AAR: Augmenting a Wearable Augmented Reality Display with an Actuated Head-Mounted Projector

Jeremy Hartmann, Yen-Ting Yeh, Daniel Vogel
School of Computer Science, University of Waterloo
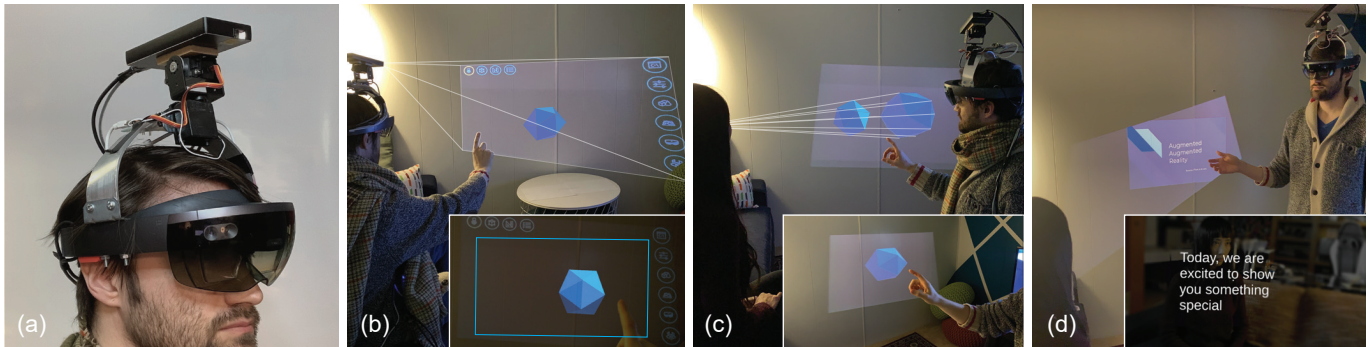{j3hartma, y6yeh, dvogel}@uwaterloo.ca

Figure 1. Augmented Augmented Reality system and use cases: (a) a pico projector on a servo-controlled pan-tilt head is mounted on a Microsoft Hololens V1 AR HMD, when calibrated precise control of the projected image relative to the HMD or the realtime scan of world geometry is possible; (b) the projector can display interactive content beside the optical display, such as a toolbar; (c) the projector can let bystanders "peek" into the HMD user's virtual world; (d) the projector can show public content while keeping private information in the HMD, such as during a presentation.

ABSTRACT

Current wearable AR devices create an isolated experience with a limited field of view, vergence-accommodation conflicts, and difficulty communicating the virtual environment to observers. To address these issues and enable new ways to visualize, manipulate, and share virtual content, we introduce Augmented Augmented Reality (AAR) by combining a wearable AR display with a wearable spatial augmented reality projector. To explore this idea, a system is constructed to combine a head-mounted actuated pico projector with a Hololens AR headset. Projector calibration uses a modified structure from motion pipeline to reconstruct the geometric structure of the pan-tilt actuator axes and offsets. A toolkit encapsulates a set of high-level functionality to manage content placement relative to each augmented display and the physical environment. Demonstrations showcase ways to utilize the projected and head-mounted displays together, such as expanding field of view, distributing content across depth surfaces, and enabling bystander collaboration.

Author Keywords

Augmented Reality; Projection Mapping; Calibration, Asymmetric Interaction, Spatial Augmented Reality

INTRODUCTION

Augmented reality (AR) has the potential to truly merge digital and physical worlds. Typically, an optical see-through head-mounted display (HMD) is used to composite virtual content into the surrounding environment [34]. While effective in many ways, it also has a limited field of view and suffers from vergence-accommodation conflicts. Further, the user experience is isolating, since the virtual environment is only visible to the HMD user. This makes collaboration and communication with external users difficult.

An alternative to creating AR with an HMD is Spatial Augmented Reality (SAR) [12], which uses projected light to directly augment physical surfaces. SAR can be used with an optical see-through HMD to alleviate some limitations, such as simulating an expanded field of view [9] and improving perceptual depth cues [12]. Another possible way to improve on AR HMD experiences is with cross-device systems that combine many conventional displays and devices with an AR HMD, like smartphones, smartwatches, and large displays. This has been used to enable external communication [54], expand the capabilities of devices [20], and enhance interaction with 3D virtual objects [42]. However, both approaches limit user mobility and do not allow for ad hoc serendipitous collaborations with external users. SAR typically requires multiple external projectors installed and carefully calibrated

Figure 2. AR persistence example application. The projector acts as an ad hoc display when the user takes off the HMD, enabling them to continue watching a video started in AR.

to a specific environment, and cross-device systems require specialized software and experiences are constrained by the physical properties of the device.

We introduce *Augmented Augmented Reality* (AAR), the combination of a see-through wearable AR display with an actuated head-mounted projector. AAR can be expressed in a concise design space, where potential user roles and projector roles intersect. Using this space, we explore how AR interfaces can be extended and combined, enabling new ways to view, manipulate, and share AR content. For example, the HMD user experience can be enhanced by using the projector to augment their view with peripheral information, such as simulating a heads-up GUI (Figure 1b). Or, the projector can share AR with external users, for instance, a view-dependent rendering on a nearby wall so a bystander can see into the HMD user's virtual world (Figure 1c). A combined HMD and projector can enable new communication opportunities, such as an impromptu presentation with projected slides for an external audience and private notes for the HMD user (Figure 1d). The system can even transition the HMD user from virtual to real worlds, when they remove the HMD, the projector could persist a portion of the virtual content (Figure 2).

To operationalize the AAR concept, we built an actuated pico projector system mounted on a HoloLens AR HMD. An important piece is how the projector, pan-tilt geometry, and HMD are calibrated: for this, we developed a novel variation of the structure from motion pipeline that utilizes a dense correspondence map between all camera and projector view combinations. Applications access the calibrated projector and HMD through a ARR software toolkit. This gives developers high level control over projector roles relative to the HMD, the user, and the physical environment, and defines content behaviour across the AR HMD and projected SAR.

In summary, we make the following contributions:

- Concept and applications for combining a wearable actuated projector with an AR HMD;

- An automated calibration process to build a geometric representation of a head-mounted projector pan-tilt structure;

- An open source toolkit to develop AAR applications.

RELATED WORK
Current optical see-through AR HMDs use diffractive grating waveguide combiners with two-dimensional eyebox expansion [35]. An alternative is a head-mounted projector display which uses a small pico projector with an optical beam splitter to reflect light off of real-world surfaces, using a retro-reflective coating [27]. A variation of this removes the optical beam splitter to use the projector directly, augmenting the environment through projected light [29]. All variants of AR offer the user unique affordances that can be thought of as complimentary. Our work investigates the duality of these approaches to AR displays, building on previous work in wearable projector displays, SAR, peripheral displays, and AR systems that span public and private contexts.

Augmented Reality Wearable Projector Systems
Using projectors to augment real-world surface geometry was first explored in Shader Lamps, later conceptualized as SAR [51]. Since then, several projects have explored SAR in complex multi-projector arrangements [25, 28, 50] and steerable projector systems [33, 49]. Beamatron [55] explored a steerable SAR environment where the projector unit was statically fixed to a location on a ceiling. They demonstrated a broad set of applications that explores SAR under this context. Most relevant, mounting a projector directly on the body has been shown to provide new opportunities for on-body [43] and context driven interaction [41, 56]. OmniTouch [24], a shoulder-worn depth-sensing and projection system, can transform an everyday surface into an interactive space, focusing solely on the the user wearing the device.

Of the many possible on-body mounting locations, on or near the head has been of particular interest. A fixed, front facing head-mounted projector can be used to directly augment the physical environment to reproduce the effect of wearing an optical see-through AR HMD [27, 31]. This has benefits. Scape [26] showed how such "head-mounted projective displays (HMPD)" can enable multi-user collaborative AR. Krum et al. found the approach allowing for more natural depth cues [36], while Kade et al. demonstrated entertainment applications like a shooting game [29]. Genç et al. showed a head-mounted projected image of static and dynamic content is effective when the user is in motion [19].

Our work builds on these concepts and extends them by actuating a head-mounted projector and using it with an optical see-through AR HMD. Different from SixthSense [43] and OmniTouch [24], we enlarge the SAR display space and enable interactions which can be independent to the user's location with the actuated projector. By combining a steerable projector with an AR HMD, we can explore the design space between an HMD and external users more broadly and under different location contexts. This contrasts other work in steerable displays that are fixed, large, and limited to a single location [55]. Together, our approach can augment the surface geometry for the user, while also enabling external ad hoc collaborations with outside observers.

Public and Private Context Sharing
The use of public and private displays to share content between users has been thoroughly explored [14]. Augmented Surfaces [52] introduced the concept of hyperdragging which allows co-located users to share private content with a shared public space. Code Space [13] investigates cross-device content sharing with a large public display in the context of code reviews, and MeetAlive [18] explores multi-device sharing in a SAR equipped meeting environments. Sharing can also be accomplished using collaborative augmented reality [10, 48], where interactive experiences are shared among multiple co-located or remote users wearing AR HMDs.

