



Fundamentals of Music Processing

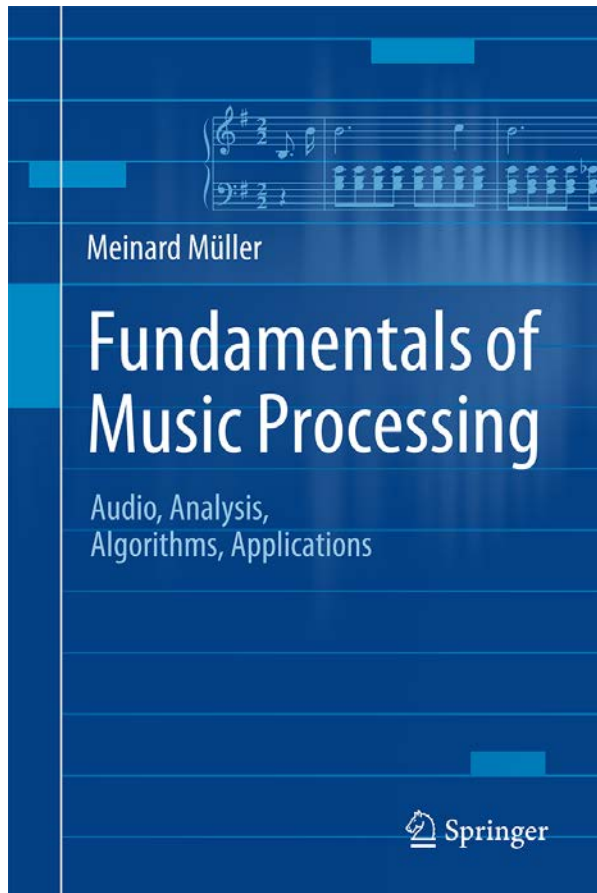
Chapter 8: Musically Informed Audio Decomposition

Meinard Müller

International Audio Laboratories Erlangen

www.music-processing.de

Book: Fundamentals of Music Processing

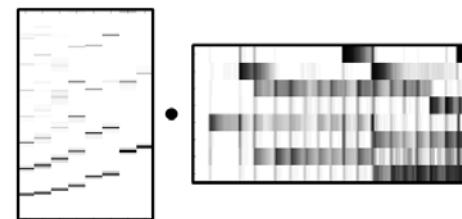


Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., 30 illus. in color, hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de

Chapter 8: Musically Informed Audio Decomposition

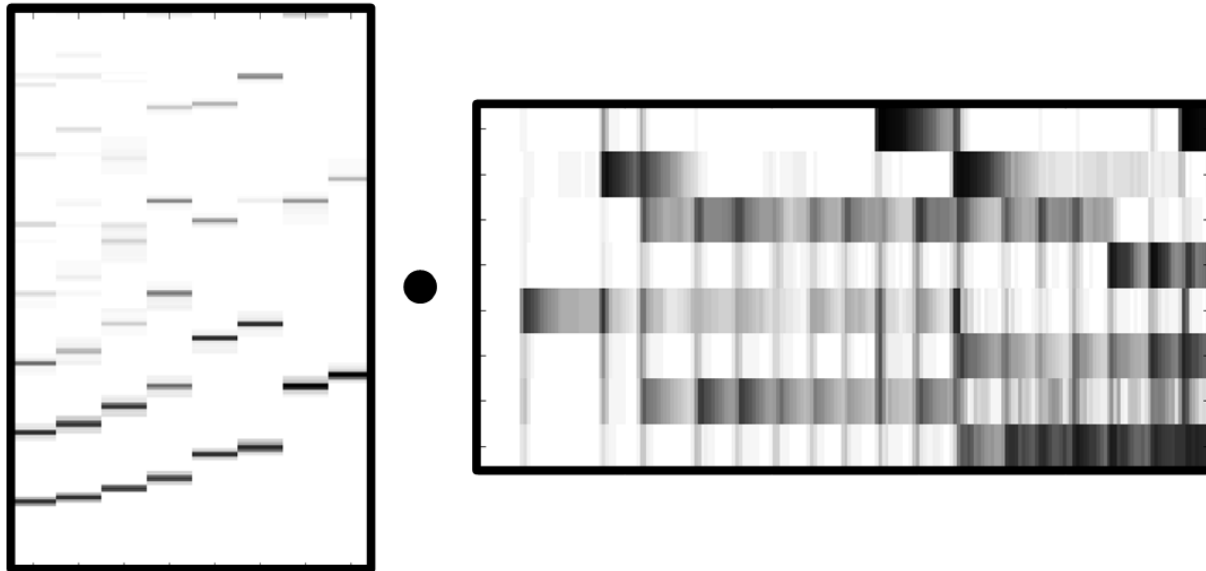
- 8.1 Harmonic-Percussive Separation
- 8.2 Melody Extraction
- 8.3 NMF-Based Audio Decomposition
- 8.4 Further Notes



In the final Chapter 8 on audio decomposition, we present a challenging research direction that is closely related to source separation. Within this wide research area, we consider three subproblems: harmonic–percussive separation, main melody extraction, and score-informed audio decomposition. Within these scenarios, we discuss a number of key techniques including instantaneous frequency estimation, fundamental frequency (F0) estimation, spectrogram inversion, and nonnegative matrix factorization (NMF). Furthermore, we encounter a number of acoustic and musical properties of audio recordings that have been introduced and discussed in previous chapters, which rounds off the book.

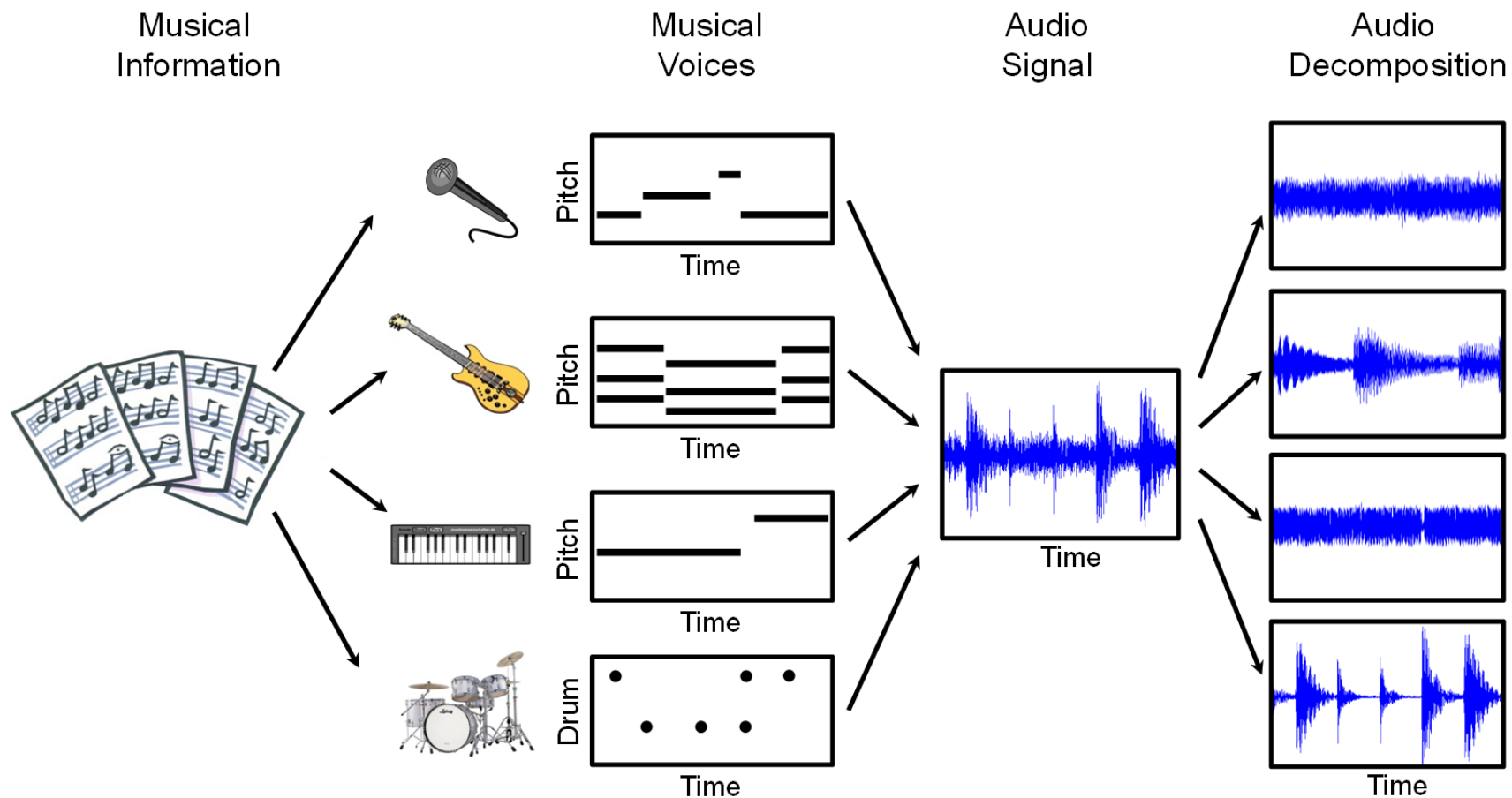
8 Content-Based Audio Retrieval

Teaser



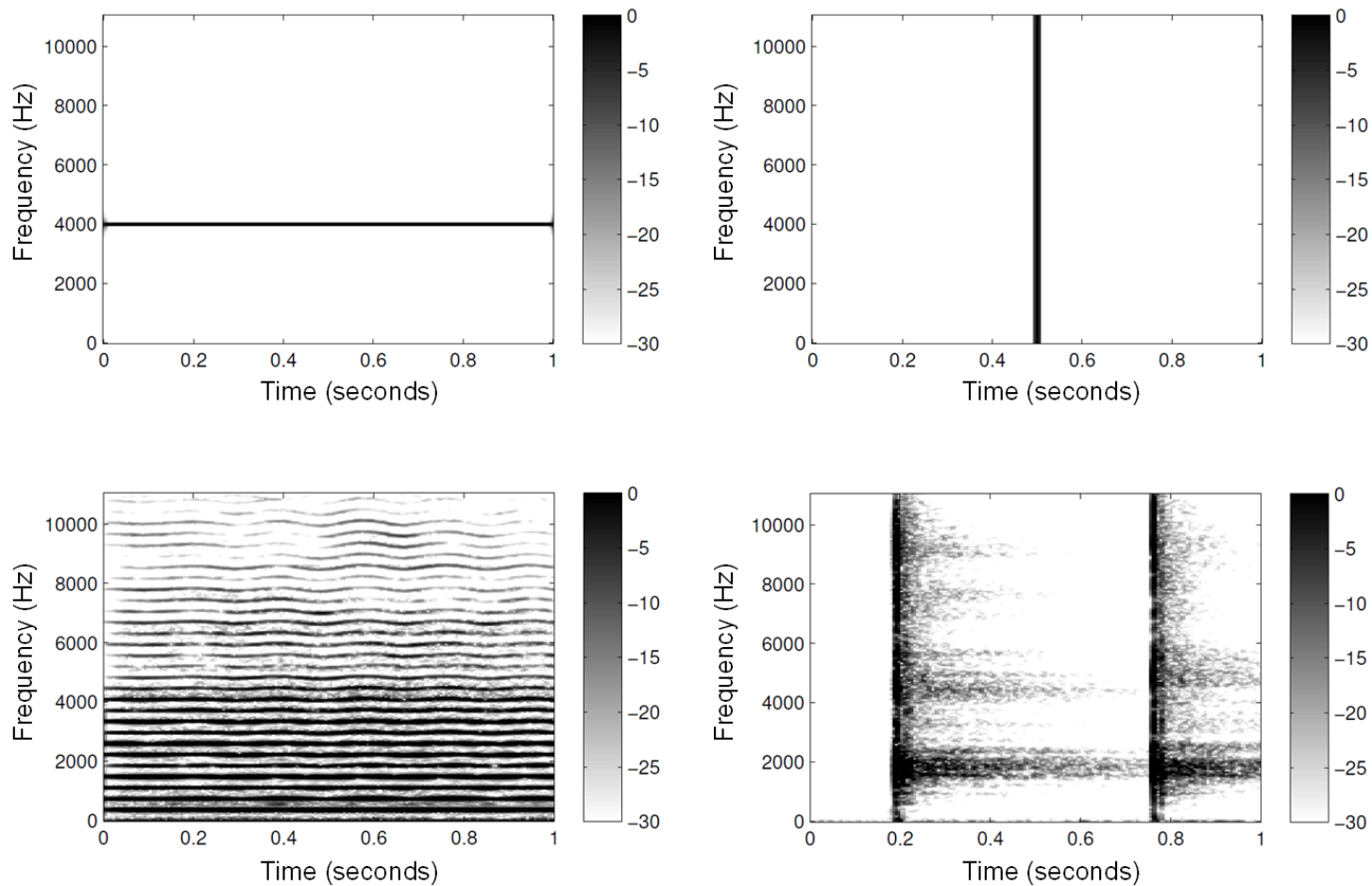
8 Content-Based Audio Retrieval

Fig. 8.1



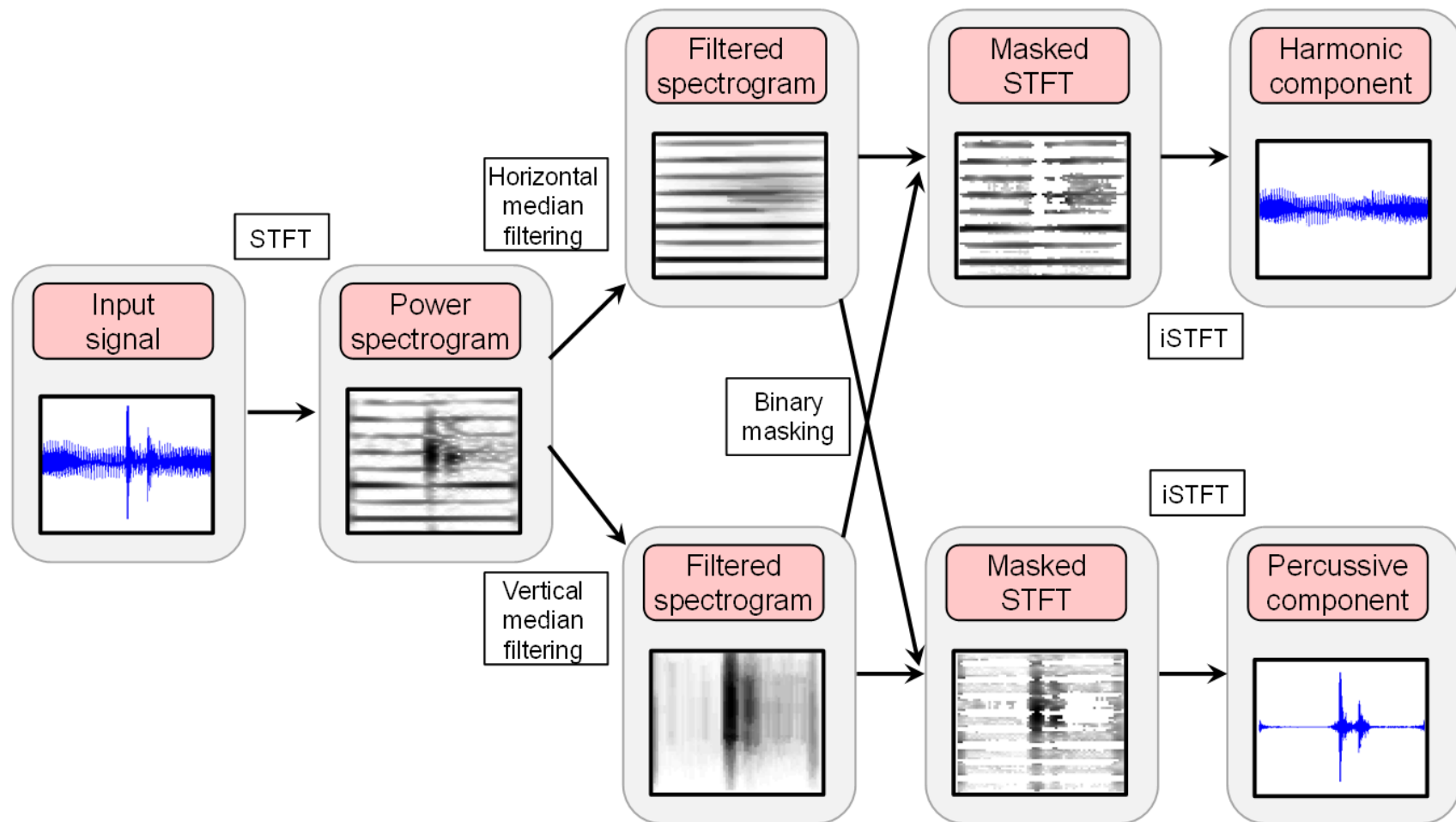
8.1 Harmonic–Percussive Separation

Fig. 8.2



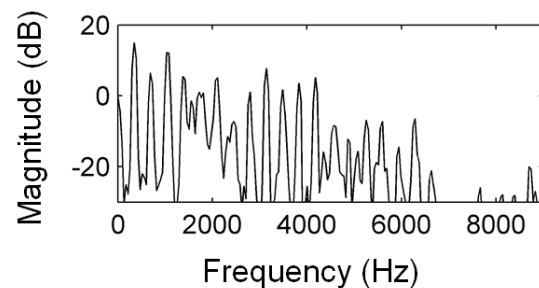
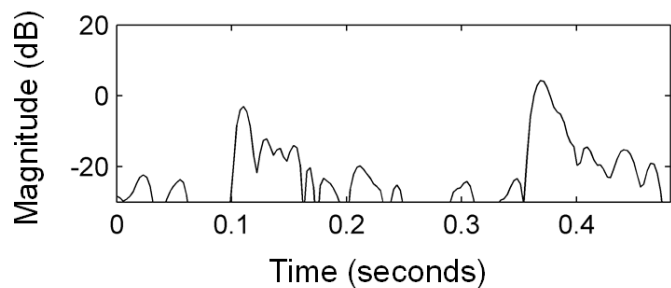
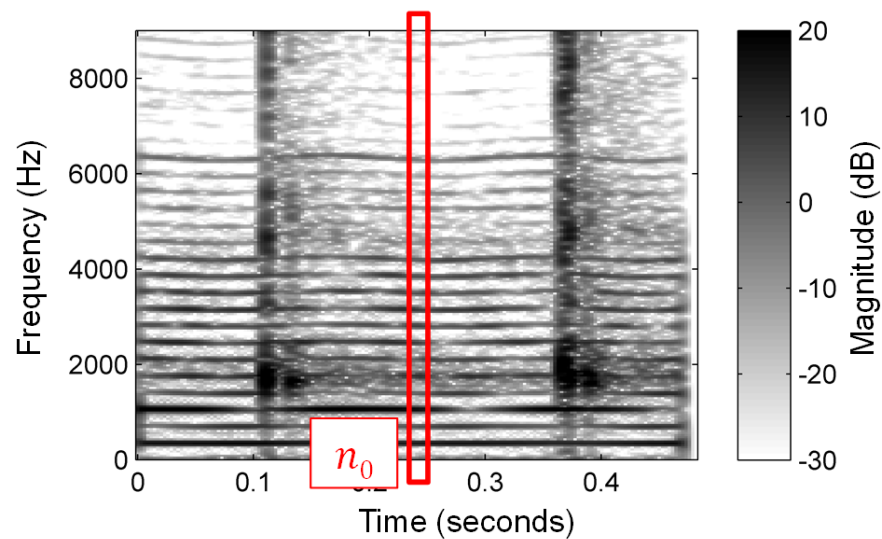
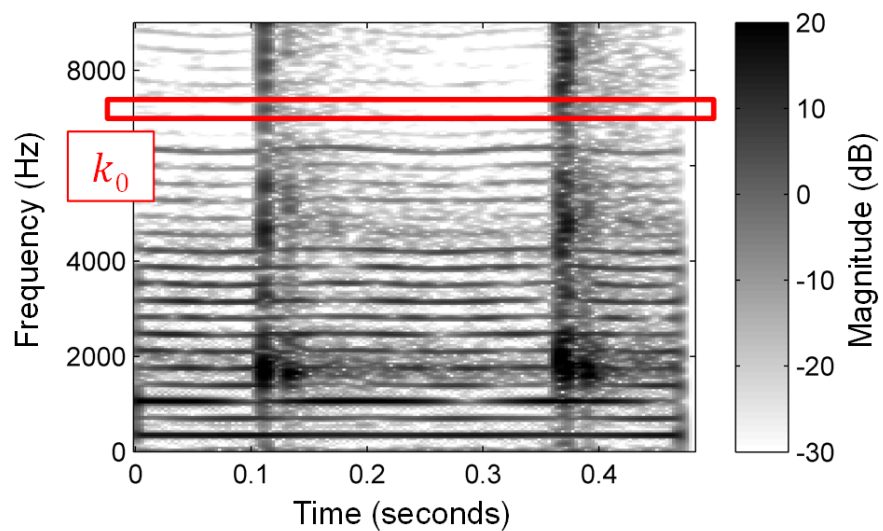
8.1 Harmonic–Percussive Separation

