2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT]

Survey on sketch based image retrieval methods

M Indu
M.Tech Student
Department of Computer Science and Engineering
SCT College of Engineering
Trivandrum, India
indu22.emp@gmail.com

Kavitha K V
Assistant Professor
Department of Computer Science and Engineering
SCT College of Engineering
Trivandrum, India
kavitha279@yahoo.co.in

***Abstract*— In many applications critical roles are played by face sketch-photo synthesis, such as law enforcement and digital entertainment. Several image based retrieval systems are text based, content based and sketch based. Image retrieval methods which use sketch content as input are referred to as Sketch Based Image Retrieval systems. In several ways an image can be retrieved from the database using user queries as input. One of the efficient methods and popular methods for retrieval is as Sketch Based Image Retrieval which is not necessary to have a high skill to draw the query sketch. Paper reviews the various sketch based image retrieval methods used in image processing and a comparison of all these methods is also done. The retrieval system using sketches can be effective and essential in our day to day life such as Medical diagnosis, digital library, search engines, crime prevention, photo sharing sites, geographical information, and sensing remote systems.**

Keywords— sketch-photo; TBIR; CBIR; SBIR; patch; dictionaries; sparse representation.

I. INTRODUCTION

Manipulation of digital images through a digital computer is familiarized as digital image processing [1]. Its a subfield of signals and systems but focus primarily on images. DIP focuses on developing a system to perform processing on an image. The input is a digital image and that image is processed using efficient algorithms to produce processed output image. example: Adobe Photoshop. On applying signal processing methods, image is treated as a two dimensional signal by the image processing system. A core research area for image processing comes in fields of computer science, business, and engineering disciplines.

Data is represented in two dimension. Digital image contains pixels arranged in rectangular array with a certain width & height. For representing brightness of an image, the content of each pixel is represented by one or more bits. Image processing can be used to solve identification problems, such as forensic medicine or creating weather maps from satellite such as refining a picture from video source.

Image processing basically has three steps. First the image is imported with optical scanner or using digital photography. Analyzing and manipulating the image includes compression of data, enhancement of image and spotting the patterns that are not visible to human eyes like satellite photographs. Finally the output digital image can be an altered image or report based on image analysis. There are various purposes for image processing which can be categorised into; visualization -Observe the objects that are not visible, image sharpen and restoring - to create a better image, image retrieval - lookout for the image of interest, measurement of pattern – various objects in an image is measured, image recognition – distinguish the objects in an image.

Due to the various advancements in image technology in the field of computer science, image retrieval has become a popular area. Retrieval systems of images were used for retrieving, searching and browsing images from large databases of digital images [2]. A specialized data search is used to find images. To search for images, a user may provide query terms such as image file/link, keyword or click on some image, and the system will return images "similar" to the query. The search criteria similarities used could be meta tags, colour distribution in images, region/shape attributes, etc.

II. OBJECTIVE OF SKETCH BASED IMAGE RETRIEVAL METHODS

Face is a prime part of human body through which an individual can be identified. It plays a major role in bearing identity and emotions of an individual. Every individual has distinctive facial features other than identical twins.

Major applications of face recognition include security systems, verification of credit card, culprit identification etc. Automatic retrieval of photos of suspects from police culprit database help them to narrow down potential suspects quickly. In cases where photo image of the suspect is not available, the best substitute is to draw a sketch based on the recall of an eyewitness. Thus automatically searching through a photo database using a query sketch drawing is very useful. Thus helping the police to locate a group of potential suspects, and also help the witness and artist to revise the sketch drawing of the culprit interactively based on the similar photos retrieved. It can also be used in other fields where photo is not available but the details of the photo can be described. This method remarkably reduces the difference between photo and sketch.

For more than twenty years human face recognition has been studying. Since faces are complex, multi-dimensional visual stimuli developing a computational model of face recognition is very difficult. Face recognition is a computer vision task, in which many early vision techniques can be involved.

978-1-5090-1277-0/16/$31.00 ©2016 IEEE

III. IMAGE RETRIEVAL SYSTEMS

Since early 90's image retrieval has become a very attractive research area. Both commercial and research image retrieval systems have been built till date. Most image retrieval systems support any one of the following options:

A. Text-based image retrieval (TBIR) system

It uses traditional database techniques to manage images. Using the text associated with an image determines the image content. Example: Google, Yahoo Image Search engines. The search engines are fast and robust but at times fail to retrieve relevant images [3]. There are various disadvantages such as for large database manual annotation is impossible, Polysemy problem (more than one object can be referred by the same word) and also nearby text also will not describe the image.

