Restricted Shape and Abbreviated Appearance based Correlation Model for Detecting Facial Landmarks

Kutiba Nanaa, Mohd Nordin Abd Rahman, and Mohamed Rizon

***Abstract*—identifying facial landmarks is a fundamental process in face recognition. Despite the abundance of studies, the published methods still suffer from some limitations because of a variety of challenges. In this paper, we realized that some landmarks could be easily detected. We labeled these landmarks as keys whereas the proposed improvements rely on them. The proposed improvements include restricting the shape based model by the key landmarks in initialization and updating phase. To overcome the limitation in boundary landmarks, an abbreviated appearance model has been combinedwith. All image comparisons are performed by composite cross-correlation to guarantee the stronger matching. The experiments are applied on PICS dataset. We compared the results with a comparative study involving various detection models. The results showed that the model's ability to reduce RMS of detection by the proposed model, whether for internal and boundary landmarks.**

***Index Terms*—Landmarks Detection, Template Matching, Shape based Model, Appearance based Model, Face recognition.**

I. Introduction

The issue of detecting and analyzing facial landmarks is a foundational process in face recognition. Though various studies have been debates to resolve this problem, it still needs to achieve a reliable detection of facial landmarks. However, facial landmarks are important to describe face shape and appearance representation. By default, facial landmarks locate on edges of the facial features and facial boundary.

There are several well-known filters of image processing which are utilized for landmarks detection. For example, Canny, Laplacian, Sobel, and Gabor are used for this purpose. Landmarks detecting by filters are debated in various comparative studies [1, 2 & 3]. However, providing global information on the geometric features still a necessary step to guarantee as mooth performance of detecting based on the filters.

Matching templates is one of the methods that can be used to detect landmarks. Typically, it is utilized to identify facial feature rather than the Landmarks. For example, it is employed to detect eye in both cases (open and closed) [4]. In subsequent enhancement, the composite template were introduced for eye detection [5]. Establishing the templates can be generalized to the rest of features. In the case of detecting landmarks, there is a need to direct the search and comparison. The search is usually done through face models that maintain consistency.

To control searching for landmarks, Active Shape-based Model (ASM) is utilized. The primary function of ASM is to represent the shape in statistically presentation. In detection term, ASM matches the model to the given image. And then it performs a recursive process of researching and updating. Searching is made around each landmark for a better position for that landmark. Updating is done for the model parameters to get the best statistical match. The compression process in searching term is performed by Mahalanob is distance [6, 7].

In subsequent improvement, initialization of ASM is performed based on the eyes center, feature map, and color information [8]. It also re-defines the shape by adding more fitting landmarks and provide profile shape model. This improvement increased the success ratio over 10.2 %, comparing to the standard ASM.

In other contribution, Active Appearance Model (AAM) has been utilized in landmark detection [9, 10 & 11]. The methods used AAM is classified as holistic methods. They used the whole texture of face area in a linear generative model. AAM requires a recursive fitting process in comparison of the given with the training set. However, AAM is sensitive to the illumination changes and bias towards the mean face. Also, it may fail to be efficient inan image which is out of the training set or taken in with a low resolution.

In this paper, we assumed that some landmarks are easy to detect. We have realized this assumption through applying composite cross-correlation for all the landmarks. Then we labeled two landmarks of each feature as keys. Accordingly, we enhanced ASM by restricting to the key landmarks. The restricted Shape Correlation Model (RSCM) ensure aproper initialization for the landmarks close to the target. The shape adjustment is made according to the result of comparison with the orthogonal regions.

Due to the lack of the general appearance in RSCM, it faces a challenge in detecting some landmarks, especially which lie on the boundary. Therefore, we combine AAM withAbbreviated Appearance based Correlation Model (AACM). The abbreviation is achieved by taking four samples of the dataset as training samples. Furthermore, the comparison is accomplished by direct composite cross-correlation and ignoring the appropriate step.

