動態時軸校正法路徑限制的研究－以語音辨識為例

羅琪, 尤廉凱

應用數學學系

工學院

chilo@chu.edu.tw

摘要

In recent years, classification and clustering of time series data have become one of the topic of great interest within the database/data mining community. Most method in classification and clustering are needed to calculate the distance measure between two series. Several popular measures of distance are Euclidean distance, squared Euclidean distance, city block distance and Minkowski distance. However, when two time series are similar but they are not aligned in the time axis, then the above distance measures are very sensitive to distortion in time axis. Although using Euclidean and other distance measures in classification and cluster analysis can make calculation efficient, we have to face the evidence of higher error rate.

The problem of distortion in the time axis can be solved by using Dynamic Time Warping (DTW) which is a method based on the dynamic programming (DP). This research will use the kth nearest neighbor rule, proposed by Fix and Hodges (1951), to discriminate and classify the speech pattern of several people speaking Chinese "open the door, sesame". The kth nearest neighbor rule is to compute the distance from an observation to all the other points then assign this observation to the group that has the highest proportion among the kth nearest neighbors. In this research, both Euclidean distance and DTW distances by various path constraints are used as the dissimilarity measures. Then the kth nearest neighbor classification rule is used to classify the speech pattern of several people speaking Chinese "open the door, sesame" and the apparent error rate and average error rate are used to compare the classification results. Finally, a discussion will be made on the effect of global path constraints to the error rate and computation speed.