Correlation between Entropy and Prediction Error in VR Head Motion Trajectories

Silvia Rossi
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
s.rossi@cwi.nl

Laura Toni
University College London (UCL)
London, UK
l.toni@ucl.ac.uk

Pablo Cesar
Centrum Wiskunde & Informatica
TU Delft
Amsterdam, The Netherlands
p.s.cesar@cwi.nl.

ABSTRACT

The general understanding of user behaviour has been often overlooked in the field of Virtual Reality (VR) and Extended Reality (XR) at large. In this work, we want to fill this gap by exploring the relationship between the way in which users navigate in immersive content and the predictability of their trajectories. Inspired by works from social science, our key assumption is that there are navigation trajectories that can be accurately predicted, while others exhibit eclectic patterns that are more challenging to anticipate. However, it is not yet clear how to effectively distinguish between these behaviours. In this context, we conduct an extensive data analysis across multiple datasets investigating users' movements in VR. The ultimate goal is to understand if a specific metric from information theory, such as the entropy of trajectory, can be adopted as a discriminating metric between predictable navigation trajectories and unpredictable ones. Our findings reveal that users with highly regular navigation styles tend to exhibit lower entropy, indicating higher predictability of their movements. Conversely, users with more diverse navigation patterns show higher entropy and lower predictability in their trajectories. Answering the question "how can we distinguish users more predictable than others?" would be crucial for different purposes in future immersive applications such as enabling new modalities for live streaming services but also for the design of more personalised and engaging VR experiences.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Virtual reality**; • **Information systems** → *Multimedia streaming*.

KEYWORDS

User Behavioural Analysis, Virtual Reality, Trajectory analysis, Information Theory, Motion Prediction

ACM Reference Format:
Silvia Rossi, Laura Toni, and Pablo Cesar. 2023. Correlation between Entropy and Prediction Error in VR Head Motion Trajectories. In *Proceedings of the 2nd International Workshop on Interactive eXtended Reality (IXR '23), October 29, 2023, Ottawa, ON, Canada.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3607546.3616805

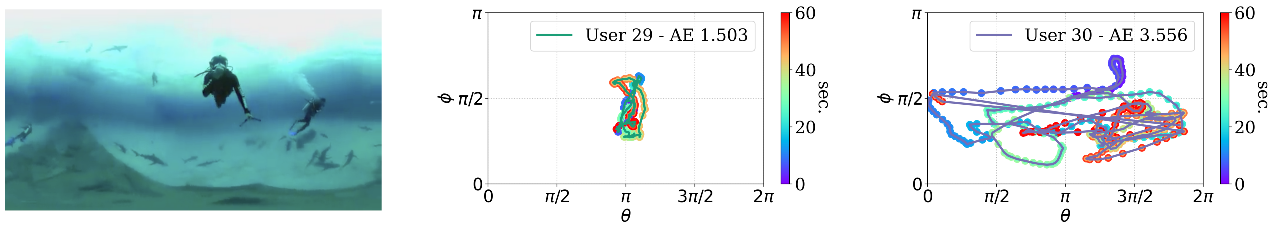
1 INTRODUCTION

Immersive technologies, such as Virtual Reality (VR), are envisioned to lead the next generation virtual worlds [13]. The objective is to offer a compelling user experience, going beyond the passive nature of traditional video and providing higher degrees of immersion and interaction. In particular, VR technology replaces the real world with a fully digital environment, typically a 360° or spherical video, in which people are immersed and feel present. To mimic a real-life scenario, only a restricted portion of the virtual environment is displayed based on the viewers' head movements. This technology has been available for a number of years and has already found numerous applications in real-world use cases, including healthcare, entertainment, and education [5, 21, 34]. However, there are still challenges that need to be overcome. For example, a critical open problem is the ability to predict users' navigation trajectories (*i.e.*, user behaviour) within the virtual space. Being able to anticipate viewers' movement is essential to ensure high-quality content and smooth navigation during the immersive experience. For instance, in a tile-based adaptive streaming scenario, each user receives at high quality only tiles that overlap the predicted displayed portion of the content [28]. This strategy, while effective from both a bandwidth and quality perspective, strongly depends on the performance of the selected prediction algorithm. An erroneous estimate would immediately lead to re-transmissions, and hence, a possible stall or quality reduction effect. While new learning models have been proposed [3, 12, 14, 16, 24], the data analysis and exploration have been overlooking, limiting the hypothesis and understanding of users' behaviour, which is crucial for enhancing the prediction process [27, 35].

The analysis of human navigation trajectories in the 3D space is a multidisciplinary challenge. For example, in social sciences and transportation research, trajectory analysis plays a significant role in several applications such as traffic control, route optimisation, and personalised advertisements but also to prevent the spread of epidemics [9, 40]. In this context, for many years, a fundamental research question has been related to the understanding of the predictability of mobility trajectory. As a pillar solution to this problem, an entropy-based metric has been proposed, able to capture the

IXR '23, October 29, 2023, Ottawa, ON, Canada

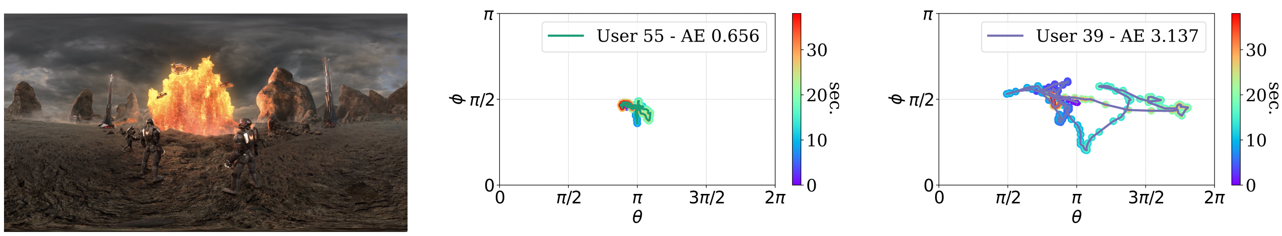
ACM ISBN 979-8-4007-0280-8/23/10.
https://doi.org/10.1145/3607546.3616805



