Biometrics for Automatic Identity Verification

Fabian Monrose
Courant Institute of Mathematical Sciences
New York University

Today there is renewed interest in reliable, rapid, and unintrusive means for automatically recognizing the identity of individuals. The key technologies behind some of the current cost-effective, reliable, and highly accurate identity based systems rely on biometrics. New developments in micro-electronics and advances in pattern matching, neural networks, and dynamic programming are enabling biometric technologies to perform more accurately at continually lower costs. In recent years, biometric technology has attained a very high degree of sophistication, and accuracy has been achieved at levels which surpass many other forms of identification. In this paper, we survey the theoretical foundations and design methodologies behind some practical identification technologies being deployed today.

Categories and Subject Descriptors: A.1 [**Surveys**]: Miscellaneous

General Terms: Verification, Experimentation, Security

Additional Key Words and Phrases: Biometrics, Pattern Recognition, Identity

1. INTRODUCTION

The increasing use of automated information systems together with our pervasive use of computers has greatly simplified our lives, while making us overwhelmingly dependent on computers and digital networks. Technological achievements over the past decade have resulted in improved network services, particularly in the areas of performance, reliability, and availability, and have significantly reduced operating costs due to the more efficient utilization of these advancements. However, the overwhelming interest in global accessibility brought about by these advances in technology have unveiled new threats to computer system security. As we press into the twenty-first century, new challenges abound. Advanced safeguards against fraud and impersonation, as well as more foolproof measures against unauthorized access to computer resources and data are now being sought. This paper surveys some practical techniques that have been applied as countermeasures to computer fraud, impersonation and unauthorized computer access. The paper is organized as follows:

Section 2 presents an introduction to biometrics and examines why biometrics are good candidates for identity verification and how they can be used as countermeasures to some of the arising computer security threats. The rest of the paper is devoted to verification methodologies that can be classified under two broad categories, namely (1) face recognition and (2) hand recognition. We do not consider biometric techniques that can be classified under voice recognition because although

Address: 251 Mercer Street, New York, NY, 10012 Tel: (212) 998-3061; Fax: (212) 995-4124

the feature set (typically a spectrogram) is quite different from that used in face and hand recognition, the identification and classification techniques are very similar to those presented herein, in particular, to an emerging non-static biometric (i.e., keystroke dynamics) .

Section 3 surveys the theoretical work done in the area of face recognition that have led to the successful implementation of access control systems. An examination of recognition techniques that involve parts of the hand which are being used to restrict access to computers and digital networks is presented in Section 4. We conclude with a summary and some closing remarks.

2. BIOMETRICS

Available countermeasures to the problem of identity verification can be categorized into three main groups: those that rely on (a) something a person knows (e.g a password), (b) something a person possesses (e.g an ID card), or (c) something about a person such as characteristics. Depending on the level of security required, these three basic methods can be combined to give greater protection.

Security based upon possession or knowledge may be compromised without discovery, since the information or article may be extorted from its rightful owner. Increasingly, attention is shifting to positive identification by biometric techniques that encompass the third class of identification (i.e., biometrics) as a solution for more foolproof methods of identification. For the foreseeable future, these biometric solutions will not eliminate the need for I.D. cards, passwords and PINs. Rather, the use of biometric technologies will provide a significantly higher level of identification and accountability than passwords and cards alone, especially in situations where security is paramount.

Biometric technologies are defined as "automated methods of verifying or recognizing the identity of a living person based on a physiological or behavioral characteristic" [Miller 1994]. Biometrics are excellent candidates for identity verification because unlike keys or passwords, biometrics cannot be lost, stolen, or overheard, and in the absence of physical damage they offer a potentially foolproof way of determining someone's identity.

Indispensable to all biometric systems is that they recognize a *living* person and encompass both physiological and behavioral characteristics. Physiological characteristics such as fingerprints are relatively stable physical features that are unalterable without causing trauma to the individual. Behavioral traits, on the other hand, have some physiological basis, but also reflect a person's psychological makeup. The degree of intra-personal variation is smaller for a physical characteristic than in a behavioral one. Barring injury, a person's fingerprint is immutable, whereas a signature is influenced by both controllable actions and unintentional psychological factors. Because of this difference, no one biometric can serve all needs. Consequently, the choice of biometric remains application specific.

Modern biometric schemes generally rely on aspects of the body and its behavior and are more adaptive to change in individuals than previous systems. While fraud is still possible the counter-measures used are usually simpler than would be expected, while difficult to fool. For instance, iris scanning technologies check for changes in the pupil because of its instinctive reaction to light, and fingerprint systems check for changes in color caused by blood flow as the finger is pressed

harder against the glass of the scanning device.

The biometric techniques presented herein, for the most part, rely on a fundamental concept — pattern recognition. Recognition is regarded as a basic attribute of human beings. Our ability to recognize people, even after being separated for some time, is remarkable. The goal is to achieve such a level of recognition capability with computers. A brief overview of the fundamental concepts related to pattern recognition is given in the following section.

2.1 Pattern Recognition: Extraction, Representation, and Classification

The design of an automatic pattern recognition system involves *representation, extraction*, and *classification*. *Representation* of input data measures characteristics of the pattern or object to be recognized. When the measurements obtained yield information in the form of real numbers, it is often useful to think of a pattern vector as a point in an n-dimensional Euclidean space.

The *extraction* of characteristic features from the input data and the reduction of the dimensionality of the resulting pattern vectors is often referred to as the preprocessing and feature extraction problem. For example, we may choose to use only a selected number of measurements from the input, either because these features are enough to identify the individual (like the eyes and mouth) or because the addition of those extra features increases the computational complexity of the problem or yields no real benefit. The number of degrees of freedom of variation in the chosen index across the human population, their immutability over time and immunity to intervention, and the computational prospects for efficiently encoding and reliably recognizing the identifying pattern must all be assessed during feature extraction.

Classification and *identification* involves the determination of optimum decision procedures. After the observed data from patterns to be recognized have been expressed in the form of measurement vectors in the pattern space, we want to decide to which pattern class these data belong [Duda 1973]. For example, given a number of face images, and an "unknown" reference template from a database of available faces, we want to be able to positively identify the "unknown" individual with a specified level of certainty.

