eMir: Digital Signs that react to Audience Emotion

Juliane Exeler, Markus Buzeck, Jörg Müller

{j.exeler, m.buzeck, joerg.mueller}@uni-muenster.de

Abstract: In this paper we present eMir, digital signage (public electronic displays) that show human faces which react to audience emotion. Using a camera installed at the sign, the system observes the audience and detects whether someone watches the display via face detection software. The face detection is able to classify facial expressions and determine gender. This information is used to let a human character on the screen react accordingly and encourage interaction with the face/sign. The system has been deployed for one month on a digital sign in a university building. We present experiences with the system, our findings from the collected interaction data and results from interviews with eight users.

1 Introduction

Public electronic displays become more and more important as one can conclude from the increasing number of displays installed in public places and buildings. Digital signage is a cheap and easy way to present customizable information. Shopkeepers are increasingly installing mid- to large-scale displays in order to attract the customers' attention and advertise new products or special offers. But there are also places where displays are installed for non-commercial purposes such as information or entertainment. These places are for example waiting areas, e.g. at the doctors or at the citizen bureau. Studies have shown that public displays are often not able to attract the passerbys' attention except some displays installed in waiting areas[1]. Although there are style guides for display size, height, color composition of the content and the content itself[2] to make displays more perceivable many digital signs are still not noticed. We started exploring how to make displays more perceivable and enjoyable that people get attracted by it. In a field study we presented content which was rated by the view time of the users collected with face detection software[3]. This was a kind of unconscious, indirect interaction because the user didn't know that he was influencing the content to be shown. In our current study we try to find out, how videos of human faces that mirror the users' emotion and gender can attract the users attention and lead to direct interaction. As a first step, we want to examine how users react to our electronic mirror system (eMir) and if the people try to interact with it. We think that in future interaction with digital signs is one way to make digital signage more attractive and perceivable.

2 Related Work

Experiences from our previous work with digital signage have shown two major shortcomings of public displays this project is supposed to overcome. First we observed the effect of Display Blindness[1]. People tend to ignore public displays intentionally or subconciously, which is influenced by negative expectations towards contents. Second we often experience that people react very reserved towards digital signs. In particular when digital signage is equiped with sensors measuring context and audience attention, people often do not feel comfortable[3]. This is partly due to privacy concerns, but often there are still negative feelings when the rational concerns are dispelled and people trust the person who installed the system.

Against this background we propose that eMir can have positive effects on user acceptance in several ways. Evidence from fields such as cognitive, evolutionary, and developmental psychology, as well as cognitive neuroscience, has pointed to the special nature of human perception of faces. Faces begin to be seen as a separate class of objects within the first six months of life[4]. Christian and Avery presented the Smart Digital Kiosk Project which is an electronic information display showing a human face that looked at the users and talked to them[5]. This project received enthusiastic feedback from the users. Faces tend to attract our attention. So we propose that a public display showing a life-sized human face with natural movement and expression will catch the people's eye in the first step. In the next step we propose that the presence of a calm and friendly face smiling back at the user in an unobstrusive way might contribute to a comfortable feeling. And finally there is the aspect of interaction with the face on the sign, that might appeal to peoples curiosity. Shawney, Wheeler and Schmandt[6] suggest to use unobstrusive means for sensing user intention and activity patters for interaction in transitional spaces. They investigated proximity and movement as an interface to maintain a casual and natural interaction. By means of simulating non-verbal interpersonal communication through facial expression we want to explore a similar approach. Advanced face detection technology allows us to make use of an even more subtle form of interaction.

3 A Concept for interactive Digital Signage

According to the capabilities of the face detection software, we define five emotional states of users the system differentiates: a neutral state, happiness, sadness, angriness and surprise. For each of these categories we also consider the gender as being either undefined, male or female. So we end up with 15 different scenarios that might occur with a neutral state and undefined gender as the default state of the system. The current state is determined by the modes of both emotion and gender detected in the camera frames within a certain time slot. According to the current state a small video sequence is selected that fits the situation. Each interaction, which we consider as a variation from the default state, is initiated by the user. The digital sign reacts passively by reflecting the same emotion through the facial expression of a human of the same gender. Happiness shall be encouraged by a happy face. A sad face shall receive empathy by another sad face. An angry face

on the screen shall make angry users laugh. The same applies to surprised faces. A user with a neutral facial expression will not be involved into any interaction by the sign. The face on the sign will remain in a default neutral state, looking around the hallway quite bored.

