Phasor-inspired Approach for Diagnosing and Analyzing Irregular XR Headset Trajectory Data

Sathaporn Hu
Faculty of Computer Science & Technology
Algoma University
Sault Ste. Marie, Canada
Email: hus@algomau.ca

Derek Reilly
Faculty of Computer Science
Dalhousie University
Halifax, Canada
Email: reilly@cs.dal.ca

***Abstract*—Having a stable of XR infrastructure and good environmental conditions is paramount for good-quality trajectory data; however, this is often not feasible outside lab settings. Microsoft HoloLens (HL) and other extended reality (XR) headsets can experience reliability (e.g., sensor accuracy) and infrastructure issues (e.g., Internet connectivity), which can produce corrupted trajectory data. Ideally, if we have sufficient trajectory information, we should try to restore or reconstruct trajectories. However, if the lack of information is severe, we must purposefully remove some information so that analyses can continue. Information removal here is distinct from dimensionality reduction, which is information-preserving. In this paper, we propose a novel approach inspired by spectral phasor analysis called Fantôme analysis. Instead of trying to restore the original trajectories, we identify the remaining usable information and expunge the rest. In this case, the remaining information is the local translation, and local rotation information, which can be used to create phasor-like vectors. Thus, machine learning techniques commonly used in spectral phasor analysis can still be applied. Finally, this paper serves as a call for the development of more resilient, standardized methods of analysis that will still yield insight even when XR infrastructure is unreliable.**

***Index Terms*—extended reality, exploratory data analysis, trajectory analysis, Microsoft HoloLens**

I. Introduction

Trajectory analysis plays an important role in extended reality (XR) research. For instance, in our work, we want to understand how players navigate through physical environments enhanced with virtual AR objects. Another example is Hong & Jung [1], who sought to understand how users of Foursquare used the application across various locations. Although trajectory analyses are helpful, they are at the mercy of the instruments used. Unfortunately, while advanced AR head-worn displays (AR-HWDs) like Microsoft HoloLens v2 (HL) can collect multiple types of sensor data (e.g., player position and gaze-tracking data), they are not always reliable. Even worse, the error characteristics of commodity devices can be poorly documented by the manufacturers themselves. Thus, researchers are often the ones to discover and document these characteristics. Soares et al. [2] found that HL is too unreliable for certain tasks which require high precision (e.g., precise hand gestures). Meanwhile, Florkowska et al. [3] found that the native method for overlaying virtual content on a quick response (QR) code is error-prone. These findings were not published until years after HL became available [4]. Although algorithms such as the Kalman filter can alleviate misalignment issues [5], obtaining suitable data can be difficult.

We first encountered the issue of abnormality after we organized AR-based interactive digital narratives (IDNs) conducted at two different sites: (1) Nova Scotia Museum of Natural History, and (2) The Cube at Virginia Tech (See [6] for more information about the Cube IDNs). In this paper, we will focus on analyzing the trajectory data collected from *Standville Museum*, an IDN hosted at the Museum of Natural History. To understand the players' behaviours, we first explored their trajectories within the installation. However, through exploratory data analysis (EDA), we found that the trajectories had been translated and rotated in a seemingly arbitrary fashion.

After performing rounds of EDAs on *Standville Museum* data, we found that there were two pieces of reliable information: (1) local translation, and (2) local rotation data. Hence, we developed a novel approach inspired by spectral phasor analysis, which we named Fantôme analysis[1]. This novel approach is similar to those used in human motion analysis, such as Boyd [7], Mejia-Romero et al. [8], and Tigrini et al. [9]. By focusing on this information, we can derive phasor-like quantities. This allows some phasor analysis techniques, like clustering, to be used. The premise of the analysis is that *even if we do not know where the participants were in the physical world, participants who performed similar actions should move and rotate in a similar manner*.

We fitted st-DBSCAN clusters and found that the transformation issues were caused by a combination of programming and HL tracking issues. However, the clusters were still informative enough to identify in-game events and general player behaviours. However, additional analysis using other types of data (e.g., video data) is necessary. Future work and refinement of the Fantôme method will require higher-quality data or data from other sources.

II. Related Work and Background Information

This section presents the foundational concepts underlying our study and discusses prior literature relevant to this work.

[1]Named after the French word for ghost and "Mint Fantôme", a pseudonym of a virtual streamer. The name is also a reference to the ghost character in *Standville Museum*, and **spectral** phasor analysis

We begin by outlining the principles of interactive digital narratives (IDNs) and different types of AR-based IDNs. Then, we discuss the reliability of AR devices, machine learning, and relevant concepts from the formal sciences. As our methods are part of exploratory data analysis (EDA), we will provide a brief introduction to EDA and its relevance. Finally, we explain the relevance of the theory of computation and trajectory analyses.

A. *Interactive Digital Narratives*

Interactive digital narratives (IDN) are narratives that not only engage in storytelling but also allow readers or viewers to influence them in some manner [10]. Examples of IDNs include "Façade", a desktop-based video game by Mateas & Stern [11], which allows players to influence a married couple's conversation by participating in it [10]. Although IDNs can be created on multiple media, we are especially interested in AR-based IDNs. AR technologies allow us to tell a story by superimposing virtual elements onto the physical world. An example of AR-enhanced storytelling is "Unforgivable: Eliza" by Toii in 2018 [12]. In the game, the player explores the White Terror period in Taiwan's history. The game relies on the Global Positioning System (GPS) and cellphones. Unlike "Unforgivable: Eliza" which requires a mobile phone, our narratives are specifically designed for HL.

