Extended Partial Angle Based Motion Compensation for Dental CBCT

Cristina Sarti[1], Mikhail Mikerov[1], and Claudio Landi[1]

See Through S.r.l, via Bolgara 2, 24060 Brusaporto BG, Italy
cristina.sarti@seethrough.one

Abstract. Motion artifacts can degrade image quality in dental cone-beam CT and complicate diagnosis. In some cases, an exam retake is necessary, resulting in additional radiation exposure for the patient, without any guarantee of improved image quality. Therefore, motion compensation methods play a crucial role. Many methods are time-consuming since they require several reconstructions. We propose a very efficient method that requires only two partial-angle reconstructions. It assumes that the patient remains still during the acquisition, except for a short interval. In this situation, two motion-free partial-angle reconstructions, one before and one after patient motion, can be reconstructed. Motion compensation is achieved by registering forward projections of the two volumes. To enhance the robustness of the registration step, we simulate an extended angular range covered by the two partial volumes using a conditioned U-Net trained on a target-specific dataset. Qualitative analysis shows that we can significantly reduce the appearance of motion artifacts even in the case of challenging motion patterns.

Keywords: Dental CBCT · Motion compensation · Partial-angle reconstruction · Angular range extension

1 Introduction

Dental cone-beam CT (CBCT) is a fully 3D imaging modality that provides accurate volumetric images of the oral cavity. Its applications include implant planning, detection of caries and periodontal diseases, orthodontic treatment, and exodontics. [9]. A typical dental CBCT device consists of a C-arm with mounted x-ray source and detector on the opposite sides, which can rotate in the horizontal plane around the patient's head. The patient can either stand or sit, although the standing position is more common. During the scan, the patient's chin is resting on a chin rest, the head is stabilized using head fixation, and the patient's arms hold a handle to minimize body movement. Additionally, the patient may be asked to bite on a bite block that is attached to the chin rest for more stability. The absence of motion is crucial for good image quality since image reconstruction algorithms rely on a well defined geometry at each acquisition step.

Despite fixation, the patient can still move during scans with a typical duration between 5 and 40 seconds. The images affected by motion artifacts are

often characterized by blurry edges, undefined contours or doubled appearance of structures [7]. Due to the ALARA[1] principle, image retakes should be avoided to reduce radiation dose even at the cost of image quality. Therefore, it becomes important to be able to reduce the appearance of motion artifacts in the image domain using non-optimal projection data.

Several solutions have been proposed in recent years to perform motion correction in dental CBCT. One class of such solutions relies on additional motion tracking hardware installed on the C-arm that enables the estimation of the motion direction and amplitude to modify the projection data [11]. These solutions are rarely used in clinical practice, since they increase device costs and complicate installation, calibration, and maintenance.

Another class of solutions relies on extracting the necessary motion parameters from the acquired projection data and imperfect reconstructions. Sun *et al.* proposed an iterative joint motion estimation and image update method [12]. Maur *et al.* estimates the motion parameters by detecting, reconstructing and forward projecting the contours of anatomical features [6]. Autofocus approaches aiming to improve some metric in the image domain, e.g., image sharpness in the bone region, have also been applied to this problem [10].

Previously, we presented an approach for motion compensation based on the registration of two volumes reconstructed using only limited angular ranges [8].We assumed that the patient remained still in an initial position, then briefly moved and remained still in a second position. We reconstructed two partial volumes corresponding to the two motion-free segments of the full scan and we added the projections affected by motion to one of the two volumes. In general, very little common information is present in both volumes since they are reconstructed using non-overlapping projection data. However, if two volumes do share common edges, it is possible to register their forward-projected data to estimate more reliably the motion parameters needed for correction. By excluding the projections affected by motion from both reconstructions can also have a positive impact on the final motion parameter estimation, especially if the motion duration exceeds one second.

In this work, we extend and optimize our previous method. We exclude the motion-affected projections from each partial reconstruction and incorporate a conditioned U-Net to modify partial-angle reconstructions to cover a larger angular range, thus increasing the amount of common features when the volumes are forward projected.

