The effect of regularization in motion compensated PET image reconstruction: a realistic numerical 4D simulation study

C Tsoumpas, I Polycarpou, K Thielemans, C Buerger, A P King, T Schaeffter and P K Marsden

Department of Biomedical Engineering, Division of Imaging Sciences and Biomedical Engineering, King’s College London, King’s Health Partners, St. Thomas’ Hospital, London, SE1 7EH, UK

E-mail: Charalampos.Tsoumpas@kcl.ac.uk

Received 16 August 2012, in final form 30 January 2013
Published 26 February 2013
Online at stacks.iop.org/PMB/58/1759

Abstract

Following continuous improvement in PET spatial resolution, respiratory motion correction has become an important task. Two of the most common approaches that utilize all detected PET events to motion-correct PET data are the reconstruct-transform-average method (RTA) and motion-compensated image reconstruction (MCIR). In RTA, separate images are reconstructed for each respiratory frame, subsequently transformed to one reference frame and finally averaged to produce a motion-corrected image. In MCIR, the projection data from all frames are reconstructed by including motion information in the system matrix so that a motion-corrected image is reconstructed directly. Previous theoretical analyses have explained why MCIR is expected to outperform RTA. It has been suggested that MCIR creates less noise than RTA because the images for each separate respiratory frame will be severely affected by noise. However, recent investigations have shown that in the unregularized case RTA images can have fewer noise artefacts, while MCIR images are more quantitatively accurate but have the common salt-and-pepper noise. In this paper, we perform a realistic numerical 4D simulation study to compare the advantages gained by including regularization within reconstruction for RTA and MCIR, in particular using the median-root-prior incorporated in the ordered subsets maximum a posteriori one-step-late algorithm. In this investigation we have demonstrated that MCIR with proper regularization parameters reconstructs lesions with less bias and root mean square error and similar CNR and standard deviation to regularized RTA. This finding is reproducible for a variety of noise levels (25, 50, 100 million counts), lesion sizes (8 mm, 14 mm diameter) and iterations. Nevertheless, regularized RTA
can also be a practical solution for motion compensation as a proper level of regularization reduces both bias and mean square error.

Online supplementary data available from stacks.iop.org/PMB/58/1759/mmedia
(Some figures may appear in colour only in the online journal)

1. Introduction

Following continuous improvements in PET spatial resolution, respiratory motion correction has become an important task (Schäfers 2008). Amongst several approaches to motion-correct PET data (Nehmeh 2013), two are the most common: reconstruct-transform-average (RTA) (Picard and Thompson 1997) and motion-compensated image reconstruction (MCIR) (Qiao et al 2006, Qi and Huesman 2002). In RTA, separate images are reconstructed for each respiratory frame, which are then transformed to one respiratory reference frame and averaged to produce a motion-corrected image. In MCIR, the projection data from all respiratory frames are reconstructed together by including the motion information into the reconstruction system matrix, so that a motion-corrected image is produced directly (Rahmim et al 2013). In both cases it is assumed that an accurate description of the motion is available, for example using non-rigid image registrations derived from independently acquired CT, MR images or PET data itself.

As shown by Thielemans et al (2006) RTA images can have lower regional standard deviation which may result in higher lesion detectability. However, this has not been communicated in recent literature and it has been assumed that RTA has inferior image quality than MCIR (Lamare et al 2007). Similar conclusions were also drawn by a theoretical comparison of RTA versus MCIR using penalized likelihood reconstruction (Asma et al 2006), where it was shown that penalized MCIR would provide lower variance than penalized RTA—at least when there are very few counts per respiratory frame. Another recent theoretical investigation of motion-compensated penalized likelihood reconstruction confirmed that the variances of MCIR methods are lower than or comparable to the variance of RTA due to the statistical weighting (Chun and Fessler 2013). However, our recently published investigation demonstrates that in the unregularized case, MCIR, even though it is quantitatively superior to RTA, produces images with higher regional standard deviation (Polycarpou et al 2012). The different conclusions may be explained by the definition of variance: regional variance versus variance over independent realizations can provide different insights on the algorithm performance. In particular, regional variance corresponds better to visual noise perception than the variance of independent realizations (Tong et al 2010).

This study aims to answer how MCIR and RTA perform with regularization and which of the two approaches has better practical performance with respect to both noise and quantification of lesions in motion-corrected images. The entire analysis in this work is based on realistic 4D simulated PET data (Tsoumpas and Gaitanis 2013).

2. Theory

Unregularized statistical iterative reconstruction methods estimate an image that is as consistent as possible with the measured data. However, in PET this problem is ill-conditioned and it generates images with a large amount of noise. As an attempt to improve this ill-conditioned situation regularization approaches have been developed that can significantly help reduce noise in the reconstructed images (Fessler 1994).
This study makes use of a well-known iterative algorithm with a common regularization approach (1): the ordered subsets maximum a posteriori one-step-late, i.e. OSMAPOSL (Green 1990a, 1990b) together with the median-root-prior (MRP) (Alenius and Ruotsalainen 1997, 2002) which is freely available within the Software for Tomographic Image Reconstruction (STIR) library (Thielemans et al 2012). This algorithm has been thoroughly examined by Bettinardi et al (2002). In our investigation we extended STIR OSMAPOSL to include motion information in order to reconstruct all respiratory gates into one reference frame (MCIR).

