International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 6, Issue 5, pp 207-212, May-2022
https://doi.org/10.47001/IRJIET/2022.605029

Performance Enhancement of Scalable Face Image Retrieval Using Multi Reference Re-Ranking

[1]A.Jaganathan, [2]N.Rupavathi, [3]Dr.K.Ramesh, [4]R.Bhuvaneswari, [5]D.Ilamparuthi

[1]PG Scholar, Applied Electronics, Jayam College of Engineering and Technology, Dharmapuri, Tamilnadu, India
[2]Associate Professor, Dept. of ECE, Jayam College of Engineering and Technology, Dharmapuri, Tamilnadu, India
[3]Professor, Dept. of ECE, Jayam College of Engineering and Technology, Dharmapuri, Tamilnadu, India
[4,5]Assistant Professor, Dept. of ECE, Jayam College of Engineering and Technology, Dharmapuri, Tamilnadu, India

***Abstract* - Image retrieval requires a system to find information relevant to a query which represents images containing faces of the same person appearing in the query image. In this paper, we aim to build a scalable face image retrieval system. For this purpose, we develop a new scalable face representation using Gray level co-occurrence matrix (GLCM) features at different orientation (45,60,90 degrees),gray level and moment invariant features, orientation histogram features and Law's texture features. The extracted features are trained and classified by feed forward back propagation neural network and Support vector machine (SVM) classifier to rank the candidate images. The performance of the designed face image retrieval system will be analyzed in terms of Accuracy and retrieval rate .The performance of the proposed retrieval system will be compared with existing system. In this proposed system we aim to build a scalable face image retrieval system.**

For this purpose, we develop a new scalable face representation using both local and global features. In the indexing stage, we exploit special properties of faces to design new component based local features, which are subsequently quantized into visual words using a novel identity-based quantization scheme. We also use a very small Hamming signature (40 bytes) to encode the discriminative global feature for each face. In the retrieval stage, candidate images are firstly retrieved from the inverted index of visual words. We then use a new multi-reference distance to re-rank the candidate images using the Hamming signature. On a one million face database, we show that our local features and global Hamming signatures are complementary the inverted index based on gray level and moment invariant features provides candidate images with good recall, while the multi-reference re-ranking with global Hamming signature leads to good precision.

Keywords: Performance Enhancement, Face Image Retrieval, Multi Reference, Re-Ranking, Matlab.

I. INTRODUCTION

In this proposed system we aim to build a scalable face image retrieval system. For this purpose, we develop a new scalable face representation using both local and global features. In the indexing stage, we exploit special properties of faces to design new component based local features, which are subsequently quantized into visual words using a novel identity-based quantization scheme. We also use a very small Hamming signature (40 bytes) to encode the discriminative global feature for each face. In the retrieval stage, candidate images are firstly retrieved from the inverted index of visual words. We then use a new multi-reference distance to re-rank the candidate images using the Hamming signature. On a one million face database, we show that our local features and global Hamming signatures are complementary—the inverted index based on local features provides candidate images with good recall, while the multi-reference re-ranking with global Hamming signature leads to good precision. As a result, our system is not only scalable but also outperforms the linear scan retrieval system using the state-of the- art face recognition feature in term of the quality.

Problem Identified

The sub problems of face recognition (like registration, illumination, classification, etc.) have been addressed in many previous works, there have been very few works only describing a complete face recognition system, State-of-the-art image retrieval systems achieve scalability by using a local representation and global methods, but their performance updates quickly in the face image domain, mainly because they produce visual words with low discriminative power for face images, and mainly focusing the special properties of faces. The leading features for face recognition can achieve good retrieval performance only in single image reference based system , but these features are not suitable for multiple image/face recognition indexing as they are high-dimensional and global, thus not scalable in either computational or storage cost. In this system the accuracy is low and searching time

ISSN (online): 2581-3048
Volume 6, Issue 5, pp 207-212, May-2022
https://doi.org/10.47001/IRJIET/2022.605029

delay is high when we are going to retrieval of the images from multiple one.

