Model-based normalization for iterative 3D PET image reconstruction

B Bai[1], Q Li[1], C H Holdsworth[2], E Asma[1], Y C Tai[3], A Chatziioannou[2] and R M Leahy[1]

[1] Signal and Image Processing Institute, 3740 McClintock Avenue EEB400, University of Southern California, Los Angeles, CA 90089, USA
[2] Crump Institute for Molecular Imaging, University of California, Los Angeles, CA 90095, USA
[3] School of Medicine, Washington University, St Louis, MO 63110, USA

E-mail: leahy@sipi.usc.edu

Abstract

We describe a method for normalization in 3D PET for use with maximum *a posteriori* (MAP) or other iterative model-based image reconstruction methods. This approach is an extension of previous factored normalization methods in which we include separate factors for detector sensitivity, geometric response, block effects and deadtime. Since our MAP reconstruction approach already models some of the geometric factors in the forward projection, the normalization factors must be modified to account only for effects not already included in the model. We describe a maximum likelihood approach to joint estimation of the count-rate independent normalization factors, which we apply to data from a uniform cylindrical source. We then compute block-wise and block-profile deadtime correction factors using singles and coincidence data, respectively, from a multiframe cylindrical source. We have applied this method for reconstruction of data from the Concorde microPET P4 scanner. Quantitative evaluation of this method using well-counter measurements of activity in a multicompartment phantom compares favourably with normalization based directly on cylindrical source measurements.

1. Introduction

Accurate normalization is essential for accurate quantitative three-dimensional (3D) PET. Inaccuracies in normalization factors can result in artefacts, poor uniformity and increased noise in the reconstructed images. Traditional solutions to the normalization problem include direct and component-based methods. In direct methods a known source of activity is scanned, then the normalization factors are estimated as the ratio between the ideal number of coincidences and those actually measured (Defrise *et al* 1991). The main problem with this

method is that it requires that a very large number of counts be detected to achieve acceptable statistical accuracy for each line of response (LOR). To maximize the number of counts over all LORs, direct approaches typically use a uniform cylindrical source. Unfortunately, this introduces its own problems since the observation model is complicated by a substantial scatter fraction.

Hoffman *et al* (1989) proposed a component-based method that divides the normalization factors into detector efficiency and spatial distortion correction, where the latter accounts for the radial mispositioning due to the geometry of the scanner. This model reduces the number of counts required by reducing the degrees of freedom in the normalization model so that the normalization factors are computed by averaging over multiple LORs. Casey *et al* (1995) and Badawi and Marsden (1999) extended this concept to develop sophisticated models accounting for a wide variety of factors affecting detection efficiency. Casey's normalization model includes intrinsic detector efficiency, a geometric factor, crystal interference and a deadtime factor. Badawi and Masden (1999) added time-alignment factors and a count-rate dependent block-profile to this model.

These models are complex and involve the sequential estimation of multiple types of normalization factors, often from different datasets. This can lead to inconsistent estimates since the normalization models are multiplicative. While optimal estimation of individual components, e.g. the detector efficiencies (Defrise *et al* 1991, Badawi 1998) have previously been investigated, joint estimation of all factors in the component-based models has not, to the best of our knowledge, previously been described. Here we present a unified model in which all factors are estimated simultaneously within a maximum likelihood framework. This model is specifically matched to the model-based maximum a *posteriori* (MAP) reconstruction method described by Qi *et al* (1998). The combination of our previously described system model with matched normalization allows us to explicitly account for the imperfections in the line-integral model using an accurate physical and statistical model for coincidence detection. In this way we build on our previous model, which included effects of detector solid angle, photon pair non-colinearity and intercrystal scatter and penetration, to also include effects arising from the block design, individual detector efficiencies, geometric effects and deadtime.

We describe our model-based approach below and present the results of experiments in which we applied this approach to data from the microPET P4 small animal scanner (Concorde Microsystems, Knoxville, TN).

