# An Automatic Landmark Localization Method for 2D and 3D Face

Junquan Liu[1,2], Feipeng Da[1,2(✉)], Xing Deng[1,2], Yi Yu[1,2], and Pu Zhang[1,2]

[1] School of Automation, Southeast University, Nanjing 210096, China
[2] Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Nanjing 210096, China
dafp@seu.edu.cn

**Abstract.** We propose an automatically and accurately facial landmark localization algorithm based on Active Shape Model (ASM) and Gabor Wavelets Transformation (GWT), which can be applied to both 2D and 3D facial data. First, ASM is implemented to acquire landmarks' coarse areas. Then similarity maps are obtained by calculating the similarity between sets of Gabor jets at initial coarse positions and sets of Gabor bunches modeled by its corresponding manually marked landmarks in the training set. The point with the maximum value in each similarity map is extracted as its final facial landmark location. It is showed in our laboratory databases and FRGC v2.0 that the algorithm could achieve accurate localization of facial landmarks with *state-of-the-art* accuracy and robustness.

**Keywords:** Facial landmark localization · Active shape model · Gabor wavelets transformation

## 1 Introduction

Exploring a method of automatic landmark detection with robustness and high accuracy is a challenging task, which plays an important role in applications such as face registration, face recognition, face segmentation, face synthesis, and facial region retrieval. Motivated by such practical applications, extensive work focusing on thinking of methods for automatic landmark localization both on 2D and 3D faces has been done. The active shape model (ASM) and active appearance model (AAM) which proposed by T.F. Cootes et al. [1, 2] can generate good performance on 2D faces. They are both based on point distribution model (PDM) [3], and principal component analysis (PCA) is applied to establish a motion model. An iterative search algorithm is implemented to seek the best location for each landmark. Compared with ASM's faster searching rate, AAM performs better on texture matching. However, both of the two algorithms suffer from illumination variation and facial expression changes.

Compared to landmark detection on 2D facial images, automatic landmark detection on 3D facial is a newer research topic. Zhao et al. [6] proposed a statistic-based algorithm, which detects landmarks by combining a global training deformation model, a local texture model of landmark, and local shape descriptor. Experiments on FRGC

v1.0 achieve a high successful rate of 99.09 %. In [7], Stefano et al. first detects nose tip based on gray value and curvature on range images, then Laplacian of Gaussian (LoG) operator and Derivative of Gaussian (DoG) operator are applied near the nose tip to locate nose ends, and Scale-Invariant Feature Transform (SIFT) descriptor is implemented to detect eye corners and mouth corners at last. However, the algorithm suffer from facial expression as well. Clement Creusot et al. [18] detected keypoints on 3D meshes which was based on a machine-learning approach. With many local features such as Local Volume (VOL), Principal Curvatures, Normals et al. combined together, Linear Discriminant Analysis (LDA) and AdaBoost algorithms are applied separately for machine-learning. The algorithm achieves state-of-the-art performance. However, the algorithm relies heavily on local descriptors. Jahanbin et al. [9] proposed a 2D and 3D multimodal algorithm based on Gabor features, which achieves high accuracy.

Inspired by Jahanbin et al., we propose a novel landmark detecting algorithm based on ASM and GWT in the rest of the paper, which can be applied to 2D and 3D facial images. The rest of this paper is organized as follows. Section 2 is the main body of our proposed algorithm. Firstly, we extend ASM to range image for landmarking; secondly, the notion of "Gabor jet" set and "Gabor bunch" set and their similarity definition are introduced in detail which are the main contributions of our work. Experimental results and evaluation are provided in Sect. 3. Several conclusions and future work are drawn in Sect. 4.

## 2 Accurate Facial Landmarks Localization Based on ASM and GWT

### 2.1 Abbreviations

Before we introduce our algorithm, several abbreviations of fiducials are listed here for convenience.

NT: nose tip; LMC: left mouth corner; RMC: right mouth corner; LEIC: left eye inner corner; LEOC: left eye outer corner; REIC: right eye inner corner; REOC: right eye outer corner.