B. *Augmented Reality Devices and Their Challenges*

1) Reliability: Although AR headsets like HL are powerful, they also have reliability issues that were not well-known at the time of their initial release. For instance, HL v2 was released in November 2019 [13], but some flaws were not known until 2021–2023. Soares et al. [2] found that HL's hand tracking was less accurate than HTC Vive's controllers. Based on this result, they advised against using hand tracking for tasks that required high accuracy. Later on, Florkowska et al. [3] found that HL had an inaccurate QR detection algorithm. While it can detect the code's position, the code's orientation should not be trusted. Navares-Vázquez et al. [14] found that, although HLs could be used as LiDAR scanners, their accuracy was also affected when used outdoors. For instance, HL's accuracy dropped during rainfall.

2) Development Difficulty: As AR researchers often work with recently released devices, they may lack a full understanding of them. This often leads to reliability issues being discovered too late in the research process. Worse, for some devices, poor documentation and tutorials affect development at all stages—from prototyping to debugging [15], [16]. This could lead to additional bugs and deployment challenges. On certain devices, like HL, researchers must consult unofficial sources—e.g., an unofficial blog post by Kiekebosch [17]. Although Takala et al. [16] state that development tools and hardware have evolved, working with headsets still remains a significant challenge.

C. *Phasors, Complex Numbers, and Quaternions*

In this paper, we rely on the concept of phasor, a type of complex number representing sinusoid behaviours [18]. Phasors are often used in physics to aid visualization and understanding of data [19]. As a complex number, a phasor can be written in this format called the rectangular form: $z = a + ib$ where $a, b \in \mathbf{R}$ and $i = \sqrt{-1}$. Related to the complex domain ($\mathbb{C}$) is the quaternion domain ($\mathbb{H}$), which represents 4D numbers. Since $\mathbb{H}$ is an extension of $\mathbb{C}$, $\mathbb{C} \in \mathbb{H}$. Quaternions are useful for their ability to represent 3D rigid-body transformations and to compactly describe 2D and 3D rotations as multiplications [20].

A complex number can also be expressed as a vector: (a, b). However, a phasor is also often expressed as:

$$z = r\angle\theta \tag{1}$$

Here, $r \in \mathbf{R}, r \geq 0$ describes the length of the number, and θ describes the rotation around the origin. r, θ are also known as modulus and argument, respectively. The phasor form is not suitable for all purposes, however. For instance, EDA software such as Tableau and Microsoft Excel relies on the Cartesian coordinate system for visualization. To make a phasor suitable for the Cartesian system, we use this formula to convert one into a 2-vector:

$$\vec{z} = (r \times \cos\theta, r \times \sin\theta) \tag{2}$$

The vector form is also useful for describing 2D cross products. The formula below is based on [21]:

$$(a, b) \times (c, d) = ad - bc \tag{3}$$

Originally, phasor analysis was developed for circuit analysis. However, they have also been adopted for other fields [22]. In biology, Mukherjee & Bhargava [22] developed a phasor-based approach to generate labels for cancer cells. First, they computed spectral phasors from tissue grayscale images. Then, the phasors were clustered into four groups, which were then assigned colours. Then, the authors applied the cluster colours back to the original images to highlight the cancer cells. Physicists also use spectral phasor clusters to understand fluorescence behaviours over time [23].

Motion and gait analysis research also uses phasors to analyze movement data [7]. Mejia-Romero et al. [8] generated phasors from gaze movement data to identify features. Meanwhile, Trigini et al. [9] developed a novel approach, PHASOR, to generate phasors from surface electromyographic (EMG) signals. Their goal was to develop machine learning features for better gait recognition. Both methods showed some effectiveness. Namely, Mejia-Romero et al. showed that phasors generated from gaze movement data could be used to identify different types of gaze directions. Meanwhile, Trigini et al. used phasors to enhance other machine learning methods. Our Fantôme approach extended Mejia-Romero et al., and Trigini et al.'s approaches by applying phasor analysis concepts to human-generated trajectory data obtained from AR-based IDNs.

D. Exploratory Data Analysis

Exploratory data analysis (EDA) is an all-encompassing framework for understanding data and performing statistical analyses. According to Tukey, the creator of EDA, EDA is not a set of procedures. Rather, EDA is an "attitude" towards the data [24]. In essence, EDA argues that we cannot gain usable insight if we simply apply statistical methods to data right away. The data must first be analyzed for anomalies, and visualized to develop a general mental model of the data [24] [25]. In some EDA, data must be transformed to ensure that they are compatible for statistical analyses. A very basic example of data transformation is rank transformation, which converts raw numbers into ranks and allows inferences to be made based on them. First pioneered by Spearman [26], the transformation method allows us to perform statistical tests (e.g., Wilcoxon signed-rank tests) on data even if they do not meet the assumptions of powerful parametric tests (e.g., t-tests). Although rank transformation leads to a loss of information, being able to make any inference is still better than gaining no insight. Machine learning techniques, such as clustering, could also be used for EDA of trajectory data. A notable example is Traclus [27], which identifies common points of visitation among many trajectories.

E. Trajectory Analysis and Uncomputability

Although machine learning algorithms are used to support trajectory analysis and restoration, as software, they are also limited in terms of computability. If a computer cannot compute a certain instance of a problem, then no machine learning algorithm can compute it either. Before discussing strategies to alleviate computational challenges, we must first define what it means for a trajectory restoration to become unsolvable. Here, we will make our argument using quaternions since they can encode all 2D and 3D position and rotation information. It is important to note that We are not arguing that all vectors used in today's software (e.g., game engines) should be replaced with quaternions.

