Digitalization versus immersion: performance and subjective evaluation of 3D perception with emulated accommodation and parallax in digital microsurgery

Siegfried Wahl
Denitsa Dragneva
Katharina Rifai

SPIE.

Digitalization versus immersion: performance and subjective evaluation of 3D perception with emulated accommodation and parallax in digital microsurgery

Siegfried Wahl,[a,b] **Denitsa Dragneva,**[a] **and Katharina Rifai**[a,b,*]
[a]University Eye Clinics, Institute for Ophthalmic Research, Tuebingen, Germany
[b]Carl Zeiss Vision International GmbH, Aalen, Germany

Abstract. In the visually challenging situation of microsurgery with many altered depth cues, digitalization of surgical systems disrupts two further depth cues, namely focus and parallax. Although in purely optical surgical systems accommodation and eye movements induce expected focus and parallax changes, they become statically fixed through digitalization. Our study evaluates the impact of static focus and parallax onto performance and subjective 3D perception. Subjects reported decreased depth realism under static parallax and focus. Thus surgeons' depth perception is impacted further through digitalization of microsurgery, increasing the potential of artificial stereo-induced fatigue. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JBO.24.10.106501]

Keywords: microsurgery; depth perception; digitalization; stereo images.

Paper 190163RR received May 20, 2019; accepted for publication Sep. 12, 2019; published online Oct. 15, 2019.

1 Introduction

The work of a surgeon is one of the most challenging ones in today's professional world: high-cognitive load combines with high-acuity manual work over a long period of time. Technological advances have shaped the surgery room to a highly specialized surrounding, impacting the surgeon's perception as well as action in performed surgery. Specifically, high-precision surgery has advanced to a new level by the application of stereo surgical microscopes. Nonetheless, stereo microscopes impose challenges on the visual system, specifically in depth perception. Depth perception is known to make use of a variety of available information, called depth cues.[1,2] The most obvious, and in many cases dominant information, is the relative differences in the images from the two eyes, called binocular disparity.[3] In addition, properties of the external world, containing information of the spatial arrangement of objects in space, such as shading, familiar size, occlusion, spatial frequency statistics, and others have been shown to contribute as well.[4] Even the consequences of the optical transformation of the image onto the retina, by the optics of the eye, are used as depth cues: focus cues of blur and depth of focus.[5–7] Stereo vision in microscopes differs from natural vision in several aspects, and these differences affect a variety of depth cues. Differences depend on the optical design of the microscope, and in this article, common main objective (CMO) microscopes are considered. In stereo microscopes, the magnification of an observed object changes depth of focus compared to the unmagnified image and affects contrast content. Both image properties are known depth cues.[5,7–9] Furthermore, viewing geometry, accommodation distance, and disparity are set by the optical design of the stereo microscope. CMO microscopes are designed for a parallel gaze,

and focus planes are often set to infinity.[10] In CMO microscopes, image projection geometry is disentangled from gaze geometry. Projection geometry is typically designed to have a viewing angle of typically 10 to 12 deg, leading to a disparity-based perceived distance of approximately arm distance.[10] Thus in stereo microscopes, many of the three-dimensional (3D) cues differ from natural 3D vision in a purely optical configuration, leading to fatigue and visual strain.[11] Fatigue induced by artificial stereo images is considered a consequence of the efficiency of the stereo system; even largely inaccurate stereo cues still lead to 3D perception.[12] Evidently, 3D perception, despite erroneous image input, comes with the cost of an increased processing load, leading to fatigue.[9,13–22] Recent advances in digitalization in the medical sector show a new disruptive technology on the horizon of microsurgery: digital surgical microscopy. With its numerous advantages in data storage, data analysis, and augmentation, it has the potential to bring microsurgery to a new level of performance. Nonetheless, it challenges the surgeon's perceptual system by presenting errant focus and parallax cues. Visual information reaching an observer's retina in natural vision contains the fingerprint of the eye's status in both cues: the observer's accommodation shifts the focus plane, and eye and head movements induce parallax cues.[5,23,24] Whereas in an optical microscope, accommodation of the surgeon's eye leads to a corresponding shift of the focus plane, in a digital stereo system the focus plane is predefined. Thus a change in accommodation is not accompanied with a corresponding change of the focus plane in the image any more. Furthermore, image parallax through either small head movements in front of an eyepiece, or even through eye movements, is disabled. The recording geometry is statically fixed in a standard digital stereo geometry. Thus additional demands are imposed onto the visual 3D system through digitalization. This study intends to evaluate the impact of disabling motion parallax cues as well as the dynamics of focus cues through digitalization in a digital stereo

---

*Address all correspondence to Katharina Rifai, E-mail: katharina.rifai@medizin.uni-tuebingen.de

microscope setup. A thorough evaluation of the two aspects requires a constancy of all other image and stereo properties. But, when comparing purely optical and digital surgical microscope systems, digitalization imposes additional image degradations: pixilation-induced reduction of resolution and limited contrast, as well as a changed color space. Thus in this study, these degradations were consistent across all conditions evaluated and the impact of focus and parallax cues were analyzed entirely in one identical digital system. To do so, a dynamic stereo microscope (DSM) of the CMO type is developed, which is able to mimic eye-induced parallax and focus changes in real time.[25] In this DSM, eye tracking is exploited to estimate the momentarily viewed depth plane as well as to estimate viewing geometry, while image content is adjusted accordingly.[26] Focus plane adjustments are implemented as focus-adjustable (Optotune AG, Dietikon, Switzerland) lens elements, and image projection changes are implemented through a moving aperture.[25]

With this DSM, behavioral performance as well as subjective 3D immersion are assessed in a standard static stereo arrangement (STAT), a focus (FOC), and a parallax (PAR) condition. Furthermore, a COMB (combined) condition estimates the cumulative effect of focus and parallax.

