

On the Interplay between Data Overlay and Real-World Context using See-through Displays

Andrea Albarelli, Augusto Celentano,
Luca Cosmo
DAIS, Università Ca' Foscari Venezia
Venice, Italy
(albarelli,auce,luca.cosmo)@unive.it

Renato Marchi
Gruppo PAM
Spinea, Venice, Italy
renato_marchi@gruppopam.it

ABSTRACT

The recent availability of affordable see-through wearable displays has fostered the development of several new interfaces and applications. Some of them take the augmented reality path, by seeking the blending of physical objects with overlaid 3D models or textual information. Some, on the other hand, are much simpler and follow a rather basic paradigm where the spatial integration between real world and data overlay is dropped. This is the case, for instance, with most applications based on Google Glass hardware, where textual data and images partially share the field of view of the user, but are not pinpointed to physical features. This is a rather important difference, since it marks the shift from a *cooperative* see-through mode, that characterizes proper augmented reality, to a *competitive* overlay, where the user attention is actually contended between real objects and displayed data. To this end, the user focus must continuously shift from one context to the other, possibly leading to both reduced productivity and usage strain. With this paper we are addressing exactly this issue. Specifically, we are assessing the role of different properties of the overlay, including the level of occlusion, the depth of the data layer, the position of the view frustum and the impact of stereo vision. Such study has been implemented by mean of a real-world evaluation which has been performed using a general purpose see-through device in a practical application scenario.

CCS Concepts

•Computing methodologies → Mixed / augmented reality; •Human-centered computing → Usability testing; Empirical studies in HCI;

Keywords

Augmented Reality, See-through display, Information overlay, User focus

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHIItaly 2015, September 28 - 30, 2015, Rome, Italy

© 2015 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-3684-0/15/09...\$15.00

DOI: <http://dx.doi.org/10.1145/2808435.2808455>

1. INTRODUCTION

The most widely accepted connotation of augmented reality implies the blending of rendered information within real scenes, according to some common reference frame. This could include several application scenarios, such as the alignment of virtual parts to actual mechanical components [2, 23], the visualization of items in architecture or interior design [22, 24], the supply of touristic information aligned with real views of the location [7, 15], or even the augmentation of physical objects with projected interfaces [6]. All these applications have one thing in common: the virtual information is designed to blend with the real scene in a seamless way. To this end, the overall user experience is designed over a combined real/virtual scene and there is no difference between the two contexts. This is true regardless of the display hardware, which can be a fully immersive helmet or a simple hand-held smartphone or tablet. The intended experience also does not depend on the type of information: in fact, while 3D models can blend in a more natural way, even icons or textual information can be rightfully pinpointed to physical objects as if they were labels or billboards.

This well-grounded paradigm breaks in many modern applications, where the goal is not actually to blend virtual and real objects, but simply to present information related to a scene that can be useful to the completion of a task. With this respect, the main feature of the see-through display is to enable the user to read the data without diverting the attention from the scene.

Such presentation model is not really new, in fact it has been adopted for decades on very specialized hardware, such as HUD displays for military vehicles [12] or automotive applications [27]. Recently, however, it is becoming very popular due to the emergence of consumer-grade see-through devices, such as the popular Google Glass platform [28], and many other offerings from several vendors, including Vuzix [20], Epson [18, 30] and Microsoft [8].

This new class of devices enables a whole new range of applications, especially in the field of connected interaction and social networking. In fact, the first use cases presented by Google for their Glass platform were focused on accessing social websites, taking pictures or getting directions while still performing daytime activities. Soon, also professional usages become apparent, including the ability of receiving information or accessing technical documentation while working in the field. The key feature about these scenarios is that the user should be able to access data without losing the focus on what he/she is doing. However, differently from pure augmented reality applications, the overlay does not blend

to a common world, but rather gets in the way. We characterize this behavior with the term *competitive see-through*, since the focus of the user must switch (even subconsciously) between the real world and the overlaid data. The term is in contrast to the *cooperative see-through* found in standard augmented reality.

1.1 Related work

Since the main goal of see-through devices is to enable practical and comfortable augmented interaction, several usability studies have been performed since their introduction. Many of them, however, focus on the so called *video see-through*, i.e., augmented reality devices where the user is not able to directly observe the world, but rather sees a composite digital image including both virtual objects and the real scene captured by one or two cameras [1, 3, 10]. This is due to the fact that the involved technology was simpler form a constructive point of view and has been available for a longer time than lightweight optical solutions, such as Google Glass and similar devices. Moreover, one clear advantage of video see-through is that no calibration of the user eye is needed [5, 11], since the augmentation happens in the camera reference frame. This of course comes at the price of an indirect view of the surrounding world and thus a less accurate perception of the real scene [25, 26].