2. Methods

2.1. Normalization within a statistical image reconstruction framework

We have developed a MAP estimation algorithm to reconstruct 3D PET images (Qi *et al* 1998). In this approach, the data are modelled as

$$\bar{\mathbf{y}} = \mathbf{P}\mathbf{x} + \bar{\mathbf{r}} + \bar{\mathbf{s}} \tag{1}$$

where $\bar{\mathbf{y}}$ is the mean of the data, $\mathbf{x}$ is the source distribution, $\bar{\mathbf{r}}$ is the mean of the randoms and $\bar{\mathbf{s}}$ is the mean of the scattered events. $\mathbf{P}$ is the system matrix describing the probability that an unscattered event is detected, which we factor as

$$\mathbf{P} = \mathbf{P}_{norm}\mathbf{P}_{blur}\mathbf{P}_{attn}\mathbf{P}_{geom} \tag{2}$$

where $\mathbf{P}_{geom}$ is the geometric projection matrix describing the probability that a photon pair reaches the front faces of a detector pair in the absence of attenuation and assuming perfect photon pair colinearity. $\mathbf{P}_{blur}$ models photon pair non-colinearity, intercrystal scatter and
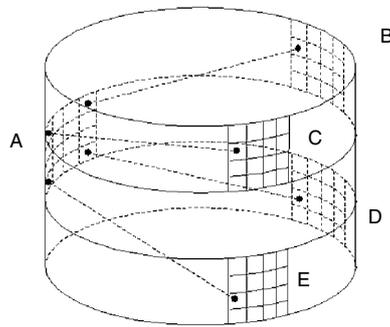
Figure 1. Illustration of the symmetries used in computing the geometric response.

crystal penetration, $\mathbf{P}_{\text{attn}}$ contains attenuation correction factors for each detector pair, and $\mathbf{P}_{\text{norm}}$ is a diagonal matrix containing the normalization factors for unscattered events.

The effects of solid angle variation at the detectors relative to the position of each voxel along a line of response, and of variations in the angle between the detector surface and the LOR, are accounted for in $\mathbf{P}_{\text{geom}}$. Similarly, the effects of crystal penetration that result in mispositioning of events towards the edge of the field of view are included in $\mathbf{P}_{\text{blur}}$. Consequently, these need not be included in the normalization factors as they are in previous factored methods (Casey *et al* 1995). Nevertheless, there are geometric factors that are not accounted for in (2), for example the gaps between detectors and the effects of the block design. These and other factors affecting sensitivity need to be included in the normalization factors.

Our approach to computing the normalization factors is to use (1) to map a known source distribution through the system matrix $\mathbf{P}$ and then choose the normalization factors to account for the differences between this mapping and the observed data. The source we use is a uniform cylinder. To account for statistical variability in the data, the normalization factors are computed in a maximum likelihood framework.

The normalization factors for unscattered events form the diagonal elements of the matrix $\mathbf{P}_{\text{norm}}$ and are modelled as

$$\mathbf{P}_{\text{norm}}(i, i) = \varepsilon_{d1^i} \varepsilon_{d2^i} g(l, j, k)^i \tau_{b1^i, b2^i} d_{d1^i} d_{d2^i} \tag{3}$$

where we have used i to index the LORs. The specific detectors forming this LOR are denoted $d1^i$ and $d2^i$; $b1^i$ and $b2^i$ are the blocks containing these detectors. The components of this model are as follows.

Geometric factors $g(l, j, k)^i$. The index $(l, j, k)^i$ represents the radial position l, the view angle j, and the sinogram index k, associated with LOR i. The sensitivity of each LOR is a function of the position of the two detectors in the block, the oblique angle of the LOR and the distance of the LOR from the centre of the field of view of the scanner. Our geometric factor combines the geometric and block-interference patterns of (Casey *et al* 1995) into a single factor. Since the scanner is highly symmetric both axially and transaxially, many LORs are equivalent in these respects. The microPET P4 has an 8×8 block design with 168 blocks arranged as 32 rings with 336 detectors in each ring. Figure 1 illustrates a four-fold symmetry of the LORs with respect to a single block A. There is another factor of two symmetry which arises from LORs originating from block A but in the detector positions shown in blocks B, C, D and E. These patterns are repeated for each block. Thus there are a maximum of $0.5 \times 8 \times 168$ equivalent LORs, where the 0.5 factor results from including each LOR twice when considering all detector blocks. The finite axial extent of the scanner results in missing

symmetries in oblique LORs. In addition, some of the symmetries are redundant within a single ring of blocks. As a result there are an average of 2.3×168 geometric symmetries for each LOR in the P4 scanner.

Detector efficiency ε_{d1^i}, ε_{d2^i}. These quantities describe the intrinsic efficiencies of the two detectors forming the LOR. The total number of such factors is equal to the number of detectors.

Time-alignment factor $\tau_{b1^i,b2^i}$. The time alignment factor is based on the model proposed by Badawi *et al* (1999) to account for differences in timing synchronization between different blocks. As the timing windows become misaligned between any pair of blocks, the detection efficiency drops. We assume an independent time-alignment factor for each pair of blocks.

Deadtime factor $d_{d1^i} d_{d2^i}$. The deadtime factors are estimated separately as described in section 2.4.

