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

Efficient Image Retrieval Using Sparse Codewords

Pooja V. Baravkar[1], Gayatri M. Bhandari[2]

[1]P.G. Student, Savitribai Phule Pune University, Department of Computer Engineering, BSIOTR, Wagholi, Pune, Maharashtra, India

[2]Professor, Savitribai Phule Pune University, Department of Computer Engineering, BSIOTR, Wagholi, Pune, Maharashtra, India

Abstract: ***Social network services such as facebook, twitter provide photo sharing facilities. Among all those photos majority photos contain human faces. People are interested to discover contents in images. Human beings are involved in uploading images, searching images in it. Hence with the exponentially growing photos, large-scale content-based face image retrieval is a facilitating technology for many emerging applications. Image retrieval is done using sparse codewords method. It is possible to build the system which automatically detects human attributes. Automatically detected human attributes are used to increase the performance of content based face image retrieval. Low level features of image are combined with high level features of image to acquire more efficient results of image retrieval. Semantic codewords are used for face image retrieval. Sparse codewords are constructed using semantic cues of the face image to have enhanced content based face image retrieval.***

Keywords: Sparse codewords, Face Image, Human Attributes

1. Introduction

Social network has revolutionized the way people distribute and access information. In communities there are many social sharing websites, which allows user to share photos. Users can interact with each other by using communities. There are large numbers of images present on internet. The popularity of social networks like facebook, twitter is mostly used by the people. Many of these social media uses human face images as their profile. Maximum of the user use the celebrities image. Due to that more than two persons have same image as their profile. It can be maintain by making use of large scale database for the image storage. An image retrieval system is a computer system which is used for browsing, searching and retrieving images from a large database of digital images. Photo search by face positions and facial attributes on the image characteristics are represented in the outline. Common attributes are used for scalable image retrieval.

In retrieval based face annotation technique is having two challenges. The first challenge is how to efficiently extract most similar facial images from a large facial image database. This typically relies on effective content based facial image retrieval. The second challenge is how to efficiently exploit the short list of candidate facial images. Images having label information for the face name annotation task. The associated labels of web facial images are often noisy and incomplete due to the environment of web images. Existing systems use the content based image retrieval method to find images. In existing systems user is unable to automatically detect the human attributes. An image is taken as input. Retrieval method includes searching similar image from the large scale image database. Image retrieval can also be done using query based image retrieval technique. Query based image retrieval is used to display the similar images. Before storing the image in the database each image is given index number. By using the index the features can be extracted. Traditional image retrieval is done by using low level characteristics to retrieve the image.

Traditional methods of image retrieval are lack of semantic meanings. Most traditional and common methods of image retrieval use some method of adding metadata such as captioning, keywords, or descriptions to the images. Due to this that retrieval can be performed over the annotation words. Manual image annotation is time consuming, and expensive; to address this, there has been a large amount of research done on automatic image annotation. Additionally, the increase in social web applications and the semantic web have motivated the development of several web-based image annotation tools. Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision techniques to the image retrieval problem. "Content-based" means that the search analyzes the contents of the image slightly than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this perspective might refer to colors, shapes, textures, or any other information that can be derived from the image itself. CBIR is desirable because most web-based image search engines rely purely on metadata. It produces a lot of garbage in the results. Thus a system that can filter images based on their content would present better indexing and return more accurate results.

To improve content based face image retrieval method namely attribute enhanced sparse coding is proposed. Image retrieval using sparse codewords method it is used in techniques such as automatic face annotation, crime investigation etc. Figure 1 shows the difference between low level features and high level features. Low level features are used to represent faces but these features are lack of semantic meanings. Face image contains rich context information such as hair colour, skin colour, gender, wearing glasswear.

High level semantics are described by using automatic attribute detection. High level features represent facial features as gender, race, hair colour, etc. It is possible to build the system that is able to retrieve images efficiently from large databaseSemantic meanings of facial images such as gender, hair, color, race. Low level features can be extracted automatically from an facial image. These features

can be extracted without considering any shape information. High-level feature extraction it concerns with finding shapes and objects in computer images. Shapes can be used for image retrieval. High level features of an image such as eyes, nose, ear are used for face image retrieval. To find these features, user can use their shape for image detection. Feature detection can be done by considering points such as the white part of the eyes are ellipsoidal; the mouth can appear as two lines. High level features can be viewed as objects. High level features use to the low-level features to define collections of points. Images or pictures with high complexity can be decomposed into arrangement of simple shapes. Facial image analysis in which user is able to find the eyes above the nose and expect to find the mouth below the nose.

