Paul Fieguth

Statistical Image Processing and Multidimensional Modeling

July 24, 2010

Springer

Berlin Heidelberg New York
Hong Kong London
Milan Paris Tokyo

1

Introduction

Images are all around us! Inexpensive digital cameras, video cameras, computer webcams, satellite imagery, and images off the Internet give us access to spatial imagery of all sorts. The vast majority of these images will be of scenes at human scales — pictures of animals / houses / people / faces and so on — relatively complex images which are not well described statistically or mathematically. Many algorithms have been developed to process / denoise / compress / segment such images, described in innumerable textbooks on image processing [36, 54, 143, 174, 210], and briefly reviewed in Appendix C.

Somewhat less common, but of great research interest, are images which do allow some sort of mathematical characterization, and to which standard image-processing algorithms may not apply. In most cases we do not necessarily have *images* here, per se, but rather spatial datasets, with one or more measurements taken over a two- or higher-dimensional space.

There are many important problems falling into this latter group of scientific images, and where this text seeks to make a contribution. Examples abound throughout remote sensing (satellite data mapping, data assimilation, sea-ice / climate-change studies, land use), medical imaging (denoising, organ segmentation, anomaly detection), computer vision (textures, image classification, segmentation), and other 2D / 3D problems (groundwater, biological imaging, porous media, etc.).

Although a great deal of research has been applied to scientific images, in most cases the resulting methods are not well documented in common textbooks, such that many experienced researchers will be unfamiliar with the use of the FFT method (Section 8.3) or of posterior sampling (Chapter 11), for example.

The goal, then, of this text is to address methods for solving multidimensional inverse problems. In particular, the text seeks to avoid the pitfall of being entirely mathematical / theoretical at one extreme, or primarily applied / algorithmic on the other, by deliberately developing the basic theory (Part I), the mathematical modelling (Part II), and the algorithmic / numerical methods (Part III) of solving a given problem.

Inverse Problems

So, to begin, why would we want to solve an inverse problem?

There are a great many spatial phenomena that a person might want to study ...

- The salinity of the ocean surface as a function of position;
- The temperature of the atmosphere as a function of position;
- The height of the grass growing in your back yard, as a function of location;
- The proportions of oil and water in an oil reservoir.

In each of these situations, you aren't just handed a map of the spatial process you wish to study, rather you have to *infer* such a map from given measurements. These measurements might be a simple function of the spatial process (such as measuring the height of the grass using a ruler) or might be complicated nonlinear functions (such as microwave spectra for inferring temperature).

The process by which measurements are generated from the spatial process is normally relatively straightforward, and is referred to as a *forward problem*. More difficult, then, is the *inverse problem*, discussed in detail in Chapter 2, which represents the mathematical inverting of the *forward problem*, allowing you to infer the process of interest from the measurements. A simple illustration is shown in Figure 1.1.

Large Multidimensional Problems

So why is it that we wish to study large multidimensional problems?

The solution to linear inverse problems (see Chapter 3) is easily formulated analytically, and even a nonlinear inverse problem can be reformulated as an optimization problem and solved. The challenge, then, is not the solving of inverse problems *in principle*, but rather actually solving them *in practice*.

For example, the solution to a linear inverse problem involves a matrix inversion. As the problem is made larger and larger, eventually the matrix becomes computationally or numerically impossible to invert. However, this is not just an abstract limit — even a modest two-dimensional problem at a resolution of 1000×1000 pixels contains one million unknowns, which would require the inversion of a one-million by one-million matrix: completely unfeasible.

