# Object Detection Systems Informing Users of Guardian Space Dangers Through Haptic Displays

Mattia Colombo[1]†, Pinar Satilmis[2], Richard Davies[3], Borislav Yordanov[4], Ian Williams[5], Carlo Harvey[6]
DMT Lab, School of Computing and Digital Technology, Birmingham City University, Birmingham, UK
orcid.org/{0000-0002-4169-2045[1], 0000-0002-5238-0725[2], 0000-0003-0656-9129[3], 0000-0002-0651-0963[5], 0000-0002-4809-1592[6]}

***Abstract*—Modern virtual reality applications rely on the guardian space to guarantee safety to the users during interactions. The current state of technology adopts virtual walls or meshes projected to the virtual space to alert users when their reach approaches the boundaries of the guardian space, requiring constant attention to the visual display. In this paper, we explore haptic displays as an alternative modality to alert users about dangers in the guardian space. We adopt object detection to analyse the scene within the guardian space, detecting unseen objects entering the space, i.e. a cat approaching the user's surroundings, causing a haptic display to signal the user the direction of the danger. We develop a proof-of-concept system capable of alerting a user of changes to the scene within the guardian space, demonstrating the feasibility of the system and its potential applications for extended reality platforms.**

***Index Terms*—haptics, computer vision, virtual reality, guardian space**

## I. Introduction

Virtual Reality (VR) has been shown to increase immersion and presence for users inside Virtual Environments (VEs) [5]. Head Mounted Displays (HMDs) that present VR, obfuscate the real surroundings from a user immersed within a VE. This presents a problem, users freely walking around real environments typically do so with some hesitance and reticence due to the possibility of colliding with obstacles within their real environment. Systems have been developed to overcome this concern, such as the HTC Chaperone [1] and Oculus Guardian [2]. These help users navigate within a predetermined volume that is considered to be safe to traverse without risk of collision. The systems simply present a wireframe mesh of the volume when coming close to a real-world boundary. This allows a user to be more confident within the virtual space and move with fewer precautionary measures.

The problem with these systems is that dynamic objects within the volume that would typically cause concern for a user are ignored. HTC Chaperone allows for using the web camera to augment a stylised edge view of the real world, so it is possible to pick up a coffee cup for example, but this system is turned on as required and doesn't monitor dynamic changes to the real environment. Any dynamic object, such as a household pet, that enters the guardian space and goes unnoticed is potentially a safety risk to the VR participant.

†Corresponding author: mattia.colombo@bcu.ac.uk

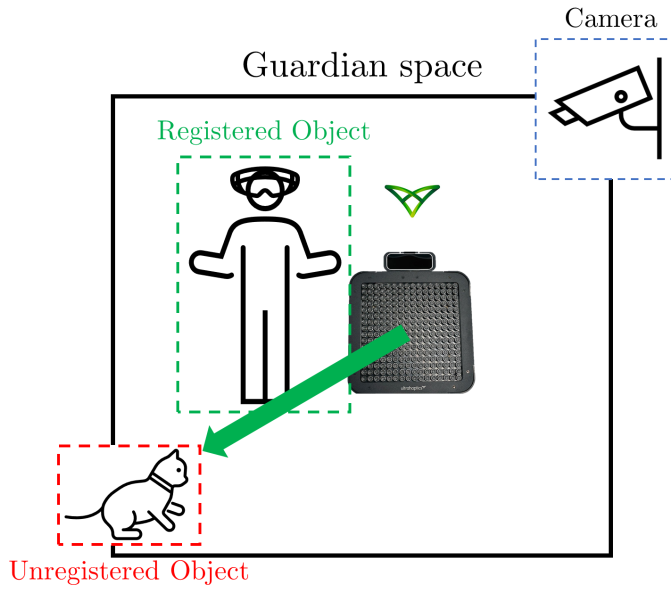


Fig. 1. The proposed pipeline for our haptic-based guardian space monitoring system: a camera monitors the space where the user is operating, feeding photographs of the scene to an object detection system to infer the presence of people and objects and register which objects are in the scene initially. When detecting an unseen object, the detection system transmits its position to a game engine that projects the object to a virtual environment which replicates the scene within the guardian space. With the detected object tracked within the virtual environment, the engine controls an Ultrahaptics STRATOS Explore to emit a sensation to the user reflecting the directionality of the detected object.

