Classifying handheld Augmented Reality: Three categories linked by spatial mappings

Thomas Vincent, Laurence Nigay, Takeshi Kurata

To cite this version:

Thomas Vincent, Laurence Nigay, Takeshi Kurata. Classifying handheld Augmented Reality: Three categories linked by spatial mappings. Workshop on Classifying the AR Presentation Space at ISMAR 2012, Nov 2012, Atlanta, GA, United States. hal-00757883

HAL Id: hal-00757883
https://hal.science/hal-00757883

Submitted on 27 Nov 2012

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.

Classifying handheld Augmented Reality: Three categories linked by spatial mappings

Thomas Vincent*
EHCI, LIG, UJF-Grenoble 1
France

Laurence Nigay†
EHCI, LIG, UJF-Grenoble 1
France

Takeshi Kurata‡
Center for Service Research, AIST
Japan

ABSTRACT

Handheld Augmented Reality (AR) relies on a spatial coupling of the on-screen content with the physical surrounding. To help the design of such systems and to classify existing AR systems, we present a framework made of three categories and two spatial relationships. Our framework highlights spatial relationships between the physical world, the representation of the physical world on screen and the augmentation on screen. Within this framework, we study the relaxing of the spatial coupling between the digital information and the physical surrounding in order to enhance interaction by breaking the constraints of physical world interaction.

Keywords: Handheld Augmented Reality, Framework, Implicit/Explicit interaction

Index Terms: H.5.2 [Information interfaces and presentation]: User Interfaces—Graphical user interfaces; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

1 INTRODUCTION

As compared to other aspects of Human-Computer Interaction (HCI), Augmented Reality (AR) constitutes a spatiotemporal relationship between the physical world and digital content. Indeed, Azuma [2] defined AR systems as systems that (i) combine real and virtual, (ii) are interactive in real time and (iii) are registered in 3D. Moreover Rekimoto et al. [22] compared HCI styles (namely Graphical User Interface, Virtual Reality, Ubiquitous Computing and Augmented Interaction) in terms of interactions between Human, Computer and the Real World: The "Augmented Interaction" style designates interaction between the three categories and supports interaction with the real world through computer augmented information.

Among the different display devices supporting AR, handheld devices used as magic lenses are becoming a popular platform and paradigm for AR applications. As defined in [23]: "The term magic lens is used here to denote augmented reality interfaces that consist of a camera-equipped mobile device being used as a see-through tool. It augments the user's view of real world objects by graphical and textual overlays". One seminal system of such handheld AR systems is the palmtop NaviCam [22] for which its authors coined the term "magnifying glass metaphor" to denote the real world enhancement with information. While offering the opportunity for AR to reach a wide audience, handheld devices also bring specific constraints [24]: the screen real estate is limited and direct touch on the screen, the de facto standard input modality on such devices is impaired by finger occlusion and an ambiguous selection point (i.e.: "fat-finger" problem).

*e-mail: thomas.vincent@imag.fr
†e-mail: laurence.nigay@imag.fr
‡e-mail: t.kurata@aist.go.jp

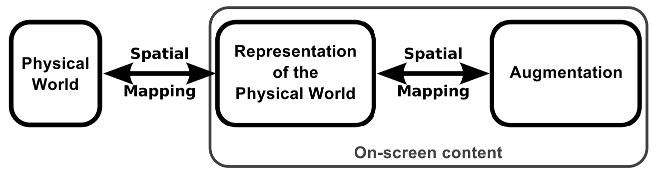

Figure 1: Handheld AR on-screen content depicted with three categories: (1) physical world, (2) video representation of the physical world and (3) digital augmentation and two spatial mappings.

Furthermore, with handheld AR, both the video representing the physical surrounding and the digital augmentation are displayed simultaneously on the screen. As a consequence the race for screen real estate is even more critical. In addition, the tight spatiotemporal coupling of the on-screen content with the physical surrounding makes touch interaction harder. Indeed the viewpoint is controlled by the device pose and its stability is impaired by hand tremor as well as motion induced by user touch input. As a result, on-screen interactive content is not stable within the touch input space. Thus, considering spatiotemporal couplings in handheld AR systems is crucial to improve on-screen content for both information presentation (i.e., outputs) and user interaction (i.e., inputs).

