User Test of Soccer Highlights Application

Lu Wang

Philips Research Europe
High Tech Campus 34
5656AE The Netherlands

+31 40 2747528

l.wang@philips.com

Pedro Fonseca

Philips Research Europe
High Tech Campus 34
5656AE The Netherlands

+31 40 2747522

pedro.fonseca@philips.com

Bas Zoetekouw

Philips Research Europe
High Tech Campus 34
5656AE The Netherlands

+31 40 2747558

bas.zoetekouw@philips.com

ABSTRACT

This paper describes an evaluation of a consumer application for soccer highlights detection, which combines advanced user interaction with an underlying automatic highlights detection algorithm. The application was developed for a home entertainment setting, i.e., for non-professional use, and was implemented on a Media Center PC with a remote control. The evaluation uses a qualitative assessment to measure the participants' satisfaction with the application, and investigates their satisfaction regarding several important factors, such as the participants' expectations and their satisfaction with the highlights. In the evaluation, we found that the participants' expectations regarding a soccer highlights detection application vary widely, and there is a gap between what participants expect and what any automatic detection application can provide. Despite this gap, the results from the qualitative assessment indicate that the application is very positively perceived by the participants. The results indicate that the participants showed positive satisfaction with the application, especially in its ease-of-learning, ease-of-use, and the feeling of being in control. Moreover, the results suggest new and further tuned user requirements that help improve satisfaction with the application.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces – *User-centered design, evaluation/methodology*.

General Terms

Design, Experimentation

Keywords

User evaluation, soccer highlights application, user requirements

1. INTRODUCTION

Worldwide, soccer is the most popular sport, both to play and to watch. Given the tremendous amount of people watching soccer, we expect that there is a large potential market for consumer devices that can detect soccer highlights in recorded matches.

Much research has been done on the topic of automatic detection of highlights in soccer and other sports. However,

only a few of these approaches (for example, the system described in [1]) have actually been implemented in consumer applications for entertainment purposes. The challenge in this topic is that it is difficult to make sure that users are given all highlights that they expect, even if the performance of the automatic highlight detector is close to 100%. As an additional challenge, users will inevitably compare the automatically detected highlights with the highlights that are manually selected and commented on by professional sports journalists. The soccer highlights application developed by Philips [2] overcomes this issue, bridges this gap by providing user-friendly interaction to users, while taking into account the gap between the accuracy of technology and users' high expectations.

Since usability [3] is an essential feature for this application, we start from understanding the essential needs of the target group: soccer fans. To understand the users' needs, we have conducted interviews (with nine participants) before implementing the prototype. In the interview, questions regarding the participants' watching experience, motivations, and wishes were asked and answered. Additionally, the participants were asked to rate approximately 20 soccer events on a five-scale preference level, from '*should be in the highlights set*' to '*no need to be in the highlights set*'. Participants were allowed to add more events if they wished. Several types of events in the soccer matches were ranked '*should be in the highlights set*' by all nine participants: goals (as well as own goals), red cards, penalties, and big scoring chances. Subsequently, we identified user requirements regarding the type of events, laid-back watching and feeling-in-control.

Finally, we have tested this prototyped application with nine participants. The test uses qualitative approach to measure the participants' satisfaction and to gather their subjective opinions and motivations. In this paper, we mainly reports the latter evaluation rather than the initial interviews.

The structure of this paper is as follows. Section 2 introduces the user interface. Section 3 describes the user evaluation and the results. Conclusions are in section 4.

2. APPLICATION INTERFACE

We prototyped an application that incorporates the algorithm [2] for automatic detection of highlights in broadcasted soccer matches. The application detects the highlights in a match and uses markers in a timeline to show where in the match the highlights are located (see Figure 1). To allow the user to be in control of the level of detail, the user can select an 'importance level' in the range of 1 to 3. If the first importance level is selected, only the most important events are shown (around 10 per match). If the user increases the level, additional, less important events will be shown (around 25 in total in level 2, and around 50 at level 3). It should be noted that 'importance' here only reflects the reaction of the crowd in the stadium, and

does not have a semantic meaning, i.e., there is no guarantee that all goals will be in the first level.

