

# Movie Character Identification using Global Face-Name Graph Matching

Miss. Swati A. Moghe

Department of Electronics & Telecommunication,  
Sinhgad College of Engineering,  
Pune, Maharashtra, India

Prof. Vaibhav V. Dixit

Department of Electronics & Telecommunication,  
Sinhgad College of Engineering,  
Pune, Maharashtra, India

**Abstract**— Identification of characters in movies, even though very intuitive to humans, still poses an important challenge to computer method. Identification of character face and labelling them with corresponding names in script is the main objective. Huge variations in appearance of each character in different frames make it a difficult task. Many methods are implanted till date and tested under clean environment. But in presence of noise generated in complex movie scenes due to the face tracking and face clustering, the accuracy of system is affected. By introducing graph matching algorithm and incorporating noise insensitive relationship, present character identification approach has been enhanced in this paper.

**Index Terms**— Graph Matching, Character Identification, Ordinal Graph, Graph Partition, Face Identification.

## I. INTRODUCTION

Huge amount of data is generated everyday due to flourishing development in movie industry. Providing media content description, indexing and organization for user to browse and retrieve content of interest, is important for media distributors. To manage this, efficient and effective technique is required which understand the movie contents and appropriately organise it. Annotating character in the movie is called movie character identification. Identification of faces in movie is a demanding task in the current research. In movies, characters are the focus center of interests for the viewers. Meaningful presentation of movie content is provided by the occurrence of characters. Automatic character identification is essential for movie indexing, retrieval, and other applications. Character identification is challenging task in computer vision. Due to some reason like, different background disturbances are present when concern with different types of movies, confusion can also create when there are multiple characters in the same scene and sometimes same character appears quite differently during the movie, there may be huge pose, wearing, clothing, expression, illumination variation, even makeup and hairstyle changes.

In related work, some methods are used like Cast list based methods [1] only utilize the case list textual resource. Names for the clusters are then manually selected from the cast list. But this method requires manual labelling and classification performance is poor due to the large intra-class variances. Subtitle or closed caption method [2] used in movie script. In this method require subtitle for face-name matching. The local matching based methods [3] require the time-fixed information, which is extract by subtitle and most of the

movies subtitle is unavailable. But Subtitle caption method and local matching based method is more sensitive to the face detection and tracking noises. In [4], used global matching based methods open the possibility of character identification without subtitle or closed caption. But this method requires very complex identification algorithm.

Global face-name graph matching with pre-cluster specified is more insensitive to noises and easy to understand compared to other methods. The accuracy of this method can be improved by making a few changes in the algorithm. In this paper; global face name graph matching without pre-cluster specified has been discussed.

## II. SYSTEM MODEL

A face-name graph matching without clusters pre-specified based method is proposed for movie character identification. In this method external script sources are required for matching purpose. Also, no cluster number is required for the face tracks clustering step and a graph partition component is added before ordinal graph representation.

In this method, K-means clustering algorithm is used for face tracks clustering, in which number of distinct speakers sets the number of clusters in that movie. In the movie, the name in script and face cluster occurs repeatedly with respect to corresponding face graph and name graph. Face track clustering plays essential role in movie character identification.

When clustering is performed number of cluster is already set. The face graph is bounded to have similar number of vertexes in name graph. In this method, modified traditional matching method by using ordinal graphs for robust representation and introduce an ECGM graph matching method.

Features are calculated for the obtaining facial regions and are stored in the database. The stored features are compared against with those that are obtained in the frame using ECGM algorithm. The matched region is displayed with the corresponding person name as text.

## III. METHODOLOGY: FACE-NAME GRAPH MATCHING WITH NUMBER CLUSTERS PRE-SECIFIED

Face-name graph matching based method is proposed for movie character identification. The flow chart of the method is shown in Figure 1. Consider a movie clip along with related script. Face tracks are extracted from the movie video using

skin detection algorithm. Gabor filter is used for feature extraction. Frequency and orientation representation of Gabor filters are similar to of human visual system and they are found to be mainly suitable for texture discrimination and representation.

In spatial domain Gabor filter is a Gaussian kernel function modulate by a sinusoidal plane wave. A set of Gabor filters with different orientations and frequency; it may useful for extracting required features from given image. Gabor filter are directly related to wavelets, designed for a number of rotations and dilations ,though expansions is not apply for Gabor wavelets as it requires compute of bi-orthogonal wavelets, which is time consuming.