In this paper emphasis is placed on the problem of determining what optimum decision procedures are needed in the identification and classification processes, and we study the fundamental ideas behind several deterministic and statistical algorithms that have been proposed. Readers interested in pattern selection and extraction are referred to [Tou and Gonzalez 1981; Duda 1973; Fu 1971; Fukanaga 1972].

The analysis of personal features has a natural range of variation; a biometric method never provides an absolutely certain identification. As such, a biometric identification systems can fail in one of two ways; either an authorized user is rejected or an illegitimate user can be incorrectly granted access to the system. Biometric systems must allow adjustments to control the error probabilities to some degree. An examination of systems based on biometric techniques which have been deployed to restrict access to computers and data networks is now presented.

3. FACE RECOGNITION

In the late 1960's many attempts to build face recognition systems resulted in limited success. However, during the early 80's new technologies rekindled interest in pattern recognition strategies for personal identification systems. The goal of the following sections is to present a general overview of the processes involved, and the techniques used, in the area of *automatic* recognition and analysis of human faces. A comprehensive survey which examines most of the latest techniques can be found in [Chellappa et al. 1995]. As outlined in section 2.1 recognizing a face in an image involves three critical steps. We discuss each of these steps in the following paragraphs.

In contemporary systems, the approach to data representation is based on either a 2D array of intensity values or a feature vector. While 2D arrays of intensity values is the simpler approach, it is not a very compact representation. Therefore, compaction and compression of the 2D arrays is typically performed. Features derived from intensity images include the size of the eye, distance between the eye, hair intensity, etc. Characteristic points such as the notch between the nose and brow, tip of the chin, tip of the nose, are some examples of features represented by a feature vector.

The first step in the identification of a face in an image is to locate the face if it is present. Several techniques have been proposed for the identification of facial images [Yulie et al. 1989; Craw et al. 1987; Perry and Carney 1990]. In some situations, the location of a face is either known *a priori* or can be easily derived from the constraints under which the image was captured. For example, the controlled conditions under which mug-shots are taken makes it relatively easy to locate the face in the image. However, if the position of the face is not known *a priori* then it must be located in the image. One approach in locating the face is by first looking for some important features, such as eyes, and then estimating the overall location of the face based on the geometric information obtained, i.e. the relative locations of the features.

Several factors make the problem of face detection difficult when its location (or even its presence) is not known *a priori*. The presence of facial hair, eye glasses and other objects that may obscure the face, make the problem quite complex. Variation in scale and orientation also pose additional obstacles. The presence of noise in the image, and occlusion (i.e. partial obstruction of the image) compound the problem even further.

Many approaches have been proposed to address these issues, most notably the work of Yulie, Cohen, and Halinan [1989], that uses deformable templates to detect the face. Craw, Ellis, and Lishman [1987] suggests a mechanism for detecting the outline of the face using a template matching scheme. Some neural network approaches [Perry and Carney 1990] have also been undertaken in recent years. In depth discussions of the problems associated with face detection can be found in [Yulie et al. 1989; Craw et al. 1987; Perry and Carney 1990].

Once the process of recognizing a face in an image begins, the tasks is either to associate a name with the image presented, or to reject the image as unknown. In addition, multiple instances, possibly with different sizes and orientation of the same face must be identified as the same person. We examine several approaches

to face recognition in the following sections.

3.1 Template Matching

Template matching is one of the oldest approaches to object recognition. Fundamentally, the approach classifies an image as the nearest neighbor in a learning set, where distances are measured using a mean square error of gray values. If all exemplar images are normalized to the zero mean and unit variance, then template matching is equivalent to choosing the image in the learning set that best correlates to the test image.

Computing the correlation function between two images amounts to comparing them against one another in every possible position, shifting both horizontally and vertically. This can be done efficiently using Fourier transforms. After image normalization, each person is represented in a database whose entries are the digitized image of the frontal view of the user and a set of masks representing the eyes, nose, and mouth. The unclassified image is compared, in turn, with all the database images, returning a vector of matching scores computed as the maximum of the normalized cross correlation. The unknown person is then classified as the one giving the highest score.

Template matching systems can be trained through the following method. When a new view of the same face is presented, if it is below a threshold correlation to other views of the same face, the entire digitized image is added as a new profile for that user. Otherwise, if the correlation is above a certain threshold, the views are combined by taking a weighted average of the digitized pixels of the images.

Unfortunately, while correlation appears to be one of the most widely used techniques in face recognition, it has several well know disadvantages: if the images in the learning and test sets are gathered under varying lighting conditions, then correlation will perform poorly as the two sets will not be tightly clustered, since we must correlate the test image with each image in the learning set, correlation is computationally expensive and a large amount of space is required because the learning set often contains numerous images of each person.

The practicality of template matching schemes depends heavily on their robustness against deformation and noise. Research suggests that preprocessing based on gradient magnitude[1] yields one of the best performance ratios and lowest illumination dependence [Poggio and Edelman 1990]. However, correlation performs poorly when subject to affine transformations, especially rotations greater that 10%. More sophisticated template matching methods have since been developed, including elastic template matching [Buhmann et al. 1989] and the use of deformable templates (where a match is obtained by deforming one surface so that its shape becomes close to that of the other surface) [Yuille 1991; Moghaddam et al. 1996] and auto-association and back-propagation neural networks [Cottrell and Fleming 1990; Bourlard and Kamp 1988]. A recent performance comparison of these techniques can be found in [Zhang et al. 1997].

In 1991 Turk and Pentland [1991a] proposed a novel scheme for recognition based on an information theoretic approach. Their approach transforms face images into a small set of characteristic feature images, called "eigenfaces" that are the principal

[1] The intensity of the gradient computed with an L_1 norm on a Gaussian image

components of the initial training set of face images [Turk and Pentland 1991a]. The novelty of this approach stems from the use of eigenvectors for recognition, and the central idea involves finding the principal components of the distribution of faces, or in other words, the eigenvectors of the covariances matrix of the set of face images. In essence, this approach attempts to extract the relevant information in a face image, encode it as efficiently as possible, and compare one face encoding with a database of models encoded similarly. The solution to extracting the information contained in an image of a face is to somehow capture the variation in a collection of face images, independent of any judgment of features, and to use this information to encode and compare individual faces [Turk and Pentland 1991b].

