Diaphragm Tracking for Respiratory Motion Compensated Cardiac C-Arm CT

Marco Bögel*, Andreas Maier*, Hannes G. Hofmann*, Joachim Hornegger*,† and Rebecca Fahrig‡

***Abstract*—Long acquisition times of several seconds lead to image artifacts in cardiac C-arm CT. While ECG gating is able to select a certain heart phase, residual artifacts are mostly caused by respiratory motion. In order to improve image quality, it is important to accurately estimate the breathing motion that occurred during image acquisition. It has been shown that diaphragm motion is correlated to the respiration-induced motion of the heart.**

We present a motion estimation and compensation method based on the tracking of the diaphragm contour in projection space. The approach utilizes a 2-D quadratic curve model to estimate and track the diaphragm with sub-pixel accuracy. Based on the tracking results, we use a motion corrected triangulation algorithm to estimate the 3-D motion of the diaphragm top. The resulting signal is used to compensate for superior-inferior respiratory motion during the reconstruction.

***Index Terms*—C-arm CT; reconstruction; respiratory motion; diaphragm tracking; motion compensation;**

I. Introduction

Cardiac C-arm CT enables reconstruction of 3-D images during medical procedures. However, the long acquisition time of several seconds, during which the heart is beating and the patient might breathe, may lead to artifacts, such as blurring or streaks. A commonly used technique to reduce breathing motion is the single breath-hold scan. The physician instructs the patient to hold his breath after exhalation. The data is then acquired during the breath-hold. Although this approach is widely used, several studies have shown that breath-holding does not eliminate breathing motion entirely. Monitoring the position of the right hemidiaphragm during breath-hold, Jahnke et al. observed residual breathing motion to a certain extent in almost half of their test group [1]. Therefore, it is necessary to develop more sophisticated methods to estimate and compensate for respiratory motion in cardiac C-arm CT.

There are many ways to acquire respiratory signals. Most are based on additional equipment, e.g. Time-of-Flight or stereo vision cameras. Other techniques aim to extract the respiratory signal directly from the projection images. Using this approach the extracted breathing signal is perfectly synchronized with the projection images. Image-based respiratory motion extraction often relies on tracking of fiducial markers in the projection images [2], [3]. Wang et al. have shown that the motion of the diaphragm is highly correlated to respiration-induced motion of the heart [4]. Sonke et al. propose to extract a 1-D breathing signal by projecting diaphragm-like features on the superior-inferior axis and selecting the features with the highest temporal change [5]. However, the downside of this approach is that the extracted signal is not the real respiration signal. Due to perspective projection, the projected amplitude depends on the C-arm rotation angle.

In this work, we propose to estimate respiratory motion by tracking the diaphragm in a set of rotational projection images. The tracked position of the diaphragm top is used to compute a 1-D respiration signal, which is then incorporated into the reconstruction algorithm to compensate for respiratory motion.

*Pattern Recognition Lab, Universität Erlangen-Nürnberg, Germany
†Erlangen Graduate School in Advanced Optical Technologies (SAOT)
‡Dept. of Radiology, Lucas MRS Center, Stanford University, CA, USA

II. Methods and Materials

The proposed method is composed of three major steps that are each discussed in the following sections. In the first step, the contour of the diaphragm is tracked throughout the entire projection image sequence. Based on this tracking, we are able to obtain the 2-D projection of the diaphragm top for each image. In the second step, a motion corrected triangulation approach is used to compute the 3-D position of the diaphragm top for each projection. Assuming superior-inferior breathing motion, the 1-D respiration signal is extracted. In the final step, the respiration signal is used to compensate for respiratory motion during reconstruction.

A. Diaphragm Tracking

We introduced a model-based tracking method that is able to accurately track the contour of a user-selected hemidiaphragm in a set of rotational projection images [6]. Compared to other tracking-based methods, e.g. fiducial markers, the shape we want to track is not unique. The diaphragm appears as two similar shaped hemidiaphragms. Therefore, it is necessary for the user to select the one to be tracked. The user selects a point roughly located at the top of the desired contour. Subsequently, we define a rectangular Region of Interest (ROI) symmetrically around the selection. The image is then preprocessed using a gaussian low-pass filter and the Canny edge detector.

