Soft-Biometrics: Soft-Computing for Biometric-Applications

Katrin Franke [1*], Javier Ruiz-del-Solar [2], Mario Köppen [1]
[1] Dept. of Pattern Recognition, Fraunhofer IPK, Berlin, Germany,
E-mail: {katrin.franke|mario.koeppen}@ipk.fhg.de
[2] Dept. of Electrical Engeneering, Universidad de Chile, Santigo, Chile,
E-mail: jruizd@cec.uchile.cl

Abstract

Biometrics, the computer-based validation of a persons' identity, is becoming more and more essential due to the increasing demand for high-security systems. A biometric system testifies the authenticity of a specific physiological or behavioral characteristic possessed by a user. New requirements over actual biometric systems as robustness, higher recognition rates, tolerance for imprecision and uncertainty, and flexibility call for the use of new computing technologies. In this context soft-computing is increasingly being used in the development of biometric applications. Soft-Biometrics correspond to a new emerging paradigm that consists in the use of soft-computing technologies for the development of biometric applications. The aim of this paper is to motivate discussions on application of soft-computing approaches in specific biometric measurements. The feasibility of soft-computing as a tool-set for biometric applications should be investigated. Finally, an application example on static signature verification is presented, providing evidence of soft-computing's impact in biometrics.

Keywords: image processing, pattern recognition, soft-computing, biometrics, soft-biometrics, signature analysis

1 Introduction

Biometrics offers new perspectives in high-security applications while supporting natural, user-friendly and fast authentication. Biometric identification considers individual physiological characteristics and/or typical behavioral patterns of a person to validate their authenticity. Compared to established methods of person identification, employing PIN-codes, passwords, magnet- or smart cards, biometric characteristics offer the following advantages:

- They are significant for each individual,
- They are always available,
- They cannot be transferred to another person,
- They cannot be forgotten or stolen,
- They always vary [1].

Although, there was a strong growth in biometric technologies during the past years [2], the introduction of biometrics into mass market applications, like telecommunications or computer-security, was comparable weak [17].

From our point of view there are three main aspects responsible for the current situation. First, soft- as well as hardware (sensor) technologies are still under development and testing, although, black sheep promising 100 percent recognition rates. Second, there is a lack of standardization and interchange of biometric systems; basically, such systems are proprietary. And third, there are only few large-scale reference projects that gave evidence of the usability and acceptance of biometrics into real worlds applications.

The work presented in this paper will contribute to the first aspect by employing soft-computing approaches to improve algorithms of biometric analysis. The aim is to provide tool-sets that are able to handle natural variations being sticking in biometrics. Also, the tool-sets should be tractable, robust and of low costs. So, the authors studied soft-computing approaches and their feasibility into biometric measurements.

Section 2 gives a short overview on biometrics where section 3 introduces soft-computing. Section 4 deals with the introduction of soft-computing into biometric application and claims the paradigm of

[1] Rem.: The presentation of two 100% identical feature sets indicates fraud.

soft-biometrics. An implemented application example, in particular for signature verification, is described in section 5. Finally, in section 6 some conclusions of this work are presented.

2 Biometrics

Biometric systems comprise the following components: data acquisition and pre-processing; feature extraction and coding; computation of reference data and validation. The systems compare an actual recorded characteristic of a person with a preregistered characteristic of the same or another person. Thereby, it has to be decided between *identification* (1 to many comparison) and *verification* (1 to 1 comparison). Then, the matching rate of the both characteristics is used to validate, whether the person is what he/she claim to be. The procedures seem to be equivalently to the traditional methods using PIN or ID-number. However, the main difference is founded by the fact that in biometrics an absolute statement *identical / not identical* cannot be given. For instance a credit card has exact that number "1234 5678 9101" or not, contrary, a biometric feature varies naturally at any acquisition.

Biometric technologies will be divided into approaches utilizing *physiological* characteristics, also referred as *passive* features, and approaches using *behavioral* characteristics that are *active* features. Behavioral characteristics, used e.g. in speaker recognition, signature verification or key-stroke analysis are always variable. On the other hand physiological characteristics employed e.g. in hand, fingerprint, face, retina or iris recognition, are more or less stabile. Variations may be caused by injuries, illness and aging, as well as variations during acquisition.

