A Bayesian Similarity Function for Segmentation using Anatomical, Shape-Based Models

FMRIB Technical Report TR05MJ1

Mark Jenkinson

Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB),
Department of Clinical Neurology, University of Oxford, John Radcliffe Hospital,
Headley Way, Headington, Oxford, UK

Abstract

Shape-based segmentation involves fitting a flexible model of anatomical shape to a measured image. It is important to be able to utilise probabilistic prior information about shape, and to combine this with data-driven likelihoods. This is naturally achieved within the Bayesian framework. In this paper a probabilistic similarity function for the anatomical, shape-based segmentation problem is derived using a fully Bayesian approach. Furthermore, it incorporates prior, probabilistic information *without* the need for additional ad-hoc parameters. Preliminary results show that this similarity function is more robust and accurate than simpler versions based on the same image formation model.

1 Introduction

Identifying anatomical structures of interest in medical images is of major importance in many areas, especially neuroimaging (e.g. [4, 2]). Due to poor contrast to noise it is advantageous to include prior information about structures as seen in the population rather than relying purely on the image intensities. However, a similarity measure is needed to combine prior information about structures, and their variation, with the image data.

Existing applications of model-based segmentation rely on explicit or implicit image similarity metrics, prior information and regularisation terms (often expressed as forces in a deformable model) which often require empirically set parameters to weight these terms (e.g. [4, 2]). Using a fully Bayesian approach naturally combines probabilistic prior shape information with a model of the image formation process in a probabilistic framework, without the need for additional, ad-hoc parameters to control the relative strength of the prior and data-driven information.

Previous Bayesian derivations of similarity functions for registration applications exist (see [5, 7]). This formulation is different because it is based on segmentation via fitting shape models, not matching two arbitrary images without any model of the image content. In addition to incorporating anatomical knowledge (via shape priors), this approach models the effects of partial volume, bias fields and changes in field of view. A somewhat similar approach was recently taken in [1], but for voxel-based tissue-type classification without the use of shape models.

The similarity function derived in this report is similar to the Correlation Ratio [6] but uses the residuals from an image formation model together with terms that normalise for the amount of partial volume and field of view of the image acquisition. The performance of the similarity function is compared to several existing functions to demonstrate the advantages of the new formulation.

2 Problem Formulation

Consider an image, Y, generated by some image generation process, G, from a known (ground truth) object, S, where Y is not spatially aligned with S, but related spatially by a transformation, T. The objective of the segmentation/registration problem is to recover the spatial transformation, T, which relates S to Y. That is, find T such that $G(T(S))$ and Y are '*most similar*'.

A Bayesian formulation of this problem is as follows:

- $p(Y|T, S, \theta)$ is the image formation/likelihood model where θ are parameters of the generation process (e.g. noise variance, tissue intensities).
- $p(T, \theta|Y, S)$ is the joint posterior probability distribution for the spatial transformation (the quantity of interest) and the unknown generation parameters.
- $p(T|Y, S)$ is the marginal posterior distribution, which integrates over the unknown parameters, θ.

Note that generally T is parameterised by its own set of parameters, and that $p(T|Y, S)$ is giving the posterior probability for these transformation parameters.

The probabilities are related by:

$$\begin{aligned} p(T, \theta|Y, S) &\propto p(Y|T, S, \theta)\, p(T, \theta|S) && (1)\\ &\propto p(Y|T, S, \theta)\, p(T|S)p(\theta|S) && (2)\\ &\propto p(Y|T, S, \theta)\, p(T)p(\theta) && (3) \end{aligned}$$

and

$$p(T|Y, S) = \int p(T, \theta|Y, S)\, d\theta \quad (4)$$

where $p(T)$ is the prior probability distribution for the transformation, T, and $p(\theta)$ is the prior for the image formation parameters, θ. It is assumed that T and S are independent, so that $p(T|S) = p(T)$. Similarly for θ and S. The constant of proportionality, $p(Y|S)$, does not depend on T and hence will be ignored for the remainder of this report.

Note that marginalising over θ is the difficult step in calculating $p(T|Y, S)$.

The problem of finding the '*best*' single segmentation/registration[1] is then equivalent to finding the maximum a-posterior probability: the MAP estimate.

That is:

$$T_{MAP} = \arg\max_T p(T|Y, S) = \arg\max_T \log\left(p(T|Y, S)\right)$$

Therefore $p(T|Y, S)$, or $\log(p(T|Y, S))$, play an equivalent role to the *similarity function* in common registration techniques.

2.1 Image Formation Model

- **Shape model:** S consists of K shapes: S_k , $k = 1, \ldots, K$. (e.g. ventricles, putamen, hippocampus, ...)
- **Transformation:** T can be any spatial transformation within a specified set. The set can include non-linear warps with arbitrarily large number of Degrees Of Freedom. This transformation applies to all of the shapes, S_k.
- **Image generation model:**

  $$G(T(S)) = \sum_{m,k} \alpha_{m,k} G_m(T(S_k))$$

  where $\alpha_{m,k}$ are intensity scaling parameters and $G_m(\cdots)$ generates the m th basis image of a shape (see section 2.3).

[1] and not a distribution of possible registrations with appropriate posterior probabilities

- **Image measurement model:**

$$Y = G(T(S)) + \epsilon$$

where ϵ is a spatially uncorrelated Gaussian noise process with $E\{\epsilon\} = 0$, $\mathrm{Cov}\{\epsilon\} = \sigma^2 I$.

Note that all images (ϵ, $G_m(T(S_k))$, Y, etc) are treated as vectors of length N, where N is the number of voxels in the image.

2.2 Anatomical Information, Intensities and Bias Field

For a model-based segmentation method, the anatomical information is encoded in a model of the anatomical shapes. This is represented by S in this formulation and in practice would take the form of something like a mesh model of the shapes of interest (e.g. ventricles and deep brain structures for our neuroimaging application).

The measured image, Y, is related to the shape model in two ways: by a spatial transformation, T; and by an intensity generation model, G. By adding prior information to what spatial transformations are more likely, the typical variation in the shapes across a population can be encoded. That is, $p(T)$, encodes the information about modes of shape variation (typically by using a multi-variate Gaussian).

The intensity generation model relates the (noise-free) intensity values to the spatially transformed shapes. This generation model includes information about how partial volume is generated during imaging.

For MRI, the actual intensity values are related in a complicated way to the tissue parameters, which is controlled by the pulse sequence, allowing many different image 'contrasts' to be formed. Consequently, without knowledge of the pulse sequence details, the intensities within each shape can take on a wide range of possible values, and so are modelled by independent parameters, $\alpha_{m,k}$.

In addition, MR images often contain bias field. This is an effect of inhomogeneities in the RF field, and usually effect both transmit and receive signals. The effect is to have a slowly varying intensity change across the image.[2] Often this is modelled as a slowly varying multiplicative change, but because the bias field also effects the trasmitted RF, there can be sharp changes between its effect at tissue boundaries, due to differing flip angles causing changes in the steady-state magnetisation. Therefore we choose to model this effect by including a linear spatially-varying intensity for each shape, which is an adequate approximation for shapes of small spatial extent relative to the effective wavelength of the bias field, and allows both effects of the bias field to be incorporated. Furthermore, such linear terms also model underlying spatial variations in the tissue parameters, which can exist (e.g. changing density within the thalamus) and give rise to spatially intensity changes in the image which are indistinguishable from bias field effects.

2.3 Image Generators

The definitions of the image generators are:

- G_1 gives an image with an intensity 1.0 for all voxels totally within the shape, 0.0 for those outside and the partial volume overlap fraction otherwise;
- G_2 gives an image with an intensity gradient along the x-axis within the shape, with zero values outside. The image is also demeaned and has partial volume modelled multiplicatively (i.e. $G_2(x, y, z) = xG_1(x, y, z)$, which is then demeaned);
- G_3 and G_4 give images with intensity gradients along the y- and z-axes respectively;
- G_5 and above can represent other expected image signals (e.g. quadratic trends, common artifacts, etc.) but will not be used in this report

Note that G_2, G_3 and G_4 represent the terms that model the combined effects of bias field and spatially varying tissue parameters. These are effectively the first terms in the Taylor series expansion of the general case for

[2] except in very high field systems

bias field and spatially varying parameters. For shapes with large spatial extent this approximation may not be sufficiently accurate, in which case it is possible to include higher-order Taylor series terms (quadratic, cubic, etc.) via extra G terms (G_5 and up).

These generators are functions, which when applied to a transformed shape, $T(S_k)$, produce an image containing N values. The images are then reshaped into vectors of length N, and assembled into a single G matrix for use in later sections.

The image generator can also be written in matrix form:

$$G(T(S)) = \sum_{m,k} \alpha_{m,k} G_m(T(S_k)) \quad (5)$$

$$= G\alpha \quad (6)$$

where G represents an N by D matrix and α is a column vector of length D (typically $D = MK$, but will be changed – see later). Note that in this form G implicitly encodes information about the transformation T and the underlying shapes, S_k. However, as neither of these will be marginalised, this dependency is not important for any of the following derivations and will therefore be left implicit.

2.3.1 Example

Consider a toy example in 1D, as shown in figure 1. In the example there are two shapes, S_1 and S_2, each of different size (length) which results in the generation of four different G vectors: $G_1(S_1)$ to $G_2(S_2)$. Note that at the interface between the shapes, the values of G_1 are between 0.0 and 1.0, representing the partial volume fraction. The transformation, T, effects both the global position (and stretching) of the shapes in the image as well as the partial volume fractions.

2.4 Probabilistic Forms

Likelihood is:

$$p(Y|T,S,\theta) = p\left(\epsilon = Y - G(T(S))\right)$$

$$= \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left(\frac{-\beta\|Y - G(T(S))\|^2}{2}\right)$$

Priors:

$$p(\beta) = \frac{1}{\beta}$$

For α there are two alternative priors that are useful.

1. Flat prior:

$$p(\alpha_j) = \begin{cases} C_1 & \text{for } 0 \le \alpha_j \le C_1^{-1} \\ 0 & \text{otherwise} \end{cases}$$

   where the range of α_j is restricted to $[0, C_1^{-1}]$.

2. Or a prior encoding prior knowledge:

$$p(\alpha|\lambda,Q) = |\det(Q)|^{1/2} \left(\frac{\lambda}{\pi}\right)^{D/2} \exp\left(-\lambda\alpha^\mathsf{T} Q\alpha\right)$$

$$p(\lambda) = C_0$$

   where Q represents the prior knowledge about the expected intensities in the image formation; λ is a parameter expressing the unknown scaling between the learnt prior distribution of α parameters and the intensities in a new image; and C_0 is a constant, representing the (improper) flat prior on λ.

Parameters:

$$\theta = \{\beta, \alpha, \lambda\}$$

where $\beta = \frac{1}{\sigma^2}$ is the precision parameter.

3 Derivation of Similarity Function

Posterior probability is:

$$\begin{aligned} p(T|Y,S,Q) &\propto \int p(Y|T,S,\beta,\alpha)p(T)p(\alpha|\lambda,Q)p(\lambda)p(\sigma)\, d\sigma\, d\lambda\, d\alpha \\ &\propto p(T)\int \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left(\frac{-\beta\|Y-G\alpha\|^2}{2}\right) p(\alpha|\lambda,Q)p(\lambda)p(\beta)\, d\beta\, d\lambda\, d\alpha \end{aligned}$$

Note that Q is considered a known parameter and not marginalised.

3.1 Intensity Parameters: α

This posterior form above can be integrated in the case where *all* the parameters, α, are associated with a column in G that is linearly independent and has a norm greater than one. However, this is not commonly the case, and the set of parameters must be split into subsets. These subsets (see figure 2) are: *interesting* (int), for shapes inside the valid field of view (i.e. contained in the observed image); *null* (null), for shapes totally outside the field of view; *uninteresting* (un), for voxels inside the field of view but not within a modelled shape (i.e. in 'areas of no interest'); and *partial volume* (pv), for shapes that overlap the field of view by less than one voxel. That is:

$$\alpha = \{\alpha_{in}, \alpha_{un}, \alpha_{null}, \alpha_{deg}, \alpha_{pv}\}$$

The dimensionality of the subsets will be denoted as: $D = \dim(\alpha), D_{in} = \dim(\alpha_{in}), D_{un} = \dim(\alpha_{un}), D_{null} = \dim(\alpha_{null}), D_{deg} = \dim(\alpha_{deg}), D_{pv} = \dim(\alpha_{pv})$ such that $D = D_{in} + D_{un} + D_{null} + D_{deg} + D_{pv}$. Note that the composition and dimension of these subsets depends heavily on the spatial transformation, T. Hence, like G, they implicitly encode dependence on T, but as T is not being marginalised, this dependence will be left implicit. It is important to remember, however, that these are not constants.

