

# Contents

<b>1</b>	<b>Continuous Bayesian Tissue Classification for Visualization</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	A Framework for Classification Solutions . . . . .	8
1.3	A Family of Solutions . . . . .	12
1.4	Results . . . . .	17
1.5	Discussion . . . . .	20
1.6	Conclusions . . . . .	22
1.7	Acknowledgements . . . . .	23
1.8	References . . . . .	23





# List of Tables

---



# List of Figures

1.1	Scientific Visualization cycle. . . . .	4
1.2	The computational framework that we use for creating static and dynamic geometric models of biological systems from MRI data. Our focus is on the second stage, tissue classification. . . . .	5
1.3	Discrete, single-material classification of the same human brain slice shown in Figure 1.12 (See Plate ??). . . . .	6
1.4	We define a sample as a scalar or vector valued element of a multi-dimensional dataset. A voxel is the region surrounding a sample. . . . .	7
1.5	Histogram calculation. The scalar samples in (i) are measurements from a dataset containing two materials, $A$ and $B$ . One material has measurement values near $v_A$ and the other near $v_B$ . In (ii) we reconstruct a continuous function from those samples, with relatively flat regions at the two material values. These flat regions correspond to the Gaussian-shaped peaks centered around $v_A$ and $v_B$ in the histogram, (iii), which is shown on its side to emphasize the axis that it shares with the spatial data. . . . .	9
1.6	Benefits of histograms of vector-valued data. We show histograms of an object with three materials. (i) is a histogram of scalar data and shows that material mean values are collinear; therefore, distinguishing among more than two materials is often ambiguous. (ii) represents a histogram of vector-valued data, with one MRI value along the axes at the bottom of the figure and one along the left side. Brighter points represent larger values of the histogram. (iii) is another representation of the same histogram. The histogram shows that mean values often move away from collinearity in higher dimensions. (iv) shows a different histogram demonstrating that the collinearity problem can occur with vector-valued data. . . . .	10

---

1.7	The three classification algorithms start from the assumption that in a real-world object each point is exactly one material, as in (i). The measurement process creates samples that mix materials together, from which we reconstruct a continuous, band-limited measurement function $\rho(x)$ . For some values of $x$ , e.g., $P_1$ and $P_2$ , $\rho(x)$ returns the signature of a pure material. For other values of $x$ , e.g., $P_3$ , $\rho(x)$ returns a combination of the pure material signatures. Algorithm A, (ii), models the measurements as a linear combination of pure material signatures and signatures of mixtures of pairs of materials. Two material mixing occurs in the region labeled “A & B,” and each voxel (delimited by grid lines) can contain some of each signature. Algorithms B and C, (iii), use a parametric model that is based on the distance from a voxel to a boundary between materials. The grid lines surround voxels. . . . .	14
1.8	Dataset parameters for a pure material histogram basis function, shown in (i) include $c$ , the mean value for the material, and $s$ , which measures the standard deviation of measurements (see Equation 1.6). (ii) shows corresponding parameters for a two-material mixture basis function. . . . .	15
1.9	Histogram basis functions, $f_{\text{boundary}}(v)$ , for scalar data. The shapes approach normal distributions as $d$ , the distance from the boundary to a voxel center, moves away from 0. . . . .	16
1.10	Histogram basis functions, $f_{\text{boundary}}(v)$ , for vector-valued data. Three different values of $d$ , the distance from the boundary to a voxel center, are represented. . . . .	16
1.11	The partial volume, boundary distance, and discrete algorithms are compared in classifying simulated MRI data. (ii) shows the simulated data, which contains three different materials. The geometry that the data measures is shown in (i). (iii) shows what an “ideal” classification algorithm should produce and (iv)-(vi) show results from different algorithms. Note that the new algorithms (v) and (vi) produce results most similar to the ideal case, and that (vi) does so even with scalar data. . . . .	18
1.12	One slice of data from a human brain. (i) shows the original two-valued data, (ii) shows four of the identified materials, white matter, gray matter, cerebrospinal fluid, and muscle, separated out into different images, and (iii) shows the results of the new classification mapped to different colors (see Plate ??). Note the smooth boundaries where materials meet and the much lower incidence of misclassified samples than in Figure 1.3. . . . .	19
1.13	A volume-rendering image of a human hand dataset. The opacity of different materials is decreased above cutting planes to show details of the classification process within the hand (see Plate ??). . . . .	20
1.14	A geometric model of tooth dentine and enamel created by collecting MRI data using a technique that images hard solid materials [GLF <sup>+</sup> 95] and classifying dentine and enamel in the volume data with our new partial-volume mixtures algorithm. Polygonal isosurfaces define the bounding surfaces of the dentine and enamel. The enamel-dentine boundary, shown in the lower images, is difficult to examine non-invasively using any other technique. . . . .	21

# Preface

Data visualization concerns the manipulation of sampled and computed data for comprehensive display. The goal of the visualization is to bring to the user a deeper understanding of the data as well as the underlying physical laws and properties. Such visualization may be used to enlighten a physicist on the complex interaction between electrons, to guide the medical practitioner in a surgery situation, or simply to view the surface of a planet which has never been seen by human eyes.

Through the presentation of massive amounts of data as images, we allow the visualization user to rapidly prune useless information, focus on necessary information, and comprehend the science behind the data. Interaction with data brings another level of understanding. Static images can be misleading and mask important features of the data. Motion in visualization brings out hidden features which are inherently dynamic. Interactive manipulation and control of visualization is an important tool which allows scientists to more quickly focus on the region of interest. In environments which are immersive, the motion is critical, to the point that delays or inconsistencies can make the viewer ill. In this case there is a desire to bound response time using time-critical techniques.

The important aspects of data visualization can be broken down into three categories:

*Computation* - the ability to speedily compute a visualization. This may include computing a polygonal approximation to an isosurface of a scalar function, or the computation of a particle trace through a time-dependent vector field, or any action which requires extracting an abstract object or representation from the data being examined.

*Display* - the ability to quickly display the computed visualization. Display encompasses both computed visualizations as listed above, as well as direct display methods such as volume visualization and ray tracing.

*Querying* - the ability to interactively probe a displayed visualization for the purpose of further understanding on a fine scale what is being displayed on a coarser scale.

The different chapters of this book address fundamental algorithmic techniques and data structures for efficient computation, display and querying of scalar, vector and tensor field data.



# 1

## Continuous Bayesian Tissue Classification for Visualization

David H. Laidlaw

*California Institute of Technology*

### ABSTRACT

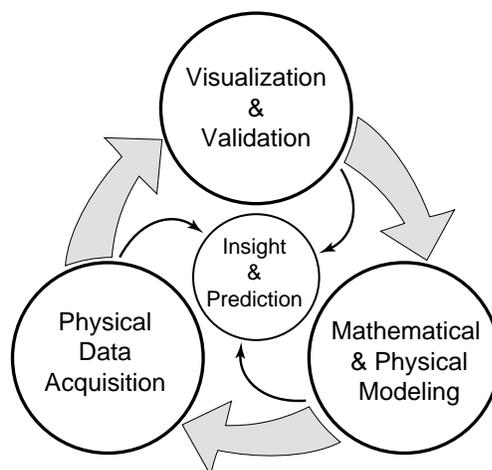
This chapter discusses tissue classification—the identification of regions of materials in volume data. Our primary motivation is creating images for understanding biological systems from medical imaging data. We advocate an approach to this problem that avoids some of the pitfalls of approaches derived directly from discrete pattern recognition. By combining sampling theory and Bayesian probability theory, images can be generated without many of the aliasing artifacts otherwise introduced.

We first discuss the classification problem, its place in the visualization cycle, and some approaches to it. We then outline the approach we advocate. It involves reconstructing a continuous function from samples and examining the function over small, voxel-sized regions rather than at single points. The behavior of the continuous function over those regions provides much richer information and is the key to improving classification results. We discuss three algorithms and a framework for developing additional ones. Results of applying the algorithms to simulated and real MRI data show their efficacy.

### 1.1 Introduction

Visualization is inherently an iterative process that must be coupled with scientific questions to be meaningful. Figure 1.1 shows the cycle, starting with a hypothetical model for some physical phenomenon, continuing with observation and collection of data, and finishing (one iteration) with visualization and analysis of the data. The ultimate goal of the process is to converge on insight and understanding of the phenomenon under study.

