

Model-based 3D Tracking for Augmented Orthopedic Surgery

Albert Murienne, Boris Labbé, Laurent Launay

▶ **To cite this version:**

Albert Murienne, Boris Labbé, Laurent Launay. Model-based 3D Tracking for Augmented Orthopedic Surgery. EPiC Series in Health Sciences, 2020. hal-03022939

HAL Id: hal-03022939
https://hal.science/hal-03022939

Submitted on 25 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Model-based 3D Tracking for Augmented Orthopedic Surgery

Albert Murienne[1], Boris Labbé[1] and Laurent Launay[1]
[1] IRT b<>com, Rennes, France.
albert.murienne@b-com.com

Abstract

Current surgical navigation systems offer sub-millimetric real-time localization, however they are expensive, require the use of invasive markers attached to the patient, and often add extra operation time. In this paper we propose an affordable markerless navigation approach, based on mid end depth sensors, as an alternative to answer medical applications needs in terms of accuracy and robustness. An algorithm called Fast Volumetric Reconstruction (FaVoR) implements a compute-efficient approach for real time 3D model registration based tracking, allowing computed 3D poses to be used for video scene augmentation. After early testing with a first proof-of-concept implementation, a preliminary accuracy evaluation was performed using a dynamic test bench, achieving an average 2mm registration error during tracking.

1 Introduction

Augmented Reality (AR) provides an intuitive way to show clinical information to guide the surgeon during his procedure. For the additional information to be consistent with the scene, it has to be accurately positioned, and should be updated in real-time with the minimal possible latency [1]. Therefore, the key point is to have efficient object localization and tracking algorithm targeting, in the case of orthopedic surgery, the surgeon tools or rigid organs like bones.

For this study, a prototype application has been built to illustrate the use case of scene augmentation to assist orthopedic knee cutting guide positioning. The performance of our algorithm is illustrated by displaying different type of hologram 3D shapes overlaid on the video stream: the planned knee alignment (Figure 1), the guide's drilling axis and the guide's cutting plane. The prototype relies on depth sensors, which have already shown good results on localization tasks [2]. The application was first evaluated visually, then a test environment was designed to validate the approach and compare the results with traditional marker based navigation systems [3].

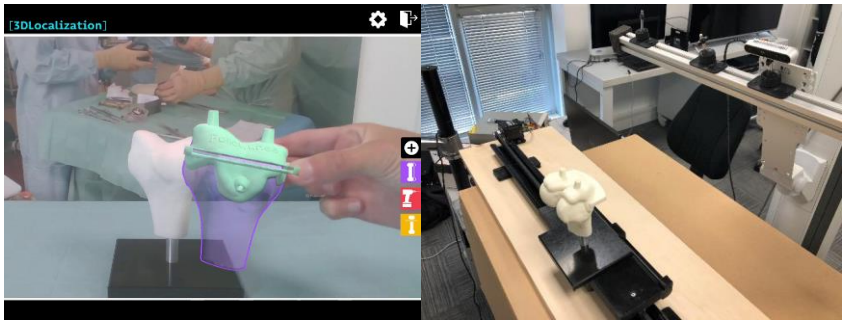


Figure 1: The prototype tracking application showing planned femur and PSI hologram, and the test bench.

2 Global Architecture

2.1 Software Architecture

FaVoR's first implementation is based on a complete geometrical paradigm, inspired by previous state-of-the-art realtime ICP tracking approaches like [4], [5] & [6], but using a 'plane-to-plane' Generalized ICP variant [7].

The key performance indicator of our approach is the frame rate which can reach, which directly impacts the convergence speed and success of the ICP: the smaller the inter-frame displacement is, the faster the ICP will converge towards the right minima. The algorithm is implemented in C++, and thanks to heavy use of vectorization and parallelization code optimizations, it can reach up to 60 frames-per-second with an input camera depth resolution of 640x480 (~300K pixels). The tracked objects are typically hand manipulated. So considering an average hand speed of 1m/s during an action, at 60fps, the inter-frame displacement of the handheld object is about 16mm, which gives us a good idea of the search area required to support this speed. This search area is configurable using a dedicated parameter, the search radius SR, which influence is further described in this paper.

Another problem is that raw computed poses can suffer from detection noise coming from the depth sensor, so exponential smoothing was applied for more realistic and less shaky results [8], as well as limited vacillation when the object remains steady.

2.2 Hardware Architecture

A prototype application was developed using an Intel RealSense D415 depth camera, running on a Microsoft Surface Pro 5 with core i7 processor. The algorithm is CPU only (no GPU processing), therefore it does not require any additional external processing server. The camera is attached to the back of the tablet, allowing the application to run in "video see-through" augmentation mode.

The goal of our test bench is to have reproducible and configurable movement (amplitude and speed), in order to ensure that our approach is robust to fast movement. The current version of the bench can reach up to 0.15m/s speed translation, with the help of a stepping motor controlled cart moving along a rail, as seen in Figure 1.

