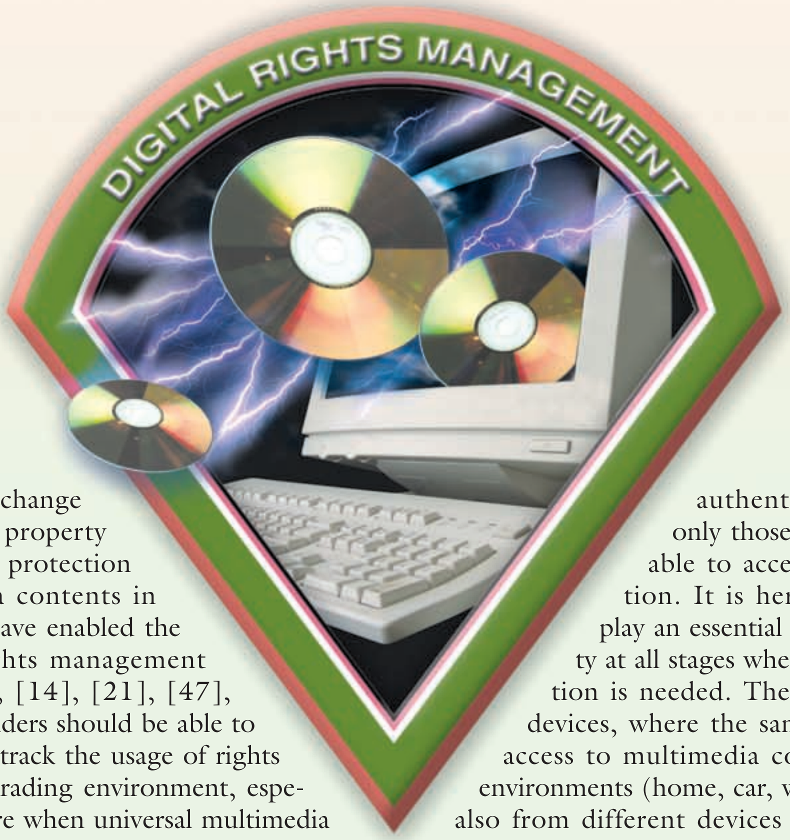
Authentication Gets Personal with Biometrics

Javier Ortega-Garcia, Josef Bigun, Douglas Reynolds,
and Joaquin Gonzalez-Rodriguez

Increasing security in DRM systems through biometric authentication.



LOGO COMPOSITE: ©1995 PHOTODISC, INC., ©DIGTIAL STOCK, ©COREL

Securing the exchange of intellectual property and providing protection to multimedia contents in distribution systems have enabled the advent of digital rights management (DRM) systems [5], [14], [21], [47], [51], [53]. Rights holders should be able to license, monitor, and track the usage of rights in a dynamic digital trading environment, especially in the near future when universal multimedia access (UMA) becomes a reality, and any multimedia content will be available anytime, anywhere. In such DRM systems, encryption algorithms, access control, key management strategies, identification and tracing of contents, or copy control will play a prominent role to supervise and restrict access to multimedia data, avoiding unauthorized or fraudulent operations.

A key component of any DRM system, also known as intellectual property management and protection (IPMP) systems in the MPEG-21 framework, is user authentication to ensure that only those with specific rights are able to access the digital information. It is here that biometrics can play an essential role, reinforcing security at all stages where customer authentication is needed. The ubiquity of users and devices, where the same user might want to access to multimedia contents from different environments (home, car, work, jogging, etc.) and also from different devices or media (CD, DVD, home computer, laptop, PDA, 2G/3G mobile phones, game consoles, etc.) strengthens the need for reliable and universal authentication of users.

Classical user authentication systems have been based in something that you have (like a key, an identification card, etc.) and/or something that you know (like a password, or a PIN). With biometrics, a new user authentication paradigm is added: something that you are (e.g., fingerprints or face) or something that you do or produce (e.g., handwritten signature or

1053-5888/04/$20.00©2004IEEE

voice). Biometric recognition, as a means of personal authentication, is an emerging signal processing area focused on increasing security and convenience of use in applications where users need to be securely identified. Biometric characteristics are inherently associated with a particular individual, making them insusceptible to being forgotten or lost.

There are many different biometric traits that can be used, each with various benefits and drawbacks, depending on the application scenarios and required accuracy. In this article, we outline the state-of-the-art of several popular biometric modalities and technologies and provide specific applications where biometric recognition may be beneficially incorporated. In addition, we discuss integration strategies of biometric authentication technologies into DRM systems so that the whole process meets the needs and requirements of consumers, content providers, and payment brokers, securing delivery channels and contents.

Personal Authentication Through Biometrics

The process of automatically associating an identity with an individual by means of some inherent personal characteristic is called biometric recognition [15], [31], [39], [52]. Traditionally, person authentication has been accomplished by associating to the person's identity something that he/she possesses (e.g., a key, a card, etc.) or knows (e.g., a password, a PIN). Biometric recognition adds a new dimension by associating a person's identity with something that he/she is (or produces). Something that a person is indicates a physiological characteristic inherently associated with the person, while something that a person produces indicates a trained act or skill that the person unconsciously does as a behavioral pattern.

Although there are several technologically mature physiological modalities, like fingerprint, iris, face, hand/finger(s) geometry, or palmprint recognition, other modalities are also described in the literature, such as retina or ear analysis, body (parts) thermogram inspection, vein structure (of the wrist), face/hand sweat pores, or objective odor measures. Regarding behavioral modalities, voice, handwriting, signature, and keystroking are the focus of most major research efforts, though modalities like gait recognition are also gaining interest. Despite its tremendous importance in modern forensic investigation, DNA-based authentication is not yet considered to be an automated means of authentication, since it requires manual intervention and is currently far from producing (near) real-time results.

Beyond sheer accuracy, there are many other factors to consider when examining a biometric solution for an application [15], [52], such as vulnerability to fraud, the degree of distinctiveness or uniqueness of the biometric, the intrinsic short- and long-term variability associated with the biometric, the intrusiveness of the system to collect the biometric sample (and the required user cooperativeness), the ease of user enrollment/recognition, and the long-term support required (database management, re-enrollment, template updating). Moreover, biometric authentication cannot be considered as producing error-free verification decisions under any application condition, making multilayered security a general condition for most applications. The need for increasing security, convenience, and accountability through biometrics has concentrated a great deal of research activities on these practical issues, which will be presented in this article.

Characterization of Biometric Systems

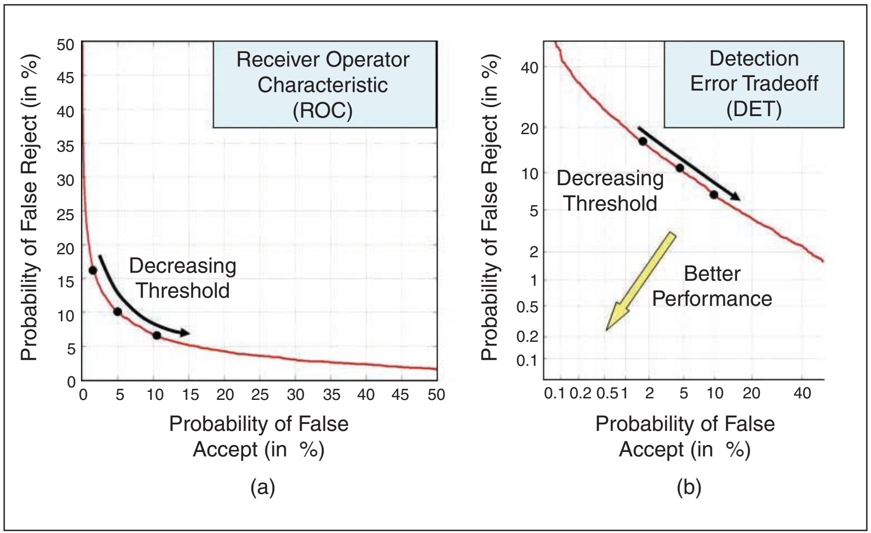
Biometric recognition is a generic term that encompasses the two main modes in which biometric systems operate: Biometric identification is the task of associating a test biometric sample with one of N patterns or models that are available from a set of known or registered individuals. It is also known as the one-to-many (specifically, one-to-N) task, and the output of this operation mode is normally a sorted list of candidate models, based on their degree of match with respect to the test sample. Biometric verification is the task of authenticating that a test biometric sample matches the pattern or model of a specific user. It is also known as the one-to-one problem, and the output is a binary decision (accept or reject), that is usually based on comparison of the match score between the test sample and the claimed user's model or pattern to a decision threshold.

