

# AU Recognition on 3D Faces Based On An Extended Statistical Facial Feature Model

Xi Zhao, Emmanuel Dellandréa, Liming Chen and Dimitris Samaras

**Abstract**—Recognition of facial action units (AU) is one of two main streams in the facial expressions analysis. Action units deform facial appearance simultaneously in landmark locations and local texture as well as geometry on 3D faces. Thus, it is necessary to extract features from multiple facial modalities to characterize these deformations comprehensively. In order to fuse the contribution of the discriminative power from all features efficiently, we propose to use our extended statistical facial feature models (SFAM) to generate feature instances corresponding to AU class for each feature. Then the similarity between each feature on a face and its instances are evaluated so that a set of similarity scores are obtained. All sets of scores on the face are then weighted for AU recognition. Experiments on the Bosphorus database show its state-of-the-art performance.

## I. INTRODUCTION

Action Units (AUs) in the Facial Action Coding System (FACS) [5] objectively describe the facial deformations in terms of visually observable muscle actions. They are considered as signals to interpret human faces and thus provide a possibility to infer expressions through a high-level decision making process, such as Emotional FACS rules and FACSaid [5]. Moreover, the recognition of AU is often preferred by researchers in affect understanding and psychology.

Most of the existing AU recognizers extract 2D spatio-temporal facial features either in geometric modality such as the shapes of the facial components (eyes, mouth, etc.) [7] and landmark location (corners of the eyes, mouth, etc.) [1] or in texture modality such as Gabor wavelet [18], [22], Haar feature [23]. It is interesting to note that features in two modalities are fused in [8], [19] for recognizing a various of AUs and measuring their intensity.

AUs originate from the contraction of facial muscles and cause the deformation on facial landmark locations, texture and facial surface. Since 3D faces contain ample information on both facial texture and geometry, they provide advantage in recognizing AUs. There exist several works on analyzing 3D facial expressions, which can be globally categorized into feature-based approaches [17], [14], [16], [21], [6], [15] and model-based approaches [10], [9], [11], [20]. However, most of them aim to identify the six universal expressions. AUs are identified only in [15] on dynamic 3D face data displaying expressions. They first use AAM to track feature points

on texture maps and extract 3D motion vectors for feature representation. HMM is learnt and used for classifying seven AUs and one AU combination.

Most of feature based approaches in 3D use line properties between landmarks, such as angles and distances [17], [14], [16], [6], [15] and principal curvatures of vertices [21]. These approaches exclude information on other modalities and thus fail to represent the comprehensive characteristics of facial deformation including both geometry and texture. Meanwhile, although experiments have proved their good performance in recognizing the universal expressions, the geometry based features adopted in them have not yet been shown to demonstrate enough discrimination power for distinguishing subtle facial AUs.

On the other hand, morphable facial models are built by learning the deformation modes on texture and geometry modality and uses the deformation parameters as features for recognition [10], [9], [11], [20]. Two major problems exist: firstly, the deformation modes learnt from whole faces describe the major variations globally and thus can not properly include local deformation patterns caused by AUs. Secondly, the learnt variation modes are not necessarily consistent with the variations among AUs and expressions, thus may not synthesize the expression accurately. In other words, AUs or expressions can only be approximated by a set of variation modes, instead of being modeled by one specific mode in the models. For certain expressions such as happiness or surprise, or action units such as AU27 (mouth opening), the deformation is prominent and thus can be approximated well enough to be distinguishable. However, for some of other moderate expressions and AUs, small variances in parameters yield different expressions.

In order to characterize the facial deformations comprehensively, features from all three modalities of facial landmark location, texture and geometry should be included, which raises the problem on how to fuse the contribution from each feature efficiently. We propose to extend our previous proposed statistical facial feature model (SFAM) for score computing and weighting for all kinds of features without presetting any thresholds or parameters. SFAM is a partial morphable facial model which learns both global variations in 3D landmark configuration and local ones around each landmark in terms of texture and geometry. SFAM is able to approximate partial regions on a new face by estimating parameters and then generating instances. Since most of the AUs influence the facial appearance locally, it is advantageous to use local based SFAM instead of deformable face models built on whole faces. In this work,

Xi Zhao, Emmanuel Dellandréa and Liming Chen are with Université de Lyon, CNRS, Ecole Centrale Lyon, LIRIS, UMR5205, 69134, France, {xi.zhao, emmanuel.dellandrea, liming.chen}@ec-lyon.fr.

