Interfacing Assessment Using Facial Expression Recognition

Andersen, Rune A.; Nasrollahi, Kamal; Moeslund, Thomas B.; Haque, Mohammad Ahsanul

Published in:
Proceedings of the 9th International Conference on Computer Vision Theory and Applications (VISAAP 2014)

DOI (link to publication from Publisher):
10.5220/0004730801860193

Publication date:
2014

Document Version
Early version, also known as pre-print

Interfacing Assessment Using Facial Expression Recognition

Rune A. Andersen, Kamal Nasrollahi and Thomas B. Moeslund
Visual Analysis of People Laboratory, Aalborg University, Sofiendalsvej 11, 9200 Aalborg, Denmark
raan11@student.aau.dk, {kn, tbm}@create.aau.dk

Keywords: Facial Expression Recognition, Interfacing Technologies, Motion Controlled, Gamepad

Abstract: One of the most important issues in gaming is deciding about the employed interfacing technology. Gamepad has traditionally been a popular interfacing technology for the gaming industry, but, recently motion controlled interfacing has been used widely in this industry. This is exactly the purpose of this paper to study whether the motion controlled interface is a feasible alternative to the gamepad, when evaluated from a user experience point of view. To do so, a custom game has been developed and 25 test subjects have been asked to play the game using both types of interfaces. To evaluate the users experiences during the game, their hedonic and pragmatic quality are assessed using both subjective and objective evaluation methods in order to cross-validate the obtained results. An application of computer vision, facial expression recognition, has been used as a non-obtrusive objective and hedonic measure. While, the score obtained by the user during the game has been used as a pragmatic quality measure. The use of facial expression recognition has, to the best of our knowledge, not been used before to assess the hedonic quality of interfaces for games. The thorough experimental results show that the user experience of the motion controlled interface is significantly better than the gamepad interface, both in terms of hedonic and pragmatic quality. The facial expression recognition system proved to be a useful non-obtrusive way to objectively evaluate the hedonic quality of the interfacing technologies.

1 INTRODUCTION

Evaluating the user experience in interactive systems, similar to many other systems, is of crucial importance. This can, for example, provide very helpful feedbacks for further improvement of such systems. One of the key issues in developing an interactive system, like a game, is deciding about the interfacing technology that the game is going to provide to its users. For the gaming purposes, traditionally gamepad interfacing technology has been around for many years. But, recently motion controlled interfacing has been widely used in many different games. Can this emerging motion controlled interfacing technology be considered as a feasible alternative to the gamepad from the user's point of view? To answer this question, we have developed a game based on kinetic interaction, which is suitable for multiple interfacing technologies, and have asked our test subjects to play the game once with a gamepad interface and once with a motion controlled interface. Following that we have evaluated and compared the users' experiences in these two different cases.

Before going into any details, we first need to answer this question: How can the user experience be assessed in interactive systems? Following Hassenzahl et al (Hassenzahl et al., 2003)'s theory on the user experience evaluation, there are two types of quality measures for such systems, hedonic and pragmatic. The hedonic quality is concerned with whether the user has a fun and pleasurable experience. But, the pragmatic quality considers an extend for which the user can complete the desired task, e.g., How easy is the interface to learn and to use? How accurate is it? For assessing these two types of quality measures, subjective methods (such as questionnaires and interviews) and objective quantitative methods (such as electroencephalography, electromyography, heartbeat rate, respiratory level) have been used. Objective quantitative methods are always good as they can complement/validate results of the subjective methods. However, there is a problem with the mentioned quantitative methods for assessing the user experience in motion controlled gaming applications. For their measurement, one needs to install a hardware/sensor on the user's body. It is obvious that such an installation restricts the user's movements in playing the game. Hence, they are not suitable for assessing the user experience when a motion controlled interface is used. To overcome this, beside employing subjective methods (using questionnaires), this paper has proposed a new quantitative measure using facial expression recognition. To do so, a camera films the test subjects during playing the game. There is obviously a distance between the camera and the test subjects, i.e., the employed sensor does not impose any restriction on the movements of the user and therefore motion controlled interfaces can be used easily.

To the best of our knowledge, facial expression recognition has not been used as a quantitative measure for evaluating game input devices and methods of interfacing in the gaming context. However, (Tan et al., 2012) conducted a feasibility study using facial expression recognition to evaluate user experience in games, which showed a correct correlation between gameplay events and the valence of the classified facial expressions.