Optical see-through is considered in [13], where the authors address the problem of mutual occlusion between real and virtual objects, due both to the incorrect depth ordering and the semi-transparent nature of the display. The effect of the display focal distance with respect to the perception of virtual objects depth has been studied in [17] using a special device with addressable focal planes using a liquid lens. The role of stereoscopic vision for see-through augmented reality has been studied in [4] and more recently in [16].

All the aforementioned studies are actually centered on *cooperative* interaction. Studies involving *competitive* augmented reality are scarce. A possible exception is [9], where the problem of occlusion between the real world and overlaid information is examined. The study, however, is only performed within a simulated environment, without using a real background and only accounting for transparency. More recently, the role of displayed information depth and of multiple data layers have been addressed in [31], finding that standard rules about contrast and color combinations could break with see-through systems. Finally, in [21], the authors investigate the use of stereoscopic vision to ease the identification of overlaid labels, especially in cluttered conditions. Still, a study accounting for many simultaneous visual factors in a real environment is currently lacking.

1.2 Contribution of this paper

With this paper we are specifically addressing the interaction problems arising from the interplay between the world the user has to interact with, and the overlaid information hindering his/her view within a *competitive* see-through setting. In the next section we shall offer a classification for the different visual factors that can influence the interaction. For each of these factor we shall describe the expected effects on the interaction. The proposed classification is useful to define a common ontology to use when analyzing competitive see-through interfaces. In Section 3 we select some of these factors and we perform a study about their influence in a real application scenario. The scenario has been cho-

sen in the context of a cooperation between Gruppo PAM, a large Italian retail channel, and DAIS - Università Ca' Foscari Venezia, aimed at the study of applications of wearable devices and reconfigurable mobile architectures for in-store operations, experimenting the use of different devices for different purposes and specific interfaces into a unique integrated enterprise environment. The adopted scenario models the re-stocking/replenishment activity performed in the stores, exhibiting characteristics that can be considered general enough from the point of view of the combined perception of the physical world and the synthetic data in a same visual space.

2. INTERPLAY FACTORS

In the following we are referring exclusively to competitive see-through interfaces. To this end, any information that is displayed is not spatially integrated in the scene observed by the user and represents a potential form of distraction for any activity requiring an interaction with the real world. Within this scenario, the design of an interface is subject to two requirements which, unfortunately, are contrasted. The first requirement is that the information should be easy to read and the effort that the user should make in order to focus his/her attention on the overlay should be minimal. The second requirement is that the operator should be able to interact with the environment in a seamless way. This obviously creates a quandary since the more prominent the overlay data is, the higher would be the obstruction of the user view. In order to find the correct balance between these two goals, several factors could be tuned. In the following section we describe each factor, while a summarization is presented in Table 1.

2.1 Stereoscopic vision

Google Glass is probably the most popular modern see-through display, even if it is not yet available for large scale distribution. Its design is strictly monocular, in fact it provides a single overlay transmitted to the right eye using an optical system made of lens and a prismatic mirror. It is reasonable to think that the original goal of the designer was to make the overlay less obtrusive. In addition this choice makes the device lighter and easy to integrate with standard eye-wears. Still, it is well understood that providing different images [14] to the right and left eye could result in a noticeable level of discomfort. This is true even for small differences, including disparate image quality [29]. Indeed, the lower amount of occlusion could fail to offer an effective advantage, since the brain would frantically try to make a sense out of the different views, hindering the ability of the occlusion-free eye to work independently.

During the last couple of years, companies with a long track on imaging technologies, namely Sony, Epson and Samsung, started to offer binocular wearable displays. Such systems are heavier to wear, however they can offer the same image (or coherent views) to both eyes. Note that the binocular design is indeed more flexible, in fact nothing prevents the interface designer to use only one display or to place the information in a peripheral section of the view frustum. It's not clear, at the current state of the knowledge, if stereoscopic vision is more well-suited for competitive augmented reality applications or if the user can get acquainted with the weird feeling induced by a monocular display.

Interplay Factor	Advantages	Limitations
Stereoscopic Vision	More natural view of the data Ability to set different depth layers	Higher level of occlusion More complex hardware required
Information depth	Ability to place the data at a give position with respect to real objects Differentiation between foreground and background	Requires stereoscopic vision
Information position	Allocation of the most prominent place to important information and use of peripheral areas for secondary data	Occlusion of the real scene Difficulties in reading peripheral data
Opaque background	Increased contrast and readability Better separation of data from real objects	Occlusion in areas with opaque background Standard behavior for some devices
Information type	Ability to use icons for simpler data	The ability to notice the information depends on its representation and position

Table 1: Synthesis of different interplay factors and of their potential impact on a see-through interface.

