Psychological and computational perspectives on recognition of moving faces

Symposium Moderator: Nicholas Costen

Symposium Presenters:

Darren Cosker
The University of Bath
Department of Computer Science,
Bath BA2 7AY UK
(D.P.Cosker@cs.bath.ac.uk)

Nicholas Costen and **Hui Fang**
Manchester Metropolitan University
Department of Computing and Mathematics,
Manchester M1 5GD UK
(n.costen@mmu.ac.uk)

Alan Johnston and **Harry Griffin**
University College London
Department of Psychology,
London, WC1E 6BT UK
(a.johnston@ucl.ac.uk)

Eva Krumhuber
Swiss Center for Affective Sciences,
University of Geneva
1205 Geneva, CH
(eva.Krumhuber@unige.ch)

Karen Lander and **Natalie Butcher**
The University of Manchester
School of Psychological Sciences,
Manchester M13 9PL UK
(karen.lander@manchester.ac.uk)

Paul Rosin and **David Marshall**
Cardiff University
School of Computer Science,
Cardiff, CF24 3AA UK
(Paul.Rosin@cs.cf.ac.uk)

Keywords: Perception, computation, face recognition, movement.

People

- **Nicholas Costen** is a Senior Lecturer in Computing at Manchester Metropolitan University. He holds degrees in both psychology and computer science, and has published extensively on the computational modeling of the perception of static and dynamic faces.
- **Darren Cosker** is a Royal Academy of Engineering / EPSRC Research Fellow at the University of Bath. His interests centre on human motion analysis and synthesis.
- **Alan Johnston** is Professor of Psychology at University College London. His group studies the perceptual mechanisms that underlie our capacity to see movement, depth and form.
- **Karen Lander** is a Senior Lecturer in Psychology at the University of Manchester. Her interests focus on the importance of face animation in the recognition of familiar faces.
- **Paul Rosin** is a Reader in Computing at Cardiff University. His interests include low-level computer vision, talking heads and 3D mesh processing.

Introduction

Our knowledge of object, scene and face recognition is largely based on studies that have used static stimuli. However, in reality, motion and change are fundamental aspects of our world. Do our perceptual/cognitive systems make use of these real world dynamics? The purpose of this symposium is to showcase a number of research projects from both Psychology and Computer Science that explore the impact of motion on our perception and representation of faces.

Overview of presentations

Darren Cosker and Eva Krumhuber

The modelling of faces is an area of intense research, and facial models - developed by computer graphics and vision researchers – have a wide range of applications. Arguably the most publicly visible result of facial research appears in movies and video games. These often astounding feats are the result of intensive labour and artistic license. Facial modelling also has major application areas in psychology and neuroscience. A powerful discovery from the application of such models in these fields is how the perception of a facial animation can change given even a small subtle variation in expression timing and duration (Krumhuber et al., 2007). An interesting relationship then exists between psychology research and the development of *human-realistic* facial models for the entertainment industry, i.e. if we are to achieve human-realism we also need to understand how faces are perceived. Under this theme, we will discuss the modelling of faces from 2D images and dynamic 3D facial data (i.e. 3D captured at 60 frames per second), parameterisation of such data for creation of facial animations, and their application (Cosker, Borkett, Marshall, & Rosin, 2008). We will also report on some example recent collaborative studies that highlight the power of manipulating photorealistic facial expressions to create different overall impressions of a person and influence decision-making.

Nicholas Costen and Hui Fang

Computational statistical models, derived from moving images, may be used in the face recognition process. As a counterpart to psychological experimental results showing a significant beneficial effect of facial non-rigid movement, two features obtained from face sequences, the central tendency and type of movement variation, are associated to improve face verification compared with single static images. By using a progressive optic-flow based correspondence algorithm (Fang & Costen, 2008), the correspondences across the sequences are learned and encoded on a combined shape and appearance model (Cootes, Edwards, & Taylor, 2001), parameterizing the face sequences. The parameters are projected to an identity-only space to find the central tendency of each subject. In addition, facial movement consistencies across different behaviors exhibited by the same subjects are recorded. These two features are fused by a confidence-based decision system for authentication applications. On a major video database, the results show that the extra information extracted from moving images significantly and efficiently improves performance.

