ISSN: 2664-9640

Model and implementation of large-scale fingerprint image retrieval

Yifan Qin[1, #], Qiming Li[2, #], Xianhao Hua[1]

[1] School of Materials Science and Engineering, Guilin University of Electronic Technology, Guilin, Guangxi, 541214, China

[2]School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin, Guangxi, 541214, China

Abstract

In the field of biometric identification, fingerprint is one of the most unique and persistent of the biological features, widely used in identity identification. The retrieval principle of fingerprint image can be viewed as an image: using several large screens to quickly and accurately screen out most of the images in the database without the "same" relationship with the query fingerprint image. After the retrieval process is completed, a small number of images and query fingerprints are highly similar, requiring further fingerprint identification algorithm to do "one by one" identification. We consider a new fingerprint matching model, through fingerprint binarization and refinement, fingerprint feature extraction and denoising, fingerprint endpoint and fork extraction, establish the fingerprint cascade decision comparison model, then on the basis of this algorithm for different fingerprint quality set fingerprint matching performance experiment, design three groups of test samples, take the highest recognition rate threshold of 97%, considering dimension for fingerprint quality, feature alignment method and recognition performance. The number of positive and negative samples of their fingerprints was 1,000 pairs, respectively. The experiments were designed considering that the proportion of unqualified fingerprint quality is 0%, 20% and 50%, respectively. The experimental results are shown in Table 6,7 and 8.The recognition performance is about 4% higher than the thread feature alignment method, 6% higher than the structure alignment method, 4% higher than the thread feature alignment method, and 6% higher than the structure alignment method.The experimental results demonstrate the feasibility and correctness of our proposed cascade decision matching method.

Keywords

Composite gradient vector, vector cluster, fingerprint cascade decision.

1. Background and Restatement

In the field of biometric identification, fingerprint is one of the most unique and the most persistent biological features in identity identification.

The fingerprint recognition process is divided into two parts: feature extraction and alignment. In the feature extraction process, the fingerprint features for fingerprint identification are extracted. Generally, the most common fingerprint feature in the world is the feature of "detail point". The visual display form is the light blue circle and the light blue short segment protruding externally in Figure 1, and the short segment is used to indicate the direction of the line at the detail point.

Generally speaking, during the image retrieval process, the "same" relationship fingerprint image is preserved after the retrieval process, but the whole identification process takes a long time; if the retrieval algorithm filters many fingerprint images, the "same" relationship fingerprint image is smaller after screening, but the whole identification process is less time-consuming. If the fingerprint of the "same" relationship is filtered out during the retrieval process, the subsequent "one by one" alignment algorithm cannot find the fingerprint of the "same" relationship again. Therefore, how to design efficient and accurate search algorithms is a key problem for large-scale fingerprint image retrieval.

Question 1: Analyze the detail features of the fingerprint image (see the three data files described in Annex 3), and give the retrieval method for quick fingerprint retrieval. Please explain the principle of the retrieval method. Key instructions: (1) avoid the mechanism of the "same" fingerprint; (2) complete and clear image retrieval model framework and implementation method; (3) give the temporal and spatial complexity analysis of the retrieval method. Estimate the memory space size occupied by the retrieval method itself and the memory space size of each fingerprint image due to the retrieval method.

Question 2: For the proposed retrieval method: all fingerprint pairs with "the same" relationship (TZ_ is the same reference) provided in the data file. See txt) verify the retrieval method for the query image in the fingerprint data set.

Question 3: The TZ_ 200_ _Data.txt data file is retrieved in the same way as question 2, and the retrieval results are given, and the result data is compressed into.zip format and uploaded to the calculation results of the competition system. Refer to Annex II for the specific requirements.

Question 4: (1) When we complete this question, since we have tried different types of mathematical models and technical routes, we need to use the data of question 2 to verify and select the best retrieval method. Secondly, introduce and evaluate the advantages and disadvantages of the considered model and technical route; (2) the highest screening volume of this question is 97%. For more than 97% screening volume, ask us what better improvement strategies in retrieval accuracy, retrieval time and memory footprint, or what new retrieval methods will be tried.

