Sparse-View 4D-CBCT Enhancement via Deep Learning with Integrated Interpolation and Motion Compensation

2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

Jie Zhang, PhD[1]; Lei Ren[1]
[1]University of Maryland School of Medicine, Baltimore, USA

ABSTRACT

Purpose: Four-dimensional cone-beam computed tomography (4D-CBCT) enables visualization of respiration-induced anatomical motion but suffers from degraded image quality due to sparse-view acquisition. This study proposes a deep learning-based approach combining interpolation and motion compensation (Interpolation+MoCo) to enhance sparse-view 4D-CBCT.
Methods: The Interpolation+MoCo consists of two stages: (a) Interpolation: an interpolation network that predicts full-view projections from sparse views using a U-Net backbone with a self-designed, angular-aware convolution that adaptively modulates kernel weights based on projection rotation angles and assigns diverse filters to different input patches. (b) MoCo: a deformable image registration (DIR) network that warps images from other phases to the target phase to mitigate undersampling artifacts. A total of 100 full-view 4D-CBCT projection sets from 20 patients were split into training (12 patients, 60 sets), validation (4 patients, 20 sets), and test (4 patients, 20 sets) cohorts. Each full-view set (Φ_{full}) was retrospectively downsampled from 2500 to 900 projections to generate sparse-view data (Φ_{sparse}). Network training consisted of two stages. (a) Interpolation network: the projections in Φ_{full}-Φ_{sparse} served as targets, each paired with its two adjacent projections in Φ_{sparse} as inputs, yielding 90190 and 30245 training and validation pairs respectively. (b) DIR network: trained in an unsupervised manner using source-target image pairs reconstructed as three-dimensional (3D) CBCT from the interpolated projection sets, with 5400 and 200 training and validation pairs respectively. The Interpolation+MoCo was tested on 20 4D-CBCT sets using root-mean-square error (RMSE), structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR).
Results: The proposed Interpolation+MoCo approach significantly enhance sparse-view 4D-CBCT quality with a RMSE of (3.25±0.51)×10^{-3} mm^{-1}, a SSIM of 0.9905±0.0061 and a PSNR of 35.09±1.65 dB, outperforming sparse-view+MoCo results (RMSE: (4.79±0.84)×10^{-3} mm^{-1}, SSIM: 0.9887±0.0055, PSNR: 31.73±1.10 dB).
Conclusion: The proposed Interpolation+MoCo approach enables high-quality 4D-CBCT from sparse views with shortened scanning-time and reduced dose.

CONTACT

Lei Ren
University of Maryland School of Medicine
655 W. Baltimore Street
Baltimore, MD 21201
LRen@som.umaryland.edu

INTRODUCTION

Four-dimensional cone-beam computed tomography (4D-CBCT) enables visualization of respiration-induced anatomical motion but suffers from degraded image quality due to sparse-view acquisition.

This study proposes a deep learning-based approach combining interpolation and motion compensation (Interpolation+MoCo) to enhance sparse-view 4D-CBCT, as shown in Figure 1.

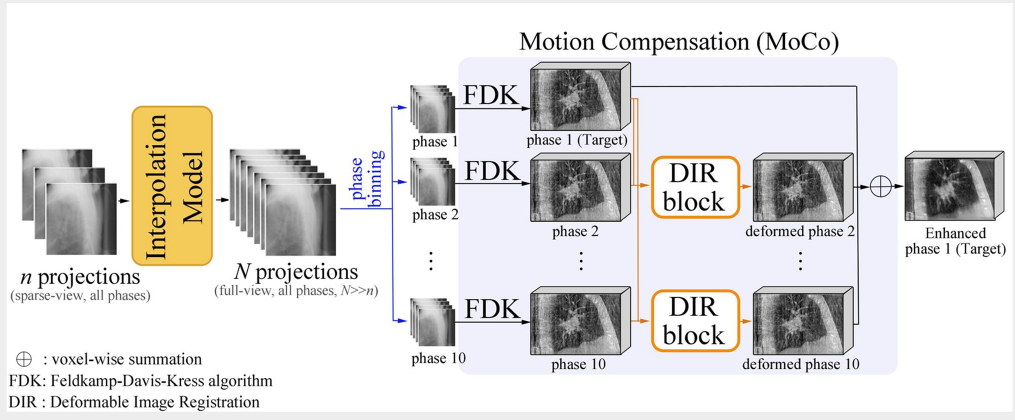

Figure 1. Illustration of the Interpolation+MoCo workflow.*
* In this figure, phase 1 is set to the target phase as an example. Adapted from [1]

METHODS AND MATERIALS

■ Model Overview

The Interpolation+MoCo consists of two stages:

(a) Interpolation: an interpolation network that predicts full-view projections from sparse views using a U-Net backbone with a self-designed, angular-aware convolution (as shown in Figure 2) that adaptively modulates kernel weights based on projection rotation angles and assigns diverse filters to different input patches.