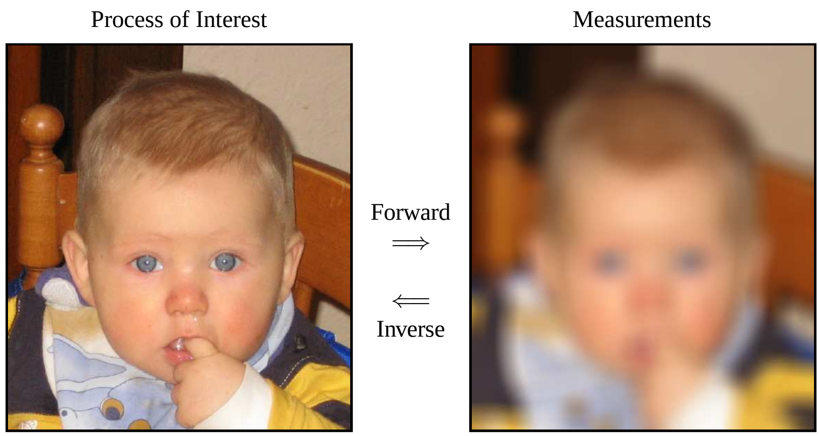


Fig. 1.1. An inverse problem: You want a nice clear photo of a face, however your camera yields blurry measurements. To solve this inverse problem requires us to mathematically invert the forward process of blurring.

Therefore even rather modestly sized two- and higher-dimensional problems become impossible to solve using straightforward techniques, yet these problems are very common. Problems having one million or more unknowns are littered throughout the fields of remote sensing, oceanography, medical imaging, and seismology, to name a few.

To be clear, a problem is considered to be multidimensional if it is a function of two or more independent variables. These variables could be spatial (as in a two-dimensional image or a three-dimensional volume), spatio-temporal (such as a video, a sequence of two-dimensional images over time), or a function of other variables under our control.

Multidimensional Methods versus Image Processing

What is it that the great diversity of algorithms in the image processing literature cannot solve?

The majority of images which are examined and processed in image processing are "real" images, pictures and scenes at human scales, where the images are not well described mathematically. Therefore the focus of image processing is on making relatively few explicit, mathematical assumptions about the image, and instead focusing on the development of algorithms that perform image-related tasks (such as compression, segmentation, edge detection, etc.).

Fig. 1.2. Which of these might be best characterized mathematically? Many natural phenomena, when viewed at an appropriate scale, have a behaviour which is sufficiently varied or irregular that it can be modelled via relatively simple equations, as opposed to a human face, which would need a rather complex model to be represented accurately.

In contrast, of great research interest are images taken at microscopic scales (cells in a Petri dish, the crystal structure of stone or metal) or at macroscopic scales (the temperature distribution of the ocean or of the atmosphere, satellite imagery of the earth) which do, in general, allow some sort of mathematical characterization, as explored in Figure 1.2. That is, the focus of this text is on the assumption or inference of rather *explicit* mathematical models of the unknown process.

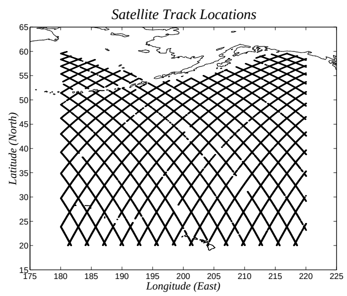
Next, in order to be able to say something about a problem, we need measurements of it. These measurements normally suffer from one of three issues, any one of which would preclude the use of standard image-processing techniques:

1. For measurements produced by a scientific instrument, acquiring a measurement normally requires time and/or money, therefore the number of measurements is constrained. Frequently this implies that the multidimensional problem of interest is only sparsely sampled, as illustrated in Figure 1.3.

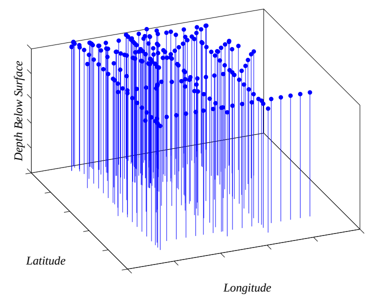
   There exist many standard methods to interpolate gaps in a sequence of data, however standard interpolation knows nothing about the underlying phenomenon being studied. That is, surely a grass-like texture should be interpolated differently from a map of ocean-surface temperature.

