Predictive Context-Awareness for Full-Immersive Multiuser Virtual Reality with Redirected Walking

Filip Lemic, Jakob Struye, Thomas Van Onsem, Jeroen Famaey, Xavier Costa-Pérez

***Abstract*—The advancement of Virtual Reality (VR) technology is focused on improving its immersiveness, supporting multiuser Virtual Experiences (VEs), and enabling the users to move freely within their VEs while remaining confined to specialized VR setups through Redirected Walking (RDW). To meet their extreme data-rate and latency requirements, future VR systems will require supporting wireless networking infrastructures operating in millimeter Wave (mmWave) frequencies that leverage highly directional communication in both transmission and reception through beamforming and beamsteering. We propose the use of predictive context-awareness to optimize transmitter and receiver-side beamforming and beamsteering. By predicting the users' short-term lateral movements in multiuser VR setups with Redirected Walking (RDW), transmitter-side beamforming and beamsteering can be optimized through Line-of-Sight (LoS) "tracking" in the users' directions. At the same time, predictions of short-term orientational movements can be utilized for receiver-side beamforming for coverage flexibility enhancements. We target two open problems in predicting these two context information instances: i) predicting lateral movements in multiuser VR settings with RDW, and ii) generating synthetic head rotation datasets for training orientational movements predictors. Our experimental results indicate that Long Short-Term Memory (LSTM) networks feature promising accuracy in predicting lateral movements, and context-awareness stemming from VEs further enhances this accuracy. Additionally, we show that a TimeGAN-based approach for orientational data generation can create synthetic samples that closely match experimentally obtained ones.**

***Index Terms*—Full-immersive multiuser Virtual Reality, predictive context-awareness, Recurrent Neural Network, Generative Adversarial Network, redirected walking**

I. Introduction

The utilization of Virtual Reality (VR) technology is transforming digital experiences and interactions of various communities [1]. To improve the immersiveness of Virtual Experiences (VEs), VR setups and content are continually being upgraded. Research efforts are primarily focused on enhancing the quality of VEs provided to the users [2], and facilitating its wireless delivery without mobility constraints, also known as "cutting the wire" [3]. Additionally, enabling multiuser experiences that allow the users to collaborate and have their actions affect the VEs of others is an important goal [4].

In the future, VR systems will have the capability to accommodate multiple users who can fully engage in immersive VEs without being limited by mobility. This advanced functionality will be made possible by high-frequency wireless networks primarily operating in the millimeter Wave (mmWave) band, ranging from 30 to 300 GHz [5]. To provide mobile VR users with high-quality content in real-time, wireless communication that supports these systems will need to be highly directional in both transmission and reception [6].

Directional mmWave beams will follow the users' movements for maintaining Line-of-Sight (LoS) connectivity with them during transmission. Meanwhile, Redirected Walking (RDW) will be utilized to prevent physical collisions between the users and VR setup boundaries or other users [4]. RDW enables the users to explore VEs freely while subtly redirecting their physical movements for collision avoidance, thus enhancing immersion. Short-term lateral movement prediction can be used to support continuous LoS connectivity, as directional mmWave beams must provide coverage for both current and near-future user locations. This requirement highlights the importance of short-term lateral movement prediction in full-immersive multiuser VR setups with RDW.

Predicting short-term movements in natural human walking is an established research topic, with Long Short-Term Memory (LSTM) networks from the family of Recurrent Neural Networks (RNNs) being particularly effective (e.g., [7]). Although these methods have proven useful for predicting natural walks, neither imperceptible nor perceptible resteering accurately mimics natural walking movements. This indicates a need to assess the suitability of RNNs in predicting VR users' lateral mobility under the constraints of RDW. Despite this, the topic has received relatively little attention in the community. Nevertheless, recent work by [8] has demonstrated that RNNs, particularly LSTM networks, can be applied for this purpose and feature promising levels of accuracy in a single-user setup.

Our work builds upon previous research by incorporating context information from VEs into the prediction. Specifically, we utilize the users' movement trajectory in a VE as an input feature, which differs from existing methods that solely rely on physical movement trajectories. Our experiments evaluate the impact of different numbers of coexisting users and types of VEs on the predictive accuracy. Our results demonstrate that incorporating virtual movement trajectory as an input context significantly enhances the accuracy of the prediction model.

It is envisioned that the users' real-world movements should be accurately reflected in VEs, allowing for seamless changes in gaze direction. To support this requirement, flexible coverage is highly advantageous for receiver-side beamforming on a Head-Mounted Device (HMD). Even a slight beam misalignment can significantly affect the Signal-to-Noise Ratio (SNR) [9], which is why a flexible beam stretching in the head

Filip Lemic and Xavier Costa-Pérez are with the i2Cat Foundation, Spain. Xavier Costa-Pérez is also with NEC Laboratories Europe GmbH, Germany and ICREA, Spain. Email: {name.surname}@i2cat.net.