This solution is achieved by using a method of reduction used is known as principal component analysis (PCA) or more formally as Karhunen-Loeve expansion. The idea is to find subspaces of face images (vectors) that best account for the variation. These vectors define the "face space". PCA uses a small set of M orthonormal vectors that best describe the distribution of the data in the face space.

Given a new face image it is projected on the "face space". The resulting vector represents the contribution (i.e. weight) of each eigenface in representing the input face image, using the M eigenfaces as a basis set for face images. Identification is announced by determining which face class minimizes the Euclidean distances between the vectors.

A drawback of using PCA for dimensionality reduction is that the scatter being maximized is not only due to the between-class scatter but also the within-class scatter that, for classification purposes, is unwanted information [Belhumeur et al. 1995]. Furthermore, since performance increases with the dimension of the Eigenspace, the Eigenface method performs no better than correlation, which is disappointing.

In recent years, some alternate approaches have been proposed to address the limitations of the eigenface model. In contrast to the eigenface approach, these approaches aim to masks out the regions of the face that are highly variable. The objective is to find a linear projection of the faces from the high-dimensional image space to a significantly lower dimensional feature space that is insensitive both to variations in lighting direction and facial expression.

These improvements are made possible by the following observation: in the absence of self shadowing the albedo and surface normal can be recovered from a Lambertian surface from the same viewpoint taken under three known, linearly independent, light source directions [Belhumeur et al. 1995]. In fact, one can reconstruct the image of the surface under arbitrary lighting directions by a linear combination of the three original images. Considering this observation more thoroughly one realizes that it suggests that for a fixed viewpoint, all images of a Lambertian surface lie in a 3-D linear subspace embedded in the high-dimensional image space. Therefore, if three or more images taken under different lighting directions for each face are used then it is possible to construct a 3-D basis for the linear subspace. To perform recognition we simply compute the distance of a new image to each linear subspace and choose the face corresponding to the shortest distance.

This approach is, however, too computationally expensive and has too high a storage requirement. Fortunately, it is possible to improve this model by per-

forming dimensionality reduction using a linear projection, in a manner that still preserves linear separability. It is interesting that error free classification under varying lighting conditions is still possible in this lower dimensional feature space using linear decision boundaries. A variation of this scheme that does not require multiple templates for every individual is given in [Moghaddam et al. 1996].

3.2 Geometric Feature Extraction

During the mid 1970's researchers interested in face recognition realized that a face can be recognized even when the details of the individual features, such as eyes, nose and mouth, are no longer resolved. This implied that the overall geometrical configuration of the face features is sufficient for discrimination over a salient population, and led to what is now being referred to as recognition by geometric feature extraction. This approach to face recognition is based on a simple observation: the values of features of a face do not change very much in different images. Furthermore, if a feature is very different in two faces, the corresponding instances in images will be very different (if the effects of lighting and shadow are minimal).

Given a set of fiducial marks (chin tip, nose, forehead, bridge, nose bottom, throat, upper lip, mouth and lower lip) a set of features can be derived from the distances between the fiducials, angles between them, and areas of some triangles formed by the fiducials [Samal and Iyengar 1992]. It is important that for these geometric features, estimation must be as easy as possible, the dependency on light conditions must be as small as possible, dependency on small changes of facial expression must be minimal, and the information content must be as high as possible. Some example fiducials are shown in figure 1.

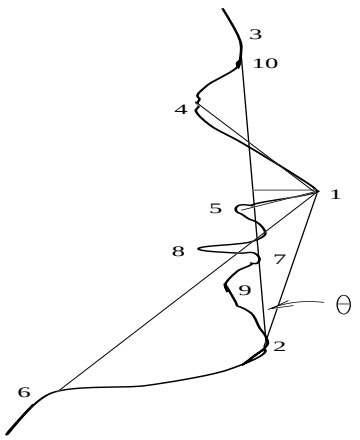

Fig. 1. Characteristic points used as features in the recognition of a facial image. The features consisted of nose & chin protrusion, area of the profile, the distances $D_{1,2}, D_{1,4}, D_{1,5}, D_{1,6}, D_{1,8}$, and $D_{7,9}$. The angles $\theta(1,2,3), \theta(1,3,4), \theta(4,1,5)$ and the ratio $D_{1,4}/D_{1,2}$ were also used.

For feature extraction bilateral symmetry and the fact that almost every face has two eyes, a nose and a mouth, can be used to detect facial features. Given the vertical and horizontal integral projections of an image $I(x,y)$ in the $[x_1,x_2]\times[y_1,y_2]$

rectangle, an estimate of the mouth and nose, for example, can be obtained by looking for peaks of the horizontal projection of the vertical gradient for the nose, and for valleys of the horizontal projection of the intensity for the mouth. Peaks and valleys can be rated using their prominence and distance from the expected location. Those with the highest ratings represent the vertical position of the nose and mouth. Features such as eyebrows can be located via similar analysis by looking for pairs of peaks of gradient intensity with opposite direction.

Once features have been identified and successfully extracted, the process of recognition begins. Some of the best performance results using geometric features for recognition have been achieved using Bayesian-like classifiers for recognition [Poggio and Edelman 1990]. These approaches aim to characterize the performance of the feature-based technique as a function of the number of classes to be discriminated. As is the case with most classifiers it is assumed that feature vectors are distributed according to a Gaussian distribution. An unknown vector is associated with the nearest neighbor in the database, i.e., the person who **maximizes** the probability of the measurement vector according to the modeled distribution. Therefore, if x_i is the feature vector, σ_i the interclass dispersion vector and w_i the weight vector, then the distance of two feature vectors x_i and $x_i^{'}$ is expressed as:

$$\Delta^{\alpha}(x, x^{'}) = \sum_{i=1}^{n} w_i \Big(\frac{|x_i - x_i^{'}|}{\sigma_i} \Big)^{\alpha} \tag{1}$$

The value of α may be adjusted to provide robust statistic properties [Huber 1981]; recognition levels may be significantly improved by using a value of α close to one (rather than the value 2 as justified by the Gaussian assumption). Two other alternatives for personal verification based on characteristics of the face include the use of facial thermograms and the use of spatial patterns in an individuals iris.

