Physics in Medicine and Biology, February 2016, 61 (4): 1416-1438

Correction for human head motion in helical X-ray CT

J-H Kim[1], T Sun[2], A R Alcheikh[1], Z Kuncic[3], J Nuyts[2], and R Fulton[1,3,4]

[1]Faculty of Health Sciences, University of Sydney, NSW 2006, Australia

[2]Department of Imaging and Pathology, Nuclear Medicine & Molecular imaging, Medical Imaging Research Center (MIRC), B-3000, KU Leuven - University of Leuven, Leuven, Belgium

[3]School of Physics, University of Sydney, NSW 2050, Australia

[4]Department of Medical Physics, Westmead Hospital, Westmead, NSW 2145, Australia

E-mail: roger.fulton@sydney.edu.au

Abstract.

Correction for rigid object motion in helical CT can be achieved by reconstructing from a modified source-detector orbit, determined by the object motion during the scan. This ensures that all projections are consistent, but it does not guarantee that the projections are complete in the sense of being sufficient for exact reconstruction. We have previously shown with phantom measurements that motion-corrected helical CT scans can suffer from data-insufficiency, in particular for severe motions and at high pitch. To study whether such data-insufficiency artefacts could also affect the motion-corrected CT images of patients undergoing head CT scans, we used an optical motion tracking system to record the head movements of 10 healthy volunteers while they executed each of 4 different types of motion (no, slight, moderate and severe) for 60 s. From these data we simulated 354 motion-affected CT scans of a voxelized human head phantom and reconstructed them with and without motion correction. For each simulation, motion-corrected (MC) images were compared with the motion-free (MF) reference, by visual inspection and with quantitative similarity metrics. Motion correction improved similarity metrics in all simulations. Of the 270 simulations performed with moderate or less motion, only 2 resulted in visible residual artefacts in the MC images. The maximum range of motion in these simulations would encompass that encountered in the vast majority of clinical scans. With severe motion, residual artefacts were observed in about 60% of the

simulations. We also evaluated a new method of mapping local data sufficiency based on the degree to which Tuys condition is locally satisfied, and observed that areas with high Tuy values corresponded to the locations of residual artefacts in the MC images. We conclude that our method can provide accurate and artefact-free MC images with most types of head motion likely to be encountered in CT imaging, provided that the motion can be accurately determined.

Submitted to: *Phys. Med. Biol.*

Keywords: Motion estimation, motion compensation, computed tomography, brain imaging, image reconstruction.

1. Introduction

Patient head motion is a common source of image artefacts in X-ray CT in patients with reduced capacity to remain motionless during the scan, including very young patients, who are often sedated or anaesthetized (Wachtel *et al.*, 2009), and patients suffering from, for example, acute stroke (Fahmi *et al.*, 2013), dementia, or head trauma (Lee and Newberg, 2005).

We have previously demonstrated the feasibility of using an optical motion tracking system to measure the rigid motion of a physical phantom in 6 degrees-of-freedom (DoF) during helical scans performed on clinical CT scanners (Kim *et al.*, 2013), and then using these data to iteratively reconstruct a motion-corrected image from a modified source-detector orbit (Kim *et al.*, 2015). The modified orbit is determined by the motion of the head during the scan, and is created by applying a transformation given by the inverse of the head motion, at the time of the projection acquisition relative to an arbitrarily chosen reference pose, to the source and detector at each projection angle. This ensures that all projections are consistent with each other, and with a stationary object. However it does not necessarily ensure that the projections are complete in the sense of being sufficient for exact reconstruction. That depends on the orbit, and hence on the actual motion of the object during the scan.

In this previous work (Kim *et al.*, 2015), the correction method was shown to be effective at compensating for rapid, oscillatory 6 DoF motion of a cylindrical head phantom rolling back and forth across the CT bed, and in simulations of the head

motion of a single volunteer. Motion correction recovered an almost undistorted image in most cases. However, in some of the phantom scans performed at high pitch (≥ 1) residual artefacts were observed after motion correction had been applied. Evidence that these artefacts were due to data insufficiency was presented in the form of a backprojected uniform sinogram which revealed that the locations where artefacts were seen corresponded to areas of reduced sampling with the modified scanning orbit. It was concluded that certain combinations of motion and helical pitch can result in data insufficiency.

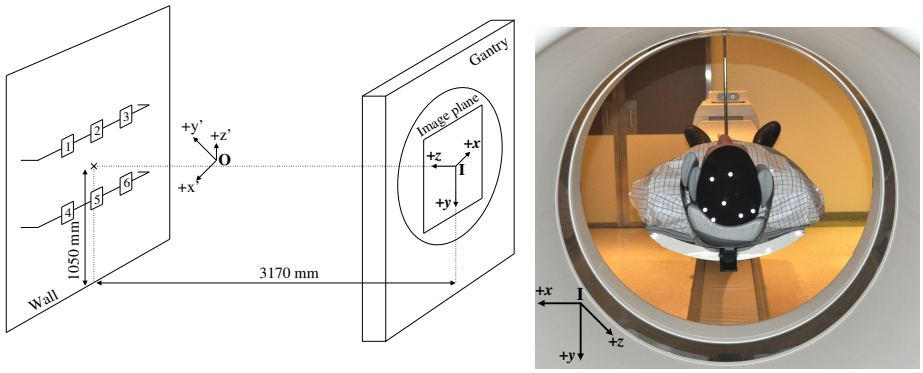
This leads to the important question of whether similar residual artefacts are likely to result when the method is applied to clinical scanning. On the one hand, most diagnostic CT scans of the head are performed at relatively low pitch (e.g. 0.8) which would reduce the likelihood of artefacts. On the other hand, while it has been shown that a fast rolling motion of a cylindrical phantom, with components in all 6 DoF, may produce artefacts that cannot be completely removed by motion correction, it was not known at the outset of the present work whether the motion that a human head can execute was capable of causing similar problems.

As a first step towards a clinical study, we have investigated the frequency with which artefacts due to data insufficiency would manifest if our motion correction method was applied to helical CT scans affected by typical human head motion. To address this question we simulated helical CT scans affected by a series of motion patterns measured from healthy volunteers. We tested motion patterns obtained when subjects were asked to remain stationary, as well as when they were asked to execute slight, moderate and severe head motion. In addition we investigated the potential for predicting the likelihood and location of residual artefacts based on the correlation between corresponding motion-free and motion-corrected slices, and a map of local data sufficiency developed by Sun *et al.* (2014).

2. Methods

2.1. Volunteer motion

2.1.1. Motion tracking An infrared optical tracking system (Optitrack, Natural Point Inc, Corvallis, OR USA) was used to acquire head motion data from ten healthy volunteers (7 male, 3 female, mean age 31.7 $\pm$10.8 y). The study was approved by the



(a) (b)

Figure 1. Set up for acquiring volunteer head motion (a) Cameras (1-6) positioned on rails behind the PET/CT gantry. **O** and **I** are the coordinate systems of the tracking system and CT scanner, respectively. (b) A volunteer positioned in the CT FOV, wearing a beanie with seven attached retro-reflective markers. Markers used to track the bed are also visible.

Human Research Ethics Committee of Western Sydney Local Health District.

The tracking system comprised six infrared cameras attached to rails on the rear wall of the scanner room (figure 1(a)), each connected via USB to a laptop computer. The system is capable of simultaneously reporting the poses of multiple rigid bodies, each comprising at least 3 retro-reflective markers, in 6 DoF, as three rotations about the x, y and z-axes (R_x, R_y, R_z) and three translations (T_x, T_y, T_z), at up to 120 Hz.

Each volunteer wore a beanie with seven retro-reflective velcro markers (spheres 14 mm diameter) attached, and lay supine on the scanning bed of a Siemens Biograph mCT PET/CT scanner (Siemens Medical Solutions USA, Inc., Knoxville, TN), which incorporates a standard Somatom Definition AS 64-row helical CT scanner. The head was centred in the CT scanner field of view (figure 1(b)).

Volunteers were asked to execute four types of motion: 'no' motion, slight, moderate and severe motion for one minute each, with a 1- min rest between each type of motion. In the 'no' motion case volunteers were asked to remain as motionless as possible. For slight and moderate motion they were asked to execute what they regarded as motion fitting those descriptions. For severe motion they were asked to move their head as fast as they could in all directions. The bed remained stationary

and no CT data were acquired during the motion data acquisition. The pose of the volunteer's head and the patient bed were simultaneously recorded at 120 Hz.

