DOI: 10.1109/MIM.2006.1634984

Scalable Distributed Biometric Systems

Advanced techniques for security and safety

Marco Gamassi, Vincenzo Piuri, Daniele Sana, Fabio Scotti, and Olga Scotti

Airports, public buildings, banks, court houses, ministry offices, and tourist attractions are frequently full of people. The ability to guarantee safety from terrorist or criminal actions would make the towns and the countryside more livable. At the same time, the public demands a guarantee that personal privacy will be maintained.

For these reasons, the use of advanced techniques for recognition and identification of people is necessary. There are three main areas of application:

- recognition and identification of terrorists or criminals
- recognition and identification for access to personal information or physical entry into an area
- studying people movements, monitoring human behavior, and detecting anomalous scenes; early detection of suspicious behaviors may prevent dangerous acts.

To guarantee the safety and privacy of citizens, institutions and enterprises are interested in using solutions based on biometric techniques to develop innovative services. For institutions, the adoption of these approaches improves services and shows responsiveness to citizens and their feelings. For enterprises, these solutions are extremely useful to maintain and increase their business and be competitive in domestic and worldwide markets.

Biometric Properties

Certain physical human features or repeatable actions are characteristics that specifically and uniquely identify a person.

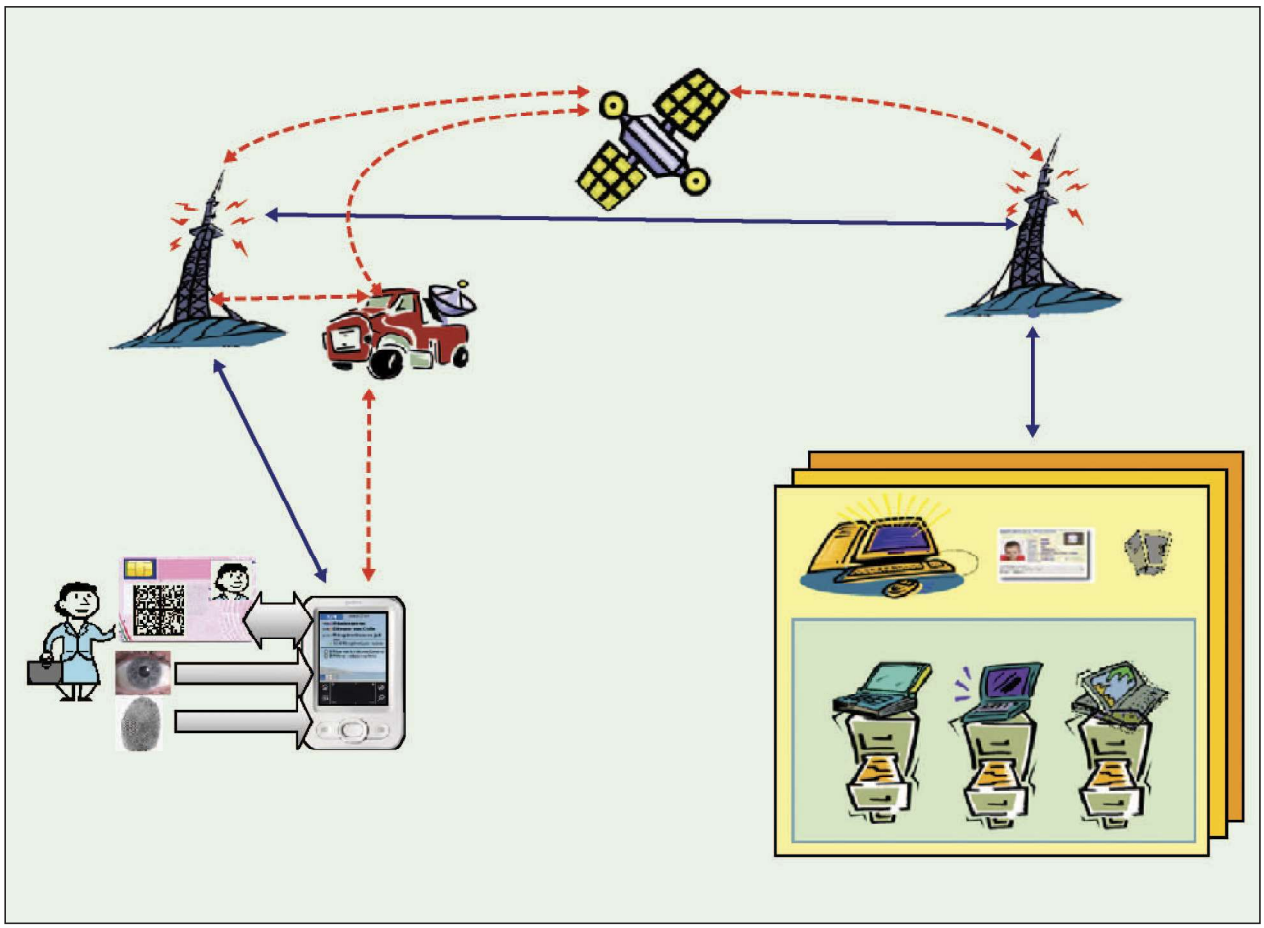

Fig. 1. The overall structure of a biometric system. Identity of the card owner can be verified by the palm top that can, if necessary, search remote archives. Connection to remote archives can follow different routes: Continuous lines indicate the normal data flow; outlined lines stand for backup connections.

Physical features that are specific identifiers include a person's retinas, iris, DNA, fingerprints, palm print, or pattern of finger lengths. Repeatable actions include voice prints, gaits, or handwriting. Biometrics are automated methods of recognizing a person based on a physiological or behavioral characteristic.

Many methods for identification and recognition have been studied. Very efficient algorithms have been developed that process biometric information from individuals when they register in a system. These algorithms produce a digital template and store it in a database. Each time any person returns and requests access, his or her biometric information is captured again; the algorithm produces a digital template and compares it with the templates in the database to ascertain matches. The system measures the Hamming distance, i.e., the similarity between two bit strings. Identical bit strings have a Hamming distance of zero. If they do not match, the Hamming distance is one.

Performance of a biometric measure is measured by the false accept rate (FAR), the false reject rate (FRR), and the failure to enroll rate. A way to compare real-world biometric systems is using the equal error rate (EER) or cross-over error rate (CER) when the FAR and FRR errors are equal. The FAR and FRR are plotted on a graph, and the intersection of the two is evaluated; the lower the CER, the more accurate the system. The EER of different systems that are available at this time ranges 60–99.9%. For recognition, good rates have been reached with face analysis [1]. The best results in identification can be obtained from iris comparison and fingerprint analysis. Other techniques (like voice and signature) still need to be improved [2].

