Available at: www.researchpublications.org

Analytical Evaluation of Multi-clue Face Retrieval Approach

Mr. Zafar G. Sheikh, Dr. V. M. Thakare, Dr. S. S. Sherekar
P.G. Department of Computer Science and Engineering
SGB Amravati University
Amravati (M.S.), India
zgsheikh@gmail.com, vilthakare@yahoo.co.in, ss_sherekar@rediffmail.com

***Abstract*- The study, analysis and investigation of recent development would leads to acquire objective of future. The proposed work is inspired from the same issue in concern face retrieval approaches based on multi clue. It would be the future demand in various applications for searching, browsing, and retrieving human face of interest from video database or in real time.**

The goal of proposed work is to systematically address the recent work of face retrieval approach based on multi-clue through evaluation that permits a meaningful objective comparison of techniques, provides the research community with sufficient data for the exploration of automatic modeling techniques. The outcome of the paper which fulfilled the three objectives-i) study and classification of recent techniques ii) analyzed the status recent techniques with respect to results and performance Finally, iii) identification of most feasible and optimized technique along with discussion for betterment. Mean while, objective evaluation and analytical study would be extremely useful to the computer vision research community for years to come.

Keywords: face retrieval; Multi-clue; detection; tracking, feature extraction.

I. INTRODUCTION

The proposed work is inspired from the issue to search and retrieve human face of interest from video. It would be the future demand for searching, browsing, and retrieving human face of interest from video database for several applications such as state-of art, security, and surveillance, personal and industrial demands. Human face retrieval is the collective work of major aspects such as detection, tracking, recognition. It is also called the process. The researcher has been successful demonstrated the implementation of detection and recognition. Still, the real life problems are yet to solve.

The proposed work has focuses on video database, as per our analysis, it has not acquire much attention and success as compare to detection and recognition because of two reasons-i) necessity is the mother of invention, i. e. researcher still not consider an important issue. ii) It is mutual facility which includes various approaches to satisfy the aim and objective which changes with time. The paper aims to analyze and study the available human face retrieval methods of multi-clue approaches for the betterment of specified application. The various media files play an important role in everyday life such as feature-length, news video, state-of-art application, movie or any other video.

Objective evaluation would be extremely useful to the computer vision research community for years to come. The outcome of the paper which fulfilled the three objectives-i) Analytical study and classification of recent techniques ii) analyzed the status recent techniques with respect to results and performance Finally, iii) Analyzed and identify the most feasible and optimized approach along with discussion for betterment.

Videos have contained information of characters or persons such as like gait, speech, motion and face. Face is unique identification of browsing and searching. The face is complex object and it turns more complex problem for video. The face detection techniques on image have been getting mature results [1] as an evidence of more than two decade of research. The recent advancement in face detection technique is briefly analyzed and described [2] on video databases and real time applications. Face retrieval is an emerging area of research with various applications. It is not innovative difficulty defined; actually it derived and combined existing feasible technique to solve real life problems.

The proposed work is inspired from the same issue and work presented by R. Kasturi *et al*.,[3], a framework for evaluating object detection and tracking in video: specifically for face, text, and vehicle objects. This framework includes the source video data, ground-truth annotations (along with guidelines for annotation), performance metrics, evaluation protocols, and tools including scoring software and baseline algorithms. The goal of proposed work is to systematically address the recent work of face retrieval in video through evaluation that permits a meaningful objective comparison of techniques, provides the research community with sufficient data for the exploration of automatic modeling techniques.

The paper is organized as follows- Section II described classification of face retrieval multi-clue approach. Section III contained important tabular analysis of recent approaches with various steps involved called significant analysis, performance evaluation or analytical evaluation and discussion has discussed in Section IV. The paper concludes with conclusion and future work in Section V.

II. STUDY AND CLASSIFICATION

This section contained study of recent techniques related with human face retrieval based on multi-clue approach specifically in video along with various steps involved in available techniques. Fourteen recent multi-clue techniques are discuss and analytical evaluated are presented as below-

1. *Character Identification in feature length films using global face –name matching*

The presented approach [4] for the identification of character in feature length films using global face and name matching. Face detected in video and clustered into groups of characters using Earth mover's Distance (EMP) for measure the distance of face tracks. Name affinity built from script. Speaking face tracks to build the face affinity network. Name and face association using matching vertices between two graphs. Multi-view face tracker to detect and track faces on each frame of the video, multi-view face exemplars in a track and finally cluster into groups corresponding to characters and built the face affinity network. Speaking face track detected using Region-of-interest and SIFT points are extracted and matched between current face image and the previous image. Face represented using Locally Linear Embedding (LLE) for dimensionality reduction. The distance measured between face track using spectral clustering and K dominant clusters from all the detected faces. The minimum distance makes them be treated as the same person due to the partial similarities. K-means clustering is performed to group the scattered face tracks which belong to the same character. The noise clear from clustering results with pruning method the marginal points which have low confidence belonging to the current cluster. The face-name association performed using affinity network in their own domain.

