ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

Facial Features Detection Using Regression

[1] G.Radha Priyadharsini, [2] Dr.K.Krishnaveni
[1] Research Scholar, [2] Head & Associate Professor
Sri S.R.N.M.College, Sattur, Tamil Nadu, INDIA

Abstract: - **Facial Feature Detection is one of the essential techniques in face recognition, face modeling, head pose estimation and facial expression recognition. Emotions may be recognized through the facial features such as eyes, nose, lip movements etc. The objective of the proposed work is to detect the face with these facial features. Cascade Object Detector is initially proposed to identify the face region from the input image. Then Regression based face alignment algorithm is employed for the feature point alignment and registration. By means of the scale-invariant feature transform and regression results, the landmark points for the facial features are estimated inside the template. Delaunay method is used to construct a triangulation that can be utilized to draw the template with the associated points. Finally the facial features are detected along with the boundary points. The WSEFEPv101lo dataset is used for evaluation and analysis. Out of 62 facial images, the facial features of 48 images are accurately detected and due to misalignment of facial landmark points due to facial emotions, 14 images are not detected correctly.**

Keywords**: Cascade Object detector, Scale invariant feature transform, regression, Delaunay, boundary.**

I. INTRODUCTION

Human visualization is the unique, complex perception mechanism. Human perception has the capability to acquire, integrate and interpret all the visual information around us. It is unidentified that how perfectly the objects can be distinguished with little effort. Day by day, the system is to be improved to mimic the human perception. Face images has rich information such as age, gender, expression, mood, health, gaze direction etc. From Fig 1, the features like female, her mood is happy, healthy, 25-35 years, expression is smiling, and foreigner may be extracted by anyone.

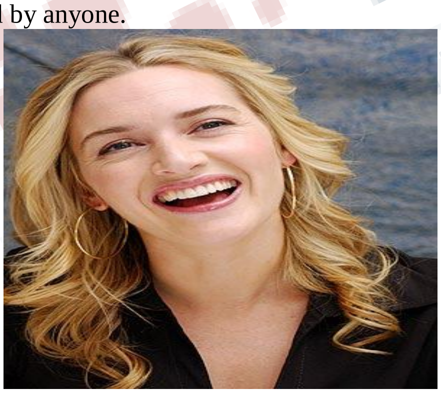


Figure 1 Sample Face image

Face detection refers to a subclass of computer technology that is able to isolate people's faces within digital images i.e., pixel recognition in the image which characterize the face. It is a special case of object recognition and common case of face localization. Face alignment is one of the computer vision problems applied to extract the facial feature points. It searches landmark points or face shapes on the face image. Facial benchmarks are the eyebrow arcs, eyes corners, mouth corners, nostril corners, nose tip, chin, ear lobes etc. For ease of analysis, most landmark detection algorithm chooses the entire facial semantic region, such as the whole region of a mouth, the region of the nose, eyes, eyebrows, cheek or chin. The facial indicator will act as a key role in the applications like lip reading, 3D face reconstruction, image editing, face tracking, verification, registration, surveillance system, head gesture understanding, sign language interpretation etc. The difficulties related with face alignment can be attributed to many variations in low resolution, illumination, facial expression, pose (out-of-plane rotation), illumination, decoration, occlusions etc. Traditional parametric methods such as Active Appearance Model [1], [2] and Active Shape Model [3], [4] are usually relying on the initialization of the landmark points. The AAM and ASM models are becoming well-known because of its extensive usage. Also most of the alignment algorithms [5], [6] depend on the face detectors, they may fail to locate the face due to some pose variations. Constrained Local Model is a successive deformable fitting model using local search and expected to pull out of the local minima [7], [8]. To handle various view points, Multi-view shape models are presented to estimate the head pose [9]. With the aim of detecting features on face images, this paper is structured as follows: Section II presents the procedure for detecting face from the given input images and also to isolate facial features of the image; Section III discusses Experimental Results and finally concludes the paper.

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

II. METHODOLOGY

The proposed technique is coarsely divided into three phases. The first phase deals with face detection from the input image and second phase presents the facial components alignment. With the results of the second phase, the last phase explains the detection of facial features. The work flow of the proposed method is shown in Fig. 2.

