Robust Facial Feature Detection for Registration

Taher KHADHRAOUI
LSTS Laboratory
National School of Engineers of Tunis (ENIT), Tunisia
khadhra.th@gmail.com

Fouazi BENZARTI
LSTS Laboratory
National School of Engineers of Tunis (ENIT), Tunisia
benzartif@yahoo.fr

Hamid AMIRI
LSTS Laboratory
National School of Engineers of Tunis (ENIT), Tunisia
hamidlamiri@yahoo.com

***Abstract*—Face image analysis, in the context of computer vision, is, in general, about acquiring a high-level knowledge about a facial image. Facial feature extraction is important in many face-related applications, such as face recognition, pose normalization, expression understanding and face tracking. Although there is not yet a general consensus, the fiduciary or first tier facial features, most often cited in the literature, are the eyes or the eye corners, the tip of the nose and the mouth corners. Similarly, second tier features are the eyebrows, the bridge of the nose, the tip of the chin, and so on. Registration based on feature point correspondence is one of the most popular methods for alignment. The importance of facial features stems from the fact that most face recognition algorithms in 2D and/or 3D rely on accurate feature localization. This step is critical not only directly for recognition techniques based on features themselves, but indirectly for the global appearance-based techniques that necessitate prior image normalization. For example, in 2D face recognition, popular recognition techniques, such as eigenfaces or Fisherfaces, are very sensitive to registration and scaling errors. For 3D face recognition, the widely used iterative closest point (ICP) registration technique requires scale-normalized faces and a fairly accurate initialization. In any case, both modalities require accurate and robust automatic landmarking. In this paper, we inspect shortcomings of existing approaches in the literature and propose a method for automatic landmarking of near-frontal faces. We show good detection results on different large image datasets under challenging imaging conditions.**

Keywords-component; face Alignment; feature extraction; face registration

I. INTRODUCTION

The face detection and facial feature extraction is instrumental for the successful performance of subsequent tasks in related computer vision applications. Many high-level vision applications such as facial feature tracking, facial modeling and animation, facial expression analysis, and face recognition, require reliable feature extraction. Facial feature points are referred to in the literature as "salient points", "anchor points", or "facial landmarks" [9]. The most frequently occurring fiduciary facial features are the four eye corners, the tip of the nose, and the two mouth corners [1]. Facial feature detection is a challenging computer vision problem due to high inter-personal changes (gender, race), the intra-personal variability (pose, expression) and acquisition conditions (lighting, scale, facial accessories). To make valid, more accurate, quantitative measurements in diverse applications, it is needed to develop automated methods for recognition. Generally, these systems include automatic feature extractors and change trackers in the facial features in static images.

The paper is organized as follows. In Section 2 we review some related work in the face image analysis. The proposed approach is given in Section 3. The experimental results are provided in Section 4. In Section 5, we discuss future work and draw our conclusions.

II. RELATED WORK

The analysis of faces has received substantial effort in recent years. In Facial feature extraction, local features on face such as nose, and then eyes are extracted and then used as input data. And it has been the central step for several applications [8]. Various approaches have been proposed in this article to extract these facial points from images or video sequences of faces. Among these approaches: Geometry-based approach, Template-based, Apprearance-based approaches.

A. Geometry-based approaches

These methods extracted features using geometric information such as relative positions and sizes of the face components.

Mauricio Hess and G. Martinez [2] used SUSAN algorithm to extract facial features such as eye corners and center, mouth corners and center, chin and cheek border, and nose corner etc. Nevertheless these techniques require threshold, which, given the prevailing sensitivity, may adversely affect the achieved performance.

