



HAL
open science

De-augmenting Visually-Augmented Reality

Eugenie Brasier, Emmanuel Pietriga, Anastasia Bezerianos, Olivier Chapuis,
Caroline Appert

► **To cite this version:**

Eugenie Brasier, Emmanuel Pietriga, Anastasia Bezerianos, Olivier Chapuis, Caroline Appert. De-augmenting Visually-Augmented Reality. RR-9494, Université Paris-Saclay; CNRS; Inria. 2022, pp.29. hal-03914418

HAL Id: hal-03914418

<https://inria.hal.science/hal-03914418>

Submitted on 3 Jan 2023

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Inria

De-augmenting **Visually-Augmented Reality**

Eugénie Brasier, Emmanuel Pietriga, Anastasia Bezerianos, Olivier Chapuis, Caroline Appert

**RESEARCH
REPORT**

N° 9494

December 2022

Project-Teams ILDA

ISRN INRIA/RR--9494--FR+ENG

ISSN 0249-6399



De-augmenting Visually-Augmented Reality

Eugénie Brasier*, Emmanuel Pietriga*, Anastasia Bezerianos*,
Olivier Chapuis*, Caroline Appert*

Project-Teams ILDA

Research Report n° 9494 — December 2022 — 30 pages

Abstract: We anticipate a future in which people frequently have virtual content displayed in their field of view to augment reality. Situations where this virtual content interferes with users' perception of the physical world will thus be more frequent, with consequences ranging from mere annoyance to serious injuries. We argue for the need to give users agency over virtual augmentations, discussing the concept of *de-augmenting* augmented reality by selectively removing virtual content from the field of view. *De-augmenting* lets users target what actually interferes with their perception of the environment while keeping what is of interest. We contribute a framework that captures the different facets of de-augmentation. We discuss what it entails in terms of technical realization and interaction design, and end with three scenarios to illustrate what the user experience could be in a sample of domestic and professional situations.

Key-words: augmented reality; visual interference; user agency

* Université Paris-Saclay, CNRS, Inria, LISN

**RESEARCH CENTRE
SACLAY – ÎLE-DE-FRANCE**

1 rue Honoré d'Estienne d'Orves
Bâtiment Alan Turing
91120 Palaiseau

Dé-augmentation de la réalité augmentée visuelle

Résumé : Nous anticipons un avenir dans lequel les gens auront fréquemment du contenu virtuel affiché dans leur champ de vision pour augmenter la réalité. Les situations où ce contenu virtuel interfère avec la perception du monde physique par les utilisateurs seront donc plus fréquentes, avec des conséquences allant du simple désagrément à des accidents graves. Nous soutenons la nécessité de donner aux utilisateurs la possibilité d'agir sur les augmentations virtuelles, en discutant du concept de dé-augmentation de la réalité augmentée par la suppression sélective du contenu virtuel du champ de vision. La dé-augmentation permet aux utilisateurs de cibler ce qui interfère réellement avec leur perception de l'environnement tout en conservant ce qui est intéressant. Nous proposons un cadre qui capture les différentes facettes de la dé-augmentation. Nous discutons de ce que cela implique en termes de réalisation technique et de conception de l'interaction, et nous terminons par trois scénarios pour illustrer ce que pourrait être l'expérience de l'utilisateur dans des situations domestiques et professionnelles.

Mots-clés : réalité augmentée ; interférence visuelle ; expérience utilisateur

Contents

1	Introduction	3
2	Related Work	4
2.1	Augmented Virtuality	5
2.2	Altered Reality	5
2.3	Augmented Reality	6
3	De-augmenting AR: a Framework	7
3.1	De-augmentation Triggers	7
3.1.1	User Command	7
3.1.2	User Rules	9
3.1.3	External Source	9
3.2	De-augmentation Scope	10
3.2.1	Spatial Scope [Entity]	10
3.2.2	Spatial Scope [Region]	11
3.2.3	Temporal Scope [Temporary]	11
3.2.4	Temporal Scope [Permanent]	11
3.3	De-augmentation Rendering	11
3.3.1	Pixel Compositing	12
3.3.2	Geometric Transformation	12
4	Implications for Technical Implementation	12
4.1	AR technologies and De-augmentation	13
4.2	Representative Sample of De-augmentation Operations	14
4.3	Lessons Learned	15
5	Implications for Interaction Design	15
6	Illustrative Scenarios	18
6.1	At Home	18
6.2	Video Call	18
6.3	Safety First	20
7	Conclusion	21
	Acknowledgment	22
	References	22

1 Introduction

Head-mounted displays remain for the time being overly conspicuous and uncomfortable for people to consider wearing unless necessary. But miniaturization and progress in terms of autonomy should eventually make it possible to embed AR display technology in consumer-oriented hardware such as prescription glasses, enabling the sort of “*continuous AR experiences*” that Grubert *et al.* envision with Pervasive Augmented Reality [30]. AR displays are actually getting integrated in domain-specific pieces of wearable equipment already, in which they are considered socially acceptable [46, 45]: motorcycle helmets, safety glasses, hard hats. Looking in the more distant future, near-eye displays [14] might involve radically different technologies such as meta-surfaces [53, 50] and other types of contact lenses [73, 76], though these are still at an early research stage [6, 37].

Anticipating such improvements, both academia and industry are contemplating an increasingly diverse range of use cases for AR. Beyond the well-known ones [7] such as gaming [77], education [81] or industry [26, 32], use cases now include domestic use [44], health and therapy [34, 38], emergency management [89], assisting people with visual impairments [52, 69, 91, 103], a large range of sports & artistic activities (see, *e.g.*, [47]), spatial orientation assistance at large [13], as well as advertising [75].

Developments on the technological side and on the usage side go hand in hand, and suggest a future in which AR has indeed become pervasive [30]. Becoming inconspicuous, the technology will disseminate beyond the workplace. Becoming practical in many contexts, it will allow for the actual realization of many of the envisioned use cases, as well as new, unanticipated ones. This can be seen as a virtuous or a vicious circle depending on the perspective we take on it; but one that leads part of society to adopt always-on AR, regardless. Matsuda [64] as well as May-raz & Lazo [65] both propose striking depictions of this in their futuristic short films. Whether we embrace such a future or not, always-on AR raises multiple questions. Many relate to ethics [6, 61, 84, 85, 90], some others relate to practical concerns [30], and yet others to both. In all cases, the more often AR content is shown, the more often it will somehow be “in the way”.

We argue that users need agency over their AR display and that to achieve proper user agency, user control over virtual content display cannot be limited to a simple on/off toggle. Some virtual elements might be interfering, while others might be critical to the task at hand or even just convenient to have, thus calling for solutions to *selectively* remove some of them from the field of view. To this end, we introduce the concept of *de-augmenting* augmented reality, which consists of removing or transforming part of the virtual content in the users’ field of view so that it no longer interferes with their perception of important elements in the physical environment. This concept is complementary to *adaptive* AR that automatically optimizes presentation based on users’ context [30]. User-initiated *de-augmentation* is rather for cases that AR designers simply could not anticipate [55], for cases where users would train an adaptive AR system, or for cases where users just want to keep control over that AR system.

After giving an overview of the research literature about techniques and displays for combining the physical and virtual worlds, we build on the literature to provide a framework that defines what a *de-augmentation* operation is and captures its different design facets. We build a proof-of-concept prototype that implements a sample of de-augmentation operations with off-the-shelf technology to probe the concept. Building on that experience, we discuss aspects related to the technical realization of the concept and its implications in terms of interaction design. We finally detail three scenarios where we used our framework to rethink what the user experience could be in a sample of domestic and professional situations. Overall, by characterizing, discussing and illustrating the concept of *de-augmentation*, our contribution aims to provide a framework that raises interaction designers’ awareness about the potential visual hindrance of AR and suggests means to actually design and implement solutions that let users selectively clear their field of view according to their needs.

2 Related Work

De-augmentation is about letting users dose the amount of virtual content on top of the physical environment. Techniques that mix virtual and physical elements are placed on the reality-virtuality continuum [67] depending on what is predominant: either the virtual environment, the physical environment, or none of them in particular. The reality-virtuality continuum may apply to multiple modalities (visual, audio, haptics), and extensions beyond the original, unidimensional, definition have also been discussed [62]. In this preliminary investigation of the *de-augmentation* concept, we focus on the visual modality.

2.1 Augmented Virtuality

At one end of the continuum, Virtual Reality (VR) completely occludes the real world, replacing it with a computer-generated 3D scene, fully isolating users from their physical environment in the process. While this helps make the experience immersive, researchers have also explored the concept of *Augmented Virtuality* where some elements from reality get integrated back into the virtual world.

Several empirical studies have investigated the user experience in augmented virtuality settings [9, 66, 94]. Overall, results suggest that, in the context of VR, seeing some reality is necessary to enable interactions with physical objects (*e.g.*, type on a keyboard or grab food) and enable safe locomotion in the physical environment. But they also suggest that the amount of elements brought from the real world should be limited so as not to impact immersion in the virtual world too much. In their study, Kanamori *et al.* [41] test whether only a light, simplified representation of nearby obstacles¹ would be enough to let participants walk and reach a specific object. However, no simplified representation was as efficient as the actual representation of obstacles. Taking a more technical perspective, the RealityCheck system [31] proposes an application-independent solution for the real-time reconstruction of people and objects in the physical space, and their graphical compositing with the virtual environment in VR.

Augmented Virtuality can be seen as a transposition of the concept of de-augmentation to a different point on the reality-virtuality continuum. One key difference between the two approaches stems from the difference in what is predominant. Augmented Virtuality integrates elements from the physical environment based on the assumption that *the system is in charge of detecting what elements from the real world should be integrated into the virtual scene*. In AR, even assuming a crowded field of view, users will see the real world through the virtual content – albeit degraded. This ability to see through the virtual layer is key to giving users agency over how to mitigate visual interference in the framework of de-augmentation.

2.2 Altered Reality

Coming back to the continuum and moving away from predominantly-virtual environments towards predominantly-physical environments, we find techniques that balance the two. With proper spatial registration [18], AR content can be seemingly integrated into the physical environment, going as far as changing the appearance of objects or even people. Such techniques can be said to *alter* reality.

But altering reality is not only about changing the visibility of physical objects. Annexing Reality [33] makes use of objects in the physical environment to create tangible proxies for manipulating virtual objects. The system scans the environment in order to identify the best match between a physical object and the virtual object that the user will manipulate. The virtual object is then used as a texture mapped to the physical object, actually modifying its appearance. Lindlbauer *et al.* [56] demonstrate that beyond direct texture mapping, AR can be used to modify the surroundings of a physical object (for instance to display virtual extensions or cast shadows). Their system can modify an object's perceived size, color, or shape. HoloFace [48] uses face tracking to change people's face. While these projects have essentially focused on the technical realization of such alterations, Rixen *et al.* [84] have investigated their acceptability from a social perspective. They studied the potential of color alterations applied to objects or clothes, focusing more particularly on situations involving interpersonal communication. They identified limits to what users would accept both as the person altering others and as the person being altered.

AR can also be used to artificially *remove* elements from the physical environment or at least alter their visibility. This is what *Diminished Reality* [62, 70] refers to. The SceneCtrl

¹In the spirit of the HTC Vive Chaperone system that displays only the objects' outline.

system [102] relies on scene reconstruction and texture patches to enable users wearing an Optical See-Through (OST) display to delete or move a real object in their field of view.

