

## **A Modeling of Singing Voice Robust to Accompaniment Sounds and Its Application to Singer Identification and Vocal-Timbre-Similarity-Based Music Information Retrieval**

Hiromasa Fujihara, Masataka Goto, Tetsuro Kitahara, and Hiroshi G. Okuno

This paper describes a method of modeling the characteristics of singing voice from polyphonic musical audio signals including sounds of various musical instruments. Because singing voices play an important role in musical pieces with vocals, such representation is useful for music information retrieval systems. The main problem in modeling the characteristics of a singing voice is the negative influences caused by accompaniment sounds. To solve this problem, we developed two methods, *accompaniment sound reduction* and *reliable frame selection*. The former makes it possible to calculate feature vectors that represent a spectral envelope of a singing voice after reducing accompaniment sounds. It first extracts the harmonic components of the predominant melody from sound mixtures and then resynthesized the melody by using a sinusoidal model driven by these components. The latter method then estimates the reliability of frame of the obtained melody (i. e., the influence of accompaniment sound) by using two Gaussian mixture models (GMMs) for vocal and nonvocal frames to select the reliable vocal portions of musical pieces. Finally, each song is represented by its GMM consisting of the reliable frames. This new presentation of the singing voice is demonstrated to improve the performance of an automatic singer identification system and to achieve an MIR system based on vocal timbre similarity.

**Index Terms** Music information retrieval (MIR), singer identification, singing voice, vocal, vocal timbre similarity