Each biometric system has to be able to handle diverse variation by using "tolerance-mechanisms." Also, it should be possible to adjust a statement about a person's identity gradually with a certain probability, and, it should allow for tuning a system not to reject a person falsely or to accept another person without permission.

Due to the variability of the biometric characteristics, a resulting error rate cannot be easily assigned. However, adaptable algorithms that are able to handle inaccuracies and uncertainty might slow down resulting error rates.

Biometric approaches have to solve the two-class problem *person accepted* or *person rejected*. So, the performance of biometric systems is measured with two basic rates: False acceptation rate (FAR) is the number of falsely accepted individuals; False rejection rate (FRR) is the number of falsely rejected individuals [22].

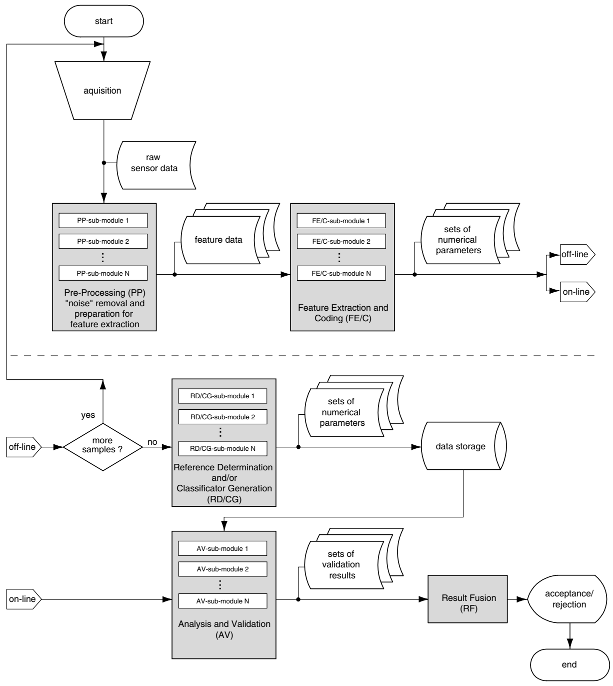


Figure 1: Block-diagram of a biometric system.

3 Soft-computing

Since the early days of Artificial Intelligence scientists and engineers have been searching for new computational paradigms capable of solving real-world problems efficiently. Soft-Computing (SC) is one of such paradigms that has emerged in the recent past as a collection of several models of computation, which work synergistically and provide the capability of flexible information processing. The principal constituents of SC are fuzzy logic, neural networks, evolutionary computing, probabilistic reasoning, chaotic theory and parts of machine learning theory. SC is more than a melange of these disciplines, it is a partnership, in which each of the partners contributes a distinct methodology for addressing problems in its domain. In this perspective, these disciplines are complementary more than competitive [24].

SC technologies are currently attracting a great deal of attention and have already found a number of practical applications ranging from industrial process control, fault diagnosis and smart appliances to speech recognition, image processing and pattern recognition. We are going to concentrate ourselves in the description of fuzzy logic, neural networks and evolutionary computing, because these are the SC technologies most used in biometric applications. In this perspective, the principal contribution of fuzzy logic is to provide algorithms for dealing with imprecision and approximate reasoning and computing with words, while neural network theory provides an effective methodology for learning from examples, and evolutionary computing a way to solve search and optimization problems.

Fuzzy Logic (FL) corresponds to a mathematical approach for translating the fuzziness of linguistic concepts into a representation that computers can understand and manipulate. Because FL can transform linguistic variables into numerical ones without jettisoning partial truth along the way, it allows the construction of improved models of human reasoning and expert knowledge. FL and in general the fuzzy sets theory provides an approximate, effective and flexible means of describing the behavior of systems which are too complex or too ill-defined to admit precise mathematical analysis by classical methods and tools. Since the theory of fuzzy sets is a generalization of classical set theory, it has greater flexibility to capture faithfully the various aspects of incompleteness or ambiguity in information of a situation. In this way, more than to operate just with linguistic variables, modern fuzzy sets systems are designed to deal with any kind of information uncertainty.

