Maximum a posteriori X-ray computed tomography using graph cuts

Shin-ichi Maeda¹, Wataru Fukuda¹, Atsunori Kanemura¹,² and Shin Ishii¹
¹ Graduate School of Informatics, Kyoto University, Kyoto 611-0011, Japan
² Department of Electrical Engineering, University of California, Santa Cruz, CA 95064, USA
E-mail: {fukuda,ichi,atsu-kan,ishii}@sys.i.kyoto-u.ac.jp

Abstract. We develop maximum a posteriori (MAP) method for X-ray computed tomography (CT). We present a mixture prior to represent the knowledge that the human body is composed of a finite number of material kinds whose CT values are roughly known in advance. The tomographic image and material classes are simultaneously estimated in an alternating manner, where a graph cut algorithm is used to minimize the MAP objective function. Experiments show that the proposed algorithm performs better than the existing methods in severe situations where samples are limited or metals are inserted into the body.

1. Introduction

Computed tomography (CT) reconstructs tomographic images from their projections, and X-ray CT is a widely used technique for e.g. medical diagnosis and noninvasive tests. A number of CT algorithms have been proposed in the literature [1–8], but there still exist difficulties that must be resolved. This article addresses two difficulties, namely, CT under fewer projections [9; 10] and metal artifact reduction (MAR) [4–7]. Our approach is based on the knowledge that the objects to be reconstructed are composed of a finite number of material kinds, e.g. air, fat, normal tissue (muscle), bone, metal, and so on, and that their CT values (X-ray attenuation coefficients) are roughly known in advance.

The background behind the difficulty of fewer projections is the need to avoid overdoses of radiation. Exposure to X-ray can cause cancer, and recent studies [9; 10] have reported the cancer risk may be higher than that expected; therefore, X-ray exposure should be minimized. However, the reduced exposure to radiation results in low signal-to-noise ratio (SNR), making the observed projection data noisy due to the fluctuation of the detected number of photons. One way to deal with low SNR is to use a good regularization on or a priori knowledge of the objects to be reconstructed. The classical filtered back projection (FBP) or maximum likelihood (ML) methods for CT do not utilize such prior knowledge, and they produce unstable solutions to this ill-posed problem of CT. In contrast, maximum a posteriori estimation incorporates a priori knowledge of the objects as a prior probability, and this is the approach we adopt in this study. Our knowledge is expressed as a mixture distribution that reflects the finite nature of the objects’ material kinds.

The need for the second difficulty, MAR, is based on the fact that high density objects such as metal prostheses and dental fillings attenuate X-ray so strongly that the projections
passing through them have almost no information. If we reconstruct the tomographic image from strongly attenuated projections by using a naive algorithm such as FBP, streak or star artifacts will be present in the reconstructed image [4–7]. The most popular approach to coping with MAR is the projection completion (PC) method [4], which tries to recover the lost information by detecting and interpolating the metal regions in the sinogram (projected domain) before applying FBP. A weak point of the PC method is that it is not easy to determine the metal regions correctly, especially when two or more metals are inserted. Moreover, the PC method is heuristic and may result in a suboptimal solution because the interpolation stage is usually independent of the reconstruction stage. Our approach to MAR is to include the metal class in our mixture prior. MAP estimation with our prior results in a statistically sound way of identifying the metal regions, in turn giving rise to artifact-reduced tomographic reconstruction by putting lower reliability to the projections passing through the identified metal regions.

The use of the mixture knowledge has been adopted by several works [3; 6–8]. Most of them [3; 6; 7] estimate the tomographic image and the material classes in a separate manner. Since the estimation of reconstructed image and the material classes are closely related, such separate estimation may deteriorate the reconstruction performance. Therefore, simultaneous estimation of a tomographic image and material classes in terms of a consistent cost function is desired. One such solution has been proposed in [8], whose probability model is simpler than our model. Our previous reports [11; 12] on CT had the same spirit as presented in this article, but they were based on variational Bayesian estimation [11] or include experimental results obtained in easier settings than conducted in this study [12]. Especially, in this study, we investigate the robustness of our algorithm to the model misspecification.