In the next step, the Random Sample Consensus (RANSAC) [7] is used to fit a parabolic curve to the obtained set of edge points. RANSAC can deal with datasets with large percentages of gross errors, and is thus the ideal choice to fit a model to our very noisy set of points. The aim of this method is to model the diaphragm as a quadratic function $v = au^2 + bu + c$, where u and v are the detector coordinates. Thus, RANSAC has to estimate the three parameters a, b, and c. In the first step, three random points are selected. The model estimation is then formulated as the following optimization

problem:

$$\sum_{i=1}^{3}(a \cdot u_i^2 + b \cdot u_i + c - v_i)^2 \rightarrow min. \quad (1)$$

A total of N models are estimated and evaluated to determine the best one. A model's quality is defined by the number of inliers. An inlier is a point that lies within a predefined distance to the model. Since an accurate model is desired, we only consider points with a one pixel distance to the model inliers. Assuming small motion between subsequent frames, the contour is tracked by calculating the current contour's vertex and using it as the start point in the subsequent frame.

One additional important optimization is made. Instead of continuing to use the rectangular ROI, we restrict it to a parabolic ROI based on the model from the previous frame. This approach decreases the number of points we have to consider in the model estimation.

To guarantee accurate tracking in projections where both hemidiaphragms are visible in the ROI, we propose additional constraints based on the small motion assumption and prior knowledge: (i) the horizontal motion of the contour is limited by the average motion, (ii) deformation of the contour is limited to 5% compared to the previous model, and (iii) the direction of horizontal motion can be derived from patient positioning and C-arm rotation.

B. Triangulation and Signal Extraction

The result of the diaphragm tracking is a parabolic model of the hemidiaphragm for each image. Our approach relies on the assumption that the projection of the 3-D diaphragm top coincides with the top of the 2-D diaphragm contour. However, this assumption is quite restrictive. Based on this assumption, we are able to reconstruct the 3-D position using multi view triangulation. However, triangulation algorithms are designed for static scenes and yield inaccurate results when used for dynamic scenes. For triangulation of dynamic scenes we propose the following four step process:

1) Select image pair
2) Rectification of the image planes [8]
3) Motion correction
4) Triangulation [9]

First, we select two images with the contour vertices $\tilde{\mathbf{g}} = (\tilde{g}_u, \tilde{g}_v, 1)^T$ and $\tilde{\mathbf{g}}' = (\tilde{g}'_u, \tilde{g}'_v, 1)^T$. Ideally, the selected images should be acquired from orthogonal views. The second step is essential for the subsequent motion correction. The rectification algorithm by Fusiello et al. transforms the image planes such that they become coplanar and their epipolar lines become parallel and horizontal [8]. The transformed images have then one very important feature: the projections of a point have the same vertical coordinate in both image planes. Thus, after transforming the point correspondences, any residual difference in their vertical coordinates must be caused by respiratory motion during image acquisition. Therefore, we can eliminate the respiratory motion of this image pair in the third step. We choose the first point $\tilde{\mathbf{g}}$ as the reference and the corresponding point in the second image is set to

$$\tilde{\mathbf{g}}' = (\tilde{g}'_u, \tilde{g}_v, 1)^T. \quad (2)$$

Algorithm 1: Motion compensated reconstruction. Respiratory motion is compensated in line 8.

```
1  forall the projections i ∈ [1, N_p] do
2      forall the voxels (x, y, z) do
3          Project voxel (x, y, z) onto detector plane
4          if point on detector plane then
5              Get update value
6          else
7              Next voxel
8          z_corr ← z + r̂_i
9          if (x, y, z_corr) in volume then
10             Update voxel (x, y, z_corr)
```

Finally, we use the transformed and motion corrected point correspondences to triangulate the corresponding 3-D point. In this work a simple iterative Linear-Eigen approach, as proposed by Hartley [9], has yielded excellent results.

After we triangulate a 3-D point corresponding to each image, we can now compute the respiration signal. Since respiratory motion is generally considered as a mainly translational motion along the superior-inferior axis, we compute the 1-D respiration signal $\hat{\mathbf{r}}$ as

$$\hat{r}_i = z_{ref} - z_i, \quad (3)$$

with z_{ref} as the z-coordinate of the reference point, and z_i as the z-coordinate of the triangulated point corresponding to image i. Finally, the resulting signal is smoothed using a gaussian low-pass filter.

