# Efficient, Pose Invariant Facial Emotion Classification using 3D Constrained Local Model and 2D Shape Information


東京大学 THE UNIVERSITY OF TOKYO

**Laszlo A. JENI**
University of Tokyo
laszlo.jeni@ieee.org

**Hideki HASHIMOTO**
Chuo University
hashimoto@elect.chuo-u.ac.jp

**Andras LORINCZ**
Eotvos Lorand University
andras.lorincz@elte.hu

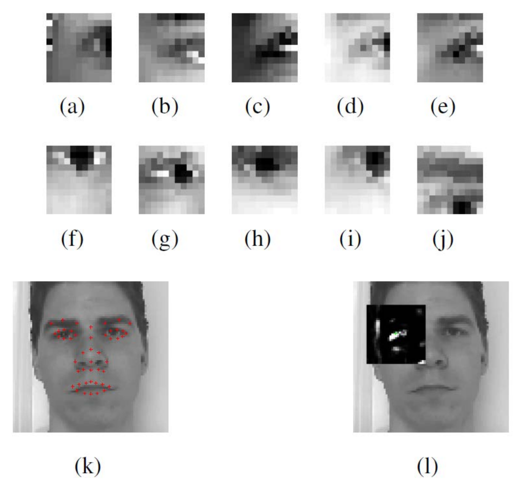

## 1. Introduction

Pose invariant facial emotion classification is important for situation analysis and for automated video annotation. We started from the raw 2D shape data of the CK+ database and used a simple Procrustes transformation and the multi-class SVM leave-one-out method for classification. We found close to 100% performance demonstrating the potentials of shape based methods. We applied a 3D constrained local model (CLM) and generated a 3D emotionally modulated database with different poses using FaceGen. We fitted 3D CLM and used it in an iterative manner to exclude the potentially occluded landmarks. We transformed the 3D shape to frontal pose and evaluated the outputs of our classifier. Excellent pose invariant performance with considerable improvement over the non-iterative method was achieved.

## 2. Methods

In this work we use a ***3D Constrained Local Model*** method, where the shape model is defined by a 3D mesh and in particular the 3D vertex locations of the mesh, called landmark points.

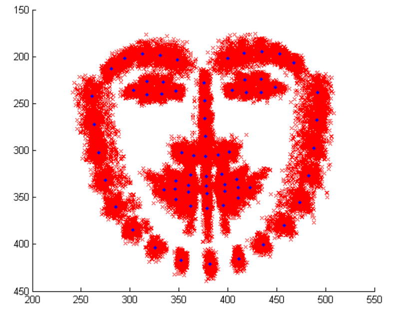
CLM is constrained through the PCA of ***Point Distribution Model***. It works with ***local experts***, whose opinion is considered independent and are multiplied to each other.




The CLM fitting can be done even if we restrict the process to a limited set of landmarks.

### Extracted Features

In 2D, we used the 2D shape coordinates of the CK+ database: we had 2*68=136 dimensional vectors for classification.



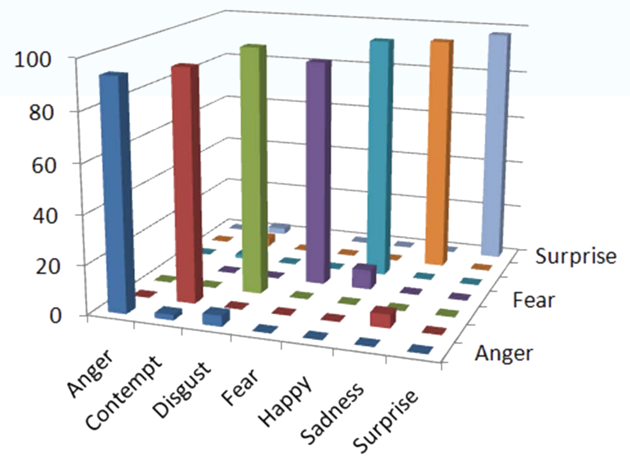
We used the ***Procrustes*** method to compute the mean shape and normalize all shapes to this mean .

For the classification task, we used multi-class SVM and the so called ***AU0 normalization***; we computed the differences between the features of the actual shape and the features of the first neutral) frame.

