Scientific Journal of Impact Factor (SJIF): 3.134
ISSN (Online): 2348-4470
ISSN (Print) : 2348-6406

International Journal of Advance Engineering and Research Development

Volume 2, Issue 3, March -2015

ADVANCED TECHNIQUE TO RETRIEVE FACE IMAGE BY USING ATTRIBUTE-ENHANCED SPARSE CODEWORDS& INVERTED INDEXING

Prof.P.Gaikwad[1],Shreekant Sanap[2],Rohit Jadhav[3] ,Sanket Kaware[4],Pratik Mhasawade[5]

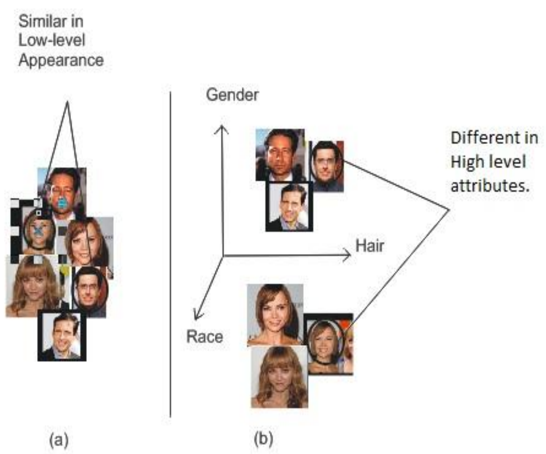
[1,2,3,4,5]*Computer Engineering,AISSMS IOIT*

Abstract—*In today's growing and competitive world we need more and exactly sure security for our applications. Many of them including face verification use face image retrieval method. It's a challenging technique since all the faces will be similar due to its similar geometrical configuration of face structure. Thus, with the exponentially increasing images, large-scale content-based face image retrieval is an enabling technology for many emerging applications.In our project, we aim to improve content-based face retrieval by constructing semantic codewords for efficient large-scale face retrieval and for that we utilize automatically detected human attributes that contain semantic cues of the face photos. In content based image retrieval we can only retrieve the image using low level features. But in our project we have use attribute-enhanced sparse coding and attribute embedded inverted indexing to retrieve the image using high level features (hair style, gender, race) and to improve the face retrieval in the offline and online stages.The results show that the proposed methods can achieve up to 42 to 43% relative improvement compared to the existing methods.*

Keywords-*Content-based image retrieval, face image, human attributes,LFW,Pubfig.***.**

I. INTRODUCTION

As we know in today's world there are largely growing consumer photos available in our life. Among all those photos, a most of them are photos with human faces. human face photos are mainly used in many security systems. Our goal in this project is to address one of the important and challenging problems – large-scale content-based face image retrieval. For a given query face image, content-based face image retrieval tries to find similar face images from a large image database.Low level features are just appearance and posing in which we cannot get the exact information whether it is similar or different faces. For this reason we have usedhigh level attributes which can differentiate the unique faces from all similar faces images. High level attributes used to consider race, hair, gender etc.



(a) Due to old low-level attributes there is lackofsemanticmeanings,face images of two different people might be close in the traditional low-level feature space.
(b) With new high-level human attributes, (e.g. gender, age…etc.) , we can provide accurate discriminability to access face image .

In order to evaluate the performance of the proposed methods, we conduct extensive experiments on two separate public datasets named LFW [1] and Pubfig [6].
Duringthe experiments, we show that the proposed methods can leverage the context information from human attributes to achieve relative improvement up to 43.55% in mean average precision on face retrieval task compared to the existing methods using local binary pattern (LBP) [8] and sparse coding [5].