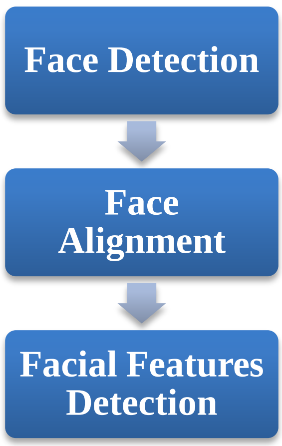


Figure 2 Outline of the proposed method

A. Face Detection

The formal definition of face detection from input image is as follows: Given an arbitrary image, the goal of face detection is to determine whether or not there are any faces in the image and if present, return the location of the image. The MATLAB system object namely "Cascade object detector" is used to identify the location of faces in the input images. This face detection process utilizes the concept of Viola-Jones algorithm which can track faces rapidly with high detection rate. It is assembled as follows:

- Haar Features – Feature Selection.
- Integral Image – Feature evaluation.
- Adaptive Boost – Learning Algorithm.
- Cascade Object Detector – Face Detector.

1. Feature Selection

Three different types of Haar features as shown in Fig. 3 and Fig.4 are proposed and these features are calculated at different sizes and orientations. The contrast difference is derived as the difference between the sum of pixels from white region and black region. The mask M from Fig. 4, the Haar feature associated with the image I behind the mask is defined by the equation (1),

$$\sum_{1\leq i\leq N}\sum_{1\leq j\leq N} I(i,j)_{white} - I(i,j)_{black} \quad \ldots.. (1)$$

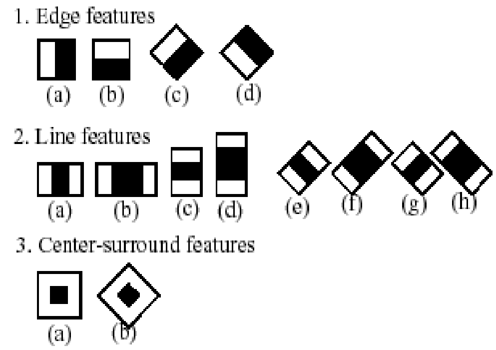


Figure 3 Types of Haar Features

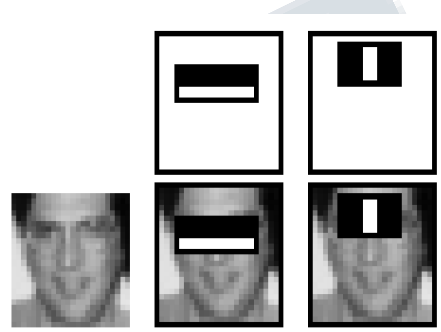


Figure 4 Haar features from which the intensity difference between the pixels from eyes region can be observed

2. Feature Evaluation

The integral image or summed area table is calculated for evaluating the Haar features. The summed area table at (x, y) is calculated by equation (2),

$$s(x,y) = i(x,y) + s(x-1,y) + s(x,y-1) - s(x-1,y-1) \quad ---- (2)$$

For example, a 2D table where each element is the sum of all elements in an input image between the lower left corner and the entry location created and shown in Fig. 5.

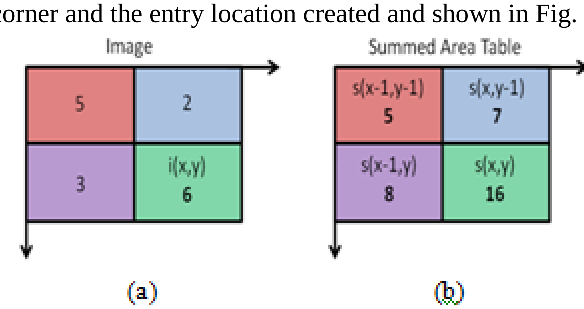


Figure 5(a) Input Image, 5(b) Summed Area Table

3. Adaptive Boost Learning Algorithm

Adaptive Boost is a machine learning boosting algorithm capable of constructing a strong classifier through a weighted combination of weak classifiers using equation (3) and the weight update process is shown in Fig. 6.

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

$$H_M(x) = \sum_{m=1}^{M} h_m(x) \quad \text{….. (3)}$$

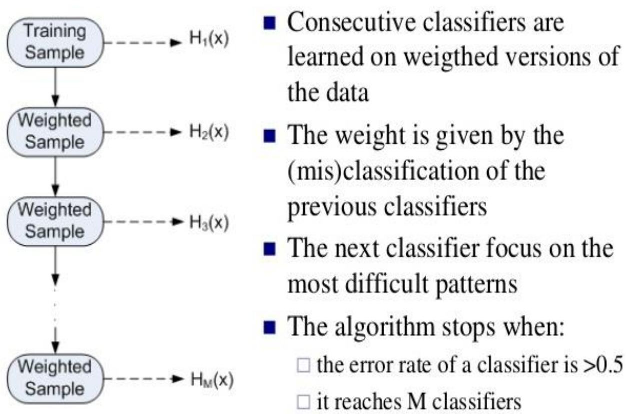


Figure 6 Weight Updating Process

The weak classifier is mathematically described as follows:

$$h(x, f, p, \theta) = \begin{cases} 1, & if \ \ pf(x) < p\theta \\ 0, & otherwise \end{cases} \quad \text{… (4)}$$

Where f is the feature, x is a 24×24 pixel image, ϴ is a threshold and p is parity. And x can be classified as positive or negative.

