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?

Multiple types of data carrying various types of information (redundant and complementary) → “Associated” or “Correlated” to the same object or event or behavior → So that estimation algorithms (mathematical techniques)—or—automated reasoning methods (artificial intelligence techniques) can produce better estimates (than based on any single type of data) → To improve estimates about those things → Multiple types of data Related to things of interest

These Basic Ideas are Transferable to Many Types of Problems
Complete Information Fusion System Model (JDL)

DATA FUSION DOMAIN

Level 0 Processing
- Sub-object Data
- Location, Movement, Identity of:
  - Parts
  - Vehicles
  - Organs; Tumors

Level 1 Processing
- Object Refinement
- Status or “Situation”:
  - Benign
  - Critical

Level 2 Processing
- Situation Refinement
- Adaptive Processing Logic, eg:
  - Active sensor control
  - Fusion process control

Level 3 Processing
- Impact Assessment

Level 4 Processing
- Process Refinement

Data Base Management System

Human Engrg:
- Decision-aiding systems
- Visualization/display

World State of Interest
- Intel. Manufacture System
- Defense/Surveillance system
- Patient-monitoring Systems

Human Computer Interaction

J. Yu, Data and Vision Seminar
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

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

- Query
- Feature Extraction
- Matching Module
- Decision Module
- Database
- User

J. Yu, Data and Vision Seminar
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

[Plenary Talk, L. O. Hall, ICPR 2006]
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, Neutral 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 work.
- No fusion scheme is guaranteed to be the best

<table>
<thead>
<tr>
<th>Fusion Method</th>
<th>Best Error Rate (Threshold ?)</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA (no fusion)</td>
<td>0.184 (0.6)</td>
<td>0.2779 (0.6)</td>
</tr>
<tr>
<td>LDA (no fusion)</td>
<td>0.159 (0.6)</td>
<td>0.25 (0.7)</td>
</tr>
<tr>
<td>&quot;AND&quot;</td>
<td>0.7175 (0.1)</td>
<td>0.3065 (0.1)</td>
</tr>
<tr>
<td>&quot;OR&quot;</td>
<td>0.6135 (0.1)</td>
<td>0.2981 (0.2)</td>
</tr>
<tr>
<td>&quot;Mean&quot;</td>
<td>0.155 (0.5)</td>
<td>0.2688 (0.7)</td>
</tr>
<tr>
<td>&quot;Max&quot;</td>
<td>0.1575 (0.7)</td>
<td>0.2682 (0.7)</td>
</tr>
<tr>
<td>&quot;Min&quot;</td>
<td>0.1735 (0.5)</td>
<td>0.2597 (0.6)</td>
</tr>
</tbody>
</table>
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
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

[A. K. Jain et al., PRL 2003]
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

J. Yu, Data and Vision Seminar
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 fingerprints

J. Yu, Data and Vision Seminar
Example of Fusion in Biometrics

- Hand Geometry

J. Yu, Data and Vision Seminar
Scores obtained from different biometric features

- Score of templates from same person
- Score of templates from diff. persons
Performance of Three Biometrics

![Graph comparing the performance of three biometrics: Fingerprint, Face, and Hand Geometry. The graph plots Genuine Accept Rate (%) against False Accept Rate (%) on a log-log scale.]
Fusion of Face and Fingerprint
Fusion of Face and Hand Geometry

J. Yu, Data and Vision Seminar
Fusion of Fingerprint and Hand Geometry

[Graph showing comparison of genuine acceptance rates and false accept rates for fingerprint, hand geometry, and their fusion]
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

Yu, Lu, Xu and Tian, ICASSP 2007
Strength of Human and Machine

Machine
- Statistical Analysis
- Scientific Visualization
- Data Mining
- Compression & Filtering

Human
- Cognitive Science
- Perception
- Visual Intelligence
- Design
- Decision Science
- User’s Vision

Information Visualization

Graphics and Rendering
Human Machine Interface

Yu, Lu, Xu and Tian, ICASSP 2007
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 an 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

Yu, Lu, Xu and Tian, ICASSP 2007
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

Yu, Lu, Xu and Tian, ICASSP 2007
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)
Muti-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.]

<table>
<thead>
<tr>
<th></th>
<th>False Alarm Rate</th>
<th>SNR</th>
<th>Detection Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single MMW</td>
<td>0.000001</td>
<td>16 dB</td>
<td>0.7</td>
</tr>
<tr>
<td>Single MMW</td>
<td>0.000001</td>
<td>10 dB</td>
<td>0.27</td>
</tr>
<tr>
<td>Tri-sensor system</td>
<td>0.000001</td>
<td>16 dB</td>
<td>0.85</td>
</tr>
<tr>
<td>(includes MMW radiometer and IR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tri-sensor system</td>
<td>0.000001</td>
<td>10 dB</td>
<td>0.63</td>
</tr>
<tr>
<td>(includes MMW radiometer and IR)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>