Singer Identification

Bertrand SCHERRER

McGill University

March 15, 2007

Outline

Outline

Singer Identification is to be (has been) applied on **pop music** mainly

- Automatically label data for which no/or not much information is available $\Rightarrow$ recognize the singer
- Distinguish between original version of a song and cover songs
- Copyright enforcement: recording companies could scan bootleg sites on the internet to check if there are any unauthorized recorded versions of a concert [Kim, 2002 and Tsai and Wang, 2006]
- Music recommendation systems could use singer identification to group singers with same voice characteristics.

- Automatically label data for which no/or not much information is available $\Rightarrow$ recognize the singer
- Distinguish between original version of a song and cover songs
- Copyright enforcement: recording companies could scan bootleg sites on the internet to check if there are any unauthorized recorded versions of a concert [Kim, 2002 and Tsai and Wang, 2006]
- Music recommendation systems could use singer identification to group singers with same voice characteristics.

- Automatically label data for which no/or not much information is available $\Rightarrow$ recognize the singer
- Distinguish between original version of a song and cover songs
- Copyright enforcement: recording companies could scan bootleg sites on the internet to check if there are any unauthorized recorded versions of a concert [Kim, 2002 and Tsai and Wang, 2006]
- Music recommendation systems could use singer identification to group singers with same voice characteristics.

- Automatically label data for which no/or not much information is available $\Rightarrow$ recognize the singer
- Distinguish between original version of a song and cover songs
- Copyright enforcement: recording companies could scan bootleg sites on the internet to check if there are any unauthorized recorded versions of a concert [Kim, 2002 and Tsai and Wang, 2006]
- Music recommendation systems could use singer identification to group singers with same voice characteristics.

- Singing Voice $=$ hybrid btw **speech** and **musical instrument** $\Rightarrow$ create specific methods of analysis.
- In pop music, voice is never heard alone: presence of **accompaniement**

- Singing Voice $=$ hybrid btw **speech** and **musical instrument** $\Rightarrow$ create specific methods of analysis.
- In pop music, voice is never heard alone: presence of **accompaniement**

Outline

- As seen in the previous diagrams: need to extract some features from the sounds.
- Features used:
  - MFCC (Mel-Frequency Cepstral Coefficient)
  - MDCT (Modified Discrete Cosine Transform)
  - LPCC (Linear Predictive Coding Coefficients)
  - WLPCC (Warped ...)
  - Cepstral Coefficients of the LPC spectrum
  - LPMFCC (MFCC of the LPC spectrum)

- As seen in the previous diagrams: need to extract some features from the sounds.
- Features used:
  - MFCC (Mel-Frequency Cepstral Coefficient)
  - MDCT (Modified Discrete Cosine Transform)
  - LPCC (Linear Predictive Coding Coefficients)
  - WLPCC (Warped ...)
  - Cepstral Coefficients of the LPC spectrum
  - LPMFCC (MFCC of the LPC spectrum)

Outline

Principle

- Difference in spectrum between voiced regions and accompaniement-only: **hamonicity** of the voice.

