Automated Facial Features Points Localization for Age Estimation Based on Ideal Frontal Symmetry and Proportion of the Face

Husniza Razalli[1], Rahmita Wirza O.K. Rahmat[2], Fatimah Khalid[2], Puteri Suhaiza Sulaiman[2]
[1]Signal Processing Lab, TM Research & Development, Cyberjaya, Selangor, Malaysia.
[2]Multimedia Department, Faculty of Computer Science & IT, Universiti Putra Malaysia, Selangor, Malaysia.
husniza@tmrnd.com.my

***Abstract*—Age Estimation Components, i.e. the extraction and localization of facial features landmark points, which represent ageing pattern are very important steps in the age estimation process. A good age estimation technique will contain enough variation of landmark points and consider all important points to express the full complexity of the problem. Major progress has been achieved recently in the area of facial landmark localization and extraction method for age estimation. Moreover, measuring facial features landmark points for age classification algorithm has become an interesting subject when dealing with automatic localization process. However, the difficulties to measure the points automatically divert to wrong result if the method unable to locate exact facial features landmark points properly at critical area with complex appearance. Such as measuring points at upper region of face when dealing with individuals with no hair or hair that covered part of the forehead. Therefore, we address this issue by proposing a new method to automatically localize optimal facial landmark points from an input of face image based on Ideal Frontal Symmetry and Proportion of the face. The performance of the proposed algorithm is evaluated with baseline localization using qualitative evaluations. The proposed method achieved a satisfying outcome, which is an average of 80% detection rate for every detected landmark points. The advantage of this method is to accurately identify the points with automatic processing. Each of the point's position localization process was learned independently so that it is suitable to be implemented in real-time face tracking application.**

***Index Terms*—Facial Feature; Facial Landmark Point; Age Estimation; Age Classification.**

I. Introduction

Age estimation in this research is defined as determining the age of a person based on biometric features, more precisely on the basis of two-dimensional images of human face. Facial landmarks can be defined as the standard reference points on the face used by scientists to recognize the face, or in this case, predict the age of a person. Anthropometry is the science dealing with measurements of the size, weight, and proportions of the human body [5]. Therefore, facial anthropometry deals with measurements of the size and proportions of human face that occur during aging.

Facial features landmark points are important key points in this research in order to create algorithm to estimate human facial age for face tracking application. However, measuring crucial facial features landmark points are becoming a prominent subject when dealing with automatic localization process. Furthermore, inaccurate points detection process will distract the structure of the facial age estimation method and will reduce the estimation accuracy, hence influencing the face tracking application.

Other than that, process to measure the facial features landmark points will generate wrong results if the method was unable to locate upper region end points or hair growth starting point, when dealing with individuals with no hair or hair that covers part of the forehead as described in Figure 1 and Figure 2. Furthermore, there will be difficulties to measure the left and right face region end points when dealing with those who cover their head or their faces with hijab or scarf.

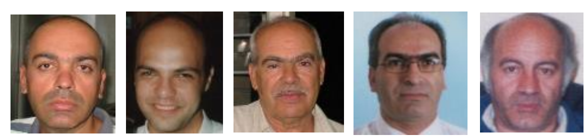

Figure 1: Sample face image from FG-NET datasets of individual with no hair

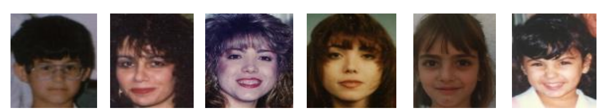

Figure 2: Sample face image from FG-NET datasets of individual with face region was covered part of the forehead

The rest of this paper is organized as follows. Section IV provides step by step methodology of the proposed work. Section V demonstrates an experimental setup, while section VI describes experimental results. Finally, conclusions are drawn in Section VII.

II. Background

In early stage of age estimation method, the facial features landmark points extraction was performed manually with the understanding of the effect of age on Facial Anthropometry. The landmark points came with the facial aging image database [6]. The points were obtained using model or tools with anthropometry perception where it has become the state of the art in the field of this research. The facial features landmark points extract the biometric ratios of human face to estimate the age of the face. The following images show facial landmarks used for age estimation algorithm by several different authors (Figure 3 to Figure 6).

