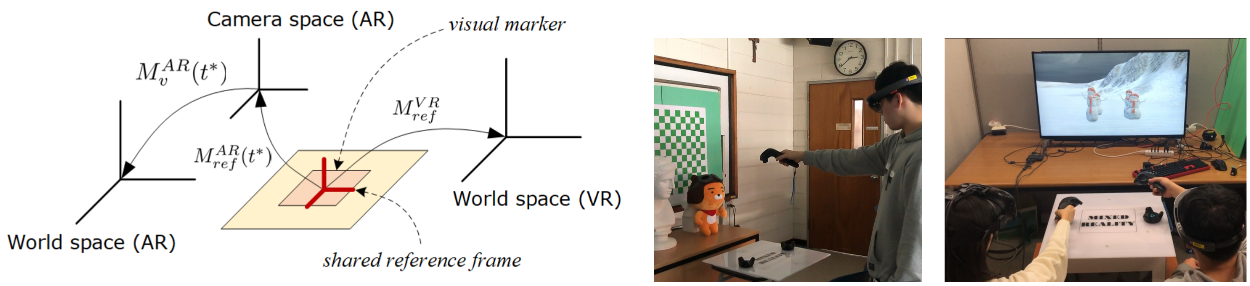
On sharing physical geometric space between augmented and virtual reality environments

Wooyoung Chun* Gyujin Choi* Jaepung An* Woong Seo* Sanghun Park† Insung Ihm*

*Dept. of Computer Sci. and Eng., Sogang University, Korea †Dept. of Multimedia, Dongguk University, Korea



(a) Geometric transformation between AR and VR spaces (b) 3D scene modeling (c) Mixed reality game

Figure 1: Exploring geometric transformation between two world spaces. (a) The world coordinate systems are independently set up in shared physical space by AR and VR systems. When a shared reality environment is initialized at time t^*, the 6-DOF camera pose $M_{ref}^{AR}(t^*)$ in the AR space with respect to the shared reference frame is first estimated by detecting a visual marker. Also, the rigid transformation M_{ref}^{VR} is estimated in the VR space using VR trackers via the method proposed in this paper. Then, combined with the viewing transformation $M_v^{AR}(t^*)$ for the AR system's camera at time t^*, the matrix $M_{V2A} = M_v^{AR}(t^*) \cdot M_{ref}^{AR}(t^*) \cdot (M_{ref}^{VR})^{-1}$ allows us to freely transform geometry between the two world spaces. (b) & (c) Two example applications have been developed to demonstrate possibilities of the integrated mixed reality system (please see the attached videos for more information).

Abstract

Despite the expected synergistic effects, augmented and virtual reality (AR and VR, respectively) technologies still tend to be discrete entities. In this paper, we describe our effort to enable both AR and VR users to share the same physical geometric space. The geometric transformation between two world spaces, defined separately by AR and VR systems, are estimated using a specially designed tracking board. Once initially obtained, the transformation will allow users to collaborate with each other within an integrated physical environment while making the best use of both AR and VR technologies.

Index Terms: Human-centered computing—Human computer interaction—Interaction paradigms—Mixed/augmented reality

1 Shared Reality Tracking Board and VR Trackers

Recently, Gottlieb presented an interesting idea of combining both AR and VR technologies to enable a user wearing a Microsoft HoloLens headset to utilize HTC Vive controllers [1]. To develop this concept further, we designed a *shared reality tracking board* that helps the geometric relationship between two world spaces independently set up in the physical space by respective VR and AR systems to be found more easily and accurately (see Fig. 2). On the center of the board, marked by a cross, a (virtual) *shared reference frame* $F = (\mathbf{p}, (\mathbf{u}, \mathbf{v}, \mathbf{n}))$ whose geometry is represented by a reference point $\mathbf{p}$ and three primary directions $(\mathbf{u}, \mathbf{v}, \mathbf{n})$ is defined. A (virtual) shared reference frame is located on the center $\mathbf{p}$ of the board where the $\mathbf{u}$ and $\mathbf{v}$ directions point right and up on the plane, respectively, and the $\mathbf{n}$ direction comes out of the board. A marker image is printed on the center of the board aligned with the frame. Around the center, there are six circles incised into the acrylic board so that the VR system trackers are placed exactly at aligned locations. During the initial setup, the frame's geometry is decided with respect to the AR system's world space by registering the visual marker using AR SDK software (see Fig. 1(a)). Then, the geometry within the VR system's world space is estimated by utilizing VR trackers. Generally, a VR tracker detects its 6-DOF pose; however, when it is put on the board, it is not easy for the user to align the local coordinate system of the tracker precisely to that of the shared reference frame. Thus, to allow easy calibration, we assume that either the position $\mathbf{p}$ or the pair of the position and upward direction $(\mathbf{p}, \mathbf{n})$ is provided by the tracker.