Our MAP estimation algorithm achieves simultaneous estimation of the material class and tomographic image by iteratively minimizing the MAP objective function in an alternating manner. In particular, when estimating the material class, we use a graph cut algorithm called \( \alpha \)-expansion [13–15] for efficient optimization.

2. Model and algorithm

Suppose we have \( T \) projections \( D = \{ y^{(1)}, \cdots, y^{(T)} \} \) and the \( t \)th projection is represented by \( y^{(t)} = \{ y_1^{(t)}, \cdots, y_J^{(t)} \} \), where \( y_i^{(t)} \) denotes the number of photons sensed by detector \( i \) when projected from direction \( \theta_i^{(t)} \) (Fig. 1). The average number of photons captured at detector \( i \) is determined by the X-ray attenuation coefficients called CT values of the object which the X-ray pass through. For simplicity, the object is discretized into pixels \( x = \{ x_1, \cdots, x_J \} \) and each pixel \( x_j \) \( (j = 1, \cdots, J) \) is assumed to have a single CT value. Our aim is to reconstruct the underlying CT values of the object, i.e., tomographic image, represented by a \( J \)-dimensional vector \( x \).

2.1. Observation model

In average, the data acquisition model is written by

\[
\hat{y}_i^{(t)} = b_i^{(t)} \exp \left\{ -\sum_{j=1}^{J} l_{ij}^{(t)} x_j \right\},
\]

(1)

where \( \hat{y}_i^{(t)} \) denotes the expected number of detected photons, \( b_i^{(t)} \) is the number of photons that would be detected in the absence of any absorber, and \( l_{ij}^{(t)} \) is the effective intersection length of projection \( i \) with pixel \( j \) when projected from direction \( \theta_i^{(t)} \) (Fig. 1).

The quantum nature of X-ray photons fluctuates \( \hat{y}_i^{(t)} \) [2; 3; 7], making the probability
distribution of the measurements the product of independent Poisson distributions:

\[
p(D|x) = \prod_{t=1}^{T} \prod_{i=1}^{I} p(y_i^{(t)}|x) = \prod_{t=1}^{T} \prod_{i=1}^{I} \frac{y_i^{(t)} e^{-y_i^{(t)}}}{y_i^{(t)}!}.
\]  

(2)

2.2. Conditional prior

Our a priori knowledge is incorporated into the prior probability of the tomographic image \( x \). We assume that each element of \( x \) is associated with a material class, and that the classes include air, fat, normal tissue (muscle), bone, and metal. We denote the number of classes by \( K (= 5) \). The material class of each pixel \( j \) is represented by hidden variables \( z_j = [z_{j1}, \ldots, z_{jK}] \in \{0,1\}^K \). We adopt 1-of-\( K \) coding for \( z_j \); that is, \( z_{jk} = 1 \) and the other elements \( \{z_{jk'} | k' \neq k\} \) are all 0 when \( x_j \) is associated with the \( k \)th material class. Then, the prior \( p(x) \) is represented using hidden variables \( z = \{z_1, \ldots, z_J\} \) as

\[
p(x) = \sum_z p(x|z)p(z) = \sum_z \prod_{j=1}^{J} p(x_j|z_j)p(z_j).
\]  

(3)

Given the material classes, we assume the attenuation coefficient \( x_j \) obeys a Gaussian distribution

\[
p(x_j|z_j) = \prod_{k=1}^{K} \mathcal{N}(x_j|\nu_k, r_k^2)^{z_{jk}},
\]  

(4)

where \( \nu_k \) and \( r_k^2 \) are the mean and variance of the Gaussian distribution, respectively.

