UCGE Reports
Number 20174

Department of Geomatics Engineering

Guided Super-Resolution

(URL: http://www.geomatics.ucalgary.ca/links/GradTheses.html)

by

Michael de Jong

July 2002

THE UNIVERSITY OF CALGARY

Guided Super-Resolution

by

Michael de Jong

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA JULY, 2002

ABSTRACT

The image formation process by a remote sensor is modeled, highlighting restrictions on spatial resolution and image formation artifacts. The super-resolution problem is introduced, describing the aim of overcoming the resolution and image quality limits imposed by the sensor optics and sampling. Problems with model inversion and conventional image restoration techniques are discussed.

The focus of this thesis is a Super-Resolution technique in which image geometry is used to guide the resampling. The dynamics considered are those of curvature flow, in which image restoration – including Fourier mode synthesis, noise-supression, and anti-aliasing – is achieved my minimzing the mean curvature of an image. Following a discussion of the curvature flow switch introduced by Malladi and Sethian, we cast their work into the neural regime, identifying a similarity to the work of Tatem et al.

The performance of the switch is compared to the performance of polynomial interpolation. Testing revealed excellent performance: the super-resolved images are sharp with high signal to noise ratio relative to polynomial-interpolated images. The peak signal-to-noise ratio is typically greater than that of polynomial-interpolated images by ten to thirty percent. The advantage in many spatial frequency bands – often the highest frequency bands – reaches several hundred percent.

Contents

List of Tables

List of Figures

Acknowledgements

For high guidance, support, and patience in this work I would like to than my supervisor, Dr. Michael Collins; for their leadership and support in an excellent project, Dr. Michael Davenport and Vicky Ressl at MacDonald, Dettwiler, and Associates. Thanks to Dr. Isabelle Couloigner, Dr. Richard Klukas, and Mr. Norman Bartley for their valuable input, criticism, and willingness to help. Thanks to Dr. Harold Zwick, Dr. Norman Goldstein, Nathan Brunner, Gabe Sentlinger, Lu-Anne Schaffland, and all those who have contributed help and insight in this work.

Chapter 1

Introduction

The ability to detect very small objects with a remote camera has been the goal of many organizations and a source of fascination to the general public. With such an ability, intelligence agencies, academic research programmes, and commercial organizations have the capacity to gather vital information for their purposes.

The governments and intelligence community is primarily interested in locations and identities of objects in the field of view. The available information, determined in part by image quality, is essential to assessment of military threats, studying land use, etc.

The academic community in various disciplines – environmental science, robotics, astronomy and astrophysics – rely on remotely sensed imagery to conduct their research. The quality of the imagery is often a severe restriction on their experimental precision.

Remotely-sensed imagery plays an important role in industry, for example, in locating mineral deposits, in urban planning and land management, in cartography, and in agriculture. In these disciplines, too, image quality imposes restrictions upon the scope and detail of a project.

The dependence on quality imagery is the common thread in these disciplines and fuels the development of improved image acquisition and enhancement techniques.

The quality of the recorded image depends on hardware characteristics: for example, the quality of the sensor optics, and the platform distance and velocity relative to the subject. Improvements to any of these characteristics typically come with large financial expense. In addition, image quality depends upon environmental conditions such as the atmosphere's coefficient of transmission, its activity, etc. Improving these conditions is typically impossible.

As an alternative, software image restoration is not only possible, it is relatively cheap. The image restoration problem is not altogether simple, however. The complications are consequences of the image formation process. In this thesis we discuss several restoration techniques in order to restore salient features in the sensor's instantaneous field of view. In this chapter we provide a brief outline of the content of the thesis, beginning with image formation.

1.1 Image Formation

In recording an image, a sensor collects light incident upon an aperture – the object signal – and focuses the light onto the image plane where it is sampled. These processes introduce distortions into the object signal, including blurring by diffraction and lensing, decimation and aliasing from sampling, and noise from optics and electronics. Most notably, decimation is a consequence of sampling a continuous function to discrete data, whereby information is typically lost. A feature smaller than the sample spacing is not discernible in the image; it is averaged with the backgound and aliased into lower frequencies. Thus, the image formation process is especially severe in the highest frequencies. In Chapter 2 we will discuss these processes in detail.

1.2 Image Restoration

Given a digital image, our goal is to devise a means of restoring the object signal. We aim to reverse decimation effects to recover features not represented in the digital image. Our biggest problem is the recovery of high Fourier modes – image details – lost in the image formation process. Also, we aim to clarify existing image features by correcting blurring, aliasing, and noise. In Chapter 3 we describe various restoration techniques in detail. We focus our attention on energy minimization – a framework for implementing robust and novel restoration techniques. We discuss interpretations of the dynamics we derive and find that they bear strong resemblance to the diffusion of a gas, the dynamics of a soap bubble, and the propagation of signals in a neural network.

With a well-defined image formation model, we might begin our restoration attempt by inverting the model. The image formation process takes a continuous function and samples it to a discrete set of data; the inverse process, if it exists, does the reverse. However, one can fit an infinite set of functions to a finite collection of data: the inverse does not exist. The inversion problem, and consequently the restoration problem, is ill-defined.

1.2.1 Underdetermination

Wherever the image formation model has zeros, the model may not be meaningfully inverted at these frequencies: the zeros diverge to infinity. This is the infinite family of functions which one may fit to the data; it implies that any member of an infinite family may have given rise to the digital image.

The restoration problem becomes a problem of selecting an appropriate member of the infinite family of solutions; an appropriate member of the family is an image that resembles the object signal. The family members from which we have to choose differ in the highest frequencies – the precise locations of edges, for example. The underdetermination of the high frequencies is apparent in digital images: cars may not be visible on roads; trees may not be distinguished. To select an appropriate member of the infinite family, we need a mechanism for synthesizing high spatial frequencies.

High spatial frequencies may be synthesized by supplying additional information. The additional information is a set of assumptions about the object signal, such as ‘the object signal is smooth’, or ‘edges appear as gradients’, or ‘the object signal is well-approximated by polynomial at scale z.’ This additional information is often referred to as *a priori* information or simply as *priors*.

1.2.2 Resampling

Conventional resampling techniques synthesize Fourier modes by fitting a function to the data; the basis functions are the prior information. Low resolution data are used to set the coefficients or parameters of the functions. The high Fourier modes are generated by evaluating the function at fine mesh points.

Instead of casting geometric quantities appearing in an image into function-analytic terms, these quantities might serve as direct guides for resampling. To guide the resampling by geometry, we take an iterative approach: we adjust pixel values according to an assessment of local image geometry. For example, we might make the adjustment to a pixel proportional to the local image gradient. With a shape-based approach, the geometry evolves; the data direct themselves toward a final configuration. The dynamics are governed by a differential equation and are analogous to physical systems like water freezing or a soap bubble collapsing.

Recovery of Fourier modes is not the entire restoration problem: it should involve the correction of blurring, aliasing, and noise. To correct these artifacts, the source data must be adjusted while restoring high frequency content. We distinguish super-resolution techniques from interpolation techniques on the basis of the treatment of source data. Interpolation techniques do not allow the correction of image formation artifacts while restoring Fourier modes; super-resolution techniques adjust the source data while inferring decimated Fourier modes. That is, in a super-resolution technique all data may be subject to evolution.

1.2.3 Energy Minimization

Correction of image formation artifacts and the employment of image geometry to guide resampling may be difficult to implement with function-fitting methods. As an alternative, energy minimization techniques may be implemented. In these techniques, the prior information is encoded in an energy functional; the minimum of the functional corresponds to the satisfaction of the prior criteria. Thus, the energy may be the difference between the priors and the current state of the image.

With energy minimization we have a means of employing non-trivial priors and of establishing a compromise among competing priors; the implementation of shape-based methods is straightforward. However, convergence, speed, and stability are issues whenever the minimization technique is iterative.

1.2.4 Switches

Complications arise in assessing image geometry: a black circle on a white background has the same geometry as a white circle on a black background. Assessing the geometry by making gradient or curvature measurements yields a relative negative sign. The result is opposite dynamics. Thus, a naïve assessment of image geometry may lead to mis-directed flow.

To encourage geometrically equivalent shapes to evolve identically, Malladi and Sethian (1) have introduced a switch. We have noticed that improvements may be made to their switch by casting it in a neural framework; our contribution is the generalization of Malladi and Sethian's switch to a neural regime. There, we have found a similarity between the work of Malladi and Sethian and of Tatem et al (2).

1.3 Testing

We have tested the performance of the generalized switch by resampling remotely sensed imagery; we have compared it with the performance of a polynomial interpolant. Our testing was conducted on a variety of images, including reflectance images and material abundance images derived from hyperspectral imagery. In Chapter 4 we outline the testing procedure and the measures of performance.

We have found encouraging results, presented in Chapter 5: the generalized switch exhibits excellent performance in the majority of tests. In particular, the fuzzy switch reliably produces images with 10% to 30% higher peak signal-to-noise ratio than polynomial resampling. The highest spatial frequencies are effectively restored, implying that the switch is excellent at resolving fine image details.

Chapter 2

Image Formation

In this chapter we describe, in detail, the image formation process by a digital system. We highlight the spatial distortions introduced by various components of the system: in particular, we discuss blurring by impulse response convolution, decimation and aliasing by sampling, and noise from sensor optics and and electronics.

2.1 The Sensor Reflectance

A remotely-sensed image is formed by recording light which has been reflected from the earth's surface. The amount of light illuminating the surface depends on the sun's radiant flux, the distance of the surface from the sun, atmospheric attenuation, and the orientation of the surface relative to the wavevector (3) (Figure 2.1). In this section we introduce a model of the digital image formation process.

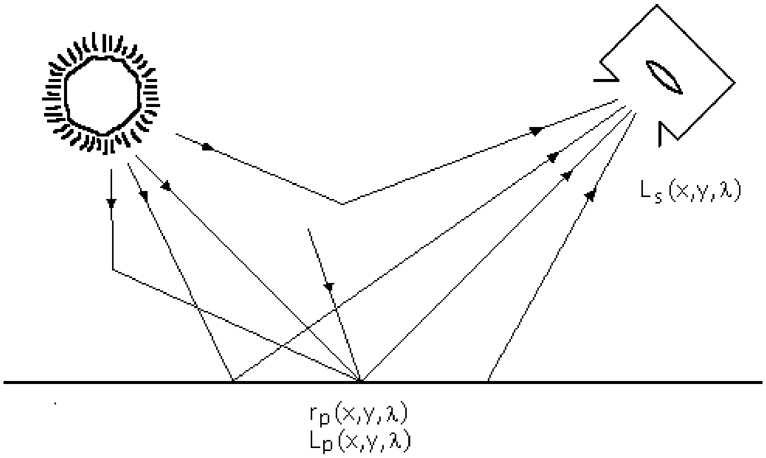


Figure 2.1: The sensor radiance

Suppose the surface of the earth is illuminated by a radiance $L_e(x, y, \lambda)$. The amount of reflected light – the pixel-leaving radiance $L_p(x, y, \lambda)$ – is determined by the earth's reflectance $r_p(x, y, \lambda)$, defined by

$$r_p(x, y, \lambda) = \frac{L_p(x, y, \lambda)}{L_e(x, y, \lambda)} \tag{2.1}$$

which implies

$$L_p(x,y,\lambda) = r_p(x,y,\lambda) L_e(x,y,\lambda). \tag{2.2}$$

The pixel-leaving radiance propagates through the atmosphere and arrives at the sensor; at this point it is called the *sensor radiance* L_s,

$$L_s(x,y,\lambda) = c(\lambda) L_p(x,y,\lambda) + n_a(x,y,\lambda). \tag{2.3}$$

We've modeled atmospheric effects to linear order, where c and n_a describe multiplicative and additive distortions to the pixel-leaving radiance; see Appendix A for a justification.

Dividing equation 2.3 by L_e gives

$$\frac{L_s(x,y,\lambda)}{L_e(x,y,\lambda)} = c(\lambda) r_p(x,y,\lambda) + \frac{n_s(x,y,\lambda)}{L_e(x,y,\lambda)} \tag{2.4}$$

We define the apparent reflectance by

$$r_s = \frac{L_s}{L_e} \tag{2.5}$$

and we treat n_s/L_e as rescaled noise,

$$n_s = \frac{n_1}{L_e} \tag{2.6}$$

giving

$$r_s(x,y,\lambda) = c(\lambda) r_p(x,y,\lambda) + n_s(x,y,\lambda) \tag{2.7}$$

We assume that the reflectance, r_p, is a superposition of the reflectance of the materials at that point,

$$r_p(x,y,\lambda) = \sum_i m_i(\lambda) \alpha_{pi}(x,y) \tag{2.8}$$

which implies

$$r_s(x,y,\lambda) = c(\lambda) \sum_i m_i(\lambda) \alpha_{pi}(x,y) + n_s(x,y,\lambda) \tag{2.9}$$

We set

$$m_{si}(\lambda) = c(\lambda) m_i(\lambda) \tag{2.10}$$

which gives

$$r_s(x,y,\lambda) = \sum_i m_{si}(\lambda) \alpha_{pi}(x,y) + n_s(x,y,\lambda) \tag{2.11}$$

Before proceeding, we note that the abundance of each pixel should be normalized:

$$\sum_i \alpha_i = |\alpha(x,y)| = 1; \tag{2.12}$$

For example, a pixel cannot be merely 20% grass and 30% asphalt; it must be 100% *something*; the remaining 50% must be other materials. This constraint will be imposed when resampling abundance imagery.

2.2 Impulse Response

The sensor radiance is incident upon an aperture (Figure 2.2) and lens which collect and focus the sensor reflectance (the object signal), r_s onto the image plane as the image signal, r_I. This process is described by the impulse response.

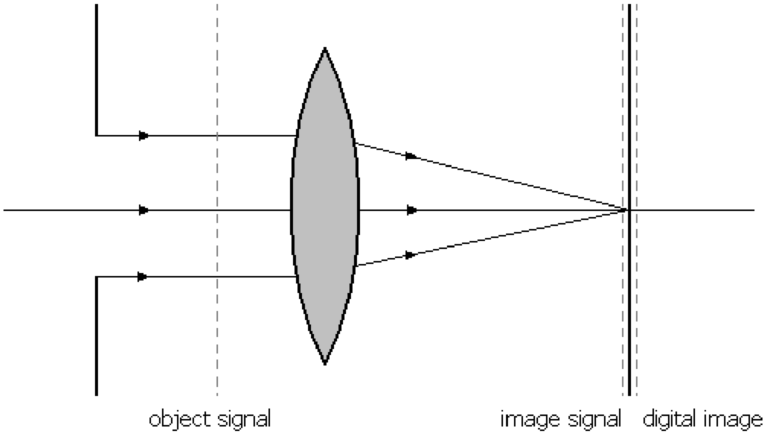


Figure 2.2: Digital image formation

To characterize the sensor impulse response, h_1, we note that each point of the image signal, treated as a function r_I, is the sum of the infinitesimal contributions from all points of the sensor radiance r_s:

$$r_I(x,y,\lambda) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_1(x\prime,y\prime,x,y;r_s(x,y,\lambda))dxdy \tag{2.13}$$

where (x,y) is a point in the object plane and where $(x\prime,y\prime)$ is a point in the image plane. We have assumed that the impulse response is independent of color − we assume chromatic aberration is absent, $h_1 = h_1(x,y;L_s)$.

Following Jiang (4) we note that if the image formation system is linear, i.e. if $h_1(x\prime,y\prime,x,y;r_s(x,y,\lambda)) = h_1(x\prime,y\prime,x,y)\cdot r_s(x,y,\lambda)$, and if h_1 acts uniformly across the object plane (the value of h_1 depends only on the separations $x\prime - x$ and $y\prime - y$),

$$h_1(x\prime,y\prime,x,y) = h_1(x\prime - x, y\prime - y),$$

then image formation is a convolution:

$$r_I(x\prime,y\prime,\lambda) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_1(x\prime - x,y\prime - y)r_s(x,y,\lambda)dxdy.$$

2.3 Impulse Response Effects

The impulse response, h_1, is a composition of functions; its form is peculiar to each sensor. However, spatial resolution is diffraction limited at the very least, a feature common to all high quality sensors (3), (5). The diffraction limit is imposed by the size of the sensor aperture and the magnification of the focusing lens(es); additional optical components such as prisms and gratings may be included for dispersion in spectrometers. A characteristic impulse response convolution effect is the blurring of edges as in Figure 2.3.

Our model for the impulse response in the spatial domain is a Gaussian (Figure 2.4),

$$h_1(x,y) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma}(x^2+y^2)} \tag{2.14}$$

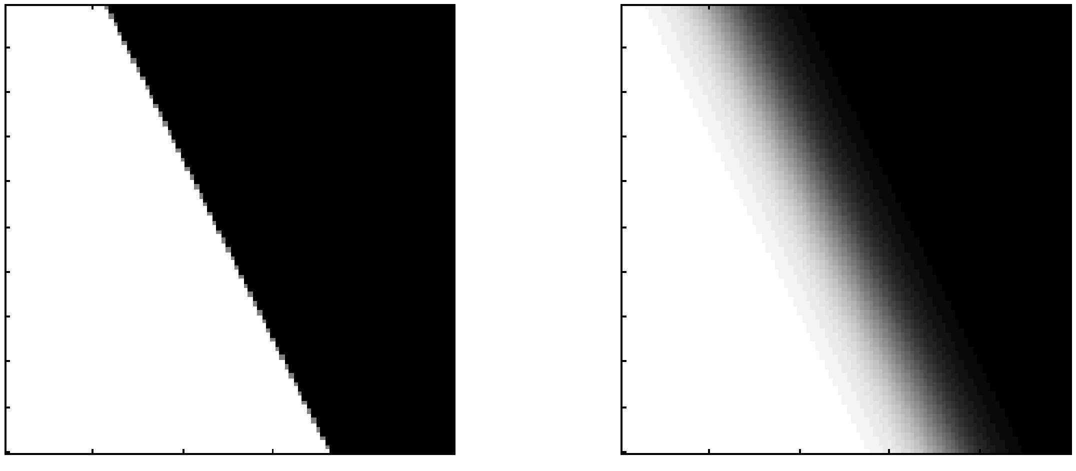


Figure 2.3: Edge blurring due to impulse response convolution

Because its Fourier transform is also a Gaussian, all Fourier modes of the object signal appear in the image signal: *the image signal is not bandlimited.*

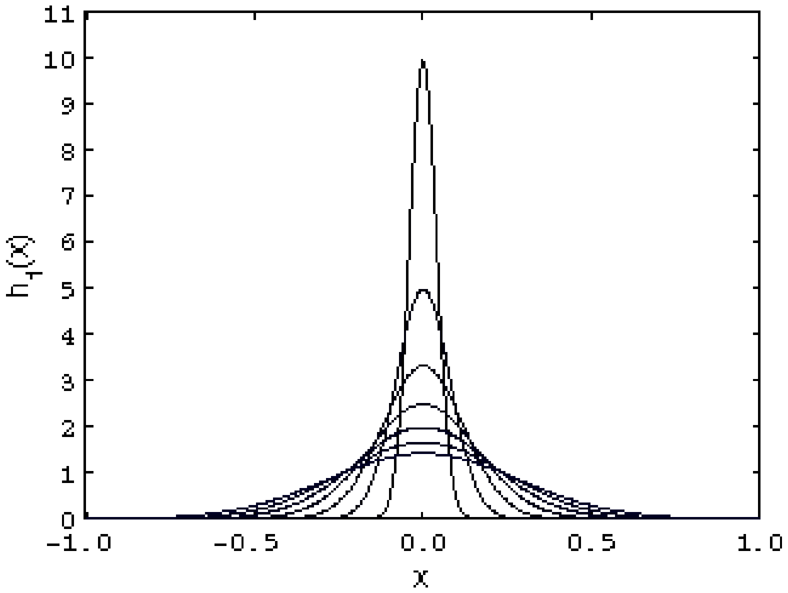


Figure 2.4: Scaled Gaussian at various standard deviations

Although high-frequency object signal modes are not lost in the convolution, to ensure that objects may be distinguished from one another, diffraction limits resolution by restricting their separation (Figure 2.5). Consider two point sources in one dimension, $\delta(x_1 - x)$ and $\delta(x_2 - x)$. The sensor reflectance is $r_s(x) = \delta(x_1 - x) + \delta(x_2 - x)$ and their image is

$$r_I(x) = h_1 \circ [\delta(x_1 - x) + \delta(x_2 - x)] = h_1(x_1 - x) + h_1(x_2 - x). \tag{2.15}$$

To distinguish the point sources we require that the amplitude at the midpoint must not exceed some fraction of the Gaussian amplitude. That is, to resolve the sources we require

$$r_I((x_1 + x_2)/2) \le \frac{a}{\sigma\sqrt{2\pi}} \tag{2.16}$$

for a suitable choice of $a \in (0, 1)$. Then, our limit is

$$\frac{a}{\sigma\sqrt{2\pi}} = \frac{1}{\sigma\sqrt{2\pi}} \left[e^{\frac{(x_1-x)^2}{2\sigma^2}} + e^{\frac{(x_2-x)^2}{2\sigma^2}} \right] \Bigg|_{x=\frac{x_1+x_2}{2}} \tag{2.17}$$

which implies

$$a = e^{\frac{(x_1-x_2)^2}{8\sigma^2}} + e^{\frac{(x_2-x_1)^2}{8\sigma^2}} \tag{2.18}$$

and

$$x_2 - x_1 = \sqrt{8\sigma^2 \ln \frac{a}{2}}. \tag{2.19}$$

Thus, if we enforce our criterion for distinguishing two point sources, then their separation in the object plane must exceed $\sqrt{8\sigma^2 \ln \frac{a}{2}}$. This resolution limit is a consequence of the finite breadth of the impulse response, rather than a (nonexistent) bandlimit.

