

# Information Fusion in Multimedia Information Retrieval: An Overview

Jie Yu

Department of Computer Science  
University of Texas at San Antonio

# Outline

- Information Fusion (I.F.)
  - Introduction
  - Model and Applications
- I.F. in Multimedia Information Retrieval
  - Multimodality and Multi-classifier
  - Case Study: Fusion in Biometrics
  - Our related study: Interactive Boosting
- Discussion and Conclusion

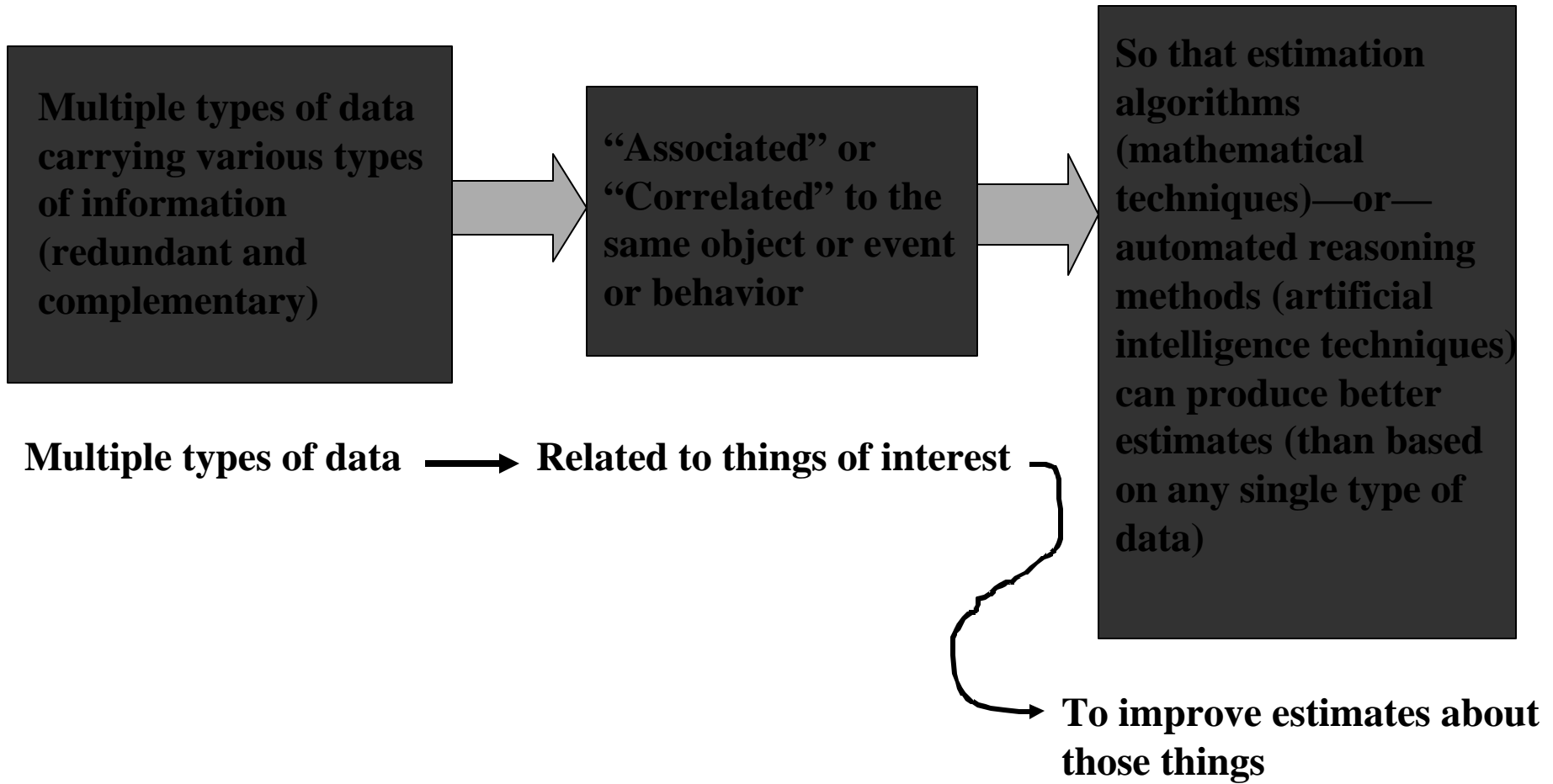
# What is Information Fusion?

- Fusion: A merging of diverse, distinct, or separate elements into a unified whole (Merriam-Webster dictionary).
- “Information fusion is an **Information Process** dealing with the:
  - [**association, correlation, and combination** of data and information] from
  - [**single and multiple sensors or sources**] to achieve
  - [**refined estimates** of parameters, characteristics, events, and behaviors] for observed entities in an observed field of view (JDL, 1999)
- It is sometimes implemented as a Fully Automatic process or as a **Human-Aiding process** for Analysis and/or Decision Support

# Why is fusion necessary?

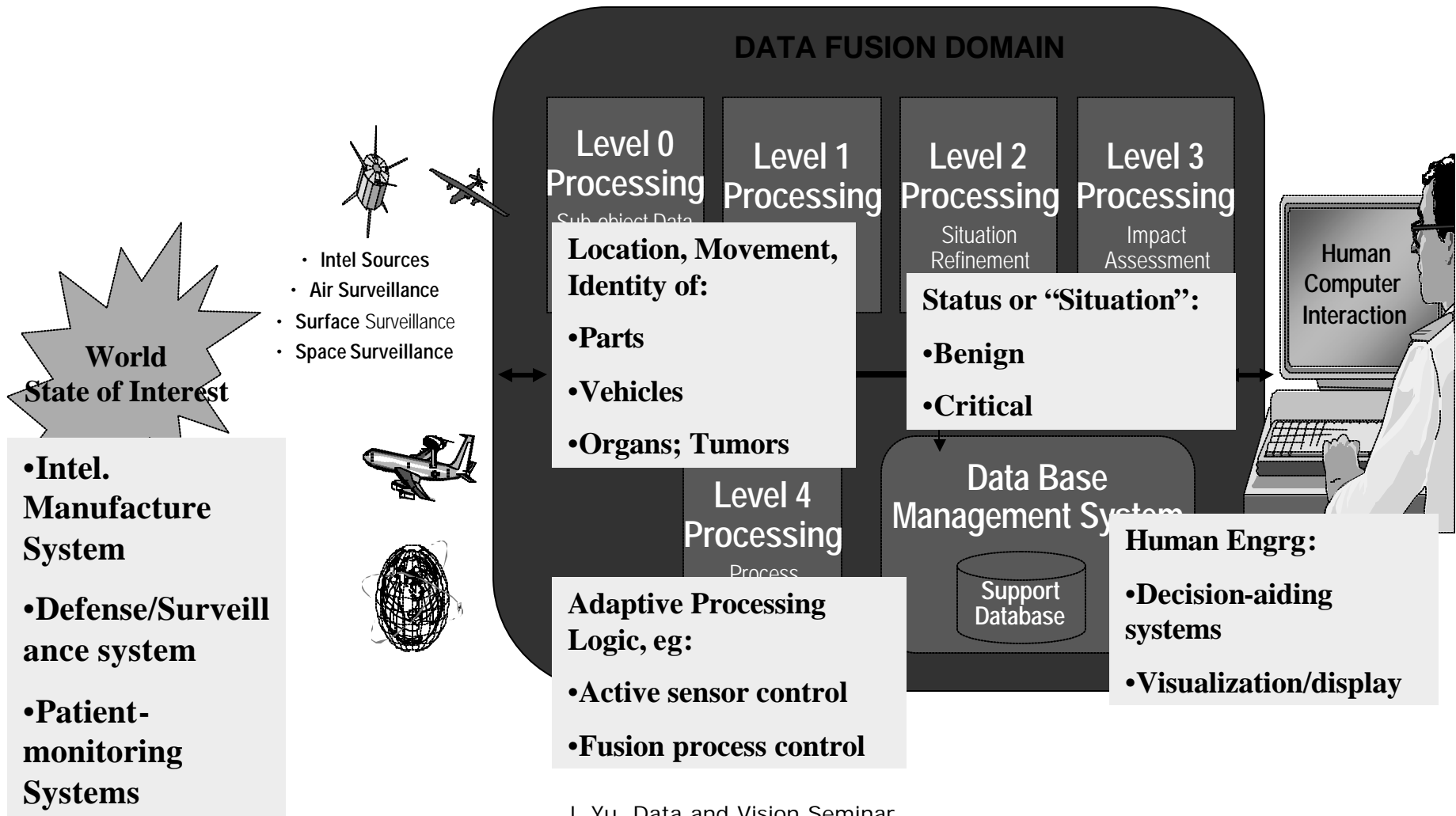
- The fusion of redundant information from different sources can **reduce overall uncertainty** and thus increase the accuracy of the system. Multiple sources providing redundant information can also increase the robustness of the system.
- The fusion of complementary information provided by different sources results in an **information gain** due to the utilization of multiple sources of information v.s. a single source.
- The fusion of information from multiple sensors may provide **more timely information** either because of the actual speed of operation of each sensor, or because of processing parallelism that may possibly be achieved as part of the integration process.

