Published in final edited form as:

IEEE Trans Med Imaging. 2016 September ; 35(9): 2005–2014. doi:10.1109/TMI.2016.2543141.

Mixed Confidence Estimation for Iterative CT Reconstruction

David S. Perlmutter, Soo Mee Kim, Paul E. Kinahan [Fellow, IEEE], and Adam M. Alessio [Senior Member, IEEE]

Imaging Research Lab, University of Washington, Seattle, WA 98195 USA

Abstract

Dynamic (4D) CT imaging is used in a variety of applications, but the two major drawbacks of the technique are its increased radiation dose and longer reconstruction time. Here we present a statistical analysis of our previously proposed Mixed Confidence Estimation (MCE) method that addresses both these issues. This method, where framed iterative reconstruction is only performed on the dynamic regions of each frame while static regions are fixed across frames to a composite image, was proposed to reduce computation time. In this work, we generalize the previous method to describe any application where a portion of the image is known with higher confidence (static, composite, lower-frequency content, etc.) and a portion of the image is known with lower confidence (dynamic, targeted, etc). We show that by splitting the image space into higher and lower confidence components, MCE can lower the estimator variance in both regions compared to conventional reconstruction. We present a theoretical argument for this reduction in estimator variance and verify this argument with proof-of-principle simulations. We also propose a fast approximation of the variance of images reconstructed with MCE and confirm that this approximation is accurate compared to analytic calculations of and multi-realization image variance. This MCE method requires less computation time and provides reduced image variance for imaging scenarios where portions of the image are known with more certainty than others allowing for potentially reduced radiation dose and/or improved dynamic imaging.

Index Terms

Dynamic imaging; iterative image reconstruction; statistical estimation; tomographic reconstruction

I. Introduction

Four dimensional computed tomography (4D CT) imaging is a useful technique for numerous medical imaging applications. The addition of temporal information provides added value in dynamic contrast enhanced exams and gated acquisitions for cardiac and lung function [1]–[3]. Multi-frame CT acquisitions require more x-ray flux and therefore more radiation dose than single-frame CT. Iterative image reconstruction for CT has the potential to provide improved signal to noise properties, enabling lower radiation dose CT

Personal use is permitted, but republication/redistribution requires IEEE permission.

Correspondence to: Adam M. Alessio.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

acquisitions [4]–[6]. A primary challenge for iterative methods in the context of 4D CT is the computation demand of reconstructing multiple high-voxel density image volumes from large 4D projection data sets. This work presents a technique that allows for reduced computation, and we demonstrate that it additionally offers improved performance compared to conventional frame-by-frame reconstruction.

Some specific CT imaging applications can have the property that certain regions in the image are known with more confidence than other regions. This is often the case in temporal acquisitions where specific regions of the field of view may be unchanging over time, allowing for information averaging across time. For example, in cardiac gated acquisitions, large regions of the field of view remain essentially static while only the cardiac region experiences temporal changes. Likewise, in dynamic contrast enhanced acquisitions, only vascular territories have varying contrast kinetics.

A common technique to suppress noise in dynamic reconstruction is to share information across time. There are numerous approaches designed for 4D imaging that make use of composite information, formed from averaging across time frames. Supanich *et al.* presented a CT formulation of the highly constrained back projection (HYPR) algorithm [7], which uses all 4D data to generate a low-noise composite image (with no dynamic information) and a low-resolution time-varying weighting image [8]. The final image is the product of the composite and weight images. This approach uses analytic reconstruction and has no model for noise in the data. Similarly, Hsieh *et al.* proposed a method that uses fractional scans for the temporal image of interest and completes the missing data with “composite” information from neighboring temporal windows [9]. Finally, Chen *et al.* used the composite image as a second term in a constrained total variation minimization algorithm in their PICCs algorithm [10]. The initial PICCs method has no explicit model of the noise content in each projection, essentially assuming noise-free data and weighting all projections equally regardless of noise content. This approach was later extended by Lauzier *et al.* to include a statistical model [11] to incorporate the heteroscedastic statistics of the tomographic problem. In general, these approaches incorporate side-information about the current reconstructed image of interest and are similar in nature to all methods that constrain or regularize the estimated image based on prior images [12], [13].

In our previous work [14], [15], we proposed the Mixed-Confidence Estimation (MCE) algorithm as a method to reduce computation time in 4D CT image reconstruction by devoting full computational resources to only the dynamic region of interest. ‘Mixed confidence’ refers to the fact in some dynamic CT acquisitions, much of the image is truly static throughout data acquisition, leading to highly redundant data. This redundant information about the static portion of an image should enable it to be estimated with higher confidence than the kinetic portion. In short, MCE updates only the time-varying elements of the 4D image sequence, while keeping the static elements constant across time. In other words, the dynamic, lower-confidence image regions are constrained by the static, higher-confidence regions.

This study extends our prior work by presenting an analytic argument, based on an estimator variance analysis [16], that MCE also lowers variance throughout the image, including the

dynamic region of interest. In Section II we introduce the CT image reconstruction problem in a statistical framework and discuss statistical performance of the maximum likelihood estimator (MLE). Section III applies this statistical methodology to our MCE algorithm and derives an expression for its mean and variance. Section IV validates these results with two simulations studies that demonstrate the performance of MCE under several imaging scenarios.

