

Tina Memo No. 2008-009

Published in Annals of the BMVA Vol 2008 no. 2(s1) p. 1-11, 2008.

<http://www.bmva.org/annals/2008/2008-0002-supplement.pdf>

# The TINA Medical Image Segmentation Algorithm: Mathematical Derivations and Proofs

P. A. Bromiley

Last updated  
21 / 01 / 2009



Imaging Science and Biomedical Engineering Division,  
Medical School, University of Manchester,  
Stopford Building, Oxford Road,  
Manchester, M13 9PT.

# The TINA Medical Image Segmentation Algorithm: Mathematical Derivations and Proofs

P. A. Bromiley  
Imaging Science and Biomedical Engineering Division  
Medical School, University of Manchester  
Manchester, M13 9PT, UK  
paul.bromiley@manchester.ac.uk

## Abstract

This document provides a number of mathematical proofs and derivations used in the development of the TINA medical image segmentation algorithm, but which have not been presented elsewhere.

## 1 Introduction

The TINA medical image segmentation algorithm is the result of an extensive development programme, described in TINA Memos no 2000-006, 2001-009, 2001-014, 2002-006, 2004-009, 2005-013, 2007-004, 2007-005, 2008-002 and 2008-003, and in several publications [3, 4, 5, 10, 11, 15, 16, 18]. The current algorithm is feature-space based, and uses a unified mathematical framework incorporating both intensity and local intensity gradients to generate a model of the feature-space distributions, which is then fitted to the image data using the EM algorithm [2, 8, 9]. The algorithm is capable of modelling distributions of arbitrary dimensionality, and so can be applied simultaneously to multiple, pre-registered images of the same biological structures.

Modelling of the intensity distributions requires components for both pure tissues and partial volumes. We follow the model first described by Santiago and Gage [13, 14], which is based on three assumptions:

- in the absence of noise and artefact generation processes, each pure tissue has a well-defined signal intensity;
- the image formation process is linear i.e. in partial volume voxels, each tissue contributes proportionately to the intensity of the voxel;
- tissue boundaries are uncorrelated with voxel boundaries.

Under these assumptions, pure tissues produce delta functions in the intensity space, which are convolved with some distribution representing both acquisition noise and any intrinsic variability of the tissue intensity. In the proposed algorithm, this distribution is assumed to be Gaussian, and so each pure tissue  $t$  has a probability density distribution given by

$$d_t(\mathbf{g}) = \alpha_t e^{-\frac{1}{2}(\mathbf{g}-\mathbf{M}_t)^T C_t^{-1}(\mathbf{g}-\mathbf{M}_t)}$$

where  $\mathbf{M}_t$  is a vector describing the tissue mean grey-level,  $C_t$  is the covariance matrix, and  $\alpha_t$  is a constant that provides unit normalisation. The intensity  $\mathbf{g}$  of a partial volume voxel is given by

$$\mathbf{g} = \sum_{i=1}^{N_T} h_i \mathbf{M}_i + \eta \quad (1)$$

where  $\mathbf{M}_i$  is the signal intensity of tissue  $i$ ,  $h_i$  is the volumetric proportion of tissue  $i$  in the voxel such that  $0 \leq h_i \leq 1$  and

$$\sum_{i=1}^{N_T} h_i = 1 \quad (2)$$

$N_T$  is the number of tissues in the image, and  $\eta$  is some noise distribution. In the case of partial volume voxels containing mixtures of only two pure tissues  $t$  and  $r$ , where  $h$  is the volumetric proportion of tissue  $t$ ,

