Eyes Do Not Lie: Spontaneous Versus Posed Smiles

Dibeklioğlu, H.; Valenti, R.; Salah, A.A.; Gevers, T.

DOI
10.1145/1873951.1874056

Publication date
2010

Document Version
Author accepted manuscript

Published in
MM '10: proceedings of the ACM Multimedia 2010 International Conference: October 25-29, 2010, Firenze, Italy

Eyes Do Not Lie: Spontaneous versus Posed Smiles

Hamdi Dibeklioğlu, Roberto Valenti, Albert Ali Salah, and Theo Gevers
Intelligent Systems Lab Amsterdam, University of Amsterdam
Science Park 107, Amsterdam, The Netherlands
h.dibeklioglu@uva.nl, r.valenti@uva.nl, a.a.salah@uva.nl, th.gevers@uva.nl

ABSTRACT

Automatic detection of spontaneous versus posed facial expressions received a lot of attention in recent years. However, almost all published work in this area use complex facial features or multiple modalities, such as head pose and body movements with facial features. Besides, the results of these studies are not given on public databases. In this paper, we focus on eyelid movements to classify spontaneous versus posed smiles and propose distance-based and angular features for eyelid movements. We assess the reliability of these features with continuous HMM, k-NN and naïve Bayes classifiers on two different public datasets. Experimentation shows that our system provides classification rates up to 91 per cent for posed smiles and up to 80 per cent for spontaneous smiles by using only eyelid movements. We additionally compare the discrimination power of movement features from different facial regions for the same task.

Categories and Subject Descriptors

I.2.10 [**Vision and Scene Understanding**]: Video analysis; H.1.2 [**User/Machine Systems**]: Human factors, Human information processing

General Terms

Human Factors, Algorithms, Experimentation

Keywords

Facial expression, Spontaneous versus posed smile detection, Eyelid movements

1. INTRODUCTION

It is well known that facial expressions are the main means of bodily communication. In fact, just by looking at one person's face, it is possible to deduce his feelings and his state of mind. Facial expression analysis can give a more natural form to human-computer interaction, similar to communication between humans. The literature on automated recognition of facial expressionsis is extensive, we refer the reader to [11] and to the more recent [16]. Newer approaches have shifted the focus to the automatic recognition of deceit, or the discrimination of spontaneous versus posed facial expressions. Specifically, [4] and [15] focused on the detection of fake smiles. However, to our knowledge, no computational study has been made for the search of what is the most important facial feature that indicates genuineness.

Recent work on spontaneous smile detection generally focuses on temporal changes in shape and facial activation information [9, 1]. Cohn and Schmidt analyzed amplitude and duration of smile onsets and showed that spontaneous smiles have smaller amplitude, but a more stable relation between amplitude and duration [4]. By using a linear discriminant classifier, they reported 93 per cent classification accuracy for spontaneous and posed smiles. [5] analyzed coordination of facial movements, head rotation, and eye motion during spontaneous smiles. Valstar et al. proposed a multimodal system to classify posed and spontaneous smiles with fusion of shoulder, head and inner facial movements, and reported 94 per cent accuracy by using all these modalities [15].

It is said that eyes are the mirror of the soul. We believe that this expression indicates the fact that much of the state of mind of one person can be seen through the eyes. Therefore, this work focuses on the use of this specific feature to detect posed smiles. As a case study, we analyze what precisely in the eyes can be used to discriminate between posed and spontaneous smiles.

In the rest of the paper, we will validate that eyes are indeed the most discriminative feature for the task. Our second contribution is a detector of genuineness of smiles, which achieves comparable results to state of the art methods by using eyelid features only. We have no claim to ecological validity, as we disregard holistic face perception in humans [10], and physiological measures do not necessarily reflect perceptual experience of observers. Yet, this doesn't preclude the possibility that some individual features may contain complete information for the classification of a given expression. The main advantage of the proposed method is in the fact that it is possible to use it to determine spontaneity even in cases of major facial occlusions (eg. when wearing a scarf) and under voluntary suppression of expression. It is known that in some cultures, people suppress their emotions by preserving a close-to-neutral face in the presence of authorities.

