Fundamentals of Music Processing

Chapter 3: Music Synchronization

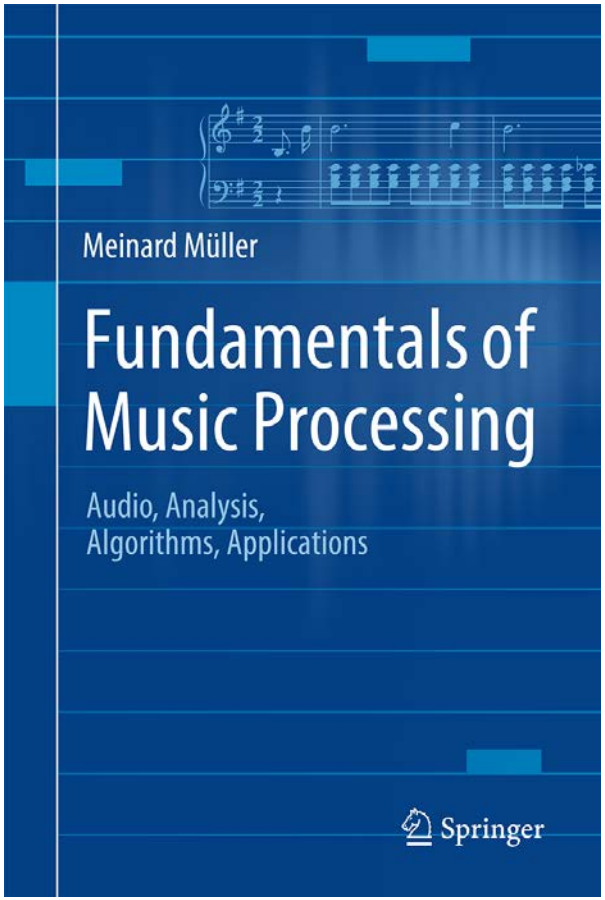
Meinard Müller
International Audio Laboratories Erlangen

www.music-processing.de

Springer

AUDIO LABS

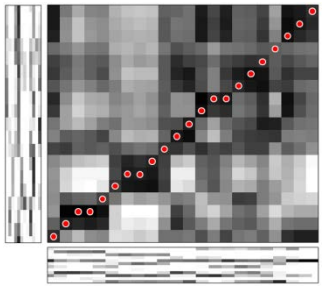
Book: Fundamentals of Music Processing



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

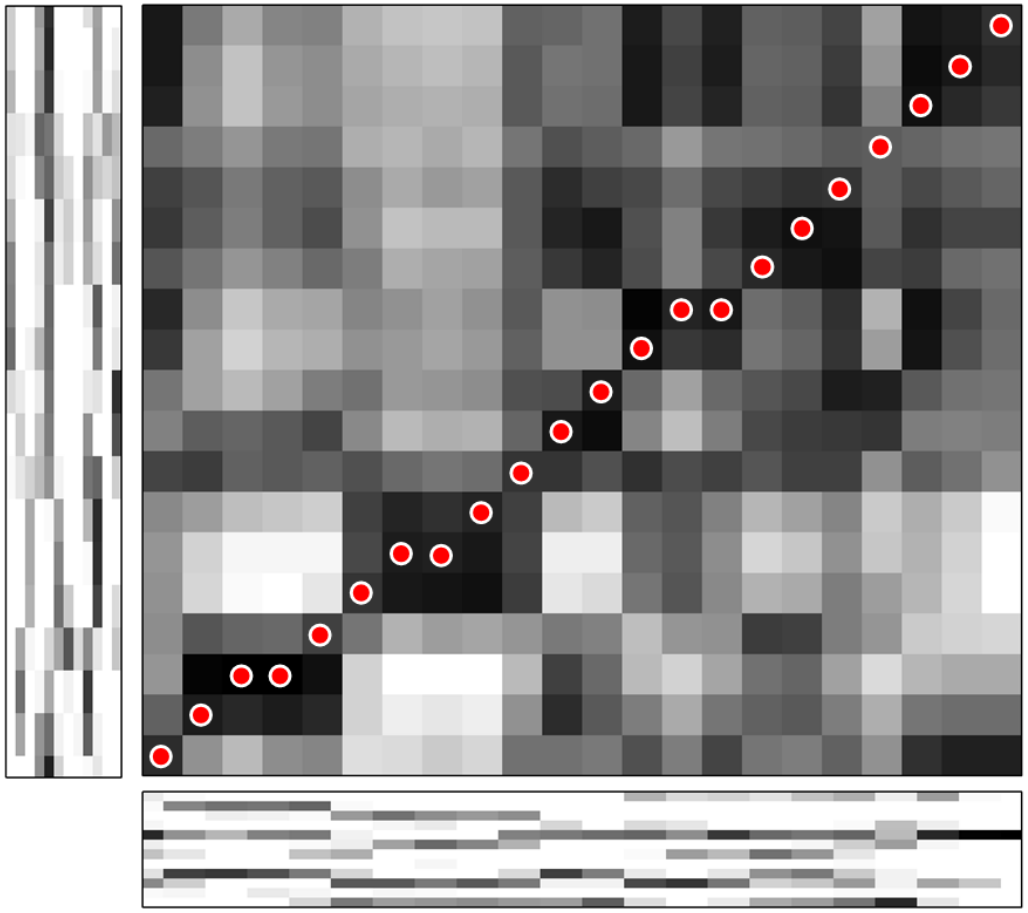
Chapter 3: Music Synchronization



As a first music processing task, we study in Chapter 3 the problem of music synchronization. The objective is to temporally align compatible representations of the same piece of music. Considering this scenario, we explain the need for musically informed audio features. In particular, we introduce the concept of chroma-based music features, which capture properties that are related to harmony and melody. Furthermore, we study an alignment technique known as dynamic time warping (DTW), a concept that is applicable for the analysis of general time series. For its efficient computation, we discuss an algorithm based on dynamic programming—a widely used method for solving a complex problem by breaking it down into a collection of simpler subproblems.