2.1.2. Motion data processing The raw tracker data reported in tracking system coordinates were converted to CT isocentre coordinates (Fig. 1(a)) using a 4×4 transformation matrix obtained from a calibration procedure described in Kim *et al.* (2013). Since the bed remained stationary during the course of volunteer head motion data acquisition, there was no need for removing the bed motion from the head motion data. The 1- min motion data obtained for each motion type were subdivided into three 15 s time segments starting at 15 s, 30 s and 45 s, respectively, to obtain three samples of each type of motion from each volunteer as shown in figure 3. Data in the first 15 s were not used to avoid possible delays in each volunteer executing the motion. From the 10 volunteers, there were a total of 120 motion samples, each 15 s in duration, comprising 30 samples of each of the 4 motion types. With severe motion, spurious tracker data were observed in 3 of the volunteers due to loss of marker visibility. In one, the spurious data were avoided by adjusting the time window slightly, but only two time segments could be extracted from each of the other two. As a result the number of samples for severe motion was reduced from 30 to 28, giving a total of 118 time segments to be simulated.

Noise due to jitter in the raw pose data was reduced by applying a 25-point polynomial (Kim *et al.*, 2015, Savitzky and Golay, 1964) to the pose estimates, except for the slowly varying no motion data, where a larger (111-point) smoothing kernel was used. Then the motion at each tracker sampling time i, M_i, relative to the mean pose was calculated as

$$M_i = S_i \bar{S}^{-1} \tag{1}$$

where S_i and $\bar{S}$ are 4×4 matrices representing the smoothed pose at time i and the mean pose, respectively. For calculation of the mean pose, x, y and z positions were averaged arithmetically, while cosine averaging (Stavdahl *et al.*, 2005) was applied to rotations. The motion data, M, were then linearly interpolated to match the timestamps of each simulated projection angle.

To characterize the amount of motion in each motion sample and each DoF we computed maximum amplitude as (maximum value) − (minimum value). We also

computed separate indices of the magnitude of rotational and translational motion for each motion sample as

$$\alpha_R = \sqrt{\sigma_{R_x}^2 + \sigma_{R_y}^2 + \sigma_{R_z}^2} \tag{2}$$

$$\alpha_T = \sqrt{\sigma_{T_x}^2 + \sigma_{T_y}^2 + \sigma_{T_z}^2} \tag{3}$$

where $\sigma_{R_{x,y,z}}$ and $\sigma_{T_{x,y,z}}$ were the standard deviations of the three rotations (in degrees) and translations (in mm), respectively.

2.1.3. Simulations CT scans affected by motion, and corresponding reference scans without motion, were simulated at 3 different pitch values (0.5, 0.8 and 1.0) by forward-projecting a 3D voxelized phantom (a reconstructed clinical brain scan) with a voxel size of $1.5\times1.5\times1.5\,\mathrm{mm}^3$. The forward-projection code simulated a Siemens Somatom Definition AS 64-row helical CT scanner using the scan parameters listed in table 1. With the three different pitch values, there were total of 354 simulation cases.

Table 1. Scanning parameters used in simulations.

Parameter	Value		
Tube voltage (kVp)	120		
Rotation time (s)	1.0		
Projections per rotation	1152		
Flying focal spot	Off		
Collimation (mm)	64×0.6		
Axial coverage (mm)	237		
Scanning direction	Caudo-cranial		
Reconstructed voxel dimensions (mm^3)	1.0×1.0×1.0		
Helical pitch	0.5	0.8	1.0
Table feed per rotation (mm)	19.2	30.72	38.4
Scan duration (s)	13.3	8.7	7.2
Total number of projections	15371	10039	8261

The upper part of figure 2 summarizes the simulation process. To simulate motion-affected scans, the appropriate 3D motion transformation derived from the acquired volunteer motion, M, was applied to the CT source and detector at each projection angle, and the CT-projections were computed with this modified orbit, using distance-driven projection, assuming a monochromatic beam and no scatter.

For the motion-free reference scans, the same procedure was used, without application of motion transformations.

The lower part of figure 2 illustrates the reconstruction scheme. A maximum-likelihood transmission (MLTR) algorithm (Nuyts *et al.*, 1998) was used for all reconstructions. Reconstructed images were dimensioned 512×512×276 with a voxel size of $1\times1\times1\text{mm}^3$. Motion-free (MF) simulations were reconstructed without motion correction, while simulations affected by motion were reconstructed using the same algorithm in two different ways: (a) using the motion-corrected orbit to obtain a motion-corrected (MC) reconstruction, and (b) with the standard orbit to obtain the conventional uncorrected (UC) reconstruction. The motion was assumed to be known exactly.

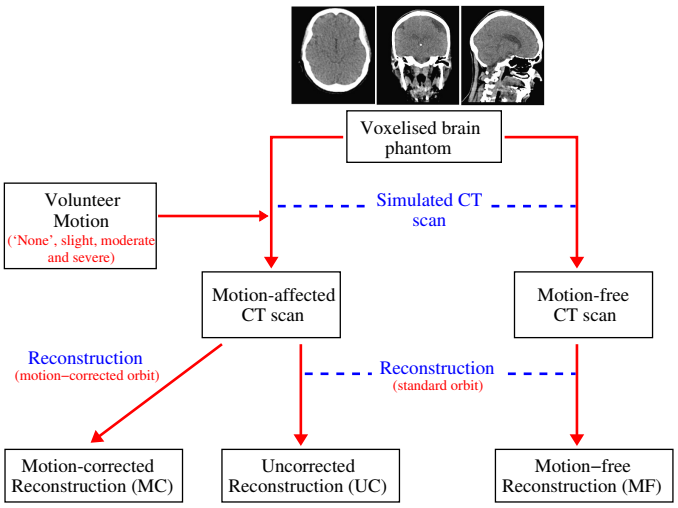


Figure 2. Simulation and reconstruction steps. In the simulations (upper part of figure) CT scans were generated from a voxelised brain phantom with and without simulated head motion. The motion-affected scans were reconstructed with and without a motion corrected orbit. Scans without simulated motion were reconstructed with the conventional orbit to obtain motion-free reconstructions for reference purposes.

2.2. Simulation accuracy

The ability of the simulation software to accurately predict the location and extent of data insufficiency artefacts was evaluated by comparing the motion-corrected images obtained from a real motion-affected CT scan with corresponding images

from a simulated CT scan. For this comparison we selected a real scan of a rolling Hoffman brain phantom acquired on a 64-row Somatom Definition AS scanner with the following scan parameters : tube voltage 120 kVp; tube current 615 mAs; collimation 64×0.6 mm; axial pitch of 1.5, in which artefacts attributable to data insufficiency were clearly seen after application of motion correction. The poses of the phantom and patient bed during this scan were recorded with a frequency of 120 Hz by a six DoF optical motion tracking system (Polaris Spectra, Northern Digital Inc., Waterloo, Canada) and converted to motion of the phantom in scanner isocentre coordinates as described in Kim *et al.* (2015).

The simulated scan was generated for the same object (a 3D reconstruction of the stationary Hoffman phantom in the same initial pose), the same scanner type, the same motion of the phantom, and the same scan parameters, as the real scan.

Motion correction was then applied to the simulated scan, and the MF, UC and MC images from both real and simulated scans real and simulated data were compared visually, and by calculating Pearson correlation coefficients between the MC images over all voxels in the 3D volume.

2.3. Effect of motion correction in simulated CT scans

The effect of motion correction on CT scans with simulated head motion was evaluated by separately comparing UC and MC images with corresponding MF images, treating the latter as a gold standard.

2.3.1. Quantitative analysis The accuracy of the motion-corrected images in each of the 354 simulations was assessed by calculating three (dis)similarity metrics between corresponding MC and MF images slice-by-slice, over 210 contiguous transaxial slices, encompassing the entire head and neck region. These metrics were root-mean square error (RMSE_c), Pearson correlation coefficient (CC_c) and mean structural similarity index (MSSIM_c) (Kim *et al.*, 2015), where the subscript 'c' denotes metrics representing comparisons between the corrected and reference (motion-free) images. Similarly, RMSE_u, CC_u and MSSIM_u were calculated from corresponding UC and MF images, as well as the means $\overline{RMSE_c}$, $\overline{CC_c}$, $\overline{MSSIM_c}$, $\overline{RMSE_u}$, $\overline{CC_u}$, and $\overline{MSSIM_u}$ over all 210 slices. The factors, F, by which motion correction improved these metrics

in each 3D reconstructed volume, were then obtained as

$$F_{RMSE} = \frac{\overline{RMSE_u}}{\overline{RMSE_c}}; \qquad F_{CC} = \frac{\overline{CC_c}}{\overline{CC_u}};$$
$$F_{MSSIM} = \frac{\overline{MSSIM_c}}{\overline{MSSIM_u}} \tag{4}$$

Finally mean improvement factors, $\overline{F}_{RMSE}$, $\overline{F}_{CC}$ and $\overline{F}_{MSSIM}$ were calculated by averaging the F values obtained over all simulations performed with the same combination of pitch and motion type.

