

WearCam: A head mounted wireless camera for monitoring gaze attention and for the diagnosis of developmental disorders in young children

Lorenzo Piccardi^{*}, Basilio Noris^{*}, Olivier Barbey^{*}, Aude Billard^{*}
Giuseppina Schiavone[†], Flavio Keller[†], Claes von Hofsten[°]

^{*} Learning Algorithms & Systems Laboratory, LASA, EPFL, 1015 Lausanne, Switzerland
{lorenzo.piccardi, basilio.noris, olivier.barbey, aude.billard}@epfl.ch

[†] DNNP Lab, Campus Bio-Medico, 00155 Roma, Italy
{g.schiavone, f.keller}@unicampus.it

[°] Department of Psychology, Uppsala University, SE-751 42 Uppsala, Sweden
claes.von_hofsten@psyk.uu.se

Abstract—Autism covers a large spectrum of disorders that affect the individual’s way of interacting socially and is often revealed by the individual’s lack of interest in gazing at human faces. Currently Autism is diagnosed in children no younger than 2 years old. This paper presents a new monitoring device, the *WearCam*, to help forming a diagnosis of this neurodevelopmental disorder at an earlier age than currently possible. The *WearCam* consists of a wireless camera located on the forefront of the child. The *WearCam* collects videos from the viewpoint of the child’s head. Color detection, face detection and gaze detection are run on the data in order to locate the approximate gaze direction of the child and determine where her attention is drawn to (persons, objects, etc.). We report on early tests of the camera within normally developing children. Firstly the technical characteristics of the current prototype of the *WearCam* will be described. Afterwards the type of data collected with this device with young children will be shown.

I. INTRODUCTION

The *WearCam* is one of the therapeutic devices developed within the TACT (Thought in Action) research Project¹, financed by the European Union’s NEST² Adventure Program. This project aims at developing non-obtrusive monitoring devices with toys appearance for the study on young children behavioral characteristics like: movement, attention, voice, grasping force, etc. The *WearCam* is a wearable wireless camera located on the forehead of the child, (Fig. 1). It collects video recordings during therapeutic sessions as well as free play session in which the child takes part. These videos are subsequently analyzed to locate and track the focus of attention of the child during the session. Such analysis may reveal developmental deficits, such as, for instance, reduced span of attention or averted attention to social stimuli (faces, people). Consequently a diagnose of syndromes such as Attention Deficit Disorder or Autism might become possible at an earlier stage than currently done. In such cases, the *WearCam* could also be used during therapy to monitor the child’s progresses.

This paper reports on the design of the *WearCam* and on the development of the analysis tools. Three prototypes of the *WearCam* have been constructed. They have been dispatched to three different collaborator institutions within the



Fig. 1. A normally developing child wearing a prototype of the *WearCam*

TACT project to conduct pilot study with normally developing children and children at risks.

II. STATE OF THE ART

The *WearCam* offers a very novel approach to systematic diagnosis of early attentional disorders. While there exists more precise instruments to monitor gaze direction, such as, for instance, the *Tobii* eye tracker, these are not wearable and, thus, constrain importantly the experimental setup. There exist wearable cameras, see, for example, the Eye Tracker at ISU [1]–[3], Spy camera glasses, head-mounted cameras [4], [5], or the recent chest-mounted camera called the "self-eye" at the MIT Media Lab [6]. These, however, are suited for adults, but are too heavy and too large for children. Compared to existing devices, the *WearCam* features a very small weight and greater portability, thanks to its wireless characteristic.

III. HARDWARE DESIGN

The *WearCam* has been designed to fit the head of children from 6 to 18 months of age. We took as reference the data from the Swiss National Institute of Health (Fig. 2) and considered the head perimeter to vary from 35 cm to 48 cm.

Requirements for the design phase were:

- The weight of the *WearCam* must not exceed 100g.

¹<http://tact.unicampus.it>

- It must be made of soft material so that it does not damage the child's head and is comfortable to wear.
- Aesthetic: the WearCam should not appear too technical to be appealing to child and parents².
- The system should be non obtrusive, so that the child's behavior is not influenced by the system. Thus, the system must be small in the order of a few cubic centimeters.
- The WearCam should send data wireless so that the child could be monitored during free play at the daycare center.
- Data sent by the WearCam should be readable by any current computer running Windows 2k/XP system and be provided with a user-friendly applet for the therapists to easily record and visualize the data.

The WearCam is composed of a TX 45 Light CCD Wireless Camera (<http://www.rc-tech.ch>), a battery that feeds the Camera with energy, a wireless receiver and an A/D converter with USB2 connection (Fig. 3).

