A Genuine Smile is Indeed in the Eyes - the Computer Aided Non-invasive Analysis of the Exact Weight Distribution of Human Smiles Across the Face

Hassan Ugail*, Ahmad Al-dahoud

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

Abstract

Understanding the detailed differences between posed and spontaneous smiles is an important topic with a range of applications such as in human-computer interaction, automatic facial emotion analysis and in awareness systems. During the past decade or so, there have been very promising solutions for accurate automatic recognition and detailed facial emotion analysis. To this end, many methods and techniques have been proposed for distinguishing between spontaneous and posed smiles. Our aim here is to go beyond the present state of the art in this field. Hence, in this work, we are concerned with understanding the exact distribution of a smile - both spontaneous and posed - across the face. To do this, we utilise a lightweight computational framework which we have developed to analyse the dynamics of human facial expressions. We utilise this framework to undertake a detailed study of the smile expression. Based on computing the optical flow across the face - especially across key parts of the face such as the mouth, the cheeks and around the eyes - we are able to accurately map the dynamic weight distribution of the smile expression. To validate our computational model, we utilise two publicly available datasets, namely the CK+ dataset in which the subjects express posed smiles and the MUG dataset in which the subjects express genuine smiles. Our results not only confirm what

*Corresponding author
Email address: h.ugail@bradford.ac.uk (Hassan Ugail)

already exists in the literature - i.e. that the spontaneous genuine smile is truly in the eyes - but it also gives further insight into the exact distribution of the smile across the face.
Keywords: Smile Analysis, Genuine Smiles, Posed Smiles, Smile Weight Distribution

---

1. Introduction

A smile is the most common facial expression that occurs among humans during our daily lives. Studies show that the smile represents a powerful signal for affirmative behaviour, social bond, health and longevity. Additionally, by way of measuring a smile one may be able to predict the outcome of a contested fight [1], whether a couple would stay married or not [2] and even gender [3, 4, 5], just to name some. In very basic terms, a smile can be considered either as a sign of true enjoyment (e.g. in the case of a spontaneous genuine smile) or as a sign of negative emotion (e.g. in the case of a posed smile). For example, posed smiles, indicating negative emotions, in social situations can be inferred as a sign of politeness, shyness, or embarrassment. An important point to bear in mind is that the predictably of certain traits and behaviour depends on understanding the nature of the smile, i.e. whether the smile is genuine or not [6, 7].

1.1. Main Contributions

The prime aim of this work is to develop an efficient computational mechanism for the analysis of smiles. More specifically, the aim here is to study the spatial and temporal distribution of the smile expression on the face. In doing so, we aim to see how the smile is weighted across the face. Through the computational engine we present in this paper, we also like to confirm the hypothesis that the genuine smile has more weight around the eyes as previously shown by psychological studies.

From the extensive literature survey we have undertaken into the topic of smile analysis especially using computational techniques we have not come across studies which try to look into the distribution of a smile across the face.

In particular, the interest here is to develop a robust computational mechanism to study the smile in order to understand the exact weight distribution around each of the corresponding facial features. In doing so, we should be able not only to affirm that the genuine smile is in the eyes but also we must be able to work the proportional distribution of the weight of a smile around the mouth and eyes. Such a computational tool must not only answer a pending theoretical question of the exact distribution weights across the face during a smile, but it should also be useful for devising efficient human computer interaction systems and also developing soft biometrics to complement existing biometric authentication systems. We believe such a tool will be equally useful to social and clinical scientists who are keen to have a deeper understanding of the behavioural and personal traits by studying the face.

The rest of this paper is organised as follows. In Section 2, we discuss some of the key literature relating to the work we present in this paper. In Section 3, we describe our proposed computational framework which we utilise for smile detection and analysis. In this framework, we use standard techniques to process and detect the face from video feed and utilise an optical flow based algorithm for analysing the temporal parts of the smile. In Section 4, we discuss the datasets we have utilised for our experiments. Section 5, is dedicated to discussing the smile analysis experiments we have carried out using our computational framework and we report some of the interesting results we have obtained following those experiments. In Section 6, we reflect on the results and finally, in Section 7, we conclude this paper.

2. Related Work

You meet an old friend after a long time. You shake hands with him and he cracks an old joke. A series of emotional data filters into very specific parts of your brain. As a result, your brain instructs to arouse two major muscles in your face - the zygomatic major and the Orbicularis Oculi. The corners of your mouth curl up, and the outer regions around your eyes squeeze with the

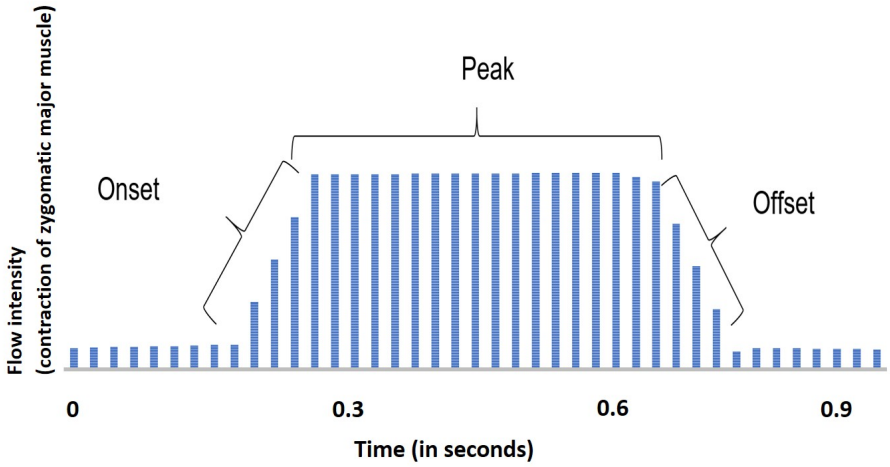


Figure 1: The dynamics of a typical smile onset to apex (peak) to offset.

appearance of visible wrinkles which depicts the shape of a crow's foot. The entire process is short lived only lasting less than two-thirds of a second.

Thus, during a smile, the zygomatic major muscle of the face contracts. The prominent visual effect, as a result, is the raising of the lip corners. As far as the dynamics of a smile is concerned, there are three distinct phases in any type of a smile as shown in Figure 1. They are called the onset, the apex or the peak and the offset. The onset starts from the neutral facial expression to the peak during which the zygomatic major muscle contracts. The apex refers to the time it takes for the smile to stay in the expressive state and the offset is when the facial expression reaches from the expressive state back to the neutral. In the case of a Duchenne (or a genuine) smile, the cheeks get raised along with a significant narrowing of the eye aperture, resulting in wrinkles (or crow's feet) on the outer sides of the eyes. In the case of a felt smile, which also may involve involuntary real laughing, the two major muscles namely zygomaticus and orbicularis oculi contacts. On the other hand, a posed smile also known as a social smile involves the isolated contraction of the zygomaticus major muscle.