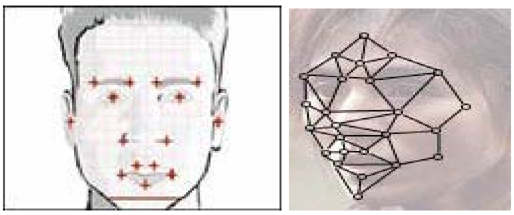


Figure 1. Geometry based approach

B. Template-based

This technique, matches the facial components to previously designed templates using appropriate energy

functional. Genetic algorithms have been proposed for more efficient searching times in template matching.
The best match of a template in the facial image will yield the minimum energy. Proposed by Yuille et al [3] these algorithms require a priori template modeling, in addition to their computational costs, which clearly affect their performance.

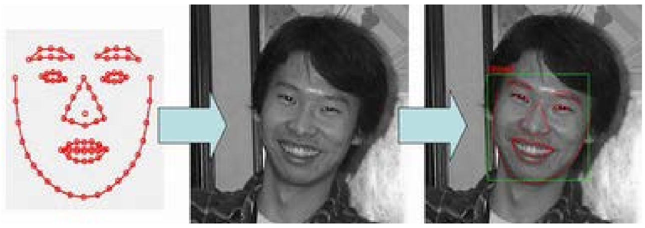


Figure 2. Template based method

C. Colour segmentation techniques

This approach makes use of skin color to isolate the face. Any non-skin color region within the face is viewed as a candidate for eyes and/or mouth. The performance of such techniques on facial image databases is rather limited, due to the diversity of ethnical backgrounds [4].

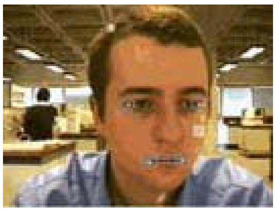


Figure 3. Colour segmentation approach

D. Apprearance-based approaches

These approaches generally use texture (intensity) information only and learn the characteristics of the landmark neighborhoods projected in a suitable subspace. Methods such as principal component analysis [5], independent component analysis [6], and Gabor wavelets [7] are used to extract the feature vector.
These approaches are commonly used for face recognition rather than person identification.

III. PROPOSED METHOD

The proposed method is summarized in figure 4. It uses four main steps:

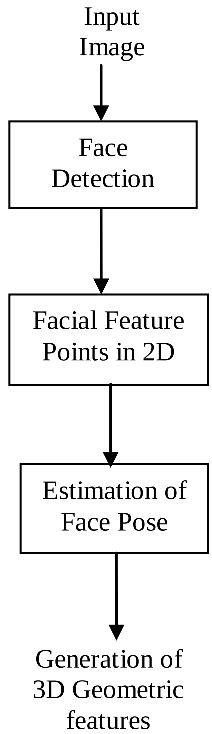


Figure 4. Framework of facial features extraction

Typical image analysis and in particular facial features detection usually consists of several steps:

A. Face detection

Face detection is a technology to determine the locations and size of a human being face in a digital image. It only detects facial expression and rest all in the image is treated as background and is subtracted from the image. Region of interest (ROI) is defined here as a facial feature candidate point or region [1]. Depending on the representation of the facial features, methods can be divided to region based and point based.
In some applications preprocessing may increase the accuracy of the localization. This applies mostly to the cases where the acquisition parameters are insufficient, for example poor lighting, noise or inadequate camera properties.
Point based ROI detection can be performed in various ways. Most of the facial features, for example eye corners, mouth corners, nostrils, are placed on the edges.
To build fully automated systems that analyze the information contained in face images, robust and efficient face detection algorithms are required. Given a single image, the goal of face detection is to identify all image regions which contain a face regardless of its three-dimensional position, orientation, and lighting conditions. Such a problem is challenging because faces are nonrigid and have a high degree of variability in size, shape, color, and texture. These algorithms aim to find structural features that exist even when the pose, viewpoint, or lighting conditions vary, and then use the these to locate faces.

B. Features extraction

In the 2D processing part of the proposed method a number of points are detected across the facial area of the current input image.

1) Nose holes.

