Motion Estimation Using a Penalized Image Sharpness Criterion for Resolution Recovery in Extremities Cone-Beam CT

A. Sisniega, J. W. Stayman, Q. Cao, J. Yorkston, J. H. Siewerdsen, W. Zbijewski

***Abstract*— Recent advances in imaging hardware, such as the development of CMOS x-ray detectors, have the potential to enhance spatial resolution of cone-beam CT (CBCT) systems to a level consistent with quantitative imaging of bone microarchitecture (~100 μm detail size). This capability would be of particular value in dedicated extremities CBCT. The accuracy in such applications will be diminished by subtle, sub-mm patient motion that cannot be managed with immobilization.**

We propose an image-based motion compensation method for high-resolution extremities CBCT that requires no fiducials or external trackers. The algorithm allows for the compensation to be applied only within a Region of Interest (RoI), so that the motion can be assumed to be locally rigid. Motion estimation is achieved by optimizing a cost-function that contains an autofocus term that favors sharp images and a penalty term that penalizes non-smooth motion. The non-convex optimization problem is solved using the CMA-ES algorithm. Following evaluation of several image sharpness metrics for application in extremities motion estimation, the variance of image gradient was chosen as the autofocus term. The effects of other parameters of the objective function (e.g. regularization strength) were evaluated in simulation studies of a hand phantom with synthetic motion patterns of variable amplitude (0.25-10 mm). Small motion amplitudes benefited from strong regularization, whereas weaker regularization was preferred for large motions.

Excellent motion compensation was obtained in the simulated data. After compensation, the structural similarity index (SSIM) computed against a static reference volume was > 0.95 for motions up to 1 mm and >0.8 for larger motions. An 80% increase in SSIM compared to uncompensated image was found for the largest motion (10 mm). Real data of a wrist phantom acquired on a CMOS testbench with 0.5 – 10 mm amplitude object motion confirmed improved visualization of the trabeculae and increased SSIM after motion compensation. The method was applied to motion contaminated patient data from the dedicated extremities CBCT, yielding visible reduction of motion artifacts.

The proposed image-based motion compensation provides robust correction of RoI motion in extremities imaging by using a simple, locally rigid motion model coupled with a penalized image sharpness criterion.

***Index Terms*—High-resolution CBCT, motion compensation, extremities imaging, autofocus.**

I. Introduction

Quantitative metrics of bone microarchitecture have been extensively studied in pre-clinical micro-CT imaging and found to provide a sensitive biomarker with potential applications in early detection and staging of osteoarthritis and osteoporosis [1]. Clinical implementation of such metrics is challenged by the relatively small size of the pertinent bone features (e.g. ~100 μm for trabeculae). Encouraging initial results in in-vivo assessment of bone microarchitecture were obtained with dedicated Flat-Panel Detector (FPD) extremities CBCT (Fig. 1A) owing to the high spatial resolution of FPDs [2]. Further improvements in spatial resolution are necessary to establish extremities CBCT as a platform for clinical evaluation of bone morphology. This will involve improvement in hardware, in particular implementation of a CMOS detector, and algorithmic developments, including compensation of patient motion.

The CMOS technology offers smaller pixel size, higher readout speed, and reduced electronic noise compared to FPDs. Fig. 1B shows an experimental testbench emulating the extremities CBCT system and equipped with a CMOS detector with 100 μm pixel pitch. Point Spread Function measurements showed >20% improvement in FWHM with CMOS-based CBCT compared to the current FPD-based CBCT (194 μm pixel pitch).

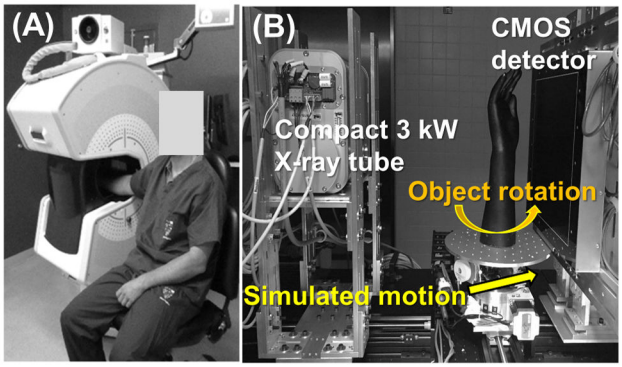

Figure 1. (A) FPD-based extremities CBCT. The system will be upgraded to a CMOS detector to enhance spatial resolution and enable in-vivo quantitative imaging of bone microarchitecture. In addition to the improved hardware, compensation of small, involuntary patient motion will be essential for achieving the required spatial resolution. (B) Experimental testbench implementing the CMOS detector in the geometric configuration of the extremities CBCT. An additional horizontal translation stage is included to enable simulated sample motion.

