Published in final edited form as:

Phys Med Biol. ; 62(23): 8813–8831. doi:10.1088/1361-6560/aa9254.

Correction of Patient Motion in Cone-Beam CT Using 3D-2D Registration

S Ouadah¹, M Jacobson¹, JW Stayman¹, T Ehtiati², C Weiss¹, and JH Siewerdsen¹

¹Department of Biomedical Engineering, Johns Hopkins University, Baltimore MD 21205, USA

²Siemens Medical Solutions USA, Inc., Imaging & Therapy Systems, Hoffman Estates IL 60192, USA

Abstract

Cone-beam CT (CBCT) is increasingly common in guidance of interventional procedures but can be subject to artifacts arising from patient motion during fairly long (~5-60s) scan times. We present a fiducial-free method to mitigate motion artifacts using 3D-2D image registration that simultaneously corrects residual errors in the intrinsic and extrinsic parameters of geometric calibration. The 3D-2D registration process registers each projection to a prior 3D image by maximizing gradient orientation using the covariance matrix adaptation-evolution strategy (CMA-ES) optimizer. The resulting rigid transforms are applied to the system projection matrices, and a 3D image is reconstructed via model-based iterative reconstruction. Phantom experiments were conducted using a Zeego robotic C-arm to image a head phantom undergoing 5–15 cm translations and 5–15° rotations. To further test the algorithm, clinical images were acquired with a CBCT head scanner in which long scan times were susceptible to significant patient motion. CBCT images were reconstructed using a penalized likelihood objective function. For phantom studies the structural similarity (SSIM) between motion-free and motion-corrected images was >0.995, with significant improvement ($p < 0.001$) compared to the SSIM values of uncorrected images. Additionally, motion-corrected images exhibited a point-spread function with full-width at half maximum comparable to that of the motion-free reference image. Qualitative comparison of the motion-corrupted and motion-corrected clinical images demonstrated a significant improvement in image quality after motion correction. This indicates that the 3D-2D registration method could provide a useful approach to motion artifact correction under assumptions of local rigidity, as in the head, pelvis, and extremities. The method is highly parallelizable, and the automatic correction of residual geometric calibration errors provides added benefit that could be valuable in routine use.

Keywords

3D-2D registration; C-arm; cone-beam CT; image-guided interventions; image-guided surgery; image-guided radiotherapy; image quality; motion correction; artifacts; geometric calibration

1. Introduction

Intraoperative C-arm cone-beam CT (CBCT) involves relatively long scan times (~5-60s) and can be subject to artifacts from patient motion. A similar challenge exists in CBCT for

image-guided radiotherapy (IGRT) where acquisition times may be even longer (~30-120s). Patient motion can arise from respiration, the cardiac cycle, gas moving in the intestines, involuntary muscle motion or twitch, etc., and patient motion artifacts are evident as blurring, streaks, and/or distortion of the image reconstruction. Such motion, even at the millimeter scale, can confound accurate localization of anatomical structures, which is critical for IGRT and image-guided interventions (IGI) using CBCT. For example, in studying motion of the head in 20 patients undergoing CT without rigid fixation, Wagner *et al* (2003) found that all 20 patients exhibited detectable motion, with 4 patients showing perceptible motion artifacts in the reconstructed image. Li *et al* (2010) similarly concluded in a study of 19 volunteers that all subjects displayed head motion while undergoing CT scans without rigid fixation, and further estimated that the head motion of patients without rigid fixation will not be smaller than 0.35 mm, which cannot be neglected in high-resolution CBCT imaging.

A variety of methods for motion correction in CBCT have been proposed. Many of these are with respect to 4D imaging, and for the purposes of this paper are referred to as “motion compensation” methods, since the goal is often to model and visualize such motion (c.f., “motion correction” methods considered below, in which any anatomical motion is undesirable and the goal is to form a 3D image free of motion artifact). Several motion compensation methods have employed gating for cyclical motion (either prospective or retrospective), most notably for periodic motion of the lungs and heart in thoracic imaging (Rit *et al* 2009, Rohkohl *et al* 2013, Muller *et al* 2014). These methods typically use a general or patient-specific motion model to divide projection data into various phases. The quality of the resulting images is therefore dependent on the accuracy of the model. Other 4D imaging techniques employ deformable image registration to solve for the motion vector fields on a voxel-wise basis (Isola *et al* 2010, Tang *et al* 2012, Brehm *et al* 2013).

For static 3D imaging - e.g., imaging of the head as in the work reported below - involuntary motion rarely follows a pre-defined model, and the 4D motion “model” is usually of secondary or no interest compared to clear visualization and static localization of the anatomy. Methods to minimize motion during the scan include instructions to the patient (e.g., breath-hold) and/or physical immobilization (frames, straps, etc.), but may be impractical or insufficient in many interventional imaging scenarios.

A number of approaches have been described to correct for motion in 3D image reconstruction using fiducial markers. Tracking systems can be used to optically track the positions of fiducial markers placed on the patient (Dinelle *et al* 2006); however, tracker-based motion measurement introduces additional hardware, may face difficulty with line-of-sight occlusion by a rotating gantry, and the geometric accuracy may not be sufficient to correct for sub-millimeter errors (i.e., at the level of the voxel size in 3D image reconstruction). The tracked motion measurements must also be accurately synchronized with projection data acquisition, and proper calibration is needed to convert tracker coordinates to image coordinates (Nazarparvar *et al* 2012).

Recognizing such drawbacks to motion estimation from externally tracked fiducials, various methods have been proposed to track radio-opaque fiducial markers within the projection

data itself. Jacobson and Stayman (2008) simultaneously estimated marker positions and their corresponding orientation using 6 degrees of freedom (DoF) for each projection, allowing the marker configuration to adjust for different patients. Marchant *et al* (2011) tracked the 3D locations of fiducial markers placed on a patient and warped the 2D projection images to match the 2D locations of their forward-projected 3D locations. Choi *et al* (2013) extended this idea by comparing simple projection shifting with projection warping using methods that prevent unrealistic warping due to noise in the marker locations. Choi *et al* (2013) also studied a 3D rigid-body warping method and found it to outperform both 2D methods. Overall, such *image* fiducial methods (including fiducials placed internal to the patient) appear advantageous to externally tracked fiducials, but require placement of fiducials that can interrupt clinical workflow and may not accurately represent the internal motion of the patient – i.e., the fiducials give a fairly sparse representation of motion and convey only the motion of structures to which the fiducials are attached.