Voice/Accompaniement Spectra

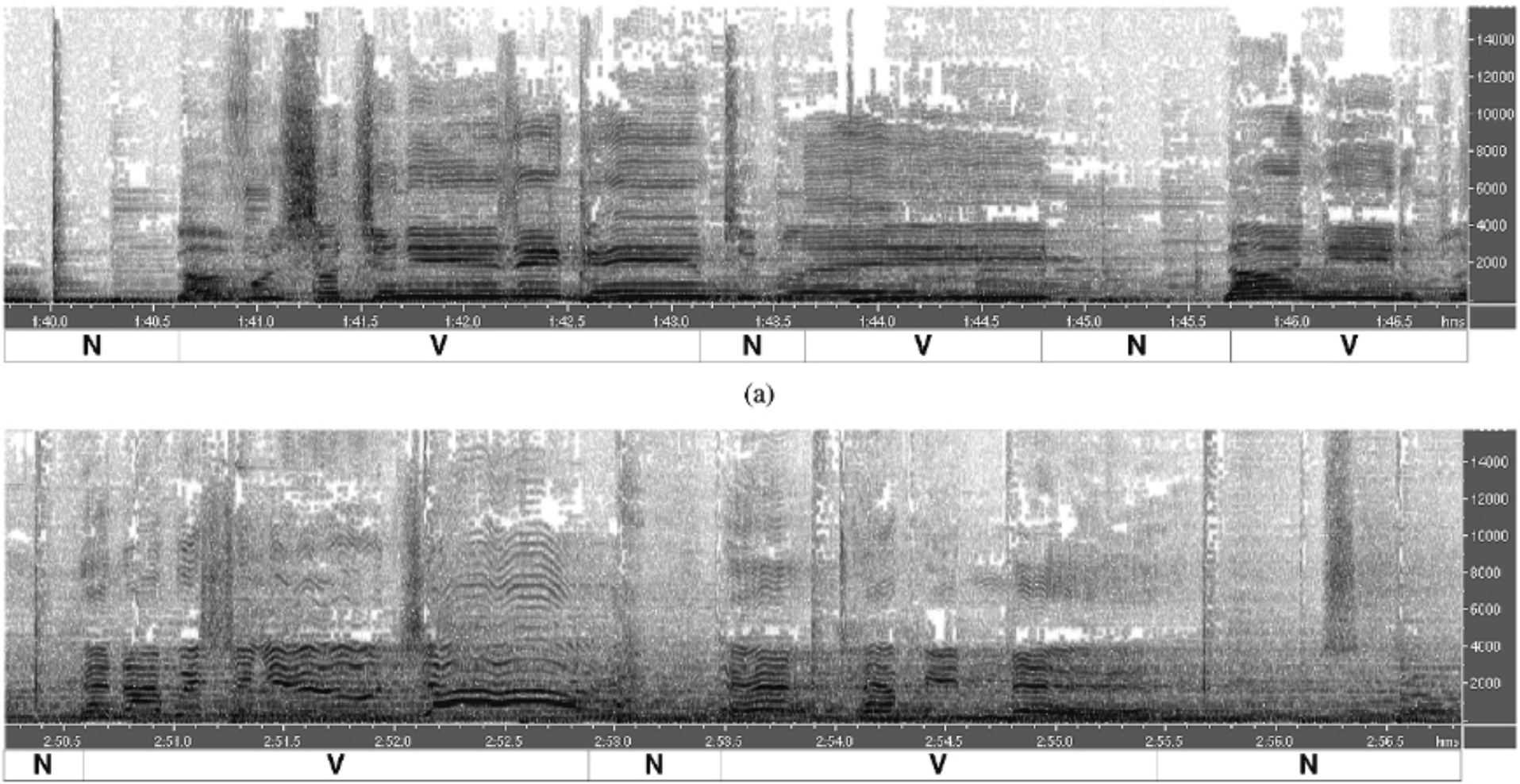


Fig.1 [Tsai and Wang, 2006]

