Sylvain Tourancheau, Mårten Sjöström, Roger Olsson, Anders Persson, Thomas Ericson, Johan Rudling and Bengt Norén, "Subjective evaluation of user experience in interactive 3D visualization in a medical context", Proc. SPIE 8318, 831814 (2012);
http://dx.doi.org/10.1117/12.910828

Subjective evaluation of user experience in interactive 3D-visualization in a medical context

Sylvain Tourancheau[a], Mårten Sjöström[a], Roger Olsson[a], Anders Persson[b], Thomas Ericson[c], Johan Rudling[b], Bengt Norén[d]

[a] Dept. of Information Technology and Media, Mid Sweden University, Sundsvall, Sweden
[b] Center for Medical Image Science and Visualization, Linköping University, Sweden
[c] Setred AB, Stockholm, Sweden
[d] Dept. of Radiology, Linköping University Hospital, Sweden

ABSTRACT

New display technologies enable the usage of 3D-visualization in a medical context. Even though user performance seems to be enhanced with respect to 2D thanks to the addition of recreated depth cues, human factors, and more particularly visual comfort and visual fatigue can still be a bridle to the widespread use of these systems. This study aimed at evaluating and comparing two different 3D visualization systems (a market stereoscopic display, and a state-of-the-art multi-view display) in terms of quality of experience (QoE), in the context of interactive medical visualization. An adapted methodology was designed in order to subjectively evaluate the experience of users. 14 medical doctors and 15 medical students took part in the experiment. After solving different tasks using the 3D reconstruction of a phantom object, they were asked to judge their quality of the experience, according to specific features. They were also asked to give their opinion about the influence of 3D-systems on their work conditions. Results suggest that medical doctors are opened to 3D-visualization techniques and are confident concerning their beneficial influence on their work. However, visual comfort and visual fatigue are still an issue of 3D-displays. Results obtained with the multi-view display suggest that the use of continuous horizontal parallax might be the future response to these current limitations.

Keywords: 3D-visualization, quality of experience, task performance, subjective assessment methodology

1. INTRODUCTION

In the last decades, new medical 3D data acquisition devices have been developed in order to provide accurate spatial information of the human body. Despite recent improvements of hardware capabilities and rendering algorithms, 3D reconstructions are not routinely used in most hospitals. This can be explained by the fact that physicians are traditionally trained to extract information from 2D image slices, and because the use of legacy 2D-displays to visualize 3D volumetric images is questionable because of ambiguities in the interpretation.[1] The fast development of new display systems[2] permit to envisage a future widespread use of 3D-visualization. The addition of recreated depth cues gives the viewer a more complete 3D experience and permits to increase user performance, with respect to 2D, in particular tasks and conditions.[2–4] However, another aspect of high importance in 3D visualization is human factors. Poorly depicted 3D can lead to visual discomfort, visual fatigue and even dizziness or headaches.[5] Comfort and quality of experience (QoE) of the physicians are as important as task performance, diagnosis reliability and accuracy when it comes to advise 3D-visualization as a new medical usage.

The aim of this study was to evaluate and compare user comfort and quality of experience when using 3D-visualization systems in a medical context. Two different 3D displays were chosen. The first one is a market stereoscopic 3D-display which provides binocular parallax (stereopsis) by presenting independently two different images to the eyes thanks to active shutter-glasses. The second one is a state-of-the-art time-multiplexed auto-stereoscopic multi-view 3D-display developed for medical usage (it will be named multi-view display in the

Further authors information: (send correspondence to M.S.)
S.T., M.S., and R.O.: [firstname].[lastname]@miun.se (without accent)
A.P.: anders.persson@cmiv.liu.se, T.E.: thomas.ericson@setred.com, B.N.: bengt.noren@lio.se

following for simplicity). It adds motion-parallax (kineopsis) to binocular parallax and provides a look-around capability by displaying multiple perspectives of the 3D scene. Users are not required to wear spectacles to see depth on this display. This evaluation took place in the context of interactive medical visualization, where users had to analyze, to measure and to act upon a scene constituted of various CT scans of a phantom object. 29 test persons took part in the experiment, 14 experimented radiologists and 15 medicine students. Their performances were measured and the user quality of experience was evaluated at the end of the experiment.

The rest of this document is organized as follows: Section 2 describes the methodology of the experiment, and more precisely the test conditions, the subjects, and the protocol. Section 3 presents and analyzes the results and discusses the differences between the two visualization systems under test. Finally, Section 4 summarizes the conclusions of this work.

2. METHODOLOGY

2.1 Experimental set-up

2.1.1 3D visualization systems

Two different visualization systems were used in this study. The first one is a time-sequential stereoscopic liquid crystal display (LCD) together with the NVIDIA 3D Vision^{TM} system.[6] This system presents left and right images sequentially to the observer thanks to liquid crystal shutter-glasses synchronized with the display via infra-red communication. The second visualization system consists in a time-multiplexed auto-stereoscopic multi-view 3D display using a state-of-the-art scanning slit system.[7,8] Contrary to spatially-multiplexed auto-stereoscopic displays (e.g. using lenticular sheet or parallax barrier in front of the pixels panel) which perform a sampling of the 3D-scene at specific discrete angular positions, this display realizes a uniform sampling of the 3D scene in both spatial and angular dimensions, which permits to present the 3D scene under the form of a *light-field* with a "continuous" motion parallax rather than a set of views. For this experiment, we chose a parametrization with a viewing zone of 30 cm at a viewing distance of 60 cm.