Other work seeks to correct patient motion by maximizing (or minimizing) various characteristics of the resulting 3D image reconstruction (e.g., image sharpness). Yu *et al* (2007) determined motion parameters by minimizing an objective function based on data redundancy in the sinogram. Similarly, Wicklein *et al* (2013) minimized an entropy criterion to optimize the underlying geometric parameters of the imaging system which adjusts for general system misalignments as well as patient motion. Sisniega *et al* (2017) combined a gradient metric and an entropy term with a penalty that encourages smooth motion for high-resolution extremity imaging. Such methods can accommodate complex source-detector trajectories as they only rely on the reconstructed image and have shown promising results for improving image quality; however, they appear to be somewhat limited to small motion magnitudes.

Other methods employ a motion-free prior image to correct for motion in a subsequent scan. For example, Berger *et al* (2016) used 3D-2D registration to estimate the motion of individual bones in the weight-bearing knee, showing that registration based on image similarity (specifically, gradient correlation (GC) and normalized gradient information (NGI)) outperformed motion correction using fiducial markers. Chen *et al* (2008) used prior image constrained compressed sensing (PICCS) to reconstruct images from highly undersampled projection data sets, providing a method to reconstruct dynamic data without motion or undersampling artifacts.

The work presented below develops and characterizes a fiducial-free method for motion correction in interventional CBCT. Particular focus is on 3D imaging of the head, as in otolaryngology/head and neck surgery, intracranial neurosurgery, or neuro-interventional radiology, where involuntary patient motion can give rise to severe artifacts. The method utilizes 3D-2D registration to register all projection views to a previously acquired, motion-free 3D image, providing the information necessary to reconstruct a motion-corrected image using model-based image reconstruction (MBIR). Using this framework, we aim to correct large-scale rigid motion that occurs during a scan and restore image quality to that of the motion-free image. The method is similar to that of Berger *et al* (2016) but uses a distinct registration framework with a variety of advantages for interventional imaging, applies the method in the context of head imaging, and demonstrates the applicability of the algorithm

in clinical cases. The similarity metric is based on gradient orientation (GO), which is robust against mismatch in image content between the prior 3D image and the current acquisition (demonstrated by De Silva *et al* 2016). This is an important characteristic when the prior image is acquired at the beginning of the case, and the current image is acquired with a variety of devices and instrumentation in the field of view (e.g., retractors, catheters, stents, coils, clips, staples, implants, etc.) that may otherwise confound the registration. We also employ a fast scan (5s, 200° orbit) for the prior image, providing a 3D CBCT image that is less susceptible to motion (assumed nominally motion-free) for phantom experiments. The fast, low-dose scan exhibits lower image quality (higher quantum noise and view sampling effects) compared to a longer, higher-dose scan and requires a registration method that is robust against such image quality degradation. Finally, we demonstrate the applicability of the algorithm in patient data exhibiting strong, realistic motion in which factors such as complexity of the motion pattern and mismatch between a prior image and the projection data present a particular challenge to the registration. The motion correction method builds upon previous work for “self-calibration” of a C-arm system geometry (Ouadah *et al* 2016). By treating the motion of the C-arm gantry and the motion of the patient simultaneously in 3D-2D registration, we obtain a method that corrects for both patient motion and geometric calibration errors, or even scan orbits for which a geometric calibration does not exist – e.g., non-circular orbits or scans with a high degree of gantry tilt.

2. Methods for Motion Correction

2.1. Overview of method

The motion-correction method is illustrated in figure 1, assuming a previously acquired, motion-free 3D image (I_M) and a current scan that is corrupted by patient motion with projection data denoted y_F . The algorithm performs a 3D-2D image registration for each projection in the current scan; alternatively, a subset of projections could be used, subject to smoothing or regularization, but in the current work, each projection was treated as a separate registration. Each 2D projection is registered to the forward projection of the previously acquired image (known as a digitally reconstructed radiograph, DRR). The nominal projection matrix (PM) for system geometric calibration (Navab *et al* 1998, Cho *et al* 2005, Ouadah *et al* 2016) is used to initialize the first registration and the result of previous registrations is used to initialize subsequent registrations. This PM used to initialize each registration is denoted PM_{init}^i . The 6 or 9 DoF transformation that optimizes each registration yields a new PM (denoted PM_{motion}^i) that encodes the motion of the object for the i^{th} projection. This process is performed for each projection in the motion-corrupted CBCT acquisition (typically 200–500 projections), and a motion-corrected 3D image is reconstructed according to the system geometry defined by PM_{motion} . Using this method, the rigid motion of the object is defined within PM_{motion} , which combines both the source and detector motion with that of the object (patient). For cases in which the patient motion is small (or more specifically, in which PM_{motion} still describes an approximately circular system geometry) conventional 3D filtered backprojection using the Feldkamp-Davis-Kress (FDK) algorithm (Feldkamp *et al* 1984) may still be applicable as a reasonable approximation. However, stronger motion amplitude, as investigated below, results in an

effective system geometry that is markedly non-circular and violates basic assumptions in conventional FDK. This invites the application of more general MBIR methods that are free from strong assumptions on system geometry. In the work reported below, a penalized likelihood (PL) estimation is used to reconstruct the motion-corrected CBCT image.

2.2. Initialization

Registration of the first projection image is initialized with a PM taken from the system geometric calibration. The second projection is initialized with the PM resulting from the first registration ($PM_{init}^2 \leftarrow PM_{motion}^1$). PM_{init}^i for the third and subsequent projections (through the N^{th} projection) is generated by bi-linear extrapolation of the extrinsic parameters of the previous two projection matrices as described in Ouadah *et al* (2016) to predict the pose of the current projection. This bilinear extrapolation used to create PM_{init}^i is also distinct from previous methods for 3D-2D registration of multiple projections, such as (Berger *et al* 2016). This approach was previously shown to trap spurious outliers and yield realistic, continuous orbits of the C-arm, introducing a slight degree of interdependence between pose estimates for each projection view. Alternative constraints on realistic motion of the C-arm could be implemented, including those that do not depend on sequential registration of projections to better parallelize the 3D-2D registration process described below.