EMMIE [15] uses an AR HMD combined with external displays to merge private and publicly viewable content. Elements of this was further explored in Focus+Context screens [8]. Machuca et al. outlined some design considerations when blending 3D content between a handheld device and a public screen [38]. Serrano et al. [54] explored the combination of an AR HMD within a distributed display environment. Rukzio and Holleis explored a design space that spans a mobile phone and a public projector [53]. Our work builds on previous explorations in context sharing by exploring the asymmetric duality provided by an AR HMD and SAR display and how the user fits within it.

Hybrid AR Displays
Researchers have investigated ways to overcome the limitations of current generation AR HMDs [3, 6] by using sparse peripheral displays, adding LED arrays surrounding AR HMD [21, 57], or by combining a SAR type environments with AR HMDs [9]. Combining a SAR environment with an AR HMD offers extra information [59] and improves the visual effect [39] in the AR world. It also provides some useful affordances since the the user's view and the environment are independent, which allows enhanced material rendering [23], shared multi-user experience [32], and fixed environments for an expanded field of view [9].

Closely related to our work is FoveAR [9] which combines a single fixed ceiling-mounted projector SAR environment with an AR-HMD as an extended peripheral display. They demonstrate their approach through a set of four experiences, 3D model animation, wide-angle immersive simulation, 3D life-size telepresence and an AR shooter game, that utilize the capabilities of both the projector and the AR HMD. Our work also utilizes a projector and AR display as well, but the projector is directly attached to the HMD and can be freely repositioned into different viewing configurations.

While these works combine an AR HMD with a fixed ceiling-mounted projector, none have investigated a compact inside-out actuated projector display combined with an AR HMD, nor do they fully explore design considerations for both the HMD user and other external users who could also benefit from projected AR content.

Summary
Our work builds and significantly extends previous concepts with a re-imagined and more comprehensive exploration, applications, and technical solutions. For example, the mobility and flexibility of a head-mounted steerable projector provides a larger set of experiences not possible in fixed projector environments, and our design space spans the asymmetric duality between the HMD user and external observers.

AUGMENTED AUGMENTED REALITY
An AR display and a SAR environment both have advantages and disadvantages in how they augment the environment and how the user interacts with the virtual content. The AR display can produce high-quality 3D holograms, but is limited to a fixed focus plane with a smaller field-of-views. SAR is able to produce realistic depth cues for surface-mapped 2D content, but 3D content is limited to a single view-dependent perspective. One goal of AAR is to create a setting where a strength of one AR device can offset a weakness in the other. For the HMD user, the two displays can work together to create an enhanced AR experience. In addition, considering the projected display as public, the two displays can provide a dynamic environment in which the HMD user can communicate their virtual environment with external observers. This provides opportunity for new AR modalities for communication. In both cases, the projector is essentially "augmenting" augmented reality.

Design Space
To generate and describe different types of AAR experiences, we developed a concise two-dimensional design space (Figure 3). It captures the two important factors: *who* is benefiting from AAR (the "User Role") and *how* the actuated projector is used with respect to the AR HMD view and physical environment (the "Projector Role"). Our design space is complementary to the technically-focused design considerations provided in FoveAR [9]. They present set of techniques to render an AR HMD with single-projector for only the HMD user. For example, surface shading textures into the environment with overlayed 3D holograms in the HMD.

User Role Dimension
(HMD User, External User, Both Users)
The projector can be used to improve the experience for the user wearing the AR HMD, one or more "external" users who are standing near the HMD user, or in some cases both types of users can benefit simultaneously.

There is a dichotomy between the HMD user and external user, as their ability to engage with virtual content is asymmetrical. Clearly, the HMD user has more affordance in the range of actions they can place on virtual objects and how they can view objects across AR views in the HMD and from the projector. For the external user, what content they can see and how they might interact with it is likely to be determined by the HMD user who is present and the dominant actor in the virtual scene. There are exceptions to the requirement of a primary HMD user, for example the HMD could be set on a table so the projector acts like a steerable ad hoc display.

Projector Role Dimension
(Augment View, Augment Environment, Simulate Display)
The interplay between the AR HMD and projector display can be thought of in terms of assistive modalities, where the projector aids the HMD display or the HMD display aids the
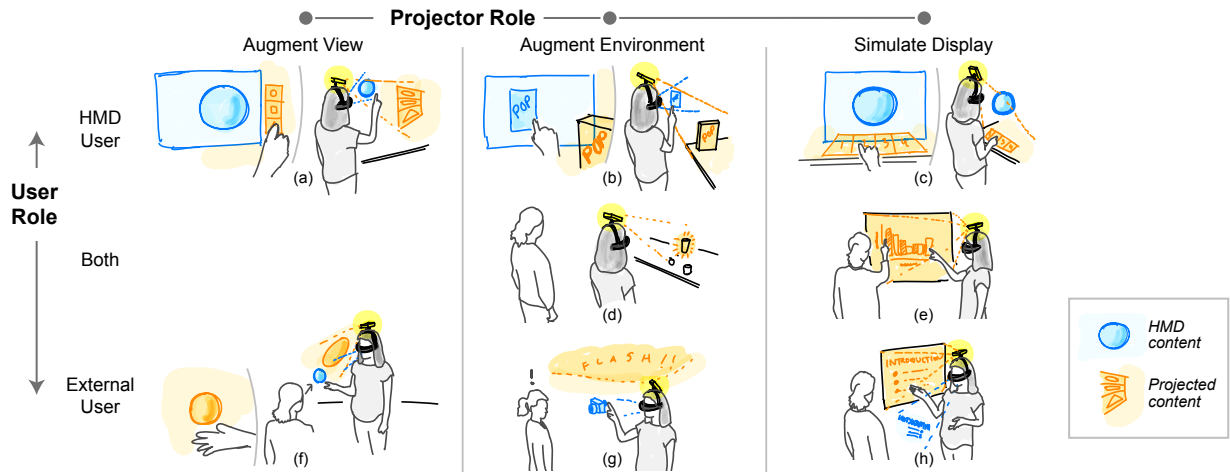
Figure 3. AAR design space considering user role and projector role: (a) HMD user's view augmented with heads-up GUI; (b) environment object augmented to show physical preview of a product design; (c) simulating a GUI display on a nearby surface; (d) highlighting an object in the environment to communicate with external user; (e) collaborating on a simulated whiteboard display; (f) augmenting an external user's view with a view-dependent rendering of an object to create the illusion that it is placed in front of the HMD user; (g) augmenting the environment with a camera flash so external user's are aware a photo was taken; (h) using a nearby wall to display a slide-show presentation.

projector. Previous work has focused on the former, such as FoveAR [9] and occlusion shadows to artificially increase AR HMD contrast [11]. In our space, the projector's role can be expressed as one of three ways in which it renders content.

The projector can be used to *augment the view* of the HMD user or in more limited cases, the external user. Using view dependent rendering with the projector frustum near the HMD view frustum, the projected image can create the illusion of peripheral content on or around the HMD view. For example, to create the illusion of an extended AR HMD for a heads-up GUI (Figure 3a). The projector can also augment the external user's view, for example creating a view-dependent rendering on a nearby wall such that they can see a location-matched 3D view into the HMD user's virtual world for the purpose of collaboration (Figure 3f).

The projector can *augment the environment* by enabling a steerable surface-mapped SAR or ambient lighting effects. For example, real objects in the environment can be texture mapped to support HMD AR tasks (Figure 3b), or ambient lighting effects like a spotlight to highlight physical objects or locations (Figure 3d), or a simulated bright flash to provide feedback when the HMD user captures a photo of the environment (Figure 3g). These can support either user, or both users, depending on the context. For example, the spotlight could be to direct the HMD user to a specific object in support of their HMD AR task, or the spotlight can be a way for the HMD user to communicate a spatial location or object to an external user. A surface mapped object like a cereal box could be solely for the HMD user to support their primary activity, or to show a design to an external user for collaboration.

The projector can also *simulate a flat digital display*, whether rectangular on a wall, or mapped to a nearby surface like a table or floor. A simulated display can benefit the HMD user, such as creating a touch GUI on a table to manipulate HMD AR content (Figure 3c). For an external user, a simulated

display can be projected on a nearby wall, for example so the HMD user can project a presentation on a nearby wall while they consult speaking notes rendered only in the HMD (Figure 3h), or both the HMD user and external user could jointly collaborate on a projected whiteboard (Figure 3e).

USAGE SCENARIOS AND APPLICATIONS
AAR can be used in a diverse set of applications to elevate the experiences of the HMD user and external collaborators around them. We explore three general categories, with demos that span our Design Space to demonstrate the range of experiences possible. See the video figure for real-time demonstrations.

