# Data-driven Correction for Head Motion In Helical X-ray CT

Tao Sun, Jung-Ha Kim, Roger Fulton, Johan Nuyts

***Abstract*— Although current CT systems can scan the head in a very short time, patient motion sometimes still induces artifacts. If motion occurs, one has to repeat the scan; to avoid motion, sedation or anesthesia is sometimes applied. We propose a data-driven method to iteratively correct this motion together with the reconstruction. In every iteration, we estimate the motion view-by-view, which then can be used to update the system matrix used during reconstruction. A multi-resolution scheme was used to speed up the convergence of the joint estimation of the motion and reconstruction. The method was evaluated on simulations and on real patient scans. The quality of the reconstructed images was improved substantially after the correction. The proposed method eliminated motion-induced artifacts in head CT scans.**

***Index Terms*—Computer Tomography (CT), iterative reconstruction, rigid motion, data-driven, motion correction.**

## I. INTRODUCTION

A slight movement of the patient will lead to a reduction of spatial resolution in CT, in severe cases resulting in corrupted images unsuitable for diagnosis or further processing. To reduce the likelihood of motion artifacts, CT manufacturers have made scanners faster by increasing the number of detector rows and the rate of rotation of the x-ray source and detector. Other ways to reduce the patient motion include general anesthesia, sedation and the use of restraining devices for head and neck imaging [1].

Despite of the effectiveness of minimizing the motion beforehand, assessment of the subject motion is of considerable general interest in tomography. A variety of methods for the estimation of motion in CT exist, including direct motion estimation using a camera system with visual markers [2] or without markers [3]. Indirect estimation methods have been proposed where motion is estimated through the minimization of errors in consistency conditions [4], or iteratively updating the motion together with the reconstruction process [5]. Once motion parameters have been estimated, the motion can be corrected for either in the measured raw data or during the reconstruction process.

While most of the retrospective methods in CT imaging addressed the problem for circular cone beam CT (CBCT), relatively few studies has been done for helical CT. Motion correction is arguably simpler in CBCT since the entire object will normally be in the field of view at all times. In contrast, in helical CT the object is always truncated in axial direction, which provides limited information to restore the consistency in projections.

In this study, we propose an approach to reduce or eliminate the motion artifacts in helical-CT reconstruction. The approach is based on our previous work [4]. The correction only needs the measured raw data, hence it is called data-driven approach. We performed simulations and a patient study to validate the proposed approach, comparing reconstructions with and without motion correction.

## II. METHOD

### A. Coordinate system

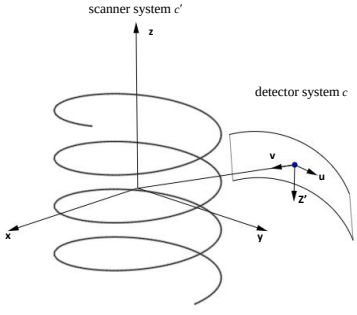
A clinical helical CT system usually has a curved detector. We define a coordinate system $c = (x, y, z) \in \mathbb{R}^3$ in Fig. 1, which is fixed with respect to the scanner, its z-axis coincides with the rotation axis of the scanner. Detector coordinate system $c' = (u, v, z') \in \mathbb{R}^3$ is fixed with respect to the rotating source-detector system: its origin is in the center of the detector, $z'$ is parallel to $z$, $u$ is tangent and $v$ is orthogonal to the detector. For one projection view, we define the rigid motion transform in the coordinate system $c$:

$$S_{world} = (\phi_x, \phi_y, \phi_z, t_x, t_y, t_z)^{\mathrm{T}} \qquad (1)$$

where $\phi_x$, $\phi_y$, $\phi_z$ are the 3 rotations, $t_x$, $t_y$, $t_z$ are the 3 translations in the world coordinate system. The transform can be mapped in the coordinate system $c'$:

$$S_{\mathrm{det}ector} = (\phi_u, \phi_v, \phi_z, t_u, t_v, t_z')^{\mathrm{T}} \qquad (2)$$

where $\phi_u$, $\phi_v$, $\phi_z$ are the rotations $t_u$, $t_v$, $t_z'$, are the translations in the detector coordinate system. A small motion $t_v$ in the direction perpendicular to the detector results in a very small magnification of the projection, which we assume




Manuscript received Mar. 31, 2016. This work was supported in part by the IWT MIRIAD SBO project and IMIR PF-project of KU Leuven.

