Statistical spectral estimation and discriminant analysis of the bivariate time series-an application to the voice recognition on Chinese speaking of several people

羅琪, 李信賢
Applied Statistics
Management
chilo@chu.edu.tw

Abstract

The extension of classical pattern-recognition techniques to experimental time series data is a problem of great practical interest. An important application in engineering is to the problem of discriminating between various speech patterns. Throughout the engineering literature, most approaches assumed very specific Gaussian additive signal and noise models and then developed the discriminant criteria to minimize errors. In general, this requires that we assume prior knowledge of the signal waveforms and spectra under each of the hypotheses, so that discriminant functions can be calculated for an observed time series. In this paper, the spectra of three speakers on speaking Chinese words "open the door, sesame" are assumed to be unknown. The estimation and hypothesis-testing problems are formulated in terms of sample spectral densities with sample approximate distributions. Finally, frequency domain approximations are used to the optimum discriminant functions to identify the speech patterns of three speakers on speaking Chinese words "open the door, sesame". Since each person has left and right two series (bivariate time series) in one record, the estimation of cross spectrum, phase spectrum and coherency between two series are also considered. The estimated quadratic discriminant function is quite well in classifying the three speakers' voices because the average apparent error rate of 84 3-people groups of 9 speakers is only about 2.524%. This indicates that it is a good way to apply the frequency domain discriminant analysis of time series to do the speech recognition on Chinese speaking of three speakers.

Keyword：bivariate time series, spectral estimation, discriminant analysis, speech recognition