Perceptual Issues in Augmented Reality Revisited

Ernst Kruijff[1] J. Edward Swan II[2] Steven Feiner[3]

[1]Institute for Computer
Graphics and Vision
Graz University of Technology

[2]Department of Computer Science
and Engineering
Mississippi State University

[3]Department of Computer Science
Columbia University

ABSTRACT

This paper provides a classification of perceptual issues in augmented reality, created with a visual processing and interpretation pipeline in mind. We organize issues into ones related to the environment, capturing, augmentation, display, and individual user differences. We also illuminate issues associated with more recent platforms such as handhelds or projector-camera systems. Throughout, we describe current approaches to addressing these problems, and suggest directions for future research.

CR Categories and Subject Descriptors: H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—*Artificial, augmented, and virtual realities*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Ergonomics, Evaluation/methodology, Screen design*
Additional Keywords: Human perception, augmented reality, handheld devices, mobile computing

1 INTRODUCTION

Over the years, research on head-worn Augmented Reality (AR) has been complemented by work on new platforms such as handheld AR and projector-camera systems. With the rapid advent of applications on cell phones, AR has become almost mainstream. However, researchers and practitioners are still attempting to solve many fundamental problems in the design of effective AR. Although many researchers are tackling registration problems caused by tracking limitations, perceptually correct augmentation remains a crucial challenge. Some of the barriers to perceptually correct augmentation can be traced to issues with depth and illumination that are often interconnected, or by issues related to the appearance of an environment. These problems may cause scene and depth distortions, and visibility issues, which can potentially lead to poor task performance.

Some of these issues result from technological limitations. However, many are caused by limited understanding or by inadequate methods for displaying information. In the mid 90s, Drascic and Milgram attempted to identify and classify these perceptual issues [8]. Focusing on stereoscopic head-worn displays (HWDs), they provided useful insights into some of the perceptual issues in AR. Since then, considerable research has provided new insights into perceptual factors. Even though HWDs are still the predominant platform for perceptual experiments, the emphasis on a broader range of AR platforms has changed the problem space, resulting in the need to address new issues. To meet this need, we have designed this paper to serve as a guide to perceptual issues in AR. We begin by providing an updated overview of the issues affecting perceptually correct AR. Next, we describe approaches that address the problems associated with these issues, and identify research directions that could be followed to gain a better understanding of possible solutions. We conclude with a discussion of the effects that different platforms may have on perception. We hope that this paper will be useful for newcomers to the field, as well as seasoned researchers.

2 BACKGROUND AND TERMINOLOGY

Perception, the recognition and interpretation of sensory stimuli, is a complex construct [7]. Each sensory modality provides a different kind of information on which we base our interpretations and decisions. While the interplay between modalities can significantly affect how we perceive our world, analyzing these interactions is difficult. We often obtain different cues from the environment we observe, and try to match those cues. Cues can override each other, or conflict—depending on the cues, conflicts may be mentally resolved or not. It is important to note that perceptually-incorrect augmentations are often a result of conflicting cues. In this article, we focus only on issues that relate to *visual perception*, ignoring the interplay with other modalities (Shimojo and Shams [52]). Perceptual issues relate to problems that arise while observing and interpreting information from the generated virtual world, and possibly the real world. A perceptual issue may not only be caused by the combination of real and virtual information, but may also originate in the representation of the real world itself.

We will relate the perceptual issues to several classes of devices used in AR: HWDs, handheld devices, and projector-camera systems. HWDs use one of two approaches to overlay virtual information: video see-through (relying on one or more cameras to view the real world) or optical see-through (using optical elements through which the real world is viewed) (Cakmakci and Rolland [6]). Handheld devices range from cell phones to ultra-mobile computers and tablet computers, contain a screen, include an internal or attached camera, and provide a small field-of-view. Finally, projector-camera systems are stationary (Bimber and Raskar [5]) or mobile systems (Karitsuka and Sato [27]) that make use of a potentially small projector and camera combo to sense and project augmenting graphics on arbitrary surfaces.

3 CLASSIFICATION

We treat perceptual problems in the context of a visual processing and interpretation pipeline (referred to as *perceptual pipeline* in this paper), describing what problems can occur from the real environment being captured up to overlaid graphics being observed by the user. As such, we identify the following categories (see Table 1 for details):

---

[1] kruijff@icg.tugraz.at, [2] swan@acm.org, [3]feiner@cs.columbia.edu

IEEE International Symposium on Mixed and Augmented Reality 2010
Science and Technolgy Proceedings
13 -16 October, Seoul, Korea
978-1-4244-9346-3/10/$26.00 ©2010 IEEE

Table 1 – Classification of perceptual issues in augmented reality. Issues that are predominant for a specific device are tagged (H=head-worn display, M= handheld mobile device, P = projector-camera system)

Issue	Problem	References
Environment		
Structure *Clutter, patterns, visibility, depth, surfaces* (H, M, P)	Visibility, depth ordering, scene distortions. object relationships, augmentation identification, surface perception	Rosenholtz et al. [49], Sandor et al. [50], Livingston et al. [36], Lappin et al. [32], Grossberg et al. [19], Bimber et al. [5], Guehring, [20], Raskar et al. [45]
Colors *Monotony, opponency* (H, M, P)	Depth distortion, depth ordering	Gabbard et al. [17], Stone [54], Gabbard and Swan [16]
Condition *Indoor, outdoor illumination* (H, M, P)	Visibility	Stauder [53]
Capturing		
Image resolution and filtering (H, M)	Object relationships, object segmentation, scene abstraction	
Lens issues *Quality, wide-angle, flares, calibration* (H, M, P)	Object relationship, scene distortion, visibility	Klein and Murray [30]
Exposure (H, M, P)	Depth distortion, object segmentation, scene abstraction	
Color correctness and contrast (H, M, P)	Depth distortion, object relationships, object segmentation	Mantiuk et al. [39], Rastogi [46], Reinhard et al. [47], Stone [54]
Capturing frame rate (H, M, P)	Scene abstraction	Thropp and Chen [58], Ellis et al. [10]
Augmentation		
Registration errors (H, M, P)	Object relationships, depth ordering	
Occlusion *Object clipping, x-ray vision* (H, M, P)	Visibility, depth ordering, scene distortion, object relationships	Ellis and Menges [9], Wloka and Anderson [61], Berger [3], Klein and Drummond [29], Feiner and MacIntyre [12], Livingston et al. [38], Tsuda et al. [59], Kjelldahl and Prime [28], Elmqvist et al. [11], Kalkofen et al. [26], Lerotic et al [33]
Layer interferences and layout *Foreground-background, clutter* (H, M, P)	Visibility, depth ordering, object segmentation, scene distortion, text readability	House et al. [22], Robinson and Robbins [48], Bell et al. [2], Azuma and Furmanski [1], Leykin and Tuceryan [34], Peterson et al. [43], Gabbard and Swan [16], Stone, [54]
Rendering and resolution mismatch *Quality, illumination, anti-aliasing, color scheme, resolution mismatch* (H, M, P)	Depth distortion, depth ordering	Thompson et al. [57], Jacobs and Loscos [23], Rastogi [46], Okumura et al. [41], Drascic and Milgram [8]
Display device		
Stereoscopy (H)	Object relationships, visibility	Livingston et al. [36], Livingston et al. [37], Jones et al. [25]
Field of view (H, M)	Scene distortion, object relationships, visibility	Knapp and Loomis [31], Ware [60], Cutting [42
Viewing angle offset (M)	Object relationships	
Display properties (H, M, P)	Visibility, object segmentation, scene abstraction, object relationships, text legibility	Livingston [37], Rastogi [46]
Color fidelity (H, M, P)	Visibility, depth distortion, color perception	Livingston et al. [37], Fraser et al. [15], Seetzen et al. [51], Gabbard et al. [17], Jefferson and Harvey [24], Ware [60], Stone [51], Gabbard et al. [17]
Reflections (H, M)	Visibility, object segmentation, scene abstraction, object relationships	
Latency (H, M)	Scene abstraction, object matching	Thropp and Chen [58], Ellis et al [10], Drascic and Milgram [8]
User		
Individual differences (H, M, P)	Object segmentation, scene abstraction	Linn and Petersen [35]
Depth perception cues *Pictorial, kinetic, physiological, Binocular* (H, M, P)	Object segmentation, scene abstraction, depth distortion	Drascic and Milgram [8], Cutting [7], Gerbino and Fantoni [18], Swan et al. [55]
Disparity planes (H, M)	Depth distortion	Gupta [21]
Accommodation *Conflict, mismatch and absence* (H)	Depth distortion, size perception	Drascic and Milgram [8], Mon-Williams and Tresilian [40], Gupta [21]