2.3.2. Visual examination All 354 simulations were inspected by one of the authors (JK) for the presence of residual artefacts visible to the eye. This involved visually comparing all corresponding MC and MF slices in all simulations. A binary classification scheme was applied to each simulation with '1' denoting that at least 1 slice contained an artefact, and '0' denoting that no artefacts were observed. In addition the number of slices containing visible artefacts in each simulation was recorded. The data were then grouped by motion type and pitch value and the number of cases exhibiting artefacts in each group, as well as the total number of slices containing artefacts, was determined.

2.4. Predicting data insufficiency

In clinical practice a MF reconstruction will not not be available to aid in the identification of residual artefacts. Two potential objective approaches to predicting the likelihood and/or location of such artefacts in MC images that could be applied in the clinical setting were explored.

2.4.1. Approach based on motion-corrected orbit The completeness of the motion-corrected orbit produced by the volunteer motion was assessed at each image voxel by computing the degree to which a local Tuy condition (Sun *et al.*, 2014) was satisfied. In this method, local data sufficiency ('Tuy') values range from 0 to 1. A low value ($\simeq 0$) indicates good sampling near the voxel. Higher Tuy values indicate that the local Tuy condition in the voxel is violated more severely. Exact reconstruction of the object is only guaranteed if the Tuy values are low everywhere. The artefacts due to data incompleteness tend to occur mostly in regions with high Tuy values. Because this evaluation was computationally intensive, Tuy maps were only computed for a

limited number of cases in which residual artefacts had previously been identified by visual analysis.

2.4.2. Approach based on calculated metrics We also examined relationships between the incidence of visually-detected residual artefacts and various objective parameters derivable from computed metric values and related to the overall accuracy of motion correction, such as the minimum value of CC_c over all slices of the reconstructed image ($minCC_c$). This was based on the theory that residual artefacts should be associated with low correlation between MF and MC slices. While this method would not be implementable directly on clinical data due to the unavailability of an MF image, it could potentially be applied to a simulation of the motion.

3. Results

3.1. Volunteer motion

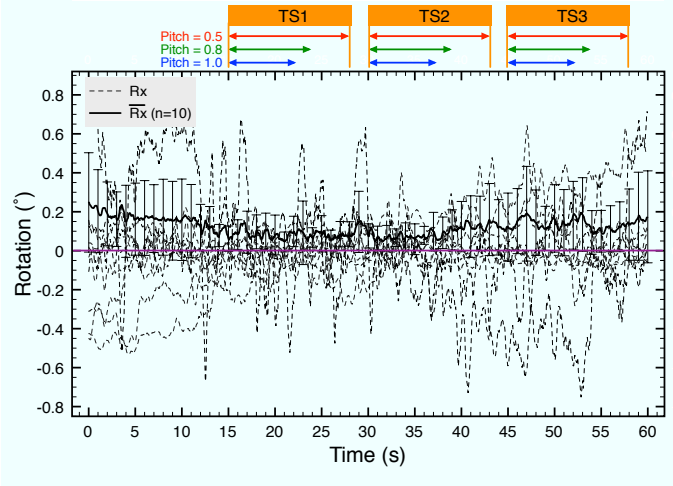
Figure 3 shows the head motion in two DoF (R_x and T_z) of the 10 volunteers for 'no' motion, over a period of 1 min. The 10 plots are superimposed. Motion of up to 1° and 1.5 mm was observed.

The slight, moderate and severe head motion patterns of a typical volunteer are shown in figure 4 for all DoF. The increase in the magnitude of the motion from slight to moderate, and from moderate to severe motion can be readily observed in the figure. Table 2 gives the mean amplitude of the motion in each DoF and motion type that was simulated at pitch 0.8. The mean amplitude increased in each degree of freedom with the amount of motion the subject was asked to execute, and there was a large increase in amplitude in the transition from moderate to severe motion. The standard deviation of the amplitude increased in a similar fashion. The corresponding mean motion indices shown at right in Table 2 followed a similar pattern.

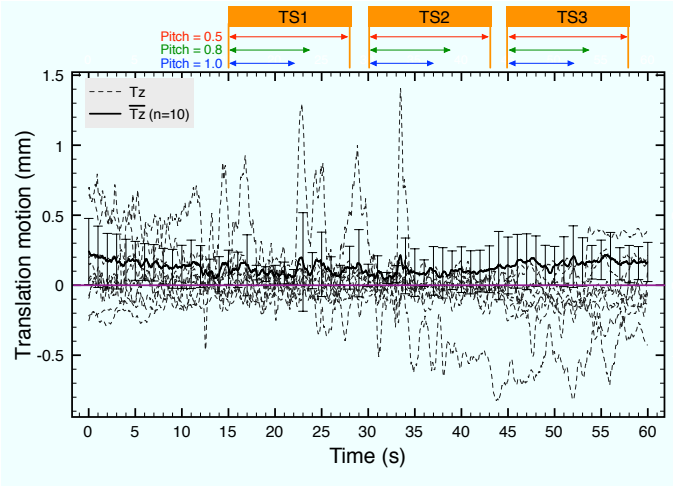
A similar increase in maximum motion amplitude (computed from each simulation as overall max - min in each DoF) can also be seen in figure 5.

3.2. Simulation accuracy

Figure 6 presents a comparison of images obtained in the real and simulated CT scans, with and without motion of the phantom. The upper two rows show images



(a) R_x

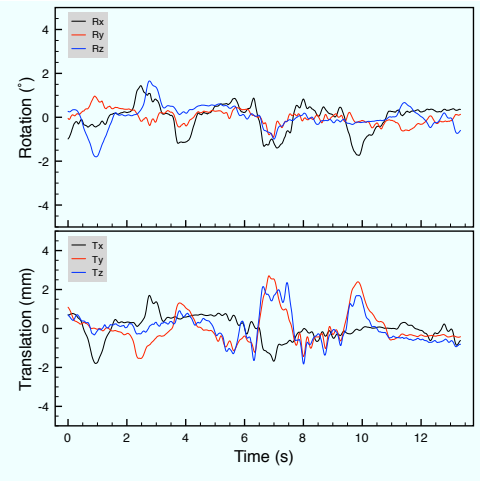


(b) T_z

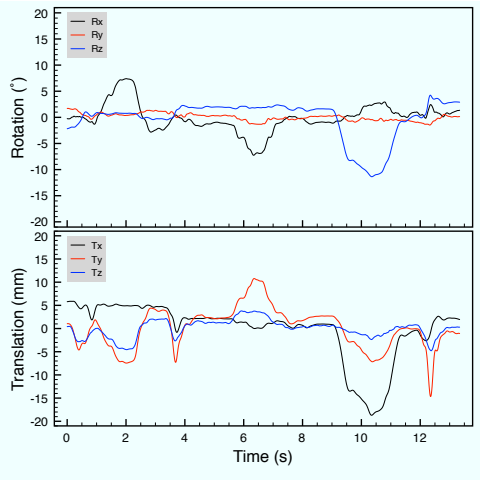
Figure 3. Measured 'no' motion data in two selected DoF of the ten volunteers (superimposed dashed lines). Solid lines are absolute means with 1 s.d. error bars. Zero-lines are shown in purple. Typical time segments (TS) used in the simulations are also shown as TS1, 2 and 3, for the three different pitch values. Six DoF pose data in isocentre coordinates for all volunteers and motion types are supplied as supplementary data.

Table 2. Mean 6 DoF motion magnitude simulated at pitch = 0.8.

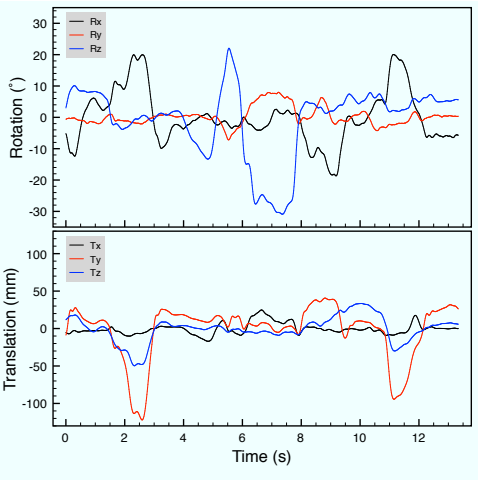
Motion	Mean Amplitude												Mean Motion Indices			
	R_x (°)		R_y (°)		R_z (°)		T_x (mm)		T_y (mm)		T_z (mm)		α_R (°)		α_T (mm)	
'No'	0.3	(0.1)	0.2	(0.1)	0.1	(0.1)	0.3	(0.1)	0.3	(0.1)	0.4	(0.1)	0.1	(0.0)	0.1	(0.1)
Slight	3.5	(0.9)	1.8	(0.4)	7.6	(2.0)	10.4	(2.6)	3.6	(0.9)	3.1	(0.7)	2.6	(2.0)	2.4	(2.3)
Moderate	8.9	(2.3)	5.1	(1.2)	18.2	(4.7)	20.8	(5.2)	13.8	(3.2)	7.9	(1.9)	6.4	(2.7)	5.7	(3.1)
Severe	29.7	(7.0)	23.9	(5.1)	72.8	(18.8)	61.3	(14.0)	92.5	(22.1)	42.5	(8.1)	24.0	(10.3)	28.7	(14.6)



(a) Slight motion



(b) Moderate motion



(c) Severe motion

Figure 4. Motion of a typical volunteer (Subject 1, TS2, pitch 0.5). Different scaling has been used for each motion type.

