Partial-Volume Bayesian Classification of Material Mixtures in MR Volume Data using Voxel Histograms

David H. Laidlaw[1], Kurt W. Fleischer[2], Alan H. Barr[1]

[1]California Institute of Technology, Pasadena, CA 91125

[2]Pixar Animation Studios, Richmond, CA 94804,

June 17, 1997

Abstract

We present a new algorithm for identifying the distribution of different material types in volumetric datasets such as those produced with Magnetic Resonance Imaging (MRI) or Computed Tomography (CT). Because we allow for mixtures of materials and treat voxels as regions, our technique reduces the classification artifacts that thresholding can create along boundaries between materials and is particularly useful for creating accurate geometric models and renderings from volume data. It also has the potential to make more-accurate volume measurements and classifies noisy, low-resolution data well.

There are two unusual aspects to our approach. First, we assume that, due to partial-volume effects, voxels can contain more than one material, e.g., both muscle and fat; we compute the relative proportion of each material in the voxels. Second, we incorporate information from neighboring voxels into the classification process by reconstructing a continuous function, $\rho(x)$, from the samples and then looking at the distribution of values that ρ takes on within the region of a voxel. This distribution of values is represented by a histogram taken over the region of the voxel; the mixture of materials that those values

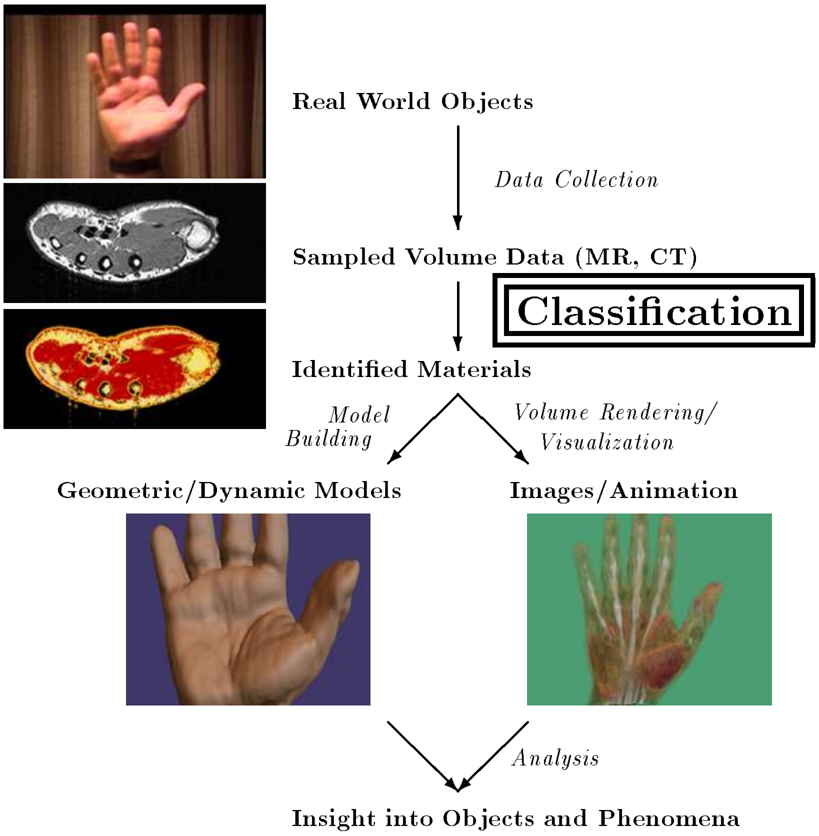


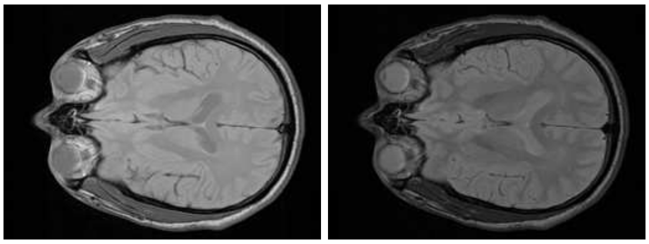
Figure 1: Classification is a key step in the process of the process of visualizing and extracting geometric information from sampled volume data. For accurate geometric results, some constraints on the classification accuracy must be met.

---

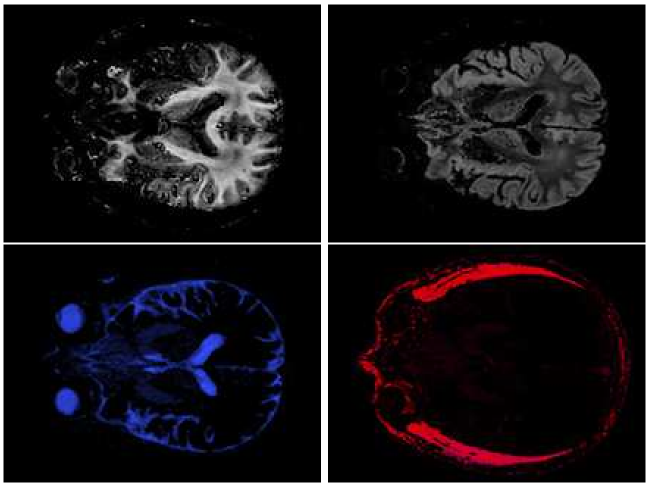
measure is identified within the voxel using a probabilistic Bayesian approach that matches the histogram by finding the mixture of materials within each voxel most likely to have created the histogram. The size of regions that we classify is chosen to match the spacing of the samples because the spacing is intrinsically related to the minimum feature size that the reconstructed continuous function can represent.

1 Introduction

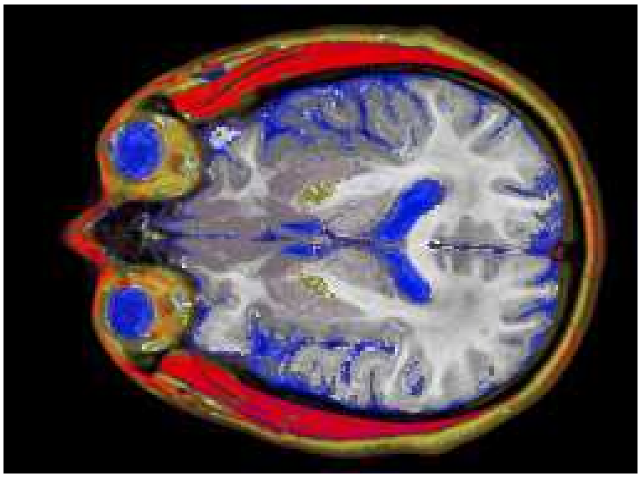
Identifying different materials within sampled datasets can be an important step in understanding the geometry, anatomy, or pathology of a subject. By accurately locating different materials, we can identify them as individual parts and measure their size and shape. We can also use the spatial location of materials



(i) Original Data



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



(iii) Combined Classified Image

Figure 2: One slice of data from a human brain. (i) The original two-valued MRI data. (ii) Four of the identified materials, white matter, gray matter, cerebro-spinal fluid and muscle, separated out into separate images. (iii) The results of the new classification mapped to different colors. Note the smooth boundaries where materials meet and the much lower incidence of misclassified samples than in Figure 5.

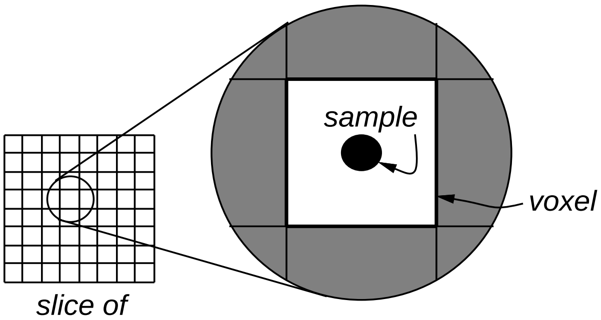


Figure 3: We define a sample as a scalar or vector valued element of a 2-D or 3-D dataset. A voxel is the region surrounding a sample.

to selectively visualize parts of the data, thus better controlling a volume-rendered image [Levoy, 1988], a surface model [Lorensen and Cline, 1987], or a volume model created from the data, and making visible otherwise obscured or subtle features. Classification is a key step towards understanding such geometry (Figure 1). Figure 2 shows an example of classified MRI data; each color represents a single material identified within the data.

