PAPER

Similarity-regulation of OS-EM for accelerated SPECT reconstruction

To cite this article: P E B Vaissier *et al* 2016 *Phys. Med. Biol.* **61** 4300

View the article online for updates and enhancements.

Similarity-regulation of OS-EM for accelerated SPECT reconstruction

P E B Vaissier[1,2], F J Beekman[1,2] and M C Goorden[1]

[1] Radiation Detection and Medical Imaging, Delft University of Technology, Delft, The Netherlands
[2] MILabs BV, Utrecht, The Netherlands

E-mail: p.e.b.vaissier@tudelft.nl

Abstract

Ordered subsets expectation maximization (OS-EM) is widely used to accelerate image reconstruction in single photon emission computed tomography (SPECT). Speedup of OS-EM over maximum likelihood expectation maximization (ML-EM) is close to the number of subsets used. Although a high number of subsets can shorten reconstruction times significantly, it can also cause severe image artifacts such as improper erasure of reconstructed activity if projections contain few counts. We recently showed that such artifacts can be prevented by using a count-regulated OS-EM (CR-OS-EM) algorithm which automatically adapts the number of subsets for each voxel based on the estimated number of counts that the voxel contributed to the projections. While CR-OS-EM reached high speed-up over ML-EM in high-activity regions of images, speed in low-activity regions could still be very slow. In this work we propose similarity-regulated OS-EM (SR-OS-EM) as a much faster alternative to CR-OS-EM. SR-OS-EM also automatically and locally adapts the number of subsets, but it uses a different criterion for subset regulation: the number of subsets that is used for updating an individual voxel depends on how similar the reconstruction algorithm would update the estimated activity in that voxel with different subsets. Reconstructions of an image quality phantom and *in vivo* scans show that SR-OS-EM retains all of the favorable properties of CR-OS-EM, while reconstruction speed can be up to an order of magnitude higher in low-activity regions. Moreover our results suggest that SR-OS-EM can be operated with identical reconstruction parameters (including the number of iterations) for a wide range of count levels, which can be an additional advantage from a user perspective since users would only have to post-filter an image to present it at an appropriate noise level.

(Some figures may appear in colour only in the online journal)

1. Introduction

Today, statistical iterative reconstruction algorithms are the method of choice for the reconstruction of single photon emission computed tomography (SPECT) images (Hutton *et al* 1997, Qi and Leahy 2006). The maximum likelihood expectation maximization (ML-EM) algorithm (Shepp and Vardi 1982, Lange and Carson 1984) is a standard in emission tomography because it has consistent convergence behavior, incorporates an intrinsic non-negativity constraint, is applicable to complex scanner geometries, and has the possibility to incorporate advanced physical models of photon transport. In contrast to analytical methods that can reconstruct an image with a single pass through the data, ML-EM reconstruction often requires many iterations to reach sufficient resolution and quantitative accuracy. This can result in prohibitively long reconstruction times. The computational load for ML-EM reconstruction increases when a more accurate model of photon transport from object to projection-space is used, which is essential for e.g. optimal resolution, quantitative accuracy, noise suppression and artifact reduction.

In order to shorten reconstruction times, several block-iterative algorithms have been introduced. These algorithms divide each iteration into a number of subsequent sub-iterations, each one using a different subset of the entire projection data to obtain a next estimate of the activity distribution. A widely used block-iterative version of ML-EM is the ordered subsets expectation maximization (OS-EM) algorithm (Hudson and Larkin 1994). Despite the absence of a mathematical proof of convergence it has been shown that OS-EM often yields almost the same reconstructed images as ML-EM, provided that the number of subsets (NS) used is modest (Kamphuis *et al* 1996, Lalush and Tsui 2000). In such cases, the acceleration factor of OS-EM over ML-EM, in terms of iterative resolution recovery, is roughly proportional to the NS (Hudson and Larkin 1994, Kamphuis *et al* 1996).

Subsets are often selected in a way that each subset contains a number of complete SPECT projections. However, it was shown that it can be a great advantage for multi-pinhole SPECT to use subsets that each contain detector pixels that are distributed over all projections, which is known as pixel-based OS-EM (Branderhorst *et al* 2010). This way, OS-EM is operated with improved subset balance (i.e. all subsets are approximately equally sensitive to a voxel), even for a high NS. As a result, for high count data pixel-based OS-EM reaches acceleration factors an order of magnitude higher than those of traditional OS-EM without significant image degradation.

Even with pixel-based OS-EM artifacts may appear when scan times are short or injected activities are low (Vaissier *et al* 2013). Because the projection data of such scans is noisy due to low numbers of counts, it may happen that all detector pixels associated with a certain voxel (i.e. nonzero detection probability in the system matrix that models photon transport) do not contain any counts. The likelihood that this happens increases if only a subset of detector pixels is used to update the image, as is the case with OS-EM. If such a subset exists, the corresponding voxel will be updated to zero activity in the sub-iteration that uses this subset. Due to the multiplicative nature of the OS-EM update equation, the reconstructed
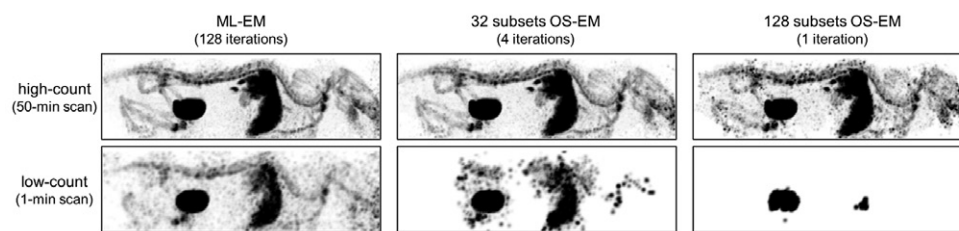
Figure 1. Maximum-intensity-projections acquired from ML-EM and OS-EM (32 and 128 pixel-based subsets) reconstructions of mouse bone scans (189 MBq ^{99m}Tc-HDP) with a high number of counts (top row) and a much lower number of counts (bottom row). While high-count scans can be reconstructed using high acceleration factors (i.e. OS-EM with high NS), low-count scans can suffer from severe quantitative inaccuracies such as erased voxel activity in reconstructed images.

activity in this voxel will remain zero in subsequent sub-iterations despite the fact that other subsets may contain counts that are associated with this particular voxel. This can result in severe image artifacts, which is illustrated in figure 1 that displays maximum-intensity-projections (MIPs) of OS-EM reconstructions of mouse bone scans. In previous work we showed that apparently obvious solutions, such as only allowing nonzero activity updates or simply excluding detector pixels that have no counts, do not solve the problem but can result in severe reconstruction artifacts too (Vaissier *et al* 2013). Operating OS-EM with fewer subsets can prevent erasure of reconstructed activity, but no generally applicable guideline exists to establish the highest NS that can still be safely used. To minimize the risk of inducing such artifacts one could always select a low NS, but this can lead to time-consuming reconstructions.

