JOURNAL OF KOREA MULTIMEDIA SOCIETY, VOL. 8, NO. 12, DECEMBER 2005(pp. 1565-1571)

Bayesian Image Reconstruction Using Edge Detecting Process for PET

Jong Seok Um†

ABSTRACT

Images reconstructed with Maximum-Likelihood Expectation-Maximization (MLEM) algorithm have been observed to have checkerboard effects and have noise artifacts near edges as iterations proceed. To compensate this ill-posed nature, numerous penalized maximum-likelihood methods have been proposed. We suggest a simple algorithm of applying edge detecting process to the MLEM and Bayesian Expectation-Maximization (BEM) to reduce the noise artifacts near edges and remove checkerboard effects. We have shown by simulation that this algorithm removes checkerboard effects and improves the clarity of the reconstructed image and has good properties based on root mean square error (RMS).

Keywords: PET, Image Reconstruction, Bayesian, Edge Detecting Process

1. INTRODUCTION

Statistical positron emission tomography (PET) image reconstruction methods produce improved spatial resolution and variance property over conventional filtered back-projection methods through accurate physical and statistical modeling. In 1982, Shepp and Vardi proposed a Poisson model for PET, known as MLEM algorithm [1]. Even though MLEM is a theoretical based approach, images reconstructed with MLEM algorithm have been observed to have checkerboard effects and have large distortions near edges as iterations proceed. To compensate this ill-posed nature, numerous penalized maximum-likelihood methods have been proposed which are known as BEM. Green proposed a MAP (Maximum a Posteriori) algorithm which has non-negative image result and caused convergence issues [2]. A modified MAP proposed by Lang was computationally expensive [3]. A fast MAP method based on ordered subsets (OS) was proposed by De Pierro and Yamagishi [4]. The OS idea is to use only one subset of the measurement data for each update instead of the total data. A class of OS algorithm has shown significantly accelerated convergence. Ahn and Fessler proposed a relaxed OS algorithm based on the OS-EM algorithm [5]. However in those OS based methods, there are some uncertainties as to how the subsets are to be chosen.

MLEM algorithm have been observed to become noisy and have noise artifacts near edges as iterations proceed. Introducing a priori probability, BEM prevents the occurrence of checkerboard effects in the image. However, images obtained from BEM have been blurred near edges since BEM uses all the neighborhood pixels. In this paper, we suggest a simple algorithm of applying edge detecting process to the MLEM and BEM to reduce the noise near edges. Once a neighbor pixel is confirmed as the edge element through edge detecting process, this pixel is not considered from reconstructing model. We use clique to remove this

※ Corresponding Author : Jong Seok Um, Address : (136-792) 389-2 Ga, Samsun-Dong, Sungbuk-Gu, Seoul, Korea, TEL: +82-2-760-4133, FAX: +82-2-760-4488
E-mail: jsum@hansung.ac.kr
Receipt date : Apr. 4, 2005, Approval date : Jun. 22, 2005
† Dept. of Multimedia Engineering, Hansung University
※ This research was financially supported by Hansung University in the year of 2004.

pixel from reconstructing model. We have shown by simulation that this algorithm improves the reconstructed image and reduces MSE compare to MLEM and BEM.

Section 2 describes the Poisson model of the PET and reconstructing algorithms, MLEM and BEM. In Section 3, we propose LEM and LBEM to which the edge detecting process is applied to MLEM and BEM respectively.

2. MODELING FOR THE PET

2.1 Likelihood for the emission counts

The data for PET consist of a vector of counts detected in the d-th detector tube $n^*(d)$, $d=1,2,\cdots,D$, where D is the total number of detector tubes. Let $\lambda(b)$ be the emission intensity at a point b in the image, $b=1,2,\cdots,B$, where B is the number of boxes in the image. Usually we choose a box as a pixel in the image. Assuming $n^*(d)$ as Poisson distribution with parameter $\lambda^*(d)$, optimization criteria is to maximize likelihood function. Let $n(b,d)$ be the random variable denoting the number of radioactive emissions occurring in the pixel b and detected at the tube d. Regarding $n(b,d)$ as unobserved data and $n^*(d)$ as observed data, EM algorithm is applicable to find $\lambda(b)$, $b=1,2,\cdots,B$, which maximize likelihood function $p(\vec{n^*}|\vec{\lambda})$. Let $p(b,d)$ be the conditional probability of detecting an emission in tube d emitted from pixel b. Then $n^*(d)$, the total number of emissions detected in tube d, has Poisson distribution with mean $\lambda^*(d)=\sum_{b=1}^{B}\lambda(b)p(b,d)$. Here $p(b,d)$'s are assumed the known non-negative constants. It depends on various factors: the geometry of the detection system, the activity of the isotope and exposure time, and the extent of attenuation and scattering between sources and detectors. Here we estimate $p(b,d)$ as the angle of view from the center of the pixel b into tube d. Log of likelihood function of the Poisson model is as follow

