Spatial-temporal Total Variation Regularization (STTVR) for 4D-CT Reconstruction

Haibo Wu[a, b], Andreas Maier[a], Rebecca Fahrig[c], and Joachim Hornegger[a, b]

[a]Pattern Recognition Lab (LME), Department of Computer Science, Friedrich-Alexander-University Erlangen-Nuremberg, Martensstr. 3, 91058 Erlangen, Germany
[b]Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander-University Erlangen-Nuremberg, Am Weichselgarten 8, 91058 Erlangen, Germany
[c] Department of Radiology, Lucas MRS Center, Stanford University, 1201 Welch Road, 94304 Palo Alto, CA, USA.

ABSTRACT

Four dimensional computed tomography (4D-CT) is very important for treatment planning in thorax or abdomen area, e.g. for guiding radiation therapy planning. The respiratory motion makes the reconstruction problem ill-posed. Recently, compressed sensing theory was introduced. It uses sparsity as a prior to solve the problem and improves image quality considerably. However, the images at each phase are reconstructed individually. The correlations between neighboring phases are not considered in the reconstruction process. In this paper, we propose the spatial-temporal total variation regularization (STTVR) method which not only employs the sparsity in the spatial domain but also in the temporal domain. The algorithm is validated with XCAT thorax phantom. The Euclidean norm of the reconstructed image and ground truth is calculated for evaluation. The results indicate that our method improves the reconstruction quality by more than 50% compared to standard ART.

Keywords: total variation, 4D-CT reconstruction, compressed sensing

1. INTRODUCTION

Four dimensional computed tomography (4D-CT) is an important tool in radiation oncology. Besides the 3D information, it also captures the movement of the body's organs over time. With the motion information, it improves the target volume definition and enhances accuracy during treatment delivery for patient with tumor in the abdomen or thorax area.

There are mainly two 4D-CT acquisition methods, namely, the retrospective acquisition method[1–5] and prospective acquisition methods.[6] In the first one, the projection images are continuously acquired. At the same time, an additional respiratory signal is recorded. The projection images are then grouped to different phases according to the amplitude or phase-angle sorting. In the second one, the CT scanner is triggered by the respiratory signal. The beam is turned on if the respiratory signal is within the tolerances to the reference phase. The projections of different respiratory phases are acquired separately.

However, to achieve clinically usable image quality, hundreds of projections are needed to reconstruct images of each respiratory phase no matter which data acquisition method is used. Therefore the radiation dose is of great concern for 4D-CT since it may increase the risk of having cancer. To reduce the radiation dose without sacrificing image quality, the theory of compressed sensing was introduced to CT reconstruction recently.[7,8] Compressed sensing employs sparsity as a prior and use L1 norm minimization method, which is a sparsity promoting algorithm, to do reconstruction.[9–11] Thus sparsity is a key point of compressed sensing. If the images are not sparse, one should apply a certain sparsifying transform. For 4D-CT, the object can be considered as a series of images or volumes. Pan's group makes use of the data redundancy in the spatial domain.[7] They apply the total variation as sparsifying transform to the volume of certain phase. However, the correlation of

Further author information:
Haibo Wu: E-mail: haibo.wu@informatik.uni-erlangen.de, Telephone: 49 9131 85 28982

volumes in the temporal domain is not considered in their method. In contrast, Jiang's group utilizes the data redundancy only in temporal domain.[12] Non-local mean operator is used to promote smooth motion in temporal domain, while the data redundancy in the spatial domain is not employed.

In this paper, we propose a novel reconstruction method which applies the sparsifying transform both in spatial and temporal domain to promote smooth image and smooth motion pattern. In our method, reconstruction problem is formulated to an optimization problem. The images of different phases are treated as one entity and they are reconstructed at same time. Consequently, the dimension of the optimization problem is incredibly high. Therefore we develop an optimization method based on the idea from Pan's group to solve the optimization problem efficiently. We used XCAT phantom[13] to validate our algorithm. The Euclidean norm of the reconstructed image and ground truth is calculated for evaluation. We compared our reconstruction results to the ones of TVR[7] and standard ART method.[14]

The outline of the paper is as follows: we describe our method in section 2 and present the experiments and results in section 3. The conclusion is in section 4.

