International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Face Image Retrieval from Sparse Codeword's

Vandana Shinde[1], S.R.Durugkar[2]

[1]Department of Computer Engineering, SNDCOE and RC Yeola, Nashik-423401

[2]Professor, Department of Computer Engineering, SNDCOE and RC Yeola, Nashik-423401

Abstract: ***The sizes of digital image collections have been changing since its progression. These images are increasing in the size and consists of human faces majority of times. Nowadays photos with people are becoming an area of interest for many users. Thus, many real time applications work on image retrieval techniques. These content-based image retrieval techniques work on a large database containing photos. The goal of the proposed system is to detect the human attributes automatically by utilizing semantic clues of face photos. This results into improved content-based image retrieval. This is done with the help of semantic codeword's. We described two orthogonal approaches named as attribute-embedded inverted indexing and attribute-enhanced sparse coding for improved face retrieval in both offline and online modes.***

Keywords: Face image, human attributes, content-based image retrieval, sparse coding, attribute embedded inverted indexing.

1. Introduction

Image retrieval is research area concerned with retrieving and searching digital images from a collection of database. The research has been started from the 1970s . Researchers started to work on interesting areas like image processing, digital libraries, astronomy, remote sensing, multimedia and other related areas.

Image retrieval system is used to retrieve the relevant images from the collection of images which is based on the query image. There are two types of research communities, First Text-based image retrieval techniques is uses text to describe the content of the image. Second Content-based image retrieval (CBIR) or Visual based image retrieval uses visual features to describe the content of the image.

The rest of the paper is organized as follows. Section II discusses content based image retrieval. Section III describes our observations on the face image retrieval problem with various state of Art methods. Section IV introduces the proposed methods including attribute-enhanced sparse coding and attribute-embedded inverted indexing. Section V shows the result of system section VI concludes this paper.

2. Content Based Image Retrieval

Content-based image retrieval systems are deal with retrieval of images and issues of automatic indexing. The basic image retrieval system is shown in figure 1. This system consist of 3 main modules.

1) **Input Module:** The feature is extracted from input. It is then stored with input image in the image database.
2) **Query Module:** When a query image enters in this module, it extracts the feature vector of the query image.
3) **Retrieval Module:** The extracted feature vector is analyzed or compared with the feature vectors stored in the image database.

As a result of query, the matching images are retrieved according to their nearest matching scores. The result image will be obtained from the retrieved images.

3. Problem Identification Based On State of Art Methods

The content-based image retrieval are divided into three phases: 1) Retrieval based on artificial notes 2) The retrieval based on vision character of image contents. 3) The retrieval based on image semantic features. First the image retrieval is based on artificial notes but it changed image retrieval into traditional keywords retrieval. There are two disadvantages of this method: 1) It brings too heavy workload 2) It still remains

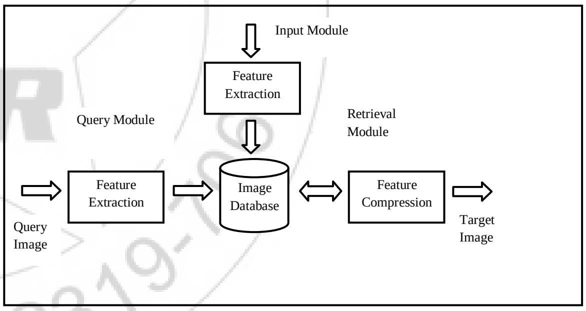

Figure 1: Block diagram of image retrieval system

Subjectivity and uncertainty. The character of this method is image feature extraction, whether the retrieval is good or not depends on the accuracy of the features extraction. so the research on vision features is good research topic in the academic community.

4. Proposed System and Methodology

Nowadays many people are sharing their lives on social sites through multimedia sharing. This includes sharing of photos, videos, etc on various social network sites like Facebook, Flickr, etc. Hence, each consumer has a large number of photo collections. In the collection, there is a big percentage of photos with human faces. These human face photos are of importance when dealing with many applications of real world. Image retrieval technique has been gaining its popularity. But there are some challenges while working on this technique. Thus, in this paper we discuss on one of the important problem known as content-based face image retrieval.