Geometry is at the heart of visualization, and extracting geometry from scientific mea-



**Figure 1.1** Scientific Visualization cycle.

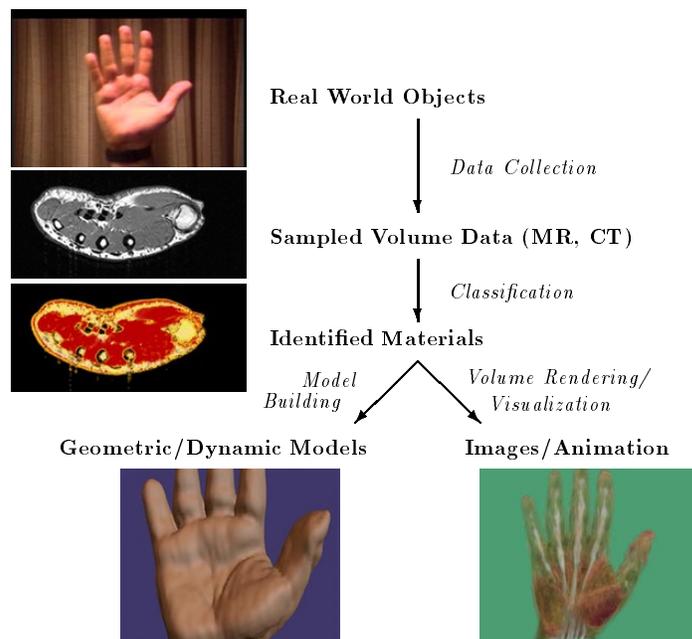
surements is a central part of scientific visualization. The goal of our visualization process is the understanding of the development and anatomy of biological systems through the study of their geometry. We describe here one stage in a three-stage computational framework for extracting geometry from sampled volume data. Within the framework, shown in Figure 1.2, we measure the biological systems, identify different regions within them, and create images and models using information about the regions and the measurement process.

Tissue classification or segmentation, the second step of the framework, is the focus of this chapter. We will describe the how the search for geometry impacts the classification process, the algorithmic changes it leads to, a framework for creating new algorithms that work well for extracting geometry, and some examples of new algorithms.

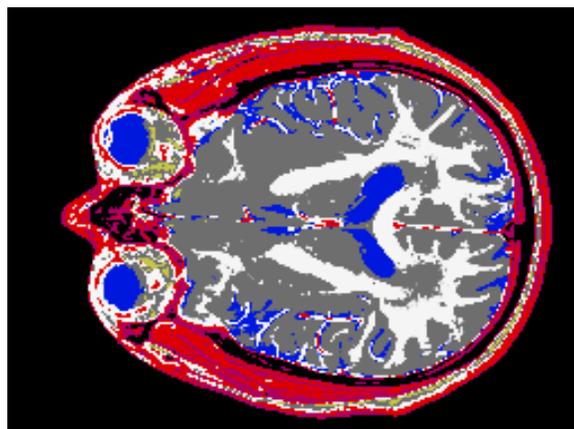
These techniques convert unclassified volume images into volume images with a physical interpretation that is more appropriate for many visualization algorithms. Resulting images may measure, for example, the local density of a material or the distance from a boundary between materials. The physical interpretation of these volume images is appropriate for volume-rendering [Lev88] (see Chapter ??), extraction of surface models [LC87] (see Chapter ??), and extraction of volume models. With these images and models we hope to be able to address anatomical and developmental questions through examination of systems and through predictive modeling of complex biological shapes and behavior. Applications also exist for surgical planning and assistance, conventional computer animation, and other imaging modalities.

### 1.1.1 Related Work

Many researchers have worked on identifying the locations of materials in sampled datasets [VBJ<sup>+</sup>85], [VSR88], [CLKJ90], [DH73]. [CVC<sup>+</sup>95] gives an extensive review of the segmentation of MRI data. However, many of these algorithms generate artifacts like those shown in Figure 1.3, an example of data classified with a technique based on discrete sample values. These techniques work well in regions where a voxel contains only a single material, but tend



**Figure 1.2** The computational framework that we use for creating static and dynamic geometric models of biological systems from MRI data. Our focus is on the second stage, tissue classification.

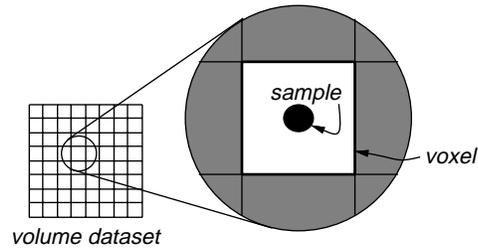


**Figure 1.3** Discrete, single-material classification of the same human brain slice shown in Figure 1.12 (See Plate ??).

to break down at boundaries between materials. This introduces both stair-step artifacts, as shown between gray matter and white matter within the brain, and thin layers of misclassified voxels, as shown by the white matter between the skull and the skin. Both types of artifacts can be ascribed to the partial-volume effects ignored by the segmentation algorithms.

A number of algorithms address this limitation of discrete classification techniques. [DCH88] demonstrates that accounting for mixtures of materials within a voxel can reduce these artifacts, and approximates the relative volume of each material represented by a sample as the probability that the sample is that material. Their technique works well for differentiating air, soft tissue, and bone in CT data, but not for differentiating materials in MR data, where the measured data value for one material may often be identical to the measured value for a mixture of two other materials. [WAAR<sup>+</sup>88] and [KSW96] avoid partial-volume artifacts by taking linear combinations of components of vector measurements. An advantage of their techniques is that the linear operations they perform preserve the partial-volume mixtures within each sample value, and so partial-volume artifacts are not created. A disadvantage is that the linear operations are not as powerful as non-linear operations, and so either more data must be acquired or classification results may not be as accurate. [CHK91] and [NFMD90] address the partial-volume issue by identifying combinations of materials for each sample value. As with many other approaches to identifying mixtures, these techniques use only a single measurement taken within a voxel to represent its contents. Without the additional information available within each voxel region, these classification algorithms are limited in their accuracy. [SG93] shares a mixture distribution for histograms with our technique. Their technique, however, estimates material amounts in an entire dataset, and does not classify the data at a voxel level.

[WCW88] presents an interesting approach to partial-volume imaging that makes assumptions similar to ours about the underlying geometry being measured and about the measurement process. The results of their algorithm are a discrete material assignment for each sub-voxel of the dataset. Taken collectively, these multiple sub-voxel results provide a measure of the mixtures of materials within a voxel but arrive at it in a very different manner than does our algorithm. This work has been applied to satellite imaging data, and so results are difficult to compare, but aspects may combine well with our technique.



**Figure 1.4** We define a sample as a scalar or vector valued element of a multi-dimensional dataset. A voxel is the region surrounding a sample.

### 1.1.2 Continuous Classification

We have developed a methodology for constructing a probabilistic Bayesian classification algorithm from a set of assumptions about the underlying data. Our algorithms start with the premise that the sampled datasets satisfy the sampling theorem, which allows us to reconstruct a continuous function  $\rho(x)$  over the entire dataset [OWY83]. We treat a voxel as a volume of space (see Figure 1.4). We examine all the values that the reconstructed function takes on over a voxel by calculating histograms of  $\rho(x)$  over small regions of the dataset and classify those histograms by fitting histogram basis functions constructed from the set of assumptions. This approach builds upon statistical classification and non-linear digital image and signal processing.

Using the Bayesian framework, we have constructed three different classification algorithms, described in more detail below. The first algorithm models the contents of each voxel as a linear combination of pure materials and mixtures of two materials. The second algorithm models each voxel as either entirely composed of a single pure material, or composed of two materials with a separating boundary. The third algorithm is substantially similar to the second, but allows the expected value, or signature, of each material to vary over a dataset, a common characteristic of MRI data. These techniques classify MRI data better than previously available techniques because they use a more accurate model of the collected data. They are also tailored to produce accurate results near boundaries between materials where extracted geometric details are most visible.

### 1.1.3 Organization of Chapter

In Section 1.2 we describe a Bayesian framework for developing new classification algorithms. As we present each step in the framework, we work through a simple example to illustrate the process. Section 1.3 gives an overview of the family of algorithms that we have developed using the framework. They are described in more detail in [Lai95], [LBJ97], [GLF<sup>+</sup>95], and [LFB97].

