4-D Generative Model for PET/MRI Reconstruction

Stefano Pedemonte[1], Alexandre Bousse[2], Brian F. Hutton[2], Simon Arridge[1], and Sebastien Ourselin[1]

[1] The Centre for Medial Image Computing, UCL, London, United Kingdom
[2] Institute of Nuclear Medicine, UCL Hospitals NHS Trust, London, United Kingdom

Abstract. We introduce a 4-dimensional joint generative probabilistic model for estimation of activity in a PET/MRI imaging system. The model is based on a mixture of Gaussians, relating time dependent activity and MRI image intensity to a hidden static variable, allowing one to estimate jointly activity, the parameters that capture the interdependence of the two images and motion parameters. An iterative algorithm for optimisation of the model is described. Noisy simulation data, modeling 3-D patient head movements, is obtained with realistic PET and MRI simulators and with a brain phantom from the BrainWeb database. Joint estimation of activity and motion parameters within the same framework allows us to use information from the MRI images to improve the activity estimate in terms of noise and recovery.

Keywords: Molecular Imaging, Emission Tomography, Motion correction, Multi-modality, Bayesian Networks.

1 Introduction

Resolution of pharmaceutical concentration in emission tomography is limited by photon count statistics and by motion of the patient [1]. The recent development of imaging systems that combine Emission Tomography and MRI in the same machine is enabling new biological and pathological analysis tools for clinical and pre-clinical research [2]. Inherent co-registration and simultaneity of the acquisitions introduce a number of advantages over the separate modalities, including improved image fusion, motion correction and enhancement of the resolution of the functional image [3], posing new algorithmic and computational challenges. Several publications have focused on improvement of the activity estimate by means of an intra-patient anatomical image, for combined or sequential imaging systems, relying on the assumption that activity is related to the underlying anatomy, which is linked to the MRI or CT image intensity. Methods in the literature fall into three main categories: methods that favor a piecewise uniform reconstruction by segmenting the anatomical image and subsequently applying a smoothing prior within each identified region [3]; methods that explicitly extract boundary information from the anatomical image and relax the effect of a global smoothing prior across the identified edges [4] and methods

based on information theoretic similarity functionals [5]. All such methods assume perfect co-registration of the two images and motion free acquisitions; however in brain imaging typical translations in the range of 5-20 mm and rotations of 1-4 deg are observed during PET and SPECT scans, concentrated in large sporadic movements occurring at discrete times [1], justifying event-driven motion compensation methods based on motion information provided by an external motion-tracking device [6]. Less successful methods to correct for patient movements involve the division of the scan into a number of frames, followed by spatial registration of the reconstructed images, relying on registration of highly photon limited image frames. In the context of combined PET/MRI we propose an event-driven motion compensation algorithm based on the acquisition of an MRI image each time large motion is detected by a tracking system or by an MRI navigator. Inspired by the Unified Segmentation [7] framework, in order to describe the interdependence of the variables in the multi-modal system we introduce a unified framework based on a 4-D joint generative model that allows us to estimate motion parameters from the MRI image and from the photon counts, obtaining a time consistent estimate of activity that accounts for its relation with the underlying anatomy, imaged by MRI.

2 Method

The model that is described is an extension to four dimensions of a bivariate Gaussian Mixture (GM) where the interdependence of the two imaging modalities is captured by assuming that image intensity y produced by the MRI sequence and activity λ are the expression of a hidden discrete variable $k = 1, 2, .., N_k$ representing anatomical/functional states (such as healthy gray matter, healthy white matter, hypoactive gray matter). k is considered the only reason of covariability of y and λ: $y \perp \lambda | k$. In order to account for a number of reasons of uncertainty of y (electronic noise, partial volume) and of λ (density of receptors, active radio-pharmaceutical molecules, perfusion), the probability distribution function (pdf) of y and λ is assumed to be a Gaussian of unknown parameters conditionally to k: $p(\lambda_b|k) = \mathcal{N}(\lambda_b, \theta_\lambda)$ and $p(y_b|k) = \mathcal{N}(y_b, \theta_y)$, where b indexes voxels and θ are the parameters of the Gaussians.