The presentation of 3D scenes was realized thanks to a 3D visualization software dedicated to medical applications. This software performs a real-time rendering of standard CT (computed tomography) medical data sets. The 3D reconstructed volume can then be displayed on one or the other of the two 3D-displays under test.

2.1.2 Displays settings and calibration

Both visualization systems were used in their best configuration in order to compare them in usual work conditions. The stereoscopic display was set to a resolution of 1680×1050 pixels with a refresh rate of 120 Hz (i.e. 60 Hz per eye). The multi-view display was configured in order to get a viewing zone of 30 cm at the chosen viewing distance. The two different 3D-display technologies were responsible for some important differences in viewing conditions. For example, 37 "perspectives" were computed in order to render one 3D frame on the multi-view display while only two views are necessary to display one 3D frame on the stereoscopic display. Since the same rendering engine with the same computer resources were used in both cases, the frame rate was much lower on the multi-view display. Another important difference concerns the peak luminance which was only 21.6 cd/m^2 on the stereoscopic display while it was up to 69.7 cd/m^2 on the multi-view display.

The viewing distance was set to 600 mm, which has been found to be the average distance used at work by the radiologists who participated to this study. At this distance, the visual resolution was 40.1 pixels per degree on the stereoscopic display and 25 pixels per degree on the multi-view display. In these conditions, the presentation subtended $40.9^{\circ} \times 30.7^{\circ}$ on the multi-view display and $41.9^{\circ} \times 26.2^{\circ}$ on the stereoscopic display

Table 1 summarizes displays parameters and viewing conditions for each system.

2.1.3 Test room

The experiment was carried out at the Center for Medical Image Science and Visualization (CMIV) in Linköping University Hospital. Tests took place in a standard workroom as used by radiologists, in which lighting and layout were kept unchanged during the whole time of the experiment. The lighting intensity was adjusted in order to optimize both displays contrast and reading comfort.

	Stereoscopic	Multi-view
Presentation format (pixels)	1680×1050	1024×768
Presentation format (mm)	446×279	440×330
Frame rate (optimal)	60	6.7
Spatial resolution (pixels/deg)	40.1	25
Viewing distance (mm)	600	600
Width of the viewing zone (mm)	-	300
View resolution (views/deg)	-	4.2
Black luminance (cd/m^2)	0.03	0.05
White luminance (cd/m^2)	21.6	69.7

Table 1: Display parameters and viewing conditions.

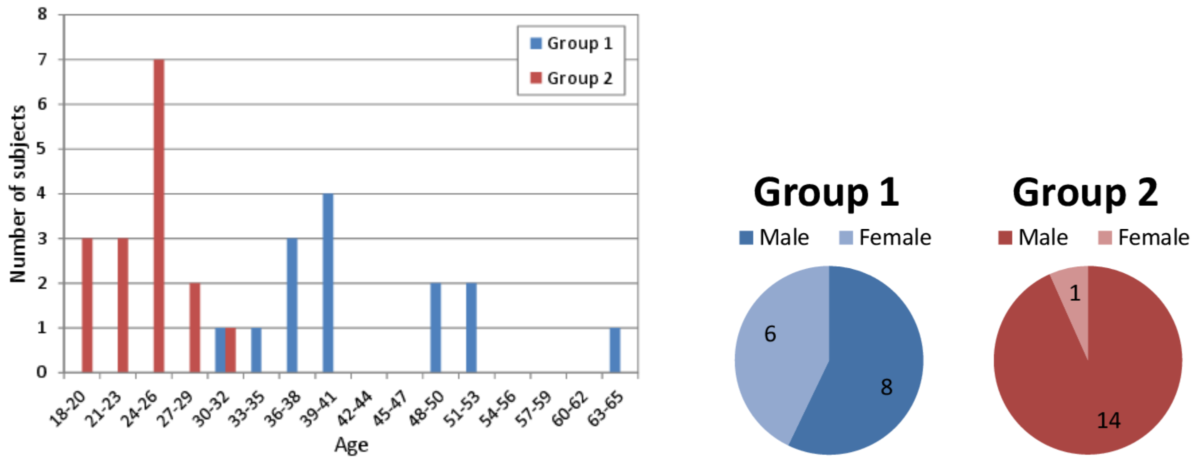


Figure 1: Age distribution and gender proportion of groups of subjects.

2.2 Test Subjects

2.2.1 Groups and sessions

14 radiologists (group 1) and 15 medicine students (group 2) from Linköping University Hospital took part in this experiment. Figure 1 presents the age and gender distributions of the subjects from both groups.

Each test subject did the experiment twice on two separated sessions, one for each visualization system. Half of each group did the experiment on the stereoscopic display first. The other half did the experiment on the multi-view display first. On average, there were 9.4 days between the two sessions (minimum: 6, maximum: 18, median: 8).

