



Is gender encoded in the smile? A computational framework for the analysis of the smile driven dynamic face for gender recognition

Hassan Ugail¹ · Ahmad Al-dahoud¹

© The Author(s) 2018. This article is an open access publication

Abstract

Automatic gender classification has become a topic of great interest to the visual computing research community in recent times. This is due to the fact that computer-based automatic gender recognition has multiple applications including, but not limited to, face perception, age, ethnicity, identity analysis, video surveillance and smart human computer interaction. In this paper, we discuss a machine learning approach for efficient identification of gender purely from the dynamics of a person's smile. Thus, we show that the complex dynamics of a smile on someone's face bear much relation to the person's gender. To do this, we first formulate a computational framework that captures the dynamic characteristics of a smile. Our dynamic framework measures changes in the face during a smile using a set of spatial features on the overall face, the area of the mouth, the geometric flow around prominent parts of the face and a set of intrinsic features based on the dynamic geometry of the face. This enables us to extract 210 distinct dynamic smile parameters which form as the contributing features for machine learning. For machine classification, we have utilised both the Support Vector Machine and the k -Nearest Neighbour algorithms. To verify the accuracy of our approach, we have tested our algorithms on two databases, namely the CK+ and the MUG, consisting of a total of 109 subjects. As a result, using the k -NN algorithm, along with tenfold cross validation, for example, we achieve an accurate gender classification rate of over 85%. Hence, through the methodology we present here, we establish proof of the existence of strong indicators of gender dimorphism, purely in the dynamics of a person's smile.

Keywords Smile dynamics · Gender recognition · Machine learning · k -Nearest neighbour

1 Introduction

It is often said that the face is a window to the soul. Bearing a metaphor of this nature in mind, one might find it intriguing to understand, if any, how the physical, behavioural as well as emotional characteristics of a person could be decoded from the face itself. With the increasing deductive power of machine learning techniques, it is becoming plausible to address such questions through the development of appropriate computational frameworks.

Computational frameworks for human face analysis have recently found their way into great many application areas. These include computer vision, psychology, biometrics, security and even healthcare. The appealing, and the practical,

nature of such face analysis techniques, are highlighted by the wealth of information it can provide in a non-invasive manner. Unsurprisingly, such applications have already found their way into furnishing useful telltale signs of an individual's health status, identity, beauty and behaviour, all of which can be enhanced by the non-invasive information that leaks directly from the face, e.g. [1,12,27].

Additionally, computer-based analysis of the human face can provide strong and useful cues for personal attributes such as age, ethnicity and more appropriately gender, in the present context. Gender classification, in this sense, can, for example, aid as an advantageous biometric feature in order to improve the accuracy of determining an identity, especially in the presence of limited information on a subject. Recent research into gender classification has faced challenging hurdles, mainly due to the reliance of static data in the form of facial images. There are many inherent issues when looking for gender clues in appearance-based facial analysis. These include variability of lighting conditions, pose and occlusions. In this regard, in this work, we departed from

✉ Hassan Ugail
h.ugail@bradford.ac.uk

¹ Centre for Visual Computing, Faculty of Engineering and Informatics, University of Bradford, Bradford BD7 1DP, UK

such appearance-based analysis of facial images. Instead, we consider the analysis of the dynamic face, in particular, the dynamics of the smile, for clues of gender. This allows us to address the very intriguing question of whether a person's sexual dimorphism is encoded in the dynamics of the smile itself.

Hence, this paper is concerned with the identification of gender from the dynamic behaviour of the face. Equally importantly, we seek to answer the crucial question of whether gender is encoded in the dynamics of a person's smile. The case for such a computer-based investigation is fuelled by an array of cognitive physiological studies showing evidence of gender variances in facial expressions, e.g. [17,22,29]. We specifically focus on studying the smile as it is considered to be a rich, complex and sophisticated facial expression, formed through the synergistic action of emotions. According to Ekman [15], there are 18 different types of smile, each of them corresponds to a specific situation and reflects a different type of emotion. Moreover, various studies show that there are differences in smiles between males and females, i.e. females tend to bear more expressive smiles than males. Furthermore, recent research indicates that females express emotions more accurately in both spontaneous and posed situations, e.g. [7,8].

Based on the findings from such psychological studies, we examine the intensity and the duration of a smile in the hope of finding a distinction between the two sexes. Hence, in this paper, we present an algorithm to measure gender solely based on the dynamics of the smile without resorting to appearance-based image analysis techniques. The dynamic framework we have developed for smile analysis has four key components. They are the spatial dynamics of the face based on geometric distances across the entire face, dynamic triangular areas of the mouth, the geometric flow across key areas of the face and statistically inspired intrinsic features which further analyse the spatial and area parameters. These purely dynamic features are then fed to a machine learning routine for classification, resulting in an algorithm for gender recognition.

This paper is structured as follows. In Sect. 2, we discuss some of the related work on gender classification from two different viewpoints, i.e. psychological and computational standpoints. In that Section, we specifically highlight the distinctiveness of our work in this area. Section 3 presents our proposed framework for identifying the 210 unique dynamic parameters for analysis of a smile. In Sect. 4, we then explain how we have utilised the computational framework we have developed to undertake analysis of smiles for gender classification. We then provide details of results we have obtained, and finally in Sect. 5, we conclude this paper.

2 Related work

In this Section, we discuss recent advances in face analysis for both smiles and gender classification. Research in these areas appears to be dominantly arising from psychological studies as well as computer aided analysis of both static and dynamic digital images.

In many psychological experiments, the use of facial electromyographic (EMG) is common, especially for studies relating to the analysis of the face. EMG is a diagnostic technique used for recording facial muscle activity by placing electrodes on the face [31]. Much work on face analysis have been undertaken using EMG. These include the study of facial reactions to auditory stimuli, gender differences in facial reactions to facial expressions and facial and emotional reactions to both genuine and induced smiles, e.g. [14,30].

For example, in [13], it is reported that, based on the facial EMG activity, happy faces evoked increased zygomatic activity and the effects were more pronounced for females. Similarly, a number of other psychological studies show that females on average bear more expressive smiles. In fact, it has been documented that females smile more often than males in a variety of social contexts, e.g. [7,10].

From a computational viewpoint, gender classification based on the analysis of the face can be divided into three main categories, namely geometric, appearance and methods comprising of a hybrid between geometric and appearance models. All these methods rely on some form of a technique for extracting features from facial images.

The geometric model relies on the spatial geometry derived from a set of facial landmarks for extracting the relevant facial features. These often include physical measurements such as the size of eyebrows, eyes, nose and the mouth. Furthermore, in the dynamic case, this model relies on measures based on the movement of the landmarks. For example, the work presented in [18] relies on the use of geometric features obtained from standard Canny edge detection. The result is a global feature set consisting of the inter ocular distance, the distance between the lips and the nose tip as well as the distance between the nose tip and the line joining the two eyes. The resulting classification was rather simple and was based on threshold values of Euclidean measures. In [19], a slightly more sophisticated model based on principle component analysis (PCA) was utilised while the classification was again based on Euclidean measures. It is claimed that they achieve an impressive gender recognition rate of 94% simply based on the geometric measures on still images.

