Enhancing Safety using AR Headsets with Motion Prediction Visualization

Vinu kamalasanan*
Institute of Cartography and Geoinformatics
Leibniz University Hannover

Ahmed Al-Taan†
Faculty of Civil Engineering and Geodesy
Leibniz University Hannover

Steffen Busch‡
Institute of Cartography and Geoinformatics
Leibniz University Hannover

Monika Sester§
Institute of Cartography and Geoinformatics
Leibniz University Hannover

Abstract

Pedestrian collisions are a serious safety issue when considering hazards to elderly persons and to distracted people due to AR gaming. The inability to anticipate an oncoming danger could put these pedestrians at a safety risk endangering their lives. However if we use an AR device to visualise future motion, then this could firstly help elderly wearing them to forsee any oncoming danger and secondly alert a distracted AR gamer if this information is included in the game. The RGBD sensors of an AR device continuously scans the environment to place virtual content. Our prototype idea focuses on applying motion detection and prediction algorithms to this scan data, and communicating it as information to the users. Then other traffic participants, who might potentially get in conflict with an AR headset user can be predicted in advance and informed. This can not only promote the use of AR headsets as a safety aid especially for elderly and disabled but also increase safety in outdoor spaces when using AR. In this paper we share our first results for our prototype based on the above idea where future pedestrian motions are detected and visualised with Unity on the Hololens 2.

1 Introduction

The ability of person to focus and respond to fast moving objects (e.g, cyclist or other pedestrians) depends highly on ones cognitive ability and the amount of attention given to the surrounding. While the former is highly depended on the mental ability of a person to act swiftly, the latter is subjective of whether the user is preoccupied indulging in parallel tasks (e.g., gaming or browsing) while walking.

Decreased activity, memory loss and poor judgement of speeds can cause elderly people to make inaccurate traffic decisions. These amongst others could be a reason why the aged feel unsafe and avoid mixed traffic environments like shared spaces [9]. Safety issues due to distractions on the other hand, are mostly attributed to gaming. A study by Ayers et al. [1] that analysed social media tweets for the safety impact of the AR game Pokemon Go reported 33 % of the hazards to the inattentiveness caused by the game. Another study for unregulated spaces [2] reported figures close to 22% where pedestrians failed to observe prior traffic. However for the traffic collisions, fatalities from pedestrian accidents could be higher when the environment is unregulated. While earlier works [4] and vision ideas [7] have explored the benefits of using AR headsets for safety of elderly and collaboration, they lack technical insights on how this can be achieved.

In this position paper, we hypothesise that an AR device with its sensors and 3D display can both perceive and augment future motion information to visually recommend averting a conflict. This we believe would not only strengthen the use of AR devices as mobility aids for elderly but also help make AR outdoor gaming safer .

*e-mail: vinu.kamalasanan@ikg.uni-hannover.de
†e-mail: al-taan@stud.uni-hannover.de
‡e-mail: steffen.busch@ikg.uni-hannover.de
§e-mail: Monika.Sester@ikg.uni-hannover.de

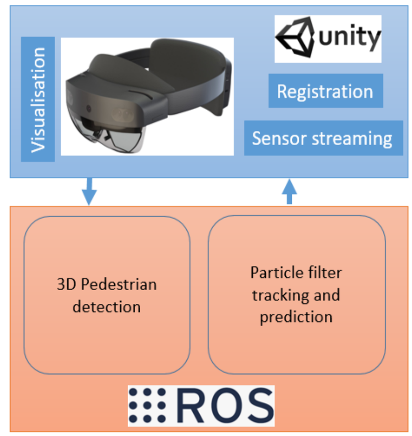

Figure 1: Block diagram of System

2 Related Works

2.1 Intelligent Mobility Aid

The white cane has been the preferred safety device to help disabled people considering its light weight and ease of use. Recent years however have explored sensor based technologies such as GPS, sonar and vision to guide elderly intelligently in outdoor spaces. Robotic researches have lately [6] focused on RGBD sensors and perception pipelines to dynamically detect other pedestrians to make navigation decisions for intelligent mobility aids like robotic wheelchairs.

2.2 Pedestrian detection and future trajectory visualisation

Deep learning based 3D pedestrian detection has been a focus topic for autonomous driving and robotic perception. Amongst other approaches, the Frustum based 3D object detector (F-Pointnet) [8] uses both the RGB and Depth to detect objects in world coordinates. As stated in [8] the F-Pointnet uses a mature image detector (e.g., YOLO) to first detect objects in image coordinates. A network is then further applied with the depth to detect objects in 3D space.

While not much research has been done for future trajectory visualisation to aid human navigation, a larger part of research is focused on the sports domain. In [3], Itoh et al. prototyped a future path visualisation for a moving ball using projection augmented reality. The visualisation approach applied a real time tracking apparatus to overlay future positions of the ball over time. Another similar work [5] focused on latency reduction in augmented visualisations, but however did not consider future motion or visualising it.

3 Proposed System

To predict and visualise future information, we propose and develop an end to end pipeline (Figure 1) that connects Hololens sensor perception with Unity visualisation. With this, our approach [a]

Figure 2: Each detected person in the view of the Hololens user is visualised in Unity with cube and future direction indicator

Figure 3: Future motion path as blurred arrow augmentation depicting uncertainty

detects and tracks pedestrians in real time using the Hololens sensors and further [b] applies 3D visualisation to the processed data with Unity to convey this information in AR.