Although the class-wise means \( \nu_k \) should be calibrated in advance, this task would not be required for every shot of CT scanning because the values of \( \nu_k \) are roughly determined from the material physical characteristics. Variations of attenuation coefficient from \( \nu_k \) due to the different compositional elements over the fine tissues or organs are assumed to be captured by the randomness of Gaussian distribution, whose uncertainty is controlled by its variance \( r_k^2 \).
Table 1. Parameter settings.

| k: material       | \( \nu_k \)  | \( r^2 \) | \( J_{\text{self}}^k \) | \( J_{\text{inter}}^k \) |
|-------------------|--------------|---------|----------------|----------------|
| 1: air            | 0.000        | \( 3 \times 10^{-5} \) | 0.333          | 0.67           |
| 2: fat            | 0.018        | \( 3 \times 10^{-5} \) | 0.167          | 0.67           |
| 3: normal tissue  | 0.022        | \( 3 \times 10^{-5} \) | 0.167          | 0.67           |
| 4: bone           | 0.050        | \( 3 \times 10^{-5} \) | 0.003          | 0.67           |
| 5: metal          | 0.120        | \( 3 \times 10^{-5} \) | 0               | 0.67           |

The mean values \( \nu_k \) are set at \( \nu_1 = 0.000 \), \( \nu_2 = 0.018 \), \( \nu_3 = 0.022 \), \( \nu_4 = 0.050 \), and \( \nu_5 = 0.120 \) for air, fat, normal tissue, bone, and metal, respectively (Table 1). Although these values are determined according to the previous studies [6; 16] and hence are reasonable to some extent, we do not pursue the reality so much here because we put our focus on examining the robustness of our model to the model misspecification.

2.3. Prior for class
The prior of the material class is given as the following Boltzmann distribution:

\[
p(z) = \frac{1}{Z} \exp\{-H(z)\},
\]

where \( Z \) is a normalizing constant and the energy is defined by

\[
H(z) = - \sum_{k=1}^{K} \left( J_{\text{self}}^k \sum_{j=1}^{J} z_{jk} + J_{\text{inter}}^k \sum_{j=1}^{J} \sum_{s \in \eta(j)} z_{jk} z_{sk} \right).
\]

Here, \( \eta(j) \) represents the set of pixels adjacent to pixel \( j \), and \( J_{\text{self}}^k \) and \( J_{\text{inter}}^k \) are nonnegative constants that control the characteristics of the class prior. The Boltzmann distribution takes a high probability when energy \( E(z) \) is low. Therefore, the first term of the energy function regulates the relative proportion of each material so that large \( J_{\text{self}}^k \) promotes \( z_{jk} = 1 \). The second term defines the degree of correlation between neighboring pixels within the material, so as to control the spatial continuity; hence \( J_{\text{inter}}^k > 0 \) enhances smoothness of the material region.

2.4. MAP estimation
According to the Bayes theorem, the joint posterior distribution for \( x \) and \( z \) is proportional to the product of the prior and likelihood functions:

\[
p(x, z|D) \propto p(D|x)p(x|z)p(z).
\]

We estimate \( x \) and \( z \) by maximizing this posterior probability. Taking the negative logarithm of (7), the following objective function is minimized with respect to \( x \) and \( z \):

\[
E(x, z) = - \ln p(D|x) - \ln p(x|z) - \ln p(z).
\]

Since the joint optimization for continuous variable \( x \) and discrete variable \( z \) is difficult, we iteratively update each component of the objective function in an alternating manner. Each update step consists of

\[
\hat{x} = \arg \min_{x} E(x, \hat{z})
\]
Input : Observation $\mathcal{D}$
Output : Estimate of CT value $\mathbf{x}$

1 : until $||\mathbf{x}|| < 10^{-6}$ do
2 : for $j = 1$ to $J$ do
3 : Update $x_j$ to minimize (9) by SCG method
4 : for $j = 1$ to $J$ do
5 : Update $z_j$ to minimize (10) by graph cuts

**Figure 2.** Proposed algorithm for CT reconstruction.

and

$$\hat{z} = \arg \min_{\mathbf{z}} E(\hat{\mathbf{x}}, \mathbf{z}). \quad (10)$$

The optimization with respect to $\hat{\mathbf{x}}$ is achieved by using the scaled conjugate gradient (SCG) method [17]. The other optimum $\hat{z}$ is obtained by $\alpha$-expansion algorithm, an instance of graph cut methods [13–15]. The entire iterations are terminated when the relative change of the $\hat{x}$’s norm is smaller than a predetermined threshold $10^{-6}$. Fig. 2 shows our proposed algorithm in pseudocode.

