

ESTIMATION OF FACIAL ACTION UNITS

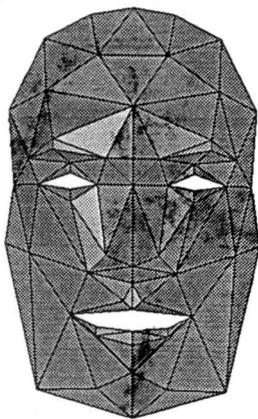
Jörn Ostermann
Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Universität Hannover, Appelstraße 9A, 3000 Hannover 1, F.R.Germany
Phone: 49-511-7625308

Abstract

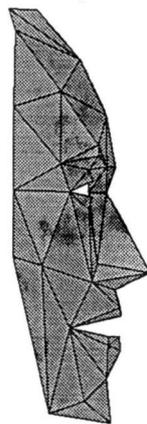
For coding of head and shoulder image sequences model-based facial image coding is investigated by different image coding groups in Japan, Europe and USA. The favorite face model is CANDIDE [1]. The specific knowledge about facial shape and facial expressions is represented in this model. The model is moved by six motion parameters affecting the model as a whole and action units AU, which describe non-facial expressions.

Experiment:

The facial mask CANDIDE [Fig. 1] is manually adjusted to the face in the first frame of the test sequence "Miss America". Using this mask a complete 3D-model MISSA representing head and shoulder is manually generated. The model object is divided into the flexibly connected rigid components shoulder, head and face. Using MISSA the following 49 images of the test sequence are analyzed automatically. The frame-to-frame motion is estimated together with the following action units: AU 10 (upper lip raiser), AU13 (sharp lip puller), AU 20 (lip stretch), AU 27 (mouth stretch), AU 46 R (right eye closes briefly) and AU 46 L (left eye closes briefly). No verification of estimated AU parameters is done.



Mask, front view



Mask, side view



Mask 'Miss America'

Fig. 1: Perspective view of the 3D facial mask

Results:

Mouth mimic: 41 images are described correctly, for three images wrong parameters were estimated, in five images mimic could not be described by implemented AUs.

Eye blink: In 42 images the right eye is modelled correctly, for the left eye only 39 estimations were correct.

Comments:

Estimation results will be improved by verification of parameters.

To the eyes one common AU-parameter should be applied.

Automatic positioning of the facial mask must be very precise (1–2 pel deviation for eyes and mouth from the correct position) and is still a major challenge[2].

References:

- [1] R. Forchheimer, T. Kronander: Image Coding – From Waveforms to Animation, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 37, No. 12, Dec. 1989.
- [2] Matthias Buck, "Construction of a 3D face model and its adaptation to a video scene", COST 211-ter, SIM 91/61.