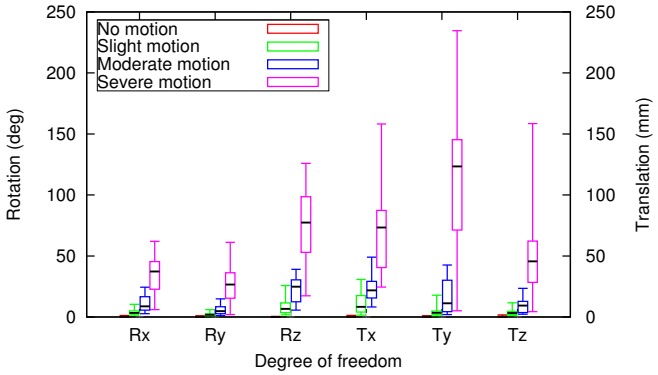
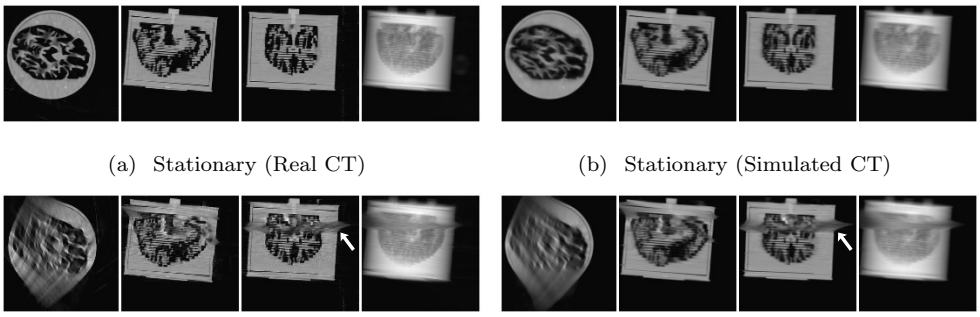


Figure 5. Maximum amplitude of motion in each DoF for each motion type. Bars represent overall maxima and minima (i.e. The minimum of the bar indicates the smallest maximum motion over all volunteers). The lower and upper bounds of the boxes represent the first and third quartiles, respectively, and the black bar in each box is the median.

obtained without motion of the phantom, and with no motion correction applied. Good correspondence was seen between transaxial, coronal, sagittal and projection images, created by forward projection using a parallel beam projector, and there were no apparent artefacts. The mean correlation coefficient over all transaxial slices covering the entire brain region ($\overline{CC}$) was 0.97. The third and fourth rows show the motion corrected images obtained by applying motion correction to real and simulated CT scans, respectively. Both were affected by the same motion. There is a close correspondence between these images in terms of the location and severity of visible artefacts, with $\overline{CC}$ of 0.96. This indicates that our simulation of the Definition AS scanner was accurate, and that the simulation software accurately predicts the location and severity of data-insufficiency artefacts in the motion-corrected image for any given motion of the object.

3.3. Effect of motion correction in simulated CT scans

Figure 7 shows motion-correction results for simulated CT scans affected by the motion of a typical volunteer (subject 4, TS3). Parts (a) and (b) of the figure show the simulated object and the reconstruction of a motion-free simulation of the object, respectively. Parts (c), (d) and (e) show simulation results with this subject's motion and pitch values of 0.5, 0.8 and 1.0, respectively. Slices reconstructed from a simulated



(a) Stationary (Real CT) (b) Stationary (Simulated CT)

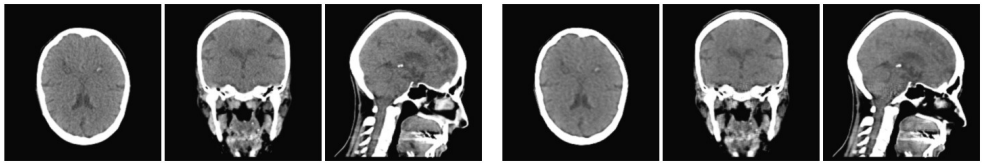
(c) Motion corrected (Real CT) (d) Motion corrected (Simulated CT)

Figure 6. Stationary (reference) and motion-corrected images for real and simulated CT scans. Each panel shows, from left to right, transaxial, coronal, and sagittal slices and a projection image. The arrows in (c) and (d) indicate a prominent residual artefact in the real and simulated scans after applying motion correction. [WL=−200 HU, WW=+2000 HU]

motion-affected scan, with and without motion correction, are shown for each motion type. At all pitch values it can be seen that the motion artefacts in the uncorrected images become progressively worse as the magnitude of the motion increases. In all cases motion correction appeared to effectively remove motion artefacts and restore images that closely resembled the MF reconstruction, without residual artefacts. However figure 7(f) shows another typical simulation of severe motion (subject 8, TS3) in which residual artefacts were observed with pitch 1.0.

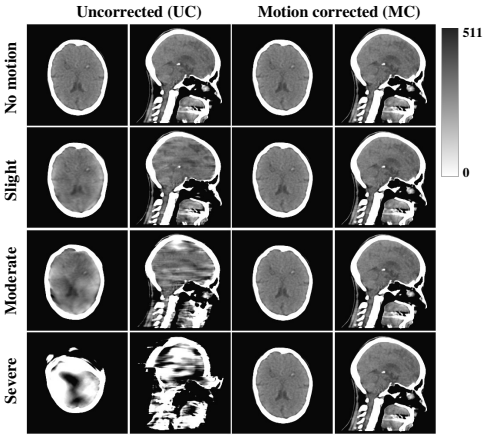
3.3.1. Quantitative image analysis The mean Pearson correlation coefficient between corresponding MC and MF slices ($\overline{CC_c}$) was high for all motion types and pitch values, as shown in figure 8. All the simulations with 'no' and slight motion, and the bulk of simulations with moderate motion, resulted in a $\overline{CC_c}$ value very close to 1.0, suggesting near-perfect motion correction at all pitch values. Motion correction was less accurate for severe motion, although $\overline{CC_c}$ was always above ∼0.988, and the deficit can be attributed to the data insufficiency rather than a deficiency in the motion correction method itself. The lowest value of $\overline{CC_c}$ observed in any of the 354 simulations resulted from a simulation of severe volunteer motion acquired with a pitch of 1.0. Median values of $\overline{CC_c}$ represented by horizontal black bars in figure 8 suggested that motion correction was generally more accurate at lower pitch values.

Table 3 shows the mean (s.d) metric values averaged over all reconstructed slices

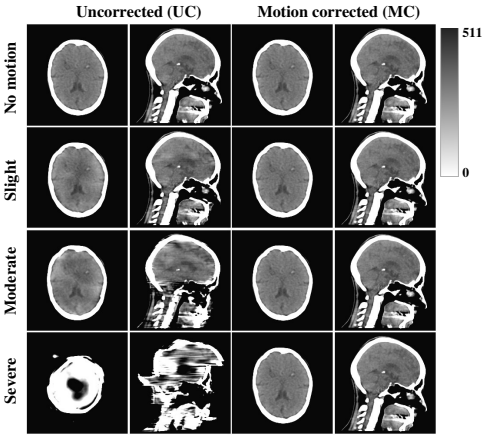


(a) Object

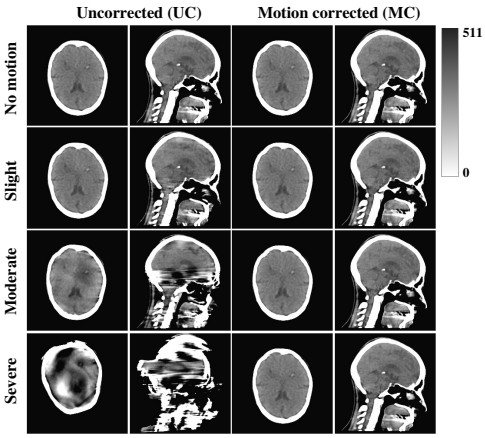
(b) Motion-free (MF) reconstruction



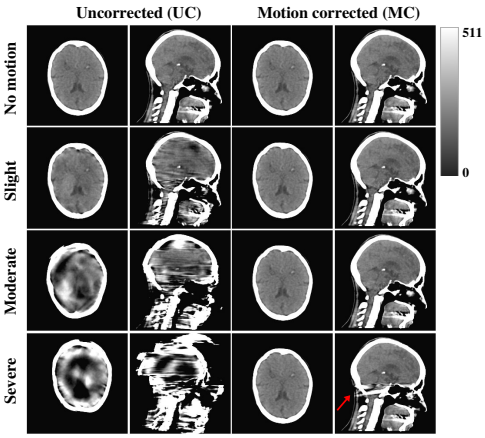
(c) Subject 4 (TS3), Pitch = 0.5



(d) Subject 4 (TS3), Pitch = 0.8



(e) Subject 4 (TS3), Pitch = 1.0



(f) Subject 8 (TS3): Pitch = 1.0

Figure 7. (a) Orthogonal slices from the phantom used in the simulated CT scans. (b) a reconstruction of a simulation with no motion of the phantom. (c, d, e) comparison of UC and MC images at 3 different pitch values with the same motion. (f) a similar comparison with different motion. The arrow indicates the location of a residual artefact after application of motion correction. [WL=+40 HU, WW=+140 HU]. The results of all other simulations in similar format, and videos of the volunteer motion in (e) and (f) are provided as supplementary data.