Dimitris Samaras is with the Department of Computer Science, Stony Brook University, 2429 Computer Science, Stony Brook, NY 11794-4400. samaras@cs.sunysb.edu.

TABLE I

DESCRIPTION OF ACTION UNITS AS DEFINED IN FACS. THE FIRST COLUMN LISTS THE AUS DETECTED IN OUR EXPERIMENTS, THE SECOND GIVES A DESCRIPTION OF THE AUS.

Action Unit	Description	Action Unit	Description
2	Outer Brow Raiser	4	Brow Lowerer
7	Lid Tightener	9	Nose Wrinkler
10	Upper Lip Raiser	12	Lip Corner Puller
14	Dimpler	17	Chin Raiser
18	Lip Puckerer	22	Lip Funneler
24	Lip Presser	26	Jaw Drop
27	Mouth Stretch	28	Lip Suck
34	Puff	43	Eyes Closure

for each type of local feature extracted based on SFAM, we learn a set of statistical feature models (SFM) from faces displaying corresponding AUs. Then we use these SFMs to approximate the feature from new faces and generate a set of corresponding instances. The similarity scores between the feature and its instances are evaluated. Then the sets of similarity scores from all features are weighted for final AU identification. 16 AUs are recognized in the experiments selected based on data availability. They are described in table I and demonstrated in fig 3. The tests achieve 94.2% positive rate on recognizing 7 AUs and 85.6% positive rate on recognizing all 16 AUs.

The rest of paper is organized as follows: In the section II we present how to build the extended SFAM, how to apply Extended SFAM on AU recognition as well as the feature extraction. The experimental results are presented in III. We draw our conclusion in IV.

## II. METHODOLOGY

The Statistical Facial Feature Model (SFAM) is a partial 3D face morphable model which contains both global variations in landmark configuration (morphology) and local ones in terms of texture and shape around each landmark [24]. In this section, we describe the extended SFAM and its learning, its application to AU recognition as well as the extraction of features.

### A. Extended Statistical Facial Feature Model

In order to learn variations on global morphology, local texture and local shape among training faces, 3D face models are first preprocessed to remove noises, e.g. spikes, holes, and to exclude such variations introduced by head pose and face scale. Local grids with 1mm resolution are used to remesh regions around landmarks and extract intensity and range data. The remesh process ensures that the same number of vertices is sampled from all faces and these vertices are aligned among them.

SFAM is learnt by applying PCA to three kinds of training features while preserving 95% of variations for each type of features. These three feature groups are respectively landmark configuration which defines a face morphology, local texture and local range around each landmark.

$$m = \bar{m} + P_m b_m, g = \bar{g} + P_g b_g, r = \bar{r} + P_r b_r \quad (1)$$



Fig. 1. First variation modes on the landmark configuration, local texture and local shape in the morphable SFAM.

where  $\bar{m}$ ,  $\bar{g}$ ,  $\bar{r}$  are respectively the mean morphology, mean intensity and mean range value while  $P_m$ ,  $P_g$ ,  $P_r$  are their learnt variation modes respectively,  $b_m$ ,  $b_g$ ,  $b_r$  are the corresponding sets of control parameters.

Partial face instances, consisting of local face regions with texture and shape configured by the landmark configuration, can be estimated and synthesized by a linear combination of these components for a given face. The model building is similar as that in [24], however in this paper we use 19 landmarks, as shown in fig. 1, so that the SFAM can be better applied to the problem of facial expression recognition.

The SFAM is further extended in this work by considering statistical feature models (SFM) for additional features which are particularly relevant to subtle facial expression analysis, including for instance local binary patterns (LBP) and shape index as explained later in the section II-C, in that they help to characterize local texture and shape changes which occur in subtle facial expressions. Each feature set  $F_l$  is divided into  $N_e$  (number of AUs) subsets. For each subset  $i_x$ , PCA is applied to learn the major modes of variations (95%). Thus, given a feature set, we therefore have  $N_e$  SFMs associated with  $N_e$  action units under study.

$$F_l^{i_x} = \bar{F}_l^{i_x} + P_l^{i_x} b_l^{i_x} \quad (2)$$

Where  $\bar{F}_l^{i_x}$  is the mean of feature vectors in the  $i_x$  subset,  $P_l^{i_x}$  is the learnt matrix of major variation modes, and  $b_l^{i_x}$  is parameters, which are supposed to follow corresponding Gaussian distributions with std  $\sigma_{l_j}^{i_x}$ .