As graphics rendering and spatial registration technologies improve, the mix of virtual content with the physical environment will be of such high quality that it will literally create “*a world in which the physical laws ordinarily governing space, time, mechanics, material properties, etc. no longer hold*” [67]. Some systems already offer such an experience in VR setups. For example, the Remixed Reality system [57] relies on depth cameras positioned in a room to capture a live 3D model of it, and then manipulate it before display in a Video See-Through (VST) headset. The system makes it possible to give users the impression that they are flying or walking on walls. Still in a VR context, it was also recently demonstrated that it is possible to perform changes to a virtual environment *in plain sight* without users actually noticing those changes² [63].

In situations where the world *perceived* by users does not match the world in which they actually *are*, methods to ensure safety must be implemented. Chaperoning techniques can for instance display grids or outlines of nearby obstacles (as does the HTC Vive Chaperone system mentioned earlier and used as a baseline in RealityCheck [31]) or indicate a safe trajectory for users to follow [42]. Some research projects also compute geometrical transformations applied to the virtual world in order to alter users’ perception and make them adopt safe walking trajectories [82] and perform safe gestures [1]. However, while those techniques can address collision issues, there are ethical issues and a broader notion of safety to take into account when considering alterations to reality [90], as Rixen *et al.* [84] have started investigating. Overall, while the possibility to alter reality yields exciting perspectives, it also raises practical and ethical concerns which, at the very least, call for a mechanism to let users remove augmentations on demand. De-augmentation capabilities offer a solution to address some of those concerns.

2.3 Augmented Reality

Still moving towards predominantly-physical environments, we enter the part of the continuum that is usually referred to as Augmented Reality. That term, however, encompasses a variety of content types and situations. Virtual content embedded in the physical environment is not necessarily about simulating physical objects such as furniture in a house [97] or underground pipes in the street. It can consist of, for instance, way-finding aids [40], data visualizations [88, 99, 100] or personal information management interfaces that display elaborate 2D components such as, *e.g.*, documents, maps, videos, address books, todo lists as well as UI widgets [24, 54, 55].

One of the main challenges of Augmented Reality lies in the combination of the physical light and virtual light so as to create visually-coherent [18, 105] representations that effectively blend the virtual and the physical; and to avoid situations where too much virtual content can clutter the user’s field of view [87], making the virtual content illegible and interfering with the perception of the physical environment.

Factors that have an adverse impact on visual coherence include poor virtual content rendering performance (field of view, resolution, frame rate); improper tracking and registration; and very importantly the lack of depth cues [49]. Such depth cues play a key role in the integration of virtual content in the scene [25]. As researchers realized very early [3], a naïve overlay of the virtual content on top of the real world yields augmentations that do not seem to be integrated at all in the physical environment, leading to serious perceptual problems [105]. Solutions involve rendering proper shadows, and managing mutual occlusion between the physical and the virtual. Mutual occlusion refers to the problem of showing and hiding both virtual and physical elements based on their depth ordering [59]. Similarly, when AR is used to provide users with so-called “X-ray vision” [27] that allows them to see *through* physical objects, multiple issues need to be addressed for users to be able to make the distinction and understand the topological relations between the virtual and the physical elements [59].

²This can be achieved using an eye-tracker to determine where users’ visual attention is in order to make changes in areas that they are not focusing on.

De-augmenting augmented reality relates to this general problem of making the virtual and the real co-exist. However, as detailed in the next section, it rather focuses on how to prevent the virtual content from visually interfering with what users *want* or *need* to perceive from the physical environment.

3 De-augmenting AR: a Framework

Following this examination of key concepts from the reality-virtuality continuum, we define a de-augmentation as *an operation that alters virtual content in order to improve the user's perception of the real world*. The definition is kept vague as we believe that the concept of de-augmentation could be relevant not only for visual content, but possibly for modalities such as audio and haptics as well. As we focus here on visual de-augmentations, we also give the more specific definition of a *visual* de-augmentation: *an operation that alters virtual content in the user's field of view in order to improve their visual perception of the real world*.

In this section, we decompose a de-augmentation into three main characteristics (Figure 1):

- what *triggers* the de-augmentation;
- what its *scope* is, *i.e.*, what part of the real-world gets de-augmented and for how long;
- how the de-augmentation is achieved, *i.e.*, how the virtual content's *rendering* is getting altered.


This section details these different characteristics and illustrate them with different examples. Some examples come from our own thinking and discussions around the concept, and others are taken or inspired by the literature about Mixed Reality in general (*i.e.*, going beyond AR-only projects by considering research prototypes for Mixed Reality in general, including VR). Figure 1 lists prior work that was key to the identification of the different categories in our framework. The idea of bringing the real back in the user's field of view has been almost exclusively considered in the literature about VR [31, 41, 66]. In an AR context, most relevant to our work are references about the concept of pervasive or adaptive AR where the amount and presentation of AR content depends on the user's context [30, 39, 55]. In terms of rendering, a de-augmentation consists of reducing the amount of virtual content in the scene that users perceive. Although the de-augmentation concept is different from the initial goal of AR from the user perspective, it still consists of rendering virtual components over the real. Our categories thus mostly align with the categories in Zollman *et al.*'s general taxonomy about visualization in AR [105].

The overall goal of our work is to provide AR designers with a general and unifying framework that helps them think about situations that call for de-augmentation and design interaction techniques to support such operations accordingly.

3.1 De-augmentation Triggers

AR de-augmentations happen in response to a trigger. We distinguish three main types of triggers. The first two are under users' control: fired either on-demand by invoking a *User Command*; or fired automatically, the system triggering *User Rules* that match the current situation. The last trigger type is also rule-based, but rules are specified by a user-independent *External Source*.

3.1.1 User Command

A de-augmentation can be triggered with a *User Command*  invoked explicitly. As illustrated in Figure 1, revealing the physical environment in response to an explicit action from the user has only been considered in a VR context for immersive gaming environments. In such a

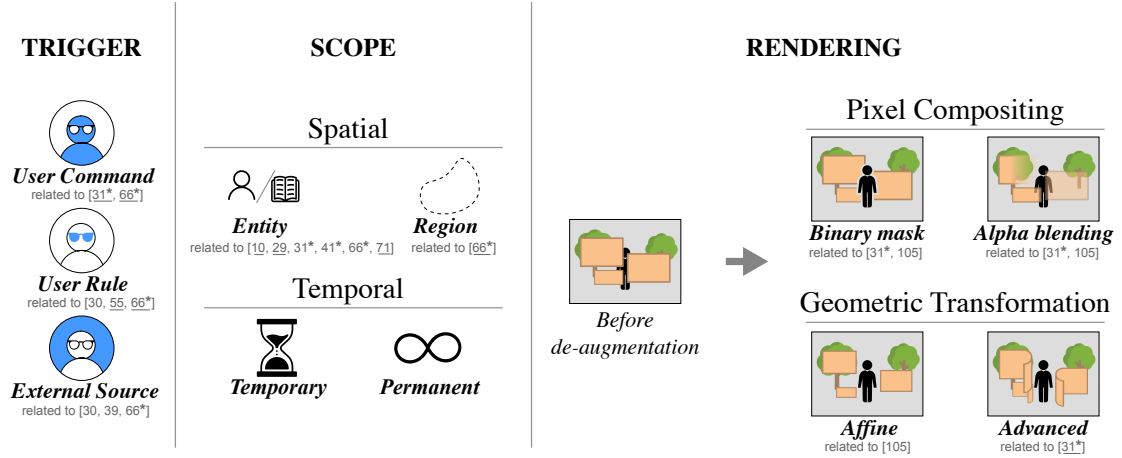


Figure 1: The three characteristics of a visual de-augmentation operation (Trigger, Scope and Rendering). Pieces of work from the literature that are directly related to those characteristics are referenced. References marked with an asterisk (*) are studies or projects in VR rather than AR. Underlined references are techniques or systems that are an illustrative instance of the general category.

VR setup that completely occludes the real, users may rely on their game controller to invoke a view of the physical world. For instance, in McGill *et al.*'s experiment [66], users can press a button on their game pad to invoke a view of the real, while in the RealityCheck system [31], users can move their controller as if it were a flashlight to expose the real-world in its cone.

In an AR context, diverse types of input modalities can be considered depending on the context and type of de-augmentation to trigger. For example, a commuter checking the news in his AR glasses while waiting for the train could point at the screen with train departure information to de-augment it, so as to always have a clear view on it but without interrupting his news reading activity. This example involves hand movements (gestures), but any other interaction technique that works in conjunction with AR technologies can be considered, such as voice or gaze. One could for instance envision giving users some control on the level of (de-)augmentation based on how they look through their glasses. Just the same as the wearer of prescription glasses equipped with progressive lenses gets a different amount of correction depending on whether they look through the lower, middle, or upper part of their eyewear, users of AR glasses could control virtual content (de-)augmentation in a similar way. The level of control would be basic, but always available, requiring only little gaze adjustments.

Depending on the other properties of the de-augmentation to be specified, more expressive input might be required, for instance to define a specific scope or a specific way to de-augment the virtual content. Simple deictic pointing combined with computer vision techniques can for instance be used to de-augment a physical object. Freeform gesture tracking enables users to delineate arbitrarily-shaped spatial scopes such as, *e.g.*, a sketch on a whiteboard or an irregular area on the floor. We can even think of using other modalities such as a handheld device (*e.g.*, a smartphone) or tangibles. For example, drawing inspiration from previous work [12, 31], a tangible could play the role of a lamp that de-augments the region it illuminates. Tangible posts could also be used to delimit a large area, in the spirit of the ShareSpace system [101].

Finer control over the de-augmentation's properties can also be provided using elaborate widgets in AR, for instance to adjust specific rendering parameters such as the transparency of de-augmented areas. However, we can expect that elaborate, time-consuming de-augmentation specifications would rarely be made on the spot, but would rather be defined offline and activated when relevant using simple interactions or automatically-triggered rules, as discussed next.

3.1.2 User Rules

Contrary to user commands, **User Rules** ☹ do not require explicit actions from the user. They have originally been configured by the user, but are triggered by the AR system automatically, whenever they apply to the current context. Context-awareness is of foremost importance to the realization of pervasive AR (PAR). As noted by Grubert *et al.*: “PAR systems are mainly controlled by the user’s current context factors and to a lesser extent by direct user control” [30].

We share this vision that, when possible, direct user input should be kept to a minimum. However, the diversity of contexts and users is such that we can neither expect designers of AR systems to anticipate all situations nor can we expect PAR systems to be smart enough to guess what users want or need in all situations. This means that, in any case, AR systems should involve some dialogue with users in order to suit individual users’ needs. For example, a user might configure their system so that a de-augmentation is systematically triggered to let that user and their relatives see each other. One way of achieving this could be to set a rule so that when at home, people’s face always gets de-augmented. This rule makes sense at home, but might not be as relevant when walking down the street. Another user might choose to systematically de-augment passers-by, while yet another one might prefer de-augmentation to apply automatically to friends and acquaintances only, and possibly to strangers meaning to initiate a conversation.

User rules can specify different situations that call for de-augmentation. For example, a rule could de-augment a window in a house whenever the user stands less than one meter from it to enjoy a stunning view. Another rule could be that any device that triggers an alarm gets de-augmented (phone, alarm clock, safety device). In McGill *et al.*’s series of experiments about VR gaming [66], one of the tested conditions to bring some of the reality in the VR world is based on a rule that triggers a de-augmentation when users’ hands are detected in the view of the headset’s camera. User Rules can also be functions of the user’s anatomic or physiological state [30] in the spirit of Lindlbauer *et al.* [55]’s system that automatically controls what amount of virtual content to show, where and when to show it, all depending on the user’s task and cognitive load. We can also envision rules that would fire at regular intervals. For instance, prior work has pointed at the actual risk associated with too much AR assistance, that could make everyday life *too facile* [44], and make dependency on technology too strong [86]. People who want to self-regulate their use of AR or just make sure that they stay in touch with (fully-unaltered) reality could configure a rule to de-augment everything for, *e.g.*, a few seconds every five minutes.