and simulations for each motion type and pitch value, with and without motion correction. All metric values calculated between corresponding UC vs MF images indicated a progressive increase in the severity of motion artefacts as the magnitude of simulated motion increased from 'no' to severe motion. Application of motion correction improved these metric values in all cases approaching the ideal values for each metric (i.e. RMSE = 0 and CC = MSSIM = 1). For slight and moderate motion, RMSE and CC values after motion correction were of the same order as those of the uncorrected 'no' motion scans, which are normally assumed to be artefact-free in clinical practice, suggesting that the motion correction was very effective.

Figure 9 shows the mean improvement factors calculated from data in table 3 for each metric, motion type and pitch value. For 'no' motion, $\overline{F}_{RMSE}$ suggested an improvement in reconstruction accuracy with motion correction, but the other two metrics indicated negligible change. With slight, moderate and severe motion, all metrics showed an improvement in reconstruction accuracy. These observations were consistent with visual examination of images such as those shown in figure 7 where no clear differences between UC and MC images could be observed visually in 'no' motion simulations, but could be readily seen in all simulations of slight, moderate and severe motion. Differences in mean improvement factors at differing pitch values were found not to be statistically significant using a Kruskal-Wallis test, except for $\overline{F}_{RMSE}$ with severe motion, in which the null hypothesis that the means were not statistically different was rejected ($p < 0.001$). The mean improvement factors increased with the amount of motion, which can be attributed to increasing distortion in the UC images combined with relatively stable performance of the motion correction algorithm.

3.3.2. Visual examination Table 4 summarizes the results of visually examining the MC images for residual artefacts. For 'no' motion and slight motion, no residual artefacts were observed in any of the 180 simulations. For moderate motion residual artefacts were observed in 2 of the 90 simulations, both performed with motion from the same subject and time segment. In these two simulations which were performed with pitch values of 0.8 and 1.0, the artefacts were observed in 15 and 24 slices, i.e. 7.1% and 11.4% of slices, respectively. In the case of severe motion, residual artefacts were seen in 50 of the 84 simulations, and covered on average 22 (i.e. 10.5% of) slices over all simulations. No artefacts were observed in the remaining slices.

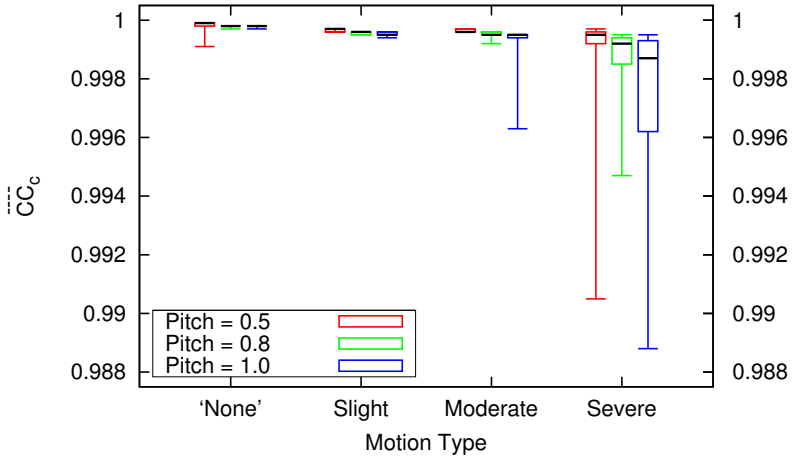


Figure 8. Mean slice-to-slice Pearson correlation coefficient between MC and MF images for all motion types and pitch values.

Table 3. Mean (s.d.) metric values, with and without motion correction, over all reconstructed slices and all simulations for each motion type and pitch value.

(a) 'No' motion (n=30)

Pitch	UC vs MF			MC vs MF		
	RMSE (HU)	CC	MSSIM	RMSE (HU)	CC	MSSIM
0.5	16.1 (8.2)	0.999 (0.001)	0.994 (0.005)	8.4 (0.9)	1.000 (0.000)	0.997 (0.001)
0.8	16.6 (6.2)	0.999 (0.001)	0.993 (0.005)	10.3 (0.6)	1.000 (0.000)	0.996 (0.002)
1.0	16.9 (6.6)	0.999 (0.001)	0.993 (0.006)	10.5 (0.7)	1.000 (0.000)	0.996 (0.002)

(b) Slight motion (n=30)

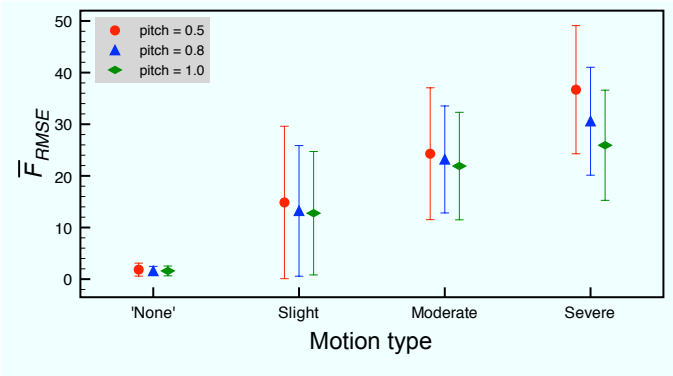
Pitch	UC vs MF			MC vs MF		
	RMSE (HU)	CC	MSSIM	RMSE (HU)	CC	MSSIM
0.5	197.5 (113.3)	0.898 (0.105)	0.799 (0.144)	13.2 (0.6)	1.000 (0.000)	0.995 (0.002)
0.8	201.0 (109.6)	0.900 (0.099)	0.776 (0.140)	15.2 (0.5)	1.000 (0.000)	0.994 (0.002)
1.0	201.3 (108.4)	0.899 (0.103)	0.771 (0.142)	15.6 (0.7)	1.000 (0.000)	0.993 (0.002)

(c) Moderate motion (n=30)

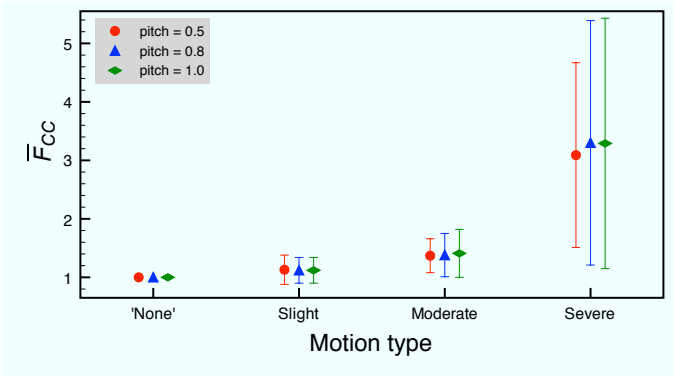
Pitch	UC vs MF			MC vs MF		
	RMSE (HU)	CC	MSSIM	RMSE (HU)	CC	MSSIM
0.5	367.0 (90.6)	0.726 (0.118)	0.566 (0.113)	13.7 (0.4)	1.000 (0.000)	0.994 (0.002)
0.8	360.3 (92.9)	0.736 (0.124)	0.570 (0.112)	15.5 (0.6)	1.000 (0.000)	0.993 (0.002)
1.0	366.8 (97.5)	0.729 (0.131)	0.562 (0.125)	16.8 (1.7)	0.999 (0.001)	0.992 (0.003)

(d) Severe motion (n=28)

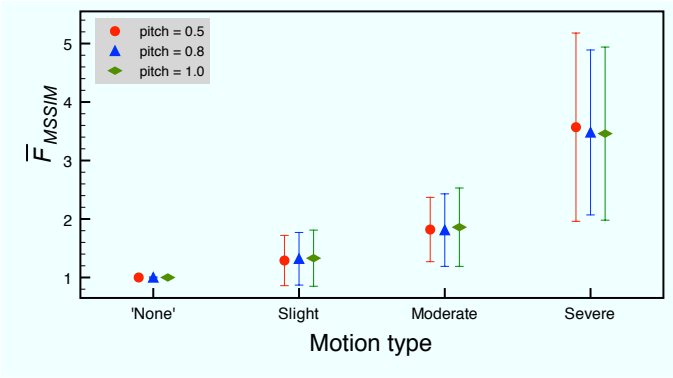
Pitch	UC vs MF			MC vs MF		
	RMSE (HU)	CC	MSSIM	RMSE (HU)	CC	MSSIM
0.5	644.0 (105.8)	0.360 (0.132)	0.309 (0.107)	18.5 (6.5)	0.999 (0.002)	0.991 (0.005)
0.8	654.7 (100.1)	0.335 (0.129)	0.317 (0.114)	22.4 (6.8)	0.999 (0.001)	0.984 (0.011)
1.0	663.3 (96.3)	0.335 (0.119)	0.317 (0.114)	27.5 (9.2)	0.997 (0.003)	0.977 (0.017)



(a) $\overline{F}_{RMSE}$



(b) $\overline{F}_{CC}$



(c) $\overline{F}_{MSSIM}$

Figure 9. Mean improvement factors with 1 s.d. error bars for RMSE, CC and MSSIM for pitch values of 0.5, 0.8 and 1.0.