Current Solutions for Biometric Data Use

The limits of current solutions and research projects are as follows:

- Each problem is specific; there is no systematic or comprehensive approach.
- They are not adaptive and do not evolve.
- There is no global optimization.
- There is no attention to specific or integrated consideration for system-level issues (e.g., multiplicity of databases, database privacy and security, algorithms privacy, scalability, the ability to integrate, interoperability, maintenance, usability, or service continuity).

A very large quantity of heterogeneous biometric databases exists. Each institution or big enterprise has one or more biometric databases that they have stored by using various sensors and collection procedures.

Our goal is to design a comprehensive system that is able to query, in parallel, all available databases with various feature extraction approaches while still receiving approval from all database holders. This would permit a check of a much wider amount of information that is already stored. Also, more data can be entered by using advanced devices that can collect "on the road" biometric data and code them before transmission. Coded data could be compared with data stored in all selected databases to contribute to recognition or identification.

The ability to guarantee safety from terrorist or criminal actions would make the towns and the countryside more livable.

Objectives

The system design we are describing deals with the following concerns:

- *personal identification:* by comparison of the code generated by biometric measurements with the one stored in a passport (one-to-one)
- *personal verification:* by code database searching (one-to-*N* search), including negative search (i.e., nonmatching approach); this approach demands interoperability of data streams incoming by terminals and heterogeneous data stored in different, multinational databases
- *personal identification:* by generating a nonreversible biometric code also from an incomplete database with heterogeneous or incomplete databases. The relationship between the biometric features and the personal identifier (template) should be nonreversible, i.e., it should not be possible to reconstruct the original biometric features from the template. This is fundamental to ensure privacy for the individual while still providing usable information. People are very concerned and have a "big brother" security syndrome. For example, taking a fingerprint has long been associated with crime and finding a criminal.

To implement these applications, the following technologies are involved: distributed systems, multiagent architectures, multiple computing paradigms, and adaptable or evolvable techniques. Cooperative and integrated use of these technologies is necessary and fundamental to reach our comprehensive target for designing and implementing the new generation of integrated biometric security systems. This innovation is proposed by our system-level design methodology. The project described here and the proposed approach studied, experimented, and evaluated the use of a modular, structured design approach at the system level.