II. PROPOSED SOLUTION

In proposed system all the problems what we find out from existing method are rectified. We can able to retrieval the images what we need from thousands of appearances by using multiple image reference system with the help of gray level and movement invariant features from the face images.

In this system, we assume face images are frontal with up to about 20 degrees of pose changes, such that the face components are visible in a given face image listed below.

- Eyes
- Nose
- Mouth

III. FACE DETECTION

State-of-the-art image retrieval systems achieve scalability by using a local representation and global methods, but their performance updates quickly in the face image domain, mainly because they produce visual words with low discriminative power for face images, and mainly focusing the special properties of faces.

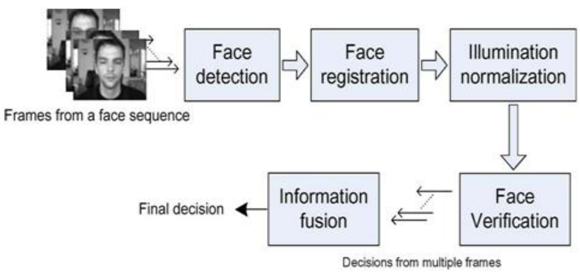

Figure 1: Face detection

The leading features for face recognition can achieve good retrieval performance only in single image reference based system , but these features are not suitable for multiple image/face recognition indexing as they are high-dimensional and global, thus not scalable in either computational or storage cost. In this system the accuracy is low and searching time delay is high when we are going to retrieval of the images from multiple one.

3.1 Face Query Pipeline

In the proposed system all the problems what we find out from existing method are rectified. We can able to retrieval the images what we need from thousands of appearances by using multiple image reference system with the help of gray level and movement invariant features from the face images. In this system, we assume face images are frontal with up to about 20 degrees of pose changes, such that the face components are visible in a given face image.

We propose a novel face image representation using both local and global features. First, we locate component-based local features that not only encode geometric constraints, but are also more robust to pose and expression variations. Second, we present a novel identity based quantization scheme to quantize local features into discriminative visual words, allowing us to index face images, a critical step to achieve scalability. Our identity based quantization can better handle intra-class variation using multiple examples from multiple identities. Finally, in addition to the local features, we compute a 40-byte Hamming signature for each face image to compactly represent a high-dimensional discriminative global (face recognition) feature.

Our face retrieval system takes advantages of the fact that local features and global features are complementary. Local features allow us to efficiently traverse the index of a large scale face image database, and return top candidate images (e.g., 1,000 candidates). While the precision may be low, we can achieve a good recall in this index traversing stage. Then, the Hamming signatures (derived from global features), which is as small as 40KB for 1,000 images, are used to re-rank the candidate images. By using a new multi-reference distance metric, the precision can be significantly improved. Overall our face retrieval system is not only scalable, but also outperforms state-of-the-art retrieval or recognition systems in terms of both precision and recall, which is demonstrated by experiments with a database containing more than one million face images.

3.2 Local Features for Indexing

In this section we describe the details of the local features, and a novel identity-based quantization for inverted indexing.

3.2.1 Component-Based Local Features

Figure shows our local feature extraction and indexing pipeline. First, five facial components (two eyes, nose tip, and two mouth corners) are located on a detected face by a neural-network based component detector.

