The Korean Communications
in Statistics Vol. 8, No. 1, 2001
pp. 299-309

Monotonic and Parallelizable Algorithm for Simultaneous Reconstruction of Activity/Attenuation using Emission Data in PET[1)]

Seung Gu Kim[2)]

Abstract

In PET(Positron Emission Tomography), it is necessary to use transmission scan data in order to estimate the attenuation map. Recently, there are several empirical studies in which one might be able to estimate attenuation map and activity distribution simultaneously with emissive sinogram alone without transmission scan. However, their algorithms are based on the model in which does not include the background counts term, and so is unrealistic. If the background counts component has been included in the model, their algorithm would introduce non-monotonic reconstruction algorithm which results in vain in practice. In this paper, we develop a monotonic and parallelizable algorithm for simultaneous reconstruction of both characteristics and present the validity through some simulations.

Keywords : EM algorithm, PET reconstruction, Activity/Attenuation, Monotonic algorithm, Background counts

1. Introduction

During about a decade, reconstruction problems of PET or SPECT(single photon emissive computed tomography) have become an important issue in applied statistics. The emissive medical imaging devices like PET or SPECT may give two different images of which one is activity distribution and the other is attenuation map. For reconstruction of the activity distribution like Natter(1993), Moore et. al.(1997), and Welsh(1997), because most of partial reconstruction algorithms come from the ML-EM(maximum likelihood expectation maximization) algorithm of Lange and Carson(1984)'s or Shepp and Vardi(1982)'s, it needs transmission scan before or after the end of emission scan in order to carry out attenuation

1) This work is supported by Sangji university research grant.
2) Associate Professor, department of Applied Statistics, SangJi university
E-mail : sgukim@chiak.sanggii.ac.kr

correction.

Recently, *simultaneous* reconstruction of activity distribution and attenuation map becomes to be required in PET because it does not use transmission scan and therefore the patient immobile time until scan is done is reduced. Among the several techniques, Nuyts et al(1999) are found to be the statistically reasonable approaches. However, their techniques are based on the model in which does not include the background counts component, and so is unrealistic because there does not exist PET device without bcakground events(eg. random coincidence, Compton scattering, crosstalks etc.). If the background counts component has been included in the model, their approaches would introduce non-monotonic algorithm which results in vain. The background counts that is must be occurred in PET although it is weaker than in SPECT. But it is well known that the background counts components are able to be estimated in accurate and is very important information to reconstruct the finer image.

Reconstruction algorithm should have several important features;

- the monotonicity of the convergence
- the global convergence
- the separability of the parameters
- the positivity of the estimates

In this paper, the goal is to provide a algorithm for simultaneous reconstruction activity and attenuation when the background counts component is included in the measurement model in PET, which satisfies above conditions. Next section presents a measurement model of PET and explains a non-concavity problem of an objective function, and in section 3, we provides a technique via the EM algorithm to avoid this problem with attention to the separability conditions. In section 4, a simulation is provided to illustrate the algorithm developed in this work to be monotonic and efficient in practical aspect.

2 Basic Notations and Problems

In PET, the observation counts at the ith projection follows

$$y_i \sim \text{Poisson}(\mu_i + \varepsilon_i), \quad i = 1, \cdots, N \tag{2.1}$$

, where

$$\mu_i = \mu_i(\alpha;\beta) = \sum_j s_{ij}\beta_i \exp\left(-\sum_j l_{ij}\alpha_j\right) \tag{2.2}$$

. And $\alpha = [\alpha_1, \cdots, \alpha_P]^T$ and $\beta = [\beta_1, \cdots, \beta_P]^T$ represent the (linear) attenuation map and the activity distribution respectively. The elements of the known system matrixes $\{s_{ij}\}$ and $\{l_{ij}\}$

indicate respectively the jth pixel's sensitivity and effective intersection length and ε_i represents a mean of the background counts component at the ith projection. And all of the symbols expressed in (2.1) and (2.2) are assumed to be nonnegative.