2. Speaker retrieval for TV show video

The retrieval of speaker in TV show programs [5] has presented using their names as the query. The framework divided into two parts, visual processing part that focuses on shot segmentation and clustering of the same person with different occurrences together and locate their faces spatially and identify association part which focuses on solving the ambiguities between present faces and who and when speaks information. For the face detection, multi-view frontal face detector implemented in OpenCv by Viola and Jones used on every frame. The face is tracked by mean shift color tracker using color histogram or kernel density estimate of model and target image.

3. Multimodal Person Search and Retrieval

Towards person Google: multimodal person search and retrieval [6] is a new search and retrieval approach. Content based on multimedia retrieval systems have been automatically indexing and retrieval multimodal person from video. The system work separately for audio and video segmentation, feature extraction and segment matching. The feature exaction from audio using mel frequency cepstral coefficient (MFCC). The temporal characteristics of audio data within segment reduce using multivariate Gaussian distribution. The audio segments are compared by applying Bayesian information Criterion (BIC) on computed the distance between two segments. For visual analysis face region determined based pupil position and anthropometric model. To handle the illumination changes, statistical normalization methods are applied globally and locally. The extracted features were reduces using PCA and the features were matched with database using Euclidean distance. The multimodal fusion consists of audio and visual score. Score has normalized with different characteristics at last the score fusion performed. It is fusion of face detection and speaker segmentation for person retrieval.

4. face indexing system for actor –based video service in an IPTV environment

Jae Young choi et al. [7] proposed system for automatic system for indexing faces of actors. It consists of two parts, first to construction of FR engine using web and second is video face recognition. Using internet connect of STB, actor names are retrieved from online drama and movie information provider. For the second part, faces are clustered using colour histogram for computing each video frame for grouping the same subject into single subject cluster and face images of different subjects included in different clusters. The Hierarchical agglomerative clustering (HAC) has used for clustering, which terminates on threshold. The extraction of colour face features performed by the query face images (RGB) converted into different colour space like YIS or HSV images and extract features from these colour component vector creations. The extracted features from different colour components are concatenated. The face recognition has performed using weighted colour feature fusion, which dealing with several defective face images in a cluster which may contain variation in terms of viewpoint, illumination and compression artifacts. The elements penalty based Minowski, includes standard deviation from sample of feature vectors.

5. Estimate discriminant coordinates

In [8], authors proposed the face under pose variation of expression and illumination detection by Kernel PCA (kPCA) and discrimination for set and the rectification of canonical pose performed by running. Each SVM over image and performed affine transformation which best map detected points tp canonical features. kPCA performed to reduce dimension and LDA to project data into space that for the discrimination task. The Nystram Approximation performs on kPCA dataset because its too large opt as kernel matrix. Finally the clustering is used to clear up the semi-super used database with error and new discrimination provide better representation of identification and re-cluster it in a modified K-means clustering.

6. Audio and visual information

Zhu Liu and Yao Wang [9] have investigated major cast detection in video by using both speaker and face information. The approach involves three steps. First step is speaker boundaries detection and clear speech extraction using GMM classifier. It has been covered the comparison of GMM classifier and SVM classifier. Simulated results are generated by GMM is better as compare to SVM. Then speaker segmentation and clustering using GMM distance metric with divergence. The second step to face detection in still image using fast template matching which detect multiple faces from still image. The face tracking and clustering involve. Video shot segmentation with distance of color histogram and face tracking within each shot using

face template. Third and last step, integrated speaker and face correlation matrix i.e. speaker face –correlation matrix.