Most biometric systems in commercial applications operate under the verification mode, as one-to-one matching is the main task facing security concerns (is the user who he/she claims to be?). Identification mode is mainly used for database searches, such as in criminal fingerprint matches, or for a small-scale user group (N on the order of 5–10) searches. All biometric systems operate in two separate stages: the enrollment phase and the testing phase. During enrollment, biometric samples from a user are used to produce, generate, or train a pattern or model from the user. This is a key process, as the resultant pattern or model represents the biometric "identity card" of each enrolled user. Care must be taken to ensure the true identity of the enrollee is established at this stage (generally using human inspection of various trusted documents), otherwise the whole system is compromised. In the testing phase a biometric sample from a person is identified or verified against the enrolled models.

Biometric verification can be considered as a detection task, involving a tradeoff between two types of errors: Type I error, also denoted as false rejection (FR), false nonmatch or miss (detection), occurring when a true user (also referred to as the client, target, genuine or authorized user) is rejected by the system, and Type II error, known as false acceptance (FA), false match or false alarm, taking place when an impostor is accepted as being a true user. These two types of errors can be traded off against each other by varying the decision threshold. A more secure system aims for low FAs at the expense of higher FRs, while a more convenient system aims for low FRs at the expense of higher FAs. The desired operating point is highly dependent on the final application and so, generally, the complete tradeoff curve over many operating points is often used to characterize biometric performance.

Performance capabilities have been traditionally shown in the form of ROC (receiver- or relative-operating characteristic) plots, in which the probability of a false-acceptance is plotted versus the probability of a false-rejection for varying decision thresholds. An example of an ROC plot is given in Figure 1(a), where the desired area is at the lower left of the plot, in which both types of errors are minimized. Unfortunately, with ROC plots, curves corresponding to well-performing systems tend to bunch together near the lower left corner, impeding a clear visualization of competitive systems. More recently, a variant of an ROC plot, the detection error tradeoff (DET) plot [26] has been used, which plots the same tradeoff using a normal deviate scale. This has the effect of moving the curves away from the lower left corner when performance is high and producing linear curves, making system comparisons easier. In Figure 1(b), the DET plot corresponding to the same data in the ROC plot in Figure 1(a) is shown.

Although the complete DET curve is needed to fully describe system error tradeoffs, it is desirable to report performance using a single number. Often the equal-error-rate (EER), the point on the DET curve where the FA rate and FR rate are equal, is used as this single summary number. However, the suitability of any system or techniques for an application must be determined by taking into account the various costs and impacts of the errors and other factors such as implementations and lifetime support costs and end-user acceptance issues.



▲ *1. Example of verification performance comparison for same hypothetical systems, A and B, for both (a) ROC and (b) DET plots.*

Overview of Biometric Modalities

This section provides a brief description of state-of-the-art biometric modalities that are considered as suitable for incorporation into DRM systems. Suitability refers to both the applicability for DRM schemes and the maturity of the technology. Modalities are divided into physiological traits, including fingerprint, face, hand geometry, palmprint, and iris biometrics, and behavioral ones, including voice, handwriting, and off- and online signature. Finally, we will also define multimodal biometrics as a promising way of enhancing performance through the combination of multiple modalities.

Physiological Biometrics

Fingerprint Biometrics

Fingerprint matching [15], [16], [23] is one of the most widespread biometric solutions and is based on the ridge structure of the epidermis of each fingertip and its peculiar distribution of minutiae points, which is preserved almost unalterable all through a person's life. Traditional ink-and-roll fingerprint images have been replaced by electronic scanning through different technologies (optical, capacitive, ultrasonic, etc.) Unfortunately, these devices may introduce some degree of variability on fingerprint patterns, due to several effects, like the sensitivity to finger temperature or sweat, the distortion due to pressing on the planar surface, fingertip placement, or the size-limited nature of the device that adds position variability to the fingerprint. Moreover, there are some population groups with special problems, like some manual workers that make intensive use of manual tools that can lead to fingerprint damaging by friction and erosion, an issue that can affect up to 5% of the population [30].

Features used to represent fingerprints for person authentication purposes are predominantly minutiae based [15], [16], [23]. These are typically end points or bifurcation ridges. The relative position of these minutiae points constitute a personal trait. A common way of extracting minutiae by image processing is by use of a lack of linear symmetry, introduced independently in [4] for (N-D) and in [17] (for 2-D). First, smoothed outer products of gradients are computed yielding 2×2 symmetric matrices (tensors). In the case of ill-defined orientations, typically caused by bifurcations, or end of lines, the smallest eigenvalue of the matrix tends to be large; otherwise, it is close to zero. A thresholding of the smallest eigenvalue usually gives satisfactory results to extract the minutiae.

For poor-quality fingerprints, however, it has been shown that the image enhancement by imposing the linear symmetry orientations back to local neighborhoods improves the fingerprint-based authentication significantly [40]. Recently [1] and [28] introduced landmark (arch, deltas) based alignment and identification results that completely bypass minutiae extraction. This is interesting for consumer electronic fingerprint sensors that tend to be small and hence the number of minutiae is necessarily less than ink-rolled-based imaging techniques.

Results of the 2000 and 2002 Fingerprint Verification Competitions (FVC2000 and FVC2002, repectively) [23] reveal that even with medium- to high-quality images, only a few technologies show good performance. This means there is still room for algorithmic improvement in terms of image processing to extract the salient features and match them despite intra-class variability. An improvement is also needed regarding acquisition devices, which provide size-reduced, rather poor-quality images.

Face Biometrics

The face of a person is considered to be the most immediate and transparent biometric modality for physical authentication applications. As far as adequate camera positioning can mitigate issues with users properly looking into the camera, cooperativeness is not in some cases a crucial issue. Nevertheless, in other uncontrolled situations, like when trying to detect a particular face in a crowd, cooperativeness is an important issue. This makes face recognition a highly desirable biometric modality, and extraordinary research efforts have been undertaken in the last decade.

A wide choice of techniques has been proposed to meet the demands of automatic person authentication by their faces. Despite its intrinsic complexity, face-based authentication still remains of particular interest because it is perceived psychologically and/or physically as noninvasive. Significant motivations for its use include the following:

- ▲ Face is a modality that humans largely depend on to authenticate other humans; consequently, every human is a putative expert in face recognition from infancy.
- ▲ Face is a modality that requires no or only weak cooperation to be useful.
- ▲ Face authentication can be advantageously included in multimodal systems, not only for authentication purposes but also to confirm the liveness of the signal source of fingerprints, voice, etc.

Unfortunately, face is a three-dimensional (3-D) modality that is usually captured through a 2-D device (photo or video camera), causing problems related to the subject's pose. Background and illumination are also problems not yet fully solved, and inherent variability, like aging, must also be considered. Of course, factors producing external variability in the face, like make-up, hairdressing (also beards or mustaches), and artifacts (glasses, jewelery, piercing, etc.), will cause degradation in performance.

Biometric recognition adds a new dimension by associating a person's identity with something that he/she is (or produces).