II. Background

A. Static Statistical Model

We adopt a standard quadratic approximation to the static, 2D transmission tomography inverse problem [17]. We wish to estimate the x-ray attenuation coefficients of each pixel in an image, $\mathbf{x} = [x_1, \dots, x_N]^T \in \mathbb{R}^N$, from sinogram measurements $\mathbf{p} = [p_1, \dots, p_M]^T \in \mathbb{R}^M$. Each sinogram bin, p_i is modeled as an independent Poisson variable whose mean, $\bar{p}_i$ is exponentially proportional to the linear attenuation of the intervening material according to Beer's law,

$$\bar{p}_i = I_0 e^{-a_i^T \mathbf{x}} \quad (1)$$

where a_i describes the intersection of a particular line integral, i , with the body and I_0 is the initial x-ray flux. Combining (1) with a Poisson log-likelihood function for $\mathbf{P}$ yields a conventional pre-log Poisson model for CT measurements,

$$P(\mathbf{p}|\mathbf{x}) \propto -\sum_i p_i(a_i^T \mathbf{x}) + I_0 e^{a_i^T \mathbf{x}}. \quad (2)$$

Maximizing (2) as a function of $\mathbf{x}$ yields the pre-log maximum likelihood reconstruction. In current practice, the computation time required to maximize this function generally makes it infeasible. Alternatively, it is common to simplify (2) by taking a second order Taylor series expansion of the Poisson likelihood function [17], in which case the problem is approximated as,

$$P(\mathbf{y}|\mathbf{x}) \approx \mathbf{N}(\mathbf{A}\mathbf{x}, \mathbf{W}^{-1}) \quad (3)$$

Where $y_i = -\log(p_i/I_0)$ is the post-log sinogram, $\mathbf{A} = [a_1, \dots, a_M] \in \mathbb{R}^{M \times N}$ is the forward tomographic model, and $\mathbf{W} = \text{diag}(\bar{p}_i)$. The maximum likelihood estimator $\hat{\mathbf{x}}_{ML}$ can be written explicitly as the weighted least squares solution,

$$\hat{\mathbf{x}}_{ML} = \arg \max_{\mathbf{x}} P(\mathbf{y}|\mathbf{x}) = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{y}. \quad (4)$$

The properties of (4) are very well understood. For one, because of its simple, linear form iterative solvers for (4) converge much faster than for the solution to (2). One can also show that $\hat{\mathbf{x}}_{ML}$ is the minimum variance unbiased estimator for (3), and is itself a Gaussian random variable with,

$$E[\hat{\mathbf{x}}_{ML}] = \mu_{\mathbf{x}} = \mathbf{x} \quad (5)$$

$$\text{Cov}[\hat{\mathbf{x}}_{ML}] = \sum_{\mathbf{x}} = (A^T W A)^{-1}. \quad (6)$$

In practice the optimal data weights, $\mathbf{p}$, are unknown and are typically approximated with the plug-in estimate, the measured sinogram $\mathbf{p}$. While this will add some small additional variance not captured in (6), its contribution is generally ignored. Interestingly, (6) can also be derived as a variance approximation to the pre-log reconstruction problem of (2) using a first-order Taylor expansion of the Poisson likelihood function and chain rule [16].

If we choose to add a quadratic prior term, $U(\mathbf{x})$, such as the standard Markov random field penalty [17], to control noise, the solution becomes,

$$\hat{\mathbf{x}}_{MAP} = \arg \max_{\mathbf{x}} P(\mathbf{x}|\mathbf{y}) = (A^T W A + R)^{-1} A^T W \mathbf{y} \quad (7)$$

where R describes the prior term (specifically $r_{ij} = \frac{1}{2} U(x_i, x_j)$). In this case, the reconstruction statistics are,

$$\mu_{\mathbf{x}} = (A^T W A + R)^{-1} A^T W A \mathbf{x} \quad (8)$$

$$\sum_{\mathbf{x}} = (A^T W A + R)^{-1} A^T W A (A^T W A + R)^{-1}. \quad (9)$$

In the following introduction to a fast variance approximation, the approach could apply to estimators with spatially invariant quadratic penalties without any major changes. We will not present the derivation with a penalty term for simplicity of presentation and because these approaches for PET [18] and fan-beam CT [19] have been presented. The following fast variance approach will be applied in Section III-C to the mixed confidence problem.

B. Fast Variance Approximation

While (6) and (9) give a simple expression for estimator variance, calculating $\Sigma_{\mathbf{x}}$ explicitly requires inverting an $N \times N$ matrix ($A^T W A$ in the non-prior case), which is intractable in

practice given the size of CT images. However fast techniques have been developed to approximately calculate Σ_x based on the assumption that $A^T W A$ is approximately a locally shift-invariant operator [18]–[20]. The motivation for these methods lies in the fact that, if $A^T W A$ is truly shift-invariant (i.e., a block-circulant matrix), it can be diagonalized quickly with a Fourier transform. Specifically, suppose F is circularly shift-invariant (i.e., a circulant matrix). Then the matrix-vector product $F x$ encodes a 1D circular convolution between x and F_1 , the first column of F . It can be shown that [21],

$$Q F Q^T = \text{diag}(\mathcal{F}\{F_1\}), \quad (10)$$

where Q is the orthonormal DFT matrix (that is, Qx returns the N -point discrete Fourier transform of x). This 1D example extends to block-circulant matrices where Q and Q^T are the Kronecker form of the 2D or 3D Fourier Transform and its inverse, respectively. Therefore, again assuming shift-invariance of $F = A^T W A$,

$$\Sigma_x = F^{-1} = Q^T Q F^{-1} Q^T Q = Q^T (Q F Q^T)^{-1} Q \quad (11)$$

$$= C \left[\mathcal{F}^{-1} \left\{ \text{diag} \left(\frac{1}{\mathcal{F}\{F_1\}} \right) \right\} \right] \quad (12)$$

where $C[\cdot]$ converts a size N kernel into its $N \times N$ circulant matrix form. Here, the matrix inversion computation time is reduced to essentially two DFTs.