Fig. 8.3



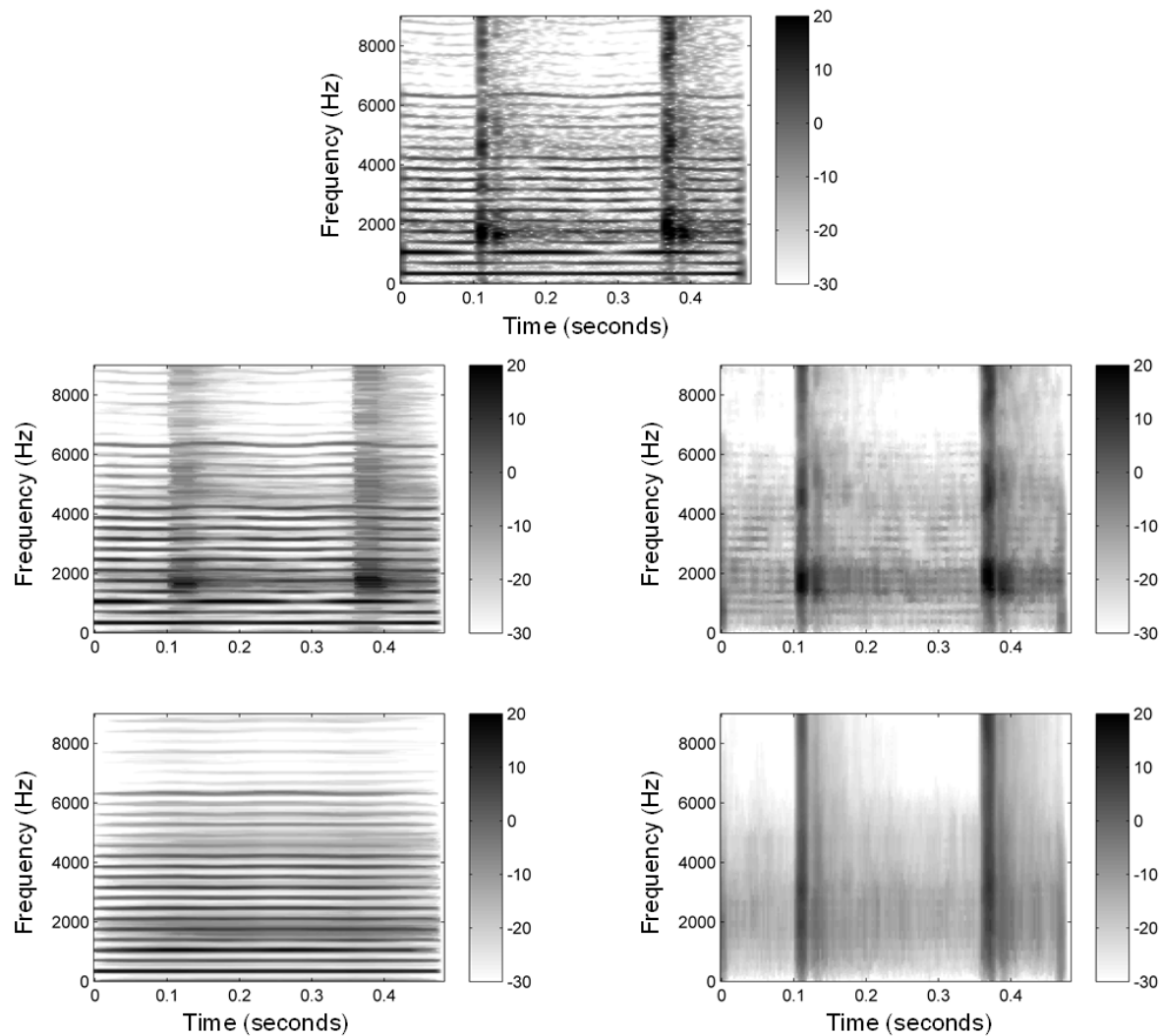
8.1 Harmonic–Percussive Separation

Fig. 8.4



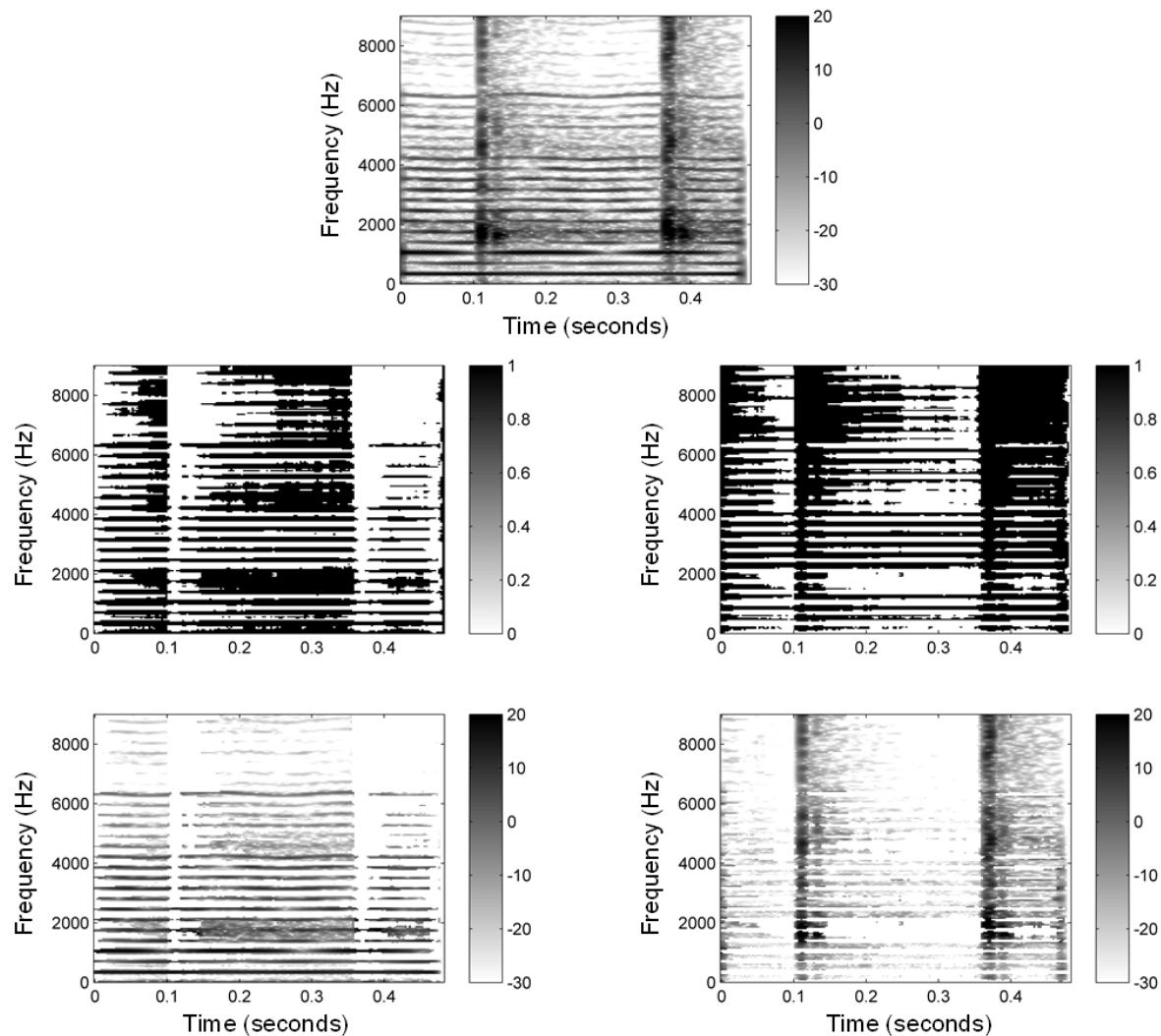
8.1 Harmonic–Percussive Separation

Fig. 8.5



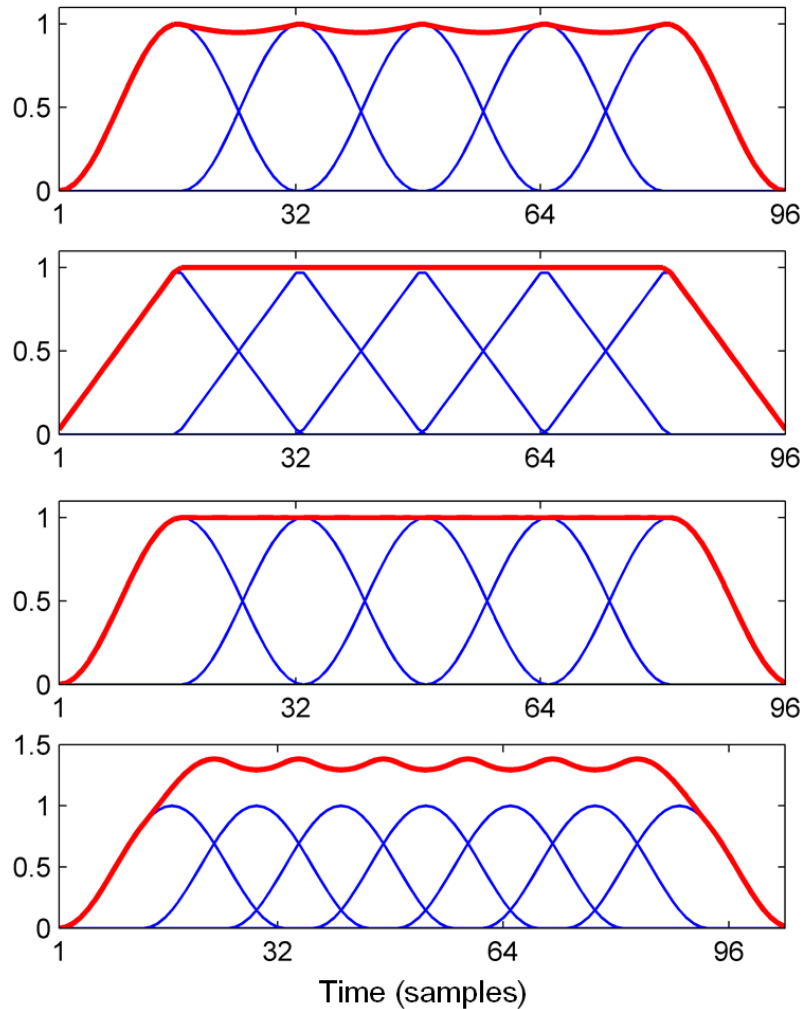
8.1 Harmonic–Percussive Separation

Fig. 8.6



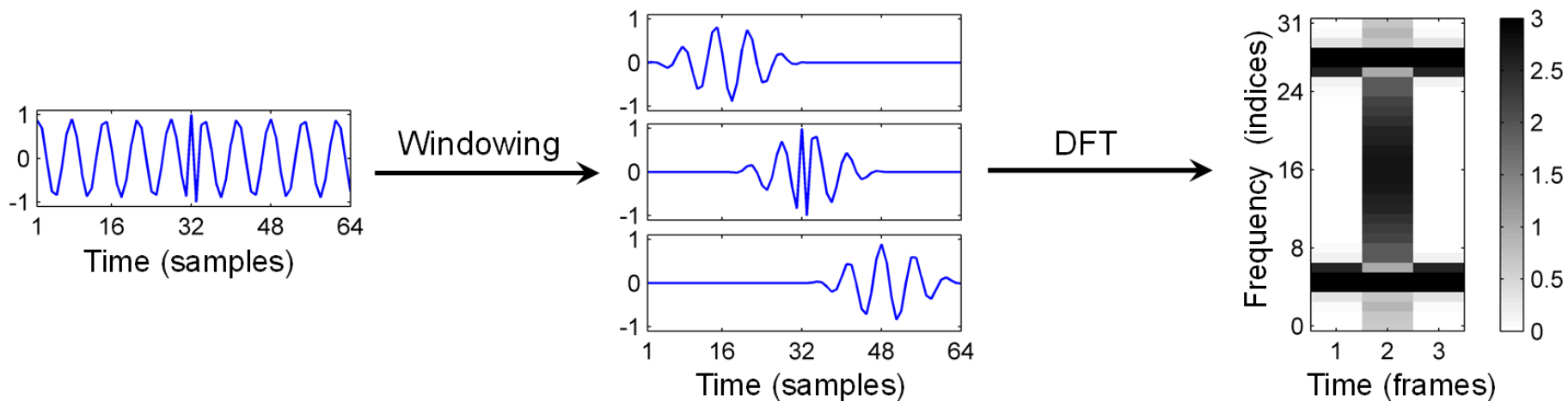
8.1 Harmonic–Percussive Separation

Fig. 8.7



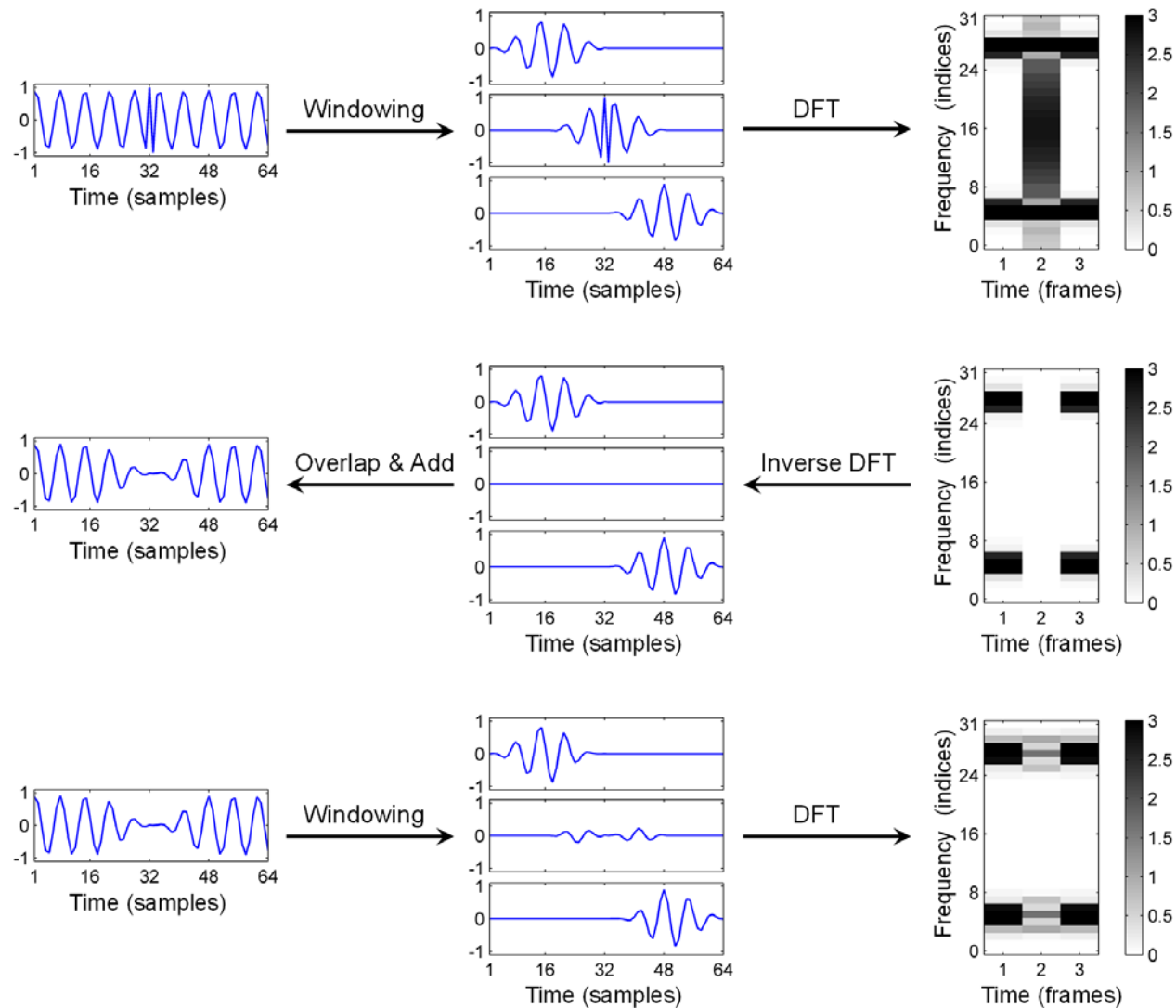
8.1 Harmonic–Percussive Separation

Fig. 8.8



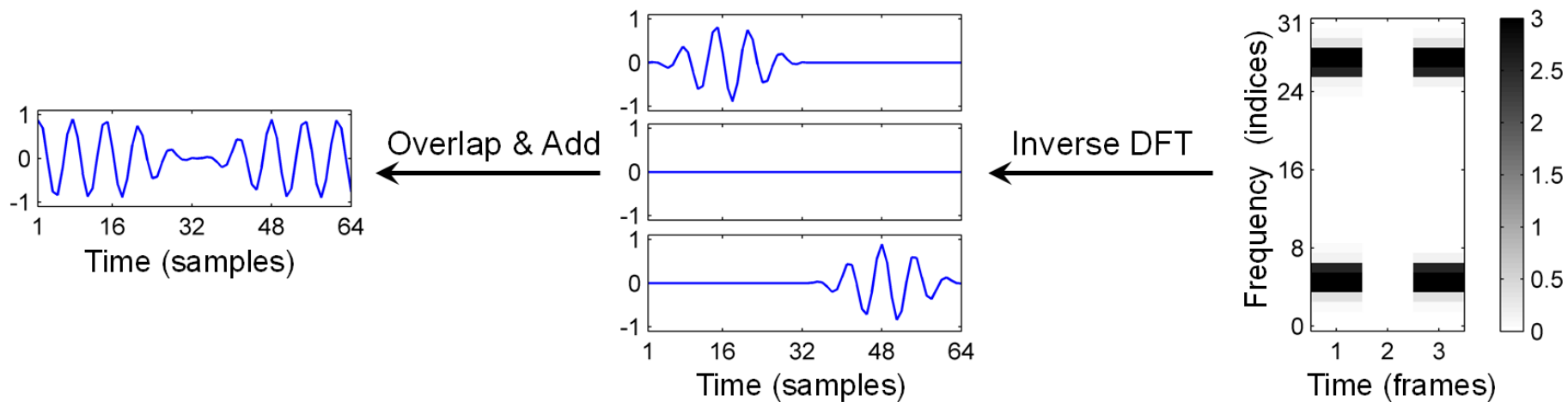
8.1 Harmonic–Percussive Separation

Fig. 8.8



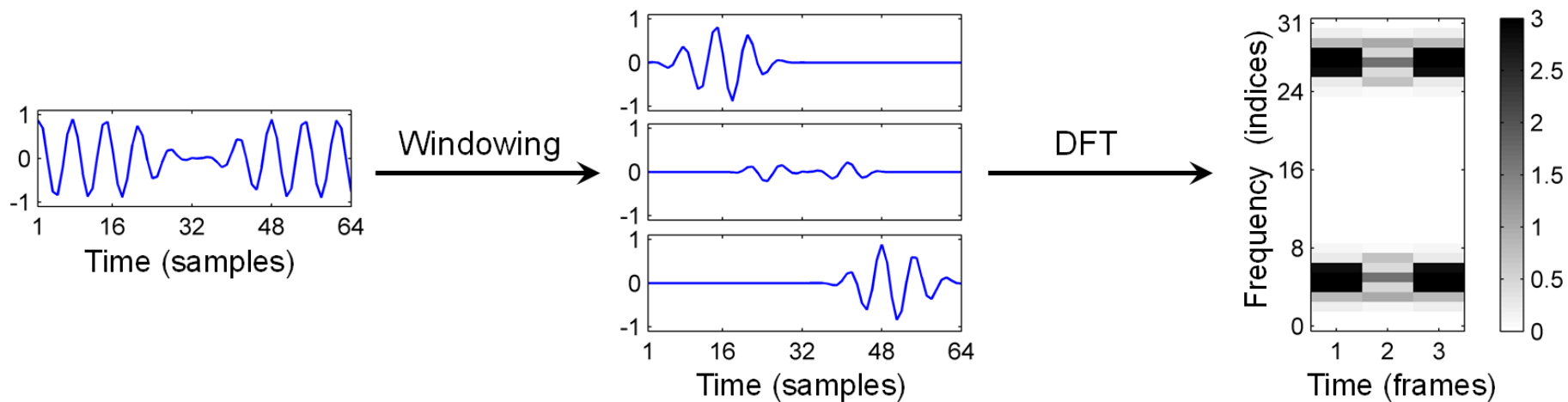
8.1 Harmonic–Percussive Separation

Fig. 8.8



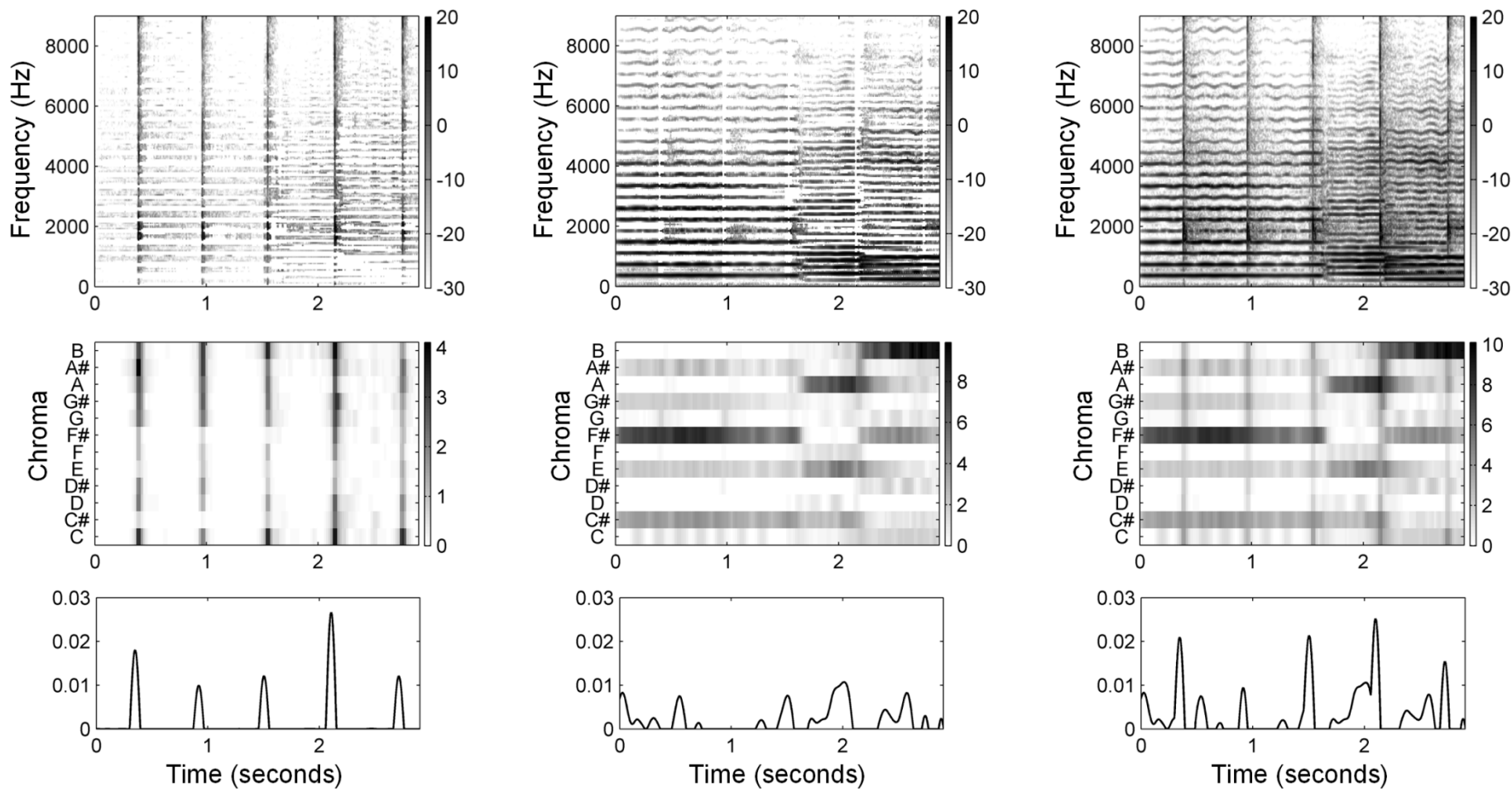
8.1 Harmonic–Percussive Separation

Fig. 8.8



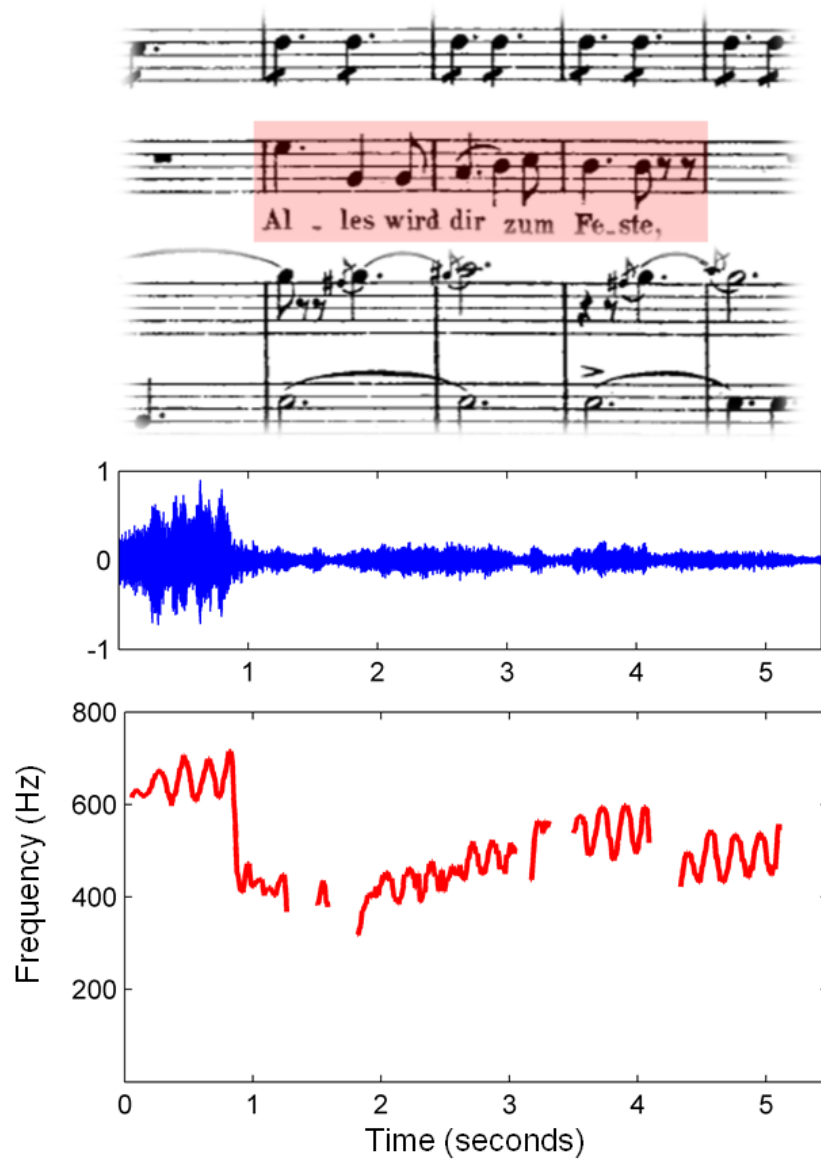
8.1 Harmonic–Percussive Separation