(a) Video "Diving" from 2_MMsys2017_2 [17]

(b) Low Entropy User in "Diving"

(c) High Entropy User in "Diving"



(d) "AcerPredator" from 6_PAMI2018 [38]

(e) Low entropy user in "AcerPredator"

(f) High Entropy User in "AcerPredator"

Figure 1: Comparison of navigation trajectories with different levels of Actual Entropy (AE in the legend). Each row corresponds to two videos selected from two distinct datasets. In the middle column, the navigation of users is characterised by low entropy values while high entropy users are given in the right column.

variability of human navigation behaviours, and thus estimate their predictability [32]. This metric, namely *actual entropy*, measures the information carried within a given trajectory, considering both the visiting rate but also the temporal order of visited areas. The information captured by the entropy is highly related to the degree of predictability of a variable, with low values of entropy for highly predictable events [11]. In the context of VR, the entropy of navigation trajectories has been applied for general behavioural studies. These preliminary investigations have demonstrated the superiority of actual entropy in detecting general patterns in navigation compared to heatmaps, primarily due to its consideration of the temporal order of navigation positions [29]. More in general, these studies have shown that viewers are typically consistent in their way of navigating: the areas that most likely will be displayed by a user do not depend only on the visual characteristic of the multimedia content but also on the personality, preferences and past history of the specific viewers [29, 30].

In this paper, we want to move a step further and investigate the role of entropy in the predictability of navigation trajectories in the VR domain. Having a holist metric capable to capture in advance the user behaviour and determining whether the viewers' navigation trajectory is more predictable could be crucial for the prediction task at large. For example, such a metric can be used to recognise outliers during the data preparation or to select the most suited prediction models based on the users' profiles. To motivate our intuition, Figure 1 shows the navigation trajectories for two different VR videos of users with different behaviours and thus opposite values of actual entropy (corresponding values of entropy are provided in the caption of each sub-figure). When individuals exhibit a highly regular pattern or limited movements, their actual entropy tends to be small, indicating high predictability in their navigation trajectories. For instance, users in Figure 1 (b,e) explore only a small portion of the content (*i.e.*, central area of the video) and focus on the first scene/object that they detected. Conversely, in the same selected videos, certain viewers exhibit a more eclectic approach to experiencing the VR content, with a tendency to navigate the entire video and not focus on specific areas as shown in Figure 1 (c,f). This behaviour suggests being more challenging to predict and more prone to high prediction errors. In this case, the actual entropy of these users is indeed higher.

To validate our intuition, we explore the correlation between the entropy of navigation trajectories in VR and their prediction error. Our analysis is based on several public collections of VR trajectories (*i.e.*, 7 different datasets). As a predictive tool for navigation movements, we consider a simple yet powerful publicly available algorithm, *deep-position-only baseline* introduced in [24]. By comparing the prediction error per trajectory with the corresponding values of entropy and distinguishing users in two classes of entropy (*i.e.*, low and high), our results confirm our initial hypothesis. Navigation trajectories characterised by low values of entropy led to small prediction errors, and thus are easier to be predicted. While users more eclectic and characterised by high values of entropy have on average high values of prediction error. Thanks to its ability to detect discontinuity and randomness in trajectories, the actual entropy seems a promising tool also in the VR context.

To summarise, the main contributions of this work are the following: (1) conducting an extensive data analysis across a wide range of VR datasets; (2) exploring novel criteria that leverage an entropy-based metric to differentiate users based on their navigation profiles, and (3) examination of the correlation between these groups of users and their predictability in navigation. Given the importance of data preparation, augmentation and exploration in every machine learning task, we believe that current works focused only on finding the right deep learning architecture to forecast users' trajectories should be paired with a proper data exploration strategy. Such strategy, including a holistic metric able to capture key behavioural features, is currently missing in the literature and this work aims at moving a step forward to fill in this gap.

2 BACKGROUND

The main objective of this study is to conduct an extensive data analysis across several publicly available VR datasets in order to understand the relationship between the way in which users navigate within the immersive content and the ability to predict their trajectory. We aim also to identify a key metric that can be utilised to improve prediction accuracy and enhance our understanding of user behaviour in immersive virtual environments. To ensure clarity and consistency, we introduce the following notation. VR dataset collects navigation trajectories of a set of user $\mathcal{U}$ who displayed a set of 360° videos $\mathcal{V}$. Viewers are provided by a VR device – typically a head-mounted display (HMD), that allows changing viewport according to their viewing direction. Therefore, the sequence of spatio-temporal points representing over time the user's viewing direction identifies the navigation within the immersive content [26]. Formally, for a given video $v \in \mathcal{V}$ of duration T seconds the VR trajectory of users $u \in \mathcal{U}$ can be denoted as $\mathbf{X_u^v} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T]$, where $\mathbf{x_t}$ denotes the 3D spatial coordinates of the viewing direction at time t. This direction can be either approximated by the head or gaze position in the VR settings. This paper focuses on the analysis of datasets containing head-motion trajectories. Thus in our scenario, $\mathbf{X}$ stores the head coordinates which can be represented in different formats, *e.g.*, Quaternion, Euler angles or Cartesian coordinates. In the case of Cartesian coordinates, $\mathbf{x}_t$ is equal to the tuple (θ_t, ϕ_t) with $0 \leq \theta < 2\pi$ and $0 \leq \phi \leq \pi$. To be compliant with most of the behavioural analysis tools, the spherical content can be also quantized into N regular block, each one with an assigned ID value $B = [b_1, b_2, .., b_N]$. Thus, the user trajectory can be also represented as the temporal sequences of the blocks to which $\mathbf{x}$ belongs at time t.