By combining RSCM and AACM, we propose Restricted shape and Abbreviated Appearance based Correlation Model (RSAACM). The obtained results of RSAACM are compared to some detection models involved in the comparative study [12]

II. Key Facial Landmarks

Facial landmarks play a discriminatory role in the issues related to face recognition. Usually, they located on the edges of facial features such as eyes, eyebrows, nose, lips and face boundary. In this paper, we defined 74 landmarks as shown in

Kutiba Nanaa, Ph.D. Candidate, Faculty of Informatics and Computing,Universiti Sultan Zainal Abdin, Terengganu, Malaysia.

the Fig. 1.

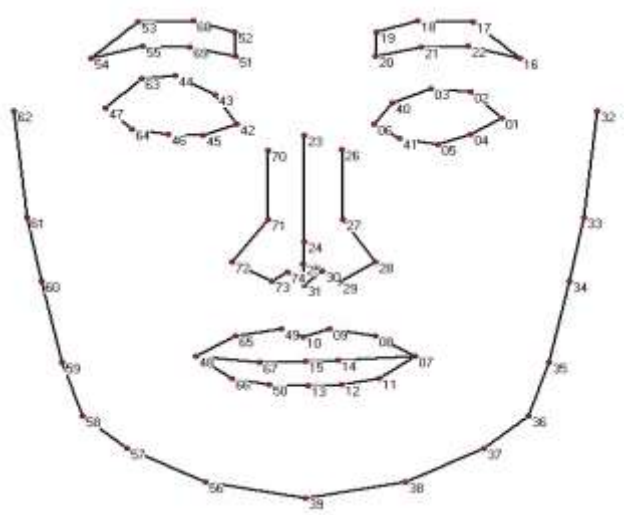

Fig. 1. The face shape and landmarks.

We assume that template matching can easily detect some landmarks. This detection is done regardless of the overall shape of the face or any prior information. For this detection, we used the composite cross correlation where it gives more accurate detection results. We create a composite template for each landmark whereas it contains four sub-templates. These sub-templates ($[T_1, T_2, T_3, T_4]$),are taken in sizes 20×20, 30×30, 40×40 and 50×50. The matching is performed according to (1).

$$Co(I,T,)_{(x,y)} = \frac{1}{2} \times \left(1 + \sqrt[4]{\prod_h^4 Co'(I,T_h)_{(x,y)}}\right) \qquad (1)$$

$$where$$

$$T = [T_1, T_2, T_3, T_4],$$

$$Co'(I,T_h)_{(x,y)} = \frac{Co''(I,T_h)_{(x,y)}}{max(Co''(I,T_h))},$$

$$Co''(I,T_h)_{(x,y)} = \frac{\sum_s \sum_t \delta_{I(x+s,y+t)}\, \delta_{T_h(s+cx,t+cy)}}{\sqrt{\sum_s \sum_t \delta^2_{I(x+s,y+t)}}\sqrt{\sum_s \sum_t \delta^2_{T_h(s+cx,t+cy)}}}$$

$$\delta_{I(x+s,y+t)} = I(x+s, y+t) - \bar{I}(x,y),$$

$$\delta_{T_h(s+cx,t+cy)} = T_h(s+cx, t+cy) - \overline{T_h},$$

$$\bar{I}(x,y) = \frac{1}{size\ of(\overline{T_h})} \sum_s \sum_t I(s+x, t+y),$$

$$\overline{T_h} = \frac{1}{size\ of(\overline{T_h})} \sum_s \sum_t T_h(s+cx, t+cy).$$

After applying the matching on all landmarks, we found that some landmarks respond smoothly to their templates. Two key landmarks are addressed of each feature, as the easiest detection and RMS of detection is no more than 3 pixels. Otherwise, all of the landmarks of face boundary are ignored because the RMS exceeded 3 pixels. Fig. 2 demonstrates the selected key landmarks in red and the rest landmarks in blue.

