# Extraction and Selection of Muscle Based Features for Facial Expression Recognition

Kristin S. Benli
FMV IŞIK University,
Şile, Istanbul, Turkey
Email: kristin@isikun.edu.tr

M. Taner Eskil
FMV IŞIK University,
Şile, Istanbul, Turkey
Email: eskil@isikun.edu.tr

***Abstract*—In this study we propose a new set of muscle activity based features for facial expression recognition. We extract muscular activities by observing the displacements of facial feature points in an expression video. The facial feature points are initialized on muscular regions of influence in the first frame of the video. These points are tracked through optical flow in sequential frames. Displacements of feature points on the image plane are used to estimate the 3D orientation of a head model and relative displacements of its vertices. We model the human skin as a linear system of equations. The estimated deformation of the wireframe model produces an over-determined system of equations that can be solved under the constraint of the facial anatomy to obtain muscle activation levels. We apply sequential forward feature selection to choose the most descriptive set of muscles for recognition of basic facial expressions.**

## I. Introduction

Humans do two-way communication through not only words but also facial expressions, gestures and posture. Mehrabian [1] reported that the feeling conveyed by the speaker in face to face communication is 7% verbal, 38% vocal and 55% facial. According to his findings tone of voice and non-verbal behavior are more effective than the "spoken words". Non-verbal behavior includes facial expression, eye, hand and head movements, posture etc. Among all gestures, facial expressions are the most direct, natural and most of the time involuntary expressions of the emotions. Due to this fact facial expressions constitute a popular field of research in varying research domains, especially in psychology. Correct analysis of human faces is also valuable for computer science, e.g. for enhancing the user experience in human computer interaction (HCI). Latest efforts in HCI focus on detecting unsuitable conditions (boredom, fatigue and stress) of the staff who work in critical positions [2], [3].

One of the earliest known studies facial expressions is by John Bulwer in 1649 [4], who hypothesized that the 'motions of the mind and muscles of the head' are directly linked, establishing the muscular basis of facial expressions. In mid 1800s famous neurologist Duchenne de Boulogne [5] studied generation of facial expressions through electrical stimulation of facial muscles on live subjects. A decade later, Charles Darwin [6] hypothesized that facial expressions anger, disgust, fear, happiness, sadness and surprise are common among all cultures.

In 1978, which is about a century later after Darwin's work, Ekman and Friesen [7] proposed a systematic method for analyzing appearance changes on human face, namely facial action coding system (FACS). They defined facial behaviors with action units (AUs) rather than muscle activations. The most current FACS categorization uses 46 action units for describing the facial actions, head and eye movements. Ekman and Friesen's study became an important milestone for computer based automated facial expression recognition.

A computer based automated facial expression recognition system includes three main stages. The first stage is detecting the human face in an input, which can be an image or the first frame of a video sequence. The second stage is extracting the features that discriminate facial expressions in the input. When the observation is video, we have the opportunity to make use of the dynamics of the expression by tracking the face and its features. Most current approaches utilize the FACS AUs as features. The last stage is classification of the facial expression using the obtained numeric values of the features.

In this study we propose a set of new and robust features that represent the muscular activities on the face. We argue that it is possible to accurately and uniquely solve muscle forces that constitute a facial expression by precise tracking of feature points that are distributed over muscular regions of influence. We show through sequential forward selection (SFS) of muscle activity based features that it is possible to classify facial expressions with high accuracy with very few number of features.

The paper is organized as follows. In Section 2, we explore the state of the art studies in facial expression recognition field. We introduce our motivation in proposing muscle activity based features in Section 3. The new set of features is introduced in Section 4. We present our experimental results in Section 5 and conclude this paper in Section 6.

## II. Facial Expression Recognition Research

The initial step of facial expression recognition study is face detection. There are a significant number of studies in this topic and it is now considered to be solved by works of Rowley et al. [8], Schneiderman and Kanade [9] and Viola and Jones [10].

The next step is extracting discriminative information for recognizing expressions. Geometric, appearance based and hybrid methods have been proposed for feature extraction.

Geometric features are derived from the coordinates of distinctive regions of the face such as eyes, nose and lips. Detected or marked facial feature points are tracked in consecutive frames and their geometrical displacements are utilized to train different classifiers [11], [12], [13].

