Whale detection - Brainstorming session

Jiří Dutkevič
Lenka Kovářová
Milan Le

Signal processing, Sampling theorem

Spojitý **signál může být nahrazen** diskrétní posloupností vzorků, aniž by došlo ke ztrátě informace, a to tak, že **vzorkovací frekvence Fs je nejméně 2x větší** než nejvyšší frekvenci v signálu.

Původní spojitý signál může být zrekonstruován ze získaných vzorků.

Signal processing, Aliasing

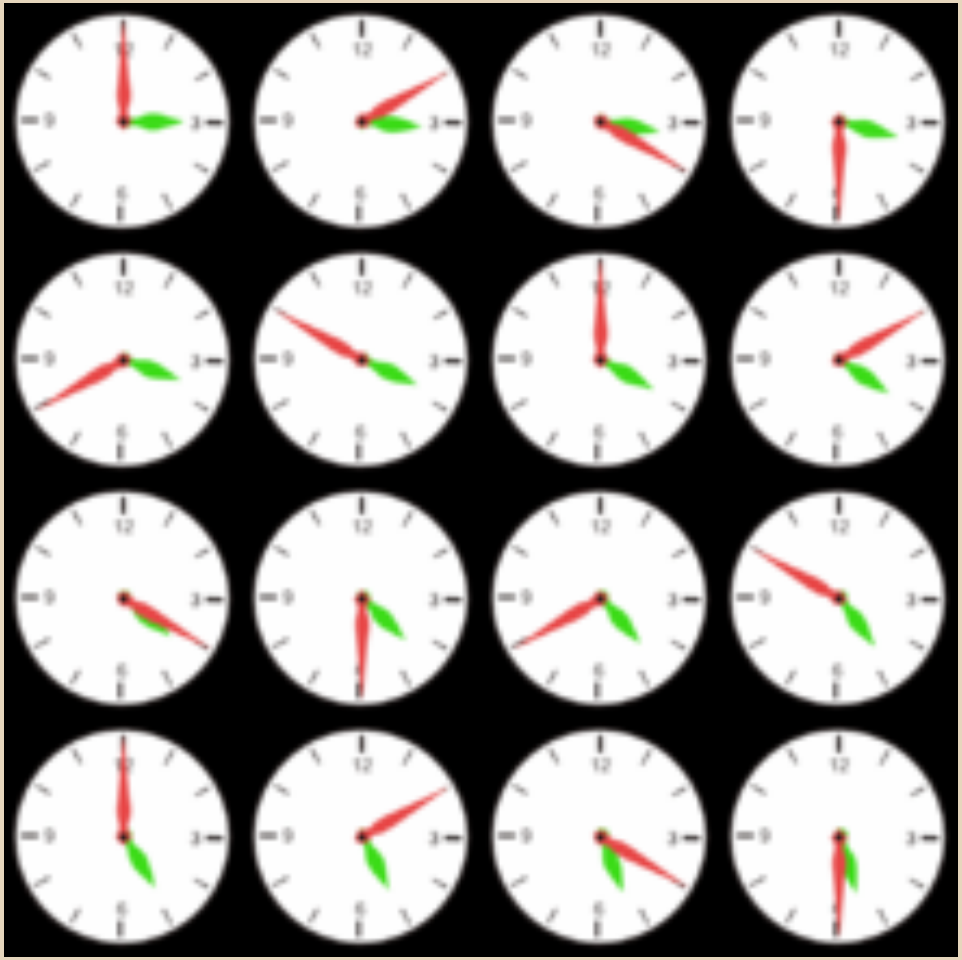
Jinými slovy: Fs/2 (*Nyquist frekvence*) je nejvyšší frekvence, které lze docílit rozkladem.