Prokoski, Riedel, and Coffin [1992] propose a recognition system based on analysis of facial thermograms. The system involves the use of "element shapes" for features that are derived from biosensor data (approximately 65,000 individual temperate measurements recorded across the face). "Element shapes" are contours in space, whose shapes are determined by the extent, intensity, and duration of output from a biosensor. At any instant these shapes are influenced by and are indicative of (1) the gross anatomy of the individual's head, (2) the ambient environment, (3) the vascular system of the person, (4) appliances (example glasses), and any immediate contact with other objects. It has been shown that thermal measurements of individuals under varying conditions are highly repeatable, and reliable enough to be used as a general identity verifier.

Automated iris recognition is yet another emerging biometric technology that has been suggested for non-invasive personal verification. It has been shown that the spatial patterns that are apparent in the human iris are unique to each individual. The uniqueness in anatomical structure is due to individual differences in the development of muscular and vascular structures [Daugman 1993]. A property the iris shares with fingerprints is the random morphogenesis of its minutiae; the phenotypic expression even of two irises with the same genetic genotype (e.g. twins) have uncorrelated minutiae. In these respects the uniqueness of every iris parallels the uniqueness of every fingerprint, common genotype or not.

The first iris recognition system dates back to the early nineteen hundreds, when visual inspection of the iris was used to distinguish inmates in a Parisian prison. More recent approaches include the work of Flom and Safir [1987], Daugman [1993] and Wildes and Asmuth [1994]. The work of Flom and Safir [1987] is theoretical and has of yet resulted in a practical implementation. These systems can be differentiated by the techniques used during the recognition process. Daugman's systems is based on a statistical decision theory approach, in which irises are efficiently recognized by executing a statistical test of independence on their codes. More precisely, the system examines whether the degree to which one iris code predicts another iris code, is compatible with the hypothesis that they arose from independent random processes.

By contrast, the approach of Wildes and Asmuth [1994] is based on spatial correlation with the match metric based on integrating pixel differences over spatial positions within each frequency band of the image. Correlation is performed over small 8×8 blocks of pixels in each spatial frequency band, and a goodness of fit is calculated for each band by summing the block correlations values via the median statistic. A solution to combining the goodness of fit for each band into a final judgment is achieved through the use of Fisher's linear discriminant [Duda 1973; Fisher 1936]. FLD is used to combine the values in a manner so that the variance within a class of iris images (i.e. various instances of the same iris) is minimized, while the variance between difference classes of iris images (i.e instances of different irises) is maximized.

4. PUTTING THE FINGER ON SECURITY

One of the most dominant biometrics used today is fingerprint identification. The process involves both classification and identification. There are two different levels of fingerprint techniques practiced, namely coarse and fine level classification. In coarse level classification, the main objective is to categorized fingerprints into one of six groups: (a) arch, (b) tented arch, (c) right loop, (d) left loop, (e) whorl and (f) twin loop. Coarse level classification greatly reduces the search time for matching fingerprints in a given database. For example, in attempting to find a match for a twin looped subject, arched looped candidates can be automatically pruned from the search.

Fine level classification involves extracting ridge endings and branching points (minutiae), and determining the similarity of fingerprints by comparing those minutiae points. It is more closely tied to the general pattern classification problems we have been examining thus far. We describe both fingerprint classification and identification in the following sections.

4.1 Classification Strategies

Today, there are two prominent techniques for fingerprint classification: (1) classification based on the number and location of extracted singular points (*cores* and *deltas*) and (2) classification based on the shape of flow lines.

However, before classification can begin, the digitized fingerprint is first transformed into a directional image pattern. The direction of each pixel in the image is computed by superimposing a 9x9 mask that computes pixel values in eight directions. Since the resulting process yields a direction at each pixel quantized to eight

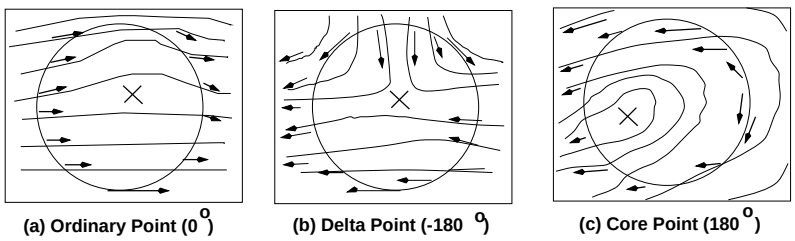

Fig. 2. Examples of ordinary, delta, and core points.

values, smoothing and averaging in the local neighborhood is performed. Once a stable directional image pattern is obtained, the process of locating *core* and *delta* points begins.

While a number of different approaches for finding *core* and *delta* points exist in the literature, the work of Karu and Jain [1996] exemplifies the most popular technique. A point is defined as *ordinary* if the angle turned, when making a full counter clockwise turn around the curve is 0^o, *core* if the angle is 180^o, and *delta* if it is -180^o. Figure 2 illustrates this further. Fingerprints are then classified based on the number of *delta* and *core* points in the fingerprint. Furthermore, left and right loops are distinguished by their relationship to *delta* points; when starting from a *core* point and moving along the directional vectors, the *delta* point remains to the left in left loops, and to the right in right loops. Figure 3 depicts the major fingerprint classes.

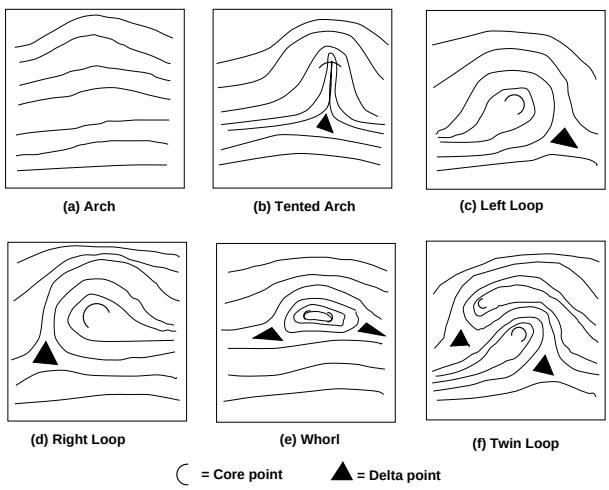

Fig. 3. The figure above depicts cores and deltas in fingerprints belonging to different classes.