In order to prevent reconstruction artifacts in image regions with low activity, while still enabling accelerated image reconstruction needed for recovering resolution in structures with high activity uptake, we recently proposed a count-regulated OS-EM (CR-OS-EM) algorithm (Vaissier *et al* 2013). CR-OS-EM uses the concepts of adaptive subset formation and spatially adaptive voxel updates proposed by (Kadrmas 2001). It applies these concepts by combining multiple subsets in order to gather sufficient counts in each subset. The number of subsets that is combined is voxel-specific and depends on the reconstructed activity in that voxel. To this end CR-OS-EM uses a count threshold value (CTV) which only allows the update of a voxel in a sub-iteration of the algorithm if the estimated number of counts that the voxel has contributed to the current (combined) subset is larger than the preset CTV. This way, voxels with low activity will not be updated every sub-iteration, rather they are updated fewer times, each time using the projection pixels of combined subsets. We validated CR-OS-EM for pinhole SPECT of mice covering a wide range of activity distributions and count levels. CR-OS-EM reached speedup factors over ML-EM that were close to the NS in high-activity regions, while at the same time preventing erasure of reconstructed activity in regions with low activity. However, contrast recovery speed with CR-OS-EM decreased sharply with decreasing activity levels and as a result CR-OS-EM was often only as fast as ML-EM.

The aim of the present work is to improve the contrast recovery rate of regulated OS-EM while maintaining its advantages. To this end we developed a new criterion for spatially adaptive subset regulation. The new criterion is based on how similar the reconstruction algorithm would update the estimated activity in a voxel when different subsets of the projection data are used. In ML-EM, the activity of every voxel is multiplied by a voxel-specific update factor

that is calculated using the activity distribution obtained in the previous iteration, the system matrix and the measured projections. If only a subset of the projection data is used in an update (as is the case in each OS-EM sub-iteration), the update factor is subset-dependent. For a pixel-based subset implementation of OS-EM every subset contains detector pixels from all projection views and different subsets are spread out rather equally over the detectors. As a consequence, one would expect that in the limit of (almost) noiseless projections, the update factors that are calculated for the same image with different subsets do not differ too much and are rather equal to the ML-EM update factor. However, when measured projections become noisier (i.e. lower numbers of counts), differences between update factors calculated with different subsets are expected to increase due to statistical variations. Therefore we may be able to use the level of similarity between these different update factors as an indicator of how many subsets have to be combined in order to prevent artifacts in the reconstructions. This similarity-criterion is implemented in a new similarity-regulated OS-EM (SR-OS-EM) algorithm. For each voxel SR-OS-EM combines subsets to a level that the deviations of the subset update factors from the ML-EM update factor are all smaller than a preset percentage called the similarity threshold value (STV). Thus, the STV determines the maximally allowed variation between subset update factors and the ML-EM update factor of each voxel. As a result, the number of updates per voxel is high for voxels that use high-count projection data for their activity update, and lower if voxels use projection data with fewer counts. This work introduces SR-OS-EM and compares its reconstruction speed and quantification accuracy to CR-OS-EM and ML-EM for data acquired with a stationary small-animal multi-pinhole SPECT system.

2. Methods

2.1. Image reconstruction algorithms

2.1.1. ML-EM. ML-EM updates the previous activity estimate $\tilde{a}_i^{(k-1)}$ in voxel i at iteration k by multiplication with an update factor that consists of a correction factor $C_{\text{MLEM},i}^{(k)}$ and a normalization factor $N_{\text{MLEM},i}$:

$$
\tilde{a}_i^{(k)} = \tilde{a}_i^{(k-1)} \frac{C_{\text{MLEM},i}^{(k)}}{N_{\text{MLEM},i}}
$$
$$
C_{\text{MLEM},i}^{(k)} = \sum_j M_{ij} \frac{p_j}{\sum_{i'} M_{i'j} \tilde{a}_{i'}^{(k-1)} + \tilde{s}_j}
$$
$$
N_{\text{MLEM},i} = \sum_j M_{ij} \tag{1}
$$

Here M_{ij} is the system matrix element representing the probability that a non-scattered photon emitted from voxel i is detected in detector pixel j. Furthermore, p_j and $\sum_{i'} M_{i'j} \tilde{a}_{i'}^{(k-1)}$ are the measured and expected number of gamma photon counts in detector pixel j with an energy within the selected photopeak window. The parameter $\tilde{s}_j$ denotes the estimated number of scattered and background counts within the photopeak in detector pixel j.

2.1.2. OS-EM. OS-EM uses the same equations as ML-EM to update the image, but for each update only a subset S_l of the detector pixels is used in each sub-iteration $l(1 \leqslant l \leqslant \text{NS})$ of iteration k.

$$\tilde{a}_i^{(k,l)} = \tilde{a}_i^{(k,l-1)} \frac{C_i^{(k,l)}}{N_i^{(l)}}$$

$$C_i^{(k,l)} = \sum_{j \in S_l} M_{ij} \frac{p_j}{\sum_{i'} M_{i'j} \tilde{a}_{i'}^{(k,l-1)} + \tilde{s}_j}$$

$$N_i^{(l)} = \sum_{j \in S_l} M_{ij} \tag{2}$$

OS-EM iteration k is defined to be completed ('full iteration') when all NS subsets have been traversed. The subsequent sub-iteration $(k+1,1)$ updates the image obtained at sub-iteration (k,NS).

2.1.3. CR-OS-EM. While OS-EM updates all voxels in each sub-iteration, CR-OS-EM only updates a voxel if the number of counts that the voxel contributed to the detector pixels of the current subset is higher than a preset count threshold value (CTV). To this end, in each sub-iteration (k,l) the algorithm estimates the number of counts $\tilde{y}_i^{(k,l)}$ that voxel i contributed to subset S_l:

$$\tilde{y}_i^{(k,l)} = \sum_{j \in S_l} M_{ij} \tilde{a}_i^{(k,l-1)} \tag{3}$$

Voxel i is only updated at sub-iteration (k,l) if $\tilde{y}_i^{(k,l)} \geqslant \text{CTV}$ or if the number of counts that voxel i contributed to all subsets that have been traversed since its last update exceeds the CTV. Once the CTV has been exceeded, the activity in voxel i is updated with an update factor that contains correction and normalization factors of all $x (1 \leqslant x \leqslant \text{NS})$ sub-iterations that have been traversed since its last update:

$$\tilde{a}_i^{(k,l)} = \tilde{a}_i^{(k,l-x)} \frac{C_i^{(k,l-x+1)} + ... + C_i^{(k,l)}}{N_i^{(l-x+1)} + ... + N_i^{(l)}} \qquad\qquad \text{if } x < l$$

$$\tilde{a}_i^{(k,l)} = \tilde{a}_i^{(k-1,\text{NS})} \frac{C_i^{(k,1)} + ... + C_i^{(k,l)}}{N_i^{(1)} + ... + N_i^{(l)}} \qquad\qquad \text{if } x = l$$

$$\tilde{a}_i^{(k,l)} = \tilde{a}_i^{(k-1,\text{NS}+l-x)} \frac{C_i^{(k-1,\text{NS}+l-x+1)} + ... + C_i^{(k-1,\text{NS})} + C_i^{(k,1)} + ... + C_i^{(k,l)}}{N_i^{(\text{NS}+l-x+1)} + ... + N_i^{(\text{NS})} + N_i^{(1)} + ... + N_i^{(l)}} \qquad \text{if } x > l \tag{4}$$