# How Information Fusion Works?



**These Basic Ideas are Transferable to Many Types of Problems**

# Complete Information Fusion System Model (JDL)



# Fusion System Applications

- Military applications:
  - Autonomous weaponry employing multiple sensors
  - Broad area surveillance systems employing single weapons platforms (e.g., ships, airborne surveillance) or distributed sensor networks
  - Fire control systems employing multiple sensors for acquisition, tracking, and command guidance
  - Intelligence collection systems
  - Indications and warning (I&W) systems, the mission of which is to assess threat and hostile intent
  - Command and control nodes for military forces

# Example: Fusion-Based Automatic Object Recognition

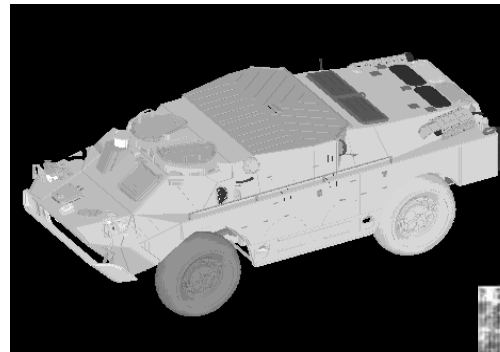
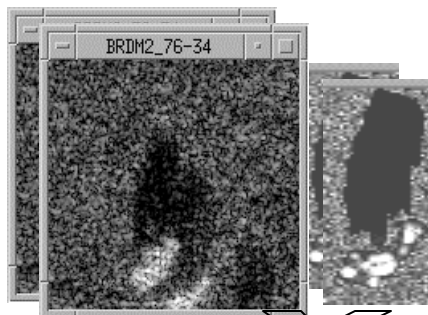
## FUSED COMPONENT MATCH SCORES

### SAR ONLY

34° pose

Collected

Predicted

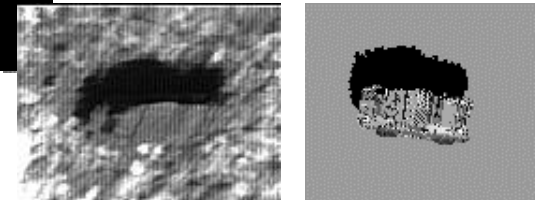


### E-O ONLY

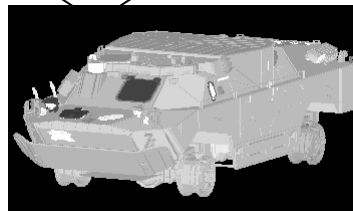
88° pose

Collected

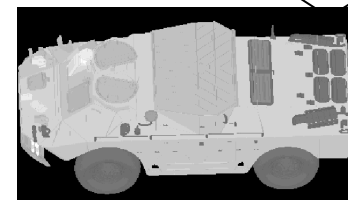
Predicted



### SAR Component Match Scores



### E-O Component Match Scores



- Generate Hypothesis (e.g. BRDM-2, 34° pose, articulation x)
- Predict Measurements
- Evaluate Component-Level Match: Actual vs. Predicted
- Select Hypothesis with Best Match