Normalization is necessary for increasing the accuracy of matching in the presence of translation, rotation, and scale changes. Normalization can be carried out by finding distinguishing points in each image. *Core* and *delta* points are good

candidates to be used for normalization. The normalization techniques are similar to those carried out in most pattern recognition systems.

We now consider classification based on shape of flow lines. Examples of work based on this approach include Kawagoe and Tojo [1984], Moayer and Fu [1975], and Rao and Balck [1980]. In this method, the flow of the ridge lines are encoded, and the representative flow-line traces are extracted by analyzing encoded ridge shapes. Fingerprints are then classified according to the shape of these flow lines. Since this method is based on global information in the fingerprint, it is more robust against missing feature points and tends to perform better than classification based on the number and location of minutiae when image quality is low. The calculation of the directional image and minutiae is similar to what was previously. However, the classification methods differ to a large extent.

While many methods for classification based on flow line tracing can be found in the literature, we examine the work of Kawagoe and Tojo [1984] as we believe their scheme to be representative of the work done in this area. As shown in figure (4), the lengths $A1, B1, A2$ and $B2$ are obtained from flow line traces starting from the singular regions to the perpendicular bisector. Once these values have been calculated, the distinction between whorl and twin fingerprints is governed by the following rules (refer to figure 4):

(1) $0 > (A1/B2 - 1)(A2/B2 - 1) \Rightarrow Cases\ \ (a), (b), (c) \rightarrow Whorl$

(2) $(A1/B2 - 1)(A2/B2 - 1) > E \Rightarrow Cases\ \ (d), (e) \rightarrow Twin$

E represents some predefined threshold. In addition, a *twinness* parameter (see [Kawagoe and Tojo 1984]) is used to further distinguish twin and whorl types. Arch and loops are determined by the direction of the flow lines.

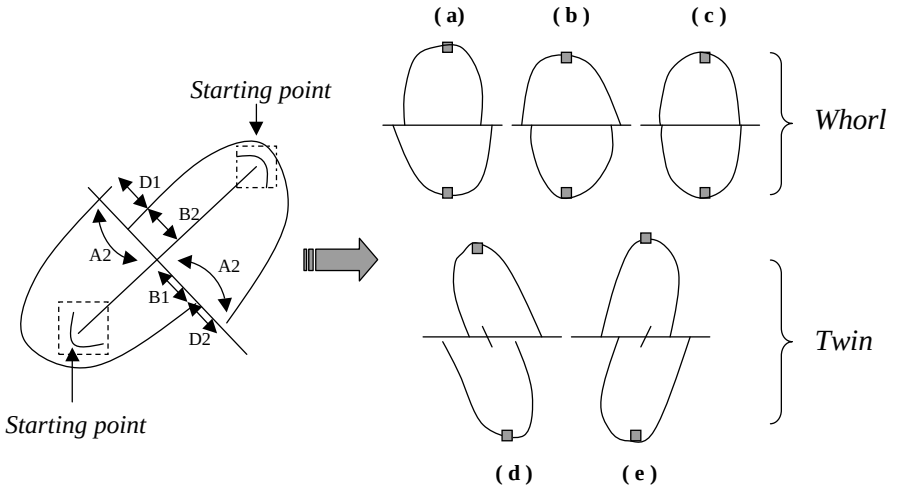

Fig. 4. Discrimination rules for whorl and twin flow lines. Also shown is the line segments that represent the lengths $A1, B1, A2$, and $B2$. By analyzing these patterns, fingerprint types, whorl and twin, can be distinguished.Arch and loop types are handled in a different manner.

4.2 Fingerprint Identification

After classifying the fingerprint type, the process of identifying the fingerprint begins. A seminal paper on fingerprint identification is the work of Isenor and Zaky [1986]. Identification is achieved by representing the fingerprint in the form of a graph, in which the nodes represent ridges and edges represent ridge adjacency information. Fingerprint matching is then reduced to the problem of matching the graphs for corresponding reference and test profiles. Since the topological relationship among the ridges is used as the means of encoding the fingerprint information, this scheme is highly tolerant of distortion, rotation, and translation of the fingerprint image.

For a given fingerprint, the corresponding graph is computed as follows. First the digitized fingerprint, in which valleys appear as light pixels and ridges as dark pixels, is subjected to a thinning algorithm, in order to reduce all ridges to a thickness of one pixel. Endings and joinings are identified by examining the state of adjacent pixels of each dark pixel. Pixels that belong to the same ridge are identified as such, and the number of pixels that constitute a ridge is taken as a measure of its length. However, since elastic distortion of the finger may cause significant changes in this parameter, it is only used as an approximate measure during the matching process [Isenor and Zaky 1986].

Once the graphs for a reference and test profile has been generated, determining whether two graphs represent the same fingerprint is accomplished by using graph matching algorithms. Graph isomorphism [Hopcroft and Wong 1974], an NP-complete problem, defines a match as being a perfect one-to-one mapping of one graph to another, whereas when matching fingerprints, each graph may have missing or extra nodes. Therefore, the matching process yields some measure of goodness of the match.

Heuristics used in practice to augment traditional isomorphism techniques are typically of the flavor of tree tracing algorithms. The general idea is similar to that of graph isomorphism, but scores for mappings are calculated based on the number and type of minutiae that have been matched for two trees grown in parallel, in the two graphs. The information recorded in each node is also used when computing scores. When all possible trees have been grown, the one producing the highest score is selected as the final tree. The resulting score is taken as a measure of the degree to which the two fingerprints match each other.

5. KEYSTROKE DYNAMICS

Another biometric technique being investigated for potential use as an identity verifier is keystroke dynamics. This method analyzes the way a user types at a terminal by monitoring the keyboard inputs to identify users based on certain habitual typing rhythm patterns.