Our goal is to estimate α and β simultaneously using the observation sinogram $y=[y_1,\cdots,y_N]$. Because two system matrixes are ill-posed and N might be not much larger than P, the penalized log-likelihood

$$L_\lambda(\theta)=L(\theta)-\lambda J(\theta) \tag{2.3}$$

is maximized with respect to $\theta=[\alpha;\beta]$ instead naive log-likelihood L is maximized in order to regularize the estimation. Here the constant $\lambda>0$ and the penalty function J is assumed to be strictly convex in order to get global maximum for θ. In addition, we define

$$J(\theta)=J(\alpha,\beta)=J(\alpha)=\sum_j J_j(\alpha_j) \tag{2.4}$$

suppressing the activity β. This definition intends that the penalty shall influence on only attenuation α and that a bias of the estimates of β due to regularization shall be removed because estimates of activity distribution is a major diagnosis tool rather than attenuation map in medical. And the function J is assumed to be separated into each of pixels. In this work, we use a roughness penalty function

$$J(\alpha)=\sum_j(1/2)\sum_{k\in\partial_j}\xi_{ik}\tau(\alpha_j-\alpha_k^n) \tag{2.5}$$

, where ∂_j represents the set of the first order neighbors of j, ξ_{ik} is a value 1 for horizontal and vertical α_k^n's site and zero otherwise, and

$$\tau(x)=\delta^2\{|x/\delta|-\log(1+|x/\delta|)\} \tag{2.6}$$

and a constant $\delta>0$ that controls the resolution of reconstruct image preserving edge. Its properties is well explained in Fessler et. al.(1997).

Consider the log-likelihood ignoring its constant terms

$$\begin{aligned} L(\theta) &= \sum_i\{-(\mu_i(\theta)+\varepsilon_i)+y_i\log(\mu_i(\theta)+\varepsilon_i)\} \\ &= \sum_i h_i\Big(\sum_j s_{ij}\beta_j,\ \sum_j l_{ij}\alpha_j\Big) \end{aligned} \tag{2.7}$$

, where

$$h_i(m_1,m_2)=-\{m_1\exp(-m_2)+\varepsilon_i\}+y_i\log\{m_1\exp(-m_2)+\varepsilon_i\}. \tag{2.8}$$

From the second partial differentiates of the function h_i those are

$$\frac{\partial^2}{\partial m_1^2} h_i(m_1, m_2) = -y_i\left[\frac{\exp(-m_2)}{m_1 \exp(-m_2) + \varepsilon_i}\right]^2, \tag{2.9}$$

$$\frac{\partial^2}{\partial m_2^2} h_i(m_1, m_2) = -m_1 \exp(-m_2) + y_i\left[\frac{\varepsilon_i m_1 \exp(-m_2)}{(m_1 \exp(-m_2) + \varepsilon_i)^2}\right] \tag{2.10}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial m_1 \partial m_2} h_i(m_1, m_2) &= \frac{\partial^2}{\partial m_2 \partial m_1} h_i(m_1, m_2)\\ &= \exp(-m_2) - y_i\left[\frac{\varepsilon_i \exp(-m_2)}{(m_1 \exp(-m_2) + \varepsilon_i)^2}\right]\end{aligned}$$

It is important to note that the two-valued function h_i is not concave for the argument (m_1, m_2), which implies that h_i is not concave for (α, β) either since m_1 and m_2 are linear with α and β respectively. At hence the log-likelihood function L and the penalized log-likelihood function L_λ is not concave for $\theta = [\alpha; \beta]$ in turn. From (2.9) and (2.10), we find that if ε_i is zero, h_i is *partially* concave in the sense that it is concave for one argument when the other is given. However, note that h_i is partially concave for m_1 but not for m_2 unless ε_i is zero, so that

$$\beta^{new} = \arg\max_\beta L_\lambda(\alpha^n, \beta | \alpha^n, \beta^n) = \arg\max_\beta L(\alpha^n, \beta | \alpha^n, \beta^n) \tag{2.11}$$

increase the L_λ but

$$\alpha^{new} = \arg\max_\alpha L_\lambda(\alpha, \beta^n | \alpha^n, \beta^n) = \arg\max_\alpha \{L(\alpha, \beta^n | \alpha^n, \beta^n) - \lambda J(\alpha)\} \tag{2.12}$$

not guarantee to increase it monotonically. As a results, this partial update strategy dose not provide the monotonicity in convergence.

3. Techniques for Surrogate of L_λ

We have to achieve two tasks; the first is to provide a technique to estimate $\{\beta_j\}$ separately. There is no need to find a monotonic algorithm for β_js because the β^{n+1} in (2.8) itself increase the L_λ in monotonic, the second is to estimate $\{\alpha_j\}$ not only be monotonic but also be separated.