In the following, we define key concepts required to understand the data analysis that we conduct in the next section. First, we introduce the entropy-based metric (*i.e.*, *actual entropy*) considered in our investigations, and how it has been applied until in the VR context. Then, after a brief overview of the main current solutions for predicting users' trajectories in VR, we summarise the predictive model (*i.e.*, *deep-position-only baseline*) that we selected in our analysis as a predictive model.

2.1 Entropy in User Navigation Analysis

Information-Theoretic (IT) metrics have been applied to a wide range of disciplines, becoming a de-facto statistical tool for data analysis in several fields such as physics, computer science, and neuroscience [33]. Specifically, the concept of *actual entropy* has been introduced as a metric to quantify the variability or predictability in human mobility behaviour by Song et al. [32]. Since the information captured by the entropy is closely related to the predictability of a variable, authors exploited this correlation by measuring the information carried within a given trajectory to estimate the predictability of human mobility as main novelty they consider both the visiting rate but also the temporal order of visited areas. More formally, the actual entropy can be estimated from the past history of the user's trajectory by Lempel-Ziv compression algorithm [41]. Let $\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_T}]$ be a trajectory of positional points in a discredited space ($x_t = b_t$ with $b \in B$ as introduced at the beginning of Section 2), and let $L_t = [x_t, x_{t+1}, \ldots x_{(t-1)+\lambda_t}]$ be a sub-sequence of X starting at time t and spanning λ_t time-slots, the actual entropy assumes the following form:

$$H^{act}(X) \approx \left(\frac{1}{n} \sum_{t=1}^{n} \lambda_t \right)^{-1} \log_2(n) \tag{1}$$

where λ_t is the length of the shortest sub-sequence in X starting at time-slot t and not appearing between time 1 and $t-1$. To be noted, the actual entropy of a given trajectory X is inversely proportional to the sum of the shortest non-repeated sub-sequences within it (*i.e.*, $\sum_{t=1}^{n} \lambda_t$). The original metric proposed in [32] has been extended to other aspects of human behaviour, including social interaction [8, 31], modelling the global spread of emerging diseases [6, 36] but also to the context of behavioural analysis in VR. Specifically, this metric has been applied to characterise the navigation of each subject individually looking for patterns over time and across different contents [29]. Recently, navigation within more challenging VR system, such as while displaying volumetric content, has also been studied through this metric to investigate the consistency of users [30]. Thanks to its versatility and successful application across diverse domains, we find the entropy of trajectory to be particularly appealing also for improving the accuracy in the prediction of VR navigation movements or their classification.

2.2 Navigation Trajectory Forecasting

The prediction of head movements in VR content has been studied since the beginning of this technology [4]. The first and among the simplest techniques were based on the past and current trajectory of a single user by simple linear regression techniques, neglecting other viewers and video content information [1, 22] Thanks to significant advancements in the field of machine learning and deep learning, these techniques found also widespread application in the VR context. For example, clustering approaches have been applied: a novel viewport prediction algorithm based on a graph-based clustering defined specifically for immersive content [25] has been proposed in [19], while authors in [18] presented a hybrid clustering approach which also takes into account the video content. Similarly, other deep learning frameworks have been augmented with the use of saliency maps [20, 24, 38]. Recently, a transformer method has been proposed as viewport prediction for 360° videos taking advantage of this novel architecture able to capture long-term dependencies and hidden patterns in the data [2]. However, the focus of this work is not to advance the state of the art in predicting head movement while experiencing VR content. Our goal is to investigate any existing correlation between the entropy of navigation trajectories in VR with the ability to predict such movements. Thus, we chose to consider a simple yet robust prediction baseline, known as the *deep-position-only baseline*, also thanks to the reproducible resources presented in [24]. Specifically, this is a simple sequence-to-sequence LSTM-based architecture that disregards the video content and takes as input only the sequence of past and current positions of the users. Authors in [24] have shown the ability

Table 1: Key features of the VR navigation datasets analysed in this work.

Name Dataset	# video	length	# participants	# training | test
1_MMsys2017_1 [7]	5	70*s.*	57	3 | 2
2_MMsys2017_2 [17]	10	60*s.*	25	8 | 2
3_MMsys2017_3 [37]	9	30-45*s.*	48	5 | 4
4_CVPR2018 [39]	208	20-60*s.*	34	134 | 74
5_MMSys2018 [10]	76	20*s.*	58	61 | 15
6_PAMI2018 [38]	19	10-80*s.*	57	14 | 5
7_MM2022 [15]	27	60*s.*	100	21 | 6

of this framework to overcome traditional baselines (*e.g.*, copy last position or linear regression) but also other current start-of-the-art deep learning architectures, even those with additional inputs such as meta-data and/or video content information. In the following, we consider the same setting for the prediction problem as presented in [24]. Specifically, we start the prediction phase after a window of 6 seconds to avoid the initial exploration phase typical of VR viewers. Finally, we set the prediction horizon and past history window equal to 5 seconds and 1 second, respectively.

3 DATA ANALYSIS

This section describes the set of datasets of VR navigation trajectories on which we based our data analysis. Then, we show the results of our in-depth data analysis aimed at understanding how the way to explore video content by users affects their prediction.