Also, many of these subsets may be empty (zero dimensional) for certain spatial transformations, T, or always empty in certain applications (e.g. where there are no areas of no interest). The way that these subsets are defined and how they are treated in the integrations will be the topic of the next sections.

3.1.1 Null Parameters and Changing FOV

Often the model of the underlying object is larger than the portion contained in the field of view of the image. In this case, the shapes S_k that are outside the field of view will not appear in the image. This leads to zero columns in G (the null space), and the parameters associated with these columns are the null parameters, α_{null}, which can be easily marginalised (integrated over).

3.1.2 Areas of No Interest

Some areas of the underlying object that are known to give rise to signal may not be explicitly modelled (e.g. cortex if only deep brain structures are of interest fot the modelling). In this case it is necessary to include these 'areas of no interest' in the model with sufficient flexibility so that they can match any observed image intensities well without adversely affecting the fit of the model parameters of interest.

One way to model these areas is to include them in the model as very small, independent subshapes. For example, if the deep brain structures were of interest, then these would cover large areas of the image with relatively few parameters (e.g. mean intensity and intensity gradients) while the areas outside this can be

tesselated with small (say 0.1mm^3 sized) voxels (subshapes), each with an independent mean intensity. See figure 3 for a 1D example. With this model, whatever the intensities were in these areas in the measured image, Y, the extra subshape parameters would fit exactly (i.e. there would be no error, or residuals, in these areas, which could otherwise degrade the overall image fit).

In terms of the model above, this would introduce three types of extra parameters: extra null parameters, parameters of no interest and also degenerate parameters. The extra null parameters α_{null} are those associated with all subshapes that are outside the field of view of the measured image. The parameters associated with voxels inside the field of view comprise both parameters of no interest, α_{un}, and degenerate parameters, α_{deg}. The degeneracy occurs because typically many small subshapes will affect the same voxel, and no others, leading to columns in G that are identical. For marginalisation, only a single parameter (at most) can be associated with each voxel in the area of no interest, and so most subshapes will be degenerate and can, with appropriate reparameterisation, be treated in the same way as null parameters.

All parameters associated with areas of no interest are taken to have a flat prior, to enable them to match any intensity with equal probability.

3.1.3 Pure Partial Volume Parameters

For some spatial transformations, T, a shape of interest (or a parameter of no interest) can overlap the field of view by less than single voxel. This leads to a column in G which has a single non-zero value and a very low norm. These parameters are called pure partial volume parameters here, α_{pv}, and require special care in the integration, especially in the case of flat priors, where some simplifying assumptions breakdown.

3.2 Flat Intensity Prior

Initially consider the case of a flat prior on α. In this case let the range of each individual α be 0 to $L = C_1^{-1}$ such that, then $p(\alpha) = C_1^D$. When included, α parameters associated with linear intensity gradients have the appropriate columns of G scaled so that the range of α is $-C_1^{-1}/2$ to $+C_1^{-1}/2$, and $p(\alpha) = C_1^D$ is still true. Note that C_1 is a constant for all α, and represents the inverse intensity range.

To start with, take the case where there are no uninteresting, degenerate or partial volume parameters. Note that there may be null parameters. In these conditions, the posterior can be simply calculated using the integrals in appendix A, giving

$$
\begin{aligned}
p(T|Y,S) &\propto p(T)\int\left(\frac{\beta}{2\pi}\right)^{N/2}\exp\left(\frac{-\beta(Y-G\alpha)^{\mathsf{T}}(Y-G\alpha)}{2}\right)p(\beta)p(\alpha)\,d\beta\,d\alpha \\
&\propto p(T)\int\left(\frac{\beta}{2\pi}\right)^{N/2}\exp\left(\frac{-\beta(Y-G\alpha)^{\mathsf{T}}(Y-G\alpha)}{2}\right)\frac{1}{\beta}C_1^D\,d\beta\,d\alpha \\
&\propto p(T)C_1^D\int\left(\frac{\beta}{2\pi}\right)^{N/2}\exp\left(\frac{-\beta(Y-G\alpha)^{\mathsf{T}}(Y-G\alpha)}{2}\right)\frac{1}{\beta}\,d\beta\,d\alpha_{null}\,d\alpha_{in} \\
&\propto p(T)C_1^D C_1^{-D_{null}}\int\left(\frac{\beta}{2\pi}\right)^{N/2}\exp\left(\frac{-\beta(Y-G_{in}\alpha)^{\mathsf{T}}(Y-G_{in}\alpha)}{2}\right)\frac{1}{\beta}\,d\beta\,d\alpha_{in} \\
&\propto p(T)C_1^{D-D_{null}}\int\left(\frac{\beta}{2\pi}\right)^{N/2}|\det(G_{in}^{\mathsf{T}}G_{in})|^{-1/2}\left(\frac{2\pi}{\beta}\right)^{D-D_{in}/2}\exp\left(\frac{-\beta Y^{\mathsf{T}}R_wY}{2}\right)\frac{1}{\beta}\,d\beta \\
&\propto p(T)C_1^{D_{in}}\,(2\pi)^{-(N-D_{in})/2}\,|\det(G_{in}^{\mathsf{T}}G_{in})|^{-1/2}\int\beta^{(N-D_{in}-2)/2}\exp\left(\frac{-\beta Y^{\mathsf{T}}R_wY}{2}\right)d\beta \\
&\propto p(T)\,C_1^{D_{in}}\,(2\pi)^{-(N-D_{in})/2}\,|\det(G_{in}^{\mathsf{T}}G_{in})|^{-1/2}\,\Gamma\left(\frac{N-D_{in}-2}{2}\right)\left(\frac{Y^{\mathsf{T}}R_wY}{2}\right)^{-(N-D_{in})/2} \\
&\propto p(T)\,C_1^{N}\,|\det(G_{in}^{\mathsf{T}}G_{in})|^{-1/2}\,(\pi)^{-(N-D_{in})/2}\,\Gamma\left(\frac{N-D_{in}-2}{2}\right)\left(C_1^2\,Y^{\mathsf{T}}R_wY\right)^{-(N-D_{in})/2} \qquad (7)
\end{aligned}
$$

where $D = D_{in} + D_{null}$, $\int_0^L d\alpha_{null} = L^{D_{null}} = C_1^{-D_{null}}$, $R_w = I - G_{in}(G_{in}^{\mathsf{T}}G_{in})^{-1}G_{in}^{\mathsf{T}}$, and G_{in} is the submatrix formed from the G matrix by only including columns associated with α_{in} parameters. That is, $G = [G_{in}\ G_{null}]$ where $\alpha^{\mathsf{T}} = [\alpha_{in}^{\mathsf{T}}\ \alpha_{null}^{\mathsf{T}}]$.

Note that D_{in} and G_{in} both depend on the transformation T. In fact, the dependence on D_{in} is a form of normalisation for the number of degrees of freedom in the model. Also note that $D_{in} < N$ in all cases so that $N - D_{in} > 0$ and that increasing the normalised residuals $C_1^2 Y^\mathsf{T} R_w Y$ causes the posterior probability to decrease, as desired.

The above integrations use the approximation that

$$\int_0^L \exp\left(\frac{-\beta(Y-G\alpha)^\mathsf{T}(Y-G\alpha)}{2}\right) d\alpha_j \approx \int_{-\infty}^{\infty} \exp\left(\frac{-\beta(Y-G\alpha)^\mathsf{T}(Y-G\alpha)}{2}\right) d\alpha_j$$

which is a good approximation when $\beta Y^\mathsf{T} G_j \gg 1$ and $\beta L G_j^\mathsf{T} G_j \gg 1$. This is true for α_j parameters where G_j (the associated column of G) has a norm greater than 1.0 (i.e. $G_j^\mathsf{T} G_j > 1$), and the true value of α is not within $1/\beta$ of the limits of the range (which can be easily ensured by slightly increasing the prior range). Note that $\beta Y_j^\mathsf{T} Y_j = SNR^2$, which is the squared Signal to Noise Ratio in voxel j and typically much greater than 1.0 in MR images.

When these conditions are not true, the above approximation breaks down and the complimentary error function erfc terms, as shown in equation 14 in appendix A, must be included. This is true for pure partial volume parameters, and will be treated in section 3.2.2 which will be the last parameters to be integrated over.

3.2.1 Marginalising over Areas of No Interest

The parameters in the area of no interest can initially be divided into those in the null space of G (associated with zero columns) and the remainder. All null parameters can be treated in the same way as null parameters formed from shapes of interest, and therefore need not be considered separately.

The remaining parameters all contribute towards the $G\alpha$ term. For small subshapes, each parameter contributes a small partial volume amount to a single image voxel. Consider the image voxel j in the area of no interest, which is modelled by N_0 subshapes of equal size. Therefore each subshape contributes a partial volume amount of $\chi = 1/N_0$ and is associated with a column in G of the form: $G_j = [0 \cdots 0\ \chi\ 0 \cdots 0]^\mathsf{T}$. Such a set of subshapes contains 1 effective parameter of no interest and $N_0 - 1$ effective degenerate parameters.

To see the degeneracy more clearly it is necessary to transform this set of parameters. Let this set of subshape parameters be α_{sv} with G_{sv} associated columns from G. Furthermore, define a transformation $\alpha_o = K^{-1}\alpha_{sv}$ where K is an N_0 by N_0 matrix such that $G_{sv}\alpha_{sv} = G_{sv}KK^{-1}\alpha_{sv} = G_o\alpha_o$ where $G_o = G_{sv}K$. These matrices are

$$G_{sv} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ \chi & \chi & \cdots & \chi \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \quad K = \begin{bmatrix} 1 & k_{1,2} & \cdots & k_{1,N_0} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & k_{m-1,2} & \cdots & k_{m-1,N_0} \\ 1 & k_{m,2} & \cdots & k_{m,N_0} \\ 1 & k_{m+1,2} & \cdots & k_{m+1,N_0} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & k_{N_0,2} & \cdots & k_{N_0,N_0} \end{bmatrix} \quad G_o = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

where the columns of K form an orthogonal basis (i.e. are mutually orthogonal) and the first column of K has a norm of $\sum_i k_{i,1}^2 = N_0$, with all other columns normalised so that $\sum_i k_{i,j}^2 = N_0^{-1/(N_0-1)}$. This normalisation ensures that $|\det(K)| = 1$. This reparameterisation expresses the voxel intensity in terms of a mean across all subshapes plus a set of demeaned fluctuations (e.g. a Fourier series of sinusoidal terms) that express the spatial variability which cannot be seen in the final image due to the summation over the voxel.

It is clear from this that the first parameter in the transformed set, α_o, associated with the first column of G_o, is the effective parameter of no interest and that all the others are null parameters. Consequently, the integration over this set of subshape parameters can be achieved using this reparameterisation. Note that this integration can be done separately for each measured voxel in the area of no interest, since the contributions to separate voxels are orthogonal in both the prior and posterior, which is the case for flat priors.

Applying this to the previously derived posterior gives:

$$
\begin{aligned}
p(T|Y,S) &\propto p(T)C_1^D \int \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left(\frac{-\beta(Y-G\alpha)^\mathsf{T}(Y-G\alpha)}{2}\right) \frac{1}{\beta}\, d\beta\, d\alpha_{null}\, d\alpha_{in}\, d\alpha_{un}\, d\alpha_{deg} \\
&\propto p(T)C_1^{D-D_{null}} \int \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left(\frac{-\beta(Y-G\alpha)^\mathsf{T}(Y-G\alpha)}{2}\right) \frac{1}{\beta}\, d\beta\, d\alpha_{in}\, d\alpha_{un}\, d\alpha_{deg} \\
&\propto p(T)C_1^{D-D_{null}-D_{deg}} \int \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left(\frac{-\beta(Y-G\alpha)^\mathsf{T}(Y-G\alpha)}{2}\right) \frac{1}{\beta}\, d\beta\, d\alpha_{in}\, d\alpha_{un} \\
&\propto p(T)\, C_1^{D-D_{null}-D_{deg}} \,|\det(G_{in}^\mathsf{T}G_{in})|^{-1/2}\, (2\pi)^{-(N-D_{in}-D_{un})/2} \,\times \\
&\qquad \int \beta^{(N-D_{in}-D_{un}-2)/2} \exp\left(\frac{-\beta Y^\mathsf{T}R_wY}{2}\right)^{-(N-D_{in}-D_{un})/2} \frac{1}{\beta}\, d\beta \\
&\propto p(T)\, C_1^N \,|\det(G_{in}^\mathsf{T}G_{in})|^{-1/2}\, (\pi)^{-(N-D_{in}-D_{un})/2}\, \Gamma\left(\frac{N-D_{in}-D_{un}-2}{2}\right) \times \\
&\qquad \left(C_1^2\, Y^\mathsf{T}R_wY\right)^{-(N-D_{in}-D_{un})/2}
\end{aligned} \tag{8}
$$

where $D_{un} = \dim(\alpha_{un})$ is equal to the number of measured voxels wholly inside the area of no interest; $R_w = I - G_{in}(G_{in}^\mathsf{T}G_{in})^{-1}G_{in}^\mathsf{T} - G_{un}(G_{un}^\mathsf{T}G_{un})^{-1}G_{un}^\mathsf{T}$, and the G submatrices are defined such that $G = [G_{in}\ G_{un}\ G_{null}\ G_{deg}]$ with $\alpha^\mathsf{T} = [\alpha_{in}^\mathsf{T}\ \alpha_{un}^\mathsf{T}\ \alpha_{null}^\mathsf{T}\ \alpha_{deg}^\mathsf{T}]$.

