

A Novel Approach for Face Pattern Identification and Illumination

Viniya.P¹,Peeroli.H²

PG scholar, Applied Electronics, Mohamed sathak Engineering college,Kilakarai,Tamilnadu, India¹

HOD, Department of ECE, MohamedSathak Engineering College,Kiakarai,Tamilnadu, India²

Abstract—Extracting facial feature is a key step in facial expression recognition (FER). Inaccurate feature extraction very often results in erroneous categorizing of facial expressions. Especially in robotic application, environmental factors such as illumination variation may cause FER system to extract feature inaccurately. In this paper, we propose a robust facial feature point extraction method to recognize facial expression in various lighting conditions. Before extracting facial features, a face is localized and segmented from a digitized image frame. Face preprocessing stage consists of face normalization and feature region localization steps to extract facial features efficiently. As regions of interest corresponding to relevant features are determined, Gabor jets are applied based on Gabor wavelet transformation to extract the facial points. Gabor jets are more invariable and reliable than gray-level values, which suffer from ambiguity as well as illumination variation while representing local features. The proposed algorithm has two advantages only one face training image is needed to train the classifier by using the facial block features with lower data dimensions, the proposed system is more computational efficiency. The main aim of this proposed work is to improve the accuracy of the face recognition system using the multiple training images.

Keywords—Facial expression recognition , Ada boost classifier, Dynamic Bayesian network,Gabor Transform.

I. INTRODUCTION

The tracking and recognition of facial activities from images or videos have attracted great attention in computer vision field. The recovery of facial activities in image sequence is an important and challenging problem. In recent years, plenty of computer vision techniques have been developed to track or recognize facial activities in three levels. First, in the bottom level, facial feature tracking, which usually detects and tracks prominent facial feature points (i.e., the facial landmarks) surrounding facial components (i.e., mouth, eyebrow, etc.), captures the detailed face shape information.

Second, facial actions recognition, i.e., recognize facial Action Units (AUs) defined in the Facial Action Coding System (FACS), try to recognize some meaningful facial activities (i.e., lid tightener, eyebrow raiser, etc). In the top level, facial expression analysis attempts to recognize facial expressions that represent the human emotional states. The facial feature tracking, AU recognition and expression recognition represent the facial activities in three levels from local to global , and they are interdependent problems.

For example, facial feature tracking can be used in the feature extraction stage in expression/AUs recognition, and expression/ AUs recognition results can provide a prior distribution for facial feature points. However, most current methods only track or recognize the facial activities in one or two levels, and track them separately, either ignoring their interactions or limiting the interaction to one way. In addition, the estimates obtained by image-based methods in each level are always uncertain and ambiguous because of noise, occlusion and the imperfect nature of the vision algorithm. In this paper, Dynamic Bayesian Network (DBN) are used to capture the facial interactions at different levels. In particular, not only the facial feature tracking can contribute to the expression/AUs recognition, but also the expression/AU recognition helps to further improve the facial feature tracking performance.

FACE RECOGNITION

Face recognition is one of the most important applications of Gabor wavelets. The face image is convolved with a set of Gabor wavelets and the resulting images are further processed for recognition purpose. The Gabor wavelets are usually called Gabor filters in

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3 , Issue 4 , April 2014

the scope of applications. There are various proposed approaches could be roughly classified into analytic and holistic approaches. Some feature points are detected from the face, especially the important facial landmarks such as eyes, noses, and mouths. These detected points are called the fiducial points, and the local features extracted on these points, distance and angle between these points, and some quantitative measures from the face are used for face recognition. In contrast to using information only from key feature points, holistic approaches extract features from the whole face image. Normalization on face size and rotation is a really important pre-processing to make the recognition robust. The eigenface based on principal component analysis (PCA) and the Fisher face based on linear discriminant analysis (LDA) are two of the most well known holistic approaches.

II. RELATED WORKS

A. Facial Feature Tracking

Facial feature points encode critical information about face shape and face shape deformation. Accurate location and tracking of facial feature points are important in the applications such as animation, computer graphics, etc. Generally, the facial feature points tracking technologies could be classified into two categories: model free and model-based tracking algorithms. Model free approaches are general purpose point trackers without the prior knowledge of the object. Each feature point is usually detected and tracked individually by performing a local search for the best matching position. However, the model free methods are susceptible to the inevitable tracking errors due to the aperture problem, noise, and occlusion.