T. Sun and J. Nuyts are with the Medical Imaging Research Center, Department of Nuclear Medicine, KU Leuven, Belgium (tao.sun@uzleuven.be).

J. Kim was with the Discipline of Medical Radiation Sciences, University of Sydney, Australia, and now is with the Faculty of Medicine, University of Sydney, Australia.

R. Fulton is with the Discipline of Medical Radiation Sciences, and School of Physics, University of Sydney, and the Department of Medical Physics, Westmead Hospital, Westmead, Australia.

Fig. 1 The scanner and detector coordinate systems on which motion correction is based.

to be negligible [5]. In every projection view, then, we set $t_v$ as zero and only 5 parameters need to be estimated in our correction scheme $\phi_u, \phi_v, \phi_z, t_u$ and $t_z'$:

$$S_{\det ector} = (\phi_u, \phi_v, \phi_z, t_u, t_z')^{\mathrm{T}} \tag{3}$$

The estimated motion in the detector coordinate system can later be transformed to the motion in the world coordinate system, as the reconstruction requires a transform in the world coordinate system:

$$\begin{aligned} S_{\det ector} &\rightarrow T_{\det ector} \\ T_{world} &= R \cdot T_{\det ector} \end{aligned} \tag{4}$$

where $T$ is the 4×4 homogenous matrix representation of the transform, $R$ is the 4×4 transformation matrix that maps the detector coordinate system to world coordinate system.

### B. OSEM reconstruction

In the presence of object motion, the helical CT-orbit is distorted into an effective orbit with arbitrary shape. Since this is problematic for analytical reconstruction, an iterative reconstruction algorithm is needed. We used Ordered Subset Expectation Maximization (OSEM) as the reconstruction algorithm [6]. We used the OSEM-algorithm for convenience, but if the use of a better noise model would be required, it can be replaced with a dedicated iterative algorithm for transmission tomography.

### C. General motion correction scheme

The following describes the basic idea of the data-driven motion correction: motion-corrected reconstruction and motion were alternately updated to increase the likelihood, the iterations were stopped when the motion estimate seemed to have converged (Fig. 2). The implementation involves: (1) a 2D-3D image registration to update the motion estimate for each view at the current iteration; (2) an image update with iterative reconstruction, incorporating the updated motion in the system matrix; (3) alternate updates of both image and motion with a multi-resolution scheme; (4) final reconstruction with a system matrix based on the last motion estimate. The following 4 paragraphs discuss the details about each part of the framework.

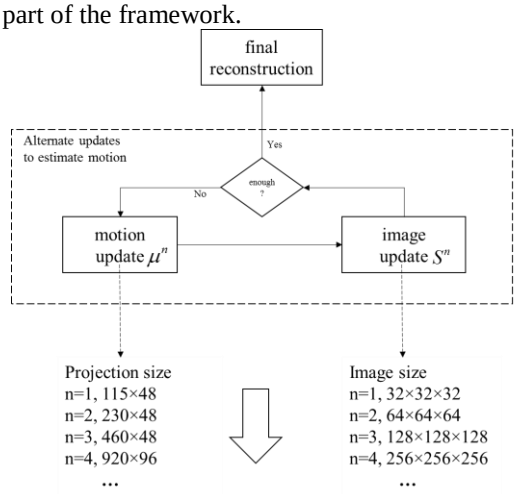



Fig. 2. General motion estimation scheme. $\mu$ is the update of the attenuation image, $s$ is the update of the rigid transform, $n$ is the iteration number. A multi-resolution scheme was applied to the motion estimation, increasing the sampling as the iteration number.

