專題報告 Pattern Recognition

第三組 電機三 B91901146 高奕豪 2005/4/14

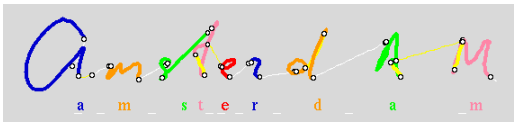
I. Introduction:

1. Definitions from the literature

"The assignment of a physical object or event to one of several pre-specified categories" –Duda and Hart

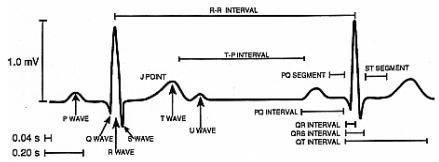
"Pattern Recognition is concerned with answering the question ***What is this?***" –Morse

2. Typical Examples:

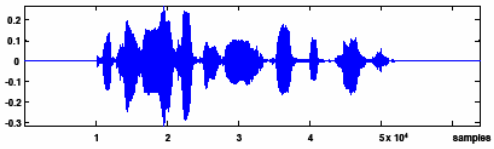


Machine vision: Visual inspection, ATR; Imaging device detects ground target; Classification into "friend" or "foe"

Character recognition: Automated mail sorting, processing bank checks; Scanner captures an image of the text; Image is converted into constituent characters



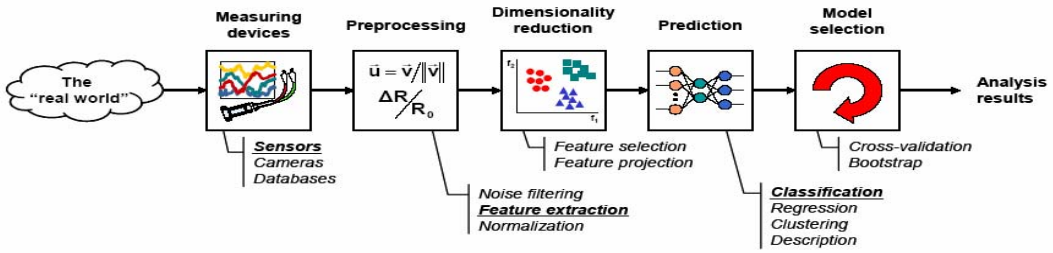
Computer aided diagnosis: Medical imaging, EEG, ECG signal analysis; Designed to assist (not replace) physicians; Example: X-ray mammography, 10-30% false negatives in x-ray mammograms;2/3 of these could be prevented with proper analysis



Speech recognition:

Human Computer Interaction, Universal Access; Microphone records acoustic signal; Speech signal is classified into phonemes and/or words

3. Pattern Recognition System



Sensing: This could be a computerized vision system or a microphone, among other types of sensor.

Segmentation: Still one of the biggest problems in pattern recognition. Segmentation occurs if you have to distinguish between two objects.

Feature Extraction: At this stage in the process we have to filter through our segmented sensed data. For example, in a speech recognition system where there is noise in the background, it would need to be filtered out.

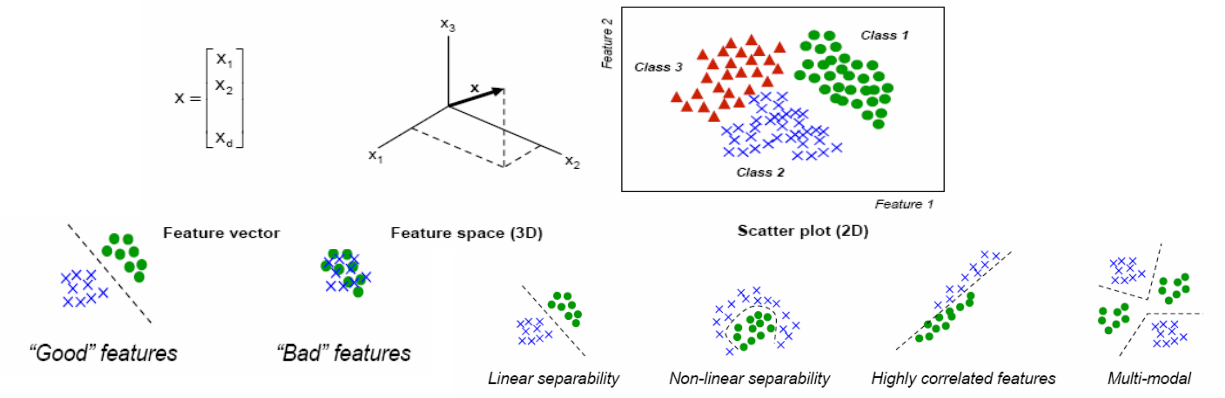
Classification: This is when the system classifies whatever was inputted into some

category. There are numerous methods of classification, and there has been no method that works universally with all problems.

