ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 8, August 2016**

# Face Recognition Using Attribute Based Tree Construction and Sparse Codewords

Swati k Chavan [1]

M.E. Student, Department of Information Technology, RMD Sinhgad School of Engineering, Warje, Pune, India[1]

**ABSTRACT**: Improvement in digital technologies makes availability of digital camera and camera smart phones easier. Many of peoples share photos clicked by them in online system. Photos with people (e.g., family, friends, celebrity , etc.) are the major interest of users. Thus, with the exponentially growing photos, large-scale content-based face image retrieval is an enabling technology for many emerging applications. Given a query face image, content-based face image retrieval tries to find similar face images from a large image database. Existing systems ignore strong, face-specific geometric constraints among different visual words in a face image. Recent works on face recognition have proposed various discriminative facial features. However, these features are typically high dimensional and global, thus not suitable for quantization and inverted indexing. It requires essentially a linear scan of the whole database in order to process a query, which is prohibitive for a web scale image database. The proposed face image Recognition uses semantic codeword generation method to utilize human face attributes that contain semantic cues of the face photos to improve content based face retrieval. It contains two methods named attribute-enhanced sparse coding and attribute embedded tree construction to improve the face retrieval in the offline and online stages. The sparse codeword of existing images from database are get compared with sparse codeword of query image and attribute based tree construction gives result of face image retrieval ordered from higher matching to lower matching. Experimenting LFW public datasets, the results show that the attribute based tree construction methods can achieve up to 75% relative improvement compared to the existing methods.

**KEYWORDS**: Face detection, Face attributes, sparse code words generation, content based face photo search, photo search.

## I. 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.

Content based image retrieval(CBIR) is a technique which use visual contents of image to search images from large scale image database according to user's interacts. CBIR also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is application of computer vision techniques to the image retrieval probe.

A new side of Content based face photo retrieval is used by incorporating high-level human attributes into face image representation and performing image search on attribute based tree construction to improve search presentation. It is useful in human face image search on social networking websites. Also, it is useful in crime investigation to find out suspects from large database of human face images [1].

Attribute-enhanced sparse coding exploits the global structure and uses several human attributes to construct semantic codewords. Attribute-embedded further considers the local attribute signature of the query image and attribute based binary tree construction ensures efficient retrieval in the online stage. The codewords generated by the proposed coding scheme helps in reducing the quantization error and improve face retrieval. Proposed system also suggest certain informative attributes for face retrieval and these attributes are also talented for other applications like face verification. Sparse modelling of image patches has been successfully applied to tasks such as image and video denoising, and segmentation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 8, August 2016**

Attribute based tree construction system is as said combination of low-level features and automatically detected human attributes for content-based face image retrieval. Attribute-enhanced sparse coding exploits the global structure and uses several human attributes to construct semantic codewords. Attribute-embedded further considers the local attribute signature of the query image and attribute based binary tree construction ensures efficient retrieval in the online stage. The codewords generated by the proposed coding scheme helps in reducing the quantization error and improve face retrieval. Proposed system also suggest certain informative attributes for face retrieval and these attributes are also promising for other applications like face verification.

## II. RELATED WORK

This work is directly related to several different research topics, including retrieval-based face annotation, identity based quantization, image ranking and retrieval based on Multi-Attribute queries and multi-Attribute Space Similarity.

**1)Retrieval-based face annotation**

Retrieval-based face annotation is a capable theory mining enormous web facial images for programmed face annotation Such an annotation concept usually encounters two key challenges. The first challenge is how to efficiently retrieve a short list of most similar facial images from facial image databases, and the second challenge is how to effectively complete annotation by exploiting these like facial images and their strong labels which are often noisy and partial.

Auto face annotation is beneficial number of real-world applications. For example, using auto face annotation techniques, online photo input sites or public networks can automatically annotate users' uploaded photos to facilitate online photo search and management tasks. Auto face annotation can also be applied in news video area to detect important persons appeared in the videos to make easy news video retrieval and summarization tasks [2].

The system flow of the Retrieval-based Face Annotation is as follows:

- It collect weakly labelled facial images from standard database.
- It perform face finding and alignment, then extract GIST features from the detected faces, and finally apply Locality Sensitive Hashing(LSH) to index the high-dimensional facial features.
- A query facial image uploaded by the user is transformed into a quality vector with the same preprocessing step, using content-based facial image retrieval engine.
- A short list of top most similar facial images and their associated names are retrieved and passed to the next learning and annotation stage.
- The weak labelled retrieval co-ordinate coding scheme is applied to go back the final list of (ranked) annotated face names.