$$\ln L(\vec{\lambda})=\sum_{d=1}^{D}\{-\sum_{b=1}^{B}\lambda(b)p(b,d)+n^*(d)\ln\left(\sum_{b=1}^{B}\lambda(b)p(b,d)\right)-n^*(d)!\} \quad (1)$$

Here $\vec{\lambda}^T=(\lambda(1),\lambda(2),\cdots,\lambda(B))$. Applying EM algorithm to equation (1), we have the following iterative formula.

$$\lambda^{new}(b)=\lambda^{old}(b)\sum_{d=1}^{D}\left(\frac{n^*(d)p(b,d)}{\sum_{b'=1}^{B}\lambda^{old}(b')p(b',d)}\right) \quad (2)$$

Start with initial estimate of $\vec{\lambda}^{old}$, satisfying $\lambda(b)>0$, $b=1,2,\cdots,B$. If $\vec{\lambda}^{old}$ denotes the current estimate of $\vec{\lambda}$, define a new estimate $\vec{\lambda}^{new}$ by equation (2). If the required accuracy for numerical convergence has been achieved, then stop. Otherwise, continue to update $\vec{\lambda}$ using equation (2).

2.2 priors for the emission intensity

Images reconstructed by using MLEM algorithm have been observed to become noisy and to have checkerboard effects. Also, these have noise artifacts near edges as the iterations proceed . Levitan and Herman use a Gaussian prior, called the penalty function, to prevent the occurrence of checkerboard effects [6]. Since the posterior is sensitive only to the local properties of priors, we use Markov Random Field (MRF) as prior. It is known that Gibbs distribution is MRF, which has following form

$$p(\vec{\lambda}) \propto \exp\left(-\sum_{c\in C}V_c(\vec{\lambda})\right), \quad (3)$$

where c is a clique and C is the collection of all cliques. A local set of point c is clique if (b,b') are neighbors for $\forall(b,b')\in c$. Shepp and Vardi suggest $V_c(\vec{\lambda})$ as follow [1];

$$V_c(\vec{\lambda}) = \beta \sum_{\{b,b'\}\in c} w_{bb'} \log\cosh\left(\frac{\lambda(b)-\lambda(b')}{\delta}\right), \quad (4)$$

where $w_{bb'}$ is a weight and $w_{bb'}=1$ if (b, b') are orthogonal neighbors, $\sqrt{1/2}$ if (b, b') are diagonal neighbors and 0 otherwise. Bouman and Sauer propose edge-preserving prior, called generalized Gaussian Markov random field (GGMRF) [7], which allows the realistic edge modeling having form

$$np(\vec{\lambda}) = -\beta^k\left\{\sum_b a_b\lambda(b)^k + \sum_{c\in C}\sum_{\{b,b'\}\in c} w_{bb'}|\lambda(b)-\lambda(b')|^k\right\} + constant \quad (5)$$

where $1\le k\le 2$. GGMRF includes Gaussian MRF when $k=2$. When $k=1$, it is similar to the median pixel prior. By choosing β and k appropriately, we can avoid negative value of $\lambda^{new}(b)$.

2.3 Bayesian EM (BEM)

With GGMRF prior, we have the posterior as follow

$$\ln p(\vec{\lambda}|\vec{n^*}) = \sum_{d=1}^{D} n^*(d)\ln\left(\sum_b \lambda(b)p(b,d)\right) - \sum_d\sum_b \lambda(b)p(b,d) - V(\vec{\lambda}) + constant \quad (6)$$

wher. $V(\vec{\lambda}) = \beta^k\left\{\sum_b a_s\lambda(b)^k + \sum_{c\in C}\sum_{\{b,b'\}\in c} w_{bb'}|\lambda(b)-\lambda(b')|^k\right\}$

Here $\vec{n^*} = (n^*(1), n^*(2), \cdots, n^*(D))^T$, is the detected X-ray counts at the detector tube.

