Enhancing Consistency in Sensible Mixed Reality Systems: a Calibration Approach Integrating Haptic and Tracking Systems

Mehdi Gorjian and Francis Quek

July 15, 2024

Enhancing Consistency in Sensible Mixed Reality Systems: A Calibration Approach Integrating Haptic and Tracking Systems

1st Mehdi Gorjian
Department of Architecture
Texas A&M University
College Station, TX United States
mgorjian@tamu.edu

2nd Francis Quek
School of Education & Human Development
Texas A&M University
College Station, TX United States
quek@tamu.edu

***Abstract*—This paper discusses the challenges of integrating close-range interactions, such as touch and physical proximity, into Mixed Reality (MR) environments. While current MR systems excel at providing visual overlays of digital content onto the physical environment, the nuanced aspects of close-range interactions remain underdeveloped. The research aims to enhance this integration by focusing on haptic and tracking systems, creating a more immersive and responsive MR experience that mimics real-world interactions. The text emphasizes the need for precise calibration using the Automatic Calculation of the Transformation Matrices method to ensure accuracy in MR environments, particularly when multiple systems are utilized. The proposed approach showcases the effectiveness of haptic and tracking camera systems in real-world applications, offering a solution for seamless and efficient MR calibration and contributing to the advancement of precise and immersive mixed-reality experiences.**

***Index Terms*—Mixed Reality, Calibration, Tangible Interactions**

I. Introduction & Background

Mixed reality (MR) technologies have brought about a new era of technology, seamlessly merging digital and physical realms to provide an immersive and engaging user experience. The efficacy of MR is seen in the seamless incorporation of digital components into the tangible environment. However, the full utilization of MR's capabilities is hindered by the existing constraints on incorporating kinesthetic and proprioceptive sensory modalities, which are especially noticeable when interacting with digital items in spatial environments.

An essential aspect is the comprehensive integration of these sensory modalities, particularly for activities that require complex design cognition [1], [2]. An ongoing obstacle in MR interfaces is the smooth integration of the tactile aspects related to interactions at close distances in the real world. Although current MR systems Fig.2 are effective at presenting visually enhanced content that is superimposed onto the real world, they do not fully handle the subtle components of interactions that occur at close distances, such as touch and physical proximity [3], [4]. This difference highlights the need for a focused attempt to combine haptic and tracking systems, thus enhancing the tangible aspects of physical interactions in the MR framework.

It is imperative to acknowledge that the incorporation of haptic and tracking technology fundamentally changes the mixed reality (MR) encounter. Therefore, a crucial obstacle arises in the necessity to accurately adjust these systems, reducing inaccuracies and guaranteeing a smooth integration of digital and physical components. Calibration is an essential part of integrating these technologies, with the goal of achieving seamless interaction without sacrificing the integrity of the mixed-reality experience. Our research aims to investigate a calibration method that effectively incorporates haptic and tracking devices into the MR framework, considering the complexities involved. This calibration method aims to improve both the physical features of close-range interactions and reduce system mistakes, resulting in a more precise and dependable mixed-reality experience.

Multiple methods for calculating the orientation and transformation matrix were investigated. Horn [5] proposed solution uses unit quaternions for rotation representation, enhancing precision in translational offset, scale, and rotation, outperforming approximate methods based on limited points. Havelock [6] analyzed the precision of estimating a target's position in digital images by considering possible digital representations, leading to indistinguishable locales. It reveals how noise tolerance for position estimation is affected by the topological properties of locales. Another method by Umeyama [7] used singular value decomposition and least-squares estimation to determine rotation, translation, and scaling between data points, rejecting reflection-based solutions and instead choosing the closest transformation matrix solution, claiming it consistently yields accurate solutions even with corrupted data. Lorusso et al. [8] compared four widely used algorithms for calculating 3-D rigid transformations in machine vision. The algorithms encode the transform using singular value de-

This work is supported by the National Science Foundation (NSF) under Grant No. 2008800

composition, orthonormal matrices, unit quaternions, and dual quaternions. The comparison evaluates precision, robustness, and computational speed in handling noise, degenerate data sets, and relative computational speed.