Features used for description of faces are either geometric metrics of the face, like distances between the nose and mouth [8], or more abstract features, like filter responses on a grid [19]. To extract geometric metrics, subparts of the face must be extracted. A common approach to do this is by using scalar products (correlate image templates) that attempt to match well-defined portions of the face (eye, mouth, etc.) with a reference template [8]. Therefore, geometric metrics are also known as being scalar product approaches. The eigenface approach [44] describes a face image in terms of linear combinations of basis images that also belongs to this class. The eigenface matching can be performed as a scalar product between the reference face image coefficients and the test image coefficients. But since the scalar products are conserved in orthogonal transformations, this result can also be obtained from a scalar product (i.e. pixelwise multiplication and summation) of the approximately reconstructed reference face image and the approximately reconstructed test face image. The number of basis images used is set empirically. Although early studies [44] indicated a face-space of dimensionality of approximately 20, later studies indicate a dimensionality of the order of hundreds [42] for relatively large face corpora. Finally, other face detection approaches related to the eigenface approach is the distance-from-feature-space [33] and Fischerfaces [2].

Methods that constrain local, highly nonlinear features by adding geometrical constraints can be considered as a mixture of both geometric and scalar product approaches. These include the dynamic link architecture (DLA) [19] and related graph-based feature matching approaches [24], as well as methods based on neural networks, and feature-based approaches where features are geometrical measures [8]. In DLA, the mechanism for assessing connections between the image and model domain turns out to be complex and time consuming. A simplified implementation called elastic graph matching (EGM) is often preferred for finding objects in the scene with a known reference [49]. However, as the attributed graph is a 2-D representation of 3-D objects, this tolerance is limited. Extensions have been proposed for rotations in facial depth [50]. In Duc et al. [11], local discriminant measures for face images are proposed. This leads to significantly improved performance for face authentication applications. In that contribution, subsets of the data are considered separately, leading to a faster training, as the solution to the optimization problem is known analytically.

The main regions of interest for the face recognition/authentication task are well known: the eyes and the mouth of a subject. A biologically inspired method to locate such facial landmarks using a Saccadic search strategy built around a rigid log-polar retina, which is used to sample the Gabor decomposition of the image [41], as shown in Figure 2, has recently been proposed. The Saccades make use of a priori knowledge of the face components in the form of appearance-based models of the eyes and the mouth. The models describe the Gabor signature of the target features. Face authentication is achieved using experts of each facial region, i.e., the two eyes and the mouth. The results report significant improvement on the Gabor features extracted on grids spanning the full face that have been employed previously [11], [19].

Iris Biometrics

Iris verification is based on the stability of the so-called trabecular pattern during a person's lifespan. This trabecular pattern is formed by an elastic connective tissue called trabecular meshwork, which gives to the iris its appearance of radial divisions. It consists of pectinate ligaments adhering into a tangled mesh revealing striations, ciliary processes, crypts, rings, furrows, freckles, vasculature, and a corona. The iris is protected by the cornea and the aqueous humor, so the iris pattern is almost unaffected by environment, except for pupillary reflex to light. This stretching of the iris tissue as the pupil dilates produces elastic deformations. Locating the inner and outer iris boundaries can lead to a linear compensation for this effect.

The segmentation of the inner and outer boundaries, together with the detection of the eyelids (if they intrude into the iris pattern), require the application of image processing techniques [10], [48]. A first approach to the problem solution [10] is accomplished by extracting concentric coronas from the iris. On each of these coronas, a 2-D Gabor wavelet transform is applied to represent image texture by the arguments of the complex filter responses. Each of these phasor angles is quantized into just the quadrant in which it lies. This operation is repeated for all filters and regularly selected spatial sites on concentric circles of the iris. In case some sites fall in nonvisible places due to an eye-lid covering the Gabor filter measurement site, the site is marked as nonvalid by a sites of interest mask. This site mask together with the Gabor-filter-based local phase measurements just described constitute the iris code.

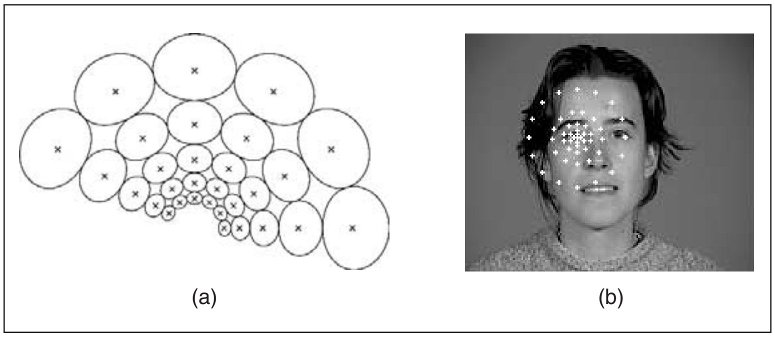
Left and right iris patterns of a given person are different. Also, iris patterns between identical twins are different. However, a nontrivial issue is securing the liveness of the signals and obtaining the high-quality images that are required to represent the iris patterns that make them so personal. Unlike face scan technology, which can leverage existing photo- or video-camera technology, iris scan deployments require specialized devices including, in some cases, infrared illumination and may be perceived as invasive by users who are required to be very collaborative. The iris code, and for that matter iris features reported by other researchers, relies on the exceptionally high-quality images that a combination of the iris image acquisition software and user collaboration allows.

Hand Biometrics

Palmprint Recognition

Palmprints are stable and show high accuracy in representing each individual's identity [52]. Thus, they have been commonly used in law enforcement and forensic environments. Since the surface of the palmprint is larger than the fingerprint itself, a higher quantity of identifying features can be extracted from the palmprint. Moreover, users consider hand biometrics as being user friendly, easy to use, and convenient. Palmprint acquisition is based on standard charge-coupled device (CCD)-based optical scanning. Although some acquisition procedures imply pressing on a glass panel (inducing an elastic distortion on the palmprint), some others do not. Those who do not however, must solve the liveness issue separately.

Palmprint features can be divided into three different categories: a) point features, which include minutiae features from ridges existing in the palm, and delta point features, from delta regions found in the finger-root region; b) line features, which include the three relevant palmprint principal lines, due to flexing the hand and wrist in the palm, and other wrinkle lines and curves (thin and irregular); and, c) texture features of the skin.



▲ *2. A biologically inspired method to locate facial landmarks using a Saccadic search strategy built around a rigid log-polar retina (a), which is used to sample the Gabor decomposition of the image (b).*

Hand Geometry

Hand geometry recognition [15] is based on the extraction of a hand pattern that incorporates parameters like finger length, width, thickness, curvatures, or relative location. To obtain these features, an image of the silhouetted hand is needed. The process of capturing this information is normally accomplished through CCD cameras and infrared illumination; the user puts his/her hand on a highly reflective surface, such as a platen, performing an orthographic scanning, consisting of top and side views of the hand shape. Surface details like texture and fingerprints are ignored for this purpose. Specific hand positioning is forced by using inter-finger pegs or locator pins.

Hand geometry requires high collaboration from the users as the hand must be kept flat while scanning. Fingernails (e.g., on females) that can potentially deteriorate intra-class variability must also be coped with by the system.

Behavioral Biometrics

Voice Biometrics

The speech signal conveys many levels of information to the listener. At the primary level, speech conveys a message via words, but at other levels speech conveys information about the language being spoken and the emotion, gender, and, generally, the identity of the speaker. While speech recognition aims at recognizing the words spoken in speech, the goal of automatic speaker recognition systems is to extract, characterize, and recognize the information in the speech signal conveying speaker identity. In this section we provide a brief overview of the area of speaker recognition, describing underlying techniques and some indications of performance. This is not a comprehensive review, and readers should see, for example, [9], [35], and their references for more details.

