

# The Reality-Virtuality Interaction Cube: A Framework for Conceptualizing Mixed-Reality Interaction Design Elements for HRI

Tom Williams  
Colorado School of Mines  
Golden, CO, USA  
twilliams@mines.edu

Daniel Szafir  
University of Colorado Boulder  
Boulder, CO, USA  
daniel.szafir@colorado.edu

Tathagata Chakraborti  
IBM Research AI  
Cambridge, MA, USA  
tathagata.chakraborti1@ibm.com

## I. INTRODUCTION

There has recently been an explosion of work in the human-robot interaction (HRI) community on the use of mixed, augmented, and virtual reality. We present a novel conceptual framework to characterize and cluster work in this new area and identify gaps for future research. We begin by introducing the *Plane of Interaction*: a framework for characterizing interactive technologies in a 2D space informed by the Model-View-Controller design pattern. We then describe how *Interactive Design Elements* that contribute to the interactivity of a technology can be characterized within this space and present a taxonomy of mixed-reality interactive design elements. We then discuss how these elements may be rendered onto both reality- and virtuality-based environments using a variety of hardware devices and introduce the *Reality-Virtuality Interaction Cube*: a three-dimensional continuum representing the design space of interactive technologies formed by combining the Plane of Interaction with the Reality-Virtuality Continuum. Finally, we demonstrate the feasibility and utility of this framework by clustering and analyzing the set of papers presented at the 2018 VAM-HRI workshop.

## II. THE PLANE OF INTERACTION

The **Model-View-Controller** design pattern separates interactive systems into three pieces: the internal *model* of the interactive system, the user's *view* into that model, and the user's *controller* for effecting changes to that model [1]. Interactive robots can be conceptualized in a similar manner. Every robot has some internal state that can include matters-of-fact such as the robot's pose and battery level as well as cognitive constructs such as beliefs, goals, and intentions. User interaction with such a robot depends upon both the expressivity of that robot's *view* (i.e., the means used by the robot to communicate its state to humans) and the flexibility of its *controller* (i.e., the means available for users to modify the robot's state). Accordingly, a robot's level of interactivity can be conceptualized as a point on a *plane of interaction*, with two axes: *expressivity of view* (hereafter EV) and *flexibility of controller* (hereafter FC).

Here, the focus on expressivity and flexibility is crucial. Simply logging additional low-level data without providing

higher-level features or summaries, or new ways of viewing that data, would yield a limited gain in expressivity. Similarly, simply enabling complete, direct joysticking without providing opportunities to influence the robot's higher-level beliefs, desires, and intentions, would yield a limited gain in flexibility.

Note that while this plane does not explicitly include the robot's model, it is implicitly represented by the scale of the plane. The more sophisticated a robot's model, the greater potential for EV and FC, but model complexity alone does not dictate interactivity. Rather, model complexity dictates the level of interactivity enabled by interaction design elements.

## III. INTERACTION DESIGN ELEMENTS

We define *interaction design elements* as those components of a robot's design that can be said to impact its interactivity. We define the impact of each design element on the robot's interactivity as  $M \begin{bmatrix} \Delta_{EV} \\ \Delta_{FC} \end{bmatrix}$ , where  $\Delta_{EV}$  is the impact a design element has on the expressivity of the user's view into the robot's model, and  $\Delta_{FC}$  is the impact a design element has on the flexibility of the user's control of the robot's model, both of which are scaled by  $M$ , a measure of complexity of the robot's internal model.

## IV. MIXED REALITY INTERACTION DESIGN ELEMENTS

Within mixed-reality environments, we consider three types of *mixed-reality* interaction design elements (MRIDEs).

**User-Anchored Interface Elements:** These are interface elements similar to those seen in traditional GUIs, anchored to points in the user's camera's coordinate system, and which do not move as the user changes their field of view.

**Environment-Anchored Interface Elements:** When interface elements are instead anchored to points in the coordinate system of a robot or some other element of the environment.

**Virtual Artifacts:** In contrast, we also consider virtual artifacts that can be manipulated by either humans or robots (or which may move under their own ostensible volition). Crucially, virtual artifacts in this category are recognizable (by humans) as additions to the underlying environment, but may not be recognizable as such to robots. For example, a robot

may render an arrow into the environment which a human may then manipulate to change the robot's intended direction; an arrow recognizable by both parties as not part of the actual environment. In contrast, a human may render a virtual wall into the environment to restrict a robot's path, but this wall may or may not be identifiable as virtual to said robot.

## V. REALITY-VIRTUALITY INTERFACE CONTINUUM

The nature of the environment onto which mixed-reality interactive design elements (hereafter simply "design elements" for brevity), if any, are rendered is itself a function of the interface in which they are presented. We consider two types of interfaces, each falling at a different point along Milgram's Reality-Virtuality Continuum.

