CIS Seminar Talk – Critical Review

Group 13, Yihao Liu

Paper: E. Azimi, L. Qian, N. Navab, and P. Kazanzides. Alignment of the Virtual Scene to the Tracking Space of a Mixed Reality Head-Mounted Display. Retireved from: https://arxiv.org/pdf/1703.05834.pdf. 2019

Paper relevance

In our project, one important problem is the registration between the rendered graphics and real objects. More specifically, we want to display the implant and skull model in the virtual space accurately so that it aligns to the actual implant. In our own approach, we have the transformation mapping in Figure 1a. In this approach, the AR tag (bar code) is the bridge between the virtual space and the real space. This is achieved by registration between the AR tag and the tracker marker, and the AR tag can be recognized and located by Vuforia tool kit. On the other hand, the paper proposed a blackbox approach, such that a 3D to 3D projection is directly calculated. In this case, the transformation mapping becomes the illustration in Figure 1b, where the virtual space and the real space can be directed linked. Internally, the relationship will be the projections matrices.

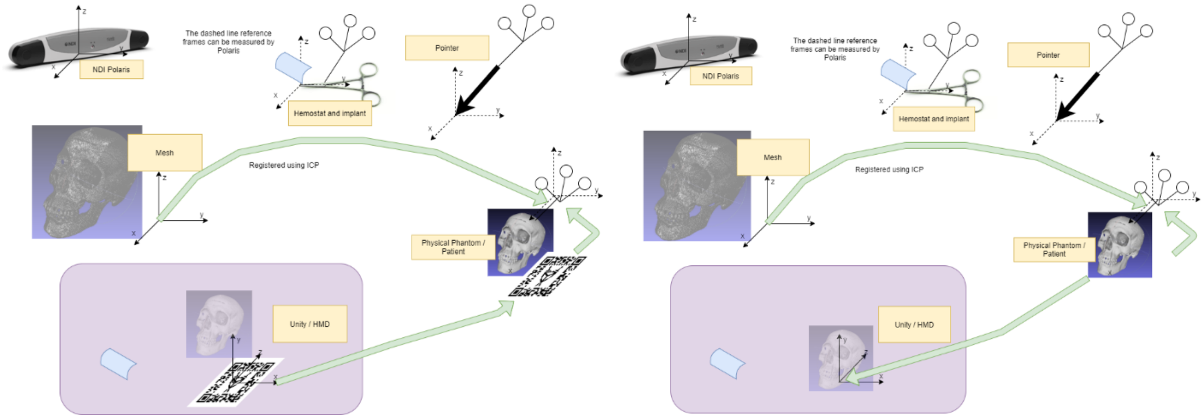


Figure 1. Comparison between CIS project approach and the paper proposed approach, integrated in the CIS project

Paper background

The authors claimed that the transformation chain to the eye of HMD's end user is incomplete. This is because that users may not have direct access to the projection matrix. Even though some flatforms provide this projection, it is still limited to a mapping from the display to some sort of tracking system coordinates, but not from display to eye.

The authors also claim that there has not been systemic work for the alignment in a dynamic setting. Although Vuforia provides image tracking, the authors believe that it does not entirely solve the incomplete chain between rendered object and user's eye. This motivates the authors to propose the calibration work that is presented in the paper.

Paper contributions

The paper proposed an end-to-end solution for the alignment between the virtual 3D object to end user's eye. The corresponding calibration method was a faster and easier multipoint alignment. The experimental procedure was conducted on different setups and tracking systems. In addition, different geometrical models were also used. These configurations are listed here:

- Different platforms: 1. HoloLens, or 2. Moverio BT-300
- Different setups: 1. inside-out: head-anchored tracking system, or 2. outside-in: world-anchored
- Different tracking systems: 1. inside-out: HoloLensARToolKit marker tracking, or BT-300 front-facing RGB camera; 2. outside-in: FusionTrack 500 optical tracker
- Geometrical models: 1. perspective, 2. affine, or 3. isometric

A novel evaluation method – Double-cube match is also proposed in this paper.