Applications of classified images and geometric models derived from them include surgical planning and assistance, diagnostic medical imaging, conventional computer animation, anatomical studies, and predictive modeling of complex biological shapes and behavior.

Our Approach. We use Bayesian probability theory to estimate the posterior probability based on conditional and prior probabilities derived from our assumptions about what we are measuring and how the measurement process works [Loredo, 1989]. With this information we identify the materials contained within each voxel based on the sample values for the voxel and its neighbors. We treat each voxel as a region (see Figure 3), not as a single point. The sampling theorem [Oppenheim et al., 1983] allows us to reconstruct a continuous function, $\rho(x)$, from the samples. We then represent all of the values that $\rho(x)$ takes on within a voxel by creating a histogram of $\rho(x)$ over the voxel. Figure 4(i) shows samples, Figure 4(ii)

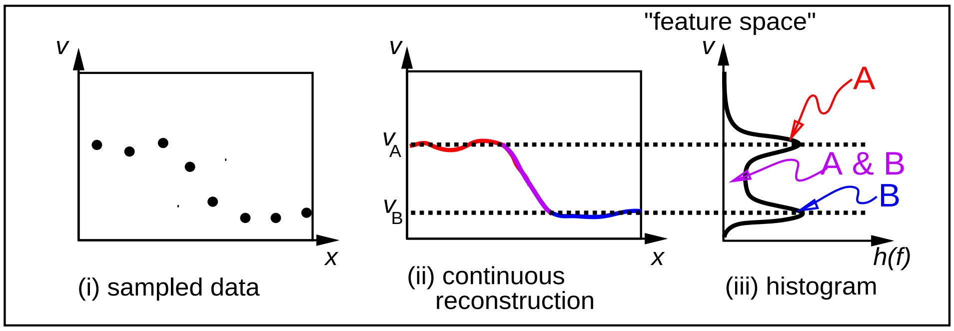


Figure 4: Continuous histograms. The scalar data in (i) and (ii) represent measurements from a dataset containing two materials, A and B, such as that shown in Figure 6. One material has measurement values near v_A and the other near v_B. These values correspond to the Gaussian-shaped peaks centered around v_A and v_B in the histograms, which are shown on their sides to emphasize the axis that they share. This shared axis is "feature space."

---

the function $\rho(x)$ reconstructed from the samples, and Figure 4(iii) a continuous histogram calculated from $\rho(x)$.

We assume that each voxel is a mixture of materials, with mixtures occurring where partial-volume effects occur, i.e., where the band-limiting process blurs pure materials together. From this assumption we derive basis functions that model histograms of voxels containing a pure material and voxels containing a mixture of two materials (Section 4). Linear combinations of these basis histograms are fit to each voxel, and the most likely combination of materials chosen probabilistically.

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. The advantages of this are discussed further in Section 8.

Related Work. Many researchers have worked on identifying the locations of materials in sampled datasets [Vannier et al., 1985], [Vannier et al., 1988], [Cline et al., 1990], [Duda and Hart, 1973]. [Clarke et al., 1995] gives an extensive review of the segmentation of MRI data. However, many of these algorithms generate artifacts like those shown in Figure 5, an example of data classified with a maximum-

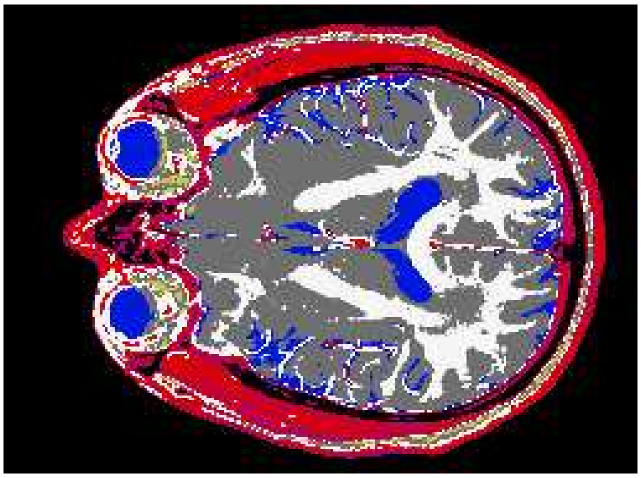


Figure 5: Discrete, single-material classification of the same slice shown in Figure 2. Note the jagged boundaries between materials within the brain and the layer of misclassified white matter outside of the skull.

---

likelihood technique based on sample values. These techniques work well in regions where the region of a voxel contains only a single material, but tend 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.

[Drebin et al., 1988] 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.

[Windham et al., 1988] and [Kao et al., 1996] 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 flexible as non-linear operations, and so either more data must be acquired or classification results will not be as accurate.

[Choi et al., 1991] and [Ney et al., 1990] 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.

[Santago and Gage, 1993] 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.

[Wu et al., 1988] 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 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.

[Laidlaw et al., 1997] gives an overview of this technique in the context of the Human Brain Project, and [Laidlaw, 1995] gives a complete description. [Ghosh et al., 1995] describes an imaging protocol for acquiring MRI data from solids and applies the technique to the extraction of a geometric model from MRI data of a human tooth (see Figure 12).

2 Overview

In this section we describe the classification problem that we solve, define terms, state assumptions we make about the data we classify, and sketch the algorithm and its derivation. Sections 3–6 give more information on each part of the process, with detailed derivations in Appendices A and B. Section 7 shows results of the application of the algorithm to simulated MR data and to real MR data of a human brain, hand, and tooth. We discuss some limitations and future extensions in Section 8 and conclude in Section 9.

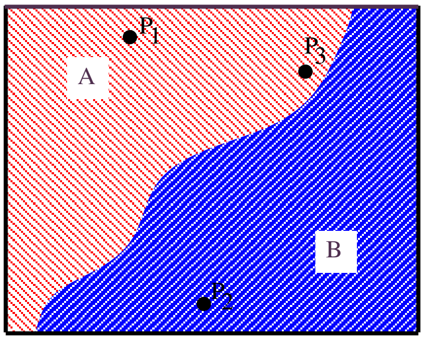
Problem Statement. The input to our process is sampled measurement data, from which we can reconstruct a continuous, band-limited function, $\rho(x)$, that measures distinguishing properties of the underlying materials. The output is sampled data measuring the relative volume of each material in each voxel.

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. This space is n_x-dimensional, where n_x is three for volume data, can be two for slices, and is one for the example in Figure 4. Each measurement, which may be a scalar or vector, lies in "feature space" (see Figure 4), 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-valued vector data, etc. Tables 2 and 3 in Appendix B may be useful for checking other definitions.

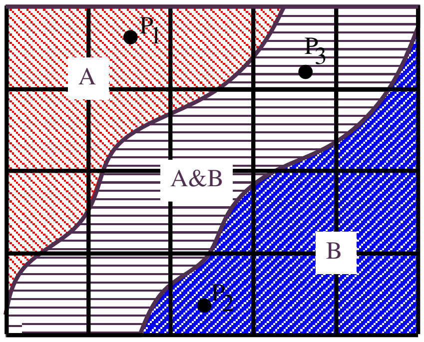
Assumptions. We make a set of assumptions about the objects that we are measuring and the measurement process.

1: **Discrete materials.** The first assumption is that materials within the objects that we measure are discrete at the resolution that we are sampling. Boundaries need not be aligned with the sampling grid. Figure 6(i) shows an object with two materials. We make this assumption because we are generally looking for boundaries between materials, and because we start from sampled data, where information about detail finer than the sampling rate is blurred.

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



(i) Real World Object



(ii) Sampled Data

Figure 6: Partial-volume effects. We 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 the samples we reconstruct a continuous, band-limited measurement function $\rho(x)$. Points P_1 and P_2 lie inside regions of a single material. Point P_3 lies near a boundary between materials, and so in (ii) lies in the A&B region where materials A and B are mixed. The grid lines show how the regions may span voxels.

---

transition from one to another over many samples or are not relatively uniformly mixed; however, our algorithm appears to degrade gracefully even in these cases.

2: **Normally-distributed noise.** We assume that noise from the measurement process 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, but seems to be satisfied sufficiently to classify data well. Note that the sample values with noise added are interpolated to reconstruct the continuous function, $\rho(x)$. The effect of this band-limited noise is discussed further in Section 6.