4 Implementation

The prototype consists of a digital sign installed in a university institute with approx. 60 employees. It is located in the hallway and is frequently passed by students and employees. Before being used for this project the sign was used for the university information system iDisplays[7] which is installed on several other signs in the building. Audience emotion is measured via camera installed on top of the sign. We use the Real Time Face Detector from Fraunhofer IIS[8] to analyze the video stream. This software is able to detect multiple faces within the camera image during runtime. It also determines gender and emotion from the facial expression. Depending on the emotion and the gender that has recently been detected most frequently, a small video sequence is selected that shows a face with similar attributes. Therefore we produced videos each showing a person in a portrait close-up. We tried to position the persons and cut the videos in a manner that allows us to play the videos in loops and one after the other without breaks looking as natural as possible (see figure 1). As these videos need to have a minimum length to reveal a certain emotion naturally,

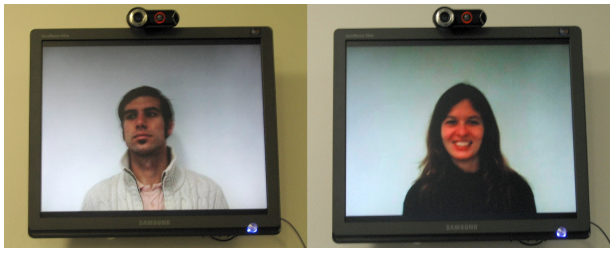

Figure 1: Faces presented on the eMir display.

we do not interrupt a video once it has started. As a consequence the system may react with a delay with maximum length equal to the length of the current video. Whenever a new video is started, either a default video or an interaction video is triggered by the face detection, the attributes of the video are stored in a database in order to be able to analyze user behavior and interaction.

5 Data Collected

During the deployment phase of one month (24/7) about 111000 videos had been played. For each displayed video an item containing an id, gender and emotion information and

the timestamp was inserted into a database. We can neglect most of the database entries for our analysis because in these cases default videos have been played. Default videos have been displayed whenever the face detection software could not identify any human face in front of the sign. Then a video with the face of a male person and a neutral emotion was shown but was inserted into the database with the attributes gender is undefined and emotion is neutral. The collected data allows us to investigate how passerbys react to the eMir system and if there is any interaction between the digital sign and a passerby who can then be considered a user. This interaction is recorded as sequences of database items representing non default videos.

In 569 cases a non default video was played. We grouped the collected data into 125 sequences of videos that indicate some interaction has taken place. An interaction sequence ends with a default video shown, a new interaction sequence starts when a default video is succeeded by a non default video. Thus as a first result we can state that during the deployment phase every day (including the weekends when there is no activity in the building) there were five interaction sequences in average. This is more than we expected from our last study[3], where we got some negative feedback by users concerning display contents and privacy concerns. Having a closer look at the duration of each interaction sequence we found out that in average every interaction takes about 30 seconds. Hence we can say that the shown faces attract attention and actually stimulate interaction. Depending on the duration of the individual displayed videos such an interaction comprises two to four videos. Classifying the 125 interaction sequences into four subsets according to the number of videos we see that in most cases (51 times) two, three or four videos build one interaction sequence. For 44 times a sequence of interaction only consists of one item. Here we have to note that the face detection software also detects a face which is just turned towards the digital sign, regardless of the fact that it may not look at it. In 19 cases an interaction sequence comprehends five up to ten video items. For interactions sequences with more than ten video items eleven blocks were identified (see table 1).

Table 1: 569 items of non default videos could be assigned to 125 sequences of interaction

Interaction sequences with...			
1 item	2 - 4 items	5 - 10 item	more than 10 items
44	51	19	11
$\sim$ 30 seconds each interaction			

The analysis of the collected data shows that there has been a trend in the frequency of interaction sequences. Splitting the deployment phase into two parts of equal length the result is that in the first phase there were 33 (75%) interaction sequences with one video item, 30 (58%) with two up to four items, 15 (78%) with five up to ten items and six (55%) with even more than ten items. This indicates that the interest for the eMir system decreases with time. One reason might be that the default video is always the same. Different default videos with alternating faces might also attract the passerbys' attention in the long run.