2.3. 3D-2D registration

The motion correction method has at its “inner loop” a 3D-2D registration method based on work by Penney *et al* (1998) and Pluim *et al* (2000) and later expanded to include alternative similarity metrics and optimizers as in Otake *et al* (2013), Uneri *et al* (2014), and De Silva *et al* (2016). The method involves a moving 3D volume (I_M) registered to a fixed 2D projection (y_F) by means of its DRR (denoted y_M). In this work we use the GO similarity metric as in De Silva *et al* (2016), defined as:

$$GO = \frac{1}{N_p} \sum_m \frac{2 - \ln(|a_m| + 1)}{2} \quad (1)$$

where $a_m = \cos^{-1}(\cos \theta_m)$ and $m \in \{S: |\nabla I_1(j)| > t_1 \cap |\nabla I_2(j)| > t_2\}$, with j indexing all pixels. The GO similarity measure describes how well image gradient vectors align, with θ_j indicating the angle between gradient vectors. To exclude small gradients associated with image noise, the summation in (1) includes only gradient magnitude exceeding thresholds t_1 and t_2 , defined as the median gradient magnitude in images I_1 and I_2 , respectively, and is normalized by the number of overlapping pixels, N_p . As shown in De Silva *et al* (2016), GO is robust against mismatch in image content (e.g., soft-tissue deformation and/or instrumentation present in y_F but not I_M), since it relies on the orientation of gradients, rather than their magnitude (as in GC, NGI, etc.).

The similarity metric was iteratively optimized using the covariance matrix adaptation-evolution strategy (CMA-ES, Hansen 2006) to solve for the 6 or 9 DoF transform:

$$\hat{T} = \underset{T \in S}{\operatorname{argmax}} GO(y_M, y_F(T)) \quad (2)$$

where $T = (t_M, \theta_M)$ for 6 DoF and $T = (t_s, t_M, \theta_M)$ for 9 DoF. The vector $t_M = [t_{M,x}, t_{M,y}, t_{M,z}]^T$ represents a 3D translation of the object (patient) with respect to the detector in mm, $\theta_M = [\theta_{M,x}, \theta_{M,y}, \theta_{M,z}]^T$ represents a 3D rotation of the object (patient) with respect to the detector in degrees, and $t_s = [t_{s,x}, t_{s,y}, t_{s,z}]^T$ represents a 3D translation of the source with respect to the detector in mm. CMA-ES is a non-linear, non-convex evolutionary optimizer that adjusts the covariance matrix of a population (set to $\lambda = 200$, consistent with previous studies (Ouadah *et al* 2016)) with each iteration. The optimization was taken to be complete (converged) when the change between iterations in each parameter in T was <0.1 mm or degrees, corresponding to sub-voxel changes in the placement of back-projected rays.

A linear forward-projector (i.e., the adjoint of a standard voxel-driven FDK back-projector) implemented on GPU was used to compute y_M , where $y_M = b_0 e^{-A(T)I_M}$, with matrix A being a linear projection operator and b_0 indicating the bare-beam image signal. A linear projector was chosen for speed, although other forms are available, such as Siddon's method (Siddon *et al* 1985, as used in Otake *et al* 2013) or a simplified separable footprints projector as shown in Wang *et al* (2014). The final PM for the motion-corrected geometry was computed as:

$$PM_{motion}^i = PM_{init}^i * \begin{pmatrix} & & t_{M,x} - t_{s,x} \\ \mathcal{R}_{3 \times 3}(\theta_{M,x}, \theta_{M,y}, \theta_{M,z}) & t_{M,y} - t_{s,y} & \\ & t_{M,z} - t_{s,z} & \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3)$$

for 6 DoF and

$$PM_{motion}^i = \begin{pmatrix} t_{s,z} & 0 & t_{s,x} & 0 \\ 0 & t_{s,z} & t_{s,y} & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} & & t_{M,x} - t_{s,x} \\ \mathcal{R}_{3 \times 3}(\theta_{M,x}, \theta_{M,y}, \theta_{M,z}) & t_{M,y} - t_{s,y} & \\ & t_{M,z} - t_{s,z} & \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

for 9 DoF, noting that the formulation of the final PM for 9 DoF does not rely on PM_{init}^i as in the 6 DoF case; rather, it is constructed using intrinsic and extrinsic parameters de novo for each projection.

These are equivalent formulations of the PM as in Ouadah *et al* (2016) in which 3D-2D registration was used for geometric calibration, showing that the method does not distinguish between rigid motion of the object (patient) and that of the source and detector.

3. Experimental Methods

3.1. Imaging systems

Phantom experiments (described in section 3.2) used a robotic C-arm (Artis Zeego, Siemens Healthcare, Forchheim Germany), shown in figure 2a. Three scan protocols were used, each based on standard head scan protocols available on the system (102 kV with a 200° circular orbit): (Scan Protocol A) $N_{\text{proj}} = 496$ projections acquired in 20 seconds (giving 37 mGy as reported by the system; reference point air kerma at 15 cm from isocenter toward the x-ray source); (Scan Protocol B) $N_{\text{proj}} = 248$ projections acquired in 10 seconds (19 mGy); and (Scan Protocol C) $N_{\text{proj}} = 133$ projections acquired in 5 seconds (10 mGy).

The motion correction method was applied in clinical studies (described in section 3.5) using a prototype CBCT head scanner (Xu *et al* 2016a) to image patients with suspected intracranial hemorrhage (ICH). The fairly long (30s) scan time combined with the sometimes unstable condition of critically ill patients (prone to involuntary motion) made the system susceptible to patient motion. Image acquisition parameters are further described in section 3.5. Scan protocols involved 450 projections acquired in a 360° rotation at 100 kV, 216 mAs (23 mGy). Each patient received a MDCT scan prior to the CBCT scan using a standard, non-contrast-enhanced head scan protocol (120 kV, 274 mAs, 44.4 mGy) and reconstructed with $(0.44 \times 0.44 \times 0.5) \text{ mm}^3$ voxels and a H20s soft tissue kernel (Somatom Definition AS, Siemens Healthcare, Forchheim Germany). The MDCT scan was taken as the 3D prior image for 3D-2D registration with the CBCT projection data.