2.2. Scattered and random events

Since we estimate the normalization factors from data acquired using a cylindrical source, it is important to accurately model the effects of random and scattered events in the calibration data using (1).

For a known source distribution, we compute the expected scatter distribution using a fast Monte Carlo code that includes a multiple scatter model (Holdsworth *et al* 2002). The observed scatter differs from the computed scatter by a normalization factor. We assume that the normalization factors for scattered events differ from those for unscattered events, since scattered events detected by each detector pair come from a large volume in the object (Ollinger 1995), which makes the geometric factors unnecessary. Consequently, we model the mean of the detected scattered events as

$$\overline{s}_i = \varepsilon_{d1^i} \varepsilon_{d2^i} \tau_{b1^i,b2^i} d_{d1^i} d_{d2^i} \tilde{s}_i \qquad (4)$$

where $\tilde{s}_i$ are the scattered events computed using the Monte Carlo scatter model.

We acquire the delayed events as a separate randoms sinogram when collecting the cylindrical source calibration data, or equivalently when histograming the listmode stream collected from the scanner. We then compute the mean of the randoms process by a maximum likelihood technique, as described by (Mumcuoglu *et al* 1994), based on estimation of the singles rates at each detector in the randoms sinogram. The estimated mean randoms sinogram can then be substituted in (1) as a known quantity, independent of normalization factors other than deadtime.

2.3. Normalization factor estimation

We compute the normalization factors described in the previous section using a joint maximum likelihood optimization procedure. This differs from the common practice of using rotating rod sources to compute geometric factors and cylinder data to compute geometric efficiencies (Casey *et al* 1995, Badawi and Marsden 1999) and provides self-consistent estimates of the unknown parameters. Moreover, by basing the estimation on the model (1), the normalization is matched to the specific forward projection model that we subsequently apply during reconstruction. We note also that this same approach can be used as part of an analytic reconstruction method where we simply substitute the adjoint of the backprojection operator in place of **P** in (1).

Since the sinogram formed from the prompt-only events does include randoms, we can accurately model the measurements as Poisson using (1) to give the log likelihood

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{N} y_i \log \overline{y}_i - \overline{y}_i \tag{5}$$

where $\boldsymbol{\theta}$ represents the set of all normalization factors. The source distribution $\mathbf{x}$ in (1) is the known cylindrical source. We estimate the parameters by maximizing $L(\boldsymbol{\theta})$ using a grouped coordinate ascent method, updating each of the three groups of factors in turn using steepest ascent with a Newton–Raphson line search (Fessler *et al* 1997). We find in practice that effective convergence is reached in five iterations with three sub-iterations of line search at each main iteration.

2.4. Deadtime estimation

We assume for the purposes of computing the count-rate independent normalization factors, that the cylindrical source is of sufficiently low activity that deadtime effects are minimal. This assumption is reasonable for the LSO detectors in the microPET P4 scanner for which deadtime factors are considerably lower than they would be for a BGO system.

Deadtime is affected by the properties of the PMTs and detection electronics (Badawi and Marsden 1999). Rather than adopt the exponential model that was developed by Casey *et al* (1995), we instead use an empirical quadratic correction method (Tai *et al* 1998) which relates observed and true singles rates at each block by

$$\lambda_t = \frac{\lambda_a}{1 + \alpha \lambda_a + \beta \lambda_a^2} \tag{6}$$

where λ_t is the true singles rate, λ_a is the detected singles rate, and α and β are experimentally determined parameters. We allow a separate deadtime calibration of this type for each detector block in the system, estimated from the measured singles rate for that detector block.

Block detectors also exhibit an additional deadtime effect, characterized by gradual mispositioning of events towards the middle of the block as the count rate increases (Germano and Hoffman 1990, Badawi and Marsden 1999). This mispositioning contributes to a count-rate dependent variation of sensitivity across the detector blocks, which we report on below. Our results indicate that these variations are significant so that we include these factors in our deadtime correction.

Thus our overall deadtime model for each detector d_i, similar to that in Badawi and Marsden (1999), is the product of the mispositioning deadtime $d_{\mathrm{mp}}(d_i)$ and PMT and electronics deadtime $d_{\mathrm{pp}}(b_i)$:

$$d_{d_i} = d_{\mathrm{mp}}(d_i) \times d_{\mathrm{pp}}(b_i) \tag{7}$$

where b_i is the block containing detector d_i. The deadtime correction factor for each LOR is then the product of the factors for the two detectors forming the LOR.