2.2 Information Depth

When stereoscopic vision is enabled, every item on the overlay is implicitly associated to a depth level. This is a result of the disparity assigned by construction to the wearable display, which, in turn, can be modified at rendering stage by increasing or decreasing the depth of single items or of the whole image plane.

The depth of information is a factor that can result in three different effects. The first one is related to occlusion: by placing items far from the user they appear smaller as a result of basic projective geometry principles. Thus, far objects are less obtrusive than near ones. The second effect is related to the attention level needed to focus on the items: objects in the background are less noticeable unless the user specifically pay attention to them [19]. Finally, the most important effect of information depth with respect to the context of this study is probably the relative position between the overlay data and the real world.

Indeed, since the information displayed is not directly pinpointed to the scene, an excessive depth of the overlay could easily result in items that would be perceived to be behind real objects, while still appearing as superimposed. In our opinion, this contravariant effect could lead to another perception problem, due to contradictory information, which could make difficult both to interact with the real world and to read easily the data. For this reason it is important to study the role of information depth in real scenarios.

2.3 Background and contrast

Some devices, such as the Vuzix M100, are not provided with a transparent display, but rather with a blinding shutter that filters out most of the light and makes the background of the data opaque (we still consider the Vuzix M100 a see-through display, since its monocular nature allows to see the real world blended through the other eye). This solution has probably been adopted to enhance the contrast of the data, however it could increase the occlusion and worsen the discomfort due to different views in the two eyes.

Still, the additional contrast could be an advantage in many situations where an interaction with the real objects is sought. This could be the case when the background offered by the real world is strongly non-uniform and makes

it difficult to read any overlaid text. Also, the presence of an opaque background could improve the perception of the overlaid information resolving the clashing between the depth of the data and the physical objects (as the background would blind out the real scene). Note that an opaque background does not require physical blindness, as for the Vuzix solution. Indeed we can simulate an occluding background even in fully transparent systems (such as Epson Moverio or Google Glass) by using solid colors and strong backlighting.

2.4 Information position

The competition for user attention between the overlay data and the real scene also depends on the position of the data with respect to the user view frustum. Specifically, information placed in central areas will be easier to read and will capture most of the attention. However, the already discussed problems about occlusion still apply. The occlusion problem could be a bit ameliorated if the information on the center of the frustum is displayed with a transparent background, however, a proper set of tests must be performed to assess if this precaution is effective or if confusion between overlay and scene still occurs. On the other hand, placing all the data on peripheral regions could force the user to continuously shift the attention from the scene to the data and vice-versa.

2.5 Information type

A last factor that might be relevant is the type of the information, i.e., textual data versus iconic representations. The former is more flexible, however it requires a higher amount of attention by the user that must be able to deliberately focus on the data to read them. By contrast, icons or pictorial information can be grasped with minimal attention since they are recognized mostly by early vision processes, especially when the user is trained or has used the application for a long time.

3. A TEST BED FOR EVALUATING COMPETITIVE SEE-THROUGH

The possible combinations of the factors described in Section 2 are many and it would be impractical to design an

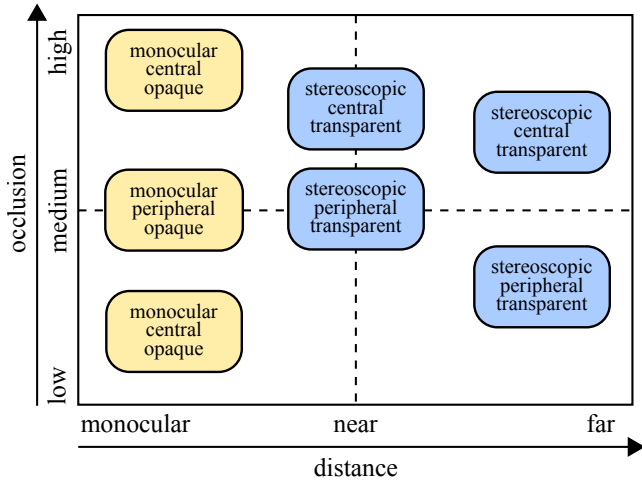


Figure 1: Relative localization of the proposed interface flavors within distance/occlusion quadrants.

evaluation scenario accounting for all the available configurations. Rather, we decided to focus on the 7 combinations shown in Figure 1. We placed the configurations on four reference quadrants that allow to highlight how they relate one to each other with respect to the factors described. We provided respectively three monocular (in yellow in the left side of Figure 1) and four binocular configurations (in blue in the middle and right side of Figure 1). We consider monocular vision to be assigned to a specific distance level, which we deem to be extremely near, as the absence of stereo vision makes the screen corresponding to the retinal imaging plane of the viewing eye.