This interpretation of F encourages us to think of each column, F_j , as a convolution kernel associated with the j th pixel in the image. Since $F = \Sigma^{-1}$ (in the unregularized case), this kernel describes how statistical certainty spreads within a pixel neighborhood. In CT, the kernel often has a circularly symmetric, $1/r$ shape due to the geometry of the Radon transform [22]. If F is shift-invariant, each kernel is simply a shifted copy of the last. However in reality F is not shift-invariant, most significantly due to the highly non-uniform weight matrix, W [22]. In other words, the kernels vary significantly across the image to reflect the changing local certainty in the data. For example, photon-starved measurements carry little statistical weight, and so create weak kernel “strength” along their line integral. This leads to asymmetric, hourglass-shaped kernels orientated perpendicular to photon starved line integrals. This phenomenon is clearly spatially-variant. We therefore make a standard assumption that the data weights are slowly varying, so that F is *locally* shift-invariant, to make (12) useful. In other words, F_j is *approximately* equivalent to a shifted copy of F_j for $j \in \mathbb{N}(i)$, a local region of pixel i . For $j \notin \mathbb{N}(i)$, the j th component of F_i , $F_{ij} \approx 0$. In this case we can approximate each column of the covariance matrix by,

$$\sum_{\mathbf{x}_j} = \mathcal{F}^{-1} \left\{ \text{diag} \left(\frac{1}{\mathcal{F}\{F_j\}} \right) \right\}. \quad (13)$$

In practice, we must modify F_j slightly before (13). To ensure local stationarity is valid, we place a Gaussian window on each F_j before computing (in this work, we use a 2D Gaussian window with standard deviation equal to $N/4$). Also, since F must be a positive semi-definite matrix, we would like our local approximation $C[F_j]$ to be so as well. We therefore follow the procedure of [18] to ensure that the Fourier transform is real and positive. In brief, to compute the j th column of F for an $N_1 \times N_2$ voxel image, we shift the image to have the j th voxel in the center at $(N_1/2+1, N_2/2+1)$. Then, we introduce the symmetry: $f(k, m) = \max\{f(k, m), f(N_1 - k + 1, N_2 - m + 1)\}$ and apply the Gaussian window. After this step, we take the 2-D FFT of the resulting symmetric image and truncate negative values to zero.

III. Methods

A. MCE Statistical Model

Returning to the imaging model, we now extend the static problem described in (4) to 4D CT. Here, we want to estimate multiple images over K time frames, or $\mathbf{x} = [\mathbf{x}^1, \dots, \mathbf{x}^K]^T$ from $\mathbf{y} = [\mathbf{y}^1, \dots, \mathbf{y}^K]^T$, where each $\mathbf{x}^k \in \mathbb{R}^N$, $\mathbf{y}^k \in \mathbb{R}^M$. A straightforward approach is to treat each frame as a separate static estimation problem, i.e.,

$$\hat{\mathbf{x}}_{ML}^k = (A^T W^k A)^{-1} A^T W^k \mathbf{y}^k \quad (14a)$$

$$\mu_{\hat{\mathbf{x}}_{ML}}^k = \mathbf{x}^k \sum_{\hat{\mathbf{x}}_{ML}}^k = (A^T W^k A)^{-1} \quad (14b)$$

but this solution ignores potentially useful information about $\mathbf{x}^k$ in $\mathbf{y}^{-k}$, where the $-k$ superscript indicates all frames besides k .

To improve on this, we separate the parameter space into a *static* and *dynamic* component, namely, $\mathbf{x}^k = [\mathbf{x}_s, \mathbf{x}_d^k]^T$. It is assumed that $\mathbf{x}_s$ is constant across all frames, and $\mathbf{x}_d^k$ constant only across frame k . This assumption allows us to factor the posterior distribution as,

$$P(\mathbf{x}^k | \mathbf{y}) = P(\mathbf{x}_s, \mathbf{x}_d^k | \mathbf{y}^k, \mathbf{y}^{-k}) \quad (15)$$

$$= P(\mathbf{x}_d^k | \mathbf{y}^k, \mathbf{x}_s) P(\mathbf{x}_s | \mathbf{y}). \quad (16)$$

The first term in (16) is the conditional of (3), and so is itself a multivariate Gaussian variable. The second term is a marginal of the composite image, which is Gaussian distributed if the composite is Gaussian distributed. Although this may not be true for arbitrary composite images, we restrict the composite image to unbiased linear estimators of (3), which we call $\hat{\mathbf{x}}^c$ with covariance $\sum_{\mathbf{x}}^c$. Thus, if:

$$\sum_{\mathbf{x}} = \begin{bmatrix} \sum_{ss} & \sum_{sd} \\ \sum_{ds} & \sum_{dd} \end{bmatrix}$$

then,

$$P(\mathbf{x}_s|\mathbf{y}) = N(\mathbf{x}_s, \sum_{ss}^c) \quad (17a)$$

$$P(\mathbf{x}_d^k|\mathbf{y}^k, \mathbf{x}_s) = N\left(\mu_{d|s}^k, \sum_{d|s}^k\right) \quad (17b)$$

$$\mu_{d|s}^k = \mu_{\mathbf{x}_d}^k + \sum_{ds}^k (\sum_{ss}^k)^{-1} (\hat{\mathbf{x}}^c - \mathbf{x}_s) \quad (17c)$$

$$\sum_{d|s}^k = \sum_{dd}^k - \sum_{ds}^k (\sum_{ss}^k)^{-1} \sum_{sd}^k \quad (17d)$$

where (17a) come directly from the marginal distribution and (17b) to (17d) from the conditional distribution of a multivariate Gaussian. Notice that (17b) depends on $\mathbf{x}_s$ only linearly through its mean via (17c). This ensures that the product of (17a) and (17b) is also Gaussian, parameterized by

$$\mu_{\mathbf{x}}^k = [\mathbf{x}_s, \mathbf{x}_d^k]^T \quad (18a)$$

$$\sum_{\mathbf{x}}^k = \begin{bmatrix} \sum_{ss}^c & \sum_{ds}^k (\sum_{ss}^k)^{-1} \sum_{ss}^c \\ \sum_{ss}^c (\sum_{ss}^k)^{-1} \sum_{sd}^k & \sum_{d|s}^k + \sum_{ds}^k (\sum_{ss}^k)^{-1} \sum_{ss}^c (\sum_{ss}^k)^{-1} \sum_{sd}^k \end{bmatrix}. \quad (18b)$$

