
**Abstract**

Speech signals convey information from many sources, but not all information sources are relevant to describe speaker identity. In fact, speech is affected by spurious events, artifacts (mouth breath, lip clicks), and noise (channel and background). Such unwanted information sources are shared by speakers and do not contribute in distinguishing between them. Furthermore, in most cases, training data are collected from different environments and it is of great importance that such data convey relevant joint information. This paper discusses a method for removing unwanted information in order to build more robust models. Two criteria are used to extract relevant information from the speech signal: the first criterion, which we call self-information criterion, is used to extract relevant information from data collected from a given environment; the second is called joint information criterion, and it is used when collected data are from different environments. Both criteria originate from information theory. Simulations on telephone speech have revealed the high efficiency of the method.