We show some results of visualizing volume data classified with our techniques in Section 1.4. Our visualizations take two basic forms, surface rendering and volume rendering. Most traditional computer graphics imagery is rendered as surfaces, although in the last decade volume rendering has emerged as a useful adjunct to the more traditional techniques. Unlike surface-rendering methods, volume rendering produces images that can show internal structure. Thus, images of solid objects can appear to consist of volumes of transparent or semi-transparent material (see Chapter ??).

## 1.2 A Framework for Classification Solutions

We define a statistical framework for creating classification algorithms. It is based on Bayesian probability theory [Lor89] and approximations of conditional and prior probabilities. Within the framework we have created a family of new algorithms that calculate the probability of a particular combination of materials given the histogram over a small region. We then find the most likely combination for the region.

In this section we outline how to construct a new classification algorithm within our framework, illustrating the process with an existing algorithm. We begin by defining terms.

### 1.2.1 Definitions

We refer to the coordinate system of the space of the object we are measuring as *spatial coordinates* and generally use  $x \in X$  to refer to points.  $X$  is  $n_x$ -dimensional, where  $n_x$  is 3 for volume data, but can be 2 for slices.

Each measurement, or *sample* (see Figure 1.4), may be a scalar or vector and lies in *feature space* (see Figure 1.5), with points frequently denoted as  $v \in V$ . Feature space is  $n_v$ -dimensional, where  $n_v$  is one for scalar-valued data, two for two-element vector data, etc.

From the samples we reconstruct a continuous function  $\rho(x)$  over  $X$  by interpolating sample values. We use tri-cubic interpolation and so incorporate information from 64 nearby samples into each interpolated measurement. A *voxel*, or *voxel region*, (see Figure 1.4) is the volume surrounding a sample. The terms are interchangeable. We use voxels that tile the volume of a dataset, but overlapping or non-adjacent voxels are also possible. We are frequently interested in the behavior of  $\rho(x)$  over the region defined by the volume of a voxel.

Classification algorithms classify a voxel based on information derived from the raw data in or near the voxel. We refer to the information as *voxel-info*, and label it  $h$ . For each of our classification techniques we use a histogram over the small region defined by a voxel to encode the information contained in the voxel. We first reconstruct a continuous function over the entire dataset from the samples and then use the continuous analog of a discrete histogram,

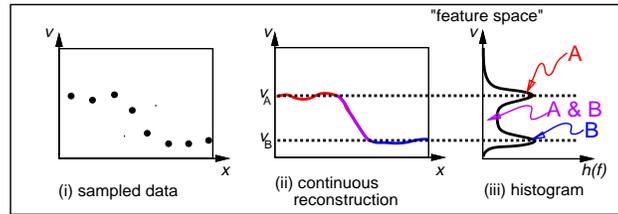
$$h^{\mathcal{R}}(v) = \int \mathcal{R}(x)\delta(\rho(x) - v)dx \quad (1.1)$$

to calculate a histogram over each voxel.  $\mathcal{R}(x)$  is non-zero within the region of interest, and integrates to 1. We define  $\mathcal{R}(x)$  to be constant in the region of interest making every spatial point contribute equally to the histogram  $h^{\mathcal{R}}(v)$ . Note also that  $h^{\mathcal{R}}(v)$  integrates to 1.  $\delta$  is the Dirac-delta function. Figure 1.5 shows an example of calculating a normalized histogram from a continuous function.

We use Equation 1.1 both as a starting point for deriving histogram basis functions and also for calculating histograms of regions of our datasets.

### 1.2.2 Construction of a New Bayesian Classification Algorithm

Construction of a new classification algorithm involves four steps: choosing voxel-info to represent the information in a voxel, selecting a set of assumptions about the collection process, defining a parameterized model of the voxel-info, and deriving material probability estimates.



**Figure 1.5** Histogram calculation. The scalar samples in (i) are measurements from a dataset containing two materials,  $A$  and  $B$ . One material has measurement values near  $v_A$  and the other near  $v_B$ . In (ii) we reconstruct a continuous function from those samples, with relatively flat regions at the two material values. These flat regions correspond to the Gaussian-shaped peaks centered around  $v_A$  and  $v_B$  in the histogram, (iii), which is shown on its side to emphasize the axis that it shares with the spatial data.

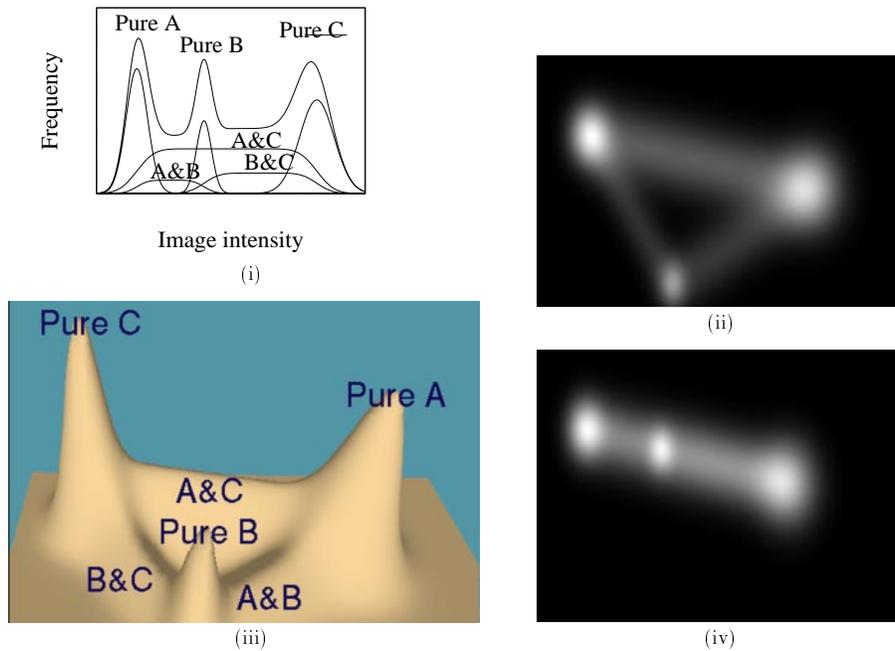
*Choose voxel-info.* Our algorithms use histograms calculated over the region of a voxel as voxel-info; other choices are possible. We have chosen histograms for a number of reasons. First, they generalize single measurements to measurements over a region, so classification concepts that apply to single measurements generalize. Second, the histograms can be calculated easily. Third, the histograms capture information about neighboring voxels, which increases the information content of the voxel-info and improves the classification results. Fourth, histograms are orientation independent; orientation independence reduces the number of parameters in the classification process hence simplifying and accelerating it.

As with many other techniques, ours works on vector-valued volume data, in which each material has a characteristic vector value rather than a characteristic scalar value. Vector-valued datasets have a number of advantages and generally give better classification results. First, they have an improved signal-to-noise ratio. Second, they frequently distinguish similar materials more effectively (see Figure 1.6). The jump from scalar to two-valued vector data is the most significant. In scalar-valued datasets it is difficult to distinguish a mixture of two pure materials with values  $v_A$  and  $v_B$  from a pure material with some intermediate value such as  $v_C = (v_A + v_B)/2$ . This is because all three material values are collinear, as they must be for such a dataset. With more measurement dimensions in the dataset, collinearity is less frequent for most combinations of three or more materials, although Figure 1.6(iv) illustrates that it can still occur. When it does occur, classification works as for scalar-valued data.

*Codify collection assumptions.* In the second step we codify a set of assumptions about the data-collection process. The assumptions embody information about:

- how sampling works on the machine we are using,
- the responses of materials or combinations of materials to the measurement process,
- the spatial uniformity of the measurements, and
- geometric characteristics of our objects.

For our example we will assume that there is a known discrete set of materials, that measurements for a single material are distributed normally, and that each voxel consists of exactly one material. Section 1.3 lists the assumptions for our algorithms, some of which are illustrated in Figure 1.7.



**Figure 1.6** Benefits of histograms of vector-valued data. We show histograms of an object with three materials. (i) is a histogram of scalar data and shows that material mean values are collinear; therefore, distinguishing among more than two materials is often ambiguous. (ii) represents a histogram of vector-valued data, with one MRI value along the axes at the bottom of the figure and one along the left side. Brighter points represent larger values of the histogram. (iii) is another representation of the same histogram. The histogram shows that mean values often move away from collinearity in higher dimensions. (iv) shows a different histogram demonstrating that the collinearity problem can occur with vector-valued data.