The French neurologist Guillaume-Benjamin Duchenne who in the late 1860s was the one who set the stage for the scientists of the 20th and 21st century

to undertake research into facial expression relating them to behaviour and personality traits. Duchenne's work, at the time, started off by applying direct electrical current to faces of people in an effort to understand how various facial muscles are related to various facial expressions, especially the smile expression. His work was later followed by Charles Darwin [8] and others, including Paul Ekman [9].

One commonly used system today to analyse facial expressions, in both human and computer based experiments, is the Facial Action Coding System (FACS). The original idea of FACS was developed by Carl-Herman Hjortsjô in 1969 [10], though much of its credit goes to Paul Ekman who has developed the idea extensively and utilised it to study facial emotions. The FACS system is effective in categorising individual facial expressions based on facial muscle movement which are referred to as Action Units (AUs) [9]. Thus, for example, it is widely accepted that a smile corresponds to the combination of AU12 and AU6. Though FACS is an effective means to categorise facial expressions and inferring the underlying emotions, it is not entirely suitable to study the dynamics of facial expressions in a detailed manner. This is because FACS does not have a mechanism to describe the temporal effects of facial expressions which can be crucial for in-depth analysis of the dynamic face. The establishment of EMFACS which groups individual AUs into categories - by Ekman and his colleagues may serve some purpose to this end, but does not go all the way in mapping the dynamics of a smile expression.

It is interesting to note that of the basic facial expressions (happiness, anger, disgust, fear, sadness, and surprise) which Ekman identifies through his FACS system, only the feeling of happiness with its expressive feature of the smile, is observably related to the underlying physiological and the facial pattern of expression. However, there exists considerable scientific controversy regarding Ekman's other basic patterns of expressions [11, 12, 13].

There is no doubt that research into human facial analysis has recently increased dramatically. In the past, researchers faced significant challenges in their studies relating to human facial expression analysis due to the limitations

of technology for automated video analysis, for example. However, this hurdle appears to be overcoming slowly with developments in automated and accurate facial analysis systems. And, of the facial expressions being studied recently, the smile appears to be one of the most studied facial expression.

Distinguishing between a genuine and posed smiles is still considered to be a challenging task, which according to [14], is associated with perceptual and attentional mechanisms, where humans pay a lack of attention to a certain cue that would help them in their judgement between these types of smiles. Recent research in analysing real (Duchenne) and posed (non-Duchenne) smiles have been carried out from three perspectives: physiological, social experiments and computational studies.

2.1. Physiological and Social Studies

Facial expression analysis and detection is also an active research field in psychology where data are collected, either by using sensors or through social experiments, including showing videos of genuine and posed smiles and using human judgements to initiate some statistical analysis. For example, in a sensor-based system, data are collected using facial electromyography (EMG) [15]. This is a diagnostic technique used for recording facial muscle activity by placing electrodes on the face [15, 16].

Surakka and Hietanen used EMG to measure the zygomaticus major movement, which is one of the facial muscles that pull the mouth into an angle to formulate the smile [14]. This is done by placing electrodes on the subjects' faces under study, thus allowing to measure the smile intensity and their duration. Using this technique, they studied the difference between genuine and posed smiles. Their results imply that spontaneous smiles not only are formed faster than posed smiles but they also use different muscle movements.

Thus, using EMG to classify Duchenne and non-Duchenne smiles, one can identify the muscles related to these smiles in the context of other emotions such as disgust, fear, anger, sadness, interest, surprise and pleasure. The results indicate that both smiles have a high occurrence in the interest, surprise and

pleasure expressions, however, the Duchenne smiles have significantly stronger EMG activity in the periocular and cheek muscle regions as compared to EMG activity in neutral faces.

From a social experimental perspective, a lot of research has been carried out on analysing people's reactions and influence on genuine and posed smiles. Research by Bernstein et al. show individuals' preferences to work with others through Duchenne versus non-Duchenne smiles [17]. In their experiment, participants performed two ostensibly unrelated tasks. First, they wrote essays about experiences of inclusion, exclusion or mundane events. Then, after successfully completing this task, participants responded to 16 items assessing their levels of belonging, control, self-esteem and meaningful existence felt during the experience. Second, participants watched videos of individuals expressing Duchenne (10 videos) and non-Duchenne smiles (10 videos) and were asked to evaluate each individual as a potential partner for a project on which they might work. The results show participants have a greater preference to work with individuals displaying genuine smiles.

Mai et al. investigated the ability of Chinese participants to look for Duchenne and non-Duchenne smiles based on focussing at the mouth and eyes [18]. In their experiment, 100 participants were asked to evaluate 20 videos and were asked to rate each video as real or fake smiles. Afterwards, participants were asked to answer the question, "what part of the face was most useful for discriminating between fake and real smiles?" The results showed that participants highly depend on information from the eyes to successfully distinguish between genuine and posed smiles and participants who focussed on the eyes can identify a genuine smile more accurately.

2.2. Computational Studies

From a computing perspective, a lot of research has been done using psychological studies as ground rules to distinguish between genuine and posed smiles using computational platforms. According to Surakka and Hietanen, a smile that raises the corners of the mouth and raises the cheeks with an appearance

of eye wrinkles is considered as a Duchenne smile, whereas a smile that only raises the corners of the mouth is known to be a non-Duchenne smile [14]. In terms of the FACS system, a Duchenne smile uses AU6 and AU12 and a non-Duchenne smile uses only AU12 [9]. Moreover, according to Manera et al., a Duchenne smile can be categorised as a true expression of happy emotion in comparison to a non-Duchenne smile [19].

Wu et al. proposed a system to detect genuine and posed smiles based on the detection of AU6 and AU12 [20]. This is done by using a Gabor filter and 2D principal component analysis. In addition, they use the well known AdaBoost algorithm for feature reduction. For classification, they used a support vector machine (SVM) and gained 85.9% correct classification. Shan looked at the problem of in-the-wild smile detection by testing pixel differencing using AdaBoost [21]. He compared his method with local binary patterns and Gabor. This pixel differencing method using AdaBoost gives an accuracy of 89.7% and this method appears to be a lot faster too.

In the work of Hoque et al., they looked into the study of smiles under the elicited conditions of frustration and delight [22]. They used the temporal pattern analysis in smiles to distinguish between the two types of smiles. For binary classifications, they used SVM, hidden Markov models and hidden-state conditional random fields and claim to have obtained an accuracy of 92% in distinguishing the smiles under frustrating and delighted stimuli using their dynamic SVM classifier. Similarly, Dibeklioglu et al. classify genuine and posed smiles using distances and angular features for eyelid movement [23]. They use continuous HMM, k-nearest neighbours (k-NN) and naive Bayes classifiers and claim to obtain an accuracy of 91% for posed smiles and 80% for genuine smiles using just the movement of eyelids.