Environment. Perceptual issues related to the environment itself, which can result in additional problems caused by the interplay between the environment and the augmentations.

Capturing. Issues related to digitizing the environment in video see-through systems, and optical and illumination problems in both video see-through and optical see-through systems.

Augmentation. Issues related to the design, layout, and registration of augmentations.

Display device. Technical issues associated with the display device.

User. Issues associated with the user perceiving the content.

As a result of the nature of human information processing, most perceptual processes also require cognitive resources. To simplify our discussion, however, we will use only the term "perception" throughout this paper.

3.1 Problems and consequences

Several problems can be identified that affect perception, and thus understanding (cognition), of augmented content. The level of impact greatly depends on the task at hand; for some tasks, partly incorrect perception of augmentation may have no effect, whereas for others, it is of utmost importance. The problems can roughly be divided into three categories:

Scene distortions and abstraction. Scenery and augmentations can become greatly distorted and partly abstracted, making correct object recognition, size perception, segmentation, and perception of inter-object (or object-augmentation) relationships difficult.

Depth distortions and object ordering. Related to the previous issue, incorrect depth interpretation is the most common perceptual problem in AR applications. Depth in AR refers to the interpretation and interplay of spatial relationships between the first-person perspective, the objects in view, and the overlaid information. These problems keep users from being able to correctly match the overlaid information to the real world.

Visibility. Users may be unable to view the content itself, mostly caused by screen problems, such as size, reflections, and brightness, or color and texture patterns that interfere with the captured environment.

Our goal in AR is the perceptually correct connection between real-world objects and digital content, supporting the correct interpretation of the spatial relationships between real and virtual objects. In general, perceptual correctness is often associated with specific sensory thresholds (Pentland [42]).

Real world objects can be overlaid with digital content (e.g., a terrain pseudocolored based on temperature), or digital content can be added to a scene (e.g., a label). The user should be able to distinguish both kinds correctly. However, incorrect depth interpretation is the most common perceptual problem in AR applications, interfering with the interpretation of spatial relationships between the first person perspectives, the objects in view, and the overlaid (embedded) information. Users are regularly unable to correctly match the overlaid information to the real world, and tend to underestimate distances in at least see-through displays (Swan [56]). Measuring perceptual problems by their level of accuracy and correctness is actually challenging (Drascic and Milgram [8]) and, outside of a few exceptions (such as the methodology of Gabbard and Swan [16]), there is no generally used framework. Nonetheless, some researchers have performed extensive perceptual tests, in particular with HWDs.

4 ISSUES AND ASSOCIATED PROBLEMS

There are many problems in the different stages of the perceptual pipeline, from the environment to which the augmentation refers, up to the interpretation by the user.

4.1 Environment

Perceptual problems associated with augmentation regularly originate in the environment to which it relates. The structure, colors and natural conditions of the environment can disturb the way in which it is recorded or perceived, creating depth problems and augmentation dependencies that must be addressed.

Environment structure. The *structure* of an environment (i.e., the arrangement of its objects) may affect all stages of the perceptual pipeline. Structure can be a great aid to provide *depth cues* (Cutting [7], see Section 4.5, depth cues). Some environments provide a richer set of cues than others, and can be used as reference points (Livingston et al. [36]), but may be biased by *context*. Both the accuracy and the precision of perceived distance may depend on the environmental context, even when familiar objects are used (Lappin et al. [32]). A key problem associated with structure is *clutter*, in which excess items lead to a degradation in task performance (Rosenholtz et al. [49]). Clutter can be difficult to segment and recognize, can cause occlusion problems, and may contain too many salient features that make general scene understanding difficult. Clutter may also obscure other problems during decision-making processes while observing data. Clutter is a problem in all further stages in the perceptual pipeline, limiting object recognition and segmentation. *Patterns* (i.e., composites of features in the environment that generally have a repeating form) can limit surface perception and augmentation identification. If an environment exhibits a pattern that resembles the pattern of an augmentation, perceptual interference will occur (see Section 4.3, layer interferences and layout). Scene understanding might also be affected by object *visibility*, referring to the occlusion relationships between objects as seen from the user's perspective. Objects may be fully or partly visible, or even completely occluded. Visibility depends on both the human-made structure (infrastructure) and the geographic features of an environment. Finally, for projector-camera systems, the environment should provide for appropriate *surfaces* on which to project. Surface angle and curvature, and characteristics like texture, fine geometric details or reflectivity may result in depth and scene distortions.

Colors. The color scheme and variety of an environment can hinder correct perception in general, and cause depth problems while interpreting it (Gabbard et al. [17]). Environments that have largely unvarying monochromatic surfaces may lose depth cues if captured at lower resolution, since the environment may end up looking amorphous. Under changing light conditions, the color scheme of an environment may also pose considerable problems (Stone [54]). Specific colors may hinder augmentation due to similarity with the chosen color scheme of, for example, labels (Gabbard and Swan [16]). Finally, surfaces with high color variances (patterns) may affect the visibility of projected images in projector-camera systems.

Environmental conditions. The state of the environment being captured can greatly influence perception: less preferable conditions bias the perception of the world around us, both through the nature of the condition, and the display of the captured image. The main variable in *indoor* environments is lighting. Lighting affects the exposure of imaging (Section 4.2, exposure), can lead to the incorrect display of color (Section 4.2, color correctness and contrast) and incorrect augmentation (Stauder [53]), and causes reflections on displays (Section 4.4, reflections) and lens flare (Section 4.2, lens issues). Furthermore, highly varying lighting (e.g., shadows on a bright wall) can make projection difficult. Lighting can also greatly affect the quality and correctness of imaging in *outdoor* scenarios. With highly varying light intensities (between 100 and 130,000 lux, a variation of three orders of magnitude), imagery can be underexposed or overexposed (Section 4.2, exposure). Furthermore, very bright environments can limit projection. Obviously, light intensity is a result of both the time of day and weather (e.g., clouds, fog and rain can limit visibility, leading to objects that are partly or fully invisible at that time). As in indoor conditions, strong light (both natural and artificial) can cause reflections and lens flare.

