ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

Facial Landmark Detection and Localisation using Explicit Model Based Approach

Naveen Kumar H N[1], Dr.Jagadeesha S[2], Amith K Jain[3]

Assistant Professor, Dept. of ECE, SDMIT, Ujire, Karnataka, India[1]

Professor, Dept. of ECE, SDMIT, Ujire, Karnataka, India[2]

Assistant Professor, Dept. of ECE, SDMIT, Ujire, Karnataka, India[3]

ABSTRACT: Facial landmark localization is one of the basic approaches to identify face alignments, achieve facial expression recognition and perform the same, even in the presence of occlusions. The acquisition conditions such as background complexity, degree of occlusion, illumination variations and expressions on the face affect landmark localization performance. Use of Eigen faces for shape modeling degrades the performance of the system as Eigen faces are not robust to variations in shape, pose and expression. In the proposed work, explicit model based method is used for landmarking approach. Landmark detection and localization is carried out using Point Distribution Model (PDM) and Active Shape Model (ASM). The proposed method is implemented on MATLAB 2015a using existing standard datasets and very own dataset.Results obtained from the experiment indicate that proposed work detects and localizes facial landmarks more effectively when encountered with images exhibiting different degree of occlusions, expressions and pose variations. This, in turn makes the system more robust.

KEYWORDS: Point Distribution Model (PDM), Active Shape Model (ASM).

I. INTRODUCTION

A landmark is a point of correspondence on each object that matches between and within populations of same class of objects. One way to describe a shape is by locating finite number of landmarks on the outline or other specific points.Landmarks are classified into following categories.Anatomical landmarks: Points assigned by an expert that correspond between organisms in some biologically meaningful way (e.g., the corner of an eye). Mathematical landmarks: Points located on an object according to some mathematical or geometrical property. Pseudo-landmarks: Constructed points on an object either on the outline or between anatomical or mathematical landmarks.Labelled landmarks: Landmarks that are associated with a label (name or number), which is used to identify the corresponding landmark.

Synonyms for landmarks include key points, homologous points, interest points, anchor points, fiducial markers, model points, markers, nodes, vertices etc. Facial landmarks are defined through detection and localization of certain key points on the face, which have an impact on resulting task focused on the face. The list of resulting tasks includes animation, face recognition, gesture understanding, gaze detection, face tracking, expression recognition etc.

Facial landmarks include nostril corners, ear lobes, nose tip, eye corners, eyebrow arcs, chin, mouth corners etc. Many of the landmark detection algorithms prefer an entire facial semantic region for ease of analysis. Facial landmarks are classified in to two groups, primary or fiducial and secondary or ancillary. This distinction is based upon the reliability of detection techniques corresponding to related image features. Fudicial group of landmarks are detected directly and play a crucial role in face identity and tracking. The fudicial landmarks such as corners of the mouth, corners of the eyes, nose tips, eyebrows can be detected easily by using low level image features such as SIFT, HOG. Ancillary landmark detection is guided by fiducial landmarks. Ancillary landmarks include nostrils, eyebrow, chin, cheek contours,and midpoints of lip region and non-extremity points. It takes up a more prominent role in recognition of a facial expression.

Regardless of a good landmark detector, there are cases where landmark values cannot be computed, due to missing data (the occlusion). At the same time, landmark approach must try to detect as many points as possible; this leads to

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

complex and more general systems. Human faces vary from one being to another; it is therefore difficult to differentiate faces under high occlusion. More so, to detect and remove occlusions on facial images quickly and automatically leads towards a largely unsolved problem. This marks out face detection and recognition as being the toughest problem in the fields of computer vision and biometrics. Moreover, aligning faces robustly and precisely is one of the significant steps to solve challenges in facial landmark detection. In order to overcome these challenges, significant contributions have been made to aid in the process of identification, in different scenarios. Some of the existing techniques for landmarking localization are reviewed in the following section.

II. RELATED WORK

Facial landmarking techniques are classified as texture based methods and template based methods [1]. Texture based methods may be transform based or template based. Existing work on transform based landmarking approaches uses Gabor transform or Discrete cosine transform (DCT) or Independent Component Analysis (ICA). Existing work on template based landmarking approaches uses fixed templates or deformable templates. Texture based approaches resulted in lower detection rates and are not robust to pose variations, occlusions and expression. Shape-guided or model-based methods consider the whole face and the ensemble of landmark as an instantiation of a shape. Of the two sub-categories of model based methods, the explicit model-based methods are by far more popular, while there are only a few research articles in the alternative implicit methods.