Fig. 8.9



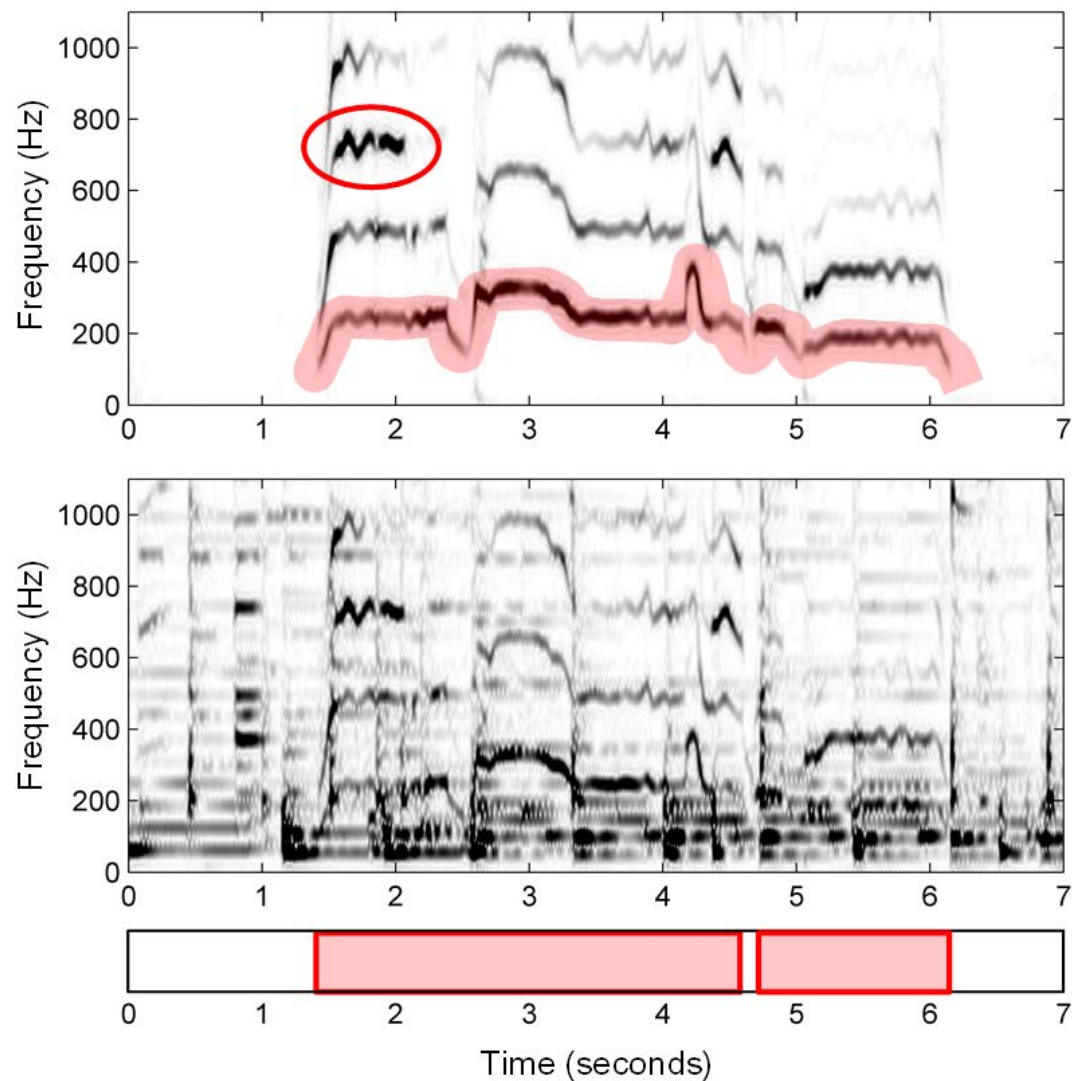
8.2 Melody Extraction

Fig. 8.10



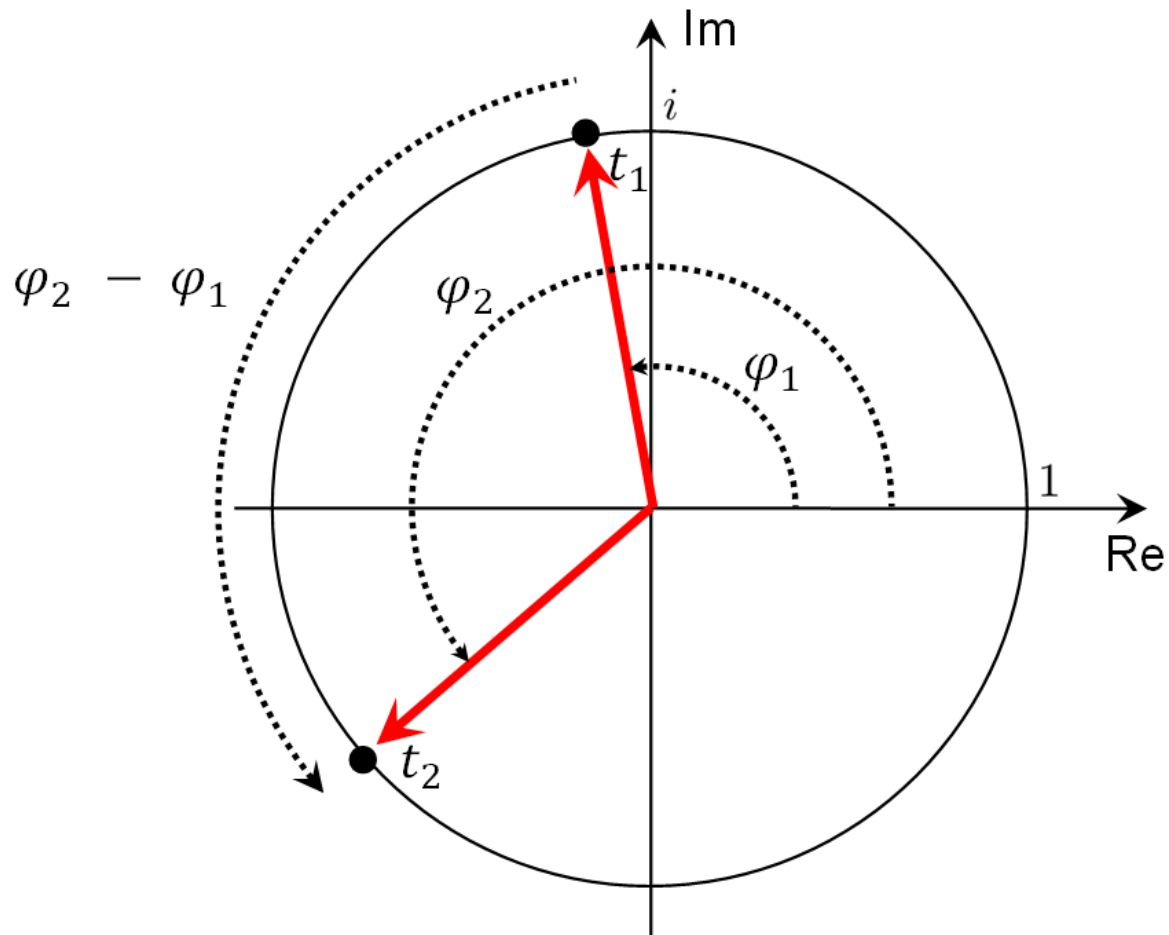
8.2 Melody Extraction

Fig. 8.11



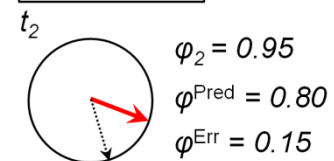
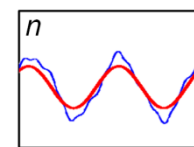
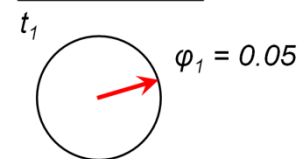
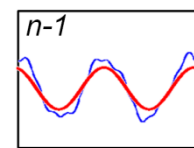
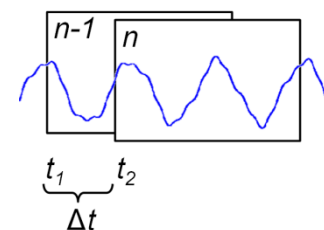
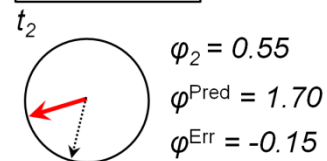
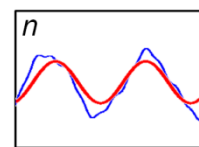
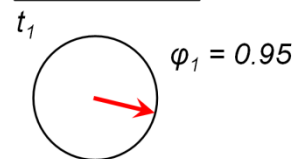
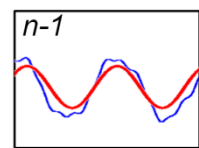
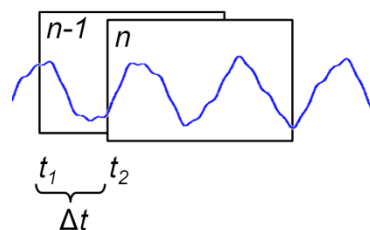
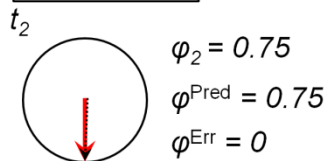
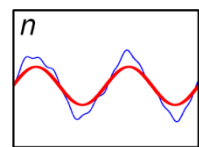
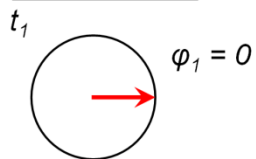
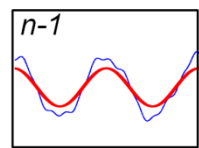
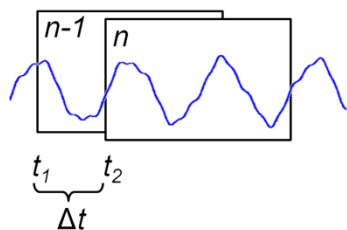
8.2 Melody Extraction

