31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Generalized Gibbs Priors Based Positron Emission Tomography Reconstruction

Jing Huang, Jianhua Ma, and Wufan Chen, *Senior Member, IEEE*

***Abstract*— Bayesian methods have been widely applied to the ill-posed problem of image reconstruction. Typically the prior information of the objective image is needed to produce reasonable reconstructions. In this paper, we propose a novel generalized Gibbs prior (GG-Prior), which exploits the basic affinity structure information in an image. The motivation for using the GG-Prior is that it has been shown to suppress noise effectively while capturing sharp edges without oscillations. This feature makes it particularly attractive for those applications of Positron Emission Tomographic (PET) where the objective is to identify the shape of objects (e.g.tumors) that are distinguished from the background by sharp edges. We show that the standard paraboloidal surrogate coordinate ascent (PSCA) algorithm can be modified to incorporate the GG-Prior using a local linearized scheme in each iteration process. The proposed GG-Prior MAP reconstruction algorithm based on PSCA algorithm has been tested on simulated, real phantom data. Comparisons the GG-Prior model with other existing prior model clearly demonstrate that the proposed GG-Prior performs better in lowering the noise, and preserving the edge and detail in the image.**

I. Introduction

Due to the low counting rates and the limited acquisition time, clinical positron emission tomography (PET) data are usually significantly affected by Poisson noise. Reconstructing the PET images is essentially an ill-posed problem [1]. Up to now, many reconstruction strategies have been proposed to address this problem. Bayesian estimation, or maximum a posteriori (MAP) estimation as a statistical approach for incorporating prior information through the choice of a prior distribution for a random field, has already been proved to be an effective solution to the ill-posed PET reconstruction problem[3]. Based on Bayesian theory, a generic contextual constraint can be transformed into some kind of prior information to regularize the solution of the original ill-posed reconstruction problem. Therefore, the regularization by such prior information can be imposed on image reconstruction to suppress noise much more effectively. A common Bayesian prior is the Gibbs distribution of the form

$$P(\lambda) = \frac{1}{\mathrm{Z}} \exp\left(-\beta U(\lambda)\right) = \frac{1}{\mathrm{Z}} \exp\left(-\beta \sum_{j} U(\lambda, j)\right) \quad (1)$$

where Z is a normalizing constant, $U(\lambda)$, as a regularization term, is prior energy function, $U(\lambda, j)$ is any function of a local group of pixel point j. β is a constant that specifies the relative strength of the prior. The specific choice of prior distribution for λ is, of course, a critical component in Bayesian reconstruction.

Usually, $U(\lambda, j)$ is chosen as a shift-invariant function that penalizes the differences between neighboring pixels. However, since the prior is shift-invariant while the likelihood is not, the MAP image, which represents the position dependent bias and resolution, can not usually produce satisfactory results. In addition, images of real world may not be globally smooth and noisy edges intensities may often vary abruptly.

Inspired by some recent studies on the image denoising and segmentation by nonlocal averages [4]–[6], We propose a family of generalized Gibbs priors (GG-Priors) to PET reconstruction. The GG-Priors can exploit the nonlocal information available in the objective image, not only density difference information between individual pixels but also nonlocal connectivity and continuity information in the objective image λ. The GG-Priors can greatly improve the reconstruction quality.

II. PET Statistical Model

We focus on the linear Poisson statistical model that has been used extensively for emission computed tomography, including PET and SPECT. Assuming usual Poisson distribution, the measurement model for emission scans is as follows[7]

$$y_i \sim Poisson\{c_i \sum_{j}^{M} a_{ij}\lambda_j + r_i\}, \quad i = 1, 2, \ldots, N \quad (2)$$

where y_i is the number of photons counted in the ith bin, N is the number of detector pairs, λ_j is the activity at the jth pixel, M is the number of unknown image pixels, r_i represents the total detected random and scattered event counts for detector pair i in emission scan, and $A = \{a_{ij}\}$ is the system matrix, a_{ij} is the geometric probability that an emission photon from image pixel j is detected by the detector pair i in ideal conditions. c_i represents the incorporate calibration factors of scan time, detector efficiencies, attenuation factors and possibly dead time correction factors for the detector pair i. The goal is to estimate the unknown activity distribution image $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_M]'$ based on the measurements $y = [y_1, y_2, \ldots, y_N]'$ with A and $r = [r_1, r_2, \ldots, r_N]'$ being known where ‘’ denotes the matrix transpose.