**2) Identity based quantization**

The scalability of extracted local features is need to be quantized into a set of discrete visual words using a visual vocabulary which is often obtained by an unsupervised clustering algorithm (e.g., *k*-means). But the unsupervised knowledge is not very good for training a vocabulary for face images, where intra-class variations are often big than interclass variations when the face undergoes pose and appearance changes. Quantization errors will degrade the retrieval performance.

The identity based quantization of face images is one person and constructed visual words. As each person has a single "person id" and each grid has a unique "position id", it define a visual word as the pair <person id, position id>and associate it with *T* local feature descriptors computed from the training samples of the "person id". In other words, each visual word is an example-based representation containing multiple examples. That is the strength of identity-based quantization. The features under various pose/lighting/expression conditions have a chance to be quantized to the same visual word [3].

**3) Image Ranking and Retrieval based on Multi-Attribute Queries**

A novel advance for ranking and retrieval of images based on multi-attribute queries. Existing image retrieval methods train separate classifiers for each word and heuristically combine their outputs for retrieving multiword queries. Moreover, these approaches also ignore the inter dependencies among the query terms. In contrast, it propose a principled approach for multi-attribute retrieval which explicitly models the correlations that are present between the attributes.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

Consider the problem of ranking/retrieval of images of people according to queries describing the physical traits of a person, including facial attributes (e.g. hair colour, presence of beard or moustache, presence of eyeglasses or sunglasses etc.), body attributes (e.g. colour of shirt and pants, striped shirt, long/short sleeves etc.), demographic attributes (e.g. age, race, gender) and even non-visual attributes (e.g. voice type, temperature and odour) which could potentially be obtained from other sensors. There are several applications that naturally fit within this attribute based ranking and retrieval framework. An example is criminal investigation. To locate a suspect, law enforcement agencies typically gather the physical traits of the suspect from eye witnesses. Based on the description obtained, entire video archives from surveillance cameras are scanned manually for persons with similar characteristics. This process is time strong and can be drastically accelerated by an effective image search mechanism.

Conventional image retrieval methods consider only the attributes that are part of the query, for retrieving relevant images. Ranking approach also takes into account the remaining set of attributes that are not a part of the query. For example, given the query "young Asian woman wearing sunspecs", the system infers that relevant images are unlikely to have a moustache, beard or blonde hair and likely to have black hair, thereby achieving superior results[4].

**4) Multi-Attribute Space Similarity**

Measuring the similarity of two images is an significant and well-studied problem. However, most approaches do so on simple trans formations of the original image space and are thus strongly biased by the "configuration" of the objects like pose, illumination, expression, etc. For instance, in face recognition, images of different individuals taken under the same pose and lighting are often more similar than those of the same character under different conditions. Since attributes are designed to capture aspects of appearance independent of imaging configuration, it should be able to measure parallel better in a multi-attribute space. Even with the use of attributes; perceptual similarity is not uniform across all values for a given attribute, or between attributes.

Since it is infeasible to directly measure these perceptual distances over an entire multi-attribute space, it must rely on a method that can estimate these directly from individual classifier outputs, without ground truth labels. [5]

Traditional techniques mainly called as CBIR uses colour, texture and gradient to represent images. To deal with large scale data, mainly leveraged inverted indexing or hash based indexing are used. These indexing systems combined with bag-of-word model (BoW) [6] and local features like SIFT [7], gives efficient similarity search. But these systems face a problem of low recall due to the semantic gap [8].

## III. SYSTEM ARCHITECTURE

When input query face photo image is given, then the goal of face image retrieval is to find the ranking result from most to least similar face images from face image database. The proposed system was useful in human face image search on social networking websites. Also, it is useful in crime investigation to find out suspects from large database of human face images.

Proposed scalable face image retrieval system has following objectives:

- The systems accept strong, face-specific geometric constraints.
- It uses different visual words in a face image. Features are typically high-dimensional and global, are suitable for quantization and inverted indexing.
- Attribute-enhanced sparse coding exploits the global structure and uses several human attributes.
- Face recognition is easy to use and in many cases it can be performed without a person even knowing.
- Sparse codeword generated by this scheme achieves salient gains in face retrieval.
- Design sparse coding method which combines automatically detected high level features with low level features for better image representation to achieve improved face retrieval result.
- Use regression technique for efficient and fast retrieval result.