### B. AU Recognition using Matching Scores

Given an AU, the matching scores of a feature set extracted from an input face model is computed through an analysis by synthesize approach. Feature instances  $\hat{F}_{l\kappa}^{i_x}$  on an input face  $\kappa$  can be generated from the eqn. 2 using the feature set  $F_{l\kappa}$  to estimate the best parameter  $b_l^{i_x}$  similar as in [3]. We set a boundary ( $\pm 0.5\sigma_{l_j}^{i_x}$ ) for the corresponding parameter to increase the distinction of instances. The matching score  $Q_{l\kappa}^{i_x}$  between the feature set  $F_l$  extracted from the input face  $\kappa$  and its instance of  $i_x$ th AU is computed as the correlation response between feature sets  $F_{l\kappa}$  and  $\hat{F}_{l\kappa}^{i_x}$ , defined in eqn. 3.

$$Q(F_{l\kappa}^{i_x}) = \left\langle \frac{F_{l\kappa}}{\|F_{l\kappa}\|}, \frac{\hat{F}_{l\kappa}^{i_x}}{\|\hat{F}_{l\kappa}^{i_x}\|} \right\rangle \quad (3)$$

The final matching score for the  $i_x$ th AU on  $\kappa$  is then a weighted sum of the matching scores produced by all feature sets.

$$Sc_{\kappa}^{i_x} = \sum_{l=1}^{N_c} A_l Q_l \kappa^{i_x} \quad (4)$$

where  $A_l$  is a set of weight (they are all set to 1 in this work). The highest matching score achieved by an AU  $i_x$  on the input face model  $\kappa$  indicate that the face model is best matched by the instances from the AU  $i_x$ . This AU  $i_x$  is then recognized as the AU produced by the input face model  $\kappa$ .

### C. Feature extraction

Local features are extracted around each of the 19 landmarks which are manually labeled or automatically located thanks to the SFAM [24]. We first concatenate the 3D coordinates of these landmarks into a single vector  $M$  which thus describe the configuration of the landmarks (face morphology). We then interpolate the locations of the eye centers using the eye corners and replace the two nose saddles by them in the  $M$  for local feature extraction. Two sets of remeshed local grids are thus created centered at landmarks in updated  $M$  and local texture features  $T$  and shape features  $R$  are extracted by concatenating all the intensity and range values on those grids. Local patches in Fig. 2 illustrate the remeshed grids formed by the local shapes and rendered using the corresponding local textures.

We also compute additional features to better capture subtle facial expressions in terms of changes in morphology, local texture and local shape. Concerning the characterization of morphology changes, we compute a landmark displacement vector  $D$  from the landmark vector  $M$ . We further form a vector  $L$  by 11 distances between facial expression sensitive landmarks as illustrated by the green lines in fig 2.

The displacement of a landmark means to capture the change of the landmark location when an expression appears

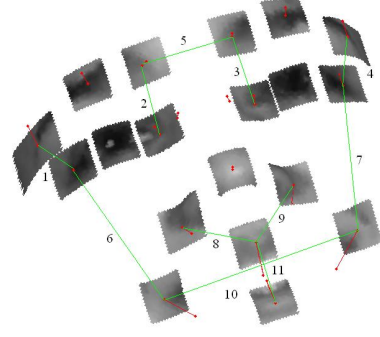


Fig. 2. Feature extraction

on a neutral face. It is informative because it represents the difference between the face with an expression and the neutral one. However, this measurement imposes a constraint that a neutral face model from each subject must be available in the gallery and is therefore subject biased. In our work, we lift this constraint and make use of a mean neutral face model in averaging the landmarks from all training neutral faces. Thus, we compute  $D$  by subtracting the mean of  $M_n$  ( $\bar{M}_n$ ) from  $M$ , shown as red lines in fig. 2.

$$D = M - \bar{M}_{PI\_neutral} \quad (5)$$

The LBP operator, a powerful texture measure used widely in 2D face analysis, extracts information which is invariant to local gray-scale variations of the image with low computational complexity. Multi-Scale LBP [13] is an improved facial representation as compared to standard LBP. We adopt multiscale LBP features for three reasons: first, LBP describes local property of images, which is consistent with the local deformations that correspond to AUs; second, the variance in the apparent AU magnitude is large, some are quite noticeable while some are subtle, thus it is necessary to analyze them under different scale; third, LBP is efficient and easy to compute.

