

# SUBSPACE MAPPING OF FACIAL ATTRIBUTES FOR FACE IMAGE RETRIEVAL

V.Padmini<sup>1</sup> and S.Suchitra<sup>2</sup>

**Abstract-** In the growing advancements of recent multimedia applications, the domain of Face images and its corresponding applications have retained its great importance. The wide array of social, real-time and crime investigation applications and their expanding field of necessities demands a far more accurate and fast face image retrieval system. Most face image retrieval techniques depend on only low-level cues to compare the face images of the people based on similar their visual identity. In this paper a face image retrieval system driven by enabled facial attributes is proposed to efficiently utilise the facial attributes information as high level semantic cues to discriminate people's identities by a semi-supervised approach of mapping the subspace labels along with the extracted low level components in the run time environment. An inverted indexing Scheme is used for the storing and retrieving face image representations. Our aim is minimise the possible marginal errors of inter-class and intra-class variations detected under different illumination conditions during the similarity search. We demonstrate relative improvement of 87% in MAP by experimenting and comparing our proposed method and other state of the art methods in LFW and PublicFig datasets.

**Keywords –** Face image, Facial attributes, High level cues, Similarity Search, Image Retrieval

## I. INTRODUCTION

Several research works had been endorsed in the last decade to develop image retrieval techniques from the multimedia databases. Although large number of indexing and retrieval techniques has been developed, there is no universally conventional feature extraction, indexing and retrieval technique obtainable. Among the vast digital images and photos shared on the internet, a big percentage of them are photos related to human faces. Many research problems and opportunities for a variety of real-world applications are based upon the face images. This has created an overriding need to provide proficient means of face image retrieval for many real-world applications.

The designing aspect of efficient face image retrieval should take account of the issues such as noise, illumination, poses and expressions, low cost of processing and time management along with elimination of misclassified face images. The Extraction of low level components by normalising the appearance of face, selecting an suitable feature description and designing matching methodologies involves the challenge of incorporation of colour, texture and shape properties of the images.

Accurate retrieval remains a difficult task for face images and is challenging due to the huge intra-class variation of same person, which is bigger than the inter-class variation between different persons. In previous works on face image retrieval, faces are typically described as low level image features and are sensitive to intra-class variations. Directly using high level attributes for retrieval, is also not precise since each individual attribute contains very limited information and is not quite enough to retrieve corresponding faces of a query subject. To improve the discriminative abilities of low level features and facial attributes, we propose a face image retrieval framework driven by enabled facial attributes.

---

<sup>1</sup> *Department of Computer Science and Engineering Easwari Engineering College, Chennai, Tamil Nadu, India*

<sup>2</sup> *Department of Computer Science and Engineering Easwari Engineering College, Chennai, Tamil Nadu, India*

The rest of the paper is organised as follows. Section II contains related works. Section III contains the proposed Face image retrieval System driven by enabled facial attributes explained in detail. Section IV includes the Experimental Column and results. Finally Section V concludes this paper and points out future work direction.

