Fast BGP Face Retrieval Based on K-means Clustering

Liwei Li¹, a, Haibin Xie²b*
¹National University of Defense Technology, Changsha 410073, China
²National University of Defense Technology, Changsha 410073, China
a 840822463@qq.com
b xhb2575 sx@sina.com

Abstract. The traditional face image retrieval method is to compare the target picture with all the pictures in the database one by one, resulting in great time consumption. This paper used the binary gradient pattern (BGP) as face feature extractor. Firstly, the database was encoded offline. Then the encoded database was offline clustered. Next, we extracted the feature of the target picture by BGP and found the nearest clustering centre from the database encoding. Finally, the target picture code for retrieval was selected in the target class.

1. Introduction
In recent years, face recognition has always been a hot issue in the field of image processing, computer vision, pattern recognition and cognitive science, and has the characteristics of obvious discrimination and easy access. It is widely used in security verification, quick payment, video survey, identification and other occasions. One of the most widely applied face recognition scenarios is to retrieve similar pictures quickly from the database with a target picture, which is given before. These problems can be summarized as fast face retrieval.

Face recognition algorithms can be summarized and divided into the following categories: 1) The recognition methods based on the local features of human faces, such as Bayesian method [1], Fisherfaces method [2], Binary Gradient Pattern (BGP) [3] and elastic graph matching [4]; 2) Recognition methods based on global features of face, such as Linear Discriminant Analysis (LDA) [6], Principle Component Analysis (PCA) [7], Independent Component Analysis (ICA) [8] and other methods; 3) Approaches based on the combination of global features and local features, such as the method based on the combination of eigenfaces and facial features [9]; 4) Methods based on deep learning are also used in face recognition, for example, the facenet is proposed and applied in the face field [10].

The fast retrieval methods include the following types: 1) Retrieval methods based on the pouch model [11][12]; 2) Methods based on KD tree [13][14]; 3) Search methods based on vector clustering and quantization [15][16]; 4) Hash-based retrieval methods [17][18].

The objects of fast retrieval are usually scene-based database pictures such as CIFAR-10 and INRIA Holidays. There are few researches on fast face image retrieval. This is due to the variable nature of the human face, and the feature extraction and description of the human face is vulnerable to expression, light, etc. What’s more, face gestures and expression changes in large-scale face databases are fierce, and it is difficult to extract stable descriptors for fast retrieval.

This paper proposes a fast retrieval method for BGP face based on K-means clustering. In the process of method development and research, the following advantages are reflected: 1) Application of
the simple and robust face descriptor BGP is conducive to stable and accurate facial feature extraction; 2) Face BGP encoding and K Mean clustering are combined to achieve quick face retrieval; 3) The Yale database is artificially expanded, which can achieve fast retrieval of large face database with relatively stable face pose.

2. Binary Gradient Pattern (BGP)

2.1. The basic thinking of BGP

BGP is a concise and efficient face descriptor proposed by Weilin and Hujun [4] in the 2017 Pattern Recognition article "Robust face recognition with structural binary gradient patterns". The idea stems from the new Gradientfaces descriptor proposed by Zhang et al. This descriptor replaces the pixel intensities with the image gradient direction (IGO) to describe the human face in order to achieve robustness against changes in illumination. Binary gradient model measures the relationship between local pixels in the image gradient domain and effectively encodes the underlying local structure into a set of binary strings, which not only increases the discriminative power, but also greatly simplifies the computational complexity.

2.2. The algorithm principle of BGP

In order to find out the potential structure of the gradient domain, we use BGP to calculate the gradients of the images from multiple directions and encode them into a series of binary strings, which can represent slight boundary changes and texture information. Therefore, BGP has strong discriminability and can achieve excellent recognition accuracy even facing occlusion, light and expression change, etc.

- Fig 1. BGP basic descriptor: (a) Eight adjacent pixels of a central pixel (value of 115) (b) Four directions (c) The main binary string (red), encoded as 0111, with label 07

The basic principles of BGP are as follows:

1. Given a central pixel and a series of local neighboring pixels (such as the eight adjacent pixels in Figure 1)

2. Based on Equation 1, a pair of binary encodings (primary and secondary) can be calculated based on two symmetric neighboring pixels in each direction. As shown in (b) (c) of fig. 1, four pairs of binary numbers are obtained from four directions G1, G2, G3, and G4

\[
B_i^+ = \begin{cases} 
1 & \text{if } G_i^+ - G_i^- \geq 0 \\
0 & \text{if } G_i^+ - G_i^- < 0
\end{cases}
\]

(1)

\[
B_i^- = 1 - B_i^+ \quad i=1,2,...,k
\]

(2)

3. Finally, the label of the center pixel is obtained through four main binary codes, that is, BGP representation as formula (2).

\[
L = B_1^* B_2^* B_3^* B_4^*
\]

(2)

Then it is transferred into a decimal number such as formula(3)
Eight binary numbers are obtained in all four directions, but the major and auxiliary binary numbers in each direction are always complementary. Therefore, only one binary bit is required for each direction. For the sake of conciseness, only the main binary is needed to calculate the label (according to equation (3)).

