Human Centered Computer Vision

Prof. Dr.-Ing. Hazım Kemal Ekenel

ekenel@itu.edu.tr

Istanbul Technical University

Outline

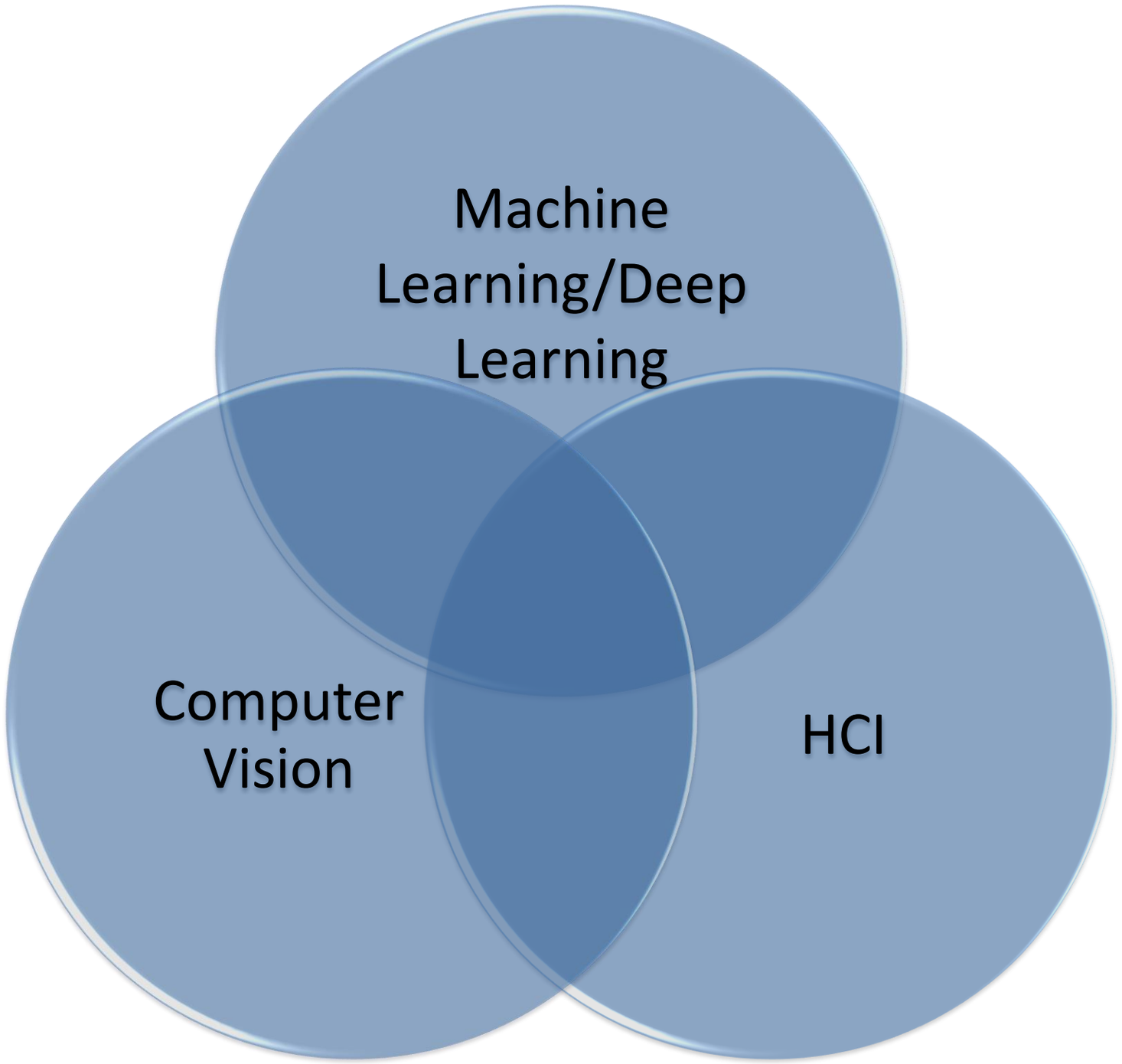
- Overview of Research Activities
- Exploiting Convolution Filter Patterns for Transfer Learning
- Domain Adaptation / Two-stage Fine-tuning
- Context Adaptation Using CycleGAN
- Future Research Directions

About myself

- Contact:
  - E-mail: ekenel@itu.edu.tr
  - Web: http://web.itu.edu.tr/ekenel/
- Prof. @ Dept. of Computer Engineering, Istanbul Technical Univ. (ITU)
- Host Prof. @ Signal Processing Lab 5 (LTS5), EPFL
- Has founded Facial Image Processing and Analysis Group (face.cs.kit.edu), Faculty of Computer Science in Karlsruhe Institute of Technology (KIT)
- Faculty member of International Center for Advanced Communication Technologies (interACT) ---CMU, HKUST, Waseda Univ.
- Awarded within the framework of the Excellence Initiative of German Research Foundation (DFG) in 2010.
- Received the EBF European Biometric Research Award 2008.
- Received PhD in Computer Science in Feb. 2009 from KIT, BS'01 and MS'03 in EE from Boğaziçi University