Artificial neural network (ANN) research takes its inspiration from biological neuronal systems. ANNs have some attributes as universal approximation, the ability to learn from and adapt to their environment, and the ability to invoke weak assumptions about the underlying physical phenomena responsible for the generation of the input data. ANNs are suitable to solve problems where no analytical model exists or where the analytical model is too complex to be applied. Basic units called artificial neurons that somehow model the working principles of their biological counterparts compose an ANN. Moreover, ANNs try not only to model the biological neurons but also their interconnection mechanisms and global functional properties.

The mechanisms that drive natural evolution are reproduction, mutation and survival of the fittest. They allow the adaptation of life forms to particular changing environments over successive generation. From a computational point of view this can be seen as an optimization process. The application of evolution mechanisms by artificial/computational systems is called Evolutionary Computing (EC). From that we can say EC takes the power of natural selection to turn computers into automatic optimization tools. EC algorithms are efficient, adaptive and robust search processes, producing near optimal solutions and have a large amount of implicit parallelism.

4 Soft-Biometrics: Soft-Computing and Biometrics

SC is increasingly being used in biometric systems whereas biometrics employing SC approaches are referred as *soft-biometrics*. SC, by its ability to consider variations and uncertainty is data, is suitable for biometric measurements due to the following reasons:

- Biometric features do not have an absolute "ground truth" and they will hardly reach this. Biometric features always vary!
- Derivations from the "ideal" biometric characteristic are difficult or even unable to describe analytically.
- High accuracy within the measurement may cause inflexibility and the loss of generalization ability.

The general biometric system whose block-diagram is shown in figure 1 is made of a pre-processing module (PP); a feature extraction and coding module (FE/C); a reference determination and/or classifier generator module (RD/CG); an analysis and validation module (AV) and a result fusion module (RF). The PP-module comprises diverse methods that treat recorded data in such a way that significant features can be extracted easily. The FE/C-module includes methods that convert treated input data into numerical parameters, which represent specific aspects of a biometric characteristic. Within the RD/CG-module the numerical parameters are used to determine reference/template-data or to generate classifiers. It is only employed in the off-line phase during enrollment/training. The AV-module is activated during the on-line phase to analyze and to validate the numerical parameters of a questioned (also called sample) characteristic. Last but not least the RF-module combines different outputs, in case there is more than one AV-module, to decide whether a person is what they claim to be.

SC can be introduced into any component/module of a biometric system. The application of SC as classifiers or decision-ruler is widely spread [18], whereas, SC in pre-processing and feature extraction is sparsely used. Nevertheless, in [16] a comparative study between different face recognition approaches was presented, which employ SC in FE/C-, RD/CG-,

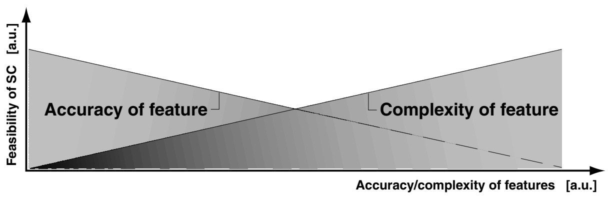


Figure 2: Feasibility of soft-computing depending on complexity and accuracy of the biometric features.

and AV-module. From our point of view the application of SC in biometrics has to be decided individually. Since SC is always data-driven, the available data has to be analyzed to decide in detail whether it is useful to employ SC or not.

To give an example. In fingerprint identification for the matching of e.g. 14 minutia the application of SC is overpowered, contrary in static signature verification, here, e.g. the number and shape of signature strokes varies. (Rem.: In case there are two almost identical sets of strokes it indicates fraud.) Consequently, SC approaches might be employed if the biometric features are of high complexity, less accurate and an analytical description is time consuming or almost impossible (see figure 2).

On the other hand it does not make sense to feed SC approaches with any available detail of a biometric characteristic. As in the traditional approach features that are considered during validation have to be significant. Here, SC can support the selection of typical details of a biometric characteristic.

5 Application Example: Signature Verification

Comparing signature verification with fingerprint, face, voice and other recognition technologies such as retina and iris scanning, signature verification has several advantages as an identity verification mechanism:

- Signing is the most traditional and social accepted method of identification.

- Signature analysis can only be applied when the person is/was conscious and disposed to write in the usual manner, although they may be under extortion to provide the handwriting. To give a counter example, a persons fingerprint may also be used when the person is in an unconscious (e.g. drugged) state of mind.