3: **Sampling theorem is satisfied.** The third assumption we make is that the sampled datasets we classify satisfy the sampling theorem [Oppenheim et al., 1983]. The sampling theorem states that if we sample a sufficiently band-limited function, we can exactly reconstruct that function from the samples, as demonstrated in Figure 4(ii). The band-limiting creates smooth transitions in $\rho(x)$ as it traverses boundaries where otherwise $\rho(x)$ would be discontinuous. The intermediate region of Figure 6(ii) shows a sampling grid and the effect of sampling that satisfies the sampling theorem. Partial-volume mixing of measurements occurs in the region labeled “A & B.”

4: **Linear mixtures.** Each voxel measurement is a linear combination of pure material measurements and measurements of their pair-wise mixtures created by band limiting (see Figure 6).

5: **Uniform tissue measurements.** Measurements of the same material have the same expected value and variance throughout a dataset.

6: **Box filtering.** The spatial measurement kernel, or point-spread function, can be approximated by a box filter for the purpose of deriving a histogram basis function. This helps the derivation remain tractable and appears to be accurate enough to classify data well.

7: **Materials identifiable in histogram of entire dataset.** The signatures for each material and mixture must be identifiable in a histogram of the entire dataset. This implies both that there must be sufficient material to recognize in the histogram, and that material signatures must be sufficiently far apart to be distinguished. This assumption is not necessary for the derivation, but is needed for the implementation to work reliably.

For many types of medical imaging data, including MRI and CT, these assumptions hold reasonably well, or can be satisfied sufficiently with preprocessing [Laidlaw, 1992]. Other types of sampled data, e.g., ultrasound, and video or film images with lighting and shading, violate these assumptions, and our technique does not apply directly.

Sketch of Derivation. Histograms represent the values taken on by $\rho(x)$ over various spatial regions. In section 3 we describe the histogram equation for a normalized histogram of data values within a region. In Section 4 we use the histogram equation to create basis functions that model histograms taken over small, voxel-sized regions. These basis functions model histograms for regions consisting of single materials and for regions consisting of mixtures of two materials. These mixtures are assumed to be partial-volume effects created by the band-limiting process accompanying sampling. The parameters for these model histograms represent the mean value, c, and variance, s, of a measurement of a pure material.

We fit histograms of both the entire dataset and the individual voxels. The dataset histograms allow us to estimate parameters of the materials that we want to identify, and the voxel histograms give us information about how much of each material is within each voxel. Using Bayes' Theorem, the histogram of the entire dataset, our histogram model basis functions, and a series of approximations, we derive an estimate of the most likely set of materials within an entire dataset in Section 5. Similarly, given the histogram of a voxel-sized region, we derive, in Section 6, an estimate of the most likely density for each material in that voxel.

Sketch of Algorithm. The algorithm produces, as its end result, a sampled dataset containing estimates of the relative amount of each material in each voxel. The process is illustrated in Figure 7. First, we collect and preprocess data to satisfy the assumptions listed above. Second, we calculate a histogram of the entire dataset, and fit a linear combination of parameterized model histograms to the dataset histogram. Third, using the fitted parameters, we process each voxel-sized region in the dataset by calculating a histogram for the voxel and then finding the combination of materials most likely to have produced the histogram.

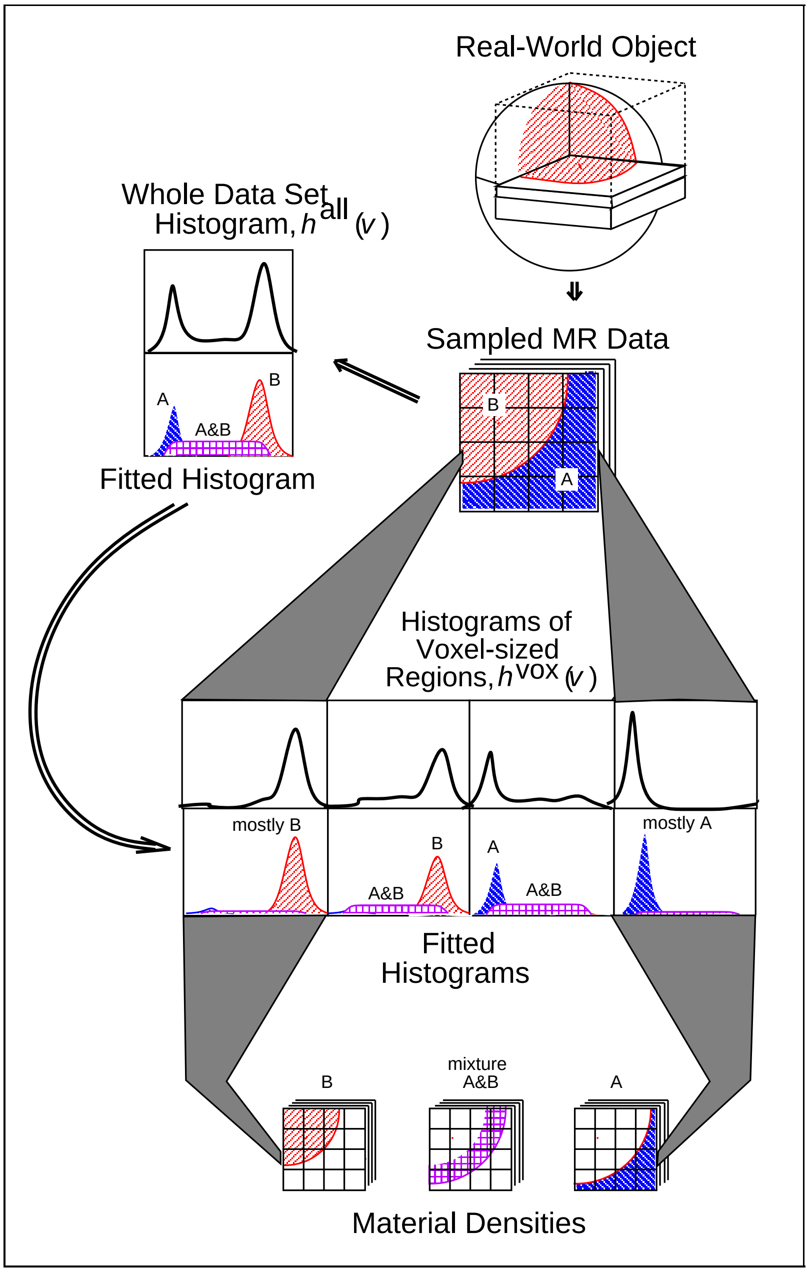


Figure 7: The classification process. We collect MR data, calculate a histogram of the entire dataset, $h^{\mathrm{all}}(v)$, and use that to determine parameters of histogram-fitting basis functions. We then calculate histograms of each voxel-sized region, $h^{\mathrm{vox}}(v)$, and identify the most likely mixture of materials for that region. The result is a sampled dataset of material densities within each voxel.

3 Normalized Histograms

In this section we present the equation for a normalized histogram of a sampled dataset over a region. We will use this equation as a building block in several later sections, with regions that vary from the size of a single voxel to the size of the entire dataset. We will also use this equation to derive basis functions that model histograms over regions containing single materials and regions containing mixtures of materials.

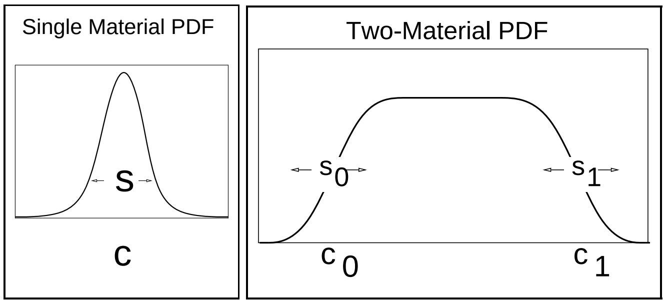
For a given region in spatial coordinates, specified by $\mathcal{R}$, the histogram $h^{\mathcal{R}}(v)$ specifies the relative portion of that region where $\rho(x) = v$, as shown in Figure 4. Because we treat a dataset as a continuous function over space, histograms, $h^{\mathcal{R}}(v) : \mathbf{R}^{n_v} \rightarrow \mathbf{R}$, are also continuous functions:

$$h^{\mathcal{R}}(v) = \int \mathcal{R}(x)\delta(\rho(x) - v)dx \tag{1}$$

This equation is the continuous analogue of a discrete histogram. $\mathcal{R}(x)$ is non-zero within the region of interest and integrates to one. We define $\mathcal{R}(x)$ to be constant in the region of interest, making every spatial point contribute equally to the histogram $h^{\mathcal{R}}(v)$, but $\mathcal{R}(x)$ can be considered a weighting function that takes on values other than zero and one to more smoothly transition between adjacent regions. Note also that $h^{\mathcal{R}}(v)$ integrates to one, which means that it can be treated as a probability density function, or PDF. δ is the Dirac-delta function.

