Vision-Based Intraoperative Cone-Beam CT Stitching for Non-overlapping Volumes

Bernhard Fuerst[1,2], Javad Fotouhi[1], and Nassir Navab[1,2]

[1] Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD
[2] Computer Aided Medical Procedures, Technische Universität München, Germany

Abstract. Cone-Beam Computed Tomography (CBCT) is one of the primary imaging modalities in radiation therapy, dentistry, and orthopedic interventions. While providing crucial intraoperative imaging, CBCT is bounded by its limited imaging volume, motivating the use of image stitching techniques. Current methods rely on overlapping volumes, leading to an excessive amount of radiation exposure, or on external tracking hardware, which may increase the setup complexity. We attach an optical camera to a CBCT enabled C-arm, and co-register the video and X-ray views. Our novel algorithm recovers the spatial alignment of non-overlapping CBCT volumes based on the observed optical views, as well as the laser projection provided by the X-ray system. First, we estimate the transformation between two volumes by automatic detection and matching of natural surface features during the patient motion. Then, we recover 3D information by reconstructing the projection of the positioning-laser onto an unknown curved surface, which enables the estimation of the unknown scale. We present a full evaluation of the methodology, by comparing vision- and registration-based stitching.

1 Introduction

Cone-Beam Computed Tomography (CBCT) enables intraoperative 3D imaging for various applications, for instance orthopedics [3], dentistry [13] or radiation therapy [4]. Consequently, CBCT is aimed at improving localization, structure identification, visualization, and patient positioning. However, the effectiveness of CBCT in orthopedic surgeries is bounded by its limited field of view, resulting in small volumes. Intraoperative surgical planning and verification could benefit of an extended field of view. Long bone fracture surgeries could be facilitated by 3D absolute measurements and multi-axis alignment in the presence of large volumes, assisting the surgeon's mental alignment.

The value of stitched fluoroscopy images for orthopedic surgery was investigated in [8]. Radio-opaque referencing markers attached to the tool were used to perform the stitching. Trajectory visualization and total length measurement were the most frequent features used by the surgeons in the stitched view. The outcome was overall promising for future development, and the usability was counted as good. Similarly, [5, 10] employed X-ray translucent references positioned under the bone for 2D X-ray mosaicing. In [15, 16], optical features

N. Navab et al. (Eds.): MICCAI 2015, Part I, LNCS 9349, pp. 387–395, 2015.
DOI: 10.1007/978-3-319-24553-9_48

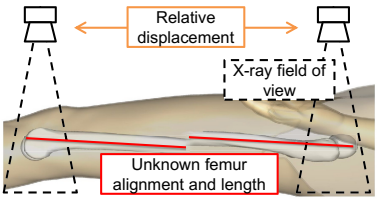

Fig. 1. The 3D misalignment of bones (red lines) may be difficult to quantify using 2D images. CBCT contributes as a valuable tool for interventions in which the 3D alignment is of importance, for instance in acute fracture treatment or joint replacement. Background image courtesy of BodyParts3D, Center for Life Science, Japan.

acquired from an adjacent camera were used to recover the transformation. The aforementioned methods all benefit from external features for 2D mosaicing, thus do not require large overlaps. However, it remains a challenge to generalize these approaches to perform 3D volume stitching, as illustrated in Fig. 1.

A validation study on using 3D rotational X-ray over conventional 2D X-rays was conducted for intra-articular fractures of the foot, wrist, elbow, and shoulder [3]. The outcome reported a reduction of indications for revision surgery. A panoramic CBCT is proposed in [4] by stitching overlapping X-rays acquired from all the views around the interest organ. Reconstruction quality is ensured by introducing a sufficient amount of overlapping regions, which in return increases the X-ray dose. Moreover, the reconstructed volume is vulnerable to artifacts introduced by image stitching. An automatic 3D image stitching technique is proposed in [6]. Under the assumption that the orientational misalignment is negligible, and sub-volumes are only translated, the stitching is performed using phase correlation as a global similarity measure, and normalized cross correlation as the local cost. Sufficient overlaps are required to support this method. To reduce the X-ray exposure, [7, 9] incorporate prior knowledge from statistical shape models to perform a 3D reconstruction.