Fig. 8.12



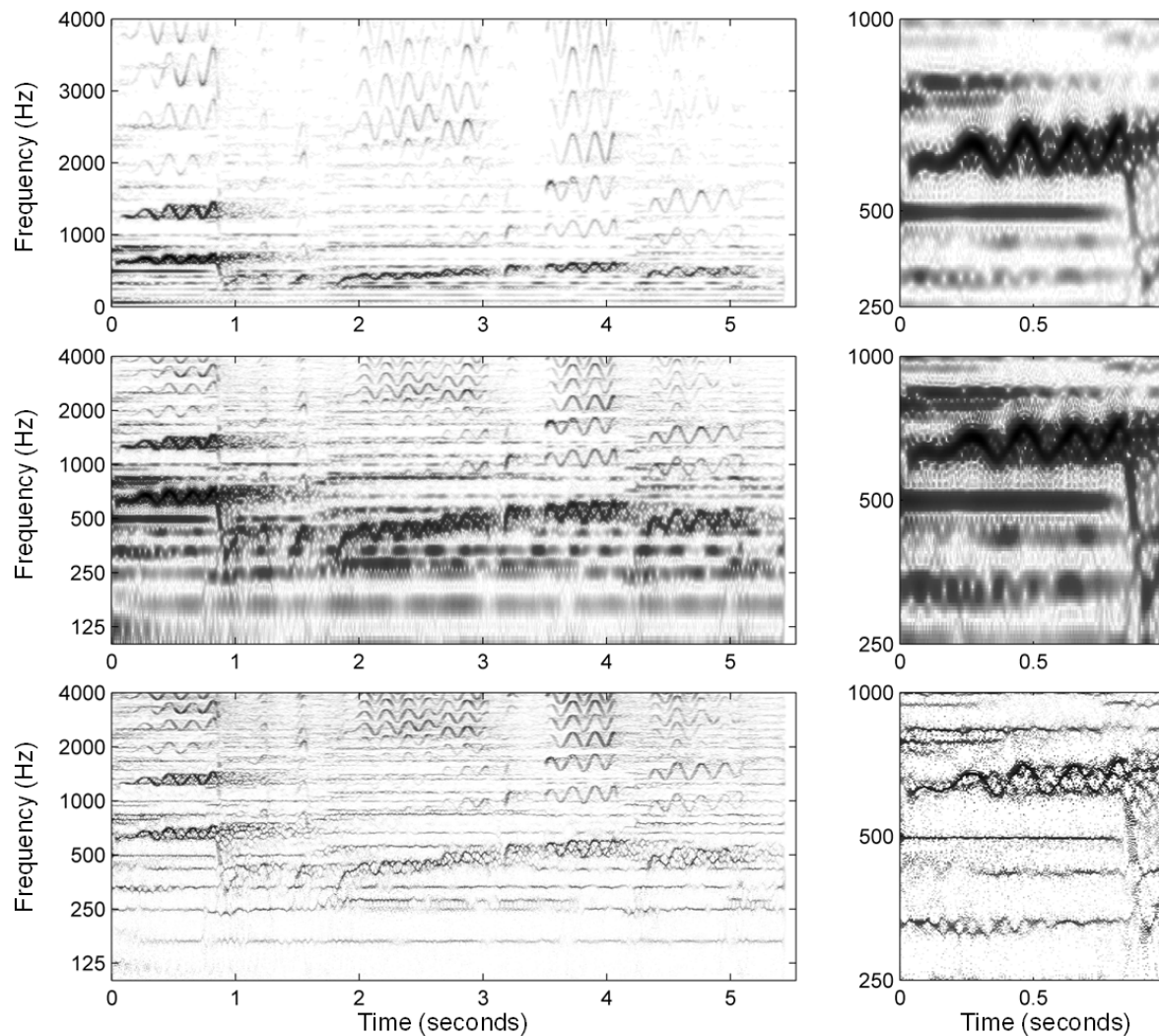
8.2 Melody Extraction

Fig. 8.13



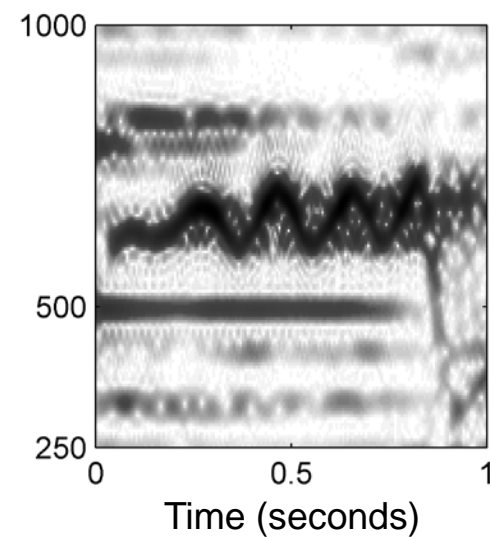
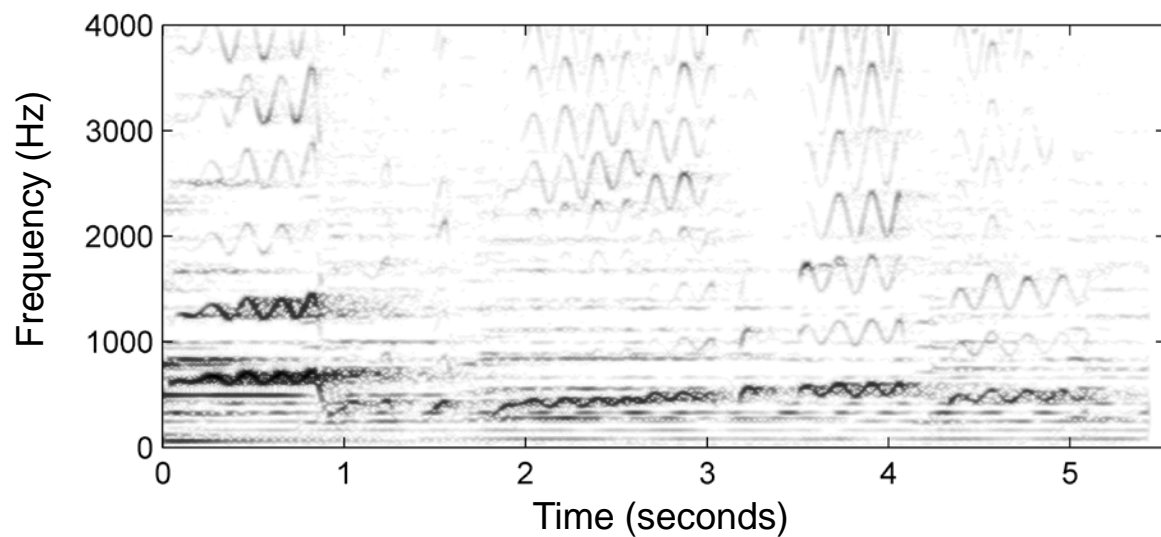
8.2 Melody Extraction

Fig. 8.14



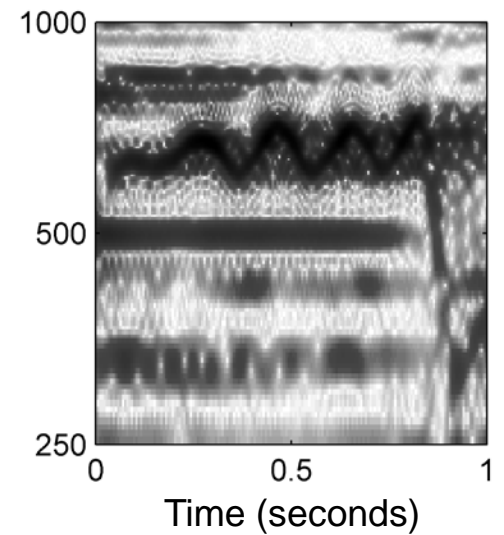
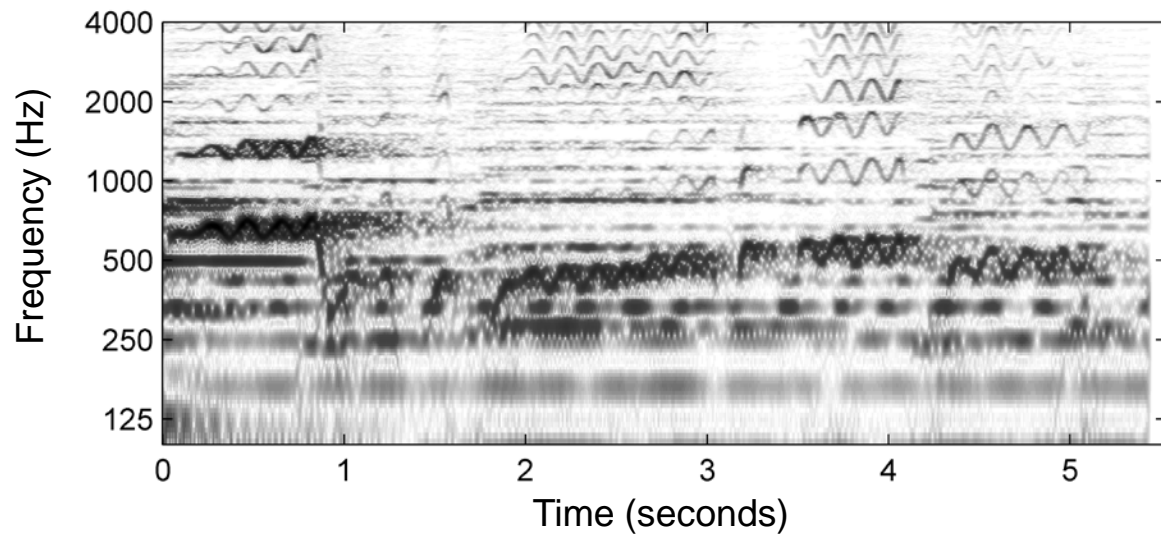
8.2 Melody Extraction