The application of a CMOS detector is a significant step toward reliable visualization of bone detail. However, even slight, sub-mm patient motion will challenge the accuracy of quantitative assessment of bone microarchitecture. For such small motions, patient immobilization (typically adequate for current applications of extremities CBCT) is not sufficient, and robust motion compensation will be necessary.

This work was supported by NIH Grant R01- EB018896.

A. Sisniega, J. W. Stayman, Q. Cao, J. H. Siewerdsen and W. Zbijewski are with the Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA.

J. Yorkston is with Carestream Health, Rochester, NY, USA.

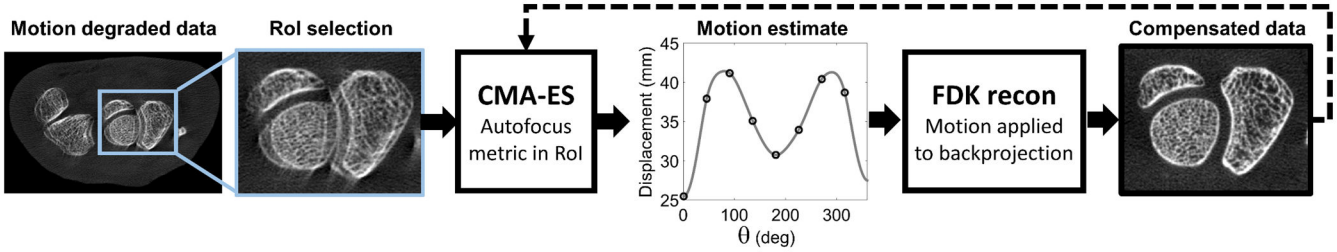

Figure 2. Workflow of the motion compensation. The compensation is applied locally to a RoI where the motion can be assumed rigid. The motion trajectory is represented using a b-spline model and estimated by CMA-ES optimization of a non-convex cost function that maximizes image sharpness.

We propose a motion compensation strategy for extremities CBCT based on the "autofocus" concept [3, 4]. The motion trajectory is estimated by maximizing a metric of image sharpness. Compared to approaches based on fiducial markers [5, 6] or external trackers [7], the proposed method does not require changes in the imaging workflow. Moreover, since the autofocus approach is purely image-based, the motion compensation can be restricted to a specific region of interest (RoI) (in contrast to algorithms relying on 3D-2D registration [8]). In application to bone microarchitecture, the RoIs will often consist predominately of bone voxels. This supports the assumption that the local motion is rigid, greatly simplifying the motion estimation compared to the complex and deformable motion of the whole extremity.

We introduce a new form of the autofocus objective that employs a novel regularization term penalizing large object displacements, and uses a spline-based model of rigid motion. The performance of a variety of image sharpness metrics in extremity motion estimation is investigated. The proposed autofocus objective is non-convex and exhibits local minima. A statistical optimization method is thus applied for motion estimation (compared to the more common choice of a simplex algorithm) and a restart strategy is introduced to homogenize performance across a wide range of motion amplitudes. The method is evaluated with simulated and experimental data. Application to patient data from current clinical CBCT prototype is presented.