### 2.2 Landmark Method

Once these definitions are given, the processing pipeline of this section is outlined as follows:

1. A set of 50 range or portrait facial images with different races, ages and genders is selected as the training set, and preprocessing steps are applied to all these images;
2. The "Gabor bunch" set features of 7 manually marked landmarks of each image in the training set are extracted;
3. A $25 \times 25$ pixel search area for each landmark is obtained based on the results of ASM;
4. Take LEIC searching for example, search counterclockwise and start from the center of the search area, and jump to step 6 if the similarity between the

"Gabor jet" set at the search point and the "Gabor bunch" set of LEIC in the training set is larger than 0.98 (experienced data), else repeat step 4 until all search points are visited and get a similarity map, as shown in Fig. 4.;

5. The point with the maximum value in each similarity map is extracted as its final facial feature location;
6. The algorithm is done.

## 2.3   Coarse Area Localization

Though general ASM algorithm which is usually applied in portrait images suffers from illumination variation and facial expression changes, it is a good choice for fast landmarking the coarse area of each landmark. However, few literatures about detecting these coarse areas on 3D face are presented. In our paper, we extend ASM to range image in order to solve this problem.

Prior to landmarks detecting, proper preprocessing steps on 2D and 3D data are appreciated. White balancing is implemented to all portrait images so as to aim a natural rendition of images [8] (see Fig. 1). 3D point clouds are converted to range images by interpolating at the integer x and y coordinates used as the horizontal and vertical indices, respectively. These indices are used to determine the corresponding z coordinate which corresponds to the pixel value on range image [17].



**Fig. 1.** White balancing algorithm: original and processed are in left and right respectively

### 2.3.1   ASM on Portrait and Range Image

As PDM [3] pointed out, an object can be described by n landmarks that are stacked in shape vectors. The shape vector of portrait and range image is constructed as:

$$x = \left[ p_{x,1}, \ldots, p_{x,n}, p_{y,1}, \ldots, p_{y,n} \right]^T \tag{1}$$

$(p_{x,i}, p_{y,i})$ refers to the coordinate in portrait and range image. In another way, a shape can be approximated by a shape model:

$$x \approx \bar{x} + \Phi b \tag{2}$$

Where $\bar{x}$ means the mean shape of all objects in the training set; the eigenvectors corresponding to the s largest eigenvalues $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s)$ $\lambda_i, i = 1, 2, \ldots, s$ are retained in a matrix $\Phi = \{\Phi_1, \Phi_2, \ldots, \Phi_s\}$; $b$ is the model parameter to be solved.

In order to get the parameter $b$, local texture models of each sample in the training set are constructed respectively as profile $g_i$, $i = 1, 2, \ldots, t$, then the mean profile $\bar{g}$ and the covariance matrix $C_g$ are computed for each landmark. Given $g_i$ obeys a multidimensional Gaussian distribution, the computation of the Mahalanobis distance between a new profile and the profile model can be defined as follows:

$$f_{Mahalanobis}(g_i) = (g_i - \bar{g})^T C_g^{-1}(g_i - \bar{g}) \tag{3}$$

Both shape model and profile model are combined to search each landmark, and points with the minimum values of $f(g_i)$ are the landmark to find.

## 2.4  Fine Landmark Detection

Since ASM performs poor in situations such as illumination variation, eye-closing and facial expression changes, we extract a $25 \times 25$ pixel area for each landmark centered at its coarse location as the fine search area. Then Gabor Wavelet Transformation (GWT) is applied to these search area to get the final location of each landmark.

### 2.4.1  Gabor Jets

Gabor filters which are composed by Gabor kernels with different frequencies and orientations can be formulated as:

$$\psi_j(\vec{x}) = \frac{k_j^2}{\sigma^2} \exp\left(\frac{-k_j^2 x^2}{2\sigma^2}\right)\left[\exp\left(i\vec{k}_j \cdot \vec{x}\right) - \exp\left(\frac{-\sigma^2}{2}\right)\right] \tag{4}$$

$$\vec{k}_j = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos\phi_u \\ k_v \sin\phi_u \end{pmatrix}, \quad k_v = 2^{\frac{-(v+2)}{2}}\pi, \quad \phi_u = \frac{\pi}{8}u \tag{5}$$

In our implementation, $\sigma = 2\pi$, $\vec{x}$ is the coordinate of a point, $u \in \{0, 1, 2, \ldots, 7\}$ and $v \in \{0, 1, 2, \ldots 4\}$ determine the orientation and scale of the Gabor filters.