The prototyped application (see Figure 1) runs on a Media Center PC, and a standard remote control (see Figure 4) is used for all interactions with the application.

The interactions that the application offers are:

- Switching between the two modes: navigation mode and auto-skip mode (see below)
- Jumping to the previous or next highlight
- Adjusting the level of highlights (which results in more or less highlight segments being shown)

In order to address the needs of laid-back watching while still allowing the users to feel in control, the application provides two modes, namely a navigation mode and an auto-skip mode. The navigation mode (see Figure 1) allows normal playback of the soccer match with the additional possibility for the user to jump to the next or previous detected highlight in the match. The auto-skip mode, on the other hand, by default only plays the detected highlight segments without manually jumping to the next segment. Furthermore, users are given the freedom of adjusting the number of highlight segments by switching between the three levels of highlights.

In the following sub-sections, the major layout elements (i.e., the navigation bar and the legend) of the application are introduced.

2.1 Navigation bar

The navigation bar, enlarged in Figure 2, shows the location of highlights for the entire recording: the detected highlight segments together with their 'importance' levels (red markers); and the current playback position (white marker). The markers' difference in intensity indicates the 'importance' level of the highlights, with higher intensity indicating more 'important' highlights.

Figure 1. The application in navigation mode.

Figure 2. Close-up of the navigation bar. The markers indicate the detected highlights, with their intensity representing their level of 'importance'.

Figure 3. Close-up of the legend below the navigation bar. The most left circle is red.

Figure 4. The remote control used in the user test. The red button on the remote control is used to switch between the modes. The left and right buttons are used to navigate to the adjacent highlight segments, while the up and down buttons are used to adjust the number of highlight segments shown.

3. USER EVALUATION

3.1 Research questions

Our user evaluation aimed to answer the following research questions:

- What is the participants' satisfaction of the application?
- What is the participants' satisfaction with the highlights delivered by the application?
- What are the usability problems of the major interaction features that are used in the application (i.e., the two modes and the navigation between highlights)?

3.2 Participants

The user test was conducted at a lab which aims to resemble a living room. A total of nine participants were invited to participate. All nine participants were male who watch soccer matches frequently. We are aware of the importance of the type of participants, and the influence of factors such as gender, interest in soccer matches, and professional background; however, due to time limitations, it was impossible to study the influence of all of these factors on the results of our study.

3.3 Material

The equipment used in the test included a 32-inch wide screen Philips TV, a Media Center PC and a remote control (Figure 4).

We used a questionnaire with a combination of open and closed questions. The closed questions were answered on a five-point Likert scale, ranging from 'disagree' to 'agree'. Nine questions target at the participants' satisfaction with the application. Two questions focus on the satisfaction with the highlights delivered by the application. Due to the space limitation, we could not list all the questions here. Some of them are introduced in the Section 3.5.

We used open questions to get feedback on the interaction features of the application, and on the general preferences and motivations of the participants. The questions about the general preferences and motivations were the following:

1. What do you most like about the user interaction in the application? Why?
2. What do you not like about the user interaction of the application? Why?
3. What do you miss about the user interaction in the application? Why?

Figure 5. A participant during the test.

3.4 Procedure

The test took about one hour. The participants took part in the test individually. The experiment leader was present in the same room during the whole test. She conducted interviews and otherwise took notes of the participants' interaction behavior. The test started with a short introduction and was followed by four sessions: a try-out session (two minutes), a limited time session (five minutes), a no-rush session (fifteen minutes) and a final interview. The motivation of designing the limited time session and the no-rush session is that we envisioned different usage scenarios, and we would like the participants to imagine themselves in the two corresponding contexts. A different soccer match was used in each session.

During the introduction, the context of the application (i.e., a feature of a consumer electronics device in the home) was explained. An interview followed in order to find out the participant's watching experience and any prior expectations he might have from such an application. It is interesting to observe that most participants already became excited during the conversation.