4.2 Capturing

Capturing refers to the process of converting an optical image to a digital signal by a camera, thus defining the first stage of providing a digital representation of an environment.

Image resolution and filtering. The *resolution* of a capturing device results in an abstracted representation of the real world by a finite number of pixels (typically arranged in a regular array at a fixed spatial frequency), each of which samples within a limited dynamic range. Low resolution sampling can lead to difficulties in visually segmenting one object from another in highly cluttered environments. With lower resolution, objects tend to merge, making correct augmentation harder, and may appear flat, losing depth cues. The problem is further exacerbated by the antialiasing performed by cameras, which generally use a Bayer color *filter* mosaic in combination with an optical anti-aliasing filter.

Lens issues. *Lens quality* varies widely in AR setups, and may cause optical aberrations such as image blurring, reduced contrast, color misalignment (chromatic aberration), and vignetting (Klein and Murray [30]). Most handheld AR platforms deploy *wide-angle lenses* whose short focal length artificially increases the size of the "window on the world," which can cause further problems. As can be seen in Figure 1, the lens shows a bigger portion of the world (*B*) in comparison to the 1:1 relationship that would be maintained using a normal focal length lens (*A*). This offset causes perspective distortion from the standpoint of the user's eye; in *B* objects are transformed in their context until they differ significantly from *A*. This leads to incorrect inter-object relationships and object sizes: objects often appear further apart (and thus smaller in the back) than they actually are. The inter-object relationships can be further biased when there is an *offset* in distance and angle between the camera lens and the display center, which may contradict the understanding of what the user thinks they are looking at "through" the display. The correction of imaging from a wide-angle lens also results in distortion of directional cues since it is artificially flattened. Finally, and similar to HWDs, handheld displays may suffer from problems related to *calibration* and *lens flare*.

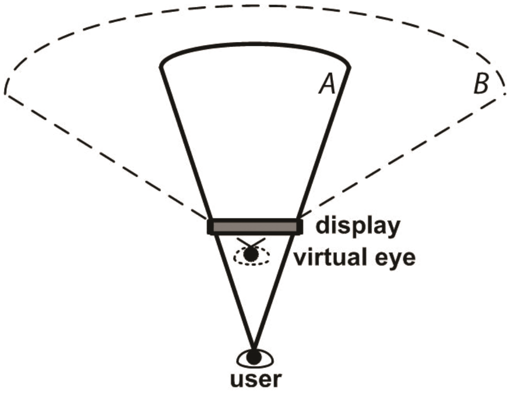

Figure 1 – Captured environment using wide-angle lens. *User* represents the actual user viewpoint and related viewing cone of a normal focal length lens, whereas *virtual eye* refers to the center of projection and viewing cone associated with the wide angle lens used in the display device. The difference causes distortion.

Exposure. *Exposure* relates to the scene luminance and the exposure time defined by the aperture and shutter speed, and hence is influenced by artificial and natural light (Section 4.1, environmental conditions). Cameras operate only within a specific range of light intensity. During capture, this can lead to under or over exposed imaging, which loses depth information, and object detail and contrast. Noise produced by the image sensor also increases as lighting decreases. Noise removes detail in shadows, and may produce incorrect atmospheric cues. Noise may make objects impossible to recognize, and can severely limit depth perception.

Color correctness and contrast. The human eye is capable of differentiating among a remarkable range of colors and contrasts. *Color correctness* refers to the fidelity of the reproduced color, which can be expressed as a variance in hue, saturation and brightness. *Contrast*, on the other hand, is defined by the difference in color and brightness of an object in comparison to other objects in the field of view. Low contrast can prevent the perception of features that may be necessary for object recognition, and result in false depth perception, since objects at different depths appear to merge in depth. Also, objects that are more blurred appear to be further away, which further distorts depth perception (Rastogi [46]). Reproducing color and contrast is often limited by the color gamut and even more limited dynamic range (contrast) that cameras can capture and that the majority of image and video formats can store (Mantiuk et al. [39]). Most image sensors only cover a part of the color gamut, resulting into tone mapping of colors available in the processed color range. Furthermore, the camera sensor capacities for white balancing and dealing with artificial light might be restricted.

Contrast limitations can be caused by the *micro-contrast* of the lens, which is the level of differentiation between smaller details that have an increasingly similar tonal value that the lens can capture. Contrast is also affected by the color capturing abilities of the image sensor, since color differentiation can create contrast.

Capture frame-rate. The *capture frame-rate* can be limited by both the camera and the display device. This can lead to visual distortions in fast moving scenes or quick display movements. Scene information will likely get lost since it cannot be captured. Lower frame rates do not seem to affect the user's situation awareness, but may decrease task performance, rendering the application useless (Thropp and Chen [58], Ellis et al. [10]). Lower frame rates seem to affect HWDs more than other platforms.

4.3 Augmentation

Augmentation refers to the registration of digital content over video imagery or on top of surfaces and can suffer from a range of problems associated with the limits of interactive technology.

Registration errors. Accurate *registration* relies on the correct localization and orientation information (pose) of a tracked device. This is often hard, particularly in outdoor environments. High-accuracy tracking is often illusory, and can only be achieved by high-quality devices. In particular, current cell phones have relatively inaccurate position and orientation sensors, resulting in far worse tracking accuracy and noticeable drifting of orientation measurements. The needed tracking accuracy depends on the environment and distance of the objects being viewed: lower accuracy tracking may be acceptable for objects far away in large scale environments where offsets are less noticeable, while accurate augmentation of nearby objects is harder. Ellis and Menges [18] found that nearby virtual objects tend to suffer from perceptual localization errors in "x-ray" or monoscopic setups. However, one may wonder if correct augmentation is not overrated, as the brain has remarkable capabilities for dealing with inconsistencies, and sometimes approximate registration may be good enough. Nevertheless, this is often not acceptable for many users.

Occlusion. *Occlusion*, the visual blocking of objects, is both a perceptual advantage for AR by providing depth cues, and a major disadvantage (Wloka and Anderson [61]). The main issue associated with occlusion is incorrect separation of foreground and background: objects that need to be rendered behind a particular object instead appear in front of it. This causes incorrect depth ordering and objects may look like they do not belong to the scene.

Once objects in the real world are fully occluded, under normal visual conditions they are (obviously) not visible anymore. Since the advent of AR, researchers have tried to make occluded or invisible objects visible again. The main method used is some form of *x-ray vision*, which allows the user to see through the objects that are in front of the occluded objects (Feiner, MacIntyre, and Seligmann [12]). However, x-ray vision is also prone to depth ordering problems, as the order of overlap is reversed (Ellis and Menges [18]). Furthermore, some of the rendering methods for visualizing occluded objects suffer from depth perception problems, in particular when used on a 2D display; for example, wireframe models are prone to the so-called Necker Cube illusion, where lines are ambiguous since they cannot be clearly assigned to either the front or the back (Kjelldahl and Prime [28]).

Layer interferences and layout. Environmental patterns can limit surface perception and augmentation identification (Section 4.1, environment structure). Depending on the features of the background and the augmentation, interference may occur where patterns intersect or visually merge, leading to foreground-background interpretation problems. These features are affected by the orientation, transparency, density and regularity of patterns, and the color schemes being used. Additionally, foreground-background pattern issues are related to problems that occur in multilayer AR systems, in which multiple layers are rendered on top of each other. A related problem is *layer clutter*, which depends on the number of labels and their opacity. Once the number of layers gets too large, labels may overlap, which may lower text readability (Leykin and Tuceryan [34]).