Depending on the level of user cooperation and control in an application, the speech used for these tasks can be either text dependent or text independent. In a text-dependent application, the recognition system has prior knowledge of the text to be spoken and it is expected that the user will cooperatively speak this text. In a text-independent application, there is no prior knowledge by the system of the text to be spoken, such as when using extemporaneous speech. Text-independent recognition is more difficult but also more flexible, for example, allowing verification of a speaker while he/she is conducting other speech interactions (background verification).

Research and development on speaker recognition methods and techniques continues to be an active area for well over four decades. Approaches have spanned from human aural and spectrogram comparisons, to simple template matching, to dynamic time-warping (DTW) approaches, to more modern statistical pattern recognition approaches, such as neural networks and hidden Markov models (HMMs). Over this same time, research and development corpora have evolved from small, private corpora (five to ten speakers) under laboratory clean, controlled conditions (single session, read speech) to large, publicly available corpora (500+ speakers) reflecting more realistic and challenging conditions (extemporaneous speech from landline and cellular telephone channels). Benchmark evaluations using common corpora and paradigms have been conducted for several years (e.g., YOHO, CAVE project, NIST [29]) allowing comparison of technical approaches and focusing effort on common challenges.

Although there are no exclusive speaker identity cues in the speech signal, information about the speaker's anatomical structure is generally conveyed in the amplitude spectrum, with the location and size of spectral peaks (formants) related to the vocal tract shape and the fine structure (pitch striations) related to the glottal source. Typically, the amplitude spectrum is estimated using 20 ms of speech and a physiologically motivated mel-scale filter-bank every 10 ms so as to capture the evolving nature of the vocal apparatus. The amplitude spectrum is then further processed through a form of cepstral analysis and appended with time derivatives to produce the feature vector that is used to create a speaker model. Recent research, which is beyond the scope of this article, is focused on using other aspects of speech, such as pronunciations, prosody, and word usage, to help better characterize and recognize speakers [36].

In most speaker recognition systems, the speaker models are generally some form of HMMs. From published results, the use of HMMs has been described as generally producing the best performance compared to other models. HMMs encode the temporal evolution of the features and efficiently model statistical variation of the features, providing a statistical representation of how a speaker produces sounds. For text-dependent applications, whole phrases or phonemes may be modeled using multistate left-to-right HMMs. For text-independent applications, single-state HMMs, also known as Gaussian mixture models (GMMs), are used.

For speaker verification systems, the decision is usually made by a likelihood ratio test computed between the claimed speaker's model and a model representing impostor or generic speech. This alternative model is often called an impostor, background, or cohort model. The use of an impostor model is widespread and can be crucial to obtaining good performance. Basically it acts as a normalization to help minimize nonspeaker related variability (e.g., text, microphone, noise) in the likelihood ratio score [35].

Speech is a natural signal to produce that is not considered threatening or intrusive by users to provide. In many applications, such as telephone applications, speech may be the main, or even the only, modality. The telephone system provides a ubiquitous, familiar network of sensors for obtaining and delivering the speech signal. Moreover, there is no need for special transducers to be installed at application access points. Even for nontelephone applications, like PC-based ones, sound cards and microphones are low cost and readily available. Combined with utterance verification, speaker verification is one of the few biometrics that supports a natural "challenge-response" to help thwart spoofing attacks. A system can present a user with a series of randomized phrases to repeat so the system can verify not only the voice matches but also the required phrases match. Additionally, it is possible to use forms of automatic knowledge verification where a person is verified by comparing the content of his/her spoken utterance against the stored information in his/her personal profile (e.g., "What is your pet's name?").

On the other hand, speech is a behavioral signal that may not be consistently reproduced by a speaker and can be affected by a speaker's health (cold or laryngitis). Also, the varied microphones and channels that people use can cause difficulties since most speaker verification systems rely on low-level spectrum features susceptible to transducer/channel effects. Furthermore, the mobility of telephones means that people are using verification systems from more uncontrolled and harsh acoustic environments (cars, crowded airports), which can stress accuracy.

Handwritten Biometrics

The convenience for paper and pen in the electronic era is the reason why people still use handwriting as a mean to convey, retain, and facilitate communication. Together with this kind of information, handwriting is also a skill that individualizes people [34]. From this point of view, the process of automatically determining who the specific author of a given handwritten text is is called writer recognition. Handwritten recognition can be accomplished from two different points of view, depending on whether there is electronic access to the handwriting process or not. If there is, one can digitize the pen's instantaneous information trajectories, and information like pressure, speed, or pen-up movements can be captured; if not, just shape-based image recognition is feasible. The former is also known as online or dynamic handwriting recognition, whereas the latter is often called offline or static recognition.

A particularly relevant component in handwriting biometrics is signature recognition, because of the social and legal acceptance and widespread use of the written signature as a personal authentication method. Regardless of the handwritten content, signature verification is entirely focused on extracting writer-specific information. It has also to be taken into account that FA refers in this modality mainly to the impostor's ability to mimic the target signature by producing a forgery. Moreover, devices like PDAs, pocket PCs, tablet PCs, or 3G mobile phones might offer handwriting capabilities, due to the fact that handwriting is considered as being more natural for humans and also to the possibility of size reduction by eliminating the keyboard.

Online Signature Verification: In online signature verification systems [20], [32], different approaches can be considered to extract signature information; they can be divided into: i) function-based approaches, in which signal processing methodology is applied to the dynamically acquired time sequences (i.e., velocity, acceleration, force, or pressure), and ii) feature-based approaches, in which statistical parameters are derived from the acquired information. One can also specify different levels of classification, so it is possible to use and combine shape-based global static (i.e., aspect ratio, center of mass, or horizontal span ratio), global dynamic (i.e., total signature time, time down ratio, or average speed) or local (stroke direction, curvature or slope tangent) parameters.

The use of complete sequences have so far yielded better results, since reducing time sequences just to statistical features diminishes our ability to make a precise characterization of this dynamic process. This time-based sequence characterization process is strongly related to the way in which a reference model is established. HMMs have shown this capability regarding other behavioral-based biometric traits, outperforming other classical approaches like distance measure, (weighted) cross correlation, or dynamic time warping (dynamic string matching). Online signature verification offers reliable identity protection, as dynamic information is not available on the signature image itself but in the process of signing. Also, pen-up dynamic information can be acquired, and these pen-up trajectories do not leave even their shapes in the final image.

Offline Signature Verification: Offline signature verification relies on extracting writer-specific information just from the shape of the image and the luminance of the trace [20], [34]. Once the signature has been extracted from the document background, several techniques have been used for offline signature verification including minimal distance classifier, nearest neighbor, dynamic programming, neural networks, and HMMs.

Unlike online signature verification, off-line signature verification cannot take full advantage of the dynamic handwriting process. Performance results of offline systems are expectedly lower than that based on online information, although false accepts rely entirely on the ability of the forger to produce a highly skilled, shape-based forgery. The extraction from offline signatures of pseudodynamic information, like recovery of the stroke sequence, or deriving the instantaneous pressure from the stroke width, is now the focus of some additional research efforts.

Multimodal Biometrics

The performance of any single-trait verification system can be improved by unimodal (or monomodal) fusion, i.e., the combination of several verification strategies applied on the same input data. Even greater verification performance improvement can be expected through the use of multiple biometric characteristics, due to their statistical independence [3], [37]. Inspired by this potential, great effort has been made to demonstrate the benefits of the multimodal fusion approach [3], [8], [18], [37].