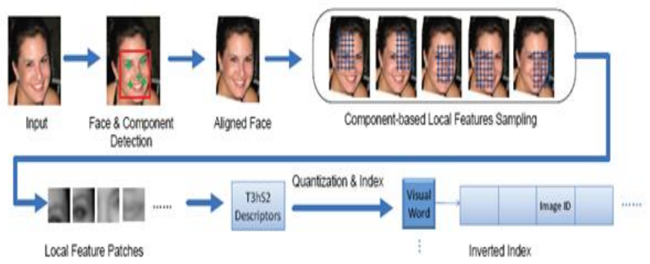

Figure 2: Local feature extraction and indexing

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 6, Issue 5, pp 207-212, May-2022

https://doi.org/10.47001/IRJIET/2022.605029

The face is then geometrically normalized by a similarity transform that maps the positions of two eyes to canonical positions. We define a 5 × 7 grid at each detected component. In total we have 175 grid points from five components. From each grid point we extract a square image patch. A T3hS2 descriptor (responses of steerable filters) is then computed for each patch. All descriptors are quantized into visual words that are subsequently indexed. Notice that the existing interest-point based local feature detectors are not suitable for face images. Such detectors tend to detect features in regions with rich textures or high contrast. They do not perform as well on face images since they contain mostly smooth textures. Compared to defining grid points over the whole face, our features are localized to the components, which allows flexible deformation among the components and are more robust to face pose and expression variations.

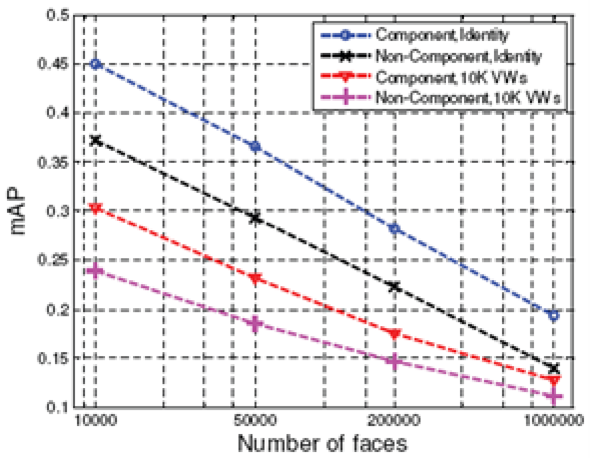

Figure 3: Comparison of "Component"-based and "Non-Component"-based local feature extraction approaches with different quantization methods

Also note that gird points from different components have some overlaps, this, together with the histogram-based T3hS2 descriptors, allows our system to tolerate some degree of errors in the component localization. To enforce geometric constraints among features, we assign each grid point a unique ID, which is called "position id". This is in contrast to existing models that allow features to match even they are coming from different grid points in the face, which performs worse in our task.

3.2.2 Identity-Based Quantization

For scalability, the extracted local features 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 unsupervised learning is not very good for training a vocabulary for face images, where intra-class variations are often larger than inter-class variations when the face undergoes pose and expression changes. Quantization errors will degrade the retrieval performance.

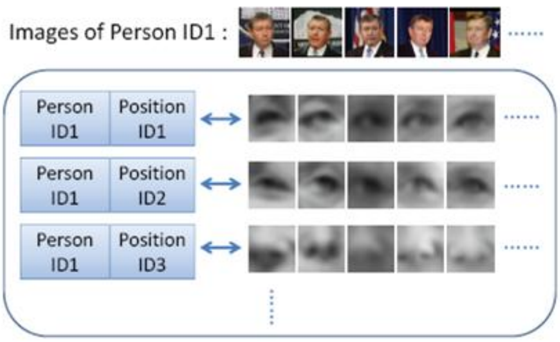

Figure 4: Identity-based vocabulary from one person

In this section, we propose an identity-based quantization scheme using supervised learning. Our training data consists of P different people and each person has T face examples, at various poses, expressions, and illumination conditions. Figure 3 shows example face images of one person and constructed visual words. Since each person has a unique "person id" and each grid point has a unique "position id", we 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 our identity-based quantization - the features under various pose/lighting/expression conditions have a chance to be quantized to the same visual word.

