Introduction-cont
Pattern classification

Biometrics
CSE 190-a
Lecture 2

How are people identified?

• People are identified by three basic means:
  – Something they have (identity document or token)
  – Something they know (password, PIN)
  – Something they are (human body)

Problems with Possession- or Knowledge-based Approaches

• Card may be lost, stolen or forgotten
  – Password or PIN may be forgotten or guessed by the imposters
• ~25% of people seem to write their PIN on their ATM card
• Estimates of annual identity fraud damages:
  – $56.6 billion in credit card transactions in U.S. alone in 2005*
  – 0.25% of internet transactions revenues, 0.08% of off-line revenues
  – $1 billion in fraudulent cellular phone use
  – $3 billion in ATM withdrawals
• The traditional approaches are unable to differentiate between an authorized person and an imposter

* Spectrum July, 2006

Requirements for an Ideal Biometric Identifier

1. Universality
   – Every person should have the biometric characteristic
2. Uniqueness
   – No two persons should be the same in terms of the biometric characteristic
3. Permanence
   – The biometric characteristic should be invariant over time
4. Collectability
   – The biometric characteristic should be measurable with some (practical) sensing device
5. Acceptability
   – One would want to minimize the objections of the users to the measuring/collection of the biometric

Behavioral vs Physical Traits

• Physical Characteristics
  – Iris
  – Retina
  – Vein Pattern
  – Hand Geometry
  – Face
  – Fingerprint
  – Ear shape

• Behavioral Characteristics
  – Keystroke dynamics
  – Signature dynamics
  – Walking Gait
  – Voice

About this Class

• See Syllabus

• Special thanks to
  – Peter Belhumeur
  – Anil Jain
There are ~500 million border crossings/year in the U.S.

As part of the enhanced procedures, most visitors traveling on visas will have two fingerprints scanned by an inkless device and a digital photograph taken. All of the data and information is then used to assist the border inspector in determining whether or not to admit the traveler. These enhanced procedures will add only seconds to the visitor’s overall processing time.

The electronic fingerprint scanner allows inspectors to check the visitors’ fingerprints against those on terrorist watch lists.

The Nine Zero hotel in Boston just installed a new system which uses digital photos of the irises of employees, vendors and VIP guests to admit them to certain areas, the same system used in high-security areas at airports such as New York’s JFK.

**Access Control**

- Automatic **personalization** of vehicle settings:
  - Seat position
  - Steering wheel position
  - Mirror positions
  - Lighting
  - Radio station preferences
  - Climate control settings

- **URLs** at your fingertips

**Applications**

<table>
<thead>
<tr>
<th>Forensic</th>
<th>Government</th>
<th>Commercial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coroner Identification, Criminal Investigation, Terrorist Identification, Parolee Identification, Missing Children, etc.</td>
<td>National ID card, Correctional Facility, Driver’s License, Social Security, Welfare Disbursement, Border Control, Passport Control, etc.</td>
<td>Computer Network Logon, Electronic Data Security, E-Commerce, Internet Access, ATM, Credit Card, Physical Access Control, Cellular Phones, Personal Digital Assistant, Medical Records Management, Distance Learning, etc.</td>
</tr>
</tbody>
</table>
What makes using biometrics difficult?

Why is Biometric Recognition Difficult?
- Large number of classes (e.g., millions of faces)
- Intra-class variability and inter-class similarity
- Segmentation
- Noisy and distorted images
- Population coverage & scalability
- System performance (error rate, speed, throughput, cost)
- Attacks on the biometric system
- Template ageing
- Non-uniqueness of biometric characteristics
- Addressing privacy concerns

Intra-class variability

Inter-class Similarity

Temporal Variations

Locating Faces in a Crowd

Games Magazine, September 2001
Noisy Images

• ~3% of the population has poor quality fingerprint images

Four impressions of a user's fingerprint

Attacks on Biometric Systems

• Spoofing a biometric trait

Dummy finger created from a lifted impression

Artificial skin/fingers (http://www.livingskin.com/)

Sensor Interoperability

• Sensors used during enrollment and verification may be different

Sensor Interoperability

A rolled inked fingerprint

Digital Biometrics optical sensor (508x480)

Veridicom capacitive sensor (300x300)

Fidelica pressure sensor (256x256)

Sensors used during enrollment and verification may be different

Performance Evaluation

• The overall performance of a biometric system is assessed in terms of its universality, accuracy, speed, and storage

• Factors like cost and ease of use also affect performance

• Biometric systems are not perfect, and can mistakenly accept an impostor as a valid user (a false match) or conversely, reject a valid individual (a false non-match)