Let $z(b, d)$ be an estimate of unobserved data $n(b, d)$ given observed data $n^*(d)$.

Then $z(b, d) = E(n(b, d)|n*(d)) =$

$$n^*(d)\frac{\lambda^{old}(b)p(b,d)}{\sum_{b=1}^{B}\lambda^{old}(b')p(b',d)} \quad (7)$$

Using this result, we have the following result at the expectation step

$$E(\ln p(\vec{\lambda}|\vec{n}, \vec{n^*})|\vec{n^*}, \vec{\lambda}^{old}) = \sum_d\sum_b z(b,d)\ln\lambda(b)p(b,d) - \sum_d\sum_b \lambda(b)p(b,d) - V(\vec{\lambda}) \quad (8)$$

Here $\vec{n} = (n(1,1), n(1,2), \cdots, n(1,D), n(2,1), \cdots, n(B,D))^T$ is the random variable denoting the number of radioactive emissions occurring in the pixel $b=1,2,\cdots,B$ and detected at the tube $d=1,2,\cdots,D$ and $\vec{\lambda}^{old} = (\lambda^{old}(1), \lambda^{old}(2), \cdots, \lambda^{old}(B))^T$ is current estimates of $\lambda(b)$, $b=1,2,\cdots,B$. At the maximization step, find a estimate of $\vec{\lambda}$ which maximize above equation and go back to expectation step putting the estimate as $\vec{\lambda}^{old}$. Instead of direct maximization, we use one-step-late (OSL) approximation proposed by [1]. We have the following iterative equation

$$\lambda^{new}(b) = \frac{\sum_{d=1}^{D} z(b,d)}{\sum_{d=1}^{D} p(b,d) + \frac{\partial V(\vec{\lambda})}{\partial\lambda(b)}\Big|_{\lambda=\lambda^{old}}}. \quad (9)$$

In equation (9), if $\lambda(b)$ has a higher intensity compare to neighbors, $\frac{\partial V(\vec{\lambda})}{\partial\lambda(b)}\Big|_{\lambda=\lambda^{old}}$ becomes positive and $\lambda^{new}(b)$ becomes smaller. If $\lambda(b)$ has a lower intensity compare to neighbors, $\frac{\partial V(\vec{\lambda})}{\partial\lambda(b)}\Big|_{\lambda=\lambda^{old}}$ becomes negative and $\lambda^{new}(b)$ becomes larger. Thus BEM has smoothing effects. Shepp and Vardi use Bayesian reconstruction for PET and there are noise artifacts in the reconstructed image near edges [1].

3. BAYESIAN EM WITH EDGE-DETECTING PROCESS (LBEM)

The objective of the edge detecting process is to detect the presence and identify the location of intensity changes in an image and then to eliminate the pixels having edge element from neighborhood. Let $l(b, b')$ = 1 if there exists an edge element between pixel b and b', 0 if not. We use the hierarchical model for the priors. Then edge-preserving priors with edge detecting process has the following form

$$p(\vec{\lambda}, \vec{l}) \propto \exp\{-V_1(\vec{\lambda}|\vec{l}) - V_2(\vec{l})\}, \quad (10)$$

Here $\vec{l} = (l(1,2), l(1,3), \cdots, l(1,B), l(2,3), \cdots, l(B-1,B))^T$

denotes edge elements between pixel b and b', where $b \neq b'$. $p(\vec{\lambda}, \vec{l})$ is a hierarchical prior for $\vec{\lambda}$ and $\vec{l}$. Conditional prior of $\vec{\lambda}$ given $\vec{l}$ is given in equation (11).

$$V_1(\vec{\lambda}|\vec{l}) = \beta^k\Big(\sum_b a_b \lambda(b)^k + \sum_{c \in C} \sum_{(b,b') \in c} (1 - l(b,b'))w_{bb'}|\lambda(b) - \lambda(b')|^k\Big) \quad (11)$$

Geman and Geman use clique of size four to construct $V_2(\vec{l})$ [8]. Applying EM algorithm with the prior in equation (10) and Poisson likelihood function in equation (1), we have the following iterative formula.

