

# Iterative reconstruction with segmentation penalty for PET

Yueyang Teng, Yaonan Zhang and Yan Kang\*

*Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang, 110004, China*

**Abstract.** Segmentation technique is widely accepted to reduce noise propagation from transmission scanning for positron emission tomography. The conventional routine is to sequentially perform reconstruction and segmentation. A smoothness penalty is also usually used to reduce noise, which can be imposed to both the ML and WLS estimators. In this paper we replace the smoothness penalty by a segmentation penalty that biases the object toward piecewise-homogeneous reconstruction. Two updating algorithms are developed to solve the penalized ML and WLS estimates, which monotonically decrease the cost functions. Experimental results on simulated phantom and real clinical data were both given to demonstrate the effectiveness and efficiency of the algorithms which were proposed.

Keywords: Segmentation penalty, fuzzy c-means clustering (FCM), space alternating descent, auxiliary function

## 1. Introduction

Many reconstruction methods, for example analytic and iterative methods, have been recently developed for positron emission tomography (PET). The iterative methods can suppress noise and model imaging physics better than the analytic methods, so that they attract more and more attention. Iterative methods mainly consist of the maximum likelihood (ML) model and the weighted least squares (WLS) estimate.

The ML-expectation maximization (MLEM) algorithm [1] has widely been used in reconstruction. It proves that the counts at each projection follow Poisson distribution. The opposite likelihood function can be viewed as the measurement between real projections and estimated ones by KL-divergence [2]. Considering the fact that a projection approximates its estimate, Fessler [3] has proposed a WLS estimator. To reduce the noise level, some smoothness penalties, which depend on some priori knowledge, have been introduced [4-6].

In transmission scan, segmentation attenuation correction is usually used to obtain low-noise attenuation coefficients since the attenuation quantity in some anatomic regions varies very little. In the conventional routine, reconstruction is firstly performed, and then it is segmented into several regions with respect to attenuation coefficients, so that average values can be calculated from these regions and assigned to the corresponding regions. By forwardly projecting the segmented image, new

---

\* Address for correspondence: Yan Kang, Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang, 110004, China. Tel.: +86-24-83689210; Fax: +86-24-83689211; E-mail: kangyan@bmie.neu.edu.cn.

attenuation coefficients can be obtained to correct emission projections. Here, fully-automated image segmentation is an important procedure. Fuzzy c-means (FCM) clustering is a powerful tool to do it, which can partition the image by thresholding the grey-levels corresponding to different regions.

In this paper we impose FCM as a segmentation penalty to both the ML and WLS estimators. Two update rules are proposed by solving the modified models according to the concepts of space alternating descent and auxiliary function. The space alternating descent is to update the parameters sequentially by alternating between several variable spaces predefined. The auxiliary function is constructed on the last iteration, and the current iteration is obtained at its minimum that can ensure the cost function decreasing. Just like the MLEM, our algorithms self-constrain in the non-negative pixels without step-size factor required. Simulated data and real clinical data can both evaluate the performance of new algorithms, which showed them effective and more robust to noise than the conventional method.

## 2. Methodology

### 2.1. ML and WLS estimates

We denote all vectors as column vectors unless specified, where  $x \in R^J$  are denoted as the image and  $y \in R^I$  are denoted as the projections;  $x_j$  is the value of the  $j$ -th pixel and  $y_j$  is the value of the  $i$ -th projection. And denote the system matrix by  $W \in R^{I \times J}$  satisfying  $\sum_{i=1}^I W_{ij} = 1$ .

Assuming that each projection was Poisson variable, the log-likelihood function, which was used for reconstruction, was introduced by Shepp and Vardi. The ML estimate consists in choosing the image to maximize it. The likelihood function may be viewed as the opposite KL-divergence between the real projections and the estimated projections. The ML estimate amounts to minimize the KL-divergence.

$$F_L(x) = \sum_{i=1}^I [(Wx)_i - y_i \ln(Wx)_i]$$

Fessler considered each projection approximating its estimate, and then the WLS estimator was introduced

$$F_w(x) = \frac{1}{2} \sum_{i=1}^I \frac{[y_i - (Wx)_i]^2}{D_{ii}}$$

Where  $D$  represents a diagonal weight matrix predetermined.

To minimize  $F_L$  and  $F_w$  under the constraints  $x_j \geq 0$  for all  $j$ , the lemma below is very important.