Setting a rule as opposed to invoking commands removes the need to take explicit action to trigger a de-augmentation. As illustrated just above, rules can also specify a scope (*e.g.*, people’s face or the window) for the de-augmentation, thus reducing user input to its minimum.

3.1.3 External Source

While many rules will have been configured, or at least activated, by the users themselves, some rules can be imposed by **External Sources** ☹. These de-augmentations are also context-related but are beyond users’ individual control. In their vision of Pervasive AR, Grubert *et al.* [30] mention how environmental factors like movements of people around the user can play a role in the presence or absence of AR. We can find such an example in a VR context in McGill *et al.*’s last experiment [66] where people coming close to the wearer of a VR headset gets automatically revealed in the VR world. Another example of external source would be if the authorities in a given area have a policy that enforces a no-AR or limited-AR policy for safety, privacy, or security reasons in a similar spirit that phone usage is prohibited at airport customs. Buildings such as police stations or any type of restricted area where sensitive activities are performed are obvious candidates for such de-augmentation rules imposed to all users.

When some higher-level knowledge about the environment is available, the AR system could take into account distance and visibility information as well as higher-level knowledge about

the environment to filter virtual content [39]. A traffic-heavy street about to be crossed would be a good example. For similar safety reasons, an AR system could also partially de-augment any object that moves above a certain velocity threshold. However, our experience with smartphones has shown that external control over individuals' personal devices is very challenging to implement because of technical and interoperability obstacles. We thus rather envision a future generation of AR systems that are configurable with respect to external triggers, empowering users to choose to receive such external notifications in order to enjoy their personal, immersive device responsibly.

A trigger from an *External Source* could also be controlled by a *User Rule* that exists on the user's side to authorize such de-augmentations initiated by third parties. For example, a de-augmentation request might be acceptable only when it originates from relatives, friends and colleagues: a person waving in the user's field of view would get de-augmented if known to the user, but would not if she was not known to the user.

3.2 De-augmentation Scope

The second defining characteristic of a de-augmentation is its scope, which can actually be broken down in two separate properties (Figure 1): the spatial scope, and the temporal scope. The spatial scope can be defined by physical entities (people, objects) no matter where they are; or it can be defined as a specific region, no matter what entities it contains. The temporal scope is orthogonal to the spatial scope. It can be either permanent or temporary.

3.2.1 Spatial Scope [Entity]

De-augmenting an entity $\mathcal{R} \boxtimes \mathcal{E}$ means de-augmenting virtual content that visually interferes with the perception of this entity, regardless of its actual location in the physical world. Typical examples include people (face or entire body); passive physical objects (*e.g.*, a workbench, a mirror); as well as physical displays such as desktop computer screens or smartphones. An entity's location can change over time, implying that the system must be able to track it in the user's field of view in order to dynamically adapt the geometry of the de-augmented area.

Although not presented as a de-augmentation operation, a typical example of an entity de-augmentation can be found in situations where an AR display is used in conjunction with a mobile device [29, 104]. The MR systems enabling this have to handle visual coherence in the sense that the virtual content is usually tightly aligned with the handheld device's surface [8], sometimes aiming to achieve a display that seamlessly extends beyond its boundaries [71]. Budhiraja *et al.* [10] investigate techniques that achieve such combinations. One of these techniques actually *de-augments* the phone when users look down at it, so as to prevent virtual objects displayed in the headset from interfering with the phone's screen. In the pure VR context studied in McGill *et al.*'s experiments [66], some conditions *cut* the VR view to reveal a physical entity such as a keyboard or the user's hands.

A scope can also consist of multiple entities. This is the situation with chaperone systems for immersive VR setups that completely isolate the user from the real world. To mitigate the risk of collision, chaperone systems reveal all objects that are close to the user as they can be dangerous obstacles [31, 41]. In collaborative AR work contexts [7], a multi-entity scope could be useful to simultaneously de-augment a set of entities that vary in nature. For example, in a setup where users share the same overall virtual scene and use physical tools to interact with it [23], only selected physical tools could be de-augmented to facilitate their manipulation, but the faces of team members could always be de-augmented to preserve the directness and naturalness of face-to-face interpersonal communication [68].

3.2.2 Spatial Scope [Region]

De-augmenting a region \odot means de-augmenting virtual content that visually interferes with the perception of a region in the user’s field of view, regardless of what physical entities it contains. A region’s location does not change over time but its contents can change, without impacting the de-augmentation. A region can be either *world-anchored* or *FoV-anchored*. A world-anchored region is fixed in the physical world, regardless of the user’s location and viewing direction. A FoV-anchored region (sometimes referred to as *view-anchored* [55]) is defined relative to the user’s field of view, as if it were attached to the eyewear.

For example, a parent in a playground could de-augment the sandbox to keep an eye on their child. A worker could de-augment an area on their desk so that AR does not visually interfere with their notebook, or their desktop input device. Traffic lights could be de-augmented. Traffic lights can be considered as entities on a conceptual level. But because those entities are stationary, handling them as fixed regions removes the need for dynamic object tracking. We can find another example of a region-based scope in the game pad used as a flashlight to reveal the real world in the context of a VR game, which we already mentioned earlier [31]. The peephole onto the real depends on what direction the user points at with their pad, independently of the nature of what is being targeted.

FoV-anchored-region scopes are perhaps less common. They can be useful to guarantee the absence of visual interference in part of the user’s field of view regardless of any other consideration. They could serve purposes such as before/after comparisons: with half of the field of view showing virtual content but not the other, users could compare, *e.g.*, actual *vs.* virtual clothes or face with/without makeup in the spirit of dual-layer image comparison techniques [58]. A FoV-anchored de-augmentation could also act as a focus+context exploration technique, “drilling a hole” through the virtual layer, akin to a magic lens [5]. This corresponds to the full view of reality condition that is tested in the second study of McGill *et al.*’s [66] in a pure VR context.

3.2.3 Temporal Scope [Temporary]

A de-augmentation whose temporal scope is temporary ⌚ will only last some pre-defined duration, after which it will automatically get cancelled, the virtual content appearing back.

For example, a smartphone could get de-augmented for three or four seconds when notifications pop up as the user only needs to look at it for a few seconds. Having to cancel the de-augmentation explicitly would be cumbersome. Temporary de-augmentations can also be useful to lower the *level of engagement* [80], something an AR user might want to do in order to eat a snack for instance or see someone who entered the room.

3.2.4 Temporal Scope [Permanent]

A de-augmentation whose temporal scope is permanent ∞ lasts until it is explicitly cancelled, either by the user or by an outside event that results in a change of context. For instance, the rule that triggered a de-augmentation might no longer apply because the user is no longer in the corresponding physical area. As discussed further in section 5, the AR system should give users the means to cancel a de-augmentation that has a permanent temporal scope.


3.3 De-augmentation Rendering


Once triggered for a specific scope, a de-augmentation is achieved by altering the virtual content that falls within that scope (as projected in the user’s field of view). This part of our taxonomy mostly aligns with the *Compositing* dimension in Zollmann *et al.* [105]’s general taxonomy for visualization in AR. Although in our case the user’s goal is to restore the visibility of the real, rendering still consists – from a technical perspective – of blending the real and the virtual in a single visual scene, as with any AR-based visualization. We only leave out the *inpainting*

condition from Zollmann *et al.* [105]’s taxonomy as it consists of actually *editing* the real by means of AR, which is somehow opposite to the spirit of de-augmenting the real.


The simplest de-augmentation consists of not showing anything in the corresponding area of the user’s field of view, as if cutting a hole [22] in the virtual content to let light from the physical environment pass through that hole. Alternative rendering strategies can also be considered (Figure 1), as detailed next.


3.3.1 Pixel Compositing

Cutting holes in a layer to reveal hidden objects has been experimented in many different situations, including window management [92], map comparisons [58], or layered 3D models [11, 17, 79]. Drilling holes in the AR layer is a form of occlusion management [59]. It can be as straightforward as applying a binary image mask  to the rasterized virtual content right before display in the eyewear (*Overlay* in [105]).

Generalizing this approach, pixels that fall within the de-augmentation’s scope or in its surroundings can have their alpha channel adjusted  without necessarily making them fully transparent (*Blending* in [105]). Content-aware techniques [35, 96] could also be used to adjust the level of transparency of individual pixels in the AR display depending on how strong the interference with the real is in a given area. However, as we discuss in Section 4, blending configurations are limited, due to the difficulty of fully blocking light from the physical environment with current Optical See-Through displays.

3.3.2 Geometric Transformation

Virtual content can also be removed from a given scope in the user’s field of view by applying transformations to its elements (*Transformation* in [105]). One of the simplest transformations consists of applying affine transforms  to individual elements that intersect the de-augmentation’s scope. By translating and downscaling them, virtual elements can be prevented from visually interfering with physical elements to be de-augmented. Early systems investigated a simplified version of the problem, but focused on the viewability of digital labels [93, 28] as opposed to the viewability of physical objects. Kawamata *et al.* [43] developed a technique for the specific situation of de-augmentation for driving, where virtual content is spatially registered with the physical environment but offset upward to avoid interfering with road signs and with the road itself. More generally, re-positioning virtual elements to de-augment a specific scope could take into account semantic associations [16] between virtual content and physical objects to optimize digital information layout.

Beyond affine transforms, more advanced transformations  that distort the geometry of elements in the AR layer can be considered. For example, content-aware focus+context techniques [78] could be adapted to distort elements in the virtual layer so as to better integrate them with the de-augmentation scope’s shape. Desktop techniques can also be a source of inspiration, based on the metaphor of rolling or folding virtual layers as if they were paper [15, 20, 21]. We could even think of gaming-like effects as used in the RealityCheck system to reveal the real world around the location of a virtual shot [31].

4 Implications for Technical Implementation

We now report on the technical implications of adding support for de-augmentation in an AR system, that are derived from a proof-of-concept prototype we built. Our goal with this prototype is two-fold. First, as user experience with AR eyewear can only be fully appreciated from a first-person view, a prototype helps understand the differences between the different points in our framework. Second but of foremost importance, we want to assess the technical feasibility of de-augmentation. While we anticipate a future where AR eyewear is much lighter than the AR

displays that are currently available on the market, the principles should essentially remain the same. This section describes our prototype and the lessons that we learned.

4.1 AR technologies and De-augmentation

AR eyewear can be implemented using either an Optical See-Through display (OST) or a Video See-Through display (VST). OST and VST displays are similar in the sense that they are both head-worn and superimpose virtual content directly in the users' field of view. But the way each of them combines physical light and virtual light is radically different:

- OST displays let physical light directly reach the user's eyes, superimposing virtual content in their field of view using some type of optical combiner [51, 37]. The physical light traverses the optical combiner, getting more or less altered in the process and blended with the virtual light generated by a source embedded in the OST device.
- VST displays are often implemented using Virtual Reality (VR) headsets augmented with cameras. These headsets block the light from the physical environment, which only gets transmitted indirectly to the user's eyes as a video stream displayed on a near-eye display together with the virtual content. As the physical light is mediated [62], users never perceive the actual light from the physical environment, but rather a *digitized* version of that physical light which gets displayed on a screen and perceived by the user's eyes.