It has been argued that use of keystroke rhythm is a natural choice for computer security. This argument stems from observations that similar neuro-physiological factors that make written signatures unique, are also exhibited in a user's typing pattern [Joyce and Gupta 1990]. When a person types, the latencies between successive keystrokes, keystroke durations, finger placement and applied pressure on the keys can be used to construct a unique signature (i.e., profile) for that individual.

For well-known, regularly typed strings, such signatures can be quite consistent. Furthermore, keystroke dynamics is not intrusive, making it quite applicable to computer access security.

Current keystroke verification techniques can be classified as either static or continuous. Static verification approaches analyze keystroke verification characteristics only at specific times, for example, during the login sequence. Although the static approach of Brown and Rogers [1993] and Monrose and Rubin [1997] provides more robust user verification than simple passwords, it does not provide continuous security. For example, it cannot detect a substitution of the user after the initial verification. Continuous verification, on the contrary, monitors the user's typing behavior throughout the course of the interaction.

Keystroke dynamics, while more of a behavioral trait than a physiological characteristic, is a good identity verifier. However, recognizing users based on habitual rhythms in their typing pattern uses dynamic (unlike fingerprint which uses static) "performance" features that depend upon an act; the rhythm is a function of the user and his environment. Thus, there exist inherent limitations which are attributed to the nature of the "signature" and its relationship to the user. The problem with keystroke recognition is that there are no known features or feature transformations which are *dedicated solely* to carrying discriminating information. However, empirical studies suggest that different individuals exhibit characteristics that are strikingly individualistic that can be used for identification purposes.

5.1 Current state of Keystroke Dynamics

The application of keystroke rhythms to computer access security is relatively new. Joyce and Gupta [1990] present a comprehensive literature review of work related to keystroke dynamics *prior* to 1990. We briefly summarize these efforts and extend our survey to cover work that has been undertaken since then.

In 1980, Gaines, Lisowski, Press, and Shapiro [1980] conducted experiments with a very small population of seven secretaries. A test of statistical independence of their profiles was carried out using the *T-Test* under the hypothesis that the means of the digraph times at both sessions were the same, and with the assumption that the two variances were equivalent. Similar experiments were conducted by Leggett, Williams, and Umphress [89] with seventeen programmers but for the continuous approach to user verification. The authors report an identity verifier that validates the results of Gaines, Lisowski, Press, and Shapiro [1980] — an identity verification system with false alarm rate of about 5.5 percent and imposter pass rate of approximately 5.0 percent.

While the approaches of [Gaines et al. 1980] and [Leggett and Williams 1988; Leggett et al. 89] address a number of problems of user verification, their methods leave considerable potential for refinement. For example, the pool variance estimate used by Leggett, Williams, and Umphress [89] is meaningful only when there is homogeneity of variance across all reference digraph latencies. However, studies by [Mahar et al. 1995] show that there is significant variability with which typist produce each digraph, and hence the use of a pooled estimate digraph latency variability is inappropriate.

An additional limitation of the Umpress and Williams digraph latency based technique is the use of a single low-pass temporal filter for all typists for the removal

of outliers. The rationale for this approach is that digraphs with abnormally long latencies are not likely to be representative of the authorized users' typing. While this seems like a reasonable assumption both Mahar, Napier, Wagner, Laverty, Henderson, and Hiron [1995] and [Monrose and Rubin 1997] have shown that one filter value for all typists does not yield optimal performance.

Furthermore, empirical data from Gentner [1993] suggests that the median inter-key latency of expert typists is approximately 96 ms, while that of novice typists is near 825 ms. Therefore, the 500 ms low-pass filter used by Leggett, Williams, and Umphress [89], excludes many keystrokes typical of a novice typists, while at the same time, includes many keystrokes which are not representative of an expert typist [Mahar et al. 1995]. The research suggest that the use of digraph-specific measures of variability instead of one low-pass filter can lead to measurable improvements in verification accuracy [Mahar et al. 1995; Monrose and Rubin 1997].

Moreover, while traditional approaches to keystroke verification use the key down-to-down time as the base unit of measure, this measure may be further delineated into two orthogonal components; total time the first key is depressed (i.e keystroke duration), and the time between first key release and second key depression (i.e keystroke latency). Brown and Rogers [1993], Mahar, Napier, Wagner, Laverty, Henderson, and Hiron [1995] and Monrose and Rubin [1997] have used these two components in their verification systems. However, the initial sample sets for both Brown et al. and Monrose et al. did not provide enough data to ascertain whether the use of the two separate orthogonal digraph components added significant predictive power to the more traditional key down-to-down measure. Substantially improved performance results based on using the bivariate measure of latency with an appropriate distance measure were achieved by [Mahar et al. 1995].

A promising research effort in applying keystroke dynamics as a static authentication method is the work of Joyce and Gupta [1990]. Their approach is relatively simple and provides very impressive results. It is based on using keystroke information obtained during the login process of a modified login sequence. The performance of their verifier is addressed in the following section.

Some neural network [Hussien et al. 1989; Brown and Rogers 1993; Alexandre 1996] approaches have been undertaken in the last few years. While the back-propagation models used yield favorable performance results on small databases, neural networks have a fundamental limitation in that each time a new user is introduced into the database, the network must be retrained. For applications such as access control, the training requirements are prohibitively expensive and time consuming. To overcome the problem of continuous retraining, most systems based on neural networks partition the database into smaller groups of subjects. However, in situations where there is a high turnover of users, the down time associated with retraining can be significant.

5.2 Classification Strategies

The problem of recognizing a given pattern as belonging to a particular person either after exhaustive search through a large database, or by simply comparing the pattern with a single authentication template can be formulated within the

framework of statistical decision theory. By this approach one can convert the problem of pattern recognition into a much more expedient task, which involves the execution of tests of statistical independence. The majority of approaches described in the following paragraphs adhere to this model.

