IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 7, 2017, accepted September 13, 2017, date of publication September 19, 2017,
date of current version October 25, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2754327

Sparse-Prior-Based Projection Distance Optimization Method for Joint CT-MRI Reconstruction

XUELIN CUI[1], LAMINE MILI[1], (Fellow, IEEE), AND HENGYONG YU[2], (Senior Member, IEEE)
[1]Department of Electrical and Computer Engineering, Virginia Tech, Falls Church, VA 22043, USA
[2]Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA

Corresponding author: Xuelin Cui (xcui@ieee.org)

The work of H. Yu was supported by the NSF CAREER Award CBET under Grant 1540898.

ABSTRACT Multimodal imaging techniques have received a great deal of attention, since their inceptions for achieving an enhanced imaging performance. In this paper, a novel joint reconstruction framework for computed tomography (CT) and magnetic resonance imaging (MRI) is implemented and evaluated. The CT and MRI data sets are synchronously acquired and registered from a hybrid CT-MRI platform. Because the image data sets are highly undersampled, the conventional methods (e.g., analytic reconstructions) are unable to generate decent results. To overcome this drawback, we employ the compressed sensing (CS) sparse priors from an application of discrete gradient transform. On the other hand, to utilize multimodal imaging information, the concept of projection distance is introduced to penalize the large divergence between images from different modalities. During the optimization process, CT and MRI images are alternately updated using the latest information from current iteration. The method exploits the structural similarities between the CT and MRI images to achieve better reconstruction quality. The entire framework is accelerated via the parallel processing techniques implemented on a nVidia M5000M Graph Processing Unit. This results in a significant decrease of the computational time (from hours to minutes). The performance of the proposed approach is demonstrated on a pair of undersampled projections CT and MRI body images. For comparison, the CT and MRI images are also reconstructed by an analytic method, and iterative methods with no exploration of structural similarity, known as independent reconstructions. Results show that the proposed joint reconstruction provides a better image quality than both analytic methods and independent reconstruction by revealing the main features of the true images. It is concluded that the structural similarities and correlations residing in images from different modalities are useful to mutually promote the quality of joint image reconstruction.

INDEX TERMS Multimodal imaging, image reconstruction, compressed sensing.

I. INTRODUCTION

The multimodal imaging techniques have been widely hailed as powerful tools to assist the modern clinical decision-making process. A claim has been made that more structural and/or functional information can be revealed by these techniques than those performing only single modality [1], [2]. Since its inception in 1960's [3], a variety of multimodal imaging commercial products have been developed, such as positron emission tomography and computed tomography (PET-CT), positron emission tomography and magnetic resonance imaging (PET-MRI), and single positron emission computed tomography and computed tomography (SPECT-CT). However, the fusion of CT and MRI has not yet been thoroughly investigated and commercialized. The idea of combining CT and MRI techniques seems so natural since both of them are the most commonly used imaging modalities today, and they are intrinsically complementary to each other in many ways. For example, CT has high contrast for bones, whereas MRI has excellent dynamic range for soft tissues. Recently, the concept of "omni-tomography" has been proposed [4], [5], and one of whose goals is to design and implement a grand fusion of CT and MRI in a single scanning machine, which has not yet been achieved. The hybrid CT-MRI platform will contain hardware and software for both systems. Under this setup, the data from CT and MRI scanners are acquired and registered synchronously

2169-3536 © 2017 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

(''all in one''), and images are reconstructed simultaneously (''all at once''). In this paper, we address the need for the planned grand fusion by proposing a projection distance-based CS method for its CT-MRI joint image reconstruction.

On the targeted hybrid CT-MRI platform, due to machanical and electrical limitations, both CT and MRI data sets are highly undersampled. These undersampled data are collected through synchronous acquisition such that online data registration could be achieved [4], [5]. From the image reconstruction perspective, the enabling technology of omni-tomography is the CS-based image reconstruction [6], [7]. Conventionally, medical imaging systems employ analytic methods such as filtered back-projection [8] for CT and inverse Fourier transform [9] for MRI to reconstruct images. However, to have decent image quality, fully-sampled data sets are required. For example, a typical commercial 64-slice CT scanner normally has about 1000×64 detector channels, and it collects approximately 1000 projections per rotation. Therefore, the sampling density is extremely high such that the Nyquist rate is guaranteed. Initiated about a decade ago, the CS-based methods have undergone a rapid growth, especially in recent years. For instance, Lustig *et al*. [10] showed that CS can be utilized to reconstruct sparse medical images from undersampled measurements. Their work makes use of the CS theory developed by Candes, Tao and Donoho, who proved that given knowledge about a signal's sparsity, the signal may be reconstructed with even fewer samples than the Nyquist-Shannon sampling theorem [11] requires [12]–[14].

Mathematically, a well-known fact in system identification is that undersampling the input and output signals leads to an under-determined mathematical system model. In theory, an undersampled inverse problem requires a regularization instead of direct inverse such that stability can be kept during the problem-solving procedure [15]. Reconstruction from undersampled data with the direct inverse generates strong artifacts in general. Obviously, we need to suppress the artifacts to avoid confusion with the real anatomic structures. The choice of regularization is essentially subjective and mainly depends on the applications. In recent years, the gradient-based methods such as total variation (TV) minimization [6], [7], [10], [16]–[18] have gained great popularity, because of their impressive results in artifacts reduction and image denoising. In many cases, medical images can be well approximated by piece-wise constant functions. A typical example is the large area of air in the CT lung imaging. Therefore, it is reasonable to add TV as a regularization, since in general the gradient-based sparse prior is very important in CS-based medical image reconstruction to suppress the streaks and to reduce noise.

From another perspective, since a multimodal platform samples the same object simultaneously, the images from different modalities must share structural similarities particularly in boundaries and edges. This is achieved despite the fact that CT and MRI images are formed under completely different physical principles. In other words, CT images and MRI images are similar in structural features even though they look quite different in terms of pixel intensity and contrast for the same anatomy. As a result, coupling some structural features between CT and MRI can potentially improve reconstruction quality for both of them. From a signal processing perspective, this is equivalent to utilizing more information to perform a better signal recovery.

The idea of performing joint reconstruction by coupling underlying models that have a common structure emerged 1990's. Haber and Oldenbourg [19] found that the joint inversion approach reduces the non-uniqueness, and may improve the quality of interpretation. However, the application concentrates on geophysical applications. Rigie and Rivire [20] employed a ''total nuclear variation'' to perform joint reconstruction on CT data acquired from different kVp. However, he did not use it for the cross-modal scenario, and thereby he did not utilize any cross-modal information. Ehrhardt *et al.* [21] investigated the PET-MRI joint reconstruction using a parallel level set. However, PET imaging is remarkably different form CT imaging in terms of its physical realization and mathematical approach. In particular, the spatial resolution of PET images is significantly lower than that of CT images [22]. Other reseraches on multimodal reconstruction can also be found in [23]–[25], but ultimately the CT-MRI joint reconstruction remains a challenging research topic. In general, the key to perform joint reconstruction lies in the following two questions. The first one is how to model a multimodal imaging system, whereas the second one is how to create a coupling among different modalities. In fact, the multimodal imaging systems can be reasonably generalized as multichannel imaging systems [20], [26], [27], which have been used in many applications, such as color and geophysics imaging [21]. On the other hand, multimodal coupling has been explored with many geometric-based [28]–[30] and theoretical information-based methods [31], [32]. However, all these methods uses offline information from other modalities. Therefore, none of these priors have actually been fit into the context of the planned omni-tomography system, which requires images to be reconstructed simultaneously.

This research is part of the planned omni-tomography project, and particularly works on the portion of image reconstruction. It requires the development of a new joint CT-MRI reconstruction from highly undersampled CT and MRI data sets. The proposed sparse-prior-based projection distance optimization method updates image signals at each iteration while it incorporates information provided by the the CT and MRI coupling. The methodology is very general, and it can be applied to any problem where images share structural correlations, for instance, dual-energy CT imaging systems. For comparison purpose, analytic reconstruction and individual CS-based reconstructions are also implemented. The latter method does not process any information from the other modality and thereby ignores the similarity and correlated information. Since the omni-tomography is currently an active research topic and has not been yet realized, our data are generated by simulating the clinical CT and MRI images. Numerical simulations reveal that our joint reconstruction

provides a significant improvement in image quality, especially when the structure of the anatomy is more complicated than the piece-wise constant model.

The rest of this paper is organized as follows. Section II provides the theoretical framework and methods of multimodal imaging systems. Section III provides simulation results and analysis. Section IV draws conclusions for this work.