Because VST headsets block the physical light, digitizing and performing any kind of compositing before rendering, this technology offers maximum flexibility in terms of the scene that users' eyes perceive. Outstanding progress has been made recently, with the introduction of commercial VST headsets that combine very high resolution, a high frame rate and high quality cameras. VST headsets are thus often used by HCI researchers to simulate an *ideal* AR experience. However, because VST headsets completely block the physical light, this approach comes at the cost of isolating users from their physical environment. It is thus fairly unlikely that VST displays gain wide adoption beyond laboratories and very specific work contexts, because of their awkwardness in many social contexts, and because of the perceived feeling of isolation inherent to the approach. On the contrary, OST displays do not isolate users from their physical environment. They let the physical light go through, combining it with some virtual light using optics.

We believe that AR de-augmentation mostly makes sense for OST technology for two reasons. First, because it does not isolate users, it is the most likely candidate for broad adoption. We actually share Itoh *et al.*'s position that “[OST displays] are now considered the ultimate hardware for augmenting our visual perception” [37]. OST displays have drawbacks. For example, they still feature a field of view that covers only a limited region of users' actual field of view [4, 60], limiting the level of immersion. Also, the virtual content that can be displayed remains semi-transparent [36] as the optical combiner only enables adding light, not subtracting it [37, 51, 105], which can create problems of accurate color rendering [37, 51, 19]. However, these “issues” can also be seen from a positive perspective. In particular, these limitations mean that users still have some (even degraded) perception of the physical environment without it being mediated by the system. This is key to the concept of giving users' agency about the de-augmentation operations to invoke on the spot. Users can rely on information they get from the physical environment at any time to make the decision of invoking a de-augmentation operation and of specifying its properties (scope and rendering) based on the information they get from both the physical and virtual environments.

We thus purposefully chose to implement our prototype using an OST display. While it made development more challenging compared to a VST display, and the end result is perhaps less convincing in terms of performance and graphics rendering, it is a good way to gain insights about particular limitations of OST display technologies [37] to inform future work about the further technical realization of de-augmentation techniques.

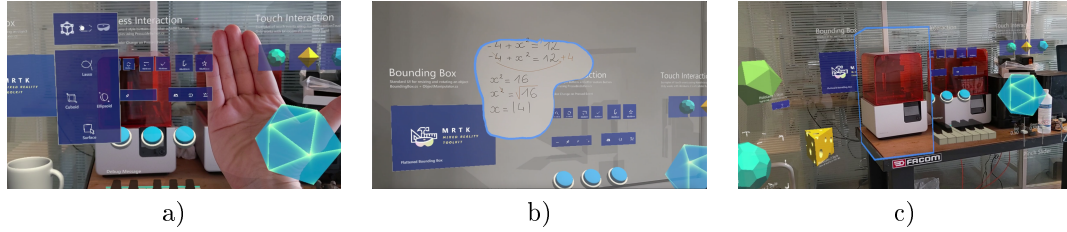


Figure 2: De-augmentations performed with our prototype on the `HandInteractionExamples` scene from MRTK. The virtual scene has been chosen for the purpose of cluttering the view with many objects that look obviously virtual to the reader. a) User invoking a hand menu with their hand flat in the HoloLens’ viewport. The user can choose how to delimit the scope of a de-augmentation to be manually created. b) Example of a world-anchored freeform flat (2D) de-augmentation to reveal equations written on a whiteboard. c) Example of a world-anchored cuboid de-augmentation to reveal a 3D printer in a fablab.

4.2 Representative Sample of De-augmentation Operations

Our proof-of-concept prototype runs on a Microsoft HoloLens 2 (OS Build 10.0.19041.1131), and was developed with Unity 2019.4.11f1 and the Microsoft Mixed Reality Toolkit (MRTK 2.5.4). We paid particular attention to illustrating different types of spatial scopes as the different possibilities in this part of the framework are more difficult to picture without illustrating them from a first-person perspective. Our prototype renders de-augmentations using pixel compositing. Pixels that fall in the spatial scope of a de-augmentation are made fully transparent – thus revealing the physical world – using one shader for each de-augmented area in the user’s field of view. The `ColorMask` property of shaders is set so that it removes any color in the area and forces the shape to be rendered first in the queue, right in front of the `near` plane. The de-augmentation mask will thus apply to all objects behind it in the virtual scene.

The prototype, illustrated in Figure 2 and demonstrated in the companion video, lets users create some de-augmentation operations by invoking a menu (Figure 2-a). They first specify the type of spatial scope: *world-anchored* or *FoV-anchored*. They can then select among multiple options to delimit that scope, which can be either 2D or 3D.

The first technique to define a world-anchored scope relies on MRTK’s `SceneUnderstanding` feature to detect surfaces in the physical environment. Users point at the surface³ they want to de-augment and perform an air tap to actually select it. Our other techniques involve drawing primitives such as lasso or rectangular selection tools. The interaction sequence differs depending on the drawing tool considered, seeking to make interaction as simple as possible. With the lasso tool, users can draw arbitrary shapes; the system will then sample points along the drawn curve and reconstruct a 2D surface in the 3D world (Figure 2-b). With our other area or volume selection tools, users have to specify only two points in 3D space. The system will automatically calculate how to constrain the 3D bounding box of the selected geometric primitive (cuboid, ellipsoid) so that two of its faces always remain parallel to the user’s sagittal plane (which is estimated based on the user’s head position). Regardless of how a 3D scope is specified, pixels that fall in its projection in the user’s field of view will be de-augmented, as illustrated in Figure 2-c.

Users define FoV-anchored scopes by delimiting an area on a 2D surface coplanar to the headset’s screen. Classic drawing tools are again used for this purpose, including lasso, rectangle and ellipse. An additional *eraser* tool lets users delimit a scope by holding one hand flat in front of them and moving it, as if they were “wiping the virtual plane”. As in the world-anchored case, pixels that fall in those areas will be de-augmented. The areas remain tied to the headset, however: de-augmentations are fixed in the user’s field of view, regardless of the incoming physical light.

³The system identifies the surface closest to the head-hand ray.

Beyond manually-specified de-augmentations, our prototype also supports a set of simple *User Rules* aimed at demonstrating automatically-triggered de-augmentations. These rules de-augment faces, the wearer’s hands, and objects that have been tagged. Hand tracking is readily available in the HoloLens. Face recognition relies on OpenCV 4.1.2 and HoloLens2ForCV [95]. Objects are tracked by attaching Vuforia tags to them and declaring them in the system.

4.3 Lessons Learned

Testing de-augmentation operations with our prototype made us quickly realize that it can be difficult to differentiate augmented from de-augmented zones. We decided to paint the border of de-augmentation masks with a solid color to help perceive a de-augmentation’s spatial scope (blue stroke around the equations and 3D printer in Figure 2-b/c). As is maybe most obvious in Figure 2-b, such a basic rendering at the boundary between the augmented and de-augmented zones fails to provide proper depth cues. This in turn makes it difficult for users to spatially situate de-augmentations relative to virtual content. Of course, motion cues alleviate this problem (see the companion video to this article) but nevertheless, this demonstrates that rendering proper depth cues will be as critical to effective de-augmentation as it is to effective augmentation. Adaptations of techniques such as ghosting, the addition of shadows, and rendering smoother transitions between augmented and de-augmented areas (as has been investigated with some “x-ray vision” techniques) should be investigated in future work. In the specific case of de-augmenting layered 3D models with cut-away views, rendering strokes where solid parts of virtual 3D content are actually cut will also be key to providing users with effective cues [11, 79] to relate the virtual and the physical.

Development of this prototype has also confirmed how sensitive the effective rendering of de-augmentations is to both latency (which impacts temporal realism [37]) and to precise spatial registration (which impacts spatial realism [37]). FoV-anchored de-augmentations should remain perfectly fixed in the user’s field of view, while world-anchored de-augmentations should remain perfectly aligned with the physical environment. This is of course much easier to handle with VST displays than OST displays. But current OST display hardware actually features tracking and rendering performance that is reasonable already, at least for the purpose of implementing de-augmentation. The video demonstration of our prototype in the companion video does not reflect the actual user experience fairly, suggesting significant lag, jitter and spatial registration offset. This, however, is due to limitations with the HoloLens 2’s Mixed Reality video capture feature used to produce those video clips. While it might not be perfect, users wearing a HoloLens 2 and trying the prototype do not experience lag, jitter and offset issues of this magnitude.

Beyond aspects related to the rendering of de-augmentation operations, our prototype helped us identify numerous challenges in terms of interaction design to specify the scope and rendering of a de-augmentation when this is done on the spot, which we detail in the next section.

5 Implications for Interaction Design

During the development and testing of our prototype, we quickly ran into issues related to interaction design. Giving users agency to specify the characteristics of a de-augmentation operation might actually require a lot of actions from them (defining the de-augmentation trigger, its spatial and temporal scope, possibly its rendering). In this section, we discuss solutions to benefit from de-augmentation without having to perform interactions that would be too cumbersome.

Trigger. One option to reduce interaction to its minimum is, obviously, to rely on *User Rules* only, which would be set upfront and then automatically triggered when a situation matches a rule. However, as mentioned above, it is neither possible for users to think of all rules upfront nor is it possible for the system to always succeed in accurately recognizing a specific situation.

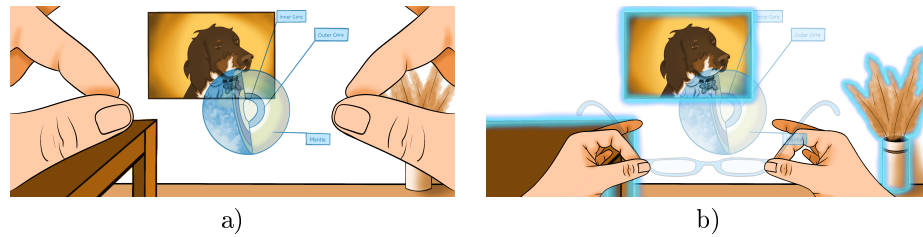


Figure 3: Entering a de-augmentation mode with a hand movement based on the metaphor of removing glasses. a) Users bring their two hands in the field-of-view with a posture that suggests taking off glasses. b) The system provides some feedback by displaying a pair of virtual glasses. Entities from the real environment are outlined with a blue halo, letting users directly point at those that they want to actually de-augment.

Below, we focus on the case where users manually specify a de-augmentation on the spot with *User commands*. It could be either to invoke a specific de-augmentation just once or to specify a rule with examples (programming by demonstration).

Spatial scope. When world-anchored, the scope of a de-augmentation is a 3D volume. Manual input for specifying a 3D volume without any knowledge about the environment is particularly challenging. Our prototype makes it possible to specify an arbitrary cuboid or ellipsoid volume with a rubber-band gesture performed in mid-air. However, not only is it cumbersome and awkward in some situations, but it also requires some walking when the volume is not fully within arms reach. A much more appropriate and quick interaction is to instead enter a “scene understanding” mode and use simple deictic pointing gestures to specify a specific surface. In our prototype, users invoke a menu to enter this scene understanding mode, but a quicker action such as adopting a specific hand posture would be more efficient. For instance, users could bring their two hands in their field of view and then remove them, mimicking a gesture to take off their glasses (Figure 3). More elaborate scene understanding capabilities (*i.e.*, surface, face and object detection) would help streamline the interactive de-augmentation process with users simply pointing at an entity in the physical environment to snap a de-augmentation to a larger set of entities [72]. However, while relying on the system’s scene understanding capabilities would work well for a single entity, it does not address the case of a region that may encompass several entities, or no entity at all.

