ISSN: 2321-8169
5287 - 5296

Large- Scale Content Based Face Image Retrieval using Attribute Enhanced Sparse Codewords.

Chaitra R,
Mtech Digital Coomunication Engineering
Acharya Institute Of Technology
Bangalore

Abstract:- Content based image retrieval (CBIR) have turn into majority dynamic exploration regions within previous couple of existence. Numerous index strategies be in light of worldwide component circulations. Be that as it may, these worldwide circulations have restricted segregating force since they are not able to catch nearby picture data. Photographs with individuals are the foremost attention of users. Consequently with exponentially increasing pictures, huge size contented base features representation recovery is a facilitating knowledge in favor of various developing applications. The main objective is to apply automatically spotted human characteristics that comprise semantic cue of facade pictures toward increase gratified base facade recovery through creating semantic codeword pro effectual huge size countenance recovery. With leveraging person characteristics into scalable as well as methodical structure, suggest and offer two orthogonal systems named attribute improved meager code and attribute entrenched upturned index toward develop facade recovery. We compare proposed method with other three methods namely LBP, ATTR and SC methods. The results illustrate that the proposed methods can attain qualified enhancement in Mean Average Precision (MAP) associated to the existing methods.

***Keywords**: Content based image retrieval, average precision, human attributes, sparse coding, embedded inverted index and low-level features.*

*****

I. INTRODUCTION

In our regular life social communities like facebook, flickr, whatsapp and so on has been given more significance which helps us to impart photographs or features to our dear ones. Popularity of computerized gadgets used to take pictures is likewise expanding in current life. Among these photographs, a major rate of them is with human faces. The significance and the sheer measure of human face photographs make controls (e.g., hunt and mining) of large scale human face pictures a truly critical examination issue and empower numerous certifiable applications CBIR can significantly improve the precision of the data being returned and is an imperative option and supplement to customary content based picture seeking. For depicting picture substance, shading, composition and shape components have been utilized. Shading is a standout amongst the most widely utilized low-level visual components and is invariant to picture size and introduction. One of the vital and testing issues of these pictures is vast scale substance based face picture recovery. Given a question face picture, substance based face picture recovery tries to discover comparative face pictures from a substantial picture database. It is an empowering innovation for some applications counting programmed face annotation, wrongdoing examination, and so forth.

In this paper, the test of extensive scale face picture recovery is getting tended to. The test like vast scale face picture recovery is found in recovery strategy as substance based face picture recovery. In conventional face recovery strategies low level elements are utilized to speak to confronts [1, 2]. However, the downside of low level elements is that they absence of semantic implications though face pictures have fluctuation in appearances, posturing, and so on. Face pictures of diverse individuals may match as indicated by low level components. This may create uncertain result. By consolidating abnormal state human traits and low level elements; better recovery results are accomplished alongside better representation of highlight.

Qualities which are removed naturally from human face picture are the wellspring of abnormal state highlight portrayal around a specific human being [3]. Programmed human quality recognition helped numerous late scientists to get proficient results in regions like check of face [3], ID of face [5], essential word based face picture recovery [6] and comparative trait seek. In spite of the fact that human properties are extremely helpful for operations identified with face picture yet it is not the situation with substance based face picture recovery undertaking. There are two noteworthy purposes for this and the first reason is that

 ISSN: 2321-8169
5287 - 5296

vector of drifting focuses won't function admirably with vast scale indexing techniques. Vectors of skimming focuses speak to human traits. This outcomes in low reaction and versatility issues as it is not totally supporting enormous measure of information. Second reason is that the constrained measurements of human facial attributes. Despite the fact that the dataset comprise of loads of pictures, however it may lose a few individuals which may appear to be comparative due to closeness in a few traits.

In this work, we give another viewpoint on substance based face picture recovery by fusing abnormal state human qualities into face picture representation and file structure. By consolidating low-level components with abnormal state human qualities, we have the capacity to discover better element representations and accomplish better recovery results.

To whole up, the contribution of this paper include:

- We consolidate naturally distinguished abnormal state human properties and low-level components to build semantic codewords.
- To offset worldwide representations in picture accumulations and by regional standards implanted facial qualities, we propose two orthogonal techniques to use consequently recognized human ascribes to enhance substance based face picture recovery under an adaptable structure.
- We further recognize enlightening and non specific human traits for face picture recovery crosswise over diverse database.

