Parameter Optimization for biometric Fingerprint Recognition using genetic Algorithms

Tobias Scheidat[1], Andreas Engel[2], Claus Vielhauer[1]

Otto-von-Guericke University of Magdeburg
Advanced Multimedia and Security Lab
Magdeburg, Germany
[1]{tobias.scheidat | claus.vielhauer}@iti.cs.uni-magdeburg.de
[2]andengel@cs.uni-magdeburg.de

ABSTRACT

In this paper, we suggest an optimization approach for fingerprint authentication using genetic algorithms. Our application was planned so that it can be used without great effort for different biometric systems. Instead of estimating the required parameters as in the case of some methods, here they are determined with the help of genetic algorithms. Our own test database consists of 1200 fingerprints of 12 persons. For the confirmation of the results, which were found out with this test set, the databases of the Fingerprint Verification Contests of the years 2000, 2002 and 2004 were examined in addition. In the best case an improvement in the recognition performance of 38% could be observed.

Categories and Subject Descriptors

C.4 [Performance of Systems]: *Measurement techniques, Performance attributes*

General Terms

Algorithms, Measurement, Performance, Security, Human Factors, Verification

Keywords

Biometrics, fingerprint, genetic algorithms, optimization

1. INTRODUCTION

The traditional ways of user authentication, secret knowledge and personal possession, have some disadvantages in the area of security. Both methods offer the possibilities of handing over, theft and losing of the authentication object and it isn't sure that the person who presents the object is also the legal owner. On the other side a biometric system authenticates the person itself by typical physiological and/or behavioral characteristics of human being. It is a disadvantage of biometric methods that there are not

two identical samples of a single biometrics of the same person. In order to achieve the desired security of a biometric application, many tests are necessary to determine the corresponding parameters. The parameters are estimated in many systems. In this paper we will introduce a method which determines optimal parameters for fingerprint recognition based on a genetic algorithm.

The use of human fingerprints for authentication purposes has a long tradition. It was already used at the end of the 19th century in the forensic field. The reasons for it are the uniqueness of an individual fingerprint and its pattern stability during a human's life.

In biometrics there are some approaches for the optimization of recognition results, which are based on the supervised data acquisition or the fusion of several biometric systems or algorithms, for example. Jain et al. describe in [1] several approaches to improve the performance of biometric fingerprint recognition. The authors show that a better authentication performance can be reached by sorting out low quality fingerprint images, habituating users, combining two fingers or two representations of the same finger. In [2] Scheidat et al. present a so-called multi-algorithmic system based on the fusion of four algorithms for online handwriting verification. By the fusion a relative improvement of the authentication performance of 12% in equal error rate (EER, see section 3.2) was reached in the best case in comparison to the best result of the single algorithms involved. A combination of several biometric modalities for performance improvement is also possible.

Today optimizations using evolutionary approaches can be found in various technical systems. The volumetric segmentation of medical brain images [3] or the optimization of aerodynamic shapes for supersonic airplanes [4] are just two of countless examples for problems that can be solved with genetic algorithms. Although there are some recent approaches to use genetic algorithms in biometric systems, as for example in face recognition [5], our approach focuses on the parameter optimization of general biometric verification algorithms that are not necessarily based on evolutionary methods. Such parameters are often very hard to estimate, yet there don't seem to be any publications available about biometric parameter optimizations in general.

The approach which we present here optimizes a fingerprint system by the use of a genetic algorithm. Through this the

automatic optimization of the parameters of the system is possible.

This paper is structured as follows: In the next section we discuss the fundamentals of the fingerprint recognition algorithm and genetic algorithm we used. Also the design of our genetic optimizer for biometric systems is presented in section 2. Section 3 gives an overview of the test database and the test methodology. The test results and a discussion of their meaning are shown in section 4. A short summary of this paper and an outlook of future work are given in section 5.