Previous approaches are either limited to the overlap size or the existing prior shape models. Providing large overlaps will significantly increase the exposure. On the other hand, the bone fractures cause large deformation, hence preoperative and postoperative structures of the region of interest are significantly different, and one cannot benefit from prior scans for alignment. Lastly, incorporating external trackers leads to an increase in surgical complexity and line of sight problem. In this work, we propose a novel stitching approach, using a co-registered X-ray source with an optical camera attached to the C-arm [11,12], and a patient positioning-laser to recover the depth scale. Therefore, the system is mobile, self-contained and independent of the OR, and the workflow remains intact. It could be deployed after a single factory calibration. The alignment transformation of volumes is computed based on the video frames, and prior models are not required. We target cases with large gaps between the volumes and focus our approach on spatial alignment of separated regions of interest.

Image quality will remain intact, and the radiation dose will be linearly proportional to the size of the individual non-overlapping sub-volumes of interest.

2 Materials and Methods

2.1 System Setup and Calibration

The CBCT-enabled motorized C-arm is positioned relative to the patient by utilizing the positioning-lasers, which are built into the image intensifier and C-arm base. To enable the stitching of multiple sub-volumes, the transformation of the patient relative to the C-arm center must be recovered. In contrast to existing techniques we do not require additional hardware setup around the C-arm, but we attach a camera to the C-arm in such manner that it does not obstruct the surgeons access to the patient. By using one mirror, the camera and the X-ray source centers are optically identical. The system setup is outlined in Fig. 2.

Our system is composed of a mobile C-arm, ARCADIS Orbic 3D, from Siemens Medical Solutions and an optical video camera, Manta G-125C, from Allied Vision Technologies. The C-arm and the camera are both connected via ethernet to the computer with custom software to store the CBCT volumes and video. The X-ray and optical images are calibrated in an offline phase [11, 12].

The positioning-laser in the base of the C-arm spans a plane, which intersects with the unknown patient surface, and can be observed as a curve in the camera image. To determine the exact position of the laser relative to the camera, we perform a camera-to-plane calibration. Multiple checkerboard poses (n) are recorded for which the projection of the positioning-laser intersects with the

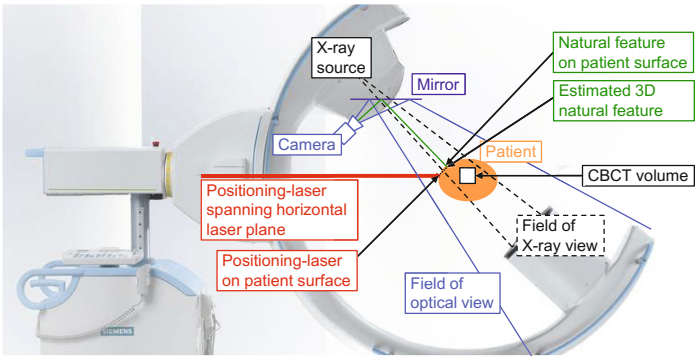

Fig. 2. A mobile C-arm, the positioning-laser (red), and an optical camera (blue) are illustrated. The mirror (purple) aligns the optical camera and X-ray source centers. The patient motion relative to the C-arm is estimated by observing both the positioning-laser and natural features (green) on the patient's surface. The 3D positions of the features are estimated using the depth of the nearest positioning-laser on the patient (black dotted line intersecting green line), of which the depth is based on calibration.

origin of the checkerboard. Once the camera intrinsics are estimated, the camera-centric 3D checkerboard poses are computed. Under the assumption that the 3D homogeneous checkerboard origins, $\mathbf{x}^{(3)} = \{\mathrm{x}_i \,|\, \mathrm{x}_i = [x, y, z, 1]^\top\}_{i=0}^{n}$ (see footnote [1] for notation), lay on the laser plane, the plane coefficients $A = [a, b, c, d]$ are determined by performing RANdom SAmple Consensus (RANSAC) based plane fitting to the observed checkerboard origins, which attempts to satisfy:

$$\arg\min_{A} \sum_{x_j \in \Omega} |Ax_j|, \tag{1}$$

where Ω is subset of checkerboard origins, which are inliers to the plane fitting.