To take steps towards a solution to this problem, we propose a two-stage system. One system uses Computer Vision (CV) to analyse image streams of the playable volume and detects the user and dynamic objects within the volume. The second stage uses this information to determine whether dynamic objects have entered the guardian system and takes measures to present this information to users via a haptic interface, specifically the Ultrahaptic touch-less haptic device.

In brief, the contributions of this work are:

- Integrating real-time object detection methods into guardian spaces to act as monitoring systems;
- Adopting touch-less haptic displays to alert users of dynamic object(s) and their directionality within a guardian space;

- A proof-of-concept system that demonstrates how these systems could interact with one another to provide multi-modal feedback to a user immersed within a VE.

## II. Related Work

In the case of emergent technologies, safety is always a concern when it comes to mass adoption, and this is also true in the case of VR. Grega *et al.* [3] discuss many potential safety issues within the field, this includes tactile subsystems, which for user interaction is documented, but the work acknowledges that existing systems for solving collision situations are reliant upon pre-programmed software-based solutions, such as to acknowledge pre-defined objects and borders. Here, there is a lack of intelligent vision-based solutions to augment the safety paradigm.

Improving the quality and realism of interactions within VR has been a key focus area for developers and while desktop haptics is common, adding haptic interaction within virtual reality is more challenging. Swapp *et al.* [9] identify the necessity of implementing haptics to build upon a user's immersion within a virtual environment. The work develops experiments that are centred around the use of haptics to boost efficiency within co-location scenarios and demonstrates the beneficial effect of haptic feedback on task performance due to the physical response when interacting with objects in a virtual space. From this, it can be surmised that the implementation of haptics will raise user awareness of boundaries within virtual spaces during exposure to virtual experiences.

Previous work has aimed at improving the safety of VR play spaces by adopting depth-sensing systems, and LiDAR scans [4]. These proposed systems augment static guardian volumes beyond Manhattan cuboid principles to accurately map the real space that a VR exposure will take place within. The information is then presented in the virtual domain. This enables users to avoid obstacles more easily in their real-world twin.

All of these approaches to VR safety augmentation suffer from the same limitation: dynamic objects that enter the proposed safe area are untracked and have the potential to provide hazards that a user will be unaware of in real time. Therefore we put forward a proof-of-concept system that can determine and provide this information to a VR participant.

## III. Methodology

We propose a system to monitor the guardian space around a user experiencing extended reality applications. The system detects objects crossing the boundaries of the guardian space and informs the users about their presence and direction via a haptic display. Aligning to the way a guardian space works, the system has a calibration phase where objects detected through the camera are registered. After the calibration phase, the system will inform and transmit data about newly-detected, *unregistered*, objects to a game engine, which replicates the current scene of the guardian space in a virtual environment, allowing the tracking of the *unregistered* object and displaying a sensation through the haptic display to reflect its direction to the user. This workflow is illustrated in Figure 1.

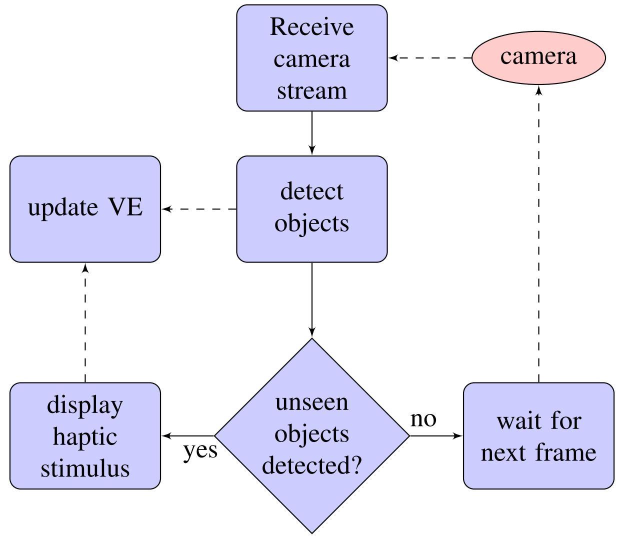


Fig. 2. A flow diagram of the proposed system: an object detection system receives frames from a camera monitoring the guardian space. Detected objects are projected into a virtual environment hosted by a game engine.