2 Materials and Methods

2.1 Study Approval

The study was approved by the Ethics Committee of the Medical Faculty of the Eberhard Karls University of Tübingen and the University Hospital.

2.2 Subjects

Twenty right-handed, eye-healthy subjects, aged between 19 and 38 years participated in the study with a prior written consent in adherence to the declaration of Helsinki. All participants were naïve to the purpose of the study; none had prior experience in microsurgery.

2.3 Experimental Setup

The experimental setup consisted of a line-stereo video system (TRENION 3D HD, Carl Zeiss Meditec AG, Jena, Germany) with the $930 \times 523 \times 1067$ mm stereo monitor (LMD-4251TD, Sony Corporation, Minato, Tokio, Japan), positioned at a height of 123 cm, and a head-worn eye tracker Dikablis Professional Glasses (Ergoneers GmbH), both connected to a prototype DSM. The setup is depicted in Fig. 1(a). The subject was standing in front of the 3D monitor at a distance of 100 cm, wearing polarized glasses enabling line-stereo-based 3D vision, as well as the head-worn eye tracker. In the subject's right hand, forceps were placed. The DSM was positioned to the right of the monitor, at approximate working distance from a subject. It was ensured that all subjects were able to reach the task object comfortably.

2.4 Dynamic Stereo Microscope

The DSM incorporated movable optical components into a CMO microscope to imitate accommodation and ocular parallax in real time, both on the basis of eye-tracking-based estimations
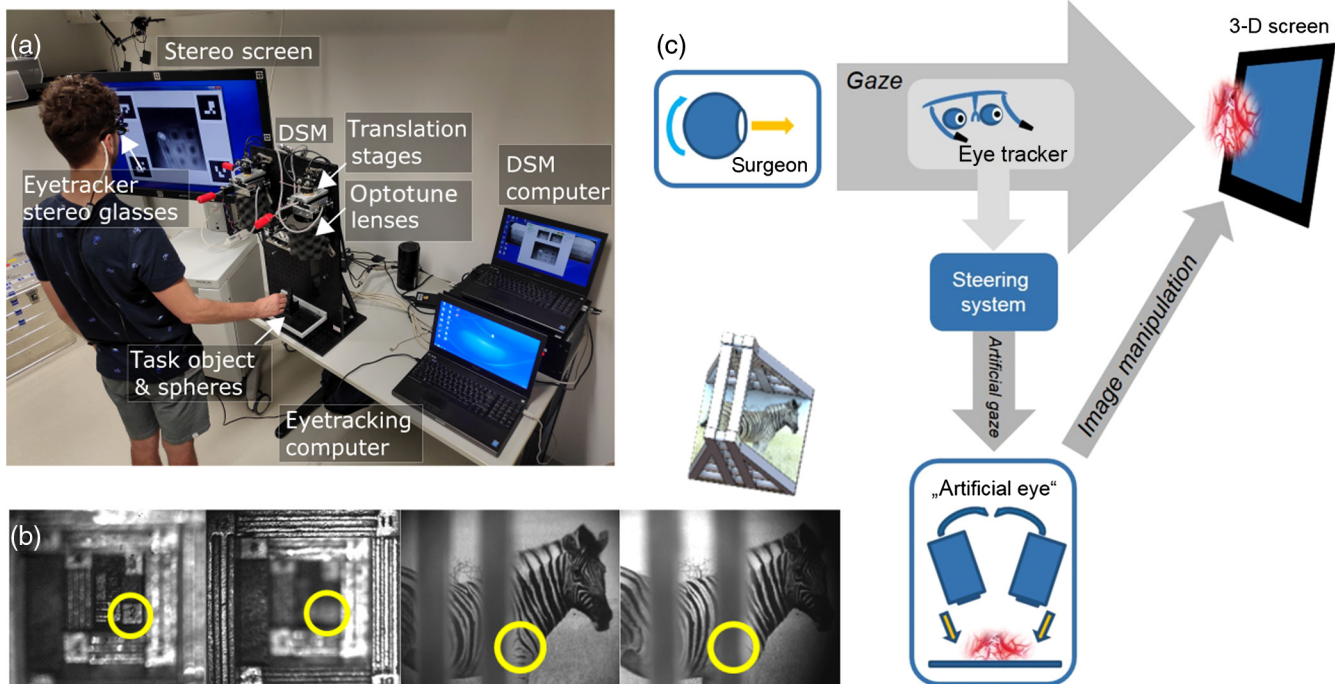