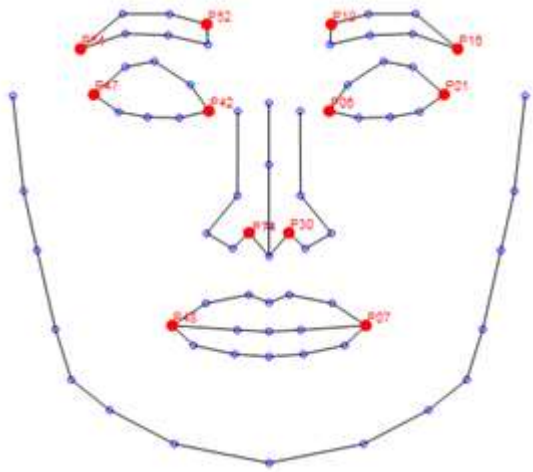

Fig. 2. The selected key landmarks.

III. Restricted Shape based Correlation Model (RSCM)

As an enhancement of ASM, Restricted Shape-based Correlation Model (RSCM) follows the same strategy to search for landmarks. The difference is that the key landmarks restrict the search direction. Also, the comparison is made by multi-resolution of composite cross-correlation.

The initialization process is performed for landmarks of each feature separately. For example, to estimate the initial position of landmark P_i we normalize it according to the two key landmarks (KPS, KPE) of the same feature. By using the relative distances $(\overline{P_i}, \overline{KPS}, \overline{KPE})$ of mean shape which are provided in Fig. 1, the estimated position is given by (2).

$$P_{i(x,y)} = KPS_{(x,y)} + \Delta KP \times \left(\overline{P_{i(x,y)}} - \overline{KPS_{(x,y)}}\right) \qquad (2)$$

$$where$$

$$\Delta KP = \frac{KPE_{(x)} - KPS_{(x)}}{\overline{KPE_{(x)}} - \overline{KPS_{(x)}}}$$

After the initialization process, all of the landmarks became close to the target. It saves time in the search process. The search process generates orthogonal lines on the Landmarks. Each landmark P_i is linked to previous $Pv(P_i)$ and next $Nx(P_i)$ landmarks. We find the angle θ between these landmarks $Pv(P_i)$, P_i , $Nx(P_i)$. And then we draw the line that halves the angle. This line is known as orthogonal line$OrtRng(P_i)$. We consider only 20 pixels of the orthogonal line whereas the center corresponds to the landmark. However these pixels are given by (3).

$$OrtRng(P_i) = \{OP_j\} \qquad (3)$$

$$where$$

$$\sqrt{\left(OP_{j(x)} - P_{i(x)}\right)^2 + \left(OP_{j(y)} - P_{i(y)}\right)^2} = j;$$

$$OP_{j(x)} = P_{i(x)} + \left(tan(\theta) \times \left(OP_{j(y)} - P_{i(y)}\right)\right);$$

$$\theta = \frac{1}{2} \times \left(tan^{-1}\frac{P_{i(x)} - Pv(P_i)_{(x)}}{P_{i(y)} - Pv(P_i)_{(y)}} + tan^{-1}\frac{Nx(P_i)_{(x)} - P_{i(x)}}{Nx(P_i)_{(y)} - P_{i(y)}}\right);$$

$$j = \{1,2,\ldots,10\}.$$

The searching process and updating shape is performed for each Landmark according to the corresponding composite template and the orthogonal lines. Each time the configuration is modified, new orthogonal lines are generated. The searching is repeated until there is no update to the last form. It should be noted that the updating excludes the key landmarks as they were detected by direct applying of composite cross-correlation on the given image. Fig. 3. shows the intermediate results of applying RSCM on a sample of the database, including the return shape after the initialization process (A), orthogonal lines (B), and the final performance form after searching and updating (C).