3. Our Computational Framework

Our computational framework is composed of three main entities. First, we process the face from a given video by detecting the face and generating

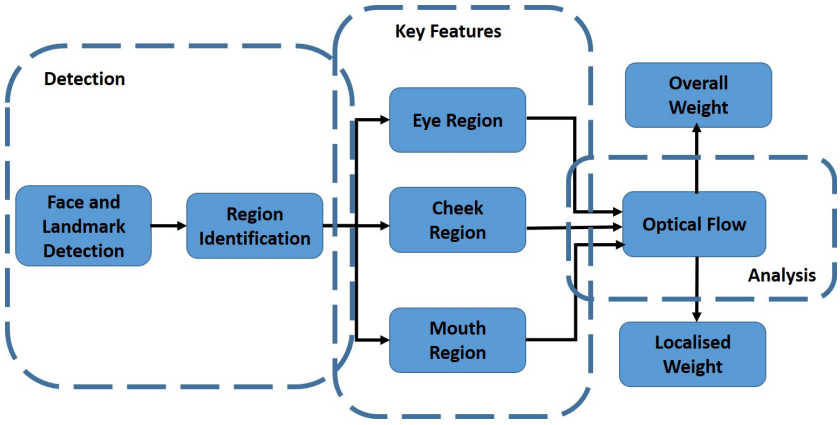
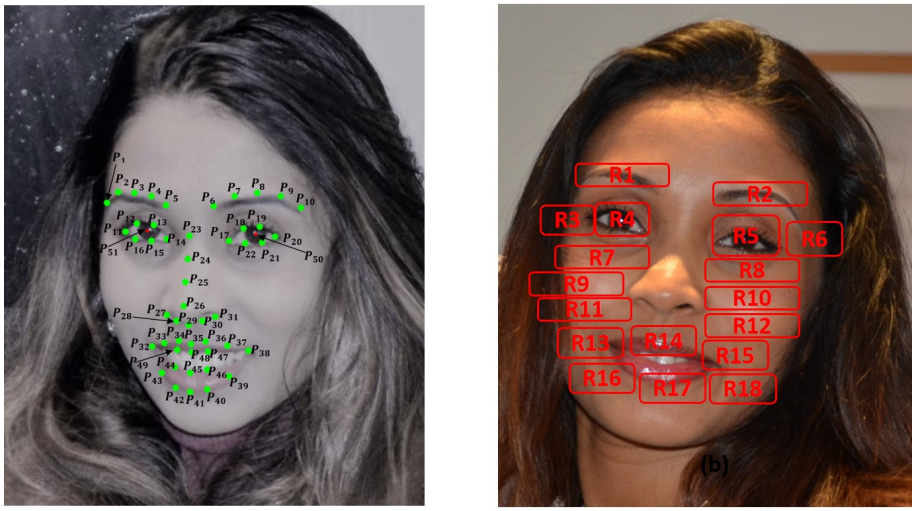


Figure 2: Description of our computational framework for analysis the smile dynamics.

the necessary landmarks to identify important regions of the face. The facial regions of concern here are the eyes, cheek and the mouth. Second, to detect the changes in the key facial regions, we designed an automated algorithm to measure the dynamics of these regions. This is done using an adaptation to the optical flow algorithm as described later. Finally, the optical flow algorithm gives us an output from which we can measure both the localised as well as the overall weight of the motion on each of the facial regions. In Figure 2, we show the flow diagram of our dynamic framework which contains these three main components, namely the face and facial landmark detection, key feature identification and analysis.

The detection phase includes face detection, face resizing, landmark detection and identifying the regions of interest (ROI). The first step in our framework is to detect and track the face within a given video sequence. For this purpose, we have used a well known algorithm for image processing called the Viola Jones algorithm [24]. It is based on Haar feature selection to create an integral image through the use of Adaboost training and cascade classifiers. We have undertaken much work on this algorithm and we have found this algorithm to be fairly stable. Thus, the ability of this algorithm to robustly detect faces under different lighting conditions has been well established and our extensive



(a) (b)

Figure 3: Description of how landmarks are detected and how the relevant regions of interest on the face are determined.

experimentations using the algorithm have demonstrated us that this algorithm is fairly robust to occlusions as well as variations in lighting conditions [25, 26].

Once the face is detected, we then crop the face to a standard size depending on the resolution of the video image. Having a standard for each image frame from the video ensures there are consistency and uniformity across. It also ensures that all the landmark detection followed by the ROI detection is undertaken appropriately.

Landmark detection is done using the CHEHRA model [27], which is a machine learning framework trained to detect the facial landmarks. The output of this algorithm is a set of facial landmark points as shown in Figure 3(a). The CHEHRA algorithm has been trained to detect facial landmarks using faces in-the-wild datasets under varying illumination, facial expressions and head pose. Again, we undertook many tests to evaluate the performance of this algorithm. We found that, in general, the algorithm is acceptable though we noticed that it is unlikely to be fully applicable for real time face analysis applications.

Facial Feature	**Regions of Interest**	**Corresponding Landmarks**
Eyebrows	R1 and R2	P_1 to P_5 and P_6 to P_{10}
Eye corners	R3 and R6	P_{11} and P_{20}
Right eye	R4	P_{12} to P_{16}
Left eye	R5	P_{17} to P_{19}, P_{21} and P_{22}
Upper right cheek	R7	P_{24}, P_{25} and P_{11}
Middle right cheek	R9	P_{25}, P_{26} and P_{11}
Lower right cheek	R11	P_{27} to P_{29} and P_{32}
Upper left cheek	R8	P_{24}, P_{25} and P_{20}
Middle left cheek	R10	P_{25}, P_{26} and P_{20}
Lower left cheek	R12	P_{27} to P_{29} and P_{38}
Upper mouth right	R13	P_{32}, P_{33} and P_{49}
Upper mouth middle	R14	P_{35} to P_{37} and P_{48}
Upper mouth left	R15	P_{37} to P_{38} and P_{47}
Lower mouth right	R16	P_{42} to P_{44}
Lower mouth middle	R17	P_{40} to P_{42} and P_{45}
Lower mouth left	R18	P_{40}, P_{46}, P_{39}

Table 1: Description of the regions of interest, the relevant facial features and the corresponding facial landmarks.

3.1. Identification of Regions of Interest

Upon detection of the landmarks, we define the relevant ROI such as the area around the eyes, cheeks and the mouth. To do this, we first identify all the landmarks in a neutral facial expression which helps us to identify the initial location of the mouth, cheeks and eyes. This process also enables us to normalise the ROI identification process.

Table 1 shows the relevant facial features (i.e. the mouth, eyes and the cheeks) and the associated ROI as well as the corresponding landmarks. The computation of a specific ROI is done through two steps. First, we locate the relevant reference landmarks which are defined as two 2-dimensional vectors (x, y) where variation across each axis is considered separately. Second, using the formulation (λ_x, λ_y), we can denote a shift distance for each axis from the reference points. Thus, using the shift distances (λ_x, λ_y) along with the origin of the central landmark point, we can compute the ROI using Equation 1 such that,

$$RO_x = P_x \pm \lambda_x, \quad RO_y = P_y \pm \lambda_y, \tag{1}$$

where P_x shows the reference point along the x-axis and P_y shows the reference point along the y-axis.