The GWT is the response of an image $I(\vec{x})$ to Gabor filters, which is obtained by the convolution:

$$J_j(\vec{x}) = \int \vec{I}(\vec{x'})\psi_j(\vec{x} - \vec{x'})d^2\vec{x'} \tag{6}$$

Where $J_j(\vec{x})$ can be expressed as $J_j = a_j \exp(i\phi_j)$, $a_j(\vec{x})$ and $\phi_j(\vec{x})$ denote the magnitude and phase respectively. The common method for reducing the computational cost for the above operation is to perform the convolution in Fourierspace [11]:

$$J_j = F^{-1}\{F(\psi_j(\vec{x}))F(I(\vec{x}))\} \tag{7}$$

Where $F$ denotes the fast Fourier transform, and $F^{-1}$ its inverse. The magnitude of GWT on range image is shown in Fig. 2.
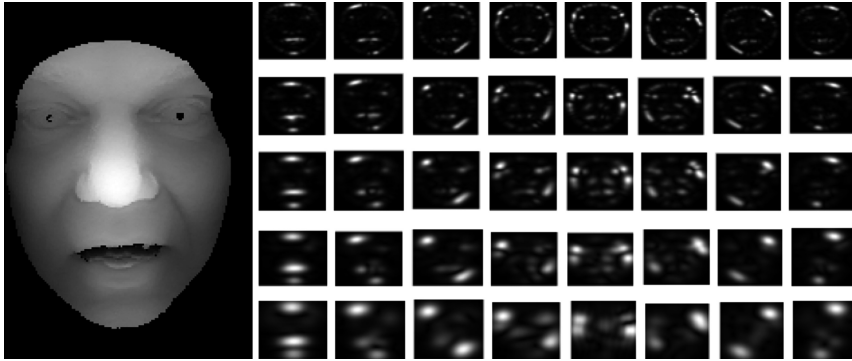
**Fig. 2.** GWT on range image: range image is in left, magnitude of GWT is in right

As proposed in [4], "Gabor jet" is a set $\vec{J} = \{J_j, j = u + 8v\}$ of 40 complex Gabor coefficients obtained from a single image point. And the similarity between two jets is defined by the phase sensitive similarity measure as follows:

$$S(\vec{J}, \vec{J}') = \frac{\sum_{i=1}^{40} a_i a_i' \cos(\phi_i - \phi_i')}{\sqrt{\sum_{i=1}^{40} a_i^2 \sum_{i=1}^{40} a_i'^2}} \tag{8}$$

Here, $S(J_j, J_j') \in [-1, 1]$ and a value which is closer to 1 means a higher similarity.

### 2.4.2 "Gabor Jet" Set and "Gabor Bunch" Set

For any given landmark $j$ in the image, its Gabor bunch is defined as $\vec{B}_j = \{J_{x,j}\}$ [12], where $x = 1, 2, \cdots, n$ denotes different images in the training set. In [13], a set of 50 facial images with different races, ages and genders was selected as the training and Gabor bunch extraction to detect landmarks. This algorithm achieves a good result on portrait image, but performs relatively poor on range image.

Inspired by "Gabor jet" and "Gabor bunch", we propose the notion of "Gabor jet" set and "Gabor bunch" set in this work. For any given landmark $j$, a "Gabor bunch" set is defined as a set constructed by the "Gabor jet" of $j$ and its eight neighborhood; "Gabor bunch" set is a set constructed by all different images' "Gabor bunch" set of $j$ in the training set. Take LEIC for example, its "Gabor jet" set is $\vec{J}_s = \{\vec{J}_0; \vec{J}_i, i = 1, 2, \ldots 8\}$, where $\vec{J}_0$ is the "Gabor jet" of LEIC, $\vec{J}_i$ is the "Gabor jet" of its 8 neighborhoods (see Fig. 3). $\vec{B} = \{\vec{J}_{s,i}, i = 1, 2, \ldots, n\}$ is the "Gabor bunch" set of LEIC, where $\vec{J}_{s,i}$ means different images' "Gabor jet" set of LEIC in the training set.