Jakob Struye, Thomas Van Onsem, and Jeroen Famaey are with the University of Antwerp, Belgium. Jakob Struye and Jeroen Famaey are also with IMEC, Belgium. Email: {name.surname}@uantwerpen.be.

Manuscript received October 31, 2022; revised XXX.

rotation direction can provide the HMD with the consistently high gain necessary for uninterrupted content delivery. This approach ensures that user motion is reflected on-screen within the motion-to-photon latency of 20 ms for avoiding nausea [5]. We argue that accurate prediction of head rotations is needed to proactively form such beams. Existing approaches for such predictions are already highly accurate, as seen in e.g., [10].

Typically, such prediction algorithms rely on Deep Learning (DL) components for transforming the users' orientation data into valuable outputs. However, training, testing, and evaluating these algorithms necessitates collecting vast amounts of orientation data. While these algorithms are the primary consumers of orientation data, other situations also require large orientation data sources, such as receiver-side beamforming and enabling RDW. Collecting these datasets is expensive and laborious, making it challenging to scale. Instead, a more effective approach would be to use synthetic data generation to supplement existing datasets with new samples, adding only minor variations to the overall dataset distribution.

To the best of our knowledge, there is currently only a single study on generating synthetic orientational data, exploring the use of Fast Fourier Transform (FFT) for the generation [11]. In this approach, the input orientation time series is treated as a signal and converted to power spectral densities, and the mean power spectral density is modeled. This method then generates synthetic time series by converting perturbed versions of the power spectral densities back to orientational series. As a result, synthetic series are generated that closely resemble the average value of the input series. Our approach, on the other hand, uses a TimeGAN model from the family of Generative Adversarial Networks (GANs) for generating synthetic orientational data. We demonstrate experimentally the model's beyond state-of-the-art capabilities.

In more general terms, this work advocates the utilization of predictive context-awareness to enhance the performance of mmWave networks that support full-immersive multiuser VR with RDW. Specifically, we demonstrate that predictive context information, such as the users' lateral mobility under the constraints of RDW and their orientational mobility, can be precisely forecasted for a brief period. Furthermore, we explore how these predictive context information instances can be utilized to optimize the performance of supporting mmWave networks. This optimization can occur along two dimensions: transmitter-side beamforming and beamsteering toward the VR users' HMDs, and receiver-side beamforming for coverage enhancements.

II. System Overview

In Figure 1, we showcase a full-immersive multiuser VR setup. Our focus lies on the deployment of this setup within a physically constrained environment that prioritizes user safety while engaging with VEs. This safe perimeter limits potential collision hazards for the users to the environmental boundaries and other users.

RDW is employed to steer the users and ensure collision avoidance between them and the environmental boundaries, as well as among themselves. Its objective is to facilitate user immersion by enabling them to freely explore VEs without constraints, while seamlessly redirecting their movements in the physical space, (ideally) without causing any noticeable disruptions.

The three ways of achieving imperceptible resterring in VEs are: i) curvature gains, which involve VE rotations, ii) translational gains, which modify the users' linear movements to change their travel distances in VEs, and iii) rotational gains that introduce additional rotations to the already rotating users. A promising algorithm for achieving this is the Artificial Potential Field (APF) [4], which generates a force vector that guides the users away from obstacles and scales inversely with the distance from each obstacle, including other users. APF respects empirically determined RDW noticeability thresholds [12], resulting in imperceptible resterring. In case of an imminent collision, the resetting algorithm called Artificial Potential Field Resetting (APF-R) [4] is triggered. The APF-R algorithm calculates the total force vector for determining the angle the users should physically turn toward, followed by instructing a 2:1 turn. During a 2:1 turn, the users' rotational speed increases, allowing them to turn 360° in the VE while turning a smaller computed angle in reality. Our study uses APF and APF-R for imperceptible resterring and resetting in situations where a collision is imminent, respectively.

To ensure optimal Quality of Experience (QoE), VR content is delivered to the users via highly directional mmWave communication. Specifically, an Access Point (AP) transmits focused beams that track the users' movements, thereby continuously maintaining LoS connectivity with each of them. This approach maximizes the link quality and enhances the users' QoE. The process involves the VR headset reporting its location to the AP, which is then used to facilitate both RDW and beamsteering. It is worth noting that modern VR headsets, such as the Vive Cosmos and Oculus Quest 2, already possess the capability to generate and share the physical locations of the HMDs via their built-in sensors and inside-out tracking.