It should be noted that there are ROI where no obvious boundaries can be defined such as the R9, R10, R11, R12 and that below the eyes, R7 and R8. In such cases, where landmarks cannot be directly identified, we use the Euclidean wise nearest reference landmark. For example, to locate the left eyebrows, using landmark P_1 we allocate a window of appropriate size (110 by 35 pixels in this case) which covers the area surrounding the eyebrows. Similarly, to locate the right cheek, we use the right mouth corner point P_{38} as the reference landmark to infer R12, R10 and R8.

Using the above formulation and by means of using the facial landmarks as reference points, we thus allocate the ROI as shown in Figure 3(b). Moreover, we identify 18 ROI through which the motion around mouth, cheeks and eyes can be monitored. In particular, we identify 4 different ROI around the eye, i.e.

eyebrow (R1 and R2), eye corners (R3 and R6) and the area beneath the eye (R7 and R8). The reason we allocate such a greater number of ROI to the eye area is that we wanted to study the regions around the eye area in greater detail. This is because previous work has shown that there are greater distinctions in the level of activity around the eye area between genuine and non-genuine smiles. Therefore, to test this as well as to fine tune our framework, to distinguish between genuine and non-genuine smiles we wanted to study the area around the eyes in greater detail.

3.2. Computing the Optical Flow

Once the relevant ROI are allocated, its motion over time (through the dynamics of a smile) can be tracked and analysed. To do this, we have implemented a technique which we derived using the optical flow algorithms proposed by Lucas Canade [28] and Farnebck [29].

In the optical flow framework, one would consider two successive image frames which can be taken at time t and $t + \Delta t$ at the allocated ROI corresponding to the facial features. Given that we represent the intensity of motion over a given feature, which can be written as $I(x, y, t)$, then following a time period of $t + \Delta t$ we can write the intensity as,

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t). \tag{2}$$

The assumption made in the optical flow computation is that the dynamic changes occurring within the ROI between successive frames are small. Hence by way of Taylor series expansion we can expand Equation 2 as,

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x}\nabla x + \frac{\partial I}{\partial y}\nabla y + \frac{\partial I}{\partial t}\nabla t + ... \tag{3}$$

Thus, we can infer,

$$\frac{\partial I}{\partial x}\nabla x + \frac{\partial I}{\partial y}\nabla y + \frac{\partial I}{\partial t}\nabla t = 0, \tag{4}$$

and dividing throughout by ∇t and representing $\frac{\nabla x}{\nabla t} = V_x$ and $\frac{\nabla y}{\nabla t} = V_y$ in Equation 2, we can write the optical flow as,

$$I_x V_x + I_y V_y + I_t = 0, \tag{5}$$

where, V_x and V_y are components of optical (velocity) flow [28].

To estimate $V_{(x,y)}$, one can employ a range of techniques including linear and non-linear polynomial fittings. For this work, we have utilised a weighted least squares fit such that we minimise, $W^2(x,y)[\Delta I(x,y,t)+I(x,y,t)]^2$, where, $W(x,y)$ denotes a window function that must give more influence to the centre of the region of interest than the periphery of a given feature.

The specific form of $W(x,y)$ is assumed to be elliptic where in this particular case we have used the discrete Harmonic mask M [30], such that,

$$M = \frac{1}{4}\begin{pmatrix} 0 & 1 & 0 \\ 1 & (x,y) & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{6}$$

The choice of this discrete form was to provide a degree of temporal smoothing, which would help mitigate potential noise in the features and to provide a real-time computational framework because harmonic discrete masks are equivalent to an averaging process [31].

Once the optical flow values are computed, they need to be normalised in order to overcome some of the challenging factors such as the face location relative to the camera, the changing size as well as the pose of the face as it moves from one frame to the other. To obtain a uniform normalisation, we use one of the most stable areas of the face namely, the triangular area formed with the tip of the nose, P_{29} and the two eye corners, P_{11} and P_{20}.

After normalising the flow values, we further eliminate some noise in the optical flow computations, which we have found affecting our results. To do this, we check the average flow value, f_{ROI}, of each region of interest and compare it to ϵ which is set to 0.1 as shown in Equation 7. Here, M_{ROI} represents the movement occurrence as a Boolean value set to be true if the flow value exceeds the ϵ which is used later in calculating the smile weight distribution, i.e.,

$$if \;\; f_{ROI} > \epsilon \;:\; M_{ROI} = True. \tag{7}$$

Note, the value of ϵ was decided by carrying out a set of explicit experiments on each ROI related to the facial features for the mouth, nose and eyes.

3.3. Computing the Smile Weight Distribution

Subsequently, as shown in Figure 2, the last part of our proposed work is the analysis, using which we study the relationship between facial features in genuine and posed smiles. We do this at two levels, i.e., we look at the overall weight distribution across the face and we look at the localised weight distribution at the feature level. The overall weight distribution represents the relationship between all the main facial features in genuine and posed smiles. Similarly, the localised weight distribution represents the relation between different parts of the facial features in genuine and posed smiles.

Computing the overall weight is based on the occurrence of flow using Equation 7, where we check whether the value in the region of interest area is higher than ϵ, which indicates a gain in flow and is considered as the flow value for the corresponding facial feature. The weight is computed for each facial feature by counting the number of flow values to the number of subjects in the dataset which indicate how many subjects used the specific facial features through the smile expression.

The localised flow is achieved by computing the flow in specific ROI, using Equation 8 where $f_{R(i)}$ represents total flow in an ROI with identity i. This is done by computing the flow value and summing up all flows in each possible direction which represent the total displacement in specific ROI. Thus,

$$f_{(R(i))} = \sum f_{(R(i),y+)} + \sum f_{(R(i),y-)} + \sum f_{(R(i),x-)} + \sum f_{(R(i),x+)}. \quad (8)$$

Equations 9, 10 and 11 represent the total flows computed using ROI for the eyes f_E, cheeks f_C and the mouth area f_M respectively. Equation 12 represents the total flow value for the face f_{Face}. Again, this is done by summing up the flow values in the eyes, cheeks and mouth.

$$f_E = \sum_{i=1}^{6} f_{R(i)} \quad (9)$$

$$f_C = \sum_{i=8}^{12} f_{R(i)} \quad (10)$$

$$f_M = \sum_{i=13}^{18} f_{R(i)} \tag{11}$$

$$f_{Face} = f_E + f_C + f_M. \tag{12}$$

After computing the above flow values, using Equations 13 to 15, we gain the flow value normalised for each facial feature. Thus, the overall weight is computed by dividing the flow value of each facial feature by the total flows as shown in Equations 9, 10 and 11 for eyes, cheeks and mouth respectively, such that,

$$Eye_{weight} = \frac{f_E}{f_{Face}}, \tag{13}$$

$$Cheek_{weight} = \frac{f_C}{f_{Face}}, \tag{14}$$

$$Mouth_{weight} = \frac{f_M}{f_{Face}}. \tag{15}$$

4. Datasets

We use two different datasets each of which has a specific type of smile - genuine or posed.