B. Content based image retrieval (CBIR) system

This technique uses user's interest to search images from large image databases. It is an alternative to TBIR. To represent and index an image, various visual contents of an image is used such as color, shape, texture, and spatial layout. From the database images visual contents are extracted and portrayed using multi-dimensional feature vectors as illustrated in Fig.1 [4]. A feature database is formed by different feature vectors of the images in the database. To retrieve images, users provide example images or sketched figures. As an internal representation, the system changes these examples into its feature vectors. Using an indexing scheme the similarities or distances between the feature vectors of the sketch and the images in the database are calculated and performed the retrieval. For searching an image from the large database a systematic way is provided by indexing scheme. To produce more meaningful retrieval result, recent retrieval systems add user's relevance feedback to modify the retrieval process.

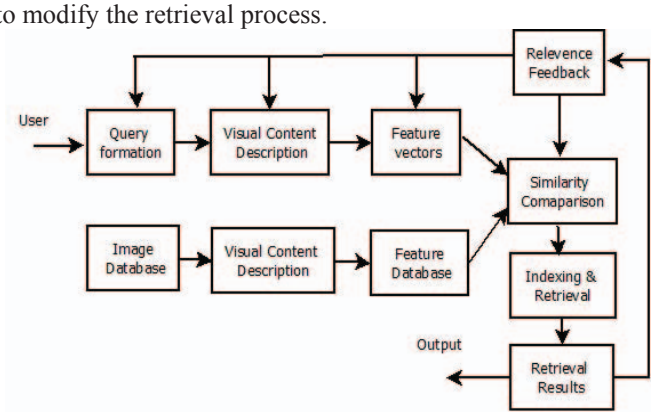

Fig.1: Content based image retrieval system.

CBIR system basically has four main sections in system realization: collection of data, builds up feature database, search in the database, orderly arrangement and deals with the result of retrieval [5]. Various features of CBIR system include automatic extraction of visual features such as color, shape information and texture of images.

C. Sketch based image retrieval (SBIR) system

It is a relevant means of querying large image databases [2]. Researches mainly focus on solving the gap between sketch and problems relating to image matching. In large scale database visual features SBIR technique is used to find query image. Using different types of descriptors many researchers tried to extract the features of images.

IV. REVIEW OF VARIOUS SKETCH BASED IMAGE RETRIEVAL METHODS

For most of the surveillance purposes such as search for wanted culprits, suspected terrorists, and missing children; face retrieval is a widely used biometric technique. Face is the key focus in the society which plays a crucial role in bearing an individual's identity and emotions. Facial recognition and retrieval system uses human faces to attempt and identify an individual or verify a person's claimed identity. It's a form of computer vision. Reduce the difference between two modalities is the key objective of this method. For making the recognition process easier, photo and sketch should be made to same mode. Various applications in the field of multimedia management, security, smart card, banking etc.

A. Principal component analysis (PCA)

In this architecture for the training phase conversion of database photo into sketch by using PCA algorithm is done primarily. That is faces from database is selected one by one and face image is extracted by removing their background details. Then face image is converted into sketch, by using PCA sketch and is then converted into Eigen Face by considering the intensity of image [6].

In the same way, average of all Eigen face is found out for the entire images in the database. Now convert artist sketch into Eigen face & mix it with average of all faces from databases. Then by comparing, two faces with maximum match will give the output. Here sketch is converted into Eigen face that is then compared with the sketch generated from photo which is obtained from the criminal record of police. The same is illustrated in Fig. 2.

The face region is divided sequentially into overlapping patches. During synthesis of sketch, a photo patch from the face to be synthesized, a similar photo patch from the training set is found. Using its corresponding sketch patch in the training set, estimates the sketch patch to be synthesized. Assumes if two photo patches are similar, their sketch patches also will be similar.

In the initial PCA method it is assumed that the source sketch and the target face photo share the same projection coefficient. Using PCA procedure coefficients are obtained. There are basically two types of photo input: target photo and training photos .The coefficients are obtained by projecting the input photos on the target photos. The target sketch is obtained by the linear combination of the training weights provided by the projection coefficients

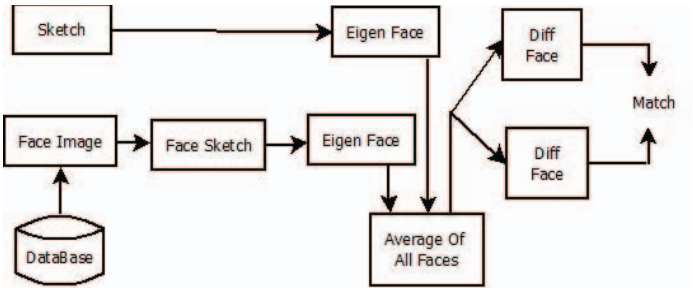

Fig.2: Face recognition with eigen face.