The AP aims at creating beams that cover the current and near-future locations of the users, facilitating LoS maintenance at present and in the upcoming period. Several techniques, including [13], have been proposed in the literature to achieve this goal for transmitter-side beamforming and beamsteering. The beamforming on the receiver side is also expected to adapt to the users' head rotations using the HMDs' built-in sensors that provide accurate orientation estimates. For those interested in this system aspect, in [14] we have introduced coVRage, a receiver beamforming technique that anticipates changes in the Angle of Arrival (AoA) from the AP using past and present orientations as references. Subsequently, the HMD adjusts the beam dynamically to encompass the AoA trajectory.

It should be noted that any possible interruptions in the LoS connectivity between the AP and its users caused by obstructions from other users are beyond the scope of our considerations. One solution to this issue could be to implement an AP handover when there is a prediction that a user's movement will obstruct the LoS path of another user. This emphasizes the need for accurately predicting VR users' short-term movements. Additionally, solutions that utilize Intelligent Reflective Surfaces (IRSs) have also been proposed, e.g., [6].

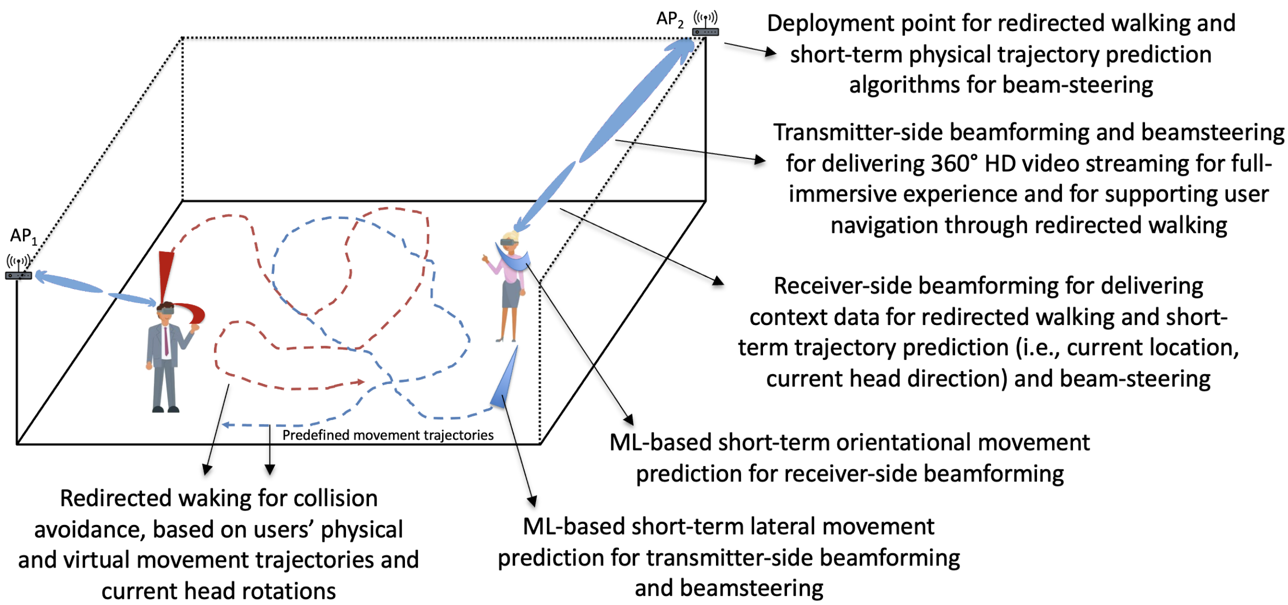


Fig. 1. Considered full-immersive multiuser Virtual Reality setup with Redirected Walking

III. Toward Predictive Context-Awareness

A. Short-Term Lateral Movement Prediction

The preceding discussion suggests that short-term prediction of users' horizontal movements will serve as a tool for RDW, transmitter-side beamsteering, and LoS obstruction avoidance. It is intuitive that the accuracy of such prediction will directly impact the effectiveness of other system components, making optimization crucial for improving the users' QoE.

RNNs have been identified as the most appropriate choice for predicting lateral movements, as mentioned earlier. RNNs are a class of artificial neural networks that consist of multiple neurons of the same kind, each of which passes a message to a succeeding one. This enables RNNs to display dynamic behavior over time, making them well-suited for tasks such as speech recognition, handwriting, and time-series prediction. We consider LSTMs as one of the most promising RNNs for predicting the trajectory of lateral movements within the RDW constraints, owing to their potential in predicting natural walking. Figure 2 illustrates a neuron of an LSTM network, as well as its key components. In LSTM, a forget gate, which is a *sigmoid* layer, is used to determine whether to retain the previous cell state. A *sigmoid* layer called the input gate, together with a *tanh* layer, is used to update the cell state. The output is then generated by filtering the cell state through a *sigmoid* layer to determine which part of the cell state should be outputted, and normalizing it using a *tanh* layer.