$$\mathbf{g} = h\mathbf{M}_t + (1-h)\mathbf{M}_r + \eta \quad (3)$$

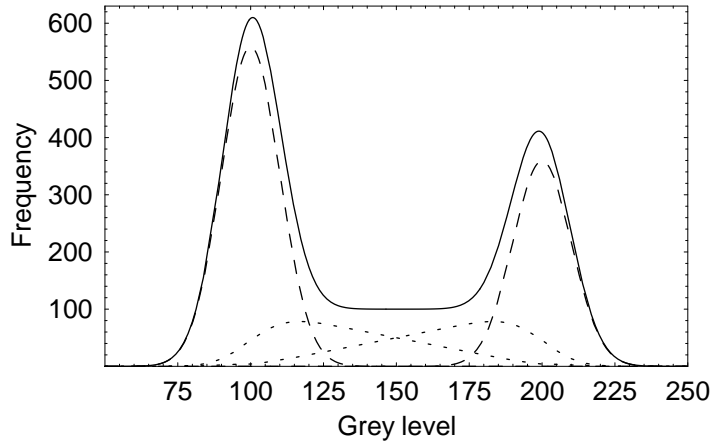


Figure 1: An example partial volume model for two pure tissues. Pure tissues have Gaussian distributions (dashed line), whereas mixtures of tissues take form of triangular distributions convolved with a Gaussian (dotted lines). The model components are summed to give the overall distribution (solid line).

Therefore, again assuming that the distribution  $\eta$  is a Gaussian distribution  $N$ , the probability density distribution for partial volume voxels is given by

$$d_{tr+rt}(\mathbf{g}; h) = N(\mathbf{g} - (h\mathbf{M}_t + (1-h)\mathbf{M}_r); C_{tr}) \quad (4)$$

and so

$$d_{tr+rt}(\mathbf{g}) = \int_0^1 d_{tr+rt}(\mathbf{g}; h) dh \quad (5)$$

If it is assumed that, following the assumptions listed above, all values of  $h$  are equally probable then the partial volume distribution takes the form of a uniform distribution convolved with a Gaussian. Unfortunately it has no closed form, and thus requires numerical integration.

The proposed algorithm considers only partial voluming between pairs of tissues i.e. triplets or quartets are ignored. In addition, in order to facilitate later volume estimation, each uniform partial volume distribution is decomposed into two complementary triangular distributions  $d_{tr}(\mathbf{g})$  and  $d_{rt}(\mathbf{g})$ , which sum to reproduce the uniform distribution. The order of the subscripts in this notation is significant:  $d_{tr}(\mathbf{g})$  represents the contribution of tissue  $t$  to the intensity distribution generated by voxels where it is partial volumed with tissue  $r$ . An example for two well-separated tissues in a 1D intensity space is shown in Fig. 1. In the case of a multi-dimensional intensity space, the model will take the form of a complementary pair of triangular distributions lying along the vector between the pure tissue intensities, convolved with a multi-dimensional Gaussian noise distribution. The covariance matrix of this noise distribution is a function of position along the partial volume vector

$$C_h = hC_t + (1-h)C_r \quad (6)$$

where  $C_t$  and  $C_r$  are the covariance matrices of the two pure tissue distributions and  $0 < h < 1$  is the fractional distance along the partial volume vector. In order to obtain  $h$  the intensity of the partial volume voxel must be projected onto the partial volume vector, using

$$h = \frac{(\mathbf{g} - \mathbf{M}_t)^T C_h^{-1} (\mathbf{M}_r - \mathbf{M}_t)}{(\mathbf{M}_r - \mathbf{M}_t)^T C_h^{-1} (\mathbf{M}_r - \mathbf{M}_t)} \quad (7)$$

The estimates of  $h$  and  $C_h$  are therefore inter-dependent and must be obtained using an iterative process. This is initiated assuming  $h = 0.5$ ;  $C_h$  is then calculated allowing re-estimation of  $h$ , and the process is iterated until convergence. Since  $C_h$  varies monotonically with  $h$  this process is stable and converges rapidly (typically within a few iterations).

If all distance measurements in the intensity space are weighted by the local covariance matrix, then we transform into an approximately variance-normalised (homoscedastic) space in which the covariance matrix of the multi-dimensional Gaussian is the identity matrix i.e. the standard deviation is unity along any of the axes. In this space the partial volume distribution can be decomposed into a product of components along any set of orthogonal axes. Decomposing into two components, one lying parallel to the vector between the pure tissue intensities and one

lying perpendicular to this, we obtain a particularly simple form; the first component is the product of a unit 1D Gaussian with a triangular distribution, and the second is a multi-dimensional unit Gaussian, so

$$d_{tr}(\mathbf{g}) = \beta_{tr} T_{tr}(h) e^{-\frac{z^2}{2}} \quad (8)$$

$\beta_{tr}$  enforces unit normalisation and  $z$  is the distance through which the partial volume intensity is projected to reach the vector between the pure tissue intensities

$$z^2 = (\mathbf{g} - \mathbf{M}_t)^T C_h^{-1} (\mathbf{g} - \mathbf{M}_t) - h^2 [(\mathbf{M}_r - \mathbf{M}_t)^T C_h^{-1} (\mathbf{M}_r - \mathbf{M}_t)] \quad (9)$$