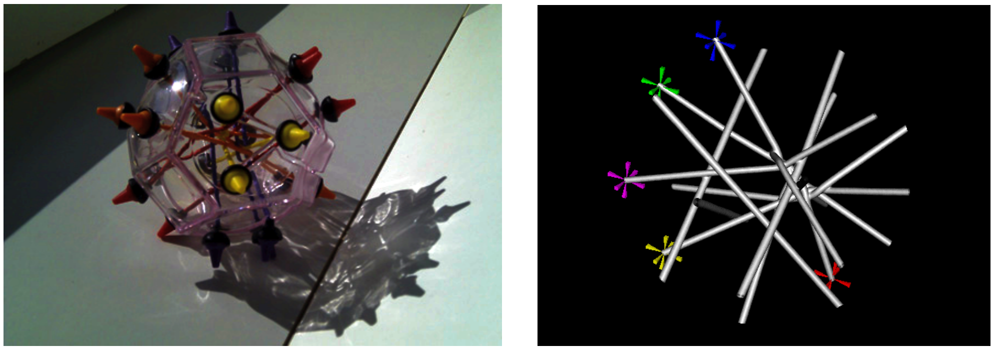
Prior to the experiment, during the first session, the vision of each subject was checked and they were asked to fill a form in order to get information concerning their experience of 3D visualization.

2.2.2 Vision check-up and questionnaire

The binocular threshold parallax of subjects were measured using to the TNO stereopsis test. Vision correction and color blindness, if any, were recorded too.

Prior to the first test session, subjects were asked to fill a questionnaire about their background concerning medical visualization and 3D visualization:

- Number of hours per week using medical visualization software?
- Number of hours per week visualizing 3D volumes?



(a) (b)

Figure 2: Puzzle-game used as a phantom (a), 3D reconstruction of the inside from CT (b).

- Number of hours per week visualizing 3D volumes on a 3D-display?
- Do you have a 3DTV at home?
- Number of 3D movies seen in theaters?
- Are you using other 3D software (other than medical visualization) ?

2.3 Test protocol

The goal of this experiment was to evaluate the influence of 3D on medical visualization in terms of task performance and user quality of experience. A medical 3D-visualization software specially designed to display 3D volumes rendered from computed tomography data sets was used (cf. 2.1.1). This software is dedicated to medical analysis and diagnosis and it provides a lot of tools to interact with the 3D volumes. In this experiment, subjects interactions was limited to rotation of the volume, and zoom-in/zoom-out. Concerning tools provided by the software, subjects were only allowed to use the tool which permits to place a marker at a specific location of the scene.

Prior to the experiment, subjects were asked to solve a training test, which had been designed in order to teach team how to use the software in the particular case of this experiment.

2.3.1 Test object

It was decided to use a real artificial object as a phantom whom geometry was known, so that the ground truth was perfectly defined. Our choice was directed towards a puzzle-game which consists in a plastic polyhedron traversed by a dozen of elastic strings the extremities of which can be moved in order to create specific patterns (Figure 2a). This phantom was scanned by Computed Tomography resulting in 475 slices with a 512×512 pixels resolution. The 3D reconstruction (Figure 2b) of the object is performed by the visualization software at a resolution adapted to the visualization system under test (cf. Table 1).

Two different configurations of the puzzle-game had been scanned. In the first one, the twelve elastic strings were positioned at the initial location, with no entanglements between them. In the second configuration, strings had been moved and tangled in order to create some "knots" inside the polyhedron.

Prior to the experiment, some colored markers had been added to the scene in order to facilitate instructions to the subjects (by identifying specific strings for example, cf. Figure 2b).

Task #	Instructions	Interaction
1	How many strings are in the scene?	No
2	How many strings are completely in front of the plane defined by the four red markers? (without intersecting the plane)	No
3	Which vertical string is the farthest from you? (right or left)	No
4	Which horizontal string is the closest to you? (top or bottom)	No
5	Which string is the closest to you?	No
6	Which string is the farthest from you?	No
7	How many strings are in the scene?	Yes
8	Place a marker on string(s) which do not touch any other string.	Yes
9	Find the place where the two marked strings are closest to each other. Place a marker halfway between them at this place.	Yes
10	Place a marker where the two marked strings are in contact.	Yes
11	Find the place where the two marked strings are closest to each other. Place a marker halfway between them at this place.	Yes
12	Place a marker at the other end of the marked string.	Yes
13	Estimate the distance between the two red markers.	Yes
14	Place a marker at the other end of the marked string.	Yes
15	Estimate the distance between the two red markers.	Yes

Table 2: Description of the 15 tasks.

	Static	Dynamic
Scene understanding	1, 2	7-12, 14
(of which assessed by marker location)		9-11
Distance estimation		13, 15
Depth perception	2, 3-6	
(of which with occlusions)	3-4	

Table 3: Classification of the tasks.

2.3.2 Tasks

The experiment consisted in 15 different tasks designed to be similar to those usually performed by radiologists in a clinical context. They are described in Table 2, in a chronological order. During tasks 1 to 6 interaction with the volume was not allowed to solve the task (static tasks), interaction using the tools provided by the software was allowed in tasks 7 to 15 (dynamic tasks). Tasks can be classified in three groups regarding the actions which were involved: object understanding, distance estimation, and depth perception. The latter can be done only for static tasks, since moving the scene would change depth order. Table 5 gives the classification of each task according to these factors.

2.3.3 Procedure

Training

The training session was designed with a growing difficulty in order to teach subject how to use the visualization software, particularly how to rotate the ball in the three dimensions using the click-and-drag mouse feature, and how to place a marker at a specific location of the scene. Stimuli used in the training session consisted of markers and simple geometric objects such as squares and circles. The phantom was not used in the training session in order to ensure that every subject starts the experiment with no prior knowledge of it.

Recordings

Solving time was recording independently for each task. Subjects were ask to proceed as follows: read the

Score	Adjective
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

(a) 5-grade subjective quality scale (questions Q1 to Q3).

