Research Article

A Study of Correlations among Image Resolution, Reaction Time, and Extent of Motion in Remote Motor Interactions

Zoltán Rusák, Adrie Kooijman, Yu Song, Jouke Verlinden, and Imre Horváth

Delft University of Technology, Landbergstraat 15, 2628CE ZH Delft, The Netherlands

Correspondence should be addressed to Zoltán Rusák; z.rusak@tudelft.nl

Motor interaction in virtual sculpting, dance trainings, and physiological rehabilitation requires close virtual proximity of users, which may be hindered by low resolution of images and system latency. This paper reports on the results of our investigation aiming to explore the pros and cons of using ultrahigh 4K resolution displays (4096 × 2160 pixels) in remote motor interaction. 4K displays are able to overcome the problem of visible pixels and they are able to show more accurate image details on the level of textures, shadows, and reflections. It was our assumption that such image details can not only satisfy visual comfort of the users, but also provide detailed visual cues and improve the reaction time of users in motor interaction. To validate this hypothesis, we explored the relationships between the reaction time of subjects responding to a series of action-reaction type of games and resolution of the image used in an experiment. The results of our experiment showed that the subjects' reaction time is significantly shorter in 4K images than in HD or VGA images in motor interaction with small motion envelope.

1. Introduction

In the past decade, dramatic developments in video and network technologies enabled the support of different types of videoconferencing applications such as conference meetings, online lecturing, and everyday communication [1]. Recent improvements of videoconferencing systems offer new experiences to users. For instance, traffic aided opportunistic scheduling algorithm may significantly reduce the video codec latency in real-time videoconferencing sessions [2], which in the end enables enhanced agility, flexibility, and efficiency of businesses [3], offers new possibilities for innovations [4], and influences social interaction [5]. Moreover, combining videoconferencing with virtual reality technology enables collaborative design and product evaluation [6]. Remote collaboration is considered to be a specific form of human-computer interaction, in which factors like frame rate, the linearity, the lip sync (synchronization of audio and video signals), the latency, the image/video quality, and the video resolution are primary factors influencing the user

experience [7]. Earlier studies have revealed that low quality audio-video communication may harm the user experience in a video-mediated interaction process [8]. High-quality video post has the risk of overloading the network, but it may also offer considerable advantages. The study reported in [9] revealed that, in a videoconferencing system, the resolution of the transferred video is more important than its size for creating proper user experience. In an extensive test of effects of screen resolution [10], it was found that most video-mediated tasks can be carried out more efficiently and sometimes more accurately with using higher resolution displays. However, it seems that the frame rate and the colour depth of the video have less influence [11] on the user experience from both the image analysis and enjoyment point of view.

Remote virtual interaction mimics physical interaction, such as remote collaboration in sculpting, dancing trainings, and physiological rehabilitation, and it sets even higher demands on the technical feasibility of videoconferencing. In many of these applications, the flow of interaction is critical

in order to achieve realistic experience of remote human-human collaboration. This phenomenon is, on the one hand, mainly influenced by the latency of the videoconferencing system, which is highly dependent on the resolution of the transferred images. On the other hand, natural interaction requires proper resolution and size of images, in order to blend virtual environments into the physical world and to mimic interaction of real world settings.

Fast developments in digital cinema systems and network technology made high-quality videoconferencing systems possible. 4K (4096 × 2160 pixels) video quadruples the number of pixels of the 1080i HD television format. It generates a compelling, true-to-life experience and is able to be projected on a big screen providing an immersive feeling without change of visual acuity [12]. Ultrahigh image resolution enables closer proximity of users to the displayed image creating an immersive experience and improving "virtual" presence. 4K displays are able to overcome the problem of visible pixels and they are able to show more accurate details of the image content on the level of textures, shadows, and reflections. In 2006, it was reported by Shimizu et al. [13] that the world's first real-time, international transmission of 4K digital cinema and 4K super high definition digital video was successfully realized by the consortium of iGrid 2005. From a technical point of view, encoding, transferring, and decoding larger images require more computing power, which in the end creates additional system latency [14, 15].

Recent studies showed that system latency over the internet has a great influence on the user experience and successful completion of tasks [16]. For example, in networked real-time strategic games, the requirement of latency can be up to 1000 ms as opposed to networked shooting games and networked racing games, in which the requirement of the latency is much tighter (100 ms) [17, 18]. For remote sensory-motoric interactions, in a networked rock-paper-scissors game [19], it is found that the mean opinion score (MOS) [20] value of the users' satisfaction largely decreases when the latency exceeds 100 ms. In general, empirical experience has shown that the maximum acceptable one-way latency for remote interactive works not perceived by users as a limiting factor is around 150 ms [21].

This paper studies the benefits and limitations of remote motoric interaction facilitated by 4K videoconferencing systems. We have investigated the correlation between display resolution and network latency in videoconferencing systems during remote motoric interaction. In a series of experiments, the users were asked to make remote sensory-motoric interactions with a distant partner. We measured the users' reaction time to images played at different screen resolutions (e.g., VGA, HD, and 4K). It was our assumption that higher resolution images can not only satisfy visual comfort of the users, but also provide visual cues to the users of videoconferencing systems. We have expected, in typical remote sensory-motoric interactions, these visual cues are able to help the user to perceive the motion of remote partners quicker or even to anticipate their next actions. The outcome of this research contributes to establishing correlations between the technical characteristics of a videoconferencing system and its application in different scenarios and domains. These

correlations may help to optimize the videoconferencing system to achieve an optimal user satisfaction. We argue that reduced user reaction time may compensate for system latency introduced by router, codec, thereby keeping the user experience on the same level of satisfaction.