$T_{tr}$  is the convolution of a triangular distribution normalised to 1/2 (since two complementary triangular distributions comprise each partial volume distribution) with a Gaussian distribution normalised to 1. The general form for a triangular distribution with start and end points  $a$  and  $b$ , intercept  $c$  and gradient  $k$  is

$$T_{tr}(x) = -\frac{kx + c}{2} \left[ \operatorname{erf}\left(\frac{x-b}{\sigma\sqrt{2}}\right) - \operatorname{erf}\left(\frac{x-a}{\sigma\sqrt{2}}\right) \right] - \frac{k\sigma}{\sqrt{2\pi}} \left[ e^{-\frac{(x-b)^2}{2\sigma^2}} - e^{-\frac{(x-a)^2}{2\sigma^2}} \right] \quad (10)$$

where  $\sigma$  is the standard deviation of the Gaussian. In the present case,  $\sigma = 1$ ,  $a = 0$ ,

$$b = \sqrt{(\mathbf{M}_r - \mathbf{M}_t)^T C_h^{-1} (\mathbf{M}_r - \mathbf{M}_t)} \quad (11)$$

and  $k$  and  $c$  are obtained trivially from  $b$  and the normalisation.

We now go on to derive or prove five important results: Eqs. 6, 7, 9, 10 and 11.

## 2 Equation 6: Interpolating the Covariance Matrix of Partial Volume Voxels

Under the linear image formation assumption adopted here (and by the majority of other authors working in the field), the intensity  $g$  of a partial volume voxel is given by a linear combination of the intensities of the pure tissues it contains, weighted by their volumetric proportions

$$\mathbf{g} = h\mathbf{M}_t + (1-h)\mathbf{M}_r \quad (12)$$

Some noise distribution must now be applied to this, and several different approaches are adopted in the literature. The simplest case is to assume that all noise is acquisition noise, in which case uniform Gaussian (or other) noise can be added after partial voluming. More complex approaches assume that the noise may be tissue dependent i.e. each pure tissue has some intrinsic intensity variability that is added to the signal prior to partial voluming. The latter approach is adopted in the proposed algorithm. Note that, since the sum of two Gaussians is itself a Gaussian (see the following derivation), the tissue-dependent noise distribution trivially incorporates acquisition noise.

There is considerable disagreement in the literature over the correct equation for the covariance matrix of a partial volume intensity given the covariance matrices of the pure tissue intensities in the tissue-dependent noise case. Using the 1D case for simplicity, and considering a partial volume vector containing a volumetric proportion  $h$  of pure tissue  $t$  with mean intensity  $M_t$  and standard deviation  $\sigma_t$ , and a volumetric proportion  $1-h$  of pure tissue  $r$  with mean intensity  $M_r$  and standard deviation  $\sigma_r$ , [14] and [1] state that the standard deviation of the partial volume intensity is

$$\sigma_h^2 = h\sigma_t^2 + (1-h)\sigma_r^2$$

However, [6], [12] and [17] state that the standard deviation of the partial volume intensity is

$$\sigma_h^2 = h^2\sigma_t^2 + (1-h)^2\sigma_r^2$$

([7] also state this result but do not use it). In this section we explain the origin of this controversy and demonstrate that the first form i.e. that originally quoted by [14] is correct, and that the second form leads to inconsistent results.