Score	Preference
+2	Much better
+1	Better
0	Similar
-1	Worst
-2	Much worst

(b) 5-grade subjective preference scale (question P1).

Score	Comparison
2	Significantly better
1	Slightly better
0	Similar
-1	Worse

(c) Comparison scale for questions C1 to C3.

Score	Comparison
1	Better
0	Similar
-1	Slightly worse
-2	Significantly worse

(d) Comparison scale for question C4.

Table 4: Subjective scales used to assess quality of experience.

instruction first, start the timer, solve the task, stop the timer, and finally write down their answer when necessary.

Result takes the form of a binary variable (correct or not) for most of the tasks, except for tasks 9, 10, and 11, for which the euclidean distance between the position of the marker set by the subjects and the correct position was computed, and for tasks 13 and 15, for which the estimation of the measure was recorded.

Questionnaire
At the end of the experiment, subjects were asked to fill a questionnaire in order to evaluate the user quality of experience.

Four questions were asked within this questionnaire:

- Q1: How would you rate the 3D experience delivered by the display? ...
- Q2: How would you rate the visual quality of the display?
- Q3: How would you rate your visual comfort during the experience?
- Q4: Did you experience visual fatigue (eye strain, headache, etc.) at the end of the experiment? (if yes, please comments)

For the three first questions, subjects provided a score according to a 5-grade quality scale (Figure 4a). A yes-no answer was expected for the last question and subjects were asked to describe their sensation if they did experience visual fatigue.

In a second questionnaire, subjects were asked to compare the usage of the visualization system under test with their current systems in terms of:

- C1: Understanding of real clinical cases
- C2: Confidence in their decisions
- C3: Performance (decision time, accuracy, reliability, etc.)

- C4: Fatigue (eye strain, headache, etc.) at the end of the day

The comparison scale in Table 4c was used for questions C1 to C3 and the comparison scale in Table 4d for question C4.

At the end of the second session, subjects were also asked to give their preference between the two visualization systems under test (stereoscopic and multi-view). The question was the following:

- P1: How would you qualify the multi-view display with respect to the stereoscopic display, regarding:
  - 3D experience (depth, realism, etc.)
  - Visual quality
  - Visual comfort
  - Globally

The 5-grade preference scale presented in Table 4b was used to answer these four questions.

3. RESULTS ANALYSIS

3.1 Task performance

3.1.1 Solving times

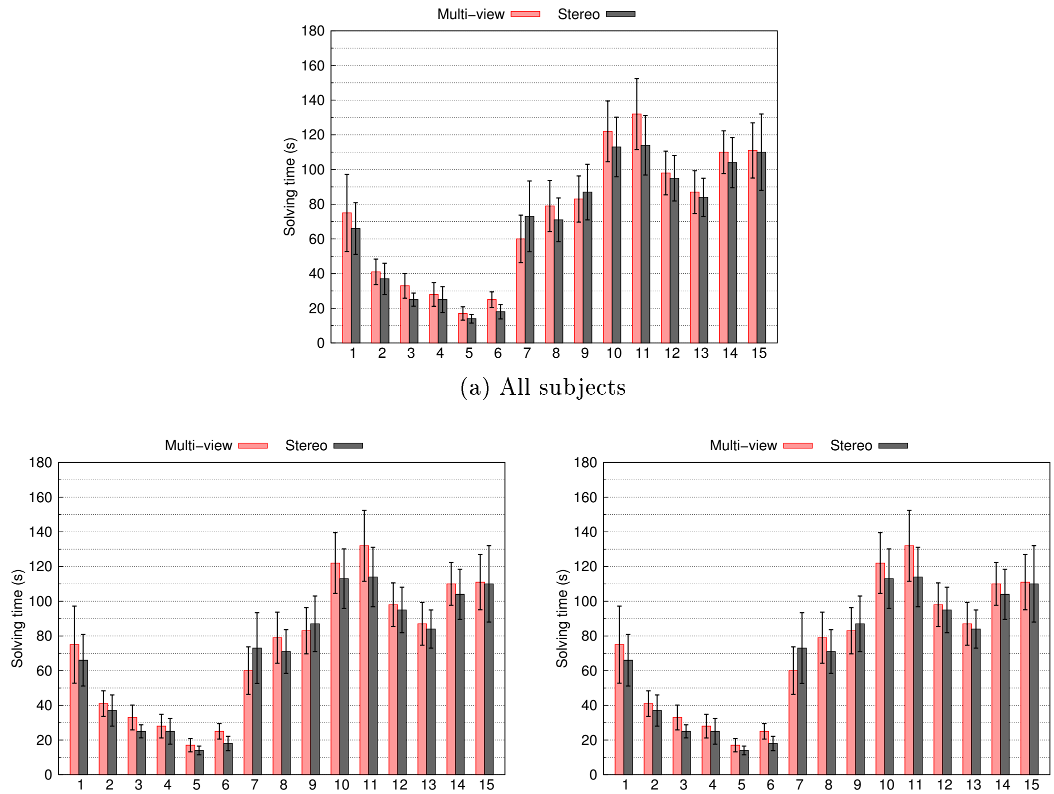
Figure 3 presents a comparison of tasks solving times between the multi-view display and the stereoscopic display, for all subjects, and separately for doctors and for students. Solving times are presented as the mean values with 95% confidence intervals.