RTA-MRP-OSL: equations (1–2) describe the separate reconstruction of each respiratory gate. Equation (3) describes the transformation step that is followed after reconstruction of all gates:

\[
\Lambda_{vg}^{(s+1)} = \Lambda_{vg}^{(s)} \frac{1}{\sum_{b \in S_l} P_{vb} A_{bg} + \beta \nabla_{\Lambda_{vg}} E_{v}} \sum_{b \in S_l} P_{vb} \sum_{\tilde{v}} P_{\tilde{v}g} A_{\tilde{v}g}^{(s)} + \frac{\beta}{\Lambda_{vg}} \nabla_{\Lambda_{vg}} E_{v} (s) Y_{vg} \sum_{b \in S_l} P_{vb} \sum_{\tilde{v}} P_{\tilde{v}g} \Lambda_{\tilde{v}g}^{(s)} + \frac{\beta}{\Lambda_{vg}} \nabla_{\Lambda_{vg}} E_{v} (s) \Lambda_{vg}^{(s)} + B_{bg} A_{bg} (1)
\]

where:

\[
\beta \nabla_{\Lambda_{vg}} E_{v} (s) \overset{df}{=} \beta \frac{\Lambda_{vg}^{(s)} - M_{vg}^{(s)}}{M_{vg}^{(s)}}
\]

\[
\Lambda_{v} = \frac{1}{G} \sum_{g} \sum_{v} \tilde{W}_{v'g\to v}^{-1} \Lambda_{v'g}. (3)
\]

While the corresponding MCIR-MRP-OSL equation (4) is:

\[
\Lambda_{vg}^{(s+1)} = \Lambda_{vg}^{(s)} \frac{1}{\sum_{b \in S_l, g} \sum_{v} \tilde{W}_{v'g\to v}^{-1} P_{vb} A_{bg} + \beta \nabla_{\Lambda_{vg}} E_{v}} \sum_{b \in S_l, g} \sum_{\tilde{v}} \tilde{W}_{v'g\to v}^{-1} P_{vb} \sum_{\tilde{v}} P_{\tilde{v}g} \Lambda_{\tilde{v}g}^{(s)} + \frac{\beta}{\Lambda_{vg}} \nabla_{\Lambda_{vg}} E_{v} (s) Y_{vg} \sum_{b \in S_l, g} \sum_{\tilde{v}} \tilde{W}_{v'g\to v}^{-1} P_{vb} \sum_{\tilde{v}} P_{\tilde{v}g} \Lambda_{\tilde{v}g}^{(s)} + \frac{\beta}{\Lambda_{vg}} \nabla_{\Lambda_{vg}} E_{v} (s) \Lambda_{vg}^{(s)} + B_{bg} A_{bg} \left(1 + \frac{1}{\Lambda_{vg}} \nabla_{\Lambda_{vg}} E_{v} (s) \right). (4)
\]

Notations:
- **$\Lambda_{vg}^{(s)}$** is the estimated radioactivity at voxel $v$ and subiteration $s$
- **$Y_{bg}$** is the number of measured coincident photons of each detector pair (bin) $b$ that belongs to the $b$th subset $S$ and gate $g$
- **$S_l$** corresponds to the $l$th subset of the projection space, which is divided into $L$ total subsets
- **$s$** is the sub-iteration number. A set of $L$ sub-iterations comprises a full iteration
- **$P_{vb}$** is the system projection matrix
- **$\tilde{W}$, $\tilde{W}^{-1}$** represent the forward /backward warping operations of the image that move the activity from one location (e.g. $v'$) to voxel $v$ using the motion fields and linear interpolation
- **$E$** is the ‘potential’ function
- **$M_{v}$** corresponds to the median $3 \times 3 \times 3$ mask width of neighbourhood voxels centred at voxel $v$ (Bettinardi et al 2002)
- **$G$** is the total number of gates
- **$\beta, \beta_g$** are the penalization factors for MCIR and RTA, respectively. Note that $\beta = G \times \beta_g$, but for simplicity all cases are displayed with respect to $\beta_g$
- **$A_{bg}$ and $B_{bg}$** are the attenuation coefficient and background (e.g. scatter) term for each bin and gate, respectively

1 The STIR library is currently available at: http://stir.sourceforge.net
2 In STIR there are two different approaches to applying the MRP: additive or multiplicative. Here, we chose the additive MRP as originally suggested by Alenius et al (1997).
Table 1. Characteristics of the lesions inserted in the computational PET-MR phantom. Note that the maximum left-right displacement is larger than the maximum posterior–anterior. This can be explained when the volunteer is not very well aligned with the direction of the bed.

| Lesion id | Diameter (mm) | Maximum inferior-superior displacement (mm) | Maximum posterior-anterior displacement (mm) | Maximum left-right displacement (mm) | SUV |
|-----------|---------------|--------------------------------------------|---------------------------------------------|-----------------------------------|-----|
| 1         | 14            | 16.9                                       | 2.7                                         | 1.9                               | 9.5 |
| 2         | 14            | 13.4                                       | 2.7                                         | 4.1                               | 7.5 |
| 3         | 14            | 3.5                                        | 2.4                                         | 4.2                               | 7.5 |
| 4         | 14            | 15.4                                       | 1.8                                         | 7.6                               | 6.5 |
| 5         | 14            | 13.8                                       | 2.9                                         | 2.3                               | 4.5 |
| 6         | 14            | 8.8                                        | 1.4                                         | 3.7                               | 4.5 |
| 7         | 8             | 11.9                                       | 3.5                                         | 8.4                               | 6.5 |
| 8         | 8             | 12.3                                       | 2.1                                         | 5.0                               | 4.5 |
| 9         | 8             | 3.7                                        | 1.2                                         | 5.8                               | 4.5 |