Figure 3: Landmark points as identified by Ramanathan et. al. [1]

Figure 4: Landmark points as identified by Dehshibi et. al. [4]

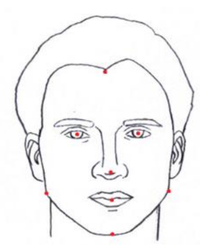

Figure 5: Landmark points as identified by Köstinger [2]

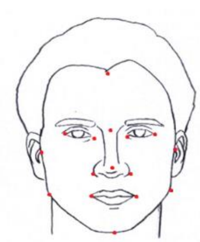

Figure 6: Landmark points as identified by Shima et. al. [3]

Face Anthropometry is the scientific study of the measurement and proportion of the biometric ratios of the human face across ages. It also provides a quantitative description of the craniofacial complex using measurements carried out over the key landmarks on human face during aging. One of the most important research in this field is the one conducted by Farkas et. al. [5]. The researcher defines 57 landmarks on human face and provides a comprehensive overview of face anthropometry studies. Research related to craniofacial growth have shown that the face shape changes from circular to oval during age progression [7]. These changes cause slight changes in the position of the primary facial features landmark points. In adults, position related to sides of the face changes slightly and the distance between the main features remains fixed. Therefore, distances between some landmarks can help distinguish immature face from others. For instance, in babies the distance between the eyes are close to the distance between the eyes and nose. In addition, the distance between the eyes and nose are close to the distance between the nose and lip. Therefore, the information regarding the location of facial feature landmark point becomes a crucial focus that could help identifying a person's age groups.

III. EXISTING METHOD

Previous works introduced various collection of facial features landmark points and some of them used more than 20 points in age estimation and classification algorithm. In this research, determination facial feature landmark points is based on two important issues. The first issue is related to the selection process of the most influential landmark points in age progression, which contributes to the ratios calculation in age estimation algorithm. The second issue is related to which efficient landmark points that needs to be significantly extracted and localized using automatic processing.

Petra et. al. [8] run several experiments to prove the usability of facial features landmark points in calculating human facial age. He finally claimed that, although a lot of landmark points are used in literature, not all points are practical in measuring human facial age or contribute in ratio calculation. Next, Ramanathan et. al. [9] identified the transformational invariants and structural invariants that are associated with facial growth, with the objective of identifying the factors that were responsible for similar change of patterns each time the event was manifested. He observed that the structural invariants associated with facial growth were responsible for preserving the identity of individuals over the course of facial growth. Shaw et al. [10] sought to identify mathematical transformations that help characterize the facial growth event. They discovered two transformations that could be applied on the outer contour of faces in the 'profile view' namely;

1. Cardioidal strain: Stretches the face downward and outward.
2. Affine shear: When applied in the right proportion, introduces a protrusion in the jaw and a backward slant in the forehead.

Later, Dehshibi et. al. [2] concluded that the studies related to craniofacial growth have shown that the face shape changes from circular to oval as a person grows. These changes caused slight changes in the position of the primary facial features. In adults, position related to sides of the face changes slightly and the distance between the main features remains fixed. Therefore, distances between some landmarks can help distinguish children from adults.

If changes that happen during facial growth and aging are taken into consideration, where eyes, nose and mouth can be detected automatically using existing tools, some of these landmarks can be omitted. The second step of purification occurs and new set of facial features landmarks points should be defined to fulfill all requirements.

Although facial age range estimation has achieved maturity and a high performance level in manual calibration algorithm, limitations and challenges still remain in automatic detection and automatic extraction of facial features landmark points [11]. Automatic landmark points localization in still image is useful in many computer vision tasks where object recognition or pose determination is needed with high reliability [12]. It aims to facilitate locating points correspondence between images and a known model where natural features, such as the texture shape or location information, are not present in sufficient quantity and uniqueness.

Old-style calibration with manual localization method in estimating human facial age is a common approach. In order to design method for face tracking application, manual fitting approach is unsuitable. Based on the objective of this research, which is to concentrate on the problems to implement age estimation method in a face tracking application, an automated extraction and localization technique are needed. Therefore, this paper proposes a novel approach for extracting and locating optimal facial features landmark points automatically based on

Ideal Frontal Symmetry and Proportion of the Face using face image as an input.