3.1 User Navigation Datasets

Our extensive data analysis is based on 7 publicly available datasets which gather head motion trajectories in 360° videos. Table 1 outlines the key information per dataset, which is described in detail in the following:

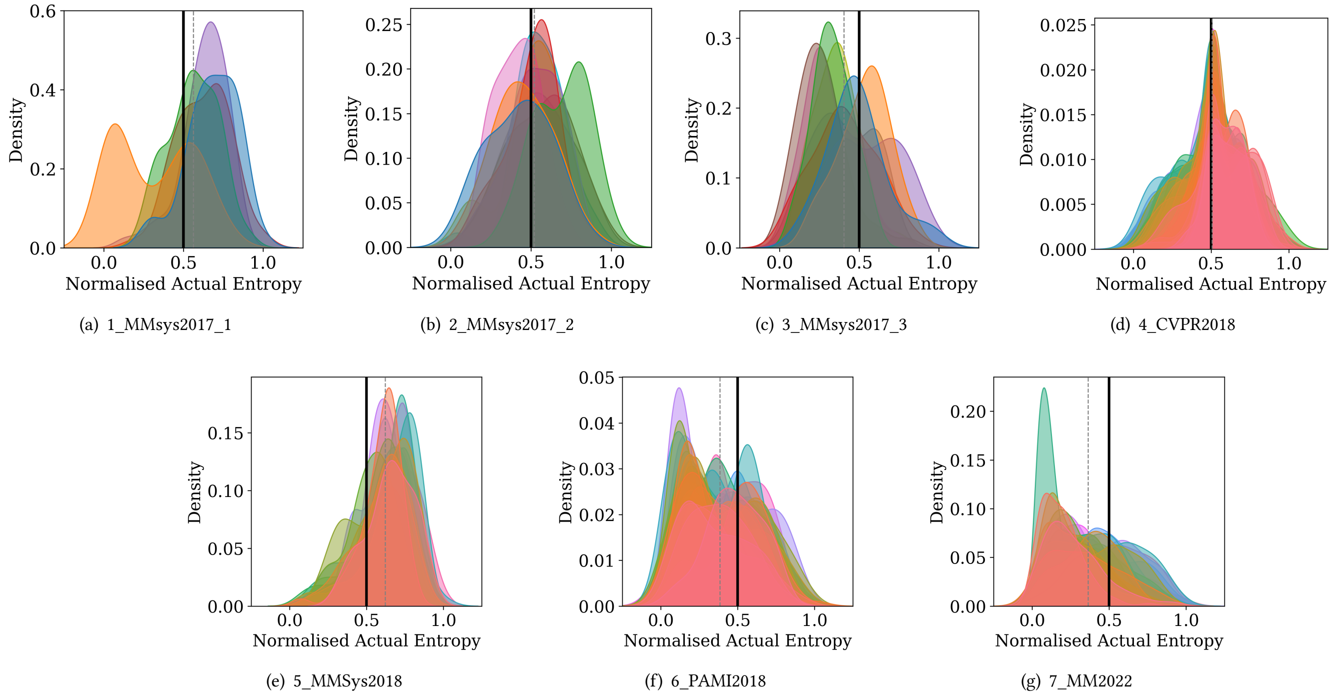
- 1_MMsys2017_1 [7]: this dataset collects navigation trajectories of 57 users who navigated within 5 sequences of the same duration (70 seconds);
- 2_MMsys2017_2 [17]: this collection is composed of 10 videos with a length of 60 seconds. Originally, the trajectories of 50 participants have been collected but only half of them have been used in the further experiments of prediction presented in [12] and in [23];
- 3_MMsys2017_3 [37]: the original dataset is composed of 18 videos displayed by 48 users. However, the first set of content has been experienced in a free-navigation experiment to identify the natural behaviour, while the second one is more specific to capture the user's attention on the video content with structured questionnaires. In our study, we consider only the first group of data;
- 4_CVPR2018 [39]: this collection has the highest number of VR content, *i.e.*, 208, characterised by a variable length ranging from 20 to 60 seconds (36 seconds on average). However, each video have been displayed only by 34 participants (at least 31 per each video);
- 5_MMSys2018 [10]: a collection of 57 navigation trajectories is presented for 19 different content. These videos, however, are short with a fixed duration of only 20 seconds,
- 6_PAMI2018 [38]: this dataset is composed of head-motion trajectories collected by 58 viewers in 76 different VR content. In this case, the videos have a variable length (*i.e.*, between 10 and 80 seconds, on average of 25 seconds),
- 7_MM2022 [15]: the most recent VR collection is characterised by the highest number of participants (*i.e.*, 100) who displayed 27 different 360° videos, each of 60 seconds.

These datasets are part of the unified collection presented in [23]. We also include one of the first VR navigation datasets (*i.e.*, 1_MMsys2017_1) and the most recent one (*i.e.*, 7_MM2022). In Table 1, the number of videos included in the training and test sets is also given. Specifically, for most of the dataset we use the same splitting as adopted in [24], while for the newly added ones (*i.e.*, 1_MMsys2017_1 and 7_MM2022), we build the training set by selecting uniformly at random 80% of the videos, leaving the remaining 20% for testing (ensuring to have at least two videos in the test set). Additionally, there is no overlap between the videos in the train and test sets.

3.2 Results

As first step of our data analysis, we explore the distribution of actual entropy for each navigation across the different datasets. Given that actual entropy ranges from 0 to infinity, to enable a fair comparison on the same scale, we normalise this metric per each individual dataset. Figure 2 shows the distributions of normalised actual entropy per video in all the selected datasets. In each subplot, the mean value of the given distribution is also represented with a grey dotted line. It can be noted that few datasets, namely 1_MMsys2017_1, 2_MMsys2017_2 and 4_CVPR2018 have a distribution with an average value very close to or equal to 0.5 represented in the figure with a black line. The remaining datasets have either a higher mean value of actual entropy (5_MMSys2018) or smaller (3_MMsys2017_3, 6_PAMI2018 and 7_MM2022) than 0.5. In the following, we choose to not aggregate the navigation trajectories from the different datasets. Instead, we continue our data analysis by considering each dataset separately. Further investigations are needed to establish a global normalisation of actual entropy across all the datasets and we leave them for future works.