Note that if $x = 1$, (4) reduces to an ordinary OS-EM update. With CR-OS-EM, a low-activity voxel will not be updated in every sub-iteration, rather a combined (larger) subset is used for each update. If a voxel is not updated in NS subsequent sub-iterations, a forced activity update is performed using detector pixels from all subsets. Thus, the number of individual voxel updates within each full iteration of CR-OS-EM lies between 1 and NS. Note that the number of updates that are performed on a voxel can change with new activity estimates and can thus be different for different iterations. To obtain an initial estimate of the activity of individual voxels CR-OS-EM commences with a single ML-EM iteration.

2.1.4. SR-OS-EM. Like CR-OS-EM, SR-OS-EM also starts with a single ML-EM iteration. While calculating the ML-EM update factor for each voxel, the subset update factors (i.e. correction and normalization terms obtained with a subset of the projection data) are calculated as well for each voxel (for the same uniform start image). The subset update factors of voxel i are not used for an update, but are used to determine the percentage deviations $D_i^{(l)}$ from the ML-EM update factor:

$$D_i^{(l)} = \left| \frac{C_i^{(1,l)}}{N_i^{(l)}} - \frac{C_{\mathrm{MLEM},i}^{(1)}}{N_{\mathrm{MLEM},i}} \right| / \frac{C_{\mathrm{MLEM},i}^{(1)}}{N_{\mathrm{MLEM},i}} \times 100\% \tag{5}$$

Note that $C_i^{(1,l)}$ in (5) is calculated using the uniform start image as the current activity estimate for all subsets, i.e. in (2) one would have to use $\tilde{a}_i^{(1,l-1)} = \tilde{a}_i^{\mathrm{start}}$. If $D_i^{(l)} \leqslant$ similarity threshold value (STV) for all subsets, similarity of subset update factors is deemed sufficiently high to justify updates of voxel i in every sub-iteration of all subsequent SR-OS-EM iterations. However, if $D_i^{(l)} >$ STV for one or more subsets, then for this voxel subsequent subsets are paired to form NS/2 larger subsets by adding correction and normalization factors and calculating new update factors. These new update factors are then again compared to the ML-EM update factor and percentage deviations are determined as in (5). This process is repeated until all subset update factors deviate less than the STV from the ML-EM update factor. This way, each voxel will be updated either NS, NS/2, NS/4,…, 4, 2 or 1 times per full iteration in all subsequent SR-OS-EM iterations. Note that if subsets are paired to form new subsets, NS has to be a power of 2. In this work we used a pixel-based subset pattern with 128 subsets as described in (Branderhorst *et al* 2010). The pattern was chosen such that subsequent subsets were spatially well separated to minimize subset unbalance. Consequently, it is most advantageous to combine subsequent subsets as the fewer larger subsets that are formed are still spatially separated.

The voxel update scheme of SR-OS-EM is determined only once in the first iteration. Our reason not to implement SR-OS-EM with update schemes that are re-determined each full iteration lies in the significant overhead that this would introduce: determining the voxel update scheme requires approximately as many calculations as an ML-EM iteration as correction and normalization factors for all subsets have to be calculated. Note that because these subset update factors are calculated using the same image they cannot be used for updating the image as this would imply a 1-subset (ML-EM)-update. As a result the algorithm would be much slower than the current implementation with an update scheme that is determined only once.

2.2. SPECT system

The U-SPECT-II system (MILabs BV, The Netherlands) is a dedicated small-animal SPECT system comprising three stationary detectors with cylindrical multi-pinhole collimators optimized for differently sized rodents. In this work, the general-purpose mouse collimator (75 pinholes, ø 0.6 mm) was used for *in vivo* mouse scans. With this specific collimator the system can reach reconstructed resolutions of 0.4 mm (Van der Have *et al* 2009). The system matrices that were used for image reconstruction of the *in vivo* scans were obtained using the method described in (Van der Have *et al* 2008). In addition, simulated SPECT scans of a digital image quality phantom were performed based on the same collimator and detector geometry. All scans in this work were acquired using the scanning focus method (Vastenhouw and Beekman 2007) that uses all projections of all bed positions together to reconstruct an image, rather than stitching reconstructions of individual bed positions.

2.3. Simulated SPECT scans of a digital image quality phantom

SPECT simulations of a digital image quality phantom were performed to find values of the CTV (CR-OS-EM) and the STV (SR-OS-EM) that maximize reconstruction speed while maintaining quantification accuracy similar to ML-EM. The digital mouse-sized cylindrical
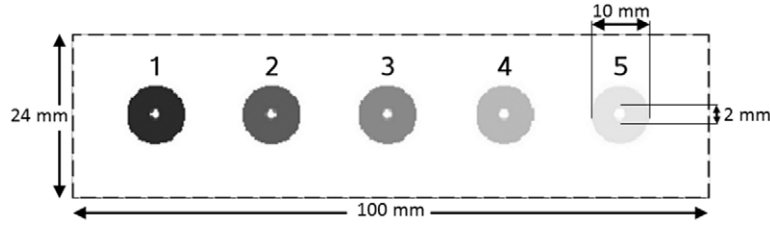
Figure 2. Cylindrical image quality phantom containing 5 activity-filled spheres with spherical cold-lesions (no activity) at their centers. Each consecutive sphere contained ten times lower activity concentration, ranging from 100 MBq ml^{-1} (sphere #1) to 10 kBq ml^{-1} (sphere #5). Background activity concentration was 1 kBq ml^{-1}. Image is shown on a log-grayscale.

phantom (figure 2, voxel size 0.15 mm) used for this purpose was the same as in our previous work (Vaissier *et al* 2013). It contained 5 spheres (ø 10 mm) filled with solution of ^{99m}Tc with different activity concentration. Each consecutive sphere contained a factor 10 lower activity concentration, ranging from 100 MBq ml^{-1} in sphere #1 to 10 kBq ml^{-1} in sphere #5. The remainder of the phantom was filled with a background activity concentration of 1 kBq ml^{-1}. Such large differences in activity concentration might not realistically occur within a single scanned object, but it represents count levels from different scans with large differences in activity and scan time. In the center of each sphere a small spherical cold-lesion (ø 2 mm, no activity) was placed for contrast measurements. Ray tracing was used to simulate noiseless phantom projections and to pre-calculate the system matrix (voxel size 0.3 mm). The simulator accounted for resolution-degrading effects of pinhole diameter, pinhole edge penetration and the intrinsic resolution of the detectors (Gieles *et al* 2002, Goorden *et al* 2011). Poisson statistics were applied to generate 20 noise realizations from the noiseless projections taking into account the total phantom activity (57.7 MBq), scan duration (60 min) and detector efficiency for 140 keV gamma rays (89%).