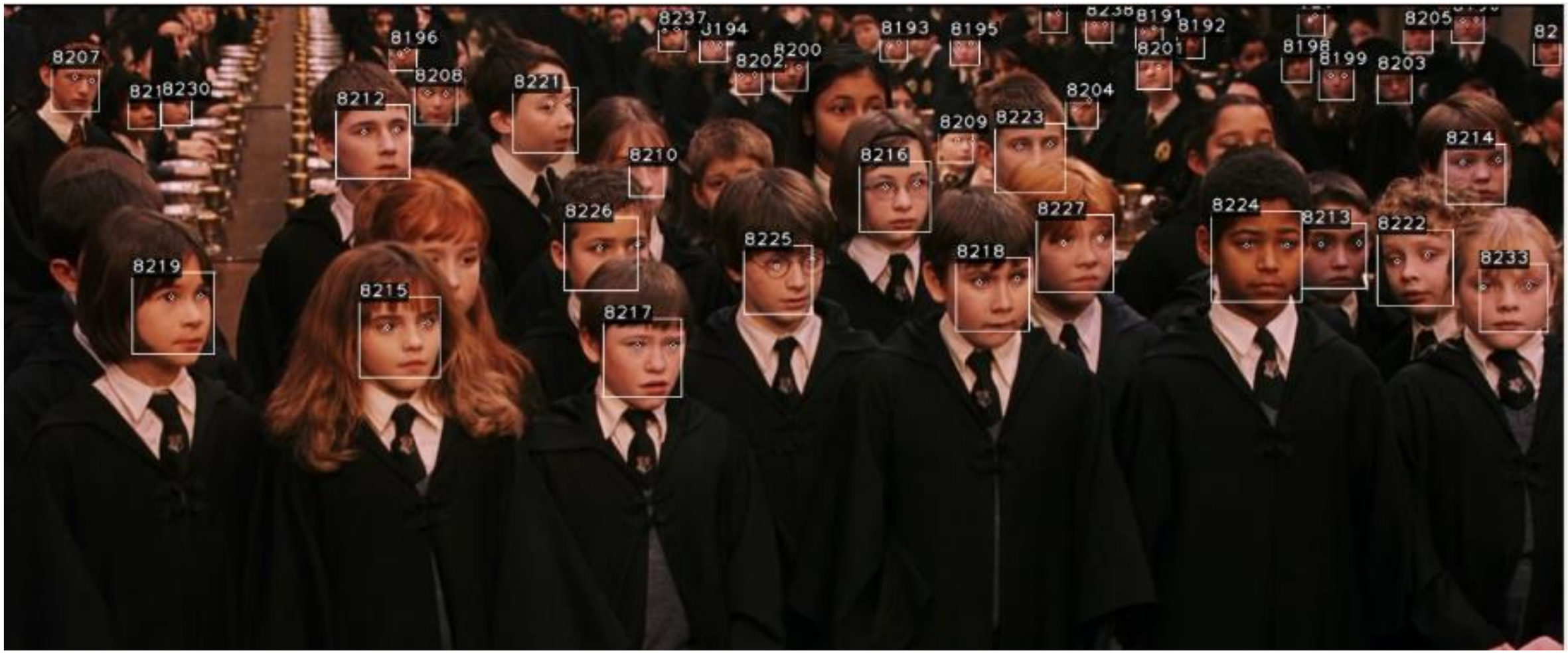
SiMiT Lab 2017 (simitlab.itu.edu.tr)