C. Motion Compensated Reconstruction

The signal is now included in the reconstruction process. Algorithm 1 shows the motion compensated reconstruction algorithm. For each projection, each voxel is projected on the detector to get the update value. Instead of regularly updating the volume, we first compensate for respiratory motion by shifting the voxel back to its reference position using the estimated signal. Then, we update the corrected voxel. Therefore, we are able to obtain a reconstruction at the reference time we selected for the respiration signal. The proposed method assumes a constant shift for the whole heart. For clinical data more sophisticated motion models are required, as the deformation of the heart is not rigid.

III. Experimental Results

The evaluation of this work was carried out on a simulated XCAT phantom [10]. The XCAT phantom was created with breathing motion only. We simulated an acquisition time of four seconds with one full respiration cycle. Both heart and diaphragm moved about 2.3 cm along the superior-inferior axis, the rest of the scene was static. A detector of size 640×480 px was simulated with a resolution of 0.616 mm/px. 200 projections were acquired with an average angular increment of $1.0°$. As ground truth we used the reconstruction of an XCAT dataset that was simulated without respiratory motion.

Fig. 1. Comparison of the extracted diaphragm motion signal and the actual breathing signal. The amplitude of the signal can not be estimated accurately, as the projections of the diaphragm top do not coincide with the 2-D contour.

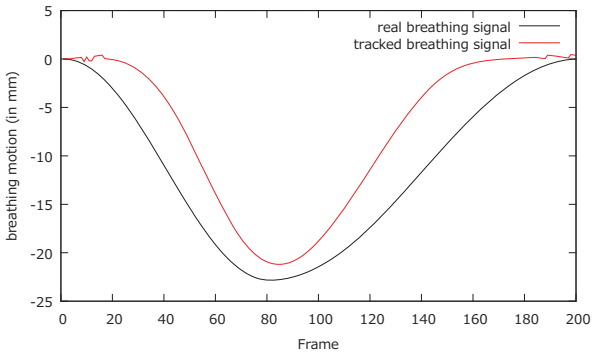


TABLE I
TRIANGULATION ERRORS (IN MM) BASED ON PROJECTIONS OF THE REAL DIAPHRAGM TOP. ANGULAR OFFSET OF THE IMAGE PAIR IN BRACKETS.

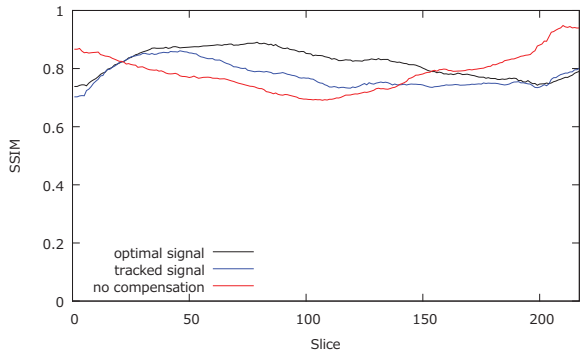
	Mean 3-D	Std. Dev. 3-D	Mean Z	Std. Dev. Z
Rect. Iter. (90°)	0.20	0.06	0.10	0.06
Rect. Iter. (30°)	0.32	0.15	0.10	0.06
Rect. Iter. (10°)	0.89	0.60	0.11	0.08
Iterative (90°)	2.22	0.97	2.22	0.96

The diaphragm tracking method was evaluated on the left and right hemidiaphragms in XCAT projection data [6]. We were able to track the vertex of the diaphragm contour with sub-pixel accuracy. We observed a Euclidean distance of the right vertex to the correct vertex of 0.45 ± 0.56 pixels, and 0.75 ± 0.84 pixels for the left vertex respectively.