The design challenge lies in the fluid and harmonious fusion of the physical and digital worlds while breaking the constraints of physical world interaction. To help the design of such handheld AR systems (and therefore no longer design and develop handheld AR systems in an ad-hoc way), we present a framework made of three categories and two spatial relationships. Our framework is useful for analysis and comparison of existing handheld AR systems as well as for design (descriptive, evaluative and generative power of an interaction model [4]). Indeed, in addition to classifying existing AR handheld systems, the underlying concepts of our framework allow generation of ideas and choice of design alternatives. In the paper we mainly focus on the descriptive and taxonomic power of our framework and give one example to illustrate its generative power.

The paper is organized as follows: We first describe the three categories of our framework and their spatial relationships. We then study the transitions between different levels of spatial couplings of the described categories. We finally expose several research axes for extending our framework.

2 FRAMEWORK: THREE CATEGORIES

Our framework articulates axes serving to distinguish between the characteristics of handheld AR applications. It is based on three main categories (or worlds), as shown in Figure 1:

1. Physical world,
2. Representation of the physical world and
3. Digital augmentation.

In this scheme, on-screen visual content of handheld AR interfaces can be characterized by the *representation of the physical world*, and the *digital augmentation*. As discussed later, while we focus on handheld AR, this framework can be relevant to a wider scope.

2.1 Representation of the Physical World

The *representation of the physical world* encompasses the displayed components that represent the physical surrounding. Such a representation allows the user to map its viewpoint and to overlay augmentation in the physical surrounding.

In handheld AR applications, this representation is commonly the live video from the rear-facing camera of the handheld device. However other modes of representation can serve the same purpose. For example, the live video can be transformed to a non-photorealistic representation of the physical world in order to have the same visual quality of representation as that of augmentation [9]. Also, a virtualized model of the physical world can be used to represent it [17]. The mode of representation can also be changed in order to support viewpoints otherwise impossible with live video or to change the style of interaction. To overcome limited cameras field of view, Alessandro et al. [1] describe animated zooming out techniques which terminate with an egocentric panoramic view of 360 degrees or with an exocentric map-like top-down view on handheld devices. With the Magic Book [6], Billinghurst et al. propose to interactively move from AR view to immersive Virtual Reality by pressing a button.

2.2 Digital Augmentation

The *augmentation* is the representation of the digital content that is not the representation of the physical world. Such content augments the physical world with extra information and interaction capabilities.

The visual quality is an important feature of the augmentation. Milgram et al. [19] grossly describe the rendering quality of the virtual content with their *Reproduction Fidelity* axis. In [16], Kruijff et al. identified different perceptual issues affecting augmentation.

The selection of the displayed digital information is also of particular importance as it can mitigate information overload and clutter and allow for a better fit to the current user's task. Julier et al. report on different approaches to filter overlaid information [15], namely physically-based methods using distance and visibility; methods based on the spatial model of interaction and rule-based filtering.

2.3 Distinguishing the Representation of the Physical World from the Digital Augmentation

A general issue of distinguishing the representation of the physical world from the augmentation is that the boundary is not always obvious and for some cases tends to be blurred. Milgram et al. [19] introduce the following definitions for clarifying the distinction between real and virtual:

- Real objects are any objects that have an actual objective existence.
- Virtual objects are objects that exist in essence or effect, but not formally or actually.

As such, real objects can be directly perceived or sampled and resynthesized, while virtual objects must be simulated. Applying this distinction is straightforward in the cases of 3D models overlaid on fiducial markers or of annotations overlaid on physical objects.

It becomes less obvious for the cases where the representation of the physical world is directly transformed. For instance ClayVision [27] aims at morphing existing buildings, changing their size or aspect. Here, such an altered building belongs both to the representation of the physical surroundings and to the augmentation. On the one hand, some characteristics like the overall appearance and texture allow the user to map the altered building to its location in the physical world and then support the *representation of the physical world*. On the other hand, some characteristics like its modified size or its highlighted color provide extra information and are thus considered as the *augmentation*. The distinction is here made on a per-characteristic rather than a per-object basis.

Another example of direct transformation of the live video feed is subtle video change magnification [29]. Such technique allows for example to render visible otherwise unnoticeable face color changes induced by blood flow. Here again a per-characteristic distinction is possible. The color of the skin can be considered as the *augmentation* as it provides extra information while the shape and appearance of the face can be considered as the *representation of the physical world* since it helps to map the augmented content into the physical world.

