Restoring functional PET Images using Anatomical MR Images

Peter Alshede Philipsen, Ulrik Kjems, Peter Toft and Lars Kai Hansen,
Department of Mathematical Modelling, Section for Digital Signal Processing,
Building 321, Technical University of Denmark, DK-2800 Lyngby, Denmark
email: pap,uk,pto,lkh@imm.dtu.dk

Abstract

In this paper we present a Bayesian method to enhance functional 3D PET images using apriori knowledge about the brain anatomy obtained from 3D MR images. We use a Markov Random Field as a prior model, which have the possibility of incorporating this information. We show both a small test example as well as a true PET-MR result, and further more show how to obtain the desired information from the MR images.

1 Introduction

Low signal to noise ratio and the low spatial resolution in functional Positron Emission Tomography (PET) images is a major limitation in brain imaging. Therefore it is desirable to use apriori knowledge about the brain to address this. We present a method using a Bayesian approach to include new structural apriori information from Magnetic Resonance (MR) images. This is possible because the tissue types in the brain, white matter, gray matter and cerebrospinal fluid (CSF), have different activations levels, in average gray matter is 4 times higher than white and CSF has virtually no activation. Using this information extracted from the MR images from the same person gives apriori information about activation edges in the PET image [Chen,84].

Magnetic Resonance Imaging is a technique using changes in the magnetic field to measure structural information, because the different tissue types have different magnetic responses. Using different input response functions it is possible to enhance specific tissue types. No functional information is obtained with MR.

Positron Emission Tomography (PET) imaging is based on radioactive tracers injected in the persons blood. The tracer can e.g. be O-15 and a part of water H_2O. The PET image gives a picture of the blood flow in the brain, which is correlated with the brain activity. In PET imaging there is an intrinsic blurring, because the positron diffuses in average up to 4 mm before annihilating with an electron.

Reconstruction of imagery of a non-ideal imaging system is a fundamental aim of computer vision. Geman and Geman [Geman and Geman,84] introduced Metropolis sampling from Gibbs distributions as a simulation tool for visual reconstruction and showed that a *Simulated Annealing* strategy could improve the efficiency of the sampling process. The sampling process implements a stochastic neural network with symmetric connections. Peterson and Anderson applied the *Mean Field* approximation, and observed substantial improvements in speed and performance [Peterson,87].

In the next section the Bayesian approach to reconstruction is outlined, as model example we study a *Markov Random Field* model.

Section three contains experiments and section four concluding remarks.

2 Bayesian Visual Reconstruction

The basic idea is to consider both the source signal and the processes of the imaging system as stochastic processes. The Bayes formula can then be used to obtain the distribution $P(\mathbf{V}|\mathbf{d})$ of the reconstructed signal $\mathbf{V}$, conditioned on the observed degraded signal $\mathbf{d}$:

$$P(\mathbf{V}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{V})P(\mathbf{V})}{P(\mathbf{d})} \tag{1}$$

This conditional distribution is the product of the distribution of the imaging system process, also called the observation model: $P(\mathbf{d}|\mathbf{V}) \equiv P(\mathbf{V} \to \mathbf{d})$, and the *prior* distribution of the reconstructed signal $P(\mathbf{V})$. $P(\mathbf{V}|\mathbf{d})$ of Equation (1) is referred to as the *posterior* distribution. Under a given observation, $\mathbf{d}$, $P(\mathbf{d})$ is constant and the posterior distribution is directly proportional to $P(\mathbf{d}|\mathbf{V})$ and $P(\mathbf{V})$.

A useful estimate of the reconstructed signal is given by the *Maximum A Posterior* estimate. Several strategies like simulated annealing and Monte Carlo simulating can be used [Blake,89]. We use a mean field approximation to the simulated annealing, to obtain a more efficient algorithm.