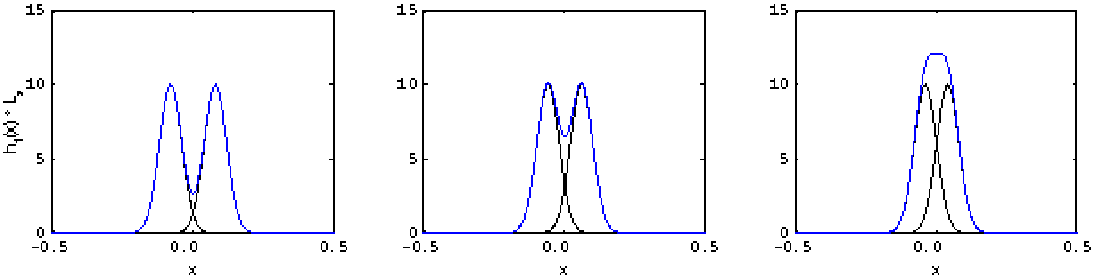


Figure 2.5: Resolvable point-sources (panel 1), just-resolvable point sources (panel 2), and not-resolvable point sources (panel 3)

Therefore, the resolution of diffraction limited optics is determined by the sensor geometry and the quality of the optics. The sampling process, however, imposes an additional restriction on the resolution of the digital image. Its restriction is a consequence of its Fourier domain bandlimiting.

2.4 Data Sampling

Crochiere et al (6) note that from a continuous function, $r_I(x, y, \lambda)$, we can define a set of samples $\mathbf{r}(x, y)^1$ specified by sampling processes

$$\lambda_i = o(x) \quad x_i = p(x) \quad y_j = q(y).$$

The sampling is uniform (periodic) if

$$o(\lambda) = \delta_{\lambda+\lambda_0 \bmod \Lambda} \quad p(x) = \delta_{(x+x_0) \bmod X} \quad q(y) = \delta_{(y+y_0) \bmod Y}$$

---

[1] We treat the samples r as a continuous function, nonzero only at sample points

for arbitrary offsets λ_0, x_0, and y_0 where

$$\delta_{s \bmod S} = \begin{cases} 1, & \text{if} s \bmod S = 0 \\ 0, & \text{otherwise} \end{cases}$$

for sampling periods Λ, X, and Y. The sampling rates conjugate to Λ, X, and Y are

$$\lambda_s = \frac{1}{\Lambda} \quad u_s = \frac{1}{X} \quad v_s = \frac{1}{Y}. \tag{2.20}$$

The image recorded by this process is

$$\mathbf{r}(x, y) = \delta_{\lambda \bmod \Lambda} \delta_{x \bmod X} \delta_{y \bmod Y} r_I(x, y, \lambda) \tag{2.21}$$

where we have set the offsets to zero.

The sampling aperture, however, is finite; we convolve the infinitesimal aperture, $\delta_{\lambda \bmod \Lambda} \delta_{x \bmod X} \delta_{y \bmod Y}$ with a boxcar function:

$$\delta_{\lambda \bmod \Lambda} \delta_{x \bmod X} \delta_{y \bmod Y} \rightarrow \delta_{\lambda \bmod \Lambda} \delta_{x \bmod X} \delta_{y \bmod Y} * h_2 = h_3 \tag{2.22}$$

where

$$h_2(x, y, \lambda) = \begin{cases} 1 & \text{if } \ x, y, \lambda < a \\ 0 & \text{otherwise} \end{cases} \tag{2.23}$$

where $2a$ is the width of the charge-coupled device (CCD). The image signal is

$$\begin{aligned} \mathbf{r}(x, y) \quad &= h_3(x, y, \lambda) r_I(x, y, \lambda) \\ &= h_3(x, y, \lambda) h_1(x, y) * r_s(x, y, \lambda) \\ &\equiv h(x, y, \lambda) * r_s(x, y, \lambda) \end{aligned}$$

where we have defined the image formation function, $h = h_3 h_1$.

2.5 Sampling Effects

The Shannon Sampling Theorem assures us that any bandlimited signal can be represented by a discrete sequence of its samples without loss of information, provided the sampling rate meets or exceeds the signal's Nyquist frequency (7). Thus, we can record all the information of r_r in digital form if we sample it at or above the Nyquist frequency.

The digital image Fourier domain of support is bounded by zero and half the sampling frequency in each dimension. On the other hand, the image signal is not bandlimited. As a result, the digital image is inequivalent to the image signal.

The Fourier transform, $\mathbf{R}(u, v)$, of $\mathbf{r}(x, y)$ is

$$\mathbf{R}(u, v) = H(u, v, \lambda) R_s(u, v, \lambda) \tag{2.24}$$

where H and $\mathbf{R}_s$ are the Fourier transforms of h and r_s.

If H is nonzero on a finite domain, then $\mathbf{R}$ is bandlimited (8). We will assume that $\mathbf{R}$ is bandlimited by

u_N and v_N. Formally,

$$\max\{u : ||\mathbf{R}(u,v)|| \neq 0\} = \max\{u : ||H(u,v,\lambda)R(u,v,\lambda)|| \neq 0\} \tag{2.25}$$

and similarly for v. We assume that the object signal is not bandlimited so that

$$\max\{u : ||\mathbf{R}(u,v)|| \neq 0\} = \max\{u : ||H(u,v,\lambda)|| \neq 0\}. \tag{2.26}$$

Because $H(u,v,\lambda) = H_1(u,v,\lambda)H_2(u,v,\lambda)$ we have

$$\max\{u : ||\mathbf{R}(u,v)|| \neq 0\} = \max\{u : ||H_1(u,v,\lambda) * H_2(u,v,\lambda)|| \neq 0\}. \tag{2.27}$$

and because the impulse response is not bandlimited,

$$\max\{u : ||\mathbf{R}(u,v)|| \neq 0\} = \max\{u : ||H_2(u,v,\lambda)|| \neq 0\}. \tag{2.28}$$

The maximum frequency of the sampling function is half the sampling rate:

$$\max\{u : ||H_2(u,v,\lambda)|| \neq 0\} = u_s/2 \tag{2.29}$$

so

$$\max\{u : ||\mathbf{R}(u,v)|| \neq 0\} = u_s/2. \tag{2.30}$$

Therefore, the digital image is bandlimited by half the sampling rate.

Sub-pixel image features are not distinguishable and, in fact, all image signal modes above the bandlimits are not properly represented. Fourier modes outside the sampling bandlimits, however, are not truncated: they are aliased. At large sample spacing, the sampling frequency is large; thus the convolution in the frequency domain results overlapping waveforms. This overlapping is aliasing: it is the illegitimate contribution by high signal frequencies to lower frequencies upon sampling.

As an example of aliasing, consider a narcoleptic cop on a stake-out. His assignment is to monitor – from a window across the street – a mob-boss' activity. In fact, the cop's only responsibility is to record how frequently the mobster leaves his apartment. In actual fact, the mobster leaves his apartment three times daily: once at 8:00am for breakfast, once at 4:00pm for an update from his colleagues on his crime syndicate, and once at midnight for a hit, a frequency of 3 times per day. The cop's narcolepsy is severe and regular, however. During each spell he sleeps for twelve hours which is followed by six hours of wakefulness. Thus, the cop begins his assignment Monday morning at 6:00am, binoculars in hand. At 8:00 he sees the boss leave for breakfast, though at 11:00am he suffers a narcoleptic attack and sleeps until 11:00pm, during which the mob boss meets with his colleagues. At midnight when the cop is awake, he sights the mob boss leaving his apartment for a hit, though at 5:00am Tuesday morning the cop falls back to sleep, missing the mobster leaving at 8:00am for his breakfast. In this way the cop's beat progresses, and he misses the boss on Tuesday altogether. His next sighting is Wednesday afternoon at 4:00, followed by Thursday at 8:00am, Saturday at 4:00pm, Sunday at 8:00am, and so on. Throughout a two week period, the cop makes only 9 sightings of the boss leaving his apartment – a paltry 0.64 times per day on average. The mob boss' $3.47 \times 10^{-5} Hz$ activity is aliased as $6.43 \times 10^{-6} Hz$ activity.

An undersampled image can be recognized by the appearance of analogous distortions in space as illustrated in Figure 2.6. On a discrete mesh smooth curves, for example, cannot be adequately represented and appear jagged: some Fourier modes of curves exceed any sampling rate.

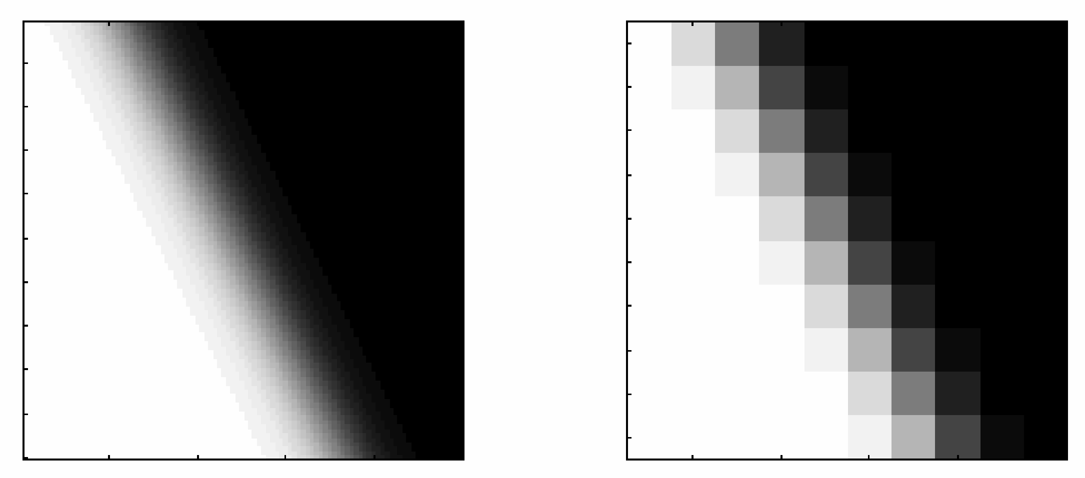


Figure 2.6: Aliasing in digital images; diagonals and curves appear jagged

Suppose the image signal, $r_I(x, y, \lambda)$, is not bandlimited. Then, image signal frequencies above $u_s/2$ and $v_s/2$ are aliased as frequency components on the domain $0 \leq u \leq u_s/2$, $0 \leq v \leq v_s/2$. For example, the component of $R_I(u, v, \lambda)$ at $u_s/2 + \Delta u$ is aliased as a component at $u_s/2 - \Delta u$ in the digital image $\mathbf{R}(u, v)$.

The illegitimate contribution of $R_I(u_s/2+\Delta u, v, \lambda)$ to $\mathbf{R}(u_s/2-\Delta u, v, \lambda)$ is inseparable from the legitimate component contributed by $R_I(u_s/2-\Delta u, v, \lambda)$ without prior knowledge of the image signal. Though filtering the image signal before sampling can eliminate aliasing effects, we are not in a position to apply such a filter in software restoration of a digital image.

Elimination with software from $\mathbf{R}(u_s/2 - \Delta u, v, \lambda)$ of $R_I(u_s/2 + \Delta u, v, \lambda)$ masquerading as $R_I(u_s/2 - \Delta u, v, \lambda)$, comes with the risk of eliminating legitimate $R_I(u_s/2-\Delta u, v, \lambda)$ contributions with the illegitimate $R_I(u_s/2 + \Delta u, v, \lambda)$ contribution.

The severity of aliasing artifacts depends on the amplitude of image signal modes above half the sampling rate. For example, let us suppose that aliasing artifacts are negligible if the sampling rate u_s is greater the frequency at which the amplitude of the Gaussian is one-tenth its maximum. This imposes a restriction on the Gaussian standard deviation, for the image signal amplitude is bounded by the impulse response amplitude. In one dimension the Fourier transform of equation 2.14 is

$$H(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\sigma^2 u^2}{2}} \tag{2.31}$$

Our aliasing restriction is

$$\frac{1}{10\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{\sigma^2 u_s^2}{2}} \tag{2.32}$$

which implies

$$\sigma = 2.14 u_s \tag{2.33}$$

Thus, as the sample spacing decreases (the sampling frequency increases) the standard deviation must increase to maintain negligible aliasing.

We note two difficulties with a large standard deviation. First, a large standard deviation implies a lot of blurring: in equation 2.15 we see that the breadth of the image of each point source increases with the impulse response standard deviation. Second, a broad impulse response implies a poor signal-to-noise ratio

(Fig 2.7): consider the one dimensional signal $r_s = \delta(x_1 - x)$. Its image is

$$r_I(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x_1-x)^2}{2\sigma^2}}. \tag{2.34}$$

The signal at the charged-coupled device (CCD) is

$$\mathbf{r}(0) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-w/2}^{w/2} dx \;\; e^{\frac{(x_1-x)^2}{2\sigma^2}} = \mathrm{erf}\left(\frac{x1-x}{\sigma\sqrt{2}}\right) \tag{2.35}$$

where w is the width of the chip, centered at $x = 0$. The function $\mathrm{erf}(x)$ is the error function, monotonically increasing with x. If n is the noise, then the signal to noise ratio is

$$SNR = \frac{erf\left(\frac{x1-x}{\sigma\sqrt{2}}\right)}{n}. \tag{2.36}$$

Thus, if the standard deviation increases, the signal-to-noise ratio decreases. Therefore, reduced aliasing artifacts come at the cost of a reduced signal-to-noise ratio.

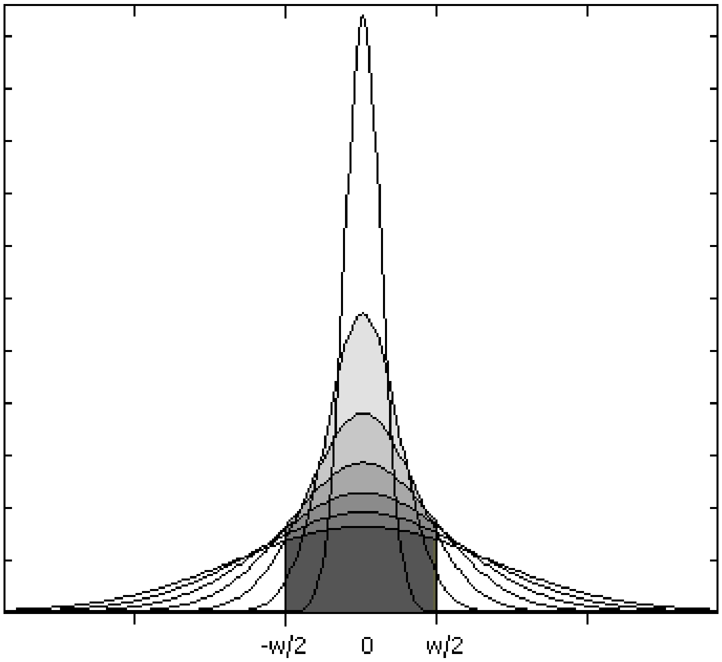


Figure 2.7: CCD SNR
The trade-off among signal-noise ratio, blurring, and aliasing: when the impulse response is broad, aliasing is low but the signal SNR is low also.

We can recognize aliasing artifacts as distortions of smooth curves. Just as the narcoleptic cop observed (illegitimate) $6.43 \times 10^{-6} Hz$ frequencies, we observe (illegitimate) spatial frequencies when the sampling rate is insufficient. The human visual system is able to recognize these illegitimate contributions because we have prior experience with smooth curves and their representation on a finite mesh. When we observe a jagged line in an image we may recognize it as the representation of a diagonal or as a circle. Encoding similar prior information into a computer program to make similar judgments is part of our goal.

We should augment the image formation described above with a noise term to represent errors in the data from sensor noise, etc. The image formed is

$$\mathbf{r}(x,y) = h(x,y,\lambda) * r_s(x,y,\lambda) + \mathbf{n}_2(x,y). \tag{2.37}$$

2.6 Reflectance and Abundance Domain Distortions

Note that h can be separated into two pieces, $h_{xy}(x,y)$ which blurs and samples in physical space (x and y coordinates), and $h_\lambda(\lambda)$ which samples in wavelength space (λ coordinate): $h(x,y,\lambda) = h_{xy}(x,y) * h_\lambda(\lambda)$. The image is

$$\mathbf{r}(x,y) = h_{xy}(x,y) * h_\lambda(\lambda) * r_s(x,y,\lambda) + \mathbf{n}_2(x,y). \tag{2.38}$$

which, with the expansion 2.11

$$\mathbf{r}(x,y) = h_{xy}(x,y) * h_\lambda(\lambda)) * \left[\sum_i m_{si}(\lambda)\alpha_{pi}(x,y) + n_s(x,y)\right] + \mathbf{n}_2(x,y). \tag{2.39}$$

h_λ converts the m_{si} from functions to vectors $\mathbf{m}_{si}$, while h_{xy} converts the α_i from functions to discrete samples in space. Ergo

$$\mathbf{r}(x,y) = h_{xy}(x,y) * \left[\sum_i \mathbf{m}_{si}(\lambda)\alpha_{pi}(x,y) + \mathbf{n}_s(x,y)\right] + \mathbf{n}_2(x,y). \tag{2.40}$$

and setting

$$\alpha_i(x,y) = h_{xy}(x,y) * \alpha_{si}(x,y) \tag{2.41}$$

we have

$$\mathbf{r}(x,y) = \sum_i \mathbf{m}_{si}(\lambda)\alpha_i(x,y) + \mathbf{n}(x,y) \tag{2.42}$$

where $h_{xy}(x,y) * \mathbf{n}_s(x,y) + \mathbf{n}_2(x,y) = \mathbf{n}(x,y)$. We note that h_{xy} is responsible for spatial blurring and aliasing; these artifacts thus appear in the abundance domain as well as in the reflectance domain. In the spatial domain the image formation processes, equations 2.37 and 2.41, and thus their distortive effects, differ only in noise.

2.7 Problem Statement

The sensor radiance samples are significantly distorted in the spatial domain: object signal features are blurred by the impulse response and the resulting image signal is bandlimited and aliased by sampling. Brightness edges, for example, appear as discretized gradients, and curves and diagonals are aliased.

Our primary requirement is an estimate of the object signal on a finer grid to distinguish signal details. This implies the requirement to generate an image with the following characteristics:

I-1 location of edges and discrete features are refined: blurring, aliasing, and noise are reduced (Figure 2.8)

I-2 the Fourier domain of support is extended to $[0, zu_N]$, $[0, zv_M]$ for $z > 1$, and

The degree to which these requirements are met should be appropriately quantified and compared with the results of an existing technique.

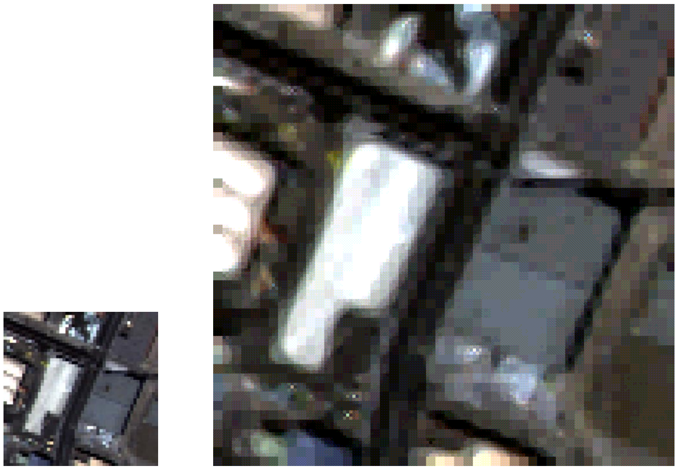


Figure 2.8: Resampled reflectance imagery

In addition, information-significant features of the object signal should be highlighted with supplementary images. Information-significant features may include such features as

II-1 boundaries between objects or materials (Figure 2.8),

II-2 the relative abundance of specified materials in each pixel, and

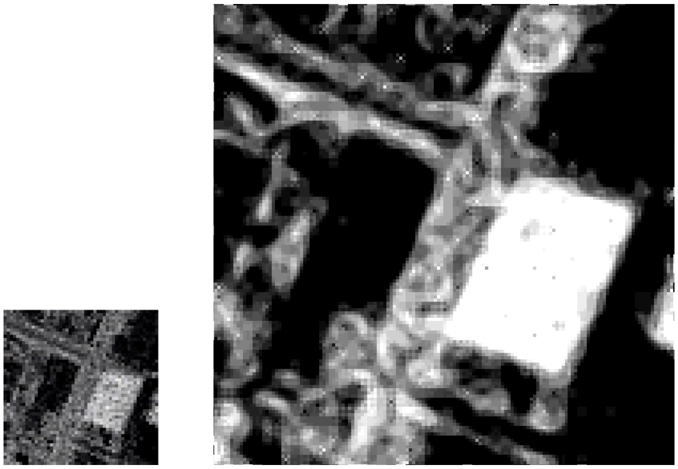


Figure 2.9: Resampled material abundance imagery: the building is highlighted in this image

Chapter 3

Model Inversion and Image Reconstruction

Frequency content restoration is required wherever $H(u,v)R_s(u,v) \ll R(u,v)$. With a well-defined image formation model, restoration may be attempted by sensor model inversions. This approach, however, proves problematic: at worst, the inverse impulse response is singular over the image frequency bands we aim to restore; most often the impulse response is nearly zero over those bands, making its inverse sensitive to aliasing and the quality of the noise model. The sampling function is always singular; if a zero sits on bands that we want to restore, the restoration is impossible by inversion.

3.1 Spectral Unmixing

If a resolution cell contains n materials, then we assume the mean reflectance is a linear combination of their individual reflectances

$$\mathbf{r}(x,y) = \mathbf{r}_1\alpha_1(x,y) + \mathbf{r}_2\alpha_2(x,y) + \ldots \mathbf{r}_n\alpha_n(x,y) \tag{3.1}$$

or in matrix notation

$$\mathbf{r}(x,y) = \mathbf{m}\alpha(x,y). \tag{3.2}$$

If the reflectivity spectrum r_i of each of the materials is known, then the abundance of each material may be estimated by inverting equation 3.2. That is materials may be detected and identified based upon their diagnostic spectral responses (9). The inversion, called unmixing, is a valuable means of displaying image data (10), while reducing the volume of data.

