Automatic Naming of Character using Video Streaming for Face Recognition with Graph Matching

Nivedita.R.Pandey PG Student at MET's IOE Bhujbal Knowledge City, Nasik, Maharashtra, India, pandeynivedita7@gmail.com

Ranjan.P.Dahake PG Student at MET's IOE Bhujbal Knowledge City, Nasik, Maharashtra, India, *dahakeranjan@gmail.com*

Abstract— Recognition of face using web camera and video is great problem and hot topic in compute vision and biometrics over many years. There is a huge variation in image appearance of each character and weakness in ambiguity of available annotation. Higher precision can be achieved by combining multiple sources from both visual and textual. The contribution of the work is: - 1] Relationship between character (images with 9 pixels are obtained using haar features) and name re obtained.2] Faces are recognized using both video streaming and web camera.3] Graphical matching of both names and character is obtained in graph format.4] Easy help for search character in video. Comparing to traditional face analysis video based face recognition has multiple advantages for more abundant information to improve accuracy and robustness.

Keywords- Character identification, face name matching, graph operation, faces name relationship.

I. INTRODUCTION

In recent years face recognition has become an active topic in field of biometrics. Pervious ally face recognition of still images was very easier but video contain abundant information than a single image. As has been noted by pervious author [1, 2] character identification is difficult due to changes in scale, pose, lighting, expression, hair style etc. In order to retrieve the image of a particular person visual information has to be recognized. The noisy and complicated nature of news videos make the face recognition on vides even more challenging. There are studies saying that the existing system human get a loop of select and relevant faces in result set .However such system usually has many result makes it prone to error and time consuming for human to extract the required name and faces. In this study we have reduce the number of results for user to examine. According to modern film theory" All films are about nothing –nothing but character [5] in which we understand why character are important for movie summarization. The occurrence and interaction of character provides meaningful representation of movie structure and its contents. The challenge is to achieve unobtrusive recognition of face in video.

Character identification, is difficult for computer point of view. The reason can be as follows 1) weakly supervised textual cues [7]. Like reaction shot where the person speaking may not be shown in the frames. 2) Face identification is difficult in video than images [8] e.g. Low resolution, occlusion, non rigid deformations, large motion, large background and other reason make the results of face detection and tracking difficult. In movies, noises are added 3) the character appears quite different during movie 4) the identification of identical faces is not trivial. The occurrence and interaction of character provides meaningful representation of movie structure and its content. As shown in Fig. [1]. The motive is to identify the faces of characters in the video and

label them with names, which can be also used as query to select character of interest and view related video. The goal is to build a real-time face recognition system (FRS) for video environments. A face recognition system is a process to identify faces present in videos automatically. It can operate in two ways given below which is the main process in face

recognition and detection in videos of both kinds of data like available data set of videos or live videos.



Figure.1. Examples of character identification from movie.

Face verification (or authentication): This process involves a one-to-one match that compares a query face image in database against the face identified.

Face identification (or recognition): This process involves identification of face and naming them manually or automatically by graph matching of face and name. Our motivation is to find solution for robust framework for movie character identification.

II. LITERATURE SURVEY

The problem of associating faces with names is to form relations between videos or images and texts in order to label the faces with name.

Name-it is the first proposal on face-name association in news videos based on the co-occurrence between the detected faces and names extracted from the video transcript [2].

in the film script, the spoken lines of different characters appearing in the same scene also represent an interaction. Thus, the names in front/bold of the name can also build a name affinity network with face affinity network.

Both the statistical properties of the faces and the names motivate us to seek a correspondence between the face affinity network and the name affinity network. The name affinity network can be straightforwardly built from the name script given during run time. For the face affinity network, we first detect face tracks in the video and cluster them into groups corresponding to the characters it can be done with non clustering method also just it produces multiply results.