The CK+ dataset [32] has 82 subjects expressing posed smiles, where each subject is sat facing the camera with a fixed light. The subjects were asked to express a happy emotion. In fact, the CK+ dataset contains metadata which identifies the AUs, landmarks and the six basic emotions - happy, surprise, anger, fear, disgust and sadness.

The MUG dataset [33] contains 52 subjects of Caucasian origin aged between 20 and 35 expressing laboratory-induced emotions. In order to create non-posed emotions, subjects were recorded while they were watching a video to help them induce emotions. For the video recording setup, subjects were sitting in a chair in front of the camera with a blue background with two 300w light sources. The camera recorded a video with 19 frames per second with size 896 by 896 pixels. Figure 4 shows a sample of the image sequence used in our experiments, Figure 4(a) shows a subject from the CK+ dataset and Figure 4(b) shows a subject from the MUG dataset.

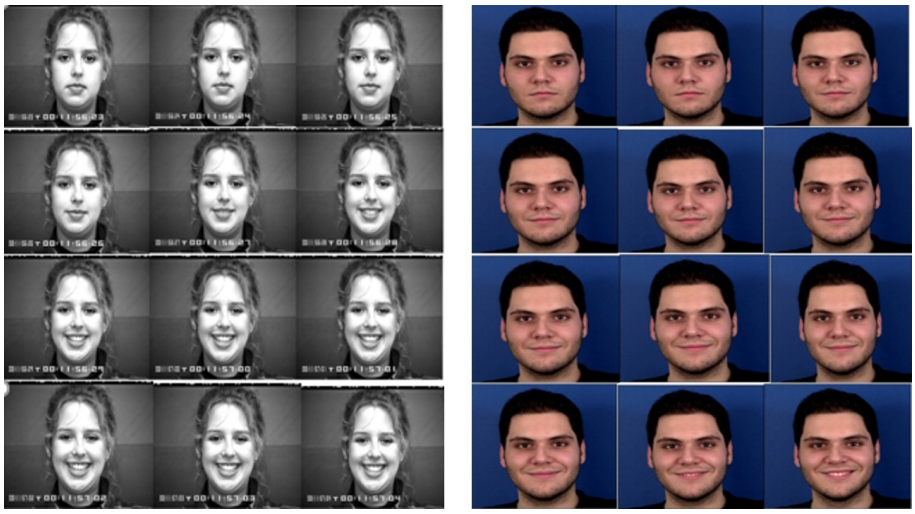


Figure 4: Sample image sequences from the datasets. **(a)** Images from the CK+ dataset and **(b)** images from the MUG dataset

.

For the purpose of balancing the number of subjects and removing potential biases due to an imbalance on the number of subjects, in all the experiments, we used an equal number of subjects from both the datasets. Thus, in this case, we have used all the 52 subjects from MUG dataset and similarly we selected 52 subjects from the CK+ dataset too.

5. Experiments and Results

In this section, we discuss some of the experiments we undertook to validate our computational framework and to obtain interesting results on the weight distribution across the face between genuine and posed smiles. The broad aims of these experiments were to show there is a distinct difference in the weight distribution computed as optical flow and also to calculate the exact weight distribution on individual facial features for genuine and posed smiles.

5.1. Flow Analysis between Neutral and Peak Smile Expressions

In this experiment, we measure the flow values to compare the flow between the neutral and peak smile expressions for both genuine and posed smiles. To do this, we utilise the first set of frames in the neutral expression and we divide them into three groups, namely, $N1$, $N2$ and $N3$. Similarly, the peak frames are divided into three groups namely, $P1$, $P2$ and $P3$. Using the optical flow algorithm described previously we measure the flow related to the ROI for each of the facial features.

In order to check if these values have a significant meaning, we compute the average flow for the neutral and peak smile expression - $f_{N(1-3)}$, $f_{P(1-3)}$ - for the 52 subjects selected from each of the datasets. Thus, we have,

$$Average(f_{N(1-3)}) = \frac{1}{52}\sum_{j=1}^{52}\frac{1}{n}\sum_{i=1}^{n} f_{N(1-3)}, \tag{16}$$

$$Average(f_{P(1-3)}) = \frac{1}{52}\sum_{j=1}^{52}\frac{1}{n}\sum_{i=1}^{n} f_{P(1-3)}, \tag{17}$$

where n represents the number of frames we have selected for each of the neutral and peak groups of the smile expressions. Note, other statistical measures such as median and standard deviation of the smile weight distribution can be computed in a similar fashion.

Figure 5 shows the average of the neutral and peak frames for the mouth, cheeks and the eyes. The results in the figure indicate that the MUG dataset produces a higher average compared to the CK+ dataset.

Furthermore, we compute the median for the neutral and peak of the smile expression, $f_{N(1-3)}$ and $f_{P(1-3)}$, for the subjects in the datasets. Figure 6 shows the median flow values for the neutral and peak smile expression for the mouth, cheeks as well as the eyes. Again, the results show that the MUG dataset has a higher median value compared to the CK+ dataset. Figure 7 shows the average and the median flow for the eyes for the selected subjects in both the datasets further indicating that there is a significant difference between the flow values for the MUG and CK+ datasets.

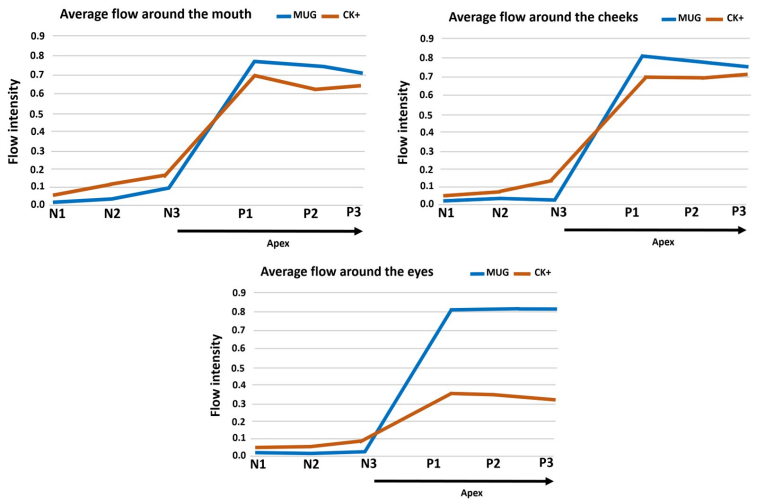


Figure 5: Average flow around the mouth, cheeks and the eyes for the selected subjects in the MUG and the CK+ datasets

.