II. THEORY AND METHODS

A. RECONSTRUCTION FRAMEWORK

Both the CT and MRI imaging systems can be modeled as a linear shift-invariant regression model as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \sigma, \tag{1}$$

where $\mathbf{x}$ denotes an unknown image to be reconstructed; $\mathbf{A}$ denotes the system matrix; σ denotes the noise; and $\mathbf{y}$ denotes the collected data. For CT imaging, the noise are dominantly Poisson, and the data are a number of projections, or sinogram, formed by passing a set of X-ray beams through an object from different angles. For MRI, the noise are Gaussian and the data, usually referred to as k-space data, are essentially the Fourier transform of the original image. The system matrix $\mathbf{A}$ in Equation (1) indicates forward projection in CT imaging and Fourier transform in MRI.

As CT and MRI are fused together, the multimodal imaging system is modeled as an N-channeled imaging system [20], [26] whose system matrix is

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1N} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{N1} & \mathbf{A}_{N2} & \dots & \mathbf{A}_{NN} \end{bmatrix}. \tag{2}$$

The diagonal terms in Equation (2) represents the within channel impulse responses, whereas the off-diagonal ones represents the cross-channel interactions. As described in Figure 1, CT and MRI imaging are performed independently within the fused platform, implying that the off-diagonal are 0s, which leads to the simplification that

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & 0 \\ 0 & \mathbf{A}_2 \end{bmatrix}. \tag{3}$$

Similarly, for multichanneled data, signal, and noise, we have

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \text{ and } \sigma = \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix}, \tag{4}$$

where the subscripts "1" and "2" indicate CT and MRI, respectively. This notation carries on throughout the rest of the paper.

The regression model given by Equation(1) is known to be ill-conditioned and encompasses a large variance noise. According to the CS theory, solving such a problem requires sparse representation of the unknown signal as an additional regularization [10], [12], [14]. Here, TV is employed because

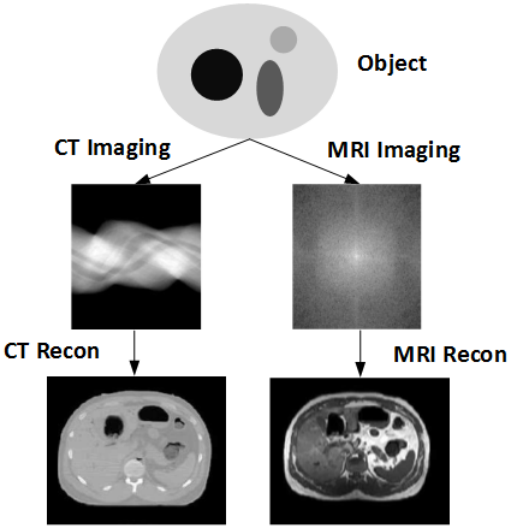

FIGURE 1. Imaging pipeline on a fused CT-MRI system. The CT and MRI signals are collected independently.

it has shown impressive results in many applications. The TV of an image $\mathbf{x}$ is defined as

$$TV(\mathbf{x}) = \int_{\Omega} \|\nabla \mathbf{x}\|_1 d\Omega, \tag{5}$$

where $\nabla \mathbf{x}$ is DGT of the image, and $\|\cdot\|_1$ denotes the L_1-norm, which is defined as

$$\|x\|_1 = (|x_1| + |x_2| + \cdots + |x_n|), \tag{6}$$

where $|\cdot|$ denotes the absolute value. Figure 2 shows the DGT calculation of an abdomen CT image (top row) and an MRI image (bottom row) using the stencil-based high-order approximation [33].

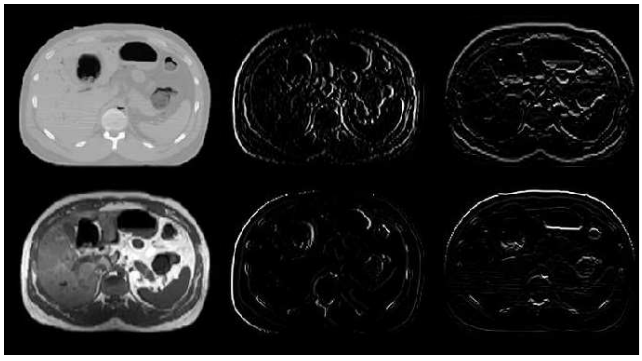

FIGURE 2. Discrete gradient transform using the 8th order approximation of an abdomen CT image on top row and MRI image on bottom row. From left to right: original image; transform in the horizontal direction; transform in the vertical direction.

The TV for a multichannel imaging system $TV_M(\mathbf{x})$ is defined as [26]

$$TV_M(\mathbf{x}) = \sqrt{\sum_{m=1}^{M} TV(\mathbf{x}_m)^2}, \tag{7}$$

where M denotes the total number of the channels in the multichannel imaging system and m denotes the channel index. As a multichannel imaging system, the multimodal reconstruction framework is formulated as a constrained optimization problem defined as

$$\mathbf{x} = \arg\min_{\mathbf{x}} \Big(TV_M(\mathbf{x})\Big), \quad s.t.\ \mathbf{y} = \mathbf{A}\mathbf{x}, \tag{8}$$

which can be solved by minimizing a cost function expressed as

$$f(\mathbf{x}) = \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \xi TV_M(\mathbf{x}), \tag{9}$$

where $\|\cdot\|_2$ denotes the L_2-norm defined as

$$\|x\|_2 = (x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{1}{2}}, \tag{10}$$

and ξ is a constant parameter. Note Equation (9) is a typical Lagrange function consisting of a data fidelity term and a regularization term. In fact, since there are two modalities in our CT-MRI system, the cost function given by Equation (9) needs to be adjusted such that

$$f(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{2}\|\mathbf{A}_1\mathbf{x}_1 - \mathbf{y}_1\|_2^2 + \frac{1}{2}\|\mathbf{A}_2\mathbf{x}_2 - \mathbf{y}_2\|_2^2 + \xi \sqrt{TV(\mathbf{x}_1)^2 + TV(\mathbf{x}_2)^2}, \tag{11}$$

where $\mathbf{x}_1$ is the CT image and $\mathbf{x}_2$ is the MRI image.

A common way to minimize $f(\mathbf{x})$ in Equation (9) is to adopt gradient-based methods. By applying the chain rule, the partial derivative of the multimodaled TV in Equation (9) with respect to (w.r.t.) the mth modality is given by

$$\frac{\partial TV_M(\mathbf{x})}{\partial \mathbf{x}_m} = \frac{TV(\mathbf{x}_m)}{TV_M(\mathbf{x})} \nabla \cdot \left(\frac{\nabla \mathbf{x}_m}{\|\nabla \mathbf{x}_m\|_2}\right). \tag{12}$$

Hence, the partial derivative of the cost function given by (9) w.r.t. the unknown of one single modality is expressed as

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}_m} = \mathbf{A}_m^*\big(\mathbf{A}_m\mathbf{x}_m - \mathbf{y}_m\big) + \xi \frac{TV(\mathbf{x}_m)}{TV_M(\mathbf{x})} \nabla \cdot \left(\frac{\nabla \mathbf{x}_m}{\|\nabla \mathbf{x}_m\|_2}\right), \tag{13}$$

where $\mathbf{A}_m^*$ is the conjugate transpose of $\mathbf{A}$.

B. MULTIMODAL COUPLING

The purpose of jointly reconstructing multimodal images is to enhance image quality by incorporating additional information like structural similarities. The design here is to utilize commonly shared features between CT and MRI images. Although the images have different dynamic range and distribution, they are acquired for the same anatomical structure, and the information encompassed in these images like edges and boundaries carries high correlation. Figure 2 displays CT and MRI images of an abdomen, and their DGTs in horizontal and vertical directions, which show obvious similarities in structures. Figure 3 shows the overlapped image gradients between CT and MRI images. Starting from perfect alignment on the upper left corner, the MRI image rotates such that the inter-vector angle increments by 22.5°. As the inter-vector angle increase, the overlap between the two decreases, and so does the similarity between the CT and MRI images.

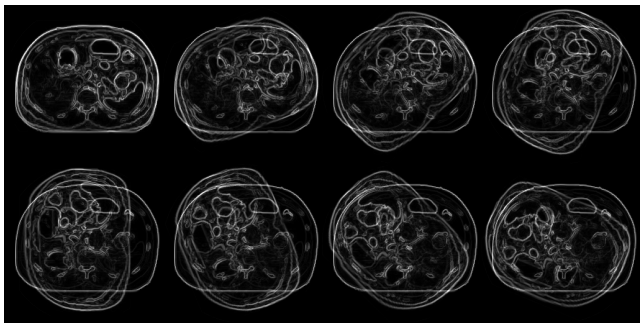

FIGURE 3. Angle between overlaid CT and MRI image gradients varying from 0° to 160°; note that when the angle is 0°, the two images are very well aligned.

