \text{if } \vec{x} \notin P \end{cases}$$

and express the steady state solution in closed form.

$$\begin{aligned}
A_P &= \int_P \mathrm{d}\vec{x} \\
\lim_{t\to\infty} \varphi(\vec{x},\vec{\omega},t) &= \frac{1}{A_P}\int_P \varphi(\vec{x},\vec{\omega},0)\mathrm{d}\vec{x} \\
&= \frac{1}{A_P}\int_P f(\vec{x})e^{-\imath\vec{\omega}\cdot\vec{x}}\mathrm{d}\vec{x} \\
&= \frac{1}{A_P}\int_I f_P(\vec{x})e^{-\imath\vec{\omega}\cdot\vec{x}}\mathrm{d}\vec{x} \\
&= \frac{1}{A_P}\mathcal{F}f_P(\vec{x}),
\end{aligned}$$

where $\mathcal{F}$ is the Fourier transform. The steady state solution within the patch is 1 over the area of the patch times the value of the Fourier transform of the luminance function defined only on the patch. *To the extent that patches are distinct and well defined, this process implements a patchwise Fourier transform.* The steady-state values within patches reflect not only the frequency characteristics of the luminance function, $f(\vec{x})$, but also the frequency characteristics of the shape of the patch P. For a perfectly flat luminance function this process generates a Fourier transform that describes only the *shapes* of patches.

6.1.5 Multi-valued diffusion

Although the argument above refers to an extreme case where patch boundaries are detectable and absolute, the goal is to employ the principles of multi-valued diffusion in order to resolve boundaries in cases where patches are ambiguous and difficult to detect. The function $\varphi : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^+ \mapsto \mathbb{C}$ describes a scale-space of an image of a range of characteristic frequencies. Instead of a continuous range of frequencies, consider a finite set of "sampled" frequencies. Let $\vec{\Omega} = \vec{\omega}_1, \ldots, \vec{\omega}_m$, and let $\vec{F}_{\vec{\omega}_i}(\vec{x},t) = \varphi(\vec{\omega}_i,\vec{x},t)$. Then $\vec{F}_{\vec{\Omega}}(\vec{x},t) = F_{\vec{\omega}_1}(\vec{x},t), \ldots, F_{\vec{\omega}_m}(\vec{x},t)$ is a multi-valued function $\vec{F}_{\vec{\Omega}} : \mathbb{R}^n \times R^+ \mapsto \mathbb{C}^m$. Now treat this new function $\vec{F}_{\vec{\Omega}}$ with the multi-valued diffusion equation from (4.1),

$$\frac{\partial \vec{F}_{\vec{\Omega}}}{\partial t} = \nabla \cdot g(\mathcal{D}(G(s) \otimes \vec{F}_{\vec{\Omega}}))\nabla \vec{F}_{\vec{\Omega}}, \tag{6.11}$$

so that

$$\begin{aligned}
\frac{\partial F_{\vec{\omega}_1}}{\partial t} &= \nabla \cdot g(\mathcal{D}_s \vec{F}_{\vec{\Omega}})\nabla F_{\vec{\omega}_1} \\
&\vdots \\
\frac{\partial F_{\vec{\omega}_m}}{\partial t} &= \nabla \cdot g(\mathcal{D}_s \vec{F}_{\vec{\Omega}})\nabla F_{\vec{\omega}_m}
\end{aligned}$$

where

$$\vec{F}_{\vec{\omega}_i}(\vec{x},0) = f(\vec{x})e^{\imath\vec{\omega}_i\cdot\vec{x}}.$$

Equation 6.11 is a system of equations in which each equation represents a distinct characteristic frequency. With perfectly defined boundaries, as described in Sect. 6.1.4, each equation computes a patchwise Fourier integral for a single frequency. This collection of equations produces a frequency-wise sampling of the Fourier transform within each patch.

This approach for segmenting images bears some similarity to conventional means for characterizing textures in images (Wechsler 1990). It utilizes a frequency-based representation of local structure and and incorporates filters that are local in both frequency and space. However, it differs from conventional approaches in some very important aspects. The narrow bandwidth in the frequency domain is obtained by blurring the *envelope* of an intensity function. The result is an incremental widening of these filters in the spatial domain. This blurring, however, is both nonstationary and nonlinear. The "spreading" of these frequency-tuned filters is limited by areas of low conductance. The result is a set of measurements that are *adapted* to local structure in images. Structure at different scales is combined by utilizing isotropic scale in the dissimilarity measurements and decreasing this scale as the process evolves.

6.1.6 Dissimilarity

The challenge is to construct a dissimilarity measure which describes the boundaries of textured patches. Patches in images are based not only on a single characteristic frequency but on the relationships between a number of frequencies. There are myriad options for this transformation ϕ associated with the dissimilarity measure. I choose a simple but effective dissimilarity operator based on the following principle: *Textural boundaries are distinguished rby the relative changes in local spectral components.* That is, boundaries are affected not by the overall quantity of spectral energy but in the relative spectral composition at each point in the image.