Considering the time dimension, with t indexing discrete time frames corresponding to detected motion events, motion is described by considering that the hidden states move. The hidden state is defined at the centre of voxel locations $X_b = (X_{b,1}, X_{b,2}, X_{b,3})$ in a reference space at time $t = 0$, assumed to be the instant the scan starts. Motion at time t warps the anatomical/functional state: $k^{[t]}(Y_b^{[t]}) = k(X_b)$, with $Y_b^{[t]} = T_{\gamma^{[t]}} X_b$ being the coordinates of body space warped by transformation $T_{\gamma^{[t]}}$ of parameter $\gamma^{[t]}$. As the body (the hidden state) deforms over time, at time t the MRI intensity in $Y_b^{[t]}$ is related to the hidden variable in X_b, being a realisation of the random process described by the GM of parameters $\theta_y^{[t]} = (\mu_{y_k^{[t]}}, \sigma_{y_k^{[t]}})$. For compactness $y_b^{[t]} = y^{[t]}(T_{\gamma^{[t]}} X_b)$; $\lambda_b = \lambda(X_b)$:

$$p(y_b^{[t]}|k^{[t]}(T_\gamma^{[t]} X_b)) = p(y_b^{[t]}|k(X_b)) = \mathcal{N}(y_b^{[t]}, \mu_{y_k^{[t]}}, \sigma_{y_k^{[t]}}) \quad (1)$$

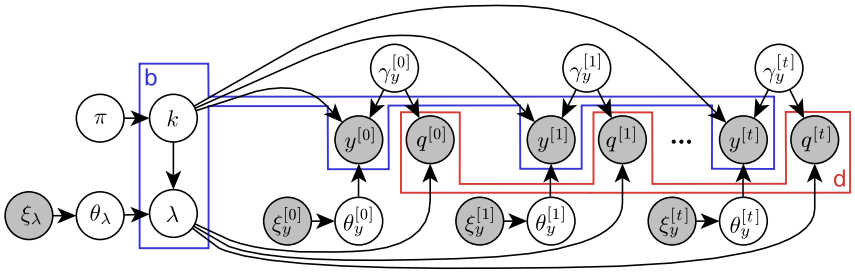

Fig. 1. Directed Acyclical Graph (DAG) of the proposed model. Observed (known) quantities are shaded. Voxels are indexed by b and lines of response (LOR) of the PET detector are indexed by d. The hidden anatomical/functional state k determines activity λ and, along with deformation parameters $\gamma^{[t]}$, determines MRI intensity $y^{[t]}$ at each time frame t. The dependence of activity and MRI intensity from the hidden state is parameterised by θ, with prior distribution of parameters ξ. Activity and deformation parameters determine photon counts $q^{[t]}$ at time t.

Activity in a voxel $Y_b^{[t]}$ at time t, assuming steady state distribution of the pharmaceutical, is a single realisation of the random process that relates it to the hidden state: $\lambda^{[1]}(T_{\gamma^{[1]}}X_b) = \lambda^{[2]}(T_{\gamma^{[2]}}X_b) = .. = \lambda(X_b)$. Photon counts $q_d^{[t]}$ along line of response (LOR) d are a realisation of the Poisson process with expected value given by the activity warped by the transformation at time t, or equivalently by activity λ_b and a time dependent transition matrix $P^{[t]} = \{p_{bd}^{[t]}\}$

$$p(\lambda_b|k) = \mathcal{N}(\lambda_b, \mu_{\lambda_k}, \sigma_{\lambda_k}) \qquad p(q^{[t]}|\lambda) = \prod_d \mathcal{P}(\sum_b p_{bd}^{[t]}\lambda_b, q_d^{[t]}) \tag{2}$$