The selected descriptors are promising for other applications (e.g., face verification) as well. The rest of the paper is organized as follows. Section II discusses the literature survey. Section III describes our implementation details which introduce the proposed methods including attribute-enhanced sparse coding and attribute embedded inverted indexing. Section IV gives the experimental results, and section V concludes this paper.

II. LITERATURE SURVEY

Yu-Heng et al proposed an efficient way for organizing and searching consumer photos effectively by positioning attributed faces at desired position and in desired sizes on a query canvas. First alignment of the detected faces into canonical position is done and then to form feature vectors, component based local binary patterns are extracted from the images. Further based on learned dictionary, sparse representation is computed from the vector features. This provided rapid online photo search by detecting the facial attributes automatically and by measuring the face similarity in the offline stage [1].

Dayong Wang et al provided a promising framework to the retrieval-based face annotation problem by mining massive weakly labeled facial images available on WWW. Using a web search engine they collected weakly labeled facial image(noisy images). Face detection and alignment are performed on these images, and then GIST features are extracted from detected faces. Finally LSH algorithm is applied to index the high dimensional facial features. A query image given by the user also undergoes the same preprocessing step. Then WLRLCC scheme is applied to obtain final list of annotated face names. WLRLCC algorithm is limited by the discriminative ability of features. So it achieves better performance on a special type of image annotation than generic image annotation [2].

Timo Ahonen et al introduced a new approach for the recognition of face, which considers the information of both shape and texture to represent the face image. First the face image is divided into several small regions; LBP features are then extracted from these regions and concatenated into a single feature histogram. This constructed feature histogram represents both the statistics of the facial micro-patterns and their spatial locations. By the construction of the face feature histogram we can recover the whole shape of the face. The subset of the images is been used to train this algorithm [3].

Consequently identified human traits have been demonstrated promising in diverse applications as of late. Kumar et al. propose a learning system to consequently

 ISSN: 2321-8169
5287 - 5296

discover describable visual characteristics [4]. Utilizing consequently recognized human traits; they accomplish fantastic execution on catchphrase based face picture recovery and face check.

Siddiquie et al. [19] further extend the structure to manage multi-characteristic inquiries for magic word based face picture recovery. Scheirer et al. [20] propose a Bayesian system way to deal with use the human traits for face ID. To further enhance the nature of qualities, Parikh et al. propose to utilize relative properties [22] and Scheirer et al. propose multi-credit space [23] to standardize the certainty scores from diverse characteristic locators for comparative quality pursuit. The works show the rising open doors for the human traits yet are not abused to produce more semantic (versatile) codewords. In spite of the fact that these works accomplish notable execution on magic word based face picture recovery and face acknowledgment, we propose to abuse viable approaches to consolidate low-level elements and naturally identified facial traits for versatile face picture recovery.

Because of the ascent of photograph sharing/informal community administrations, there rise the solid requirements for expansive scale substance based face picture recovery. Substance based face picture recovery is firmly identified with face acknowledgment issues yet they concentrate on discovering suitable element representations for versatile indexing frameworks. Since face acknowledgment as a rule requires significant processing expense for managing high dimensional components and producing unequivocal arrangement models, it is non-insignificant to specifically apply it to face recovery undertakings. In the mean time, the photograph quality in shopper photographs is more various and stances more visual differences. Wu et al. [5] propose a face recovery system utilizing part based neighborhood highlights with character based quantization to manage versatility issues. To remunerate the quantization misfortune, they further propose to utilize best in class highlights with main segment examination for re-ranking.

III. IMPLEMENTATION DETAILS

1 EXISTING SYSTEM

At the point when searching for a photograph at the top of the priority list, it is troublesome and in-proficient for clients to demonstrate the careful document area in the capacity despite the fact that they are very much classified by time or geo-areas. Some overarching photograph sharing sites utilize group sourcing to acquire free labels semantically related to pictures; however the system can't be copied to individual photograph coordinator in light of the fact that clients are not anticipated that would effectively comment their photographs. Existing frameworks overlook solid, face-particular geometric imperatives among diverse visual words in a face picture. At the end of the day, utilizing such worldwide elements as a part of a recovery sys-tem requires basically a straight output of the entire database so as to process an inquiry, which is restrictive for a web-scale picture database.

2 PROPOSED SYSTEM

Albeit momentum business photograph administration programming start to adventure innovations in face acknowledgment and face bunching, such arrangements still do not have the ability of scanning for scenes with countenances sent in a particular design. In light of this perception, our proposed framework endeavors to make shopper photograph administration speedier and less demanding.