Enhanced AR
The two displays create opportunities to expand the utility and visual quality of AR for the HMD user.

Steerable Expanded HMD Field of View — Similar to FoveAR [9], if the virtual scene is too large to fit within the AR HMD's field of view, the projector can function to artificially widen the user's effective view of the scene (Figure 4a). With an actuated projector, utility is further expanded by allowing the user to switch focus to particulars parts of the object, or by locking onto a target object, continually rendering it in the periphery even as they move around. *(Design space: Augment view, HMD user)*.

Secondary Heads-Up Display — A view-dependent render of virtual objects floating around the viewing frustum of the AR HMD can be used to create a heads-up display for the HMD user without interfering with any content within the AR display. For example, a GUI can be created, such as a toolbar, menu, or clipboard, that can expand the users ability to work (Figure 4b). Another use can be for peripheral awareness. For example, if an object of interest is outside the user's current field of view, an arrow can be used to point to in the direction
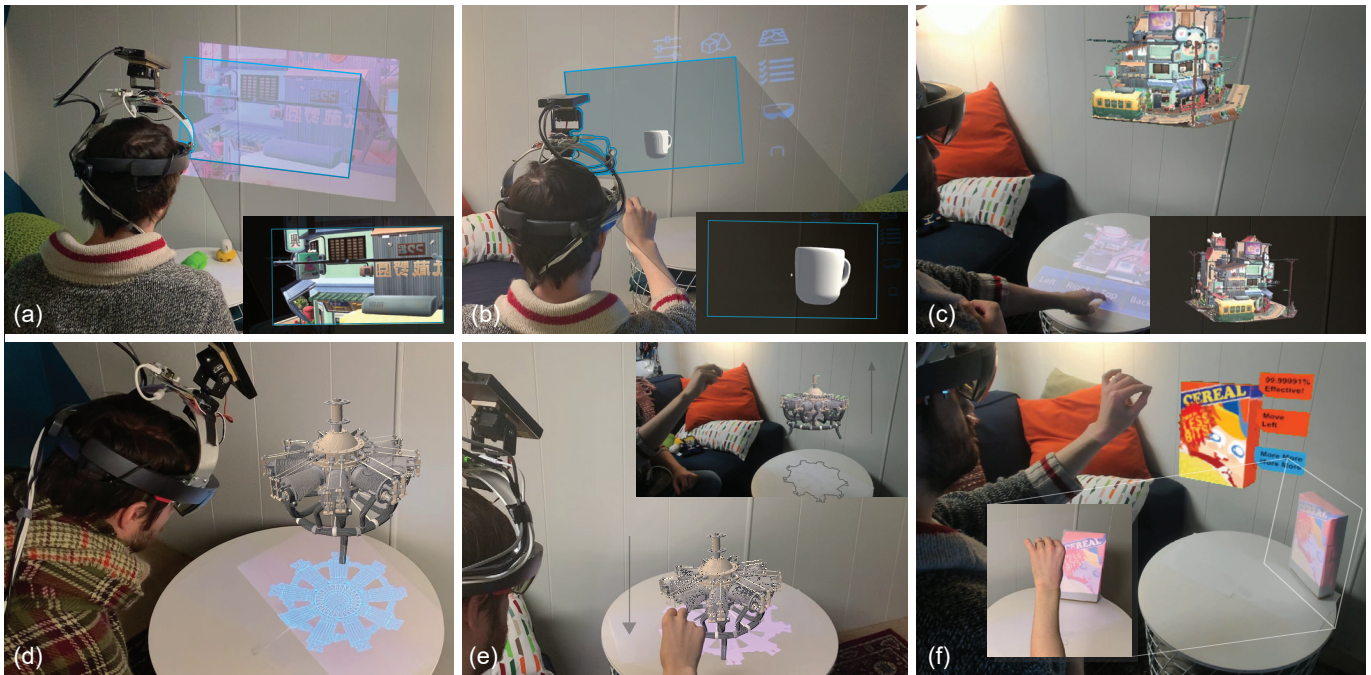
Figure 4. Enhanced AR applications: (a) steerable expanded FoV; (b) GUI heads-Up secondary AR display; (c) surface mapped GUI for peripheral interaction; (d) simulated display in physical environment to show orthographic projection of CAD model; (e) simulating virtual object shadows in physical environment; (f) augmenting a physical box in the environment with a surface mapped texture.

the object, guiding the user to find it. *(Design space: Augment view, HMD user)*.

Secondary Environment Display — A secondary display can be rendered onto nearby surfaces to provide new utility for the HMD user. For example, a GUI can be rendered on a nearby table that allows the user to interact with virtual content while maintaining their focus on a object they are working with (Figure 4c). Another use case is to provide an alternative perspective on a virtual object. For example, an orthographic projection of a plane engine can be rendered so that its schematics can be projected onto a table in front of an engineer. Because the projector is attached to their head, they can move in closer and study its finer details (Figure 4d). *(Design space: Simulate display, HMD user)*.

Simulate Physical Phenomenon in Environment — A simulated shadow of a virtual object can be projected onto a physically realistic location in the environment. For example, when an engineer is examining a virtual 3D model, an inverted shadow cast on the nearby table or wall could communicate its physical height off the surface below (Figure 4e). The projector can simulate the flash of a camera when the HMD user captures a "photo" of the physical environment. This adds additional meaningful feedback and increases realism. Both of these physical phenomenon examples provide some benefit to the external user as well. Object shadows can provide ambient awareness to external users, communicating that the HMD user is editing some type of object. A flash effect lets any external users nearby know their image may have been captured. *(Design space: Augment environment, Both users)*.
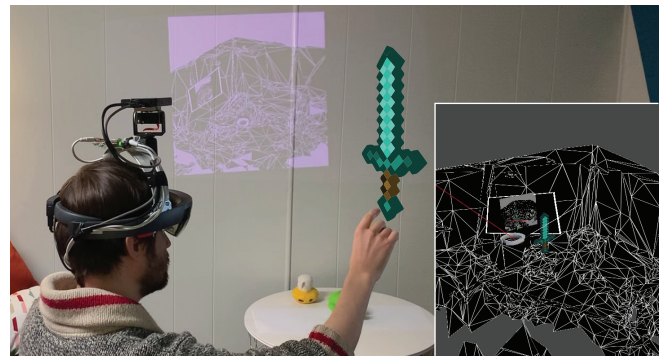

Figure 5. Sharing AR example: an external user viewing the virtual world of HMD user on a simulated display created by the projector.

Physical Object Augmentation — The projector can directly augment a physical object in the space around the user. For example, if a graphic designer is iterating on a product box design, the projector can surface map the virtual box onto a physical prop in the real environment, giving them an idea of what the final product will physically look like (Figure 4f). This would also benefit external users as well, enabling them to monitor progress or critique design choices. *(Design space: Augment environment, HMD user)*.

Sharing AR

The HMD user can utilize the projector to explicitly share virtual content with the external users around them, enabling new forms of interaction and collaboration.

Explicit Content Sharing — The projector can be used to display curated content for external users. A simple example is

Figure 6. Ambient display examples: (a) spotlight as a contextual environment display; (b) mood-lights as ambient environment lighting

when the HMD user and external user jointly interact with a simulated display, like a whiteboard brainstorm or card sorting task. During collaboration, the HMD user may be focusing entirely on the shared projected display. *(Design space: Simulate display, Both users).* A more interesting example is when the HMD user focuses on AR content while external users focus on the projected display. For example, during a meeting, a slideshow presentation could be rendered on an adjacent wall for external users to view, while the HMD user presents using speaker notes rendered in the HMD (Figure 1d). *(Design space: Simulate display, External user).*

Virtual World Camera — The HMD user can share a 2D rendering of the virtual world as viewed from an arbitrary virtual camera position. This would be a simulated display projected on an available surface near the HMD user, enabling external users to see the virtual world as they would from a typical desktop display. The camera could even be controlled by the external user, allowing them to explore the HMD user's virtual environment. (Figure 5) *(Design space: Simulate display, External user).*

Window to Virtual World — When the HMD user is working with a virtual 3D object, they may need to show an external user what it looks like or how they are interacting with it. This can be made possible by projecting onto a wall opposite of the external user and rendering the scene from their perspective. This will create an illusion where the scene objects appear to be at the correct location and have the same shape and size relative to the HMD user (Figure 1c). *(Design space: Augment view, External user).*

Ambient Display

There are circumstances where rendering scene objects or projection mapping is not needed, but an external light source may be desirable for notification, awareness, or navigation. We explore these across two related cases.