3 FRAMEWORK: TWO SPATIAL MAPPINGS

The three presented categories are coupled by spatial mappings. We identify two spatial mappings in our framework.

3.1 Spatial mapping between the physical world and its representation

This spatial mapping describes the coupling of the viewpoint of the representation with the handheld device pose in the physical world. Such a coupling can be relaxed along the axis of Figure 2 that extends from *conformal mapping* where the viewpoint is controlled by the handheld device pose in an absolute manner to *no mapping* where there is no relation between the device pose and the viewpoint. This spatial mapping can also be *relaxed* when the viewpoint is only partially controlled by the device pose [11, 13].

Figure 2: Different spatial mappings between the *physical world* and its on-screen *representation*.

A second aspect of this spatial mapping is the characterization of the projection performed to represent the physical world on screen. When using camera images to represent the physical world, this projection is characterized by the camera parameters. However other projection like an orthographic projection can possibly be used in the case of a 3D model representing the physical world. Also further transformation like dynamic zoom or fish-eye view can be applied.

For example, the physical magic lens [22] approach has a conformal spatial mapping and a fixed projection (the one of the camera). Güven et al. [10] propose handheld AR interaction techniques relying on freezing of the frame in order to edit it. Similarly, Lee et al. proposed and evaluated the Freeze-Set-Go [18] technique, which lets the user freeze the video and continue to manipulate virtual content within the scene. Such video freeze techniques break the spatial mapping in order to improve user interaction.

TouchProjector [7] enables users to move pictures displayed on remote screens through direct touch on the live video image of their handheld device. To improve interaction, TouchProjector includes video freeze as well as zooming capabilities. Zooming allows a more detailed view and a dynamic ratio between the size of a real object and its on-screen representation, but it also increases instability of the camera image.

3.2 Spatial mapping between the augmentation and the representation of the physical world

It describes the spatial coupling of the augmentation with the representation of the physical world. This axis, presented in Figure 3, goes from *conformal mapping* where augmentation is exactly mapped onto the physical object representation to *no mapping* where augmentation has no spatial relationship with the representation of the physical world. In between, there are different forms of *relaxed* spatial mappings.

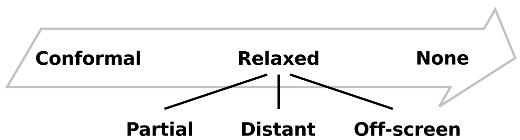

Figure 3: Different spatial mappings between the *representation of the physical world* and the *augmentation*.

Partial mapping corresponds to the case where some degrees of freedom between the augmentation and the representation of the corresponding physical object are exactly matched while others are relaxed. This is the case for example for annotations displayed with billboarding and/or a fixed on-screen size. *Distant mapping* depicts augmentations like annotations that are displayed at a distance from the physical objects they refer to, but are visually linked to them with lines for example. *Off-screen mapping* includes visualization techniques of off-screen points of interests such as Arrows [25].

Using a relaxed spatial mapping is useful to improve text readability and to avoid clutter. The main advantage is to allow extra degrees of freedom for on-screen layout but this might reduce the feeling of collocation with the physical world.

4 Characterizing the dynamicity of the spatial mappings

The different values of the previously presented spatial mappings describe different levels of coupling between the on-screen content and the physical surrounding.

On the one hand, these values define a static snapshot at a given time of the level of coupling supported by a handheld AR application. On the other hand, studying the transitions along the two spatial mapping axes is essential in order to support improved interaction (e.g., for pointing accuracy) but also to allow "magic" like transitions to other modes of representation [6, 1] or movement within the augmented space while not moving in the physical world [26]. Indeed interaction with AR settings is constrained by the spatial relationship with the physical world. Yet this is not the physical world the users are interacting with, so such constraints can be relaxed, at least temporarily. We expressed these as transitions along the two axes in our framework. We characterize such transitions in the light of previous studies on mode switching in terms of:

- Initiative: extending from *explicit* user's interaction to *automatic* (system's initiative), through *implicit* interaction (system's initiative upon indirect interpretation of user's action); and
- Sustainability: extending from *transient* to *sustained-mode*.

Classical interaction modes (e.g., drawing mode of a graphical editor) are explicit and sustained, while quasi-modes (e.g., holding the Shift key for upper case typing) are explicit and transient. Proxemic interaction [3, 14], which is based on spatiotemporal relationships between users and devices in order to adapt on-screen content and interaction capabilities is characterized as implicit and transient.