Both RTA-MRP-OSL and MCIR-MRP-OSL made use of the same routines for warping each respiratory gate to the reference gate. The motion correction operation was approximated by warping the image with the ‘inverted’ motion fields that were numerically estimated from the motion vectors (Crum et al. 2007) which were used to simulate the motion. The operator $\hat{W}^{-1}$ acts to warp all respiratory gates to the reference gate, while the operator $\hat{W}$ ‘unwarps’ the reference gate to each respiratory gate so that they can be compared with the corresponding projection gates in the numerator.

3. Experiments and methods

Generation of realistic 4D PET data

Realistic PET data for the chest were simulated using segmentations of real MR images (Tsoumpas et al. 2011). We acquired two different MR sequences from a volunteer to generate the 4D attenuation and emission maps (Buerger et al. 2012). The first acquisition used an ultra-short time-echo (UTE) sequence in order to separate structures such as liver, lung, cortical bone and soft tissues from the background using a semi-automatic segmentation approach with local thresholding. These were combined to generate the 3D emission and attenuation maps. The second acquisition was a dynamic MR sequence (35 dynamic frames, 0.7 s duration each) acquired during normal tidal breathing. This acquisition was used to estimate motion fields with a non-rigid registration algorithm (Buerger et al. 2011) which are essential to simulate realistic respiratory motion of the 3D emission and attenuation maps. Eight respiratory frames were selected based on displacement gating of a 1D signal navigator placed at the diaphragm of the dynamic MR.

Realistic attenuation values were assigned to the segmented regions leading to the 3D attenuation map $A_v$ with the following attenuation values: air 0 cm$^{-1}$, lung tissue 0.03 cm$^{-1}$, soft tissue 0.099 cm$^{-1}$, bone 0.15 cm$^{-1}$. Similarly, realistic emission values were assigned leading to the 3D emission map $\Lambda_v$ with the following emission values: air 0, lung tissue 0.5, soft tissue 1, bone 2.3, liver 2.5, and myocardium 3.2. Nine soft tissue tumour lesions with their characteristics and standardized uptake values (SUV) shown in table 1 were manually inserted at different positions of the 3D computational phantom as shown in figure 1. The attenuation and emission maps ($A_v$, $\Lambda_v$) were then combined with the motion fields to compute eight respiratory-gated maps ($A_{vg}$, $\Lambda_{vg}$). The reference gate was the one at the mid-respiration position. The resulting attenuation and emission maps were used to simulate a 4D PET.
acquisition with a fast analytic simulation technique (Tsoumpas et al 2011). PET projection data accounted for attenuation, scatter (approximately 33%) and resolution effects (Gaussian kernel of 5 mm), but not for randoms. Each of the eight gates was forward-projected simulating a fully 3D acquisition (maximum ring difference: 43, span: 3) with a range of different total unscattered photon-pair count measurements for the Philips Gemini PET scanner: $25 \times 10^6$, $50 \times 10^6$, and $100 \times 10^6$ Poisson-distributed counts which correspond to approximately 2.5, 5 and 10 min of a standard clinical PET acquisition. Counts are reported as the total number for the eight gates with each gate having one eighth of the total counts.

Iterative reconstruction

One of the main questions in iterative reconstruction methods is how many iteration steps should be performed to converge? Or at least, how far from convergence can we afford to stop? In particular, for OSEM a higher number of iterations is generally suggested for better quantification otherwise the value is biased towards the uniform initial image (Barrett et al 1994). The noise amplification with increase in iteration number (Wilson et al 1994) can be handled with some form of regularization, as for example in this paper with the MRP regularization. In particular, this type of regularization can also improve convergence rate (Bettinardi et al 2002). As this study focuses on the regularization term in relationship with motion compensated reconstruction, one specific question is how the regularization term affects convergence of MCIR and RTA respectively? Ideally, we should study the results at (numerical) convergence, where the required number of iterations would vary for the different cases. However, this may require a large number of iterations (Nuyts and Fessler 2003), which is impractical for such type of investigation, as each 4D equivalent iteration requires about 1.5 h (single process on computational platform with two Intel® Xeon® L5420 quad core processors running at 2.5 GHz). In order to minimize the computational cost, the main part of the investigation compares the figures of merit for a variety of settings fixed at ten iterations (23 subsets). Then, in order to study the impact of iteration number on RTA-MRP-OSL and MCIR-MRP-OSL images, a selected number of reconstructions were performed up to 40 iterations. This helps to characterize how the figures of merit are affected at larger number of iterations and if the main conclusions would have been affected.

Each reconstructed slice consisted of $250 \times 250$ pixels with size $2 \times 2$ mm each, and the entire volume consisted of 87 slices with 2 mm thickness. The kernel size for the regularization was $3 \times 3 \times 3$ voxels and a range of different weights ($0, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100$) were selected for the regularization term, also referred to as the penalization weighting factor ($\beta_n$). These weights were constant and spatially invariant throughout each reconstruction.