Let $p' \in \mathbb{H}$ be a specific point in a trajectory time series that requires restoration. Let $q_1, q_2, q_3, \ldots \in \mathbb{H}$ representing transformation factor, and $\star_1, \star_2, \star_3, \ldots \in \{+, \times\}$ represent translation and rotation operations that cause the p to become p'. Then, trajectory restoration can be defined as an algorithm that substitutes all q's and $\star$'s in the equation below so that p can be found:

$$p' = q_1 \star_1 q_2 \star_2 q_3 \star_3 \ldots \star_n p \ldots \tag{4}$$

To restore p, we must be able to replace all non-p values using the data. Thus, in essence, we have an inverse problem that involves finding the variables from the observed data [28]. To express this notion in computer science terms, we develop the following proof of contradiction:

Proof 1: A trajectory cannot be restored without sufficient information. This is a proof by contradiction, which involves reducing the trajectory restoration to solving for p. Let $p' \in \mathbb{H}$ be a quaternion variable from a linear equation. Then, p can be expressed as $p' = q_1 \star_1 q_2 \star_2 q_3 \star_3 \ldots \star_n p \ldots$. We let $Res(t')$ be a function that takes a transformed trajectory t' and returns the original trajectory *without fail*. Since p' is a quaternion, it can be considered a single-point trajectory. Thus, $Res(p')$ must correctly return p in the equation regardless of the available information. However, this behaviour violates the fundamental rule of quaternion algebra. Thus, Res cannot exist. ■

It is important to note that the above proof, like any proof of uncomputability, is a statement about a class of problems, and may not apply to all instances of the members of the class [29]. For instance, while Turing [30] shows that we cannot always decide whether an arbitrary program will output a specific symbol, it does *not* mean that we can never decide. In fact, the problem can easily be decided for many programs. We can simply test this by asking an AI chatbot whether a short one-line program will output a specific symbol [31]. Still, if researchers find that their machine learning algorithms repeatedly fail to restore a trajectory despite repeated tuning, they should pause and think whether they are implicitly trying to construct an algorithm like Res. If so, they may be dealing with an unsolvable instance of trajectory restoration, and they should reconsider their process. If restoration is impossible, then we may need to consider a reconstruction. A reconstruction will not produce the original trajectory; however, the result should still be useful for many analyses. An example of this is the Kalman filter [5]. If all else fails, we must redefine a trajectory into a problem that requires less information.

III. Fantôme Analysis

We developed a novel method named the Fantôme technique by first analyzing *Sentiments* and *Standville Museum* data. After performing multiple EDAs on the HL data, we conclude that while the HL could have shifted trajectories and rotated them in an undetermined fashion, HL did not scale the player trajectory data. Therefore, a translation of 1 metre in the physical world would have been observed as 1 metre in the data. We can also know how much rotation has occurred since the previous time step and since the beginning of the trajectory, but the actual direction is unknown. Complex numbers and 2-vectors cannot be used to convey this information, because they contain directional information with a global and shared axis of rotation. Thus, we define a new quantity called a Fantôme quantity. A Fantôme quantity is a phasor-like quantity with two components: modulus and argument. However, both components are only defined relative to the first modulus and argument in the trajectory, and not to the universal spanning axes like complex numbers (the real and imaginary axes) or 2-vectors (the x- and y- axes).

A Fantôme trajectory can be recursively defined as:

$$d_t \angle \theta_t = \begin{cases} 0\angle 0 & \text{if } t = 0 \\ |d_{t-1} + \Delta d_{t-1}| \angle (\theta_{t-1} + \Delta\theta_{t-1}) & \text{if } t > 0 \end{cases} \tag{5}$$

d_t represents the distance from the origin at timestep t. We assume that the first point of the trajectory is simply $(0,0)$. Meanwhile, θ_t represents a rotation around an axis. Unlike complex numbers and 2-vectors which assume that all trajectories share the same axis of rotation, each Fantôme trajectory has its own unshared axis defined at the beginning of the trajectory. When $t = 0$, $\theta = 0$ because we assume that no rotation exists. d_{t-1} represents the previous distance in the trajectory. Δd_{t-1} represents the change of distance. θ_{t-1} represents the current rotation around the axis. $\Delta\theta_{t-1}$ represents the change in rotation from the previous timestep. Let $\vec{p}_t, \vec{p}_{t-1}$ be the points that must be transformed, then formulae for Δd_{t-1} and $\Delta\theta_{t-1}$ are as follows:

$$\Delta d_{t-1} = |\vec{p}_t| - |\vec{p}_{t-1}| \tag{6}$$

$$\Delta\theta_{t-1} = \sin^{-1} \frac{|\vec{p}_{t-1} \times \vec{p}_t|}{|\vec{p}_{t-1}||\vec{p}_t|} \tag{7}$$

In essence, Eq. 5 serves as an approximation of the integration of velocities in the Fantôme domain to obtain displacement from the velocities. Let $v_d(t)$ be the 1D velocity at t, and $v_\theta(t)$ be the angular velocity at t, then:

$$d_t \approx \int_0^t v_d(t)dt \tag{8}$$

$$\theta_t \approx \int_0^t v_\theta(t)dt \tag{9}$$

By applying Eqs. 5,6, & 7 to all collected points in a trajectory, we have converted the trajectory into a Fantôme one. The original trajectory has now been re-interpreted as a series of jumps between multiple orbits (Fig. 1), and rotations around the origin. In essence, the process "reset" the translation and the rotation.