$$\lambda^{new}(b) = \frac{\sum_{d=1}^{D} z(b,d)}{\sum_{d=1}^{D} p(b,d) + \frac{\partial V_1(\vec{\lambda}|\vec{l})}{\partial \lambda(b)}\Big|_{\lambda = \lambda^{old}, l = l^{old}}}, \quad (12)$$

$l(b,b')^{new} = 1$ or 0 which maximize

$$p(\vec{\lambda}, \vec{l}) L(\vec{\lambda})|_{\lambda = \lambda^{old}} \quad (13)$$

Iterative formular for LBEM in equation (12) includes edge elements. If $\lambda(b)$ has a higher intensity because of edge relation with some neighboring pixels, then $\frac{\partial V_1(\vec{\lambda}|\vec{l})}{\partial \lambda(b)}\Big|_{\lambda = \lambda^{old}, l = l^{old}}$ becomes small and $\lambda^{new}(b)$ changes a little. Smoothing effect of BEM is removed in LBEM. Edge elements are determined to maximize the posterior probability $p(\vec{\lambda}, \vec{l}) L(\vec{\lambda})|_{\lambda = \lambda^{old}}$ as given in equation (13).

Here we propose to determine $l(b,b')$ using empirical distribution of the difference of the pixel intensities of b and b'. Here is the procedure for the edge detecting process. We call it as an empirical edge detecting process. First, we consider the clique of order 2 with size four and four types of pair interactions: horizontal, vertical, right-skewed diagonal and left-skewed diagonal. Let F_i be the empirical distribution of $d_i(b,b')$ with $(b,b') \in c$ where $d_i(b,b')$ is the difference of the two pixel intensities of i type pair interaction, $i = 1,2,3,4$. Second, determine the p_i which is the proportion of pixels having i type edge relations in the image. The p_i is determined by examining the empirical distribution of $d_i(b,b')$. If there is the first point from which the right side is flat, then p_i is the right-tail probability from this point. Third, find $(1-p_i)^{th}$ percentile point of F_i, $F_i^{-1}(1-p_i)$, and determine $l(b,b') = 1$ if $d_i(b,b') \geq F_i^{-1}(1-p_i)$ and 0 otherwise. The LBEM algorithm is as follows.

Step 1: Apply BEM to the data with initial estimate $\vec{\lambda}^{(0)}$.

Step 2: Applying the empirical edge detecting process to the result obtained after K iterations of BEM, determine $l(b,b')$.

Step 3: Calculate $\lambda^{new}(b)$ and go to Step 2 with the data $\lambda^{new}(b)$ until it satisfies the convergence criteria.

When this algorithm is applied to the result obtained after K iteration of MLEM, we call it LEM (MLEM with edge detecting process).

4. SIMULATIONS AND RESULTS

We study two-dimensional reconstruction. We assume a single ring of radius $\sqrt{2}$ with 128 equally spaced detectors around the phantom with $B = 100 \times 100$ and $D = 64 \times 65$. The phantom is made up of 5 ellipses with 4 different emission intensities λ as 0, 5, 10 and 15. Synthetic image, Fig. 1(a), was created by linearly mapping emission intensities onto a linearly gray scale from 0 to 240. We generate 10^6 emissions that agrees with the Poisson model. The details of the data generation methods are as in Shepp and Vardi [1]. The histogram of the 10^6 counts drawn from the phantom of Fig. 1(a) at a rate proportional to $\lambda(b)$ is in Fig. 1(d). The displays in Fig. 1 and Fig. 2 were created by linearly mapping each estimated intensities $\hat{\lambda}(b)$ onto a linearly gray scale from 0 to 255. We use the uniform values as initial estimate of $\vec{\lambda}$. Fig. 1(b), (c), (e) and (f) are the results of the MLEM, BEM, LEM and LBEM algorithm

after 32 iterations respectively. Fig. 1(b) has checkerboard effects as expected. To apply BEM, we use $\beta = 0.01$, $k = 1.05$ and $a_b = 0$ for $b = 1, 2, \cdots, B$ as the parameter values for the prior. Checkerboard effects are suppressed in Fig. 1(c). The reconstructed images are settled down around 20 iterations according to our simulation. However, it shows noise artifacts around edges. We adapt the empirical edge detecting process to the result of MLEM and BEM after 16 iterations respectively. We choose K=16. The result of LEM after 32 iteratons (including 16 iteration of MLEM) shows suppression of the noise around edges. Since empirical edge detecting process introduce a prior in the model, this algorithm has smoothing effects on the image except edges.