3 Results

A first study has been carried out, targeting registration residual error, during static or dynamic translations phases, at different speeds and different algorithmic configurations, and tracking a prototype femoral PSI (~8cm span) at 40cm distance. The influence of FaVoR's search radius (SR) parameter was evaluated, using an Intel Realsense D415 and an Occipital Structure Core running in 640x480@60fps depth mode. During dynamic runs, a 20cm back and forth translation was performed on the bench, with configurable speed. The results are illustrated in Figure 2, where the RMSE registration error is reported for each test configuration. Blank cells indicate tracking was lost during test.

Intel Realsense D415	**0 cm/s**	**5 cm/s**	**10 cm/s**	**15 cm/s**
SR=**3mm**	1.68mm ±0.02mm	1.71mm ±0.03mm	-	-
SR=**5mm**	2.12mm ±0.03mm	2.16mm ±0.06mm	2.23mm ±0.14mm	-
SR=**8mm**	2.83mm ±0.08mm	3.02mm ±0.16mm	3.14mm ±0.18mm	3.16mm ±0.19mm
Occipital Structure Core	**0 cm/s**	**5 cm/s**	**10 cm/s**	**15 cm/s**
SR=**3mm**	1.59mm ±0.01mm	1.60mm ±0.03mm	-	-
SR=**5mm**	1.92mm ±0.04mm	1.87mm ±0.07mm	1.93mm ±0.085mm	2.05mm ±0.15mm
SR=**8mm**	2.34mm ±0.09mm	2.46mm ±0.15mm	2.55mm ±0.165mm	2.75mm ±0.295mm

Figure 2: Influence of FaVoR search radius on tracking robustness and accuracy.

These results show that search radius has direct influence on the tradeoff between robustness and accuracy: high search radius means less accuracy due to more potential outliers included in the search path, small search radius means less robustness to fast inter-frame movement.. The direct correlation between search radius and speed is not a straight forward computation: 0.15mm/s should be supported with SR 3mm, but that would be neglecting the inherent sensors noise and lag. When comparing the two cameras, the Structure Core is overall slightly more accurate (up to 0.5mm at highest speed) and more robust (SR 5mm@15cm/s didn't lose tracking). The Structure Core has a global shutter, whereas the D415 has a rolling shutter, so it's more sensitive to moving scenes. The overall mean accuracy level is around 2-3mm. Further work should be done to assess the orientation error by collecting ground truth data.

4 Conclusion

In this paper, a fast volumetric registration technique embedded in a depth sensor equipped tablet has been presented and evaluated on a dynamic test bench. The system can reach 60fps on the tablet CPU and a tracking registration accuracy of 2mm. Robustness improvement is still needed to support higher speed movements. Future work will include ground truth based accuracy qualification and further code optimizations to address cameras 90fps mode. Machine learning based pose inference [9] could also be investigated to initialize registration.

Acknowledgment

This work benefited from State aid managed by the National Research Agency under the future investment program bearing the reference ANR-17-RHUS-0005 (FollowKnee project).

References

[1] N Navab, J Traub, T Sielhorst, M Feuerstein, C Bichlmeier. IEEE computer graphics and applications, 27. 10-4. 10.1109/MCG.2007.117. (2007). Action- and Workflow-Driven Augmented Reality for Computer-Aided Medical Procedures.

[2] J Ogor, G Dardenne, S Sta, J Bert, H Letissier, E Stindel, C Hamitouche. CAOS 19th Annual Meeting of the International Society for Computer Assisted Orthopaedic Surgery, vol 3, p.289-292. (2019). 3D Pose Estimation with Depth Camera for Markerless Computer Assisted Orthopaedic Surgery.

[3] R Elfring, M de la Fuente, K Radermacher. Computer Aided Surgery, Volume 15, Issue 1-3, p.1-12. (2010). Assessment of optical localizer accuracy for computer aided surgery systems.

[4] M Belshaw, M Greenspan. IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops 2009, p.1449-1456. (2009). A high speed iterative closest point tracker on an FPGA platform.

[5] O Korkalo, S Kahn. Journal of Virtual Reality and Broadcasting. (2016). Real-time depth camera tracking with CAD models and ICP.

[6] SC Lee, B Fuerst, K Tateno, A Johnson, J Fotouhi, G Osgood, F Tombari, N Navab. Healthcare technology letters 4 (5), p.168-173. (2017). Multi-modal imaging, model-based tracking, and mixed reality visualisation for orthopaedic surgery.

[7] A Segal, D Hähnel, S Thrun. (2009). Proceedings of Robotics: Science and Systems. 10.15607/RSS.2009.V.021. Generalized-ICP.

[8] JJ LaViola. Eurographics Workshop on Virtual Environments, p.199-206. (2003). Double Exponential Smoothing: An Alternative to Kalman Filter-Based Predictive Tracking.

[9] S Peng, Y Liu, Q Huang, X Zhou, H Bao. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2018). Pvnet: Pixel-wise voting network for 6dof pose estimation.