Figures of merit

First, we performed a bias versus standard deviation characterization for all lesions, penalty factors and three different noise levels. Then, for a selected noise level (50 million counts)
we plotted the root mean square error (RMSE) and contrast to noise ratio (CNR) against the penalty factor for each lesion. These figures of merit are estimated with the following equations:

\[
\begin{align*}
    mL_i &= \frac{\sum_{v \in L_i} \Lambda_v}{\sum_{v \in L_i} 1} \\
    sL_i &= \sqrt{\frac{\sum_{v \in L_i} (\Lambda_v - mL_i)^2}{\sum_{v \in L_i} 1}} \\
    bL_i &= mL_{i,\text{ideal}} - mL_i \\
    rL_i &= \sqrt{bL_i^2 + sL_i^2} \\
    cL_i &= \frac{mL_i - mB}{mB \cdot sB}
\end{align*}
\]

where \(mL\), \(sL\), \(bL\), \(rL\), \(cL\) correspond to values of the mean, standard deviation, bias, root mean square error and contrast to noise ratio of the voxels \(v\) of each lesion \(L_i\) and \(mB, sB\) correspond to the mean and standard deviation values of their background respectively, and \(mL_{i,\text{ideal}}\) is the mean value of the corresponding lesion in the noiseless PET image. In all cases the background regions were placed in the tissue that surrounds the lesion and selected to have the same size and shape with each corresponding lesion. Finally, based on these comparisons we select the best case of RMSE for MCIR and at matched regional standard deviation levels for RTA we provide results from twenty independent noise realizations in order to confirm that the regional comparisons are representative to make valid conclusions. The following equations (10)–(13) were used to calculate the mean \((M)\), standard deviation \((S)\), bias \((B)\) from the ideal noiseless PET image \((\Lambda_{\text{ideal}})\) and root mean square error images \((R)\):

\[
\begin{align*}
    M &= \sqrt[20]{\sum_{r=1}^{20} \Lambda_r} \\
    S &= \sqrt[20]{\sum_{r=1}^{20} (\Lambda_r - M)^2} \\
    B &= \Lambda_{\text{ideal}} - M \\
    R &= \sqrt{B^2 + S^2}
\end{align*}
\]

Finally, we calculate the average and standard error values for \(mL, sL, bL, rL, cL\) overall the twenty independent noise realizations.

4. Results

In order to show how \(\beta_g\) affects the images visually we present in figure 2 coronal planes through the three lesions located over the liver dome for \(50 \times 10^6\) counts (typical for a 5 min FDG scan) and for a selection of \(\beta_g\) (i.e. 0, 1, 2, 5, 10, 20, 50, 100) for MCIR-MRP-OSL and RTA-MRP-OSL. Figure 3 shows the visual improvement of images for the simulations of
Regularization in motion incorporated PET reconstruction

**Figure 2.** Coronal planes through lesions 2, 5, 8 for $50 \times 10^6$ counts and different penalization factors ($\beta_g$) after ten iterations of RTA-MRP-OSL and MCIR-MRP-OSL. All images are displayed in SUV scale from 0 to 5.

**Figure 3.** Coronal planes through lesions 2, 5, 8 for $25 \times 10^6$ and $100 \times 10^6$ counts reconstructed without penalization and with penalization ($\beta_g = 50$) after 10 iterations of RTA-MRP-OSL and MCIR-MRP-OSL. All images are displayed in SUV scale from 0 to 5.

$25 \times 10^6$ and $100 \times 10^6$ counts for $\beta_g = 50$ that produces high CNR. To study the effect of regularization in a systematic manner we firstly explore how the weight factor $\beta_g$ affects regional bias, $b_L$ and standard deviation $s_L$ values of the nine lesions. This is provided in figure 4 in the form of a graph of bias versus standard deviation as a function of $\beta_g$ for three noise levels.

The multiple plots in figure 5 (RMSE) and figure 6 (CNR) show the corresponding values for each lesion for the simulation of the $50 \times 10^6$ counts as reconstructed with RTA-MRP-OSL and MCIR-MRP-OSL at ten iterations with different $\beta_g$. For the same number of counts, the
Bias versus standard deviation for different noise levels and penalisation factors for RTA and MCIR for each lesion

**Figure 4.** Bias ($b$; y-axis) versus regional standard deviation ($s$; x-axis) for each lesion as computed from three different noise realizations (25 $\times$ $10^6$ counts: thin lines, 50 $\times$ $10^6$ counts: lines with medium thickness, 100 $\times$ $10^6$ counts: thick lines) reconstructed with RTA-MRP-OSL and MCIR-MRP-OSL at 10 iterations with different penalization. The eleven points correspond to $\beta_g = 0, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, and 100$. Rows correspond to lesion size and columns to lesion organ location.