The expression for R_w is derived using that fact that the columns of G_{in} and G_{un} are all mutually orthogonal – that is, $G_{un}^\mathsf{T}G_{in} = 0$. This is true since G_{un} only have non-zero entries in voxels inside the area of no interest.

This expression for the posterior is exactly the same form as previously derived (equation 7) without having an area of no interest. The difference between the two forms is that D_{in} is replaced by $D_{in} + D_{un}$ and the residual operator removes signals related to both G_{in} and G_{un}, rather than just G_{un}. As the signals spanned by G_{un} are voxels in the area of no interest, of which there are D_{un}, it is completely equivalent to ignoring these area completely – that is, by effectively removing them from the measured image and doing the analysis as if only $N - D_{un}$ voxels had been measured.

Finally, the total number of parameters used to model the area of no interest does not appear in the final posterior. Only the number of uninteresting parameters (equal to the number of measured voxels totally in the area of no interest) is important. Consequently, there is no limit on the number of subshapes which may be used to model area of interest, and the subshapes can become infinitesimal, justifying the implicit assumption that it is possible to treat subshapes as only ever contributing to one image voxel at a time, regardless of the spatial transformation, T.

Note that a similar analysis applies to voxels which only partially cover the area of no interest. In this case, only $N_{pv} < N_0$ subshapes will contribute towards the voxel intensity, such that $\chi N_{pv} < 1$ and the non-null parameter becomes a pure partial volume parameter, with $q = \chi N_{pv}$ being the partial volume fraction. These will be treated in the next section.

3.2.2 Pure Partial Volume Parameters

Some parameters are associated only with a small partial volume effect, normally at a single voxel, which can occur when a shape of interest or the area of no interest partially overlaps a measured voxel. The parameters will be called pure partial volume parameters, α_{pv} here. In this case the associated column of G has a norm less than 1.0, and the finite range of α can no longer be ignored. Hence the approximation of using equation 13 (in appendix A) must be replaced by the more accurate integral of equation 14. However, in this case the region of integration also becomes important. In the preceding sections the variables are often transformed in order to compute the integrals (the substitution $F = KG$ which is done in appendix A).

For typical models, most columns of G, apart from those associated with α_{pv}, have norms that are much larger than 1.0, and are almost orthogonal with the columns associated with these partial volume parameters. Consequently, the effect of the change of variables is to induce a slight rotation in the effective region of integration. However, this effect is small and will be ignored here.

Applying the previous marginalisations of other α parameters first, leads to a posterior of the form

$$p(T|Y,S) \propto p(T)\, C_1^{D-D_{null}-D_{deg}} \,|\det(G_{in}^{\mathsf{T}}G_{in})|^{-1/2}\, (2\pi)^{-(N-D_{in}-D_{un})/2} \times \int \beta^{(N-D_{in}-D_{un}-2)/2} \exp\left(\frac{-\beta(Y-G_{pv}\alpha_{pv})^{\mathsf{T}}R_w(Y-G_{pv}\alpha_{pv})}{2}\right)\frac{1}{\beta}\,d\beta.$$

Using equation 14 and assuming that $G_{pv}^{\mathsf{T}}R_wG_{pv}$ has negligible off-diagonal terms gives

$$\begin{aligned} p(T|Y,S,\beta) &\propto p(T)\, C_1^{D-D_{null}-D_{deg}} \,|\det(G_{in}^{\mathsf{T}}G_{in})|^{-1/2}\, \left(\frac{2\pi}{\beta}\right)^{-(N-D_{in}-D_{un}-D_{pv})/2} \times \\ &\qquad \exp\left(\frac{-\beta Y^{\mathsf{T}}R_wR_{pv}R_wY}{2}\right)\,|\det(G_{pv}^{\mathsf{T}}R_wG_{pv})|^{-1/2}\,\left(\prod_{j=1}^{D_{pv}} w_j\right) \\ &\propto p(T)\, C_1^{N}\,|\det(G_{in}^{\mathsf{T}}G_{in})|^{-1/2}\,\left(\frac{2\pi C_1^2}{\beta}\right)^{-N_{eff}/2}\exp\left(\frac{-\beta Y^{\mathsf{T}}R_wR_{pv}R_wY}{2}\right)\,\left(\prod_{j=1}^{D_{pv}} q_j^{-1}w_j\right) \quad (9) \end{aligned}$$

where

$$w_j = \frac{1}{2}\mathrm{erfc}\left(\frac{-\beta^{1/2}Y^{\mathsf{T}}R_wG_{pv,j}}{(2G_{pv,j}^{\mathsf{T}}R_wG_{pv,j})^{1/2}}\right) - \frac{1}{2}\mathrm{erfc}\left(\frac{-\beta^{1/2}Y^{\mathsf{T}}R_wG_{pv,j}+\beta^{1/2}C_1^{-1}(G_{pv,j}^{\mathsf{T}}R_wG_{pv,j})}{(2G_{pv,j}^{\mathsf{T}}R_wG_{pv,j})^{1/2}}\right)$$

represents a general weighting factor for the jth voxel (associated with the jth partial volume parameter); q_j^2 is the jth diagonal element of $(G_{pv}^{\mathsf{T}}R_wG_{pv})$ which represents the squared partial volume fraction of the jth partial volume parameter; $N_{eff} = N - D_{in} - D_{un} - D_{pv}$ are the effective number of degrees of freedom for this model; and $R_{pv} = I - R_wG_{pv}(G_{pv}^{\mathsf{T}}R_wG_{pv})^{-1}G_{pv}^{\mathsf{T}}R_w$ is the residual projection matrix including G_{in}, G_{un} and G_{pv}. The projection matrix R_{pv} can also be calculated more straightforwardly from the pseudo-inverse of G as $R_{pv} = I - GG^{\dagger}$, where $G^{\dagger}$ is the pseudo-inverse of G.

Note that it is now no longer possible to easily marginalise over β as it appears in the complimentary error functions. Consequently, the posterior has been left in the form $p(T|Y,S,\beta)$ which can either be used with a pre-specified value of β (e.g. by estimating the SNR from the image) or numerically marginalised.

Using the fact that $SNR \approx \beta^{1/2}Y_j \gg 1$ and $Y^{\mathsf{T}}R_wG_{pv,j} \approx q_j(R_wY)_j$ gives

$$w_j = \frac{1}{2}\mathrm{erfc}\left(-\left(\frac{\beta}{2}\right)^{1/2}(R_wY)_j\right) - \frac{1}{2}\mathrm{erfc}\left(\left(\frac{\beta}{2}\right)^{1/2}\left(q_j\,C_1^{-1} - (R_wY)_j\right)\right)$$

which is a monotonic function of q.

Furthermore, as $q \to 1$ then $q_j^{-1}w_j \to 1$, and as $q \to 0$ then

$$q_j^{-1}w_j \to \left(\frac{2\pi C_1^2}{\beta}\right)^{-1/2}\exp\left(\frac{-\beta(R_wY)_j^2}{2}\right)$$

The consequence of this is that when the partial volume overlap becomes large ($q \to 1$) the posterior assumes the form that would result if this parameter was part of α_{un}, that is a voxel wholly inside the area of no interest. This is what would occur as the transformation moved in such a way as to move this measured voxel wholly into the area of no interest. Therefore the posterior is continuous with respect to this change in transformation.

Also, as the partial volume overlap becomes small ($q \to 0$) then the posterior assumes the form that would result by increasing N_{eff} by 1 and including the residual intensity mismatch from voxel j back into the general residual term. This is what would happen as this voxel makes the transition fully into the measured area of interest, resulting in an extra effective voxel in the measurement (the increase in N_{eff}) and the full inclusion of the residual intensity error.

In between these extreme values the contribution becomes partial. This can be compared to the form of de-weighting used for voxels at the edge of the valid field of view, in [3], which can be expressed in this notation as

$$\text{weighting} = \left(\frac{2\pi C_1^2}{\beta}\right)^{-(1-q)/2}\exp\left(\frac{-\beta(1-q)(R_wY)_j^2}{2}\right).$$

This takes the same extreme values and is also monotonic, and thus performs the same function, albeit with a slightly different rate of change.

Note that the careful use of the finite range of α and the introduction of the complimentary error functions in the integrations is crucial for making the posterior a continuous function of the spatial transformation.

3.3 Multi-Variate Gaussian Intensity Prior

Using the Gaussian prior for the parameters of interest, α_{in}

$$p(\alpha_{in}|\lambda, Q) = |\det(Q)|^{1/2} \left(\frac{\lambda}{\pi}\right)^{D_{in}/2} \exp\left(-\lambda \alpha_{in}^{\mathsf{T}} Q \alpha_{in}\right)$$

using an inverse covariance matrix, Q, of dimension D_{in} by D_{in}.

Neglecting all but the interesting parameters (for now) gives a posterior of the form

$$\begin{aligned}
p(T|Y,S,Q) &\propto \int p(Y|T,S,\beta,\alpha)p(T)p(\alpha|\lambda,Q)p(\lambda)p(\sigma)\, d\sigma\, d\lambda\, d\alpha \\
&\propto p(T) \int \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left(\frac{-\beta\|Y - G\alpha\|^2}{2}\right) p(\alpha|\lambda,Q)p(\lambda)p(\beta)\, d\beta\, d\lambda\, d\alpha \\
&\propto p(T) \int \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left(\frac{-\beta\|Y - G_{in}\alpha_{in}\|^2}{2}\right) |\det(Q)|^{1/2} \left(\frac{\lambda}{\pi}\right)^{D_{in}/2} \exp\left(-\lambda \alpha_{in}^{\mathsf{T}} Q \alpha_{in}\right) C_0 \frac{1}{\beta}\, d\beta\, d\lambda\, d\alpha_{in} \\
&\propto p(T)|\det(Q)|^{1/2} C_0 \int \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left(\frac{-\beta(Y - G_{in}\alpha_{in})^{\mathsf{T}}(Y - G_{in}\alpha_{in})}{2}\right) \left(\frac{\lambda}{\pi}\right)^{D_{in}/2} \exp\left(-\lambda \alpha_{in}^{\mathsf{T}} Q \alpha_{in}\right) \frac{1}{\beta}\, d\beta\, d\lambda\, d
\end{aligned}$$

When $Q \neq 0$ then the part of the posterior that depends on α_{in} can be written as

$$\begin{aligned}
&\int \exp\left(\frac{-\beta}{2}(Y - G_{in}\alpha_{in})^{\mathsf{T}}(Y - G_{in}\alpha_{in}) - \lambda \alpha_{in}^{\mathsf{T}} Q \alpha_{in}\right) d\alpha_{in} \\
&\quad = \int \exp\left(\frac{-\beta}{2}(Y^{\mathsf{T}}Y - 2Y^{\mathsf{T}}G_{in}\alpha_{in} + \alpha_{in}^{\mathsf{T}}G_{in}^{\mathsf{T}}G_{in}\alpha_{in}) - \lambda \alpha_{in}^{\mathsf{T}} Q \alpha_{in}\right) d\alpha_{in} \\
&\quad = \int \exp\left(-(\alpha_{in}^{\mathsf{T}} A_1 \alpha_{in} + 2A_2\alpha_{in} + A_3)\right) d\alpha_{in} \\
&\quad = (\pi)^{D_{in}/2} \, |\det(A_1)|^{-1/2} \exp\left(A_2 A_1^{-1} A_2^{\mathsf{T}} - A_3\right) \\
&\quad = (\pi)^{D_{in}/2} \left|\det\left(\frac{\beta}{2}G_{in}^{\mathsf{T}}G_{in} + \lambda Q\right)\right|^{-1/2} \exp\left(\frac{\beta^2}{4} Y^{\mathsf{T}} G_{in} \left[\frac{\beta}{2}G_{in}^{\mathsf{T}}G_{in} + \lambda Q\right]^{-1} G_{in}^{\mathsf{T}} Y - \frac{\beta}{2} Y^{\mathsf{T}} Y\right) \\
&\quad = \left(\frac{\beta}{2\pi}\right)^{-D_{in}/2} \left|\det\left(G_{in}^{\mathsf{T}}G_{in} + \frac{2\lambda}{\beta} Q\right)\right|^{-1/2} \exp\left(\frac{-\beta Y^{\mathsf{T}} R Y}{2}\right)
\end{aligned}$$

where $A_1 = \frac{\beta}{2}G_{in}^{\mathsf{T}}G_{in} + \lambda Q$, $A_2 = -\frac{\beta}{2}Y^{\mathsf{T}}G_{in}$, $A_3 = \frac{\beta}{2}Y^{\mathsf{T}}Y$ and $R = I - G_{in}\left[G_{in}^{\mathsf{T}}G_{in} + \frac{2\lambda}{\beta}Q\right]^{-1} G_{in}^{\mathsf{T}}$ which is a residual forming matrix (no longer a projection matrix) and it also depends on β and λ.