Consider the range of $\vec{F}_{\vec{\Omega}}$, which is a $2m$-dimensional feature space D. The points in this feature space can be represented in hyper-spherical coordinates. Each point in the space D can be treated as a vector with both a magnitude and a direction. The principle stated above requires that the dissimilarity measure should reflect changes in the direction of vector $\vec{F}_{\vec{\Omega}}$, not its magnitude. In order to achieve a dissimilarity measure which does not depend on the magnitude of $\vec{F}_{\vec{\Omega}}$, use a set of normalized intermediate features $\vec{G}(\vec{F}_{\vec{\Omega}})$. The normalization is another application of dimensionless metrics as discussed in Section 4.5.2. This dissimilarity can also be expressed as a linear transformation $\phi(\vec{F}_{\vec{\Omega}})$ applied to Jacobian of $\vec{F}_{\vec{\Omega}}$, so that $\mathcal{D}\vec{F}_{\vec{\Omega}} = \|\phi(\vec{F}_{\vec{\Omega}})\mathrm{D}\vec{F}_{\vec{\Omega}}\|$. However, it is more convenient to write it as

$$\mathcal{D}\vec{F}_{\vec{\Omega}} = \left\|\nabla\vec{G}(\vec{F}_{\vec{\Omega}})\right\| = \left\|\nabla\frac{\vec{F}_{\vec{\Omega}}}{|\vec{F}_{\vec{\Omega}}|}\right\|, \tag{6.12}$$

where

$$|\vec{F}_{\vec{\Omega}}| = \left(\sum_{i=1}^{m} F_{\vec{\omega}_i}F_{\vec{\omega}_i^*}\right)^{\frac{1}{2}},$$

and F^* is the complex conjugate of F. This dissimilarity operator has the property that it is insensitive to linear intensity transformations in the original image and thus it does not favor areas in an image that have high contrast. Boundaries in low contrast images are detected on the basis of their frequency composition and are not penalized for having low energy across all spectral features. Lee et al. (1992) use a similar measure for calculating relative changes in frequency decomposition.

6.1.7 Results

The spectral features $F_{\vec{\omega}_1}, \ldots, F_{\vec{\omega}_m}$ are *samples* taken in the frequency domain that encode textures in the image. Choosing the number and placement of these samples proved to be very difficult and appears to affect the results. Choosing too few samples results in processes that are unable to segment simple textures. I have experimented with different grid sizes and choose a sampling strategy that is a compromise between sample density and computation time. For this study I sample the frequency domain on a regular grid that includes the half plane minus a ray from the origin as shown in the Fig. 6.7.

The variable conductance diffusion is computed by the method of finite differences using a rectilinear grid (pixels) in the image space as described in the Appendix. Dissimilarities are computed

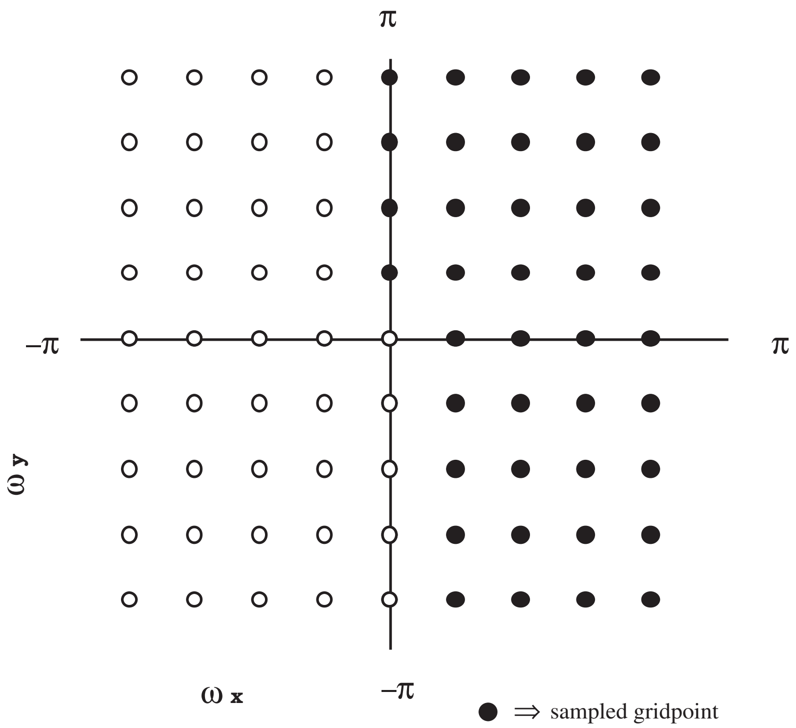
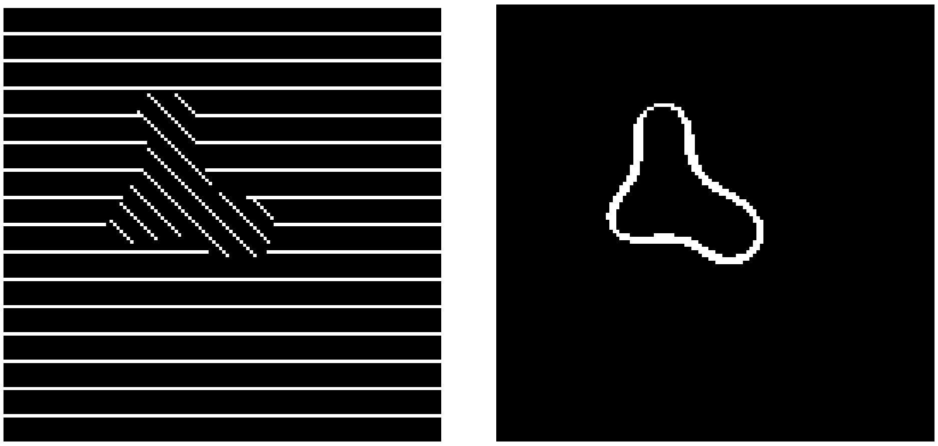