Yuchi Huang et al [2], presented component based deformable model for generalized face alignment, they used a novel bi-stage statistical framework to account for both local and global shape characteristics. It uses separate Gaussian models for shape components in place of statistical analysis on the entire shape, which preserves more detailed local shape deformations. Each model of components used the Markov Network as search strategy. They used Gaussian Process Latent Variable Model to gain control over full range shape variations; hence it makes out better description of the nonlinear interrelationships over shape components. This approach allows system to preserve the full range low frequency shape variations, and also the high frequency local deformations caused by exaggerated expression. Their work is implemented on YALE FACE DATABASE B images only and subsequently 79 key points were detected.

David Cristinacce et al [3] developed CLM which produces landmark templates iteratively and uses shape constrained search technique. The position vectors of the landmark templates are estimated using Bayesian formulation. The posterior distribution in Bayesian formula incorporates image information via template matching scores and statistical shape information. Hence, positions of new landmarks are predicted in the joint shape model and light of the image. Furthermore, templates are updated through sampling from training images. The work is implemented on BioID and XM2VTS database and subsequently 22 key points were detected.

Fernando De la Torre et al [4], presented the Parameterized Kernel Principle Component Analysis (PKPCA) model by extending KPCA to incorporate geometric transformation into formulation and applying gradient descent algorithm for fast alignment. This model differs from PCA, as it can model non-linear structure in data variant to rigid or non-rigid deformations. It does not require manually labeled training data. Their work is implemented on CMU Multi-PIE database for images only and subsequently 46 key points were detected.

None of the existing work addresses landmark localization for images which have certain degree of occlusion, pose variation, expression or blur. There is still need for significant attention in terms of missing data, control points (labeling), misalignment, restoration, illumination variations and expression. Information about variations in shape is usually collected for building a model. The model represents a predefined number of landmark points, which depends on the object's shape complexity and desired level of detailed descriptions, wherever needed. The proposed work highlights facial landmark detection and estimation under various expressions and occlusions. The approach is aimed at detecting the position of face and facial features (eye, mouth, and nose) despite object being occluded by hand, hair, scarf, and so on. It is also tested to fit on faces with different expressions such as smile, open mouth etc.

III. PROPOSED ALGORITHM

The fundamental block diagram of Facial Landmark Localization system consists of face detection, PDM and ASM block as shown in Fig.1. Viola jones algorithm is employed for detection of frontal face. Use of integral image representation in viola jones face detector makes the detection faster. Pre-processing stage is optional and operations such as geometric correction, histogram equalization, contrast enhancement, resizing and low pass filtering may be employed for the acquired image prior to feature extraction so as to improve the landmark detection rate. Explicit model based approach employs PDM and ASM for facial landmarking.

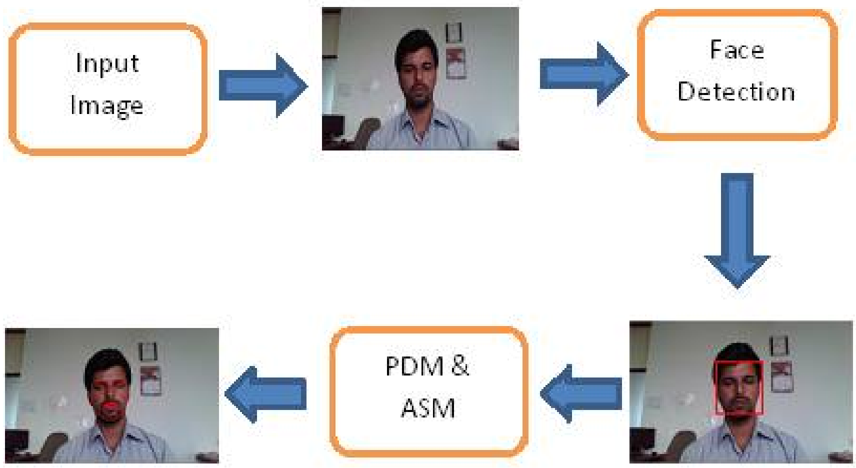

Fig.1: Block diagram of the proposed system.