The support system is composed of two belts easily adjustable to fit the child's head using Velcro bands.

The Wireless Camera is mounted in a plastic cylinder. The camera can be inclined vertically for a total of 180°, by rotating the cylinder along its axis from 45° to -135° around the horizontal (Fig. 4).

The battery and the associated electronics is contained in a plastic box located on the back of the child's head and is attached on one of the elastic bands (Fig. 3).

The TX45Light camera is typically used in miniature aircrafts for its low weight and wireless property. This camera has a good image quality and can provide 30 frames per second with a resolution of 640x480.

The camera dimensions are 27x27x38 mm. The diagonal FOV is 92° for an average of 56° and 74° vertically and horizontally respectively. The battery is a NiMH 8.4 V/ 250

²This aspect will be taken into account only in the final prototype of the WearCam.

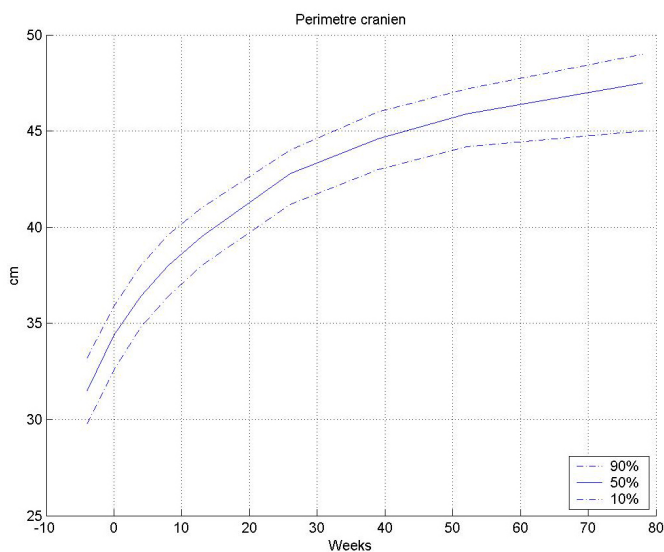


Fig. 2. Head perimeter for children aged 0 to 100 weeks

| Camera | Sensor | Weight | Fps | Resolution |
|------------------------|--------|------------|----------------------------|------------|
| TX45 Light | CCD | 16g | 30 | 640x480 |
| FOV (vert/horiz/diag) | | Wireless | Price(CHF) | |
| 56°/74°/92° | | Yes | 579.- | |
| Battery | Weight | Price(CHF) | A/D Converter | Price(CHF) |
| NiMH 8.4 V/ 250 mAh | 42g | 20.- | Dazzle Video Compressor | 100.- |

TABLE I

TECHNICAL SPECIFICATION OF THE CURRENT WEARCAM PROTOTYPE.

mAh rechargeable battery that can give autonomy to the system of 1 hour of use.

Data is transmitted to the computer via a radio receiver on the 2.4 GHz frequency. The outdoor range of the transmission is 500m (the indoor transmission has not been measured systematically, but was found quite proficient for at least a good 50 meters). The receiver uses a Cinch connection that feeds into an A/D converter, the Dazzle Video Compressor, to output a USB signal directly pluggable and interpretable by the PC. The receiver needs a power supply of 12V / 250mA.

The total cost of the prototype is 699.- CHF (~435 euros). This does not take into account the manufacturing of the electronic board and the mechanical pieces (belt, battery and camera support), which have been produced internally at EPFL.

The acquisition software is provided with user friendly interface that allows the therapist to easily record and review the WearCam data. It offers the therapist the option to optimize the tracking algorithms, see Section IV for details, by manually selecting a region of the image that contains a particular color (e.g. the skin color). The selected color can then be localized throughout the whole video and might be used to improve the face detection (Fig. 5).

In spring 2006, preliminary tests have first been conducted at EPFL and in a nearby kindergarten with a 24 months child in free-play settings. These early tests were very successful and promising. The child accepted the WearCam without problem



Fig. 3. The WearCam mounted on a pair of elastic bands (top left/ bottom left-right), the radio receiver and the A/D converter (top right)