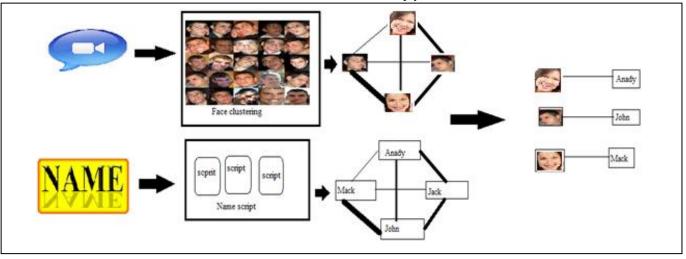


Figure 2.Framework for face name relationship.

Named Faces system built a database of named faces by video optical character recognition (VOCR) [3]. Yang et al. developed closed caption and speech transcript, and built model. Thus which is improved their methods in dynamic captioning by using multiple instance learning for partial labeled faces to reduce to collect data by users. In the speech transcript was also used to find people frequently appearing in the news videos. In news videos, candidate names are available from local matching based while in TV and movies, the names of characters are seldom directly produced in the subtitle or closed caption which containing names and no time stamps to align to the video. The proposed system frame work as shown in Fig. [2] using clustering method.

III. OVERVIEW OF OUR APPROACH

In a film, the interactions among the characters resemble their relationship between each other, which makes a film be treated as a small society.Every character has his/her social position in films and keeps a certain relationship with others. In the video, faces can stand for characters and the co-occurrence of the faces in a scene can represent an interaction between characters hence obtaining there name and showing relationship with name is necessary. Thus, the statistical properties of faces can preserve the mutual relationship in the character network. As the same way During the clustering, the Earth Mover's Distance (EMD) is utilized to measure the face track distance. Since we try to keep as same as possible with the name statistics in the script, we select the speaking face tracks to build the face affinity network, which is based on the co-occurrence of the speaking face tracks.

For name and face association i.e. relationship, it is formulated with matching with vertices using edges. A spectral clustering method, which has been used in registration of faces and names object recognition, is introduced here to build name-face association. Especially, during the matching process, priors can be incorporated for improvement. After assigning names to faces, we also determine the leading characters and find cliques based on the affinity network of face and name using social network provided using films. A platform is presented for character-centred film browsing to the user, which enables users to easily use the name as a query to search related name and face which helps to digest the film content. Compared with the previous work, the contributions of our work include: 1) A graph matching method is introduced to build face-name association. 2) An EMD-based measure of face track distance is presented for face track clustering. 3) Based on character identification, the relationship between characters is mined and a platform is provided for face name browsing. Although the graph matching method and the EMD measure are derived from the existing work, to the best of our knowledge, they have not focused on the face naming problem which we have discussed.

IV. UNSUPERVISED FACE METRIC LEARNING

Face recognition is a visual pattern recognition method. A face is a three-dimensional object subject to varying illumination, pose, expression is to be identified based on its two-dimensional image. A face recognition system generally consists of 4 modules which are-detection, alignment, feature extraction, and matching.

Localization and normalization (face detection and alignment) are processing steps before face recognition (facial feature extraction and matching) is performed.

A. Face Recognition processes:

For face detection we have to segments the face areas from the background. In the case of video, the detected faces may need to be tracked using a face tracking component. Face alignment is aimed at achieving more accurate locating and normalizing faces, whereas face detection provides coarse estimates of the location and scale of each face in the video present. Facial components and facial outline are located; based on the location points using haar features. The face is further normalized with respect to photometrical properties such as illumination below and gray scaling process. Face detection segments means finding the face areas from the background. In the case of video, the detected faces may need to be tracked using a face tracking component like haar feature. Face alignment is aimed at achieving more accurate localization and at normalizing faces, whereas face detection provides coarse estimates of the location and scale of each face in the frame. Facial components and facial outline are located; based on the location 9- pixel points of face. The input face image is normalized in respect to geometrical properties, such as size and pose, using geometrical transforms or morphing of the face is done to extract of the face. The face is further normalized with respect to photometrical properties such as illumination and gray scale which is the process of converting RGB into grey scale. After a face is normalized, feature extraction is performed to provide effective information that is useful for distinguishing between faces of different persons and stable with respect to the geometrical and photometrical variations.