2.1 Markov Random Field Model

In this section a description of the prior distribution is given. The prior model is a Markov Random Field, which is a Gibbs-Markov distribution. The signal is indexed V_i where $i = 0, .., I-1$. Introducing edges p_i, q_i and r_i connected to edges in the x, y and z direction respectively. The edge element p_i have the label {Edge} if the signal elements V_i and V_{i+dx} are independent and similar for q_i in the y direction between V_i and V_{i+dy}, and r_i in z-direction. A 2D part of the model can have the following outline.

$V_{i-dx+dy}$	$p_{i-dx+dy}$	V_{i+dy}	p_{i+dy}	$V_{i+dx+dy}$
q_{i-dx}		q_i		q_{i+dx}
V_{i-dx}	p_{i-dx}	V_i	p_i	V_{i+dx}
$q_{i-dx-dy}$		q_{i-dy}		$q_{i+dx-dy}$
$V_{i-dx-dy}$	$p_{i-dx-dy}$	V_{i-dy}	p_{i-dy}	$V_{i+dx-dy}$

The binary labels for the edges is {Edge, Non-Edge}, implemented as $\{+1, -1\}$.

For simplicity the prior model is first derived without dependence between the edges. In the entire model a simple smoothness between the neighbouring intensity values is used, unless an edge-element is active and allows a discontinuity. In this paper text a simple second order smoothness between neighbour voxels is used.

The prior MRF distribution can be more or less complicated depending on the complexity of the observed model [Hertz,91, Chen,84].

The Gibbs distribution has the following form:

$$P(\mathbf{V},\mathbf{p},\mathbf{q},\mathbf{r}\,|\,\theta) = Z_1^{-1}\exp\left(-E_{\text{prior}}(\mathbf{V},\mathbf{p},\mathbf{q},\mathbf{r}|\theta)\right), \tag{2}$$

where $E_{\text{prior}}(\mathbf{V},\mathbf{p},\mathbf{q},\mathbf{r}|\theta)$ is the energy or cost function for the MRF model, θ are the model parameters and Z_1^{-1} is a normalisation constant.

In 3D the prior energy function has the following form

$$\begin{aligned} E_{\text{prior}}(\mathbf{V},\mathbf{p},\mathbf{q},\mathbf{r}\,|\,\boldsymbol{\theta}) &= \frac{\nu}{2}\sum_i (1-p_i)(V_i - V_{i+dx})^2 + \frac{\nu}{2}\sum_i (1-q_i)(V_i - V_{i+dy})^2 \\ &+ \frac{\nu}{2}\sum_i (1-r_i)(V_i - V_{i+dz})^2 + \mu\sum_i p_i + \mu\sum_i q_i + \mu\sum_i r_i \end{aligned} \tag{3}$$

where the model parameters, $\boldsymbol{\theta} = \{\sigma^2, \nu, \mu\}$, all are positive. The last 3 terms are called chemical potentials and are incorporated to control the number of active edges. In this case the value of edge elements p_i , q_i and r_i are set using information from the MR image.

2.2 Observation model

The observation model: $P(\mathbf{d}|\mathbf{V}) \equiv P(\mathbf{V} \to \mathbf{d})$ in (1) is the model describing how the observed data is created from $\mathbf{V}$. In general a non-linear transformation of $\mathbf{V}$ is examined, called $\vartheta(\cdot)$ and describing the mapping $\mathcal{R}^I \xrightarrow{\vartheta} \mathcal{R}^J$.

The distribution may be calculated as

$$p(\mathbf{d}|\mathbf{V}) = p_{noise}(\mathbf{d} - \vartheta(\mathbf{V})) \tag{4}$$

where $p_{noise}(\mathbf{e})$ is the noise distribution of the residual $\mathbf{e} = \mathbf{d} - \vartheta(\mathbf{V})$.

In this paper we only consider a linear transformation $\varphi(\cdot)$ and additive spatially independent Gaussian noise.

We look at the observations model for each point d_j of $\mathbf{d}$.

$$d_j = \varphi_j(\mathbf{V}) + e_j \tag{5}$$

where $\varphi_j(\mathbf{V}) = \sum_i \varphi_{ji} V_i$ and $e_j \in \mathcal{N}(0, \sigma_j^2)$ are independent for each d_j.