The similarity between any two "Gabor jet" sets is defined as:

$$S(\vec{J}_s, \vec{J}_s') = \frac{1}{9} \sum_{i=0}^{i=8} S(\vec{J}_i, \vec{J}_i') \tag{9}$$

Fig. 3. "Gabor jet" set of LEIC


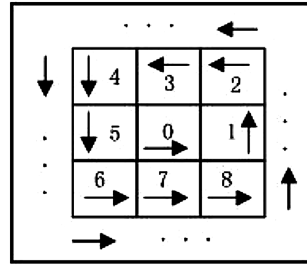
Fig. 4. Searching path

Considering that probes are often not contained in the training set, we define the similarity between a "Gabor jet" set and a "Gabor bunch" set as:

$$S_{B'}(\vec{J}, \vec{\vec{B'}}) = SORT_n^{\alpha}\{S(\vec{J}_s, \vec{J}'_{s,i}), i = 1, 2, \ldots, n\} \tag{10}$$

Where $SORT_n^{\alpha}\{f_i, i = 1, 2, \ldots, n\}$ denotes the sum of $\alpha$ largest values of $f_i$.

## 3    Experimental Results and Evaluation

In this section, we present the performance of our proposed approach when tested both on the FRGC v2.0 and our laboratory database (OLD).

FRGC v2.0 [14] is one of the largest available public human face datasets, which consists of about 5,000 recordings covering a variety of facial appearances from surprise, happy, puffy cheeks, and to anger divided into training and validation partitions. The training partition which is acquired in Spring 2003 includes 273 individuals with 943 scans. The validation partition for assessing performance of an approach includes 4007 scans belonging to 466 individuals (Fall 2003 and Spring 2004). Since there was a significant time-lapse between the optical camera and the operation of the laser range finder in the data acquisition, 2D and 3D images are usually out of correspondence [9].

OLD consists of 247 scans from 83 individuals with expression, eye-closing, and pose variation. All data are acquired by our laboratory self-developed 3D shape measurement system based on grating projection, where a certain map exists between 2D and 3D data.

### 3.1    Experiment on Portraits

Two training sets are needed in our work. The ASM training set for finding search areas which consist of 1040 portraits with different expressions, pose, and illumination is selected from FRGC v2.0 and OLD at the rate of 25:1; the Gabor training set for fine landmark detection includes 50 portraits covering facial appearance from subjects with different gender, age, expression, and illumination, which is selected in the same ratio

as the ASM training set. The probe set consists of 895 portraits selected at random from the FRGC v2.0 and the remaining portraits in the OLD. All portraits are normalised to $320 \times 240$ pixel.

The Euclidean distance between the manually marked and the automatically detected landmarks is measured as $d_e$ so as to evaluate the accuracy of the algorithm. *mean* and std are the mean and standard deviation of $d_e$ respectively. The positional error of each detected fiducial is normalized by dividing the error by the interocular distance of that face. The normalized positional errors averaged over a group of fiducials is denoted $m_e$, adopting the same notation as in [5].

Compared with general ASM algorithm which suffers from eye-closing, illumination and expression, the algorithm we proposed performs well in such situations as shown in Fig. 5. It is evident from the statistic in Table 1 that LEIC, LEOC, and LMC which have significant texture features perform best in sharply contrast with NT that performs worst. Moreover, the algorithm we proposed achieves a high successful ratio of 96.4 % with just 1.41 mm of *mean* and 1.78 mm of std in FRGC v2.0. By contrast, it performs even better in OLD that the successful ratio is 97.7 %, that *mean* and std are 1.31 mm and 1.49 mm respectively. The reason why the results in OLD outperform



**Fig. 5.** Landmark detection on portraits: the first row is FRGC v2.0, the second row is OLD

FRGC v2.0 may be that there are less probes and facial variation in OLD.