Fig. 1 (a) Labeled display of the setup with the microscope breadboard on the right, together with the steering unit, the screen in the center, and system computers on the right. (b) Sample images demonstrating focus shift and ocular parallax. Circles indicate selected spatial positions in which image changes through accommodation and parallax become specifically obvious. On the top of the parallax images, a 3D representation of the object shown in the parallax example is shown. (c) Schematic representation of the principle of the DSM. Eye movements of a subject viewing the object under the DSM on the screen are recorded. Based on the gaze position, autofocus and parallax-inducing aperture are adjusted, which adjust screen content.

of the gaze position on the screen. The accommodation mimicking autofocus was implemented by tuneable lenses (EL-10-30-C, Optotune AG, Dietikon, Switzerland), ocular parallax was mimicked by a translating circular aperture, realized through linear actuators (LS2818L0604-T5x5-75, Nanotec Electronic GmbH & Co. KG, Feldkirchen, Germany). The aperture selects a sub-bundle of the light to be imaged to the camera sensor, just like the ocular pupil would in natural vision. Thereby, the image projection of a 3D object onto the camera sensor changes as it would through eye movements, restoring ocular parallax.[23] Figure 1(b) shows examples of according focus adjustment and ocular parallax. Gaze positions of both eyes on screen were recorded by the Dikablis hardware, combined with customwritten software (EyeRecToo).[27] Upon saccades, according new gaze directions were estimated, and focus planes for each eye were drawn from a predefined depth map of the object. The software adjusted the lens focus and shifted the motorized stage accordingly, so that the adjusted image content appeared on screen 205 ms after each saccadic eye movement in the autofocus condition and 295 ms after shifting the motor stage. The delay was measured end-to-end by videorecording a blinking LED together with the videostream of the same LED on the DSM screen on one video. Synchronization of eyetracking data and command logs of the steering system allowed an isolation of autofocus delay and translation stage, which equalled on 95 ms for the translation stage and 185 ms for the autofocus, whereas 110 ms of the overall delay was attributed to the pure video delay, present in a static scenario as well. These adjustments were based on the gaze estimation of the eye tracker, as well as a predetermined depth map estimation of the scene. Eye movements were recorded with the Dikablis Professional Glasses (Ergoneers GmbH) in combination with the EyeRecToo, which also took care of the pupil detection. Further descriptions of the system are published by Fuhl et al.,[25] a schematic representation of the function principle is shown in Fig. 1(c).

2.5 Task Design

A standard task of high-accuracy manual interaction was developed. The task included the placement of spheres into a topographically challenging sphere holder, shown in Fig. 2. The sphere holder included different depths to mimic the multiple depth plane nature of surgical tasks. It was designed in Autodesk 123D Design (Autodesk, Inc., San Rafael) and 3D printed on an Ultimaker2 (Ultimaker, Geldermalsen, Netherlands) out of black polylactic acid. The 3D object for the haptic task condition contained 12 pockets of 3.4-mm diameter with a circular opening at 12 different planes, arranged in a 4×3 grid. During the haptic task participants were asked to place small spheres out of white polyoxymethylen of the diameter of 3 mm (Hoch KG, Haßfurt, Germany) into those pockets with the help of forceps. The pocket top surfaces were tilted by 20 deg in order to increase the difficulty of the manual task and enforce the use of visual information. It was ensured that the spheres fitted into the pockets comfortably, once placed correctly on top of the opening. With one attempt per pocket subjects were asked to fill the 12 pockets, row by row, from left to right starting in the bottom row, repeating the task, as often as possible, for 5 min. A sufficient supply of spheres was stored in the bowl-shaped socket of the 3D object.
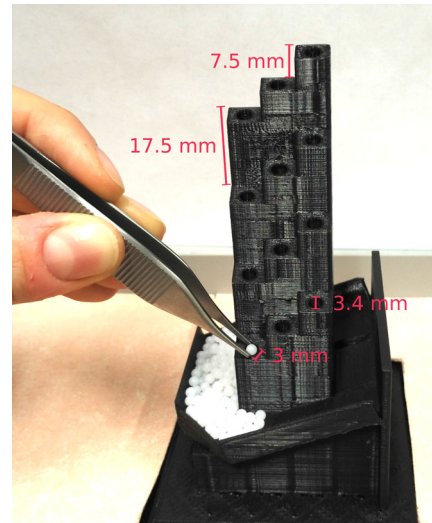
Fig. 2 Topographically challenging sphere holder for the manual task, together with measurement readings, and a subject's hand holding the forceps. On the bottom of the holder, a bowl-shaped socket for sphere supply is visible.