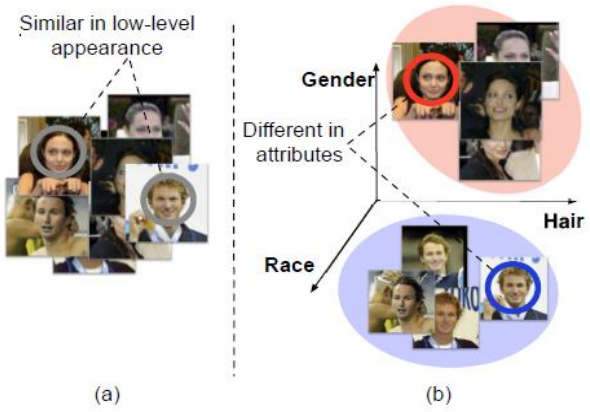


Figure 1: Difference between low level features and high level features [1].

2. Literature Review

Several studies have been contributed to make image retrieval efficient reduce and to increase the performance retrieving result.

2.1 Sparse Codewords

Sparse codewords has been used in many information retrieval and data mining tasks, such as image retrieval and object recognition. Sparse representation exploits the global structure of feature space. It uses several important human attributes combined with low-level features to build semantic codewords in the offline stage. Automatically detected high level human attributes and low level features are used to construct semantic codewords.

2.2 Scalable Face Image Retrieval with Identity-Based Quantization and Multi-reference Re-ranking

Face image is taken as a query aim is to retrieve images containing faces of the same person . Faces appearing in the query image from a web-scale image database containing tens of millions face images [2]. By using a novel face image representation by means of both local and global features. First, locate component-based local features that encode geometric constraints [2]. Using identity based quantization that can better hold intra class variation. Using multiple examples from multiple identities.

2.3 Semisupervised Face Image Retrieval With Identity Constraint

The face image retrieval is used to discover the ranking result. From most to least similar face images in a face image database. Such has many applications in different areas. For instance, when it applies on personal multimedia, it can enable automatic face tagging and face image clustering [6]. When it is applied in forensics, it can help with crime investigation. Face retrieval task is closely related to face recognition task. The descriptor proposed is efficient and has superior performance. The dissimilarity between face recognition and face retrieval is that face recognition task requires completely labeled data in the training set. Sparse coding with identity constraint framework and it is combined with inverted indexing results in an efficient retrieval framework [6]. Sparse coding with identity constraint is used to retrieve similar images with the same identity.

2.4 Boosting Sex Identification Performance

This method is intended to identify the sex of a person from a low resolution grayscale image of their face. Boosting sex identification performance is exploited to create an efficient system. This method is simple to implement and maintain a learning framework for image retrieval. The queries for people vary from specific individuals, such as celebrities, actors, musicians to general queries such as adult-content and stock-photography image [2]. AdaBoost works by choosing and combining weak classifiers together. Weak classifiers are used to form a more accurate strong classifier. The weak classifiers used to distinguish pose, expression. A weak classifier gives have accuracy slightly better than random chance [2]. Numerically outputs of these are represented as 1 and 0 respectively. Even accounting for symmetries it considers extremely large number of classifiers.