For the monocular configurations we consider one case where the data is in central position (with respect to the frustum of a single eye) and drawn on a transparent background and two conditions with an opaque background, respectively with central and peripheral positioned data. The rationale is that with monocular view, even data drawn centrally and with an opaque background, while exhibiting high occlusion, would not render the user unable to see the scene.

Differently, the available feasible choices with binocular setup only comprise central data with transparent background and peripheral data with opaque background. Note that, in Figure 1, we deem the peripheral data with opaque background to be more or less as obtrusive as the monocular configuration. By contrast we have assigned a higher potential occlusion score to the central and transparent case because of the coverage of both eyes. Also, the far variants of the binocular setups are placed in an area of the quadrants associated to a reduced occlusion.

Note, finally, that in our tests we did not consider the information type, since this factor varies too much with respect to the application. In the following tests we only used textual information.

3.1 Evaluation Scenario

The actual influence of each described factor depends also on the specific scenario it is applied to. In order to adopt a test setup that could be considered general enough we defined an application interleaving operations on real objects and reading of overlay data.

Specifically, we are simulating the restocking operation that could be performed in a store by an employee in charge of the shelves resupply. This has been done within a controlled environment by simulating the store with a life size picture of three shelves and letting the user interact with them (Figure 2). The size and colors of the products has been chosen to offer a high variability of the background. Even if the picture is flat compared to the real shelves, thus introducing a simplification in the depth perception, this simulation does not change the meaningfulness of the test since the alignment of the products in the selected type of shelves reduces to a minimum the depth differences among them.

The interaction model has been designed as follows:

- The name of a product appears on the wearable display. The user must find the product on the shelves and read its barcode with a dedicated reader;
- If the correct product has been selected, the quantity to be restocked appears on the display as a circular icon containing the number of units to be ordered;
- The user must press buttons on the barcode reader to input the correct quantity. This is done using just two buttons that respectively increment and decrement the resupply quantity. Once satisfied the user must press a third button to confirm;
- If the reordered quantity is correct another product is presented and the process repeats.

The idea underlying this evaluation scenario is that the user is required to be able to read easily the data on the display and, at the same time, to interact in a comfortable manner with the environment. In fact, the display shows information that is required to perform the overall task; however, the ability to search among the products in an effective and timely manner is critical to perform the restocking operation. In addition, the barcode reading step requires a correct alignment between the reading device and the product label, which could also be hindered by an obtrusive interface.

3.2 Technical Setup and Viewing Conditions

We implemented the evaluation application on an Epson Moverio BT-200 wearable display. Such device exhibits two independent virtual screens, one for each eye. This way we were able to present to the user both monocular and stereoscopic interfaces. Monocular interfaces were simulated by



Figure 2: The setup used to perform our evaluation



Figure 3: Three of the tested viewing conditions (see the text for details). These shots have been taken by placing a camera behind the device lens. Only the monocular variants are displayed, since it is not possible to properly and meaningfully represent stereoscopic views.

turning off one of the screens. While this can be considered a reasonable approximation for testing purposes, it should be noted that this is not exactly the same condition offered by true monocular displays such as Google Glasses or Vuzix M100.

This is due to two main differences. The first one is that the build of the system is different: the device we used is symmetric and balanced, while true monocular displays are located on a side of the user head. According to our experience this results in an improved wearing comfort and a more stable alignment between eyes and display. Since we are interested only in the visual aspects of the user experience this is not a real concern.

The second difference is more relevant. In fact, while a single screen can be cleared to simulate monocular vision, its presence can still be perceived, due to a minimal glowing of the display and to the actual lens assembly. We were not able to completely eliminate this effect, however we reduced its effect by tuning contrast and brightness parameters. A possible alternative could be to use real monocular devices, however this would result in several relevant differences, including resolution, virtual distance of the display and relative position with respect to the view frustum. We felt that all these hindrance would be a bigger hurdle in order to obtain a fair comparison between different viewing conditions, for this reason we resorted to use the same device for the whole set of tests.