2.6 Experimental Procedure

Prior to the experiment, subjects were introduced to the microscope and the task. After familiarization, subjects were positioned standing in front of the monitor at a distance of 100 cm and equipped with polarized 3D glasses and an eye-tracker on top. Viewing was binocular. Subsequently, a custom-made eye tracking calibration was performed, in which a calibration marker (ArUco) was moved over the screen.[28] Subjects followed the marker with their gaze, and calibration was performed through the scene camera of the head-worn eye tracker. Calibration accuracy was tested in four reference positions in the corners of the microscope image. Thereafter, the training phase started. Participants were instructed on how to place the spheres, as well as the procedure of the task. Before the task was started, subjects were trained the task for 5 min to reach an initial saturation level of training effects of the manual task. The task started when the subjects had the forceps in the field of view of the microscope. After that they put as many spheres in the pockets as possible within 5 min. The number of spheres successfully placed was counted thereafter. At the end of each task, a task-specific questionnaire, aiming to evaluate their experience, was filled out by the participants (see Supplemental Fig. S1). Four separate conditions, consisting of three enhanced conditions (FOC, PAR, and COMB), as well as the static control condition (STAT) were tested in a random order on separate days, with one condition tested per day. In the STAT condition, variable focus and variable parallax were disabled, the focus plane was set to the average level of pocket openings, so that the highest and lowest pockets were blurred. In the FOC and the COMB conditions, the variable focus was activated, so that the system gaze-contingently kept the focus at fixation. In the PAR and COMB conditions, the moving aperture of the system mimicked the gaze direction changes induced through eye movements, and thus successfully imitated eye movement-induced parallax.

2.7 Data Analysis

Separately for each condition, task performance was estimated based on the number of spheres successfully placed within the

5 min. Average task performance was estimated in spheres per minute. In the questionnaire, different perceptual properties were evaluated with separate sets of questions: general perception, interaction, immersion, and depth. For details, refer to Supplemental Fig. S1. In each question, subjects rated a specific aspect on a Likert-type scale of 21 discrete steps. Individual questionnaire ratings were collected for each condition. Ratings were normalized to the maximum rating.
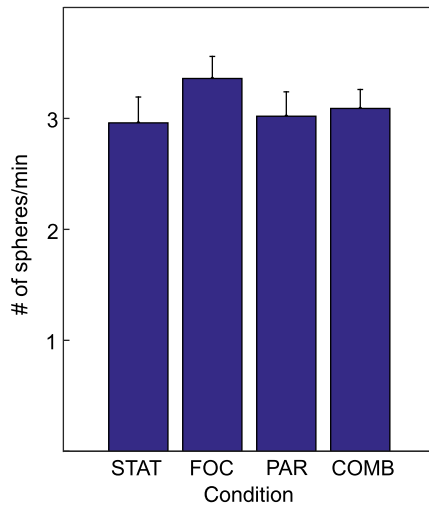
Fig. 3 Manual task performance in the four conditions: STAT, FOC, PAR, and COMB. Error bars indicate standard errors.

Data analysis of performance as well as questionnaire data focused on the difference of the three enhanced conditions (PAR, FOC, and COMB) to the baseline condition (STAT). Thus in the following, performance and ratings of any of the enhanced conditions are subtracted from condition STAT. As first step, an analysis of variance (ANOVA) is performed on relative performance changes as well as changes questionnaire ratings relative to baseline (STAT). Thereafter, in separate tests, the difference of each of the three enhanced conditions (PAR, FOC, and COMB) to the STAT condition was tested with student's t-test.

3 Results

3.1 Task Performance

Figure 3 summarizes task performance in the four measured conditions; error bars show standard errors. Average task performance is approximately three spheres per minute. An ANOVA indicated no significant differences in performance [$F(2, 59) = 1.3698$, $p = 0.26$]. When comparing performance between the different sockets of the object, no distinct pattern occurred. Thus there does not seem to be a large behavioral difference between the conditions.