3.2. Phantom experiments

A custom motion phantom was built by attaching an anthropomorphic head phantom to a motion controller (figure 2b). The head phantom consisted of a human skull encased in tissue-equivalent plastic (RANDO, The Phantom Laboratory, Greenwich NY) with a 0.13 mm diameter tungsten wire placed in the trachea for spatial resolution assessment. The CIRS Dynamic Thorax Phantom (Model 008A) and software (CIRS, Norfolk VA) were used for computer-controlled motion of the head phantom during the scan.

The motion phantom provided 2 degrees of freedom under computer control: longitudinal rotation in the θ direction (approximately about the z -axis of the object and opposite to the Zeego gantry rotation) following a sigmoidal profile; and periodic superior-inferior (SI) translation along the z -axis following a sinusoidal profile that persisted for the duration of the scan. The amplitude (A_θ , units of degrees) and duration (P_θ , units of seconds) of rotational motion and the amplitude (A_z , units of mm) of translational motion were variable. For experiments containing periodic translational motion, the period was held constant at 4 seconds, comparable to the respiratory cycle. The phase of rotational motion during the scan (ϕ_θ , units of degrees) was adjusted so that the midpoint of the motion profile occurred near the beginning (50°), middle (100°), or end (150°) of the 200° scan. The motion profiles transferred to the anthropomorphic head were not precisely reproducible, owing to friction between the phantom and its support base and instability or backlash in the attachment piece, so the requested motion profiles were not taken as “truth.” However this was not a factor in experiments described below, since the performance of the motion correction

method was assessed with respect to the quality of the resulting 3D image reconstruction and not in comparison to the motion profiles.

The performance of the (6 DoF) motion correction method was tested for motion profiles defined by the amplitude of longitudinal rotation (A_θ), the duration of longitudinal rotation (P_θ), the angular phase of the motion relative to the scan (ϕ_θ), the amplitude of SI translation (A_z), and the scan time (T_{scan} as defined by protocols A, B, and C described in section 3.1). For each experiment, the values of each variable are shown in table I, and the corresponding motion profiles are shown in figure 3.

One motion-free scan was acquired for each of the three scanning protocols as a basis of comparison and taken as the “true” reference image without motion artifact. The motion-free image from protocol C was used as the 3D volume input (I_M) to the motion correction algorithm for all experiments. Using the fast scan image from protocol C as the motion-free prior corresponds to a clinical scenario in which a fast ($T_{scan} = 5s$) CBCT is acquired at the beginning of a case and is subsequently used as the 3D input for slower scans (higher quality, but susceptible to motion) taken later in the procedure. This also stresses the algorithm in that the 3D input image from protocol C is of fairly low quality, containing a high degree of noise due to low dose, and angular sampling artifacts within the image due to the low angular sampling of 133 projections over 200° .

3.3. Image reconstruction

Projection data were read from the Artis Zeego system in DICOM format at isotropic pixel size of 0.308 mm. Since automatic exposure control (AEC) was implemented on this system, a frame-to-frame intensity correction was performed by dividing each frame by the mAs for each projection as recorded in the DICOM header. A simple scatter correction was performed by subtracting a fraction of the minimum pixel value for each frame. A fraction equal to 0.97 was found to minimize cupping artifact and was used throughout phantom experiments shown below.

The FDK algorithm was used to reconstruct motion-corrupted images to illustrate the magnitude of motion artifacts without corrections. A $(512 \times 512 \times 512)$ voxel grid size was used with 0.75 mm isotropic voxel size, apodization filter cutoff at 50% of the Nyquist frequency, and 200 mm lateral extrapolation to reduce truncation effects. A MBIR method based on PL was used to reconstruct images using the motion-corrected projection matrices. The method is irrespective of the non-circular geometry implicit in PM_{motion} for motion correction and formulates the reconstructed image (μ) by maximizing an objective function comprising the log-likelihood of the projection data $L(\mu; y)$ and a regularization term with roughness penalty $R(\mu)$ and penalty strength β :

$$\hat{\mu} = \underset{\mu}{\operatorname{argmax}} L(\mu; y) - \beta R(\mu) \quad (5)$$

The log-likelihood is:

$$L(\mu; y) \cong - \sum_i [b_0 \exp(-A\mu)]_i + y_i [A\mu]_i \quad (6)$$

with i indicating all pixel locations and relates the consistency of μ with the projection measurements y_i through the forward-projection operator A . For simplicity, a quadratic penalty was used;

$$R(\mu) = \sum_j \sum_{k \in N_j} \frac{(\mu_j - \mu_k)^2}{2}, \quad (7)$$

with j indexing all voxels, and k indexing the 6 nearest-neighbor voxels around voxel j in 3D (Wang *et al* 2014). For phantom experiments, the penalty strength β was set to a low value of 10^2 relative to a bare-beam b_0 value of 10^5 . The relatively low penalty strength decreased the strength of regularization (smoothing) in the reconstruction, which in turn emphasized high-spatial resolution characteristics of the reconstructed images. The MBIR was initialized using a FDK image reconstructed using the motion-corrected projection matrices. A $(512 \times 512 \times 512)$ voxel grid was used with 0.75 mm isotropic voxel size, and the optimization used 50 iterations with 10 subsets each. Linear forward- and back-projectors implemented on GPU were used for both FDK and PL reconstructions.

3.4. Figures of merit

Image reconstructions were qualitatively and quantitatively assessed to evaluate performance of the motion correction method. Qualitative assessment of motion artifacts included visualization of high-contrast bone details in regions of the temporal bone and sinuses. Quantitative assessment involved analysis of the structural similarity (SSIM) measured between the motion-corrected images and the corresponding motion-free reference images as in Wang *et al* (2004):

$$SSIM(I_1, I_2) = \frac{(2\mu_1\mu_2)(2\sigma_{12})}{(\mu_1^2 + \mu_2^2)(\sigma_1^2 + \sigma_2^2)} \quad (8)$$

where μ_1 and μ_2 are the mean value of images I_1 and I_2 , respectively, σ_1^2 and σ_2^2 are the variance of I_1 and I_2 , respectively, and σ_{12} is the covariance of I_1 and I_2 . SSIM was measured over 100 corresponding axial slices about the central axial slice. Only voxels falling within the head (i.e., excluding air; voxel value $> 0.01 \text{ mm}^{-1}$) were used for calculating SSIM. In addition, the full-width at half-maximum (FWHM) of the point spread function (PSF) formed by the tungsten wire placed in the trachea of the head phantom was measured over 10 axial slices in the motion-corrected images and compared to that of the motion-free reference images (using PL reconstruction). To calculate FWHM, images were reconstructed at finer voxel size $(0.1 \times 0.1 \times 0.5 \text{ mm}^3)$ to over-sample the image within limitations of GPU memory. Line profiles through the center of the wire were sampled

radially over 360°. A Gaussian distribution was fit to each line profile, and the FWHM was averaged over all line profiles and slices. The motion-corrupted images were not suitable to analysis of FWHM, because the wire PSF was too severely degraded by motion artifact.