We propose two orthogonal techniques named characteristic upgraded scanty coding and trait installed altered indexing. Property upgraded meager coding endeavors the worldwide structure of highlight space and uses a few critical human traits consolidated with low-level elements to build semantic code words in the disconnected from the net stage. Then again, property inserted transformed indexing by regional standards considers human traits of the assigned question picture in a paired mark and gives effective recovery in the online stage.

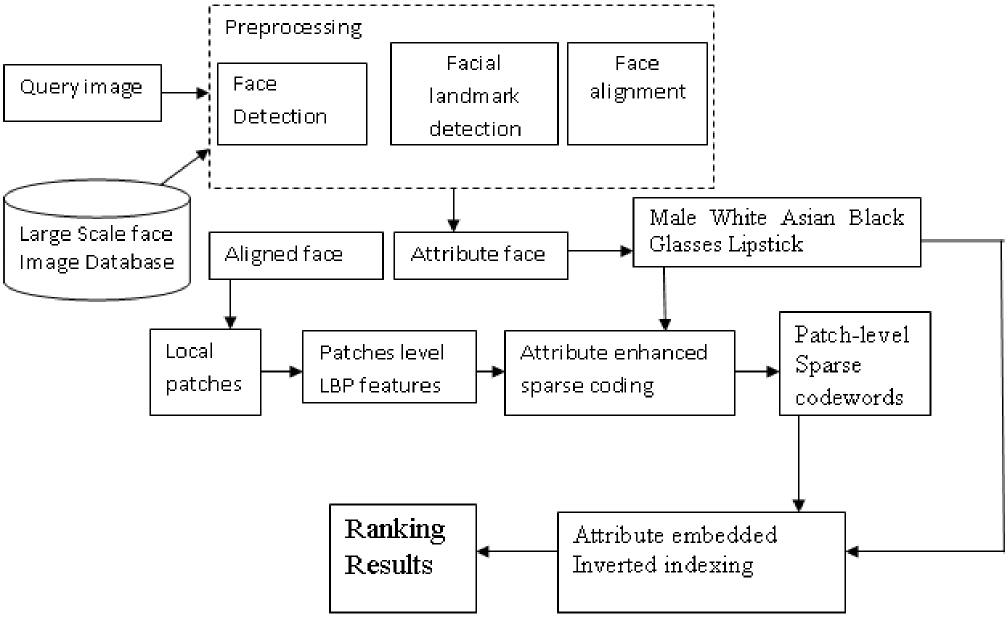

Fig 1: Basic block diagram of the proposed framework.

As we can see in the figure both the query and database images will go through the same procedures which include face detection, facial landmark detection, face alignment, attribute detection and LBP feature extraction. Each and every image in the database as well as the query image first go through the preprocessing step, which includes face detection, facial landmark detection and face alignment. For each aligned faces local patches are generated from which LBP features are extracted. Sparse codewords are generated for these low-level features, then it is combined with high-level attributes in attribute embedded inverted indexing after which gives the ranking results.

Firstly, Viola-Jones face detector algorithm is applied for every image in the database to find the locations of faces. We then find 73 different attribute scores using describable visual attributes framework. Active shape model is applied to locate 68 different facial landmarks on the image. Using these facial landmarks, we apply barycentric coordinate based mapping process to align every face with the face mean shape. For each detected facial component, we will extract 7x5 grids where each grid is a square patch. For each and every image in the database, we extract LBP features and assign a sparse codewords. Each image has separate and unique codewords and stored in matrix form in mat function. When the consumer provides an image as query, codeword is generated to this after extracting features. It will undergo the same steps as the database images. The codeword generated for the query image is then compared with the codewords of the database images. Using attribute embedded inverted indexing method; the top matching images are displayed as ranking results.