Rendering and resolution mismatch. The *rendering quality* defines the fidelity with which digital objects are displayed on the screen. Surprisingly, no direct relationship has been found between the level of fidelity and the judgment of depth in digitally reproduced graphics (Thompson et al. [57]). In addition to the rendering quality, *illumination* can affect both the fidelity of the augmented objects and their correct perception. Jacobs and Loscos [23] provide an excellent overview of illumination issues. Using *antialiasing* methods can also improve fidelity, but may lead to perceptual distortions. Differences in both resolution (rendering quality) and clarity (antialiasing) could be interpreted as a difference in accommodation, leading to false stereoscopic disparity (Rastogi [46]). A similar effect can be noticed between the different resolutions of the captured video background and the rendered objects (Drascic and Milgram [8]). Finally, the *color scheme* of an augmentation may affect at which depth level the augmentation is perceived to reside (Klein and Murray [30]).

4.4 Display device

The *display device* shows the augmented environment to the user and, like the other stages, can give rise to perceptual problems. Most of the problems can be associated with the screen, but some problems also arise from the relatively modest capabilities of the processor and graphics unit.

Stereoscopy. Focusing primarily on HWDs, numerous researchers have identified the main issues and problems of correctly displaying stereoscopic content. Typical problems include differences between real and assumed inter-pupillary distances (Livingston et al. [36]), visual acuity and contrast effects (Livingston et al. [37]), alignment and calibration issues (Jones et al. [25]), and issues associated with accommodation (see Section 4.5, accommodation). However, some perceptual issues that arise when an HWD is used to view a fully synthetic virtual environment may be mitigated with AR (e.g., depth perception (Jones et al. [25])). Stereoscopic display issues currently are of less importance for handheld devices and projection-camera systems. However, this may change in the future: already some commercially available stereo hand-held displays resemble binoculars and display stereoscopic content.

Field of view. *Field of view* (FOV) refers to the extent of the observable world. In video see-through displays, FOV obviously restricts how much of the real world can be seen. Although human foveal vision comprises less than 1° of the visual field, humans rely heavily upon peripheral vision, and a limited FOV makes many visual tasks very difficult (Ware [60]). However, a limited FOV does not necessarily cause depth estimation failures (Knapp and Loomis [31]). In optical see-through and handheld setups, the issue becomes complex, since the information space is not unified anymore, but separated. Humans have a horizontal FOV of over 180°, while video see-through HWDs typically support between 30°–100° horizontal FOV (although some go up to almost 180°). With optical see-through displays and handheld devices, a relatively small FOV is used for the digitized information. This leads to two variations of a *dual-view* situation. In some optical see-through displays in which the optics are frameless or surrounded by a very thin frame, users can observe the real world in a much larger portion of their FOV than what is devoted to overlaid graphics: users see the real world at the correct scale in both portions. Similarly, most handheld video see-through displays allow the user to view the real world around the bezel of the display that shows the augmented world. However, in these displays the wide FOV lens used by the camera (see Section 4.2), combined with the lens offset from the center of the display, typically creates a significant disparity between the small, incorrectly scaled augmented view and the much larger, full scale unaugmented view that surrounds it. In addition, a frame can have a profound effect on how the scene inside the frame is perceived (Cutting [42]). This raises interesting questions as to the advantages and disadvantages of both of these approaches.

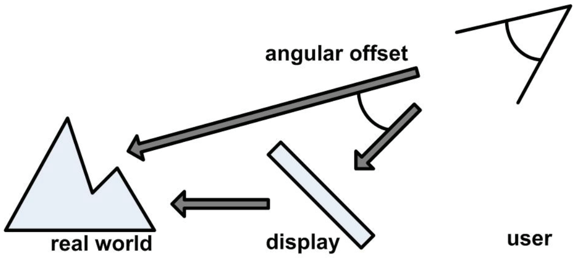

Figure 2 –Offset caused by object location and indirect display-camera angle observing the object.

Viewing angle offset. HWDs are placed directly in front of the eye, and hence there is often relatively little angular offset between the real world being observed, and the display through which it is seen. However, when using a handheld device, the angle at which the display is held can result in an angular offset (Figure 2), which can be strengthened by a possible further offset caused by the camera attached to the display. Whereas cell phones can be held relatively close to the apex of the viewing cone, most other handheld devices are typically held lower to support a more ergonomic pose. Depending on the device weight, this angular offset can be large and dynamic over time. The offset results in an indirect view of the world. This may lead to alignment problems: users may not readily understand the relationship between what is seen directly in the real world and what is shown on the screen when comparing both, which may require difficult mental rotation and scaling. In addition, the viewing angle offset can be further exacerbated by the angle and placement of the camera relative to the display (see Section 4.2, lens issues).

Display properties. Display brightness and contrast affect the visibility of content when blended with ambient light. Display *brightness* refers to the luminance of a display, and varies roughly between 250–500 candelas per square meter (cd/m^2). *Contrast* can be expressed by the ratio of the luminance of the brightest color (hence, white) to that of the darkest color (black) that the display is capable of producing. Particularly in outdoor applications, contrast is still limited due to the effects of ambient light. Ambient light will lower the contrast of the display, which leads to the inability of the human visual system to differentiate between finer nuances of color (Ware [60]). Consequently, colors, and possibly objects, may start to blend visually. Currently, there is no display technology with the dynamic range needed to show content correctly under all outdoor light intensities.

The *resolution*, the number of pixels a screen can display, has a strong effect on the perception of objects and is closely related to the pixel density (expressed as pixels per inch (PPI)). Handheld devices tend to have small screens, but are now capable of delivering high pixel density. This results in an image that is perceived as "sharp". However, users may perceive sharp objects as being closer than they actually are, affecting depth perception (Rastogi [46], also see Section 4.3, rendering and resolution mismatch). Furthermore, with high pixel density displays, very small objects may be displayed, which can result in object recognition and segmentation problems. Using a larger display is often constrained by ergonomics and form factors, since users may have difficulty carrying the device.

In projector-camera systems, the display characteristics depend on the brightness and contrast of the projector, and the albedo of the projected surface. Current handheld projector-camera systems suffer from very low brightness and hence are not usable in many daytime outdoor situations.

Color fidelity. Within AR, *color fidelity* refers to the color resemblance between the real world and what is displayed on the screen. Whereas in print media there are standard conversions between the color representations of different devices (Fraser et al. [15]), in AR such conversions are typically not addressed: current practice typically does not address mapping between sampled real-world colors and how they are represented. Also, there is usually no adjustment for color blindness (Ware [60]). Color space conversions use gamut mapping methods to shift colors into a range that is displayable on a given device (Stone [51]). The full range of natural colors cannot be represented faithfully on existing displays; in particular, highly saturated colors cannot be reproduced. This can distort color-based perceptual cues, and affect the interpretation of color-coded information (Gabbard et al. [17]). Color fidelity in outdoor environments is a highly complex issue. Changing outdoor conditions affect optical see-through displays to a greater extent than video see-through displays, because in video see-through both the real world and the overlays are displayed in the same color gamut. In projector-camera systems, texture variation across the projection surface can disturb color representation.