7. Local features and Statistical-Structural learning

In [10], an integration of statistical and structural information that uses the local feature constructed from coefficient of quantized block transforms which is use in video compression. Under quantization and performing statistical histogram of the local features treated as vectors and similarity measure. The image is decomposed into subarea called as local feature using quantized block transforms. Quantized coefficients of block transform are used for construction of local features and description called as feature vector. Ternary feature vector (TFV) structure from the collection of same order transform coefficient boring transformation blocks. Statistical information compared using TFV histogram based on 0th and 4th transform coefficient which represent different types of information about local feature. Structural description of pattern is represented by sub area histograms. Three aspects for pattern retrieval 1st is the set of local features which is robust from perceptual point of view is not selected arbitrarily but by adjusting the quantization level of block transform. 2nd size of selected feature histogram and last is scope of structural information.

8. Video Shot Retrieval for Face Sets

The work suggested by J Sivic et al., [11] for frontal face detection has been performed on every frame of the movie. This is achieved by first running a general purpose region tracker called affine covariant region tracker by J. Sivic et al this tracking algorithm can develop tracks on deforming objects. Resultant, person's face can be tracked with significant pose variation and expression changes, but tracking is done offline. A single shot contains hundreds of frames with possibly one or more detected faces and the detected faces generally connected by several region tracks. A single-link agglomerative grouping strategy is used to merges face detection into larger group starting from most closest(most connected).

Each face sets (face-track) are described by a collection of five affine transformed local spatial orientation fields based around facial features. The entire set represented as a single distribution over local feature descriptors. The facial features(left and right eyes, tip of nose and center of the mouth) localized allows to local face descriptors and affinely deform support regions to normalize for pose variation. A probabilistic part-based "constellation" model of faces is used to modal the joint position(shape) and appearance of the facial features. Facial features are located by searching for joint position of the features which maximizes the posterior probability of the feature positions and appearance. Each face in the set is represented as a collection of five local overlapping parts(SIFT descriptors)placed at the detected feature locations(eyes, mouth, nose and mid point between eyes)[11].

9. Film Character Retrieval In Feature-Length Films

Ognjen Arandjelovie et al. [12] has proposed approach to recognize all the frontal faces of a character in the movies or situation comedy . given a small number of query faces. The recognition method based on cascade of processing steps that normalized the effect of changing imaging environment, particularly in three areas suppress the background of face, pose refinement to optimize the registration and used robust distance to a sub-space to allow for partial occlusion and expression changes.

Ognjen Arandjelovie et al. [12] considered content-based multimedia retrieval setup to retrieval and rank by confidence from film shots. A query consists of user choosing the person of interest in one or more key-frame. A face detection stage has performed by local implementation of 3D object detection based on correctly detection of both eyes and mouth, are visible. The proposed approach consists of computing the low distance in numeric value, a distance, expressing the degree of belief that two images belongs to same person and computing a series of transformation of the original image.

10. Scalable Face Image Retrieval

In [13], authors aims to build scalable face image retrieval system by using both local and global features for face representation. The first step is locate component-based local features that encode geometrical constraint and robust to pose and expression variation. The special properties of faces has exploit to design new component-based local features in indexing stage, which are subsequently quantized into visual words using identify-based quantization scheme. By using a very small Hamming signature (40 bytes) encode performed on the discriminative global features for each face. In the retrieval stage, candidate images are first retrieved from the inverted index of visual words and than new multireference distance to rerank the candidate images using the Hamming signature.

11. Face Verification and Image Search

Neeraj Kumar *et al.,* [14] proposed face verification and image search using describable visual attributes. The face images were collected from the varieties of online sources used for collecting face images from search engines like Yahoo Images, flicker. OKAO face detector has been applied to downloaded images to extract faces. The attributes were collected and labels were identified by using Amazon Mechanical Turk service. Learning an attribute or simile classifier consists of fitting a function to set of labeled training data. The face images are first aligned using affine transformation and mask out the background prior to feature extraction to avoid contaminating the classifier. Each region extracted has different information on feature type such as pixel value types(RGB, HSV, intensity, edge Magnitude and edge Orientation), normalization(mean , energy) and aggregation(histogram, mean/variance). All types of low level features extracted from whole face. The SVM classifier with RBF Kernels, trained using libsvm performed grid search. The simile classifiers are used to recognized similarity to part of a reference person's face in many images. The trained SVM used to distinguish a region (e.g., eyebrows, eyes, etc) on their face from the same region on other faces. Face verification performed by training a verification classifier by SVM with an RBF kernel. To

search results are ranked by confidence by computing distance to classifier decision boundary as a measure of confidence. Combined confidence and final ranking on relevance used for multiple queries. The confidence converted into probabilities by fitting Gaussian distributions on attribute score. Images with high confidences for all attributes are shown first.

