

么性质和特点才能更好地发挥模式识别技术的优势,这也许是模式识别理论中应该研究的问题之一,遗憾的是目前这方面的工作尚未得到充分研究。

预处理一般有两种情况:一是使数据的质量更好,比如用一些数字信号处理的方法去除信号中的噪声,或者对一幅模糊的图像进行图像增强等,需要注意的是要确保这种预处理是有利于后期的模式识别工作的,另一种预处理相对没有得到足够的重视,这就是样本集的预处理,比如样本集中野值的剔除、类别的合并或分裂等。这一工作一般可以根据领域的专门知识进行,也可以采用模式识别中的一些技术,比如在进行后续工作之前先对样本集进行一次聚类分析(参见第 10 章)。

3. 特征提取和选择

这是本章和上一章讨论的内容,指在已经得到数据样本之后如何用数学的办法对数据进行必要的变换和选择,使所得的特征更易于分类。应当指出,从这两章的介绍可以看出,虽然人们对特征提取和选择问题已经进行了很多研究,但这仍然是一个相对不成熟的领域,多数方法仍具有很大的经验性。

4. 分类器设计

本书的大多数章节都用来讨论这一内容。前面几章已经介绍了在有已知类别标号的训练样本时的分类器设计方法,即监督学习方法。下一章将介绍样本类别标号未知时的分类器设计方法,即非监督学习方法,此时用来设计分类器的样本集可能就是设计好分类器后要进行分类的样本集。

5. 分类及结果解释

对于监督模式识别情况,这一阶段就是用设计好的分类器对新的或者类别标号未知的样本进行分类;而对于非监督模式识别情况,则往往需要将得到的分类(聚类)进行解释,赋予各类一定的专业含义,同时也判断所得分类是否符合问题需要。有些情况下,可能还需要根据所用的分类器给出为什么把某个未知样本或新样本划分为某一类的解释,以利于人利用这些分类结果进行后续的决策。

习 题

9.1 若有下列两类样本集:

ω_1	ω_2
$x_1^1 = (0, 0, 0)^T$	$x_1^2 = (0, 0, 1)^T$
$x_2^1 = (1, 0, 0)^T$	$x_2^2 = (0, 1, 0)^T$
$x_3^1 = (1, 0, 1)^T$	$x_3^2 = (0, 1, 1)^T$
$x_4^1 = (1, 1, 0)^T$	$x_4^2 = (1, 1, 1)^T$

用 K-L 变换,分别把特征空间维数降到 $d=2$ 和 $d=1$ 并用图画出样本在该特征空间中的位置。

9.2 对上列样本,分别用(9.5),(9.6)和(9.7)三节中所述方法降低特征空间的维数。

9.3 令 Σ_i 和 P_i 分别是 ω_i 类($i=1,2$)的协方差矩阵和先验概率。假定对数据进行了白化变换,即使 $B^T S_w B = I$ 。这里 $S_w = \sum_i P_i \Sigma_i$, I 是单位阵。证明矩阵 $P_1 B^T \Sigma_1 B$ 和 $P_2 B^T \Sigma_2 B$ 所产生的 K-L 坐标轴是相同的。若用 Λ_i 表示矩阵 $P_i B^T \Sigma_i B$ 的本征值矩阵,求证

$$\Lambda_1 = I - \Lambda_2$$