2. Model hypothesis

In order to ensure the accuracy and operability of the model, and to exclude the interference of some small factors, several reasonable assumptions are made based on the actual situation.

Forget about the disadvantage of the relatively small sample, and think that the sample number is sufficient. Ignoring the effect that the line edge boundaries are not very clear.When performing preprocessing, the accuracy is considered sufficient.

3. Model establishment and solution

3.1. Question 1: A fingerprint matching algorithm based on composite gradient vectors

The composite gradient vector fingerprint matching method is proposed in the direction of biological "adaptive" feature constraint identification. It combines the features of fingerprint dispersion to form the most obvious feature distribution constraint information in the biological visual effect, which is used as the basis for target matching. The fingerprint feature distribution constraint information extracted by the composite gradient vector method can adapt to the nonlinear change of the fingerprint image. When the deformation causes the feature rotation and offset of the fingerprint image, the method can make a relatively stable matching, reflecting the strong non-linear deformation processing ability. At the same time, because the hierarchical labeling rules are adopted to mark the composite gradient vector information, the process of searching the fingerprint matching process is accelerated and the fingerprint matching speed is improved. As we know, fingerprint features are obvious and dense, while fingerprint image features are less and scattered, so some concepts of composite gradient vector need to be redefined and adjusted, specifically described as follows:

Because the collected fingerprint image will have different degrees of noise interference, it is necessary to process the fingerprint image to remove the noise in the image, so as to extract accurate fingerprint features in the following steps, so the algorithm needs to carry out fingerprint image preprocessing first.

In the process of fingerprint sampling, the different contact parts of the fingertip and the acquisition plane will cause the rotation and offset of the collected fingerprint image at different angles, that is, the positions of the same fingerprint details collected at different moments are quite different. This is a common phenomenon of nonlinear deformation, which has some interference to the feature acquisition and matching of the later fingerprint image. Therefore, the detail point coordinates in the fingerprint image should be adjusted before the feature acquisition to solve the rotation problem in the fingerprint image.

Let the original coordinate of the initial point C be (x_0， y_0), and the original coordinate of the singularity is (x'', y'') . Since the new coordinate axis is established, both the initial point and the singular point are translated along the straight line, let the coordinate of the initial point be (0,0), and the coordinate of the singular point is (x', y') .

$$C_{(0,0)} = \begin{cases} 0 = x_0 - x_0 \\ 0 = y_0 - y_0 \end{cases} \tag{1}$$

$$\sigma_{(x',y')} = \begin{cases} x' = x'' - x_0 \\ y' = y'' - y_0 \end{cases} \tag{2}$$

The fingerprint image rotates clockwise along the origin, equivalent to the coordinate axis rotating counterclockwise along the origin, setting the rotation angle as θ， Coordinates of the singularity before the rotation (x', y')， The coordinates of the singularity after the rotation are(x， y). Rotating the fingerprint image is shown in Figure 1. There are

$$\sigma_{(x,y)} = \begin{cases} x = x'\cos\theta + y'\sin\theta \\ y = y'\cos\theta - x'\sin\theta \end{cases} \tag{3}$$

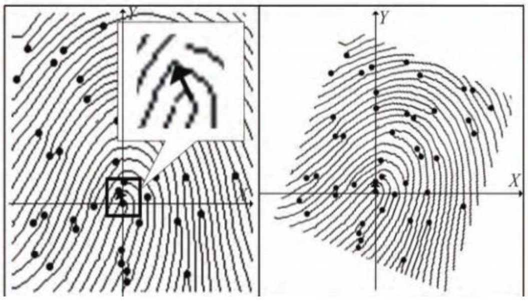

Figure 1 Rotating the fingerprint image

Subsequent vectors of the resulting base vectors were searched in order from small to large and vector labeled at the size of the clip angle. When all the singularities are the beginning or end point of the base vector, the search is stopped, and all the base vectors of the fingerprint image have been sampled. Basal vector acquisition is shown in Figure 2.