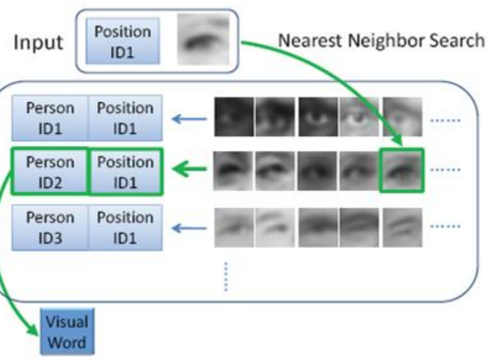

Figure 5: Extracted at the "Position ID1"

IV. RESULTS AND DISCUSSION

The candidate images returned from traversing index are initially ranked based on the number of matched visual words, i.e., it is solely based on the query image. Images of one face contain variations induced by changes in pose, expression, and illumination. We account for such intra class variations by using a set of reference images to re-rank the candidate images. At each iteration, we select a reference image that is

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 6, Issue 5, pp 207-212, May-2022
https://doi.org/10.47001/IRJIET/2022.605029

close to both the query image and the reference images from the previous iteration. More specifically, at each iteration we select an image I that minimizes the following cost:

$$D = d(Q, I) + \alpha \cdot \frac{1}{|\mathbf{R}|} \sum_i d(R_i, I)$$

Where Q is the query image, R = {Ri} is the current reference set, d(·, ·) is the Hamming distance between two faces, and α is a weighting parameter. I is then added to R. The iterative process stops when the expected number of reference images are chosen, or the distance D is larger than a threshold. By using a conservative threshold, the majority of the reference images are expected to be from the same person in the query image.

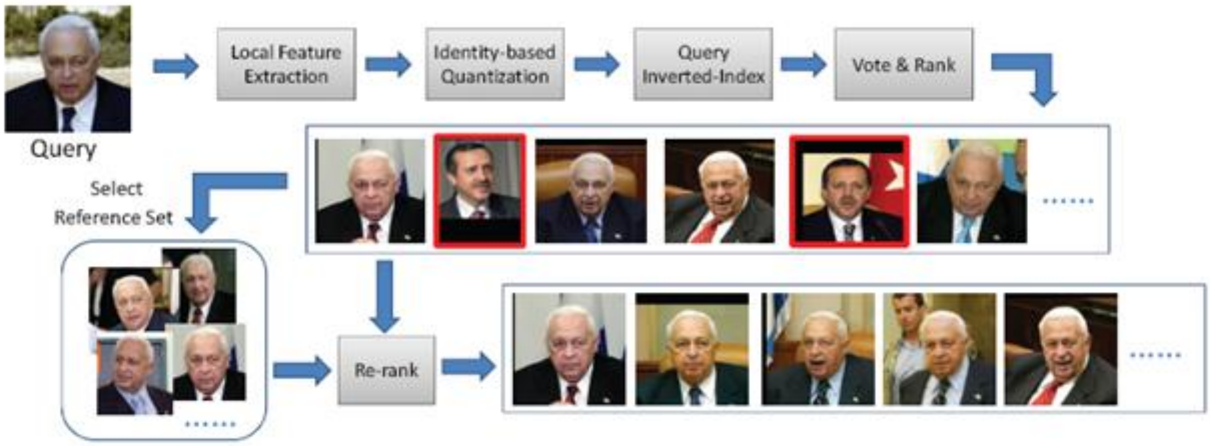

Figure 6: Face query pipeline

Figure shows the basic process of the multi-reference re-ranking. In our approach, we use a fixed size of reference set for all the queries. One alternative is to set up some thresholds for choosing the reference set adaptively. Faces in a dense region have smaller distances among each others. Applying a threshold to decide the number of reference images may bring many un-related images. Vice versa, faces in a sparse region of the feature space may result in insufficient reference images.