Then, the corresponding log-likelihood function can be

This work was supported in part by the National Basic Research Program of China under Grant 2003CB716101, in part by the Southern Medical University of China under Grant B1000369.

J. Huang, J. Ma, and W. Chen are with the Institute of Medical Information and Technology, School of Biomedical Engineering, Southern Medical University, Guangzhou, 510515, China. jhma@fimmu.com

written, ignoring constants independent of the λ, as follows

$$L(y,\lambda)=\sum_{i}^{N} h_i([A\lambda]_i) \tag{3}$$

where $h_i(l)=y_i\log(c_il+r_i)-(c_il+r_i)$ and $l_i(\lambda)=[A\lambda]_i=\sum_j^M a_{ij}\lambda_j$.

Therefore the mathematical formulation of PET is the following maximization problem:

$$\max_{\lambda} L(y,\lambda) \tag{4}$$

However, this problem is ill-posed and solving this alone will produce oscillatory solutions. This is why prior information is needed.

III. Generalized Gibbs Priors

Motivated by the nonlocal averages in the image denoising and segmentation by nonlocal averages [4]–[6], we propose a family of generalized Gibbs priors for Bayesian image reconstruction.

The construction of the proposed GG-Priors is generalized as follows

$$U_{GG}(\lambda)=\frac{1}{2}\sum_j\sum_{k\in\mathcal{N}_j} w_{kj}\phi_\alpha(\lambda_k-\lambda_j) \tag{5}$$

$$w_{kj}=\exp(\frac{-|k-j|^2}{2\sigma^2})\exp(\frac{-D(k,j)}{h^2}) \tag{6}$$

$$D(k,j) = \parallel \lambda(\mathcal{V}_k)-\lambda(\mathcal{V}_j)\parallel_2^2 \tag{7}$$
$$\lambda(\mathcal{V}_k) = \{\lambda(l): l\in\mathcal{V}_k\} \tag{8}$$
$$\lambda(\mathcal{V}_j) = \{\lambda(l): l\in\mathcal{V}_j\} \tag{9}$$

where $\mathcal{N}_j$ is a large search neighborhood set to incorporate geometrical configuration information in the image, the weight w_{kj} , which reflects the degree of connectivity affinity between pixel k and pixel j, is defined as a decreasing function of the similarity of the two neighborhoods $\mathcal{V}_k$ and $\mathcal{V}_j$ (named similarity neighborhoods) centered on pixel k and pixel j, respectively. The parameter σ controls how much one wishes to penalize distant of two grid points in the weight, while h controls how much one wishes to penalize similarity of the two patches. Larger σ allows one make use of more remote information, while larger h gives results with sharper features. Fig.1 shows the illustration of weight w_{kj} computation.

Choices of weight w_{kj} and similarity function $D(k,j)$ in (6) can be made as done in [6]. In our experiments, we simply measure the L_2 distance of two cubical patches without doing Gaussian smoothing first. Hence the direct L_2 distance gives a good measurement of similarity, which saves computation time. The function ϕ_α should satisfy (i) $\phi_\alpha\in\mathcal{C}^1$, and (ii) ϕ_α is strongly convex on any bounded interval.Examples of edge-preserving functions ϕ_α that satisfy