The core of our data analysis is to verify if there is any relationship between the complexity of the navigation trajectories (measured by the entropy of the trajectory) and their ability to be predicted. To do so, the actual entropy per user is compared with the accuracy of their prediction. Since users are typically consistent in their way of navigating within immersive content [29, 30],



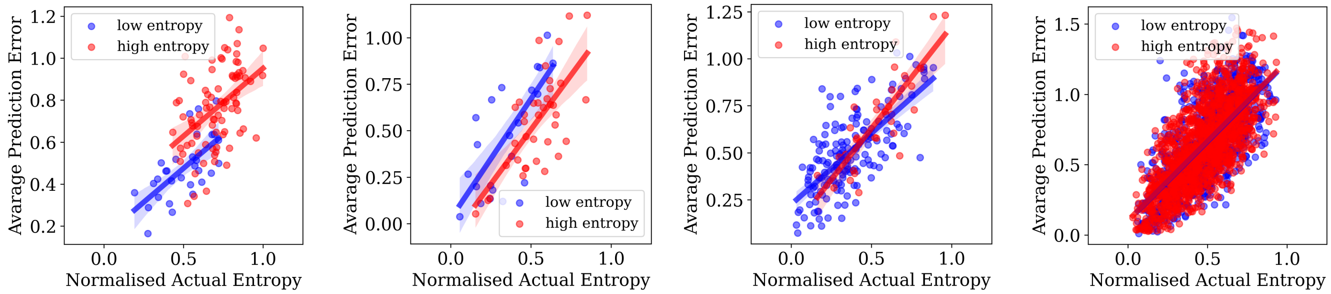
(a) 1_MMsys2017_1 (b) 2_MMsys2017_2 (c) 3_MMsys2017_3 (d) 4_CVPR2018

(e) 5_MMSys2018 (f) 6_PAMI2018 (g) 7_MM2022

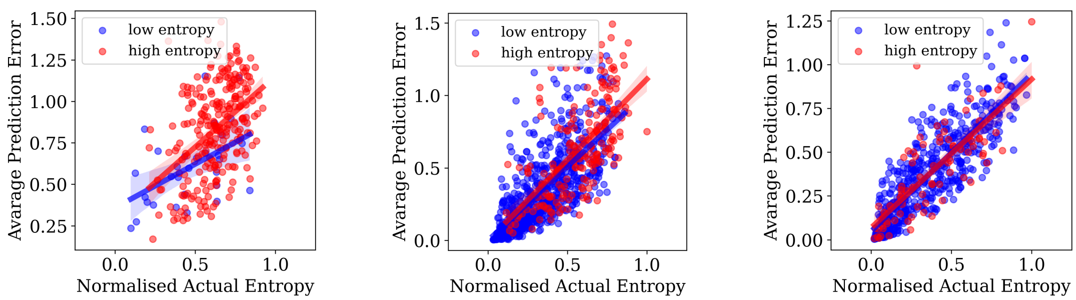
Figure 2: Density distribution of actual entropy per video in all the selected datasets. The entropy is normalised per single dataset. The black line represent the threshold that will be used in the following while the grey dotted line is the averaged value across each dataset.

we create two classes of users based on their mean value of actual entropy in each dataset. To be as generic as possible, we use a threshold of 0.5 of the normalised actual entropy: users with a mean value of entropy above 0.5 are considered to belong to the "high entropy" category; while those with an averaged value below 0.5 are classified as "low entropy" users. In terms of the prediction model, as introduced in Section 2.2, we use *deep-position-only baseline* to predict the navigation trajectories. For evaluating the final prediction, we choose the orthodromic distance as a metric for prediction error following the evaluation done in [24]. For each video in the test set, Figure 3 shows the scatterplot between the actual entropy and average prediction error for the aforementioned categories of users. The presence of a linear correlation between these variables is quite evident in all the analysed datasets: when the actual entropy increases, the averaged prediction error also grows, indicating that more predictable trajectories have lower entropy. Our initial intuition is therefore confirmed. Similar observations can be done by looking at the two categories of users. The regression line for "low entropy" users (*i.e.*, blue line in Figure 3) in most of the case is below the one for the "high entropy" category (*i.e.*, red line). This means that the prediction error of users with "high entropy" is on average higher, indicating that this category is more difficult to be correctly predicted with high precision. Moreover, in Figure 3 (c) and (e), the slope of the red line is steeper than the blue one: as the entropy value increases, the prediction error increases more quickly for "high entropy" users than for the other group. However, further investigation are needed since there are exceptions of this general behaviour in 2_MMsys2017_2, 4_CVPR2018 and 7_MM2022. In the latter two cases, the regression lines for the two groups are actually equivalent both in terms of slope and range of values. To be noted as reported in Table 1, these two datasets have the highest number of videos and users, respectively. On the contrary, Figure 3 (b) shows an opposite trend: the regression line of "high entropy" users is below the one for the "low entropy" category. In this particular dataset, we have only two videos in the test set and the lowest number of navigation trajectories as reported in Table 1. Thus, this contrasting behaviour could be attributed to the limited amount of data available for the analysis in 2_MMsys2017_2.

