

***i*KnowU – Exploring the Potential of Multimodal AR Smart Glasses for the Decoding and Rehabilitation of Face Processing in Clinical Populations**

Simon Ruffieux^{1(✉)}, Nicolas Ruffieux², Roberto Caldara²,
and Denis Lalanne³

¹ HumanTech, University of Applied Sciences and Arts, Fribourg, Switzerland
simon.ruffieux@hes-so.ch

² iBMLab, University of Fribourg, Fribourg, Switzerland
{nicolas.ruffieux, roberto.caldara}@unifr.ch

³ Human-IST, University of Fribourg, Fribourg, Switzerland
denis.lalanne@unifr.ch

Abstract. This article presents an explorative study with a smart glasses application developed to help visually impaired individuals to identify faces and facial expressions of emotion. The paper discusses three experiments in which different patients, suffering from distinct pathologies impairing vision, tested our application. These preliminary studies demonstrate the feasibility and usefulness of visual prostheses for face and emotion identification, and offer novel and interesting directions for future wearable see-through devices.

Keywords: Wearable · Smart glasses · Face identification · Emotion recognition · Prosthesis · Rehabilitation · User-study

1 Introduction

Our ability to effectively perceive and identify faces, to decode their emotions and social signals is crucial for everyday life social interactions. This specific aptitude can be impaired in a wide range of clinical conditions, such as low vision, dementia or autism, only to name a few. Such an impairment leads to suboptimal social interactions and can lead to potentially dramatic consequences on psychological well-being of the visually impaired. Worldwide, approximately 32 million people are blind and 191 million live with moderate to severe visual impairment [1]. In Switzerland alone, about 10'000 individuals are blind and over 300'000 have a severe visual handicap [2]. In this population, one of the most frequent complaint relates to the difficulty to identify faces and decode facial expressions of emotion [3, 4].

While automatic recognition of faces and facial expressions of emotion is already well developed and implemented on different static computing platforms (e.g. websites, cameras, desktop computers, etc.) and through online cloud-based services (Amazon Rekognition, Google Cloud Vision API, Microsoft Face API), its application on wearable devices in the wild is more complex (multiple faces to process simultaneously,

targets in motion, changes in luminosity, limited connectivity, etc.) and requires additional developments. The very recent advances in multimodal processing and wearable technologies provide a novel and unique opportunity to create perceptual and cognitive wearable prosthesis. The aim of our project is to develop always available augmented-reality smart glasses that support patients with face and emotion recognition aids. This can be achieved via multimodal information (e.g. sounds, images, texts, vibrations) provided by the glasses to the person wearing them.

The technology of smart glasses has been evolving rapidly these last years and several well-known companies are now working on eyewear computing devices for augmented reality (Microsoft HoloLens, Epson Moverio, Vuzix, Google Magic Leap). Naturally, the emergence of such devices has aroused much hope for people with visual and/or cognitive impairment. Hence, the need to fill the enormous gap between this already partially available technology and the potential benefit for individuals who present with visual and/or cognitive impairment. The potential impact of this project for the population is therefore considerable, knowing the lack of wearable vision-based technologies designed to compensate for such impairments. More specifically, there is a huge potential benefit in the elderly population, as they represent a major proportion of the visually impaired.

We thus aim to use smart glasses to support the decoding and rehabilitation of face identification and facial expressions of emotion, by using multimodal solutions in order to provide tailored feedback depending on the specific disabilities of the user. In this study, we describe the initial explorations with a prototype application implemented on the Epson Moverio BT-200 glasses. This prototype has been tested with three participants suffering from various visual pathologies. These experiments performed in indoor conditions allowed us to gather the impressions from participants and identify their specific needs in terms of tailored multimodal feedback provided by the device according to their respective pathology.

2 State of the Art

In recent years, eyewear computing devices have been catching more attention from research, thanks to the affordable and increased hardware availability. Researchers have outlined the potential of these devices in terms of Human-Computer Interaction, notably in terms of multimodal inputs and outputs. The strong inherent advantage of such device being privacy, usability and hand-free access to the digital information thanks to their location on the head of the subject which are very important features for the target population of this study [5, 6].