Fig. 8.14



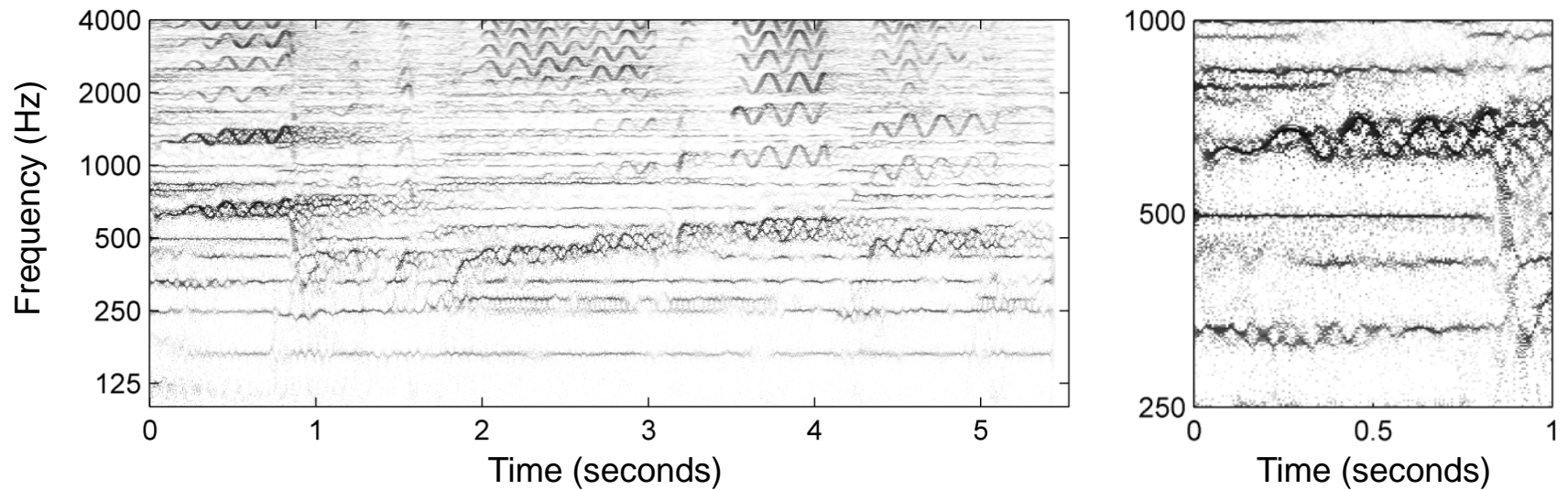
8.2 Melody Extraction

Fig. 8.14



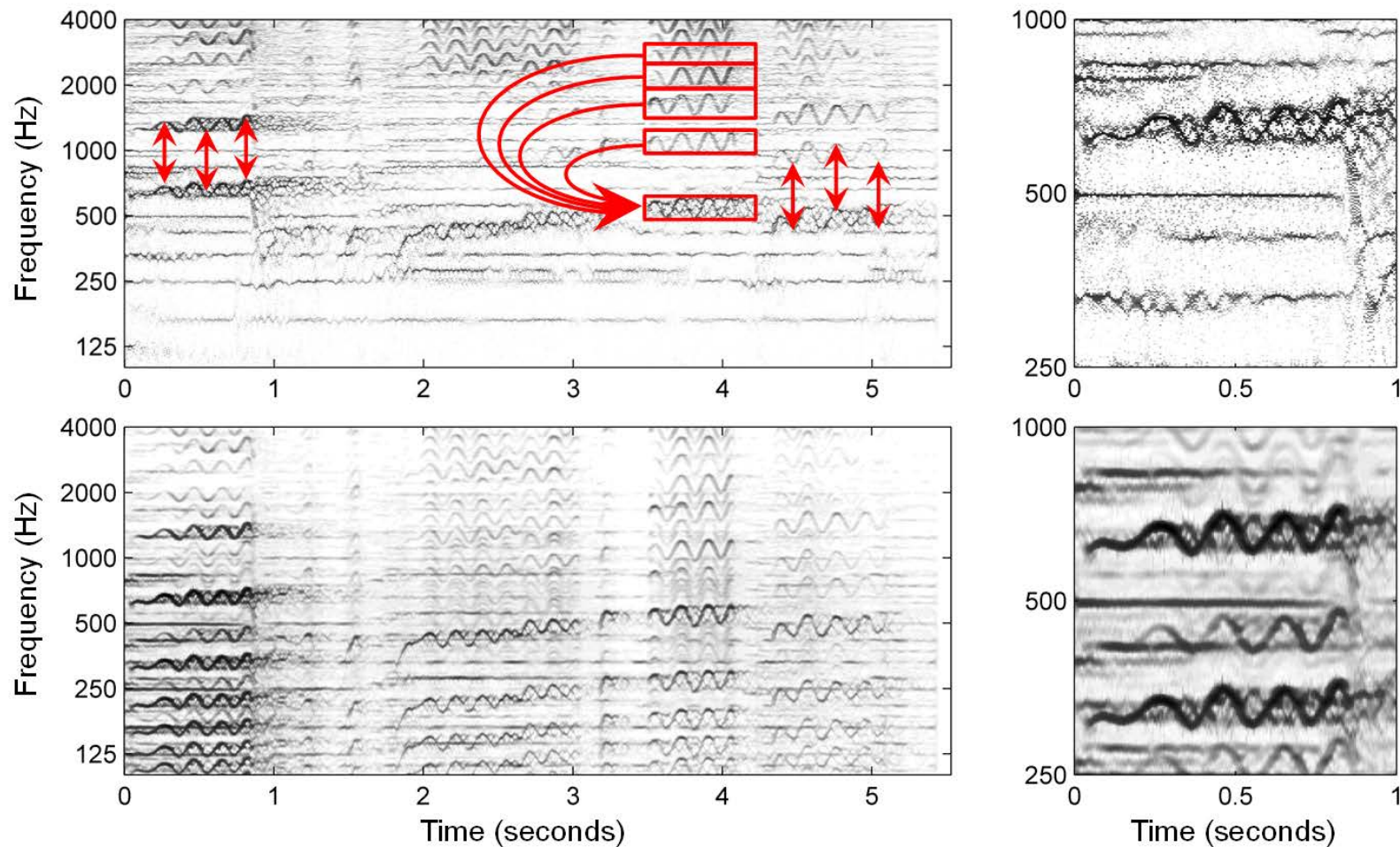
8.2 Melody Extraction

Fig. 8.14



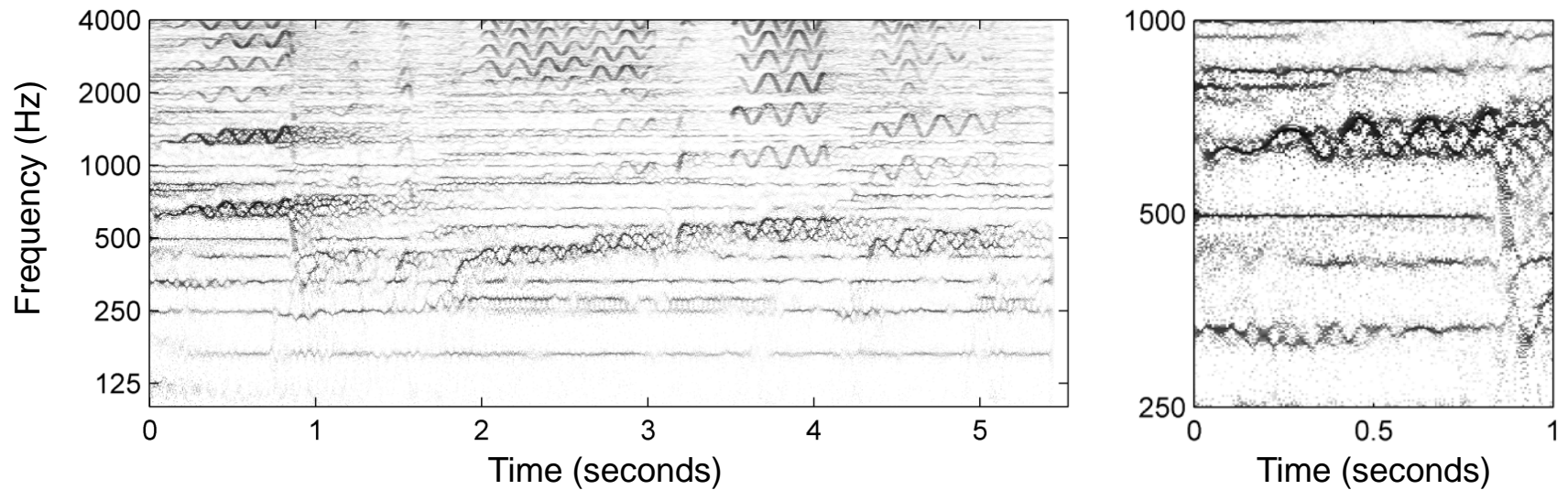
8.2 Melody Extraction

Fig. 8.15



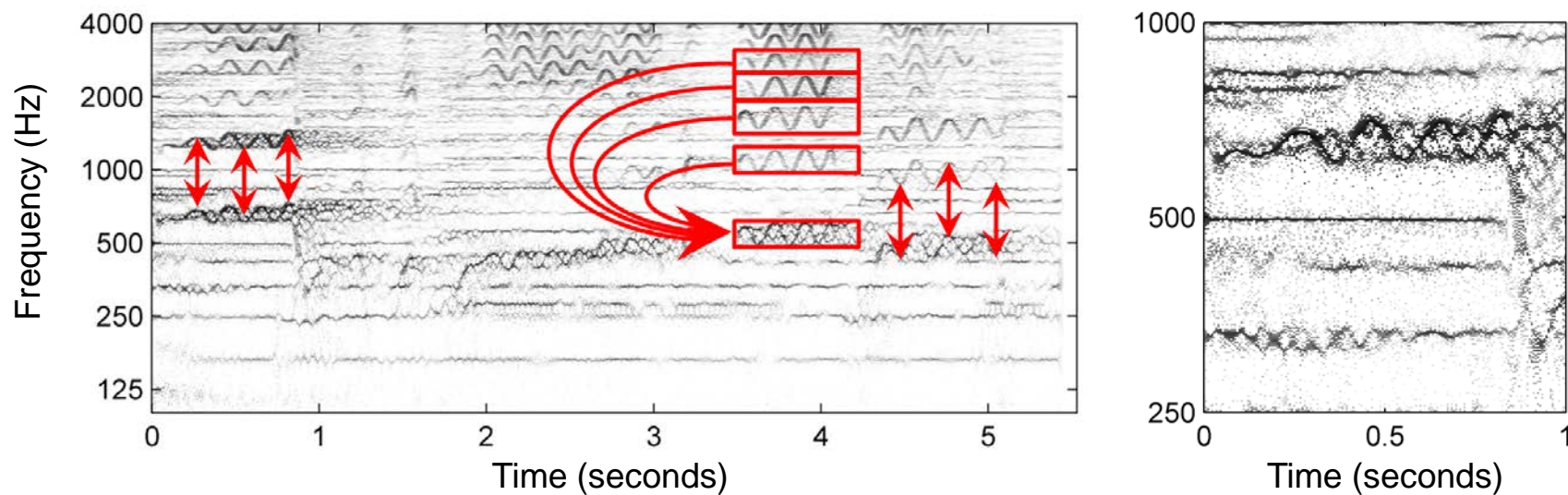
8.2 Melody Extraction

Fig. 8.15



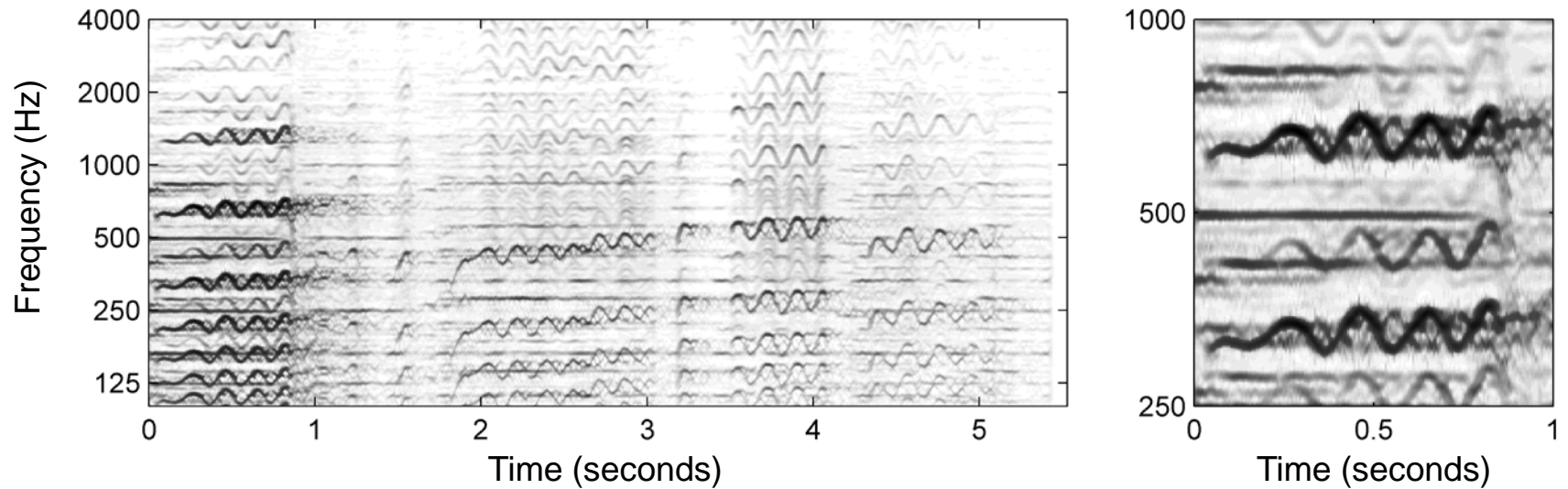
8.2 Melody Extraction

Fig. 8.15



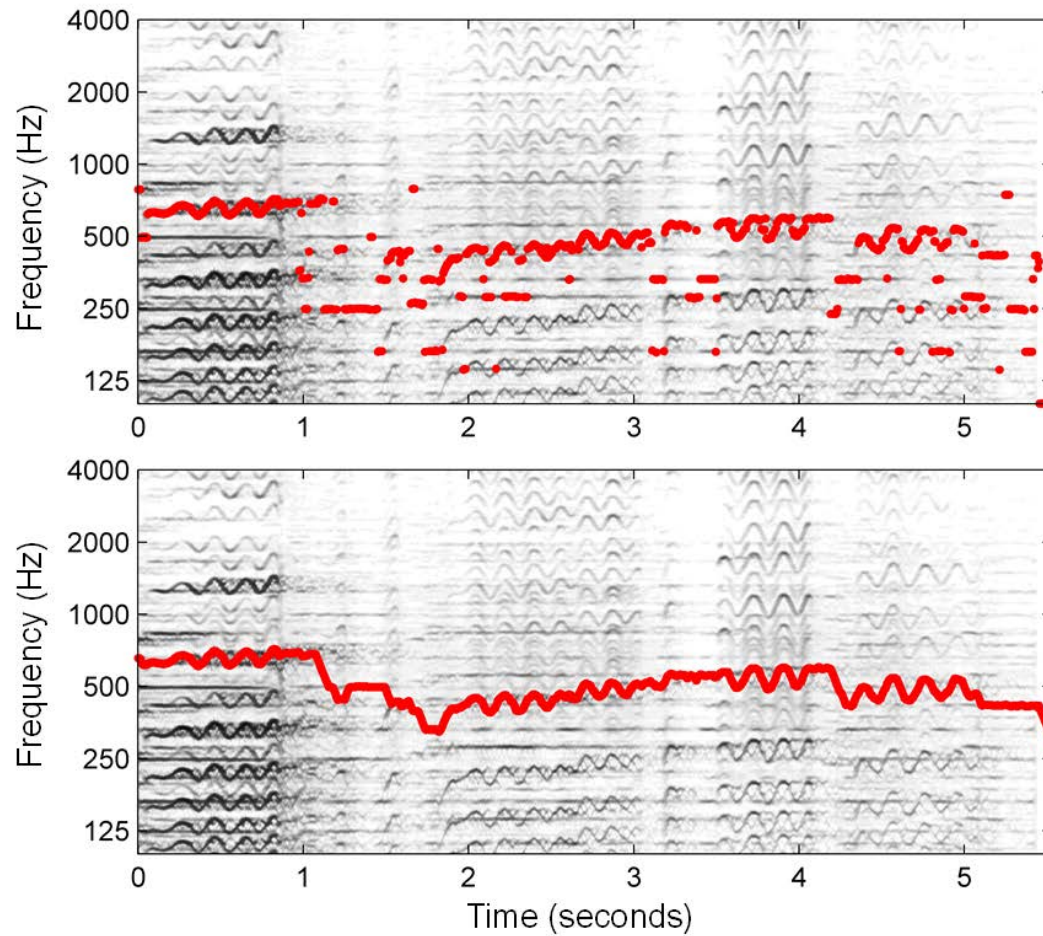
8.2 Melody Extraction

Fig. 8.15



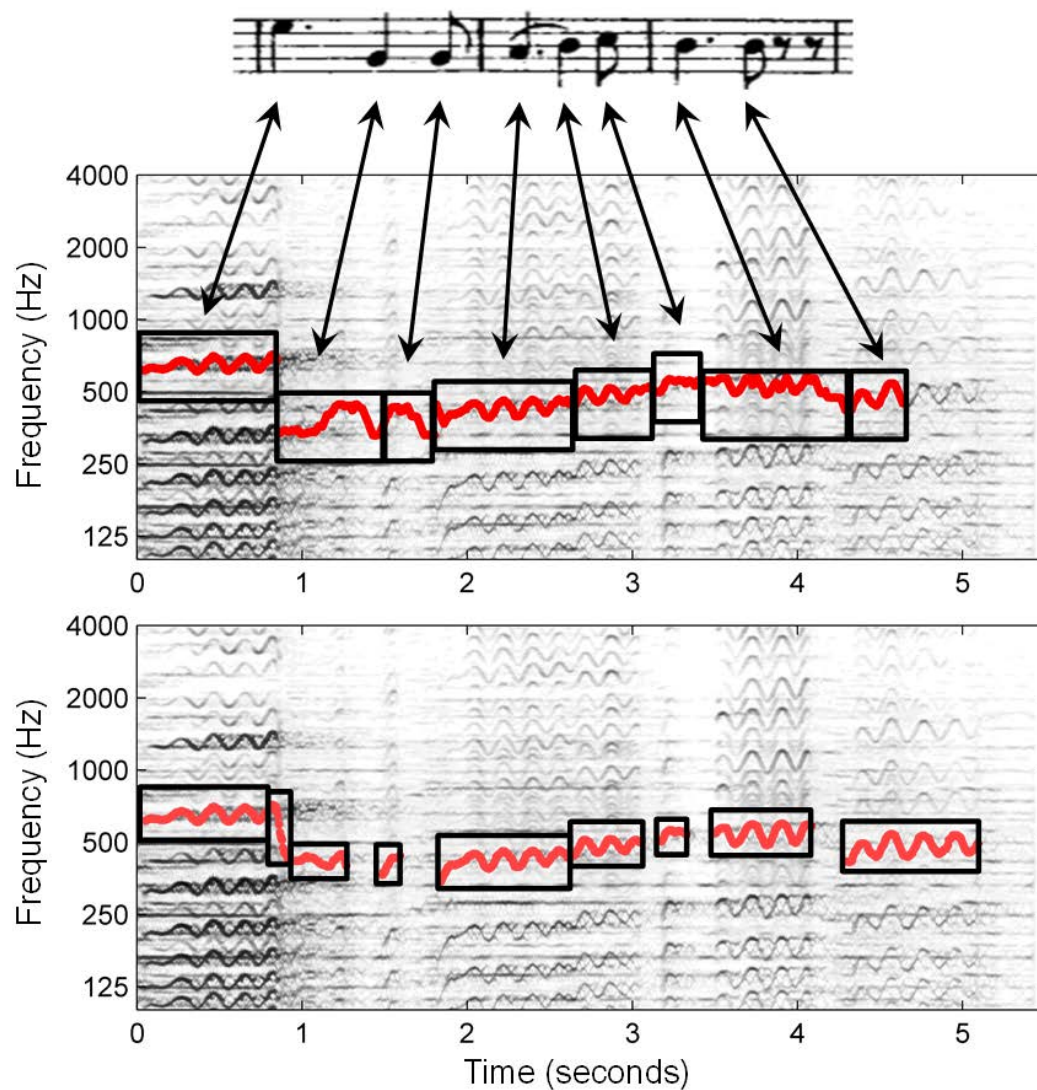
8.2 Melody Extraction

Fig. 8.16



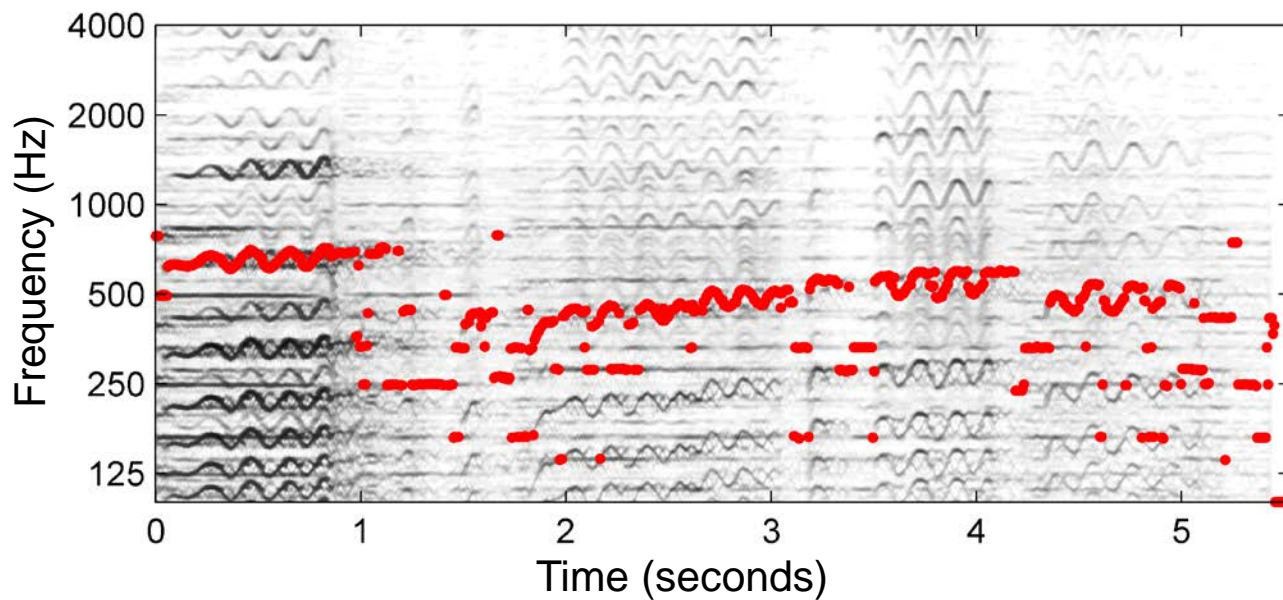
8.2 Melody Extraction

Fig. 8.17



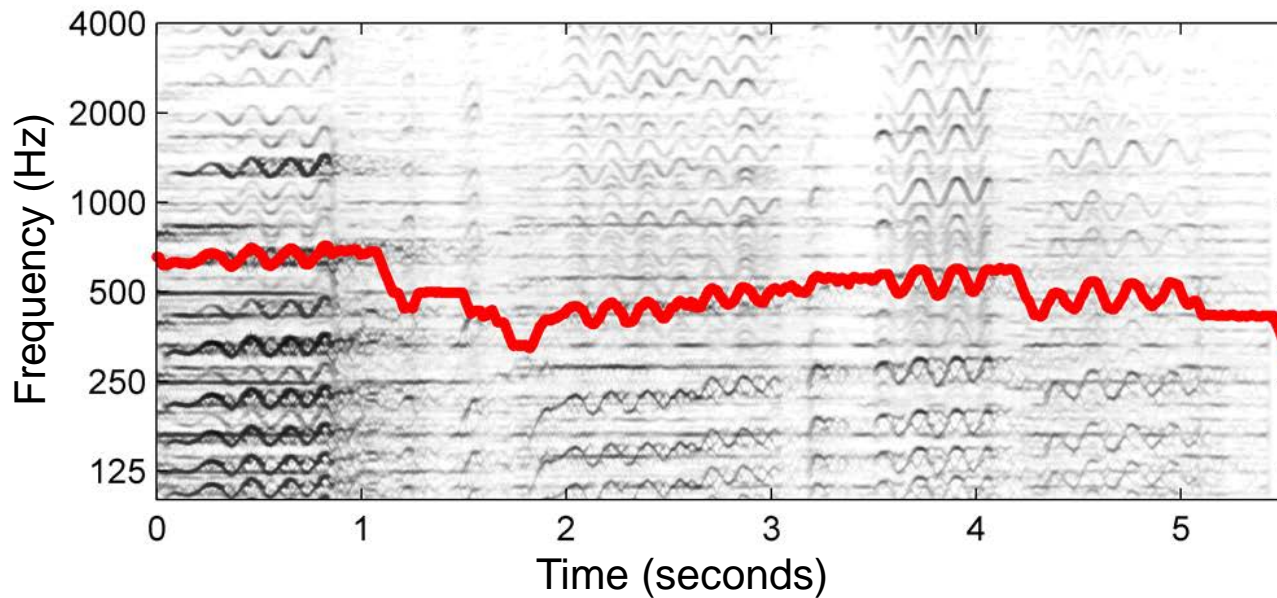
8.2 Melody Extraction

Fig. 8.16



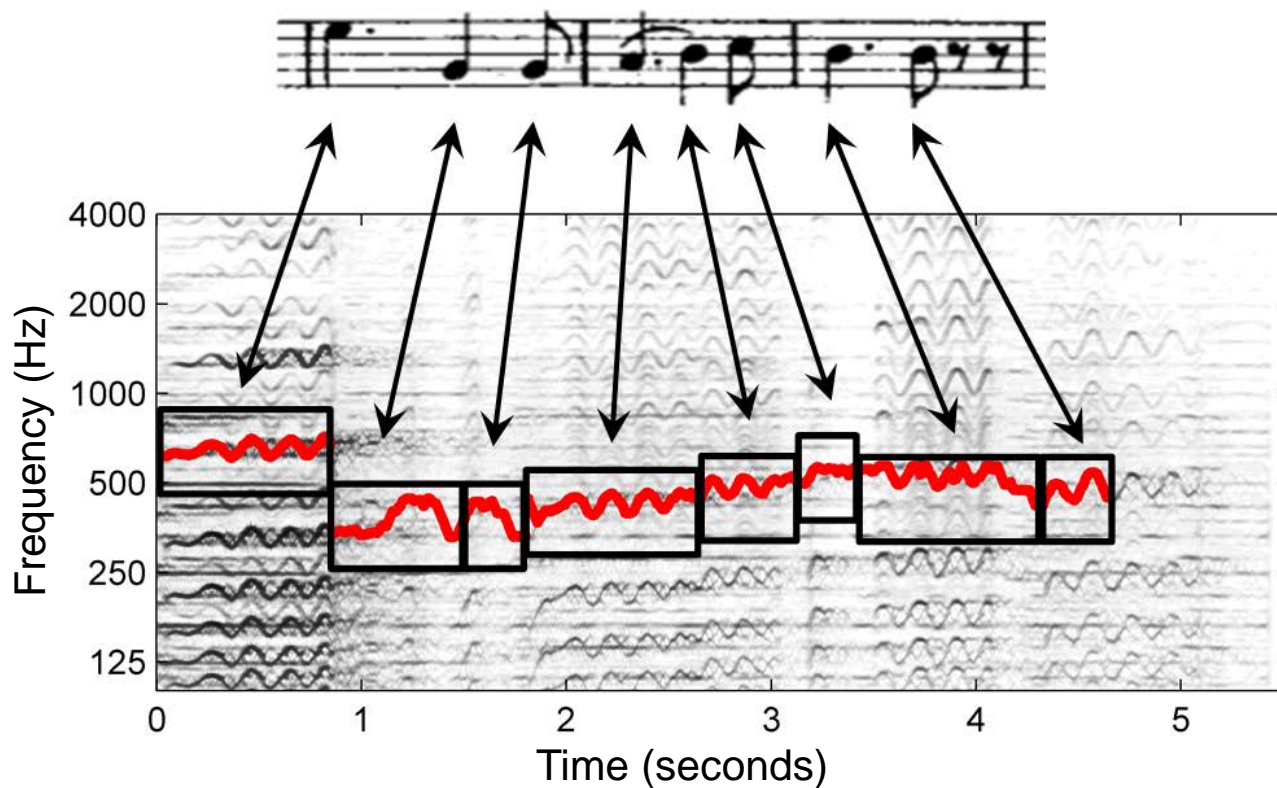
