

Digital Image Processing

S. Aravindhana¹

¹ Saveetha School of Engineering

Abstract— Most systems that control access to financial transactions, computer networks, or secured locations identify authorized persons by recognizing passwords or personal identification numbers. The weakness of these systems is that unauthorized persons can discover others' passwords and numbers quite easily and use them without detection. Biometric identification systems, which use physical features to check a person's identity, ensure much greater security than password and number systems. Biometric features such as the face or a fingerprint can be stored on a microchip in a credit card, for example. If someone steals the card and tries to use it, the impostor's biometric features will not match the features stored in the card, and the system will prevent the transaction. Biometrics are automated methods of recognizing a person based on a physiological or behavioral characteristic. Biometric technologies are becoming the foundation of an extensive array of highly secure identification and personal verification solutions. A single feature, however, sometimes fails to be exact enough for identification. Consider identical twins, for example. Their faces alone may not distinguish them. Another disadvantage of using only one feature is that the chosen feature is not always readable. For example, some five percent of people have fingerprints that cannot be recorded because they are obscured by a cut or a scar or are too fine to show up well in a photograph. This paper presents a system called BioID which is developed to identify a person using different features—Face, voice, lip movement, iris recognition, finger and palm geometry. With its three modalities, BioID achieves much greater accuracy than single feature systems.

Key words: lip movement, iris recognition, finger and palm geometry.

I. INTRODUCTION

Biometric (Biological features as a measure) recognition refers to the use of distinctive physiological and behavioral characteristics (e.g., fingerprints, face, hand geometry, iris, gait, signature), called biometric identifiers or simply biometrics, for automatically recognizing a person. In multimodal biometric identification systems even if one modality is somehow disturbed, for example, if a noisy environment drowns out the voice—the other two modalities still lead to an accurate identification. BioID is the first identification system that uses a dynamic feature, lip movement. This feature makes BioID more secure against fraud than systems using only static features such as fingerprints.

A. BioID'S:

BioID is suitable for any application in which people require access to a technical system such as computer networks, Internet commerce and banking systems, and ATMs, for example. In addition, this system secures access to rooms and buildings. So far, most BioID installations serve physical structures and small office-computer networks. Depending on the application, BioID authorizes people

either through identification or verification. In identification mode, the system identifies a person exclusively through biometric traits. In verification mode user name or a number are used, system then verifies by means of biometric traits. Figure shows a user interacting with the system



Fig. 1: user interacting with the system

B. Interacting with BioID:

Seeking access to a computer Network, user would pose in front of the PC camera and speaks his name.

II. SYSTEM FUNCTIONS

Figure shows BioID's functions. The system acquires (records), preprocesses, and classifies each biometric feature separately. During the training (enrollment) of the system, biometric templates are generated for each feature. For classification, the system compares these templates with the newly recorded pattern. Then, using a strategy that depends on the level of security required by the application, it combines the classification results into one result by which it recognizes persons.

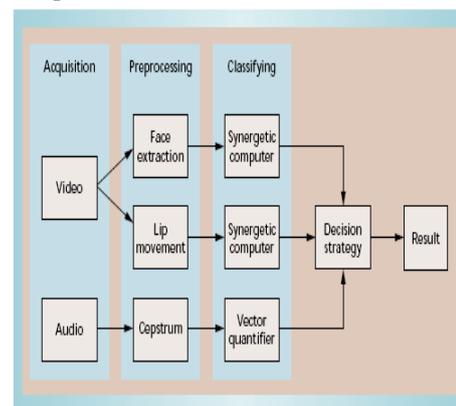


Fig. 2: BioID Function

A. BioID's main functions:

From video and audio samplings of a person speaking, the system extracts facial, lip movement, and voice features (a cepstrum is a special form of the frequency spectrum). Synergetic computers and a vector quantifier classify the recorded pattern and combine the results.

III. DATA ACQUISITION AND PREPROCESSING

The input to the system is a recorded sample of a person speaking. The one-second sample consists of a 25- frame video sequence and an audio signal. From the video sequence, the preprocessing module extracts two optical biometric traits: face and lip movement while speaking a word. To extract those features, the preprocessing module must have exact knowledge of the face's position. Since this recognition system should be able to function in any arbitrary environment with off-the- shelf video equipment, the face-finding process is one of the most important steps in feature extraction

IV. FACIAL FEATURES

For face recognition, the preprocessing module uses the first image in the video sequence that shows the person with eyes open. Once the eyes are in position, the preprocessing module uses anthropomorphic knowledge to extract a normalized portion of the face. That is, it scales all faces to a uniform size, as shown



Fig. 3: Facial Features

A. Samples of different faces:

BioID scales all faces to the same size and crops the images uniformly for easier comparison. This photograph collection shows 12 individuals; here we note the uniformity that the system achieves.