Finding nose holes in an area given from face's geometry depends on the angle between camera and face. If there isn't a direct line of sight between nose holes and camera, it is obviously not possible to track them.
Nose holes color have a significant saturation, depending on its color black.
The threshold must be defined and over geometry or clustering two centers of saturation can be found.

2) Mouth

Detecting the middle of the mouth isn't as simple as it is thought. There are a lot of possibilities, going over gradient horizontal and/or vertical decent, hue or saturation. At the moment it is implemented utilizing the distinct hue of lips. Light reflects on lips and this point is fetched by a defined hue value. In contrast to the methods, this method is not light independent, thus intensity and direction of the light can influence results. A better method should be included in the future.

3) Eyes and pupils

A lot of ways can be developed to find pupils in the given area surrounding the eyes, a Gabor eye-corner filter is constructed to detect the corner point. It is more robust than projection methods or edge detection methods.

C. Estimation of Face Pose

In order to correct those geometrically violated facial features that deviate far from their actual positions, the geometry constraint among the detected facial features is imposed. However, in practice, the geometry variations among all the thirty facial features under changes in individuals, facial expressions and face orientations are too complicated to be modelled.
Estimation of face pose is a fundamental task in computer vision. We infer face pose from geometric alignment of face model and coarse face mesh reconstruction.
Homogeneous model transformation matrix T_{4x4} formulates the alignment (Eq. 1). The accurate triangular face mesh model is gained from a previous range scan of the person observed.

$$T_{4\times4} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (1)$$

We apply a least squares error metric E (Eq. 1) that minimizes the sum of squared distances d_i from each point m_i of model to the plane containing the destination point S_i, oriented perpendicular to normal

$$E = \sum_i (d_i) = \left((T_{4x4} * m_i - S_i) * \vec{n_i} \right)^2 \qquad (2)$$

$T_{4\times4}$: Homogeneous Model Transformation Matrix
m_i, S_i: Points of model and Reconstruction
Geometric alignment of mesh M_{3D} (3) and point set W_{3D} (4) is realized with a variant of the Iterative Closest Point (ICP) algorithm. In the ICP procedure we determine pose vector T (5), which represents the optimal model transformation parameters, with respect to an error metric.

$$M_{3D} = (a_1, b_1, \dots, a_n, b_n), a_j \in 3D, b_j \in V, n > 1 \qquad (3)$$

$$W_{3D} = (p_1, p_2, \dots, p_n), p_i \in 3D, n > 1 \qquad (4)$$

$$T = (t_x, t_y, t_z, \varphi, k, \omega) \qquad (5)$$

The ICP principle applied is as follows:

- Let cluster W_{3D} (4) be a set of n points p_i and M_{3D} (3) a surface model with m vertices a_j and normals b_j
- Let CP(p_i, a_j) be the closest vertex a_j to a point p_i

  1. Let $T^{[0]}$ be an initial transformation estimate (5).
  2. Repeat for k = 1...k_{max} or until convergence:

- Compute the set of corresponding pairs S

$$S = \bigcup_{i=1}^{m} \left\{ \left(p_i, CP\left(T^{[k-1]}(p_i), a_j \right) \right) \right\} \qquad (6)$$

- Compute the new transformation $T^{[k]}$ that minimizes Error metric E (2) w.r.t. all pairs S.

Finally, we present a feature preserving Delaunay refinement algorithm which can be used to generate 3D geometric feature.