Most LSTM-based lateral movement predictors rely on past physical locations and historical movement patterns to predict users' movements, taking inspiration from predictions of natural walking trajectories. However, in full-immersive VR setups with RDW, users' movements are not entirely natural. RDW techniques steer the users to avoid collisions by directing the delivery of VR content towards collision-free physical locations. Therefore, the redirections that are expected to happen in the VE to prevent collisions are a crucial aspect of VR users' mobility.

We argue that the RDW-related inputs stemming from the VEs can be useful for the optimization of predicting near-future movement trajectories of full-immersive VR users. To incorporate this information, we introduce the concept of virtual locations, which represent the users' locations in the VEs. In addition to historical physical locations, we use a stream of historical virtual locations as input features to the proposed LSTM network. Depending on the utilized sources of input information, we distinguish between the "baseline" and "virtual" versions of our approach. It is worth noting that the VR users' virtual coordinates are presumed to be available one time step ahead of their physical counterparts, as depicted in Figure 2. We consider this assumption to be intuitive, since the RDW-derived virtual coordinates of a near-future time step are based on the physical coordinates at current time.

B. Short-term Orientational Movement Prediction

To predict orientational movements, a sufficient amount of experimentally obtained input data is required (i.e., distributions of yaw, pitch, and roll movements). To make existing predictors more effective, synthetic data needs to be generated using the limited experimental data available. Current methods of generating synthetic data are mostly model based, which necessitates expert input and significant modifications to the original design to transfer the generator across sources. Ideally, a model-free approach with higher generalizability would be preferable. GANs are a type of a general DL agent design that involves two sub-systems interacting adversarially to generate samples that can be integrated into the original dataset without significantly altering its overall distribution. As a result, GANs are extremely useful for developing model-free synthetic data generation techniques.

The GAN-based system starts with the generator utilizing random noise to create synthetic samples. Meanwhile, the discriminator acts as a supervised learning-based classifier, identifying whether a presented sample is genuine (coming

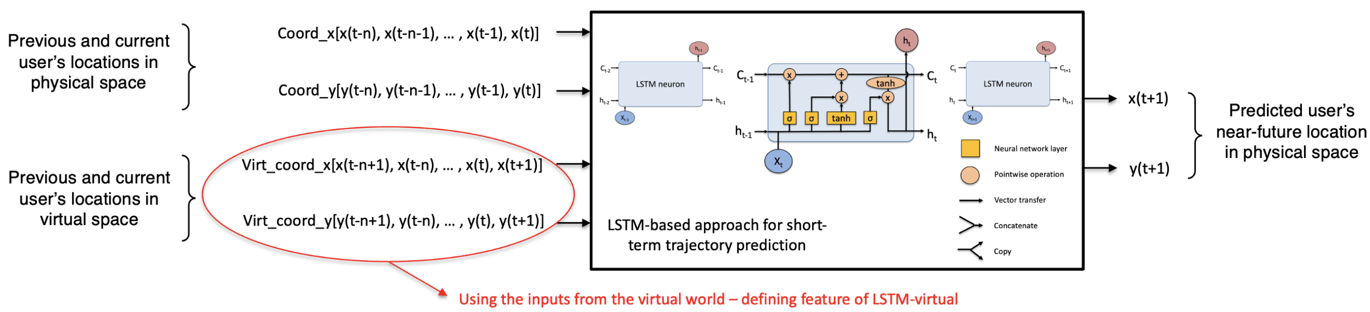


Fig. 2. Input features of the considered LSTM approaches for short-term lateral movement prediction

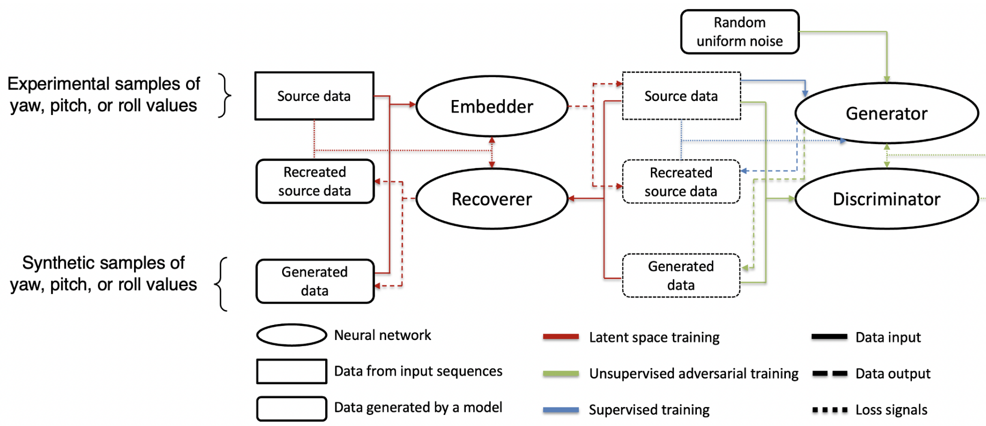


Fig. 3. Proposed design of the TimeGAN-based generation of synthetic orientational movement datasets

from the source dataset) or fake (coming from the generator). The generator cannot access the source samples and it only has access to the discriminator's loss function. Therefore, the generator focuses on optimizing the loss to improve its output. This iterative process drives the generator to generate increasingly realistic synthetic samples that resemble the original distribution, unveiling more nuanced differences between real and synthetic samples.