Appearance based methods deal with the texture of the skin including wrinkles, bulges and furrows. Gabor filters [14], [15], [16], Haar features [17], multilevel motion history images [18] and local binary pattern descriptors [19] are a few methods for extracting appearance based features.

Researchers also proposed hybrid approaches that combine the geometric and appearance based features at the feature or decision level. It has been shown that hybrid features consistently achieve better recognition performances than individual features [20], [21], [22].

Model based methods are proposed to derive a mathematical model of variation modes of geometric or appearance based features. Typical examples are active shape model (ASM) [23] and active appearance model (AAM) [24]. ASM is a parametric deformable model. A statistical shape model of the face object is built using a set of training examples. Pose and shape parameters are iteratively modified for a better fit. AAM combines the statistical model of the shape and the gray-level appearance of the object of interest. The synthesized model is projected onto the face image and matching is done iteratively. Lucey et al. [21], [25] and Cheon and Kim [26] derived features based on AAM and employed them for facial action and facial expression recognition tasks.

The last stage of facial expression recognition is classification. Kotsia et al. [20] classified six basic expressions and neutral and obtained 74.3% accuracy with texture based, and 84.8% accuracy with shape based features. They improved the classification accuracy to 92.3% by combining texture and shape based features. Cheon and Kim [26] achieved 86.5% accuracy in classification of six basic expressions with differential AAM features. In our previous work [27], we obtained 89.0% accuracy in classification of six basic expressions and neutral through muscle based features and Adaboost. These studies were carried out using leave-sequence-out strategy on the Cohn Kanade (CK) [28] dataset.

Lucey et al. [21] classified six basic expressions and contempt with shape and appearance based features, obtaining 50.4% and 66.7% classification rates, respectively. When combined, shape and appearance based features resulted in 83.3% classification accuracy. This study was done on the extended Cohn Kanade (CK+) [21] dataset. These classification performances are close to human recognition ceiling that is believed to be in range 87.0% [29] to 91.7% [30].

In this paper we propose the activation levels of a subset of facial muscles as novel features. Muscle forces are the ultimate base functions that compose all facial expressions under the constraint of facial anatomy. We utilize a semi-automatic customization method to fit a generic face model to subject's face. The projection of vertices that lay on the influence regions of muscles onto the image plane are identified as facial features and are tracked in consecutive frames. Coordinates of tracked feature points are utilized to estimate head orientation of the face model and relative displacements of vertices. We model human face with a system of springs and solve muscle forces through convex optimization.

## III. Motivation

FACS defines 46 AUs based on psychological studies of expressions. Each AU is defined with the motivation to represent a head, eye or face activity that is empirically known to relate to an emotion. Since the basis of FACS coding is humans' perception of emotions and not the anatomical structure of the face, an AU may refer to the action of a unique muscle or the compound visual effect of a set of muscles.

Action units compose facial expressions individually or in different combinations. The combinations can be additive or non-additive. In non-additive combinations, the compound effect alters the appearance of individual action units. Once AUs are compounded, it is extremely difficult to decompose an expression back to AUs unless a large rule base is made available. Investigation of facial expressions revealed more than 7,000 possible combinations of AUs [31].

We argue that the muscular activities can be uniquely and accurately solved by observing displacements of feature points that are dispersed in muscular regions of influence. Our reasoning can be outlined as follows:

1) Depending on the anatomical structure of the human face, we can estimate the layout of muscles for a person through customization of an anatomy based generic model to the detected face in an observed scene.
2) Once the model is customized, the estimated layout of muscles completely defines the muscular regions of influence on the skin.
3) The displacement of each feature point that lies in the region of influence of a muscle is an evidence of related muscular activity.
4) A set of $n$ feature points carefully distributed in regions of influence of $m$ muscles generates an over-determined system of equations if $n > m$. An error-minimizing solution for this system is achievable through convex optimization methods provided that the condition number of the coefficient matrix is low.

## IV. Muscle Force Based Features

### A. Generic wireframe model

In this study we use a generic wireframe model that conforms to the human face anatomy. High polygon generic face model (HIGEM) comprises of 612 nodes and 1128 polygonal surfaces [32]. HIGEM includes 18 major muscles of the human face. Each muscle is represented by an insertion point (on the skin) and an attachment point (on the skull). The structure of muscles on HIGEM is illustrated in Figure 1.