4. Cascaded Classifier

The attentional cascade structure discards the majority of the sub-windows in early layers of the detector. It leads to the construction of the detection process very efficient. Positive response from the first classifier triggers the evaluation of a second (more complex) classifier, and so on. A negative outcome at any point will lead to the immediate rejection of the sub-window. The classifier structure of this algorithm is shown in Fig 7.

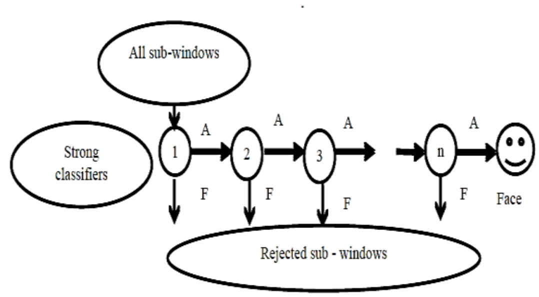


Figure 7 Cascade Classifier

The classification model property controls the type of the object to detect. By default, the detector is configured to detect faces with the shape measurement property called 'Bounding Box' [10]. The output of the Cascade Object Detector using bounding box method is shown in the Fig 8.

(a) **(b)**

Figure 8(a) Input Image, 8(b) Face Detection

B. Feature Point Alignment

According to the computer vision technology, face alignment is used to discover the geometric configuration of the human face in digital image. The Regression based face alignment is used for feature point alignment. In general, regressors study from the training data. For the given test image, constructing with the use of the learned data or model as well as testing to predict the contour, targeting at aligning the assessed shape to the true shape. Supervised Descent Method is employed to align the landmarks in the input images. It uses scale invariant features for feature extraction. It results better than Active Shape Models or Active Appearance Model. The alignment algorithm is non-parametric in appearance and shape. It uses part-based depiction and learns different regressors at different stages. Supervised descent method frames the face alignment problem as a minimization problem,

$$\Delta X^* = \underset{\Delta X}{\operatorname{argmin}} f(\Delta X) = \underset{\Delta X}{\operatorname{argmin}} \left\| h\left(I(X_0 + \Delta X)\right) - \phi_* \right\|_2^2 \quad \text{..(5)}$$

Where h (I(x)) is the SIFT feature extraction function that analyzes feature values of image I at benchmarks x and ϕ* = h (I(x*)) signifies the SIFT values at the ground truth landmarks x*. The procedure acquires descent directions from training data. The system learns a sequence of mapping matrix Rk that maps the local features ϕ at landmark x to the motion of landmarks Δx from training dataset {Ii} with known landmarks {xi*}.

$$\underset{R^k, b^k}{\operatorname{argmin}} \sum_{I_i} \sum_{x_k^i} \left\| \Delta x_*^{ki} - R^k \phi_i^k - b^k \right\|^2 \quad \text{… (6)}$$

And updates the location of landmarks by

$$X^k = X^{k-1} + R^{k-1} \phi^{k-1} + b^{k-1} \quad \text{… (7)}$$

Where Δx*ki is the displacement between the ground truth landmarks and the estimated landmarks at the kth stage for image i [11]. As soon as the set of interest points are detected from an input image at p(x, y), scale s, and orientation θ, their image construction in a neighbourhood of p desires to be prearranged in a suitable descriptor for discriminative matching and unresponsive to local image distortions. The descriptor should be associated with θ and comparative to the scale s. Scale invariant feature transform is a scheme for discovering salient, rigid feature points in an image. For all such point, it also delivers a set of features that describe a minor image area around the point. These features are invariant to rotation and scale [12]. Figure 9 demonstrates the graphical representation of the scale invariant feature transform – an image feature descriptor algorithm; where the gradient orientations and magnitudes are calculated at every pixel and then weighted by a Gaussian falloff (showed by overlapped circle). A weighted gradient orientation histogram is then computed for each sub region. There are 4 × 4 histograms every one with 8 bins; the feature vector has 4 × 4 × 8 = 128 elements for every key point. In this research, a 300w shape model is used for training with 49 landmark points. The initial position of the landmark points generated using the mean shape of the model points. And scale invariant features (SIFT) are extracted from the areas around the landmark points of patch size 16. The new landmark points are estimated using the outcomes of the equation (7). The learning process is repeated. And the values of the generic descent directions and the bias term are stored. Using these values the updated landmark positions are predicted. These steps are shown in the Fig. 10.