2. Most measurements are not exact, but suffer from some degree of noise. Ideally we would like to remove this noise, to infer a more precise version of the underlying multidimensional phenomenon.

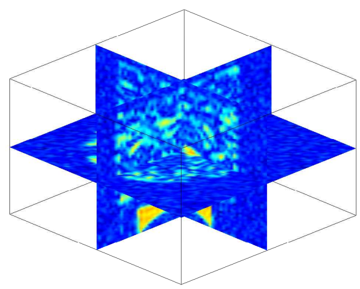
   There exist many algorithms for noise reduction in images, however these are necessarily heuristic, because they are designed to work on photographic images, which might contain images of faces / cars / trees and the like. Given a scientific dataset, surely we would wish to undertake denoising in a more systematic (ideally optimal) manner, somehow dependent on the behaviour of the underlying phenomenon.



Satellite Altimetry
(each point measures ocean height)



Ship-Based Measurements
(each point measures ocean temperature)



Slices from a 3D MRI
(each point represents the concentration of water in a block of concrete)

Fig. 1.3. Multidimensional measurements: Three examples of two- or three-dimensional measurements which could not be processed by conventional means of image processing. The altimetric measurements are sparse, following the orbital path of a satellite; the ship-based measurements are irregular and highly sparse, based on the paths that a ship followed in towing an instrument array; the MRI measurements are dense, but at poor resolution and with substantial noise.

3. In many cases of scientific imaging, the raw measurement produced by an instrument is *not* a direct measurement of the multidimensional field, but rather some function of it. For example, in Application 3 we wish to study atmospheric temperature based on radiometric measurements of microwave intensities: the air temperature and microwave intensity are indeed related, but are very different quantities.

   Standard methods in image processing normally assume that the measurements (possibly noisy, possibly blurred) form an image. However, having measurements being some complicated function of the field of interest (an inverse problem) is more subtle and requires a careful formulation.

Statistics and Random Fields

What is it that makes a problem statistical, and why do we choose to focus on statistical methods?

An interest in *spatial statistics* goes considerably beyond the modelling of phenomena which are inherently *random*. In particular, multidimensional random fields offer the following advantages:

1. Even if an underlying process is not random, in most cases measurements of the process are corrupted by noise, and therefore a statistical representation may be appropriate.

2. Many processes exhibit a degree of irregularity or complexity that would be extremely difficult to model deterministically. Two examples are shown in Figure 1.4; although there are physics which govern the behaviour of both of these examples (e.g., the Navier–Stokes differential equation for water flow) the models are typically highly complex, containing a great number of unknown parameters, and are computationally difficult to simulate.

   A random-fields approach, on the other hand, would implicitly approximate these complex models on the basis of observed statistics.

A random field[1] X is nothing but a large collection of random variables arranged on some set of points (possibly a two- or three-dimensional grid, perhaps on a sphere, or perhaps irregularly distributed in a high-dimensional space). The random field is characterized by the statistical interrelationships between its random variables.

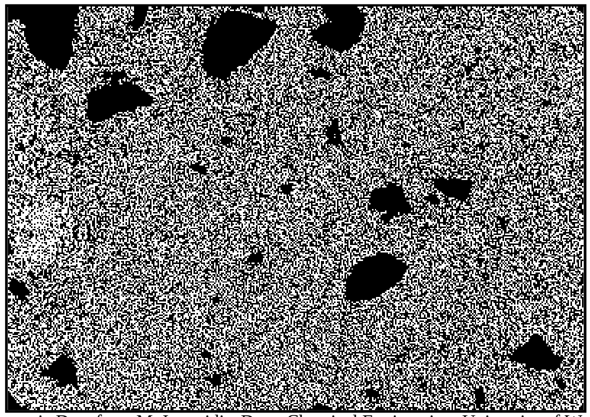
The main problem associated with a statistical formulation is the computational complexity of the resulting solution. However, as we shall see, there exists a comprehensive set of methods and algorithms for the manipulation and efficient solving of problems involving random fields. The development of this theory and of associated algorithms is the fundamental goal of this text.