For cases where the user cannot rely on the system’s understanding of the physical environment, several options are available. One interesting solution consists of relying on *tangible* de-augmenters. The number and size of tangible de-augmenters would depend on the nature of the region to delimit. For example, one might attach a small de-augmenter token to their keychain to de-augment a region of a few centimeters in radius around that token; a parent might place a series of tangible posts to delimit the area where their child is playing (Figure 4-a). Another option consists of using *hand postures* for a FoV-anchored scope. For example, a flat hand is a posture that is easy to recognize with any hand tracking system. It could be used to de-augment a region of the user’s field of view. In our prototype, the *eraser* tool lets users wipe out the AR canvas by moving their hand flat with the back of the hand visible. Such a simple posture could also be used for other types of FoV-anchored scopes. For example, users could move their hand flat with the edge of the hand visible to de-augment their field of view along one specific direction, following a *drawer* metaphor (Figure 4-b) or even use two hands to delimit a specific vertical or horizontal corridor in their field of view.

Temporal scope. Our framework makes the distinction between permanent and temporary de-augmentations. Testing the de-augmentation concept with our prototype made it clear that,

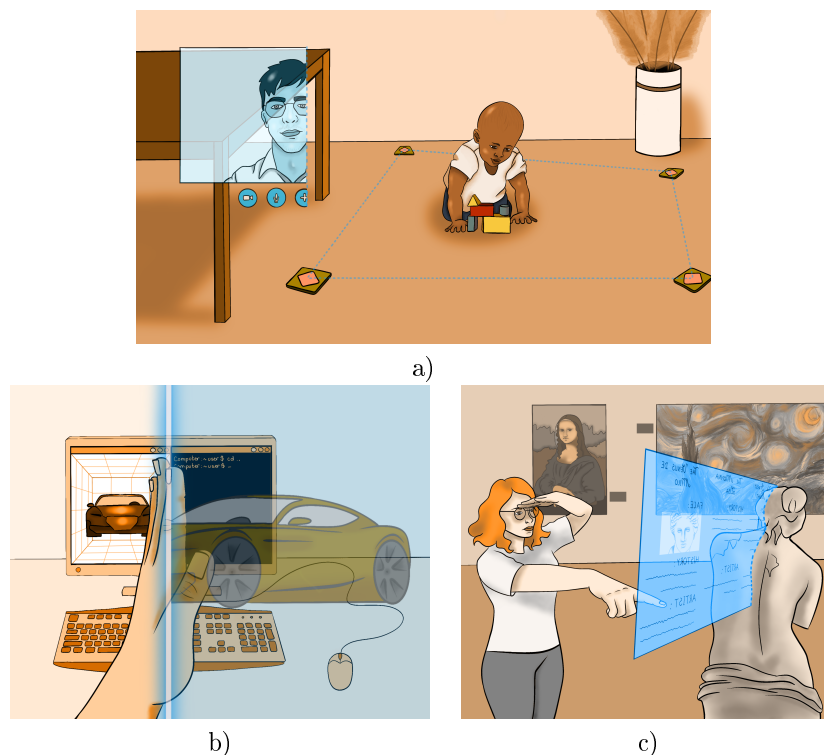


Figure 4: Examples of interaction for de-augmenting. a) A parent has placed tangible de-augmenters to delimit the volume in which the kid is playing. b) A software developer de-augments a part of their field of view to take a look at their code and understand why the orientation of the hologram is wrong. c) A user is reading summary information about the different pieces of art in this room of the museum. She enters the de-augmentation mode to clear virtual content actually in front of the Venus de Milo and relate what she reads in AR to her observation of the actual statue.

while users need agency regarding the possibility to de-augment, they should also be able to *re-augment*, *i.e.*, cancel a de-augmentation that they initially had set as permanent. Some techniques that we discuss above actually feature an obvious reciprocal version that can be performed to re-augment the field of view. For example, a tangible post can be removed; a discrete hand posture can be released. But the reciprocal action for techniques that rely on deictic pointing or on continuous motion – such as the *eraser* or *laser* – is less obvious. In such cases, a specific cancel action should be defined. Users could for instance bring their hand flat over a de-augmentation in their field-of-view and then close their fist as if they were crumpling the de-augmentation stencil.

Specifying a duration for a temporary de-augmentation does not seem to be something that can be easily done on the spot. Relying on a tensed hand posture to set a quasi-mode during which a given de-augmentation is active seems to be a more viable option. For instance, users could put their non-dominant hand flat on their forehead (as one would to avoid being dazzled by the sun) and then point at an entity in their environment with their dominant hand to de-augment that entity for the time they hold a flat posture with their non-dominant hand (Figure 4-c). In the same spirit, they could also hold a quasi-mode with their gaze by specifying a de-augmentation while looking through the upper part of their glasses. The de-augmentation would then be canceled as soon as they look through the lower part of their glasses again.

Rendering. Same as setting a specific duration, specifying a rendering strategy on the spot would be quite cumbersome. Some interaction techniques provide a good metaphor that matches

the rendering technique well, such as the *eraser* tool that will progressively clear a region, making pixels' transparency depend on how much they have been "wiped" by the user's hand. But finding such a direct interaction / rendering metaphor is an exception. Generally speaking, the specification of an arbitrary geometric transformation would be much more challenging. We rather envision the rendering strategy of de-augmentations as either a choice by design or as a parameter that users could set at the application, context or system level. For example, a de-augmentation would by default be rendered as a simple binary mask, except for notifications which would be scaled down and translated to the bottom-right corner of the field of view. A user could also decide to configure a paper-fold metaphor around the de-augmentation's spatial scope, folding all virtual content around a physical object of interest, such as a movie screen.

6 Illustrative Scenarios

Taking a leap into the future in which AR has indeed become pervasive, we consider a series of scenarios in which users are interacting with both virtual content and the physical environment and might want to remove part of the virtual content. We rely on the de-augmentation framework to propose solutions that would effectively let users achieve this.

6.1 At Home

Alex is a data analyst working from home. As their office space is not very large, their desk cannot accommodate more than one screen. Nevertheless, they are able to display multiple pieces of content beyond the boundaries of that single screen [16, 54] using their high-resolution wearable AR glasses. Alex is deeply focused on their work when their phone alarm goes off. This automatically triggers a de-augmentation of the corresponding physical area around the phone, enabling Alex to check and discard the alarm. They see that it is time to prepare dinner for the family. They exit work mode – which immediately resets their AR eyewear to home mode. In this mode, incoming AR notifications appear on the left side of their field of view. Alex likes this setting as they can check messages while they move around the house.

They start cooking dinner and launch an AR cooking application to get virtual instructions about how to prepare a new dish they have been wanting to try for a while. Meanwhile they receive a message – also displayed in AR – from their partner that they will be home soon. In the kitchen, the induction cooking plates and the smart oven – appliances for which safety precautions are automatically enforced – are always de-augmented. This is not only for safety but for practical reasons as well, since it enables direct monitoring of the cooking process even when both hands are busy. Figure 5-1 illustrates two de-augmentations triggered while looking at the oven: the AR recipe gets partially hidden where it intersects the oven; and the AR notifications (cooking plate countdown and message from partner) get translated to a part of Alex's field of view where they do not interfere with the physical environment.

About 15 minutes later, Alex's partner has come home. Dinner is ready. Alex goes upstairs to fetch their kids. One of the kids is playing an AR game that involves catching butterflies flying around in the bedroom. As Alex enters, their face gets automatically de-augmented (Figure 5-2). This way they can talk to each other without AR interfering and adversely impacting social interactions [68].

6.2 Video Call

Jen is taking a stroll in the park. She just noticed a poster advertising a live music concert when her partner, Sam, calls her (Figure 6-1). He is at the bakery and wants to know what type of bread Jen would prefer. One of the two *baguettes* definitely looks tastier than the other. Once the matter is settled, Jen asks Sam to remain on the call. She thinks the concert advertised on the poster is by his favorite band, and she wants to discuss the possibility of going.



Figure 5: Some elements from the At Home scenario, illustrated with and without de-augmentation. (1) While Alex is cooking in the kitchen, the oven beeps, leading them to look at it. Left: Alex's field of view without de-augmentation. The video recipe and notifications displayed in AR interfere with elements of the oven, such as the main control knob. Right: same field of view, with de-augmentation enabled. The oven is always de-augmented. The video recipe and the notifications get de-augmented differently: the notifications get translated to the bottom right corner, while the recipe gets cropped. A red dashed stroke is painted on the border of the de-augmented area to further emphasize the safety aspect. (2) One of Alex's kids is playing an AR game about catching butterflies. Left: Alex enters the kid's room. The kid looks at them, but AR butterflies flying around sometimes obscure their face. Right: Alex's face is automatically de-augmented, cropping the butterflies when necessary. This time, the border of the de-augmentation area is rendered as a smooth transition (using mask feathering) as there is no particular need to emphasize that boundary.

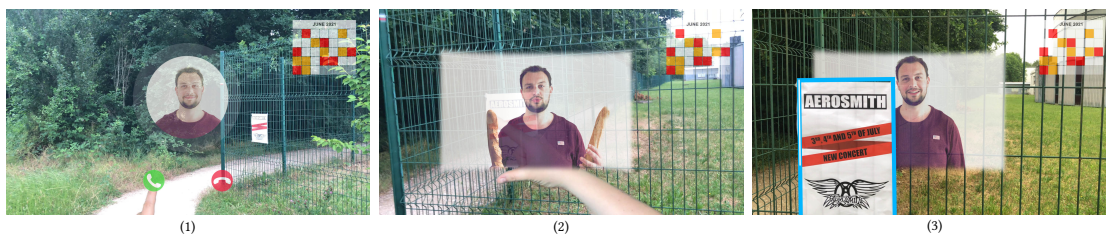


Figure 6: De-augmentation from the Video Call scenario. (1) Incoming call from Sam. (2) A concert poster catches Jen's attention, but she cannot see the details. She points at the poster to de-augment it while talking with Sam. (3) The poster is now de-augmented: Sam's video feed is still visible, but cropped so that it no longer interferes with the poster.

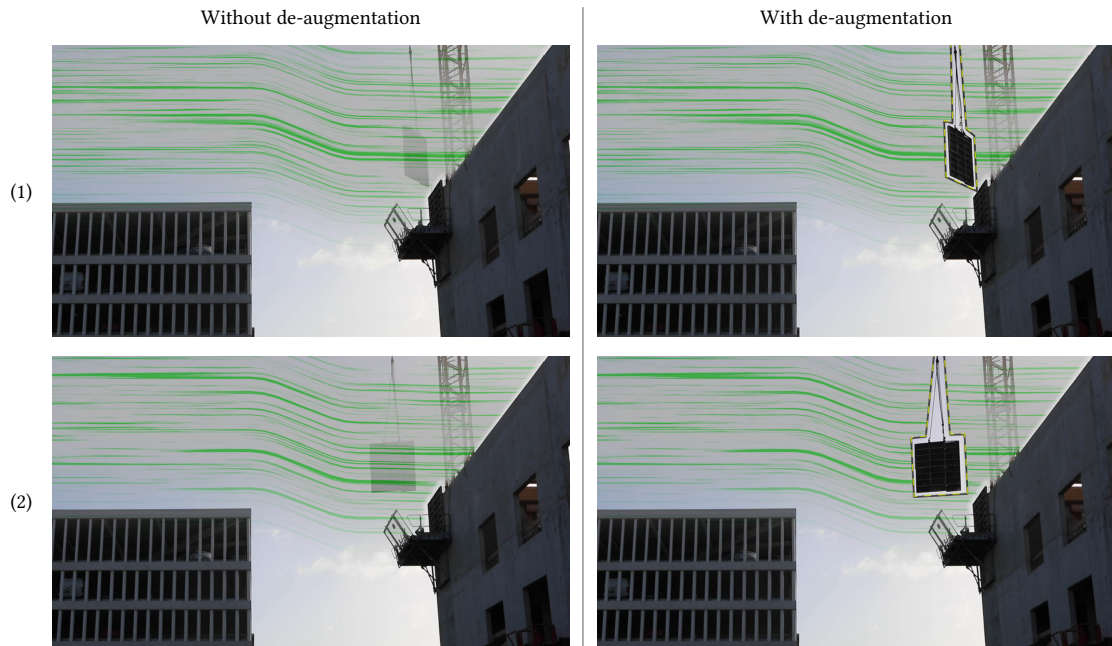


Figure 7: Some elements from the Safety First scenario. Kim is visiting the construction site. She is visualizing the wind flow simulation above one of the buildings. Another building is under construction on the right-hand side. A tower crane has just started lifting a heavy object, which is obscured by the virtual content. Left: without de-augmentation, the object’s movement is barely noticeable. Right: with de-augmentation, the object gets automatically de-augmented as its speed increases, making it easy for Kim to spot.