MM'10, October 25–29, 2010, Firenze, Italy.

2. FACIAL FEATURES

2.1 Facial Muscles

Smiles can be simply distinguished from other facial expressions by analyzing movement of mouth corners and cheeks. But as it is known, there are two distinct types of smiles, which are distinguished as spontaneous (felt) and posed. This distinction was predicted by Guillaume Duchenne in the mid-nineteenth century, so spontaneous smiles are also called Duchenne smiles. Researchers confirmed Duchenne's observation with empirical findings after 120 years [7]. In this study, Ekman proposed that spontaneous smiles are formed by the contraction of both the zygomatic major and the orbicularis oculi muscles, where posed smiles involve only the zygomatic major muscle. Zygomatic major muscle raises the corners of the mouth, and orbicularis oculi muscle raises the cheeks and forms crows-feet around the eyes. Ekman also reported that asymmetry and timing of smiles are discriminative to classify different types of smiles [7]. We hypothesize that the eye openings can become smaller by the activity of orbicularis oculi muscle in spontaneous smiles, which can cause lowered eyelids.

2.2 Features

In this work we focus on eyelid movements to distinguish posed and spontaneous smiles. First we track 21 points on eyebrows, eyelids, eye corners, nose tip and mouth corners, as shown in Figure 1, to assess the reliability of eye region over other face regions for smile expressions. Facial feature points are grouped into four main regions as eyebrows, eyelids & eye corners, cheeks and mouth corners. Then, landmarks in each group are manually initialized at the first frame, and tracked. Estimated movements are normalized with respect to tilt rotation, translation and scale of the face, and used for classification. Preliminary results showed that the eye region is the most reliable one. We then focus on eye region of the face and analyzed temporal changes in this region. We extract distance-based and angular features to discriminate the movements of eyelids. For this purpose, we track inner and outer eye corners, eyelids and nose tip (15 points, see Figure 2). We manually initialize the tracker on the first frames of the videos for reliable and accurate tracking. Before feature extraction, all faces are normalized in terms of rotation, scale and translation. For normalization, eye centers are estimated as middle points between inner and outer eye corners. The tilt rotation of the face is estimated and normalized using the line between eye centers. After normalization of rotation, face is translated to origin with respect to nose tip. Then, inter-occular distance d_{io} (distance between eye centers) is calculated and the face is scaled with a factor of $100/d_{io}$.

We use an angular measure, β_m, to determine the amount of eye opening. β_m is the angle between v_1 and v_2 where v_1 and v_2 denote the vectors from outer eye corner to closest eyelid landmark and from outer eye corner to inner eye corner, respectively (see Figure 3). We calculate β_m for left and right eyes, seperately.

If d_m denotes the Euclidean distance between eyelid and the eye center, displacement of eyelids can be defined as change in d_m (see Figure 3). To estimate displacement of eyelids, d_m values are normalized by subtracting the d_m value of the first frame. Displacement of eyelids are calculated for both left and right eyes seperately. As indicated

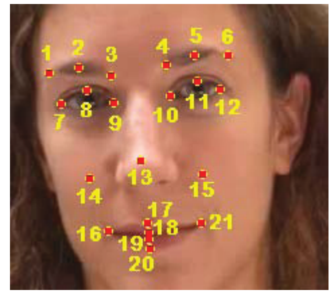

Figure 1: Feature points on different face regions.

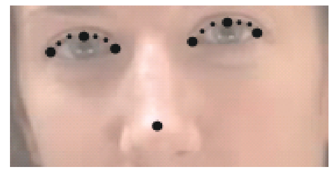

Figure 2: Feature points on eye corners, eyelids and nose tip.

in recent studies [12], asymmetry in movements of different sides of the face can be discriminative. Consequently, difference between left and right eyelid displacements is used as an asymmetry feature.