3.2 Subjective Evaluation

In the next step, subjective ratings were evaluated. In Fig. 4, questions were grouped into three categories "interaction," "immersion," and "depth." Each subgraph shows average
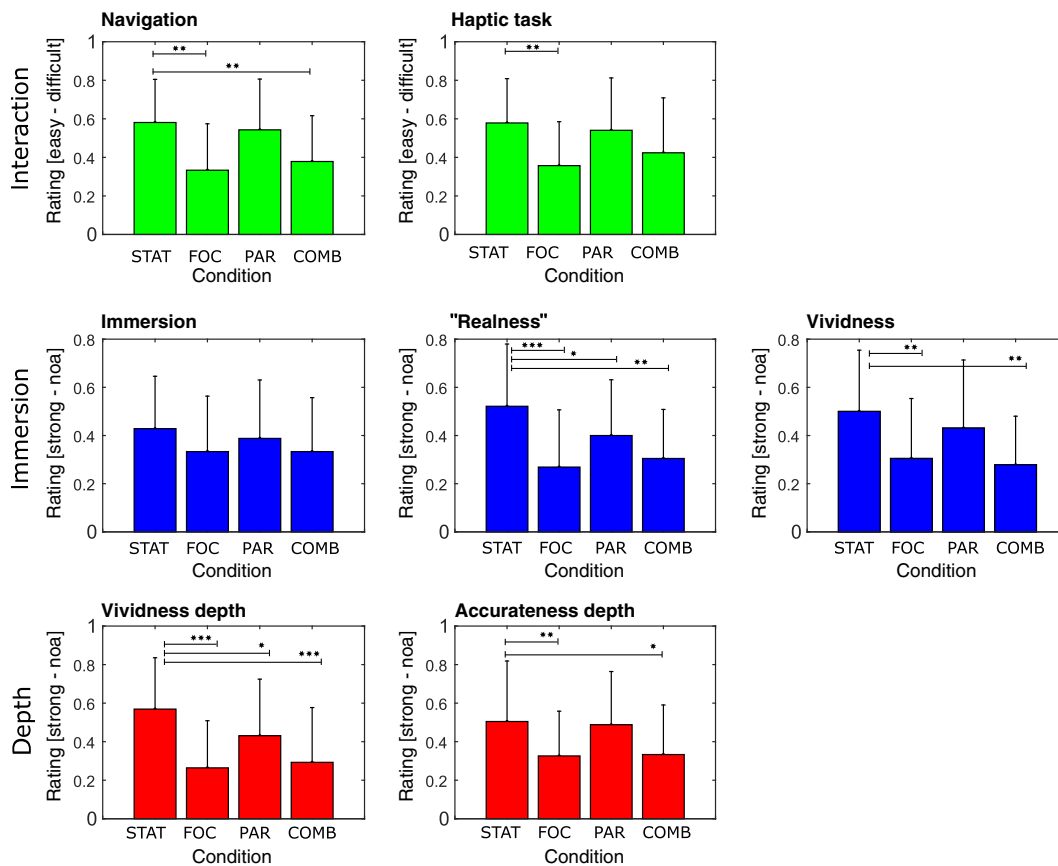
Fig. 4 Subjective ratings of performance and perception, grouped into topics of interaction, immersion, and depth. Error bars indicate standard deviations. Significance levels are indicated by asterisks.

Table 1 Summary of statistics of subjective evaluation.

Questions	ANOVA		Post hoc Δ_{STAT}		
	$F(2, 59)$	p	FOC	PAR	COMB
Navigation	8.4487	0.0009	0.0019	0.4588	0.0035
Haptic task	6.6281	0.0034	0.0035	0.5390	0.0504
Immersion	1.3171	0.2799	—	—	—
Realness	7.1601	0.0023	0.0001	0.0229	0.0014
Vividness	6.7238	0.0032	0.0036	0.2839	0.0034
Vividness depth	5.7316	0.0067	0.0001	0.0277	0.0008
Accurateness depth	6.7804	0.0030	0.0081	0.8141	0.0252

subjective ratings over all subjects, with zero indicating a maximally positive response, and one indicating a maximally negative response. Overall, enhancement conditions show significantly decreased rating scores, compared to the STAT condition in all fields: interaction, immersion, and depth. Decreased scores indicate a higher level of satisfaction. All but one of the questions showed significant differences in ratings to condition STAT in the ANOVA, results are summarized in Table 1. Solely the question related to immersion did not show a significant effect in the ANOVA, probably the term was difficult to interpret for the subjects [$F(2,59) = 1.3171$, $p = 0.2799$].

In those questions with significantly different ratings, each of the conditions was tested against zero. Conditions FOC and COMB show significantly improved ratings in almost all questions, p values are summarized in Table 1. Thus subjects indicate that they found it easier to navigate in the FOC ($p_{FOC} = 0.0019$) and COMB ($p_{COMB} = 0.0035$) conditions. They rated the haptic task better in the FOC condition ($p_{FOC} = 0.0035$). The COMB condition did not reach significance ($p_{COMB} = 0.0504$), well matching with the fact, that performance did not increase significantly. They furthermore rated "realness" of the presented object ($p_{FOC} = 0.0001$ and $p_{COMB} = 0.0014$) and vividness of presentation ($p_{FOC} = 0.0036$ and $p_{COMB} = 0.0034$) significantly better. When asked about their subjective rating of depth perception, their vividness of depth ($p_{FOC} = 0.0001$ and $p_{COMB} = 0.0008$), as well as accurateness in depth ($p_{FOC} = 0.0081$ and $p_{COMB} = 0.0252$) significantly increased, as soon as the autofocus of the system was activated. But interestingly, also condition PAR showed significantly improved ratings in two questions, even though on much lower significance levels. Realness of the observed object was rated more positively in the PAR condition ($p_{PAR} = 0.0229$), and vividness of depth was rated significantly higher in the PAR condition ($p_{PAR} = 0.0277$).

4 Discussion

This study demonstrates the positive contribution of focus cues and eye-movement-induced parallax cues in the perception of 3D content in a surgical microscope. In the FOC condition, in which the subject changed focus freely in the scene by accommodation, subjects rated interaction, immersion, and depth perception more positively. Furthermore, eye-movement-induced parallax cues played a role in depth perception as well. In sum, in surgical microscopes with static stereo, missing

accommodation and parallax cues are likely responsible for sub-optimal interaction, immersion, and depth perception.