II. SYSTEM MODEL

We have lot of images in database hence for every image we apply some techniques on them. First of all we need to detect the face from image so we apply Viola-Jones facedetector [12]tofindthelocationsoffaces. After it we have to find 73 different attributes scores for that we use framework proposed in [7]. Now to locate 68 different facial landmarks on the image we make use of Active shape model [13].Using these facial landmarks, we apply barycentric coordinate based mapping process to align every face with the face mean shape [3]. Now it's time to extract grids, for each detected facial component, we will extract 75 grids, where each grid is a square patch [4]. In total we have 175 grids from five components like 2 eyes, tip of nose , and 2 mouth corners etc..on the aligned image using similar methods proposed in [4].Fromeachgrid, we extractanimagepatchandcomputea 59-dimensional uniform LBP feature descriptor as our local feature. Once we get local attribute/featuredescriptor , then we quantizeevery descriptorintocodewordsusingattribute-enhancedsparsecoding described in Section IV-B. Attribute-embedded inverted index described in Section IV-C is then built for efficient retrieval.

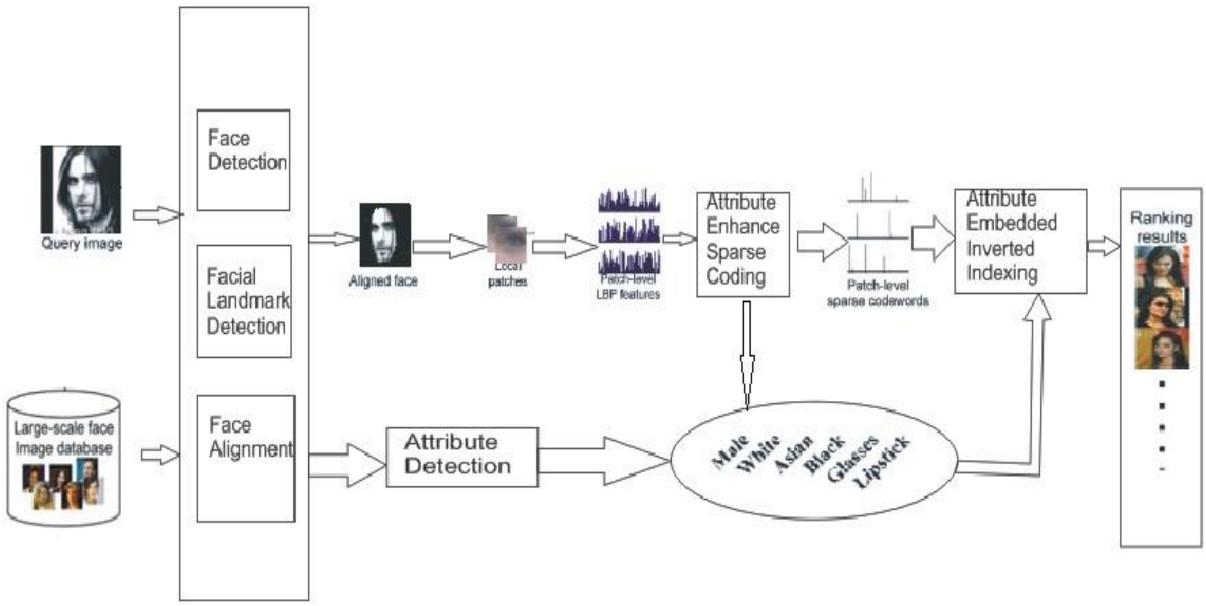


Fig. 3. The proposed system framework.

When a query image arrives, it will go through the same procedure to obtain sparse codewords and human attributes, and use these codewords with binary attribute signature to retrieve images in the index system. Fig. 3 illustrates the overview of our system

