8-4

MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

Identifying Hand Gesture Images by Using Genetic Algorithms

Tetsuji KOBAYASHI and Norifumi MACHIDA

Nippon Institute of Technology
Department of Computer and Information Engineering
4-1-1 Gakuendai, Miyashiro-machi, Saitama-ken, 345-8501 Japan

Abstract

A genetic algorithm (GA) is an optimization algorithm that simulates the hereditary phenomenon of natural life. Although GA has been applied to image processing, it has not been studied to apply for matching hand gesture images. In this paper, a method to use GA for the matching process of hand gesture images expressing numbers is proposed. The processing algorithm including GA for hand gesture images is evaluated experimentally by changing the parameters of GA. It is shown that each person can be distinguished from other persons by the fitness value of GA. Hand gesture images to express each number are used for evaluation. Next, the personal identification to search the most suitable hand gesture image in the database of hand gesture images by using GA is proposed and evaluated experimentally. As a result, the effectiveness of the proposed method using GA for matching hand gesture images and identifying each person is confirmed.

1 Introduction

A genetic algorithm (GA) applies the evolutionary process of genes to optimization problems. Although GA has been applied to image processing, it has not been studied to apply for the processing of hand gesture images [1], [2]. The template matching is one of widely used techniques for matching images.

In this paper, GA is applied for processing hand gesture images. The parameters of GA are changed to many values, and the processing algorithm including GA for hand gesture images is evaluated. As for hand gesture images expressing numbers, each number is expressed by using each hand pattern. As an application for personal identification, a person can be distinguished from other persons by the fitness value of GA by searching the database of hand gesture images.

2 The Proposed Method by Using GA

2.1 The overview of GA for matching images

Although GA has been studied for matching images, the application for hand gesture images has not been studied. Therefore, we propose a method to apply GA for matching hand gesture images expressing numbers.

The overview of GA for matching hand gesture images is shown in Fig. 1. The judgement means that the fitness value is compared to the threshold value in Fig. 1.

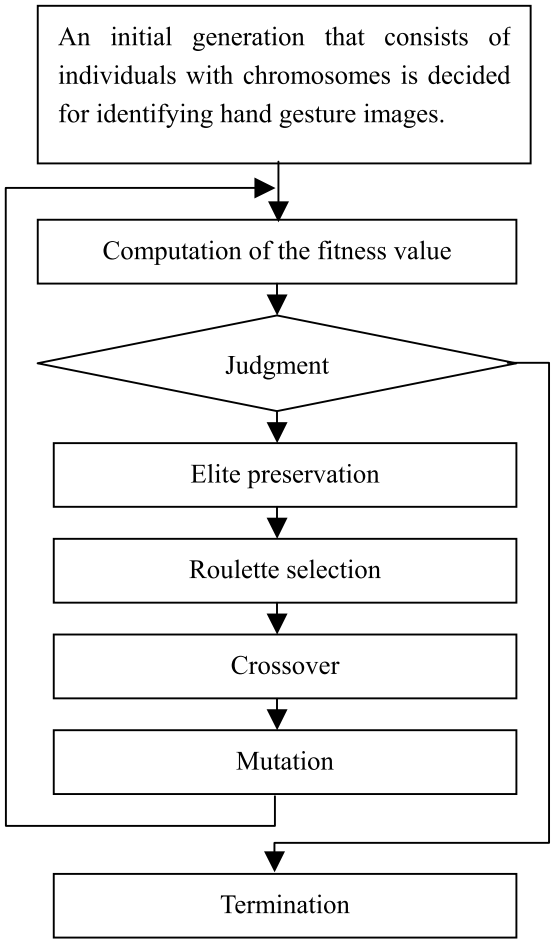

Figure 1. The overview of the genetic algorithm (GA) for matching hand gesture images.

2.2 Parameters of GA for hand gesture images

The algorithm of the generation change of GA uses the crossover rate, the mutation rate, the rate of elite preservation, and the rate of roulette selection. The total number of generations and the total number of individuals in each generation are also considered as the parameters of GA.