The term content-based face image retrieval is the process of finding similar face images from a huge database containing images. The input image is called as a query face image, which is to be searched. This technique is gaining popularity due to its use in many applications like face annotation, crime investigation, etc. Conventional methods for face image retrieval were using low-level features for representing faces. But these low-level features lacked semantic meaning, resulting in unsatisfactory images. For dealing with this problem the authors Wu et al. and Chen et al. proposed different methods like identity based quantization and identity constrained sparse coding respectively. But these methods, required clean training data and massive human annotations.

In the proposed work we provided an approach with content-based face image retrieval by assimilating high-level features for image representation and index structure. In many situations, the two different face images might look very close enough because of the low-level feature spaces. In this situation, if we use high-level attributes, then it is possible to represent the image features in a better way, and achieve good retrieval results.

Every human have various attributes like gender, race, hair style, etc. These attributes can give a high-level semantic description of a person. Figure 2 (a) shows some examples of human attributes. Nowadays, these attributes are used in many real world applications like face verification, keyword-based face image retrieval, face identification, etc. Researchers have worked on these techniques.

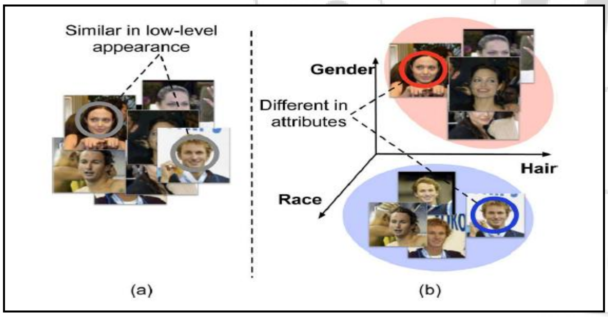

Figure 2: (a) Similar in low level appearance (b) High level human attributes (Different in Attributes) e.g. Gender

In the proposed system, we described two orthogonal approaches named as attribute-embedded inverted indexing and attribute-enhanced sparse coding. When the user uses attribute-enhanced sparse coding approach, many important human attributes and global structure of feature space are used in conjunction with low-level features for creating semantic code words in offline mode. But while working with the other approach, attribute-embedded inverted indexing uses human attributes of the query image in a binary signature which results into retrieval of images in online mode. Using both the approaches benefits with a large-scale content-based face image retrieval system. This system includes advantages, low-level features and high-level semantics. For examining the proposed system, we executed experiments on two datasets. The system can hold up to 43.55 percent of improvement in face retrieval task as compared to existing systems.

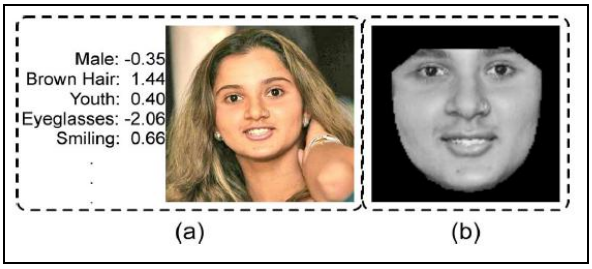

Figure 3: (a) A face image contains rich context information such as hair color, skin color, race, gender, etc.). (b) The same image after pre-processing steps in prior works for face image retrieval or recognition.

We propose a scalable content-based face image retrieval system, with two proposed approaches: attribute-enhanced sparse coding and attribute-embedded inverted indexing. We start our work on a database containing images by applying Viola-Jones face detector. There are different facial marks on every image. The use of active shape model is required for detecting the 68 different facial marks. After detecting the facial marks, we then perform barycentric coordinate mapping process for aligning the faces. There will be exact 75 grids in each facial component making a total of 175 grids consisting of two eyes, nose tip, etc. A grid is a square patch.

After aligning the faces, we then obtain local feature descriptor and then make codeword's by using attribute-enhanced sparse coding. Finally we image retrieval is performed by using attribute-embedded indexing. Each and every query image goes through the above discussed steps as illustrated in figure 3.

