Image Quality Assessment for Rigid Motion Compensation

Alexander Preuhs[1], Michael Manhart[2], Philipp Roser[1], Bernhard Stimpel[1], Christopher Syben[1], Marios Psychogios[3], Markus Kowarschik[2], Andreas Maier[1]
[1]Pattern Recognition Lab, Friedrich-Alexander Universität, Erlangen, Germany
[2]Siemens Healthcare GmbH, Forchheim, Germany
[3]Department of Neuroradiology, Universitätsspital Basel, Basel, Switzerland
`alexander.preuhs@fau.de`

Abstract

Diagnostic stroke imaging with C-arm *cone-beam computed tomography* (CBCT) enables reduction of time-to-therapy for endovascular procedures. However, the prolonged acquisition time compared to helical CT increases the likelihood of rigid patient motion. Rigid motion corrupts the geometry alignment assumed during reconstruction, resulting in image blurring or streaking artifacts. To reestablish the geometry, we estimate the motion trajectory by an autofocus method guided by a neural network, which was trained to regress the reprojection error, based on the image information of a reconstructed slice. The network was trained with CBCT scans from 19 patients and evaluated using an additional test patient. It adapts well to unseen motion amplitudes and achieves superior results in a motion estimation benchmark compared to the commonly used entropy-based method.

1 Introduction

Mechanical thrombectomy [16, 2] is guided by an interventional C-arm system capable of 3-D imaging. Although its soft tissue contrast is comparable to helical CT scans, the prolonged acquisition time pose C-arm CBCT more susceptible to rigid head motion artifacts [13]. In the clinical workflow, however, it is desirable to reduce the time-to-therapy by avoiding prior patient transfers to helical CT or MR scanners [18]. To this end, robust motion compensation methods are desirable.

Methods for rigid motion compensation can be clustered in four categories: 1) image-based autofocus [19, 21], 2) registration-based [15], 3) consistency-based [5, 17] and 4) data-driven [3, 12, 11].

Recent data-driven approaches use image-to-image translation methods based on GANs [12, 11] or aim to estimate anatomical landmarks in order to minimize a *reprojection error* (RPE) [3]. The latter approach does not provide the required accuracy, whereas GAN-based approach are deceptive for clinical applications, as the data-integrity cannot be assured [8].

We propose a learning-based approach for rigid motion compensation ensuring data integrity. An image-based autofocus method is introduced, where a regression network predicts the RPE directly from reconstructed slice images. The motion parameters are found by iteratively minimizing the predicted RPE using the Nelder-Mead simplex method [14].

2 Motion Estimation and Compensation Framework

Autofocus Framework: Rigid motion is compensated by estimating a motion trajectory $\mathcal{M}$ which samples the motion at each of the N acquired views within the trajectory [10]. $\mathcal{M}$ contains the motion

matrices $\mathbf{M}_i$, where each motion matrix $\mathbf{M}_i \in \mathbb{SE}(3)$ — with $\mathbb{SE}(3)$ being the special Euclidean group — describes the patient movement at view $i \in [1, N]$. The motion matrices can be incorporated in the backprojection operator of a *filtered backprojection*-type (FBP) reconstruction algorithm. We denote the reconstructed image in dependence of the motion trajectory by $\text{FBP}_y(\mathcal{M})$, where FBP_y is the FDK-reconstruction [4] from projection data y. In the following, FBP_y will reconstruct the central slice on a 512^2 pixel grid using a sharp filter kernel to emphasize motion artifacts.

Typical autofocus frameworks (cf. [19]) estimate the motion trajectory based on an *image quality metric* (IQM) evaluated on the reconstructed image by minimizing

$$\underset{\mathcal{M}}{\arg\min} \quad \text{IQM}(\text{FBP}_y(\mathcal{M})) \ . \tag{1}$$

A common problem in solving (1) is the non-convexity of the IQM, which is typically chosen to be the image histogram entropy or total variation of the reconstructed slice. To overcome this limitation, we propose to replace the IQM by a network architecture that is trained to regress the RPE, which was shown to be quasi convex for a geometric reconstruction problem [9].