## 3. Results on the Cohn-Kanade+ Dataset

### Experiment on the CK+ dataset with Procrustes method and original landmarks

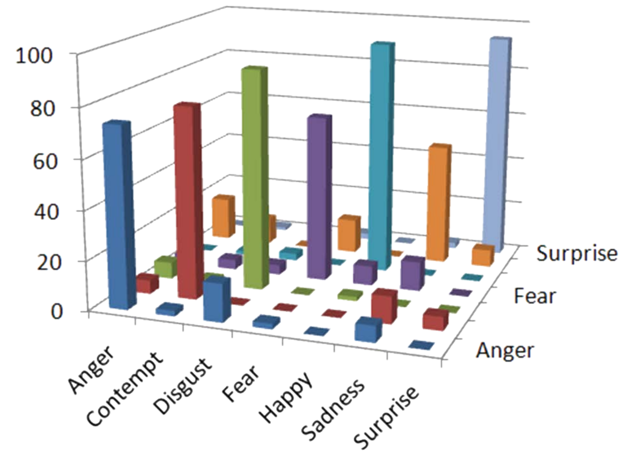
We used the Cohn-Kanade Extended Facial Expression Database (CK+) with the original 68 CK+ landmarks, calculated the mean shape and normalized all shapes by minimizing the Procrustes distance to it. We trained a multi-class SVM using the leave-one-subject-out cross validation method. Results show that shape is an excellent gauge if information is precise in 2D.



| | An. | Co. | Di. | Fe. | Ha. | Sa. | Su. |
|---|---|---|---|---|---|---|---|
| **Anger** | **93.3** | 2.2 | 4.4 | 0 | 0 | 0 | 0 |
| **Contempt** | 0 | **94.4** | 0 | 0 | 0 | 5.6 | 0 |
| **Disgust** | 0 | 0 | **100** | 0 | 0 | 0 | 0 |
| **Fear** | 0 | 0 | 0 | **92** | 8 | 0 | 0 |
| **Happy** | 0 | 1. 5 | 0 | 0 | **98.5** | 0 | 0 |
| **Sadness** | 0 | 3.6 | 0 | 0 | 0 | **96.4** | 0 |
| **Surprise** | 0 | 2.6 | 0 | 0 | 0 | 0 | **97.4** |

### Experiment on the CLM-tracked CK+ dataset

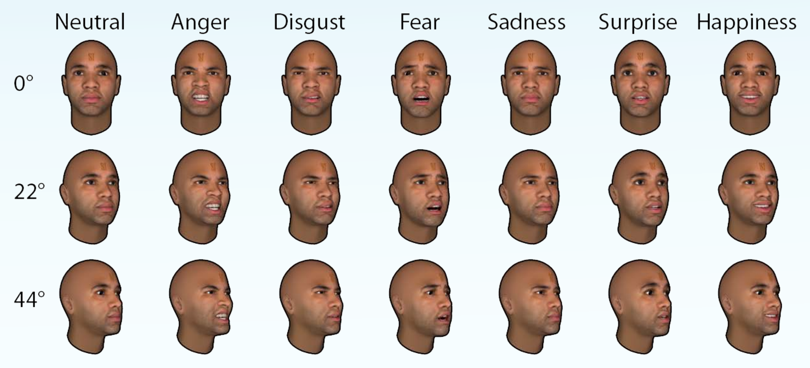
In this experiment we tracked the CK+ facial expressions with the CLM tracker and annotated all image sequences starting from the neutral expression to the peak of the emotion. 3D CLM estimates the rigid and non-rigid transformations. We removed the rigid ones from the faces and projected the frontal view to 2D.



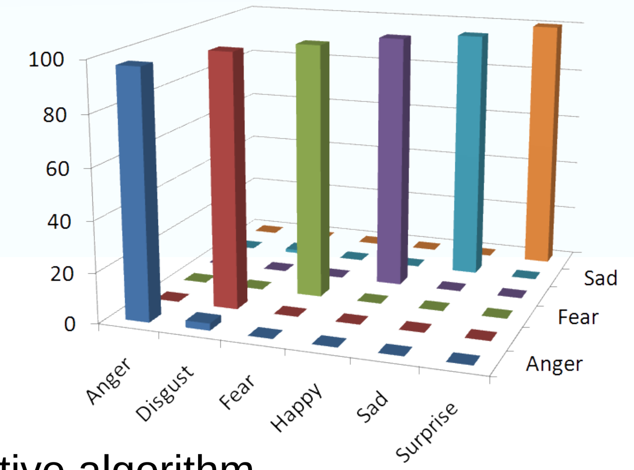
| | An. | Co. | Di. | Fe. | Ha. | Sa. | Su. |
|---|---|---|---|---|---|---|---|
| **Anger** | **73.3** | 2.2 | 15.6 | 2.2 | 0 | 6.7 | 0 |
| **Contempt** | 5.6 | **77.8** | 0 | 0 | 0 | 11.1 | 5.6 |
| **Disgust** | 6.8 | 1.7 | **89.8** | 0 | 1.7 | 0 | 0 |
| **Fear** | 4 | 4 | 4 | **68** | 8 | 12 | 0 |
| **Happy** | 0 | 1.5 | 2.9 | 0 | **95.7** | 0 | 0 |
| **Sadness** | 17.9 | 10.7 | 0 | 14.3 | 0 | **50** | 7.14 |
| **Surprise** | 0 | 1.2 | 0 | 2.4 | 0 | 2.4 | **94.0** |

## 4. FaceGen Generated Dataset

We used FaceGen to generate faces with different degree of yaw rotation an with six putative emotions built into the software (anger, disgust, fear, happiness, sadness, and surprise). We generated frontal image sequences displaying transitions from the neutral expression to the six emotional expression.