Post-processing: A lot of different things can be done here after the object has been classified. This all depends on the demands of the system.

4. Pattern Features

Feature is any distinctive aspect, quality or characteristic; may be symbolic (i.e., color) or numeric (i.e., height). The combination of d features is represented as a d-dimensional column vector is called a ***feature vector*** ; The d-dimensional space defined by the feature vector is called the ***feature space;*** Objects are represented as points in feature space are called ***scatter plot.***

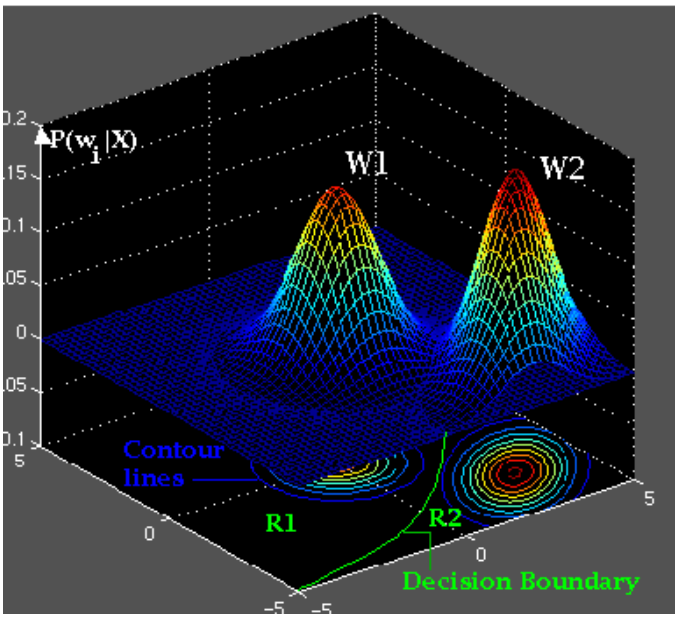


II. Approach

1. Bayesian Decision

Decide w_{org} *if* $p(x|w_{org})P(w_{org})>p(x|w_{app})P(w_{app})$,

otherwise decide w_{app}



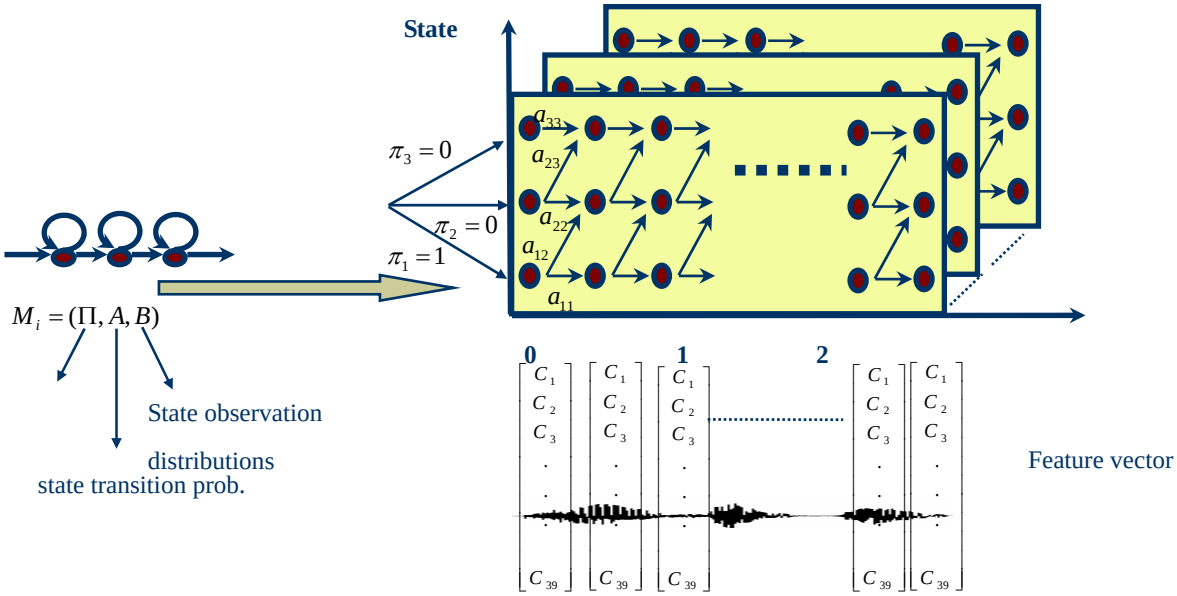
When any decision rule is applied to the d-dimentional feature space R^d, the result is that the space is split up into c decision regions R_1, ..., R_c. In general, if x lies in decision region R_i then it means that the pattern classifier selected the function $g_i(x)$ to be the maximum of all the discriminant functions. The decision regions are any subset of the space R^d. For example, if the feature vector is a 2-dimentional vector, then the discriminant functions $g_i(\mathbf{x})$ will be functions of 2 variables and will be mapped in 3-D. The decision regions for this case will be subsets of the x-y plane.