3 SPARSE CODING

Inadequate codes of neighborhood elements in a picture at various spatial scales speak to picture with more precision by utilizing area requirements. Conventional BoW model utilization k-means bunching calculation to learn word reference for quantization. The calculation takes care of the accompanying advancement issue:

$$\min_{D,V} \sum_{i=1}^{n} \left\| x_i - Dv_i \right\|^2 \quad [1]$$

Subject to card $(v_i) = 1, \|v_i\|_1 = 1, v_i \geqslant 0, \forall i$

where $D = [d_1, \ldots, d_k]$ is a dictionary matrix with the size of 59×*K*, each column represents a centroid. $V = [v_1, \ldots, v_n]$ is the centroid indicator matrix, each *vi* indicates that the original feature *xi* belongs to a centroid in *D*. The constraint $Card(v_i) = 1$ means each feature can only be assigned to one centroid. This constraint is

 ISSN: 2321-8169
5287 - 5296

considered to be too strict because some features might be at boundary of two or more centroids, therefore many people suggest relaxing the constraint and putting L1 regularization term on *vi* instead. This then turn into another optimization problem known as sparse coding:

$$\min_{D,V} \sum_{i=1}^{n} \left\| x_i - Dv_i \right\|^2 + \lambda \left\| v_i \right\|_1 \qquad [2]$$

This advancement issue can be productively fathomed by an online improvement calculation. Thus, there are up to 175 word references learned. Once the word reference is found out, we can settle D in the above definition and minimize the target work alongside vi independently to locate the inadequate representation of every element. At the point when D is altered, the advancement issue transforms into a minimum square issue with L1 regularization. Since the L1 regularization term makes the target work non-differentiable when vi contains zero components, we can't comprehend it by standard unconstraint advancement technique. Due to the L1 regularization term, each scanty representation vi will just contain a few nonzero components out of K measurement. These nonzero passages are then considered as the visual words from the descriptor xi. Since there are 175 lattice areas, all out lexicon size is 175×K.

3.1 Attribute-enhanced sparse coding:

Keeping in mind the end goal to consider human characteristics in the scanty representation, we first propose to utilize lexicon determination to drive pictures with distinctive credit qualities to contain diverse code words. For a solitary human characteristic, we partition word reference centroids into two distinct subsets, pictures with positive quality scores will utilize one of the subset and pictures with negative trait scores will utilize the other. Case in point, if a picture has a positive male property score, it will utilize the first a large portion of the word reference centroids.

On the off chance that it has a negative male characteristic score, it will utilize the second a large portion of the word reference centroids. By doing these, pictures with diverse qualities will without a doubt have distinctive code words. For the instances of different qualities, we separate the meager representation into various portions in light of the quantity of traits, and every section of scanty representation is created relying upon single property.

The main a large portion of the inadequate representation is prone to contain non-zero worth for the picture patches, and the other way around. To streamline the documentation, the above mathematical statement just considers single quality. For different qualities, we can basically characterize z(i) contingent upon various properties. Case in point, if there are two properties and picture i contains positive scores for both qualities, z(i) will get to be [1; : ; 1;1; : ;1; 1; : ; 1;1; : ;1]T . Effectively encode the human characteristics into the meager representation, it has two fundamental issues. To begin with, it is not hearty to conceivable trait discovery slips.

Since our human qualities are naturally recognized, they are not mistake free. At the point when the identified quality isn't right, it will constrain pictures of the same individual to be connected with very surprising code words. Second, it just encodes the human qualities as paired markers however we really have relative certainty scores. Hence, we further propose to coordinate the relative scores of human credits to unwind into a delicate weighted form by characterizing z(i) in view of the quality scores of pictures.

Inspired by, we first appoint an a large portion of the word reference centroids to have +1 characteristic score and utilization them to speak to pictures with the positive quality; the other 50% of the lexicon centroids are allocated with - 1 to speak to pictures with the negative trait. After the task, we can utilize the separation between property scores of the picture and the credit scores alloted to the lexicon centroids as the weights for selecting code words.

Since the weights are chosen by property scores, two pictures with comparative quality scores will have comparable weight vector, and in this manner have a higher opportunity to be relegated with comparative code words and result in comparative inadequate representations. The above strategy, pictures with comparable qualities will be alloted with comparable centroids, yet pictures with wrong traits may in any case have the capacity to recover right pictures if their unique elements are comparable base right.

3.2 Image ranking and inverted indexing:

For every picture, subsequent to processing the meager representation utilizing the technique, we can utilize codeword set c(i) to speak to it by taking non-zero sections in the inadequate representation as codewords. The closeness between two pictures are then processed as takes after,

$$S(i,j) = ||c^{(i)} \cap c^{(j)}|| \quad [3]$$

The picture positioning as indicated by this closeness score can be productively discovered utilizing transformed list structure.