3.1 Separability Technique for Activity Estimation

Postulate $y_i = \sum_j x_{ij} + e_i$, where x_{ij} and e_i are independent and follows

$$x_{ij} \sim \text{Poisson}(\mu_{ij}) \text{ and } e_i \sim \text{Poisson}(\varepsilon_i) \tag{3.1}$$

respectively indicating $(x_{i1}, \cdots, x_{iP}, e_i)$ the complete data and y_i the incomplete data and $\mu_{ij} = s_{ij}\beta_j \exp\left(-\sum_k l_{ik}\alpha_k\right)$.

Then ignoring constant terms

$$\begin{aligned} Q(\alpha^n, \beta|\alpha^n, \beta^n) &= E[\log f(x, e; \alpha^n, \beta)|\alpha^n, \beta^n, y] \\ &= E[\log f(x; \alpha^n, \beta)|\alpha^n, \beta^n, y] E[\log f(e)|\alpha^n, \beta^n, y] \\ &\propto E[\log f(x; \alpha^n, \beta)|\alpha^n, \beta^n, y] \\ &= \sum_i \sum_j \left\{-s_{ij}\beta_j \exp\left(-\sum_k l_{ik}\alpha_k^n\right) + x_{ij}^n \log(-s_{ij}\beta_j)\right\} \end{aligned} \tag{3.2}$$

, where

$$x_{ij}^n = E[x_{ij}|y, \alpha^n, \beta^n] = y_i\left(\frac{\mu_{ij}^n}{\mu_i^n + \varepsilon_i}\right). \tag{3.3}$$

and $\mu_i^n = \mu_i(\alpha^n, \beta^n)$. From $Q(\alpha^n, \beta|\alpha^n, \beta^n) = 0$ the update algorithm is given as

$$\beta_j^{n+1} = \frac{\sum_i x_{ij}^n}{\sum_i s_{ij} \exp\left(-\sum_k l_{ik}\alpha_k^n\right)} = \beta_j^n \sum_i \mu_{ij}^n \left(\frac{y_i}{\mu_i^n + \varepsilon_i}\right) \Big/ \sum_i \mu_{ij}^n \tag{3.4}$$

$$, \quad j = 1, \cdots, P$$

, which increase monotonically Q_λ, and so the penalized log-likelihood L_λ.

Note that if $\varepsilon_i = 0$, then the algorithm (3.4) becomes Lange and Carson(1984)'s ML-EM algorithm when the activity distribution is solely(not partially) estimated.

3.2 Monotonicity and Separability Technique for Attenuation Estimation

It has becomes to be conventional for many authors(e.g Fessler et. al.,1997) to use the De Pierro(1995)'s convexity technique to provide a separable algorithm to estimate attenuation map α. It begins from setting

$$\sum_j l_{ij}\alpha_j = \sum_j w_{ij}\left[\frac{l_{ij}}{w_{ij}}(\alpha_j - \alpha_j^n) + \sum_k l_{ik}\alpha_k^n\right]$$

for the weights $0 < w_{ij} < 1$ and $\sum_j w_{ij} = 1$, and then the Jensen's inequality says that for a concave function f it holds

$$\begin{aligned} f\left(\sum_j l_{ij}\alpha_j\right) &= f\left(\sum_j w_{ij}\left[\frac{l_{ij}}{w_{ij}}(\alpha_j - \alpha_j^n) + \sum_k l_{ik}\alpha_k^n\right]\right) \\ &\geq \sum_j w_{ij} f\left(\frac{l_{ij}}{w_{ij}}(\alpha_j - \alpha_j^n) + \sum_k l_{ik}\alpha_k^n\right). \end{aligned} \tag{3.5}$$

For separability, one cannot directly apply the inequality (3.5) to the log likelihood L because the function h_i is not concave as long as ε_i terms is present as it has been mentioned in previous section. So that, not only for the separable update but for the monotonic convergence, we need a surrogate function of the L(so the L_λ) that is always concave regardless of ε_i.

To avoid the nonconcavity problems due to ε_i term in h_i, one can use the paraboloidal technique(Erdogan and Fessler, 1999 and Sotthivirat and Fessler, 2000). But in this paper, we choose the Kim(1999)'s the EM technique because the latter is simpler and more statistically reasonable than the former in nature. While the paraboloidal surrogate use a quadratic approximation of log likelihood based on numerically analytic approach, the Kim's technique provide a surrogate of log likelihood using EM property.

Consider

$$\alpha^{new} = \arg\max_\alpha Q_\lambda(\alpha, \beta^n | \alpha^n, \beta^n) = \arg\max_\alpha\{Q(\alpha, \beta^n | \alpha^n, \beta^n) - \lambda J(\alpha)\} \tag{3.6}$$

, then we know that α^{new} should increase the eq. (2.3) monotonically by the property of EM.

