Multi-dimensional Medical Image Segmentation with Partial Volume and Gradient Modelling Supplement 1: Mathematical Derivations and Proofs

P.A. Bromiley and N.A. Thacker

Division of Imaging Science and Biomedical Engineering,
School of Medicine, University of Manchester, M13 9PT, U.K
⟨paul.bromiley@manchester.ac.uk⟩

Abstract

This document constitutes supplementary material for the paper "Multi-dimensional Medical Image Segmentation with Partial Volume and Gradient Modelling", and contains mathematical derivations and proofs for several aspects of the image intensity model developed therein.

1 Introduction

Modelling of the intensity distributions of magnetic resonance (MR) images requires components for both pure tissue and partial volume voxels. We adopt the model first described by Santago and Gage [1993, 1995], 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 α_t is a constant that provides unit normalisation. The intensity $\mathbf{g}$ of a partial volume voxel is given by

$$\mathbf{g} = \sum_{i=1}^{N_T} h_i \mathbf{M_i} + \eta \tag{1}$$

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

$$\sum_{i=1}^{N_T} h_i = 1 \tag{2}$$

N_T is the number of tissues in the image, and η 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,

$$\mathbf{g} = h\mathbf{M_t} + (1-h)\mathbf{M_r} + \eta \tag{3}$$

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

$$d_{tr+rt}(\mathbf{g};h) = N(\mathbf{g} - (h\mathbf{M_t} + (1-h)\mathbf{M_r}); C_{tr}) \tag{4}$$

and so

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

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
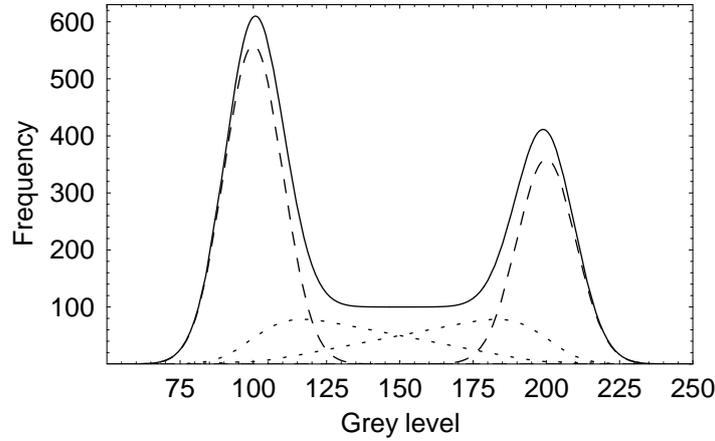
$$C_h = hC_t + (1-h)C_r \tag{6}$$

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).

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

$$h = \frac{(\mathbf{g} - \mathbf{M_t})^T C_h^{-1} (\mathbf{M_r} - \mathbf{M_t})}{(\mathbf{M_r} - \mathbf{M_t})^T C_h^{-1} (\mathbf{M_r} - \mathbf{M_t})} \tag{7}$$

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

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

β_{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

$$z^2 = (\mathbf{g} - \mathbf{M_t})^T C_h^{-1} (\mathbf{g} - \mathbf{M_t}) - h^2 [(\mathbf{M_r} - \mathbf{M_t})^T C_h^{-1} (\mathbf{M_r} - \mathbf{M_t})] \tag{9}$$

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

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

where σ 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 σ_t, and a volumetric proportion $1-h$ of pure tissue r with mean intensity M_r and standard deviation σ_r, Santago and Gage [1995] and Ballester et al. [2002] state that the standard deviation of the partial volume intensity is

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

However, Chiverton [2006], Ruan et al. [2000] and Tohka et al. [2007] state that the standard deviation of the partial volume intensity is

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

(Choi et al. [1991] 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 Santago and Gage [1995] 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

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

where e'_r is the rth raw moment of the PDF. Note that, for independent X and Y, the moment generating function satisfies

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

So, for a weighted sum of N_T independent variables

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

$$\epsilon(t) = <exp(t\sum_{i=1}^{N_T} h_i x_i)> = <e^{(th_1 x_1)}e^{(th_1 x_1)}e^{(th_1 x_1)}> = \prod_{i=1}^{N_T} <e^{(th_i x_i)}>$$

The moment generating function of a normal distribution is given by

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

giving

$$\epsilon(t) = \prod_{i=1}^{N_T} <e^{(th_i x_i)}> = \prod_{i=1}^{N_T} exp(\mu h_i t + \frac{\sigma^2 h_i^2 t^2}{2}) \tag{14}$$

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

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

and standard deviation

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

Returning to the consideration of the partial volume voxel described above, it is tempting to substitute the pure tissue standard deviations σ_t and σ_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 σ_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

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

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, σ_t and σ_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}. σ_t and σ_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

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

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

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

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

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

Therefore,

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

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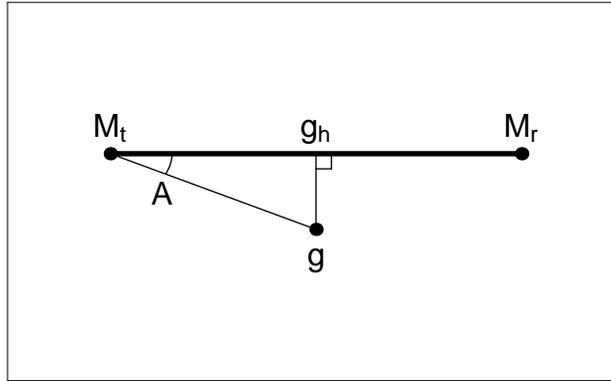
Figure 2: Projection of the intensity g of a partial volume voxel onto the vector between pure tissues with intensities M_t and M_r.

