Legibility of Letters in Reality, 2D and 3D Projection

Elisabeth Dittrich[1], Stefan Brandenburg[2], and Boris Beckmann-Dobrev[3]

[1] Research training group Prometei, Berlin, Germany
elisabeth.dittrich@zmms.tu-berlin.de
[2] University of Technology of Berlin, Chair of cognitive psychology
and cognitive ergonomics, Berlin, Germany
stefan.brandenburg@tu-berlin.de
[3] University of Technology of Berlin,
Chair of Industrial Information Technology, Berlin, Germany
bdobrev@mailbox.tu-berlin.de

Abstract. Virtual prototypes are essential for engineers to understand the complex structures and arrangements of mechatronic products like automobiles. Currently, Virtual Environments (VE) are used for visual analysis and interaction with virtual models. In the next years more supplementary information will be integrated in the VE, completing the 3D-model. This includes names of single parts, corresponding materials or masses. However, up till now there is little explicit research on the psychological effects of additional text visualization in VE's. For example it unclear if it is possible to visualize the textual information like on paper prints or on 2D displays. The current study empirically compares these types of different output mediums to advise rules for visualization of text in 3D Virtual Environments. Results show, that textual information has to be slightly enlarged for the 3D Virtual Environment. In addition, subjects performed better in conditions with projected textual information compared to real text.

Keywords: 2D and 3D text, Virtual Environments, legibility of letters, information visualization.

1 Introduction

Virtual Reality is a growing technology in different areas of science and industry. In the engineering context it is an "[…] effective tool for a number of purposes."[1]. Moreover, in the industrial engineering process Virtual Reality is the key technology to validate mechatronic products in early stages of the engineering lifecycle [2]. Its importance becomes evident in virtual design reviews. Within such a review engineers can interact with a three-dimensional model of the desired product in its original size [3]. There the product reacts like a real one. Hence the simulation of products in an early stage of the engineering process helps to find a lot of mistakes in the construction, the dependency of functions and problems with the arrangement of modules. To get there the product engineering lifecycle starts with the definition of product requirements. These requirements are used to design the initial virtual

R. Shumaker (Ed.): VAMR/HCII 2013, Part I, LNCS 8021, pp. 149–158, 2013.

prototype with geometrical masses and nearly real functions. This virtual prototype can be tested and validated in an immersive virtual environment (VE) [4].

This so called virtual design review is an in-progress validation of the product. It controls the state of the construction and the realization of the requirements [5][6]. Hence this review needs a lot of essential information, for example text-documents, the product-structure, sketches and descriptive statistics [7]. To date each piece of information is displayed on a single resource. Whereas the virtual prototype is mostly realized in an immersive VE, documents, sketches or descriptive statistics are printed or on a separate computer. Therefore participants of a virtual design review constantly need to switch between different sources of information. In addition they cannot interact with the textual information in the design review. For example, viewing and manipulating components of the product-structure inside of the VE would be helpful to access the desired information without interrupting the review. Thus, the visualization of the 3D-model and additional information on separate screens complicates and lengthened the procedure of a virtual design review significantly.

Therefore, research has started to find out how to integrate words in virtual environments (see [8] for an overview).

1.1 Information Rich Virtual Environments (IRVE)

The integration of words in virtual scenes is called Information Rich Virtual Environments (IRVE) [9]. Polys [10] define an IRVE "[…] as a class of visual analysis tools for integrated information spaces [that] start[s] with realistic and perceptual information and enhance it with abstract and temporal information." [10, p.31]. Therefore a more or less complex data structure is visualized as a three-dimensional object and enhanced with additional information, so called annotations. Annotations might be simple labels, graphs, or other multimedia [11].

As Pick [7] point out, it is much of a challenge to include supplementary information to the virtual object under investigation. Several problems come up when visually integrating annotations to visualized objects. For example, unfavorable annotation positions might occlude data or vice versa [7]. Moreover, if annotations are placed in IRVEs they need to be clearly structured to avoid information clutter or confusion by subjects.

Additionally, [11] provides an overview regarding the parameters that affect the quality of textual annotations in IRVEs. These parameters are: color, fonts, size, background and transparency. Polys and Bowman [11] even conclude that the legibility of words is the most important attribute of annotations in IRVEs. Finally Polys [8] explicitly highlights the size of annotations as a major determinant with respect to their legibility.