4.1 Limitations of the Study and Application to Microsurgery

The generation of a stereo microscope with autofocus and parallax restauration in hardware represented a challenging task, which came with imminent limitations, namely an unevitable presentation delay in the system used. Specifically, in the conditions which restored ocular parallax, delays of up to 295 ms occurred. These might explain decreased performance and ratings in the COMB condition, compared to the FOC condition where solely the autofocus was in operation. Furthermore, task performance estimation is limited in its sensitivity to changes in stereo perception. This becomes specifically evident in the parallax restoration condition, where subjective ratings do show improvements, but performance is not significantly improved. One reason might be that for the haptic interaction the increased delay upon restoration of ocular parallax plays a role in the PAR condition as well as in the COMB condition. But furthermore, the changes in subjective perception are more subtle compared to improvements induced by an autofocus. Thus haptic performance might just not have been affected by it significantly. Haptic performance depends on manual feedback and repetition learning. Thus a moderate degradation of visual depth estimation might not necessarily strongly affect accuracy of reaching in depth. A task design in which no feedback on the task is given as it is here through the successful placement of a sphere might increase accuracy. In that case, manual depth accuracy would have to be tracked by a high-accuracy tracking device. A further study limitation concerns the study cohort of subjects in a task. Thus task as well as haptic strategies significantly differ from the surgical situation. But this study does not intend to estimate the extent of degradation in accuracy in a surgical situation, it instead proves the existence of a degradation of stereo perception. This degradation is universal to any human system and can thus be assumed to be well transferable to the surgery situation. In a surgery situation, the accuracy of depth estimations needed might even be higher than in the described haptic task performed in this study, thus the impact of the observed mechanism is expected to be more prominent. An comparison of systems in a long-term evaluation in long-term use would thus reveal great insights into the actual burden imposed onto the surgeon.

4.2 Digitalization in Microscopes

Specifically in the surgical context, accurate perception is critical. But technical solutions to circumvent each of the digital stereo-induced inaccuracies are challenging and expensive.[29] Furthermore, they seldomly restore more than one of the necessary depth cues. It is therefore useful, in the context of a stereo microscope, to classify alterations in depth cues according to their occurrence within the microscope and to enable an efficient search for technical solutions. In the following, we will therefore distinguish between image acquisition related depth cues and image presentation related depth cues.

4.3 Image Presentation-Related Depth Cues: Accommodation-Vergence Conflict

In most digital stereo systems, image presentation is bound to one specific surface at a fixed distance. Although stereo images

encourage vergence changes within the image, based on relative parallax between image details at different depths, accommodation is fixed to a plane. The two properties are naturally coupled; thus a constant mismatch is known to induce discomfort, termed accommodation-vergence conflict.[15,17,30–32] Various technological concepts have been developed to allow digital content presentation, in which accommodation changes together with vergence.[29] The approaches can roughly be classified into two groups – concepts, in which information is presented in different depths in parallel, and concepts, in which the presentation plane is changed gaze-contingently. Both represent cost extensive expansions to the standard two plane arrangement and suffer from high technological complexity. At a close look, the accommodation-vergence conflict consists of two types of components: motor and visual components. On the motor side, a certain status of contraction of the ciliary muscles is connected to a certain degree of vergence. But furthermore, both lead to a specific footprint in the image, namely a specific image projection and a specific blur distribution. This study shows that already the gaze-correct visual presentation provides considerable benefit, a solution in which image projection and blur distribution match vergence, although accommodation is still bound to the monitor plane.

4.4 Image Acquisition-Related Depth Cues: Focus Cues and Temporal Relative Parallax

Thus important depth cue alterations through digitalization originate in the image acquisition. Digitalization differs from optical microscopes in the fact that the light experiences a digitalization in an interim image plane. The light field thus collapses to the image plane, and the achieved image is presented on a presentation device. In an optical surgical microscope, this image plane undergoes changes through the surgeon's eye, which is part of the optical system. Thus the main additional impact on depth cues through digitalization is the removal of relative changes of depth cues through eye movements and accommodation. Depth cues of absolute disparity, spatial frequency content, familiar size, and occlusions are already affected by the optical transformation, they thus impact the surgeon's 3D perception already in purely optical systems. But relative changes of focus cues through accommodation, as well as parallax cues through eye and small head movements are conserved in an optical stereo microscope, but abolished in a static digital stereo microscope. Thus 3D perception is impacted further through digitalization.

4.5 Focus Cues

The contribution of focus cues to 3D perception has been evaluated in detail.[5,6,9,17,33,34] Although suffering from ambiguity, a significant contribution of blur as depth cue has been shown.[5,7] In this study, specifically the fixation position induced change in focus cues has been shown to significantly contribute to realism of depth. Its dominant role in the surgical context might originate in two aspects. As mentioned above, the magnification disrupts a variety of depth cues, thus leading to a more fragile situation, in which the visual system must rely on remaining cues. Furthermore, in the current situation, correct focus cues equal the situation of an autofocus, meaning, the focus is always set to the fixated plane, allowing clear vision. The comparison situation of a fixed depth plane, providing blurred vision upon fixations in depth planes different from the set focus

plane, is rather artificial and disturbing. But in the context of a digital stereo microscope, extending a stereo microscope with a static digital stereo would potentially become the new standard. Thus this study can be treated as a motivation to consider autofocus systems in digital microscopes.