II. Method

We utilized OptiTrack, an advanced Optical-passive technology known for its precise tracking of retroreflective markers using a series of infrared cameras. The purpose of this optical tracking system is to accurately record and monitor the motions of these markers, offering real-time data on their positions and rotation angles. The system functions by strategically positioning infrared cameras in the surroundings, which subsequently monitor the reflected markers on a certain object or subject. OptiTrack offers a significant benefit by enabling highly accurate monitoring of stiff bodies. This implies that the system has the capability to not only identify the current position of the markers but also furnish comprehensive data regarding their rotational angles. The results obtained from the rigid body tracking system are shown in a customizable coordinate system, which improves the system's versatility and capacity to adapt to different experimental configurations. The real-time data generated by OptiTrack enables an in-depth understanding of the dynamic motions and positions of the tracked markers inside the designated coordinate system. OptiTrack's high level of precision and ability to track in real-time makes it an invaluable tool for virtual reality applications and augmented reality experiences. The dependable and adaptable characteristics of the coordinate system enhance the accuracy and flexibility of the data acquired using this optical tracking technique. Conversely, the Touch haptic device, in conjunction

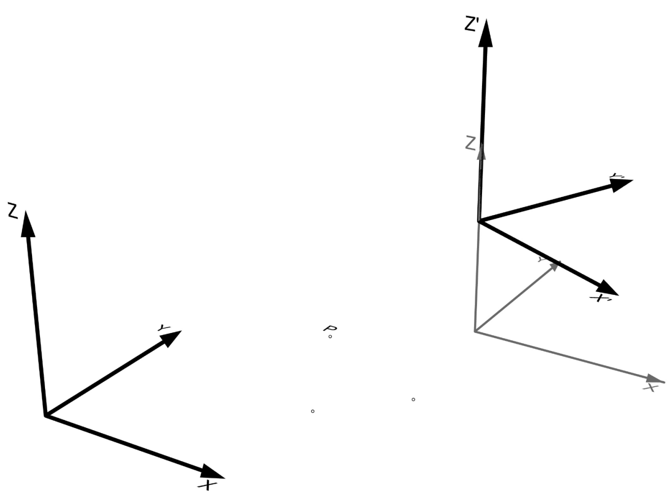

Fig. 1: Haptic and OptiTrack coordinate sytems

with its OpenHaptics API, provides a fully immersive virtual environment interaction that is tailored for a workstation that is roughly the size of a letter-sized piece of paper. With the help of this haptic device, users can feel as though they are physically connected to virtual objects within this specific

Fig. 2: Mixed Reality environment setup and calibration

spatial range, which increases engagement and realism. It is easy to integrate a virtual object into the haptic workflow. But there is a problem when you want to integrate the Touch haptic device with other systems, especially because each of them uses a different coordinate system Fig.1. Notably, the Touch haptic device's preset coordinate system limits its ability to be modified to meet the needs of an external tracking system or particular application. When attempting to synchronize the Touch haptic device with other tracking technologies, like an optical tracking system that uses its own coordinate framework, this becomes more relevant. A unified coordinate system is essential to overcoming this gap in coordination and creating a smooth connection between different systems. This unified system operates as a common reference point, bridging the gap between the external tracking system's coordinate system—which may be based on camera data or the haptic device's intrinsic coordinate system—and the native coordinate system of the Touch haptic device. A number of variables, including the particular application requirements, precision considerations, and the desired degree of synchronization between the haptic and tracking systems, influence the decision between the camera and the haptic coordinate system.