**Table 2.** Indicative mean value of bias ($b$), standard deviation ($s$), RMSE ($r$) and CNR ($c$) of the SUV over all the nine lesions for one realization.

| Units in | RTA  | MCIR | RTA  | MCIR | RTA  | MCIR | RTA  | MCIR | RTA  | MCIR |
|----------|------|------|------|------|------|------|------|------|------|------|
| SUV scale | $\beta_g = 2$ | $\beta_g = 5$ | $\beta_g = 10$ | $\beta_g = 20$ | $\beta_g = 50$ |
| Bias ($b$) | 2.35 | **2.07** | 2.36 | 2.12 | 2.41 | 2.19 | 2.50 | 2.32 | 2.72 | 2.59 |
| Standard deviation ($s$) | 1.14 | 1.73 | 1.03 | 1.37 | 0.93 | 1.14 | 0.83 | 0.99 | **0.71** | 0.84 |
| RMSE ($r$) | 2.62 | 2.71 | 2.58 | 2.53 | 2.59 | **2.48** | 2.64 | 2.53 | 2.82 | 2.74 |
| CNR ($c$) | 13.7 | 11 | 15.4 | 13.2 | 16.8 | 15.7 | 18.3 | 18.3 | 19.2 | **20.77** |

average regional values over all nine lesions for the bias, standard deviation and RMSE of the SUVs are shown in table 2. Inspection of the table values indicates that the RMSE is minimum for MCIR ($\beta = 10 \times 8$). The regional standard deviation ($s$) of this MCIR setting matches the one of RTA-MRP-OSL at $\beta_g = 2$. For these two particular settings we repeated twenty independent noise realizations and display coronal planes of the $M$, $B$, $S$, and $R$ (figure 7). Figure 8 shows the average values for $m$, $s$, $r$ and $c$ over all realizations for each lesion.

Finally, we investigate the impact of iterations in RTA-MRP-OSL and MCIR-MRP-OSL and in particular what is the average bias difference ($\delta b$) and the average standard deviation difference ($\delta s$) between iterations 10 and 40, as defined in the equations (14) and (15). We plot these in figure 9 for three penalty factors ($\beta_g = 1, 10, 100$) and three count-levels
Regularization in motion incorporated PET reconstruction

Figure 5. RMSE values ($r$: y-axis) for each lesion as computed from one realization ($50 \times 10^6$ counts) reconstructed with RTA-MRP-OSL (thin line) and MCIR-MRP-OSL (thick line) at 10 iterations with different penalization weights ($\beta_g$: x-axis). Rows correspond to lesion size and columns to lesion organ location.

(25 $\times$ $10^6$, 50 $\times$ $10^6$, 100 $\times$ $10^6$). Furthermore, in the supplementary data (available at http://stacks.iop.org/PMB/58/1759/mmedia), we provide additional plots to demonstrate how the mean $m_i$ and standard deviation $s_i$ values of each lesion $l$ are changing over iterations (5, ..., +5, ..., 40) for the same three penalization factors and noise levels:

$$\delta b = \sum_{i=1}^{9} (b_{Li}^{10it} - b_{Li}^{0it})$$

$$\delta s = \sum_{i=1}^{9} (s_{Li}^{10it} - s_{Li}^{0it})$$

5. Discussion

Figures 2 and 3 provide a visual demonstration of the images reconstructed with MCIR-MRP-OSL and RTA-MRP-OSL with different $\beta_g$ and noise levels. It can be noticed that $\beta_g$ affects MCIR-MRP-OSL and RTA-MRP-OSL differently: A stronger penalty is needed in MCIR-MRP-OSL to obtain images of similar noise level comparing to RTA-MRP-OSL. At larger $\beta_g$ e.g. 50 or 100, MCIR-MRP-OSL and RTA-MRP-OSL images become similar. This is also confirmed quantitatively by the results in table 2 and figure 4. Figure 3 demonstrates substantial visual improvement with $\beta_g = 50$ for both low and high counts. The low count experiment corresponds to 2–3 min of PET acquisition, which provides a realistic simulation setup as in current whole body PET/CT scanning protocols.

In figure 4 the quantitative comparison shows that the regional relative bias for the nine lesions is almost unaltered with MCIR-MRP-OSL until $\beta$ is between $2 \times 8$ and $10 \times 8$. This does not appear to be dependent on the lesion location, size or number of simulated
Figure 6. CNR values (c; y-axis) for each lesion as computed from one realization (50 × 10^6 counts) reconstructed with RTA-MRP-OSL (thin line) and MCIR-MRP-OSL (thick line) at 10 iterations with different penalization weights (β; x-axis).

counts. On the other hand, RTA-MRP-OSL always demonstrates considerable improvement with respect to bias for small β (e.g. 0.1) with minimum change—even small increase in a few instances—in standard deviation for any type of lesion. The improvement is more apparent at lower number of counts. For 50 × 10^6 counts, the RTA bias in the nine lesions is 13.7% higher than MCIR, but following regularization it drops to 7.2%. This is an important finding as the RTA bias can be substantially improved with regularization possibly because it makes the convergence faster, as discussed in a previous section (Bettinardi et al 2002). The remaining bias is probably a result of the image-based transformations (Dikaios and Fryer 2011, Polycarpou et al 2012). This bias slightly changes for β up to 5 for any of the lesions and noise levels. Within this range of β, RTA-MRP-OSL remains less accurate than MCIR-MRP-OSL. The observed bias/standard deviation trade-off depends on the SUV of the background tissue. For example, the lesions 1 and 4 that are located in the liver have very high standard deviation, which can be explained by their location being near to the edge of the axial field of view and because the liver has high noise due to its high radioactivity. Further investigation of figure 4 can help understand better the effects of MCIR-MRP-OSL on lesion quantification with different size, SUV and background. In particular lesions 7, 8 and 9 have larger bias than the lesions 4, 5, 6 respectively but similar standard deviation. All these lesion pairs (7/4, 8/5, 9/6) have the same SUV and background but different size. Another expected observation is that the larger the SUV is, the larger are the standard deviation and the bias, which occurs in lesions 1, 2, 3 and 4. Still, while the trend of improvement is the same for both RTA-MRP-OSL and MCIR-MRP-OSL, the standard deviation of the latter is significantly reduced without any practical increase in bias. In summary, figure 4 shows that MCIR-MRP-OSL can provide better quantitative information as long as the penalization factor
Figure 7. Coronal planes of two selected cases of $\beta_g$ for RTA-MRP-OSL and MCIR-MRP-OSL of the mean ($M$), bias ($B$), standard deviation ($S$) and root mean square error images ($R$) that have been derived from twenty multiple realizations ($50 \times 10^6$ counts). The images are displayed in SUV scale from: 0 to 4 for the mean, standard deviation and RMSE; and –2 to 2 for the bias with values greater than 2 corresponding to white, lower than –2 to black and 0 to grey.