## 3.2  3D Landmark Localization Experiment

Due to the definitely corresponding relationship between 2D and 3D facial datum in OLD, accurate landmark localization can be realized by mapping the 2D localization results to 3D point cloud according to the corresponding relation-ship. However, the corresponding relationship between 2D and 3D faces does not exist in FRGC v2.0. Therefore, to achieve a precise 3D landmark localization, the 3D point cloud is converted into a range image at first, and then ASM is used to locate landmarks coarsely in the range image. Finally, the best landmarks can be found by extracting Gabor wavelet feature in the coarse location area. Both the training set and test set are range images and the selection rules are as described in Sect. 3.1. The results of 3D landmark

**Table 1.**  statistic landmark result on portraits

| Landmark | FRGC v2.0 | | | OLD | | |
|---|---|---|---|---|---|---|
| | mean | std | $m_e \leq 0.06$ | mean | std | $m_e \leq 0.06$ |
| **LEIC** | **1.39** | **1.44** | **98.1 %** | **1.43** | **1.66** | **98.5 %** |
| **LEOC** | **1.47** | **1.57** | **97.0 %** | **1.57** | **1.48** | **98.8 %** |
| REIC | 1.64 | 2.41 | 94.2 % | 1.61 | 2.39 | 95.7 % |
| REOC | 1.41 | 1.83 | 96.1 % | 1.04 | 1.27 | 97.9 % |
| NT | 1.33 | 1.86 | 94.3 % | 1.36 | 1.13 | 97.1 % |
| **LMC** | **1.28** | **1.55** | **98.4 %** | **0.99** | **1.07** | **99.1 %** |
| RMC | 1.35 | 1.90 | 97.1 % | 1.19 | 1.49 | 97.4 % |
| average | 1.41 | 1.78 | 96.4 % | 1.31 | 1.49 | 97.7 % |

localization in FRGC v2.0 are as shown in Fig. 6 and Table 2. In OLD, we just need to map the 2D localization results to 3D point cloud according to the corresponding relation-ship, so the localization results are as shown in Table 1.

The first row is in FRGC v2.0, the second row is that in OLD. Tables 1 and 2 show that the algorithm in this paper can also achieve better results in 3D landmark localization: in the case of a small error $m_e \leq 0.06$, the accuracy can be 94.2 % in FRGC v2.0. Meanwhile, in OLD, because of the corresponding relationship, we can directly locate landmarks on portrait images which have richer texture information, then we can
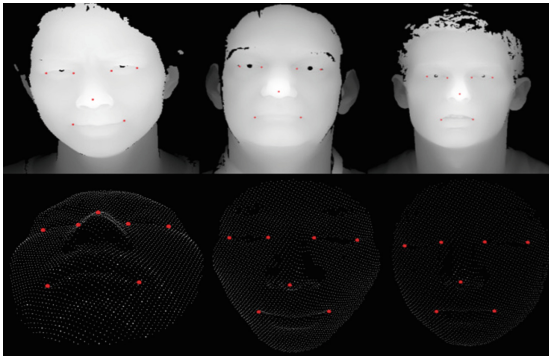


**Fig. 6.**  3D landmark localization in FRGC v2.0 (first row) and OLD (second row)

**Table 2.**  3D landmark localization results in FRGC v2.0

| Landmark | LEIC | LEOC | REIC | REOC | NT | LMC | RMC | Average |
|---|---|---|---|---|---|---|---|---|
| mean | 1.54 | 2.74 | 1.94 | 2.55 | 1.29 | 1.78 | 1.66 | 1.92 |
| std | 1.62 | 2.72 | 2.50 | 2.61 | 1.00 | 1.45 | 1.30 | 1.88 |
| $m_e \leq 0.06$ | 96.1 % | 91.7 % | 96.3 % | 92.4 % | 96.9 % | 91.9 % | 94.7 % | 94.2 % |

map the localization results to the 3D point cloud. Finally, the accuracy we achieve is 97.7 %.