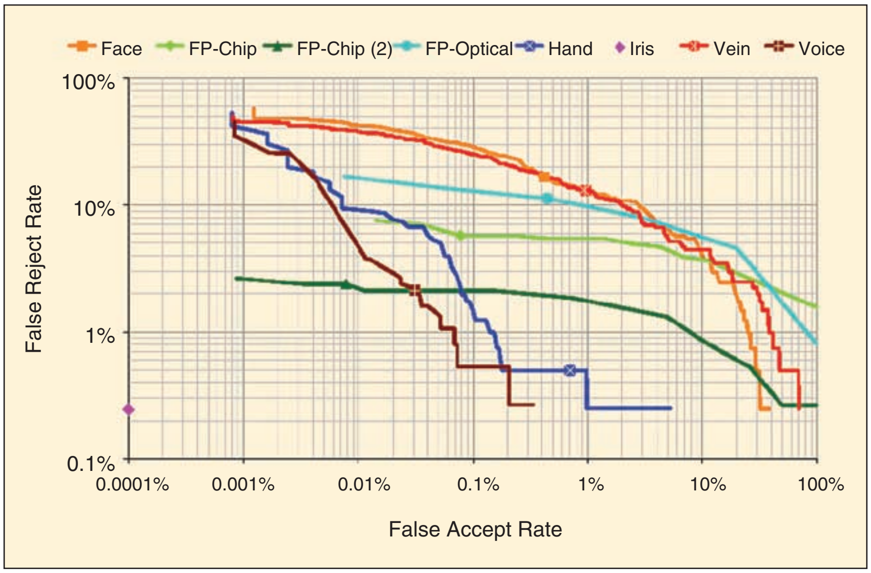
Biometric multimodality can be studied under the general field of data fusion or under the particular frame of classifier combination. Bigun et al. in [3] combined machine expert opinions (applied to the frontal face and speech) in a probabilistic Bayesian framework, whereas Kittler et al. in [18] showed that weighted summation of machine expert opinions outperforms other strategies, including product rule, by using Bayesian arguments. Interestingly, both approaches independently show results in favor of weighted machine expert summation rules, via Bayesian methodologies, by using both real data (speech, frontal, and profile images modalities) and simulated data. The approach of [3] is based on Bayesian results obtained from catastrophe studies where human expert opinions must be combined to do assessment on probability of (undesired) events. In an analogous manner, (machine) experts' future opinions where no true data are available (operational state of an authentication system when the true identity is not available) are predicted by using the Bayes' theorem from the past performance of each expert by weight optimization. In this approach expert opinions are automatically calibrated by their "historical bias" they have shown in the past. In the approach of [18], R modalities, two classes (ω_1 for clients and ω_2 for impostors), and a given pattern Z that generates the feature vector $\mathbf{x}_i$ for modality i, the classifiers (or experts) are considered to give the a posteriori probability for each class k: $P(\omega_k \mid \mathbf{x}_i)$. Several ways to implement the fusion of the modalities are then obtained (sum, product, max, min, etc.), based on the Bayes' theorem. The sum rule outperformed the remainder in the experimental comparison, due to its robustness to errors in the estimation of $P(\omega_k \mid \mathbf{x}_i)$ made by the individual classifiers.

Multimodal fusion can also be treated as a pattern classification problem. Under this point of view, the scores given by individual expert modalities are considered as input patterns to be labeled as accepted/rejected (for the verification task). Verlinde et al. followed this approach and compared in [45] the following pattern classification techniques for multimodal fusion (sorted by relative decreasing performance): logistic regression, maximum a posteriori, k-nearest neighbors classifiers, multilayer perceptrons, binary decision trees, maximum likelihood, quadratic classifiers, and linear classifiers. More recently [13], support vector machines (SVMs) have been compared with all the above-mentioned techniques carrying out the same experiments, showing very promising results. From now on, this perspective will be referred to as learning-based (or trained) fusion, because it requires sample outputs from the experts to train the pattern classifiers.

Performance Characterization in Biometrics

Although there have been some attempts to directly compare performance of different biometric modalities (for example, see Figure 3, used with permission from [25]), it is still difficult to characterize the performance of different biometric recognition systems in a consistent way. One reason for this is that there are many factors that produce degradation in recognition performance, and these factors are not homogeneous throughout biometric modalities. Factors like the type of application, enrollment and/or testing scenario, size of the population under study, controlled situations versus uncontrolled situations, etc., introduce a heterogeneous assessment framework. There are also intra-modality factors (e.g., type of acquisition device, fingertip position, or finger humidity in fingerprint matching; illumination, pose, face artifacts, or background in face recognition; transmission channel, noise, or type of handset in speaker verification; online versus offline acquisition, or degree of skill of forgeries in signature biometrics, etc.) that are almost impossible to equalize between modalities to make absolute comparisons.

One observed theme in these cases is that performance tends to improve with increasing constraints on the application (more biometric samples, less distortion/noise/artifacts, well-performing acquisition devices, cooperative users, etc.). Determining acceptable performance for a particular application will depend on the benefit of replacing any current verification procedure, the threat model (claimant to impostor attempts), and the relative costs of errors.



▲ *3. From [39], DET plot showing biometric performance for six different modalities, namely, face, fingerprint, hand geometry, iris, veins, and voice. Regarding fingerprints, results include performance with three different sensors, two of them being capacitive chips and the third being optical.*

Integrating Biometrics in DRM Systems

Security Analysis of Nonbiometric DRM Approaches

A content provider has mainly two ways to assure that content is delivered only to appropriate authorized clients [47]. The first way is to set up a secure connection with the client and to authenticate him as a trusted client before delivering any content. This type of content delivery is most suited to the scenario of a consumer obtaining content in a session, which is an important but very limited part of the DRM world and has the problems associated with key management (e.g., storing, password-based access to the encryption keys). The second way to deliver content only to authorized clients is to deliver contents in an encrypted form that can only be decrypted by a trusted client. This approach is particularly suited to store-and-forward content delivery systems, allowing for super-distribution (any user can redistribute with no rights violation) while the consumer remains anonymous to the content provider. A further advantage is that it separates the content payment and authorization from the content delivery, simplifying the content distribution. However, there are issues with ensuring that only authorized clients obtain and use the decryption keys.

Most current DRM systems are realized in software, so there is nearly no protection against serious attacks. Technical components of DRM systems consist of special adapted and well-known IT security functions. As shown in [12], pirates can reverse-engineer DRM systems, making them independent of hardware/software specific devices/tools (e.g., useable on other platforms) and proliferating this knowledge via freely available tools.

Existing DRM systems base their security structure on different combinations of the following technologies.

▲ *Watermarking:* This technique enables the inclusion of imperceptible information within audio/video/document digital content. It is primarily used for copy control and illegal distribution detection, usually including right-holder information. A secret key must be used to reveal the watermark allowing later extraction by the content provider.

▲ *Fingerprinting:* The digital content is watermarked with the consumer's identity for every client. Usually used in multicast encrypted streaming of content, users have individual decryption keys containing fingerprints (personal data). It enables fair use copies (e.g., copies between members of a family) providing control of the origin in case of illegal copies.

▲ *Tamper-resistant software/hardware:* This is responsible for securely handling the decryption and rendering of the content for the end consumer. It must prevent the consumer from gaining access either to the decryption key or to the decrypted digital content.

▲ *Encryption:* This is the process of data locking based on encryption/decryption keys. Both symmetric and public key infrastructure (PKI) approaches have been shown extraordinary useful in a wide range of applications from secure communications to e-commerce.

Many successful DRM and non-DRM systems rely on these technologies, but they are still susceptible to serious attacks to crack or circumvent them. The main vulnerability is in key management. First, there are many keys (for encryption, decryption, watermarking, and fingerprinting) needed by many people (individual consumers, system administrators, dealers and distributors, payment brokers, etc.) increasing the potential for compromise. Second, since cryptographically strong keys are not amenable for people to remember and enter, at one point, key entry is reduced to some easily remembered passcode. Thus, the security of the encryption system, and so that of the DRM, is only as good as the passcode, with all its well-known problems (the FBI ranks problems with "accounts with no passwords or weak passwords" as second in its "Top 20 Security Holes").

One of the main problems with passcodes is the lack of direct connection between users and passcodes, so a legitimate user and an impostor who fraudulently acquires the user passcode are indistinguishable to the system. Biometric authentication provides this linkage between users and passcodes for key security using a biometric technology to secure the cryptographic key. Apart from the different biometric technologies that can be adopted, each one of them can be applied in different ways, as discussed below, adding another security layer to the system giving access to the encryption keys, or directly generating keys from biometric data.