There are no significant differences in solving times between both display systems under test. Furthermore, students and doctors have spent in average the same amount of time to solve tasks. When compared on the same display, significant differences between doctors and students were found for only three tasks over fifteen: students were faster on task 3 with the multi-view display, they were faster on tasks 2 and 14 on the stereoscopic display.

3.1.2 Tasks responses

Results of the tasks are presented in Figure 4 for each category of tasks (cf. Table 3). Results are compared between both visualization systems, for all subjects and separately for doctors and students. Results of the depth perception tasks (tasks 2 to 6) are presented as the percentage of correct responses in Figure 4a. Results of the precision tasks (tasks 9 to 11) are presented as the averaged distance between the marker placed by the test subjects and the correct position, in Figure 4b. The average of the results of tasks 1, 7, 8, 12, and 14, during which test subjects had to figure out the 3D scene in order to understand the position of strings between them, are presented in Figure 4c as the percentage of correct responses. Finally, results of the precision tasks (tasks 13 and 15), as the relative error (in percentage) between subjects estimation and ground truth, are presented in Figure 4d.

Concerning the tasks' responses, no significant differences are observed between both 3D-visualization systems. Overall results show that depth perception tasks were solved by more than 90% of the subjects in average, on both displays. Similarly, scene understanding tasks were solved by 80% of the subjects in average with no significant differences between displays or subjects' group. Averaged results of precision tasks are also very similar on both displays, with an averaged distance to target of about 2.5 mm. Finally, distance estimation was globally overestimated by 10% on the multi-view display and by 12% on the stereoscopic display.



(a) All subjects

(b) Doctors

(c) Students

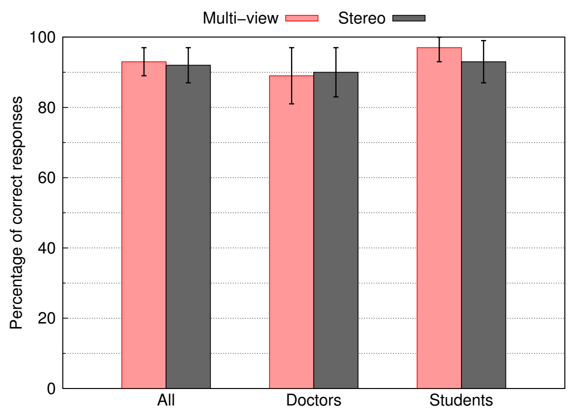
Figure 3: Comparison of task solving times between the multi-view display and the stereoscopic display, for all subjects (a), and for doctors (b) and students (c) separately. Error bars are 95% confidence intervals.

3.2 User experience

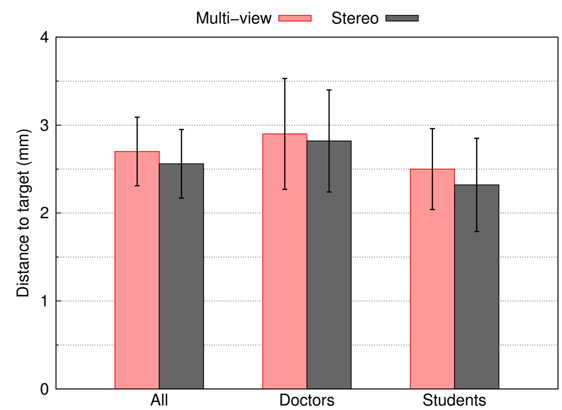
3.2.1 Quality evaluation

The results of the subjective evaluation of the quality of experience (questions Q1 to Q4) are presented and compared between both visualization systems in Figure 5. The mean opinion scores (MOS) were found equivalent (no statistically significant differences) between both displays for the 3D experience, with an average score of "Good". Visual quality was found significantly better (p=0.0002) on the stereoscopic display than on the multi-view display. This can be explained by the fact that both the spatial resolution and the frame rate were higher on the stereoscopic display. Visual comfort was found quite low (between "Fair" and "Good") for both systems, with no significant difference (p=0.2252). However, a difference in terms of visual fatigue was observed (p=0.089) in favor of the multi-view system: 45% of the subjects experienced visual fatigue at the end of the experiment on the stereoscopic display against only 20% on the multi-view display. It is noteworthy that this difference in terms of visual fatigue does not have more impact on the visual comfort. It may be assumed from this observation that both visual quality and visual fatigue equally contribute to the final visual comfort experienced by the users.

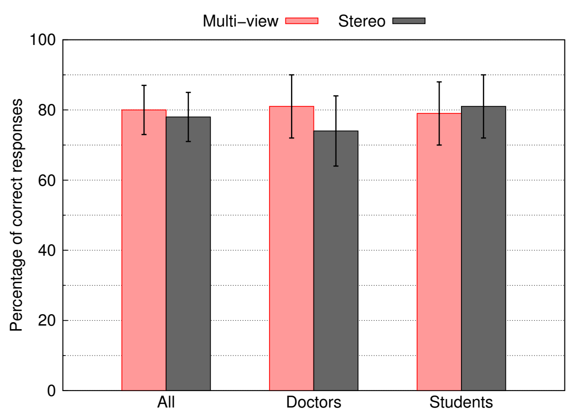
Figure 6 presents the MOS of test subjects when comparing the quality of the auto-stereoscopic multi-view 3D-display to the quality of the time-sequential stereoscopic LCD with active shutter-glasses, in terms of 3D experience, visual quality, visual comfort, and from global point of view. Results here follow the same trend as those obtained in the subjective evaluation of user quality of experience. That is to say, both visualization systems were found to be equivalent in terms of 3D experience (depth, realism, etc.) and the auto-stereoscopic multi-view display has been found to be worse in terms of picture quality, which again can probably be explained by the differences in terms of resolution and frame rate. Concerning the visual comfort, the auto-stereoscopic multi-view 3D-display was considered significantly better than the shutter-glasses stereoscopic display. This