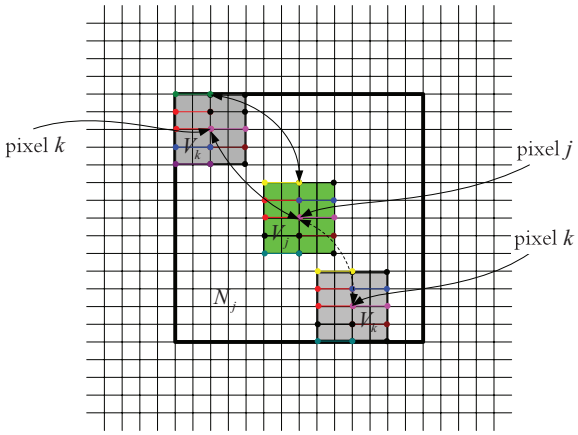


Fig. 1. Illustration of weight w_{kj} computation

the two requirements are:

$$\phi_\alpha = |t|^\alpha, 1<\alpha\le 2, \tag{10}$$
$$\phi_\alpha = 1+\frac{|t|}{\alpha}-\log(1+\frac{|t|}{\alpha}), \alpha>0, \tag{11}$$
$$\phi_\alpha = \log(cosh(\frac{|t|}{\alpha})), \alpha>0, \tag{12}$$
$$\phi_\alpha = \sqrt{\alpha+t^2}, \alpha>0. \tag{13}$$

see[8].

In our experiments we only focused on PET reconstruction based on (10) with $\phi_\alpha(s)=|s|$ and s^2. If we set $\phi_\alpha(s)=s^2$, this GG-prior may be similar to the nonlocal Gibbs Prior[9].

IV. Generalized Gibbs Priors Based Binary Optimal Reconstruction Algorithm

With the analysis of Section 3, choosing the GG-Prior, the PET image λ based on the MAP estimation can be obtained through an iterative maximization of the cost function $\Phi(\lambda)$

$$\hat{\lambda}=\arg\max_{\lambda\ge 0}\Phi(\lambda) \tag{14}$$

$$\Phi(\lambda)=L(y,\lambda)-\beta U_{GG}(\lambda) \tag{15}$$

But, (6)-(9) show that the weight term w_{kj} is the function of objective image λ, which can keep U_{GG} neither quadratic nor convex. Thus, we propose the following binary optimal algorithm to solve the optimization.

After setting an initial estimate $\hat{\lambda}$ for the first iteration, we update weights w_{kj} and the image λ alternatively by the following binary optimal strategy in each iteration phase until convergence.

- ***Weight update***. When $\hat{\lambda}$ is fixed, compute w_{kj} using (6)-(9),
- ***Image update***. For the second stage of the maximization, we hold w_{kj} fixed at its previous estimate and maximize $\Phi(\lambda)$ with regard to λ.

Because the $Hessian$ Matrix of $L(y,\lambda)$ is strictly negative definite[7] and according (6)-(9), we get

$$0<w_{kj}(\hat{\lambda})=\exp(-\frac{|k-j|^2}{2\sigma^2})\exp(-\frac{D(k,j)}{h^2})\le 1. \tag{16}$$

Fig. 2. The *Shepp-Logan* phantom

Considering the ensured convexity for ϕ_α in (5) and the positivity for penalty parameter β, we can reach a conclusion that, given the computed value for $w_{kj}(\hat{\lambda})$, the second derivatives of $\Phi(\lambda)$ in (15) is definitely concave in each iteration. Thus based on the theory of local linearization[10], we can monotonically maximize the posterior energy function $\Phi(\lambda)$ using iterative algorithms such as the paraboloidal surrogate coordinate ascent (PSCA) iterative algorithm[11]. In our study, the standard PSCA algorithm was modified to incorporate the GG-Prior using a local linearized scheme in each iteration process.

V. Experimentation and Analysis

In this section, we present some experiments for PET reconstruction with simulated and real phantom data.