Contextual Environment Display — The projector can be used as a controllable spotlight, directed by the HMD user

to highlight objects or locations in their environment. For example, the HMD user could pin an object with light, keeping track of where it is, or they could highlight an object to direct external users to it (Figure 6a). *(Design space: Augment environment, Both users).*

Ambient Environment Lighting — The projector can act as a generalized source of light. Reflecting it off a ceiling could add illumination to the surround environment or artificially adjust its colour temperature. The projector can also be used to enhance multi-media experiences by producing ambient RGB lighting effects. For example, it could be used to enhance a music listening experience or to elevate PC gaming sessions during a live-stream (Figure 6b). *(Design space: Augment environment, Both users).*

Persistent AR

Its likely that the HMD user may wish to remove the HMD from time-to-time to take break, have a snack, or other real world tasks. Current HMDs have no self-contained way to transition from a virtual AR task to one in the real world. The projector can enable such a transition for limited, but potentially useful, interaction and awareness of the virtual world. When the HMD user removes the headset, the projector can create an ad hoc inside-out SAR environment. For example, if the HMD user is watching a video in the virtual world, then removes the HMD to place it on a table, the projector can automatically transition the video to a simulated display on an adjacent wall (Figure 2). In this example, the HMD user becomes an external user in a unique "external users only" AAR usage context. *(Design space: Simulate display, External user).*

ACTUATED PROJECTOR FOR AN AR HMD

Prototyping the AAR design space with real applications requires precise control over the projector's movement relative to the AR HMD. We present a novel approach to calibrate the positions and offsets for the projector and pan-tilt structure.

Hardware

A Celluon PicoBit laser projector [5] (Fig. 7a) is mounted onto two linked servo motors that form the pan-tilt mechanism (Fig. 7b). An Arduino Pro Mini ATmega328 acts as their controller (Fig. 7c). Both the Arduino and pan-tilt mechanism are attached to a custom aluminum mounting bracket which is bolted onto a Hololens V1 [3] (Fig. 7d).

The projector has a resolution of 1280×720 pixels with a brightness of 63 ANSI lumens. The laser projection module is infinite focus, which eliminates the need to manually adjust the projector's focal plane. Considering a left-handed coordinate frame, the Hololens points along positive Z, the top servo tilts the projector along its X-axis, and the bottom servo pans the projector along its Y-axis. A Kuman [2] 17 Kg high torque 270° motor is used for the bottom servo and a DFRobot DSS-M15 180° motor is used for the top. Both servos are hard-limited in range to ensure no damage to the Arduino or projector can occur. The bottom servo is limited to a range between 15° and 255°. The top servo is limited to a range between 30° and 125°. When the servos are set to 135° and
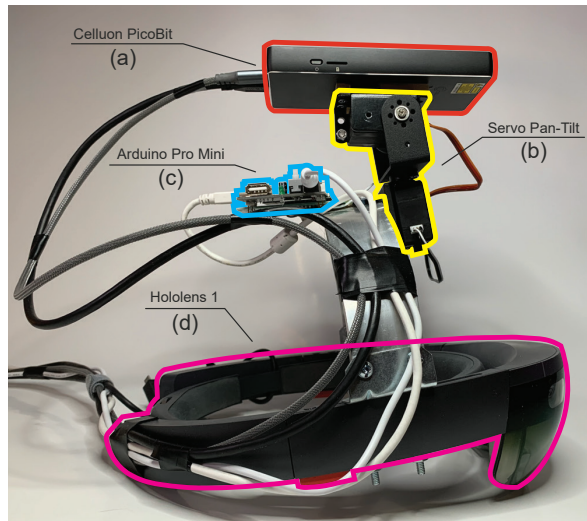
Figure 7. Actuated Projector HMD: (a) Celluon PicoBit laser projector; (b) Kuman and DFRobot pan-tilt servos; (c) Arduino Pro Mini; (d) Microsoft Hololens V1.
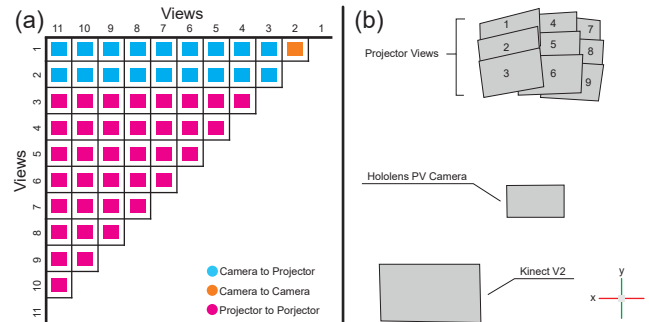

Figure 8. View-pair correlation matrix and view frustums: (a) point correspondence pairs between the camera views (1 and 2) and projector views (3-11); (b) reconstructed view frustums for the projector, Kinect, and Hololens viewed in the direction of the Z-axis.

$90°$ respectively, we consider the projector to be in its default position, pointing forward along the Z-axis with the Hololens.

Automatic Calibration

In order to enable the range of experiences outlined in the usage scenarios, a one-time calibration is required for the hardware. Steerable projector systems have previously discussed calibration techniques for projector-camera units through the physical repositioning of a checkerboard in an environment [47, 55]. However, these current approaches are labour-intensive and the internal geometric structure of the actuators are not fully captured. Another concern is the coupling between the projector and the HMD device, where the projector's world pose changes with the transformation of the HMD.

We frame projector and camera pose reconstruction as a Structure from Motion (SfM) [45] problem, which is commonly used in large scale computer vision problems, like photogrammetry. The reconstructed poses are used in an optimization routine over a Denavit-Hartenberg (DH) [17] parameterization of the servos' kinematic chain, finding their axes of rotation and offsets. The resulting geometric relationships can be used for precise movement control.

An SfM pipeline typically has of four phases: (1) data acquisition, (2) feature point detection, (3) putative point correspondence matching, and (4) pose and point cloud reconstruction [45]. We adapt this pipeline for an expanded set of correspondences required by our pose reconstruction problem that utilizes projector-projector, projector-camera, and camera-camera correspondence pairs.

Data acquisition — Previous work has used structured light to create dense camera-projector maps using Gray codes [28]. In contrast, we adapt the approach from Yamazaki et al. [58] which combines Gray codes with phase-shifting sinusoidal codes to increase the sub-pixel accuracy of the resulting maps. During data aquistion, the projector is repositioned 9 times

using the pan-tilt servos, moving in increments of $10°$ for the X-axis and $15°$ for the Y-axis to create a 3×3 grid of projector views (Fig. 8b). For each of these views, the built in Photo-Video (PV) camera from the Hololens V1 and a Kinect V2 camera is used to capture the projected structured light. We use a Kinect camera to capture a wider FoV of the scene, which allows larger projector movement during acquisition. We do not use depth data in our pipeline, other wide FoV RGB cameras could also work. The total number of views is 11, (9 repositioned projector views and 2 stationary cameras views), resulting in 18 camera-projector pairs with a total of 1134 captured images across all structured light sequences.

Putative Correspondences — Calibrating with structured light does not require feature point detection to build view correspondences. Instead a dense point-to-point correspondence between the camera and projector can be achieved by decoding the captured structured light during data acquisition. By utilizing both a forward mapping (a pixel-point in the image to sub-pixel in the projector) and a reverse mapping (pixel-point in projector to camera pixel), a complete set of putative correspondence pairs can be created for all 11 represented views. The complete set of view-pair regions is 55, composed of 18 camera-projector, 1 camera-camera, and 36 projector-projector putative correspondence pairs (Fig. 8a).

Structure from Motion — Using the complete set of correspondence pairs, we extend the open source implementation of OpenMVG [46] to account for the different view pair regions discussed above. We solve for all 11 views using sequential SfM [45] with AContrario RANSAC [44]. The reconstruction process solves for: (1) the extrinsic (i.e. poses) and intrinsic parameters of the views, and (2) the 3D point cloud of the environment (Fig. 8b). A 25mm × 25mm checkerboard is used to solve the scale ambiguity of the resulting reconstruction, no repositioning is required. In actuality, any known point-to-point distance in the scene could be used instead, the checkerboard is used for convenience, it is not a requirement for our calibration. The final RMSE is 0.51mm on 2.3 million residuals taking 131 seconds.

Optimization Solver to Recover Pan-Tilt Geometry

We take the 9 transformation matrices representing the views of the projector when actuated, and solve for the
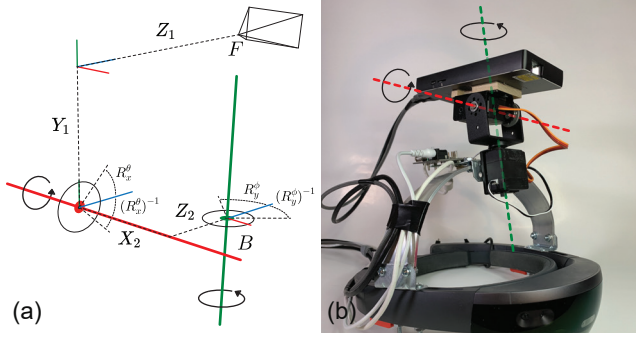
Figure 9. Projector mount geometry: (a) DH parameters and pan-tilt axes forming a kinematic chain from projector (F) to base servo (B). (b) depiction of the pan-tilt axes.