2.4 Describable Visual Attributes For Face Verification and Image Search

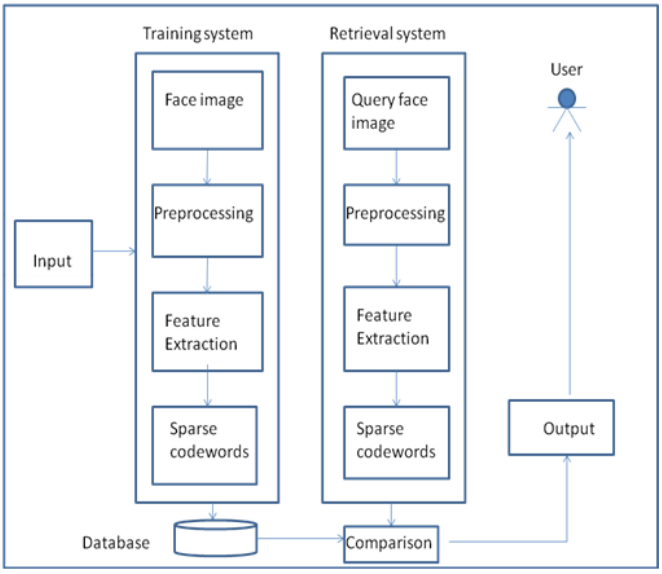
N. Kumar et al., [7] proposed describable visual attributes for face verification. This method uses an extensive vocabulary of visual attributes. It is used to label a large data set of images. This dataset is then used to train classifiers that automatically recognize the presence, absence, or degree to which these attributes are in new images. Describable visual attributes are one of the most natural ways of describing faces. A person look can also be described in terms of the similarity of a part of their face to the same part of another individual [7]. The classifier output can then be used to recognize faces and search through large image collections. Any user can learn a set of attributes from large image collection.

3. System Implementation

3.1 System Architecture

The Figure 2 shows the overall system architecture. As per shown in figure the process flow goes according to the

system architecture. The system architecture includes following components



1) Input
In the process of facial image retrieval input is taken from the user. User can give input in the form of celebrities image, actors images.

2) Training System
In the process of facial image retrieval at first face image is detected. Once facial image is detected on that image cropping operation is performed. Preprocessing phase includes various operations on facial image such as grayscale function, edge detection on image. Grayscale is used for an image in which the value of each pixel is a single sample. Edge detection is used for detecting edges, corner points of facial images. For edge detection sobel algorithm is used. Feature extraction is perfomed on facial images. Features such as male, hair, race etc. are extracted from facial image. Extracted features are used for feature encoding. Sparse codewords are constructed using semantic codewords.

3) Retrieval System
Input for retrieval process is taken through query processing. User can give an input image and search for similar images in database. Preprocessing phase of retrieval system includes various operations on facial image such as grayscale function, edge detection on image. Grayscale is used for an image in which the value of each pixel is a single sample. Edge detection is used for detecting edges, corner points of facial images. Feature extraction is also performed in retrieval system. Features such as male, hair, race etc. are extracted from facial image.

Euclidean distance is used for comparison between stored image and processed query image. Euclidean distance is average distance between two points. Through this distance, Euclidean space becomes a metric space.

4) Output
Output is given to user on the basis of query face image is retrieved or not.

3.2 Proposed System

The proposed system is built using this equation. It describes how sparse coding is used for face image retrieval. Different codewords are finding out by applying the same procedures to all regions in a single image [1]. All these codewords are combined together to represent the face image. Sparse coding is used for facial image retrieval by solving the following optimization problem:

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

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

Where x^(i) is the original feature of image extracted from a facial image , D∈ $R^{\wedge}(d * K)$ is to be learned dictionary. V=[v^(1),v^(2),...,v^(n)] it is the sparse representation of the image patches.

The sobel edge detection algorithm is used for edge detection. Spatial edge detection filter that discover edges by finding the gradient of an image. This means we are finding a remarkable difference between the value of pixels. The input image is first converted to grayscaled image.

- Traverse through entire image.
- For each pixel in the image. A window of 3*3 pixel is taken and multiply it the given template for matrix.
- Then we have calculated the Sobel gradient G using formula