Figure 6.7: Features represent samples from a rectilinear grid in the frequency domain.

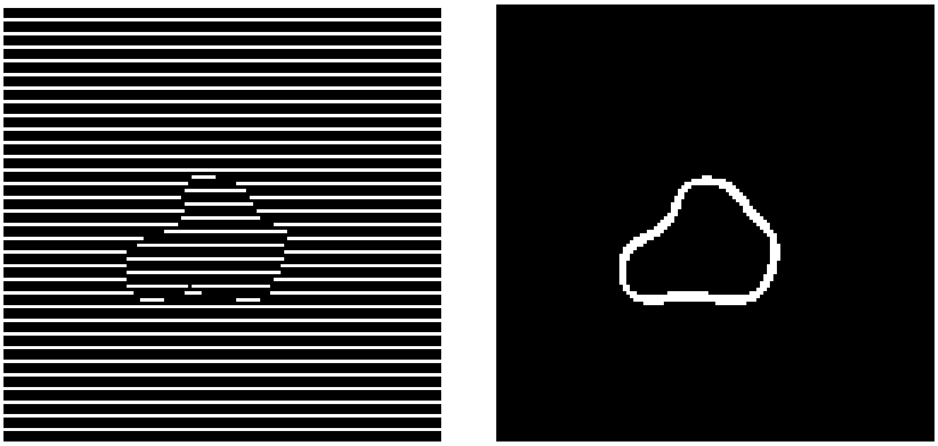
from the derivatives of Gaussian blurred versions of the features at every time step. This approach requires, for each time step and all 40 features, the convolution with a Gaussian kernel, the computation of conductance (including normalization, derivative, and exponent), and the calculation of second-order finite differences. A typical run requires 100 time steps. This computational burden is addressed by utilizing the parallel nature of this diffusion process and using a highly parallel single-instruction-multiple-datastream (SIMD) computer. Examples in this paper were computed on the MasPar MP-1, a SIMD computer made up of a 64×128, eight-connected, processor grid. Because the diffusion requires only nearest neighbor differences, communication costs are low, making this application extremely efficient on this type of machine.

The following figures show the results of this process using the sampling pattern shown in Fig. 6.7, which contains 40 features. In each example the process runs 100 iterations and requires about 15 minutes to complete. These pictures show the original images and the results of computing a threshold on the dissimilarity measure, $\mathcal{D}\vec{F}_{\vec{\Omega}}$, after the final iteration. The dissimilarity measure controls the conductance throughout the process. A threshold of the dissimilarity measure provides a very simple way of locating boundaries with this technique. The process does not work as well on more complicated textures. I believe that this problem is related to the sampling strategy in the frequency domain, as well as the limitation on image size and resolution (128×128 pixels) which was imposed by the working memory available on the Maspar MP-1.



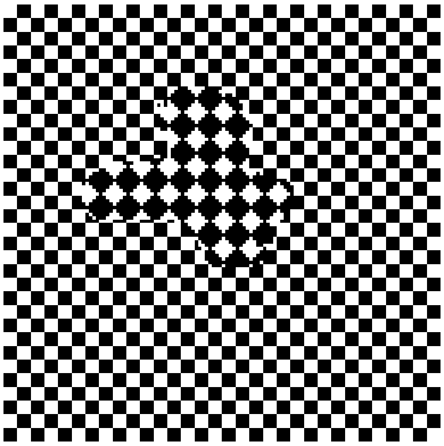
Original image Dissimilarity after processing

Figure 6.8: A patch in an image is distinguished from the background by the orientation of the stripes.