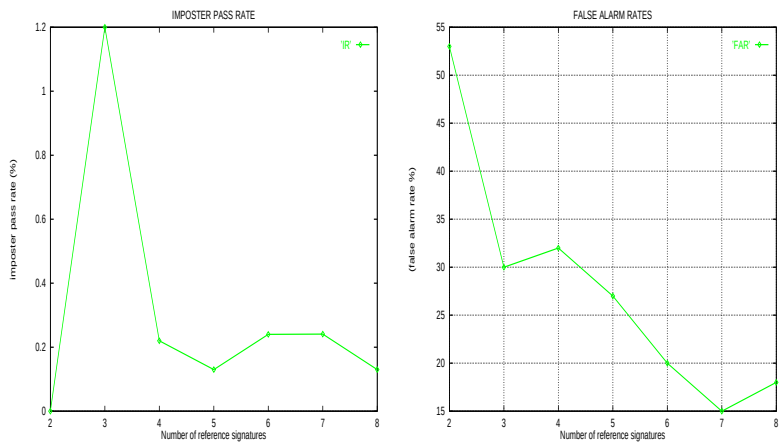

Fig. 5. Impact of the number of reference signatures on the false alarm and imposter pass rates for 33 users

The classification technique employed by Joyce and Gupta [1990] represents the mean reference signature for a given user as $\mathcal{M} = \{\mathcal{M}_{username}, \mathcal{M}_{password}, \mathcal{M}_{firstname}, \mathcal{M}_{lastname}\}$. Verification is performed by comparing the test signature $\mathcal{T}$ (acquired at login time) with $\mathcal{M}$ and determining the magnitude of difference between the two profiles. Given $\mathcal{M} = (m_1, m_2, ..., m_n)$ and $\mathcal{T} = (t_1, t_2, ..., t_n)$ where n is the total number of latencies in the signature, the verifier computes the magnitude of difference using an L_1 norm. Positive identification is declared when this difference is within a threshold variability of the reference signature. The mean and standard deviation of the norms $\|M - S_i\|$, where S_i is one of the eight training signatures, are used to decide the threshold for an acceptable difference vector between a given $\mathcal{T}$ and $\mathcal{M}$. The verifier is reported in figure (5).

The performance of a keystroke verification system by Mahar et al. can be found in [Mahar et al. 1995; Napier et al. 1995]. In their experiments users undertook a practice session in which stimuli consisted of four words, four sentences and a name, and were allowed as much practice as desired before commencing the experimental trials. The false alarm and false reject rates ranged from 24% to 2% with poorest performance under a univariate measure (i.e using only traditional down-to-down time) and a z-test statistic condition, while the best performance was achieve using a multivariate measure and a chi-square test as the decision function. Their results show that the decomposition of latency, in terms of its two orthogonal components, increases recognition performance.

While the absolute verification rates achieved here are encouraging, [Mahar et al. 1995; Napier et al. 1995] and [Joyce and Gupta 1990] test using a replacement methodology, which means that the distribution of the training set is necessarily

representative of the learning set. The use of separate data sets, recorded at different times, would be more reliable. More realistic results are reported by Monrose and Rubin [1997] based on studies where participants were allowed to participate in experiments conducted at varied times under no supervision. Assuming that each feature for a user is distributed according to a Normal distribution, a likelihood score between a reference profile R and unknown profile U is calculated based on the probability of observing a feature value in the reference profile R, given the mean and standard deviation for that feature in R. Scores are weighted an the "unknown" profile is then associated with the *nearest neighbor* in the database, i.e., to the person who maximizes the score of the feature vector.

The correct identification rate using a weighted probabilistic classifier was approximately 90%, which represents improvement with respect to the performance of a Euclidean distance and a non-weighted scoring approach (see [Monrose and Rubin 1997]). Additionally, their research argues in favor of the use of structured text instead of allowing users to type arbitrary text (i.e., text-independent or "free-text"). While recognition based on free-text may be more desirable, free-text recognition did not perform as well as recognition based on fixed-text. Recognition based on free-text may be expected to vary greatly under operational conditions in which the user may be absorbed in a task or involved in an emotionally charged situation. The fact that the input is unconstrained, that the user is may be uncooperative, and that environmental parameters are uncontrolled impose limitations on what can be achieved with free-text recognition.

While it is difficult to give a meaningful comparison of the approaches outlined above, as there is no unified data set under which the approaches can be compared, overall, the results suggest that it is possible to use keystroke dynamics to accurately verify computer users, albeit in somewhat of a controlled environment.

6. SUMMARY

In this paper we survey the theoretical foundations and design methodologies behind some practical biometric identification technologies being deployed today. Issues pertaining to the application of biometrics to general security models are discussed, with a primary focus on the theoretical foundations for some emerging technologies that offer advanced security countermeasures in the face of intentional disguises by intruders.

We survey identification methods that can be classified under two *broad* categories, namely (1) face recognition and (2) hand recognition. Four different, but related aspects of the problem of recognition, namely representation, detection, identification and classification are examined with particular focus on the latter two problems.

The practical importance of using biometrics to restrict access to computers and digital networks is also addressed. It is our opinion that for the foreseeable future, regardless of their success, these biometric solutions will not eliminate the need for passwords and PINs — rather, the use of biometric technologies will provide a significantly higher level of identification and accountability than passwords and cards alone. For this reason we expect that biometrics will play a crucial role in situations where security is paramount.

REFERENCES

ALEXANDRE, T. J. 1996. Biometrics on Smartcards: An approach to keyboard behavioral signature. In *Second Smart Card Research & Advanced Applications Conference* (1996).

BELHUMEUR, P., HESPANHA, J., AND KRIEGMAN, D. 1995. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *Center for Systems Science 9506, Yale University, PO box 208267, New Haven, CT 06520*.

BOURLARD, H. AND KAMP, Y. 1988. Auto-association by multilayer perceptrons and singular value decomposition. *Biological Cybernetics 59*, 291–294.

BROWN, M. AND ROGERS, S. J. 1993. User Identification via Keystroke Characteristics of Typed Names using Neural Networks. *International Journal of Man-Machine Studies 39*, 6, 999–1014.

BUHMANN, J., LANGE, J., AND VON DER MALSBURG, C. 1989. Distortion invariant object recognition by matching heirachically labelled graphs. *Proc. IJCNN*, 151–159.