Alternatively, the recursive definition could also be used iteratively to transform a trajectory. Let T be a list of timestamped 2-vectors, sorted by timestamps in ascending order, then the iterative algorithm creates and returns F, a Fantôme trajectory, using the pseudocode below:

Require: $T = [\{t: 0, \vec{p}: (x,y)\}, \ldots]$
Require: $T[1].t < T[2].t < T[3].t < \ldots < T[|T|].t$
Require: $T[1].t = 0$
$F \leftarrow [\{t: 0, d: 0, \theta: 0\}]$
for $i \leftarrow 2$ **to** $|T|$ **do**
  $\vec{p}_i = T[i].\vec{p}$
  $\vec{p}_{i-1} = T[i-1].\vec{p}$
  $\Delta d = |\vec{p}_i| - |\vec{p}_{i-1}|$
  $\Delta\theta = \sin^{-1}(|\vec{p}_i \times \vec{p}_{i-1}|/(|\vec{p}_i||\vec{p}_{i-1}|))$
  $F \cup \{t: T[i].t, d: F[i-1].d + \Delta d, \theta: F[i-1].\theta + \Delta\theta\}$
end for
return F

Since Fantôme quantities are similar to phasors, we can apply the methods used in a spectral phasor analysis like clustering [23], after we convert them into their rectangular forms using Eq. 2. In the spectral phasor analysis, clustering of phasors can be used to generate a label for the original data. For example, Mukherjee & Bhargava [22] generated phasors representing colour spectra of infrared images of cancer tissues. The phasors were then clustered into multiple groups. The group labels were then used to highlight different types of cells in the original samples. Similar to Mukherjee & Bhargava [22], a Fantôme cluster can then be used to label trajectory segments where changes in rotation and distance from the relative origin are similar. Although Vallmitjana, Torrado, & Gratton [23] argue that GMM is the most appropriate technique for clustering phasors due to its ability to deal with sparse data, we argue that st-DBSCAN is more suitable for a Fantôme analysis, as outlier elimination is more important in our case. Furthermore, st-DBSCAN can handle temporal data. However, the application of machine learning techniques is a relatively recent development. In the past, all analyses were done visually [23]. We argue that, whether or not machine learning is used, inspection of data from unreliable sources must be done carefully and critically. The *attitude* is more important than the method. Sec. IV-D provides a practical example of how the Fantôme analysis could be applied.

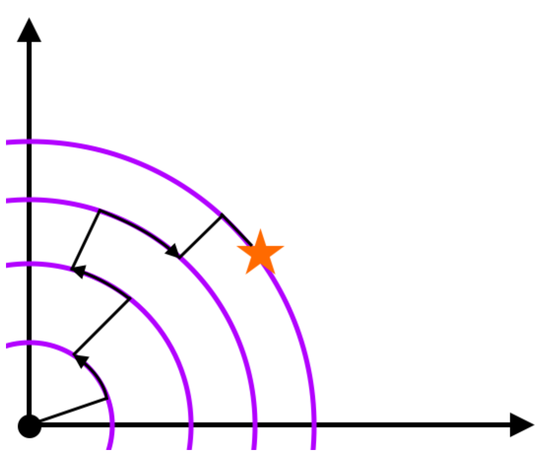

Fig. 1. An example of a Fantôme trajectory.

IV. Analysis of the Standville Museum Data

This section describes our attempt to analyze the Standville Museum data. First, it outlines how more conventional methods are used. At the same time, the failures of the conventional methods are highlighted. Finally, we describe how the Fantôme analysis, along with methods used in spectral phasor analysis, still allows us to perform useful analysis.

A. Description of Standville Museum

Standville Museum was an IDN implemented using Unity, C#, and MRTK library. It was held in a multipurpose room at the Natural History Museum of Nova Scotia during Nocturne 2021, which was a public city-wide art event. In each session, two players visited a museum and explored it in order to uncover its secrets. The IDN was designed to contain some horror elements. For example, one room contained an ominous

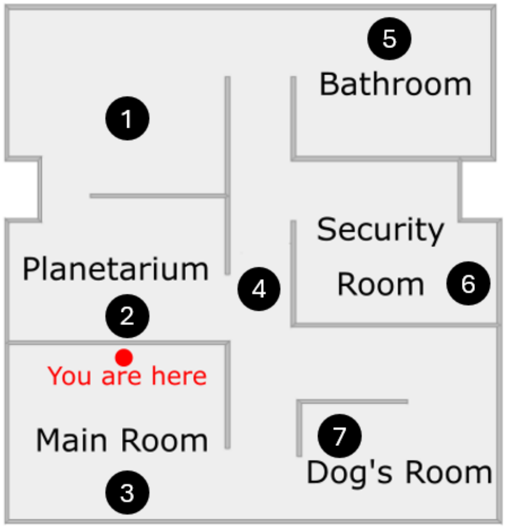

Fig. 2. The map of the room where *Standville Museum*. The players were able to see the map as soon as they entered. The room numbers were not part of the original map, and were only added for the reader's convenience. **Room 1:** A room with a pentagram. **Room 2:** A planetarium. **Room 3:** The entry room. **Room 4:** The central corridor where a ghost character would appear at one point. **Room 5:** The toilet. **Room 6:** The security room. **Room 7:** The room with a virtual dog.

pentagram with incomprehensible chanting. Certain players were told that their son was lost in the museum. In order to find the son, they must navigate through the museum while contending with supernatural elements. This included being chased by a ghost-like entity and interacting with a talking dog. The players were not aware that there would be no penalty for being caught by the ghost, interacting with a pentagram, or not following the dog to their son. To assist the players, a physical actor would be assigned to play a docent—i.e. a museum guide. They provided play instructions, explained the story's premise, and initiated the IDN by guiding the players to the toilet area. Once the narrative was completed, they debriefed the players and assisted them in exiting the room. If the players were unable to complete the game, they would help to terminate the narratives and guide the player out.

B. Description of the Data

This is to provide some context for the data. From each *Standville Museum* play session, we collected the following data:

- **Timestamp**: The time when the trajectory points were being recorded. In the original format, we collected the date-time stamp. The smallest unit is a millisecond. To simplify the calculation, we converted the date-time timestamps to timestamps representing elapsed times during a session.
- **Position**: The HL collected 2D position data during play sessions using. Only x- and y-positions[2] were considered in the trajectory analysis, and we treated each point in the trajectory as a single complex number.