Inserting in (4) gives

$$\begin{aligned} p(d_j|\mathbf{V}) &= p_{noise}(\delta(d_j - \varphi_j(\mathbf{V}))) &(6)\\ &= \frac{1}{\sqrt{2\pi\sigma_j^2}}\exp\left(-\frac{(d_j - \varphi_j(\mathbf{V}))^2}{2\sigma_j^2}\right) &(7) \end{aligned}$$

The $p(d_j|\mathbf{V})$ are independent and therefore

$$\begin{aligned} p(\mathbf{d}|\mathbf{V}) &= \prod_j p(d_j|\mathbf{V}) && (8)\\ &= Z_2 \exp\left(-\sum_j \frac{(d_j-\varphi_j(\mathbf{V}))^2}{2\sigma_j^2}\right) && (9)\\ &= Z_2 \exp\left(-E_{noise}(\mathbf{d}\,|\,\mathbf{V}\,,\boldsymbol{\theta})\right) && (10)\end{aligned}$$

where $Z_2 = \prod_j \frac{1}{\sqrt{2\pi\sigma_j^2}}$ and $E_{noise}(\mathbf{d}\,|\,\mathbf{V}\,,\boldsymbol{\theta}) = \sum_j \frac{(d_j-\varphi_j(\mathbf{V}))^2}{2\sigma_j^2}$

2.3 3D system model

The system model in 3D then have the following form:

$$\begin{aligned} P(\mathbf{V}\,|\mathbf{d},\,\mathbf{p},\mathbf{q},\mathbf{r}\,,\boldsymbol{\theta}) &\propto \exp\left(-E_{\text{prior}}(\mathbf{V},\mathbf{p},\mathbf{q},\mathbf{r}\,|\,\boldsymbol{\theta})\right)\cdot\exp\left(-E_{\text{noise}}(\mathbf{d}\,|\,\mathbf{V}\,,\boldsymbol{\theta})\right) && (11)\\ &= \exp\left(-E_{\text{total}}(\mathbf{V},\mathbf{d},\mathbf{p},\mathbf{q},\mathbf{r}\,|\,\boldsymbol{\theta})\right) && (12)\end{aligned}$$

where $E_{\text{total}}(\mathbf{V},\mathbf{d},\mathbf{p},\mathbf{q},\mathbf{r}\,|\,\boldsymbol{\theta}) = E_{\text{prior}}(\mathbf{V},\mathbf{p},\mathbf{q},\mathbf{r}\,|\,\boldsymbol{\theta}) + E_{\text{noise}}(\mathbf{V}\,|\,\mathbf{d}\,,\boldsymbol{\theta})$ is the energy or cost function for the system model and $\boldsymbol{\theta} = \{\sigma^2,\nu,\mu\}$ are the model parameters, all positive.

In this case the value of edge elements p_i, q_i and r_i are set using information from the MR image.

Since the edge elements are fixed based on the MR scan, only an updating formula for $\mathbf{V}$ is needed. The MAP estimate is where E_{total} have its minimum. A simplified description of E_{total} gives an easier derivation.

$$E_{total} = \mathbf{V}^T\boldsymbol{\Sigma}_{\mathbf{MRF}}\mathbf{V} + E_{MRF-edge-control} + (\mathbf{d}-\mathbf{AV})^T\boldsymbol{\Sigma}_{\mathbf{gauss}}(\mathbf{d}-\mathbf{AV}) \qquad (13)$$

where $E_{MRF-edge-control} = \sum_i \mu_i^p p_i + \sum_i \mu_i^q q_i + \sum_i \mu_i^r r_i$ and $\boldsymbol{\Sigma}_{\mathbf{MRF}}$ describes the MRF model, and the linear mapping $\varphi(\mathbf{V}) = \mathbf{AV}$. If each voxel is observed with the same noise level, then $\boldsymbol{\Sigma}_{\mathbf{gauss}} = \frac{1}{\sigma^2}\mathbf{I}$.