Postulate $y_i = x_i + e_i$, where x_i and e_i are independent and follows

$$x_i \sim \text{Poisson}(\mu_i) \text{ and } e_i \sim \text{Poisson}(\varepsilon_i) \tag{3.5}$$

respectively. Here (x_i, e_i) are regarded as the complete data and y_i to the incomplete data, then the evaluation function for the complete data given y, in similar to (3.2), is

$$Q(\alpha, \beta^n | \alpha^n, \beta^n) = E[\log f(x; \alpha, \beta^n) | y, \alpha^n, \beta^n] = \sum_i g_i\left(\sum_j l_{ij}\alpha_j\right) \tag{3.7}$$

, where

$$g_i(m) = -\sum_k s_{ik}\beta_k^n \exp(-m) - x_i^n m \tag{3.8}$$

ignoring constant terms and

$$x_i^n = E[x_i | y, \alpha^n, \beta^n] = y_i\left(\frac{\mu_i^n}{\mu_i^n + \varepsilon_i} \right) . \tag{3.9}$$

It is necessary to note that the function g_i is concave regardless whether ε_i is zero or not, and Q and Q_λ are also concave in turn. Hence, the α^{new} of (3.6) will monotonically increase Q_λ and so L_λ.

Now, applying De Pierro's technique to (3.7),

$$\begin{aligned} Q(\alpha, \beta^n | \alpha^n, \beta^n) &= \sum_i g_i\left(\sum_j l_{ij}\alpha_j \right) \\ &\geq \sum_j \sum_i w_{ij} g_i\left(\frac{l_{ij}}{w_{ij}} (\alpha_j - \alpha_j^n) + \sum_k l_{ik}\alpha_k^n \right) \\ &= \sum_j \phi_j(\alpha_j | \alpha^n) \equiv \phi(\alpha | \alpha^n) \end{aligned} \tag{3.10}$$

ignoring constant terms, where

$$\begin{aligned} \phi_j(\alpha_j | \alpha^n) = -\sum_i w_{ij}\left(\sum_k s_{ik}\beta_k^n \right) \exp\left(-\frac{l_{ij}}{w_{ij}} (\alpha_j - \alpha_j^n) - \sum_k l_{ik}\alpha_k^n \right) \\ -(\alpha_j - \alpha_j^n) \sum_i y_i l_{ij}\left(\frac{\mu_i^n}{\mu_i^n + \varepsilon_i} \right) . \end{aligned} \tag{3.11}$$

So far in this section, we find the monotonic and separable function $\phi(\alpha|\alpha^n)$ that is surrogate the L, which implies that

$$\alpha_j^{n+1} = \arg\max_\alpha \{ \phi(\alpha|\alpha^n) - \lambda J(\alpha) \} = \arg\max_\alpha \{ \phi_j(\alpha_j|\alpha^n) - \lambda J_j(\alpha_j) \} \quad , \quad j = 1, \cdots, P \tag{3.12}$$

increase L_λ in monotonic. Therefore, using one-step Newton-Raphson method, we provide a gradient algorithm

$$\alpha_j^{n+1} = \left[\alpha_j^n + \frac{\dot{\phi}_j(\alpha_j^n) - \lambda \dot{J}_j(\alpha_j^n)}{-\{ \ddot{\phi}_j(\alpha_j^n) - \lambda \ddot{J}_j(\alpha_j^n) \}} \right]_+ , \quad j = 1, \cdots, P \tag{3.13}$$

, where one dot and double dot indicate the first and second partial derivatives at $\alpha_j = \alpha_j^n$ respectively, so that

$$\dot{\phi}_j(a_j^n) = \sum_i \left(1 - \frac{y_i}{\mu_i^n + \varepsilon_i}\right) l_{ij}\mu_i^n \quad , \tag{3.14}$$

$$\ddot{\phi}_j(a_j^n) = -\sum_i l_{ij}^2 \mu_i^n / w_{ij} = -\sum_i \sum_k l_{ik}^2 \mu_i^n \tag{3.15}$$

in case that we choose $w_{ij} = l_{ij}^2 / \sum_k l_{ik}^2$ and

$$\dot{J}_j(a_j^n) = \sum_{k \in \partial_j} (a_j^n - a_k^n) \Big/ \left(1 + \left| \frac{a_j^n - a_k^n}{\delta} \right| \right) , \tag{3.16}$$

$$\ddot{J}_j(a_j^n) = \sum_{k \in \partial_j} \left(1 + \left| \frac{a_j^n - a_k^n}{\delta} \right| \right)^{-1} \tag{3.17}$$

and the operator $[\,]_+$ means for any real valued function d

$$[d(x)]_+ = \begin{cases} x \ , & d(x) < 0 \\ d(x) \ , & d(x) \geq 0 \end{cases} \tag{3.18}$$