The full posterior is then

$$p(T|Y,S,Q,\beta,\lambda) \propto p(T)|\det(Q)|^{1/2} \left(\frac{\beta}{2\pi}\right)^{(N-D_{in})/2} \left(\frac{\lambda}{\pi}\right)^{D_{in}/2} \left|\det\left(G_{in}^{\mathsf{T}}G_{in} + \frac{2\lambda}{\beta}Q\right)\right|^{-1/2} \exp\left(\frac{-\beta Y^{\mathsf{T}} R Y}{2}\right)$$

This form is extremely difficult (probably impossible) to integrate analytically with respect to β and λ. Hence it is left in this semi-marginalised form.

3.3.1 Marginalising over Null Parameters

As long as the null parameters are associated with areas of no interest, then they can be assigned a flat prior and integrated as shown in the previous section. However, when a shape of interest moves out of the field of

view and its parameters become null parameters, then this requires special attention as the prior is no longer separable into null and non-null terms.

Specifically, consider the prior term, $\exp(-\lambda\alpha^{\mathsf{T}}Q\alpha)$, and rewrite the prior as:

$$\alpha = \begin{bmatrix} \alpha_{in} \\ \alpha_{null} \end{bmatrix} \qquad Q = \begin{bmatrix} Q_{in} & Q_{cross}^{\mathsf{T}} \\ Q_{cross} & Q_{null} \end{bmatrix}.$$

This gives

$$\begin{aligned}
\int \exp\left(-\lambda\alpha^{\mathsf{T}}Q\alpha\right)\, d\alpha_{null} &= \int \exp\left(-\lambda(\alpha_{in}^{\mathsf{T}}Q_{in}\alpha_{in} + 2\alpha_{in}^{\mathsf{T}}Q_{cross}^{\mathsf{T}}\alpha_{null} + \alpha_{null}^{\mathsf{T}}Q_{null}\alpha_{null})\right)\, d\alpha_{null} \\
&= \int \exp\left(-\lambda(\alpha_{null} + Q_{null}^{-1}Q_{cross}\alpha_{in})^{\mathsf{T}}Q_{null}(\alpha_{null} + Q_{null}^{-1}Q_{cross}\alpha_{in})\right) \\
&\qquad \exp\left(-\lambda(\alpha_{in}^{\mathsf{T}}Q_{in}\alpha_{in} - \alpha_{in}^{\mathsf{T}}Q_{cross}^{\mathsf{T}}Q_{null}^{-1}Q_{cross}\alpha_{in})\right)\, d\alpha_{null} \\
&= C_1^{-D_{null,u}} \left(\frac{\lambda}{\pi}\right)^{-D_{null,i}/2} |\det(Q_{null})|^{-1/2} \exp\left(-\lambda\alpha_{in}^{\mathsf{T}}(Q_{in} - Q_{cross}^{\mathsf{T}}Q_{null}^{-1}Q_{cross})\alpha_{in}\right)
\end{aligned}$$

where $D_{null,i}$ are the number of null parameters associated with shapes of interest, and $D_{null,u}$ are the number of null parameters associated with areas of no interest, such that $D_{null} = D_{null,i} + D_{null,u}$.

The resulting form is still a multi-variate Gaussian, and the effect on the previous formula is to modify Q by replacing it with its reduced form, $Q' = Q_{in} - Q_{cross}^{\mathsf{T}}Q_{null}^{-1}Q_{cross}$ and to pre-multiply the posterior by the factor

$$C_1^{-D_{null,u}} \left(\frac{\lambda}{\pi}\right)^{-D_{null,i}/2} |\det(Q_{null})|^{-1/2}.$$

where $Q_{null} = Q_{null}$ is the submatrix of Q associated with the null parameters.

Note that because the null space of G depends on both the underlying shape models, S_k, and the transformation, T, both D_{null} and Q_{null} depend on these and so this factor will not be a constant in the similarity function.

The posterior is now in the form

$$\begin{aligned}
p(T|Y,S,Q,\beta,\lambda) \quad \propto \quad & p(T)|\det(Q)|^{1/2}|\det(Q_{null})|^{-1/2} \left(\frac{\beta}{2\pi}\right)^{(N-D_{in})/2} \left(\frac{\lambda}{\pi}\right)^{D_{in}/2} \times \\
& \left|\det\left(G_{in}^{\mathsf{T}}G_{in} + \frac{2\lambda}{\beta}Q'\right)\right|^{-1/2} \exp\left(\frac{-\beta Y^{\mathsf{T}}RY}{2}\right)
\end{aligned}$$

where $Q' = Q_{in} - Q_{cross}^{\mathsf{T}}Q_{null}^{-1}Q_{cross}$ and $R = I - G_{in}\left[G_{in}^{\mathsf{T}}G_{in} + \frac{2\lambda}{\beta}Q'\right]^{-1}G_{in}^{\mathsf{T}}$.

The factor depending on C_1 does not appear, as it is cancelled by the prior, $p(\alpha_{null,u}) = C_1^{D_{null,u}}$, and the normalisation of the prior $p(\alpha|\lambda,Q)$ includes $(\lambda/\pi)^{(D_{in}+D_{null,i})/2}$ which cancels the other term.

3.3.2 Marginalisation over Areas of No Interest

Marginalisation over voxels wholly in the areas of no interest proceeds exactly as in the case of a flat intensity prior, since the multi-variate prior here does not include these parameters. These parameters therefore split into null, degenerate, and uninteresting. The degenerate and uninteresting parameters integrate in the same way as for the flat prior case and the null parameters integrate as part of $D_{null,u}$ in the previous section.

Specifically, this gives

$$\begin{aligned}
&\int \exp\left(\frac{-\beta}{2}(Y-G\alpha)^{\mathsf{T}}(Y-G\alpha)-\lambda\alpha_{in}^{\mathsf{T}}Q\alpha_{in}\right)d\alpha_{null}\,d\alpha_{un}\,d\alpha_{deg}\\
&\quad= C_1^{-D_{null,u}-D_{deg}}\left(\frac{\lambda}{\pi}\right)^{-D_{null,i}/2}|\det(Q_{null})|^{-1/2}\int\exp\left(\frac{-\beta}{2}(Y-G\alpha)^{\mathsf{T}}(Y-G\alpha)-\lambda\alpha_{in}^{\mathsf{T}}Q'\alpha_{in}\right)d\alpha_{un}\\
&\quad= C_1^{-D_{null,u}-D_{deg}}\left(\frac{\lambda}{\pi}\right)^{-D_{null,i}/2}|\det(Q_{null})|^{-1/2}\left(\frac{\beta}{2\pi}\right)^{-D_{un}/2}\left|\det\left(G_{un}^{\mathsf{T}}G_{un}\right)\right|^{-1/2}\times\\
&\qquad\exp\left(\frac{-\beta}{2}(Y-G\alpha)^{\mathsf{T}}R_{un}(Y-G\alpha)-\lambda\alpha_{in}^{\mathsf{T}}Q'\alpha_{in}\right)
\end{aligned}\tag{10}$$

where Q_{null} stands for the square part of Q associated with the shapes of interest that are currently in the null space (totally outside the field of view), and $R_{un}=I-G_{un}(G_{un}^{\mathsf{T}}G_{un})^{-1}G_{un}^{\mathsf{T}}$.

$$\begin{aligned}
p(T|Y,S,Q,\beta,\lambda)\quad\propto\quad & p(T)\,C_1^{D_{un}}\,|\det(Q)|^{1/2}|\det(Q_{null})|^{-1/2}\left(\frac{\beta}{2\pi}\right)^{(N-D_{in}-D_{un})/2}\left(\frac{\lambda}{\pi}\right)^{D_{in}/2}\times\\
&\left|\det\left(G_{un}^{\mathsf{T}}G_{un}\right)\right|^{-1/2}\left|\det\left(G_{in}^{\mathsf{T}}R_{un}G_{in}+\frac{2\lambda}{\beta}Q'\right)\right|^{-1/2}\exp\left(\frac{-\beta Y^{\mathsf{T}}R_{un}RR_{un}Y}{2}\right)
\end{aligned}$$

where $Q'=Q_{in}-Q_{cross}^{\mathsf{T}}Q_{null}^{-1}Q_{cross}$, $R_{un}=I-G_{un}(G_{un}^{\mathsf{T}}G_{un})^{-1}G_{un}^{\mathsf{T}}$ and
$R=I-R_{un}G_{in}\left[G_{in}^{\mathsf{T}}R_{un}G_{in}+\frac{2\lambda}{\beta}Q'\right]^{-1}G_{in}^{\mathsf{T}}R_{un}$.

3.3.3 Marginalisation over Partial Volume Parameters

Pure partial volume parameters can again be split into those associated with shapes of interest, $\alpha_{pv,i}$, and those associated with areas of no interest, $\alpha_{pv,u}$. The former have non-flat priors and the latter have flat priors. Integration of the former requires more care, as they interact with the multi-variate prior. In order to do this, rewrite the appropriate matrices as

$$\alpha=\begin{bmatrix}\alpha_{pv,i}\\ \alpha_{in}\\ \alpha_{null,i}\end{bmatrix}\qquad G=[\,G_{pv,i}\;\;G_{in}\;\;0\,]\qquad Q=\begin{bmatrix}Q_{pv}&Q_a^{\mathsf{T}}&Q_b^{\mathsf{T}}\\ Q_a&Q_{in}&Q_{cross}^{\mathsf{T}}\\ Q_b&Q_{cross}&Q_{null}\end{bmatrix}.$$

Therefore the posterior integrations take the form

$$\begin{aligned}
&\int\exp\left(\frac{-\beta}{2}(Y-G\alpha)^{\mathsf{T}}(Y-G\alpha)-\lambda\alpha^{\mathsf{T}}Q\alpha\right)d\alpha\\
&\quad=\int\exp\left(\frac{-\beta}{2}(Y-G_{pv,i}\alpha_{pv,i})^{\mathsf{T}}(Y-G_{pv,i}\alpha_{pv,i})-\lambda\alpha_{pv,i}^{\mathsf{T}}Q_{pv}\alpha_{pv,i}\right)\times\\
&\qquad\exp\left(\frac{-\beta}{2}(\alpha_{in}^{\mathsf{T}}G_{in}^{\mathsf{T}}+2\alpha_{pv,i}G_{pv,i}^{\mathsf{T}}-2Y^{\mathsf{T}})G_{in}\alpha_{in}-\lambda(\alpha_{in}^{\mathsf{T}}Q_{in}+2\alpha_{pv,i}^{\mathsf{T}}Q_a^{\mathsf{T}}+2\alpha_{null}^{\mathsf{T}}Q_c)\alpha_{in}\right)\times\\
&\qquad\exp\left(-\lambda(2\alpha_{pv,i}^{\mathsf{T}}Q_b^{\mathsf{T}}+\alpha_{null}^{\mathsf{T}}Q_{null})\alpha_{null}\right)\;d\alpha_{pv,i}\,d\alpha_{in}\,d\alpha_{null}
\end{aligned}$$

Integrating with respect to α_{in} gives

$$\begin{aligned}
&\int\exp\left(\frac{-\beta}{2}(\alpha_{in}^{\mathsf{T}}G_{in}^{\mathsf{T}}+2\alpha_{pv,i}G_{pv,i}^{\mathsf{T}}-2Y^{\mathsf{T}})G_{in}\alpha_{in}-\lambda(\alpha_{in}^{\mathsf{T}}Q_{in}+2\alpha_{pv,i}^{\mathsf{T}}Q_a^{\mathsf{T}}+2\alpha_{null}^{\mathsf{T}}Q_c)\alpha_{in}\right)d\alpha_{in}\\
&\quad=\int\exp\left(\frac{-\beta}{2}(\alpha_{in}^{\mathsf{T}}A_0\alpha_{in}+2B_0^{\mathsf{T}}\alpha_{in})\right)d\alpha_{in}\\
&\quad=\left(\frac{\beta}{2}\right)^{-D_{in}/2}|\det(A_0)|^{-1/2}\exp\left(\frac{\beta}{2}B_0^{\mathsf{T}}A_0^{-1}B_0\right)
\end{aligned}$$

where $A_0=G_{in}^{\mathsf{T}}G_{in}+\frac{2\lambda}{\beta}Q_{in}$ and $B_0=G_{in}^{\mathsf{T}}G_{pv,i}\alpha_{pv,i}-G_{in}^{\mathsf{T}}Y+\frac{2\lambda}{\beta}(Q_a\alpha_{pv,i}+Q_c^{\mathsf{T}}\alpha_{null})$.

