Implementation of Real-Time Multiple Sound Source Localization and Separation

Authors : Jeng-Shin Sheu, Qi-Xun Zheng

Abstract : This paper mainly discusses a method of separating speech when using a microphone array without knowing the number and direction of sound sources. In recent years, there have been many studies on the method of separating signals by using masking, but most of the separation methods must be operated under the condition of a known number of sound sources. Such methods cannot be used for real-time applications. In our method, this paper uses Circular-Integrated-Cross-Spectrum to estimate the statistical histogram distribution of the direction of arrival (DOA) to obtain the number of sound sources and sound in the mixed-signal Source direction. In calculating the relevant parameters of the ring integrated cross-spectrum, the phase (Phase of the Cross-Power Spectrum) and phase rotation factors (Phase Rotation Factors) calculated by the cross power spectrum of each microphone pair are used. In the part of separating speech, it uses the DOA weighting and shielding separation method to calculate the sound source direction (DOA) according to each T-F unit (time-frequency point). The weight corresponding to each T-F unit can be used to strengthen the intensity of each sound source from the T-F unit and reduce the influence of the remaining sound sources, thereby achieving voice separation.