12. Face Image Retrieval Using Genetic Algorithm

In [15], authors proposed the method using genetic algorithm and bag of pixel for face image retrieval. The Maximum of likelihood Gaussian covariance estimator gives rise to an estimate of permutation matrixes that aligned images into minimally low-dimensional subspace which is preprocessing for PCA. The 9 color features are selected in 2 RGB and HSV environments. The 18 features for any image studied into two different color environments. The similarity comparison for input images with image database using Euclidean distance and for changed images with respect to rotation and enlargement is SPIRMAN cohesion coefficient. The genetic algorithm is based on fitness function to reduce number of features and choose more efficient features.

13. Interactive person re-identification

The system which is analyses video using shot boundary detection and face tracking component has presented [16], and then extracts features which is used for retrieval. The four main features of system, track the person successfully using face tracker by M. Fischer across pose changes. Second, robust face recognition algorithm based on local features appearance to match the tracks. In third, retrieval process, a query built by collecting all face images of the track and user selected by clicking in face images. System searches very close matches to any images in the query sets, new search performed using enlarged query set. The features are extracted using local appearance from cropped face image to 64 X 64 pixels without alignment step. Discrete cosine transform (DCT) is applied to each blocks and the DCT coefficients are ordered using zig-zag scanning. The order coefficients are used to build local feature vector, which is normalized to unit norm in order to reduce the illumination effect. Person re-identification has been done with interactive feedback, system automatically searching for track with a distance to one of the tracks in the query set that is lower than some threshold. Track distance is the minimal L1-distance of any two feature vector extracted from the tracks.

14. Face Image Retrieval Based on Vertical Web Image Retrieval

Ran Zheng *et al.,*[17] has demonstrated vertical web retrieval framework using combined semantic and visual features. It divided into three parts: data resources, algorithm resources and processing layer. Data resource stored databases including raw website database, web images image feature, semantic database and knowledge database. The algorithm provides processing such as design of crawler, filtering, word segmentation, feature vector extraction and others. The processing layer includes work such as data gathering, exclusion of image, index construction and other. The special techniques used such as data gathering of special field, word segmentation in special area, automatic filtering of non-specific image.

III. PERFORMANCE EVALUATION AND DISCUSSION:

The analysis of the different approaches concerned in table1, the approaches have different objectives to deal in real life problems. The major concern of the analysis is to evaluate the performance. For the evaluation as a part, consider the objective that the human face retrieval suing multi-clue features.

The recent available techniques has been studied and classified according to types of approach has been used. There are approaches like Feature invariant, template and appearance based. The features invariant based approach based on structural features; however template based approach based on defined or trained template. In contrast to template matching, the models (or templates) are learned from a set of training images which capture the representative variability of facial appearance. These learned models are then used for detection called as appearance based methods [2].

The techno-critical analysis (Table 1) presents the involvement of different major and minor aspects of human face retrieval from video. The detection involved the location of face image if present. Whereas, tracking approach, tracks the face image (template) or points in video sequences. The different features of face images were extraction and represented after normalization. The face image has classified from the background or other effects. However the matching and recognition are the final parameter for the identification. Since sometime training images need to make a final face image representation. According to the user requirements, the different parameter involves to make a complete face retrieval process successful.

We investigated that multi clue based approaches leads to more accurate results as compared to others classified techniques, however it requires more processing speed. It may leads to fail to real life situation. The chronological amalgamation of holistic and part-based approach may leads to good results with best performance.

The Multi-clue approach has the highest number of approaches because of wide varieties of combination. In future, it can be identified sub categorized according to (table 1) or by considering other widespread and significant specification. It is find out that, the 11 different performances out of 14 available approaches for mean precision rate. The three approaches did not have precision rate, so the average precision rate in accordance with the 11 available approaches. The highest precision rate is 93% and lowest is 85.99% and the average precision rate 89.59%.

IV. CONCLUSION AND FUTURE WORK

The study and analysis would help to identify problems in current techniques and improved by eliminating it. The proposed work provides a review and analysis of latest different available approaches for human face retrieval with multi-clue. The study leads towards development of improved approach or method which has reduced the existing limitation of system. With the help of above study, the decision has been made for making the framework which provides feasible solution.

The work has been extended for the various objectives such as searching, browsing and indexing. Meaningful objective comparison of techniques would provide the research community with sufficient data for the exploration of automatic modeling techniques. We explored the study and analysis for any other available and latest relevant approaches with suitable classification. This work would be extended with using scientific tools and data for evaluation and performance analysis. The useful investigation has sum up with better results in terms of accuracy and speed under the assumption that human face retrieval from video databases.