In the try-out section, the major features of the application were verbally introduced without showing the actual layout on the screen. The try-out session was used to let the participants get familiar with the interface and the remote control. All participants watched the same soccer match in the try-out session.

The try-out session was followed by two sessions, in which the participants were asked to interact with the application under two real-life situations: 1) limited time session: the user is in a hurry and wants to quickly watch the most exciting moments in the match; 2) no-rush session: the user does not need to rush and wants to enjoy the freedom of navigating through and exploring the whole match. Five matches from the UEFA Champions League 2006 and FIFA World Cup 2006 were used in the test. Two matches were used for the limited time session, and two different matches for the no-rush session. For each of the session, the participants could choose either one of the two available matches. After the no-rush session, participants were asked to fill in the questionnaire about their satisfaction on various aspects.

At the end of the session, a short interview was conducted to acquire a general impression of the participants' preference and motivations. The experiment leader took the notes so that the participants can verbalize their feelings and thinking.

Figure 5 shows the test setup and a participant in the test, controlling the application with the remote control.

3.5 Results and Analysis

3.5.1 Participant satisfaction

The evaluation shows that the participants were very satisfied with the application, especially in its ease-of-learning, ease-of-use, and the feeling of being in control.

'Ease-of-learning' is rated positive¹ by all participants. Also, all participants ranked positively on the question: 'It is clear to achieve my goals'. In the try-out section, all participants verbally conveyed that they understood the interaction. The participants commented that the control is easy and that there was no need to learn the function of all buttons on the remote control (only five out of 44 buttons were used by the application).

The participants showed positive satisfaction using the application. Eight out of the nine participants ranked positive on the question 'I would like to have this application'. Additionally, six out of the nine participants ranked positive on the pleasantness of operating the application. The easiness to correct interaction mistakes is perceived positively. Six out of the seven participants ranked positive on the question 'I can easily correct my interaction mistakes'. Two participants commented that they did not make any mistakes and therefore cannot judge this aspect. In the final interview, the participants commented that they liked the simple interface, easy interactions and freedom of control (e.g., adjustment on the level of highlights).

Regarding to the conformity with expectation question: 'The application works according to my expectations', the results show an interesting finding. Only four participants ranked positive on the question. This can be explained from the interviews. All but two of the participants held prior expectations for a soccer highlight application, namely:

1. The application should not miss any goals.
2. All highlights should be accessible with easy interaction.
3. It should be easy to choose different soccer matches.
4. Studio talk before and after the match should be included.
5. Advanced interactive features should be present, e.g., the application can indicate all goals scored by a specific team.

It is clear that, except for the first expectation, these expectations are either not addressed (3, 4 and 5) by the application at all or for a few participants not addressed in a completely satisfactory manner (2). This is explained in the next section. Nevertheless, it suggests that users' expectation seems to be one of the key influencing factors of their satisfaction.

3.5.2 Participant satisfaction with the highlights

During the first user interviews, we found that the users' satisfaction with the highlights is an important factor influencing their overall satisfaction with the application. Therefore, we investigated how they are satisfied with the highlights produced by the application and what the rationales behind it are. Two relevant questions were included in the questionnaire:

¹ We consider as a positive scale the scales four and five from a five-point Likert scale.

1. 'I have the impression that I have watched all the highlights'
2. 'I have the feeling that some highlights are missing'

The result of these two questions shows a discrepancy. Whereas in the limited time session only two out of the nine participants agreed with the first question, agreement with this question increased to six out of nine in the no-rush session. Likewise, after the limited time session only one out of nine participants had the impression they did not miss any highlights, whereas after the no-rush session, this increased to five out of nine. This can be explained by the fact that given more time and freedom in navigation and exploration, they could watch all the highlights they expected; therefore, the satisfaction with the offered highlights increases. In other words, the application should be positioned as a system that provides freedom of control, rather than one that gives an automatic summary.