An example of the GA parameter values that are used for the change of generations in the experiment are as follows. The total number of each individual is 100, the total number of generations to stop the change of generations is 100, the mutation rate is 0.1, the crossover rate is 0.8, and the crossover scheme is the two-point crossover.

In order to decrease the matching time, the initial set of chromosome values is set at the center of the image plane, because the optimal position is near the center of the image plane for the matching of two hand gesture images.

When GA is used for matching a pair of two hand gesture images, the chromosome is a set of {x-coordinate, y-coordinate, scaling ratio, rotation angle} for an image, and each of them has 8 bits.

2.3 Fitness values of GA for matching images

A fitness value of GA is used to check the similarity degree between a registered grayscale hand gesture image and an input grayscale hand gesture image. The fitness value can change from 0 to 1. To compute the fitness value, the residual correlation and the normalized correlation are independently used for comparison. When the fitness value between two grayscale hand gesture images is greater than the threshold value, these two grayscale hand gesture images are regarded as the same hand gesture pattern expressing the same number of the same person.

(1) The residual correlation for the fitness value

The fitness value defined by the residual correlation is shown in equation 1.

$$Fitness\ value = 1 - \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left|g(x+i, y+j) - f(i,j)\right|}{M \times N \times 255} \quad (1)$$

where, $f(i, j)$ is a pixel value of the registered image, $g(x+i, y+j)$ is a pixel value of the input testing grayscale image, M is the width and N is the height of the registered grayscale image.

(2) The normalized correlation for the fitness value

The fitness value defined by the normalized correlation is shown in equation 2.

$$Fitness\ \ value = \frac{\sum_{i,j}\left\{\left(f_{ij} - f_m\right)\times\left(g_{ij} - g_m\right)\right\}}{\sqrt{\sum_{i,j}\left(f_{ij} - f_m\right)^2} \times \sqrt{\sum_{i,j}\left(g_{ij} - g_m\right)^2}} \quad (2)$$

where f_{ij} is a brightness value of pixel address (i, j) of a registered grayscale hand gesture image, g_{ij} is a brightness value of pixel address (i, j) of an input testing grayscale image, f_m is an average brightness value of a registered grayscale image, and g_m is an average brightness value of an input testing grayscale image.

2.4 Expressing numbers by hand gestures

It is possible to consider many sets of finger usages to indicate each number by each hand gesture. The requirements for the usage to express numbers by means of hand gestures are as follows. Almost every user can easily express all numbers, and learn how to express numbers. The system should be easily implemented. The error rate should be low for personal identification. All of the numbers of 0-9 should be available. However, each person should be able to decide each set of hand gestures expressing numbers. An example of the set of hand gestures expressing numbers is shown in Fig. 2.

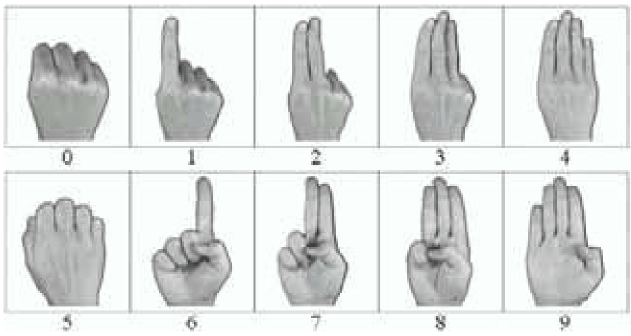

Figure 2. Hand gestures expressing numbers.

3 The Evaluation of the Proposed Method Using GA

3.1 The variation of the maximum fitness values

A personal computer is used for the evaluation of the proposed matching method using GA. An example of the size of a hand gesture image is 201 by 375 pixels, and 76 kilobytes in PGM (portable graymap) format. Figure 3 shows the variation of the maximum fitness values when the residual correlation is used. Figure 4 shows the variation of the maximum fitness values when the normalized correlation is used. In these figures, the solid line shows the matching between hand gesture images of the same person, and the dotted line shows the matching between hand gesture images of different persons.