[2]A.k.a. z-positions in Unity, the game engine used in our IDN development. To obtain y-positions, we simply re-label z to y without any conversion.

Fig. 3. **Left:** A video recording of an actual play session. The docent sits in Room 3 after briefing the players. The players, in Room 7, were looking at the virtual dog. **Right:** The virtual dog in Room 7. The image is from a simulated play session.

In addition to the sensor data, we collected other types of data, such as video and audio data. Preliminary EDA indicated that the video data could not be used to reconstruct the original trajectories: some footage was missing, and there was insufficient camera coverage in the room. Furthermore, to maximize HLs' battery power and to optimize the turnaround process, we only collected audio data from the HLs, and not the video data.

C. Initial EDA Attempts

Earlier in our trajectory analyses of *Standvill Museum* data, we were aware that our trajectory data required extensive data wrangling through EDA. This led us to deploy multiple machine learning techniques. First, we explored applying a dimension reduction technique. We started out using PCA on the trajectory, as some prior research had used it for trajectory analyses. For instance, Li et al. [32] used it to analyze trajectories of marine vessels. When PCA proved ineffective, we turned to t-SNE, which also failed to yield any insight. Lastly, we turned to a more advanced technique called UMAP [33], which was also ineffective. As it turns out, dimension reduction was not appropriate, because it still aims to preserve information [34]. Thus, any unusable information introduced by the HL would not have been eliminated.

We then explored clustering approaches, which are often used for trajectory analysis [27]. Some clustering techniques, like DBSCAN, can be used to eliminate outliers [35]. However, clustering also proved inappropriate for the raw trajectory data. This is because expunging outliers may not necessarily restore the original trajectories. Instead of proceeding, we started to suspect the trajectory restoration was impossible.

Based on Florkowska et al. [3] and visual inspection of individual trajectories, we suspected our trajectories could have been rotated and translated by a flawed QR detection code. Furthermore, we were aware that there could also be translations of points as well. For each point in the trajectory, we were essentially trying to recover $p \in \mathbb{C}$ from:

$$p' = r \times t + r \times p \tag{10}$$

where $r \in \mathbb{C}$ represents rotation and $t \in \mathbb{C}$ represents translation. Thus, we must solve for r and t. Since only a

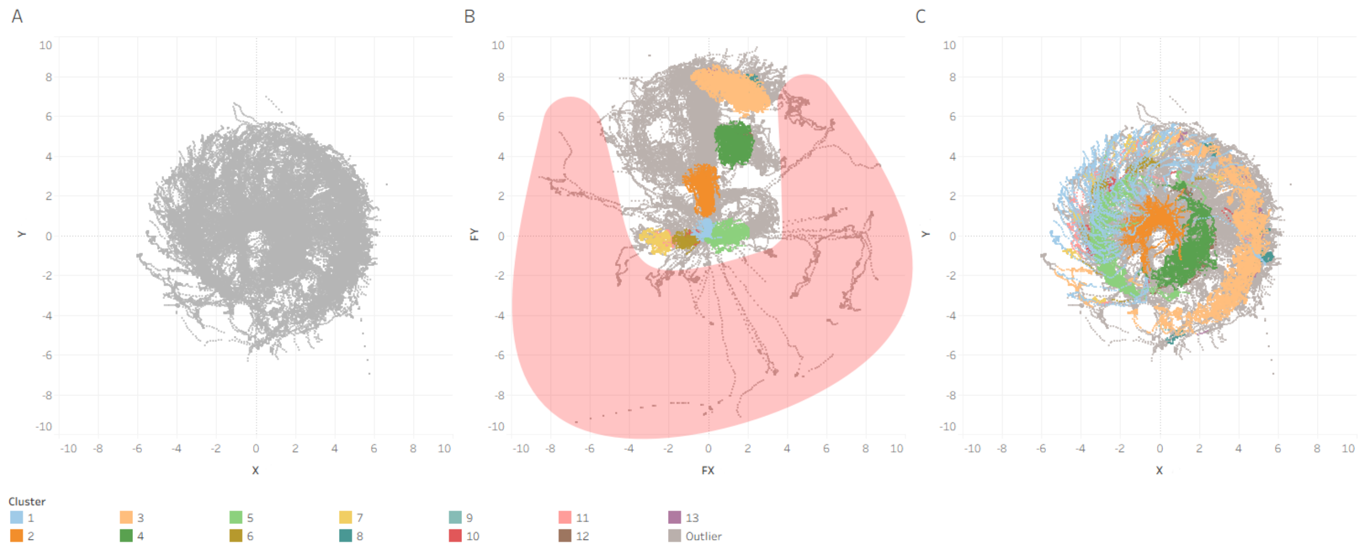

Fig. 4. Examples of the techniques being applied on the *Standville Museum* data. **A** is the original trajectory data. **B** shows trajectories that have been converted to Fantôme ones. st-DBSCAN was then applied with the following parameters: spatial $\epsilon = 0.4$, temporal $\epsilon = 0.5$, minimum number of points = 750. Some clusters are hidden by other points; however, these are generally uninformative. The red transparent bands represent trajectories that were not properly "reset." This indicates hardware errors in the HoloLens. **C** shows **A** with the cluster colour from **B** applied to it.