We estimate the factors $d_{\mathrm{pp}}(b_i)$ by observing the singles rate at each detector block over a series of I frames, taken as an F-18 source decays over the expected range of activities for the scanner. Since the LSO detectors have a natural background activity, we model the true activity at the block as

$$\lambda(t) = A \, e^{-\phi t} + C \tag{8}$$
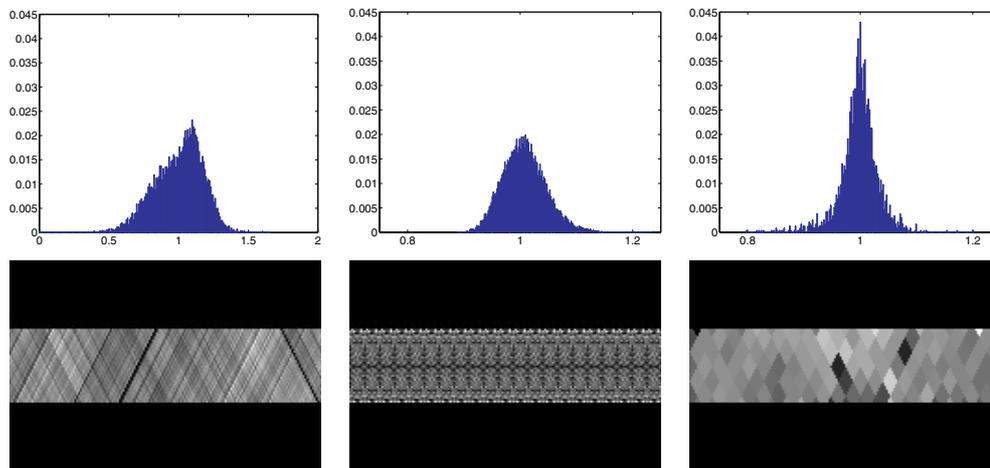
Figure 2. Illustration of the components of the count-rate independent normalization factors. Top row: normalized histograms showing distribution of the values of each factor; bottom row: normalization factors for a single sinogram. Left: detector efficiency; middle: geometric factors; right: time-alignment factors.
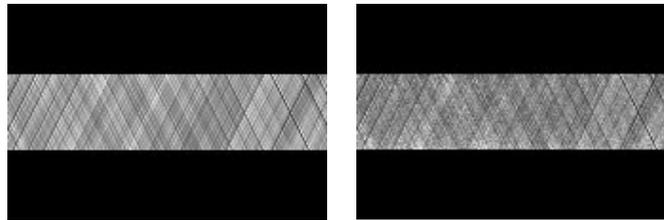
Figure 3. Normalization factors. Left: combined intrinsic, geometric and time-alignment factors jointly estimated from 70M count dataset; right: factors computed directly from a 600M count dataset.

where A is the initial singles rate, ϕ is the decay constant, and C is the background activity. Integrating this activity over the duration of each frame from time t_i to $t_i + T$ and applying the deadtime model (6) we obtain the series of equations

$$\frac{A}{-\phi T}[(e^{-\phi(t_i+T)} - e^{-\phi t_i})] + C = \frac{\lambda_{a,i}}{1 + \alpha\lambda_{a,i} + \beta\lambda_{a,i}^2} \tag{9}$$

for $i = 1, \ldots I$ where $\lambda_{a,i}$ is the observed total singles rate at the ith frame and α, β, C and A are the constants to be estimated for each block. This set of equations is solved using nonlinear least squares to obtain a separate pair of parameters for each detector block. The mispositioning deadtime parameters were computed as described by Badawi (1998).