Learning to Assess Image Quality: Let $\mathcal{X}$ be a set of 3-D points $\mathbf{x} \in \mathbb{P}^3$ uniformly sampled from a sphere surface and let the acquisition trajectory associated to a dataset y be defined by projection matrices $\mathbf{P}_i \in \mathbb{R}^{3 \times 4}$ mapping world points on the detector of a CBCT system at view i [6], then the RPE is computed as

$$\text{RPE}(\mathcal{M}) = \frac{1}{|\mathcal{X}|N} \sum_{\mathbf{x} \in \mathcal{X}, i \in N} ||\mathbf{P_i M_i x} - \mathbf{P_i x}||_2^2 \ . \tag{2}$$

This metric measures the reconstruction-relevant deviations induced by motion [20] and can thus be expected to be estimated directly from the reconstruction images. To this end, we device a regression network learning the RPE directly from a reconstructed image. Our regression network R consists of a feature extraction stage, pretrained on ImageNet and realized by the first 33 layers from a residual network, [7] followed by a densely connected layer defining the regression stage. The cost function L of the network is defined by the difference between the network-predicted RPE from a reconstruction slice with simulated motion trajectory $\mathcal{M}$ and the corresponding RPE as defined by Eq. (2)

$$L = ||R(\text{FBP}_y(\mathcal{M})) - \text{RPE}(\mathcal{M})||_2^2 \ . \tag{3}$$

For training, the projection data y is ensured to be motion free, such that motion artifacts solely source from the virtual motion trajectory $\mathcal{M}$. For training and testing, we use CBCT acquisitions (Artis zee Q, Siemens Healthcare GmbH, Germany) of the head ($N = 496$) acquired from 20 patients which were split in 16 for training 3 for validation and 1 for testing. For each patient we simulate 450 random motion trajectories resulting in a training set of 7650 reconstructions.

3   Experiments and Results

For motion generation, we use rotational movements along the patient's longitudinal axis. The motion trajectory is modeled by an Akima spline [1] with 15 equally distributed nodes inducing RPEs ranging from 0 mm to 0.6 mm. With the RPE measurement being sensitive to constant offsets, not inducing motion artifacts, we further only use motions affecting a third of the acquisition.

First, we inspect how well the network is able to regress the RPE on test and validation data. Then, in in an inverse crime scenario — i.e. the modeling capacity of the spline used for motion generation is equal to the spline used for motion compensation — we inspect the behavior for motion types significantly varying in their shape from any motion seen during training. In a last experiment we compare the performance of the network with a state-of-the-art IQM utilizing the histogram entropy. Therefore, we deploy an inverse crime scenario and a more realistic case where we use 10 spline nodes for motion generation and 20 nodes for compensation.

Regression Network: We use Adam optimization with learning rate of 0.00001 and select the network parameters that achieved the best RPE prediction on our validation dataset. Our network achieves an average RPE deviation from the Gt of 0.031 mm on the test dataset, as depicted in Fig. 1.
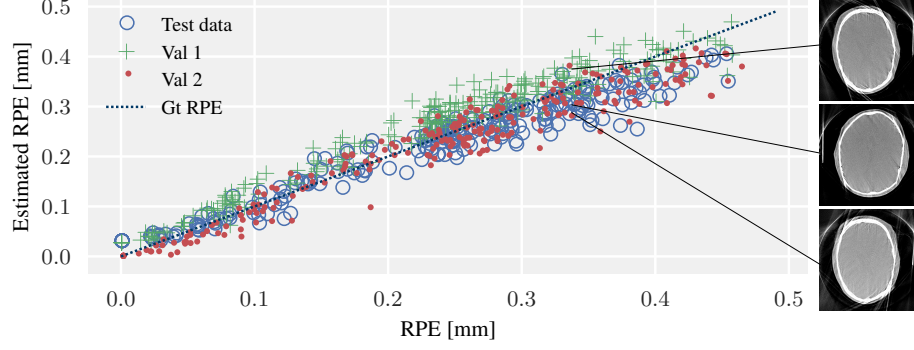
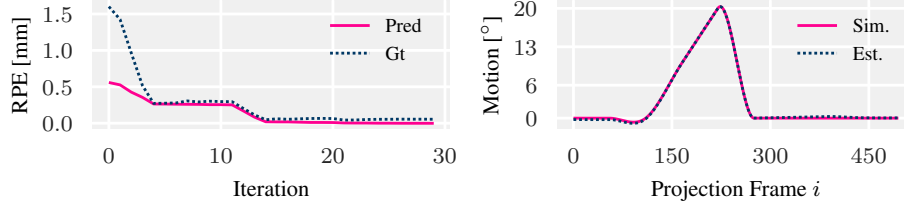Figure 1: Network estimated RPE and different reconstructions, all revealing a RPE of ≈ 0.34 mm.