**Lemma1** Let  $g(x, x')$  be an auxiliary function for  $G(x)$  on  $x'$  if  $g(x', x') = G(x')$  and  $g(x, x') \geq G(x)$ , then  $G$  decreases under the update  $x'' = \min_x g(x, x')$ .

It is easy to be proven by  $G(x'') \leq g(x'', x') \leq g(x', x') = G(x')$ . The lemma has been used in many places [7, 8].

Let  $\lambda_{ij} = W_{ij}x_j^t / (Wx^t)_i$ , it is well-known that  $F_L$  and  $F_w$  has an auxiliary function on the  $t$ -th

iteration  $x^t$  as the following [2, 9].

$$f_L(x, x^t) = \sum_{i=1}^I [(Wx)_i - y_i \sum_{j=1}^J \lambda_{ij} \ln \frac{W_{ij}x_j}{\lambda_{ij}}]$$

$$f_W(x, x^t) = \frac{1}{2} \sum_{i=1}^I \frac{1}{D_{ii}} [y_i^2 - 2y_i(Wx)_i + \sum_{j=1}^J \lambda_{ij} (\frac{W_{ij}x_j}{\lambda_{ij}})^2]$$

### 2.2. Fuzzy c-means clustering

For the data set  $\{x_1, \dots, x_J\}$ , we assign the data into the cluster centers  $\{c_1, \dots, c_L\}$ . FCM is formulated as follows:

$$\min_{u,c} V(x, u, c) = \frac{1}{2} \sum_{j=1}^J \sum_{l=1}^L u_{jl}^2 \|x_j - c_l\|^2$$

$$s. t. \sum_{l=1}^L u_{jl} = 1, u_{jl} \geq 0, j = 1, \dots, J, l = 1, \dots, L$$

Where  $u$  are called fuzzy classifier that indicates to assign the pixel  $j$  to the  $l$ -th class if  $u_{jl}$  is the biggest in  $\{u_{jk}, k = 1, \dots, L\}$ . In fact,  $u_{jl}$  could have any power ( $\geq 1$ ) besides twice, however, these are not in the scope of this paper. To solve it, the usual method is to alternate between solving a quadratic program in the variable  $u$  and solving another quadratic program in the variables  $c$ . The updates to  $u$  and  $c$  can be written as,

$$u_{jl}^{t+1} = [\sum_{m=1}^L \frac{\|x_j - c_m^t\|^2}{\|x_j - c_l^t\|^2}]^{-1}, c_l^{t+1} = \frac{\sum_{j=1}^J (u_{jl}^{t+1})^2 x_j}{\sum_{j=1}^J (u_{jl}^{t+1})^2} \tag{1}$$

For image segmentation, the input data are restricted to the grey-levels. Then the histogram-based FCM algorithm has been developed to save training time, and the updates can be reformulated [10].

$$u_{jl}^{t+1} = [\sum_{m=1}^L \frac{\|j - c_m^t\|^2}{\|j - c_l^t\|^2}]^{-1}, c_l^{t+1} = \frac{\sum_{j=1}^J (u_{jl}^{t+1})^2 H(j)j}{\sum_{j=1}^J (u_{jl}^{t+1})^2 H(j)} \tag{2}$$

Where  $H$  represents the histogram of image of  $J$  grey-levels.

### 2.3. Penalized estimator and update rule

Regardless of  $F_L$  and  $F_W$ , we impose the penalty  $V$  to them.

$$\min_{x,u,c} \Phi(x, u, c) = F(x) + \beta V(x, u, c) \tag{3}$$

$$\begin{aligned} & \text{s. t. } x_j \geq 0, j = 1, \dots, J \\ & \sum_{l=1}^L u_{jl} = 1, u_{jl} \geq 0, j = 1, \dots, J, l = 1, \dots, L \end{aligned}$$

Where  $F_L$  and  $F_W$  are collectively called  $F$ , and  $\beta > 0$  is a penalized parameter.

To solve the model, we alternate among solving  $x$ ,  $u$  and  $c$ . Fixed  $x$ , we can sequentially update  $u$  and  $c$  by Eq. (1). The next we focus on solving Eq. (3) on  $x$ .