The current research on eyewear devices to help visually impaired is predominantly an extension of the existing research based on mobile devices such as smartphones or small embedded systems. In their survey, Terven et al. reviewed eight different portable computer-vision-based travel aids assistive technology systems that have been developed this last decade for mobile devices, highlighting three different types of feedbacks: acoustic, electric stimulation or tactile [7]. The authors notably point out the new opportunities for those types of systems with the increased computing power of mobile devices and wide smartphone availability. They conclude by highlighting the need to

gather better knowledge of visually impaired user's needs in order to guide the development from prototypes to products. They mention particularly the problem of the acoustic feedback acceptance, which interferes with the remaining senses of the visually impaired individuals. In their survey, Sujith and Safeeda instead provided an overview of computer-vision based projects for visually impaired persons [8]. They compared the various functionalities of the surveyed projects: door detection, obstacle detection & identification, finding specific objects and path finding. They conclude that most of the studied projects focused on a single task, although visually impaired persons would need a single system encompassing all the different mentioned functionalities. In their studies, Jafri et al. focused mostly on object recognition to help visually impaired individuals [9] and draw attention of the research community on the potential of eyewear devices to help visually impaired [6]. In a recent study, Jafri proposed the use of CUDA-based GPU computing in order to perform real-time people and objects detection and recognition on the "Tango Tablet" device [10].

Recently, several studies investigated face recognition with eyewear and wearable devices to improve social interactions. Mandal et al. proposed a system to assist people during social interactions using a face detection system implemented on Google Glass paired with a smartphone. The algorithmic computations are performed on the smartphone, while the feedback is provided through the transparent display of the Google Glasses [11]. Li et al. extended this project by developing a multi-threaded architecture on the smartphone to better handle the constant stream of images and reduce the time required by the system to recognize faces. They notably established that an acceptable delay for face recognition should be below 2.3 s in order to be accepted by users [12]. Finally, Xu et al. proposed the "SocioGlass" project, which combines the two precedent projects and provides dynamic face recognition and adaptive retrieval of personal information in an Android application [13]. Importantly, they proposed to adapt the information provided to the user depending on work or personal context. The use of augmented-reality smart glasses to directly enhance vision of visually impaired has also been investigated in several studies. The idea is to provide a simplified or modified vision of the real world using the transparent display of the glasses in order to compensate for the visual impairment. Hu et al. proposed to enhance night vision for individuals suffering nyctalopia [14]. Hwang and Peli investigated the possibility to use edge enhancement techniques on Google Glasses to enhance the vision of visually impaired [15].

On the commercial side, the company ORCAM developed a pair of smart glasses with a mounted-camera in order to acoustically translate text in real-time for the visually impaired. They use bone conduction to transmit the information to the users, thus reducing the interferences with the auditory sense of the users. The new version of their product also provides face recognition and identification of a wide range of consumer products. Hicks et al. have developed a hardware AR see-through device specifically designed for visually impaired individuals. This device is based on a Moverio BT-200 modified with a depth-based camera [16]. They recently created the OxSight company to promote and develop new software for their device [17].

Importantly, only limited literature exists on research conducted on the usability, acceptance and real needs of clinical populations for those types of systems, notably with see-through smart glasses. Based on a study on a cohort of 364 individuals

suffering from age-related macular degeneration, Cimarolli et al. report that more than half of the patients mentioned the inability to recognize people as the most disabling difficulty for social interactions [4]. Krishna et al. investigated the needs of blind and visually impaired in the context of social interactions [18]. Their work highlighted specific needs such as knowing the facial expression of their interlocutor, identifying the names of the surrounding persons and the direction of their attention, their appearance and if their look had changed since the last meeting. More recently, Sandnes et al. conducted open interviews with visually impaired individuals to better identify their needs, specifically focusing on eyewear computing devices [3]. The authors highlighted the reluctance of visually impaired individuals for highly recognizable and stigmatizing devices visible to all. The interviewed individuals clearly expressed that face recognition and text reading would be the most important feature for a system based on smart glasses. Importantly, to the best of our knowledge, while many devices have been developing in recent years, none has been validated among clinical populations yet.

3 *iKnowU*

The current study takes place in the context of the multi-disciplinary “*iKnowU*” project that originates from a collaboration between the Human-IST institute and the *iBMLab* (Eye and Brain Mapping Laboratory), both from the University of Fribourg. The intended goal is to create a solution on see-through smart glasses to aid visually impaired individuals to compensate for their difficulties during social interactions. Using collected information about the user’s relatives, friends and colleagues, the system should be able to automatically recognize the presence of these persons in the visual field of the camera, their identity and emotion, and report this information in the most adapted way to the user.