From the data $\mathbf{r}(x,y)$ we can recover abundance maps:

$$\alpha(x,y) = \mathbf{m}^{-1}(\mathbf{r}(x,y) - \mathbf{n}(x,y)) \tag{3.3}$$

This treatment is merely schematic because m may not be invertible.

With pseudo-inversion formalism we can estimate the abundance maps:

$$\mathbf{r}(x,y) = \mathbf{m}\alpha(x,y) + \mathbf{n}(x,y) \Rightarrow \mathbf{m}^T\mathbf{r}(x,y) = \mathbf{m}^T\mathbf{m}\alpha(x,y) + \mathbf{m}^T\mathbf{n}(x,y) \tag{3.4}$$

The matrix $\mathbf{m}^T\mathbf{m}$ is invertible so we have

$$\mathbf{m}^T\mathbf{m}\alpha(x,y) = \mathbf{m}^T(\mathbf{r}(x,y) - \mathbf{n}(x,y)) \tag{3.5}$$

$$\Rightarrow \alpha(x,y) = (\mathbf{m}^T\mathbf{m})^{-1}\mathbf{m}^T(\mathbf{r}(x,y) - \mathbf{n}(x,y)). \tag{3.6}$$

To improve unmixing performance, we may generalize the unmixing operator by using the Mahalanomis distance, d_m, to rescale length in λ-space:

$$d^2 = (r_i) - \bar{r}_i)^T\mathcal{V}_{ij}^{-1}(r_j - \bar{r}_j) \tag{3.7}$$

where $\mathcal{V}_ij$ is the covariance between the i^{th} and j^{th} bands and where $\bar{r}_i$ is the mean in the i^{th} band. With this metric, distances are short when the covariance is large. Thus, pixels which previously may have been identified as one material are now identified as another whose covariance is larger. The corresponding unmixing matrix is replaced:

$$(m^Tm)^{-1}m^Tm \rightarrow G = (m^T\mathcal{V}^{-1}m)^{-1}m^T\mathcal{V}^{-1} \tag{3.8}$$

We estimate the abundances by computing

$$\alpha(x,y) = G(\mathbf{r}(x,y) - \mathbf{n}(x,y)). \tag{3.9}$$

Spectral unmixing separates various regions of an image into different bands. This process identifies targets which may have previously been hidden in various reflectance bands (Figure 3.1). By identifying these targets, however, edges are introduced. Edges, as always, are the source of difficulty in image resampling.

Note that equations 2.38 and 2.41 imply that the spatial distortions induced into the reflectance domain by the image formation process are similar to those induced into the abudnace domain; they differ only in noise. As a result, the techniques we develop for one will be applicable to the other. The interpolants and energy functionals associated with one will be suitable for the other. Therefore, we designate samples $\psi(x,y)$ and develop our techniques with this notation. Both $\mathbf{r}$ and α are vectors, and so, ψ is a vector.

In analogy with equations 2.38 and 2.41, the samples are derived from their corresponding signal, in the obvious way:

$$\psi_s(x,y) = h(x,y) * \psi(x,y) + \mathbf{n}(x,y) \tag{3.10}$$

with the auxiliary condition

$$|\psi| = 1 \quad \text{if} \quad \psi = \alpha. \tag{3.11}$$

The Fourier domain representation is

$$\Psi(u,v) = H(u,v)\Psi_s(u,v) + N(u,v). \tag{3.12}$$

3.2 Model Inversion and Underdetermination

When the imaging system is bandlimiting, object signal reconstruction is underdetermined; the image signal frequency content is restricted to $0 \leq u \leq u_s/2$, $0 \leq v \leq v_s/2$. These restrictions are imposed by the convolution of the object signal with the impulse response and by the sampling function. Restoration of the

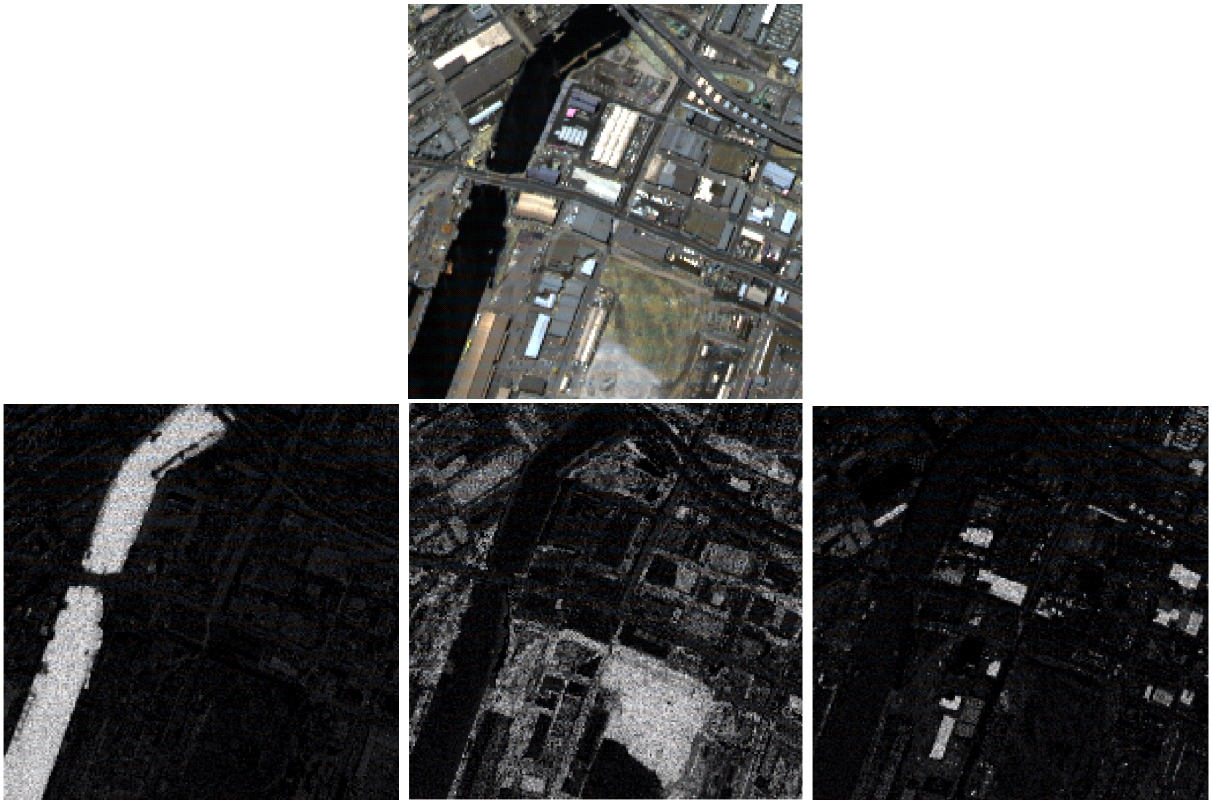


Figure 3.1: From a reflectance image (top) abundance images may be derived (bottom three panels). The distribution of water in the scene is depicted in panel 2; brightness of each pixel indicates the abundance of the water in that area

object signal may be attempted by inverting the image formation model, 3.12.

$$\Psi_s(u,v) = H^{-1}(u,v)[\Psi(u,v) - N(u,v)]. \tag{3.13}$$

If H is zero for any frequency we wish to construct, equation 3.13 implies a division by zero. As a result, some authors claim that signal reconstruction from bandlimited systems is impossible (11).

Alternatively, we can interpret the consequences of equation 3.13 differently: we recognize that *equation 3.13 supports an infinite family of solutions beyond the bandlimits* – we need only pick the right one! The non-invertability of the image formation function does not imply that ψ_s does not exist beyond the passband; it implies that we cannot invert H to find the solution. As an alternative, we seek a solution by using the information below the cutoff frequency and with a priori information.

Hunt (12) notes, however, that "the statement that division by zero prevents inferring information about $[\Psi_s]$ is equivalent to restricting the solution for $[\Psi_s]$ in terms of linear systems that can be implemented as Fourier filters. However, there are many rationales for the estimation of $[\Psi_s]$ that do not have a Fourier implementation... there is another facet to success in the recovery of information beyond the diffraction limit cut-off, namely the presence of information below the diffraction limit cut-off that is directly related to the information above the cutoff".

Before proceeding to Fourier mode restoration, we note that the impulse response may be inverted: we may attempt to de-blur the image before restoring its Fourier modes. The mathematical ability to apply deblurring effects is afforded by the existence of the inverse impulse response at all frequencies. We will see, however, that there are practical obstacles which interefere with deblurring methods.

3.3 Impulse Response Deconvolution

In response to requirement I-1, impulse response deconvolution may be attempted first by specifying a noise model $\hat{\mathbf{n}}$ and subtracting it from the digital image (equation 3.10),

$$\psi(x,y) - \hat{\mathbf{n}}(x,y) = h_3(x,y,\lambda)h_1(x,y) * \psi_s(x,y,\lambda) + \mathbf{n}(x,y) - \hat{\mathbf{n}}(x,y). \tag{3.14}$$

If $\hat{\mathbf{n}}$ is a very good noise model then

$$\mathbf{n}(x,y) - \hat{\mathbf{n}}(x,y) \approx 0 \quad \forall(x,y), \tag{3.15}$$

so we have

$$\psi(x,y) - \hat{\mathbf{n}}(x,y) = h_3(x,y,\lambda)h_1(x,y) * \psi_s(x,y,\lambda). \tag{3.16}$$

Now, operating on the left side of equation 3.16, by h_1^{-1}

$$h_1^{-1}(x,y) * [\psi(x,y) - \hat{\mathbf{n}}(x,y)] = h_1^{-1}(x,y) * h_3(x,y,\lambda)h_1(x,y) * \psi_s(x,y,\lambda). \tag{3.17}$$

In the absence of noise and aliasing artifacts $h_3(x,y,\lambda)$ and $h_1(x,y)*$ commute,

$$h_3(x,y,\lambda)h_1(x,y)* \approx h_1(x,y) * h_3(x,y,\lambda) \tag{3.18}$$

which implies

$$h_1^{-1}(x,y) * [\psi(x,y) - \hat{\mathbf{n}}(x,y)] = h_3(x,y,\lambda)\psi_s(x,y,\lambda). \tag{3.19}$$

Equation 3.19 is an estimate of the object signal satisfying requirement I-1, pending the legitimacy of approximations 3.15 and 3.18.

The approximations presented above, however, are often poor: the noise model may be globally good but locally poor; where noise or aliasing appears h_1* and h_3 do not commute. These errors are compounded by the sensitivity of the inverse impulse response: wherever h_1 is small, its inverse is large, as in the highest frequency bands. If the difference $\mathbf{r}(x,y) - \hat{\mathbf{n}}(x,y)$ is large and if the difference $\mathbf{r}(x,y) - \mathbf{n}(x,y)$ is small, then this error is amplified. Similarly, aliasing artifacts are amplified by the inverse impulse response.

Maximum entropy is a famous approach to image deconvolution. In this technique, the solution is sought by extracting much information from an image as is justified by its signal–to–noise ratio. In doing so, targets which have have been blurred by the impulse response are restored to their original compact versions. Errors in the noise model and aliasing artifacts receive the same treatment, however, and are restored as legitimate signal features.

3.4 Image Resampling

Image resampling is a means of rendering the model inversion problem well-defined by specifying additional prior information. The role of this information is synthesis of Fourier modes beyond the passband of the imaging system; additionally this information may encode means of reducing aliasing, blurring and noise.

The reciprocity between a signal's sample spacing, X and its sampling frequency, u_s,

$$Xu_s \propto 1, \tag{3.20}$$

implies that as we extend its frequency content, we confine its position estimate. Formally, *data resampling* is the process of transforming a collection of data ψ^0 with sample spacings X^0 and Y^0 to a collection of data ψ with sample spacings $X = X^0/z$ and $Y = Y^0/z$. If $0 < z < 1$, then the process is called *decimation*; if $z > 1$ and if the source data are fixed,

$$\psi(x_i, y_i) = \psi^0(x_i, y_i) \tag{3.21}$$

at all low-resolution grid points (x_i, y_i), then the process is called *interpolation*. We refer to equation 3.21 as *the interpolation constraint*. When we interpolate an image, then, we make no attempt to correct the source data ψ located at the grid points: we use them as fixed reference points to determine the coefficients in an interpolating function. We label those techniques in which $z > 1$ and in which the interpolation constraint is not enforced *super-resolution* techniques.

In what follows, we identify the source data – the digital image – with a superscript, ψ^0; resampled data are denoted by ψ.

3.4.1 Conservation

We postulate a conservation rule:

$$\sum_{\mathcal{R}} \psi(x, y) = \sum_{\mathcal{R}} \psi^0(x, y) \tag{3.22}$$

This condition stipulates that the radiance or abundance of a region must be conserved during the resampling process, where ψ is an estimate of ψ_s. For example, if a low-resolution pixel in band i, $\alpha_i^0(x, y) = 0.4$ is split into four, $\alpha_i(x + \frac{1}{2z}, y + \frac{1}{2z})$, $\alpha_i(x - \frac{1}{2z}, y + \frac{1}{2z})$, $\alpha_i(x + \frac{1}{2z}, y - \frac{1}{2z})$, and $\alpha_i(x - \frac{1}{2z}, y - \frac{1}{2z})$, then we require

$$\begin{aligned} &\alpha_i(x + \frac{1}{2z}, y + \frac{1}{2z}) + \alpha_i(x - \frac{1}{2z}, y + \frac{1}{2z}) \\ &\qquad + \alpha_i(x + \frac{1}{2z}, y - \frac{1}{2z}) + \alpha_i(x - \frac{1}{2z}, y - \frac{1}{2z}) = 0.4 \end{aligned} \tag{3.23}$$

The conservation of ψ may be required to ensure that the final image resembles the source data.

3.4.2 Interpolation

Conventional approaches to the image resampling problem are interpolation techniques. From the source data ψ^0, interpolation estimates ψ as approximations of ψ_s at non-grid points by assuming a functional relationship among the data. As an example, consider a polynomial approximation of one band, ψ_j^0, of ψ^0 at (x, y):

$$\psi_j(x, y) \approx a_{00} + a_{01}x + a_{10}y + a_{11}xy + \ldots a_{mn}x^m y^n$$

The coefficients a_{ij} are computed with ψ_j^0. After establishing values for the coefficients, $\psi_j(x + \Delta x, y + \Delta y)$ can be evaluated.

By imposing a functional relationship among the data, sufficient prior information has been supplied to render the image restoration problem well-defined. This does not grantee that this formulation is an adequate representation of the inverse sampling function. Moreover, interpolation makes no attempt to address noise, aliasing artifacts, and blurring: interpolation treats these as negligible effects.

Optimal Linear Interpolation

When the data sampling rate meets or exceeds the Nyquist criterion, the image signal is determined by its samples. According to E. T. Whittaker (13), (7) in one dimension the image signal is

$$\psi(x) = \sum_{n=-\infty}^{\infty} \psi^0(x) \frac{\sin \pi/X(x-nX)}{\pi/X(x-nX)} \tag{3.24}$$

Optimal linear interpolation gives us a means of recovering the image signal $h_1 * \psi_s = \psi_I$ because it simulates h_3^{-1} for critically sampled or oversampled signals. The signal ψ_s is yet undetermined. For undersamped signals, optimal linear interpolation is of little use. As we have seen, the impulse response is not bandlimited and so the image signal is undersampled. We consider other interpolants in the following.

Bilinear Interpolation

We might assume that the spatial variation of a scene's reflectivity is sufficiently small that the apparent reflectivity at one point is linearly related to the measured data. In this case, we approximate $\psi_s(x+\Delta x, y+\Delta y)$ by a linear function of the samples $\psi^0(x,y)$.

Given gridded data $\psi^0(x_i, y_j)$, $\psi^0(x_{i+1}, y_j)$, $\psi^0(x_{i+1}, y_{j+1})$, and $\psi^0(x_i, y_{j+1})$, we may estimate data at nongrid points: we make the following assignments:

$$\begin{aligned} x &= x_i + \Delta x & \psi_1 &= \psi^0(x_i, y_j) \\ y &= y_j + \Delta y & \psi_2 &= \psi^0(x_{i+1}, y_j) \\ t &= \tfrac{x-x_i}{x_{i+1}-x_i} & \psi_3 &= \psi^0(x_{i+1}, y_{j+1}) \\ u &= \tfrac{y-y_j}{y_{j+1}-y_j} & \psi_4 &= \psi^0(x_i, y_{j+1}) \end{aligned}$$

The linearly interpolated value is

$$\psi(x+\Delta x, y+\Delta y) = (1-t)(1-u)\psi_1 + t(1-u)\psi_2 + tu\psi_3 + (1-t)u\psi_4.$$

Because the high frequency power of the bilinear interpolation kernel is meager, it has limited success in restoring high frequency image content. To improve this approach we can augment the interpolatory function from linear to polynomial, whose high frequency kernel power depends on the polynomial order.

Polynomial Interpolation

A polynomial interpolatory kernel offers an arbitrarily high power spectrum. Noting that a polynomial of degree N requires $N+1$ points in the neighborhood of $\psi^0(x,y)$ to set the coefficients Press et al (14) author a polynomial of the form

$$\begin{aligned} P(\psi)(x) = {} & \frac{(x-x_2)(x-x_3)\dots(x-x_{N+1})}{(x_1-x_2)(x_1-x_3)\dots(x_1-x_{N+1})}\psi_1 + \frac{(x-x_1)(x-x_2)\dots(x-x_{N+1})}{(x_2-x1)(x_2-x_3)\dots(x_2-x_{N+1})}\psi_2 \\ & + \dots \frac{(x-x_1)(x-x_2)\dots(x-x_N)}{(x_{N+1})(x_{N+1}-x_2)\dots(x_{N+1}-x_N)}\psi_N \end{aligned} \tag{3.25}$$

in one dimension. To interpolate with this polynomial in two dimensions one could interpolate along rows, then along columns.

While polynomial interpolation may supply arbitrarily high power over any frequency band (provided we're willing to implement a polynomial interpolant of arbitrarily high order), it does not adequately treat image degradation by the imaging system. Moreover, interpolation does not address the constraints, equations 2.12 and 3.22.

To properly treat image degradation by the imaging system, we need to relieve the interpolation constraint, equation 3.21, and re-estimate the source data. We will, however, require an alternate constraint.

3.4.3 Super-Resolution

Among the alternate approaches to resampling is Bayesian inference. In this regime, the conditional probabilities of two states, ψ_s and ψ are related by Bayes' Theorem,

$$P(\psi^0|\psi_s)P(\psi_s) = P(\psi_s|\psi^0)P(\psi^0) \tag{3.26}$$

Thus, we can estimate the probability of an image signal ψ_s, given sample ψ^0:

$$P(\psi_s|\psi^0) = \frac{P(\psi^0|\psi_s)P(\psi^0)}{P(\psi_s)} \tag{3.27}$$

To evaluate this probability, we require a model of the probabilities $P(\psi_s)$ and $P(\psi^0)$ and an assessment of $P(\psi^0|\psi_s)$. With these models we can calculate the most likely state ψ_s giving rise to ψ^0 by maximizing 3.27.

Yet another approach to the resampling problem is energy minimization. We suppose that ψ^0 and ψ_s are the initial and final states, respectively, of a system, ψ. Associated with the system is an energy E which is minimized in seeking the final state given the initial conditions. We have the differential equation

$$\frac{\partial E}{\partial \psi} = -\frac{d\psi}{dt} \quad \text{with} \quad \psi(t=0) = \psi^0 \quad \text{and} \quad \psi(t \to \infty) = \psi_s \tag{3.28}$$

The restoration problem is reduced to formulating an appropriate energy functional of the system. Data evolution according to equation 3.28 is called *gradient descent*.

We deliberately use the ambiguous term 'system' because we treat energy minimization as dynamics which can be assumed by a variety of physical systems. We will explore several physical systems that provide insight into the dynamics we employ. We will find that at various stages of development our dynamics resemble diffusion of gas in a room, data propagation in a neural network, or the evolution of soap bubble.

Shape-based Methods

Interpolation techniques such as polynomial interpolation, outlined above, are described by Morse and Schwartzwald (15) as "fitting the function" methods. As an alternative, "fitting the geometry" methods have been proposed. In these techniques, low resolution image geometry is measured and used as a guide for the resampling. The essential feature of these techniques, claim Morse and Schwartzwald, is the use of "visually-significant geometric properties, not function-analytic properties of the image".

The approach of Alleback and Wong (16) is to estimate sub-pixel edge locations, then to use these estimates to guide a modified bilinear interpolation scheme. They estimate sub-pixel edge locations by first calculating an edge map with a centre-on surround-off filter. To determine the edge map on the fine mesh they linearly interpolate the filter output between points on the low resolution mesh. Using this map, they determine which pixel centers are separated by an edge. If two pixels are separated by an edge then they resample the image with one of these pixels values adjusted, rather than with the recorded pixel value. They have found that artifacts appear during their process which they mitigate by imposing auxiliary constraints on the resampling process.

To employ image geometry for resampling Carrato et al (17) introduce an interpolant based on a non-linear mean of a pixel's neighbourhood. As opposed to Allebach and Wong's approach, Carrato's technique incorporates edge estimates directly into the interpolant. Both techniques, however, introduce a nonlinearity into a conventional linear interpolant.

We note that the methods of Allebach and Wong and of Carrato et al treat the restoration problem as an interpolation problem, whereby the source data are not corrected.

Perhaps a natural means of exploiting image geometry is to evolve all data by a differential equation. Doing so, we can subject all data to evolution, guided by intensity gradient measurements. In particular we may evolve an image $\psi(x, y)$ according to

$$\frac{\partial \psi}{\partial t} = \nabla^2 \psi. \tag{3.29}$$

This is the diffusion equation. It minimizes the difference between pixel values, and thus smoothes an image (Figures 3.2 and 3.3).