It can be observed that both  $f_L$  and  $f_W$  are separable on  $x$ , i.e., to minimize two functions is equivalent to solve each  $x_j$ , respectively. The penalty  $V$  is naturally separable on  $x$ . Moreover,  $f_L + \beta V$  and  $f_W + \beta V$  are certainly the auxiliary functions of  $F_L + \beta V$  and  $F_W + \beta V$ . We take the partial derivatives of  $f_L$ ,  $f_W$  and  $V$  with respect to  $x_j$ , then we can solve the following equations in  $R^1$ .

$$\frac{\partial f_L(x, x^t)}{\partial x_j} + \beta \frac{\partial V(x, u, c)}{\partial x_j} = 0 \quad (4)$$

$$\frac{\partial f_W(x, x^t)}{\partial x_j} + \beta \frac{\partial V(x, u, c)}{\partial x_j} = 0 \quad (5)$$

For Eq. (4), denote  $a_j = \beta \sum_{l=1}^L u_{jl}^2$ ,  $b_j = 1 - \beta \sum_{l=1}^L u_{jl}^2 c_l$  and  $d_j = -x_j^t \sum_{i=1}^I W_{ij} y_i / (W x^t)_i$ , then the sole positive solution is,

$$x_j^{t+1} = \frac{-b_j + \sqrt{b_j^2 - 4a_j d_j}}{2a_j} \quad (6)$$

For Eq. (5), let  $a_j = (1/x_j^t) \sum_{i=1}^I W_{ij} (W x^t)_i / D_{ii} + \beta \sum_{l=1}^L u_{jl}^2$  and  $b_j = \sum_{i=1}^I W_{ij} y_i / D_{ii} + \beta \sum_{l=1}^L u_{jl}^2 c_l$ , and the solution below is also positive.

$$x_j^{t+1} = \frac{b_j}{a_j} \quad (7)$$

Thus, we obtain the updating algorithm for Eq. (3). Given penalized parameter and cluster number, start from the initial image  $x^0 (x^0 > 0)$ , the clustering centers  $c^0 (c_l^0 > 0)$ , and the fuzzy classifiers  $u^0$  satisfying the constraints of Eq. (3), where  $j = 1, \dots, J$  and  $l = 1, \dots, L$ .

- (1) update  $x$  by Eq. (6) (or Eq. (7));
- (2) update  $u$  and  $c$  by Eq. (1).

As can be seen, these update monotonically decrease the cost functions, so a feasible convergence can be obtained.

### 3. Experimental results

We had studied the proposed methods' behaviors, of which were compared to the conventional routine that segmentation after reconstruction. For convenience, the penalized ML and WLS

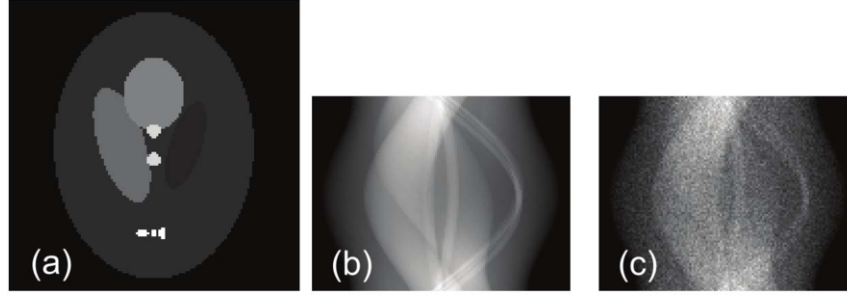


Fig. 1. (a) Shepp-Logan phantom, (b) noisy-free projections, and (c) noisy projections.

algorithms were called ML+SEG and WLS+SEG. Simulated phantom and real clinical data were both used to evaluate their performance. The simulated one was a Shepp-Logan phantom shown in Figure 1(a). The real PET data could be obtained from a transmission scan using mPower scanner of positron. Since total pixel values had effect on the selection of the penalized parameter, we rescaled them to  $10^6$  for two data sets. The system matrix was computed using the "angle of view" method in [1]. In the WLS estimator, the diagonal matrix  $D$  was calculated by Fessler's "data-plugin" technique [3]. The reconstructions in the conventional method were performed by both filtered backprojection (FBP) and MLEM, where the former used a hann filter with cut-off frequency 0.95 (normalized). The segmentation in the conventional method was performed by the histogram-based FCM with 500 bins. MLEM and two proposed algorithms were both initiated using the same initial image whose total counts equaled these of the projections. The experiments were performed on the PC of Lenovo with 3.00 GHz Pentium 4 (Dual Core) and 1.5 GB memory. The algorithms were implemented in MATLAB.