3. Experimental studies

3.1. Cylinder calibration experiments

Using the method described above, we estimated normalization factors for the microPET P4 scanner. A 262 cc cylinder, diameter 5.5 cm and length 11.4 cm, containing 2.5 mCi F-18
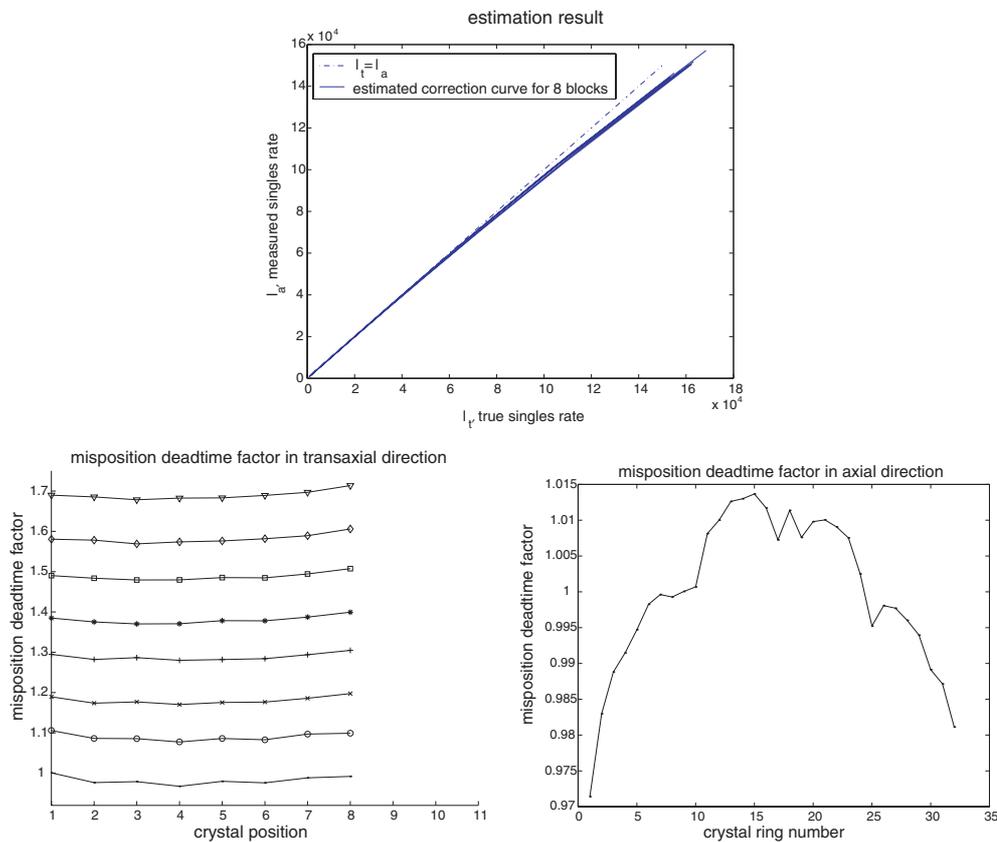
Figure 4. Deadtime correction factors. Top: results from fitting the quadratic deadtime model to singles data from a 2.5 mCi cylinder; shown are the measured versus quadratic fit to observed singles rates for eight different block detectors. Here a true singles rate of 16×10^4 corresponds to about 2.5 mCi activity; bottom left: transaxial block profile factors computed from coincidence data for a measured singles rate of 20 K for eight different detector rings. The eight profiles are displayed separately by adding multiples of 0.1 to each profile so that they do not overlap. Bottom right: axial block profile from same data as left.

solution was scanned for 10 h. Data were initially collected in listmode format and then re-sorted into prompt and delayed event sinograms for each frame. Singles rates for each block were also obtained for each frame. The data from the last two hours of the study were used to form the prompt and delayed sinograms used in estimating the count-rate independent normalization factors.

For comparison, we generated normalization factors directly from a frame of 2 h duration with start time 2 half-lives after the start of acquisition. This frame contained approximately 600M counts compared to 70M in the frame used to estimate the model-based normalization factors. The higher counts are necessary for the direct normalization method since the factor for each LOR is proportional to the total detected counts at that LOR and hence more sensitive to photon counting noise than the model-based factor computed using our method.