6 Interviews

After running eMir for one month, we conducted semi-structured interviews with 8 employees and regular visitors of the institute. The interviews were partially transcribed and evaluated using affinity analysis[9]. Perception and acceptance of the system was mixed. The presentation of a human face on the public display was appreciated by seven users. These stated that they found it funny and amusing. This was mainly due to the fact that the presented faces were familiar to the users. Five users said it made a big difference if they knew the face or it was a foreign face. "The first time it was appealing to see him there. That made me stop at the sign to see what happens." (U2) "It would be curious to see more familiar faces." (U7) Two interviewees stated that any human face attracted attention. We also wanted to investigate if the presence of a human face on a screen compared to textual inforation conveys a different atmosphere. Five interviewees said that the system had a positive effect on the atmosphere. "It enlightens the mood." (U4) Drawing a comparison to other systems which were previously installed on the same device, one user said: "There is an increase in the attraction of user attention from the iDisplay information system, over video presentation to faces being presented. It is appealing on an emotional level." (U2) "It is a warmer, more cordial atmosphere, as if there was just information running on the screen." (U5) When it comes to interaction, four out of eight interviewees stated they had intentionally interacted with the system. "It is funny to stop for some minutes in front of the display and to try it. It is exciting and nice playing with it." (U4) "I tried it. I was curious to see more images." (U7) Two users did not recognize it was an interactive system. They were not stimulated to watch the display long enough to evoke any reaction from the system. Another two users deliberately did not interact. One of these two gave her negative attitude towards cameras as a reason. "I don't like standing in front of cameras. I feel like being observed." (U3) Moreover six out of eight interviewed users said to have witnessed or have personally been involved in situations in which several people were interacting with the system and with each other. There were discussions of people altogether exploring the functionality of the system. "We were standing there several times and tried to imagine how it works. We were talking about it, because we found it really funny." (U2) Figure 2 illustrates a similar situation.

7 Conclusion

In this paper we have presented eMir, a digital signage system that shows human characters reacting to audience emotion. The system was deployed for one month and evaluated through analysis of logging data and interviews with users. We wanted to investigate how such a system is accepted by the user, whether it stimulates interaction and whether the mere presence of a human face on a screen influences the atmosphere. The analysis shows that most users react with curiosity to the display and try to interact. This effect is even stronger when the presented face is familiar to the user. Interaction sequences took place every day and many users spent some time to trigger the different emotions to be shown. However there are also some users who reject such a system because of distrust towards

Figure 2: People observing the face on the eMir display.

the technology. In particular the camera being visible and obviously running is a big issue, although the sign suggests for what it is used. To those users who do not have any privacy concerns, the face on the screen looking around the hallway is funny and effects their mood in a positive way. At least compared to pure information being presented on the sign the system appeals to the users in a more emotional way and thus more effectively. Perhaps this could be an approach to make digital signage more perceivable in general.

References

[1] J. Müller, D. Wilmsmann, J. Exeler, M. Buzeck, A. Schmidt, T. Jay, and A. Krüger. Display Blindness: The Effect of Expectations on Attention towards Digital Signage. In *Proc. of Pervasive 2009*

[2] E. Huang, A. Koster, and J. Borchers. Overcoming assumptions and uncovering practices: When does the public really look at public displays? In *Proc. of Pervasive 2008*

[3] J. Müller, J. Exeler, M. Buzeck, and A. Krüger. ReflectiveSigns: Digital Signs that Adapt to Audience Attention. In *Proc. of Pervasive 2009*

[4] C. A. Nelson. The development and neural bases of face recognition. In *Infant and Child Development*, Volume 10, Issue 1-2, pages 3–18, 2001.

[5] Andrew D. Christian and Brian L. Avery. Digital smart kiosk project, In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems, 1998*

[6] N. Shawney, S. Wheeler, and C. Schmandt. Community Aware Portals: Shared Information Appliances for Transitional Spaces. In *Personal and Ubiquitous Computing*, 5 (1):66-70, 2001.

[7] J. Müller, O. Paczkowski, and A. Krüger. Situated public news and reminder displays. In *European Conference/Symposium on Ambient Intelligence*, pages 248–265, 2007.

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

[9] H. Beyer and K. Holtzblatt. Contextual Design : A Customer-Centered Approach to Systems Designs, 1997