3.3 Attribute Embedded Inverted Indexing (AEI):

The systems depicted expect to build code words improved by human characteristics. To insert characteristic data into record structure, for every picture, notwithstanding meager codewords $c^{(i)}$ computed from the facial appearance, we use a d_b dimension binary signature to represent its human attribute $b^{(i)}$:

$$b_j^{(i)} = \begin{cases} 1 & \text{if } f_a^{(i)}(j) > 0 \\ 0 & otherwise, \end{cases} \quad [4]$$

The closeness score is then adjusted into,

$$S(i,j) = \begin{cases} ||c^{(i)} \cap c^{(j)}|| & \text{if } h(b^{(i)}, b^{(j)}) \leq T \\ 0 & otherwise, \end{cases} \quad [5]$$

where h(i; j) signifies hamming separation in the middle of i and j, and T is a settled edge such that $0 \leq T \leq d_b$. As indicated in Figure 5, characteristic inserted rearranged list is manufactured utilizing the first codewords and the double property marks connected with all database pictures. The picture positioning as per Equation 2 can even now be productively processed utilizing reversed file by basically doing a XOR operation to check the hamming separation before upgrading the closeness scores. Since XOR operation is speedier than redesigning scores, by skipping pictures with high hamming separation in characteristic hamming space, the general recovery time altogether reductions. Note that stockpiling of the transformed record can be further packed utilizing a wide range of strategies as a part of data recovery.

.

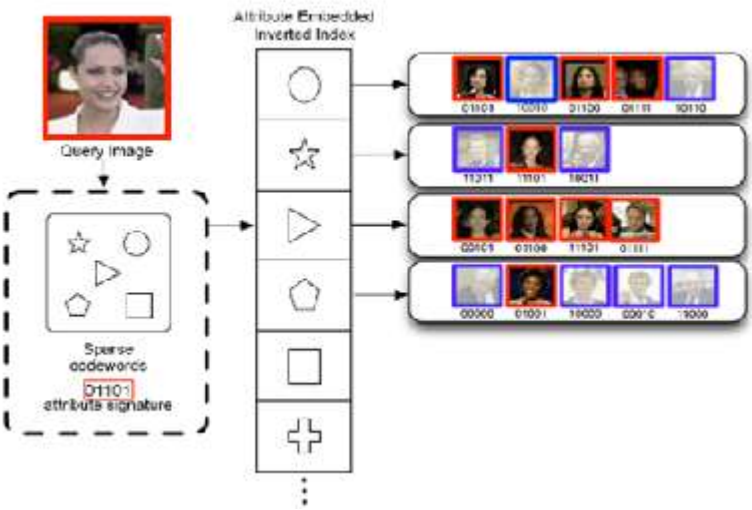

Fig 2: Illustration of attribute-embedded inverted indexing.

.

ISSN: 2321-8169
5287 - 5296

Estimating face similarities:

To empower seek through face appearance, we adjust the face recovery structure. The benefit of this system incorporates:

1. Efficiency, which is accomplished by utilizing scanty representation of face picture with transformed indexing, and
2. Leveraging personality data, which is finished by joining the character data into the improvement process for codebook development, both of the over two focuses are suitable for our framework. In points of interest, distinguished appearances are initially adjusted into sanctioned position, and afterward part based neighborhood parallel examples are removed from the pictures to shape highlight vectors.

Meager representations are further processed from these component vectors in view of a scholarly lexicon consolidated with additional character data. By fusing such structure into our framework, the client cannot just indicate positions and traits of the face additionally utilize a face picture itself with position as the question. The genuine esteemed similitude scores are standardized to the interim (0, 1) preceding they are utilized.

IV. RESULTS

The proposed system makes use of different algorithms at each stage so as to achieve better performance. The proposed system will assure to retrieve similar images. When at input side an query image is provided for similar images as result then after all image processing operations are done on image; set of similar images is get as output . This is shown in figure 3.

Fig 3: Retrieval results using ASC method where we can find only one false image which shows higher precision.

Table 2: Performance of all four methods in terms of MAP. The precision for top 20 images is calculated for all the four methods and tabulated. Then MAP and time is calculated. As we can see the combined ASC (proposed) method shows higher precision compared to other methods.