The appearance model can also be referred to as template matching or using an exemplar of the object. Appearance models imply that objects look different under changes in lighting, colour, direction and can be at different scales.

Furthermore, it describes the texture of the facial features. The work presented in [21,26] utilises the Gabor function to extract the relevant texture from which facial features are derived. The Gabor function uses a set of wavelets with specific orientations and directions in order to represent a given texture. It is computationally intensive and so hardly applicable for real-time applications. A trick often employed to speed up the computations is to use PCA or local binary pattern (LBP) to minimise the number of features extracted. More sophisticated methods such as the use of convolutional neural networks for appearance-based gender recognition from facial images are also increasingly becoming popular, e.g. [5].

The hybrid model which takes advantage of both the appearance and geometric parameters uses a variety of techniques. The work presented in [33] makes use of a hybrid model consisting of active appearance model (AAM) and Haar wavelets for gender classification from faces. In their model, a total of 3403 geometric distances were chosen from which 10 significant features were extracted. They are then fed to an SVM classifier. The work described in [28] utilises case-based reasoning (CBR) whereby 8 geometric distances across the face along with 6 different facial feature ratios were used to train a hybrid of k -NN and Linear discriminant analysis (LDA) classifier. In [25], the use of discrete cosine transform (DCT) and LBP algorithms as appearance models and geometric distance features are described. Similarly, the work described in [24] divides the face into regions of interest to apply PCA and utilise SVM for classification.

It is entirely appropriate for us to place the work presented in this paper in the category of geometric models. However, the distinct difference between our work and the rest is that we concentrate purely on the dynamic facial features and more specifically the dynamics arising from the smile. The closest to our work in the present literature is the interesting piece of research in the area recently presented in [11] in which the use of smile for gender classification has been

discussed. Their framework makes use of 49 facial landmarks produced by the cascade of linear regression, and they track them using sparse optical flow, which is used to measure 27 geometric distances across the face. For classification, they have used a pattern classifier on labelled data with an SVM. At a superficial level, it might appear that this work is rather similar to ours, though upon closer examination one would uncover distinct differences in that they have in fact utilised an appearance-based model along with smile dynamics to enhance the gender classification. On the other hand, as highlighted earlier, in our work, we resort to the pure dynamic features of the smile, and hence, we present an analysis framework to solely process the dynamics of the face for gender identification.

3 A computational framework for smile dynamics

It has been hypothesised and evidenced by various psychological experiments that there exist differences in smiles between the two genders. To verify this computationally and at the same time to develop a tool for gender classification solely based on the smiles, we propose a framework which can track the dynamic variations in the face from neutral to the peak of a smile. Our framework is based upon four key components. They are (1) the spatial features which are based on dynamic geometric distances on the overall face, (2) the changes that occur in the area of the mouth, (3) the geometric flow around prominent parts of the face and (4) a set of intrinsic features based on the dynamic geometry of the face. Note, all of the dynamic features described here are intuitive extensions of the relevant physical experimentations and are based on the reported literature on facial emotions, especially on the dynamics of the smile, eg. [8,10].

Figure 1 presents a block diagram showing the key components of our framework for the analysis of the dynamics

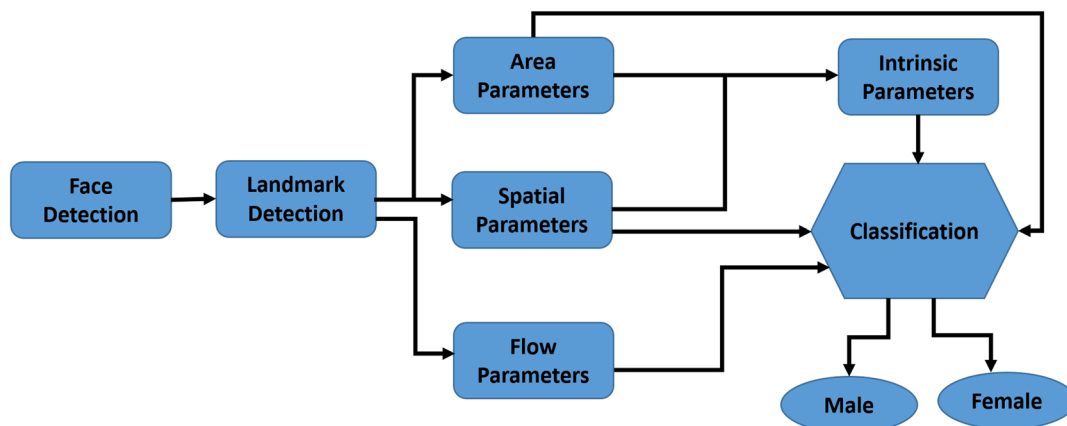


Fig. 1 Block diagram showing the key components of the computational framework for automatic analysis of smile dynamics

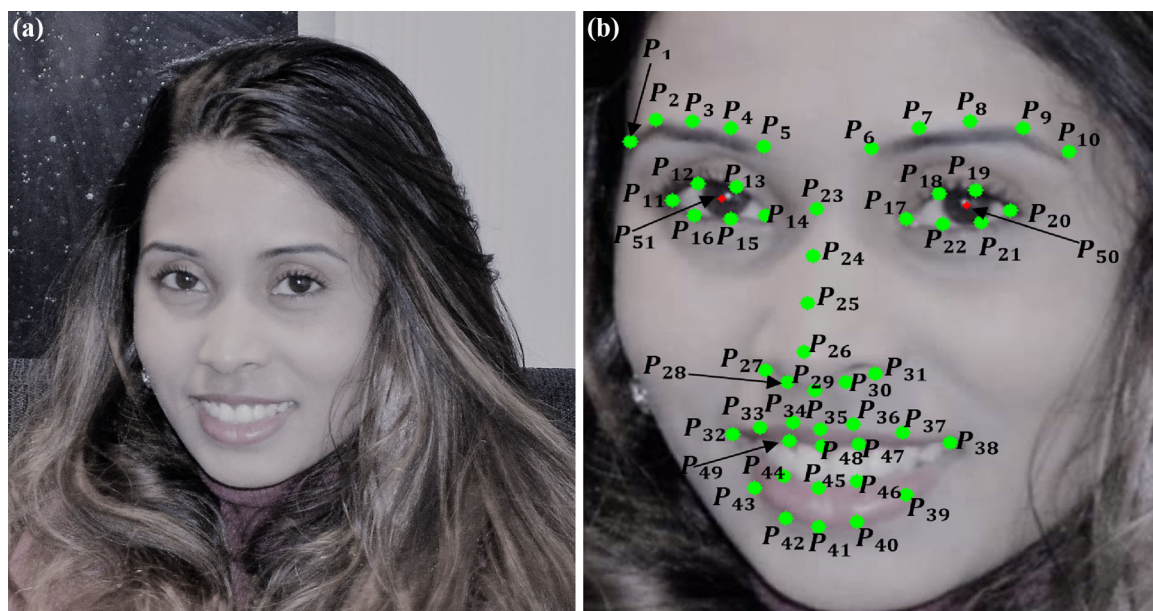


Fig. 2 Automatic landmark detection. **a** An example input face, **b** landmarks detected using the CHEHRA model

of a smile. The first step in our framework is to detect and track the face within a given video sequence. To do this, we have used a well-known Viola-Jones algorithm. It is based on Haar feature selection to create an integral image through the use of Adaboost training and cascade classifiers [32]. The ability of this algorithm to robustly detect faces under different lighting conditions is well established, and we have also demonstrated this in previous work [2].