is chosen properly. It also demonstrates that the results are reproducible for different noise levels and the optimum penalization factor is within similar range. For example with respect to RMSE, as shown in table 2, the optimal $\beta$ for MCIR is $10 \times 8$ while for RTA $\beta_g$ is 5–10.

Figure 5 shows how the regional RMSE values change with penalty. Without penalty, RTA-MRP-OSL comparing to MCIR-MRP-OSL has worse RMSE only in lesion 9. The values of RMSE are minimized for RTA-MRP-OSL with $\beta_g$ between 5 and 20 while for MCIR-MRP-OSL with $\beta_g$ 10 for all lesions apart from the lesion 8. On the other hand, figure 6 shows that the CNR is maximized at $\beta_g = 20–50$ for both methods and all lesions apart from the first. Figure 6 demonstrates also that the CNR with small and large $\beta_g$ is similar for both MCIR-MRP-OSL and RTA-MRP-OSL. However, if we compare the CNR for matched standard deviation (figure 8), MCIR-MRP-OSL is better than RTA-MRP-OSL in all lesions apart from lesion 8, where both methods perform equally well.

Figure 7 shows the corresponding mean ($M$), bias ($B$), standard deviation ($S$) and RMSE ($R$) images for the two methods with $\beta_g$ at matched regional standard deviation ($\sigma$). All the images are visually similar but following the quantitative comparison in figure 8 MCIR-MRP-OSL appears to outperform RTA-MRP-OSL. Bias with RTA-MRP-OSL is probably due to low counts, interpolation errors, regularization, reconstruction inaccuracies (e.g. no resolution modelling) and errors in the numerical inversion of the motion fields. Similarly, in MCIR-MRP-OSL any bias is due to errors in the numerical inversion of the motion fields,
Mean ($m$), standard deviation ($s$), RSME ($r$) and CNR ($c$) for each lesion over twenty independent noise realisations (50x10^6 counts) for RTA and MCIR.

Figure 8. Average values that have been derived from twenty independent noise realizations for each lesion (id: 1–9) for different figures of merits: mean ($m$), standard deviation ($s$), root mean square error ($r$) and contrast to noise ratio ($c$). The error bars indicate the standard error estimated from twenty realizations. Two selected cases of $\beta_g$ for RTA-MRP-OSL ($\beta_g = 2$) and MCIR-MRP-OSL ($\beta = 10 \times 8$) are compared. On the left of each pair of bars is the RTA value and on the right is the MCIR value. The $m$, $s$ and $r$ are in SUV scale, while $c$ is in inverse SUV scale.

regularization and reconstruction inaccuracies. However, an additional advantage of MCIR-MRP-OSL is that the regularization level can be fine-tuned according to the properties being optimized while with RTA-MRP-OSL bias is reduced but only up to a certain extent.

Table 2 demonstrates that the selection of $\beta = 10 \times 8$ is on average the best for MCIR-MRP-OSL with respect to bias, standard deviation and RMSE over all the lesions. In particular, the comparison with $\beta_g = 2$ (matched standard deviation with MCIR-MRP-OSL, $\beta = 10 \times 8$) shows that MCIR-MRP-OSL is on average: 7.0%, 5.5% and 13.8% better in terms of bias, RMSE and CNR than RTA-MRP-OSL. Furthermore, the quantitative statistical comparison derived from twenty independent noise realizations in figure 8 at matched noise, i.e. regional standard deviation $s$, demonstrates clearly that MCIR-MRP-OSL is better than RTA-MRP-OSL with respect to RMSE and CNR for all lesions (apart from lesion 8 for CNR) independent of their size and background. Also, all mean values are higher for MCIR-MRP-OSL indicating that recovery of the SUV is more accurate. Finally, the standard deviation values of the lesions located in the liver are lower for MCIR-MRP-OSL while for all the other lesions they are higher.

Finally, figure 9 demonstrates the impact of regularization effects on bias and standard deviation when iterating longer. Comparing 40 to 10 MCIR-MRP-OSL iterations, the standard deviation is affected only for $\beta_g = 1$ (i.e. 0.3 – 0.5 in SUV), but it is almost insensitive for $\beta_g = 10$ and $\beta_g = 100$. This result is observable in all noise levels. Bias shows only minor reduction for all instances. These results support that iterating longer for MCIR-MRP-OSL is
Figure 9. $\delta b$ and $\delta s$ for three penalization factors ($100, 10, 1$: with higher penalty corresponding to smaller $\delta s$) and three noise levels ($25 \times 10^6, 50 \times 10^6, 100 \times 10^6$ counts) for RTA and MCIR. This graph displays the differences between iterations 40 and 10.