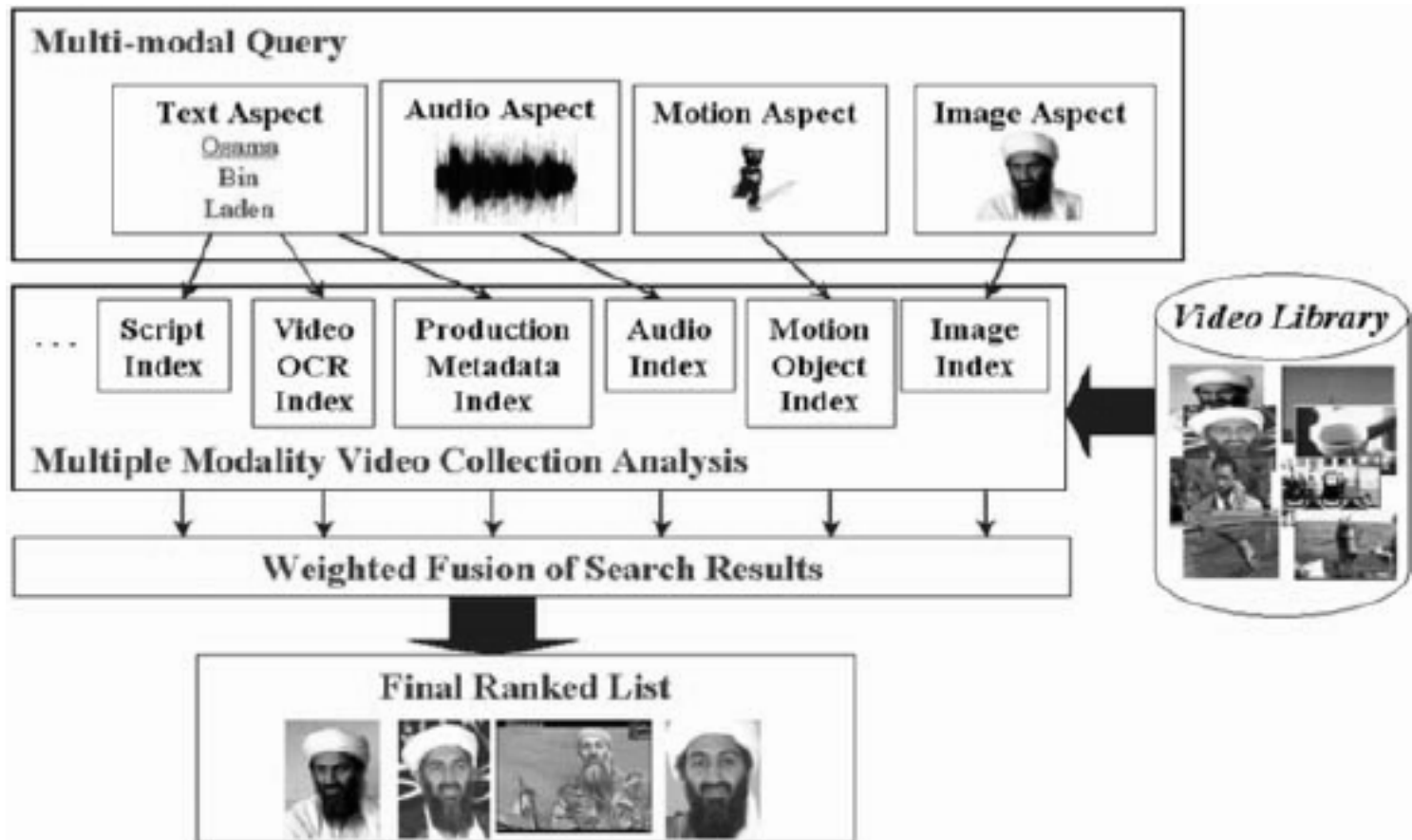
### Match Scores

← Low High →



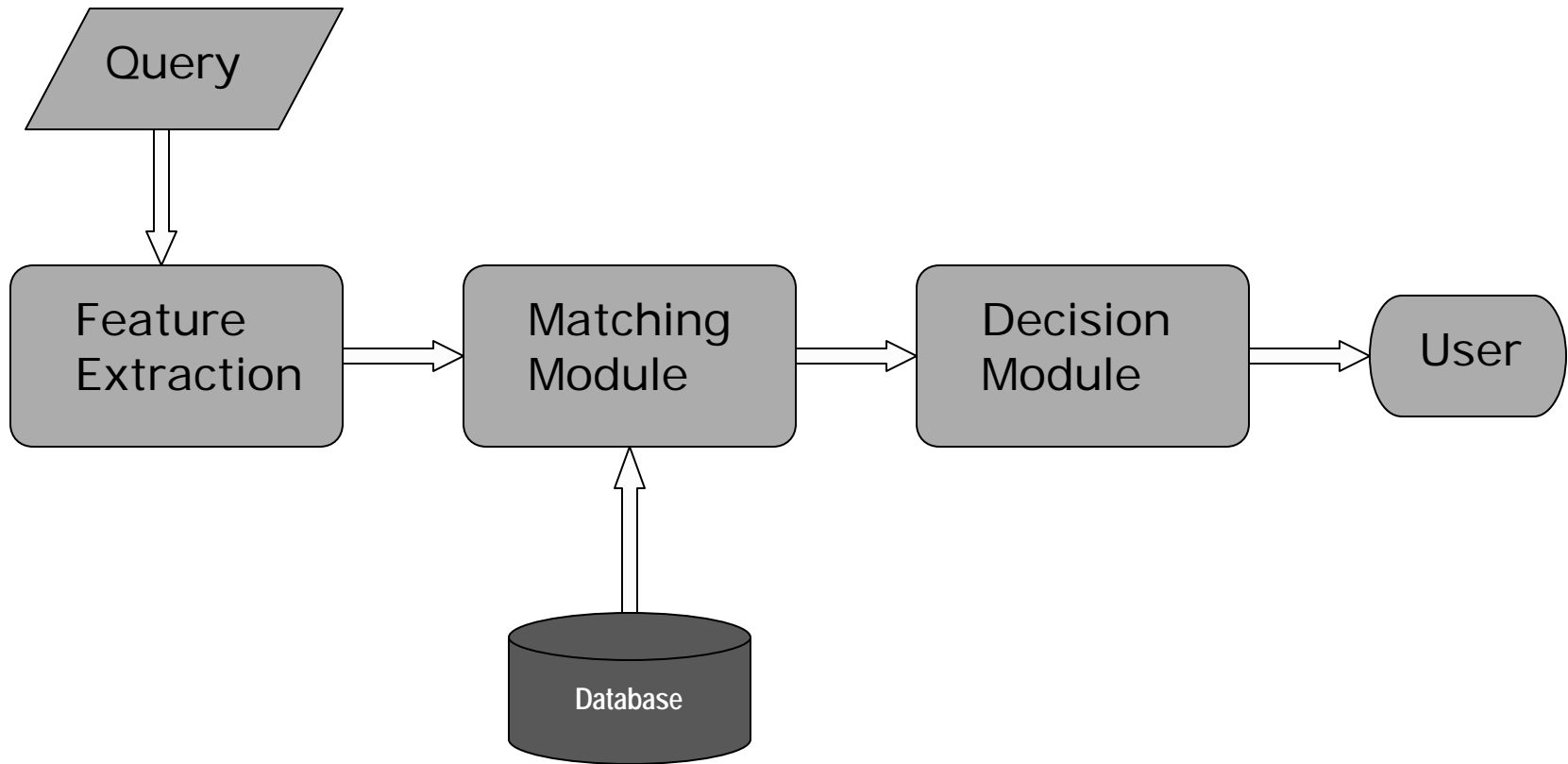
# Fusion System Applications

- Industrial and commercial applications --- robotics, machine intelligence, remote sensing, image processing, medical systems.
  - E.g. Video retrieval



# Multimedia Information Retrieval

- Simple model:



# Information Fusion in MIR

- Feature Extraction Module:
  - Multiple features -> vectors
  - Concatenated vector
  - Feature Fusion: more discriminating hyperspace can be found in the new vector
- Matching Module:
  - One types of classifiers for multiple features or
  - Multiple types of classifiers for one feature or
  - Both
  - The output score can be combined
- Decision Module:
  - The output decision of each classifier can be combined

# Two Forms of Information Fusion in MIR

## ■ Multi-Modality

- Modality refers to the type of *communication channel* used to convey or acquire information. It also covers the way an idea is expressed or perceived, or the manner an action is performed (audio, video, text)
  - E.g. Video clip->visual information, audio information, textural information
- Multi-modality fusion occurs at feature extraction module.
- Single source information may be represented by multiple features,
  - E.g. Color image->color, texture, shape

# Two Forms of Information Fusion in MIR

- Multi-Classifiers (Ensemble of Classifiers, EOS)
  - A set of classifiers are trained to solve the same problem
  - Applied on single or multiple source of information
  - A single type of base classifiers or
  - Different types of classifiers (Bayesian, K-NN, SVM)
  - Multiple classifiers can be generated by
    - Randomize training set
    - Randomize classifier structure
    - Or both
- Advantages:
  - High accuracy ??
  - Can deal with extremely large data sets e.g. biology data
  - Weighting schemes may provide an approach for dealing with skewed data (1000:1), business fraud, internet security

# Fusion Schemes

- The prediction of multiple classifiers need to be integrated into one fused decision by ***Fusion Scheme***.
  - The output of different classifiers need to be normalized.
- Rule-based
  - Decision is made by a simple operation on the output of all classifiers
  - E.g. Max, Min, Sum, Mean (Matching Module)
  - E.g. AND, OR (Decision Module)
- Learning-based
  - The output of all classifiers is fed into a ***learning process*** to obtain the final decision
  - E.g. Decision Tree, Neural Network
- No one is guaranteed to be the best empirically or theoretically.

# Comparison of Simple Fusion Schemes

- Test conducted on 3 UCI benchmark data sets
- PCA and LDA + Bayesian classifiers are fused.
- Simple fusion scheme does not always works.
- No fusion scheme is guaranteed to be the best

Probability Based PCA and LDA Fusion Tests			
Best Error Rate (Threshold ? )	Benchmark		
Fusion Method	Heart	Breast-Cancer	Banana
PCA (no fusion)	0.184 (0.6)	0.2779 (0.6)	0.4272 (0.6)
LDA (no fusion)	0.159 (0.6)	0.25 (0.7)	0.4425 (0.6)
"AND"	0.7175 (0.1)	0.3065 (0.1)	0.4473 (0.2)
"OR"	0.6135 (0.1)	0.2981 (0.2)	0.4432 (0.4)
"Mean"	0.155 (0.5)	0.2688 (0.7)	0.4333 (0.6)
"Max"	0.1575 (0.7)	0.2682 (0.7)	0.4303 (0.6)
"Min"	0.1735 (0.5)	0.2597 (0.6)	0.4194 (0.6)

# Information Fusion in Biometrics

## What is Biometrics?

-(General definition) the science and technology of measuring and analyzing biological data

-(In information technology) biometrics refers to technologies that measure and analyze unique human body characteristics, e.g.

- Fingerprint
- Facial image
- Hand Geometry
- Handwriting
- Iris image
- Voice



(a)



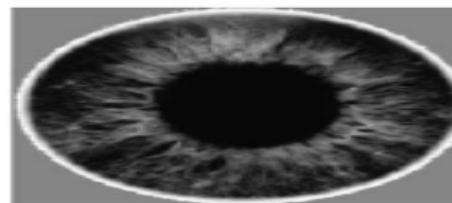
(b)



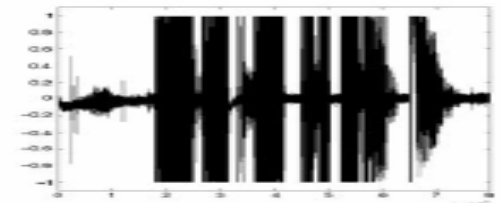
(c)



(d)



(e)

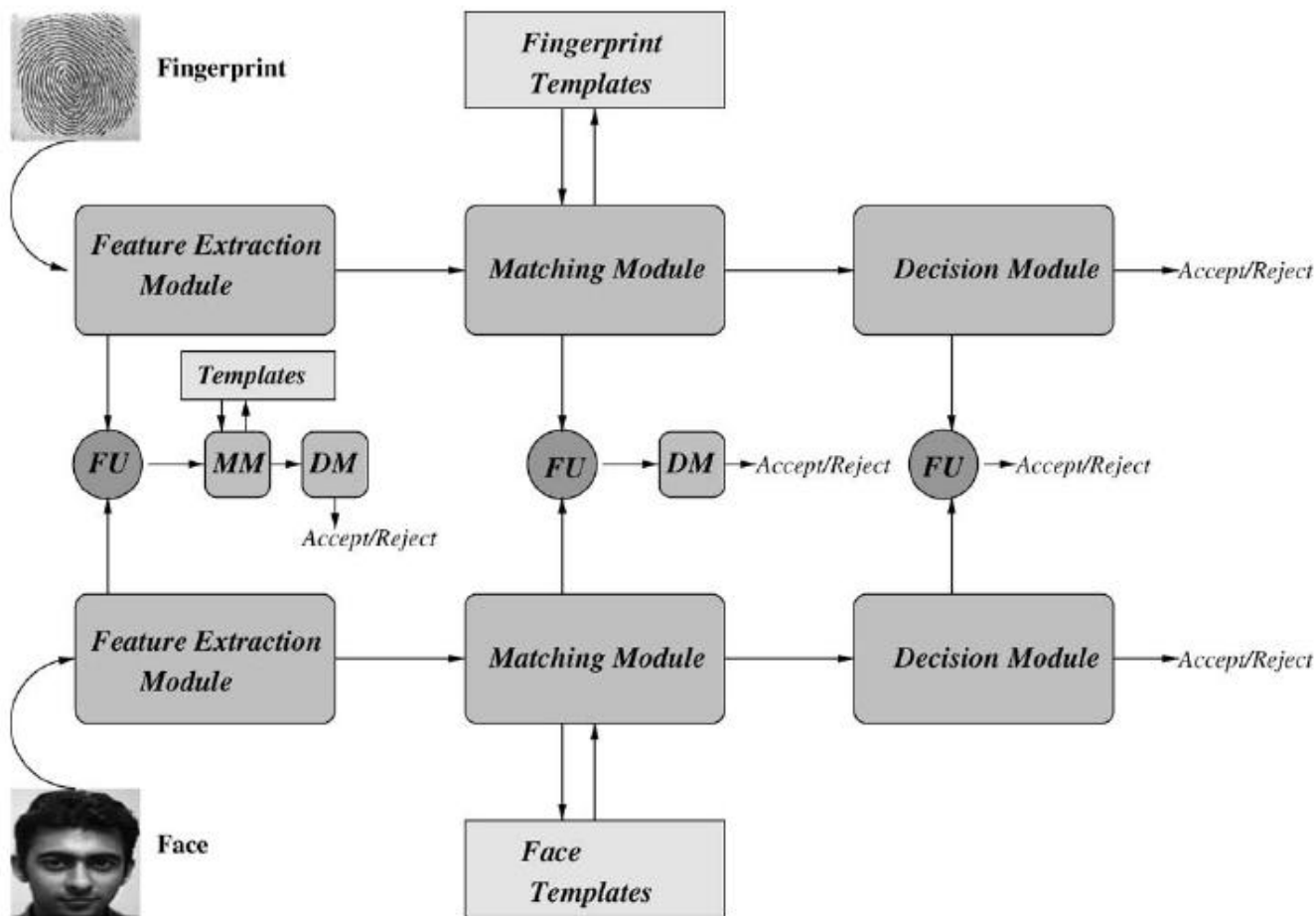


(f)