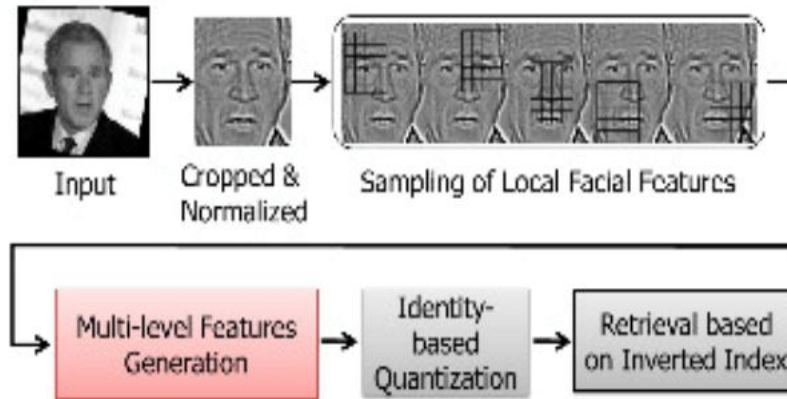


Figure 1.1 Face Image Retrieval System

## II. RELATED WORKS

For image retrieval, low-level features have been adopted in early work. Local Binary Pattern (LBP) is a powerful feature extractor in which histograms are extracted and articulated into a single, spatially enhanced feature histogram representing face image proposed by Timo Ahonen et. al[9]. The LBP cannot adequately deal with the presence of variations in the faces. This problem was addressed with the development of the local ternary pattern (LTP) by introducing a third agent and was followed by variations such as LTrP and LOP. Bayesian age-difference classifier is built over a probabilistic eigenspaces framework and the work of Active Shape model to locate features in frontal views of upright faces further extensions of the Facial descriptor techniques in Low level Features. An approach to combine multiple attributes for multiple attribute query search was investigated by Walter J. Scheirer et. al [10]. Here Normalized multi-attribute space is constructed from raw classifier outputs. Chen-Kuo Chiang et. al [3] proposed a multi-attributed dictionary learning algorithm for sparse coding by jointly incorporating data and attribute similarities. Bor-Chun Chen et. al [2] proposed to use human attributes as high-level semantic cues to determine people's identities. Discriminative image features with attribute information encoded to achieve more accurate face image retrieval is developed. A Binary encoding Scheme and Re-ranking approach was used in this approach.

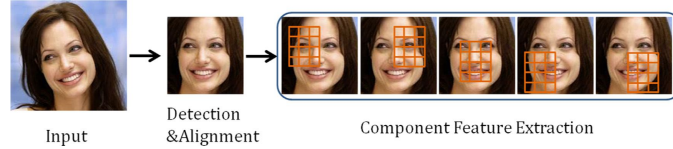
## III. PROPOSED SYSTEM

### A. System Overview

Human face image are trained in the Offline stage. In image preprocessing, Detection of face, Face landmark detection and Face alignment takes place to get the aligned face for the further processing. The component of face images such as eyes, nose and mouth is extracted and identity based vocabulary is added for each component. For each component Facial Features are extracted. The attribute score is calculated from the aligned face. The extracted feature and attribute scores are combined to generate codewords using sparse coding for each face images in the run time environment where the additional attributes are detected. Similarity computation is done with sparse codewords and attribute signature to get set of retrieval images in attribute embedded inverted indexing and result of similar images from database image are retrieved.

**B. Face detection and alignment (Preprocessing)**

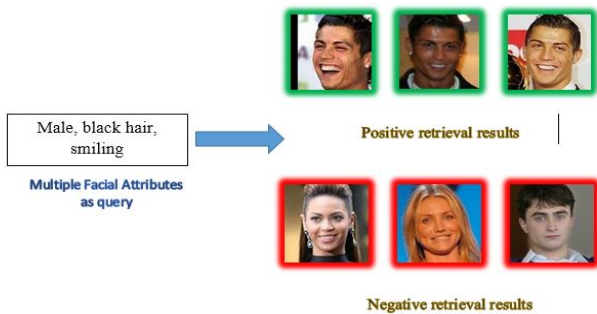
The First, system detects the location of the faces using Viola-Jones face detector for all image in the database. For each detected face image, we will extract the facial components of face which are two eyes, nose tip, and two mouth corners. A 5x7 grid is defined at each detected component, where each grid is a square patch . Thereby a total of 175 grids are extracted from five components. We align all the faces to a canonical position and the faces are re-scaled to 200x200 using bicubic interpolation.



**Figure. 3.2 Face detection and alignment**

**C. Attribute detection and Feature Projection**

The enabled facial attributes are derived by projecting the low level feature components into attribute discriminative subspace via mapping attribute labels.



**Fig. 3.3 Facial Attributes Detection**

Proposed algorithm selects optimal set of attributes from the original set of attributes.

Given data  $X = \{x_1, x_2, \dots, x_n\}$  consisting of n samples with feature dimension d and possible class labels  $\{1, 2, \dots, C\}$ , we aim to find a transformation T such that the transformation that maximizes the between-class scatter while minimizing the within-class scatter of the projected data.

The transformation matrix T is obtained through

$$T_{FDA} = \arg \max_{T \in R^{d \times d}} [\text{tr}((T^T S^W T)^{-1} T^T S^B)]$$

In our method, for each attribute k we learn a projection of transformation matrix to transform the original low-level feature x in each cell of a face component to the attributed encoded feature. The attributed encoded features using K different attributes are then concatenated.