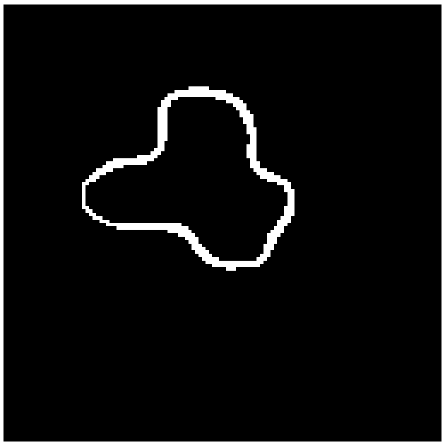


Original image Dissimilarity after processing

Figure 6.9: A patch in an image is distinguished from the background by the relative phase of the stripes.



Original image



Dissimilarity after processing

Figure 6.10: Textures are distinguished on the basis of orientation and size.

6.2 A supervised clustering approach

Previous sections of this dissertation (Sects. 4.1.3–4.6.2, 6.1) propose dissimilarity measures that are based on the geometry of the intensity surface, the desired patch characteristics, and some general requirements such as group invariance (Sect. 4.5.1) and dimensionless metrics (Sect. 4.5.2). In this section I explore the possibility that dissimilarity could be based on statistical rather than geometric considerations.

A frequent scenario in computer vision is as follows. Several "measurements", $F_1, \ldots F_m$, exist at each point (or pixel) in an image. These measurements might result from some imaging device (or devices) which measures several properties at one point. This device could measure some kind of radiation at a number of different wavelengths. These measurements could also result from a set of filters applied to a single-valued image as discussed in Chapt. 4. In addition to these measurements one is given a set of "classes" or "categories" and sets of measurements which consist of representative samples from each of these classes. The problem is to "classify" pixels in an image into one or more of these categories on the basis of the measurements at each pixel.

One approach is to treat the set of measurements, $F_1, \ldots F_m$, as a feature space and to assume that the position of a pixel in this feature space is indicative of its membership in one of the classes (Fig. 6.11). In addition, many algorithms assume that the likelihood of membership in a class depends only on the position in the feature space; membership is independent of the position in the image space. The probability distributions associated which each class have mean and covariance. These metrics provide a means for making decisions about the *most likely* membership of a particular pixel, and they can provide information for computing the confidence of such classifications. This approach (and some others like it) suffers from a problem. This algorithm ignores the spatial homogeneity (or inhomogeneity) of pixel classifications. If a pixel i is surrounded by neighbors that are in class A, then should not the likelihood of its membership in A reflect this fact? Blurring feature values with low-pass filters helps to achieve this property of spatial interdependence but with many of the same drawbacks discussed in Chapt. 2. The multi-valued diffusion equation of Section A.18 offers the opportunity to combine information about proximity in the feature space with homogeneity in the image space.

Assume a set of classes $C_1, \ldots, C_l$ and a multivariate probability density distribution for each class $p_1(\vec{w}), \ldots, p_l(\vec{w})$, where $\vec{w} \in D \subset \mathbb{R}^m$. Each distribution is normalized so that $\int_D p_i(\vec{w})d\vec{w} = 1 \;\; \forall C_i$.

Finally, assume that one knows the relative *total* likelihood of each class, $S_1, \ldots, S_l$. That is, a priori, without knowledge about location in the image space, a pixel has a probability S_i of being a member of C_i. Note that $\sum_i S_i = 1$. The class probability information, $S_1 \ldots S_l$, would likely be obtained at the same time information about the individual probability density distributions is gathered. For instance, given a pre-segmented image, random samples from that image could give information about the shapes of the distributions $p_1(\vec{w}), \ldots, p_l(\vec{w})$, as well as the total likelihood of each distribution. As a notational convenience, define a set of distributions that reflect this weighting $P_1(\vec{w}), \ldots, P_l(\vec{w})$, where $P_i(\vec{w}) = S_i p_i(\vec{w})$.

The strategy is to find a $\phi(\vec{w})$ that measures dissimilarity in the feature space and takes advantage of the statistical information that is captured in the probability distributions $P_1(\vec{w}), \ldots, P_l(\vec{w})$. I construct such a metric using the principle that *distance in the feature space should reflect changes in the likelihood of membership in the classes and should* not *capture distances within classes that do not change the likelihood of membership.*

I adopt a strategy that uses the statistical information about the classes in order to compute a dissimilarity measure. The dissimilarity measure can be used in the multi-valued diffusion equation in order to achieve a patchwise classification scheme. The principle driving this analysis is that "dissimilarity" should reflect the variation in conditional probabilities associated with a neighborhood in the image space.