3.5. Clinical data

The method was further tested in clinical data for patients with known or suspected ICH under an IRB-approved study. Two cases were selected that exhibited strong patient motion artifacts and presented several challenges beyond the phantom studies. First, of course, was that motion patterns were unknown, uncontrolled, and realistic – with one of the cases exhibiting up to ~50 mm motion (noncompliant or involuntary raising of the head during the CBCT scan). Overall motion of the head, jaw, and neck was complex and non-rigid. The CBCT scan data were also collimated to ~12 cm field of view (FOV) along the SI axis to limit patient dose, with the reduced FOV presenting an additional challenge to 3D-2D registration. Registration parameters were as described in section 2.3, and pose was estimated using both 6 and 9 DoF motion models for each projection in the CBCT scan. 3D images were reconstructed with and without motion correction using 250 projections covering a 200° arc (half-scan) centered on the occurrence of motion. The quadratic PL algorithm was used with 200 iterations and 10 subsets to reach convergence ($0.44 \times 0.44 \times 0.44$ mm³ voxels).

To further evaluate the performance of the GO metric under such challenging conditions, the motion correction algorithm was repeated (for Patient #2, who demonstrated a large motion with limited FOV and, therefore, a more difficult registration) using GC and GI as a similarity metric, leaving other parameters unchanged. Images were reconstructed with the resulting PMs for all projections using the PL algorithm and compared with the motion-corrected image obtained using the GO similarity metric.

4. Results

4.1. Phantom images with and without motion correction

Images of the anthropomorphic head phantom with and without motion correction are shown in figure 4. All of the motion profiles described in section 3.2 resulted in strong motion artifacts, and none of the motion-corrupted images would be clinically acceptable. The motion-corrected images are shown in the right column of figure 4 for the most severe motion profile. Motion-corrected images for less severe motion profiles (not shown in figure 4) were qualitatively indistinguishable from the worst case (right column) and are quantified below. The motion-corrected images appear to be free from motion artifact and are qualitatively indistinguishable from the motion-free image for even the most extreme cases (15° lateral rotation, 7.5 mm SI translation, etc.).

Considering the results in figure 4 more closely, we note the different types of motion artifacts resulting from the various motion profiles imparted in the 5 experiments. In Exp 1, the severity of motion artifact increases with the magnitude of A_θ as expected. As evident in Exp 2, the P_θ has a less obvious effect, but it appears that a shorter rotation period (i.e., faster speed of rotation, as in a sudden twitch of the head) results in more severe artifact than

a slower rotation (as in a slow drift of head angulation). The ϕ_θ with respect to the scan orbit is shown in Exp 3 to affect the appearance of motion artifact, and one may distinguish to some extent the pose of the source and detector during object motion based on the directionality of the motion artifacts. Moreover, motion at the midpoint of the scan results in the strongest artifacts, associated with sampling over a short scan. In Exp 4 in-plane features become nearly indistinguishable due to out-of-plane motion contamination that worsens with increasing A_z as expected. The severity of motion artifact in Exp 5 appears to be greatest for the longest scan time, which corresponds with the highest number of projections obtained.

4.2. SSIM

Measurements of SSIM for motion-corrupted and motion-corrected images are shown in figure 5. In each case, SSIM was restored to >0.995 following motion correction, a statistically significant difference (increase) compared to the uncorrected image ($p < 0.001$) in each case. For the motion-corrupted images, many of the trends in SSIM match the qualitative observations of artifact magnitude discussed in section 4.1. For example, SSIM decreases with increasing A_θ , decreases with increasing A_z , and increases with P_θ . For Exp 4, SSIM for the motion-corrupted images appears better (perhaps surprisingly, given the strong artifacts exhibited in figure 4) than in other experiments, indicating that SSIM as a figure of merit is more sensitive to strong gradients about high-contrast bones and the skin line than it is to an overall loss of resolution. The SSIM measured for Exp 3 shows a somewhat unexpected trend increasing with ϕ_θ , compared to visual impression of images in figure 4 suggesting worst performance for motion centered about $\phi_\theta = 100^\circ$. The SSIM measurements for Exp 5 also show a somewhat unexpected trend in that the 20 second scan time might be expected to exhibit the worst SSIM, rather than the 5 second scan. However, these trends have more to do with limitations in SSIM as a metric of visual image quality than the performance of the motion correction method. In each case, the correction method restored SSIM to a level that was equivalent to (or at least statistically indistinguishable from) the motion-free case.

4.3. FWHM

Figure 6 shows the FWHM of the wire PSF following motion correction. The mean FWHM measured for the motion-free reference image was $(0.94 \pm 0.01 \text{ mm})$. By comparison, the FWHM for the various cases of object motion was $\sim 0.94 \text{ mm}$, indicating that the motion correction algorithm was able to recover spatial resolution to approximately the same level as the motion-free case. Without motion correction, the PSF images (not shown) were entirely confounded by artifact and did not permit FWHM analysis.