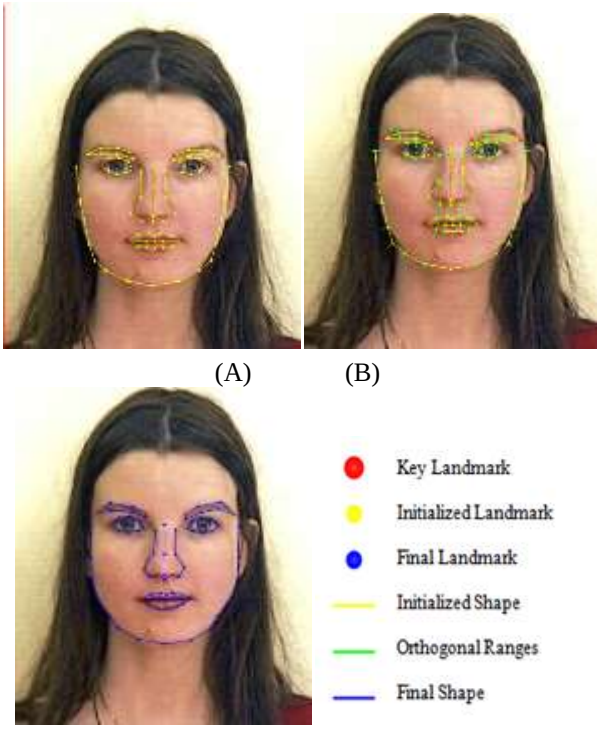

Fig. 3: the intermediate results of applying RSCM on a sample of the database.

IV. Abbreviated Appearance based Correlation Model (AACM)

The performance of RSCM is robust to detect most of the facial landmarks. However, it has limitation especially in identifying the boundary landmarks. The reason is that we did not determine easy-to-detect key landmarks through composite cross-correlation. Consequently, there is no proper initialization has occurred for these landmarks. Furthermore, the lack of general information on facial appearance has also had an adverse effect. In addition to RSCM searches for landmarks in a narrow range, only orthogonal areas with their current location. So, the idea is to reinforce the performance with a model that handles the overall appearance such as AAM.

The limitation of AAM lies in the fact that it takes a long time to implement, requires additional fitting operations, and fails in treatment an image from out of the training samples. So the idea of AACM is to abbreviate AAM model by choosing only four images as training examples. And then test the sample which is most closely associated with the given image by using composite cross- correlation. Based on this, the locations of the landmarks are estimated.

After selecting the most closely related sample, a displacement of the correlation matrix is performed by the landmarks position in the sample image. We, therefore, used the displacement parameter F_i as shown in (4). whereas the displacement indicates to the distance between the center of the sample and the landmark position in the sample

$$AACM(I,P_i)_{(x,y)} = Co(I,T,F_i)_{(x,y)} \quad (4)$$
$$where$$
$$T = [T_1, T_2, T_3, T_4],$$
$$F_i = [F_{T_1,i}, F_{T_2,i}, F_{T_3,i}, F_{T_4,i}]$$
$$F_{h,i} = (fx_{h,i}, fy_{h,i}).$$

The term of $Co(I,T,F_i)_{(x,y)}$ in Eq. 4, can be realized by (1), after upgrading to (5).It includes the increment parameter while retaining the rest of the parameters

$$Co(I,T,F)_{(x,y)} = \frac{1}{2}\left(1 + \sqrt[4]{\prod_h^4 Co'(I,T_h)_{(x+fx_h,y+fy_h)}}\right) \quad (5)$$

Fig. 4, shows sample result of applying AACM, (A) indicates a given image, (B) indicates facial area which AACM applied on, (C) indicates the correlation image, (D, E, F, G, H) indicates the correlation images after increment according to five landmarks P_{14},P_{33},P_{37},P_{39}, and P_{43}, respectively, and (I) indicated the result of suggested positions in yellow circles while the red circles indicate the annotated landmarks.