Ve filmu se vrtule točí:

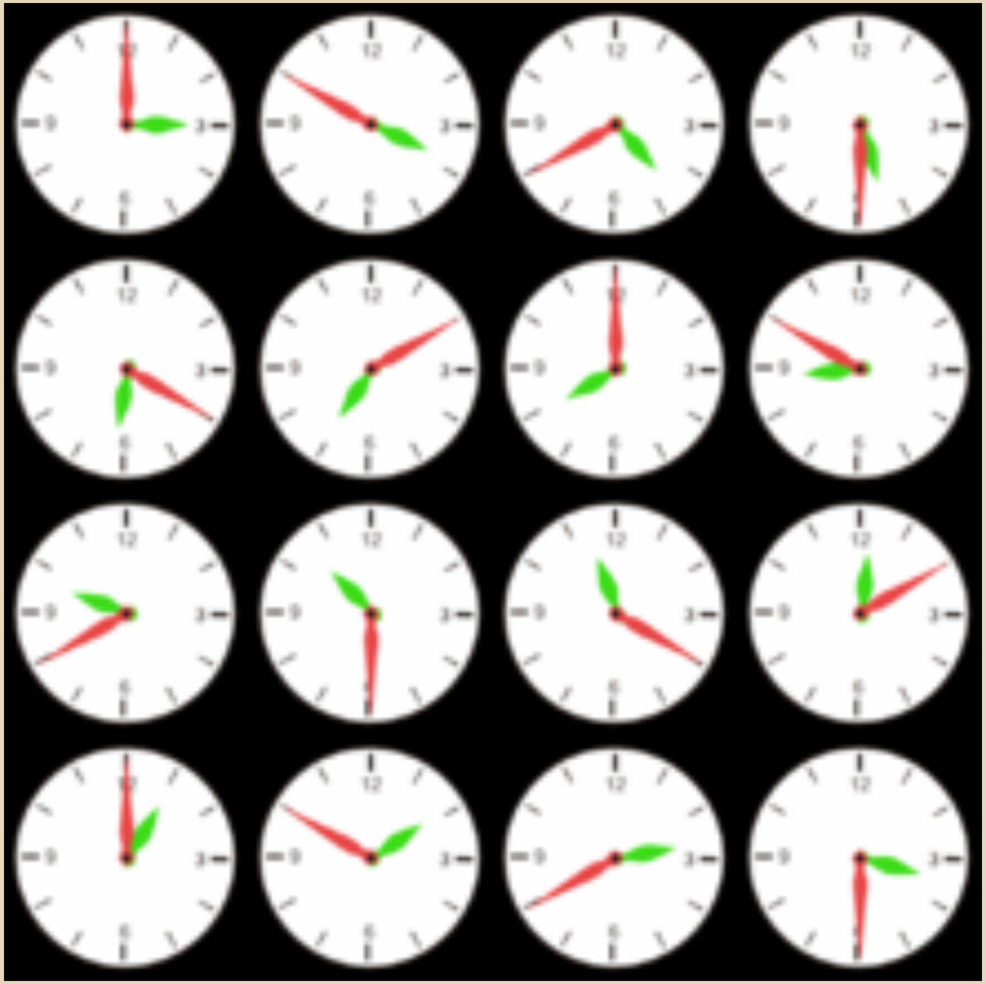
- V opačném směru než ve skutečnosti
- Nepřirozeně pomalu

Správně by byl pohyb zachycen jenom kdyby kamera snímala alespoň dvakrát tak rychleji

Signal processing, Aliasing



Signal processing, Aliasing



Antialiasing

Spojitý signál -> antialising filter -> diskrétní popis signálu

Filtruje frekvence vyšší než odpovídají Shannonovu teorému

Framing - rozdělení na "okénka".

