

# One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities

Joan Sol Roo, Martin Hachet

## ► To cite this version:

Joan Sol Roo, Martin Hachet. One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities. UIST 2017 - 30th ACM Symposium on User Interface Software and Technology, ACM, Oct 2017, Quebec City, Canada. 10.1145/3126594.3126638 . hal-01572490

**HAL Id: hal-01572490**

**<https://hal.archives-ouvertes.fr/hal-01572490>**

Submitted on 20 Dec 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities

Joan Sol Roo  
Inria  
Bordeaux, France  
joan-sol.roo@inria.fr

Martin Hachet  
Inria  
Bordeaux, France  
martin.hachet@inria.fr

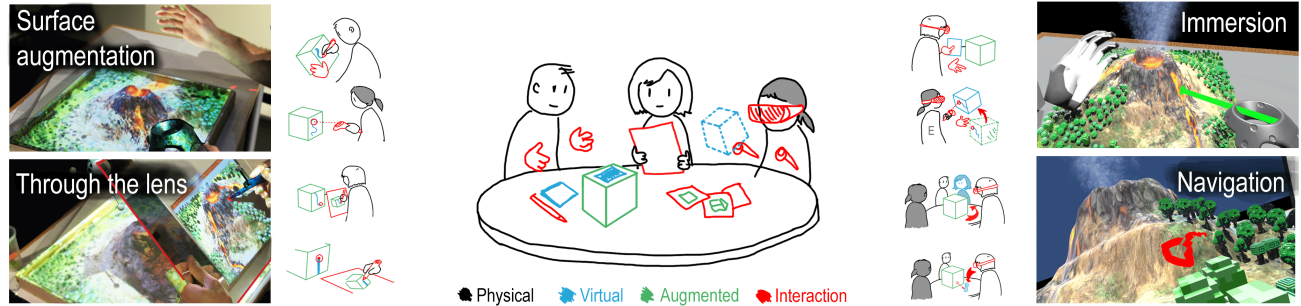


Figure 1. The presented system, exemplified here by an augmented volcano mock-up, allows one or more users to use and transition between multiple mixed reality modalities while interacting with augmented artifacts. The increase in instrumentation provides increasing flexibility, while keeping the interaction framed in the physical world.

## ABSTRACT

Most of our daily activities take place in the physical world, which inherently imposes physical constraints. In contrast, the digital world is very flexible, but usually isolated from its physical counterpart. To combine these two realms, many Mixed Reality (MR) techniques have been explored, at different levels in the continuum. In this work we present an integrated Mixed Reality ecosystem that allows users to incrementally transition from pure physical to pure virtual experiences in a unique reality. This system stands on a conceptual framework composed of 6 levels. This paper presents these levels as well as the related interaction techniques.

## ACM Classification Keywords

H.5.1 Multimedia Information Systems: Artificial, augmented, and virtual realities; H.5.2 User Interfaces: Input devices and strategies

## Author Keywords

Augmented Reality; Mixed Reality; Spatial Augmented Reality; Virtual Reality; Through-the-lens Technique; Tangible User Interfaces; Organic User Interfaces

"One of the things our grandchildren will find quaintest about us is that we distinguish the digital from the real, the virtual from the real." - William Gibson, 2007

## INTRODUCTION

The physical world in which we live provides rich interactions that can hardly be replaced by digital artifacts. In particular, what we see, touch, or smell, and our social collective experiences with face-to-face Humans constitute a unique experience that cannot (should not?) be totally replaced by simulated environments. On the other hand, the flexibility of digital tools has shown undeniable benefits that allow overtaking the constraints of the physical world. As an example, Virtual Reality (VR) allow engineers to immerse themselves inside a simulation and to navigate virtual mock-ups in a way that would not be possible to do in the real life.

In order to complement the physical and the digital worlds, researchers have studied the creation of mixed spaces where digital information can be used in combination with its physical counterpart. This is notably the case of Augmented Reality (AR), where digital and physical information are co-located in a seamless space. The resulting hybrid spaces benefit from digital information anchored onto the physical world. AR is definitively a powerful medium for a number of fields. On the other hand, AR experiences still remain constrained by the physical environment in which the illusion takes place. As an example, neither See-Through AR, nor Spatial AR allows one to see a scene from a point of view that is different from the current viewpoint of the observer.

In our approach, we propose a conceptual framework and its implementation where the physical world that stands in front of the users can be progressively augmented and distorted, with the final goal of augmenting the users' perception. By providing augmented experiences, anchored in the physical reality, we seek for a hybrid space where users benefit from both the force of the physical sensing, and the flexibility of digital interaction.

Our contributions in this work are i) the design of a conceptual framework, where the user can smoothly travel between the physical and the virtual worlds, ii) the implementation of this framework on a unified system, and iii) the exploration of the interaction possibilities. The result builds towards a unified way to look at the intersection between physical and digital realms, combined on a single reality for the user.

### Examples

To better understand this concept, let's consider two concrete scenarios: collaborative learning and hardware maintenance.

Imagine a group of students working around an augmented volcano mock-up (Figure 1). They can discuss and observe the simulated behaviour, while physically touching and moving around and maybe dissecting the mock-up. Some information can be difficult to understand from an egocentric perspective, so they can travel inside the mock-up for more information, using VR. In this case, they can "jump" into the physical model, follow the tubes, and experience the volcano from a different point of view. While immersed, they can still discuss with her classmates, their bodies visible as giants around the mock-up. This simple example, inspired from TV shows like "Cosmos" [40], "the Magic School bus" [11] or "Once upon the time... Life" [2], can be generalized to any pedagogical content, where the changing of viewpoint of the physical environment may improve the learning and understanding of the studied phenomena. Using the shared physical space as starting point fosters discussion.