Substituting this back into the previous expression and integrating over α_{null} gives

$$
\begin{aligned}
&\int \exp\left(\frac{-\beta}{2}(Y-G\alpha)^\mathsf{T}(Y-G\alpha)-\lambda\alpha^\mathsf{T}Q\alpha\right)\,d\alpha \\
&= \left(\frac{\beta}{2}\right)^{-D_{in}/2} |\det(A_0)|^{-1/2}\int \exp\left(\frac{-\beta}{2}(Y^\mathsf{T}Y-2Y^\mathsf{T}G_{pv,i}\alpha_{pv,i}+\alpha_{pv,i}^\mathsf{T}G_{pv,i}^\mathsf{T}G_{pv,i}\alpha_{pv,i})-\lambda\alpha_{pv,i}^\mathsf{T}Q_{pv}\alpha_{pv,i}\right)\times \\
&\quad \exp\left(\frac{\beta}{2}(Y^\mathsf{T}G_{in}A_0^{-1}G_{in}^\mathsf{T}Y-2Y^\mathsf{T}G_{in}A_0^{-1}(G_{in}^\mathsf{T}G_{pv,i}+\frac{2\lambda}{\beta}Q_a)\alpha_{pv,i}+\alpha_{pv,i}^\mathsf{T}(G_{in}^\mathsf{T}G_{pv,i}+\frac{2\lambda}{\beta}Q_a)^\mathsf{T}A_0^{-1}(G_{in}^\mathsf{T}G_{pv,i}+\frac{2\lambda}{\beta}Q_a)\alpha_{pv,}\right. \\
&\quad \exp\left(-\lambda\left[2Y^\mathsf{T}G_{in}A_0^{-1}Q_c^\mathsf{T}+2\alpha_{pv,i}^\mathsf{T}(Q_b^\mathsf{T}-G_{pv,i}^\mathsf{T}G_{in}A_0^{-1}Q_c^\mathsf{T}-\frac{2\lambda}{\beta}Q_a^\mathsf{T}A_0^{-1}Q_c^\mathsf{T})+\alpha_{null}^\mathsf{T}(Q_{null}-\frac{2\lambda}{\beta}Q_cA_0^{-1}Q_c^\mathsf{T})\right]\alpha_{null}\right)\,d\alpha_{pv,} \\
&= \left(\frac{\beta}{2}\right)^{-D_{in}/2} |\det(A_0)|^{-1/2}\int \exp\left(\frac{-\beta}{2}(Y^\mathsf{T}R_0Y+\alpha_{pv,i}^\mathsf{T}A_2\alpha_{pv,i}+2B_2^\mathsf{T}\alpha_{pv,i})-\lambda(\alpha_{null}^\mathsf{T}A_1\alpha_{null}+2B_1^\mathsf{T}\alpha_{null})\right)\,d\alpha_{pv,i}\,d\alpha_{n} \\
&= \left(\frac{\beta}{2}\right)^{-D_{in}/2} |\det(A_0)|^{-1/2}\int \exp\left(\frac{-\beta}{2}(Y^\mathsf{T}R_0Y+\alpha_{pv,i}^\mathsf{T}A_2\alpha_{pv,i}+2B_2^\mathsf{T}\alpha_{pv,i})\right)\lambda^{-D_{null}/2}|\det(A_1)|^{-1/2}\exp\left(\lambda B_1^\mathsf{T}A_1^{-1}B_1\right) \\
&= \left(\frac{\beta}{2}\right)^{-D_{in}/2} \lambda^{-D_{null}/2}|\det(A_0\,A_1)|^{-1/2}\int \exp\left(\frac{-\beta}{2}(Y^\mathsf{T}R_1Y+\alpha_{pv,i}^\mathsf{T}A_3\alpha_{pv,i}+2B_3^\mathsf{T}\alpha_{pv,i})\right)\,d\alpha_{pv,i}
\end{aligned}
$$

where

$$
\begin{aligned}
R_0 &= I-G_{in}A_0^{-1}G_{in}^\mathsf{T} \\
A_0 &= G_{in}^\mathsf{T}G_{in}+\frac{2\lambda}{\beta}Q_{in} \\
A_1 &= Q_{null}-\frac{2\lambda}{\beta}Q_cA_0^{-1}Q_c^\mathsf{T} \\
B_1^\mathsf{T} &= Y^\mathsf{T}G_{in}A_0^{-1}Q_c^\mathsf{T}+\alpha_{pv,i}^\mathsf{T}\left(Q_b^\mathsf{T}-G_{pv,i}^\mathsf{T}G_{in}A_0^{-1}Q_c^\mathsf{T}-\frac{2\lambda}{\beta}Q_a^\mathsf{T}A_0^{-1}Q_c^\mathsf{T}\right) \\
A_2 &= G_{pv,i}^\mathsf{T}G_{pv,i}+\frac{2\lambda}{\beta}Q_{pv}-(G_{in}^\mathsf{T}G_{pv,i}+\frac{2\lambda}{\beta}Q_a)^\mathsf{T}A_0^{-1}(G_{in}^\mathsf{T}G_{pv,i}+\frac{2\lambda}{\beta}Q_a) \\
B_2^\mathsf{T} &= Y^\mathsf{T}G_{in}A_0^{-1}(G_{in}^\mathsf{T}G_{pv,i}+\frac{2\lambda}{\beta}Q_a)-Y^\mathsf{T}G_{pv,i} \\
A_3 &= A_2-\frac{2\lambda}{\beta}\left(Q_b^\mathsf{T}-G_{pv,i}^\mathsf{T}G_{in}A_0^{-1}Q_c^\mathsf{T}-\frac{2\lambda}{\beta}Q_a^\mathsf{T}A_0^{-1}Q_c^\mathsf{T}\right)A_1^{-1}\left(Q_b^\mathsf{T}-G_{pv,i}^\mathsf{T}G_{in}A_0^{-1}Q_c^\mathsf{T}-\frac{2\lambda}{\beta}Q_a^\mathsf{T}A_0^{-1}Q_c^\mathsf{T}\right)^\mathsf{T} \\
B_3^\mathsf{T} &= B_2^\mathsf{T}+Y^\mathsf{T}G_{in}A_0^{-1}Q_c^\mathsf{T}A_1^{-1}\left(Q_b^\mathsf{T}-G_{pv,i}^\mathsf{T}G_{in}A_0^{-1}Q_c^\mathsf{T}-\frac{2\lambda}{\beta}Q_a^\mathsf{T}A_0^{-1}Q_c^\mathsf{T}\right) \\
R_1 &= R_0-\frac{2\lambda}{\beta}G_{in}A_0^{-1}Q_c^\mathsf{T}A_1^{-1}Q_cA_0^{-1}G_{in}^\mathsf{T}
\end{aligned}
$$

However, as shown above, if all the previous integrations are performed first, then the remaining posterior takes the form

$$
\begin{aligned}
p(T|Y,S,Q,\beta,\lambda) \quad\propto\quad & p(T)\,C_1^{D_{un}}\,|\det(Q)|^{1/2}|\det(Q_{null})|^{-1/2}\left(\frac{\beta}{2\pi}\right)^{(N-D_{in}-D_{un})/2}\left(\frac{\lambda}{\pi}\right)^{D_{in}/2}\times \\
& \left|\det\left(G_{un}^\mathsf{T}G_{un}\right)\right|^{-1/2}\left|\det\left(G_1^\mathsf{T}R_{un}G_1+\frac{2\lambda}{\beta}Q'\right)\right|^{-1/2}\times \\
& \int\exp\left(\frac{-\beta}{2}(Y-G_{pv}\alpha_{pv})^\mathsf{T}R_{un}RR_{un}(Y-G_{pv}\alpha_{pv})-\lambda\alpha_{pv,i}^\mathsf{T}Q''\alpha_{pv,i}\right)\,d\alpha_{pv,i}\,d\alpha_{pv,u}
\end{aligned}
$$

where $R_{un}=I-G_{un}(G_{un}^\mathsf{T}G_{un})^{-1}G_{un}^\mathsf{T}$; $R=I-R_{un}G_1\left[G_1^\mathsf{T}R_{un}G_1+\frac{2\lambda}{\beta}Q'\right]^{-1}G_1^\mathsf{T}R_{un}$; $G_1=[\,G_{pv,i}\;\;G_{in}\,]$; and

$$
Q'=\begin{bmatrix} Q_{pv}-Q_b^\mathsf{T}Q_{null}^{-1}Q_b & Q_a^\mathsf{T}-Q_b^\mathsf{T}Q_{null}^{-1}Q_{cross} \\ Q_a-Q_{cross}^\mathsf{T}Q_{null}^{-1}Q_b & Q_{in}-Q_{cross}^\mathsf{T}Q_{null}^{-1}Q_{cross}\end{bmatrix}
$$

and $Q''=(Q_{pv}-Q_b^\mathsf{T}Q_{null}^{-1}Q_b)-(Q_a-Q_{cross}^\mathsf{T}Q_{null}^{-1}Q_b)^\mathsf{T}(Q_{in}-Q_{cross}^\mathsf{T}Q_{null}^{-1}Q_{cross})^{-1}(Q_a-Q_{cross}^\mathsf{T}Q_{null}^{-1}Q_b)$.

Assuming that $G_{pv,u}^{\mathsf{T}} R_{un} R R_{un} G_{pv,u}$ and $(G_{pv,i}^{\mathsf{T}} R_{un} R R_{un} G_{pv,i} + (2\lambda/\beta) {Q'}_{pv})$ both have negligible off-diagonal terms gives

$$
\begin{aligned}
p(T|Y,S,Q,\beta,\lambda) \quad \propto \quad & p(T)\, C_1^{D_{un}+D_{pv,u}} \,|\det(Q)|^{1/2} |\det(Q_{null})|^{-1/2} \left(\frac{\beta}{2\pi}\right)^{(N-D_{in}-D_{un}-D_{pv,i}-D_{pv,u})/2} \left(\frac{\lambda}{\pi}\right)^{(D_{in}+D_{pv,i})/2} \times \\
& \left|\det\left(G_{un}^{\mathsf{T}} G_{un}\right)\right|^{-1/2} \left|\det\left(G_{in}^{\mathsf{T}} R_{un} G_{in} + \frac{2\lambda}{\beta} Q'\right)\right|^{-1/2} \times \\
& \exp\left(\frac{-\beta}{2} Y^{\mathsf{T}} R_{un} R^{1/2} R_{pv,u} R_{pv} R_{pv,u} R^{1/2} R_{un} Y\right) \left(\prod_{j=1}^{D_{pv,u}} q_j^{-1} w_j\right) \left(\prod_{j=1}^{D_{pv,i}} {q'}_j^{-1} w'_j\right) \\
\propto \quad & p(T)\, C_1^{N} \,|\det(Q)|^{1/2} |\det(Q_{null})|^{-1/2} \left(\frac{2\pi C_1^2}{\beta}\right)^{-(N-D_{un}-D_{pv,u})/2} \left(\frac{2\lambda}{\beta}\right)^{(D_{in}+D_{pv,i})/2} \times \\
& \left|\det\left(G_{un}^{\mathsf{T}} G_{un}\right)\right|^{-1/2} \left|\det\left(G_{in}^{\mathsf{T}} R_{un} G_{in} + \frac{2\lambda}{\beta} Q'\right)\right|^{-1/2} \times \\
& \exp\left(\frac{-\beta}{2} Y^{\mathsf{T}} R_{un} R^{1/2} R_{pv,u} R_{pv} R_{pv,u} R^{1/2} R_{un} Y\right) \left(\prod_{j=1}^{D_{pv,u}} q_j^{-1} w_j\right) \left(\prod_{j=1}^{D_{pv,i}} {q'}_j^{-1} w'_j\right)
\end{aligned}
$$

where
$R_{pv,u} = I - R^{1/2} R_{un} G_{pv,u} (G_{pv,u}^{\mathsf{T}} R_{un} R R_{un} G_{pv,u})^{-1} G_{pv,u}^{\mathsf{T}} R^{1/2} R_{un}$;
$R_{pv} = I - R_{pv,u} R^{1/2} R_{un} G_{pv,i} (G_{pv,i}^{\mathsf{T}} R_{un} R^{1/2} R_{pv,u} R^{1/2} R_{un} G_{pv,i})^{-1} G_{pv,i}^{\mathsf{T}} R_{pv,u} R^{1/2} R_{un}$;
q_j^2 is the jth diagonal of $(G_{pv,u}^{\mathsf{T}} R_{un} R R_{un} G_{pv,u})$;
${q'}_j^2$ is the jth diagonal of $(G_{pv,i}^{\mathsf{T}} R_{un} R^{1/2} R_{pv,u} R^{1/2} R_{un} G_{pv,i} + (2\lambda/\beta) Q'')$;

$$
w_j = \frac{1}{2} - \frac{1}{2}\mathrm{erfc}\left(\left(\frac{\beta}{2}\right)^{1/2} \left(q_j C_1^{-1} - (R^{1/2} R_{un} Y)_j\right)\right)
$$

and

$$
w'_j = \frac{1}{2} - \frac{1}{2}\mathrm{erfc}\left(\left(\frac{\beta}{2}\right)^{1/2} \left({q'}_j C_1^{-1} - (R_{pv,u} R^{1/2} R_{un} Y)_j\right)\right).
$$

3.4 Summary

The derived posterior probabilities are as follows.