Model based methods, such as Active Shape Model (ASM), Active Appearance Model (AAM), Direct Appearance Model (DAM) etc., on the other hand, focus on explicitly modeling the shape of the objects. The ASM proposed by Cootes et al. is a popular statistical model-based approach to represent deformable objects, where shapes are represented by a set of feature points. Feature points are first searched individually, and then Principal Component Analysis (PCA) is applied to analyze the models of shape variation so that the object shape can only deform in specific ways found in the training data. Robust parameter estimation and Gabor wavelets have also been employed in ASM to improve the robustness and accuracy of feature point search. The AAM and DAM are subsequently proposed to combine

constraints of both shape variation and texture variation. In the conventional statistical models, e.g. ASM, the feature points positions are updated (or projected) simultaneously. Intuitively, human faces have a sophisticated structure, and a simple parallel mechanism may not be adequate to describe the interactions among facial feature points.

B. Expression/AUs Recognition

Facial expression recognition systems usually try to recognize either six expressions or the AUs. Over the past decades, there has been extensive research on facial expression analysis. Current methods in this area can be grouped into two categories: image-based methods and model-based methods. Image-based approaches, which focus on recognizing facial actions by observing the representative facial appearance changes, usually try to classify expression or AUs independently and statically. This kind of method usually consists of two key stages. First, various facial features, such as optical flow, explicit feature measurement (e.g., length of wrinkles and degree of eye opening), Haar features, Local Binary Patterns (LBP) features, independent component analysis (ICA), feature points, Gabor wavelets etc., are extracted to represent the facial gestures or facial movements. Given the extracted facial features, the expression/AUs are identified by recognition engines, such as Neural Networks, Support Vector Machines (SVM), rule-based approach, AdaBoost classifiers, Sparse Representation (SR) classifiers etc. The common weakness of image-based methods for AU recognition is that they tend to recognize each AU or certain AU combinations individually and statically directly from the image data, ignoring the semantic and dynamic relationships among AUs, although some of them analyze the temporal properties of facial features. Model-based methods overcome this weakness by making use of the relationships among AUs, and recognize the AUs simultaneously.

C. Simultaneous facial activity tracking / recognition

The idea of combining tracking with recognition has been attempted before, such as simultaneous facial feature tracking and expression recognition and integrating face tracking with video coding. However, in most of these works, the interaction between facial feature tracking and facial expression recognition is one-way, i.e., facial feature tracking results are fed to facial expression recognition. There is no feedback from the recognition results to facial feature tracking. Most recently, Dornaika et al. and Chen & Ji

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3 , Issue 4 , April 2014

improved the facial feature tracking performance by involving the facial expression recognition results. However, in Simultaneous facial action tracking and expression recognition in the presence of head motion, they only modeled six expressions and they need to retrain the model for a new subject, while in a hierarchical framework for simultaneous facial activity tracking, they represented all upper facial action units in one vector node and in such a way, they ignored the semantic relationships among AUs, which is a key point to improve the AU recognition accuracy.

III. PROPOSED SYSTEM

A. Facial Activity Model

Dynamic Bayesian network is a directed graphical model, DBN is more general to capture complex relationships among variables. Specifically, the global facial expression is the main cause to produce certain AU configurations, which in turn causes local muscle movements, and hence feature points movements. For example, a global facial expression (e.g. Happiness) dictates the AU configurations, which in turn dictates the facial muscle movement and hence the facial feature point positions. For the facial expression in the top level, we will focus on recognizing six basic facial expressions, i.e., happiness, surprise, sadness, fear, disgust and anger. Though psychologists agree presently that there are ten basic human emotions, most current research in facial expression recognition mainly focuses on six major emotions, partially because they are the most basic, and culturally and ethnically independent expressions and partially because most current facial expression databases provide the six emotion labels.

various computer vision techniques are used to track the facial feature points, and to get the measurements of facial motions, i.e., AUs. These measurements are then used as evidence to infer the true states of the three level facial activities simultaneously.