As a final analysis, we go more in detail for two selected content such as "Diving" from 1_MMsys2017_1 and "Guitar" from 6_PAMI2018. Figure 4 shows per each participant of the datasets the distribution of their prediction error over time. The box of each user is coloured according to the belonging class of entropy, and users are also arranged in ascending order based on their mean value of actual entropy across the entire dataset, with lower values on the left and higher on the right side. Finally, the black line represents the mean value of prediction error across all viewers in the two selected content. The analysis reveals a distinct pattern and confirms the previous observations: the prediction error is consistently below the average for users in the "low entropy" category, while it tends to be above the average for users in the "high entropy" category. This trend becomes especially apparent in the extreme cases where participants are characterised by the lowest and highest values of entropy. However, some outlier behaviour can be detected,



(a) 1_MMsys2017_1 (b) 2_MMsys2017_2 (c) 3_MMsys2017_3 (d) 4_CVPR2018



(e) 5_MMSys2018 (f) 6_PAMI2018 (g) 7_MM2022

Figure 3: Scatterplot of the prediction error versus actual entropy per user in each dataset. The entropy is normalised per dataset. Users are also divided in classes of "low" and "high entropy" based on their mean value of entropy in the entire dataset.

especially for users who fall in between the two defined classes of entropy. For example, the participant with ID 51 Figure 4 (a) has a prediction error above the average even if classified in the category of "low entropy"; on the contrary, user 08 in Figure 4 (b) belongs to the group of "high entropy" but is predicted with a quite small error over time. This observation suggests that certain users may exhibit less consistency in their way of navigating, leading them to be more static or dynamic depending on the content. Introducing an additional category of users (*i.e.*, "medium entropy") could better capture this inconsistency. However, further investigations regarding the classification of users based on their entropy will be explored in future works.

4 CONCLUSION

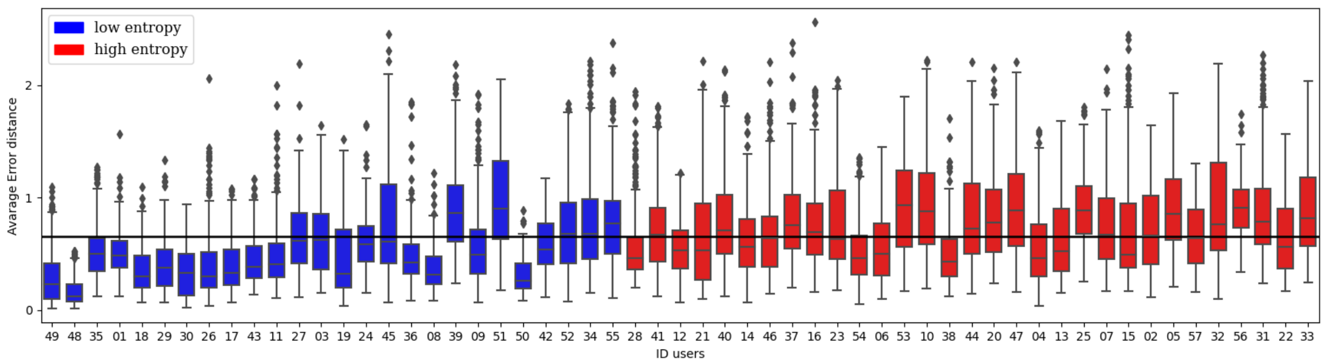
In this paper, we presented an extensive data analysis across several publicly available VR datasets in order to understand correlation between users' head motion and trajectory predictability. To do so, we had to identify a key metric that can be utilised to improve prediction accuracy and enhance our understanding of user behaviour in immersive virtual environments. Our results have confirmed a consistent correlation between the entropy of VR trajectories and their prediction error. Individuals who show a highly regular style of navigating tend to have a low entropy of their trajectory, indicating thus a more predictable nature in their navigation. On the other hand, users with a high value of entropy present less predictability in their movements. However, our findings also highlight some limitations on the applicability of this entropy-based metric that needs further investigation. To be a more generic metric, applicable across different VR datasets and capable of generating meaningful user classes based on their level of entropy, additional research and refinement are necessary. This will be the focus of our future work. We will also enhance existing predictive algorithms by incorporating this metric to improve their performance and unlock the full potential of entropy-based metrics in immersive technology.

ACKNOWLEDGMENTS

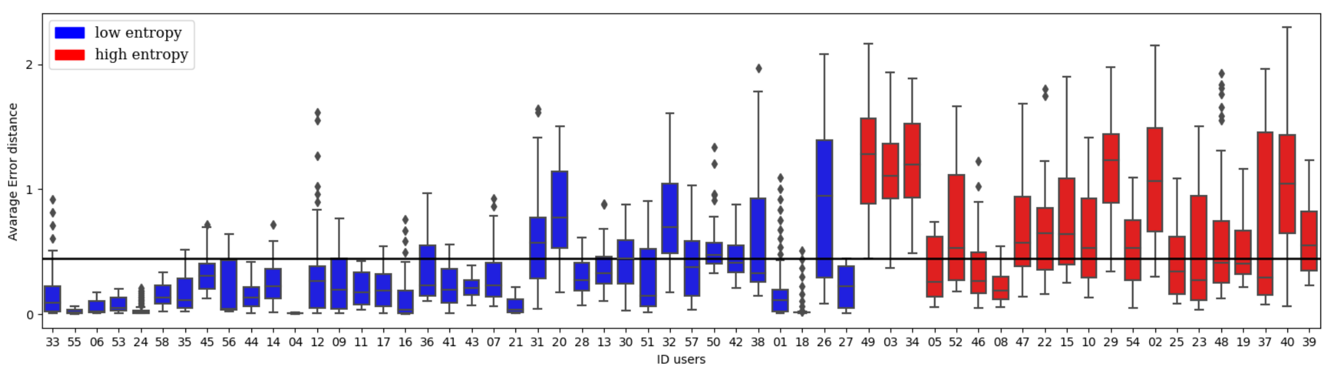
This work was partially carried out during the tenure of an ERCIM "Alain Bensoussan" Fellowship Programme and in part supported through the European Commission Horizon Europe program under grant 101070109 TRANSMIXR (https:// transmixr.eu/).