We used an approach to solve the matrix problem that was based on the work of Cashbaugh and Kittis [2], utilizing linear regression optimization techniques. According to this methodology, our implementation makes use of their insights to maximize the matrix equation's solution, improving the computing process's overall correctness and efficiency. Our method seeks to further improve the matrix equation's resolution by extending and modifying these ideas, guaranteeing the system's stable calibration. This application of optimization methods is consistent with previous work in the field but also represents a deliberate incorporation of previous understanding to improve the computational features of our investigation.

Our goal in this case is to find the transformation matrix T so that $V_B = T \cdot V_A$ will be used to explain the relationship between the vectors V_B and V_A. The transformation that translates the coordinates from the original space defined by V_A to the altered space represented by V_B is what we are trying to compute.

$$V_B = \begin{bmatrix} x_B \\ y_B \\ z_B \\ 1 \end{bmatrix} T = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} V_A = \begin{bmatrix} x_A \\ y_A \\ z_A \\ 1 \end{bmatrix} \quad (1)$$

To precisely characterize this geometric transformation, this entails determining the rotation and translation operations contained in the matrix T (1). The resulting matrix T is an essential element for smoothly transferring between these two vector spaces, allowing accurate and effective transformations inside the framework of mathematics.

$$\begin{aligned} x'_B &= r_{11}x_A + r_{12}y_A + r_{13}z_A + t_x \\ y'_B &= r_{21}x_A + r_{22}y_A + r_{23}z_A + t_y \\ z'_B &= r_{31}x_A + r_{32}y_A + r_{33}z_A + t_z \end{aligned} \quad (2)$$

The terms x'_B, y'_B, z'_B on the left side of the equations (2) are called the expected values in this context. The aim is to optimize these values, which are produced by a process of calculation. In order to accomplish this optimization, the first derivative of the equations (2) is subjected to the Mean Square Error (MSE) E (3):

$$E = \sum_{i=1}^{n} (V_B - V'_B) \quad (3)$$

Minimizing these derivatives and bringing them closer to *zero* is the goal. In order to get a precise fit between the expected and actual values, it is imperative to minimize the first derivatives. Deviations between the expected and actual values are quantified by using Mean Squared Error (MSE), a statistical measure frequently employed in optimization problems. The focus in this particular instance is on the equations' first derivatives, and the goal is to get these derivatives as near to zero as feasible. Through this optimization procedure, the calculated values for the coefficients r_{ij} are established. These coefficients are essential for improving the model's ability to predict outcomes. The optimization process, which is driven by the first derivatives' minimization, helps improve the model's overall performance by improving the accuracy of the predictions. We express r_{ij} and t_x, t_y, t_z in terms of a matrix-vector multiplication, with the matrix A (4) taking the non-singular form:

$$A = \begin{bmatrix} \sum(x_{Ai}^2) & \sum(x_{Ai}y_{Ai}) & \sum(x_{Ai}z_{Ai}) & \sum(x_{Ai}) \\ \sum(x_{Ai}y_{Ai}) & \sum(y_{Ai}^2) & \sum(y_{Ai}z_{Ai}) & \sum(y_{Ai}) \\ \sum(x_{Ai}z_{Ai}) & \sum(y_{Ai}z_{Ai}) & \sum(z_{Ai}^2) & \sum(z_{Ai}) \\ \sum(x_{Ai}) & \sum(y_{Ai}) & \sum(z_{Ai}) & n \end{bmatrix} \quad (4)$$

Meanwhile, r_{ij} and t_x, t_y, t_z can be obtained from these equations:

$$\begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ t_x \end{bmatrix} = A^{-1} \begin{bmatrix} \sum_{i=1}^{n} x_{Bi} \cdot x_{Ai} \\ \sum_{i=1}^{n} x_{Bi} \cdot y_{Ai} \\ \sum_{i=1}^{n} x_{Bi} \cdot z_{Ai} \\ \sum_{i=1}^{n} x_{Bi} \end{bmatrix}$$

