Supporting Children with Special Needs through Multi-Perspective Behavior Analysis

Florian Alt,
Andreas Bungert
University of Duisburg-Essen
paluno - Ruhr Institute for Software Technology
Schützenbahn 70
45117 Essen, Germany
{firstname.lastname}@uni-due.de

Bastian Pfleging,
Albrecht Schmidt
University of Stuttgart
Institute for Visualization and Interactive Systems
Pfaffenwaldring 5a
70569 Stuttgart, Germany
{firstname.lastname}@vis.uni-stuttgart.de

Meindert Havemann
TU Dortmund
Faculty for Rehabilitation Sciences
Emil-Figge-Str. 50
44227 Dortmund, Germany
meindert.haveman@tu-dortmund.de

ABSTRACT

In past years, ubiquitous computing technologies have been successfully deployed for supporting children with special needs. One focus of current research has been on post-hoc behavior analysis based on video footage where one or multiple cameras were used to review situations in which children behaved in a certain way. As miniaturized cameras as well as portable devices are becoming available at low costs, we envision a new quality in supporting the diagnosis, observation, and education of children with special needs. In contrast to existing approaches that use cameras in fix locations, we suggest to use multiple mobile camera perspectives. In this way observation data from fellow classmates, teachers, and caregivers can be considered, even in highly dynamic outdoor situations. In this paper we present *MuPerBeAn*, a platform that allows multi-perspective video footage from mobile cameras to be collected, synchronously reviewed, and annotated. We report on interviews with caregivers and parents and present a qualitative study based on two scenarios involving a total of seven children with autism (CWA). Our findings show that observing multiple mobile perspectives can help children as well as teachers to better reflect on situations, particularly during education.

Keywords

Autism, ubiquitous computing, cameras, mobile devices.

Categories and Subject Descriptors

H.5.1 [**Information interfaces and presentation (e.g., HCI)**]: Multimedia Information Systems — Video (e.g., tape, disk, DVI); K.4.2 [**Computers and Society**]: Social Issues — Assistive technologies for persons with disabilities

General Terms

Human Factors.

MUM '11, December 7-9, 2011, Beijing, China

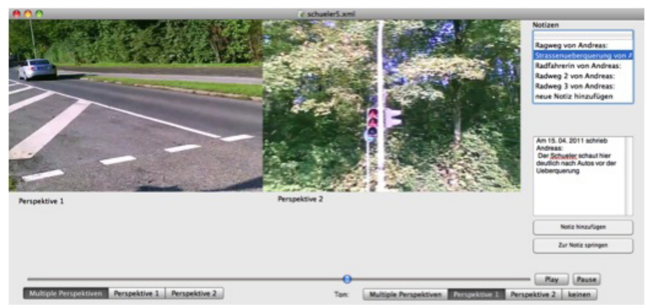

Figure 1. Interface of the multi-perspective behavior analyzer (*MuPerBeAn*): We used the prototype in a traffic scenario to assess the traffic competency of children with autism.

1. INTRODUCTION

Behavior of children with special needs is often difficult to understand without knowing the context of a given situation. In a sample scenario, a mother might be in a playground with her child suffering from autism. As she tells her son that she loves him, the child starts crying – an unexpected reaction unless one is aware of the fact, that the mother just came back from a stressful business meeting causing a rather aggressive body language. In another scenario a child in class, during a phase of silent work, all of a sudden loses concentration of the task at hand and is being distracted.

Such situations are currently difficult to analyze as the complex context of a situation is not being captured, However, especially in the context of special education this is highly desirable in order to better understand the functions of inappropriate behavior. Previous work looked into how this issue could be tackled based on a post-hoc analysis of data obtained from capture technologies, such as cameras [10][4]. So far mainly a single camera perspective has been subject to research, or, if using multiple cameras, investigations were conducted with cameras mounted in fix locations. However, we see large potential in using video footage from multiple *mobile* perspectives that can later be reviewed in a time-synchronized manner. This allows the context of indoor (e.g., class room) as well as highly dynamic outdoor situations (e.g., pedestrian crossing and approaching traffic lights) to be reviewed and analyzed. With miniaturized cameras embedded into glasses becoming available at low costs it is finally feasible to use capture technologies in a very unobtrusive, yet powerful manner.