*Model voxel-info.* From our choice of voxel-info and the set of assumptions about the data-collection process, we define a parameterized model of the voxel-info,  $f(\alpha)$ . The parameters,  $\alpha$ , for the voxel-info model are divided into two classes. The first, *dataset parameters*, consists of those that are known before the voxel classification process. The second, *voxel parameters*, vary from voxel to voxel.

For our example the dataset parameters are a mean and deviation for measurements of each materials; the single voxel parameter is discrete-valued and each value corresponds to one of the possible materials.

*Estimate material probabilities.* Given voxel-info,  $h$ , which encodes information from a single voxel and a parameterized model of the voxel-info,  $f(\alpha)$ , we want to find the most likely set of parameters  $\alpha$ . The *posterior probability* defines how likely a set of parameters  $\alpha$  is given an observed voxel-info,  $h$ :

$$P(\alpha|h) \tag{1.2}$$

By maximizing the posterior probability we find the most likely set of parameters. Equation 1.2 cannot, in general, be calculated directly, so we use Bayes' Theorem to decompose it into pieces that we can either calculate directly or estimate.

$$P(\alpha|h) = \frac{P(\alpha)P(h|\alpha)}{P(h)} \tag{1.3}$$

$P(h|\alpha)$  is the *likelihood* of a particular instance of voxel-info for a given set of voxel parameters. We calculate it by comparing the parameterized model of the voxel-info to the actual voxel-info and quantifying the difference.

$P(\alpha)$  is the *prior probability* and tells us how likely each set of parameters is. We estimate the prior probability from the model of the voxel-info and from the assumptions that we make about the data-collection process.

$P(h)$  is the *global likelihood* of a particular instance of voxel-info. We assume that it is a constant function of  $h$ . It becomes a normalization factor for Equation 1.3.

### 1.2.3 Classification

*Estimate dataset parameters.* Any dataset parameters must be estimated before the classifier can be used. We estimate them by calculating their values for a training set of voxels with known voxel parameters. In our example we would calculate the mean and variance of a set of measurements known to be from each discrete material.

*Classify voxels.* Finally, we calculate the voxel-info,  $h$ , for each voxel and use the classifier to estimate the voxel parameters.

### 1.2.4 Example of Classification Algorithm Construction

In this section we construct the Bayesian classifier for the example we introduced in Section 1.2.2. This classifier is not new [DH73], but its construction within our framework illustrates how to create a classifier.

*Example voxel-info.* For our example we define voxel-info  $h_\epsilon$  as the single data measurement at the center of a voxel.

*Example assumptions.*

- $e_1$ : Each measurement comes from exactly one material.
- $e_2$ : The measurements from each material are normally distributed.
- $e_3$ : We know the number of materials and can identify samples from each material within the data.
- $e_4$ : All materials are equally likely.

*Example model of voxel-info.* Our model of the voxel data,  $f_\epsilon(\alpha_\epsilon)$ , has a single discrete voxel parameter,  $\alpha_\epsilon$ , that specifies the material within the voxel.

$$f_\epsilon(\alpha_\epsilon) = \mu_{\alpha_\epsilon} \quad (1.4)$$

For each material  $i$ , our model has two dataset parameters,  $\mu_i$  and  $\sigma_i$ , defining the expected value and the standard deviation of measurements.

*Example material probabilities.* From assumptions  $e_1$  and  $e_2$ , the likelihood,  $P(h_\epsilon|\alpha_\epsilon)$ , can be calculated by evaluating a normal distribution with mean  $\mu_{\alpha_\epsilon}$  and variance  $\sigma_{\alpha_\epsilon}^2$ :

$$P_\epsilon(h_\epsilon|\alpha_\epsilon) = \frac{1}{\sigma_{\alpha_\epsilon}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\alpha_\epsilon - \mu_{\alpha_\epsilon}}{\sigma_{\alpha_\epsilon}}\right)^2\right) \quad (1.5)$$

From assumption  $e_4$ , the prior probability,  $P_\epsilon(\alpha_\epsilon)$ , is  $\frac{1}{n_m}$  where  $n_m$  is the number of materials.

*Example dataset parameter estimation.* The dataset parameters consist of the mean and variance of measurements of each discrete material. From assumption  $e_3$  we find the dataset parameters by interactively selecting a set of points in the dataset for each material. We define each set as measurements of the single material they represent; from them we calculate the mean and variance for each material.

*Example classification.* We iterate over each voxel calculating the most likely value for the single voxel parameter,  $\alpha_\epsilon$ ; for each voxel we measure  $h_\epsilon$ , the value at the center of the voxel. For each possible material we calculate the corresponding posterior probability,  $P(\alpha_\epsilon|h_\epsilon)$ , and choose the largest of these values. This gives us the most likely material.

### 1.3 A Family of Solutions

In this section we give an overview of three new classification algorithms constructed within our framework and compare and contrast them with one another. We first list the assumptions that are common to all three algorithms and the dataset parameters that these assumptions imply. We then present the assumptions unique to each algorithm and summarize both the dataset and voxel parameters.

### 1.3.1 Assumptions Common to New Algorithms

We make several assumptions that are consistent among the new algorithms that we have developed. Each algorithm also makes additional assumptions detailed in subsequent sections.

$e_{c1}$ : **Discrete materials.** The first assumption is that materials within the objects that we measure are discrete at the resolution that we are sampling. Material boundaries are not assumed to be aligned with the sampling grid. We make this assumption because we are generally looking for boundaries between materials, and because we are starting from sampled data, which loses information about detail that is finer than the sampling rate.

This assumption does not preclude homogeneous combinations of sub-materials that can be treated as a single material at our sampling resolution. For example, muscle may contain some water, and yet be treated as a separate material from water. This assumption is not satisfied where materials gradually transition from one to another over many samples or are not relatively uniformly mixed. Section 1.5 discusses cases where this assumption is not satisfied.

$e_{c2}$ : **Normally distributed noise.** The second assumption is that noise is added to each discrete sample and that the noise is normally distributed. We assume a different variance in the noise for each material. This assumption is not strictly satisfied for MRI data in some cases, but seems to be satisfied sufficiently to classify data well.

$e_{c3}$ : **Data satisfies sampling theorem.** The third assumption we make is that the sampled datasets we classify satisfy the sampling theorem [OWY83]. The sampling theorem states that if we sample a sufficiently band-limited function, we can exactly reconstruct that function from the samples.

From assumption  $e_{c1}$  the underlying physical object has discontinuous boundaries between materials, and an infinite-precision MRI machine would generate a dataset with discontinuities at material boundaries. At finite resolutions, the measurement function must be band limited so that it can be reconstructed from the samples.

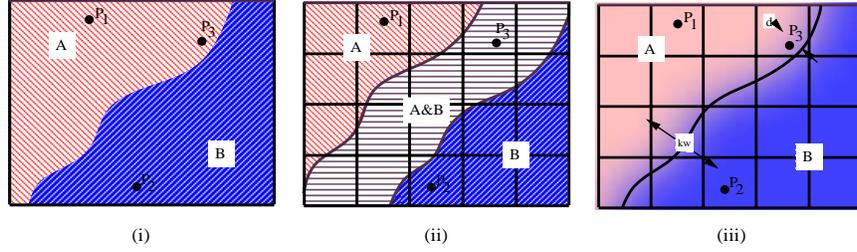
MRI slice data generally satisfies this assumption or can be pre-processed to satisfy it [Lai92], [GLS<sup>+</sup>97] within the slice. Without data that satisfy the sampling theorem, we cannot reconstruct a continuous function, and without a continuous function, we cannot extract geometric models as shown in Section 1.4.

### 1.3.2 Overview of Algorithm A: Partial Volume Mixtures

Our new partial volume mixtures algorithm was developed to create classified data with fewer boundary artifacts so that we could produce better geometric models. The choice of voxel-info, the model of the voxel-info, and some of the assumptions are formulated to capture and identify information about the boundaries. The remainder of the assumptions help make some of the probability calculations more tractable.

We assume, as in Figure 1.7(i), that each voxel is a mixture of materials, with mixtures like “A & B” occurring where the band-limiting effects of the data collection process blur pure materials together, as in Figure 1.7(ii). From this assumption we derive basis functions that model histograms for pure materials and for mixtures of two materials.