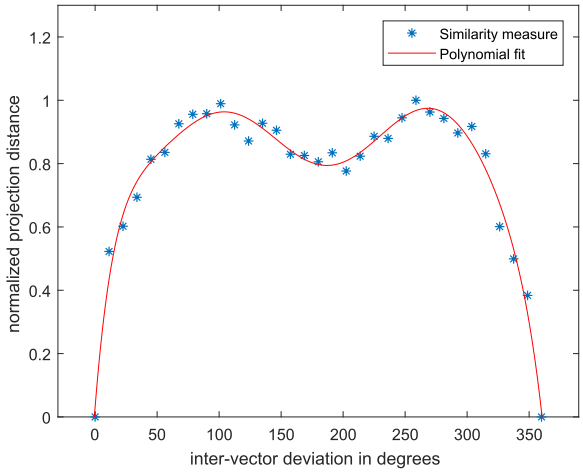

FIGURE 4. Fitting curve defined in Equation (15) shows projection distance as a function of the angle between CT and MRI image TVs; note that it reaches its maximum when the two images are orthogonal and its minimum when the two images are perfectly aligned.

To incorporate the structural similarity of CT and MRI images, one needs to measure the divergence between the two images in a feature space. To this end, we introduce a projection distance between the TVs of two images as a quantitative metric. Specifically, given two vectors $\mathbf{u}$ and $\mathbf{v}$, the projection distance $d_{proj}(\mathbf{u}, \mathbf{v})$ is defined as

$$d_{proj}(\mathbf{u}, \mathbf{v}) = 1 - \cos(\mathbf{u}, \mathbf{v}) = 1 - \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2}, \tag{14}$$

where $\cos(\mathbf{u}, \mathbf{v})$ is the cosine similarity, and $\mathbf{u} \cdot \mathbf{v}$ is the inner product of $\mathbf{u}$ and $\mathbf{v}$. The projection distance measures the dissimilarity between two non-zero vectors of an inner product space by computing the cosine of the angle between them. Consequently, it becomes larger as the two images get less similar, and it becomes smaller as they get more alike. Figure 4 shows how the projection distance changes as a function of the deviation changing between the two vectors. The function curve is approximated by an 8*th* order

polynomial function given by

$$p(x) = \sum_{n=0}^{8} (a_n x^n). \tag{15}$$

Note that $1 \geqslant d_{proj}(\mathbf{u}, \mathbf{v}) \geqslant 0$ and $d_{proj}(\mathbf{u}, \mathbf{v}) = d_{proj}(\mathbf{v}, \mathbf{u})$. It is also straightforward to prove that $d_{proj}(\mathbf{u}, \mathbf{v}) = 0$ when two vectors are identical, and $d_{proj}(\mathbf{u}, \mathbf{v}) = 1$ when two vectors are orthogonal.

In order to embed the projection distance into our gradient-based optimization framework, its derivative w.r.t. the TV of the unknown images is necessary. Here we only show the derivation w.r.t. one of the modalities, and the other simply follows the symmetric principle. By definition, the derivative of $d_{proj}(\mathbf{u}, \mathbf{v})$ w.r.t. $\mathbf{u}$ is given by

$$\frac{\partial}{\partial \mathbf{u}} d_{proj}(\mathbf{u}, \mathbf{v}) = -\frac{\partial}{\partial \mathbf{u}} cos(\mathbf{u}, \mathbf{v}),$$

where

$$\frac{\partial}{\partial \mathbf{u}} cos(\mathbf{u}, \mathbf{v}) = \lim_{\partial \mathbf{u} \to 0} \frac{cos(\mathbf{u} + \partial \mathbf{u}, \mathbf{v}) - cos(\mathbf{u}, \mathbf{v})}{\partial \mathbf{u}},$$

and

$$cos(\mathbf{u} + \partial \mathbf{u}, \mathbf{v}) = \frac{(\mathbf{u} + \partial \mathbf{u}) \cdot \mathbf{v}}{\|\mathbf{u} + \partial \mathbf{u}\|_2 \|\mathbf{v}\|_2}.$$

Particularly, we have

$$\begin{aligned} \|\mathbf{u} + \partial \mathbf{u}\|_2 &= \sqrt{(\mathbf{u} + \partial \mathbf{u}) \cdot (\mathbf{u} + \partial \mathbf{u})} \\ &= \|\mathbf{u}\|_2 \sqrt{1 + 2\frac{\mathbf{u}}{\|\mathbf{u}\|_2^2} \cdot \partial \mathbf{u} + \frac{\|\partial \mathbf{u}\|_2^2}{\|\mathbf{u}\|_2^2}} \\ &\approx \|\mathbf{u}\|_2 \left(1 + \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2} \cdot \partial \mathbf{u}\right). \end{aligned}$$

It follows that

$$\begin{aligned} &cos(\mathbf{u} + \partial \mathbf{u}, \mathbf{v}) \\ &\approx \frac{\mathbf{u} \cdot \mathbf{v} + \partial \mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|_2 \left(1 + \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2} \cdot \partial \mathbf{u}\right) \|\mathbf{v}\|_2} \\ &\approx \frac{\mathbf{u} \cdot \mathbf{v} + \partial \mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2} \left(1 - \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2} \cdot \partial \mathbf{u}\right) \\ &= \left(\frac{\mathbf{u}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2} - \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2} \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2}\right) \cdot \partial \mathbf{u} \\ &\quad + \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2} + O(\partial \mathbf{u}^2) \\ &\approx \cos(\mathbf{u}, \mathbf{v}) + \left(\frac{\mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2} - \cos(\mathbf{u}, \mathbf{v}) \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2}\right) \cdot \partial \mathbf{u}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial \mathbf{u}} cos(\mathbf{u}, \mathbf{v}) &= \lim_{\partial \mathbf{u} \to 0} \frac{cos(\mathbf{u} + \partial \mathbf{u}, \mathbf{v}) - cos(\mathbf{u}, \mathbf{v})}{\partial \mathbf{u}} \\ &= \frac{\mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2} - \cos(\mathbf{u}, \mathbf{v}) \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2}. \end{aligned}$$

Hence,

$$\frac{\partial}{\partial \mathbf{u}} d_{proj}(\mathbf{u}, \mathbf{v}) = \cos(\mathbf{u}, \mathbf{v}) \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2} - \frac{\mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2}. \tag{16}$$

Similarly, due to the symmetric form of projection distance, its derivative w.r.t. $\mathbf{v}$ is given by

$$\frac{\partial}{\partial \mathbf{v}} d_{proj}(\mathbf{u}, \mathbf{v}) = \cos(\mathbf{u}, \mathbf{v}) \frac{\mathbf{v}}{\|\mathbf{v}\|_2^2} - \frac{\mathbf{u}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2}. \tag{17}$$

Note that in our optimization framework $\mathbf{u}$ and $\mathbf{v}$ here are TV terms of the CT and the MRI image, therefore, by applying chain rule we have

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{x}_1} d_{proj}(\mathbf{u}, \mathbf{v}) \\ &= \left(\cos(\mathbf{u}, \mathbf{v}) \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2} - \frac{\mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2}\right) \nabla \cdot \left(\frac{\nabla \mathbf{x}_1}{\|\nabla \mathbf{x}_1\|_2}\right) \end{aligned} \tag{18}$$

for CT imaging, and

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{x}_2} d_{proj}(\mathbf{u}, \mathbf{v}) \\ &= \left(\cos(\mathbf{u}, \mathbf{v}) \frac{\mathbf{v}}{\|\mathbf{v}\|_2^2} - \frac{\mathbf{u}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2}\right) \nabla \cdot \left(\frac{\nabla \mathbf{x}_2}{\|\nabla \mathbf{x}_2\|_2}\right) \end{aligned} \tag{19}$$

for MRI. In Equation (18) and Equation (19), $\mathbf{u} = TV(\mathbf{x}_1)$ and $\mathbf{v} = TV(\mathbf{x}_2)$.

C. MULTIMODAL OPTIMIZATION

To restrict the divergence between CT and MRI images, the projection distance is added to Equation (9). The cost function thereby becomes

$$f(\mathbf{x}) = \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 + \xi TV_M(\mathbf{x}) + \lambda d_{proj}(TV(\mathbf{x})), \tag{20}$$

where λ is the tuning factor of the projection distance term. By expressing Equation (20) as a cost function of a joint CT-MRI optimization, we get

$$\begin{aligned} f(\mathbf{x}_1, \mathbf{x}_2) = &\frac{1}{2}\|\mathbf{A}_1\mathbf{x}_1 - \mathbf{y}_1\|^2 + \frac{1}{2}\|\mathbf{A}_2\mathbf{x}_2 - \mathbf{y}_2\|^2 \\ &+ \xi \sqrt{TV(\mathbf{x}_1)^2 + TV(\mathbf{x}_2)^2} \\ &+ \lambda d_{proj}\big(TV(\mathbf{x}_1), TV(\mathbf{x}_2)\big), \end{aligned} \tag{21}$$

where $d_{proj}\big(TV(\mathbf{x}_1), TV(\mathbf{x}_2)\big)$ is the projection distance of TVs of the CT and the MRI images, and it is defined in Equation (14).