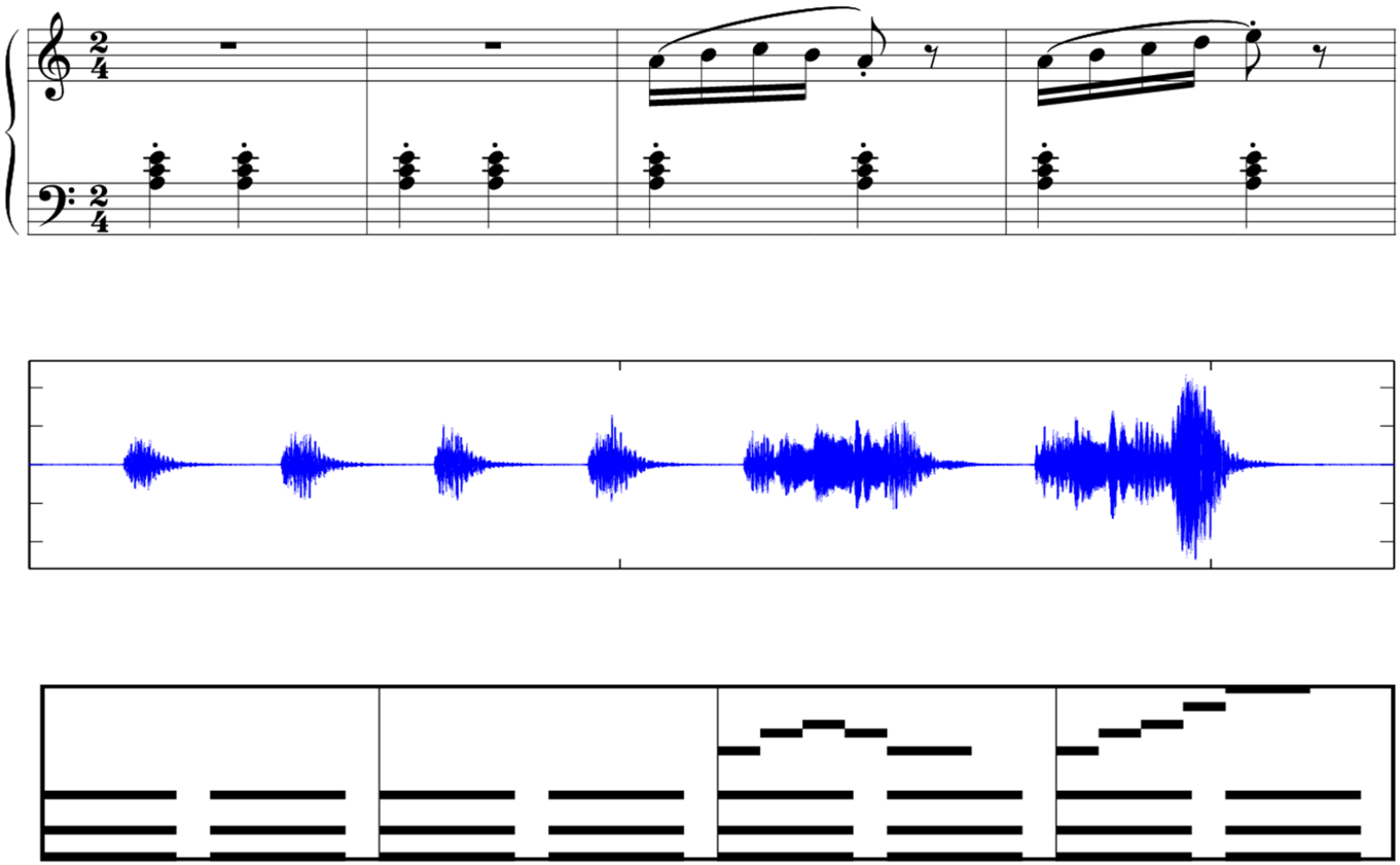
3 Music Synchronization

Teaser



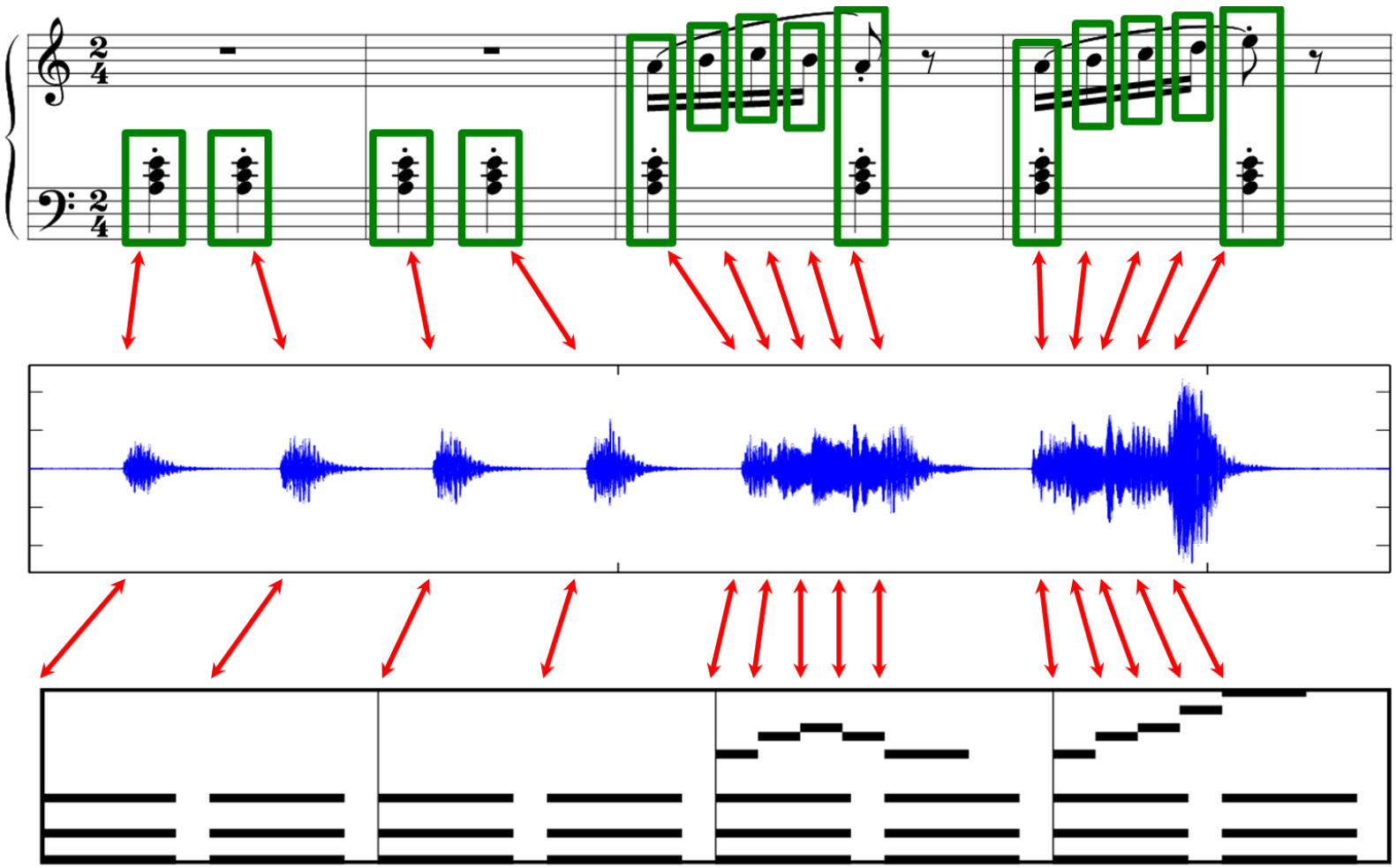
3 Music Synchronization

Fig. 3.1



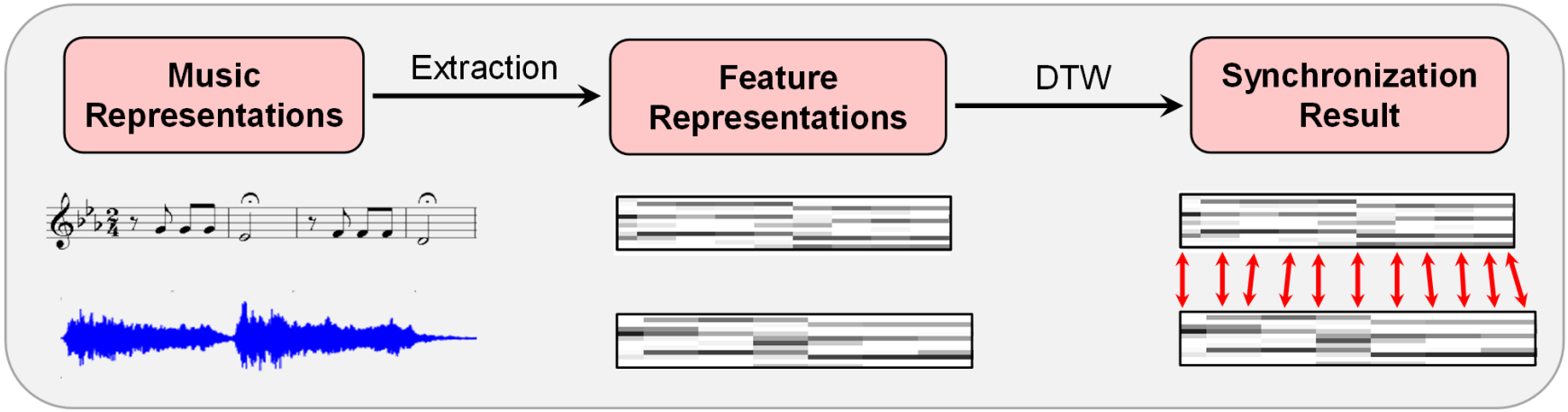
3 Music Synchronization

Fig. 3.1



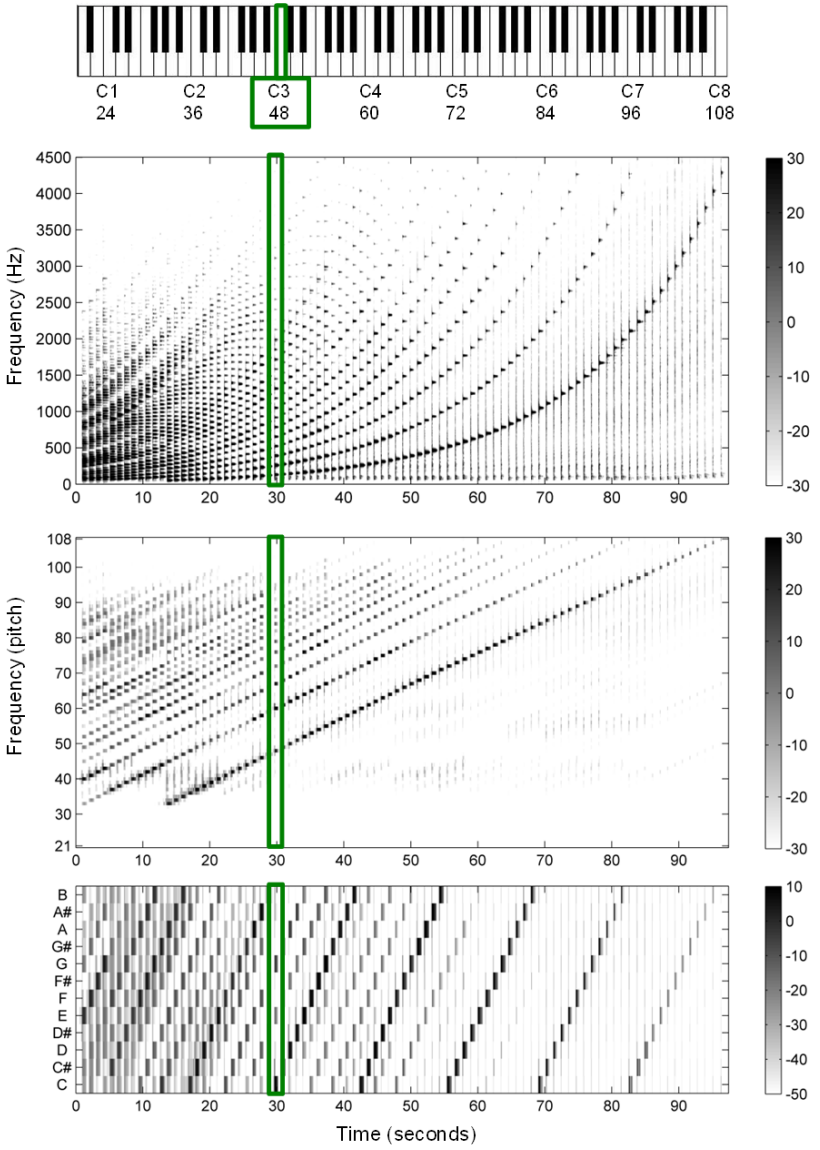
3 Music Synchronization

Fig. 3.2



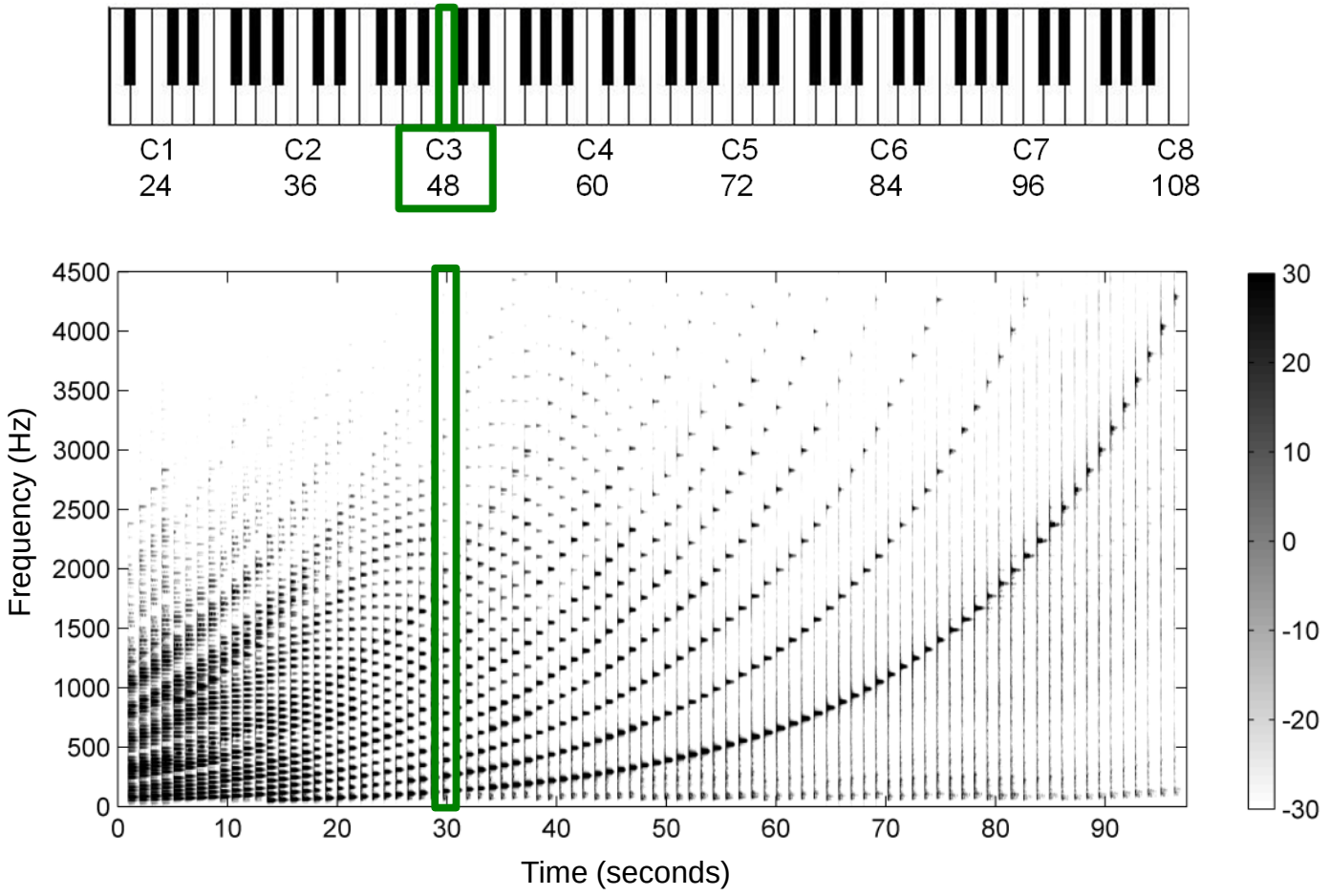
3.1 Audio Features

Fig. 3.3



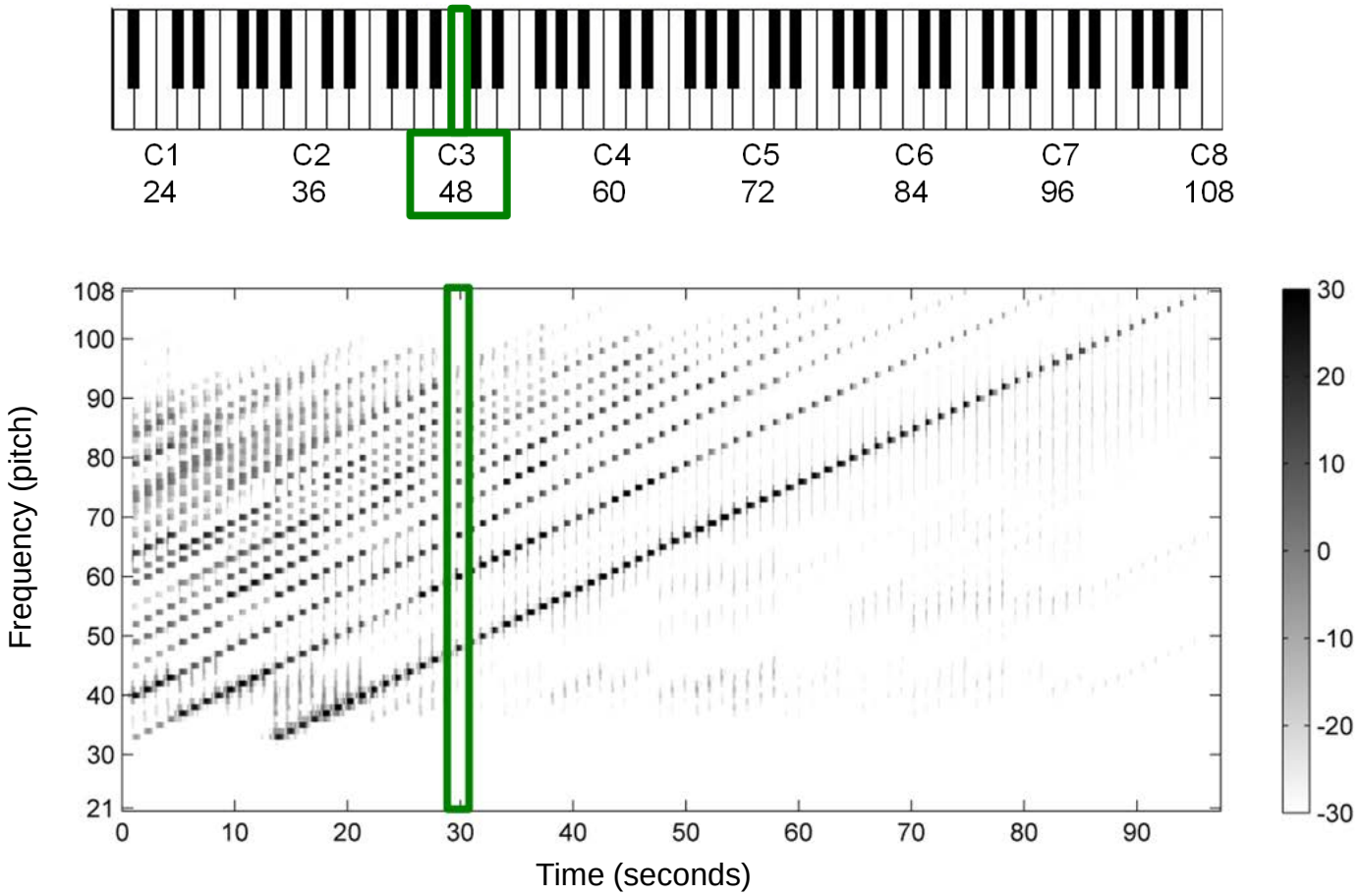
3.1 Audio Features

Fig. 3.3



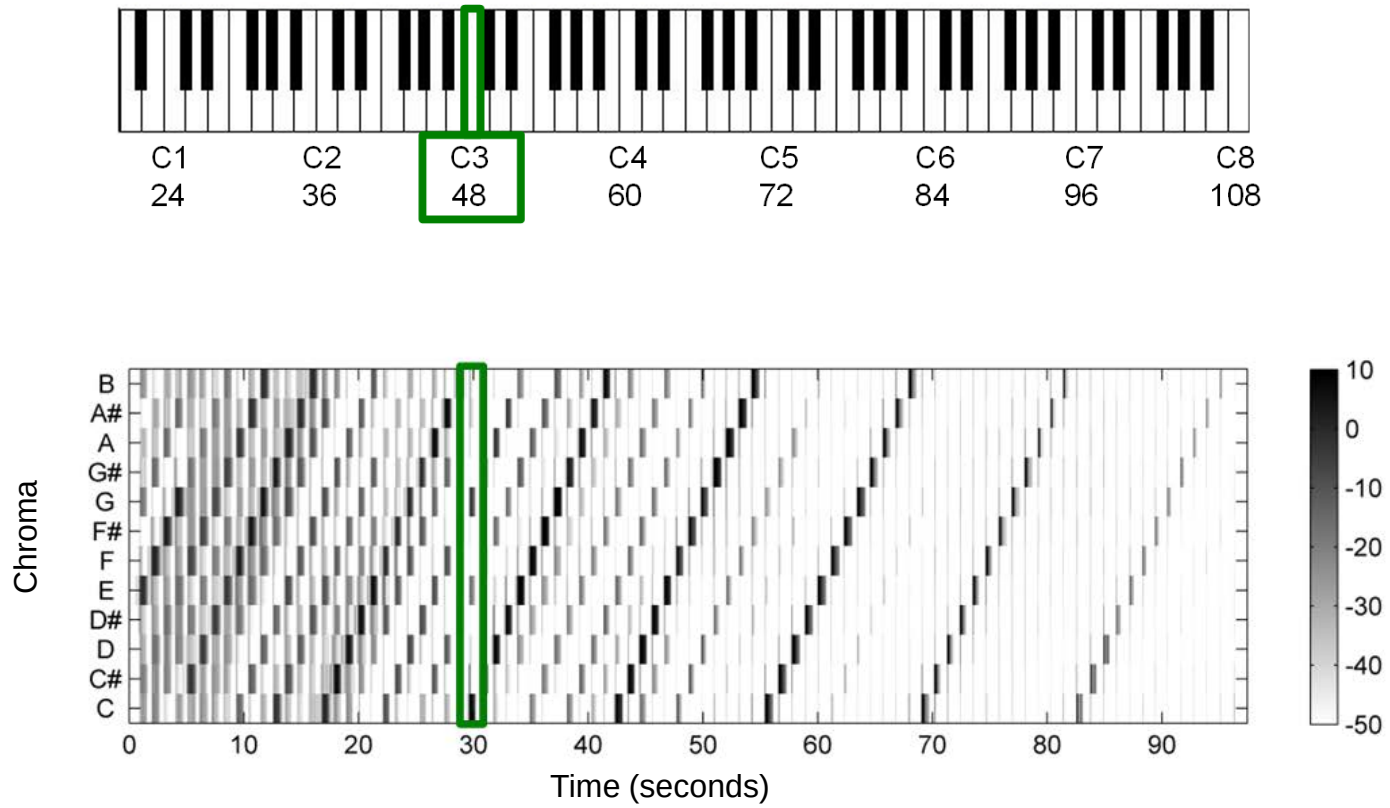
3.1 Audio Features

Fig. 3.3



3.1 Audio Features

Fig. 3.3



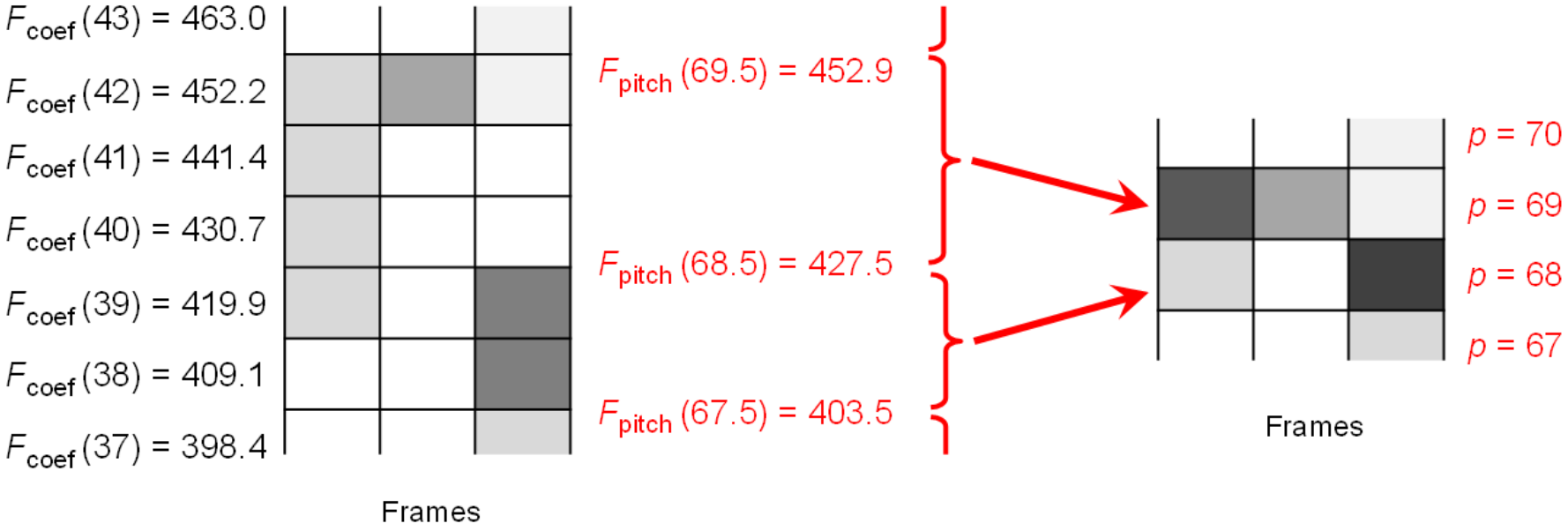
3.1 Audio Features

Table 3.1

Note	p	$F_{\text{pitch}}(p)$	$F_{\text{pitch}}(p-0.5)$	$F_{\text{pitch}}(p+0.5)$	$\text{BW}(p)$
C4	60	261.63	254.18	269.29	15.11
C♯4	61	277.18	269.29	285.30	16.01
D4	62	293.66	285.30	302.27	16.97
D♯4	63	311.13	302.27	320.24	17.97
E4	64	329.63	320.24	339.29	19.04
F4	65	349.23	339.29	359.46	20.18
F♯4	66	369.99	359.46	380.84	21.37
G4	67	392.00	380.84	403.48	22.65
G♯4	68	415.30	403.48	427.47	23.99
A4	69	440.00	427.47	452.89	25.41
A♯4	70	466.16	452.89	479.82	26.93
B4	71	493.88	479.82	508.36	28.53
C5	72	523.25	508.36	538.58	30.23