Now let's consider a maintenance task for a car engine: the engineer is in a process where she has to work physically with the object standing in front of her. Thanks to situated projection, she can benefit from digital support that will guide her during the process (e.g. highlight a given piece) as done in [54]. Now, the engineer needs to observe the engine from the back, or have a closer view of a specific part. Because she cannot manipulate or move around the physical engine, she decides to virtually change her point of view. To do so, she puts on an HMD, and navigate in and around the virtual object. She can also observe a virtual engineer performing the required task, and take the viewpoint of this expert. Numerical simulation can be launched, too, in order to observe for instance the flow in and around the engine. These situated interactions, which would not be possible with standard SAR approaches, may help the engineer to better understand the physical engine she is working on.

Before exploring in detail the different levels present in our framework and their implementation, the next section briefly reviews the previous work that enabled their conception.

### RELATED WORK

The ideas presented in this work build upon 1) the vision that the digital realm can be integrated into the physical one, 2) and the different technologies that were used in the past to explore this possibility.

Since the dawn of digital technologies, researchers and sci-fi writers envisioned a future where digital and physical will be indistinguishable from each other [43]. Science fiction has been used in the past both as inspiration [26, 8] and research tool [7] to study not just how that future can be, but also about the dangers of not designing consciously. Mark Weiser emphasized that ubiquitous technology [48] must be also calm [49]: the digital tools should be available when needed, in a non-disruptive way; for this reason, we propose that the increase in immersion should be progressive and at the user's discretion. Bret Victor with his seminal talk "a humane representation of thought" [46] reflects on how the representation of information frames our way of thinking, and he argues that we must actively channel the advance of technology so it is empowering, instead of limiting. A common theme arises: when designing technology, we should create interfaces that support and collaborate with the users, to help humanity to flourish. In this direction, Jacob et al. formalized the Reality-Based interfaces [22]: computational interfaces that use our already available skills to understand and interact with the external world and each other.

As a first step towards this vision, Milgram presented physical and digital realms not as a dichotomy, but as a Mixed Reality continuum [30]. Ullmer and Ishii presented Tangible User Interfaces (TUIs) [45], which provide physical handles to digital information. This concept was later extended with the vision of Radical Atoms [21]: computational matter that can change its physical properties as needed, an indivisible emulsion of digital and physical. This leads to the question of "what is real?": Eissele et al. [14] reminded us the philosophical distinction between "reality" (*realitas*, what is constructed in the mind) from "actuality" (*actualitas*, the invariant truth), a specially relevant separation when studying mixed reality, which is performed by altering the information received by the senses. By using this definition, when digital and physical work together, they are just counterparts of the same reality for the user.

Rekimoto et al. [36] studied the different ways these realms interact with each other, a classification that could be extended to include spatial augmentation [6]. The latter has been particularly explored in the construction of interactive spaces. Wellner et al. created DigitalDesk [50], a projector-augmented desktop that supports paper based interaction. The Office of the Future [33] envisioned the extension of CAVEs [12] to everyday non-flat surfaces, and soon after was implemented using Spatial Augmented Reality (SAR) [34]. Initially, the focus of the interaction was on the digital information (and still required to wear stereo glasses), but then the focus moved towards the physical space [35]. Since then, it was used to support physical-virtual tools [29], improvised interfaces using objects [47], touch [52], and Organic User Interfaces [17, 32], among others.

In order to take advantage of the complementary characteristics of mixed reality technologies, hybrid systems have been studied. For instance, see-through displays and SAR have been combined, notably in order to complement the HMD’s high resolution with the projectors’ scalable FOV (field of view) [20, 4]. Transitioning between see-through AR and VR have been also explored by Kiyokawa et al. [25]. In the context of multi-display environments, the combination of screens and projection has been studied, both with [10] and without see-through technologies [37, 15]. Dedual et al. [13] proposed a system that combines interactive tabletops with video see-through AR. Smarter Objects [16] and exTouch [24] use video ST-AR to control embedded systems; even when the physical artifact was the focus of attention, no spatial augmentation was presented, except the electronic behaviour itself. Closer to our work is Metadesk [44] by Ullmer and Ishii, which combines tabletop interaction, tangible tokens and a see-through AR window. Hybrid systems not just complement technological limitations, but also combine well with physical artifacts.

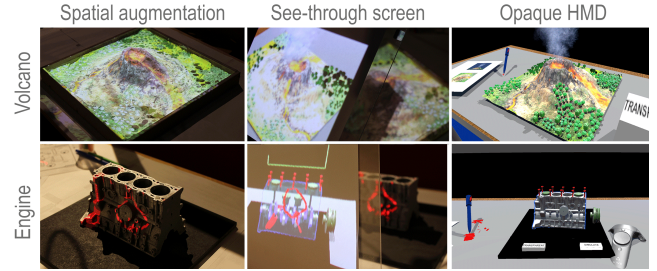
Mixed Reality was also studied as a mean to provide different perspectives of a scene. Magicbook [5] augmented a physical book using an HMD, while also supporting transitions between AR and VR. Benko et al. also studied the use of mixed reality to explore digital locations collaboratively [3]. Komiyama et al. [26] used video see-through to enable the transition between viewpoints (objects, users and spaces) inside the physical world. In all these works, user instrumentation is a requirement to access the augmentation; this can be overcome with always-available viewpoints using screen or projectors. Ibayashi et al. created Dollhouse VR [19], where users can interact with a virtual space, either with a top view (displayed on a tabletop), or from inside (using an HMD-VR), keeping the virtual space as the center of attention. In a preliminary work, we started exploring the use of the virtual interaction with the physical space, supporting not only see-through and immersed modalities, but also spatial augmentation [39, 38].