Figure 1 shows the overall structure of a biometric system designed by using the comprehensive system-level methodology that the authors propose.

Example: The Airport

Consider the policy agents in charge of airport safety. If they need to find a person or find out if a person is a suspect in an incident, a biometric data gatherer could be very helpful in completing that task. It is a handheld device, similar to a palmtop, which contains multiple single-biometry sensors for collecting biometrics data (fingerprints, iris, or signature), a camera, a slot for a smart card, and a CPU [11].

Each security agent assigned to identities control would be supplied with this palmtop device. After collecting the biometrics data on the person to be identified, they would codify the information using all of the techniques loaded on the palmtop at the moment (techniques can be added when additional technologies are developed).

Multiple biometric data are accessed from intelligent terminals in a type of biological code [nonreversible coding (NRC)] that is a unique identifier for each person. The unique identifier is based on the combination of data from the sensors so that it may not be traced back to any of the biometrics or sensors used to generate it. Generation of the personal identifier is shown in Figure 2. The NRC can be generated from multiple biometric measurements and also from incomplete, heterogeneous databases. Development of algorithms to generate NRC from an existing limited database allows use of data already stored and to manage data in a transition phase.

Coded data can be transmitted without violating the privacy of the individual, since the union of all the possible coding of the same type of biometric information does not go back to the original data. The communication would use secure channels [11]. If conventional channels are interrupted, a backup procedure using satellite communication could be used.

Implementing agent systems based on information technologies and advanced hardware platforms (i.e., parallel computing or clustering) allows managing the complexity of heterogeneous databases and performing all database operations (queries and new data introduction from multiple sources) by satisfying stringent time requirements [13], [14]. The multiagent structure supporting all features is described in the following sections and is shown in Figure 3.

The Multiagent Structure

Multiparadigm Application

Every technique that operates on only one biometric characteristic has specific limits. The use of more techniques for the analysis of one biometric characteristic constitutes the multiparadigm approach to system design. The multiparadigm approach allows for the following:

- the same biometric information to be analyzed from various points of view, possibly overcoming the limitations and the problems that are intrinsic in the individual technique [3], [4];