Table 4. Number of cases (slices) with visually-detected residual artefacts after motion correction.

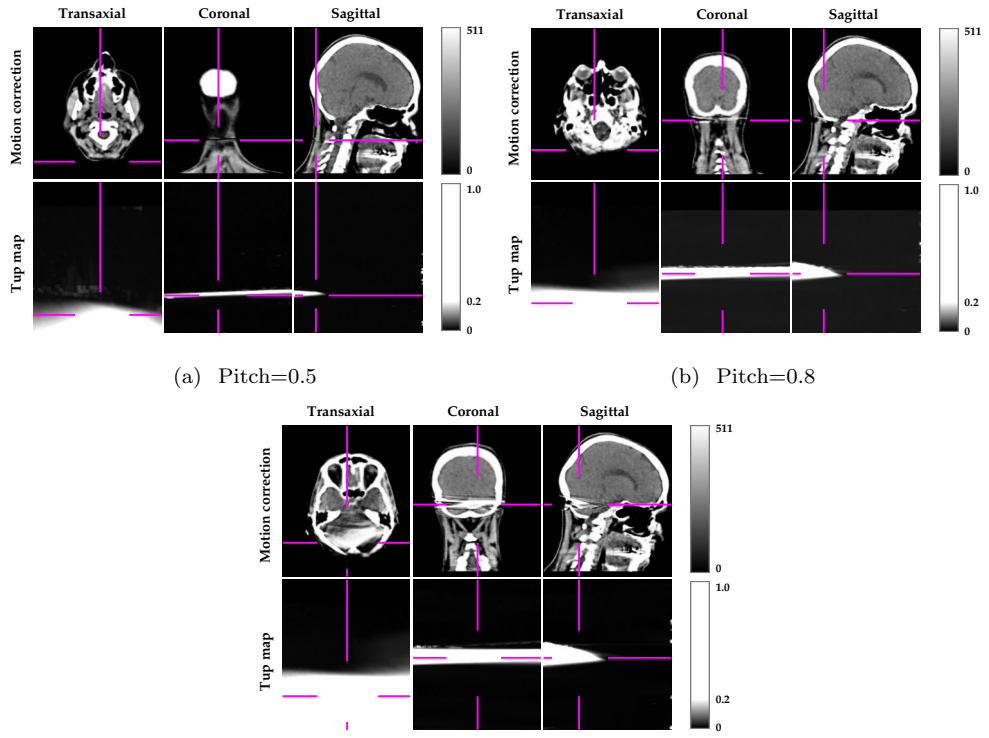
Pitch	No motion	Slight	Moderate	Severe
0.5	0	0	0	12 (316)
0.8	0	0	1 (15)	17 (561)
1.0	0	0	1 (24)	21 (961)

The number of slices containing residual artefacts tended to increase with increasing pitch, even though less motion was simulated due to the shortening duration of the scan as pitch was increased. In general, when residual artefacts were visible in the reconstruction after motion correction, they were limited in their extent and other parts of the reconstructed volume appeared to be free of artefacts. For motion of the type considered here, this suggests that data insufficiency may have a local rather than global effect on reconstruction accuracy.

3.4. Predicting data insufficiency

3.4.1. Approach based on motion-corrected orbit From the 52 simulations with visually identified residual artefacts we selected one moderate motion time segment for which residual artefacts were observed at pitch values of 0.8 and 1.0, and one severe motion time segment for which an artefact was seen at all 3 pitch values, and generated Tuy maps from these two time segments for all pitch values. The Tuy maps were then compared visually with the corresponding motion-corrected images (figures 10 and 11). Regions with reduced sampling (indicated by high Tuy values) were shown to correspond to regions affected by residual artefacts in the motion corrected images. In figure 10 the crosshairs are centred on a region of increased local Tuy value, and it can be seen in the corresponding orthogonal slices of the MC image that residual artefacts occur in the same region. Interestingly, at pitch 0.5, this revealed a subtle artefact in the neck muscle that was not detected during visual examination. It was also observed that the axial extent of the region of data insufficiency increased with increasing pitch, probably due to the motion that caused the insufficiency affecting more slices when the speed of bed motion was increased. In figure 11, areas of increased Tuy value at pitch values of 0.8 and 1.0 corresponded to artefacts in the region of the buccal cavity

that had been detected during visual examination. Again, the axial extent of the area of reduced sampling increased with increasing pitch. The corresponding region in the Tuy map for pitch 0.5 did not suggest data insufficiency in that region, but a new region of reduced sampling appeared near the top of the head. This was associated with motion that was not present in the shorter scans performed at pitch 0.8 and 1.0.



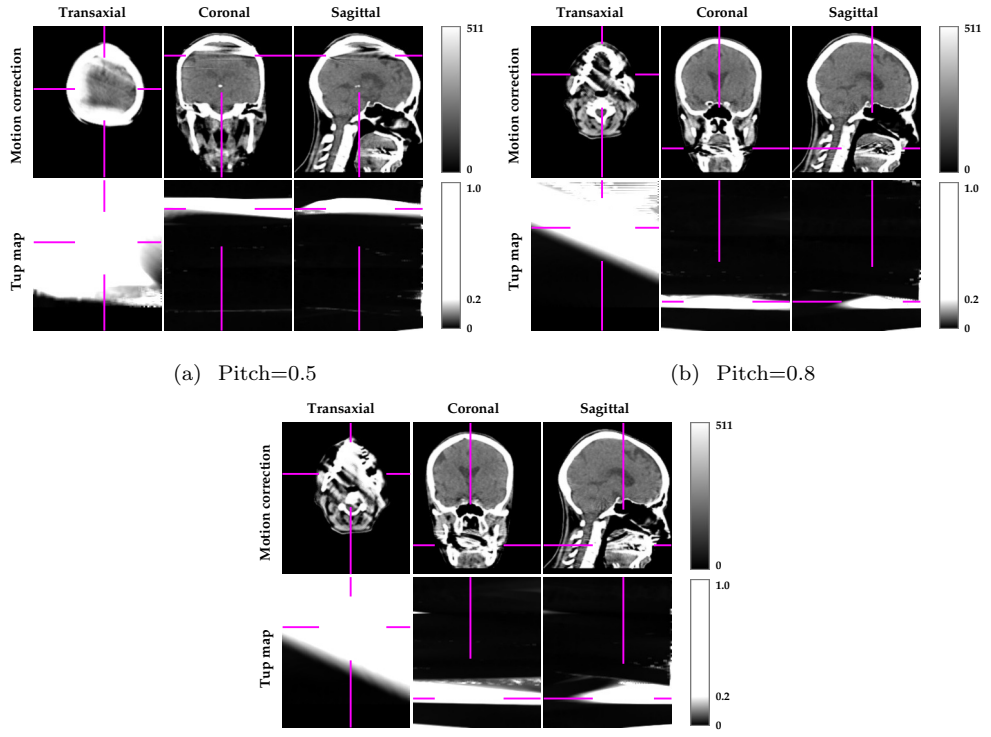
(a) Pitch=0.5

(b) Pitch=0.8

(c) Pitch=1.0

Figure 10. Orthogonal motion-corrected slice images (upper row in each panel) and corresponding Tuy map (lower row each panel) for the only moderate motion time segment that produced visible residual artefacts (Subject 5, TS2). A video of the volunteer motion is provided as supplementary data.