(b) MoCo: a deformable image registration (DIR) network that warps images from other phases to the target phase to mitigate undersampling artifacts.

■ Data

100 full-view 4D-CBCT projection sets (Φ_{full}) from 20 real patients provided by The Cancer Imaging Archive (TCIA):

- Training data: 60 sets from 12 patients
- Validation data: 20 sets from 4 patients.
- Test data: 20 sets from 4 patients.

■ Training Strategy

(a) Interpolation network: each Φ_{full} was retrospectively downsampled from 2500 to 900 projections to generate sparse-view data (Φ_{sparse}). The projections in Φ_{full} - Φ_{sparse} served as targets, each paired with its two adjacent projections in Φ_{sparse} as inputs, yielding 90,190 and 30,245 training and validation pairs respectively.

(b) DIR network: 5,400 training and 200 validation source–target 3D CBCT pairs reconstructed from interpolated projections.

■ Performance Evaluation

The Interpolation+MoCo was tested using the following metrics:

- root-mean-square error (RMSE)
- structural similarity index measure (SSIM)
- peak signal-to-noise ratio (PSNR)

RESULTS

The proposed Interpolation+MoCo approach significantly enhances the quality of sparse-view 4D-CBCT reconstruction, as summarized in Table 1.

Figure 3 shows an example of the improvement achieved using the Interpolation+MoCo method.

Chart 1 provides a visual comparison of three metrics between the "Sparse-view+MoCo" and "Interpolation+MoCo" reconstructions, demonstrating the overall effectiveness of the proposed workflow.

Table 1. Quantitative analysis of the enhanced results of CBCT volumes*

	sparse-view+MoCo	Interpolation+MoCo
RMSE (×10^{-3} mm^{-1})	4.79±0.84	**3.25±0.51**
SSIM	0.9887±0.0055	**0.9905±0.0061**
PSNR (dB)	31.73±1.10	**35.09±1.65**

* The best results were printed in bold

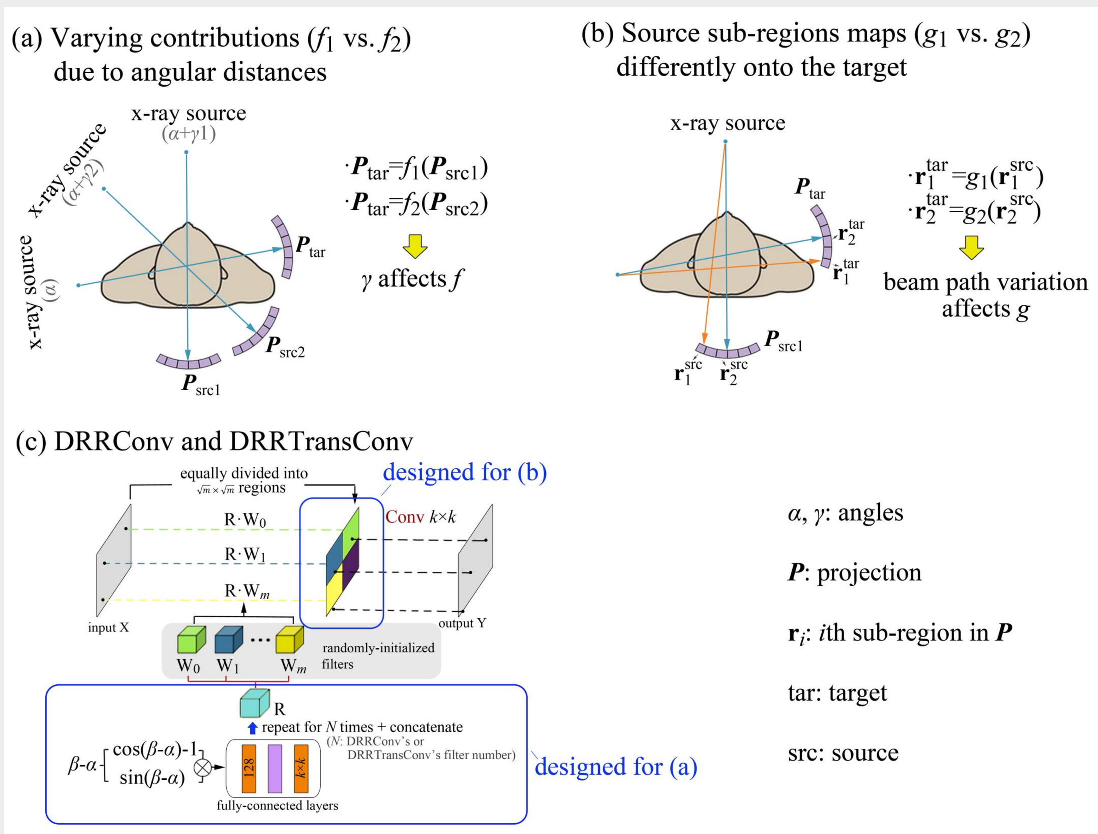

Figure 2. Illustrations of (a-b) rotation-induced data regularities and (c) customized rotation-aware convolutions (abbr. DRRConv and DRRTransConv) in the Interpolation Network. Adapted from [1].