#### 1) Motion update

From the measured raw data an initial image is reconstructed. This image can be reconstruction produced with the system software (postprocessed to convert Hounsfield units back to attenuation integrals), or a first iterative reconstruction (Eq. 5) from the measured data.

For one projection line $i$, we integrate along the projection line to define the forward projection of the current estimate $\mu$:

$$f_i = \sum_j a_{ij} \mu_j \tag{5}$$

In helical CT, we organize the line integrals in views, where view θ contains all line integrals associated with a single source position:

$$f_\theta = \{ f_i \} \tag{6}$$

For each view, the 5 motion parameters are estimated one after the other. Suppose the general motion correction scheme (Fig. 2) is at the iteration n, hence the current motion update is $s^n$. Assuming that the motion represented by the (rotation or translation) parameter $\hat{s}$ is small, the derivative of projection $f$ with respect to $\hat{s}$ can be approximated as a finite difference of the intensities:

$$\frac{\partial f_\theta}{\partial s} \approx \frac{f_\theta(\hat{s}) - f_\theta(s^n)}{\hat{s}} \tag{7}$$

where $f_\theta(s^n)$ is the calculated re-projection (using the current estimates for the image and motion) and $f_\theta(\hat{s})$ is the measured projection for view $\theta$. To estimate $\hat{s}$ in Eq. 7, we need to know the derivative on the left hand side, hence we introduce another equation which is very similar to Eq. 7:

$$\frac{\partial f_\theta}{\partial s} \approx \frac{f_\theta(\Delta s) - f_\theta(s^n)}{\Delta s} \tag{8}$$

where $\Delta s$ is known small increment of the parameter to be estimated. When $\Delta s$ represents a translation, $f_\theta(\Delta s)$ can be approximated as simple translation of the current re-projection $f_\theta(s^n)$; for in-plane rotation, again $f_\theta(\Delta s)$ can be approximated as a simple rotation of the re-projection $f_\theta$, as shown in Fig.3. For the two out-of-plane rotations, we calculate $f_\theta(\Delta s)$ with a forward projection using a system matrix adjusted with $\Delta s$.

Eq. 7 and 8 assume that a small increment of one degree-of-freedom rigid motion only results in a linear change of the intensities in the projection. All above lead to a least squares minimization problem for current view at current iteration :

$$s^n_{incre} = \arg\min_{\hat{s}} \left\| \begin{matrix} \Delta s \cdot \left[ f_\theta(s^n + \hat{s}) - f_\theta(s^n) \right] - \\ \hat{s} \cdot \left[ f_\theta(s^n + \Delta s) - f_\theta(s^n) \right] \end{matrix} \right\|^2 \tag{9}$$

To find $s^n_{incre}$, Eq. 9 was solved analytically. Defining

$$\begin{aligned} P_\theta &= f_\theta(s^n + \hat{s}) - f_\theta(s^n) \\ Q_\theta &= f_\theta(s^n + \Delta s) - f_\theta(s^n) \end{aligned} \tag{10}$$

and setting the derivative of the Eq. 9 with respect to $\hat{s}$ to zero, one obtains:

$$s_{incre}^{n} = \frac{\sum_{N} P_\theta \cdot \sum_{N} Q_\theta}{\left\| \sum_{N} Q_\theta^{\,2} \right\|} \cdot \Delta s \qquad (11)$$

where $N$ is total number of voxels in one projection view $\theta$.

This procedure was applied to estimate all five parameters in Eq. 3. The sequence of the estimation is translation first, then rotation. The newly estimated parameter values were used immediately when estimating the value of next parameter. This sequential estimation of the five motion parameters for all projection views completes the update of the rigid motion in the current iteration. Then the rigid motion parameters for each projection view were transformed into a homogenous matrix in the detector coordinate system (Fig.1). Applying the Eq. 4, the matrix was mapped into world coordinate system. The transformation matrix obtained in the current iteration (*n*) was then used to update the previous motion estimate for every view, which will be used in the next iteration (*n+1*):

$$\begin{aligned} S_{incre}^{n} &\rightarrow \Delta T_\theta^{n} \\ T_\theta^{n+1} &= T_\theta^{n} \cdot \Delta T_\theta^{n} \end{aligned} \qquad (12)$$