II. Motion Compensation Framework

A flowchart of the algorithm is shown in Fig 2. An initial, motion-contaminated reconstruction is obtained and the RoI to be compensated is selected. It is assumed that the motion of the RoI is rigid, even if the extremity as a whole undergoes a more complex transformation. The motion trajectory T consists of a 6 DoF rigid transformation of the RoI at each projection angle θ. Each DoF is represented as a cubic b-spline (B):

$$T(\theta, j) = \sum_{i=0}^{N} c_{ij} B(\theta - \theta_i) \qquad (1)$$

where j enumerates the DoF (j = 1,…,6) and N is the number of spline knots. The motion trajectory of the RoI is estimated by finding the b-spline coefficients c_{ij} through maximization of the following objective function:

$$\hat{T} = \arg\min_T S(T, \mu) - \beta R(T) \qquad (2)$$

where μ is the reconstructed RoI, and $S(T,\mu)$ is an image sharpness metric. In each iteration, S is computed on a volume obtained from a reconstruction for which the current motion estimate T was applied during the backprojection. $R(T)$ is a penalty (regularization) term encouraging smooth motion trajectories, and β is a scalar penalty strength. The regularization penalizes the first order difference of the positions of the RoI in subsequent projections:

$$R(T) = \sum_{k=1}^{8} \sum_{q=2}^{N_\theta} \sqrt{(x_{k,q}(T) - x_{k,q-1}(T))^2 + (y_{k,q}(T) - y_{k,q-1}(T))^2 + (z_{k,q}(T) - z_{k,q-1}(T))^2} \qquad (3)$$

where x_{kq}, y_{kq}, and z_{kq} are the coordinates of the k-th corner of the RoI in projection q.

The optimization in Eq. 2 is not convex and exhibits multiple minima that challenge conventional gradient-based methods. Instead, the minimization was performed with the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [8]. At each iteration, a population of 20 RoI volumes was generated in parallel for a set of candidate motion trajectories using a GPU implementation of Feldkamp (FDK) reconstruction. Failure to converge or residual motion after a fixed number of iterations was handled by restart of the CMA-ES iterations with increased size of the solution space to be explored (σ) [9].

Several image sharpness metrics have been proposed for autofocus compensation of motion and geometric misalignment. In the latter context, entropy and gradient based metrics were shown to be the most appropriate [10]. Here, four metrics were evaluated for the task of extremity motion correction: (i) image variance, previously used in microscopy autofocus applications [11]; (ii) Entropy $E(T,\mu) = \sum_l h_l(\mu) \cdot \log h_l(\mu)$, where h_i in an intensity histogram with 256 bins; (ii) the (negative) squared spatial gradient [3]; and (iv) the (negative) gradient variance.

The performance of the metrics was evaluated in a population of motion-contaminated bone images obtained by simulating 3000 random motion trajectories over a 0-50 mm range of average motion amplitudes (where the average amplitude is equivalent to mean displacement in each motion trajectory).

III. Experimental Evaluation

The evaluation used data acquired on a CMOS-based x-ray testbench (Fig. 1B). The system geometry emulated the extremities CBCT prototype. The CMOS detector was a Dalsa Xineos 3030 (Eindhoven, NL) with a pixel size of 0.1 mm and 600 µm-thick CsI columnar scintillator. A 3kW, small focal spot (0.3 FS), rotating anode x-ray source (IMD RTM 37, Italy) was operated at 90 kV (+0.2 mm Cu), and 0.12 mAs per projection; 720 projections were acquired over 360º. The bench included a linear translation stage that was synchronized with the rotation stage to simulate patient motion. Reconstruction voxel size was 0.075 mm for image evaluation and 0.5 mm for motion

estimation; FDK algorithm with Hann apodization and cutoff at the Nyquist frequency was used.

A simulation study was performed to explore the performance of motion compensation as a function of number of spline knots N and regularization strength β. A cadaveric wrist was imaged on the testbench and reconstructed to obtain a static image. Motion contaminated projections were simulated by applying a projection-wise rigid transformation to the volume followed by a forward projection. Simulated motion trajectories involved translations in the transaxial (x-y) plane with amplitudes ranging from 0.25 mm to 10.0 mm. Each translation was performed as linear motion that began at 90° gantry rotation and finished at 150° gantry rotation. Motion was modelled with spline interpolation with 360 control points to obtain smooth trajectories. Motion compensation involved a maximum of 4000 iterations of CMA-ES applied to a 200x200x20 voxels RoI including 3 carpal bones (see Fig. 5). The optimization was considered converged for changes in the cost function smaller than 10^{-4}. A restart with 4-fold increase in σ was performed for cases for which convergence was not achieved. Maximum runtime was ~30 min if convergence was not reached earlier.