Regarding all parameters as random variables, this model is represented by the Directed Acyclical Graph (DAG) in figure 1, where π is a multinomial distribution of the hidden state k and is here assumed to be spatially independent and unknown; ξ are hyper-parameters of the conjugate prior distribution for each of the mixture parameters. The deformation parameters $\gamma^{[t]}$ are assumed to be independent from one another conditionally to y and q as motion is considered unpredictable. Solving for the most likely state of the system provides motion parameters, activity at time $t = 0$ (and at all times when warped by $T_{\gamma^{[t]}}$) and fuzzy multi-modal segmentation $p(k_b|\lambda_b, y_b^{[t]})$ according to any combination of the MRI time frames, in fact any of $y_b^{[t]}$ (and λ) can be ignored when estimating the probability of the discrete states in a voxel. To summarise, the model accounts for the existence of an underlying tissue variable that explains the dependence of the two imaging modalities; the warping of tissue due to motion is described by a parametric deformation model. The deformation parameters, activity (in the reference space) and the parameters of the bivariate GM that describes the MR imaging system and the uncertainty about radio-pharmaceutical concentration are unknown. For a given estimate of the parameters, the marginal $p(k_b|\lambda_b, y_b^{[t]}, \theta_y, \gamma_y, \theta_\lambda)$ represents a probabilistic classification of the tissue in the reference space, according to the MRI and the projections at all time frames.

2.1 Inference

We aim at computing a point estimate of the unknown variables that maximises the joint *pdf*. Reasoning on the graph (Markov blankets), marginalising over k, all variables are independent of motion parameter $\gamma^{[t]}$ conditionally to $\lambda, \theta_y^{[t]}, \pi$; λ is independent of all other variables given $\theta_\lambda, \pi, q, \gamma$; $\theta_y^{[t]}$ is independent of all other variables conditionally to $y^{[t]}, \gamma^{[t]}, \pi$ and π is independent of all other variables given $\lambda, y, \gamma, \theta$. Starting from an initial estimate of each variable, always increasing joint *pdf* is obtained by updating the unknowns iteratively by the Iterated Conditional Modes (ICM) algorithm [8], converging to a local minimum. For compactness the expected value of the hidden state in b given λ_b, $y_b^{[t]}$ and all the parameters is defined as z_{bk}

$$z_{bk} = p(k_b|\lambda_b, y_b^{[1]}, \dots, y_b^{[t]}) = \frac{\pi_k \mathcal{N}(\lambda_b, \mu_{\lambda_k}, \sigma_{\lambda_k}) \prod_t \mathcal{N}(y_b^{[t]}, \mu_{y_k^{[t]}}, \sigma_{y_k^{[t]}})}{\sum_{k=1}^{N_k} \pi_k \mathcal{N}(\lambda_b, \mu_{\lambda_k}, \sigma_{\lambda_k}) \prod_t \mathcal{N}(y_b^{[t]}, \mu_{y_k^{[t]}}, \sigma_{y_k^{[t]}})}$$

(i) $p(\gamma^{[t]}|y, q, \theta, \lambda, \pi)$ Given activity and the GM parameters, the deformation parameters are updated by maximising the partial probability distribution by gradient ascent. For each t: $\gamma_g^{[t](n+1)} = \gamma_g^{[t](n)} + \beta \frac{\partial \log p(\gamma^{[t]}|y,q,\theta,\lambda,\pi)}{\partial \gamma_g^{[t]}}\Big|_{\gamma^{[t](n)}}$

$$p(\gamma^{[t]}|y, q, \theta, \lambda, \pi) = \prod_b \left[\sum_k p(k) p(y_b^{[t]}|k, \theta) p(\lambda_b|k, \theta) \prod_{r \neq t} p(y_b^{[r]}|k, \theta) \right] p(q^{[t]}|\lambda, \gamma^{[t]})$$