Performance	P@20	MAP	Time(s)
LBP	63.4%	59.91%	3.43
ATTR	76.7%	66.14%	3.32
SC	84.5%	73.12%	3.44
ASC-combined	93.2%	89.05%	3.21

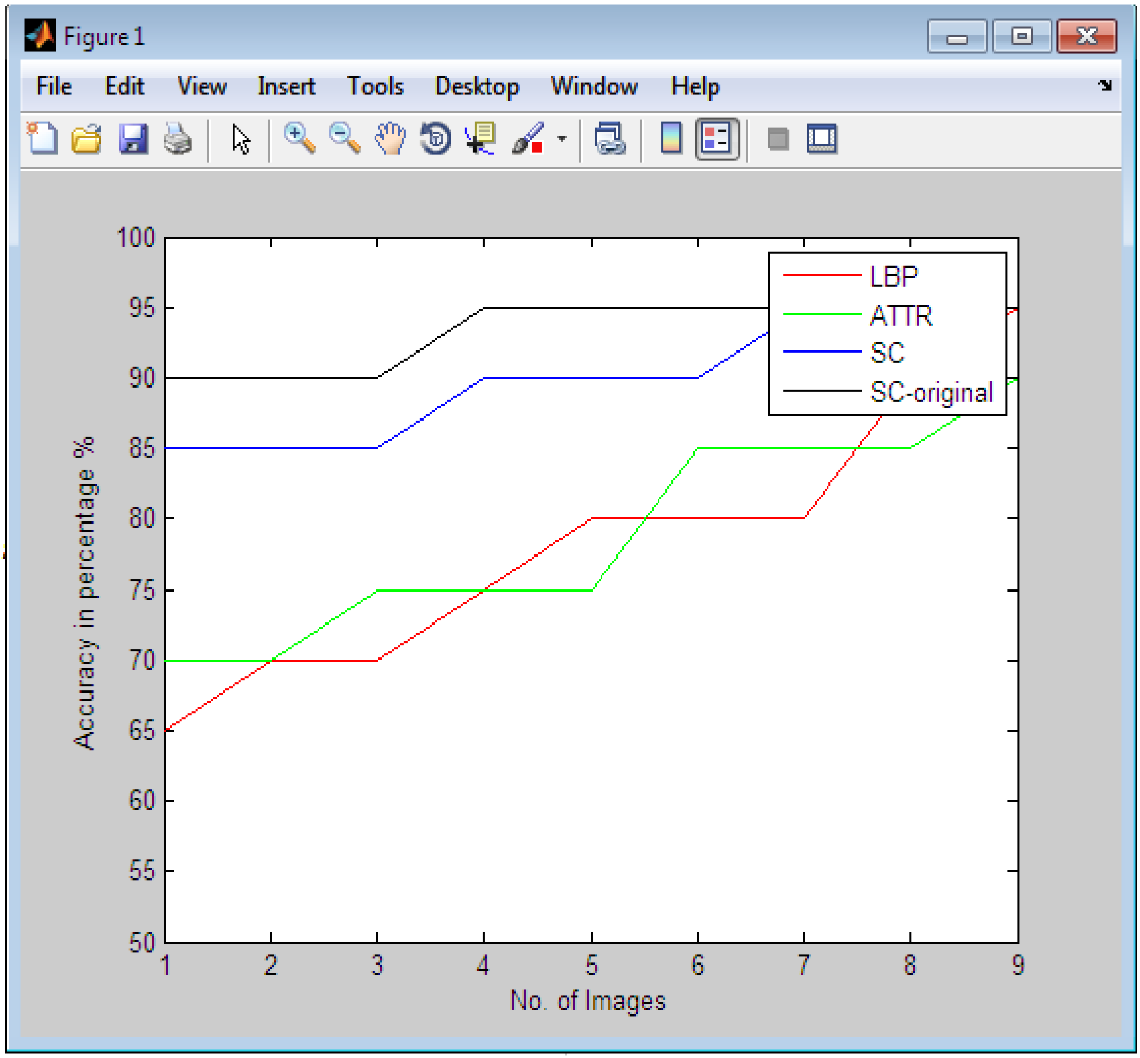

Fig 4: Accuracy plot of all the 4 methods. By comparing the performance of the proposed method (ASC) with other three methods we can see that the retrieval result of our method is more accurate than other methods.

V. CONCLUSION

Another system which consolidates naturally recognized human characteristics with low-level components (LPB) is proposed. Two orthogonal techniques are proposed to use these high-level attributes which enhances the content-based face image retrieval. The two techniques are attribute-enhanced sparse coding and attribute-embedded inverted indexing. The principal strategy utilizes human credits to develop semantic codewords. Second technique further considers the characteristic marks of the question picture which guarantees productive recovery. The proposed framework likewise demonstrates that the recovery results are enhanced contrasted with different techniques.

