

Facial Gesture Recognition from Static Dual-view Face Images

Maja Pantic
Delft University of Technology
Faculty of ITS / Mediamatics Dept.
Delft, the Netherlands
M.Pantic@cs.tudelft.nl

ABSTRACT

The research presented here pertains to the problem of automatic facial gesture analysis. Our major impulse to investigate this problem comes from the significance of the information that the face provides about the human behaviour. Facial gestures (facial muscle activity underlying a facial expression) regulate our social interactions: they clarify whether our current focus of attention (a person, an object or what has been said) is important, funny or unpleasant for us. They are our direct, naturally preeminent means of communicating emotions. Automatic analysers of subtle facial changes, therefore, seem to have a natural place in various vision systems including the automated tools for psychological research, lip reading, bimodal speech analysis, affective computing, videoconferencing, face and visual speech synthesis, and human-behaviour-aware next-generation interfaces.

Within our research on automatic facial gesture analysis, we first investigated whether and to which extent human facial gestures could be recognised automatically from a static face image. We developed an automatic system to analyse subtle changes in facial expressions based on both fiducial points of the contours of the permanent facial features (eyes, mouth, etc.) and fiducial points of the face-profile contour in a static dual view (i.e., frontal and profile view) of the face. Fig. 1 outlines the developed system.

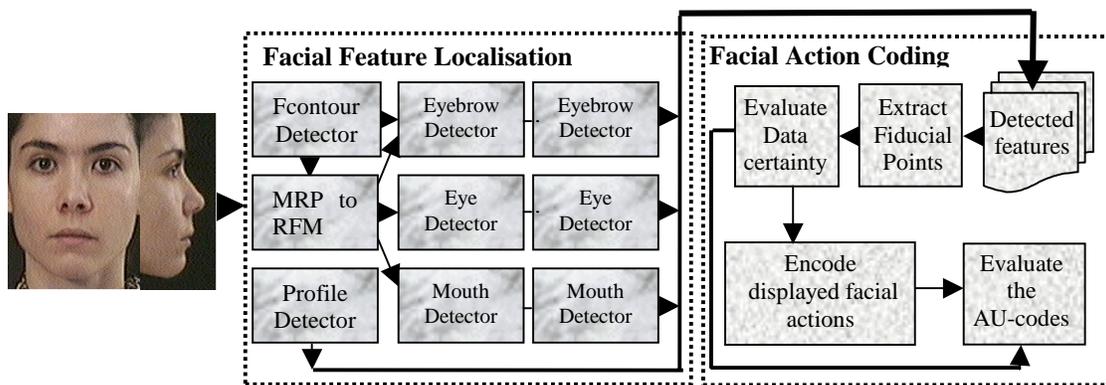


Fig 1: Outline of the dual-view-based AU recognition

A hybrid approach to facial feature localisation has been proposed for spatial sampling of the profile contour and the contours of the permanent facial features. This hybrid facial-feature detector employs various image processing techniques such as active contours, fuzzy edge detection and artificial neural networks. From the extracted contours of the facial features, we extract 10 profile-contour fiducial points and 19 fiducial points of the contours of the permanent facial features (Fig. 2).

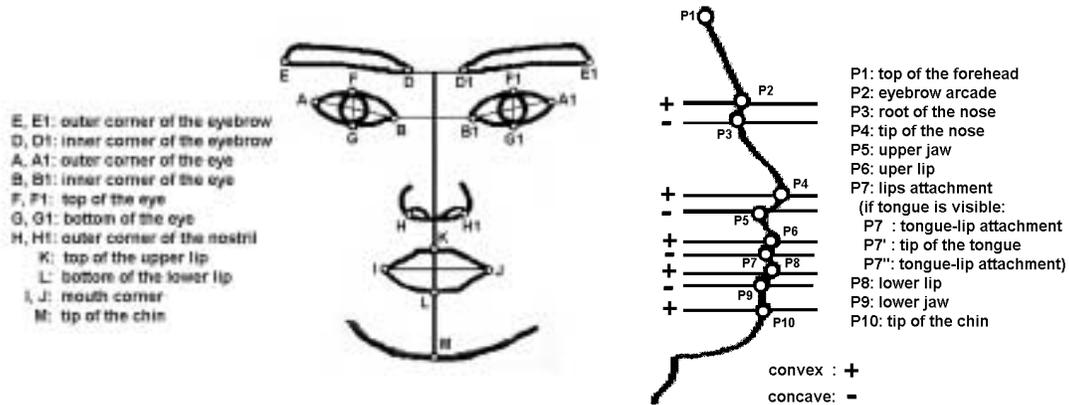


Fig. 2: Fiducial points

By comparing the examined facial expression with a neutral facial expression of the currently observed subject, subtle changes in the examined facial expression are measured next. Motivated by the facial action units (AUs) of the FACS system, these changes are represented as a set of mid-level feature parameters describing the state and motion of the fiducial points and the shapes formed between certain fiducial points. We defined six mid-level feature parameters in total: two describing the motion of the fiducial points, two describing their state, and two describing shapes formed between certain fiducial points. The definitions of the parameters are given in Fig. 3.

Fiducial points motion	
$up/down(P) = y_{P_{neutral}} - y_{P_{current}}$ If $up/down(P) < 0$, P moves up.	$in/out(P) = x_{P_{neutral}} - x_{P_{current}}$ If $in/out(P) > 0$, P moves outward.
Fiducial points state	
If $P9$ equals $P7$, $absent(P9)$. If there is no maximum of f' between $P5$ and $P7$, $absent(P6)$. Similarly for $P7'$, $P7''$ and $P8$ (see Fig. 2).	$increase/decrease(AB) = AB_{neutral} - AB_{current}$, where $AB = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}$ If $increase/decrease(AB) < 0$, distance AB increases.
Shapes formed by fiducial points	
The physic meanings of $angular(P6P8) = true$ and $increased_curvature(P5P6)$ are shown below.	

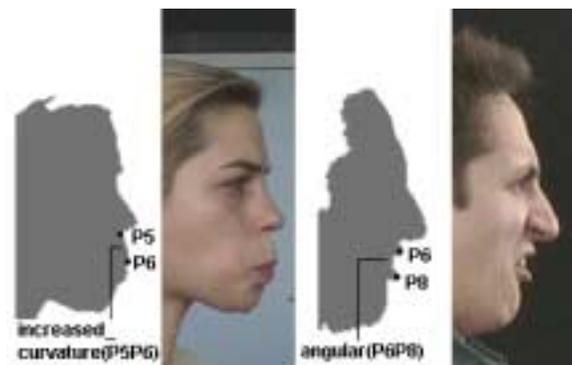


Fig. 3: Mid-level feature parameters for AU recognition

Based on these feature parameters, a fuzzy-rule-based algorithm interprets the extracted facial information in terms of 32 AUs (from a total of 44 possible AUs) occurring alone or in a combination. With each scored AU, the utilised algorithm associates a factor denoting the certainty with which the pertinent AU has been scored.

Our AU-encoder demonstrates rather high concurrent validity with manual FACS coding of test images; it achieves a correct recognition rate of 85%.