### B. Muscle Model

Based on Waters' research on 3 dimensional animation of facial expressions [33], we define our muscles as linear springs with distributed forces in their regions of influence.

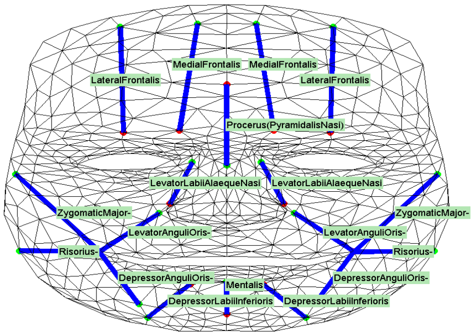


Fig. 1. The muscle structure on HIGEM.

We consolidate the effects of muscles on wireframe vertices in matrix $\mathbf{A}$, which serves as our muscle map [27]. The muscle map is solely dependent on the anatomical structure of human face. The product of $\mathbf{A}$ with the vector of muscle activations $\vec{\mathbf{f}}_{\mathbf{m}}$ produces the vector of muscle forces on each vertex in each axis $\vec{\mathbf{f}}_{\mathbf{s}}$ (Equation 1).

$$\mathbf{A}\vec{\mathbf{f}}_{\mathbf{m}} = \vec{\mathbf{f}}_{\mathbf{s}} \tag{1}$$

### *C. Semi-automatic customization*

Our feature extraction method commences with customization of the generic wireframe model (HIGEM) to a subject's face. We use the nearest neighbor weighted average customization (NNWA) in this stage, which is computationally efficient for customization of high-polygon models and sufficiently accurate for our purposes [32]. This procedure is done once for a subject and only on the first frame of the video.

### *D. Tracking facial features*

Given a model that has been customized for a subject, we can identify its vertices that are in the region of influence of a muscle using Waters' muscle model. We identify the facial feature points through projection of these vertices onto the image plane, as shown in Figure 2. These facial feature points will be tracked on the image plane using optical flow [34] .

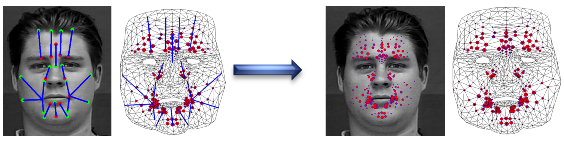

Fig. 2. Identifying feature points to be tracked.

### *E. Estimating head orientation*

Precise alignment of the face model with the observed face image is mandatory for estimation of relative displacements of vertices, i.e. the deformation due to the performed facial expression. The orientation of the subject's head is determined by greedy search on feature points and their corresponding vertices in 6 degrees of freedom. A similar greedy search algorithm for finding the head orientation was implemented by Dornaika and Ahlberg [35].

### *F. Estimating deformations*

Once the estimation of head orientation is complete, we have the face model aligned with the observed face on the image plane. Note that the projections of the wireframe vertices still would not precisely overlap with the facial feature points. The deviations between these tracking points and the projections of corresponding vertices serve as indicators of facial expressions. We apply ray tracing to extract the relative displacements of vertices.

Figure 3 depicts a landmark vertex $\mathbf{x}_0$ and its neighbors on the wireframe model. Assuming that the surfaces are small enough so that they do not bulge or wrinkle, this vertex hypothetically moves on one of the faces it resides on. In this illustration, $\mathbf{x}_0$ moves on the surface defined by $\mathbf{x}_0$, $\mathbf{x}_1$ and $\mathbf{x}_2$.

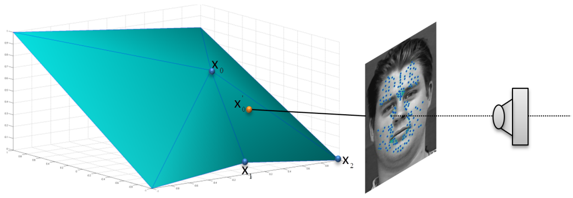

Fig. 3. Estimating the new coordinates of vertices through ray tracing.