The rest of this paper is organized as follows: in the next section the literature on assessing the user experience of interfacing technologies in games is reviewed, in section 3 the developed game and the chosen devices for both interfacing methods are introduced, the employed measurement factors for assessing the user experience are explained in section 4, the experiential results are discussed in section 5, and finally, the paper is concluded in section 6.

2 RELATED WORK

This section provides an overview on existing approaches on evaluating the effect of input devices on the user experience in games. In (Natapov et al., 2009) the user experience of playing a game using Wii® remote and a gamepad are compared. It is shown that the throughput of the Wii® remote is significantly higher than the gamepad. At the same time, the error rate of Wii® remote is higher than the gamepad. Despite this, the Wii® remote was the preferred controller for playing games amongst the test subjects. In (Bateman et al., 2011) the performance difference of game input devices for steering a car has been studied, and it is shown that a thumbstick performs significantly better in terms of game completion time than a steering wheel and a mouse.

For examining the user experience of different game input devices, self-determination theory (Deci and Ryan, 2002) has been commonly used to study factors that influence motivation. The Player Experience of Needs Satisfaction (PENS) evaluation instrument encompasses evaluation of concepts such as intrinsic motivation, competence, autonomy, relatedness, presence and intuitive controls through the use of questionnaires (Ryan et al., 2006). PENS was used in a study showing that realistic and tangible mapping of controllers with game tasks increases the experience of autonomy and presence, but not competence (McEwan et al., 2012). PENS has further been used in (Birk and Mandryk, 2013) with self-discrepancy theory (Higgins, 1987) to study self-perception during playing of a game, including the effect of extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience. In this study three input devices, the Kinect®, the PlayStation Move® and the gamepad are compared. It is shown that Kinect® produces significantly higher positive affect than the Move and gamepad. Furthermore, Kinect® showed to be more enjoyable than the gamepad. Moreover, it was shown that Kinect® provided higher autonomy, relatedness, immersion and agreeableness than the gamepad, while the gamepad increased the neuroticism of the user compared to Kinect®.

Regarding the effect of motion control in a collaborative multi-user game, (Lindley et al., 2008) shows that input based on motion increases the engagement and social interaction in the game. In (Dahlgren and Lyck, 2011) grounded theory is used to examine the effect of motion controllers on the gameplay experience, compared to a gamepad. The interviews showed that motion controllers can enhance the gameplay, making the players more immersed, if the control mapping is sufficiently realistic and natural. Studies have further shown that interactivity in the form of natural controller mapping increases spatial presence (Skalski et al., 2007) and immersion (Pietschmann et al., 2012), which has been shown to predict enjoyment of games (Shafer et al., 2011). The studies show that motion controlled input devices that are mapped naturally can increase spatial presence and immersion, which leads to greater entertainment value, compared to when a gamepad is used.

3 THE DEVELOPED GAME

To evaluate the user experience of the two interfacing technologies, a game was developed that features a humanoid avatar, which is controlled by the user. The avatar needs to be navigated through a game course using kinetic interaction, while objects are moving towards the user. There are two types of objects; target objects that need to be hit, and obstacle objects that need to be dodged. The two types of objects and the avatar are shown on Figure 1 (top). Every time an object is successfully hit or dodged, points are scored. The purpose of the game is to score as many points as possible.