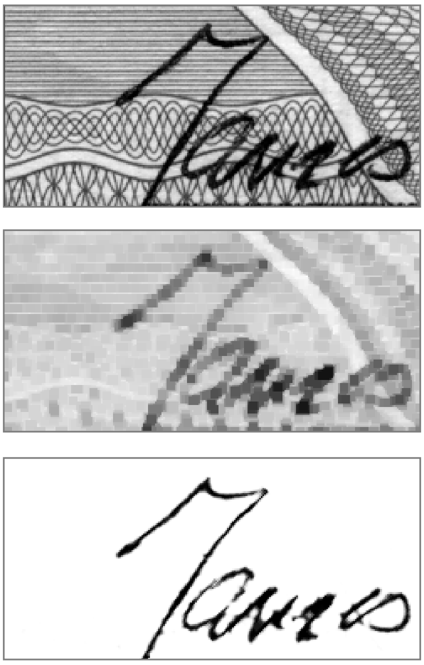
- Forging a signature is deemed to be more difficult than a fingerprint given the availability of sophisticated analyses.

There are two fundamental approaches for signature validation, which are determined by the manner of signature digitalization. The first, dynamic signature verification, uses a special electronic device (electronic pen and/or an electronic pad) for the acquisition during the signing process. Pen movements are stored as time signals primarily representing the pen displacement in the x,y-writing plane. A person must be present and must explicitly use this special devise for applying signature analysis. For static signature verification, the second approach, the final handwritten signature pattern is handled as a graphical image and used for the analysis. So, static signature verification is able to process signatures written on paper documents as well as signature images generated by the time series of electronic devices. Static signature validation can be applied in a persons absence.

Static signature verification is a great challenge among biometrics. Due to the involvement of various writing utensils such as diverse kinds of pen/ink , kinds of paper and kinds of physical texture of the writing pad, static handwriting specimens underlie great variations. The challenge of static signature verification is to employ sophisticated methods of digital image processing and pattern recognition to compensate influences by writing utensils and to validate writer specific characteristics objectively and robustly.

The impact of SC to biometrics and particularly to static signature analysis becomes obvious by introducing SC technologies in different processing stages of the verification process. Former signature verification approaches employed SC, in particular ANN, as trainable classifiers [15][18]. Until now there are only few systems that use other SC methods (e.g. FL and/or EC), too [23]. During the past years we have developed the *SIC Natura* system [6][4] for signature verification that includes SC methods in the document-PP-module, in the RD/CG-module and within the AV-module. SC is used to adapt image processing operators, to design image filter algorithms, to derive feature classifier rules, and, to evaluate signature stroke and signature shape characteristics.

Due to the influences of writing utensils the preprocessing of the digitized paper document has to be highly adaptive. The texturing of the background of e.g. Euro-checks or traveller checks is a great problem for the extraction of the handwriting on the bank check. Nevertheless, it has to be performed before



a)

b)

c)

Figure 3: Application of the standard gray value (b) and the Dubois and Prade fuzzy morphology [5] (c) for background removal of bankchecks.

any further processing can be started. In [3], a framework was presented, which is able to derive general foreground-background separating algorithms on textured images by using genetic programming of image processing operations. However, this framework is as good as the collection of image processing operations provided to the framework. A set of operations capable of background removal had to be selected.

The first example of SCs application are image processing operations based on the Dubois and Prade proposal of a fuzzy triangular norm [5].

Triangular norms, which were initially developed for metrics of statistical spaces [19], have recently found a high interest in fuzzy theory [10]. Especially they are used for providing formal definitions of fuzzy set union and intersection operations. So far, triangular norms did not find much resonance in the pattern recognition community. This is an unlucky situation, since, at least in a formal sense, triangular norms allow for the formal extension of standard image processing operations and for the definition of new ones as well. From an image processing laymans point of view, triangular norms could be used whereever maximum or minimum formally appears in a definition.

For using the Dubois and Prade fuzzy norm [1], morphological operations [20][21] are derived from it. Such operations can be formally qualified as fuzzy morphology operations [12], since they give replacements for the infimum and supremum of a set of data, and basic definitions of mathematical morphology are derived from the requirement for an operation to commute with the supremum or infimum [9].