# Information Fusion in Biometrics

- Two modes:
  - Enrollment mode:
    - Biometrical information is obtained and stored in database with identification label
  - Authentication mode
    - Identification: Is he/she in the database?
    - Authentication: Is he/she the one claimed?
  
- Performance metrics:
  - Given a threshold
  - False accept rate
  - False reject rate

# Information Fusion in Biometric System



# Experiment Settings

- Multi-Modality:
  - Face, Fingerprint, Hand Geometry of
  - 50 persons
  - Independence among different biometrics is assumed.
- Classifier:
  - Distance-based
  - Scores have to be normalized
- Performance measure:
  - ROC curve

# Example of Fusion in Biometrics

- Face Verification



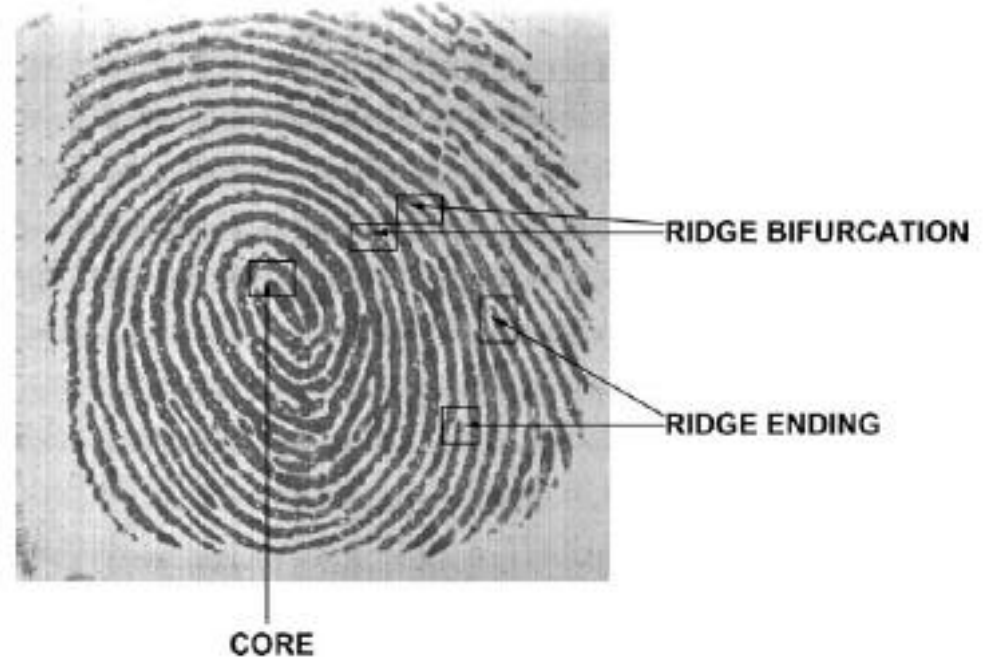
- Challenges:

- illumination, expression, head pose, background

# Example of Fusion in Biometrics

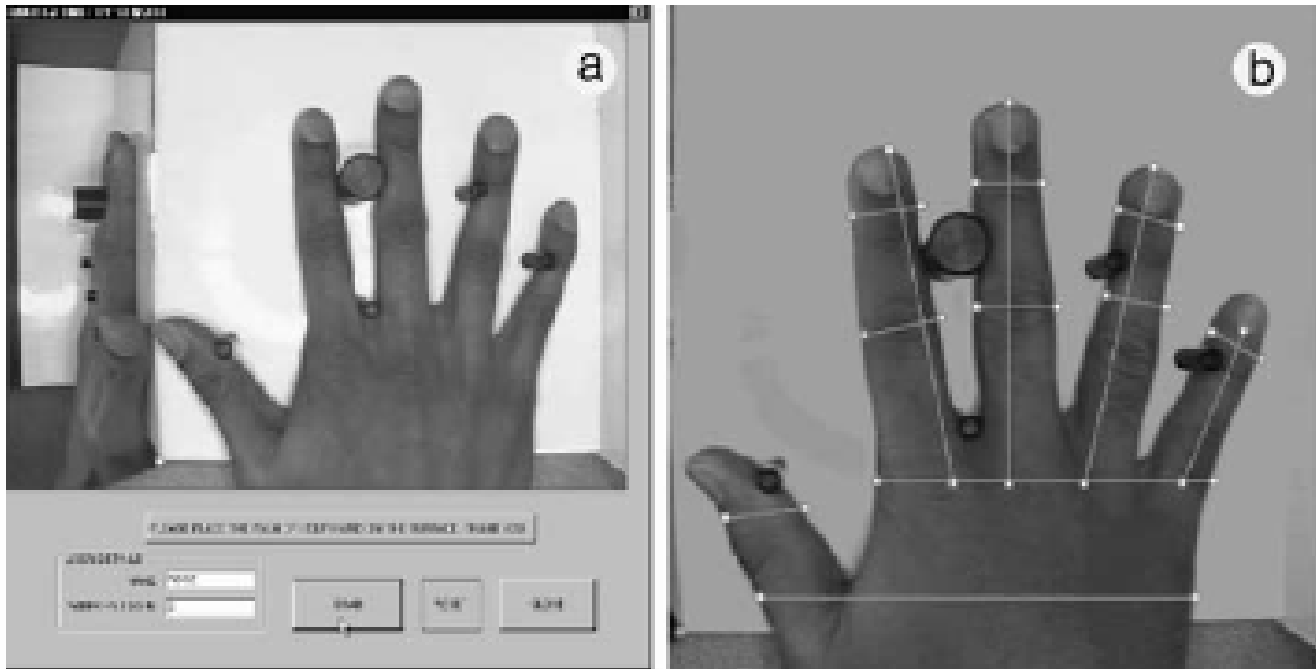
## ■ Fingerprint Verification

- (a) Preprocess to get ridge map
- (b) Detect ridge bifurcation and ending
- (c) Find Minutiae point
- (d) Align and find matching minutiae points between two finger prints