Research Areas



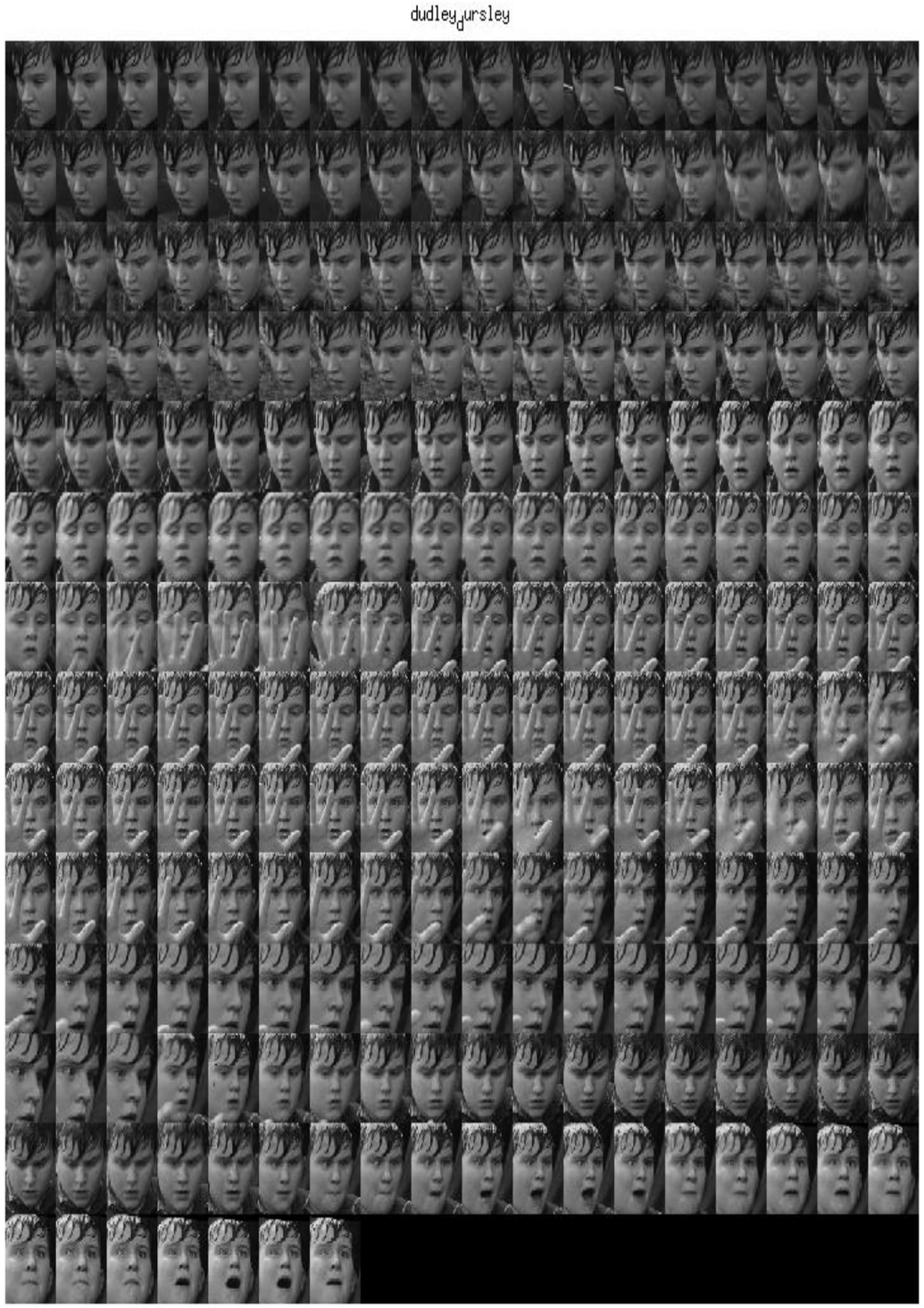
Overview of Research Activities

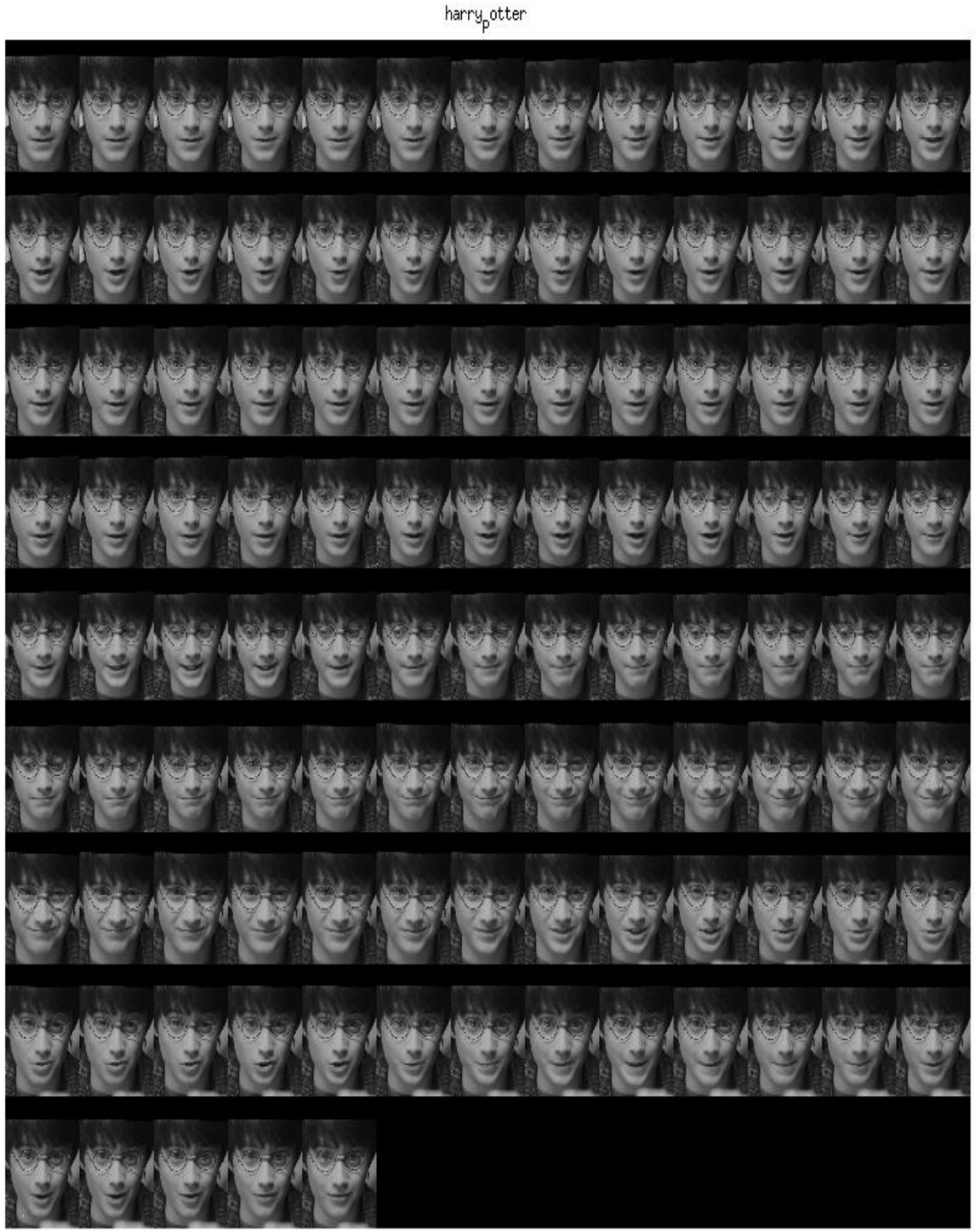
Face Detection & Tracking



* E. Ghaleb, M. Tapaswi, Z. Al-Halah, H.K. Ekenel, R. Stiefelhagen, "Accio: A Data Set for Face Track Retrieval in Movies Across Age", ACM International Conference on Multimedia Retrieval (ICMR), Shanghai, China, June 2015.

Face Tracks Examples





* Face tracks are available @ https://cvhci.anthropomatik.kit.edu/~mtapaswi/publications.html

