Janice Kerfoot
MUMT 621
Nov. 13, 2009

Summary: Performer Identification and Similarity

The development of automatic performer identification methods is an example of researchers trying to devise ways to get computers to be passably competent at something humans are inexplicably good at. Theoretical fascination is thus one motivation and the hope of finding ways to execute a simple task on a high volume of data another. When people identify the artist performing a piece of music, they are probably calling a combination of several kinds of knowledge into service. Just by recognizing some coarse classes like kind of instrument, gender of vocalist, broad music genre (e.g. opera, heavy rock), people can narrow their guess by considering those observations against background knowledge about what kind of sounds artists are typically responsible for. Oftentimes, identifying the voice of a singer takes only a few seconds, and exactly what features allow people to do that is what the researchers discussed here are concerned with, along with developing machine–learning algorithms that can effectively be taught to generalize from training on those features to allow them to identify voices as similar for purposes of voice type classification or artists identification.

Performer recognition can also refer to efforts is designing technology that can discern between the sounds produced by different instrumentalists playing the same songs. This is a logical extension of methods developed for musicological ends; facilitating close analysis of the playing styles of particular performers. Though many people can recognize the characteristic sound of some instrumentalists, it usually implies more training in the form of interested, repeated listening, and timbral features alone do not play as important a role as they do with voice recognition.

Extracting features from just the target "voice" (vocal or instrumental) in polyphonic recordings is a serious challenge. Researchers have approached this challenge by finding ways to reduce the influence of the accompaniment sounds, or by identifying segments of the music where the target voice is present or especially strong (e.g. Fujihara et al. 2005; Berenzweig, Ellis, and Lawrence 2002). Another challenge is to prevent overfitting so that the system will be robust against variations in the sound of the voice due to different recording situations and production styles. The "album effect"—referring to how songs from one album are usually easier to automatically identify as belonging to the same artist than songs from a different albums—has been a particular concern of Mandel and Ellis (2005). In their experiments with using SVMs to identify singers in pop and rock songs, they trained the algorithms using songs from the same album and tested them on a mixture of songs from that album and different albums. While the album effect was still evident, their technique of using song–level features and SVMs showed improvement over methods involving artist–level features and/or non–SMV algorithms.

Along with the title of the work, the name of the artist responsible for a piece of music is one of the most crucial pieces of information to meeting user needs in most contexts. Performer recognition systems could have useful application in cases where metadata is absent or nearly absent, as in mobile services that allow users to identify songs from audio samples they provide, or to aid archival work of processing large collections of recorded audio. In the latter example, performer identification could constitute a useful form of quality control to assess the likelihood that metadata is correct, especially in systems that use other automated techniques of identification, or those that employ user–generated information (e.g. tagging). Developers of performer identification systems are likely to find very lucrative applications in the area of copyright defense, by introducing a means of crawling the World Wide Web "listening" to audio for which metadata is deliberately evasive as a way to conceal illegal bootlegs.

Bibliography

Berenzweig, A., D. Ellis, and S. Lawrence. 2002. Using voice segments to improve artist classification of music. In *Proceedings of the 22nd AES International Conference*, Espoo, Finland.

Fujihara, H., T. Kitahara, M. Goto, K. Komatani, T. Ogata, and H.G. Okuno . 2005. Singer identification based on accompaniment sound reduction and reliable frame selection. In *Proceedings of the 6th International Conference on Music Information Retrieval*.

Kako, T., Y. Ohishi, H. Kameoka, K. Kashino, and K. Takeda. 2009. Automatic identification for singing style based on sung melodic contour characterized in phase plane. In *Proceedings of the 10th International Conference on Music Information Retrieval*.

Kim, Y.E. and B.Whitman. 2002. Singer identification in popular music recordings using voice coding features. In *Proceedings of the 3rd International Conference on Music Information Retrieval*.

Mandel, M., and D. Ellis. 2005. Song–level features and support vector machines for music classification. In *Proceedings of the 6th International Conference on Music Information Retrieval*.

Mesaros, A., and J. Astola. 2005. The mel–frequency cepstral coefficients in the context of singer identification. In *Proceedings of the 6th International Conference on Music Information Retrieval*.

Ramirez, R., A. Perez, and S. Kersten. 2008. Performer identification in Celtic violin recordings. In *Proceedings of the 9th International Conference on Music Information Retrieval*.

Tsai, W.H. and H.M. Wang. 2006. Automatic singer recognition of popular music recordings via estimation and modeling of solo vocal signals. *IEEE Transactions on Audio, Speech and Language Processing*, vol.14: 330–341.

Zhang, T.. 2003. Automatic singer identification. In *Proceedings of the 2003 International Conference on Multimedia and Expo*, vol. 1., pp.33–36.