Substituting (1) and (2) and differentiating its logarithm, by the chain rule, it simplifies to the following:

$$\frac{\partial \log p(\gamma^{[t]}|y, q, \theta, \lambda, \pi)}{\partial \gamma_g^{[t]}} = \sum_b \left[\frac{\partial y^{[t]}(T_{\gamma^{[t]}} X_b)}{\partial \gamma_g^{[t]}} \sum_k z_{bk} \frac{\mu_{y_k} - y^{[t]}(T_{\gamma^{[t]}} X_b)}{\sigma_{y_k}^2} \right] + \sum_d \left[\sum_b -p_{bd} \frac{\partial \lambda(T_{\gamma^{[t]}}^{-1} X_b)}{\partial \gamma_g^{[t]}} + q_d^{[t]} \frac{p_{bd} \frac{\partial \lambda(T_{\gamma^{[t]}}^{-1} X_b)}{\partial \gamma_g^{[t]}}}{\sum_{b'} p_{b'd} \lambda(T_{\gamma^{[t]}}^{-1} X_{b'})} \right]$$

where the gradient of the MRI image intensity with respect to the transformation parameters is, by the chain rule, the dot product of the gradient of the transformed image and the gradient of the coordinates Y_b with respect to the transformation parameters:

$$\frac{\partial y^{[t]}(T_{\gamma^{[t]}} X_b)}{\partial \gamma_g^{[t]}} = \nabla y^{[t]}(T_{\gamma^{[t]}} X_b) \cdot \left(\frac{\partial}{\partial \gamma_g^{[t]}} T_{\gamma^{[t]}} \right) X_b$$

The same derivation applies to the gradient of the activity with respect to the transformation parameters.

(ii) $p(\lambda|\gamma, \pi, q, \theta) = p(\lambda|\gamma, y, \pi, \theta) p(q|\lambda, \gamma)$ Given the deformations and the GM parameters, activity is estimated by the EM algorithm for Poisson likelihood,

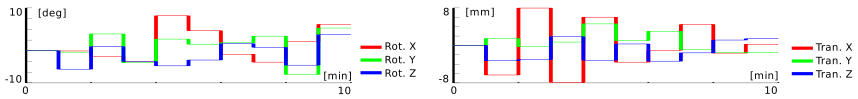

Fig. 2. Simulated random rigid motion parameters. Rotation in degrees (left) and translation im mm (right) along the three axes.

which is preferred over other gradient type optimisation algorithms as it guarantees positivity. The One Step Late approximation is adopted as the M step isn't otherwise solvable in closed form for any $p(\lambda|\gamma, y, \pi, \theta)$, as discussed in [9]:

$$\hat{\lambda}_b^{(n+1)} = \hat{\lambda}_b^{(n)} \frac{1}{\sum_{d,t} p_{bd}^{[t]} + \frac{\partial}{\partial \lambda_b} \log p(\lambda|\gamma, y, \pi, \theta)\Big|_{\lambda_b^{(n)}}} \sum_{d,t} \frac{p_{bd}^{[t]}\, q_d^{[t]}}{\sum_{b'=1}^{N_b} p_{b'd}^{[t]}\, \hat{\lambda}_{b'}^{(n)}} \tag{3}$$

$$p(\lambda|\gamma, y, \pi, \theta) = \prod_b \left[\sum_k p(k) p(\lambda_b|k, \theta) \prod_t p(y_b^{[t]}|k, \theta, \pi) \right]$$

Differentiating and expanding, the gradient in (3) simplifies to the following:

$$\frac{\partial \log p(\lambda|\gamma, y, \pi, \theta)}{\partial \lambda_b} = \sum_k z_{bk} \frac{\mu_{\lambda_k} - \lambda_b}{\sigma^2_{\lambda_k}}$$