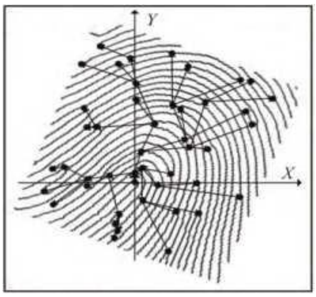

Figure 2 The calibration base vector

The basis vector is the local feature of the fingerprint image, which is less affected by the nonlinear deformation and can describe the local fingerprint information stably. However, the basis vector can not express the fingerprint features comprehensively. Therefore, it is necessary to combine the scattered local features to form the most obvious feature distribution constraint information in the biological visual effect.

In the constraint matrix $\mathbf{RES}_{(m-1)\times n}$, All basis vectors without subsequent vectors are searched for $\boldsymbol{x}_{\alpha,\beta}$, And search for its precursor vector until you find the root vector $\boldsymbol{x}_{1,1}$; Record the search path $\boldsymbol{x}_{1,1}$, $\boldsymbol{x}_{2,1}$, …, $\boldsymbol{x}_{\alpha,\beta}$, All base vectors on the path are combined multidimensional to obtain maximal gradient vectors for all leaf vectors $\boldsymbol{X}_{\alpha,\beta}^{\max}$, as

$$\boldsymbol{X}_{\alpha,\beta}^{\max} = \sum_{i=1,\, j=1}^{\alpha,\beta}{}^{\Theta}\, \boldsymbol{x}_{i,j} = \boldsymbol{x}_{1,1}\Theta\boldsymbol{x}_{2,1}\Theta\cdots\Theta\boldsymbol{x}_{\alpha,\beta} \tag{4}$$

$\varphi(\boldsymbol{X}_{\alpha,\beta}^{\max})$、$\psi(\boldsymbol{X}_{\alpha,\beta}^{\max})$ respectively

$$\varphi(\boldsymbol{X}_{\alpha,\beta}^{\max}) = \sum_{i=1,\, j=1}^{\alpha,\beta} \varphi(\boldsymbol{x}_{i,j}) = \varphi(x_{1,1}) + \cdots + \varphi(\boldsymbol{x}_{\alpha,\beta}) \tag{5}$$

$$\psi(\boldsymbol{X}_{\alpha,\beta}^{\max}) = \sum_{i=1,\, j=1}^{\alpha,\beta} \psi(\boldsymbol{x}_{i,j}) = \psi(x_{1,1}) + \cdots + \psi(\boldsymbol{x}_{\alpha,\beta}) \tag{6}$$

All the maximal gradient vectors in the vector cluster are fused as elements to obtain the composite gradient vector of the fingerprint image $\Delta(\boldsymbol{X}_{\alpha,\beta}^{\max})$, as

$$\Delta(\boldsymbol{X}_{\alpha,\beta}^{\max}) = \left[\boldsymbol{X}_{\alpha_1,\beta_1}^{\max}, \boldsymbol{X}_{\alpha_2,\beta_2}^{\max}, \cdots, \boldsymbol{X}_{\alpha_3,\beta_3}^{\max}\right]^{-1} \tag{7}$$

$$\varphi(\Delta(\boldsymbol{X}_{\alpha,\beta}^{\max})) = \sum_{i=\alpha_1,\, j=\beta_1}^{\alpha_n,\beta_n} \varphi(\boldsymbol{X}_{i,j}^{\max}) = \varphi(\boldsymbol{X}_{\alpha_1,\beta_1}^{\max}) + \cdots + \varphi(\boldsymbol{X}_{\alpha_n,\beta_n}^{\max}) \tag{8}$$

$$\psi(\Delta(\boldsymbol{X}_{\alpha,\beta}^{\max})) = \sum_{i=\alpha_1,\, j=\beta_1}^{\alpha_n,\beta_n} \psi(\boldsymbol{X}_{i,j}^{\max}) = \psi(\boldsymbol{X}_{\alpha_1,\beta_1}^{\max}) + \cdots + \psi(\boldsymbol{X}_{\alpha_n,\beta_n}^{\max}) \tag{9}$$