This article describes the first explorative step of the project, in which we developed a prototype application and evaluated it with several patients in order to better understand and differentiate their needs and interests depending on their specific disabilities. As illustrated on Fig. 1, difficulties in face and/or emotion processing can result from damages to peripheral visual organs (e.g. age-related macular degeneration), from brain lesions (e.g., prosopagnosia, visual agnosia, dementia) or from psychopathological conditions (e.g. autism, schizophrenia). Whereas most research in this field is performed with specific tools for each population and each deficit, our project intends to use of a unique solution with multimodal augmented reality feedback (visual, auditory, tactile) tailored to maximize the user’s residual visual and cognitive abilities for face and emotion processing. Several studies have shown limited adoption of those types of devices, mainly due to false positive/negative and improper handling of the feedback [3]. Therefore understanding the need of each patient and tailoring the feedback is crucial towards a better adoption. Similarly, better ‘error handling’ strategies can be used to avoid providing unrequested information to the user or to convey algorithms’ confidence information directly to the user through the feedback mechanisms.

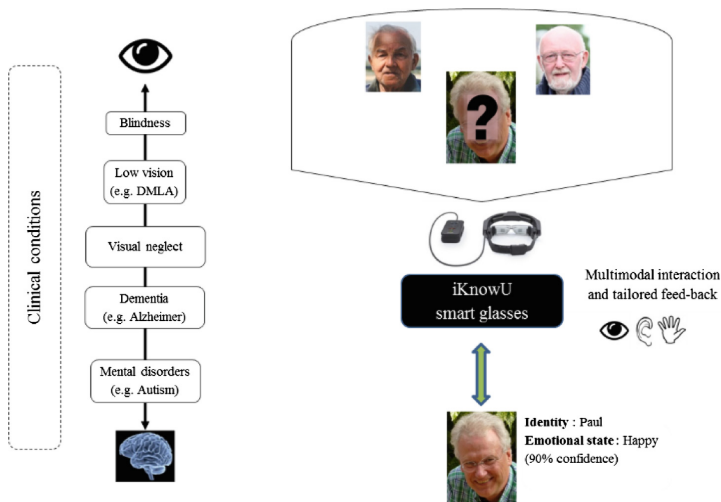


Fig. 1. Left: Clinical conditions with potential difficulties in face and emotion processing. Right: Smart glasses for multimodal interaction and tailored feedback.

As an explorative step and proof of concept for the project, we implemented a prototype of face and emotion recognition application on the BT-200 Epson Moverio smart glasses. Face detection and recognition are based on the standard FaceRecognizer algorithm provided by the OpenCV library [19]. With these algorithms, the system was able to distinguish quite accurately up to five persons in laboratory conditions. The target persons' faces were inserted in the database by using the interface of the developed application. Before each experiment, twenty pictures of each target persons' face were taken from different angles by a member of the team. Emotion recognition has been implemented using the AFFDEX SDK developed by Affectiva [20]. The implemented system is able to recognize five different facial expressions (anger, joy, fear, sadness and surprise) and joint attention. Note that exact accuracies of the implemented algorithms have not been quantified since the goal of the study was to gather first impressions and feedbacks from a clinical population.

The system offers various options to feed information back to the user; in this experiment we used different text and vocal feedbacks. These options could be enabled and disabled according to the capacities and desires of each patient. For face recognition, the user interface highlights detected faces with a green rectangle and recognized persons with a red rectangle around their eyes (potentially flashing to attract user attention) and their name textually presented above their head as shown in Fig. 3. When a new person is recognized, its name is also uttered vocally by the system. For emotions, the interface displays the recognized expressions and emotions with their confidence index textually on the left side of the screen, as shown in Fig. 2. When a new emotion is detected, the system vocally informs the user with a volume that varies according to the confidence index of the algorithm.

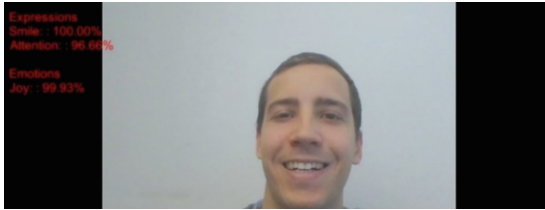


Fig. 2. The user interface with expressions and emotions confidence indexes displayed textually to the user on the upper-left of the display.