(iii) $p(\theta_y^{[t]}|y^{[t]}, \gamma^{[t]}, \pi)$ Given activity and the transformation parameters, the parameters of the GM and the latent prior probability of the hidden state are updated by EM as it has better convergence properties than other gradient based algorithms: the expected value of the unobserved variables z_{bd} is updated using the provisional estimate of the parameters and the parameters are updated by:

$$\hat{\mu}_{\lambda_k}^{(n+1)} = \frac{1}{N_b} \frac{\sum_{b=1}^{N_b} z_{bk}^{(n+1)} \lambda_b}{\hat{\pi}_k^{(n+1)}} \qquad \hat{\sigma}^{2^{(n+1)}}_{\lambda_k} = \frac{1}{N_b} \frac{\sum_{b=1}^{N_b} z_{bk}^{(n+1)} (\hat{\mu}_{\lambda_k}^{(n+1)} - \lambda_b)^2}{\hat{\pi}_k^{(n+1)}}$$

$$\hat{\mu}_{y_k^{[t]}}^{(n+1)} = \frac{1}{N_b} \frac{\sum_{b=1}^{N_b} z_{bk}^{(n+1)} y_b^{[t]}}{\hat{\pi}_k^{(n+1)}} \qquad \hat{\sigma}^{2^{(n+1)}}_{y_k^{[t]}} = \frac{1}{N_b} \frac{\sum_{b=1}^{N_b} z_{bk}^{(n+1)} (\hat{\mu}_{y_k^{[t]}}^{(n+1)} - y_b^{[t]})^2}{\hat{\pi}_k^{(n+1)}}$$

$$\hat{\pi}_k = \frac{1}{N_b} \sum_b z_{bk}$$

3 Validation Study

Noisy simulation data modeling 3-D patient head movements were constructed by projecting at various orientations a brain phantom based on the BrainWeb database [11] and by warping the corresponding MRI image. Rigid motion was simulated by randomly generating rotation and translation along the three axes,

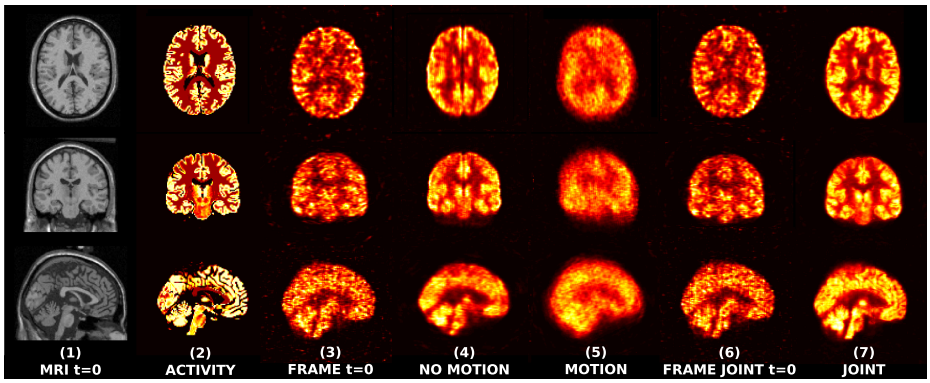

Fig. 3. From left to right: (1) MRI image frame at time $t = 0$; (2) activity phantom; (3) MLEM reconstruction of single frame ($t = 0$); (4) MLEM reconstruction of motion-free simulation; (5) MLEM reconstruction with no motion compensation; (6) reconstruction of single frame using joint generative model; (7) joint estimation of motion and activity. The method estimates correctly the rigid motion parameters, and reduces noise.