This Procedure ensures that the appropriate facial features are analyzed (but not, for example, the head size, the hairstyle, a tie, or a piece of jewelry). After rotating and scaling the image, the preprocessing module extracts a grayscale image. Some further preprocessing steps take care of lighting conditions and color variance.

B. Using Hausdorff distance for face location:

To detect the location of a face in an arbitrary image, identification systems often use neural-net-based algorithms, but these approaches are very time-consuming. Instead, BioID uses a newly developed, model-based algorithm that matches a binary model of a typical human face to a binarized, edge-extracted version of the video image. The figure illustrates this process. The face extractor bases its comparison on the modified

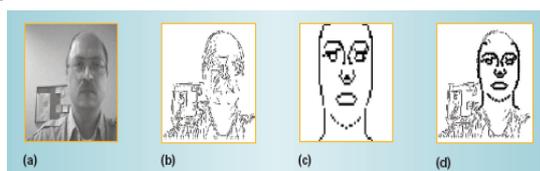


Fig. 4: Face location (a) original image, (b) edge-extracted image, (c) face model, and (d) face model overlaid on the edge-extracted image

C. Hausdorff distance:

which determines the model's optimal location, scaling, and rotation. The Hausdorff distance uses two point sets, A and B. To obtain the Hausdorff distance, we calculate the minimum distance from each point of set A to a point of set B and vice versa. The maximum of these minimum distances is the Hausdorff distance. Point set A represents the face model, and point set B is a part of the image. The minimum of the calculated maximum distances determines the part of the image where the face is located. After detecting the face boundaries, the preprocessing module locates the eyes from the first three images of the video sequence, under the assumption that a person often closes his eyes when beginning to speak.

As with face location, eye location also relies on an image model and the Hausdorff distance.

Classification:

- (1) Lip movement and face classification
- (2) Finger and palm geometry
- (3) Voice recognition
- (4) Iris recognition

V. LIP MOVEMENT AND FACE RECOGNITION

The synergetic computer serves as a learning classifier for optical biometric- pattern recognition. In the training phase, BioID records several characteristic patterns of one person's face and lip movement, and assigns them to a class. Each class represents one person. During the training process, all patterns are orthogonalized and normalized. The resulting vectors, called adjunct prototypes, are compressed in each class. This leads to one prototype for each class (person), representing all patterns initially stored in the class without any loss of information. This prototype is called as biometric template. The classification process is fairly easy: Firstly preprocess and multiply a newly recorded pattern with each biometric template and then rank the obtained scalar products, and the highest one (as an absolute value) leads to the resulting class. This strategy is known as winner-takes-all. Because this principle always leads to a classification that is, no pattern is rejected. We also take the second highest scalar product into account. If the difference between the highest and the second highest is smaller than a given threshold, we reject the pattern.

VI. CLASSIFICATION OF RESULTS

If the two highest scalar products have nearly the same value, the two classes (two people) are indistinguishable, and the classification is "insecure." The training process for the optical features of 30 persons with five learning patterns echo takes about 15 minutes on an Intel Pentium II. The classification time is very short (several milliseconds) since there are only 30 scalar Products to calculate.

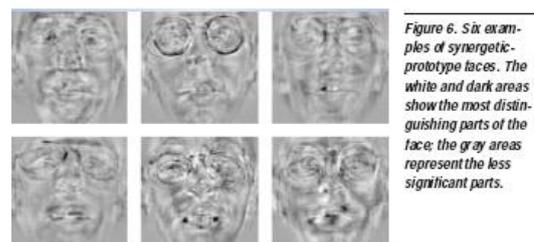


Figure 6. Six examples of synergetic-prototype faces. The white and dark areas show the most distinguishing parts of the face; the gray areas represent the less significant parts.

Fig. 5: facial biometric templates of six classes (six people).
The figure shows the facial biometric templates of six classes (six people). In this each template consists of several overlying patterns

VII. OPTICAL FLOW TECHNIQUE

BioID collects lip movements by means of an optical-flow technique that calculates a vector field representing the local movement of each image part to the next image in the video sequence of several smaller, overlapping windows. For each window, it calculates the cepstral coefficients, which form the audio feature vector. The vector quantifier uses this feature vector for classifying audio patterns.

For this process, the preprocessing module cuts the mouth area out of the first 17 images of the video sequence. It gathers the lip movements in 16 vector fields, which represent the movement of identifiable points on the lip from frame to frame. shows the optical-flow vector field of two consecutive images.