3. METHOD

3.1 Tracking

The face tracking used in our system is based on the system proposed in [3], which is in turn based on the system developed by Tao and Huang [13], called the Piecewise Bézier Volume Deformation (PBVD) tracker. This tracker constructs an explicit 3D wireframe model of the face. The generic face model consists of 16 surface patches embedded in Bézier volumes and is warped to fit selected facial features (such as the eye and mouth corners) manually selected in the first frame of the image sequence. Given a set of $n+1$ control points $b_0, b_1, \ldots, b_n$, the corresponding Bézier curve (or Bernstein-Bézier curve) is given by

$$x(u) = \sum_{i=0}^{n} b_i B_i^n(u) = \sum_{i=0}^{n} b_i \binom{n}{i} u^i (1-u)^{n-i}, \qquad (1)$$

where the shape of the curve is controlled by the control points b_i and $u \in [0,1]$. As the control points are moved, a new shape is obtained according to the Bernstein polynomials $B_i^n(u)$ in Eq. (1). The displacement of a point on the

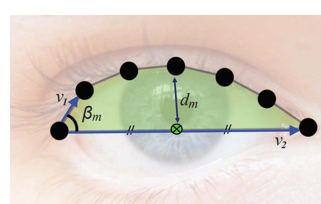

Figure 3: Movement features of eyelids.

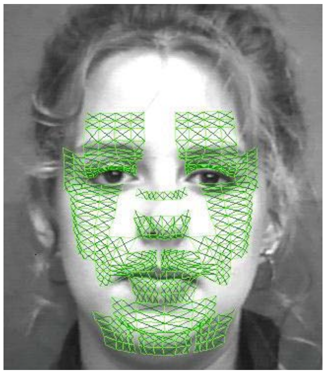

Figure 4: The Bézier volume model.

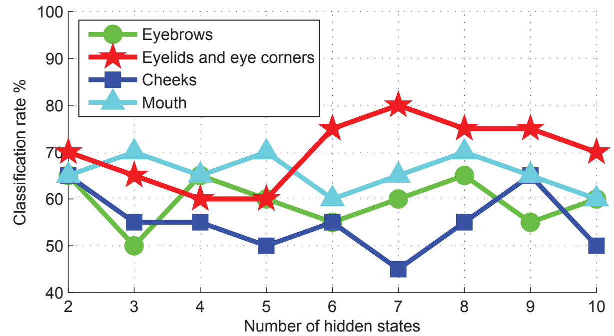

Figure 5: Classification rates with different features and different number of hidden states.

curve can be described in terms of linear combinations of displacements of the control points.

The Bézier volume is a straight-forward extension of the Bézier curve and is defined as $V = BD$, where V is the displacement of the mesh nodes, D is a matrix whose columns are the control point displacement vectors of the Bézier volume, and B is the mapping in terms of Bernstein polynomials. In other words, the change in the shape of the face model can be described in terms of the deformations in D.

Once the model is constructed and fitted (see Figure 4), both general head motion and local deformations of the facial features, such as the eyebrows, eyelids, and mouth, can be tracked. First 2D image motions are measured using template matching between frames at different resolutions. Image templates from previous frames are used for more robust tracking. The measured 2D image motions are modelled as projections of the true 3D motions onto the image plane. From the 2D motions of several points on the mesh, the 3D motion can be estimated. In our specific case we need to focus on fine details that change between spontaneous and posed smiles. To this end, instead of collecting these motions into Ekman AU's as in [14], each tracked point is used as a feature in the final estimation in our system.

3.2 Classification

To analyze the discriminative power of eyelid movements over other regions for classification of spontaneous versus posed smiles, we model movement features of eyebrows, eyelids & eye corners, nose tip and mouth corners by continuous hidden Markov models (CHMM) [6]. Each group is modelled by two left-to-right CHMMs (one for spontaneous, one for posed smiles), separately. For continuous input structure of HMM we use a mixture of Gaussians to represent observations for each region. Six Gaussians are used for each mixture. For classification, log-likelihood scores of spontaneous and posed CHMMs are checked and the class with the highest score is selected.