The difference between the reconstructed and original images could be demonstrated by the mean absolute error (MAE), where  $n$  was the total number of pixels of image.

$$MAE(t) = \frac{1}{n} \|x^t - x^{original}\|_1$$

### 3.1. Shepp-Logan phantom simulation

Tests were first realized on the Shepp-Logan phantom with  $128 \times 128$  grids and 3 mm pixels. We projected the phantom to obtain the projections with 128 radial bins (bin size 3 mm) and 96 angular views evenly spaced over  $\pi$  by the system matrix. Using Fessler's pseudo-random Poisson variants formulation, we can obtain the noisy projections.

$$y = c_i \text{Poisson}\{c_i^{-1}(y_i^* + a_i y_i^*)\} - c_i \text{Poisson}\{c_i^{-1} a_i y_i^*\}$$

Where  $y_i^*$  represented the observed projection,  $a_i = 20\%$  simulated random events, and  $c_i$  was the efficiency of  $i$ -th detector. We selected  $c_i = 1$  since the detector's efficiency was ignored. The noisy-free and noisy projections are shown in Figures 1(b) and 1(c).

In order to look for a suitable penalized parameter  $\beta$  for different cluster number  $L$ , we investigate the MAE changes of ML+SEG and WLS+SEG with both of them varying. Figure 2(a) shows these of ML+SEG after 200 iterations. We can see that the minimum is obtained at  $\beta = 10^{-3}$  and  $L = 3$ . Figure 2(b) shows these of the WLS+SEG, which indicates that the minimal MAE is obtained at the

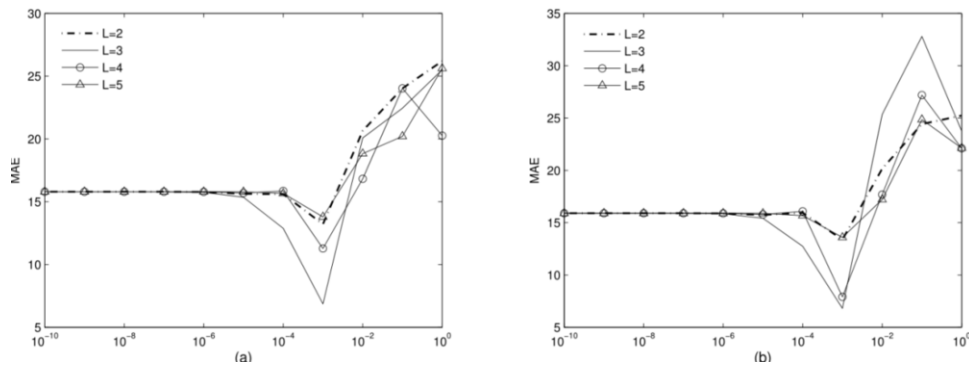


Fig. 2. MAE change with different  $\beta$  and  $L$ : (a) ML+SEG and (b) WLS+SEG.

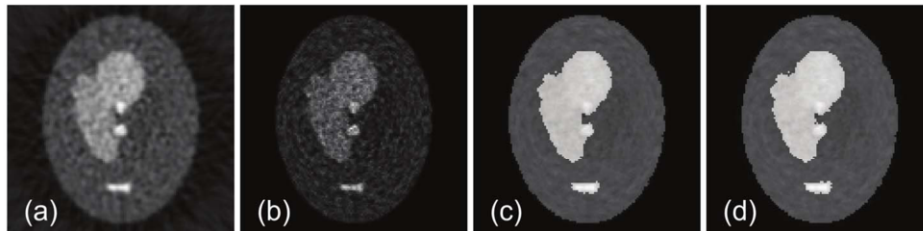


Fig. 3. Reconstructed images by (a) FBP, (b) MLEM, (c) ML+SEG, and (d) WLS+SEG.

same  $\beta$  and  $L$ .

Following Figure 2, we chose  $\beta = 10^{-3}$  and  $L = 3$  for two proposed algorithms. Figure 3 displays the reconstructions by FBP, MLEM, ML+SEG and WLS+SEG, where all iterative algorithms execute 100 iterations. We can see that the FBP and the MLEM provide rough images. But in order to make the image noise effectively suppressed, the proposed algorithms clearly segment the pixels into 3 classes and improve homogeneity within every region.