The rest of this paper is structured as follows. Section 2 presents a review of the state of the art in image forming and videoconferencing technologies together with methods of quality of videoconferencing system and sessions. Section 3 describes the research instruments and the setup of our experiment. In Section 4, the design of the content of the experiment is revealed. The results of the experiment are analysed in Section 5 and discussed in Section 6. Finally, the paper ends with conclusion.

2. State of the Art

With the proliferation of videoconferencing techniques, many aspects of remote communication and interaction have been studied. Media Richness Theory provides insights into how people adapt their behaviours to make use of various media systems. Though the claims of Walther and Parks [22] are still to be tested, it was already pointed out that if users select richer media for equivalent messages, then the efficiency of their communication and interaction will be increased.

Neustaedter et al. [23] have shown that image quality transferred during videoconferencing sessions has a significant influence on the persons' perception of the image content. They have studied blurring techniques to reduce image quality in order to preserve privacy and obscure awareness during videoconferencing. Their research discovered that video obfuscation (e.g. blurring) is not always an efficient filtration technique for preserving privacy during videoconferencing. Even videoconferencing sessions with relatively low image quality may provide enough information for identifying presence of persons in a room and the activities they perform. Motoric interaction of remote partners, on the other hand, requires fast and accurate perception of image content about remote partners in order to be able to produce realistic reactions. We argue that higher image quality provides a better platform for users to achieve realistic remote motoric interaction as it enables monitoring of movement and perception of the position of objects and people in real time.

Remote motoric interaction requires videoconferencing technology that facilitates interaction in close proximity. A key contribution to this issue can be found in the work of Hall and his notion of proxemics [24] (i.e., man's use of space as a specialized elaboration of culture). One of the key themes of Hall's work concerns the notion of physical distances that people maintain between each other according to their relationships and types of interaction. He characterized four main spatial distances that exist around a person: (i) intimate distance, (ii) personal distance, (iii) social distance, and (iv) public distance. Intimate distance is somewhere in the range up to 45 cm and is reserved to interactions with lovers, family, and close friends. Personal distance is in the range of 45 cm

to 1.2 m and is the distance that we would naturally maintain in the case of strangers in everyday life. These results suggest that remote motoric interaction requires technology (e.g., ultrahigh resolution displays) capable of imitating usage of spaces with personal proximity at intimate and personal distances.

Mediated interactions have been compared with face-to-face interactions by many researchers, with the goal of exploring relative dissatisfaction of people [25, 26]. Mediated interactions have been found inferior to face-to-face inter-action as they diminish some cues relevant to perception of vital information about interacting partners. The missing cues may force interacting partners to fill in the blanks based on past experiences, which may in the end result in a cognitive overload. This assumption has also been proven by some studies on efficiency prediction revealing. It has been revealed that successful interactions using cue lean media take more time than using cue rich or face-to-face interactions [27]. On the other hand, overload of cues may also be counterproductive and may result in a dissatisfying experience in videoconferencing. The efficiency framework introduced by Nowak et al. [28] proposes that when effort is relatively demanding, a more difficult medium may initially be less satisfying, but still useful. This increased effort, on the other hand, requires more attention to the interaction which might increase the sense of connection, or copresence, with the interaction partner [29]. To compensate increased attention and cognitive overload, people use multiple chan-nels (i.e., verbal, tactile, and visual physical behaviours that are simultaneously perceived via vision, hearing, and touch sensory systems and decoded at the same time) to trans-mit and receive information simultaneously. Simultaneous processing of verbal and nonverbal information can be achieved with little conscious effort and at the same time the information captured through these multiple channels is more complementary than redundant. Delays, however, can severely disturb this process and may result in unfavourable effects, such as displeasure, stress, and/or decrease in per-formance [30–32]. Recent research by Szameitat et al. [33] showed that delay not only decreases performance but also results in increased error rates and reaction times, as well as negatively affecting the emotional states. Most of the research summarized above employed computers with longer system response time than contemporary videoconferencing systems and therefore does not reflect the delay characteristics of modern HCI.

In our work, we investigated if system delays of videocon-ferencing systems can be compensated by using ultrahigh res-olution images. It is our assumption that ultrahigh resolution images reduce the reaction time of users and therefore can compensate for system latency in motoric interaction. Taking the reaction time of users as a quantitative measure, we aim to identify the relationship between the image resolution and the reaction time in remote sensory-motoric interaction. This relationship may offer new ways to improve quality of expe-rience by enabling the usage of ultrahigh resolution images in remote collaborative environments without increasing the perceived system latency.

3. Setup of Experiment

Our experiment was designed with the intent to minimize the effect of the biases on technical systems, offering deep engage-ment to collaborating partners and reducing the effect of learnability of the problem solving strategies. The experiment was conducted in an offline setup in order to get strict control over the system latency and to avoid any technical biases coming from the delays in image compression and fluctuation in data transfer in the network connection. Figure 1 illustrates the steps taken in the experiment. The offline setup of the experiment meant that the users were involved in an action-reaction activity, in which only one of the users made a motoric action, and the subject of the experiment reacted on his movement.

(i) In Step 1, we recorded the motoric activities of actors (such as dropping and catching an object, hitting the hand of the other person, and drawing a card from a pack). The motions of the actors were captured by a 4K camera from a frontal view representing the point of view of the remote partner, an HD camera in the side view for reconstructing interaction in front of a green screen, and a motion capture system to reconstruct and analyse the 3D motions of the actors and objects.

(ii) In Step 2, we played the recorded material (frontal view) on a 65-inch display with different resolutions (i.e., VGA, HD, and 4K) to the subjects included in the experiment. We measured the response times of the participants of the experiment with a high speed camera (200 fps) and we recorded their motoric reac-tion with an HD camera in the side view of a green screen and a motion tracking system. The high speed camera viewport included both the played back image (i.e., the stimulus) and the reaction of the participants. The viewport of the high speed camera is illustrated by Figure 2(b). The motion capture system recorded the full upper limb motion of each participant and the HD camera captured the reaction from a side view.