III. PREVIOUS WORK

Our project work is closely related to several different research topics, including content based image retrieval (CBIR), human attribute detection, and content-based face image retrieval. Traditional CBIR techniques use different face image attributes/content like texture, color and gradient to represent images. To deal with large-scale data we mainly use two kinds of indexing systems, to achieve efficient similarity search. Although these methods can achieve high precision on rigid object retrieval, they suffer from low recall problem due to the semantic gap [9]. When dealing with face images, prior works [2],[4],[5] usually. Crop only the facial region and normalize the face into the same position and illumination to reduce intra class variance caused by poses and lighting variations. Doing these pre-processing steps, they ignore the rich semantic cues for a designated face such as skin color, gender, hair style. After preprocessing steps, the information loss causes difficulty in identifying attributes (e.g., gender) of the face. Therefore, we propose to use automatically detected human attributes to compensate the information loss. Recently, some researchers have focused on bridging the semantic gap by finding semantic image representations to improve the CBIR performance. [10] And [11] propose to use extra textual in formation to construct semantic code- words; Kumar et al. propose a learning framework to automatically find describable visual attributes [7]. Taking advantages of the effectiveness and simplicity of LBP feature with the superior characteristics of sparse coding on face images, we adopt a similar approach as Chen et al. used in[5] using component-based LBP combined with sparse coding to construct sparse codewords for efficient content-based face image retrieval. However, instead of using identity in-formation that might need manual annotations, we focus on utilizing automatically detected human attributes to construct semantic-aware sparse codewords using attribute-enhanced sparse coding. In addition, we propose another orthogonal approach to further leverage attribute in formation by constructing attribute-embedded inverted index in the online ranking stage. Note that the proposed methods can be easily combined with the method proposed in [5] to take advantage of both identity information and automatically detected human attributes. we propose to exploit effective ways to combine low-level features and automatically detected facial attributes for scalable face image retrieval. To the best of our knowledge, very few works aim to deal with this problem.

IV. PROPOSED METHODOLOGY

Attribute-Enhanced Sparse Coding (ASC)

As we are using sparse coding in our project for face image retrieval, let's see how to use sparse coding Wethendescribedetailsoftheproposed attribute-enhanced sparse coding. Here we apply the same procedures to all patches in a single image to find different codewords and we need to combine all these codewords together to represent the image.

1) Sparse Coding for Face Image Retrieval (SC):

We can solve the following optimization problem Using sparse coding for face image retrieval :

$$\min_{D,V} \sum_{i=1}^{n} \|x^{(i)} - Dv^{(i)}\|_2^2 + \lambda\|v^{(i)}\|_1$$

$$\text{subject to } \|D_{*j}\|_2^2 = 1, \quad \forall j \quad (2)$$

Where,

$x^{(i)}$: Istheoriginalfeaturesextractedfromapatch of face image i .

$D \in R^{d*K}$: Is a to be-learned dictionary contains K centroids with d dimensions.

$V=[v^{(1)},v^{(2)},\ldots\ldots,v^{(n)}]$: Is the sparse representation of the face image spots/patches. The constraint on eachcolumnof D (D_{*j})istokeep 'D' from becoming large. By utilizing sparse coding, a feature is linear combination of the column vectors of the dictionary. [14] Provides an efficient online algorithm for solving the above problem. Note that we apply the above process to 175 different spatial grids separately, so codewords from different grids will never match. Accordingly, we can encode the important spatial information of faces into sparse coding.

2) Attribute-Enhanced Sparse Coding (ASC):

As we know that the images with different attributes will surely have different codewords. For the casesofmultipleattributes, wedividethesparserepresentation into multiple segments based on the number of attributes, and each segment of sparse representation is generated depending on single attribute. The above goal can be achieved by solving the following optimization problem modified from (2):

$$\min_{V} \sum_{i=1}^{n} \|x^{(i)} - Dv^{(i)}\|_2^2 + \lambda\|z^{(i)} \circ v^{(i)}\|_1$$

$$z_j^{(i)} = \begin{cases} \infty, & \text{if } (1)\ j \geq \lfloor \frac{K}{2} \rfloor \text{ and } f_a(i) \geq 0 \\ & \quad (2)\ j < \lfloor \frac{K}{2} \rfloor \text{ and } f_a(i) < 0 \\ 1, & \text{otherwise,} \end{cases} \quad (3)$$

Where,

'o': Is pairwise multiplication between two vectors.