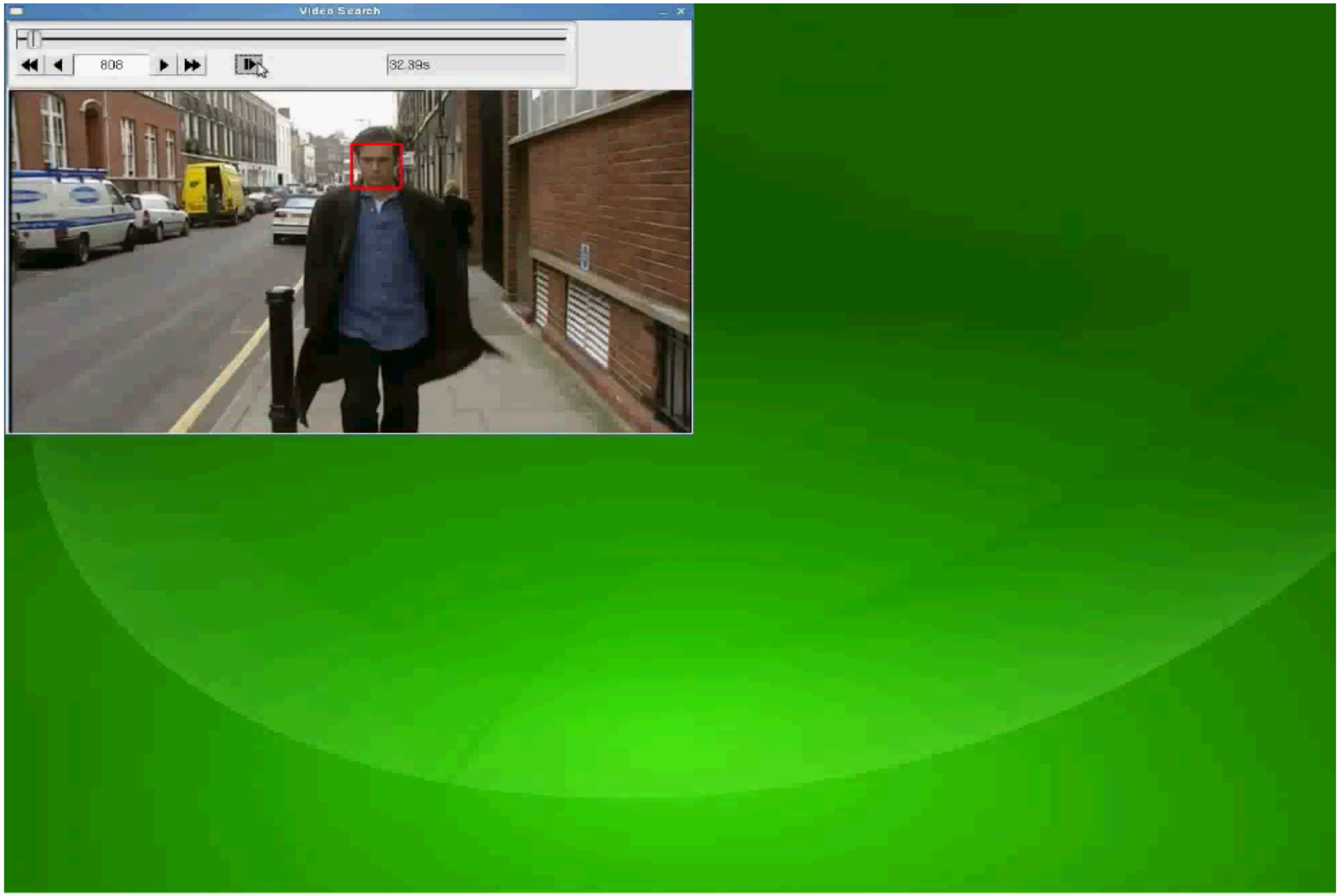
Person Tracking in TV Series

* M. Tapaswi, C.Ç. Çörez, M. Bäuml, H.K. Ekenel, R. Stiefelhagen, "Cleaning Up After a Face Tracker: False Positive Removal", IEEE International Conference on Image Processing (ICIP), Paris, France, October 2014.

Video-based Face Recognition

* J. Stallkamp, H.K. Ekenel, R. Stiefelhagen, "Video-based Face Recognition on Real-World Data", In Proc. of International Conference on Computer Vision (ICCV'07), pp. 1-8, Rio de Jenario, Brasil, October 2007.