2.3. The steps of BGP algorithms
After obtaining the BGP image of the face picture, non-overlapping segmentation is performed and each block histogram feature is counted, finally the feature vector of the face is got. The specific algorithm steps are shown in Figure 2.

![Diagram showing the steps of BGP algorithms](image)

**3. K-means Clustering**

3.1. The basic thinking
The K-means clustering algorithm problem can be described as: Given an integer k and a sample set with n data objects, the goal of the problem is to select k cluster centres (i=1, 2,...,k) in order to minimize the objective function F, where the function F is calculated as shown in Equation (4):

$$ F = \sum_{x \in X} \min \| x - C_i \|^2 $$

The algorithm steps
K initial centres are randomly selected from n data samples. According to the distance between the samples and the centres, they are divided into the classes to which the centres belong, forming k initial clusters. The cluster centre is recalculated for each cluster, and then the data is re-divided according to the new cluster centre, and iteration is performed until the number of iterations reaches the maximum value or clustering centres no longer change.

The specific steps are as follows:
Step1: Select k data points randomly as the initial clustering centres, which is shown as follow:
$$ C^{(0)} = \{C_1^{(0)}, C_2^{(0)}, \ldots, C_k^{(0)}\} $$

Step2: Calculate the distances between each sample and the centers of k clusters:
d(x_p, C_j) = \sqrt{\sum_{q=1}^{m} (x_p^{(q)} - C_j^{(q)})^2}, \ p=1,2,...,n, j=1,2,...,k. If d(x_p, C_j) = \min(d(x_p, C_j)) (j=1,2,...,k), then x_p \in C_j, j=1,2,...,k.

Step3: Calculate k new cluster centres:

\[ C_j = \frac{1}{n_{j \in C_j}} \sum_{x \in C_j} x, \ j=1,2,...,k. \]

Step4: Determine if the maximum number of iterations has been reached. If so, the algorithm ends. Otherwise, determine if there is a change in cluster center. If there is a change, go back to Step2. Otherwise, the algorithm ends.

The flow chart of K-means clustering algorithm is shown in Figure 3

Fig 3. The flow chart of K-means

4. BGP face retrieval based on K-means clustering

4.1. The basic thinking

BGP can describe face and extract facial features in a robust and concise manner. It can be used to encode faces to ensure the accuracy and conciseness of coding. The K-means clustering method gathers the codes with higher similarity into one class, which reduces the search space and improves the retrieval efficiency. The combination of BGP face descriptors and K-means clustering methods can be used to accurately describe the face and achieve fast retrieval.

4.2. Retrieval framework

The retrieval process includes two parts: offline operation of the database and online operation of the target picture. In the offline operation part, all the pictures in the face database are firstly encoded by BGP, then the resulting codes are clustered to obtain K categories and their respective cluster centers; in the online operation part, BGP coding is performed on the retrieved pictures first, then the distances between the encoding and the K cluster centers are calculated, and the class that has the smallest
distance is obtained, next this class is regarded as a similar face image result list, finally the results are got by comparing the code of target picture with the similar face image encodings.

The overall search framework is shown in Figure 4:

**Fig 4. Overall retrieval framework**

5. **Experimental results and analysis**

5.1. **Standard Face Library Introduction**
Yale library contains 15 subjects, each with 11 positive face images, a total of 165 face images, including the light, facial expression, occlusion and other changes. A part of Yale face image is shown in Figure 5.

**Fig 5. A Part of images of one person in Yale**

5.2. **Sample expansion**
Yale face database contains various changes in face, but the sample size is limited. Only 165 pictures cannot meet the data volume requirements for large-scale search. However, a great many large-scale face databases, such as LFW[19] and CASIA[20], contain face images with different postures, even including side faces. It is difficult to extract facial features by a single face descriptor.

In view of the above contradictions, this paper expands the samples based on the Yale database. The expansion method is to perform pixel intensity transformation or add some noises to each picture. Finally, the number of samples will be expanded to 495, 1650, 4950, and 16500.

5.3. **Experimental results and analysis**
On Yale and the expanded sample database, we apply the method proposed in this paper and the traditional one-by-one retrieval method. The numbers of samples N are 165, 495, 1650, 4950, and 16500, and the K values are 3, 5, 7, and 9, respectively. The experimental steps are as follows:

Step1: A picture is chosen from the database as the target image.