Reflections. *Reflections* are among the most significant effects for disturbing the perception of AR content. In HWDs, shiny objects may disturb perception. In handheld systems with an exposed screen, content may become almost invisible. This generally depends on both the ambient light conditions, such as the brightness and orientation towards the sun or artificial lights, and the objects being reflected. Reflections also introduce the problem of *multiple disparity planes,* since reflected objects are usually at a different depth than the screen content. Reflections may also be an issue in projector-camera systems, when content is projected on specularly reflective surfaces.

Latency. *Latency* relates to the possible delay of capturing or showing content, and is directly dependant on the number of frames per second the display device is able to generate. Mostly, this is dependent on the performance capacities of the processor and graphics board, which is in direct relation to the complexity of content. The performance may affect both the capturing of content (Section 4.2) and rendering quality (Section 4.4). Latency may include dynamic registration effects, in which camera imagery is updated quickly, but overlays lag behind (Drascic and Milgram [8]). Latency seems to affect the user experience and direct interaction with content more than the perception of what is being viewed. Many AR applications involve static, or at least slowly changing, content, which may not be as affected by rendering speed. Usability of applications that are dependent on fast graphics (such as games) or dexterous motor tasks that depend on overlays may suffer from perceptual limitations caused by latency (Ellis et al. [10]).

4.5 User

The user is the final stage of the perceptual pipeline and is affected differently by the various platforms.

Individual differences. The perception of the digital content presented at the display screen can be highly influenced by individual differences between users. These differences may require noticeable modifications of the way we represent information, such as icons or text. Individual differences include the user's ability to perceive detail (visual acuity), which can be corrected by prescription eyewear; eye dominance; color vision capabilities; and differences in spatial abilities (Linn and Petersen [35]).

Depth cues. Depth cues play a crucial role in the success or failure of interpreting augmented content. *Pictorial depth cues* are the features in drawings and photographs that give the impression of objects being at different depths (Cutting [7]). These cues include occlusion (opposition), height in the visual field, relative size, aerial perspective, relative density, relative brightness, and shadows. *Kinetic depth cues* can provide depth information obtained by changing the viewpoint, such as relative motion parallax and motion perspective. *Physiological depth cues* come from the eyes' muscular control systems, and comprise *vergence* (rotations of the eyes in opposite directions to fixate at a certain depth), *accommodation* (which counteracts blurring by changing the shape of the eye's lens), and pupil diameter (which counteracts blurring by changing the eye's depth of field, but which is also affected by ambient illumination levels). Finally, *binocular disparity* provides depth cues by combining the two horizontally-offset views of the scene that are provided by the eyes. Of all of these depth cues, occlusion is the most dominant (Cutting [7]), and this drives the most pervasive depth cue problem in AR: the incorrect depth ordering of augmentations. This problem becomes even more problematic when only a limited number of depth cues are available, which may lead to the *underspecified depth* of objects (Gerbino and Fantoni [18]), or even contradiction or biasing (Lappin et al. [32]), see Section 4.1, environment structure).

Disparity planes. In relation to Section 4.4 (field of view), both real-world and virtual objects can have different binocular disparities, and result in perceptual problems related to disparity planes and disparity areas. A *disparity plane* defines the depth disparity at which content is observed. Focal depth often relates to *disparity areas*—groups of objects that are in similar disparity planes. In the case of dual-view AR systems, a depth disparity will often occur: the augmentations exist in one disparity area, and the real world in another. Since these areas are at different focal depths, users may need to continuously switch their vergence (eye rotation) between these areas to compare content, or because their attention is drawn to the other area. Furthermore, in HWDs there may be an offset in depth between user interface elements rendered in the front plane of the viewing cone and the actual AR content. When users often need to use the interface, it will result in regularly switching between these different depth planes, which may lead to visual fatigue (Gupta [21]).

Accommodation. Users of stereo displays typically experience what is known as a *vergence-accommodation conflict*. This conflict occurs when the eyes converge on an object that is seen in two spatially offset views provided to the left and right eyes, but the eyes' lenses accommodate at a different (typically constant) depth—that of the display. The human visual system has the ability to tolerate this mismatch, but depth perception is distorted (Mon-Williams and Tresilian [40]). In monoscopic, video-see-through and projector-camera systems, all content is displayed and viewed on a single depth plane, and hence this problem does not exist (at the expense of losing both vergence and accommodation depth cues). Projector-camera systems will likely have no focal plane (disparity) problems. However, because all except laser projectors have a fixed focal depth, multiple non-connected surfaces that are disparate in depth will cause problems.

5 MITIGATION

Researchers have come up with various approaches to address the problems of the perceptual pipeline. In this section, we define the main directions in all stages except the user.

5.1 Environment

Augmentation of objects in *cluttered* scenes often requires a way of uniquely binding augmentations to an object using visual aids. In particular, when an augmentation overlaps several objects, a correct layout can aid this binding (Section 5.3). Once augmentation is also interfered by *pattern interferences* between objects, the visualization method can be modified (Section 5.3) to separate foreground and background layers. Augmentations may also require some *color* opponency to avoid the label visually merging with the object over which it is overlaid (Gabbard and Swan [16]). However, the object and the augmentation may become separated if the color of the object changes.

When virtual objects are *occluded*, x-ray vision methods can be used to view them. However, users often mix up the spatial relationships between virtual and real objects, in both direction and distance (Sandor et al. [50], also see Section 4.2, occlusion).

With regard to projection on surfaces, geometric and photometric methods are provided by Grossberg et al. [19] and Bimber et al. [4] to solve color pattern correction (pixel to pigment correction) and angular or curvature corrections; this research relates to work on the"Office of the Future" (Raskar et al. [45]). Similarly, illumination problems such as patterns caused by shadows on surfaces can also be addressed (Guehring, [20]).

5.2 Capturing

Capturing can be constrained by both lens and camera parameters. Solving problems caused by lenses, however, is often hard, and only a few solutions exist (Klein and Murray [30]). With respect to *wide-angle* lenses, an observer does not necessarily notice the introduced distortions, due to the dual-view condition: the view of the real world may correct potential cue conflicts or misleading perceptions, including those caused by low resolution. The dual-view situation, though, may increase disparity plane switching (see Section 4.5, disparity planes and areas) and cognitive load, and be ineffective when objects are moving fast. Theoretically, the user may also move closer towards the display, hence lowering the angular difference between *A* and *B*, to minimize the distortion (see Figure 1). Similar effects have been noticed by Cutting [7], who observed users looking at photographs; however, most users will not move closer to the display (towards the virtual eye), often being constrained by ergonomic limitations (Section 4.4, viewing angle offset).

Often, problems can be solved by using a different or improved hardware and software. The problems caused by the limited *sensitivity* of current cameras will likely be reduced with improved image sensor sensitivity and noise reduction methods**.** *Color* and *contrast,* and potential *depth problems* can be improved by using a better lens and a higher resolution sensor, or by using high dynamic range (HDR) imaging (Reinhard et al. [47]). HDR allows for a greater dynamic range between luminance in the darkest and lightest areas in a scene being captured, thus making it possible to display a wider range of intensity levels. The result is high-contrast imagery, where objects can easily be identified, but which may have a compressed color range that can affect perception. Significant ameliorating perceptual phenomena include simultaneous color contrast and simultaneous luminance contrast: the human visual system changes the perceived color of an object according to the colors that surround the object (Stone [54]).