2.2 CBCT Volume and Video Acquisition

To acquire a CBCT volume, the patient is positioned under guidance of the lasers. Then, the motorized C-arm orbits 190° around the center visualized by the laser lines, and automatically acquires a total of 100 2D X-ray images. The reconstruction is performed using the Feldkamp method, which utilizes filtered back-projection, resulting in a cubic volume with a 256 voxels along each axis and an isometric resolution of 0.5 mm. During the re-arrangement of C-arm and patient for the next CBCT acquisition, the positioning-laser is projected at the patient, and each video frame is recorded. For simplicity, we will assume that in the following the C-arm is static, while the patient is moving. However, as only the relative movement of patient to C-arm is recorded, there are no limitations on allowed motions.

2.3 Two-Dimensional Feature Detection and Matching

The transformation describing the relative patient motion observed between two video frames is estimated by detecting and matching a set of natural surface features and the recovery of their scale. For each frame, we automatically detect Speeded Up Robust Features (SURF) as described in [2], which are well suited to track natural shapes and blob-like structures. To match the features in frame k to the features in frame $k + 1$, we find the nearest neighbor by exhaustively comparing the features, and removing weak or ambiguous matches. Outliers are removed by estimating the Fundamental Matrix, $\mathbf{F}_k$, using a least trimmed squares formulation and rejecting up to 50% of the features, resulting in a set of n_k features $\mathbf{f}_k^{(2)} = \{\mathrm{f}_{k,j} \,|\, \mathrm{f}_{k,j} = [x, y, 1]^\top\}_{j=1}^{n_k}$ in frame k (see Fig. 3). To estimate the 3D transformation, the 3D coordinates of this set of features need to be estimated.

2.4 Recovering Three-Dimensional Coordinates

In each frame k, the laser is automatically detected. First the color channel corresponding to the laser's color is thresholded and noise is removed by analyzing connected components. To find the m_k 2D points, $\mathbf{p}_k^{(2)} = \{\mathrm{p}_{k,i} \,|\, \mathrm{p}_{k,i} = [x, y, 1]^\top\}_{i=1}^{m_k}$,

[1] Superscripts $^{(2)}$ and $^{(3)}$ denote 2D and 3D points; $^{(s)}$ denotes points up to a scale.

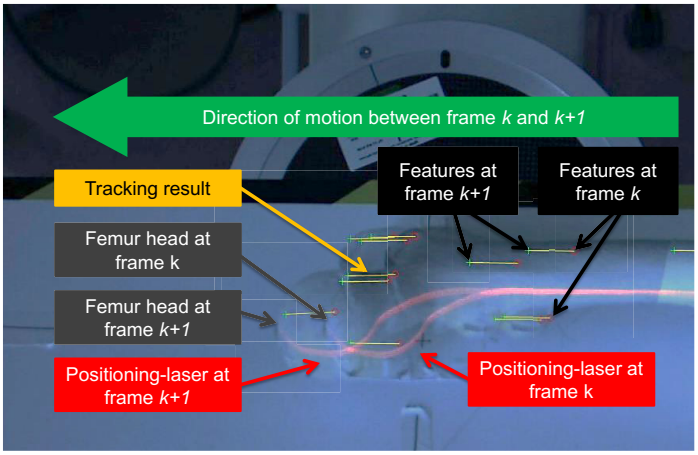

Fig. 3. The figure shows the overlay of two frames to illustrate the feature correspondences to estimate the movement of the patient. From both frames, the positioning-laser (red) and natural surface features are extracted. The tracking results of the matched features in frame k (+) and frame $k+1$ (∘) are illustrated as yellow lines.