We also compare the conventional and proposed methods by a labeling technique that the pixel  $j$  is labeled  $l$  if it belongs to the  $l$ -th class. The images of the FBP and MLEM are segmented into 3 classes by the FCM that also executes 100 iterations. Figure 4 displays the labeling images. It can be seen that the conventional method provides the corrupted images with salt and pepper noise, which implies that they inaccurately extract the objects. The FBP produces slightly better result than the MLEM. Our algorithms, however, overcome these shortcomings and correctly recognize the homogeneous regions.

Figure 5 presents the comparison of MAE changes among the MLEM, ML+SEG and WLS+SEG with the selected  $\beta$  and  $L$ . As can be seen, two proposed algorithms decrease the MAE curves more

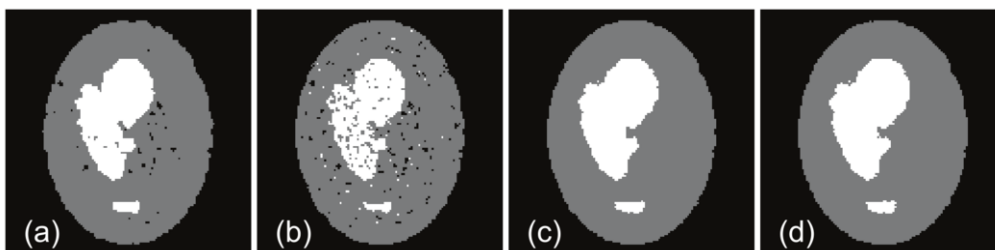


Fig. 4. Labeling images by (a) FBP and FCM, (b) MLEM and FCM, (c) ML+SEG, and (d) WLS+SEG.

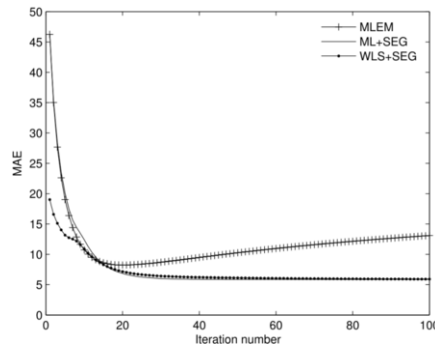


Fig. 5. MAE change as the iteration number's function.

rapidly than the MLEM, which shows our results much closer to the true phantom. It can also be seen that our algorithms deliver similar curves after 60 iterations making it hard to find out a difference each other from a point of view.

### 3.2. Real clinical data simulation

The real PET transmission data were obtained using a  $5\text{ uCi Ge}^{68}$  rotating rod source. The acquisition lasted for 10 minutes and 70 M counts in 61 slices were obtained. The normalization correction was performed by measurements obtained from a calibration scan. The original projections were 192 angles, 256 radial bins (bin size 2 mm). These angles, however, were not evenly spaced, for which (also to reduce computation complexity) we resampled them to 128 bins and 96 even angles. The reconstructed grid was  $128 \times 128$ , with 4 mm pixels. Figure 6 shows the resampled sinogram and the reconstructions by the FBP and MLEM. In practical applications, we need extract air, the lungs and soft tissue (for mPower scanner, it is generally difficult to distinguish the lungs from bed because their intensities are close), then we compute the mean of pixels for each class and assign them to pixels with respect to the classes. We away execute 100 iterations for the MLEM, ML+SEG, WLS+SEG and FCM.

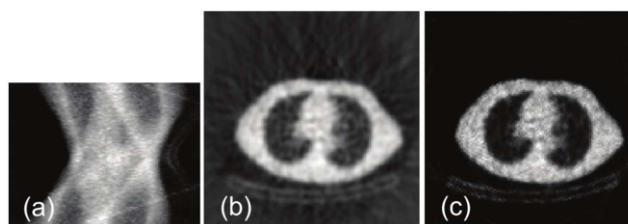


Fig. 6. (a) Real clinical projections, (b) reconstruction by FBP, and (c) reconstruction by MLEM.