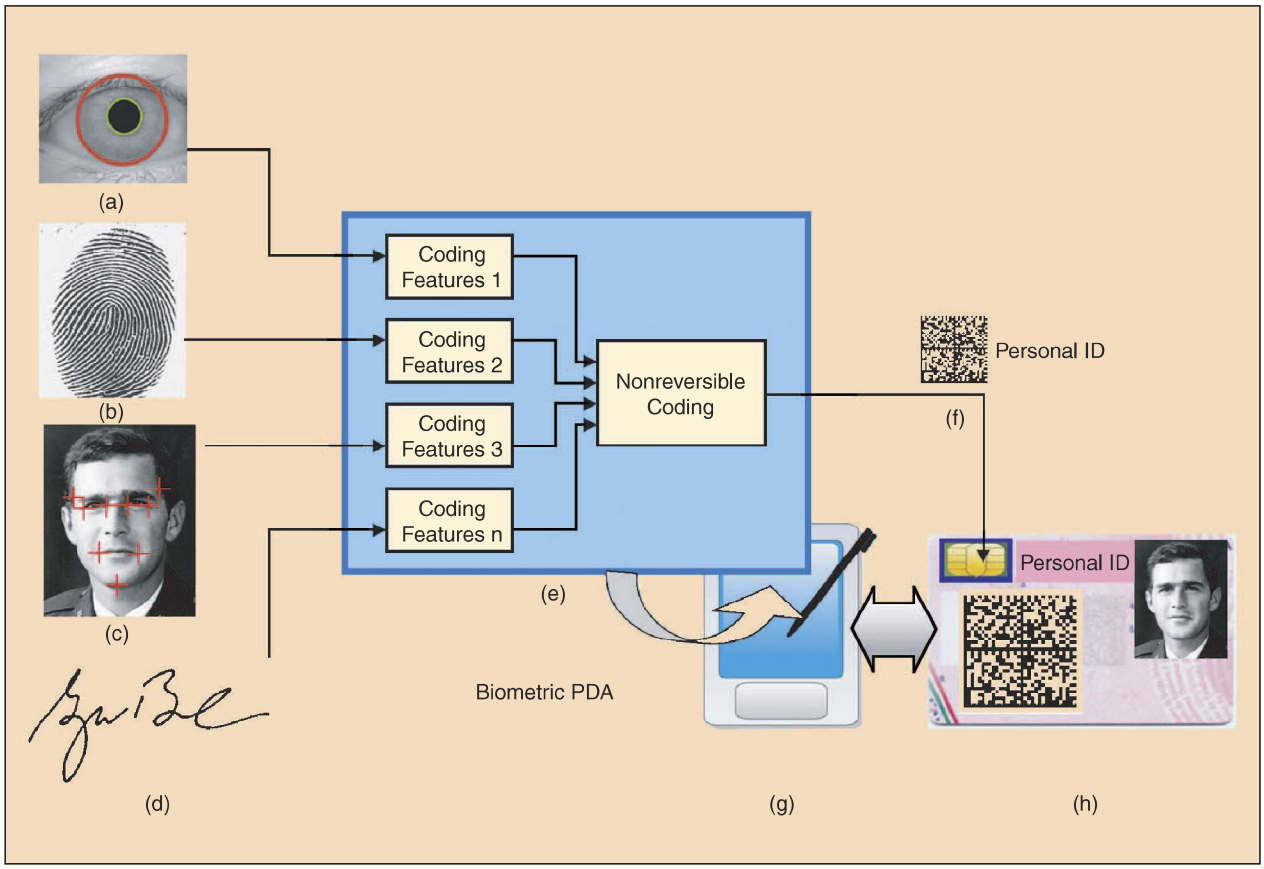

Fig. 2. All biometric features available (iris, fingerprint, face points, signature) can be combined by the palm top in a nonreversible code that can be stored in the identifier document that can be stored in the document.

- exploiting the characteristics of each alternative technique to process the same biometric information;
- dealing with different characteristics of the information stored in the various databases (also due to the various sensors and collection procedures used during biometric information acquisition). For example, fingerprints can be coded by looking at local or global features. The method used depends on the kind of image available, the sensor, and the algorithm, but the match with all databases can be possible only if each database receives data that are compatible with the ones stored in the database itself;
- for supporting natively the inclusion and deployment of more advanced algorithms and new solutions as soon as the research and the industry make them available without changing the overall biometric system structure and services.

Multimodal Application

Each method, or mode, of identification has different effectiveness from the others. Success or failure can be caused by the quantity of available information, the operating conditions, and the intrinsic characteristics of each technique. These considerations suggest not limiting the recognition process by only using one biometric technique. When one method fails or is not sufficiently adequate for accuracy and confidence, other techniques or other biometrics modalities might offer better results.

The multimodal approach is based on the use of more than one biometric characteristic to carry out identification, recognition, or monitoring. This increases the probability of correct identification and recognition by using different biometric information. A biometric feature that is accurate for recognition and identification does not exactly overlap with other good biometric features. So the union of all biometric features is better than individual solutions, thus providing better performances overall [5], [6]. The most efficient system will be created by fusing the partial results achieved by each of them. Solutions will also need the ability to adapt to changes in environmental conditions and to the quality of the analyzed information to overcome the limits of solutions that are tailored to specific conditions.

Adapting and Evolving Techniques

Computer technology evolves very quickly and biometric fields are improving thanks to better sensors, speed and capacity of calculation, and research progress. System characteristics and desired features that are changeable or not completely specified make it difficult to implement solutions