5.3 Augmentation

One of the longstanding problems associated with augmentation, *registration*, can be mitigated by new or improved tracking methods. However, this topic falls outside the scope of this paper.

With regard to problems associated with *occluded* objects, most researchers have avoided *object clipping* to correct depth ordering, although multiple clipping solutions have appeared. Most of these approaches take a contour-based approach to clip parts of the occluded virtual object, including those of Berger [3] and Klein and Drummond [29]. Furthermore, a number of techniques have appeared that improve the x-ray visualization of occluded objects, including rendering of wireframe models or top-views by Tsuda et al. [59], distortion of the real space by "melting" by Sandor et al. [50], dynamic transparency methods by Elmqvist et al. [11], focus and context methods by Kalkofen et al. [26], non-photorealistic rendering methods by Lerotic et al. [33] and optimized wire-frame rendering by Livingston et al. [38]. Correct *illumination* may also aid in depth ordering associated with occluded objects: shadows can be helpful, providing an important depth cue. Correct illumination can also help make the scenery more believable, preventing augmented objects from looking like cardboard mock-ups. Additionally, artificial depth cues such as grids or depth labels (distance indicators) can be used.

Both House et al. [22] and Robinson and Robbins [48] provide some directions for dealing with *pattern interferences*, by changing parameters of the visualization (like stripping a texture apart); however, these methods are not typically used in AR. Other solutions are offered by Livingston et al. [38], including varying the opacity of layers, which improved the wireframe-only rendering methods by simulating the depth cue of aerial perspective.

To alleviate label *clutter* and improve text readability, Bell et al. developed *view management* methods [2], whereas Peterson et al. focused on *depth-based partitioning* methods [43]. In addition, highly saturated labels might be needed to separate them from the background, but may conflict with the rules of atmospheric perspective: such labels may be interpreted as being closer than they actually are (Stone, [54]).

Finally, dealing with the *offset* between video and rendering fidelity, Okumura et al. focused on blurring the scenery and the augmentations [41]. Similarly, applications could simply adapt the rendering resolution to that of the video background.

5.4 Display device

Display quality improves continuously. New display technologies are expected to emerge that may better cope with *brightness* and *contrast* issues. Displays often make use of backlighting and anti-reflective coatings to make content more visible, although content is often still not visible under sunny conditions. *Reflections* can be minimized by coatings, which may reduce the brightness of the screen. Similarly, reflective surfaces should be avoided on the interior of HWD enclosures. Matching the dynamic range of outdoor illumination is a problem. The head-up displays used in aircraft can match this dynamic range, and laser-based display technologies (e.g., those of MicroVision) could potentially match it, but are not widely used.

General guidelines to improve the perceptual qualities of visualizations also aid in addressing *color correctness* problems (Ware [60]). To date, color correction methods have mostly been applied in projector-camera systems (Section 4.2), and in a limited extent in handheld AR [30], but all platforms can benefit. The same thing applies to *color blindness* [14], some work on which was performed by Jefferson and Harvey [24].

Whereas content may get lost due to *latency*, handheld device users can at least retrieve information from the captured environment by direct view. The dual-view allows the user to relate the real-world in full detail with the content represented on the screen, even when the difference in disparity planes can make this hard. Furthermore, a monoscopic view that is seen and focused on biocularly in close range can make it difficult to determine the actual distance of objects. Finally, performance can still be a bottleneck, affecting all stages in the pipeline.

6 FURTHER RESEARCH

Though the approaches that we have discussed to mitigate perceptual problems can bring us closer to achieving perceptually-correct augmentation, many problems remain to be solved. In this section, we identify several research questions that deserve further work. The questions focus on the various aspects of the perceptual pipeline, thereby also covering multiple stages at once.

Environment: *How can we deal with dynamic aspects (color, illumination) of environments?* While (indirectly) some work has been performed on visual patterns, in general the structure, colors, and illumination conditions in an environment are ignored or adapted for manually. For example, dynamically adaptable color schemes that adjust to the environment conditions could be of great benefit to solve some of the object segmentation and depth problems that are caused by the environment.

Capturing: *How do high-definition and HDR cameras coupled with improved display resolution change perception on small devices?* These camera types are currently attracting interest: they are suitable for solving perceptual problems associated with resolution mismatches, and the improvement of the color gamut and contrast. However, the perceptual consequences of using HDR cameras with non-HDR displays should be carefully studied, since skewed colors can be counterproductive.

Capturing: *How can we design systems with dynamic FOV, and what effects do they have?* The FOV mismatch introduced by using wide-angle lenses with small FOV displays causes scene distortion. This could be addressed through dynamic FOV (e.g., by using liquid lens technology). Similarly, (software) methods that adapt to the actual position of the eye relative to the display could prove useful. It is unknown, though, if such methods are achievable and if they will cause perceptual disturbances.

Augmentation: *How can we further improve AR methods to minimize depth-ordering problems?* X-ray vision is useful to look through objects in the real scene. However, depth ordering and scene understanding in such systems still requires improvement: one direction that may yield benefits is multi-view perception. Similarly, label placement in highly cluttered environments still suffers from depth ordering problems. Layout and design can also be improved—apt associations need to be implemented that uniquely bind a label to an object. Cues that specify potentially disambiguating information related to the real world (e.g., a street address) might be one possibility in cluttered city environments.

Display: *Can we parameterize video and rendering quality to pixel density, to support "perceptually correct" AR?* In particular, improvements in camera capturing quality and pixel density will make it possible to use very high resolution imagery on very small screens, but, to what extent do we need to change the image's visual representation to maximize its understandability? Additionally, what is the maximum disparity between video and rendering resolution before noticeable perceptual problems arise? And, is it possible to parameterize the offset effects between video and rendering, for example with respect to mismatches or abstractions? Finally, how much rendering fidelity is truly needed? For example, depth does not seem to be affected much by fidelity (see Section 4.3, rendering and resolution mismatch).

Display: *What is the weighting of perceptual issues among different display devices?* One of the most pressing questions is the actual effect each problem has on the various display types: comparative evaluations are required to generate a per-device weighting of perceptual problems, which would be particularly useful for determining those problems that should be tackled first. In the next section, we provide an initial overview of the differences between the various platforms.

User: *What are the effects of the dual-view situation on perception and cognition in AR systems?* In particular, handheld and see-through devices introduce a dual view situation, which may help to verify ambiguous cues obtained from display content. However, its true effects are unknown; for example, disparity plane switching is expected to be counterproductive, but are the advantages of dual-view more important, and, how could we possibly minimize the effects of disparity plane switching?

User: *What are the effects of combinations of these problems on the perceptual pipeline?* A single problem can have effects on different stages, as evidenced by our repeated mentions of some issues in multiple sections; for example, sunlight can make capturing, display, and user perception difficult. What may be even more important is the actual combination of problems that accumulate through the pipeline: for instance, low-resolution capturing may affect multiple subsequent stages in the perceptual pipeline, and problems may become worse at each stage. The question is how much the accumulation affects perceptual problems on different platforms.