The next step in our proposed framework involves automatic detection and tracking of a stable set of landmarks on the dynamic face. Automated Landmark detection is done using the CHEHRA model [6]. The algorithm has been trained to detect facial landmarks using in-the-wild datasets under various illumination, facial expressions and head poses. It is based on cascade linear regression for discriminative face alignment. This is done by applying Incremental Parallel Cascade of Linear Regression (iPar-CLR) method. The tests we have carried using the CHEHRA model appear to be acceptable though we noticed it is likely to suffer when it comes to real-time applications. The algorithm has been utilised to detect 49 landmarks on the face, marked as $P_1 \dots P_{49}$ as shown in Fig. 2b for the face shown in Fig. 2a. Note, in addition to the 49 landmarks which CHEHRA detects, we also include the centres of the eyes as two additional landmark points, as shown in Fig. 2b, marked as P_{50} and P_{51} .

3.1 Dynamics of the spatial parameters

Based on some of the positions of the 49 landmarks we obtain through the CHEHRA model, we identify 6 dynamic geomet-

Table 1 Description of the geometric distances from which dynamic spatial parameters are derived

Distance	Description	Landmarks
d_1	Mouth corners	P_{32} to P_{38}
d_2	Upper and lower lip	P_{45} to P_{48}
d_3	Mouth to nose (left corners)	P_{32} to P_{27}
d_4	Mouth to nose (right corners)	P_{38} to P_{31}
d_5	Mouth to eye (left corners)	P_{32} to P_{11}
d_6	Mouth to eye (right corners)	P_{38} to P_{20}

ric Euclidean distances across the face which are utilised to compute our dynamic spatial parameters. Further details of these spatial parameters are given in Table 1. The general form of a given spatial parameter is,

$$\delta d_i = \frac{d_i}{N_i} + \sum_{n=1}^t \frac{d_i}{N_i} - \frac{d_{in}}{N_{in}}, \quad (1)$$

where t is the total number of video frames corresponding to each $\frac{1}{10}$ th increment of the total time T for the smile, from neutral to the peak. Here N_i is the length of the nose, for a given video frame, computed as the distance between P_{23} and P_{26} . Thus, by dividing the spatial parameters by the length of the nose N_i , we normalise these parameters to the given dynamic facial image. It is noteworthy to point out that for a given smile, from neutral to the peak, we divide the time it takes into ten partitions and therefore for each of the d_i we have 10 times d_i parameters which are fed to the machine

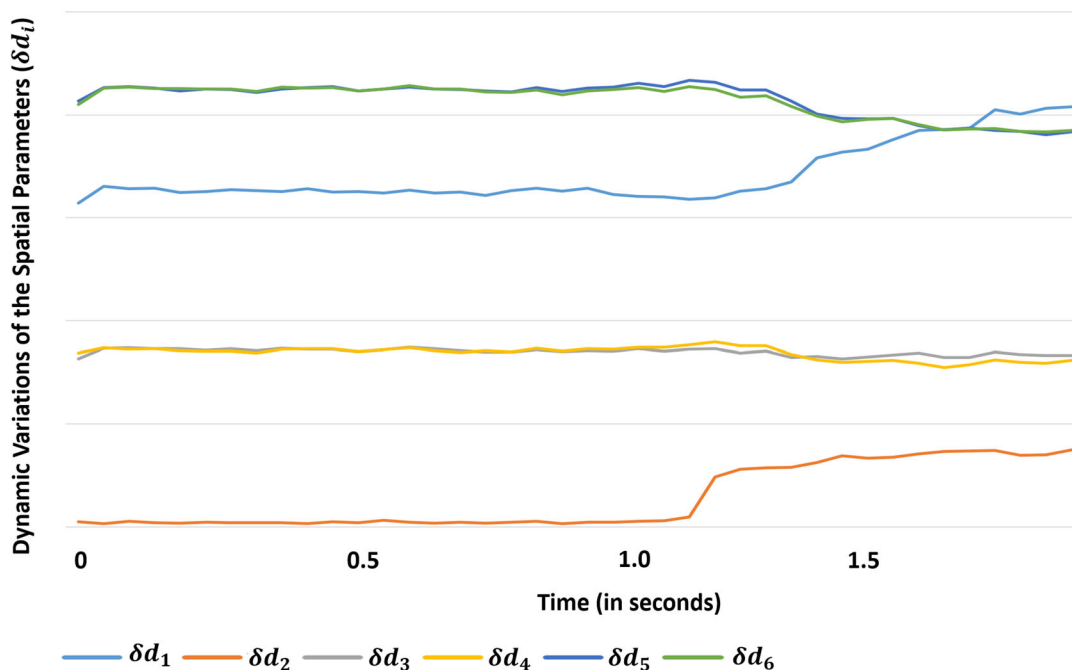


Fig. 3 Variation in the dynamic spatial parameters δd_i across the 10 partitions of time, for a typical smile, from neutral to the peak

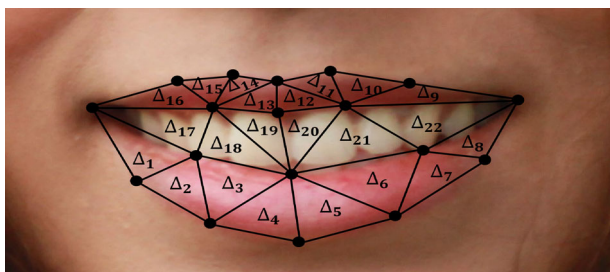


Fig. 4 Description of triangular mouth areas used to form the dynamic area parameters on the mouth

learning. Hence, in our dynamic smile framework, we have a total of 60 dynamic spatial parameters.

Figure 3 shows the variation of δd_i across the 10 time partitions for a typical smile. As can be observed, there is a variation in each parameter as the smile proceeds from neutral to its peak.