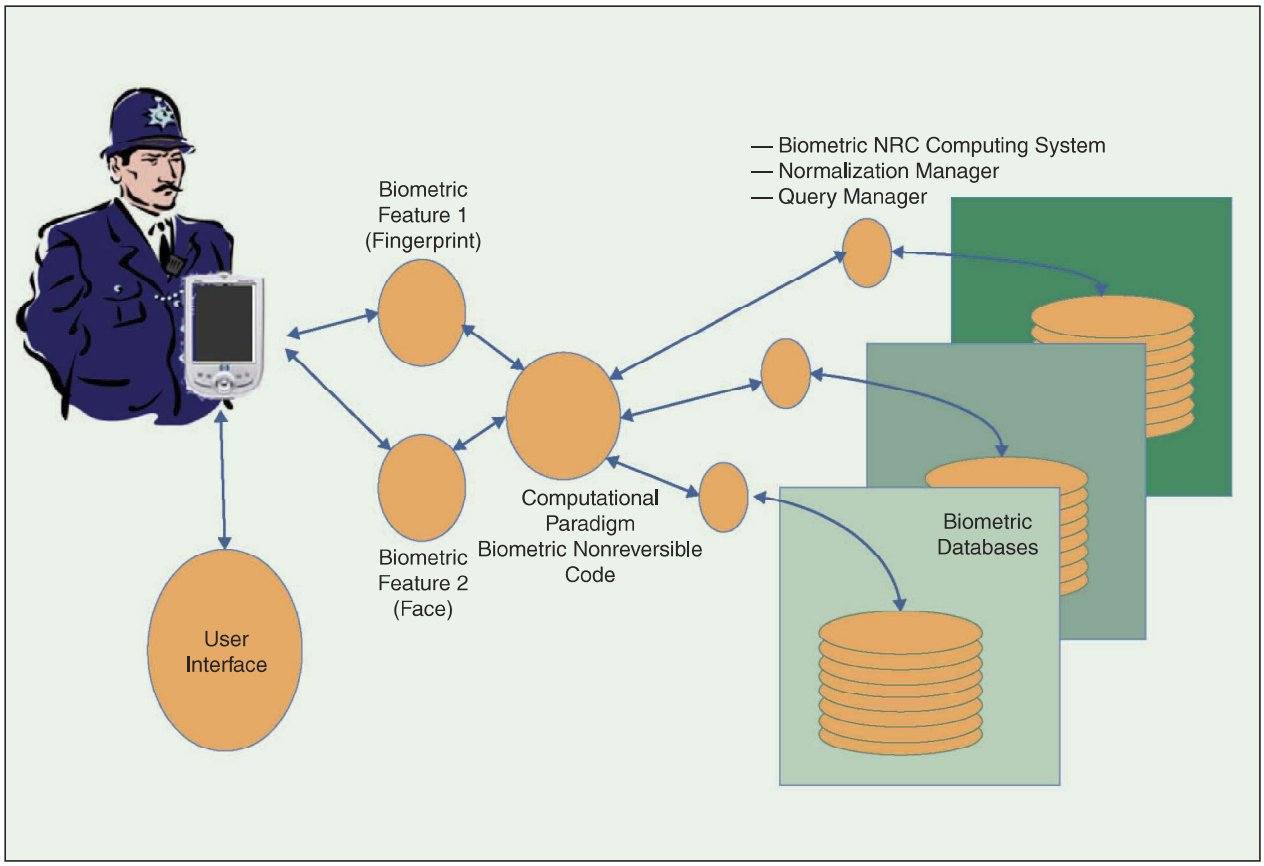

Fig. 3. The multiagent structure supports all features: multiparadigmatic, multimodal approaches, adaptive and evolving techniques, etherogeneous archives in distributed systems.

that are optimum, evolve by following the environmental and information changes, and remain optimum even after the evolution of the application environment.

Continuous learning by a biometric system is possible by using neural networks, fuzzy systems, evolutionary computing, and granular computing. These are adaptive and evolutionary techniques that provide computational intelligence [7]. Intelligent biometric systems can be created by defining their objectives by means of an appropriate set of examples, and by allowing for their self-adaptation to changing or not completely defined operating conditions. This type of system will exploit both the interpolation of the desired behaviors and extrapolation of them from the known database entries.

Distributed Systems

The use of distributed systems in this application brings up at lease two questions.

- Can biometric information be protected?
- Can the databases be integrated [8], [9]?

The various entities holding biometric databases (e.g., national and international police, security services, ministries of interior, and banks) are usually eager and very firm in protecting their own biometric information for security and privacy reasons and to reassuring the public opinion. Sharing of information is perceived as necessary to increase protection but is usually feared for possible information leakage or database intrusion and manipulation. In addition, centralizing all information could result in loss of quality and ownership, as well as in reduced responsibility. On the other hand, the heterogeneity of sources, collection procedures, and sensors make it difficult to integrate different databases into a single environment by using conventional database technologies.

Distributed systems are a technology solution that integrates the biometric resources available in the various existing databases without the need for physically merging the databases themselves. The parallel use of biometric modalities and analysis paradigms make it feasible to implement the parallelisms of the hardware architecture of the distributed information processing system [10]. It is possible to physically execute the computation of single tasks in parallel on different processors when these tasks are sufficiently independent of each other, thus reducing the total elapsed time for processing for a better real-time response. To this end, we can use multiprocessor structures, distributed architectures in local or geographic networks, graphic interactive display structures, local clusters, or systems based on blade processors. The use of parallelism based on the concurrent execution of the paradigms for the biometric modalities should be

much easier since it is not related to the parallelization of a single algorithm. It also allows for a dynamic computation allocation and, thus, a better balancing of workloads and use of resources.