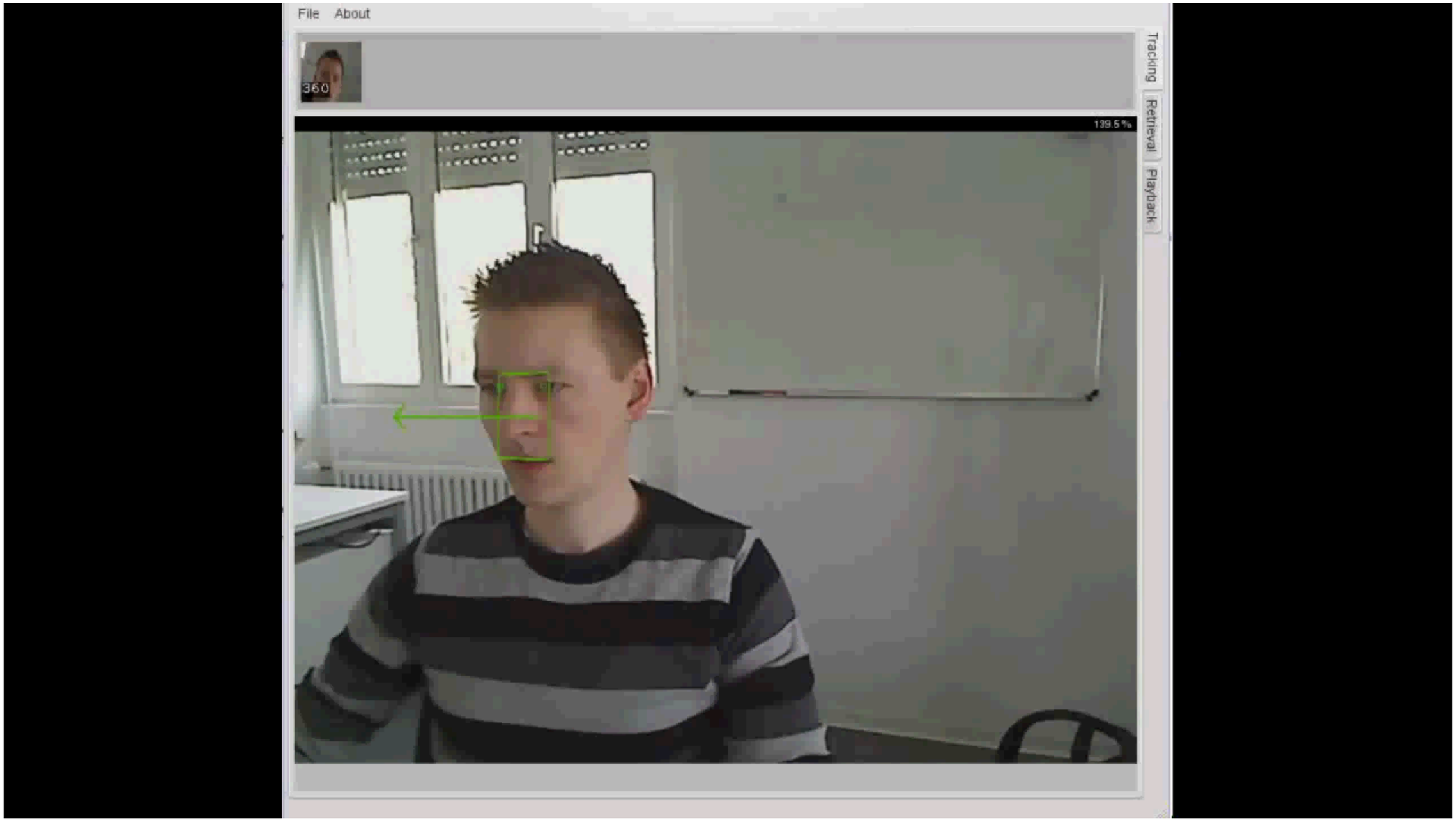
Interactive Person Retrieval in Movies



* M. Fischer, H.K. Ekenel, R. Stiefelhagen, “Person Re-identification in TV Series Using Robust Face Recognition and User Feedback”, Springer Multimedia Tools and Applications, Vol. 55, No. 1, pp. 83-104, 2011.

* M. Fischer, H.K. Ekenel, R. Stiefelhagen, “Interactive Person Re-identification in TV Series”, In Proc. of Intl. Workshop on Content-based Multimedia Indexing (CBMI’10), Grenoble, France, June 2010.

Interactive Person Retrieval in Camera Networks

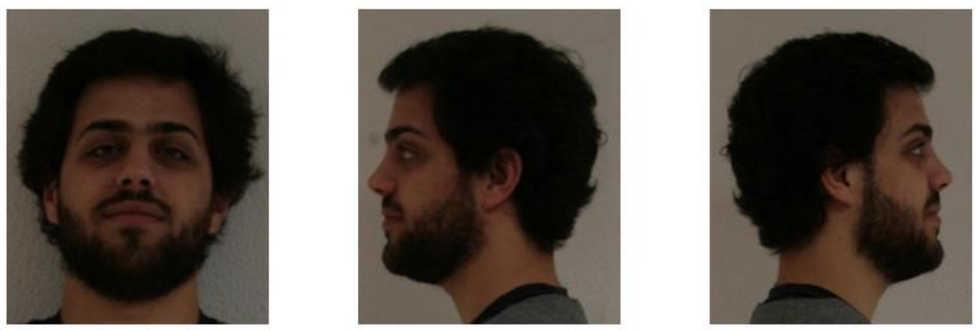


* M. Bäuml, K. Bernardin, M. Fischer, H.K. Ekenel, R. Stiefelhagen, “Multi-Pose Face Recognition for Person Retrieval in Camera Networks”, In Proc. of 7th IEEE Intl. Conf. on Advanced Video and Signal-Based Surveillance (AVSS), 2010.

Face Recognition in the Wild

- International Challenge on Biometric Recognition in the Wild 2016
- 90 Subjects: 3 images for each subject in gallery set and
  5 images in probe set

Gallery Set: Example



1) Frontal Image 2) Left-side Image 3) Right-side Image

Probe Set: Example

003_01.jpg 003_02.jpg 003_03.jpg

003_04.jpg 003_05.jpg

* E. Ghaleb, G. Özbulak, H. Gao, H.K. Ekenel, “Deep Representation and Score Normalization for Face Recognition Under Mismatched Conditions”, IEEE Intelligent Systems, accepted for publication.

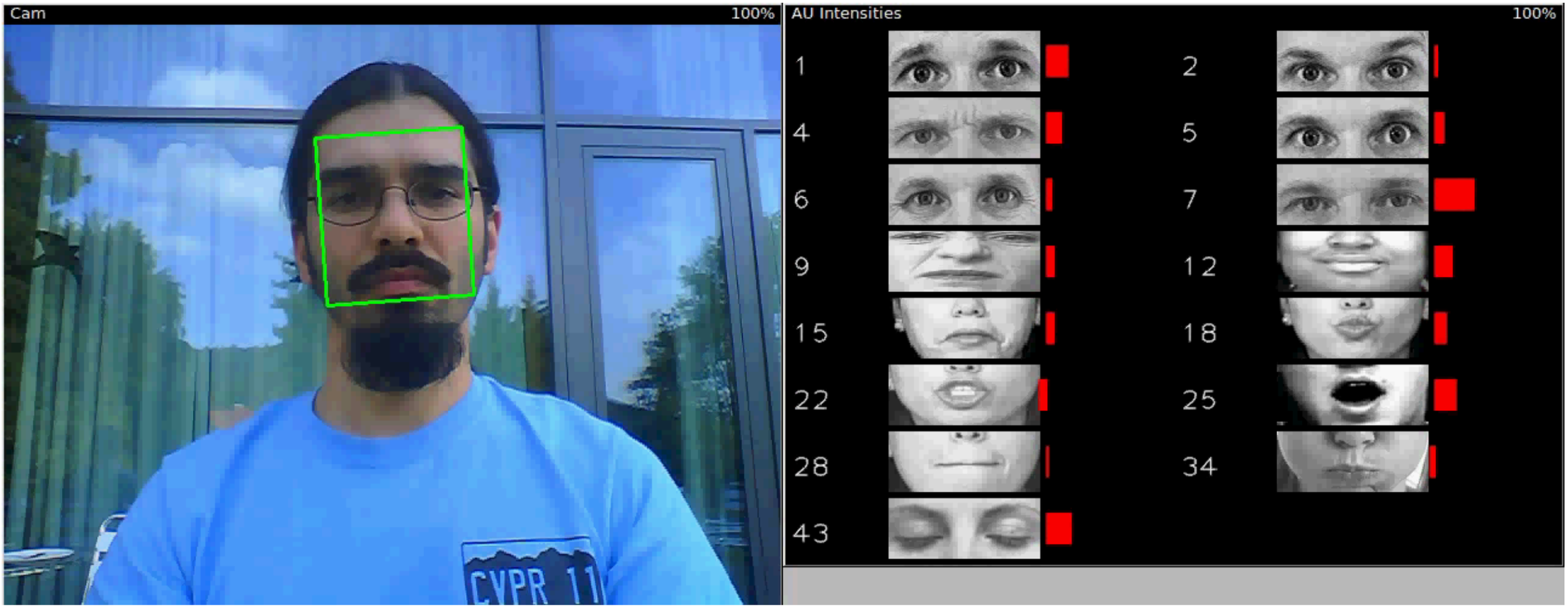
Table 1. **Final results for the ICB-RW competition.** The nine valid submissions of ICB-RW are listed according to their rank in the challenge. Methods are ranked according to the AUC of the CMC curve. Also, both rank-1 and rank-5 identification rate (IR) are presented, along with a brief description provided by the authors.

