Gaze-based Online Face Learning and Recognition in Augmented Reality

Takumi Toyama
German Research Center for Artificial Intelligence (DFKI) GmbH
Kaiserslautern, Germany
Takumi.Toyama@dfki.de

Daniel Sonntag
German Research Center for Artificial Intelligence (DFKI) GmbH
Saarbrücken, Germany
Daniel.Sonntag@dfki.de

Markus Weber
German Research Center for Artificial Intelligence (DFKI) GmbH
Kaiserslautern, Germany
Markus.Weber@dfki.de

Christian Schulz
German Research Center for Artificial Intelligence (DFKI) GmbH
Saarbrücken, Germany
Christian.Schulz@dfki.de

IUI'13

Abstract
We propose a new online face learning and recognition approach using user gaze and augmented displays. User gaze is used to select a face in focus in a scene image whereupon visual feedback and information about the detected person is presented in a head mounted display. Our specific medical application leverages the doctor's capabilities of recalling the specific patient context.

Author Keywords
Eye Tracking; Augmented Reality; Face Recognition; Real-time Interaction

ACM Classification Keywords
H.5.2. User Interfaces: Input Devices and Strategies, Graphical HCIs, Prototyping

General Terms
Experimentation, Human Factors, Performance

Introduction
Interaction technologies have become mobile and hence thoroughly integrated into everyday objects and activities. Embedded computing devices allow us to access digital resources in information systems more easily even in professional life. In fact, augmented reality has aroused people's attention [1]. Recent see-

through head mounted display (HMD) technology has a great potential for the future of augmented reality. We propose a new interactive machine learning (ML) environment for them: online face learning by combining a see-through HMD and a wearable eye tracker. The combined system allows a doctor to learn new patient faces online: he or she provides the name labeling by looking at the current patient whose features are extracted in real-time. At a later stage, our face classifier functions as an "external brain" for the doctor in order to recall the specific patient.

System Description

First, the eye tracker captures the scene image in front of the doctor and computes the gaze position on the image. The image and the gaze data is piped to the face detection module and the nearest detected face to the gaze position is selected as the face in focus. Then, the face's ML features are sent to the learning module or the recognition module (figure 1).

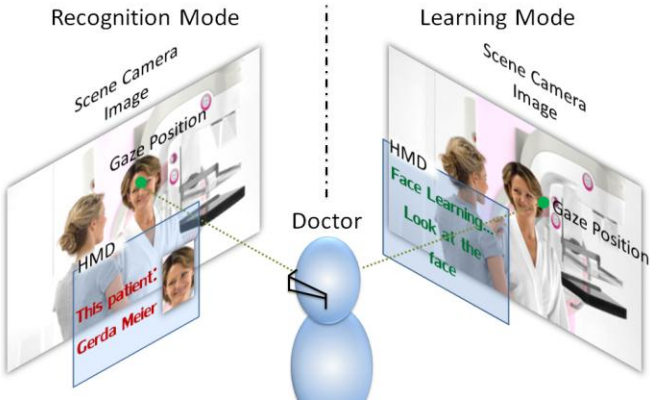

Figure 1: Left: Recognition mode providing patient information Right: Learning Mode.

Learning mode

Texts are prompted in the HMD to navigate the doctor to look at the face that he or she wants to add to the face database. After a new face in focus is detected (the learning mode is contextualized), the local binary pattern (LBP) features [2] from the face images are stored in the database.

Recognition mode

The LBP features of the face in focus are extracted (similar to learning mode). Nearest neighbor (*NN)* search is then applied to the database in order to find the person. After checking that the doctor is looking at the same person for a while, the retrieval results are displayed in the HMD.

Learning Results

In a preliminary test, we used 5 images (n=5) from 8 different persons for training, and test images were taken in the same lighting conditions. We achieved 100% of precision rate and 68% of recall rate. This result indicates that the system performs reasonably in our context where the doctors have to detect patients in the hospital in quite stable lighting conditions for small samples. The precise recognition of 100-500 faces is the next goal.

Acknowledgements

This work is part of THESEUS-RadSpeech (see www.dfki.de/RadSpeech/), supported by the German Federal Ministry of Economics and Technology (01MQ07016) and Nuance Foundation (see www.nuancefoundation.org/).

I/O devices used in this system

SMI Eye Tracking Glasses: The scene camera is located in the center.

Brother Airscouter (HMD)

Person wearing the devices

References

[1] D. W. F. van Krevelen, R. Poelman. A Survey of Augmented Reality Technologies, Applications and Limitations. *The International Journal of Virtual Reality*, Vol. 9, No. 2. (June 2010), pp. 1-20.

[2] Ahonen, T., Hadid, A., and Pietikäinen, M. Face description with local binary patterns: Application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell*. 28, 12 (2006), 2037–2041.