4.6 Parallax Cues

Motion parallax is the only depth cue inherently defined in time. It describes the relative position change of objects at different distances upon global change of the environment, e.g., through movement of the observer. It has been demonstrated to be a rather strong depth cue.[35] Parallax through eye movements is a rather subtle image change, and the system is not able to perceive the relative motion of the object's details at different depths directly through saccadic suppression. Nonetheless, it has been shown to function as depth cue.[23] Specifically for near distances, a benefit of dynamic projection changes has been demonstrated in fusion times.[36] The present subjective ratings of depth realism confirm that parallax as a depth cue provides a subjectively perceivable benefit.

4.7 Fatigue

Inaccuracies of stereo content as demonstrated in this study play a major role in 3D induced fatigue and 3D-induced discomfort.[37] It can thus be assumed that decreased realism of 3D perception in static stereo images shows the potential to increase fatigue in digital surgical systems. Artificial stereo-induced fatigue has been measured extensively.[13,17,19,22,30,31,34,37–42] Target applications evaluated were mainly the commodity use of stereo content, in TV, cinema, or gaming VR.[20,21,43–46] Fatigue after viewing of 3D content is a well-reproduced phenomenon; its occurrence is temporary and limited to the duration of use and some hours thereafter. Long-term detrimental effects of viewing artificial 3D stereo content have not been found.[47] But in professional use, exposure to 3D stereo content differs from occasional commodity use of 3D: surgeons will use digital stereo microscopes for a vastly extended duration and more frequently. Furthermore, their role differs distinctly from that of a passive viewer. A surgeon interacts with the viewed content and needs to be able to perform high-acuity tasks. Thus the current results give insights into potential visual drawbacks of future digital surgical systems.