7 DISCUSSION AND CONCLUSIONS

Throughout this paper, we presented the main issues that affect the correct perception of augmentations on a range of AR platforms. We deliberately chose to use the "perceptual pipeline" to structure the issues involved. In this final section, we focus on the perceptual differences among the platforms, both positive and negative. Though all platforms discussed in this article support AR, there are substantial differences in how they achieve it. The parameters of every platform likely have a considerable effect on the perceptual problems they may induce. These differences affect both the suitability of a platform for a specific task and the future research that may need to be performed to improve the platform. As stated in Section 6, it is useful to identify how much a perceptual problem affects a display platform: in Table 2, we provide a first indication of the dominant factors and their effects (advantages and disadvantages), largely caused by two factors.

Table 2 – Differences in perceptual trade-offs across platforms

Platform	Advantages	Disadvantages
Head-worn		
Video see-through	Quality of depth cues, potential fidelity and speed of rendering, single disparity/focal plane, registration	No real-world verification, potentially limited FOV, vergence-accommodation conflict for real and virtual objects
Optical see-through	Real-world verification, quality of depth cues, potential fidelity and speed of rendering, potentially full FOV (real and augmented world), registration	Different disparity planes, brightness and contrast, reflections, potentially limited FOV, vergence-accommodation conflict for virtual objects
Handheld	Real-world verification	Brightness, contrast, and visibility of screens, wide-angle lens, lower fidelity, different disparity planes, higher latency, screen size, limited FOV, reflections, registration
Projector-camera		
Stationary	Potentially bright and full contrast projection, high fidelity, potentially single disparity plane, lower latency, larger FOV of projected display	Surface-based distortions
Mobile	Potentially single disparity plane	Surface-based distortions, brightness, contrast, and visibility of projection, lower fidelity, higher latency, limited FOV of projector, registration

Differences in the *sensor and processing technology* used in the various platforms have a significant effect on the perceptual trade-offs. Based on where the platform is located on the *mobility continuum* (from stationary, to backpack or belt-worn, to handheld), sensor technology in particular seems to deteriorate.

As ergonomic issues related to wearability are taken into consideration, one can notice a rapid decrease in quality of image capturing and tracking: camera sensitivity, lens quality, and both position and orientation sensor quality are constrained by size and weight needs.

Commercial interests are a strong driver here. The quality of sensors in cell phones could be significantly better, but this would increase the price of the device. For professional users, hence, bigger platforms such as those supported by a backpack system (Feiner et al. [13], Piekarski et al. [44]) may still be the first choice from a perceptual point of view. Nonetheless, these platforms are often overruled by ergonomic considerations: handheld platforms have quite a few disadvantages, but are the current platform of choice for the vast majority of consumer applications due to portability, price, and prevalence.

Display type also has a major influence on perceptual issues. Differences in FOV, screen size, and ability to verify potentially contradicting cues in the real world affect the level of influence of a perceptual problem. Furthermore, stereoscopy can provide additional information needed to disambiguate contradicting depth cues. There are also noticeable differences in brightness and contrast that can seriously limit perception in outdoor situations. With the rise of new display types, further work is required to uncover problems.

To conclude, we hope this paper will serve as a guide to understanding and tackling perceptual issues in AR. We have identified and categorized the main perceptual problems and showed how they are currently being mitigated by researchers. As we have made clear, perceptually-correct augmentation remains a difficult problem. Technically and methodologically, there is much room for improvement. In particular, this will require considerable work on evaluation to better understand the effects of different techniques on different platforms. Perceptually-correct augmentation, while challenging, will be accomplished through improved hardware and software, the development of which will form the basis for future research.

ACKNOWLEDGEMENTS

This work was funded in part through the HYDROSYS project (EU FP7/DGINFSO grant 224416) and NSF Grant 0905569.

REFERENCES

[1] Azuma, R. and C. Furmanski. Evaluating Label Placement for Augmented Reality View Management, in Proc. of the 2nd IEEE/ACM Int. Symp. on Mixed and Augmented Reality, 2003.

[2] Bell, B., S. Feiner, and T. Höllerer. View Management for Virtual and Augmented Reality, in Proc. of the annual ACM Symp. on User Interface Software and Technology, 2001.

[3] Berger, M. Resolving Occlusion in Augmented Reality: a Contour Based Approach without 3D Reconstruction, in Proc. the IEEE Conf. on Computer Vision and Pattern Recognition, 1997.

[4] Bimber, O., A. Emmerling, and T. Klemmer, Embedded entertainment with smart projectors, Computer, 2005, 38(1): 48-55.

[5] Bimber, O. and R. Raskar, Spatial Augmented Reality: Merging Real and Virtual Worlds. 2005: A K Peters, Ltd. .

[6] Cakmakci, O. and J. Rolland, Head-worn displays: a review, Journal of Display Technology, 2006, 2(3): 99-216.

[7] Cutting, J., Reconceiving Perceptual Space, in Perceiving Pictures: An Interdisciplinary Approach to Pictorial Space, H. Hecht, M. Atherton, and R. Schwartz, Editors, 2003, MITPress: Cambridge. 215-238.

[8] Drascic, D. and P. Milgram, Perceptual Issues in Augmented Reality, SPIE 1996, 2653: Stereoscopic Displays and Virtual Reality Systems III: 123-134.

[9] Ellis, S. and B. Menges, Localization of virtual objects in the near visual field., Human Factors, 1988, 40(3): 415-431.

[10] Ellis, S., A. Wolfram, and B. Adelstein. Three Dimensional Tracking in Augmented Environments: User Performance Trade-Offs Between System Latency and Update Rate, in Proc. of the 46th Annual Meeting of the Human Factors and Ergonomics Society, 2002.

[11] Elmqvist, N., U. Assarsson, and P. Tsigas. Employing Dynamic Transparency for 3D Occlusion Management: Design Issues and Evaluation, in Proc. of INTERACT 2007, 2007.

[12] Feiner, S., B. Macintyre, and D. Seligmann. Knowledge-Based Augmented Reality, Communications of the ACM, 1993, 36(7): 53-61.

[13] Feiner, S., et al. A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment, in Proc. of the First Int. Symp. on Wearable Computers 1997.

[14] Flatla, D. and C. Gutwin. Individual models of color differentiation to improve interpretability of information visualization, in Proc. of the 28th International Conf. on Human Factors in Computing Systems, 2010.

[15] Fraser, B., C. Murphy, and F. Bunting, Real World Color Management. 2004: Peachpit Press.

[16] Gabbard, J. and J. Swan II, Usability Engineering for Augmented Reality: Employing User-Based Studies to Inform Design, IEEE

Transactions on Visualization and Computer Graphics, 2008, 14(3): 513-525.

[17] Gabbard, J., et al., More Than Meets the Eye: An Engineering Study to Empirically Examine the Blending of Real and Virtual Color Spaces, in Proc. of IEEE Virtual Reality 2010, 2010.

[18] Gerbino, W. and C. Fantoni, Interposition, perspective, and the principle of minimal depth, Perception, 2007, 36: ECVP Abstract Supplement: 181.

[19] Grossberg, M., et al. Making one object look like another: controlling appearance using a projector-camera system, in IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2004.

[20] Guehring, J., Dense 3D Surface Acquisition by Structured Light Using Off-the-Shelf Components in Proc. of Videometrics and Optical Methods for 3D Shape Measurings (SPIE), 2001.

[21] Gupta, D., An Empirical Study of the Effects of Context-Switch, Object Distance, and Focus Depth on Human Performance in Augmented Reality. 2004, Virginia Polytechnic Institute and State University.