4.4. Application to clinical CBCT data

Application of the motion correction algorithm to clinical data is illustrated in figures 7 and 8. Patient motion artifacts are evident in the uncorrected images as distortion (blur, doubling, and streaks), particularly in relation to high-contrast structures. The motion correction method is seen to significantly improve image quality for both cases. For such complex motion, the 9 DoF model outperformed the 6 DoF (not shown for brevity) due to the ability to solve for both the source and object positions simultaneously, resulting in greater

accuracy in translation and rotation parameters. This is similar to results found in previous studies comparing 3D-2D registration using 6 DoF with 9 DoF (Otake *et al* 2015). For Patient #1, the motion profile resulting from the registration suggested ~10 mm motion at approximately the midpoint of the scan due to involuntary tremor. The most noticeable improvement is increased delineation of the anterior skull base (figure 7e) and hard palate (figure 7f). Patient #2 exhibited even stronger motion artifacts (figure 8a) that render the uncorrected image largely unusable. In this case, the registration suggested ~50 mm motion during the latter half of the scan, during which the patient lifted his/her head from the support. The motion correction algorithm resolved most of the motion artifacts from the image (figure 8d) with strongly improved visualization of a metal mesh placed within a craniotomy (figure 8e) and an intraventricular shunt within the cranium (figure 8f). Residual artifacts (e.g., at the posterior skin line) may be due to motion/truncation at the edge of the FOV and possibly to non-rigidity of superficial soft tissues and/or patient support.

The results of the clinical study demonstrate the robustness of the algorithm under realistic conditions, including challenges of complex motion and mismatch between MDCT and CBCT images. Both cases exhibited a mismatch in FOV between MDCT and CBCT – with CBCT exhibiting less coverage in the SI direction due to patient positioning and/or narrow collimation – such that the shoulders, neck, and jaw (including dental implants) appear in the DRRs but not in the CBCT projections. The additional gradients can confound the registration by driving the solution to false local maxima associated with high contrast features.

Such sensitivity was especially evident for the GI and GC similarity metrics, which rely on gradient magnitude. As shown in figure 9 for Patient #2, the GI metric (figure 9a) failed to give reliable registration for such strong and complex motion. The GC metric (figure 9b) performed better but was subject to spurious registration failure (false maxima) for numerous views, and the resulting reconstructions demonstrated correspondingly severe artifacts. The GO metric (figure 9c), on the other hand, better incorporated the broad distribution of strong and subtle gradients across the image. These results are consistent with previous studies (De Silva *et al* 2016) in which GO demonstrated superior robustness to image mismatch compared to GI, GC, and other similarity measures under conditions of challenging registration.

5. Discussion and conclusions

The motion correction method presents a promising framework for the mitigation of large-amplitude motion artifacts in CBCT images acquired during IGI, specifically for non-periodic, involuntary motion (e.g., the head) that can be assumed to be locally rigid and well described by 6 or 9 DoF. The method yields a fiducial-free approach that has been shown to be robust against misregistration in the presence of image mismatch between the prior and current scan (e.g., introduction of surgical instruments, collimation, and/or additional anatomy) by virtue of the GO similarity metric. The resulting transform incorporates the intrinsic and extrinsic parameters of the system geometry with patient motion and can therefore be used in combination with systems that have a non-reproducible orbit and/or vibration in the rotation of the gantry. The method demonstrated correction of motion-

corrupted images with image quality that was quantitatively and qualitatively indistinguishable from the motion-free CBCT image. The registration method underlying the correction technique is also highly parallelizable in that multiple projections can be registered simultaneously. The current work initialized each pose with bilinear extrapolation of the two preceding poses, which is not parallelizable but provides a good estimate and fast convergence. Alternatively, one may envision a fully parallelized implementation without such initialization. In such an implementation (depending on the form and search range of the optimizer), the time required for each pose estimate would likely increase due to coarser initialization; however, the total time for registration of all projections would likely decrease and be more compatible with clinical workflow than the current sequential implementation. The method could in principle be run with every scan to remove image artifact stemming from both geometric errors and patient motion; alternatively, the method could run on-demand in cases where motion artifacts appear to confound the imaging task.

The method is well suited to image-guided procedures in which a prior 3D image is available (e.g., MDCT acquired for diagnosis and/or planning, or a CBCT scan acquired at the beginning of the case for surgical navigation). Similarly in IGRT, treatment planning normally involves an MDCT acquired with the patient in the treatment position, and many therapies involve repeat CBCT on successive treatment fractions. In such scenarios, the prior MDCT or CBCT may be suitable as input to the 3D-2D registration underlying the motion correction method.

Note that the motion-corrected CBCT image is reconstructed in the coordinates of the input 3D image, since the 3D-2D registration algorithm takes the origin of the transformation to be the center of the prior 3D image. This may be beneficial for purposes of image guidance, since the motion correction yields an image that is automatically registered to the prior 3D image, which is often the pertinent coordinate system for the surgical navigation system, surgical robot, or linear accelerator. Put another way, the motion correction method described here can be alternatively viewed as a means of automatically registering a CBCT image to a prior MDCT (or CBCT) via 3D-2D registration.

The phantom study used a fast, low-dose, short scan CBCT as the motion-free prior, and results were unchanged (i.e., not improved) when using a high-quality MDCT as the prior image, indicating the robustness of the registration framework to image quality (quantum noise) in the 3D image. Additionally the clinical cases demonstrated realistic situations of moderate to severe motion, as well as mismatch between a prior MDCT image and the projection data, indicating the robustness of the registration framework using the GO metric.

The primary objective of this study was to assess the feasibility of motion correction in conjunction with geometric calibration in cases of strong, involuntary, non-periodic motion, and therefore the run-time of the algorithm was not fully optimized. The framework was implemented in MATLAB (The MathWorks, Inc., Natick MA) and the average registration time per projection was 3 seconds, with total runtime scaling according to the number of projections in the motion-corrupted dataset. Although the method in its current form registers projections sequentially, a parallel registration scheme (all projections registered

simultaneously) could certainly be implemented and decrease total runtime for practical clinical implementation.

Limitations to the motion correction algorithm are similar to those mentioned in Ouadah *et al* (2016) in that the accuracy of the registration and subsequent artifact correction is somewhat dependent on the quality of the 3D volume and the 2D images. The method appears to be robust against quantum noise, evidenced by performance achieved using the lowest-dose CBCT image as the input 3D image for phantom studies. Although the algorithm is robust to many types of artifact present in either of the images (e.g., cone-beam artifacts, scatter, truncation, angular sampling, etc.), anatomical gradients must be sufficiently distinct for their orientation to be matched between DRRs and projection views. Another limitation is the assumption of local rigid motion, which may be applicable for the skull, pelvis, and extremities, but limits the algorithm in application to scenarios such as the abdomen and thorax. However, even in cases of non-rigid deformation of soft tissue, the method may still be applicable to local ROIs where motion can be approximated as rigid.