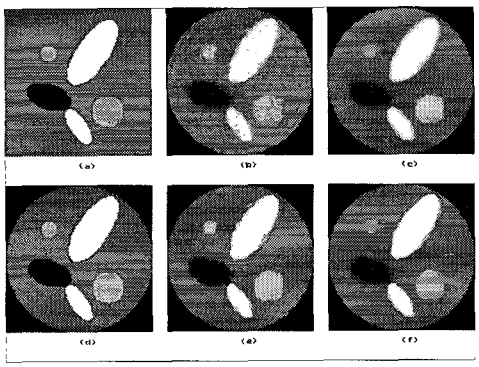


Fig. 1. (a) Synthetic image (b) MLEM after 32 iterations (c) BEM after 32 iterations (d) histogram of 10^6 emissions (e),(f) LEM and LBEM after 32 iterations resp.

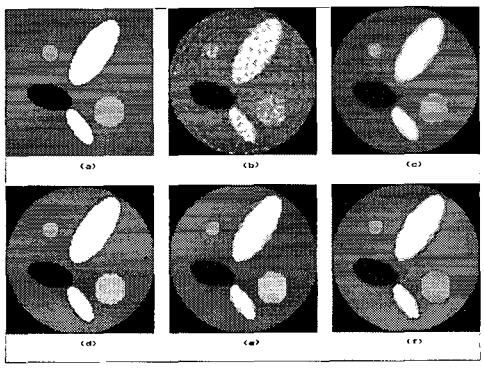


Fig. 2. (a) Synthetic image (b) MLEM after 64 iterations (c) BEM after 64 iterations (d) histogram of 10^6 emissions (e),(f) LEM and LBEM after 64 iterations resp.

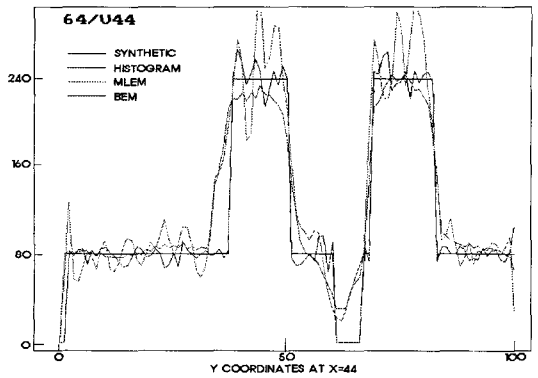


Fig. 3. Line plot of synthetic, histogram, MLEM and BEM through 44th column.

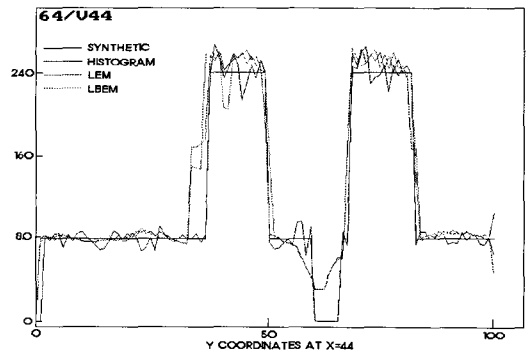


Fig. 4. Line plot of synthetic, histogram, LEM and LBEM through the 44th column.

The result of LBEM after 32 iterations (including 16 iteration of BEM) shows improvement around edges. The results of applying empirical edge detecting process (LEM and LBEM) show clear edge elements compare to the result of MLEM and of BEM. The differences between Fig.1(b), (c) and Fig. 1(e), (f) occur at the edge points. The results of 64 iteration are given in Fig. 2. Fig. 2(a), (d) are the same as Fig 1.(a), (d). The images in Fig. 2(b), (c), (e) and (f) are the results of the MLEM, BEM, LEM and LBEM algorithm after 64 iterations respectively.