*Additions to common assumptions.*



**Figure 1.7** The three classification algorithms start from the assumption that in a real-world object each point is exactly one material, as in (i). The measurement process creates samples that mix materials together, from which we reconstruct a continuous, band-limited measurement function  $\rho(x)$ . For some values of  $x$ , e.g.,  $P_1$  and  $P_2$ ,  $\rho(x)$  returns the signature of a pure material. For other values of  $x$ , e.g.,  $P_3$ ,  $\rho(x)$  returns a combination of the pure material signatures. Algorithm A, (ii), models the measurements as a linear combination of pure material signatures and signatures of mixtures of pairs of materials. Two material mixing occurs in the region labeled “A & B,” and each voxel (delimited by grid lines) can contain some of each signature. Algorithms B and C, (iii), use a parametric model that is based on the distance from a voxel to a boundary between materials. The grid lines surround voxels.

$e_{m4}$ : **Linear mixtures.** Each voxel measurement is a linear combination of pure material measurements and measurements of their pair-wise mixtures.

$e_{m5}$ : **Uniform tissue measurements.** Measurements for the same material have the same expected value throughout a dataset.

$e_{m6}$ : **Box filtering.** The spatial measurement process can be approximated by a box filter for the purpose of calculating histogram basis functions.

$e_{m7}$ : **Materials identifiable in histogram of entire dataset.** The signatures for each material and mixture must be identifiable in a histogram of the entire dataset.

*Description.* The parameters for each voxel in this algorithm are density values for each pure material and for each pair-wise combination of materials and an estimate of the low-frequency noise within the voxel. The densities sum to one, and each density weights a histogram basis function for either a pure material or a mixture. The basis function for pure materials is a normal distribution.

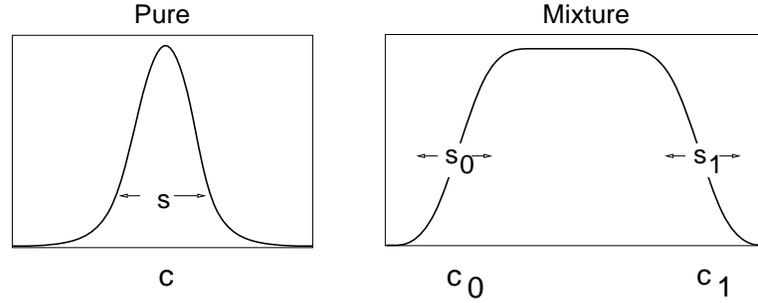
$$f_{\text{single}}(v; c, s) = \prod_{i=1}^{n_v} \frac{1}{s_i \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{v_i - c_i}{s_i}\right)^2\right) \quad (1.6)$$

The basis function for a mixture is

$$f_{\text{double}}(v; c, s) = \int_0^1 f_{\text{single}}(v; (1-t)c_1 + tc_2, s) dt. \quad (1.7)$$

Both basis functions are shown in Figure 1.8.

The dataset parameters are the mean,  $c$ , and standard deviation,  $s$ , for each pure material, as well as an expected deviation of the model histogram from actual histograms. The parameters are estimated by analyzing a histogram taken over the entire dataset and fitting a combination of materials to that histogram.



**Figure 1.8** Dataset parameters for a pure material histogram basis function, shown in (i) include  $c$ , the mean value for the material, and  $s$ , which measures the standard deviation of measurements (see Equation 1.6). (ii) shows corresponding parameters for a two-material mixture basis function.

### 1.3.3 Overview of Algorithm B: Boundary Distance

The boundary distance algorithm addresses some of the limitations we discovered in the partial volume mixtures algorithm. The voxel-info and most of the assumptions are the same, but the histogram basis functions differ. The main change is that the distance from a boundary is explicitly incorporated into the histogram basis function for mixtures, as shown in Figure 1.7(iii). The explicit model better fits histograms of voxels near boundaries. A second change is that the histogram basis functions are derived with the more-accurate assumption of Gaussian filtering.

*Additions to common assumptions.*

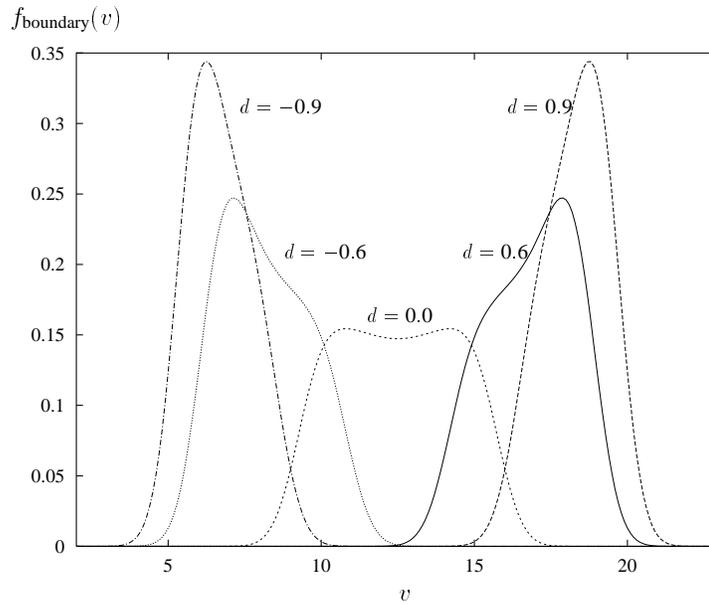
- $\epsilon_{b4}$ : **Only pair-wise mixtures.** Each voxel measurement is of either a pure material or a mixture of exactly two materials near a boundary.
- $\epsilon_{b5}$ : **Uniform tissue measurements.** Measurements for the same material have the same expected value throughout a dataset.
- $\epsilon_{b6}$ : **Gaussian filtering.** The measurement process can be approximated by a Gaussian filter for the purpose of calculating histogram basis functions.
- $\epsilon_{b7}$ : **Known materials.** We know the number of materials and can identify samples from each material and mixture within the data.

*Description.* The voxel parameters for this algorithm are a discrete parameter,  $\alpha_b$  that determines the material or mixture, a signed distance,  $d$ , from a boundary for mixtures, and an estimate,  $\bar{N}$ , of the low-frequency noise within the voxel. Once again, the histogram basis functions for pure materials are normal distributions. The basis functions for mixtures are

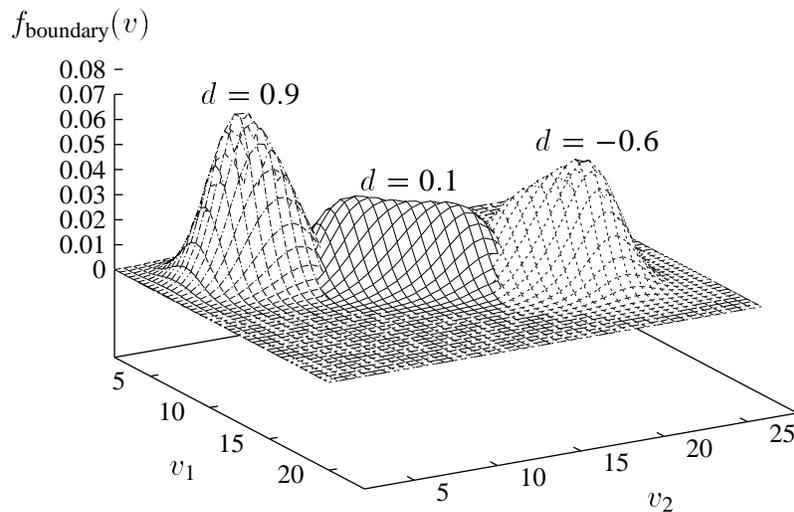
$$f_{\text{boundary}}(v; d, \bar{N}, c, s, k_w) = k_n(v; s) * \left( \left( H\left(d + \frac{k_h}{2} - \frac{k_e(v)}{k_w}\right) - H\left(d - \frac{k_h}{2} - \frac{k_e(v)}{k_w}\right) \right) \left| \frac{e^{k_e(v)} \sqrt{\pi}}{(c_2 - c_1) k_w} \right| \right) \quad (1.8)$$

where  $k_e(v) = \text{erf}^{-1}\left(\frac{c_1 + c_2 - 2v}{c_1 - c_2}\right)$  and  $H(x)$  is the Heaviside, or step, function. The basis functions are shown in Figure 1.9 and 1.10.