Equation (18) is the expression for the statistical performance of our Mixed Confidence estimator, and is the main theoretical result of this paper. Notice that $\hat{\mathbf{x}}_{MCE} = \mu_{\mathbf{x}}^k$ is an unbiased estimate, as might be expected (this, however, would not be true if we included a prior term). Next we compare the MCE covariance with (14). The static pixel variance under MCE is simply $\sum_{ss}^c$. A reasonable choice for the composite image would be the average of each frame estimate, $\mathbf{x}^k$, in which case $\sum_{\mathbf{x}}^c \approx (1/K) \sum_{\mathbf{x}}^k$. Finally, we note that the dynamic pixel variance (lower right term of (18b)), is the sum of two terms; $\sum_{dd}^k$, and a correction factor, Σ^* , where,

$$\Sigma^* = \sum_{ds}^k (\sum_{ss}^k)^{-1} \sum_{ss}^c (\sum_{ss}^k)^{-1} \sum_{sd}^k - \sum_{ds}^k (\sum_{ss}^k)^{-1} \sum_{sd}^k.$$

In our application, we choose a composite image such that the covariance of the static region of the composite image, $\sum_{ss}^c$ is less than or equal to the covariance of the static region of a single frame estimate, $\sum_{ss}^k$. Therefore, $\Sigma^* \leq 0$ and the covariance of the dynamic portion of the MCE estimate (lower right term of (18b)) is less than or equal to the covariance of the dynamic portion of a conventional single frame estimate, $\sum_{dd}^k$.

To better understand (18b), consider two cases; case 1: $\sum_{ss}^c = 0$ and case 2: $\sum_{ss}^c = \sum_{ss}^k$. In the first case, the full dataset $\mathbf{y}$ provides perfect information about $\mathbf{x}_s$, and the covariance of the dynamic portion of the image $\mathbf{x}_d^k$ (lower right quadrant of (18b)) reduces to the conditional covariance $\sum_{d|s}^k$, which is guaranteed to be smaller than $\sum_{dd}^k$. In the second case, the covariance of the static portion of the composite image equals the covariance of the static portion of a single frame; i.e., the composite image did not improve the estimate of $\mathbf{x}_s$. Then $\Sigma^* = 0$, and the covariance is simply $\sum_{dd}^k$. In practice, we expect to lie somewhere in between these two bounds; the composite provides some extra, but not perfect, information about $\mathbf{x}_s$, which helps lower the covariance of *both* dynamic and static pixel estimates.

B. Implementation of MCE

MCE applied to dynamic CT data comprises the following general steps (specific details and variants will be discussed in Section IV):

1. Classify each image element as either *static* or *dynamic* across frames.
2. Form a low-noise, low-temporal resolution “composite image” from the entire data set to initialize all frames.
3. Reconstruct each frame, updating only the dynamic pixels.

Previous work on MCE [14] suggests practical ways to classify the image elements as static or dynamic. For example, a variance image could be computed across frames and image

elements with variance below a certain threshold could be deemed static and those above the threshold could be deemed dynamic. Alternatively, in certain applications, it is known a priori which regions of the image will change across time. For example, in dynamic contrast enhanced cardiac imaging, the heart region has rapidly changing contrast across frames and the extra-cardiac region is relatively constant; These regions could be classified as dynamic and static respectively.

In our application, we reconstruct a composite image from the average of the post-log

sinograms across frames, $\bar{y} = \sum_{k=1}^K y^k$. This provides a reconstruction of the static component, $\mathbf{x}_s$, and a reasonable initial image for the dynamic component, $\mathbf{x}_d$. Finally, for each frame, a convergent iterative reconstruction is applied to the dynamic component voxels while the static component is kept constant. The resulting computation reduction scales linearly with the percentage of dynamic pixels, minus the time to form the composite image.

C. Fast Mixed Confidence Variance Estimation

To calculate the variance reduction of mixed confidence estimation for frame k we adopt the fast FFT variance approximation technique described above. First, we form each F_i explicitly by forward projecting a shifted 2D impulse function, multiply the resulting sinogram by the measured data weights, W^k , then backproject. We then use the variance

approximation method described above to calculate $\sum_{\mathbf{x}_{ML}}^k$. We repeat the process to form

$\sum_{\mathbf{x}}^c$ although it is possible to approximate $\sum_{\mathbf{x}}^c$ as either $(1/K) \sum_{\mathbf{x}_{ML}}^k$ or $(1/K^2) \sum_k \sum_{\mathbf{x}_{ML}}^k$ for simplicity. We could partition these matrices and calculate (18b)

directly. However, since the static region is typically large and larger than the dynamic region, inverting $\sum_{ss}^k$ is computationally demanding. Instead, we substitute the matrix equality

$$\left(\sum_{ss}\right)^{-1} = F_{ss} - F_{ds} F_{dd}^{-1} F_{sd} \quad (19)$$

into (18b), where $F_* = \sum_*^{-1}$ as before. This requires inverting the smaller matrix F_{dd} .