To summarize, the research involving the combination of physical and digital is vast, both focusing on a specific technology or combining complementary alternatives. In this line of research, we propose a smooth transition through progressive immersion, as described in the next section.

## ONE REALITY

This work focuses on reducing the gap between physical and digital, without limiting the interaction to neither of them. We propose to provide the users the liberty to increment the degree of immersion when needed, without losing contact with the physical world.

In order to explore this incremental immersion, we created a hybrid working space that supports multiple mixed reality modalities simultaneously, and enables the user to free transition between them (Figure 2). In it, the environment and the users are scanned and tracked in real time, enabling their augmentation, movement logging and digital reproduction. To interact with the system, physical tools are rendered available even in virtual spaces, and vice versa.



**Figure 2.** Example scenes: volcano mock-up made out of sand (top), 3D printed Toyota engine (bottom). Each scene can be interacted with different display technologies: spatial augmentation (left), see-through displays (middle), and opaque head mounted displays (right).

When designing the progressive transition between the physical and digital realms, we considered 6 incremental levels where digital characteristics are progressively included (Figure 3). The following subsections describe, in an incremental way, the features of each of the levels in combination with their implementation and interaction consequences.

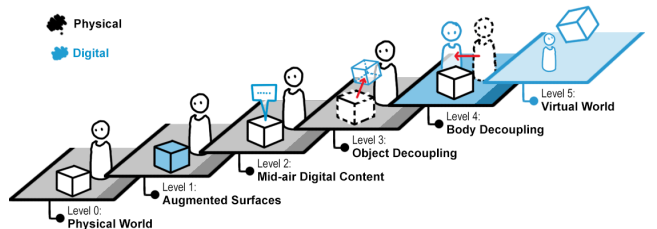
### Level Zero: Physical World

The starting point of interaction is the physical world, where participants can interact with each other and with physical artifacts. This level is comprised of both natural elements and technological devices that alter physical properties in a non-cosmetic and persistent way (i.e., the modifications stay even when the artifact is not working, for example mechanical artifacts or pens). In the physical world, objects can be manipulated when they are not too heavy and not too big, and people can move their heads and bodies to change their point of view, as long as nothing is blocking them.

### Level One: Augmented Surfaces

At the first level, digital information can be placed at the surface of physical objects in the user’s environment, as in the classical work of Rekimoto et al. [37]. This can be used to display complementary information (e.g., show text, change the object perceived texture or show annotations on its surface), or arbitrary information (i.e., using the surface as a screen).

We implemented the augmentation using Spatial Augmented Reality, which allows us to place the digital information directly onto physical surfaces. When the surface of a given object is not suited to support the digital information, the content is then leaked to adjacent surfaces, such as tables or



**Figure 3.** Conceptual framework with 6 levels of incremental augmentation, from Physical to Virtual worlds.



walls. With SAR, digital information is equally available for all the users and viewpoint-independent, as it is the case with physical information.

By tracking objects of interest in the scene, the system supports various interaction modalities (Figure 4). It is possible to directly manipulate the augmented objects, and to use both direct and indirect interaction using pens or pointers. Given the physically nature of the display supports, non-augmented tools can be used naturally.

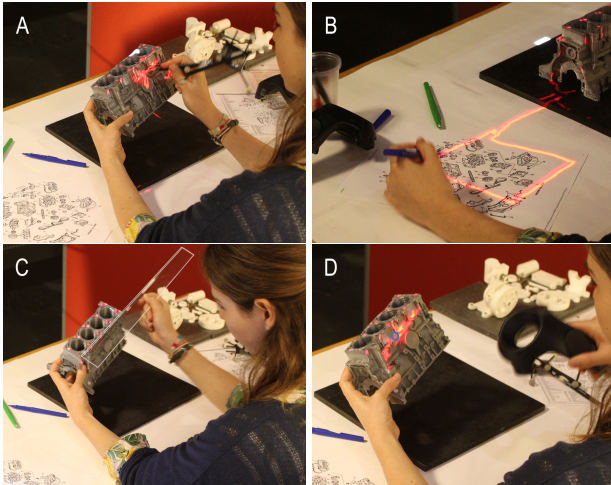


Figure 4. Using spatial augmentation in combination with object tracking enables us: to use augmented pens to draw directly onto artifacts (A) and their surroundings (B), to manipulate the augmented artifacts to use in combination with traditional tools such as rulers (C), or to use of indirect interaction (D).

A widely distributed support of information is paper. When paper is tracked, any digital information displayed on it will follow accordingly. The content of augmented paper can be created using augmented pens; it is also possible to use augmented paper as support for digitally created content – a paper window [18] –, bringing computer screens closer to the physical environment [27]. An additional benefit of using projection is that normal paper can be placed over the augmented counterpart, much like tracing paper. Among the possible digital content, paper windows can display a 3D render of the augmented scene, which can be interacted via ray-casting like traditional displays. As a result, users can interact with their surroundings from a given point of view (Figure 5). We call these windows “interactive pictures”.

### Level Two: Mid-Air Digital Content

While surface augmentation can give dynamism and interactivity to passive objects, it is unable to directly display content in mid-air. In order to provide support for this it is possible to use see-through devices to create the illusion of floating elements.

We prototyped hand-held see-through displays using the interactive pictures presented in the previous level, by attaching the camera position to the paper position (Figure 6). The user can then interact with the digital elements in their field of view using a tracked pen and ray-casting.