All noise realizations were reconstructed with ML-EM (gold standard), CR-OS-EM (128 pixel-based subsets) with different CTVs (80k, 40k and 20k counts ml^{-1}) and SR-OS-EM (128 pixel-based subsets) with different STVs (80%, 40% and 20%). OS-EM reconstructions of the image quality phantom are shown for a range of NS (16, 32, 64 and 128) in the supplementary data (stacks.iop.org/PMB/61/4300/mmedia). Images were post-filtered with a 3D Gaussian kernel (0.4 mm full-width-at-half-maximum).

The quantification accuracy of each algorithm was evaluated in each sphere by quantifying the reconstructed activity in a spherical volume-of-interest (ø 12 mm) centered on the sphere. From these values we calculated for each sphere the mean and standard deviation of the activity over the noise realizations, which were expressed as percentages of the true activity.

Cold-lesion contrast recovery in the spheres was used as a performance measure for iterative reconstruction speed. Cold-lesion contrast inside sphere #n at iteration k was defined as the contrast between the average reconstructed activity per voxel $\langle A_{hot}^{(n,k)} \rangle$ in a spherical shell (ø 4–8 mm) inside the sphere and the average reconstructed activity per voxel $\langle A_{cold}^{(n,k)} \rangle$ inside a spherical volume (ø 1 mm) centered on the cold-lesion:

$$\text{Contrast}^{(n,k)} = \frac{\langle A_{hot}^{(n,k)} \rangle - \langle A_{cold}^{(n,k)} \rangle}{\langle A_{hot}^{(n,k)} \rangle} \tag{6}$$

The normalized standard deviation of the reconstructed activity in all Q voxels that were used to calculate $\langle A_{cold}^{(n,k)} \rangle$ was a measure of the noise in sphere #n at iteration k:

$$\text{Noise}^{(n,k)} = \frac{1}{\left\langle A_{\text{cold}}^{(n,k)} \right\rangle} \sqrt{\frac{\sum\limits_{q=1}^{Q} (A_{\text{cold}}^{(n,k)}(q) - \left\langle A_{\text{cold}}^{(n,k)} \right\rangle)^2}{Q - 1}} \tag{7}$$

The average contrast and the average noise at iteration k were obtained by averaging contrast and noise over all noise realizations.

We determined the contrast recovery speeds of CR-OS-EM and SR-OS-EM with respect to ML-EM by comparing the number of iterations that were required to achieve approximately equal contrast and noise (different algorithms never achieve the exact same contrast-noise). Iterations that were compared this way were selected in the following manner. Let $\text{Contrast}_{\text{max}}^{(n)}$ be the maximum contrast value in sphere $\#n$ obtained by ML-EM over all iterations that were performed. We first determined the lowest iteration numbers at which ML-EM recovered a contrast of at least 95% and 99% (sphere $\#1$), or 80%, 95% and 99% (spheres $\#3$ and $\#5$) of $\text{Contrast}_{\text{max}}^{(n)}$. Subsequently, the corresponding ML-EM noise levels at these iteration numbers were determined. Then, the iteration numbers with a noise value closest to the ML-EM noise levels were determined for CR-OS-EM and SR-OS-EM.

2.4. Mouse SPECT

SPECT scans of mice that cover a wide range of count levels were reconstructed to evaluate the performance of CR-OS-EM and SR-OS-EM on *in vivo* data. ML-EM images with 128 iterations served as references to CR-OS-EM and SR-OS-EM images with 2, 4, 6 and 8 iterations. CR-OS-EM was operated with CTV = 40k counts ml^{-1} and SR-OS-EM with STV = 40% and both were operated with 128 pixel-based subsets. These settings were selected based on the results of the simulated image quality phantom scans. OS-EM reconstructions of the *in vivo* scans are shown for a range of NS (16, 32, 64 and 128) in the supplementary data. Reconstructions were post-filtered with a 3D Gaussian kernel (1 mm full-width-at-half-maximum).

A 30 g male mouse was anesthetized with isoflurane and ^{99m}Tc-hydroxymethylene diphosphonate (^{99m}Tc-HDP, 189 MBq) was administered via injection in the tail vein. Scanning started right before injection. The mouse was scanned for 60 min in frames of 1 min. High-count projection data was obtained by summing the projection data of the final 50 frames (50 min scan). Projection data with few counts was obtained by using only the projection data from the last frame (1 min scan). This experiment was conducted following protocols approved by the Animal Research Committee of the University Medical Center Utrecht.

Furthermore, a low-count mouse tumor scan was performed. A 20 g male mouse was inoculated in the shoulder with 22Rv1 human prostatic carcinoma cells. The tumor developed for 3 weeks and the mouse was then intravenously injected with a proprietary ^{99m}Tc-prostate-specific-membrane-antigen-targeted (^{99m}Tc-PSMA, 16.7 MBq) radioligand (Kularatne *et al* 2009). Four hours after injection the mouse was sacrificed and a 32 min scan was performed. This experiment was conducted in accordance with Purdue University Animal Care and User Committee guidelines.

Scatter correction was performed using a triple-energy-window technique (King *et al* 1997). The estimated scatter projections were scaled according to the ratio of the window widths, and were incorporated in the reconstructions according to the method of Bowsher *et al* (1996). Table 1 provides for each scan the number of counts in the photopeak and scatter windows.

Table 1. Number of counts in photopeak and scatter windows of *in vivo* scans.

	# counts in photopeak window	# counts in scatter windows
50 min bone scan	15.9×10^6	2.62×10^6
1 min bone scan	0.29×10^6	0.04×10^6
32 min tumor scan	0.26×10^6	0.16×10^6

Table 2. Reconstructed activities in spheres of image quality phantom as percentages of true activities (mean and standard deviation over 20 noise realizations).