In our case, LBP are computed and extracted from the local grids on both texture maps and range maps with scale 1 to 5 respectively ( $LBP_{(16,1)}^{U2}t, LBP_{(16,2)}^{U2}t, \dots, LBP_{(16,5)}^{U2}t$  and  $LBP_{(16,1)}^{U2}r, LBP_{(16,2)}^{U2}r, \dots, LBP_{(16,5)}^{U2}r$ ). Superscript  $u2$  indicates that the definition relates to uniform patterns with a U value of at most 2 [2]. The extracted LBP values on local grids using each (P,R) pair parameter are concatenated into a vector respectively. In total 10 LBP features are extracted ( $LBP_t1-5, LBP_r1-5$ ).

To describe local surface properties, we compute the shape index [4] of all points on the local grids and concatenate into vector  $SI$ . We choose shape index because it has been proven to be an efficient feature to describe local curvature information and is independent of the coordinate system. The shape index is computed on each vertex on local grids and feature  $SI$  is constructed by concatenating those values into a vector.

In summary, three sets of features are extracted corresponding to the three facial modalities. In landmark configu-

ration modality, person independent landmark displacement  $D$  and distance between landmarks  $L$  are extracted based on landmark locations. In facial texture modality, local texture  $T$  as well as multi-scale LBP  $LBP_{p,1-5}$  are extracted from local texture maps. In facial geometry modality, local range  $R$ , shape index  $SI$  as well as multi-scale LBP  $LBP_{p,1-5}$  are extracted from local range maps.

### III. EXPERIMENTAL RESULTS

The Bosphorus dataset [12] consists of 105 subjects in various poses, expressions and occlusion conditions. The number of face scans from each subject varies from 31 to 54 and the number of total face scans is 4652. Face scans which relate to facial expression analysis are those displaying the six universal facial expressions and those displaying Action Units. The number of scans displaying AUs from each subject varies from less than 7 to more than 20. The Bosphorus dataset is chosen because it is the only public database which contains 3D face scans displaying separate AUs.

In the first experiments for facial AU recognition, 100 subjects who have displayed 7 AUs are used (AU2, AU4, AU9, AU12, AU27, AU28 and AU34). This test follows a ten-fold cross-validation process, where subjects are independent between training sets and testing sets. In the ten-fold person-independent cross-validation method, subjects are partitioned into two subsets in each round (totally 10 rounds): one with 90% subjects for training and the other with 10% subjects for testing. This experiment setup guarantees that each subject appears once in testing set and 9 times in training set and any subject used for testing does not appear in the training set because the partition is based on the subjects rather than the individual expression.

TABLE II

AVERAGE RECOGNITION RATES FOR 7 ACTION UNITS USING DIFFERENT FEATURE SET

	M	T (2D)	G	M+T	M+G	T+G	M+T+G
RR	74.60%	87.01%	91.92%	88.74%	90.19%	93.36%	94.23%

Table II shows the overall average recognition rates for 7 AUs using the different feature sets. The first three columns show the recognition rates from each modality, (landmark (M), texture (T), geometry (G)) respectively, the following three columns are the results combining features from any two modalities, and the last column shows the result for features from all modalities. We can observe that by adopting more features, our approach is able to enhance the overall recognition rate from 87.01% achieved on 2D facial textures to 94.23% using features from all three facial modalities.

Table III shows the confusion matrix on recognizing 7 AUs. It is notable that the identification of AU27 (Mouth Stretch) achieves 100% recognition rate using our method. AU27, AU4 (Brow Lowerer) and AU28 (Lip Suck) has an average recognition rate over 95%, while the relative lower recognition rates are still around 90% which are of AU2 (Outer Brow Raiser), AU34 (Puff) and AU9 (Nose Wrinkler).

TABLE III

CONFUSION MATRIX OF THE PERSON-INDEPENDENT AU RECOGNITION.

Input \ Output	AU12	AU27	AU28	AU34	AU9	AU2	AU4
AU12	94,95%	0,00%	0,00%	0,00%	0,00%	1,01%	4,04%
AU27	0,00%	100,00%	0,00%	0,00%	0,00%	0,00%	0,00%
AU28	0,00%	0,00%	95,96%	0,00%	0,00%	1,01%	3,03%
AU34	0,00%	0,00%	1,01%	89,90%	0,00%	0,00%	9,09%
AU9	1,01%	0,00%	0,00%	0,00%	89,90%	0,00%	9,09%
AU2	3,03%	0,00%	1,01%	1,01%	0,00%	90,91%	4,04%
AU4	0,00%	0,00%	1,01%	0,00%	1,01%	0,00%	97,98%

AU4 is often confused with other AUs because its appearance is close to neutral faces.