(iii) In Step 3, the two HD videos and captured motion data were merged into one scene in order to evaluate if the reaction of the participants resulted in a successful performance of the tasks. In addition, different laten-cies have been manually introduced to evaluate their effect on the completion of the tasks.

(iv) In Step 4, we extracted and analysed the reaction time of the users from the high speed videos.

4. Research Instrumentation

Figure 2(a) illustrates the basic setup of the experiment. In this setup, a 65-inch (width 1439 mm) screen is placed 1 meter in front of the user. The screen is able to display VGA ($640 * 480$), HD ($1920 * 1080$), and 4K ($4096 * 2160$) videos. A motion tracking system (yellow spot in Figure 2(a)), which is deployed to track the motion of the hands of the user, is placed
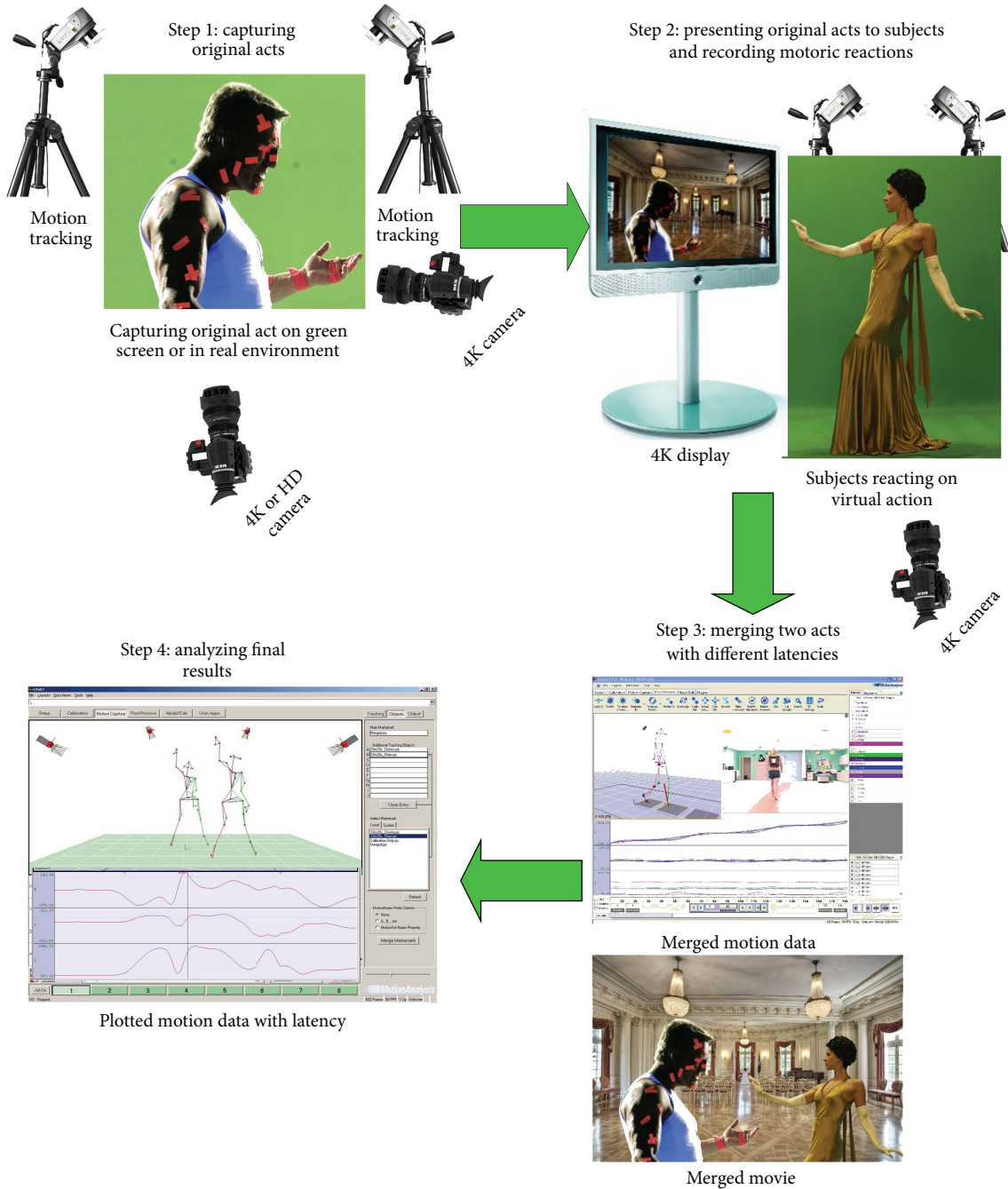
FIGURE 1: Work flow of the experiment (photograph that is courtesy of Dimension Films and Troublemaker Studios has been edited and used for illustrative purposes).

around the user. In the 4K configuration of this setup, the distance between pixels 1 and 2 at the boundary of the screen is about 0.35 mm. The visual angle $\theta 1$ of these two pixels is 1 arc minute. Using HD and VGA configurations, the visual angle $\theta 1$ of two pixels at the boundary of the screen will be enlarged to 2.2 and 6.6 arc minutes, respectively. That is, for an average human, pixels of the screen in 4K configuration are nearly visible. The pixels are visible in the case of HD images, and the pixels are clearly visible in the case of VGA images.