The dataset parameters are the mean,  $c$ , and standard deviation,  $s$ , values for each pure



**Figure 1.9** Histogram basis functions,  $f_{\text{boundary}}(v)$ , for scalar data. The shapes approach normal distributions as  $d$ , the distance from the boundary to a voxel center, moves away from 0.



**Figure 1.10** Histogram basis functions,  $f_{\text{boundary}}(v)$ , for vector-valued data. Three different values of  $d$ , the distance from the boundary to a voxel center, are represented.

material, the width,  $k_w$  (in samples), of the acquisition sampling kernel, and an expected deviation of the model histogram from actual histograms. They are estimated from a training set of points interactively chosen for each material and mixture.

### 1.3.4 Overview of Algorithm C: Boundary Distance with Non-Uniform Material Signatures

Our third algorithm augments the boundary distance algorithm to handle a common characteristic of MRI data that often complicates classification: MRI measurements of the same material can be different at different spatial locations. There are a number of factors that can cause these intensity distortions, from antenna coils that produce spatially dependent RF radiation to different amounts of RF absorption in different parts of an object being imaged. The algorithm relaxes the assumption that the expected value for a material is constant; instead, the expected value is a function of spatial location.

*Additions to common assumptions.* Only  $e_{v5}$  differs from the assumptions for the boundary distance algorithm.

$e_{v4}$ : **Only pair-wise mixtures.**

$e_{v5}$ : **Predictable tissue measurements.** Measurements for the same material have an expected value that can be modeled with a small number of parameters across a dataset.

$e_{v6}$ : **Gaussian filtering.**

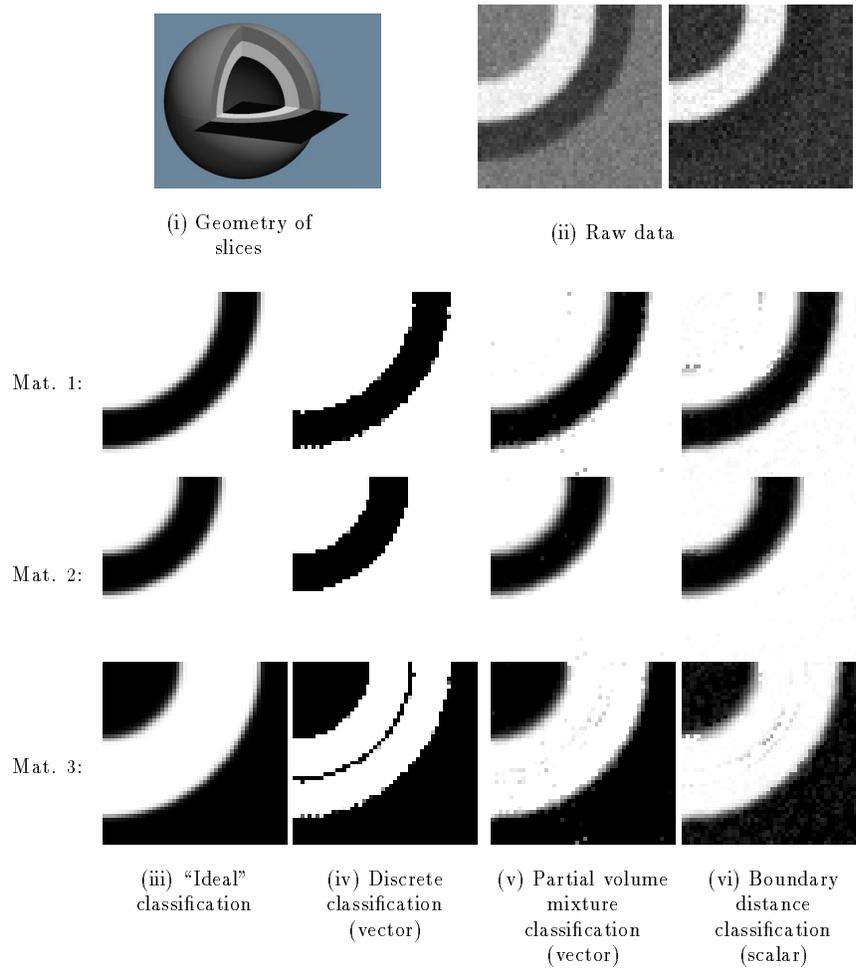
$e_{v7}$ : **Known materials.**

*Description.* This algorithm is very similar to the boundary distance algorithm. The basis functions and voxel parameters are the same. Only the dataset parameters that determine the expected value for a tissue measurement are different. In this case, the expected value is a parameterized function of space. Its parameters are calculated from a set of interactively-specified points for each material. The calculation is similar to calculating a mean and variance from a set of points, but the mean is now a function of spatial location.

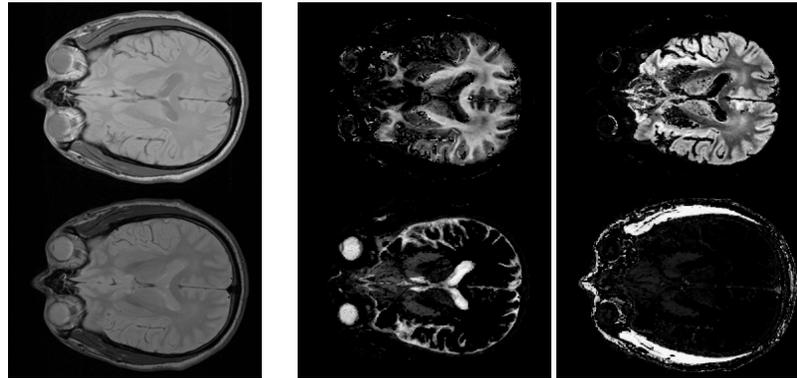
## 1.4 Results

We have applied our new technique to several MRI datasets, both simulated and real. Figure 1.11 shows results generated from simulated imaging data of a tennis ball-like object – a hollow ball with two concentric surface layers. Figures 1.3 and 1.12 show classification results from an MRI section of a human brain.

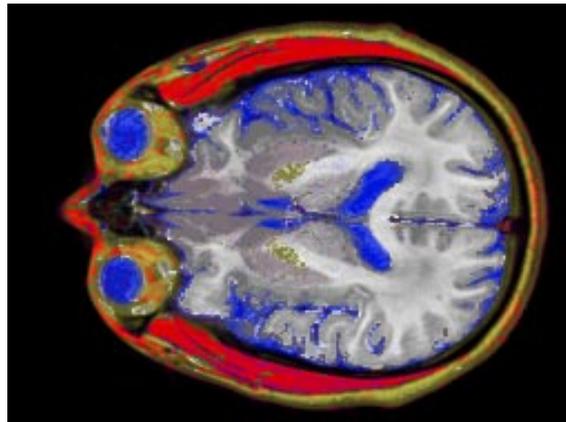
We first compare our techniques with a probabilistic approach that uses pure materials only and only a single measurement value per voxel. The new technique produces many fewer misclassified voxels, particularly in regions where materials are mixed due to filtering. The difference is illustrated in Figures 1.11(iii) and (iv) where an incorrect layer of background material has been introduced between the white and gray regions. This layer occurs because multiple materials are present in each voxel. The simulated data that we classified is shown in Figure 1.11(ii) with Figure 1.11(iii) illustrating what an ideal classification algorithm would produce. Discrete classification, using vector data, but only a single measurement point per voxel and assuming only pure materials, produces the results in Figure 1.11(iv). Note the jaggy



**Figure 1.11** The partial volume, boundary distance, and discrete algorithms are compared in classifying simulated MRI data. (ii) shows the simulated data, which contains three different materials. The geometry that the data measures is shown in (i). (iii) shows what an "ideal" classification algorithm should produce and (iv)-(vi) show results from different algorithms. Note that the new algorithms (v) and (vi) produce results most similar to the ideal case, and that (vi) does so even with scalar data.



(i) Original Data

(ii) Results of Algorithm  
Classified White Matter, Gray Matter  
Cerebro-Spinal Fluid, Muscle

(iii) Combined Classified Image

**Figure 1.12** One slice of data from a human brain. (i) shows the original two-valued data, (ii) shows four of the identified materials, white matter, gray matter, cerebro-spinal fluid, and muscle, separated out into different images, and (iii) shows the results of the new classification mapped to different colors (see Plate ??). Note the smooth boundaries where materials meet and the much lower incidence of misclassified samples than in Figure 1.3.