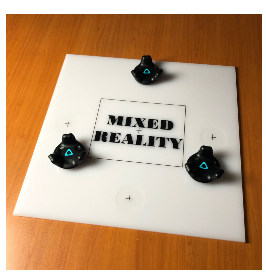


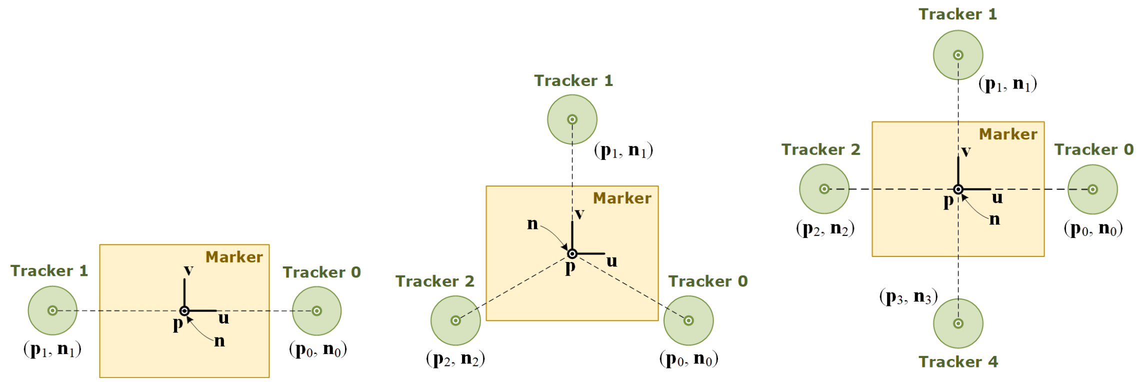
Figure 2: Shared reality tracking board.

In extensive experiments, we carried out testing with three geometry estimation models classified by how many VR trackers are involved. Fig. 3 illustrates the actual layout of the tracking board corresponding to each of the three models (compare the figures to Fig. 2). Around the shared reference frame $F = (\mathbf{p}, (\mathbf{u}, \mathbf{v}, \mathbf{n}))$, VR trackers are placed on the board so that their centers are positioned precisely on a circle with center $\mathbf{p}$ and a radius of 200 mm. The center positions of the two, three, and four trackers on the circle form a line, an equilateral triangle, and a square, respectively. Then, depending on whether the upward direction $\mathbf{n}$ is utilized, we have five cases in total, each with a numerical procedure to estimate the frame geometry in the VR space, as summarized in Table 1. Once all needed transformations are known, we can now compose the rigid transformation from the VR to AR environments (and vice versa), enabling both AR and VR technologies to be explored in an integrated physical space (see Fig. 1(a) again).

2 Experiments on the Accuracy of Frame Geometry

In order to (indirectly) evaluate the accuracy of the geometry estimated for the shared reference frame in the VR space, we also manufactured a *VR tracking-error estimation board* on which two

*{dndud0802,chlrbwls95,ajp5050,wng0620,ihm}@sogang.ac.kr
†mshpark@dongguk.edu



(a) Two trackers (b) Three trackers (c) Four trackers

Figure 3: Design plans for the shared reality tracking board according the number of VR trackers used.

interrelated frames were defined (see Fig. 4). Given the two estimates $\bar{F}_a = (\bar{\mathbf{p}}_a, (\bar{\mathbf{u}}_a, \bar{\mathbf{v}}_a, \bar{\mathbf{n}}_a))$ and $\bar{F}_b = (\bar{\mathbf{p}}_b, (\bar{\mathbf{u}}_b, \bar{\mathbf{v}}_b, \bar{\mathbf{n}}_b))$ obtained respectively by the presented method, the geometric relationship between the two frames suggests that the translational and rotational error metrics ε_{trans} and ε_{rot} defined in the caption of Table 2 provide good (although not absolute) insights into the geometric errors that occurred while tracking the shared reference frame within the VR system's environment.

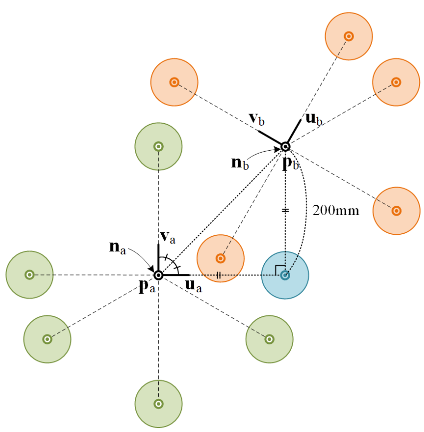


Figure 4: Design plan for the VR tracking-error estimation board. The two shared reference frames F_a and F_b are 60° rotated to each other around the upward direction and $200\sqrt{2}$ mm apart.

In this test, we selected four locations near the center of the VR space, and averaged the errors obtained while estimating the geometries of the two frames eight times per location. The preliminary test results given in Table 2 indicate that fairly accurate estimation was achieved by each of the five numerical procedures. A more sophisticated accuracy analysis that also reflects the marker tracking process in the AR system's environment remains a future work.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF2017R1D1A1B03029625).

References

[1] D. Gottlieb. Mixing reality with virtual reality. http://drewgottlieb.net-/2017/01/31/mixing-reality-with-vr.html, 2017.

