Comp 254, Image Processing and Analysis
Spring 2002, Prof. Guido Gerig

Assignment 1: Statistical Pattern Recognition

Out: Wednesday Jan-29-2002
Due: Wednesday Feb-12-2002 (theoretical part), Feb-14-2002 (practical part)

Required Readings: Duda and Hart, Chapter 2

I. Theoretical Problems

Problem 1: Normal Density Functions

Consider a 1-D two-category problem with $p(x|\omega_i)$ being a normal density function

$$N(\mu_i, \sigma^2), \qquad P(\omega_1) \;=\; P(\omega_2)$$

Determine the minimal error probability P_e. Describe P_e if

$$\frac{|\mu_2 - \mu_1|}{\sigma} \;\rightarrow\; \infty$$

$$\frac{1}{\sqrt{2\,\pi}} \int_a^{\infty} e^{\frac{-t^2}{2}} dt \;\leq\; \frac{1}{\sqrt{2\,\pi} \cdot a} e^{\frac{-a^2}{2}}$$

Problem 2: Decision with minimum risk

In many pattern classification problems one has the option either to assign the pattern to one of c classes, or to *reject* it as being unrecognizable. If the cost for rejects is not too high, rejection may be a desirable action.

$$\lambda(\alpha_i|w_j) \;=\; \begin{cases} 0, & i = j \qquad i, j = 1, \ldots, c \\ \lambda_r, & i = c + 1 \\ \lambda_s, & \text{otherwise.} \end{cases}$$

where λ_r is the loss incurred for choosing the $(c+1)$th action of rejection, and λ_s the loss incurred for making a substitution error. Show that the minimum risk is obtained if we decide

a) Show that the minimum risk for a decision on w_i is obtained if we decide

$$P(w_i|x) \;\geq\; P(w_j|x) \quad \forall j, \quad P(w_i|x) \geq 1 - \frac{\lambda_r}{\lambda_s},$$

and reject otherwise.

b) What is the decision with minimum risk for $\lambda_r \;=\; 0$?.

c) What is the decision with minimum risk for $\lambda_r > \lambda_s$?

d) Show that the following discriminant functions are optimal:

$$g_i(x) = \begin{cases} p(x|w_i)P(w_i) & i = 1, \cdots, c \\ \frac{\lambda_s - \lambda_r}{\lambda_s} \sum_{j=1}^{c} p(x|w_j)P(wj) & i = c+1 \end{cases}$$

e) Sketch these discriminant functions and the decision regions for the two-category one-dimensional case with $p(x|w_1) \approx N(1,1)$, $p(x|w_2) \approx N(-1,1)$, $P(w_1) = P(w_2) = 0.5$, and $\frac{\lambda_r}{\lambda_s} = 1/4$.

I. Practical Problems

Problem 3: Multivariate Classification

Double-echo MRI images of the human head allow a reasonable segmentation with pixelwise multivariate statistical classification. In fact, this is the most common segmentation method applied in medical image analysis labs, e.g. at Duke hospital. The feature space is two-dimensional, and the feature vector $\underline{v}$ is the pair of intensity values $\underline{v} = [v_1(x,y), v_2(x,y)]$ at each pixel location (x,y).
Categories which can be discriminated most often are ω_1 = brain gray matter, ω_2 = brain white matter, and ω_3 =liquid-filled cerebro-spinal fluid spaces. The class-specific probability density distributions $p(\underline{v}|\omega_i)$ can be described by 2-D normal density functions. In supervised classification, a user defines training regions to obtain estimates of the $p(\underline{v}|\omega_i) = N(\underline{\mu}, \Sigma)$ for each category. The a priori probabilities $P(\omega_i)$ are obtained using anatomical knowledge about soft tissue or by rough estimates.

- Input images are the image pairs of a patient (mri-echo1, mri-echo2) and of an MRI phantom (phan8_1/phan8_2).
- Calculate the 2-D histograms (joint histogram, scatterplot) and display them as images. (you can choose the module JointHistogram in /usr/Image).
- Manually choose a typical region of interest (ROI) for the 3 categories white and gray matter, and fluid. Calculate the 2-D probability density functions $p(\underline{v}|\omega_i) = N(\underline{\mu}, \Sigma)$ for each category by writing a program that reads in the two images and calculates the 2-D statistics within a given ROI.
- Give rough estimates of the three $P(\omega_i)$ for the tissue categories that you observe in the image pair by visual inspection.
- Calculate discriminant functions for the normal density (see Duda and Hart p. 24ff) by using the simplest case 1 where the individual covariance matrices are approximated by the variance σ^2 times the identity matrix I. Discuss two cases, first a minimum error-rate classification (MAP) and second a maximum likelyhood classification (ML).
- Classify the images using MAP and ML classification and represent the classification map as an image with labeled pixels.
  Alternatively to applying the set of discriminant functions to each image pixel to find the most probable category you can also think of classifying each possible feature vector

in the 2-D feature space (which nicely shows the decision regions) and using it as a 2-D look-up-table (Feel free to choose the option you like most for your program).

** Voluntary, only for sportsmen: Classify the images using case 2 or case 3 as discussed in Duda and Hart pages 26-31.

- Write a short discussion to describe a) the difference between MAP and ML classification and b) your opinion of the appropriateness of the simplification (case 1) used for calculating discrimination functions.