Figure 2: Left: Network-predicted and Gt RPE in each iteration step of the optimization. Right: Simulated motion trajectory and estimated motion trajectory after optimization.

Network Inference for Motion Compensation: Using the test patient, the network behavior for motion exceeding the RPE of the training process is inspected in an inverse crime scenario. The simulated motion trajectory is depicted in Fig. 2 together with the estimated motion trajectory after optimization using the network as IQM (cf. Eq. 1). For each iteration of the optimization process the network predicted RPE together with the corresponding Gt RPE is depicted. While the RPE is underestimated within the first iterations, the proportionality is still kept, guiding the optimization to a motion free reconstruction.

Figure 3 compares the proposed network-based IQM with the entropy-based IQM. The optimization process is identically for both metrics. In an inverse crime scenario both methods can restore the original image quality, however, in a more realistic setting the image entropy is stuck in a local minimum, whereas the network is able to lead the optimization to a nearby motion-free solution.

4   Conclusion and Discussion

We present a novel data driven autofocus approach lead by a convolutional neural network. The network is trained to predict the RPE given a slice of a CBCT reconstruction. The final motion
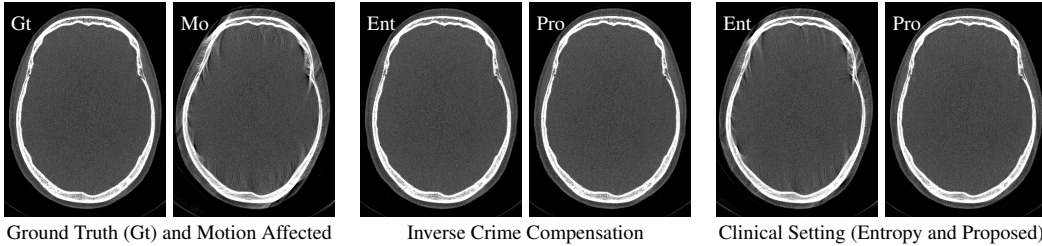

Ground Truth (Gt) and Motion Affected   Inverse Crime Compensation   Clinical Setting (Entropy and Proposed)

Figure 3: Reconstructions of the test patient using [500-2000] HU window. In the inverse crime scenario, the SSIM to the Gt is 0.84 (Ent/Gt) and 0.95 (Pro/Gt), respectively for the entropy (Ent) and proposed (Pro) measure. For the more realistic setting (Clinical Setting) the SSIM is 0.65 (Ent/Gt) and 0.84 (Pro/Gt), respectively.

compensated reconstruction is solely based on the projection raw-data and the estimated motion trajectory. This allows us to device a learning-based motion compensation approach while ensuring data integrity. We showed that the network is capable of generalizing well to unseen motion shapes and achieves higher SSIM compared to a state-of-the-art IQM measure.

Disclaimer: The concepts and information presented in this paper are based on research and are not commercially available.

References

[1] H. Akima. A new method of interpolation and smooth curve fitting based on local procedures. *Journal of the ACM (JACM)*, 17(4):589–602, 1970.

[2] O. A. Berkhemer, P. S. Fransen, D. Beumer, L. A. van den Berg, H. F. Lingsma, A. J. Yoo, W. J. Schonewille, J. A. Vos, P. J. Nederkoorn, M. J. Wermer, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. *N Engl J Med*, 372(1):11–20, 2015.