Jen got nearer the poster while she was talking with Sam. She is now close-enough to read the details, but Sam’s video feed is in her way. Jen points at the poster (Figure 6-2) with her dominant hand while holding her non-dominant hand on her forehead to temporarily de-augment the poster. She confirms that this is indeed the band, and finds out when they will be performing (Figure 6-3). She asks Sam whether he would like to go, and suggests one of the three performance dates. Sam seems to remember that they are already invited to a party on that day. Jen checks her calendar in AR, and confirms that there is a conflict indeed. Having both her calendar and the concert poster in her field of view, she finds another date that works. Jen removes her non-dominant hand from her forehead to cancel that temporary de-augmentation. Jen and Sam keep chatting, Jen asking Sam to book tickets as soon as he is home.

6.3 Safety First

Kim is a civil engineer visiting a construction site with her team. Like all other members of the group, she wears an AR-enabled hard hat. The primary purpose of these hats is safety, but the group also uses them to display two types of information that are central to the ongoing discussion: 3D models of concrete structures that are yet to be constructed; and a visualization of wind flow simulated for this specific urban environment [83] – see Figure 16 in that paper for an example of such an embedded [100] wind flow simulation visualization.

As in the *At Home* scenario, the MR system makes sure that virtual content is never displayed in front of other team members’ face. But other elements from the physical environment should be de-augmented as well. The 3D flow visualization is pretty dense and clutters a significant part of Kim’s field of view (Figure 7). The visualization is highly relevant to the task at hand, but

it also obscures much of the physical environment and its potential hazards. Still, Kim and her team feel safe, as their hats are configured with rules that prevent them from potential hazards. For instance, one rule automatically de-augments moving objects such as, *e.g.*, vehicles passing nearby or heavy weights lifted by cranes.

Back to the office after this site visit, Kim suggests they go get a coffee. By company policy, AR is not allowed in the cafeteria to ensure healthy, relaxing breaks promoting truly-social face-to-face interaction. Management had originally considered asking people to remove their eyewear in the cafeteria, but quickly realized that there were legitimate reasons to keep wearing it in there too. For instance, some employees have an AR display integrated in their prescription glasses. One of Kim's co-workers recently bought the latest technological innovation: true-color AR contact lenses. As any other lenses, those would be quite impractical to remove just to have a coffee break. Another co-worker has a color vision deficiency, and uses their AR eyewear to help them differentiate colors [91]. Management eventually opted for another solution: people are still allowed to keep their eyewear, but the cafeteria area is delimited by specific posts. These posts can be active and automatically instruct AR devices to de-augment everything that falls within the delimited region and that is not safety-critical or related to a medical condition. Management can also decide on a less directive approach where users simply receive notifications encouraging *de-augmentation* within the cafeteria area. Relying on tangible posts also has the advantage of clearly indicating to users the area where de-augmentation is strongly encouraged.

7 Conclusion

Overall, this paper provides a common framework to talk about de-augmentation, organize de-augmentation techniques, and inspire future designs. It also aims at helping designers integrate de-augmentation in their AR experiences.

Technological advances let us foresee a future where AR eyewear has become so lightweight and inconspicuous that reality can be augmented most of the time, if not all the time. The purpose of this paper is not to advocate for such a future, but to highlight the need to give users some agency over the display of virtual content, enabling them to *de-augment* augmented reality. While toggling the full AR display off is a special case of de-augmentation, we are also interested in locally-bounded de-augmentations, that remove virtual content in some parts of the users' field of view, where it actually interferes with their perception of reality. The concept of de-augmentation that we discuss here builds over the concept of pervasive AR where the AR system would automatically adapt to the user's context. With AR *de-augmentation*, users can also keep control over what is shown and what is hidden rather than rely on automatic adaptation only. If we take the example of users viewing a recipe in AR, they might just as well want to de-augment a parent who shows up in the kitchen to have face-to-face communication as they might want to keep the recipe in AR over that parent's face to relay the recipe's instructions. The *de-augmentation* framework is flexible enough to cope with both cases.

The notion of triggers that originate from an *External Source* that we explicitly list in our framework can raise questions though. In particular, what sort of control should AR and non-AR users have over the virtual content? What sort of control should authorities have in areas under their jurisdiction? The notion of sensitive areas with restricted physical access, or where taking pictures is not allowed, could be extended to limit possible augmentations. Safety-related de-augmentations such as automatically removing virtual content interfering with emergency exits would seem to make sense, but what if the user had special needs that made the virtual content key to their safety? Could these external triggers be overridden by the user's own system? These questions have safety and societal dimensions (privacy, ethics) that require investigation but that obviously go well beyond interaction design. The concept of de-augmentation seems relevant to the broader matter of AR-related laws and regulations [74, 98]. Keeping in mind that in many scenarios a full AR deactivation will be neither desirable nor necessary, and drawing a parallel

with accept *vs.* block lists, we can formulate that *while some cases are best handled by preventing an augmentation in the first place, other cases are best handled by de-augmenting part of what was augmented in the first place.*

Taking a more prospective look at future uses of de-augmentation, the concept might find interesting applications as we get closer to the ultimate goal of “*generat[ing] virtual objects that are so realistic that they are virtually indistinguishable from the real environment*” [2], possibly helping locally differentiate the virtual from the real. An important direction for future work will be to investigate the concept of de-augmentation applied to other AR modalities such as sound and haptics. Technological advances (audio spatialization, noise-cancelling techniques and ultrasound-based haptics just to mention a few) yield exciting opportunities to manipulate these signals. Here again, and as per the general definition of de-augmentation given in Section 3, the idea would be to investigate the feasibility and benefits of *partially* de-augmenting the sounds and haptic feedback generated by the AR system – as opposed to fully deactivating them – so as to improve users’ perception of the real world across more senses than vision alone.

Acknowledgment

We would like to extend our sincere thanks to Margaux Brasier for participating in the cooking scenario and to Valentin Sauvage for his fantastic help with the video and some figures.

This research was supported by French government funding managed by the National Research Agency (ANR) under: (i) the Investments for the Future program (PIA) grant ANR-21-ESRE-0030 (CONTINUUM); and (ii) the Collaborative Research Project (PRC) grant ANR-21-CE33-0022 (InterPlay).

References

- [1] Mahdi Azmandian, Mark Hancock, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2016. Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, 1968–1979. <https://doi.org/10.1145/2858036.2858226>
- [2] Ronald T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoper. Virtual Environ.* 6, 4 (Aug. 1997), 355–385. <https://doi.org/10.1162/pres.1997.6.4.355>
- [3] Michael Bajura, Henry Fuchs, and Ryutarou Ohbuchi. 1992. Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery within the Patient. *SIGGRAPH Comput. Graph.* 26, 2 (jul 1992), 203–210. <https://doi.org/10.1145/142920.134061>
- [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 *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (UIST '15)*. ACM, 129–135. <https://doi.org/10.1145/2807442.2807493>
- [5] Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony D. DeRose. 1993. Toolglass and Magic Lenses: The See-through Interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '93)*. ACM, 73–80. <https://doi.org/10.1145/166117.166126>
- [6] Mark Billinghurst. 2021. Grand Challenges for Augmented Reality. *Frontiers in Virtual Reality* 2 (2021). <https://doi.org/10.3389/frvir.2021.578080>
- [7] Mark Billinghurst, Adrian Clark, and Gun Lee. 2015. A Survey of Augmented Reality. *Found. Trends Hum.-Comput. Interact.* 8, 2–3 (March 2015), 73–272. <https://doi.org/10.1561/1100000049>

- [8] Eugenie Brasier, Emmanuel Pietriga, and Caroline Appert. 2021. AR-Enhanced Widgets for Smartphone-Centric Interaction. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*. ACM, Article 32, 12 pages. <https://doi.org/10.1145/3447526.3472019>
- [9] Pulkit Budhiraja, Rajinder Sodhi, Brett Jones, Kevin Karsch, Brian Bailey, and David Forsyth. 2015. Where's my drink? Enabling peripheral real world interactions while using HMDs. *arXiv preprint arXiv:1502.04744* (2015).
- [10] R. Budhiraja, G. A. Lee, and M. Billinghurst. 2013. Interaction techniques for HMD-HHD hybrid AR systems. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 243–244.
- [11] Michael Burns and Adam Finkelstein. 2008. Adaptive cutaways for comprehensible rendering of polygonal scenes. *ACM Transactions on Graphics* 27, 5 (2008). <https://doi.org/10.1145/1409060.1409107>
- [12] Andreas Butz, Clifford Beshers, and Steven Feiner. 1998. Of Vampire Mirrors and Privacy Lamps: Privacy Management in Multi-User Augmented Environments. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology (UIST '98)*. ACM, 171–172. <https://doi.org/10.1145/288392.288598>
- [13] Carlos Carbonell Carrera and Luis A. Bermejo Asensio. 2017. Landscape interpretation with augmented reality and maps to improve spatial orientation skill. *Journal of Geography in Higher Education* 41, 1 (2017), 119–133. <https://doi.org/10.1080/03098265.2016.1260530>
- [14] Chenliang Chang, Kiseung Bang, Gordon Wetzstein, Byoung-ho Lee, and Liang Gao. 2020. Toward the next-generation VR/AR optics: a review of holographic near-eye displays from a human-centric perspective. *Optica* 7, 11 (Nov 2020), 1563–1578. <https://doi.org/10.1364/OPTICA.406004>
- [15] Olivier Chapuis and Nicolas Roussel. 2007. Copy-and-Paste between Overlapping Windows. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, 201–210. <https://doi.org/10.1145/1240624.1240657>
- [16] Yifei Cheng, Yukang Yan, Xin Yi, Yuanchun Shi, and David Lindlbauer. 2021. SemanticAdapt: Optimization-Based Adaptation of Mixed Reality Layouts Leveraging Virtual-Physical Semantic Connections. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21)*. ACM, 282–297. <https://doi.org/10.1145/3472749.3474750>
- [17] Chris Coffin and Tobias Höllerer. 2006. Interactive perspective cut-away views for general 3D scenes. *3DUI 2006: IEEE Symposium on 3D User Interfaces 2006 - Proceedings 2006* (2006), 25–28. <https://doi.org/10.1109/VR.2006.88>
- [18] Jonny Collins, Holger Regenbrecht, and Tobias Langlotz. 2017. Visual Coherence in Mixed Reality: A Systematic Enquiry. *Presence* 26, 1 (2017), 16–41. https://doi.org/10.1162/PRES_a_00284
- [19] Juan David Hincapié-Ramos, Levko Ivanchuk, Srikanth K. Sridharan, and Pourang P. Irani. 2015. SmartColor: Real-Time Color and Contrast Correction for Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 21, 12 (2015), 1336–1348. <https://doi.org/10.1109/TVCG.2015.2450745>
- [20] Pierre Dragicevic. 2004. Combining Crossing-Based and Paper-Based Interaction Paradigms for Dragging and Dropping between Overlapping Windows. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*. ACM, 193–196. <https://doi.org/10.1145/1029632.1029667>
- [21] Niklas Elmqvist, Nathalie Henry, Yann Riche, and Jean-Daniel Fekete. 2008. Melange: Space Folding for Multi-Focus Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, 1333–1342. <https://doi.org/10.1145/1357054.1357263>