according to amplitudes and frequency that resemble typical measurements of head movement within a scanner for Emission Tomography [1] (figure 2). The MRI and functional imaging processes were decoupled by running independent simulations based on the ground truth normal brain tissue model from BrainWeb. The MRI image was generated with the BrainWeb simulator, which realistically accounts for noise of the imaging system. The parameters of the simulator were set for T1-weighted imaging with noise standard deviation set at 3% of the brightest tissue and perfect uniformity of the magnetic field (in accordance with the simplistic GM model). Activity of $^{18}F\text{-}FDG$ was simulated by associating typical activity levels to different tissue types, proportionally to partial voxel occupation. Specifically the activity in gray matter was set to a value 4 times higher than in all other tissues. The total number of counts was set to 32 Million. A $1.9M$ LOR PET imaging system was simulated by means of a rotation-based projector with realistic position dependent point source response [10] and ideal pre-reconstruction correction for randoms, scatter and attenuation. The MRI and activity images were defined on a cubic grid of $(128 \times 128 \times 128)$ voxels. The number of tissue types was assumed to be $N_k = 4$; for each t, $\mu_{y_k^{[t]}}$ were initialised to evenly spaced values in the range of intensity of the MRI image; $\sigma_{y_k^{[t]}}$ were initialised to $1/N_k$ of the image intensity range; μ_{λ_k} were initialised to evenly spaced values between 0 and the maximum activity assigned to the phantom; σ_{λ_k} to $1/N_k$ of the maximum activity assigned to the phantom; the mixing coefficients to $\pi_k = 1/N_k \; \forall \; k \in N_k$. Though ICM can be applied in any order, in this validation study the transformations and the parameters of the GM were updated for all time points at the same time and activity was estimated considering projections at all time frames, repeating 5 iterations of (i), (ii), (iii) for 20 iterations (early termination criterium). Results of the reconstruction are

	GM RC	GM SNR	WM RC	WM SNR
SINGLE FRAME	0.71	2.4	0.76	2.2
NO MOTION	0.71	4.0	0.76	3.0
MOTION	0.68	2.9	0.60	3.5
JOINT FRAME	0.72	2.7	0.77	2.3
JOINT	0.78	4.3	0.82	3.2

Fig. 4. Left: multi-modal anatomical/functional classification ($p(k_b|\lambda_b, y_b^{[t]})$). Right: Recovery Coefficient (RC) and Signal to Noise Ratio (SNR) in the gray matter and white matter regions. The table refers to the images in figure 3. The proposed method improves recovery and reduces noise by virtue of the additional information provided by the MRI image.

reported in figures 3 and 4. The activity estimate produced by applying the joint generative model presents higher coefficients of recovery both in gray matter and white matter when compared to standard MLEM reconstruction of the motion affected data. Even when compared to motion free MLEM reconstruction, the activity obtained by the generative model presents slightly higher recovery and signal to noise ratio, due to the MRI images improving the activity estimate.

3.1 Computational complexity

Computing the registration, reconstruction and segmentation jointly is costly. The proposed algorithm is essentially based on the recursive computation of projections and back-projections, exponentials (Gaussian likelihoods), spatial gradients, linear interpolations and 3-D transformations. Projections and back-projections constitute by far the most demanding part. Computing and storing the projection matrix $P = \{p_{bd}\}$ is not feasible. However the matrix is sparse and can be applied efficiently to an activity vector by a series of convolutions and summations if activity is defined on a regularly spaced grid [10]. If $P = \{p_{bd}\}$ is the probability that an event emitted in X_b in the reference space is detected in LOR d, then the expected photon counts at time t is $\bar{n}^{[t]} = \Delta t P \lambda^{[t]}(X_b)$, where Δt is the length of the time frame

$$\lambda^{[t]}(X_b) = \lambda^{[t]}(T^{-1}TX_b) = \lambda(T^{-1}X_b) \qquad \bar{n}^{[t]} = P\Delta t\lambda(T^{-1}X_b)$$

Projection and backprojection in (3) are then computed efficiently by convolution on a ragular grid as in [10] by warping the current activity estimate by each transformation, projecting, back-projecting and transforming back. This applies to any invertible transformation. A fast GPU based algorithm was used [10] along with an ordered subset scheme that updates activity considering only subsets of the lines of response in space and time, achieving convergence in 2 h.