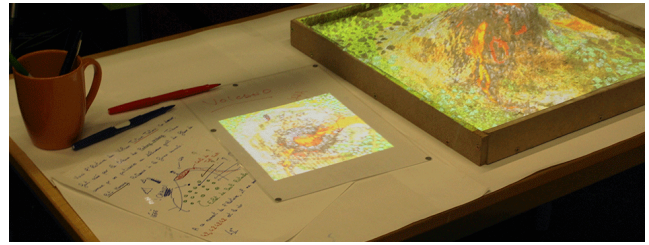


Figure 5. Interactive pictures can display and support interaction with the augmented scene from a different point of view. They can be used in combination with traditional pen and paper.

Compared to head-worn displays (e.g. HoloLens), such an approach has several advantages. First, users do not need to wear specific equipment. Then, the proximity to the physical world is strengthened; it is possible to switch very quickly and easily between the two visual modalities. Both the physical-augmented object (level 1) and the through-the-lens object (level 2) can also be observed at the same time. Finally, multiple observer can use the same through-the-lens image. This reinforces collaborative and social interaction.

The main drawbacks are that, as with other hand-held technologies, they require the device to be held, and they have a limited field of view. For the cases where comfort is relevant or both hands are needed (e.g., long or precise tasks) it is possible to detach the viewpoint and use indirect interaction, as seen in the previous level.

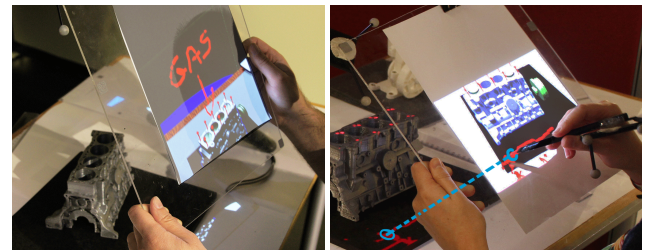


Figure 6. See-through displays, implemented using interactive pictures, allow to create the illusion of mid-air information (red text and arrow) floating over the engine (left), while also supporting indirect interaction with the scene via tracked pens and ray-casting (right).

### Back to the Examples

To better understand the first three levels and how the system supports them, let’s explore the examples in more detail.

The students (Figure 1) around the volcano mock-up (level 0) can observe superficial simulation (level 1) of its activity and in-place information, while also being able to move around and touch the mock-up, sharing a unified experience. A see-through screen can help the students to visualize mid air information (level 2) such as steam and lava coming from the top of the volcano, or its interaction with virtual trees. They can use the screen to directly take notes of what they see through it, or place different views of the volcano side-by-side to discuss and make their own drawings.

In the case of the engineer performing maintenance tasks on the engine (level 0), the augmentation can be bi-directional.

She can both use previously stored information to guide the process, or create new annotations, creating a tutorial on how to perform the task (level 1). She can also use the see-through approach (level 2) to visualize mid air annotations, measurements and spatial instructions, such as "screw here, then drill there". Once again, this can be bidirectional, either following or creating instructions through graphical annotations or by example.

### Level Three: Object Decoupling

So far, level 1 and 2 aimed at keeping the augmentation as close as possible to the physical world. Starting at level three, we propose to soften the physical constraints in order to give more flexibility to the users. This can be useful when trying to understand complex processes (simulation), when needing to perform tasks that are not normally possible, such as lifting a heavy object (overcoming physical constraints), or trying to observe actions performed in the past (replay).

To this end, beyond the see-through displays we described in the previous section, we introduced a fully immersive modality that reproduces the physical world standing in front of the users, and that is perceived through a VR Head-Mounted display. Both worlds (physical and virtual) are mapped one-to-one. This means that the users can physically touch the objects they are observing virtually (Figure 7). This gives the immersed user a strong anchorage with the physical world. To increase the immersion, we use hand tracking based on Leap Motion, which provide not only feedback but also the possibility to directly interact with purely digital content.

At this level, the users can now free themselves from strict mapping and perform actions that would not be possible in reality (e.g. manipulate a heavy or fragile object). To do so, they can use virtual controllers. In addition to the features we already described, the users have this new ability consisting in distorting the reality.

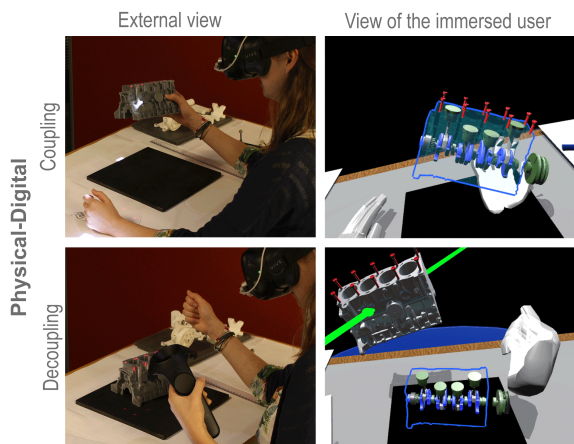


Figure 7. While immersed, the physical and digital counterparts of augmented objects can be mapped one-to-one (top). This coupling can be relaxed for additional flexibility (bottom).

#### Awareness considerations

When physical and digital are decoupled, it is necessary to take additional considerations to help the user to prevent acci-

dentally hitting physical objects. For this reason, when users gets near a physical object (either with their hands or tools) we display its wire-frame or bounding box.

It is also important to know where the users are (for themselves, and for others). For this, we display user avatars. The representation depends on the modality: augmented surfaces display the position and orientation in an iconic way (arrow), while see-through and HMD-VR show a 3D reconstruction of the users, enabling them to see themselves and others. Currently the 3D reconstruction is only based on the Microsoft Kinect point-cloud, yet it could be possible to extend it to rigged meshes [41], animated using Kinect skeleton tracking.