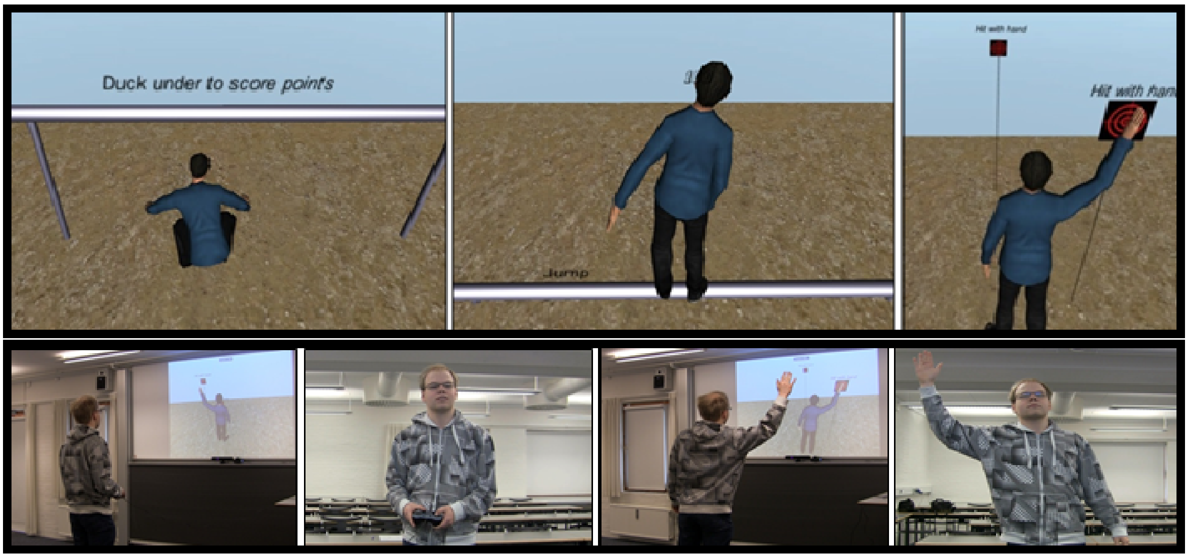

Figure 1: Top: The environment of the game showing the avatar and the two types of objects. The obstacle objects in the first two images on the left should be dodged (duck under or jump over) and the target object in the right most image should be hit. Bottom: The player with the gamepad interface (two left images), and with the motion controlled interface (two right images).

As mentioned before, two interfacing methods have been used in this paper. One interfacing method uses the fine motor skills of the hand (gamepad), and the other one uses gross motor skills of the entire body (motion controlled). Historically, video games have employed a gamepad in some form as input device, where a gamepad is defined as a controller held with one or two hands, using the fingers to provide input. As such, the gamepad is the standard interfacing technology for games. The chosen gamepad input device in this paper is a Logitech® RumblePad 2®, because it is a modern gamepad with features equal to those found on the current generation of gaming consoles.

The motion controllers that are commercially popular today as gaming controllers, and thus are candidates for input devices for game developers, fall into the categories of the hand held device, such as the Wii remote® and the PlayStation Move®, and whole-body tracking devices, such as the Kinect®. The Kinect for Windows® is chosen as the input device for the motion controlled interface. It is selected because it allows for the widest range of movements of the user, encompassing both movements of the hands that the hand held devices also cover, and movements of the rest of the body. Thus, interfacing through the Kinect® allows a wider range of body movements as input to the game, compared to a hand held device. This will enable an interface that fully utilizes motion of the user's body as input mechanism, in order to fully examine the possibilities of motion controlled gaming.

4 USER EXPERIENCE ASSESSMENT

For comparing the two different interfacing methods, using the chosen devices described in the previous section, both objective and subjective evaluation methods are used in accordance with the principle of triangulation (Jick, 1979), such that the strength in one method can compensate for the weakness in another, leading to a more valid result, without scientific artifacts caused by the use of a single evaluation method only. The conclusion of this study will therefore be cross-validated by comparing the results of the objective and subjective methods.

The hedonic and pragmatic quality measures are evaluated subjectively by asking test subjects to choose the interface they find most entertaining and easiest to learn to use. The pragmatic quality is further evaluated objectively by logging the score of the game. The easier the interface is to learn and the more accurate it is, the more points the user will score in the game, as the user has an easier time hitting and dodging the game objects. The hedonic quality is further evaluated objectively by a facial expression analysis system (Ghijsen, 2004), where the SHORE® library (Fraunhofer IIS, 2013), (Kublbeck and Ernst, 2006) is used to detect and classify the emotional valence of the expression of a recognized face in a video stream. The method is chosen because it is capable of inferring affective states of the user in a non-obtrusive way, compared to the more intrusive alternative of facial electromyography, where the muscle activity in the face is measured through attached electrodes, which to a greater extend will bias the evaluation. By applying this SHORE® facial expression recognition system to an image, first the faces are detected, then for each detected face its facial expression is expressed by four numbers. Each of these numbers are associated to one of these facial expressions: happiness, angry, sad, and surprised. Examples of the output of this system will be shown later in the paper.