Hololens spatial mapping [1] enables scanning the environment while detecting surfaces to place virtual hololgrams. With the release of the research mode, the device also provides access to this data as raw sensor streams. Among the available sensor streams detailed in [10], the RGBD data from the far-depth sensors and Hololens PV Image sensor has been used in our work.

In the calibration step, QR codes placed in the environment are scanned using the Mixed Reality Toolkit (MRTK). This enables to align the Unity coordinate system to the world space. The aligned virtual space is locked to the real world using World Locking Tools (WLT) [2]. This step allows both spatially and temporally aligned 3D content to be placed into the field of view of the Hololens user.

The RGB images and raw point clouds from the Hololens are input to the F-Pointnet [8]. The 3D detections from the network are further tracked over multiple frames using the particle filter. The positions of the tracked surrounding pedestrians are communicated back to the Hololens via the network, which is further visualised using the Unity Game Engine. For this work, we use a simple arrow based visualisation based on current 3D position of person(Figure 2).

4 Implementation and Conclusion

For our demonstration, we stream the RGBD data from the Hololens 2 to a Titan X GPU for F-Pointnet inference running at 9 fps. The particle filter tracked pedestrians trajectories are further smoothed and communicated back to Hololens Unity. For visualisation, we used a unit cube with a direction indicator that would symbolise the future path in the subsequent steps for the detected user. The complete video demonstration for the results can be viewed on the link https://youtu.be/3wYByD5zrbQ .

Even when our early prototype has proven the concept of visualising the future path, the limited sensing field of the Hololens device restricted the detection and visualisation to only a few meters. Also, our current tracking approach that worked for single pedestrians, needs to be extended to multi pedestrian scenarios and benchmarked. Furthermore it might be important to improve latency in AR visualisations and optimise our approach for smaller footprint devices (e.g., directly on a AR headset) for inference.

Figure 3 shows some of our next steps towards a more intuitive visualisation that would also account for the predicted uncertainty of the potential path (as a blurred arrow). Also it might be good to generalize the visualisation and framework to support tracking of several different types of objects. For instance, showing a moving person in one representation, and showing a moving vehicle with a different AR cue.

[1]https://learn.microsoft.com/en-us/windows/mixed-reality/design/spatial-mapping

[2]https://microsoft.github.io/MixedReality-WorldLockingTools-Unity/README.html

Acknowledgments

This research is funded by the DAAD under the GSSP Program with the German Research Foundation (DFG) through the Research Training Group SocialCars (GRK 1931).

References

[1] J. W. Ayers, E. C. Leas, M. Dredze, J.-P. Allem, J. G. Grabowski, and L. Hill. Pokémon go—a new distraction for drivers and pedestrians. *JAMA internal medicine*, 176(12):1865–1866, 2016.

[2] P.-L. Chen and C.-W. Pai. Smartphone gaming is associated with pedestrians' head-turning performances: An observational study of street-crossing behaviors at uncontrolled intersection in taipei. *International journal of sustainable transportation*, 12(1):12–18, 2018.

[3] Y. Itoh, J. Orlosky, K. Kiyokawa, and G. Klinker. Laplacian vision: Augmenting motion prediction via optical see-through head-mounted displays. In *Proceedings of the 7th Augmented Human International Conference 2016*, pp. 1–8, 2016.

[4] V. Kamalasanan, F. Schewe, M. Sester, and M. Vollrath. Exploratory study on the use of augmentation for behavioural control in shared spaces. In *Virtual, Augmented and Mixed Reality: 13th International Conference, VAMR 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings*, pp. 18–31. Springer, 2021.

[5] J. Knibbe, H. Benko, and A. D. Wilson. Juggling the effects of latency: Software approaches to minimizing latency in dynamic projector-camera systems. In *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pp. 93–94, 2015.

[6] M. Kollmitz, A. Eitel, A. Vasquez, and W. Burgard. Deep 3d perception of people and their mobility aids. *Robotics and Autonomous Systems*, 114:29–40, 2019.

[7] Y. Li, V. Kamalasanan, M. Batista, and M. Sester. Improving pedestrians traffic priority via grouping and virtual lanes in shared spaces (short paper). In *15th International Conference on Spatial Information Theory (COSIT 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022.

[8] C. R. Qi, W. Liu, C. Wu, H. Su, and L. J. Guibas. Frustum pointnets for 3d object detection from rgb-d data. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 918–927, 2018.

[9] C. Thomas. Briefing: Guide dogs announces results of shared surfaces research. In *Proceedings of the Institution of Civil Engineers-Municipal Engineer*, vol. 161, pp. 77–79. Thomas Telford Ltd, 2008.

[10] D. Ungureanu, F. Bogo, S. Galliani, P. Sama, X. Duan, C. Meekhof, J. Stühmer, T. J. Cashman, B. Tekin, J. L. Schönberger, et al. Hololens 2 research mode as a tool for computer vision research. *arXiv preprint arXiv:2008.11239*, 2020.