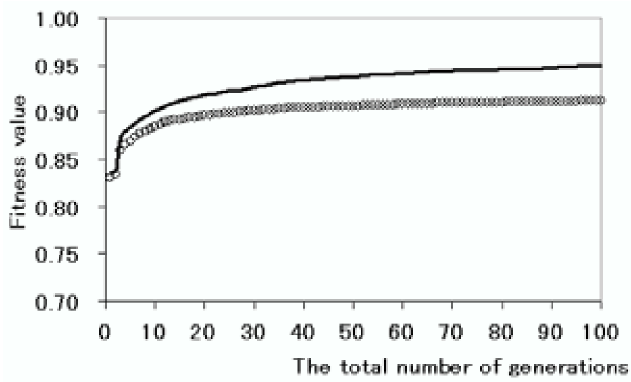

Figure 3. The variation of the maximum fitness values when the residual correlation is used.

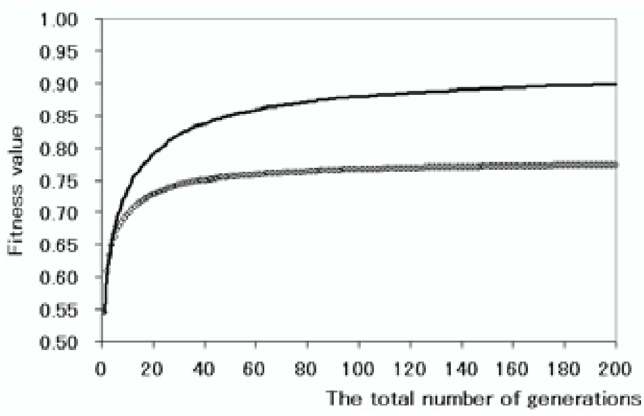

Figure 4. The variation of the maximum fitness values when the normalized correlation is used.

According to Fig. 3 and Fig. 4, the maximum fitness value between a pair of hand gesture images of the same person is larger than that of different persons. This is because the similarity between hand gesture images of the same person is larger than that of hand gesture images between different persons. The convergence time depends on each number expressed by each hand gesture image. The convergence time when the residual correlation is used for the fitness value is smaller than the convergence time when the normalized correlation is used for the fitness value. The matching time depends on the CPU speed. For example, when the residual correlation is used for the fitness value, the matching time for a pair of hand gesture images is 4.777 milliseconds, and when the normalized correlation is used, the matching time is 14.941 milliseconds.

3.2 Error rates

FAR (False Acceptance Rate) and FRR (False Rejection Rate) are used for error rates. The false acceptance error means that a non-registered user is falsely determined as a registered user. The false rejection error means that a registered user is falsely determined as a non-registered user. Figure 5 shows the error rate versus the threshold value for matching when the residual correlation is used for the fitness value. Figure 6 shows the error rate versus the threshold value for matching when the normalized correlation is used for the fitness value.

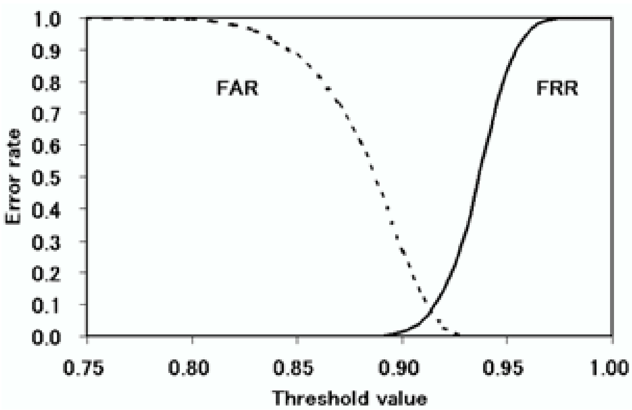

Figure 5. The error rate versus the threshold value for matching when the residual correlation is used for the fitness value.