Ten additional frames were also generated from the listmode data, with each frame of 20 min duration and start point 0.5 half-lives after the start of the previous frame. The average
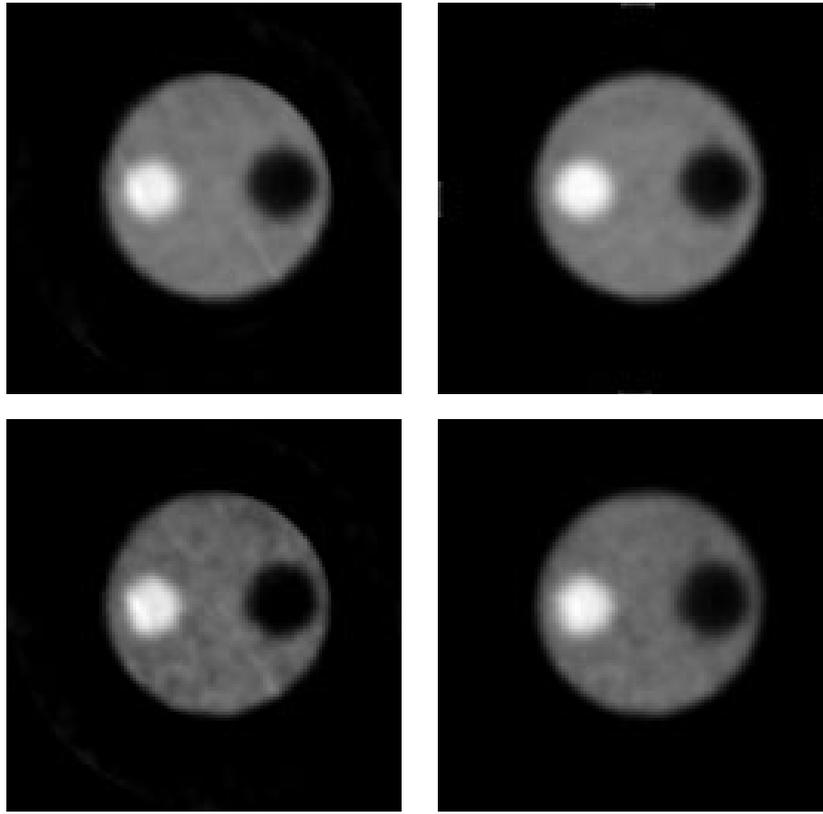
Figure 5. Comparison of reconstructed cylindrical phantom data. Top left: direct normalization, high count (139M); top right: model-based normalization, high count; bottom left: direct normalization, low count (42M); bottom right: model-based normalization, low count.
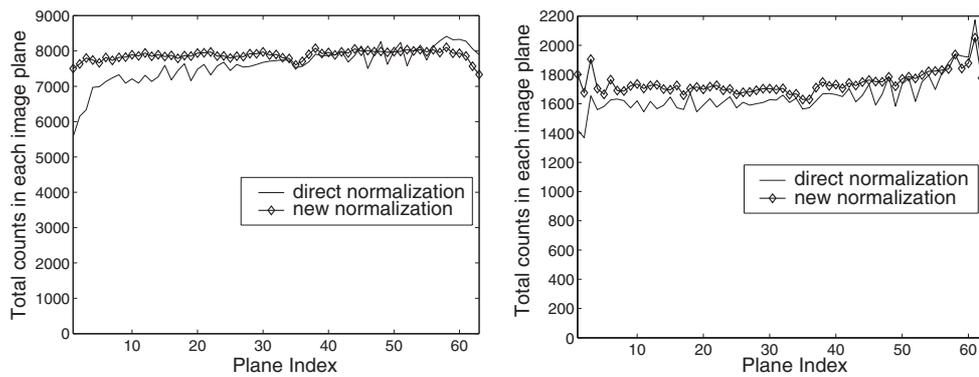
Figure 6. Axial profiles formed by integrating over transaxial images in each plane. Left: high count frame; right: low count frame.

singles rate for the first frame for each block was 1.43×10^5, the average singles rate for the last frame was 827, which was largely due to background radiation from the LSO detectors. These data were used to compute the deadtime factors.

Table 1. Means and standard deviations of absolute ROI values from reconstruction of the high count frame.

Method	Hot (μCi cc^{-1})	Cold (μCi cc^{-1})	Back 1 (μCi cc^{-1})	Back 2 (μCi cc^{-1})
Direct norm	3.41 ± 0.14	0.22 ± 0.04	1.77 ± 0.10	1.79 ± 0.11
Model-based norm	3.70 ± 0.06	0.20 ± 0.03	1.84 ± 0.04	1.87 ± 0.04
Well counter	3.87	0	1.83	1.83

3.2. Cylindrical phantom experiment

Using the model-based and direct normalization factors, we reconstructed a 4.5 cm diameter cylindrical phantom that contained hot and cold cylindrical inserts using the MAP approach described in Qi *et al* (1998) with a system matrix adapted to the size and LSO detector properties of the microPET P4 scanner. Scatter correction was based on the fast Monte Carlo method (Holdsworth 2002) applied to a preliminary reconstruction. The phantom listmode data was re-sorted into frames of 30 min duration each. One early frame (total counts: 139M) and one late frame (total counts: 42M) were reconstructed. Deadtime correction was applied to the earlier frame but not the latter frame in which deadtime factors were insignificant.

4. Results and discussion

In figure 2 we show the individual factors in our normalization model obtained from low activity cylinder scan data using joint maximum likelihood estimation. Figure 3 shows the combined normalization factors, formed as the product of these individual factors, in comparison to the higher count direct method. While the model-based approach uses far fewer counts, the resulting normalization files are less noisy because of the fewer effective number of parameters that must be estimated.