### 2.5. Graph cut

Graph cut is an efficient optimization algorithm for solving a particular class of discrete optimization problems. Although a general class of discrete optimization problems is NP hard, the graph cut finds a global optimum in a polynomial time by restricting a class of the optimization problems such to satisfy certain conditions like submodularity [15].

$\alpha$-expansion algorithm, which we use here, repeatedly utilizes graph cuts [13] so that it reduces the optimization of the multi-valued variable to the optimization of the binary variable for which the original graph cut can efficiently find the global optimum. This modification makes $\alpha$-expansion algorithm applicable to a wider class of optimization problems, e.g., image restoration, stereo and segmentation [13; 14; 18]; on the other hand, $\alpha$-expansion does not necessarily find a global optimum, but finds a local optimum.

Basically, the graph cut solves the objective function of the following type:

$$E(\mathbf{z}) = \sum_{i} g_{i}(z_{i}) + \sum_{(i,j) \in E} h_{ij}(z_{i}, z_{j}), \quad (11)$$

where $\mathbf{z} \in L$ takes discrete values either binary or multi-valued, $g_{i}$ is any function of $z_{i}$, and $h_{ij}$ is a particular function of $z_{i}$ and $z_{j}$. For the $\alpha$-expansion algorithm, $h_{ij}$ should satisfy the following condition.

For any $\alpha, \beta, \gamma \in L$, $h_{ij}(\alpha, \alpha) + h_{ij}(\beta, \gamma) \leq h_{ij}(\alpha, \gamma) + h_{ij}(\beta, \alpha). \quad (12)$

In our case, $h_{ij}$ corresponds with the second term of the right hand side of (6), and satisfies the above condition as long as $J_{k}^{\text{inter}} \geq 0$.

### 3. Experimental results

We test our method by reconstructing phantom images in two severe situations. In the first experiment, the number of projections is severely restricted, as to simulate a low-exposure and
Figure 3. Model and True CT values. All model CT values are represented by a Gaussian distribution with mean $\nu_k$ and variance $\nu_k^2$ (written in capital letter). True CT values are (i) air: not changed, (ii) bone and metal: deviated from $\nu_k$ by $|r_k|$, and (iii) fat and normal tissue: deviated to be (case A) more distant from or (case B) closer to each other.

low-SNR situation. This setting is prepared to see the reconstruction performance when the X-ray exposure is minimized. In the second experiment, the phantom includes metal regions. This setting is to see how metal artifacts can be reduced by our method.

To see the robustness to the parameter setting of our algorithm, all the parameters are fixed throughout the experiments (see Table 1), and the true CT values are deviated from the assumed CT values $\nu_k$, except for air whose CT value is assumed to be known and not so fluctuated. For bone and metal, the true CT values are set to be smaller than $\nu_k$ by the standard deviations $|r_k|$. For fat and normal tissue, the true CT values are set to be more distant from (case A) or closer to (case B) each other as shown in Fig. 3. This makes a more difficult setting than that conducted in [11].

The CT values we used here can be transformed into the Hounsfield unit (HU) by the following transformation: $1000(x - x_0)/x_0$, where $x_0 (= 0.02)$ is the attenuation coefficient of water (H$_2$O). The average and the standard deviation of CT values, $\nu_k$ and $|r_k|$, are determined according to the known HU for several materials. The parameters $J^\text{self}_k$ and $J^\text{inter}_k$ are set at hand-tuned values.

3.1. CT with fewer projections

A phantom to be reconstructed is shown in Fig. 4(a) (471 $\times$ 353 mm ellipse). Parallel beam acquisition is simulated using 367 detectors and 32 projection angles over 180$^\circ$. The blank scan value $b_i$ is set to $10^5$. Images of $256 \times 256$ pixels are reconstructed. The following three different approaches are compared: filtered back projection (FBP), maximum likelihood (ML) [2], and our proposed MAP method.