Finally, what the helmet shows can be displayed onto an interactive picture. This way, non-immersed users can see what the immersed user does. Given the behaviour of interactive pictures, users are also able to *interact* with the scene in front of the immersed user, indicating for example a point of interest, or moving elements through the window. Similar results can be obtained for non immersed users using the estimation of head's position and orientation using skeleton tracking.

#### Back to the Examples

At this level, the students can virtually modify the geometry of the objects. For example, they can grab the volcano standing in front of them, and stretch it to observe what would happen if the volcano was higher. They can also extend the frontiers of the current physical mock-up by using 3D terrains around the volcano, or by visualizing internal structure above it. They can also copy and paste the volcano to experiment interactions between multi-objects.

In the case of the engineer, she can observe the engine in isolation by wearing an HMD, removing occluding elements or only showing their wire-frame. She can also virtually move a piece mid-air, in order to study it and annotate it. The annotations are reflected onto the physical engine in real time, since both views are synchronized. Finally, she can simulate the proper working behaviour of the engine, to clarify the function of the individual pieces.

This level gives users the illusion of changes in the physical world, while they keep their body as frame of reference: for a given user, the physical landmarks (objects and other users) stay still relatively to him or herself. The next level explores the separation between physical and digital body.

### Level Four: Body Decoupling

The fourth level explores the immersive virtual navigation of the physical space. The users can then change their perceived position, orientation and *scale* in space (Figure 8).

For this, HMD-VR helmets are ideal, since the users are presented with a 3D scene overriding their senses, with an arbitrary point of view. In order to control the navigation, we use a teleportation-based interaction using a wand controller, where the resulting scale depends on the target surface (and in our case is configured manually). A simple fade transition was implemented; since both a change of position and scale are involved, travelling could lead to cyber-sickness. Besides being able to teleport, users can also rotate either around a point,

or on their current location. Since it is possible to ray-cast through the interactive pictures, users can pick the target position *through* a window, both before and during the immersive session.

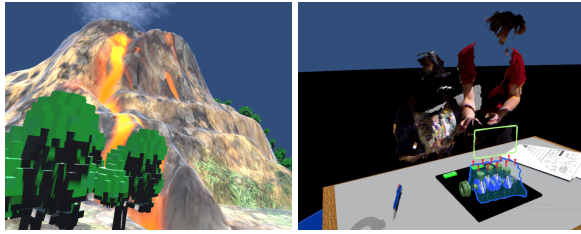


Figure 8. Users can navigate the augmented scene virtually. Left: the user jumped inside the scene and experience the volcano from an ego-centric point of view. Right: the engineer rotated to her right, to see herself and her colleague.

The navigation is then not restricted to the current table and their surroundings, but could be used for tele-presence. Indeed, it is possible to transition to distant places (e.g., coworkers office, Mars surface), one step at the time (similar to Google Street View [1]), or jumping there through an interactive picture. Once the digital environment is framed in relationship with the physical one, it is possible to then transition to purely virtual spaces (the fifth level).

*Back to the Examples*

Interacting with the mock-up is enriching, but in the end it is a representation of something bigger. If the students want to experience or study the surface of the volcano, they can do so by wearing an HMD-VRs and teleporting on its surface (Figure 8). The virtual model can then be navigated by jumping from place to place, experiencing the represented scale. While on the surface, the students can see their schoolmates and themselves as giants around the table, allowing to asymmetrically collaborate, not just speaking but also using the aforementioned tools. They can use both the volcano surface and the interactive pictures to communicate with each other.

If the engineer needs help to repair the engine, she can observe another engineer performing the task, either using a recording or tele-presence. She can see her colleague working from his or her point of view, or virtually move around for a better view.

**Overview**

In this section we described the 6 levels of our conceptual framework, accompanied by the implementation of the 4 central levels that provide augmentation. This framework was conceived to incrementally provide digital tools to interact with the physical world, up to the point where purely digital applications can be framed within the physical space.

The selection of a given modality involves a trade-off (Table 1). As digitality increases, so does the flexibility of interaction and simulation, while the connection with the physical world gets thinner by the instrumentation. We took awareness considerations to keep the user anchored in the physical world as much as possible, yet we argue that immersion should not be increased because it is possible, but only when it is needed.

The result moves towards a space where physical and digital do not compete with each other for the users attention, but instead complement in the construction of a unified experience. This work builds on top of the vast Mixed Reality literature, providing a new lens to look at the diverse projects in the field – such as augmented surfaces [37], magicbook [5] and physical-virtual tools [29]–, while also providing a unified ecosystem where they can coexist.

**SYSTEM IMPLEMENTATION**

The system runs using two Alienware desktop computers, one worked as a server (Intel i7-3820, 8GiB RAM with an NVIDIA GTX 660 Ti) the second one as a client (Intel i7-3820, 24GiB RAM with dual NVIDIA GTX 690s). The use of multiple computers was necessary given the USB BUS bandwidth required by the sensors. The client performs dense acquisition (Kinect) and rendering (HTC Vive, projectors), while the server handles sparse tracking (Optitrack, Leap Motion) and stream it using UDP and OSC. We selected this distribution in order to minimize the intranet usage, yet dense acquisition and rendering could be distributed among several computers.