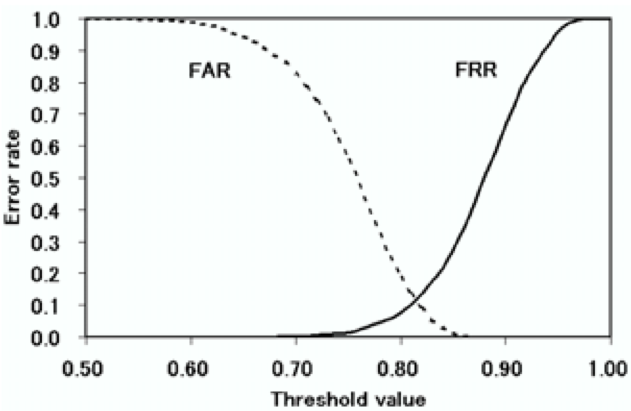

Figure 6. The error rates versus the threshold value for matching when the normalized correlation is used for the fitness value.

By means of Fig. 5 and Fig. 6, we obtain Table 1 and Table 2. Table 1 shows the FRR when FAR=0, the FAR and FRR when FAR=FRR, and threshold values, in the case that the residual correlation is used for the fitness value. Table 2 shows the FRR when FAR=0, the FAR and FRR when FAR=FRR, and threshold values, in the case that the normalized correlation is used for the fitness value. The threshold values in Table 1 and Table 2 can be used to decide the threshold values for personal identification.

According to Table 1 and Table 2, the error rate when the residual correlation is used for the fitness value is larger than the error rate when the normalized correlation is used for the fitness value.

Table 1. FRR, FAR and threshold values, when the residual correlation is used for the fitness value.

Each number expressed by each finger.	FRR when FAR = 0.	Threshold value.	FAR and FRR, when FAR=FRR.	Threshold value.
0	0.296	0.930	0.053	0.915
1	0.720	0.950	0.093	0.925
2	0.452	0.940	0.119	0.915
3	0.404	0.930	0.081	0.915
4	0.280	0.930	0.098	0.915
5	0.272	0.930	0.090	0.915
6	0.180	0.920	0.062	0.905
7	0.188	0.920	0.052	0.905
8	0.264	0.920	0.058	0.895
9	0.600	0.940	0.094	0.915
Average value.	0.366	0.931	0.080	0.912

Table 2. FRR, FAR, and threshold values, when the normalized correlation is used for the fitness value.

Each number expressed by each finger.	FRR when FAR = 0.	Threshold value.	FAR and FRR, when FAR=FRR.	Threshold value.
0	0.400	0.840	0.130	0.765
1	0.120	0.870	0.043	0.855
2	0.096	0.880	0.033	0.865
3	0.112	0.850	0.070	0.825
4	0.200	0.850	0.051	0.825
5	0.336	0.840	0.072	0.785
6	0.144	0.860	0.065	0.835
7	0.120	0.860	0.033	0.825
8	0.104	0.850	0.043	0.825
9	0.224	0.890	0.098	0.865
Average value.	0.186	0.859	0.064	0.827

4 Personal Identification by Using GA

4.1 The proposed method for identifying persons

The proposed procedures for personal identification consist of the registration and identification procedures. At first, a user registers the personal identification data to the system by using the registration procedure. Hereafter, the user can be identified in the identification procedure.

An example of the personal identification procedures are described as follows.

[Registration procedure]

Step R1: The administrator of the system inputs the user ID (identification) number with more than one digit by using the keyboard.

Step R2: The user shows a hand gesture expressing the last digit of the user ID number. The system captures the hand gesture image by using a video camera, stores the image file, and sets the file name as the user ID number.

[Identification procedure]

Step P1: The user shows only a hand gesture expressing the last digit of the user ID number. The system captures the hand gesture image by using the video camera.

Step P2: The system sequentially compares the input hand gesture image to the registered hand gesture images one by one in the database by using GA, and finds the most similar registered hand gesture image. As a result, the user who has inputted the hand gesture image is decided.

4.2 Evaluation of the personal identification

In order to investigate the effect of GA for one-to-many matching that is used in the identification procedure, the following experiment has been executed. The system has many hand gesture images in the database. When a new hand gesture image is shown, the system finds the hand gesture image with the best fitness value for it. As a result, the person with the new hand gesture image is decided.