6.2.1 Two-class problem

Consider a simplified case with two classes, C_1 and C_2. Assume that that the classes are mutually exclusive, that is $p(C_1 \wedge C_2) = 0$. For a point $\vec{w} \in D$, Bayes rule states that the probability of $\vec{w} \in C_1$ is

$$p(C_1|\vec{w}) = \frac{p(\vec{w}|C_1)p(C_1)}{p(\vec{w})} \tag{6.13}$$

$$= \frac{p(\vec{w}|C_1)p(C_1)}{p(\vec{w}|C_1)p(C_1) + p(\vec{w}|C_2)p(C_2)} \tag{6.14}$$

$$= \frac{p_1(\vec{w})S_1}{p_1(\vec{w})S_1 + p_2(\vec{w})S_2} \tag{6.15}$$

$$= \frac{P_1(\vec{w})}{P_1(\vec{w}) + P_2(\vec{w})} \tag{6.16}$$

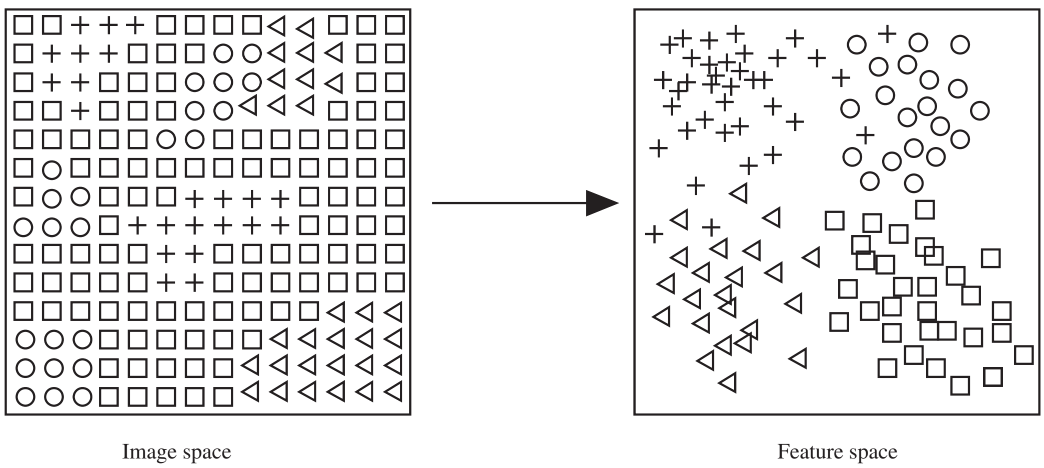


Figure 6.11: Locations in the image space are classified based on their position in the feature space.

(6.17)

The derivative of this probability describes the way in which that probability varies in the feature space:

$$\begin{aligned} \frac{\partial}{\partial \vec{w}} p(C_1|\vec{w}) &= \frac{\partial}{\partial \vec{w}} \frac{P_1}{P_1 + P_2} \\ &= \frac{\frac{\partial}{\partial \vec{w}} P_1}{P_1 + P_2} - \frac{P_1(\frac{\partial}{\partial \vec{w}} P_1 + \frac{\partial}{\partial \vec{w}} P_2)}{(P_1 + P_2)^2} \\ &= \frac{P_2 \frac{\partial}{\partial \vec{w}} P_1 - P_1 \frac{\partial}{\partial \vec{w}} P_2}{(P_1 + P_2)^2}, \end{aligned} \quad (6.18)$$

where the probability functions are understood to be dependent on the position $\vec{w}$ in the feature space. The partial derivative with respect to the vector $\vec{w}$ indicates the derivative in the form of a vector, so that

$$\frac{\partial}{\partial \vec{w}} \stackrel{\text{def}}{=} \begin{matrix} \frac{\partial}{\partial w_1} \\ \vdots \\ \frac{\partial}{\partial w_n} \end{matrix} .$$

The symbol ∇ is reserved for partial derivative with respect to coordinates in the image space, that is $\nabla \stackrel{\text{def}}{=} \partial / \partial \vec{x}$. Let $\vec{\gamma}_i = (\partial / \partial \vec{w}) p(C_i|\vec{w})$ for $i = 1, 2$. Notice that $\vec{\gamma}_1 = -\vec{\gamma}_2$.

The behavior of $\vec{\gamma}_i(\vec{w})$ is understood by treating it as a vector located at a point $\vec{w}$ in the feature space. This function defines a vector field which is at every point normal to the set of surfaces defined by $P_1/P_2 = \alpha$. The direction of $\vec{\gamma}_i$ indicates the direction of maximal change in the ratio P_1/P_2. All of the directions perpendicular to $\vec{\gamma}_i(\vec{w})$ lie in one of the surfaces $P_1/P_2 = \alpha$. The magnitude of this vector is

$$|\vec{\gamma}_1| = |\vec{\gamma}_2| = \frac{(P_1^2 \frac{\partial}{\partial \vec{w}} P_2 \cdot \frac{\partial}{\partial \vec{w}} P_2 - 2 P_1 P_2 \frac{\partial}{\partial \vec{w}} P_1 \cdot \frac{\partial}{\partial \vec{w}} P_2 + P_2^2 \frac{\partial}{\partial \vec{w}} P_1 \cdot \frac{\partial}{\partial \vec{w}} P_1)^{\frac{1}{2}}}{(P_1 + P_2)^2} \quad (6.19)$$