2. Hidden Markov Model

The Hidden Markov Model is a finite set of *states*, each of which is associated with a (generally multidimensional) probability distribution. Transitions among the states are governed by a set of probabilities called *transition probabilities.* In a particular state an outcome or *observation* can be generated, according to the associated probability

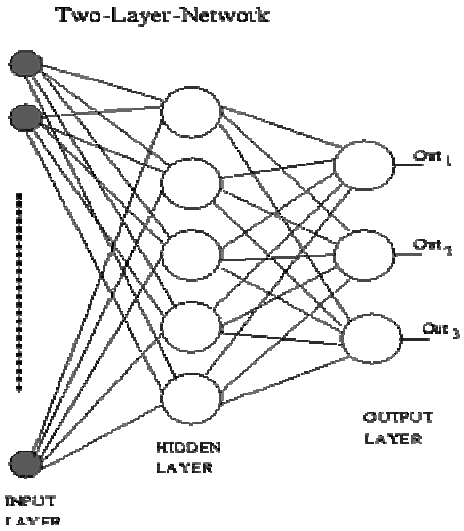
distribution. It is only the outcome, not the state visible to an external observer and therefore states are ``hidden" to the outside.

Viterbi algorithm:



3. Multilayer Neural Network

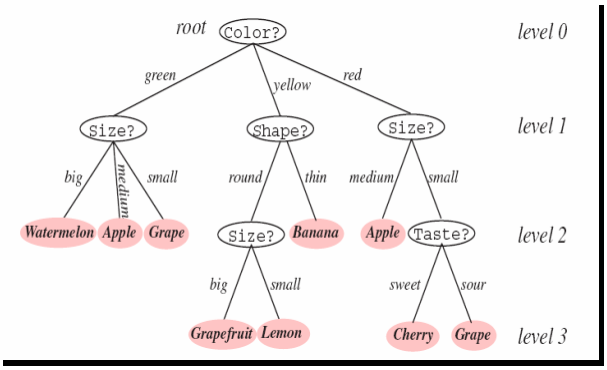
Multilayer networks solve the classification problem for nonlinear sets by employing *hidden layers*, whose neurons are not directly connected to the output. The additional hidden layers can be interpreted geometrically as additional hyper-planes, which enhance the separation capacity of the network.



4. Decision Tree

Robust to noisy data and capable of learning disjunctive expressions, **decision tree learning**, a method for approximating discrete-valued target functions, is one of the most widely used and practical methods for inductive inference. Decision tree learning is generally best suited to problems with the following characteristics:

- Instances are represented by attribute-value pairs.
- The target function has discrete output values.
- Disjunctive descriptions may be required.
- The training data may contain errors.
- The training data may contain missing attribute values.



III. Modern Applications

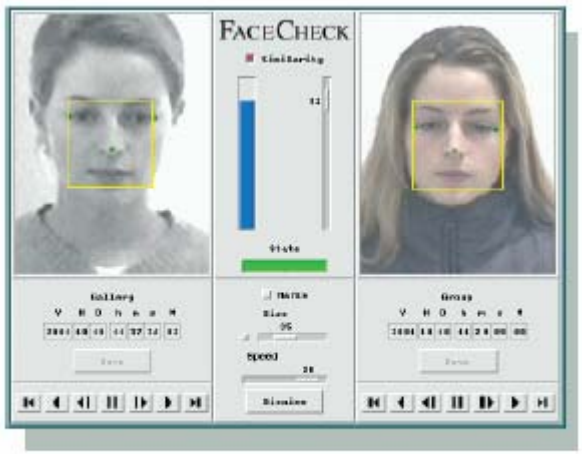
1. Face Recognition: C-VIS

C-VIS GmbH was founded in 1992 by a team of scientists from the internationally renowned Institute for Neurocomputing at Bochum University in Germany.

C-VIS has taken biometrics out of the lab and into the workplace by translating vision technologies into user-friendly applications that meet real-world needs. C-VIS products are currently in use for video surveillance, security systems, airports, casinos, etc.

In recording mode, the FACESNAP RECORDER screen simultaneously shows the live camera shot and the latest sequence of captured facial images.