Table 1: Numerical procedures to estimate the shared reference frame $F = (\mathbf{p}, (\mathbf{u}, \mathbf{v}, \mathbf{n}))$. We handle both cases where the VR trackers do and do not provide the upward direction $\mathbf{n}_i$ in addition to the reference point $\mathbf{p}_i$. Here, $\mathbf{q}_i = \mathbf{p}_{(i+1)\%n_{tr}} - \mathbf{p}_i$, in which n_{tr} is the number of trackers used and '%' denotes the modular operation. Also, the upward directions $\mathbf{n}_i$ are assumed to be unit vectors.

n_{tr}	$\mathbf{n}_i$	numerical procedure
2	No	N/A
	Yes	$\mathbf{p} \leftarrow \frac{1}{2}(\mathbf{p}_0 + \mathbf{p}_1)$; $\mathbf{n} \leftarrow \mathbf{normalize}(\mathbf{n}_0 + \mathbf{n}_1)$; $\bar{\mathbf{u}} \leftarrow \mathbf{normalize}(-\mathbf{q}_0)$; $\mathbf{v} \leftarrow \mathbf{n} \times \bar{\mathbf{u}}$; $\mathbf{u} \leftarrow \mathbf{v} \times \mathbf{n}$;
3	No	$\mathbf{p} \leftarrow \frac{1}{3}(\mathbf{p}_0 + \mathbf{p}_1 + \mathbf{p}_2)$; $\mathbf{n} \leftarrow \mathbf{normalize}(\mathbf{q}_0 \times \mathbf{q}_1)$; $\mathbf{u} \leftarrow \mathbf{normalize}(-\mathbf{q}_0 - \mathbf{q}_1 + \mathbf{q}_2)$; $\mathbf{v} \leftarrow \mathbf{n} \times \mathbf{u}$;
	Yes	$\mathbf{p} \leftarrow \frac{1}{3}(\mathbf{p}_0 + \mathbf{p}_1 + \mathbf{p}_2)$; $\mathbf{n} \leftarrow \mathbf{normalize}(\mathbf{n}_0 + \mathbf{n}_1 + \mathbf{n}_2)$; $\bar{\mathbf{u}} \leftarrow \mathbf{normalize}(-\mathbf{q}_0 - \mathbf{q}_1 + \mathbf{q}_2)$; $\mathbf{v} \leftarrow \mathbf{n} \times \bar{\mathbf{u}}$; $\mathbf{u} \leftarrow \mathbf{v} \times \mathbf{n}$;
4	No	$\mathbf{p} \leftarrow \frac{1}{4}(\mathbf{p}_0 + \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3)$; **for** $(i = 0$ **to** $3)$ $\bar{\mathbf{n}}_i \leftarrow \mathbf{normalize}(\mathbf{q}_{(i+1)\%4} \times \mathbf{q}_i)$; $\mathbf{n} \leftarrow \mathbf{normalize}(\bar{\mathbf{n}}_0 + \bar{\mathbf{n}}_1 + \bar{\mathbf{n}}_2 + \bar{\mathbf{n}}_3)$; $\bar{\mathbf{u}} \leftarrow \mathbf{normalize}(-\mathbf{q}_0 - \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3)$; $\mathbf{v} \leftarrow \mathbf{n} \times \bar{\mathbf{u}}$; $\mathbf{u} \leftarrow \mathbf{v} \times \mathbf{n}$;
	Yes	$\mathbf{p} \leftarrow \frac{1}{4}(\mathbf{p}_0 + \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3)$; $\mathbf{n} \leftarrow \mathbf{normalize}(\mathbf{n}_0 + \mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_3)$; $\bar{\mathbf{u}} \leftarrow \mathbf{normalize}(-\mathbf{q}_0 - \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3)$; $\mathbf{v} \leftarrow \mathbf{n} \times \bar{\mathbf{u}}$; $\mathbf{u} \leftarrow \mathbf{v} \times \mathbf{n}$;

Table 2: Accuracy evaluation for the shared reference frame estimated in the VR space. Each pair of numbers represents the translational and rotational errors induced between the two shared reference frames, which are defined as $\varepsilon_{trans} = ||\bar{\mathbf{p}}_a - \bar{\mathbf{p}}_b|| - 200\sqrt{2}$ (mm) and $\varepsilon_{rot} = \cos^{-1}\frac{trace(R_\varepsilon)-1}{2} - 60$ (deg), respectively. Here, $R_\varepsilon = R(\bar{\mathbf{u}}_b, \bar{\mathbf{v}}_b, \bar{\mathbf{n}}_b) \cdot R(\bar{\mathbf{u}}_a, \bar{\mathbf{v}}_a, \bar{\mathbf{n}}_a)^t$ where $R(\mathbf{x}, \mathbf{y}, \mathbf{z})$ is a 3×3 matrix made of three column vectors, $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$.

$\mathbf{n}_i$	n_{tr}		
	2	3	4
No	N/A	0.680mm / 0.535°	1.076mm / 0.518°
Yes	0.869mm / 0.376°	0.589mm / 0.531°	0.865mm / 0.589°