The proposed system contains mainly three modules:

A) Image Pre-processing.

B) Sparse Codeword Generation.

C) Attribute based Tree Construction.

The face photo images from facial database and the input query face photo image, both undergo same modules but

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

with different intentions. The database face photo images undergo the process flow of first two modules, image preprocessing and sparse codeword generation. After codewords generation modules, the generated sparse codewords from face photo image database are stored. When input query face photo image was submitted by user, it undergoes first two modules same as database face photo images. After that, attribute based tree was constructed against this sparse codewords and already generated sparse codewords from database. The generated sparse codewords are only stored if and only if they are not already present in the database. The described process is represented in block diagram as shown in below fig.1.

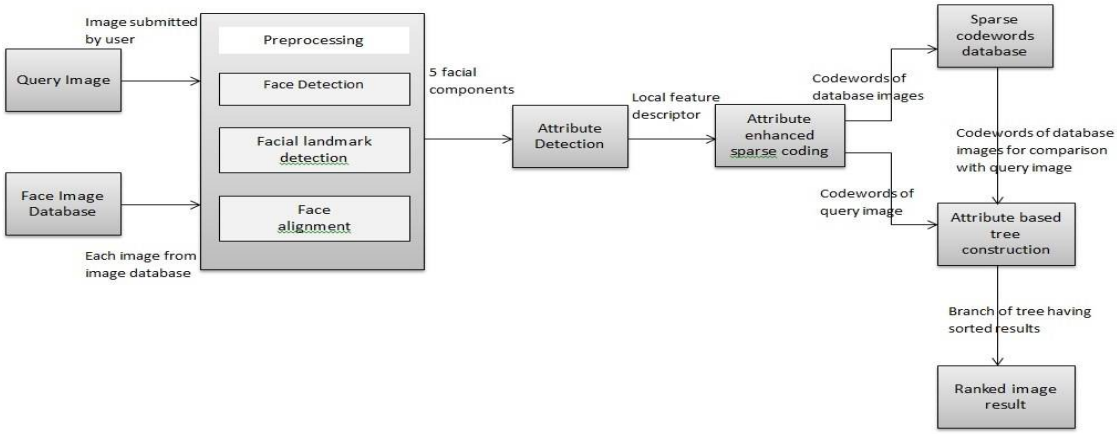


Fig 1. System Framework

**A) Preprocessing**

The Image Preprocessing module will process the image to find out facial attributes. For this, the image will be converted into grayscale as processing image in grayscale is fast and less memory consumption than processing colour image. After preprocessing face alignment and attribute detection will be performed. Face alignment is performed using position of eyes as eyes are detected well in a face.

To locate 68 different facial landmarks on the faces, Active shape model is applied. Using these facial landmarks, barycentre coordinate based mapping process is applied to align every face with the face mean shape. Proposed attribute-enhanced sparse coding method is applied to construct sparse codewords for that image. Query image will go through the same procedure to obtain sparse codewords and human attributes, and use these codewords to retrieve images from dataset. For efficient retrieval from dataset images, attribute based tree construction technique is used.

Attribute-enhanced sparse coding is used to find sparse codewords of database images globally in the offline stage. Codewords of the query image are combined locally with binary attribute signature to traverse the attribute-embedded inverted index in the online stage and derive real-time ranking results over database images [5].

Database images and query image will go through same procedure as shown in fig. 1. First step is image preprocessing. In pre-processing stage, first step is to find location of face from input image and then find components of face. Human attributes are detected from given face. After that, face alignment is done to extract low level features. From detected facial components local patches extracted and 9-dim LBP features computed. These are called local feature descriptors. After obtaining local low-level LBP features and attributes scores. To combine these features to semantically represent image.