Tsai's Approach

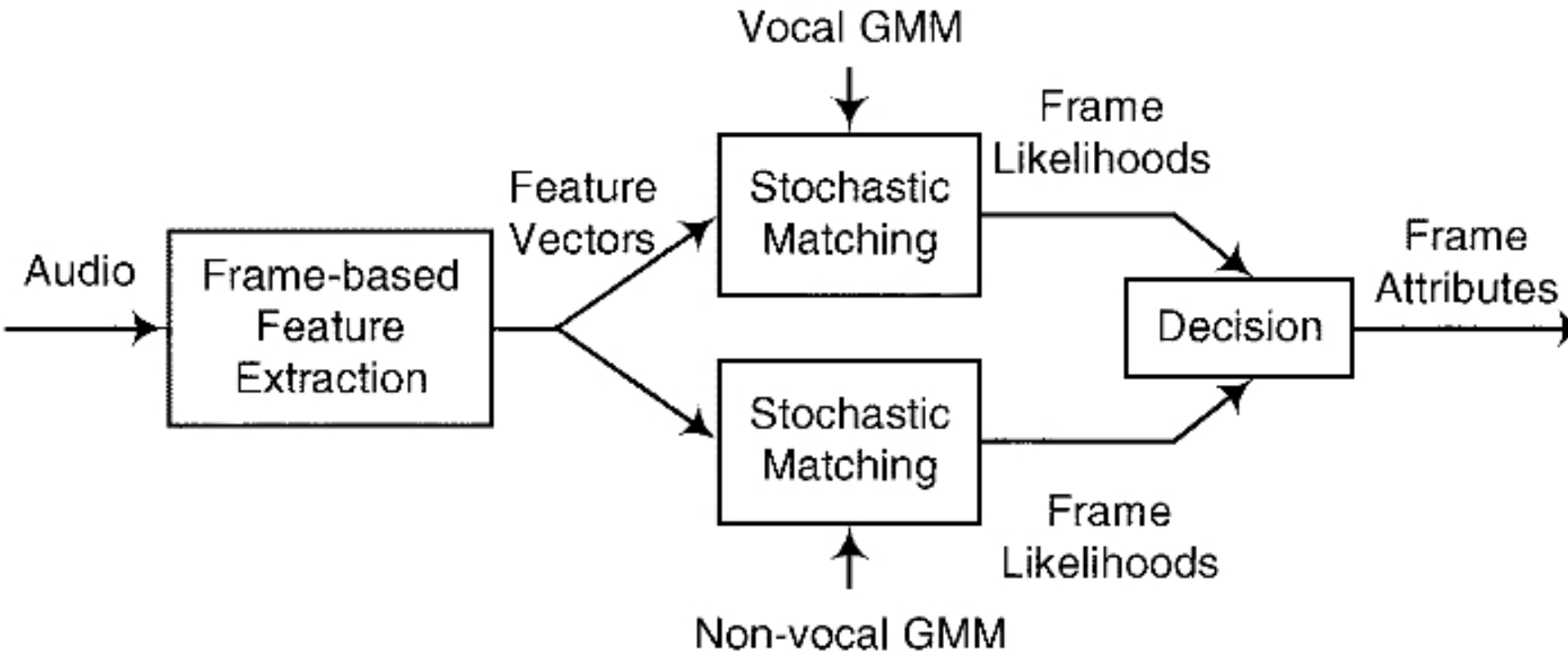
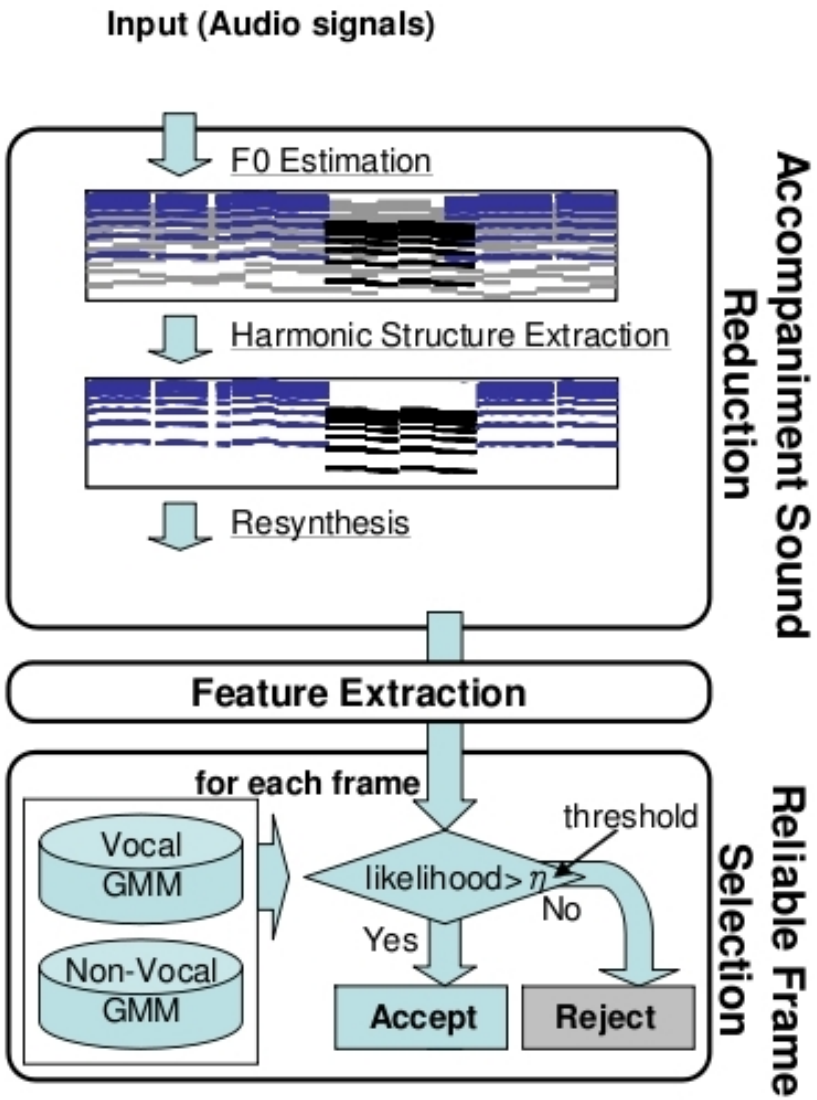