Overall, the motion correction method produced CBCT images of the same quality as motion-free images, demonstrating robustness in cases of strong non-periodic motion of the object (patient) using either a low-dose, fast, short-scan prior CBCT image or high quality MDCT. Future work includes incorporating multiple rigid components into the registration and increasing the robustness to deformable motion.

Acknowledgments

Research supported by National Institutes of Health Grant No. R01-EB-017226 and academic-industry partnership with Siemens Healthcare (AX Division, Forchheim, Germany). Clinical data were acquired using a CBCT prototype scanner developed in collaboration with Carestream Health (Rochester, NY). Many thanks to Dr. Tharindu De Silva and Dr. Ali Uneri (Biomedical Engineering, Johns Hopkins University) for help with the 3D-2D registration method and Dr. Alejandro Sisniega and Mr. Pengwei Wu (Biomedical Engineering, Johns Hopkins University) for the clinical imaging studies. The authors thank Ms. Robin Belcher (Department of Radiology, Johns Hopkins University) as well as Mr. Robert Meyer (Siemens Medical Solutions USA, Inc., Customer Solutions Group, Baltimore/Washington DC) for assistance with the Zeego imaging system.

References

- Berger M, Müller K, Aichert A, Unberath M, Thies J, Choi JH, Fahrig R, Maier A. Marker-free motion correction in weight-bearing cone-beam CT of the knee joint. *Med Phys*. 2016; 43:1235–48. [PubMed: 26936708]
- Brehm M, Paysan P, Oelhafen M, Kachelrieß M. Artifact-resistant motion estimation with a patient-specific artifact model for motion-compensated cone-beam CT. *Med Phys*. 2013; 40:101913. [PubMed: 24089915]
- Chen GH, Tang J, Leng S. Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med Phys*. 2008; 35:660. [PubMed: 18383687]
- Cho Y, Moseley DJ, Siewerdsen JH, Jaffray DA. Accurate technique for complete geometric calibration of cone-beam computed tomography systems. *Med Phys*. 2005; 32:968–83. [PubMed: 15895580]
- Choi JH, Fahrig R, Keil A, Besier TF, Pal S, McWalter EJ, Beaupré GS, Maier A. Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. Part I. Numerical model-based optimization. *Med Phys*. 2013; 40:091905. [PubMed: 24007156]
- De Silva T, Uneri A, Ketcha MD, Reaungamornrat S, Kleinszig G, Vogt S, Aygun N, Lo SF, Wolinsky JP, Siewerdsen JH. 3D-2D image registration for target localization in spine surgery: investigation of

similarity metrics providing robustness to content mismatch. *Phys Med Biol.* 2016; 61:3009. [PubMed: 26992245]

Dinelle K, Blinder S, Cheng J, Lidstone S, Buckley K, Ruth TJ, Sossi V. Investigation of subject motion encountered during typical positron emission tomography scan. *IEEE Nuc Sci Symp Conf Rec.* 2006; 6:3283–7.

Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. *JOSA A.* 1984; 1(6):612–9.

Hansen, N. The CMA evolution strategy: a comparing review Towards a New Evolutionary Computation. Berlin: Springer; 2006. p. 75-102.

Isola AA, Grass M, Niessen WJ. Fully automatic nonrigid registration-based local motion estimation for motion-corrected iterative cardiac CT reconstruction. *Med Phys.* 2010; 37:1093–109. [PubMed: 20384245]

Jacobson MW, Stayman JW. Compensating for head motion in slowly-rotating cone beam CT systems with optimization transfer based motion estimation. *IEEE Nucl Sci Symp Conf Rec.* 2008:5240–45.

Li B, Chen Z, Jin X, Yu H, Wang G. Experimental measurement of human head motion for high-resolution computed tomography system design. *Optical Engineering.* 2010; 49:063201.

Marchant TE, Price GJ, Matuszewski BJ, Moore CJ. Reduction of motion artefacts in on-board cone beam CT by warping of projection images. *Brit J Radiol.* 2011; 84:251–64. [PubMed: 21081580]

Müller K, Maier AK, Schwemmer C, Lauritsch G, De Buck S, Wielandts JY, Hornegger J, Fahrig R. Image artefact propagation in motion estimation and reconstruction in interventional cardiac C-arm CT. *Phys Med Biol.* 2014; 59:3121–38. [PubMed: 24840084]

Navab, N., Bani-Hashemi, A., Nadar, MS., Wiesent, K., Durlak, P., Brunner, T., Barth, K., Graumann, R. 3D reconstruction from projection matrices in a C-arm based 3D-angiography system Medical Image Computing and Computer-Assisted Intervention—MICCAI'98. Berlin: Springer; 1998. p. 119-29.

Nazarparvar B, Shamsaei M, Rajabi H. Correction of head movements in positron emission tomography using point source tracking system: a simulation study. *Annals of Nuc Med.* 2012; 26(1):7–15.

Otake Y, Wang AS, Stayman JW, Uneri A, Kleinszig G, Vogt S, Khanna AJ, Gokaslan ZL, Siewerdsen JH. Robust 3D-2D image registration: application to spine interventions and vertebral labeling in the presence of anatomical deformation. *Phys Med Biol.* 2013; 58:8535–53. [PubMed: 24246386]

Otake Y, Wang AS, Uneri A, Kleinszig G, Vogt S, Aygun N, Lo SL, Wolinsky JP, Gokaslan ZL, Siewerdsen JH. 3D-2D registration in mobile radiographs: algorithm development and preliminary clinical evaluation. *Phys Med Biol.* 2015; 60:2075. [PubMed: 25674851]

Ouadah S, Stayman JW, Gang GJ, Ehtiati T, Siewerdsen JH. Self-calibration of cone-beam CT geometry using 3D-2D image registration. *Phys Med Biol.* 2016; 61:2613. [PubMed: 26961687]

Penney GP, Weese J, Little JA, Desmedt P, Hill DL, Hawkes DJ. A comparison of similarity measures for use in 2-D-3-D medical image registration. *IEEE Trans Med Imaging.* 1998; 17:586–95. [PubMed: 9845314]