IV. Evaluation Studies

To validate our analysis we conducted two simulation studies comparing MCE to standard framed reconstruction. In this work, all reconstructions were run to convergence defined as having reached image updates with a maximum pixel change of less than epsilon, where epsilon was 1% of the smallest non-zero value in the image.

Proof-of-Concept Simulation

The first study compared our predicted covariance with those measured from simulations of a 20×20 pixel proof-of-concept image, which is shown in the upper left of Figs. 1 and 2. We used a monoenergetic, distance driven fan-beam CT projector [23] to simulate a 35 detector $\times$ 38 view sinogram to make explicit computation of the WLS estimate (4) tractable. We added Gaussian noise to the sinogram data according to the approximate signal model in (3). The noise level was chosen to match a “ultra-low” dose flux regime, in which the mean/minimum sinogram bin measurements were 50 and 5 detected photon counts, respectively. While there is no motion in the true image, we nonetheless treat the inner pixels (yellow pixels in Fig. 1, upper right) as “dynamic”, and collected $K = 3$ frames of data. We then estimated the composite image $\hat{\mathbf{x}}_c$ and framed image $\hat{\mathbf{x}}_{ML}^1$ using the ML estimator (4) and $\hat{\mathbf{x}}_{MCE}$ using the MCE method. The choice of composite image is important, and must balance low-noise and computation time. We estimate $\mathbf{x}^c$ using the framed-averaged data set, $(1/K) \sum_k \mathbf{y}_k$.

When using a prior, the prior strength of the composite, framed, and MCE reconstructions was identical. Therefore while the composite data is less noisy, its statistical weight, relative to regularization, is the same as in framed recon. All reconstruction used the post-log separable quadratic surrogates (SQS) algorithm [24], and the data weights were drawn from the noiseless, pre-log sinogram so as to conform with (14). The lower plots of Fig. 1 show an example of both the framed and MCE reconstructions.

We perform data generation and image reconstruction for 1000 independent noise realizations in order to calculate the sample variance of reconstructed images. This sample variance is compared with the the predicted variance map of both framed and MCE recon from (14) and (18), respectively, and the FFT variance approximation of both methods with (12) and (19).

A. Dynamic Cardiac CT Simulation

In the second study we used a 12 frame dynamic cardiac CT patient perfusion image volume to generate sinogram data. This image volume was acquired from an end-diastolic cardiac gated acquisition across 25 seconds (50 cm transaxial field of view, 512×512 pixel image). These images show the iodine-base contrast agent passing from the right to the left ventricle of the heart, with a small amount of gross body motion. The dynamic region (Fig. 3) was conservatively selected by hand to enclose the entire heart across all frames. We projected each frame of the image into a 222 detector $\times$ 246 view sinogram and added pre-log Poisson noise. We again held the minimum pre-log sinogram bin to 5 detected photons. The mean across the entire sinogram was ~ 450 photon counts. The post-log SQS algorithm was again used to compute 128×128 reconstructions, but this time the noisy plug-in estimate was used for the weights. To quantify reconstruction image quality, we calculated two metrics: the FFT variance approximation, and the mean squared error (MSE) from a noiseless 128×128 reconstruction using the same algorithm.

V. Results

Proof-of-Concept Simulation Results

The second row of Figs. 1 and 2 show the reconstructed images from conventional framed reconstruction and MCE reconstruction for two different selections of dynamic regions. The variance images for the conventional and MCE reconstructions are presented in Figs. 4 and 5. These figures show the variance image a) analytically calculated from direct expressions, b) from the sample variance over 1000 noise realizations, and c) from the FFT approximation of Section III-C. Both the analytical expression calculation and the FFT approximation of the magnitude and distribution of the variance matches well with the sample variance over 1000 noise realizations.

Fig. 6 presents a profile through the center of the variance images in Fig. 4. As expected, the composite image has a variance that is K times lower than the conventional framed reconstruction. The MCE reconstruction uses the static portion from the composite image (marked as pixels 1–7 and 14–20 in this profile plot), so it has identical values as the composite image in the static region. In the dynamic region, the MCE reconstruction yields a $\sim 10\%$ reduction in variance compared to framed reconstruction (central region; pixels 8–13 in profile) in this proof-of-concept example.

Fig. 7 shows how these relative variance gains trend as a function of different reconstruction parameters. As the number of frames increases, the top plots shows the framed reconstruction is constant since it doesn't share information across frames, while the composite image shows the typical $1/K$ variance reduction. The MCE variance initially

decreases rapidly, then plateaus to the conditional variance, $\sum_{dd'}^k$

The size and shape of the dynamic region is also an important factor determining MCE performance. The variance reduction column of Fig. 4 shows that higher performance gains are achieved in those dynamic pixels adjacent to static regions. This is further exemplified in Fig. 5, in which the dynamic region is replaced with three smaller ROIs, each of which achieve a higher variance reduction. Intuition would suggest that potential MCE variance reduction is inversely proportional to dynamic region size; The middle plot of Fig. 7 demonstrates this by changing the radius of the centered dynamic region of Fig. 1. Finally, Fig. 7 also shows that MCE continues to reduce variance over framed reconstruction even when applied with a moderate strength quadratic MRF prior.

Table I summarizes the variance estimates for the different reconstruction techniques and static regions. This small simulation image allowed direct comparison of a multi-realization estimate of variance (“sample”, truth for this study), our proposed expression for the variance in MCE (“analytic”, (18b)), and our fast variance calculation using FFT ((19)). The three columns: Sample, Analytic, FFT, in Table I present these estimates and demonstrate that our proposed expression for variance in MCE is essentially identical to the sample variance. And, this table shows that the fast variance method leads to equally accurate variance estimates supporting this approach as a pragmatic method to potentially estimate noise in resulting MCE images.