To solve $\mathbf{x}_1$ and $\mathbf{x}_2$ in Equation (21), we generalize the conjugate gradient method to the nonlinear optimization problems. The nonlinear conjugate gradient (NCG) method [34], [35] only requires the gradient of the cost function to seek a local minimum of a quadratic function in the neighborhood of the critical points. By taking the derivatives on both sides of Equation (21) and letting $\mathbf{u} = TV(\mathbf{x}_1)$ and

$\mathbf{v} = TV(\mathbf{x}_2)$, the gradient of $f(\mathbf{x}_1, \mathbf{x}_2)$ in Equation (21) w.r.t. $\mathbf{x}_1$ is

$$\begin{aligned}\frac{\partial f(\mathbf{x}_1, \mathbf{x}_2)}{\partial \mathbf{x}_1} &= \mathbf{A}_1^*(\mathbf{A}_1\mathbf{x}_1 - \mathbf{y}_1) \\ &+ \xi \frac{\mathbf{u}}{TV_M(\mathbf{x})} \nabla \cdot \left(\frac{\nabla \mathbf{x}_1}{\|\nabla \mathbf{x}_1\|_2}\right) \\ &+ \lambda \left(\cos(\mathbf{u}, \mathbf{v}) \frac{\mathbf{u}}{\|\mathbf{u}\|_2^2} - \frac{\mathbf{u}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2}\right) \nabla \cdot \left(\frac{\nabla \mathbf{x}_1}{\|\nabla \mathbf{x}_1\|_2}\right)\end{aligned} \tag{22}$$

and the counterpart w.r.t. $\mathbf{x}_2$ is

$$\begin{aligned}\frac{\partial f(\mathbf{x}_1, \mathbf{x}_2)}{\partial \mathbf{x}_2} &= \mathbf{A}_2^*(\mathbf{A}_2\mathbf{x}_2 - \mathbf{y}_2) \\ &+ \xi \frac{\mathbf{v}}{TV_M(\mathbf{x})} \nabla \cdot \left(\frac{\nabla \mathbf{x}_2}{\|\nabla \mathbf{x}_2\|_2}\right) \\ &+ \lambda \left(\cos(\mathbf{u}, \mathbf{v}) \frac{\mathbf{v}}{\|\mathbf{v}\|_2^2} - \frac{\mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2}\right) \nabla \cdot \left(\frac{\nabla \mathbf{x}_2}{\|\nabla \mathbf{x}_2\|_2}\right),\end{aligned} \tag{23}$$

where

$$\cos(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|_2 \, \|\mathbf{v}\|_2}. \tag{24}$$

In practical implementation, in order to make L_1-norm differentiable throughout its entire support, an infinitesimal term ϵ, e.g. 1×10^{-15}, is added such that

$$\nabla \cdot \left(\frac{\nabla \mathbf{x}_m}{\|\nabla \mathbf{x}_m\|_2}\right) \rightarrow \nabla \cdot \left(\frac{\nabla \mathbf{x}_m}{\|\nabla \mathbf{x}_m + \epsilon\|_2}\right). \tag{25}$$

The joint CT-MRI system essentially extends the NCG optimization problem to a multivariate level. To solve it, the updates of Algorithm 1 are calculated at each iteration. Here, the superscript n denotes the nth iteration; $\mathbf{d}^{(n)}$ denotes conjugate gradient; $\alpha^{(n)}$ denotes step size found by the line search algorithm satisfying the Wolfe condition [36], [37]; $\beta^{(n)}$ is computed by the Fletcher-Reeves method [38]; and $\cdot$ depicts the dot product. The entire iterative algorithm is summarized in Algorithm 1. Note that in Algorithm 1, $\beta_1^{(n)}$ and $\beta_2^{(n)}$ are the same since the NCG method accounts for the union of $\mathbf{x}_1$ and $\mathbf{x}_2$ expressed as $[\mathbf{x}_1, \mathbf{x}_2]$.

Algorithm 1 CS-Based Joint Reconstruction

Initialization:
$\mathbf{x}_1^{(0)} = 0$; $\mathbf{x}_2^{(0)} = 0$;
$\mathbf{d}_1^{(0)} = -\nabla_{\mathbf{x}_1} f(\mathbf{x}_1^{(0)}, \mathbf{x}_2^{(0)})$;
$\mathbf{d}_2^{(0)} = -\nabla_{\mathbf{x}_2} f(\mathbf{x}_1^{(0)}, \mathbf{x}_2^{(0)})$;
$k = 0$;
while *true* **do**
  Line search:
  $\alpha^{(n)} = \arg\min_\alpha f(\mathbf{x}_1^{(n)} + \alpha \mathbf{d}_1^{(n)}, \mathbf{x}_2^{(n)} + \alpha \mathbf{d}_2^{(n)})$.
  Update results
  $\mathbf{x}_1^{(n+1)} = \mathbf{x}_1^{(n)} + \alpha^{(n)} \mathbf{d}_1^{(n)}$;
  $\mathbf{x}_2^{(n+1)} = \mathbf{x}_2^{(n)} + \alpha^{(n)} \mathbf{d}_2^{(n)}$;
  Update steepest decent direction
  $\mathbf{g}_1^{(n+1)} = -\nabla_{\mathbf{x}_1} f(\mathbf{x}_1^{(n+1)}, \mathbf{x}_2^{(n+1)})$;
  $\mathbf{g}_2^{(n+1)} = -\nabla_{\mathbf{x}_2} f(\mathbf{x}_1^{(n+1)}, \mathbf{x}_2^{(n+1)})$;
  Update Fletcher-Reeves value
  $\mathbf{g}^{(n)} = [\mathbf{g}_1^{(n)}, \mathbf{g}_2^{(n)}]$;
  $\beta_1^{(n)} = \frac{\mathbf{g}^{(n+1)} \cdot \mathbf{g}^{(n+1)}}{\mathbf{g}^{(n)} \cdot \mathbf{g}^{(n)}}$;
  $\beta_2^{(n)} = \frac{\mathbf{g}^{(n+1)} \cdot \mathbf{g}^{(n+1)}}{\mathbf{g}^{(n)} \cdot \mathbf{g}^{(n)}}$;
  Update conjugate gradient
  $\mathbf{d}_1^{(n+1)} = \mathbf{g}_1^{(n+1)} + \beta_1^{(n)} \mathbf{d}_1^{(n)}$;
  $\mathbf{d}_2^{(n+1)} = \mathbf{g}_2^{(n+1)} + \beta_2^{(n)} \mathbf{d}_2^{(n)}$;
  $k = k + 1$;
  if $k > K$ or $|d\mathbf{x}_1| < \epsilon_1$ or $|d\mathbf{x}_2| < \epsilon_2$ **then**
    break;
  end if
end while

III. RESULTS AND DISCUSSION

The joint reconstruction framework is tested on a pair of body images, including a CT image and an MRI image. The images are acquired for the same anatomy and are both sized 256×256 pixels. Image intensity is normalized to [0 1] for both images. The undersampled CT sinogram is created by simulating a forward projection from 51 different directions out of a full rotation ranging from 0° to 360°. Similarly, the undersampled MRI k-space data are collected by applying an undersampling mask to the fully sampled k-space data, which are generated by applying the Fourier transform to the original image.

The algorithm is implemented on a PC with Microsoft Windows 10 Pro powered by Exon CPU. The PC has 32 Gigabyte of memory such that all programmings are coded under x64 mode. The computer is also equipped with an nVidia M5000M Graphic Processing Unit (GPU) provided with an 8-Gigabyte on-chip memory driven by Compute Unified Device Architecture (CUDA) 8.0. The software is developed in MATLAB, C/C++, and CUDA, combined by the MEX interface.

The forward projection and back-projection for the CT imaging is programmed in CUDA-based C code to further accelerate the computation. The CT projection follows an equi-angular fanbeam geometry with no iso-center offset. The number of the detectors is 1024, which is a typical value emulating major vendors in current market. The projection implementation employs a ray-driven model for forward projection and a pixel-driven [39] model for back-projection. As for the MRI, the 2D-Fourier operator and its inverse are performed with built-in MATLAB functions. The Fourier transform is not carried out by GPU code, but it is sufficiently fast.

The evaluation is conducted by comparing the results obtained from all experiments. The three tested methods are listed as following scenarios:

1) Analytic reconstructions;
2) Reconstruction based on Equation (11);
3) Reconstruction based on Equation (21).

The comparison among the three reconstruction strategies are also compared in Table 1.

TABLE 1. Comparison of different reconstruction strategies.