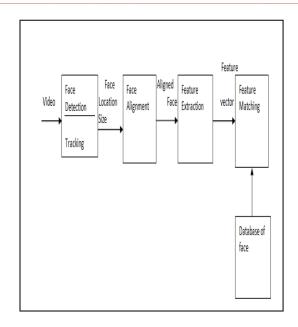


Figure 3. Framework for face name

For face matching, the extracted feature vector of the input face is matched against those of enrolled faces in the database; it outputs then identity of the face when a match is found with sufficient confidence or indicates an unknown face which can be name.

B. Analysis in face subspaces

Subspace analysis techniques for face recognition are based on the fact that a class of patterns of interest, such as the face, resides in a subspace of the input image space. A small image let be any has 64×64 having 4096 pixels total which can be express a large number of pattern classes, consider trees, houses and faces. Among the $256^{4096} > 10^{9864}$ in which all possible "configurations", only a few correspond to faces. Therefore, the original image representation is always less of face, and the dimensionality of this representation of face can reduced greatly by just considering haar feature. With the eigenface method of haar feature or PCA (principal components Analysis) approach, a small number (40 or lower) of eigenfaces are obtained which from a set of training face images .A face image is efficiently represented in grey scale. These features in such subspace provide more salient and richer information for recognition than the raw image. The manifold (i.e. distribution) of all faces accounts for variation in face appearance whereas the non face manifold (distribution) accounts for everything else except the faces in the frame of the video accessed. The videos can be any live videos or data set provided by the user which are then converted into frames of videos.

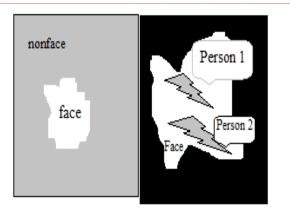


Figure 4. Subspace with face and non face

This manifolds of two individuals in the entire face manifold or non face manifolds. Face detection is a task of distinguishing between the face and non face manifolds in the image (sub window) space and face recognition between those of individuals in the face manifold.

C. Face-name association

After faces are identified with non face in the frame it should be name and relationship in face and name should be obtained. In the video and in the name script given during runtime or compile time, the faces and the names can both stand for the same characters. By treating all the characters as a small society, we can, respectively, build a name affinity network and a face affinity network in their own domains (name script and video). For face-name association, we want to seek a matching between the two networks i.e. face and name. From the social network point of view, our work is to find the structural equivalence actors between the two networks or two domains. Two actors, respectively, from the two networks are defined to be structural equivalence if they have the same profile of relationship to other actors in their own networks.

V. TECHNICAL CHALLENGES

The performance of much state-of-the-art face recognition in videos goes hand in hand with changes like lighting, pose and other factors. The key technical challenges are broadly classified as: Large Variability in Facial Appearance: Whereas shape and reflectance are intrinsic properties of a face object, the face (i.e. texture) is subject to several other factors, including the facial pose, illumination, facial expression. The Highly Complex Nonlinear Manifolds: The entire face manifold (distribution) is highly non convex and so is the face manifold of any individual under various changes. Linear methods such as PCA, independent component analysis (ICA) and linear discriminant analysis (LDA) focused the data linearly from a highdimensional space (e.g. the image space) to a lowdimensional subspace. As such, they are unable to preserve the non convex variations of face manifolds necessary to differentiate among individuals. In a linear subspace, Euclidean distance and Mahalanobis distance do not perform well for classifying between face and non face manifolds and between manifolds of individuals in existing system. This limits the power of the linear methods to achieve highly accurate face detection and recognition.High Dimensionality and Small Sample Size: Another challenge is the ability to generalize as illustrated in figure 4. A canonical face image of 112×92 resides in a 10,304-dimensional feature space. Nevertheless, the number of examples per person (typically fewer than 10) available for learning the manifold is usually much smaller than the dimensionality of the image space; a system trained on so few examples may not generalize well to unseen instances of the face.

VI] ALGORITHMS

A. Clustering

The features extracted from face regions has to group together for easy recognition of face thus we use cluster for similar faces. Ideally, all faces of a particular person is collected in a single cluster.

