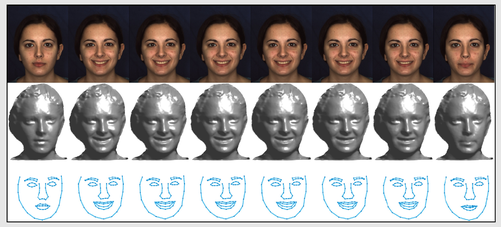
Unsupervised features for facial expression intensity estimation over time

Maren Awiszus · Stella Graßhof · Felix Kuhnke · Jörn Ostermann
Gottfried Wilhelm Leibniz Universität Hannover

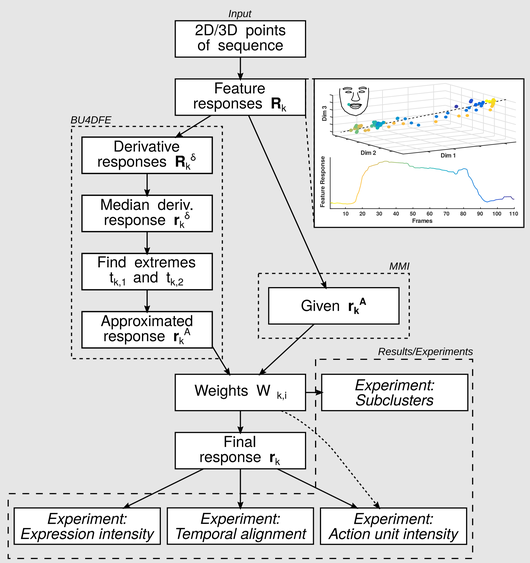
Introduction

- Goal: Automatically analyze facial expressions over time (e.g. for statistical models)
- Problems: Different persons, speed, performance
- Solution: Temporal alignment through expression intensity



Example of a sequence from BU4DFE [1]

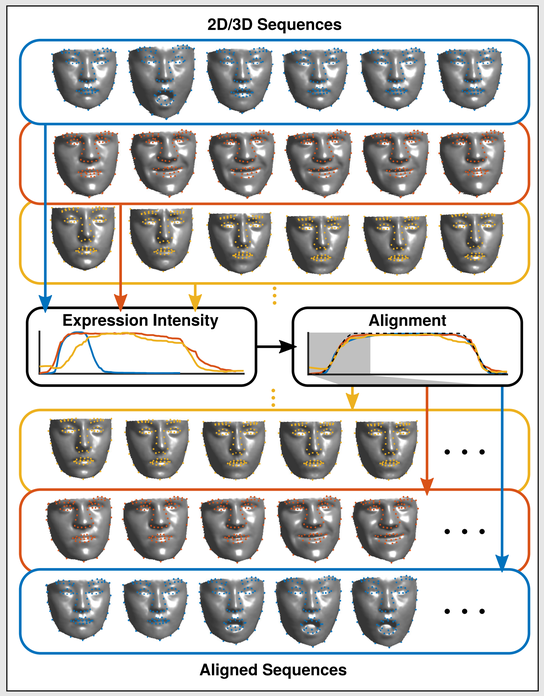
Method



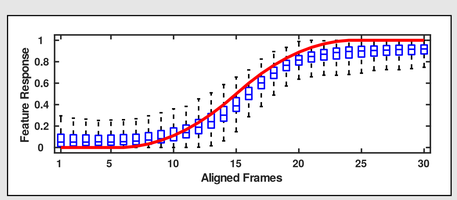
Results

Temporal alignment

Having calculated the facial expression intensities for a given dataset, we can now temporally align the sequences.



Our method is robust against difficult sequences in the dataset. In contrast to a simple, global PCA, it can deal with noisy and erroneous points automatically. This allows us to find and align at least one expression transition in all 606 sequnces of the BU4DFE [1] dataset, for which a boxplot is shown below.



Expression intensity estimation

We can compare our intensity estimation to state-of-the-art method OSVR [3]:

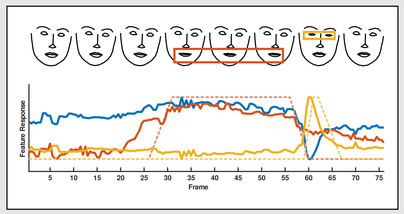
	MAE	PCC	ICC
OSVR-L1 [3]	2.5736	0.5393	0.4808
OSVR L2 [3]	2 2424	0.5453	0.5025
Ours	**1.354**	**0.9354**	**0.9013**
Ours (606 seq.)	1.476	0.9125	0.8765

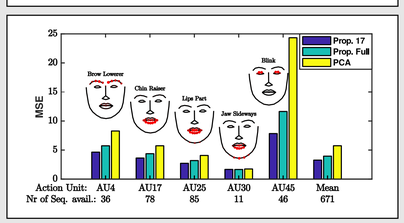
This table shows the comparison of expression intensity estimation results on BU4DFE [1]. Our method outperforms OSVR in all metrics introduced by [3].

Facial action unit intensity estimation

Experiment on the MMI [2] dataset:

- Given approximated response through onset, apex, offset (AU17 and AU45 in orange and yellow, dotted lines)
- Calculated intensities (solid lines) in contrast to simple, global PCA (blue)
- Our method is able to disentangle AUs



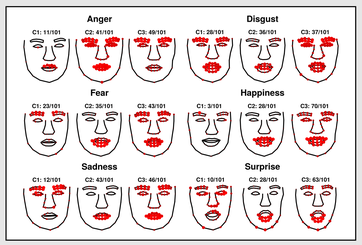


The above barplot illustrates the MSEs between the different feature responses and their corresponding approximated responses for specific AUs. The face shapes above the bars visualize the mean weights for the AUs.

Summary and Conclusions

We propose a method to generate one-dimensional time-varying feature responses from facial motion sequences based on feature points which describe the intensity of an expression. Our experiments include:

- Facial expression intensity estimation on BU4DFE [1], outperforming a state-of-the-art method [3]
- Temporal alignment of all BU4DFE [1] sequences
- Intensity estimation for specific AUs in the MMI database [2], with which we prove our feature can find the intensity of AUs impacting small areas
- Using our generated weights to cluster within the six prototypic emotions, revealing how the motion of the facial feature points used to perform one emotion differ between persons



Find our resulting alignments at:

www.tnt.uni-hannover.de/project/facialanimation/bu4dfe-alignment

Literature

[1] L. Yin, X. Chen, Y. Sun, T. Worm, and M. Reale. *A high-resolution 3d dynamic facial expression database*. FG 2008.

[2] M. Pantic, M. V., R. Rademaker, and L. Maat. *Web-based database for facial expression analysis*. ICME 2005.

[3] R. Zhao, Q. Gan, S. Wang, and Q. Ji. *Facial Expression Intensity Estimation Using Ordinal Information*. CVPR 2016.

Institut für Informationsverarbeitung
Maren Awiszus
awiszus@tnt.uni-hannover.de
http://www.tnt.uni-hannover.de/~awiszus

3D HUMANS 2018: 1st International Workshop on HUman pose, Motion, Activities aNd Shape in 3D
in conjunction with CVPR 2018, Salt Lake City

Leibniz Universität Hannover