A. Emission Reconstruction of Simulated Data

In this experiment, the synthetic simulated phantom data with 128×128 square pixels was used for emission PET reconstruction. a *Shepp-Logan* head phantom with pixel values from 0 to 8, and the total counts amount to 3×10^6 was designed as shown in Fig.2. The simulated sinogram data for the *Shepp-Logan* phantom was Poisson distributed and the percentages of simulated delayed coincidences r_i factor (scatter effects are ignored) was set to be 10%. The c_i factor was generated using pseudo-random log-normal variates with a standard deviation of 0.3 to account for detector efficiency variations. Generated by the ASPIRE software system [12], the transition probability matrixe used in the reconstructions of the *Shepp-Logan* phantom corresponded to parallel strip-integral geometry with 128 radial samples and 128 angular samples distributed uniformly over 180 degrees. The code was run on a PC with Intel(R) Pentium(R) 4 3.16GHz 3.00GHz processor and 2GB of memory.

FBP reconstruction and Bayesian reconstructions based on the Huber prior (Huber-Prior), the proposed GG-Prior were performed, respectively. For FBP method, we chose the ramp filter with cutoff frequency equal to the Nyquist frequency. We did not consider methods for choosing the penalty parameter β, decay parameter h and threshold parameter δ in either algorithm, but rather studied the results obtained by a broad range of parameter values by hand to give the reconstructions in terms of maximum of signal-to-noise ratio(*SNR*), which was computed by (17)

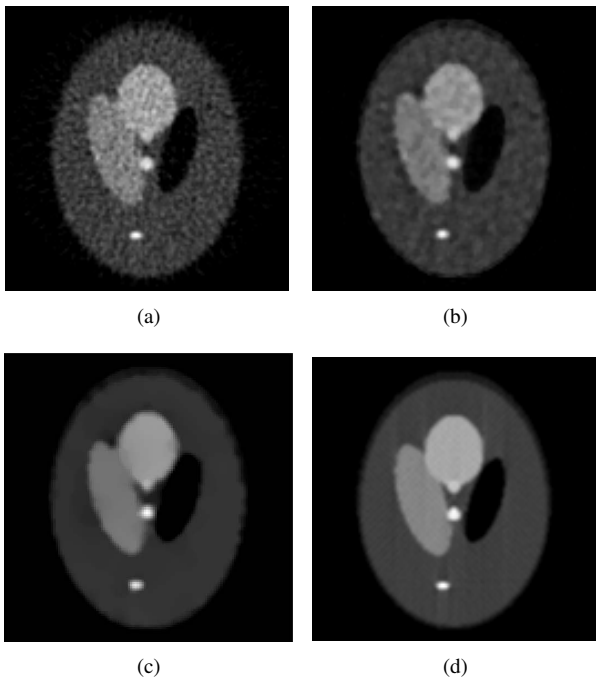


Fig. 3. FBP reconstruction and Huber-Prior and GG-Priors based Bayesian reconstructions in the *Shepp-Logan* phantoms study. (a) is the FBP reconstruction, (b)is the Huber-Prior reconstruction, (c)is the GG-Prior reconstruction with $(\phi_\alpha = t^2)$, and (d)is the GG-Prior reconstruction with $(\phi_\alpha = |t|)$

$$SNR = 10\log_{10}\left(\frac{\sum_{i,j}(\lambda(i,j)-\bar{\lambda})^2}{\sum_{i,j}(\lambda(i,j)-\lambda_{phantom}(i,j))^2}\right) \quad (17)$$

where $\lambda(i,j)$, $\bar{\lambda}$ and $\lambda_{phantom}(i,j)$ denote the reconstructed image, the mean of the reconstructed image λ and the original true phantom image, respectively.