5 EXPERIMENTAL RESULTS

To evaluate the user experience of the two interfacing methods, 25 test subjects have been asked to participate in the experiments. There were 22 males and 3 females between the ages of 20 and 32 years. Among them, 21 were experienced with playing video games on a gamepad-like input device, four were inexperienced with video games using a gamepad-like device. Each test subject was led into the test room alone by the test facilitator. It was explained to them that they were going to play a video game twice, once with a gamepad interface and once with a motion controlled interface through Kinect®. The game was demonstrated to them by the facilitator, so they knew how to play it. The test subject tried both versions, the order of which was switched for every new subject. The setup is shown in Figure 1 (bottom).

During the experiments the following data were gathered from each test subject:

- Data from the algorithm that registers and classifies facial expressions of the users as an index of hedonic quality.
- The score attained while completing the game as an index of pragmatic quality.
- The questionnaire data evaluating perceived entertainment value and learnability.

The facial expression recognition algorithm classifies and records the mean percentage value per video frame of both the positive and negative facial expressions during the use of each interface. The data is analyzed using a two-sample design with interfacing method as the treatment factor. It shows whether there is a difference in each facial expression category during the use of the two interfaces. The role of gaming experience level is examined visually through box plots depicting facial expression values with both versions of the game and gaming experience level as factors.

The game score in each prototype version is equally analyzed in a two-sample design. This determines whether there is a difference in game score using the two interfaces. Box plots are used to visually inspect whether gaming experience level is a factor on game score.

The questionnaire is designed as two AB experiments to evaluate the hedonic and pragmatic quality of the interfaces. The test subjects are asked to select the version that they find most entertaining, and the version they find easiest to learn to use. The dichotomous response variable is the preferred prototype version. Each AB experiment is analyzed in an exact binomial test (Crawley, 2005) in order to determine if one version is chosen significantly more often than the other. To avoid biasing the AB experiment, the sequence that the selectable options A and B are presented in the questionnaire are switched for every new test subject.

Following subsections describe the obtained results.

5.1 Facial expression analysis

The first step is to objectively evaluate the hedonic quality of the two interfaces by studying the valence of the facial expressions. Figure 2 shows the facial expression graphs of one of the test subjects, while using the gamepad. For each test subject using each prototype version an average value per video frame is computed for each facial expression. It can be seen in Figure 2 how the facial expression values in percentage fluctuate during a game course play-through using the gamepad interface. Based on these fluctuations the average value is computed for each facial expression. Taking the blue happy graph as an example, it has modest spikes for the entire play-through, resulting in a 4.59% average value.

The facial expression graphs of the same test subjects while using the Kinect® prototype is shown in Figure 3. As can be seen, the blue happy graph has significant spikes for most of the play-through, resulting in a 20.17% average happy rating. It can be seen that this test subject smiles a lot while using the Kinect®. When compared to Figure 2, she doesn't smile while using the gamepad. The happy rating her face is given while using the gamepad isn't based on a smile, which is also indicated by the fact that the happy rating is mostly accompanied by a parallel angry or sad rating. In Figure 3, spikes in the blue happy curve is rarely accompanied by spikes in the other curves, as an indication of a smiling face. The smiling face while using the Kinect®, resulted in an average happy rating of 20.17%, compared to the 4.59%

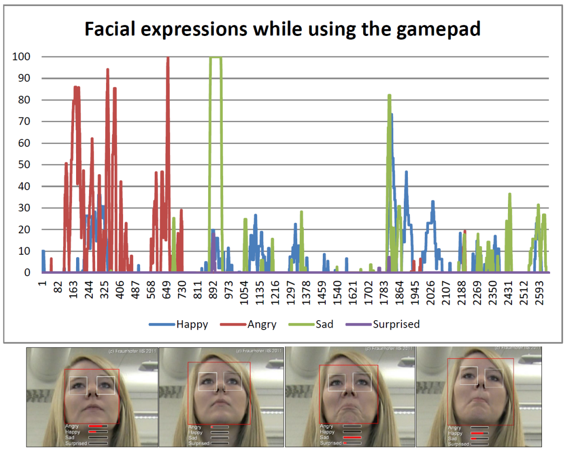

Figure 2: The results of the facial expression recognition system for a test subjects using the gamepad interfacing method. On the y-axis is the facial expression values in percentage, and on the x-axis is the video frame numbers. Below is shown the main facial expressions.