The barcode reader was a Motorola/Symbol CS4070 connected to the Moverio display using bluetooth. This is a reliable dedicated device with a very high reading capability. Still, a correct alignment between the device and the barcode is needed, which in turn requires the user to be able to correctly perceive the orientation and position of the target barcode and the reader itself.

The test shelves have been printed on six A0 posters, organized in a 3×2 grid representing 3 shelves for a total size of about 1.7×3.5 meters. Each product was reproduced in color and at real-world size. Each user has been asked to perform the actions described in Section 3.1 for a total of five random products and the time required for each reorder was recorded. The test has been repeated for each one of the following viewing configurations.

3.2.1 Monocular/Central/Transparent

For all the monocular versions of the interface, data is presented only on the right display. This is similar to what happens with native monocular displays, albeit with the aforementioned limitations. In this configuration (abbreviated

in MCT in the following discussion) the data is centered. This means that, in principle, it should be aligned with the principal direction of the user’s view frustum. We noticed, however, that subjective differences in the way the device is worn can result in a slight offset. The background of the data is completely transparent. Figure 3a shows how this view configuration appears with respect to the background. The image has been captured by placing a camera behind the device lens, thus simulating the human point of view. The slight blur around the edges of the view is due to visual artifact generated by the prismatic lens and is actually experienced by the user as well as captured by the camera.

The rationale behind this configuration is that a centrally placed interface should be more comfortable to read and in general it should be easier for the user to focus the attention on it without losing the sight of the environment. This, of course, comes at the cost of more occlusion, which is a bit mitigated by the transparent background.

3.2.2 Monocular/Peripheral/Opaque

This view configuration (abbreviated in MPO), depicted in Figure 3b, tries to enhance the contrast of the data by placing it on a solid white background. Of course this will result in even more occlusion. For this reason the data is moved to a peripheral position. The idea is that this configuration should provide a balance between occlusion and data readability.

3.2.3 Monocular/Central/Opaque

One of the design goals of monocular displays is to occlude just a single eye and let the brain merge the displayed information with the real scene. The practical aspects of this configuration (abbreviated in MCO) have been debated a lot [14], still, for the sake of completeness, we feel that it can be useful to test also a viewing condition when one eye is almost fully covered by the displayed information.

In this viewing configuration, shown in Figure 3c we placed the text in central position and on a solid white background. Data readability should be maximum and we are interested in studying if the non occluded eye gathers enough information about the real world and if the experience is still comfortable.

3.2.4 Stereoscopic viewing conditions

For the stereoscopic viewing conditions we only tested the peripheral placement with opaque background and the central location with transparent background. This is because the central position with opaque background leads to a complete occlusion and thus to the inability to perform any task

Evaluation criteria	Average							Standard Deviation						
	MCO	MPO	NPO	FPO	MCT	NCT	FCT	MCO	MPO	NPO	FPO	MCT	NCT	FCT
1. Ease of reading the displayed information	2,60	2,90	3,90	3,20	3,40	3,90	2,80	0,84	0,99	0,57	0,79	0,70	0,74	1,03
2. Ease of switching between displayed information and real world scene	2,50	2,90	3,60	2,70	4,00	4,20	3,10	0,71	1,10	0,84	0,82	0,67	0,63	0,99
3. Contrast between displayed information and real world background	3,00	3,10	4,00	3,30	3,40	3,80	2,80	1,05	1,29	0,67	0,82	1,07	0,42	0,79
4. Ease of localization of products in the real world	3,40	3,40	3,30	3,20	4,30	4,20	3,70	0,97	1,07	0,95	1,03	0,48	0,42	0,82
5. Alignment between the barcode reader and the product tags	3,80	3,80	4,00	3,50	4,50	4,10	4,10	1,14	1,14	0,94	1,27	0,53	0,99	0,88
6. Orientation and movement in the real world	3,50	3,60	3,60	3,30	4,40	4,10	3,80	0,85	1,17	1,07	1,06	0,52	0,57	1,03
7. Overall quality level and comfort of the visual experience	2,50	2,80	3,40	2,80	3,70	4,00	3,10	0,85	0,92	0,97	0,63	0,48	0,94	0,88
8. Overall ease of the task execution	3,70	3,70	3,90	3,30	4,10	4,10	3,40	1,06	1,06	0,88	0,82	0,57	0,57	0,84
Objective measurement														
Time required to complete one resupply task of five products (in seconds)	97	102	106	91	88	83	107	26	43	48	33	23	20	53

Table 2: Survey results for the different viewing configurations.

involving actions in the real environment. In addition, differently from the monocular configuration, stereoscopic visualization allows to set a distance of the data plane from the user. We think that this distance could produce a noticeable effect in terms of comfort and performance since the data, albeit not integrated, is perceived in a specific spatial position and, for this reason, it could interfere with the perception of real objects. Specifically, we tested two distances for the data plane. The first one, labeled *near*, is about 1.5 meters far from the user. The second one, labeled *far*, is at about 3 meters from the user.