B. Gabor Wavelet Representation of Faces

Daugman generalized the Gabor function to the following 2D form in order to model the receptive fields of the orientation selective simple cells:

$$\Psi_i(\vec{x}) = \frac{|\hat{k}_i|^2}{\sigma^2} c \frac{\|\hat{k}_i\|^2 \|\vec{x}\|^2}{2\sigma^2} \left[c^{i\hat{k}_i \cdot \vec{x}} - c^{\frac{\sigma^2}{2}} \right] \quad (1)$$

Each Ψ_i is a plane wave characterized by the vector \hat{k}_i enveloped by a Gaussian function, where σ is the standard deviation of this Gaussian. The center frequency of i th filter is given by the characteristic wave vector,

$$\vec{k}_i = \begin{pmatrix} k_{ix} \\ k_{iy} \end{pmatrix} = \begin{pmatrix} k_v \cos \theta_\mu \\ k_v \sin \theta_\mu \end{pmatrix}$$

having a scale and orientation given by (k_v, θ_μ) . The first term in the brackets determines the oscillatory part of the kernel, and the second term compensates for the DC value of the kernel. Subtracting the DC response, Gabor filters become insensitive to the overall level of illumination. Daugman and others have proposed that an ensemble of simple cells is best modeled as a family of 2D Gabor wavelets sampling the frequency domain in a log-polar manner. This class is equivalent to a family of affine coherent states generated by rotation and dilation. The decomposition of an image I into these states is called the wavelet transform of the image:

$$R_i(\vec{x}) = \int I(\vec{x}') \Psi_i(\vec{x} - \vec{x}') d\vec{x}' \quad (3)$$

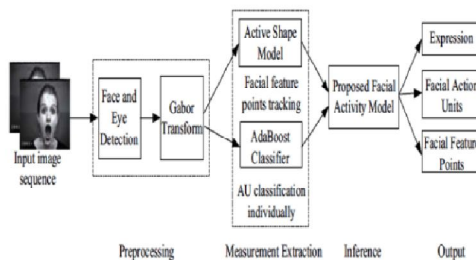


Fig 1: Block Diagram

The proposed facial activity recognition system consists of two main stages: offline facial activity model construction and online facial motion measurement and inference. Specifically, using training data and subjective domain knowledge, the facial activity model is constructed offline. During the online recognition,

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3 , Issue 4 , April 2014

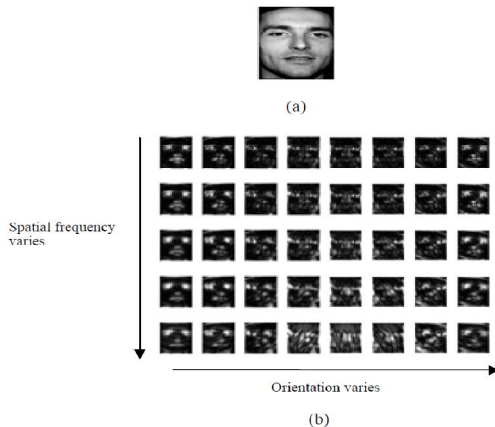


Fig 2: Example of A Facial Image Response to Above Gabor Filters
A)Original Face Image (From Stirling Database) B) Filter Responses

Input face image and the amplitude of the Gabor filter responses are shown in figure 2. One of the techniques used in the literature for Gabor based face recognition is based on using the response of a grid representing the facial topography for coding the face. Instead of using the graph nodes, high-energized points can be used in comparisons which forms the basis of this work. This approach not only reduces computational complexity, but also improves the performance in the presence of occlusions.

C. Ada Boost Classifier

Boosting is a method to combine a collection of weak classification functions (weak learner) to form a stronger classifier. AdaBoost is an adaptive algorithm to boost a sequence of classifiers. AdaBoost is a kind of large margin classifiers. A weak classifier in simple terms is a decision rule that classifies a test sample as either a positive or a negative sample. A weighted linear combination of these weak classifiers forms a strong classifier with improved detection performance. The boosting process has two stages corresponding to training and detection. In the training stage a very large set of labeled samples is used to identify the better performing weak classifiers, and a strong classifier network is constructed by a weighted linear combination of these weak classifiers. The output of the training stage is a trained classifier network that can be used in the detection phase to classify samples as positive or negative. The idea of boosting is to combine a set of simple rules or weak classifiers to form an ensemble such that the performance of a single ensemble member is improved. For example given a family of weak

classifiers and a set of training data consisting of positive and negative samples, the adaboost approach can be used to select a subset of weak classifiers and the classification function.