References

1. B. G. Cumming, E. B. Johnston, and A. C. Hurlbert, "Multiple cues for three-dimensional shape," in *The Cognitive Neurosciences*, M. Gazzaniga, Ed., pp. 351–364, The MIT Press, Cambridge, Massachusetts (1995).
2. J. E. Cutting and P. M. Vishton, "Perceiving layout and knowing distances: the integration, relative potency, and contextual use of different information about depth," in *Perception of Space and Motion*, W. Epstein and S. Rogers, Eds., pp. 69–117, Elsevier, San Diego, California (1995).
3. A. J. Parker, "Vision in our three-dimensional world," *Philos. Trans. R. Soc. B: Biol. Sci.* **371**(1697) (2016).
4. M. S. Landy et al., "Measurement and modeling of depth cue combination: in defense of weak fusion," *Vision Res.* **35**(3), 389–412 (1995).
5. M. S. Langer and R. A. Siciliano, "Are blur and disparity complementary cues to depth?," *Vision Res.* **107**, 15–21 (2015).
6. G. Mather, "The use of image blur as a depth cue," *Perception* **26**, 1147–1158 (1997).
7. S. J. Watt et al., "Focus cues affect perceived depth," *J. Vision* **5**(10), 7 (2005).
8. R. P. O'shea, S. G. Blackburn, and H. Ono, "Contrast as a depth cue," *Vision Res.* **34**, 1595–1604 (1994).
9. M. Vinnikov and R. S. Allison, "Gaze-contingent depth of field in realistic scenes," in *Proc. Symp. Eye Tracking Res. and Appl. (ETRA '14)*, pp. 119–126 (2014).
10. K.-P. Zimmer, "Optical designs for stereomicroscopes," *Proc. SPIE* **3482**, 690–697 (1998).
11. I. Söderberg et al., "Investigation of visual strain experienced by microscope operators at an electronics plant," *Appl. Ergon.* **14**(4), 297–305 (1983).
12. Y. Y. Yeh and L. D. Silverstein, "Limits of fusion and depth judgment in stereoscopic color displays," *Hum. Factors* **32**, 45–60 (1990).
13. T. Bando, A. Iijima, and S. Yano, "Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: a review," *Displays* **33**(2), 76–83 (2012).
14. D. M. Hoffman et al., "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *J. Vision* **8**(3), 33–33 (2008).
15. J. Kim, D. Kane, and M. S. Banks, "The rate of change of vergence-accommodation conflict affects visual discomfort," *Vision Res.* **105**, 159–165 (2014).
16. F. L. Kooi and A. Toet, "Visual comfort of binocular and 3D displays," *Displays* **25**(2–3), 99–108 (2004).
17. M. Lambooij et al., "Visual discomfort and visual fatigue of stereoscopic displays: a review," *J. Imaging Sci. Technol.* **53**(3), 30201 (2009).
18. C. J. Lin and B. H. Woldegiorgis, "Interaction and visual performance in stereoscopic displays: a review," *J. Soc. Inf. Disp.* **23**(7), 319–332 (2015).
19. J. C. Read and I. Bohr, "User experience while viewing stereoscopic 3D television," *Ergonomics* **57**(8), 1140–1153 (2014).
20. W. J. Tam et al., "Stereoscopic 3D-TV: visual comfort," *IEEE Trans. Broadcast.* **57**(2 PART 2), 335–346 (2011).
21. S. N. Yang et al., "Stereoscopic viewing and reported perceived immersion and symptoms," *Optometry Vision Sci.* **89**(7), 1068–1080 (2012).
22. L. Zhang et al., "Visual fatigue and discomfort after stereoscopic display viewing," *Acta Ophthalmol.* **91**(2), e149–e153 (2013).
23. G. P. Bingham, "Optical flow from eye movement with head immobilized: "Ocular occlusion" beyond the nose," *Vision Res.* **33**(5–6), 777–789 (1993).
24. M. Wexler and J. J. Van Boxtel, "Depth perception by the active observer," *Trends Cognit. Sci.* **9**(9), 431–438 (2005).
25. W. Fuhl et al., "Non-intrusive practitioner pupil detection for unmodified microscope oculars," *Comput. Biol. Med.* **79**, 36–44 (2016).
26. A. T. Duchowski et al., "Measuring gaze depth with an eye tracker during stereoscopic display," in *Proc. ACM SIGGRAPH Symp. Appl. Percept. in Graphics and Vis. (APGV '11)* (2011).
27. T. Santini et al., "EyeRecToo: open-source software for real-time pervasive head-mounted eye tracking," in *Proc. 12th Int. Joint Conf. Comput. Vision, Imaging and Comput. Graphics Theory and Appl.* (2017).
28. S. Garrido-Jurado et al., "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognit.* **47**(6), 2280–2292 (2014).
29. G. Kramida, "Resolving the vergence-accommodation conflict in head-mounted displays," *IEEE Trans. Vis. Comput. Graphics* **22**, 1912–1931 (2016).
30. D. Kim and K. Sohn, "Visual fatigue prediction for stereoscopic image," *IEEE Trans. Circuits Syst. Video Technol.* **21**(2), 231–236 (2011).
31. M. Lambooij et al., "Measuring visual fatigue and visual discomfort associated with 3-D displays," *J. Soc. Inf. Disp.* **18**(11), 931 (2010).
32. G. P. Mylonas et al., "Assessment of perceptual quality for gaze-contingent motion stabilization in robotic assisted minimally invasive surgery," in *Int. Conf. Med. Image Comput. and Comput.-Assist. Interv.*, Springer, pp. 660–667 (2007).
33. G. Maiello et al., "Simulated disparity and peripheral blur interact during binocular fusion," *J. Vision* **14**(8), 13 (2014).
34. M. Vinnikov, R. S. Allison, and S. Fernandes, "Impact of depth of field simulation on visual fatigue: who are impacted? and how?" *Int. J. Hum.-Comput. Stud.* **91**, 37–51 (2016).
35. M. F. Bradshaw et al., "The task-dependent use of binocular disparity and motion parallax under natural viewing conditions," *Perception* **26**(suppl), 48 (1997).
36. M. Bernhard et al., "The effects of fast disparity adjustment in gaze-controlled stereoscopic applications," in *Proc. Symp. Eye Tracking Res. and Appl.* (2014).
37. J. Park et al., "3D visual discomfort predictor: analysis of horizontal disparity and neural activity statistics," *IEEE Trans. Image Process.* **24**(3), 1101–1114 (2015).
38. D. Kane, R. T. Held, and M. S. Banks, "Visual discomfort with stereo 3D displays when the head is not upright," *Proc. SPIE* **8288**, 828814 (2012).
39. Y. Nojiri et al., "Visual comfort/discomfort and visual fatigue caused by stereoscopic HDTV viewing," *Proc. SPIE* **5291**, 303–313 (2004).
40. T. Shibata et al., "The zone of comfort: predicting visual discomfort with stereo displays," *J. Vision* **11**(8), 11 (2011).
41. C. Tosha et al., "Accommodation response and visual discomfort," *Ophthal. Physiol. Opt.* **29**(6), 625–633 (2009).
42. S. Yano, M. Emoto, and T. Mitsuhashi, "Two factors in visual fatigue caused by stereoscopic HDTV images," *Displays* **25**(4), 141–150 (2004).
43. P. A. Howarth, "Potential hazards of viewing 3-D stereoscopic television, cinema and computer games: a review," *Ophthal. Physiol. Opt.* **31**(2), 111–122 (2011).
44. M. Lambooij, W. A. Ijsselsteijn, and I. Heynderickx, "Visual discomfort of 3D TV: assessment methods and modeling," *Displays* **32**(4), 209–218 (2011).
45. A. Turner, J. Berry, and N. Holliman, "Can the perception of depth in stereoscopic images be influenced by 3D sound?" *Proc. SPIE* **7863**, 786307 (2011).
46. M. Urvoy, M. Barkowsky, and P. Le Callet, "How visual fatigue and discomfort impact 3D-TV quality of experience: a comprehensive review of technological, psychophysical, and psychological factors," *Ann. Telecommun.* **68**(11–12), 641–655 (2013).
47. J. C. Read et al., "Viewing 3D TV over two months produces no discernible effects on balance, coordination or eyesight," *Ergonomics* **59**, 1073–1088 (2016).

Biographies of the authors are not available.