Figure 10. Movement of the projected image (red dots) to locations around the HMD (blue dots) in Unity3D: (a) left; (b) centre; (c) right.

kinematic chain and rotation axes of the servo motors. A Denavit–Hartenberg (DH) parameter representation is used to represent this structure relative to the projector's frustum (F). Each rotation axis is represented by a rotation matrix (R_X and R_Y) that is parameterized by a single rotation value θ and ϕ in radians (Fig. 9a). A DH parameter is 4 $\times$ 4 transformation matrix defined by a translational offset and rotational displacement along a single axis (e.g. the DH parameter for the transformation along the z-axis is $Z = \{z^t, z^\theta\}$ where z^t represents translational offset and z^θ rotation around the z-axis). In this way, a homogeneous point ($\hat{x}$) in the projector's coordinate space (F) can be transformed to the coordinate space of the base servo (B) by

$$\hat{x}_B = R_Y^\phi Z_2 X_2 R_X^\theta Y_1 Z_1 \hat{x}_F \quad (1)$$

where the unknown parameters are: (1) the rotation around the x- and y-axis (R_X^θ and R_Y^ϕ), (2) the DH parameters (Z_1 and Y_1) describing the link between projector's view to the x-axis, and (3) the DH parameters (X_2 and Z_2) describing the link between the x- and y-axis, which is the base servo. The knowns are the point observations ($\mathcal{X}_{obs}$) and the 3 $\times$ 3 grid of projector view transformations.

Further, if we consider the relationship between the projector's centre transformation matrix (T_5) with the other 8 surrounding transformation matrices (T_i) in the 3 $\times$ 3 grid of projector views (Fig. 8b), we can use a variation of equation 1 to relate a homogeneous world point ($\hat{x}$) observed from the centre projector view to any other projector view through the following equality constraint:

$$R_Y^\phi Z_2 X_2 R_X^\theta Y_1 Z_1 T_i \hat{x} = Z_2 X_2 Y_1 Z_1 T_5 \hat{x} \quad (2)$$

where we consider the centre view (T_5) to be the calibrated view and default projector transformation.

With the equality outlined in equation 2, a cost function is constructed to solve for the unknowns, enumerated as: $\Phi = \{z_1^t, z_1^\theta, y_1^t, y_1^\theta, x_2^t, x_2^\theta, z_2^t, z_2^\theta, \theta, \phi\}$. The cost function contains two parts, one describing the constraint for the servo rotating around the x-axis ($f_1^{(i)}(\hat{x})$) and one describing the constraint for the servo rotating around the y-axis ($f_2^{(i)}(\hat{x})$), where both are parameterized by the view transformation i:

$$f_1^{(i)}(\hat{x}) = Y_1 Z_1 T_5 \hat{x} - R_x^\theta(i) Y_1 Z_1 T_i \hat{x} \quad (3)$$
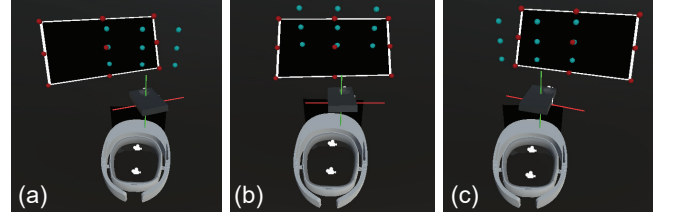
$$f_2^{(i)}(\hat{x}) = Z_2 X_2 Y_1 Z_1 T_5 \hat{x} - R_y^\phi(i) Z_2 X_2 R_x^\theta(i) Y_1 Z_1 T_i \hat{x} \quad (4)$$

Given that the rotations around the x- and y-axis are symmetrical, we need to determine whether the inverse of the rotation matrix is appropriate based on what view transformation is used within the equality constraint. We define a function $R_x^\theta(i)$ and $R_y^\phi(i)$ to provide the correct rotation matrix based on whether the transformation is from the top, bottom, left, or right column of the 3 $\times$ 3 grid of projector views.

$$R_x^\theta(i) = \begin{cases} R_x^\theta & i \in T_{1,*} \\ (R_x^\theta)^{-1} & i \in T_{3,*} \\ I & \text{otherwise} \end{cases} \quad R_y^\phi(i) = \begin{cases} R_y^\phi & i \in T_{*,1} \\ (R_y^\phi)^{-1} & i \in T_{*,3} \\ I & \text{otherwise} \end{cases} \quad (5)$$

If we consider $\mathbf{x} \in \mathcal{X}_{obs}$ an n-dimensional vector of homogeneous world point observations and we utilize equations 3, 4, and 5, we can minimize the following through a non-linear least square Levenberg-Marquardt [37, 40] trust region [16] method with Cauchy loss:

$$\underset{\Phi}{\text{argmin}} \frac{1}{2} \left\| \left[f_1^{(1)}(\mathbf{x}), ..., f_1^{(9)}(\mathbf{x}), f_2^{(1)}(\mathbf{x}), ..., f_2^{(9)}(\mathbf{x}) \right]^\top \right\|^2 \quad (6)$$

The resulting solution is used to reconstruct the projector's centre view (T_5) and the servos' kinematic chain relative to the Hololens PV camera. A final step rectifies all reconstructed transforms to the origin point of the Hololens device. We use these final transforms in our software toolkit.

AAR SOFTWARE TOOLKIT

To assist in the creation and design of AAR experiences, we developed an open source software toolkit[1] for Unity3D [7]. Our toolkit works in conjunction with the Microsoft Mixed Reality Toolkit (MRTK) and provides a streamlined development experience for creating and iterating on AAR design concepts.

Toolkit functionality is divided into two parts: (1) a native C++ plugin encapsulating the servo, projector, and calibration controllers; (2) a C# unity package that interfaces with our native plugin and provides high-level APIs for AAR services. Each service is associated with a Unity Prefab Asset, encapsulating 3D object information, editor metadata, and the toolkit's C# scripting components. These can be easily drag-and-dropped into an active scene. The toolkit is able to simulate all the functionality of the real hardware, enabling development without the need for a physical device.

Projector Control — This provides interfaces to manipulate the projector and servo hardware, giving developers high-level control over where the projector is pointing in the virtual
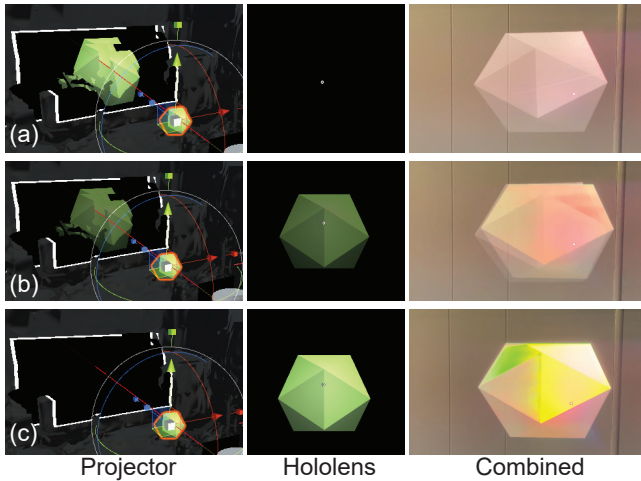
---

[1] https://github.com/exii-uw/AARToolkit

Figure 11. Blending a virtual object between the projector and Hololens at: (a) 100% and 0%; (b) 50% and 50%; (c) 0% and 100%.

Figure 12. AR and project field of view (FoV) with pan-tilt servo range: a) vertical FoV and tilt range, b) horizontal FoV and pan range.

environment. This can be to locations relative to the frustum of the Hololens (Fig. 10), or can be to specific points in the world coordinate system. The toolkit uses the servos' calibrated axes to calculate all necessary rotations for both software simulation and hardware control.

Spatial Awareness — A unified 3D model of the environment is provided by the Hololens. This is represented inside Unity as mesh object from which ray intersection and object collision is possible. Similar to RoomAlive [28], our toolkit reconstructs the planes in the scene and provides them with meaningful semantic names (i.e. floor, wall, ceiling, and unknown) based on their surface normal, size and position.

Rendering Spatial Content
We built a rendering engine that is able to handle different configurations of the AR display and projector.

Rendering as an External Display — Static textures, shaders, or videos are rendered and displayed in the real-world environment. This is useful for both situations where projection mapping is not needed, such as ambient lighting or using the projector as a spotlight. With surface mapping enabled, the view can be used as a secondary display, showing contextual information or an interface to interact with, all rendered onto a physical surface in the environment. We developed a custom material object to abstract all possible combinations.