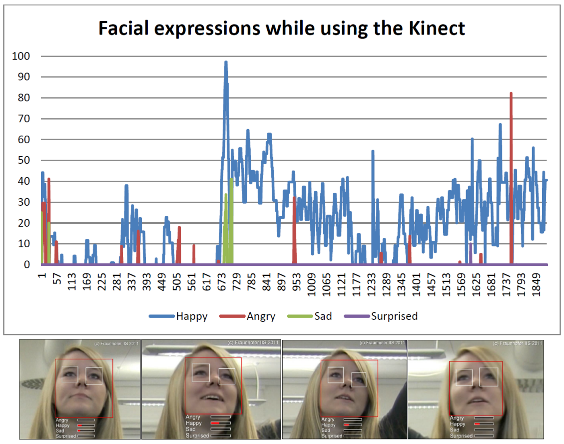

Figure 3: The results of the facial expression recognition system for a test subjects using the Kinect® interfacing method. On the y-axis is the facial expression values in percentage, and on the x-axis is the video frame numbers. Below is shown the main facial expressions.

happy rating while using the gamepad. This means that the employed facial expression recognition system does its job well of giving a higher average happy rating during a video sequence where a person smiles, compared to when not smiling. As such it is judged as a useful tool to classify and analyze valence of emotions based on facial expressions, particularly happy facial expression caused by being entertained.

Now, the facial expression data is analyzed in a two-sample design to determine whether there is a difference between the facial expressions in the two prototype versions, signifying a difference in hedonic quality between the two interfacing methods. First, the happy facial expression is analyzed. To be able to submit the two happy samples to a Student's t-test (Crawley, 2005), each has to follow normal distribution. As can be seen in Figure 4 (left), the happy percentage values of both samples exhibit exponential growth, which rules out normal distribution. However, a logarithmic data transformation results in approximately linear growth, which is illustrated on the right graph of Figure 4.

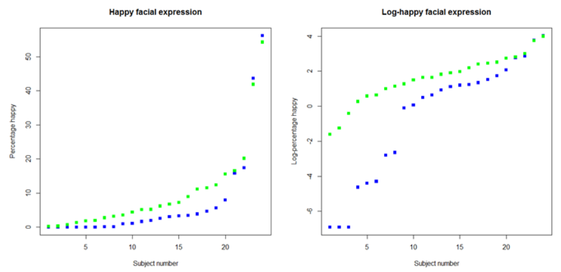

Figure 4: Left: The sorted observations of average happy facial expression per video frame. The blue are from the gamepad, the green from the Kinect®. Both exhibit an exponential growth pattern. Right: The logarithm of the same values. The growth is now close to linear, especially for the Kinect® values.

To determine whether the happy facial expression data follows log-normal distribution, Q-Q plots of the logarithm of the two samples are shown in Figure 5. Both samples appear to follow log-normal distribution. When the log-values are submitted to the Kolmogorov-Smirnov normality-test (Crawley, 2005), the gamepad sample comes out with a p-value at 0.1882 and the Kinect® sample at 0.9453. As such, the Kolmogorov-Smirnov null-hypothesis cannot be rejected, and the assumption of log-normal distribution can be accepted for both samples.

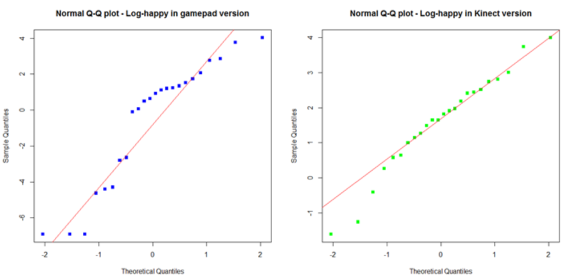

Figure 5: Q-Q plots for the logarithm of percentage happy facial expression. To follow log-normal distribution they are to lie on a straight line. The Kinect® sample does so very well, while the gamepad sample also does so, but not as well.

In Figure 6 (left) box plots of the log-samples are shown. Based on the box plots, it appears that the test subjects have more happy facial expressions when using the Kinect®, compared to the gamepad.

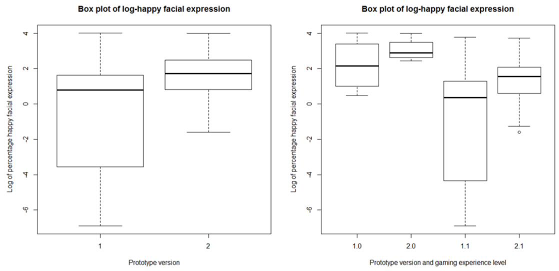

Figure 6: Left) Box plot of the logarithm of average percentage happy facial expression per video frame. Prototype version 1 is the gamepad, version 2 is the Kinect®. Right) Box plot of the logarithm of average percentage happy facial expression per video frame. On the x-axis 1.0 is gamepad version with non-gamers, 2.0 is Kinect® version with non-gamers, 1.1 is gamepad version with gamers, and 2.1 is Kinect® version with gamers.