**Figure 1.13** A volume-rendering image of a human hand dataset. The opacity of different materials is decreased above cutting planes to show details of the classification process within the hand (see Plate ??).

edges and the band of misclassified data for material 3 along the boundary between materials 1 and 2. Figure 1.11(v) and Figure 1.11(vi) show the partial volume mixtures algorithm and the boundary distance algorithm. Even with scalar data the boundary distance algorithm achieves results very close to the ideal case.

Figure 1.3 and 1.12(iii) also show comparative results for classification of an MR section through a human brain. Note in the new classification results the reduction of jagged edges between materials and the virtual elimination of the layer of misclassified brain material just under the skin.

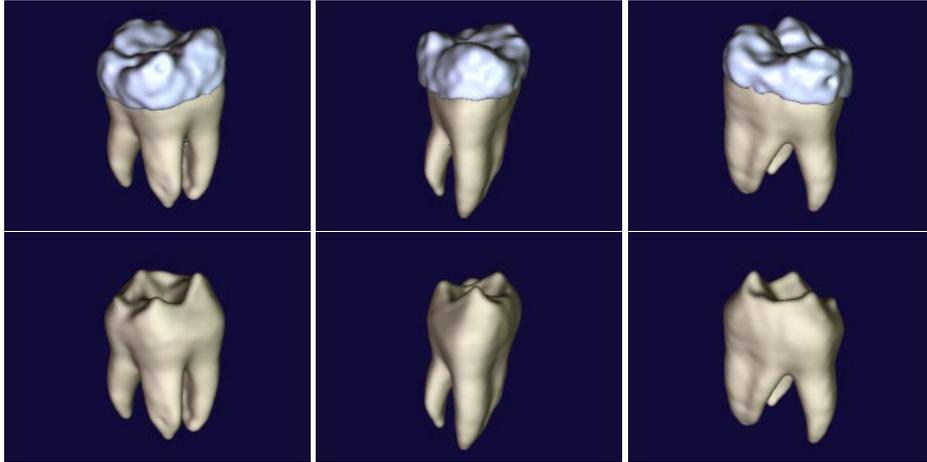
Models and volume rendered images, as shown in Figure 1.13 and 1.14, benefit from our approach because less incorrect information is introduced into the classified datasets, and so the images and models more accurately depict the objects they are representing. With other classification techniques, models and images contain jaggy artifacts, particularly along surfaces where materials meet, because the algorithms are less accurate there.

## 1.5 Discussion

We have made several assumptions and approximations while developing and implementing the continuous Bayesian algorithms. This section discusses some of the tradeoffs and suggests some possible directions for future work.

### 1.5.1 Mixtures of Three or More Materials.

We assume that each measurement contains values from at most two materials, although the approach easily extends to mixtures with more materials. We chose two-material mixtures



**Figure 1.14** A geometric model of tooth dentine and enamel created by collecting MRI data using a technique that images hard solid materials [GLF<sup>+</sup>95] and classifying dentine and enamel in the volume data with our new partial-volume mixtures algorithm. Polygonal isosurfaces define the bounding surfaces of the dentine and enamel. The enamel-dentine boundary, shown in the lower images, is difficult to examine non-invasively using any other technique.

because surfaces between boundaries of pure materials are one of the most important parts of computer graphics models. Voxels containing three-material mixtures happen near lines where three materials meet, and are generally much less common because the dimensionality of the lines is smaller than the dimensionality of surfaces where two materials meet.

Our algorithm chooses a classification for voxels containing more than two materials from the set of 2-material mixtures. Generally, the two most prevalent materials in the voxel influence the choice, producing a dataset with small artifacts where three or more materials come together.

### 1.5.2 More sophisticated geometric basis functions.

The basis functions that we have developed model the two most common geometric cases: samples within regions of pure material and samples near surface boundaries. Additional basis functions, however, could model other geometries and create more accurate models. Examples include samples near edges where three materials come together, or points near membranes that are thinner than the sample spacing, where, again, three materials would have an effect on the measurement.

### 1.5.3 Ambiguous classification.

For a voxel that is well within a region of pure material  $A$ , the algorithms sometimes correctly classify the voxel as pure material  $A$ , and sometimes classify it as a mixture of  $A$  and a small amount of some other material. Both solutions are physically reasonable because the mixture basis functions approach a normal distribution as the boundary distance parameter  $d$  moves away from zero.

Similarly, two different mixtures, each containing material  $A$ , can match a voxel that is within a region of pure material  $A$ . Again, the choice is not critical.

#### **1.5.4 Sensitivity to interactively selected material classes.**

The results of the algorithm are highly dependent on the material points selected interactively to represent each pure material and each pair of materials. These points must be selected carefully, and should come from a set of points that actually represent a single consistent material. Representing points from two materials as one material can create a situation where the distributions of the sample values do not match a normal distribution, and the classification results are less accurate.

#### **1.5.5 Sensitivity to contrast between materials.**

The classification is sensitive to the contrast-to-noise ratio between different materials. If this ratio is too small, materials cannot be distinguished effectively. We pass this requirement back to the data acquisition process, and use goal-directed techniques to ensure that we collect data that can be classified effectively [LBJ97].

#### **1.5.6 Computational expense.**

The implementations described in this chapter are computationally expensive. The optimization process must be run on each voxel in a dataset. At ten voxels per second, a medium-sized dataset of  $256 \times 256 \times 64$  voxels runs in about 5 days. Through approximations it may be possible to reduce this time significantly.

The algorithm processes each voxel independently, and so is highly amenable to a domain-decomposition parallel solution. In fact, we have run it on a network of ten HP 9000/700 and DEC Alpha workstations and gotten a speedup of almost ten in classifying medium to large datasets.

#### **1.5.7 Incorporating additional global information.**

Except for the interpolation of samples, we currently classify each voxel without regard to its neighbors and without directly using the interactively selected representative points for each material. Both types of information could be incorporated into the prior probability estimates to influence the classification process.

### **1.6 Conclusions**

The algorithms described here for classifying scalar- and vector-valued volume data produces more-accurate results than existing techniques in many cases, particularly at boundaries between materials. The aspects responsible for this improvement are: 1) the reconstruction of a continuous function from the samples, 2) the use of histograms taken over voxel-sized regions to represent the contents of the voxels, 3) the modeling of sub-voxel partial-volume effects caused by the band-limiting nature of the acquisition process, and 4) the use of a Bayesian classification approach. We have demonstrated the technique on both simulated and real data, and it correctly classifies many voxels containing multiple materials.

The construction of a continuous function is based on the sampling theorem, and while it does not introduce new information, it provides a richer context for the information that classification algorithms such as ours can use. It incorporates information about neighboring samples into the classification process for a voxel in a natural and mathematically rigorous way and thereby greatly increases classification accuracy. In addition, because the operations that can be safely performed directly on sampled data are so limited, treating the data as a continuous function helps to avoid introducing artifacts.

Histograms are a natural choice for representing voxel contents for a number of reasons. First, they generalize single measurements to measurements over a region, allowing classification concepts that apply to single measurements to generalize. Second, the histograms can be calculated easily. Third, the histograms capture information about neighboring voxels; this increases the information content over single measurements and improves classification results. Fourth, histograms are orientation independent; orientation independence reduces the number of parameters in the classification process hence simplifying and accelerating it.

Partial-volume effects are a nemesis of classification algorithms, which traditionally have drawn from techniques that classify isolated measurements. These techniques do not take into account the related nature of spatially-correlated measurements. Many attempts have been made to model partial-volume effects, and ours continues that trend, with results that suggest that further study is warranted.

We believe that the Bayesian approach we describe is a useful formalism for capturing the assumptions and information gleaned from the continuous representation of the sample values, the histograms calculated from them, and the partial-volume effects of imaging. Together, these allow a generalization of many sample-based classification techniques, three of which we have demonstrated.

## 1.7 Acknowledgements

This work was supported in part by grants from Apple, DEC, Hewlett Packard, and IBM. Additional support was provided by NSF (ASC-89-20219) as part of the NSF STC for Computer Graphics and Scientific Visualization, by the DOE (DE-FG03-92ER25134) as part of the Center for Research in Computational Biology, and by the National Institute on Drug Abuse, the National Institute of Mental Health, and the National Science Foundation as part of the Human Brain Project.