$$\begin{bmatrix} r_{21} \\ r_{22} \\ r_{23} \\ t_y \end{bmatrix} = A^{-1} \begin{bmatrix} \sum_{i=1}^{n} y_{Bi} \cdot x_{Ai} \\ \sum_{i=1}^{n} y_{Bi} \cdot y_{Ai} \\ \sum_{i=1}^{n} y_{Bi} \cdot z_{Ai} \\ \sum_{i=1}^{n} y_{Bi} \end{bmatrix}$$

$$\begin{bmatrix} r_{31} \\ r_{32} \\ r_{33} \\ t_y \end{bmatrix} = A^{-1} \begin{bmatrix} \sum_{i=1}^{n} z_{Bi} \cdot x_{Ai} \\ \sum_{i=1}^{n} z_{Bi} \cdot y_{Ai} \\ \sum_{i=1}^{n} z_{Bi} \cdot z_{Ai} \\ \sum_{i=1}^{n} z_{Bi} \end{bmatrix}$$

III. Results

In our comprehensive evaluation, we rigorously tested the system's performance by subjecting it to the analysis of 68 distinct point groups, each comprising six sample points Fig.3. To gauge the accuracy of our system, we employed the Mean Squared Error (MSE) metric (measured in mm), comparing the Haptic-captured data against the corresponding data calculated using the transformation matrix. The outcomes of this meticulous analysis furnish valuable insights into the precision of our system. The calculated mean MSE across all tested point groups was determined to be 0.203141, providing a central measure of the overall accuracy.

TABLE I: Statistics metrics

Metrics	***Value (mm)***
Mean	0.203141
Median	0.141731
Standard Deviation	0.199265
Variance	0.039706
Minimum	0.022448
Maximum	0.898051
Range	0.875603
1st Quartile (Q1)	0.073249
2nd Quartile (Q2, Median)	0.141731
3rd Quartile (Q3)	0.217622

The median, representing the middle value of the MSE dataset, was found to be 0.141731. Further, the standard deviation, a measure of the variability in MSE values, was computed as 0.199265, while the corresponding variance was 0.039706. The analysis also revealed specific data points such as the minimum MSE (0.022448) and maximum MSE (0.898051), providing a range of 0.875603. Quartiles, including the first quartile (Q1) at 0.073249, median (Q2) at 0.141731, and third quartile (Q3) at 0.217622, contribute to a comprehensive understanding of the distribution of MSE values. This statistical approach affords a nuanced perspective on the accuracy and variability exhibited by our system across diverse point groups.

Given that 50% of the MSE values are below this cutoff, the box plot's 4 median value of 0.141731 indicates that the data's central tendency is reflected in this finding. The following is the interquartile range (IQR): The box, which spans from the first quartile (Q1), which is at 0.073249, to the third quartile (Q3), which is at 0.217622, represents the interquartile range (IQR). The spread, which represents the middle 50% of the data, is described by the IQR. Outliers are defined as any points that deviate from the data range and have whiskers that