There is a need to maintain correlation between the samples as they represent a time series. We suggest using TimeGANs to create synthetic orientational datasets, as they are particularly well-suited for this task due to their ability to maintain time-dependencies. To achieve this aim, our proposed TimeGAN training process consists of a discriminator and generator, both of which use Gated Recurrent Units (GRUs) to handle time-dependencies. Additionally, we introduce an embedder and recoverer subsystems for data encoding into a lower-dimensional latent space, followed by decoding back to the original dimension. We first train these two subsystems before the generator produces samples in the latent space, which are then converted to time series through the recovery process. To generate complete latent representations of incomplete series from the source dataset, we use supervised learning. The generator is encouraged to capture time-correlation within the series through a loss function that measures the distance between synthetically generated data and the actual source data at similar time steps. For achieving this, we alternate between adversarial and supervised learning, using cross-entropy and Mean-Square Error (MSE) losses, respectively. Please refer to Figure 3 for a visual representation of our proposed design.

IV. Evaluation Setup and Results

We assess the performance of LSTM-based short-term lateral movements prediction in potentially multiuser VR setups with RDW. This is followed by the performance assessment of TimeGAN for generating synthetic orientational movement datasets, envisaged as a primer for predicting orientational movements of fully immersed VR users.

A. Short-term Lateral Mobility Prediction

The deployment environment utilized in the experiment was a square with dimensions of 15 meters. These dimensions were determined through experimentation to find the optimal size that would ensure maintaining acceptable level of noticeability, while still being practical for future deployments (e.g., in residential settings). To create this environment and prevent testers from colliding with obstacles, a 106×65 m^2 outdoor field near the University of Antwerp was utilized. The server running the RDW algorithm (i.e., APF) was a Windows 10-based MSI GS66 laptop with an Intel i7 processor, 16 GB of RAM, and WiFi 6. The HMDs used were Android-based Oculus Quests 2 with Qualcomm Snapdragon XR, a 120 Hz refresh rate, and 6 GB of RAM. Connectivity between the server and the HMDs was provided by a wireless hotspot using a Samsung Galaxy S8.

Two VEs were designed in Unity to evaluate the prediction accuracy. In the "straight path" experience, the testers were instructed to follow a straight path throughout the VE, representing the worst-case scenario for the noticeability and performance of the RDW algorithm. In the "random path"