This is further improved by separating shape from the texture. Here the target sketch is obtained by fusing the shape and texture using the Eigen transformation.

B. Multiscale Markov random field (MRF)

To synthesize local face structures in different scales this approach is used [7]. Here training set consists of photo-sketch pairs. The faces to be studied should be in frontal pose, with a normal lighting, neutral expression and without occlusion. As learning global face structure is complicated, simple structured local patches are considered. A Face region is sequentially divided into overlapping patches. During sketch synthesis, to process a photo patch of the face, similar photo patch from the training set is need to be found out and its corresponding sketch patch is generated. The assumption is, if two patches are similar then their corresponding sketch patch will also be similar. In addition, smoothness is required between adjacent patches of the synthesized sketch. The scale of face structure is decided by its patch size. To learn face structures at different scales MRF model is widely used. All face photos in the gallery are transformed to sketches using sketch synthesis algorithms and to the synthesized sketches a query sketch is matched. A query sketch is transformed to a photo and the synthesized photo is then matched with the existing real photos in the gallery.

This approach is based on local patches. It does not require a face photo constructed by PCA from training set. It can construct more complicated face structures such as hair. A sketch patch not only requires information from neighbouring patches but also from patches that are far away by belief propagation [8]. It can synthesize sketch patches over different scales. Here sketch synthesis algorithm may not work properly if the input photo is taken under different conditions than the photos in the training set.

C. Local linear embedding (LLE)

There are various problem encountering if a photo image is recovered from a sketch [9]. For a sketch image often main facial features, and many detailed features are ignored. A different approach is to generate a pseudo-sketch from a photo image which is being used in this scheme. To setting up a mapping relation between a photo and a sketch the method of pseudo-sketch generation is equivalent. Focus of this method is on sketches of plain style. It is inevitable to have some distortions when sketches are drawn by artists. Additionally, the weight combination of local nearest neighbours brings some blurs to the pseudo-sketches. LLE is a learning method. It is widely used for nonlinear dimension reduction of high-dimensional data and image analysis [10], [11], [12]. The fundamental idea of LLE is to compute neighbour-preserving mapping between a high-dimensional original data space and a low-dimensional feature space, on the basis of simple geometric intuition that each data and its neighbours lie on or close to a locally patch.

A patch-based strategy is used and dividing the photo and sketch images into N small overlapping image patches. For each photo image patch first fit it with its K nearest neighbours from training samples and calculate the reconstruction weights. Then its corresponding sketch patch can be estimated from training sketch samples by preserving the local geometry. Face sketch recognition is used to measure similarities between probe sketch and pseudo-sketches from photos. In sketch recognition, for better describing nonlinear variations caused by distortions and blurs in the real sketches drawn by artists and pseudo-sketches, KNDA based nonlinear discriminative classifier is used. A nonlinear version of LDA is KNDA. It is widely used in pattern recognition [13], [14], [15].

When comparing to PCA the experimental results show that the recognition accuracy and performance of KNDA is better. PCA has the worst performance, but it is optimal for reconstruction but not for discriminating one class from the other.

D. Multi dictionary sparse representation framework

Sketches or photos generated using LLE method produces lower definition and blurred outputs that consequently reduces the visual quality and recognition rate across the heterogeneous images. So in order to improve the quality of the synthesized images a novel multi dictionary based sparse based sketch-photo synthesis model is constructed.

This method also consists of N sketch-photo pairs as training set. Since the method works at several patch level, each sketch and photo is partitioned into even patches with some overlapping. Using clustering method such as k means approach, these patches are clustered into classes. The sketch patch intensity subtracting the mean intensity of this patch acts as the feature for clustering, this determines the cluster to which that patch has to be clustered. K nearest neighbours is found with its corresponding optimized weights [16]. Fuse these synthesized patches into an initial sketch with overlapping areas averaged. Here two feature dictionary of sketch path and photo patch is there. A joint learning strategy is applied to each cluster for two dictionaries [17]. For any testing photo patch a sparse representation coefficient vector is computed by projecting the feature of patch to the photo patch feature dictionary learned from the cluster. The features are used in the initial estimate stage and image enhancement stage was different, clustering algorithm is used twice for two stages which improve the results.

From the experimental results it has got that, this method has higher definition and richer detail information. For a sketch of size 64x64 it takes about 2.5minutes to synthesis a sketch running on a 3GHz CPU computer. Sketch-photo recognition rate is better when compared to other methods.