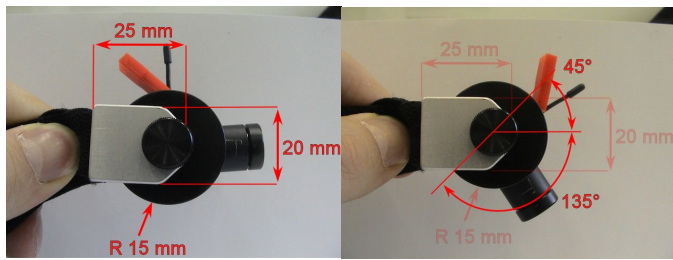


Fig. 4. The support system to incline manually the camera

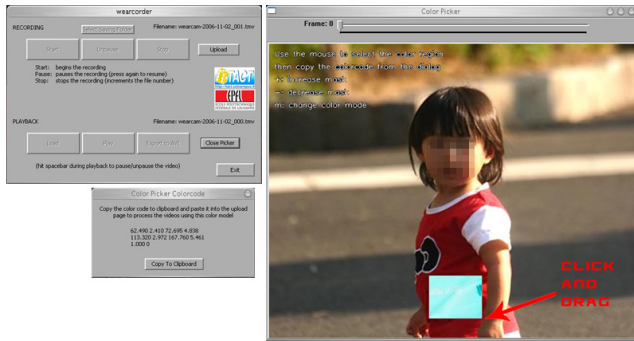


Fig. 5. The data acquisition software

and the weight did not seem to be an issue. We observed none of the blurry effects typically seen with webcams, even when the child moved the head very rapidly while walking around. In early fall 2006, a set of pilot studies with normally developing children, children at risks and children with Autism have been started by three institutions (UU, MEDEA and CBM)³, member of the TACT project. UU tests normally developing children and children with autism, aged 2 to 6 years old. CBM conducts study with normally developing children, age 18 months to 36 months. MEDEA tests 6-month old babies, normally developing and at risk of neurodevelopment disorders. Preliminary feedback from the collaborators informed us that the therapists and researchers using the WearCam were satisfied with the user-friendliness of the software and the quality of the data acquired by the camera

Currently, we are developing a second prototype of the WearCam using higher quality but WIRED cameras. This new WearCam will provide higher quality videos for use in more constrained experiments, as the child will be forced to sit on a chair and will no longer be able to roam about the room. This is adequate to monitor one-to-one interaction during typical diagnosis tests. Having a non-wireless camera will also reduce importantly the weight of the system, which in the current prototype is mostly due to the battery. This will allow us to consider adding a second camera to widen the total field of view of the system.

A major assumption in the design of the prototype was that

³Dr. Claes von Hofsten Research Team at University of Uppsala (UU), Flavio Keller, Domenico Campolo and Giuseppina Schiavone at Università Campus Bio-Medico (CBM), Sara Forti and Maria Nobile at MEDEA Associazione "La Nostra Famiglia" (3 different hospital are involved, Monza, Lecco and Como).



Fig. 6. Experimental setup with Wearcam and other TACT devices (CBM)



Fig. 7. 24 months child in free-play settings wearing the WearCam in a kindergarten

the child's gaze direction would be strongly correlated with the head direction, and thus, placing the camera on the child's forehead would be sufficient to extract that information. Early tests conducted at EPFL with a 2-year old normally developing child showed that this was hardly the case when the child was interacting close-up with either objects or people. Thus, unless the experimental set-up was configured in such a way that all objects and people were at a distance of 1 meter minimum, the above hypothesis would not hold and the information returned by the camera could not be trusted.

Not only the camera does not follow the child's gaze, but, sometimes it may even return an image that does not correspond to what the child is looking at because of its small field of view. For instance, if the child looks downwards and the camera is placed horizontally on the child's forehead, it will miss at least half of the child's field of view. As a result, further development had been made to provide a better localization of the gaze. This was possible by using a mirror pointing directly at the eyes. The mirror reduces slightly the field of view of the camera since the lower part of the image

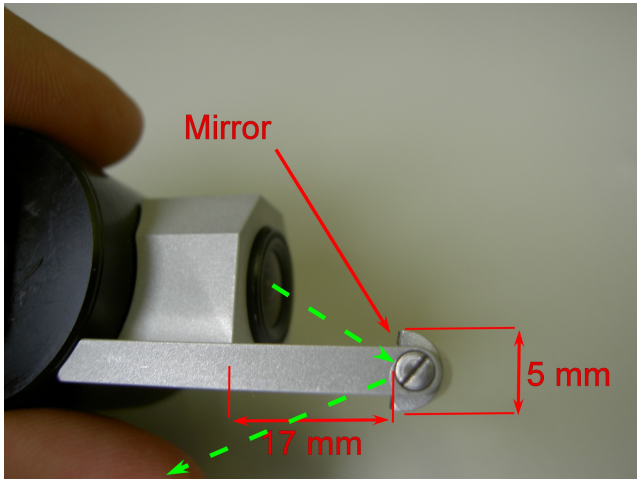


Fig. 8. The mirror for the eye gaze detection

is occupied with the reflected image of the eyes. The efficacy of the mirror has not yet been tested with children. However, we report here on software development for better detecting the gaze direction.