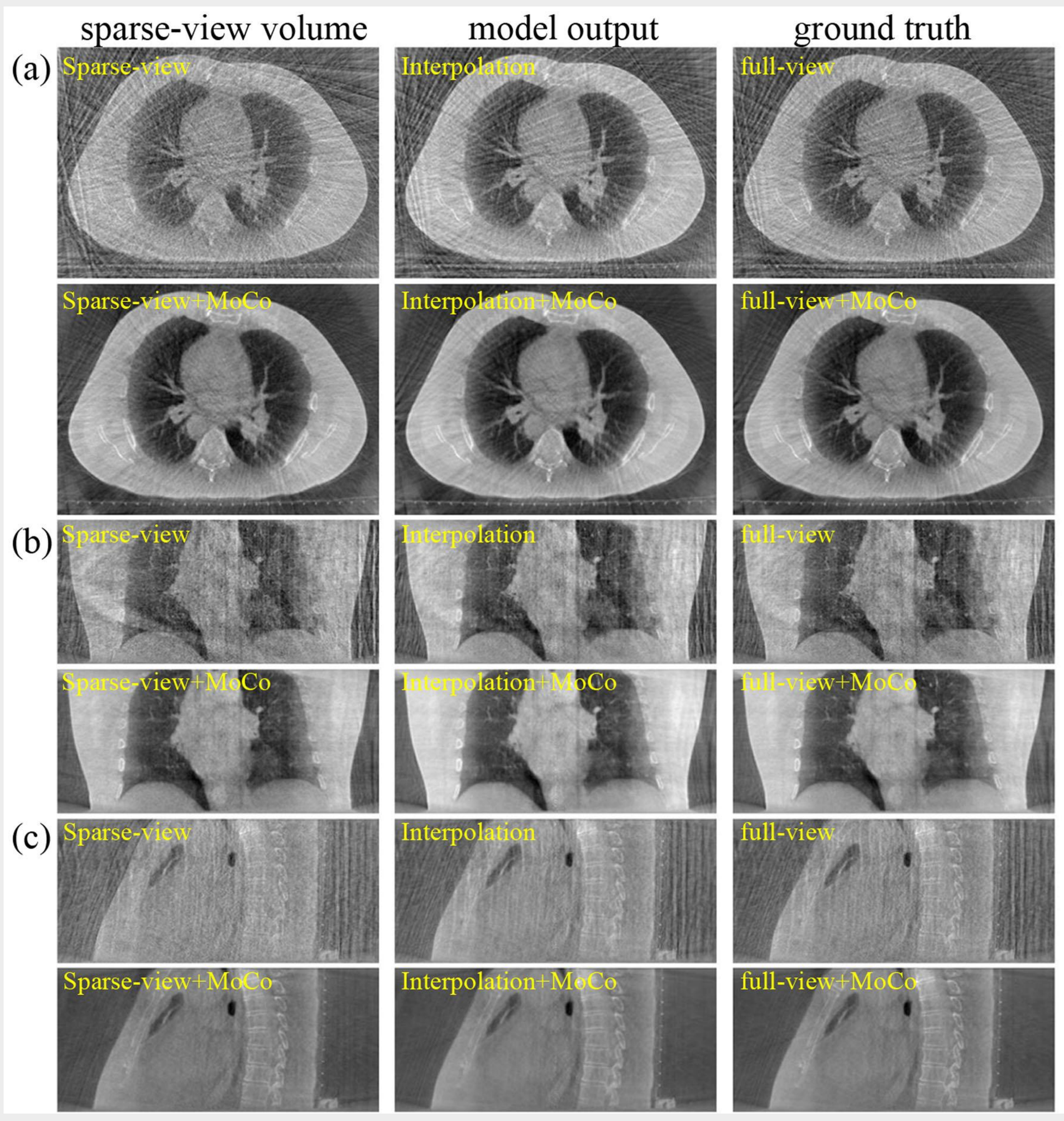

Figure 3. Qualitative example for Interpolation and Interpolation+MoCo. (a-c) exhibit transverse, coronal and sagittal planes respectively.