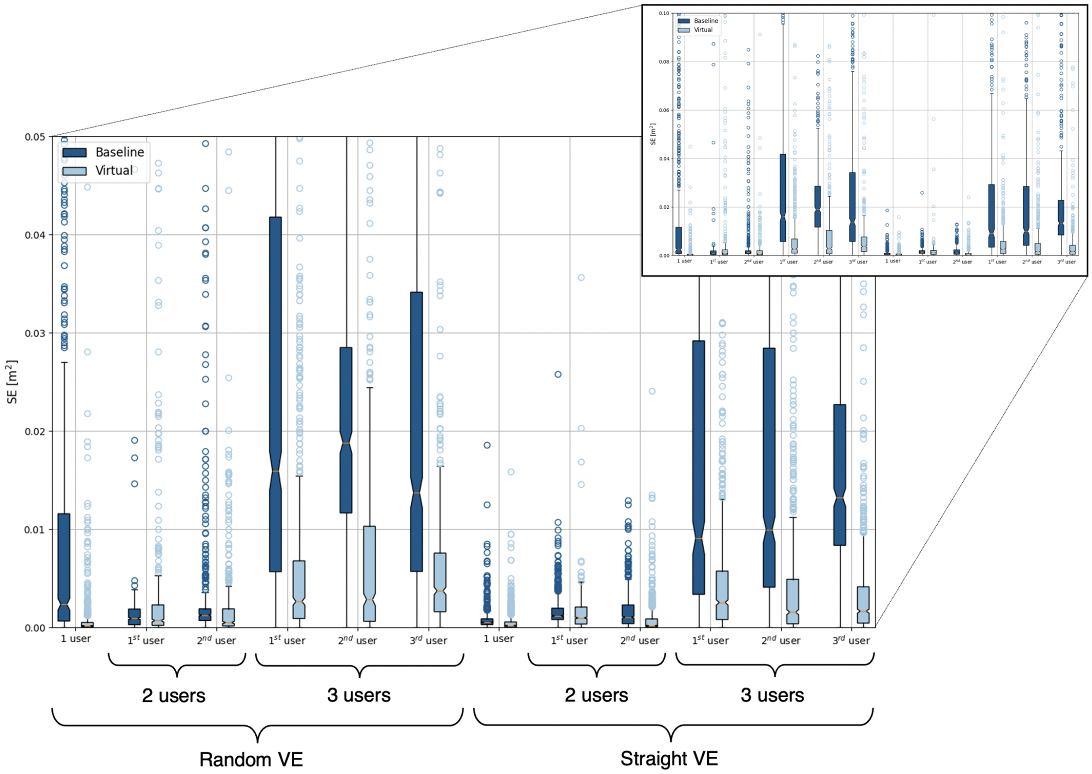


Fig. 4. Squared errors achieved by the considered LSTM-based approaches for short-term lateral movement prediction

experience, the testers were directed to follow a randomly curved path in an open environment. The expectation was that the curved path introduced by the VE would be less noticeable and benefit the RDW algorithm compared to the straight path.

The experiments involved the testers walking in an unbounded VE while being physically confined to a restricted environment. The positional data from the HMDs was sent to the server, where the RDW algorithm guided the testers within the physical boundaries to avoid collisions with other testers and environmental borders. Each experiment involved up to three testers coexisting in the environment and fully immersed in the VE. At the beginning of each experiment, the testers were instructed to follow a predefined path. They were informed that a reset might occur based on recommendations from the RDW algorithm (i.e., APF-R). If a reset occurs during the testing process, the testers would encounter a stop sign, followed by the VE rotating to provide guidance in the suggested direction. To maintain engagement, the duration of each experiment was limited to 5 minutes because the testers would lose interest afterwards as there were no distractions nor interactions within the VEs.

Figure 4 illustrates the performance of the LSTM-based prediction using Squared Errors (SEs) as the metric of interest. Comparing the baseline with the version that incorporates virtual coordinates alongside physical ones, it is evident that the latter generally outperforms the former. In a single-user system and for both VEs, the average per-user SE of the prediction is reduced from around 0.001 m^2 in the baseline to less than 0.0005 m^2 when utilizing additional context from the VE, resulting in a twofold increase in prediction accuracy.

The improvements become more significant with an increasing number of users. For instance, in a three-user system, utilizing context instances from VEs leads to improvements of up to 75% in the worst case, as depicted in the figure. These findings demonstrate the potential of leveraging specific context information from VEs, such as virtual coordinates of the users, to enhance the performance of LSTM-based predictors of near-term lateral movements. Moreover, the introduction of a second user does not notably affect prediction accuracy in the considered VEs, indicating the usefulness of short-term lateral movement prediction in *multiuser* VR systems with RDW. However, the accuracy decreases considerably when a third tester is introduced, regardless of the VE type. This observation highlights the significance of adequately scaling the physical environment according to the number of users and their patterns of mobility. Notably, the prediction accuracy in the straight path VE, which is the worst-case scenario for RDW, is better than that in the random path VE. This is likely because the straight path VE introduces less curvature in the testers' movements, and these more linear movements can be more accurately predicted.

B. Training of Orientational Mobility Predictors

During six two-minute-long sessions in an immersive VE from [15], three testers were free to navigate and have their full poses, including lateral and orientational movements, sampled at a frequency of 250 Hz. The orientational traces from this dataset were used to evaluate our approach for generating synthetic head rotation datasets, with a focus on

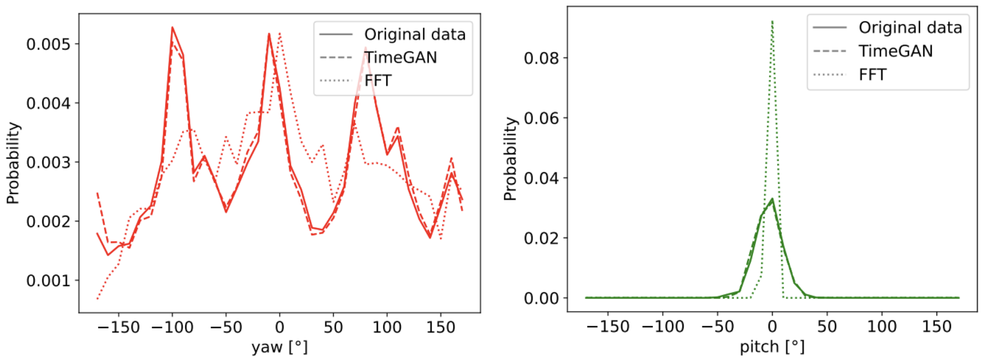


Fig. 5. Yaw and pitch values generated using the considered approaches for generating synthetic orientational datasets

the Probability Density Functions (PDFs) of the yaw, pitch, and roll movements. We quantized the PDFs into 10°-wide buckets centered around 0°. The yaw motion PDF was found to be substantially more complex than the other two types of movements, presumably due to the points of interest being spread across the testers' horizontal plane. As the VE was placed indoors, the users' gaze was primarily directed toward one of the walls, which explains the local maxima observed in Figure 5 around -90, 0, and 90°.