3.2 Dynamic area parameters on the mouth

The second set of dynamic parameters concern the mouth. Here we compute the changes in the area of 22 triangular regions that occupy the total area of the mouth. This is shown in Fig. 4. Again these areas are computed using the corresponding landmarks obtained from the CHEHRA model. The general form of how the changes in the mouth area are computed is described as,

$$\Delta_{\text{area}}^i = \sum_{n=1}^{22} \frac{\Delta_i}{\Delta N_i}, \tag{2}$$

and,

$$\delta \Delta_i = \sum_{n=1}^t \Delta_{\text{area}}^i, \tag{3}$$

where t is the total number of video frames corresponding to each $\frac{1}{10}$ th increment of the total time T for the smile, from neutral to the peak. Here ΔN_i is the invariant triangle area determined by the landmarks defining the outer corners of the eyes and the tip of the nose, namely P_{11}, P_{20} and P_{26} . Again we divide the total time of the smile, from neutral to peak, into ten partitions, and therefore we obtain 10 parameters from the $\delta \Delta_i$, though time, which are fed to the machine learning. Thus, in our dynamic smile framework, we have a total of 10 parameters which capture dynamics of the mouth.

For the purpose of illustration, in Fig. 5 we show the distribution of areas of the triangular regions, Δ_i , across a typical smile.

3.3 Dynamic geometric flow parameters

The third component of our framework for smile dynamics is the computation of flow around the face during a smile. More specifically, we compute the flow around the mouth, cheeks and around the eyes. To do this, we have utilised the dense optical flow developed by Farneback [16]. It is a two-frame

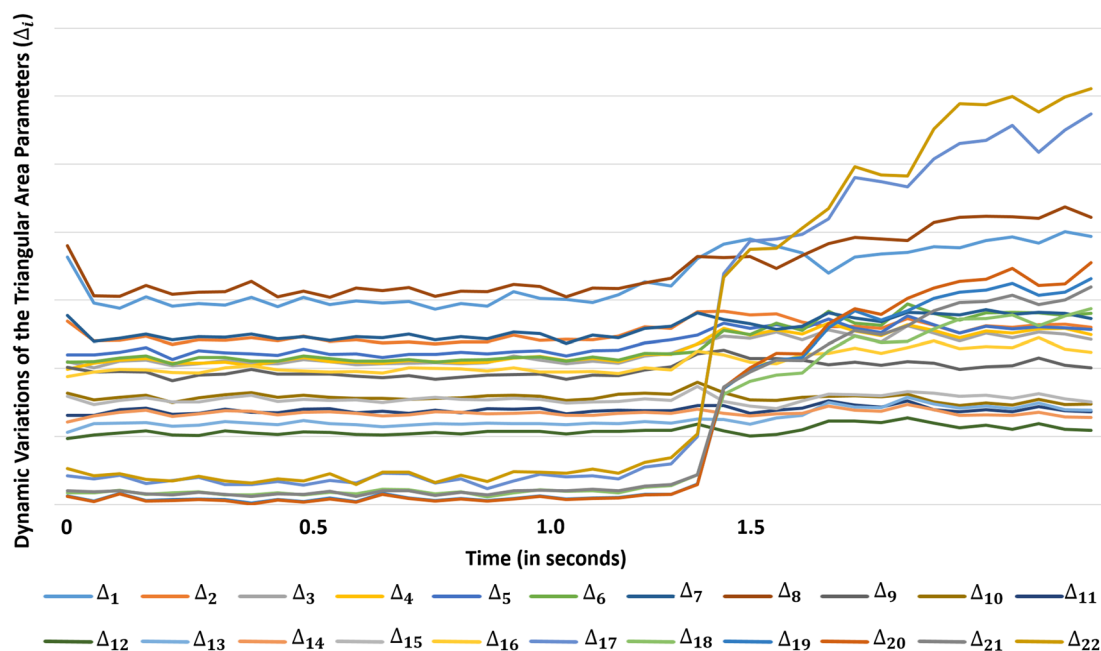


Fig. 5 Variation in the dynamic area parameters Δ_i on the mouth, across the 10 partitions of time, for a typical smile, from neutral to the peak

Table 2 Description of how the optical flow parameters around the face are derived

Optical flow	Description	Landmarks/regions
δf_1	Mouth	Landmarks P_{32} to P_{49}
δf_2	Left eye	f_6, f_7, f_8, f_9
δf_3	Right eye	f_1, f_2, f_3, f_4
δf_4	Left cheek	f_{10}
δf_5	Right cheek	f_5

motion estimation algorithm in which quadratic polynomials are used to approximate the motion between two subsequent frames in order to approximate motion of neighbourhood pixels for both the frames. Using this algorithm, we are able to estimate the successive displacement of each of the landmarks during the smile.

Table 2 shows how the various landmarks and regions of the face are utilised to compute the optical flows around the face. The relevant facial regions and landmarks are given in Figs. 2b and 6 respectively. We also show the variations in the dynamic optical flows, δf_i , around the face for a typical smile in Fig. 7.

Note, the geometric flow, δf_i , for each of the regions is normalised upon computation by means of the corresponding flow around the invariant triangle area of the face determined by the landmarks defining the outer corners of the eyes and the tip of the nose, namely P_{11}, P_{20} and P_{26} . Again, each of the geometric flow parameters δf_i is computed across the 10 time interval through which the smile is measured, resulting

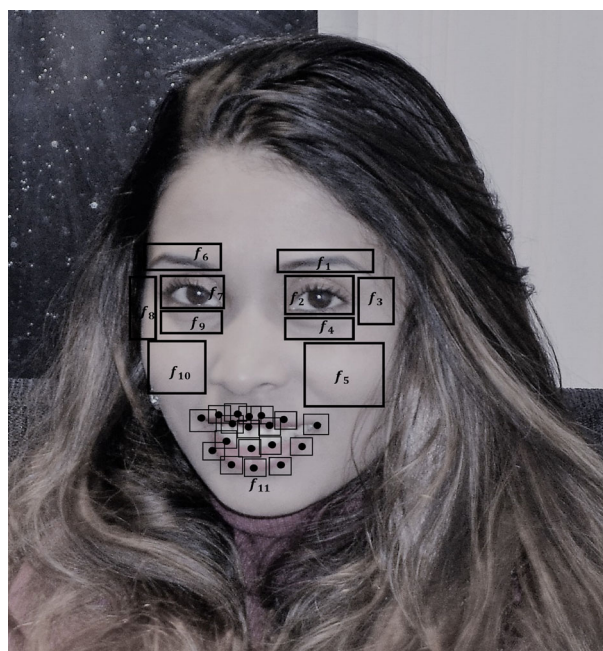


Fig. 6 Regions of the face identified for dynamic optical flow computation

in a total of 50 dynamic geometric flow parameters which are then fed to machine learning.