[22] House, D., A. Blair, and C. Ware, An Approach to the Perceptual Optimization of Complex Visualizations, IEEE Transactions on Visualization and Computer Graphics, 2006, 12(4): 509-521.

[23] Jacobs, K. and C. Loscos, Classification of illumination methods for mixed-reality, Computer Graphics Forum, 2006, 25(1): 29-51.

[24] Jefferson, L. and R. Harvey, An Interface to support color blind computer users, in Proc. of the ACM Conf. on Human Factors in Computing Systems, 2007.

[25] Jones, J., et al. The Effects of Virtual Reality, Augmented Reality, and Motion Parallax on Egocentric Depth Perception, in Proc. of the Symp. on Applied Perception in Graphics and Visualization, 2008.

[26] Kalkofen, D., E. Mendez, and D. Schmalstieg. Interactive Focus and Context Visualization for Augmented Reality, in Proc. of 6th IEEE and ACM Int. Symp. on Mixed and Augmented Reality 2007.

[27] Karitsuka, T. and K. Sato. A Wearable Mixed Reality with an On-Board Projector, in Proc. of the 2nd IEEE/ACM Int. Symp. on Mixed and Augmented Reality, 2003.

[28] Kjelldahl, L. and M. Prime, A study on how depth perception is affected by different presentation methods of 3D objects on a 2D display Computers and Graphics, 1995, 19(2): 199-202.

[29] Klein, G. and T. Drummond. Sensor Fusion and Occlusion Refinement for Tablet-Based AR, in Proc. of the 3rd IEEE/ACM Int. Symp. on Mixed and Augmented Reality, 2004.

[30] Klein, G. and D. Murray. Compositing for Small Cameras, in Proc. of the 7th Int. Symp. on Mixed and Augmented Reality 2008.

[31] Knapp, J. and J. Loomis, Limited Field of View of Head-Mounted Displays Is Not the Cause of Distance Underestimation in Virtual Environments, Presence: Teleoperators and Virtual Environments, 2004, 13(5): 572-577.

[32] Lappin, J., A. Shelton, and J. Rieser, Environmental context influences visually perceived distance, Perception and Psychophysics, 2006, 68: 571-581.

[33] Lerotics, M., et al. Pq-space based non-photorealistic rendering for augmented reality, in Proc. of the 10th international Conf. on Medical image computing and computer-assisted intervention, 2007.

[34] Leykin, A. and M. Tuceryan, Automatic Determination of Text Readability over Textured Backgrounds for Augmented Reality Systems, in Proc. of the 3rd IEEE/ACM Int. Symp. on Mixed and Augmented Reality, 2004.

[35] Linn, M. and A. Petersen, Emergence and characterization of sex differences in spatial ability: A meta-analysis, Child development, 1985, 56: 1479-1498.

[36] Livingston, M., Z. Ai, and J. Decker. A user study towards understanding stereo perception in head-worn augmented reality displays, in Proc. of the 8th IEEE Int. Symp. on Mixed and Augmented Reality, 2009.

[37] Livingston, M., J. Barrow, and C. Sibley, Quantification of Contrast Sensitivity and Color Perception using Head-worn Augmented Reality Displays, in Proc. of IEEE Virtual Reality, 2009.

[38] Livingston, M., et al. Resolving Multiple Occluded Layers in Augmented Reality in Proc. of the 2nd IEEE/ACM Int. Symp. on Mixed and Augmented Reality, 2003.

[39] Mantiuk, R., et al. High Dynamic Range Imaging Pipeline: Perception-motivated Representation of Visual Content, in Proc. of the Human Vision and Electronic Imaging XII, IS&T/SPIE's Symp. on Electronic Imaging, 2007.

[40] Mon-Williams, M. and J. Treselian, Ordinal Depth Information from Accommodation?, Ergonomics, 2000, 43(3): 391-404.

[41] Okumura, B., M. Kanbara, and N. Yokoya, Augmented Reality Based on Estimation of Defocusing and Motion Blurring from Captured Images, in Proc. of the 5th IEEE/ACM Int. Symp. on Mixed and Augmented Reality, 2006.

[42] Pentland, A., Maximum likelihood estimation: The best PEST, Perception & Psychophysics, 1980, 28(4): 377-379.

[43] Peterson, S., M. Axholt, and S. Ellis. Label segregation by remapping stereoscopic depth in far-field augmented reality, in Proc. of the 7th IEEE/ACM Int. Symp. on Mixed and Augmented Reality, 2008.

[44] Piekarski, W., R. Smith, and B. Thomas. Designing Backpacks for High Fidelity Mobile Outdoor Augmented Reality, in Proc. of the 3rd IEEE/ACM Int. Symp. on Mixed and Augmented Reality, 2004.

[45] Raskar, R., Greg Welch, et al. The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays, in Proc. of SIGGRAPH'98, 1998.

[46] Rastogi, A., Design of an interface for teleoperation in unstructured environments using augmented reality displays. 1996, University of Toronto

[47] Reinhard, E., et al., High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting. 2006: Morgan Kaufmann.

[48] Robinson, M. and K. Robbins, Toward perceptual enhancement of multiple intersecting surfaces, in Proc. of SPIE 2005, 2005.

[49] Rosenholtz, R., Y. Li, and L. Nakano, Measuring visual clutter, Journal of Vision, 2007, 7(2): 1-22.

[50] Sandor, C., et al. Egocentric Space-Distorting Visualizations for Rapid Environment Exploration in Mobile Mixed Reality, in Proc. of the IEEE Conf. on Virtual Reality, 2010.

[51] Seetzen, H., et al., High dynamic range display systems, ACM Transactions on Graphics, 2004, 23(3): 760-768.

[52] Shimojo, S. and L. Shams, Sensory Modalities are not Separate Modalities: Plasticity and Interactions, Current Opinion in Neurobiology, 2001, 11: 505-509.

[53] Stauder, J., Augmented Reality with Automatic Illumination Control Incorporating Ellipsoidal Models, IEEE Transactions on Multimedia, 1999, 1(2): 136-143.

[54] Stone, M., A Field Guide to Digital Color. 2003: A K Peters.

[55] Swan II, J., D. Hix, and J. Gabbard, Perceptual and Ergonomic Issues in Mobile Augmented Reality for Urban Operations. 2003, Naval Research Laboratory.

[56] Swan II, J., et al., Egocentric Depth Judgments in Optical, See-Through Augmented Reality, IEEE Transactions on Visualization and Computer Graphics, 2007, 13(3): 429-442.

[57] Thompson, W., et al., Does the quality of the computer graphics matter when judging distances in visually immersive environments?, Presence: Teleoperatore and Virtual Environments, 2004, 13(5): 560-571.

[58] Thropp, J. and J. Chen, The Effects of Slow Frame Rates on Human Performance. 2006, Army Research Laboratory.

[59] Tsuda, T., et al. Visualization methods for outdoor see-through vision, in Proc. of the 2005Int. Conf. on Augmented Tele-Existence, 2005.

[60] Ware, C., Information Visualization: Perception for Design. 2000, New York: Morgan Kauffman.

[61] Wloka, M. and B. Anderson. Resolving occlusion in augmented reality, in Proc. of the 1995 Symp. on interactive 3D Graphics, 1995.