Note that it guarantees to satisfy the positivity condition whenever one put the initial estimates to be non-zero, and that it never break the monotonicity condition since it choose $a_j^{n+1} = a_j^n$ if $d(a_j^n)$ is negative, hence it would not decrease the L_λ.

4. Simulation Study

In this section, we investigate whether a couple of algorithm (3.4) and (3.14) is monotonically converge to some reconstruction image. and it provides actually the locational information of disease positions presumed in real image. An experiment use a coupe of artificial images. 64×64 attenuation map image, A. in Figure 1, simulates a human thorax in which the body, the heart, the lung, the spine and the tubes with the attenuation coefficients 0.030, 0.050, 0.050, 0.120 and 0.070 respectively are included. And

A. True Attenuation map *B. True Activity distribution*

Figure 1. *64×64 Simulated Images that are inverted, i.e., from white to black and black to white in grey intensity.*

64×64 activity distribution image, B. in Figure 1, simulates the concentration of radiopharmaceutical on its own positions of disease(a larger tumor in the left lung and a very small tumor in the right lung) and skin sources from which positron radiates to a pair of outer detectors.

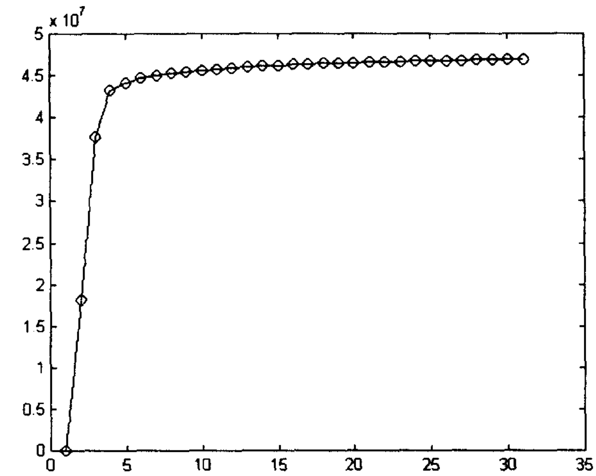

Figure 2. *Plot of* $L_\lambda(\theta^n) - L_\lambda(\theta^0)$ *with 30 iterations illustrating the sequence of estimated image is quickly and monotonically converged to a reconstructed image.*

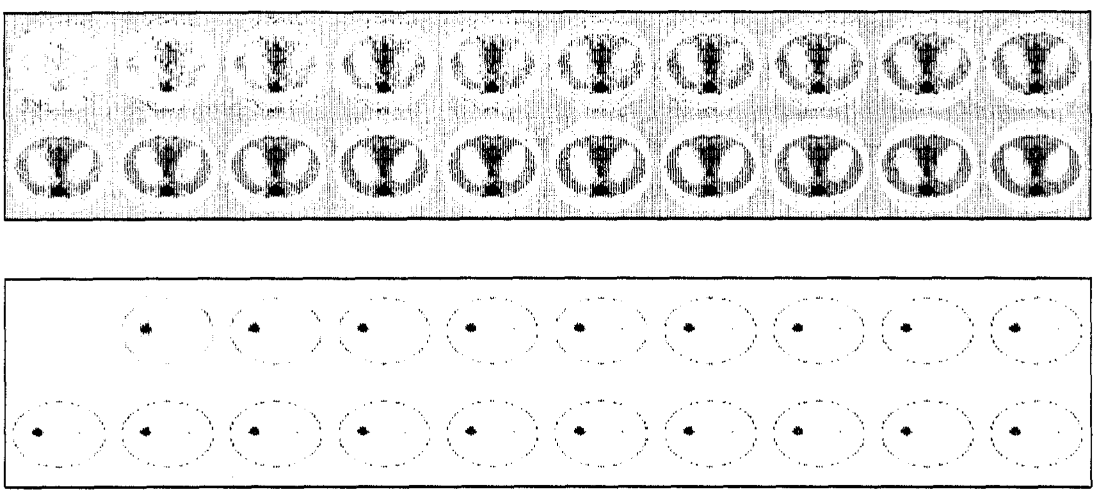

Figure 3. *20 iterates of algorithm developed in this paper. the Upper : Attenuation Maps and the Lower : Activity Distributions*