REFERENCES

[1] Yanan Bao, Huasen Wu, Tianxiao Zhang, Albara Ah Ramli, and Xin Liu. 2016. Shooting a moving target: Motion-prediction-based transmission for 360-degree videos. In *International Conference on Big Data*. IEEE. https://doi.org/10.1109/BigData.2016.7840720
[2] Fang-Yi Chao, Cagri Ozcinar, and Aljosa Smolic. 2021. Transformer-based Long-Term Viewport Prediction in 360° Video: Scanpath is All You Need. In *International Workshop on Multimedia Signal Processing*. IEEE. https://doi.org/10.1109/MMSP53017.2021.9733647
[3] Fang-Yi Chao, Cagri Ozcinar, and Aljosa Smolic. 2022. Privacy-Preserving Viewport Prediction using Federated Learning for 360° Live Video Streaming. In *IEEE 24th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE. https://doi.org/10.1109/MMSP55362.2022.9950044
[4] Federico Chiariotti. 2021. A survey on 360-degree video: Coding, quality of experience and streaming. *Computer Communications* 177 (2021). https://doi.org/10.1016/j.comcom.2021.06.029
[5] Pietro Cipresso, Irene Alice Chicchi Giglioli, Mariano Alcañiz Raya, and Giuseppe Riva. 2018. The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. *Frontiers in psychology* (2018). https://doi.org/10.3389/fpsyg.2018.02086



(a) "Diving" from 1_MMsys2017_1



(b) "Guitar" from 6_PAMI2018

Figure 4: Distribution of prediction error per user in two different content. The users are arranged in ascending order based on their mean value of actual entropy across the entire dataset, with lower values on the left and higher values on the right. The black line shows the mean value of prediction error across all viewers in the selected video.

[6] Vittoria Colizza, Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. 2006. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences* (2006). https://doi.org/10.1073/pnas.0510525103

[7] Xavier Corbillon, Francesca De Simone, and Gwendal Simon. 2017. 360-degree video head movement dataset. In *Proceedings of ACM Multimedia Systems Conference*. 199–204. https://doi.org/10.1145/3193701.

[8] Justin Cranshaw, Eran Toch, Jason Hong, Aniket Kittur, and Norman Sadeh. 2010. Bridging the gap between physical location and online social networks. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. https://doi.org/10.1145/1864349.1864380

[9] Andrea Cuttone, Sune Lehmann, and Marta C González. 2018. Understanding predictability and exploration in human mobility. *EPJ Data Science* (2018). https://doi.org/10.1140/epjds/s13688-017-0129-1

[10] Erwan J. David, Jesús Gutiérrez, Antoine Coutrot, Matthieu Perreira Da Silva, and Patrick Le Callet. 2018. A dataset of head and eye movements for 360° videos. In *Proceedings of ACM Multimedia Systems Conference*. ACM. https://doi.org/10.1145/3204949.3208139

[11] Douglas do Couto Teixeira, Jussara M Almeida, and Aline Carneiro Viana. 2021. On estimating the predictability of human mobility: the role of routine. *EPJ Data Science* (2021). https://doi.org/10.1140/epjds/s13688-021-00304-8

[12] Ching-Ling Fan, Jean Lee, Wen-Chih Lo, Chun-Ying Huang, Kuan-Ta Chen, and Cheng-Hsin Hsu. 2017. Fixation Prediction for 360 Video Streaming in Head-Mounted Virtual Reality. In *Proceedings of ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*. https://doi.org/10.1145/3083165.3083180

[13] Torres I. Hupont, V. Charisi, G. De Prato, K. Pogorzelska, S. Schade, A. Kotsev, M. Sobolewski, N. Duch Brown, E. Calza, C. Dunker, F. Di Girolamo, M. Bellia, J. Hledik, I. Nai Fovino, and M. Vespe. 2023. Next Generation Virtual Worlds: Societal, Technological, Economic and Policy Challenges for the EU. *Publications Office of the European Union* (2023). https://doi.org/10.2760/51579

[14] Gazi Karam Illahi, Ashutosh Vaishnav, Teemu Kämäräinen, Matti Siekkinen, and Mario Di Francesco. 2023. Learning to Predict Head Pose in Remotely-Rendered Virtual Reality. In *Proceedings of the 14th Conference on ACM Multimedia Systems*. 27–38. https://doi.org/10.1145/3587819.3590972

[15] Yili Jin, Junhua Liu, Fangxin Wang, and Shuguang Cui. 2022. Where Are You Looking? A Large-Scale Dataset of Head and Gaze Behavior for 360-Degree Videos and a Pilot Study. In *Proceedings of ACM International Conference on Multimedia*. https://doi.org/10.1145/3503161.3548200

[16] Mu Li, Kanglong Fan, and Kede Ma. 2023. Scanpath Prediction in Panoramic Videos via Expected Code Length Minimization. *arXiv preprint arXiv:2305.02536* (2023).

[17] Wen-Chih Lo, Ching-Ling Fan, Jean Lee, Chun-Ying Huang, Kuan-Ta Chen, and Cheng-Hsin Hsu. 2017. 360° video viewing dataset in head-mounted virtual reality. In *Proceedings of ACM Multimedia Systems Conference*. https://doi.org/10.1145/3192927

[18] Dario D. R. Morais, Lucas S. Althoff, Ravi Prakash, Marcelo M. Carvalho, and Myléne C.Q. Farias. 2021. A Content-Based Viewport Prediction Model. In *Electronic Imaging*. https://doi.org/10.2352/ISSN.2470-1173.2021.9.IQSP-255

[19] Afshin Taghavi Nasrabadi, Aliehsan Samiei, and Ravi Prakash. 2020. Viewport prediction for 360° videos: a clustering approach. In *Proceedings of ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*. https://doi.org/10.1145/3386290.3396934