which are most likely on the plane, the resulting binary image is thinned [17]. Each 2D laser point $\mathrm{p}_{k,i}^{(2)}$ is projected back to a point up to a scale $\mathrm{p}_{k,i}^{(s)} = [x_{k,i}^{(s)}, y_{k,i}^{(s)}, 1, 1]^\top$ using the Moore-Penrose pseudo-inverse of the camera projection matrix, $\mathbf{P}$:

$$\mathrm{p}_{k,i}^{(3)} = s_{k,i}\mathrm{p}_{k,i}^{(s)} = s_{k,i}\mathbf{P}^{+}\mathrm{p}_{k,i}^{(2)}, \tag{2}$$

where the scale $s_{k,i}$ is recovered by intersecting the point up to a scale $\mathrm{p}_{k,i}^{(s)}$ with the plane:

$$s_{k,i} = \frac{-d}{ax_{k,i}^{(s)} + by_{k,i}^{(s)} + c}. \tag{3}$$

Once the 3D laser points are recovered, the scale for each feature, $\mathrm{f}_{k,j}^{(s)} = s_{k,j}\mathbf{P}^{+}\mathrm{f}_{k,j}^{(2)}$, can be estimated by interpolating the scales of the closest points $\mathrm{p}_{k,i}^{(3)}$.

2.5 Estimating 3D Transformation and CBCT Volume Stitching

After the estimation of the 3D coordinates of the matched features, the transformation for the frames k and $k+1$ is computed by solving the least squares fitting for two sets of 3D points [1], obtaining the transformation matrix T_k. Note that, only features in a small neighborhood of the laser line, < 1 cm, are used. Hence, features on other body parts, e.g. the opposite leg, are discarded. To verify the estimated transformation, the Iterative Closest Point (ICP) algorithm is used to perform a redundancy test using the laser points. In other words, ICP is applied after transforming the laser points $\mathrm{p}_i^{(3)}$ from frame k to the next $k+1$ only for verification. Consequently, for long bones, translation along the laser line is not

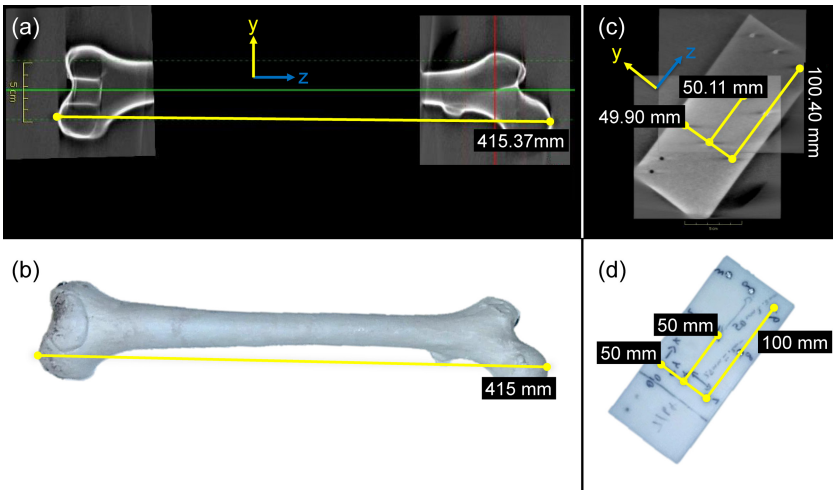

Fig. 4. Absolute distance of the aligned sub-volumes in (a) is measured (415.37 mm), and compared to the real world measurements (415 mm) of the femur phantom in (b). Similarly, a fiducial phantom was scanned and the vision-based stitching (c) compared to the real world object (d). For visualization purposes and volumetric appearance in (a) and (c), multiple parallel slices are averaged.