To exploit this idea, we present a prototype called *MuPerBeAn* (Multiple Perspective Behavior Analyzer). The system takes multiple video streams from mobile cameras, time-synchronizes them, and allows them to be played back in parallel as well as to be annotated on a time-framed basis using multiple devices. To develop and evaluate our system we cooperated with a school for children with special needs in Duisburg (Germany) over the course of 6 months. The system was developed iteratively through expert evaluations and discussions with teachers, caregivers, and parents. Finally, we deployed MuPerBeAn in a real-world setting.

The contribution of this paper is twofold:

1. We present and discuss an approach as to how mobile, multi-perspective (post-hoc) video analysis can be used to analyze the behavior and support the observation and the education of children with special needs.

2. We present a prototype implementation and report on a qualitative evaluation among seven pupils based on two scenarios.The evaluation included a field trial and discussions with teachers, caregivers, and parents, uncovering potential application domains, strengths, and challenges.

2. RELATED WORK

Supporting people with special needs through ubiquitous technologies is a highly active research area. Examples of how caregivers of children with special needs can be supported are presented in [3]. Hayes *et al.* presented social and technical design issues associated with using tracking technologies to assess the effectiveness of educational and therapeutic interventions for children with special needs [4]. Technological aids such as visual or tactile output devices have been explored in [2]. CareLog supports functional behavior assessment by selectively recording audio and video data and allowing material to be manually metatagged [5].

More specifically looking into camera-based approaches, Patel *et al.* presented the ContextCam [8], a video annotation system using multiple sensing techniques to add metadata about when, where, and who has been recorded in a video. Further single-camera-based approaches include [7] where a portable system consisting of a camera and a portable PC is presented that allows various facial expressions to be tracked, recorded, analyzed, and presented either in real-time or post-hoc. In contrast, self-cam [9] is a wearable camera system that allows the facial expressions and head movements of the wearer to be analyzed. The wearable, sensor-equipped SenseCam [6] provides means to capture and later analyze series of still images of the wearer's surrounding.

Previous work showed the potential of using video analysis in children care. However, to the best of our knowledge, no approach exists that supports the examination from multiple mobile perspectives. In the remainder of this paper we show, how a further perspective could help to assess the context of a given situation and hence provide valuable additional information – for example *"What caused a reaction or distraction?"* or *"Who else was present and what was everybody doing"*.

3. CONCEPT

In the following section we sketch our idea and report on expert interviews we conducted in order to identify potential application domains and requirements.

3.1 Idea and Concept

The belated development of language skills as well as problems with regard to social interaction often lead to situations, where behavior of children with special needs are difficult to understand. To support special education we hence suggest taking into account both the personal view of the child as well as of the situation's context. This can be done based on the view of people nearby, such as other children, parents, teachers, and caregivers as it is very likely that the person or object that might cause the distraction can be identified in the visual field. For a post-hoc video analysis we suggest to use cameras that can be worn in an unobtrusive way (e.g., glasses) and which are, in the best case, unperceivable. To obtain usable data, the captured video should cover the field of vision of the involved persons. For the post-hoc analysis we suggest a user interface that allows a time-synchronized presentation of multiple perspectives and a way to easily annotate the situation to later retrieve similar or related scenes.

3.2 Informing the Design

Prior to the realization of our concept idea, we conducted a series of expert interviews to inform our design. The aim was to unveil potential application domains and tailor our prototype to the needs of the different stakeholders. We interviewed 3 psychologists from the University of Dortmund (Germany) as well as 6 teachers, caregivers, and parents at a schools for children with special needs in Duisburg (Germany). To sketch our concept and illustrate the envisioned benefit, we presented technology probes and potentially suitable devices, and showed sample video we previously recorded. Based on the subsequent discussions we identified the following application domains:

Diagnosis. A very basic problem is to unveil what causes certain reactions. As mentioned earlier, it is often not possible to easily tell which factors caused a child to behave in a certain way (e.g., a child taking a wrong bus several times). Teachers and parents felt that a post-hoc analysis of multi-perspective video material could be valuable as on one hand the view of the child could help to identify his focus and orientation point, on the other hand additional perspectives could reveal peripheral clues and the situation's context.