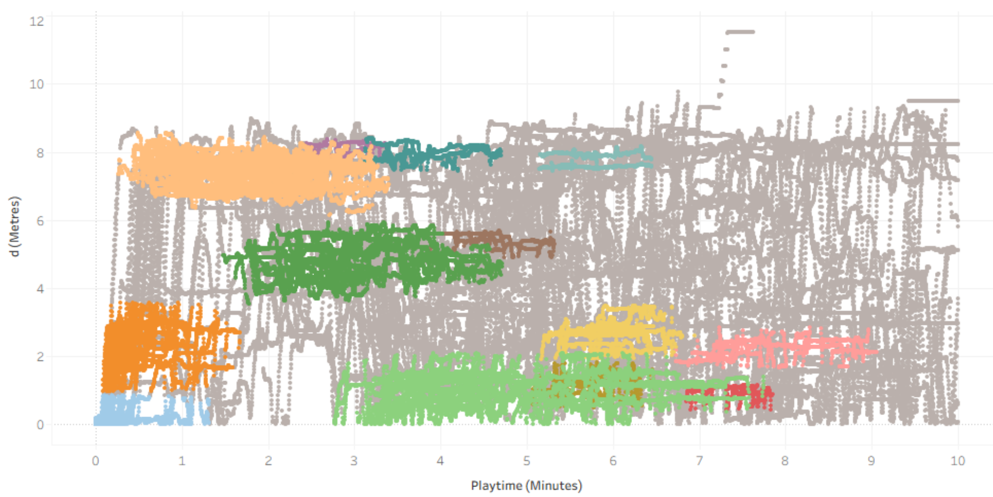

Fig. 5. To further aid the EDA process, Fantôme trajectories could be visualized just by its distances vs. playtime. Please refer to Fig. 4 for the colour legend.

single HL was used to collect data for a single trajectory, there is not sufficient information to find or estimate r, t. Worse, we later found that the HL also had additional sensor issues, which would make Eq. 10 even more complex and more unsolvable.

As such, instead of dealing directly with r, t, we must expunge them. Originally, we considered gesture recognizers. However, such algorithms may have difficulty labelling trajectories with unknown axes of rotation. To better understand this, let's examine two Chinese characters: 上 and 下. These characters have the meaning of *up* and *down* respectively. Since they resemble the reflected versions of each other, they must share the same axis of rotation for correct interpretation. If the axes of rotation of both characters are unknown, then 下 may simply be a reflected 上 with a slight deformation. Thus, the gesture recognizers are not useful in our cases.

D. Second EDA Attempt with Fantôme Analysis

To resolve the issue, we applied the Fantôme analysis with st-DBSCAN. We used the following parameters: spatial $\epsilon = 0.4$, temporal $\epsilon = 0.5$, minimum number of points = 750. The parameters were selected using scatterplots as a visual aid. The goal was to gain a basic overall understanding of participants' actions. We then visualized the clusters as Fig. 4 and Fig. 5. By reviewing the script of the story, we were able to indicate what each cluster represented:

- **Cluster 1:** Players entered the room.

- **Cluster 2:** Players were led by the docent into the washroom.
- **Cluster 3:** Players left the washroom to investigate a disturbance in the security room.
- **Cluster 4:** Players approached the dog, who guided them to the son character.
- **Clusters 5-7:** The story ending and debrief with the docent.
- **Cluster 8:** Players exited from the room through a different door.
- **Other Clusters:** Unusual behaviours specific to certain players which require further examination using other EDA methods.

Even though the Fantôme space is not a Euclidean one, we can still deduce where the participants were by comparing the clusters against the map (Fig. 3). Despite the distortion, the clusters in Fig. 4 could still be visually correlated to the original map. Fig. 4-C indicates that the trajectories have been arbitrarily rotated by HLs during collection. This means that each trajectory has its own axis of rotation. Such a finding is consistent with that of Florkowska et al. [3], who state that using a QR code as an anchor could introduce arbitrary rotation. However, the QR code calibration may not be the sole issue as we suspected. Since the participants did not experience objects being strongly offset from their original positions, the QR code-based calibration must have its potentially erroneous rotation corrected by the other method. As it turns out, the setup process created two positional anchors. The first anchor, used by the position logging software, was calibrated solely by the QR code synchronization. The second anchor, used by the IDN, was also initially calibrated with a QR code; however, additional steps largely eliminated the errors associated with the code.

Although programming caused the major issues, this does not mean HL did not contribute to the trajectory transformation issue. The process that converts trajectories to Fantôme trajectories should have produced trajectories that are independent of global rotation and translation. However, some trajectories were not "reset" as expected. Thus, it is clear that there were some tracking issues with the headset itself. As seen in Fig. 4-B., there are still some irregular trajectories. If these trajectories were accurate, we would have observed participants walking through walls and entering the raised soil around the museum.

V. Limitations, Discussion and Future Work

A. Artificial Intelligence and Trajectory Analysis

When there is a large amount of trajectory data, AI tools are indispensable. However, as Proof 1 suggests, AI tools may fail when we encounter uncomputable instances of trajectory analysis. For some, this may signal the end of data analysis. However, through reframing, simplifying the analyses, and removing information, we have demonstrated that some insights can still be gained. While the information will not be as complete, it may be sufficient to answer research questions and test hypotheses. For our digital narratives, we selectively removed certain trajectory information to create Fantôme trajectories. While AI has become increasingly powerful, we argue that EDA, a human-assisted process, still plays an important role.AI techniques must be chosen with care and with some reservation.

B. Comparison with Ground Truth and Other Data

We were able to use our method to reconstruct trajectories that, upon visual inspection, resembled a distorted version of the original trajectories. However and unfortunately, we could never obtain an undistorted reconstruction, as this involves solving Eq. 10 with insufficient information. Since we lack ground truth, we cannot fully verify the effectiveness of the Fantôme analysis.