The dimension of the composite gradient vector is the dimension of its elements (maximal gradient vector). From the perspective of geometry, the dimension of the composite gradient vector is also the constraint set of its fingerprint image singular points, that is, the dimension of the composite gradient vector has not only numerical information, but also has spatial geometry derived relations.Let there be a maximal gradient vector present $\boldsymbol{X}_{\alpha,\beta}^{\max}$ 和 $\boldsymbol{X}_{\alpha',\beta'}^{\max}, \alpha, \alpha' \in [\alpha_1, \alpha_2, \cdots, \alpha_n]$, $\beta, \beta' \in [\beta_1, \beta_2, \cdots, \beta_n]$, The dimensions are respectively recorded as follows $\varphi(\boldsymbol{X}_{\alpha,\beta}^{\max})$ 和 $\varphi(\boldsymbol{X}_{\alpha',\beta'}^{\max})$， The maximal gradient vector is used $\boldsymbol{X}_{\alpha,\beta}^{\max}$ and $\boldsymbol{X}_{\alpha',\beta'}^{\max}$ following relationships

$$\varphi(\boldsymbol{X}_{\alpha,\beta}^{\max})+\varphi(\boldsymbol{X}_{\alpha',\beta'}^{\max})=\varphi(\boldsymbol{X}_{\alpha,\beta}^{\max}+\boldsymbol{X}_{\alpha',\beta'}^{\max})+\varphi(\boldsymbol{X}_{\alpha,\beta}^{\max}\cap\boldsymbol{X}_{\alpha',\beta'}^{\max}) \tag{10}$$

Therefore, there is

$$\boldsymbol{X}_{\alpha,\beta}^{\max}\cap\boldsymbol{X}_{\alpha',\beta'}^{\max}=r(C,\sigma^1,\cdots,\sigma^\kappa) \tag{11}$$

$$\boldsymbol{X}_{\alpha,\beta}^{\max}+\boldsymbol{X}_{\alpha',\beta'}^{\max}=r(C,\sigma^1,\cdots,\sigma^\kappa,\sigma^\lambda,\cdots,\sigma^m,\sigma^\eta,\cdots,\sigma^n) \tag{12}$$

$$\begin{aligned}\varphi(\boldsymbol{X}_{\alpha,\beta}^{\max}+\boldsymbol{X}_{\alpha',\beta'}^{\max})&=\varphi(r(C,\sigma^1,\cdots,\sigma^\kappa,\sigma^\lambda,\cdots,\sigma^m,\sigma^\eta,\cdots,\sigma^n))\\&=\varphi(r(C,\sigma^1,\cdots,\sigma^\kappa,\sigma^\lambda,\cdots,\sigma^m))\\&+\varphi(r(C,\sigma^1,\cdots,\sigma^\kappa,\sigma^\lambda,\cdots,\sigma^n))\\&+\varphi(r(C,\sigma^1,\cdots,\sigma^m,\sigma^\eta,\cdots,\sigma^n))\\&+\varphi(r(\sigma^\lambda,\cdots,\sigma^m,\sigma^\eta,\cdots,\sigma^m))\\&-\varphi(r(C,\sigma^1,\cdots,\sigma^\kappa))-2\varphi(r(\sigma^\lambda,\cdots,\sigma^m))\\&-2\varphi(r(\sigma^\eta,\cdots,\sigma^n))\end{aligned} \tag{13}$$

Because the maximal gradient vector of the $\boldsymbol{X}_{\alpha,\beta}^{\max}$ leaf vector is $\boldsymbol{x}_{\alpha,\beta}=\sigma^{m-1}\sigma^m$， leaf vector $\boldsymbol{x}_{\alpha,\beta}$ no posterior relay vectors，That's singular σ^m is not the starting point of any base vector，then the base vector $\sigma^m\sigma^n$ non-existent

$$\begin{aligned}&\varphi(r(\sigma^\lambda,\cdots,\sigma^m,\sigma^\eta,\cdots,\sigma^n))\\&=\varphi(r(\sigma^\lambda,\cdots,\sigma^m))+\varphi(r(\sigma^\eta,\cdots,\sigma^n))\end{aligned} \tag{14}$$

According to the test data in Table 1, the average equal error rate EER of CGV method is 2. 52%, 0. 36% higher than the first algorithm P066; the average registration time is 0. 35s, 0. 23s less than P066; the average matching time is 0. 25 s, 0. 55 s less than P066, and the average registration time and average matching time rank third, only higher than the algorithms P045, P017 and P074.