8.2 Melody Extraction

Fig. 8.16



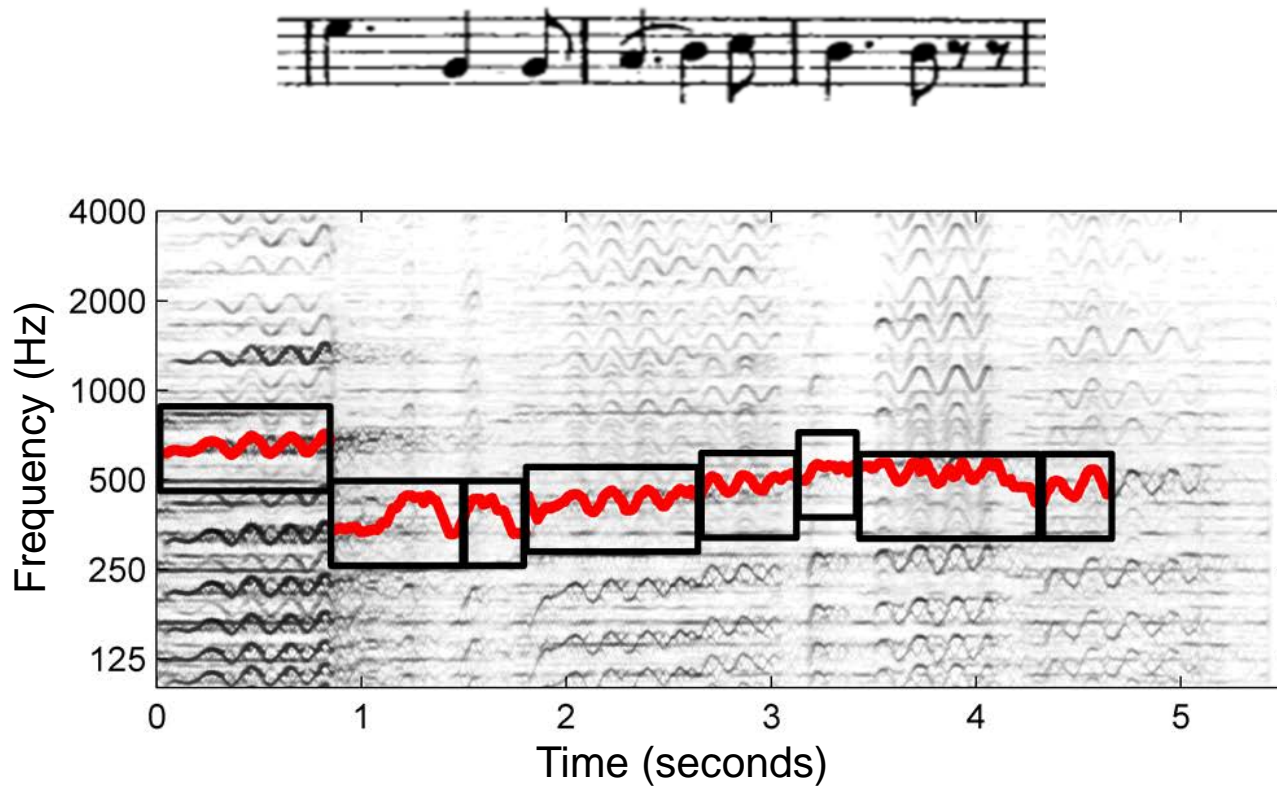
8.2 Melody Extraction

Fig. 8.17



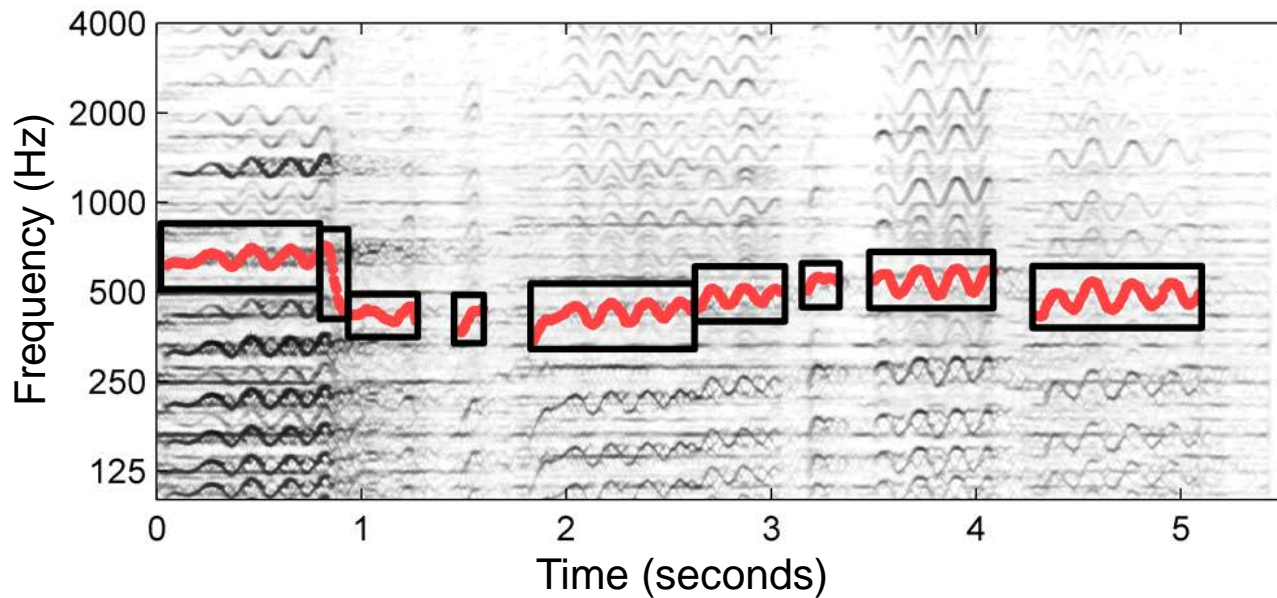
8.2 Melody Extraction

Fig. 8.17



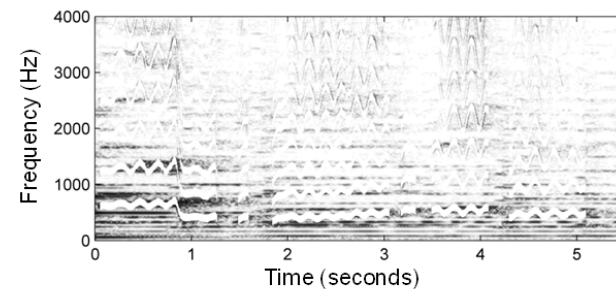
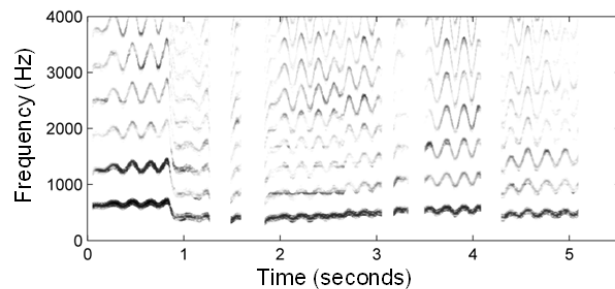
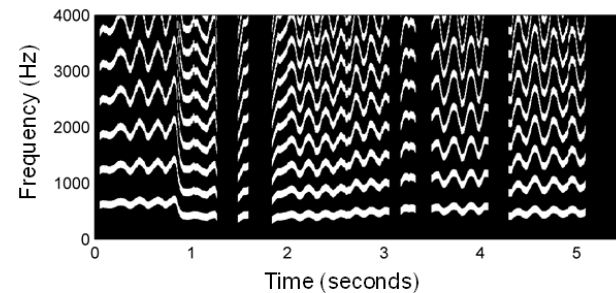
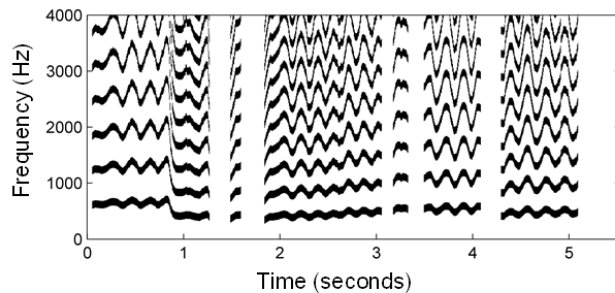
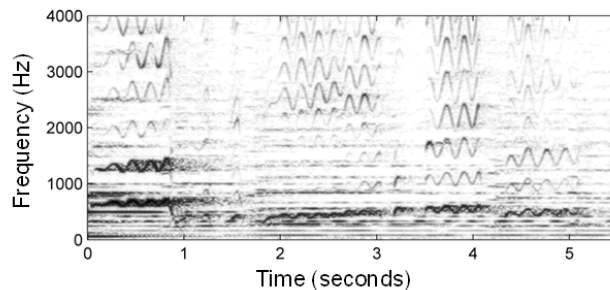
8.2 Melody Extraction

Fig. 8.17



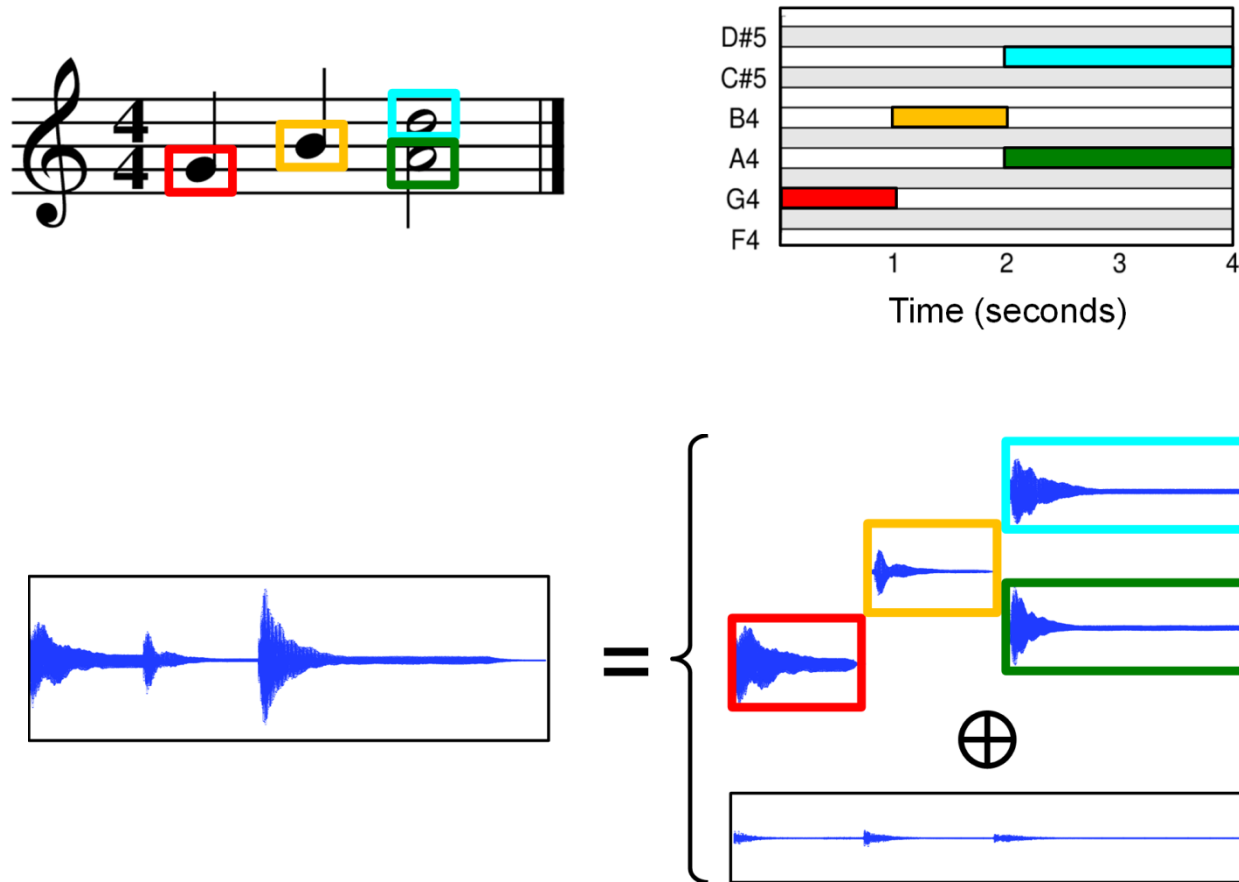
8.2 Melody Extraction

Fig. 8.18



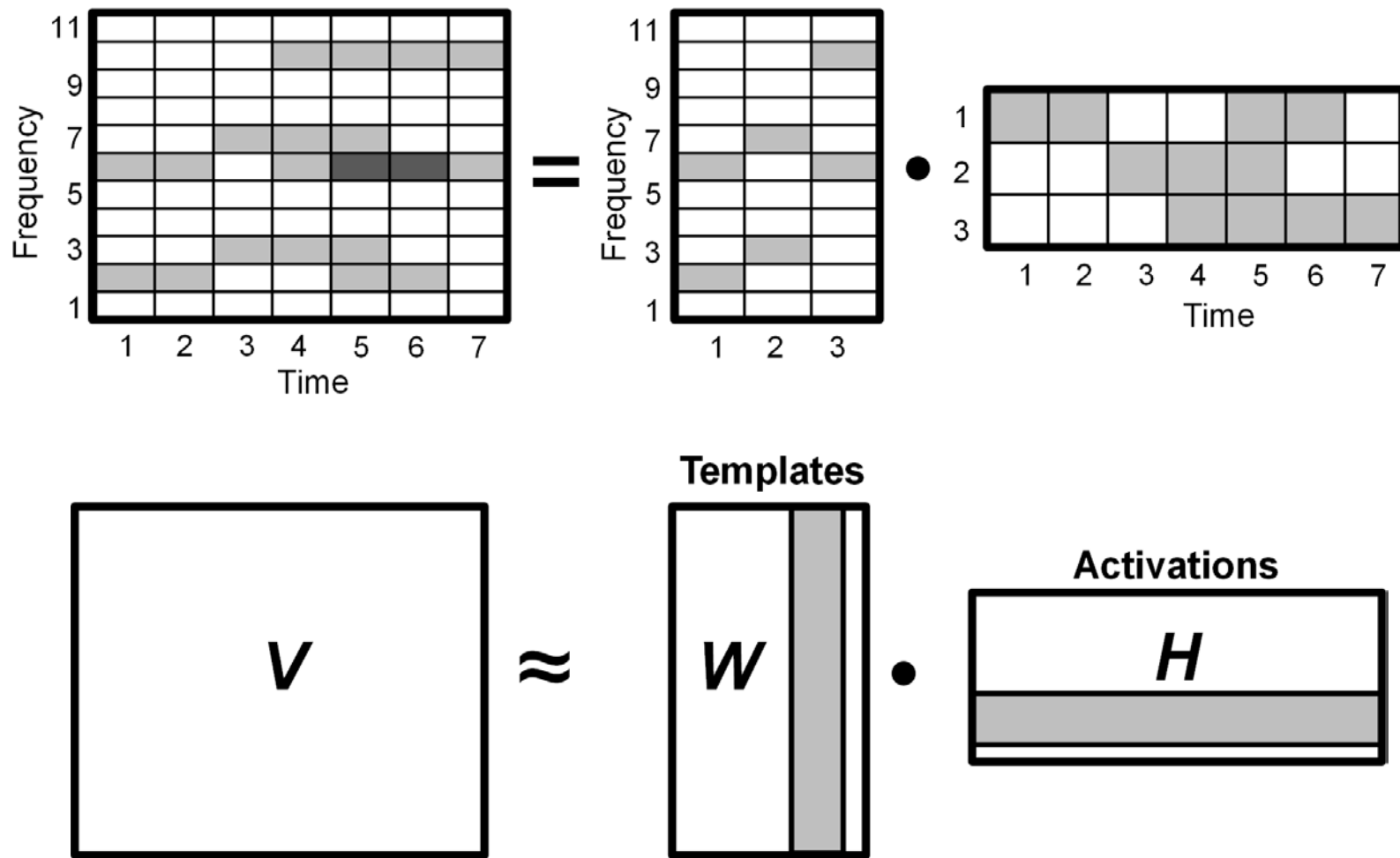
8.3 NMF-Based Audio Decomposition

Fig. 8.19



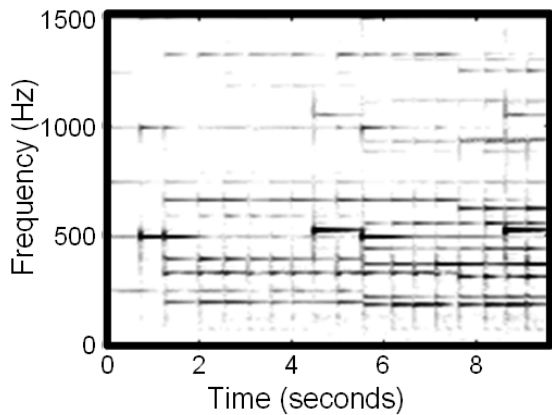
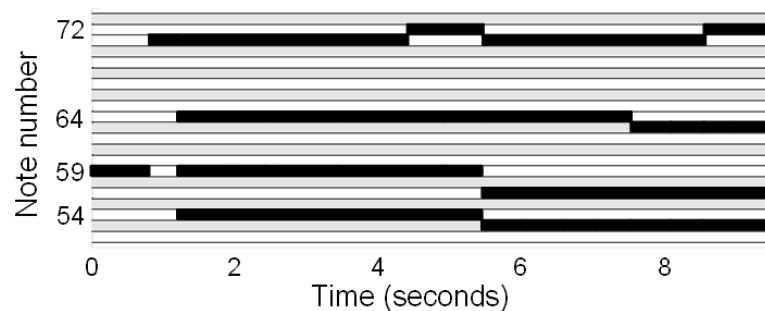
8.3 NMF-Based Audio Decomposition