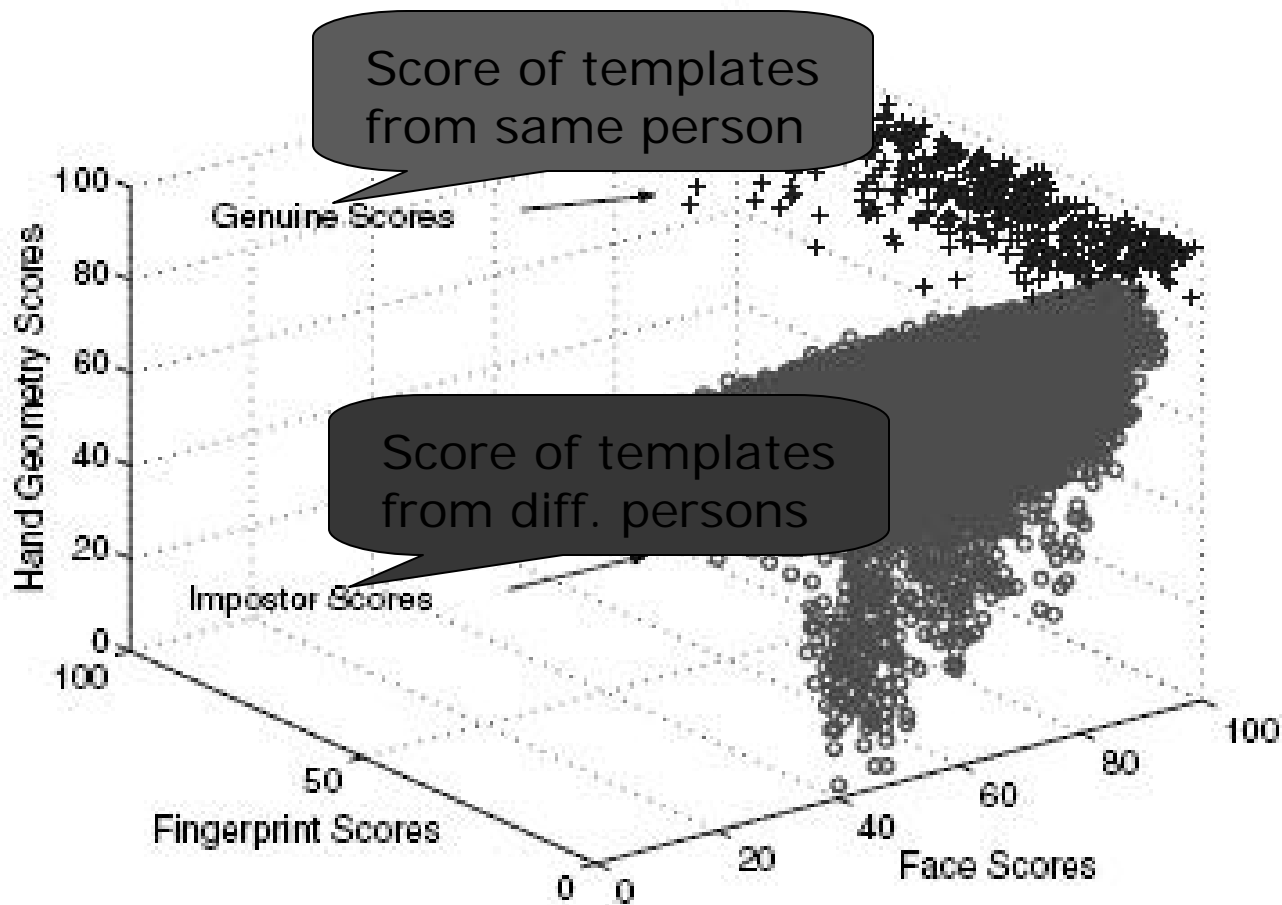


# Example of Fusion in Biometrics

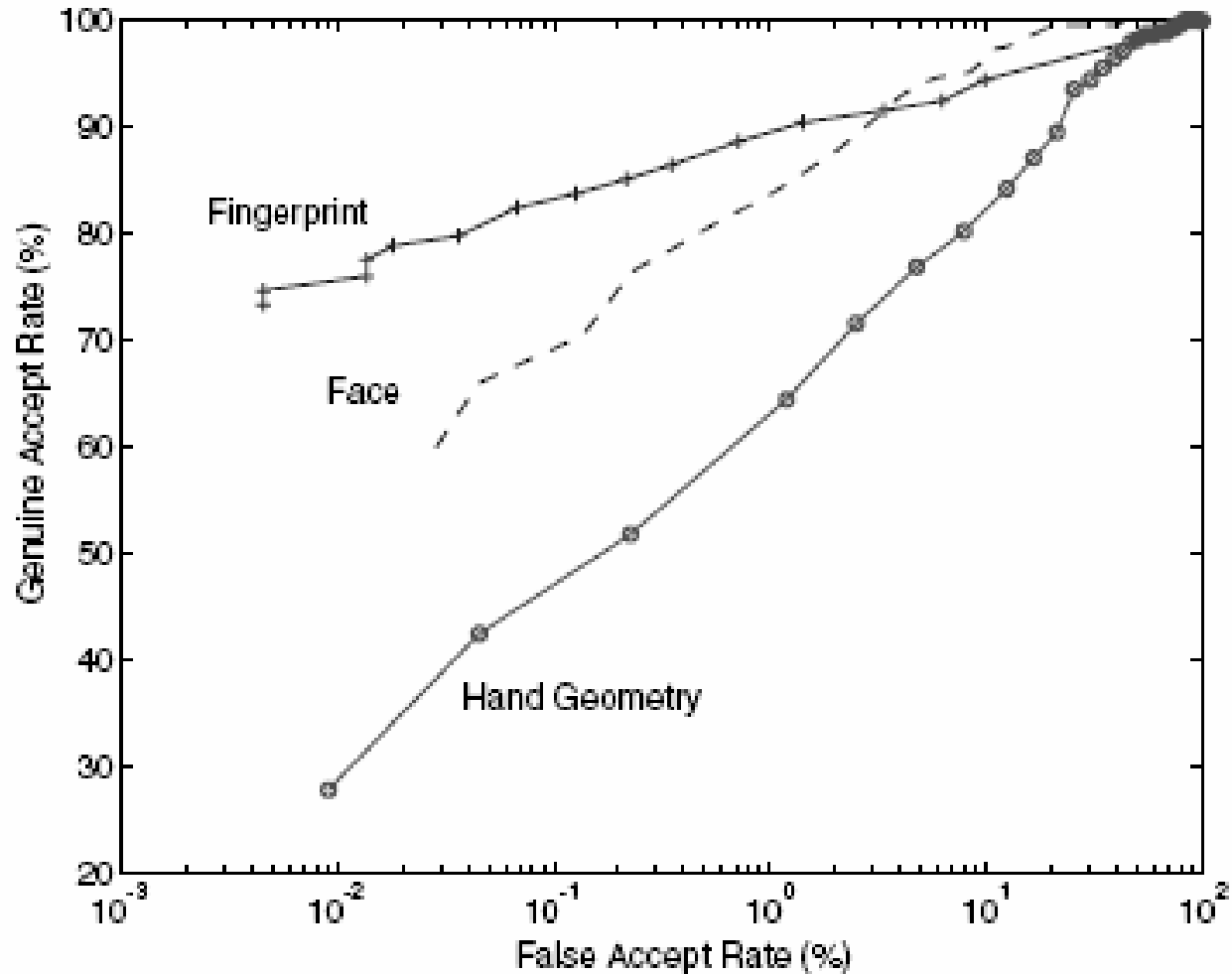
- Hand Geometry



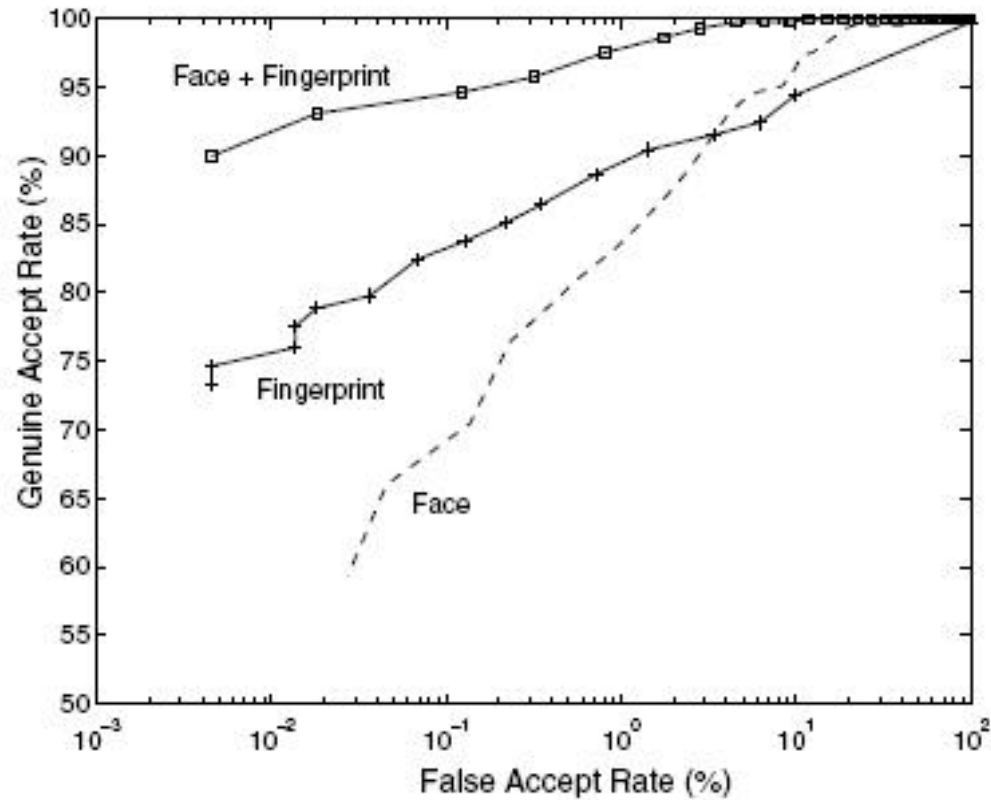
# Scores obtained from different biometric features