Every technique that operates on only one biometric characteristic has specific limits.

Parallelism can also be used to improve performance limits imposed by the size of the databases and ensure high scalability to biometric databases. In particular, database partitioning can be used to limit the size of the archive in which a search must be performed; distribution of the partitions on various processing systems allows for search parallelization, while limiting the size of each portion of a global database to be managed by the security system.

Modeling biometric information processing on a distributed system allows an intrinsic mobility in the operations, service continuity, and fault tolerance. Appropriate replication of the databases allows for designing strategies that prevent the lack of accessibility due to network or processing nodes failures; the application will search for an available copy of the desired database in the accessible network.

Distributed systems allow for querying various databases by ensuring their individual privacy, security, integrity, and ownership, without actually sharing their information within the whole security system. At the same time, the distributed systems allow the use of private algorithms for feature extraction and matching, by ensuring their secrecy even within the security system. In particular, databases and biometric analysis algorithms can be made available remotely, without actually sharing them publicly within the security system. The system can access remote processing systems containing the biometric information obtained from the sensors and perform the database search remotely on the concise biometric features extracted locally. This will avoid publication of algorithms and database information within the security system with a high degree of privacy.

Distributed structures allow easy scalable increases in available database information and the integration of heterogeneous systems and legacy systems. Scalability can be achieved, for example, by expanding the individual database, restructuring the database allocation by means of splitting various host computers, merging various databases, or connecting new separate databases. Interoperability of heterogeneous databases also can be supported since individual databases are connected into the integrated environment.

Multiagent Architecture

Essential requirements for the system are architectures that are flexible, adaptable, and modularly scalable. Organizing the biometric systems as multiagent systems allows

- efficient software construction to enable an agent to carry out the analysis of a specific biometric modality with a specific technique on a specific database
- natively supporting the use of all of the above characteristics for advanced biometric information processing, namely, distributed information sources, multimodal approaches, and multiparadigmatic techniques, by encapsulating the various sources and the various queries in a homogeneous framework
- merging the results of the various processing to give a comprehensive result. The use of multiple modalities will be encapsulated by spawning parallel searches on the collected biometric features and then, by fusing the various partial results.

The intrinsically parallel structure of the computation based on multiple agents can use the parallelisms offered by the hardware architecture of the information processing system. The use of various databases containing biometric information in the same modality will be hidden by running separate agents querying the various databases in parallel and then, by merging the retrieved partial results. In particular, it is possible to physically execute the computation of single agents in parallel on different processors, thus reducing the total elapsed processing time. This allows also for exploiting possible parallelisms in the analysis of biometric information and in database searching, thus reducing the overall processing time. The agent-based approach also makes it easier to transparently interrogate several information sources, thus increasing the system usability and usage, as well as reducing the overall response time for a better real-time operation of the whole biometric system.

Modeling the biometric information processing as a multiagent system allows an intrinsic mobility in the operations, supporting service continuity, and fault tolerance. Agents will search for an available copy of the desired database in the accessible network. This allows for ensuring fault tolerance and continuity of service through automatic redistribution of the queries. The intrinsic parallelism of computation in multiagent systems allows for

- supporting the distribution of information and services, thus also guaranteeing high levels of security, protection, and privacy of information and processing algorithms
- easily supporting the scalable increase of databases and the integration of heterogeneous systems and legacy systems by encapsulation in suited agents.

The synergic cooperation of agents allows for finding the optimal solution by compensating for possible deficiencies and limits of individual paradigms, modalities, and databases, by operating in parallel. The use of multiagent systems will, in particular, be very useful to address the interoperability of biometric databases and biometric sensors outputs. The various differences among them can be encapsulated in the agent-based framework, and

appropriate analysis algorithms can be devised to take into account the differences between the quality and the characteristics of the collected biometric information. The multiagent structure can hide the various heterogeneous components by providing the interconnection glue and the homogenization infrastructure for database access.

Biometrics are automated methods of recognizing a person based on a physiological or behavioral characteristic.

On the other hand, the availability of efficient interoperability support will be highly valuable to realize multinational security systems since it will natively and easily allow for querying in the databases of the various cooperating nations without the actual need of sharing the real biometric information and analysis algorithms.

The System Level Design Approach

The construction of a biometric system with high abilities of integrating existing components and information databases, preserving ownership and privacy, scalability, and interoperability, as well as the native ability of supporting multimodal and multiparadigmatic methodologies, is challenging. It would have inestimable value for advanced biometric systems and their effective and efficient use in a heterogeneous distributed environment.