We use Equation 1 both as a basis for calculating histograms of regions of our datasets and also as a starting point for deriving models of those dataset and voxel histograms based on the materials they contain. The equations are described in Sections 4 and 5 and derived in Appendix A. We will now discuss some implementation considerations for calculating histograms.

Computing Voxel Histograms. We calculate histograms in constant-sized rectangular "bins," sized such that the width of a bin is smaller than the standard deviation of the noise within the dataset. This ensures that we do not lose significant features in the histogram.



(i) (ii)

Figure 8: Parameters for histogram basis function. (i) Single-material histogram parameters include c, the mean value for the material, and s, which measures the variance of the noise (see Equation 2). (ii) Corresponding parameters for a two-material mixture basis function. s_0 and s_1 affect the slopes of the two-material histogram basis function at either end. For vector-valued data, c and s are vectors and are the mean values and variances of the noise for the two constituent materials (see Equation 3).

We first initialize the bins to zero. We subdivide each voxel into sub-voxels, usually 4 for 2-D data or 8 for 3-D data, and evaluate $\rho(x)$ and its derivative at the center of each sub-voxel. From the value and derivative we use Equation 1 to calculate the contribution of a linear approximation of $\rho(x)$ over the sub-voxel to each histogram bin, accumulating the contributions from all sub-voxels. This gives us a more-accurate histogram than would evaluating just the function values at the same number of points.

4 Histogram Basis Functions for Pure Materials and Mixtures

In this section we describe basis functions that model histograms of regions consisting of pure materials and of regions consisting of pairwise mixtures of materials. Other voxel contents are also possible. See Section 8 for a discussion. The parameters of the basis functions specify the expected value, c, and variance, s, of each material's measurements (see Figure 8).

We use Equation 1 to derive these basis functions, which we fit to histograms of the data. We then verify that the equations provide reasonable fits to typical MR data, which gives us confidence that our assumptions about the measurement function, $\rho(x)$, are reasonable. The details of the derivations are in Appendix A.

For a single material, the histogram produced is a Gaussian distribution:

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

We derive this equation by manipulating Equation 1 evaluated over a region of constant material, where the measurement function, $\rho(x)$, is a constant value plus additive, normally-distributed noise. Because the noise in different channels of multi-valued MRI images is not correlated, the general vector-valued normal distribution reduces to this equation with zero co-variances.

For mixtures along a boundary between two materials, we derive another equation similarly:

$$f_{\text{double}}(v; c, s) = \int_0^1 k_n((1 - t)c_1 + tc_2 - v; s) dt \tag{3}$$

As with the single material, this derivation follows from Equation 1 evaluated over a region where two materials mix. In this case, we approximate the band-limiting filter causing partial-volume effects with a box filter and make the assumption that the variance of the additive noise is constant across the region. This basis function is a superposition of normal distributions representing different amounts of the two constituent pure materials. k_n is the normal distribution, centered at zero, t the relative quantity of the second material, c_1 and c_2 the expected values of the two materials, and s the variance of measurements.

The assumption of a box filter affects the shape of the resulting histogram basis function. We derived similar equations for different filters (triangle, Gaussian, and Hamming), but chose the box filter derivation because we found it sufficiently accurate in practice and because its numerical tractability saves significant computation.

5 Estimating Histogram Basis Function Parameters

In this section we describe parameter-estimation procedures for fitting histogram basis functions to a histogram of an entire dataset. For a given dataset we first calculate the histogram, $h^{\text{all}}(v)$, of the entire dataset. We then combine an interactive process of specifying the number of materials and approximate

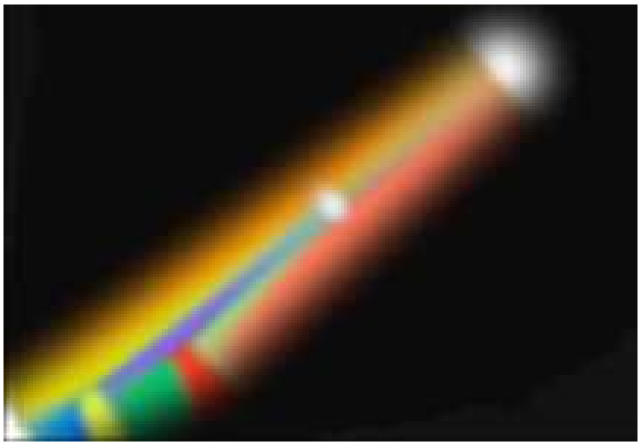


Figure 9: Basis functions fit to histogram of entire dataset. This figure illustrates the results of fitting basis functions to the histogram of the hand dataset. The five circular dots represent pure materials, while the lines connecting them represent mixtures. The rightmost two white dots are pure fat and bone marrow in the hand. The lower yellow and red dot are pure skin and muscle, respectively. The mixture between muscle (red) and fat (white) is a salmon-colored streak. The green streak between the red and yellow dots is a mixture of skin and muscle. These fitted basis functions were used to produce the classified data used in Figures 1 and 11.

---

feature-space locations for them with an automated optimization [Laidlaw, 1992] to refine the parameter estimates. Under some circumstances, users may wish to group materials with similar measurements into a single "material," whereas in other cases they may wish the materials to be separate. The result of this process is a set of parameterized histogram basis functions, together with values for their parameters. The parameters describe the various materials and mixtures of interest in the dataset. Figure 9 shows the results of fitting a histogram. Each colored region represents one distribution, with spots representing pure materials and connecting shapes representing mixtures.

To fit a groups of histogram basis functions to a histogram, as in Figure 9, the optimization process estimates the relative volume of each pure material or mixture (vector α^{all}), and the mean value (vector c) and variance (vector s) of measurements of each material. The process is derived from the assumption that all values were produced by pure materials and two-material mixtures. We define n_m as the number of pure materials in a dataset, and n_f as the number of histogram basis functions. Note that $n_f \geq n_m$, since n_f includes any basis functions for mixtures, as well as those for pure materials.

The optimization minimizes the function

$$\mathcal{E}(\alpha^{\text{all}}, c, s) = \int \left(\frac{q(v; \alpha^{\text{all}}, c, s)}{2w(v)} \right)^2 dv \tag{4}$$

where:

$$q(v; \alpha^{\text{all}}, c, s) = h^{\text{all}}(v) - \sum_{j=1}^{n_f} \alpha_j^{\text{all}} f_j(v; c_j, s_j) \tag{5}$$

Note that f_j may be a pure or a mixture basis function and that its parameter c will be a single feature-space point for a pure material or a pair for a mixture. The function $w(v)$ is analogous to a variance at each point, v, in feature space, and gives the expected value of $|q(v)|$. We approximate $w(v)$ as a constant, and discuss it further in Section 8.

This equation is derived in Appendix B using Bayesian probability theory with estimates of prior and conditional probabilities.

6 Classification

In this section we describe the process of classifying each voxel. This process is similar to that described in Section 5 for fitting the histogram basis functions to the entire dataset histogram, but now we are fitting histograms taken over small, voxel-sized regions. We use the previously computed histogram basis functions calculated from the entire dataset histogram and no longer vary the mean vector, c, and variance, s. The only parameters allowed to vary are the relative material volumes (vector α_j^{vox}), and an estimate of the local noise in the local region (vector $\bar{N}$) (see Equations 6 and 7).