4.1. The Games. The interaction in our experiment has been purposefully designed with the goal of mimicking coopera-tion or competition between two persons being at a personal distance (0.45–1.2 meters). The goal was to imitate real world scenarios in a controlled manner, in which people are in close contact and they are reacting on each other's motion. It is expected that in the future more and more distant motor interaction will appear with the advancement and proliferation of videoconferencing, virtual and augmented
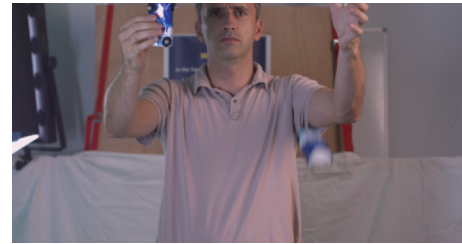
(a)
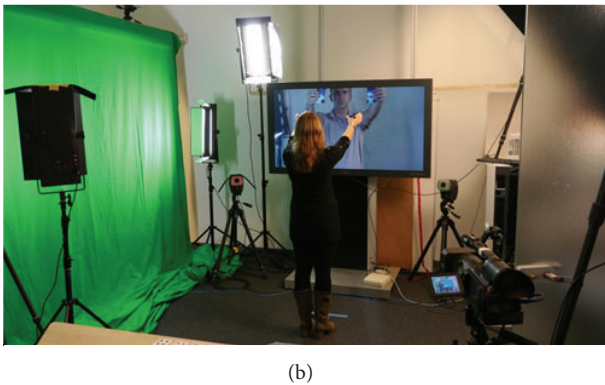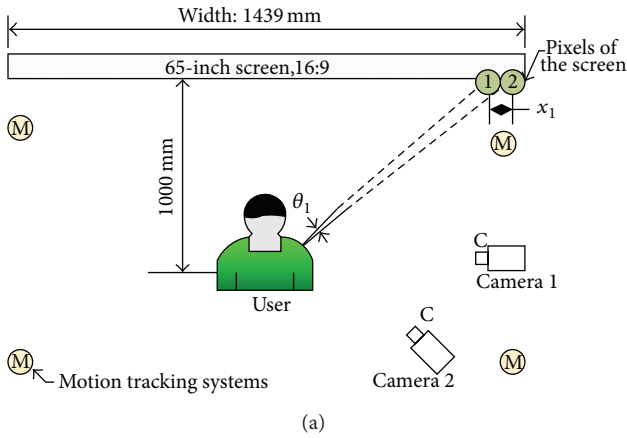
(b)

Figure 2: The setup of the experiment.

(a) Red hand game

(b) Drop test

(c) Card game

Figure 3: Games applied in the experiment.

reality technologies, robotics, and cyber-physical systems in applications. For example, with the spread of assisted living environments, service robots and telerobotic solutions are being introduced in the homes of the elderly, in order to support their daily living activities. In these applications, remote control of service robots and virtual interaction between the elderly and care givers takes places, in which the elderly are remotely monitored and assisted. Other applications, such as teleoperation of remote equipment (remote minimal invasive surgery, telerobotics in space and unsafe environments) or collaborative virtual design environments, may also demand quick motor reactions from remote users based not only on tactile and haptic feedback but also on the basis of visual information.

To embed realistic motoric interaction in an experimental setup, stimuli have been generated as games engaging remotely interacting partners. In our experiment, neither the actors performing the action nor the subjects reacting on the actions were constrained in their motions and expressions. Realism of the stimuli was necessary in order to reduce the chance of predicting the next moves of actors. The following criteria have been considered when selecting games for the experiments: (a) engagement of users in a monotonous exercise should be maintained as much as possible, (b) the motion envelope of the game should be consistent from the start of the action until 200 ms (i.e., the average reaction time to visual stimuli), (c) the subjects of the experiment should

react to motion only and no other modality of the stimuli should trigger their reaction, (d) the games should be reactive and not interactive due to the offline setup, and (e) the type of games should not impose cognitive challenges on the subjects engaging only the perceptive and motor channels.

Considering these criteria, we have designed three games (red hand, drop test, and card game shown in Figure 3), each of them triggering motoric reaction from the participants as response to a stimulus presented on the display. In the red hand game, a recorded actor holding his/her hands with the palm facing upward tries to hit the hand of the participant, whose hand is placed "into his hands." The task of the actor is to reach as fast as possible the hand of the participant, while he/she can pull back his/her arm to avoid getting touched. In this test, the motion envelope of the stimulus is determined by the complete arm movement of the actor. In the drop test game, the actor held two objects in his hand at the height of his shoulder and released one or both of them at his will. The task of the participant was to catch these object(s). In this case the motion envelope of the stimulus is generated by the finger movements of the actor as well as by the size of the dropped object. In the third game, 9 similar cards were laid facing downward on a table in front of the actor. The actor picked up one of them and showed it to the camera. The task of the participant is to point at the same card among the cards

laid in front of him/her. In this case the motion envelope of the stimulus is determined by the size of the card and the arm movement of the actor. The reaction time of the participants on the other hand is influenced not only by the extent of the stimulus but also by the visibility of details on the cards.

4.2. Experiment Design. Our experiment was conducted in a Motion Capture Laboratory of the Delft University of Technology. It was composed based on a within-subjects design, in which 20 participants were involved playing all three games at three resolutions (VGA, HD, and 4K). As a result of our setup, all participants were engaged 9 times for 1 to 2 minutes with a short break between each session. Each participant had been trained before the start of the experiment by exercising the games with a real person, in order to make sure that they have a good level of understanding of the games from the first moment of the experiment. In addition, the training session helped to bring the participants into an engaging attitude. Nevertheless, it should be noted that three factors influenced the engagement of the participants compared to real life situation. First, the participants did not share the same physical space and time with the actor. As a result the actor did not receive any feedback on the reaction of the participants and did not express any feelings (empathy, anger, etc.) towards them. Second, the participants did not get real-time feedback on their success of performing a task. Though, we have to note that the tasks were simplified to very low level actions such as moving the arm up or down. Third, the participants perceived the actors and the objects as two-dimensional images, which might have had an influence on the outcome of the drop test, but no influence on the reaction time. The displayed images of the actor and the objects were exactly the same as they are in real life; no image distortion due to resizing and cropping was introduced. We have used the same display with the same settings in all experiments and the recorded 4K material was converted to the same file format for all image resolutions. The games and the resolutions were played in a random order, in order to reduce the effect of learning. The participants were mostly students in the age group of 18 to 25 years and 30–70% female-male ratio. The independent variables of our experiment were the image resolution of the played back videos and the type of game. The dependent variable of our study was the reaction time of the participants.