Applied to transitions between spatial mappings, we observe that:

- Modifications of the *spatial mapping between the physical world and its representation* have been mostly explicit and sustained: Indeed the video freeze technique [10, 18, 7] has been implemented in numerous systems as an explicit transition (from conformal to none) triggered by a button between two sustained modes. In contrast to this explicit transition, TouchProjector [7] is a special case since the system includes an automatic zoom, in order to maintain a fixed control-to-display ratio between the touch-screen and the controlled remote screen. This system zooms in when a remote screen is recognized and zooms out when there is no remote screen in the live video. This is one example of implicit and transient transition in order to enhance the interaction.
- Modifications of the *spatial mapping between the representation of the physical world and the augmentation* are mostly implicit and automatic: Indeed view management [5] enhances the mapping between the augmentation and the representation of the physical world by taking into account visual constraints of the projection of objects onto the view plane. Such techniques avoid label clutter and can prevent augmented content from occluding interesting areas of the live video image. To do so, augmentation is automatically laid out according to both augmented objects position in 3D and on-screen footprint. Annotations mapping with objects is dynamically and automatically adapted from partial mapping (billboarding) to a distant one (linked with a line).

In AR settings, implicit and temporary relaxing of spatiotemporal constraints are of particular interest. Temporary transitions allow for a best fit of the visual content to the current user's focus and task. Moreover implicit transitions do not require extra user's action to benefit from such transitions.

At the same time, such temporary relaxing of spatiotemporal constraints in order to improve the interaction in AR settings comes with some drawbacks that need to be further studied. Indeed after a constraint has been relaxed for a specific purpose (e.g., freezing the video to support stable interaction), it should be restored when it is no longer necessary. Breaking and restoring constraints can disorient users as observed in [18]. An animation from frozen view to live video as used in [6, 1] and suggested in [18] can minimize such a spatial discontinuity problem.

5 Framework: Its generative power

While describing the three categories and the two spatial relationships of our framework, we showed how existing handheld AR systems are described within our framework. It enables us to highlight the descriptive and taxonomic powers of our framework. We now illustrate its generative power by considering the design of an AR system that we developed: AR TapTap.

Based on our framework, our design goal was to explore techniques for explicitly relaxing the spatial mapping between the physical world and its representation. But as opposed to existing handheld AR techniques that implement explicit transitions between two sustained modes, we implemented a transient mode.

AR TapTap uses video freeze and zoom to ease the placement of digital marks on a physical map (Figure 4). It builds on TapTap [24], a target acquisition technique dedicated to one-handed touch interaction on handheld devices. With AR TapTap, placing a mark on the physical map is performed with two taps on the touch-screen. The first tap selects an area of the live video that is displayed frozen and zoomed at the center of the screen. The second tap places the mark in this frame, thus improving pointing accuracy. In comparison with the original TapTap application, AR TapTap adds video freeze at no extra user's action.

Figure 4: AR TapTap: First tap (left) to freeze and zoom the video (center); Second tap to place the mark (center) and automatically close the frozen and zoomed view (right).

Inherited from TapTap, the interaction is very fast, making it practically like a transient (or temporary) transition. The first selection tap provokes a transition along the axis *Spatial mapping with the physical world* (from conformal to none in Figure 2). The second tap for placing a mark also terminates the frozen and zoomed view returning thereafter to the initial state along the axis *Spatial mapping with the physical world* (i.e., conformal mapping - live video playback). In order to allow accurate placement of marks, AR TapTap therefore implements an explicit modification of the *spatial mapping between the physical world and its representation* with a first selection tap. This modification is transient since the second selection tap is not dedicated to changing the current mode (from none to conformal in Figure 2) but rather to placing a mark. As such, with AR TapTap, the frozen mode is only maintained for one mark placement. On the one hand, by placing the mark, the user also modifies the *spatial mapping between the physical world and its representation*: It is therefore a transient mode since no extra action from the user is required to explicitly change the mode. On the other hand, an additional third tap in order to change the mode after placing the mark would be a case of explicit transition between two sustained modes as in [10, 18, 7].

With AR TapTap the frozen view is not displayed full screen, so the live video is still visible on the edges of the screen. This is an example of on-screen multiplexing of two views with different spatial mappings with the physical world. By minimizing the spatial discontinuity, we expected such multiplexing to help users to map the viewpoint of the camera when the frozen view was closed. However informal tests were inconclusive and this was not further evaluated.