2. MODEL AND ALGORITHM

2.1 MODEL

The 4D object to be reconstructed can be viewed as a temporal sequence of 3D spatial images, i.e.,

$$\vec{X} = \{\vec{x}_j, 1 \leq j \leq n_t\}. \tag{1}$$

where $\vec{X}$ is the 4D object consisted by $\vec{x}_j$ with temporal index j, which corresponds to one of n_t phases. Consequently, $\vec{Y}$ is the complete set of projection data.

$$\vec{Y} = \{\vec{y}_j = \vec{A}_j\vec{x}_j, j \leq n_t\}, \tag{2}$$

where $\vec{y}_j$ stands for the projection images of $\vec{x}_j$ and $\vec{A}_j$ corresponds to the system matrix for the jth phase. The reconstruction problem is ill-posed in 4D-CT reconstruction due to lack of projection images. To overcome this, compressed sensing introduces sparsity as a prior to do the reconstruction. It formulates the reconstruction problem as the following optimization problem:

$$\vec{x}_j = \min_{\vec{x_j}} ||\vec{A}_j\vec{x}_j - \vec{y}_j||_2^2 + \beta R(\vec{x}_j), j \leq n_t, \tag{3}$$

where $R(\bullet)$ stands for the regularization term and β is a weighting factor. There are two parts in the cost function. The first term enforces the data fidelity and second term is a regularization term which promotes sparsity after applying the sparsifying transform on the images. We use total variation as regularization term to benchmark our results in the experiments, since it is widely used in compressed sensing based reconstruction methods.[7] The equation can be written as:

$$\vec{x}_j = \min_{\vec{x_j}} ||\vec{A}_j\vec{x}_j - \vec{y}_j||_2^2 + \beta||\vec{x}_j||_{TV}, j \leq n_t, \tag{4}$$

The total variation of a 2D image can be calculated using the following equation.

$$||\vec{X}||_{TV} = \sum_{a,b=1}^{N-1} |(\bigtriangledown(\vec{X})_{a,b})| \quad with \quad (\bigtriangledown(\vec{X})_{a,b}) = \begin{pmatrix} \vec{X}_{a,b} - \vec{X}_{a+1,b} \\ \vec{X}_{a,b} - \vec{X}_{a,b+1} \end{pmatrix} \tag{5}$$

Here $\vec{X}$ stands for the 2D image. (a, b) are its index in vertical and horizontal direction respectively. Total variation of a image is in fact the L1 norm of the image gradient. For simplicity, we call the method using the total variation as the regularizor TVR. Noting that, this formulation only favors the sparse solution in the spatial domain and a smooth image is actually a sparse image after the total variation transform.

2.2 PROPOSED METHOD (STTVR)

As mentioned above, we propose to make use of the data redundancy in temporal domain as well as in spatial domain. We treat the images of different phases as one entity and reconstruct them at the same time. The idea is to reconstruct the images by solving the following optimization problem:

$$\vec{x}_j = \min_{\vec{x}_j} ||\vec{A}_j\vec{x}_j - \vec{y}_j||_2^2 + \beta||\vec{x}_j||_{TV} + \alpha R(\vec{x}_j - \vec{x}_{j-1}) + \alpha R(\vec{x}_j - \vec{x}_{j+1}), j \leq n_t, \tag{6}$$

where α and β are weighting factors. We use total variation as the regularizer in both spacial and temporal domain as sparsifying transform. Then, the equation can be written as:

$$\vec{x}_j = \min_{\vec{x}_j} ||\vec{A}_j\vec{x}_j - \vec{y}_j||_2^2 + \beta||\vec{x}_j||_{TV} + \alpha||\vec{x}_j - \vec{x}_{j-1}||_{TV} + \alpha||\vec{x}_j - \vec{x}_{j+1}||_{TV}, j \leq n_t, \tag{7}$$

As in (4), the first term enforces the data fidelity and second term promotes the sparsity after the total variation transform. The third and the fourth term promote the sparsity in temporal domain. They try to make the difference of images from neighboring phases smooth, such that the motion or the changes of the pixels are similar to the ones of their neighboring pixels. Because of the smooth breathing pattern, it is very reasonable to promote a smooth motion field.

2.3 OPTIMIZATION METHOD

It is not easy to solve (7) directly because of the large dimension. Pan's group proposes an efficient method.[15] They minimize the different parts in (4) separately. The algorithm can be summarized as

- Step 1) $\vec{v}_j^k = \vec{x}_j^k - \gamma\vec{A}_j^T(\vec{A}_j\vec{x}_j^k - \vec{y}_j)\ \forall j$
- Step 2) $\vec{t}_j^k = \min_{\vec{x}_j} ||\vec{x}_j^k - \vec{v}_j^k||_2^2 + \beta||\vec{x}_j^k||_{TV}\ \forall j$
- Step 3) Repeat step 1 to step 2 until $||\vec{X}^{(k)} - \vec{X}^{(k+1)}||_2^2$ is less than a certain value or the maximum iteration number is reached.