A. Dynamic Cardiac CT Experimental Results

Fig. 8 shows both framed and MCE image reconstructions of the dynamic phantom experiment. The “truth” images are SQS image reconstruction using noise free data and no data weights. While the MCE images are clearly less noisy in the static region, the visual appearance of the dynamic region of both reconstructions is very similar. To quantify this performance, the normalized image MSE averaged over the dynamic region only is presented above each plot. There is a 10%–30% MSE reduction in the dynamic region using MCE. Fig. 9 shows an image of the first frame MSE and FFT variance approximation for both reconstructions, and Fig. 10 shows a profile through the variance approximation. Table II quantifies the framed-averaged performance. The mean squared error (MSE) image shows a bias in the static region of MCE due to patient motion which is not present in the framed images. However, visual assessment and MSE suggest that this error does not propagate into the dynamic region.

In this example, the conventional framed reconstruction required 25 iterations to reach convergence for each of the 12 frames. The MCE reconstruction focused on the dynamic region that was 15% of the area of the full field of view. The MCE updates of the dynamic region required 15 iterations to reach convergence for each frame and the total computation time was 12% of the framed reconstruction time. In our non-optimized implementation on a multi-core Linux system, the MCE reconstruction completed in 6 minutes and the framed reconstruction completed in 50 minutes.

VI. Discussion

We show through statistical analysis and simulation that MCE leverages the inherent redundancy of dynamic data such that both computation time and image noise, which can be traded for reduced radiation dose, grow sub-linearly with the number of frames. MCE significantly decreases the size of the estimation problem in each frame, by a factors of 2-20, depending on the dynamics being imaged. This leads to an equivalent decrease in the projector operation complexity, the main bottleneck in all statistical reconstruction algorithms. For most applications, these computational savings easily offset the additional cost of computing the composite image. Also, for a fixed optimization algorithm, fewer unknowns generally require fewer iterations to converge, further reducing computation time. Previous work [14] showed two applications, retrospective gated CT angiography and dynamic perfusion CT, in which MCE provided similar image quality to conventional maximum likelihood reconstruction with ~50% dynamic pixels, and therefore ~50% compute time.

Beyond computational considerations, exploiting the redundancy of dynamic data using MCE reduces noise throughout the image in both the static and dynamic regions. This noise reduction was shown to be theoretically unbiased by analyzing the underlying statistical assumptions of the MCE model using the standard Gaussian noise model approximation to post-log CT noise. Likewise, many optimization algorithms have been proposed to solve the quadratic post-log reconstruction problem [17], [24] and MCE and the variance reduction analysis presented here would pertain to all such methods as long as they are run to convergence. In summary, in most 4D applications, the primary goal is to generate high-

quality images in the dynamic regions; MCE offers an approach that is computationally faster and has reduced variance in dynamic regions compared to conventional frame-by-frame reconstruction.

Fig. 7 presents anticipated trends that the variance reduction with MCE compared to framed reconstruction is diminished when the total number of frames is smaller, when the portion of dynamic pixels compared to static pixels increases, and as the penalty strength increases. The last row of Fig. 7 shows the general trend of variance with increasing quadratic penalty strength. This comparison is limited in that we used a fixed prior strength across all methods even though the relative weighting of the data fidelity term and penalty term has different values for each method (there is a different number and magnitude of data elements and image elements between composite, framed, and the MCE method). A thorough performance comparison for different penalty terms would require a more thorough analysis. This does show the expected trend that in the limit as the penalty strength approaches infinity, the variance reduction of MCE compared to framed reconstruction become smaller.

We used a small Monte Carlo simulation to show the the amount of MCE variance suppression depends on the size and shape of the dynamic region, the number of temporal frames, and strength of regularization(Fig. 7). The effects on these factors generally follow the intuitive trend that information sharing across the CT image is generally localized due to the nature of the Radon transform and that a higher confidence static region can propagate noise reduction only up to the intrinsic, conditional covariance limit. In other words, where the covariance between the static and dynamic regions, Σ_{ds} , is large, the variance reduction is greater; This leads to more relative variance reduction along the boundaries of the dynamic region.

We also measured the performance of MCE in terms of mean square error and variance reduction in a patient-derived dynamic contrast enhanced cardiac CT simulation. While noise reduction of MCE is only visually appreciable in the static region, the variance estimates demonstrate an approximately 5% reduction in variance and 20% reduction in MSE on average across the dynamic region in this scenario. The clinical significance of this reduction will depend on the clinical task and in most cases this minor reduction will likely not provide substantial performance gains. That said, the benefit of less computation remains and would allow for faster image generation and the potential to devote the saved computation to more physically realistic system models or objective functions [25], [26]. This approach might have clinical value in dynamic CT perfusion studies where A) computation is challenging due to large data set sizes from multiple frames, B) noise is a issue due to the need for low-dose multi-frame acquisitions, and C) quantitative accuracy and precision are critical for kinetic model analysis [27], [28].

The use of the fast-variance approximation, along with demonstrating the value of the MCE algorithm, could be used in dynamic data processing as weighting terms for kinetic model fitting [29]. Nonetheless, the FFT variance approximation depends strongly on the local space-invariance assumption, which is not equally valid across the entire image. For example, since the DFT is inherently circular, local stationarity is a poor approximation at

the periphery of the image. Also any location where F_j is quickly varying, such as on the boundary of dense regions, will lead to poorer variance estimation with this approximation.

The MCE algorithm was applied here in a relatively simple approach that fixes pixel values in a high-confidence regions and updates only the lower-confidence regions. This algorithm could be extended to other applications with multiple, varying degrees of confidence (not just binary high and low). For example, a cardiac gated and dynamic-contrast CT acquisition could have image regions a) shared across all gates and frames (high confidence); b) shared only across dynamic frames (several degrees of medium confidence); and c) unique to all gates and frames (low confidence). In this application, the higher-confidence regions could be applied with varying strength to constrain the lower-confidence regions. The argument presented in our work for reduced variance in the lower-confidence regions would be applicable to these expanded applications.

