

Computerized-Eyewear Based Face Recognition System for Improving Social Lives of Prosopagnosics

Xi Wang, Xi Zhao*, Varun Prakash⁺, Weidong Shi, Omprakash Gnawali

Department of Computer Science, University of Houston, Houston, TX 77004, U.S.A

Email: {xiwang,larryshi,gnawali}@cs.uh.edu, xzhao21@central.uh.edu*, vsprakash@uh.edu⁺

Abstract—Prosopagnosia (PA), or the inability to recognize faces, is widespread around the world, with no effective cure for the disease. We propose a wearable system to improve prosopagnosic’s daily life. The system utilizes a real-time face recognition application that runs on a smartphone, and a portable eyewear that displays descriptive information when a person appears in range of camera mounted on the eyewear. The system helps these patients identify faces, reminds them to make social reactions and ultimately prevents them from developing social anxiety disorders, characterized by fear and avoidance of social situations that may cause embarrassment. Lastly, we implement the system in a real life demonstration to test its performance.

Index Terms—Prosopagnosia, face recognition, computerized-eyewear

I. INTRODUCTION

Prosopagnosia [1] is characterized by a person’s inability to recognize faces. People with Prosopagnosia, referred to as Prosopagnosics, find it difficult to distinguish facial features and differentiate people in their social lives [2]. Dr. Bate found that indifference is easily generated between patients and their families and friends [3]. They struggle with interpersonal relationships, become depressed and may even end up avoiding social interaction. In recent years, there have been a rapid growth in the number of prosopagnosic cases reported. In August 2006, a German study found that as many as 1 out of 50 people may have had some form of prosopagnosia [4]. A separate web survey of 1,600 people conducted jointly between researchers in Harvard University and University College of London substantiates this finding [5]. Recently, a hereditary sub-type of congenital prosopagnosia with a high prevalence rate of 2.5% has been identified [6]. If these figures are accurate, there may be 6.26 million prosopagnosics in the United State of America alone.

The most prevailing strategy for prosopagnosics is to learn to identify people by their clothing, gait, and other distinctive features [7], [8]. However, memorizing various details puts heavy burden on the patients. And the variability of these details can easily confuse the patients. Since the strategy does not adopt the countenance information, which historically has been the defining factor to distinguish people [9], it is not able to adequately improve a patient’s social life from a real-world perspective. Instead of transient identity details, we propose a novel computerized-eyewear based face recognition system which directly uses faces as the identity feature, as seen in Fig. 1. Since patients’ visual processing and intellectual functioning remain intact [9], the assistance provided by our system to compensate their inability in face recognition can help them in living a normal life. Specifically, the system localizes faces in front of the users in the video captured by the eyewear, matches the face in the captured image to the

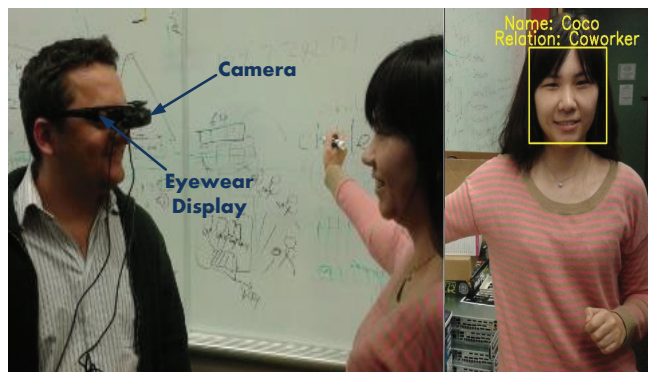


Fig. 1. An typical scenario of the usage of the computerized-eyewear based face-recognition system. The left figure depicts an user wearing the system and talking with his coworker; the right figure simulates the display of the tagged identity information from the system on the eyewear.

most similar face in the database to identify and verify them, and displays the tagged identity information in a proper region in the eyewear. Researchers in Arizona State University have designed a wearable face recognition system which identifies people using the attached camera for the user with visual impairment [10]. The identities are informed to users vocally via headphones. However, when multiple people are present in the scene, the user may encounter the difficulty of associating the vocal identity tag with visually presented people. Different from their work, our system leverages the emerging eyewear devices that are equipped with see-through displays (e.g., Google Glass), and clearly display the identities by overlapping the tags on the people’s faces.

Our system enables prosopagnosic patients to function properly in a social setting by helping them identify the people they come across. Essential features of our solution include: firstly, Our system is wearable making it possible for the patients to wear it during their daily activities. Secondly, smartphones, as the essential computation component in the system, offer a high computational power with relatively low cost.

