International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol.3, Issue.3, May-June. 2013 pp-1339-1342 ISSN: 2249-6645

A Novel Method for Movie Character Identification and its Facial Expression Recognition

M. Dharmateja Purna, [1] N. Praveen[2]
[1]*M.Tech, Sri Sunflower College of Engineering & Technology, Lankapalli*
[2]*Asst. Professor, Dept. of ECE, Sri Sunflower College of Engineering & Technology, Lankapalli*

***Abstract*:** *Image processing is a method to convert an image into a digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. Face recognition is an important concept in image processing. The input of a face recognition system is always an image or video (movie). The output is an identification or verification of the subject or subjects that appear in the image or movie. Another important concept in image processing is facial expression recognition. Human facial expression recognition has attracted much attention in recent years because of its importance in realizing highly intelligent human machine interfaces. Facial expression plays important role in cognition of human emotions and facial expression recognition is the base of emotions understanding. In this paper, we propose a novel method for both face recognition and its facial expression. This is very helpful in situation when users want to identify human characters in movies and their facial expressions.*

***Keywords*:** *Face recognition, Facial expression, FCP, Template.*

I. INTRODUCTION

There are number of applications where face recognition [1] can play an important role including biometric authentication, high technology surveillance and security systems image retrieval and passive demographical data collections .it is observable that the behavior and social interaction are face recognition system could have great impact in improving human computer interaction systems in such a way as to make them be more user-friendly and acting more human-like. It is unarguable that the face is one the most important feature that characterizes human beings. By only looking ones faces, we are not only able to tell who they are but also perceive a lot of information such as their ages, emotions and names. This is why face recognition has received much interest in computer vision research community over past two decades. Figure 1 shows an example of movie character identification.

There are two main steps involved in recognizing names of humans presented in an image .These are- face detection and name classification [2], which are applied consecutively. In order to exploit uniqueness of faces in name recognition, the first step is to detect and localize those faces in the given images. This is the task achieved by face detection systems.

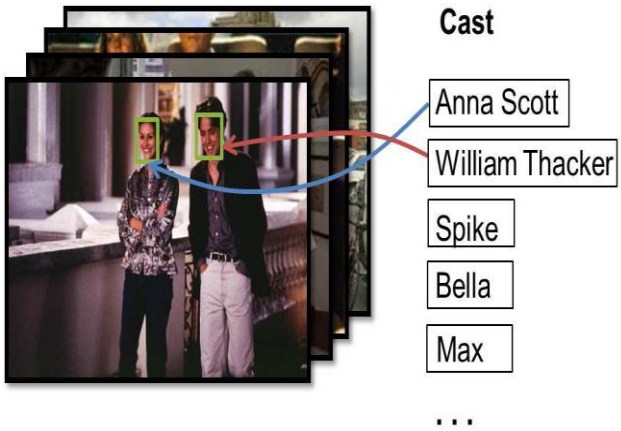

Figure 1: Example of movie character identification

Sometimes we want to know the facial expression (happy, sad, smiley, fear) of the recognized face. Facial expression recognition [3][4] is a process performed by humans or computers, which consists of(figure 2):

1. Locating faces in the scene (e.g., in an image- this step is also referred to as face detection),
2. Extracting facial features from the detected face region (e.g., detecting the shape of the facial components or describing the texture of the skin in a facial area; this step is referred to as facial feature extraction),
3. Analyzing the motion of facial features and the changes in the appearance of facial features and classifying this information into some facial expression- interpretative categories such as facial muscle activations like smile or frown, emotion (affect) categories like happiness or anger, attitude categories like (dis)liking or ambivalence, etc. (this step is also referred to as facial expression interpretation).