2 Materials and Methods

In this section, we provide details about all the steps of our motion compensation method. These steps include generation of two partial-angle reconstructions, optimization of the motion compensation parameters with 2D-3D registration and training and application of a conditioned U-Net for volume modification in the image domain.

[1] As Low As Reasonably Achievable

2.1 Geometric calibration

During installation, a CBCT device is usually calibrated with an offline calibration process. A geometric calibration using specially designed phantoms is performed to calculate so-called *projection matrices* corresponding to each detector position i of a total of N positions assumed by the C-arm during the X-ray scan. They contain the parameters describing the projection geometry of the 3D patient head onto the 2D detector. Each matrix $P_i \in \mathbb{R}^{3\times 4}$ can be written as a product of two matrices:

$$P_i = K \cdot G_i, \quad i = 0, \ldots, N. \tag{1}$$

The matrix $G_i = [R_i|t_i] \in \mathbb{R}^{4\times 4}$ contains the extrinsic parameters, i.e., the parameters describing the relative translation and rotation of the X-ray source with respect to the patient's position. The matrix $K \in \mathbb{R}^{3\times 4}$ contains the intrinsic parameters of the projections such as the source-detector-distance, the detector origin and the pixel size [4].

Patient motion during acquisition can partially or entirely invalidate the parameters estimated during calibration. The goal of motion compensation is to restore a correct geometry and to output an artifact-free reconstruction. To reach this goal, our method calculates a transformation matrix ${G_i}' = [{R_i}'|{t_i}'] \in \mathbb{R}^{4\times 4}$ for each device position i such that

$${P_i}' = K \cdot G_i \cdot {G_i}', \quad i = 0, \ldots, N. \tag{2}$$

The matrix ${G_i}'$ contains the rotation and translation parameters that best compensate for patient motion.

2.2 Partial-angle reconstruction

Similarly to our perviously proposed method [8], we perform motion compensation by partial-angle volume registration. With *partial-angle reconstruction* we describe a reconstruction that only includes a small subset of the full set of projections collected during a scan. Figure 1 compares a complete reconstruction and a partial-angle reconstruction covering 60 degrees. Both volumes are reconstructed using the Feldkamp-Devis-Kress (FDK) algorithm [3]. The partial-angle reconstruction shows artifacts due to the fact that only the information contained in a small subset of the projections is reconstructed.

In our approach, we assume that the patient remains still during the first portion of the scan, then moves and remains still in a new position during the final part of the scan. In our previous method, we registered the two partial-angle volumes assuming that the portion of the scan affected by motion was small compared to the one or the other partial reconstruction. So, we included the projections affected by motion in one or in the other reconstruction. The method turned out to be very robust for abrupt motion covering 0.5 seconds but reliability decreased for motion exceeding 1 second.

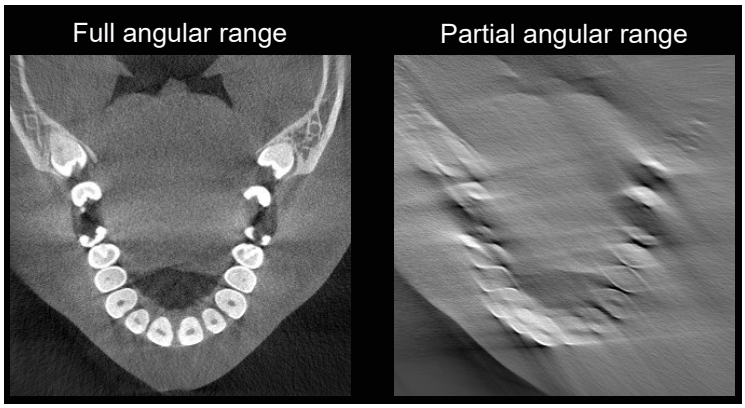

Fig. 1: Complete reconstruction and partial-angle reconstruction (60 degrees).