4 Discussion

In this paper we presented a novel unified framework to reconstruct brain PET images taking into account potential motion during the acquisition and prior information coming from a second modality. With the growing field of PET/MRI

scanners, we believe that such paradigm can have a tremendous impact on the quality of the reconstructed PET images. The model has been evaluated with synthetic data, proving feasibility of the method; it is difficult, however, to envisage at this stage implementation strategies on real imaging systems such as defining the most appropriate MRI sequence and frequency of the frames. Given the general acceptance of probabilistic atlas based segmentation algorithm [7][12], we believe that a joint generative model for PET/MRI might prove useful when coupled with population based spatially varying priors of the hidden states and hyperpriors for the parameters of the mixture model in order to include experience in the reconstruction process. Since the presented model is valid for any invertible transformation, the method could be applied to other parts of the body by using a non-rigid transformation. Application to continuously moving structures however would require continuous MR acquisition, generating a tremendous amount of data and requiring a conspicuous amount of computational resources. Possible extensions of the model may include accounting for farmacokinetics and modeling lesions as outliers.

References

1. Fulton, R.R. Meikle, S.R. Eberl, S. Pfeiffer, J. Constable, R.T. and Fulham, M.J.: Correction for head movements in positron emission tomography using an optical motion-tracking system. In: IEEE Trans. Nucl. Sci., 49, 1, 116–123, February 2002.
2. Cherry, S.R.: Multimodality Imaging: Beyond PET/CT and SPECT/CT. Semin Nucl Med, 39(5), 348–353, September 2009.
3. Atre, A. Vunckx, K., Baete, K., Reilhac, A., Nuyts, J.: Evaluation of different MRI-based anatomical priors for PET brain imaging. In: IEEE Nucl. Sci. Sym. Conf. 1–7, Orlando, October 2009.
4. Leahy, R., Yan, X.: Incorporation of Anatomical MR Data for Improved Functional Imaging with PET. In: Inf. Proc. in Med. Imag. 105–120, Springer, 1991.
5. Somayajula, S., Rangarajan, A., Leahy, R.M.: PET image reconstruction using anatomical information through mutual information based priors: a scale space approach. In: Int. Sym. on Biomed. Imag. 165–168, April 2007.
6. Rahmim, A. Dinelle, K. Cheng, J.C. Shilov, M.A. Segars, W.P. Lidstone, S.C. Blinder, S. Rousset, O.G Vajihollahi, H. Tsui, B.M.W. Wong, D.F. Sossi, V.: Accurate Event-Driven Motion Compensation in High-Resolution PET Incorporating Scattered and Random Events. IEEE Trans. Med. Imag., 27, 8, August 2008.
7. Ashburner, J., Friston, K.: Unified segmentation. Neuroimage, 26(3), 839–51, 2005.
8. Besag, J.: On the Statistical Analysis of Dirty Pictures. J. of the Royal Stat. Soc. Series B (Methodological), 48(3), 259-302, 1986.
9. Green, P.G., Bayesian Reconstructions From Emission Tomography Data Using a Modified EM Algorithm. IEEE Trans. on Med. Imag. 9(1), 84–93, 1990.
10. Pedemonte, S., Bousse, A., Erlandsson, K., Modat, M., Arridge, S., Hutton, B.F., Ourselin, S.: GPU accelerated rotation-based emission tomography reconstruction. IEEE Nucl. Sci. Sym. Conf. M13-287, pp. 2657–2661, 2010.
11. Brain Web, `http://mouldy.bic.mni.mcgill.ca/brainweb/`
12. Menze, B.H. Leemput K.V., Lashkari, D. Weber, M.A. Ayache, N. Golland, P.: A Generative Model for Brain Tumor Segmentation in Multi-Modal Images. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (Eds.) MICCAI 2010, Part II, LNCS 6362, pp. 151–159. Springer-Verlag Heidelberg 2010.