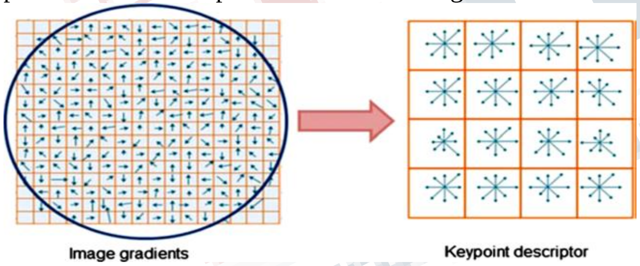


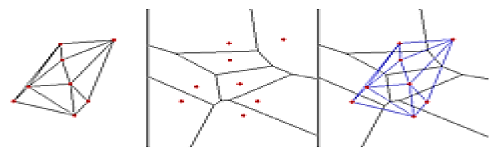
Figure 9 SIFT for a 16 × 16 pixel patch and a 4 × 4 Descriptor array

(a) (b)

Figure 10 (a) Initialization (b) Aligned Landmark Points

C. Facial Features Detection

The boundary of the facial feature points are aligned and extracted using the Delaunay Triangulation and its convex hull property. A triangulation T of a point set S of N points is a collection of non-overlapping triangles cornered at these N points that covers the convex hull of S. For a set S of points in the Euclidean plane, the unique triangulation DT(S) of S such that no point in S is inside the circumcircle of any triangle in DT(S) then DT(S) is the dual of the Voronoi diagram of S. If n is the number of points in S, the Voronoi diagram of S is the partitioning of the plane containing S points into n convex polygons such that each polygon contains exactly one point and every point in a given polygon is closer to its central point than to any other. A Voronoi diagram is also known as a Dirichlet Tessellation [13]. Delaunay uses an existing vector point's map (input) to create a Delaunay triangulation vector map (output). Delaunay triangulation and Voronoi diagram example are shown in the Fig. 11. The exterior face of the Delaunay triangulation is the convex hull of the point set. In the trial, the aligned and updated landmark points are triangulated using Delaunay triangulation method.



(a) (b) (c)

Figure 11 (a) Delaunay Triangulation (left), (b) Voronoi diagram (center), (c) both (right)

The method will construct the triangulation structure using the aligned data points as input. The boundary of the triangulation is estimated using convex hull property of the Delaunay and it is shown in the Fig. 12 in green colour.

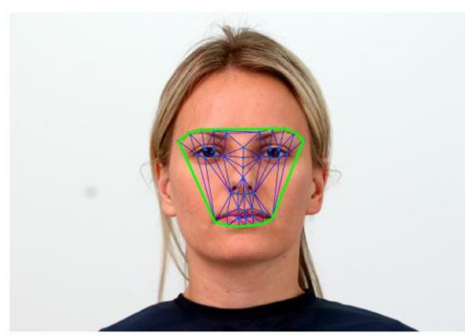


Figure 12 Boundaries of the Triangulation Points

Then, the mask is created using the convex hull points using the function roipoly () in the MATLAB as follows:

$$BW = roipoly\ (im, x\ (k), y\ (k));\qquad \text{------- (8)}$$

Where im is the input image and x (k), y (k) are convex hull points of the input image. Then crop out the masked region for facial features detection as shown in Fig. 13.

Figure 13 Facial Features detection using Face Alignment

III EXPERIMENTAL RESULTS

The facial features detection process is analyzed in MATLAB environment. The WSEFEP_v101-lo dataset is utilized for the experimental analysis. The dataset contains color images 800×542 dimensions with sRGB color representation [14]. Some of the accurate and invalid faces are shown in Table I and Table II respectively.

