Facial Action Unit Recognition using Temporal Templates and Particle Filtering with Factorized Likelihoods

Michel Valstar Maja Pantic Ioannis Patras

Delft University of Technology
Electrical Engineering, Mathematics and Computer Science
Man-Machine Interaction Group, Delft, the Netherlands

{M.F.Valstar,M.Pantic,I.Patras}@ewi.tudelft.nl

Abstract

Automatic recognition of human facial expressions is a challenging problem with many applications in human-computer interaction. Most of the existing facial expression analyzers succeed only in recognizing a few basic emotions, such as anger or happiness. In contrast, the system we wish to demonstrate recognizes a large range of facial behavior by recognizing facial action units (AUs, i.e. atomic facial signals). Our system performs AU recognition using temporal templates as input data and a combined kNN-rulebase two-stage classifier. Besides demonstrating the facial action recognizer, we will demonstrate a new point-tracking algorithm based on particle filtering with factorized likelihoods that we use for the registration of the input data.

1. Introduction

Interaction between man and machine is presently far from perfect. It is even perceived by a broad user audience as the bottleneck for the utilization of the available information flow. Of all modalities available to humans, the face carries the most information. Thus, being able to read the face as humans do would greatly benefit a machine to interpret the messages send by a human.

The system we wish to demonstrate is able to recognize 15 AUs occurring alone or in combination in an input face video displaying a facial expression [1]. The proposed method performs AU recognition using temporal templates as input data. All face images are first registered using point tracking based on particle filtering with factorized likelihoods [2]. It classifies a new face video by

first finding *k* nearest neighbors and subsequently compensating for common confusions by applying a rulebase on the results of the kNN classifier.

2. Temporal Representation

Temporal templates are images constructed from image sequences, showing areas where motion occurred in white and areas that remained static in black. This way, a 2 dimensional motion representation is extracted from 3 dimensional input data.

3. Two Stage Learning Machine

The classifier used consists of a kNN algorithm for the initial classification followed by a set of rules to overcome some common confusions made by the kNN classifier. When two AUs produce activity in globally the same part of the MHI, we still can make a distinction by comparing the motion activation in specified facial regions, defined by an expert in facial expression recognition.

4. Demonstration

The first part of the demonstration shows how the tracking algorithm that is used for image registration works. A test subject shows a facial expression in front of a webcam connected to a PC. In the first frame of the recorded face video 9 facial points have to be selected. The demonstrator then shows how the likelihood of the particles is distributed and the resulting tracking of the points.

The second demonstration shows the registration of the image sequence and subsequently the creation of the temporal template. Next, the input temporal template is classified by the combined kNN rulebased classifier (i.e. activated AUs are recognized).

For our demonstration we would like to use a bigger display device than our laptop display (plasma screen or a screen/projector for instance).

References

- [1] M.F. Valstar, I. Patras and M. Pantic, "Recognizing Facial Action Units using Temporal Templates", *proceedings IEEE Intl. workshop on Robot-Human Interaction,*, in print.
- [2] I. Patras and M. Pantic, "Particle Filtering with Factorized Likelihoods for Tracking Facial Features", *proceedings Automatic Face and Gesture Recognition 2004*, pp 97-102