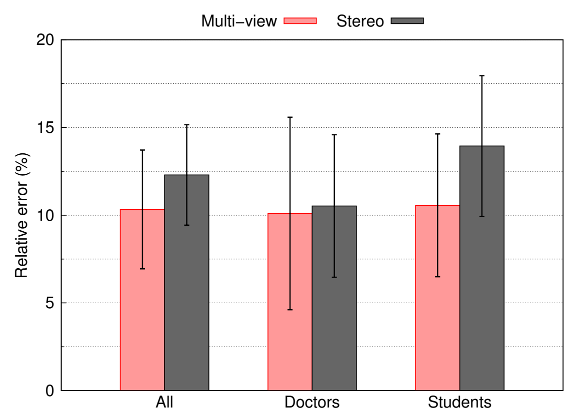
(a) Depth perception



(b) Precision (marker placement)

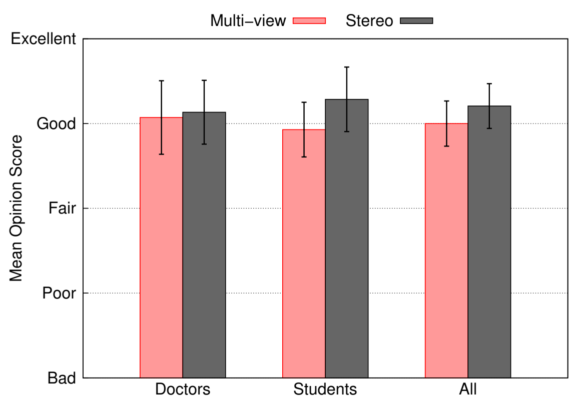


(c) Scene understanding.

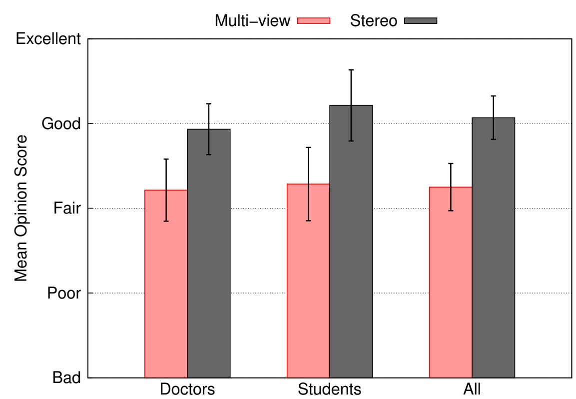


(d) Distance estimation

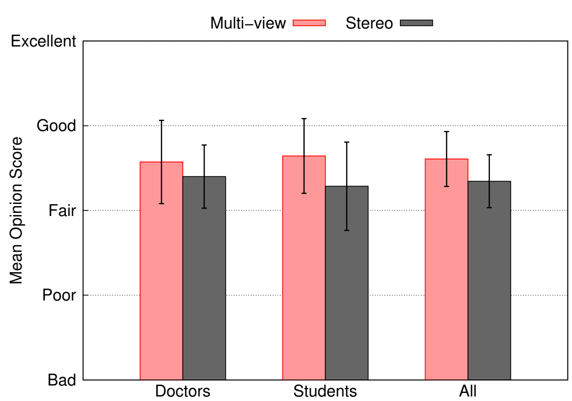
Figure 4: Task performance averaged over all tasks for the four different types of tasks. Error bars are 95% confidence intervals.



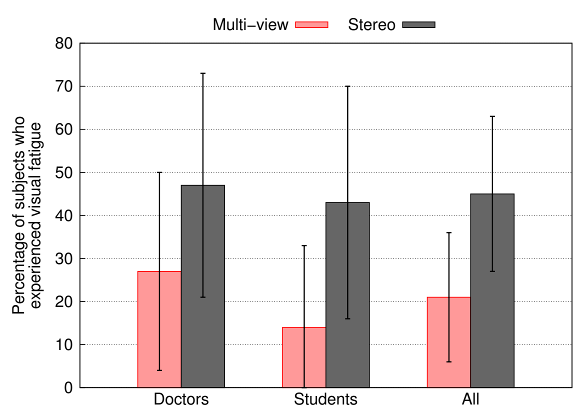
(a) 3D experience.



(b) Visual quality.



(c) Visual comfort.



(d) Visual fatigue.

Figure 5: Results of the subjective evaluation of the quality of experience. Comparison between both 3D-visualization systems. Error bars are 95% confidence intervals.