One of the simplest algorithms for clustering is Kmeans. K-means is a component for the spectral clustering, which is used to cluster not only objects in arbitrary metric spaces, but also Euclidean points lying on nonlinear manifolds and having local cluster structure.

Algorithm (k-means--) Input: Set of points. $X = (x_1, x_2, x_3, \dots, x_n)$ A distance function d: $X * X \rightarrow R$ Numbers k and l Output: A set of k cluster with centres C A set of l outliers L subset of X 1: $C_0 \leftarrow \{k \text{ random points of } X\}$ 2: i ←1 3: while (no convergence achieved) do 4: Compute $d(x | C_{i-1})$, for all $x \in X$ 5: Re-order the points in X such that $d(x_1 \mid C_{i-1}) \geq \cdots \geq d(x_n \mid C_{i-1})$ $6: L_i \leftarrow (x_1 \dots x_l)$ $7{:}\;X_i \hspace{0.1in} \leftarrow \hspace{-0.1in} X \setminus L_i = \{X_{l+1} \hspace{0.1in} , \hspace{-0.1in} ..., \hspace{-0.1in} X_n \hspace{0.1in} \}$ for $(j \in \{1, ..., k\})$ do 9: $P_i \{x \in X_i | c(x | C_{i-1}) = c_{i-1,j}\}$ 10: $c_i \leftarrow \text{mean}(P_i)$

- $10: c_i \leftarrow \text{mean}(P_j)$
- 11: $C_i \leftarrow \{c_{i,1},...,c_i,k\}$ 12: $i \leftarrow i + 1$

B. NON CLUSTERING:

We future extend our work for face name graph matching without clustering specified. The basic idea behind non clustering method is that appearances of the same character vary in the significantly and it is difficult to group them in a unique cluster. High cluster purity with large number of the clusters is expected. Since one character name may correspond to several face clusters, graph partition is introduced before graph matching in non clustering methodology. Therefore, it can provide certain robustness to the intra-class variance, which is very in movies where characters change appearance significantly or go through a long time period. Instead of the separately performing graph partition and graph matching, and using the partitioned face graph as input for graph matching, graph partition and graph matching are optimized in a unique framework .We then first define the graph partition P with respect to the original face graph G^{face}. Consider N character names and M face track clusters, it divides G^{face} into N disjoint sub graph.

 $\mathbf{P} = \{ \mathbf{g}_1^{\text{face}}, \mathbf{g}_2^{\text{face}}, \dots, \mathbf{g}_n^{\text{face}} \}.$

The similarity input s(i, k) is set between face tracks. All the face tracked are equally suitable as exemplars and the preferences s(k, k) are input of median similarities. There are two kind of message, "availability" and "responsibility", changed between face tracks. Were availability means present in database set and responsibility means face has to be track due to upcoming of new face in the videos. With "availability" a(i,k) initialized to be zero, the "responsibilities" r(i,k) are computed and updated as follows initially

$$r(i,k) \leftarrow s(i,k) - \max \left[\hat{k}_{s,t}, \hat{k}_{\neq k} \left\{ a(i,k) + s(i,k) \right\} \right]$$

While, a(i,k) is updated using the rule give below $a(i,k) \leftarrow \min\{0, r(k,k) + \sum_{i,s,t,i \notin i',k} \max 0, r\{i'k\}\}$

C.ERROR CORRECTING GRAPH MATCHING

ECGM is a powerful tool for graph matching with distorted inputs. It has various applications in pattern recognition and computer vision. To measures the similarity of two graphs edits operations, such as the deletion, insertion and substitution of vertexes and edges. Each of these operations is required certain cost. The cost mainly dependent on graph distortions. If less distortion occurs smaller is its cost. Cost function to improve the performance. Let L be a finite alphabet of labels for vertexes and edges.

Notation: A graph is a triple $g = (V, \alpha, \beta)$, where V is the finite set of vertexes, $\alpha : V \rightarrow L$ is vertex labeling function, and $\beta : E \rightarrow L$ is edge labeling function. The set of edges E is implicitly graphs are fully connected, i.e., $E = V \times V$.