The novel finding demonstrated in this paper is that regularization improves the images in terms of bias mainly for RTA-MRP-OSL and in terms of noise when reconstructed with both MCIR-MRP-OSL and RTA-MRP-OSL. This was not obvious, as it was expected that images reconstructed with MCIR have less noise than those reconstructed with RTA. The latter is more biased but less noisy with the most likely explanations for these being the post-reconstruction transformations (Dikaios and Fryer 2011) and the ill-behaved convergence occurring due to the highly noisy data within each gate, which enhances the low statistics bias and non-negativity constraints (Polycarpou et al 2012). This problem is expected to be larger in the case of less measured counts per respiratory position (Dawood et al 2009) but may improve when we incorporate resolution modelling (Walker et al 2011) and time of flight information (Westerwoudt et al 2012) in the reconstruction. The usefulness of regularization is likely to be similar for any reconstruction algorithm at low count rates. Another way to achieve similar bias-variance performance with regularization is to smooth the images with a filter not necessary, as noise increases with minimum decrease in bias especially for $\beta_g \geq 10$. In the supplementary material (available at http://stacks.iop.org/PMB/58/1759/mmedia), the results show how the mean and the standard deviation are modified with iterations. On the other hand, in RTA-MRP-OSL ($\beta_g < 10$) the results clearly have not converged. Thus, iterating 40 iterations will reduce bias ($0.15 – 0.25$ in SUV) with small increase in the standard deviation (less $0.1 – 0.2$ in SUV) but then only minor improvement in the RMSE (less than $3\%$ reduction for $\beta_g = 1$). Therefore, at relatively late iterations (e.g. 40), the findings of this investigation would mainly show similar performance for both MCIR-MRP-OSL and RTA-MRP-OSL with respect to bias, standard deviation, root mean square error and contrast to noise ratio for $\beta_g \geq 10$. For $\beta_g = 1$, the metrics are affected for both RTA-MRP-OSL and MCIR-MRP-OSL but differently: RTA-MRP-OSL would have slightly better RMSE while MCIR-MRP-OSL would become worse. In the case of small $\beta_g$ values, the results have not stabilized even after 40 iterations. Even if ideally the comparison should have been performed at convergence, iterating for many hours is not practical.
after their reconstruction. This is expected to be a good alternative approach to ‘regularize’
MCIR especially when quantification is not the main issue.

As a side observation, this study highlights that regularization in both MCIR-MRP-OSL
and RTA-MRP-OSL can help balance bias and noise thus improve mean square error and
contrast to noise ratio. On the one hand proper modifications of the iterative reconstruction
methods may be useful to successfully deal with the low statistics bias (Byrne 1998). On the
other hand further investigation on the type of regularization can be useful to minimize noise
but also maintain good resolution in small lesions. One recent example is the incorporation
of non-conventional priors such as patches (Wang and Qi 2012). Further consideration of
regularization may help achieve isotropic uniform resolution in areas moving with different
magnitudes (Chun and Fessler 2012). This approach might also be useful as a means to
account for the effects of inaccuracies in the estimation of motion, which are expected to occur
in clinical data.

6. Conclusion

In this investigation we have demonstrated that MCIR has higher regional standard deviation
and RMSE than RTA unless proper regularization parameters are included within the
reconstruction. As we have shown, MCIR-MRP-OSL with proper regularization outperforms
RTA-MRP-OSL both in terms of bias and RMSE for the optimal penalization factors.
Nevertheless, regularized RTA can also be a practical solution for motion compensation as a
proper level of regularization improves both bias and mean square error.

Acknowledgments

We are thankful to Dr M Jacobson, Dr A Solomon, Dr V Schulz and Dr M Walker for
discussions and Mr A Cantell and Mr D Poccecai (King’s College London) for invaluable
IT support. This work is part of the HYPERImage and SUBLIMA projects supported by
the European Union under the seventh framework program (no: 201651, 241711). Dr K
Thielemans has received financial support from General Electric Research. The authors
acknowledge additional financial support from: (1) the Department of Health via the National
Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to
Guy’s & St Thomas’ NHS Foundation Trust in partnership with King’s College London
and King’s College Hospital NHS Foundation Trust; and (2) the Centre of Excellence in
Medical Engineering funded by the Wellcome Trust and EPSRC under grant number WT
088641/Z/09/Z. The views expressed are those of the authors and not necessarily those of
the NHS, the NIHR or the Department of Health.