4 Experiments

In order to better understand and differentiate the needs of the users, we performed three different experiments in indoor conditions using the developed prototype system with patients suffering from distinct pathologies affecting visual processing. We chose those different pathologies to span over the range of clinical conditions shown in Fig. 1. The three patients described below all have in common severe difficulties to identify faces and facial expressions of emotion, which they report as a major handicap for everyday life social interactions. In those experiments, the patients were asked to identify “actors” or relatives (i.e. up to five persons) while wearing the smart glasses in a room with a good lighting level. In addition, the patients were asked to periodically identify the facial expressions of emotion acted by the person facing them. The patient was asked to look at each of the “actors” in turn and utter what he was seeing/hearing through the device. The “actors” were aligned at about 2 m in front of the patient, who was free to move if necessary. At the end of each experiment, we conducted an informal interview with the participant to collect its impressions and feedbacks about the proposed system and the device.

The first experiment was performed with a patient suffering from prosopagnosia (i.e., the inability to recognize faces due to brain damage) consecutive to a severe head injury in 1992. While her general intellectual abilities remained intact, since the accident the patient suffers from major difficulties to visually identify familiar faces, such as her husband or her children [21]. Using our prototype, we observed that this prosthesis improved her ability to identify faces.

In her case, the visual feedbacks (i.e. name of the person shown on the screen) were the most effective, while she found the audio feedbacks to be too intrusive. Furthermore, it has previously been reported that this patient uses a suboptimal strategy to process face identity and expressions, focusing her attention on the mouth instead of the eyes [22, 23]. Whereas automatic visual attention processing is extremely hard to modify through classical rehabilitation techniques, a visual feedback allowed to highlight the eyes (e.g., a flashing red square around the eyes, as shown on Fig. 3) and to train the patient to better focus on this informative part of the face. It can be hypothesized that if the patient is able to integrate such an efficient compensatory technique, she would ultimately be able to better cope without the smart glasses in everyday life situations.

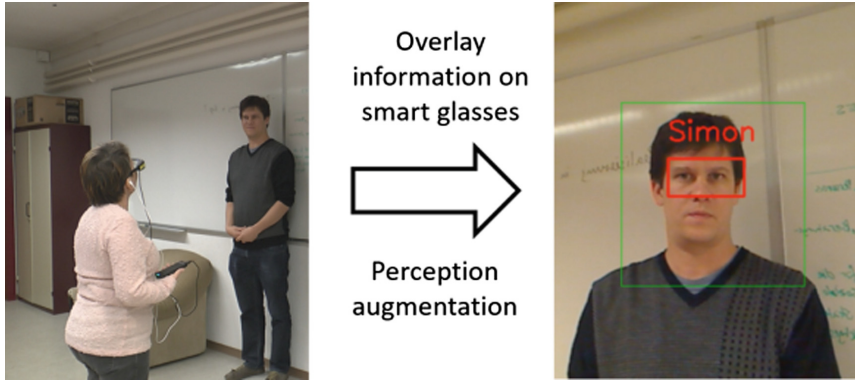


Fig. 3. Overlaying information on the smart glasses and eye focus rehabilitation concept (flashing red square around the eyes). (Color figure online)

The second experiment was conducted with a patient suffering from cortical blindness (i.e. severe loss of vision-despite normal oculomotor behavior- caused by damage to the brain's occipital cortex). This patient is unable to visually identify objects, colors, letters or shapes and is therefore severely incapacitated for everyday life activities [24]. His major complaint concerns his inability to visually identify his relatives.

We performed the experiment at the patient's home, due to the difficulty for him to come to our laboratory. The subject was asked to recognize his wife amongst three people that were sitting on a couch, as the patient had difficulties standing up. Using the smart glasses, he was able to rapidly identify his wife, although the system was not as reliable as usual due to low luminosity conditions. Then we tested the emotion recognition system; his wife sat in front of him and acted different emotions. He successfully detected most emotions relying on the auditory feedback provided by the system. He notably appreciated the fact that the volume of the auditory feedback was modulated according to the confidence indexes inferred by the algorithm; such indexes are visible on the textual feedback shown in Fig. 2.