VII. Conclusion

We have presented an iterative method for reconstructing dynamic CT images that reduces both computation time and image noise throughout the image over conventional, frame-by-frame reconstruction. These benefits are achieved by segmenting the image into distinct regions that can be estimated with varying levels of statistical confidence. Reconstructing each confidence region separately allows us to exploit the inherent redundancy in dynamic CT by sharing information across frames. This sharing is done in a simple, efficient manner that reduces computation time that scales with the size of the dynamic region. We developed a fast variance approximation for the MCE problem and demonstrated through simulation studies that this approximation is accurate compared to analytic expressions for variance and multi-realization sample variance. The MCE method leads to a reduction in variance in the dynamic region on the order of 1–20% depending on the size and shape of the dynamic region. The MCE method could be used to provide faster iterative image reconstruction of dynamic data with better noise performance in the kinetic region of interest compared to conventional frame-by-frame reconstruction allowing for improved dynamic imaging and radiation dose reduction.

Acknowledgments

The authors thank Drs. P. La Riviere and B. De Man for helpful conversations.

This work is supported by the National Institutes of Health under grants R01-HL109327 and R01-CA115870.

References

1. Miles K, Griffiths M. Perfusion CT: A worthwhile enhancement? *Br J Radiol.* 2003; 76(904):220–231. [PubMed: 12711641]
2. Ding K, Bayouth JE, Buatti JM, Christensen GE, Reinhardt JM. 4DCT-based measurement of changes in pulmonary function following a course of radiation therapy. *Med Phys.* 2010; 37(3): 1261–1272. [PubMed: 20384264]
3. Yamamuro M, et al. Cardiac functional analysis with multidetector row CT and segmental reconstruction algorithm: Comparison with echocardiography, SPECT, and MR imaging. *Radiology.* 2005; 234(2):381–390. [PubMed: 15670995]

4. Thibault JB, Sauer KD, Bouman CA, Hsieh J. A three-dimensional statistical approach to improved image quality for multislice helical CT. *Med Phys.* 2007; 34(11):4526–4544. [PubMed: 18072519]
5. De Man B, Nuyts J, Dupont P, Marchal G, Suetens P. An iterative maximum-likelihood polychromatic algorithm for CT. *IEEE Trans Med Imag.* Oct; 2001 20(10):999–1008.
6. Sunnegardh J, Danielsson PE. Regularized iterative weighted filtered backprojection for helical cone-beam CT. *Med Phys.* 2008; 35(9):4173–4185. [PubMed: 18841870]
7. Johnson KM, et al. Improved waveform fidelity using local HYPR reconstruction (HYPR LR). *Magn Reson Med.* 2008; 59(3):456–462. [PubMed: 18306397]
8. Supanich M, et al. Radiation dose reduction in time-resolved CT angiography using highly constrained back projection reconstruction. *Phys Med Biol.* 2009; 54(14):4575–4593. [PubMed: 19567941]
9. Hsieh J, Wei Y, Wang G. Fractional scan algorithms for low-dose perfusion CT. *Med Phys.* 2004; 31(5):1254–1257. [PubMed: 15191317]
10. Chen GH, Tang J, Leng S. Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med Phys.* 2008; 35(2):660–663. [PubMed: 18383687]
11. Lauzier PT, Chen GH. Characterization of statistical prior image constrained compressed sensing. I. Applications to time-resolved contrast-enhanced CT. *Med Phys.* 2012; 39(10):5930–5948. [PubMed: 23039632]
12. Stayman JW, Zbijewski W, Otake Y, Uneri A, Schafer S, Lee J, Prince JL, Siewerdsen JH. Penalized-likelihood reconstruction for sparse data acquisitions with unregistered prior images and compressed sensing penalties. *Proc SPIE.* 2011; 7961:79611L–79611L–6.
13. Rashed EA, Kudo EA. Statistical image reconstruction from limited projection data with intensity priors. *Phys Med Biol.* 2012; 57(7):2039. [PubMed: 22430037]
14. Alessio A, La Riviere P. Constrain static target kinetic iterative image reconstruction for 4D cardiac CT imaging. *Proc SPIE.* 2011; 7873:78730S–78730S–7.
15. Alessio A, Kinahan P. Statistical reconstruction of targeted regions with application to CT and PET cardiac imaging. *Am J Roentgenol (Suppl).* 2006; 186(4):A46.
16. Fessler J. Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography. *IEEE Trans Image Process.* Mar; 1996 5(3):493–506. [PubMed: 18285134]
17. Sauer K, Bouman C. A local update strategy for iterative reconstruction from projections. *IEEE Trans Signal Process.* Feb; 1993 41(2):534–548.
18. Qi J, Leahy RM. Resolution and noise properties of MAP reconstruction for fully 3-D PET. *IEEE Trans Med Imag.* May; 2000 19(5):493–506.
19. Zhang-O'Connor Y, Fessler JA. Fast predictions of variance images for fan-beam transmission tomography with quadratic regularization. *IEEE Trans Med Imag.* Mar; 2007 26(3):335–346.
20. Fessler JA, Booth SD. Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction. *IEEE Trans Image Process.* May; 1999 8(5):688–699. [PubMed: 18267484]
21. Gray, RM. *Toeplitz and Circulant Matrices: A Review.* Breda, The Netherlands: now publishers; 2006.
22. Fessler JA, Rogers WL. Spatial resolution properties of penalized-likelihood image reconstruction: Space-invariant tomographs. *IEEE Trans Image Process.* Sep; 1996 5(9):1346–1358. [PubMed: 18285223]
23. De Man B, Basu S. Distance-driven projection and backprojection. *IEEE Nucl Sci Symp Conf Rec.* 2002; 3:1477–1480.
24. Erdogan H, Fessler JA. Ordered subsets algorithms for transmission tomography. *Phys Med Biol.* 1999; 44(11):2835. [PubMed: 10588288]
25. Do S, et al. High fidelity system modeling for high quality image reconstruction in clinical CT. *PLoS One.* 2014; 9(11):e111625. [PubMed: 25390888]
26. Lasio GM, Whiting BR, Williamson JF. Statistical reconstruction for x-ray computed tomography using energy-integrating detectors. *Phys Med Biol.* 2007; 52(8):2247. [PubMed: 17404467]