[3] B. Bier, K. Aschoff, C. Syben, M. Unberath, M. Levenston, G. Gold, R. Fahrig, and A. Maier. Detecting anatomical landmarks for motion estimation in weight-bearing imaging of knees. *MICCAI Workshop on MLMIR*, pages 83–90, 2018.

[4] L. A. Feldkamp, L. Davis, and J. W. Kress. Practical cone-beam algorithm. *J Opt Soc Am A*, 1(6):612–619, 1984.

[5] R. Frysch and G. Rose. Rigid motion compensation in C-arm CT using consistency measure on projection data. *MICCAI*, pages 298–306, 2015.

[6] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[7] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. *CVPR*, pages 770–778, 2016.

[8] Y. Huang, A. Preuhs, G. Lauritsch, M. Manhart, X. Huang, and A. Maier. Data consistent artifact reduction for limited angle tomography with deep learning prior. *MICCAI Workshop on MLMIR*, 2019.

[9] Q. Ke and T. Kanade. Quasiconvex optimization for robust geometric reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(10):1834–1847, 2007.

[10] J. Kim, T. Sun, J. Nuyts, Z. Kuncic, and R. Fulton. Feasibility of correcting for realistic head motion in helical CT. *Fully 3-D*, pages 542–545, 2014.

[11] T. Küstner, K. Armanious, J. Yang, B. Yang, F. Schick, and S. Gatidis. Retrospective correction of motion-affected mr images using deep learning frameworks. *Magn Reson Med*, 82(4):1527–1540, 2019.

[12] S. Latif, M. Asim, M. Usman, J. Qadir, and R. Rana. Automating motion correction in multishot mri using generative adversarial networks. *MedNeurips*, 2018.

[13] J. R. Leyhe, I. Tsogkas, A. C. Hesse, D. Behme, K. Schregel, I. Papageorgiou, J. Liman, M. Knauth, and M.-N. Psychogios. Latest generation of flat detector ct as a peri-interventional diagnostic tool: a comparative study with multidetector ct. *JNIS*, 9(12):1253–1257, 2017.

[14] D. M. Olsson and L. S. Nelson. The nelder-mead simplex procedure for function minimization. *Technometrics*, 17(1):45–51, 1975.

[15] S. Ouadah, W. Stayman, J. Gang, T. Ehtiati, and J. Siewerdsen. Self-calibration of cone-beam CT geometry using 3D–2D image registration. *Phys Med Biol*, 61(7):2613, 2016.

[16] W. J. Powers and et al. 2015 aha/asa focused update of the 2013 guidelines for the early management of patients with acute ischemic stroke regarding endovascular treatment. *Stroke*, 46(10):3020–3035, 2015.

[17] A. Preuhs, A. Maier, M. Manhart, J. Fotouhi, N. Navab, and M. Unberath. Double your views – exploiting symmetry in transmission imaging. *MICCAI*, pages 356–364, 2018.

[18] M. Psychogios, D. Behme, K. Schregel, I. Tsogkas, and M. Knauth. One-stop management of acute stroke patients: Minimizing door-to-reperfusion times. *Stroke*, 48(11):3152–3155, 2017.

[19] A. Sisniega, J. W. Stayman, J. Yorkston, J. Siewerdsen, and W. Zbijewski. Motion compensation in extremity cone-beam ct using a penalized image sharpness criterion. *Phys Med Biol*, 62(9):3712–3734, 2017.

[20] N. K. Strobel, B. Heigl, T. M. Brunner, O. Schuetz, M. M. Mitschke, K. Wiesent, and T. Mertelmeier. Improving 3d image quality of x-ray c-arm imaging systems by using properly designed pose determination systems for calibrating the projection geometry. *Medical Imaging*, pages 943–954, 2003.

[21] J. Wicklein, H. Kunze, W. A. Kalender, and Y. Kyriakou. Image features for misalignment correction in medical flat-detector ct. *Med Phys*, 39(8):4918–4931, 2012.