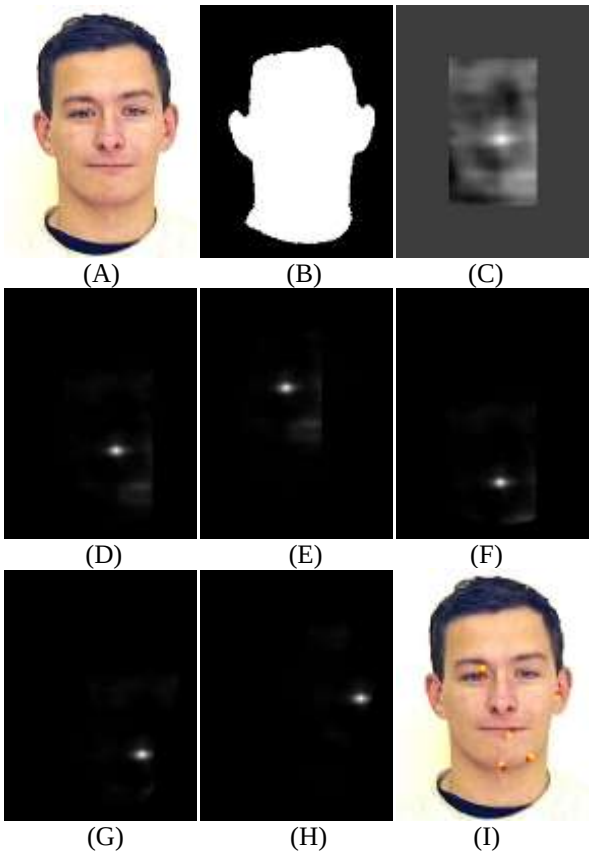

Fig.4.the sample result of applying AACM.

V. Restricted Shape and Abbreviated Appearance based Correlation Model (RSAACM)

In general, the performance of RSCM is much robust than the performance of AACM. The main reason due to the ability of RSCM to handle samples from our training set. But it lacks general information of the face appearance which may return a distorted shape or out of the target. Hence, the idea is to enhance the performance by utilizing AACM. As a similar result, AAM is more durable than AACM, but we abbreviated it to save time as its task is to guarantees that the search for the target is not lost. However, the most limitation of RSCM has encountered in detecting boundary landmarks. Thus the AACM is much activated for these landmarks. On the other hand, AACM activation is reduced for the rest landmarks. The contribution of AACM in detection formula is shown in (6), which discusses the compatibility of the landmarks of its surrounding area to update the shape.

$$Lct(I,P_i) = OrtRng(I,P_i) \times \left((1-\propto_{P_i})C_{P_i} + \propto_{P_i} {C_{P_i}}^2 AACM(I,P_i)\right) \quad (6)$$

$$where$$

$$C_{P_i} = Co(I,T_{P_i});$$

$$\propto_{P_i} = \begin{cases} 0.9 & P_i \in boundry \\ 0.1 & otherwise \end{cases}$$

The flowchart of an overall process of RSAACM is shown in Fig. 5. After determining the key landmarks, RSCM and AACM are performed, in parallel. Updating the shape is carried out based on the combination of the output of RSCM and AACM as shown in Eq. 6. After each updating, RSAACM re-searches in the orthogonal regions. Updating and searching continue until there is no change in the last shape.

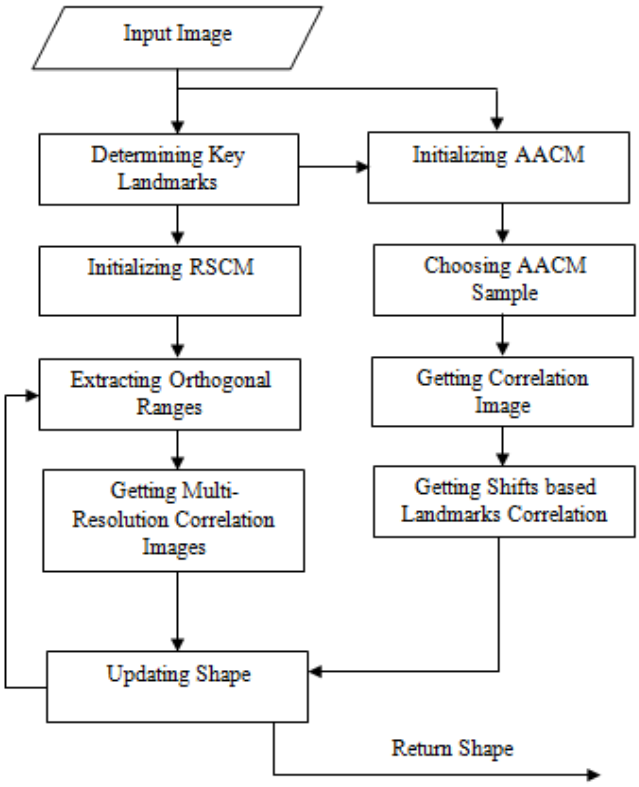

Fig.5. Flowchart of RSAACM.