Education. An important part of daily school life is the education of children for safely moving outside the school environment. A central skill is traffic competency. Children are instructed how to behave correctly at a traffic light or crossing. In this context, teachers considered it to be helpful to find out how children orient themselves when crossing a street, whether they, e.g., established eye contact with the driver of a car, and if it was possible to rate the traffic competency of a child based on video data only. Additionally, cycling classes were expected to be promising.

Observation. Parents and teachers were also interested in long term observations of children in order to understand behavior patterns. In this way it could, e.g., be analyzed, how a child behaves in a certain situation based on a given context. Possible application scenarios were breaks or situations were children had to solve tasks in classes and on the campus.

4. PROTOTYPE

To investigate the potential of our concept we implemented *MuPerBeAn* (Multiple Perspective Behavior Analyzer). Our software allows videos and sequences of still images from arbitrary sources to be processed. Information on the video footage and notes are stored in XML files (title, content, author, time-stamp, synchronization). Videos of multiple perspective are synchronized by aligning them at a frame which shows an object (e.g., a wrist watch) that has been captured at the same time. As we envisioned the system to be used for a long-term analysis it supports storing multiple situations, which can easily be retrieved based on time-stamp or associated keywords. MuPerBean allows content to be annotated on a frame level. The annotations can be exported into a text file.

The user interface is depicted in Figure 1. The main content window shows the available video files for the current scene. Buttons at the bottom allow for switching between the multiple-perspective mode and the full screen view of the different perspectives. Similarly, audio can be turned on/off independently. As it is sometimes helpful to closer analyze certain elements of the scene, a slider can be used to zoom into the video. By using the slider at the bottom, the screen can easily be browsed forward and backward. On the right side, the annotations available for the video and the content of the selected annotation can be seen. A button is available to generate a new annotation, which is associated with the current frame of the scene. The button "jump to scene" allows to instantly jump to the point the selected annotation refers to.

MuPerBeAn is currently available for MacOS X, the iPad and the iPhone. Especially mobile devices such as the iPhone or iPad allow the data to be easily brought along and instantly be reviewed, e.g., by specialists or also be review them in front of the class.

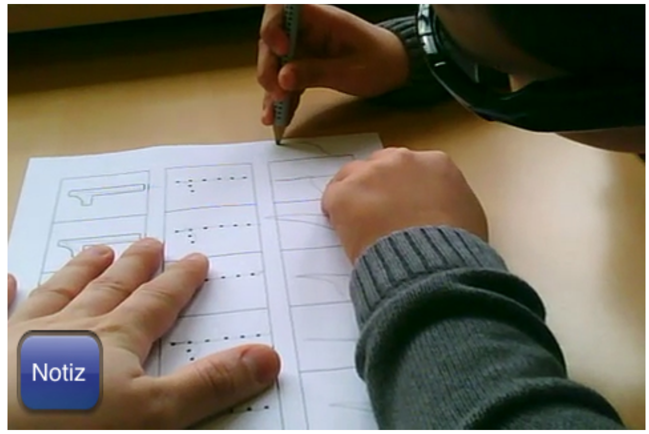

Figure 2. MuPerBeAn video analysis of scenario 2: Child and caregiver in a math class using the camera glasses.

5. EVALUATION

In the following we report on a field study with our prototype, which aimed at uncovering potential benefits and pitfalls, to explore uptake and acceptance, and to gather qualitative feedback from involved people. We were especially interested in whether our system provided helpful means to support teachers during their work with the children and how they used the system. Hence we were interested if the teachers could easily use it in the intended way.

5.1 Preparations

Prior to the study we tested different cameras for recording video footage, including cameras attached to clothes, head-mounted cameras, and camera glasses. Therefore we conducted a pilot study with four participants (students from our courses) whom we used to collect sample data with the different devices as they walked on a pre-defined route through the surrounding neighborhood of the University campus. We analyzed the videos with regard to quality (audio, video) and covered field of vision. Our subjects reported that camera glasses were most comfortable. The resolution (720x480p) and audio quality were satisfactory and microSD cards (16 GB) allowed several hours of video footage to be recorded.