3.4 Intrinsic dynamic parameters

In addition to the spatial parameters, the area parameters and geometric flow parameters, we compute a family of intrinsic

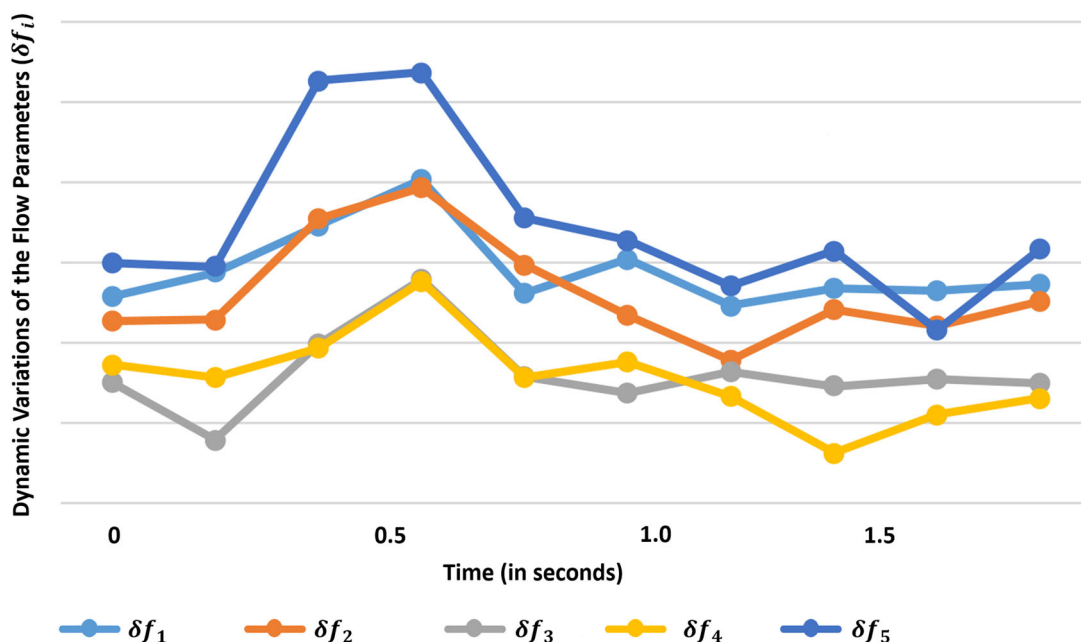


Fig. 7 Variations in the dynamic optical flows δf_i around the face, for a typical smile, from neutral to the peak

dynamic parameters on the face to further enhance the analysis of the dynamics of the smile. These intrinsic parameters are mainly based on the computation of the variations in the slopes and the growth rates of various features across the face. We identify these features as s_1, s_2, s_3 and s_4 , details of which we describe as follows.

The first parameter family in this category relates to the computation of the overall slope variation around the mouth during a smile. To compute this, we use,

$$s_{1i} = \frac{N \sum_{n=1}^N P_{ix} P_{iy} - \sum_{n=1}^N P_{ix} \sum_{n=1}^N P_{iy}}{\sum_{n=1}^N P_{ix}^2 - \left(\sum_{n=1}^N P_{ix}\right)^2}, \tag{4}$$

where N is the number of video frames comprising the whole smile, from neutral to the peak, P_{ix} and P_{iy} are the Cartesian coordinate equivalents in the image space corresponding to the landmark point P_i . Hence, a total of 12 parameters are identified for the variations in slopes around mouth corresponding to the mouth landmarks P_{32} to P_{43} .

The second family of parameters, s_2 , in this category corresponds to the growth rates across smile corresponding to the spatial parameters as well as area parameters on the mouth. The growth rates arising from the spatial parameters are defined as,

$$s_{2i(\text{spatial})} = \sum_{n=1}^N \frac{\delta d_i^t - \delta d_i^{t+1}}{\delta d_i^t}, \tag{5}$$

and for the area parameters on the mouth are,

$$s_{2i(\text{area})} = \sum_{n=1}^N \frac{\Delta_i^t - \Delta_i^{t+1}}{\Delta_i^t}, \tag{6}$$

where N is identified as the total number of frames in the video sequence of the smile while t to $t + 1$ defines two successive video frames. In addition to the growth rates $s_{2i(\text{area})}$, for each of the 22 triangular regions of the mouth, we also compute the total growth rate for the mouth, by using Eq. (6) along with the 22 triangular mouth area information. This means we have a total of $6 + 22 + 1 = 29$ parameters of dynamic intrinsic type s_2 .

The third family of parameters, s_3 , in this category we have identified is for both spatial parameters across the face and area parameters in the mouth. These are defined as compound growth rates given as,

$$s_{3i(\text{spatial})} = \left(\frac{\delta d_i^{\text{neutral}}}{\delta d_i^{\text{peak}}}\right)^{1/N} - 1, \tag{7}$$

and,

$$s_{3i(\text{area})} = \left(\frac{\Delta_i^{\text{neutral}}}{\Delta_i^{\text{peak}}}\right)^{1/N} - 1, \tag{8}$$

where N , like previously, is the total number of frames in the video sequence of the smile. The compound growth rate is measured simply using the neutral and peak of the smile. Again, like previously, in addition to the compound growth rates $s_{3i(\text{area})}$ we also compute the compound growth for the entire mouth by means of the utilising the total area of the

Table 3 Parameter description for the computational framework for smile dynamics

Parameter	Description	Number of parameters
δd_i	Spatial—involving 6 geometric distances across the face	60
$\delta \Delta_i$	Mouth area—derived from the total area for the 22 parts of the mouth	10
δf_i	Geometric flow around the mouth, eyes and cheeks	50
s_{1i}	Slope measurements around mouth landmarks P_{32} to P_{43}	12
s_{2i}	Growth rates of the spatial parameters and mouth areas	29
s_{3i}	Compound growth rates of the spatial parameters and mouth areas	29
s_{4i}	Gradient orientations for the mouth corners and the mouth area	20

mouth. This implies that we obtain a total of 29 parameters of dynamic intrinsic type s_3 too.

For the final family of parameters, s_4 , in this category, we compute the gradient orientation of the mouth based on the two mouth corner landmarks P_{32} and P_{38} which provides us with a line m passing δd_1 at the neutral and the peak of the smile. We then use,

$$s_{4i} = \sum_{t=1}^T \delta d_1^t - m^t, \quad (9)$$

to compute the rate of deviation of the mouth corners against the gradient m over the 10 time partitions where T is the total time from neutral to the peak of the smile. Similarly, we compute the gradient orientation of the mouth area based on the combined 22 triangular areas of the mouth between the neutral frame and the peak of the smile.

These parameters provide us with a sense of the *smoothness* of the smile and forms an additional $10 + 10 = 20$ parameters for machine learning.

Table 3 provides a summary and brief description of various parameters associated with our computational framework for smile dynamics.