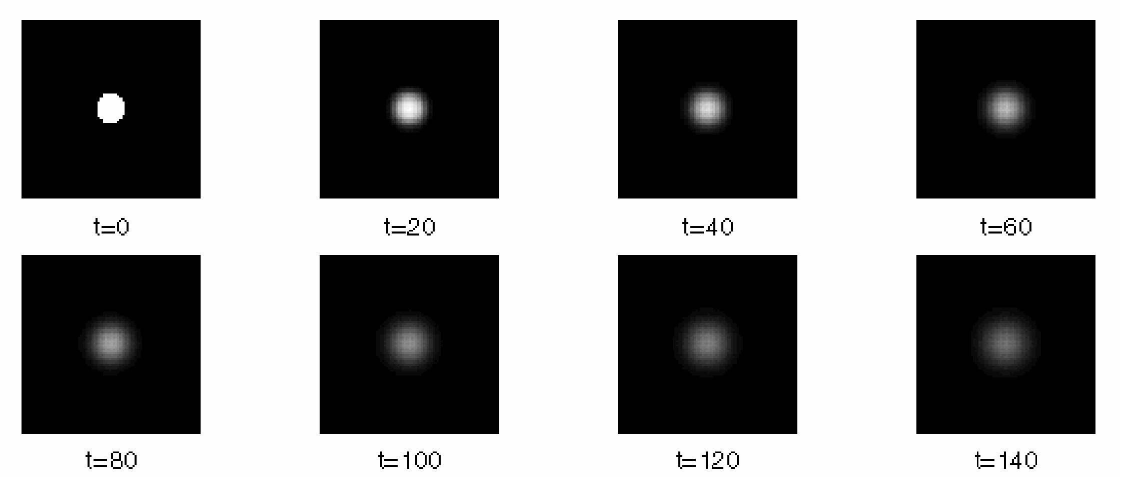


Figure 3.2: Time sequence of data diffusion: image defects are diffused away. If diffusion continues, then all image features are lost

If, for example, $\psi(-2X) = 0.75$, $\psi(-X) = 0.75$, $\psi(x) = 0$, $\psi(X) = 0.25$, and $\psi(2X) = 0.25$, are values of pixels at time t, we compute $\nabla^2 \psi(x = 0)$ by discretizing the derivative,

$$\nabla \psi(x) = \frac{\psi(x + X) - \psi(x - X)}{2X} \hat{x} \tag{3.30}$$

where $\hat{x}$ is the normal in the x-direction. Then

$$\nabla^2 \psi(x) = \frac{\nabla \psi(x + X) - \nabla \psi(x - \Delta X)}{\Delta X} \hat{x} \cdot \hat{x} \tag{3.31}$$

and we have

$$\nabla^2 \psi(x) = \frac{\frac{\psi(x+2X) - \psi(x)}{2X} - \frac{\psi(x) - \psi(x-2X)}{2X}}{2X} \tag{3.32}$$

with $X = 1$ this implies

$$\nabla^2 \psi(x) = \frac{1}{4} [\psi(x + 2X) - 2\psi(x) + \psi(x - 2X)] \tag{3.33}$$

and in this case,

$$\nabla^2\psi(0) = \frac{1}{4}[0.25 - 2 \times 0 + 0.75] = 1.0. \tag{3.34}$$

Now, discretizing the time derivative in equation 3.29,

$$\frac{d\psi}{dt} \approx \frac{\Delta\psi}{\Delta t} \tag{3.35}$$

We have

$$\Delta\psi(x) = \nabla^2\psi(x)\Delta t \tag{3.36}$$

If $\Delta t = 0.1$, then with result 3.34 we have

$$\Delta\psi(0) = 0.1 \tag{3.37}$$

and therefore $\psi(x)$ is incremented from 0 to 0.1. It will continue to be incremented until it reaches 0.5; if it exceeds 0.5 then $\Delta\psi$ will be negative and it will return toward 0.5. The outcome of this resampling resembles bilinear interpolation, except that noise is suppressed.

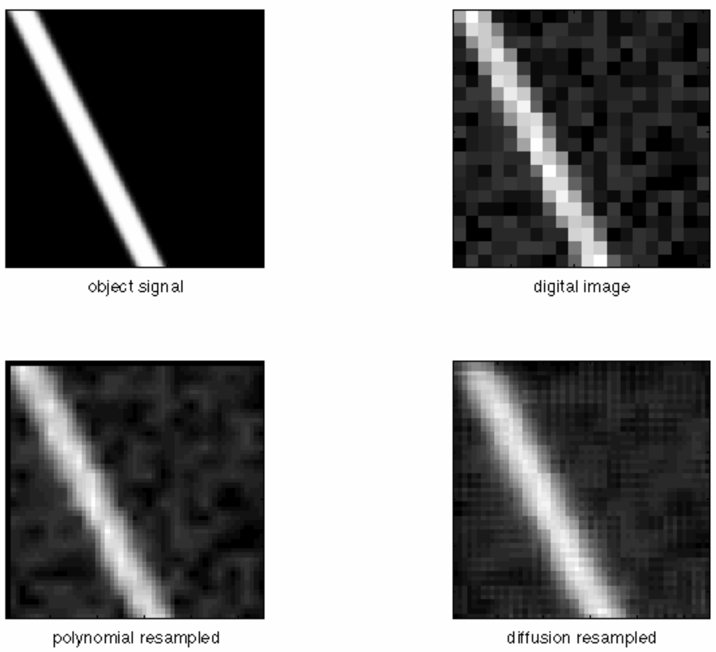


Figure 3.3: Object signal (panel 1), digital image (panel 2), polynomial resampled digital image (panel 3), diffusion resampled image (panel 4). Note that noise and aliasing are mitigated by diffusion, in contrast with interpolation

We find benefit in data diffusion when we subject all data, including the source data, to equation 3.29. Doing this, we have the opportunity to correct noise and other image formation artifacts. Noise, for example, is smoothed and deleted.

To see why we've selected the Laplacian rather than the gradient, consider a noisy pixel at location (x, y) on an otherwise flat background. Presumably, the pixel should be set to the background value by the evolution scheme we pick. However, the intensity gradient at (x, y) is zero, and in fact, would remain zero if the right hand side of equation 3.29 read $|\nabla\psi|$. To see this, note that the four adjacent pixels have nonzero, equal gradients. Each of these pixel values would be equally increased toward the value of the noisy pixel; with this symmetric increase, the gradient at (x,y) remains zero. In contrast, the Laplacian at (x, y) is nonzero and the noisy pixel is adjusted toward the background value under diffusion.

To fully justify the diffusion equation we begin by relating the image energy to its smoothness

$$E_1 = \frac{1}{2} \sum_{(x,y)\in\mathcal{I}} \left[(\psi_i(x+X,y) - \psi_i(x,y))^2 + (\psi_i(x,y+Y) - \psi_i(x,y))^2 \right], \tag{3.38}$$

In addition, we will need energy functional terms to halt the evolution

$$E_2 = \frac{1}{2} \sum_{(x,y)\in\mathcal{I}} \left(1 - \sum_j \psi_j(x,y) \right)^2, \tag{3.39}$$

and conservation,

$$E_3 = \frac{1}{2} \sum_{\mathcal{R}\in\mathcal{I}} \left(\sum_{(x,y)\in\mathcal{R}} \psi_i^0(x,y) - \sum_{(x,y)\in\mathcal{R}} \psi_i(x,y) \right)^2 \tag{3.40}$$

where $\mathcal{I}$ refers to the image spatial domain, and where $\mathcal{R}$ is a subset of the image domain. We have denoted the source data by ψ^0.

Each energy term is a measure of the state of the image: the first term, equation 3.38, is a measure of image smoothness. The second term, equation 3.39, measures the departure of a pixel from unit norm. The third term, equation 3.40, measures the degree to which reflectance or abundance is locally conserved. The region $\mathcal{R}$ is the area assigned to a low-resolution pixel. Because unit normalization is a requirement only of abundance image pixels, the coefficient a_2 should be nonzero only when the image is an abundance image.

As an example of a conservation energy measurement, consider the abundance band, $\alpha_i^0(x,y)$, of the i^{th} material. Suppose the abundance in this band obtains the value 0.4 on the course grid. Resampling to a finer grid is the process of splitting this pixel into several pixels, say four, $\alpha_i(x-X,y-Y)$, $\alpha_i(x+X,y-Y)$, $\alpha_i(x-X,y+Y)$, $\alpha_i(x+X,y+Y)$, We'll call the original pixel the parent pixel and the new pixels the children. If at some point during the resampling process the children have the values $\alpha_i(x-X,y-Y) = 0.2$, $\alpha_i(x+X,y-Y) = 0.1$, $\alpha_i(x-X,y+Y) = 0.1$, and $\alpha_i(x+X,y+Y) = 0.2$, then conservation energy in this band is

$$\begin{aligned} E_3 = {} & [\alpha_i^0(x,y) - \alpha_i(x-X,y-Y) - \alpha_i(x+X,y-Y) \\ & -\alpha_i(x-X,y+Y) - \alpha_i(x+X,y+Y)]^2 \\ = {} & [0.4 - 0.2 - 0.1 - 0.1 - 0.2]^2 \\ = {} & 0.04 \end{aligned} \tag{3.41}$$

which is not a minimum. The minimum of this energy term is 0, which is achieved when the sum of the children equals the parent pixel.

To proceed, we move to the continuous model, noting that the difference $\psi(x_i,y) - \psi(x_{i+1},y)$ in equation 3.38 is the discrete form of the derivative of ψ in the $x-$direction, with $X = 1$:

$$\psi_i(x+X,y) - \psi_i(x,y) = \frac{\partial\psi}{\partial x} \tag{3.42}$$

$$\psi_i(x,y+Y) - \psi_i(x,y) = \frac{\partial\psi}{\partial y}. \tag{3.43}$$

Then,

$$E_1 = \frac{1}{2} \sum_{(x,y)\in\mathcal{I}} \left[\left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial y}\right)^2 \right] \tag{3.44}$$

We replace the sums by integrals,

$$\sum_{\mathcal{I}} \to \int\int_{\mathcal{I}} \tag{3.45}$$

Thus, we direct the flow by evaluating the Energy gradient, equation 3.28.

3.4.4 Diffusion

Our energy terms are

$$E_1 = \frac{1}{2} \int\int_{\mathcal{I}} dxdy \left[\left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial y}\right)^2 \right] \tag{3.46}$$

$$E_2 = \frac{1}{2} \int\int_{\mathcal{I}} dxdy \left(1 - \sum_j \psi_j\right)^2 \tag{3.47}$$

$$E_3 = \frac{1}{2} \sum_{\mathcal{R}\in\mathcal{I}} \left(\int\int_{\mathcal{R}} dxdy\psi^0 - \int\int_{\mathcal{R}} dxdy\psi \right)^2 \tag{3.48}$$

With this model, the energy of the image at any time during the reconstruction is

$$E[\psi] = a_1 E_1[\psi] + a_2 E_2[\psi] + a_3 E_3[\psi]. \tag{3.49}$$

The coefficients a_1, a_2, and a_3 determine the relative priority of the smoothness, normalization, and conservation terms. The compromise among smoothness, normalization, and conservation is directed by these coefficients. To super-resolve reflectivity images, we set $a_2 = 0$; to super-resolve abundance images we ensure $0 < a_2 < 1$.

Differentiating our smoothness term with respect to ψ according to 3.28 we have

$$\frac{\partial E_s[\psi]}{\partial\alpha} = \frac{1}{2} \int\int_{\mathcal{D}} dxdy \left[\frac{\partial}{\partial\tilde{\psi}} \left(\frac{\partial\psi}{\partial x}\right)^2 + \frac{\partial}{\partial\psi} \left(\frac{\partial\psi}{\partial y}\right)^2 \right] \tag{3.50}$$

$$= \int\int_{\mathcal{D}} dxdy \left[\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} \right] \tag{3.51}$$

where we've treated the derivatives as follows:

$$\frac{\partial}{\partial\psi} \left(\frac{\partial\psi}{\partial x} \frac{\partial\psi}{\partial x} \right) = 2\frac{\partial^2\psi}{\partial x^2} \tag{3.52}$$

The derivative operator,

$$\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial y^2}$$

is the Laplacian operator, ∇^2.

Differentiating the normalization term gives

$$\frac{\partial E_2}{\partial \psi} = \int\!\!\int_{\mathcal{I}} \left(1 - \sum_j \psi \right) \equiv -\Delta_n \psi \tag{3.53}$$

where $\sum_j = 1$. Finally, differentiating the conservation term gives

$$\frac{\partial E_3}{\partial \psi} = -\sum_{\mathcal{R} \in \mathcal{I}} \left(\int\!\!\int_{\mathcal{R}} dxdy\psi - \int\!\!\int_{\mathcal{R}} dxdy\psi \right) \int\!\!\int_{\mathcal{R}} dxdy \equiv -\Delta_3 \psi \tag{3.54}$$

We have set the high resolution pixel size to 1: the region $\mathcal{R}$ is the area occupied by a single low-resolution pixel, whose area is $\int\int_{\mathcal{R}} dxdy = z^2$.

Thus, we have the differential equation for each pixel

$$\frac{\partial \psi}{\partial t} = a_1 \nabla^2 \psi + a_2 \Delta_2 \psi + a_3 \Delta_3 \psi \tag{3.55}$$

Equation 3.55 is an the diffusion equation where the abundance constraints operate as a source. If the coefficients a_2 and a_3 are zero then the data evolve analogously to the gas in a room or the temperature on a conducting plate. If a_2 and a_3 are nonzero then the dynamics resemble a plate with points maintained at constant temperature.

3.4.5 Diffusion in Nature

Consider a collection of gas molecules in a room (18). Let us suppose that the gas is not in equilibrium so that the number of molecules per unit volume, ψ, is not constant, $\psi = \psi(\mathbf{x})$. The molecules will diffuse until they reach an equilibrium state where ψ is constant. The flux $\mathbf{J}$ of molecules through a surface is defined as the number of molecules passing through that surface per unit area per unit time (Figure 3.4). We expect that the flux should be proportional to the concentration gradient:

$$\mathbf{J} = -a_1 \nabla \psi(\mathbf{x}) \tag{3.56}$$

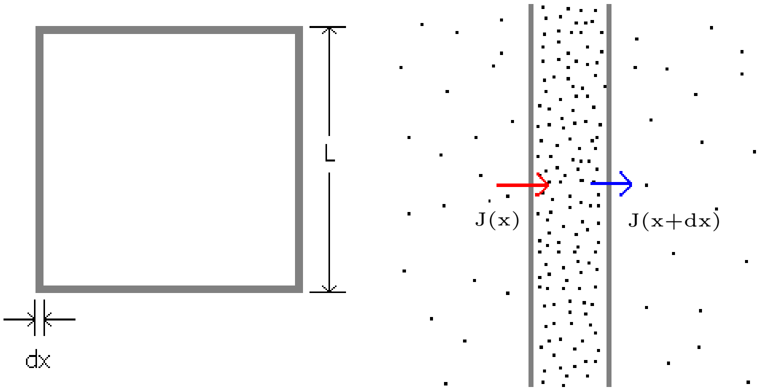


Figure 3.4: Diffusion of a gas. The rate of diffusion of molecules through a side of the enclosure depends on the concentration gradient.

In particular, consider a square in two dimensions aligned with coordinates x and y The face aligned perpendicular to the x-axis has thickness dx and length L. The number of molecules contained in that face is $nLdx$ where n is the number of molecules per unit area; the rate of change of this number is the total flux through the surface at x minus the flux through the surface at $x + dx$:

$$\frac{\partial}{\partial t}(\psi nLdx) = nLJ_x(x) - nLJ_x(x + dx) \tag{3.57}$$

We Taylor expand the second term on the right side to first order:

$$J_x(x + dx) \approx J_x(x) + \frac{\partial J_x}{\partial x}dx \tag{3.58}$$

giving

$$\frac{\partial}{\partial t}(nLdx) = nLJ_x(x) - nLJ_x + nL\frac{\partial J_x}{\partial x}dx = nL\frac{\partial J_x}{\partial x}dx \tag{3.59}$$

with the x-component of equation 3.56,

$$J_x = -a_1\frac{\partial \psi}{\partial x}, \tag{3.60}$$

equation 3.59 implies

$$\frac{\partial}{\partial t}(\psi dx) = -a_1\frac{\partial^2 \psi}{\partial x^2}. \tag{3.61}$$

In general,

$$\frac{\partial \psi}{\partial t} = -a_1\nabla^2\psi. \tag{3.62}$$

Equation 3.62 is called the *diffusion equation*; the coefficient a_1 is called the *coefficient of self-diffusion.*

3.4.6 Curvature Flow

Edges are not well preserved in the diffusion of data by application of equation 3.55. As a solution we include a coefficient to inhibit diffusion across edges. The equation of motion is

$$\frac{\partial \psi}{\partial t} = a_1\nabla\cdot(s\nabla\psi) + a_2\Delta_n\psi + a_3\Delta_c\psi \tag{3.63}$$

where

$$s(x,y) = \begin{cases} 1 & \text{if} \quad |\nabla\psi| < \phi \\ 0 & \text{otherwise} \end{cases} \tag{3.64}$$

for some appropriate choice of ϕ. We note that if the gradient is sufficiently high, then s is 0; when the gradient is low the coefficient is 1. That is, only features of magnitude ϕ guide the diffusion (19).

To guide the diffusion by all features in the image we replace s with the inverted gradient magnitude. The super-resolution will be directed by the gradient at each point in the image. Equation 3.63 becomes

$$\frac{\partial \psi}{\partial t} = a_1\nabla\cdot\left(\frac{\nabla\psi}{|\nabla\psi|}\right) + a_2\Delta_n\psi + a_3\Delta_c\psi \tag{3.65}$$

We explicitly compute the derivatives below:

$$\nabla\cdot\left(\frac{\nabla\psi}{|\nabla\psi|}\right) = \frac{\nabla\cdot\nabla\psi|\nabla\psi| - \nabla\psi\cdot\nabla|\nabla\psi|}{|\nabla\psi|^2} \tag{3.66}$$

In 2D Cartesian coordinates with the abbreviations

$$\frac{\partial\psi}{\partial x} = \psi_x \quad \frac{\partial\psi}{\partial y} = \psi_y \tag{3.67}$$

$$\frac{\partial^2\psi}{\partial x^2} = \psi_{xx} \quad \frac{\partial^2\psi}{\partial y^2} = \psi_{yy} \tag{3.68}$$

we have

$$\nabla\psi = \psi_x\hat{x} + \psi_y\hat{y}, \tag{3.69}$$

$$\nabla\cdot\nabla\psi = \psi_{xx} + \psi_{yy}, \tag{3.70}$$

$$|\nabla\psi| = \left[\psi_x^2 + \psi_y^2\right]^{\frac{1}{2}}, \tag{3.71}$$

and

$$\nabla|\nabla\psi| = \left[\psi_x^2 + \psi_y^2\right]^{-\frac{1}{2}}\left[(\psi_x\psi_{xx} + \psi_y\psi_{xy})\hat{x} + (\psi_x\psi_{xy} + \psi_y\psi_{yy})\hat{y}\right], \tag{3.72}$$

with

$$\psi_{xy} = \psi_{yx}.$$

Equations 3.69 - 3.72 imply

$$\nabla\psi\cdot\nabla|\nabla\psi| = \left[\psi_x^2 + \psi_y^2\right]^{-\frac{1}{2}}\left[\psi_x^2\psi_{xx} + 2\psi_x\psi_y\psi_{xy} + \psi_y^2\psi_{yy}\right] \tag{3.73}$$

and

$$\nabla^2\psi|\nabla\psi| = \left[\psi_{xx} + \psi_{yy}\right]\left[\psi_x^2 + \psi_y^2\right]^{\frac{1}{2}}.. \tag{3.74}$$

We rewrite equation 3.74 as

$$\nabla^2\psi|\nabla\psi| = \left[\psi_{xx} + \psi_{yy}\right]\left[\psi_x^2 + \psi_y^2\right]\cdot\left[\psi_x^2 + \psi_y^2\right]^{-\frac{1}{2}}.. \tag{3.75}$$

Substituting equations 3.73 and 3.75 into equation 3.66 we have

$$\nabla\cdot\left(\frac{\nabla\psi}{|\nabla\psi|}\right) = \frac{\psi_{xx}\psi_y^2 - 2\psi_x\psi_y\psi_{xy} + \psi_{yy}\psi_x^2}{\left[\psi_x^2 + \psi_y^2\right]^{\frac{3}{2}}} \equiv \kappa. \tag{3.76}$$

κ is called the *curvature* of ψ (Figure 3.5). We have, then,

$$\frac{d\tilde{\alpha}}{dt} = a_1\kappa + a_2\Delta_2\psi + a_3\Delta_3\psi \tag{3.77}$$

A surface evolving according to equation 3.77 is said to be undergoing (constrained) curvature flow[1]. A white square subject to curvature flow is depicted in Figure 3.6. If we set $\Delta_2\psi(x,y) = \Delta_3\psi(x,y) = 0$, the evolution is unconstrained curvature flow

3.4.7 Curvature Flow in Nature

If we model the interaction potential between molecules in an ensemble by the Morse potential,

$$U(x) = a[1 - e^{-bx}]^2 \tag{3.78}$$

---

[1] The expression for curvature in (15), appearing there as equation 2, is incorrect. It should read as equation 3.76

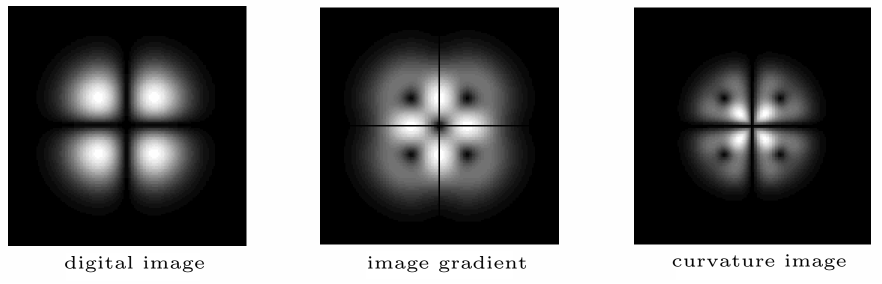


Figure 3.5: Geometric quantities – gradient and curvature

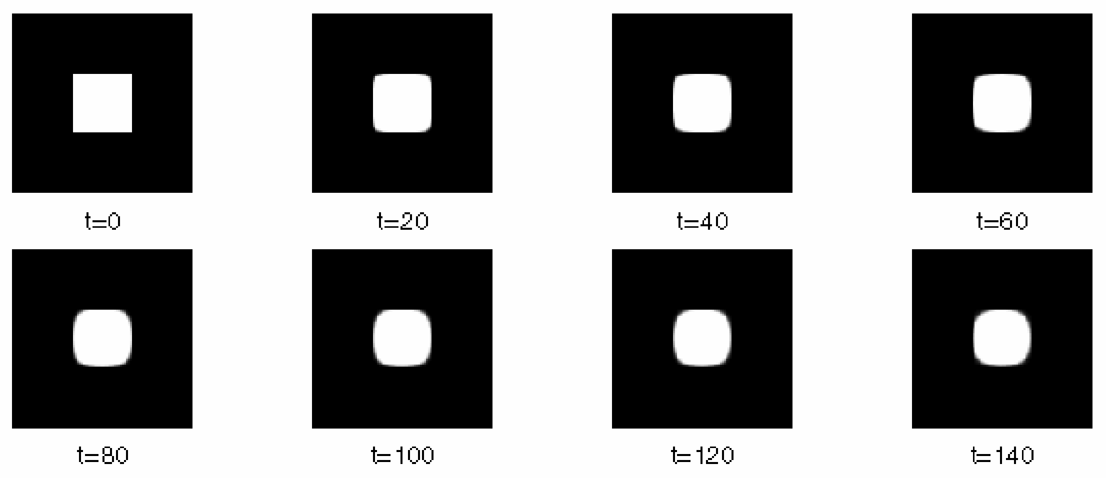


Figure 3.6: Time series of curvature flow: the shape collapses as time proceeds

then the force between molecules is attractive at large distances and repulsive at short distances. As a result, molecules on the boundary of the ensemble experience a net force directed to the interior. Until the net force acting on a molecule is zero, it will continue to move, resulting in a depopulation of molecules on the boundary; the molecules on the surface experience surface tension because their average separation is increased.