Treating the noise as tissue dependent, Eq. 12 becomes a linear combination of two Gaussians. The simplest derivation of the result relies on the moment generating function; for a random variable  $x$  with a probability density function  $p(x)$ , this is given by

$$\epsilon(t) = \langle e^{tx} \rangle = \int_{-\infty}^{\infty} e^{tx} p(x) dx = 1 + te'_1 + \frac{1}{2!} t^2 e'_2 + \dots$$

where  $e_r'$  is the  $r^{th}$  raw moment of the PDF. Note that, for independent  $X$  and  $Y$ , the moment generating function satisfies

$$\epsilon_{x+y}(t) = \langle e^{t(x+y)} \rangle = \langle e^{tx} \rangle \langle e^{ty} \rangle = \epsilon_x(t) \epsilon_y(t)$$

So, for a weighted sum of  $N_T$  independent variables

$$y = \sum_{i=1}^{N_T} h_i x_i$$

$$\epsilon(t) = \langle \exp(t \sum_{i=1}^{N_T} h_i x_i) \rangle = \langle e^{(th_1 x_1)} e^{(th_2 x_2)} \dots e^{(th_{N_T} x_{N_T})} \rangle = \prod_{i=1}^{N_T} \langle e^{(th_i x_i)} \rangle$$

The moment generating function of a normal distribution is given by

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \exp(tx) dx = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \quad (13)$$

giving

$$\epsilon(t) = \prod_{i=1}^{N_T} \langle e^{(th_i x_i)} \rangle = \prod_{i=1}^{N_T} \exp\left(\mu h_i t + \frac{\sigma_i^2 h_i^2 t^2}{2}\right) \quad (14)$$

Comparing Eqs. 13 and 14 shows that the weighted linear combination of several Gaussian distributions with means  $\mu_i$  and standard deviations  $\sigma_i$  is itself a Gaussian distribution with mean

$$\mu_{sum} = \sum_{i=1}^{N_T} h_i \mu_i$$

and standard deviation

$$\sigma_{sum}^2 = \sum_{i=1}^{N_T} h_i^2 \sigma_i^2 \quad (15)$$

Returning to the consideration of the partial volume voxel described above, it is tempting to substitute the pure tissue standard deviations  $\sigma_t$  and  $\sigma_r$  (or equivalently  $C_t$  and  $C_r$ ) directly into this equation to obtain

$$\sigma_h^2 = h^2 \sigma_t^2 + (1-h)^2 \sigma_r^2$$

However, consider applying this to the case of a pure tissue voxel treated as consisting of 50% tissue  $t$  with standard deviation  $\sigma_t$  and another 50% of the same pure tissue  $t$ . We should expect to obtain the result that  $\sigma_h = \sigma_t$ , but instead we obtain

$$\sigma_h^2 = 0.5^2 \sigma_t^2 + 0.5^2 \sigma_t^2 = \frac{\sigma_t^2}{2} \Rightarrow \sigma_h = \frac{\sigma_t}{\sqrt{2}} \quad (16)$$

which is clearly erroneous. The error arises because no consideration was made of the sampling process intrinsic to the signal generation process. The signal from an MR voxel is the average of the signal generated by all of the tissue within that voxel i.e. all of the magnetised spins. In the notation used here,  $\sigma_t$  and  $\sigma_r$  represent the noise on the signal from one voxel of pure tissues  $t$  and  $r$  respectively; let the number of spins in such voxels be given by  $N_{vs}$ .  $\sigma_t$  and  $\sigma_r$  are the standard deviations of a sample of  $N_{vs}$  spins drawn from an infinite population, and so relate to the standard deviation of the populations  $\sigma_{t,pop}$  and  $\sigma_{r,pop}$  by

$$\sigma_t = \frac{\sigma_{t,pop}}{\sqrt{N_{vs}}} \quad \text{and} \quad \sigma_r = \frac{\sigma_{r,pop}}{\sqrt{N_{vs}}}$$

In a partial volume voxel we have a smaller sample of  $N_h$  spins from pure tissue  $t$  and  $N_{(1-h)}$  spins from pure tissue  $r$ , where

$$h = \frac{N_h}{N_{vs}} \quad \text{and} \quad 1-h = \frac{N_{(1-h)}}{N_{vs}}$$

and so the standard deviations  $\sigma_{th}$  and  $\sigma_{r(1-h)}$  on the samples of each pure tissue within the partial volume voxel are given by

$$\sigma_{th} = \frac{\sigma_{t,pop}}{\sqrt{N_h}} \quad \text{and} \quad \sigma_{r(1-h)} = \frac{\sigma_{r,pop}}{\sqrt{N_{(1-h)}}}$$