1.2 Text Legibility

The international organization of standardization (ISO) defines text legibility as the „[…] ability of unambiguous identification of single characters or symbols that may be presented in a non-contextual format.“ [12, p.25]. To ensure legibility a lot of parameters are to be considered. In ISO [13] ten parameters describing among other things the contrast, size, thickness of lines of the letters. One of them is the symbol height. The minimum height for digital symbols is determined at 16 minutes of arc, which is larger than the minimum for paper prints. This difference in minimum character size is justified by the fact that the legibility of characters on digital displays is also influenced by "[…] pixel density or resolution, contrast and character font and matrix, as well as viewing distance." [13, p.12].

However, in an IRVE text legibility is also affected by its resolution. Finally in a virtual environment like a CAVE the viewing distance is dynamic [8], because the position of the viewer is changing while watching the virtual object or text. Therefore this dynamics influences the perception of the virtual objects resolution [11]. Empirical studies showed that reading time of words increased when text was rotated, or the viewing position changed in the IRVE [14], [15]. Additionally this effect is greater for smaller than for larger fonts.

1.3 Perception in (IR)VEs

Stereoscopic view in reality means that the two pictures from the two eyes influence the impression of an object. One of the depth cues is the disparity of these two eye pictures. Regarding the effect of disparity on object perception one can state that the wider the disparity, the closer is the focused object. For simulating this three dimensional impression in virtual environments the flat image of the object is projected twice on the screen and the distance both pictures (= disparity) specifies the perceived depth [16].

However, compared to reality the stereoscopic projection in IRVEs might result in the convergence-accommodation-conflict that is connected to simulator sickness [17] and viewing fatigue [18]. In reality people set their focal point of view directly on the object. We obtain a tree-dimensional image because both eyes set this focal viewpoint slightly different. In contrast in a virtual environment, the projected object has no depth itself and people need to focus on the display. However the computer varies the disparity between the two projections of an object in this way that the focus point of the object lays in front of or behind the screen. This difference between the physical (i.e. the screen) and the virtual (i.e. the object) point of focus leads to the convergence-accommodation conflict which in turn is related to an increase in users strain and a decrease in information acquisition.

2 Research Question

Previous studies investigated a lot of parameters that affect perception in Virtual Environments. In most of them resolution and the depending font size are not explicitly listed, sometimes not mentioned at all. Moreover there are no standards for text legibility in (IR)VEs. For example, [13] defines larger letter sizes for digital (2D) letters compared to print media. In addition, we pointed out that VEs impose special demands on the reader regarding text legibility.

Hence we want to find out, which minimal resolution and font size is needed to ensure legibility in VEs. Based this theoretical background we would assume that if the same letters are visualized in normal 2D projection and in a stereoscopic 3D projection, than subjects' performance with respect to legibility is better in the 2D condition. In other words, letter size needs to be larger in 3D compared to 2D. To test this hypothesis we chose a two-step procedure. First, we exactly accessed subjects' eyesight using a real and a projected standard eyesight examination plate. Second, we tested whether resolution needs to be different for 2D and 3D projection of letters.

3 Empirical User-Study

3.1 Sample

A total number of N = 21 subjects participated in this experiment, 13 (61%) females and 8 (39%) male. Subjects age ranged between 14 and 60 M = 31.5 (SD = 13) years. Most of them (71%) were students, 1 person (5%) was a retiree and another 5 participants (24%) were employees. Almost half of the sample (47%) had corrected to normal vision. All of them wore glasses and no contact lenses. Their mean correction was M = 1.53 (SD = 2.03) diopter. Eight of these 10 participants with corrected to normal vision showed near-sightedness. One participant had color deficiency, one was night-blind and another one had dyslexia.

3.2 Technical Equipment

In the experiment we used a portable powerwall. This powerwall was operated through an active stereo projector (DepthQ HDs3D) with 1280*720 ppi at 120 Hz. The stimulus material was rear-projected on a silver screen that was optimized for 3D environments. The projector setup was also optimized for higher pixel density. This means it was possible to realize 1 mm pixel height and thus a picture of 1280*720 millimeters. Two computers with Nvidia Quadro FX 3800 graphic cards presented the stimulus material. The visualization software for the stereoscopic view was VDP by ICIDO (now ESI-group). The shutter glasses and the shutter emitter were named APG600. The eye chart was designed in NX (Siemens PLM) and modulated with Deep Exploration.

3.3 Experimental Design

The first independent variable was the kind of letter projection, with two manifestations: a) normal 2D and b) stereoscopic 3D projection. All settings were done in the software for 3D visualization. For the 2D condition the disparity was adjusted on zero percent. Therefore the stimulus material and all other conditions were absolutely the same in both conditions.