Specifically, the key problem explored in this text is representational and computational efficiency in the solving of large problems. The question of efficiency is easily motivated: even a very modestly sized 256×256 image has 65 536 elements, and the glass beads image in Figure 1.4 contains in excess of 100 million elements! It comes as no surprise that a great part of the research into random fields involves the discovery or definition of *implicit* statistical forms which lead to effective or faithful representations of the true statistics, while admitting computationally efficient algorithms.

Broadly speaking there are four typical problems associated with random fields [112]:

[1] Random variables, random vectors, and random fields are reviewed in Appendix B.1.

A Porous Medium of Packed Glass Beads



(Microscopic Data from M. Ioannidis, Dept. Chemical Engineering, University of Waterloo)

Global Ocean Surface Temperature

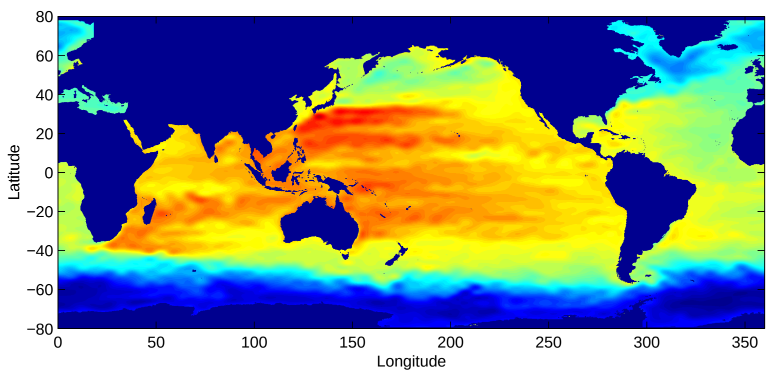


Fig. 1.4. Two examples of phenomena which may be modelled via random fields: packed glass beads (top), and the ocean surface temperature (bottom). Alternatives to random fields do exist to model these phenomena, such as ballistics methods for the glass beads, and coupled differential equations for the ocean, however such approaches would be greatly more complex than approximating the observed phenomena on the basis of inferred spatial statistics.

1. Representation: how is the random field represented and parametrized?
2. Synthesis: how can we generate “typical” realizations of the random field?
3. Parameter estimation: given a parametrized statistical model and sample image, how can we estimate the unknown parameters in the model?

4. Random fields estimation: given noisy observations of the random field, how can the unknown random field be estimated?

All four of these issues are of interest to us, and are developed throughout the text.

For each of these there are separate questions of formulation,

How do I write down the equations that need to be solved?

as opposed to those of solution,

How do I actually find a solution to these equations?

Part I of this text focuses mostly on the former question, establishing the mathematical fundamentals that are needed to express a solution, *in principle*. This gives us a solution which we might call

1. **Brute Force:** The direct implementation of the solution equations, irrespective of computational storage, complexity, and numerical robustness issues.

Parts II and III then examine the latter question, seeking practical, elegant, or indirect solutions to the problems of interest. However, *practical* should not be interpreted to mean that the material is only of dry interest to the specialist sitting at a computer, about to develop a computer program. Many of the most fundamental ideas expressed in this text are particularly in Part II, where deep insights into the nature of spatial random fields are explored.

A few kinds of efficient solutions, alternatives to the direct implementations from Part I, are summarized as follows:

2. **Dimensionality Reduction:** Transforming a problem into one or more lower-dimensional problems.
3. **Change of Basis:** A mathematical transformation of the problem which simplifies its computational or numerical complexity.
4. **Approximate Solution:** An approximation to the exact analytical solution.
5. **Approximated Problem:** Rather than solving the given problem, identifying a similar problem which can be solved exactly.
6. **Special Cases:** Circumstances in which the statistics or symmetry of the problem gives rise to special, efficient solutions.