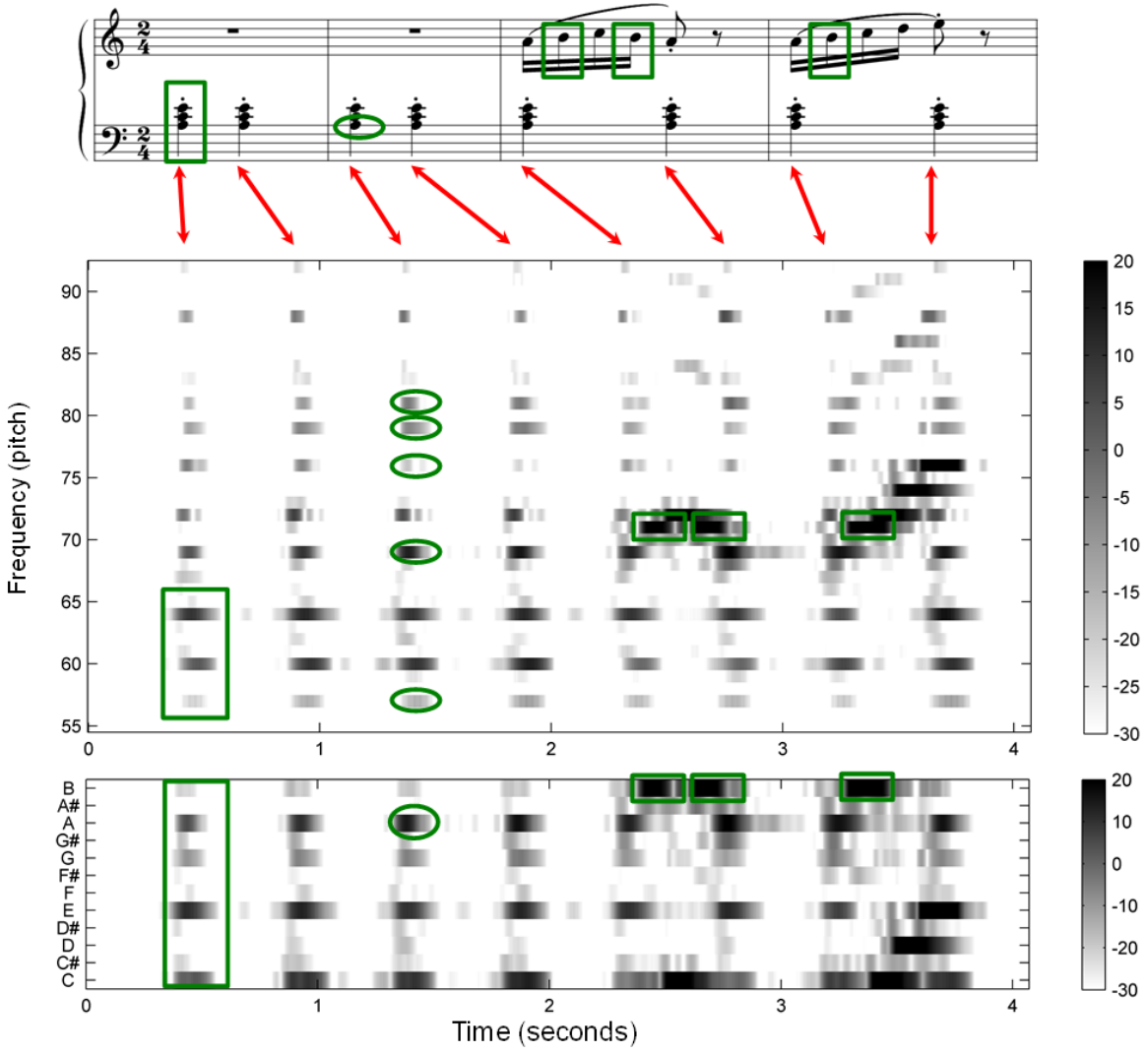
3.1 Audio Features

Fig. 3.4



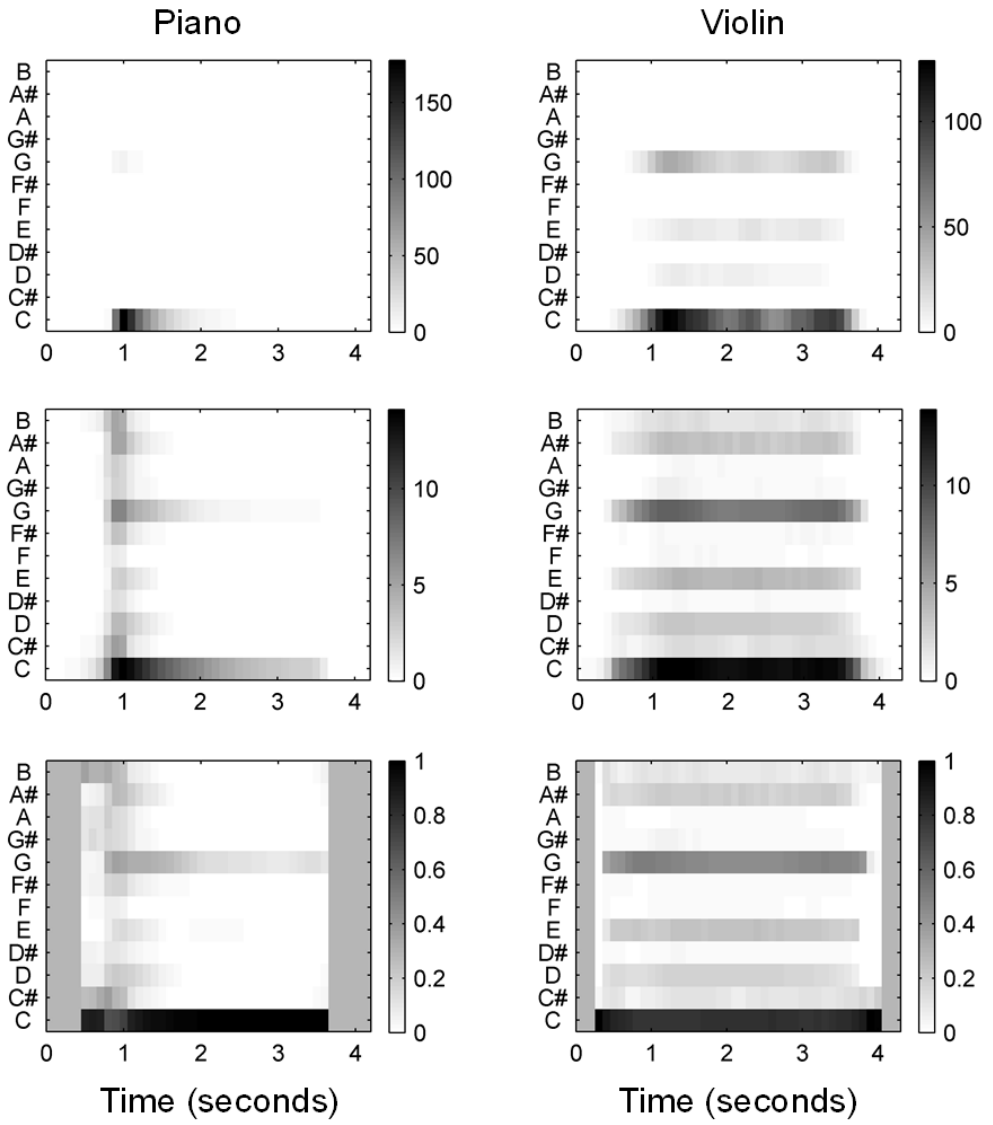
3.1 Audio Features

Fig. 3.5



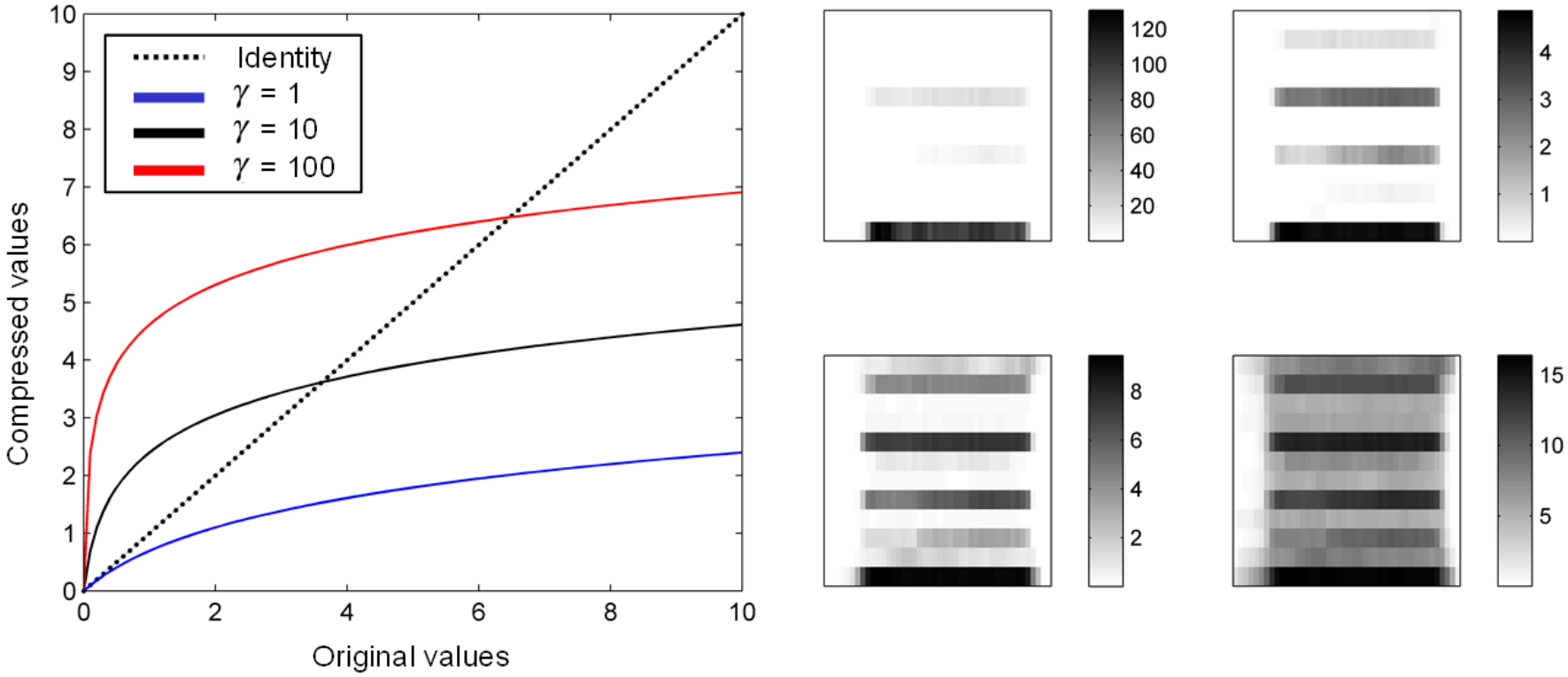
3.1 Audio Features

Fig. 3.6



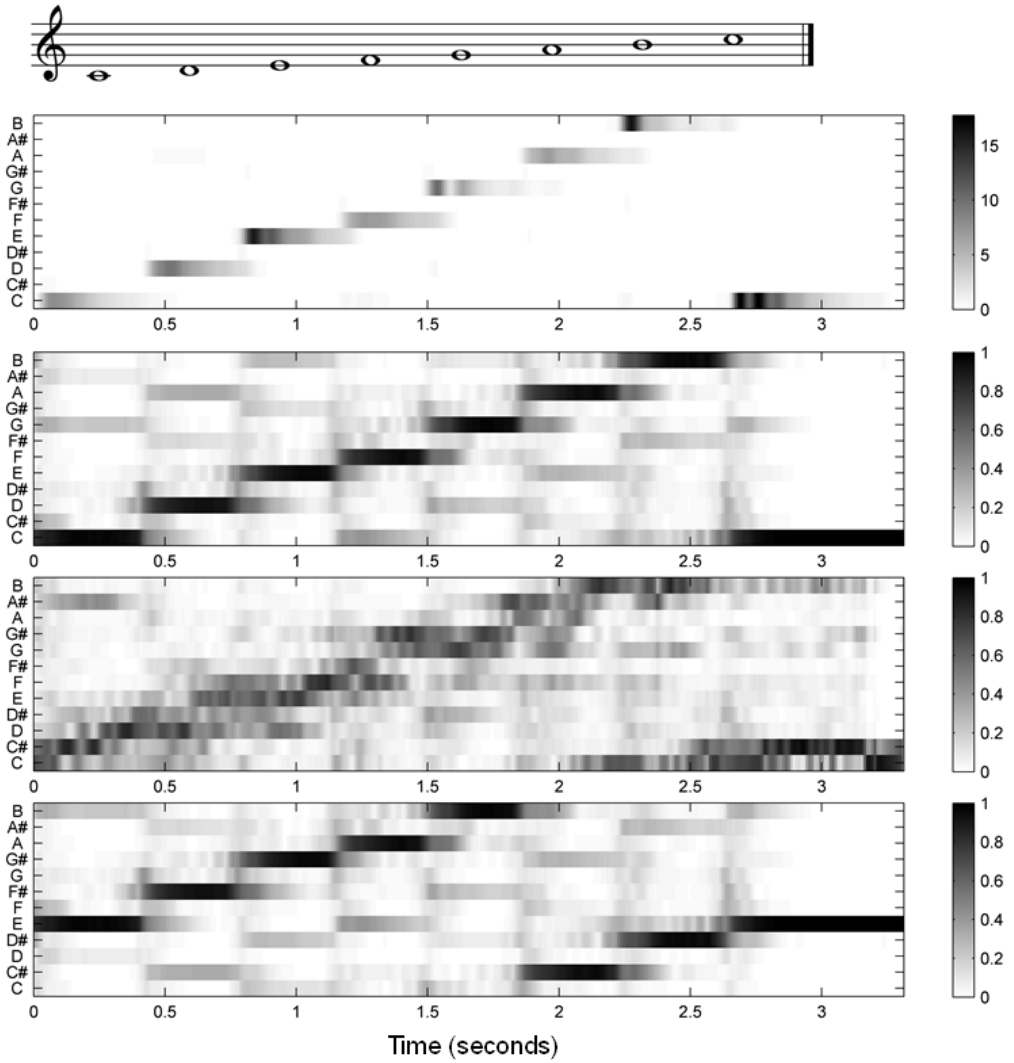
3.1 Audio Features

Fig. 3.7



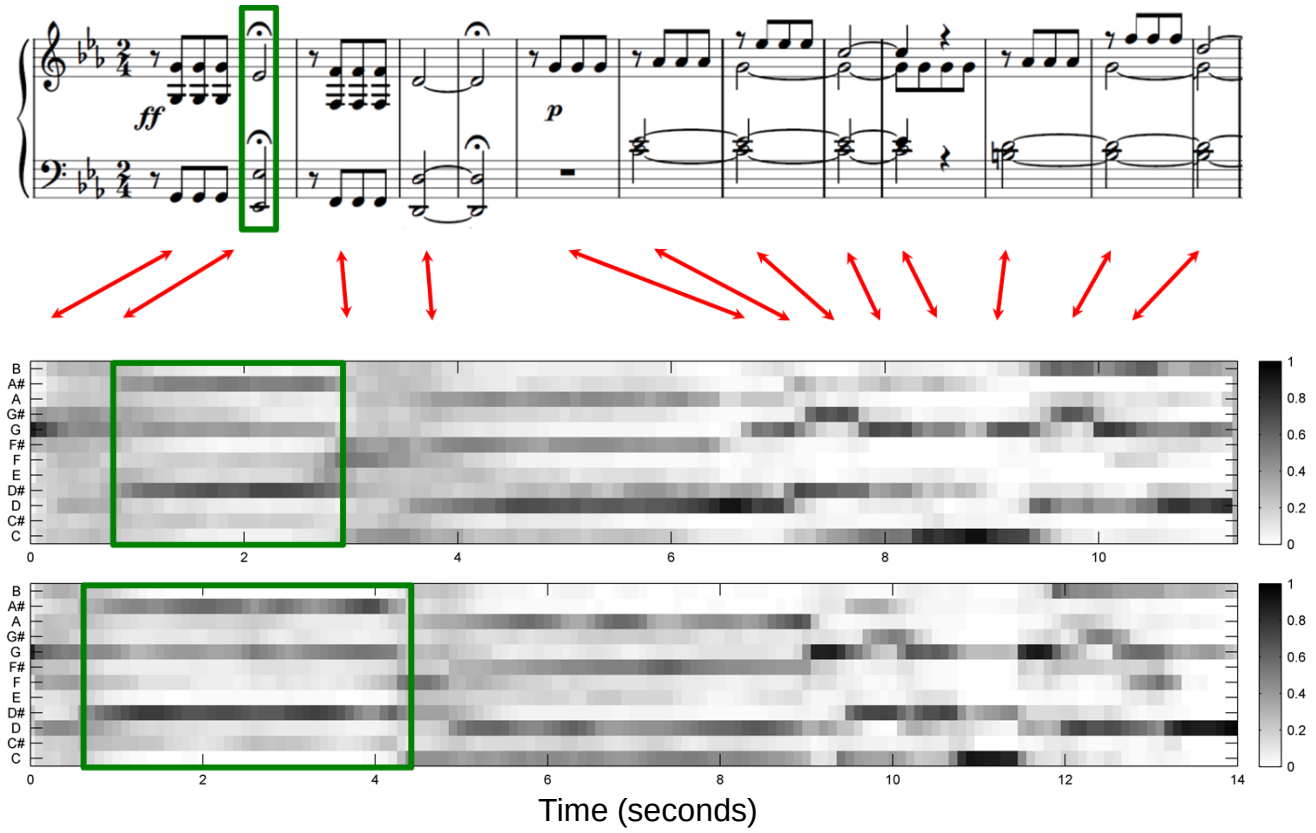
3.1 Audio Features

Fig. 3.8



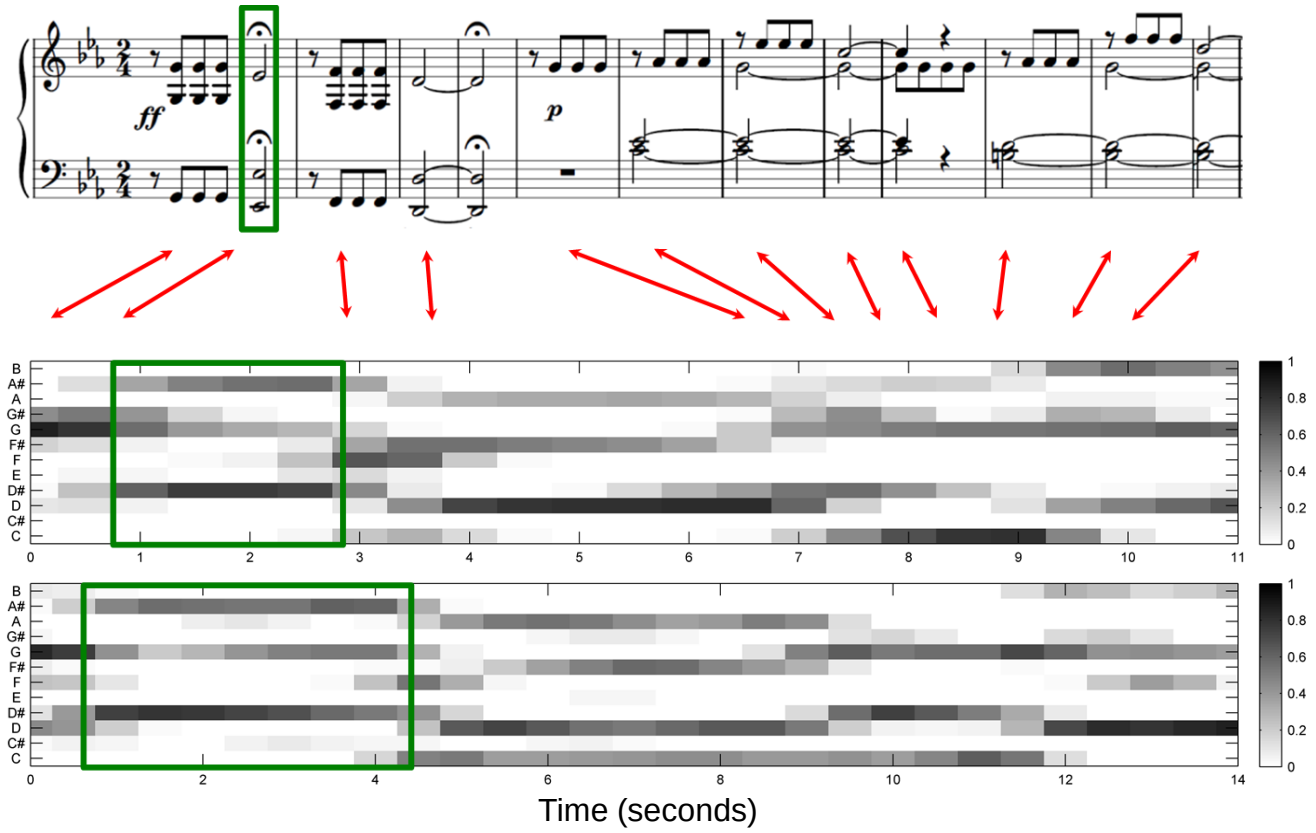
3.1 Audio Features

Fig. 3.9



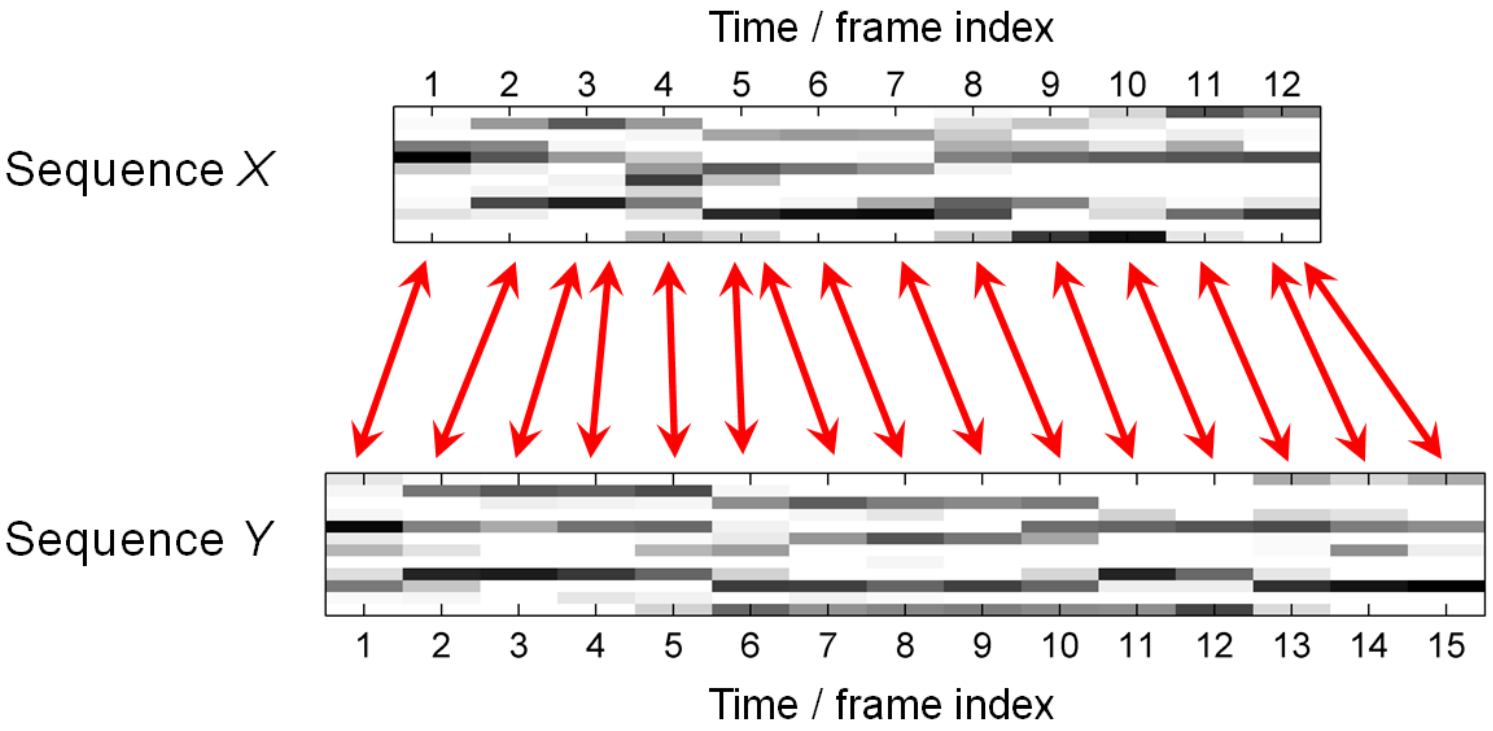
3.1 Audio Features

Fig. 3.9



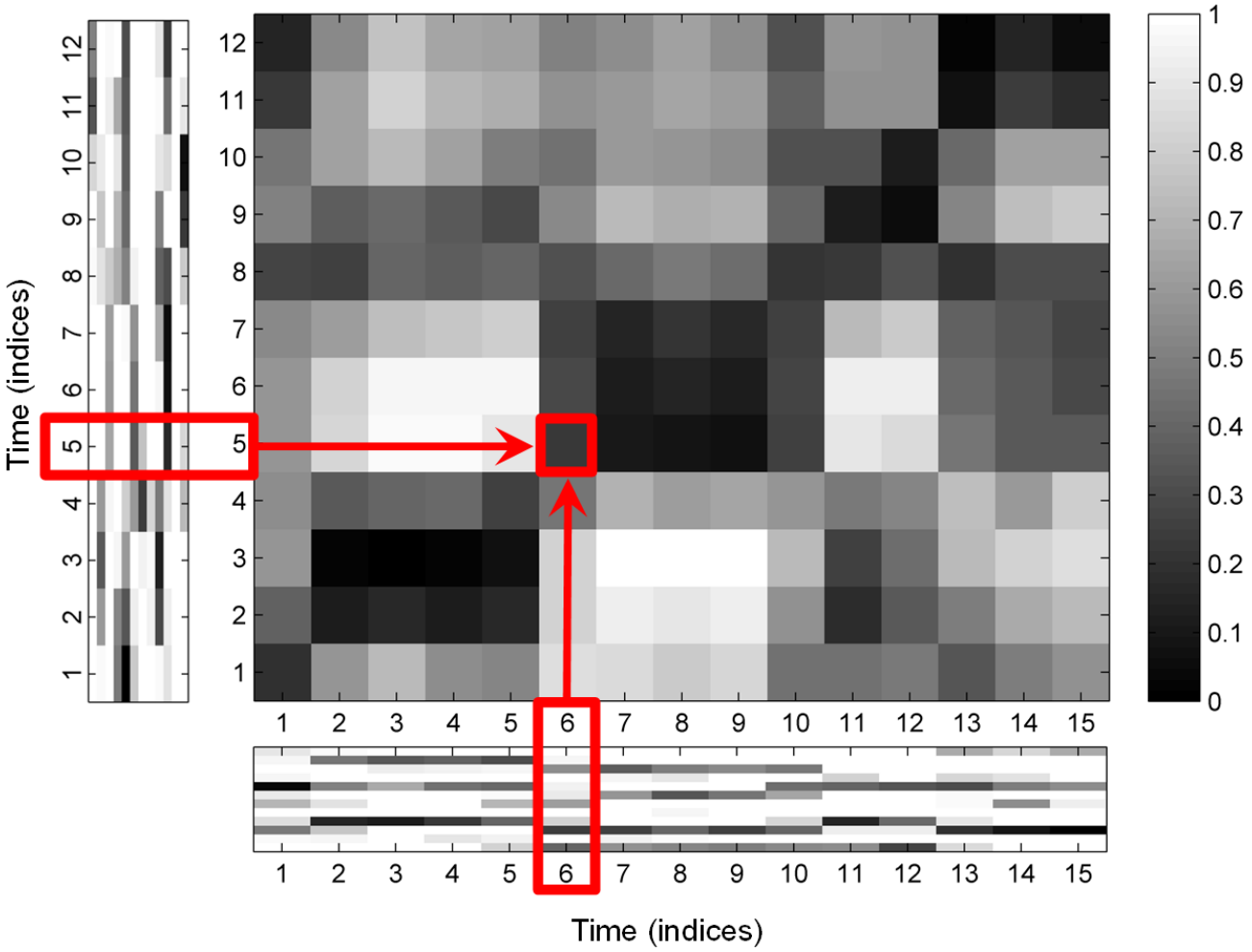
3.2 Dynamic Time Warping

Fig. 3.10



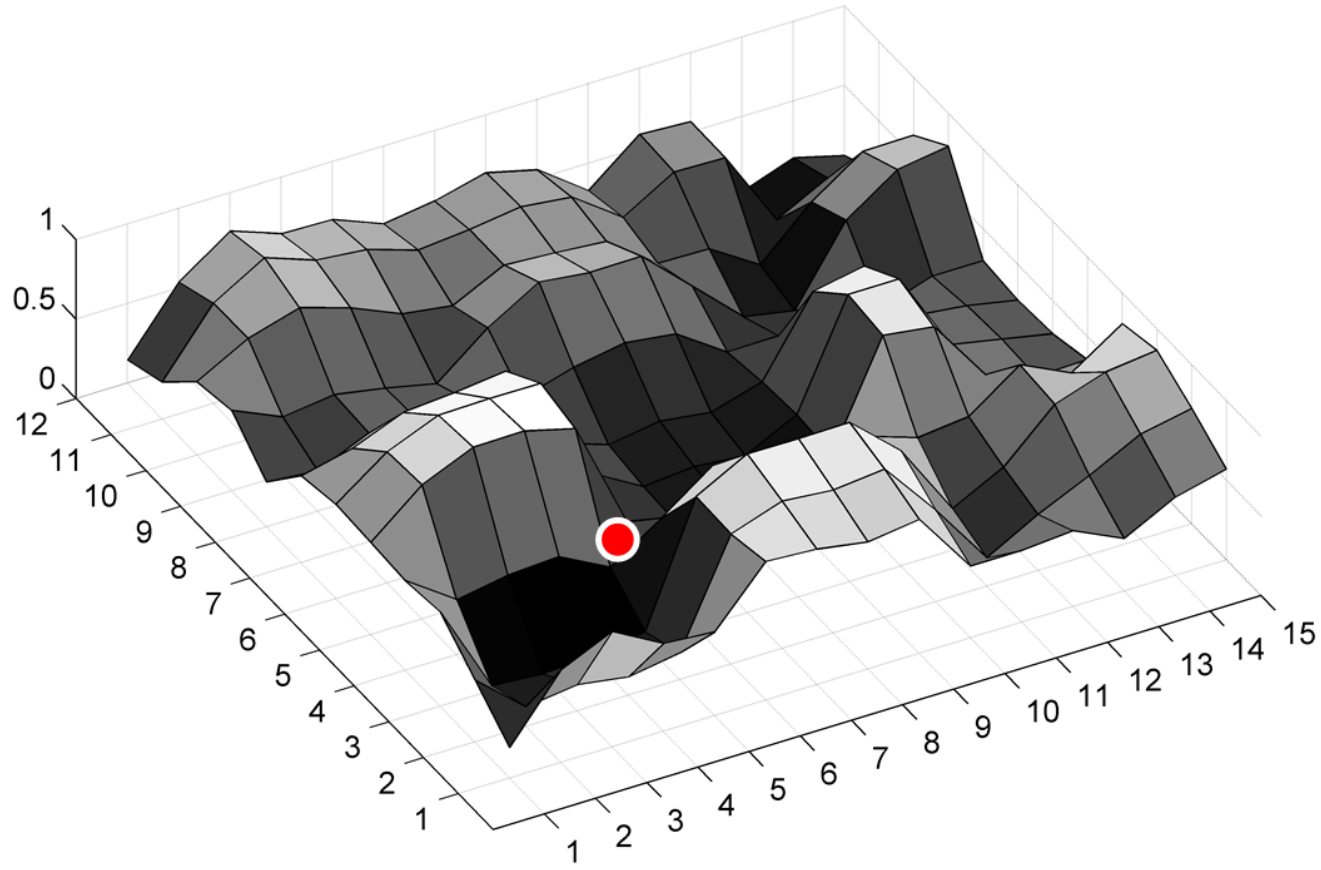
3.2 Dynamic Time Warping

Fig. 3.11



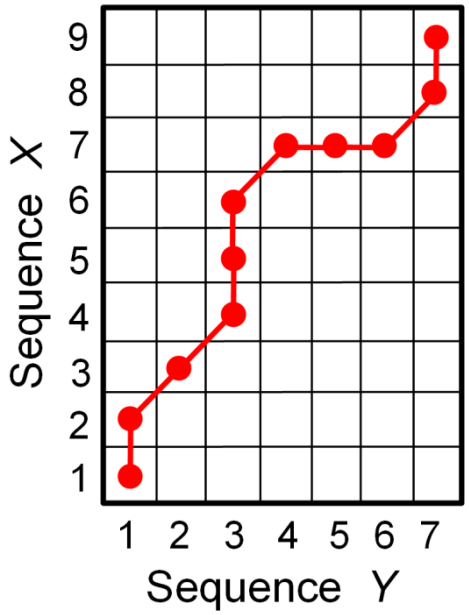