Adaboost requires no prior knowledge, that is no information about the structure or features of the face are required when used for face detection. Given a set of training samples and weak classifiers the boosting process automatically chooses the optimal feature set and the classifier function. The approach is adaptive in the sense that misclassified samples are given higher weight in order to increase the discriminative power of the classifier. As a result, easily classified samples detected in the burst iteration and have less weight and harder samples with higher weights are used to train the later iterations. The theoretical training error converges to zero as proved by Freund and Schapire (1995). The training for a set of positive and negative samples reaches zero after a finite number of iterations. Given a feature set

$$T = \{ (x_1, y_1), (x_2, y_2), (x_3, y_3), \dots \}$$

where x_i is the training sample and y_i is a binary value of the sample class (1 is positive, 0 negative). A final boosted classifier network is formed from the subset of given features after an arbitrary number of iterations as shown in Equation .

$$f(x) = \sum_{t=1}^T \alpha_t h_t(x) \tag{5}$$

where α_t is the weight assigned to the t th classifier, and h_t is the classifier decision. Adaboost training is an iterative process and the accuracy of the final classifier function depends on the number of iterations and whether the training error converges to zero after finite number of iterations. A classifier is trained as follows by Viola et.

- Given example images $(x_1; y_1), (x_2; y_2) \dots, (x_n; y_n)$ where $y_i = [0; 1]$ for negative and positive samples respectively.
- Initialize weights $w_1; i = 1/2m; 1/2l$ for $y_i = [0; 1]$ respectively, where m and l are the number of negative and positive samples respectively.

IV. EXPERIMENTAL RESULT

In the experimental result showing the two conditions first one is illumination condition another one is facial expression condition. Test image and equivalent image of the illumination image is shown in the figure.

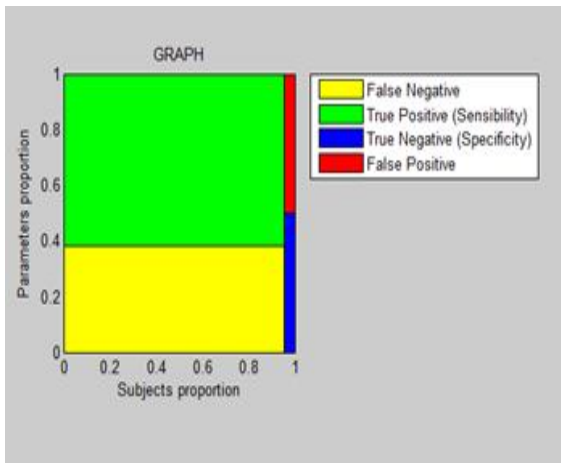
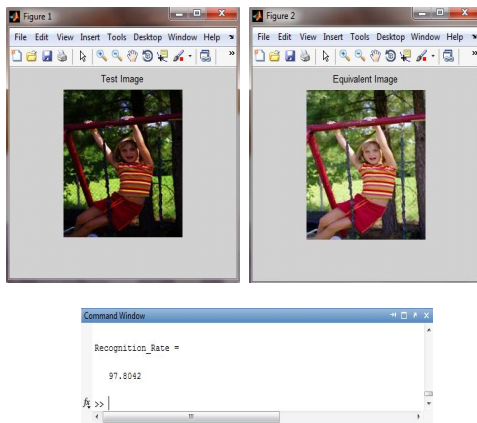
International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