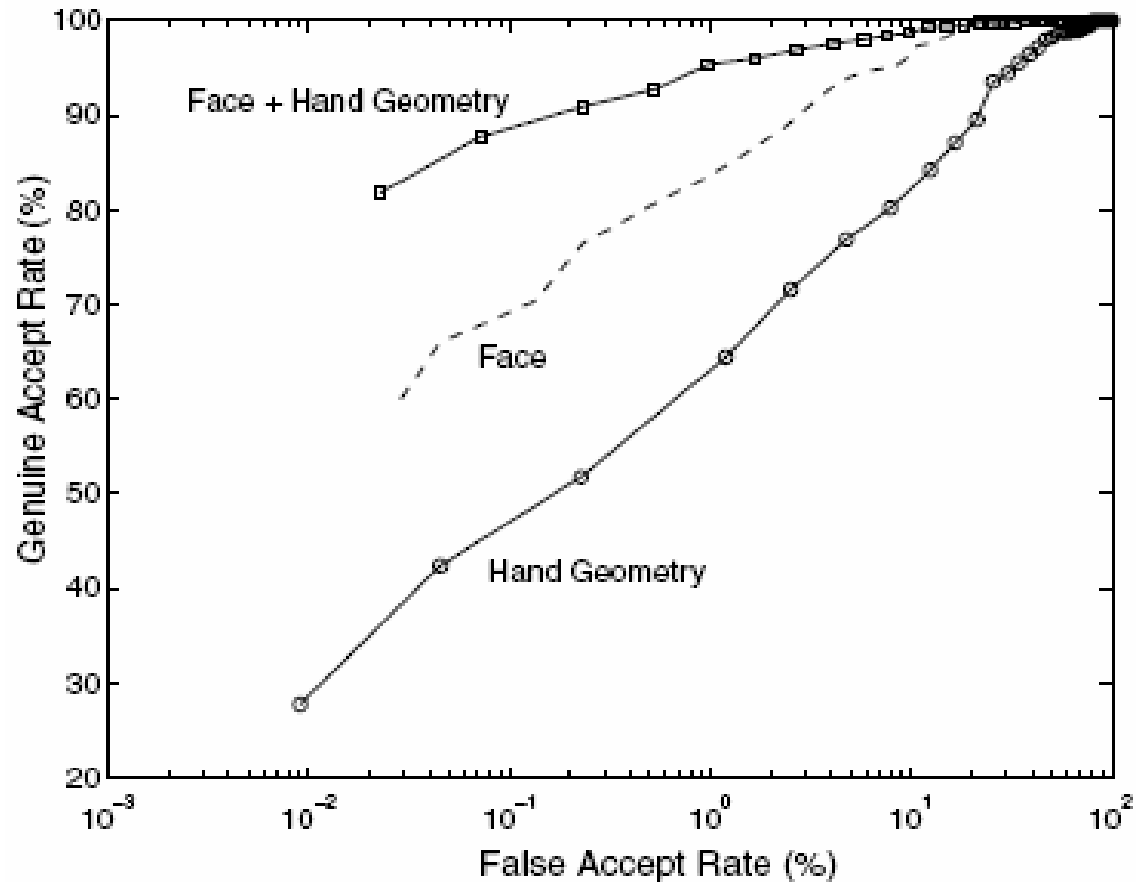
# Performance of Three Biometrics



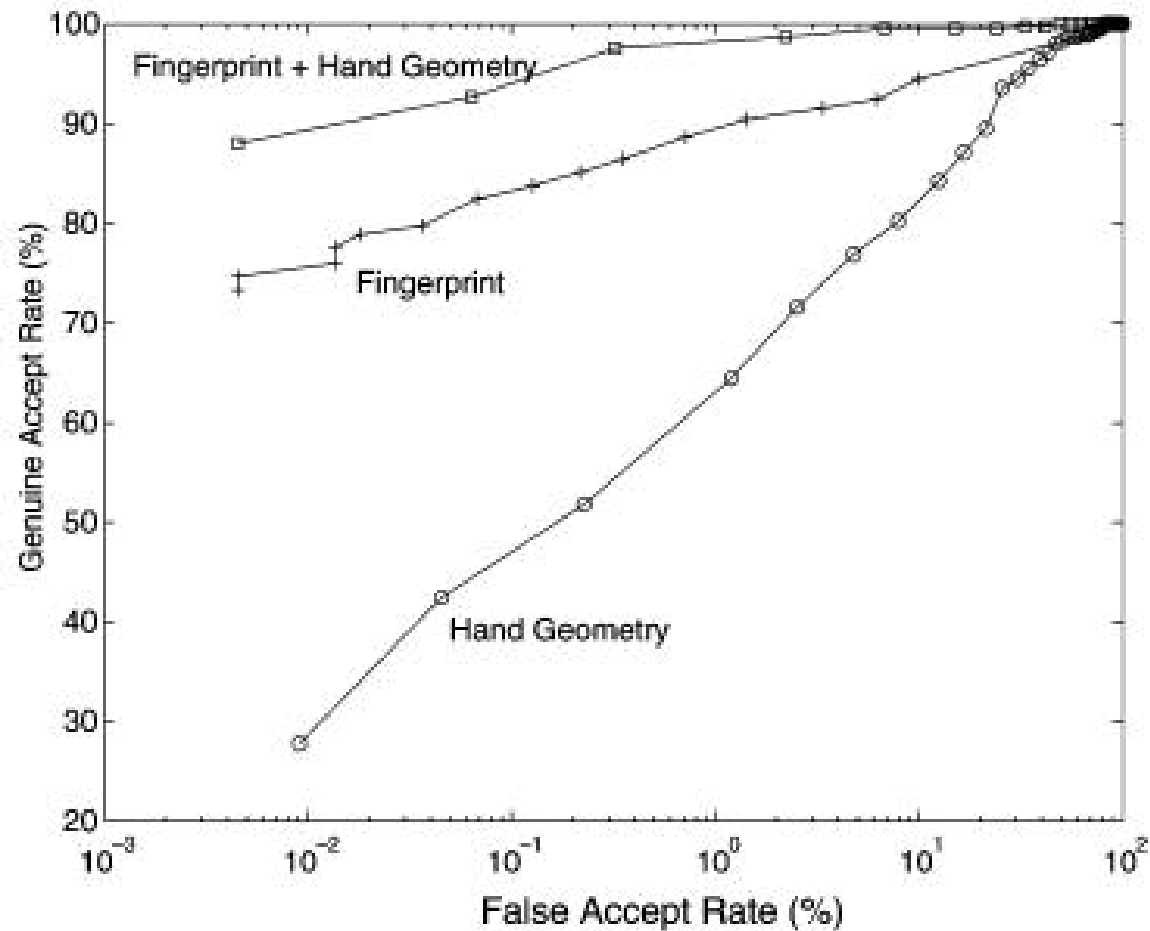
# Fusion of Face and Fingerprint



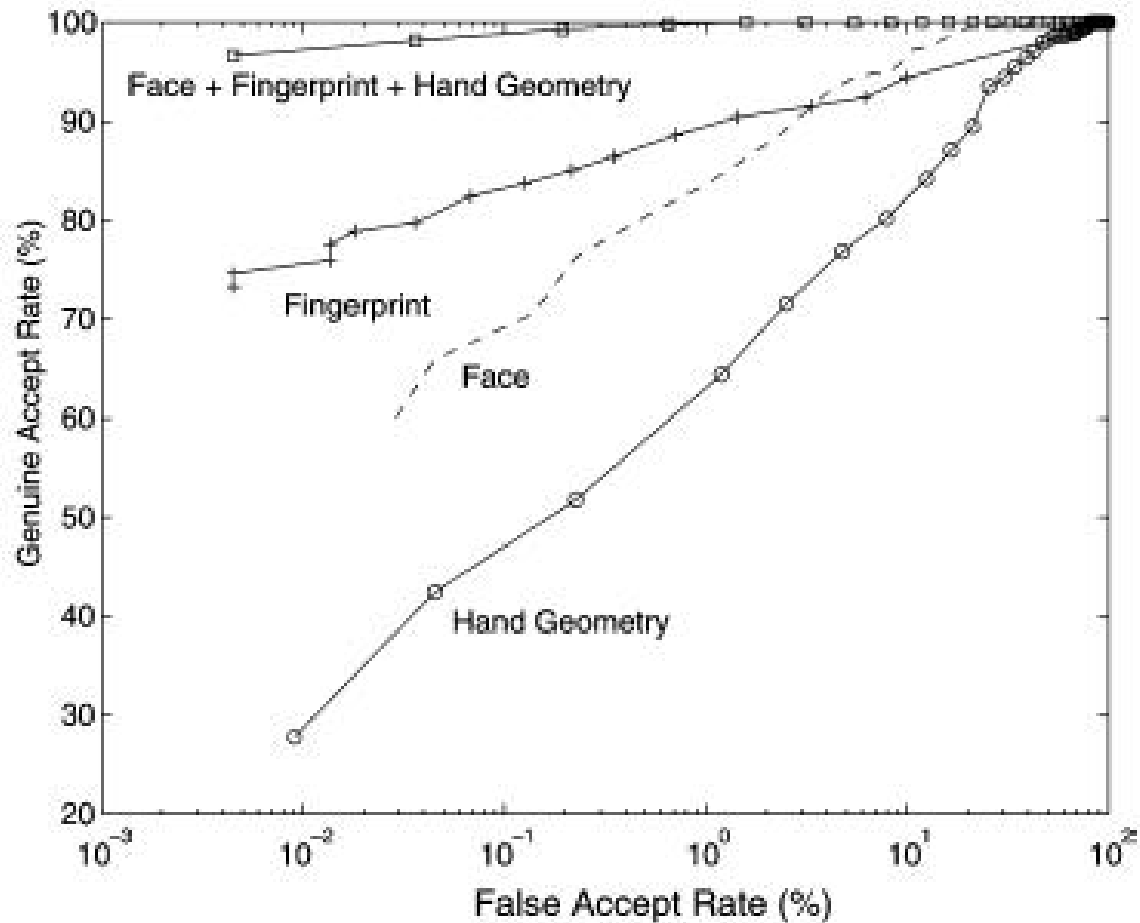
# Fusion of Face and Hand Geometry



# Fusion of Fingerprint and Hand Geometry



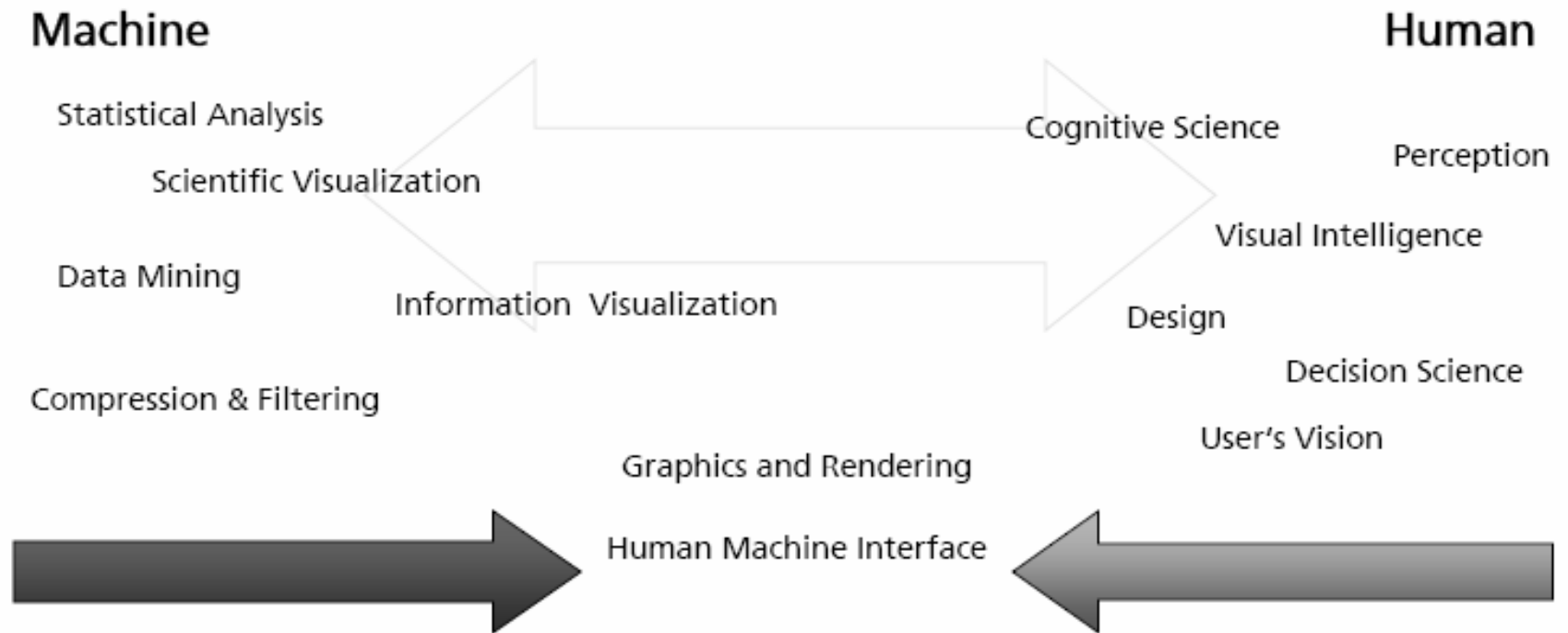
# Fusion of Three Biometrics



# Our Related Study: Interactive Boosting

- Motivation
  - Multi-modality and Multi-classifier system performs better in many MIR systems.
    - Compared with the single-source single-classifier system
  - Fusion schemes selection is difficult and may not be efficient.
    - Some fusion scheme cannot improve the performance
  - Human aid in a fusion process may provide more information gain in a guided direction.
- Idea: to introduce human-centered computing in information fusion

# Strength of Human and Machine



# Human Centered Computing and Information Fusion

- Human Centered Computing