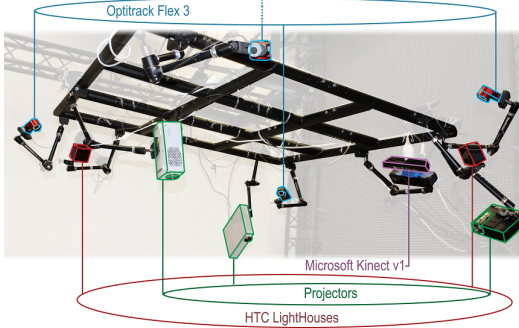
The whole setup was organized around a rectangular table of 130cm by 80 cm, which was used as the reference for calibration. All the sensors and projectors were placed on an overhead platform (Figure 9). Regarding the displays, we use: 1) an HTC Vive as HMD, and 2) three off-the-shelf projectors, providing 360° augmentation. The projectors are: an LG PF80G, an LG PF1500 and an Asus B1M; this selection was

PROPERTY	LEVEL			
	1. Augmented Surfaces	2. Mid-Air Digital Content	3. Object Decoupling	4. Body Decoupling
Location of the physical display	On the surface of the objects (main objects, table, tacked panel)		Head Mounted Display	
Nature of the display	- Appearance of the objects - Additional info (text, pics)	Virtual objects or information located in the 3D space		
Link to the physical world	One-to- one mapping (POV and objects)		Global scene and viewpoint coherency	Independent
Main interaction	- Natural interaction (object manipulation, paper and pen) - Digital tool (virtual ray)	-Natural manipulation of the See Through panel - Window-based technique (picking a virtual object in the panel)	Natural manipulation of the co-located objects or dedicated technique when disconnecting the objects	VR interaction techniques
Main POV	Natural POV		Natural simulated POV	VR navigation techniques

Table 1. Each of the augmentation levels and their properties.



based on the available hardware, and the setup could be greatly improved by using 3 projectors of the same model, preferably with higher brightness and frame-rate. The tracking uses 4 Optitrack Flex cameras covering the volume over and around the table. A single Microsoft Kinect v1 provides a partial point-cloud of the table contents and the users, including their skeleton tracking. Given that we used HTC Vive, the tracking lighthouses are also placed on the overhead platform. For hand tracking we attached a Leap Motion to the helmet.



**Figure 9.** The setup supports surrounding tracking and augmentation. The tracking is performed using 4 Optitrack Flex 3 infrared cameras (sparse information), a Microsoft Kinect v1 (dense information point-cloud), and 2 HTC lighthouses (off the shelf HTC infrared emitters, used to track the HTC components). The augmentation is performed using three overhead projectors.

### Infrared Interference

Besides the USB bandwidth, the main issue of this setup is infrared interference between sensors: all of them – except the Kinect – use the same wavelength. To address this, we synchronized the Optitrack cameras with the lighthouses<sup>1</sup>; Leap Motion emitters can also create interference, but the impact was diminished with Optitrack parametrization.

The use of additional Kinects would greatly improve the point-cloud coverage, yet additional considerations are required to prevent interference [9]. Regarding Microsoft Kinect v2, we were unable to include it in the system since it uses the same wavelength as the other sensors, leading to interference that could not be mitigated.

### Calibration

The alignment of all the subsystems was performed in stages. Both Optitrack and Kinect were calibrated offline, computing the alignment between coordinate systems (3D to 3D calibration, translation and rotation), using the table as frame of reference. The projectors were calibrated using OpenCV camera calibration (2D to 3D calibration, intrinsic and extrinsic estimation), by matching reference pixels with their 3D position, using the already calibrated Optitrack. The HTC Vive was calibrated online, by placing the controller (most reliable transform) on a previously determined place in the table, and computing the transform between coordinate systems (3D to 3D calibration). Finally, the Leap Motion was manually calibrated, by finding the offset with HMD center.

<sup>1</sup>The method used was adapted from [http://wiki.optitrack.com/index.php?title=Sync\\_Configuration\\_with\\_an\\_HTC\\_Vive\\_System](http://wiki.optitrack.com/index.php?title=Sync_Configuration_with_an_HTC_Vive_System)

### Technical Future Work

This work moves towards the construction of spaces that support mixed reality, similarly on how CAVEs support virtual reality. To this end, further work could take inspiration from literature. First, the interaction space could be extended as with *Steerable Augmented Reality* [51], Roomalive [23] and CAVE [12] technologies. In the same line, full nonrigid reconstruction [31] could be used to scan and track physical elements. This could also be used to support touch interaction everywhere [53]. The use of projectors to create paper windows has clear limitations (occlusions, tracking), but we look forward towards the promising technology of Organic User Interfaces [17], which could replace the projectors partially or completely. In the mean time, the SAR augmentation could be improved by preventing overlapping between projectors.

We also consider that the use of embedded components would enrich the system. Light sensors were used in the past to perform precise projector calibration [42] or perform projector based tracking [28]. Another use for embedded systems is actuation and sensing [16], which could be used to render the physical world (level 0 of our conceptual framework) more dynamic and interactive, as with smarter objects [16].

### CONCLUSION AND DIRECTIONS TO FUTURE WORK

In this paper we presented a hybrid mixed reality conceptual framework, providing incremental augmentation and instrumentation. The framework was implemented using a combination of multiple display technologies. The first one is Spatial augmentation, always available augmentation where only the surfaces' appearance is modified. See-through devices can display mid-air information, and partially override physical information. Finally, HMDs provide a virtual replica of the physical scene, taking advantage of the freedom of virtual spaces without losing connection with the environment.

Even when the different technologies coexist simultaneously in our system, the interaction was designed to perform the tasks while keeping the connection with the physical space and other users; when a task is not possible, then the amount of digital support is increased, in combination with awareness considerations to help the users keep the link with their environment. The resulting system focuses on the smooth transition from a purely physical to a purely digital experience.