IV. METHODOLOGY

The aim of the proposed method is to localize facial features landmark points accurately based on automatic extraction processes in order to interpret them as an object features to calculate ratios in age estimation algorithm. The proposed method comprises of several stages as illustrated in Figure 7. Firstly, the method starts with pre-processing steps, which is a standard procedure of detecting face region and facial features from the input image. The Viola Jones face detection algorithm [13] is used in this stage. In general, Viola Jones face detection algorithm is further divided into three basic steps. The three basic steps include feature extraction, boosting and multi scale detection. It is followed by algorithm implementation to measure selected facial features landmark points based on extracted facial features.

The Facial feature landmark points are mainly located based on key facial feature components such as eyes, mouth, nose and chin. Usually, the process starts with detection of a rectangular bounding boxes returned by the Viola Jones face detector [14]. The detector implies the location of a face and key facial components. Then, this bounding box can be employed to initialize the position of facial features landmark points.

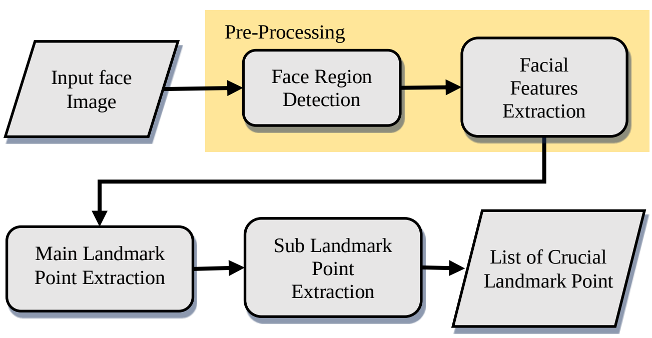

Figure 7: The steps of the Proposed Facial Landmark Point Extraction and Localization Algorithm

In this work, nine crucial facial features landmark points are obtained to propose a new age estimation algorithm as illustrated in Figure 7. Four landmark points are labeled as a main landmark points that localize based on extracted facial features region discussed previously. Five landmark points which are labeled as a sub landmark points are localize based on Ideal Frontal Symmetry and Proportion of the Face. The details about these five points will be discussed in the next sub topic.

A. Identification of Main Facial Features Landmark Points

In order to reduce the process of the localization algorithm and computation cost, the extracted region is divided into searching perimeter. Center point of all extracted facial features regions are examined as a main facial features landmark point described in Figure 8. For example, after obtaining the values of $x_1, x_2, y_1,$ and $y_2,$ the the center position of the eye described in Figure 9, (x_0, y_0) can be determined by the equation below:

$$x^0 = \frac{x^1 - x^2}{2} \qquad y^0 = \frac{y^1 - y^2}{2} \tag{1}$$

The same equation and the same projection function were applied to all extracted facial features region in order to localize main landmark points. The other main landmark points as previously stated is the center point of the other side of the eye, center point of the nose and center point of the mouth.

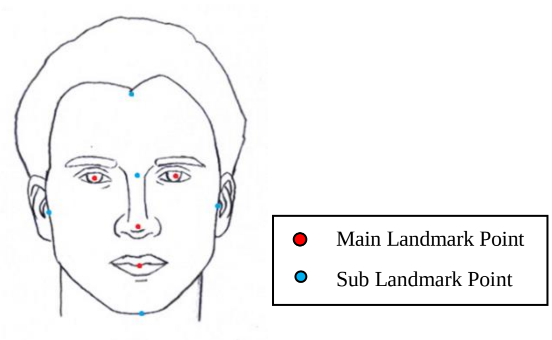

Figure 8: The Proposed Facial Features Landmark Points

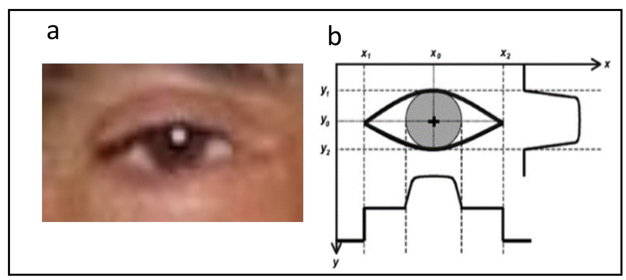

Figure 9: (a) Eye Region. (b) Eye Model with center point examination

B. Identification of Sub Facial Features Landmark Points

As discussed by previous researcher in [15] The face can be divided into three portions horizontally described in Figure 10. The upper third extends from the hairline to the glabella, the middle third from the glabella to the subnasale, and the lower third from the subnasale to the menton. The lower face also can be divided into thirds, the upper lip length from subnasale to stomion equaling one third, and the lower lip and the chin encompassing two thirds.