6 Conclusion and Research Agenda

This paper introduces a framework for studying handheld AR on-screen content that emphasizes spatial relationships between the *physical world*, the *representation of the physical world* on screen and the *augmentation* on screen. Along the axes of our framework we highlighted transitions for relaxing the tight coupling between the on-screen content and the physical surroundings. Such transitions are studied in the light of previous studies on mode switching in Graphical User Interface (implicit/explicit transition and transient/sustained mode). While we focused on spatial mappings and their dynamicity in the scope of handheld AR, this work can be further continued and extended in the following directions.

6.1 Validation

The framework has been used to describe and compare existing handheld AR systems. It enables us to describe in detail the systems according to the three categories and two spatial relationships and to make a fine distinction between them. To further validate the framework, we need to consider more existing handheld systems, in particular to check that no existing handheld AR systems are left out by our framework.

6.2 Input modalities

While our framework describes the visual output modality on screen, we need to extend it in order to include the input modalities and thus the input spaces. This will allow further depicting of how users control the viewpoint in the augmented scene. For instance, with handheld AR applications, the viewpoint is classically controlled by the device pose, but it can also be partially controlled by head-coupled display [11] or touch input [13]. Moreover focusing on different input modalities will enable us to focus on the spatial relationships between the input spaces and the three categories that form our framework. This should help to clearly depict the strengths and weaknesses of different input modalities.

6.3 Generalization to other AR display devices

The framework is dedicated to handheld AR on-screen content. Its categories and its axes can be nevertheless relevant for other AR settings. Indeed, different display devices used for AR can be compared in the light of the 3 categories of our framework as presented in Table 1.

Display device	Physical World	Representation Physical World	Augmentation
HMD			
- Video		✓	✓
- Video Miniat.	✓	✓	✓
- Optical	✓		✓
Projection-based	✓		✓
Handheld device	✓	✓	✓

Table 1: AR display comparison

- With video see-through Head-Mounted Displays (HMDs), a representation of the physical world exists: the live video sampled by the cameras. However, as users cannot directly observe the physical world, modification of its representation is limited as it impacts user's actions in the physical world. For example, freezing the frame might prevent the user from operating safely in the physical world. This limitation does not hold for miniaturized HMDs allowing direct observation of the physical world.
- With optical see-through HMDs, there is no representation of the physical world as it is observed directly. Also, users cannot observe the physical world un-augmented.
- With projection-based systems, there is also no representation of the physical world and the physical world cannot be observed simultaneously augmented and un-augmented.
- Handheld devices allow both direct observation of the physical world un-augmented and observation of the augmented scene on the screen. It thus allows for more design possibilities for modifying the representation of the physical world. Such differences encouraged us to first focus more specifically on handheld AR.

6.4 Positioning with respect to other existing classification schemes

In [12], Hugues et al. briefly review existing taxonomies of interactive mixed systems. They categorize such taxonomies as being either technical, functional or conceptual. In [21], Normand et al. organize AR taxonomies with four categories: technique-centered, user-centered, information-centered and interaction-centered. In this scope, the work presented in this paper is a conceptual framework. The description of the spatial mappings is information-centered while the description of the transitions in this framework is interaction-centered. In the following we present some relation between our work and some existing classifications.

On the one hand, our classification is dedicated to on-screen content for the case of handheld AR. As a consequence and in comparison with other taxonomies of AR applications, the scope of our framework is therefore more focused. For instance our previous classification space of mixed reality systems [8] is general. It identifies two types of augmentation: *augmented execution* and/or *augmented evaluation* applied to Augmented Reality (where the target of the task belongs to the physical world) and Augmented Virtuality (where the target of the task belongs to the digital world). Within this framework for mixed reality systems, the classification of this paper details the case of *augmented evaluation* in the context of Augmented Reality.