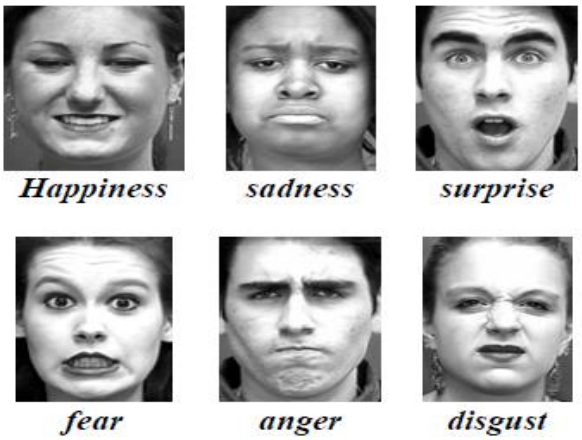

Figure 2: Different facial expressions

Monitoring and interpreting facial expressions can also provide important information to police, lawyers, security, and intelligence agents regarding person's identity (research in psychology suggests that facial expression recognition is much easier in familiar persons because it seems that people display the same, ''typical'' patterns of facial behavior in the same situations), deception (relevant studies in psychology suggest that visual features of facial expression function as cues to deception), and attitude (research in psychology indicates that social signals including accord and mirroring – mimicry of facial expressions, postures, etc., of one's interaction partner – are typical, usually unconscious gestures of wanting to get along with and be liked by the interaction partner). Automated facial reaction monitoring systems could form a valuable tool in law enforcement, as now only informal interpretations are typically used.

II. RELATED WORK

In [5], [6], the authors proposed, the faces are clustered by appearance and faces of a particular character are expected to be collected in a few pure clusters. Names for the identified clusters are then manually selected from the cast list. In [7], the authors proposed to manually label an initial set of face clusters and further cluster the rest face instances based on clothing within scenes. In [8], the authors have addressed the problem of finding particular characters by building a model/classifier of the character's appearance from user-provided training data. An interesting work combining character identification with web image retrieval is proposed in [9]. The character names in the cast are used as queries to search the face images and constitute gallery set. The probe face tracks in the movie are then identified as one of the characters by multi task joint sparse representation and classification.

In [10], the authors proposed to combine the film script with the subtitle for local face-name matching. Researchers from University of Pennsylvania utilized the readily available time-stamped resource, the closed captions, which is demonstrated more reliable than the OCR-based subtitles [11]. They investigated on the ambiguity issues in the local alignment between the video, screenplay and closed captions. A partially-supervised multi class classification problem is formulated. Recently, they attempted to address the character identification problem without the use of screenplay [12].

In [13], the authors proposed the facial action coding system (FACS) which represents the facial expression by a set of facial action units. In [14], the authors proposed an approach for analyzing and representing the dynamics of facial expression. Their system consists of locating of tracking the prominent facial features, optical flow analysis, and the classification. In [15], the authors extended the work of [14] by using connectionist architecture. Individual emotion networks were trained by viewing a set of sequences of one emotion for many of the objects. The trained neural network was then tested for the emotion recognition. In [16], the authors provided a facial expression representation by characterizing facial muscle activation. The facial motion estimation is operated by fitting the 3D deformable facial model to the face in an image for the muscle based representation. In [17], the authors developed a facial expression recognition method by using a synergetic pattern recognition approach. In [18], the authors proposed a facial expression recognition method to identify the shape of the mouth feature only. In[19], the authors used simple measurements (0 or 1) of the forehead wrinkle, eye opening, nostril furrow deepening, mouth opening, and eyebrow motion to recognize human facial expression.

III. PROPOSED WORK

The proposed work is shown in Figure 3. Our proposed work for the character recognition was motivated by the Bag-of-Features method [20]. The Bag-of-Features method extracts the feature points (i.e., image points that are described not necessarily by their color/intensity values, but by their local neighborhood based on, e.g., gradient information) from a set of training images. In the feature space, the feature points are grouped by a clustering algorithm. Based on the resulting clusters (all clusters together are referred to as code book and one cluster is referred to as visual word), occurrence histograms are then generated for each body part image. A classifier is then trained on the obtained histograms. Occurrence histograms reflect how many feature points are assigned to each of the visual word. Our approach is build on SIFT- [21] and CIE L*u*v* color-based code books that are obtained by clustering with k-means. A non-linear multi-class Support Vector Machine (SVM) is learned on the occurrence histograms.