The estimation results for case A (the true CT values of fat and normal tissues are more distant than the assumed $\nu_k$ values) are shown in Fig. 4. The panels show (a) the ground truth image and reconstructed images by (b) FBP, (c) ML, and (d) the proposed method. The performances are measured by peak signal-to-noise ratio (PSNR), which is shown at the bottom of each panel. PSNR is defined as $20 \log_{10} \left( \left(J \max_{j \in \{1, \ldots, J\}} |x^*_j| \right)/\left(\sum_{j=1}^{J} |x^*_j - \hat{x}_j|\right) \right) \text{(dB)}$ where $x^*$ and $\hat{x}$ are true and estimated CT values, respectively. The PSNR of our algorithm (21.13 dB) is higher than those by the existing algorithms. Although the reconstruction results of FBP and ML are very noisy due to the limited number of projections, our result is reasonably smooth thanks to the prior incorporating the knowledge of material classes. Also, the class
Figure 4. Results of CT reconstruction with fewer projections (case A). The window used is $[-500, 500]$ HU so that the corresponding $x_j$ values are within $[0.01, 0.03]$.

Figure 5. Estimated class regions in the experiment with fewer projections (case A).

regions estimated by our algorithm are shown in Fig. 5. As shown in Fig. 5, a smooth class segmentation is achieved.

The estimation results for case B (the true CT values of fat and normal tissues are closer than the assumed $v_k$ values) are shown in Fig. 6. Similar to the case A, our proposed method yields the best reconstruction result while the other algorithms estimate noisy CT images. Corresponding to the good reconstruction performance, our proposed method gives a reasonably good segmentation as shown in Fig. 7. Overall, the quality of the estimated CT image in case B is better than that obtained in case A. It is considered that the closer radiodensities of rather high density materials (normal class and bone class) make the problem in case A more difficult than that in case B.

3.2. Phantom data with metal inserted

In this second experiment, we reconstruct a head phantom ($472 \times 436$ mm ellipse) shown in Fig. 8 that includes three dental fillings made of metal (three disks with diameters 18, 19, and
Figure 6. Results of CT reconstruction with fewer projections (case B). The window used is $[-500, 500]$ HU so that the corresponding $x_j$ values are within $[0.01, 0.03]$.

Figure 7. Estimated class regions in the experiment with fewer projections (case B).

23 mm). Parallel beam acquisition is simulated using 367 detectors and 180 projection angles over $180^\circ$. The blank scan value $b_i$ is set to $10^6$. The size of reconstructed images is $256 \times 256$ pixels. Reconstruction is performed by FBP, ML, the projection completion method based on linear interpolation in the sinogram (PCLIS) [4], and our proposed MAP method.

The reconstruction results in case A are shown in Fig. 8. The panels show (a) the ground truth and reconstructed images by (b) FBP, (c) PCLIS, (d) ML, and (e) the proposed method. The corresponding PSNR value is shown at the bottom of each panel. Our algorithm achieved the highest PSNR ($30.98$ dB) among the algorithms we compared. Good smoothing within each region is obtained by our method due to the successful segmentation. The class regions estimated by our algorithm are shown in Fig. 9.

3.3. Relation between the number of projections and PSNR

We investigate the relation between the number of projections and PSNR. The same phantom and experimental condition as those in Section 3.1 were used. To reduce the computation cost,
Figure 8. Results of CT reconstruction of a phantom with metal regions. The window used is $[-500, 500]$ HU so that the corresponding $x_j$ values are within $[0.01, 0.03]$.

Figure 9. Estimated class regions in the experiment with metal.

we test on a smaller image of $64 \times 64$ pixels. In this experiment, we add the results of FBP (FBP2) where the blank scan value is twice with the one used in the other algorithms. Six runs are performed for each algorithm. Fig. 10 shows the mean and standard deviation of PSNR estimated by each algorithm. As seen in the figure, our algorithm almost always achieves the best performance. The result of our algorithm is even better than that of FBP2, whose PSNR saturates as the number of projections increases. It is considered that FBP yields a biased estimation due to the filtering operation, which removes blurs that are arisen in the CT image obtained by a naive back projection algorithm. Overall, our algorithm achieves the excellent reconstruction performance, and could reduce the exposure less than the half of that achieved by the other algorithms when the PSNR of the CT image is fixed.