According to the test data in Table 1, in the results, the CGV method we used had similar recognition accuracy to the top ten algorithms, while the registration time and matching time of the algorithm were reduced. Mainly because the composite gradient vector method of the discrete features in the fingerprint image combined as the basis of target matching, maintain the high recognition accuracy, at the same time using the hierarchical marker rules for the composite gradient vector, reduce the fingerprint retrieval time, and the composite gradient vector itself less information, the matching time is short. It can be seen that using the composite gradient vector method for fingerprint recognition can achieve a fast calculation speed while ensuring a high recognition accuracy.

3.2. Verification of the algorithm based on Model 1

The title requires the proportion of 80%, 90%, 95% and 97% images in the dataset, the number of "same" fingerprint matching pairs and the total number of "same" fingerprint matching pairs (500 pairs, TZ_. t t x t).Second, question 3 requires us

The _Data.txt after T Z file 200_ is retrieved in the same way as problem 2, and the retrieval results are given, and the result data is compressed into.zip format and uploaded to the calculation results of the competition system.As the algorithm is the same, here we treat problems two and three together.

In order to quantify the actual effect of the algorithm, we use two important parameters to evaluate the performance of the fingerprint recognition algorithm: error rejection rate FRR (false rejection rate) and error acceptance rate FAR (false acceptance rate), and formulas (22) and (23) give two parameters.

After multiple validation, we list the validation results as follows Table 1.

Table 1 Validation Results

Filter out the image proportions in the dataset（%）	The number of "same" fingerprint matches remains as a proportion of the total number of "same" fingerprint matches（%）
80	1.42
90	1.31
95	1.27
97	1.26

The following identification accuracy verification, which can be calculated with the following formula:

$$A_{CC}=\left(1-\frac{FAR+FRR}{2}\right)\times 100\% \qquad (15)$$

Both the error rejection rate FRR and the error acceptance rate FAR are relatively low, and the system identification accuracy is high, which illustrates the effectiveness and robustness of the algorithm.For question 3, we also performed an algorithm test and saved the results in "result_90", "result_95", and "result_97".At this point, the problem two, three is solved.

3.3. The best retrieval method of exploration

For this question, we reconsider a new fingerprint matching model.For the fingerprint image to be matched, the image is first preprocessed to refine the image.The refined fingerprint image can be extracted and compared.

Fingerprint preprocessing is the basic link of fingerprint recognition, mainly including fingerprint image prospect extraction, image enhancement, binarization, denoising and refinement and other processes.We establish block-based methods and adaptive ideas for fingerprint preprocessing, dividing the whole image into blocks and preprocessing it block by block.Compared with the traditional method, the simple method can effectively enhance the effect of fingerprint preprocessing.An overview of the entire preprocessing method is given below by Figure 3.

Figure 3 Original image of fingerprint (left) and extracted foreground (right)

The ultimate goal of fingerprint preprocessing is to obtain the macroscopic description of the fingerprint topology and the structure of the fingerprint detail points, so the fingerprint image

ISSN: 2664-9640

needs to be binarized and refined.We first calculate the direction field to obtain the binarized image of the fingerprint, and finally use the table checking method again to refine the binarized image.

According to statistics, for local details, the frequency of the fork and endpoint features of fingerprints accounts for more than 90% of the sum of all feature frequencies.For the global feature, the center point and the triangle point of the fingerprint play a very important role in the alignment and positioning of the fingerprint.Based on this, this study focuses on the above characteristics.

The fingerprint fork refers to the fork point when a line divides into two, while the end point is the beginning or end point of a line.Through a large number of summaries, the description model of the fingerprint endpoint and fork point can be better established. As showing in Figure 4.