3.4.1 Flat Intensity Prior

$$
p(T|Y,S,\beta) \propto p(T)\, C_1^N \,|\det(G_{in}^{\mathsf{T}} G_{in})|^{-1/2} \left(\frac{2\pi C_1^2}{\beta}\right)^{-N_{eff}/2} \exp\left(\frac{-\beta Y^{\mathsf{T}} R_w R_{pv} R_w Y}{2}\right) \left(\prod_{j=1}^{D_{pv}} q_j^{-1} w_j\right) \qquad (11)
$$

where

$$
w_j = \frac{1}{2}\mathrm{erfc}\left(-\left(\frac{\beta}{2}\right)^{1/2} (R_w Y)_j\right) - \frac{1}{2}\mathrm{erfc}\left(\left(\frac{\beta}{2}\right)^{1/2} \left(q_j\, C_1^{-1} - (R_w Y)_j\right)\right);
$$

q_j^2 is the jth diagonal element of $(G_{pv}^{\mathsf{T}} R_w G_{pv})$; $N_{eff} = N - D_{in} - D_{un} - D_{pv}$; $R_{pv} = I - R_w G_{pv} (G_{pv}^{\mathsf{T}} R_w G_{pv})^{-1} G_{pv}^{\mathsf{T}} R_w$; and $R_w = I - G_{in}(G_{in}^{\mathsf{T}} G_{in})^{-1} G_{in}^{\mathsf{T}} - G_{un}(G_{un}^{\mathsf{T}} G_{un})^{-1} G_{un}^{\mathsf{T}}$.

3.4.2 Multi-variate Gaussian Prior

$$p(T|Y,S,Q,\beta,\lambda) \propto p(T)\, C_1^N\, |\det(Q)|^{1/2} |\det(Q_{null})|^{-1/2} \left(\frac{2\pi C_1^2}{\beta}\right)^{-(N-D_{un}-D_{pv,u})/2} \left(\frac{2\lambda}{\beta}\right)^{(D_{in}+D_{pv,i})/2} \times$$
$$\left|\det\left(G_{un}^{\mathsf{T}} G_{un}\right)\right|^{-1/2} \left|\det\left(G_{in}^{\mathsf{T}} R_{un} G_{in} + \frac{2\lambda}{\beta} Q'\right)\right|^{-1/2} \times$$
$$\exp\left(\frac{-\beta}{2} Y^{\mathsf{T}} R_{un} R^{1/2} R_{pv,u} R_{pv} R_{pv,u} R^{1/2} R_{un} Y\right) \left(\prod_{j=1}^{D_{pv,u}} q_j^{-1} w_j\right) \left(\prod_{j=1}^{D_{pv,i}} {q'}_j^{-1} w'_j\right)$$

where
$R_{pv,u} = I - R^{1/2} R_{un} G_{pv,u} (G_{pv,u}^{\mathsf{T}} R_{un} R R_{un} G_{pv,u})^{-1} G_{pv,u}^{\mathsf{T}} R^{1/2} R_{un}$;
$R_{pv} = I - R_{pv,u} R^{1/2} R_{un} G_{pv,i} (G_{pv,i}^{\mathsf{T}} R_{un} R^{1/2} R_{pv,u} R^{1/2} R_{un} G_{pv,i})^{-1} G_{pv,i}^{\mathsf{T}} R_{pv,u} R^{1/2} R_{un}$;
q_j^2 is the jth diagonal of $(G_{pv,u}^{\mathsf{T}} R_{un} R R_{un} G_{pv,u})$;
${q'}_j^2$ is the jth diagonal of $(G_{pv,i}^{\mathsf{T}} R_{un} R^{1/2} R_{pv,u} R^{1/2} R_{un} G_{pv,i} + (2\lambda/\beta) Q'')$;

$$w_j = \frac{1}{2}\mathrm{erfc}\left(-\left(\frac{\beta}{2}\right)^{1/2} (R^{1/2} R_{un} Y)_j\right) - \frac{1}{2}\mathrm{erfc}\left(\left(\frac{\beta}{2}\right)^{1/2} \left(q_j C_1^{-1} - (R^{1/2} R_{un} Y)_j\right)\right)$$

and

$$w'_j = \frac{1}{2}\mathrm{erfc}\left(-\left(\frac{\beta}{2}\right)^{1/2} (R_{pv,u} R^{1/2} R_{un} Y)_j\right) - \frac{1}{2}\mathrm{erfc}\left(\left(\frac{\beta}{2}\right)^{1/2} \left({q'}_j C_1^{-1} - (R_{pv,u} R^{1/2} R_{un} Y)_j\right)\right).$$

3.5 Shape-specific Intensity Distributions

The above model may be easily extended to include some non-deterministic intensity characteristics for the shapes. For instance, typical shapes are characterised not only by a mean intensity and a spatially-linear intensity gradient, but also by a distribution of intensities about this deterministic intensity model. This distribution can be characterised empirically and used in the similarity model, as done in [**?**, **?**]. Alternatively, the distribution can be approximated by a Gaussian and these variance properties inserted into the above model.

Consider that each shape (S_j) is associated with a Gaussian noise process (of length N), η_j, where $\mathrm{Cov}(\eta_j) = E(\eta_j \eta_j^{\mathsf{T}}) = \sigma_j^2 I$. The model of image formation is now $Y = G\alpha + \sum_j W_j \eta_j + \epsilon$, where W_j is an N by N weighting matrix given by $\mathrm{diag}(G_1(T(S_j)))$ – i.e. a diagonal matrix where the diagonal elements are taken from the vector $G_1(T(S_j))$. This weighting is such that the noise process η_j will only affect voxels that overlap S_j and will not affect other voxels.

The random component of this model is $\sum_j W_j \eta_j + \epsilon = Y - G\alpha$ and is a multivariate Gaussian with covariance of $\mathrm{Cov}(Y - G\alpha) = \sum_j \sigma_j^2 W_j W_j^{\mathsf{T}} + \sigma^2 I = V$. As a consequence the likelihood becomes

$$\begin{aligned} p(Y|T,S,\theta) &= (2\pi)^{-N/2} |\det(V)|^{-1/2} \exp\left(\frac{-(Y-G\alpha)^{\mathsf{T}} V^{-1} (Y-G\alpha)}{2}\right) \\ &= (2\pi)^{-N/2} |\det(V)|^{-1/2} \exp\left(\frac{-(V^{-1/2}Y - V^{-1/2}G\alpha)^{\mathsf{T}} (V^{-1/2}Y - V^{-1/2}G\alpha)}{2}\right) \end{aligned}$$

which is the same form as before, but with $\beta = 1$, Y replaced by $V^{-1/2}Y$ G replaced by $V^{-1/2}G$ and the prefactor $|\det(V)|^{-1/2}$ inserted. Therefore all previous results hold with these substitutions. Note that this is only true when V has a nearly block diagonal structure (with few cross-terms between shapes) which is the case for this modelling with non-overlapping shapes.

The posterior for this new model now also depends upon the parameters, σ_j (or $\beta_j = \sigma_j^{-2}$) which can be either treated as known parameters, or marginalised numerically. Their effect on the projection matrices and determinants is such that analytical marginalisation is intractable. Also, it is often more convenient to subsume the measurement noise, ϵ, with the new random processes, η_j, in order to simplify the model and marginalisation. This simply has the effect of changing the values of σ_j that will be used (or integrated over) in practice, since all of these processes are considered independent.

3.6 Implementation

Both forms of priors lead to posteriors that depend on β and possibly λ where these cannot be easily integrated over analytically. Therefore the alternatives are: (1) to approximate the integration numerically; (2) to simplify the models/assumptions (e.g. flat priors on α); or (3) to set β and λ to be known constants (pragmatically they can be measured from the data Y).

The most expensive computation is that of the residuals, as this requires the accumulation of intensities over many voxels and the appropriate updating of summary statistics to do the effective planar fit. Once these statistics have been generated, the remaining matrix computations are relatively fast and so it is feasible to integrate over β numerically, as the residual term is easily and cheaply recalculated.

4 Validation Tests

4.1 Similarity Measures

In order to test the usefulness of the above derivation, the posterior-based similarity function for the flat prior case is investigated. That is

$$\begin{aligned} F_0 &= \log\left(p(T|Y,S,\beta)\right) + \text{constant} \\ &= \log(p(T)) - \frac{1}{2}\log\Big(|\det(G_{in}^{\mathsf{T}}G_{in})|\Big) + \frac{N_{eff}}{2}\log\left(\frac{\beta}{2\pi C_1^2}\right) - \left(\frac{\beta Y^{\mathsf{T}}R_w R_{pv} R_w Y}{2}\right) + \sum_{j=1}^{D_{pv}}(\log(w_j) - \log(q_j)) \end{aligned}$$

was compared with simpler similarity functions

$$\begin{aligned} F_1 &= Y^{\mathsf{T}}R_w R_{pv} R_w Y \\ F_2 &= \frac{1}{N_{eff}}\left(Y^{\mathsf{T}}R_w R_{pv} R_w Y\right) \end{aligned}$$

in the performance of some simple segmentation tasks.

In addition, the performance is compared with registration to an atlas image (generated as a separate noiseless measurement of the model).

4.2 Test Model

The test model is one dimensional and consists of 2 known shapes, an area of unknown intensity and some background. Only global translation will be considered for the spatial transformation, T. The sizes and intensities of the shapes are shown in figure 3. Note that the contrast between shapes 1 and 2 is quite small, which is typical in many applications. The measured image is taken to have 256 voxels, each of 1mm width. The position of the field of view of the measured image is varied across the object, over a 100mm range, to demonstrate the effect of partial object-image overlap on the similarity functions.

The portion of the object to the left of shape 1 is modelled using a set of small 'voxel-shapes' (each 1mm wide) to account for unknown object structure in this region. To the right of shape 3 there is no object and this is therefore modelled by a single extra shape to account for potential constant background intensity (common in MRI where the background noise is strictly Rician, not Gaussian, and hence not zero mean).

4.3 Experimental Tests

Three separate tests of the similarity functions and registrations will be performed:

1. Each shape having a constant intensity.
2. Shape 2 having a linear intensity gradient (representing variation of the object's density).

	SNR = 5		SNR = 15		SNR = 50	
F_0	MR = 0.194	MA = 0.71	MR = 0.065	MA = 0.26	MR = 0.000	MA = 0.20
F_1	MR = 1.000	-	MR = 0.909	MA = 0.35	MR = 0.807	MA = 0.97
F_2	MR = 0.516	MA = 0.86	MR = 0.494	MA = 0.48	MR = 0.129	MA = 0.54

Table 1: Results of robustness and accuracy measures for a simple object (see figure 3b) and a range of SNR values, for each of the three cost functions. $\Delta = 2$mm.

3. As for test 2 but with bias field (extra linear intensity gradients) added to the measured image, Y. The amount of bias field was chosen randomly for each trial from a uniform distribution such that the intensity variation did not exceed $\pm 10\%$ of the mean intensity of shape 2.

In each case a range of SNR values is tried, each using a set of 100 instances of measurement noise.

Furthermore, the field of view of the measurement, Y, was shifted along the model object, as illustrated in figure **??**. This allowed the sensitivity of the similarity functions with respect to changing fields of view to be measured.

4.4 Performance Criteria

Performance was assessed by comparing the correct translation value, T_{true}, used to generate the measured image, to the recovered translation parameter estimate $T_{MAP,j}$ in the jth trial.