### 3.3 Landmark Localization Results Comparing

The landmark detection algorithm proposed in this paper first locates the landmarks coarsely basing on ASM, then the accurate localization is accomplished by the use of Gabor features. The algorithm can overcome the shortage of the ASM that the reduced accuracy is caused by eyes slightly open, eyes closed, large expressions and obvious illumination changes and so on. Furthermore, the landmark localization performance is greatly improved compared with ASM. AAM is also a sophisticated 2D facial landmark localization algorithm. As is pointed out in literature [15], the accuracy of AAM algorithm to locate the landmarks is only 70 % under the circumstance that error satisfies $m_e \leq 0.1$. However, the accuracy of the algorithm locating landmarks in 2D faces in both literature [13] and this paper is over 99 % under the same circumstance. But compared with the automatic landmark localization algorithm pro-posed in this paper, the algorithm in literature [13] is only semi-automatic because its coarse landmark localization area cannot be completely determined automatically. Figure 7 describes the relationship between permissible error $m_e \leq 0.1$ and the accuracy of 2D landmark localization resulting from the algorithm in literature [13] and this paper. We can also conclude that the algorithm in this paper has a better performance than that of the algorithm in literature [13] from it.

   Table 3 displays the results of the comparison among three 3D landmark detection algorithms. The literature [16] used the Shape Index and Spin Image as local descriptors to locate landmarks coarsely, and FLM is used as topological constraint among landmarks, then the group of landmarks which meets FLM model in the coarse location area is chosen as the best landmarks. This method can also locate each landmark well in the case of large deflection in faces, but its average error is larger than that of the method in this paper from Table 3. The 3D landmark-locating algorithm in literature [13] also shows a good performance, and has a higher locating accu-
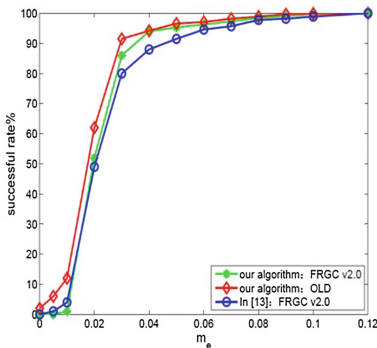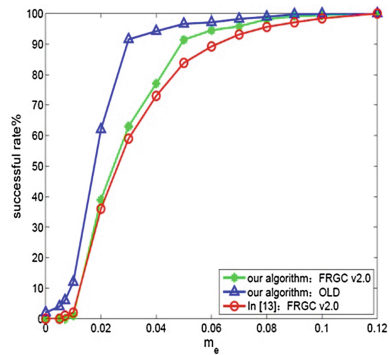


**Fig. 7.** The relationship between permissible

**Fig. 8.** The relationship between permissible error and in accuracy error and accuracy in 3D

racy than that in literature [16]. However, compared with the automatic landmark localization algorithm pro-posed in this paper, the 3D landmark detection algorithm in literature [16] is only semi-automatic. What's more, Fig. 8 shows that the algorithm in

**Table 3.** the comparison of average error of three 3D landmark detection algorithms

|       | LEIC | LEOC | REIC | REOC | NT   | LMC  | RMC  |
|-------|------|------|------|------|------|------|------|
| In[16] | 4.41 | 5.83 | 5.03 | 5.79 | 4.09 | 5.42 | 5.65 |
| In[13] | 1.65 | 3.93 | 1.63 | 3.77 | 1.40 | 1.89 | 1.81 |
| Our   | 1.54 | 2.74 | 1.94 | 2.55 | 1.29 | 1.78 | 1.66 |

this paper has a higher locating accuracy than that in literature [13]. Figure 9 also shows 3D landmark localization can be realized by transforming it into corresponding 2D landmark localization to improve locating accuracy due to the mapping relationship between 2D and 3D images.