If we can identify the plane of motion for the vertex, we can estimate its new coordinates through a line-plane intersection. The plane of motion can be any of the faces the vertex resides on. We find the intersection of the ray with each of these faces. The intersection point $\mathbf{x}'_0$ may be found within or outside the boundaries of a triangular face as depicted in Figure 4.

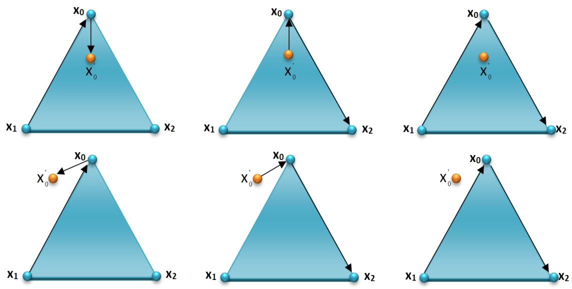

Fig. 4. Identifying the plane of motion.

To eliminate those intersection points that do not lie in the plane of motion, we determine three normal unit vectors for each face;

$$\begin{aligned}
\vec{\mathbf{n}}_1 &= \overrightarrow{\mathbf{x}_1\mathbf{x}_0} \times \overrightarrow{\mathbf{x}_0\mathbf{x}'_{\mathbf{0}}} / ||\overrightarrow{\mathbf{x}_1\mathbf{x}_0} \times \overrightarrow{\mathbf{x}_0\mathbf{x}'_{\mathbf{0}}}|| \\
\vec{\mathbf{n}}_2 &= \overrightarrow{\mathbf{x}'_{\mathbf{0}}\mathbf{x}_0} \times \overrightarrow{\mathbf{x}_0\mathbf{x}_2} / ||\overrightarrow{\mathbf{x}'_{\mathbf{0}}\mathbf{x}_0} \times \overrightarrow{\mathbf{x}_0\mathbf{x}_2}|| \\
\vec{\mathbf{n}}_3 &= \overrightarrow{\mathbf{x}_1\mathbf{x}_0} \times \overrightarrow{\mathbf{x}_0\mathbf{x}_2} / ||\overrightarrow{\mathbf{x}_1\mathbf{x}_0} \times \overrightarrow{\mathbf{x}_0\mathbf{x}_2}||
\end{aligned} \tag{2}$$

where $\mathbf{x}_0$, $\mathbf{x}_1$ and $\mathbf{x}_2$ are the vertex and its neighbors on the *aligned* wireframe model. The intersection point $\mathbf{x}'_0$ is the back projection of the tracked feature point found through line plane intersection. Note that for $\mathbf{x}'_0$ to be in the region bounded by the face, normal vectors must point to the same direction;

$$\vec{\mathbf{n}}_1 \cdot \vec{\mathbf{n}}_3 > 1 - \epsilon \quad and \quad \vec{\mathbf{n}}_2 \cdot \vec{\mathbf{n}}_3 > 1 - \epsilon \tag{3}$$

These two conditions enable us to identify the plane of motion and the new coordinates of the corresponding vertices.

### G. Solving muscle forces

The wireframe is modeled as a 3D surface that is composed of polygons. Each face on the wireframe model is defined by three vertices, representing a triangular plane in the 3D space. The edges between each neighboring vertices are modeled with springs as illustrated in Figure 5.

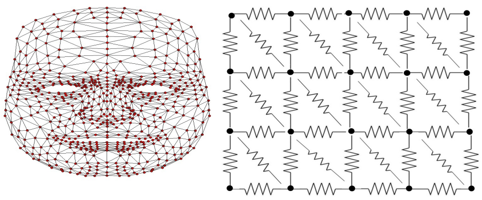

Fig. 5. Representing the edges of the wireframe model with springs.

The spring force vector on vertex $i$, exerted by the contraction of the spring between vertices $i$ and $j$, can be represented using Hooke's elasticity law;

$$\vec{\mathbf{f}}_{ij} = k_{ij}(l_{ij} - ||\mathbf{x}_i - \mathbf{x}_j||)\frac{\mathbf{x}_i - \mathbf{x}_j}{||\mathbf{x}_i - \mathbf{x}_j||} \tag{4}$$

In Equation 4, $k_{ij}$ is the stiffness of the spring attached to vertices $i$ and $j$, $l_{ij}$ is the rest length of this spring, $\mathbf{x}_i$ and $\mathbf{x}_j$ are the 3D coordinates of the vertices.