The total number of stereoscopic viewed condition tested is 4. We labeled them as follows:

- Near stereoscopic/Peripheral/Opaque (NPO)
- Near stereoscopic/Central/Transparent (NCT)
- Far stereoscopic/Peripheral/Opaque (FPO)
- Far stereoscopic/Central/Transparent (FCT)

4. SURVEY ANALYSIS AND DISCUSSION

We have conducted the survey with 10 users: 7 men and 3 women, aged 23–34, average age 28.3, standard deviation 3.2. They performed the task presented in Section 3 for each of the viewing configurations described in Section 3.2. For each viewing configuration the task involved five products chosen at random within a panel of 60 different products. All the users exhibited normal binocular vision; some of them were using prescription glasses, but this was not considered a factor influencing the test since the Epson Moverio device is large enough to accommodate them without changes in the device pose. After testing each viewing configuration we asked the users to fill out a survey about his/her experience.

The survey, whose results are summarized in Table 2, included three sets of evaluation criteria. The first set included

criteria 1–3 and was directed to measure the readability and quality of perception of data overlaid to the real market shelf. The second set included criteria 4–6 and was directed at assessing the quality of perception and interaction with the real world while wearing the Moverio glasses and performing the task. The last set included criteria 7–8 and inquires the general usage comfort experienced during the execution of the task.

For each criterium the user expressed an evaluation on a five points Lickert scale, where higher values mean more satisfaction and comfort. In addition to these subjective evaluation we also collected an objective quantitative measure, i.e., the average time required to complete a single resupply task of five products for each viewing condition. This measure is coarse and embeds several different factors that influence the time: for example, the simulated shelves required some movement of the user to locate the products; some of them were similar in aspect and required reading the product labels accurately to avoid mistakes. We think however that such measure, compared with the user subjective evaluation, could offer a reasonable insight about the actual suitability of each viewing configuration. Indeed, the results confirm our hypothesis.

In Table 2 the first four columns are related to opaque viewing configurations, the last three columns to the transparent configurations. We have used grey levels together with numeric values to make visually more evident the relations between the different viewing configurations and the different evaluation criteria. Even if there are variations and disagreements between the users, that result in the lack of large differences in the average values, Table 2 shows some evidences.

First, the users were not unanimous in their evaluation; the average values are in the restricted range 2.50–4.50,

with standard deviations generally lower for higher average scores but higher for lower average scores. A partition in two sets of viewing configuration is however visible. At one side, the better configurations are, orderly, NCT (Near focus stereo, Central, Transparent background), MCT (Mono, Central, Transparent background) and NPO (Near focus stereo, Peripheral, Opaque background), that show high average scores in almost all the evaluation criteria. Of the three, the NPO configuration suffers in the localization of the real products due to the view occlusion given by the opaque background. At the other side, the other three configurations with opaque background and the FCT configuration (Far focus stereo, Central, Transparent background) show lower scores, with the exception of the evaluation criterium #5 related to the barcode reader alignment. The low performance of the FCT configuration was expected, as we noted in Section 2.2. Indeed, also the ease of reading the displayed information (criterium #1) is marked with a low average score, due to the unnatural spatial relation between the real shelf that should act as a background and the text which instead of floating in front it appears behind it in perspective.

The two “winning” central transparent configurations NCT and MCT are very close, with low values for the standard deviation in most criteria. The overall quality and comfort of the visual experience (criterium #8) is better with the stereo configuration, as is the ease of reading the displayed information (criterium #1); the result is coherent with our comment in Section 2.1 about the drawbacks of monocular vision in integrating real and virtual images. The monocular view is however preferred in the evaluation criteria dealing with the real world perception (#4-6).

The evaluation criteria that are most influenced by changes in viewing configurations are the ease of reading the displayed information (#1) and the overall quality level and comfort (#7). On the contrary, the barcode reader alignment (#5), the overall ease of task execution (#8) and the ease of localization of products in the real world (#4) are not severely impacted by configurations visually less suitable.

A pairwise comparison between opposite viewing configurations reveals that, on the average, binocular vision is evaluated better than monocular vision, due to a lack of a focus reference in the latter. Comparing central with peripheral vision, the former is judged superior, but the occlusion conditions are not the same for the two cases, so the result is only partially meaningful. Finally, in stereo vision far focus is evaluated worse than near focus, for the already recalled reasons related to the background-foreground spatial perception.