Parameterize the surface of the ensemble by ψ; the pressure on the surface is proportional to the surface tension and the curvature of the surface κ,

$$P(\psi(x)) = \gamma\kappa(\psi(x)). \tag{3.79}$$

The pressure is also the force per unit area:

$$P(\psi(x)) = f(\psi(x))/A \tag{3.80}$$

Equations 3.79 and 3.80 imply

$$f(\psi(x))/A = \gamma\kappa(\psi(x)). \tag{3.81}$$

According to Newton, the total force is the mass m times the acceleration,

$$f(\psi(x)) = ma\frac{\partial^2\psi(x)}{\partial t^2} \tag{3.82}$$

The mass of an element of the surface is the area mass density ρ times the area

$$m = \rho A \tag{3.83}$$

so we have

$$f(\psi(x)) = \rho A\frac{\partial^2\psi(x)}{\partial t^2}. \tag{3.84}$$

Equations 3.81 and 3.84 imply

$$\rho a(\psi(x)) = \gamma\kappa(\psi(x)) \tag{3.85}$$

which implies

$$\frac{\partial^2\psi(x)}{\partial t^2} = \frac{\gamma\kappa(\psi(x))}{\rho}. \tag{3.86}$$

The velocity is the anti-derivative,

$$\frac{\partial\psi(x)}{\partial t} = \frac{da(\psi(x))}{dt} = \int \frac{\gamma\kappa(\psi(x))}{\rho} + C. \tag{3.87}$$

For short times, the surface tension, curvature, and mass are approximately constant giving

$$\frac{d\psi(x)}{dt} = \frac{\gamma\kappa(\psi(x))}{\rho}\Delta t \tag{3.88}$$

where we've set the constant C to zero because we start with a system at rest. Equation 3.88 is curvature flow with $a_1 = \gamma\Delta t/\rho$. Curvature flow is the dynamics of a volume of fluid with surface tension, such as a raindrop or a soap bubble.

Curvature flow is exhibited by a variety of physical systems: for example, grain growth in metal is directed by curvature. There, molecules on the boundary of a grain have higher potential energy than those on the interior (20) (21). With higher potential energy on the surface, the grain has a surface tension and is subject to curvature flow as it grows.

3.4.8 Curvature Flow Phenomenology

The curvature flow evolution described by equation 3.77 is designed to inhibit smoothing where edges may appear. That is, diffusion is low where the gradient is high. With these dynamics, diffusion on either side of an edge may take place, while the edge, itself, is sharpened. To see how an edge may be enhanced, consider a straight edge cutting across an image (Figure 3.7). On one side of the gradient, call it side 0, where the value is low the diffusion is influenced, for the most part, by the data on side 0. Those pixels at the low end of the gradient are reduced in value to the mean of side 0. The pixels on the other side of the gradient, call it side 1, where the values are, high have little influence on the diffusion taking place in side 0 on account of the factor of $1/|\nabla\psi|$ in the definition of curvature. This edge enhancement, then, resembles deblurring.

To see this in another way, note that the curvature κ is a measure of the flux density of a vector field $B = \nabla\psi$. Now, reconsider the edge discussed above. In particular, consider a point x_0 at the midpoint of the gradient: to calculate the flux of B, i.e. the curvature at x_0, center an infinitesimal Gaussian box on x_0. The flux through the faces perpendicular to B are equal and opposite; the flux through each remaining side is zero. The mean flux through the surface of the box – zero – is the curvature at x_0. The flux density of B at points not on the midpoint of the gradient is nonzero. Consider a second point located not at the midpoint of the gradient: the flux through the face of the Gaussian box nearest the midpoint of the gradient is greater than the flux through the farthest face. Only when the mean value on the two faces is equal and opposite does the diffusion halt.

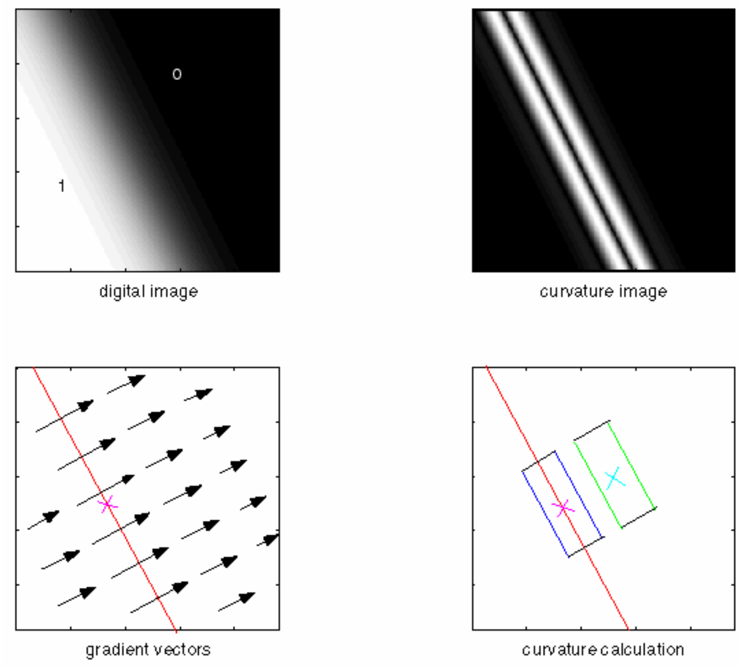
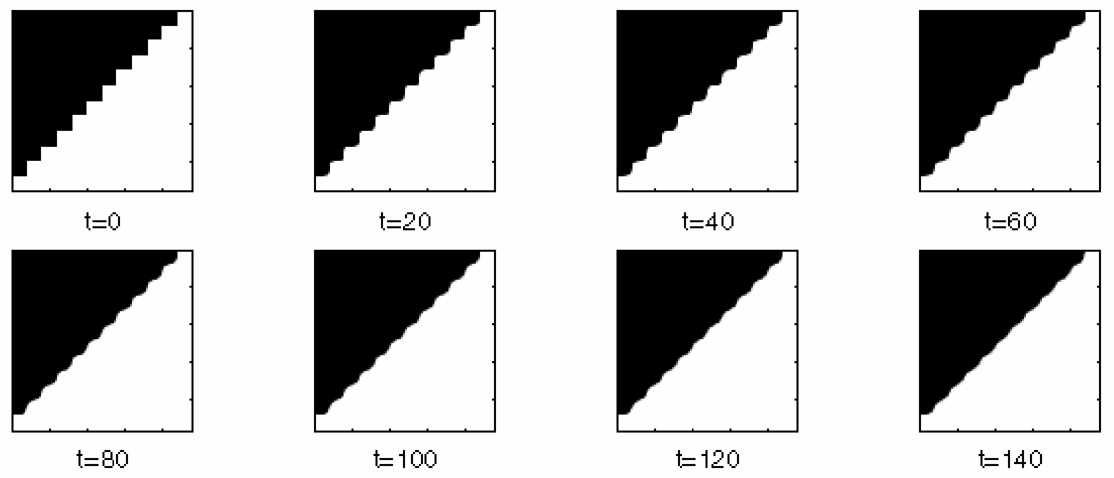


Figure 3.7: Curvature calculation; see section 3.4.8 for description

We find additional benefit in the curvature flow dynamics. Consider a point a the center of curved edge, that is, a point in a non-uniform vector field B. The flux density through sides of the Gaussian box oriented perpendicular to B are unequal; the flux through each remaining side is zero. Therefore, the curvature at that point is nonzero; the result is that the gradient sharpens and migrates. Until the edge is locally straight, i.e. until B is locally uniform, diffusion takes place. Therefore, aliasing artifacts on a straight edge are minimized, as the vector field nears uniformity. If the curve is closed, B is never uniform so the curve collapses under its own curvature. In this way, noise is removed. However, all image features are deleted unless auxiliary constraints are imposed. The auxiliary constraints balance the curvature flow force to make shape representation as geometrically economical as possible. This phenomenology is summarized by Grayson's Theorem(22): any non-convex curve will collapse to a circular point under unconstrained curvature flow. The equilibrium surface resulting from constrained curvature flow is one of minimum curvature. Feature representation is, then, as achieved with the least possible perimeter; for example a curve would be as non-jagged as possible.

By inhibiting diffusion across high gradients, the resampling process is guided by the geometry supplied by the image signal samples. With these cues the image is pulled up by its own bootstraps: the prior information is, in part, extracted from the data, so the resampling is adaptive.

We find one difficulty with curvature flow. Suppose the image is a disk with the value 0 on the disk and 1 elsewhere. Then, the normal on the boundary of the disk points outward and the curvature on the boundary is positive. Therefore, the boundary will move inward. On the other hand, if the interior of the

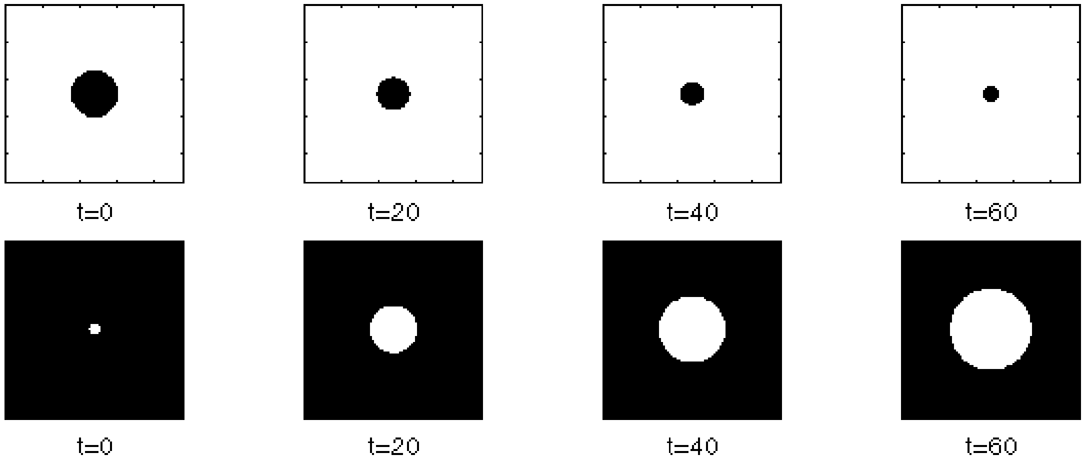


t=0 t=20 t=40 t=60

t=80 t=100 t=120 t=140

Figure 3.8: Correction of aliasing by curvature flow

disk is 1 and the exterior is zero, then the disk will grow (Figure 3.9).



t=0 t=20 t=40 t=60

t=0 t=20 t=40 t=60

Figure 3.9: Geometrically equivalent features may undergo opposite dynamics

This has been treated by Malladi and Sethian by the introduction of a switch. In contrast with the auxiliary constraints employed by Allebach and Wong, for example, we will see that the introduction of the switch is natural: it leads us to a neural network interpretation of the dynamics.

3.4.9 Curvature Flow Switching

Curvature flow data evolution suffers a deficiency: the direction of curvature flow depends on intensity. For instance, suppose we have only two materials in the scene. A region of abundance 1 in one of the abundance maps corresponds to a region of abundance 0 in the other map. This region in each map should collapse to the *same* shape. Under curvature flow, however, the shape will evolve in one direction in one map, and in another direction in the other map. The opposite dynamics are a consequence of the sign (positive or negative) of the curvature measurement. Where it is positive in one map, it may be negative in the other.

To encourage geometrically equivalent shapes to evolve in the same direction, regardless of intensity, Malladi and Sethian introduce a curvature flow switch as follows: if $\mu_r(x, y)$ denotes the average of pixel values in a neighborhood or radius r centered on pixel (x, y) then we replace the first term on the right side of equation 3.77 by the following:

$$F_\kappa(\kappa, \mu_r, \phi) = \begin{cases} \min(\kappa, 0)\nabla\psi & \text{if} \mu_r(x, y) < \phi \\ \max(\kappa, 0)\nabla\psi & \text{otherwise} \end{cases} \tag{3.89}$$

for some choice of ϕ. We note the inclusion of the gradient-dependent velocity factor $\nabla\psi$. With this inclusion, pixel adjustments are largest when the gradient is high, and directed according to the sign of the gradient. However, adjustments are made independently of the difference $\mu_r(x, y) - \phi$.

By incorporating a dependence on the difference $\mu_r(x, y) - \phi$ we might encourage the system to focus on a solution when the difference is minimized. That is, when the difference is small, adjustments should be small, when the difference is large, adjustments may be large.

If we generalize Malladi and Sethian's curvature flow switch we will recover neural network-esque dynamics. Our contribution is this generalization: first we note that we can use the step function S to rewrite the switch. Note

$$\left.\begin{array}{cl} a & \text{if } \ p < q \\ b & \text{otherwise} \end{array}\right\} = S(q - p)a + S(p - q)b \tag{3.90}$$

where

$$S(x) = \begin{cases} 1 & \text{if } \ x > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.91}$$

so we have

$$\begin{aligned} F_\kappa(\kappa, \mu_r, \phi) \quad &= \begin{cases} \min(\kappa, 0)\nabla\psi & \text{if } \ \mu_r(x, y) < \phi \\ \max(\kappa, 0)\nabla\psi & \text{otherwise} \end{cases} \\ &= S(\phi - \mu_r(x, y)) \min(\kappa, 0)\nabla\psi + S(\mu_r(x, y) - \phi) \max(\kappa, 0)\nabla\psi \end{aligned} \tag{3.92}$$

Note that the step function is the limit of the sigmoid $\mathbf{S}$:

$$\lim_{\beta\to\infty} \mathbf{S}(\beta, x) = \lim_{\beta\to\infty} \frac{1}{1 + e^{-2\beta x}} = S(x) \tag{3.93}$$

Denoting the sigmoid by $\mathbf{S}(\beta, u)$, the finite temperature switch is

$$F_\kappa(\kappa, \mu_r, \phi, \beta) = \mathbf{S}(\beta, \mu_r - \phi) \max(\kappa, 0)\nabla\psi + \mathbf{S}(\beta, \phi - \mu_r) \min(\kappa, 0)\nabla\psi. \tag{3.94}$$

With finite β, the switch is fuzzy and the dynamics are

$$\frac{d\psi}{dt} = a_1 F_\kappa(\psi, \mu_r, \phi, \beta) + a_2\Delta_n\psi + a_3\Delta_c\psi \tag{3.95}$$

These are the dynamics of a modified Hopfield Neural Network (23). The value of ϕ determines the degree to which a pixel's neighborhood influences its evolution. When ϕ is near zero, the neighborhood dictates, in part, the magnitude of the adjustment to a pixel.

3.4.10 Neighborhood Switch

Tatem et al have designed an energy functional term for abundance image resampling. Their switch adjusts the value of a pixel by measuring the average abundance in its neighborhood: if the average is high, then the adjustment is positive, if the average is low then the adjustment is negative. They have demonstrated its effectiveness in reconstructing discrete targets in abundance imagery.

In addition the generalization of Malladi and Sethian's we find a similarity between their work and Tantem's. In what follows we illustrate this similarity, and suggest modifications to their approach.

Instead of the first term in equation 3.95, Tatem et al use

$$F_\alpha(\tilde{\alpha}, \mu_r, \beta) = \frac{1}{2}\left[1 + \tanh \beta(\mu_r - 0.5)\right]\left[\alpha(x, y) - 1\right] + \frac{1}{2}\left[1 - \tanh \beta\mu_r\right]\alpha(x, y). \tag{3.96}$$

We can express this in terms of the sigmoid:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.97}$$

so that

$$\frac{1}{2}(1 + \tanh(x)) = \frac{e^x}{e^x + e^{-x}} = \mathbf{S}(x) \tag{3.98}$$

and

$$\frac{1}{2}(1 - \tanh(x)) = \frac{e^{-x}}{e^x + e^{-x}} = \mathbf{S}(-x) \tag{3.99}$$

Then

$$F_\alpha(\alpha, \mu_r, \beta) = \mathbf{S}(\beta, \mu_r - 0.5)(\alpha - 1) + \mathbf{S}(\beta, -\mu_r)\alpha \tag{3.100}$$

In analogy to the curvature flow switch, we introduce the gradient dependent velocity factor $\nabla\alpha$ to improve the convergence rate:

$$F_\alpha(\alpha, \mu_r, \beta) = \mathbf{S}(\beta, \mu_r - 0.5)(\alpha - 1)\nabla\alpha + \mathbf{S}(\beta, -\mu_r)\alpha\nabla\alpha \tag{3.101}$$

We now have two resampling techniques, the curvature flow switch and the neighborhood switch. The first was derived from a smoothness energy term. We saw that this smoothness energy implied dynamics governed by the diffusion equation. This was modified by inhibiting diffusion across regions of high gradient, yielding the dynamics of curvature flow. In making this modification, the resampling was now guided by known image geometry; the prior information was extracted from the image so the system was, in this sense, adaptive. Two further modifications ensued: the first was the inclusion of a switch to direct the evolution according to shape, rather than intensity; the second was a generalization of the switch designed to allow finer adjustments to pixel values when a solution was near, while retaining coarse adjustments far from a solution. This generalization was equivalent to casting the dynamics into a neural network in which a neuron's response is the fuzzy switch.

3.4.11 Neural Interpretation

Energy minimization may be interpreted as the dynamics of a Hopfield Neural Network (24), (25). Artificial neural networks are arrangements of processing units (neurons) connected to their neighbours. Artificial neural networks[2] are structures designed to harness the information processing capabilities of biological neural networks (26).

---

[2] 'artificial neural network' is commonly abbreviated to 'neural network'

Each neuron receives input from one or more of its neighbors. The total input, p_i, to a neuron is a linear combination of the input, ψ_j from each of its neighbours:

$$p_i = T_{i1}\psi_1 + T_{i2}\psi_2 + \dots T_{ip}\psi_p \tag{3.102}$$

where there are p neurons and where the matrix T is symmetric and with $T_{ii} = 0$. The output, ψ_i of the i^{th} neuron is a function of its input. In early models of neural activity this function was the step function,

$$\psi_i = S(p_i) = \begin{cases} 1 & \text{if } \ p_i > \phi \\ 0 & \text{otherwise} \end{cases} \tag{3.103}$$

for an appropriately chosen ϕ. In later models the sigmoid was chosen,

$$\psi_i = \mathbf{S}(h_i) = \frac{1}{1 + e^{-2\beta h_i}} \tag{3.104}$$

where β is called the *pseudotemperature* of the neuron. The step function is the infinite pseudotemperature limit of the sigmoid (Figure 3.10).

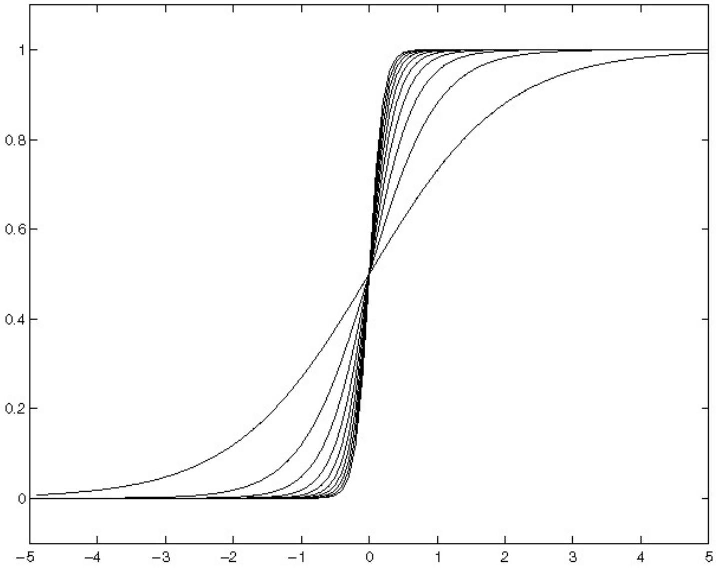


Figure 3.10: Sigmoid at various pseudotemperatures: when the psuedotemperature is infinite, the sigmoid is identical to the step function

A neuron at position x_i, initialized to ψ^0 is updated to $\mathbf{S}(p_i)$ upon receiving total input p_i from all other neurons:

$$\psi^0 \rightarrow \mathbf{S}(p_i) \tag{3.105}$$

Instead of computing the new neural value $\mathbf{S}(p_i)$ by summing all weighted neural inputs a là equation 3.102 we assume

$$\mathbf{S}(p_i) = \psi^0 + \frac{d\psi^0}{dt}. \tag{3.106}$$

We calculate the derivative with the use of an energy functional.