Given the 3D coordinates of wireframe vertices (Section IV-F) and the stiffness matrix we can compute the external forces on each vertex using Equation 4. Our aim is to extract muscle activations from the external forces under the constraint of human anatomy. Under the anatomic constraints, external forces are represented in terms of muscular activations as shown in Equation 1. Equations 1 and 4 constitute a linear, over-determined system of equations. We use constrained least squares optimization ($\vec{\mathbf{f}}_{\mathbf{m}} > 0$) to solve this system.

## V. Experimental Results

We carried out our classification experiments using support vector machine (SVM) classifier on the original CK [28] and extended CK+ [21] datasets. The original CK dataset contains 228 image sequences labeled as one of the six basic expressions. The extended version, CK+, contains 327 image sequences of seven expressions including contempt. The input to our classifier is a set of muscle activations for a snapshot image. We consistently used leave-sequence-out cross-validation scheme in our experiments.

Table I presents our results for the multi-class SVM classifier on the CK dataset. We obtained the lowest classification accuracy in fear (64.7%). It was confused with anger, happiness and sadness. We achieved highest performance in the surprise expression (98.6%). The overall classification performance is found as 84.9%.

TABLE I
Classification of six basic expressions on the CK dataset with leave-sequence-out cross-validation. A: Anger, D: Disgust, F: Fear, H: Happy, Sa: Sad, Su: Surprise

| | **A** | **D** | **F** | **H** | **Sa** | **Su** | % |
|---|---|---|---|---|---|---|---|
| **A** | **26** | 1 | 0 | 0 | 2 | 0 | 89.7 |
| **D** | 3 | **30** | 0 | 0 | 0 | 1 | 88.2 |
| **F** | 1 | 0 | **11** | 3 | 2 | 0 | 64.7 |
| **H** | 0 | 1 | 3 | **57** | 0 | 0 | 93.4 |
| **Sa** | 2 | 0 | 2 | 0 | **12** | 0 | 75.0 |
| **Su** | 0 | 0 | 1 | 0 | 0 | **70** | 98.6 |

Table II presents our results for the multi-class SVM classifier on CK+ dataset. We obtained the lowest classification rate in contempt (44.4%). A significant percentage of contempt examples were misclassified as anger (27.7%). Once again we obtained the highest accuracy in the surprise expression (94.0%). The overall classification performance is found as 75.5%.

TABLE II
Classification of six basic expressions and contempt on the CK+ dataset with leave-sequence-out cross-validation. A: Anger, D: Disgust, F: Fear, H: Happy, Sa: Sad, Su: Surprise, C: Contempt.

| | **A** | **D** | **F** | **H** | **Sa** | **Su** | **C** | % |
|---|---|---|---|---|---|---|---|---|
| **A** | **32** | 1 | 3 | 0 | 5 | 0 | 4 | 71.1 |
| **D** | 2 | **54** | 1 | 0 | 0 | 0 | 2 | 91.5 |
| **F** | 3 | 1 | **15** | 2 | 3 | 0 | 1 | 60.0 |
| **H** | 2 | 1 | 4 | **59** | 0 | 1 | 2 | 85.5 |
| **Sa** | 3 | 0 | 2 | 0 | **23** | 0 | 0 | 82.1 |
| **Su** | 1 | 0 | 3 | 0 | 0 | **78** | 1 | 94.0 |
| **C** | 5 | 2 | 0 | 2 | 1 | 0 | **8** | 44.4 |

We utilized sequential forward selection (SFS) strategy to obtain the best subset of muscle based features. SFS starts search with an empty set and sequentially selects the most significant feature. This procedure is repeated until there is no improvement in classification performance.

Table III presents the results of SFS on the CK dataset. SFS selects 11 features as the most descriptive. New feature set increases the performance of fear expression from 64.7% to 82.4%. The overall classification performance increases from 84.9% to 85.9%.