REFERENCES

[1] Y.-H. Lei, Y.-Y. Chen, L. Iida, B.-C. Chen, H.-H. Su, and W. H. Hsu, “Photo search by face positions and facial attributes on touch devices,” ACM Multimedia, 2011.

[2] D. Wang, S. C. Hoi, Y. He, and J. Zhu, “Retrieval-based face annotation by weak label regularized local coordinate coding,” ACM Multimedia, 2011.

[3] T. Ahonen, A. Hadid, and M. Pietikainen, “Face recognition with local binary patterns,” European Conference on Computer Vision, 2004.

[4] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, “Attribute and simile classifiers for face verification,” International Conference on Computer Vision, 2009.

[5] Z. Wu, Q. Ke, J. Sun, and H.-Y. Shum, “Scalable face image retrieval with identity-based quantization and multi-reference re-ranking,” IEEE Conference on Computer Vision and Pattern Recognition, 2010.

[6] B.-C. Chen, Y.-H. Kuo, Y.-Y. Chen, K.-Y. Chu, and W. Hsu, “Semi-supervised face image retrieval using sparse coding with identity constraint,” ACM Multimedia, 2011.

[7] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, “Robust face recognition via sparse representation,” IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2009.

[8] H. Jegou, M. Douze, and C. Schmid, “Hamming embedding and weak geometric consistency for large scale image search,” European Conference on Computer Vision, 2008.

[9] Pranali Prakash Lokhande, P. A. Tijare, “Feature Extraction Approach for Content Based Image Retrieval,” International Journal of Advanced Research in Computer Science and Software Engineering,2012.

[10] M. Douze and A. Ramisa and C. Schmid, “Combining Attributes and Fisher Vectors for Efficient Image Retrieval,” IEEE Conference on Computer Vision and Pattern Recognition, 2011.

[11] S. Mangijao Singh , K. Hemachandran,“Content-Based Image Retrieval using Color Moment and Gabor Texture Feature,” IJCSI International Journal of Computer Science Issues, 2012.

[12] Bor-Chun Chen, Yan-Ying Chen, Yin-Hsi Kuo, Winston H. Hsu,”Scalable Face Image Retrieval Using Attribute-Enhanced Sparse Codewords,” IEEE Transactions On Multimedia, 2014.

[13] [13] Munmun N. Bhagat and Prof. B. B. Gite,”Image Retrieval using Sparse Codewords with Cryptography for Enhanced Security,” IOSR Journal of Computer Engineering , 2014.

[14] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, “Describable visual attributes for face verification and image search,” in IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), Special Issue on Real-World Face Recognition, Oct 2011.

[15] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, “Labeled faces in the wild: A database for studying face recognition in unconstrained environments,” University of Massachusetts, Amherst, Tech. Rep. 07-49, October 2007.

[16] M.Gopi Krishna, A. Srinivasulu, “Face Detection System On AdaBoost Algorithm Using Haar Classifiers,” International Journal of Modern Engineering Research, 2012.

[17] Paul Viola and Michael Jones,”Rapid Object Detection using a Boosted Cascade of Simple Features,” Accepted Conference on Computer Vision and Pattern Recognition 2001.

[18] M.Risvana Fathima, M.Kaja Mohaideen,S.Sakkaravarthi, “Content Based Image Retrieval Algorithm Using Local Tetra Texture Features”, International Journal Of Advanced Computational Engineering And Networking,2014.

ISSN: 2321-8169
5287 - 5296

[19] B. Siddiquie, R. S. Feris, and L. S. Davis, "Image ranking and retrieval based on multi attribute queries," IEEE Conference on Computer Vision and Pattern Recognition, 2011.

[20] W. Scheirer, N. Kumar, K. Ricanek, T. E. Boult, and P. N. Belhumeur, "Fusing with context: a bayesian approach to combining descriptive attributes," International Joint Conference on Biometrics, 2011.

[21] D. Parikh and K. Grauman, "Relative Attributes," IEEE International Conference on Computer Vision, 2011.

[22] W. Scheirer and N. Kumar and P. Belhumeur and T. Boult, "Multi- Attribute Spaces: Calibration for Attribute Fusion and Similarity Search," IEEE Conference on Computer Vision and Pattern Recognition, 2012.

Author details

Chaitra R
Mtech, digital communication engineering
Acharya Institute of Technology
Bangalore