2. FINGERPRINT AND GENETIC ALGORITHMS

It is the aim of this work to optimize parameters of biometric authentication systems using a framework, based on a genetic algorithm, which can optimize parameters of biometric systems independently from their mode of operation. No extensive knowledge about the system shall be necessary to use it with the genetic optimizer. Initial tests with the genetic algorithm were already carried out which have led to encouraging results. These results are based on an algorithm for fingerprint recognition which was improved and implemented as an operating system login application.

2.1 BioGINA-Algorithm

Several different properties of fingerprint images can be used as features for verification. The matching of whirl structures and image correlation are commonly found in different biometric algorithms as well as minutiae based approaches which are used by the BioGINA login system. Its matching algorithm uses a set of 15 parameters to define allowed deviations for minutiae locations and orientations. In the matching process pairs of two minutiae in the verification sample are compared against minutiae pairs in the enrollment template. The first subset of the BioGINA parameters defines percentage values for the maximum allowed deviation for relative properties of minutiae like distance, ridge count between minutiae and angle between minutiae orientations. Measured deviations are then scaled into partial scoring values using a linear equation whose coefficients are defined in the second parameter subset. Further parameters define minimum partial score values for minutiae to be counted for the overall score value as well as maximum allowed deviations for global orientations of the fingerprint images. A more detailed description of the BioGINA algorithm and parameter set would go beyond the scope of this work but should not be necessary for understanding, since the presented method is supposed to work independently from the biometric matching algorithm and its parameters.

Overall the parameters are highly correlated and optimal values cannot be calculated analytically. Extensive experiments would be required for best possible calibration. Therefore the BioGINA parameter set has been based on intuitive estimation only. Performance evaluation using these default settings has shown an equal error rate (EER) of 3.13 %. Due to the high effort to implement new features or enhance the algorithm itself, ways had to be found to improve the system performance by parameter tuning.

The underlying algorithm based on the fingerprint matching method proposed in [6] by Yang and Verbauwhede. This method was optimized by an additional parameter describing the number of ridges between two minutiae points. For further and more detailed information regarding BioGINA and its parameters the interested reader is referred to [7].

2.2 Genetic Algorithm

The mechanisms of natural evolution such as natural selection, inheritance, mutation and recombination allow life forms to adapt to their environmental conditions by gradual random modifications of their genetic code. Similar strategies can be used to adapt parameters of a dynamic system to certain specified conditions. Genetic algorithms, first presented by Holland in 1975 ([8]), are a method based on natural evolution strategies to solve problems where analytic solutions do not exist or cannot be determined with realizable effort and time.

The first component for a genetic algorithm is a binary representation of a solution candidate, in this case any numeric parameter vector. In analogy to nature vector components are referred to as *genes* and vector representations as *chromosomes*. The *population* of a genetic algorithm is a set of such solution candidates, called *individuals*, which are initialized with random values. The population size has a significant influence on the convergence behavior. Small populations tend to produce local optima in very short time while higher population sizes increase the computational effort but also raise the chance to find global optimal values or better parameter sets in general.

A so-called *fitness function* calculates the quality of an individual as a numeric value by determining its conformance to the user specified conditions. The actual optimization is accomplished by an iterative reproduction process. In the first step a small set of individuals with higher fitness values is selected from the current population. Using these solution candidates a new generation will be created through genetic operators. Mutation operators randomly change single or multiple values in the chromosome, therefore providing a wide exploration of the search space. Crossover mechanisms exchange genes between two or more parent chromosomes to create new individuals. This way with a good chance two potentially good partial solutions are combined. The inheritance of genes that have a positive influence on the solution quality and the preferred selection of individuals with higher fitness values results in a slowly rising average fitness in the population by iteration of the reproduction process.

The selection and reproduction cycle is repeated until a user specified termination condition is reached, such as a solution is found that satisfies the target criteria, an allocated amount of time has been spend or the algorithm did not create any new better solutions for a certain period of time.

2.3 Design of the genetic Optimizer

In order to allow optimization of the BioGINA algorithm the biometric interface has to be extended by two functions to control its parameter settings. First the genetic algorithm has to be informed about number and range of values for all parameters. A second function is provided to change the current parameter set. This allows the genetic algorithm to be implemented in an independent external application so it will be possible to exchange the biometric interface with minimal effort enabling parameter optimizations on other biometric systems and other modalities. Figure 1 shows the basic architecture of the genetic optimization system for general biometric verification algorithms.