	Sphere #1	Sphere #2	Sphere #3	Sphere #4	Sphere #5
ML-EM	99.9 ± 0.02	100 ± 0.06	100 ± 0.26	99.8 ± 0.90	90.9 ± 2.35
CR-OS-EM (CTV = 80k)	100 ± 0.22	99.8 ± 0.35	99.8 ± 0.52	99.3 ± 0.91	90.7 ± 2.38
CR-OS-EM (CTV = 40k)	100 ± 0.23	99.8 ± 0.40	99.5 ± 0.57	99.1 ± 0.94	90.5 ± 2.29
CR-OS-EM (CTV = 20k)	99.9 ± 0.23	99.8 ± 0.51	99.1 ± 0.72	98.5 ± 1.27	89.7 ± 2.38
SR-OS-EM (STV = 80)	100 ± 0.23	99.9 ± 0.82	100 ± 1.92	100 ± 2.67	91.3 ± 3.45
SR-OS-EM (STV = 40)	100 ± 0.23	99.9 ± 0.74	100 ± 1.03	100 ± 1.90	91.5 ± 2.73
SR-OS-EM (STV = 20)	100 ± 0.23	99.9 ± 0.42	99.8 ± 0.65	99.9 ± 1.24	91.2 ± 2.46

Note: activity in each sphere was determined in the iteration that achieved a contrast level similar to the ML-EM iteration with 95% of the maximum recovered contrast. For all algorithms activity quantification became less reliable with decreasing number of counts. For CR-OS-EM (NS = 128) we found that the quantification accuracy became worse for lower (less strict) CTVs as activities deviated more from the true activity and standard deviations increased. Quantification accuracy of SR-OS-EM (NS = 128) became worse with higher (less strict) STVs as standard deviations increased.

3. Results

3.1. Simulated SPECT scans of a digital image quality phantom

Figure S1 in the supplementary data shows OS-EM image slices for different NS. These images show that OS-EM resulted in artifacts in the low-activity spheres for all NS, and even led to erasure of reconstructed activity in spheres #4 and #5. The severity of the artifacts is quantified in table S1 by determining reconstructed activities in all spheres (mean and standard deviation over 20 noise realizations) and comparing these to the true sphere activities. The performance of ML-EM, CR-OS-EM and SR-OS-EM was quantified in the same way (table 2) and corresponding image slices are displayed in figure 3. Figure 3(a) shows that CR-OS-EM and SR-OS-EM images did not have the low-count induced artifacts that were observed for OS-EM. Figure 3(b) shows the number of voxel updates per full iteration for CR-OS-EM and SR-OS-EM for the same image slices that are displayed in figure 3(a). This figure illustrates that CR-OS-EM and SR-OS-EM generally allow most updates per full iteration (i.e. more subsets) in high-activity regions of the image. For CR-OS-EM the number of updates in a voxel appears to be directly related to the level of the reconstructed voxel activity while in SR-OS-EM low-activity voxels close to high-activity regions are also frequently updated. As expected, for CR-OS-EM a lower CTV allows for more voxel updates per full iteration, while for SR-OS-EM a higher STV allows for more voxel updates.

Contrast recovery speed of CR-OS-EM and SR-OS-EM was evaluated with a value of the CTV (CR-OS-EM) and STV (SR-OS-EM) that resulted in similar quantification accuracy as ML-EM for all spheres of the image quality phantom. To this end, we first selected from the tested CTVs and STVs those that resulted in reconstructed sphere activities that were at most
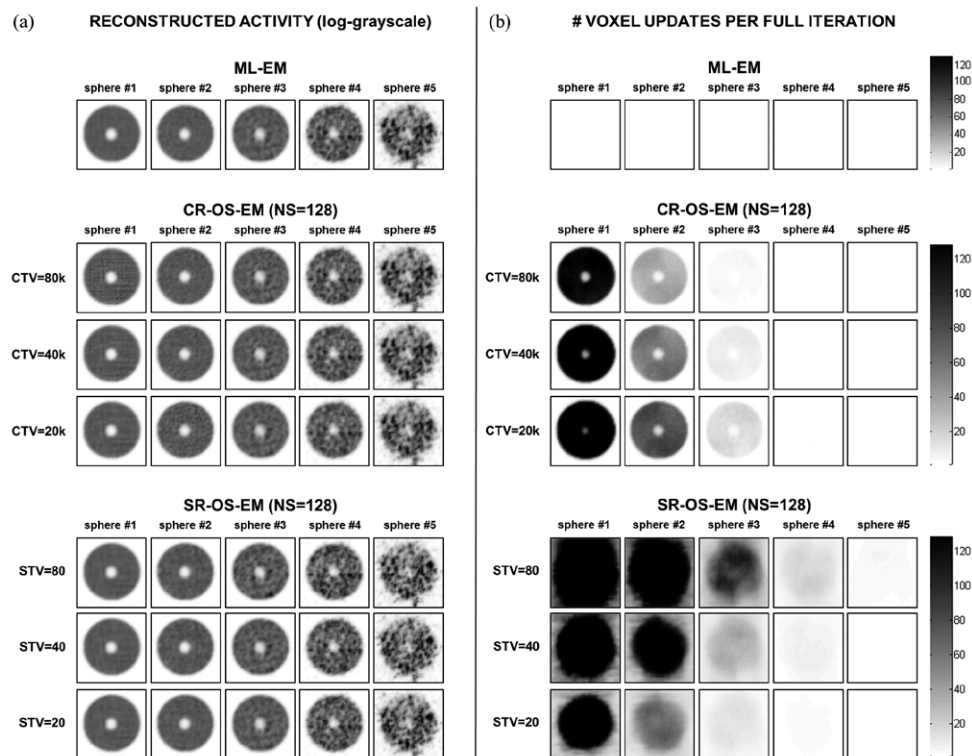
Figure 3. (a) Slices through sphere centers of average images (average over 20 noise realizations) of image quality phantom (57.7 MBq, 60 min) for ML-EM, CR-OS-EM and SR-OS-EM. Spheres #1–5 are represented on different grayscales to allow for simultaneous visualization. Every sphere is shown at the iteration number which achieved a contrast level similar to 95% of the maximum contrast recovered by ML-EM in that sphere. Quantified sphere activities are found in table 2. (b) Corresponding images showing the number of voxel updates per full iteration in the spheres. With CR-OS-EM a lower CTV allows for more voxel updates, while with SR-OS-EM a higher STV allows for more voxel updates.

1% different from those obtained with ML-EM in terms of mean and standard deviation of the sphere activities over 20 noise realizations (table 2). CR-OS-EM met this requirement for CTV = 40k and 80k, SR-OS-EM for STV = 40% and 20%. From these values we selected CTV = 40k and STV = 40% because these allowed the highest number of voxel updates and therefore are expected to result in the highest reconstruction speeds.

Figure 4 displays plots of contrast versus noise of CR-OS-EM and SR-OS-EM with these parameter settings for sphere #1 (highest activity), sphere #3 (medium activity) and sphere #5 (lowest activity). The solid lines are contrast-noise curves obtained with ML-EM. The symbols inside each of the dashed rectangles denote contrast-noise obtained with CR-OS-EM and SR-OS-EM closest to a specific contrast level of ML-EM, namely 95% and 99% of the maximum ML-EM contrast (sphere #1) or 80%, 95% and 99% of the maximum ML-EM contrast (spheres #3 and #5). The corresponding iteration numbers are provided between square brackets in each graph.