To be able to compensate for longer motion, we propose a new strategy and modify some of the steps to generate the reference partial-angle reconstructions and to perform registration: (1) The projections affected by motion are no longer included in one of the partial reconstructions. This means that two partial-angle reconstructions are obtained using only motion-free projection data acquired before and after patient motion. The angular gap between the two partial reconstruction corresponds to the motion extent. (2) In order to extend the angular range covered by each partial volume, we apply a suitably trained U-Net. We simulate the addition of projections within the angular gap to each of the two original partial volumes. This leads to a higher number of common structures and stabilizes the subsequent 2D-3D registration step.

2.3 Conditioned U-Net

U-Nets are widely used for image-to-image translations in medical imaging. We apply a U-Net architecture to transform partial-angle reconstructions obtained from data acquired over a limited angular range $\Delta\theta$, so that they approximate reconstructions covering a slightly extended range $\Delta\theta + \Delta\phi$. In general, the appearance of the partial volumes will vary depending on whether the additional angular range $\Delta\phi$ is added before the first or after the last angle in the initial reconstruction. We control the output of the network by defining the side to which the additional angular range should be added.

Architecture For this task, we added a control mechanism in form of simple Feature-wise Linear Modulation (simple FiLM) layers to the encoder block of a standard 3D U-Net working on 3D patches [2]. The simple FiLM layer is an additional fully connected layer in each double convolution block that receives an indicator of extension direction as either 0 or 1 and outputs two scalars, the multiplicative factor γ and the offset β. They are used to modify the output of the block. The network's architecture is displayed in Figure 2.

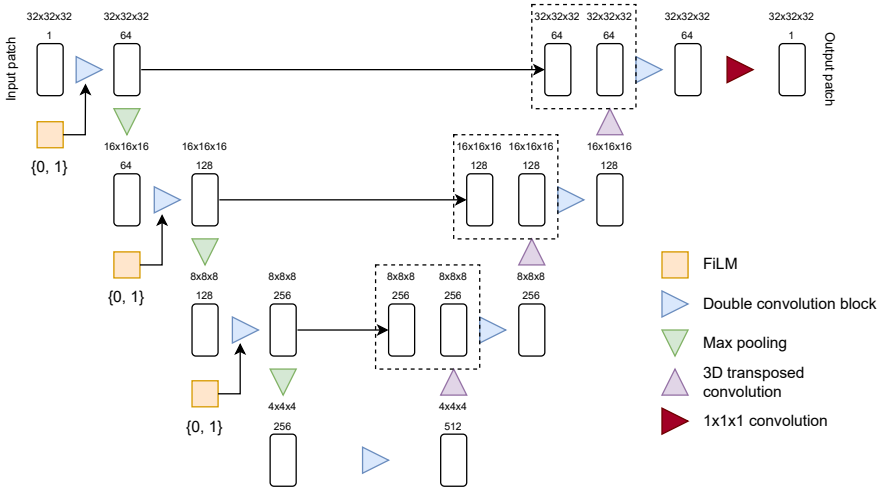

Fig. 2: Architecture of a 3D U-Net with simple FiLM modulation layers. Black arrows represent skip connections; dashed rectangles represent data concatenation. Patch size and number of channels are displayed above each patch.

Dataset preparation Dental cone-beam CT data is characterized by high variability. For example, some patients may have implants, crones, or other metal inserts which can lead to significant image artifacts even without motion. Moreover, missing teeth are a common pathology. The operator can influence the image, too, by instructing the patient to bite on a bite block, thereby introducing an additional visible structure into the scan. Finally, another source of variability is the time point at which the patient moves and the motion extent. We constructed our dataset to account for as much variability as possible.

We started by selecting seven highly variable patients. Then, we randomly sampled the starting projection number of a partial-angle reconstruction from a uniform distribution of possible projection numbers. If the projection number was too close to an already added projection number in the dataset, it was not included. Finally, for each angle, we randomly assigned one of the seven patients. Figure 3 displays the distribution of the input volumes corresponding to the selected patients and projection numbers in the dataset. The same reconstruction is included twice in the dataset if it can be extended in both directions. Otherwise, it appears in the dataset only once.