Overview of DRM Systems Integrating Biometrics

In the first part of this article, different biometric modalities and technologies, most of them commercially available, together with their respective levels of performance for different application scenarios, were presented. However, it is not easy to select a single biometric trait for a target application. For instance, end users might not accept invasive authentication techniques, or system vendors might not desire expensive acquisition sensors. Moreover, a specific biometric trait cannot be considered as a standalone technology but as a component in the overall global solution, adapted to the target application by the biometric technology provider, and easy to integrate into the DRM system.

The selection of the adequate biometric modality is just the first task to integrate biometrics in this new framework. DRM systems are extremely complex, with both local and distributed resources and processes and different types of clients and providers. For instance, a content distributor (such as a TV station, an Internet pay-per-download music site, or a wireless 3G entertainment provider) acts simultaneously as a client from content creators (e.g., a Hollywood film studio or a game software company) and as a provider for end customers or additional distribution layers. Consequently, there are two main kinds of consumers of intellectual rights, namely people (as individuals, groups, or companies) acting as end users, and distribution clients acting as intermediaries, and both kinds of entities need to be securely authenticated.

Moreover, some DRM applications concern transactions between companies, which have to be authorized and signed off by chief executives (for example, a new catalog from a content provider that is to be offered by a content distributor). The use of biometric tokens as electronic signatures in electronic documents, as well as the access to reserved executive keys by means of biometric verification, are both viable ways to offer instantaneous access to protected content just offered to your company.

However, the biggest area of interest for biometrics into DRM systems is that of key management. In complex structures such as DRM systems, many keys, passwords, digital certificates, and electronic signatures, each one of them with different levels of permissions, have to be managed properly, and here biometrics provides a natural way to guarantee that only the real individual has access or can generate his correspondent specific key.

Biometric-Based DRM Approaches

Both the "secure the pipe" and the "secure the content" scenarios apply to biometric user authentication when users are humans. Biometric technologies will work to provide secure authentication of the user in every task or target application where a human consumer of multimedia content is involved, either as a standalone technology or as passport to other well-established technologies (i.e., by using biometrics to unlock your Internet digital certificate).

Several commercial biometric solutions exist that are applicable to those target scenarios, ranging from online signature recognizers in tablet PC environments, to fingerprint readers in USB keys, to integrated cameras on 3G mobile phones, to multimodal speaker verification integrated in vocal portals and dialogue systems. This is just the beginning, as many other applications will be available in the near future when new devices will integrate high processing capabilities with multimodal acquisition devices. Applications are countless, and system designers must look for the best biometric solution (in terms of performance, acceptability, and cost-effectiveness) for each specific target application.

Below we outline several examples of biometric technologies in DRM systems, but a complete listing of all existing systems or proposals is out of the scope of this work. The selection is based both on covering the wide variety of biometric technologies and easy-to-access documentation of the reported systems.

We will differentiate two kinds of biometric system applications in the two next subsections. The first set of applications involves the use of conventional biometric systems as an additional security layer for access to restricted data. The second set of applications involves the use of biometrics for encryption key generation, which gives an extra security layer compared to mere access control, as the encryption key is never stored and is dynamically generated from the biometric data both for encryption and decryption of content.

DRM Systems with Biometric-Added Security

In this case, when a user wishes to access a secured key, he/she will be prompted to provide a biometric sample. If the user is accepted after matching of the input biometric sample with the enrollment data, the key is released and can be used to encrypt or decrypt the desired data. Biometric authentication is generally coupled with password and, occasionally, token authentication.

Here any existing biometric technique or system can be adapted to work in the DRM environment as an additional security layer. The usability of biometrics for the creation and retrieval of electronic signatures, strongly linked with DRM and encryption, has been extensively studied in [38]. In this work most existing biometric technologies are analyzed in the environment of smartcard technologies, one of the most prominent representatives of technologies for secure signature creation devices (SSCDs), as it is capable to execute signature algorithms and to provide storage and access to certificates. The pros and cons of each technology are thoroughly studied, with conclusions oriented to the smartcard and electronic signature environment, which can directly be extended to DRM systems.

However, even if providing an added security layer relative to nonbiometric password-based systems, biometric-based systems are still not perfect. First, the bio-protected key is (securely) stored somewhere in a password-protected system, so it could probably be accessed by a software attacker or by physical access to hardware/disks. Second, a compromised biometric pattern (fingerprint spoofing, voice recording, etc.) is impossible to replace.

To mitigate the compromise problem, a form of an interactive response system is often used. These procedures cannot be applied to static biometrics, such as fingerprints or iris recognition, but are primarily useful for dynamic biometrics, such as voice or online writing. A very interesting solution in the voice domain is known as conversational biometrics, as shown, for example, where the speaker is authenticated not just from the acoustic characteristics of his/her voice, as in conventional voice biometric systems, but also from his/her particular way of interaction with a dialogue manager capable of natural language understanding in a restricted domain (e.g., access to the catalogue of a content provider). In this sense, the process of authentication is never the same, and even recordings of the true speaker from previous access attempts will not give access to the system in a different trial.

The face of a person is considered to be the most immediate and transparent biometric modality for physical authentication applications.

DRM Systems with Biometric Key Generation

A first approach to integrate biometric patterns and encryption keys is to hide the cryptographic key in the enrollment template via a secret bit-replacement algorithm [43]. If the user is authenticated through biometrics, then it looks into the enrollment template in the specified positions for the bits that constitute the key. However, an attacker can learn from different user enrollment templates the positions or algorithm that gives access to the key bits, thus compromising the keys.

A second approach is discussed in [6], where the key is directly obtained from the biometric template, from direct template coding. In this proposal, two main problems appear. The first one is that as a result of changes in the biometric template due to environmental and physiological factors, the biometric template is generally not consistent enough to be used as a cryptographic key. The second problem is that if the key is ever compromised, then the use of that particular biometric is irrevocably lost. Note that this is the same consequence as above but from different origin.

A similar approach in [46] uses statistical features of online signatures instead of biometric images. This method allows obtaining a biometric hash vector based on an individual interval matrix, where cryptographic keys can be directly obtained from that hash vector. The idea relies on establishing confidence intervals, based in the user intra-variability, in 24 different features of a signature, and checking with the test signature if every feature is or is not between the corresponding limits, accepting (bit = 1) or rejecting (bit = 0) the feature as original from the author, obtaining finally the 24-bit hash vector if the signature comes from the true signer. Another interesting example using voice [27] uses text-dependent (fixed) passwords to generate the cryptographic key. Thus, segmental vector quantization is used to define m consecutive clusters that are represented by their centroids. During tests, every aligned segment is distance-based checked as accepting (1) or rejecting (0) the segment. An additional error correction procedure is implemented to allow for insertions/deletions up to 3 or 4 bits, for a total length of $m = 46$ bits. So, a double objective is achieved, obtaining unpredictable and reproducible keys. Unpredictability is achieved from the entropy from how the user speaks the password; that is, an impostor knowing the password will not obtain the (correct) key after uttering it. As happened in [6], users are generally not consistent in all features, so the challenge to obtain reproducible keys from biometrics relies on accommodating variations in those features in which a user is inconsistent while still generating the same key each time.

The same idea regarding the generation of the biometric key is presented for fingerprints in [43], allowing for reproducible keys with different images (rotation, translation, elastic skin distortions) from the same fingerprint. But the innovation in this approach comes from the fact that during enrollment the biometric image is combined with a digital key to create a secure block of data. This data block is secure in that neither the fingerprint nor the key can be independently obtained from it. During verification, the cryptographic key is retrieved by combining the biometric image with the secure block of data. Thus, it does not simply provide an accept/reject authentication decision to release the key but instead retrieves a key that can only be recreated by combining the input biometric image with the secure block of data. In this way, the key is completely independent of the biometric data, which means that, first, the use of the biometric is not forfeited if the key is ever compromised, and, second, the key can be easily modified or updated at a later date.