A. *Point Distribution Model (PDM) & Active Shape Model (ASM)*

The Point Distribution Model (PDM) or Statistical Shape Model is a shape description technique that is used in locating new instances of shapes in images. As opposed to just building rigid model, PDM tries to "understand" the shape. It is very useful for describing features that have well understood general shape, but which cannot be easily described by a rigid model. Mean geometry of a shape along with statistical modes of geometric variation inferred from a training set of shapes are represented by PDM. It is a combination of model based approach and local edge feature detection, which makes PDM a fast and simple method for representing an object. PDM relies on landmark points.

The PDM approach assumes existence of a set of N examples (a training set) from which a statistical description of the shape and its variation has to be derived. In this context, it refers toward some number of instances of the shape represented by a boundary (a sequence of pixel co-ordinates). In addition, some number N of landmark points is selected on each boundary; these points are chosen to correspond to a feature of the underlying object. Variations in the positions of these points would then be attributable to natural variation between individuals. The PDM approach allows for modelling these small differences (and, indeed, to identify which are truly small, and which are more significant). ASM is used to interpret shape model for new images by using statistical shape model derived from PDM. ASM assures fast, simple and accurate shape modelling for face. ASM performs more precisely on shape localization, and is quite more robust to illumination and bad initialization. After landmarks alignment and Principal Component Analysis, we construct gray-level profile for each landmark in all multi-resolution versions of a training image. In search procedure, we give the model's position an initial estimate. Then it can compute the suggested movements through an iteration approach using the gray-level profile. When convergence is established, we get a final matching result.

IV. PSEUDO CODE

Step 1: The input to PDM consists of M training samples with N points for each sample, represented by equation (1).

$$x^i = (x^i_1, y^i_1, x^i_2, y^i_2, \dots\dots\dots x^i_N, y^i_N)^T \text{............ (1)}$$

Step 2: Align each x^i with x^1, for i=2,3,…..,M

Step 3: The mean $\bar{x} = [\overline{x_1}, \overline{y_1}, \overline{x_2}, \overline{y_2}, \dots, \overline{x_N}, \overline{y_N}]$ of the aligned shapes $\{\hat{x}^1, \hat{x}^2, \dots\dots, \hat{x}^M\}$ is represented in equation (2).

$$\overline{x_m} = \frac{1}{M}\sum_{i=1}^{M}\hat{x}^i_j \quad \overline{y_m} = \frac{1}{M}\sum_{i=1}^{M}\hat{y}^i_j \text{............ (2)}$$

Step 4: Align the mean shape $\bar{x}$ with x^1.

Step 5: Align $\hat{x}^2, \dots\dots, \hat{x}^M$ to the adjusted mean.

Step 6: Repeat until convergence to align all the training shapes.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

Step 7: From the available M boundaries $\hat{x}^1, \hat{x}^2, \dots\dots, \hat{x}^M$ and the mean $\bar{x}$and the variation from the mean for each training shape iscalculated byequation (3).

$$\Delta x^i = \hat{x}^i - \bar{x} \dots\dots\dots\dots\dots\dots\dots\dots \quad (3)$$

Step 8: Covariance matrix Cm (2N x 2N) is represented by equation (4)

$$C_m = \frac{1}{M}\sum_{i=1}^{M} \Delta x^i (\Delta x^i)^T \dots\dots\dots\dots\dots \quad (4)$$

Step 9: Eigen decomposition is given by equation (5), where $q = [q^1 q^2 q^3 \dots\dots q^{2N}]$

$$C_m q_i = \lambda_i q_i \dots\dots\dots\dots\dots\dots\dots\dots \quad (5)$$

Step 10: As C_m is symmetric and positive definite, the eigenvalues λ_i are real. Eigen vectors q_i are orthogonal, so q is Basis vector and any vector x can be represented as $x = \bar{x} + qb$

Step 11: Consider only k largest Eigen valuesas in equation (6).

$$x \approx \bar{x} + q_k b_k \dots\dots\dots\dots\dots\dots\dots\dots \quad (6)$$

with$q_k = [q^1 q^2 q^3 \dots\dots q^k]$ and $b_t = [b_1, b_2, \dots\dots . b_k]^T$

Step 12: The outputs of PDM are mean $\bar{x}$ and reduced Eigen vector matrixq_k.

Step 13: The new shape generated by ASM model is represented in equation (7)

$$y = \bar{x} + q_k b_k \dots\dots\dots\dots\dots\dots\dots\dots \quad (7)$$

Step 14: Fitted model is represented by pose parameters (θ, s, t_x ,t_y) and shape parameter b.

$$\hat{q} = qb + \bar{q} \dots\dots\dots\dots\dots\dots\dots \quad (8)$$

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = s \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} \hat{x}_i \\ \hat{y}_i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \dots\dots \quad (9)$$

$$q = [x_1, y_1, x_2, y_2 \dots\dots . x_N, y_N] \dots\dots \quad (10)$$

Step 15: Calculate the edge map of the image.