### A. Overview

We adopt an Ultrahaptics STRATOS Explore Development Kit as a haptic display within the user's reach. An object detection system analyses frames from a camera monitoring the guardian space, inferring the position and category of objects in the scene, see Figure 2.

### B. Object Detection

We adopt YOLOv5 to infer the presence and position of a variety of objects within a scene, including people [6], [8]. We place a camera to observe the guardian space, feeding frames to the model pre-trained on the COCO dataset [7], loaded on an Nvidia GTX 1060 CUDA GPU device. Objects are inferred for every frame grabbed from the camera stream, expressed as a series of bounding boxes within the frame with associated categories and confidence scores. With the camera monitoring the space, detected objects are marked as *registered* during the calibration phase and as *unregistered* during the running phase.

### C. Game Engine

We adopt Unity as a game engine to construct a virtual environment that hosts the guardian space, the haptic display, and objects detected by the camera system. The guardian space is implemented as a box trigger causing the haptic display to emit sensations upon collisions with an unseen object.

### D. Haptic Display

The haptic display is linked to the game engine and positioned within the virtual guardian space during the calibration phase. Upon collision with the box trigger, the engine causes the haptic display to emit a line sensation spanning between

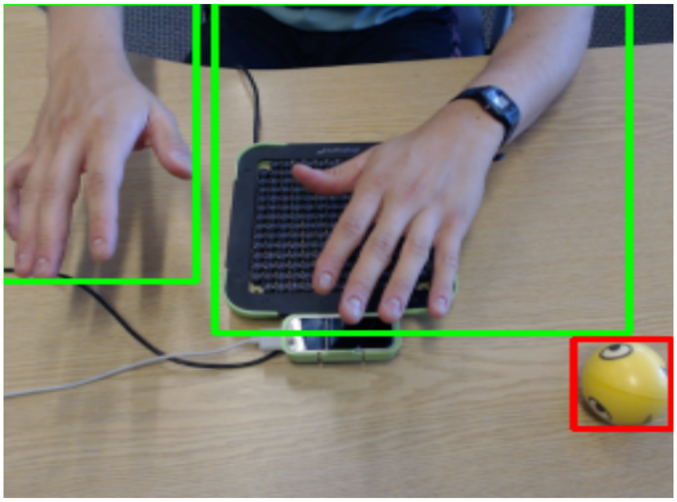

Fig. 3. A frame captured from a webcam, showing bounding boxes determined by the inference obtained from the object detection system. Green bounding boxes indicate objects that are registered at the calibration phase of the system. Red boxes indicate objects that enter the guardian space and pose a potential danger to the user.

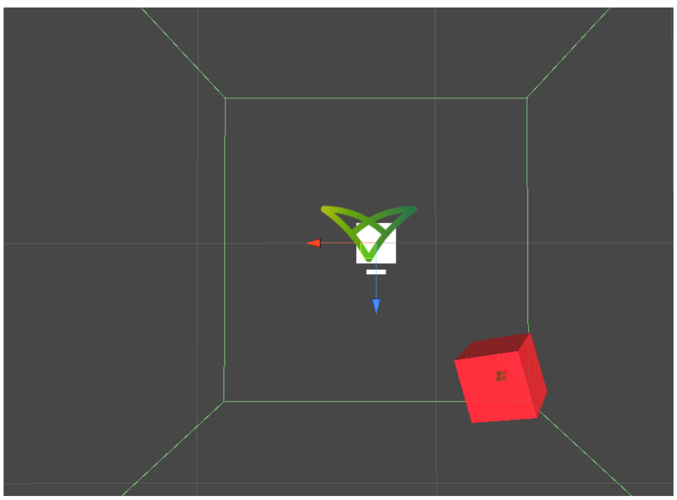

Fig. 4. A top view of the virtual environment showing the haptic display at the centre of the guardian space, as depicted in Figure 3. The red cube at the bottom right mirrors the danger detected by the vision system, generating a haptic sensation projected according to the position and direction of the object from the haptic display.

the centre of the device and the centre of the bounds of the colliding, *unregistered*, object. This vector creates a sensation of direction toward the hazard and indicates attention should be given to the new stimulus. The line is displayed as far as the delivery space of the display, allowing the user to track the directionality of the potential new dynamic hazard via the sensation of the haptic delivery.