Rendering for the HMD User — The projector can be treated as a camera in the virtual environment, enabling it to render virtual scene objects from multiple different viewing locations. When rendering for the HMD user, the developer must decide how much of each object is visible in each display. This is realized through a custom shader post-processing stack that can be accessed through a single C# script attached to the object being blended. The developer can individually assign how much the object is visible in each display, allowing detailed control (see Fig. 11).
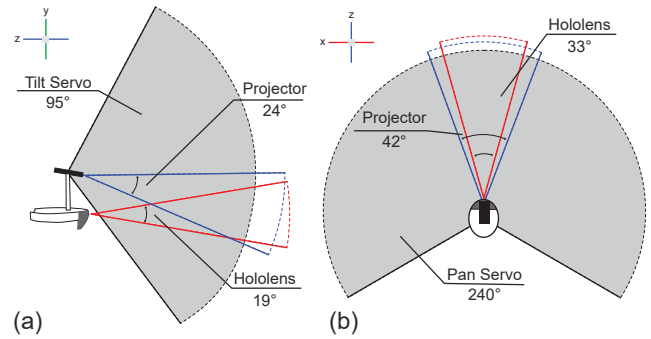
Rendering for External Users — Our toolkit gives developers access to specific projection mapping functionality, allowing a perspectively correct view form any external location.

Performance and Quality Analysis
In this section, we report on toolkit performance for the rendering pipeline and servo control. Note the calibration process is only for initial hardware setup, it has no impact on runtime performance. Appendix A in the supplementary materials also reports on an initial user study with AR developers to evaluate the toolkit API and their general understanding of the AAR concept.

Field of view — The vertical FoV for the projector and Hololens is $24°$ and $19°$ (Fig. 12a), while the horizontal FoV is $42°$ and $33°$ (Fig. 12b).

Servo control — The servos are able to relocate the projector within a range of $95°$ vertically and $240°$ horizontally. There is a 74ms latency from the moment an action occurs in the toolkit to the physical movement of the servo. The servo takes 180ms for every $60°$ rotated.

Rendering performance and latency — The toolkit updates at a frame rate of 60 FPS. The point-to-point latency from the moment the toolkit serves a rendered frame to the moment the projector displays that same frame is 116ms. This latency results from the projector's hardware and firmware design.

DISCUSSION
Our AAR concept is highly dependent on hardware capabilities and usability. In this section, we enumerate current limitations with possible solutions, and present the most compelling future enhancements.

AAR Hardware and Usability
Our proof-of-concept prototype was adequate for demonstrating applications, but there remain aspects that could be refined.

Projector Latency and Refresh Rate — The PicoBit projector scan process is 60-hertz interlaced with no persistence, its input to output latency is reported as 116ms. This can cause notable image lag during movement, even when the projector is fixed and not moving relative to the HMD. This problem can also be observed in commercial HMDs, where the rendering

pipeline, and latency and refresh rate of the HMD can cause the image to lag, creating a mismatch between the in-game state and what the user can see. Improving one more sources of lag, like a 90 or 120 Hz projector with low latency output, will help minimize mismatch and improve the overall experience.

Servo Accuracy — We actuate the projector with servos commonly used by model and electronic hobbyist. While sufficient for small robots, we found their movement sometimes inconsistent, and lacking some precision. This is especially noticeable when projecting over a large distance (>5m): if the servo is off by $1°$, that equates to 9cm displacement of projected content. To help alleviate this, our calibration process recovers the "real" degrees moved, and scales output in the toolkit accordingly. Further improvements could be achieved with higher-accuracy actuators.

Projected Image Stability — Related to latency and servo accuracy, is projected image stability. Currently, the toolkit knows where the HMD is relative to world geometry, which allows the calibrated pan-tilt mechanism to compensate for any head movement in the world and can keep the projected image steady for different positions, movements, and angles. However, tracking and environment scanning is only as good as what the HMD device provides, and more critically, the servos are limited by their rotation speed. Projector output latency further compounds this. In most cases, instability is not very perceptible, such as when augmenting the view of the HMD user where the projected image is anchored to the HMD frame of reference. In other cases, like surface mapping a 3D object or projecting onto a distant display for an external user, projector instability can be noticeable. Although this was acceptable for our usage, faster and more accurate servos, along with improved refresh rates and low latency projectors, could further improve image stability and the user experience.

Weight and Ergonomics — The original Hololens V1 is notorious for being uncomfortable to wear for long periods of time. It weighs 579 grams which is ideally distributed around the crown of the head. Our mounting bracket, servos, and pico projector add an additional 585 grams, resulting in just more than 1kg total. It is not to the point where a pulley counter balance is needed, but it can cause noticeable discomfort for periods longer than 10 or 20 minutes. Improvements in weight distribution and minimization of the hardware through OEM LBS projection engines [4] and custom circuit designs will reduce the weight and improve comfort.

Projector Eye Safety — Our PicoBit is a laser beam scanning (LBS) projector [22], which has a potential safety issue due to the "IEC-60825-1: Class 3R Laser" classification: prolonged direct eye contact into the beam can be harmful [1]. In practice, staring into projectors using other technologies, like LCoS and DLP, should generally be avoided too. Methods have been proposed to automatically block projected light from entering people's eyes [30], and these could be incorporated into our system. A key advantage for LBS in an ad hoc SAR setting such as ours is infinite focus, the image is crisp no matter where its pointed within an environment.

System Extensions
We enumerate on some possible future extensions.

External user tracking — Tracking external users in the space around the HMD user is a potentially exciting avenue for future work. The expanded utility could open up interesting new experiences for AAR interaction, like adding motion parallax to a view-dependent projection or creating new immersive gaming experiences.

Improved Toolkit Integration — The toolkit is specific to a Hololens V1 with a head-mounted projector. Extending it to work with other AR HMDs is one direction for future work that will broaden its generalizability.

Future Work
In its current form, the AAR hardware and toolkit allow a user to freely move around a physical space and augment their environment through either the projector, AR display, or both. Here, we discuss two topics for future work.

Hardware minimization — As outlined above, there are many directions to improve the hardware implementation. These include using custom projection engines with low latency output, to smaller and faster servo motors. Minimizing and improving hardware could reveal new interaction modalities and alternative mounting locations on the HMD.

User-centric Studies — AAR specific input and interaction could be further extended and investigated, and the user-centric impacts of the system pipeline could be further explored. For example. specialized interaction techniques could be created and evaluated in experiments for critical tasks like pointing and selection. Studies could investigate the perceived user affordances of imagery when presented on the projector compared to the HMD. Another topic are studies that look at specific asymmetric interactions between an external user and HMD user, during co-located collaborative tasks.

CONCLUSION
We presented the concept of Augmented Augmented Reality for a wearable augmented reality HMD and an actuated head-mounted projector. We constructed a working hardware and software system, calibrated through a modified structure from motion algorithm and a novel optimization solver to reconstruct the kinematic chain and rotation axes of the actuators. Our Unity3D toolkit encapsulates a set of high-level functionality for the iteration of AAR experiences. We hope our work inspires more investigations into combining different AR devices in new ways, and the pursuit of ever more immersive experiences that can still remain grounded in our physical and social world.

REFERENCES

[1] 2008. Safety of laser products. In *IEC TR60825-3*. International Electrotechnical Commission, Chapter Part 3: Guidance for laser displays and shows. `https://webstore.iec.ch/publication/3598`

[2] 2020. Kuman 17Kg 270 Degree Metal Gear Digital Servo with U Bracket & Side mount for RC Robot Helicopter Airplane Car Boat KY72-1,Servos. `http://www.kumantech.com/kuman-17kg-270-degree-metal-gear-digital-servo-with-u-bracket-amp-side-mount-for-rc-robot-helicopter-airplane-car-boat-ky72-1_p0398.html`. (2020). (Accessed on 04/12/2020).

[3] 2020. Microsoft HoloLens | Mixed Reality Technology for Business. `https://www.microsoft.com/en-us/hololens`. (2020). (Accessed on 02/10/2020).

[4] 2020. MicroVision Technology. `https://www.microvision.com/technology/`. (2020). (Accessed on 05/01/2020).

[5] 2020. PicoBit – Celluon Inc. `https://celluon.com/picobit/`. (2020). (Accessed on 04/12/2020).

[6] 2020. Spatial Computing for Enterprise | Magic Leap. `https://www.magicleap.com/`. (2020). (Accessed on 02/10/2020).

[7] 2020. Unity Real-Time Development Platform | 3D, 2D VR & AR Visualizations. `https://unity.com/`. (2020). (Accessed on 04/14/2020).

[8] Patrick Baudisch, Nathaniel Good, and Paul Stewart. 2001. Focus plus context screens: combining display technology with visualization techniques. In *Proceedings of the 14th annual ACM symposium on User interface software and technology - UIST '01*. ACM Press, New York, New York, USA, 31. DOI: `http://dx.doi.org/10.1145/502348.502354`

[9] Hrvoje Benko, Eyal Ofek, Feng Zheng, and Andrew D. Wilson. 2015. FoveAR: Combining an Optically See-Through Near-Eye Display with Spatial Augmented Reality Projections. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*. ACM Press, New York, New York, USA, 129–135. DOI: `http://dx.doi.org/10.1145/2807442.2807493`

[10] Mark Billinghurst and Hirokazu Kato. 2002. Collaborative augmented reality. *Commun. ACM* 45, 7 (2002), 64–70.

