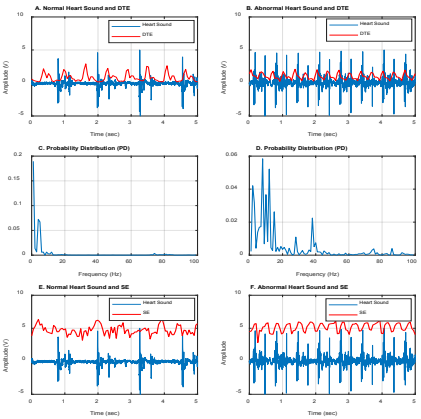
Feature Extraction and Classification of Heart Sounds Signals Based on Time-Dependent Entropy and Spectral Entropy Estimation

Rosario Ríos-Prado, Álvaro Anzueto-Ríos, Blanca Tovar-Corona
Instituto Politécnico Nacional (UPIITA-IPN)
Mexico City, Mexico

Nowadays, signal processing techniques can be used to extract features that provide auxiliary information to assist the detection of abnormalities in heart valves, mainly where medical specialists are not available. This is carried out using phonocardiographic recordings where the mechanical activity can be analyzed.



Estimation of Time-Dependent Entropy, Probability Distribution and Spectral Entropy in heart sound signals.

In this paper, two entropy methods based on Shannon Entropy are exploited, the Time-Dependent Entropy and the Spectral Entropy, calculated in the time domain and frequency domain, respectively. The two calculated entropies together with the Probability Distribution were obtained from a database, as shown in the figure, that contains simultaneous recordings from the four main auscultation areas with two main purposes: 1) to test if the use of the channels increases de probability of detecting the abnormality in any of the heart valves and, 2) to compare the results per area respect to signals randomly selected from the four areas. The three parameters obtained from 20 randomly selected signals of the database were used as input features for the K-Nearest Neighbor classifier to present the classification results, obtaining accuracies of 90% and 80% for pathologic and normal sounds classification, respectively.

Finally, the parameters calculated from all the database were separated and presented for each of the auscultation areas in 3D-distributions where a visible separability is shown. Results suggest that some variations in the sounds associated to valve disfunction is reflected in the entropy values, where the greatest changes occur in the main sounds, S1 and S2, in the time series and the highest values were associated with the probability of occurrence of the data in the frequency domain. In addition, results show that comparing the four areas might improve the correct classification suggesting that the information by area is different depending on the heart condition.