Figure 4 shows the block-wise and block-profile components of the deadtime correction factors. Figure 4(a) shows the excellent count-rate linearity for the LSO block detectors in the scanner over a wide activity range. The block-profile results indicate that these factors introduce significant count-rate dependent variations in sensitivity as a function of position in the block which should be included as part of the normalization process.

In figures 5 and 6 we show the effects of applying the model-based and direct normalization factors to the high- and low-count frame data from the phantom. In both cases, the images were reconstructed using 30 MAP iterations with all algorithm parameters otherwise equal. Figure 5 shows a single transaxial plane from the reconstructed phantom image. These results indicate improvement in transaxial uniformity and reduction in noise. Figure 6 shows axial uniformity in terms of the total activity of each of the 63 reconstructed transaxial sections. The model-based normalization produces improved axial uniformity compared to the direct normalization method. The axial uniformity of the direct normalization method could be improved by scatter correcting the cylindrical calibration scan. We will examine the impact of this in future work. We note however that while scatter in the normalization file may produce slow variations in axial uniformity, it is unlikely to produce the slice-to-slice variability exhibited in figure 6.

To quantify the uniformity of the image, four ROIs were chosen in each plane of the reconstructed image: one at the centre of the hot spot, another at the centre of the cold spot and two in the background. The mean and standard deviation of the activity in each ROI were computed, as were the aggregate mean and standard deviations for the four ROIs across all transaxial sections. We compared the values to decay-corrected well-counter measurements

Table 2. Means and standard deviations of absolute ROI values from reconstruction of the low count frame.

Method	Hot (μCi cc^{-1})	Cold (μCi cc^{-1})	Back 1 (μCi cc^{-1})	Back 2 (μCi cc^{-1})
Direct norm	0.87 ± 0.034	0.01 ± 0.006	0.40 ± 0.033	0.41 ± 0.034
Model-based norm	0.97 ± 0.025	0.04 ± 0.007	0.43 ± 0.018	0.44 ± 0.018
Well counter	1.03	0	0.49	0.49
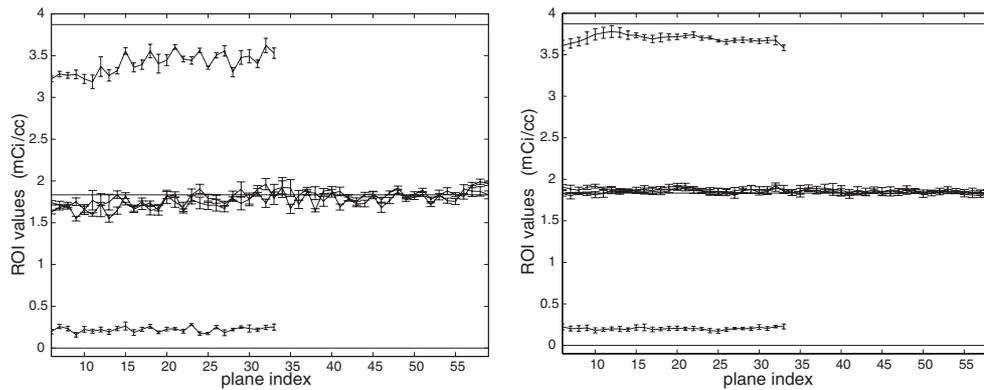
Figure 7. Plot of means and standard deviations of each ROI as a function of axial position, high count frame. Left: direct normalization; right: model-based normalization. Also shown as solid lines are the well counter values for each ROI. Note: the hot and cold inserts occupy approximately half the length of the scanner (planes 1–34).
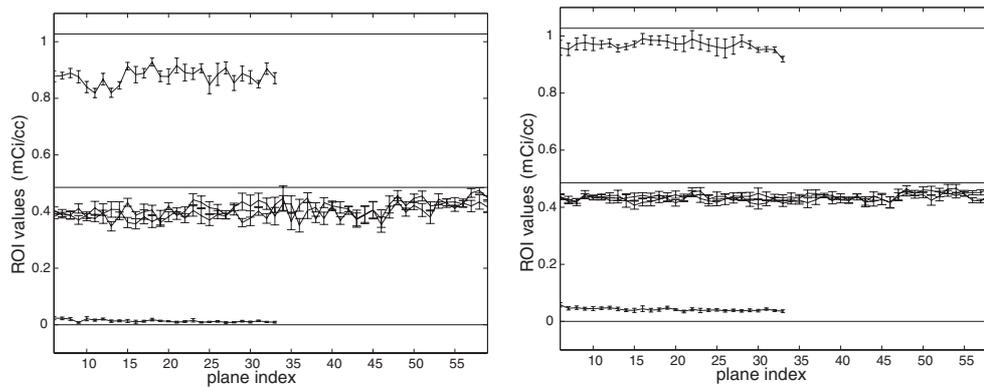
Figure 8. Plot of means and standard deviations of each ROI as a function of axial position, low count frame. Left: direct normalization; right: model-based normalization. Also shown as solid lines are the well counter values for each ROI.