Experimental evaluation involved an anthropomorphic hand phantom acquired on the testbench with motion implemented as a lateral translation of the linear stage with a slope of 1 mm per degree of rotation. A static volume (no motion) and scans with 0.5, 1.0 and 10 mm amplitude motions were acquired

The performance of motion compensation in the simulation and experimental studies was quantified using the structural similarity index (SSIM), with the static image as reference [12]:

$$SSIM = \frac{(2\bar{\mu}_{ref}\bar{\mu}_{MC} + c_1)(2\sigma_{ref-MC} + c_2)}{(\bar{\mu}_{ref}^2 + \bar{\mu}_{MC}^2 + c_1)(\sigma_{ref}^2 + \sigma_{MC}^2 + c_2)} \quad (4)$$

where $\bar{\mu}_j$ is the average attenuation and σ_j is the variance of the attenuation values in image j. The index ref denotes the reference static image, MC denotes the motion compensated image, and $\sigma_{ref\text{-}MC}$ is the covariance between the two images.

The method was also applied to a motion contaminated patient knee scan obtained on the current generation FPD-based extremities CBCT.

IV. Results

The normalized value of the various image sharpness metrics is shown as a function of average motion amplitude in Fig. 3 (the metrics were normalized by subtracting their minimum value and dividing by maximum value so that 0 represents a perfectly focused image).

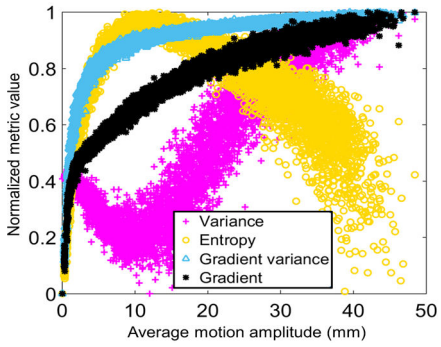

Figure 3. Normalized autofocus sharpness metrics as a function of average motion amplitude for an ensemble of fifty random realizations of motion. Gradient based metrics showed a lower number of local minima and monotonic increase with motion.

An ideal sharpness metric for motion estimation should monotonically decrease towards a global minimum corresponding to a static image. At a fixed motion amplitude, narrower dispersion of the metric is preferred, as it indicates weaker local minima (the metric is consistent across motions with the same average displacement but different trajectories). Image variance is not monotonic, showing a strong local minimum at ~10 mm amplitude. Entropy is monotonically increasing for motion amplitudes of up to 10 mm, but decreases with increasing motion at larger amplitudes. This reflects the fact that images with significant motion blur are relatively uniform and thus exhibit small entropy. While successful motion compensation has been shown with entropy-based metrics [4], this particular objective may thus not be well suited for large motions. Both gradient-based metrics are monotonically increasing throughout the investigated range of motions. Gradient variance exhibits lower dispersion and was chosen for the studies presented here.

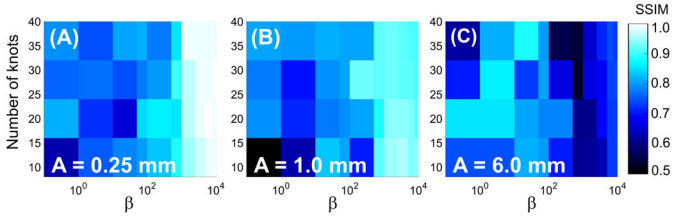

Figure 4. SSIM as a function of number of knots in the b-spline motion model (N) and regularization strength (β) for small (A), moderate (B) and large (C) motions.