Method	Description	Rank-1 IR (%)	Rank-5 IR (%)	AUC (CMC curve)
H. Ekenel, G. Ozbulak, E. Ghaleb Istanbul Technical University	The probe and gallery face images are aligned with respect to eye centers. Only the frontal images are used as gallery. Face representation is extracted from a convolutional neural network (CNN) with a VGG face model [16]. In the test phase, the nearest neighbour classifier is used with the correlation distance as the similarity score.	69.8	85.3	0.954
K. Grm, S. Dobrisek, V. Struc University of Ljubljana	An augmented dataset was generated through oversampling the training images via bounding box noise and horizontal flipping. The pre-trained VGG face deep convolutional network [16] was used to extract features from the images. Then, a softmax classifier was trained on the features.	62.0	78.7	0.952
H. Shi, X. Zhu, S. Liao, Z. Lei, S. Li Institute of Automation, Chinese Academy of Sciences	Features are extracted from a deep convolutional network model trained on the CASIA-Webface database and the cosine similarity is used as score. Ten models learned from different facial parts are fused, and the gallery images of different poses are synthesized to ease the matching phase.	57.6	75.8	0.921
W. Gutfeter NASK, Warsaw University of Technology	The algorithm builds similarity scores by merging results obtained from a set of convolutional neural networks trained for recognizing faces from different angles.	42.9	64.4	0.918
J. Brogan University of Notre Dame	Gallery and probe images are frontalized using a modified version of [2]. Data features are extracted from a SLMSimple Neural Network [1] and four bins are created containing different versions of the gallery images. Probe descriptors are matched with one of the four bins according to yaw angle of the head, and the resulting pairs of feature vectors are input into an SVM trained with LFW [7] data.	11.6	30.4	0.755
E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez University Autonoma de Madrid	The LBPs [13] of nine facial regions are extracted from a frontalized image [22] followed by illumination normalization. A fused distance score is determined by only considering the five best individual facial regions at each trial.	24.0	39.1	0.725
D. Riccio, M. Nappi University of Salerno	The algorithm locates facial points through an extended Active Shape Model and remaps the face region to a 64x100 image. It applies a local normalization process to correct illumination variations, and the matching is performed with a	11.1	25.1	0.694

Paintball/Lasertag Style Game Using Face Recognition

* U. Demir, E. Ghaleb, H.K. Ekenel, "A Face Recognition Based Multiplayer Mobile Game Application", 10th Intl. Conference on Artificial Intelligence Applications and Innovations (AIAI), Rhodes, Greece, September 2014.

Facial Expression Analysis



✓ Best Performing European System

Facial Expression Recognition and Analysis Challenge in 2011 (FERA 2011)

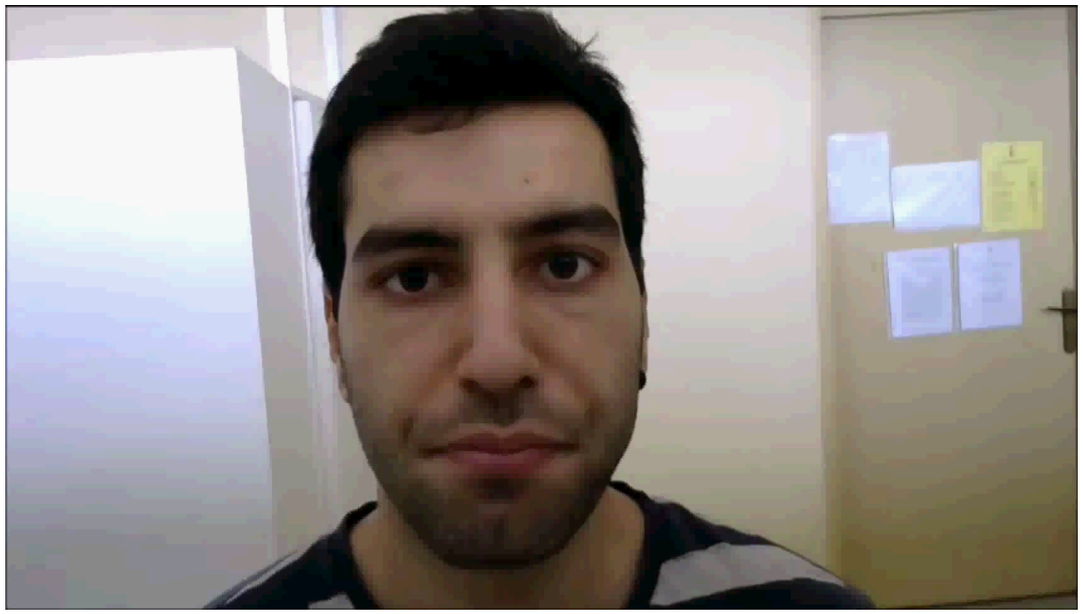
Emotion Recognition in the Wild Challenge (EmotiW 2013)