5.2 Scenario 1: Traffic competency

Teachers and parents at the school where we had previously conducted the interviews showed large interest in testing the application. After demonstrating the prototype, several scenarios were discussed with interested parents and teachers as to where such a system could be usefully deployed. The teachers saw large potential in assessing and supporting the children's traffic competency. They were especially interested in whether it was possible by using the mobile multi-perspective video (a) to rate the traffic competency of a child (observation), (b) to determine a child's points of orientation at a traffic light to understand a certain behavior (diagnosis), and (c) to post-hoc discuss the video material with other children in class to outline potentially wrong or dangerous behavior (education).

For the study the teachers chose, given the permission of the parents, six children (age 13-16 years) to participate in the study: two with high, two with moderate, and two with low traffic competency. The children and an accompanying teacher were equipped with camera glasses and went for a 700 meter walk around the area of the school, which included crossing a traffic light and a street. The teacher was requested to ask each child whether it checked both directions prior to crossing and whether an approaching car has been realized. The study was conducted on 5 consecutive school days. Completing the track took 9 to 13 minutes.

During the walk, the first person view of each child as well as of the accompanying teacher were recorded. The videos were time synchronized and had a duration of 9 to 13 minutes. After the study the video files were imported to the application. The teachers then used the MuPerBeAn interface to review the recorded video streams together with the children and discussed potential issues with them.

We found that the teachers were indeed able to extract valuable information from the application. The video data allowed the orientation points of the children to be detected (e.g., children used different parts of the traffic lights to orient themselves before crossing). Furthermore, for the teachers it was easily possible to detect and highlight dangerous behavior to the class, which led to discussions about what was wrong and why. The teachers stated that the system definitely allows for assessing the traffic competency of the children. Especially the second perspective gave highly valuable hints about uncertainties and teachers felt that errors can be recognized more precisely as they envisioned.

5.3 Scenario 2: Behavior in Class

The parents of one of the children asked to use the system for a longer-term observation. This gave us the opportunity to investigate the prototype over a longer period in time. Videos were recorded in several sessions over the course of four weeks in math classes (see Figure 2), during lunch breaks, during cycling lessons, and on the way to the train station. Each session lasted around one hour and 9 teachers (all female, avg. age. 38.4 years) were involved. The video material was reviewed after each session using the MuPerBeAn interface and discussed with the teacher(s) and the child. Apart from the teachers that were present, also teachers that did not participate in the very session participated in the analysis.

After four weeks we had the involved teachers fill in a brief questionnaire with regard to the perceived value as well as the usability of the system and performed in-depth interviews. All teachers considered the chosen situations to be suitable or very suitable (avg=1.22, SD=0.44, 5-Point Likert scale, 1=very suitable, 5=not suitable at all) and all of them felt that the system was helpful or very helpful to analyze the child's behavior (avg=1.22, SD=0.44, 5-Point Likert scale, 1=very helpful, 5=not helpful at all). The majority also liked the opportunity to annotate the videos (avg=1.66, SD=0.50) and to watch multiple perspectives in parallel (avg=1.44, SD=0.53). To assess the usability, we had the teachers fill in a standardized SUS [1], which revealed a high usability (84.6 points).

The following interviews showed that the teachers considered the view of the entire environment to be especially useful. One teacher reported, e.g., that the child was talking a lot to himself and that it was often not clear what he was referring to. Here, the second perspective gave valuable additional cues. Although we only used 2 mobile perspectives during the study, two teachers expressed their desire to have a third perspective showing the entire room, which they considered to be particularly useful for persons not having been present personally. Finally, the interviews revealed further scenarios, e.g., working phases, orientation on the school grounds, playing (e.g., selecting toys), and conflict situations.

6. DISCUSSION

Summarizing the results, the teachers found the system to be highly valuable when analyzing how long children were concentrated, when they lost concentration, what caused this behavior, and what happened then. However, they felt that the system should not be used permanently. For the use in a classroom scenario, teachers felt it was essential to use even more than two perspectives. Additional application scenarios include the observation of social interaction within smaller groups, especially among very young children. The overall goal would be to individually adapt the situation in class for the children.