From the iteration numbers in the graphs of figure 4 one can infer that CR-OS-EM resulted in speedup factors over ML-EM ranging from about 28 to 33 in sphere #1, while for the same
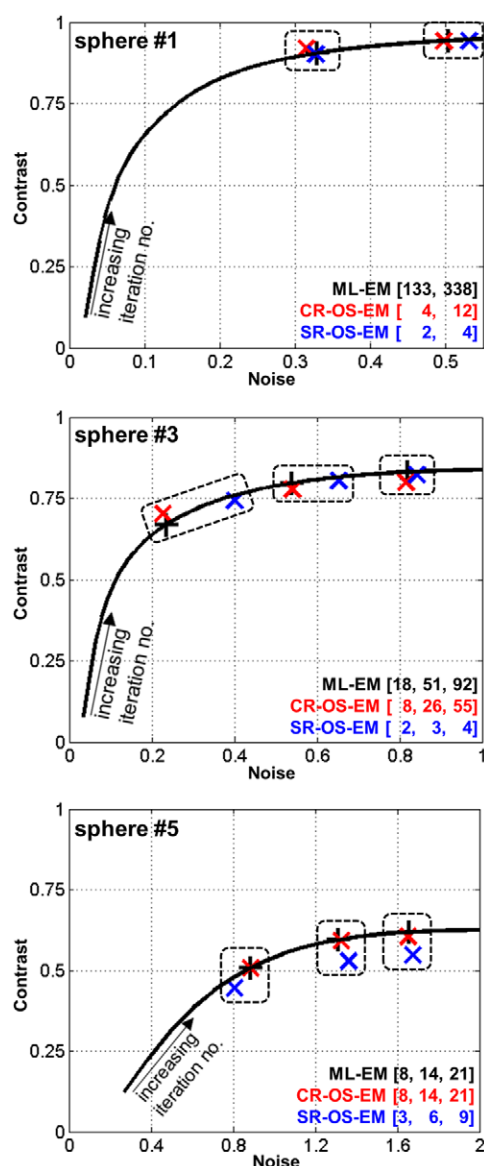
Figure 4. Plots of contrast versus noise in sphere #1 (top graph), sphere #3 (middle graph) and sphere #5 (bottom graph) of image quality phantom (57.7 MBq, 60 min, 20 noise realizations). Black crosses represent ML-EM iterations at which 95% and 99% (sphere #1), or 80%, 95% and 99% (spheres #3 and #5) of the maximum ML-EM contrast was recovered. As a reference, full ML-EM curves are also plotted (solid lines). Red and blue crosses in the same dashed rectangles represent iterations of CR-OS-EM (NS = 128, CTV = 40k counts ml^{-1}) and SR-OS-EM (NS = 128, STV = 40%) with noise values closest to the selected ML-EM points. Corresponding iteration numbers are provided between square brackets. From these iteration numbers follows that speedup factors of SR-OS-EM over ML-EM ranged from 85 in sphere #1 down to >2.0 in sphere #5. Contrast recovery with SR-OS-EM was always faster than with CR-OS-EM, with SR-OS-EM reaching up to an order of magnitude faster contrast recovery in sphere #3.
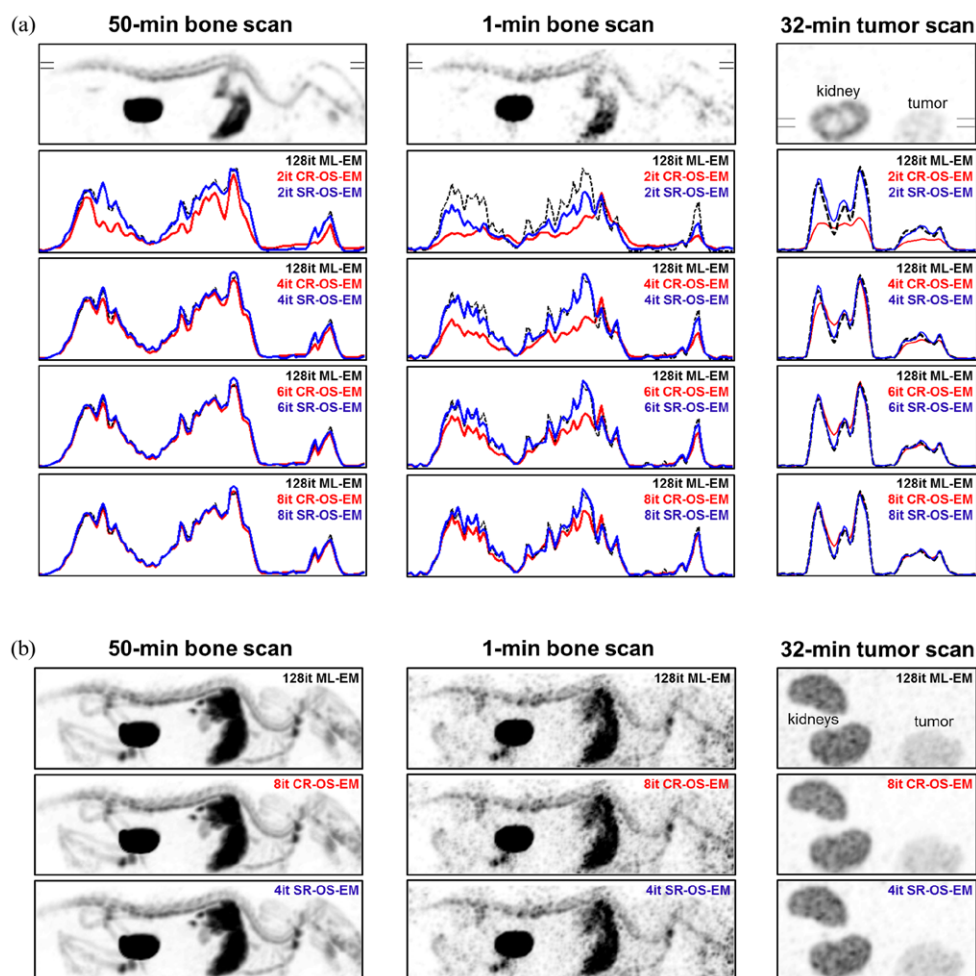
Figure 5. (a) Top row: image slices through ML-EM images (128 iterations) of mouse bone scans (189 MBq ^{99m}Tc-HDP, 50 min and 1 min) and mouse tumor scan (16.7 MBq ^{99m}Tc-PSMA (Kularatne *et al* 2009), 32 min). Lower rows: for each scan profiles for CR-OS-EM (NS = 128, CTV = 40k counts ml^{-1}) and SR-OS-EM (NS = 128, STV = 40%) at different numbers of iterations together with ML-EM profile with 128 iterations (reference profile). Profile regions are indicated in ML-EM image slices (top row). CR-OS-EM required 8 iterations to reduce profile deviations of all scans to <5% with respect to ML-EM, while SR-OS-EM required only 4 iterations (table 3). (b) MIPs for ML-EM (128 iterations), CR-OS-EM (8 iterations) and SR-OS-EM (4 iterations).