IV. DATA ANALYSIS

An important part of the WearCam project consists in the analysis of the data collected from the WearCam. Data analysis should be automatic so as to reduce significantly the workload on the therapists. Nowadays, the analysis of videotapes of therapeutic sessions is usually done manually, by having independent raters watching the videos frame by frame (e.g. T.I.M.E. scale [8]). Analysis of the WearCam data aims at determining the focus of attention of the child and to quantify the span of attention given to looking at faces. This is achieved by using consecutively software for:

- Face Detection
- Gaze Direction Detection

Face detection is performed using Cascade of weak Haar-like Classifiers with AdaBoost ([9]). The classifier is trained on a database of faces to learn several features typical of faces (e.g. a long vertical or horizontal stripe corresponding roughly to the eyes or nose region). If a region of the image contains the necessary features it is considered as a face. Depending on where the faces detected are located and for how long they are detected, we can make some assumptions as to whether the child's attention is focused on a person or not. A face detected in the center of the image for a long moment will mean that the child is interacting with the person whereas if the face stays at the edges of the field of view, the child's attention is probably focused on something else. The main limitation of the method currently used is the sensibility to the rotation of the face. As the method uses the relationship between horizontal, vertical and diagonal lines composing a face, when a face is tilted towards one side the detection fails (Fig. 9). While this is usually of little concern in usual applications of face recognition, in-plane rotation occurs very often with the WearCam, as the child adapts its gaze direction

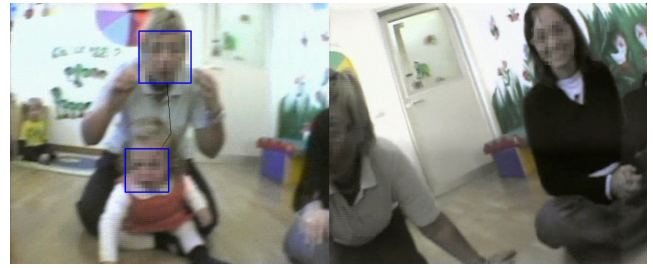


Fig. 9. Multiple face detections (left). Face detection failure due to head tilting (right).

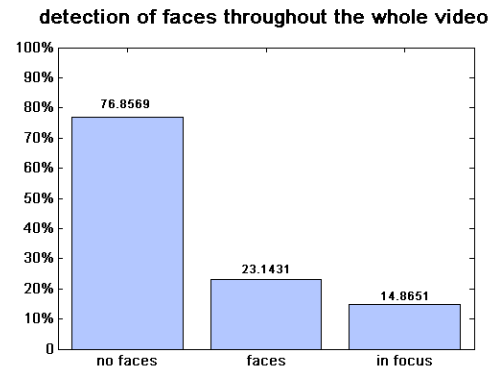


Fig. 10. Quantification of the time (percentage) in a ~7 min long video during which no faces were detected (left), at least one face was detected (center) and at least one face was located in the focus of attention, i.e. was detected within a 230pixels radius from the centre of the image (right).

to the height of the adults it is looking at. We are currently expanding the Haar-Cascade to become Rotation-invariant, so as to overcome this problem.

Gaze direction detection is performed on the part of the image returned by the mirror that looks directly into the eyes of the person wearing the WearCam⁴ (Fig. 11). As of now, conventional computer vision methods such as correlation and pattern matching are used to detect the position of the pupils. The direction of the gaze is, then, extracted by combining the information of the positions of the pupils in both eyes. The region of the eyes returned by the WearCam' image is too small to allow detection of the vertical movement of the eyes; thus, sole movement along the horizontal plane with respect to the child's head orientation can be detected. As it stands, the method needs a fair amount of calibration, the eyes position/size must be configured manually and the position of the eyes at rest (the center of the gaze) must be gauged by the user. To circumvent these problems, automatic methods to detect the eye position and direction are being investigated, using Back Propagation Multi-Layer Perceptron, and Support Vector Machines [10] trained on a set of videos captured with the WearCam and manually labeled.

The data collected through the methods described above can be used in several ways. For instance, in a preset experiment

⁴Note that the prototypes of WearCam used by the therapists and psychologists do not yet have the mirror, but they will be provided with one once the first sets of tests will have been completed.



Fig. 11. Detection of the gaze direction. (top) The eyes of the wearer are reflected in the bottom portion of the WearCam image, the black line shows the estimated direction of the gaze. (middle) Result of the illumination normalization. (bottom) Results of the eye template matching with best candidates. The white and red circles superimposed to the top and middle images show respectively the center of the eyes and the detected position of the pupil in the eye.

involving social interaction and object manipulation (as per the protocol of standard attention-directed experiments [11]), it will be possible to gather statistics about the time the child is looking at the persons around her and the time the child is looking at the objects involved in the experiment (Fig. 11). Additionally it will be possible to analyze the correlation between eyes and head movement (this is a research direction that has not yet been explored and no assumptions about the results can be made at this stage).