$F_a(i)$: Is the attribute score for i_{th}image.

$Z^{(i)}$: is a mask vector for deciding which codewords are allowed to be used by image i .

By using the mask vector $z^{(i)}$,it forces the sparse representation$v_j^{(i)}$ to be zero if $z_j^{(i)}$ is ∞ because any other values in these dimensions will cause the objectivefunction to become infinity. The final sparse representation $v^{(i)}$can be found by solving a L1 regularized least square problem and only considering the dimensions where $z_j^{(i)}$=1 .

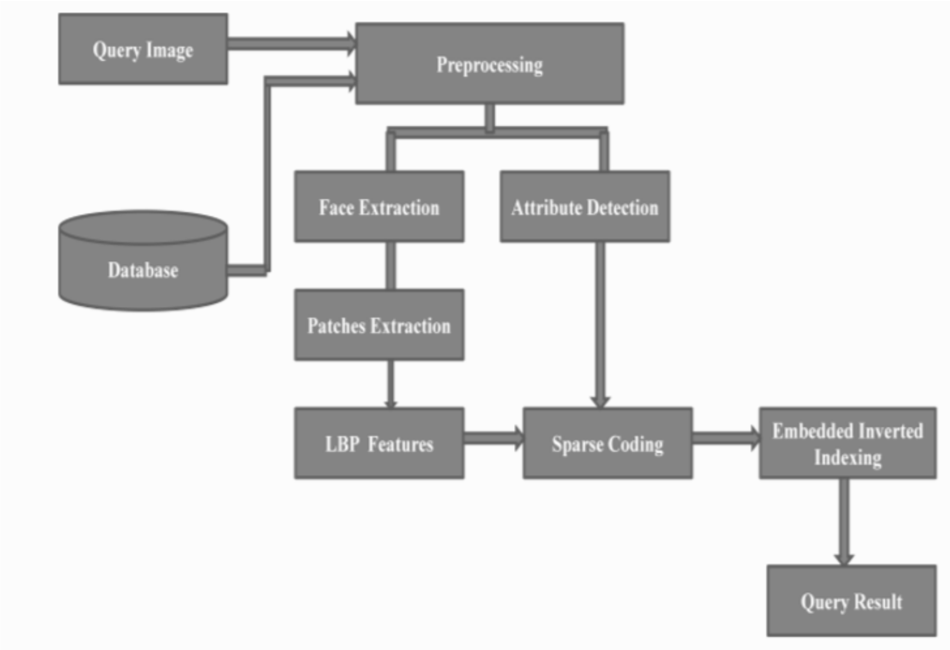


Fig.Proposed architecture of image retrieval system with enhanced sparse coding.

A) *Attribute Embedded Inverted Indexing (AEI):*

As we have describedmethods inSection IV-A thataimtoconstructcodewordsenhancedbyhumanattributes.Now wedescribe thesecondmethodthatcanutilizehumanattributesbyadjusting the inverted index structure

1. Face image Ranking with Inverted Indexing:

After finding thesparserepresentation for each imageusingthemethod describedin Section IV-A,wecanusecodewordset to representitbytakingnon-zeroentriesinthesparserepresentationas codewords. Hence we can compute the similarity between two images as follows,

$$S(i,j) = \|c^{(i)} \cap c^{(j)}\|.$$

The image ranking according to this similarity score can be efficiently found using inverted index structure.

2. Inverted Indexing:

Now we have to manage attribute information into index structure, for each image, in additiontosparsecode words $c^{(i)}$computed from the facial appearance, we use a d_b a dimension binary signature to represent its human attribute,$b^{(i)}$:

$$b_j^{(i)} = \begin{cases} 1 & \text{if } f_a^{(i)}(j) > 0 \\ 0 & \text{otherwise.} \end{cases} \qquad (7)$$

The similarity score is then modified into,

$$S(i,j) = \begin{cases} \|c^{(i)} \cap c^{(j)}\| & \text{if } h(b^{(i)}, b^{(j)}) \le T \\ 0 & \text{otherwise,} \end{cases} \qquad (8)$$

Where,h(i , j) denotes hamming distance between i and j, and T is a fixed threshold such that $0 \le T \le d_b$.