During on-line processing when the query image is supplied the features are extracted and concurrently additional attributes are detected in run time and quantized as sparse codewords and combined with binary representation of the additional feature projection in the enabled attributes subspace.

**D. Indexing**

For embedding attribute information into index structure, for each image, in addition to sparse codewords c (i) computed from the facial appearance, we used a dimension binary signature to represent its human attribute, b(i)

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

The similarity score is then modified into,

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

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

$$t_{h_i} = LB_i + \frac{(h_i - L_i)}{(H_i - L_i)} * (UB_i - LB_i)$$

Conditional entropy in the pre-classified data set for the descriptors or values of the current attribute are calculated. One way to achieve this is to make Attribute Importance score as an inverse of the corresponding Conditional Entropy. Attribute Importance score is then transformed using Linear Transformation in the Interval of Variation for current attribute. Normalise the Attribute Importance scores of all the attributes so that they sum up to the number of attributes. The normalised values are the dynamic weights of each of the attributes.

#### IV. EXPERIMENTAL COLUMN

##### A. Experimental Setup

Two public datasets are used to evaluate our proposed method. The first dataset is Labeled Faces in the Wild (LFW) contains 13,233 images of 5749 different people, among which 1680 people have more than one image. These images are of uncontrolled environments and vary radically in pose, illumination, expression. 10 people that have more than 50 images each are used as query subjects. For each person, 10 images are randomly sampled as query images. In total there are 100 query images and the remaining 13,133 images are used as gallery images.

The Pubfig dataset is used in a second set of experiments. The Pubfig dataset contains 58,797 real-world images of 200 people collected from Internet. We randomly select 100 images for each subject and in total 40 subjects are chosen for our gallery and query. 10 images for each person are used as query images and the rest of the images are used as gallery. Thus there are 400 query images and 3600 gallery images in total. Following Table summarizes the datasets used in the experiments

Table 4.1 Experimental Dataset Information

Category	LFW	Pubfig
Gallery Subjects	5749	40
Gallery Images	13133	3600
Query Subjects	10	40
Query images	100	400

##### B. Evaluation Metrics

We use three metrics for evaluating our Experimental Results. The Mean average precision is the mean of the average precision scores for each query image and is calculated by

$$MAP = \frac{1}{N} \sum_{i=1}^N \int_0^1 P_i(r) dr$$

where  $N$  is the number of queries, and  $p_i(r)$  is the precision at recall  $r$  for query image  $i$ .  
The precision is defined as follows

$$Precision = \frac{A}{A + C} \times 100\%$$

The recall is defined as follows

$$Recall = \frac{A}{A + B} \times 100\%$$

where

A = Number of relevant images retrieved

B = Number of relevant images not retrieved

C = Number of irrelevant images retrieved

### C. Retrieval Results



Figure. 4.1 Retrieved Results of ASW System



Figure. 4.2 Retrieved Results for Proposed System

The implementation of our method is done in Matlab on a laptop with Intel i7 2.4 GHz CPU and 4 GB RAM. On average for one query image, the low-level feature extraction takes 0.007 s. Subspace mapping using Feature label by semi supervised learning takes about 0.0047 s. On the LFW dataset, the initial search for one query takes 0.0697 s and on the Pubfig dataset it completes in 0.0056 s using proposed method.

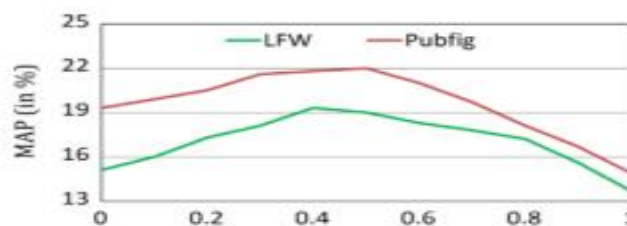


Figure 4.3 Comparison of LFW and Pubfig MAP values with  $\alpha$  in the Equation