Fig. 4 shows the performance of motion compensation as a function of regularization strength β and number of knots in the motion model N. Only CMA-ES runs with no restart were included in this investigation. Strong regularization is preferred for small motion amplitudes, achieving almost perfect correspondence with the reference image (SSIM > 0.9). As motion amplitude increases, the value of β yielding maximum SSIM decreases. Optimal β for 6 mm motion is 10^2-10^3x smaller than that for sub-mm motion. This is likely because large β encourages solutions that smooth out the large motions. The trend in the number of knots is weaker, with a combination of a low number of knots and small β yielding sub-optimal performance, which may involve solutions with oscillatory b-spline motion patterns. For the 6 mm motion, combination of a large number of knots with moderate β (~10^3) resulted in slow convergence (not reached within the 4000 CMA-ES iterations), indicating the need for a restart. Fig. 5 shows a selection of image results from the simulation study in Fig 4. Motion-compensated images corresponding to parameters of the objective function yielding maximum SSIM are compared to uncompensated reconstructions and the reference static volume for 0.5 mm and 10 mm motion. Significant reduction in motion-induced artifacts and recovery of trabecular detail are achieved for both small and large motions. The plot of maximum SSIM (obtained at "optimal" values of β and N) as a function of motion amplitude shows almost perfect structure recovery (SSIM > 0.95) for motions <= 1 mm and significant improvement compared to no compensation for larger motions (~2x increase in SSIM). For motions > 5 mm, CMA-ES restart is essential for optimal results.

Fig. 6 shows the results of benchtop experiments. Even for the relatively small 1 mm motion, significant deterioration in the visualization of the trabeculae is found (arrows). Motion compensation successfully recovers the trabecular structure. A 30% improvement in SSIM was achieved, confirming the results of the simulations in real data.

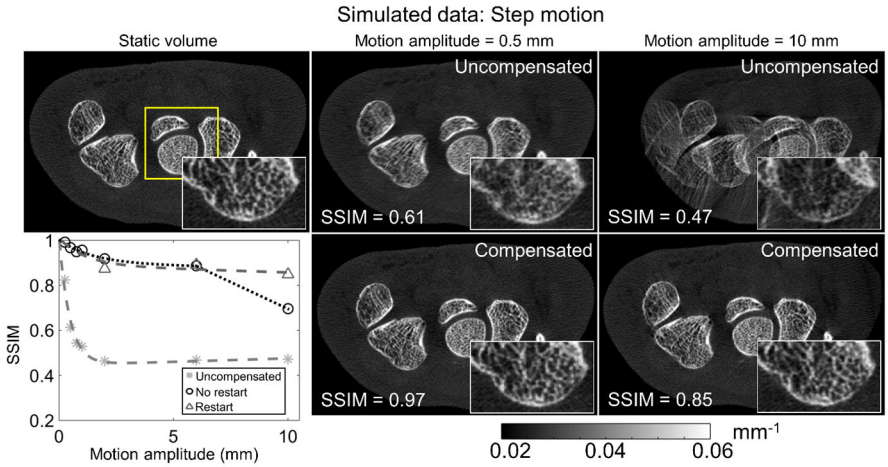

Figure 5. Motion compensation in simulated data for small (0.5 mm) and large (10 mm) motion amplitudes using optimization parameters yielding best SSIM for a given motion amplitude. The performance of the method improves for smaller motion amplitudes, but recovery of the trabecular structure and significant reduction of artifacts is apparent in both cases. This is quantified by the plot of SSIM (computed against the reference static volume) as a function of motion amplitude. Compensation using CMA-ES without restart (circles) is compared to that with restart (triangles), showing the benefits of restart for cases with large motion amplitudes (>5 mm). The square in the static image marks the RoI used for motion estimation.

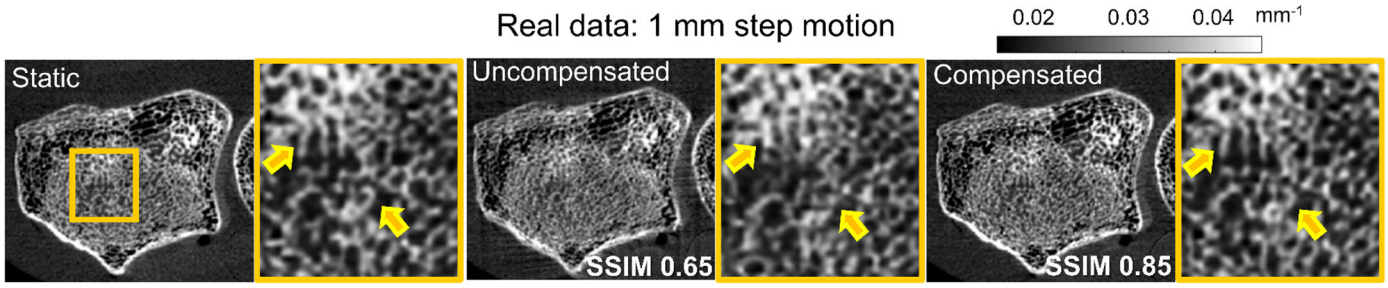

Figure 6. Motion compensation in experimental testbench data with step motion generated by a linear translation of the volume during the acquisition. Motion-induced artifacts (double contours) are reduced and details of the trabecular architecture are recovered after compensation (arrows).

