Computerized Medical Imaging and Graphics 37 (2013) 304–312

Contents lists available at SciVerse ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Parameter optimization of relaxed Ordered Subsets Pre-computed Back Projection (BP) based Penalized-Likelihood (OS-PPL) reconstruction in limited-angle X-ray tomography

Shiyu Xu [a], Henri Schurz [b], Ying Chen [a,c,*]

[a] Department of Electrical and Computer Engineering, Southern Illinois University Carbondale, IL 62901, USA
[b] Department of Mathematics, Southern Illinois University Carbondale, IL 62901, USA
[c] Biomedical Engineering Graduate Program, Southern Illinois University Carbondale, IL 62901, USA

ARTICLE INFO

Article history:
Received 28 September 2012
Received in revised form 28 April 2013
Accepted 30 April 2013

Keywords:
Limited angle tomography
Mono-energetic X-ray
Penalized-likelihood
Predictable resolution properties
Relaxed Ordered Subsets
Smoothing parameter

ABSTRACT

This paper presents a two-step strategy to provide a quality-predictable image reconstruction. A Pre-computed Back Projection based Penalized-Likelihood (PPL) method is proposed in the strategy to generate consistent image quality. To solve PPL efficiently, relaxed Ordered Subsets (OS) is applied. A training sets based evaluation is performed to quantify the effect of the undetermined parameters in OS, which lets the results as consistent as possible with the theoretical one.

1. Introduction

Limited angle X-ray tomography is increasingly used in a range of non-invasive anatomical imaging applications. On one hand, few sampling is inevitable due to the constraints of geometric configuration, limitations on image acquisition time, or the necessity to reduce patient radiation dose. On the other hand, the benefits from three-dimensional (3D) reconstruction can be obtained by the limited angle configuration. These benefits include the feasibility in detecting anatomical structure with overlaps and localizing the region of interest. Current clinical applications include: intra-operative imaging for reference with a pre-operative planning CT, angiography, chest tomosynthesis, dental tomosynthesis, cardiac CT, orthopedic imaging, and, most recently, Digital Breast Tomosynthesis (DBT) [7,26].

Among current reconstruction techniques, both analytical reconstruction and iterative methods [13] are widely used. One classical analytical reconstruction technique is Filtered Back Projection (FBP) [23] based on Fourier Slice Theorem, which can yield a precise signal reconstruction at a sampling rate satisfying Nyquist–Shannon Theorem, but it may induce reconstruction error from highly incomplete frequency information [5]. Several revised versions of FBP such as adding post-processing filters and interpolating frequency information by a Total-Variation (TV) framework [6] were proposed. One of the iterative methods is Simultaneous Algebraic Reconstruction Technique (SART) developed based on the ART-type (Algebraic Reconstruction Technique type) method in [4,3]. SART actually applies a Ordered Subsets (OS) method to solve a unweighted least square model, which may lead to over-fitting to the noise data and non-convergence to the optimal value. Signal statistics in X-ray Computed Tomography (CT) follows Poisson distribution for mono-energetic CT and compound Poisson for polyenergetic CT [24,9]. Reconstruction methods such as Maximum Likelihood (ML), Penalized Weighted Least Squares (PWLS) and Penalized Likelihood (PL) with Poisson model were strongly proposed and well studied in [16–18,10,13,9,8]. The main benefit from these methods is that the missing data in highly incomplete sampling could be "guessed" at the maximum probability according to the observation. However, the computational intensity attacks statistical methods. Some accelerated strategies successfully speed up the convergence, such as relaxed Ordered Subsets (OS) [15,11], Transmission Incremental Optimization Transfer (TRIOT) [2] and recent Alternating Direction Method of Multipliers (ADMM) [20].

In the classical model of X-ray imaging, the Poisson distribution of incident photon number dominates the physical process. Although X-ray detectors are not quanta counters, Poisson distribution still confirms the signal statistics of mono-energetic X-ray

* Corresponding author at: Department of Electrical and Computer Engineering, Southern Illinois University Carbondale, IL 62901, USA.
E-mail address: adachen@siu.edu (Y. Chen).

0895-6111/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compmedimag.2013.04.005

detection [9]. The number of photons generated and ultimately detected along a projection follows Poisson distribution which can be described mathematically as

$$P(Y_i = y_i) = \frac{\theta_i^{y_i} e^{-\theta_i}}{y_i!}, \tag{1}$$

where Y_i is a random variable counting the observed photons on the detector along ith X-ray; y_i is the observation of Y_i; θ_i is the expectation value of the random variable Y_i. In the classical physical model, θ_i can be expressed as

$$\theta_i = d_i e^{-\langle \mu, l_i \rangle}, \tag{2}$$

where d_i is the intensity of the incident X-ray beam; μ is a linear attenuation coefficient vector to be estimated. Each voxel is assigned an attenuation coefficient and the l_i denotes the vector of the intersection length between the ih X-ray and each voxel. The negative log-likelihood function of all observed photons on the detector can be written as [17]

$$L(\mu) = \sum_{i}^{M} \{d_i e^{-\langle \mu, l_i \rangle} + y_i \langle \mu, l_i \rangle\} + c, \tag{3}$$

by the assumption that $\{Y_i\}_{i \in [1,M]}$ are *i.i.d*, where c is a constant and M is the number of X-ray beams. Through minimizing (3), the optimal μ can be estimated.