Over large regions the noise is normally distributed, with zero mean; however, for small regions the mean noise is generally non-zero due to the band-limiting introduced in the data-collection process. We label this local mean voxel noise value $\bar{N}$. As derived in Appendix B the equation that we minimize is:

$$\mathcal{E}(\alpha^{\text{vox}}, \bar{N}) = \sum_{i=1}^{n_v} \left(\frac{\bar{N}_i}{2\sigma_i} \right)^2 + \int \left(\frac{q(v; \alpha^{\text{vox}}, \bar{N})}{2w(v)} \right)^2 dv \tag{6}$$

where

$$q(v; \alpha^{\text{vox}}, \bar{N}) = h^{\text{vox}}(v - \bar{N}) - \sum_{j=1}^{n_f} \alpha_j^{\text{vox}} f_j(v) \tag{7}$$

and the minimization is subject to the constraints

$$0 \leq \alpha_j^{\text{vox}} \leq 1, \text{ and } \sum_{j=1}^{n_f} \alpha_j^{\text{vox}} = 1.$$

Vector σ is the variance of the noise over the entire dataset. For MR data the variances in the signals for different materials are reasonably similar, and so we estimate σ to be an average of the variances of the histogram basis functions.

With vector α^{vox} for a given voxel-sized region and the mean value, vector $\bar{v}$, within that region, we solve for the amount of each pure material contributed by each mixture to the voxel. This is our output, the estimates of the amount of each pure material in the voxel-sized region.

$$\bar{v} = \int \left(h^{\text{vox}}(v) - \sum_{i=1}^{n_m} \alpha_i f_{\text{single}}(v) \right) dv \tag{8}$$

$\bar{v}$ contains the mean signature of the portion of the histogram that arises only from regions with partial-volume effects. We determine how much of each pure component of pairwise mixture materials would be needed to generate $\bar{v}$, given the amount of each mixture that α^{vox} indicates is in the voxel. t_k represents this relative amount for mixture k, with $t_k = 0$ indicating that the mixture is comprised of only the first pure component, $t_k = 1$ indicating that it is comprised only of its second component, and intermediate values of t_k indicating intermediate mixtures. The t_k values are calculated by minimizing the following equation subject to the constraint $0 \leq t_k \leq 1$.

$$\mathcal{E}_{\bar{v}}(t) = \left(\bar{v} - \sum_{k=n_m+1}^{n_f} \alpha_k \left(t_k c_{ka} + (1 - t_k) c_{kb} \right) \right)^2 \tag{9}$$

Vector c_{ka} is the mean value for the first pure material component of mixture k, and vector c_{kb} the mean value for the second component. The amount of each material is the amount of pure material added to the t_k-weighted portion of each mixture.

7 Results

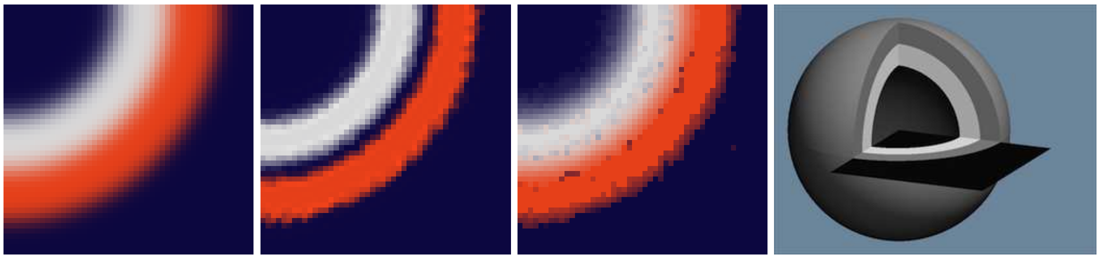
We have applied our new technique to both simulated and collected MRI datasets. When results can be verified and conditions are controlled, as with the results based on simulated data, the algorithm comes very close to "ground truth," or perfect classification. The results based on collected data illustrate that the algorithm works well on real data, with a geometric model of a tooth showing boundaries between materials, a section of a human brain showing classification results mapped on to colors, and a volume-rendered image of a human hand showing complex geometric relationships between different tissues.

This technique significantly reduces artifacts introduced by some other techniques at boundaries between materials. In Figures 2, 5, and 10 we compare our technique 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 partial-volume effects. In Figure 10(ii) and (iii) the difference is particularly noticeable where an incorrect layer of background material has been introduced between the two lighter regions, and where jagged boundaries occur between each pair of materials. In both cases this is caused by partial-volume effects, where multiple materials are present in the same voxel. Figures 2 and 5 also show comparative results between the two methods, and illustrate that the same artifacts occur with real data and are reduced by our technique.

Because of the lack of a "gold standard" against which classification algorithms can be measured, it is difficult to compare our technique with others. Each technique presents a set of results from some application area, and so anecdotal comparisons can be made, but quantitative comparisons cannot, in general. Work in generating a standard would greatly assist in the search for effective and accurate classification techniques. Our technique appears to achieve a given level of accuracy with fewer vector elements than the eigenimages

Object	Machine	Voxel Size	Figs.
shells	simulated	1x1x10 mm	10
hand	GE 1.5T	0.7x0.7x3 mm	1, 11
brain	GE 1.5T	0.8x0.8x3 mm	2, 5
tooth	Bruker AMX500	312x312x312 μm	12

Table 1: MRI dataset sources, resolutions, and figure references.



(i) "Perfect" Classification

(ii) Discrete Classification (no mixtures)

(iii) Continuous, Multi-material Classification

(iv) Slice Geometry

Figure 10: Comparison of discrete, single-material classification (ii), and the new classification (iii). (i) is a reference for what "ideal" classification should produce. Note the band of background material in (ii) between the two curved regions. This band is incorrectly classified, and could lead to errors in models or images produced from the classified data. The original dataset is simulated, two-valued data of two concentric shells, as shown in (iv).

of [Windham et al., 1988] or the classification results of [Choi et al., 1991]. Their results are visually similar to ours, and underscore the need for quantitative comparison. Because we interpolate neighboring sample values, we are able to work with two-valued or even scalar data, while their technique is likely to require more vector components. [Kao et al., 1996] shows good results for a human brain dataset, but we believe their technique will be less robust in the presence of material mixture signatures that overlap, a situation their examples do not include.

Models and volume-rendered images, as in Figures 1, 11, and 12, benefit from our new techniques because less incorrect information is introduced into the classified datasets, and so the images and models more accurately depict the objects they are representing. These models and images illustrate the underlying geometries, and are particularly sensitive to errors at geometric boundaries.

Figure 11: 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.

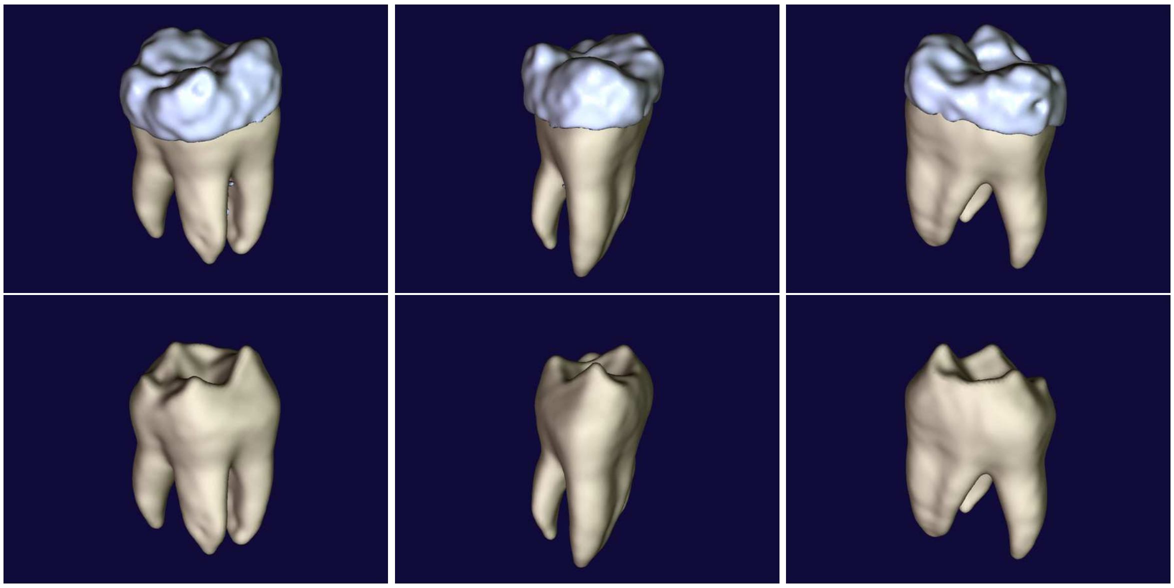


Figure 12: A geometric model of tooth dentine and enamel created by collecting MRI data using a technique that images hard solid materials [Ghosh et al., 1995] and classifying dentine and enamel in the volume data with our new partial-volume mixtures algorithm. Polygonal isosurfaces define the bounding 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.