To achieve effective usability of biometric technologies in institutions and enterprises, solutions must be easily integrated into the usual operating processes and products. They must have good effectiveness, reasonable cost, small managerial impact, ability to be integrated, good scalability, good interoperability, and good adaptability to changeable conditions. In summary, biometric solutions need to be studied and implemented with a system-level perspective and not only as a technological add-on to existing services and products. The cooperative coordinated use of multiple modality, multiple paradigms, distributed systems, and multiagent frameworks allows for supporting such an integrated perspective in system design.

Results

In [3], a new design methodology for multimodal distributed biometric systems has been proposed, applying evolutionary high-level system design techniques to better structure the design procedure. The methodology aimed to avoid evident drawbacks in current designing procedures for biometric chains (e.g., time consuming and nonoptimality), by means of a comprehensive approach, thus automating the repetitive design tasks and preventing designers from an exhaustive design space exploration (trial-and-error approach).

The high-level optimum synthesis approach proposed in [3] can be summarized in the following main activities:

- to model the possible hardware architectures that are available in the design environment to implement the biometric systems
- to specify the behavioral description of the biometric system for the envisioned application
- to map the behavioral description for the specific application into a hardware model by means an iterative procedure.

In [3], we verify the feasibility and the usability of the methodology by implementing a prototype of the methodology by means of a object-oriented toolbox written in Matlab. To guarantee the viability of the methodology we introduced well-known biometric chains available in literature.

The cooperative multiagent system technologies are arrived to mature stage and a large number of real-world applications are now present in the literature [13], [14]. In [12], we presented a specific application. In particular, we describe issues about communication between perceptive agents and two implemented multiagent systems. As in this case, heterogeneous components can synergically cooperate to pursue an applicative goal. The use of commercial-off-the-shelf communication infrastructure technologies guarantees advantages of reliable mature technologies and allows supported, scaleable, and, generally, low cost components.

Conclusions

To achieve these goals, we studied, experimented, and evaluated the technologies and the design methodologies suited to realize an intelligent distributed system for personal identification, recognition, and monitoring. They are based on multimodal, multiparadigmatic, multiagents, and adaptive-evolutionary technologies that take advantage of the synergy offered by every technology.

Since the main focus is on the system-level perspective, the proposed approach is open to the possible enhancement of the individual components that are used in the security system. These components include sensors, portable devices for collecting the biometric features, biometric feature extraction algorithms, biometric identification, recognition, monitoring algorithms, database structure and query processing, communication technologies, and security for information and communication.

Future research will study the issues of optimum application of this comprehensive system-level design approach to satisfy the requirements of real applications. In particular, optimization will consider accuracy, confidence, reliability, performances, security, and cost. These characteristics can be very useful for governmental agencies, enterprises, banks, and all other organizations that need to protect people or control access to critical resources.

References

[1] P. Stéphane and V. Luc, "Image-based multimodal face authentication," *Signal Process.*, vol. 69, no. 1, pp. 59–79, Aug. 31, 1998.

[2] J.J. van Oosterhout, H. Dolfing, and E. Aarts, "On-line signature verification with hidden markov models," in *Proc. 14th Int. Conf. Pattern Recognition*, 16–20 Aug. 1998, vol. 2, pp. 1309–1312.

[3] M. Gamassi, V. Piuri, D. Sana, and F. Scotti, "A high-level optimum design methodology for multimodal biometric systems," in *Proc. Int. Conf. Computational Intelligence Homeland Security Personal Safety CIHSPS*, Venice, Italy, July 2004, pp. 21–22.

[4] S.A. Israel, W.T. Scruggs, W.J. Worek, and J.M. Irvine, "Fusing face and ECG for personal identification" in *Proc. 32nd Applied Imagery Pattern Recognition Workshop*, Oct. 15–17, 2003, pp. 226–231.

[5] M. Gamassi, M. Lazzaroni, M. Misino, V. Piuri, D. Sana, and F. Scotti, "Accuracy and performance of biometric systems," in *Proc. IEEE Instrumentation Measurement Technology Conf.*, Como, Maggio Italy, 2004, pp. 18–20.

[6] M. Tistarelli, A. Lagorio, M. Jentile, and E. Grosso, "Design of a vision system for identity verification," in *Proc. 32nd Annu. Hawaii Int. Conf. System Sciences, HICSS-32*, Jan. 5–8, 1999, vol. Track 3, p. 9.