This magnitude is greatest when $\frac{\partial}{\partial \vec{w}} P_1$ and $\frac{\partial}{\partial \vec{w}} P_2$ point away from each other (i.e. angles greater than $\pi/2$ so that $\frac{\partial}{\partial \vec{w}} P_1 \cdot \frac{\partial}{\partial \vec{w}} P_2 < 0$). The vector $\vec{\gamma}_i$ combined with a multi-valued image $F(\vec{x}) : I \mapsto D$ describes how movements in the image space affect the conditional probability.

$$\nabla p(C_1|\vec{F}(\vec{x})) = \frac{\partial p(C_1|\vec{F}(\vec{x}))}{\partial \vec{w}} \cdot \nabla \vec{F}(\vec{x}) = \vec{\gamma}_1(\vec{w}) \cdot \nabla \vec{F}(\vec{x}) \quad (6.20)$$

The derivative of the conditional probability in the image space describes the *degree of inhomogeneity in the probability of a classification* C_i $i = 1, 2$, *in the neighborhood of a point* $\vec{x}$. Movements in the image space which are perpendicular to this vector do not change the conditional probability of being in a class C_i. This vector points in a direction in the image space which produces a maximal change in the conditional probability. The magnitude of this derivative is the maximum probability change that can be achieved for a small movement in the image space.

This suggests a strategy for combining a supervised classification scheme with the multi-valued diffusion equation. Use the information from a supervised classification scheme to produce a probability distribution for each of the classes. From the set of probability distributions compute the function $\vec{\gamma}(\vec{w}) = \vec{\gamma}_1(\vec{w}) = -\vec{\gamma}_2(\vec{w})$ and use it as the metric ϕ in the multi-valued diffusion equation, so that the diffusion process combines information about the classes with spatial homogeneity. The result is a multi-valued diffusion equation

$$\frac{\partial \vec{F}}{\partial t} = \nabla \cdot g(\|\vec{\gamma}(G(s) \otimes \vec{F}) \nabla_s \vec{F}\|) \vec{F} \quad (6.21)$$

6.2.2 Normal distributions

This analysis applies to any set of probability distributions in a feature space. However, if these probability distributions are normal, the metric used in the dissimilarity can be described analytically.

$$P_1 = A_1 \exp(-(\vec{w} - \vec{\mu_1})^T \Sigma_1 (\vec{w} - \vec{\mu_1})) \text{ and } P_2 = A_2 \exp(-(\vec{w} - \vec{\mu_2})^T \Sigma_2 (\vec{w} - \vec{\mu_2}))$$

The derivatives are

$$\begin{aligned} \frac{\partial}{\partial \vec{w}} P_1 &= -A_1 (\vec{w} - \vec{\mu_1})^T \Sigma_1 \exp\left(-(\vec{w} - \vec{\mu_1})^T \Sigma_1 (\vec{w} - \vec{\mu_1})\right) & (6.22) \\ \text{and } \frac{\partial}{\partial \vec{w}} P_2 &= -A_2 (\vec{w} - \vec{\mu_2})^T \Sigma_2 \exp\left(-(\vec{w} - \vec{\mu_2})^T \Sigma_2 (\vec{w} - \vec{\mu_2})\right) & (6.23) \end{aligned}$$

The feature-space position of each point in the image is determined by the multi-valued image, $\vec{w} = \vec{F}(\vec{x})$, so the metric $\phi(\vec{F}(\vec{x})) = \vec{\gamma}(\vec{F}(\vec{x}))$ is

$$\begin{aligned} \vec{\gamma}(\vec{F}(\vec{x})) = \; - \; & \Big(A_1 A_2 (\vec{F}(\vec{x}) - \vec{\mu}_1)^T \Sigma_1 - (\vec{F}(\vec{x}) - \vec{\mu}_2)^T \Sigma_2\Big) \\ & \Big(A_1 \exp(-(\vec{F}(\vec{x}) - \vec{\mu}_1)^T \Sigma_1 (\vec{F}(\vec{x}) - \vec{\mu}_1)) & (6.24) \\ & + A_2 \exp(-(\vec{F}(\vec{x}) - \vec{\mu}_2)^T \Sigma_2 (\vec{F}(\vec{x}) - \vec{\mu}_2))\Big)^{-2} \\ & \times \exp\Big(-(\vec{F}(\vec{x}) - \vec{\mu}_1)^T \Sigma_1 (\vec{F}(\vec{x}) - \vec{\mu}_1) & (6.25) \\ & -(\vec{F}(\vec{x}) - \vec{\mu}_2)^T \Sigma_2 (\vec{F}(\vec{x}) - \vec{\mu}_2)\Big) \end{aligned}$$