In transmission tomography, Compton scattering which makes X-ray photon deflected from its original path could yield a photon noise randomly adding on the detector. The reconstructed results may over fit the noise data at the convergence. Xu and Chen [25] proposed a novel reconstruction method aiming to suppress the Compton scattering. PL method suggesting to insert a penalty function was also proposed. One can revise (3) by appending a scaled penalty function to the likelihood function. The negative log PL function can be written as

$$\Phi(\mu) = L(\mu) + \lambda R(\mu), \tag{4}$$

where the smoothing parameter λ controls the strength of the penalty function. By minimizing (4), the optimal μ is estimated. Generally speaking, to solve the optimization directly is intractable. But, by using the method in [18] or the separable parabolic surrogate in [10,11], the optimal μ in $\Phi(\mu)$ can be approached monotonically with optimal solutions of surrogate functions $h_n(\mu, \mu_n)$, each of which is bounded by $\Phi(\mu)$ at μ_n.

The main obstacle to apply PL method into the application is the unpredictable effect of the smoothing parameter on resolution properties. But thanks to the discussion in [14,22], the authors proposed a modified quadratic penalty function to eliminate the data-dependent terms in impulse response and noise, such that the effect of the smoothing parameter on resolution properties can be evaluated in advance by studying simulated data. Along with the spirit of the research, we present a simplified version of the modified penalty, which is Pre-computed Back Projection based PL (PPL). A two-step procedure is proposed to perform 3D reconstructions along with the desired resolution properties. To solve the optimization, relaxed OS separable parabolic surrogate (OS-SPS) algorithm is applied and forms our relaxed OS-PPL. But the undetermined parameters such as relaxation and subsets make the practical resolution properties deviated from the theoretical one. To conquer it, a training set based semi-quantitative evaluation is presented, by which the parameters in OS are tuned to make the results as consistent as possible with the theoretical with less computational cost.

2. Method for characterizing the smoothing parameter λ

In (4), the penalty $R(\mu)$ can take a general form

$$R(\mu) = \sum_{j=1}^{N} \sum_{k \in N_j} \psi(\mu_j - \mu_k), \tag{5}$$

where N_j is the neighbors of the jth voxel. The function $\psi(t)$ denotes the spatial constraint for adjacent voxels. For a quadratic penalty, ψ can be formulized as follows

$$\psi(\mu_j - \mu_k) = \frac{1}{2}(\mu_j - \mu_k)^2, \tag{6}$$

which results in a consistent smoothing on adjacent voxels. Through minimizing (4), the optimal estimation of μ can be shown in the following form

$$u^* = \underset{\mu \geq 0}{argmin} \Phi(\mu). \tag{7}$$

It is intractable to solve it directly. However, SPS introduced by [10,11] leads to an iterative solution, which is parallel, monotonic, but suffers slow convergence. The basic idea of SPS is that by constructing a series of separable parabolic functions lower bounded with the objective function, the optimal value can be approached by the solution of the surrogate one at each iteration. By applying SPS on (4) with the quadratic penalty (5), the approximation of (7) at the $(n+1)$th iteration can be written as

$$\mu_j^{(n+1)} = \mu_j^{(n)} - \frac{\left[\sum_{i=1}^{M} l_{ij}(-d_i e^{-\langle \mu^{(n)}, l_i \rangle} + y_i) + \lambda \sum_{k \in N_j} (\mu_j^{(n)} - \mu_k^{(n)})\right]}{\left[\sum_{i=1}^{M} \left(l_{ij} \sum_{j=1}^{N} l_{ij} d_i e^{-\langle \mu^{(n)}, l_i \rangle}\right) + 2\lambda |N_j|\right]}, \tag{8}$$

where $|N_j|$ is the cardinality of the subset N_j. The solution sequence of the surrogates converges to the optimal value of the objective function monotonically. Compare to a large curvature of the surrogate function, a small one can yield a faster convergence with “bounded” condition. By replacing $d_i e^{-\langle \mu^{(n)}, li \rangle}$ in the denominator of (8) as y_i, a precomputed curvature in [10] is conceived, which may lead to a faster convergence, yet “almost always” monotonic decreasing.

In practical application, to find a proper smoothing parameter λ in (8) is not trivial. The main reason is that the impulse response and the noise from PL reconstruction are data-sensitive, which means small difference in datasets will yield huge difference on resolution properties, such that the effect of λ is unpredictable. To reduce the data dependence, Fessler et al. proposed a modified penalty function [14] and demonstrated that the impulse response of the reconstructed results is only dominated by λ. The modified penalty is written as

$$R_m(\mu) = \sum_{j=1}^{N} \kappa_j \sum_{k \in N_j} \omega_{jk} \kappa_k \psi(\mu_j - \mu_k), \tag{9}$$

where ω is a weighted coefficient assigned to ψ. κ_j is formulized for emission tomography as follows:

$$\kappa_j = s_j \sqrt{\frac{\sum_{i=1} g_{ij}^2 q_i}{\sum_{i=1} g_{ij}^2}}, \tag{10}$$