Training The conditioned U-Net was trained using PyTorch's automatic mixed precision package for 13 epochs on $32 \times 32 \times 32$ patches with stride 16 in all three directions. The patches were extracted from $316 \times 316 \times 344$ volumes with 320 µm voxel size. The training dataset contained reconstructions from the first

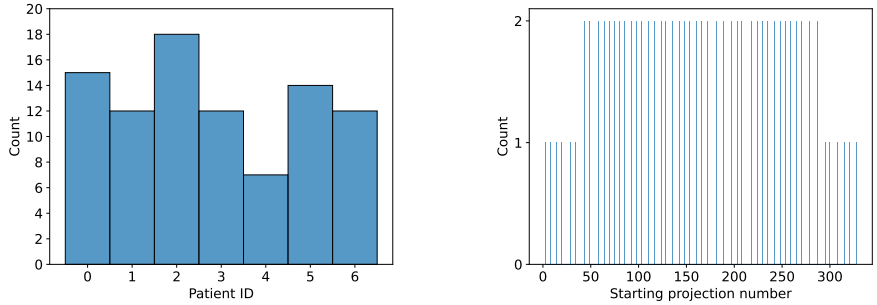

Fig. 3: Left: number of input images in the dataset per patient. Right: number of input images in the dataset per starting projection number in the partial-angle reconstruction.

six patients. The learning rate of the Adam optimizer was set to 0.0001. No learning scheduler was used. We used MSE loss as a cost function for this task. The data of the seventh patient were used for validation.

2.4 Partial-angle reconstruction registration

After modification using the U-Net, the two partial-angle volumes share common structures corresponding to the data within the angular gap. The volumes are normalized and binarized by applying an experimentally estimated threshold. This operation enables the extraction of hard tissue structures like bones and teeth and the removal of soft tissues and possible artifacts that could affect the subsequent registration process. The binarized volumes are forward projected at M sampling positions within the gap. For each position j, we have pair of projections (r_j, r'_j) defined as

$$r_j = P_j \cdot V_I, \quad {r_j}' = P_j \cdot {G_j}' \cdot V_{II}, \quad j = 1, \ldots, M. \tag{3}$$

The volumes V_I and V_{II} are the partial volumes. The matrix P_j is the projection matrix that corresponds to the acquisition position j. The compensation matrix ${G_j}'$ is defined in Equation (2).

The process of retrieving the parameters of ${G_j}'$ is formulated as a minimization problem. Since small translations can be well approximated by small rotations [5] we limit our estimation to the rotation parameters (rX, rY, rZ). As explained in [8] we start the minimization process by extracting hard tissue edges on the forward-projected images. We apply a gradient operator $\boldsymbol{\nabla} = (\nabla_x, \nabla_y)$ and generate corresponding gradient images $(\boldsymbol{\nabla} r_j, \boldsymbol{\nabla} {r_j}')$. We define the cost function $\Phi(\boldsymbol{\nabla} r_j, \boldsymbol{\nabla} r'_j)$ as explained in [8] and for each sampling position j we retrieve the entries of the matrix ${G_j}'$ repeating the 2D-3D registration step until the function Φ reaches a minimum.

The final matrix used to align V_I and V_{II} is the average of the matrices estimated for each position $j = 1, \ldots, M$. To correct the motion-affected projections within the angular gap and to ensure a smooth transition between the first and the second part of the final volume, we linearly interpolate the values from 0 to (rX, rY, rZ) across the angular gap.