Augmented evaluation is also called *augmented perception* in the AR taxonomy presented in [12]. In this taxonomy, the authors divide augmented perception into five sub-functionalities, namely (1) *Documented reality and virtuality*, (2) *Reality with augmented perception or understanding*, (3) *Perceptual association of the real and the virtual*, (4) *Behavioral association of the real and the virtual* and (5) *Substituting the real by the virtual*. In our framework, such functionalities describe the different relationships that the information of the *augmentation* maintains with the *physical world* or with the *representation of the physical world*. Such functionalities have a direct impact on the type of *spatial mappings between the augmentation and the representation of the physical world* (Figure 3). For instance the functionality *perceptual association of the real and the virtual* implies a conformal mapping while the functionality *reality with augmented perception or understanding* implies a relaxed or a conformal spatial mapping according to the considered levels (The first level - *Reality with augmented understanding* corresponds to relaxed mapping; The second level - *Reality with augmented visibility* corresponds to conformal mapping).

On the other hand, the previous section 6.3 shows that we can generalize the categories and axes of our framework and therefore extend the scope of our framework to other AR settings. By considering our three categories we are able to classify AR displays in Table 1. In comparison with the axis *Augmentation type* of the recent taxonomy presented in [21] that distinguishes *Mediated augmentation* from *Direct augmentation*, our framework makes a clear distinction between mediated and direct augmentation by considering the presence or not of the *representation of the physical world*. Our framework also distinguishes optical see-through devices from video see-through devices by considering, or not considering the direct perception of the *physical world*. Furthermore our framework enables us to consider optical see-through AR settings and projected-based AR settings in the same category, while they belong to two distinct classes in [21]. Optical see-through AR settings such as navigation systems based on head-up displays in cars or the SixthSense projected-based system [20] share the design issue of the *spatial relationships between the augmentation and the physical world*.

Tönnis et al. propose six classes to classify the AR presentation space [28]. The *Registration* class is related to the *Spatial relation between the representation of the physical world and the augmentation*, and the *Frame of Reference* class is related to the *Spatial relation between the physical world and its representation*. Two other classes, *Referencing* and *Mounting* are also at least partially related to spatial relations and positions. This highlights the importance of spatial relations in AR classification. The two remaining classes are related to the augmentation. The *Dimensionality* is related to the augmentation's visual aspect. The *Temporality* as well as the already mentioned *Referencing* are related to the selection of the displayed content. As focusing on AR presentation, those classes does not cover transitions and interaction.

ACKNOWLEDGEMENTS

This work has been supported by the ANR/JST AMIE project (http://amie.imag.fr/). We thank Matthieu Riegler for his implementation of AR TapTap.

REFERENCES

[1] M. Alessandro, A. Dünser, and D. Schmalstieg. Zooming interfaces for augmented reality browsers. In *Proceedings of the 12th international conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, pages 161–170. ACM, 2010.

[2] R. T. Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, August 1997.

[3] T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: designing for a proximity and orientation-aware environment. In *Proceedings of the 2010 international conference on Interactive Tabletops and Surfaces*, ITS '10, pages 121–130. ACM, 2010.

[4] M. Beaudouin-Lafon. Designing interaction, not interfaces. In *Proceedings of the 2004 working conference on Advanced Visual Interfaces*, AVI '04, pages 15–22. ACM, 2004.

[5] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. In *Proceedings of the 14th symposium on User Interface Software and Technology*, UIST '01, pages 101–110. ACM, 2001.

[6] M. Billinghurst, H. Kato, and I. Poupyrev. The magicbook: A transitional ar interface. *Computers and Graphics*, 25(5):745–753, 2001.

[7] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch projector: Mobile interaction through video. In *Proceedings of the 28th international conference on Human Factors in Computing Systems*, CHI '10, pages 2287–2296. ACM, 2010.

[8] E. Dubois, L. Nigay, J. Troccaz, O. Chavanon, and L. Carrat. Classification space for augmented surgery, an augmented reality case study. In *Proceedings of the 7th IFIP Conference on Human-Computer Interaction*, INTERACT '99, pages 353–359. IOS Press, 1999.

[9] J. Fischer, D. Bartz, and W. Straßer. Stylized augmented reality for improved immersion. In *Proceedings of the IEEE Conference on Virtual Reality 2005*, VR '05, pages 195–202, 325. IEEE Computer Society, 2005.

[10] S. Guven, S. Feiner, and O. Oda. Mobile augmented reality interaction techniques for authoring situated media on-site. In *Proceedings of the 5th International Symposium on Mixed and Augmented Reality*, ISMAR '06, pages 235–236. IEEE Computer Society, 2006.