Conclusion

We have outlined some sample applications where biometric technologies can successfully be applied to DRM applications. However, extreme care has to be taken with respect to the customers' rights, in particular when dealing with sensitive personal data and specifically regarding biometric data. Identification patterns cannot be transmitted, saved, or watermarked into digital contents without extreme security measures (secured communication channels, encryption, data hiding, etc.) and according to applicable personal data regulatory laws. Consumer-rights organizations are sensitive to biometric data management, and all preventive measures should be addressed to hold back serious planning errors in the first systems to be deployed, which could provoke greater damages to the DRM industry in the future.

DRM systems are ready to be extensively used, as several proprietary solutions already exist, and open standards are ready, like the MPEG-21 IPMP and the full standard [7]. In this deployment process, biometrics should play a crucial role to guarantee the necessary mutual trust in the chain from right holders to content distributors and consumers. We believe that a proper use of biometric technologies will constitute one of the key issues for DRM technologies to succeed in the near future.

Acknowledgment

J. Ortega-Garcia and J. Gonzalez-Rodriguez wish to acknowledge the partial support of the Spanish Ministry of Science and Technology under projects TIC2000-1669-C04-01 and TIC2000-1683-C03-02, respectively. D. Reynolds was sponsored by the U.S. Department of Defense under Air Force contract F19628-00-C-0002.

Javier Ortega-Garcia received the Ph.D. degree "cum laude" in electrical engineering (Doctor Ingeniero de Telecomunicación) in 1996 from Universidad Politécnica de Madrid, Spain. He has been an associate professor since 1999 at the Audio-Visual and Communications Engineering Department, Universidad Politécnica de Madrid, where he is director of the Biometrics Research Laboratory. His research interests are focused on biometric signal processing: speaker recognition, fingerprint recognition, online signature verification, data fusion, and multimodality in biometrics. He has published over 60 international contributions. He is general chair for Odyssey'04—The ISCA Speaker and Language Recognition Workshop to be held in Toledo, Spain, in June 2004.

Josef Bigun obtained his Ph.D. degree from Linkoeping University in 1988. In 1988, he joined the Swiss Federal Institute of Technology in Lausanne (EPFL) where he worked until 1998 as adjoint scientifique. He was elected professor to the signal analysis chair, his current position, at Halmstad University and Chalmers Institute of Technology in 1998. He is a Fellow of the IEEE and IAPR. He has contributed significantly to the formation of biometric research in Europe, including the M2VTS consortium. He cochaired the First International Conference on Audio and Video-Based Biometric Person Authentication (AVBPA) in 1997. He has been an editorial member of several journals in pattern recognition. His interests include biometrics, texture analysis, motion analysis, and understanding of the biological processing of audio-visual signals.

Douglas Reynolds received the B.E.E. degree and Ph.D. degree in electrical engineering, both from the Georgia Institute of Technology. He joined the Information Systems Technology Group at the Massachusetts Institute of Technology Lincoln Laboratory in 1992. Currently, he is a senior member of the Technical Staff and his research areas cover robust speaker identification and verification, language recognition, speech recognition, and speech-content-based information retrieval. He has over 50 publications in the area of speech processing and three patents related to secure voice authentication. He is a Senior Member of the IEEE Signal Processing Society and has served on the Speech Technical Committee.

Joaquin Gonzalez-Rodriguez received the Ph.D. degree "cum laude" in electrical engineering (Doctor Ingeniero de Telecomunicación) in 1999 from Universidad Politécnica de Madrid, Spain. He has been an associate professor since 2002 at the Audio-Visual and Communications Engineering Department, Universidad Politécnica de Madrid, where he is head of the Speech and Signal Processing Group (ATVS). His research interests are focused on signal processing, biometrics and forensics: speaker recognition, forensic biometrics, data fusion in biometrics, robustness in speech/speaker recognition, and speech enhancement (auditory criteria, microphone arrays, binaural modeling). He has published over 50 international contributions. He is an invited member of ENFSI (European Network of Forensic Science Institutes) and will be acting as vice-chairman for Odyssey'04—The ISCA Speaker and Language Recognition Workshop, to be held in Toledo, Spain, in June 2004.

References

[1] M. Bazen and S.H. Gerez, "Systematic methods for the computation of the directional fields and singular points of fingerprints," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, no. 7, pp. 905–919, July 2002.

[2] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. fisherfaces: Using class specific linear projection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 7, pp. 711–720, July 1997.

[3] E.S. Bigun, J. Bigun, B. Duc, and S. Fischer, "Expert conciliation for multi modal person authentication systems by Bayesian statistics," in *Proc. Intl. Conf. Audio and Video Based Person Authentication, AVBPA'97*, LNCS-1206, Springer, 1997, pp. 291–300.

[4] J. Bigun and G.H. Granlund, "Optimal orientation detection of linear symmetry," *First Int. Conf. Computer Vision, ICCV-87*, Washington, DC, June 8–11, 1987, pp. 433–438.

[5] J.A. Bloom, I.J. Cox, T. Kalker, J.-P. M.G. Linnartz, M.L. Miller, and C.B.S. Traw, "Copy protection for DVD video," *Proc. IEEE*, vol. 87, no. 7, pp. 1267–1276, July 1999.

[6] A. Bodo, "Method for producing a digital signature with aid of a biometric feature," German patent DE 42 43 908 A1, 1994.

[7] J. Bormans, J. Gelissen, and A. Perkis, "MPEG-21: The 21st century multimedia framework," *IEEE Signal Proc. Mag.*, vol. 20 no. 2, pp. 53–62, Mar. 2003

[8] R. Brunelli and T. Poggio, "Face recognition: features versus templates," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1042–1043, Oct. 1993.

[9] J. Campbell and J. Schroeder, Eds., Special issue on speaker recognition, *Digital Signal Proces.*, vol. 10, Jan. 2000 [Online]. Available: http://www.sciencedirect.com/science/journals

[10] J. Daugman, "Recognizing persons by their iris patterns," in *Biometrics—Personal Identification in Networkwed Society*, K. Jain, R. Bolle, and S. Pankanti, Eds. Norwell, MA: Kluwer, 1999, pp. 103–121.

[11] B. Duc, S. Fischer, and J. Bigun, "Face authentication with Gabor information on deformable graphs," *IEEE Trans. Image Proc.*, vol. 8, no. 4, pp. 504–516, 1999.

[12] H. Federrath, "Scientific evaluation of DRM systems," [Online]. Available: http://www.inf.tu-dresden.de/~hf2/

[13] B. Gutschoven and P. Verlinde, "Multi-modal identity verification using support vector machines (SVM)," in *Proc. 3rd Int. Conf. Information Fusion*, July 2000, vol. 2, pp. 3–8.

[14] F. Hartung and F. Ramme, "Digital rights management and watermarking of multimedia content for m-commerce applications," *IEEE Commun. Mag.*, vol. 38, pp. 78–84, Nov. 2000.

[15] A.K. Jain, R. Bolle, and S. Pankanti (Eds.), *Biometrics—Personal Identification in Networkwed Society*. Norwell, MA: Kluwer, 1999.

[16] A.K. Jain and S. Pankanti, "Automated fingerprint identification and

imaging systems," in *Advances in Fingerprint Technology*, 2nd ed. New York: Elsevier Science, 2001.

[17] M. Kass and A. Witkin, "Analyzing oriented patterns," *Computer Vision, Graphics, Image Proc.*, vol. 37, 1987, pp. 362–385.

[18] J. Kittler, Y.P. Li, J. Matas, and M.U. Ramos-Sanchez, "Combining evidence in multimodal personal identity recognition systems," in *Proc. 1st Int. Conf. Audio and Video-Based Person Authentication, AVBPA'97,* vol. LNCS-1206, pp. 327–334, 1997.