Step 1 is one gradient descent step to minimize the first term ($||\vec{A}_j\vec{x}_j - \vec{y}_j||_2^2$) in the (4), where γ is a relaxing factor to control the step size. To further speed up the algorithm, they use one step of the standard ART update instead. They use the gradient descent method to minimize the objective function in step 2. The objective function takes the results from step 1 as a constraint and promotes the spatial sparsity. It tries to find a smooth image near the solution of step 1. Although they did not show the solid mathematical proof for the convergence of the optimization method, their numerical experiments show the stability of method.

Inspired by their work, we also minimize the cost function in (7) separately. The algorithm can be summarized as

- Step 1) $\vec{v}_j^k = \vec{x}_j^k - \gamma\vec{A}_j^T(\vec{A}_j\vec{x}_j^k - \vec{y}_j)\ \forall j$
- Step 2) $\vec{t}_j^k = \min_{\vec{x}_j} ||\vec{x}_j^k - \vec{v}_j^k||_2^2 + \beta||\vec{x}_j^k||_{TV}\ \forall j$
- Step 3) $\vec{x}_j^{k+1} = \min_{\vec{x}_j} ||\vec{x}_j^k - \vec{t}_j^k||_2^2 + \alpha||\vec{x}_j^k - \vec{x}_{j-1}^k||_{TV} + \alpha||\vec{x}_j^k - \vec{x}_{j+1}^k||_{TV}\ \forall j$
- Step 4) Repeat step 1 to step 3 until $||\vec{X}^{(k)} - \vec{X}^{(k+1)}||_2^2$ is less than a certain value or the maximum iteration number is reached.

Step 1 and step 2 are the same with the algorithm of Pan's group. In these two steps, the images at different phases are reconstructed individually, the redundancy in temporal domain is not taken into account. In step 3, we use the gradient descent method to minimize the objective function. The objective function considers the results from step 2 as a constraint and the images of neighboring phases are used to promote smooth changes or motions between each other. It tries to find a image which has a smooth motion pattern near the solution of step 2. We test the optimization method with different parameter settings to show the stability of the method in our experiments. The results can be seen in the next section.

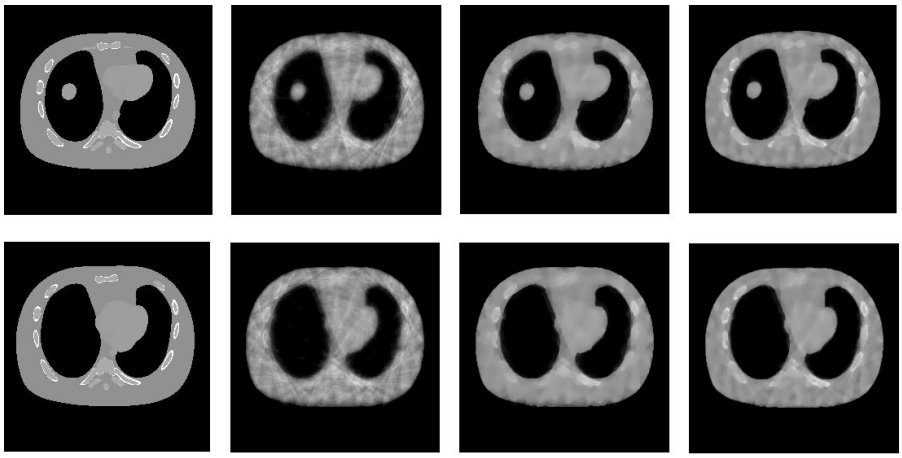

Figure 1. Reconstruction Results
The first row shows the image of phase 0 and the second row shows the image of phase 7. The first column is the ground truth. The second to the fourth column list the results reconstructed by standard ART, TVR and our method. Only 20 iterations are used for all the method.