DISCUSSION

Innovation/Impact: The proposed Interpolation+MoCo approach introduces two innovations:

(1) a customized rotation-aware convolution in the interpolation stage for accurate projection estimation.

(2) in the MoCo stage, automated deformable registration with extrathoracic DVF zeroed out to better mitigate respiratory phase-sorting artifacts.

Specifically,

(1) Customized rotation-aware convolution: This convolution is designed to capture rotation-induced signal variations, as shown in Figure 2.

- It adaptively modulates kernel weights based on the projection angle (γ), since the mapping from the same projection ($\boldsymbol{u}_{\alpha}$) to different target projections ($\boldsymbol{u}_{\alpha+\gamma1}$ and $\boldsymbol{u}_{\alpha+\gamma2}$) varies with γ, as illustrated in Figure 2(a).
- It assigns diverse filters to different input patches to account for spatial variations in X-ray attenuation, even under the same rotation angle, as illustrated in Figure 2(b).

(2) Automated deformable registration with extrathoracic DVF zeroed out: DVFs are estimated from 4D-CBCT reconstructed using interpolated full-view projections, rather than from 4D-CT with mismatched acquisition timing or from sparse-view 4D-CBCT with degraded image quality that introduces registration uncertainty. Extrathoracic DVFs are zeroed out, consistent with the absence of respiratory motion outside the thorax.

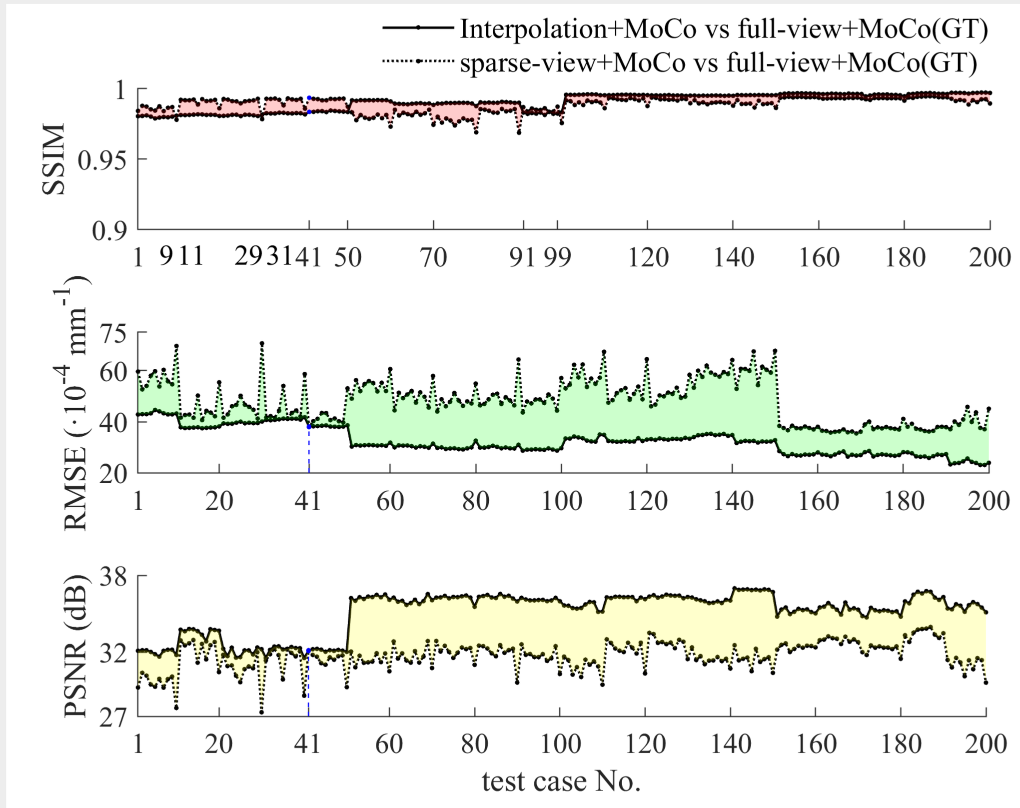

Chart 1. Comparison between Interpolation+MoCo and Sparse-view+MoCo reconstructions. Differences between the two methods are highlighted by the shaded regions. GT represents ground truth. Adapted from [1].

CONCLUSIONS

The proposed Interpolation+MoCo approach enables high-quality 4D-CBCT from sparse views with shortened scanning-time and reduced dose.

REFERENCES

1. J. Zhang, L. Ren, "Enhance Four-dimensional Cone-beam Computed Tomography (4D-CBCT) From Sparse View Acquisitions Using a Novel Deep Learning Model", Biomedical Signal Processing and Control, 119, 109935, 2026.