For the minimum of E_{total} the derivative $\frac{\partial E_{total}}{\partial V_i} = 0$ for all values of i. Using a mean field approach, and assuming that $V_{i'\neq i}$ is independent of V_i during update of V_i. Isolation of V_i gives.

$$\begin{aligned} V_i &= k^{-1}\left[\mathbf{1}_i^T\frac{1}{2\sigma^2}\mathbf{A}^T(\mathbf{d}-\mathbf{AV}) + kV_i - \mathbf{1}_i^T\boldsymbol{\Sigma}_{\mathbf{MRF}}\mathbf{V}\right] && (14)\\ \text{where } k &= \mathbf{1}_i^T\left(\boldsymbol{\Sigma}_{\mathbf{MRF}} + \frac{1}{2\sigma^2}\mathbf{A}^T\mathbf{A}\right)\mathbf{1}_i && (15)\end{aligned}$$

and $\mathbf{1}_i$ is a zero vector with only element i set to 1. In (14) the right side is independent of V_i. Assuming that the linear transformation is a Gaussian blurring $\mathbf{A}^T = \mathbf{A}$ the updating scheme can be formulated as a double Gauss filtering.

3 Experiments

In this section a few examples using the restoration technique are described. And some of the considerations needed to extract desired edge information. First a small toy example.

3.1 A small example.

In this example a $6 \times 6 \times 6$ cube with intensity value of 100 on a $16 \times 16 \times 16$ background, is blurred with a Gaussian kernel and Gaussian noise is added. The blurring kernel is symmetric with standard deviation 1.5 and the standard deviation of the noise is 5. A slice of the observed signal is in figure (1), and the corresponding restored slice is in figure (2). The edges are fixed on the true edges. Profiles through the centre of the cube can be seen in figure (3).

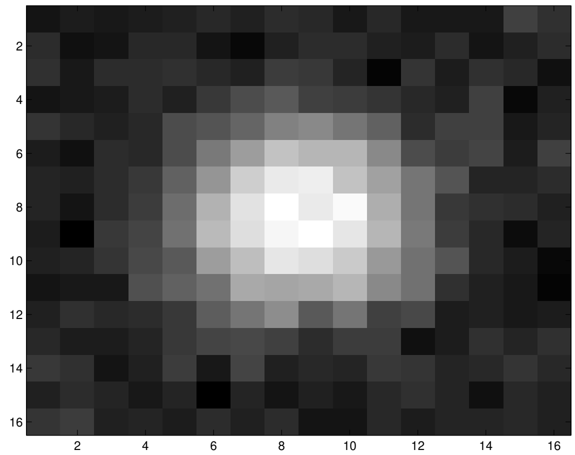

Figure 1: Slice of observed cube, blurred and noise are added.

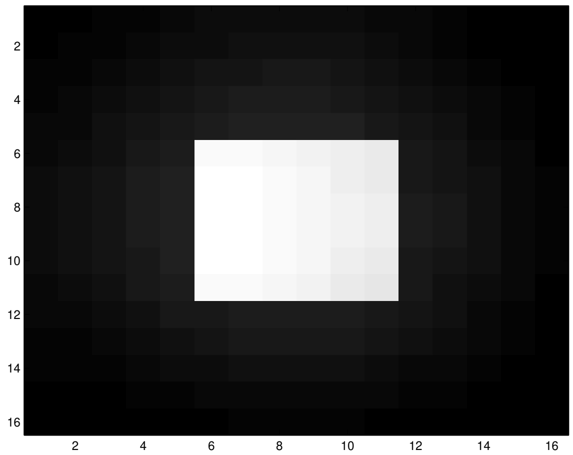

Figure 2: Slice of the restored image.