4 Experiments

Once an appropriate framework for the analysis of the dynamics of the smiles, as described above, is in place, we carried out a series of experiments to further analyse the pattern of smile and more importantly to look for clues of gender in the smile. For this purpose, we utilised two well-known datasets to carry out an initial set of experiments. We then utilised the same datasets to extract the parameters described in Table 3 and fed them to machine learning routines.

4.1 Datasets

We tested our approach on two main datasets namely, the CK+ [23] and the MUG [4] datasets. The CK+ dataset has a

total of 83 subjects, consisting of 56 females and 27 males. The smile of each of the subjects went from the neutral expression to the peak of the smile. On the other hand, the MUG dataset contains a total of 26 subjects, consisting of 13 females and 13 males. The smile of each subject, in this case, went from the neutral expression through to the peak and finally returning to the neutral. Since our framework has been developed to analyse smiles from neutral to the peak, we modified the MUG dataset so as it only contained the relevant parts of the smile for each subject. In addition to this, for each smile, we also ensured that within the two datasets there contained an equal number of video frames. Thus, a total of 109 unique subjects were available to us for training and testing.

4.2 Initial experiments

Here we report an initial set of interesting experiments that we undertook to further understand the dynamics of smiles and to seek for clues of gender in smiles.

In our first experiment, we tried a rather brute force approach to identify the areas of the face that contain most information of the smile that relates to the gender. Figure 8 shows some results based on the changes in the areas of the mouth region for 54 subjects (27 females and 27 males) in the CK+ database for the peak of the smile. As can be observed, there appears to be no significant difference between genders when one considers this simple form of analysis.

Next, we considered a computation relating to the product value for the areas for the upper lip and lower lip using the $\delta \Delta_i$ given in Equation (2) throughout the smile expression. This was done by multiplying each feature value for the changes in the mouth areas from a given video frame by the corresponding values from the next frame, in order to obtain the product of the features (POF) through the smile expression, i.e.

$$\text{POF}_i = \sum_{t=1}^N \Delta_i^t \Delta_i^{t+1}, \quad (10)$$

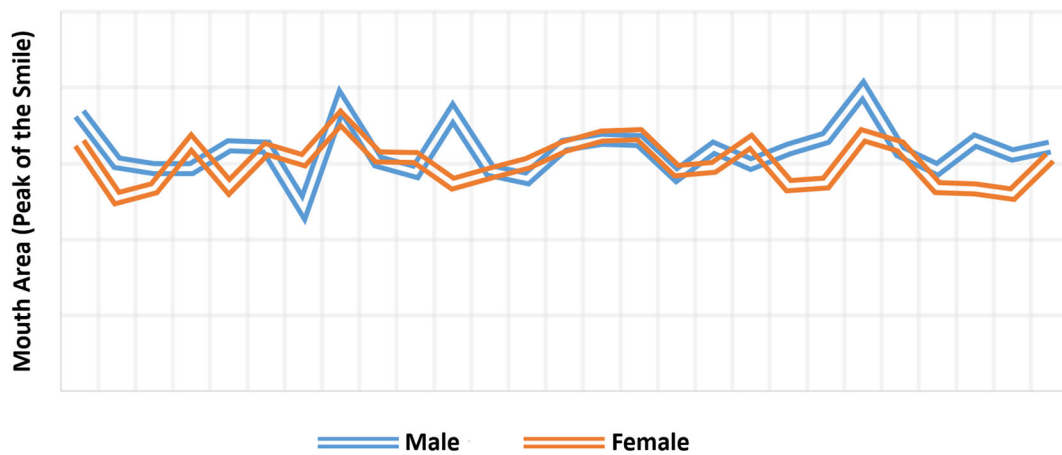


Fig. 8 Variations in the area of the mouth at the peak of the smile for 54 subjects in CK+ database

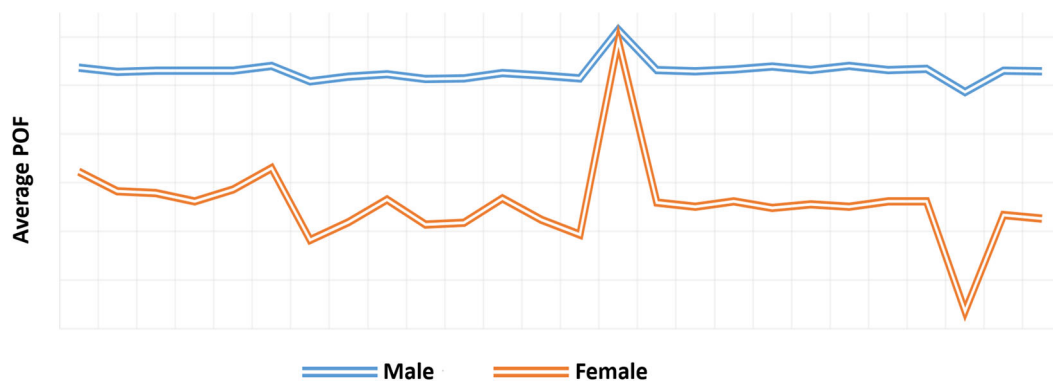


Fig. 9 Average POF plots for the 54 subjects in the CK+ dataset

where N is the number of video frames containing the smile expression from neutral to the peak.

Analysis of the POF data gave us some clues for gender difference between the smiles. Thus, by taking the average of each attribute, for both the genders, Fig. 9 shows the POF for the females and the males in the CK+ dataset and Fig. 10 shows the corresponding POF plots for the subjects in the MUG dataset. As can be inferred, from the average POF results shown in Figs. 9 and 10, there appears to be a distinguishable difference in the smile of the males and females in terms of the POF computations.

Furthermore, from a first glance at these results, one might infer that males have a more intense smile than females which directly conflicts with the various psychological studies. However, that is indeed not the case. In fact, we note that in this experiment we computed the POF for each triangular features whose values are always less than 1. Additionally, for normalisation, we divided the POF values with the invariant area of the eyes-nose triangle. The result is a very small number, less than 1. Since the product of smaller num-

bers is smaller too, the POF values for females are smaller than that for males. Hence, it indeed confirms the smiles of females expand more through time in comparison to males.

Though rather simple, using the above approach, we were able to classify the data, through the median POF value computed from the mouth triangular attributes. This lends us a 60% correct classification for gender. That, however, is just slightly above chance and hence would not be considered an acceptable method of classification. We then used all the features described in our computational framework for smile dynamics (Sect. 3) to train and test a machine learning classifier.

With this knowledge, we constructed a suitable machine learning approach for classification. Following the procedure described in the block diagram shown in Fig. 1, the following algorithmic form further elaborates the computational framework we have developed for the automatic analysis of smile dynamics.