Alan Johnston and Harry Griffin

Much of the psychological research on face perception and recognition has used static pictures of faces. However these are frozen configurations of a dynamic object that continually changes it's expression, to supply a communicative stream of information to the viewer, and adjusts its pose, indicating, to some degree, where the person's attention lies. To investigate the perception of facial motion and dynamic face recognition ideally we would like to separate out the motion from the form. Past techniques have included viewing photographic negatives, which disrupts 3D shape (Knight & Johnston, 1997) but leaves 2D motion intact or performance-driven animation of rendered avatars controlled by markers or feature points on an actors face (Hill & Johnston, 2001). The problems of marker driven animation is that the motion vectors are typically sparse and the resulting stimulus may not accurately represent the subtle motion of the face that project essential cues to facial identity or expression semantics. We will review the conclusions of this work and describe a new markerless approach we are taking to constructing performance-driven photorealistic avatars and how we intend to extend this work to dynamic 3D facial structure.

Karen Lander and Natalie Butcher

Previous psychological research suggests that most benefit, for familiar face recognition, is gained when the observed motion retains its original characteristics compared with speeded up, slowed down or rhythm disrupted motion (Lander & Bruce, 2000). It is interesting to speculate how this 'dynamic' information might be stored in memory. One possibility is that the stored representation of familiar faces may themselves be dynamic in nature, intrinsically linking specific motion characteristics to specific identities (Lander & Bruce, 2004). Our recent work aims to expand our current knowledge about which parameters of the to-be-recognised face influence the motion advantage, and how these parameters may be interlinked. Here we describe three experiments designed to investigate the interrelationships between motion distinctiveness, face distinctiveness and the amount of motion exhibited by a face, and how they impact the identification of individual faces. Finally we investigate whether these motion parameter ratings are linked to the clip shown or related to famous face identity (consistent across video clips).

Paul Rosin and David Marshall

Recent research (Knappmeyer, Thornton, & Blthoff, 2003) has analysed the effects of facial motion on face processing (e.g. recognition of gender, age, identity). Computer vision has also experimented with acquisition and analysis of 3D facial data. However, it is only in the last few years that hardware technology has become available to easily obtain 4D facial data, i.e. 3D meshes captured at video rate. Our current work is investigating face dynamics as a physiological biometric. The dynamics are performed by the subjects uttering a single spoken word to create "verbal facial actions". The face is modelled by a 3D Active Shape Model (Cootes et al., 2001) which enables the face to be described as a dynamic signal consisting of its trajectory in face space. The similarity between two faces is then be computed using a modification of Derivative Dynamic Time Warping (Keogh & Pazzani, 2001) between the two dynamic signals. Our preliminary results provide evidence that the dynamics of even very short facial actions contain sufficient information for identity recognition (Benedikt, Kajic, Cosker, Rosin, & Marshall, 2008).

References

Benedikt, L., Kajic, V., Cosker, D., Rosin, P. L., & Marshall, D. (2008). Facial dynamics in biometric identification. In *BMVC* (p. 1075-1084).

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE PAMI*, *23*(6), 681–685.

Cosker, D., Borkett, R., Marshall, D., & Rosin, P. (2008, September). Towards automatic performance-driven animation between multiple types of facial model. *Computer Vision, IET*, *2*(3), 129-141.

Fang, H., & Costen, N. P. (2008). Tracking face localization with a hierarchical progressive face model. In J. Gonzàliz (Ed.), *Tracking humans for the evaluation of their motion in image sequences* (pp. 89–99).

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology*, *11*(11), 880–885.

Keogh, E. J., & Pazzani, M. J. (2001). Derivative dynamic time warping. In *First SIAM International Conference on Data Mining.*

Knappmeyer, B., Thornton, I., & Blthoff, H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research*, *43*(18), 1921-1936.

Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition*, *4*, 265–273.

Krumhuber, E., Manstead, A., Cosker, D., Marshall, D., Rosin, P., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion*, *7*, 730-735.

Lander, K., & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology*, *12*(4), 259–272.

Lander, K., & Bruce, V. (2004). Repetition priming from moving faces. *Memory and Cognition*, *32*(4), 604–647.