Therefore,

$$\sigma_{th} = \frac{\sigma_t}{\sqrt{h}} \quad \text{and} \quad \sigma_{r(1-h)} = \frac{\sigma_r}{\sqrt{1-h}}$$

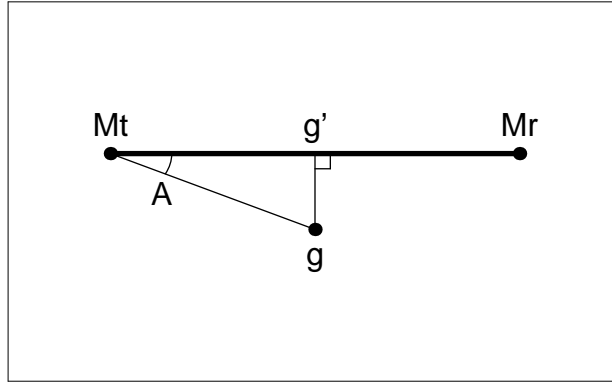


Figure 2: Projection of the intensity  $g$  of a partial volume voxel onto the vector between pure tissues with intensities  $M_t$  and  $M_r$ .

Substituting  $\sigma_{th}$  and  $\sigma_{r(1-h)}$  into Eq. 15 gives

$$\sigma_h^2 = h\sigma_t^2 + (1-h)\sigma_r^2$$

Applying this equation to a pure tissue voxel treated as a partial volume combination of the same tissues gives

$$\sigma_h^2 = 0.5\sigma_t^2 + 0.5\sigma_t^2 = \sigma_t^2 \quad i.e. \quad \sigma_h = \sigma_t$$

which is the correct solution. The multi-dimensional equivalent is therefore

$$C_h = hC_t + (1-h)C_r$$

### 3 Equations 7, 9 and 11: Projecting Partial Volume Voxels in Intensity Space

In one-dimensional intensity space, the noise on the partial volume voxels can only move the voxel intensity along the vector between the pure tissue intensities. However, in multi-dimensional intensity spaces, the noise can move the intensity off of this vector as well as along it. The situation is shown in Fig. 2. We need to find the position along the vector in order to calculate the mixing proportion  $h$ , and so we need to find the projection of the intensity back onto the vector.

Consider first the case of an equal-variance (homoscedastic) intensity space i.e. an intensity space in which the covariance matrix of the data is equal and circular at all points, and so does not need to be taken into account in order to perform the projection. The points  $\mathbf{g}$ ,  $\mathbf{M}_t$  and  $\mathbf{M}_r$  represent vectors in this space<sup>1</sup>. Let

$$\mathbf{k} = \mathbf{g} - \mathbf{M}_t \quad and \quad \mathbf{l} = \mathbf{M}_r - \mathbf{M}_t \quad (17)$$

Let the projection of  $\mathbf{k}$  onto  $\mathbf{l}$  i.e. the vector from  $\mathbf{M}_t$  to  $\mathbf{g}_h$  be called  $\mathbf{m}$ . The aim is to find

$$h = \frac{|\mathbf{m}|}{|\mathbf{l}|} \quad (18)$$

The dot product between  $\mathbf{k}$  and  $\mathbf{l}$  is

$$\mathbf{k} \cdot \mathbf{l} = |\mathbf{k}| |\mathbf{l}| \cos A \quad (19)$$

and, since  $\mathbf{g}$ ,  $\mathbf{g}_h$  and  $\mathbf{M}_t$  form a right-angled triangle,

$$|\mathbf{k}| \cos A = |\mathbf{m}| \quad (20)$$

<sup>1</sup>The homoscedastic space considered here is a variance-normalised transform of the heteroscedastic space encountered in practical image segmentation problems. In order to highlight the difference, we introduce at this point a clarification of the notation used in the main paper and in Section 1 of this document, such that vectors with the prime symbol represent those in the heteroscedastic space i.e. intensities measured from the images in the general case, and vectors without the prime symbol represent their equivalents in the homoscedastic i.e. variance-normalised space.