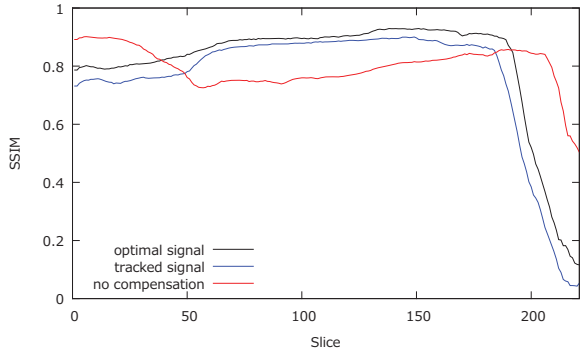
Figure 1 shows the extracted signal based on the diaphragm tracking results. As previously noted, our approach depends on the assumption that the projection of the diaphragm top lies on the 2-D contour. However, this is a strong assumption that is not always fulfilled. In fact, the correct projection of the diaphragm top is often located below the contour, due to perspective projection. This results in inaccuracies in the estimated amplitude of the signal, caused by triangulation with false point correspondences. In order to assess the accuracy of the triangulation approaches without the effect of false point correspondences, we tested the methods using the correct projections of the diaphragm top as input. Therefore, we can test the performance of our algorithm if the assumption is met. As results in TABLE I show, our rectified iterative approach provides sub-millimeter accuracy even for image pairs with low angular offset, whereas the average error of the standard approach without rectification and motion correction is about 10% of the total breathing motion.

For the evaluation of reconstruction quality we used the structural similarity index (SSIM) by Wang et al. [11]. SSIM measures the similarity of two images based on structural information. Two images are compared and a value between -1.0 and 1.0 is returned, with 1.0 for a perfect match and -1.0 for completely different images. In order to reduce the influence of the static background on the quality evaluation, the reconstructed volume was cropped to the bounding box that contains the heart. In total, we evaluated the quality of three different reconstructions: (i) a compensated reconstruction using the proposed tracking methods, (ii) a compensated reconstruction with the correct 2-D projections of the diaphragm top (simulating an optimal diaphragm tracking), and (iii) an uncompensated reconstruction. Figure 2 shows the evaluation results for xy and xz-slices. Both compensated reconstructions show highly improved image quality. As expected, the diaphragm tracking approach is slightly below the quality of the optimal reconstruction. However, it shows significant improvement when compared to the uncompensated reconstruction. The uncompensated reconstruction seems to be superior in the first and last slices. The heart is only of small size in these slices. Therefore, the static background has a larger influence on the evaluation. Compensation blurs the static background, whereas it is perfectly reconstructed without compensation. Figures 3 and 4 show the results for two example slices.

Fig. 2. Structural similarity index of the heart volume for xy and xz-slices. The uncompensated reconstruction shows better results in the beginning and the end, as the heart is only of small size in these slices.



(a) xy-slices



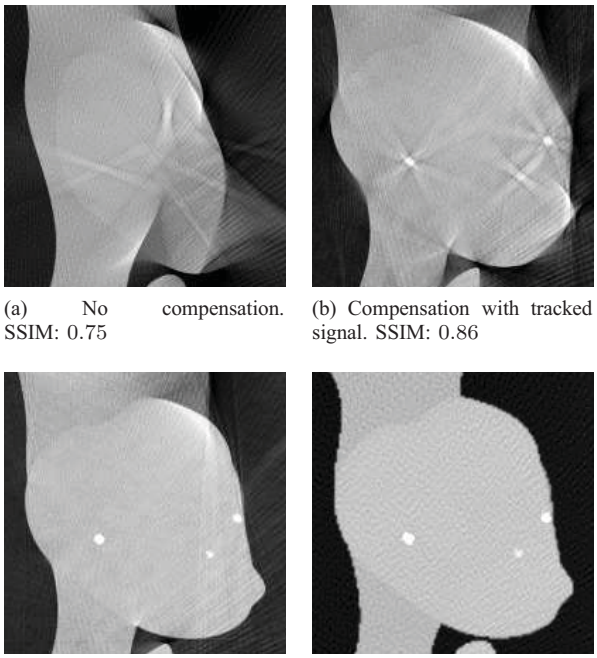
(b) xz-slices

IV. CONCLUSION AND OUTLOOK

Results of respiratory motion compensated reconstruction already show promising results in image quality improvement close to the optimal solution. Still, there is one important issue that has to be solved. The proposed method works on the very restrictive assumption that the projection of the 3-D diaphragm top coincides with the top of the 2-D contour.

Future work will be focused on this problem. Due to perspective projection, this assumption is not always fulfilled.

Fig. 3. Comparison of xy-slice 70 of compensated and uncompensated volumes (cf. Fig. 2a). Simulated high-contrast heart lesions further illustrate the improved image quality.