4.3. Data Analysis and Processing. From the recorded material of high speed video and synchronized HD movie and motion tracking system we have extracted the following data:

(1) time of trigger: the beginning of the application of the stimulus that was introduced by the recorded actor;

(2) time of reaction: the beginning of the participants' motoric reaction to the stimulus.

In the data analysis phase we treated the outliers of our data based on the recommendation of Ratcliff [34] titled "Methods for dealing with reaction time outliers." In his study he claimed that the best way to treat outliers for data with high variability among subject means is to use standard deviation

TABLE 1: Summary of test means and standard deviations.

	4K		HD		VGA	
	Mean	STD	Mean	STD	Mean	STD
Red hand	0.198	0.07	0.211	0.055	0.199	0.049
Drop test	0.182	0.065	0.219	0.054	0.213	0.054
Card game	0.341	0.146	0.396	0.108	0.414	0.116

cutoffs. Following these recommendations we have applied this procedure on our data before the application of statistical methods and removed data that was at a larger distance than three times the standard deviation from the mean value. Figure 4 presents the complete data sets in the form of box diagrams for the red hand, drop test, and card games.

4.4. Analysis. Table 1 summarizes the results of the experiment. A one-way within-subjects ANOVA test was applied to compare the effect of image resolution on the reaction time in remote motoric interaction at VGA, HD, and 4K image resolutions. We have found that in the case of the red hand game there is no significant effect of the image resolution on the reaction time (F_{red} (2, 250) = 1.117, P = 0.311). However, we have found that the image resolution has a significant effect on the reaction time of users in case of the drop test and card game at the $P < 0.05$ level for the conditions (F_{drop} (2, 196) = 7.674, P = 0.001 and F_{card} (2, 163) = 5.163, P = 0.007, resp.). Post hoc comparisons using the Tukey HSD test indicated that in the case of the card game the mean score for the 4K resolution condition (M = 0.341, SD = 0.146) was significantly different than in case of HD resolution (M = 0.396, SD = 0.108). However, the HD resolution condition (M = 0.396, SD = 0.108) did not significantly differ from the VGA resolution condition (M = 0.414, SD = 0.116). Tukey HSD test showed similar results for the drop test. 4K resolution condition (M = 0.182, SD = 0.065) was significantly different than the HD (M = 0.219, SD = 0.054) and VGA (M = 0.213, SD = 0.054). In conclusion, these results suggest that ultrahigh resolutions do have an effect on reaction time of users when fine motoric interaction (i.e., stimuli with small motion envelope) is performed between remotely collaborating partners.

In order to validate this assumption, we have analysed the extent of stimuli for the red hand and drop test games using computer vision algorithms of Matlab. We have applied blob detection for identifying regions in the digital images based on differences in properties, such as brightness or colour, compared to areas surrounding those regions. Figure 5 illustrates an example of the original sequence of images together with changed pixels compared to the original image captured at the start of the stimuli. In our analysis the start of the motion has been considered as a reference image and each frame of the consecutive sequences of images was compared to it. The extent of stimuli is computed as the number of changed pixels divided by the total number of pixels.

The results of blob analysis algorithm on the extent of stimuli for the drop test and red hand game are summarized by Figure 6. This diagram shows how the area of changed
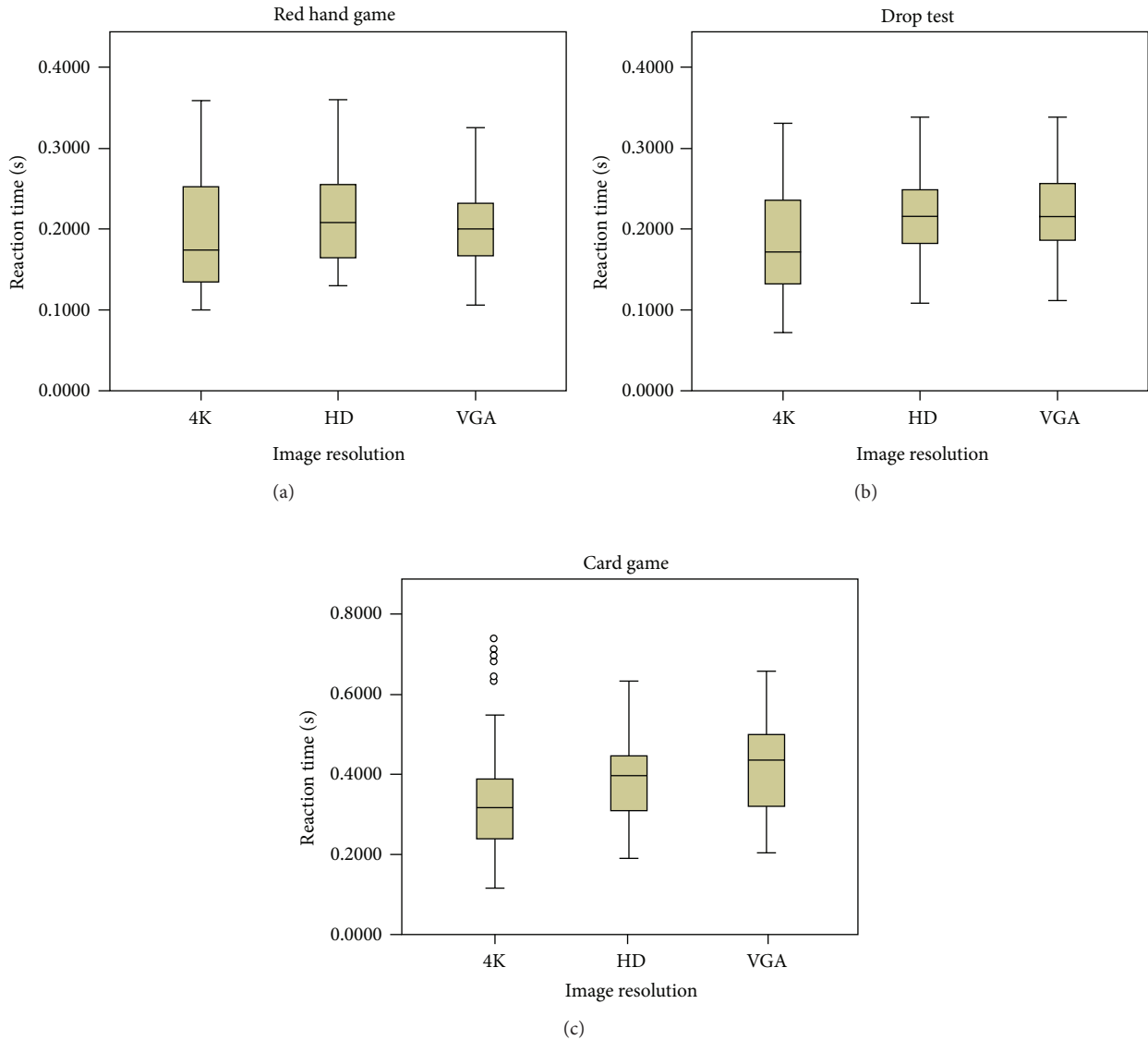
(a)