Fig. 7 shows the application of the compensation algorithm to patient data from FPD-based extremities CBCT with artifacts due to insufficient immobilization. The scan was processed using the proposed method using a 9x9x1.2 cm RoI centered at the femoral head. Significant reduction of motion artifacts is apparent. This indicates that the motion found in an RoI can in some cases be applied outside of the RoI to yield partial correction of the artifacts throughout the volume.

VI. Conclusion

A purely image-based motion compensation framework for high resolution CBCT extremities imaging was presented and evaluated. The performance of the method as a function of the autofocus metric, regularization strength and motion model was investigated. The algorithm recovered trabecular structure and suppressed motion artifacts across a broad range of motion amplitudes. In particular, almost perfect correction was achieved for sub-mm motions representative of small, involuntary patient drift that cannot be controlled with immobilization. Compensation of such motion is essential for emerging applications in quantitative imaging of bone microarchitecture. The robust motion correction framework will be essential for realizing the gains in spatial resolution expected with new imaging hardware for extremities CBCT.

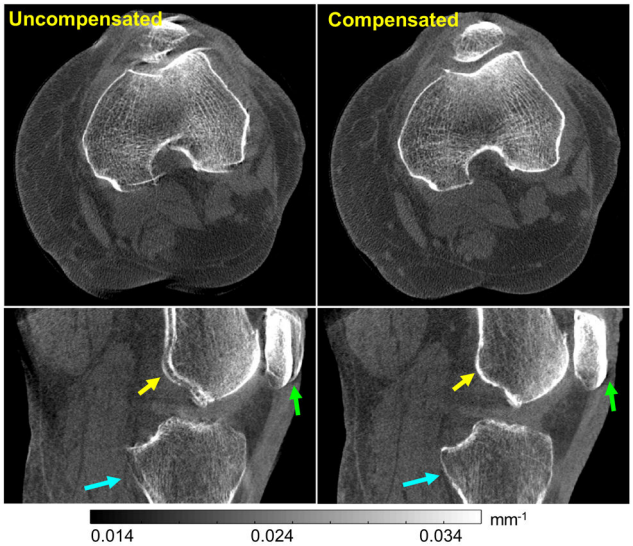

Figure 7.Motion compensation in patient data.

References

[1] H. Weinans, et al., *Bone* 51(2), pp. 190–6, 2012.
[2] A. Muhit, et al. *Proc. SPIE Medical Imaging* vol. 8672, 867203, 2013.
[3] A. Kingston, et al. *Med. Phys.*, vol. 38, no. 9, pp. 4934, 2011.
[4] J. Wicklein, et al., *Proc. SPIE Medical Imaging* vol. 8668, 86681S, 2013.
[5] J.-H. Choi, et al., *Med. Phys.*, vol. 41, no. 6, pp. 061902, 2014.
[6] M. W. Jacobson and J. W. Stayman, *IEEE Nucl. Sci. Symp. Conf. Rec.*, pp. 5240, 2008.
[7] J.-H. Kim, et al., *Phys. Med. Biol.*, vol. 60, no. 5, pp. 2047–2073, 2015.
[8] M. Unberath, et al., *Proc. SPIE Medical Imaging* vol. 9413, pp. 94130D, 2015.
[9] N. Hansen and S. Kern, *Proc. 8th Int. Conf. Parallel Probl. Solving from Nat. - PPSN VIII*, vol. 3242/2004, pp. 282–291, 2004.
[10] J. Wicklein, et al., *Med. Phys.*, vol. 39, no. 8, pp. 4918–31, 2012.
[11] J. M. Mateos, et al., *Cytometry Part A*, vol. 81A, pp. 213-221, 2012.
[12] Z. Wang, et al., *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.