Globally, the best evaluation in all the evaluation criteria, with the exception of the contrast between displayed information and real world background (#3) comes from the configurations with central display and transparent background, followed by the peripheral opaque configuration. The contrast between the displayed information and the background is better in this last case due to the partial occlusion that puts into evidence the displayed text.

The analysis of the execution times matches the subjective evaluations provided by the users for the two preferred viewing configurations, assigning to the NCT and MCT configurations the lowest average execution times and lowest deviations in time. Differences do however appear in the other configurations: in particular, the third preferred configura-

Criteria	1	2	3	4	5	6	7	8
P-value	0.002	4×10^{-5}	0.051	0.022	0.449	0.146	0.001	0.236

Table 3: P-values resulting from an ANOVA test performed over the different evaluation criteria.

tion, NPO, shows an execution time exceeding the time of the less appreciated configurations, while the FPO configuration, even if evaluated poor by the users, shown an average execution time close to that of the MCT configuration.

Finally, in order to assess the significance of the obtained insights, we performed an ANOVA test over the user answers to each evaluation criteria, even if the number of collected samples is small to guarantee the statistical reliability of the test; the large separation of averages, however, still supplies a strong hint about the perceived differences between the proposed view modes, and the ANOVA test should confirm this perception. The results are shown in Table 3. It can be seen that the first two criteria are really significant. The third one is also quite discriminant (p -value = 0.051). Regarding the second set of criteria, since criteria #5 and #6 appear to be less discriminant, we can confirm that the different viewing conditions do not make a relevant impact with respect to the ability to handle the barcode reader or to move in the real world. Also, we can analytically confirm that the perceived usage comfort (#7) is significantly more differentiated among viewing conditions than the perceived overall easiness of the task (#8).

5. CONCLUSION AND FUTURE WORK

In Section 4 we detailed several insights resulting from our study. The most relevant ones can be summarized in two observations. The first one is that the largest impact of different visualization factors for competitive see-through occurs with respect to the ability to perceive overlaid data. In fact, the interaction with the real world is less affected, with the possible exception of activity requiring detailed visual inspection, such as product identification in clutter. The second observation is that central overlay positioning with transparent background is a generally preferred solution, regardless of the system being binocular or monocular, providing that the depth of the overlay in binocular setups is not too far from the user point of view.

The current analysis has been done in a simulated environment using pictures of a few market shelves, random selection of products and simulated resupply information. A further test is planned in next weeks, that will be performed in the real market environment with the personnel of the market devoted to this task, testing only the three best evaluated configurations in a real-life resupply situation.

6. REFERENCES

- [1] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. In *14th Annual ACM Symposium on User Interface Software and Technology*, UIST '01, pages 101–110, New York, NY, USA, 2001. ACM.
- [2] J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo, E. Damiani, and M. Ivkovic. Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 51(1):341–377, 2011.