Definition 1. Let $g_1 = (V_1, \alpha_1, \beta_1)$ and $g_2 = (V_2, \alpha_2, \beta_2)$ be two graphs. An ECGM from g_1 to g_2 is a bijective function given below

 $F:\widehat{V1} \rightarrow: \widehat{V2}$ where $\widehat{V1} \subseteq \widehat{V1}$ and $\widehat{V2} \subseteq \widehat{V2}$

It that vertex $x \in V_1$ is substituted by vertex $y \in V_2$ if f (x) = y. If $\alpha_1 (x) = \alpha_2$ (f (x)), the substitution is called an identical substitution. The cost is zero, and edit operation is greater than zero. There is vertex substitution edge substitution and edge deletion or creation performed in graph matching using above notations.

Definition 2. The cost of an ECGM f: $\widehat{V1} \rightarrow : \widehat{V2}$ from

Graph $g_1 = (V_1, \alpha_1, \beta_1)$ and $g_2 = (V_2, \alpha_2, \beta_2)$ is given by $Y_1 = (f_1, \alpha_1, \beta_1)$ is $(f_2, \alpha_2, \beta_2) = (V_2, \alpha_2, \beta_2)$

 $\begin{array}{lll} \gamma & (\mathbf{f}, & \mathbf{g}_1, & \mathbf{g}_2) = & \sum_{x \in V1 - V1} \operatorname{Cvd}(x) & + \\ \sum_{x \in V2 - V2} \operatorname{Cvi}(x) & + & \sum_{x \in V1} \operatorname{Cvs}(x) & + \\ \sum_{e \in E1} \operatorname{Ces}(x) & \end{array}$

D.EARTH MOVER DISTANCE:

It's a novel approach of using Earth Mover's Distance for video-based face recognition. General methods can be classified into sequential approach and batch approach. Batch approach is to compute a similarity function between two faces. There are two classical batch methods. This paper considers a most straightforward method of using distance for matching. We propose a metric based on an average Euclidean distance between two videos as the classifier. This metric makes use of Earth Mover's Distance (EMD) as the underlying similarity measurement between two distributions of face images. During matching, the distance between two signatures is computed by EMD. Earth Mover's Distance for Recognition .It is a general metric to compare two distributions that have the same weights in the video In order to define the similarity function f(A;B) between two videos, we introduce the notion of Earth Mover's Distance (EMD).By the definition of EMD is:

$EMD(A;B) = min f \in F EMD(A; f(B))$

where A and B are two distributions of the faces. The purpose of this equation is to seek a transformation f that minimizes EMD(A;B). In this paper, considering its application to video-based face recognition, we define it as:

 $EMD(A;B) = \max g \in GEMD(g(A); g(B))$

where g is a linear transformation to project two distributions onto feature space so as to maximize EMD(A;B). The EMD is a metric to evaluate the dissimilarity between two faces.

It reflects the minimal amount of work that must be performed to transform one distribution into the other by moving "distribution mass" around. The EMD shows the dissimilarity by increasing the amount of transportation work. It is represented as a distribution on certain dominant clusters, which are called signature/faces of videos. The signatures do not necessarily have the same mass, thus it allows for partial matches. Hence, the EMD is adequate to measure face track distance. The face tracks can be represented as follows: Let $P = \{(C_{p1}, W_{p1})..., (C_{pm}, W_{pm})\}$ cluster where be the signature of the first face track with $M(M \le K)$) clusters, where C_{p1} is the cluster centre W_{pl}is the number of faces belonging to this and cluster;

 $Q=\{(C_{q1}, W_{q1})..(C_{qm}, W_{qm})\}$ be the signature of the second face track with $N(N \le K)$ clusters. The EMD between two face tracks is defined as follows:

$$MD(p,q) = \frac{\sum_{i}^{m} \sum_{j=1}^{n} dij fij}{\sum_{i}^{m} \sum_{j=1}^{n} fij}$$

F

Where dij is the ground distance between cluster centres and C_{pi} and C_{qj} . Note that the distance is calculated on the feature vectors of points derived from clustering. The fij is the flow between C_{qj} and C_{qj} . This

way relationship between two face are obtained in a video.