In the second test, we use face scans displaying 16 AUs (table I) from the 60 subjects who have performed all of these AUs. Note that the reduction in subject numbers is because other subjects were not scanned with all of these AUs. Totally 960 face scans are involved in this test. The test also follows a 10-fold cross-validation process.

Table IV shows the average positive rates and average false-alarm rates of all AUs. Recognizing each  $AU_i$  can be considered as a two-class classification of the  $AU_i$  and the non- $AU_i$ . The positive rate is defined as  $PR = \frac{TP}{TP+FN}$  and the false-alarm rate is  $FAR = \frac{FP}{FP+FN}$ . Note that the recognition rate mentioned in the previous test is equal to positive rate here.

TABLE IV

AVERAGE POSITIVE RATES (PR) AND AVERAGE FALSE-ALARM RATES (FAR) OF AUS.

	AU10	AU12	AU14	AU17	AU18	AU22	AU24	AU26
PR	95.0 %	85.0 %	75.0 %	80.0%	91.7%	90.0%	76.7%	91.7%
FAR	10.9%	19.0%	23.7%	7.7%	14.1%	3.6%	40.3%	12.7%
	AU27	AU28	AU34	AU9	AU2	AU43	AU7	AU4
PR	91.7%	81.7%	88.3%	81.7%	90.0%	98.3%	78.3%	75.0%
FAR	3.5%	7.5%	20.9%	5.8%	3.6%	4.8%	13.0%	26.2%

Among 16 AUs, 7 of them (AU10, AU18, AU22, AU26, AU27, AU2, AU43) have average positive rates over 90%, while 4 of them (AU14, AU24, AU7, AU4) have average positive rates below 80%. Meanwhile, AU24 has the highest false alarm rate, which implies that it is easier to be confused with other AUs. AU34 and AU4 also have false alarm rate above 20%. AU43, AU27, AU22 can not be easily confused with other AUs, since they have the FAR below 5%. Our method achieves an overall average positive rate for all 16 AUs of 85.6% with the overall average FAR 13.6%. Compared with the previous test, the overall average positive rate decreases 9% due to the increase of AU classes and the confusion between the former 7 AUs and the added 9 AUs. Specifically, the positive rates of AU4 and AU28 decrease more than 14%. Those of AU12, AU27 and AU9 decrease around 9% while the positive rates of AU34 and AU2 do not have obvious variance.

To our best knowledge, there is no similar work on AU recognition using the Bosphorus dataset. Thus, it is difficult to compare our work directly with existing studies on AU recognition. We provide here the results from two studies on AU recognition for reference. In [15], seven AUs and an

TABLE V  
EXPLANATION OF TP AND FAR DEFINITION.

	$AU_i$	non- $AU_i$
$AU_i$	True Positive (TP)	False Negative (FN)
non- $AU_i$	False Positive (FP)	True Negative (TN)

AU combination are recognized on their own database with a recognition rate of 89.1%. In [18], a Dynamic Bayesian Net is used to learn the relationship between AUs on 2D Cohn-Kanade database in order to enhance the recognition performance using gabor features and ababoost classifier. They achieve an 85.8% positive rates on 14 AUs. Compared with them, our approach achieves recognition rates of 94.2% for 7 AUs and 85.6% for 16 AUs respectively.

#### IV. CONCLUSION

In this paper we propose an AU recognition method based on an extended statistical facial feature model. The SFAM is suitable for extracting local characteristics of facial deformations in facial landmark configurations, local texture and local geometry. In order to combine the contribution from all extracted features, we build statistical feature models and weight the similarity scores computed between feature instance sets and facial features of investigated faces. 15 features are extracted from three facial modalities, including multi-scale LBP, shape index, distances between landmarks and landmark displacement. Experiments on recognizing 7 AUs and 16 AUs have achieved 94.2% and 85.6% recognition rates respectively. Although this work is among the first attempts to recognize AUs on 3D faces, it achieves a consistent result with the highly optimized 2D method [18]. In the future, we will build a probabilistic latent semantic space of AUs and recognize spontaneous expressions based on this space.

#### REFERENCES

[1] T. Brick, M. Hunter, and J. Cohn. Get the facts fast: Automated face analysis benefits from the addition of velocity. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–7, 10–12 2009.