[20] Anh Nguyen, Zhisheng Yan, and Klara Nahrstedt. 2018. Your Attention is Unique: Detecting 360-Degree Video Saliency in Head-Mounted Display for Head Movement Prediction. In *Proceedings of ACM International Conference on Multimedia*. https://doi.org/10.1145/3240508.3240669

[21] Sara Qamar, Zahid Anwar, and Mehreen Afzal. 2023. A systematic threat analysis and defense strategies for the metaverse and extended reality systems. *Computers & Security* (2023). https://doi.org/10.1016/j.cose.2023.103127

[22] Feng Qian, Lusheng Ji, Bo Han, and Vijay Gopalakrishnan. 2016. Optimizing 360 video delivery over cellular networks. In *Proceedings of ACM Workshop on All Things Cellular: Operations, Applications and Challenges*. https://doi.org/10.1145/2980055.2980056

[23] Miguel Fabián Romero Rondón, Lucile Sassatelli, Ramon Aparicio-Pardo, and Frédéric Precioso. 2020. A unified evaluation framework for head motion prediction methods in 360° videos. In *Proceedings of ACM Multimedia Systems Conference*. https://doi.org/10.1145/3339825.3394934

[24] Miguel Fabian Romero Rondon, Lucile Sassatelli, Ramon Aparicio-Pardo, and Frederic Precioso. 2022. TRACK: A New Method From a Re-Examination of Deep Architectures for Head Motion Prediction in 360° Videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022). https://doi.org/10.1109/TPAMI.2021.3070520

[25] Silvia Rossi, Francesca De Simone, Pascal Frossard, and Laura Toni. 2019. Spherical clustering of users navigating 360° content. In *International Conference on Acoustics, Speech and Signal Processing*. IEEE. https://doi.org/10.1109/ICASSP.2019.8683854

[26] Silvia Rossi, Alan Guedes, and Laura Toni. 2022. *Coding, Streaming, and User Behaviour in Omnidirectional Video*. Elsevier.

[27] Silvia Rossi, Cagri Ozcinar, Aljosa Smolic, and Laura Toni. 2020. Do users behave similarly in VR? Investigation of the user influence on the system design. *ACM Transactions on Multimedia Computing, Communications, and Applications* (2020). https://doi.org/10.1145/3381846

[28] Silvia Rossi and Laura Toni. 2017. Navigation-aware adaptive streaming strategies for omnidirectional video. In *IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE. https://doi.org/10.1109/MMSP.2017.8122230

[29] Silvia Rossi and Laura Toni. 2020. Understanding user navigation in immersive experience: an information-theoretic analysis. In *Proceedings of the 12th International Workshop on Immersive Mixed and Virtual Environment Systems*. ACM. https://doi.org/10.1145/3386293.3397115

[30] Silvia Rossi, Irene Viola, and Pablo Cesar. 2022. Behavioural Analysis in a 6-DoF VR System: Influence of Content, Quality and User Disposition. In *Proceedings of the 1st Workshop on Interactive eXtended Reality*. https://doi.org/10.1145/3552483.3556454

[31] Filippo Simini, Gianni Barlacchi, Massimilano Luca, and Luca Pappalardo. 2021. A deep gravity model for mobility flows generation. *Nature communications* (2021). https://doi.org/10.1038/s41467-021-26752-4

[32] Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. 2010. Limits of predictability in human mobility. *Science* (2010). https://doi.org/10.1126/science.1177170

[33] Nicholas M Timme and Christopher Lapish. 2018. A tutorial for information theory in neuroscience. *eNeuro* (2018). https://doi.org/10.1523/ENEURO.0052-18.2018

[34] Jeroen van der Hooft, Hadi Amirpour, Maria Torres Vega, Yago Sanchez, Raimund Schatz, Thomas Schierl, and Christian Timmerer. 2023. A Tutorial on Immersive Video Delivery: From Omnidirectional Video to Holography. *IEEE Communications Surveys & Tutorials* (2023).

[35] Jeroen Van Der Hooft, Tim Wauters, Filip De Turck, Christian Timmerer, and Hermann Hellwagner. 2019. Towards 6DoF HTTP adaptive streaming through point cloud compression. In *Proceedings of ACM International Conference on Multimedia*. https://doi.org/10.1145/3343031.3350917

[36] Lin Wang and Joseph T Wu. 2018. Characterizing the dynamics underlying global spread of epidemics. *Nature communications* (2018). https://doi.org/10.1038/s41467-017-02344-z

[37] Chenglei Wu, Zhihao Tan, Zhi Wang, and Shiqiang Yang. 2017. A dataset for exploring user behaviors in VR spherical video streaming. In *Proceedings of ACM Multimedia Systems Conference*. ACM. https://doi.org/10.1145/3192423

[38] Mai Xu, Yuhang Song, Jianyi Wang, MingLang Qiao, Liangyu Huo, and Zulin Wang. 2019. Predicting head movement in panoramic video: A deep reinforcement learning approach. *IEEE Transactions on pattern analysis and machine intelligence* (2019). https://doi.org/10.1109/TPAMI.2018.2858783

[39] Yanyu Xu, Yanbing Dong, Junru Wu, Zhengzhong Sun, Zhiru Shi, Jingyi Yu, and Shenghua Gao. 2018. Gaze prediction in dynamic 360 immersive videos. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*. IEEE. https://doi.org/10.1109/CVPR.2018.00559

[40] Chao Zhang, Kai Zhao, and Meng Chen. 2022. Beyond the limits of predictability in human mobility prediction: context-transition predictability. *IEEE Transactions on Knowledge and Data Engineering* (2022). https://doi.org/10.1109/TKDE.2022.3148300

[41] Jacob Ziv and Abraham Lempel. 1978. Compression of individual sequences via variable-rate coding. *IEEE Transactions on Information Theory* (1978). https://doi.org/10.1109/TIT.1978.1055934