For each point in the image space, this equation describes the local coordinate transformation in the feature space which is used to compute the dissimilarity. These vectors are perpendicular to the family of curves $Q(\vec{w}) - \kappa = 0$ defined by

$$Q(\vec{w}) - \kappa = (\vec{w} - \vec{\mu_1})^T \Sigma_1 (\vec{w} - \vec{\mu_1}) - (\vec{w} - \vec{\mu_2})^T \Sigma_2 (\vec{w} - \vec{\mu_2}) - \kappa = 0. \quad (6.26)$$

These curves are quadratics. Figure 6.12 illustrates the situation at a single point $\vec{x}_0$ in a two-dimensional image which is described by two features. The gradient of $\vec{F}(\vec{x})$ can be represented as a pair of vectors V_x, V_y existing at a point $\vec{w}_0 = \vec{F}(\vec{x}_0)$ in the feature space. Through that point, there is a manifold (quadratic) defined by Eq. (6.26). Perpendicular to that manifold is a vector representing the metric defined by Eq. (6.24). This metric is used to measure the "size" of the vectors V_x and V_y. The root-sum-squared length of these vectors becomes the dissimilarity an is used to control the conductance in the multi-valued diffusion equation.

6.2.3 Generalizing to more classes

The approach of Sect. 6.2.1 generalizes to more than two pixel classes after resolving one important question. The conditional probability for each class changes by a different amount. How does one combine each of these differences in order to obtain a single dissimilarity measure?

Suppose a point $\vec{x}_0 \in I$ belongs to a class S_j. Then the dissimilarity should reflect the way in which the probability $p(C_1|\vec{F}(\vec{x}_0))$ varies (within a neighborhood in I) from class S_j to any of the other classes. This probability shifts most dramatically in a single direction given by the gradient of this conditional probability. The magnitude of that gradient is the amount of variation in the probability of class S_j. However, the image is not yet segmented, so it is impossible to say that $\vec{x}_0 \in S_j$ with complete certainty. One approach to combining information from a set of class

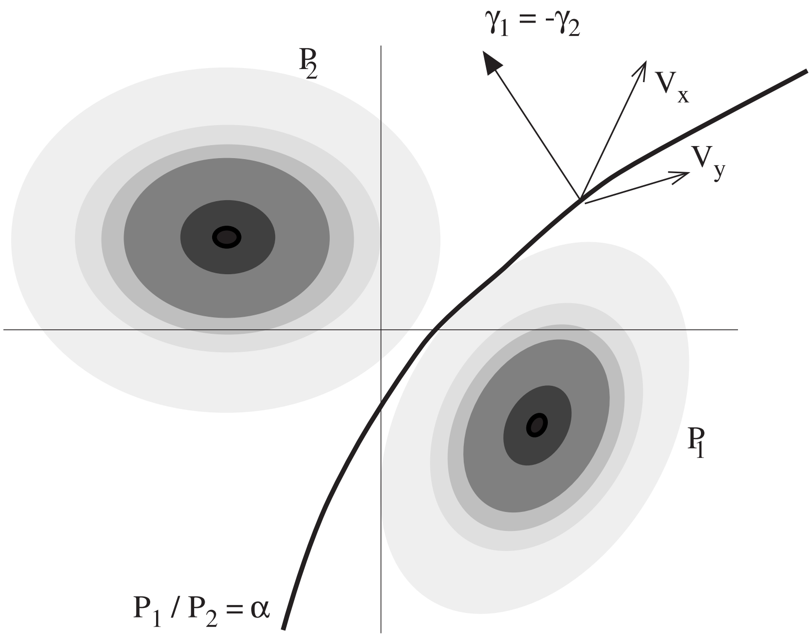


Figure 6.12: Dissimilarity is computed in the direction of vectors that are perpendicular to curves of equiproportionality ($P_1/P_2 = \alpha$).

distributions is to compute an *expected value* in the change of probability by computing a weighted sum across all of the classes.

For each of l classes, the conditional probability is

$$\begin{aligned} p(C_i|\vec{w}) &= \frac{p(\vec{w}|C_i)p(C_i)}{p(\vec{w})} \\ &= \frac{p(\vec{w}|C_i)p(C_i)}{\sum_{j=1}^{l} p(\vec{w}|C_j)p(C_j)} \\ &= \frac{P_i(\vec{w})}{\sum_{j=1}^{l} P_j(\vec{w})}. \end{aligned} \tag{6.27}$$

The first derivative of this conditional probability (with respect to feature space positions) is

$$\vec{\gamma}_i(\vec{w}) \equiv \frac{\partial}{\partial \vec{w}} p(C_i|\vec{w}) = \frac{\sum_{j \neq i} P_j \frac{\partial}{\partial \vec{w}} P_i - P_i \sum_{j \neq i} \frac{\partial}{\partial \vec{w}} P_j}{(\sum_{j=1}^{l} P_j)^2}. \tag{6.28}$$