## V. CONCLUSION

Face image retrieval method using Subspace mapping of Feature Labels has been proposed in this paper. Instead of relying on low-level visual features like the previous methods to retrieve similar images from a gallery given a query image, we use attribute information to transform the low-level features to high level feature projections in a semi-supervised learning approach. The transformation is learned by mapping the feature projections on subspace labels on Discriminant Analysis model. The highest MAP is achieved when  $\alpha$  remains close to mean on both datasets. When  $\alpha$  is 0 the results are purely based on the distance between high level image features. When  $\alpha$  is 1 the results are purely based on distance between low-level image features. We compared the retrieval performance of our methods against state-of-the-art work which also utilised both low-level features and attribute information on two publicly available real-world face image datasets. Experimental results showed that the proposed method achieved superior retrieval accuracy of a relative improvement of 87% on MAP. Future work includes further investigation of the retrieval performance using different types of additional attributes. Effective feature and attribute selection are to be studied to further improve the retrieval accuracy

## REFERENCES

1. Bor-Chun Chen., Winston H. Hsu., Yan-Ying Chen., Yin-Hsi Kuo., "Scalable Face Image Retrieval Using Attribute- Enhanced Sparse Codewords," IEEE Transactions on Multimedia, Volume:15, Issue: 5, 1163 – 1173, August 2013.
2. Chen-Kuo Chiang., Te-Feng Su., Chih Yen., Shang-Hong Lai "Multi-Attributed Dictionary Learning for Sparse Coding," ICCV2013,IEEE XPLore
3. Arijit Biswas., Amar Parkash ., Devi Parikh., "Attributes for Classifier Feedback," in European Conference on Computer Vision (ECCV), 2012.
4. Neeraj Kumar., Alexander C. Berg., Peter N. Belhumeur., Shree K. Nayar., "Describable Visual Attributes for Face Verification and Image Search," IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 10, october 2011.
5. Jun Liu,Xiaojun Jing,Songlin Sun,Zifeng Lian," Local Gabor Dominant Direction Pattern for Face Recognition", Chinese Journal of Electronics,2015.
6. Siddique B., R. S. Feris., L. S. Davis., "Image ranking and retrieval based on multi-attribute queries," in Proc. IEEE Conf. Computer Vision and Pattern Recognition,2011
7. Stephen Milborrow., Fred Nicolls., "'Locating facial features with an extended active shape model," Springer-Verlag Berlin Heidelberg 2008
8. Timo Ahonen., Abdenour Hadid., Matti Pietikainen., "Face recognition with local binary patterns," Springer-Verlag Berlin Heidelberg 2004
9. Walter J. Scheirer., Neeraj Kumar.,Peter N. Belhumeur .,Terrance E. Boult.,"Multi-Attribute Spaces: Calibration for Attribute Fusion and Similarity Search", University of Colorado at Colorado Springs
10. Yu-Heng Lei., Yan-Ying Chen., Lime Iida., Bor-Chun Chen., Hsiao-Hang Su., Winston H. Hsu., "Photo Search by Face Positions and Facial Attributes on Touch Devices," November 28–December 1, 2011.
11. Wu Z, Ke Q, Sun J, and Shum H Y, "Scalable face image retrieval with identity-based quantization and multi-reference re-ranking," in Proc. IEEE Conf. Computer Vision and Pattern Recognit., 2010.
12. Xiaohui Shen., Zhe Lin., Jonathan Brandt., Ying Wu., "Detecting and aligning faces by image retrieval,"in Proc. CVPR, 2013, pp.3460-3467
13. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," IEEE Conference on Computer Vision and Pattern Recognition, 2001.
14. S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," European Conference on Computer Vision, 2008.
15. N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and simile classifiers for face verification," International Conference on Computer Vision, 2009.
16. J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learningfor sparse coding," ICML, 2009.
17. Md. Abdur Rahim., Md. Najmul Hossain., Tanzillah Wahid ., Md. Shafiul Azam.,"Face Recognition using Local Binary Patterns (LBP)," Global Journal of Computer Science and Technology Graphics & Vision Volume 13 Issue 4 Version 1.0 Year 2013
18. Hao Shao., T. Svoboda., V. Ferrari., T. Tuytelaars and L. Van Gool.,"Fast Indexing For Image Retrieval Based On Local Appearance With Re-Ranking," Computer Vision Lab, Swiss Federal Institute of Technology.