Fig. 8.20

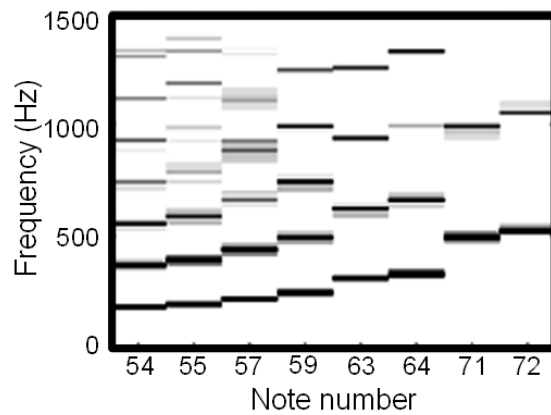


8.3 NMF-Based Audio Decomposition

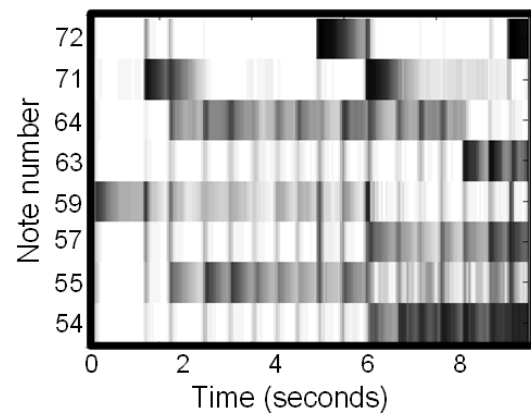
Fig. 8.21



≈

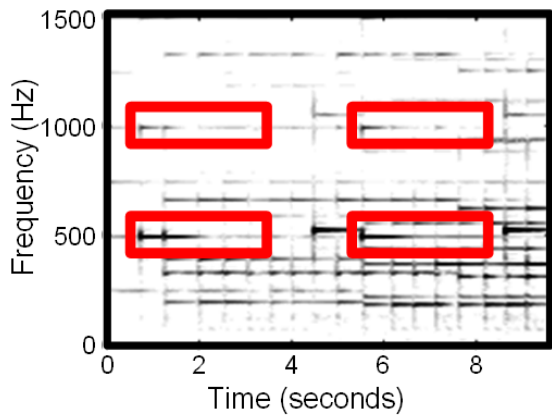
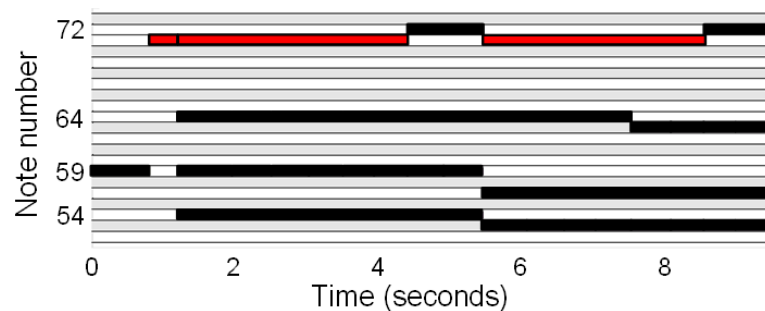


•

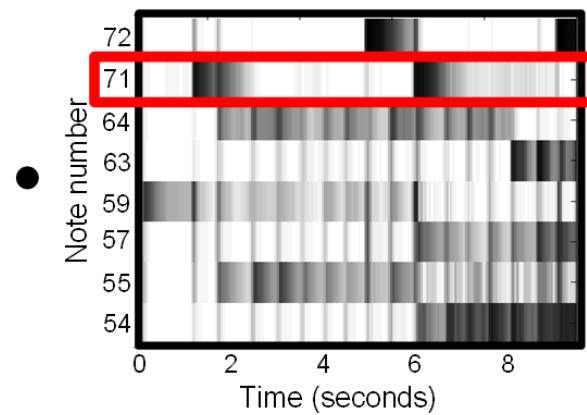
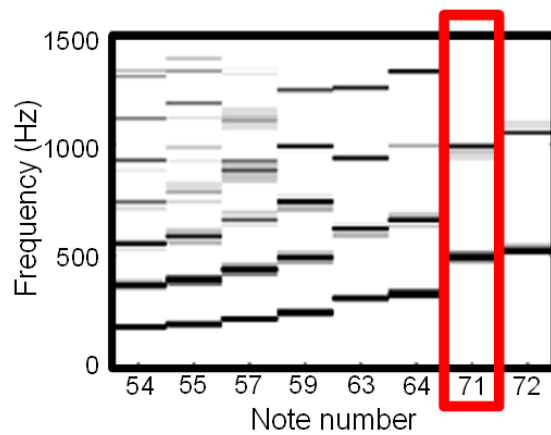


8.3 NMF-Based Audio Decomposition

Fig. 8.21



\approx



8.3 NMF-Based Audio Decomposition

Table 8.1

Algorithm: NMF ($V \approx WH$)

Input: Nonnegative matrix V of size $K \times N$
Rank parameter $R \in \mathbb{N}$
Threshold ε used as stop criterion

Output: Nonnegative template matrix W of size $K \times R$
Nonnegative activation matrix H of size $R \times N$

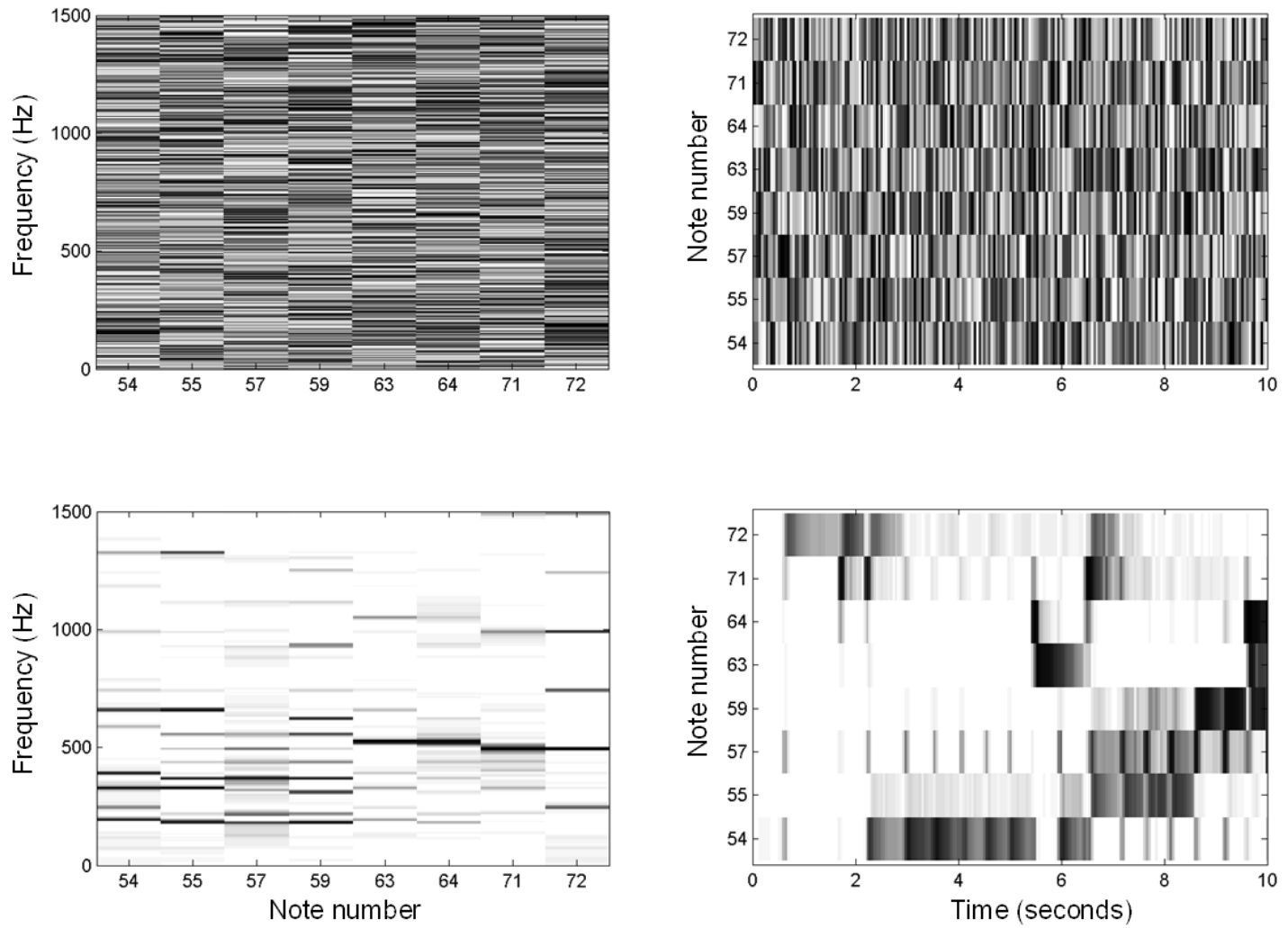
Procedure: Define nonnegative matrices $W^{(0)}$ and $H^{(0)}$ by some random or informed initialization. Furthermore set $\ell = 0$. Apply the following update rules (written in matrix notation):

- (1) $H^{(\ell+1)} = H^{(\ell)} \odot (((W^{(\ell)})^\top V) \oslash ((W^{(\ell)})^\top W^{(\ell)} H^{(\ell)}))$
- (2) $W^{(\ell+1)} = W^{(\ell)} \odot ((V(H^{(\ell+1)})^\top) \oslash (W^{(\ell)} H^{(\ell+1)} (H^{(\ell+1)})^\top))$
- (3) Increase ℓ by one.

Repeat the steps (1) to (3) until $\|H^{(\ell)} - H^{(\ell-1)}\| \leq \varepsilon$ and $\|W^{(\ell)} - W^{(\ell-1)}\| \leq \varepsilon$ (or until some other stop criterion is fulfilled). Finally, set $H = H^{(\ell)}$ and $W = W^{(\ell)}$.

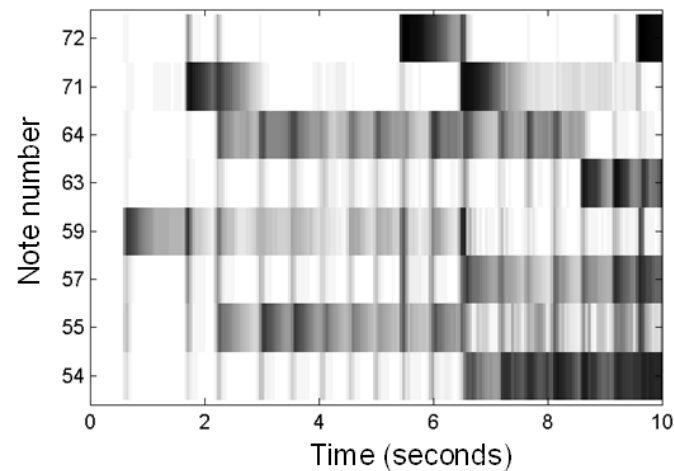
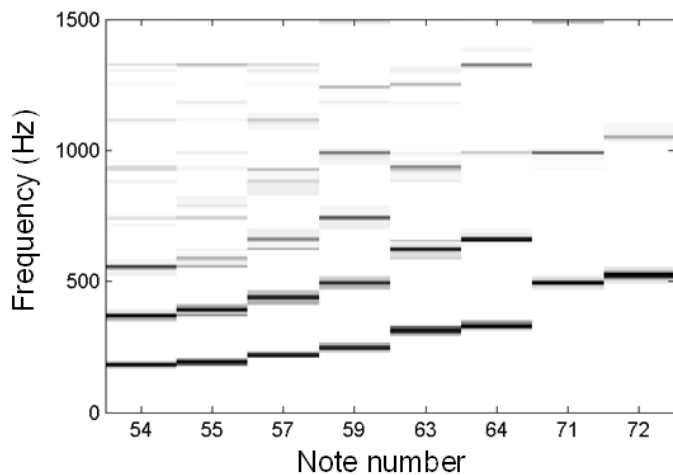
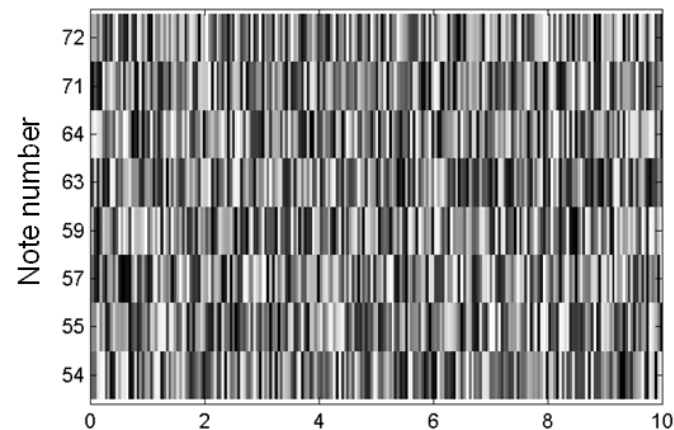
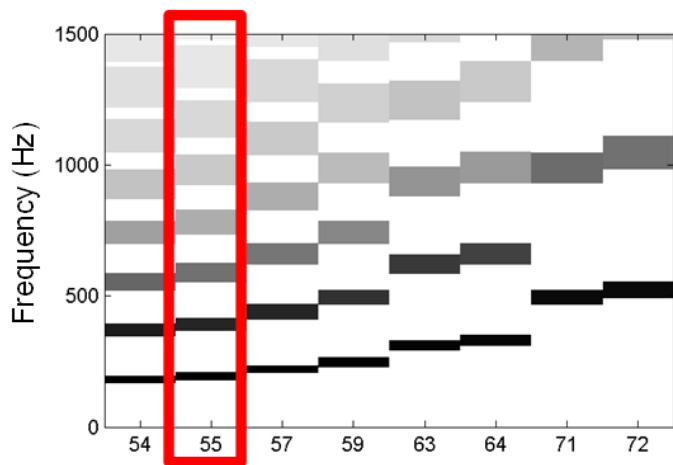
8.3 NMF-Based Audio Decomposition

Fig. 8.22



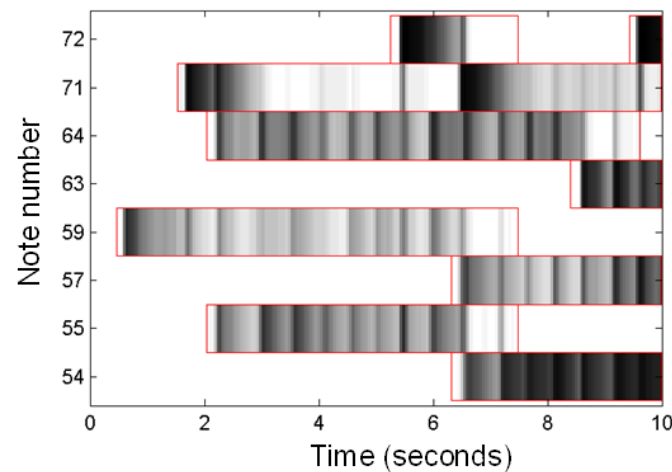
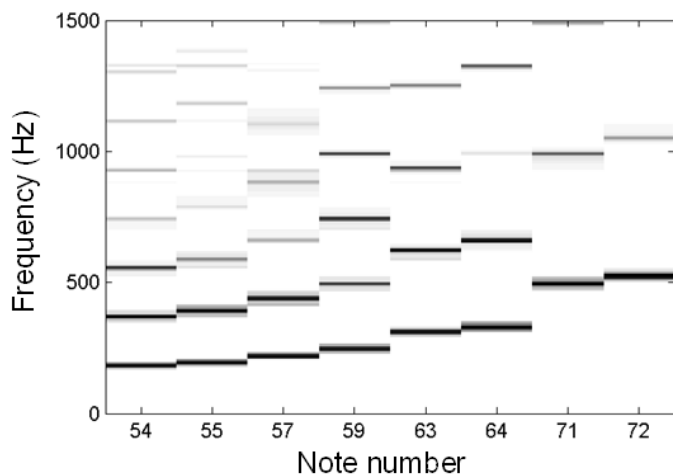
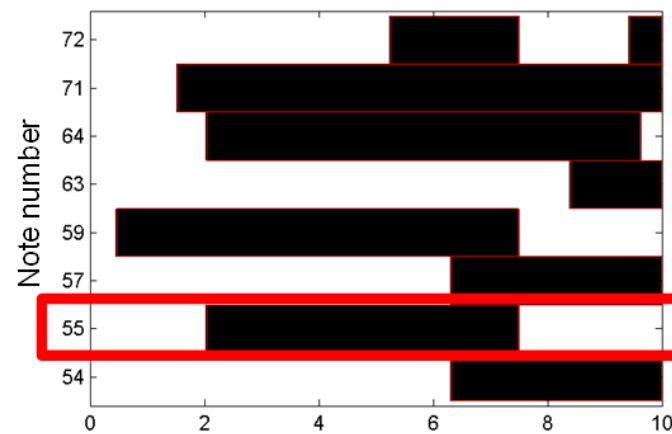
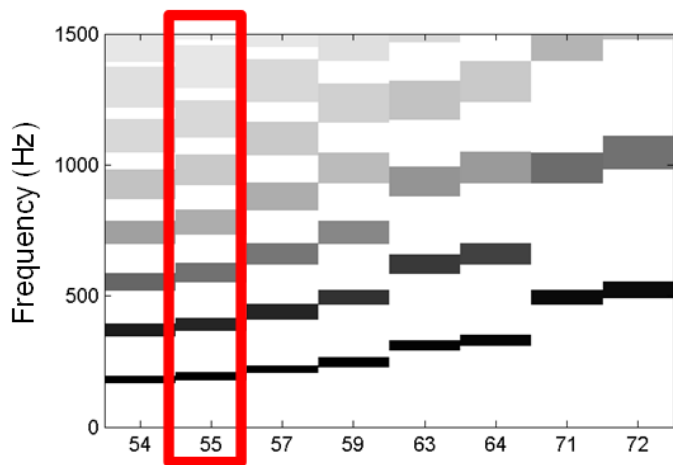
8.3 NMF-Based Audio Decomposition