Since this work only focuses on the Fantôme analysis, we excluded other types of analyses. However, despite the logistical challenges of implementing the IDNs, we were still able to collect other types of data, such as video recordings, interviews, and more. Although these are insufficient to reconstruct the original trajectories, analyses of the data could potentially further enhance the results of the Fantôme analysis. For instance, although the clusters did not clearly indicate that participants were trying to run away from the ghost character, audio and video recordings of participants reacting to the character could shed more light.

C. Resilient Data Analysis

Failures are common in XR deployments [36]. XR devices may fail due to various reasons. The devices may have design issues, which prevent data from being correctly collected. In our case, some position data were not properly detected. Other possible includes hand gestures not being properly detected, or the weather affecting the LiDAR's accuracy [14]. Additionally, the infrastructure that supports XR devices and their activities (e.g., Wi-Fi, mobile Internet, power) may not be reliable themselves. In our case, the Wifi security measure put in place by the museum adversely affected our XR devices as well as the IDN. In certain cases, such as in-the-wild data collection, Internet connectivity may not exist at all [37]. Setting up a stable XR infrastructure itself can be extremely difficult, especially for public events (e.g., [38]). In addition to the challenges related to failures of XR devices and their infrastructure, there could also be other challenges. For instance, *Standville Museum* experienced a surge of attendees, which forced us to expedite the turnaround procedure to ensure that as many people as possible could experience the IDN.

When XR devices and their infrastructures fail or are under strain, the devices can produce corrupted or incomplete data. In such scenarios, we argue that the data should still be analyzed in order to better understand the failure, and to generate useful insights. Therefore, we argue for robust, resilient statistical analyses and EDA processes that can withstand XR failures.

VI. Conclusion

This paper demonstrates a myriad of challenges when it comes to XR research and development. In an ideal scenario, all sensors and measurement instruments would have been properly installed and calibrated. However, in our experience, several challenges disrupted the process. First, we operated our exhibition within a public space and tried to maximize the number of participants. This meant not all equipment could be set up properly. Our HLs were also pushed to their limits to ensure that everyone could participate. Secondly, the documentation of HL was poor, and the software development process was overly difficult, which led to a challenging debugging process. Despite the challenges, we were still able to make a contribution in terms of the Fantôme analysis. Going forward, we argue that having a resilient XR infrastructure is insufficient. As failures are inevitable, we must also have a robust method of data analysis with a standard that respects the interdisciplinary nature of XR work.

Acknowledgment

Standville Museum began as an IDN authored by several anonymous participants who participated in Singh et al. [39]'s study. We also would like to acknowledge Wallace Lages, Thiago Porcino, and Ryan Chown for their input. Lages provided access to The Cube data, which was used to develop the earlier version of the Fantôme analysis. Porcino developed the play session simulation software, which significantly aided the EDA process. Chown provided his insight into physics and mathematics. All three contributed to the earlier versions of the paper. We would finally like to thank all anonymous reviewers for their valuable feedback. The data used in the analyses were collected from the studies approved by the research ethics board at Dalhousie University (REB #2021-5745), and the institutional review board at Virginia Tech (Protocol #21-788).

References

[1] I. Hong and J.-K. Jung, "What is so "hot" in heatmap? qualitative code cluster analysis with foursquare venue," *Cartographica*, vol. 52, no. 4, pp. 332–348, 2017.

[2] I. Soares, R. B. Sousa, M. Petry, and A. P. Moreira, "Accuracy and repeatability tests on hololens 2 and htc vive," *Multimodal Technologies and Interaction*, vol. 5, no. 8, 2021.

[3] A. Florkowska, M. Trojak, W. Celniak, M. Stanuch, M. Daniol, and A. Skalski, "Metrological analysis of hololens 2 for visual marker-based surgical navigation," in *2023 IEEE International Conference on Imaging Systems and Techniques (IST)*, pp. 1–6, 2023.

[4] J. White, "Microsoft at mwc barcelona: Introducing microsoft hololens 2," Feb. 2023. Blog.

[5] M. B. Rhudy, R. A. Salguero, and K. Holappa, "A kalman filtering tutorial for undergraduate students," *International Journal of Computer Science & Engineering Survey*, vol. 8, no. 1, pp. 1–9, 2017.

[6] C. Connor, E. C. Schoenborn, S. Hu, T. M. Porcino, C. Moore, D. Reilly, and W. S. Lages, "Examining pair dynamics in shared, co-located augmented reality narratives," in *Proceedings of the 2024 ACM Symposium on Spatial User Interaction*, SUI '24, (New York, NY, USA), Association for Computing Machinery, 2024.

[7] J. E. Boyd, "Synchronization of oscillations for machine perception of gaits," *Computer Vision and Image Understanding*, vol. 96, no. 1, pp. 35–59, 2004.

[8] S. Mejia-Romero, J. Eduardo Lugo, D. Bernardin, and J. Faubert, "The dynamic performance of gaze movement, using spectral decomposition and phasor representation," in *Proceedings of International Conference on Data Science and Applications* (K. Ray, K. C. Roy, S. K. Toshniwal, H. Sharma, and A. Bandyopadhyay, eds.), (Singapore), pp. 133–144, Springer Singapore, 2021.

[9] A. Tigrini, R. Mobarak, A. Mengarelli, R. N. Khushaba, A. H. Al-Timemy, F. Verdini, E. Gambi, S. Fioretti, and L. Burattini, "Phasor-based myoelectric synergy features: A fast hand-crafted feature extraction scheme for boosting performance in gait phase recognition," *Sensors*, vol. 24, no. 17, 2024.

[10] H. Koenitz, "Towards a theoretical framework for interactive digital narrative," in *Interactive Storytelling* (R. Aylett, M. Y. Lim, S. Louchart, P. Petta, and M. Riedl, eds.), (Berlin, Heidelberg), pp. 176–185, Springer Berlin Heidelberg, 2010.