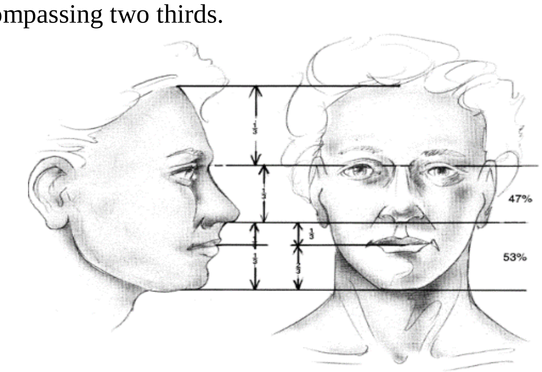

Figure 10: Illustration of Facial Proportion Horizontally [15]

On the frontal view, the symmetry of the face is assessed about the vertical midline, and the width-to-length ratio of the face should approximate a 3:4 ratio. Further assessment of the symmetry and proportion can be accomplished by dividing the face into fifths vertically, with each fifth approximating one eye width as shown in Figure 11.

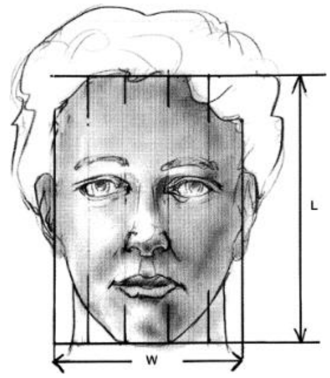

Figure 11: Illustration of Facial Proportion Vertically [15]

Based on these Ideal Frontal Symmetry and Proportion of the Face, we formulate an equation to localize sub points of facial features to estimate human face. Figure 12.

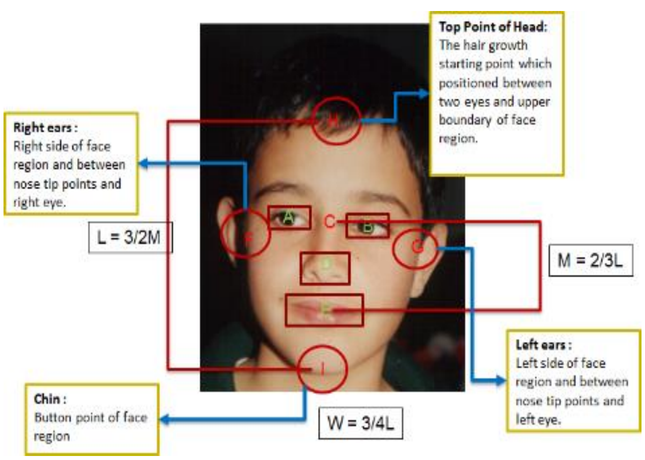

Figure 12: Position and Label of Sub Landmark Points that Used in This Research (Image obtained from [6])

Given point of; Right eye A = (xa, ya), Left eye B = (xb, yb), Nose D = (xd, yd) and Mouth E = (xe, ye). Following are the calculation for those points.

1. Middle point of two eye C:
$$xc = \frac{xa + xb}{2} \qquad yc = \frac{ya + yb}{2} \tag{2}$$
2. Point of Chin I:
$$xi = xe + \frac{xe - xc}{2} \qquad yi = ye + \frac{ye - yc}{2} \tag{3}$$
3. Point of Forehead H:
$$xh = xc - (xe - xc) \qquad yh = yc - (ye - yc) \tag{4}$$
4. Point of Right Ears F:
$$xf = xa - (xc - xa) \qquad yf = ya - (yc - ya) \tag{5}$$
5. Point of Left Ears G:
$$xg = xb + (xb + xc) \quad yg = yb + (yb - yc) \tag{6}$$

V. EXPERIMENTAL SETUP

The localization of optimal facial landmark points from an input face image using Ideal Frontal Symmetry and Proportion of the Face is evaluated on the FG-NET aging database. The database contains 1002 face images from 82 subjects with the age range between newborn to 69 years. However, only 180 images with frontal view were chosen as samples in the experiment. With the aim of identifying and addressing the challenge that is associated with such automatic face tracking application, the selected sample image must be limited to several scopes as follows;

1. The sample image must contain image with less expression or normal expression
2. The face in the image comprises a rotation tolerance of ±15° of head pose and full frontal view of the image.
3. The sample normal face, which refers to the person in the image should be in a normal facial feature orientation appearance, or in other words, the person does not have any facial disabilities.
4. The sample is not wearing any facial accessories such as spectacles, nose rings, studs or jewelry.