Step 16: For each landmark q_i find a line normal to the shape contour.

Step 17: Align pose parameter θ, S, t_x and t_y.

Step 18: Adjust pose parameter b and repeat until convergence.

V. SIMULATION RESULTS

In the proposed work, face model is divided into 5 regions as shown in Fig.2. Region 1 represents the right eye and a right eye brow, region 2 represents the left eye and a left eye brow, region 3 represents nose, region 4 represents mouth and region 5 represents face contour. PDM for an example image is as shown in Fig.3.

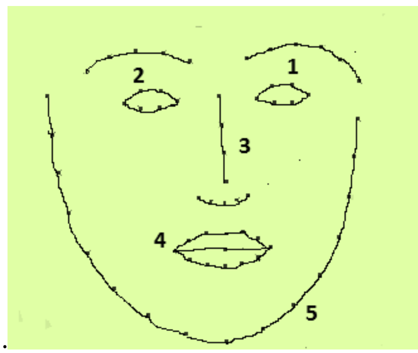

Fig.2: PDM for face.

Fig.3: PDM for face; An example.

The number of landmarks detected in each region of the Fig.3 is tabulated in TABLE I. The total number of landmark points detected is 66. The number of key points detected in each eye region is 11. The number of key points detected in nose, mouth and face contour are 9, 18 and 17 respectively.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

TABLE I. NUMBER OF LANDMARK POINTS DETECTED IN EACH REGION.

Region number	Number of landmark points
Region 1	11
Region 2	11
Region 3	9
Region 4	18
Region 5	17

The proposed work is implemented on MATLAB 2015a. It is carried out on standard data sets like JAFFE, Cohn-kanade, YALE, CMU Multi-PIE and is also experimented on own data set. The data set being considered is varied with respect to age, gender, ethnicity, expressions and occlusions. The images exhibiting small pose variations have been experimented and results of these scenarios are as shown in Fig.4. The proposed algorithm has given good landmark detection and localization results for pose varied images.

Fig.4: Results of the proposed work for images that comprise pose variations.

The images exhibiting various expressions like yawning, laughing, surprise, fear, anger, sad etc. have been experimented and results of these scenarios are as shown in Fig.5. The proposed algorithm has achieved good landmark detection and localization results for different expression images.

Fig.5: Results of the proposed work for images that comprisedifferent expressions.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

The images exhibiting various occlusions like mouth region being covered, eye region being covered, eyes covered with goggle have been experimented and results of these scenarios are as shown in Fig.6. Results indicate that the proposed work is robust to pose variations, expressions and occlusions.

Fig.6: Results of the proposed work for images that comprise different degree of occlusions.

The comparison of the proposed work with the state of art techniques is summarized in TABLE II. The existing methods for landmark detection and localization like Component based Deformable Model, Constrained Local Model (CLM) and Parameterized Kernel Principal Component Analysis (PKPCA) are not robust to pose variations, expressions on face and facial occlusions. Whereas this model works robustly, even if more than half of the face is entirely covered. Moreover, the proposed work gives good localization results when implemented with images representing a combination of occlusion, pose variation and expression.

TABLE II. COMPARISON OF THE PROPOSED WORK WITH THE OTHER APPROACHES.

Method	Result	Summary
Component based Deformable Model implemented on YALE Face Data base B		The number of key points detected is 79. Implemented only for frontal face images. The model is not robust to pose variations, occlusions and expressions.
Constrained Local Model (CLM) implemented on BioID and XM2VTS Data base		The number of key points detected is 22. Because of less number of key points the facial region segmentation was unsuccessful.Implemented only for frontal face images.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

Parameterized Kernel Principal ComponentAnalysis (PKPCA) implemented on CMU Multi-PIE Data base		The number of key points detected is 46. Implemented only for frontal face images. The model is not robust to pose variations, high occlusions and expressions.
Proposed work (PDM + ASM) implemented on JAFFE, Cohn-kanade, YALE, CMU Multi-PIE Data base		The number of key points detected is 66. The model is robust to pose variations, occlusions and expressions.