It is assumed that photon counts are emitted by 1000 per a white pixel in image B. And the initial estimates of α and β are used with the filtered back projection(FBP) image based on a smoothed sinogram and with an uniform-valued image respectively. Observed photon counts y_is are generated independently from Poisson distribution with mean $\mu_i + \varepsilon_i$ where ε_i is also randomly generated for uniform distribution on [10,15]. And we choose $\delta = 0.05$ in (3.16) and (3.17) and smoothing constant $\lambda = 6600$ in (3.13).

Figure 2 confirms again that a couple of algorithm (3.4) and (3.14) developed in this paper increase the penalized log-likelihood in monotonic and it illustrates convergence to a reconstructed image is very quickly in iterations (and it takes less than 30 seconds until 20 iterations on *Pentium III/600MHz*).

It is very interesting to note that, in Figure 3, activity distribution converges faster than attenuation does; we are able to recognize apparently it as actual activity distribution within 5 iterations while attenuation map is not apparent at even 20 iterations, which is saying that to reconstruct activity distribution does not require excessively accurate information for attenuation map in PET!

5. Conclusion and Further Study

This paper provides a simultaneous reconstruction algorithm for activity and attenuation which guarantees to converge in monotonic when background counts terms is included in measurement model of PET, and attempt to avoid non-monotonicity caused by background counts using familiar EM surrogate but not paraboloidal surrogate of the Erdogan and Fessler(1999)'s. Our method is also perform well in a simulation study and easier, simpler and more reasonable in statistical aspects than paraboloidal technique.

This paper develop a simultaneous reconstruction algorithm in PET in which the measurement model might be quite simpler than in SPECT. In (2.2), both projection of activity and attenuation are separated independently, which makes the EM construction be easier. In SPECT, however, every attenuation projection is associated by activity projection therefore both are independent which is very complicate to construct EM surrogate. Our next work will be to provide the simultaneous algorithm in SPECT via EM surrogate technique like this paper did.

References

[1] De Pierro, A. R.(1995), A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Trans. Medical Imaging*, vol 14, No. 2, 532-541.

[2] Erdogan, H. and J. A. Fessler(1999), Monotonic Algorithms for Transmission Tomography, *IEEE trans. Medical Imaging*, vol. 18, No. 9, 801-814.

[3] Fessler, J. A. E., P. Ficaro(1997), N. H. Clinthorne and K. Lange, Grouped Coordinate ascent algorithms for penalized-likelihood transmission image reconstruction, *IEEE Trans., Medical Imaging.*, vol. 16, 166-175.

[4] Kim, S. G(1999), GCA Reconstruction Algorithm within the EM model for Transmission Computed Tomography, *The Korean Journal of Applied Statistics*, vol. 12, No. 2, 537-551.

[5] Lange, K., and R. Carson(1984), EM reconstruction algorithms for emission and transmission tomography, *Journal of Computer Assistant Tomography.*, vol. 8, 306-316.

[6] Moore, S. C., M. F. Kijewski and S. P. Mueller(1997), A general approach to nonuniform attenuation correction using emission data alone, *Journal of. Nuclear Medicine.*, vol 41, 1777-1807.

[7] Natterer, F.(1993), Determination of tissue attenuation in emission tomography of optically dense media, *Inverse Problems*, vol. 9, 731-736.

[8] Nuyts, J., P. Dupont, S. Stroobants, R Benninck, L Mortelmans and P. Suentens(1999) Simultaneous Maximum A Posteriori Reconstruction of Attenuation and Activity Distribution form Emission Sinograms, *IEEE Trans. Medical Imaging*, vol 18, 393-403.

[10] Shepp, L., A. and Y. Vardi(1982), Maximum Likelihood Reconstruction for Emission Tomography, *IEEE Trans. Medical Imaging*, vol 1, 113-122.

[11] Sotthivirat, S and J. A. Fessler(2000), Partitioned Saparable Paraboloidal Surrogate Coordinate Ascent Algorithm for Image Restoration, *Proceeding IEEE International Conference on Image Processing*, Vol. 1, 109-112.

[12] Welch, A., R. Clack, F. Natterer and G. T. Gullberg(1997), Toward accurate attenuation correction in SPECT without transmission measurements, *IEEE Trans. Medical Imaging*, vol 16, 532-541.