Figure 4 Endpoint model (left) and fork point model (right)

If there is a singularity in a region, the directional field around this point will change dramatically.Therefore, calculating the size of the directional field around the point can determine whether it is a singularity.This method is evolved from the Poincare formula invented by the famous mathematician Poincare (Poincare).The calculation formula is as follows:

$$\mathrm{pioncare}(t_x, t_y) = \frac{1}{2p}\sum_{k=0}^{N}\Delta(k) \tag{16}$$

Inside

$$\Delta(k) = \begin{cases} \delta(k), & if\ |\delta(k)| < \dfrac{\pi}{2} \\ \pi + \delta(k), & f\ |\delta(k)| \le -\dfrac{\pi}{2} \\ \pi - \delta(k), & otherwise \end{cases}$$

$$\delta(k) = \mathrm{Ori}(p_x(i'), p_y(i')) - \mathrm{Ori}(p_x(i), p_y(i))$$

$$i' = (i+1) \bmod N$$

Since the highest screening volume of this question is 97%, and for more than 97% of the screening volume, we can make better improvements in the retrieval accuracy, retrieval time and memory footprint.In order to achieve a better comparison effect, the method of cascade decision comparison is adopted here. The central point matching is conducted first. If the central point matching is not successful, the middle region matching is continued, and if it fails, the global matching is conducted finally.In this way, the unsuccessful phenomenon of fingerprint matching of the same finger is effectively avoided.

For pattern matching, the quality of matching objects (fingerprint, face, iris, etc.) will have a great impact on the matching results, especially for fingerprint image, if the fingerprint texture is not clear or the fingerprint key details are damaged, it will cause serious interference to fingerprint-based authentication.Our new method has good performance advantages over other fingerprint alignment algorithms.

It can be intuitively seen that the cascade alignment method proposed in this paper can achieve the best results in all three cases.The identification performance is about 4% higher than the thread feature alignment method and about 6% higher than the structure alignment method.The experimental results demonstrate the feasibility and correctness of our proposed cascade decision matching method.

4. Analysis of the model advantages and disadvantages

4.1. Advantages

For general fingerprint identification, the offset and rotation phenomenon of misidentification problem has good identification effect, at the same time for the nonlinear deformation caused by feature compression, image stretching has strong adaptability, at the same time, due to the hierarchical marking rules for the composite gradient vector information, accelerate the speed of fingerprint retrieval process, improve the processing speed of fingerprint matching algorithm.

It has the advantages of fast recognition speed and high accuracy in the process of fingerprint recognition.For the fingerprint image with large deformation, it has strong adaptability, not only high recognition accuracy, but also fast processing speed, and strong resistance to nonlinear deformation.

4.2. Disadvantages

In the actual use, you should register a fingerprint first. If the registered fingerprint does not judge the quality, it is easy to cause the situation that always fails to pass during the verification.From this perspective, fingerprint quality judgment has a very practical value and application significance.The line structure information considers the similarity between lines, but because the fingerprint itself is composed of a large number of lines, and the edge boundary of the line is not very clear, it is easy to cause the integration of different lines or agree to divide the lines into two parts.

References

[1] Wang Chao. Research on the wireless fingerprint matching localization algorithm based on machine learning [D]. Xidian University, 2020.DOI:10.27389/d.cnki.gxadu. 2020.000785.

[2] Wang YiTeng. Study on small-area fingerprint image recognition and matching algorithm [D]. Xidian University, 2019.DOI:10.27389/d.cnki.gxadu. 2019.003021.

[3] Wang Yiming. Fingerprint quality assessment and matching algorithm study [D]. Shandong University, 2017.

[4] Bian Zhi. Research on fingerprint matching localization technique based on machine learning algorithm [D]. Beijing University of Posts and Telecommunications, 2017.

[5] Che Ronglu. Feature extraction and matching algorithm study of fingerprint images [D]. Yanshan University, 2014.

[6] Pingjianchuang. Research on fingerprint recognition preprocessing and matching algorithms [D]. University of Electronic Science and Technology, 2013.

[7] Hu Chaofeng. Fingerprint image preprocessing and matching algorithm research [D]. Wuyi University, 2012.

[8] Guike. Research on Fingerprint Image Processing and Matching algorithm [D]. Wuhan University of Technology, 2010.

[9] Xu Dandan. Deep learning-based multifeature fingerprint image recognition [D]. Fuyang Normal University, 2021.DOI:10.27846/d.cnki.gfysf. 2021.000121.

ISSN: 2664-9640

[10] Ren Yi. Research and implementation of the fingerprint recognition system [D]. Nanjing University of Posts and Telecommunications, 2019.DOI:10.27251/d.cnki.gnjdc. 2019.01376.