TABLE III
Feature selection results on the CK database.

| | **A** | **D** | **F** | **H** | **Sa** | **Su** | % |
|---|---|---|---|---|---|---|---|
| **A** | **22** | 1 | 1 | 0 | 5 | 0 | 75.9 |
| **D** | 3 | **30** | 0 | 0 | 0 | 1 | 88.2 |
| **F** | 1 | 0 | **14** | 2 | 0 | 0 | 82.4 |
| **H** | 1 | 0 | 7 | **53** | 0 | 0 | 86.9 |
| **Sa** | 1 | 0 | 1 | 0 | **14** | 0 | 87.5 |
| **Su** | 0 | 0 | 2 | 1 | 1 | **67** | 94.4 |

Table IV presents the results of SFS on the CK+ dataset. This time SFS chooses a subset of size 9. The performance of contempt expression is increased to 55.6% with the selected feature set. The overall classification performance increases from 75.5% to 77.6%.

TABLE IV
FEATURE SELECTION RESULTS ON THE CK+ DATABASE.

| | A | D | F | H | Sa | Su | C | % |
|---|---|---|---|---|---|---|---|---|
| **A** | **33** | 1 | 2 | 1 | 7 | 0 | 1 | 73.3 |
| **D** | 5 | **49** | 1 | 0 | 0 | 1 | 3 | 83.1 |
| **F** | 3 | 1 | **15** | 4 | 2 | 0 | 0 | 60.0 |
| **H** | 2 | 1 | 4 | **58** | 0 | 1 | 3 | 84.1 |
| **Sa** | 0 | 0 | 1 | 0 | **27** | 0 | 0 | 94.4 |
| **Su** | 2 | 1 | 1 | 1 | 0 | **77** | 1 | 92.8 |
| **C** | 4 | 1 | 1 | 1 | 1 | 0 | **10** | 55.6 |

We obtained slightly higher classification performances both on CK (+1.0%) and CK+ (+2.1%) datasets with SFS. Muscle based features proposed in this paper competes with the state-of-the-art algorithms that use geometric, appearance or FACS based features in classification.

Table V presents comparative evaluation of our approach with our previous work [27], Kotsia et al. [20] and Cheon and Kim [26]. These studies were carried out on the CK dataset. Texture based features utilized by Kotsia et al. are inherently more complex than geometric or muscle based features, although the actual number of dimensions is not available. In this study 2D displacements of grid nodes (104 for Candide model) were used as shape based features. Muscle activities based feature subset outperforms both inherently complex texture based, and high number of shape based features.

Cheon and Kim [26] chose six basis vectors for each of the shape, appearance and AAM modalities on the avearage, making a total of approximately 18 features. They obtained slightly better results from the proposed feature set with differential AAM. Note that the AAM approach requires an iterative fit for orientation, shape and appearance of the face, which implies that extraction of AAM based features is significantly more complex than muscle based features.

Table VI compares our approach with our previous work [27] and Lucey et al. [21]. These studies were carried out on the CK+ dataset. Lucey et al. represent shape based features with a 136 dimensional feature vector. Gray values of 8,091 pixels ($87 \times 93$ window) denote the appearance based features. Muscle activities based feature subset achieves higher performance than shape based and appearance based features when they are used alone. Note also that the number muscle activity based features (9) is orders of magnitude lower than shape based (136) and appearance based (8,091) features.

## VI. Conclusion

In this study we argue that facial muscles are the ultimate basis functions that compose all facial expressions. We propose extraction and selection of muscle activity based features for recognition of facial expressions. The activity levels of facial muscles are extracted unintrusively, through tracking of facial feature points, mapping their relative displacements onto a 3D model, and finding an error-minimizing solution for muscle forces in an over-determined system of linear equations.

We chose the most descriptive features among all muscle activities through SFS. On the CK database, we obtained 85.9% classification performance on six basic expressions with only 11 features. On the CK+ database our classification performance is 77.6% over 7 basic expressions and using only 9 features. These performance figures are better than that of shape and texture based features when they are used alone. The number of muscle based features we used in our experiments is an order of magnitude lower than shape based, and two orders of magnitude lower than texture based features that are commonly used in the literature. This result implies that muscle based features are significantly more descriptive of facial features, as our starting intuition suggests.

## Acknowledgment

This research is part of project “Expression Recognition based on Facial Anatomy”, grant number 109E061, supported by The Support Programme for Scientific and Technological Research Projects of The Scientific and Technological Research Council of Turkey (TÜBİTAK).

## References

[1] A. Mehrabian, *Silent messages: implicit communication of emotions and attitudes*. Wadsworth Pub. Co., 1981.