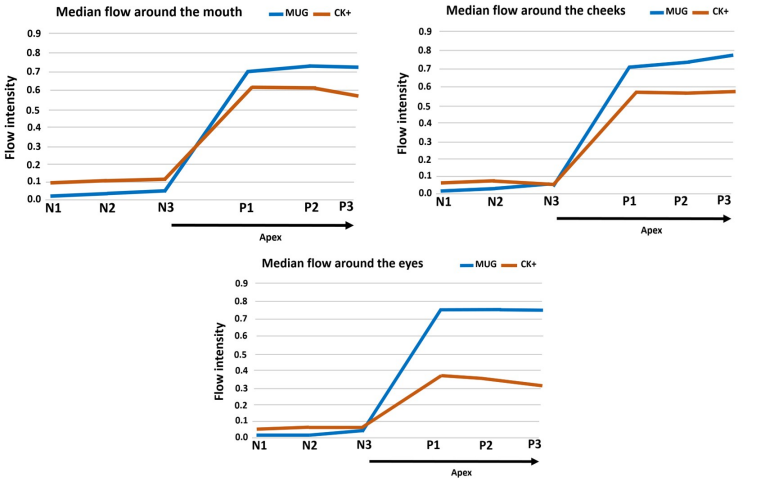


Figure 6: Median flow around the mouth, cheeks and the eyes for the selected subjects in the MUG and the CK+ datasets

.

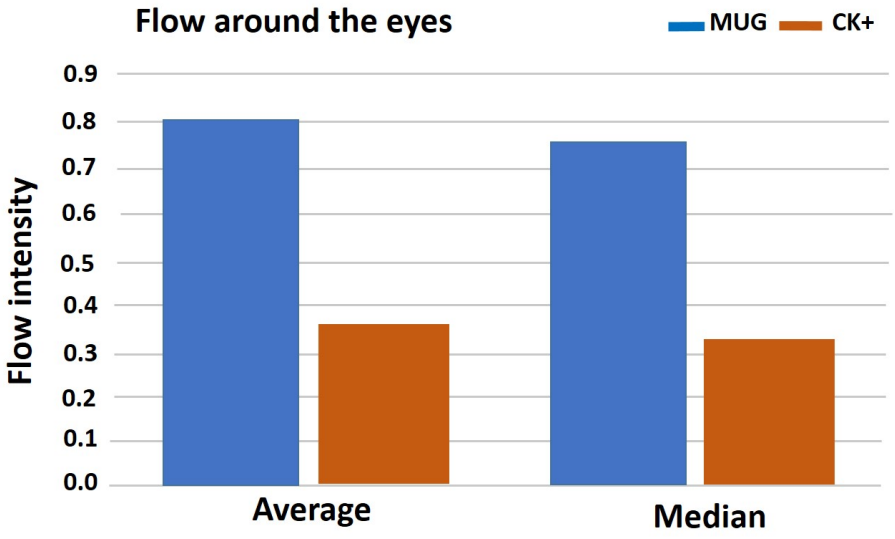


Figure 7: Average and median flow around the eyes for the selected subjects in the MUG and the CK+ datasets

.

To investigate the above results further, in the second experiment, we compute the flow for the peak of the smile expression, $f_{P(1-3)}$, for the selected subjects in both the datasets.

Figures 8 shows the $f_{P(1-3)}$ averages and the standard deviations for the subjects in both datasets for the mouth, cheeks and the eyes. As shown in these figures, there are differences between the flow for the subjects in the MUG and CK+ datasets in the mouth and cheeks areas where the flow around mouth in the MUG dataset has an increase of 79% and the cheeks show 53% increase in flow compared to the corresponding flow values for the subjects in CK+ dataset. Finally, for the eyes, for the subjects in the MUG dataset, we see an increase in 122% compared to the CK+.

As a final experiment, we wanted to measure the weight of the flow distribution across individual facial features under consideration. Thus, we computed the flow around the 18 different ROI. Figure 9 and 10 summarises the results.

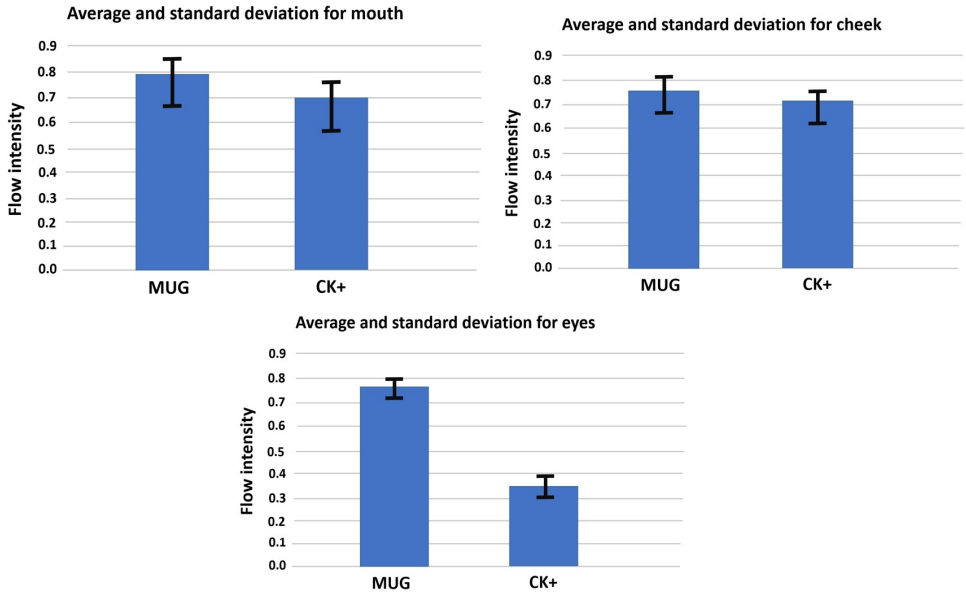


Figure 8: Average and the standard deviation for flow around the eyes, cheeks and mouth for the peak of the smile expression for the selected subjects in the MUG and the CK+ datasets.

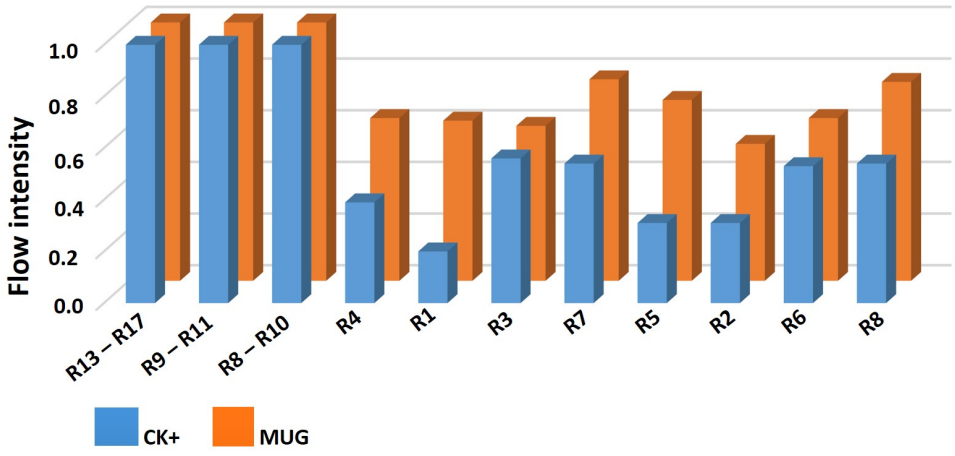


Figure 9: Smile weight distribution for individual ROI for the subjects in the MUG and the CK+ datasets.