To determine the fitness of a specific parameter set, biometric error rates (see section 3.2) of the system can be directly transferred to a fitness value. This way it represents an inverted quality measure. Therefore the goal of the optimization process will in this case be the minimization of the fitness value. For all tests performed within the scope of this article the equal error rate (EER) will be used.

To estimate error rates for fitness calculation a biometric database is needed. Only a sufficient count of biometric samples allows a reliable training process. To increase the number of tests for the error rate estimation, every possible combination of authentic and imposter samples will be considered. To reduce computation time different methods are used to estimate fitness values with highest possible precision without the necessity to calculate matching scores for all available samples. Therefore sample pairs with very low effects on the resulting error characteristics are dynamically skipped, limiting calculation time to the matching duration for samples with expected high impact on the fitness value.
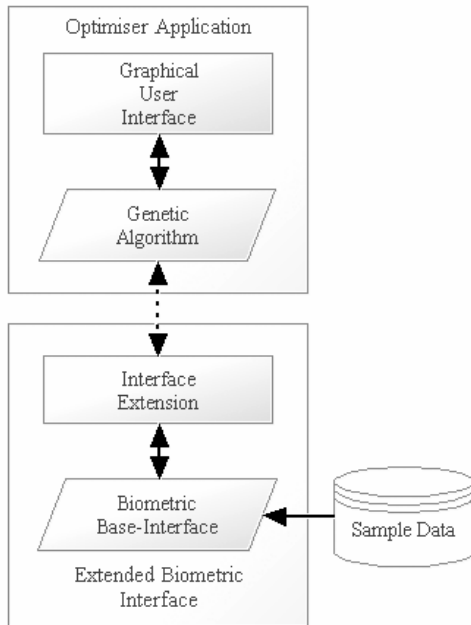

Figure 1. Architecture of the genetic optimizer for biometric systems

The optimization process starts with the creation of an initial population with randomly selected parameter sets. For a wide exploration of the search space, about 10.000 individuals should be created. After determining fitness values for the initial population, a smaller set of the best solution candidates is selected to be the population for the actual optimization iteration. The optimal population size depends on the speed of the biometric matching algorithm and was set to 100 for the BioGINA algorithm.

The final optimization process is done by a modified form of the steady state genetic algorithm which creates exactly one new individual in each generation using the following scheme:

1. Fitness based selection of one or two parent individuals

   A predefined number of individuals of existing population is selected to generate one new individual for next generation, suitable candidates are determined by a fitness function.

2. Mutation or crossover of the selected parents

   Creation of one new individual based on one parent (mutation) or two parents (crossover) using reproduction rules.

3. Calculate new fitness value

   The fitness value of the new individual will be determined.

4. Replace a low quality individual in the current population

   In order to create the next generation a low quality individual will be replaced by the new one.

5. Repeat steps 1 – 5 until termination condition is reached

   Depending on the fulfilling of the predefined termination condition(s) steps 1 to 5 are repeated.

The creation of exactly one individual in each generation by using only one genetic operator enables analysis of the current efficiency of each operator and could be used for adaptation of their selection probability and other related parameters. Strict implementation of the 4th step also ensures that already found current best solutions will not accidentally get lost until at least one better candidate is found.

Sorting all individuals by their fitness value and definition of a monotonic decreasing probability distribution over them allows a simple rank based selection mechanism. Unlike direct fitness proportional approaches, this technique eliminates the problem of dominant individuals with much higher-than-average fitness values preventing other individuals from being selected.

After selection one of five different genetic operators is used, which all have the property never to leave the search space. The first is a single mutation that randomly selects and modifies one gene. A full mutation operator modifies all genes by smaller values. The primary recombination operator is a uniform crossover which randomly exchanges genes between two parent individuals. A second arithmetic crossover operator calculates the weighted sum of genes of two parents. A fifths pseudo operator randomly creates new solution candidates without using any parent individuals in order to continuously increase the exploration of the search space.