The second independent variable was the number of pixels per letter (resolution). In both conditions, 2D and 3D, the resolution of letters differed in 6 steps. For each step of this manipulation the virtual distance of the plate was decreased for one meter. The algorithm for the 3D illusion of a rearward moved letter made them smaller and their disparity higher, so the letters seemed to be more behind the screen. In total we moved the letters in six steps rearward, each one resembled a virtual meter. Each time we placed the letter one more virtual meter backwards; the participant went one physical meter forward. Therefore the letter size was constant (see also minute of arc in Table 1) and the amount of pixels for one letter decreased. The following table shows the six steps of reducing the resolution from the best to the worst resolution.

The dependent variable was text legibility, which is dichotomous with the manifestations non legible and legible. Hence a border of legibility for every participant and condition was quantified. This border served as the dependent variable for subsequent analysis.

Table 1. Manipulation of the number of pixel (second column), distance between the participant and the projection screen (3rd column) and (in last column) the post hoc calculated minute of arc of the letters

	number of pixels	distance to screen	post hoc calculated minute of arc
1. best resolution	11	6 m	6,3‘
2.	9	5 m	6,1‘
3.	7	4 m	6,0‘
4.	6	3 m	6,8‘
5.	5	2 m	8,5‘
6. worst resolution	4	1 m	13,7‘

Contrary to our expectations table 1 shows that the virtual and the physical meter were not equivalent, because the depending minutes of arc are not constant. We would consider this as a bug of the virtual visualization software. It remains unclear why the algorithm produced this error.

3.4 Procedure

First of all participants completed the questionnaire assessing demographic variables. Now, subjects completed the pre-test. The goal of this pre-test was to check the sightedness of the participants. Hence all subjects completed a standardized eye examination procedure using a plate (see Fig.1 on the right side). For this eye examination, the standardized procedure is to read each line of letters from top to the bottom of the plate. The line of letters where subjects were not able to read the characters anymore defined the sightedness in percent.

After finishing reading the letters of the physical eye examination plate, participants repeated the same eye examination procedure with a projected plate of letters (see middle of Fig.1). Here they were placed 20 centimeters closer to the projection screen to compensate the different distances of both plates. Every participant accomplished both tasks in the same order and the eyesight in percent was noted for both conditions. However this second trial of the pre-test was to guarantee transferability from the physical to the projected examination plate. Therefore no differences in eyesight were expected between both conditions (pre-test hypotheses).

Fig. 1. Shows the physical standardized eye examination plate on the right hand side and the projected plate on the left hand side

In the experimental part of this study we asked participants again to read out loud one line of letters of the 2D projected examination plate. More in detail, they were requested to read the line of letters that resembled 100% eyesight. Each participant repeated this procedure (i.e. reading the line of letters that resembled 100% eyesight) 16 times, 8 times in the 2D condition and 8 times in the 3D condition. We designed different digital plates, so that letters in the same line were randomized from the whole alphabet. The participants began with the best resolution in the 2D condition. When the subjects spelled the presented letter correctly, the next smaller resolution was shown. In case of an error, the measurement stopped and the experimenter noted the number of pixels. This procedure was repeated eight times. Trials starting from a decreasing number of pixels were alternately presented to trials increasing beginning with an increasing number of pixels. After finishing all 8 trials for the 2D condition, we added disparity and repeated the procedure for the 3D condition.

4 Results

A paired t-test was computed for the two hypotheses, the difference between the real and the projected eye-examination plate and the difference between 2D and 3D stereoscopic view. In addition confidence intervals (CI) and effect sizes (d) after [19] were reported if applicable. Regarding the pre-test question, we assumed that there is no difference in subjects' performance depending on the real and the projected eye chart. In contrast to the hypothesis the statistical test did reveal a medium and significant difference between the real and the projected version of the chart, $t(1,20) = -3.23$, $p = 0.004$, $d = 0.7$. As visualized in figure 2, subjects eyesight increased about $M = 12.85\%$ (95% CI ranging from 4.57% to 21.14%) when changing from a real eye chart to the projected version.

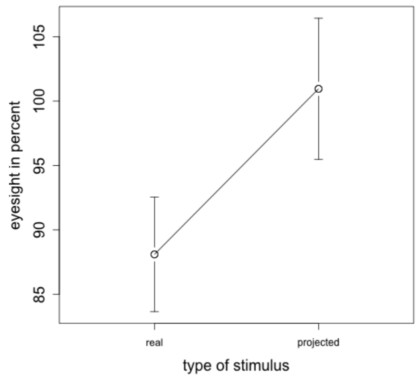

Fig. 2. Subjects' eyesight depending on the type of eye chart used. Bars indicate the 80% confidence intervals.