An underlying question driving this work is "where are these virtual spaces for the mind?". We consider that the cost of multitasking and traditional computers is that our attention constantly switches between two locations perceived as remote, and by easing the transition we can reduce the cognitive impact. If that is the case, then the dissociation created by modern computers and mobile devices could be greatly reduced by rendering them more aware and in synchrony with their surroundings. The next step, is to test these hypothesis with dedicated user studies using the presented ecosystem, and the answers will in term help the improvement of the system.

### ACKNOWLEDGEMENTS

This work was possible thanks to the ISAR project (ANR-14-CE24-0013), and the collaboration from Diota. We would also like to thank our colleagues for their enriching feedback.



## REFERENCES

1. Dragomir Anguelov, Carole Dulong, Daniel Filip, Christian Frueh, Stéphane Lafon, Richard Lyon, Abhijit Ogale, Luc Vincent, and Josh Weaver. 2010. Google street view: Capturing the world at street level. *Computer* 43, 6 (2010), 32–38.
2. Albert Barille. 1987. Once Upon a Time... Life. (1987).
3. Hrvoje Benko, Edward W Ishak, and Steven Feiner. 2004. Collaborative mixed reality visualization of an archaeological excavation. In *Proceedings of the 3rd IEEE/ACM international Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 132–140.
4. Hrvoje Benko, Eyal Ofek, Feng Zheng, and Andrew D Wilson. 2015. Fovear: Combining an optically see-through near-eye display with projector-based spatial augmented reality. In *UIST '15*. ACM, 129–135.
5. Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. 2001. The magicbook-moving seamlessly between reality and virtuality. *IEEE CGandA* 21, 3 (2001), 6–8.
6. Oliver Bimber and Ramesh Raskar. 2006. Modern approaches to augmented reality. In *ACM SIGGRAPH 2006 Courses*. ACM, 1.
7. Mark Blythe. 2014. Research through design fiction: narrative in real and imaginary abstracts. In *CHI '14*. ACM, 703–712.
8. Sophia Agnes Brueckner. 2014. *Out of network: technologies to connect with strangers*. Ph.D. Dissertation. Massachusetts Institute of Technology.
9. D Alex Butler, Shahram Izadi, Otmar Hilliges, David Molyneaux, Steve Hodges, and David Kim. 2012. Shake'n'sense: reducing interference for overlapping structured light depth cameras. In *CHI '12*. ACM, 1933–1936.
10. Andreas Butz, Tobias Hollerer, Steven Feiner, Blair MacIntyre, and Clifford Beshers. 1999. Enveloping users and computers in a collaborative 3D augmented reality. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*. IEEE, 35–44.
11. Joanna Cole and Bruce Degen. 1994. The Magic School Bus. (1994).
12. Carolina Cruz-Neira, Daniel J Sandin, and Thomas A DeFanti. 1993. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*. ACM, 135–142.
13. Nicolas J Dedual, Ohan Oda, and Steven K Feiner. 2011. Creating hybrid user interfaces with a 2D multi-touch tabletop and a 3D see-through head-worn display. In *ISMAR 2011*. IEEE, 231–232.
14. Mike Eissele, Oliver Siemoneit, and Thomas Ertl. 2006. Transition of mixed, virtual, and augmented reality in smart production environments-an interdisciplinary view. In *Robotics, Automation and Mechatronics, 2006 IEEE Conference on*. IEEE, 1–6.
15. Renaud Gervais, Joan Sol Roo, and Martin Hachet. 2016. Tangible Viewports: Getting Out of Flatland in Desktop Environments. In *TEI'16*. ACM, 176–184.
16. Valentin Heun, Shunichi Kasahara, and Pattie Maes. 2013. Smarter objects: using AR technology to program physical objects and their interactions. In *CHI'13 Extended Abstracts*. ACM, 961–966.
17. David Holman and Roel Vertegaal. 2008. Organic User Interfaces: Designing Computers in Any Way, Shape, or Form. *Commun. ACM* 51, 6 (June 2008), 48–55. DOI: <http://dx.doi.org/10.1145/1349026.1349037>
18. David Holman, Roel Vertegaal, Mark Altosaar, Nikolaus Troje, and Derek Johns. 2005. Paper windows: interaction techniques for digital paper. In *CHI '05*. ACM, 591–599.
19. Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, and Takeo Igarashi. 2015. Dollhouse VR: a multi-view, multi-user collaborative design workspace with VR technology. In *SIGGRAPH Asia 2015 Emerging Technologies*. ACM, 8.
20. Adrian Ilie, Kok-Lim Low, Greg Welch, Anselmo Lastra, Henry Fuchs, and Bruce Cairns. 2004. Combining head-mounted and projector-based displays for surgical training. *Presence: Teleoperators and Virtual Environments* 13, 2 (2004), 128–145.
21. Hiroshi Ishii, Dávid Lakatos, Leonardo Bonanni, and Jean-Baptiste Labrune. 2012. Radical atoms: beyond tangible bits, toward transformable materials. *interactions* 19, 1 (2012), 38–51.
22. Robert JK Jacob, Audrey Girouard, Leanne M Hirshfield, Michael S Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. Reality-based interaction: a framework for post-WIMP interfaces. In *CHI '08*. ACM, 201–210.
23. Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 2014. RoomAlive: magical experiences enabled by scalable, adaptive projector-camera units. In *UIST '14*. ACM, 637–644.
24. Shunichi Kasahara, Ryuma Niiyama, Valentin Heun, and Hiroshi Ishii. 2013. exTouch: Spatially-aware Embodied Manipulation of Actuated Objects Mediated by Augmented Reality. In *TEI '13 (TEI '13)*. ACM, New York, NY, USA, 223–228.
25. Kiyoshi Kiyokawa, Haruo Takemura, and Naokazu Yokoya. 2000. SeamlessDesign for 3D object creation. *IEEE multimedia* 7, 1 (2000), 22–33.
26. Ryohei Komiyama, Takashi Miyaki, and Jun Rekimoto. 2017. JackIn Space: Designing a Seamless Transition Between First and Third Person View for Effective Telepresence Collaborations. In *Proceedings of the 8th Augmented Human International Conference (AH '17)*. ACM, New York, NY, USA, Article 14, 9 pages. DOI: <http://dx.doi.org/10.1145/3041164.3041183>