Table1 lists the datasets, the MRI machine they were collected on, the protocol used (with some collection parameters), the voxel size, and the figures in which each dataset appears. All datasets except the tooth were collected with a spin-echo or fast spin-echo protocol, with one proton-weighted and one T_2-weighted acquisition. The tooth was acquired with a technique described in [Ghosh et al., 1995].

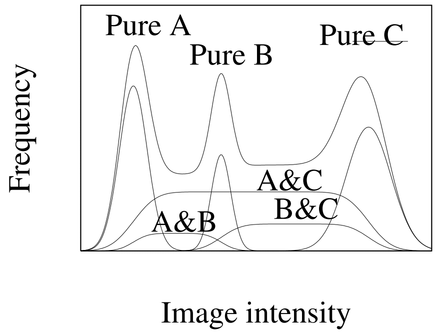
8 Discussion

We have made several assumptions and approximations while developing and implementing this algorithm. This section will discuss some of the tradeoffs and suggest some possible directions for further work.

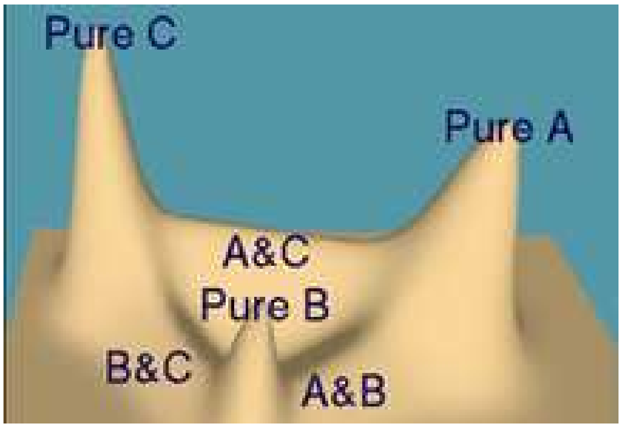
Mixtures of Three or More Materials. We assume that each measurement contains values from at most two materials. We chose two-material mixtures based on a dimensionality argument. In an object that consists of regions of pure materials, as shown in Figure 6, voxels containing one material will be most prevalent because they correspond to volumes. Voxels containing two materials will be next most prevalent, because they correspond to surfaces where two materials meet. As such, they are the first choice to model after those containing a single material. Our approach can be extended in a straightforward manner to handle the three-material case as well as cases with other less-frequent geometries, such as skin, tubes, or points where four materials meet.

Benefits of Vector-Valued Data. 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 improved signal-to-noise ratio. Second, they frequently distinguish similar materials more effectively (see Figure 13).

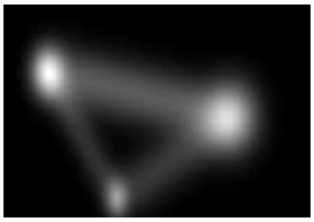
In particular, the jump from scalar to two-valued vector data is very important. In scalar-valued datasets it is difficult to distinguish a mixture of two pure materials with values v_A and v_B from a pure material with



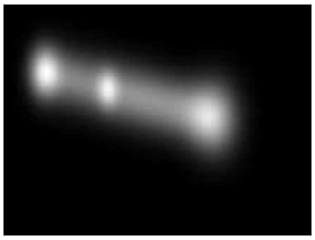
(i)



(ii)



(iii)



(iv)

Figure 13: Benefits of histograms of vector-valued data. We show histograms of an object with three materials. (i) This histogram of scalar data shows that material mean values are collinear. Distinguishing among more than two materials is often ambiguous. (ii) and (iii) are two representations of histograms of vector-valued data and show that mean values often move away from collinearity in higher dimensions, and so materials are easier to distinguish. While less likely, (iv) shows that the collinearity problem can exist with vector-valued data.

some intermediate value such as $v_C = (v_A + v_B)/2$. This is because the 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 materials, although Figure 13 illustrates that it can still occur. When it does occur, classification works as for scalar-valued data.

Why "Voxel-Sized" Regions? The regions that we classify could be smaller or larger than voxels. Smaller regions would include less information, and so the context for the classification would be reduced and accuracy would suffer. Larger regions would contain more complicated geometry because the features that could be represented would be smaller relative to the size of the region. Again, accuracy would suffer. Because the spacing of sample values is intrinsically related to the minimum feature size that the reconstructed continuous function, $\rho(x)$, can represent, that spacing is a natural choice for the size of regions that are classified.

Partial Mixtures. We note that the histograms, $h^{\text{vox}}(v)$, for some voxel-sized regions are not ideally matched by a linear sum of basis functions. We discuss two possible sources of this mismatch.

The first source is the assumption that within a small region we still have normally distributed noise. $\bar{N}$ models the fact that the noise no longer averages to zero, but we do not attempt to model the change in shape of the distribution as the region size shrinks.

The second source is related. A small region may not contain the full range of values that the mixture of materials can produce. As a result, the histogram over that small region is not modeled ideally by a linear combination of pure material and mixture distributions. We are investigating model histogram basis functions with additional parameters to better match histograms [Laidlaw, 1995], [Laidlaw et al., 1997]. Modeling the histogram shape as a function of the distance of a voxel from a boundary between materials is likely to address both of these effects and give a result with a physical interpretation that will make geometric model extraction more justifiable and the resulting models more accurate.

We postulate that these two effects weight the optimization process such that it tends to make $\bar{N}$ much larger than we expect. As a result, we have found that setting $w(v)$ to approximately 30 times the maximum value in $h^{\mathrm{vox}}(v)$ gives good classification results. Smaller values tend to allow $\bar{N}$ to move too much, and larger values hold it constant. Without these problems we would expect $w(v)$ to take on values equal to some small percentage of the maximum of $h^{\mathrm{vox}}(v)$.

Mixtures of Materials within an Object. Based on our assumptions, voxels only contain mixtures of materials when those mixtures are caused by partial-volume effects. These assumptions are not true in many cases. By relaxing them and then introducing varying concentrations of given materials within an object, one can derive histogram basis functions parameterized by the concentrations and can fit them to measured data. The derivation would be substantially similar to that presented here.

Non-uniform Spatial Intensities. Spatial intensity in MRI datasets can vary due to inhomogeneities in the RF or gradient fields. We assume that they are small enough to be negligible for our algorithm, but it would be possible to incorporate them into the histogram basis functions by making the parameters c and possibly s be functions of space. Deriving appropriate spatial functions and calibrating them is likely to be somewhat time consuming but may pay off with more-accurate classification results.

Implementation. Our implementation is written in C and C++ on Unix workstations. We use a sequential-quadratic-programming constrained-optimization algorithm [NAG, 1993] to fit h^{vox} for each voxel-sized region, and a quasi-Newton optimization algorithm for fitting h^{all}. The algorithm classifies approximately 10 voxels per second on a single HP9000/730, IBM RS6000/550E, or DEC Alpha AXP 3000 Model 500 workstation. We have implemented this algorithm in parallel on these machines and get a corresponding speedup on multiple machines.

9 Conclusions

Our new algorithm 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 neighbor information 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 the 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 continued 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, one of which we have demonstrated.

10 Acknowledgments

Many thanks to Matthew Avalos, who has been instrumental in implementation. Thanks to Barbara Meier, David Kirk, John Snyder, Bena Currin, and Mark Montague for reviewing early drafts and making suggestions. Thanks also to Allen Corcoran, Constance Villani, Cindy Ball, and Eric Winfree for production help, and Jose Jimenez for the late-night MR sessions. Our data was collected in collaboration with the Huntington Magnetic Resonance Center and the Caltech Biological Imaging Center, both in Pasadena.