Theory

The paper solves the problem of 3D to 3D calibration. The basic equation for this problem is:

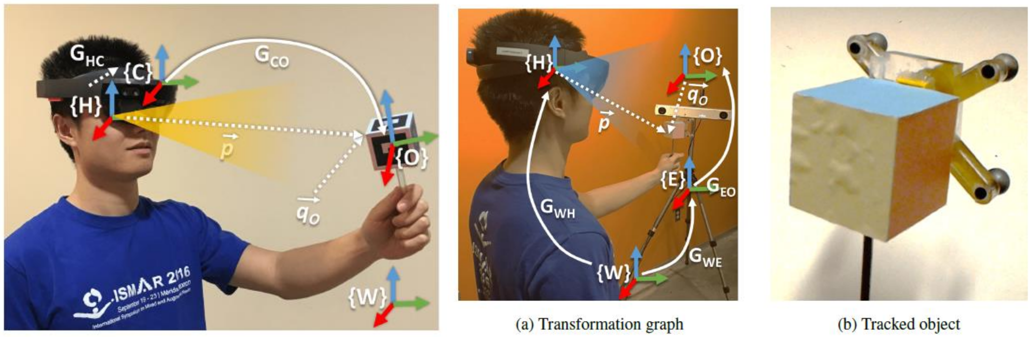
$$p_i = T(q_i)$$

The paper proposes to use three different transformation models: affine with 12 degrees of freedom, isometric with 6 degrees of freedom, and prospective with 15 degrees of freedom. Solving for the transformation then becomes an optimization problem that can be solved by Direct Linear Transformation (DLT) algorithm etc. The problem is defined as:

$$\min_{T} E_{reproj}, \text{ where } E_{reproj} = \frac{\sum_{i=1}^{n} \sqrt{(p_i - T(q_i))^2}}{n}$$

Experiment

The calibration system has two different tracking configurations: head-anchored tracking, and world-anchored tracking. The transformation maps are illustrated in the following figures.



(a) Transformation graph

(b) Tracked object

Figure 2a. The transformation map for the head-anchored tracking. 2b. The transformation map for the world-anchored tracking. Figure adapted from Azimi et al. (2019)

The coordinate systems of the head-anchored system are {C} tracker, {O} object, and {H} HMD, respectively. To avoid ambiguity, the world-anchored tracking coordinate systems used a different set of symbols: {E} tracker, {O} object, {H} HMD, and {W} world, respectively.

During the calibration, the user is asked to align the calibration cube to the HMD prompt, so the coordinates of the cube corners are captured to calculate the transformation. The alignment can be single point alignment or multipoint alignment. For multipoint alignment, 5 points (corners) on the calibration are captured in a single align, as can be seen in the following image.

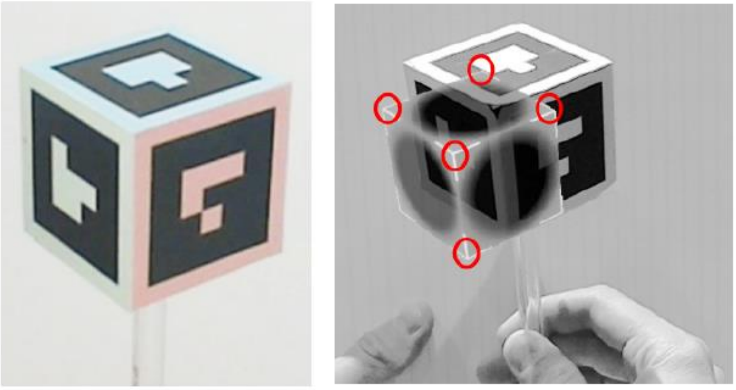


Figure 3. Multipoint alignment method. Figure adapted from Azimi et al. (2019)

The performance of the calibration is evaluated using two different method: calibrate-and-test method, and the innovative double-cube-match method. In the calibrate-and-test method, the user is asked to capture additional samples to calculate reprojection error. On the other hand, in the double-cube-match, the user uses a second cube, and the HMD display a virtual cube in the virtual space. This virtual cube has a predefined offset from the first cube. The user moves the second cube to the virtual cube, and the error can be calculated by the difference between the predetermined pose offset and the observed pose. In this case, rotation error is also evaluated. The experiment workflow is depicted in Figure 4.