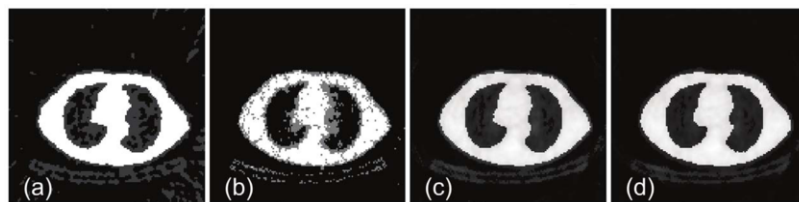


Fig. 7. 3-class segmented images by (a) FBP and FCM, (b) MLEM and FCM, (c) ML+SEG, and (d) WLS+SEG.

Considering 3-class segmentation with  $\beta = 10^{-3}$ , Figure 7 shows the segmented mean-assigned images and reconstructions by ML+SEG and WLS+SEG. For Figures 7(a) and 7(b), the results show much misidentification especially when identifying air for the FBP reconstruction and soft tissue for the MLEM reconstruction. However, our algorithms well bring out three parts that are less speckled and smoother for everyone. We can also see that the conventional method, which assigns the same value to the pixels belonging to a class, blurs out distinction even in the same organ. The proposed algorithms pursue both homogeneity and diversity.

We also study the overestimation of the actual regions, which optimizes 5 classes as proposed in [11]. Figure 8 shows the corresponding images with  $\beta = 10^{-3}$  (proposed by Figure 2). The conventional method repeats the same mistakes to recognize air for the FBP and soft tissue for the MLEM. Our algorithms can still overcome this weakness and reduce noise in the same organ. It should be noted that all the algorithms segment the lungs into several classes, which is mainly due to the large scope of intensities between air and soft tissue.

We plot the histograms of the images of FBP, MLEM, ML+SEG and WLS+SEG in Figure 9. This figure further verifies that ML+SEG and WLS+SEG provide clear classification with respect to the cluster number predetermined, and the reconstructed pixels in each class give consideration to the original distribution. We don't show the histograms for the segmented images because they are so simple to hold single value for each class.

The overall CPU times are listed in Table 1. It should be pointed out that our algorithms cost about 25 times CPU time than the conventional routine using the FBP reconstruction. It is not surprise to us that the analytic methods save much time than the iterative ones. However, new algorithms require comparable time costs with that of the MLEM. Actually, the updates to  $u$  and  $c$  in ML+SEG and WLS+SEG contribute very little (about 4%) to the total cost.

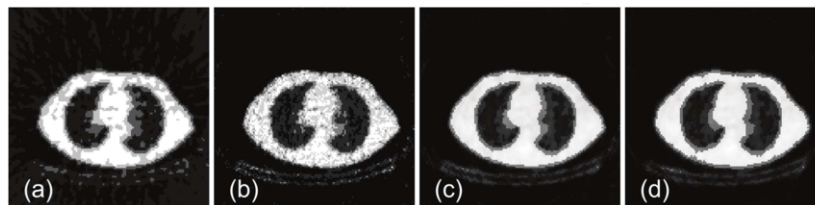


Fig. 8. 5-class segmented images by (a) FBP and FCM, (b) MLEM and FCM, (c) ML+SEG, and (d) WLS+SEG.

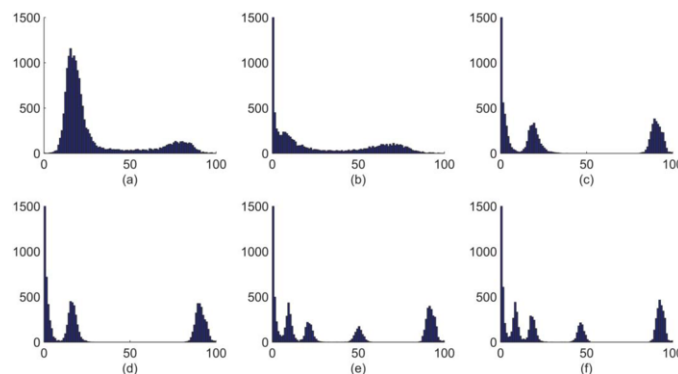


Fig. 9. Histograms of the reconstructions by (a) FBP, (b) MLEM, (c) ML+SEG with 3 classes, (d) WLS+SEG with 3 classes, (e) ML+SEG with 5 classes, and (f) WLS+SEG with 5 classes. It should be noted that the grey-levels are rescaled to 100 evenly bins according to their own maximum and minimum, and (b)-(f) are truncated above 1500 for a good visualization.