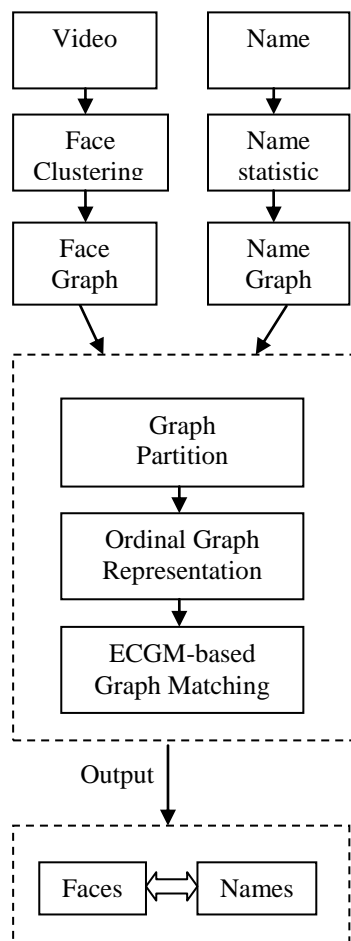


Figure 1: Face name graph matching without clusters pre-specified

As a result generally a filter bank consists of Gabor filters with a variety of rotations and scales are created. Once feature extraction is achieved, the next step is face clustering, namely k-means clustering.

In k-means clustering algorithm, the dataset is partitioned into K cluster and data points are randomly assigned to that particular clusters resulting in clusters that have generally the same number of data points. For each data point, compute the distance from the data point to each cluster. If the data point is close to its own cluster, leave it where it is. If the data point is not close to its own cluster, then shift it into the close cluster. Above procedure is repeated till it completely passes through

all the data point so as to group each point into a cluster. At this stage clusters are stable and clustering process ends.

After clustering, face graph and name graph is constructed using radial basis function (RBF).RBF is a real value function that value depend only on distance from origin or on the distance from any other point centre. A radial basis- function is defined on a Euclidean space whose value at each point depends only on the distance between point and the origin. The Euclidean distance between points p and q is the length of the line segment connecting them (p,q).

If  $p = (p_1, p_2, \dots, p_n)$  and  $q = (q_1, q_2, \dots, q_n)$  are two points in Euclidean  $n$ -space, then the distance from p to q, or from q to p is given by

$$d_{(p,q)} = d_{(q,p)} = \sqrt{\sum_{i=1}^n (q_i^2 - p_i^2)} \quad (1)$$

Face graph and name graph consist of similar names in script and face clusters. The initial affinity graph is constructed so as to link between the face tracks and names which can associate using temporal link or frame-wise assignment. The graph is pruned based upon the repetition of any of the nodes. The graph is mathematically represented by  $G(v,e)$ , where, v defines a particular node and e defined the edge between two nodes. Certain visual primitives or key points are extracted from the face track which acts as feature points. These feature points are used for the initial clustering of graph. The clustered graph is now supposed to consist of that group which contains each of the characters present in the movie.

For matching purpose error correcting graph matching algorithm (ECGM) is used. In movie, the interactions among characters appear into a relationship network. Co-occurrence of names in script and faces in videos can represent such interactions. The more scenes where two characters appear together, the closer they are, and the larger the edge weights between them. The final graph is modified based on the ECGM algorithm where the priority is given in such a manner so as to minimize the clustering error. After getting character-wise clustering, the matching of the character and the names can be achieved by combining the face graph and the name graph. By overlapping the graphs can find the minimum edge distance between two nodes. For this scheme, in advance no cluster number is required and face-tracks are clustered based on their intrinsic data structure. Therefore, this scheme provide certain robustness to intra-class variance, which is very general in movies where movie characters change look significantly or go through a lengthy time period. Concerning the movie cast cannot contain pedestrians whose face is detected and face track is added , restrict the number of face tracks cluster the same as that of name from movie cast will be deteriorate the clustering process. In movie there is a few chance that movie cast does not cover up all the characters. In this case, pre-specification for the face clusters is dangerous: face tracks from different characters will be mix together and graph matching may be fail. So we avoid pre-specification for the face clusters.

IV. RESULTS

Consider movie along with related script. Features of face track get extracted by Gabor filter. It gives the promising results. After extracting the features from face track, certain visual primitives or key points are extracted from the face track which acts as feature points.

These feature points are used for the initial clustering of graph. For the formation of face graph and name graph used radial basis function. The clusters graph consists of all the characters present in the movie.

The final output is based on the ECGM algorithm where minimization of clustering error is prioritized. By combining the face graph and the name graph, will find the character-wise clustering.

But the previous system fails to identify the character where there is change in illumination, pose and expression as shown in figure 2. This has been corrected in propose system.



Figure 2: Character identification from movie in previous method

The proposed system identifies the characters starting from the left hand side of the frame. Fig. 2 shows the results. First frame has three characters, each character is identified and their names are displayed. Similarly, second frame contains only one character which is labelled after identification.

The matching is done by overlapping the graphs and finds the minimum edge distance between two nodes. Figure 2 shows that face-name matching of character identification using ECGM algorithm and display the recognize character name as text.