V. DISCUSSION AND CONCLUSION

This paper presented a novel tool, the WearCam a wearable wireless non obtrusive CCD camera, for monitoring directed-attention of 6 months to 2 years old children in free play settings. We reported on the technical specifications set for the design of the WearCam. Particular attention was given to ensure that the weight and size of the camera would be sufficiently light and small to be wearable by very young infants and that the device would be non-obtrusive.

The WearCam is provided with user-friendly data acquisition and analysis software for use by the therapists and psychologists conducting the tests. Different types of data analysis can be performed, such as face detection, object/color recognition and gaze detection. Such data analysis may reveal attentional disorders, indicating deficits in social interaction such as those known in Autism [7].

Validation tests of the current prototypes have started at

three institutions of TACT project partners. The data collected at these places will set the ground to develop standards for comparison between children with neurodevelopmental disorders and normally developing ones.

ACKNOWLEDGEMENT

This work is funded by the European Union's NEST Adventure Program as part of the Thought in Action (TACT) research project (<http://tact.unicampus.it>).

We would like to thank Frederic Magnard, MSc Student at the LASA laboratory during the summer 2006, for his contribution on the development of the data analysis software and also the electronic and mechanical Workshops at EPFL where the different WearCam components have been manufactured.

REFERENCES

- [1] D. Winfield, Dongheng Li, J. Babcock, D.J. Parkhurst, "Towards an open-hardware open-software toolkit for robust low-cost eye tracking in HCI applications", Iowa State University Human Computer Interaction Technical Report ISU-HCI, April 2005.
- [2] Dongheng Li, J. Babcock, D.J. Parkhurst, "openEyes: a low-cost head-mounted eye-tracking solution", *Eye Tracking Research & Application, Proceedings of the 2006 symposium on Eye tracking research & applications*, pp. 95-100, San Diego, California, 2006.
- [3] Dongheng Li, D. Winfield, D.J. Parkhurst, "Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches", *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 3, pp. 79, June 2005.
- [4] A. Iijima, M. Haida, N. Ishikawa, H. Minamitani, Z. Shinohara, "Head Mounted Goggle System with Liquid Crystal Display for Evaluation of Eye Tracking Functions on Neurological Disease Patients", *Engineering in Medicine and Biology Society, 2003, Proceedings of the 25th Annual International Conference of the IEEE*, Volume4, 17-21 Sept. 2003, pp 3225-3228.
- [5] R.S. Allison, M. Eiyenman, B.S.K. Cheung, "Combined Head and Eye Tracking System for Dynamic Testing of the Vestibular System", *IEEE Transaction on Biomedical Engineering*, Volume 43, Issue 11, Nov. 1996, pp 1073-1082.
- [6] R. el Kaliouby, A. Teeters, R.W. Picard, "An Exploratory Social-Emotional Prosthetic for Autism Spectrum Disorders", in *Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks*, BSN 2006.
- [7] K. Pierce and E. Courchesne, "Evidence for a Cerebellar Role in Reduced Exploration and Stereotyped Behavior in Autism", *Biological Psychiatry*, Vol. 49, Issue 8, pp. 655-664, 2001.
- [8] L.J. Miller, G.H. Roid, "Sequence comparison methodology for the analysis of movement patterns in infants and toddlers with and without motor delays." *Am J Occup Ther.* Apr;47(4):339-47, 1993.
- [9] R.E. Schapire and Y. Singer, "Improved Boosting Algorithms Using Confidence-rated Predictions", *Machine Learning*, Vol. 37, pp. 297-336, 1999.
- [10] X. Zhang, H. Zhan, "An Illumination Independent Eye Detection Algorithm", *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, Vol. 1, pp.392-395, 2006.
- [11] P.A. Filipek, P.J. Accardo, G.T. Baranek, E.H.Jr. Cook, G. Dawson, B. Gordon, J.S. Gravel, C.P. Johnson, R.J. Kallen, S.E. Levy, N.J. Minshew, B.M. Prizant, I. Rapin, S.J. Rogers, W.L. Stone, S. Teplin, R.F. Tuchman, F.R. Volkmar, "The Screening and Diagnosis of Autistic Spectrum Disorders", *Journal of Autism and Developmental Disorders*, Vol. 29, Issue 6, pp.439-484, 1999.