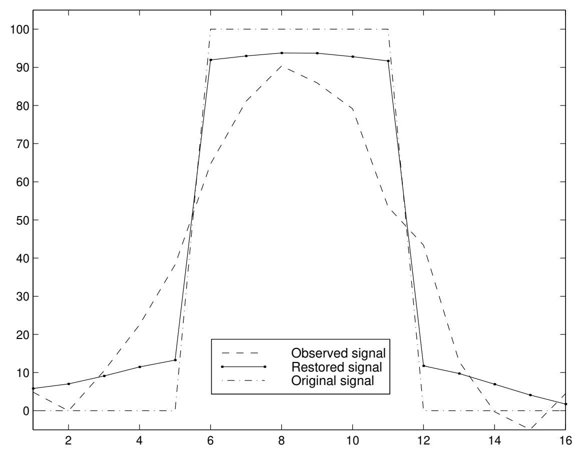

Figure 3: Profiles through the centre of the observed and restored image for the toy example.

3.2 Pre-processing MR Images

To obtain the desired edge information from the raw MR images shown in figure(4), several steps are needed. Focusing on edges between different tissue types based the on knowledge that tissue types in the brain have different activations levels. To obtain a tissue segmented brain several steps are necessary.

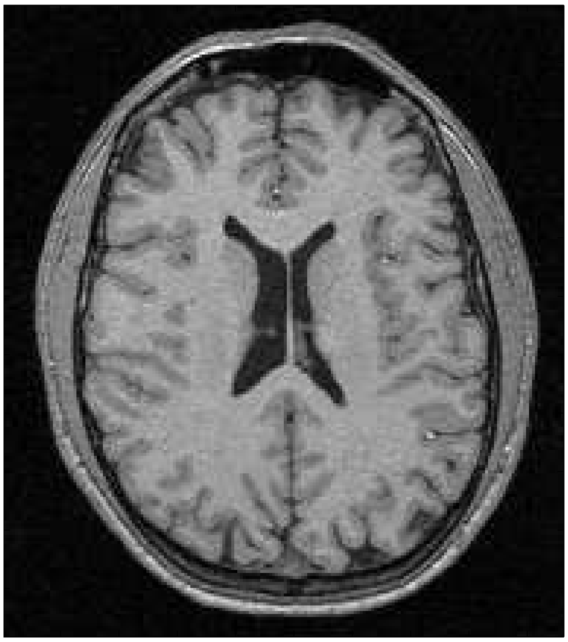

Figure 4: Slice of the Raw MR image, through centre of the brain.

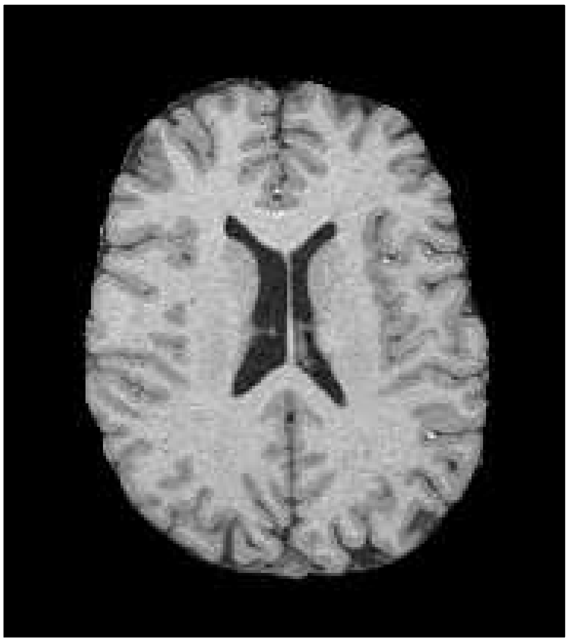

Figure 5: The same slice only containing brain voxels.

First a segmentation in brain/non-brain is done as shown in figure(5). This is needed to obtain a better alignment between PET and MR and to improve segmentation/classification of the brain. Alignment is done using AIR, [Woods,93], which is a 6 parameter transformation determining translation and rotation.

Next step is a segmentation of the brain to obtain volumes of the different tissue types. In this example we segment into CSF, white matter, cortex gray matter and subcortical gray matter, figure(6), using a Gaussian classification. To improve classification gray matter is separated into 2 types, but in the following these are merged again, because we have no information about difference in PET activation in these 2 types of gray matter.

From the segmented MR image a set of edge volumes $(\mathbf{p}, \mathbf{q}, \mathbf{r})$ are created based only on gray matter and white matter, figure(7).