## 4　Conclusion

This paper proposes an automatic facial landmark localization algorithm on the basis of ASM and GWT. The method is based on the coarse localization using ASM, then the concepts of "Gabor jet" set and "Gabor bunch" set are introduced. Hence, the precise localization of the facial landmarks can be realized. The experimental results show that the landmark localization algorithm proposed in this paper has a higher accuracy and is robust to facial expressions and illumination changes. The algorithm has the following characteristics: (1) it is suitable for both 2D and 3D facial data; (2) it is based on ASM coarse localization, then the accurate localization is realized because of the use of "Gabor jet" set and "Gabor bunch" set. The localization accuracy is significantly increased compared with ASM.

GWT algorithm which has been used to locate landmarks accurately in this paper is rotation invariant. If ASM algorithm can be improved to realize coarse localization, the accurate landmark localization in faces having arbitrary pose can be achieved. This paper uses 5 directions, 8 scales and 40 Gabor filters to make up a filter group. If coupling filters in it can be reduced, the instantaneity of the algorithm in this paper will be better.

# References

1. Cootes, T.F., Taylor, C.J., Hill, A., Halsam, J.: The use of active shape models for locating structures in media images. Image Vis. Comput. **12**(6), 355–366 (1994)
2. Cootes, T.F., Edwards, G., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 681–685 (2001)
3. Cootes, T.F., Taylor, C.J.: Active shape models-smartsnakes. In: Proceedings of British Machine Vision Conference, Leeds, pp. 266–275. BMVC, UK (1992)
4. Clark, M., Bovik, A., Geisler, W.: Texture segmentation using Gabor modulation/demodulation. Pattern Recognit. Lett. **6**(4), 261–267 (1987)
5. Cristinacce, T., Cootes, D.: A comparison of shape constrained facial feature detectors. In: Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004, pp. 375–380 May 17–19 (2004)
6. Zhao, X., Dellandréa, E., Chen, L.: A 3D statistical facial feature model and its application on locating facial landmarks. In: Blanc-Talon, J., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2009. LNCS, vol. 5807, pp. 686–697. Springer, Heidelberg (2009)
7. Berretti, S., Amor, B.B., Daoudi, M., et al.: 3D facial expression recognition using SIFT descriptor so automatically detected keypoints. Visual Comput. **27**(11), 1021–1036 (2011)
8. Ancuti, C.O., Ancuti, C.: Single image dehazing by multi-scale fusion. IEEE Trans. Image Process. **22**(8), 3271–3282 (2013)
9. Jahanbin, S., Choi, H., Bovik, A.C.: Passive multimodal 2-D + 3-D face recognition using Gabor features and landmark distances. IEEE Trans. Inf. Forensics Secur. **6**(4), 1287–1304 (2011)
10. Daugman, J.G.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. B **2**(7), 1160–1169 (1985)
11. Hj Wan Yussof, W.N.J., Hitam, M.S.: Invariant gabor-based interest points detector under geometric transformation. Digt. Sig. Proc. **25**(10), 190–197 (2014)
12. Wiskott, L., Fellous, J.M., Kuiger, N., der Malsburg von, C.: Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 775–779 (1997)
13. Jahanbin, S., Jahanbin, R., Bovik, A.C.: Passive three dimensional face recognition using iso-geodesic contours and procrustes analysis. Int. J. Comput. Vis. **105**(1), 87–108 (2013)
14. Phillips, P.J., Flynn, P.J., Scruggs, T. et al.: Overview of the face recognition grand challenge. In: Computer Vision and Pattern Recognition CVPR, 2005, pp. 947–954. San Diego, CA, USA (2005)
15. Cristinacce. T., Cootes, D.A.: Comparison of shape constrained facial feature detectors. In: Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 375–380. Los Alamitos, CA, USA (2004)
16. Perakis, P., Passalis, G., Theoharis, T., et al.: 3D facial landmark detection under large yaw and expression variations. IEEE Trans. Pattern Anal. Mach. Intell. **35**(7), 1552–1564 (2013)
17. Lei, Y., Bennamoun, M., Hayat, M., et al.: An efficient 3D face recognition approach using local geometrical signatures. Pattern Recognit. **47**(2), 509–524 (2014)
18. Creusot, C., Pears, N., Austin, J.: A machine-learning approach to keypoint detection and landmarking on 3D meshes. Int. J. Comput. Vis. **102**(1–3), 146–179 (2013)