  - Combining the strength of human and machine
    - Human strength: creativity, solving of ambiguity, strategies and principles, etc
    - System strength: accuracy without fatigue, data storage, deterministic processing
  - “Division of Labor” between human and labor
  - Enhancement of overall system performance
- Goal: To amplify and extend of human cognitive and perceptual ability in information fusion

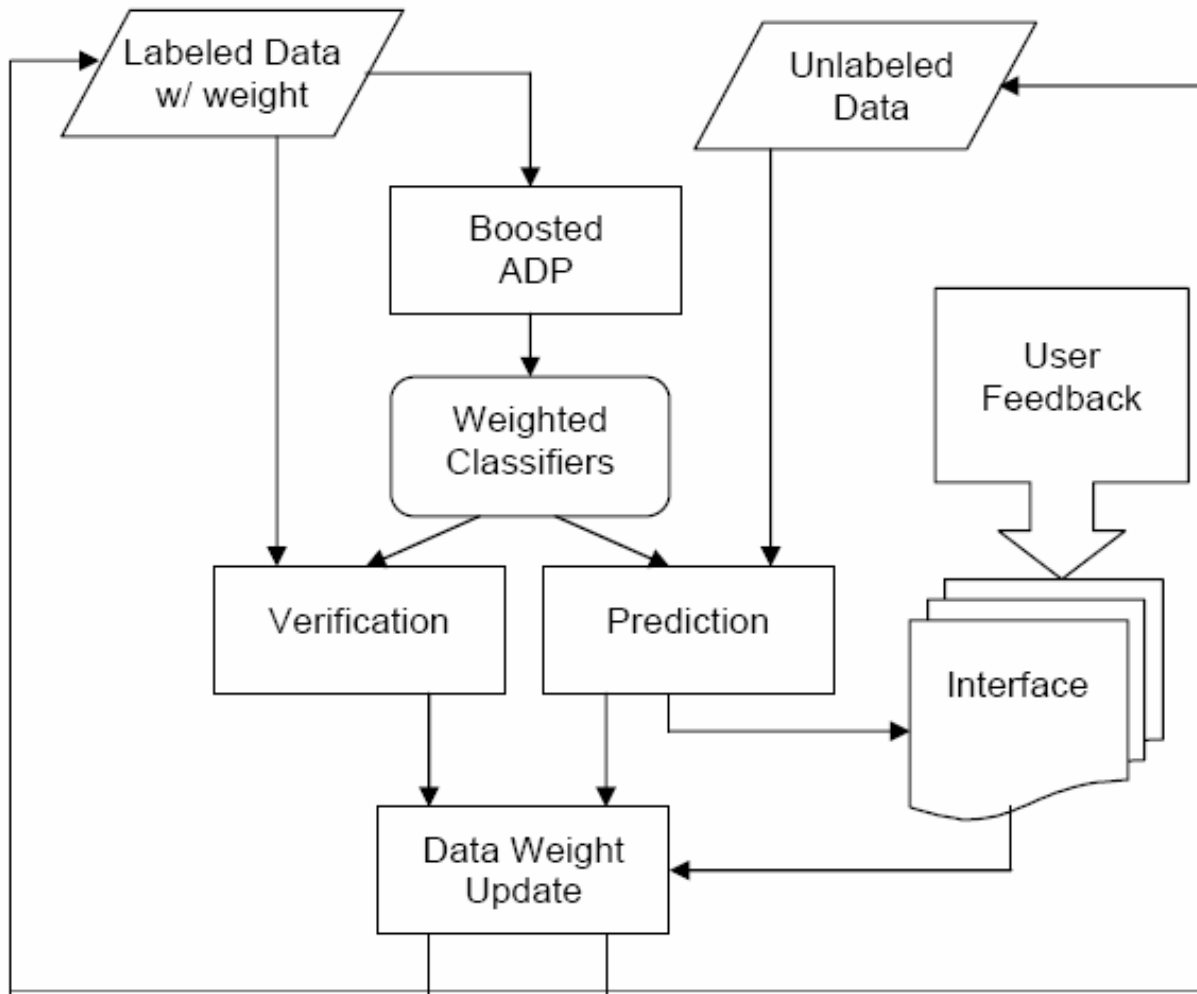
# Interactive Boosting

- Idea of Boosting:
  - Reinforcement training
  - Weight combination
- Basic Idea:
  - User relevance feedback on unlabeled data is incorporated into a iterative boosting process
    - Semi-supervised vs. fully-supervised boosting
  - Adaptive Discriminant Projection (ADP) is used to generate multiple representation of data [Yu and Tian, ICPR 2006].
  - Multiple-classifiers can be applied on the features generated by ADP.
    - K-NN, Bayesian, SVM