FT jsou stacionární sinusoidy, skutečný signál není

U lidí je framing 25ms, u velryb jsem zvolil 250ms

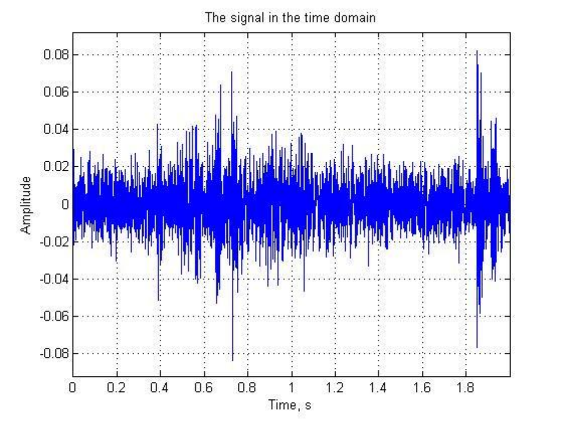
Vážení signálu s okénkovou funkcí

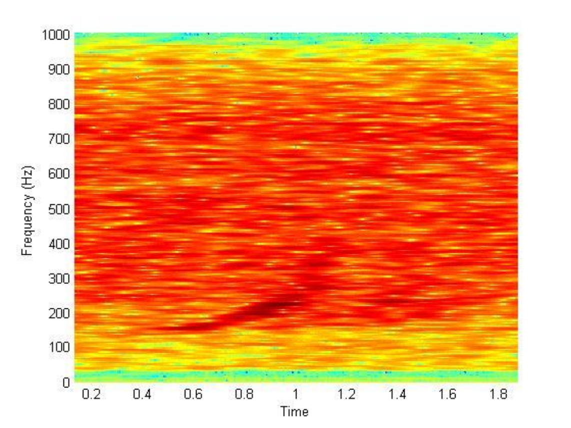
$$\mathbf{STFT}\{x[n]\}(m,\omega) \equiv X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n}$$