Best Practices:  www.cesg.gov.uk/technology/biometrics
FRVT2002:  www.frvt.org/
FVC 2004:  bias.csr.unibo.it/fvc2004
NIST SV:  www.nist.gov/speech/tests/spk

(c) Jain 2004
Performance Characterization

- Impostor Distribution
- Genuine Distribution
- Threshold
- False Accept Rate (FAR) or False Match Rate
- False Reject Rate (FRR) or False Non-match Rate
- Receiver Operating Characteristic (ROC) curve
- Equal Error Rate or Crossover Rate
- Failure to Enroll (FTE)
- Failure to Acquire (FTA) or Failure to Capture

Error Rates

- False Match (False Accept): Mistaking biometric measurements from two different persons to be from the same person; False Non-match (False reject): Mistaking two biometric measurements from the same person to be from two different persons.

Error vs Threshold

- FAR: False accept rate
- FRR: False reject rate

ROC Curve

- Accuracy requirements of a biometric system are application dependent.

Evaluation Protocol

- Define a protocol to test the system, select the data and measure the performance; performance depends on the test set.
- Evaluations should be conducted by an independent organization (that is not involved in the design of the system).
- Test on biometric data previously unseen by the system.
- Size of the data-set and representative examples of the data set should be provided for tuning algorithmic parameters.
- Face, Fingerprint and Voice systems have undergone the most study and testing.

"State-of-the-art" Error Rates

<table>
<thead>
<tr>
<th>Test</th>
<th>Test Parameter</th>
<th>False Reject Rate</th>
<th>False Accept Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fingerprint</td>
<td>FVC (2002) 20 years (average age)</td>
<td>0.2%</td>
<td>0.2%</td>
</tr>
<tr>
<td></td>
<td>FVC (2004) Deliberate perturbations</td>
<td>2.07%</td>
<td>2.07%</td>
</tr>
<tr>
<td>Face</td>
<td>FRVT (2002) Varied lighting, outdoor/indoor</td>
<td>10%</td>
<td>1%</td>
</tr>
<tr>
<td>Voice</td>
<td>NIST (2000) Text Independent</td>
<td>10-20%</td>
<td>2-5%</td>
</tr>
</tbody>
</table>

At NY airports, an average of ~ 300,000 passengers pass through daily. If all of these used biometric-authenticated smart cards for identification, there would be 600 falsely rejected (and inconvenienced) passengers per day for fingerprints, 30,000 for face and 45,000 for voice. Similar numbers can be computed for false accepts.
Biometrics: A Pattern Recognition System

Chapter 1: Introduction to Pattern Recognition (Sections 1.1-1.6)

- Machine Perception
- An Example
- Pattern Recognition Systems
- The Design Cycle
- Learning and Adaptation

An Example

“Sorting incoming Fish on a conveyor according to species using optical sensing”

Species
- Sea bass
- Salmon

Problem Analysis

- Set up a camera and take some sample images to extract features
  - Length
  - Lightness
  - Width
  - Number and shape of fins
  - Position of the mouth, etc...

  This is the set of all suggested features to explore for use in our classifier!

Preprocessing

- Use a segmentation operation to isolate fishes from one another and from the background
- Information from a single fish is sent to a feature extractor whose purpose is to reduce the data by measuring certain features
- The features are passed to a classifier

Reading List

Performance Metrics and Evaluation

   http://bias.csr.unibo.it/fvc2004/
2. Fingerprint Vendor Technology Evaluation (FpVTE 2003):
   http://fpvte.nist.gov/
   http://www.frvt.org/
4. Face Verification Contest on the BANCA dataset (2004):
   http://www.ee.surrey.ac.uk/banca/icpr2004/
   http://www.nist.gov/speech/tests/spk/
6. Signature Verification Competition (SVC 2004):
• Classification

  • Select the length of the fish as a possible feature for discrimination

The length is a poor feature alone!

Select the lightness as a possible feature.

• Threshold decision boundary and cost relationship

  • Move our decision boundary toward smaller values of lightness in order to minimize the cost (reduce the number of sea bass that are classified salmon!)

Task of decision theory
• Adopt the lightness and add the width of the fish

$$\mathbf{x}^T = [x_1, x_2]$$

Lightness  Width

• We might add other features that are not correlated with the ones we already have. A precaution should be taken not to reduce the performance by adding such “noisy features”

• Ideally, the best decision boundary should be the one which provides an optimal performance such as in the following figure:

• However, our satisfaction is premature because the central aim of designing a classifier is to correctly classify novel input

Issue of generalization!