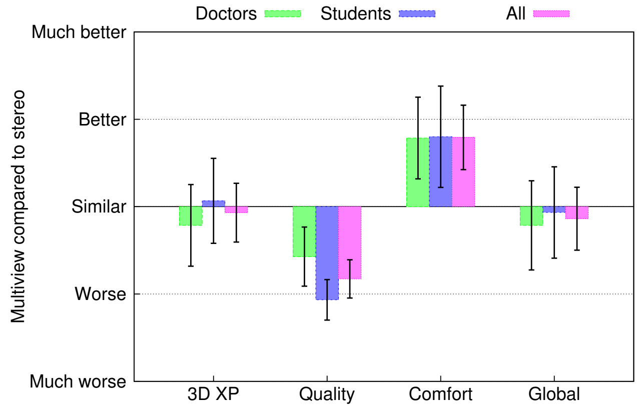


Figure 6: Subjective quality of the auto-stereoscopic multi-view 3D-display compared to the shutter-glasses stereoscopic display, in terms of 3D experience, visual quality, visual comfort and from a global point of view. Error bars are 95% confidence intervals.

result is in accordance with the previous results concerning visual fatigue. From a global point of view, both 3D-visualization systems were found similar, which suggests that the lowest visual quality (in terms of spatial resolution and display frame rate) of the multi-view display was "compensated" with a better visual comfort.

3.2.2 Comparison with legacy 2D-systems

In the second part of the subjective evaluation, subjects were asked to state if the 3D-visualization systems under test would make their work better, similar or worse, according to four specific features: understanding of the case, confidence in the decision, performance, fatigue (cf. Section 2.3.3). The multi-view 3D-display (Figure 7) and the stereoscopic 3D-display (Figure 8) were compared to legacy 2D-visualization systems. Results were very similar for both systems when speaking about the understanding of real clinical cases, the confidence in the decisions, and the performance. Both doctors and students thought in majority that these three aspects would be improved by the use of a 3D visualization system (compared to legacy 2D systems). When it comes to visual fatigue though, the same antagonism as previously was observed: 69% of the subjects judged that their fatigue at the end of a work day would be similar or better while using the auto-stereoscopic multi-view display; on the contrary 62% of them thought that it would be worse while using the stereoscopic display. These results seem to suggest that medical doctors are opened to 3D-visualization techniques and think that they could eventually have a beneficial influence on their work conditions and performance. This is a very important step towards the acceptance by professionals of these new technologies. However, visual comfort and visual fatigue are still major issues of 3D-displays,[5,9] but the results obtained with the automultiscopic display suggest that the use of continuous horizontal parallax might be the future response to these current limitations.

3.3 Influence of user-centered factors

The influence of user-centered factors on the performance and quality of experience has been investigated, but no particular relationship has been found.

The prior experience in medical visualization (stressed by the two different groups of subjects: radiologists where used to visualization of 3D-volumes while most of the students did not have any experience of it) did not affect the results of the tasks nor the solving times as it could have been expected. This might be due to the fact that students, unexperienced but younger, compensate this lack of experience with some better interaction skills for most of them are used to use a computer since their childhood.

Subjects' binocular parallax threshold did not seem to affect their performance nor their quality of experience, differently to what has been reported in a previous study.[10]

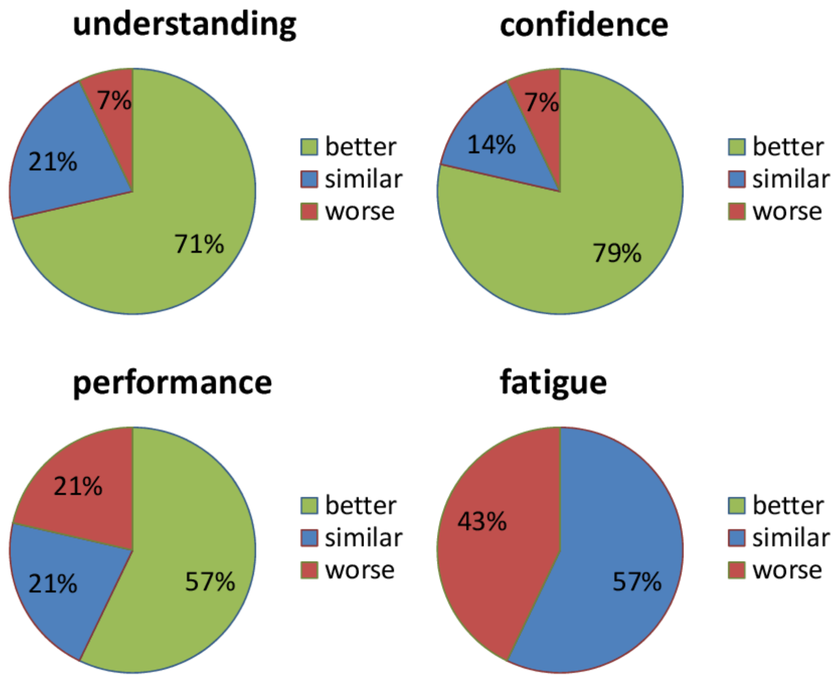


Figure 7: Radiologists' self-evaluation of their user experience. Evaluation of the auto-stereoscopic multi-view display according to four features (see text) with respect to currently used visualization systems.