With respect to the research question, we obtained a tendency for the difference between the 2D and the 3D visualization of letters, $t(1,20) = 1.57$, $p = 0.06$, $d = 0.2$. Figure 3 shows that stimulus size increased about M=0.4 pixels (6.7%).

For practical reasons we would suggest to report the recognition border in whole numbers. Therefore the 2D-legibility-border is M = 5.9 pixels, which resembles 6 pixels height for one letter. The 3D-legibility-border is M = 6.3 pixels which should be rounded to 7 pixel height of a small letter.

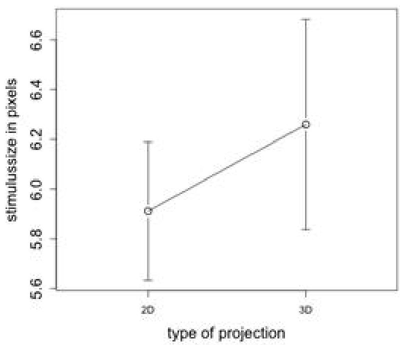

Fig. 3. Subjects' recognition border depending on the type of projection. Bars indicate the 80% confidence intervals.

5 Discussion

Text in Virtual Environments is needed in engineering contexts. For example in virtual design reviews information like measurements of a distance or properties of the designed objects are discussed. If at all, a small fraction of this textual information is mostly arranged near by the object in form of an annotation. However up till now it was not clear how large these annotations need to be. Hence we were interested in the general question how the legibility of the text depends on the resolution and which font size in VEs ensures text legibility. Results show that subjects' performance is better when letters are a little bit larger in stereoscopic projection compared to 2D projection. However, statistics revealed a small tendency and not a very strong effect. Possibly this result might depending on the rather small amount of participants that were used. In addition, it is possible that the difference between 2D and 3D conditions has other reasons. For example, in the 3D condition participants have worn shutter-glasses. This might have altered their performance in the letter-reading task. Moreover we might have elicited a ceiling effect in terms of the readability of letters at all. Hence the difference between 2D and 3D projection would have been larger under different circumstances. Results of related studies undermine this position. For example [20] discovered that a font size of 9 pixels is optimal for the legibility of virtual lowercase letters. In our study, it was around 6 pixels and therefore a lot lower than Sheedys' numbers [20]. Possibly other parameters of text legibility like luminance, contrast or thickness of lines of the letters were very well set in our experiment. Therefore subjects' showed best performance and the difference of type of projection was at its minimum. Besides related work the result of our pre-test also supports this position. In contrast to our hypothesis, subjects' eyesight was better in the 2D projection compared to the real eye-examination plate. On the one hand this result points out the potential for projections to improve peoples' sightedness. On the other hand, we might have shown the importance of comparable testing situations. We invested a great amount of effort regarding the standardization of subjects' eye examination. Of course we tried to balance light and contrast conditions for the use of the real and projected eye examination plate. Again, results indicate that lighting conditions, contrast or other indicators of legibility was better for the projected eye examination compared to the real one. The direction of the difference between real world and projected test-plate contradicts the definition of larger fonts for digital texts compared to print media. However, 2D and 3D projection was extremely comparable since we used the same method and instruments for the visualization of letters for both projections.

Finally practical implications and limitations of our work need to be mentioned. Regarding the practical implications we were lucky to find a small difference between 2D and 3D projection only. Since the amount of information that is used in a virtual design review is usually quite large [7], small fonts enable engineers to include much information in the Virtual Environment. However, this advantage directly relates to the problems of text visualization in Virtual Environments. As stated above, unfavorable annotation positions might occlude data or vice versa [7]. Moreover, since we found that annotations do not need to be large in terms of their size, programmers might be tempted to include many of them in IRVEs. This would directly lead to information clutter or confusion of subjects. Additionally, other parameters of annotation in IRVEs might reduce generalizability of our results. For example if programmers chose unfavorable colors, fonts, backgrounds or transparency settings a small increase in font size might not be enough to ensure text legibility. In line with [15] and [11], we would assume these text legibility-determining factors to interact with each other. With respect to these interactions many open questions are left at the end of this work. Future research should therefore replicate our small effect of the 2D versus 3D comparison with a larger sample. If it is possible to replicate our finding, one could think of expanding the experimental setting to more realistic examples. In the present work, we only used extremely simple stimuli with maximum contrast, black letters on white background. Contrasting this approach, real world mockups are colorful. And so are annotations. On a more complex level of research one could further address questions of occlusion or visual clutter [8].