lost. This results in a transformation $\hat{\mathrm{T}}_k$. If $\hat{\mathrm{T}}_k$ is not nearly identity, the frame $k+1$ is rejected and the frames k and $k+2$ are used to compute $\hat{\mathrm{T}}_k$. To obtain the overall transformation $\mathrm{T}_{\mathrm{CBCT}}$, all transformations $\mathrm{T}_k \in \mathit{\Gamma}$ are accumulated, where $\mathit{\Gamma}$ is the domain of all valid transformations:

$$\mathrm{T}_{\mathrm{CBCT}} = {}^{\mathrm{CBCT}}\mathrm{T}_{\mathrm{camera}} \prod_{\mathrm{T}_k \in \mathit{\Gamma}} \mathrm{T}_k, \tag{4}$$

where ${}^{\mathrm{CBCT}}\mathrm{T}_{\mathrm{camera}}$ is the transformation from camera coordinate system to the CBCT coordinate system obtained during calibration.

3 Experiments and Results

The novel laser-guided stitching method is evaluated in two different, but realistic scenarios. For each phantom, we performed vision-based stitching and evaluated the quality by measuring 3D distances in the stitched volumes and real object. In addition, the stitching quality was compared to intensity-based mosaicing using overlapping CBCT volumes, indicating the accuracy of the overall 3D transformation $\mathrm{T}_{\mathrm{CBCT}}$.

The result of vision-based stitching is illustrated in Fig. 4 (a) on the long bone phantom in the absence of overlaps, and in Fig. 4 (c) on the fiducial phantom with overlaps. The absolute distances are compared to real world measurements which are illustrated in Fig. 4 (b) and (d). Detailed results are reported in

table 1, which shows the differences of measurements of the vision-based stitched CBCT volumes and real objects. The errors are apportioned according to the coordinate frames illustrated in Fig. 4, while the *norm* reflects the overall error. In addition, the *absolute distance error* reports the percentage of error with respect to the absolute distances measured. Average errors are in the range of 0.65 ± 0.28 mm and 0.15 ± 0.11 mm for long bone and fiducial phantom stitching, respectively. Lastly, for overlapping volumes, we have compared the vision- and intensity-based stitching by performing rigid registration using normalized cross correlation as similarity measure. The intensity-based stitching deviated from the vision-based stitching by 0.23 mm, indicating an overall good alignment.

Table 1. Errors are computed by comparing the vision-based stitched CBCT to the real objects. The final row presents the difference to the intensity-based stitching.

Error	X	Y	Z	Norm
Long Bone (Femur) Phantom				
Alignment error (mm)	0.75	0.83	0.37	1.18
Absolute distance error (%)	1.30	1.11	0.10	n/a
Fiducial Phantom				
Alignment error (mm)	0.08	0.10	0.26	0.29
Absolute distance error (%)	0.52	2.00	2.60	n/a
Vision- vs. intensity-based (mm)	0.11	0.08	0.18	0.23

4 Discussion and Conclusion

The proposed technique is an overlap-independent, low dose, and accurate stitching method for CBCT sub-volumes with minimal increase of workflow complexity. We attached an optical camera to a mobile C-arm, and used the positioning-laser to recover the 3D depth scales, and consequently aligned the sub-volumes. As a result of this method, the stitching is performed with low dose radiation, linearly proportional to the size of non-overlapping sub-volumes. We expect this to be applicable to intraoperative planning and validation for long bone fracture or joint replacement interventions, where multi-axis alignment and absolute distances are difficult to visualize and measure from the 2D X-ray views.

Our approach does not limit the working space, nor does it require any additional hardware besides a simple camera. The C-arm remains mobile and independent of the OR. One requirement is that the C-arm does not move during the CBCT acquisition, but we believe that the use of external markers could solve this problem and may yield a higher accuracy. However, in our scenario we intentionally did not rely on markers, as they would increase complexity and alter the surgical workflow. Our approach uses frame-to-frame tracking, which can cause drift. In fact, the ICP verification helps us to detect such drifts as it is based on points which were not used for motion estimation. Therefore, if the estimated motion from ICP increases over time, we can detect the drift and use

ICP to correct if necessary. Alternatively, the transformations could be refined using bundle adjustments [14]. Further studies on the effectiveness during interventions are underway. Also, the reconstruction of the patient surface during the CBCT acquisition may assist during the tracking of the patient motion.