3.3 Restoring PET image.

In figure (8) a slice of the resampled PET image is shown. The PET image is smoothed during the alignment and the resampling to MR resolution.

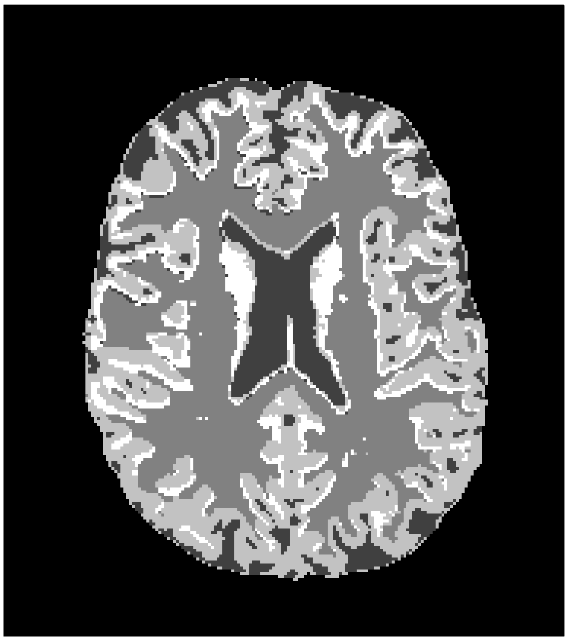

Figure 6: Slice of the segmented MR image, gray matter are the two brightest, then white matter and CSF.

Figure 7: Slice of edges created from MR image, ignoring CSF and merging the two types of gray matter.

Figure 8: Slice of the original PET image, smooth due to resampling to MR resolution.

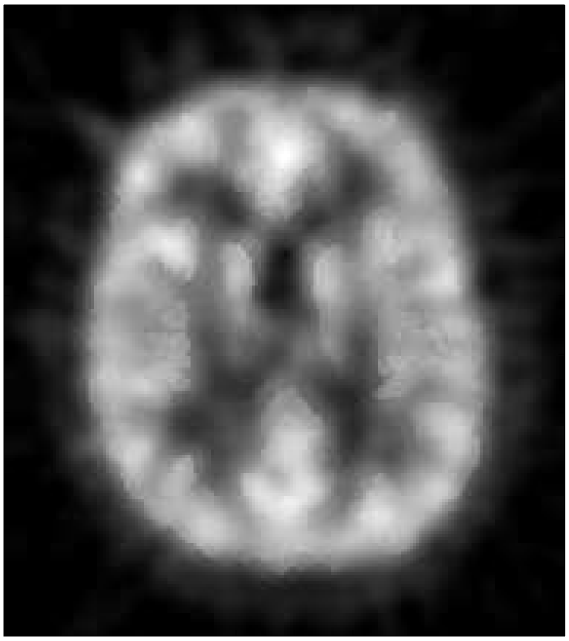

Figure 9: Slice of Restored PET image with a Gauss blurring on 3 mm and a noise level on 2.

A result of this strategy can be seen in figure (9). The Gaussian blurring was set to 3 mm with a noise level on 2. Running 100 iterations with an image size of $190 \times 215 \times 176$ voxels uses 220 Mb ram on a Silicon Graphics ONYX and less than 1 cpu hour. It was possible to preserve/enhance edges and avoid edge blurring on the outside of the brain.