3.4. Relation between the number of projections and computation time

We also investigate the relation between the number of projections and PSNR under the conditions given in Section 3.1 of case A. All the algorithms are implemented in Matlab, and they are run on the machine which has four Intel Dual-Core Xeon 5470 3.33 GHz processors. The result is summarized as follows. First, the computation time of FBP is always the smallest.
Even when the image size is 256x256, the computation time of FBP is less than one second. Second, the computation time of ML and our proposed MAP increases as the image size grows, and the computation time of our MAP method is larger than that of ML. The comparison of the computation time between ML and our proposed MAP is shown in Fig. 11. The fact that our proposed MAP requires more computation time than ML is natural because MAP performs ML for the initial estimate of the CT values. Rather, we take notice that the difference between ML and MAP is not so large than we expected, which suggests the quick convergence of the coordinate descent optimization for the CT values and the material classes. We also observe the relatively low computation cost for the optimization of the material class compared with that of CT values thanks to the efficient algorithm, graph cuts.

4. Conclusion
In this study, we have described a CT reconstruction method based on the optimization of the MAP objective function. The key point of our method is the introduction of the knowledge of material classes as a mixture prior that regularizes the solution with appropriate smoothing within each class region and allows the existence of highly attenuating objects like metal. Our method works well under fewer projections and achieves reduction of metal artifacts compared to the existing algorithms. Furthermore, it shows better generalization even when the true tomographic image has different class-wise means from the assumed mean values. Our material class model can be straightforwardly extended to what includes more classes. And, it can be applied to detection of tumor classes or to identification of anatomical structures, because segmentation of CT values is performed as a by-product of our algorithm.

Acknowledgments
This study was supported in part by Grant-in-Aid for Scientific Research on Priority Areas, “Deepening and Expansion of Statistical Mechanical Informatics,” from MEXT, Japan. A. Kanemura was supported by Grant-in-Aid for JSPS Fellow 20-8856 and by Excellent Young Researchers Overseas Visit Program from JSPS.
Figure 11. Image size and computation time

References

[1] Kak A C and Slaney M 1988 *Principles of Computed Tomographic Imaging* (New York: IEEE Press)

[2] Nuyts J, De Man B, Dupont P, Defrise M, Suetens P and Mortelmans L 1998 *Phys. Med. Biol.* 43 729

[3] Hsiao I T, Rangarajan A and Gindi G 2002 *IEEE Trans. Image Process.* 11 1466

[4] Kalender W A, Hebel R and Ebersberger J 1987 *Radiology* 164 576

[5] Watzke O and Kalender 2004 *Eur. Radiol.* 14 849

[6] Bal M and Spies L 2006 *Med. Phys.* 33 2852

[7] Lemmens C, Faul D and Nuyts J 2009 *IEEE Trans. Med. Imag.* 28 250

[8] Mohammad-Djafari A 2008 *Int. J. Tomogr. Stat.* 11 76

[9] Berrington de González A, Mahesh M, Kim K P, Bhargavan M, Lewis R, Mettler F and Land C 2009 *Arch. Intern. Med.* 169 2071

[10] Smith-Bindman R, Lipson J, Marcus R, Kim K P, Mahesh M, Gould R, Berrington de González A and Miglioretti D L 2009 *Arch. Intern. Med.* 169 2078

[11] Fukuda W, Maeda S, Kanemura A and Ishii S 2010 *IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*

[12] Fukuda W, Maeda S, Kanemura A and Ishii S 2010 *Int. Symp. Artificial Life and Robotics (AROB)*

[13] Boykov Y, Veksler O and Zabih R 2001 *IEEE Trans. Pattern Anal. Mach. Intell.* 20 1222

[14] Boykov Y and Kolmogorov V 2004 *IEEE Trans. Pattern Anal. Mach. Intell.* 26 1124

[15] Kolmogorov V and Zabih R 2004 *IEEE Trans. Pattern Anal. Mach. Intell.* 26 147

[16] Wei J, Chen L, Sandison G A, Liang Y and Xu L X 2004 *Phys. Med. Biol.* 49 5407

[17] Bishop C M 1995 *Neural Networks for Pattern Recognition* (New York: Oxford University Press)

[18] Richard S, Ramin Z, Daniel S, Olga V, Vladimir K, Aseem A, Marshall T and Carsten R 2008 *IEEE Trans. Pattern Anal. Mach. Intell.* 30 1068