3.2 Dynamic Time Warping

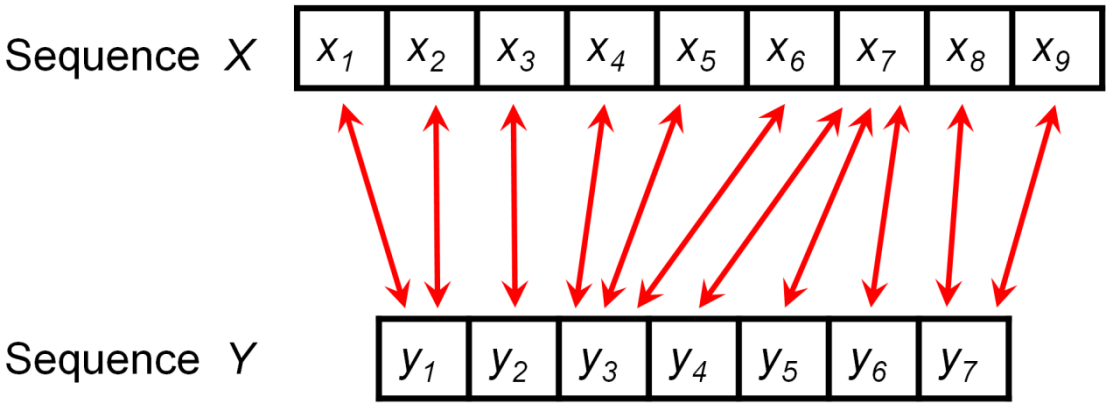
Fig. 3.11



3.2 Dynamic Time Warping

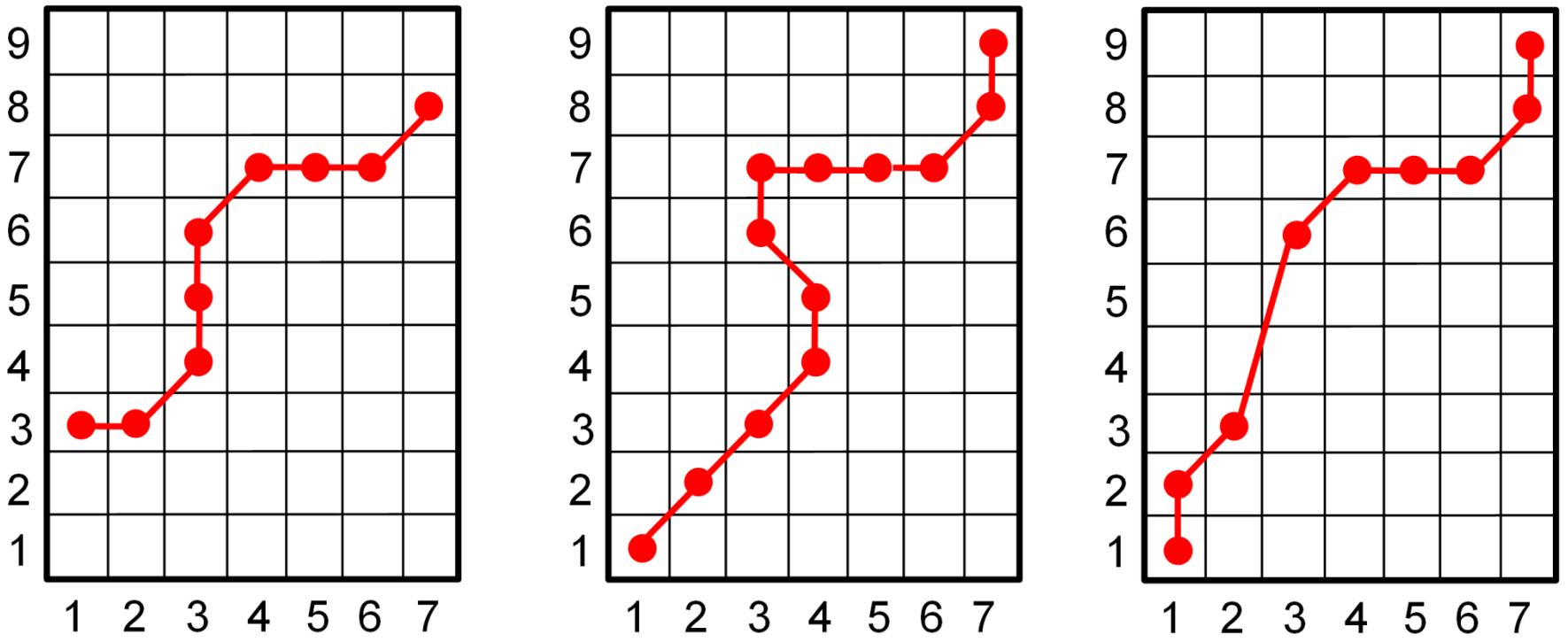
Fig. 3.12





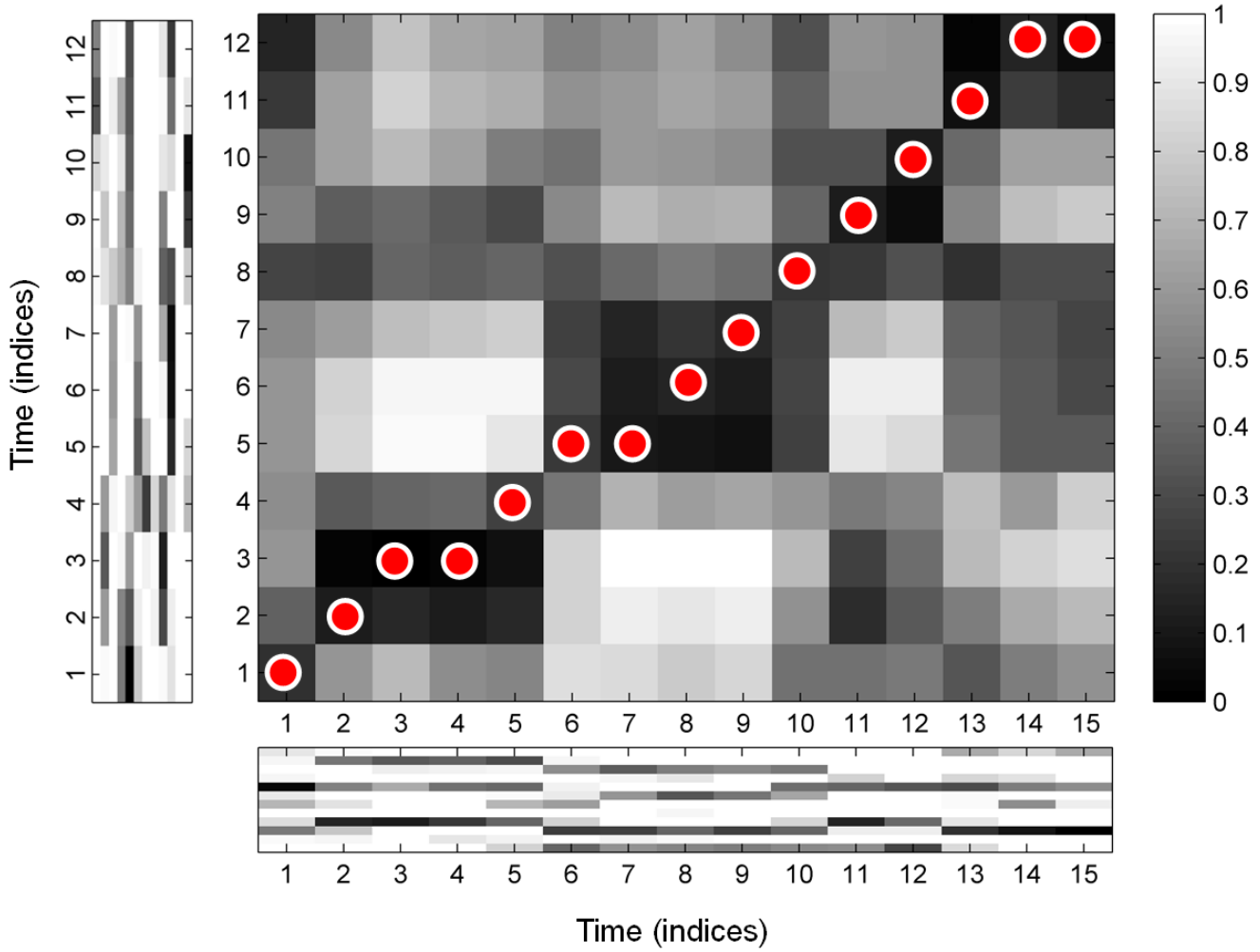
3.2 Dynamic Time Warping

Fig. 3.13



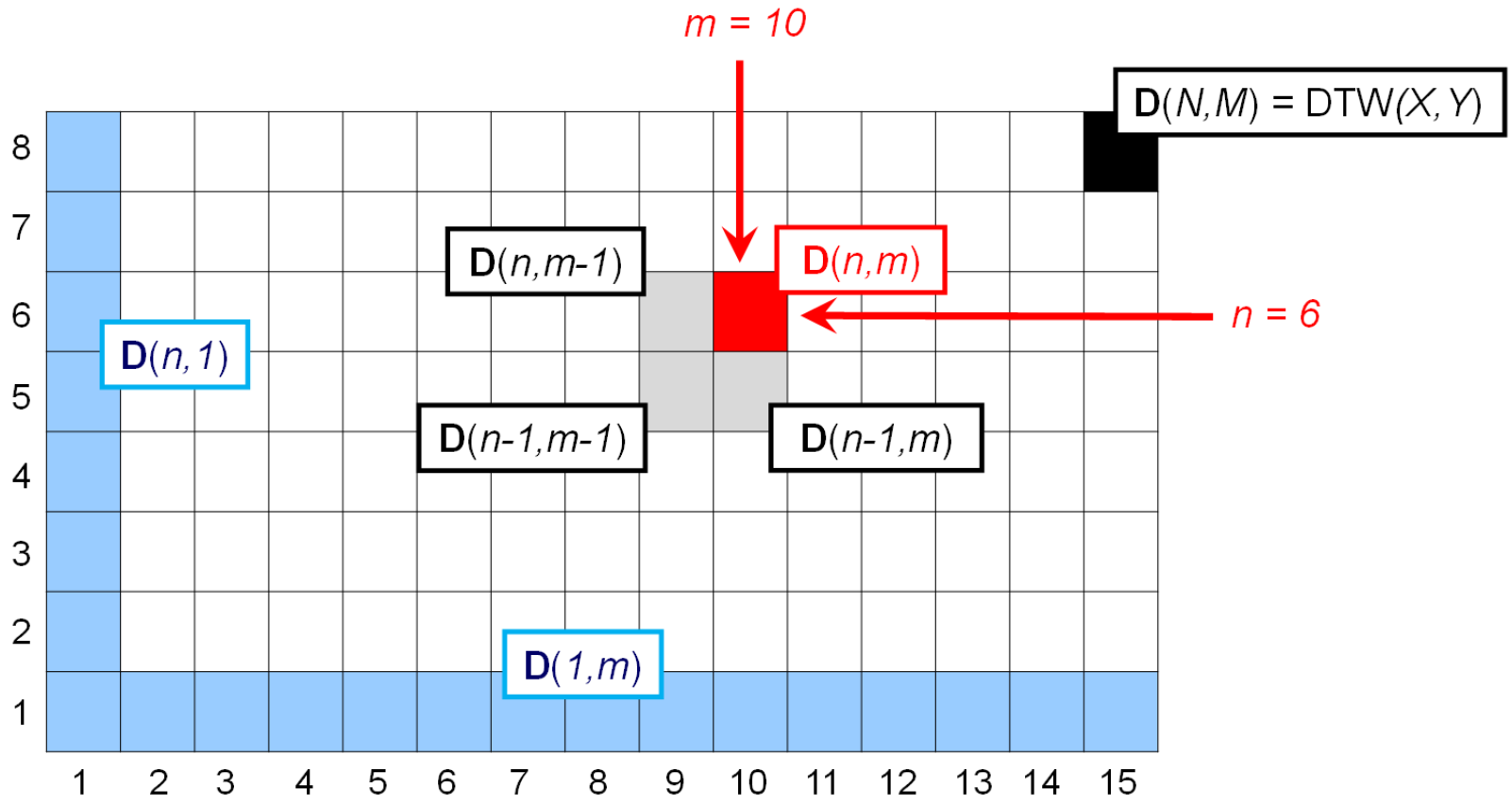
3.2 Dynamic Time Warping

Fig. 3.14



3.2 Dynamic Time Warping

Fig. 3.15



3.2 Dynamic Time Warping

Table 3.2

Algorithm: DTW

Input: Cost matrix $\mathbf{C}$ of size $N \times M$
Output: Accumulated cost matrix $\mathbf{D}$
Optimal warping path P^*

Procedure: Initialize $(N \times M)$ matrix $\mathbf{D}$ by $\mathbf{D}(n,1) = \sum_{k=1}^{n} \mathbf{C}(k,1)$ for $n \in [1:N]$ and $\mathbf{D}(1,m) = \sum_{k=1}^{m} \mathbf{C}(1,k)$ for $m \in [1:M]$. Then compute in a nested loop for $n = 2,\ldots,N$ and $m = 2,\ldots,M$:

$$\mathbf{D}(n,m) = \mathbf{C}(n,m) + \min\{\mathbf{D}(n-1,m-1), \mathbf{D}(n-1,m), \mathbf{D}(n,m-1)\}.$$

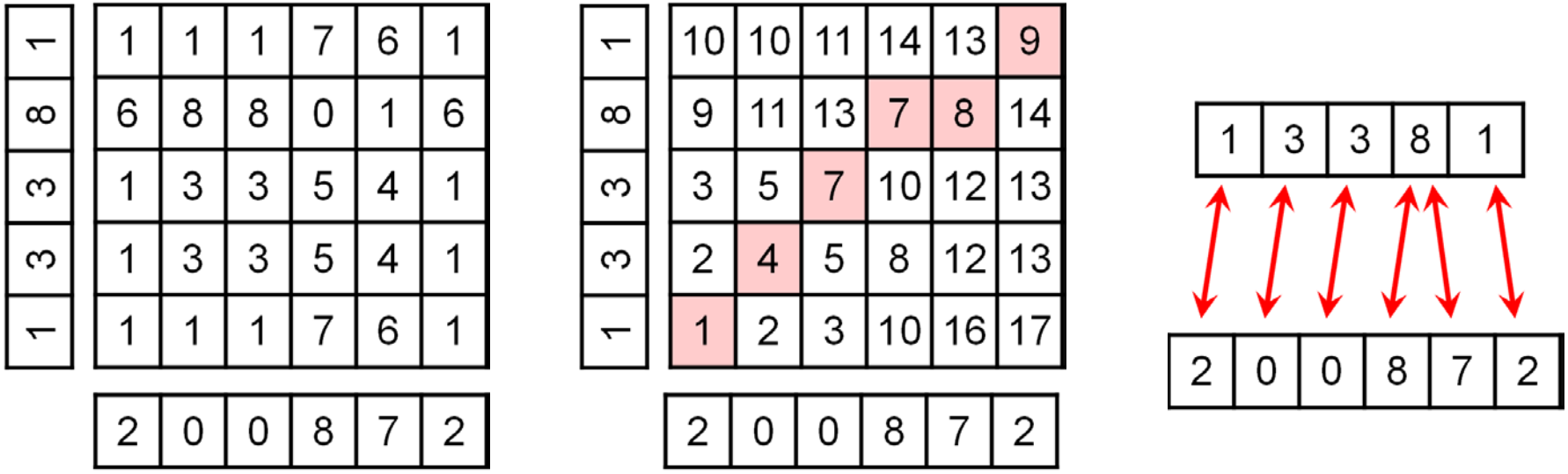
Set $\ell = 1$ and $q_\ell = (N,M)$. Then repeat the following steps until $q_\ell = (1,1)$:

Increase ℓ by one and let $(n,m) = q_{\ell-1}$.
If $n = 1$, then $q_\ell = (1,m-1)$,
else if $m = 1$, then $q_\ell = (n-1,m)$,
else $q_\ell = \operatorname{argmin}\{\mathbf{D}(n-1,m-1), \mathbf{D}(n-1,m), \mathbf{D}(n,m-1)\}$.
(If 'argmin' is not unique, take lexicographically smallest cell.)

Set $L = \ell$ and return $P^* = (q_L, q_{L-1}, \ldots, q_1)$ as well as $\mathbf{D}$.