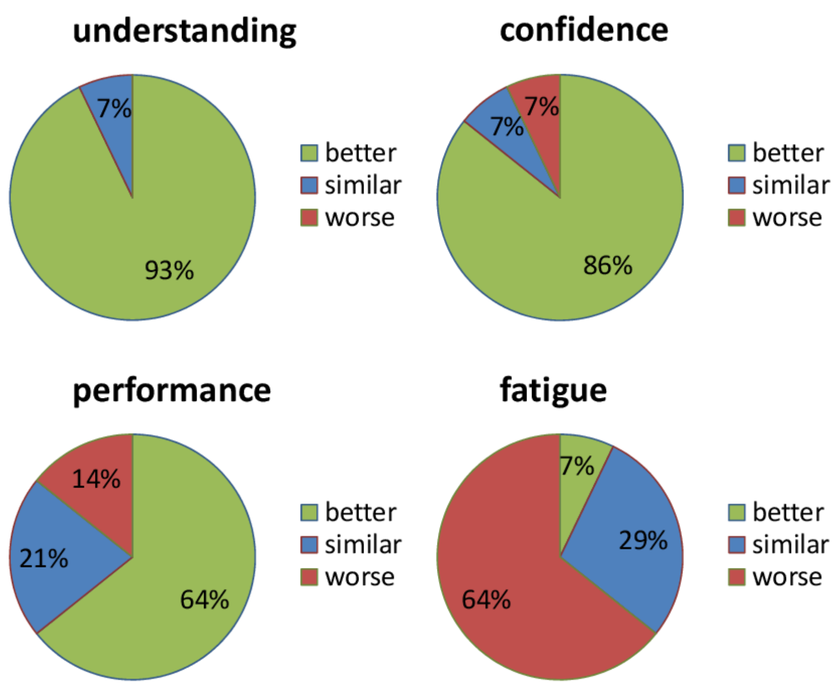


Figure 8: Radiologists' self-evaluation of their user experience. Evaluation of the stereoscopic display according to four features (see text) with respect to currently used visualization systems.

4. CONCLUSION

This study aimed to compare task performance and quality of experience on two different 3D-visualization systems, in a medical 3D-visualization context. 29 test subjects (radiologists and medicine students) took part in a subjective experiment in which they were asked to solve some tasks in a 3D scene, on both displays. Results demonstrates that for these 15 tasks, performance was similar on both visualization systems. No significant differences were found in accuracy, correctness, or time.

An evaluation of user experience was conducted at the end of each test session. 3D experience delivered by both systems was found similar, but picture quality was judged significantly better on the stereoscopic display, while significantly more test subjects reported visual fatigue after having used this display, than after having used the multi-view display. When asked to compare directly both visualization systems, test subjects did not express any preference from a general point of view, but in terms of comfort the multi-view display was significantly preferred. According to previous studies on that aspect,[5,9] visual comfort might be a key feature in the acceptation of 3D-displays' usage. Therefore, it is interesting to observe that for an equivalent user performance the auto-stereoscopic multi-view was preferred over the shutter-glasses stereoscopic display in terms of comfort.

In a second evaluation, radiologists were asked to compare these two 3D-systems with their current 2D visualization systems. Results suggest that medical doctors would agree to work with 3D-displays, and are confident concerning the beneficial influence of 3D-visualization on their work (better confidence in their decision, better understanding of the clinical cases, and better performance in general). However visual comfort and visual fatigue are still issues of 3D-displays, but results obtained with the auto-stereoscopic multi-view display suggest that the use of continuous horizontal parallax might be the future response to these current limitations.

5. ACKNOWLEDGMENT

This work has been supported by grant 2007/0390 of the Swedish Visualization programme (funded by the Swedish Knowledge Foundation, the Vårdal Foundation, Vinnova, Swedish Foundation for Strategic Research, and Invest in Sweden Agency), by grant 00156702 of the EU European Regional Development Fund, Mellersta Norrland, Sweden, and by grant 00155148 of Länsstyrelsen Västernorrland, Sweden

REFERENCES

[1] M. A. Kersten, A. J. Stewart, N. Troje, and R. Ellis, "Enhancing depth perception in translucent volumes," *IEEE Transactions on Visualization and Computer Graphics Journal* **12**(6), pp. 1117–1123, 2006.

[2] M. Agus, F. Bettio, E. Gobbetti, and G. Pintore, "Medical visualization with new generation spatial 3d displays," in *Eurographics Italian Chapter Conference*, R. D. Amicis and G. Conti, eds., 2007.

[3] M. StJohn, M. Cowen, H. Smallman, and H. Oonk, "The use of 2d and 3d displays for shape-understanding versus relative-position tasks," *Human Factors* **43**(1), pp. 79–98, 2001.

[4] M. Agus, E. Gobbetti, J. A. I. Guitián, and F. Marton, "Evaluating layout discrimination capabilities of continuous and discrete automultiscopic displays," in *3DPVT Conference*, (100), 2010.

[5] M. Lambooij and W. Ijsselsteijn, "Visual discomfort and visual fatigue of stereoscopic displays: A review," *Journal of Imaging Science and Technology* **53**(3), 2009.

[6] NVIDIA, "Nvidia 3D Vision™ system - Overview." http://www.nvidia.com/object/3d-vision-main.html.

[7] O. Cossairt, C. Møller, A. L. Travis, and S. A. Benton, "Novel view sequential display based on DMD™ technology," in *Stereoscopic Displays and Virtual Reality Systems XI*, *Proceedings of SPIE-IS&T Electronic Imaging* **5291**, 2004.

[8] C. Møller, *Scanning Slit 3D Displays*. PhD thesis, University of Cambridge, 2005.

[9] W. A. Ijsselsteijn, P. J. Seuntiëns, and L. M. Meesters, *Human Factors of 3D Displays*, ch. 12, pp. 217–233. John Wiley & Sons, Ltd, 2006.

[10] S. Tourancheau, M. Sjöström, R. Olsson, A. Persson, and T. Ericson, "Evaluation of quality of experience in interactive 3d visualization: methodology and results," in *Stereoscopic Displays and Applications XXIII, Proc. of SPIE*, 2012.