- [22] Niklas Elmqvist and Philippos Tsigas. 2008. A Taxonomy of 3D Occlusion Management for Visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 5 (Sept. 2008), 1095–1109. <https://doi.org/10.1109/TVCG.2008.59>
- [23] Barrett Ens, Sarah Goodwin, Arnaud Prouzeau, Fraser Anderson, Florence Y. Wang, Samuel Gratzl, Zac Lucarelli, Brendan Moyle, Jim Smiley, and Tim Dwyer. 2021. Uplift: A Tangible and Immersive Tabletop System for Casual Collaborative Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1193–1203. <https://doi.org/10.1109/TVCG.2020.3030334>
- [24] Barrett M. Ens, Rory Finnegan, and Pourang P. Irani. 2014. The Personal Cockpit: A Spatial Interface for Effective Task Switching on Head-Worn Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, 3171–3180. <https://doi.org/10.1145/2556288.2557058>
- [25] Taiki Fukiage, Takeshi Oishi, and Katsushi Ikeuchi. 2012. Reduction of Contradictory Partial Occlusion in Mixed Reality by Using Characteristics of Transparency Perception. In *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (ISMAR '12)*. IEEE Computer Society, 129–139. <https://doi.org/10.1109/ISMAR.2012.6402549>
- [26] Markus Funk, Andreas Bächler, Liane Bächler, Thomas Kosch, Thomas Heidenreich, and Albrecht Schmidt. 2017. Working with Augmented Reality? A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. In *Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '17)*. ACM, 222–229. <https://doi.org/10.1145/3056540.3056548>
- [27] C. Furmanski, R. Azuma, and M. Daily. 2002. Augmented-reality visualizations guided by cognition: perceptual heuristics for combining visible and obscured information. In *Proceedings. International Symposium on Mixed and Augmented Reality*, 215–320. <https://doi.org/10.1109/ISMAR.2002.1115091>
- [28] Raphaël Grasset, Tobias Langlotz, Denis Kalkofen, Markus Tatzgern, and Dieter Schmalstieg. 2012. Image-driven view management for augmented reality browsers. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 177–186. <https://doi.org/10.1109/ISMAR.2012.6402555>
- [29] 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)*. ACM, 3933–3942. <https://doi.org/10.1145/2702123.2702331>
- [30] Jens Grubert, Tobias Langlotz, Stefanie Zollmann, and Holger Regenbrecht. 2017. Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 23, 6 (2017), 1706–1724. <https://doi.org/10.1109/TVCG.2016.2543720>
- [31] 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, 1–12. <https://doi.org/10.1145/3290605.3300577>
- [32] Steven Henderson and Steven Feiner. 2011. Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. *IEEE Transactions on Visualization and Computer Graphics* 17, 10 (2011), 1355–1368. <https://doi.org/10.1109/TVCG.2010.245>
- [33] Anuruddha Hettiarachchi and Daniel Wigdor. 2016. Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, 1957–1967. <https://doi.org/10.1145/2858036.2858134>

- [34] Kenny Hong, Yumiko Sakamoto, and Pourang Irani. 2019. The Use of Head-Worn Augmented Reality Displays in Health Communications.. In *Information Technology and Communications in Health Conference*. IOS Press, 163–169. <https://doi.org/10.3233/978-1-61499-951-5-163>
- [35] Edward W. Ishak and Steven K. Feiner. 2004. Interacting with Hidden Content Using Content-Aware Free-Space Transparency. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*. ACM, 189–192. <https://doi.org/10.1145/1029632.1029666>
- [36] Yuta Itoh, Takumi Hamasaki, and Maki Sugimoto. 2017. Occlusion Leak Compensation for Optical See-Through Displays Using a Single-Layer Transmissive Spatial Light Modulator. *IEEE Transactions on Visualization and Computer Graphics* 23, 11 (2017), 2463–2473. <https://doi.org/10.1109/TVCG.2017.2734427>
- [37] Yuta Itoh, Tobias Langlotz, Jonathan Sutton, and Alexander Plopski. 2021. Towards Indistinguishable Augmented Reality: A Survey on Optical See-through Head-Mounted Displays. *ACM Comput. Surv.* 54, 6, Article 120 (jul 2021), 36 pages. <https://doi.org/10.1145/3453157>
- [38] M. Carmen Juan, Mariano Alcaniz, Carlos Monserrat, Cristina Botella, Rosa M. Banos, and Belen Guerrero. 2005. Using augmented reality to treat phobias. *IEEE Computer Graphics and Applications* 25, 6 (2005), 31–37. <https://doi.org/10.1109/MCG.2005.143>
- [39] S. Julier, Y. Baillot, D. Brown, and M. Lanzagorta. 2002. Information filtering for mobile augmented reality. *IEEE Computer Graphics and Applications* 22, 5 (2002), 12–15. <https://doi.org/10.1109/MCG.2002.1028721>
- [40] Jinki Jung, Hyeopwoo Lee, Jeehye Choi, Abhilasha Nanda, Uwe Gruenefeld, Tim Claudius Strattmann, and Wilko Heuten. 2018. Ensuring Safety in Augmented Reality from Trade-off Between Immersion and Situation Awareness. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 70–79. <https://doi.org/10.1109/ISMAR.2018.00032>
- [41] Kohei Kanamori, Nobuchika Sakata, Tomu Tominaga, Yoshinori Hijikata, Kensuke Harada, and Kiyoshi Kiyokawa. 2018. Obstacle Avoidance Method in Real Space for Virtual Reality Immersion. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 80–89. <https://doi.org/10.1109/ISMAR.2018.00033>
- [42] Kohei Kanamori, Nobuchika Sakata, Tomu Tominaga, Yoshinori Hijikata, Kensuke Harada, and Kiyoshi Kiyokawa. 2019. Walking Assist Method for VR Zombie. In *2019 12th Asia Pacific Workshop on Mixed and Augmented Reality (APMAR)*. 1–7. <https://doi.org/10.1109/APMAR.2019.8709270>
- [43] Takaya Kawamata, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta. 2013. Lifted road map view on windshield display. In *IEEE Symposium on 3D User Interfaces (3DUI)*. 139–140. <https://doi.org/10.1109/3DUI.2013.6550217>
- [44] Pascal Knierim, Paweł W. Woundefinedniak, Yomna Abdelrahman, and Albrecht Schmidt. 2019. Exploring the Potential of Augmented Reality in Domestic Environments. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '19)*. ACM, Article 31, 12 pages. <https://doi.org/10.1145/3338286.3340142>
- [45] Marion Koelle, Abdallah El Ali, Vanessa Cobus, Wilko Heuten, and Susanne CJ Boll. 2017. All about Acceptability? Identifying Factors for the Adoption of Data Glasses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, 295–300. <https://doi.org/10.1145/3025453.3025749>
- [46] Marion Koelle, Matthias Kranz, and Andreas Möller. 2015. Don't Look at Me That Way! Understanding User Attitudes Towards Data Glasses Usage. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, 362–372. <https://doi.org/10.1145/2785830.2785842>

- [47] Shalva Kohen, Carmine Elvezio, and Steven Feiner. 2020. MiXR: A Hybrid AR Sheet Music Interface for Live Performance. In *2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 76–77. <https://doi.org/10.1109/ISMAR-Adjunct51615.2020.00035>
- [48] Marek Kowalski, Zbigniew Nasarzewski, Grzegorz Galinski, and Piotr Garbat. 2018. HoloFace: Augmenting Human-to-Human Interactions on HoloLens. *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018* 2018-Janua (2018), 141–149. <https://doi.org/10.1109/WACV.2018.00022> arXiv:1802.00278
- [49] Ernst Kruijff, J. Edward Swan, and Steven Feiner. 2010. Perceptual issues in augmented reality revisited. In *IEEE International Symposium on Mixed and Augmented Reality*. 3–12. <https://doi.org/10.1109/ISMAR.2010.5643530>
- [50] Shoufeng Lan, Xueyue Zhang, Mohammad Taghinejad, Sean Rodrigues, Kyu-Tae Lee, Zhaocheng Liu, and Wenshan Cai. 2019. Metasurfaces for Near-Eye Augmented Reality. *ACS Photonics* 6, 4 (04 2019), 864–870. <https://doi.org/10.1021/acsp Photonics.9b00180>
- [51] Tobias Langlotz, Matthew Cook, and Holger Regenbrecht. 2016. Real-Time Radiometric Compensation for Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 22, 11 (2016), 2385–2394. <https://doi.org/10.1109/TVCG.2016.2593781>
- [52] Tobias Langlotz, Jonathan Sutton, Stefanie Zollmann, Yuta Itoh, and Holger Regenbrecht. 2018. ChromaGlasses: Computational Glasses for Compensating Colour Blindness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 1–12. <https://doi.org/10.1145/3173574.3173964>
- [53] Gun-Yeal Lee, Jong-Young Hong, SoonHyoung Hwang, Seokil Moon, Hyeokjung Kang, Sohee Jeon, Hwi Kim, Jun-Ho Jeong, and ByoungHo Lee. 2018. Metasurface eyepiece for augmented reality. *Nature Communications* 9, 1 (2018), 4562. <https://doi.org/10.1038/s41467-018-07011-5>
- [54] Zhen Li, Michelle Annett, Ken Hinckley, Karan Singh, and Daniel Wigdor. 2019. HoloDoc: Enabling Mixed Reality Workspaces That Harness Physical and Digital Content. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, 1–14. <https://doi.org/10.1145/3290605.3300917>
- [55] David Lindlbauer, Anna Maria Feit, and Otmar Hilliges. 2019. Context-Aware Online Adaptation of Mixed Reality Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. ACM, 147–160. <https://doi.org/10.1145/3332165.3347945>
- [56] David Lindlbauer, Jörg Mueller, and Marc Alexa. 2017. Changing the Appearance of Real-World Objects By Modifying Their Surroundings. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, 3954–3965. <https://doi.org/10.1145/3025453.3025795>
- [57] David Lindlbauer and Andy D. Wilson. 2018. Remixed Reality: Manipulating Space and Time in Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 1–13. <https://doi.org/10.1145/3173574.3173703>
- [58] María-Jesús Lobo, Emmanuel Pietriga, and Caroline Appert. 2015. An Evaluation of Interactive Map Comparison Techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, 3573–3582. <https://doi.org/10.1145/2702123.2702130>
- [59] Marcio C. de F. Macedo and Antonio L. Apolinario. 2021. Occlusion Handling in Augmented Reality: Past, Present and Future. *IEEE Transactions on Visualization and Computer Graphics* (2021). <https://doi.org/10.1109/TVCG.2021.3117866>
- [60] Andrew Maimone, Douglas Lanman, Kishore Rathinavel, Kurtis Keller, David Luebke, and Henry Fuchs. 2014. Pinlight Displays: Wide Field of View Augmented Reality Eyeglasses Using Defocused Point Light Sources. *ACM Trans. Graph.* 33, 4, Article 89 (jul 2014), 11 pages. <https://doi.org/10.1145/2601097.2601141>