[2] C. Chan, J. Kittler, and K. Messer. Multi-scale local binary pattern histograms for face recognition. *Proceedings of International Conference on Advances in Biometrics*, pages 809–818, 2007.

[3] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. *ECCV*, 2, 1998.

[4] C. Dorai and A. Jain. Cosmos - a representation scheme for 3d free-form objects. *IEEE Trans. PAMI*, 19(10):1115–1130, 1997.

[5] P. Ekman, W. V. Friesen, and J. C. Haper. Facial action coding system, the manual on cd rom. 2002.



Fig. 3. Examples of Facial AUs.

[6] Y. Hu, Z. Zeng, L. Yin, X. Wei, X. Zhou, and T. Huang. Multi-view facial expression recognition. *Face & Gesture*, pages 1–6, 2008.

[7] S. Koelstra and M. Pantic. Non-rigid registration using free-form deformations for recognition of facial actions and their temporal dynamics. In *Automatic Face Gesture Recognition, 2008. FG '08. 8th IEEE International Conference on*, pages 1–8, 17–19 2008.

[8] M. Mahoor, S. Cadavid, D. Messinger, and J. Cohn. A framework for automated measurement of the intensity of non-posed facial action units. *Computer Vision and Pattern Recognition Workshop, 0:74–80*, 2009.

[9] I. Mpiperis, S. Malassiotis, and M. Srinivas. Bilinear models for 3-d face and facial expression recognition. *IEEE Trans. Info. Fore. and Secur.*, 3(3):498–511, 2008.

[10] S. Ramanathan, A. Kassim, Y. Venkatesh, and W. S. Wah. Human facial expression recognition using a 3d morphable model. *ICIP*, pages 661–664, 2006.

[11] M. Rosato, X. Chen, and L. Yin. Automatic registration

- of vertex correspondences for 3d facial expression analysis. *BTAS*, pages 1–7, 2008.
- [12] A. Savran, N. Alyuz, H. Dibeklioglu, O. Celiktutan, B. Gokberk, B. Sankur, and L. Akarun. Bosphorus database for 3d face analysis. *The First COST 2101 Workshop on Biometrics and Identity Management*, 2008.
- [13] C. Shan and T. Gritti. Learning discriminative lbp-histogram bins for facial expression recognition. *BMVC*, 2008.
- [14] H. Soyel and H. Demirel. 3d facial expression recognition with geometrically localized facial features. *Sympo. on Comp. and Info. Sci.*, pages 1–4, 2008.
- [15] Y. Sun, M. Reale, and L. Yin. Recognizing partial facial action units based on 3d dynamic range data for facial expression recognition. In *Automatic Face Gesture Recognition, 2008. FG '08. 8th IEEE International Conference on*, pages 1 –8, 17-19 2008.
- [16] H. Tang and T. Huang. 3d facial expression recognition based on properties of line segments connecting facial feature points. *Face & Gesture*, pages 1–6, 2008.
- [17] H. Tang and T. S. Huang. 3d facial expression recognition based on automatically selected features. *CVPR workshop*, pages 1–8, 2008.
- [18] Y. Tong, J. Chen, and Q. Ji. A unified probabilistic framework for spontaneous facial action modeling and understanding. *IEEE Transaction of Pattern Analysis and Machine Intelligence*, 32(2):258–273, 2010.
- [19] F. Tsalakanidou and S. Malassiotis. Robust facial action recognition from real-time 3d streams. In *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*, pages 4 –11, 20-25 2009.
- [20] Y. V. Venkatesh, A. A. Kassim, and O. V. R. Murthy. A novel approach to classification of facial expressions from 3d-mesh datasets using modified pca. *Patt. Recogn. Lett.*, 30(12):1128–1137, 2009.
- [21] J. Wang, L. Yin, X. Wei, and Y. Sun. 3d facial expression recognition based on primitive surface feature distribution. *CVPR*, pages 1399–1406, 2006.
- [22] J. Whitehill, M. Bartlett, and J. Movellan. Automatic facial expression recognition for intelligent tutoring systems. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on*, pages 1 –6, 23-28 2008.
- [23] P. Yang, Q. Liu, and D. Metaxas. Boosting coded dynamic features for facial action units and facial expression recognition. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1 –6, 17-22 2007.
- [24] X. Zhao, E. Dellandréa, and L. Chen. A 3d statistical facial feature model and its application on locating facial landmarks. *ACIVS*, pages 686–697, 2009.