[11] Oliver Bimber and Bernd Frohlich. 2002. Occlusion shadows: using projected light to generate realistic occlusion effects for view-dependent optical see-through displays. In *Proceedings. International Symposium on Mixed and Augmented Reality*. IEEE Comput. Soc, 186–319. DOI: `http://dx.doi.org/10.1109/ISMAR.2002.1115088`

[12] Oliver Bimber and Ramesh Raskar. 2005. *Spatial Augmented Reality Merging Real and Virtual Worlds*. 1–392 pages.

[13] Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris. 2011. Code Space: Touch + Air Gesture Hybrid Interactions for Supporting Developer Meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS '11*, Vol. 16. ACM Press, New York, New York, USA, 212. DOI: `http://dx.doi.org/10.1145/2076354.2076393`

[14] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. 2019. Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, New York, New York, USA, 1–28. DOI: `http://dx.doi.org/10.1145/3290605.3300792`

[15] Andreas Butz, T. Hollerer, Steven Feiner, B. MacIntyre, and Clifford Beshers. Enveloping Users and Computers in a Collaborative 3D Augmented Reality. In *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)*. IEEE Comput. Soc, 35–44. DOI: `http://dx.doi.org/10.1109/IWAR.1999.803804`

[16] Gerald A Byrd, Richard H and Schnabel, Robert B and Shultz. 1987. A trust region algorithm for nonlinearly constrained optimization. *SIAM J. Numer. Anal.* 24 (nov 1987), 1152—-1170.

[17] J. Denavit and R.S. Hartenberg. 1955. A Kinematic Notation for Lower-Pair Mechanisms Based on Matrices. *Journal of Applied Mechanics* 22, 2 (1955), 215–221.

[18] Andreas Rene Fender, Hrvoje Benko, and Andy Wilson. 2017. MeetAlive: Room-Scale Omni-Directional Display System for Multi-User Content and Control Sharing. In *Proceedings of the Interactive Surfaces and Spaces on ZZZ - ISS '17*. ACM Press, New York, New York, USA, 106–115. DOI: `http://dx.doi.org/10.1145/3132272.3134117`

[19] Çağlar Genç, Shoaib Soomro, Yalçın Duyan, Selim Ölçer, Fuat Balcı, Hakan Ürey, and Oğuzhan Özcan. 2016. Head Mounted Projection Display & Visual Attention: Visual attentional processing of head referenced static and dynamic displays while in motion and standing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1538–1547. DOI: `http://dx.doi.org/10.1145/2858036.2858449`

[20] Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. 2015. MultiFi: Multi-Fidelity Interaction with Displays On and Around the Body. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, Vol. 2015-April. ACM Press, New York, New York, USA, 3933–3942. DOI: `http://dx.doi.org/10.1145/2702123.2702331`

[21] Uwe Gruenefeld, Tim Claudius Stratmann, Lars Prädel, and Wilko Heuten. 2018. MonoculAR: A radial light

display to point towards out-of-view objects on augmented reality devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI '18*. ACM Press, New York, New York, USA, 16–22. DOI: `http://dx.doi.org/10.1145/3236112.3236115`

[22] Karl Guttag. 2020. Celluon Laser Beam Scanning Projector Technical Analysis – Part 1 – Karl Guttag on Technology. `https://www.kguttag.com/2015/06/01/celluon-laser-beam-scanning-projector-part-1/`. (2020). (Accessed on 05/01/2020).

[23] Takumi Hamasaki, Yuta Itoh, Yuichi Hiroi, Daisuke Iwai, and Maki Sugimoto. 2018. HySAR: Hybrid Material Rendering by an Optical See-Through Head-Mounted Display with Spatial Augmented Reality Projection. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (apr 2018), 1457–1466. DOI: `http://dx.doi.org/10.1109/TVCG.2018.2793659`

[24] Chris Harrison, Hrvoje Benko, and Andrew Wilson. 2011. OmniTouch: Wearable Multitouch Interaction Everywhere. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*. ACM Press, New York, New York, USA, 441. DOI:`http://dx.doi.org/10.1145/2047196.2047255`

[25] Jeremy Hartmann, Christian Holz, Eyal Ofek, and Andrew D. Wilson. 2019. RealityCheck: Blending Virtual Environments with Situated Physical Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, New York, New York, USA, 1–12. DOI: `http://dx.doi.org/10.1145/3290605.3300577`

[26] Hong Hua, L.D. Brown, and Chunyu Gao. 2004. Scape: Supporting Stereoscopic Collaboration in Augmented and Projective Environments. *IEEE Computer Graphics and Applications* 24, 1 (jan 2004), 66–75. DOI: `http://dx.doi.org/10.1109/MCG.2004.1255811`

[27] Hong Hua, Axelle Girardot, Chunyu Gao, and Jannick P Rolland. 2000. Engineering of head-mounted projective displays. *Applied Optics* 39, 22 (aug 2000), 3814. DOI: `http://dx.doi.org/10.1364/AO.39.003814`

[28] Brett Jones, Lior Shapira, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, and Nikunj Raghuvanshi. 2014. RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-camera Units. In *Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14*. ACM Press, New York, New York, USA, 637–644. DOI: `http://dx.doi.org/10.1145/2642918.2647383`

[29] Daniel Kade, Kaan Akşit, Hakan Ürey, and Oğuzhan Özcan. 2015. Head-mounted mixed reality projection display for games production and entertainment. *Personal and Ubiquitous Computing* 19, 3-4 (jul 2015), 509–521. DOI: `http://dx.doi.org/10.1007/s00779-015-0847-y`

[30] Bonifaz Kaufmann and Martin Hitz. 2011. Eye-Shield: Protecting Bystanders from Being Blinded by Mobile Projectors. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '11)*. Association for Computing Machinery, New York, NY, USA, 31–34. DOI: `http://dx.doi.org/10.1145/2076354.2076359`

[31] Ryugo Kijima and Takeo Ojika. 1997. Transition between virtual environment and workstation environment with projective head mounted display. In *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*. IEEE Comput. Soc. Press, 130–137. DOI: `http://dx.doi.org/10.1109/VRAIS.1997.583062`

[32] Pascal Knierim, Markus Funk, Thomas Kosch, Anton Fedosov, Tamara Müller, Benjamin Schopf, Marc Weise, and Albrecht Schmidt. 2016. UbiBeam++: Augmenting interactive projection with head-mounted displays. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction - NordiCHI '16*, Vol. 23-27-Octo. ACM Press, New York, New York, USA, 1–6. DOI:`http://dx.doi.org/10.1145/2971485.2996747`

[33] Pascal Knierim, Steffen Maurer, Katrin Wolf, and Markus Funk. 2018. Quadcopter-Projected In-Situ Navigation Cues for Improved Location Awareness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, Vol. 2018-April. ACM Press, New York, New York, USA, 1–6. DOI:`http://dx.doi.org/10.1145/3173574.3174007`

[34] Bernard C. Kress. 2020a. *Optical Architectures for Augmented-, Virtual-, and Mixed-Reality Headsets*. SPIE PRESS, Bellingham, Washington. 270 pages.

[35] Bernard C. Kress. 2020b. Waveguide Combiners. In *Optical Architectures for Augmented-, Virtual-, and Mixed-Reality Headsets*. SPIE PRESS, Bellingham, Washington, Chapter 14, 127–150.