sphere SR-OS-EM reached speedup factors over ML-EM ranging from 67 to 85. Note that if the first iteration of SR-OS-EM (i.e. the ML-EM iteration) would be disregarded, speedup factors of SR-OS-EM over ML-EM would be very close to the NS (128), which is the highest acceleration one can expect. In sphere #3 contrast recovery of CR-OS-EM and SR-OS-EM was slower: CR-OS-EM resulted in speedup factors over ML-EM that ranged from about 1.7 to 2.3, while SR-OS-EM reached speedup factors over ML-EM that ranged from 9.0 to 23. Thus, contrast recovery of SR-OS-EM was up to an order of magnitude faster than that of CR-OS-EM. Note that the clusters of points with similar noise and contrast in sphere #3 were

Table 3. Mean deviations of CR-OS-EM (NS = 128, CTV = 40k) and SR-OS-EM (NS = 128, STV = 40%) image profiles from ML-EM image profiles (128 iterations).

Iterations	(High-count) 50 min bone scan		(Low-count) 1 min bone scan		(Low-count) 32 min tumor scan	
	CR-OS-EM	SR-OS-EM	CR-OS-EM	SR-OS-EM	CR-OS-EM	SR-OS-EM
2	11%	1.6%	17%	9.9%	11%	3.4%
4	4.0%	1.5%	13%	3.3%	4.1%	1.8%
6	2.5%	1.5%	7.9%	2.6%	2.7%	1.4%
8	1.9%	1.5%	4.9%	2.3%	2.0%	1.2%

Note: Values are percentages of maximum value of ML-EM profiles.

sometimes rather far apart which is a result of the high contrast recovery speed of SR-OS-EM: only 2 iterations SR-OS-EM recovered ≫ 80% of the maximum ML-EM contrast and 3 iterations SR-OS-EM recovered ≫ 95% of the maximum ML-EM contrast. So, actual speedup of SR-OS-EM over ML-EM and CR-OS-EM in sphere #3 was even higher than aforementioned. Finally, from figure 4 follows that contrast recovery speed of CR-OS-EM in sphere #5 was equal to that of ML-EM. SR-OS-EM achieved 2–3 times higher speed than ML-EM and CR-OS-EM in this sphere. Figure 4 also shows that contrast recovered by SR-OS-EM in sphere #5 was lower than that of ML-EM and CR-OS-EM. However, the differences in contrast between the algorithms were smaller than the standard deviations of the contrast over the noise realizations. Such high standard deviations were a result of the very low-count projections associated with sphere #5. Results of contrast recovery in spheres #1, #3 and #5 for the other CTVs and STVs can be found in figure S2 in the supplementary data.

3.2. Mouse SPECT

The top row of figure 5(a) shows slices through the ML-EM (128 iterations) images of the mouse bone scans and mouse tumor scan. Profile regions are indicated in the image slices. These ML-EM profiles are used as references and are compared to the image profiles of CR-OS-EM (NS = 128, CTV = 40k counts ml^{-1}) and SR-OS-EM (NS = 128, STV = 40%) with 2, 4, 6 and 8 iterations. This way, the differences in speed at which small anatomical features are recovered can be visualized and quantified.

Table 3 contains the quantitative profile deviations from the reference (ML-EM) profiles. Profile deviation is defined as the mean of the absolute intensity differences expressed as percentages of the maximum ML-EM profile intensity. For an equal number of iterations, the deviations of SR-OS-EM were smaller than those of CR-OS-EM for all scans. Moreover, if we take the deviation of 8 iterations CR-OS-EM as a benchmark, we find that SR-OS-EM reached smaller deviations after only 2 iterations (50 min bone scan) or 4 iterations (1 min bone scan and 32 min tumor scan). Thus, while CR-OS-EM required a maximum of 8 iterations to reduce the profile deviations of all scans to <5%, SR-OS-EM only required a maximum of 4 iterations to achieve this. Figure 5(b) displays MIPs of ML-EM images with 128 iterations together with MIPs of CR-OS-EM and SR-OS-EM images with 8 and 4 iterations, respectively. The results of the *in vivo* scans suggest that SR-OS-EM can be operated with a single choice of the reconstruction parameters (NS = 128, STV = 40% and 4 iterations) for a wide range of count levels while CR-OS-EM should be operated with more iterations (NS = 128, CTV = 40k and 8 iterations) to achieve similar results.

OS-EM reconstructions of the *in vivo* scans are shown for a range of NS in the supplementary data (figure S3). These images show that OS-EM resulted in low-count induced artifacts which were more severe for higher NS.

4. Conclusion and discussion

Recently we showed that OS-EM can cause severe quantitative inaccuracies such as improper erasure of reconstructed activity in image regions with low activity. As an initial solution we proposed CR-OS-EM which can attain high speedup factors over ML-EM in image regions with high activity and prevents erasure of reconstructed activity in image regions with low activity. However, contrast recovery with CR-OS-EM in regions with low activity can still be very slow. In this paper we introduced and validated SR-OS-EM as a fast alternative to CR-OS-EM.

The *in vivo* scans were all acquired with the same collimator and radioisotope. For these scans we showed that quantitative differences between SR-OS-EM and ML-EM (128 iterations) were reduced <5% within four iterations of SR-OS-EM. These results suggest that SR-OS-EM can be operated with an identical parameter setting for a wide range of count levels. If this parameter setting results in similar performance with other collimators and radioisotopes remains to be investigated. The advantage of a single parameter setting for a user is that they do not have to compromise between resolution and quantitative accuracy as they only have to use a post-filter to present the image at an appropriate noise level.

The NS determines the maximum possible number of individual voxel updates per full SR-OS-EM iteration. A lower or higher NS would respectively reduce or increase the maximum possible number of voxel updates which could result in different reconstruction speeds for high-count data, as for high-count data SR-OS-EM may actually use all subsets individually instead of combining them into larger subsets. Therefore, ideally one would always like to choose a large enough NS such that the reconstruction speed is never limited by NS. In this work we chose to use NS = 128 because it has been shown by (Branderhorst *et al* 2010) that this high NS still results in balanced subsets.