D. Generation of 3D Geometric Features

3D structure of a face is estimated using the face feature points. 3D measurements of any three points on a face can be computed based on the perspective projection of a triangle. These three feature points derived from the eyes and the middle of the mouth

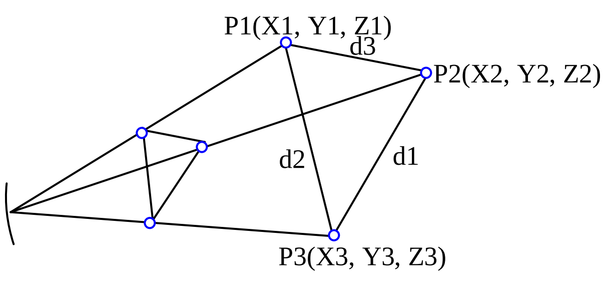


Figure 5. Illustration of Perspective Projection of a 3D Triangle

The 3D measurement of lengths d_1, d_2 and d_3 of all edges of the triangles are computed from the Equation (7).

$$d_1 = |P_2 - P_3|$$
$$d_2 = |P_3 - P_1|$$
$$d_3 = |P_1 - P_2| \tag{7}$$

$$d_1^2 = (X_2 - X_3)^2 + (Y_2 - Y_3)^2 + (Z_2 - Z_3)^2$$
$$d_1^2 = X_2^2 + X_3^2 - 2X_2X_3 + Y_2^2 + Y_3^2 - 2Y_2Y_3 + Z_2^2 + Z_3^2 - 2Z_2Z_3$$

$$d_2^2 = (X_3 - X_1)^2 + (Y_3 - Y_1)^2 + (Z_3 - Z_1)^2$$
$$d_2^2 = X_3^2 + X_1^2 - 2X_3X_1 + Y_3^2 + Y_1^2 - 2Y_3Y_1 + Z_3^2 + Z_1^2 - 2Z_1Z_3$$

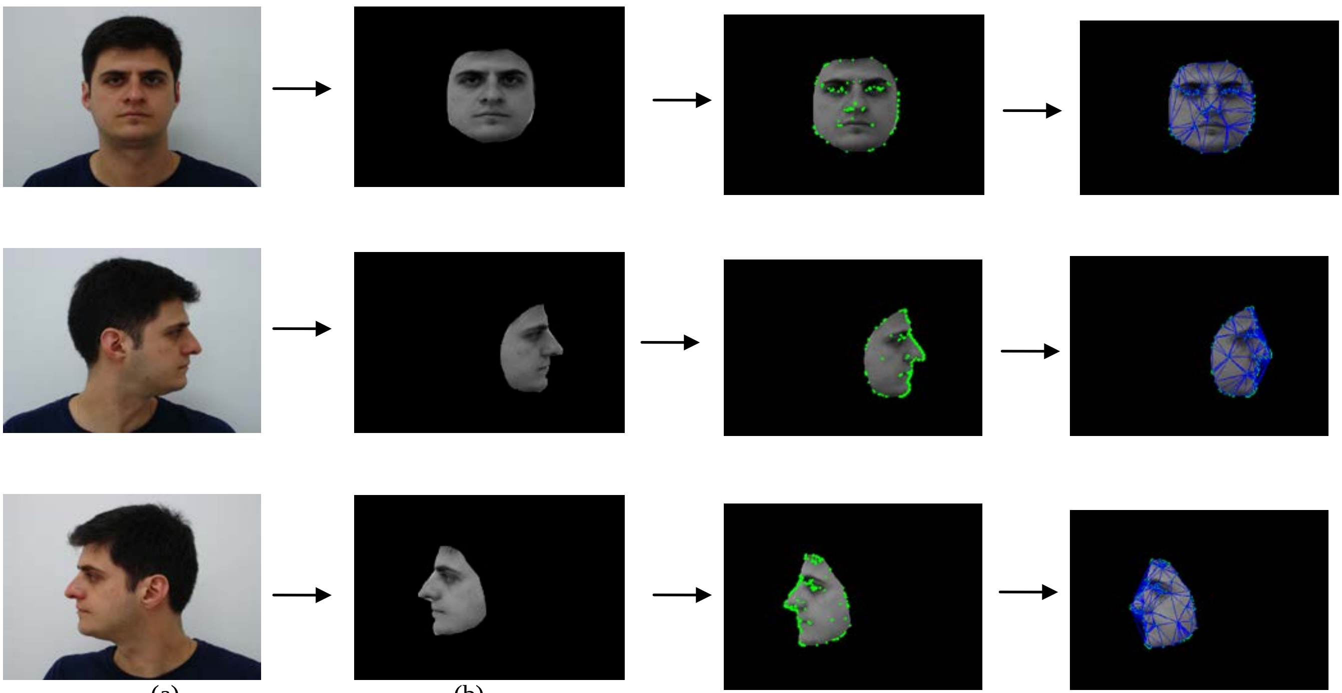
$$d_3^2 = (X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2$$
$$d_3^2 = X_2^2 + X_1^2 - 2X_1X_2 + Y_1^2 + Y_2^2 - 2Y_1Y_2 + Z_1^2 + Z_2^2 - 2Z_1Z_2$$