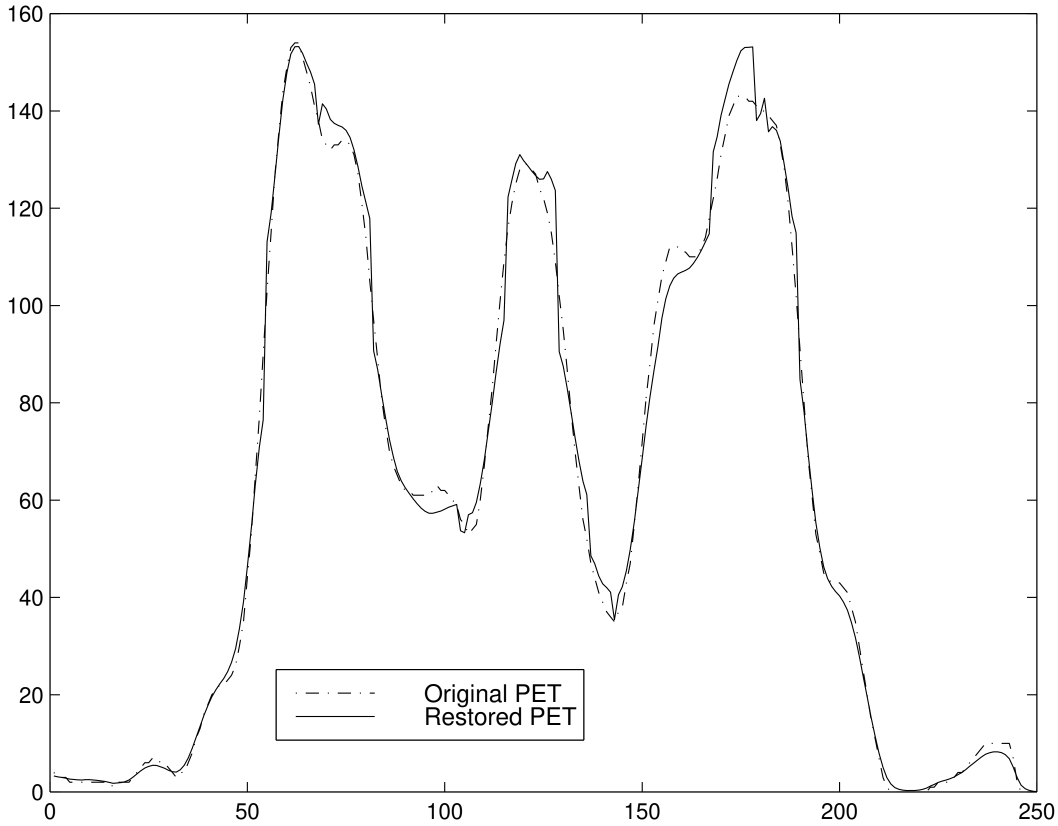

Figure 10: Profile through pet scans.

4 Concluding remarks

In this article we presented a method to enhance the resolution in functional positron emission tomography (PET) images, using anatomical information of the brain from MR images of a higher resolution by using a Markov Random Field (MRF) model which have the possibility of incorporating this information. This gives a new possibility of smoothing the very noisy PET images without blurring edges, and at the same time recovering the intrinsic blurring. In this scheme we do not preserve edges within tissue types, but an extension of the prior model can overcome this problem.

References

[Blake,89] A. Blake: *Comparison of the Efficiency of Deterministic and Stochastic Algorithms for Visual Reconstruction.* IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-11** 2-12 (1989).

[Chen,84] X. Ouyang, W.H. Wong, V.E. Johnson, X. Hu, C.-T. Chen: *Incorporation of Correlated Structural Images in PET Image Reconstruction.* IEEE Transactions on Medical Imaging **13** 627-640 (1994).

[Geman and Geman,84] D. Geman and S. Geman: *Stochastic Relaxation, Gibbs distributions and the Bayesian restoration of images.* IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-6** 721-741 (1984).

[Hertz,91] J. Hertz, A. Krogh and R.G. Palmer: *Introduction to the Theory of Neural Computation.* Addison Wesley, New York (1991).

[Peterson,87] C. Peterson and J.R. Anderson: *A Mean Field Learning Algorithm for Neural Networks.* Complex Systems **1** 995-1019, (1987).

[Roth,90] M.W. Roth: *Survey of Neural Network Technology for Automatic Target Recognition.* IEEE Transactions on Neural Networks **1**, 28-43, (1990).

[Woods,93] R. P. Woods and S. R. Cherry: *MRI-PET registration with automated algorithm.* Journal of Computer Assisted Tomography, 17(4):536-546, (1993).