[2] H. Gu and Q. Ji, “Information extraction from image sequences of real-world facial expressions,” *Machine Vision and Applications*, vol. 16, no. 2, pp. 105+, 2005.

[3] E. Vural, M. S. Bartlett, G. Littlewort, M. etin, A. Eril, and J. R. Movellan, “Discrimination of moderate and acute drowsiness based on spontaneous facial expressions,” in *20th International Conference on Pattern Recognition, ICPR 2010, Istanbul, Turkey, 23-26 August 2010*. IEEE, 2010, pp. 3874–3877.

[4] J. Bulwer, *Pathomyotamia, or, A dissection of the significative muscles of the affections of the minde being an essay to a new method of observing the most important movings of the muscles of the head*. EEBO Editions, ProQuest, 1649.

[5] G. de Boulogne and R. Cuthbertson, *The Mechanism of Human Facial Expression*, ser. Studies in Emotion and Social Interaction. Cambridge University Press, 1990.

[6] C. Darwin, *The Expression of the Emotions in Man and Animals*, anniversary ed., P. Ekman, Ed. Harper Perennial, 1872/2009.

[7] P. Ekman and W. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto: Consulting Psychologists Press, 1978.

[8] H. A. Rowley, S. Member, S. Baluja, and T. Kanade, “Neural network-based face detection,” *IEEE Transactions On Pattern Analysis and Machine intelligence*, vol. 20, pp. 23–38, 1998.

[9] H. Schneiderman and T. Kanade, “A statistical model for 3d object detection applied to faces and cars,” in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2000, pp. 1746–1759.

[10] P. Viola and M. Jones, “Robust real-time face detection,” *International Journal of Computer Vision*, vol. 57, pp. 137–154, 2004.

[11] A. A. Seyedarabi, H. and S. Khanmohammadi, “Analysis and synthesis of facial expressions by feature-points tracking and deformable model.” *Journal of Iranian Association of Electrical and Electronic Engieers (IAEEE)*, vol. 4, no. 1, 2007.

[12] I. Kotsia and I. Pitas, “Facial expression recognition in image sequences using geometric deformation features and support vector machines,” *Image Processing, IEEE Transactions on*, vol. 16, no. 1, pp. 172–187, 2007.

[13] K. Lu and X. Zhang, “Facial expression recognition from image sequences based on feature points and canonical correlations,” in *Artificial Intelligence and Computational Intelligence (AICI), 2010 International Conference on*, vol. 1, 2010, pp. 219–223.

TABLE V

COMPARATIVE EVALUATION OF THE PROPOSED FEATURES ON THE CK DATASET. SUPERSCRIPT $n$ INDICATES THAT NEUTRAL IS INCLUDED IN EXPRESSION CLASSES.

| Study | Feature Type | No.of Features | No.of Classes | Methodology | Success Rate |
|---|---|---|---|---|---|
| **Muscle based features** | Geometry | 11 | 6 | SVM | 85.9% |
| Eskil and Benli [27] | Geometry | 18 | $7^n$ | NB | 77.8% |
| | | | | SVM | 87.0% |
| | | | | Adaboost | 89.0% |
| Kotsia et al. [20] | Texture | N/A | $7^n$ | DNMF | 74.3% |
| | Shape | 208 | | SVM | 84.8% |
| | Combined | | | MRBF NN | 92.3% |
| Cheon and Kim [26] | AAM | 18 | 6 | k-NNS | 73.9% |
| | Diff. AAM | 18 | | | 86.5% |

TABLE VI

COMPARATIVE EVALUATION OF THE PROPOSED FEATURES ON THE CK+ DATASET

| Study | Feature Type | No.of Features | No.of Classes | Methodology | Success Rate |
|---|---|---|---|---|---|
| **Muscle based features** | Geometry | 9 | 7 | SVM | 77.6% |
| Eskil and Benli [27] | Geometry | 18 | 7 | SVM | 75.5% |
| | | | | Adaboost | 76.0% |
| Lucey et al. [21] | Shape | 136 | 7 | SVM | 50.4% |
| | Appearance | 8,091 | | | 66.7% |
| | Combined | | | | 83.3% |


[14] G. Guo and C. R. Dyer, “Learning from examples in the small sample case: face expression recognition,” *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, vol. 35, no. 3, pp. 477–488, 2005.