E. HAAR FEATURES

The area of the image with the video has to be analyzed for a facial feature needs to be regionalized to the location with the highest probability of containing the feature. By regionalizing the detection area, false positives areas are eliminated and thus speed up the detection which is increased thus reduces the area examined which is done using haar feature. The human face poses has more problems than other objects. Since the human face is a dynamic object that comes in many forms and colours. However, facial detection and tracking provides many benefits in various applications. Facial recognition is not possible if the face is not isolated from the background. These features, rather than using the intensity values of a pixel, use the change in contrast values between adjacent rectangular groups of pixels. The contrast variances between the pixel groups are used to determine relative light and dark areas. Two or three adjacent groups with a relative contrast variance form a Haar-feature. Haar features can easily be scaled by increasing or decreasing the size of the pixel group being examined as the data set.

This allows features to be used to detect objects of various sizes thus help to detect all types and angles of faces.

• INTEGRAL IMAGES

The integral image is an array containing the sums of the pixels' intensity values located directly to the left of a pixel and directly above the pixel at location (x,y) inclusive. So if A[x,y] is the original image and AI[x,y] is the integral image then the integral image is computed as shown in equation below.

$$AI[x,y] = \sum_{x' \le xy \le y} A(x', y')$$

VI. EXPERIMENT:

Data sets

To evaluate our global character identification approach, the experiments are conducted on three videos "National anthem",:"aarti", "Mr. lonely ".The information of these are shown in table

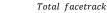
TABLE I

FLIM INFORMATION				
ID	VIDEOS	Length	Character	
1	National	1.51	20	
	Anthem	mins		
2	Aarti	2.6	25	
		mins		
3	Mr.Lonely	4.26	36	
		mins		

A. COST FUNCTION

As the preliminary we start with the face detection. After face detection we cluster them into corresponding to the character. For each cluster the cluster pruning mechanism is the used to refine the results. To demonstrate the result of our method on face track clustering we use precision/recall method. Recall here means the proportion of face track which are assigned a name and precision is the proportion of correctly labeled tracks. The face track classification precision is used as parameter for cost function.

 $\begin{array}{l} \text{Precision} = \frac{Facetracks \ correctly \ classified}{facetracks \ correctly \ classified} \\ \text{Recall} = \frac{Facetracks \ correctly \ classified}{m_{\text{rec}}} \end{array}$



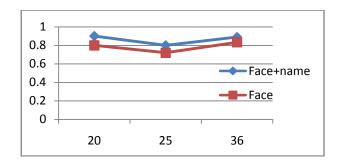


Figure.5 Recall of face track classification.

The result are demonstrated in fig.5.As expected, the performance of global matching features (face + name) is better than single feature(face) on the marginal points.

B. FACE TRACK DETECTION

We utilized a multi-view face tracker to detect and track faces on each frame of the video.

The statistics of the performance are shown in Table II, where "#Face track" and "#Track detected" are the numbers on face track and track detected.

TABLE II

FACE TRACK DETECTION ACCURACY

Videos	#Face	#Track	Accuracy
Id	track	detected	
1	20	18	90.2%
2	25	22	88.1%
3	26	23	89.6%

C.FACE TRACK CLUSTERING

The clustering process is actually divided into two parts face track clustering by face recognised and script.

PURITY: Cluster purity is used to evaluate the performance of face clustering.

$$purity = \frac{1}{N} \sum_{i=1}^{N} \frac{\max_{j \in C_i} \bigcap Name_{j \in I}}{|C_i|}$$

Where C_i is the set of face tracks in the ith cluster and Name j^{th} is the set of face tracks with the j^{th} label (character name).Since high purity is easy to achieve when the number of clusters for large dataset. It mainly uses global matching by exploring a diverse set of technologies including face detection and recognition by lip motion analysis, visual saliency analysis. Visual saliency analyses are mainly used in global captioning method. Names for the clusters are then manually selected from the global list name script location determine the region in which name scripts will be presented. Recently, metric learning is introduced into character identification in uncontrolled videos. Castspecific metrics are adapted to the people appearing in a particular video in an unsupervised manner. It is easy for understanding and implementation.