$$\mathbf{G} = \sqrt{\mathbf{G}_x{}^2 + \mathbf{G}_y{}^2}$$

Template is given as

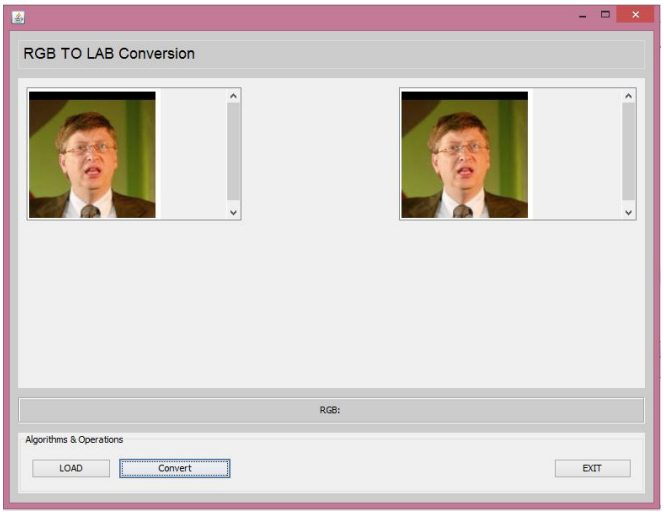
-1 0 1 1 2 1
-2 0 2 0 0 0
-1 0 1 -1 -2 -1

Apply the templates to a 3*3 filter window.
a1 a2 a3
a4 a5 a6
a7 a8 a9
where a1 .. a9 are grey levels of each pixel in the filter window.
X = -1*a1 + 1*a3 - 2*a4 + 2*a6 - 1*a7 + 1*a9
Y = 1*a1 + 2*a2 + 1*a3 - 1*a7 - 2*a8 - 1*a9
Sobel Gradient = sqrt(X*X + Y*Y)
Sobel gradient is calculated using this formula.

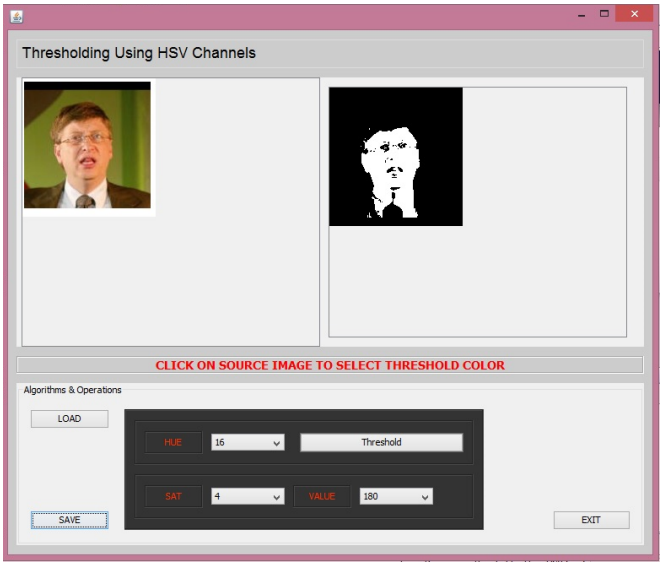
4. Results

The system is developed by using JAVA (Version JDK). The development tool used is NetBeans for desktop application. The experiments are performed on Core2Duo Intel processor 2 GB RAM. The results are shown as the report whether the query image is retrieved or not.

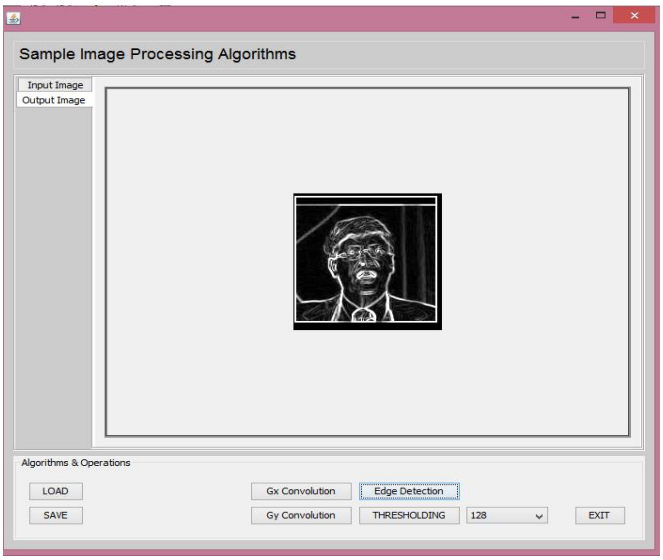
4.1 Facial Image conversion from RGB To LAB



4.2 Result of Thresholding using HSV Channels



4.3 Result of Edge Detection



Features such as gender, lips, beard or not etc are found. Attribute weight is assigned to each attribute. Sparse codewords are constructed for these features. Desired image is retrieved from collection of images utilizing sparse codewords.