[36] David M Krum, Evan A Suma, and Mark Bolas. 2012. Augmented reality using personal projection and retroreflection. *Personal and Ubiquitous Computing* 16, 1 (jan 2012), 17–26. DOI: `http://dx.doi.org/10.1007/s00779-011-0374-4`

[37] Kenneth Levenberg. 1944. A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.* 2, 2 (jul 1944), 164–168. DOI: `http://dx.doi.org/10.1090/qam/10666`

[38] Mayra Donaji Barrera Machuca, Winyu Chinthammit, Yi Yang, and Henry Duh. 2014. 3D mobile interactions for public displays. In *SIGGRAPH Asia 2014 Mobile Graphics and Interactive Applications on - SA '14*. ACM Press, New York, New York, USA, 1–4. DOI: `http://dx.doi.org/10.1145/2669062.2669074`

[39] Andrew Maimone, Xubo Yang, Nate Dierk, Andrei State, Mingsong Dou, and Henry Fuchs. 2013. General-purpose telepresence with head-worn optical see-through displays and projector-based lighting. In

2013 IEEE Virtual Reality (VR). IEEE, 23–26. DOI:
`http://dx.doi.org/10.1109/VR.2013.6549352`

[40] Donald W Marquardt. 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J. Soc. Indust. Appl. Math.* 11, 2 (jun 1963), 431–441. DOI:
`http://dx.doi.org/10.1137/0111030`

[41] Daniel C. McFarlane and Steven M. Wilder. 2009. Interactive dirt: Increasing mobile work performance with a wearable projector-camera system. In *Proceedings of the 11th international conference on Ubiquitous computing*. ACM, New York, NY, USA, 205–214. DOI:
`http://dx.doi.org/10.1145/1620545.1620577`

[42] Alexandre Millette and Michael J. McGuffin. 2016. DualCAD: Integrating Augmented Reality with a Desktop GUI and Smartphone Interaction. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, 21–26. DOI:
`http://dx.doi.org/10.1109/ISMAR-Adjunct.2016.0030`

[43] Pranav Mistry and Pattie Maes. 2009. SixthSense: A Wearable Gestural Interface. In *ACM SIGGRAPH ASIA 2009 Art Gallery & Emerging Technologies: Adaptation (SIGGRAPH ASIA '09)*. Association for Computing Machinery, New York, NY, USA, 85. DOI:
`http://dx.doi.org/10.1145/1665137.1665204`

[44] Lionel Moisan, Pierre Moulon, and Pascal Monasse. 2012. Automatic homographic registration of a pair of images, with a contrario elimination of outliers. *Image Processing On Line* 2 (2012), 56–73.

[45] Pierre Moulon, Pascal Monasse, and Renaud Marlet. 2012. Adaptive Structure from Motion with a Contrario Model Estimation. In *Proceedings of the Asian Computer Vision Conference (ACCV 2012)*. Springer Berlin Heidelberg, 257–270. DOI:
`http://dx.doi.org/10.1007/978-3-642-37447-0_20`

[46] Pierre Moulon, Pascal Monasse, Romuald Perrot, and Renaud Marlet. 2016. Openmvg: Open multiple view geometry. In *International Workshop on Reproducible Research in Pattern Recognition*. Springer, 60–74.

[47] P M Ngan and R J Valkenburg. 1995. *Calibrating a pan-tilt camera head*. Technical Report.
`http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.6.3733`

[48] Danakorn Nincarean Eh Phon, Mohamad Bilal Ali, and Noor Dayana Abd Halim. 2014. Collaborative augmented reality in education: A review. In *2014 International Conference on Teaching and Learning in Computing and Engineering*. IEEE, 78–83.

[49] Claudio Pinhanez. 2001. The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 2201. Springer Verlag, 315–331. DOI:
`http://dx.doi.org/10.1007/3-540-45427-6_27`

[50] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. 1998. The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques - SIGGRAPH '98 (SIGGRAPH '98)*. ACM Press, New York, New York, USA, 179–188. DOI:`http://dx.doi.org/10.1145/280814.280861`

[51] Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay. 2001. Shader Lamps: Animating Real Objects With Image-Based Illumination. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*. 89–102. DOI:
`http://dx.doi.org/10.1007/978-3-7091-6242-2_9`

[52] Jun Rekimoto and Masanori Saitoh. 1999. Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99*. ACM Press, New York, New York, USA, 378–385. DOI:
`http://dx.doi.org/10.1145/302979.303113`

[53] Enrico Rukzio and Paul Holleis. 2016. Projector Phone Interactions: Design Space and Survey. *Workshop on coupled display visual interfaces at AVI* April (2016).
`https://www.researchgate.net/publication/228370421`

[54] Marcos Serrano, Barrett Ens, Xing-Dong Yang, and Pourang Irani. 2015. Gluey: Developing a Head-Worn Display Interface to Unify the Interaction Experience in Distributed Display Environments. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15*. ACM Press, New York, New York, USA, 161–171. DOI:
`http://dx.doi.org/10.1145/2785830.2785838`

[55] Andrew Wilson, Hrvoje Benko, Shahram Izadi, and Otmar Hilliges. 2012. Steerable Augmented Reality with the Beamatron. In *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*. ACM Press, New York, New York, USA, 413. DOI:`http://dx.doi.org/10.1145/2380116.2380169`

[56] Christian Winkler, Julian Seifert, David Dobbelstein, and Enrico Rukzio. 2014. Pervasive Information through Constant Personal Projection: The Ambient Mobile Pervasive Display (AMP-D). In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, 4117–4126. DOI:
`http://dx.doi.org/10.1145/2556288.2557365`

[57] Robert Xiao and Hrvoje Benko. 2016. Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1221–1232. DOI:
`http://dx.doi.org/10.1145/2858036.2858212`

[58] Shuntaro Yamazaki, Masaaki Mochimaru, and Takeo Kanade. 2011. Simultaneous self-calibration of a

projector and a camera using structured light. In *CVPR 2011 WORKSHOPS*. IEEE, 60–67. DOI:
`http://dx.doi.org/10.1109/CVPRW.2011.5981781`

[59] Jianlong Zhou, Ivan Lee, Bruce Thomas, Roland Menassa, Anthony Farrant, and Andrew Sansome. 2012. In-Situ Support for Automotive Manufacturing Using Spatial Augmented Reality. *International Journal of Virtual Reality* 11, 1 (jan 2012), 33–41. DOI:
`http://dx.doi.org/10.20870/IJVR.2012.11.1.2835`

APPENDIX A: DEVELOPER TOOLKIT STUDY

The goals of this open ended remote study were to verify whether developers understood the concept of ARR and whether the toolkit provided adequate tools to create AAR experiences. Over 7 days, invited AR developers familiarized themselves with the AAR concept and toolkit, and created their own AAR applications. They did not have access to our hardware, but the toolkit simulates the HMD and projector inside Unity3D.

We recruited 6 developers through social media websites, ages 20 to 38, 2 male and 1 female. Three become unresponsive after initial on-boarding, and so where dropped from the study. All reported experience with AR and VR development, and all have familiarity with Unity and the Microsoft Mixed Reality Toolkit (MRTK). Over 7 days, the participants spent 10, 3, and 6 hours using the toolkit, totalling 19 hours. Each received a $50 gift card for their time.

Our study included 3 stages: (1) tutorial and pre-experiment questionnaire, (2) independent development and (3) post-experiment questionnaire.

On day 1, participants completed a demographic and development experience survey, then participated in an one-hour live-streaming tutorial with a questions and answer period. During the tutorial, the concept of AAR was introduced and the hardware was demonstrated to make participants familiar with the structure of the actuated projector and Hololens. Followed by an introduction to our toolkit, including software environment, installation, and functions.

On days 2 to 7, the participants developed an AAR applications at their own pace. They used Slack to ask us questions and comment on progress.

After the seventh day, each participant submitted their Unity3D project and completed a questionnaire about how they worked with the toolkit, what applications they created, and their thoughts when designing for AAR.

Results

Although our study was small in terms of participants, we believe results from even a few expert developers over this longer usage period are reasonable as a first validation of the AAR concept and toolkit.

AAR Applications

Each participant was tasked to create an AAR application. P1 created a virtual gallery where the HMD user can see descriptive text of gallery objects projected on the wall with the projector. They describe it as: "a "virtual gallery" wherein the AR headset wearer acts as a docent and can view a script which can prompt them with information on what art piece is being viewed, either for lecture or other use." P2 explored the use of AAR for human perception experiments. They continue to state the "The external observer's task would be to judge which plate has a larger portion of food." P3 created a card game in which the HMD and external user can both view their cards through their respective displays. The projector's direction would be changed according to external user's viewpoint. They described it as: " This scene ties into AAR since both an HMD user and an external user are being part of an AR experience."

Feedback and Discussion

Participants generally enjoyed working with the toolkit. P1 said that "It was straightforward to work with the library and functionality ... I understood the concepts being presented, and I can see a lot of potential use for this type of setup." P3 commented "It's an interesting and unique concept with its own set of UX considerations to think about. I think it could be really useful in providing AR experiences to large environments with small groups of people."

Most of the issues and suggested improvements are related to our toolkit or the MRTK directly. P1 said "The tooling was not straightforward to set up at first, better documentation will help." P3 echoed this, "I found the documentation a little confusing at first." P1 suggested that "Simplification of which camera mode and which layers can be viewed will help with this process." P3 commented on the need for MRTK profiles, suggesting "a lot of the project settings/configuration [can] be automated, just like the MRTK." For the spatial awareness feature, they suggested additional features like a plane destruction function. They continued to suggest future support for other HMDs and multi-HMD and multi-projector setups.

Overall, participant comments and applications suggest they all understood the concept of AAR. Without knowing our proposed usage scenarios and applications, participants developed similar ideas. The virtual gallery (P1) was similar to our explicit content sharing applications and the card game (P3) was the extension of window to virtual world usage. This showed how practical and essential AAR can be. All had a positive experience and agreed that our toolkit is useful: *"excited by the possibilities"* [P1], *"a lot of potential"* [P2], and *"unique concept"* [P3].