V. Conclusion

We propose and combine two orthogonal methods to utilize automatically detected human attributes to significantly improve content-based face image retrieval (up to 43% relatively).To the best of our knowledge, this is the first proposal of combining low-level features and automatically detected human attributes for content-based face image retrieval. Current methods treat all different attributes as same. We have used methods to dynamically decide the importance of the attributes and further exploit the contextual relationships between them. Attribute enhanced sparse coding will extract less amount of images which exactly match with the query images. This is an efficient method compared to the existing methods.

VI. FUTURE SCOPES

- By using sparse codewords we directly compare this attributes to match any face images.
- Due to presence of sparse attributes we do not require any other details of person to match his identity.
- We can provide better discriminability for face image retrieval (gender, race, hairstyle, etc.).

VII. REFERENCES

[1] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, Labeled Faces in the Wild: A Database for Studying Face Recognition in Un- constrainedEnvironments, Univ.Massachusetts,Amherst,MA, USA, 2007, Tech. Rep. 07-49.

[2] D. Wang, S. C. Hoi, Y. He, and J. Zhu, “Retrieval-based face annotationbyweaklabelregularizedlocalcoordinatecoding,”inProc.ACM Multimedia, 2011.

[3] U.ParkandA.K.Jain,“Facematchingandretrievalusingsoft biometrics,” IEEE Trans. Inf. Forensics Security, vol. 5, no. 3, pp. 406–415, Sep 2010.

[4] Z. Wu, Q.Ke, J. Sun, and H.-Y.Shum, “Scalable faceimageretrieval with identity-based quantization and multi-reference re-ranking,” in Proc.IEEEConf.ComputerVisionandPatternRecognit.,2010.

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

[6] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, “Attribute and simile classifiers for face verification,” in Proc. Int. Conf. Computer Vision, 2009.

[7] N.Kumar,A.C.Berg,P.N.Belhumeur,andS.K.Nayar,“Describable visual attributes for face verification and image search,” IEEE Trans. PatternAnal.Mach.Intell.,SpecialIssueonReal-WorldFace Recognition, vol. 33, no. 10, pp. 1962–1977, Oct. 2011.

[8] T.Ahonen,A.Hadid,andM.Pietikainen,“Facerecognitionwithlocalbinarypatterns,”inProc.Eur.Conf.ComputerVision,2004.

[9] O.Chum,J.Philbin,J.Sivic,M.Isard,andA.Zisserman,“Totalrecall:Automaticqueryexpansionwithagenerativefeaturemodelforobject retrieval,” in Proc.IEEEInt.Conf.ComputerVision,2007.

[10] L. Wu, S. C. H. Hoi, and N. Yu, “Semantics-preserving bag-of-words modelsandapplications,”IEEETrans.ImageProcess., vol.19, no.7, pp. 1908–1920, Jul. 2010.

[11] Y.-H. Kuo, H.-T. Lin, W.-H. Cheng, Y.-H. Yang, and W. H. Hsu, “Unsupervised auxiliary visual words discovery for large-scale image object retrieval,” in Proc. IEEE Conf. Computer Vision and Pattern Recognit., 2011.

[12] P.ViolaandM.Jones,“Rapidobjectdetectionusingaboostedcascadeofsimplefeatures,”inProc.IEEEConf.ComputerVisionandPattern Recognit., 2001.

[13] S.MilborrowandF.Nicolls,“Locatingfacialfeatureswithanextendedactiveshapemodel,”inProc.Eur.Conf.ComputerVision,2008.

[14] J.Mairal,F.Bach,J.Ponce,andG.Sapiro,“Onlinedictionarylearning for sparse coding,” in Proc. ICML, 2009.