so

$$\mathbf{k} \cdot \mathbf{l} = |\mathbf{m}| |\mathbf{l}| \quad (21)$$

The dot product of  $\mathbf{l}$  with itself gives the square of the magnitude of  $\mathbf{l}$

$$\mathbf{l} \cdot \mathbf{l} = |\mathbf{l}|^2 \quad (22)$$

so

$$\frac{\mathbf{k} \cdot \mathbf{l}}{\mathbf{l} \cdot \mathbf{l}} = \frac{|\mathbf{m}| |\mathbf{l}|}{|\mathbf{l}|^2} = \frac{|\mathbf{m}|}{|\mathbf{l}|} = h \quad (23)$$

and so

$$h = \frac{(\mathbf{g} - \mathbf{M}_t) \cdot (\mathbf{M}_r - \mathbf{M}_t)}{(\mathbf{M}_r - \mathbf{M}_t) \cdot (\mathbf{M}_r - \mathbf{M}_t)} \quad (24)$$

Now consider the more general case in which the space is not homoscedastic. We wish to find the most probable projection of the partial volume intensity onto the vector between the pure tissue intensities rather than just the normal projection i.e. the point from which it is most probable that the particular partial volume intensity could have originated. The simplest way to do this is to transform into a homoscedastic space i.e. to weight the distance metric with the covariance matrix.  $C_h$  is a positive semi-definite matrix (since by definition variances and covariances must be real and lie in the range 0 to 1), and so must have an eigenvector decomposition. So, if  $\mathbf{E}$  is the matrix of eigenvectors,

$$\mathbf{C}_h \mathbf{E} = \mathbf{E} \mathbf{D} \quad (25)$$

where  $\mathbf{D}$  is a diagonal matrix whose diagonal elements are the variances in the rotated space. The aim is to rotate into this space, and then divide by the square-root of the matrix  $\mathbf{D}$  i.e. normalise distances along each axis of the space using the standard deviation along the axis. Let  $\mathbf{k}$ ,  $\mathbf{l}$  and  $\mathbf{m}$  represent vectors in the homoscedastic space as above, and  $\mathbf{k}'$ ,  $\mathbf{l}'$  and  $\mathbf{m}'$  represent the equivalent vectors in the heteroscedastic space. Multiplying the vectors in the heteroscedastic space by the matrix of eigenvectors rotates them into the new space, and dividing by the square-root of the matrix  $\mathbf{D}$  normalises them, so

$$\mathbf{k} = \mathbf{k}'^T \mathbf{E} \mathbf{D}^{-1/2} \quad \text{and} \quad \mathbf{l} = \mathbf{l}'^T \mathbf{E} \mathbf{D}^{-1/2} \quad \text{and} \quad \mathbf{m} = \mathbf{m}'^T \mathbf{E} \mathbf{D}^{-1/2}$$

Therefore

$$h = \frac{\mathbf{k} \cdot \mathbf{l}}{\mathbf{l} \cdot \mathbf{l}} = \frac{\mathbf{k}'^T \mathbf{E} \mathbf{D}^{-1} \mathbf{E}^T \mathbf{l}'}{\mathbf{l}'^T \mathbf{E} \mathbf{D}^{-1} \mathbf{E}^T \mathbf{l}'} \quad (26)$$

However,

$$\mathbf{C}_h = \mathbf{E} \mathbf{D} \mathbf{E}^{-1} \quad \text{so} \quad \mathbf{C}_h^{-1} = \mathbf{E} \mathbf{D}^{-1} \mathbf{E}^{-1} \quad (27)$$

and, since  $\mathbf{D}$  is by definition an orthogonal matrix, its inverse is equal to its transpose, and so the term

$$\mathbf{E} \mathbf{D}^{-1} \mathbf{E}^T = \mathbf{E} \mathbf{D}^{-1} \mathbf{E}^{-1} = \mathbf{C}_h^{-1} \quad (28)$$

and so

$$h = \frac{(\mathbf{g}' - \mathbf{M}'_t)^T \mathbf{C}_h^{-1} (\mathbf{M}'_r - \mathbf{M}'_t)}{(\mathbf{M}'_r - \mathbf{M}'_t)^T \mathbf{C}_h^{-1} (\mathbf{M}'_r - \mathbf{M}'_t)} \quad (29)$$

where  $\mathbf{g}'$ ,  $\mathbf{M}'_t$  and  $\mathbf{M}'_r$  represent intensities in the heteroscedastic space i.e. the intensities measured from the images in the general case.