Fig.1 [Tsai, 2004]

Tsai's Approach

- This method is supposed to yield 82.3% accuracy [Tsai and Wang, 2006]

Fujihara's Approach



from Fig.1 [Fujihara 2005]

- The GMM classification between Vocal and Non Vocal is done on the resynthesized signal.

Outline

3 main strategies

- GMM
- SVM
- k-NN

GMM Method with Solo Voice Modeling

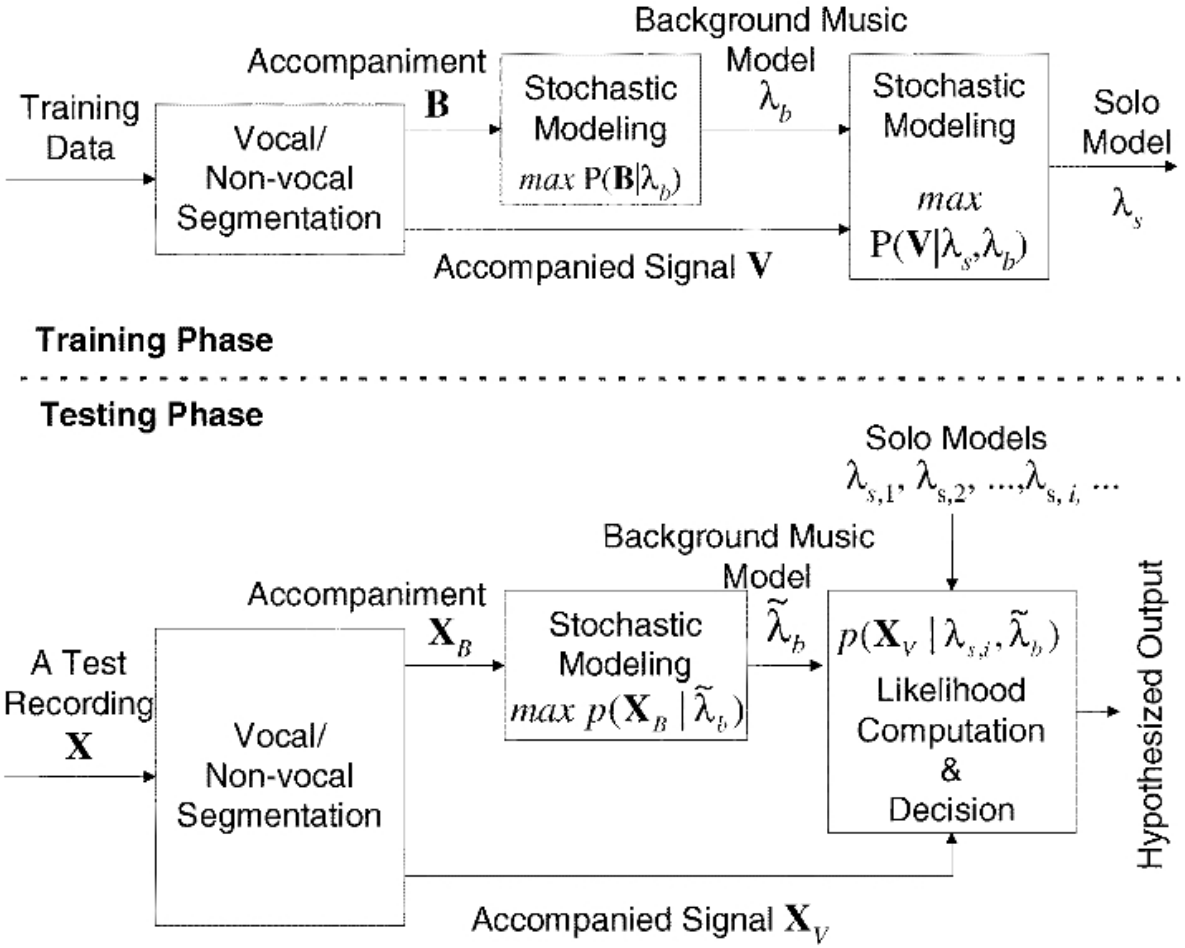


Fig.3 [Tsai and Wang, 2006]

Outline

Performance

- Kim and Whitman 2002 $\Rightarrow$ 45%
- Liu and Huang, 2002 $\Rightarrow$ 80 %
- Tsai and Wang, 2006, Fujihara et al., 2005 $\Rightarrow$ 95%

Outline

Good

- Singer identification yields satisfactory results.

But ...

- Only one article tackles Target Singer Detection or Target Singer Tracking: [Tsai and Wang 2006]. $\Rightarrow$ results are not perfect for duet but are better than doing GMM without solo modeling.
- Specific to pop music $\Rightarrow$ what happens with **a cappela** singer**s**?
- Specific to on geographical area (Asia) $\Rightarrow$ important because of voice mix

But ...

- Only one article tackles Target Singer Detection or Target Singer Tracking: [Tsai and Wang 2006]. $\Rightarrow$ results are not perfect for duet but are better than doing GMM without solo modeling.
- Specific to pop music $\Rightarrow$ what happens with **a cappela** singer**s**?
- Specific to on geographical area (Asia) $\Rightarrow$ important because of voice mix

But ...