Dealing with the face images includes pre-processing on the images. This includes cropping of face region in the image, normalizing the face to the same position, performing illumination on the image for reducing intra-class variance. This pre-processing may ignore the semantic cues. For example figure 4 illustrates a face image that is before and after the pre-processing steps. The pre-processing steps may lose face identification attributes. For dealing with this issue, in our system we used automatic human attribute detection for recompensing the data loss.

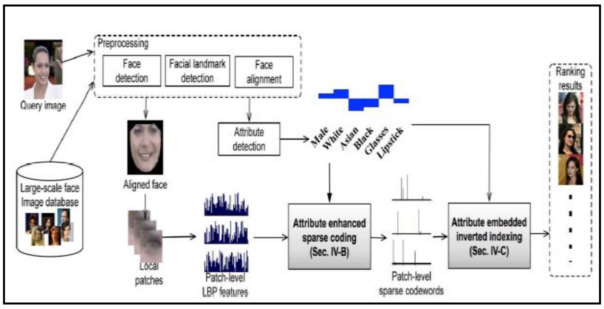

Figure 4: Framework of proposed system

Algorithm:

For implementation of this system following algorithm have been used.

1. Haar-cascade algorithm to find the location of face and detecting face.

2. Locating facial landmark.
3. Extract image patch.
4. LBP feature for each patch.
5. Quantize every descriptor into codewords using attribute enhanced sparse coding.
6. Use this codewords with binary signature to retrieve image in the index system.

- **Attribute-Enhanced Sparse Coding (ASC)**

Sparse coding is one of the approaches for face image retrieval. In our proposed work, we used an attribute-enhanced sparse coding applied to all the patches of a single image. The codeword's are then combined together for representing a single image. Let us discuss the process of sparse coding for image retrieval.

a) Sparse Coding for Face Image Retrieval (SC): The sparse coding technique for face image retrieval is a combination of two processes: dictionary learning and sparse feature encoding. Learning the dictionary which is a large vocabulary is a time consuming process. Hence, the author Coa tes et al. has suggested an idea of using randomly sampled image patches as dictionary. The use of the sampled image patches produces similar result than using a learned dictionary. After learning the dictionary, we solve the problem by LARS algorithm. Then we identify the nonzero entries as codeword's of images. The codeword's of 175 different grids would never match to each other enabling to encode the spatial information into sparse coding.

b) Attribute-Enhanced Sparse Coding (ASC):
For sparse representation of human attributes, we enforce the process of dictionary selection (ASC-D), for assuring every image with different attributes should contain different codeword's. As shown in figure 5, for every human attribute, the dictionary is divided into two dissimilar subsets. One subset is used by the images with positive attribute scores and the other subset is used by the images with negative attribute scores. Due to these dissimilar subsets, we can surely have non-identical codeword's.

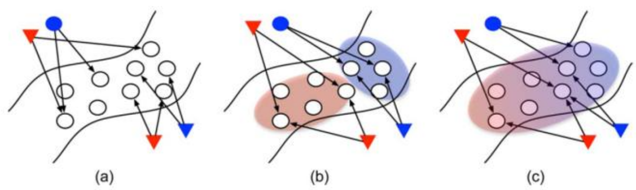

Figure 5: Comparison between attribute enhancing coding methods: SC, ASC-D and ASC-W.

- **Attribute Embedded Inverted Indexing (AEI)**

The above approach included the construction of codeword's from human attributes. The second approach for utilizing the human attributes is discussed in this section. The human attributes are used by adjusting the inverted index structure.

a) Image Ranking and Inverted Indexing: After the images are represented in sparse representation, the codeword's can be used for representing. There is a set of codeword's say c(i). The computation of similarity between two images is done as follows:
S(i,j)=||c(i)\c(j)||
The ranking of image based on these similarities can be efficiently found using inverted index structure.

b) Attribute-Embedded Inverted Indexing*:* We can use a binary signature for embedding the attribute information into index structure. The signature is a dimension binary signature used for representing the human attributes. The sparse codeword's set c(i) with combination with the signature b(i) represents the attributes as:
b(i) = 1 if f(i) (j)>0
0 otherwise.
The attribute-embedded inverted index is constructed with the features of database images. These features include the codeword's and binary signature of the image.