When submitting the samples to the F-test, the F-value comes out at 6.1 with a p-value at 5.262×10^{-5}, meaning that the variance isn't equal between the samples. This was also indicated by the box plots, where the spread of the gamepad values was much greater than the Kinect® values.

The data is submitted to a paired t-test using unequal variance. The t-value comes out at -3.475 with a p-value at 0.002051. The 95% confidence interval is at -3.443 to -0.8733. As the p-value is well below the critical p-value of 0.05, the null hypothesis can be rejected stating that the mean log-happy values are equal in the samples. The Kinect® and gamepad mean values are 1.581 and -0.5769, respectively. It therefore can be concluded that the test subjects had significantly more happy facial expressions when using the Kinect®, compared to the gamepad. As an index of hedonic quality, it signifies that the test subjects had a more pleasurable experience using the motion controlled interface of the Kinect®, compared to the gamepad interface. To examine whether gaming experience level is a factor in happy facial expression level, the box plot including gaming experience level is shown in Figure 6 (right). When examining the box plots, it is evident that the non-gamers in general had more happy facial expressions when playing, compared to the gamers. However, both the gamers and the non-gamers looked happier when using the Kinect®, compared to the gamepad.

Similar analysis has been done for the other three facial expressions. While the angry and sad facial expression values didn't exhibit a clear pattern when analyzed, the samples for surprised facial expression showed that the Kinect® values were significantly higher than the gamepad values. The t-value was -4.126 and the p-value 0.0004117 with a 95 % confidence interval from -3.524 to -1.17. As an index of hedonic quality, the increase in surprised facial expression signifies that the Kinect® version produces a more pleasurable experience, which to a greater extend is capable of surprising the test subjects, compared to the gamepad version. Regarding the effect of gaming experience level on the amount of surprised facial expression, it appeared that non-gamers didn't look significantly more surprised when using the Kinect®, compared to the gamepad version, meaning that it is only gamers that look significantly more surprised when using the Kinect®.

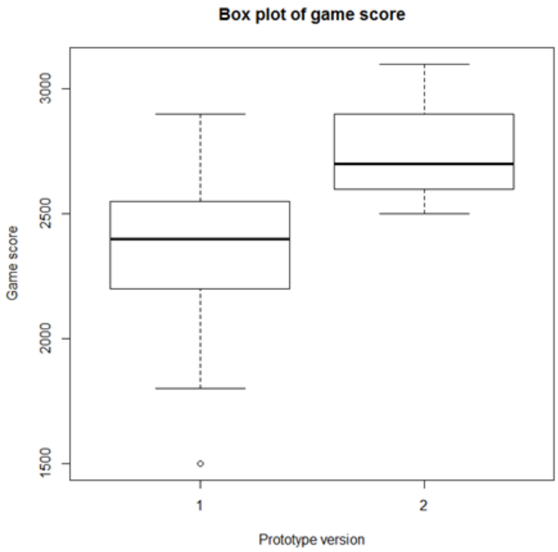

Figure 7: Box plot of the game score data. On x-axis is the prototype version, 1 is the gamepad, and 2 is the Kinect®, on the y-axis is the obtained game score.

5.2 Game score analysis

Next, the pragmatic quality is analyzed through the game score. The game score was obtained for each prototype version for each participant. Figure 7 shows the box plot of the game score data. It can be seen from this figure that the gamepad scores were spread out more widely from 1500 to 2900. The middle 50% of the gamepad values (the box in the box plot) are all below the middle 50% of the Kinect® values, indicating that the gamepad sample has significantly lower values. The mean game scores of each sample for gamepad and Kinect® are 2365.22 and 2756.52, respectively.