To sum up, a lot of rules for legibility are accessible for printed documents. However, there are none for stereoscopic texts. In contrast to current practice we would propose to only slightly enlarge (plus one pixel) text size in information rich virtual environments. Hence practitioners would be enabled to include larger amounts of data in their virtual environments. Hence, the present work is important since it shows three things. First, texts only need to be a little larger in (IRVEs) than in 2D. Second, projected text bears the potential to enhance subjects' performance compared to real world stimulus material. And third, we did not obtain any evidence for the convergence-accommodation-conflict [17], [18].

References

1. Wickens, C., Holland, J.: Engineering Psychology and Human Performance. Prentice Hall, New Jersey (2000)
2. Krause, F.-L., Franke, H.-J., Gausemeier, J.: Innovationspotenziale der Produktentwicklung. Carl Hanser Verlag, München (2007)
3. Schmid, D.: Konstruktionslehre Maschinenbau. Verlag Europa- Lehrmittel, Haan-Gruiten (2009)
4. Blümel, E., Straßburger, S., Sturek, R., Kimura, I.: Pragmatic Approach to Apply Virtual Realty Technology in Accelerating a Product Life Cycle. In: Proceedings of the International Conference INNOVATIONS, pp. 199–207 (2004)
5. Kamiske, G.F., Ehrhart, K.J., Jacobi, H.-J., Pfeifer, T., Ritter, A., Zink, K.J.: Bausteine des innovativen Qualitätsmanagement. Hanser, München (1997)
6. Masing, W.: Handbuch Qualitätsmanagement. Hanser, München (1994)
7. Pick, S., Hentschel, B., Wolter, M., Tedjo-Palczynski, I., Kuhlen, T.: Automated positioning of annotations in immersive virtual environments. In: Kuhlen, T. (ed.) Joint Virtual Reality Conference of EGVE-EuroVR-VEC, pp. 1–8. The Eurographics Association (2010)
8. Polys, N.: Display Techniques in Information-Rich Virtual Environments. Doctoral dissertation, Virginia Polytechnic Institute and State University (2006)
9. Bowman, D., Hodges, L., Bolter, J.: The Virtual Venue: User-Computer Interaction in Information-Rich Virtual Environments. Presence: Teleoperators and Virtual Environments 7(5), 478–493 (1998)
10. Polys, N., Bowman, D., North, C.: The role of Depth and Gestalt cues in information-rich virtual environments. International Journal of Human-Computer Studies 69, 30–51 (2011)
11. Polys, N., Bowman, D.: Design and display of enhancing information in desktop information-rich environments: challenges and techniques. Virtual Reality 8, 41–54 (2004)
12. ISO 9241-302: Ergonomics of human-system interaction – Terminology for electronic visual displays. ISO, Geneva (2008)

13. ISO 9241-303: Ergonomics of human-system interaction – Requirements for electronic visual displays. ISO, Geneva (2011)
14. Larson, K., van Dantzich, M., Czerwinski, M., Robertson, G.: Text in 3D: some legibility results. In: CHI 2000 Extended Abstracts on Human Factors in Computing Systems, pp. 145–146. ACM, New York (2000)
15. Grossman, T., Wigdor, D., Balakrishnan, R.: Exploring and reducing the effects of orientation on text readability in volumetric displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 483–492. ACM, New York (2007)
16. Goldstein, B.E.: Wahrnehmungspsychologie. Spekrum, Heidelberg (2001)
17. Häkkinen, J., Pölönen, M., Takatalo, J., Nyman, G.: Simulator sickness in virtual display gaming: a comparison of stereoscopic and non-stereoscopic situations. In: Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services, pp. 227–230. ACM, New York (2006)
18. Stern, A., Javidi, B.: Three dimensional image sensing, visualization and processing using integral imaging. Proceedings of the IEEE 94(3), 591–607 (2006)
19. Cohen, J.: Statistical power analysis for the behavioral sciences, 2nd edn. Erlbaum, New York (1988)
20. Sheedy, J.E., Subbaram, M.V., Zimmerman, A.B., Hayes, J.R.: Text legibility and the letter superiority effect. Human Factors: The Journal of the Human Factors and Ergonomics Society 47(4), 797–815 (2005)