**Reality:** Typically, design elements are overlaid onto a user's local reality. This may be achieved by using AR Head-Mounted Displays (AR-HMDs) (or similar technologies such as AR displays in car windshields) to render design elements over a user's field of view; using passthrough from a robot's camera into a VR Head-Mounted Display (VR-HMD); or using projectors to render design elements directly onto the environment for multiple users to view simultaneously without the need for HMDs (although this approach does not typically allow for the rendering of three-dimensional Virtual Artifacts and is limited in the placement of design elements).

Using HMDs, especially VR-HMDs, design elements can also be overlaid onto a reality other than a user's local reality. This may be the visual perspective of a robot, another human, or a camera not associated with any particular agent.

**Virtuality:** Finally, design elements may be overlaid onto completely virtual environments. This is typically the case when VR is used for training and simulation of robots before they are moved to the real world.

## VI. THE REALITY-VIRTUALITY CUBE OF INTERACTION

By combining the interaction plane and the reality-virtuality interface continuum, we produce a three-dimensional space which we term the *Reality-Virtuality Cube of Interaction*. In this survey, we will use this cube to describe and categorize recent research efforts at the intersection of HRI and AR/VR.

## VII. SURVEY: VAM-HRI 2018

To demonstrate the applicability of the proposed framework, we analyze papers presented at the VAM-HRI workshop<sup>1</sup> at HRI 2018 [2], [3] in Table I. According to our framework, approximately two-thirds of papers leveraged AR or VR to enhance interactions in reality, while third of the papers leveraged AR or VR to enhance interactions in virtuality.

Of the approaches enhancing interactions in reality, six enhanced both view and control without the use of mixed reality design elements. These all used a VR HMD for remote teleoperation, without displaying any additional graphics within the HMD. The VR HMD necessarily increased view (through direct visualization of remote robots' sensor data),

	No MRIDEs used			MRIDEs used			Total
	$\emptyset$	C+	V+C+	C+	V+	V+C+	
Reality	0	0	6	2	8	4	20
Virtuality	3	3	0	0	3	0	9

TABLE I: VAM-HRI'18 Papers, categorized by (1) whether Mixed Reality Interaction Design Elements (MRIDEs) are overlaid over the environment, (2) whether that environment is real or virtual, and (3) whether the presented technologies improve flexibility of control (C+), expressivity of view (V+), both (V+C+), or neither ( $\emptyset$ ).

and all approaches used either the HMD or other hardware to also increase flexibility of control. Two approaches used AR/VR for teleoperation, displaying interface elements for controlling the robot. One used AR for local teleoperation; one used VR for remote teleoperation. Eight approaches used view-enhancing augmentations: five provided passive displays of robot trajectories or sensor data; two provided or proposed active communicative displays; and one displayed virtual robots. Finally, four approaches used augmentations enhancing both view and control. Two used AR UI elements to control and calibrate a virtual robot, and two provided virtual objects that could be interacted with to affect robot behavior.

Of the approaches enhancing interactions in virtuality, six did so without using mixed reality interaction design elements, simply using VR as a window into a virtual environment. Three of these increased control by enabling humans to use control virtual robots, while three allowed observation of uncontrollable virtual robots. These six approaches focused either on using VR to allow humans to train robots or on using VR to train humans to interact with robots or study perceptions of virtual robots. Finally, three approaches enhanced expressivity of view in virtuality. These involved large-scale maritime or aviation contexts in which it was more helpful to see a top-down view of the larger maritime or aerial region than just the perspective of the single unmanned surface vehicle or drone, and in which helpful information was overlaid on the canvas of the open maritime or air space.

## VIII. CONCLUSION

In this paper we have presented a novel framework for categorizing AR and VR approaches intended to improve interactive technologies, and demonstrated the utility of this framework for HRI by using it to categorize papers from VAM-HRI 2018. In future work, we hope to use this framework as a jumping off point for a full survey of work using VR and AR in HRI. Furthermore, we hope that this framework serves as a useful tool to allow researchers working in this area to better communicate the contributions of their research.

## REFERENCES

- [1] Glenn E Krasner, Stephen T Pope, et al. A description of the model-view-controller user interface paradigm in the smalltalk-80 system. *Journal of object oriented programming*, 1(3):26–49, 1988.
- [2] Tom Williams, Daniel Szafir, Tathagata Chakraborti, and Heni Ben Amor. Report on the 1st international workshop on virtual, augmented, and mixed reality for human-robot interaction (VAM-HRI). *AI Mag.*, 2018.
- [3] Tom Williams, Daniel Szafir, Tathagata Chakraborti, and Heni Ben Amor. Virtual, augmented, and mixed reality for human-robot interaction. In *Companion of HRI 2018*, pages 403–404. ACM, 2018.

<sup>1</sup>VAM-HRI Series: <http://vam-hri.xyz/>