The mean opinion score (MOS) in the field of multimedia is also higher than the other methods.

V. Comparison Of Various Sketch Based Image Retrieval Methods

In this review paper various image retrieval methods based on query sketch is discussed and a comparison of these methods are discussed in TABLEI.

TABLE I. Comparison of Various retrieval Methods

Method	Feature/ Methodology	Advantages	Disadvantages
PCA	Mapping between a photo and its corresponding sketch via Linear transformation.	-Simple and efficient -Low dimensional subspace representation.	Very sensitive to scale, so a low-level pre-processing is needed.
Multiscale MRF	The relationship of adjacent local patches and the face structures at different scales.	Boundaries can be identified easily.	Unsupervised parameter estimation of the MRF is difficult.
LLE	Nonlinear process of face sketch synthesis.	Relationship between face photos and face-sketches can be easily estimated.	Need more training samples.
Multi dictionary sparse representation framework	Using LLE method and multi-dictionary sparse representation model.	- Enhance the quality of the initial image -Sketch-photo recognition rate is better.	Multiple sketches cannot be retrieved.

VI. Conclusion

A study of various image retrieval work is done in this survey. A wide variety of researches is taking place in the field of image retrieval. Each work has its own technique, contribution and limitations. As a review paper, it cannot incorporate each and every detail of individual works, however this paper attempts to deal with an elaborate review of the most usual traditional and modern image retrieval systems. This review mainly focuses on the methods or approaches to come up with an efficient retrieval system together with the limitations or challenges.

REFERENCES

[1] http://wwwelec.inaoep.mx/~jmram/Digital_Image_Processing_GONZALEZ.pdf.

[2] Huda Abdulaali Abdulbaqi, Ghazali Sulong,Soukaena Hassan Hashem, "A sketch based image retrieval: a review of literature," Journal of Theoretical and Applied Information Technology, 2014.

[3] GulfishanFirdose Ahmed, RajuBarskar,"A Study on Different Image Retrieval Techniques in Image Processing," International Journal of soft Computing and Engineering (IJSCE), ISSN: 2231-2307,Volume-1, Issue-4, September 2011.

[4] Dr. Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng "Fundamentals of content-based image retrieval" Springer,2003.

[5] NidhiSinghai, Prof. Shishir K. Shandilya,"A Survey On: Content Based Image Retrieval Systems," International Journal of Computer Applications (0975 – 8887) Volume 4 – No.2, July 2010.

[6] Amit R. Sharma and Prakash. R. Devale, "An application to human face photo-sketch synthesis and recognition," International Journal of Advances in Engineering & Technology, May 2012.

[7] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 11, pp. 1955–1967, Nov. 2009.

[8] J.S. Yedidia, W.T. Freeman, and Y. Weiss, "Understanding Belief Propagation and Its Generalizations," Exploring Artificial Intelligence in the New Millennium, Morgan Kaufmann, 2003.

[9] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A nonlinear approach for face sketch synthesis and recognition," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Jun. 2005, pp. 1005–1010.

[10] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding", Science, 290(5500), pp.2323-2326, 2000.

[11] D. Ridder and R. Duin, "Locally Linear Embedding For Classification, Tech-report", PH-2002-01, Imaging Science & Technology Department, Delft University of Technology, Netherlands.

[12] H. Chang, D. Yeung, and Y. Xiong, "Super- Resolution Through Neigbhor Embedding", Proc. of. Int. Conf. Computer Vision and Pattern Recognition, 2004.

[13] Q. S. Liu, R. Huang, H. Q. Lu, S. D. Ma, "Face Recognition Using Kernel Based Fisher Discriminant Analysis", Proc. of. Int. Conf. Automatic Face and Gesture Recognition, 2002.

[14] M. H. Yang, Kernel Eigenfaces vs. Kernel Fisherfaces: "Face Recognition Using Kernel Methods", Proc. of. Int. Conf. Automatic Face and Gesture Recognition, pp. 215-220, 2002.

[15] Y. Wu, T. S. Huang, and K. Toyama, "Self- Supervised Learning for Object based on Kernel Discriminant-EM Algorithm", Proc. of Int. Conf. Computer Vision, 2001.

[16] Nannan Wang, Xinbo Gao, "Face sketch photo synthesis under multi dictionary sparse representation framework", sixth international Conf. on image ang graphics,2011.

[17] J. Yang, J. Wright, T. Huang and Y. Ma," Image super resolution via sparse representation," IEEE Trans. Image processing, 2010, vol. 19, no. 11,pp.2861-2873.