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

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
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 expression 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 (e.g. 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.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM ’10, October 25–29, 2010, Firenze, Italy.
Copyright 2010 ACM 978-1-60558-933-6/10/10 ...$10.00.
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 crow’s 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-ocular 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, \( \beta_m \), to determine the amount of eye opening. \( \beta_m \) 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 \( \beta_m \) for left and right eyes separately.

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 separately. As indicated 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

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

(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 1: Feature points on different face regions.

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

Figure 3: Movement features of eyelids.
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 naive 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.
fier with standard 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 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>
<thead>
<tr>
<th>Classifier</th>
<th>Classified Class</th>
<th>Real Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHMM</td>
<td>8</td>
<td>38</td>
</tr>
<tr>
<td>1-NN</td>
<td>6</td>
<td>40</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>4</td>
<td>42</td>
</tr>
</tbody>
</table>

**6. REFERENCES**