In order to determine whether there is a significant difference between the mean values of the gamepad and Kinect® samples, they are submitted to a paired t-test, where the variance is approximated for each sample. Prior to that, the assumption of normal distribution has been validated through Q-Q plots and

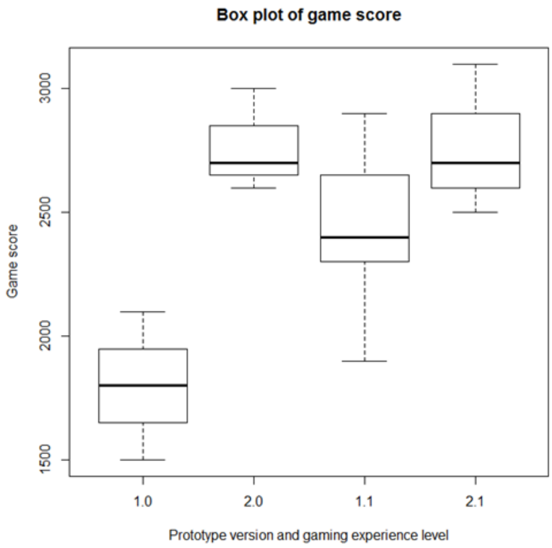

Figure 8: Box plot of the game score data. On the x-axis 1.0 is gamepad version with non-gamers, 2.0 is Kinect® version with non-gamers, 1.1 is gamepad version with gamers, and 2.1 is Kinect® version with gamers.

the Kolmogorov-Smirnov normality-test. The t-value comes out at -5.964 with a p-value at 5.301×10^{-6}. The 95% confidence interval is -527.37 to -255.23. As the Kinect® score is greater than the gamepad score, it can be concluded that the Kinect® interface has a higher pragmatic quality than the gamepad, as the test subjects was able to score higher when using the Kinect® interface as an indication of an interface that is easier to learn to use, and one that has greater accuracy. By studying the box plot in Figure 8, it can be seen that non-gamers score much lower than gamers when using the gamepad. This was expected as gamers are previously trained in using a gamepad-like input device for playing games like the prototype, where non-gamers are not. In terms of Kinect® scores, both gamers and non-gamers score equally, indicating that gaming experience level doesn't influence the ability to use the Kinect® interface. The bottom-line is that the Kinect® interface has greater pragmatic quality than the gamepad interface for both gamers and non-gamers, but for non-gamers the difference in pragmatic quality is much greater than for gamers.

5.3 Questionnaires analysis

Finally, the questionnaire data is analyzed. The purpose is to find out how the self-reported evaluation of hedonic and pragmatic quality comes out for the two prototype versions. Out of the 25 test subjects, 22 chose the Kinect® as the prototype version that they found most entertaining to use. When submitted to an exact binomial test, the p-value comes out at 0.0001565 with 95% confidence interval from 0.6878 to 0.9745. The null-hypothesis can therefore be rejected, stating that each prototype version is equally likely to be selected, with the Kinect® version being selected 88% of the time, and the gamepad 12% of the time. It can therefore be accepted that the Kinect® version is selected significantly more often than the gamepad version, when being asked to select the version that provides the most entertaining experience. As such, the Kinect® version is perceived as having significantly higher hedonic quality by the test subjects, providing a more pleasurable experience when used.

When examining the role of gaming experience level on the perceived hedonic quality of the two interfaces, it can be seen that the three test subjects that selected the gamepad version as the most entertaining all were gamers. All non-gamers selected the Kinect® version as most entertaining. It indicates that people who are inexperienced in playing video games using a gamepad-like device are even more likely to prefer the Kinect® over the gamepad, when considering entertainment value.

When examining the pragmatic quality based on the questionnaire data, 23 out of the 25 chose the Kinect® as the version that was easiest to learn to use. The binomial test comes out with a p-value of 1.943×10^{-5} with a 95% confidence interval at 0.7397 to 0.9902. This means that the Kinect® version is chosen significantly more often than the gamepad version, when asked to select the version that is easiest to learn to use. The Kinect® version is selected 92% of the time, with the gamepad 8% of the time. Thus, the Kinect® version is perceived as having significantly higher pragmatic quality, than the gamepad version. The two test subjects that find the gamepad version easier to learn to use are both gamers, who are experienced in playing video games with a gamepad-like device. The test subjects who are not previously trained in using the gamepad, all choose the Kinect® as easier to learn to use. This indicates that non-gamers are more likely to prefer the motion controlled interface of the Kinect®, when considering pragmatic quality, compared to gamers.

6 CONCLUSION AND DISCUSSIONS

In the general context of gaming, this paper has compared two interfacing technologies, traditional gamepad and motion controlled techniques. To do so, a game has been developed which can be played using both of the interfacing technologies. Then, a group of 25 test subjects have been asked to play the game twice using the available interfacing methods. During the gaming process both hedonic and pragmatic quality measures of the test subjects have been monitored and analyzed using subjective and objective assessment methods.