(b)

(c)

FIGURE 4: Box diagram of reaction time of participants.

pixels is changing in time during the first 200 ms from the start of the stimulation. The stimuli of the red hand game are characterized by a peak within 30 ms followed by a steady state or decreasing inclination. In the case of the drop test, the peak of the stimuli is more to the end of the standard reaction time of the users that is around 180 to 200 ms. The extent of motion nevertheless has also reached a representative value of 1% within 30 ms. Statistical analysis of the stimuli in Figure 7 shows that the mean of the percentage of the changed pixels is around 3% and 1% for the red hand and drop test games, respectively. The standard deviation of the stimuli of red hand game is in the range of 2–4 percent, while for the drop test these values are 0.5 and 1.5 percent. This means that the motion envelope of the stimuli in our setup was in the range of 10–25 cm^2 and 35–80 cm^2 on the display for the drop test and red hand games, respectively. These results imply that extent of motion in remote motoric interaction

has significant influence on the reaction time during remote collaboration. It can be concluded that fine motion (e.g., fine finger movements) can be faster detected on a 4K resolution videoconferencing system than on VGA or HD resolution.

5. Discussion

The results of our experiment showed that application of 4K technologies in videoconferencing sessions can reduce the reaction time of users in cases of fine motoric interaction. However, further studies are needed in order to explore this relationship between the extent (motion envelope) of the stimulus and the actual reduction of reaction time as well as their relationship to the image resolution. The results of our ANOVA test imply a relationship between the envelope of motion and the reaction time of the users. We found that the smaller the motion envelope is, the larger the gap between

FIGURE 5: Analysis of extent of stimuli for the drop test game.

the mean values of the reaction time gets between the 4K and HD/VGA images.

Based on the findings of our experiment, we can argue that 4K images in videoconferencing can be beneficial in those situations, in which motoric interaction is performed with a small motion envelope. As it is illustrated in Figure 8, the reaction time of remote partners is 30–40 ms faster and even with 80 ms latency the participant would be able to perform this task successfully. Larger latency will cause failure in the completion of these tasks. One may argue that a 4K videoconferencing setup is beneficial for a drop test if the latency is smaller than 40 ms, that is, the average difference in

reaction time for 4K and HD/VGA images in case of a drop test.

In other applications, such as designing teleoperations and telerobotics, it is necessary to consider not only the technical aspects of the designed system but also the human aspects [35]. By identifying relationship between the extent of motion, reaction time of users, and image resolution of stimuli, the results of our experiment are providing guidelines for system designers and engineers.

In the current setup of our experiment, the sensory-motoric interaction was simplified to scenarios with one-way communication between remote partners. This means that an
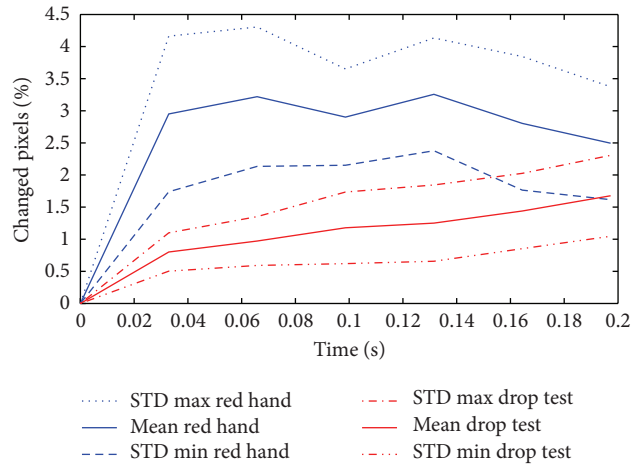
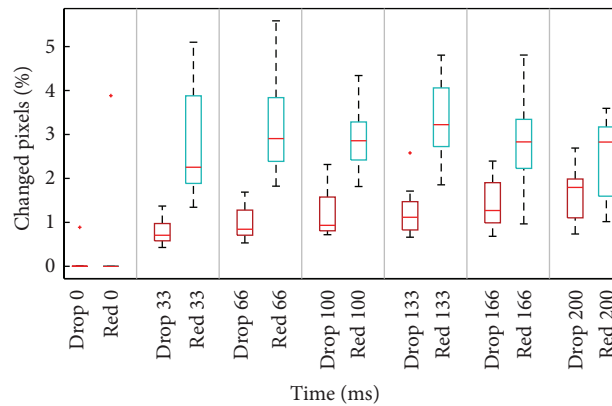FIGURE 6: Percentage of changed pixels of stimuli.