[15] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. R. Fasel, and J. R. Movellan, “Recognizing facial expression: Machine learning and application to spontaneous behavior,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), 20-26 June 2005, San Diego, CA, USA*, 2005, pp. 568–573.

[16] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, “Fully automatic facial action recognition in spontaneous behavior,” in *in 7th International Conference on Automatic Face and Gesture Recognition*, 2006, pp. 223–230.

[17] J. Whitehill and C. W. Omlin, “Haar features for facs au recognition,” in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, ser. FGR ’06, 2006, pp. 97–101.

[18] M. F. Valstar, M. Pantic, and I. Patras, “Motion history for facial action detection in video,” in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics: The Hague, Netherlands, 10-13 October 2004*, 2004, pp. 635–640.

[19] B. Jiang, M. F. Valstar, and M. Pantic, “Action unit detection using sparse appearance descriptors in space-time video volumes,” in *Ninth IEEE International Conference on Automatic Face and Gesture Recognition (FG 2011), Santa Barbara, CA, USA, 21-25 March 2011*. IEEE, 2011, pp. 314–321.

[20] I. Kotsia, S. Zafeiriou, and I. Pitas, “Texture and shape information fusion for facial expression and facial action unit recognition,” *Pattern Recognition*, vol. 41, no. 3, pp. 833–851, 2008.

[21] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, “The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression,” in *Computer Vision and Pattern Recognition*, 2010.

[22] J. Chen, D. Chen, Y. Gong, M. Yu, K. Zhang, and L. Wang, “Facial expression recognition using geometric and appearance features,” in *Proceedings of the 4th International Conference on Internet Multimedia Computing and Service*, ser. ICIMCS ’12, 2012, pp. 29–33.

[23] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, “Active shape models-their training and application,” *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[24] T. F. Cootes, G. J. Edwards, and C. J. Taylor, “Active appearance models,” in *Computer Vision - ECCV 98, 5th European Conference on Computer Vision, Freiburg, Germany, June 2-6, 1998, Proceedings, Volume II*, ser. Lecture Notes in Computer Science, vol. 1407, 1998, pp. 484–498.

[25] S. Lucey, I. Matthews, C. Hu, Z. Ambadar, F. D. la Torre, and J. F. Cohn, “AAM derived face representations for robust facial action recognition,” in *Seventh IEEE International Conference on Automatic Face and Gesture Recognition (FG 2006), 10-12 April 2006, Southampton, UK*, 2006, pp. 155–162.

[26] Y. Cheon and D. Kim, “Natural facial expression recognition using differential-aam and manifold learning,” *Pattern Recognition*, vol. 42, no. 7, pp. 1340–1350, 2009.

[27] M. Eskil and K. Benli, “Facial expression recognition based on anatomy,” *Computer Vision and Image Understanding*, accepted.

[28] T. Kanade, J. Cohn, and Y.-L. Tian, “Comprehensive database for facial expression analysis,” in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition (FG’00)*, March 2000, pp. 46 – 53.

[29] J. N. Bassili, “Facial motion in the perception of faces and of emotional expression.” *J. Exp. Psychol. Human*, vol. 4, no. 3, pp. 373–379, Aug. 1978.

[30] P. Ekman and W. V. Friesen, *Pictures of facial affect*. Consulting Psychologists Press, 1976.

[31] K. R. Scherer and P. Ekman, *Handbook of methods in nonverbal behavior research*. Cambridge and New York: Cambridge University Press, 1982.

[32] H. P. Abeysundera, K. S. Benli, and M. T. Eskil, “Nearest neighbor weighted average customization for modeling faces,” *Machine Vision and Applications*, vol. 24, no. 7, pp. 1525–1537, 2013.

[33] K. Waters, “A muscle model for animation three-dimensional facial expression,” in *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH ’87, 1987, pp. 17–24.

[34] B. D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI 81), Vancouver, BC, Canada, August 1981*, 1981, pp. 674–679.

[35] F. Dornaika and J. Ahlberg, “Fitting 3d face models for tracking and active appearance model training,” *Image Vision Comput.*, vol. 24, no. 9, pp. 1010–1024, 2006.