[11] A. Hill, J. Schiefer, J. Wilson, B. Davidson, M. Gandy, and B. MacIntyre. Virtual transparency: Introducing parallax view into video see-through ar. In *Proceedings of the 10th International Symposium on Mixed and Augmented Reality*, ISMAR '11, pages 239–240. IEEE Computer Society, 2011.

[12] O. Hugues, P. Fuchs, and O. Nannipieri. *New Augmented Reality Taxonomy: Technologies and Features of Augmented Environment*, chapter 1, pages 47–63. Springer, August 2011.

[13] S. Hwang, H. Jo, and J. hee Ryu. Exmar: Expanded view of mobile augmented reality. In *Proceedings of the 9th International Symposium on Mixed and Augmented Reality*, ISMAR '10, pages 235–236. IEEE Computer Society, 2010.

[14] W. Ju, B. A. Lee, and S. R. Klemmer. Range: exploring implicit interaction through electronic whiteboard design. In *Proceedings of the 2008 conference on Computer Supported Cooperative Work*, CSCW '08, pages 17–26. ACM, 2008.

[15] S. Julier, Y. Baillot, D. G. Brown, and M. Lanzagorta. Information filtering for mobile augmented reality. *IEEE Computer Graphics and Applications*, 22(5):12–15, September 2002.

[16] E. Kruijff, J. E. Swan II, and S. Feiner. Perceptual issues in augmented reality revisited. In *Proceedings of the 9th International Symposium on Mixed and Augmented Reality*, ISMAR '10, pages 3–12. IEEE Computer Society, 2010.

[17] T. Kurata, M. Kourogi, T. Ishikawa, J. Hyun, and A. Park. Service cooperation and co-creative intelligence cycles based on mixed-reality technology. In *Proceedings of the 8th International Conference on Industrial Informatics*, INDIN '10, pages 967–972. IEEE, 2010.

[18] G. A. Lee, U. Yang, Y. Kim, D. Jo, K.-H. Kim, J. H. Kim, and J. S. Choi. Freeze-set-go interaction method for handheld mobile aug-

mented reality environments. In *Proceedings of the 16th symposium on Virtual Reality Software and Technology*, VRST '09, pages 143–146. ACM, 2009.

[19] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems*, E77-D(12), 1994.

[20] P. Misry and P. Maes. Sixthsense: a wearable gestural interface. In *SIGGRAPH ASIA 2009 Sketches*, SIGGRAPH ASIA '09, pages 11:1–11:1. ACM, 2009.

[21] J.-M. Normand, M. Servières, and G. Moreau. A new typology of augmented reality applications. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, pages 18:1–18:8. ACM, 2012.

[22] J. Rekimoto and K. Nagao. The world through the computer: Computer augmented interaction with real world environments. In *Proceedings of the 8th symposium on User Interface and Software Technology*, UIST '95, pages 29–36. ACM, 1995.

[23] M. Rohs and A. Oulasvitra. Target acquisition with camera phones when used as magic lens. In *Proceedings of the 26th international conference on Human Factors in Computing Systems*, CHI '08, pages 1409–1418. ACM, 2008.

[24] A. Roudaut, S. Huot, and E. Lecolinet. Taptap and magstick: Improving one-handed target acquisition on small touch-screens. In *Proceedings of the 2008 working conference on Advanced Visual Interfaces*, AVI '08, pages 146–153. ACM, 2008.

[25] T. Schinke, N. Henze, and S. Boll. Visualization of off-screen objects in mobile augmented reality. In *Proceedings of the 12th international conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, pages 313–316. ACM, 2010.

[26] M. Sukan and S. Feiner. Using augmented snapshots for viewpoint switching and manipulation in augmented reality. In *Proceedings of the 2012 conference on Human Factors in Computing Systems Extended Abstracts*, CHI EA '12, pages 1095–1098. ACM, 2012.

[27] Y. Takeuchi and K. Perlin. Clayvision: The (elastic) image of the city. In *Proceedings of the 2012 conference on Human Factors in Computing Systems*, CHI '12, pages 2411–2420. ACM, 2012.

[28] M. Tönnis and D. A. Plecher. Presentation principles in augmented reality - classification and categorization guidelines. Technical Report TUM-INFO-06-I11-0/1.-FI, Institut für Informatik der Technischen Universität München, 2011.

[29] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. *Transactions on Graphics - SIGGRAPH 2012 Conference Proceedings*, 31(4):65:1–65:8, 2012.