The Gaussian distributions of the roll and pitch movements are easily reproduced by neural networks. However, the distribution of yaw samples is more complex due to the multiple peaks and discontinuities that arise over time, since these samples are limited to the range between -180 and 180°. To overcome this challenge, we employ a quantile transformer to non-linearly transform the data and address its non-normality. In addition, we shift the remainder of a time series by 360° to avoid discontinuities. These transformations mean that the data range cannot be predetermined anymore, but the samples still fall within practically useful boundaries. Furthermore, these transformations are reversible, enabling synthetic data to be backtransformed to their original representations.

The time and complexity of GAN training increase in proportion to the input data size. To accommodate the demand for a large number of distinct samples in GAN-based DL, we divided each time series into instances of 25 samples using a sliding window. We also applied data downsampling, which did not significantly impact its utility. Based on the law of inertia, humans can perform a limited number of distinct head rotations within a brief timeframe, resulting in relatively smooth motions that can be accurately reconstructed using simple interpolation. With most of the energy concentrated in the lowest frequencies, over 90% below 5 Hz, the Shannon-Nyquist sampling theorem confirms that this information can be preserved by downsampling to 10 Hz.

The dataset is, therefore, divided into 23,700 samples of 25 points, each sample being 1.5 s long. The data was provided to TimeGAN for fitting the quantile transformer, followed by generating a synthetic dataset 10 times larger than the original one, once every 10 epochs. This was done because GANs are challenging to train and known to significantly degrade when overtrained. The procedure yielded a well-distributed synthetic dataset, despite the fact that we did not optimize the epoch hyperparameter nor repeat the resource intensive training procedure. The generated dataset is the final result of the procedure, which means that the system will not encounter inputs with slightly different distributions in the future. This is because the input comes from samples of uniformly distributed noise, implying that data overfitting can be excluded. Finally, we generated a baseline synthetic dataset using the FFT approach from [11]. The result is a 30,000 steps-long series, which we further downsample and divide into shorter samples using the above-discussed sliding window approach.

Figure 5 presents the distributions of yaw and pitch values. The roll distribution has been omitted because it is range-constrained as it captures uncomfortable head tilting. Hence, both the FFT- and TimeGAN-generated synthetic datasets closely match the roll distribution. However, the pitch distribution is slightly wider, and the FFT approach already fails to match it accurately, in contrast to TimeGAN. This trend is further pronounced in the yaw distribution, where TimeGAN closely matches its three local maxima, while FFT does so to a lesser extent. For example, the Kullback-Leibler divergence of the target yaw distribution compared to the TimeGAN-generated one is 0.00235, while the same compared to the FFT-resulting distribution is 0.0447 (i.e., almost 19 times higher). It is also worth noting that only FFT was hand-crafted to match this distribution. Based on these results, we argue that utilizing TimeGANs for generating synthetic head rotation datasets shows great promise.

V. Conclusion

Our work has shown that Long Short-Term Memory (LSTM)-based Recurrent Neural Networks (RNNs) hold promise for accurately predicting lateral movements in multiuser full-immersive Virtual Reality (VR) environments enabled through Redirected Walking (RDW). We have showcased the advantages of incorporating virtual context, such as the virtual movement trajectory in a Virtual Experience (VE), as an input source for enhancing the prediction accuracy. Moreover, we have proposed a TimeGAN approach for generating synthetic head rotation data, envisioned to serve as a primer for training orientational movement predictors in full-immersive VR setups. On a high level, we advocate for the utilization of predictive context-awareness in optimizing the connectivity in next generation VR. We expect this approach to be beneficial in applications ranging from dynamic multimedia encoding to millimeter Wave (mmWave) beamforming.

Note that we did not optimize the duration of the prediction window, but this duration should intuitively depend on the transmitter-side beamforming and beamsteering operating in the 100 ms timeframe considered in this work. Deriving optimal hyperparameterizations of the presented approaches was also not in scope, as the goal was to demonstrate the feasibility of predictive context-awareness. We consider addressing these limitations as a part of our future efforts. We argue that other context instances, for example the users' full 3-dimensional (3D) pose estimates, might be of interest in future VR systems, e.g., for enabling touch-like feedback or mobility-wise unconstrained portrayal of users in VEs.

Acknowledgments

This work was supported by the MCIN/AEI/10.13039/01100011033/FEDER/EU HoloMit 2.0 (nr. PID2021-126551OB-C21). This work also received support within the framework of the Recovery Plan, Transformation and Resilience (UNICO I+D 5G 2021, nr. TSI-063000-2021-6-Open6G Joint Open 6G Communications and Sensing), funded by the Spanish Ministry of Economic Affairs and Digital Transformation and European Union - NextGeneration EU. Finally, the work was supported by the Research Foundation-Flanders (FWO, nr. G034322N and 1SB0719N).