These six points give a broad sense of what this text is about.

Interpolation as a Multidimensional Statistical Problem

We conclude the Introduction by developing a simple canonical example, to which we frequently refer throughout the text. We have chosen this problem because it is intuitive and simple to understand, yet possesses most of the features of a large, challenging, estimation problem.

Suppose you had sparse, three-dimensional measurements of some scalar quantity, such as the temperature throughout some part of an ocean. You wish to produce a dense map (really, a volume) of the temperature, based on the observed measurements.

Essentially this is an interpolation problem, in that we wish to take sparse measurements of temperature, and infer from them a dense grid of temperature values. However by *interpolation* we do not mean standard deterministic approaches such as linear, bilinear, or B-spline interpolation, in which a given set of points is deterministically projected onto a finer grid. Rather, we mean the *statistical* problem, in which we have a three-dimensional random field Z with associated measurements M, where the measurements are subject to noise V, such that

$$m_i = z_{j_i} + v_i, \tag{1.1}$$

where j_i is an index, describing the location of the ith measurement. Thus (1.1) gives the forward model, which we wish to invert (Chapter 2).

Given the definition of the inverse problem, we can formulate the analytical solution, depending on whether this is a static problem, a single snapshot in time (Chapter 3), or a more complicated time-dynamic problem, in which the temperature evolves and is estimated over time (Chapter 4).

However, so far we haven't said anything about the mathematics or statistics governing Z. What distinguishes statistical interpolation from deterministic methods, such as linear or bilinear interpolation, is the ability to take into account specific properties of Z (Chapter 5). Thus is Z smooth, on what length scales does it exhibit variability, and what happens at its boundaries? Furthermore, are the statistics of Z spatially stationary (not varying from one location to another) or not, and are the statistics best characterized by looking at correlations of Z or at the inverse correlations (Chapter 6)? Finally are there hidden underlying aspects to the problem, such that the model in one location may be different from that in another (Chapter 7)?

If the problem is particularly large, would it be possible to collapse it along one dimension, or possibly to solve the problem in pieces, rather than as a whole? One could also imagine transforming the problem, for example using a Fourier or wavelet transform (Chapter 8).

At this point we have determined what sort of problem we have, whether reduced in dimensionality, whether transformed, whether stationary. We are left with two basic approaches for solving the inverse problem: we can convert the inverse problem to a linear system, and use one of a number of linear systems solvers (mostly iterative) to find the desired map (Chapter 9), or we could use a domain-decomposition approach that tackles the problem row-by-row, column-by-column, block-by-block, or scale-by-scale (Chapter 10). We may also wish to understand the model better by generating random samples from it (Chapter 11).

How to Read This Text

The preceding interpolation example has been very short and many details are omitted, but it is hoped that it gives the reader a sense of the scope of the ideas developed in this text. The reader wishing to follow up on interpolation in more detail is encouraged to move directly to the three interpolation examples developed in Chapter 2 on pages 20, 32, and 36.

Those readers unfamiliar with the contents of this text may wish to survey the book by glancing through the worked applications at the end of every chapter, which cover a variety of topics in remote sensing and scientific imaging. These applications, and also the various examples throughout the text, are all listed beginning on page XIII.

Any reader who wishes to explore multidimensional random fields and processes in some depth should focus on the chapters on inverse problems and modelling, Chapters 2, 5, 6, and 8, which form the core of this text.

Readers who are interested in numerical implementations of the methods in this text should consult the list of MATLAB[2] code samples on page XV. The code samples are cross-referenced to figures and examples throughout the text, and all of the listed samples are available online at

http://ocho.uwaterloo.ca/book

[2] MATLAB® is a registered trademark of The MarhWorks Inc.