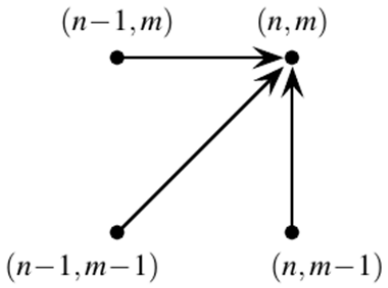
3.2 Dynamic Time Warping

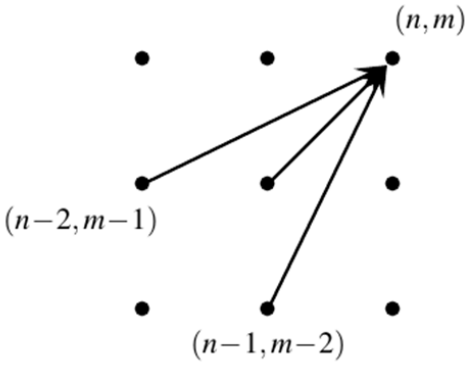
Fig. 3.16

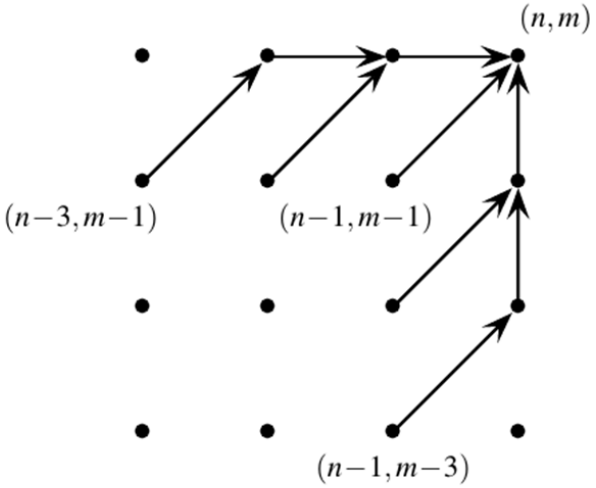


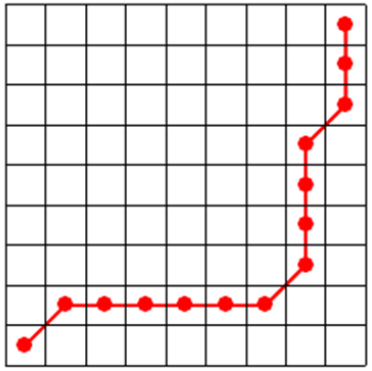
3.2 Dynamic Time Warping

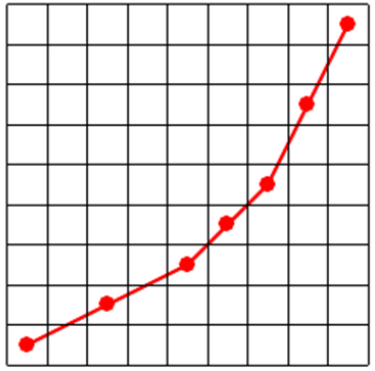
Fig. 3.17

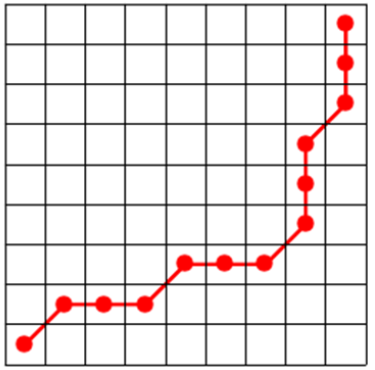






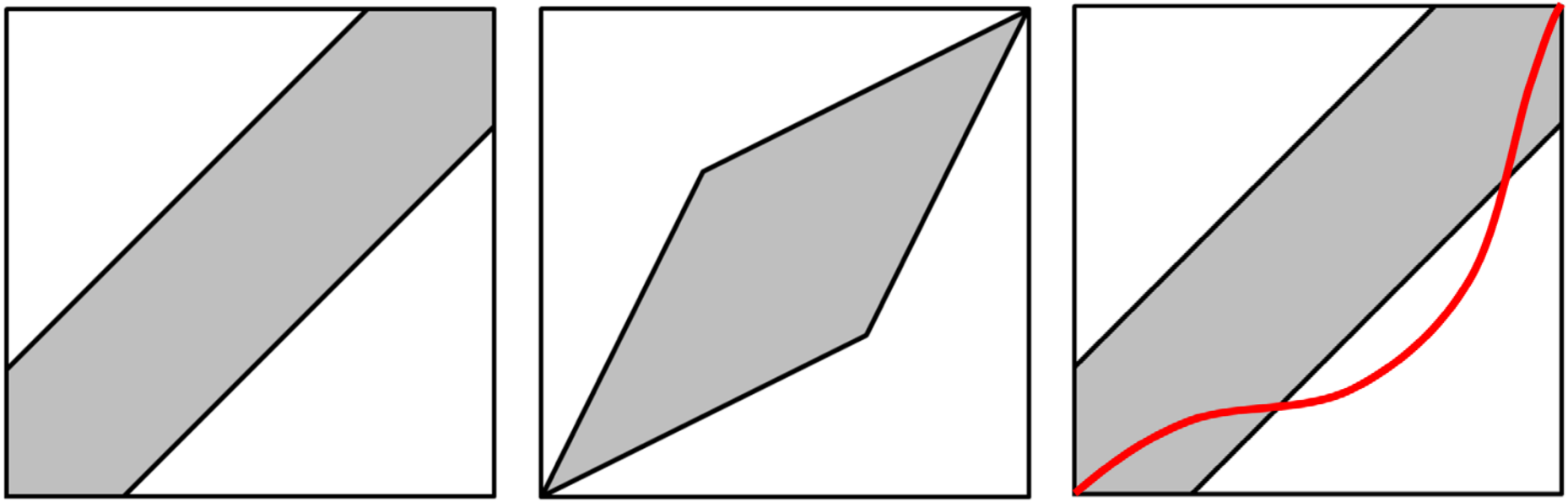






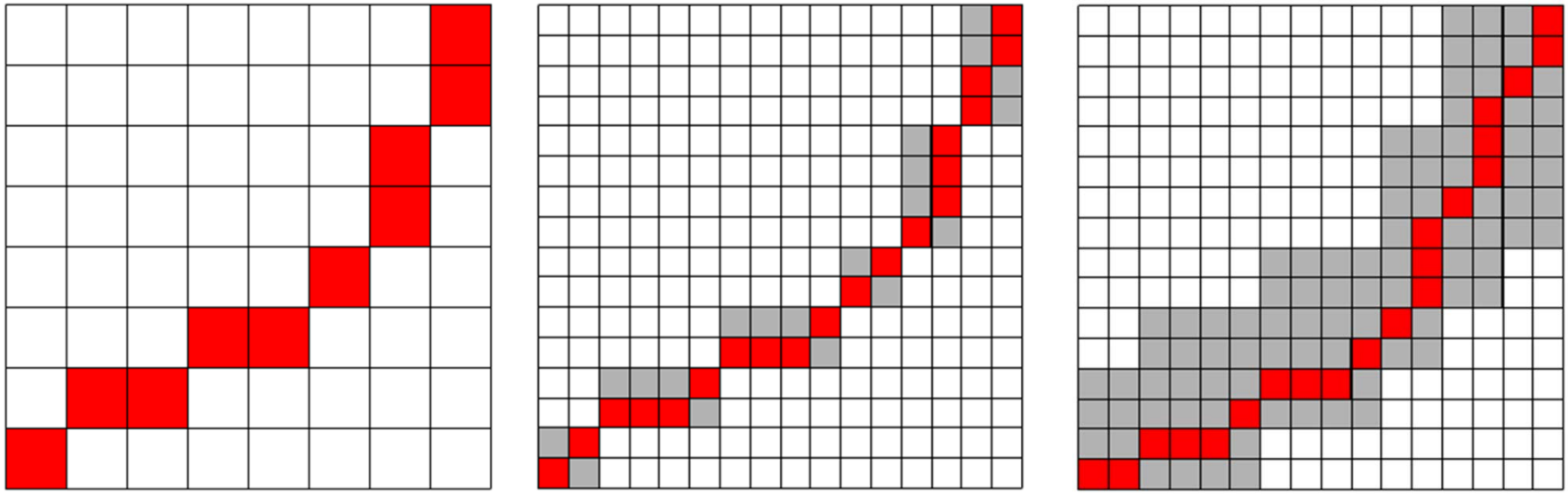
3.2 Dynamic Time Warping

Fig. 3.18



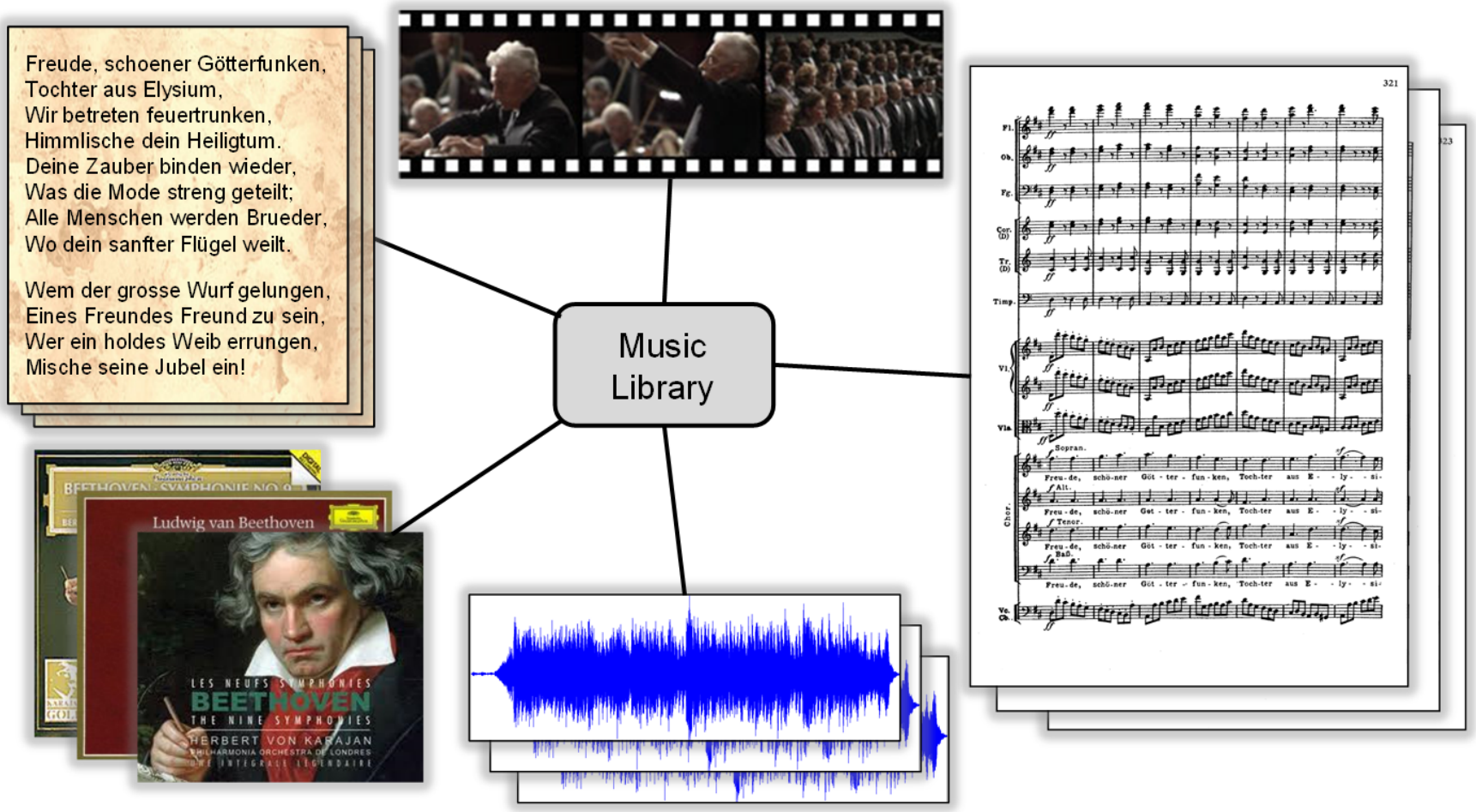