**B) Attribute detection**

Before face alignment in preprocessing stage attribute detection is needed. Attribute detection framework is used to find human attributes from located face. For automatic attribute detection, attribute classifier is trained using various labeled images from internet which measures the presence, absence, or degree to which an attribute is expressed in Images.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

**C) Sparse codewords generation**

Sparse codewords are generated by attribute enhanced sparse coding. For detected five facial components including two eyes, nose tip, and two mouth corners, 7 * 5 grids, are extracted. Total 7 * 5 * 5 means 175 grids are extracted. From each grid, application can extract an image patch and compute a 59-dimensional uniform LBP feature descriptor as local feature. After obtaining local feature descriptors, application can quantize every descriptor into codewords using attribute enhanced sparse coding. The generated sparse codewords with an image index are stored in sparse codewords database. Every face photo image from database undergoes these two modules and sparse codewords are generated for each and every face photo image in database. Sparse codewords for each image with image index was stored in sparse codewords database. When an input query face photo image was submitted by user, the image undergoes same process as described above. After sparse codewords generation from query image, Attribute based tree construction is used to find out and retrieve most similar face photo images starting from higher matching to lower matching of face photo image [1].

For example, if an image has a positive female attribute score, it will use the first half of the dictionary centroids. If it has a negative female attribute score, it will use the second half of the dictionary centroids. With help of this, images with different attributes will definitely have different codewords. For considering multiple attributes, sparse representation is divided into multiple segments based on the number of attributes, and each segment of sparse representation is generated depending on single attribute. Attribute embedded inverted indexing creates inverted index structure by incorporating binary attribute signature in addition to the sparse codewords. Certain facial attributes are not correlated to human identity. So these attributes are degrading retrieval performance.

**D) Attribute based tree construction**

To build attribute based binary tree, the attributes of query image such that query image is considered as root of the tree. The images from image database having matching attributes are bind to right branch of the tree while unmatched images are ignored. The right branch of tree was get modified each time new image get bound. If image under bounding have more attributes than last node of branch then it get bound to parent of last node and if not then it get bind to last node. In this way of binding, along with finding similar face photo images from database for query, ranking of images was also performed. Ranking of resultant images is in order from highest matching to lowest. Since, finding and ranking of images get performed in single process by ignoring matching images, it works fast as compared to existing method of attribute embedded inverted index based search. It is used to retrieve the most scalable image database against submitted query image. Before proceeding to scalable face image retrieval for query image, training and sparse codewords generation from image datbabse is performed.

**ALGORITHM :**

PCA algorithm are used to identify the query image to database image. following steps are used in PCA algorithm.

1) Convert the face images in training set to face vectors
2) Normalize the face vectors.
   a) Calculate the average face vector.
   b) Subtract the average face vector from each face vector

   Calculate Eigen vectors?

   We have to calculate the covariance vector to get the Eigen vectors

   $C = A.A^T$

   $A = N^2 X M$ (m means total no of images in the database)

   $C = N^2 X N^2$

   If the image size is 50 then $N^2$ 50X50= 2500
3) Reduce the dimensionality of the training set.
4) Calculate the eigenvectors from covariance matrix c.
5) Select the k best eigenfaces such that k < m and can represent the whole training set.
6) Convert the lower dimensionality k eigenvectors to original face dimensionality.
7) Represent each face image a linear combination of all k eigenvectors.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

## IV. EXPERIMENTAL RESULT

**Retrieved Image Experiments**

As shown in the below table 1, the standard images are in the size of 250 X 250 and all these images are given as input to the query images. These Images is called as a query face image. It shows the query images from the database. As a result of query, the matching images are retrieved according to their nearest matching scores. The result image is obtaining from the retrieved images. Attribute embedded inverted indexing technique shows retrieve images less than the images available in the database. And attribute tree construction technique shows retrieve images equal to or approx matching images with available in the database. It means that the attribute tree construction technique is more accurate than the attribute embedded inverted indexing technique.

**For e.g. Person Name : Scott Peterson**

The total images of Scott Peterson in the database is 5 and during image retrieving from the database with attribute embedded inverted indexing technique images retrieved only 2 images. But same image retrieve from database with attribute based tree construction technique images retrieved 4 images.

Accuracy is calculated by following Formula:

Person name = (Retrieved images/Total no of images)*100

| Person Name | Total no of images in db | Attribute embedded in % | Tree construction in % |
|---|---|---|---|
| Scott Peterson | 5 | 40 | 80 |
| Roman Polanski | 6 | 33 | 67 |
| Robe Lowe | 4 | 50 | 100 |
| Robe Marshall | 6 | 67 | 100 |
| Rick Pitino | 4 | 25 | 75 |
| Rachel Hunter | 4 | 50 | 100 |
| Princess Caroline | 5 | 60 | 100 |

Table 1. Comparison between the retrieved images in %.