The trained Support Vector Machines (or SVM models) are then used to predict the identity of the detected person. Probabilistic votes of connected body parts (i.e., body parts that belong to one and the same person) are combined for a more stable prediction. The training data is generated from an annotation data set in which the name of the corresponding character is noted for each of the body part. Based on the obtained annotation data, codebooks are generated and SVM models are learned. The codebooks and the SVM models are then applied subsequently on the entire video file. In this way, particular (human) characters are recognized at different points in time in a given video file. After obtaining a particular character, we need to identify the facial expression of that character.

To identify facial expression, template matching is being carried out by making use of convolution and correlation coefficients for the highest and perfect matching. The desired eyes, eyebrows and mouth templates are being excerpt from the image and the extracted results are shown in the form of bounded rectangles. The Facial characteristics points (FCP's) is being computed by knowing the top left coordinate of each template bounded by rectangles. Once we obtained the parameters from FCP's we set the threshold value and then proceed for creation of Decision tree. A decision tree is a classifier in the form of a tree structure. Information gain (IG) is used to select the most useful attributes for classification:

- The entropy of total data set is calculated
- The dataset is then split on the different attributes.
- The entropy of each branch is calculated then it is added proportionally to get the total entropy for the split.
- The resulting entropy is subtracted from the entropy before the split; with the result is the information gain.
- The attribute that have the largest information gain is chosen for the decision node.
- A branch set with a entropy of zero is the leaf node.
- Otherwise, further splitting to classify its data set.
- The ID3 algorithm is run recursively on the non leaf branches until all data is classified.

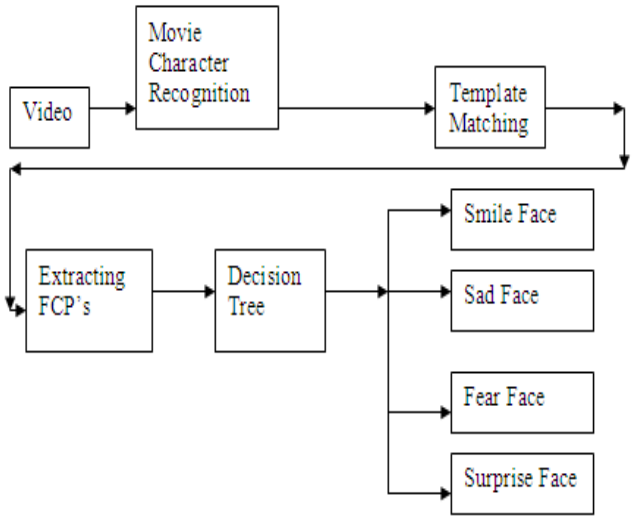

Figure 3: Proposed Method

IV. CONCLUSIONS

The proliferation of TV and movie provides large amount of digital video data. This has led to the requirement of efficient and effective techniques for movie or image content understanding and organization. Automatic image or video annotation is one of such key techniques. In this paper our focus is on annotating characters in the movies, which is called movie character identification and their facial expressions. The movie character identification is performed based on Bag-of-Features method extracts the feature points and SIFT and CIE L*u*v* color-based code books that are obtained by clustering with k-means. To identify facial expression, template matching is being carried out by making use of convolution and correlation coefficients for the highest and perfect matching. The Facial characteristics points (FCP's) is being computed by knowing the top left coordinate of each template bounded by rectangles. Once we obtained the parameters from FCP's we set the threshold value and then proceed for creation of Decision tree. After classification, we obtain the required facial expression of the identified character.