- Only one article tackles Target Singer Detection or Target Singer Tracking: [Tsai and Wang 2006]. $\Rightarrow$ results are not perfect for duet but are better than doing GMM without solo modeling.
- Specific to pop music $\Rightarrow$ what happens with **a cappela** singer**s**?
- Specific to on geographical area (Asia) $\Rightarrow$ important because of voice mix

But ...

- Only one article tackles Target Singer Detection or Target Singer Tracking: [Tsai and Wang 2006]. $\Rightarrow$ results are not perfect for duet but are better than doing GMM without solo modeling.
- Specific to pop music $\Rightarrow$ what happens with **a cappela** singer**s**?
- Specific to on geographical area (Asia) $\Rightarrow$ important because of voice mix

Bibliography I

- 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 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 International Conference on Music Information Retrieval*.
- Liu, C.-C. and C.-S. Huang, 2002. A singer identification technique for content-based clas- sification of MP3 music objects. *In Proceedings of the eleventh International Conference on Information and Knowledge Management.*

Bibliography II

- Tsai, W.-H. and H.-M. Wang, 2004. Automatic detection and tracking of target singer in multi-singer music recordings. In Proceedings of the 2004 IEEE International Conferecence on Acoustics, Speech and Signal Processing, vol. 4. pp. 221–224.
- 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.

Questions ?