References

1. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of two 3-D point sets. IEEE Transactions on PAMI (5), 698–700 (1987)
2. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded Up Robust Features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part I. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
3. Carelsen, B., Haverlag, R., Ubbink, D., Luitse, J., Goslings, J.: Does intraoperative fluoroscopic 3D imaging provide extra information for fracture surgery? Archives of Orthopaedic and Trauma Surgery 128(12), 1419–1424 (2008)
4. Chang, J., Zhou, L., Wang, S., Clifford Chao, K.S.: Panoramic cone beam computed tomography. Medical Physics 39(5) (2012)
5. Chen, C., Kojcev, R., Haschtmann, D., Fekete, T., Nolte, L., Zheng, G.: Ruler Based Automatic C-Arm Image Stitching Without Overlapping Constraint. Journal of Digital Imaging, 1–7 (2015)
6. Emmenlauer, M., Ronneberger, O., Ponti, A., Schwarb, P., Griffa, A., Filippi, A., Nitschke, R., Driever, W., Burkhardt, H.: XuvTools: free, fast and reliable stitching of large 3D datasets. Journal of Microscopy 233(1), 42–60 (2009)
7. Fleute, M., Lavallée, S.: Nonrigid 3-D/2-D registration of images using statistical models. In: Taylor, C., Colchester, A. (eds.) MICCAI 1999. LNCS, vol. 1679, pp. 138–147. Springer, Heidelberg (1999)
8. Kraus, M.: von dem Berge, S., Schoell, H., Krischak, G., Gebhard, F.: Integration of fluoroscopy-based guidance in orthopaedic trauma surgery - a prospective cohort study. Injury 44(11), 1486–1492 (2013)
9. Lamecker, H., Wenckebach, T., Hege, H.C.: Atlas-based 3D-Shape Reconstruction from X-Ray Images. In: 18th International Conference on Pattern Recognition, ICPR 2006, vol. 1, pp. 371–374 (2006)
10. Messmer, P., Matthews, F., Wullschleger, C., Hgli, R., Regazzoni, P., Jacob, A.: Image Fusion for Intraoperative Control of Axis in Long Bone Fracture Treatment. European Journal of Trauma 32(6), 555–561 (2006)
11. Navab, N., Heining, S.M., Traub, J.: Camera Augmented Mobile C-Arm (CAMC): Calibration, Accuracy Study, and Clinical Applications. IEEE Transactions on Medical Imaging 29(7), 1412–1423 (2010)
12. Nicolau, S., Lee, P., Wu, H., Huang, M., Lukang, R., Soler, L., Marescaux, J.: Fusion of C-arm X-ray image on video view to reduce radiation exposure and improve orthopedic surgery planning: first in-vivo evaluation. Proceedings of Computer Assisted Radiology and Surgery 6, 115–116 (2011)
13. Pauwels, R., Araki, K., Siewerdsen, J., Thongvigitmanee, S.S.: Technical aspects of dental CBCT: state of the art. Dentomaxillofacial Radiology 44(1) (2014)
14. Triggs, B., McLauchlan, P.F., Hartley, R.I., Fitzgibbon, A.W.: Bundle adjustment – a modern synthesis. In: Triggs, B., Zisserman, A., Szeliski, R. (eds.) ICCV-WS 1999. LNCS, vol. 1883, pp. 298–372. Springer, Heidelberg (2000)

15. Wang, L., Traub, J., Heining, S.M., Benhimane, S., Euler, E., Graumann, R., Navab, N.: Long bone X-Ray image stitching using camera augmented mobile C-arm. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 578–586. Springer, Heidelberg (2008)
16. Wang, L., Traub, J., Weidert, S., Heining, S.M., Euler, E., Navab, N.: Parallax-free intra-operative X-ray image stitching. Medical Image Analysis 14(5)
17. Zhang, T., Suen, C.Y.: A fast parallel algorithm for thinning digital patterns. Communications of the ACM 27(3), 236–239 (1984)