VI. CONCLUSION AND FUTURE WORK

The various challenges in facial landmark detection and localization have been discussed and a new approach is proposed to address the same, using PDM and ASM. Herein, PDM is employed with a total number of 66 points, which fully describes shape variations in accordance with five divided regions. Shape of the object is represented by a set of points using PDM, and ASM aims to match the model to a new image. The results of the experiment carried out suggest that, combination of PDM and ASM makes the proposed system more robust to images with different degree of occlusions, pose variations and expressions. The model works robustly, even if more than half of the face is entirely covered. Moreover, the proposed work gives good localization results when implemented with images representing a combination of occlusion, pose variation and expression. In future the present work can be enhanced for blurred images.

REFERENCES

1. OyaCeliktutan, SezerUlukaya and BulentSankur, “A comaprative study of facial landmarking techniques”, EURASIP Journal on Image and Video Processing 2013.
2. YuchiHuang, Qingshan Liu1and Dimitris Metaxas, “A Component Based Deformable Model for Generalized Face Alignment”, IEEE Conference on Computer Vision & Pattern Recognition, 978-1-4244-1631-8/07/$25.00 ©2007.
3. David Cristinacce and Tim Cootes, “Automatic feature localisation with constrained local models”, Pattern Recognition Elsevier, doi:10.1016/j.patcog.2008.01.024.
4. Fernando De la Torre and Minh Hoai Nguyen, “Parameterized Kernel Principal Component Analysis: Theory and Applications to Supervised and Unsupervised Image Alignment”, conference proceedings Carnegie Mellon University.
5. Paul Viola and Michael Jones, “Rapid Object Detection using a Boosted Cascade of Simple Classifiers”, Conference paper on Computer Vision and Pattern Recognition 2001.
6. Y Tie, L Guan, “Automatic landmark point detection and tracking for human facial expressions”, Image Video Process. 2013.
7. F. de la Torre, A. Collet, J. Cohn, and T. Kanade.,“Filtered component analysis to increase robustness to local minima in appearance models”, International Conference on Computer Vision and Pattern Recognition, 2007.
8. I. Kookinosand A. Yuille., “Unsupervised learning of object deformation models”, International Conference on Computer Vision, 2007.
9. F. T. Haar andR. C. R. Velkamp, “Expression Modelling for Expression-invariant Face Recognition”, Computers and Graphics, vol. 34, no. 3, pp. 231-241, 2010.
10. P. Perakis, T. Theoharis, G. Passalis, and I. Kakadiaris, “3D Facial Landmark Detection Under Large Yaw and Expression Variations”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 7, pp. 1552–1564, July 2013.
11. P. J Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, “Overview of the Face Recognition Grand Challenge”, In Proceeding of International Conference on Computer Vision and Pattern Recognition, vol. 1, 2005, pp. 947–954.
12. P. Nair and A. Cavallaro,“3D Face Detection, Landmark Localization, and Registration using a Point Distribution Model”, IEEE Transactions on Multimedia, vol. 11, no. 4, pp. 611–623, June 2009.
13. G. Passalis, P. Perakis, T. Theoharis, and I. Kakadiaris, “Using Facial Symmetry to handle Pose Variations in Real-world 3D Face Recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 10, Oct. 2011, pp. 1938–1951.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

14. M. V. Gupta and D. Sharma, “A Study of Various Face Detection Methods”, International Journal of Computer and Communication Engineering, vol. 3, no. 5, pp. 3–6, 2014.
15. D. N. Parmar and B. B. Mehta, “Face Recognition Methods & Applications”, International Journal of Computer Technology & Applications, vol. 4, no. 1, pp. 84–86, 2013.

BIOGRAPHY

Naveen Kumar H N has received the B.E degree in Electronics and Communication and M.Tech degree in Digital Electronics and Communication from VTU Belagavi. His research interests include Image Processing & Computer Vision, Facial Expression Analysis and Human Computer Interaction.

Dr.Jagadeesha S has received the B.E degree in Electronics & Communication from Kuvempu University and M.E degree in Power Electronics from Gulbarga University. He has received Doctor of Philosophy (Ph.D) in the Department of Applied Electronics from Gulbarga University, Gulbarga. His research interests include Antennas, Wireless communication, Image Processing & Computer vision. He has published more than 40 research papers in various international and national journals.

Amith K Jain has received the B.E degree in Electronics and Communication and M.Tech degree in Control Systems from VTU Belagavi. His research interests include Image Processing, Computer Vision and Facial Expression Analysis.