Figure 7 shows an example of the change of the fitness values between the registered hand gesture image and the input hand gesture image for each generation.

According to Fig. 7, the converged fitness value is obtained at the 65th generation for number 0, the 30th generation for number 1, the 30th generation for number 2, the 65th generation for number 3, the 65th generation for number 4, the 45th generation for number 5, the 30th generation for number 6, the 30th generation for number 7, the 40th generation for number 8, and the 40th generation for number 9. There appears an effect to speed up the convergence of generation change by using the number expressed by hand gestures.

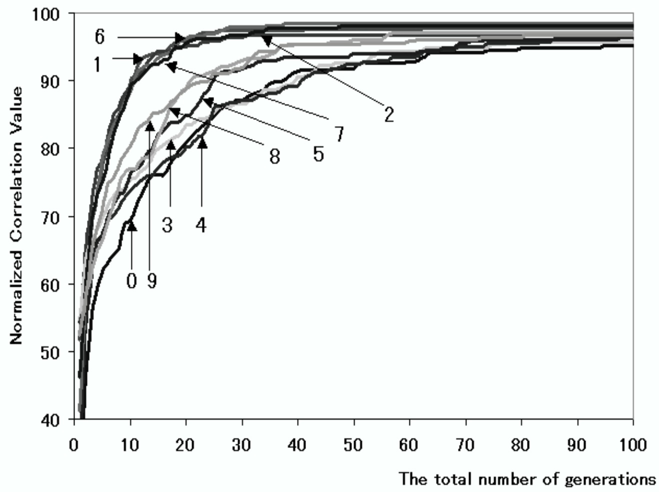

Figure 7. The change of the fitness values between a registered grayscale hand gesture image and the input grayscale hand gesture image for each generation, where each number is expressed by each hand gesture image, and the normalized correlation is used for the fitness value.

5 Discussion

5.1 Merits of the proposed method using GA

The general difficulties of optimization by GA are resolved in the proposed method using GA as follows.

(1) In general, when there are multiple extreme values, the local maximum value may be obtained in the optimization by GA. In the proposed method using GA to hand gesture images, even if a local maximum is obtained, there is no difficulty, because the final decision is judged by using the threshold value.

(2) In general, the convergence time for matching is long in the optimization by GA. In the proposed method using GA to hand gesture images. The number of individuals that consists of one generation of GA is decreased in each generation. If a similar individual pattern of GA was appeared before, it is not used for matching.

(3) In general, it is difficult to decide when the generation change shall be stopped in the optimization by GA. In the proposed method using GA to hand gesture images, either of the following cases is used to stop the generation change, the case when the fitness value becomes larger than the threshold value, or the case when the total number of the generation changes becomes larger than the pre-specified value.

5.2 Comparison of fitness values

Either the residual correlation or the normalized correlation can be used for the fitness value. The matching time of the residual correlation is smaller than that of the normalized correlation. However, the error rates (FAR, FRR) become smaller when the normalized correlation is used. Therefore, when the system has enough CPU speed, the normalized correlation is better than the residual correlation for the fitness value of GA.

6 Conclusion

A method to use genetic algorithms (GA) has been proposed for the matching process of hand gesture images expressing numbers. The fitness values and parameters for using GA to the matching process are defined. The GA can be applied to both the one-to-one matching for positioning a pair of two hand gesture images and the one-to-many matching for finding the best suitable hand gesture image in the database for personal identification. As a result, the effectiveness of the proposed method using GA has been confirmed experimentally.

References

[1] M.Takagi and H.Shimoda (ed.): Handbook of Image Analysis, Revised Edition, Tokyo University Publishing, 2004.

[2] N. Machida and T.Kobayashi: Proceedings of the Ninth IAPR Machine Vision Applications (MVA2005), IAPR(International Association for Pattern Recognition) & AIST(National Institute of Advanced Industrial Science and Technology), Tsukuba, Japan, pp.294-297, 2005.