The patient clearly outlined that, in his case, visual feedbacks were not as useful as the auditory feedback as he was not able to easily process the information shown on the smart glasses display, although he mentioned that some flashing elements on the screen attracted his attention during the experiments (i.e. the red square when a person was recognized). He also mentioned the glasses were quite uncomfortable and that he would probably not wear them during a whole day.

The third experiment was conducted with a 48 years old patient presenting with quasi-blindness due to retinitis pigmentosa. This inherited eye disease causes a progressive degeneration of the rod cells in the retina. Consequently, this patient presents with tubular vision (visual field: 15°), low visual acuity (0.2), defective light-dark adaptation and photophobia. The patient was extremely enthusiastic when he tried the prototype. As the two previous patients, he was able to rapidly identify faces and emotions using the smart glasses. An interesting discovery was made during the

experiment because of a trial with a dark filter overlaid on the glasses. This filter obfuscated the view of the real world and left only the view from the glasses' camera visible to the subject. This setup greatly helped him identify people and objects in the room due to the high contrast and luminosity of the live video shown on the display. Globally, the user's feedback was very positive and he reported that he would enjoy trying the prototype in everyday life conditions. To compensate for his severe visual impairment, this patient already uses various technologies (e.g. white cane, guide dog, closed-circuit TV, screen reader, handheld magnifier, etc.). He notably mentioned his major interest in having a single device solving most of his difficulties instead of carrying multiple devices with him, each solving a particular problem.

5 Conclusion

We developed a novel smart glasses application providing assistance for face and facial expressions of emotion and performed an evaluation in indoor conditions with three patients presenting with distinct pathologies impairing their vision.

During the interviews conducted after the trials, the patients' feedback were very positive, as they all reported they would like to use the prototype in everyday life conditions. One patient mentioned that the glasses were a bit uncomfortable and that he would not wear them during a whole day. The experiments and interviews clearly outlined the need to be able to tailor the feedbacks provided by the device to each patient's specific needs. The experiments also highlighted the current limitation of the application and the need to use more advanced algorithms and image processing heuristics to improve the robustness in the wild. In this study, we did not consider online cloud-based detection and recognition solutions, as we aim at providing an independent device that is always functional; indeed, network connectivity is often limited in numerous locations (underground locations, remote areas, etc.). A hybrid approach could however be considered in a future step as cloud-based solutions provide access to powerful algorithms by externalizing computations from the device.

In the future, we plan to further improve the system, notably by developing algorithms that are more robust to light conditions. We also plan on extending output and input feedback mechanisms by adding additional output modalities and by letting users inform the system of misdetections. False positives/negatives handling is indeed crucial for user acceptance of such devices. Such system could also be used for rehabilitation in some cases, as highlighted by the approach used in the first experiment. We will further investigate this axis in the next steps of the project by working in close collaborations with neurologists and clinical populations, assessing the effectiveness of the smart glasses through controlled experiments. In parallel, we are currently conducting a questionnaire survey with clinical populations to further investigate the needs according to the range of patients' disabilities. Our ultimate aim is to develop an effective visual prosthesis for the recognition of face identity and expressions of emotion. While the technology is still young and developing, we believe that in the long term smart glasses could have a similar game-changing effect on the visually impaired population, as the auditory prosthesis did since the 1970's.

