Noname manuscript No.
(will be inserted by the editor)

Tracking Better, Tracking Longer: Automatic Keyframe Selection in Model-based Laparoscopic Augmented Reality

Kilian Chandelon[1,2] · **Adrien Bartoli**[3,1,2]

Abstract *Purpose.* We present a novel automatic system for markerless real-time augmented reality. Our system uses a dynamic keyframe database, which is required to track previously unseen or appearance-changing anatomical structures. *Methods.* It works with an offline stage which constructs the initial keyframe database and an online stage which dynamically updates the database with new keyframes selected from the video stream. We propose keyframe selection criteria ensuring tracking stability and a database management scheme ensuring real-time performance. *Results.* Experimental results show that our system outperforms the baseline system with a static keyframe database, increasing the number of tracked frames, without requiring surgeon input. *Conclusion.* Our software-based system copes with new viewpoints and appearance changes. It improves surgical organ tracking performance.

1 Introduction

A major challenge in laparoscopy is to accurately localise and visualise sub-surface anatomical structures, which can be alleviated by Computer-Assisted Surgery (CAS). A key tool in CAS is Augmented Reality (AR), which consists in displaying visual information such as the tumours available from preoperative data –MRI or CT– directly in the intraoperative view. Implementing AR efficiently requires markerless real-time organ tracking, which is yet unresolved in the general case and forms our main focus. Specifically, we endeavour to improve the state-of-the-art model-based tracking system already in use in several CAS-AR tools such as [1]. This system works by computing keypoint correspondences between the current frame and a set of reference frames called keyframes, then computing camera pose from the correspondences. Because

[1] EnCoV, Institut Pascal, UMR6602 CNRS, UCA, Clermont-Ferrand University Hospital
[2] SurgAR, Surgical Augmented Reality, Clermont-Ferrand, France
[3] Department of Clinical Research and Innovation, Clermont-Ferrand University Hospital

the keyframes are registered to preoperative data, chaining the transformations then allows one to transfer information from the preoperative data to the current frame, and realise AR. The system uses a database of keyframes constructed at the beginning of surgery. A major limitation is that the keyframe database is static, whereas new parts of the organ become visible and the organ appearance changes during surgery. This hinders the performance of organ tracking. Our objective is to track the organ more accurately and on a longer timeframe through the surgery. We propose a dynamical keyframe database system, where the keyframes are automatically updated as and when needed. This is a challenging problem because the selection of keyframes is multi-criteria and critical, for adding the wrong keyframe causes tracking drift. The size of the database must be scalable, with the ability to forget keyframes, in order to maintain the real-time tracking performance.

2 Materials and Methods

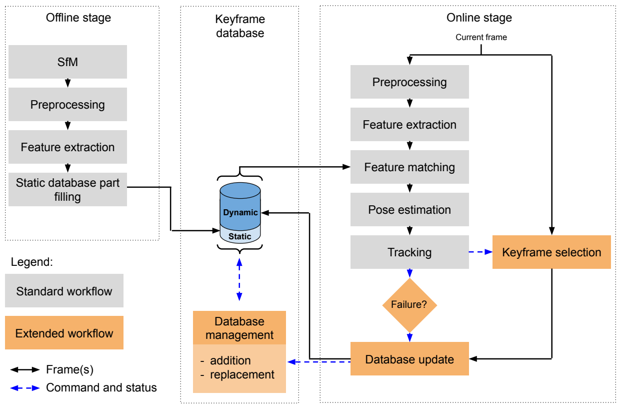

Fig. 1 Existing system with a static keyframe database and its proposed extension with a dynamic keyframe database.

***StaticDB*: baseline static keyframe database system.** Our system is designed very closely to [1], which has an offline and an online stages, as shown in figure 1. The *offline stage* constructs the static keyframe database. It assumes the organ remains rigid and uses Structure-from-Motion (SfM) to compute the keyframe poses and Iterative Closest Point (ICP) to compute the deformable transformation to the preoperative data. Keypoints extracted using POPSIFT [2], a GPU implementation of SIFT, and the descriptors stored in the database. The *online stage* processes the current frame of the video stream in real-time. Keypoints are extracted with POPSIFT and matched to the database using Brute Force Matching (BFM). Finally, camera pose is computed with RANSAC and AR displayed.