Thanks to Al Barr for much advice, to Kurt Fleischer for his help in developing Algorithm A, to Matt Avalos for extensive programming support, to Jim Arvo, Al Barr, and Peter Schröder for brainstorming on Figure 1.1, and to Pratik Ghosh and Russ Jacobs for the collaborative effort behind Figure 1.14.

## 1.8 References

- [CHK91] H. S. Choi, D. R. Haynor, and Y. M. Kim. Partial volume tissue classification of multi-channel magnetic resonance images — a mixel model. *IEEE Transactions on Medical Images*, 10(3):395–407, 1991.
- [CLKJ90] Harvey E. Cline, William E. Lorensen, Ron Kikinis, and Ferenc Jolesz. Three-dimensional segmentation of MR images of the head using probability and connectivity. *Journal of Computer Assisted Tomography*, 14(6):1037–1045, November, December 1990.

- [CVC<sup>+</sup> 95] L. P. Clarke, R. P. Velthuizen, M. A. Camacho, J. J. Neine, M. Vaidyanathan, L. O. Hall, R. W. Thatcher, and M. L. Silbiger. MRI segmentation: Methods and applications. *Magnetic Resonance Imaging*, 13(3):343–368, 1995.
- [DCH88] Robert A. Drebin, Loren Carpenter, and Pat Hanrahan. Volume rendering. In John Dill, editor, *Computer Graphics (SIGGRAPH '88 Proceedings)*, volume 22, pages 65–74, August 1988.
- [DH73] Richard P. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, New York, 1973.
- [GLF<sup>+</sup> 95] Pratik R. Ghosh, David H. Laidlaw, Kurt W. Fleischer, Alan H. Barr, and Russell E. Jacobs. Pure phase-encoded MRI and classification of solids. *IEEE Transactions on Medical Imaging*, 14(3):616–620, 1995.
- [GLS<sup>+</sup> 97] Galen G. Gornowicz, David H. Laidlaw, Jerry W. Shan, Daniel B. Lang, and Paul E. Dimotakis. De-aliasing undersampled volume images for visualization. Technical Report CS-TR-97-11, Caltech, 1997.
- [KSW96] Yi-Hsuan Kao, James A. Sorenson, and Stefan S. Winkler. MR image segmentation using vector decomposition and probability techniques: A general model and its application to dual-echo images. *Magnetic Resonance in Medicine*, 35:114–125, 1996.
- [Lai92] David H. Laidlaw. Tissue classification of magnetic resonance volume data. Master's project, California Institute of Technology, 1992.
- [Lai95] David H. Laidlaw. *Geometric Model Extraction from Magnetic Resonance Volume Data*. PhD thesis, Caltech, 1995.
- [LBJ97] David H. Laidlaw, Alan H. Barr, and Russell E. Jacobs. Goal-directed brain micro-imaging. In Steven H. Koslow and Michael F. Huerta, editors, *Neuroinformatics: An Overview of the Human Brain Project*, volume 1, chapter 6. Kluwer, February 1997.
- [LC87] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In Maureen C. Stone, editor, *Computer Graphics (SIGGRAPH '87 Proceedings)*, volume 21, pages 163–169, July 1987.
- [Lev88] Marc Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(3):29–37, May 1988.
- [LFB97] David H. Laidlaw, Kurt W. Fleischer, and Alan H. Barr. Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms. Technical Report CS-TR-97-09, Caltech, 1997.
- [Lor89] T. J. Lored. From Laplace to supernova SN1987A: Bayesian inference in astrophysics. In P. Fougere, editor, *Maximum Entropy and Bayesian Methods*. Kluwer Academic Publishers, Denmark, 1989.
- [NFMD90] Derek R. Ney, Elliot K. Fishman, Donna Magid, and Robert A. Drebin. Volumetric rendering of computed tomography data: Principles and techniques. *IEEE Computer Graphics and Applications*, 10(2):24–32, March 1990.
- [OWY83] Alan V. Oppenheim, Alan S. Willsky, and Ian T. Young. *Signals and Systems*. Prentice-Hall, Inc., New Jersey, 1983.
- [SG93] Peter Santago and H. Donald Gage. Quantification of MR brain images by mixture density and partial volume modeling. *IEEE Transactions on Medical Imaging*, 12(3):566–574, september 1993.
- [VBJ<sup>+</sup> 85] Michael W. Vannier, Robert L. Butterfield, Douglas Jordan, William A. Murphy, Robert G. Levitt, and Mokhtar Gado. Multispectral analysis of magnetic resonance images. *Radiology*, 154:221–224, 1985.
- [VSR88] Michael W. Vannier, Christopher M. Speidel, and Douglas L. Rickman. Magnetic resonance imaging multispectral tissue classification. In *Proc. Neural Information Processing Systems (NIPS)*, August 1988.
- [WAAR<sup>+</sup> 88] Joe P. Windham, Mahmoud A. Abd-Allah, David A. Reimann, Jerry W. Froelich, and Allan M. Hagggar. Eigenimage filtering in MR imaging. *Journal of Computer Assisted Tomography*, 12(1):1–9, 1988.
- [WCW88] Zhenyu Wu, Hsiao-Wen Chung, and Felix W. Wehrli. A bayesian approach to subvoxel tissue classification in NMR microscopic images of trabecular bone. *Journal of Computer Assisted Tomography*, 12(1):1–9, 1988.

# References

- [Aam86] L. A. Aamport. The gnats and gnu document preparation system. *G-Animal's Journal*, 41(7):73–79, 1986. This is a full ARTICLE entry.
- [GAJ86] *G-Animal's Journal*, 41(7), July 1986. The entire issue is devoted to gnats and gnu. This entry is a cross-referenced ARTICLE (journal).
- [Knu73] Donald E. Knuth. *Fundamental Algorithms*, volume 1 of *The Art of Computer Programming*, section 1.2, pages 10–119. Addison-Wesley, Reading, Massachusetts, second edition, 1973. This is a full INBOOK entry.
- [Knu81] Donald E. Knuth. *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*. Addison-Wesley, Reading, Massachusetts, second edition, 1981. This is a full BOOK entry.
- [Lin77] Daniel D. Lincoll. Semigroups of recurrences. In David J. Lipcoll, D. H. Lawrie, and A. H. Sameh, editors, *High Speed Computer and Algorithm Organization*, number 23 in *Fast Computers*, part 3, pages 179–183. Academic Press, New York, third edition, 1977. This is a full INCOLLECTION entry.
- [Man86] Larry Manmaker. *The Definitive Computer Manual*. Chips-R-Us, Silicon Valley, silver edition, 1986. This is a full MANUAL entry.
- [Mas88] Édouard Masterly. Mastering thesis writing. Master's project, Stanford University, English Department, June-August 1988. This is a full MASTERSTHESIS entry.
- [Oah83] Alfred V. Oaho, Jeffrey D. Ullman, and Mihalis Yannakakis. On notions of information transfer in VLSI circuits. In Wizard V. Oz and Mihalis Yannakakis, editors, *Proc. Fifteenth Annual ACM Symposium on the Theory of Computing*, number 17 in *All ACM Conferences*, pages 133–139, Boston, March 1983. The OX Association for Computing Machinery, Academic Press, London. This is a full INPROCEEDINGS entry.
- [Oz83] Wizard V. Oz and Mihalis Yannakakis, editors. *Proc. Fifteenth Annual ACM Symposium on the Theory of Computing*, number 17 in *All ACM Conferences*, Boston, March 1983. The OX Association for Computing Machinery, Academic Press, London. This is a full PROCEEDINGS entry.
- [Pho88] F. Phidias Phony-Baloney. *Fighting fire with fire: festooning French phrases*. PhD dissertation, Fanstord University, Department of French, June-August 1988. This is a full PHDTHESIS entry.
- [Tér88] Tom Tèrrific. An  $O(n \log n / \log \log n)$  sorting algorithm. Wishful Research Result 7, Fanstord University, Computer Science Department, Fanstord, California, October 1988. This is a full TECHREPORT entry.
- [Ünd88] Ulrich Ünderwood, Ned Ñet, and Paul P̄ot. Lower bounds for wishful research results. Talk at Fanstord University, November, December 1988. This is a full UNPUBLISHED entry.