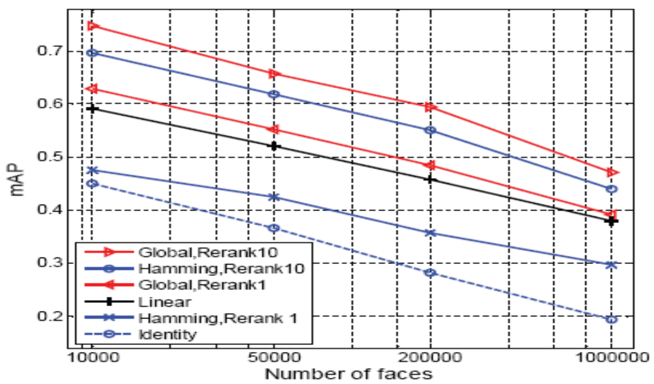

Figure 7: Comparison of different re-ranking methods

4.1 Face Detection

Face detection is the process of automatically locating human faces in visual media (digital images or video). The appearance of faces is governed by a large number of factors, and can therefore vary widely from one face to the next and one image to the next. Some of these complicating factors include variations in age, gender, ethnicity, head pose, facial expressions, lighting, and image quality and compression artifacts.

4.2 Face Tracking

Face tracking extends face detection to video sequences. Any individual who appears in a video for any length of time generates a face track – that is, a sequence of face instances across time. The goal of face tracking then is to aggregate single-frame detection results into a collection of face tracks. Our face tracking technology makes no a priori assumptions about the background, camera motion, or the number of individuals in a scene. We can track an arbitrary number of faces simultaneously, for any length of time, and require no “lock-on” phase or other type of initialization.

4.3 Face Registration

Generally speaking, the location of the detected face is not precise enough for further analysis of the face content. It has been emphasized in the literature that face registration is an essential step after face detection for subsequent face interpretation task.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 6, Issue 5, pp 207-212, May-2022

https://doi.org/10.47001/IRJIET/2022.605029

The following two popular ways of doing the face registration can be found in the literature:

- Holistic methods and
- Local methods.

Face recognition is the process of automatically determining whether two faces are the same person. A number of factors make this a challenging problem for computers. Faces in images and video can be captured at various resolutions, quality, and lighting conditions.

Figure 8: Unparalleled Accuracy on Real-World Data

By design, our face recognition technology performs accurately on real-world, uncontrolled data. For example, our recognizers correctly identify the 24 images of our VP of Research and Development (Michael Nechyba) despite some challenging factors:

We propose a real-time robust facial feature detector based on the Viola–Jones method, in corporating a novel error model that is precise and concise. Given that it is, in reality, impossible to build a reliable facial feature detector with both low FAR and low FRR, we guarantee the detection of the facial features in the first place, at the cost of a high number of false detections. The facial feature detection problem is therefore converted into a post selection problem of the multiple detected facial features.

4.4 Illumination Normalization

The variability on the face images brought by illumination changes is one of the biggest obstacles for face verification. It has been suggested that the variability caused by illumination changes easily exceeds the variability caused by identity changes. Illumination normalization, therefore, is a very important preprocessing step before verification.

There are several advantages brought by the simplified LBP filtering. First, the LBP is a local measure, so the LBPs in a small region are not affected by the illumination conditions in other regions. Second, the measure is relative, and therefore, is strictly invariant to any monotonic transformation, such as shifting, scaling, or logarithm, of the pixel values. Third, we largely reduce the sensitivity of the LBP value to noise by assigning uniform weights in all eight directions. Finally, even for the MPD with limited computation resources, the proposed filtering operation is extremely fast.

4.5 Face Verification

In the verification case, the user class and the impostor class are more closely distributed in space than the two classes in the detection case. In other words, the chance that an impostor face resembles a user face is much higher than the chance that a random background patch resembles a face patch.

In the detection case, a number of support vectors, are sufficient to “support” the decision boundary. In the recognition case, however, the distributions of the two classes are intermingled in a more complex way. This implies that the boundary-based classification methods that work well on the face detection problem, like the support vector machine method.