for each region. Tables 1 and 2 show the aggregate mean and standard deviation of absolute ROI values for the high and low count frames, respectively. Figure 7 shows the ROI values of the high count frame as a function of axial position. Figure 8 shows the corresponding values for the low-count frame. Note that the inserts extend approximately half of the axial extent of the scanner so that hot and cold ROIs are shown only for the first 34 planes.

These results indicate both improved axial uniformity and reduced standard deviation using the model-based approach. The warm background and hot regions also show better quantitation since they are closer to the well counter measured activity. The new normalization does however indicate increased bias in the cold region for the low-count frame. Again, we suspect that scatter has a significant impact here, negatively impacting on the hot and background regions for the direct method (in which the normalization factors are not scatter corrected) while resulting in superior quantitation in the cold region. These results highlight the importance of accurate scatter handling for quantitation, even in small animal scanners where scatter is a lesser problem than in human studies.

5. Conclusion

We have described a component-based approach to normalization which can be specifically tailored to a model-based reconstruction method. By jointly estimating all components we obtain self-consistent estimates of these parameters. Calibration and phantom experiments indicate significant improvements in axial uniformity, noise due to normalization and quantitative accuracy when compared to normalization based directly on a high-count uniform cylinder. The method also has the advantage over the direct approach in that a shorter duration scan can be used since the factored model has far fewer parameters than the total number of LORs in the data.

One potential disadvantage of our approach is that the true source distribution and its position in the scanner must be known accurately since this distribution is forward projected and compared to the measured data in order to compute the normalization factors. The method is also dependent on the accuracy of the scatter model, but this is also the case for the direct method in which scatter will affect the measured sensitivity of each LOR. With our model-based approach we could use alternative source distributions such as rotating planar or line sources in which scatter fractions are far lower. Provided the parameters remain identifiable, which does not necessarily require that all LORs are illuminated, the same approach can be used. These alternative source distributions, however, present other practical problems including the difficulties of producing a planar source with sufficient uniformity and of providing an appropriate trajectory for a line source to ensure that all parameters are identifiable.

Acknowledgments

References

Badawi R D 1998 Aspects of optimisation and quantification in three-dimensional positron emission tomography *PhD Thesis* University of London
Badawi R D and Marsden P K 1999 Developments in component-based normalization for 3D PET *Phys. Med. Biol.* **44** 571–94
Casey M E, Gadagkar H and Newport D 1995 A component based method for normalization in volume PET *Proc. 3rd Int. Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* pp 67–71
Defrise M, Townsend D, Bailey D, Geissbuhler A, Michel C and Jones T 1991 A normalization technique for 3D PET data *Phys. Med. Biol.* **36** 939–52

Fessler J A, Ficaro E P, Clinthorne N H and Lange K 1997 Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction *IEEE Trans. Med. Imaging* **16** 166–75

Germano G and Hoffman E J 1990 A study of data loss and mispositioning due to pileup in 2-D detectors in PET *IEEE Trans. Nucl. Sci.* **37** 671–5

Hoffman E J, Guerrero T, Germano G, Digby W M and Dahlbom M 1989 PET system calibrations and corrections for quantitative and spatially accurate images *IEEE Trans. Nucl. Sci.* **33** 1108–12

Holdsworth C H, Levin C S, Janecek M, Dahlbom M and Hoffman E J 2002 Performance analysis of an improved 3D PET Monte Carlo simulation and scatter correction *IEEE Trans. Nucl. Sci.* **48** 83–9

Mumcuoglu E U, Leahy R M, Cherry S R and Zhou Z 1994 Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images *IEEE Trans. Med. Imaging* **13** 687–701

Ollinger J M 1995 Detector efficiency and Compton scatter in fully 3D PET *IEEE Trans. Nucl. Sci.* **42** 1168–73

Qi J, Leahy R M, Cherry S R, Chatziioannou A and Farquhar T H 1998 High resolution 3D Bayesian image reconstruction using the microPET small-animal scanner *Phys. Med. Biol.* **43** 1001–13

Tai Y C, Chatziioannou A, Dahlbom M and Hoffman E J 1998 Investigation of deadtime characteristics for simultaneous emission-transmission data acquisition in PET *IEEE Trans. Nucl. Sci.* **45** 2200–4