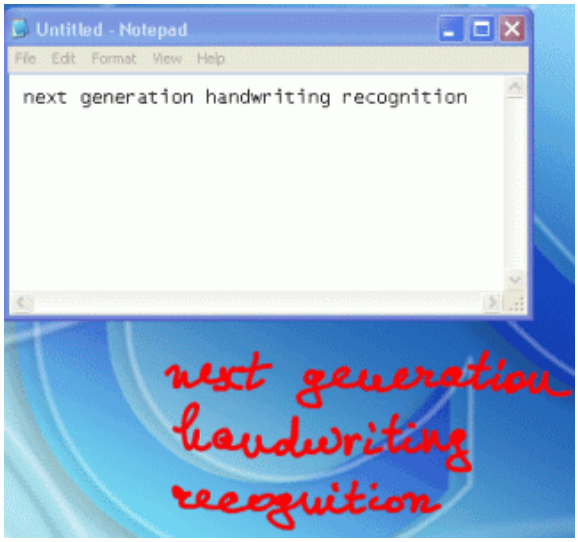
In search mode, the user can define image groups and search criteria for face identification. Automatic functions for image pattern recognition can also be used to search the image database.



In the comparison window, images can be viewed, analysed and compared in detail.

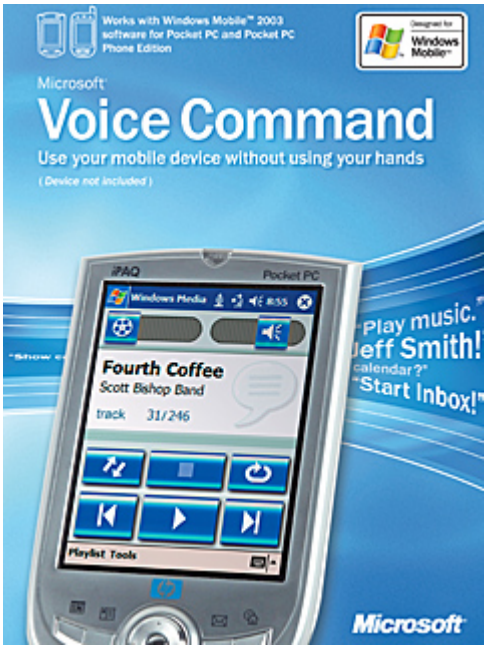
2. Handwriting Recognition: EverNote

The current version of riteScript runs on EverNote's servers and is designed as a component of the upcoming Web and wireless services for online forms recognition and for advanced processing of notes, taken on pen-enabled handhelds, Webpads, intelligent pens and other devices with handwritten input and transmitted via wireless networks. The interaction of client side applications with riteScript back-end is performed via an XML-based data transmission



protocol, which includes electronic ink encoding and compression.

3. Speech Recognition: Microsoft Voice Command



Winner of the Pocket PC Magazine Best Software Award 2004 for Speech Applications!! Voice Command transforms your Pocket PC into your own virtual personal assistant, letting you use your voice to look up contacts and place phone calls, get calendar information, play and control your music, and get key device status and information. Voice Command makes it easier and more convenient than ever to take your digital lifestyle with you wherever you go. The state-of-the-art speech technology means you will never have to pre-record important phone numbers or use difficult commands to access the information you need. Simply select one of the many commands that is most natural to you and let Voice Command do the rest.

4. Fingerprint Recognition: Microsoft Mouse Reader

A mouse that recognizes you. This stylish wireless optical mouse offers the Fingerprint Reader to eliminate password hassles—now you can log on to Web sites and your computer with the touch of a finger. Tilt wheel technology makes navigation easy, and more than six months of battery life lets you stay productive

IV. 心得感想

Pattern Recognition 一直是我在計算機領域中最感興趣的課題，又因爲本學期修習李琳山教授數位語音專題有所接觸，使我決定藉這次專題報告的機會去整理一些該領域的其它方法與應用。我覺得這領域最有趣的地方在於：爲什麼人類可以輕易做到的事，電腦卻要做得這麼辛苦？我有一個夢，有一天，人類能在自己的大腦裡找出 Pattern Recognition 的解答，並且套入電腦實際應用。值得一提的是，世界計算機科學界的超級龍頭 MIT，最近大力投入神經科學的研究，就連校長也換上該領域的專家。這是否正暗示著：神經認知科學與計算機科學的下一波革命已密不可分了？

Reference (*in addition to those listed in my powerpoint!)*

http://www2.cs.uregina.ca/~hamilton/courses/831/notes/ml/dtrees/4_dtrees1.html
http://www.microsoft.com/windowsmobile/downloads/voicecommand/default.mspx
http://www.microsoft.com/hardware/mouseandkeyboard/productdetails.aspx?pid=035
http://www.cs.mcgill.ca/~mcleish/644/main.html