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/ARPA STC for Computer Graphics and Scientific Visualization, by the DOE (DE-FG03-92ER25134) as part of the Center for Research in Computational Biology, by the National Institute on Drug Abuse and the National Institute of Mental Health as part of the Human Brain Project, and by the Beckman Institute Foundation. All opinions, findings, conclusions, or recommendations expressed in this document are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

Appendices

A Derivation of Histogram Basis Functions

In this appendix we derive parameterized model histograms that we use as basis functions, f_i, for fitting histograms of data. We derive two forms of basis functions: one for single, pure materials; another for two-material mixtures that arise due to partial-volume effects in sampling. Equation 1, the histogram equation, is:

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

Note that if $\rho(x)$ contains additive noise, $n(x; s)$, with a particular distribution, $k_n(v; s)$, then the histogram of ρ with noise is the convolution of $k_n(v; s)$ with $\rho(x) - n(x; s)$ (i.e, $\rho(x)$ without noise). $k_n(v; s)$ is, in general, a normal distribution. Thus,

$$h^{\mathcal{R}}(v) \quad = \quad k_n(v; s) * \int \mathcal{R}(x)\delta((\rho(x) - n(x; s)) - v)dx \tag{10}$$

A.1 Pure Materials

For a single pure material we assume that the measurement function has the form:

$$\rho_{\text{single}}(x) = c + n(x; s) \tag{11}$$

where c is the constant expected value of a measurement of the pure material, and s is the variance of additive, normally-distributed noise.

The basis function we use to fit the histogram of the measurements of a pure material is

$$
\begin{aligned}
f_{\text{single}}(v; c, s) &= \int \mathcal{R}(x)\delta(\rho_{\text{single}}(x) - v)dx \\
&= \int \mathcal{R}(x)\delta(c + n(x; s) - v)dx \\
&= k_n(v; s) * \int \mathcal{R}(x)\delta(c - v)dx \\
&= k_n(v; s) * \left(\delta(c - v) \int \mathcal{R}(x)dx\right) \\
&= k_n(v; s) * \delta(c - v) \\
&= k_n(v - c; s) \\
&= \left(\prod_{i=1}^{n_v} \frac{1}{s_i\sqrt{\pi}}\right) \exp\left(-\sum_{i=1}^{n_v}\left(\frac{v_i - c_i}{2s_i}\right)^2\right),
\end{aligned}
\tag{12}
$$

Thus, $f_{\text{single}}(v; c, s)$ is a Gaussian distribution with mean c and variance s. We assume the noise is independent in each element of vector-valued data, which for MRI appears to be reasonable.

A.2 Mixtures

For a mixture of two pure materials, we assume the measurement function has the form:

$$
\rho_{\text{double}}(x) = \ell_{\text{double}}(x; c_1, c_2) + n(x; s) \tag{13}
$$

where ℓ_{double} approximates the band-limiting filtering process, a convolution with a box filter, by interpolating the values within the region of mixtures linearly between c_1 and c_2, the mean values for the two materials.

$$
\ell_{\text{double}} = (1 - t)c_1 + tc_2 \tag{14}
$$

$$
\begin{aligned}
f_{\text{double}}(v; c, s) &= \int \mathcal{R}(x)\delta(\rho_{\text{double}}(x) - v)dx \\
&= \int \mathcal{R}(x)\delta(\ell_{\text{double}}(x; c_1, c_2) + n(x; s) - v)dx \\
&= k_n(v; s) * \int \mathcal{R}(x)\delta(\ell_{\text{double}}(x; c_1, c_2) - v)dx \\
&= \int_0^1 k_n(v; s) * \delta((1 - t)c_1 + tc_2 - v)dt \\
&= \int_0^1 k_n((1 - t)c_1 + tc_2 - v; s)dt
\end{aligned}
\tag{15}
$$

B Derivation of Classification Parameter Estimation

In this appendix we derive the equations that we optimize to find model histogram parameters and to classify voxel-sized regions. We use Bayesian probability theory [Loredo, 1989] to derive an expression for the probability that a given histogram was produced by a particular set of parameter values in our model. We maximize an approximation to this "posterior probability" to estimate the best-fit parameters.

$$\text{maximize } P(\text{ parameters } \mid \text{histogram }) \tag{16}$$

We use this optimization procedure for two purposes:

- **Find model histogram parameters**. Initially, we find parameters of basis functions to fit histograms of the entire dataset h^{all}. This gives us a set of basis functions that describes histograms of voxels containing pure materials or pairwise mixtures.

- **Classify voxel-sized regions**. We fit a weighted sum of the basis functions to the histogram of a voxel-sized region h^{vox}. This gives us our classification (in terms of the weights α).

The posterior probabilities P^{all} and P^{vox} share many common terms. In the following derivation we distinguish them only where necessary, using P where their definitions coincide.

B.1 Definitions

Table 3 lists Bayesian probability terminology as used in [Loredo, 1989] and in our derivations. Table 2 defines additional terms used in this section.

B.2 Optimization

We perform the following optimization to find the best-fit parameters:

$$\text{maximize } P(\alpha, c, s, \bar{N} | h) \tag{17}$$

Term	Dimensionality	Definition
n_m	scalar	number of pure materials
n_f	scalar	number of pure materials and mixtures
n_v	scalar	dimensions of measurement (feature space)
α	n_f	relative volume of each mixture and material within the region
c	$n_f \times n_v$	mean of material measurements for each material
s	$n_f \times n_v$	variance of material measurements (chosen by procedure discussed in Section 5) for each material
$\bar{N}$	n_v	mean value of noise over the region
p_{1-6}	scalars	arbitrary constants
$h^{\text{all}}(v)$	$\mathbf{R}^{n_v} \to \mathbf{R}$	histogram of an entire dataset
$h^{\text{vox}}(v)$	$\mathbf{R}^{n_v} \to \mathbf{R}$	histogram of a tiny, voxel-sized region

Table 2: Definitions of terms used in the derivations.

$P(\alpha, c, s, \bar{N}\|h)$	posterior probability (we maximize this)
$P(\alpha, c, s, \bar{N})$	prior probability
$P(h\|\alpha, c, s, \bar{N})$	likelihood
$P(h)$	global likelihood

Table 3: Probabilities, using Bayesian terminology from [Loredo, 1989].

With $P \equiv P^{\text{all}}$, we fit histogram basis function parameters $c, s, \alpha^{\text{all}}$ to the histogram of an entire dataset, $h^{\text{all}}(v)$. With $P \equiv P^{\text{vox}}$, we fit $\alpha^{\text{vox}}, \bar{N}$ to classify the histogram of a voxel-sized region, $h^{\text{vox}}(v)$.

B.3 Derivation of the posterior probability, $P(\alpha, c, s, \bar{N}|h)$

We start with Bayes' Theorem, expressing the posterior probability in term of the likelihood, the prior probability, and the global likelihood.

$$P(\alpha, c, s, \bar{N}|h) = \frac{P(\alpha, c, s, \bar{N})P(h|\alpha, c, s, \bar{N})}{P(h)} \tag{18}$$

Each of the terms on the right side is approximated below, using p_{1-6} to denote positive constants (which can be ignored during the optimization process).

Prior Probabilities. We assume that α, c, s, and $\bar{N}$ are independent, so

$$P(\alpha, c, s, \bar{N}) = P(\alpha)P(c, s)P(\bar{N}) \tag{19}$$

Because the elements of α represent relative volumes, we require that they sum to 1 and are positive.

$$P(\alpha) = \begin{cases} 0 & \text{if } \sum_{j=1}^{n_f} \alpha_j \neq 1 \\ 0 & \text{if } \alpha_j < 0 \text{ or } \alpha_j > 1 \\ p_1 & \text{(constant) otherwise} \end{cases} \tag{20}$$

We use a different assumption for $P(c, s)$ depending on which fit we are doing (h^{all} or h^{vox}). For fitting $h^{\text{all}}(v)$, we consider all values of c, s equally likely:

$$P^{\text{all}}(c, s) = p_6 \tag{21}$$

For fitting h^{vox}, the means and variances, c, s, are fixed at c^0, s^0 (the values determined by the earlier fit to the entire data set):

$$P^{\text{vox}}(c, s) = \delta(c - c^0, s - s^0) \tag{22}$$

For a small region, we assume that the noise vector, $\bar{N}$, has normal distribution with variance σ:

$$P^{\text{vox}}(\bar{N}) = p_2 e^{-\sum_{i=1}^{n_v} \left(\frac{\bar{N}_i}{2\sigma_i}\right)^2} \tag{23}$$

For a large region, the mean noise, $\bar{N}$, should be very close to zero; hence, $P^{\text{all}}(\bar{N})$ will be a delta function at $\bar{N} = 0$.