We employ both the indicated CHMM structure, k-nearest neighbor (k-NN) algorithm and naïve Bayes classifier with eyelid-focused features to classify smiles. For naïve Bayes classifier and k-NN, we use standard deviation, maximum, minimum and mean values of each stream as features, instead of using feature sequences. So we have four parameters for each feature stream. Neighborhood size of k-NN is selected as one (k=1), empirically.

4. EXPERIMENTS

In this section, the accuracy of the proposed system is evaluated on two different public datasets.

4.1 Datasets

We use the BBC smile dataset [2] and the Cohn-Kanade AU-Coded Facial Expression Database [8] for our experiments. BBC smile dataset was gathered from "Spot the fake smile" test of Paul Ekman on BBC website. There are videos of 20 subjects, each starting and ending with a neutral face and showing a posed/spontaneous smile. Image resolution is 314×286 pixels. The Cohn-Kanade AU-Coded Facial Expression Database has approximately 500 image sequences from 100 subjects, each starting with a neutral face and showing a basic expression. Only frontal images are open to public use, and we only used smile sequences in those (46 sequences). Image resolution is 640×480 pixels. These datasets were manually landmarked for tracking initialization. 10-fold (trained with nine folds and tested on the remaining fold) cross-validation is used for BBC smile dataset. The Cohn-Kanade database has only posed smiles, hence the tests we run with it involve training on BBC, and measure cross-database generalization.

4.2 Results and Discussion

First, we test the reliability of features on eye region with respect to other regions for classification. As indicated in Section 3.2, CHMM-based classification is tested on eyebrows, eyelids & eye corners, nose tip and mouth corners with different number of hidden states. Our results show that for each feature group, eyelid movements provide the highest classification rates (up to 80 per cent) with six and more hidden states, as shown in Figure 5. After the verification of our hypothesis, we test the proposed distance-based and angular features which focus on eyelids. Classification results with CHMM, 1-NN and naïve Bayes classifiers on BBC and Cohn-Kanade datasets are given in Table 1 and Table 2, respectively. Confusion of CHMM, 1-NN and naïve Bayes classifiers on BBC are same and the classification rate is 85.0 per cent. There are only posed smiles on Cohn-Kanade dataset, so the results on this dataset are given just for posed smiles with a training on BBC dataset. Classification rates on the Cohn-Kanade dataset with CHMM, 1-NN and naïve Bayes classifiers are 82.6 per cent, 87.0 per cent, and 91.3 per cent, respectively. As it is reported, the highest classification rates are provided by naïve Bayes classifier with standart deviation, minimum, maximum and mean values of the feature sequences. Besides, it is obvious that proposed features are usable with simple classifiers as k-NN and naïve Bayes, instead of using CHMM with its time consuming training phase.

It is interesting that the highest results are obtained on the Cohn-Kanade dataset with a training on the BBC dataset. To find the reason of this increase we analyzed the eyelid features. Extracted features showed that eyelid positions are higher (eyes are more opened), and more stationary in both datasets, except blinks in posed smiles. As a result, we can say that people tend to lower their eyelids while they are smiling spontaneously. Image sequences are much shorter in the Cohn-Kanade dataset, which means there are less eye blinks in the smile sequences of this dataset with respect to the BBC dataset. Indicated change might raise the mean position of the eyelids, which in turn might increase the classification rate of posed smiles on the Cohn-Kanade dataset. Additionally, classification of posed smiles (90 per cent on the BBC dataset) are more accurate than spontaneous smiles (80 per cent on the BBC dataset), because eyelids are more stationary (except for blinks) in posed smiles.

Table 1: Classification results of spontaneous versus posed smiles on BBC dataset. All three classifiers give the same results.

Classified Class	Real Class	
	Spontaneous	Posed
Spontaneous	8	2
Posed	1	9

Table 2: Classification results of posed smiles on Cohn-Kanade dataset. There are no spontaneous smiles in Cohn-Kanade dataset.