Algo. Descrip. \ Recon. Strategy	Analytic Recon.	Ind. Recon.	Joint Recon.
Algorithm	CT: FBP MRI: IFT	NCG	NCG
Iterative	No	Yes	Yes
TV	N/A	Yes	Yes
PD	N/A	No	Yes

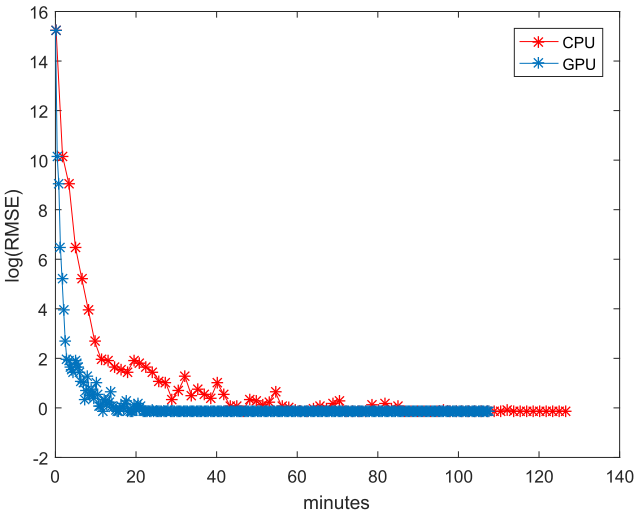

FIGURE 5. Convergence curves with CPU implementation and accelerated GPU implementation. Both follow the strategy described in Algorithm 1.

Figure 5 shows the convergence of reconstruction implemented in GPU and CPU, both of which follow Algorithm 1. The GPU implementation significantly decease the time of convergence as the plot shows. The NCG algorithm approaches its minimal at approximately 600 iterations, and the cost function decreases with a very steady behavior afterwards. The forward and backprojection of CT imaging are computationally expensive, so they are accelerated with CUDA implementation with massive parallelism. Consequently, each iteration takes roughly 0.4 seconds to update both CT and MRI images. To get final reconstruction results, it takes approximately 400 seconds.

The results of different reconstructions are displayed in Figure 6 and Figure 7 in the following order: top left depicts the original image as the truth; top right depicts the CS-based joint reconstruction with projection distance regularization; bottom left depicts the CS-based reconstruction independently on single modality without projection distance regularization; and bottom right depicts the analytic reconstruction of the undersampled sinogram with 51 projections using

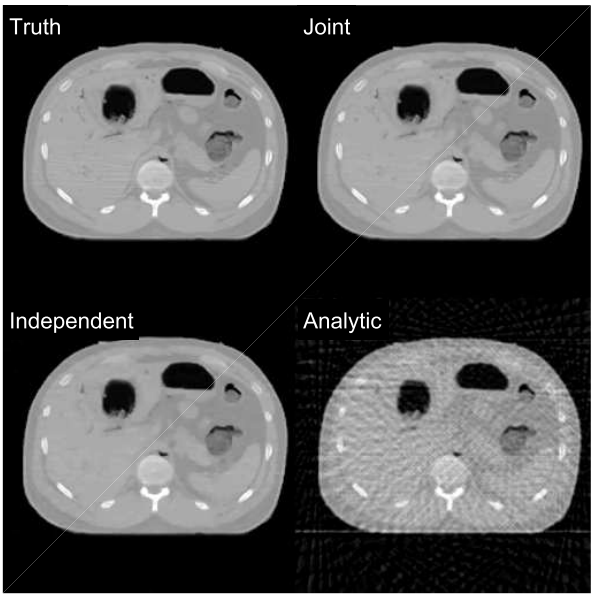

FIGURE 6. Comparison of CT reconstructions: top left depicts the original image as the truth; top right depicts the joint reconstruction; bottom left depicts the independent reconstruction; and bottom right depicts the analytic reconstruction. The display windows is [0 1].

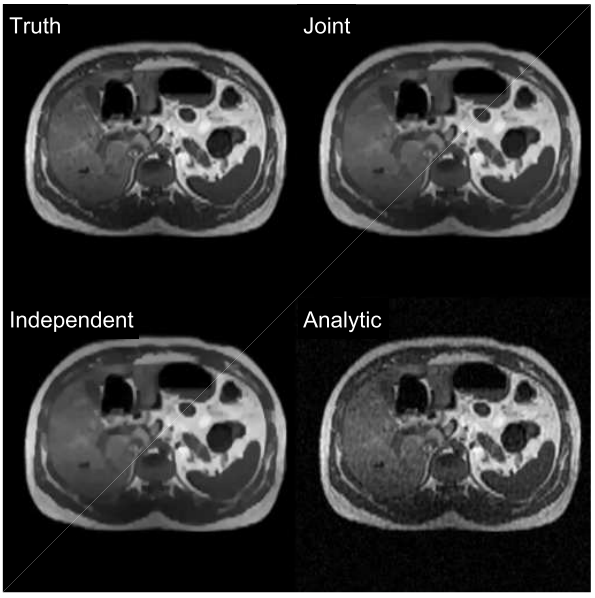

FIGURE 7. Comparison of MRI reconstructions: top left depicts the original image as truth truth; top right depicts the joint reconstruction; bottom left depicts the independent reconstruction; and bottom right depicts an analytic reconstruction. The display windows is [0 1].

native fanbeam geometry algorithm [8] for CT, and inverse Fourier transform with 40% k-space sampling rate for MRI. The analytic reconstruction exhibits strong artifacts due to severe undersampling. Streaks and distortion are quite obvious in both CT and MRI images. In contrast, the joint reconstruction shows a similar image quality as the independent reconstruction. Both of the reconstructed images are very similar to the truth. The tuning parameters here are chosen

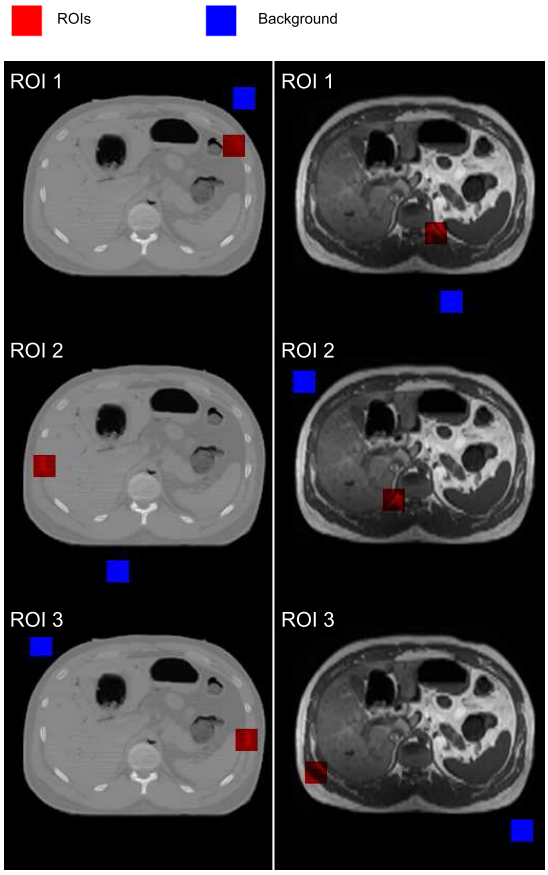

FIGURE 8. Three ROIs are chosen from both CT and MRI images. ROIs are higjlighred in red color, whereas, regions of background are also chosen, but labeled in blue color. Display window is [0 1].

as $\xi = 0.01$ and $\lambda = 0.05$. The choice of ξ and λ is experimental, and is not under the scope of this work.

A detailed analysis is conducted by observing zoomed-in region of interests (ROIs). The ROIs are marked in red color as depicted in Figure 8. Figure 8 also defines background regions where noise is calculated. This will be further explained in the paragraph where statistics of ROIs are discussed.

Figure 9 and Figure 10 display ROIs for CT and MRI images, respectively. CT ROIs focus on bony structures which absorb a large amount of X-ray photons, and therefore show much higher intensity. Similarly, MRI ROIs are chosen with tissues containing high density of protons, such as lipids, as they show stronger intensity in images. The evaluation of the images concentrates on the texture, artifacts, resolution and noise. By visual inspection, it is observed that the joint reconstruction (top right panel) in both CT and MRI provides better intensity and focus of texture on bone structure as compared with the independent and analytic reconstructions. It also achieves the best resolution among the tested reconstruction methods by comparison with the original image. In addition, the joint reconstruction also exhibits the best smoothness in background tissue, especially in MRI ROIs.