Likelihood. We approximate the likelihood, $P(h|\alpha, c, s, \bar{N})$, by analogy to a discrete normal distribution. We define $q(v)$ to measure the difference between the "expected" or "mean" histogram for particular $\alpha, c, s, \bar{N}$ and a given histogram $h(v)$:

$$q(v; \alpha, c, s, \bar{N}) = h(v - \bar{N}) - \sum_{j=1}^{n_f} \alpha_j f_j(v; c, s) \tag{24}$$

Now we create a normal-distribution-like function. $w(v)$ is analogous to the variance of q at each point of feature space:

$$P(h|\alpha, c, s, \bar{N}) = p_3 e^{-\int \left(\frac{q(v;\alpha,c,s,\bar{N})}{2w(v)}\right)^2 dv} \tag{25}$$

Global Likelihood. Note that the denominator of Equation 18 is a constant normalization of the numerator:

$$\begin{aligned} P(h) &= \int P(\hat{\alpha}, \hat{c}, \hat{s}, \hat{N}) P(h|\hat{\alpha}, \hat{c}, \hat{s}, \hat{N}) d\hat{\alpha} d\hat{c} d\hat{s} d\hat{N} && (26) \\ &= p_4 && (27) \end{aligned}$$

B.3.1 Assembly

Using the approximations discussed above, we arrive at the following expression for the posterior probability:

$$P(\alpha, c, s, \bar{N}|h) = p_5 P(\alpha) P(c, s) \exp\left(-\sum_{i=1}^{n_v} \left(\frac{\bar{N}_i}{2\sigma_i}\right)^2\right) \exp\left(-\int \left(\frac{q(v;\alpha, c, s, \bar{N})}{2w(v)}\right)^2 dv\right) \tag{28}$$

For fitting h^{all}, the mean noise is assumed to be zero, so maximizing equation 28 is equivalent to minimizing $\mathcal{E}^{\text{all}}$ to find the free parameters $(\alpha^{\text{all}}, c, s)$:

$$\mathcal{E}^{\text{all}}(\alpha^{\text{all}}, c, s) = \int \left(\frac{q(v;\alpha^{\text{all}}, c, s)}{2w(v)}\right)^2 dv \tag{29}$$

subject to $P(\alpha^{\text{all}}) \neq 0$.

For fitting h^{vox}, the parameters c and s are fixed, so maximizing equation 28 is equivalent to minimizing $\mathcal{E}^{\text{vox}}$ to find the free parameters $(\alpha^{\text{vox}}, \bar{N})$:

$$\mathcal{E}^{\text{vox}}(\alpha^{\text{vox}}, \bar{N}) = \sum_{i=1}^{n_v} \left(\frac{\bar{N}_i}{2\sigma_i}\right)^2 + \int \left(\frac{q(v;\alpha^{\text{vox}}, \bar{N})}{2w(v)}\right)^2 dv \tag{30}$$

subject to $P(\alpha^{\text{vox}}) \neq 0$.

As stated in Equation 6, Section 6, Equation 30 is minimized to estimate relative material volumes, α^{vox}, and the mean noise vector, $\bar{N}$.

References

[Choi et al., 1991] Choi, H. S., Haynor, D. R., and Kim, Y. M. (1991). Partial volume tissue classification of multichannel magnetic resonance images — a mixel model. *IEEE Transactions on Medical Images*, 10(3):395–407.

[Clarke et al., 1995] Clarke, L. P., Velthuizen, R. P., Camacho, M. A., Neine, J. J., Vaidyanathan, M., Hall, L. O., Thatcher, R. W., and Silbiger, M. L. (1995). MRI segmentation: Methods and applications. *Magnetic Resonance Imaging*, 13(3):343–368.

[Cline et al., 1990] Cline, H. E., Lorensen, W. E., Kikinis, R., and Jolesz, F. (1990). Three-dimensional segmentation of MR images of the head using probability and connectivity. *Journal of Computer Assisted Tomography*, 14(6):1037–1045.

[Drebin et al., 1988] Drebin, R. A., Carpenter, L., and Hanrahan, P. (1988). Volume rendering. In Dill, J., editor, *Computer Graphics (SIGGRAPH '88 Proceedings)*, volume 22, pages 65–74.

[Duda and Hart, 1973] Duda, R. P. and Hart, P. E. (1973). *Pattern Classification and Scene Analysis*. John Wiley and Sons, New York.

[Ghosh et al., 1995] Ghosh, P. R., Laidlaw, D. H., Fleischer, K. W., Barr, A. H., and Jacobs, R. E. (1995). Pure phase-encoded MRI and classification of solids. *IEEE Transactions on Medical Imaging*, 14(3):616–620.

[Kao et al., 1996] Kao, Y.-H., Sorenson, J. A., and Winkler, S. S. (1996). MR image segmentation using vector decompsition and probability techniques: A general model and its application to dual-echo images. *Magnetic Resonance in Medicine*, 35:114–125.

[Laidlaw, 1992] Laidlaw, D. H. (1992). Tissue classification of magnetic resonance volume data. Master's project, California Institute of Technology.

[Laidlaw, 1995] Laidlaw, D. H. (1995). Geometric model extraction from magnetic resonance volume data. Technical Report CS-TR-95-05, Caltech.

[Laidlaw et al., 1997] Laidlaw, D. H., Barr, A. H., and Jacobs, R. E. (1997). *Goal-directed Brain Micro-imaging*, volume 1, chapter 6. *in press*.

[Levoy, 1988] Levoy, M. (1988). Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(3):29–37.

[Loredo, 1989] Loredo, T. J. (1989). From Laplace to supernova SN1987A: Bayesian inference in astrophysics. In Fougere, P., editor, *Maximum Entropy and Bayesian Methods*. Kluwer Academic Publishers, Denmark.

[Lorensen and Cline, 1987] Lorensen, W. E. and Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. In Stone, M. C., editor, *Computer Graphics (SIGGRAPH '87 Proceedings)*, volume 21, pages 163–169.

[NAG, 1993] NAG (1993). *NAG Fortran Library*. Numerical Algorithms Group, 1400 Opus Place, Suite 200, Downers Grove, Illinois 60515.

[Ney et al., 1990] Ney, D. R., Fishman, E. K., Magid, D., and Drebin, R. A. (1990). Volumetric rendering of computed tomography data: Principles and techniques. *IEEE Computer Graphics and Applications*, 10(2):24–32.

[Oppenheim et al., 1983] Oppenheim, A. V., Willsky, A. S., and Young, I. T. (1983). *Signals and Systems*. Prentice-Hall, Inc., New Jersey.

[Santago and Gage, 1993] Santago, P. and Gage, H. D. (1993). Quantification of MR brain images by mixture density and partial volume modeling. *IEEE Transactions on Medical Imaging*, 12(3):566–574.

[Vannier et al., 1985] Vannier, M. W., Butterfield, R. L., Jordan, D., Murphy, W. A., Levitt, R. G., and Gado, M. (1985). Multispectral analysis of magnetic resonance images. *Radiology*, 154:221–224.

[Vannier et al., 1988] Vannier, M. W., Speidel, C. M., and Rickman, D. L. (1988). Magnetic resonance imaging multispectral tissue classification. In *Proc. Neural Information Processing Systems (NIPS)*.

[Windham et al., 1988] Windham, J. P., Abd-Allah, M. A., Reimann, D. A., Froelich, J. W., and Haggar, A. M. (1988). Eigenimage filtering in MR imaging. *Journal of Computer Assisted Tomography*, 12(1):1–9.

[Wu et al., 1988] Wu, Z., Chung, H.-W., and Wehrli, F. W. (1988). A bayesian approach to subvoxel tissue classification in NMR microscopic images of trabecular bone. *Journal of Computer Assisted Tomography*, 12(1):1–9.