Our main contributions are: (i) exploration and analysis of the pathology of Prosopagnosia and the needs from the prosopagnosic patients, and proposal for a portable face-recognizer-based solution to improve the prosopagnosics’ life quality; (ii) utilization of advanced computerized-eyewear, which provides functions of real-time video capture and information display; (iii) design of a software application that uses face detection and recognition.

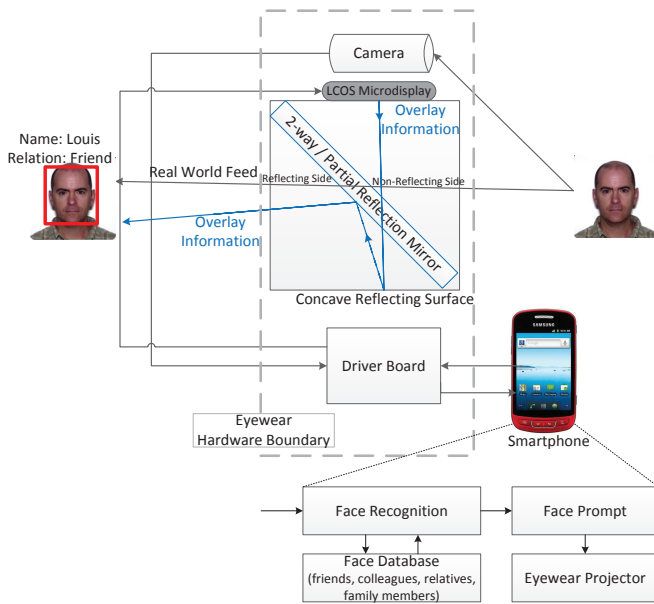


Fig. 2. Block diagram of the proposed system. The computation is carried out by the smartphone which contains the face database and performs the face recognition activity. The boundaries of the eyewear encompass the visual units such as the camera and LCOS micro display which projects the image in the display window.

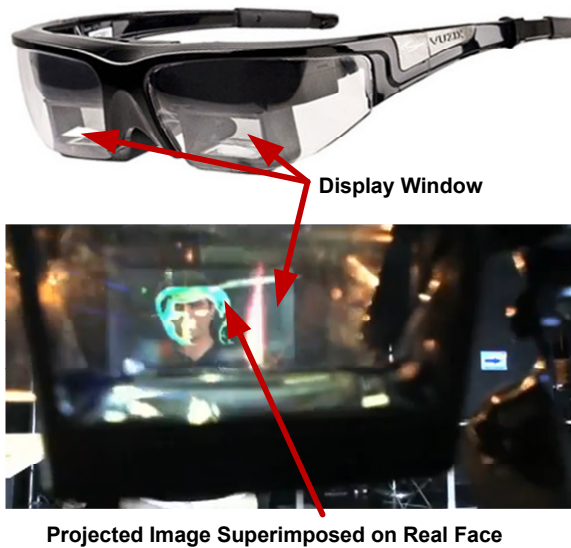


Fig. 3. The figure on top shows the location of the display window in the eyewear. The display window represents a container holding the partial mirror upon which the resulting information is displayed. The bottom figure demonstrates a sample view produced during the working of the device.

II. SYSTEM DESIGN

In this section, we present the hardware and software components of our system.

A. Hardware

The real-time system discussed in this paper requires a hardware setup that is portable and can accompany the patient without causing any discomfort. The wearable hardware used for experiments is similar to the widely known Google Glass in terms of assembly, functionality and operation. It comprises of an eyewear which is worn like spectacles by the patient, a smartphone which runs the face recognition and other

computation modules to support the function of the components mounted on the eyewear. The eyewear in Fig. 3, Vuzix STAR 1200XL third generation augmented reality device [11], comprises the following individual units, a forward pointing camera, an eyepiece including a LCOS (Liquid crystal on silicon) micro display and an interface/driver board that offers a wide range of views, and an eyepiece adjustment mechanism to suit a variety of users.

The eyewear functions as a data collection and information display unit. The arrangement of the eyepiece is reproduced using the widely known HUD display designs. The eyepiece consists of a partial mirror which has one surface coated with reflective material while the other side is a plain glass which allows light to pass through. We can see how the setup works in Fig. 2. The real world images and the raw camera feed pass through the plain surface of the partial mirror in the eyepiece while the information overlay generated by an embedded display is delivered by a process of internal reflection.