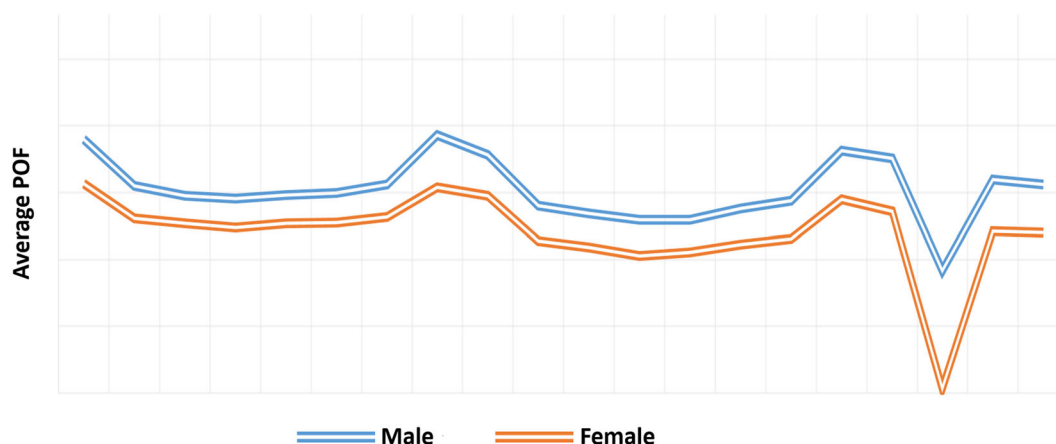


Fig. 10 Average POF plots for the 26 subjects in the MUG dataset

Algorithm 1 Gender from the dynamics of a smile.

M_j = Matrix of features in the smile
 N_i = Sequence of video frames in the smile
 H_k = Matrix of computed features from each video frame

while $k \neq N_i$ **do**

 Detect the face
 Detect the landmarks, P_k
 Compute mouth areas, Δ_k
 Compute geometric distances, d_k
 Compute geometric flows, f_k
 Populate H_k
 $k=k+1$

end while

 Compute spatial parameter values, $H_k \rightarrow \delta d_j$
 Compute area parameter values, $H_k \rightarrow \delta \Delta_j$
 Compute flow parameter values, $H_k \rightarrow \delta f_j$
 Compute intrinsic parameter values, s_1, s_2, s_3 and s_4

 Form data clusters $M_j \rightarrow D_n, n = 10$
 Apply classifier to D_n
 Output gender

4.3 Classification using machine learning

For our machine learning based classification, we have utilised two well-known classification algorithms namely, the support vector machine (SVM) and the k -nearest neighbour (KNN).

First, we tried to use PCA as a pre-step before applying SVM. The results indicated this approach yields a very low classification rate. This is probably due to the fact that PCA reduces the number of features, while at the same time eliminating some distinguishing features. Second, we used SVM on its own, without the PCA. We had a mild improvement in the classification rate of 69%.

Finally, we used the k -NN algorithm which is a nonparametric method used for classification and regression [3]. The

Table 4 Results using the k -NN classification

	CK+	MUG
k -NN distance	Correlation	Cosine
k value	3	14
Classification	78%	86%

output of k -NN algorithm is a class relationship. The object can be assigned a class by k nearest neighbours where k is a positive integer. We utilised all the 210 features described in Table 3 to train our classifier. Additionally, we used a ten-fold cross validation scheme to validate our k -NN classifier. The results were tested on several distance functions namely, Euclidean, Cosine, Minkowsky and Correlation. In Table 4, we report the best results we have obtained using the k -NN classifier.

5 Conclusions

This paper is concerned with the identification of gender from the dynamic behaviour of the face. In this sense, we wanted to answer the crucial question of whether gender is encoded in the dynamics of a person's smile. To do this, we have developed a computational framework which can analyse the dynamic variations of the face from the neutral pose to the peak of the smile. Our framework is based upon four key components. They are the spatial features which are based on dynamic geometric distances on the overall face, the changes that occur in the area of the mouth, the geometric flow around some of the prominent parts of the face and a set of intrinsic features based on the dynamic geometry of the face. This dynamic framework enables us to compute 210 unique features which can then be fed to a k -NN classifier for gender recognition.

We ran our experiments on a total of 109 subjects (69 females and 40 males) from two datasets, namely the CK+ and the MUG datasets. Firstly, our results do agree with that of various psychological studies, indicating that females are more expressive in their smiles. For example, this became evident to us by simply looking at the changes in the lip area during a smile in which the lip area of female subjects expands more in comparison with the male subjects. Further, and more importantly, using machine learning approaches, we can also classify gender from smiles. In particular, by means of the standard k -NN algorithm, we are able to obtain a classification rate of up to 86%, purely based on the dynamics of smiles.

We understand from the presently available literature that some of the recent work carried in gender classification can achieve over 90% recognition rates using hybrid models with a combination of geometric and appearance features which are both static and dynamic. This is particularly clear from the work presented in [11]. It is, however, noteworthy that our work is geared to study the gender classification rates purely based on the dynamics of a smile. In fact, some of the results reported in [11] indicate that using their chosen dynamic smile features they obtain a classification rate of 60%, whereas using the smile dynamics framework we have proposed, we are able to obtain a higher gender classification rate of over 75%. There is also an added advantage of using the dynamic features, as opposed to static images, for gender identification since it presents with the opportunity to infer gender from certain parts of the face such as the mouth and the eyes areas.

Going forward into the future, there are a number of directions in which this work can be further taken forward. It will be useful to see if it would be possible to enhance the classification rates using other correlation and sophisticated statistical analysis techniques. In this paper, we have only used simple machine classification techniques such as SVM and K -NN, since our prime aim here was to demonstrate the power of smile dynamics in gender identification. We believe the utilisation of sophisticated machine learning techniques will further improve the results. We also believe this will be the case if novel machine learning techniques such as convolutional neural networks based deep learning (eg. [9,20]) can be adopted to the problem at hand. However, having said that, we must also highlight the fact that such sophisticated machine learning techniques usually require sufficient and significant training data which, as far as smiles are concerned, are scarce at present.