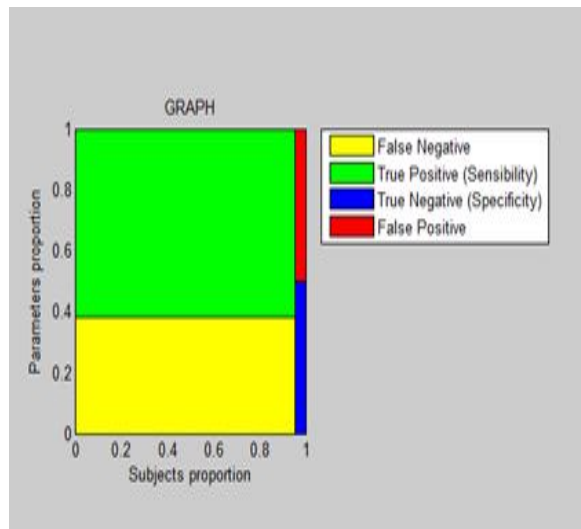
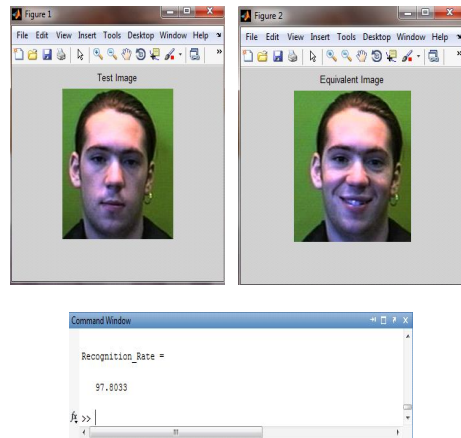
Vol. 3 , Issue 4 , April 2014

Recognition Rate of the original image is 97.8042. The original image and the test image is rotated to achieve the original test image.

ILLUMINATION



FACIAL EXPRESSION



Test image and equivalent image of the facial expression image is shown in the figure. The facial expression Recognition Rate is 97.8033. The graph is plot between parameters proportion and subjects proportion. Four samples are used to construct the plot. The four samples are False Negative, True Positive (Sensibility), True Negative (Specificity), False Positive. Sensitivity is the probability that test is positive on unhealthy data. specificity is the probability that test is negative on healthy data.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3 , Issue 4 , April 2014

V.CONCLUSION

In this paper, we proposed a hierarchical framework based on Dynamic Bayesian Network for simultaneous facial feature tracking and facial expression recognition. By systematically representing and modeling inter relationships among different levels of facial activities, as well as the temporal evolution information, the proposed model achieved significant improvement for both facial feature tracking and AU recognition. The improvements for facial feature points and AUs come mainly from combining the facial action model with the image measurements. Specifically, the erroneous facial feature measurements and the AU measurements can be compensated by the model's build-in relationships among different levels of facial activities, and the build-in temporal relationships. Since our model systematically captures and combines the prior knowledge with the image measurements, with improved image-based computer vision technology, our system may achieve better results with little changes to the model. In this paper, we evaluate our model on posed expression databases from frontal view images. In the future work, we plan to introduce the rigid head movements, i.e., head pose, into the model to handle multi view faces.

REFERENCES

- [1] M. Valstar and M. Pantic, "Fully automatic recognition of the temporal phases of facial actions," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 1, pp. 28–43, Feb. 2012.
- [2] H. Dibeklioglu, A. A. Salah, and T. Gevers, "A statistical method for 2-D facial landmarking," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 844–858, Feb. 2012.
- [3] L. Zhang and D. Tjondronegoro, "Facial expression recognition using facial movement features," *IEEE Trans. Affect. Comput.*, vol. 2, no. 4, pp. 219–229, Oct.–Dec. 2011.
- [4] J. Chen and Q. Ji, "A hierarchical framework for simultaneous facial activity tracking," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recognit.*, Mar. 2011, pp. 679–686.
- [5] S. Koelstra, M. Pantic, and I. Patras, "A dynamic texture-based approach to recognition of facial actions and their temporal models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 1940–1954, Nov. 2010.
- [6] Y. Tong, J. Chen, and Q. Ji, "A unified probabilistic framework for spontaneous facial activity modeling and understanding," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 2, pp. 258–273, Feb. 2010.
- [7] F. Dornaika and F. Davoine, "Simultaneous facial action tracking and expression recognition in the presence of head motion," *Int. J. Comput. Vis.*, vol. 76, no. 3, pp. 257–281, 2008.
- [8] Y. Tong, W. Liao, and Q. Ji, "Facial action unit recognition by exploiting Their dynamic and semantic relationships," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1683–1699, Oct. 2007.
- [9] Y. Tong, Y. Wang, Z. Zhu, and Q. Ji, "Robust facial feature tracking under varying face pose and facial expression," *Pattern Recognit.*, vol. 40, no. 11, pp. 3195–3208, 2007.
- [10] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun. 2007.