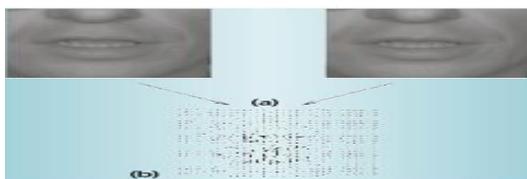


Fig 6: The lip movement between the two images is defined by the vector field
Example of an optical-flow vector field. The lip movement between (a) the two images is defined by (b), the vector field.

VIII. REDUCING AMOUNT OF DATA

To reduce the amount of data, we reduce the optical flow resolution to a factor of four through averaging. Finally, a 3D fast Fourier transformation of the 16 vector fields takes place. The result is a one-dimensional lip movement feature vector, which the system uses for training and classification of lip movement. Essentially, we are condensing the detailed movement defined by several vector fields to a single vector. The following figure presents an overview of the optical preprocessing steps.

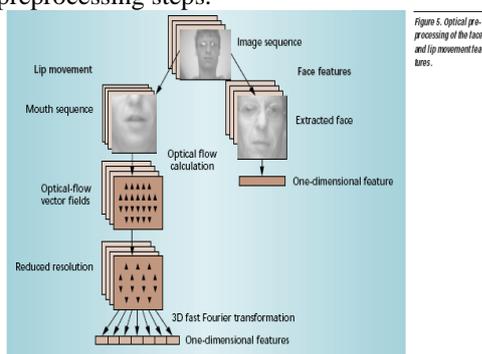


Fig. 7: overview of the optical preprocessing steps.

IX. FINGER GEOMETRY

Finger geometry biometric is very closely related to hand geometry. The use of just one or two fingers means more robustness, smaller devices and even higher throughput. Two variations of capture processes are used, first being

similar to hand geometry presented above. The second technique requires the user to insert a finger into a tunnel so that three-dimensional measurements of the finger can be made.

X. PALM RECOGNITION

Palm biometrics is close to finger scanning and in particular AFIS technology. Ridges, valleys and other minutiae data are found on the palm as with finger images. Main interest in palm biometrics industry is law enforcement as latent images - "palm prints" - found from the crime scenes are equally useful as latent fingerprints. Certain vendors are also looking at the access control market and hope to follow the footsteps of finger scanning

XI. VOICE RECOGNITION

Voice biometrics examines particularly the sound of the voice. Speech recognition can be defined as a system that recognizes words and phrases that are spoken. Voice identification has been derived from the basic principles of speech recognition.

A. Speaker recognition:

focuses on recognizing the speaker, and is accomplished either by speaker verification or speaker identification.

B. Speaker verification:

is a means of accepting or rejecting the claimed identity of a speaker.

C. Speaker identification:

is the process of determining which speaker is present based solely on the speaker's utterance. The speaker identification application evaluates the input with models stored in a database to determine the speaker's identity.

The sound of a human voice is caused by resonance in the vocal tract. The length of the vocal tract, the shape of the mouth and nasal cavities are all important. Sound is measured, as affected by these specific characteristics. The technique of measuring the voice is discussed below.

We use vector quantification to classify the audio sequence. In the system-training phase, the audio preprocessing module analyzes several recordings of a single person's voice. From each voice pattern, it creates a matrix, and the vector quantifier combines these matrices into one matrix. This matrix serves as a prototype (or codebook) that displays the reference voice pattern. Using this voice pattern, a minimum distance classifier assigns the current pattern to the class showing the smallest distance.

XII. ACOUSTIC PREPROCESSING

We record the speech sample using a 22-kHz sampling rate with 16-bit resolution. After channel estimation and normalization, the preprocessing module divides the time signal into several smaller, overlapping windows. For each window, it calculates the cepstral coefficients, which form the audio feature vector. The vector quantifier uses this feature vector for classifying audio patterns.

XIII. IRIS RECOGNITION

Iris is the colored part of the eye that consists of a muscular diaphragm surrounding the pupil and regulating the light entering the eye by expanding and contracting the pupil. Iris-recognition technology was designed to be less intrusive than retina scans, which often require infrared rays or bright light to get an accurate reading. Scientists also say a person's retina can change with age, while an iris remains intact. And no two iris blueprints are mathematically alike, even between identical twins and triplets.

To record an individual's iris code, a black-and-white video camera uses 30 frames per second to zoom in on the eye and "grab" a sharp image of the iris. A low-level incandescent light illuminates the iris so the video camera can focus on it, but the light is rarely noticeable and used strictly to assist the camera. One of the frames is then digitized and stored in pc database of enrolled users. The whole procedure takes less than a few seconds, and can be fully computerized using voice prompts and auto focus.

An iris has a mesh-like texture to it, with numerous overlays and patterns that can be measured by the computer, said Johnston. The iris-recognition software uses about 260 "degree of freedom," or points of reference, to search the data for a match. By comparison, the best fingerprint technology uses about 60 to 70 degrees of the freedom, he noted.