3 Data

We tested our method with real patient data acquired using the CBCT device Seethrough Max (See Through s.r.l, Brusaporto, Italy). The acquisitions lasted 14 seconds and covered an angle slightly larger than 180 degrees. The size of the reconstructed field of view is 10×11 centimeters. Since the collected data did not present motion artifacts, we simulated different motion patterns (nodding, tilting, axial rotation, translation) and durations modifying the projection matrices at some angular points. We used the strategy described in [1]. We particularly focused on movements of about $1.5 - 2.0$ seconds duration for which our previously implemented method did not perform reliably. Furthermore, we focused on movements that occurred either at the very beginning or the very end of the acquisition, as these are especially challenging to compensate for. In the next section, we present exemplary results for three cases: (1) nodding and axial rotation with translation (1.5s); (2) combined nodding, tilting, axial rotation, and translation (1.5s); (3) nodding only (2s). A comparison with our previous method [8] is also provided.

4 Results

On the left, Figure 4 shows the results of our motion compensation method for Case 1. Strong motion as a combination of a single rotation (nodding, axial rotation) and a translation was simulated. In both cases artifacts are almost fully corrected or strongly attenuated even in the case that the patient has metal inserts. The same can be observed for more complicated types of motion resulting from a combination of rotations (nodding, tilting, axial rotation) and a translation (Case 2). Also in this case, artifacts could be almost completely corrected or significantly attenuated, as shown in Figure 4 on the right. Finally, we compared our current approach with our previous method. Figure 5 shows that the current method also performs very well in the case of a longer nodding motion of about 2 seconds (Case 3).

5 Discussion and Conclusion

In this paper, we present a robust optimized method to compensate for motion in dental CBCT images. The current method preserves the advantages of the previous method [8]. It is robust and fast, since only forward-projected data are registered and only two partial-angle volumes are reconstructed. Very few

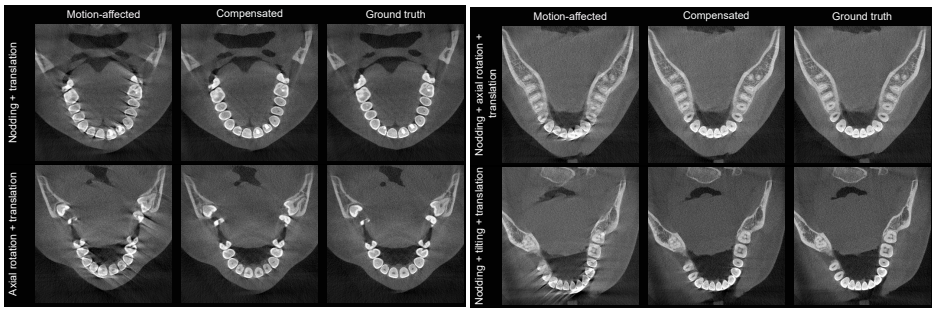

Fig. 4: Left: Case 1. Right: Case 2.

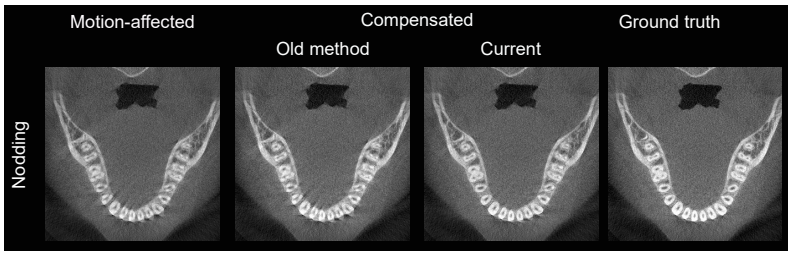

Fig. 5: Case 3.

projections, taken at predefined sampling angles within the angular gap, are necessary for the 2D-3D registration process to recover the motion compensation parameters. The novelty of the current approach relies on a combination of our previous strategy with a deep learning approach. With a suitably trained U-Net, we could modify each partial-angle volume to cover a larger angular range. In this way, we were able to compensate for longer and more complex motion patterns.