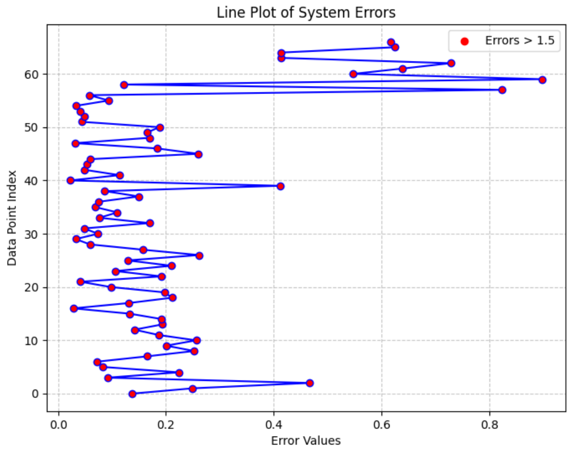

Fig. 3: Line Plot of System Errors

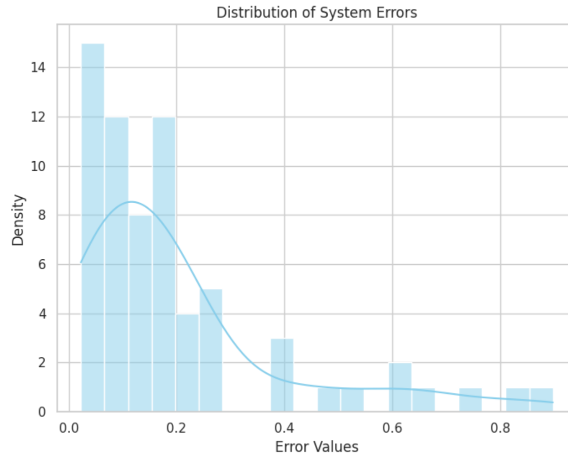

Fig. 5: Distribution of System Errors

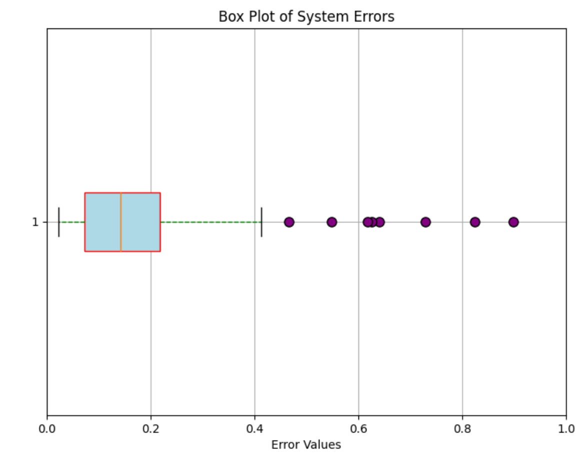

Fig. 4: Box Plot of System Errors

extend beyond the box. In this case, the data points peak at 0.898051, indicating the potential for outliers in the higher MSE value range.

Central tendency can be measured using the mean (0.203141) and median (0.141731) values 5. The median denotes the distribution's middle point, and the mean shows the MSE's average value. The distribution may be skewed to the right because the median is smaller than the mean. The variance (0.039706) and standard deviation (0.199265) measure how far apart the MSE values are. Understanding the range between the lowest (0.022448) and highest (0.898051) MSE values is possible thanks to the range (0.875603). The large range, with values ranging from relatively low to high mistakes, points to a varied distribution of system errors.

The comment highlights how important it is to understand that system errors include both intrinsic system errors and, most importantly, user errors. User errors, particularly those that transpire during the point-picking process, are quite important. This is especially important when taking into account the possibility that the haptic tip and tracker center are not aligned properly. Conversely, "user errors" highlight faults induced by human contact with the system, especially during the crucial point-picking process. In order to do this, the user must choose particular locations in the real or virtual space. These inaccuracies have a complex nature and can be caused by a variety of factors, including the accuracy of the user's motions, visual perception, and the alignment of the haptic tip with the center of the tracker. One particular area of potential mistake is highlighted by the statement that the haptic tip is not always at the center of the tracker. Users' interaction haptic tip may or may not line up with the tracker's center, which introduces unpredictability into the spots that are chosen. This misalignment may lead to variations in the overall point-picking accuracy and, as a result, compromise the output reliability of the system.

IV. Conclusion

In conclusion, our research addresses the challenges and limitations in the current state of mixed reality (MR) systems, specifically focusing on integrating kinesthetic and proprioceptive sensory modalities for precise spatial interactions. While MR has successfully merged digital and physical environments, the finer aspects of close-range interactions, such as touch and physical proximity, still need to be developed. The paper identifies the need for a more comprehensive integration of haptic and tracking systems to enhance the tangible aspects of physical interactions within the MR framework. This is particularly crucial for tasks involving design cognition. The integration of technologies like OptiTrack, known for its precise optical tracking, and the Touch haptic device provides a tangible virtual experience and forms the basis of our approach. Calibration emerges as a critical aspect of seamlessly combining haptic and tracking technologies.