27. J r my Laviolle and Martin Hachet. 2012. PapARt: interactive 3D graphics and multi-touch augmented paper for artistic creation. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on*. IEEE, 3–6.
28. Johnny C Lee, Scott E Hudson, Jay W Summet, and Paul H Dietz. 2005. Moveable interactive projected displays using projector based tracking. In *UIST '05*. ACM, 63–72.
29. Michael R Marner, Bruce H Thomas, and Christian Sandor. 2009. Physical-virtual tools for spatial augmented reality user interfaces. In *ISMAR*, Vol. 9.
30. Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1995. Augmented reality: A class of displays on the reality-virtuality continuum. In *Photonics for industrial applications*. International Society for Optics and Photonics, 282–292.
31. Richard A Newcombe, Dieter Fox, and Steven M Seitz. 2015. Dynamicfusion: Reconstruction and tracking of non-rigid scenes in real-time. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.
32. Ben Piper, Carlo Ratti, and Hiroshi Ishii. 2002. Illuminating clay: a 3-D tangible interface for landscape analysis. In *CHI '02*. ACM, 355–362.
33. Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. 1998b. 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*. ACM, 179–188.
34. Ramesh Raskar, Greg Welch, and Henry Fuchs. 1998a. Spatially augmented reality. In *First IEEE Workshop on Augmented Reality (IWAR'98)*. Citeseer, 11–20.
35. Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay. 2001. Shader lamps: Animating real objects with image-based illumination. In *Rendering Techniques 2001*. Springer, 89–102.
36. Jun Rekimoto and Katashi Nagao. 1995. The world through the computer: Computer augmented interaction with real world environments. In *UIST '95*. ACM, 29–36.
37. Jun Rekimoto and Masanori Saitoh. 1999. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *CHI'99*. ACM, 378–385.
38. Joan Sol Roo, Renaud Gervais, J r my Frey, and Martin Hachet. 2017. Inner Garden: Connecting Inner States to a Mixed Reality Sandbox for Mindfulness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1459–1470.
39. Joan Sol Roo and Martin Hachet. 2017. Towards a hybrid space combining Spatial Augmented Reality and virtual reality. In *3DUI'17*. IEEE, 195–198.
40. Carl Sagan, Steven Soter, and Ann Druyan. 1989. *Cosmos: A personal voyage*. (1989).
41. Ari Shapiro, Andrew Feng, Ruizhe Wang, Hao Li, Mark Bolas, Gerard Medioni, and Evan Suma. 2014. Rapid avatar capture and simulation using commodity depth sensors. *Computer Animation and Virtual Worlds* 25 (2014).
42. Ross T Smith, Guy Webber, Maki Sugimoto, Michael Marner, and Bruce H Thomas. 2013. [Invited Paper] Automatic Sub-pixel Projector Calibration. *ITE Transactions on Media Technology and Applications* 1, 3 (2013), 204–213.
43. Ivan E Sutherland. 1965. The ultimate display. *Multimedia: From Wagner to virtual reality* (1965).
44. Brygg Ullmer and Hiroshi Ishii. 1997. The metaDESK: models and prototypes for tangible user interfaces. In *UIST '97*. ACM, 223–232.
45. Brygg Ullmer and Hiroshi Ishii. 2000. Emerging frameworks for tangible user interfaces. *IBM systems journal* 39, 3.4 (2000), 915–931.
46. Bret Victor. 2014. Humane Representation of Thought: A Trail Map for the 21st Century. In *UIST '14*. ACM, New York, NY, USA, 1.
47. James A Walsh, Stewart von Itzstein, and Bruce H Thomas. 2014. Ephemeral interaction using everyday objects. In *Proceedings of the Fifteenth Australasian User Interface Conference-Volume 150*. Australian Computer Society, Inc., 29–37.
48. Mark Weiser. 1993. Some computer science issues in ubiquitous computing. *Commun. ACM* 36, 7 (1993).
49. Mark Weiser and John Seely Brown. 1997. The coming age of calm technology. In *Beyond calculation*. Springer.
50. Pierre Wellner. 1993. Interacting with Paper on the DigitalDesk. *Commun. ACM* 36, 7 (July 1993), 87–96.
51. Andrew Wilson, Hrvoje Benko, Shahram Izadi, and Otmar Hilliges. 2012. Steerable augmented reality with the beamatron. In *UIST '12*. ACM, 413–422.
52. Robert Xiao, Chris Harrison, and Scott E. Hudson. 2013a. WorldKit: Rapid and Easy Creation of Ad-hoc Interactive Applications on Everyday Surfaces. In *CHI '13 (CHI '13)*. ACM, New York, NY, USA, 879–888.
53. Robert Xiao, Chris Harrison, and Scott E Hudson. 2013b. WorldKit: rapid and easy creation of ad-hoc interactive applications on everyday surfaces. In *CHI '13*. ACM, 879–888.
54. Jianlong Zhou, Ivan Lee, Bruce Thomas, Roland Menassa, Anthony Farrant, and Andrew Sansome. 2011. Applying Spatial Augmented Reality to Facilitate In-situ Support for Automotive Spot Welding Inspection. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry (VRCAI '11)*. ACM, New York, NY, USA, 195–200.