## IV. Discussion

In our pilot study, it became clear that the relatively small delivery space of the haptic presentation, constraining the user to a small play area, would provide limitations when paralleled to the comparatively larger volumes of 6 degrees of freedom VR exploration. However, with future hardware advancements and the nature of the touchless technology alongside accurate hand and finger tracking, this realm of haptic modality augmentation within VR has significant benefits to be explored. In addition, occlusion could cause undetected unseen objects due to obstructions or adverse camera positioning.

There are numerous applications that can benefit from this approach, not just in the virtual world, but in the physical world too. For example, using Computer Vision paradigms to object detect and track everyday objects, such as bicycles externally to a vehicle may be used to provide decision-making and warning delivery mechanisms to drivers. Here, it would be possible to exploit onboard cameras commonly mounted on modern vehicles to scan the spaces surrounding a driver, improving safety in situations where other vehicles or objects are present within the driver's blind spots.

## V. Conclusion

To summarise, we have presented a proof-of-concept system that integrates CV Object Detection with haptic interfaces to augment the information presented to a user of VR. This system has the potential to increase safety and provide users with even greater confidence for real-world navigation tasks. The proposed system provides more detailed information about the surroundings and works for single or many dynamic objects. With this design, even though users cannot anticipate changes to their surroundings, they can still avoid collisions with the accompaniment of this real-time monitoring system. However, there are areas to further research and limitations to overcome.

### A. Future Work

Further work will study whether users can engage with VR and use the haptic information presented effectively, in order to improve safety in guardian spaces that have dynamic objects within them by addressing the following points of expansion. These include:

- determining which haptic interaction paradigms best provide a sense of directionality to a user, in addition to engaging user attention via alert mechanisms;
- investigating the effect on presence and immersion when using the modality of haptics to present this information to a user in VR;
- exploring alternative modalities, such as acoustics to present information about the real-world surroundings behind a user;
- solving the problem of the static nature of the haptic device given the 6 degrees of freedom nature of typical VR usage;
- employing multi-camera systems to address potential occlusion of unseen objects.

## References


[1] Changing the chaperone settings. https://www.vive.com/uk/support/vive/category_howto/changing-the-chaperone-settings.html. Accessed: 2022-08-10.

[2] Oculus guardian system. https://developer.oculus.com/documentation/native/pc/dg-guardian-system/?locale=en_GB. Accessed: 2022-08-10.

[3] Matúš Grega, Pavel Nečas, and Branislav Lancik. Virtual reality safety limitations. *INCAS Bulletin*, 13:75–86, 2021.

[4] Shaoyan Huang, Huidong Bai, VBH Mandalika, and Robert W Lindeman. Improving virtual reality safety precautions with depth sensing. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, pages 528–531, 2018.

[5] Jason Jerald. *The VR Book: Human-Centered Design for Virtual Reality*. Association for Computing Machinery and Morgan & Claypool, 2015.

[6] Glenn Jocher, Ayush Chaurasia, Alex Stoken, Jirka Borovec, NanoCode012, Yonghye Kwon, TaoXie, Jiacong Fang, imyhxy, Kalen Michael, Lorna, Abhiram V, Diego Montes, Jebastin Nadar, Laughing, tkianai, yxNONG, Piotr Skalski, Zhiqiang Wang, Adam Hogan, Cristi Fati, Lorenzo Mammana, AlexWang1900, Deep Patel, Ding Yiwei, Felix You, Jan Hajek, Laurentiu Diaconu, and Mai Thanh Minh. ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference, February 2022.

[7] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[8] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[9] David Swapp, Vijay Pawar, and Céline Loscos. Interaction with co-located haptic feedback in virtual reality. *Virtual Reality*, 10(1):24–30, May 2006.