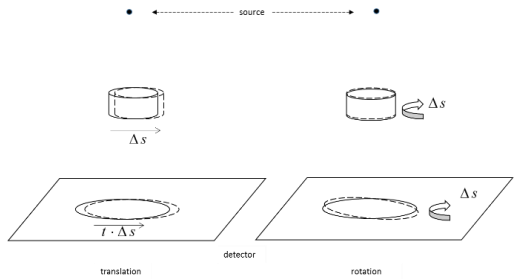



Fig. 3 Cartoon illustrating that the effect of object translation or rotation parallel to the detector can be well approximated as translation and rotation of the projection. For simplicity, the curvature of the detector was ignored. In the figure at left, t is the magnification factor from the object to detector.

*2) Image update*

After obtaining the motion, the image representing the attenuation coefficients can be updated with iterative reconstruction. We used OSEM as the reconstruction algorithm.

Instead of moving the reconstruction image, rigid motion correction was done by considering a coordinate system fixed to the object and incorporating the motion (now associated to the source-detector pair) into the system matrix. This corresponds to an arbitrary 3-dimensional (3D) motion of the virtual gantry around the object being scanned [2]. Motion correction was enabled by introducing a modified version of standard OSEM:

$$\begin{aligned} T_\theta^{n+1} &= invert(T_\theta^{n+1}) \\ \mu_j^{n+1} &= \frac{\mu_j^{n}}{\sum_i T_i^{n+1}(a_{ij})} \sum_i T_i^{n+1}(a_{ij}) \frac{m_i}{\sum_k T_i^{n+1}(a_{ij}) \mu_j^{n}} \end{aligned} \qquad (13)$$

where $i$ is the projection line index, $j$ is the voxel index, $a_{ij}$ is the effective intersection length of line $i$ with voxel $j$, m is the log converted sinogram, $S_b$ is one subset of projections of b subsets. $T_i$ is a $4 \times 4$ rigid transformation matrix applied to the projection line $i$. If $T_i$ is the identity matrix for all projection lines, then Eq. 13 is the same as original OSEM. In helical CT, $T_i$ is constant for all projection lines in one projection view. Because of the high rotation speed and the large number of views, the motion within a single view is negligible.

Distance-driven projectors were used for interpolation during the (back) projection [7]. The new estimation of the attenuation from step 2 can then be used for next motion update step (step 1).

*3) Multi-resolution alternate updates*

By repeating steps 1 and 2 we can estimate the reconstruction (Eq. 13) and motion (Eq.14) alternately. Because the correction of the image and the correction of the transform parameters are jointly estimated from the measured data, the problem of error propagation is mitigated.

An approach to reduce the computation time is to apply a multi-resolution techniques. We utilize this technique by running the algorithm with a coarse to fine representation of the raw data and the image. As in Fig. 2, the image update is reconstructed at coarse resolution at early iterations, while the resolution increases as the iteration numbers increase. Similarly, the projections in Eq. 10 are computed with gradually increased resolution. An possible additional advantage of the multi-resolution scheme is that it may help to avoid convergence to an undesired local maximum. Since these computations are the most expensive ones in the multi-resolution scheme, we stopped the scheme at the one but finest resolution.

As proposed in [2], the motion estimates were smoothed (by filtering each component independently) to remove outliers. We chose the Savitzky-Golay filter [8] to do the smoothing right after every motion update. We found that selection of the 15-point and 201-point kernel can achieve satisfactory jitter suppression, in simulations and patient scans respectively.

It is not obvious to define good stopping criteria for the motion estimation, especially considering the ground truth image is missing for the clinical data. In the motion estimation scheme, a maximum number of iterations was chosen for each resolution level. In addition, the summation of projection errors between the re-projected and measured data over all the views was computed, and at each resolution level, the iterations were stopped earlier when the relative change of this error measure did not exceed 5% of the summed error between last re-projected and measured data.