Another important part of the eyewear is the interface/driver board that connects the individual units mounted on the eyewear to the I/O module of the application running on the smartphone. The driver board is needed to convert the format of data when it is transferred between the eyewear and smartphone, and provide power to drive the camera and the LCOS microdisplay. The mounted camera records the video from the user's viewpoint and sends it to the driver board, which buffers and sends all or a configured number of frames periodically to the smartphone. The smartphone processes the video frames using the face detection and recognition algorithms to detect and identify the faces in each frame. After the identity information has been determined, the smartphone passes this information to the driver board which in turn converts it into a format which can be displayed by the LCOS display. This information is reflected off the partial mirror and delivered to the user along with the stream of real world images.

B. Software

As seen in Fig. 4, the software application consists of two components: (1) an offline face recognizer training algorithm on the smartphone or a computer, and (2) an online face recognition algorithm running on the smartphone. The two parts are explained as follows:

Offline: In order to train the face recognizer for a specific patient, training data from his family members and friends are needed. They can be captured by the camera in the eyewear as video data during the patient's daily life when the patient sees a frontal face of his/her family members or friends presented with a reasonable scale. The video frames are first preprocessed by cutting facial parts out of the entire picture. These facial parts are detected by Viola-Jones method [12] for each frame. Then the user needs to label the detected faces using the application's user interface on the smartphone or the computer, i.e. which person these faces belong to. We further align all the labeled facial patches by normalizing their sizes into 100×100 pixels and converting color facial patches to greyscale for simplicity and computation efficiency. The recognizer based on Local Binary Patterns [13] (LBP) is generated as an xml file. The patient can press "train recognizer" button to train a new recognizer with new data and associated labels, and replace the old ones if they exist. Furthermore, the patient can press an "update recognizer" button to update the existing face recognizer by registering the

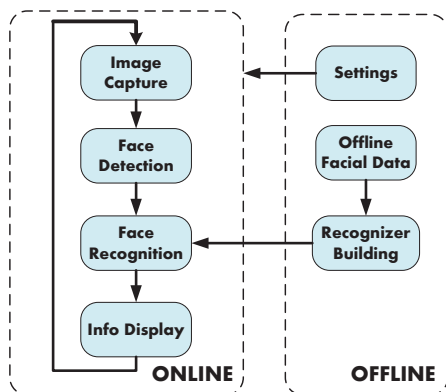


Fig. 4. The flow chart of the face recognition algorithm.

new subject data into the existing recognizer. After training, facial features are extracted and stored in the xml file and thus the facial patches used can be erased from the memory of the computer or the smartphone for efficiency. The trained xml file is cross-platform with support for both computer and smartphone.

The patient can add or edit information about a person, e.g. name, email, relation to the patient, so the corresponding information about a person can be displayed on eyewear when the patient meets the person. A patient can also configure other settings: (1) a verification threshold to set how easily an unknown face can be verified as a person who appeared in the training set; and (2) what information should be displayed on the eyewear in the online phase when a particular person’s face is detected and recognized, e.g. name, relation, birth date or matching distance. The patient can choose the default settings: (1) the threshold is 2985, a value that we obtained according to the empirical results presented in the next section, (2) name and relations are shown on the eyewear.

Online: In the online phase, the application detects faces on the video sent from the camera. If no face is detected in the current frame, the application will continue its detection for the new incoming frames. If faces are detected and positions of them are located in a frame, the system crops the detected faces. Then the application applies the same image pre-processing techniques as in the offline phase in order to increase the recognition performance. The application extracts facial features from the input images, feeds these features into the face recognizer, and obtains a matching distance. We set a threshold for the matching distance. If the distance is larger than or equal to the threshold, the person is considered as a “stranger”; if not, the person may be someone who the patient is familiar with. The information related to the recognized person, whose distance is the minimal among all the registered people, is displayed on the eyewear. The patient then may use the displayed information to decide how to respond to the presence of the person, e.g., displaying a proper facial expression, making friendly gestures, shaking hands, or hugging.

III. SYSTEM EVALUATION

In this section, we report and analyze experimental results of our prototyping system based on the approach detailed in the last section.

A. Dataset

To test the performance of the proposed approach, we used both Yale dataset [14] and our own dataset named

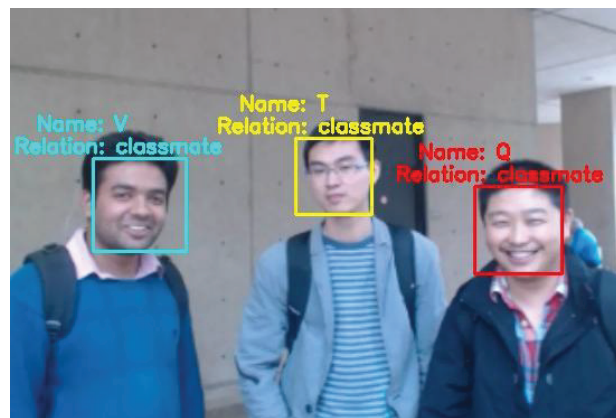


Fig. 5. Demonstration of the system display on the lenses when multiple people appeared in front of the user. The tagged identity information includes three rectangles around the detected faces, subjects’ name (replaced with a letter to keep privacy for participants) and relationship to the user.