References

1. Stevens, G.A., White, R.A., Flaxman, S.R., Price, H., Jonas, J.B., Keeffe, J., Leasher, J., Naidoo, K., Pesudovs, K., Resnikoff, S., Taylor, H., Bourne, R.R.A.: Global prevalence of vision impairment and blindness: magnitude and temporal trends, 1990-2010. *Ophthalmology* **120**, 2377–2384 (2013)
2. Spring, S.: Handicap visuel et cécité : évolution en suisse, pp. 5–22 (2012)
3. Sandnes, F.E.: What do low-vision users really want from smart glasses? Faces, text and perhaps no glasses at all. In: Miesenberger, K., Bühler, C., Penaz, P. (eds.) ICCHP 2016. LNCS, vol. 9758, pp. 187–194. Springer, Cham (2016). doi:[10.1007/978-3-319-41264-1_25](https://doi.org/10.1007/978-3-319-41264-1_25)
4. Cimarolli, V.R., Boerner, K., Brennan-Ing, M., Reinhardt, J.P., Horowitz, A.: Challenges faced by older adults with vision loss: a qualitative study with implications for rehabilitation. *Clin. Rehabil.* **26**, 748–757 (2012)
5. Bulling, A., Kunze, K.: Eyewear computers for human-computer interaction. *Interactions* **23**, 70–73 (2016)
6. Jafri, R., Ali, S.A.: Exploring the potential of eyewear-based wearable display devices for use by the visually impaired. In: Proceedings - 2014 3rd International Conference User Science Engineering Experience Engineer Engage, i-USEr 2014, pp. 119–124 (2015)
7. Terven, J.R., Salas, J., Raducanu, B.: New opportunities for computer vision-based assistive technology systems for the visually impaired. *Comput. (Long. Beach. Calif)* **47**, 52–58 (2014)
8. Sujith, B., Safeeda, V.: Computer vision-based aid for the visually impaired persons - a survey and proposing. *Int. J. Innov. Res. Comput. Commun. Eng.* 365–370 (2014)
9. Jafri, R., Ali, S.A., Arabnia, H.R., Fatima, S.: Computer vision-based object recognition for the visually impaired in an indoors environment: a survey. *Vis. Comput.* **30**, 1197–1222 (2014)
10. Jafri, R.: A GPU-accelerated real-time contextual awareness application for the visually impaired on Google's project Tango device. *J. Supercomput.* **73**, 887–899 (2017)
11. Mandal, B., Chia, S.-C., Li, L., Chandrasekhar, V., Tan, C., Lim, J.-H.: A wearable face recognition system on Google glass for assisting social interactions. In: Jawahar, C.V., Shan, S. (eds.) ACCV 2014. LNCS, vol. 9010, pp. 419–433. Springer, Cham (2015). doi:[10.1007/978-3-319-16634-6_31](https://doi.org/10.1007/978-3-319-16634-6_31)
12. Chia, S., Mandal, B., Xu, Q., Li, L., Lim, J.: Enhancing social interaction with seamless face recognition on Google glass. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI 2015, pp. 750–757. ACM Press, New York (2015)
13. Xu, Q., Chia, S.C., Mandal, B., Li, L., Lim, J.-H., Mukawa, M.A., Tan, C.: SocioGlass: social interaction assistance with face recognition on Google glass. *Sci. Phone Apps Mob. Devices* **2**, 7 (2016)
14. Hu, C., Zhai, G., Li, D.: An augmented-reality night vision enhancement application for see-through glasses. In: 2015 IEEE International Conference Multimedia Expo Work, ICMEW 2015 (2015)
15. Hwang, A.D., Peli, E.: An augmented-reality edge enhancement application for Google glass. *Optom. Vis. Sci.* **91**, 1021–1030 (2014)
16. Hicks, S.L., Wilson, I., Muhammed, L., Worsfold, J., Downes, S.M., Kennard, C.: A depth-based head-mounted visual display to aid navigation in partially sighted individuals. *PLoS One* **8**, e67695 (2013)
17. Oxsight: Oxsight company. <http://smartspeccs.co/>

18. Krishna, S., Colbry, D., Black, J., Balasubramanian, V., Panchanathan, S.: A systematic requirements analysis and development of an assistive device to enhance the social interaction of people who are blind or visually impaired. In: *Work Computer Vision Applications for Visually Impaired* (2008)
19. Wagner, P.: Face recognition with opencv. OpenCV 2.4. 9.0 Doc (2012)
20. McDuff, D.: AFFDEX SDK : a cross-platform real - time multi-face expression recognition toolkit. In: *CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 3723–3726 (2016)
21. Rossion, B.: A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* **126**, 2381–2395 (2003)
22. Caldara, R., Schyns, P., Mayer, E., Smith, M.L., Gosselin, F., Rossion, B.: Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *J. Cogn. Neurosci.* **17**, 1652–1666 (2005)
23. Fiset, D., Blais, C., Royer, J., Richoz, A.-R., Dugas, G., Caldara, R.: Mapping the impairment in decoding static facial expressions of emotion in prosopagnosia. *Soc. Cogn. Affect. Neurosci.* (2017). (in press)
24. Ruffieux, N., Ramon, M., Lao, J., Colombo, F., Stacchi, L., Borruat, F.-X., Accolla, E., Annoni, J.-M., Caldara, R.: Residual perception of biological motion in cortical blindness. *Neuropsychologia* **93**, 301–311 (2016)