[11] M. Mateas and A. Stern, "Façade," 2005.

[12] Toii, "Unforgivable: Eliza," 2018.

[13] A. Palumbo, "Microsoft hololens 2 in medical and healthcare context: State of the art and future prospects," *Sensors*, vol. 22, no. 20, 2022.

[14] J. C. Navares-Vázquez, Z. Qiu, P. Arias, and J. Balado, "Hololens 2 performance analysis for indoor/outdoor 3d mapping," *Journal of Building Engineering*, vol. 108, p. 112826, 2025.

[15] N. Ashtari, A. Bunt, J. McGrenere, M. Nebeling, and P. K. Chilana, "Creating augmented and virtual reality applications: Current practices, challenges, and opportunities," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, (New York, NY, USA), p. 1–13, Association for Computing Machinery, 2020.

[16] T. M. Takala, R. P. McMahan, N. Chawla, E. Kruijff, and T. Kawai, "Xr application development in the 21st century: a survey spanning two decades of xr developers, applications, and challenges," *Frontiers in Virtual Reality*, vol. Volume 7 - 2026, 2026.

[17] S. Kiekebosch, "Qr code tracking in the hololens," Mar. 2021. Blog Post.

[18] L. Schraf, *Phasor Representation of Signals*, ch. 3.2. California, USA: LibreTexts, 2009.

[19] L. Malacrida, S. Ranjit, D. M. Jameson, and E. Gratton, "The phasor plot: A universal circle to advance fluorescence lifetime analysis and interpretation," *Annual Review of Biophysics*, vol. 50, no. Volume 50, 2021, pp. 575–593, 2021.

[20] R. Mukundan, *Quaternions*, pp. 77–112. London: Springer London, 2012.

[21] R. L. Coffie, *Functions of a Complex Variable*, ch. 1. Boca Raton, USA: CRC Press, Sept. 2021.

[22] S. S. Mukherjee and R. Bhargava, "Phasor representation approach for rapid exploratory analysis of large infrared spectroscopic imaging data sets," *Analytical Chemistry*, vol. 95, no. 30, pp. 11365–11374, 2023.

[23] A. Vallmitjana, B. Torrado, and E. Gratton, "Phasor-based image segmentation: machine learning clustering techniques," *Biomed. Opt. Express*, vol. 12, pp. 3410–3422, Jun 2021.

[24] J. T. Behrens, "Principles and procedures of exploratory data analysis," *Psychological Methods*, vol. 2, no. 2, pp. 131–160, 1997.

[25] M. Komorowski, D. C. Marshall, J. D. Salciccioli, and Y. Crutain, *Exploratory Data Analysis*, pp. 185–203. Cham: Springer International Publishing, 2016.

[26] C. Spearman, "The proof and measurement of association between two things," *The American Journal of Psychology*, vol. 100, no. 3/4, pp. 441–471, 1987.

[27] J.-G. Lee, J. Han, and K.-Y. Whang, "Trajectory clustering: a partition-and-group framework," in *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, SIGMOD '07, (New York, NY, USA), p. 593–604, Association for Computing Machinery, 2007.

[28] H. C. Nguyen, R. Zecchina, and J. Berg, "Inverse statistical problems: from the inverse ising problem to data science," *Advances in Physics*, vol. 66, no. 3, pp. 197–261, 2017.

[29] U. Boker and N. Dershowitz, *The Church-Turing Thesis over Arbitrary Domains*, pp. 199–229. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.

[30] A. M. Turing, "On computable numbers, with an application to the entscheidungsproblem," *Proceedings of the London Mathematical Society*, vol. s2-42, no. 1, pp. 230–265, 1937.

[31] S. Hu, "Chatbots as turing machines," in *32nd IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking*

and Parallel/Distributed Computing, IEEE/ACIS SNPD-2025 Summer IV, Institute of Electrical and Electronics Engineers, 2025.

[32] H. Li, J. Liu, R. W. Liu, N. Xiong, K. Wu, and T.-h. Kim, "A dimensionality reduction-based multi-step clustering method for robust vessel trajectory analysis," *Sensors*, vol. 17, no. 8, 2017.

[33] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," 2020.

[34] M. D. Giudice, "Effective dimensionality: A tutorial," *Multivariate Behavioral Research*, vol. 56, no. 3, pp. 527–542, 2021.

[35] M. Çelik, F. Dadaşer-Çelik, and A. Ş. Dokuz, "Anomaly detection in temperature data using dbscan algorithm," in *2011 International Symposium on Innovations in Intelligent Systems and Applications*, pp. 91–95, 2011.

[36] R. Schatz, H. Schrom-Feiertag, G. Papagni, F. Steinicke, L. Skorin-Kapov, M. Billinghurst, L. Oppermann, and G. Aumayr, "Xr for challenging environments - enabling human performance and agency under stress," in *Proceedings of the Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems*, CHI EA '26, (New York, NY, USA), Association for Computing Machinery, 2026.

[37] M. Whitlock, K. Wu, and D. A. Szafir, "Designing for mobile and immersive visual analytics in the field," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 503–513, 2020.

[38] D. Yule, B. MacKay, and D. Reilly, "Operation citadel: Exploring the role of docents in mixed reality," in *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, (New York, NY, USA), p. 285–294, Association for Computing Machinery, 2015.

[39] A. Singh, R. Kaur, P. Haltner, M. Peachey, M. Gonzalez-Franco, J. Malloch, and D. Reilly, "Story creatar: a toolkit for spatially-adaptive augmented reality storytelling," in *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 713–722, 2021.