In Fig.6., there are the returned shapes by applying RSCM, AACM, and RSAACM. As noted, (A) the RSCM detect the internal landmarks efficiently. In contrast, (B) AACM quickly estimate the boundary landmarks. As a result, (C) RSAACM combines the advantages of RSCM and AACM in the detection of all the landmarks.

(A) (B) (C)

Fig. 6. The returned facial shape by RSCM, AACM, and RSAACM.

VI. Comparison Results

Landmarks play a vital role in facial shape representation. The number and locations of landmarks that can be labeled are flexible through previous studies. The number of landmarks is usually 8 to 194 [13&14]. To reconcile our proposed model with previous models, the results of the previous models have been obtained on the 74 landmarks that are labeled in this paper. For further details on the results, the landmarks have been categorized according to their location into key, eyebrows, eyes, nose, lips, and boundary.

Results were made on the PICS [15]. The dataset was acquired by the frontal images of 50 subjects. The size of images ranges from 336*480 to 624*544.

In comparison term, the comparative study [12] was revised, as it shows a comparison between several models to detect landmarks. It includes boosted regression and graph models (Borman) [16], Optimized Part mixtures and Cascaded deformable shape Model (OPM-CLM) [17], Graphical Model (GM) [18] and Fast Fitting Appearance Model (FF-AAM) [19]. The results of applying these models on PICS showed a close approximation of the results obtained by the comparison study. Table I. shows RMS of detecting the Landmarks through these models in addition to the proposed model, RSAAM.

Table I: RMS of detection landmarks across RSAACM and previous methods

Landmark	BorMan	OPM-CLM	GM	FF-AAM	RSAACM
Keys	12.34	07.57	10.15	05.31	02.99
Eyes	11.74	06.08	06.74	05.68	05.52
Eyebrows	11.43	07.14	09.53	07.57	06.61
Nose	15.68	11.07	16.84	08.09	09.64
Lips	15.02	12.10	17.26	08.16	08.43
Boundary	19.75	18.72	19.23	15.04	14.68
Overall	14.63	10.92	13.76	08.65	08.12

As observed in Table I, the direct detection by composite cross-correlation of the key landmarks by RSAACM, returns lesser RMCM. This is resulting in an appropriate initialization that produces a more accurate detection of the rest of the landmarks. However, FF-AAM detects boundary landmarks better than the rest of the previous models because it depends on the overall facial appearance. Despite this, RSAACM overcame it because of combining AACM to RSCM.

VII. Conclusion

In this paper, we proposed RSAACM for detecting facial landmarks. This model was based on determining the easiest-to-detect landmarks. Improvements of previous models are included in the RSCM. These improvements were accomplished by restricting ASM by RSCM and abbreviate AAM by AACM. The experimental results on a data set of PICS showed the effectiveness of RSAACM in the detecting both internal and frame landmarks. The results were compared with a range of modern detection models. The paper focused on detection in case of the frontal pose. It is therefore advisable to generalize RSCM to include multi-view as a future work. It involves to using a method for estimating the pose and choosing the shape. Also, AACM use only four training samples. It is useful to select the training based on analyzing the dataset. Itmay is achieved by using classification methods or employs ethnic and gender classification during selection.

References

[1] N. Chandrakar, and D. Bhonsle. "Study and comparison of various image edge detection techniques."*International Journal ofManagment, IT and Engineering* vol. 2, no. 5, pp. 499-509, 2012.