Which relies explicitly on the support vectors, or the Viola–Jones Adaboost method, which relies implicitly on the highly weighted samples, is no longer suitable for the verification problem. The better solution, instead, is to classify

ISSN (online): 2581-3048
Volume 6, Issue 5, pp 207-212, May-2022
https://doi.org/10.47001/IRJIET/2022.605029

in the overlapped regions with a minimal possible error, from a statistical point of view. For this reason, we propose to verify the feature vectors in a statistically optimal way using the likelihood ratio.

We notice that the face verification is normally in a very high dimensional space. A small-enough face images, for example, of size 32 × 32, already has 1024 pixels, which implies a 1024- dimensional feature vector. High-dimensional space potentially has great power of discrimination but is relatively difficult to generalize.

V. CONCLUSION

This system proposes a content-based indexing and retrieval (CBIR) system based on query-by-visual-example using hierarchical binary signatures. Binary signatures are obtained through a described binarization process of classical features (color, texture and shape). The Hamming binary distance (based on binary XOR operation) is used for computing distances. This technique was tested on a real natural image collection containing 10 000 images and on a virtual collection of one million images. Results are very good both in terms of speed and accuracy allowing near real-time image retrieval in very large image collections.

REFERENCES

[1] Zhimin Cao, Qi Yin, Jian Sun, and Xiaoou Tang. Face recognition with learning-based descriptor and pose-adaptive matching. In Proc. IEEE International Conference on Computer Vision and Pattern Recognition, 2021.

[2] Jaime g. Carbonell, yiming yang, robert e. Frederking, ralf d. Brown, yibing geng, and danny lee. Translingual information retrieval: a comparative evaluation. In in proceedings of the fifteenth international joint conference on artificial intelligence, pages 708–714, 2020.

[3] Jian Chen, Rui Ma, and Zhong Su. Weighting visual features with pseudo relevance feedback for cbir. In Conference on Image and Video Retrieval, 2021.

[4] J. Friedman, J. Bentley, and R. Finkel. An algorithm for finding best matches in logarithmic expected time. ACM Trans. Math. Software, 3:209226, 2021.

[5] A.Gionis, P. Indyk, and R.Motwani. Similarity search in high dimensions via hashing. In International Conference on Very Large Data Bases, 2020.

[6] G. Hua and A. Akbarzadeh. A robust elastic and partial matching metric for face recognition. In Proc. IEEE International Conference on Computer Vision and Pattern Recognition, 2021.

[7] Gary B. Huang, Marwan Mattar, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. In European Conference on Computer Vision, 2020.

[8] Herve Jegou, Matthijs Douze, and Cordelia Schmid. Hamming embedding and weak geometric consistency for large scale image search. In European Conference on Computer Vision, 2020.

[9] Herve Jegou, Matthijs Douze, and Cordelia Schmid. Packing bagof-features. In Proc. IEEE International Conference on Computer Vision, 2021.

[10] Brian Kulis and Kristen Grauman. Kernelized locality-sensitive hashing for scalable image search. In Proc. IEEE International Conference on Computer Vision, 2020.

[11] Neeraj Kumar, Alexander C.Berg, Peter N.Belhumeur, and Shree K.Nayar. Attribute and simile classifiers for face verification. In Proc.IEEE International Conference on Computer Vision, 2021.

[12] Ping-Han Lee, Gee-Sern Hsu, and Yi-Ping Hung. Face verification and identification using facial trait code. In Proc. IEEE International Conference on Computer Vision and Pattern Recognition, 2021.

[13] Z. Lei, S. Li, R. Chu, and X. Zhu. Face recognition with local gabor textons. Lecture Notes in Computer Science, 4642:49, 2021.

Citation of this Article:

A.Jaganathan, N.Rupavathi, Dr.K.Ramesh, R.Bhuvaneswari, D.Ilamparuthi, “Performance Enhancement of Scalable Face Image Retrieval Using Multi Reference Re-Ranking” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 6, Issue 5, pp 207-212, May 2022. Article DOI https://doi.org/10.47001/IRJIET/2022.605029

****