***DynamicDB*: proposed dynamic keyframe database system.** We extend *StaticDB* with a dynamic database part using two new steps: keyframe selection and database management. The static part ensures that the overall system does not drift by completely losing long-term focus on the organ. The *keyframe selection step* selects new keyframes from the video stream. It tests the current frame with five criteria. The first two criteria determine if the database lacks knowledge of the surgical field covered by the frame. C1 is passed if the number of keypoint matches to the keyframe database is lower than 50 and C2 is passed if the number of RANSAC pose estimation inliers is lower than 8. The last three criteria determine if the frame is good enough in terms of pose and focus. The pose quality involves two criteria. C3 is passed if the pose estimation root mean square reprojection residuals is lower than 0.4 % of the image diagonal and C4 is passed if camera jerk in normalised SfM units is lower than 1 s^{-3} (high jerk, hence high frequency, if typical of pose instability). The focus quality involves C5, which is passed if the frame Laplacian variance is lower than 2.9 % of the maximum intensity. Lastly, a frame is added to the database if tracking failed for three consecutive frames, which reduces false selections. All the thresholds were set empirically. The *database management step* updates the dynamic database part with the selected keyframes so as to maintain real-time tracking performance. Obviously, the more keyframes, the slower the processing. We thus determine a maximal size N of the database a priori. This size depends on many factors, including the system implementation and the hardware capacity; we determine it experimentally. The database management step adds the selected frame and drops out a past keyframe if the database size exceeds N. The drop out criterion takes into account the amount of time which passed since a keyframe was last used, allowing the replacement of the no longer relevant keyframes.

3 Experimental Results and Discussion

We have compared *DynamicDB* to our re-implementation of *StaticDB* on seven laparoscopic videos, showing the uterus with rigid and non-rigid movements with instruments generating smoke and bleeding. Our main evaluation criterion is the number of frames for which the uterus is tracked. We ran the tests on a PC with Linux, CPU Intel i9-10900K and GPU Nvidia RTX 2080Ti. We found the maximum size of the dynamic database to be $N = 30$ for a framerate greater than 25 fps. The experimental results are shown in Table 1. A general increase in number of tracked frames is observed with *DynamicDB*, with an average improvement margin over *StaticDB* of 11.19 %. We have observed a stronger improvement when the organ deforms and a larger area is explored. As these experiments are retrospective, they do not take the user reaction to the tracking failure into account, which could lead to an even stronger increase.

Figure 2 gives details on video 3, where the laparoscope and the uterus are moved to extreme positions. *StaticDB* failed after 45 seconds while *DynamicDB* succeeded for all frames where the organ is sufficiently rigid. Both

Table 1 Comparison statistics in number of tracked frames.

Video index	Number of frames	Frames tracked with *StaticDB* (%)	Frames tracked with *DynamicDB* (%)	Difference (%)
1	1892	55.57	75.15	+19.58
2	1554	95.75	95.91	+0.19
3	2207	32.02	73.17	+41.15
4	1475	80.24	80.30	+0.06
5	5484	73.09	77.94	+4.84
6	514	85.78	96.09	+10.31
7	3250	96.12	98.34	+2.22

methods are equivalent in terms of framerate, with averages of 47 fps and 46 fps respectively, well above the 25 fps of the video stream.

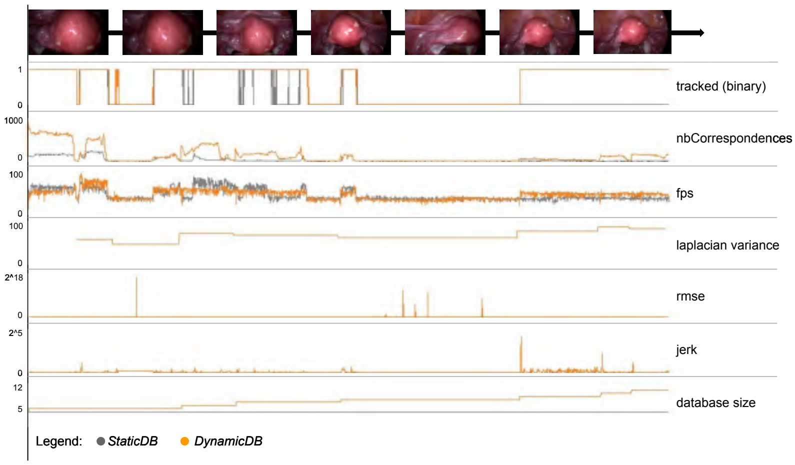

Fig. 2 Detailed comparison between StaticDB and DynamicDB for laparoscopic video 3.

4 Conclusion

Dynamical model update improves organ tracking. We plan to improve our system with additional selection criteria and a representation of the spatial coverage of keyframes, and to run in vivo laparoscopic tests.

References

1. Collins T, Pizarro D, Gasparini S et al. Augmented Reality Guided Laparoscopic Surgery of the Uterus. IEEE Transactions on Medical Imaging, 40(1):371-380, 2021.
2. Griwodz C, Calvet L, Halvorsen P. Popsift: a faithful SIFT implementation for real-time applications. ACM Multimedia Systems Conference, 2018.