Figure 3 shows the result of the application of standard morphology and Dubois and Prade morphology to a bankcheck image. The image (a) was opened with standard gray value morphology and dilated afterwards, to obtain the image (b) (using a 3×3 structuring element), and it was Dubois and Prade opened with parameter $\alpha = 0.8$ and dilated with $\alpha = 0.5$, using a 4-neighbor structuring element to obtain image (c). Despite of the interactive setting of reasonable values for α (which may not be suitable for other bankcheck images), the image (c) clearly shows the ability of Dubois and Prade morphology to completely remove the background while preserving a good quality of the handwriting. Such operations are a valuable addition to the set of basic operations needed for adaptive textured image processing according to the *LUCIFER*-approach[3].

The *LUCIFER*-system, being the second example of SC-application, considers EC to design document background filters on the fly [3],[4]. The system (see figure 4) is composed of (user-supplied) original image, filter generator, filter output images 1 & 2, result image, (user-supplied) goal image, 2D-Lookup matrix, comparing unit and filter design signal. An evolutionary algorithm maintains a population of individuals, each of which specifies a setting of the framework. By applying a 2D-Lookup [4] and measuring the quality of coincidence of goal and result image, a fitness value can be assigned to each individual. These fitness measures are used for the standard genetic operations of an evolutionary algorithm. Essential parts of the framework are the two image processing operations. In order to generate them, genetic programming is used [14]. Genetic programming maintains a population of expression trees (shortly: trees). Every tree equals a program by its structure. "Performing" the tree may be used for evaluating a quality value for the suitability of the program represented by the tree. An evolutionary algorithm can use these quality measures as fitness values. In this case, the evolutionary algorithm is referred to as genetic programming.

The *Lucifer*-system was applied to a patch image of a filled-in Euro-check. Euro-checks are considered to be the most important family of bankchecks in Europe, and also to comprise the most challenging background filtering problem. The patch image and

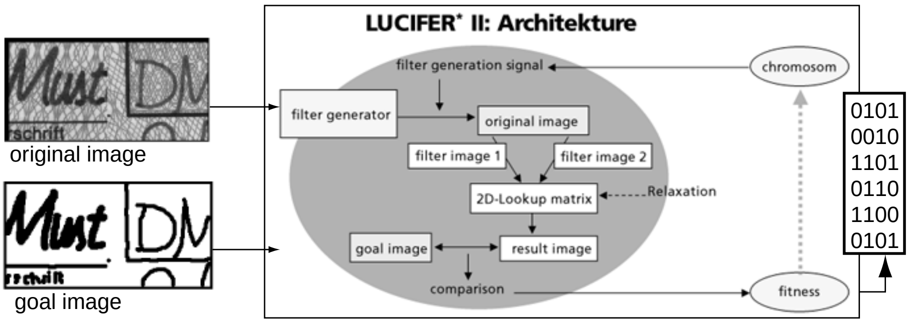


Figure 4: LUCIFER-module for background filter generation.

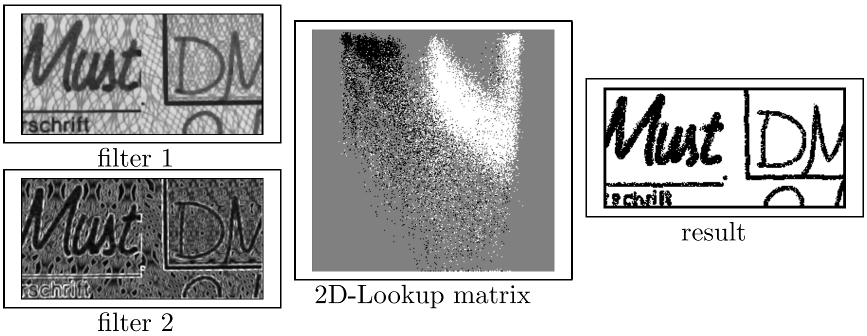


Figure 5: Example of *LUCIFER*-result with fitness 0.0118.

its binary goal image are shown in figure 4. The GP runs were configured as follows:

- parent trees: 20
- children trees: 50
- maximal initial tree depth: 2
- maximal tree depth during the run: 3
- crossover selection rule: tournament selection

Most runs gave promising results. One example with the corresponding filters and 2D-Lookup matrices is given in figures 5.