- [3] S. Debernardis, M. Fiorentino, M. Gattullo, G. Monno, and A. E. Uva. Text readability in head-worn displays: Color and style optimization in video versus optical see-through devices. *Visualization and Computer Graphics*, 20(1):125–139, 2014.
- [4] S. R. Ellis and U. J. Bucher. Distance perception of stereoscopically presented virtual objects optically superimposed on physical objects by a head-mounted see-through display. In *Human Factors and Ergonomics Society*, volume 2, pages 1300–1304, 1994. cited By 8.
- [5] Y. Genc, M. Tuceryan, and N. Navab. Practical solutions for calibration of optical see-through devices. In *1st International Symposium on Mixed and Augmented Reality*, ISMAR '02, pages 169–, Washington, DC, USA, 2002. IEEE Computer Society.
- [6] M. Gervautz and D. Schmalstieg. Anywhere interfaces using handheld augmented reality. *Computer*, 45(7):26–31, 2012.
- [7] D. A. Guttentag. Virtual reality: Applications and implications for tourism. *Tourism Management*, 31(5):637–651, 2010.
- [8] M. Halsey. What's new for end users? In *Windows 10 Primer*, pages 49–62. Springer, 2015.
- [9] B. L. Harrison, H. Ishii, K. J. Vicente, and W. A. S. Buxton. Transparent layered user interfaces: An evaluation of a display design to enhance focused and divided attention. In *SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, pages 317–324, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.
- [10] O. Hilliges, D. Kim, S. Izadi, M. Weiss, and A. Wilson. Holodesk: direct 3d interactions with a situated see-through display. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 2421–2430. ACM, 2012.
- [11] F. Kellner, B. Bolte, G. Bruder, U. Rautenberg, F. Steinicke, M. Lappe, and R. Koch. Geometric calibration of head-mounted displays and its effects on distance estimation. *Visualization and Computer Graphics*, 18(4):589–596, 2012.
- [12] S.-H. Kim, L. J. Prinzel, D. B. Kaber, A. L. Alexander, E. M. Stelzer, K. Kaufmann, and T. Veil. Multidimensional measure of display clutter and pilot performance for advanced head-up display. *Aviation, space, and environmental medicine*, 82(11):1013–1022, 2011.
- [13] K. Kiyokawa, Y. Kurata, and H. Ohno. An optical see-through display for mutual occlusion with a real-time stereovision system. *Computers & Graphics*, pages 765–779, 2001.
- [14] F. L. Kooi and A. Toet. Visual comfort of binocular and 3D displays. *Displays*, 25(2–3):99–108, 2004.
- [15] C. D. Kounavis, A. E. Kasimati, E. D. Zamani, and G. Giaglis. Enhancing the tourism experience through mobile augmented reality: Challenges and prospects. *International Journal of Engineering Business Management*, 4(10):1–6, 2012.
- [16] M. Kytö, A. Mäkinen, T. Tossavainen, and P. Oittinen. Stereoscopic depth perception in video see-through augmented reality within action space. *Journal of Electronic Imaging*, 23(1):011006, 2014.
- [17] S. Liu, D. Cheng, and H. Hua. An optical see-through head mounted display with addressable focal planes. In *ISMAR 2008*, pages 33–42, Sept 2008.
- [18] A. Lucero, K. Lyons, A. Vetek, T. Järvenpää, S. White, and M. Salmimaa. Exploring the interaction design space for interactive glasses. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 1341–1346. ACM, 2013.
- [19] V. Mazza, M. Turatto, and C. Umiltà. Foreground-background segmentation and attention: A change blindness study. *Psychological Research*, 69(3):201–210, 2005.
- [20] A. Nee, S. Ong, G. Chryssoulouris, and D. Mourtzis. Augmented reality applications in design and manufacturing. *CIRP Annals-Manufacturing Technology*, 61(2):657–679, 2012.
- [21] S. D. Peterson, M. Axholt, and S. R. Ellis. Objective and subjective assessment of stereoscopically separated labels in augmented reality. *Computers & Graphics*, 33(1):23 – 33, 2009.
- [22] V. T. Phan and S. Y. Choo. Interior design in augmented reality environment. *International Journal of Computer Applications*, 5(5), 2010.
- [23] R. Radkowski and C. Stritzke. Interactive hand gesture-based assembly for augmented reality applications. In *ACHI 2012, The Fifth International Conference on Advances in Computer-Human Interactions*, pages 303–308, 2012.
- [24] E. Redondo, I. Navarro, A. Sánchez Riera, and D. Fonseca. Augmented reality on architectural and building engineering learning processes. two study cases. *UbiCC Journal*, pages 1269–1279, 2012.
- [25] J. P. Rolland and H. Fuchs. Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoper. Virtual Environ.*, 9(3):287–309, June 2000.
- [26] J. P. Rolland, R. L. Holloway, and H. Fuchs. Comparison of optical and video see-through, head-mounted displays. In *Proc. SPIE*, volume 2351, pages 293–307, 1995.
- [27] S. Smith and S.-H. Fu. The relationships between automobile head-up display presentation images and drivers' Kansei. *Displays*, 32(2):58–68, 2011.
- [28] T. Starner. Project glass: An extension of the self. *Pervasive Computing, IEEE*, 12(2):14–16, 2013.
- [29] L. B. Stelmach and W. J. Tam. Stereoscopic image coding: Effect of disparate image-quality in left- and right-eye views. *Signal Processing: Image Communication*, 14(1–2):111–117, 1998.
- [30] C. Sulisz and P. Seeling. An off-the-shelf wearable hud system for support in indoor environments. In *11th International Conference on Mobile and Ubiquitous Multimedia*, page 60. ACM, 2012.
- [31] B. L. William Wong, R. Joyekurun, H. Mansour, P. Amaldi, A. Nees, and R. Villanueva. Depth, layering and transparency: Developing design techniques. In *17th Australia Conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future*, OZCHI '05, pages 1–10, Narrabundah, Australia, Australia, 2005. Computer-Human Interaction Special Interest Group (CHISIG) of Australia.