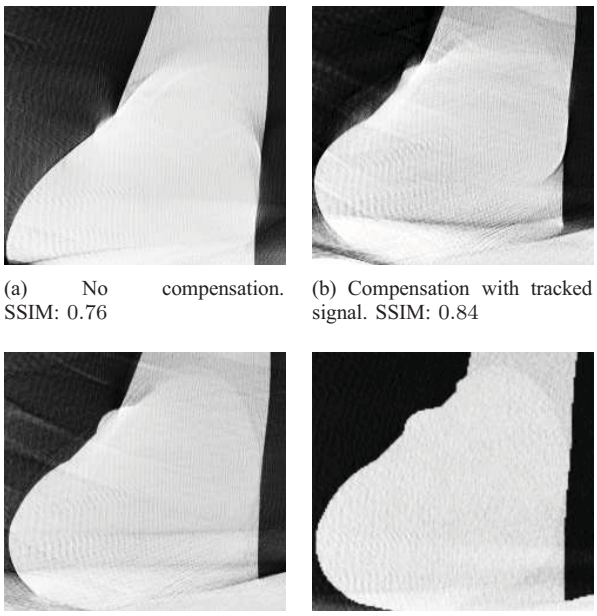
(a) No compensation. SSIM: 0.75

(b) Compensation with tracked signal. SSIM: 0.86

(c) Compensation with optimal signal. SSIM: 0.90

(d) Ground truth

Fig. 4. Comparison of xz-slice 60 of compensated and uncompensated volumes (cf. Fig. 2b).



(a) No compensation. SSIM: 0.76

(b) Compensation with tracked signal. SSIM: 0.84

(c) Compensation with optimal signal. SSIM: 0.88

(d) Ground truth

This results in inaccurate amplitudes of the extracted respiration signal. One interesting approach would be to combine the current approach with other existing methods to estimate respiratory motion, e.g. Time-of-Flight cameras. This way, we could possibly improve the extracted diaphragm motion signal by correlating it to the externally measured respiration signal.

Acknowledgement

This work is supported by NIH grant R01 HL087917.

References

[1] C. Jahnke, I. Paetsch, S. Achenbach, B. Schnackenburg, R. Gebker, E. Fleck, and E. Nagel, "Coronary MR Imaging: Breath-hold Capability and Patterns, Coronary Artery Rest Periods, and beta-Blocker Use," *Radiology*, vol. 239, no. 1, pp. 71–78, 2006.

[2] S. Wiesner and Z. Yaniv, "Respiratory Signal Generation for Retrospective Gating of Cone-Beam CT Images," *Proc SPIE*, vol. 6918, pp. 691 817–1 – 691 817–12, 2008.

[3] T. Marchant, G. Price, B. Matuszewiski, and C. Moore, "Reduction of motion artefacts in on-board cone beam CT by warping of projection images," *Br J Radiol*, vol. 84, no. 999, pp. 251–264, March 2011.

[4] Y. Wang, S. Riederer, and R. Ehman, "Respiratory motion of the heart: Kinematics and the implications for spatial resolution in coronary imaging," *Magn Reson Med*, vol. 33, pp. 716–719, 1995.

[5] J.-J. Sonke, L. Zijp, P. Remeijer, and M. van Herk, "Respiratory correlated cone beam CT," *Med Phys*, vol. 32, no. 4, pp. 1176–1186, 2005.

[6] M. Bögel, A. Maier, H. G. Hofmann, J. Hornegger, and R. Fahrig, "Diaphragm Tracking in Cardiac C-Arm Projection Data," *Proc BVM*, pp. 33–38, 2012.

[7] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun ACM*, vol. 24, pp. 381–395, June 1981.

[8] A. Fusiello, E. Trucco, and A. Verri, "A compact algorithm for rectification of stereo pairs," *Mach Vis Appl*, vol. 12, no. 1, pp. 16–22, 2000.

[9] R. I. Hartley and P. Sturm, "Triangulation," *Comput Vis Image Underst*, vol. 68, no. 2, pp. 146–157, 1997.

[10] W. P. Segars, M. Mahesh, T. J. Beck, E. C. Frey, and B. M. W. Tsui, "Realistic CT Simulation Using the 4D XCAT Phantom," *Med Phys*, vol. 35, no. 8, pp. 3800–3808, 2008.

[11] 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, 2004.