Fig.3 shows the reconstructed images using the *Shepp-Logan* phantoms data with FBP and different priors. For the PACS based GG-Prior algorithm the search neighborhood $\mathcal{N}_j$ in (5) was set to be 21×21 neighborhood and the two similarity neighborhoods $\mathcal{V}_k$ and $\mathcal{V}_j$ in (7) were both set to 7×7 neighborhood. For the reconstruction using the Huber-Prior, the 8-neighborhood was used. The values of h for the GG-Prior and δ for the Huber-Prior were fixed to 0.8 and 0.2, respectively, β was set to 1.40 in all the reconstructions. Fig.3.(a)(b)(c)(d) are the reconstructed images using the FBP with the ramp filter,the Huber-Prior, the proposed the GG-Prior reconstruction with $(\phi_\alpha = t^2)$, and the GG-Prior reconstruction with $(\phi_\alpha = |t|)$, respectively.

Table I displays *SNR* comparisons with regards to *Shepp-Logan* phantoms data in Fig.3 for above reconstructions. It indicates that the images from the reconstructions using the proposed GG-Priors reconstructed images with much higher *SNRs* than other reconstruction methods.

B. Emission Reconstruction of Real Phantom Data

In this study, a real *hotlesion* phantom emission scan data was used. 192 radial samples and 192 angular samples over

TABLE I

SNRS FOR THE RECONSTRUCTED IMAGES IN FIG.3

SNRs(dB)	*Shepp-Logan* phantom
FBP	11.63
Huber-Prior	15.14
GG-Prior($\phi_\alpha = t^2$)	16.33
GG-Prior($\phi_\alpha = \|t\|$)	16.91

180 degrees. The sinogram was precorrected for scatter, attenuation and geometric arc correction etc. The 5th 2D slice data of the total 9 slices was used. The objective image consisted of 128×128 pixels with $0.356cm$ pixel resolution. The total counts amounted to 4.4×10^6. All c_i factors and r_i factors were assumed to be ones and zeros, respectively.

The reconstructed image from the FBP method was used as the initial image in the iterative algorithms. In order to test the GG-priors, we compared it with the Huber-Prior. Iteration numbers for all Bayesian reconstructions were set to be 50. The PSCA based GG-Prior algorithm was also used here.

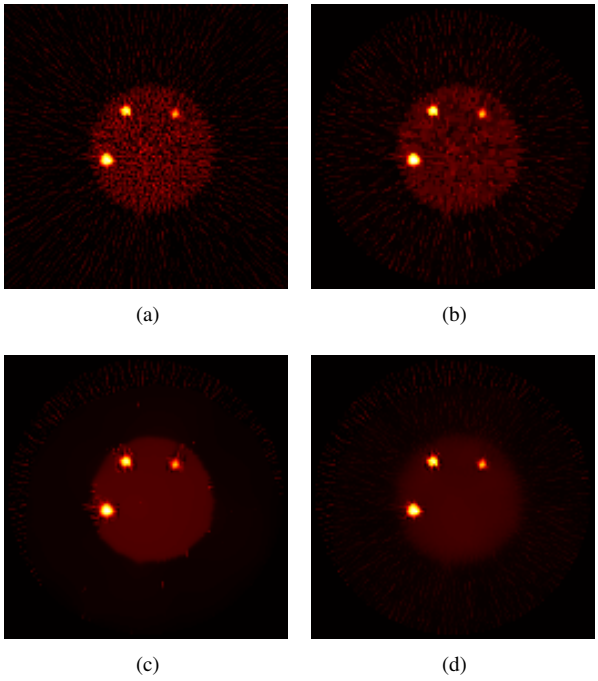
Fig.4 shows the reconstructed results from different methods with different parameter settings. For all of Bayesian reconstruction problems, we manually selected parameters β, δ and h to cover a range of parameter settings. Obviously, when suitable parameter values are set, the reconstructions using the proposed GG-Priors outperformed the FBP reconstruction and reconstructions using the Huber-Prior in terms of noise suppression and edge preserving, and able to reconstruct more appealing emission images.