We trained a multi-class SVM on the FaceGen test dataset using the leave-one subject-out cross validation method: every emotion gave rise to 100% classification performance in the frontal case.

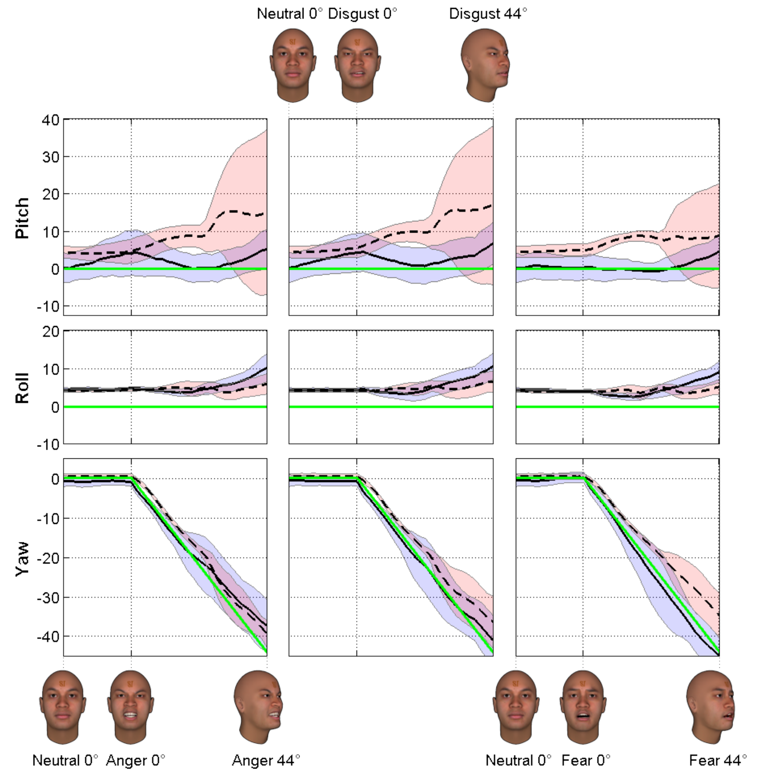


**Algorithmic design.** We designed an iterative algorithm.

Step 1. Pose estimation using stable landmarks.
Step 2. Excluding occluded landmarks as determined by the previous step.
Step 3. Filling in missing information using the PCA method.
Step 4. Apply SVM trained on FaceGen emotions.

## 4. Pose Estimation Experiment

Pitch/Yaw/Roll estimation of the head using every landmark (red region, dashed line), and using only the selected stable landmarks (blue region, solid line).



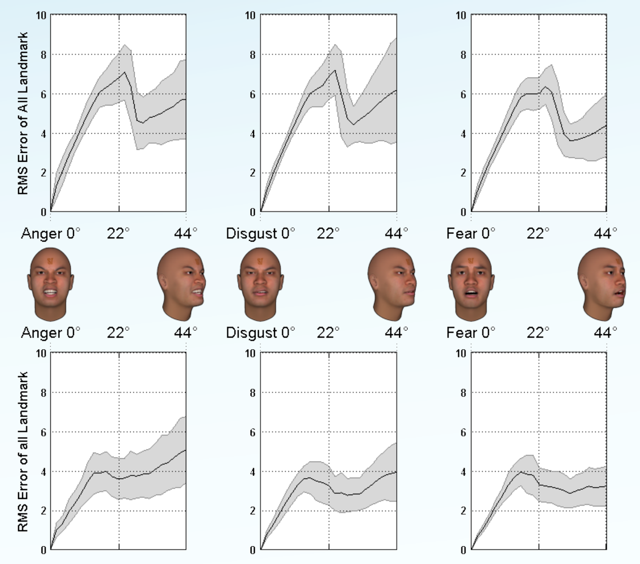

Bold lines denote the mean, the region shows the standard deviation.

Only yaw angle changes, but pitch estimation becomes erroneous (red region) unless stable landmarks are used exclusively (blue region).

Solid green line: ground truth according to FaceGen.