The method was programmatic based on each process illustrated and explained in the proposed methodology section and runs on MATLAB R2013(a). FG-NET database adopted from [6] is used for this implementation. It is also been resided in the application for testing and validation process. The experiment should be able to accurately identify the facial features landmark points from the input of face image.

VI. RESULT AND DISCUSSION

The several experiments conducted involved the localization of facial features landmark points from extracted facial feature region of interest. The performance of the proposed method is evaluated with baseline localization. In fact, the detection and the localization of the landmark points have been continuously tested with more than seventy datasets for qualitative evaluation in order to monitor the improvement in the method applied in [12]. The output of the experiments is the location of 9 facial features landmark points in the test images. In order to evaluate the accuracy of detection, the detection error of facial points is defined as follows:

$$error\_i = \frac{||(x_i, y_i) - (xo_i, yo_i)||}{\mathrm{ED}} \tag{7}$$

Where, ED stands for the points distance; (xi, yi) is the detected position of the facial point i; (xoi, yoi) is the manually labeled position of the facial point i.

Based on the detection error, a correct detection rate is further defined as:

$$ri(m) = \begin{cases} 1, \text{if } error\_i_{(\mathrm{m})} < \mathrm{T\,h}; \\ 0, \text{otherwise} \end{cases} \tag{8}$$

$$DRi = \frac{\sum_{m=1}^{M} \mathrm{ri(m)}}{\mathrm{M}} \tag{9}$$

where, M is the number of test images; $error\_i_{(m)}$ is the error of detecting the facial feature point i in the m T h test image; T h denotes a distance threshold in pixels. When $error\_i_{(m)}$ is greater thanT h, the detection is considered unsuccessful.

The results of the cross-validation test are listed in Table 1. The detection errors of all 9 facial features landmark points are calculated based on equation (7). The mean and standard deviation of $error\_i$ (i ∈ [A, B, C, D, E, F, G, H, I) are denoted as $error\_\mu$ and $error\_\alpha$. The detection rates DR respective to T h = 0.03, 0.08, and 0.12 (see equation (7)) are shown in Table 1. In Figure 2, the average detection rate of the 9 facial feature points has been indicated respective to T h ∈ [0.01, 1.6].

Table 1
The error of the cross-validation test

Point	A	B	C	D	E	F	G	H	I
error_µ (%)	3.77	3.35	3.78	3.83	2.62	2.46	1.73	5.94	1.57
error_α (%)	2.58	2.86	3.81	3.65	2.82	2.52	1.07	1.01	4.84

Table 2
The detection rate of the cross-validation test

Point	A	B	C	D	E	F	G	H	I
DR(%) Th = 0.03	43.33	57.92	57.08	52.08	78.75	77.50	87.08	91.67	25.42
Point	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
DR(%) T h = 0.08	94.17	94.58	91.25	87.92	92.92	93.75	100	100	82.92
Point	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
DR(%) T h = 0.012	98.33	97.08	94.17	95	97.92	98.33	100	100	92.92

The observation can be made based on Table 2. The table shows the implication of the relationship between the detection rate of each facial feature points and the distance threshold T h. While T h = 0.03 is a strict standard, the detection rate of D is still high, which has further proven the previous analysis. When T h = 0.08, the detection of most facial feature points has been regarded as achieving a satisfying outcome of roughly 80% detection rate, except for the left and right face region points.