References

[1] L. Muñoz-Saavedra *et al.*, "Augmented and virtual reality evolution and future tendency," *Applied sciences*, vol. 10, no. 1, p. 322, 2020.

[2] H. Zhang *et al.*, "Wireless access to ultimate virtual reality 360-degree video," in *IoT Design and Implementation*, 2019, pp. 271–272.

[3] M. Chen *et al.*, "Virtual reality over wireless networks: Quality-of-service model and learning-based resource management," *IEEE Trans. on Communications*, vol. 66, no. 11, pp. 5621–5635, 2018.

[4] E. R. Bachmann *et al.*, "Multi-user redirected walking and resetting using artificial potential fields," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 5, pp. 2022–2031, 2019.

[5] M. S. Elbamby *et al.*, "Toward low-latency and ultra-reliable virtual reality," *IEEE Network*, vol. 32, no. 2, pp. 78–84, 2018.

[6] X. Liu *et al.*, "Learning-based prediction, rendering and transmission for interactive virtual reality in ris-assisted terahertz networks," *IEEE Journal on Selected Areas in Communications*, 2021.

[7] X. Song *et al.*, "Pedestrian trajectory prediction based on deep convolutional lstm network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3285–3302, 2020.

[8] T. Nescher and A. Kunz, "Analysis of short term path prediction of human locomotion for augmented and virtual reality applications," in *International C. on Cyberworlds*, IEEE, 2012, pp. 15–22.

[9] O. Abari *et al.*, "Enabling high-quality untethered virtual reality," in *USENIX Symposium on Networked Systems Design and Implementation*, 2017, pp. 531–544.

[10] T. Aykut *et al.*, "Realtime 3d 360-degree telepresence with deep-learning-based head-motion prediction," *IEEE Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 231–244, 2019.

[11] S. Blandino *et al.*, "Head rotation model for virtual reality system level simulations," in *IEEE International Symposium on Multimedia*, IEEE, 2021, pp. 43–49.

[12] F. Steinicke, G. Bruder, *et al.*, "Estimation of detection thresholds for redirected walking techniques," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 1, pp. 17–27, 2009.

[13] J. Lim *et al.*, "Beam tracking under highly nonlinear mobile millimeter-wave channel," *IEEE Communications Letters*, vol. 23, no. 3, pp. 450–453, 2019.

[14] J. Struye *et al.*, "Covrage: Millimeter-wave beamforming for mobile interactive virtual reality," *IEEE Transactions on Wireless Communications*, 2022.

[15] J. Chakareski *et al.*, "6dof virtual reality dataset and performance evaluation of millimeter wave vs. free-space-optical indoor communications systems for lifelike mobile vr streaming," in *Signals, Systems, and Computers*, IEEE, 2020, pp. 1051–1058.

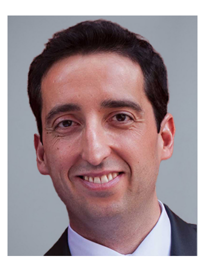
Biographies

Filip Lemic is a senior researcher at the i2Cat Foundation. He held positions at the University of Antwerp, imec, Universitat Politècnica de Catalunya, University of California at Berkeley, Shanghai Jiao Tong University, FIWARE Foundation, and Technische Universität Berlin. He received his M.Sc. and Ph.D. from the University of Zagreb and Technische Universität Berlin, respectively.

Jakob Struye is a Ph.D. researcher in the field of wireless networking at the IDLab research group (University of Antwerp) and imec research institute, Belgium. He obtained his B.Sc (2015) and M.Sc. (2017) in Computer Science at the University of Antwerp. His current research focuses on leveraging extremely high frequency wireless networks in the millimeter wave bands to improve the performance of truly wireless Virtual and Augmented Reality experiences, and he has experience in applying Artificial Intelligence to time series prediction problems.

Thomas Van Onsem is a graduated M.Sc. in software engineering from the University of Antwerp. He has an interest in influential and impactful technologies. Currently, he assists companies' digital transformations as a technical consultant for Exellys.

Jeroen Famaey is a research professor at IDLab research group of the University of Antwerp and imec, Belgium. He leads a team of researchers at IDLab, focusing on future wireless network technologies and protocols. His research has led to the publication of over 160 peer-reviewed journal articles and conference papers, and 7 granted patents.



Xavier Costa Perez is a research professor at ICREA, scientific director at i2Cat and head of 5G/6G R&D at NEC Laboratories Europe. He received his M.Sc. and Ph.D. degrees from the Polytechnic University of Catalonia and was the recipient of a national award for his Ph.D. thesis.