3. EXPERIMENTS AND RESULTS

We use the digital XCAT thorax phantom in our experiments. The projections are generated in a fan beam geometry and are equally spaced throughout an entire 360 degree rotation. The image resolution is 256×256 and only 24 projections are used to reconstruct each phase. We reconstruct 15 phases in a breathing cycle and the reconstruction results are listed in Fig. 1. Two different phases are shown in the figure. The first column is the ground truth. The second to fourth column are images reconstructed by standard ART, TVR and our method respectively. The image quality of standard ART is not acceptable for clinical use due to the severe streak artifacts and motion artifacts. Although the TVR method improves the image quality dramatically compared to the results of standard ART, the blurring artifacts are still visible and the contrast of the image is reduced. The results from our method show much more details. The corresponding errors of all the method at each phases are listed in Table 1. The error is calculated by

$$Error = ||\vec{x}_j - \vec{x}_j^{true}||_2 \tag{8}$$

The same conclusion can be drawn from Table 1 too. Our method's reconstruction error are the smallest and less than half of the ones of standard ART. In the experiments, all the parameters were chosen for the optimal results. For our method, we use one standard ART update in step 1. Five gradient descent updates for step 2 and step 3. The maximal iteration number is set to 20. α and β are both 0.1 for the best performance. For the TVR method, the step 3 is ignored and the other parts are the same as in our method.

The convergence maps of TVR and our method are shown in Fig 2. In Fig 2, we set β to 0.1 for TVR and our method. α are set to two different values. First time we set α to 0.1 and second time we set α to 0.5. Clearly, step 3 which makes use of the correlation between neighboring phases does improve the image quality in every iteration step. Although we set different values to α, the convergence curves look similar which shows the stability of our optimization method. To further study the stability of our optimization method, we use different settings for our algorithm. The results are shown in Fig 3. The difference of results between different parameter settings is small, which implies the robustness of our optimization method.

4. CONCLUSION

In this paper, we have presented a novel 4D-CT reconstruction method based on compressed sensing. The method employs the redundancy in the temporal domain as well as in the spatial domain. The experiments indicate that our method outperforms the conventional reconstruction methods by using the correlation between the neighboring phases.

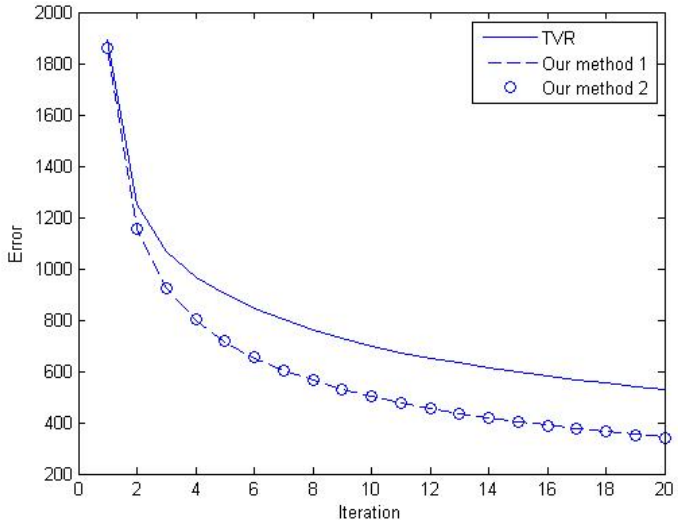

Figure 2. Converge map.

The only difference between our method and TVR is step 3 introduced in section 2. We set β to 0.1 for both TVR and our method. And we set α to two different values. In our method 1, α is 0.1 and in our method 2, α is 0.5. The only difference between TVR and our method is step 3. The figure shows that step 3 does improve the reconstruction quality at each iteration.

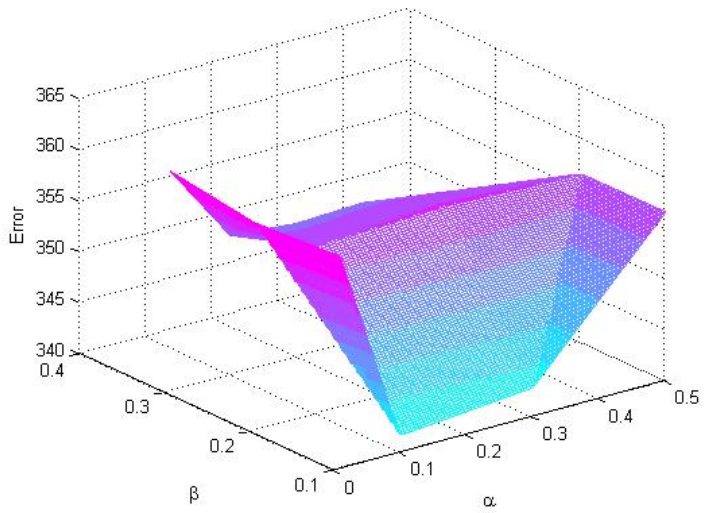

Figure 3. Varying α and β test.