Fig. 8.23



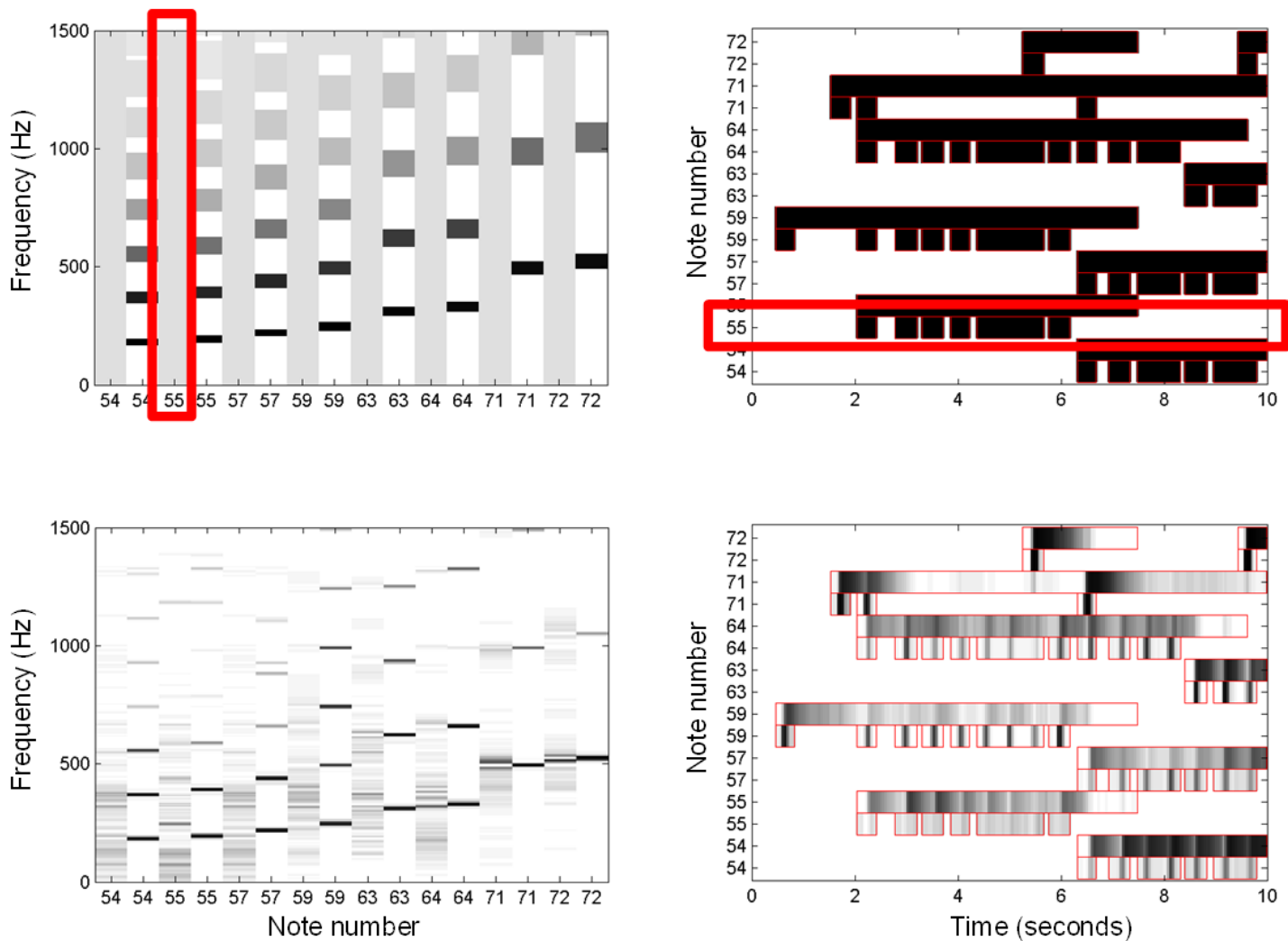
8.3 NMF-Based Audio Decomposition

Fig. 8.24



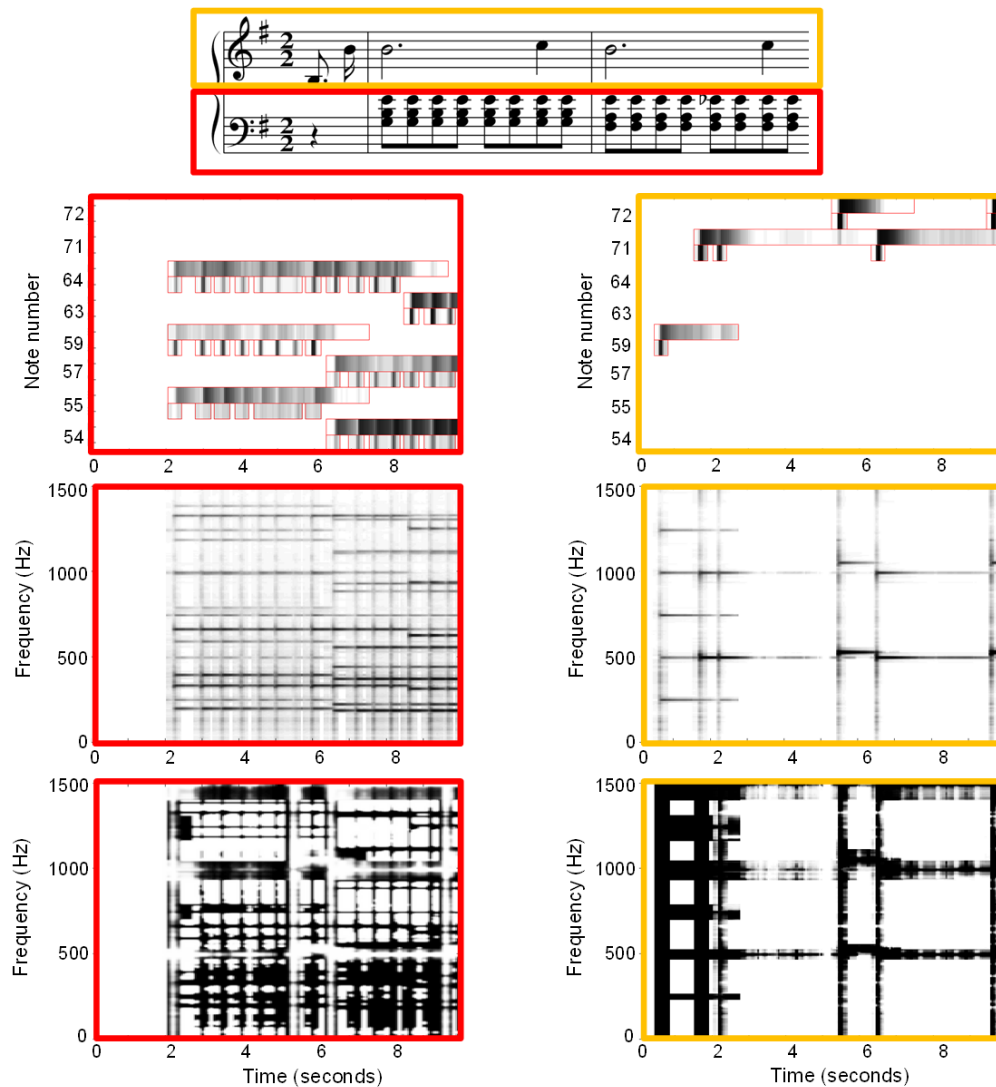
8.3 NMF-Based Audio Decomposition

Fig. 8.25



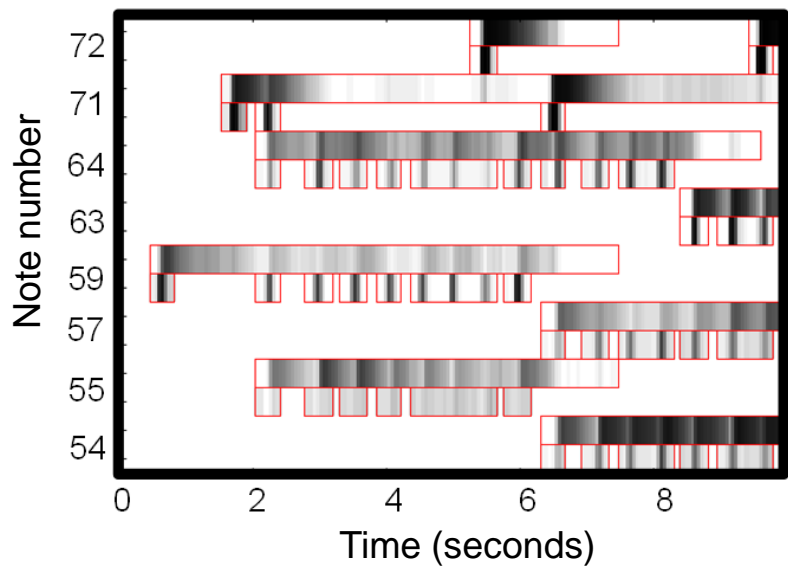
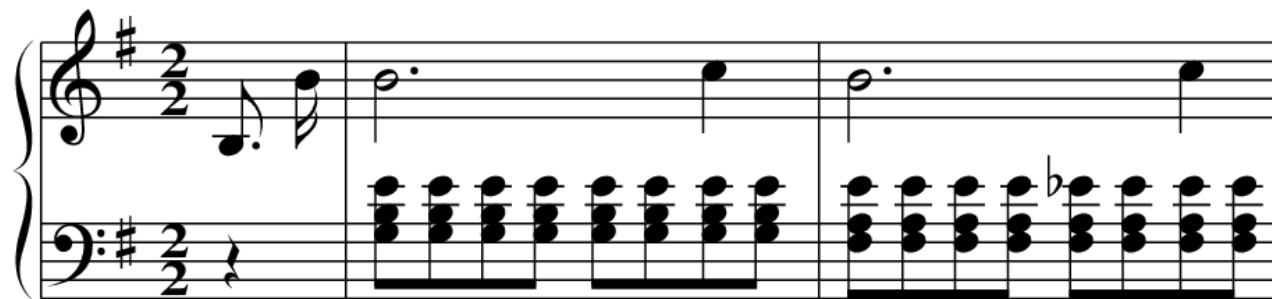
8.3 NMF-Based Audio Decomposition

Fig. 8.26



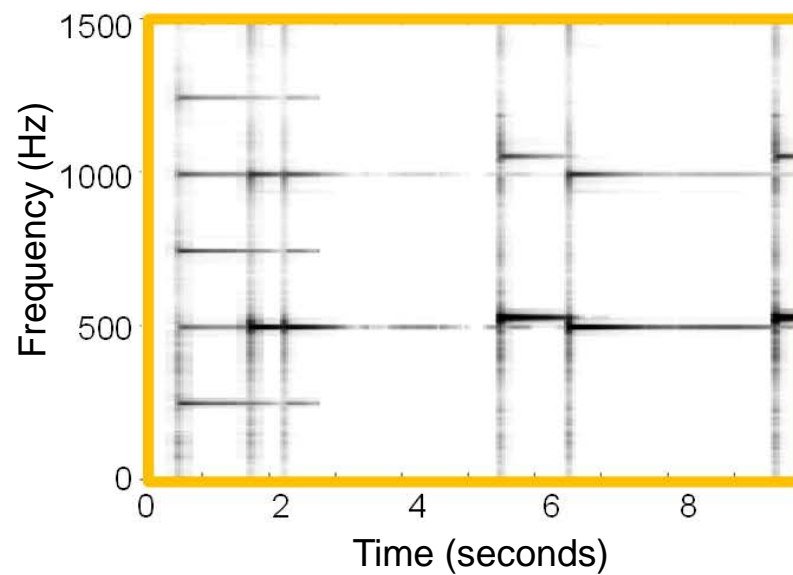
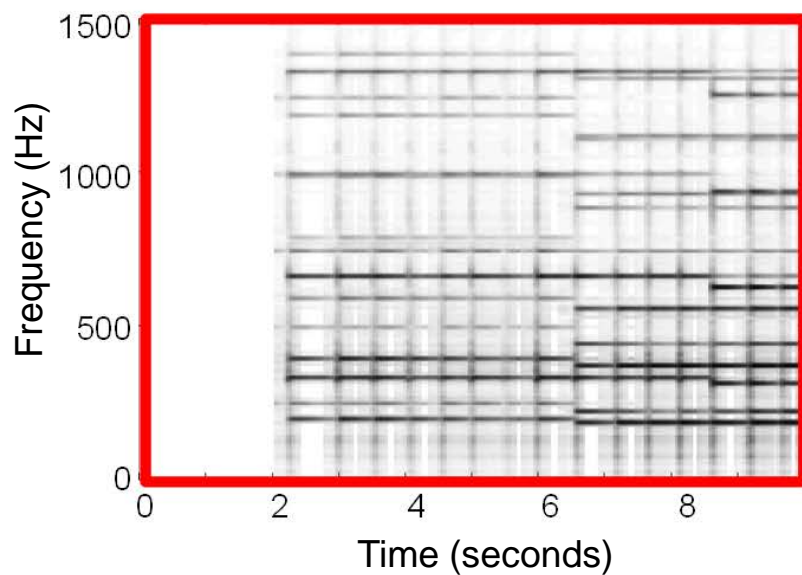
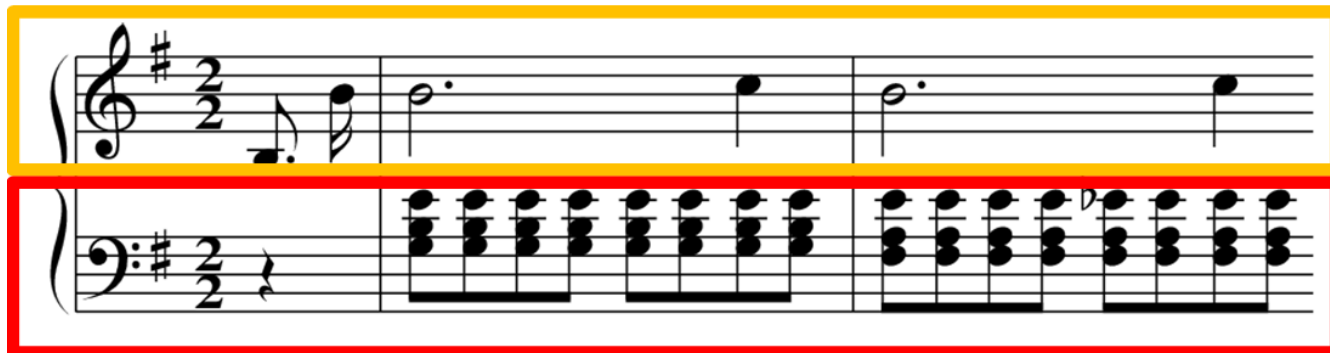
8.3 NMF-Based Audio Decomposition

Fig. 8.26



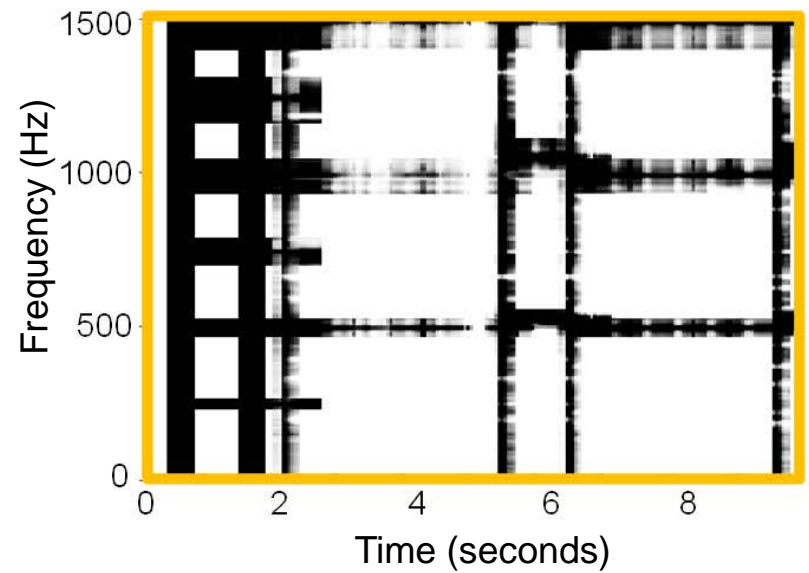
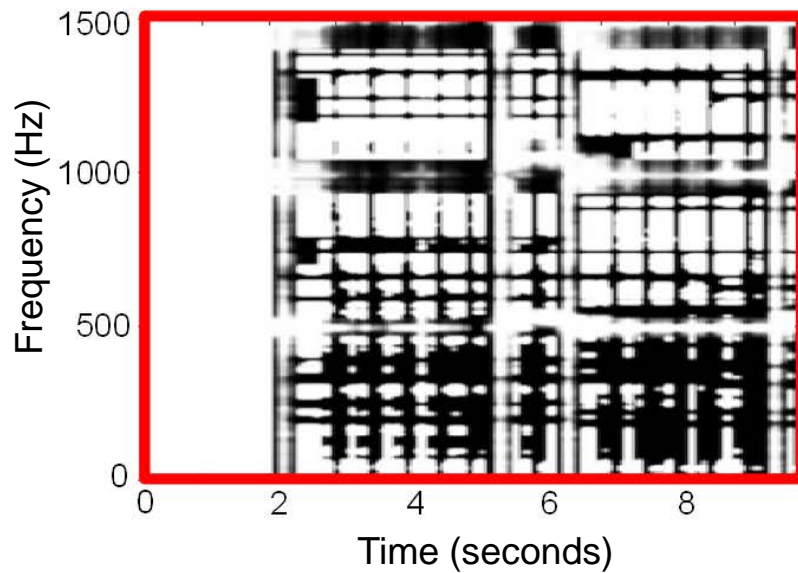
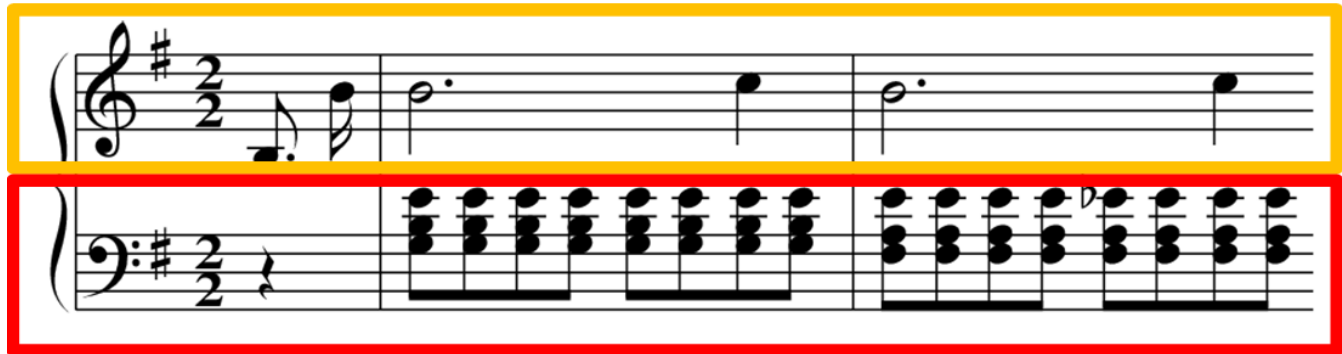
8.3 NMF-Based Audio Decomposition

Fig. 8.26



8.3 NMF-Based Audio Decomposition

Fig. 8.26



8.3 NMF-Based Audio Decomposition

Fig. 8.27