References

Alenius S and Ruotsalainen U 1997 Bayesian image reconstruction for emission tomography based on median root
prior Eur. J. Nucl. Med. 24 258–65
Alenius S and Ruotsalainen U 2002 Generalization of median root prior reconstruction IEEE Trans. Med.
Imaging 21 1413–20
Asma E, Manjeshwar R and Thielemans K 2006 Theoretical comparison of motion correction techniques for PET
image reconstruction IEEE Nuclear Science Symp. Conf. Record pp 1762–7
Barrett H H, Wilson D W and Tsui B M W 1994 Noise properties of the EM algorithm: I. Theory Phys. Med.
Biol. 39 833–46
Bettinardi V, Pagani E, Gilardi M, Alenius S, Thielemans K, Teras M and Fazio F 2002 Implementation and evaluation
of a 3D one-step late reconstruction algorithm for 3D positron emission tomography brain studies using median
root prior Eur. J. Nucl. Med. 29 7–18
Regularization in motion incorporated PET reconstruction

Buerger C, Schaefter T and King A P 2011 Hierarchical adaptive local affine registration for fast and robust respiratory motion estimation Med. Image Anal. 15 551–64

Buerger C, Tsoumpas C, Attik A, King A P, Schleyer P, Schulz V, Marsden P K and Schaeffer T 2012 Investigation of MR-based attenuation correction and motion compensation for hybrid PET/MR IEEE Trans. Nucl. Sci. 59 1967–76

Byrne C 1998 Iterative algorithms for deblurring and deconvolution with constraints Inverse Problems 14 1455–67

Chun S Y and Fessler J A 2013 Noise properties of motion-compensated tomographic image reconstruction methods IEEE Trans. Med. Imaging 32 141–52

Chun S Y and Fessler J A 2012 Spatial resolution properties of motion-compensated tomographic image reconstruction methods IEEE Trans. Med. Imaging 31 1413–25

Crum W, Camara O and Hawkes D 2007 Methods for inverting dense displacement fields: evaluation in brain image registration Medical Image Computing and Computer-Assisted Intervention (MICCAI) ed N Ayache et al (Berlin: Springer) pp 900–7

Dawood M, Buther F, Stegger L, Jiang X, Schober O, Schafer M and Schafer K P 2009 Optimal number of respiratory gates in positron emission tomography: a cardiac patient study Med. Phys. 36 1775–84

Dikaios N and Fryer T D 2011 Improved motion-compensated image reconstruction for PET using sensitivity correction per respiratory gate and an approximate tube-of-response backprojector Med. Phys. 38 4958–70

Fessler J A 1994 Penalized weighted least-squares image reconstruction for positron emission tomography IEEE Trans. Med. Imaging 13 290–300

Green P J 1990a Bayesian reconstructions from emission tomography data using a modified EM algorithm IEEE Trans. Med. Imaging 9 84–93

Green P J 1990b On use of the EM algorithm for penalized likelihood estimation R. Stat. Soc. 52 443–52

Lamare F, Carbajo M J L, Cresson T, Kontaxakis G, Santos A, Le Rest C C, Reader A J and Visvikis D 2007 List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations Phys. Med. Biol. 52 5187–204

Nehmeh S A 2013 Respiratory motion correction strategies in thoracic PET-CT PET Clinics 8 29–36

Nuyts J and Fessler J A 2003 A penalized-likelihood image reconstruction method for emission tomography, compared to postsmoothed maximum-likelihood with matched spatial resolution IEEE Trans. Med. Imaging 22 1042–52

Picard Y and Thompson C J 1997 Motion correction of PET images using multiple acquisition frames IEEE Trans. Med. Imaging 16 137–44

Polycarpoiu I, Tsoumpas C and Marsden P K 2012 Analysis and comparison of two methods for motion correction in PET imaging Med. Phys. 39 6474–83

Qi J and Huesman R H 2002 List mode reconstruction for PET with motion compensation: a simulation study IEEE Int. Symp. on Biomedical Imaging: From Nano To Macro pp 413–6

Qiao F, Pan T, Clark J W and Mawlawi O R 2006 A motion-incorporated reconstruction method for gated PET studies Phys. Med. Biol. 51 3769–83

Rahim M, Tang J and Zaidi H 2013 Four-dimensional image reconstruction strategies in cardiac-gated and respiratory-gated PET imaging PET Clinics 8 51–67

Schafer K P 2008 The promise of nuclear medicine technology: Status and future perspective of high-resolution whole-body PET Phys. Medica 24 57–62

Tong S, Alessio A M and Kinahan P E 2010 Noise and signal properties in PSF-based fully 3D PET image reconstruction: an experimental evaluation Phys. Med. Biol. 55 1453

Thielemans K, Manjeshwar R M, Xiaodong T and Asma E 2012 Lesion detectability in motion compensated image IEEE Trans. Med. Imaging 31 2194–204

Thielemans K, Tsoumpas C, Mustafovic S, Beisel T, Aguiar P, Dikaios N and Jacobson M W 2012 STIR: software for tomographic image reconstruction release 2 Phys. Med. Biol. 57 867

Tsoumpas C and Gairtanis A 2013 Modeling and simulation of 4D PET-CT and PET-MR images PET Clinics 8 95–110

Tsoumpas C et al 2011 Fast generation of 4D PET-MR data from real dynamic MR acquisitions Phys. Med. Biol. 56 6597–613

Walker M D, Asselin M C, Jultyan P J, Feldmann M, Talbot P S, Jones T and Matthews J C 2011 Bias in iterative reconstruction of low-statistics PET data: Benefits of a resolution model Phys. Med. Biol. 56 931–49

Wang G and Qi J 2012 Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization IEEE Trans. Med. Imaging 31 2194–204

Westerwoudt V, Conti M and Eriksson L 2012 Comparison of TOF and non-TOF iterative reconstruction at low statistics IEEE Nuclear Science Symp and Medical Imaging Conf. (M08-6) (Anaheim, California, USA)

Wilson D W, Tsui B M W and Barrett H H 1994 Noise properties of the EM algorithm. II. monte carlo simulations Phys. Med. Biol. 39 847–71