Using the subjective evaluation method of a questionnaire designed as an AB experiment, significantly more test subjects chose the Kinect® interface as the most entertaining, compared to the gamepad interface. In terms of pragmatic quality, significantly more test subject chose the Kinect® interface as easier to learn to use, compared to the gamepad interface. Furthermore, an objective evaluation method was applied by analyzing the facial expressions of the test subjects. The result was that the test subjects had significantly more happy and surprised facial expressions when using the Kinect®, compared to the gamepad. As an index of hedonic quality, it signifies that the test subjects had a more pleasurable experience using the motion controlled interface of the Kinect®, compared to the classical gamepad interface, causing them to look happier and more surprised. This cross-validates the subjective result. Furthermore, an objective evaluation method for pragmatic quality, using the logged game scores of the test subjects, cross-validated the obtained subjective result of a higher pragmatic quality of the Kinect® interface.

Facial expression recognition proved to be a useful application of computer vision, suited to evaluate the hedonic quality of the interfacing technologies in an objective, non-obtrusive manner.

REFERENCES

Bateman, S., Doucette, A., Xiao, R., Gutwin, C., Mandryk, R., and Cockburn, A. (2011). Effects of view, input device, and track width on video game driving. In *Graphics Interface 2011*, pages 207–214, St. John's, Canada.

Birk, M. and Mandryk, R. L. (2013). Control your game-self: effects of controller type on enjoyment, motivation, and personality in game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 685–694, New York, NY, USA. ACM.

Crawley, M. J. (2005). *Statistics: An Introduction using R*. Wiley. ISBN 0-470-02297-3.

Dahlgren, K. and Lyck, M. (2011). Motion control: In control or out of control? Master's thesis, Stockholm University.

Deci, E. and Ryan, R. (2002). *Handbook of Self-Determination Research*. University of Rochester's Press.

Fraunhofer IIS (2013). Intelligent Systems - Fraunhofer Institute for Integrated Circuits. http://www.iis.fraunhofer.de/de/bf/bsy/fue/isyst.html.

Ghijsen, M. (2004). Facial expression analysis for human computer interaction. *IEEE Transactions on Visualization and Computer Graphics*, 2(3):147–161.

Hassenzahl, M., Burmester, M., and Koller, F. (2003). Attrakdiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualitt.

Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychol Rev*, 94(3):319–340.

Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24(4):pp. 602–611.

Kublbeck, C. and Ernst, A. (2006). Face detection and tracking in video sequences using the modified census transformation. *Image and Vision Computing*, 24(6):564 – 572.

Lindley, S. E., Le Couteur, J., and Berthouze, N. L. (2008). Stirring up experience through movement in game play: effects on engagement and social behaviour. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 511–514, New York, NY, USA. ACM.

McEwan, M., Johnson, D., Wyeth, P., and Blackler, A. (2012). Videogame control device impact on the play experience. In *Proceedings of The 8th Australasian Conference on Interactive Entertainment: Playing the System*, IE '12, pages 18:1–18:3, New York, NY, USA. ACM.

Natapov, D., Castellucci, S. J., and MacKenzie, I. S. (2009). Iso 9241-9 evaluation of video game controllers. In *Proceedings of Graphics Interface 2009*, GI '09, pages 223–230, Toronto, Ont., Canada, Canada. Canadian Information Processing Society.

Pietschmann, D., Valtin, G., and Ohler, P. (2012). The effect of authentic input devices on computer game immersion. pages 279–292.

Ryan, R., Rigby, and Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, 30(4):344–360.

Shafer, D. M., Carbonara, C. P., and Popova, L. (2011). Spatial presence and perceived reality as predictors of motion-based video game enjoyment. *Presence: Teleoper. Virtual Environ.*, 20(6):591–619.

Skalski, P., Lange, R., Tamborini, R., and Shelton, A. (2007). Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. In *57th Annual Conference of the International Communication Association*.

Tan, C. T., Rosser, D., Bakkes, S., and Pisan, Y. (2012). A feasibility study in using facial expressions analysis to evaluate player experiences. In *Proceedings of The 8th Australasian Conference on Interactive Entertainment: Playing the System*, IE '12, pages 5:1–5:10, New York, NY, USA. ACM.