Our current results are very promising and suggest the possibility of using the optimized method to compensate for motions longer than two seconds. It is also conceivable to apply the method to compensate for artifacts due to multiple motion patters. Multiple partial-angle volumes could be reconstructed and iteratively aligned using the motion compensation approach. Once motion artifacts are sufficiently compensated, the diagnostic quality of the final fully reconstructed volume may be restored. Finally, additional experiments are necessary involving various patient pathologies and various types of motion to determine whether some data types and artifacts are more challenging to compensate for than others.

Acknowledgments The authors thank Lorenzo Arici, Andrea Delmiglio, Luca Fracassetti, and Ivan Tomba for their valuable contributions to the implementation of the reconstruction library.

Disclosure of Interests. C. Sarti, M. Mikerov, and C. Landi are employed by See Through s.r.l.

References

1. Asraf Ali, A.S.R., Fusiello, A., Landi, C., Sarti, C., Siswadi, A.A.P.: Motion artifacts detection in short-scan dental cbct reconstructions. arXiv preprint arXiv:2304.10154 (2023)
2. Brocal, G.M., Peeters, G.: Conditioned-u-net: Introducing a control mechanism in the u-net for multiple source separations (2019). https://doi.org/10.5281/ZENODO.3527766, https://zenodo.org/record/3527766
3. Feldkamp, L.A., Davis, L.C., Kress, J.W.: Practical cone-beam algorithm. Josa a **1**(6), 612–619 (1984)
4. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2004), https://books.google.fr/books?id=e30hAwAAQBAJ
5. Hernandez, J., Eldib, M., Hegazy, M., Cho, M., Cho, M., Lee, S.: A head-motion estimation algorithm for motion artifact correction in dental ct imaging. Physics in Medicine and Biology **63** (03 2018). https://doi.org/10.1088/1361-6560/aab17e
6. Maur, S., Stsepankou, D., Hesser, J.: CBCT auto-calibration by contour registration. In: Schmidt, T.G., Chen, G.H., Bosmans, H. (eds.) Medical Imaging 2019: Physics of Medical Imaging. vol. 10948, p. 109481N. International Society for Optics and Photonics, SPIE (2019). https://doi.org/10.1117/12.2512181, https://doi.org/10.1117/12.2512181
7. Moratin, J., Berger, M., Rückschloss, T., Metzger, K., Berger, H., Gottsauner, M., Engel, M., Hoffmann, J., Freudlsperger, C., Ristow, O.: Head motion during cone-beam computed tomography: Analysis of frequency and influence on image quality. Imaging Sci Dent **50**(3), 227–236 (2020). https://doi.org/https://doi.org/10.5624/isd.2020.50.3.227, https://pmc.ncbi.nlm.nih.gov/articles/PMC7506095/
8. Sarti, C., Mikerov, M., Landi, C.: Single motion compensation by fast partial volume registration in dental cbct. In: Proceedings of the 18th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (2025)
9. Schulze, R., Drage, N.: Cone-beam computed tomography and its applications in dental and maxillofacial radiology. Clinical Radiology **75**(9), 647–657 (2020). https://doi.org/https://doi.org/10.1016/j.crad.2020.04.006, https://www.sciencedirect.com/science/article/pii/S0009926020301586
10. Sisniega, A., Stayman, J.W., Yorkston, J., Siewerdsen, J., Zbijewski, W.: Motion compensation in extremity cone-beam ct using a penalized image sharpness criterion. Physics in Medicine & Biology **62**(9), 3712 (2017)
11. Spin-Neto, R., Matzen, L.H., Schropp, L.W., Sørensen, T.S., Wenzel, A.: An ex vivo study of automated motion artefact correction and the impact on cone beam ct image quality and interpretability. Dentomaxillofacial Radiology **47**(5), 20180013 (2018). https://doi.org/10.1259/dmfr.20180013
12. Sun, T., Jacobs, R., Pauwels, R., Tijskens, E., Fulton, R., Nuyts, J.: A motion correction approach for oral and maxillofacial cone-beam ct imaging. Physics in Medicine & Biology **66**(12), 125008 (2021). https://doi.org/10.1088/1361-6560/abfa38