TABLE I ACCURATE RESULTS

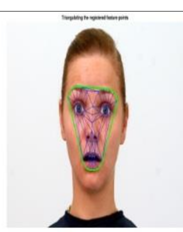
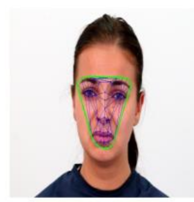
S.No	Aligned Image	Feature Detection
1		
2		
3		
4		

TABLE II MISIDENTIFIED FACES

S. NO	Misaligned Images	Inaccurate Features
1		
2		
3		
4		

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

IV. CONCLUSION

The face is detected in the frontal view images with facial features such as eyebrow, eye, nose and lip. The dataset has 62 input images and 14 images are misaligned due to multiple variations of features. The next logical step is to propose the scheme for recognizing emotions and it will be the future enhancement of this research.

REFERENCES

[1] T. Cootes, G. Edwards, and C. Taylor.''Active appearance models'', In ECCV, 1998.

[2] I. Matthews and S. Baker,''Active appearance models revisited'', IJCV, 2004.

[3] D. Cristinacce and T. Cootes,''Boosted regression active shape models'', In BMVC, 2007.

[4] T. Cootes, C. Taylor, D. Cooper, and J. Graham, ''Active shape models-their training and application'', CVIU, 1995.

[5] P. Dollar, P. Welinder, and P. Perona,'' Cascaded pose regression'', In CVPR, 2010.

[6] J. Saragih, S. Lucey, and J. Cohn,'' Deformable model fitting by regularized landmark mean-shift'', IJCV,2010.

[7] Xudong Cao , Yichen Wei, Fang Wen , Jian Sun,'' Face Alignment by Explicit Shape Regression'', DOI 10.1007/s11263-013-0667-3,Springer, 2013.

[8] D. Cristinacce and T. Cootes,'' Automatic feature localization with constrained local models'', PR, 2007.

[9] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. Metaxas, and X. Zhou,'' Sparse shape composition: A new framework for shape prior modeling'', In CVPR, 2011.

[10] Elena Alionte, Corneliu Lazar, Member, IEEE,'' A Practical Implementation of Face Detection by Using Matlab Cascade Object Detector'', 19th International Conference on System Theory, Control and Computing (ICSTCC), 2015.

[11] X. Xiong and F. De la Torre,'' Supervised descent method and its applications to face alignment'', In Computer Vision and Pattern Recognition (CVPR), 2013.

[12] D. G. Lowe,'' Object recognition from local scale invariant features'', In Computer vision, 1999. The proceedings of the seventh IEEE international conference on, volume 2, pages 1150–1157. IEEE, 1999.

[13]https://root.cern.ch/root/html516/TGraph2D.html.

[14] Olszanowski, M., Pochwatko, G., Kukliński, K., Ścibor-Rylski, M., Lewinski, P., Ohme, R. (2015). Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. Frontiers in Psychology, 5:1516. doi: 10.3389/fpsyg.2014.01516.

[15] Hlaing Htake Khaung Tin, '' Robust Algorithm for Face Detection in Color Images'', I.J.Modern Education and Computer Science, 2012, 2, 31-37.

[16] Md. Hafizur Rahman & Jinia Afrin," Human Face Detection in Color Images with Complex Background using Triangular Approach", Global Journal of Computer Science and Technology, Graphics & Vision Volume 13, Issue 4, Version 1.0, 2013.

[17] Mohamed A. Berbar, Hamdy M. Kelash, and Amany A. Kandeel," Faces and Facial Features Detection in Color Images", IEEE, 2006.

[18] Chandni Sharma, Shreya Patel, Abhishek More," Review of Face Detection based on Color Image and Binary Image'', International Journal of Computer Applications ,Volume 134,January 2016.

ACKNOWLEDGMENT

G.Radha Priyadharsini has received her M.Sc (CS&IT) degree from Devanga Arts College affiliated to Madurai Kamaraj University, Madurai, TamilNadu and M.Phil degree in Computer Science from Alagappa University, Karaikudi, TamilNadu. She is pursuing Ph.D in Computer Science. Her area of interests are Image processing and Data mining.

Dr.K.Krishnaveni, is an Associate Professor and Head of the Department of Computer

Science, Sri S.Ramasamy Naidu Memorial College, Sattur, India. She has 25 years of teaching experience. She received her B.E. degree in Computer Science and Engineering from Bharathiar University, Coimbatore, India in 1990 and M.Tech. degree in Computer and Information Technology from Center for Information Technology and Engineering of Manonmaniam Sundaranar University, Tirunelveli, India in 2004 and received Ph.D degree from the same University. Currently her research interests are Medical Image Processing, Data mining, and Mathematical Morphology. She is a member of IEEE and Editor of various reputed journals