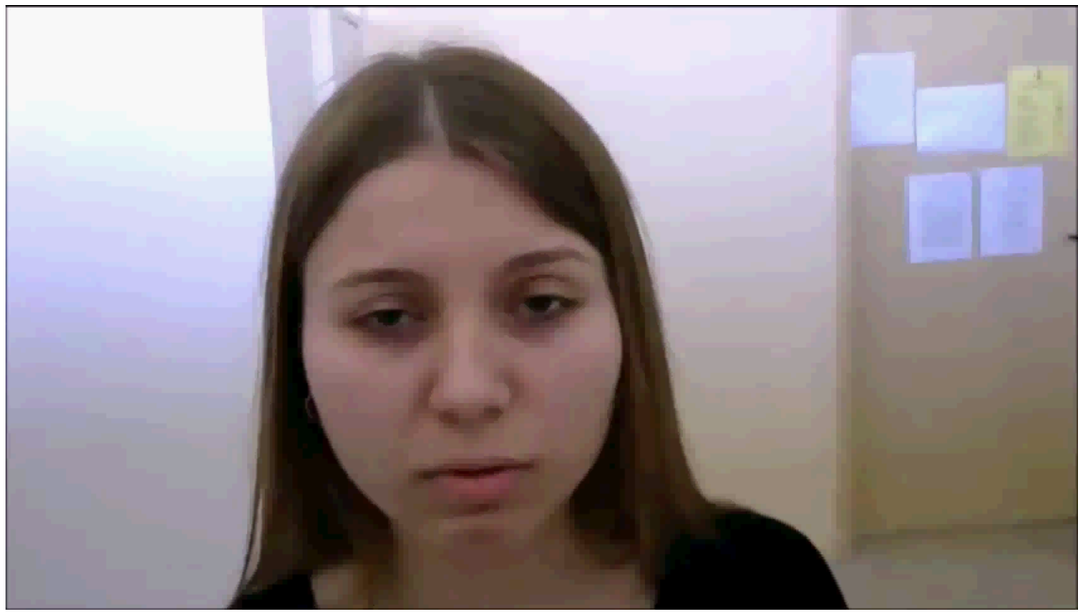
✓ Fast and practical

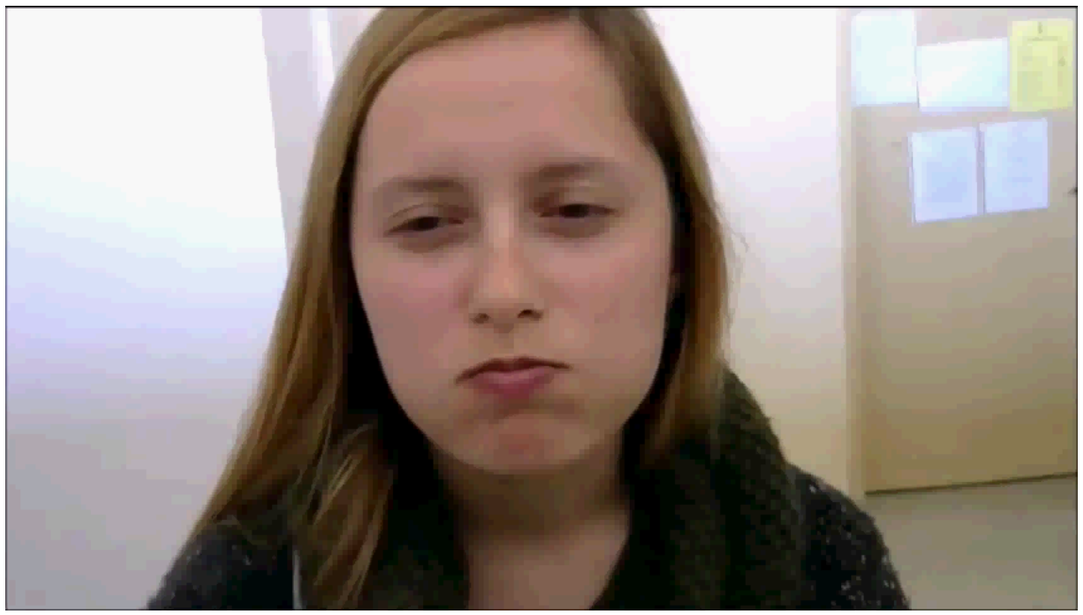
A Mobile Game to Teach Facial Expressions

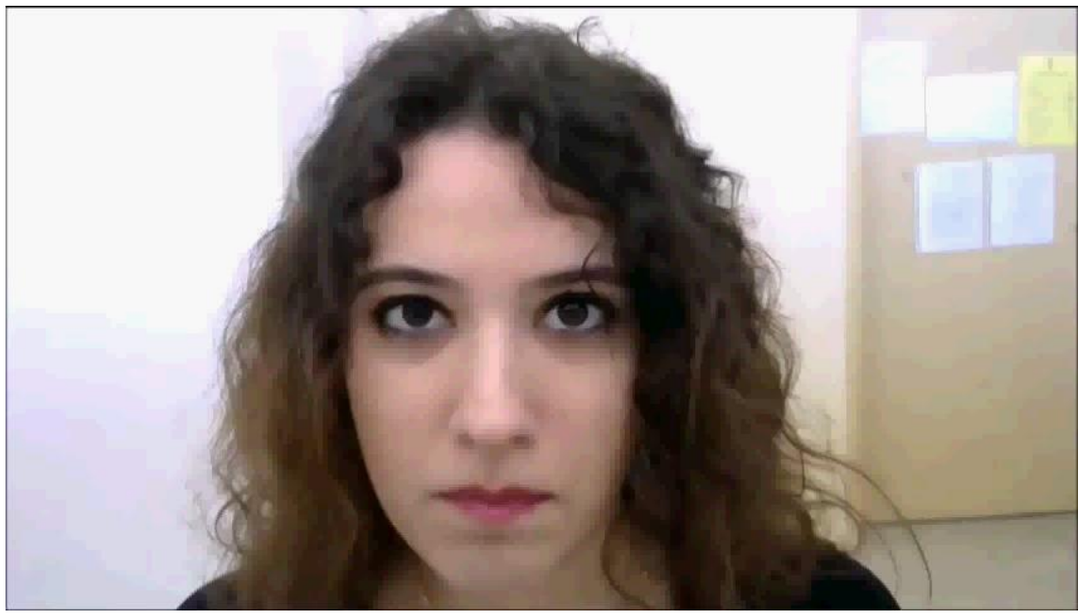
Facial Expression Recognition as a Pair Matching Problem