Figure 3: Character identification from movie

For face track detection accuracy performed a simple experiment by randomly selecting three clips from three movies. The statistics of performance are shown in Table I. Table I show the face track detection accuracy. The results of three movies clips are shown in table. It can be seen that the accuracy of the clip 1 and clip 2 is higher than clip 3. It is due more variation of expression, face position, illumination variation and clothing.

TABLE I. FACE TRACK DETECTION ACCURACY

Clips	Face track	Track detected	Accuracy
1	370	352	95.13%
2	189	177	93.65%
3	279	255	91.39%

Precision is the proportion of correctly labelled tracks. Precision is ratio of face tracks correctly classified to the face tracks classified. Precision is the fraction of retrieve instance that are significant. Figure 4 shows the Precision graph. Precision graph shows the how many face tracks are correctly classified from total face track classified.

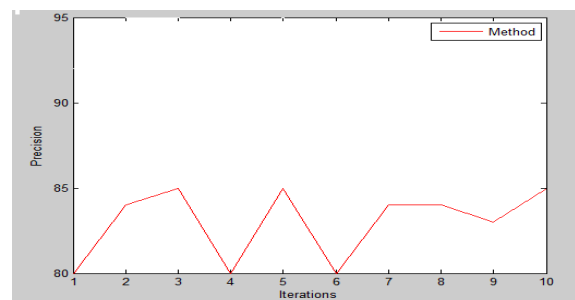


Figure 4: Precision graph

Recall here means the proportion of tracks which are assigned a name. Recall is ratio of face tracks classified to the total face tracks. Recall is the fraction of significant instance that are retrieve. Figure 5 shows the recall graph. Recall graph shows the how many face tracks are classified from total face tracks.

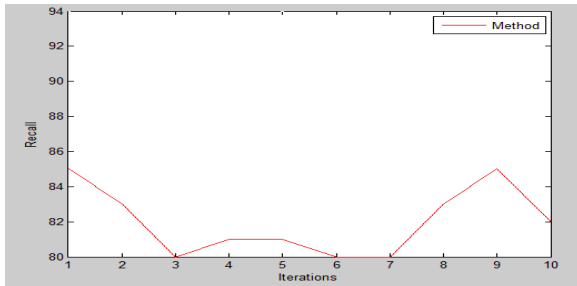


Figure 5: Recall graph

## V. CONCLUSIONS

Novel framework for character identification in movies is proposed in this paper. Previous work mostly relied on local matching where as the proposed system is based on global matching method. Name-face association between the name affinity network and face affinity network is built utilizing a graph matching method. As an application relationship between characters is mined and a platform for character specific movie browsing is provided. The proposed method is helpful to achieve better results for identification. An experimental result shows that this method gives improved performance. Character identification is easy using face- name graph matching. In future, goal is to exploiting more character relationships to improve the robustness.

## REFERENCES

- [1] J. Sang, C. Liang, C. Xu, and J. Cheng, "Robust video character identification and the sensitivity analysis," in *ICME*, 2011, pp. 1–6
- [2] M. Everingham and A. Zisserman, "Identifying individuals in video by combining "generative" and discriminative head models," in *ICCV*, 2005, pp. 1103–1110.
- [3] J. Stallkamp, H. K. Ekenel, and R. Stiefelhagen, "Video-based face recognition on real-world data." in *ICCV*, 2007, pp. 1–8.
- [4] M. Xu, X. Yuan, J. Shen, and S. Yan, "Cast2face: character identification in movie with actor-character.
- [5] J. Stallkamp, H. K. Ekenel, and R. Stiefelhagen, "Video-based face recognition on real-world data." in *ICCV*, 2007, pp. 1–8.
- [6] Jitao. Sang, Changsheng Xu, "Robust face-name graph matching for movie character identification" *IEEE Transactions on Multimedia*, vol. 10, no. 10, 2012.
- [7] Y. Zhang, C. Xu, H. Lu, and Y. Huang, "Character identification in feature-length movies using global face-name matching," *IEEE Trans. Multimedia*, vol. 11, no. 7, pp. 1276–1288, November 2009.
- [8] M. Everingham, J. Sivic, and A. Zisserman, "Taking the bite out of automated naming of characters in tv video," in *Journal of Image and Vision Computing*, 2009, pp. 545–559.
- [9] C. Liang, C. Xu, J. Cheng, and H. Lu, "Tvparser: An automatic tv video parsing method," in *CVPR*, 2011, pp. 3377–3384.
- [10] O. Arandjelovic and R. Cipolla, "Automatic cast listing in feature-length movies with anisotropic manifold space," in *CVPR (2)*, 2006, pp. 1513–1520
- [11] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, pp. 972–977, 2007.
- [12] J. Sang and C. Xu, "Character-based movie summarization," in *ACM*, 2010.