[2] L. Bin, and M. S.Yeganeh. "Comparison for image edge detection algorithms."*IOSR Journal of Computer Engineering* vol. 2, no. 6,pp. 1-4, 2012.

[3] H. Kaur, and L. Kaur. "Performance Comparison of Different Feature Detection Methods with Gabor Filter." *International Journal of Science and Research (IJSR)* vol. 3, pp. 1880-1886, 2014.

[4] N. Bhoi, and M. N. Mohanty. "Template matching based eye detection in facial image." *International Journal of Computer Applications (0975–8887)* pp 15-18, Dec. 2010.

[5] K. Nanaa, M. Rizon, M. N. Abd Rahman, A. Almejrad, A. Z. Abd Aziz, and S. B. Mohamed. "Eye detection using composite cross-correlation." *American Journal of Applied Sciences* vol. 10, no. 11, pp. 1448. Nov. 2013.

[6] K. Seshadriand M. Savvides. "Robust modified active shape model for automatic facial landmark annotation of frontal faces." In*Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on*, pp. 1-8. IEEE, Sep. 2009.

[7] T. H. Le, and T. N. Vo. "Face alignment using active shape model and support vector machine." *arXiv preprint arXiv:1209.6151*. Sep. 2012.

[8] Y. H Lee, C. G Kim, Y. Kim, and T. K. Whangbo. "Facial landmarks detection using improved active shape model on android platform."*Multimedia Tools and Applications* vol. 74, no. 20,pp. 8821-8830, Oct. 2015.

[9] S. Baker, and I. Matthews. "Lucas-kanade 20 years on: A unifying framework." *International journal of computer vision* vol. 56, no. 3, pp. 221-255, Feb. 2004.

[10] T. F. Cootes, , G. J. Edwards, and C. J. Taylor. "Active appearance models." *IEEE Transactions on pattern analysis and machine intelligence* vol. 23, no. 6, pp. 681-685, Jun. 2001.

[11] A. Criminisi, D. Robertson, E. Konukoglu, J. Shotton, S. Pathak, S. White, and K. Siddiqui. "Regression forests for efficient anatomy detection and localization in computed tomography scans." *Medical image analysis* vol. 17, no. 8, pp. 1293-1303, Dec. 2013.

[12] N. Wang, , X. Gao, D. Tao, H. Yang, and X. Li. "Facial feature point detection: A comprehensive survey." *Neurocomputing.* Jan. 2017.

[13] M. Uricar, V. Franc, and V. Hlavac. "Detector of facial landmarks learned by the structured output SVM." *VIsAPP* vol. 12, pp. 547-556, Feb. 2012.

[14] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. Huang. "Interactive facial feature localization." *Computer Vision–ECCV 2012* , pp. 679-692, 2012.

[15] P. Hancock. "Psychological image collection at stirling (pics)." *Web address: http://pics. psych. stir. ac. uk* , Mar. 2008.

[16] M. Valstar, B. Martinez, X. Binefa, and M. Pantic. "Facial point detection using boosted regression and graph models." In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 2729-2736, Jun. 2010.

[17] X. Yu, J. Huang, S. Zhang, W. Yan, and D. N. Metaxas. "Pose-free facial landmark fitting via optimized part mixtures and cascaded deformable shape model." In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1944-1951. 2013.

[18] X. Zhu, and D. Ramanan. "Face detection, pose estimation, and landmark localization in the wild." In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 2879-2886. IEEE, 2012.

[19] G. Tzimiropoulos, and M.Pantic. "Optimization problems for fast aam fitting in-the-wild." In *Proceedings of the IEEE international conference on computer vision*, pp. 593-600. 2013.

Kutiba Nanaa was born in Kingdom of the Saudi Arabia. He had Bachelors Degree in Informatics Engineering, Department of Artificial Intelligence from University of Aleppo, Syria in 2006. Masters in intelligence system from Universiti Utara Malaysia, Malaysia in 2010. Presently a Ph. D. candidate at Universiti Sultan Zainal Abidin in Kuala Terenganu, Malaysia.

The area of interest is computer vision, robotics, expert system, and natural language.