

Correlates of facial expressions in the primary visual cortex

Face detection and recognition should be complemented by recognition of facial expression, for example for social robots which must react to human emotions. Our framework is based on two multi-scale representations in cortical area V1: keypoints at eyes, nose and mouth are grouped for face detection [1]; lines and edges provide information for face recognition [2]. We assume that keypoints play a key role in the where system, lines and edges being exploited in the what system. This dichotomy, together with coarse-to-fine-scale processing, yields translation and rotation invariance, refining object categorisations until recognition, assuming that objects are represented by normalised templates in memory. Faces are processed the following way: (1) Keypoints at coarse scales are used to translate and rotate the entire input face, using a generic face template with neutral expression. (2) At medium scales, cells with dendritic fields at corners of mouth and eyebrows of the generic template collect evidence for expressions using the line and edge information of the (globally normalised) input face at those scales. Big structures, including mouth and eyebrows, are further normalised using keypoints and first categorizations (gender, race) are obtained using lines and edges. (3) The latter process continues until the finest scale, with normalisation of the expression to neutral for final face recognition. The advantage of this framework is that only one frontal view of a person's face with neutral expression must be stored in memory.

This procedure resulted from an analysis of the multi-scale line/edge representation of normalised faces with seven expressions: neutral, anger, disgust, fear, happy, sad and surprise. Following [3], where Action Units (AUs) are related to facial muscles, we analysed the line/edge representation in all AUs. We found that positions of lines and edges at one medium scale, and only at AUs covering the mouth and eyebrows, relative to positions in the neutral face at the same scale, suffice to extract the right expression. Moreover, by implementing AUs by means of six groups of vertically aligned summation cells with a dendritic field size related to that scale (sizes of simple and complex cells), covering a range of positions above and below the corners of mouth and eyebrows in the neutral face, the summation cell with maximum response of each of the six cell groups can be detected, and it is possible to estimate the degree of the expression, from mild to extreme. This work is in progress, since the method must still be tested using big databases with many faces and their natural variations. Perhaps some expressions detected at one medium scale must be validated at one or more finer scales. Nevertheless, in this framework detection of expression occurs before face recognition, which may be an advantage in the development of social robots.

[1] Rodrigues and du Buf 2006. *BioSystems* 86,75-90 [2] Rodrigues and du Buf 2009. *BioSystems* 95,206-26 [3] Ekman and Friesen 1978. *FACS*, Consulting Psychologists Press, Palo Alto

Acknowledgements: FCT funding of ISR-IST with POS-Conhecimento and FEDER; FCT projects PTDC-PSI-67381-2006 and PTDC-EIA-73633-2006

Correlates of facial expressions in the primary visual cortex

Ricardo Sousa^{1,2}, João Rodrigues^{1,2}, J.M.H. du Buf¹

¹Vision Laboratory – Institute for Systems and Robotics (ISR), ²Instituto Superior de Engenharia, University of the Algarve, Faro, Portugal

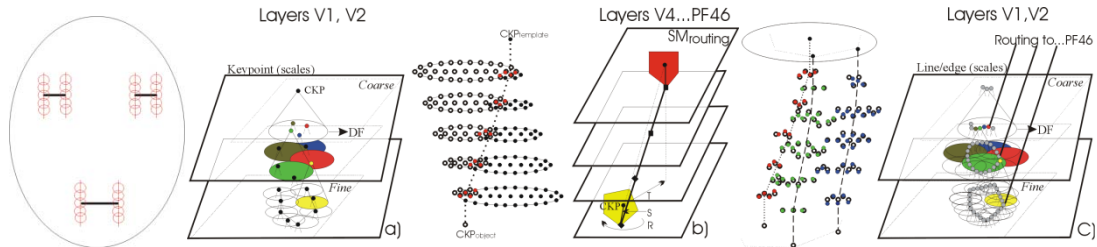


Fig. 1 – Left to right: six summation cell clusters at the corners of mouth and eyebrows; keypoint representation from coarse to fine scales; neural routing of keypoints from input to template in memory; routing of major keypoints for translation and rotation invariance; the same showing dynamic routing; the dynamic routing established by keypoints is applied to lines and edges.

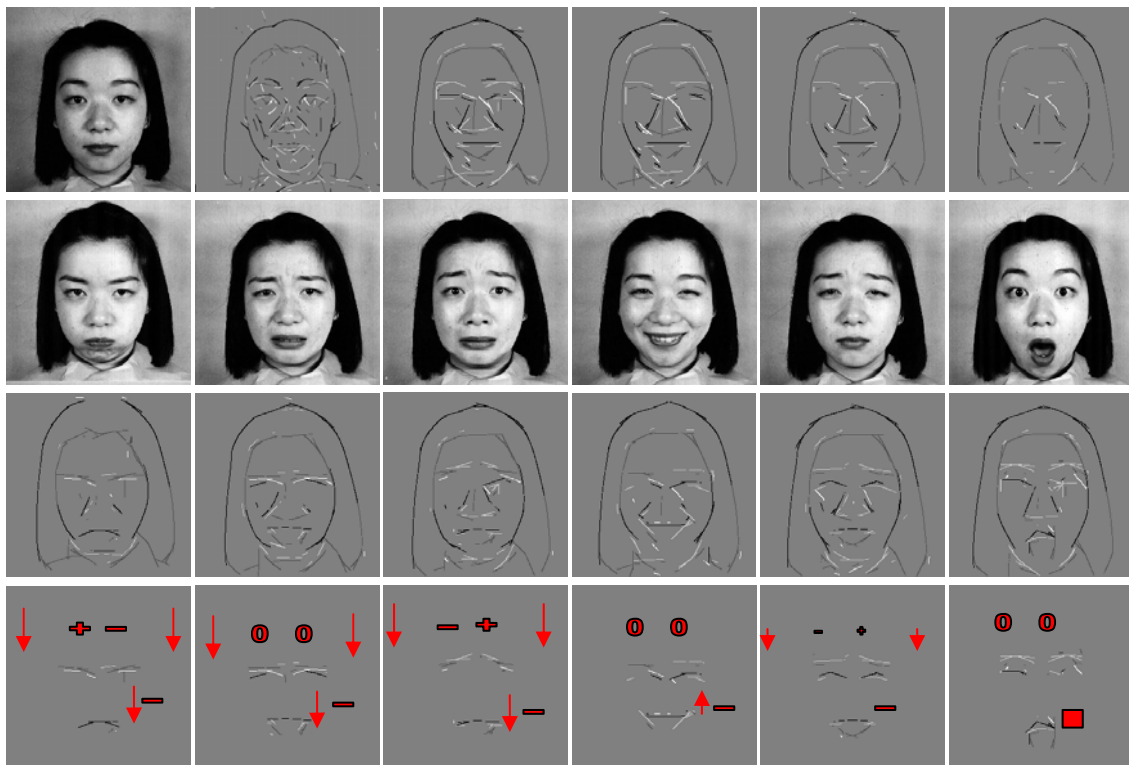


Fig. 2 – First row: neutral face and lines and edges detected at fine (left) and coarse (right) scales. Second row: anger, disgust, fear, happy, sad and surprise. Third row: corresponding lines and edges at a medium scale. Fourth row: displacements detected at mouth and eyebrows (+ and – refer to inclination angle; 0 is the same as neutral; □ is open mouth). Images are from the Japanese Female Facial Expression (JAFPE) database (<http://www.kasrl.org/jaffe.html>).