TABLE I. SIGNIFICANT ANALYSIS OF FACE RETRIEVAL(MULTI-CLUE) TECHNIQUES

Approach of implementation	Steps used in Approach									
	Face detection	**Tracking**	**Feature extraction**	**Normalization**	**Matching**	**Representation**	**Classification**	**Training data**	**Recognition**	**Other techniques**
1. Affinity network based global face-name matching[4]	Earth mover's Distance(EMP)	Multi-view face tracker	spectral clustering and K dominant clusters	K-means clustering	Matching face tracks with (Euclidean distance)	Locally Linear Embedding (LLE)	K-means clustering between face track	Did not used training data	Gaussian Mixture Models(GMM)	Faces grouped in each shot
2.IDU system[5]	Viola and Jones detector enhancement by Lienhart	Kernel-based tracker using radial basis function(RBF)	13 facial features	Affine transformation	Set of posteriors (Max-max or max-sum)	13 facial features	Random-ferns classifier	Cascade face detector	Pictorial structure model(9)+4 additional facial features	40 ferns of 17 levels
3. Multimodal fusion system[6]	pupil positions and anthropometric face model	Did not used tracker	PCA	z-score	Euclidean distance	PCA	pupil position and anthropometric model	multivariate Gaussian distribution	Score level fusion(product, sum, min, max)	score fusion performed
4. Weighted feature fusion scheme[7]	Face clustering : Hierarchical agglomerative clustering (HAC)	faces are clustered using colour histogram	Color component (YIQ) face features	z-score	Nearest Neighbor classifier	Color component vector creation	Hierarchical agglomerative clustering (HAC)	Did not used training data	Weighted color features fusion	Shot segmentation by color histogram
5. Estimate discriminant coordinates [8]	Nystrom Approximation	Did not used tracker	kPCA and LDA	Cholesky Decomposition	Low likelihood with threshold	Cluster	Modified K-means clustering	Did not used training data	Nearest Neighbor	Merging clusters
6. Audio and visual information [9]	Face template matching	Skin color distribution	Skin tone color	Generating Average face template	Average face template matching(maximum Matching Value)	Face template	Face track clustering	Average face template	χ^2 Distance in color histogram	Shot segmentation by color histogram
7. local features and Statistical-Structural learning [10]	Searching best subarea	Did not used tracker	discrete cosine transform (DCT) block transforms	lower-order AC coefficients	Histogram matching	Histogram of Ternary feature vector (TFV)	TFV Histogram using single subarea	Training images	Histogram matching	Statistical-Structural learning Approach
8.Spatial–temporal system[11]	Propagation probability	affine covariant region tracker	Principal Component Analysis	local face descriptors and affinely deform support regions	Histograms compared using χ^2	overlapping five SIFT descriptors	histogram	Principal Component Analysis	probability density function	Face exemplars (Region tracking)
9. Cascade approach[12]	SVM based detector	Did not used tracker	SVM	Background removal, band-pass filtering, rank matching	Threshold	Face registration using facial features	Background removal	SVM	Comparing face signature image	illumination is normalized by using band-pass filtering
10. Local and global	Viola and Jones Detector	Did not used tracker	total 175 grid points	similarity transfor	histogram based T3hs2	. Local and global	Did not used	Did not used training	nearest-neighbor search using	Hamming signature, Rerank

features[13]	and facial features using NN based componen t approach		from five compone nts	m	descriptors	features		data	k-trees	
11. Describable Visual Attributes[14]	OKAO face detector	Did not used tracker	All types of low level features	mean , energy	SVM with an RBF kernel	histogram, mean/varia nce	mask out the background	Libsvm	measure of confidence	fitting Gaussian distributions on attribute score
12. Genetic Algorithm and Bags of Pixels[15]	Consider face database	Did not used tracker	9 color features	PCA	Euclidean distance	18 features selected in 2 RGB and HSV	Did not used	Did not used	genetic algorithm	SPIRMAN cohesion coefficient
13. Interactive person re-identification [16]	shot boundary detection	M. Fischer across pose changes	local feature vector	order coefficie nts	match the tracks	cropped face image to 64 X 64 pixels	Did not used	Did not used	minimal L1-distance	re-identification
14. combined semantic and visual features[17]	Viola and Jones AdaBoost	Did not used tracker	LBP operator and color histogra m intersecti on	Gabor wavelet transfor m	LGBPHS for similarity match	256*256	Did not used	Did not used	LGBPHS for similarity match	each image divided into 4*4 blocks

REFERENCES

[1] M. Yang, D. J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey*", IEEE trans. pattern analysis and machine intelligence*, Vol. 24, No. 1, pp-36-58, Jan 2002.