VI. CONCLUSIONS AND FUTURES

Based on Bayesian theory, the proposed GG-Prior is theoretically reasonable and straight forward. Though choosing a large search neighborhood $\mathcal{N}$ and a new weighting strategy that incorporates the connectivity affinity between pixels in $\mathcal{N}$, the proposed GG-Prior is able to provide more image nonlocal similarity information for the original ill-posed reconstruction.

We also introduced an algorithm, called the PSCA+GG-Prior algorithm, that can be feasibly and effectively implemented using the proposed binary optimal reconstruction strategy. Because the proposed PSCA+GG-Prior algorithm can be considered as a version of OSL algorithm, it is easily straightforward to implement, but nonnegative estimates cannot be guaranteed and, like many existing algorithms, convergence is an open issue. Although the tradeoff between resolution and noise can be controlled by certain regularization parameters(i.e. β and h), like many researchers, we have not determined a way to choose the parameters so that the data-fit and a prior information are optimally "balanced".

Future work includes analyzing the convergence property of the reconstruction algorithm, exploring effective ways to reduce the computational cost and study automatic parameter methods.



(a) (b) (c) (d)

Fig. 4. FBP reconstruction and Bayesian reconstructions using different priors in the real hotlesion phantom studies. (a)is the FBP reconstruction, (b) is Huber-Prior reconstruction, (c)is the GG-Prior reconstruction with ($\phi_\alpha = t^2$), and (d)is the GG-Prior reconstruction with ($\phi_\alpha = |t|$)

REFERENCES

[1] M. Bertero, T. Poggio, and V. Torre, Ill posed problems in early vision, *Proc. IEEE*, vol. 76, 1988, pp 869-889.

[2] K. Lange and R.Carson, EM reconstruction algorithm for emission and transmission tomography, *J.Comp.Assisted.Tomo.*,vol. 8, 1984, pp 306-316.

[3] K. Lange, Convergence of EM image reconstruction algorithms with Gibbs smoothness, *IEEE Trans. Med. Imag.*, vol .9, 1990, pp 439-446.

[4] A. Buades, B. Coll, and J. M. Morel, A nonlocal algorithm for image denoising, in *Proc. IEEE Int. Conf. Computer Vision Pattern Recognition*, vol. 2, 2005, pp 60-65.

[5] A. Buades, B. Coll, J.M Morel, A review of image denoising algorithms, with a new one, *Multiscale Modeling and Simulation* (SIAM interdisciplinary journal), vol. 4, 2005, pp 490-530.

[6] G. Gilboa and S. Osher, Nonlocal linear image regularization and supervised segmentation, *UCLA CAM*, 2006, Report 06-47.

[7] L.A. Shepp and Y. Vardi, Maximum likehood reconstruction for emission tomography, *IEEE Trans. Med. Imag.*, vol.1, 1982, pp 113-121.

[8] R.H. Chan and C. Ho, Convergence of Newton's method for a minimization problem in impulse noise removal ,*Journal of Computational Mathematics*, vol. 22, 2004 , pp 168-177.

[9] Yang Chen, Wufan Chen, Pengchen Shi, Yanqiu Feng, Qianjin Feng, Qingqi Wang, and Zhiyong Huang, Bayesian reconstruction using a new nonlocal MRF prior for count-limited PET transmission scans, *Lecture Notes in Computer Science*, vol. 4633, 2007, pp 972-981.

[10] Wufan Chen, Ming Chen, and Jie Zhou, Adaptively regularized constrained total least-squares Image Restoration, *IEEE Trans. Image Process.*, vol. 9, 2000, pp 588-596.

[11] J. A. Fessler and H. Erdoğan, A paraboloidal surrogates algorithm for convergent penalized-likelihood emission reconstruction, in *Proc. IEEE Nuc. Sci. Symp. Med. Imag. Conf*, vol. 2, 1998, pp 1132-1135.

[12] J. A. Fessler, Aspire 3.0 user's guide: A sparse reconstruction library, *Communication & Signal Processing Laboratory Technical Report* No. 293, Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, 1998.