## 5. CLM Fitting Experiment

RMSE of the reconstructed landmark positions in pixels during the movement using every landmark (upper subfigures), and using the modified algorithm (bottom subfigures).

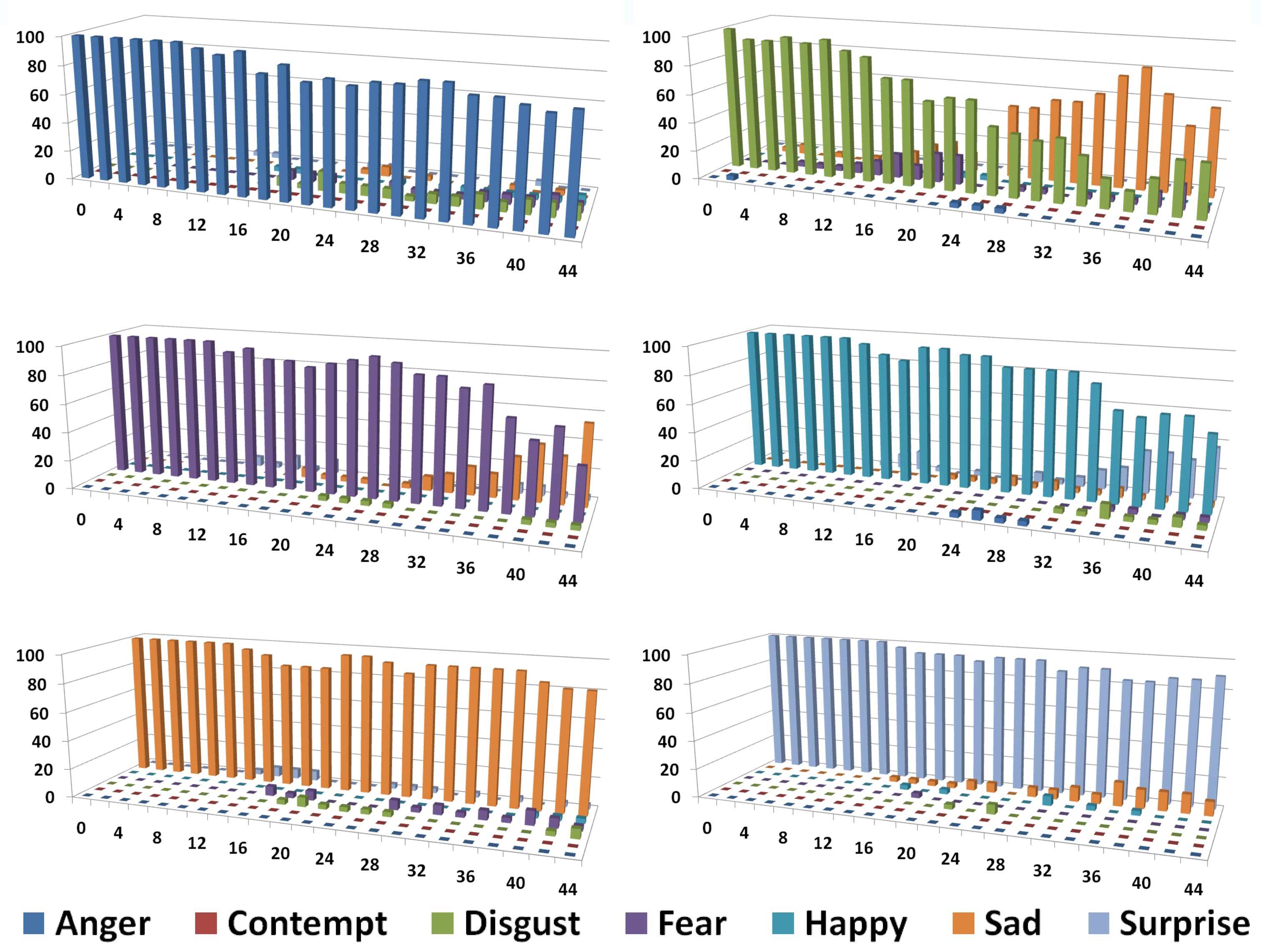


The bold lines denote the mean, the region shows the standard deviation.

The distortion was compared to the initial frame (0 degrees of yaw rotation). 1 pixel error for all landmarks corresponds to 1 RMSE unit.

## 6. Pose Invariant Emotion Classification

Recognition of different FaceGen emotions as a function of yaw angle.





## Acknowledgement

We are grateful to Jason Saragih for providing his CLM code for our studies. The Project is supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1/B-09/1/KMR-2010-0003).

Nemzeti Fejlesztési Ügynökség
www.ujszechenyiterv.gov.hu
06 40 638 638

MAGYARORSZÁG MEGÚJUL

The Project is supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003).