*4) Final reconstruction*

When the motion estimate has converged, a final reconstruction of diagnostic quality must be produced. In simulation studies we start the final reconstruction with the last image update from the alternate updates. To achieve a similar speedup in the clinical study, we start the final reconstruction from an approximate helical Feldkamp-Davis-Kress reconstruction (motion correction was enabled in the backprojection step), provided the image is not affected much by the motion artifacts.

To further accelerate the final reconstructions, the forward and backward projection operations were implemented in OpenCL and run on a GPU (NIVIDIA Tesla C2075).

## III. Experiments and Results

### A. Simulation

In the simulation, a segment of measured motion from a volunteer was applied to a voxelized phantom to generate a simulated scan. Reconstructions from this scan were analyzed quantitatively to assess the performance of the motion correction algorithm. The phantom is a 3D voxelized phantom from the Visible Human Project [9]. The unit was converted from Hounsfield (HU) to attenuation coefficient ($cm^{-1}$) at peak kilovoltage of 70 kVp. The image size was $256 \times 256 \times 240$, pixel size was $1 \times 1 \times 1$ $mm^3$.

A helical scan was simulated as being scanned with a Siemens Definition AS CT scanner (Siemens Medical Solutions USA, Inc., Malvern, PA), with reduced angular sampling to reduce the computation times. The crucial parameters were: angles per rotation 250, pitch 1.0, collimation $32 \times 1.2$ mm. The motion was applied to the phantom for the simulated helical scan. To avoid the cone-beam artifacts, all simulated helical scans covered a bit more than the entire object.

OSEM was used for all reconstructions, with motion correction enabled (Eq. 13). During the joint estimation of the attenuation image and the motion, the attenuation image was updated using 1 OSEM iteration with 40 subsets. Reconstruction pixel size was $1 \times 1 \times 1$ $mm^3$ at the finest resolution. Alternate updates of both image and motion were performed within a multi-resolution scheme to obtain the optimal motion. For the final reconstruction, 4 iterations and 60 subsets were applied. Fig. 4 and Fig. 5 show the estimated reconstruction and motion.

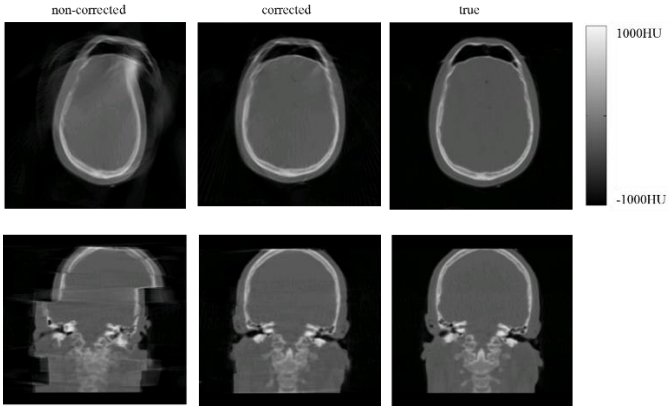



Fig. 4. Selected transaxial and coronal slices from reconstructions without and with motion correction, and also from the true image.

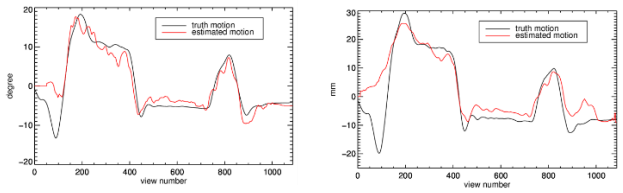



Fig. 5. The estimated motion values as a function of the view angle for the two most prominent motion parameters. Left: $\phi_x$, right: $t_x$.

### B. Real scan

The method has been applied to clinical studies in which motion artifacts had been observed. The outcome was evaluated by assessing the image visually.

The anonymized raw data of one patient who had previously undergone a head CT scan in the Department of Radiology at Westmead Hospital, Sydney, Australia, were collected with the approval of the Human Research Ethics Committee of the Western Sydney Local Health District. The scan was performed on a Siemens Force scanner (pitch 0.55, tube voltage 120 kVp, tube current 150 mAs, angles per rotation 1050, collimation $64 \times 0.6$ mm). Flying focus was turned on in both *z* and *phi* directions.