The reconstruction times with 16 CPUs (16 × AMD Opteron 6344, 2.6 GHz) of a single iteration ML-EM, CR-OS-EM and SR-OS-EM were determined for the image quality phantom. These amounted to 6 min 3 s (ML-EM), 11 min 55 s (CR-OS-EM) and 9 min 33 s (SR-OS-EM). The extra time that was required for a CR-OS-EM or SR-OS-EM iteration compared to ML-EM depended on the number of CPUs used; reconstruction times of CR-OS-EM and SR-OS-EM compared to ML-EM were both only 13% longer if reconstruction was performed on a single CPU. The larger relative differences in reconstruction times between ML-EM and CR-OS-EM/SR-OS-EM with 16 CPUs compared to 1 CPU are probably a result of the way the algorithms were parallelized. Note that although the time per iteration of CR-OS-EM and SR-OS-EM is larger than that of ML-EM, both CR-OS-EM and SR-OS-EM require many fewer iterations than ML-EM to reconstruct high-resolution images.

Contrast recovery speedup of CR-OS-EM over ML-EM decreased with increasing numbers of iterations in spheres #1 and #3 of the phantom (figure 4). We believe that the reason for this is that in CR-OS-EM the number of updates of a voxel directly depends on the reconstructed activity values of voxels, which is illustrated in figure 3(b): CR-OS-EM updates the cold lesions less often than the hot spheres. With higher iterations the reconstructed activity in the cold-lesion voxels becomes lower and therefore the number of updates decreases, which slows down contrast recovery. The number of voxel updates with SR-OS-EM does not directly depend on the reconstructed activity values of voxels, rather it is directly related to the quality

of the projection data, which is comparable for neighboring cold and hot voxels. As a result, cold and hot voxels that lie in the same neighborhood are updated at a similar rate. This can also be inferred from figure 3(b): SR-OS-EM updates the cold lesions approximately as frequently as the hot surrounding.

Like CR-OS-EM, SR-OS-EM uses the concepts of adaptive subset formation and spatially adaptive voxel updates. The number of voxel updates is therefore likely to be non-uniform over the image, which can result in non-uniform image resolution. Users should be aware of non-uniform resolution as this may introduce errors in quantitative comparisons, although we showed in this work that quantitative differences between SR-OS-EM and ML-EM were small.

Besides CR-OS-EM and SR-OS-EM, other algorithms have been proposed that also handle subsets differently than OS-EM, such as StatREM (Kadrmas 2001), RAMLA (Browne and De Pierro 1996) and accelerated C-OS-EM (Hsiao and Huang 2010). We already tested StatREM in our previous work (Vaissier *et al* 2013) and found that it did not always prevent erasure of reconstructed activity in image regions with low activity. RAMLA uses relaxation within a modified version of OS-EM. There are no rules for finding optimal relaxation sequences in terms of reconstruction speed. Therefore, a separate optimization study is required and a comparison between RAMLA and SR-OS-EM remains a topic for future investigations. Accelerated C-OS-EM does not require a relaxation schedule but it requires an acceleration factor to be set similar to accelerated ML-EM (Hwang and Zeng 2006). Although accelerated C-OS-EM was found to be faster than (parameter-free) C-OS-EM, there was no obvious additional speed increase of accelerated C-OS-EM for NS higher than 16.

Acknowledgments

References

Branderhorst W, Vastenhouw B and Beekman F J 2010 Pixel-based subsets for rapid multi-pinhole SPECT reconstruction *Phys. Med. Biol.* **55** 2023–34

Bowsher J E, Johnson V E, Turkington T G, Jaszczak R J, Floyd C R and Coleman R E 1996 Bayesian reconstruction and use of anatomical *a priori* information for emission tomography *IEEE Trans. Med. Imaging* **15** 673–86

Browne J and De Pierro A R 1996 A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography *IEEE Trans. Med. Imaging* **15** 687–99

Gieles M, de Jong H W and Beekman F J 2002 Monte Carlo simulations of pinhole imaging accelerated by kernel-based forced detection *Phys. Med. Biol.* **47** 1853–67

Goorden M C, Van der Have F, Kreuger R and Beekman F J 2011 An efficient simulator for pinhole imaging of PET isotopes *Phys. Med. Biol.* **56** 1617–34

Hudson H M and Larkin R S 1994 Accelerated image-reconstruction using ordered subsets of projection data *IEEE Trans. Med. Imaging* **13** 601–9

Hsiao I T and Huang H M 2010 An accelerated ordered subsets reconstruction algorithm using an accelerating power factor for emission tomography *Phys. Med. Biol.* **55** 599–614

Hutton B F, Hudson H M and Beekman F J 1997 A clinical perspective of accelerated statistical reconstruction *Eur. J. Nucl. Med.* **24** 797–808 (PMID: 9211768)

Hwang D and Zeng G L 2006 Convergence study of an accelerated ML-EM algorithm using bigger step size *Phys. Med. Biol.* **51** 237–52

Kadrmas D J 2001 Statistically regulated and adaptive EM reconstruction for emission computed tomography *IEEE Trans. Nucl. Sci.* **48** 790–8

Kamphuis C, Beekman F J and Viergever M A 1996 Evaluation of OS-EM vs ML-EM for 1D, 2D and fully 3D SPECT reconstruction *IEEE Trans. Nucl. Sci.* **43** 2018–24

King M A, deVries D J, Pan T S, Pretorius P H and Case J A 1997 An investigation of the filtering of TEW scatter estimates used to compensate for scatter with ordered subset reconstructions *IEEE Trans. Nucl. Sci.* **44** 1140–5

Kularatne S A, Zhou Z, Yang J, Post C B and Low P S 2009 Design, synthesis, and preclinical evaluation of prostate-specific membrane antigen targeted (99m)Tc-radioimaging agents *Mol. Pharm.* **6** 790–800

Lalush D S and Tsui B M 2000 Performance of ordered-subset reconstruction algorithms under conditions of extreme attenuation and truncation in myocardial SPECT *J. Nucl. Med.* **41** 737–44 (PMID: 10768577)

Lange K and Carson R 1984 EM reconstruction algorithms for emission and transmission tomography *J. Comput. Assist. Tomogr.* **8** 306–16 (PMID: 6608535)

Qi J and Leahy R M 2006 Iterative reconstruction techniques in emission computed tomography *Phys. Med. Biol.* **51** R541–78

Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **1** 113–22

Vaissier P E B, Goorden M C, Taylor A B and Beekman F J 2013 Fast count-regulated OSEM reconstruction with adaptive resolution recovery *IEEE Trans. Med. Imaging* **32** 2250–61

Van der Have F, Vastenhouw B, Rentmeester M and Beekman F J 2008 System calibration and statistical image reconstruction for ultra-high resolution stationary pinhole SPECT *IEEE Trans. Med. Imaging* **27** 960–71

Van der Have F, Vastenhouw B, Ramakers R M, Branderhorst W, Krah J O, Ji C, Staelens S G and Beekman F J 2009 U-SPECT-II: an ultra-high-resolution device for molecular small-animal imaging *J. Nucl. Med.* **50** 599–605

Vastenhouw B and Beekman F 2007 Submillimeter total-body murine imaging with U-SPECT-I *J. Nucl. Med.* **48** 487–93 (PMID: 17332628)