Classifier	Classified Class		Real
	Spontaneous	Posed	Class
CHMM	8	38	Posed
1-NN	6	40	Posed
Naïve Bayes	4	42	Posed

5. CONCLUSIONS

In this paper, we have presented a smile classifier, which can distinguish posed and spontaneous smiles. The method is based on the hypothesis that eyelid movements can identify the smiles. First, we compared discrimination power of the eye region movements with other facial movements and showed the reliability of the eyelid movements for classification of smiles. Then, we assessed the performance of the system with proposed eyelid-focused features. Our system reached 85 per cent and 91 per cent classification rates on BBC and Cohn-Kanade datasets, respectively. The obtained results are very promising, as they only use eyelid movements. Additionally, the proposed system is fast enough for real time usage.

Currently, a more extensive spontaneous/posed smile dataset is being collected for further experiments and more detailed analysis of eyelid movements. The proposed method is being extended for lower eyelids as well.

6. REFERENCES

[1] M. S. Bartlett, G. Donato, J. R. Movellan, J. C. Hager, P. Ekman, and T. J. Sejnowski. Face image analysis for expression measurement and detection of deceit. In *Proc. Joint Symposium on Neural Computation*, pages 8–15, 1999.

[2] BBC-Dataset. `http://www.bbc.co.uk/science/humanbody/mind/surveys/smiles/`.

[3] I. Cohen, N. Sebe, L. Chen, A. Garg, and T. S. Huang. Facial expression recognition from video sequences: Temporal and static modelling. *Computer Vision and Image Understanding*, pages 160–187, 2003.

[4] J. Cohn and K. Schmidt. The timing of facial motion in posed and spontaneous smiles. *Int. Journal of Wavelets, Multiresolution and Information Processing*, 2:1–12, March 2004.

[5] J. F. Cohn, L. I. Reed, T. Moriyama, J. Xiao, K. Schmidt, and Z. Ambadar. Multimodal coordination of facial action, head rotation, and eye motion during spontaneous smiles. In *Proc. AFGR*, pages 129–135, 2004.

[6] S. David, M. A. Ferrer, C. M. Travieso, and J. B. Alonso. gpdsHMM: A hidden Markov model toolbox in the MATLAB environment. In *Proc. Complex Systems Intelligence and Modern Tech. App.*, 2004.

[7] P. Ekman and W. V. Friesen. Felt, false, and miserable smiles. *Journal of Nonverbal Behavior*, 6:238–252, 1982.

[8] T. Kanade, J. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. In *Proc. AFGR*, 2000.

[9] E. Krumhuber and A. Kappas. The role of dynamic components for the perception of the genuineness of smiles. *Journal of Nonverbal Behavior*, 29:3–24, 2005.

[10] H. Leder and C. Carbon. When context hinders! Learn–test compatibility in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, 58(2):235–250, 2005.

[11] M. Pantic and L. Rothkrantz. Automatic analysis of facial expressions: The state of the art. *IEEE TPAMI*, 22(12):1424–1445, 2000.

[12] K. L. Schmidt, Z. Ambadar, J. F. Cohn, and L. I. Reed. Movement differences between deliberate and spontaneous facial expressions: Zygomaticus major action in smiling. *Journal of Nonverbal Behavior*, 30:37–52, 2006.

[13] H. Tao and T. Huang. Connected vibrations: a modal analysis approach for non-rigid motion tracking. In *Proc. CVPR*, pages 735–740, 1998.

[14] Y. Tong, W. Liao, and Q. Ji. Facial action unit recognition by exploiting their dynamic and semantic relationships. *IEEE TPAMI*, 29(10):1683–1699, 2007.

[15] M. F. Valstar and M. Pantic. How to distinguish posed from spontaneous smiles using geometric features. In *Proc. ICMI*, pages 38–45, 2007.

[16] Z. Zeng, M. Pantic, G. Roisman, and T. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE TPAMI*, pages 39–58, 2009.