The third example of SC's application is the derivation of fuzzy matching rules [7]. In [13] Kosko presented a framework for adapting fuzzy rules to data. The approach is based on the geometric interpretation of fuzzy rules as so-called fuzzy patches, and decomposes the rule adaptation into two stages: stage 1 is the unsupervised learning for a rough placement of the rule parameters, stage 2 a fine-tuning of the so-placed rules by means of a supervised learning.

In our approach to static signature verification, questioned and reference signatures are threaded morphologically to derive so-called regions (compare figure 6). Then, human labeled region-training-sets are used to select and to adapt fuzzy rules for upcoming region matching. So, SC technologies are employed to derive a sophisticated schema for the region approach validating the spatial organization of signatures. A comparative study with first empirically determined rules as well as the increasing of the achievable matching rate by 20% for the manual designed rules up to 98% for the automatically approximated rules punctuate the performance of SC technologies in real-life applications. As it can be seen in table 1, the studied SOM-net [11] providing 35 rules is mostly effective, compared to the 45 rules the applied 35 rules provide the same error of 1.98% (117 falsely classified region pairs out of 5914 region pairs).

6 Conclusions

Soft-Biometrics, the application of soft-computing in biometrics is motivated. Taking the example of signature verification it is shown that SC approaches can be employed within all components of the biometric system, e.g. for data pre-processing itself or for designing of adapted pre-processing filters. Also, it can be applied for extracting significant features, for reference determination, as classifiers or as decision-

Net nodes	Number of rules	MSE 1. stage	MSE 2. stage	Err. train. data	Err. all data
3×8	19	0.0327	0.0286	129	255
3×12	29	0.0281	0.0217	98	202
3×16	35	0.0216	0.0171	55	117
3×24	54	0.0180	0.0131	45	117

Table 1: Result for SOM-architectures providing different numbers of adapted fuzzy rules. The total number of samples was 2929 for training and 5914 for testing.

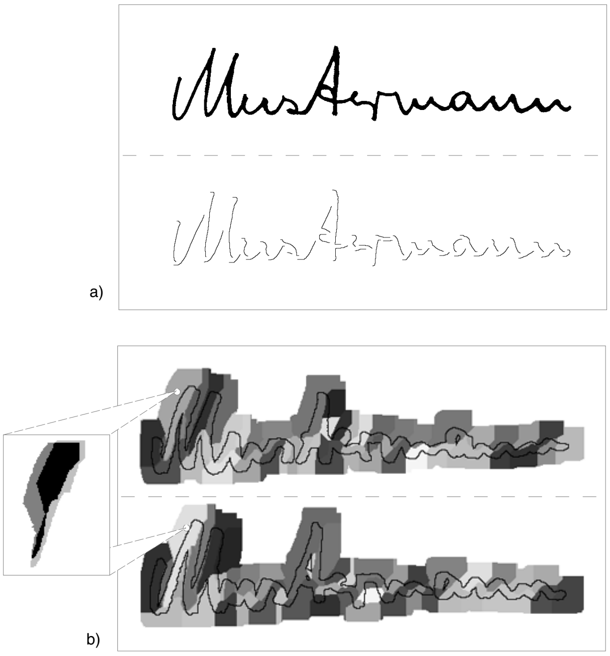


Figure 6: Region production and matching: a) binary signature image and corresponding shadow image; b) regarding their centers of gravity, regions of reference and pattern signature were laid one upon the other.

ruler. However, the feasibility of SC as a tool-set in biometric applications has to be studied, in particular by investigating the accuracy and complexity of the given data. More detailed, the usage of SC approaches within the *SIC Natura* system are presented. Fuzzy triangular norms, proposed by Dubois and Prade [1] are employed to design fuzzy morphology operators being able to remove textured background imprints on bankchecks. Evolutionary computation, in particular genetic programming, is used within the *LUCIFER* framework to design complete bankcheck filter algorithms on the fly[3]. And, a hybrid neuronal system, proposed by Kosko [13], was employed to derive rules for signature region matching. Since there is a bundle of further feature extraction and validation methods within the *SIC Natura* system upcoming work will be devoted to the fusion of sub-module-results by using fuzzy fusion (e.g. Fuzzy Integral [8]).