Pluim JP, Maintz JB, Viergever MA. Image registration by maximization of combined mutual information and gradient information. *IEEE Trans Med Imaging.* 2000; 19:809–14. [PubMed: 11055805]

Rit S, Sarrut D, Desbat L. Comparison of analytic and algebraic methods for motion-compensated cone-beam CT reconstruction of the thorax. *IEEE Trans Med Imaging.* 2009; 28:1513–25. [PubMed: 19211348]

Rohkohl C, Bruder H, Stierstorfer K, Flohr T. Improving best-phase image quality in cardiac CT by motion correction with MAM optimization. *Med Phys.* 2013; 40:031901. [PubMed: 23464316]

Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med Phys.* 1985; 12:252–5. [PubMed: 4000088]

Sisniega A, Stayman JW, Yorkston J, Siewerdsen JH, Zbijewski W. Motion compensation in extremity cone-beam CT using a penalized image sharpness criterion. *Phys Med Biol.* 2017; 62:3712. [PubMed: 28327471]

Tang Q, Cammin J, Srivastava S, Taguchi K. A fully four-dimensional, iterative motion estimation and compensation method for cardiac CT. *Med Phys.* 2012; 39:4291–305. [PubMed: 22830763]

- Uneri A, Otake Y, Wang AS, Kleinszig G, Vogt S, Khanna AJ, Siewerdsen JH. 3D-2D registration for surgical guidance: effect of projection view angles on registration accuracy. *Phys Med Biol.* 2014; 59:271–87. [PubMed: 24351769]
- Wagner A, Schicho K, Kainberger F, Kirkfellner W, Grampp S, Ewers R. Quantification and clinical relevance of head motion during computed tomography. *Invest Radiol.* 2003; 38:733–41. [PubMed: 14566184]
- Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process.* 2004; 13:600–12. [PubMed: 15376593]
- Wang AS, Stayman JW, Otake Y, Kleinszig G, Vogt S, Gallia GL, Khanna AJ, Siewerdsen JH. Soft-tissue imaging with C-arm cone-beam CT using statistical reconstruction. *Phys Med Biol.* 2014; 59:1005–26. [PubMed: 24504126]
- Wicklein, J., Kyriakou, Y., Kalender, WA., Kunze, H. SPIE Med Imaging. 2013. An online motion- and misalignment-correction method for medical flat-detector CT; p. 86681S
- Xu J, Sisniega A, Zbijewski W, Dang H, Stayman JW, Wang X, Foos DH, Aygun N, Koliatsos VE, Siewerdsen JH. Modeling and design of a cone-beam CT head scanner using task-based imaging performance optimization. *Phys Med Biol.* 2016; 61:3180. [PubMed: 27025783]
- Yu H, Wang G. Data consistency based rigid motion artifact reduction in fan-beam CT. *IEEE Trans Med Imaging.* 2007; 26:249–60. [PubMed: 17304738]

Figure 1.

Overview of the motion correction method. Each 2D projection in the CBCT scan is registered to the DRR of a previously acquired 3D volume to determine a rigid transformation. The transform is encoded in PM_{motion} (6 DoF transform shown), combining the motion of source, detector, and object (patient) within a combined description of system geometry for 3D image reconstruction.

Figure 2. Experimental setup. (a) Artis Zeego robotic C-arm. (b) Custom motion phantom consisting of an anthropomorphic head attached to a motion controller showing direction of motion in θ and z .

Figure 3.

Motion profiles for each experiment. (a) Exp 1 involved three amplitudes of rotational motion. (b) Exp 2 involved three durations of rotational motion. (c) Exp 3 involved three phases (relative to the scan orbit) at which motion occurred. (d) Exp 4 involved three amplitudes of periodic longitudinal translation. (e) Exp 5 involved three scan durations according to the three scan protocols.

Figure 4.

Axial slices from the motion-corrupted reconstructions (left three columns) for five experiments exhibiting various motion profiles in A_θ , P_θ , ϕ_θ , A_z , and T_{scan} . The right column shows the same axial slice reconstructed using the motion correction method for the worst case motion pattern (corresponding to the 3rd column from the left).

Figure 5.

SSIM (measured across 100 axial slices) for each of the five motion experiments. The SSIM for the motion-corrupted images is shown on the left of each plot, and the SSIM for the motion-corrected images is shown on right (>0.995 for all cases).

Figure 6. FWHM of the PSF of a thin (0.13 mm) tungsten wire placed in the head phantom. The FWHM for the motion-corrected images is shown for each of the five motion studies, compared to the motion-free (Reference) image. Each plot indicates the full range (whiskers), interquartile range (box), and median (horizontal line) in FWHM analyzed in 10 axial slices.

Figure 7. Motion correction method applied to clinical data exhibiting moderate motion artifacts (Patient #1). (a-c) Motion-corrupted and (d-f) motion-corrected CBCT images. Images (b) and (e) show a magnified axial view of the anterior skull base marked by the dashed box in (a) and (d). Images (c) and (f) show a magnified sagittal view of the hard palate.

Figure 8.

Motion correction method applied to clinical data exhibiting severe motion artifacts (Patient #2). (a-c) Motion-corrupted and (d-f) motion-corrected CBCT images. Images (b) and (e) show a magnified axial view in the region of a craniotomy and implanted mesh marked by the dashed box in (a) and (d). Images (c) and (f) show a magnified axial view in the region of an intraventricular shunt.

Figure 9.

Comparison of (a) GI, (b) GC, and (c) GO similarity metrics in motion correction applied to clinical data exhibiting strong motion (patient 2). The GI and GC metrics were subject to false local maxima arising from high-contrast features, whereas GO was more robust and better incorporated the full distribution of strong and subtle gradients in the registration process.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table I

Experimental variables. Corresponding motion profiles are shown in figure 3.

Exp	$A_{\theta} (^{\circ})$	$P_{\theta}(s)$	$\phi_{\theta} (^{\circ})$	$A_z (mm)$	$T_{scan} (s)$
1	{ 5, 10, 15 }	6	100	0	20
2	10	{ 2, 6, 10 }	100	0	20
3	10	2	{ 50, 100, 150 }	0	20
4	5	6	100	{ 2.5, 5, 7.5 }	20
5	5	2	100	5	{ 20, 10, 5 }