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[1]. Let

$$\mathbf{k} = \mathbf{g} - \mathbf{M_t} \quad and \quad \mathbf{l} = \mathbf{M_r} - \mathbf{M_t} \tag{17}$$

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

$$h = \frac{|\mathbf{m}|}{|\mathbf{l}|} \tag{18}$$

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

$$\mathbf{k}.\mathbf{l} = |\mathbf{k}||\mathbf{l}| \cos A \tag{19}$$

and, since $\mathbf{g}$, $\mathbf{g_h}$ and $\mathbf{M_t}$ form a right-angled triangle,

$$|\mathbf{k}| \cos A = |\mathbf{m}| \tag{20}$$

so

$$\mathbf{k}.\mathbf{l} = |\mathbf{m}||\mathbf{l}| \tag{21}$$

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

$$\mathbf{l}.\mathbf{l} = |\mathbf{l}|^2 \tag{22}$$

so

$$\frac{\mathbf{k}.\mathbf{l}}{\mathbf{l}.\mathbf{l}} = \frac{|\mathbf{m}||\mathbf{l}|}{|\mathbf{l}|^2} = \frac{|\mathbf{m}|}{|\mathbf{l}|} = h \tag{23}$$

---

[1]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.

and so

$$h = \frac{(\mathbf{g} - \mathbf{M_t}).(\mathbf{M_r} - \mathbf{M_t})}{(\mathbf{M_r} - \mathbf{M_t}).(\mathbf{M_r} - \mathbf{M_t})} \tag{24}$$

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,

$$\mathbf{C_h E} = \mathbf{ED} \tag{25}$$

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

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

Therefore

$$h = \frac{\mathbf{k}.\mathbf{l}}{\mathbf{l}.\mathbf{l}} = \frac{\mathbf{k}'^T \mathbf{ED}^{-1} \mathbf{E}^T \mathbf{l}'}{\mathbf{l}'^T \mathbf{ED}^{-1} \mathbf{E}^T \mathbf{l}'} \tag{26}$$

However,

$$\mathbf{C_h} = \mathbf{EDE}^{-1} \quad so \quad \mathbf{C_h}^{-1} = \mathbf{ED}^{-1} \mathbf{E}^{-1} \tag{27}$$

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

$$\mathbf{ED}^{-1} \mathbf{E}^T = \mathbf{ED}^{-1} \mathbf{E}^{-1} = \mathbf{C_h}^{-1} \tag{28}$$

and so

$$h = \frac{(\mathbf{g}' - \mathbf{M_t'})^T \mathbf{C_h}^{-1} (\mathbf{M_r'} - \mathbf{M_t'})}{(\mathbf{M_r'} - \mathbf{M_t'})^T \mathbf{C_h}^{-1} (\mathbf{M_r'} - \mathbf{M_t'})} \tag{29}$$

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

$$|\mathbf{l}| = \sqrt{\mathbf{l}.\mathbf{l}} = \sqrt{(\mathbf{M_r'} - \mathbf{M_t'})^T C_h^{-1} (\mathbf{M_r'} - \mathbf{M_t'})}$$

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 $|\mathbf{m}| = h|\mathbf{l}|$

$$|\mathbf{n}| = \sqrt{[(\mathbf{g}' - \mathbf{M_t'})^T C_h^{-1} (\mathbf{g}' - \mathbf{M_t'})] - h^2[(\mathbf{M_r'} - \mathbf{M_t'})^T C_h^{-1} (\mathbf{M_r'} - \mathbf{M_t'})]}$$
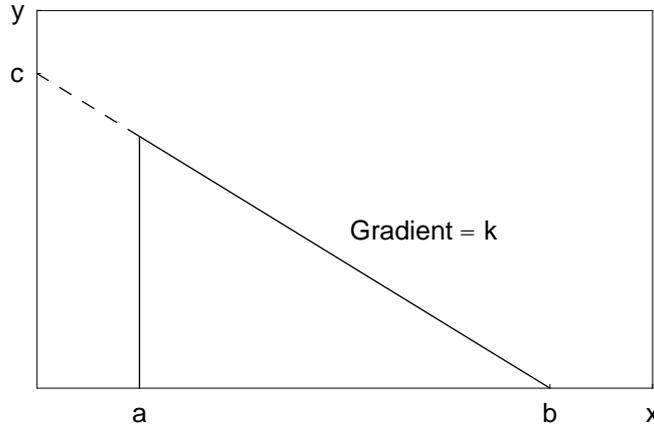
Figure 3: The parameters of a triangular distribution.

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 σ represent the standard deviation of the normal distribution. The triangular distribution can then be written as

$$kx + c \tag{30}$$

and the normal distribution as

$$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}} \tag{31}$$

The convolution of the two distributions is then given by

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

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

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

The integral of the erf function is given by

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

Therefore,

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

Evaluating the limits and rearranging gives

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

References

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.

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.

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.

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.

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.

P Santago and H D Gage. Statistical models of partial volume effect. *IEEE Trans Image Processing*, 4: 1531–1540, 1995.

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.