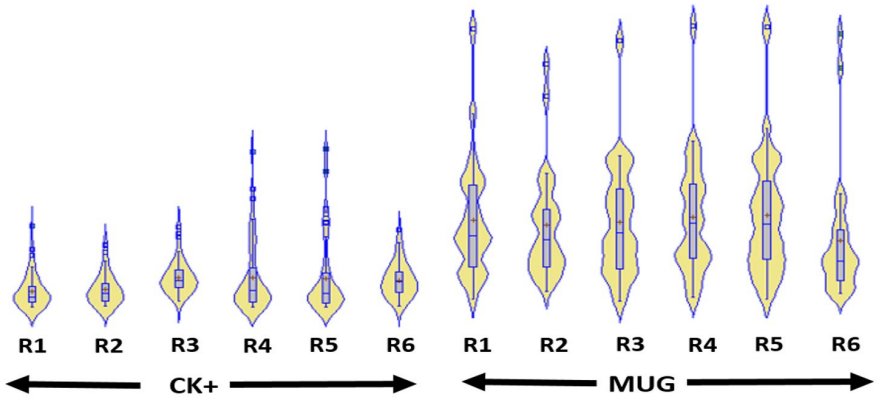


Figure 10: A violin plot showing the relative weight distribution of smiles for some selected ROI for the selected subjects in the MUG and the CK+ datasets.

6. Discussions

As can be seen from the above results, we found there is a significant difference in the flow values between genuine and posed smiles. In the first experiment, where we looked at the neutral and peak portion of the smile expressions, the results show that both the average and the median of the flow around the selected facial features are significantly different. Furthermore, the results also indicate that there is a significant difference between the flow around the eyes for genuine and posed smiles. In fact, in some cases, we observed a 4-fold increase in the average flow values around the eyes in the case of a genuine smile.

Additionally, we computed the flow distribution for the subjects expressing genuine and posed smiles for each of the related ROI in the areas around the eyes. The results show that subjects expressing genuine smiles have a higher average flow distribution and higher median values in the case of a genuine smile.

In order to test the statistical significance of our results for flow distribution across various regions of the face for genuine and posed smiles, we carried out an unpaired t test with the mean and standard deviations as illustrated in Figure

Facial Feature	Error of Difference	p Value
Mouth	0.050	0.0774
Cheeks	0.038	0.4364
Eyes	0.020	0.0011

Table 2: Results of unpaired t-test showing that the flow distribution around the eyes have statistically significant difference among the subjects in the MUG and CK+ datasets.

8. As previously highlighted, the sample size for each of the datasets is 52. With these settings, Table 2 shows the error of difference and the p values for the mouth, cheeks and the eyes. As can be seen, the results do indicate that the flow distribution around the eyes is highly statistically significant.

In general, from the experiments we have undertaken, we can assume there is at least a 10% higher rate of activity around the eye in the case of a genuine smile. This measure thus can be broadly applied for distinguishing between a genuine and a posed smile.

7. Conclusions

In this work, we have looked at identifying the genuineness of a smile by means of analysing the dynamic components of the facial expression associated with the smile. As shown, the weight distribution of the smile implies that it can be detected accurately in both the mouth and cheeks since the flow around these areas are significant among all the test subjects we studied for both genuine and posed smiles. Additionally, there is significant flow distribution in the eyes of the subjects expressing genuine smiles when compared to those expressing posed smiles.

Hence, to answer the question of which part of the facial feature contains most information with regard to a genuine smile, our studies conclude that it is indeed the eyes, as previously shown, e.g. in [11, 20, 23]. From our experiments, we can infer that there is at least a 10% increase in activity around the eye in the case of genuine smiles. Furthermore, genuine smiles show a higher average

and median values when compared to posed smiles, in all regions of the face, especially in the eye area. Thus, our studies confirm what has been previously uncovered. Besides, our framework can be utilised to compute the exact weight distribution over individual facial features in a smile - genuine or posed.

To our knowledge, there is no literature for computationally measuring the exact weight distribution of facial expressions such as the smile expression across the face. Hence, in some sense, it has been difficult to compare and contrast our results. Nonetheless, the computational framework presented in this paper serves to add weight to the psychological and social studies relating to the expressions of the smile, e.g. the genuine smile is indeed in the eyes. Equally, our framework can be used to infer the exact weight distribution of the smile expression, or for that matter other emotional expressions, across the face.

Knowing the distribution of weights across the individual facial features is important in devising future human machine interaction systems. Such systems can assist in accurate measurement of emotional levels of individuals in a non-invasive manner, automated analysis of interviews, detection of deceit as well as intent. We also believe such a tool will be equally useful to social and clinical scientists who are keen to seek a deeper understanding of the behavioural and personal traits of people through facial emotional expressions.

The work we have presented here has clear limitations. For example, we have utilised datasets which are relatively small in size to undertake our experimentation and to draw conclusions from. It will be useful to know the scalability of our results for larger datasets. In addition to this, we have intuitively chosen three areas of the face to focus our study upon. However, there could be other areas of the face which could equally be significant - the eyelid, for example, being one such plausible area.

Computational methods coupled with modern machine learning techniques might be one possible direction to take the studies relating to the smile weight distribution further. Finally, a further area of importance that could be investigated is to see the uniqueness of the smiles and the uniqueness of the weight distribution of the smiles, at the individual level, across the face which we be-

lieve can have potential further applications in person identification and in soft biometrics.

Acknowledgements

This work was supported in part by the European Union's Horizon 2020 Programme H2020-MSCA-RISE-2017, under the project PDE-GIR with grant number 778035.

References

[1] M. W. Kraus, T. W. D. Chen, A winning smile? Smile intensity, physical dominance, and fighter performance, Emotion 13 (2) (2013) 270–279. doi:10.1037/a0030745.

[2] M. J. Hertenstein, C. A. Hansel, A. M. Butts, S. N. Hile, Smile intensity in photographs predicts divorce later in life, Motivation and Emotion 33 (2) (2009) 99–105. doi:10.1007/s11031-009-9124-6.