A measure of the robustness of the estimated translation is:

$$M_R = \frac{1}{N_{trials}} \#\{T_{MAP,j} \mid (T_{MAP,j} - T_{true})^2 > \Delta^2\}$$

where $\#\{A\}$ represents the number of elements in set A, and Δ is a moderate tolerance, taken to be 2mm in this case. This measure represents the fraction of solutions which were not considered "close", where Δ sets the threshold by which something is considered "close".

A measure of the accuracy of the estimated translation is:

$$M_A = \frac{1}{\Delta} \left(\frac{1}{N_{trials}} \sum_{j=1}^{N_{trials}} \min\left((T_{MAP,j} - T_{true})^2, \Delta^2 \right) - M_R \Delta^2 \right)^{1/2}$$

This measure is a modified RMS error measurement, where any errors greater than Δ are replaced by Δ. In this way it downweights the contribution for very large errors which a non-robust technique can produce and gives a normalised measure between 0 and 1.

For an ideal method, both measures, M_R and M_A, would be zero.

5 Results

Some sample plots of the similarity functions F_0, F_1 and F_2 are shown in figure 4. These show the functions over a relatively large range of translations, and demonstrate two important features: that the secondary (incorrect) peak is downweighted in F_0, reducing the chances large mis-matches; and in the close-up results, the improved continuity of the similarity functions (which is smoother and has less discontinuities for F_0).

Table 1 shows the values of M_R and M_A for 3 different SNR values and 2000 trials in each case. The low values of M_R for F_0 are a result of the downweighting of the secondary maxima, significantly reducing bad mismatches. Also, F_0 was the most accurate (having the smallest value of M_A) which is likely to be due to the decrease in discontinuities in F_0.

6 Discussion

The derived similarity function, F_0, demonstrates superior accuracy and robustness when compared with the simpler similarity functions, F_1 and F_2. The partial volume terms decrease the number and extent of discontinuities, and the terms that normalise for the number of model parameters, N_{eff}, de-weight erroneous local maxima. In addition, this similarity function automatically includes prior shape information, via $p(T)$, as well as partial volume and bias field effects without needing adjustable, ad-hoc parameters.

Future work will apply this similarity function to our intended application of anatomical shape segmentation of the human brain and extend the above to incorporate intensity priors, and modelling intrinsic tissue parameter distributions with one variance per shape.

A Useful Integrals

A.1 Marginalisation of Precision:

$$\int_0^\infty \beta^c e^{-H\beta}\, d\beta = \Gamma(c) H^{-(c+1)} \tag{12}$$

A.2 Uni-variate Gaussian Integration:

$$\int_{-\infty}^{\infty} \exp\left(-(ax^2+bx+c)\right)\, dx = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2-4ac}{4a}\right) \tag{13}$$

or in the finite case

$$\int_0^{L} \exp\left(-(ax^2+bx+c)\right)\, dx = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2-4ac}{4a}\right) \left(\frac{1}{2}\mathrm{erfc}\left(\frac{b}{2\sqrt{a}}\right) - \frac{1}{2}\mathrm{erfc}\left(\frac{b+2La}{2\sqrt{a}}\right)\right) \tag{14}$$

where erfc is the complimentary error function, defined as

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2)\, dt$$

A.3 Multi-variate Gaussian Integration:

A.3.1 Case 1:

$$\begin{aligned} \int \exp\left(-(Y-A)^{\mathsf{T}} V^{-1} (Y-A)\right)\, dY &= \int \exp\left(-Z^{\mathsf{T}} Z\right) |\det(V^{-1/2})|^{-1}\, dZ \\ &= (\pi)^{N/2} |\det(V)|^{1/2} \end{aligned} \tag{15}$$

where $\dim(Y) = N$ and using $Z = V^{-1/2}(Y-A)$ so that $dZ = |\det(V^{-1/2})| dY$.

A.3.2 Case 2:

$$\begin{aligned} \int \exp\left(-(Y^{\mathsf{T}} A Y + BY + C)\right)\, dY &= \int \exp\left(-(Z^{\mathsf{T}} A Z + C - \frac{1}{4} B A^{-1} B^{\mathsf{T}})\right)\, dZ \\ &= (\pi)^{N/2} |\det(A)|^{-1/2} \exp\left(\frac{1}{4} B A^{-1} B^{\mathsf{T}} - C\right) \end{aligned} \tag{16}$$

where $Z = Y + \frac{1}{2}A^{-1}B^{\mathsf{T}}$ and A^{-1} exists.

A.3.3 Case 3:

$$\begin{aligned}
\int \exp\left(-(Y-BF)^{\mathsf{T}}V^{-1}(Y-BF)\right)\,dF &= \int \exp\left(-Y^{\mathsf{T}}V^{-1}Y + 2Y^{\mathsf{T}}V^{-1}BF - F^{\mathsf{T}}B^{\mathsf{T}}V^{-1}BF\right)\,dF \\
&= |\det(K)| \int \exp\left(-Y^{\mathsf{T}}V^{-1}Y + 2Y^{\mathsf{T}}V^{-1}BKG - G^{\mathsf{T}}K^{\mathsf{T}}B^{\mathsf{T}}V^{-1}BKG\right)\,dG \\
&= |\det(K)| \int \exp\left(-Y^{\mathsf{T}}V^{-1}Y + 2Y^{\mathsf{T}}V^{-1}BKG - G^{\mathsf{T}}G\right)\,dG \\
&= |\det(K)| \int \exp\left(-(G-K^{\mathsf{T}}B^{\mathsf{T}}V^{-1}Y)^{\mathsf{T}}(G-K^{\mathsf{T}}B^{\mathsf{T}}V^{-1}Y)\right)\,dG \\
&\qquad \exp\left(-Y^{\mathsf{T}}V^{-1}Y + Y^{\mathsf{T}}V^{-1}BKK^{\mathsf{T}}B^{\mathsf{T}}V^{-1}Y\right) \\
&= |\det(K)|\,(\pi)^{M/2} \exp\left(-Y^{\mathsf{T}}V^{-1/2}(I - V^{-1/2}BKK^{\mathsf{T}}B^{\mathsf{T}}V^{-1/2})V^{-1/2}Y\right)
\end{aligned}$$

where $\dim(Y) = N$, $\dim(F) = \dim(G) = M \leq N$, $F = KG$ and $K^{\mathsf{T}}B^{\mathsf{T}}V^{-1}BK = I$. This assumes that $\mathrm{rank}(B^{\mathsf{T}}V^{-1}B) = M$ so that $K = (B^{\mathsf{T}}V^{-1}B)^{-1/2}$.

For $M = N = \mathrm{rank}(B)$

$\mathrm{rank}(B^{\mathsf{T}}V^{-1}B) = N$ implies that $\mathrm{rank}(B) = N$ and $\mathrm{rank}(V) = N$. Consequently, using the SVD decompositions $B = U_B D_B W_B^{\mathsf{T}}$ and $V = U_V D_V U_V^{\mathsf{T}}$ gives $K = W_B D_B^{-1} U_B^{\mathsf{T}} U_V D_V^{1/2}$. Therefore $BK = U_V D_V^{1/2}$ and so $V^{-1/2}BKK^{\mathsf{T}}B^{\mathsf{T}}V^{-1/2} = I$, so that the integral is given by:

$$\int \exp\left(-(Y-BF)^{\mathsf{T}}V^{-1}(Y-BF)\right)\,dF = |\det(B^{\mathsf{T}}V^{-1}B)|^{-1/2}(\pi)^{M/2} \tag{17}$$

For $\mathrm{rank}(B) = M < N$**,**

$B^{\mathsf{T}}V^{-1}B$ is taken to have full rank, but $BKK^{\mathsf{T}}B^{\mathsf{T}}$ is not, as D_B is not a square matrix, so that D_B^{-1} does not exist. Instead, $K = W(D^{\mathsf{T}}D)^{-1/2}$ where $V^{-1/2}B = UDW^{\mathsf{T}}$ by SVD decomposition. As D has dimensions N by M (same as B) then $D^{\mathsf{T}}D$ is M by M and hence invertible (by virtue of $B^{\mathsf{T}}V^{-1}B$ being full rank). Therefore, $K^{\mathsf{T}}B^{\mathsf{T}}V^{-1}BK = (D^{\mathsf{T}}D)^{-1/2}(D^{\mathsf{T}}D)(D^{\mathsf{T}}D)^{-1/2} = I$ as desired, but $V^{-1/2}BKK^{\mathsf{T}}B^{\mathsf{T}}V^{-1/2} = UD(D^{\mathsf{T}}D)^{-1}D^{\mathsf{T}}U^{\mathsf{T}} = P_w$ which is a projection matrix. Consequently, $R_w = I - P_w = I - V^{-1/2}BKK^{\mathsf{T}}B^{\mathsf{T}}V^{-1/2} = I - V^{-1/2}B(B^{\mathsf{T}}V^{-1}B)^{-1}B^{\mathsf{T}}V^{-1/2}$ is not zero, but is the residual projection matix (onto the null space of $V^{-1/2}B$ – the prewhitened version of B). The integral can then be written as:

$$\begin{aligned}
\int \exp\left(-(Y-BF)^{\mathsf{T}}V^{-1}(Y-BF)\right)\,dF &= |\det(K)|(\pi)^{M/2} \exp\left(-Y^{\mathsf{T}}V^{-1/2}R_w V^{-1/2}Y\right) \\
&= |\det(B^{\mathsf{T}}V^{-1}B)|^{-1/2}(\pi)^{M/2} \exp\left(-Y^{\mathsf{T}}R_c Y\right) \qquad (18)
\end{aligned}$$

where $R_c = V^{-1/2}R_w V^{-1/2}$ is the residual projection matrix in the coloured space, and $R_w = I - V^{-1/2}B(B^{\mathsf{T}}V^{-1}B)^{-1}B^{\mathsf{T}}V^{-1/2}$ is the residual projection matrix in the whitened space.

For $\mathrm{rank}(B) < M$**,**

In this case the matrix B will have many linearly dependent columns and $K^{\mathsf{T}}B^{\mathsf{T}}V^{-1}BK = I$ cannot be achieved. Instead, let the number of independent columns be $N_1 \leq N$. Furthermore, let $V^{-1/2}B = UDW^{\mathsf{T}}$ by SVD, where

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix}$$

where D_1 is an N_1 by N_1 diagonal matrix with all diagonal entries being non-zero. Integration of the parameters associated with the zero singular values can be carried out if they have finite extents. Letting $F = KG$, and $K = W$ gives

$$\begin{aligned}
\int \exp\left(-(Y-BF)^{\mathsf{T}}V^{-1}(Y-BF)\right)\,dF &= \int \exp\left(-(V^{-1/2}Y - UDG)^{\mathsf{T}}(V^{-1/2}Y - UDG)\right)\,dG \\
&= L^{M-N_1}|\det(D_1^{\mathsf{T}}D_1)|^{-1/2}(\pi)^{M/2} \exp\left(-Y^{\mathsf{T}}R_c Y\right) \qquad (19)
\end{aligned}$$

where the null parameters are integrated over $[0, L]$, and $R_c = V^{-1} - V^{-1/2} U D_0 (D_1^\mathsf{T} D_1)^{-1} D_0^\mathsf{T} U^\mathsf{T} V^{-1/2}$, with

$$D_0 = \begin{bmatrix} D_1 \\ 0 \end{bmatrix}$$

which is an N by N_1 matrix.

A.3.4 Special case: White noise

White noise with variance σ^2 gives $V^{-1} = \frac{1}{2\sigma^2} I$. In this case the above integral becomes:

$$\int \exp\left(\frac{-1}{2\sigma^2}(Y - BF)^\mathsf{T}(Y - BF)\right) dF \quad = \quad |\det(B^\mathsf{T}B)|^{-1/2} (2\pi\sigma^2)^{M/2} \exp\left(\frac{-Y^\mathsf{T} R_w Y}{2\sigma^2}\right) \tag{20}$$

where $R_c = \frac{1}{2\sigma^2} R_w$ and $R_w = I - \frac{1}{2\sigma^2} BKK^\mathsf{T}B^\mathsf{T} = I - B(B^\mathsf{T}B)^{-1}B^\mathsf{T}$.

A.3.5 Special case: Scalar F and White noise

Consider F as a scalar, and $B = 1^\mathsf{T}$ (a column vector of ones). This represents BF being a constant (or mean) vector. Therefore, $B^\mathsf{T}B = N$ which gives the integral as

$$\int \exp\left(\frac{-1}{2\sigma^2}(Y - 1^\mathsf{T}F)^\mathsf{T}(Y - 1^\mathsf{T}F)\right) dF \quad = \quad (N)^{-1/2} (2\pi\sigma^2)^{1/2} \exp\left(\frac{-N\mathrm{Var}(Y)}{2\sigma^2}\right) \tag{21}$$

with $R_c = \frac{1}{2\sigma^2}(I - \frac{1}{N} BB^\mathsf{T})$. Note that $Y^\mathsf{T} R_c Y = \frac{N\mathrm{Var}(Y)}{2\sigma^2}$ where $\mathrm{Var}(Y) = \frac{1}{N}\sum_j (Y_j^2) - (\frac{1}{N}\sum_j Y_j)^2 = \frac{1}{N}(Y^\mathsf{T}Y - \frac{1}{N} Y^\mathsf{T} BB^\mathsf{T} Y)$ represents the estimated variance of the data vector, Y.