FIGURE 7: Descriptive statistics of extent of stimuli.

actor (the person creating the stimulus for the experiment) completed a move (made an action), and the subject of the experiment reacted on this movement. The actor in our setup did not see the subjects and could not react on the movements of the subjects. For this reason, it would be interesting to investigate more complex and interactive scenarios of sensory-motoric interaction.

However, this would require setting up 4K videoconferencing sessions in which the system latency is under 40 ms. Such setup would offer the opportunity to evaluate the correlation of image resolution system latency and motion intensity in more complex interaction scenarios such as dancing and boxing.

The following should also be further investigated: (i) what types of sensory-motoric actions are dominating in various application domains (e.g., healthcare: remote physiotherapy; industrial design: remote codesign and concept evaluation, education, and training) and (ii) how they can be performed in videoconferencing sessions. This scenario would enable us to propose the most relevant requirements for a videoconferencing system in these specific application domains.

The present and, in particular, the expected future results of our research may also provide guidelines for the development of new video compression and streaming technologies. Next generation video compressing and streaming technologies would adapt themselves not only to the technical conditions (actual bandwidth of networking, image resolution, and system latency), but also to the characteristics and types of sensory-motoric interactions, such as speed and extent of motion. For instance, if the user performs large and rough movements, then it is not necessary to use 4K resolution to successfully complete a task. On the other hand, fine motoric interaction may benefit from the usage of 4K resolution.

6. Conclusions

The paper investigated possible relationships between image resolution and reaction time (latency) in remote motoric interaction in videoconferencing sessions. We have conducted an experiment in which participants had to react to prerecorded games played at VGA (640 ∗ 480), HD (1920 ∗ 1080), and 4K (4120 ∗ 2160) resolutions. Their

FIGURE 8: Snapshot of merged videos of the drop test game with average reaction time of users.

reaction times have been measured with a motion tracking system and an HD camera, respectively. The collected data have been analysed by various statistical methods. We have done ANOVA and Tukey tests to see if there is a significant difference in the reaction time of the users to images played at different resolutions. The conducted ANOVA test showed that the difference of the mean values of the reaction times is not significant in the case of the red hand game, in which the motion envelope of the stimulus is relatively large. On the other hand, the mean value of the reaction times was 30 ms smaller for the 4K resolution video compared to the HD and VGA resolutions in the drop test game. There was no difference between the HD and VGA resolutions in this game. We have found similar trends in the case of the card game, but the difference among the mean values of 4K and HD/VGA was 60 ms. We are expecting that the correlations found by this explorative research will provide information about the technical requirements and implications of setting up videoconferencing sessions for sensory-motoric interactions, among others, in the fields of distance education, collaborative healthcare, product design, and remote user-product interaction.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] M. Credé and J. A. Sniezek, "Group judgment processes and outcomes in video-conferencing versus face-to-face groups," *International Journal of Human Computer Studies*, vol. 59, no. 6, pp. 875–897, 2003.

[2] M. Hu, J. Zhang, and J. Sadowsky, "Traffic aided opportunistic scheduling for downlink transmissions: algorithms and performance bounds," in *Proceedings of the 23rd Conference of the IEEE Communications Society (INFOCOM '04)*, pp. 1652–1661, Hong Kong, March 2004.

[3] T. E. Julsrud, R. Hjorthol, and J. M. Denstadli, "Business meetings: do new videoconferencing technologies change communication patterns?" *Journal of Transport Geography*, vol. 24, pp. 396–403, 2012.

[4] M. Kadama, "Creating new businesses through a strategic innovation community-case study of a new interactive video service in Japan," *International Journal of Project Management*, vol. 20, no. 4, pp. 289–302, 2002.

[5] A. Becker, A. Mladenow, N. Kryvinska, and C. Strauss, "Evolving taxonomy of business models for mobile service delivery platform," in *Proceedings of the 9th International Conference on Mobile Web Information Systems (MobiWIS '12)*, pp. 650–657, August 2012.

[6] R. Poelman, Z. Rusák, A. Verbraeck, and L. S. Alcubilla, "The effect of visual feedback on learnability and usability of design methods," *Strojniski Vestnik/Journal of Mechanical Engineering*, vol. 56, no. 11, pp. 744–753, 2010.

[7] R. Finger, "Measuring Video Quality in Videoconferencing Systems," EETimes, 2012, http://www.eetimes.com/design/other/4018049/Measuring-Video-Quality-in-Videoconferencing-Systems.

[8] S. Whittaker, "Rethinking video as a technology for interpersonal communications: theory and design implications," *International Journal of Human-Computer Studies*, vol. 42, no. 5, pp. 501–529, 1995.

[9] A. H. Anderson, A. Newlands, J. Mullin, A. M. Fleming, G. Doherty-Sneddon, and J. van der Velden, "Impact of video-mediated communication on simulated service encounters," *Interacting with Computers*, vol. 8, no. 2, pp. 193–206, 1996.