3.2 Dynamic Time Warping

Fig. 3.19



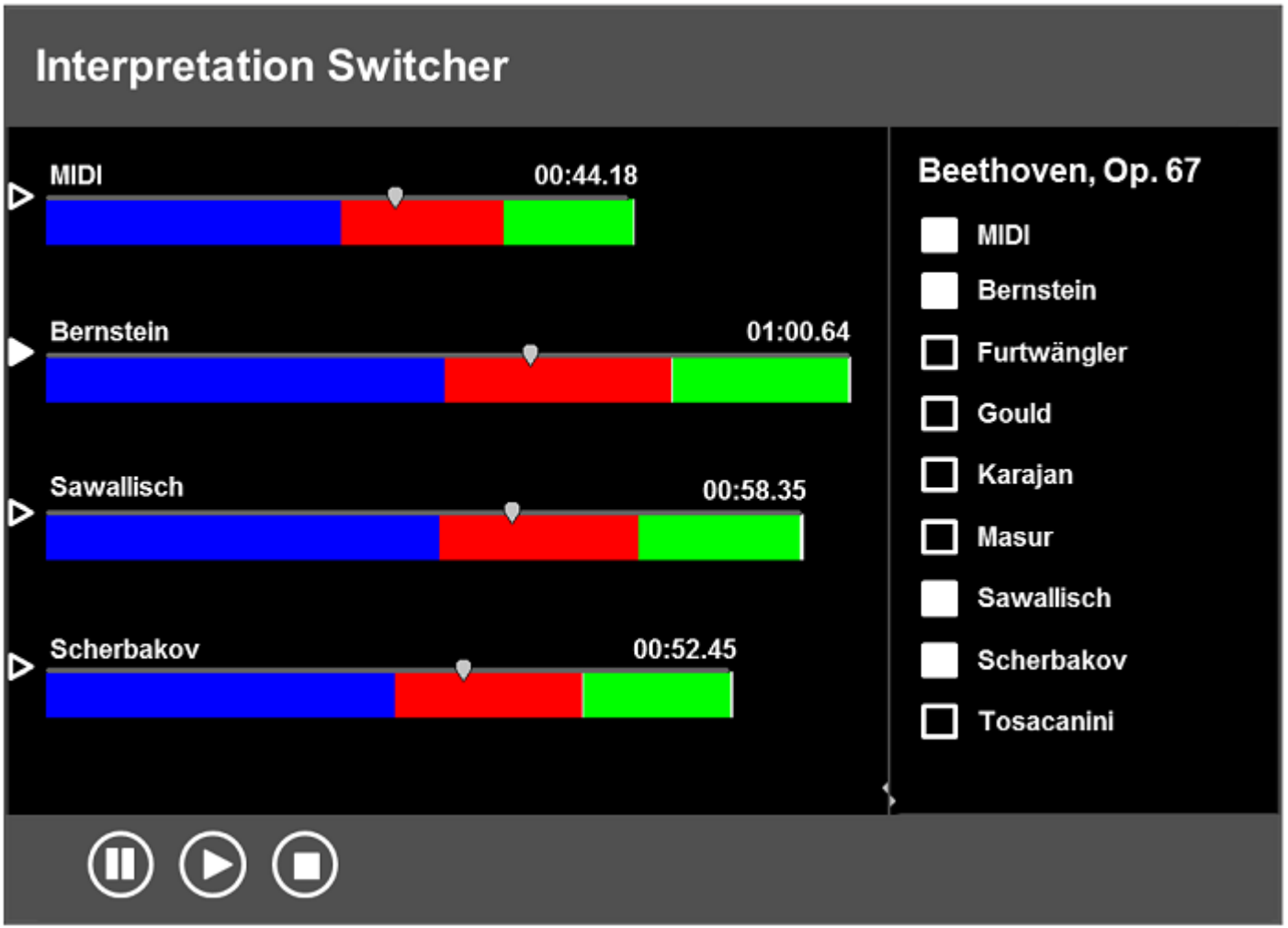
3.3 Applications

Fig. 3.20



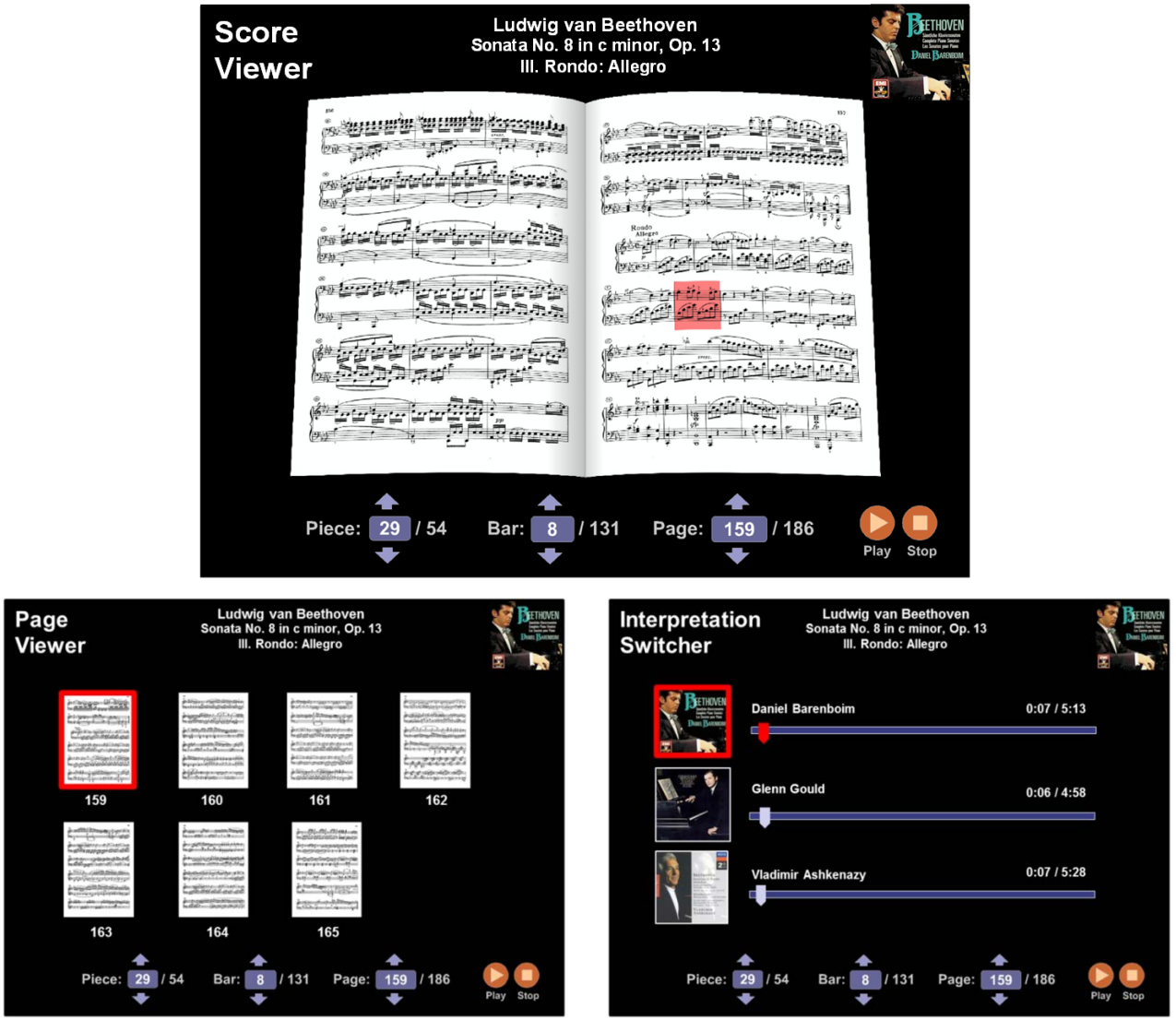
3.3 Applications

Fig. 3.21



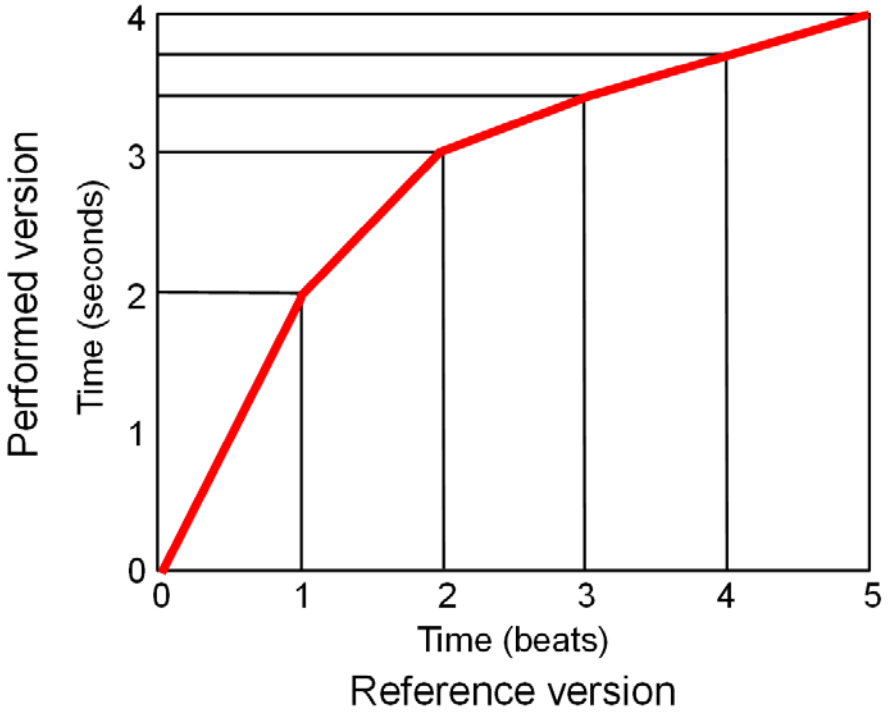
3.3 Applications

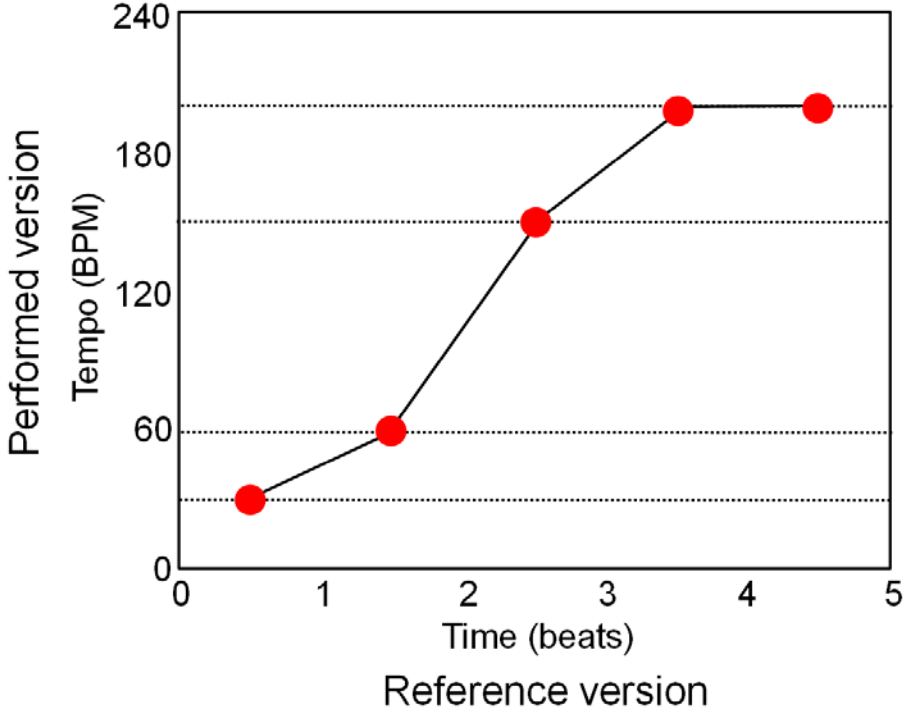
Fig. 3.22



3.3 Applications

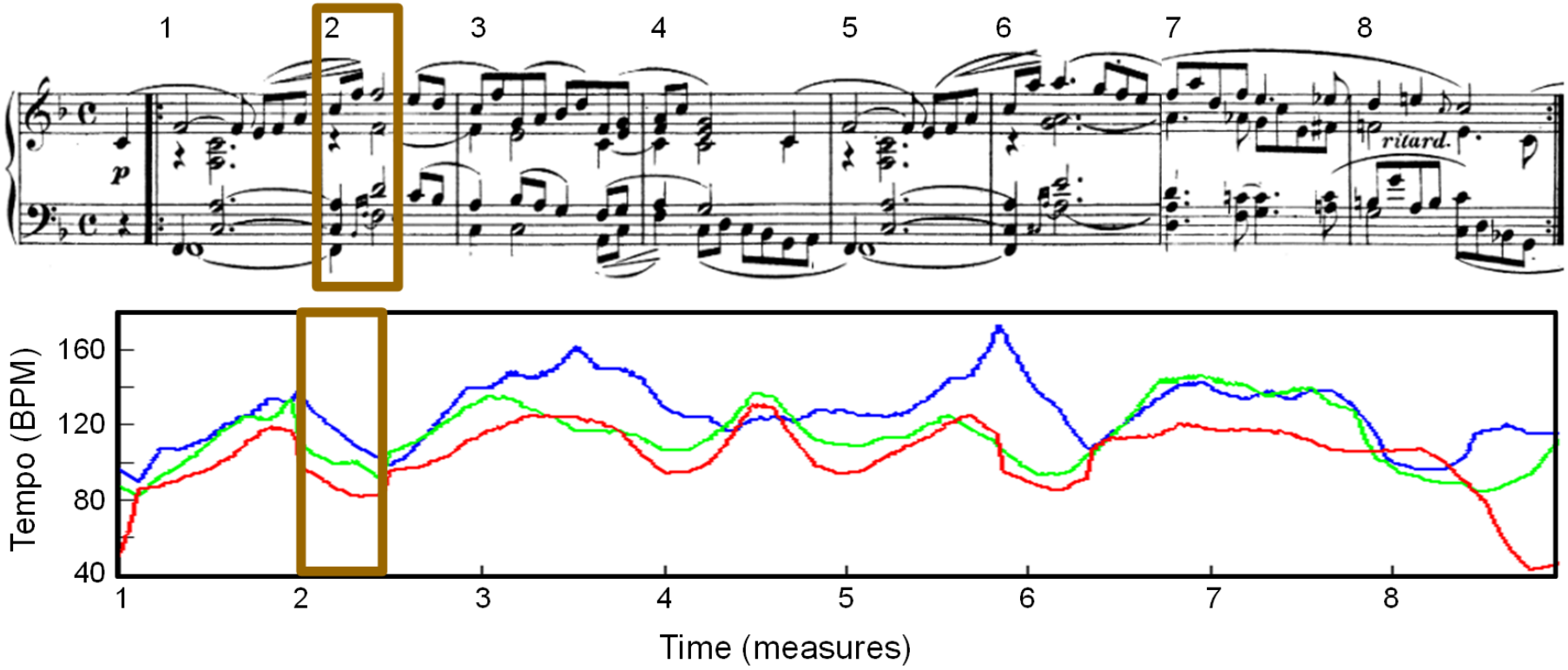
Fig. 3.23





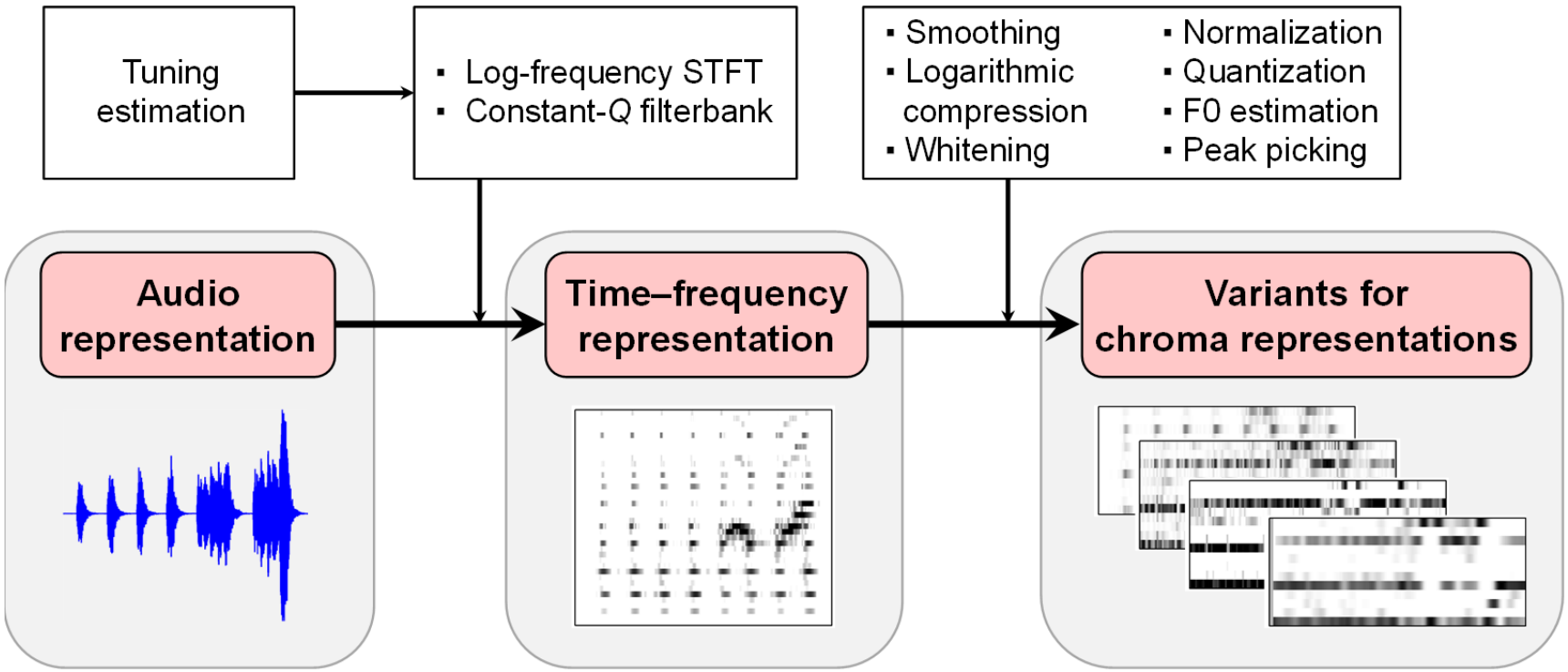
3.3 Applications

Fig. 3.24



3.4 Further Notes

Fig. 3.25



3.4 Further Notes

Fig. 3.26