IV. EXPERIMENTS RESULTS

In our preliminary experiments, we have obtained some promising results for different facial expressions (see Figure. 6). Thus, in order to determine a six degrees of freedom pose vector, at least three points need to be found in the face, which firstly, have to be visible also in a range of perspectives and secondly, do not change during expression.
Even more, these points need to be well distributed in space and must be robustly and accurately detectable in the image.

V. CONCLUSION

In this paper we have proposed an approach for the facial feature detection which is efficient and fast to implement.
It provides a practical solution to the recognition problem. We are currently investigating in more detail the issues of robustness to changes in head size and orientation. Also we are trying to recognize the gender of a person using the same algorithm.



(a) (b) (c) (d)

Figure 6. Results of facial features extraction, (a) face image input, (b) face detection, (c) facial features extraction, (d) generation of 3D geometric features

REFERENCES

[1] S.P. Khandait, P.D. Khandait, and Dr.R.C.Thool, “An Efficient Approach to Facial Feature Detection for Expression Recognition”, International Journal of Recent Trends in Engineering, Vol 2, No. 1, November 2009.

[2] Mauricio Hess and Geovanni Martinez, “Facial Feature Extraction based on Smallest Univalue Assimilating Nucleus (SUSAN) Algorithm”.

[3] A. Yuille, D. Cohen, and P. Hallinan, “Facial feature extraction from faces using deformable templates”, Proc. IEEE Computer Soc. Conf. On Computer Vision and Pattern Recognition, pp. 104-109, 1989.

[4] Yingjie XIA, Kuang MAO, and Wei WANG, “A Novel Facial Feature Localization and Extraction Method”, Journal of Computational Information Systems 6:8(2010) 2779-2786
Available at http://www.Jofcis.com

[5] K. Susheel Kumar, Shitala Prasad, Vijay Bhaskar Semwal, and R C Tripathi, “*Real Time Face Recognition Using Adaboost Improved Fast PCA Algorithm*", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.2, No.3, July 2011.

[6] Antonini, G., V. Popovici and J. P. Thiran, ”*Independent Component Analysis and Support Vector Machine for Face Feature Extraction*”, Int. Conf. on Audio and Video-based Biometric Person Authentication, pp. 111-118, Guildford, UK, 2003.

[7] Vukadinovic, D. and M. Pantic, ”*Fully Automatic Facial Feature Point Detection using Gabor Feature Based Boosted Classifiers*”, IEEE Int. Conf. on Systems, Man and Cybernetics, Hawaii, October 2005.

[8] Chai Tong Yuen, M. Rizon, Woo San San, and M.Sugisaka, “*Automatic Detection of Face and Facial Features*”, Proceedings of the 7th WSEAS International Conference on SIGNAL PROCESSING, ROBOTICS and AUTOMATION (ISPRA '08).

[9] Sanqiang ZHAO and Yongsheng GAO, “*Towards Robust and Automated Head Pose Estimation: Elastic Energy Model*”, Biomedical Soft Computing and Human Sciences, Vol. 14, No. 1, pp. 21-26 (2009).