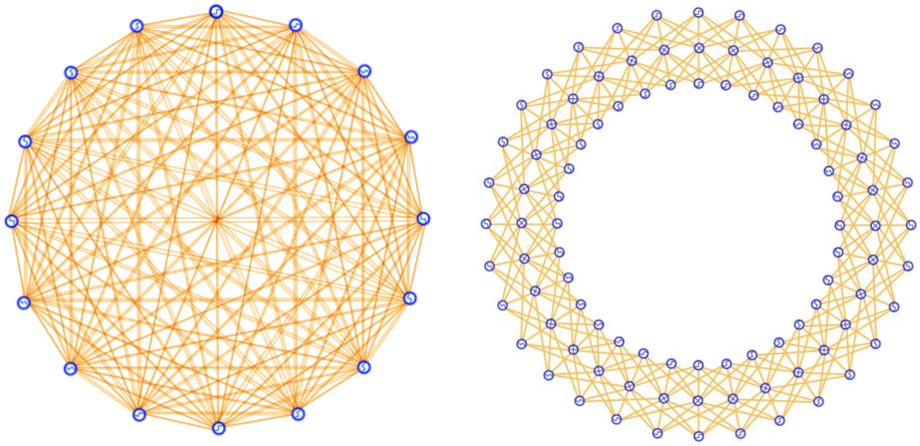


Figure 3.11: Traditional Hopfield Neural Network (panel 1) and the neural interpretation of the fuzzy switch, equation 3.94 (panel 2).

Hopfield introduces an energy, E, associated with the network

$$E = -\frac{1}{2}\sum_{i,j} T_{ij}\psi_i\psi_j + E_\Delta(\psi) \tag{3.107}$$

where E_Δ is an energy associated with a conserved "charge" on the network. With an energy function of this form he demonstrates that the energy decreases in time toward a stable state: "the convergence of the neuronal state of the continuous, deterministic model to its stable state is based on the existence of an energy function that directs the flow in state space" (25). By specifying an energy functional, we set the network weight matrix T_{ij} (equation 3.107) and we can be certain that the network will evolve toward a minimum of the energy functional.

According to Hopfield,

$$\frac{d\psi}{dt} = -\frac{\partial E}{\partial \psi}. \tag{3.108}$$

According to equation 3.106, then, a neuron is updated to

$$\mathbf{S}(p_i) = \psi^0 - \frac{\partial E}{\partial \psi}. \tag{3.109}$$

Thus, the process of passing data among neurons is the process of computation in a neural network: "Useful computations in this system involve the change of state of the system with time."

Further, Hopfield notes a dependence on the transfer function $\mathbf{S}$, "When the response is less steep, the continuous-response model can have fewer stable states." Thus, when the transfer function is mild, the number of spurious minima is reduced. We expect an advantage, then, when the finite temperature sigmoid is used rather than the step function.

We make two observations: first, the step-neural response S is reminiscent of the curvature flow switch. Second, the replacement of the step function with the sigmoid lead to fewer spurious minima in the neural network energy function. This suggests that replacing the curvature flow switch with a sigmoid could lead to improvements in the curvature flow performance.

When we don't know the network weight matrix T we need a means of calculating it. Often networks are trained on collections of data in order to set the elements of the matrix. Various techniques have been proposed for the training regime, such as back-propagation (23). Our super-resolution technique uses a hard-wired network; i.e. we know the connection matrix and it doesn't change. The connection strengths are the coefficients in the discretized curvature calculation.

Chapter 4

Tests and Methods

To evaluate the performance of the neural network resampling techniques presented in Chapter 3 we have used a testing procedure designed to measure the network performance relative to a conventional resampling technique.

Our choice of a representative from the group of conventional techniques is directed by two issues. First, in most cases of interest the image signal has not been sampled at the Nyquist rate. Also, we aim to restore Fourier content which has been damped by impulse response convolution. For these reasons, optimal linear interpolation is inapproprate. Second, we select a standard resampling technique with a track-record of good performance for most applications, polynomial interpolation.

In this chapter we describe the tests we used to characterize the performance of the super-resolution techniques presented in Chapter 3. We describe network operational modes, test, and standards of evaluation.

We adopt the following testing procedure:

A Select object signal

B Prepare image samples

C Select resampling technique, set parameters, and resample image signal data

D Measure network preformance

The object signal has been simulated by a collection of data with sample spacings X and Y. The image samples have been created by subjecting the object signal to a simulated sensor response; the image signal then had sample spacing zX and zY. The resampling technique may be either a conventional interpolation technique or a neural resampling technique; the paramters are set as appropriate. The resampled image signal has sample spacings X and Y. Our measures of performance will be comparisons between the resampled image signal and the object signal. In what follows we discuss each of the steps in this testing procedure.

4.1 Object Signal

Our object signal may be a reflectance signal or an abundance signal, where $|\vec{\alpha}| = 1$. In each case we test artificial and sensor-derived imagery. The object signal options are

A-1 Artificial reflectivity object signal, $\psi = \mathbf{r}$ (Figure 4.1, panel 1)

A-2 Sensor-derived reflectivity object signal, $\psi = \mathbf{r}$, (Figure 4.2, panel 1)

A-3 Artificial abundance object signal, $\psi = \alpha$, (Figure 4.1, panels 2 and 3)

A-4 Sensor-derived abundance object signal, $\psi = \alpha$. (Figure 4.2, panels 2, 3, and 4)

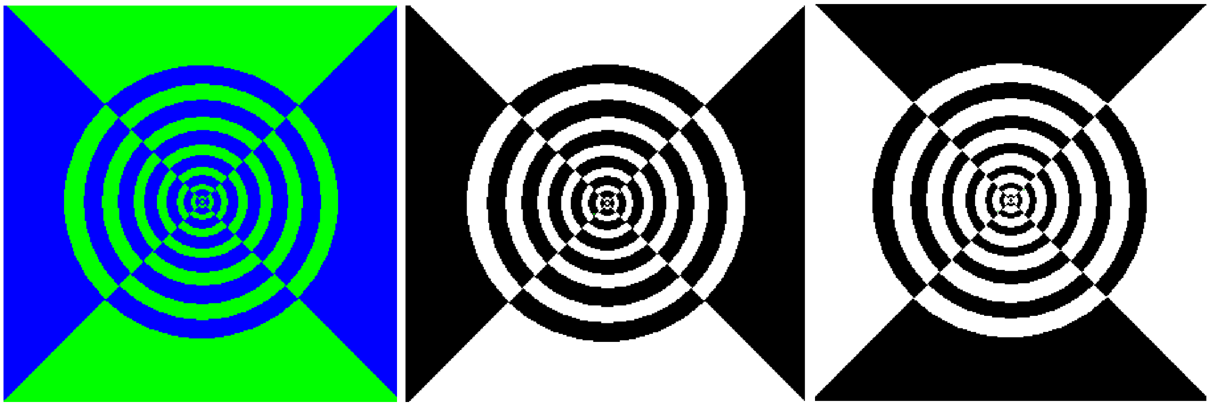


Figure 4.1: Arificial reflectance image (panel 1) and abundance images (panels 2 and 3).

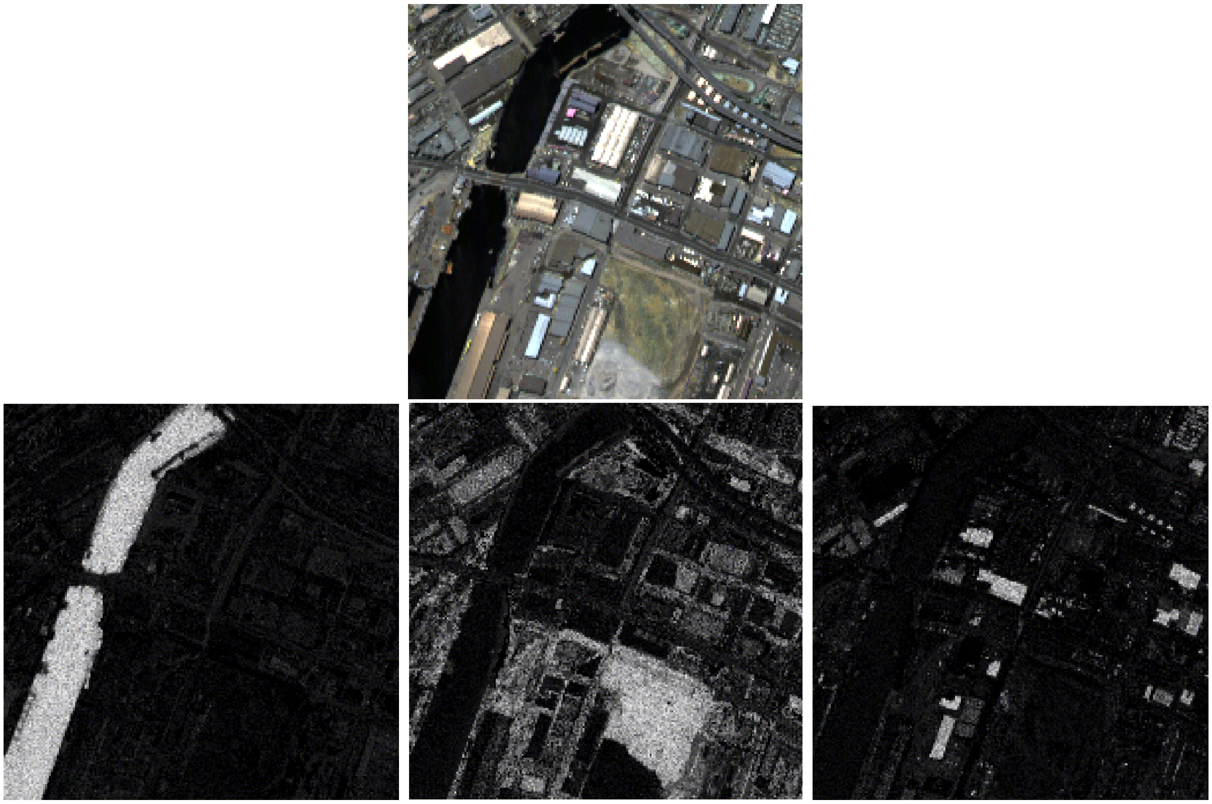


Figure 4.2: Sensor-derived reflectance image (top) and abundance images (panels 2,3, and 4).

4.1.1 Image Signal Samples

The image signal samples are derived from the object signal by simulating a sensor response. Our model for the impuse response is a Gaussian,

$$h_1(x, y) = N \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right). \tag{4.1}$$

where N is chosen so that the Gaussian is normalized to 1. We parametrize the noise as

$$n = k\hat{n} \tag{4.2}$$

where $\hat{n}$ is a random number between 0 and 1, and k is a constant between 0 and 1.

We will test resampling performance with $z = 3$ by varying the parameters σ_y and k independently. We index the stage of testing by the number $n = 1, \ldots 10$; the sensor model options are

B-1 $\sigma_x = 0.5$, $\sigma_y = n/5$, $k = 0.10$

B-2 $\sigma_x = 0.5$, $\sigma_y = 0.5$, $k = n/20$

We select these ranges because they span the range of reasonable impulse response standard deviation and noise ranges: the impulse response standard deviation cannot be zero and rarely exceeds twice the ground sample spacing; noise cannot be negative, and rarely exceeds 50%. In method B-1, then, the standard deviation in the x-direction is fixed at 0.5, the noise coefficient is fixed at 0.10, while the standard deveiation in the y-direction varies from 0.0 to 2.0. Similarly, in method B-2 the standard deviations are fixed at 0.5 while the noise varies from 0.05 to 0.5. Now, we specify which resampling technique is to be used.

4.2 Resampling Techniques

Resampling may be carried out by conventional interpolation techniques or by the neural techniques. We identify each of these possibilities as steps in the testing procedure.

The general data evolution model is

$$\frac{d\psi}{dt} = a_1 F_\kappa(\psi, \mu_r, \phi, \beta) + a_2 \Delta_n \psi + a_3 \Delta_c \psi. \tag{4.3}$$

We have the luxury of selecting the remaining parameter values, a_1, a_2, a_3, ϕ, and β. We require that $a_1 + a_2 + a_3 = 1$. Also, we select $a_2 \neq 0$ only when we resample abundance data. As above, we parameterize the tests by n. For reflectivity signals we must set a_2 to zero for abundance signals we tested the role of the normalization cofficient a_2:

C-1 polynomial interpolation, with polynomial order 3

C-2 $a_1 = n/10$, $a_2 = 0$, $a_3 = 1 - a_1$, $\phi = 0.5$, $\beta = 5.0$

C-3 $a_1 = 0.4$, $a_2 = 0.6n$, $a_3 = 0.6 - a_2$, $\phi = 0.5$, $\beta = 5.0$

C-4 $a_1 = 0.4$, $a_2 = 0$, $a_3 = 0.6$, $\phi = 0.1n$, $\beta = 5.0$

C-5 $a_1 = 0.4$, $a_2 = 0$, $a_3 = 0.6$, $\phi = 0.5$, $\beta = n$

Tests $C - 4$ and $C - 5$ have been conducted to assess the roles of β and ϕ

4.3 Performance Measures

In this section we introduce the quantities which have been measured in evaluating the algorithm performance.

The mean-squared error is a measure of the confidence that a resampled image accurately represents the object signal. The peak signal-to-noise ratio is a means of assessing the quality of an image independent of its intensity: a restored image may appear very dim relative to its object signal, yeilding a high mean-squared error. The peak signal-to-noise accounts for the scale of the image intensity to provide a meaningful assessment of the image information content.

Mean-squared error and peak signal-to-noise ratio do not accurately characterize the quality of edge restoration. Image sharpness is a description of the quality with which edges are represented; it is most easily measured in the spatial frequency domain where a measure of the signal-to-noise ratio at any frequency can be made. A signal to noise ratio measurement at a frequency u characterizes the quality with which $1/u$-wavelength features have been restored across the entire image.

4.3.1 Spatial Domain Assessment

We measured the mean-squared error of the restored image $\psi(x, y)$ relative to the object signal ψ_s,

$$e = \frac{1}{MN} \sum_{n=0}^{N-1} \sum_{M=0}^{M-1} [\psi_s(x, y) - \psi(x, y)]^2 \tag{4.4}$$

where the image has M rows and N columns. The mean-squared error depends on the image intensity scale; the Peak Signal-To-Noise Ratio is a standard characterization of the image quality, independen of intensity.

In the Frequency (Fourier) Domain we may characterize the performance of the fuzzy switch at any specified frequency. Thus we, we have a means of assessing the performance of the switch in restoring fine image details by measuring the signal-to-noise ratio in the highest spatial frequencies.

D1-1 **Peak Signal-to-Noise Ratio** avoids the dependence of the measure of image quality on intensity range by scaling the measure according to the maximum datum, s:

$$PSNR = -10 \log_{10} \frac{e_{ms}}{s^2}. \tag{4.5}$$

D1-2 **Frequency Domain Signal-to-Noise Ratio**: Let d be the difference image, $d(x, y) = \psi_s(x, y) - \psi(x, y)$ and let D be the Fourier Transform of d. The power of d on the annulus in the frequency domain, $R_i \equiv r_i \leq \sqrt{u^2 + v^2} < r_{i+1}$, is

$$P_d(R_i) = \frac{1}{\pi(r_{i+1}^2 - r_i^2)} \int_{r_i}^{r_{i+1}} \int_0^{2\pi} r^2 dr d\theta |D(r, \theta)|^2. \tag{4.6}$$

Similarly,

$$P_\psi(R_i) = \frac{1}{\pi(r_{i+1}^2 - r_i^2)} \int_{r_i}^{r_{i+1}} \int_0^{2\pi} r^2 dr d\theta |\Psi(r, \theta)|^2. \tag{4.7}$$

The signal-to-noise ratio over R_i is

$$SNR_{R_i} = \frac{P_\psi(R_i)}{P_d(R_i)}. \tag{4.8}$$

where r and θ are polar coordinates in the Fourier Domain.

Note that the frequency domain signal to noise ratio is not a logarithmic scale: because it is a quotient the frequncy domain signal to noise ratios may achieve very large values when the difference image intensiy is small. We should expect, therefore, to see large values on occasion. The frequency domain signal to noise ratio is a quotient of powers and is, therefore, unitless.

4.3.2 Additional Measurements

Additional observations were made to characterize the neural network. We take not of the following:

D2-1 **Final State** is the the state of the network - the image - when the absolute values of the difference between successive measurements of the mean-squared error is less than a tolerance, ϵ.

D2-2 **Convergence** is achieved if the system reaches a final state

D2-3 **Stablility** is achieved if the system converges and if the convergence is monotonic.

4.4 Performance Standards

We define the terms 'excellent', 'good','satisfactory', and 'poor' as follows: if the PSNR (or frequency domain SNR) of a Hopfield–resampled image is a, and if the PSNR (frequency domain SNR) of a polynomial–resampled image is b then define the quotient

$$r = \frac{a}{b}.$$

This quotient characterize the relative performance of the Hopfield network:

	excellent	good	satisfactory	poor
r	> 1.0	[0.8,1.0)	[0.8,0.6]	<0.6

Table 4.1: Image reampling measures of peformance. For example, if the network PSNR is 1.0 times polynomial PSNR, or higher, then the performance is excellent

Chapter 5

Results

Testing has been carried out in order to assess the performance of various of network configurations under reasonable variations in the sensor model. With the results we can anticipate and network performance when sensor model details are unknown. In this chapter we summarize the results of the testing sequences. At the end of the chapter several resampled images are displayed (Figures 5.7 and 5.8).

The figures which follow depict the fuzzy switch-to-polynomial performance ratio versus impulse response standard deviation, noise, normalized spatial frequency, time, and various neural parameters, The alternating grey and white bands denote regions of performance standards; the uppermost white band in each figure, for instance, represents excellent performance.

5.1 Sensor Model Variation Response

5.1.1 Reflectance Imagery

Impulse Response Width Variation

In artificial imagery we observed a decrease in performance ratio with an increase in the impulse response width, illustrated in Figure 5.1a. This implies that as aliasing decreases and blurring increases, relative performance decreases; that is, as aliasing increases, the advantage of the switch increases. The performance is excellent even at its lowest point where $\sigma_y = 2.0$.

Fourier domain measurements indicate excellent performance in most bands when resampling artificial imagery. The trend at low standard deviation ($\sigma_y = 0.2$, Figure 5.1c) is repeated at moderate ($\sigma_y = 1.0$, e) and high standard deviation ($\sigma_y = 2.0$, g): most frequency bands receive excellent treatment; the upper-mid frequency bands (normalized wavelength range of 12 - 15) have consistently lower performance ratio. These measurements indicate that broad features and fine details of the artificial image have been restored effectively. Only features on the order of 10 pixels in breadth are not treated adequately.

The decrease in performance ratio as impulse response breadth increases is repeated in sensor-derived imagery (b). The magnitude of the decrease is, again, on the order of 0.015. We note that the performance ratio obtains about the same values in artificial and sensor-derived imagery.

The frequency domain trend in sensor derived imagery at low standard deviation (d), is repeated in moderate (f), and high (h) standard deviations: nearly all features are recovered well. The performance ratio is much greater than 1.0 in many bands; only in mid-upper bands have we observed less-than-excellent performance. Spikes consistently appear among the performance ratio troughs in the mid-frequency bands.

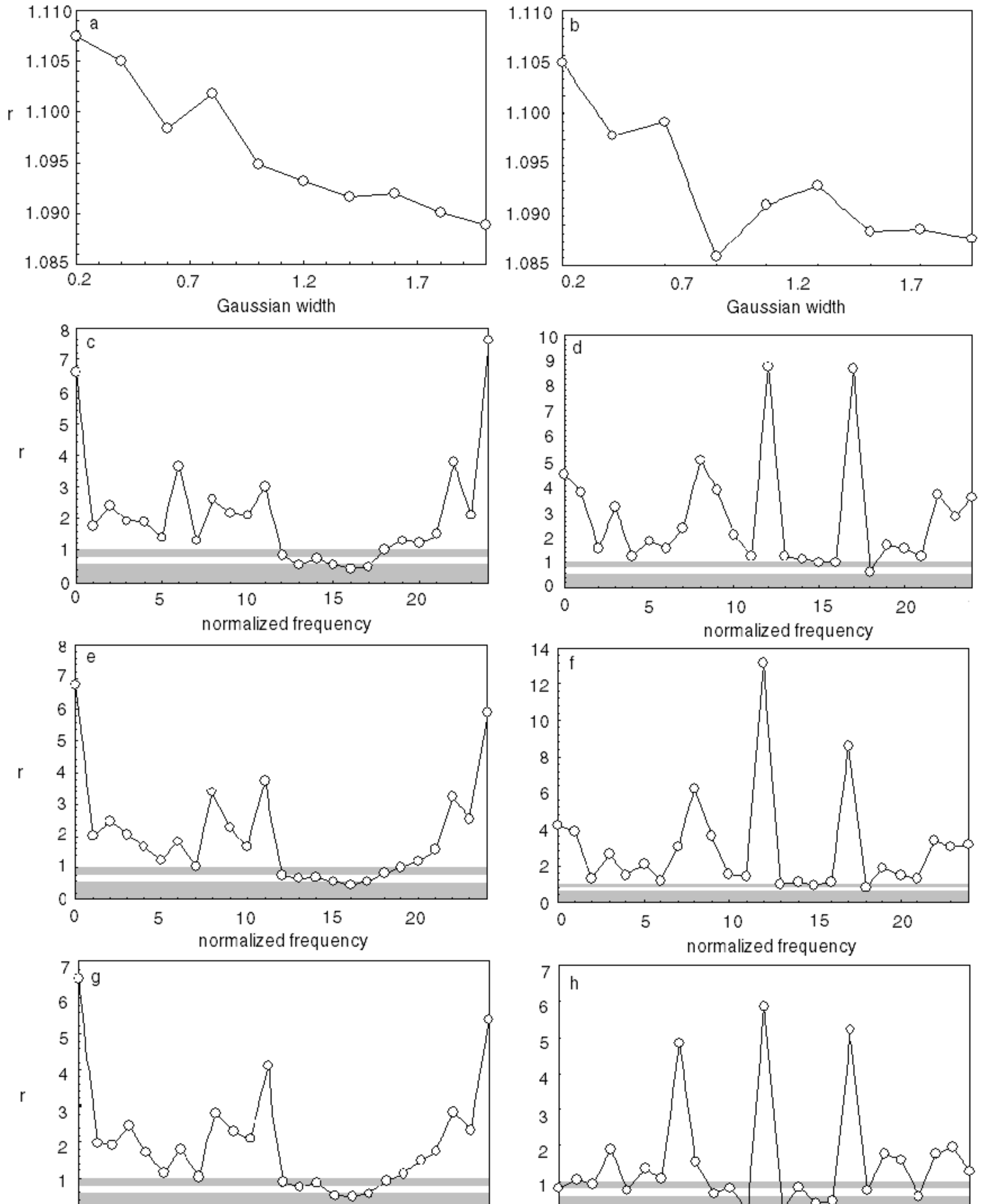


Figure 5.1: Impulse response width dependence: reflectance imagery

As observed in the artificial imagery, the performance ratio decreases as the impulse response breadth increases, consistent with the measurements indicated in graph (b).