Termination of the genetic optimization process will be performed manually after at least 500 generations without measurable improvements.

3. EXPERIMENTAL SETUP
In this section the fingerprint databases and the test methodology are explained, which are used for parameter optimization and evaluation. Furthermore the measurement of error rates in the biometric context is described.

3.1 Test Database
For the tests a new fingerprint database for the BioGINA algorithm was created. In this database 10 fingerprints each were taken by all 10 fingers of 12 test persons. The data were acquired with a capacitive Infineon FingerTIP sensor. The graphical data have a size of 224 x 288 pixels and a resolution of 513 dpi. For the tests the database was subdivided into two test sets, set *A* and set *B*, containing 60 fingerprints each with 10 Samples. Set *B* was basis for the calculation of an optimal parameter vector using genetic algorithms. Set *A* on the other side was used for the evaluation of the generated parameter set. Through this a strict

separation of the training data and the evaluation data was reached and an over fitting avoided.

In order to cross-validate our initial results, the optimized parameters were also tested with the data of the Fingerprint Verification Competitions (FVC) from the years 2000, 2002 and 2004. For each of the single competitions four databases were produced. The first three are based on real fingerprints which were taken by three different sensors with varying characteristics. The data of the last database were generated by a fingerprint software synthetically. The databases are also subdivided into two subsets, where set *B* (8 samples of 10 fingerprints) contains data for parameter tuning and set *A* (8 samples of 100 fingerprints) contains the evaluation data. Further information about the Fingerprint Verification Competitions can be obtained from [9], [10] and [11].

3.2 Methodology
Since the training sets of the Fingerprint Verification Competition databases turned out to be too small to provide reliable results, the parameter set trained on the BioGINA database has been used for performance evaluation on all databases.

In order to compare the authentication performance of the algorithm using the different parameters, biometric error rates are used. The false non-match rate (FNMR) specifies how often authentic persons are rejected from the algorithm. How frequently non-authentic persons are accepted by the algorithm is indicated by the false match rate (FMR). The point of our interest in the error rate characteristics is the equal error rate (EER), where both error rates, FNMR and FMR, yield the same value.

4. EXPERIMENTAL RESULTS
On an average computer system where the BioGINA matching algorithm is able to execute about 200 verifications per second, results with lesser required solution quality can be obtained within a few hours. The runtime for one full optimization process adds up to about 1 - 2 days. For multiple independent training processes to minimize the EER all resulting values were very close, allowing to conclude that the optimizer was actually nearing the global optimum. On the other hand major differences in the resulting parameter set values prove the functionality of the genetic algorithm even for systems with very high parameter correlation.

The second row of Table 1 titled *standard* shows the EERs of the training set *B* and the evaluation set *A* using standard parameter set without genetic optimization on the BioGINA database. In the third row (*optimized*) the EERs for set *A* and set *B* are presented, which are determined using the parameter set optimized by our genetic algorithm. For each set the EER based on the optimized parameters is better than the EER without optimization. The relative improvement is approximately 25% for set *A* and 41% for set *B*, respectively.

Table 1. EERs using BioGINA database

Parameter set	Set B (training)	Set A (evaluation)
standard	0.0283	0.0313
optimized	**0.0167**	**0.0233**

Table 2. EER using FVC databases of 2000, 2002 and 2004

Parameter set	DB1	DB2	DB3	DB4
FVC 2000 database				
standard	0.0620	0.0281	0.0754	0.0793
optimized	**0.0369**	**0.0232**	**0.0671**	**0.0552**
FVC 2002 database				
standard	0.0898	0.0739	0.1517	0.0505
optimized	**0.0759**	**0.0704**	**0.1165**	**0.0379**
FVC 2004 database				
standard	0.0783	0.0763	0.0495	0.1074
optimized	**0.0633**	**0.0611**	**0.0469**	**0.0820**