- [61] Steve Mann. 2002. Mediated reality with implementations for everyday life. *Presence Connect* (2002).
- [62] Steve Mann, Tom Furness, Yu Yuan, Jay Iorio, and Zixin Wang. 2018. All Reality: Virtual, Augmented, Mixed (X), Mediated (X, Y), and Multimeditated Reality. *CoRR* abs/1804.08386 (2018). <http://arxiv.org/abs/1804.08386>
- [63] Sebastian Marwecki, Andrew D. Wilson, Eyal Ofek, Mar Gonzalez Franco, and Christian Holz. 2019. Mise-Unseen: Using Eye Tracking to Hide Virtual Reality Scene Changes in Plain Sight. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. ACM, 777–789. <https://doi.org/10.1145/3332165.3347919>
- [64] Keiichi Matsuda. 2016. Hyper-Reality. Short film - <https://vimeo.com/166807261>.
- [65] Eran May-raz and Daniel Lazo. 2012. Sight. Short film - <https://vimeo.com/46304267>.
- [66] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, 2143–2152. <https://doi.org/10.1145/2702123.2702382>
- [67] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329.
- [68] Mark Roman Miller, Hanseul Jun, Fernanda Herrera, Jacob Yu Villa, Greg Welch, and Jeremy N. Bailenson. 2019. Social interaction in augmented reality. *PLOS ONE* 14, 5 (05 2019), 1–26. <https://doi.org/10.1371/journal.pone.0216290>
- [69] Hein Min Htike, Tom H. Margrain, Yu-Kun Lai, and Parisa Eslambolchilar. 2021. Augmented Reality Glasses as an Orientation and Mobility Aid for People with Low Vision: A Feasibility Study of Experiences and Requirements. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, New York, NY, USA, Article 729, 15 pages. <https://doi.org/10.1145/3411764.3445327>
- [70] Shohei Mori, Sei Ikeda, and Hideo Saito. 2017. A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects. *IPSJ Transactions on Computer Vision and Applications* 9, 1 (2017), 17. <https://doi.org/10.1186/s41074-017-0028-1>
- [71] Erwan Normand and Michael J. McGuffin. 2018. Enlarging a Smartphone with AR to Create a Handheld VESAD (Virtually Extended Screen-Aligned Display). In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 123–133. <https://doi.org/10.1109/ISMAR.2018.00043>
- [72] Benjamin Nuernberger, Eyal Ofek, Hrvoje Benko, and Andrew D. Wilson. 2016. SnapToReality: Aligning Augmented Reality to the Real World. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, 1233–1244. <https://doi.org/10.1145/2858036.2858250>
- [73] Babak A. Parviz. 2009. Augmented Reality in a Contact Lens. *IEEE Spectrum* (September 2009).
- [74] Shane Pase. 2012. Ethical considerations in augmented reality applications. In *Proceedings of the International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE)*.
- [75] Jon Peddie. 2017. *Augmented reality: Where we will all live*. Springer.
- [76] Tekla S. Perry. 2020. Augmented Reality in a Contact Lens: It's the Real Deal. *IEEE Spectrum* (January 2020).
- [77] Wayne Piekarski and Bruce Thomas. 2002. ARQuake: The Outdoor Augmented Reality Gaming System. *Commun. ACM* 45, 1 (Jan. 2002), 36–38. <https://doi.org/10.1145/502269.502291>

- [78] Cyprien Pindat, Emmanuel Pietriga, Olivier Chapuis, and Claude Puech. 2012. JellyLens: Content-Aware Adaptive Lenses. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, 261–270. <https://doi.org/10.1145/2380116.2380150>
- [79] Cyprien Pindat, Emmanuel Pietriga, Oliver Chapuis, and Claude Puech. 2013. Drilling into Complex 3D Models with Gimlenses. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology (VRST '13)*. ACM, 223–230. <https://doi.org/10.1145/2503713.2503714>
- [80] Henning Pohl and Roderick Murray-Smith. 2013. Focused and Casual Interactions: Allowing Users to Vary Their Level of Engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, 2223–2232. <https://doi.org/10.1145/2470654.2481307>
- [81] Iulian Radu. 2012. Why should my students use AR? A comparative review of the educational impacts of augmented-reality. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 313–314. <https://doi.org/10.1109/ISMAR.2012.6402590>
- [82] Sharif Razzaque, Zachariah Kohn, and Mary C. Whitton. 2001. Redirected Walking. In *Eurographics 2001 - Short Presentations*. Eurographics Association. <https://doi.org/10.2312/egs.20011036>
- [83] Sebastian Ritterbusch, Staffan Ronnäs, Irina Waltschläger, Philipp Gerstner, and Vincent Heuveline. 2013. Augmented reality visualization of numerical simulations in urban environments. *International Journal of Advances in Systems and Measurements* 6, 1 (2013), 26–39.
- [84] Jan Ole Rixen, Teresa Hirzle, Mark Colley, Yannick Etzel, Enrico Rukzio, and Jan Gugenheimer. 2021. Exploring Augmented Visual Alterations in Interpersonal Communication. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, Article 730, 11 pages. <https://doi.org/10.1145/3411764.3445597>
- [85] Franziska Roesner, Tadayoshi Kohno, and David Molnar. 2014. Security and Privacy for Augmented Reality Systems. *Commun. ACM* 57, 4 (April 2014), 88–96. <https://doi.org/10.1145/2580723.2580730>
- [86] Yvonne Rogers. 2006. Moving on from Weiser's Vision of Calm Computing: Engaging Ubicomp Experiences. In *Proceedings of the 8th International Conference on Ubiquitous Computing (UbiComp'06)*. Springer-Verlag, 404–421. https://doi.org/10.1007/11853565_24
- [87] Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. 2007. Measuring visual clutter. *Journal of Vision* 7, 2 (08 2007), 17–17. <https://doi.org/10.1167/7.2.17>
- [88] Marc Satkowski and Raimund Dachsel. 2021. Investigating the Impact of Real-World Environments on the Perception of 2D Visualizations in Augmented Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Article 522, 15 pages. <https://doi.org/10.1145/3411764.3445330>
- [89] Paul Schlosser, Ben Matthews, Isaac Salisbury, Penelope Sanderson, and Sass Hayes. 2021. Head-Worn Displays for Emergency Medical Services Staff: Properties of Prehospital Work, Use Cases, and Design Considerations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, Article 40, 14 pages. <https://doi.org/10.1145/3411764.3445614>
- [90] Mel Slater, Cristina Gonzalez-Liencre, Patrick Haggard, Charlotte Vinkers, Rebecca Gregory-Clarke, Steve Jelley, Zillah Watson, Graham Breen, Raz Schwarz, William Steptoe, Dalila Szostak, Shivashankar Halan, Deborah Fox, and Jeremy Silver. 2020. The Ethics of Realism in Virtual and Augmented Reality. *Frontiers in Virtual Reality* 1 (2020), 1. <https://doi.org/10.3389/frvir.2020.00001>

- [91] Jonathan Sutton, Tobias Langlotz, and Alexander Plopski. 2022. Seeing Colours: Addressing Colour Vision Deficiency with Vision Augmentations Using Computational Glasses. *ACM Trans. Comput.-Hum. Interact.* 29, 3, Article 26 (jan 2022), 53 pages. <https://doi.org/10.1145/3486899>
- [92] Desney S. Tan, Brian Meyers, and Mary Czerwinski. 2004. WinCuts: Manipulating Arbitrary Window Regions for More Effective Use of Screen Space. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*. ACM, 1525–1528. <https://doi.org/10.1145/985921.986106>
- [93] Kohei Tanaka, Yasue Kishino, Masakazu Miyamae, Tsutomu Terada, and Shojiro Nishio. 2008. An information layout method for an optical see-through head mounted display focusing on the viewability. In *7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. 139–142. <https://doi.org/10.1109/ISMAR.2008.4637340>
- [94] Franco Tecchia, Giovanni Avveduto, Raffaello Brondi, Marcello Carrozzino, Massimo Bergamasco, and Leila Alem. 2014. I'm in VR! Using Your Own Hands in a Fully Immersive MR System. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology (VRST '14)*. ACM, 73–76. <https://doi.org/10.1145/2671015.2671123>
- [95] Dorin Ungureanu, Federica Bogo, Silvano Galliani, Pooja Sama, Xin Duan, Casey Meekhof, Jan Stühmer, Thomas J. Cashman, Bugra Tekin, Johannes L. Schönberger, Bugra Tekin, Pawel Olszta, and Marc Pollefeys. 2020. HoloLens 2 Research Mode as a Tool for Computer Vision Research. *arXiv:2008.11239* (2020).
- [96] Manuela Waldner, Markus Steinberger, Raphael Grasset, and Dieter Schmalstieg. 2011. Importance-Driven Compositing Window Management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, 959–968. <https://doi.org/10.1145/1978942.1979085>
- [97] Zeyu Wang, Cuong Nguyen, Paul Asente, and Julie Dorsey. 2021. DistanciAR: Authoring Site-Specific Augmented Reality Experiences for Remote Environments. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, Article 411, 12 pages. <https://doi.org/10.1145/3411764.3445552>
- [98] Brian Wassom. 2014. *Augmented reality law, privacy, and ethics: Law, society, and emerging AR technologies*. Syngress.
- [99] Sean White and Steven Feiner. 2009. SiteLens: Situated Visualization Techniques for Urban Site Visits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, 1117–1120. <https://doi.org/10.1145/1518701.1518871>
- [100] Wesley Willett, Yvonne Jansen, and Pierre Dragicevic. 2017. Embedded Data Representations. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 461–470. <https://doi.org/10.1109/TVCG.2016.2598608>
- [101] Keng-Ta Yang, Chiu-Hsuan Wang, and Liwei Chan. 2018. ShareSpace: Facilitating Shared Use of the Physical Space by Both VR Head-Mounted Display and External Users. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. ACM, 499–509. <https://doi.org/10.1145/3242587.3242630>
- [102] Ya-Ting Yue, Yong-Liang Yang, Gang Ren, and Wenping Wang. 2017. SceneCtrl: Mixed Reality Enhancement via Efficient Scene Editing. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, 427–436. <https://doi.org/10.1145/3126594.3126601>
- [103] Yuhang Zhao, Elizabeth Kupferstein, Brenda Veronica Castro, Steven Feiner, and Shiri Azenkot. 2019. Designing AR Visualizations to Facilitate Stair Navigation for People with Low Vision. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. ACM, 387–402. <https://doi.org/10.1145/3332165.3347906>

-
- [104] Fengyuan Zhu and Tovi Grossman. 2020. BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, 1–14. <https://doi.org/10.1145/3313831.3376233>
- [105] Stefanie Zollmann, Tobias Langlotz, Raphael Grasset, Wei Hong Lo, Shohei Mori, and Holger Regenbrecht. 2021. Visualization Techniques in Augmented Reality: A Taxonomy, Methods and Patterns. *IEEE Transactions on Visualization and Computer Graphics* 27, 9 (2021), 3808–3825. <https://doi.org/10.1109/TVCG.2020.2986247>



**RESEARCH CENTRE
SACLAY – ÎLE-DE-FRANCE**

1 rue Honoré d'Estienne d'Orves
Bâtiment Alan Turing
91120 Palaiseau

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399