Table 1: Retrieved Objects

Dataset Name	*Actual Objects*	*Retrieved Objects*	*Correct Retrieved Objects*
Sparse codewords system A A Apps	20	18	17
Existing System	25	24	

Table 2: Total Accuracy

Dataset Name	*Precision*	*Recall*
Sparse codewords system	0.94444444	0.85
Existing Systems	0.95833333	0.92
Total	0.951388889	0.885
Accuracy percentage	0.885	-

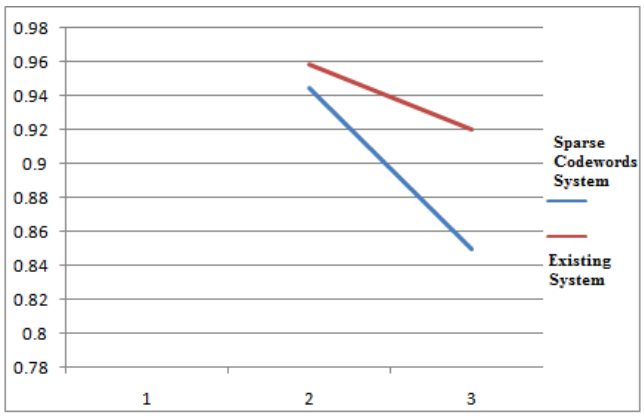


Figure 1: Accuracy Plotted

Table 1 shows the correct retrieved objects i.e. Using sparse codewords how many images are retrieved.

Table 2 shows the accuracy plotted in terms of Precision and recall.

5. Conclusion

Attribute enhanced sparse codewords exploit automatically detected human attributes to construct semantic codewords. Sparse codewords are stored in binary format. It requires low storage space to store. Efficient image retrieval with sparse codewords discovers informative attributes for face retrieval from dataset. In sparse codewords method attribute weight is used. Each attribute is given a specific weight.

As a future work sparse codewords technique can be used for object detection. There is a scope to retrieve specific part of object for detailed devise.

6. Acknowledgement

I would like to express my sincere gratitude to my guide and Head of Computer Engineering Prof. G. M. Bhandari for her continuous support, patience, motivation, enthusiasm, and

immense knowledge. Her guidance helped me in all the time of research and writing of this paper.

References

[1] Bor-Chun Chen, Yan-Ying Chen, Yin-HsiKuo, Winston H. Hsu "Scalable face image retrieval using attribute enhanced sparse codewords", IEEE TRANSACTIONS ON MULTIMEDIA VOL:PP NO:99 YEAR 2013.
[2] Shumeet Baluja and H. Rowley,"Boosting sex identification performance", Int. J. Comput.Vision, 2007.
[3] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," ICML, 2009.
[4] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, Robust face recognition via sparse representation", IEEE Trans. Pattern Anal.Mach. Intell., vol. 31, no. 2, pp. 210-.227, Feb. 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 in Proc. IEEE Conf. Computer Vision and Pattern Recognit., 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", in Proc. ACM Multimedia, 2011.
[7] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Describable visual attributes for face verification and image search", IEEE Trans. Pattern Pattern Anal. Mach. Intell., Special Issue on Real-World Face Recognition,vol. 33, no. 10, pp. 1962–1977, Oct. 2011.

Author Profile

Ms. Pooja V. Baravkar received the Bachelors degree (B.E.) Computer Engineering in 2013 from SVPMCOE, Baramati. She is now pursuing Masters degree (Computer Engineering), from BSIOTR, Wagholi, Pune, Maharashtra.

Prof. Gayatri M. Bhandari received her M.Tech(Computer Engineering) and now pursuing Ph.D. She is also working as Head Of Computer Engineering Department, Bhivarabai Sawant Institute of Technology & Research, Wagholi, Pune, Maharashtra. Her research areas Signal Processing, Computer Network