In the below graph retrieved images are converted into the percentile. The above Table 1 and Graph shows that attribute embedded inverted indexing technique is less retrieved images from the database as compared to the attribute based tree construction technique.

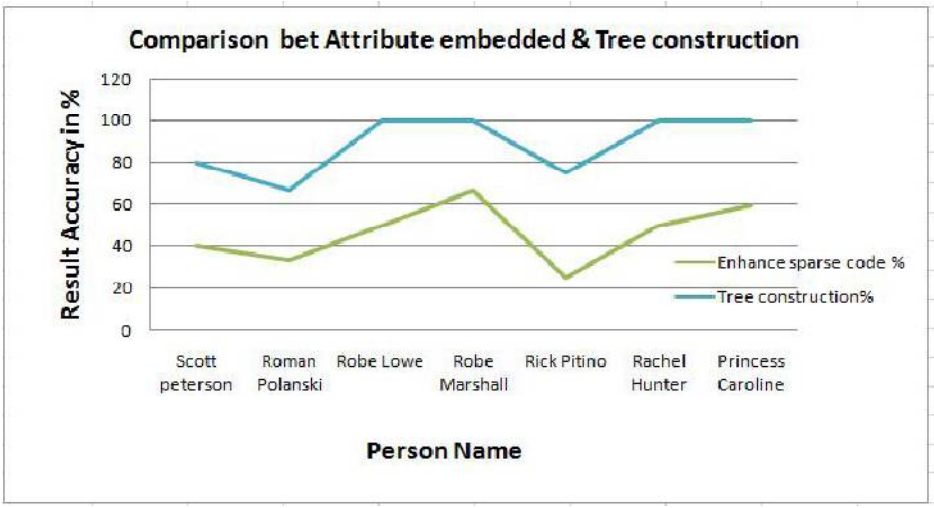


Fig. 2. Comparison between the retrieved images in %.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 8, August 2016**

## V. CONCLUSION AND FUTURE WORK

The attribute based tree construction system utilizes the human visual facial attributes like eye pairs, mouth, etc. to generate semantic key words called as sparse codewords. Sparse codewords improves discriminative image representation. The proposed system first off all generates sparse codewords from available large scale image database and stores it. This sparse codewords are then used to construct a sparse codewords attribute based binary tree for fast and efficient retrieval of query image search result from large scale image database. The attribute based binary tree construction ensures efficient retrieval in the online stage.

Thus attribute based tree construction system having a new perspective on content based face image retrieval by incorporating high-level human attributes into face image representation and image search on attribute based tree construction to improve search performance.

## VI. ACKNOWLEDGMENT

I would like to express my gratitude towards Mrs. PRACHI SORTE for their persistence guidance throughout the project. I would like to thank Mr. D. A. Meshram for their constructive criticism and Mrs. Sweta Kale for their valuable support.

## REFERENCES

1. Chen, B. C., Chen, Y. Y., Kuo, H., and Winston, H. Hsu., "Scalable Face Image Retrieval Using Attribute Enhanced Sparse Codewords", IEEE Transactions on Multimedia, vol. 15,pp. 99 ,Aug 2013.
2. Wang, D., Hoi, S. C., He, Y., and Zhu, J., "Retrieval based face annotation by weak label regularized local coordinate coding," ACM Multimedia, pp. 351-362, 2011.
3. Wu, Z., Ke, Q., Sun, J. and Shum, H.-Y., "Scalable face image retrieval with identity-based and multi reference re ranking," IEEE Conference on Computer Vision and Pattern Recognition, 2010.
4. Siddiquie, B. R., Feris, S., and Davis, L. S., "Image ranking and retrieval based on multi-attribute queries," IEEE Conference on Computer Vision and Pattern Recognition, pp.801-808, 2011.
5. Scheirer, W., Kumar, N., Belhumeur, P., and Boult, T., "Multi-Attribute Spaces: Calibration for Attribute Fusion and Similarity Search," IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2012.
6. Wu ,L., . Hoi, S. C., Hand N. Yu, "Semantics-preserving bag-of-words models and applications," Journal of IEEE Transactions on image processing, VOL. 1, NO. 1, 2010.
7. Lowe, D., "Distinctive image features from scale-invariant keypoints," international Journal of Computer Vision, 2003.
8. Chum, O., Philbin, J., Sivic, J., Isard , M., and A. Zisserman, "Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval," IEEE International Conference on Computer Vision, 2007