Appreciation Estimation



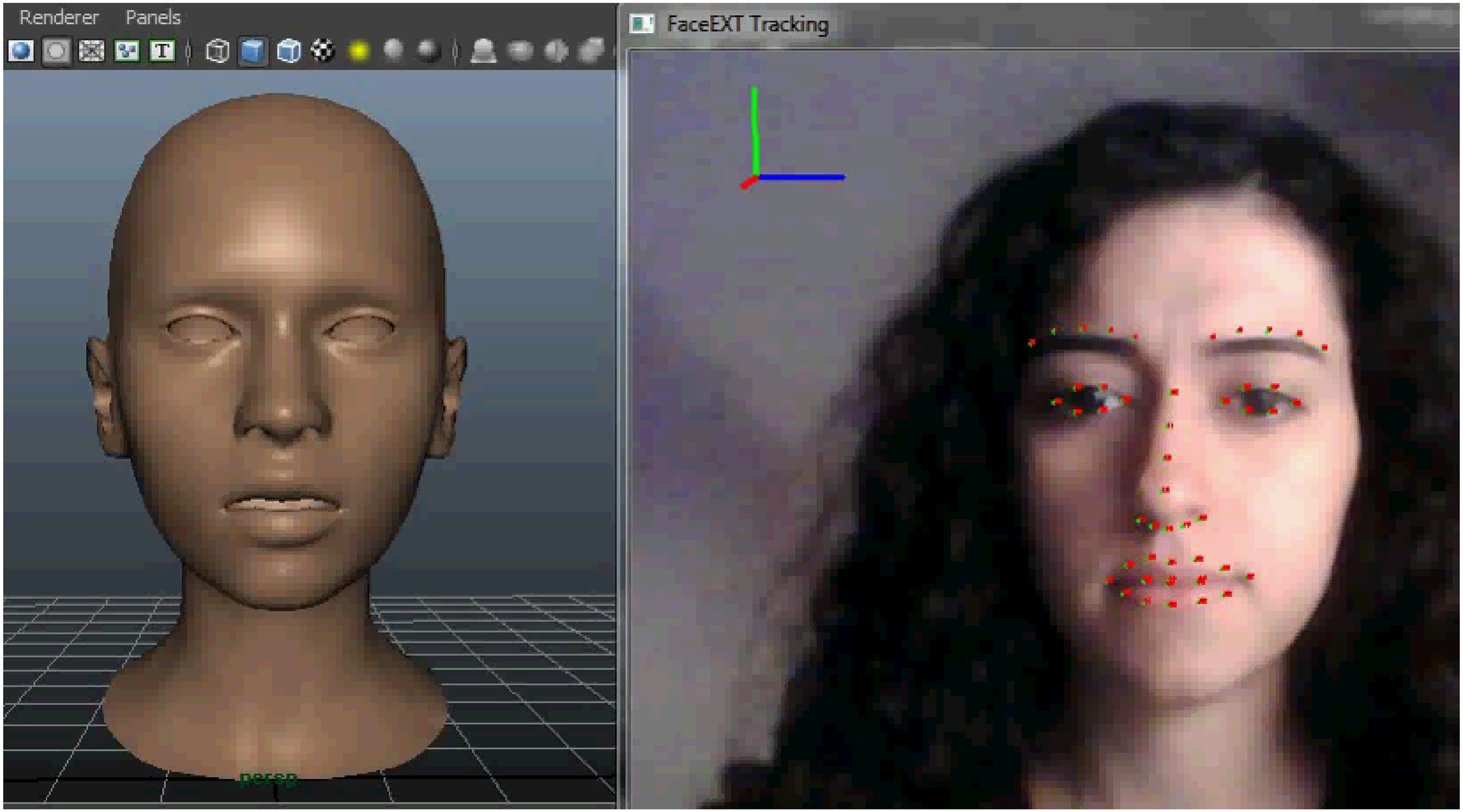






* İ.E. Zülfikar, H. Dibeklioğlu, H.K. Ekenel, "A Preliminary Study on Visual Estimation of Taste Appreciation", IEEE ICME International Workshop on Multimedia for Cooking and Eating Activities", Seattle, USA, July 2016.

Face Animation / Expression Transfer



* B. Ekmen, H.K. Ekenel, "Real Time Animated Facial Expression Tranfer", IEEE Signal Processing and Communications Applications Conference, May 2016.

Real-time Face Swapping

* N.M. Arar, F. Güney, N.K. Bekmezci, H. Gao, H.K. Ekenel, “Real-time Face Swapping in Video Sequences - Magic Mirror”, International Conference on Computer Vision Theory and Applications (VISAPP), Rome, Italy, February 2012.

Ethical Concerns

- Privacy
- Data Bias
- Security

Human Centered Computer Vision

Prof. Dr.-Ing. Hazım Kemal Ekenel

ekenel@itu.edu.tr

Istanbul Technical University