[3] A. Dantcheva, F. Brémond, Gender estimation based on smile-dynamics, IEEE Transactions on Information Forensics and Security 12 (3) (2017) 719–729. doi:10.1109/TIFS.2016.2632070.

[4] H. Ugail, A. Al-dahoud, Is gender encoded in the smile? a computational framework for the analysis of the smile driven dynamic face for gender recognition, The Visual Computer 34 (9) (2018) 1243–1254. doi:10.1007/s00371-018-1494-x.

[5] H. Ugail, Secrets of a smile? your gender and perhaps your biometric identity, Biometric Technology Today 2018 (6) (2018) 5–7. doi:10.1016/S0969-4765(18)30081-X.

[6] B. Duchenne, The mechanism of human facial expression, 1st Edition, Cambridge University Press, Cambridge, UK, 1990.

[7] M. J. Bernstein, D. F. Sacco, C. M. Brown, S. G. Young, H. M. Claypool, A preference for genuine smiles following social exclusion, Journal of Exper-

imental Social Psychology 46 (1) (2010) 196–199. doi:10.1016/j.jesp.2009.08.010.

[8] C. Darwin, The expression of the emotions in man and animals, John Murray, London, UK, 1872.

[9] P. Ekman, R. Davidson, W. Friesen, The duchenne smile: Emotional expression and brain physiology II, Journal of Personality Social Psychology 57 (2) (1990) 342–353. doi:10.1037/0022-3514.58.2.342.

[10] C.-H. Hjortsjô, Man's face and mimic language, Studen litteratur, 1969.

[11] K. Wolf, Measuring facial expression of emotion, Dialogues in Clinical Neuroscience 17 (4) (2015) 257–262.

[12] E. Poljac, B. Montagne, E. De Haan, Reduced recognition of fear and sadness in post-traumatic stress disorder, Cortex 47 (8) (2011) 974–980. doi:10.1016/j.cortex.2010.10.002.

[13] V. Bruce, A. Young, Face Perception, Psychology Press, Hove, UK, 2011.

[14] V. Surakka, J. K. Hietanen, Facial and emotional reactions to Duchenne and non-Duchenne smiles, International Journal of Psychophysiology 29 (1) (1998) 23–33. doi:10.1016/S0167-8760(97)00088-3.

[15] A. van Boxtel, Facial EMG as a tool for inferring affective states, in: A. Spink, F. Grieco, O. Krips, L. Loijens, L. Noldus, P. Zimmerman (Eds.), Proceedings of Measuring Behavior 2010, Noldus Information technology, 2010, pp. 104–108.

[16] U. Dimberg, Facial electromyographic reactions and autonomic activity to auditory stimuli, Biological Psychology 31 (2) (1990) 137–147. doi:10.1016/0301-0511(90)90013-M.

[17] M. J. Bernstein, D. F. Sacco, C. M. Brown, S. G. Young, H. M. Claypool, A preference for genuine smiles following social exclusion, Journal of Exper-

imental Social Psychology 46 (1) (2010) 196–199. doi:10.1016/j.jesp.2009.08.010.

[18] X. Mai, Y. Ge, L. Tao, H. Tang, C. Liu, Y.-J. Luo, Eyes are windows to the chinese soul: evidence from the detection of real and fake smiles, PloS One 6 (5) (2011) 1–6. doi:10.1371/journal.pone.0019903.

[19] V. Manera, M. Del Giudice, E. Grandi, L. Colle, Individual differences in the recognition of enjoyment smiles: no role for perceptualattentional factors and autistic-like traits, Frontiers in Psychology 2 (2011) 143–152. doi:10.3389/fpsyg.2011.00143.

[20] P. Wu, W. Wang, H. Liu, Methods of recognizing true and fake smiles by using au6 and au12 in a holistic way, in: Z. Sun, Z. Deng (Eds.), Proceedings of 2013 Chinese Intelligent Automation Conference, Springer, Berlin, Heidelberg, 2013, pp. 603–613. doi:10.1007/978-3-642-38466-0_67.

[21] C. Shan, Smile detection by boosting pixel differences, IEEE Transactions on Image Processing 21 (1) (2012) 431–436. doi:10.1109/TIP.2011.2161587.

[22] M. E. Hoque, D. J. McDuff, R. W. Picard, Exploring temporal patterns in classifying frustrated and delighted smiles, IEEE Transactions on Affective Computing 3 (3) (2012) 323–334. doi:10.1109/T-AFFC.2012.11.

[23] H. Dibeklioglu, R. Valenti, A. A. Salah, T. Gevers, Eyes do not lie: spontaneous versus posed smiles, in: Proceedings of the 18th ACM international conference on Multimedia, ACM, 2010, pp. 703–706. doi:10.1145/1873951.1874056.

[24] P. Viola, M. J. Jones, Robust real-time face detection, International Journal of Computer Vision 57 (2) (2004) 137–154. doi:10.1023/B:VISI.0000013087.49260.fb.

[25] G. G. Chrysos, E. Antonakos, P. Snape, A. Asthana, S. Zafeiriou, A comprehensive performance evaluation of deformable face tracking in-the-

wild, International Journal of Computer Vision 126 (2-4) (2018) 198–232. doi:10.1007/s11263-017-0999-5.

[26] A. Al-dahoud, H. Ugail, A method for location based search for enhancing facial feature detection, in: P. Angelov, A. Gegov, C. Jayne, Q. Shen (Eds.), Advances in Computational Intelligence Systems, Springer, Cham, 2017, pp. 421–432. doi:10.1007/978-3-319-46562-3_27.

[27] A. Asthana, S. Zafeiriou, S. Cheng, M. Pantic, Incremental face alignment in the wild, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2014. doi:10.1109/CVPR.2014.240.

[28] B. D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in: Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI'81, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1981, pp. 674–679.

[29] G. Farnebâck, Two-frame motion estimation based on polynomial expansion, in: J. Bigun, T. Gustavsson (Eds.), Scandinavian Conference on Image Analysis, Springer, Berlin, Heidelberg, 2003, pp. 363–370. doi:10.1007/3-540-45103-X_50.

[30] J. Monterde, H. Ugail, On harmonic and biharmonic bézier surfaces, Computer Aided Geometric Design 21 (2004) 697–715. doi:10.1016/j.cagd.2004.07.003.

[31] R. K. Jha, H. Ugail, H. Haron, A. Iglesias, Multiresolution discrete finite difference masks for rapid solution approximation of the poissons equation, 2018 12th International Conference on Software, Knowledge, Information Management & Applications (SKIMA) (2018) 1–7doi:10.1109/skima.2018.8631514.

[32] P. Lucey, J. F. Cohn, T. Kanade, J. M. Saragih, Z. Ambadar, I. A. Matthews, The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression, in: 2010 IEEE Computer

Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010, pp. 94–101. doi:10.1109/CVPRW.2010.5543262.

[33] N. Aifanti, C. Papachristou, A. Delopoulos, The mug facial expression database, 11th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 10 (2010) 1–4.