5. Contribution

We have added some more function to existinig system .We have provided other retrieval system by using the query. By using this query we can retrieve the image on the basis of attributes. This query base retrieval gives the different result as compare to the image base retrieval. This query base retrieval use the same database as the image base retrieval database.

6. Result Set

For every image in database we are applying Haar-cascade algorithm which detect the face As shown in the below figure query image is seleced then the face is detected after detecting the face face get crop from query image patches are find.

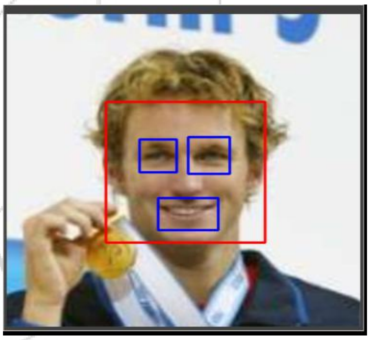
Figure 6: Query image with detected face

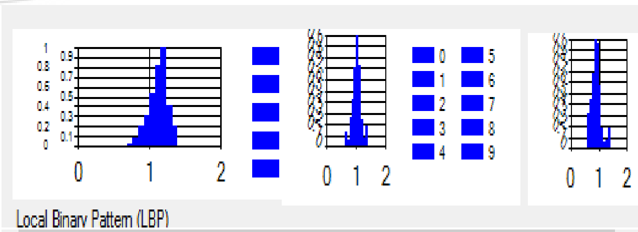

Figure 7: Local Binary Patterns.


Figure 8: Retrieval result.

Local binary patterns are find for each patch of image .It goes through sparse codeword's and human attributes and use these codeword's with binary attribute signature to retrieve image in the system.

7. Conclusion

We combine two methods to utilize automatically human detected attributes to significantly improve content-based face image retrieval. Attribute-enhanced sparse coding uses several human attributes and exploits the global structure. These human attributes used to construct semantic aware codeword's in offline situation. Attribute-embedded inverted indexing considers local attributes of the query image and used for retrieval in online situation. The proposed scheme can easily integrated into inverted index to maintain a scalable framework. We invented methods which dynamically decide the importance of the attributes to exploit the contextual relationships between them.

References

[1] W. Scheirer and N. Kumar and P. Belhumeur and T. Boult, "MultiAttribute Spaces: Calibration for Attribute Fusion and Similarity Search," IEEE Conference on Computer Vision and Pattern Recognition, 2012.
[2] 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.
[3] D. Wang, S. C. Hoi, Y. He, and J. Zhu, "Retrieval-based face annotation by weak label regularized local coordinate coding," ACM Multimedia, 2011.
[4] 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.
[5] B.-C. Chen, Y.-H. Kuo, Y.-Y. Chen, K.-Y. Chu, and W. Hsu, "Semisupervised face image retrieval using sparse coding with identity constraint," ACM Multimedia, 2011.
[6] 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 Vi sion and Pattern Recognition, 2010.
[7] U. Park and A. K. Jain, "Face matching and retrieval using soft biometrics," IEEE Transactions on Information Forensics and Security, 2010.
[8] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Localityconstrained linear coding for image classification," in Proc. IEEE Conf. Computer Vision and Pattern Recognit., 2010.
[9] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in Proc. IEEE Conf. Computer Vision and Pattern Recognit., 2009.
[10] R. Brunelli and O. Mich, "Histograms Analysis for Image Retrieval," Pattern Recognition, Vol.34, No.8, pp1625–1637,2001.
[11] B. Brandshaw. "Semantic based image retrieval: aprobabilistic approach," proc, ACM Multimedia, October 2000. http://www.cs.virginia.edu/papers/MIS03.pdf
[12] M. Worring and G. Schreiber. Semantic image and video indexing in broad domains. IEEE Transactions on Multimedia, 9(5):909{911,August 2007
[13] Shneiderman B and Kang H (2000) Direct Annotation: A Drag-and-Drop Strategy for Labeling Photos. In: Proc. International Conference Information Visualisation (IV2000). London, England.

Author Profile

V. K. Shinde , Post Graduate Student, was with Pune University , Maharashtra, India. She is now with the Department of Computer Engineering, SND COE, Pune University, Maharashtra. India.