[19] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. Comput.*, vol. 42, pp. 300–311, Mar. 1993.

[20] F. Leclerc and R. Plamondon, "Automatic signature verification: The state of the art, 1989–1993," *Intl. J. Pattern Rec. Machine Intell.*, vol. 8, no. 3, pp. 643–660, 1994.

[21] M. Maes, T. Kalker, J.-P.M.G. Linnartz, J. Talstra, G.F.G. Depovere, and J. Haitsma, "Digital watermarking for DVD video copy protection," *IEEE Signal Proc. Mag.,* vol. 17, pp. 47–57, Sept. 2000, .

[22] D. Maio and D. Maltoni, "Direct gray-scale minutiae detection in fingerprints," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 1, pp. 27–40, 1997.

[23] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*. New York: Springer-Verlag, 2003.

[24] S. Manjunath, R. Chellappa, and C.v.d. Malsburg, "A feature based approach to face recognition," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, June 1992, pp. 373–378.

[25] T. Mansfield, G. Kelly, D. Chandler, and J. Kane, "Biometric product testing final report," *Technical Report, CESG contract X92A/4009309.* Available : http://www.cesg.gov.uk/site/ast/biometrics/media/BiometricTestReportpt1.pdf

[26] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki, "The DET curve in assessment of decision task performance," in *Proc. ESCA 5th Eur. Conf. Speech Comm. and Tech., EuroSpeech '97*, Rhodes, Greece, 1997, pp. 1895–1898.

[27] F. Monrose, M.K. Reiter, Q. Li, and S. Wetzel, "Using voice to generate cryptographic keys," in *Proc. Odyssey 2001*, 2001, pp. 237–242.

[28] K. Nilsson and J. Bigun, "Prominent symmetry points as landmarks in fingerprint images for alignment," in *Proc. ICPR-16, Int. Conf. Pattern Recognition*, Quebec, Canada, Aug. 11–15, 2002, vol. III, pp. 395–398.

[29] NIST, National Institute of Standards and Technology [Online]. Available: http://www.nist.gov/speech/tests/spk/index.htm

[30] NIST, "Summary of NIST Standards for Biometric Accuracy, Tamper Resistance, and Interoperability," Report to U.S. Congress, Nov. 13, 2003. [Online]. Available: http://www.itl.nist.gov/iad/894.03/NISTAPP_Nov02.pdf

[31] J. Ortega-Garcia, J. Gonzalez-Rodriguez, D. Simon-Zorita, and S. Cruz-Llanas, "From biometrics technology to applications regarding face, voice, signature and fingerprint recognition systems," in *Biometric Solutions for Authentication in an E-World*, D.D. Zhang, Ed. Norwell, MA: Kluwer, July 2002, pp. 289–337, ch. 12.

[32] J. Ortega-Garcia, J. Fierrez-Aguilar, J. Martin-Rello, and J. Gonzalez-Rodriguez, "Complete signal modeling and score normalization for function-based dynamic signature verification," in *Proc. 4th Int. Conf. Audio- and Video-Based Person Authentication, AVBPA 2003*, LNCS 2688, June 2003, pp. 658–667.

[33] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in *Proc. 1994 IEEE Computer Soc. Conf. Computer Vision and Pattern Recognition*, Seattle, WA, June 1994, pp. 84–90.

[34] R. Plamondon and S.N. Srihari, "On-line and off-line handwritting recognition: A comprehensive survey," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 1, pp. 63–84, Jan. 2000.

[35] D.A. Reynolds, T.F. Quatieri, and R.B. Dunn, "Speaker verification using adapted gaussian mixture models," *Dig. Signal Proc.*, vol. 10, pp. 181–202, Jan. 2000.

[36] D.A. Reynolds et al., "The SuperSID project: Exploiting high-level information for high-accuracy speaker recognition," in *Proc. IEEE Int. Conf. Acous. Speech Signal Processing*, *ICASSP 2003*, Hong Kong, 6–10 April 2003, vol. 4, pp. 784–787.

[37] A. Ross, A.K. Jain, and J.Z. Qian, "Information fusion in biometrics," in *Proc. 3rd Audio and Video-Based Person Authentication, AVBPA'01*, 2001, LNCS-2091, pp. 354–359.

[38] D. Scheuermann, S. Schwiderski-Grosche, and B. Struif, "Usability of biometrics in relation to electronic signatures," EU Study 502533/8, GMD Report 118, July 2000 [Online]. Available: http://www.sit.fraunhofer.de/english/SICA/sica_projects/project_pdfs/eubiosig.pdf

[39] W. Shen and R. Khanna (Eds.), "Special issue on automated biometrics," *Proc. IEEE*, vol. 85, no. 9, Sept. 1997.

[40] D. Simon-Zorita, J. Ortega-Garcia, S. Cruz-Llanas, J.-L. Sanchez-Bote, and J. Gonzalez-Rodriguez, "An improved image enhancement scheme for fingerprint minutiae extraction in biometric identification," in *Proc. 3rd Audio and Video-Based Person Authentication*, *AVBPA'01,* LNCS-2091, Halmstad, Sweden, 6–8 June 2001.

[41] F. Smeraldi and J. Bigun, "Retinal vision applied to facial features detection and face authentication," *Pattern Recognit. Lett.*, vol. 23, pp. 463–475, 2002.

[42] F. Smeraldi, J. Bigun, and W. Gerstner, "On the role of dimensionality in face recognition," in *Proc. 2002 Workshop on SVM's*, S.W. Lee and A. Verri, Eds., LNCS-2388, 2002, pp. 249–259.

[43] C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and B.V.K. Vijaya Kumar, "Biometric encryption" [Online]. Available: http://www.bioscrypt.com

[44] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, pp. 71–86, 1991.

[45] P. Verlinde, G. Chollet, and M. Acheroy, "Multi-modal identity verification using expert fusion," *Inf. Fusion*, no. 1, pp. 17–33, 2000.

[46] C. Vielhauer, R. Steinmetz, and A. Mayerhofer, "Biometric hash based on statistical features of online signatures," in *Proc. 16th Conf. Pattern Recognition*, vol. 1, Aug. 2002, pp. 123–126.

[47] A.O. Waller, G. Jones, T. Whitley, J. Edwards, D. Kaleshi, A. Munro, B. MacFarlane, and A. Wood, "Securing the delivery of digital content over the Internet," *Electron. Comm. Eng. J.*, vol. 14, no. 5, pp. 239–248, Oct. 2002.

[48] R.P. Wildes, "Iris recognition: An emerging biometrics technology," *Proc. IEEE* (Special Issue on Automated Biometrics), vol. 85, no. 9, pp. 1348–1363, 1997.

[49] L. Wiskott, "Labeled graphs and dynamic link matching for face recognition and scene analysis," *Reihe Physik,* 53, 1995.

[50] L. Wiskott, J.-M. Fellous, N. Kruger, and C.v.d. Malsburg, "Face recognition and gender determination," in *Proc. Int. Workshop Automatic Face- and Gesture Recognition*, M. Bichsel, Ed. Zurich, Switzerland: MultiMedia Lab., Dept. Comput. Sci., Univ. Zurich, June 1995, pp. 92–97.

[51] H. Yu, D. Kundur, and C.-Y. Lin, "Spies, thieves, and lies: The battle for multimedia in the digital era," *IEEE Multimedia*, vol. 8, no. 3, pp. 8–12, Jul.-Sep. 2001.

[52] D.D. Zhang, Ed., *Biometrics Solutions for Authentication in an E-World.* Norwell, MA: Kluwer, July 2002.

[53] T. Zhang and C.-C. Jay Kuo, "Audio content analysis for online audiovisual data segmentation and classification," *IEEE Trans. Speech and Audio Processing*, vol. 9, no. 4, pp. 441–457, May 2001.