We also need to obtain the length of the partial volume vector  $|\mathbf{M}'_r - \mathbf{M}'_t|$  (Eq. 11) and the distance through which the partial volume intensity is projected  $|\mathbf{g}_h - \mathbf{g}|$  (Eq. 9) in the homoscedastic space; let the latter be called  $|\mathbf{n}|$ . The first of these is the length of  $\mathbf{l}$  and is trivial to obtain from the dot product of the vector with itself

$$|\mathbf{l}| = \sqrt{\mathbf{l} \cdot \mathbf{l}} = \sqrt{(\mathbf{M}'_r - \mathbf{M}'_t)^T \mathbf{C}_h^{-1} (\mathbf{M}'_r - \mathbf{M}'_t)}$$

The second can be obtained by observing that, in the homoscedastic space, the vectors  $\mathbf{n}$ ,  $\mathbf{m}$  and  $\mathbf{k}$  form a right-angled triangle in which  $\mathbf{k}$  is the hypotenuse, and that  $|\mathbf{m}| = h |\mathbf{l}|$

$$|\mathbf{n}| = \sqrt{[(\mathbf{g}' - \mathbf{M}'_t)^T \mathbf{C}_h^{-1} (\mathbf{g}' - \mathbf{M}'_t)] - h^2 [(\mathbf{M}'_r - \mathbf{M}'_t)^T \mathbf{C}_h^{-1} (\mathbf{M}'_r - \mathbf{M}'_t)]}$$

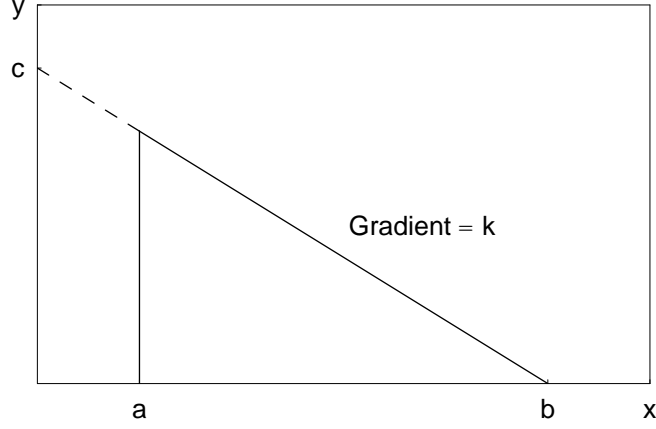


Figure 3: The parameters of a triangular distribution.

## 4 Equation 10: Convolution of a Gaussian Distribution with a Triangular Distribution

Let  $a$  and  $b$  represent the non-zero range of a triangular distribution,  $c$  its intercept and  $k$  its gradient, as shown in Fig. 3. Let  $\sigma$  represent the standard deviation of the normal distribution. The triangular distribution can then be written as

$$kx + c \quad (30)$$

and the normal distribution as

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (31)$$

The convolution of the two distributions is then given by

$$\int_a^b (kt + c) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-x)^2}{2\sigma^2}} dt \quad (32)$$

where the limits  $a$  and  $b$  can be imposed on the integral since this is the non-zero range of the triangular distribution. Integrating by parts (letting the triangular term be the term to be differentiated and the normal distribution be the term to be integrated) gives

$$\frac{kt + c}{2} \operatorname{erf} \left( \frac{t-x}{\sqrt{2}\sigma} \right) \Big|_a^b - \int_a^b \frac{k}{2} \operatorname{erf} \left( \frac{t-x}{\sqrt{2}\sigma} \right) dt \quad (33)$$

The integral of the erf function is given by

$$\int \operatorname{erf}(u) du = u \operatorname{erf}(u) + \frac{e^{-u^2}}{\sqrt{\pi}} \quad (34)$$

Therefore,

$$\frac{kt + c}{2} \operatorname{erf} \left( \frac{t-x}{\sqrt{2}\sigma} \right) \Big|_a^b - \frac{k\sigma}{\sqrt{2}} \left[ \frac{t-x}{\sqrt{2}\sigma} \operatorname{erf} \left( \frac{t-x}{\sqrt{2}\sigma} \right) + \frac{1}{\sqrt{\pi}} e^{-\frac{(t-x)^2}{2\sigma^2}} \right] \Big|_a^b \quad (35)$$