This biometric technology could also be used to secure your computer files. By mounting a Webcam to your computer and installing the facial recognition software, your face can become the password you use to get into your computer. IBM has incorporated the technology into a screensaver for its A, T and X-series Thinkpad laptops.

Advantages:

- (1) Glasses and contact lenses wont interfere in this.
- (2) Blind peoples are also involved in this.
- (3) Cataract, cornea transplant, surgery wont disturbs this process.

XIV. SENSOR FUSION

To analyze the classification results, BioID chooses from different strategies to obtain various security levels. Figure shows the available sensor fusion options, that is, the combinations of the three results.

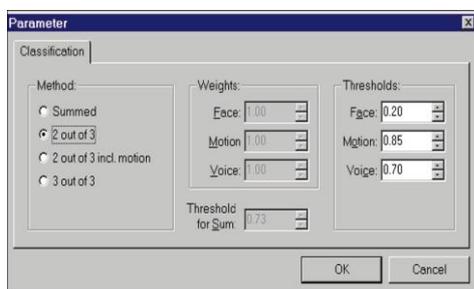


Fig. 8: Sensor Function

For normal operations, the system uses a two-out-of-three strategy, which classifies two of the three biometric features to an enrolled class (person), without falling below threshold values set in advance. The threshold values apply to the relative distances of the best and the second-best scalar products--that is, the two classes that best match and can be

determined by the system administrator. For a higher security level, the system can demand agreement of all three traits, a three-out-of-three strategy.

With this strategy, the probability that the system will accept an unauthorized person decreases, but one must live with the possibility that it will reject an authorized person. Additional methods make the sum of the classification results of all traits available. These methods allow us to weight individual traits differently. For example, if the system always correctly identifies a person by lip movement, this feature will be more significant than the others.

A. APPLICATIONS:

- (1) Electronic commerce
- (2) Information security
- (3) Entitlements authorization
- (4) Building entry
- (5) Automobile ignition
- (6) Forensic and police applications
- (7) Network access

XV. FUTURE BIOMETRICS

A system that analyses the chemical make-up of body odor is currently in development. In this systems sensors are capable of capturing body odor from non-intrusive parts of the body such as the back of the hand. Each unique human smell consists of different amount of volatiles. These are extracted by the system and converted into a biometric template.

All testing and fastest possible analysis of the human DNA takes at least 10 minutes to complete and it needs human assistance. Thus, it cannot be considered as biometric technology in its sense of being fast and automatic.

Additionally current DNA capture mechanisms, taking a blood sample or a test swab inside of the mouth, are extremely intrusive compared to other biometric systems. Apart from these problems DNA, as a concept, has a lot of potential.

Ear shape biometrics research is based on law enforcement needs to collect ear markings and shape information from crime scenes. It has some potential in some access control applications in similar use as hand geometry. There are not excessive research activities going on with the subject.

XVI. KEYSTROKE DYNAMICS

keystroke dynamics is a strongly behavioural, learnt biometric. As being behavioural, it evolves significantly as the user gets older. One of the many problems include that highly sophisticated measuring software and statistical calculations have to be made real time if the user actions should be constantly verified. Standard keyboard could be used in simplest cases.

XVII. VEINCHECK

Vein check is a technique where infrared camera is used to extract vein pattern from the back of the hand. The pattern is very unique and the capture method is user friendly and

non-intrusive as hand geometry check. Maybe combining them could result very accurate and easy-to-use biometric.

XVIII. CONCLUSION

With its multimodal concept, BioID guarantees a high degree of security from falsification and unauthorized access. It also protects the privacy rights of system users, who must speak their name or a key phrase, and therefore cannot be identified without their knowledge. To guard against the threat of unauthorized use, users can invalidate their stored reference template at any time, simply by speaking a new word and thus creating a new reference template. In a test involving 150 persons for three months, BioID reduced the false-acceptance rate significantly below 1 percent, depending on the security level.

The higher the security level, the higher the false-rejection rate. Thus, system administrators must find an acceptable false-rejection rate without letting the false-acceptance rate increase too much. The security level depends on the purpose of the biometric system.

Biometric templates of people provide a reference from one human being unique to just one identity. This can be too tempting target to link different personal datas to if stored on a central database. Solutions with central databases are reasoned for better service to John Doe customer. For example replacement of smart card which has biometric information inside is time consuming and inconvenient if the biometric data cannot be recovered from anywhere else than the users body itself.

REFERENCE

- [1] R. Frischholz, and U. Dieckmann, BioID: A Multimodal Biometric Identification System, in *IEEE Computer*, Vol. 33, No. 2, pp. 64-68, February 2000.
- [2] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network Based Face Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Jan. 1998
- [3] D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge, "Comparing Images Using the Hausdorff Distance," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Sept. 1993,