REFERENCES

[1] J. Stallkamp, H. K. Ekenel, and R. Stiefelhagen, "Video-based face recognition on real-world data," in Proc. Int. Conf. Comput. Vis., 2007, pp. 1–8.
[2] T. L. Berg, A. C. Berg, J. Edwards, M. Maire, R.White, Y. W. Teh, E. G. Learned-Miller, and D. A. Forsyth, "Names and faces in the news," in Proc. Comput. Vis. Pattern. Recognit., 2004.
[3] Tian, Y.L., Kanade, T., Cohn, J.F.: Facial expression analysis. In: Li, S.Z., Jain, A.K. (eds.) Handbook of Face Recognition, pp. 247–276. Springer, New York (2005)

[4] Pantic, M., Bartlett, M.S.: Machine analysis of facial expressions. In: Delac, K., Grgic, M. (eds.) Face Recognition, pp. 377–416. I-Tech Education and Publishing, Vienna, Austria (2007)
[5] A. W. Fitzgibbon and A. Zisserman, "On affine invariant clustering and automatic cast listing in movies," in ECCV (3), 2002, pp. 304–320.
[6] O. Arandjelovic and R. Cipolla, "Automatic cast listing in feature-length films with anisotropic manifold space," in CVPR (2), 2006, pp. 1513– 1520.
[7] D. Ramanan, S. Baker, and S. Kakade, "Leveraging archival video for building face datasets," in ICCV, 2007, pp. 1–8.
[8] M. Everingham and A. Zisserman, "Identifying individuals in video by combining"generative" and discriminative head models," in ICCV, 2005, pp. 1103–1110.
[9] M. Xu, X. Yuan, J. Shen, and S. Yan, "Cast2face: character identification in movie with actor-character correspondence," in ACM Multimedia, 2010, pp. 831–834.
[10] M. Everingham, J. Sivic, and A. Zissserman, "Hello! my name is... buffy automatic naming of characters in tv video," in Proceedings of BMVC, 2006, pp. 889–908.
[11] rT. Cour, C. Jordan, E. Miltsakaki, and B. Taskar, "Movie/script: Alignment and parsing of video and text transcription," in ECCV (4), 2008, pp. 158–171.
[12] T. Cour, B. Sapp, A. Nagle, and B. Taskar, "Talking pictures: Temporal grouping and dialog-supervised person recognition," in CVPR, 2010, pp. 1014–1021.
[13] P. Ekman and W. V. Friesen, Measuring facial movement with facial action coding system, in Emotion in Human Face (P. Ekman, Ed.), Cambridge Univ. Press, Cambridge, 1982.
[14] Y. Yacoob and L. Davis, Computing spatio-temporal representation of human face, in CVPR'94, June 21–23, 1994, Seattle, pp. 70–75. USA.
[15] M. Rosenblum, Y. Yacoob, and L. Davis,Human emotion recognition from motion using a radial basis function network architecture, in Proc. of IEEE Workshop on Motion of Non-Ridgid and Articulated Objects, Nov. 11–12, 1994, Austin, pp. 43–49, USA.
[16] I. A. Essa and A. Pentland, Facial expression recognition using virtually extracted facial action parameters, in Proc. of Int. Workshop on Auto. Face- and Gesture Recognition, pp. 35–40, Zurich, 1995.
[17] P. Vanger, R. Honlinger, and H. Haken, Applications of synergetics in decoding facial expression of emotion, in Proc. of Int. Workshop on Auto. Face- and Gesture Recognition, pp. 24–29, Zurich, 1995.
[18] Y. Moses, D. Reynard, and A. Blake, Determining facial expression in real-time, in Proc. of Int. Workshop on Automatic Face- and Gesture Recognition, pp. 332–337, Zurich, 1995.
[19] Y. Kitamura, J. Ohya, N. Ahuja, and F. Kishino, Computational taxonomy and recognition of facial expression, in Proc. of ACCV'93 Nov. 23–25, 1993, Osaka, Japan.
[20] Chris Dance, Jutta Willamowski, Lixin Fan, Cedric Bray, and Gabriela Csurka. Visual categorization with bags of keypoints. In ECCV International Workshop on Statistical Learning in Computer Vision, 2004.
[21] D. G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110, 2004.