The reconstructed image with MLEM tends to get worse as the iterations go on. The results of LEM and LBEM show the clear improvement in the noises and edges. To check the accuracy, we draw a line plot of the 44th column through the

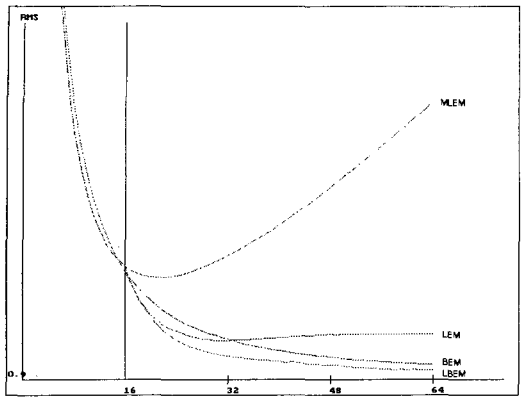


Fig. 5. Root mean square error of MLEM, BEM, LEM and LBEM

y axis of the reconstructed images obtained after 64 iterations. In Fig. 3, the line plot of synthetic, histogram, MLEM and BEM are drawn. In Fig. 4, the line plot of synthetic, histogram, LEM and LBEM are drawn. The x-axis of Fig. 3 and Fig. 4 is the pixel location of y coordinate at x=44 and y-axis is the gray level. These show that MLEM has the noise artifacts and BEM suppresses noise artifacts but has the smoothing effect at the edges. LEM and LBEM detect the edge elements correctly at the edges. We compare root mean square error (RMS) as an overall measure of reconstruction accuracy.

$$\mathrm{RMS} = \sqrt{\frac{\sum_{b=1}^{B}(\lambda(b)-\hat{\lambda}(b))^2}{B}} \qquad (14)$$

In Fig. 5, RMS of MLEM is increased as iterations go on. This indicates that reconstructed image is deteriorated by noise as iterations go on. Since BEM has smoothing effects, RMS is decreased. RMS of LEM and LBEM are decreased until 32 iterations and then those are stabilized as iterations proceed. Since priors introduce blurring effects, RMS of BEM is smaller than that of LEM. Because of edge elements, there is not much changes in the reconstructed images as iterations proceed.

5. CONCLUSIONS

In this paper we have proposed an edge detecting process for MLEM and BEM, we call LEM and LBEM respectively, for image reconstruction of PET. These algorithms suppress the checkerboard effects occurring in the MLEM algorithm and overcome the smoothing effects near edges occurring in the BEM. Based on RMS, the reconstructed images using these algorithms have become stable as iterations proceed. Since priors introduce blurring effects, RMS of BEM is smaller than that of LEM. Appropriate choice of priors is important since the results are sensitive to the priors.

6. REFERENCES

[1] L.A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging,* Vol. MI-1. pp. 113-122, Oct. 1982.

[2] P. J. Green, "Bayesian reconstruction from emission tomography data using a modified EM algorithm," *IEEE Transactions on Medical Imaging,* Vol. 9. pp. 84-93, Mar. 1990.

[3] K. Lange, "Convergence of EM image reconstruction algorithm with Gibbs smoothing," *IEEE Transactions on Medical Imaging,* Vol. 9. pp. 439-446, Dec. 1990.

[4] A. R. De Pierro and M. E. B. Yamagishi, "Fast EM-like methods for maximum a posteriori estimates in emission tomography," *IEEE Transactions on Medical Imaging,* Vol. 20. pp. 280-288, Apr. 2001.

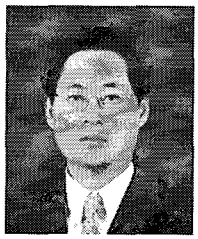
[5] S. Ahn and J.A. Fessler. "Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithm," *IEEE Transactions on Medical Imaging,* Vol. 22. pp. 613-626, May 2003.

[6] Levitan, E. and Herman, G. T., "A Maximum A Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography," *IEEE Transactions on Medical Imaging*, Vol. MI-6, No. 3, 1987.

[7] Bouman, C. and Sauer, K., "A Generalized

Gaussian Image Model for Edge_Preserving MAP Estimation," *IEEE Transactions on Image Processing*, Vol. 2, No. 3, pp. 296-310, 1993.

[8] Geman, S., and Geman, D., "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 6, pp. 721-741, 1984.



Jong-Seok Um

1982 Yonsei Univ. Bachelor of Applied Statistics
1984 Yonsei Univ. Master of Applied Statistics
1991 Ohio State Univ. Dept. of Statistics(Ph.D in Statistics)
1992~Present Dept. of Multimedia, Engineering, Hansung Univ.
Research Interest : Data Mining, Image Recognition