In addition to this, the results could be further tested and validated on other datasets. One deficiency of this present study is that we did not look deeply into the gender variation between posed and spontaneous smiles. We believe our framework has merit in providing much room for such detailed analysis to seek gender differences between the two

types of smiles. Additionally, aside from the expression of a smile, other basic emotional attributes such as surprise, fear, anger and disgust can be studied in detail to look for cues to enhance gender recognition from facial expressions in general. We believe the framework we have presented in this paper can easily be adapted to undertake such studies.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

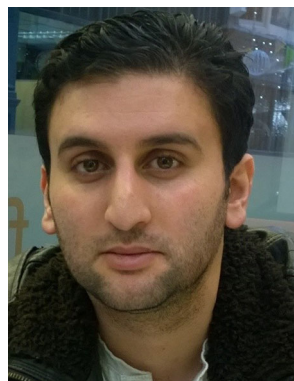
References

1. Abel, E.L., Kruger, M.L.: Smile intensity in photographs predicts longevity. *Psychol. Sci.* **21**(4), 542–544 (2010)
2. Al-dahoud, A., Ugail, H.: A method for location based search for enhancing facial feature detection. In: The proceedings of the international conference on advances in computational intelligence systems, AISC 2016, pp. 421–432 (2016)
3. Altman, N.S.: An introduction to kernel and nearest-neighbor non-parametric regression. *Am. Stat.* **46**(3), 175–185 (1992)
4. Aifanti, N., Papachristou, C., Delopoulos A.: The MUG facial expression database. In: The 11th international workshop on image analysis for multimedia interactive services, WIAMIS 2010, pp. 1–4. IEEE (2010)
5. Antipov, G., Baccouch, M., Berrani, S.-A., Dugelay, J.-L.: Effective training of convolutional neural networks for face-based gender and age prediction. *Pattern Recognit.* **72**(C), 15–26 (2017)
6. Asthana, A., Zafeiriou, S., Cheng, S., Pantic, M.: Incremental face alignment in the wild. In: Proceedings of the IEEE conference on computer vision and pattern recognition CVPR 2014, pp. 1859–1866 (2014)
7. Briton, N.J., Hall, J.A.: Gender-based expectancies and observer judgements of smiling. *J. Nonverbal Behav.* **19**(1), 49–65 (1995)
8. Brody, L.R., Hall, J.A., Stokes, L.R.: Gender and emotion: theory, findings, and content. In: Barrett, L.F., Lewis, M., Haviland-Jones, J.M. (eds.) *Handbook of Emotions*, 4th edn, pp. 369–392. The Guildford Press, New York (2016)
9. Bukar, A.M., Ugail, H.: Automatic age estimation from facial profile view. *IET Comput. Vis.* **11**(8), 650–655 (2017)
10. Cashdan, E.: Smiles, speech, and body posture: how women and men display sociometric status and power. *J. Nonverbal Behav.* **22**(4), 209–228 (1998)
11. Dantcheva, A., Brémond, F.: Gender estimation based on smile-dynamics. *IEEE Trans. Inf. Forensics Secur.* **12**(3), 719–729 (2017)
12. Deutsch, F.M., LeBaron, D., Fryer, M.M.: What is in a smile? *Psychol. Women Q.* **11**(3), 341–352 (1987)
13. Dimberg, U., Lundquist, L.O.: Gender differences in facial reactions to facial expressions. *Biol. Psychol.* **30**(2), 151–159 (1990)
14. Dimberg, U.: Facial electromyographic reactions and autonomic activity to auditory stimuli. *Biol. Psychol.* **31**(2), 137–147 (1990)
15. Ekman, P.: *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage* (revised edition). WW Norton & Company, New York (2009)
16. Farneback, G.: Two-frame motion estimation based on polynomial expansion. In: Proceedings of the Scandinavian conference on image analysis, SCIA 2003, pp. 363–370 (2003)

17. Hess, U., Adams Jr., R.B., Kleck, R.E.: Facial appearance, gender, and emotion expression. *Emotion* **4**(4), 378–388 (2004)
18. Kalam, S., Guttikonda, G.: Gender classification using geometric facial features. *Int. J. Comput. Appl.* **85**(7), 32–37 (2014)
19. Lale, D.P., Karande, K.J.: Gender classification using facial features. *Int. J. Adv. Res. Electron. Commun. Eng. (IJARECE)* **5**(9), 2227–2231 (2016)
20. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**, 436444 (2015)
21. Lian, H.C., Lu, B.L.: Multi-view gender classification using local binary patterns and support vector machines. In: Wang, J., Yi, Z., Zurada, J.M., Lu, B.L., Yin, H. (eds.) *Advances in Neural Networks-ISNN 2006*, pp. 202–209. Springer, Berlin (2006)
22. Liébart, M.F., Fouque-Deruelle, C., Santini, A., Dillier, F.L., Monnet-Corti, V., Glise, J.M., Borghetti, A.: Smile line and periodontium visibility. *Periodontol Pract. Today* **1**(1), 17–25 (2004)
23. Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The extended Cohn–Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression. In: *IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW) 2010*, pp. 94–101, IEEE (2010)
24. Lu, L., Shi, P.: A novel fusion-based method for expression-invariant gender classification. In: *Proceedings of IEEE international conference on acoustics, speech and signal processing, ICASSP 2009*, pp. 1065–1068, IEEE (2009)
25. Mozaffari, S., Behravan, H., Akbari, R.: Gender classification using single frontal image per person: combination of appearance and geometric based features. In: *Proceedings of the 20th international conference on pattern recognition, ICPR 2010*, pp. 1192–1195. IEEE (2010)
26. Rai, P., Khanna, P.: Appearance based gender classification with PCA and $(2D)^2$ PCA on approximation face image. In: *Proceedings of the 9th international conference on industrial and information systems (ICIIS) 2014*, pp. 1–6. IEEE (2014)
27. Rubin, L.R.: The anatomy of a smile: its importance in the treatment of facial paralysis. *Plast. Reconstr. Surg.* **53**(4), 384–387 (1974)
28. Saleem, M.A., Tamoor, M., Asif, S.: An efficient method for gender classification using hybrid CBR. In: *Proceedings of future technologies conference (FTC) 2016*, pp 116–120. IEEE (2016)
29. Simon, R.W., Nath, L.E.: Gender and emotion in the United States: Do men and women differ in self-reports of feelings and expressive behavior? *Am. J. Sociol.* **109**(5), 1137–1176 (2004)
30. Surakka, V., Hietanen, J.K.: Facial and emotional reactions to Duchenne and non-Duchenne smiles. *Int. J. Psychophysiol.* **29**(1), 23–33 (1998)
31. Van Boxtel, A.: Facial EMG as a tool for inferring affective states. In: *Proceedings of measuring behavior, 2010*, pp. 104–108. Noldus Information Technology, Wageningen (2010)
32. Viola, P., Jones, M.J.: Robust real-time face detection. *Int. J. Comput. Vis.* **57**(2), 137–154 (2004)
33. Xu, Z., Lu, L., Shi, P.: A hybrid approach to gender classification from face images. In: *Proceedings of the 19th international conference pattern recognition, ICPR 2008*, pp. 1–4. IEEE (2008)



Hassan Ugail is the Director of the Centre for Visual Computing at the Faculty of Engineering and Informatics at University of Bradford. He has a first-class BSc Honours degree in Mathematics from King's College, London, and a PhD in the field of Geometric Design from the School of Mathematics at University of Leeds. Professor Ugail's research interests include computer-based geometric and functional design, and imaging and machine learning.



Ahmad Al-dahoud is a research student at the Centre for Visual Computing at the Faculty of Engineering and Informatics at University of Bradford. He received a BSc degree in computer information system and an MSc degree in computer science both from Jordan University of Science and Technology. His research interests are in image and video processing, pattern recognition, and machine learning.