[7] T. Lynda, C. Claude, and B. Mohand, "Multiple query evaluation based on an enhanced genetic algorithm," *Inform. Processing Manage.*, vol. 39, no. 2, pp. 215–231, Mar. 2003.

[8] R. Clarke, "Biometrics and privacy," in *Proc. Computers, Freedom Privacy 2002*, San Francisco, Apr. 2002 [Online]. Available: http://www.anu.edu.au/people/Roger.Clarke/DV/Biometrics.html

[9] M. Crompton, "Biometrics and privacy—The end of the world as we know it or the white knight of privacy?" Mar. 2002 [Online]. Available: http//www.privacy.gov.au/news/speeches/sp80notes.pdf

[10] J. You, D. Zhang, J. Cao, and G. Minyi, "Parallel biometrics computing using mobile agents" in *Proc. 2003 Int. Conf. Parallel Processing 2003*, Oct. 6–9, 2003, pp. 305–312.

[11] S. Cherukuri, K.K. Venkatasubramanian, and S.K.S. Gupta, "Biosec: A biometric based approach for securing communication in wireless networks of biosensors implanted in the human body" in *Proc. 2003 Int. Conf. Parallel Processing Workshops*, 2003, Oct. 6–9, 2003, pp. 432–439.

[12] F. Amigoni, A. Brandolini, V. Caglioti, V. Di Lecce, A. Guerriero, M. Lazzaroni, F. Lombardo, R. Ottoboni, E. Pasero, V. Piuri, and D. Somenzi "Agencies for perception in environmental monitoring," in *Proc. IEEE Instrumentation Measurement Technology Conf.*, Como, Italy, 18-20 May 2004, pp. 1266-1271

[13] G. Weiss, *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, MA: MIT Press, 1999.

[14] M. Wooldridge, *An Introduction to Multiagent Systems*. Chichester, UK: Wiley, 2002.

Marco Gamassi received the Ing. degree in computer engineering from the Politecnico di Milano, Italy, in 2003. Since 2003, he has been a research associate with the Department of Information Technologies, University of Milan, Crema, Italy. His research interests include biometrics identification systems, signal and image processing, and soft-computing technologies for high-level system design.

Vincenzo Piuri received the Ph.D. degree in computer engineering from Politecnico di Milano, Milan, Italy, in 1989. From 1992 to 2000, he was an associate professor in operating systems at Politecnico di Milano. Since October 2000, he has been a full professor in computer engineering at the University of Milan, Crema, Italy. He was a visiting professor at the University of Texas at Austin during the summers of 1993–1999. His research interests include distributed and parallel computing systems, computer arithmetic, application-specific processing architectures, digital signal-processing architectures, fault tolerance, and neural network architectures. His original results have been published in more than 200 book chapters, international journals, and proceedings of international conferences. He is Fellow of the IEEE and member of ACM, INNS, and AEI. He was associate editor of *IEEE Transactions on Instrumentation and Measurement* and *IEEE Transactions on Neural Networks*. He is president of the IEEE Computational Intelligence Society and was vice president for Publications of the IEEE Instrumentation and Measurement Society, vice president for Member Activities of the IEEE Neural Networks Society, and member of the Administrative Committee of the IEEE Instrumentation and Measurement Society and the IEEE Computational Intelligence Society. In 2003, he received the IEEE Instrumentation and Measurement Society Technical Award for his contributions to the advancement of computational intelligence theory and practice in measurement systems and industrial applications.

Daniele Sana received the Ing. degree in computer engineering from the Politecnico di Milano, Milan, Italy in 2003. Since 2003, he has been a research associate with the Department of Information Technologies, University of Milan, Crema, Italy. His research interests include multimodal biometrics identification systems, signal and image processing, and multiagent systems.

Fabio Scotti received the Ing. degree in electronic engineering in 1998 and the Ph.D. degree in computer engineering in 2003 from Politecnico di Milano, Italy. Since 2003, he is an assistant professor at the Department of Information Technologies, University of Milan. His research interests include high-level system design, signal and image processing, and computational intelligence algorithms and their applications in the industrial field. His current research focuses on design methodologies and algorithms for multimodal biometric systems.

Olga Scotti (oscotti@dti.unimi.it) received the computer science doctor degree from Università degli Studi di Milano, Italy in 2001. Since 2001, she has been a research associate with the Department of Information Technologies, a branch of Università degli Studi di Milano. Her research interests include face identification and recognition, multimodal systems, and multiagent systems.