3.4.2. Approach based on calculated metrics Figure 12 shows the distribution of minCC_c, the lowest correlation coefficient in any slice for a given simulation calculated from corresponding MC and MF slices, for all 354 simulations, subdivided into two categories, depending on whether or not a residual artefact was observed in the MC image during visual examination. In all of the 302 simulations in which no residual artefacts were observed, minCC_c was greater than 0.996. In the majority (44 of 52) of



(a) Pitch=0.5

(b) Pitch=0.8

(c) Pitch=1.0

Figure 11. Orthogonal motion-corrected slice images (upper row in each panel) and corresponding Tuy map (lower row each panel) for a typical severe motion time segment (Subject 1, TS1). A video of the volunteer motion is provided as supplementary data.

the remaining simulations in which residual artefacts were observed, minCC_c was $<$ 0.996. Assuming that the subjective visual examination represented truth, and using $\text{minCC}_c > 0.996$ as a predictor for the presence of a residual artefact, this would have resulted in eight of the 354 simulations (2.3%) being wrongly classified as artefact-free (see overlap region in figure 12).

3.5. Simulations of 'no' motion

Motion artefacts were readily observable in the UC reconstructions of all slight, moderate and severe motion simulations. However in simulations of 'no' motion, artefacts in UC images were only observed in 3 of the 90 simulations. Two of these were performed using the same time segment and pitch values of 0.8 and 1.0. The third,

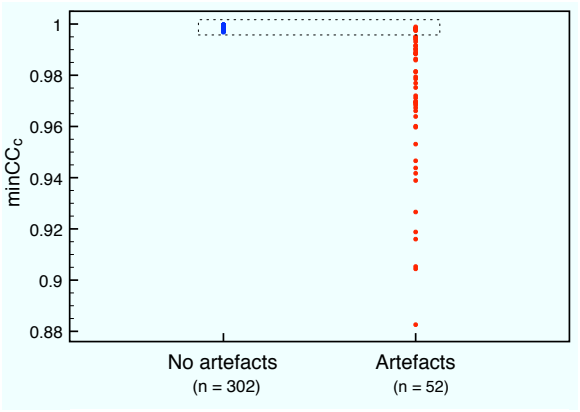
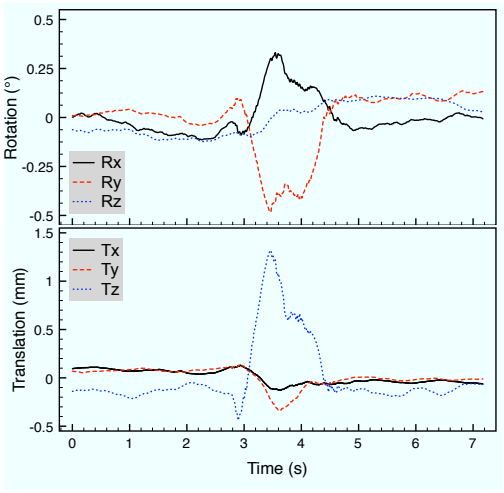


Figure 12. minCC$_c$ values for simulations classified by visual examination as free of residual artefacts (blue symbols) and containing residual artefacts (red symbols).

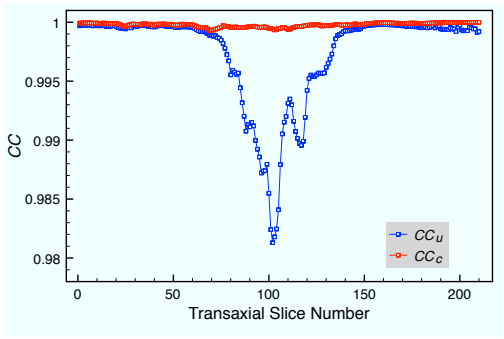
performed at pitch 1.0 is shown in figure 13. The motion simulated, shown in figure 13(a), was very small, with all rotations $< \pm 0.5°$, and all translations $< \pm 0.5$ mm, except for a transient excursion of ~ 1.3 mm in the z direction midway through the scan. This translation coincided with similarly brief rotations of up to $\sim 0.5°$ about the x- and y-axes. Without the aid of motion tracking such small movements would not normally be noticed in the clinical setting.

Figure 13(b) shows that even this very slight motion can cause a demonstrable loss of accuracy. A plot of CC_u versus slice number (blue curve) shows that reconstruction accuracy was degraded in about 85 slices without motion correction, peaking at the mid-slice which was acquired at about the same time as excursions were observed in figure 13(a). With motion correction, reconstruction accuracy was almost fully restored (red curve), with $CC_c \approx 1.0$ in all slices.

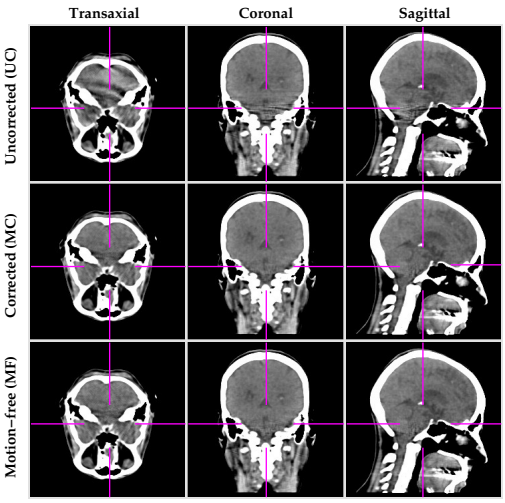
The corresponding degradation and improvement in reconstruction accuracy can be seen in the images of figure 13(c). Comparing the UC images (upper row) with the MF images (lower row), motion artefacts are readily seen in the mid-axial slices of the UC reconstruction. However these artefacts are absent after motion correction (middle row). The crosshairs on these images are linked in the 3 planes.



(a)



(b)



(c)

Figure 13. A 'no' motion simulation in which slight motion degraded the accuracy of the reconstructed image, and motion correction recovered an accurate image (Subject 7, TS2). (a) simulated motion, (b) slice-by-slice correlation coefficient before (blue) and after (red) motion correction, (c) UC, MC, and MF reconstructions. [WL=+40 HU, WW=+140 HU]. A video of the motion is provided as supplementary data.

4. Discussion

We have shown in this simulation study that effective motion correction without the introduction of data-insufficiency artefacts is feasible for a wide range of human head movements likely to occur during clinical helical CT scanning. In the 270 simulations performed with moderate motion or less, motion correction produced reconstructions that were imperceptibly different from the MF reconstruction in virtually all cases. Only 2 of the simulations (0.74%) exhibited residual data-insufficiency artefacts detectable by eye. The maximum range of motion in this group would encompass that encountered in the vast majority of clinical scans.

Residual artefacts were much more prevalent in simulations of severe motion, in which the motion was considerably greater on average, with mean rotational and translational amplitudes in some DoF of up to 72.8°and 92.5 mm, respectively. Motion of this magnitude would be rarely encountered clinically, but it is interesting to note that despite the large magnitude of motion, about 40% of these simulations did not have any visually detected residual artefacts after correction. In the remaining simulations that did, only 17.5% of slices on average contained visible residual artefacts. This suggests that clinically useful images can be obtained when patients move extensively during a scan, as might occur for example when attempting to scan a paediatric patient without general anaesthesia or sedation. In such settings the method could reduce the number of instances when a repeat scan is required due to movement of the patient, and thereby avoid unnecessary repeat doses of radiation.

In simulations the head motion is precisely known, but in clinical practice the accuracy of motion correction will depend on the accuracy with which head motion can be estimated. With the optical motion tracking method, we previously demonstrated accurate motion correction in real CT scans of a moving physical brain phantom (Kim *et al.*, 2015). In these studies the reflective disk markers used for motion tracking were rigidly attached to the phantom. However when the method is applied clinically we will adopt the same motion tracking methods described here with a tight fitting cap attached to the patient's head. There will be potential for the markers to move relative to the head, and introduce errors in the motion data.

Optical motion tracking is also subject to other sources of error. Measurement jitter, errors in the spatial calibration of the tracker and CT scanner coordinate

systems, and synchronisation error impose a finite limit on the accuracy of motion correction achievable. The true efficacy in patients will not be known until the method is tested clinically. Recent preliminary work on a data-driven motion estimation method (Sun *et al.*, 2015) is promising, and if shown to be feasible could eliminate the need for motion tracking provided that the motion is not too extreme. For larger motion, the tracker could be used to correct most of the motion, while the data driven approach would then be used to reduce the residual motion, in which case, the problems of jitter, tracker calibration, cap slippage, and synchronisation error would be eliminated.