Because of the huge size of the raw data, we read and average every 8 projection of them for the motion estimation. This accelerated both the motion updates and the image updates. We used OSEM as the reconstruction algorithm. The motion correction was enabled for all reconstructions. Unlike in the simulation, 2 OSEM iterations with 40 subsets were done for the image updates. The final reconstruction pixel size was $0.40039 \times 0.40039 \times 0.75$ $mm^3$, image size is $512 \times 512 \times 436$. To accelerate the motion estimation, the multi-resolution scheme was applied as in II. C. Stopping criteria were also described in II. C. For the final reconstruction with motion correction, the starting image was set as the image from helical FDK reconstruction. Six OSEM iterations with 60 subsets were applied in combination with Nesterov's acceleration [10] on the GPU. Fig. 6 shows the original reconstructed image and motion-corrected reconstructed image.

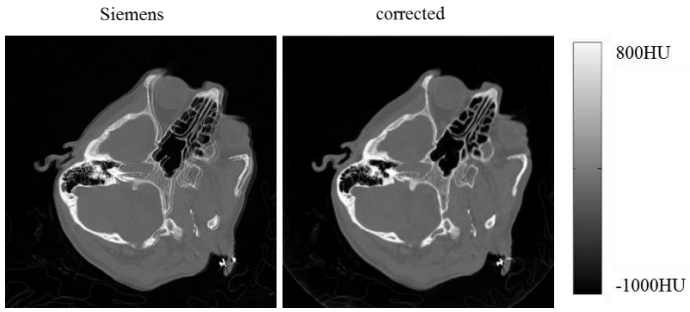



Fig. 6. Selected transaxial plane, without (left) and with correction (right) for motion artifacts in a clinical scan.

## IV. Conclusion

In this paper, we proposed a motion estimation and correction approach for helical X-ray CT of the head, only based on the measured raw data. Since no additional measurements are needed, it can be applied retrospectively to standard helical CT data. Further testing of the method with clinical data is ongoing.


## Acknowledgment

The authors want to thank Krystal Moore from the Department of Radiology at Westmead Hospital for collecting the datasets.


## References


[1] J. F. Barrett and N. Keat, "Artifacts in CT: recognition and avoidance.," *Radiographics*, vol. 24, no. 6, pp. 1679–91, 2004.

[2] J.-H. Kim, J. Nuyts, A. Kyme, Z. Kuncic, and R. Fulton, "A rigid motion correction method for helical computed tomography (CT).," *Phys. Med. Biol.*, vol. 60, no. 5, pp. 2047–73, Mar. 2015.

[3] A. Kyme, S. Se, S. Meikle, G. Angelis, W. Ryder, K. Popovich, D. Yatigammana, and R. Fulton, "Markerless Motion Tracking of Awake Animals in Positron Emission Tomography.," *IEEE Trans. Med. Imaging*, vol. 33, no. 11, Jun. 2014.

[4] T. Sun, R. Fulton, J.-H. Kim, and J. Nuyts, "Data driven helical CT rigid motion correction," in *Proc. of the 13th*

*Int. Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2015, pp. 444–447.

[5] G. T. Gullberg, B. M. W. Tsui, C. R. Craford, B. J.Glen, and J. T. Hagius, "Estimation of Geometrical Parameters for Cone Beam Tomography," *Med. Phys.*, vol. 17, no. 2, pp. 264–272, 1989.

[6] H. M. Hudson and R. S. Larkin, "Ordered subsets of projection data," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[7] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Phys. Med. Biol.*, vol. 49, no. 11, pp. 2463–2475, Jun. 2004.

[8] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures.," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.

[9] M. J. Ackerman, "The Visible Human Project," *Proceedings of the IEEE*, vol. 86, no. 3. pp. 504–511, 1998.

[10] Y. Nesterov, "A method of solving a convex programming problem with convergence rate O(1/sqr(k))," *Sov. Math. Dokl.*, vol. 27, pp. 372–376, 1983.