Acknowledgement

This research was supported by the join "Program of Scientific Cooperation" of CONICYT (Chile) and BMBF (Germany), under the project "Face Recognition and Signature Verification using Soft-Computing" (CHL 01/011).

References

[1] D. Dubois and H. Prade. *Fuzzy Sets and Systems: Theory and Applications.* Academic Press, New York, 1980.

[2] Elsevier Advanced Technology. Biometrics technology today, 01/1997-02/2000.

[3] K. Franke and M. Köppen. Towards an universal approach to background removal in images of bankchecks. In *Proc. 6th International Workshop on Frontiers in Handwriting Recognition (IWFHR)*, pages 55–66, Tajon, Korea, 1998.

[4] K. Franke and M. Köppen. A computer-based system to support forensic studies on handwritten documents. *International Journal on Document Analysis and Recognition*, 3(4):218–231, 2001.

[5] K. Franke, M. Köppen, and B. Nickolay. Fuzzy image processing by using Dubois and Prade fuzzy norm. In *Proc. 15th International Conference on Pattern Recognition (ICPR)*, pages 518–521, Barcelona, Spain, 2000.

[6] K. Franke, M. Köppen, B. Nickolay, and S. Unger. Machbarkeits- und Konzeptstudie automaisches System zur Unterschriftenverifikation. Technical report, Fraunhofer IPK Berlin, 1996. (in German).

[7] K. Franke, Y.-N. Zhang, and M. Köppen. Static signature verification employing a Kosko-Neuro-Fuzzy approach. In N.R. Pal and M. Sugeno, editors, *Advances in Soft Computing - AFSS 2002, LNAI 2275*, pages 185–190. Springer Verlag, 2002.

[8] M. Grabisch and J.-M. Nicolas. Classification by fuzzy integral: Performance and tests. *Fuzzy Sets and Systems*, 65:255–271, 1994.

[9] H.J.A.M. Heijmans. *Morphological Image Operators*. Academic Press, Boston, MA, 1994.

[10] George J. Klir and Bo Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, New York, 1995.

[11] T. Kohonen. *Self-Organizing Maps*. Springer, 1995.

[12] Mario Köppen and Katrin Franke. Fuzzy Morphology revisited. In *Proc. 3rd International Workshop on Softcomputing in Industry (IWSCI)*, pages 258–263, Muroran, Japan, 1999.

[13] B. Kosko. *Fuzzy Engineering*. Prentice Hall Internationall, Inc., 1997.

[14] J. Koza. *Genetic Programming – On the Programming of Computers by Means of Natural Selection*. MIT Press, 1992.

[15] F. Leclerc and R. Plamondon. Automatic signature verification: the state of the art 1989-1993. *International Journal of Pattern Recognition and Artificial Intelligence*, 8:643–660, 1994.

[16] P. Navarrete and J. Ruiz del Solar. Comparative study between different Eigenspace-based approaches for face recognition. In N.R. Pal and M. Sugeno, editors, *Advances in Soft Computing - AFSS 2002, LNAI 2275*, pages 178–184. Springer Verlag, 2002.

[17] E. Newham, C. Bunney, and C.Mearns. Biometrics report, 1999.

[18] J. Schneider, K. Franke, and B. Nickolay. Konzeptstudie - Biometrische Authentifikation. Technical report, Fraunhofer IPK Berlin, 2000. (in German).

[19] B. Schweizer and A. Sklar. Statistical metric spaces. *Pacific J. of Mathematics*, 10:313–334, 1960.

[20] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, 1982.

[21] J. Serra. *Image Analysis and Mathematical Morphology. Volume 2: Theoretical Advances*. Academic Press, 1988.

[22] Tele Trust - AG 6. Biometrische Identifikation - Bewertungskriterien zur Vergleichbarkeit biometrischer Verfahren, 1998. (in German).

[23] X. Yang, T. Furuhashi, K. Obata, and Y. Uchikawa. Constructing a high performance signature verification system using a GA method. In *Proc. 2nd International Two-Stream Conference on Artificial Neural Networks and Expert Systems*, pages 170–173, Dunedin, New Zealand, 1995.

[24] L.A. Zadeh. Some reflections on soft computing, granular computing and their role in the conception, design and utilization of information/intelligent systems. *Soft Computing*, 2:23–25, 1998.