In addition to this evaluation, quantitative measurement is also applied. An accuracy measurement used by [16] with two metrics, which are *precision* and *recall* is applied to evaluate the correctness. *Precision* can be defined as exactness of measurement, whereas *Recall* is a measure of completeness. Let TP (True Positive) represents the number of true detected landmark point; FP (False positive) is the number of falsely detected landmark point; FN (False Negative) serve as missing detection of landmark point. The ideal value for *Accuracy* should be 1. The ground truth or landmark point that is manually marked is formed from TP+FN. The metrics are as following:

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$Accuracy = \frac{(Precision + Recall)}{2} \tag{12}$$

The evaluation started by manually creating a set of ground truth facial images obtained from FG_NET aging database. The correct landmark points are labeled in the facial image. The labeled landmarks in the ground truth are compared with the proposed method. The landmarks that are provided by FG-NET [6] are frequently used in the literature which used AAM tools to localize the landmark and closest method to this work. For fair evaluation, the selected landmark point should comply with the same evaluation criteria. What is important is how the method performs according to the target criteria, not which method is employed in the design of that detector.

VII. Conclusion

The advantage of the proposed method as described in this paper is to accurately identify the location of facial features landmark point, which lies on a person’s face image using automatic processing. Each landmark point position estimation structure was learned independently in our method. However, the available database such as FG-NET dataset provides facial images with numerous constraints. Those constraints gave difficulties in measuring the points automatically which could generate wrong results. However, the proposed method have the ability to deal with these issues.

Previous works focused on multilevel preprocessing and complicated functionality to estimate facial age, which involves complex computation. Hence more time is needed to detect, extract and process numerous facial features landmark points. Based on this case, those available age estimation methods that are currently available in the market, are inappropriate to be implemented in a real-time face tracking application. By minimizing the involvement of image processing work particularly to obtain facial features landmark points, the minimization will reduce processing time in the real-time tracking application.

Although the proposed method produces positive results, these localization and extraction techniques still faces several limitations. One of the crucial limitations is, the process to extract and locate facial features that totally depends on the detection of face region and facial features. If the detection failed to detect an appropriate region, the method will produce a negative result. For future enhancement, we will focus on improving the accuracy of the extraction and localization of the method, and enhance the process by implementing it in real-time facial age estimate application.

References

[1] N. Ramanathan, and R. Chellappa, "Modeling age progression in young faces." pp. 387-394, 2006.

[2] M. M. Dehshibi, and A. Bastanfard, “A new algorithm for age recognition from facial images,” *Signal Processing*, vol. 90, no. 8, pp. 2431-2444, 2010.

[3] S. Izadpanahi, and Ö. Toygar, “Human Age Classification with Optimal Geometric Ratios and Wrinkle Analysis,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 28, no. 02, pp. 1456003, 2014.

[4] M. Köstinger, P. Wohlhart, P. M. Roth *et al.*, "Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization." pp. 2144-2151, 2011

[5] L. G. Farkas, *Anthropometry of the Head and Face*: Raven Press, 1994.

[6] FGNet. "The FGNet Aging Database," 14th June, 2013; http://www-prima.inrialpes.fr/FGnet/html/benchmarks.html.

[7] N. Ramanathan, and R. Chellappa, “Face verification across age progression,” *Image Processing, IEEE Transactions on*, vol. 15, no. 11, pp. 3349-3361, 2006.

[8] K. Petra, B. Miroslav, and S. Markus, "Analysis of craniofacial morphology changes during aging and their connection with facial age estimation." pp. 481-486, 2011.

[9] N. Ramanathan, R. Chellappa, and S. Biswas, “Age progression in human faces: A survey,” *Journal of Visual Languages and Computing*, vol. 15, pp. 3349-3361, 2009.

[10] R. Shaw, M. McIntyre, and W. Mace, “The role of symmetry in event perception,” 1974.

[11] C.-T. Lin, D.-L. Li, J.-H. Lai *et al.*, “Automatic Age Estimation System for Face Images,” *International Journal of Advanced Robotic Systems*, pp. 1, 2012.

[12] N. Markuš, M. Frljak, I. S. Pandžić *et al.*, “Fast Localization of Facial Landmark Points,” *arXiv preprint arXiv:1403.6888*, 2014.

[13] Y.-Q. Wang, “An Analysis of the Viola-Jones face detection algorithm,” *Image Processing On Line*, vol. 4, pp. 128-148, 2014.

[14] P. Viola, and M. J. Jones, “Robust real-time face detection,” *International journal of computer vision*, vol. 57, no. 2, pp. 137-154, 2004.

[15] P. M. Prendergast, “Facial Proportions,” pp. 15-22, 2012.

[16] T. Le, U. Prabhu, and M. Savvides, "A novel eyebrow segmentation and eyebrow shape-based identification." pp. 1-8, 2014.