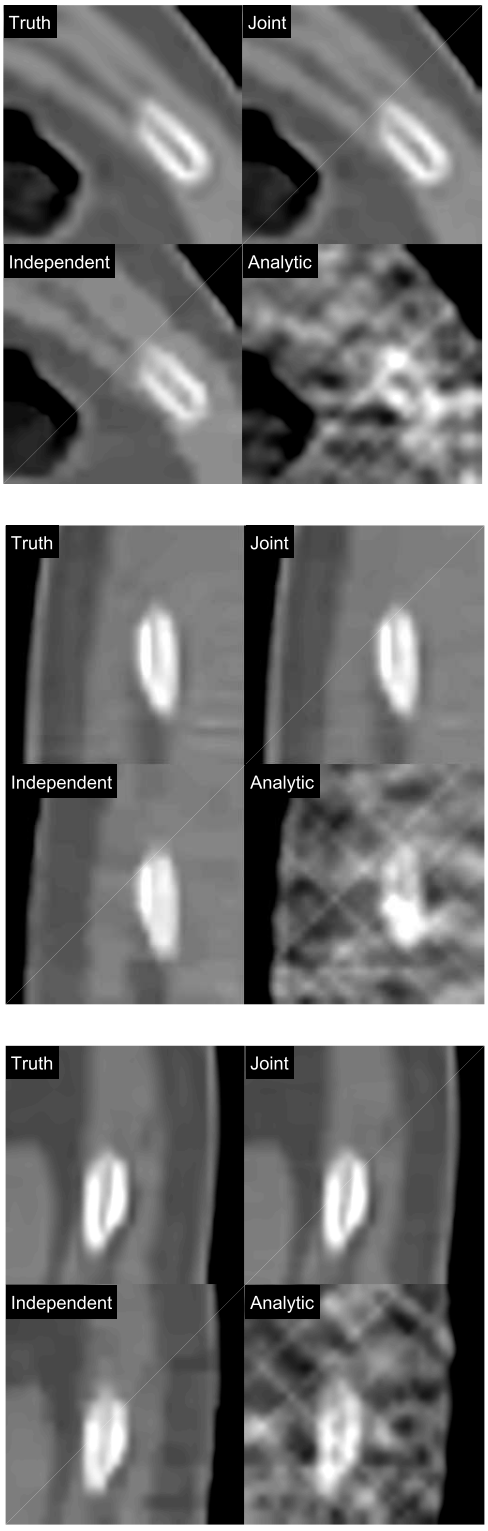

FIGURE 9. ROI comparison for CT imageswith all three ROIs being defined in Fig. 8 in red color; All ROIs are zoomed in with zoom factor of 8×. Display window is [0 1].

Image difference is also examined. Figure 11 and Figure 12 show the image differences for different reconstructions compared to the truth. In particular, image difference gives a much clearer distinction between joint

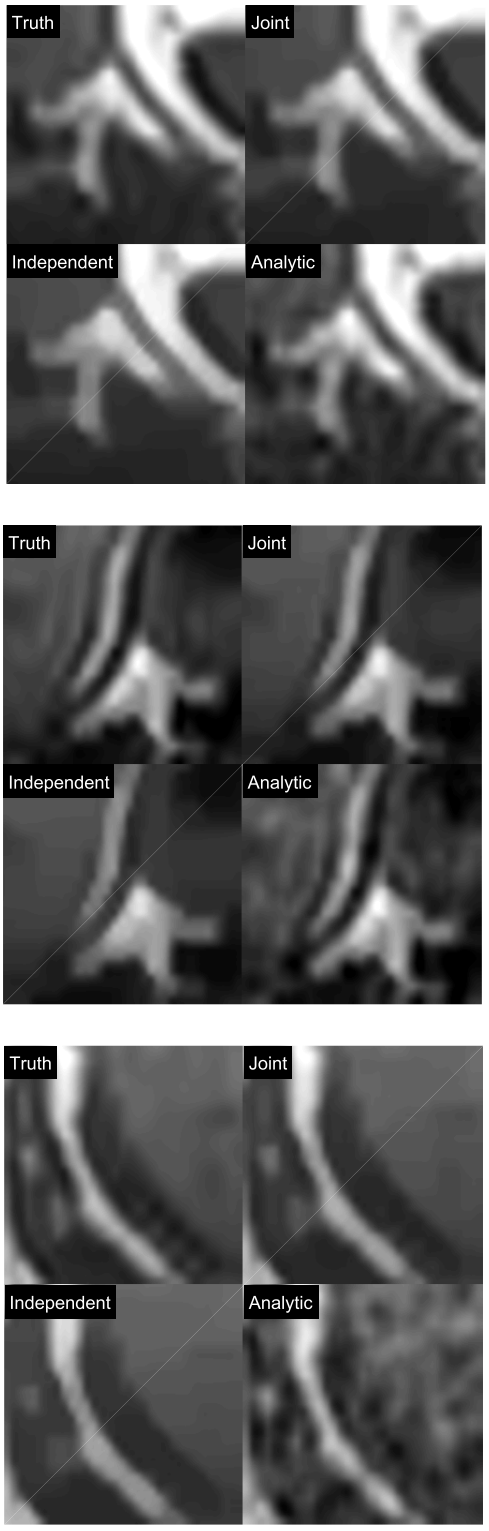

FIGURE 10. ROI comparison for MRI images with all three ROIs being defined in Fig. 8 in red color; All ROIs are zoomed in with zoom factor of 8×. Display window is [0 1].

and independent reconstruction, both of which are iterative methods. Under the proper window level, joint reconstruction shows much less difference than independent reconstruction when both are compared to the truth.

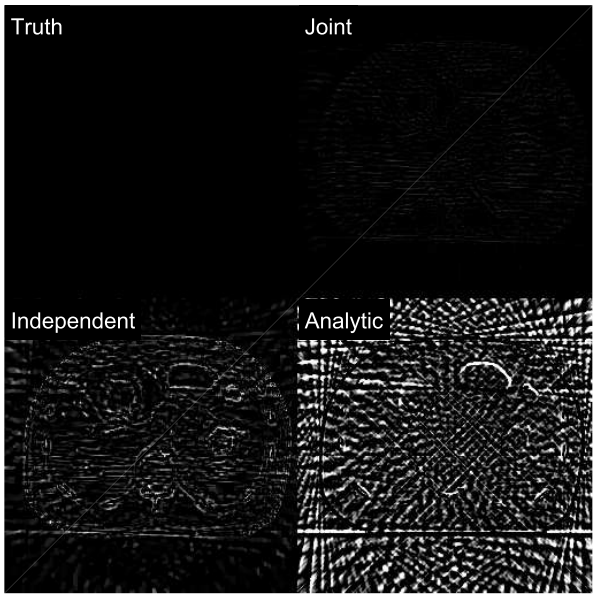

FIGURE 11. Image difference with respect to ground truth: truth on the top left; joint reconstruction on the top right; independent reconstruction on the bottom left; analytic reconstruction on the bottom right. The display windows is [0 0.08].

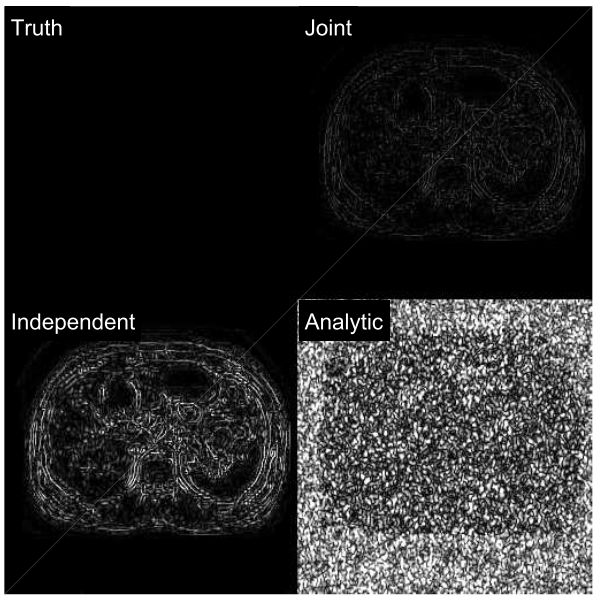

FIGURE 12. Image difference with respect to ground truth: truth on the top left; joint reconstruction on the top right; independent reconstruction on the bottom left; analytic reconstruction on the bottom right. The display windows is [0 0.08].

Quantitative analysis using statistics are also performed for each ROI. All ROIs are also considered region of signal, yet the back-group airial region is excluded. The background regions, highlighted in blue color in Figure 8, are chosen to evaluate noise. All regions of signal contain high intensity objects, and all backgroud regions lie where no anatomy being imaged. As larger intensity does not necessarily mean good image nor does mean low noise,

TABLE 2. Statistics CT ROIs.

ROI	Stats	Truth	Joint	Ind.	Ana.
ROI 1	Max	0.666	0.664	0.652	0.679
	Min	0.519	0.522	0.514	0.425
	Mean	0.590	0.590	0.588	0.563
	Noise	0.0010	0.0015	0.0023	0.0609
	SNR	590.0000	393.3333	255.6522	9.2447
ROI 2	Max	0.803	0.788	0.776	0.778
	Min	0.529	0.523	0.524	0.479
	Mean	0.656	0.627	0.623	0.589
	Noise	0.0011	0.0016	0.00528	0.0318
	SNR	596.3636	391.8750	117.9924	18.5220
ROI 2	Max	0.825	0.822	0.809	0.761
	Min	0.470	0.469	0.486	0.447
	Mean	0.632	0.632	0.630	0.625
	Noise	0.0011	0.0014	0.0051	0.0306
	SNR	574.5455	451.4286	123.5294	20.4248
Ave.	Max	0.7647	0.7580	0.7457	0.7393
	Min	0.5060	0.5047	0.5080	0.4503
	Mean	0.6260	0.6163	0.6137	0.5923
	Noise	0.0011	0.0015	0.0042	0.0411
	SNR	586.9697	412.2123	165.7247	16.0638

TABLE 3. Statistics of MRI ROIs.