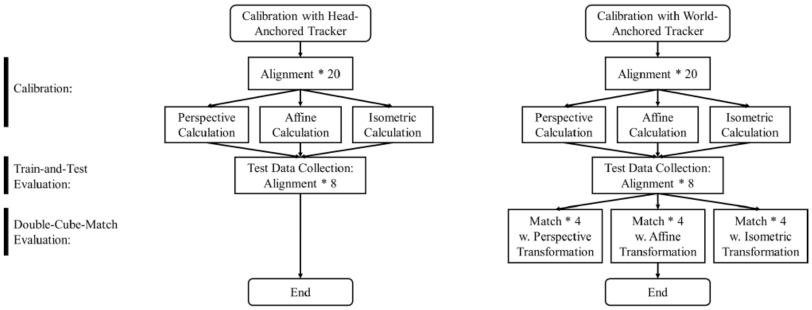


Figure 4. The workflow of the calibration with head-anchored tracker and world-anchored tracker. Figure adapted from Azimi et al. (2019)

Results and discussion

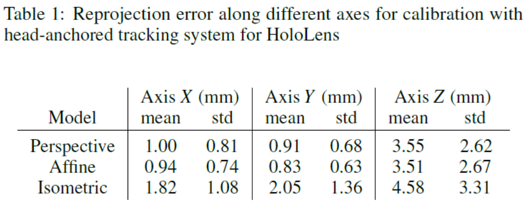


Table 1: Reprojection error along different axes for calibration with head-anchored tracking system for HoloLens

Model	Axis X (mm) mean	Axis X (mm) std	Axis Y (mm) mean	Axis Y (mm) std	Axis Z (mm) mean	Axis Z (mm) std
Perspective	1.00	0.81	0.91	0.68	3.55	2.62
Affine	0.94	0.74	0.83	0.63	3.51	2.67
Isometric	1.82	1.08	2.05	1.36	4.58	3.31

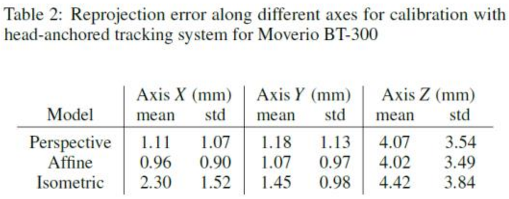
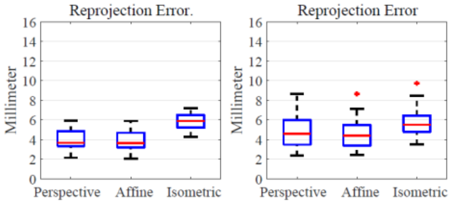


Table 2: Reprojection error along different axes for calibration with head-anchored tracking system for Moverio BT-300

Model	Axis X (mm) mean	Axis X (mm) std	Axis Y (mm) mean	Axis Y (mm) std	Axis Z (mm) mean	Axis Z (mm) std
Perspective	1.11	1.07	1.18	1.13	4.07	3.54
Affine	0.96	0.90	1.07	0.97	4.02	3.49
Isometric	2.30	1.52	1.45	0.98	4.42	3.84

Table 3: Reprojection error along different axes for calibration with world-anchored tracking system for HoloLens

Model	Axis X (mm)		Axis Y (mm)		Axis Z (mm)	
	mean	std	mean	std	mean	std
Perspective	2.47	2.04	3.01	2.49	3.20	3.01
Affine	2.44	1.98	2.98	2.52	3.21	3.01
Isometric	3.64	2.75	6.14	3.88	3.43	2.93



(a) HoloLens

(b) Moverio BT-300

Fig. 10: Evaluation result of Calibrate-and-Test for the calibration with head-anchored tracker with two different HMDs, based on different geometrical models.