[10] B. Yost, Y. Haciahmetoglu, and C. North, "Beyond visual acuity: the perceptual scalability of information visualizations for large displays," in *Proceedings of the Computer/Human Interaction Conference*, pp. 101–110, San Jose, Calif, USA, May 2007.

[11] S. Y. Chen, G. Ghinea, and R. D. Macredie, "A cognitive approach to user perception of multimedia quality: an empirical investigation," *International Journal of Human Computer Studies*, vol. 64, no. 12, pp. 1200–1213, 2006.

[12] V. Dobson and D. Y. Teller, "Visual acuity in human infants: a review and comparison of behavioral and electrophysiological studies," *Vision Research*, vol. 18, no. 11, pp. 1469–1483, 1978.

[13] T. Shimizu, D. Shirai, H. Takahashi et al., "International real-time streaming of 4K digital cinema," *Future Generation Computer Systems*, vol. 22, no. 8, pp. 929–939, 2006.

[14] R. Serral-Graci, E. Cerqueira, M. Curado, M. Yannuzzi, E. Monteiro, and X. Masip-Bruin, "An overview of quality of experience measurement challenges for video applications in IP networks," in *Proceedings of the 8th International Conference on Wired/Wireless Internet Communications (WWIC '10)*, pp. 252–263, 2010.

[15] J.-R. Yang, R.-M. Hu, J. Chen, and J.-B. Zhong, "A service-oriented framework of distributed QoS measurement based on multi-agent for overlay network," in *Proceedings of the International Conference on Communication Software and Networks (ICCSN '09)*, pp. 158–162, Chengdu, China, February 2009.

[16] M. Claypool and K. Claypool, "Latency and player actions in online games," *Communications of the ACM*, vol. 49, no. 11, pp. 40–45, 2006.

[17] M. Claypool, "The effect of latency on user performance in Real-Time Strategy games," *Computer Networks*, vol. 49, no. 1, pp. 52–70, 2005.

[18] W. Wu, Y. Dong, and A. Hoover, "Measuring digital system latency from sensing to actuation at continuous 1-ms resolution," *Presence*, vol. 22, no. 1, pp. 20–35, 2013.

[19] Y. Hashimoto and Y. Ishibashi, "Influences of network latency on interactivity in networked rock-paper-scissors," in *Proceedings of the 5th ACM SIGCOMM Workshop on Network and System Support for Games (NetGames '06)*, Singapore, October 2006.

[20] M. Viswanathan and M. Viswanathanb, "Measuring speech quality for text-to-speech systems: development and assessment of a modified mean opinion score (MOS) scale," *Computer Speech and Language*, vol. 19, no. 1, pp. 55–83, 2005.

[21] R. Hill, C. Madden, A. van den Hengel, H. Detmold, and A. Dick, "Measuring latency for video surveillance systems," in *Proceedings of the Digital Image Computing: Techniques and Applications (DICTA '09)*, pp. 89–95, December 2009.

[22] J. B. Walther and M. R. Parks, "Cues filtered out, cues filtered," in *Handbook of Interpersonal Communication*, M. L. Knapp and J. A. Daly, Eds., pp. 529–563, Sage, Thousand Oaks, Calif, USA, 2002.

[23] C. Neustaedter, S. Greenberg, and M. Boyle, "Blur filtration fails to preserve privacy for home-based video conferencing," *ACM Transactions on Computer-Human Interaction*, vol. 13, no. 1, pp. 1–36, 2006.

[24] E. T. Hall, *The Hidden Dimension*, Doubleday, New York, NY, USA, 1966.

[25] B. Caldwell, S. Uang, and L. Taha, "Appropriateness of communications media use in organizations: situation requirements and media characteristics," *Behavior and Information Technology*, vol. 14, pp. 199–207, 1995.

[26] J. Schmitz and F. Fulk, "Organizational colleagues, media richness, and electronic mail; a test of the social influence model of technology use," *Communication Research*, vol. 18, no. 4, pp. 487–523, 1991.

[27] J. B. Walther, "Interpersonal effects in computer-mediated interaction. A relational perspective," *Communication Research*, vol. 19, no. 1, pp. 52–90, 1992.

[28] K. L. Nowak, J. Watt, and J. B. Walther, "The influence of synchrony and sensory modality on the person perception process in computer-mediated groups," *Journal of Computer-Mediated Communication*, vol. 10, no. 3, 2005.

[29] K. L. Nowak, J. Watt, and J. B. Walther, "Computer mediated teamwork and the efficiency framework: exploring the influence of synchrony and cues on media satisfaction and outcome success," *Computers in Human Behavior*, vol. 25, no. 5, pp. 1108–1119, 2009.

[30] R. E. Barber and H. C. Lucas Jr., "System response time—operator productivity, and job satisfaction," *Communications of the ACM*, vol. 26, no. 11, pp. 972–986, 1983.

[31] J. L. Guynes, "Impact of system response time on state anxiety," *Communications of the ACM*, vol. 31, no. 3, pp. 342–347, 1988.

[32] L. M. Schleifer and B. C. Amick III, "System response time and method of pay: stress effects in computer-based tasks," *International Journal of Human-Computer Interaction*, vol. 1, no. 1, pp. 23–39, 1989.

[33] A. J. Szameitat, J. Rummel, D. P. Szameitat, and A. Sterr, "Behavioral and emotional consequences of brief delays in human-computer interaction," *International Journal of Human Computer Studies*, vol. 67, no. 7, pp. 561–570, 2009.

[34] R. Ratcliff, "Methods for dealing with reaction time outliers," *Psychological Bulletin*, vol. 114, no. 3, pp. 510–532, 1993.

[35] R. G. Julian, "Considerations for the operator-robot interface in teleoperation," in *Proceedings of the IEEE International Conference on Systems Engineering*, pp. 90–93, 1990.