References

[1] J. Ashburner and K. Friston. Segmentation by integrating intensity-based classification and template matching. In *10th Int. Conf. on Human Brain Mapping*, 2004.

[2] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(6), 2001.

[3] M. Jenkinson, P.R. Bannister, J.M. Brady, and S.M. Smith. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841, 2002.

[4] A. Pitiot, H. Delingette, N. Ayache, and P. Thompson. Expert knowledge guided segmentation system for brain MR. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, number 2879 / 2003, 2003.

[5] A. Roche, G. Malandain, and N. Ayache. Unifying maximum likelihood approaches in medical image registration. *International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging*, 11(1):71–80, 2000.

[6] A. Roche, G. Malandain, X. Pennec, and N. Ayache. Multimodal image registration by maximization of the correlation ratio. Technical Report 3378, INRIA Sophia-Antipolis, 1998.

[7] J. Zhang and A. Rangarajan. Bayesian multimodality non-rigid image registration via conditional density estimation. In *International Conference on Information Processing in Medical Imaging (IPMI)*, pages 499–511, July 2003.

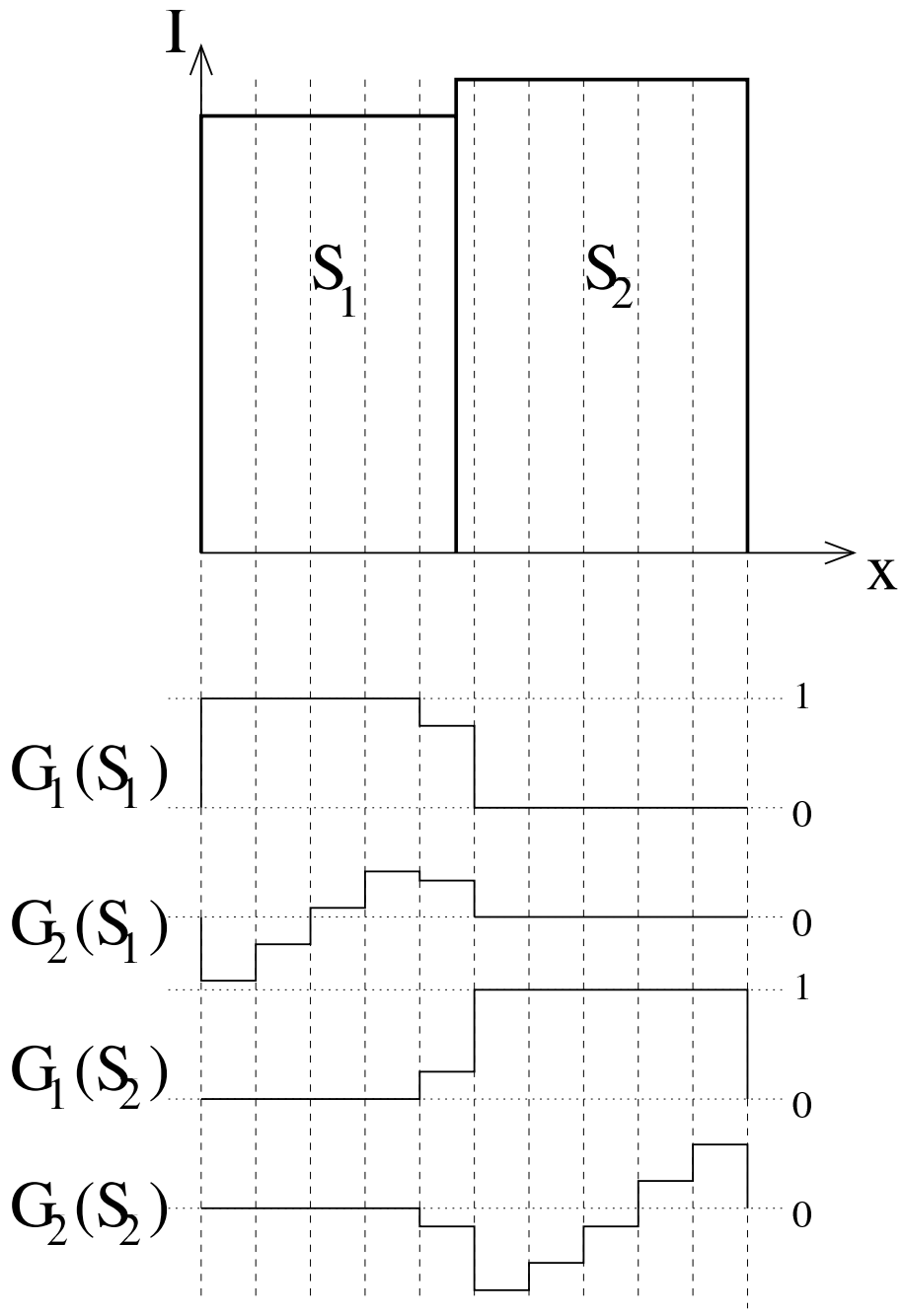

Figure 1: Example of image model formation in 1D. The two shapes, S_1 and S_2, cover ten image voxels (indicated by vertical dashed lines) and generate four different G vectors: $G_1(S_1)$ which represents the mean intensity (partial volume component) of S_1; $G_2(S_1)$ which represents the a linear intensity change across S_1 (in the x-direction); and similarly for S_2. The middle voxel is a partial volume voxel and shows how both the mean and linear components are multiplied by the appropriate partial volume fraction. Furthermore, note that the linear components are zero mean. Note that in 3D there would also be linear intensity changes across the y-direction and z-direction, represented by G_3 and G_4.

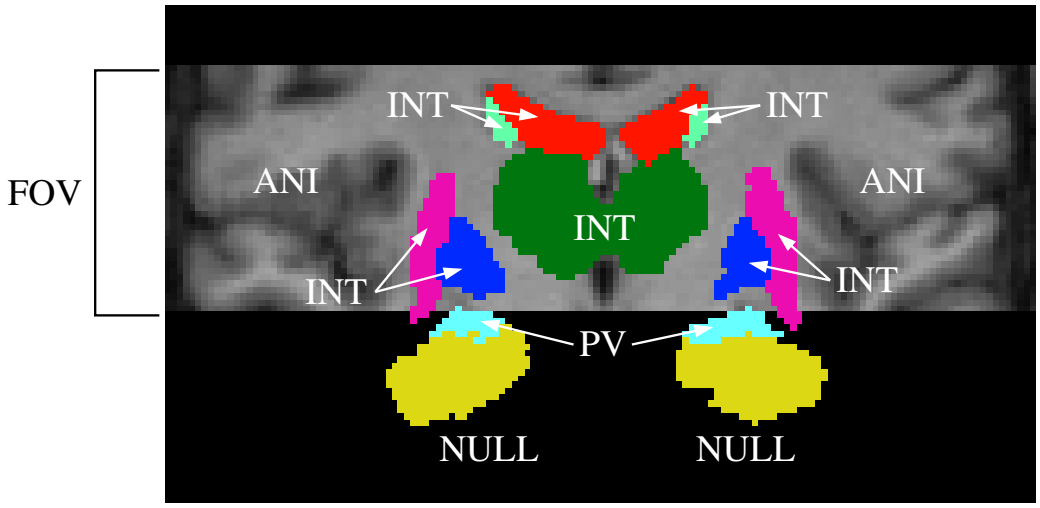

Figure 2: Example of shape models superimposed on an image showing (in one 2D coronal slice); Field of View (FOV), Shapes of Interest (INT), Areas of No Interest (ANI), Null Shapes (NULL), and Partial Volume Shapes (PV). Note that all parts of the observed image not covered by a shape model is part of the Area of No Interest.

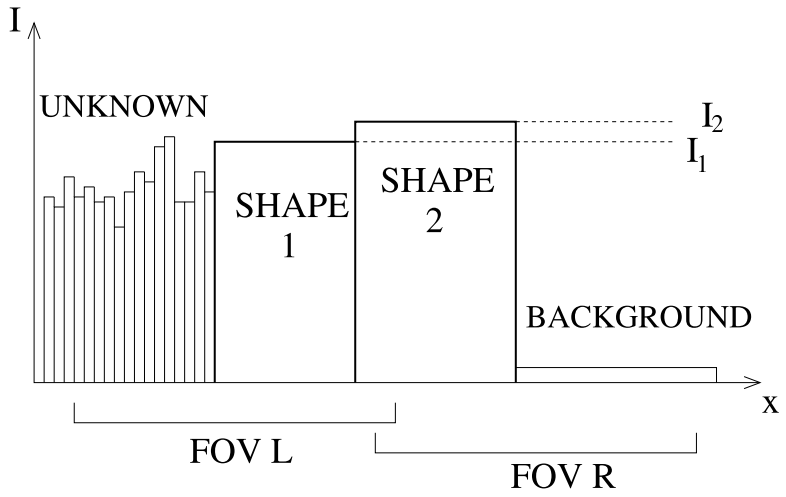

Figure 3: Shape model used for the experimental tests.

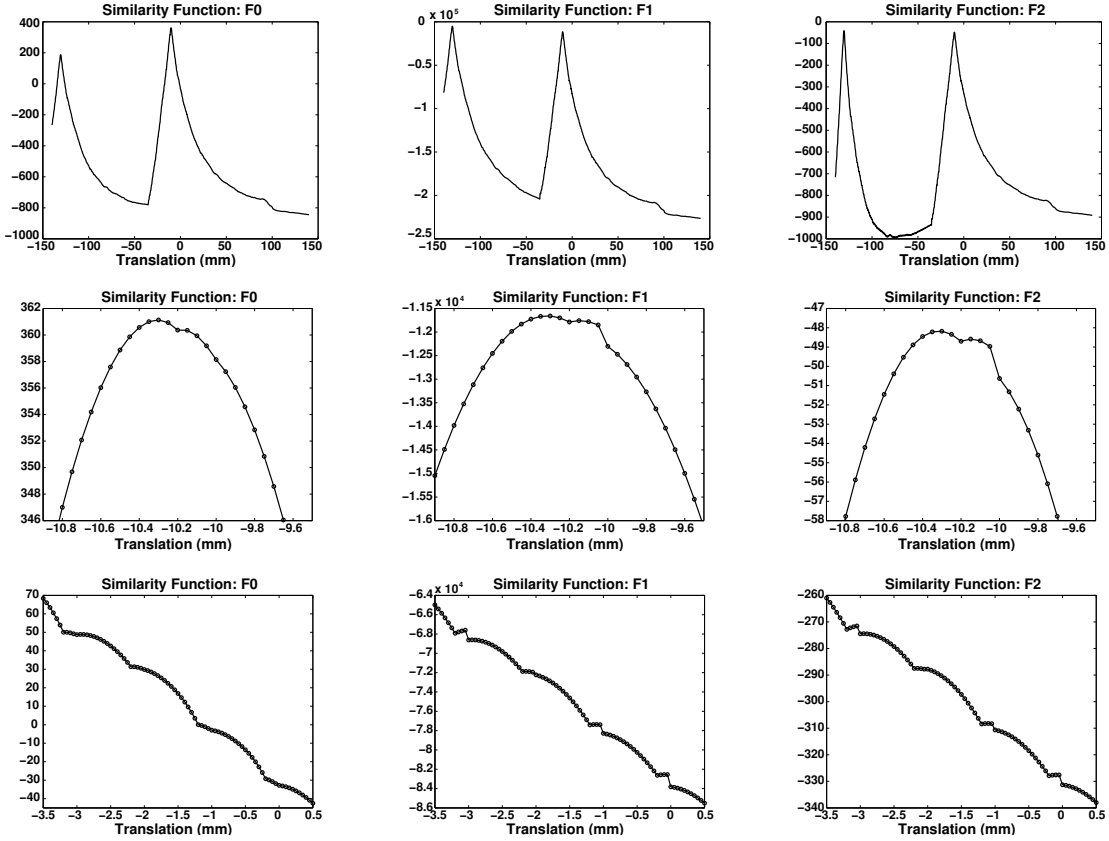

Figure 4: Example similarity functions for F_0 (left), F_1 (middle) and F_2 (right). The top row shows the functions over a large range of translations, the second and third rows show a close-up of the functions, with the actual calculated values represented by circles. The true translation value is -10.2mm for all of these plots.