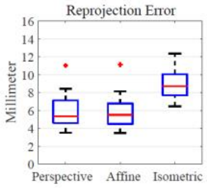
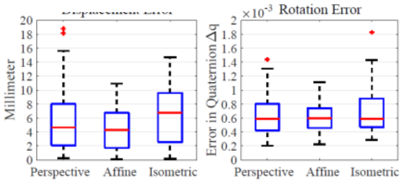


Fig. 11: Evaluation result of Calibrate-and-Test for the calibration with world-anchored tracker using the HoloLens, based on different geometrical models.



(a) Displacement error $E_{\Delta P}$

(b) Rotational error $E_{|\Delta q|}$

Fig. 12: Evaluation result using Double-Cube-Match method, with world-anchored tracker and HoloLens

The worst accuracy is observed to be around 4 mm (in reprojection error). The authors reported in single corner calibration, perspective and affine models are better than isometric model, while in multipoint calibration, isometric model performed better. This is because the different calibration methods retain different geometric relation, so the observations are expected. In Double-cube-match evaluation, affine transformation works better, because in the experimented configuration, specifically the world-anchored configuration, the transformation between HoloLens and the Atracsys tracker is affine. Depth axis has the largest error in head-anchored tracking. This observation is also expected because the manual alignment is hard to align in the depth axis. However, the user is able to move around and make alignments in the case of world-anchored tracking. Thus, the consistent errors in all axis in Table 3 is also expected.

Future work

The authors proposed to integrate position tracking system to HMDs using sensor fusion, so that it can solve the limited line of sight problem in outside-in tracking. The ambiguity of the corners of the calibration cube can pose a problem, so the authors also proposed to create asymmetric calibration rigs. In terms of the whole system, the authors think a user study is necessary in order to investigate the effect of fatigue on the performance, and to investigate user-friendliness.

Critics

This paper proposed a very innovative idea of calibrating HMD system end-to-end. Specifically, the user can expect to compute the transformation between the virtual object and the user’s eye direction, without considering the internal features of HMD. This idea can provide some advantages. The most straightforward advantage is that the calibration method is independent to platform. Regardless of the internal features of HMD, the calibration process should be able to work universally as the problem is simplified to a transformation calculation problem, which can be solved by optimization algorithms like DLT. However, this also pose a problem: the calibration is vulnerable to relative movement between the user’s eye and the HMD. If the HMD is not tightly fixed on user’s head, it can slip or move. As a result, the calibration will no longer be valid.

There are some other pros and cons regarding to this paper. The authors provided an innovative evaluation method, Double-cube Match, which can provide rotation error, in addition to the translational error (as specified in the paper, reprojection error). With this calibration method, the transformation chain is greatly simplified. Nonetheless, the calibration can prone to subjective errors, because during the calibration, the alignment is made manually. This error cannot be well-defined and analyzed because it can change for different testing individuals. On the other hand, if the HMD and real world is bridged by a barcode, the error can be defined so that error propagation is possible.

In terms of the writing structure, I was not very clear about the description to the effect of the 3D to 2D projection. From the paper’s description, I understand that the calibration is a process of 3D to 3D transformation. However, the depth cue is made by two 2D displays on HMD, as the paper mentions. I did not fully understand what role would the transformation between the rendered 3D object and the two 2D displays play in the calibration process.

In conclusion, the paper is a great work for its attempt to the simplification transformation process and the innovation of verification method, although there are some disadvantages, they can be addressed by proper future works such as a method of fixing the HMD tightly to user’s head, or maybe some calibration process that can cancel out the movement of the HMD.

Reference

E. Azimi, L. Qian, N. Navab, and P. Kazanzides. Alignment of the Virtual Scene to the Tracking Space of a Mixed Reality Head-Mounted Display. Retireved from: https://arxiv.org/pdf/1703.05834.pdf. 2019