We attribute the low performance ratio in the upper-mid frequencies to the regularization effects of diffusion: when the gradient is low, diffusion dominates. When the variations are on the order of 10 pixels, the diffusion is sufficiently strong to penetrate the gradient; those regions are illegitimately regularized. In contrast the only gradient in low-frequency regions is noise, and regularization by diffusion is appropriate. In high frequency regions, the gradient is sufficiently high to inhibit diffusion.

Noise Variation

As depicted in Figure 5.2a, the performance ratio when resampling artificial reflectance imagery increases as noise increases. The increase is on the order of 0.03. This suggests that noise removal by curvature flow is effective. The advantage of the fuzzy switch over polynomial resampling increases with increasing noise.

We have observed excellent artificial image Fourier mode restoration except in upper-mid frequencies; there the performance is poor. This trend, indicated in Figure 5.2c, is repeated as the noise level increases (graphs (e) and (g)). This trend implies that fine image details are effectively restored; features roughly of size 10 pixels are not well-restored. In addition, large features are restored well, as indicated by the lowest modes.

A monotonic increase in performance ratio as noise increases was observed when resampling sensor-derived imagery (Figure 5.2 (b)). As with artificial imagery, this indicates that the fuzzy switch offers increasing advantage as noise increases.

This increase in the performance ratio is echoed in the Fourier domain measurements of graphs d, f, and h. At only a few modes is the performance ratio below one; at most mode the ratio is in the neighborhood of three. The greatest advantage was observed in the highest bands, implying that fine details are restored well (relatively) amidst high noise.

5.1.2 Abundance Imagery

Impulse Response Width Variation

As indicated in Figure 5.3a, performance is consistently excellent in PSNR performance ratio measurements.: there is an advantage over polynomial resampling as indicated by the peak signal-to-noise ratio – a global image quality measurement. The monotonic decrease with increasing impulse response width implies greatest advantage is observed at high aliasing and low blurring. This advantage decreases by 0.04 as the width increases from 0.2 to 2.0.

The Fourier domain measurements exhibit an enormous advantage in many bands, most notably in the highest frequencies. This implies that fine detail reconstruction is superb.

A similar trend in PSNR performance ratio measurements is observed in sensor-derived abundance imagery resampling: the performance is always excellent, with a monotonic decrease in the performance ratio as aliasing decreases and blurring increases. The change in the performance ratio is about 0.1.

Fourier domain measurements made on the sensor-derived abundance images as impulse response width increases (Figure 5.3d, f, and h) indicate that the fuzzy switch offers advantage over polynomial resampling in all bands except the zero mode. This advantage persist as with width increases; the greatest advantage is observed in the highest Fourier modes. Again, this implies that the fuzzy switch is excellent at restoring fine image details.

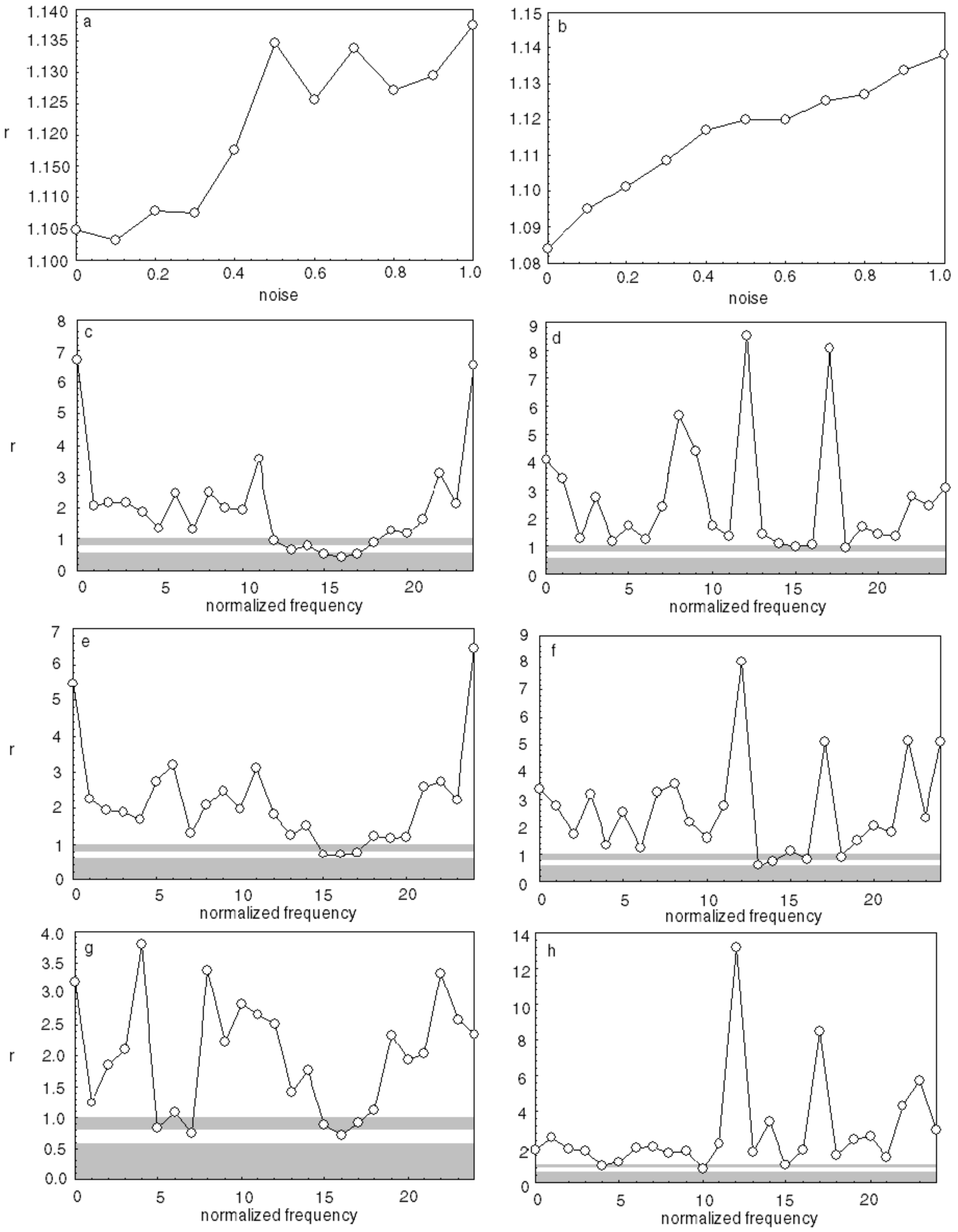


Figure 5.2: Noise dependence: reflectance imagery

Noise Variation

Measurements on artificial abundance imagery (Figure 5.4a) indicate that the performance was consistently excellent; the performance ratio increased by about 0.02 as noise increased. Though this indicates consistently

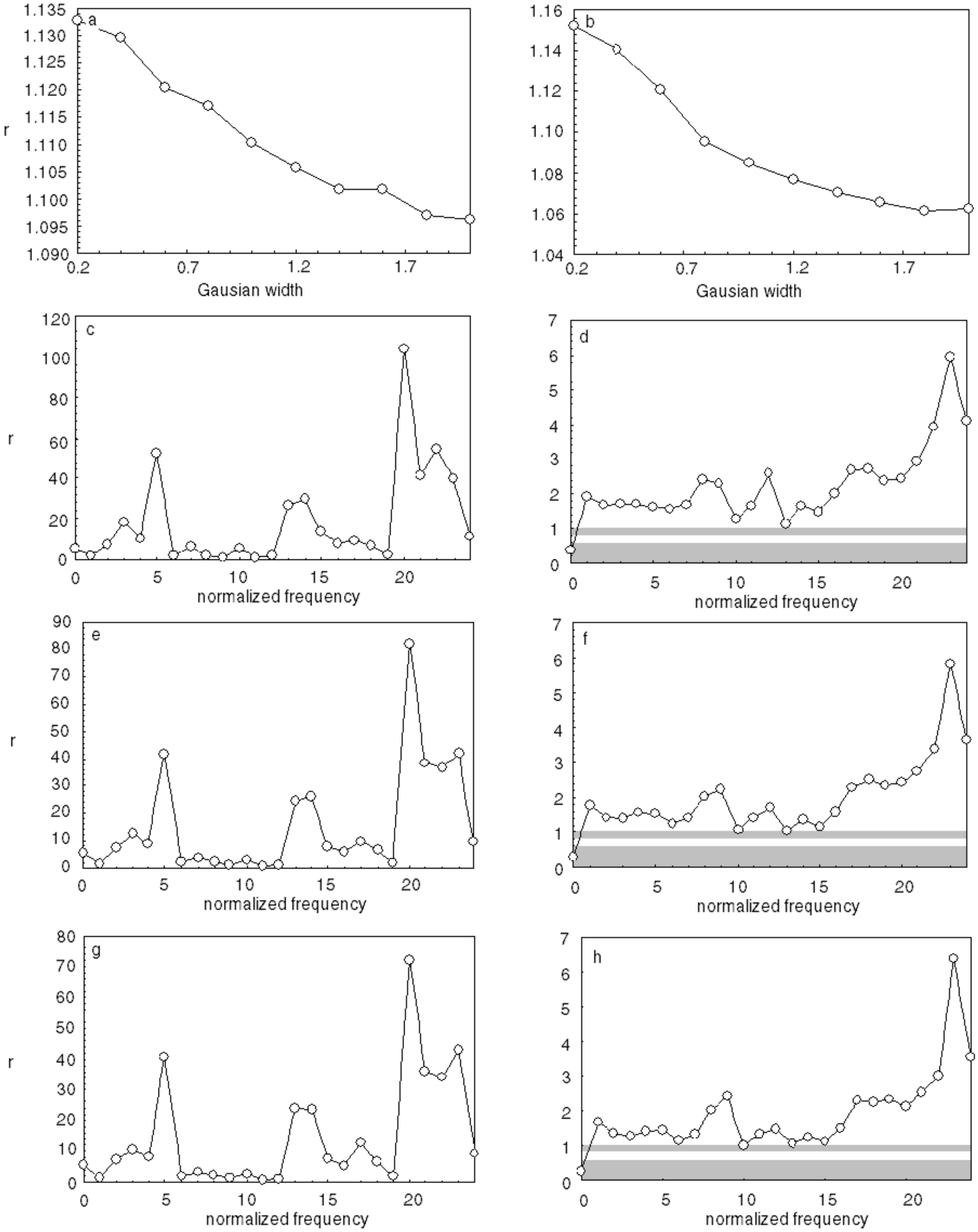


Figure 5.3: Impulse response width dependence: abundance imagery

excellent performance, the trend as noise increases is somewhat irradic.

Artificial abundance image Fourier domain measurements indicate excellent performance in majority of bands (Figure 5.4c, e, g). We observe the greatest performance ratio in the highest bands, indicating excellent detail restoration. The performance ratio in low and moderate frequency modes is lower. We observe this trend as blurring increases/aliasing decreases, though the advantage in each Fourier band decreases as blurring increases. We note, also, that the performance in the mid-bands is somewhat irradic.

In contrast with other noise tests, we have observed a decrease in the performance ratio with increasing noise When resampling sensor-derived abundance imagery. The decrease amounts to about 0.03, though this decrease is irradic. We attribute this behaviour to irregular image geometry: as curvature is not well-corellated spatially, it does not serve as an adequate guide for resampling.

Fourier domain measurements indicate that the performance ratio is excellent in all bands except the zero mode. Again, this indicates that fine details are restored well by the fuzzy switch. This trend persists as noise increases (graphs d, f, and h).

5.2 Neural Characterization

5.2.1 Reflectance Imagery

Figure 5.5, graphs a and b depict the increasing performance ratio versus time when resampling artificial and sensor-derived reflectance imagery. The increase in performance ratio with time indicates an increasing PSNR and decreasing mean-squared error. The time series is monotonic on the range [0,30] iterations, with a final state typically reached after about 40 iterations. On occasion, instability occurred in preliminary tests when the time step was set very large. In addition, the conservation term played a role in creating instability by conflicting with the fuzzy switch; the conflict occasionally induced oscillations until the stepsize or a_3 was decreased.

Graphs c and d depict the dependence on the coefficient of self-diffusion a_1. Both indicate a maximum near $a_1 = 0.4$, suggesting that conservation plays a role in improving performance. This improvement in performance, we note, is achieved when the time step is sufficiently small (0.1). The role of the conservation term and its influence on instability deserves further study.

Graphs e and f indicate dependence on the global neural bias ϕ. The dependence is weak with no clear trend: observations e indicate greater values of ϕ give higher performance ratio while observations f indicate lower values of ϕ give higher performance ratio. We make no conclusions and direct the reader to variable neural bias, as discussed by Malladi and Sethian, for further development options.

Graphs g and h indicate dependence on neural pseudotemperature; here, too, the dependence is modest with no clear trend. The absence of significant dependence on pseudotemperature indicates little advantage in replacing step function with sigmoid. That is, we observe little performance advantage in moving to a neural regime. The neural regime, however, has highlighted a relationship between Malladi's work and Tatem's work.

5.2.2 Abundance Imagery

We have observed excellent performance throughout the entire time series when resampling artificial and sensor-derived abundance imagery. The time series we have observed are monotonic in the vast majority of cases, as indicated in Figure 5.6a and b. A final state is typically reached within 45 iterations, such as in graph b. Instability occurred when the time step was large: the fuzzy switch conflicted with conservation and normalization terms, inducing oscillations

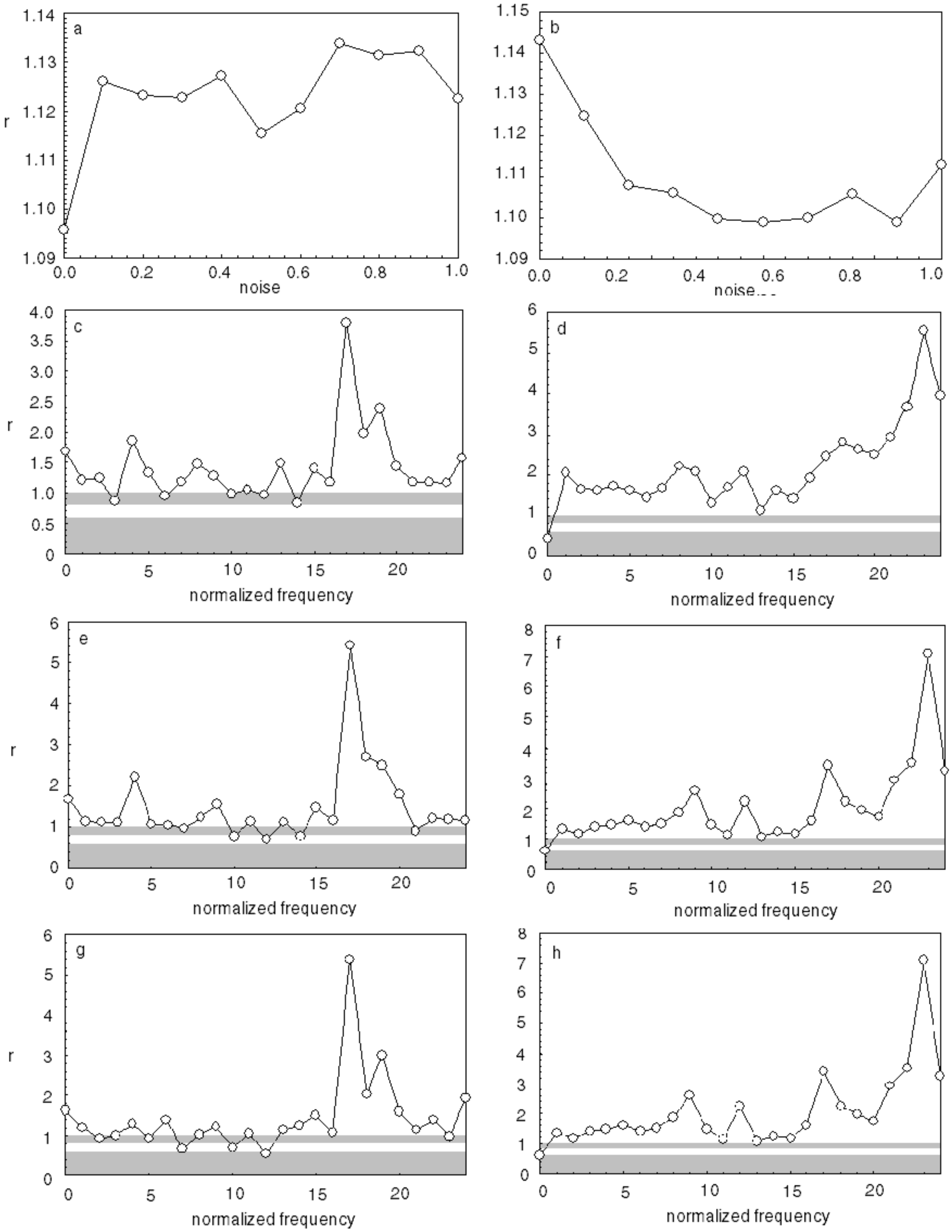


Figure 5.4: Noise dependence: abundance imagery

Graphs c and d indicate maximum performance ratio is achieved when the normalization term is in the neighborhood of 0.1. Normalization thus appears to be an important constraint in evolving abundance

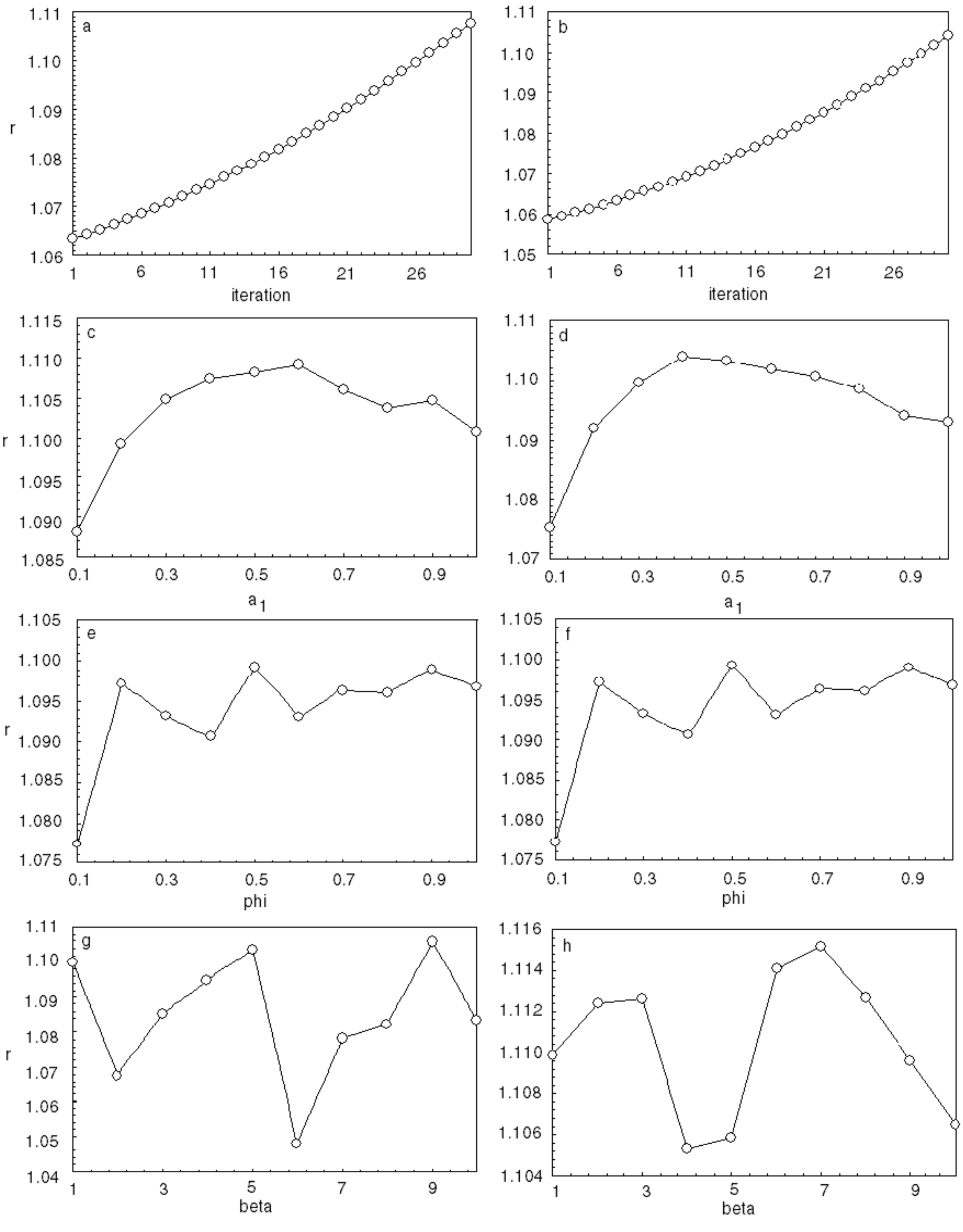


Figure 5.5: Neural characterization: reflectance imagery

imagery. However, very strong normalization conflicts with fuzzy-switch: oscillations we have observed in the time series led to growing errors.