Table 2 presents the EERs based on the databases of the biometric Fingerprint Verification Competitions (FVC) of the years 2000, 2002 and 2004 using standard parameters (rows *standard*) and the parameters optimized genetically (rows *optimized*). The results are divided according the used individual databases of each competition. The optimized parameter set for each evaluation is based on the BioGINA training set *B*, because the size of the FVC training sets are not adequate for the determination of optimal parameters using the genetic algorithm. For every single database of every fingerprint competition an improvement in the results could be reached by using the genetic optimizer. The best result based on the optimized parameter set was determined on the database 1 (DB1) of the FVC 2000. Here the equal error rate amounts 0.0620 for standard parameters and 0.0369 for the optimized parameter set. This corresponds to a relative improvement of approximately 40%. The lowest relative improvement is calculated on database 2 (DB2) of the FVC 2002 with about 4.7%. The individual EERs are 0.0739 for the standard and 0.0704 for the optimized parameter set.

5. CONCLUSIONS AND FUTURE WORK
An optimization strategy for biometric parameters using genetic algorithms was presented in this paper. For all test databases, new acquired BioGINA database and the databases of the past Fingerprint Verification Contests, an improvement was reached. Our genetic optimizer initially uses a training set based on the BioGINA database to calculate a new parameter set and test have been additionally performed for public FVC databases of the years 2000, 2002 and 2004. These parameters were used for verification tests on the individual databases. For the BioGINA evaluation a relative improvement of 25% was reached. In the best case a relative improvement of the equal error rate was obtained of 40% for database 1 of the FVC 2000.

In our further work we will test the genetic algorithm, introduced here, at additional biometric systems. After the encouraging results of the initial tests based on fingerprints, we will carry out evaluations on dynamic biometric modalities (handwriting and speech) soon.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Jain, A.K., Prabhakar, S., Ross, A. Fingerprint Matching: Data Acquisition and Performance Evaluation. MSU Technical Report TR99-14, 1999.

[2] Scheidat, T., Vielhauer, C., Dittmann, J. Distance-Level Fusion Strategies for Online Signature Verification. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Amsterdam, The Netherlands, 2005.

[3] Fan, Y., Jiang, T., Evans, D.J. Volumetric Segmentation of Brain Images Using Parallel Genetic Algorithms. In IEEE Transactions On Medical Imaging, Vol. 21, no. 8, August 2002, pp. 904-909.

[4] Sasaki, D., Morikawa, M., Obayashi, S., Nakahashi, K. Aerodynamic Shape Optimization of Supersonic Wings by Adaptive Range Multiobjective Genetic Algorithms. Lecture Notes In Computer Science; Vol. 1993, Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization 2001, pp. 639- 652.

[5] Sun, Y., Yin, L. A Genetic Algorithm Based Feature Selection Approach for 3D Face Recognition. In Proceedings of Biometric Consortium Conference BC 2005, Arlington, VA USA, 2005.

[6] Yang, S., Verbauwhede, I.M. A Secure Fingerprint Matching Technique, In: Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications, pp. 89 94, 2003.

[7] Engel, A. Genetische Algorithmen zur Parameteroptimierung biometrischer Verfahren - Konzeption, Implementierung und Evaluierung, Master Thesis, Otto-von-Guericke University of Magdeburg, Department of Computer Science, February 2006 (in German).

[8] Holland, J.H. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology. Control and Artificial Intelligence, MIT Press, 1995, First Published by University of Michigan Press 1975.

[9] Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K. FVC2000: Fingerprint Verification Competition, http://bias.csr.unibo.it/fvc2000/Downloads/fvc2000_report.p df

[10] Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K. FVC2002: Second Fingerprint Verification Competition, in proceedings of International Conference on Pattern Recognition (Quebec City, August 2002), vol.3, pp. 811-814.

[11] Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K., FVC2004: Third Fingerprint Verification Competition, in proceedings of Biometric Authentication, First International Conference, ICBA (Hong Kong, China, July 2004), Lecture Notes in Computer Science, Volume 3072, pp. 1-7.