Eyewear2013A. To test the application’s face recognition performance under real-world scenarios, we built the video dataset Eyewear2013A using the camera mounted on eyewear at a framerate of 20fps and a resolution of 640×480 . The dataset contains a rich collection of examples with variations in poses, illumination conditions, resolutions, expressions and subject races. In order to mimic daily life, every subject performs a series of actions during capture, such as staying still looking towards the camera, having a conversation, walking towards or passing by the camera etc. Subjects may display various facial expressions including normal, happy, angry and a few exaggerated ones. Eyewear2013A Dataset contains 41 sessions of video data collected from 20 subjects. Every video lasts from 20 to 40 seconds. We recorded 3 video sessions for 10 people at three different scenes respectively: office, aisle and outdoors. We also recorded 1 session for another 10 people respectively at random scenes. Three individuals are present in the 41th video. We assume that in our experiments, the first 10 subjects are family members or friends and the rest 10 subjects are “strangers”. Thus, a part of the data from the first 10 people are used to train a face recognizer, while the rest data from these 10 people and the data from the second 10 people are used to test the face recognizer.

B. Experiments and Results

Figure. 5 illustrates the display on the see-through eyewear glass with the tagged identity information, including three rectangles around the detected faces, subjects’ name (replaced with a letter to keep privacy for participants) and relationship to the patient. It demonstrates that when more than one individual appears in the view field of the camera, each face can be clearly distinguished. Our system is able to successfully perform face detection and recognition on videos that contain multiple faces.

We evaluated the face detection performance by calculating the percentage of frames in which face is successfully detected. Most of the frames are used for evaluation, while a few frames are excluded where subjects did not always show frontal face, or they were too far away until moving towards the camera. This is because realistically no social action is necessary when the people show side faces to the patient or they are too far away from the patient. According to Fig. 7, when a frontal face is in the view field and with a distance less than 10 meters from the camera, it can be detected with averagely

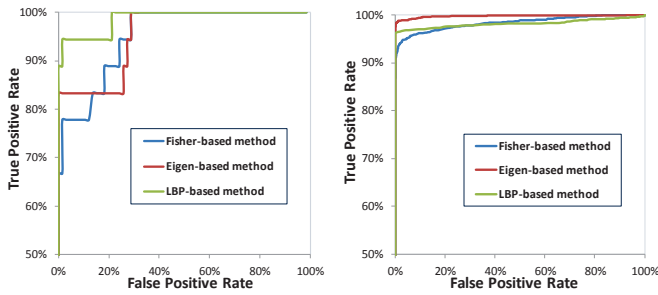


Fig. 6. The ROC Curves depicting the face recognition performance evaluated respectively on the (a) Yale Dataset and (b) Eyewear2013A Dataset.

99% detection rate. We observed that even in some cases the face recognizer failed to recognize the faces in a few frames, it is able to recover from the failure when more input images are processed.

We also investigated the performance of three face-recognition algorithms based on Fisher [15], Eigen [15] and LBP [13] respectively. Figure. 6 shows that LBP based method works best on the Yale Dataset, LBP and Eigen based methods yield better results than Fisher one on Eyewear2013A. We opt to adopt LBP based method in our system. Our reasons are: firstly, unlike Fisher and Eigen based methods, LBP based method supports update of the face recognizer. It requires less running time and memory space compared with the other two methods. Secondly, LBP based method has a better performance on average on both datasets. Its accuracy rates of face verification and recognition are 95% and 100% on the Yale Dataset, 97% and 100% on the Eyewear2013A Dataset when the distance threshold is empirically chosen as 2985. Thirdly, LBP based method can work reliably under different illuminations, because LBP is an illumination-robust feature [16].

Furthermore, we evaluated the system latency, which is the time delay from the time that an image is captured to the time of face detection, face recognition, and showing the result on the eyewear display. Figure. 8 plots the system latencies on all the 41 video sessions. The mean latency is around 0.80s. This is much less than the minimum response time when a patient meets another person in the real world scenarios. In cases that a person moves fast enough to disappear in one second and the system happens to not detect that person, the patient does not have to make any social interaction with that person. Thus our system fits for daily life very well.