From the view of the children we found that camera glasses are highly accepted, and that they perceived it as being fun to wear them. When showing the videos obtained during class, the children were highly focused and further classmates wanted to use the glasses immediately. Though we did not elaborate on this, the enthusiasm of the children might well be the effect of curiosity. However, this is highly encouraging as our study revealed a useful way of supporting the education of children with special needs.

7. CONCLUSION

In this paper we presented *MuPerBeAn*, a tool, which allows behavior to be analyzed based on multiple mobile camera perspectives. We built a prototype and evaluated it in a school environment. Initial results and discussions unveiled a large potential: The results of the study and the statements and reactions of both teachers and children indicate that the approach has indeed the potential to support working with children having special needs. We see the biggest benefit for education and believe that also long-term observations and diagnosis might be useful application domains. The positive feedback from all people involved as well as requests to provide the system for a long-term study are highly encouraging and motivate us to further enhance and explore the system.

Future work will focus on including additional context information (e.g., bio feedback sensors and location). This would, in combination with the multi-perspective approach, allow peaks in the stress level and what causes them to be detected. Furthermore, positioning and orientation information could be used to obtain the spatial context. Finally, it could be interesting to obtain more fine-grained data by using a mobile eye tracker.

8. ACKNOWLEDGEMENTS

We thank the teachers, caregivers, and parents of the Buchholzer Waldschule in Duisburg, Germany for their valuable feedback and their support to evaluate the potential of MuPerBeAn. Special thanks go to Barbara Wedekind (headmaster) who made the application of MuPerBeAn in classes possible. Thanks go also to all children for their cooperation during the study and their interest in using the camera glasses, which strongly motivated us to continue this research.

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 244011 and from the German Research Foundation within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.

9. REFERENCES

[1] Bangor, A., Kortum, P., and Miller, J. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction 24*, 6 (2008), 574–594.

[2] Goldsmith, T. R., and Leblanc, L. A. Use of technology in interventions for children with autism. *Journal of Early and Intensive Behavior Intervention 1* (2004).

[3] Hayes, G., Kientz, J., Truong, K., , D., Abowd, G., and Pering, T. Designing capture applications to support the education of children with autism. In *Proceedings of the Sixth International Conference on Ubiquitous Computing*, Ubicomp 2004. Springer, Heidelberg, 2004, 161–178.

[4] Hayes, G. R., and Abowd, G. D. Tensions in designing capture technologies for an evidence-based care community. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '06, ACM (New York, NY, USA, 2006), 937–946.

[5] Hayes, G. R., Gardere, L. M., Abowd, G. D., and Truong, K. N. Carelog: a selective archiving tool for behavior management in schools. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '08, ACM (New York, NY, USA, 2008), 685–694.

[6] Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N., and Wood, K. Sensecam: A retrospective memory aid. In *Proceedings of the Eights International Conference on Ubiquitous Computing*, Ubicomp 2006. Springer Heidelberg, 2006, 177–193.

[7] Madsen, M., el Kaliouby, R., Goodwin, M., and Picard, R. Technology for just-in-time in-situ learning of facial affect for persons diagnosed with an autism spectrum disorder. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*, ACM (New York, NY, USA, 2008), 19–26.

[8] Patel, S., and Abowd, G. The contextcam: Automated point of capture video annotation. In *Proceedings of the Sixth International Conference on Ubiquitous Computing*, Ubicomp 2004. Springer Heidelberg, 2004, 301–318.

[9] Teeters, A., El Kaliouby, R., and Picard, R. Self-cam: feedback from what would be your social partner. In *Proceedings of the 33rd International Conference and Exhibition on Computer Graphics and Interactive Techniques*, ACM SIGGRAPH 2006, ACM (New York, NY, USA, 2006).

[10] White, D., Camacho-Guerrero, J., Truong, K., Abowd, G., Morrier, M., Vekaria, P., and Gromala, D. Mobile capture and access for assessing language and social development in children with autism. In *Extended Abstracts of the Fifth International Conference on Ubiquitous Computing*, Ubicomp 2003. Springer, Heidelberg, 2003.