The comprehensive evaluation of our system, involving 68 point groups and utilizing Mean Squared Error analysis, demonstrates promising results with an MSE of 0.203141, indicating improved accuracy. However, it is crucial to acknowledge the inherent complexities in the system, encompassing both system and user errors. User errors, especially during point picking, are highlighted, emphasizing the potential misalignment of the haptic tip with the tracker's center. This recognition underscores the need for a holistic understanding of inaccuracies in MR systems, encompassing both technical and user-related factors.

Our research contributes to the ongoing efforts to advance MR technologies by proposing a calibrated system that effectively integrates haptic and tracking technologies, addressing challenges in close-range interactions and user engagement. The findings open avenues for further exploration and refinement, paving the way for more immersive and accurate mixed-reality experiences.

ACKNOWLEDGMENT

This work would not have been possible without the generous support of the National Science Foundation (NSF) under Grant No. 2008800. We extend our sincere gratitude to the NSF for their financial backing, which enabled us to conduct the research and experiments presented in this paper. We would like to express our appreciation to the program officers and reviewers at the NSF for their valuable insights and feedback throughout the duration of the project. Their guidance has been instrumental in shaping the direction of our research and ensuring its quality. Additionally, we want to thank our colleagues and collaborators for their contributions and support during the course of this investigation. Their expertise and collaboration enriched the overall research process and outcomes. Lastly, we extend our heartfelt thanks to all individuals who directly or indirectly contributed to the success of this project. Your support, encouragement, and expertise have played a significant role in the completion of this work.

REFERENCES

[1] A. D. Cheok, X. Yang, Z. Z. Ying, M. Billinghurst, and H. Kato, "Touch-space: Mixed reality game space based on ubiquitous, tangible, and social computing," *Personal and ubiquitous computing*, vol. 6, pp. 430–442, 2002. Publisher: Springer.

[2] J. Cashbaugh and C. Kitts, "Automatic calculation of a transformation matrix between two frames," *IEEE Access*, vol. 6, pp. 9614–9622, 2018. Publisher: IEEE.

[3] S. Hormoz and M. P. Brenner, "Design principles for self-assembly with short-range interactions," *Proceedings of the National Academy of Sciences*, vol. 108, no. 13, pp. 5193–5198, 2011. Publisher: National Acad Sciences.

[4] A. Palacios-Ibáñez, M. Alonso-García, M. Contero, and J. D. Camba, "The influence of hand tracking and haptic feedback for virtual prototype evaluation in the product design process," *Journal of Mechanical Design*, vol. 145, no. 4, p. 041403, 2023. Publisher: American Society of Mechanical Engineers.

[5] B. K. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Josa a*, vol. 4, no. 4, pp. 629–642, 1987. Publisher: Optica Publishing Group.

[6] D. I. Havelock, "The topology of locales and its effects on position uncertainty," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 13, no. 04, pp. 380–386, 1991. Publisher: IEEE Computer Society.

[7] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 13, no. 04, pp. 376–380, 1991. Publisher: IEEE Computer Society.

[8] A. Lorusso, D. W. Eggert, and R. B. Fisher, *A comparison of four algorithms for estimating 3-D rigid transformations*. University of Edinburgh, Department of Artificial Intelligence, 1995.

[9] D. Cermak-Sassenrath, M. Faust, and H. Rosch, "AirKanoid—visual presentation vs. physical proximity in mixed reality entertainment applications," *PerGames*, vol. 5, pp. 2–11, 2005.

[10] R. R. Mohanty, A. S. Raina, S. Chaudhuri, F. Quek, S. Sueda, and V. R. Krishnamurthy, "Spatial Manipulation in Virtual Peripersonal Space: A Study of Motor Strategies," *Journal of Computing and Information Science in Engineering*, vol. 23, no. 2, p. 021004, 2023. Publisher: American Society of Mechanical Engineers.