Step2: We retrieve the most similar ten pictures and record the retrieval time using the traditional method and the method proposed in this paper respectively.

Step3: The above two steps are repeated for ten times.

Step4: The average retrieval time for ten times is considered as the retrieval time of the method on the corresponding database.

The final retrieval time comparison is shown in Table 1.

| Method | N=165 | N=495 | N=1650 | N=4950 | N=16500 |
|--------|-------|-------|--------|--------|---------|
| Traditional | 0.021 | 0.045 | 0.134 | 0.494 | 1.332 |
| K=3 | 0.009 | 0.022 | 0.077 | 0.205 | 0.732 |
| K=5 | 0.008 | 0.021 | 0.063 | 0.191 | 0.641 |
| K=7 | 0.008 | 0.019 | 0.064 | 0.208 | 0.571 |
| K=9 | 0.008 | 0.020 | 0.063 | 0.202 | 0.620 |

It can be seen from the table 1 that the BGP face retrieval method based on K-means clustering shortens the retrieval time and improves the retrieval efficiency, compared to the traditional retrieval method.

We select an image randomly as the target image from one experiment where N is 165 and K is 9. The retrieval results obtained by the two methods are shown in Figure 6 and Figure 7, respectively.

![Fig 6. Traditional method retrieval results: (a)target picture (b)the list of retrieval results](image1)

![Fig 7. K-means clustering method retrieval results: (a)target picture (b)the list of retrieval results](image2)

We can find that the retrieval process of K-means clustering method retains high retrieval accuracy compared with the traditional methods. This paper focuses on the retrieval time and efficiency, so there is no over-depth discussion of the accuracy.

6. Conclusion

This paper proposed a fast BGP face retrieval based on K-means clustering method, which combined the advantages of the BGP face feature extractor and the K-means clustering method. In addition to
the robust and accurate description of human faces, the retrieval speed and efficiency had also been improved significantly.

References
[1] Moghaddam and A. Pentland 1997 *J. IEEE Transactions on Pattern Analysis and Machine Intelligence* *19* 696-710
[2] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld 2003 *J. ACM Computing Surveys (CSUR)* *35* 399-458
[3] Weilin Huang and Hujun Yin 2017 *J. Pattern Recognition* *68* 126-140
[4] L. Wiskott, N. Krüger, N. Kuiger and C. von der Malsburg 1997 *J. IEEE Transactions on Pattern Analysis and Machine Intelligence* *19* 775-779
[5] Belhumeur P N, Hespanha J P and Kriegman D J 1997 *J. IEEE Tram on PAMI* *19* 711-720
[6] Turk M and Pentland 1991 *J. Journal of Cognitive Neuroscience* *3* 71-86
[7] Comon P 1994 *J. A New Concept Signal Processing* *36* 287-314
[8] Pentland A, Moghaddam B and Stamer T 1994 *Proc of IEEE Conf on Computer Vision and Pattern Recognition* pp 84-91
[9] Karen Simonyan and Andrew Zisserman 2014 *J. Computer Science* *1409-1556*
[10] F. Schrofl, D. Kalenichenko, and J. Philbin 2015 *CVPR*
[11] Sivic J. and A. Zisserman 2003 *Proceedings. Ninth IEEE International Conference on Computer Vision* pp 1470-1477
[12] Nister D. and H. Stewenius 2006 *IEEE Computer Society Conference* pp 2161-2168
[13] Friedman Jerome H, Jon Louis Bentley and Raphael Ari Finkel 1977 *J. ACM Trans. Math. Softw* *3* 209-226
[14] Arya Sunil, David M. Mount, Nathan S. Netanyahu, Ruth Silverman and Angela Y. Wu 1998 *J. J. ACM* *45* 891-923
[15] Fukunage K. and P. M. Narendra 1975 *J. Computers, IEEE Transactions* *24* 750-753
[16] Muja M. and D. G. Lowe 2009 *International Conference on Computer Vision Theory and Application Vissapp* *1* 331-340
[17] Gionis Aristides, Piotr Indyk and Rajeev Motwani 1999 *Proceedings of the 25th International Conference on Very Large Data Bases* pp 518-529
[18] Alexandr Andoni and Indyk Piotr 2006 *J. IEEE Symposium on Foundations of Computer Science* *51* 459-468
[19] Huang G B, Mattar M, Berg T, et al 2008 *J. Month*
[20] Yi, D., Lei, Z., Liao, S. and Li, S. Z. 2014 *J. Computer Science*