The translational excursions observed in our volunteers while executing 'no' motion were similar in magnitude to values previously reported by Li *et al.* (2010) who observed absolute excursions < 1.5 mm in 16 out of 19 healthy volunteers, and larger movements of up to 4 mm in the other 3 volunteers. In a similar study of 20 patients referred for CT scans of the head for a variety of clinical indications, Wagner *et al.* (2003) observed mean translational and rotational excursions of 2.5 mm and 1.06°, respectively. It is not unexpected that the motion amplitude should be greater in patients than in normal healthy volunteers. Motion amplitude could also be reasonably expected to increase with the duration of the study and a reduction in the ability of the patient cohort to remain still. A recent study (Fahmi *et al.*, 2013) reported the head motion of 103 consecutive patients admitted with suspected acute ischaemic stroke, who underwent 1 min CT perfusion scans. In this cohort, motion classified as 'moderate' or 'severe' by these authors was observed in 24% of all patients, with mean R_z and T_z of 14° and 22.6 mm, and maximum amplitudes of 62.9° and 69.3 mm, respectively. Referring to figure 5, motion classed as moderate and extreme by these authors would both fit within our moderate motion range.

While motion of the magnitude of the severe motion simulated in this study may be rarely seen in some clinics, there are some patients (e.g. unsedated paediatric, or agitated and uncooperative patients, or patients with debilitating neurological disorders) in whom the ability to cooperate is compromised, and consequently large motion may occur. The ability to correct for it without having to repeat the scan is of significant clinical value. Often a clinical decision is made not to scan the patient at all, due to the high probability of severe motion. Our work suggests that scanning patients under such circumstances could deliver clinically useful information despite

the motion.

The possibility of large head movements resulting in insufficient data for exact reconstruction appears to be the only limitation on the accuracy of the motion correction method. Indeed this is not a limitation of the method but rather a limitation of the data available for reconstruction when motion of this kind occurs. Although movements of this magnitude would only be anticipated in particular patient groups prone to such motion, such as young children imaged without sedation/anaesthesia, and trauma and dementia patients, it would be useful to have a means to predict the sufficiency or otherwise of the acquired data in such cases. If the data were found to be sufficient, the need for a repeat scan and additional radiation dose could be avoided by applying motion correction. For all cases in which we computed a Tuy map, regions with elevated Tuy values corresponded to locations where data-insufficiency artefacts were observed visually. The Tuy map therefore appears to be an accurate means of predicting the presence, and location, of data-insufficiency artefacts in the clinical setting. Generating the Tuy map is however highly compute-intensive and the software would need to be accelerated by an order of magnitude to make this a practical technique for routine use.

An alternative prediction method examined here was to calculate the minimum CC between corresponding slices of the MC and MF reconstructions, on the basis that this statistic would identify the MC slice that differed most from the corresponding MF slice. By setting a cutoff level of 0.996 for minimum CC we achieved a good separation between simulations with residual artefacts and those without. In clinical practice however it will not be possible to calculate CC, as the MF reconstruction will be unavailable and this method will not be directly implementable. However one could perform a simulation, similar to those performed in this work, using the patient motion and a suitable phantom (Sun *et al.*, 2014). A low minimum CC value in such a simulation would be suggestive of residual artefact, and could assist in a decision as to whether a repeat scan is necessary. Although not as time consuming as generating a Tuy map, this method of prediction would still require considerable computation and would need to be accelerated for routine clinical use. Fortunately, based on our results with no, slight and moderate motion, the question of data sufficiency is unlikely to arise in the majority of clinical CT scans.

An unexpected finding was that small movements that are normally undetected

in clinical scanning have the potential to cause inaccuracies and artefacts in the reconstructed images, and that motion correction can effectively remove these effects. In clinical CT scanning, unless the motion is large enough to be noticed at the time of scanning, or to produce obvious motion artefacts, it is normally assumed that the reconstructed image is an accurate depiction of the anatomy. Our simulations suggest that it may be possible to improve image accuracy in many cases where motion is currently assumed to have been absent or to have had a negligible effect on the image. Whether this could have a clinically significant impact on the interpretation of the such scans can, however, only be established in a clinical study.

5. Conclusions

In this study we have used software that accurately simulates a state-of-the-art helical CT scanner, and human head motion data derived from volunteers, to examine the feasibility of correcting for a wide range of human head movements during helical CT imaging. In all 354 simulated helical CT scans with 'no', slight, moderate and severe head motion, image accuracy was measurably improved after applying motion correction. This suggests that if this method was applied to clinical CT scanning it would provide more accurate images, free of motion and data-insufficiency artefacts in the majority of scans affected by head motion. It could enable successful imaging of patients whose images would otherwise be non-diagnostic with conventional reconstruction methods, thus avoiding the need for repeat scans and additional radiation doses.

The results also suggest that residual artefacts for which no correction is available can be expected in some patients exhibiting moderate and severe motion. However data from the visual analysis of the scans with residual artefacts suggests that these artefacts will be limited in extent, and that the number of slices affected will generally increase with increasing pitch. At a pitch value commonly used for head CT scanning, 0.8, residual artefacts were observed in only 15 of 6300 (0.2%) and 561 of 5880 (9.5%) of reconstructed slices for moderate and severe motion, respectively. The remaining slices (99.8% and 90.5%, respectively) were free of visible artefacts and highly correlated with corresponding motion-free slices. It is therefore possible that motion correction could yield a diagnostically useful image even in cases of the most

severe motion as simulated in this study. This could for example enable successful imaging in the most challenging of clinical situations, such as imaging paediatric patients without anaesthesia or sedation. In cases where the magnitude of motion suggests the possibility of data-insufficiency artefacts, the Tuy map could be used to identify regions susceptible to such artefacts, and which should be treated with caution diagnostically.

We have also observed that motion correction may be of value in the many scans performed clinically, in which head motion sufficient to affect the reconstructed image occurs, but is too small to be noticed at the time of the scan. It would indeed be interesting to apply motion correction to a large sample of such scans and perform a thorough evaluation of the impact of motion correction on clinical interpretation and subsequent management, but we defer this to a future clinical study.

These preliminary data convincingly demonstrate the potential of this recently reported method to correct for the motion likely to be encountered in a range of clinical CT studies. To facilitate incorporation of the method into routine clinical imaging procedures, we plan in future work to accelerate our algorithms for motion corrected reconstruction, Tuy map generation, and data-driven motion estimation, using GPUs. It is hoped that these methods will benefit patients by making their CT images more accurate, and improve the safety of CT imaging by eliminating the need for repeat scans when data acquisition is compromised by patient head motion.

Acknowledgments

This work was supported by National Health and Medical Research Council Australia Project Grant 632677, by the KU Leuven IMIR project and by the MIRIAD SBO project of IWT, Flanders. J.-H. Kim was supported by a University Postgraduate Award from the University of Sydney. We are grateful to Dr K. Stierstorfer of Siemens AG, Healthcare Sector, Forchheim, Germany for assistance with raw CT data formats.

References

Fahmi F, Beenen L F M, Streekstra G J, Janssen N Y, Jong H W de, Riorda A, Roos Y B, Majoie C B, Bavel E Van and Marquering H A 2013 Head movement

during CT brain perfusion acquisition of patients with suspected acute ischemic stroke *Eur. J. Radiol.* **82** 2334–41

Kim J-H, Nuyts J, Kuncic Z and Fulton R 2013 The feasibility of head motion tracking in helical CT: A step toward motion correction *Med. Phys.* **40**(4) 041903-1–4

Kim J-H, Kyme A, Nuyts J, Kuncic Z and Fulton R 2015 A rigid motion correction method for helical computed tomography (CT) *Phys. Med. Biol.* **60** 2047–73

Lee B and Newberg A 2005 Neuroimaging in traumatic brain imaging *NeuroRX* **2**(2) 372–83

Li L, Chen Z, and Jin X 2010 Experimental measurement of human head motion for high-resolution computed tomography system design *Optical Engineering* **49**(6), 063201-1– 6

Nuyts J, Man B De, Dupont P, Defrise M, Suetens P, and Mortelmans L 1998 Iterative reconstruction for helical CT: A simulation study *Phys. Med. Biol.* **43** 729–37

Savitzky A and Golay M J E 1964 Smoothing and differentiation of data by simplified least squares procedures *Anal. Chem.* **36**(8)1627–39

Stavdahl O, Bondhus A K, Pettersen K Y and Malvig K E 2005 Optimal statistical operators for 3-dimensional rotational data: geometric interpretations and application to prosthesis kinematics *Robotica* **23**(3) 283–92

Sun T, Clackdoyle R, Fulton R and Nuyts J 2014 Quantification of local reconstruction accuracy for helical CT with motion correction *2014 IEEE Nuclear Science Symposium Record*

Sun T, Kim J-H, Fulton R and Nuyts J 2015 Data-driven rigid motion correction for helical CT *Proc. of the 13th Int. Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (New Port, RI, USA, Jun 2015)* pp 444–7.

Wachtel R E, Dexter F and Dow A J 2009 Growth rates in pediatric diagnostic imaging and sedation *Anesth. Analg.* **108**(5) 1616–21

Wagner A, Schicho K, Kainberger F, Birkfellner W, Grampp S, and Ewers R 2003 Quantification and clinical relevance of head motion during computed tomography *Invest. Radiol.* **38**, 733–41