27. Bindschadler M, Modgil D, Branch KR, La Riviere PJ, Alessio AM. Comparison of blood flow models and acquisitions for quantitative myocardial perfusion estimation from dynamic CT. *Phys Med Biol.* 2014; 59(7):1533–56. [PubMed: 24614352]
28. Lee T. Functional CT: Physiological models. *Trends Biotechnol.* 2002; 20(8):S3–S10. [PubMed: 12570152]
29. Dell RB, Sciacca R, Lieberman K, Case DB, Cannon PJ. A weighted least-squares technique for the analysis of kinetic data and its application to the study of renal ¹³³Xenon washout in dogs and man. *Circulat Res.* 1973; 32(1):71–84. [PubMed: 4684130]

Fig. 1.

First proof-of-concept simulation images using a single circular dynamic region. Top, upper left: truth image featuring a circular phantom with three small circular ROIs of varying density. Top, upper right: The dynamic region chosen for MCE. Top, lower row: a single realization of framed and MCE reconstructions from noisy data.

Fig. 2.

Second proof-of-concept simulation study images, which used the same truth image as the first study but three smaller, targeted dynamic regions (multiple disjoint regions). Top, upper left: truth image featuring a circular phantom with three small circular ROIs of varying density. Top, upper right: Three small dynamic regions chosen for MCE. Top, lower row: a single realization of framed and MCE reconstructions from noisy data.

Fig. 3.

Left: Composite image reconstruction the from dynamic cardiac CT simulation, reconstructed using post-log SQS on the frame-averaged sinogram across all 12 frames of data. Right: Manually selected dynamic region, chosen to cover the full extent of heart motion, presented as a binary mask on the composite image.

Fig. 4.

Variance results for the first proof-of-concept simulation study in Fig. 1. Reconstruction variance images using the framed (left) and MCE (center) methods, and the ratio of MCE to framed (right), calculated three different ways: direct analytic calculation using (14), sample variance over 1000 noise realizations, and using the FFT approximation method of Section III-C.

Fig. 5.

Variance results for the second proof-of-concept simulation study in Fig. 2. Reconstruction variance images using the framed (left) and MCE (center) methods, and the ratio of MCE to framed (right), calculated three different ways: direct analytic calculation using (14), sample variance over 1000 noise realizations, and using the FFT approximation method of Section III-C.

Fig. 6. Horizontal profile through the center of the variance images in Fig. 4. There is good agreement between the analytic (solid line), FFT approximation (dashed line), and sample variance (data points) for each of the three reconstruction methods.

Fig. 7.

Average variance in the dynamic region reconstructed with three different methods from the first proof-of-concept simulation as a function of different simulation parameters: changing the number of frames contributing to composite image (top), relative size of the dynamic region compared to entire field of view (middle), and strength of the quadratic MRF prior (bottom).

Fig. 8.

Image reconstructions using the framed (middle) and MCE (bottom) methods for odd-numbered frames of the dynamic cardiac CT experiment. The normalized mean squared error in the entire dynamic region is shown above each figure. The top row shows the “truth” image of each frame, which is an 128×128 reconstruction from noiseless data.

Fig. 9.

Two image quality metrics from image reconstructions of the first frame of the dynamic cardiac CT experiment using both framed (top) and MCE (bottom) methods. Left: Error from truth image. In the static region, the MCE image shows boundary errors due to gross patient motion, but an otherwise lower noise level, and slightly lower error in the dynamic region, especially near the static-dynamic boundary. Right: FFT-approximation variance map.

Fig. 10.

Profile through the variance images in Fig. 9 showing the improved variance performance of MCE, particularly near the static-dynamic boundary.

Comparison of Variance Estimates for 20×20 Simulation Study. Demonstrates That Analytic Calculation and FFT Approximation are Close to Sample Variance for Several Variations of Methods

Table I

Dynamic Target	β	Method	Variance Estimate (a.u.)		
			Sample	Analytic	FFT
Circular Region	0	Composite	0.056	0.054	0.054
		Framed	0.162	0.163	0.163
		MCE	0.149	0.148	0.152
	10^3	Composite	0.043	0.043	0.041
		Framed	0.129	0.129	0.127
		MCE	0.121	0.120	0.120
Multiple Disjoint Regions	0	Composite	0.045	0.045	0.043
		Framed	0.135	0.134	0.128
		MCE	0.076	0.075	0.075
	10^3	Composite	0.036	0.036	0.034
		Framed	0.107	0.106	0.103
		MCE	0.097	0.093	0.092

Table II
Comparison of Variance Estimates for Dynamic Cardiac CT Simulation Study in Two
Different Regions of the Heart

Dynamic Target	β	Method	FFT Variance Estimate (a.u.)
Heart Wall	0	Composite	0.0189
		Framed	0.1420
		MCE	0.1343
Ventricle	0	Composite	0.0135
		Framed	0.1866
		MCE	0.1823