CHELLAPPA, R., WILSON, C. L., AND SIROHEY, S. 1995. Human and Machine Recognition of Human Face Images: A survey. In *Proceeding of the IEEE*, Volume 83 (1995), pp. 705–741.

COTTRELL, G. W. AND FLEMING, M. 1990. Face recognition using unsupervised feature extraction. In *International Neural Network Conference*, Volume 1 (1990), pp. 322–325.

CRAW, I., ELLIS, H., AND LISHMAN, J. R. 1987. Automatic extraction of face-features. *Pattern Recognition Letters 5*, 183–187.

DAUGMAN, J. G. 1993. High Confidence Visual Recognition of Persons by a Test of Statistical Independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence 15*, 11 (November).

DUDA, R. 1973. *Pattern Classification and Scene Analysis*. John Wiley & Sons, New York.

FISHER, R. A. 1936. The use of multiple measurements in taxonomic problems. *Ann. Eug. 7*, 2, 179–188.

FLOM, L. AND SAFIR, A. 1987. Iris Recognition System. *US Pat. 4,641,349*.

FU, K. S. 1971. *Pattern Recognition and Machine Learning*. Academic Press, New York.

FUKANAGA, K. 1972. *Introduction to Statistical Pattern Recognition*. Academic Press, New York.

GAINES, R., LISOWSKI, W., PRESS, S., AND SHAPIRO, N. 1980. Authentication by Keystroke Timing: some prelimary results. *Rand Report R-256-NSF. Rand Corporation*.

GENTNER. 1993. Keystroke timing in transcription typing. *Cognitive Aspects of skilled typewritting*, 95–120.

HOPCROFT, J. AND WONG, J. 1974. Linear time algorithm for isomorphism of planar graphs. *Sixth annual ACM Symposium on Theory of Computing*.

HUBER, P. J. 1981. *Robust Statistics*. Wiley.

HUSSIEN, B., MCLAREN, R., AND BLEHA, S. 1989. An application of fuzzy algorithms in a computer access security system. *Pattern Recognition Letters 9*, 39–43.

ISENOR, D. K. AND ZAKY, S. 1986. Fingerprint Identification using Graph Matching. *Pattern Recognition 19*, 2, 113–222.

JOYCE, R. AND GUPTA, G. 1990. Identity Authorization Based on Keystroke Latencies. *Communications of the ACM 33*, 2 (Feb.), 168–176.

KARU, K. AND JAIN, A. K. 1996. Fingerprint Classification. *Pattern Recognition 29*, 3, 389–404.

KAWAGOE, M. AND TOJO, A. 1984. Fingerprint Pattern Classification. *Pattern Recognition 17*, 3, 295–303.

LEGGETT, J. AND WILLIAMS, G. 1988. Verifying Identity via Keystroke Characteristics. *International Journal of Man-Machine Studies 28*, 1, 67–76.

LEGGETT, J., WILLIAMS, G., AND UMPHRESS, D. 89. Verification of User Identity via Keystroke Characteristics. *Human Factors in Management Information Systems*.

MAHAR, D., NAPIER, R., WAGNER, M., LAVERTY, W., HENDERSON, R., AND HIRON, M. 1995. Optimizing digraph-latency based biometric typist verification systems: inter and intra

typists differences in digraph latency distributions. *Int. Journal of Human-Computer Studies 43*, 579–592.

MILLER, B. 1994. Vital Sings of Identity. *IEEE Spectrum*, 22–30.

MOAYER, B. AND FU, K. S. 1975. A Syntactic Approach to Fingerprint Pattern Recognition. *Pattern Recognition 7*, 1–23.

MOGHADDAM, B., NASTAR, C., AND PENTLAND, A. 1996. Bayesian Face Recognition using Deformable Intensity surfaces. *IEEE conference on Computer Vision & Pattern Recognition*.

MONROSE, F. AND RUBIN, A. 1997. Authentication via Keystroke Dynamics. *Fourth ACM Conference on Computer and Commmunications Security*, 48–56.

NAPIER, R., MAHAR, L. D., HENDERSON, R., HIRON, M., AND WAGNER, M. 1995. Keyboard user verification: toward an accurate, efficient, and ecologically valid algorithm. *Int. Journal of Human-Computer Studies 43*, 213–222.

PERRY, J. L. AND CARNEY, J. M. 1990. Human face recognition using a multilayer perceptron. *International Conference on Neural Networks 2*, 413.

POGGIO, T. AND EDELMAN, S. 1990. A Network that learns to recognize three-dimensional objects. *Nature 343*, 6225, 1–3.

PROKOSKI, F. J., RIEDEL, R. B., AND COFFIN, J. S. 1992. Identification of Individuals by Means of Facial Thermography.

RAO, K. AND BALCK, K. 1980. Type classification of fingerprints:a syntactic approach. *IEEE Transaction on Pattern Analysis and Machine Intelligence 2*, 223–231.

SAMAL, A. AND IYENGAR, P. A. 1992. Automatic recognition and analysis of human faces and facial expressions: A survey. *Pattern Recognition 25*, 65–77.

TOU, J. T. AND GONZALEZ, R. 1981. *Pattern Recognition Principles.* Addison-Wesley.

TURK, M. AND PENTLAND, A. 1991a. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience 3*, 1, 71–86.

TURK, M. AND PENTLAND, A. 1991b. Face Recognition using Eigenfaces. *Computer Vision and Pattern Recognition*, 586–591.

WILDES, R. AND ASMUTH, J. 1994. A System for Automatic Iris Recognition. *Pattern Recognition*, 121–128.

YUILLE, A. L. 1991. Deformable templates for face recognition. *Journal of Cognitive Neuroscience 3*, 1, 59–70.

YULIE, A., COHEN, D. S., AND HALINAN, P. 1989. Feature extraction using a multilayer perceptron. *Computer Vision and Pattern Recognition*, 104–109.

ZHANG, J., YAN, Y., AND LADES, M. 1997. Face Recognition:Eigenface, Elastic Matching and Neural Nets. *Proceedings of the IEEE 85*, 9 (Sept.), 1423–1435.