Recall that the gradient of the conditional probability is

$$\nabla p(C_i|\vec{F}(\vec{x})) = \frac{\partial p(C_1|\vec{F}(\vec{x}))}{\partial \vec{w}} \cdot \nabla \vec{F}(\vec{x}) = \vec{\gamma}_i(\vec{w}) \cdot \nabla \vec{F}(\vec{x}). \tag{6.29}$$

The magnitude of this vector (l_2 norm) is

$$\left| \frac{\partial p(C_i|\vec{F}(\vec{x}))}{\partial \vec{x}} \right|_2 = \left((\vec{\gamma}_i(\vec{w}) \cdot \frac{\partial \vec{F}(\vec{x})}{\partial x_1})^2 + \cdots + (\vec{\gamma}_i(\vec{w}) \cdot \frac{\partial \vec{F}(\vec{x})}{\partial x_n})^2 \right)^{\frac{1}{2}}. \tag{6.30}$$

The magnitudes of the $(\partial/\partial\vec{x})p(C_i|\vec{F}(\vec{x}))$ terms for all of the classes must be combined in order to produce the dissimilarity at $\vec{x}$. The measure in 6.30 gives the desired metric in the event that the point in question belongs to class C_i. Because the class is not known, it is appropriate to use the *expected value* of this measure over the l possible classes.

$$\sum_{i=1}^{l} p(C_i|\vec{F}(\vec{x})) \left| \frac{\partial p(C_i|\vec{F}(\vec{x}))}{\partial \vec{x}} \right|_2 \tag{6.31}$$

$$= \sum_{i=1}^{l} \frac{P_i(\vec{F}(\vec{x}))}{\sum_{j=1}^{l} P_j} \left((\vec{\gamma}_i(\vec{w}) \frac{\partial \vec{F}(\vec{x})}{\partial x_1})^2 + \cdots + (\vec{\gamma}_i(\vec{w}) \frac{\partial \vec{F}(\vec{x})}{\partial x_n})^2 \right)^{\frac{1}{2}}, \tag{6.32}$$

were the $|\cdot|_2$ indicates the l_2 norm on a vector. Equation 6.31 is the dissimilarity measure. In order to be consistent, express this measure in terms of the gradient of $\vec{F}$ times the feature space transformation ϕ,

$$\phi(\vec{F}(\vec{x})) = \begin{bmatrix} \frac{P_1}{\sum_{j=1}^{l} P_j} \vec{\gamma}_1 \\ \vdots \\ \frac{P_l}{\sum_{j=1}^{l} P_j} \vec{\gamma}_l \end{bmatrix} \quad \text{and} \quad \mathcal{D}\vec{F} \equiv ||\phi \nabla \vec{F}||_{2,1} \tag{6.33}$$

where the probabilities and depend on $\vec{F}(\vec{x})$ so that $P_i = P_i(\vec{F}(\vec{x}))$ and $\vec{\gamma}_i = \vec{\gamma}_i(\vec{F}(\vec{x}))$. The matrix norm $||\cdot||_{2,1}$ indicates that the l_2 norm (root sum of squares) is used within the rows, and the l_1 norm is used to sum the resulting column. For a matrix M with elements M_{ij} this norm is

$$||M||_{2,1} \equiv \sum_i (\sum_j M_{ij}^2)^{\frac{1}{2}}.$$

6.2.4 Remarks

The theoretical foundation described in this section promises to be an interesting technique for classifying pixels in images. This technique has several useful properties.

First, the dissimilarity as defined in Eq. (6.33) is a boundary measure which depends on the statistical structure of a feature space. It tends to produce large measurements where local variations in the image correspond to variations between classes in the feature space and not within those classes. Assuming that the probabilities $p_1, \ldots, p_l$ incorporate the uncertainty in the feature values as of result of noise, then *this boundary measure reflects expected noise characteristics of images.* In cases where the noise is substantial, these probability density distributions have large variance measures, and the transformation ϕ is small (its elements are small). In that case large image gradients $\nabla \vec{F}$ are necessary to produce large dissimilarity values. Tighter class distributions produce larger dissimilarity values for the same $\nabla \vec{F}$.

Second, the analysis of dissimilarity in the case of known probability distributions does not preclude the possibility that this technique could be applied to features that are derived from the local structure of a single-valued image. This could apply to the geometric feature spaces of Chapter 4 as well as the "spectral" feature space described in Section 6.1. This analysis represents a general technique for constructing dissimilarity measures on the basis of the statistical properties of known classes, but assumes nothing about the origin of the features or the classification scheme.

Finally, the presence of "pre-segmented" images for computing the classifications might not be necessary. It is reasonable that "statistics" could be computed on images using some automated but less definitive classification scheme. As an example consider the following segmentation strategy. First, analyze a multi-valued image to find tentative locations and sizes of objects. This analysis would be rough, and it could result from some extensively blurred version of the image. Then, treat each object as a class and compute the appropriate statistics. Finally, use these statistics in the supervised clustering strategy outlined here in order to find more precise locations and boundaries. The result is a segmentation strategy that adapts to the noise characteristics of individual images.