Table 1

Time costs in second, where: represents to sequentially execute two operations

$L$	FBP:FCM	MLEM:FCM	ML+SEG	WLS+SEG
3	2.30	51.81	50.38	45.09
5	2.58	52.46	54.12	47.50

#### 4. Discussion and conclusion

Two reconstruction models have been developed by modifying the objective functions with a penalty term that takes into account the influence of classification of pixels. A variety of data sets, including simulated and real data, are used to compare the performance of the conventional and new methods. The experiments demonstrate that the incorporation of the segmentation penalty into the reconstruction estimator can significantly improve the reconstructed result in presence of noise. We guess it due to the conventional method viewing reconstruction and segmentation in isolation, however, our methods combine two phases in a model and optimize them, simultaneously. In addition, the conventional method assigns the same value (calculated mean) to the pixels corresponding to classification, which put emphasis on the homogeneity and ignores the diversity even in the same organ. Our algorithms are capable of coping with both of them, and we can control the trade-off by adapting the penalized parameter. It is worth noting that our algorithms require no extra time cost compared to the conventional method with the MLEM reconstruction.

In recent years, MR-based attenuation correction is becoming the research hot points [12, 13], which our methods are also suitable for it. The future work is to address them to meet further clinical demands. In addition, there are many methods that use different regularizations, so we will research the combination of them with our method to obtain a better segmentation result in the future.

#### Acknowledgment

The work was supported by the National Natural Science Foundation of China (61302013, 61372014) and the Fundamental Research Funds for the Central Universities of China (N130319003).

#### References

- [1] L.A. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography, *IEEE Transactions on Medical Imaging* **1** (1982), 113-122.
- [2] C.L. Byrne, Iterative image reconstruction algorithms based on cross-entropy minimization, *IEEE Transactions on Medical Imaging* **2** (1993), 96-103.
- [3] J.A. Fessler, Penalized weighted least-squares image reconstruction for positron emission tomography, *IEEE Transactions on Medical Imaging* **13** (1994), 290-300.
- [4] J.G. Peter, Bayesian reconstructions from emission tomography data using a modified em algorithm, *IEEE Transactions on Medical Imaging* **9** (1990), 84-93.
- [5] J. Zhou, J.L. Coatrieux, A. Bousse, et al., A bayesian map-em algorithm for pet image reconstruction using wavelet transform, *IEEE Transactions on Nuclear Science* **54** (2007), 1660-1669.
- [6] J. Ma, Q. Feng, Y. Feng, et al., Generalized gibbs priors based positron emission tomography reconstruction, *Computers in Biology & Medicine* **40** (2010), 565-571.
- [7] Daniel D. Lee and H. Sebastian Seung, Algorithms for non-negative matrix factorization, *Proceedings of the 2000 Conference on Advances in Neural Information Processing Systems 13*, MIT Press, 2001, pp. 556-562.

- [8] A. Cichockia, H. Lee, Y.D. Kimb, et al., Non-negative matrix factorization with  $\alpha$ -divergence, *Pattern Recognition Letters* **29** (2008), 1433-1440.
- [9] C.M. Chen, H.H. Lu and Y.P. Hsu, Cross-reference maximum likelihood estimate reconstruction for positron emission tomography, *Biomedical Engineering Applications Basis & Communications* **13** (2002), 190-198.
- [10] A.O. Boudraa, S.M. Dehak, Y.M. Zhu, et al., Automated segmentation of multiple sclerosis lesions in multispectral MR imaging using fuzzy clustering, *Computers in Biology & Medicine* **30** (2000), 23-40.
- [11] H. Zaidi, J. Diaz-Gomez, A.O. Boudraa, et al., Fuzzy clustering-based segmented attenuation correction in whole-body pet imaging, *Physics in Medicine & Biology* **47** (2002), 1143-1160.
- [12] Y. Chen, M. Juttukonda M, Y. Su Y, et al., Probabilistic air segmentation and sparse regression estimated pseudo CT for PET/MR attenuation correction, *Radiology* **275** (2015), 562-569.
- [13] H. Kawaguchi, Y. Hirano, E. Yoshida, et al., A proposal for PET/MRI attenuation correction with mu-values measured using a fixed-position radiation source and MRI segmentation, *Nuclear Instruments & Methods in Physics Research B* **734** (2014), 156-161.