VII. CONCLUSIONS:

In this paper, we have proposed a novel framework for character identification in feature-length films. Different from the previous work on naming faces in the videos, most of which relied on local matching, method we provided global matching. As the amount of the stage implemented at this stage we got 50-70% accurate result in both face recognition and detection due to name script provided and relationship between face and name.

The proposed methodologies are useful to implement for clustering and identification of face tracked from uncontrolled movie videos in both live videos and data set.

VIII. FUTURE CONTRIBUTION:

In future we expand our work to investigate sequential statistics for speaker in real time videos application for more accurate result.

REFERENCES

[1] J. Sang, C. Liang, C. Xu, and J. Cheng, "Robust movie character identification and the sensitivity analysis," in ICME, 2011 ,pp 1-6. [2] Y. Zhang, C. Xu, H. Lu, and Y. Huang, "Character identification in feature-length films using global face-name matching," IEEE Trans.Multimedia, vol. 11, no. 7, pp. 1276-1288,November,2009.

[3] M. Everingham, J. Sivic, and A. Zissserman, "Taking the bite out of automated naming of characters in tv video," in Jounal of Image and Vision Computing, 2009, pp. 545-559.

[4] C. Liang, C. Xu, J. Cheng, and H. Lu, "Tvparser: An automatic tvvideoparsing method," in CVPR, 2011, pp. 3377-3384.

[5] J. Sang and C. Xu, "Character-based movie summarization," in ACMMM, 2010.

[6] R. Hong, M. Wang, M. Xu, S. Yan, and T.-S. Chua, "Dynamiccaptioning: video accessibility enhancement for hearing impairment,"in ACM Multimedia, 2010, pp. 421-430.

[7] T. Cour, B. Sapp, C. Jordan, and B. Taskar, "Learning from ambiguouslylabeled images," in CVPR, 2009, pp. 919-926.

IJRITCC | May 2014, Available @ http://www.ijritcc.org

H. K. [8] J. Stallkamp, Ekenel. and R. Stiefelhagen, "Video-based facerecognition on real-ICCV, world data." in 2007, 1-8. pp. [9] S. Satoh and T. Kanade, "Name-it: Association of face and name invideo," in Proceedings of CVPR, 1997, pp. 368-373. [10] T. L. Berg, A. C. Berg, J. Edwards, M. Maire, R. White, Y. W. Teh, E. G.Learned-Miller, and D. A.

White, Y. W. Teh, E. G.Learned-Miller, and D. A. Forsyth, "Names and faces in the news," inCVPR, 2004, pp. 848-854.

[11]J. Yang and A. Hauptmann, "Multiple instance learning for labelling faces in broadcasting news video," in ACM Int. Conf. Multimedia, 2005,pp. 31-40. [12]A. W. Fitzgibbon and A. Zisserman, "On affine invariant clustering andautomatic cast listing in movies," in ECCV (3), 2002, pp. 304-320.

AUTHOR:



NiveditaPandey was born in Mumbai, Maharashtra, India in year 1990. She is currently pursuing Master Degree in Computer Engineering from MET Bhujbal Knowledge Nasik. City, She has completed her degree in Computer Engineering from Sir Visvesvaraya Institute of Technology, Nasik.



R. P. Dahake She is currently Assistant working as Professor in Department of Computer Engineering, MET's Bhujbal IOE Knowledge City, Nasik, Maharashtra, India. She has completed her Post Graduation in Computer Engineering from Govt. College of Engineering Aurangabad Maharashtra.She has presented papers at National and International conferences and also published papers in National and International Journals on various aspects of Computer Engineering and Networks. Her areas of interest include Computer Networks Security and Embedded Systems.