Evaluating the limits and rearranging gives

$$-\frac{kx + c}{2} \left[ \operatorname{erf} \left( \frac{x-b}{\sqrt{2}\sigma} \right) - \operatorname{erf} \left( \frac{x-a}{\sqrt{2}\sigma} \right) \right] - \frac{k\sigma}{\sqrt{2\pi}} \left[ e^{-\frac{(x-b)^2}{2\sigma^2}} - e^{-\frac{(x-a)^2}{2\sigma^2}} \right] \quad (36)$$

## References

- [1] M A G Ballester, A P Zisserman, and M Brady. Estimation of the partial volume effect in MRI. *Med. Im. Anal.*, 6:389–405, 2002.
- [2] C M Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.



- [3] P A Bromiley and N A Thacker. When less is more: Improvements in medical image segmentation through spatial sub-sampling. In *Proc. MIUA 2007*, pages 131–135, Aberystwyth, Wales, 17-18 July 2007, 2007.
- [4] P A Bromiley and N A Thacker. Improving the accuracy of mr image segmentation through the use of local gradient information. In *Proc. MIUA 2008*, pages 84–88, Dundee, Scotland, 2-3 July 2008, 2008.
- [5] P A Bromiley, N A Thacker, M L J Scott, M Pokrić, A J Lacey, and T F Cootes. Bayesian and non-Bayesian probabilistic models for medical image analysis. *Image and Vision Computing*, 21(10):851–864, 2003.
- [6] J P Chiverton. *Probabilistic Partial Volume Modelling of Biomedical Tomographic Image Data*. PhD thesis, Center for Vision, Speech and Signal Processing, University of Surrey, Guildford, Surrey GU2 7HX, U.K., Aug 2006.
- [7] H S Choi, D R Haynor, and Y Kim. Partial volume tissue classification of multichannel magnetic resonance images—a mixel model. *IEEE Trans Med Imaging*, 10(3):395–407, 1991.
- [8] A P Dempster, N M Laird, and D B Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Society*, 39:1–38, 1977.
- [9] R Guillemaud and J M Brady. Estimating the bias field of MR images. *IEEE Trans Med Imaging*, 16(3):238–251, 1997.
- [10] M Pokrić, N A Thacker, and A Jackson. The importance of partial voluming in multi-dimensional medical image segmentation. In *Proc. MICCAI’01*, pages 1293–1294, 2001.
- [11] M Pokrić, N A Thacker, M L J Scott, and A Jackson. Multi-dimensional medical image segmentation with partial voluming. In *Proc. MIUA’01*, pages 77–80, 2001.
- [12] S Ruan, C Jaggi, J Xue, J Fadili, and D Bloyet. Brain tissue classification of magnetic resonance images using partial volume modeling. *IEEE Trans Med Imaging*, 19(12):1179–1187, 2000.
- [13] P Santago and H D Gage. Quantification of MR brain images by mixture density and partial volume modelling. *IEEE Trans Med Imaging*, 12:566–574, 1993.
- [14] P Santago and H D Gage. Statistical models of partial volume effect. *IEEE Trans Image Processing*, 4:1531–1540, 1995.
- [15] N A Thacker, M Pokrić, and D C Williamson. Noise filtering and testing illustrated using a multi-dimensional partial volume model of MR data. In *Proc. BMVC*, pages 909–919, Kingston, London, 2004.
- [16] N A Thacker, D C Williamson, and M Pokric. Voxel based analysis of tissue volume from mri data. *British Journal of Radiology*, 77:S114–S125, 2004.
- [17] J Tohka, E Krestyannikov, I D Dinov, A M Graham, D W Shattuck, U Ruotsalainen, and A W Toga. Genetic algorithms for finite mixture model based voxel classification in neuroimaging. *IEEE Trans Med Imaging*, 26(5):696–711, 2007.
- [18] D C Williamson, N A Thacker, S R Williams, and M Pokrić. Partial volume tissue segmentation using grey-level gradient. In *Proc. MIUA’02*, pages 17–20, 2002.