ROI	Stats	Truth	Joint	Ind.	Ana.
ROI 1	Max	0.739	0.716	0.709	0.783
	Min	0.098	0.122	0.123	0.022
	Mean	0.321	0.320	0.318	0.322
	Noise	0.0010	0.0015	0.0023	0.0609
	SNR	321.0000	213.3333	138.2609	5.2874
ROI 2	Max	0.687	0.650	0.637	0.718
	Min	0.078	0.107	0.109	0.047
	Mean	0.248	0.248	0.246	0.251
	Noise	0.0011	0.0016	0.00528	0.0318
	SNR	225.4545	155.0000	46.5909	7.8931
ROI 3	Max	0.622	0.577	0.567	0.618
	Min	0.120	0.156	0.157	0.088
	Mean	0.289	0.288	0.286	0.291
	Noise	0.0011	0.0014	0.0051	0.0306
	SNR	262.7273	205.7143	56.0784	9.5098
Ave.	Max	0.6827	0.6477	0.6377	0.7063
	Min	0.0987	0.1283	0.1297	0.0523
	Mean	0.2860	0.2853	0.2833	0.2880
	Noise	0.0011	0.0015	0.0042	0.0411
	SNR	269.7273	191.3492	80.3101	7.5634

signal-to-noise ratio (SNR) is calculated for each reconstruction. For digital images, the SNR is defined as [40]

$$SNR = \frac{\mu_{signal}}{\sigma_{background}}, \tag{26}$$

where μ_{signal} is the mean value of the signal region, and $\sigma_{background}$ is the standard deviation of the noise of the background region. As Figure (8) shows, the signal region is represented with red color while the noise region is represented with blue color. Maximal value, minimal value, and mean value are computed for each signal region, and noise is computed as the standard deviation of each background region. The results of the CT and MRI ROIs are listed in Table 2 in Table 3, respectively.

The statistical results confirm the previous visual inspection: joint reconstruction reveals the truth with the highest confidence. In particular, it approximates the true signal

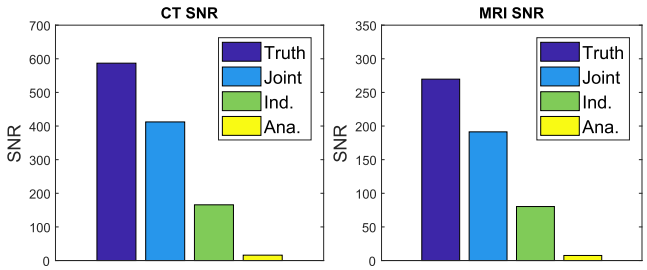

FIGURE 13. Averaged SNR comparison for CT and MRI reconstruction using the joint, individual and analytic methods.

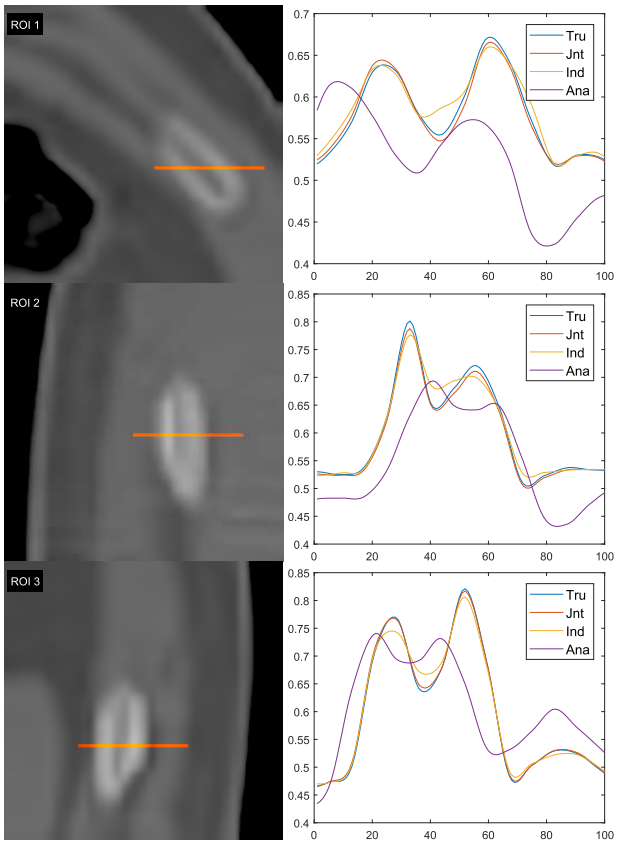

FIGURE 14. CT reconstruction comparison in line profile: top depicts the line ROI on CT image; bottom left depicts the line profile comparison; bottom right depicts a zoom-in of high frequency part of the line ROI; display windows is normalized to [0 1].

better in high intensity location and generates less noise in the background. The independent reconstruction has close quality to joint reconstruction, but ignoring the multimodal information induces small accuracy as compared with the joint reconstruction. It is also noticed that the analytic reconstruction sometimes gives stronger signal intensity than the joint reconstruction, but its noise in the background is extremely large, and this eventually leads to a very poor SNR. Figure 13 displays values of SNR according to Equation (26) for different reconstructions. The calculation is based on normalized image pixel value. SNRs of both CT and MRI images follow a descending order from truth to analytic as the graph shows.

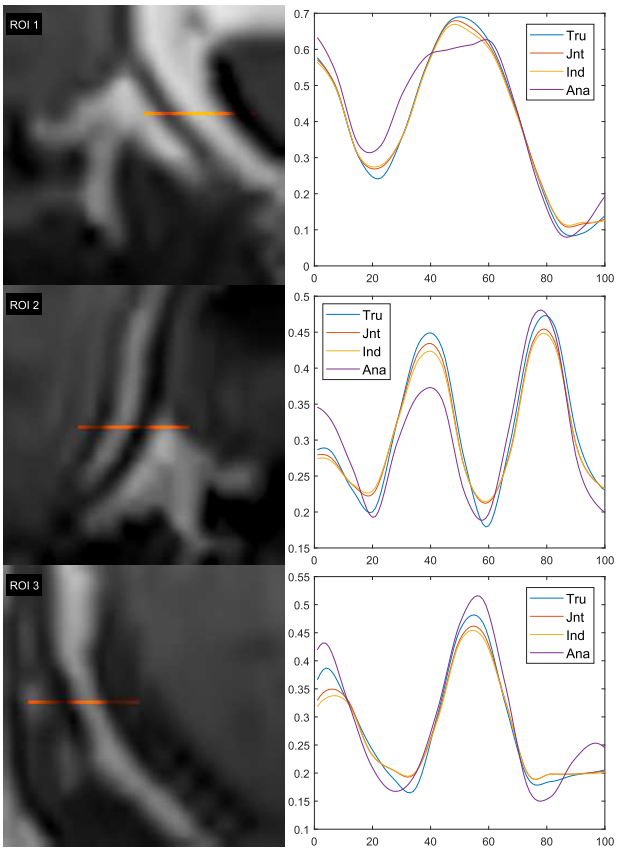

FIGURE 15. CT reconstruction comparison in line profile; top depicts the line ROI on CT image; bottom left depicts the line profile comparison; bottom right depicts a zoom-in of high frequency part of the line ROI; display windows is normalized to [0 1].

Aside from the statistic comparison, line profile is another way to investigate the reconstruction quality. The line profiles are also drawn in the same ROIs as defined in Figure 8, and they are across areas containing high frequency components such as boundaries and edges. Figure 14 and Figure 15 displays how each reconstruction performs comparing to the original image. The results are consistent with those from visual inspection and statistic analysis. Joint reconstruction, among the three, gives the best reconstruction quality of reconstruction followed by individual reconstruction. Analytic reconstruction, however, holds a far more inferior result to the aforementioned two.

TABLE 4. Similarity measurement for CT reconstruction.

Strategy	RMSE	SSIM	Correlation
Analytic	13.5862	0.4135	0.9786
Independent	1.7551	0.9690	0.9996
Joint Recon	0.7578	1.0000	0.9999
Truth	0.0000	1.0000	1.0000

Similarity metrics between the original image and its reconstructed versions are computed as the third mean for verification. The metrics here include the Root-Mean-Square Error (RMSE), the structural similarity (SSIM) [41], and the correlation coefficients. The results are displayed in Table 4 for CT imaging and in Table 5 for MRI. All calculations are

TABLE 5. Similarity measurement for MRI reconstruction.