Graphs e and f indicate a weak dependence on global bias, with no clear trend in improved performance. The variable neural bias suggested by Malladi and Sethian may be investigated as a possible improvement to the technique,

Graphs g and h indicate very weak dependence on pseudotemperature. Little performance benefit observed on $0 < \beta \leq 10$. Therefore, the neural regime offers little performance increase. That is when the sigmoid is smooth, we see no consistent improvement in the performance ratio.

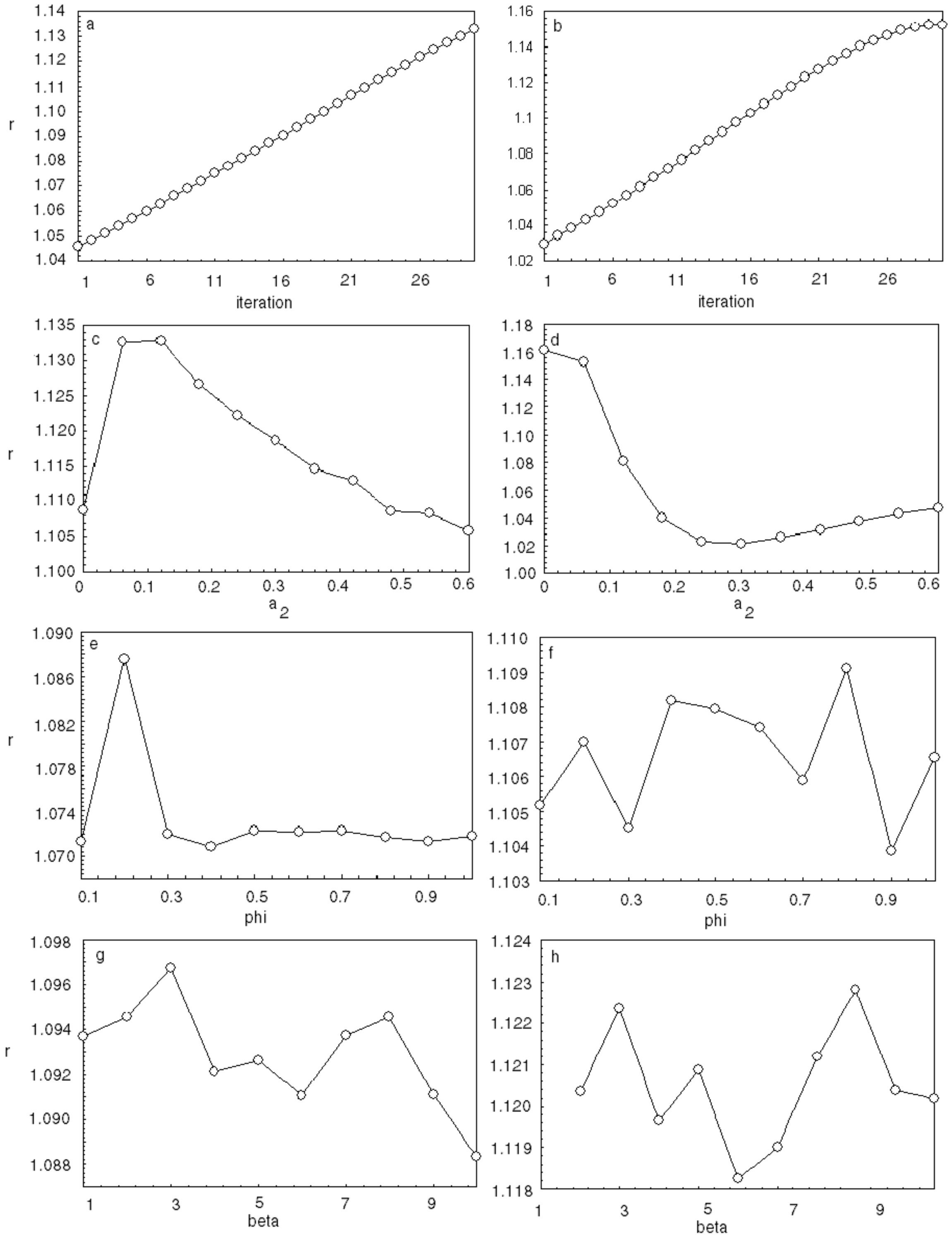


Figure 5.6: Neural characterization: abundance imagery

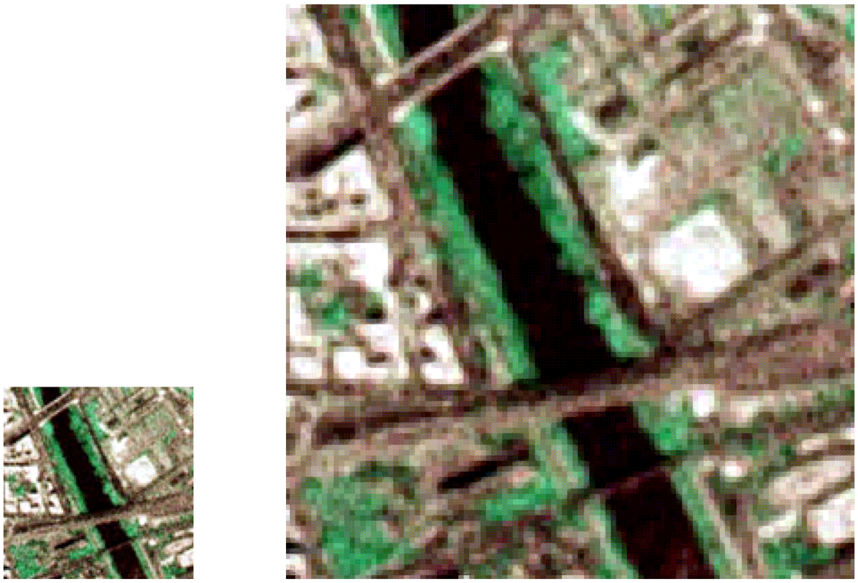


Figure 5.7: Rome

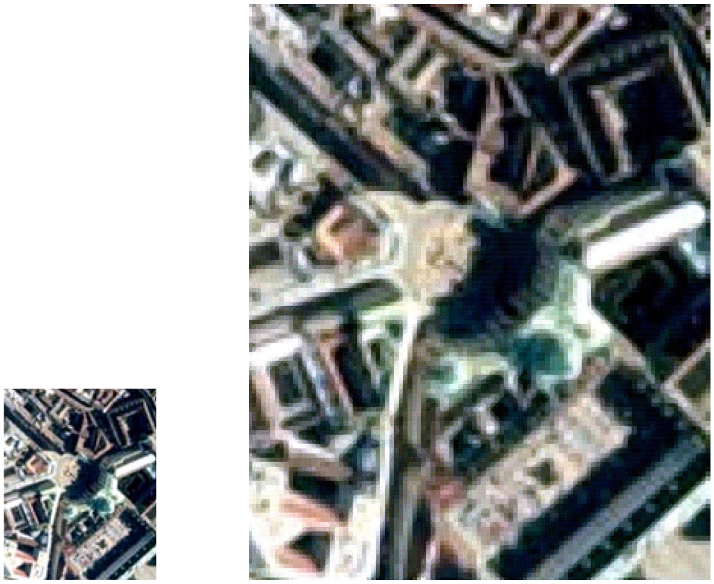


Figure 5.8: Vienna

Chapter 6

Discussion

6.1 Summary

We summarize the main components of the fuzzy switch resampling technique:

- **Image geometry guides resampling** by measuring image curvature. In curvature flow, the adjustment to a pixel depends on the local image curvature; image curvature, in turn, is computed by making gradient measurements. Curvature, however, is insufficient to guide the resampling

- **The curvature flow switch directs the flow** because image geometry is not properly assessed by curvature measurements. The direction of curvature flow depends on the sign of the curvature; thus, geometrically equivalent shapes may evolve in opposite directions.

- **The fuzzy switch is the curvature flow switch cast into the neural regime.** In an attempt to avoid spurious minima, we promoted the step function in the curvature flow switch to the sigmoid – a model of the neural response. Doing so, we found a similarity between the work of Malladi and Sethian and of Tatem et al.

We have seen that the image reconstruction problem is underdetermined as a result of the singular sampling function. Image formation artifacts such as blurring, aliasing, and noise complicate the problem. To render the resampling/restoration problem well-defined, we need to supply prior information. Prior information in convention techniques is a functional relationship imposed on the data: in many conventional techniques, we describe the recorded data by function fitting, and resample the image by evaluating the function at the fine mesh points.

An alternative method, adopted in this thesis, is to employ image geometry to guide the resampling. Early models guided the resampling by the image gradient. In those cases, adjustments to a pixel we proportional to the Laplacian of the image intensity; this approach was equivalent to simulating data diffusion. Noise and aliasing were treated as defects in the image gradient and were diffused away; the diffusion was halted by employing auxiliary constraints.

Later models emerged when diffusion was inhibited across regions of high gradient, where an edge was suspected. In doing so, the diffusion was anisotropic – the curvature flow dynamics. Curvature flow techniques were promising for resampling and restoration for several reasons: Fourier modes were synthesized as diffusion took place on either side of a gradient, while the gradient sharpened as a result of mean curvature minimization; blurring was treated by Grayson's theorem – regions collapse to their minimum area allowed by the auxiliary constraints; local diffusion treated noise; and aliasing was treated as planar mean curvature was minimized.

Curvature flow suffers at least one problem: geometrically equivalent shapes may evolve in opposite directions. The direction of the flow depends on the sign of the gradients. Malladi and Sethian have introduced a switch to encourage geometrically equivalent shapes to evolve identically. We found that we can make sense of their technique by casting the switch into a neural framework with the use of the sigmoid. In doing so, we found that their switch resembles the work of Tatem et al. The similarity between these techniques suggests future improvements and courses of research.

We have tested the fuzzy curvature flow switch to evaluate its performance relative to polynomial resampling. Because polynomial interpolation is common with adequate performance, it has served as a suitable standard of comparison. We have found that the performance of the fuzzy switch is consistently excellent relative to polynomial resampling.

- **Global image quality** is excellent in all tests. It was assessed by measuring the peak signal-to-noise ratio of fuzzy switch resampled and polynomial resampled images. The advantage of the fuzzy switch it typically a minimum of 10% over polynomial resampling in PSNR.
- **Image details are restored** as indicated by image quality measurements in the highest Fourier bands. The greatest advantage of the fuzzy switch was frequently observed in the highest Fourier modes. This implies that edged and discrete features are effectively restored by the fuzzy switch.
- **Convergence toward a final state is rapid**, usually on the order of 45 iterations. In comparison, Tatem et al quote several thousand iterations required to reach a final state. Instabilities may occur in any iterative technique: when the time step is too large, the system may not reach a minimum and oscillate. We have observed such oscillations in this programme; the auxiliary constraints have occasionally induced oscillations in the system.
- **Artifact are effectively treated**: we observed that the fuzzy switch offers greatest advantage when image formation artifacts are severe. We conclude that the phenomenolgy of curvature flow plays an important role in restoring images.

6.2 Recommendations

Future work may be directed in at least one of at least two directions: one may further study the performance of the switch as presented in this thesis, or one may investigate improvements to the switch. Performance studies may include the following:

- the results in this thesis may be supported by other tests,
- the range of parameter (β, ϕ, neighborhood size) testing may be extended or studied in finer detail,
- the role of the conservation and normalization terms in inducing oscillations,
- a detailed performance comparison between the fuzzy switch and Tatem's switch,
- the role of the switch in halting the evolution, and
- different iterative and non-iterative implementations of the technique

Further research and development may include

- devising a generalized switch, thereby establishing a concrete relationship between Malladi's work and Tatem's work,
- improvements to the fuzzy switch or the neighborhood switch, such as symmetrization, variable neural bias, gradient-dependent velocity factors, etc., and
- implementing learning regimes.

Appendix A

Sensor Radiance Model

The radiant intensity of the sun is its flux Φ per unit solid angle ω (3)

$$I = \frac{d\Phi}{d\omega}. \tag{A.1}$$

If the sun emits light equally in all directions, then the radiant intensity is constant on the surface of any sphere enclosing the sun. The radiant intensity is independent of the distance between the sun and the surface of the sphere. The flux per unit area, the irradiance,

$$E = \frac{\Phi}{A} \tag{A.2}$$

where A is the area of the sphere, depends on the distance between the sun and the sphere. Noting that $r^2 d\omega = dA$ we have

$$\frac{d\Phi}{d\omega} = r^2 \frac{d\Phi}{dA} \tag{A.3}$$

and

$$I = r^2 E. \tag{A.4}$$

I.e. the radiant intensity is the angular density of radiation, while the irradiance is the angular density scaled by the distance separating the distance to the surface.

The radiance L describes the radiant intensity per unit are on the sphere; it depends on the brightness of the sun and the separation between the sun and the surface of the sphere. It is defined by

$$L_\perp = \frac{dI}{dA} \tag{A.5}$$

In general, the radiance on a surface oriented at an angle θ_1 relative to the wavevector is

$$L = \frac{dI}{dA \cos\theta_1} \tag{A.6}$$

The signal recorded by the sensor is the radiance from the sun reflected off the earth and subject to extinction by the atmosphere. The total extinction depends on the optical thickness, τ, of the atmosphere and the pathlength through the atmosphere. If the atmosphere is uniform with a thickness t then the direct radiance at the earth's surface is

$$L_e(x,y,\lambda) = L_T(x,y,\lambda)e^{-\tau(\lambda)t\sec\theta_1}. \tag{A.7}$$

The reflectivity of the surface is defined as the ratio of the reflected (pixel-leaving) radiance to the incident radiance

$$r(x,y,\lambda) = \frac{L_p(x,y,\lambda)}{L_e(x,y,\lambda)} \tag{A.8}$$

This implies that the contribution L_1 of this direct reflection to the sensor radiance is

$$L_1(x,y,\lambda) = r(x,y,\lambda)L_e(x,y,\lambda)e^{-taut(\lambda)\sec\theta_2} \tag{A.9}$$

where θ_2 is the zenith angle of the satellite.

In addition, the environmental surrounding on the surface contribute the sensor radiance. The sensor radiance is modeled at the contributions of eight components(27):

1 **The direct solar radiation reflected from the pixel**,

$$L_1(x,y,\lambda) = r(x,y,\lambda)L_e(x,y,\lambda)e^{-\tau(\lambda)t\sec\theta_2} \tag{A.10}$$

2 **The solar radiation scattered downward and reflected from the pixel**

$$L_2(x,y,\lambda) = r(x,y,\lambda)F(x,y,\lambda)L_d(x,y,\lambda)e^{-\tau(\lambda)t\sec\theta_2} \tag{A.11}$$

where L_d is the downwelling radiance onto the pixel from skylight and where F is the fractional size of the sky

3 **solar radiation scattered upward into the sensor aperture**

$$L_3(x,y,\lambda) = L_u(x,y,\lambda) \tag{A.12}$$

4 **emitted radiation from the pixel** at an emissivity of ϵ_p,

$$L_4(x,y,\lambda) = L_{bb}(x,y,\lambda)\epsilon_p(x,y,\lambda)e^{-\tau(\theta)t\sec\theta_2} \tag{A.13}$$

where $L_{bb}(x,y,\lambda)$ is the blackbody radiance at wavelength λ

5 **radiation emitted by the atmosphere downward and reflected upward** contributes an amount proportional to the solid angular size of the sky and the reflectivity of the earth,

$$L_5(x,y,\lambda) = r(x,y,\lambda)FL_d(x,y,\lambda)e^{-tau(\lambda)t\sec\theta_2} \tag{A.14}$$

6 **radiaton emitted by the atmosphere upward into the sensor aperture**,

$$L_6(x,y,\lambda) = L_{bb}(x,y,\lambda)\epsilon_a(x,y,\lambda) \tag{A.15}$$

where ϵ_a is the emissivity of the atmosphere

7 **solar radiation reflected from the background into the pixel**,

$$L_7(x, y, \lambda) = (1 - F)L_b(x, y, \lambda)e^{-\tau(\lambda)t\sec\theta_2} \tag{A.16}$$

8 **radiation emitted by the background and reflected from the pixel**

$$L_8 = (1 - F)\epsilon(x, y, \lambda)rL_{bb}\lambda e^{-\tau(\lambda)t\sec\theta_2} \tag{A.17}$$

Altogether, the sensor radiance is

$$\begin{aligned} L_s &= rL_e e^{-\tau t\sec\theta_2} + rFL_d e^{-\tau t\sec\theta_2} + L_u + L_{bb}\epsilon_p e^{-\tau t\sec\theta_2} \\ &+ rFL_d e^{-\tau t\sec\theta_2} + L_{bb}(T_a)\epsilon_a + (1 - F)L_b e^{-\tau t\sec\theta_2} + (1 - F)\epsilon rL_{bb}(T_b)e^{-\tau t\sec\theta_2} \end{aligned} \tag{A.18}$$

This equation is linear in L_e; we abbreviate it as

$$L_s(x, y, \lambda) = c(\lambda)r_p(x, y, \lambda)L_e(x, y, \lambda) + n_1(x, y, \lambda) \tag{A.19}$$

Bibliography

[1] Malladi, R., J. A. Sethian, *Image Processing: Flows Under Min/Max Curvature and Mean Curvature*, Graphical Models and Image Processing, **58**, 2, p. 127-141, 1996

[2] Tatem, A. J., et al, *Land Cover Mapping at the Sub-Pixel Scale Using a Hopfield Neural Network*, Proc. 28th Symposium on Remote Sensing of the Environment, p. 47-51, 2000

[3] Pedrotti, F. L., L. S. Pedrotti, *Introduction to Optics, Second Edition*, Prentice Hall, Inc, 1993

[4] Jiang, M., *Mathematical Models in Computer Vision and Image Processing*, http://ct.radiology.uiowa.edu/jiangm/courses/mm-cv-ip/

[5] Basedow R., et al, Proceedings of the International Symposium on Spectral Sensing Research, **1**, 430, 1992

[6] Crochiere, R. E., L. R. Rabiner, *Interpolation and Decimation of Digital Signals - A Tutorial Review*, Proceedings of the IEEE, 1981, **69**, 3, p. 300-331

[7] Jerri, A. J., *The Shannon Sampling Theorem - Its Various Extensions and Applications: A Tutorial Review*, Proceedings of the IEEE, 1977, **65**, 11, p. 1565- 1596

[8] Brigham, E. O., *The Fast Fourier Transform and its Applications*, McGraw-Hill Book Company, 1988

[9] Kruse, F. A., et al, *Mineral Mapping at Cuprite, Nevada, with a 63 channel imaging spectrometer*, Photogram. Eng. Remote Sensing, **40**, 1989

[10] Bellucci, G., *Imaging Spectroscopy of Planetary Surfaces: Improving the Spatial Contrast*, Astronomy and Astrophysics Supplement Ser. **134**, p. 187-192

[11] Andrews, H. C., B. R. Hunt, *Digital Image Restoration*, Prentice-Hall, 1976

[12] Hunt, B.R., *Super-Resolution of Imagery: Understanding the Basis for Recovery of Spatial Fequencies Beyond the Diffraction Limit*, Proc. IEEE, 1999

[13] Whittaker, E. T., *On the functions which are represented by the expansion of interpolating theory*, Proc. Roy. Soc. Edinburgh, **35**, p. 181-194, 1915

[14] Press, W. H., et al, *Numerical Recipes in C The Art of Scientific Computing*, Cambridge University Press, 1997

[15] Morse, B. S., D. Schwartzwald, *Isophote-Based Interpolation*; Proc. IEEE 1998

[16] Allebach, J., P. W. Wong, *Edge Directed Interpolation* Proc. IEEE, 1996

[17] Carrato, S., et al, *A Simple Edge-Sensitive Image Interpolation Filter*, Proc. IEEE, 1996

[18] Reif, R., *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill Inc., 1965

[19] Koch, C., J. Marroquin, A. Yuille, *Analog "neuronal" networks in early vision*, Proceedings of the National Academy of Sciences, USA, **83**, p. 4263-4267, 1996

[20] Brakke, K. E., *The Motion of a Surface by its Mean Curvature*, Princeton University Press, 1978

[21] Rhines, F. N., et al, *Mechanism of Steady-State Grain Growth in Aluminum, Metallurgical Transactions*, **5**, p. 413-425, 1974

[22] Grayson, M. A. *The heat equation shrinks plane curves to round points*, Journal of Differential Geometry, **26**, 1997

[23] Hertz, J., A. Krough, R. G. Palmer, *Introduction to the theory of neural computation*, Addison-Wesley, 1991

[24] Hopfield, J. J., *Neural networks and physical systems with emergent collective computational abilities*, Proceedings of the National Academy of Sciences, USA, **79**, p. 2554-2558, 1982

[25] Hopfield, J. J., *Neurons with graded response have collective computational properties like those of two-state neurons*, Proceedings of the National Academy of Sciences, USA, **81**, p. 3088-3092, 1984

[26] Rojas, R., *Neural networks: a systematic introduction*, Springer-Verlag, 1996

[27] Collins, M. J., *Quantitative Remote Sensing*, unpublished

[28] Veldhuizen, T. L., *Grid Filters for Local Nonlinear Image Restoration*, http://osl.iu.edu/t̃veldhui/papers/MAScThesis/thesis.html

[29] Wu, H.-H. P., R. Schowengerdt, *Improved Estimation of Fraction Images Using Partial Image Restoration*, IEEE Transactions on Geoscience and Remote Sensing, **31**, 4, 1993.

[30] Young, I. T., J.J. Gerbreands, L,T, van Vliet, *Image Processing Fundamentals*, http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip.html