We set parameter α and β for different values to test the robustness of our method. The range of α is [0.01 0.5] and the range of β is [0.1 0.3]. The figure shows that the difference of results for different parameter settings is not large, which implies that our method is quite stable.

Table 1. Error of different methods at all phases. Low value indicates a small error. Clearly, our method outperforms the others.

	ART	TVR	Our method
Phase 0	954.129	528.115	341.590
Phase 1	955.064	530.937	348.814
Phase 2	952.191	519.982	355.063
Phase 3	924.856	514.082	362.180
Phase 4	909.119	509.091	357.419
Phase 5	906.057	507.531	346.261
Phase 6	899.210	537.518	345.506
Phase 7	891.567	501.844	334.336
Phase 8	896.414	496.971	346.209
Phase 9	892.580	493.052	341.404
Phase 10	903.200	520.333	352.169
Phase 11	921.861	518.745	345.804
Phase 12	940.355	528.281	338.815
Phase 13	944.477	527.099	336.582
Phase 14	952.982	527.087	340.677
mean of all 15 phases	922.938	517.378	346.189

ACKNOWLEDGMENTS

The first and last authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) in the framework of the German excellence initiative. The first author thanks the financal support from Chinese Scholarship Council (CSC). The authors thank the financial support from NIH grant R01 HL087917.

REFERENCES

[1] E. Hansis, D. Schaefer, O. Doessel, and M. Grass, "Evaluation of iterative sparse object reconstruction from few projections for 3-d rotational coronary angiography," *IEEE Trans. Med. Imag.* **27**(11), pp. 1548 – 55, 2008.

[2] E. Ford, G. Mageras, E. Yorke, and C. Ling, "Respiration-correlated spiral ct: a method of measuring respiratory-induced anatomic motion for radiation treatment planning," *Med. Phys.* **30**(1), pp. 88–97, 2003.

[3] D. Low, M. Nystrom, E. Kalinin, P. Parikh, J. Dempsey, J. Bradley, S. Mutic, S. Wahab, T. Islam, G. Christensen, D. Politte, and B. Whiting, "A method for the reconstruction of four-dimensional synchronized ct scans acquired during free breathing," *Med. Phys.* **30**(6), pp. 1254–63, 2003.

[4] T. Pan, T. Lee, E. Rietzel, and G. Chen, "4d-ct imaging of a volume influenced by respiratory motion on multi-slice ct," *Med. Phys.* **31**(2), pp. 333–40, 2004.

[5] J. W. L. K. H. F. R. Lauritsch, G. Boese, "Towards cardiac c-arm computed tomography," *Medical Imaging, IEEE Transactions on* **25**(7), pp. 922 – 934, 2006.

[6] U. W. Langner and P. J. Keall, "Prospective displacement and velocity-based cine 4d ct," *Med. Phys.* **35**(10), pp. 4501–4513, 2008.

[7] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained total-variation minimization," *Phys. Med. Bio.* **53**, pp. 4777–4807, 2008.

[8] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**, pp. 1289 – 1306, 2006.

[9] R. Baraniuk, "Compressive sampling," *IEEE Signal Processing Magazine* **24**(4), pp. 118–121, 2007.

[10] E. Cands, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory* **52**(2), pp. 489 – 509, 2006.

[11] H. Wu, C. Rohkohl, and J. Hornegger, "Total variation regularization method for 3-d rotational coronary angiography," *Bildverarbeitung fr die Medizin* , pp. 434–438, 2011.

[12] Z. Tian, X. Jia, B. Dong, Y. Lou, and S. B. Jiang, "Low-dose 4dct reconstruction via temporal nonlocal means," *Med. Phys.* **38**(3), pp. 1359–65, 2008.

[13] W. P. Segars, "A realistic spline-based dynamic heart phantom," *IEEE Trans. Nucl. Sci.* **46**, pp. 503–506, 2011.

[14] G. L. Zeng, *Medical Image Reconstruction*, High Education Press, China, 2009.

[15] E. Y. Sidky, C. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam ct," *Journal of X-Ray Science and Technology* **14**, pp. 119–139, 2006.