IV. CONCLUSIONS

This paper proposes an effective assistive system to perform face recognition in real-time for prosopagnosic patients. It has the capacity of providing face recognition assistance to the patients and improving their social interactions via displaying the tagged identity information using smartphone controlled eyewear. By its nature, the proposed wearable system is different from the current treatments for prosopagnosics that uses “secondary clue” strategies. We tested our solution on the video data of 20 different subjects interacting in their daily lives. The system is able to detect faces with an average accuracy of 99%, verify and recognize faces respectively with an average accuracy of 96% and 100% with acceptable latency of 0.80s. The experimental results have demonstrated the accuracy of the technologies and the overall usability of our system.

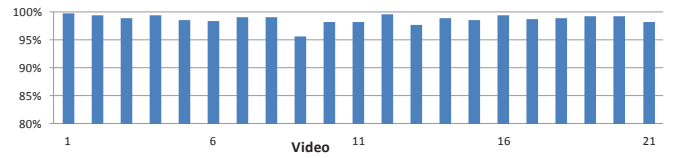


Fig. 7. The face detection rate of the software component in the proposed system evaluated on the Eyewear2013A dataset. Each bar corresponds to a video session in the dataset.

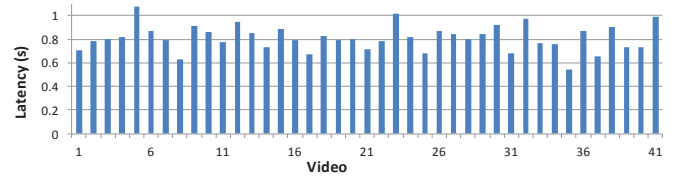


Fig. 8. The running latency of the software component in the proposed system evaluated on the Eyewear2013A dataset. Each bar corresponds to a video session in the dataset (41 in total).

REFERENCES

- [1] E. D. Haan and R. Campbell, “A fifteen year follow-up of a case of developmental prosopagnosia,” *Cortex, a Journal Devoted to the Study of the Nervous System and Behavior*, vol. 27, no. 4, pp. 489–509, December 1991.
- [2] J. M. Davis, E. McKone, H. Dennett, K. B. O’Connor, R. O’Kearney, and R. Palermo, “Social perception of faces in acquired prosopagnosia,” *the Public Library of Science*, vol. 6, no. 12, pp. 1–7, December 2011.
- [3] S. Bate, *Face recognition and its disorders*. UK: Palgrave Macmillan, 2013.
- [4] I. Kennerknecht, T. Grueter, B. Welling, S. Wentzek, J. Horst, S. Edwards, and M. Grueter, “First report of prevalence of non-syndromic hereditary prosopagnosia (HPA),” *American Journal of Medical Genetics*, vol. 140, no. 15, pp. 1617–1622, August 2006.
- [5] R. Russell, B. Duchaine, and K. Nakayama, “Super-recognizers: People with extraordinary face recognition ability,” *Psychonomic Bulletin & Review*, vol. 16, no. 2, pp. 252–257, May 2009.
- [6] T. Grter, M. Grter, and C. Carbon, “Neural and genetic foundations of face recognition and prosopagnosia,” *Journal of Neuropsychol*, vol. 2, no. 1, pp. 79–97, March 2008.
- [7] O. Sacks, “Why are some of us terrible at recognizing faces,” *The New Yorker*, pp. 36–43, August 2010.
- [8] E. Mayer and B. Rossion, “Prosopagnosia,” *The Behavioral and Cognitive Neurology of Stroke*, pp. 316–335, 2006.
- [9] A. C. Little, B. C. Jones, and L. M. DeBruine, “The many faces of research on face perception,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 366, no. 1571, pp. 1634–1637, May 2011.
- [10] J. B. Sreekar Krishna, Greg Little and S. Panchanathan, “A wearable face recognition system for individuals with visual impairments,” in *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility*, NY,USA, October 2005, pp. 106–113.
- [11] Wearable glasses technology. http://www.vuzix.com/augmented-reality/products_star1200xl.html.
- [12] P. Viola and M. Jones, “Robust real-time object detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, December 2006.
- [13] T. Ahonen, A. Hadid, and M. Pietikainen, “Face description with local binary patterns: Application to face recognition,” *ICCV 2001 Workshop on Statistical and Computation Theories of Vision*, vol. 28, no. 12, October 2001.
- [14] Yale face database. <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.
- [15] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, July 1997.
- [16] C. ho Chan, J. Kittler, and K. Messer, “Multi-scale local binary pattern histogram for face recognition,” in *Proceedings of the 2007 international conference on Advances in Biometrics*, vol. 4647, 2007, pp. 809–818.