[2] Zafar G. Sheikh, Dr. V. M. Thakare, Dr. S. S. Sherekar, "Advances in face detection techniques in video" , *2nd IEEE International Conference parallel and distributed computing (PDGC-12)* on, pp-278 - 283 , Dec. 2012.

[3] R. Kasturi, D. Goldgof, P. Soundararajan, Manohar, J. Garofolo, R. Bowers, M. Boonstra, V. Korzhova, and J. Zhang, "Framework for Performance Evaluation of Face, Text, and Vehicle Detection and Tracking in Video: Data, Metrics, and Protocol" , *IEEE trans. pattern analysis and machine intelligence*, Vol. 31, No. 2, pp-319-336 FEB 2009.

[4] Yi-Fan Zhang, Changsheng Xu et al, "Character Identification in feature length films using global face –name matching". *IEEE transaction on multimedia*, Vol. 11, No. 7, pp-1276-1288, Nov 2009.

[5] Yina Han et al. , "Speaker retrieval for TV show video by associating audio speaker recognition result to visual faces", Proceedings of 2nd International conference the K-Space , 2008.

[6] Lutz Goldmann , Amjad Samour and Thomas Sikora "Towards Person Google: Multimodal Person Search and Retrieval". *In Proceedings of International conference SAMT*, pp.292-295 , 2007.

[7] Jae Young Choi, Wesley De Neve and Young Man Ro, "Towards an automatic face indexing system for actor –based video service in an IPTV environment". *IEEE Transaction on Consumer Electronics* , Vol. 56, Issue 1, pp-147-155, Feb 2010.

[8] Tamara L. Berg et al , "Names and faces in the news" *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (CVPR'04) - Volume 2, pp-848-854, July 2004, ISSN: 1063-6919 , Print ISBN: 0-7695-2158-4.

[9] Zhu Liu and Yao Wang, "Major cast detection in video using both speaker and face information" . *IEEE transaction on Multimedia* , Vol. 9, No. 1, pp-89-101, Jan 2007.

[10] Daidi Zhong and Irek Defee , " Face retrieval based on robust local features and statistical structure learning approach" . *EURASIP Journal on Advances in Signal Processing* , Volume 12, pp- 2008.

[11] J. Sivic, M. Everingham and A. Zisserman, "Person Spotting: Video Shot Retrieval for Face Sets". *Proceeding of ACM International conference on image and video retrieval*, (CIVR'05), pp.226-236, 2005.

[12] Ognjen Arandjelovie et al. , "Automatic face recognition for film character retrieval in feature-length films*". In proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp-860-867, Jun 2005.

[13] Zhong Wu, Qifa Ke, Jian Sun, and Heung-Yeung Shum. "Scalable Face Image Retrieval with Identity-Based Quantization and Multireference Reranking", *IEEE Trans. on pattern analysis and machine intelligence*, Vol. 33, No. 10, pp-1991- 2001, Oct 2011.

[14] Neeraj Kumar, Alexander C. Berg, Peter N. Belhumeur, and Shree K. Nayar, "Describable Visual Attributes for Face Verification and Image Search". *IEEE Trans. on pattern analysis and machine intelligence*, Vol. 33, No. 10, pp-1962-1977, Oct 2011.

[15] Mohammad Ali Azimi kashani and Mohammad reza ramezanpour, "Face Image Retrieval Using Genetic Algorithm and Bags of Pixels". *IEEE Computer Society Eighth International Conference on Information Technology: New Generations*, pp-490-493, 2011.

[16] Mika Fischer Hazım, Kemal Ekenel and Rainer Stiefelhagen, "Interactive person re-identification in TV series". *IEEE International Workshop on Content-Based Multimedia Indexing* (CBMI),pp-1-6, July 2010.

[17] Ran Zheng, Shilei Wen, Qin Zhang, Hai Jin, Xia Xie , "Compound Face Image Retrieval Based on Vertical Web Image Retrieval". IEEE Computer Society Sixth Annual International Conference on Grid computing , pp-130-135, Oct 2010.