Strategy	RMSE	SSIM	Correlation
Analytic	11.2517	0.4717	0.9860
Independent	2.3990	0.9755	0.9991
Joint Recon	2.0061	0.9869	0.9993
Truth	0.0000	1.0000	1.0000

based on normalized images as the aforementioned. For both the CT and MRI images, the joint reconstruction exhibits the highest similarity and correlation with the original image. The independent reconstruction follows the joint reconstruction. The analytic methods rank the lowest among all the tested methods here.

IV. CONCLUSIONS

In this paper, it is demonstrated that integration of multimodal information of structural similarity can enhance reconstruction quality. The CS technique employed here gives an example of how to achieve multimodal reconstruction with highly undersampled data. The choice of the joint priors is instrumental and needs to be handled properly to balance the image quality. TV is included so that artifacts are suppressed while projection distance is utilized to govern the divergence between different modalities. The results reveal that the joint reconstruction framework produces superior image reconstruction quality when using the iterative-based methods that incorporates multimodal structural similarities. Future work will include the extension of the proposed joint reconstruction concept to more scenarios for which structural similarity is available. It will also include the application of different sparse transformations, the development of 3D-imaging framework, new ways of coupling multimodality information, and faster and more robust optimization algorithm to further reduce data sampling.

ACKNOWLEDGMENT

The authors are grateful for Dr. Ge Wang from Rensselaer Polytechnic Institute for his constructive suggestions.

REFERENCES

[1] J. Siewerdsen, "MO-B-352-02: Multi-modality imaging: Physics and technology considerations in diagnostic and image-guided procedures," *Med. Phys.*, vol. 35, no. 6, p. 2862, 2008.
[2] M. Vannier, "TU-A-201C-01: Multi-modality imaging: Current state of the art," *Med. Phys.*, vol. 37, no. 6, p. 3371, 2010.
[3] L. Marti-Bonmati, R. Sopena, P. Bartumeus, and P. Sopena, *Multimodality Imaging Techniques*, vol. 5. Hoboken, NJ, USA: Wiley, 2010, pp. 180–189.
[4] G. Wang *et al.*, "Towards omni-tomography—Grand fusion of multiple modalities for simultaneous interior tomography," *PLoS ONE*, vol. 7, no. 6, p. e39700, 2012.
[5] G. Wang *et al.*, "Vision 20/20: Simultaneous CT-MRI—Next chapter of multimodality imaging," *Med. Phys.*, vol. 42, no. 10, pp. 5879–5889, 2015.
[6] H. Yu and G. Wang, "Compressed sensing based interior tomography," *Phys. Med. Biol.*, vol. 54, no. 9, pp. 2791–2805, May 2009.
[7] H. Yu *et al.*, "Compressive sensing-based interior tomography: Preliminary clinical application," *J. Comput. Assist. Tomogr.*, vol. 35, no. 6, pp. 762–764, 2011.
[8] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Philadelphia, PA, USA: SIAM, 2001.
[9] D. G. Nishimura, *Principles of Magnetic Resonance Imaging*. Stanford, CA, USA: Stanford Univ. Press, 2010.

[10] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008.
[11] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. New York, NY, USA: Pearson, 1975.
[12] E. J. Candés, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
[13] E. J. Candés, J. K. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
[14] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[15] M. Bertero, *Inverse Problems in Scattering and Imaging*, 1st ed. Boca Raton, FL, USA: CRC Press, 1992.
[16] E. Y. Sidky and X. C. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.*, vol. 53, no. 17, pp. 4777–4807, 2008.
[17] G.-H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.*, vol. 35, no. 2, pp. 660–663, 2008.
[18] H. Kudo, T. Suzuki, and E. A. Rashed, "Image reconstruction for sparse-view CT and interior CT—Introduction to compressed sensing and differentiated backprojection," *Quant. Imag.Med. Surg.*, vol. 3, no. 3, pp. 147–161, 2013.
[19] E. Haber and D. Oldenburg, "Joint inversion: A structural approach," *Inverse Problems*, vol. 13, no. 1, pp. 63–77, 1997.
[20] D. S. Rigie and P. J. L. Rivière, "Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization," *Phys. Med. Biol.*, vol. 60, no. 5, pp. 1741–1762, 2015.
[21] M. J. Ehrhardt *et al.*, "Joint reconstruction of PET-MRI by exploiting structural similarity," *Inverse Problems*, vol. 13, no. 1, pp. 15001–15023, 2015.
[22] J. J. Vaquero and P. Kinahan, "Positron emission tomography: Current challenges and opportunities for technological advances in clinical and preclinical imaging systems," *Annu. Rev. Biomed. Eng.*, vol. 17, pp. 385–414, Dec. 2015.
[23] X. Cui, H. Yu, G. Wang, and L. Mili, "Total variation minimization-based multimodality medical image reconstruction," *Proc. SPIE*, vol. 9212, Aug. 2014, Art. no. 92121D.
[24] Y. Xi, J. Zhao, J. R. Bennett, M. R. Stacy, A. J. Sinusas, and G. Wang, "Simultaneous CT-MRI reconstruction for constrained imaging geometries using structural coupling and compressive sensing," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 6, pp. 1301–1309, Jun. 2015.
[25] L. Gjesteby, Y. Xi, M. K. Kalra, Q. Yang, and G. Wang, "Hybrid imaging system for simultaneous spiral MR and X-ray (MRX) scans," *IEEE Access*, vol. 5, pp. 1050–1061, 2016.
[26] X. Bresson and T. F. Chan, "Fast dual minimization of the vectorial total variation norm and applications to color image processing," *Inverse Problems Imag.*, vol. 2, no. 4, pp. 455–484, 2008.
[27] F. Knoll, M. Holler, T. Koesters, R. Otazo, K. Bredies, and D. K. Sodickson, "Joint MR-PET reconstruction using a multi-channel image regularizer," *IEEE Trans. Med. Imag.*, vol. 36, no. 1, pp. 1–16, Jan. 2017.
[28] L. A. Gallardo and M. A. Meju, "Structure-coupled multiphysics imaging in geophysical sciences," *Rev. Geophys.*, vol. 49, no. 1, pp. 1–19, 2011.
[29] E. Haber and M. H. Gazit, "Model fusion and joint inversion," *Surv. Geophys.*, vol. 34, no. 5, pp. 675–695, 2013.
[30] A. Aghasi, I. Mendoza-Sanchez, E. L. Miller, C. A. Ramsburg, and L. M. Abriola, "A geometric approach to joint inversion with applications to contaminant source zone characterization," *Inverse Problems*, vol. 29, no. 11, p. 115014, 2013.
[31] J. Tang and A. Rahmim, "Bayesian PET image reconstruction incorporating anato-functional joint entropy," *Phys. Med. Biol.*, vol. 54, no. 23, p. 7063–7075, 2010.
[32] S. Somayajula, C. Panagiotou, A. Rangarajan, Q. Li, S. R. Arridge, and R. M. Leahy, "PET image reconstruction using information theoretic anatomical priors," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 537–549, Mar. 2010.
[33] X. Cui and H. Guo, "Stencil-based discrete gradient transform using GPU device in compressed sensing MRI," in *Proc. Fully3D*, 2017, pp. 433–437.
[34] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C—The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.
[35] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, NY, USA: Springer-Verlag, 2000.
[36] P. Wolfe, "Convergence conditions for ascent methods," *SIAM Rev.*, vol. 11, no. 2, pp. 226–235, 1969.
[37] P. Wolfe, "Convergence conditions for ascent methods. II: Some corrections," *SIAM Rev.*, vol. 13, no. 2, pp. 185–188, 1969.
[38] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *Comput. J*, vol. 7, pp. 149–254, 1964.
[39] B. De Man and S. Basu, "Distance-driven projection and backprojection," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 3. Nov. 2002, pp. 1477–1480.
[40] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. London, U.K.: Pearson, 2007.
[41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

XUELIN CUI received the B.S. degree in mechanical engineering from Southeast University, China, in 2001, and the M.S. degree in electrical engineering from the University of Hawaii, Manoa, USA, in 2009, and the M.S. degree in biomedical engineering from Vanderbilt University, USA, in 2011. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Virginia Tech, USA. His research interests include medical imaging, image reconstruction, and signal processing.

LAMINE MILI (F'82) received the Diploma degree in electrical engineering from the Swiss Federal Institute of Technology, Lausanne, in 1976, and the Ph.D. degree from the University of Liege, Belgium, in 1987. He is currently a Professor of electrical and computer engineering with Virginia Tech. His research interests include robust statistics, robust statistical signal processing, radar systems, and power system analysis and control.

HENGYONG YU received the bachelor's degree in information science and technology in 1998 and computational mathematics in 1998, respectively, and the Ph.D. degree in information and communication engineering from Xi'an Jiaotong University in 2003. He is currently an Associate Professor and the Director of the Imaging and Informatics Laboratory, Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA, USA. His research interests include medical imaging (computed tomography) and medical image analysis.