The following remarks from the interviews explain the reason why not every subject is satisfied:

- In the case of red and yellow cards, the highlight picked by the algorithm corresponds to the moment that the referee shows the card rather than to the event that caused the action of the referee. However, most of the test participants stated that they would like to see the events leading up to the card being shown. This shows a further tuned insight.
- Participants have diverse definitions of what 'highlights' are². Half of the participants associate highlights with events, such as goals, missed chances, controversial decisions and more. Others gave open definitions, such as 'anything that is interesting'. One subject stated that background stories are highlights as well.
- Some participants have already seen summaries of the same matches we used in the test on dedicated broadcasted TV programs, in which a panel of experts discussed hand-picked highlights. Based on their memory of these broadcasts, they held prior expectations of what the application should deliver.

3.5.3 Interaction features

The two modes (i.e., navigation mode and auto-skip mode) are seen as useful. However, due to the overlapping interactions (i.e., navigating through the highlights is allowed in both modes) in the two modes and lack of easily perceived feedback (i.e., the icon indicating the mode is not recognizable), half of the participants found it confusing to distinguish between the two modes. However, three participants appreciate the two modes if given distinguishing features and clearer feedback.

The participants commented positively on the highlight navigation control. Its main advantages are flexibility by adjusting the highlight levels and efficiency for watching exciting moments within a short time. Five out of nine participants commented that the markers' difference in intensity is either confusing or not needed at all. Our initial motivation for using the difference in intensity was to indicate the importance level of the detected highlight segment. However, the participants associated the marker with highest intensity with a goal segment, and they got disappointed if it turned out not to be the case. This indicates an important new user

requirement regarding the use of color intensity on the markers. One subject commented that he does not want to see the markers at all because it kills the surprise of the match.

4. CONCLUSIONS

In this paper, we have presented the results of the user evaluation of a prototyped consumer-level soccer highlight application.

The application was designed using a user-centered approach. A preliminary interview was conducted to identify user needs, from which the essential user requirements were used to guide the design of the application. Subsequently, the application offered users the option of switching between two modes (navigation mode and auto-skip mode), jumping to adjacent highlights, and adjusting the level of highlights.

In order to evaluate participants' satisfaction and the usability of individual interaction features, a user evaluation followed. The evaluation was performed with nine soccer fans. The participants were asked to interact with the application in two real-life situations: 'in a hurry' and with 'plenty of time'. Closed questions were used to measure the participants' satisfaction on a Likert scale. Interviews at the beginning and end of the evaluation were used to determine their watching experience, expectation, subjective opinions and motivations.

We found that the participants' satisfaction with the actual highlights depends on the time they have available. When given limited time, some participants have the feeling that they did not watch all the highlights. However, this changes drastically as soon as they are given unlimited time. In that case, the majority had the impression they had watched all the important moments of the match. This shows that this is a valuable application for navigation and exploration of soccer recordings.

Despite the difficulty in bridging the gap between what users expect of a highlight detection application and what it can deliver in practice, the application is very positively perceived by the participants, as has become clear from the qualitative assessment. Although the group of test participants was small, it was still clear that they positively appreciate the ease-of-learning, ease-of-use, and the feeling of being in control. In fact, nearly all participants answered positively when asked whether they would like to have this application.

Important new user requirements were found in the evaluation, which were not fully addressed in the preliminary interview. The new findings indicate that in addition to the major user requirements identified from the preliminary interviews, the application needs to resonate by optimizing expectation management and using attractive design. Furthermore, existing requirements resulting from the preliminary interview were further tuned by the evaluation.

5. REFERENCES

- [1] Otsuka, I. et al. A Highlight Scene Detection and Video Summarization System using Audio Feature for a Personal Video Recorder. *IEEE Transactions on Consumer Electronics*, 51, 1 (Feb. 2005), 112-116.
- [2] Zoetekouw, B., Fonseca, P. and Funken, R. A Fast and Efficient Highlight Detector for Soccer. *Proceedings of the International Conference on Consumer Electronics (ICCE 2008)*, Las Vegas, USA, 2008.
- [3] Nielsen, J. *Usability Engineering*. Morgan Kaufmann, San Francisco, 1994.

² Each subject was asked to give a definition of highlight during the interview.