# Interactive Boosting

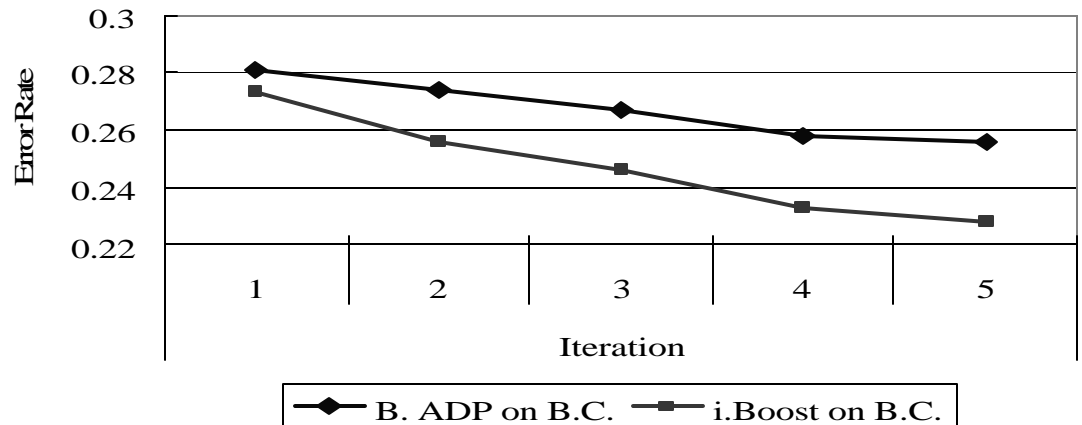
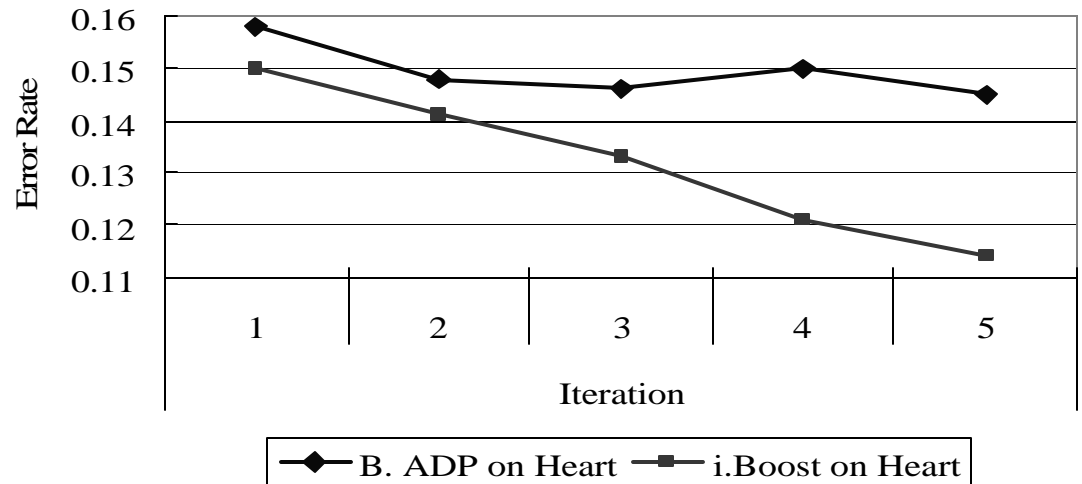
- *Step 1:* Train weak classifiers on the original labeled data set and assign weights to classifiers based on their performance.
- *Step 2:* Predict the labels of unlabelled data and present a subset of unlabeled data with their predicted labels to the user.
- *Step 3:* User gives feedback on the retrieved data.
- *Step 4:* Data obtained from user relevance feedback is added to construct a new labeled data set and removed from unlabeled data set.
- *Step 5:* The labeled data are weighted according to their predicted label correctness.
- *Step 6:* Go back to Step 1.

# Interactive Boosting



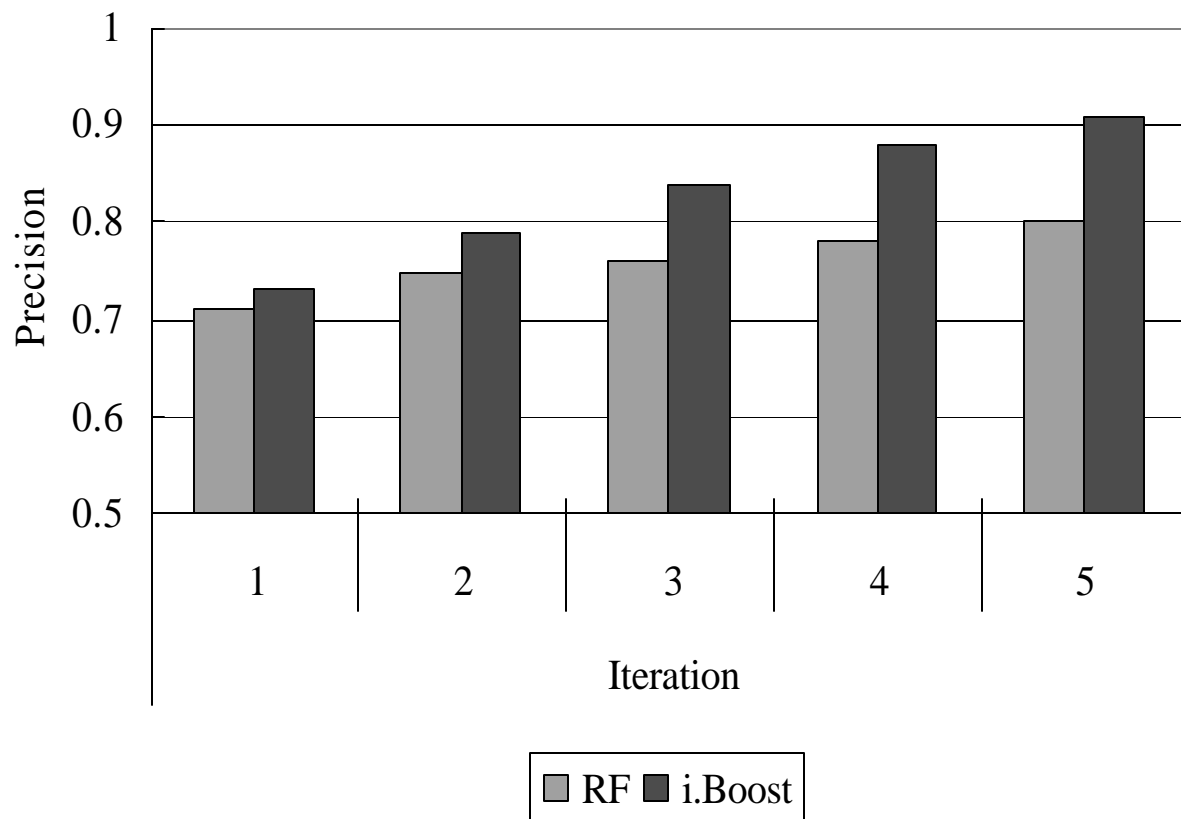
# Interactive Boosting: Experiment on Benchmark Data Sets

- Data set: Heart and Breast-Cancer (B.C.) from UCI database
- Performance of 5 iterations
- Relevance feedback of 5 data is obtained
- Compared with classic boosting+ADP
- Classifier: K-NN
- Lower overall prediction error rate is obtained on both datasets



# Interactive Boosting: Experiment on Corel Data Sets

- Data set: Corel
- Performance of 5 runs
- Relevance feedback of 5 data is obtained
- Compared with classic relevance feedback (RF)
- Classifier: K-NN
- Overall prediction precision is higher for i.Boost



# Discussion and Conclusion

- Information fusion fits the nature of multimedia retrieval system.
  - Multimedia data integrates information from multimodality.
- More robust and efficient retrieval can be achieved due to the redundant and complementary information.
- Multiple-classifier system can provide higher accuracy and handle extreme data set better.
- Human-centered computing may be a useful tool to further improve information fusion in MIR.

# Extra Slides

# The State-of-the-Arts

- Single Biometric Multiple Representation
- Single Biometric Multiple Matcher
- Multiple Biometric Fusion (Multimodality)

# Multi-source fusion system v.s. Single source system

A quantitative example of benefits of multisources fusion systems [L. A. Klein, Sensor and data Fusion Concepts and Applications.]

	False Alarm Rate	SNR	Detection Rate
Single MMW	0.000001	16 dB	0.7
Single MMW	0.000001	10 dB	0.27
Tri-sensor system (includes MMW radiometer and IR)	0.000001	16 dB	0.85
Tri-sensor system (includes MMW radiometer and IR)	0.000001	10 dB	0.63