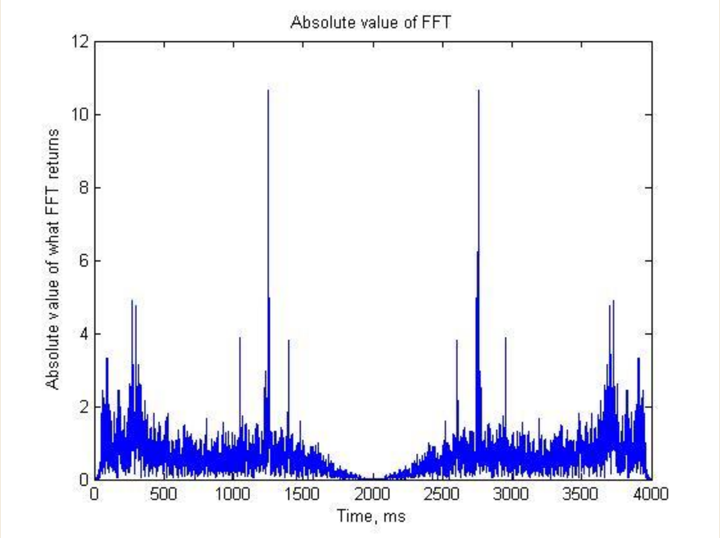
Požadujeme

$$\int_{-\infty}^{\infty} w(\tau)\,d\tau = 1.$$

DFT, Spectral analysis





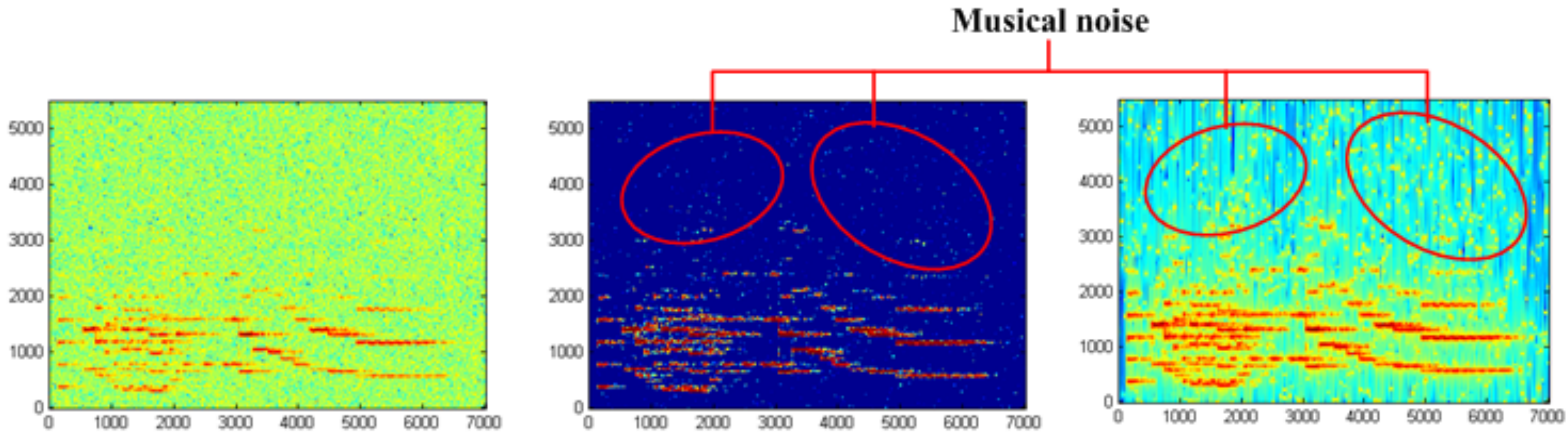


Denoising "odšumění"

Audio Denoising by Time-Frequency Block Thresholding. Volba všech parametrů adaptivně na vlastnostech signálu minimalizací Steinovým odhadem risku.

"Hudební šum" způsobený denoising algoritmem jako *power subtraction*, algoritmus block thresholding tento šum reguluje

Power subtraction

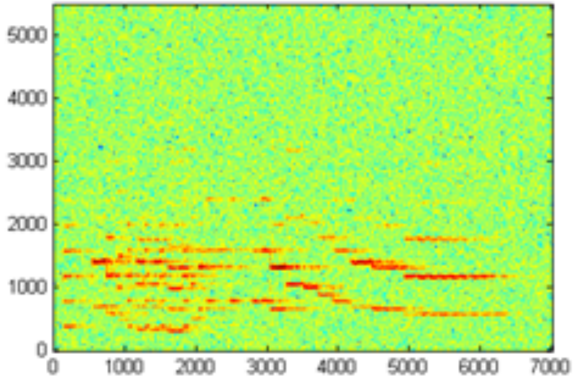


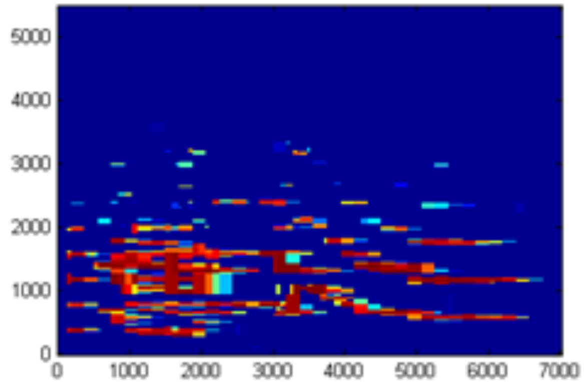
Noisy spectrogram

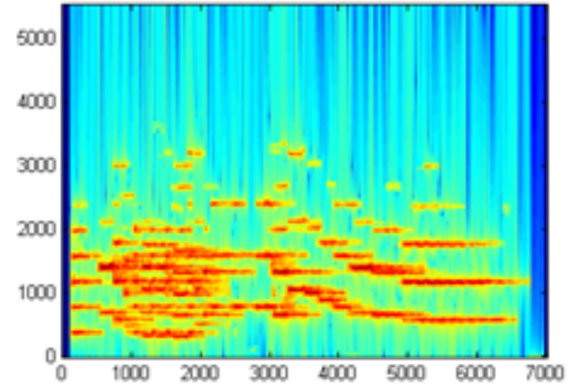
Attenuation factor of power subtraction. The isolated coefficients restore isolated short-time Fourier structures perceived as musical noise.

Spectrogram of denoised signal. The isolated coefficients correspond to musical noise.

Block thresholding



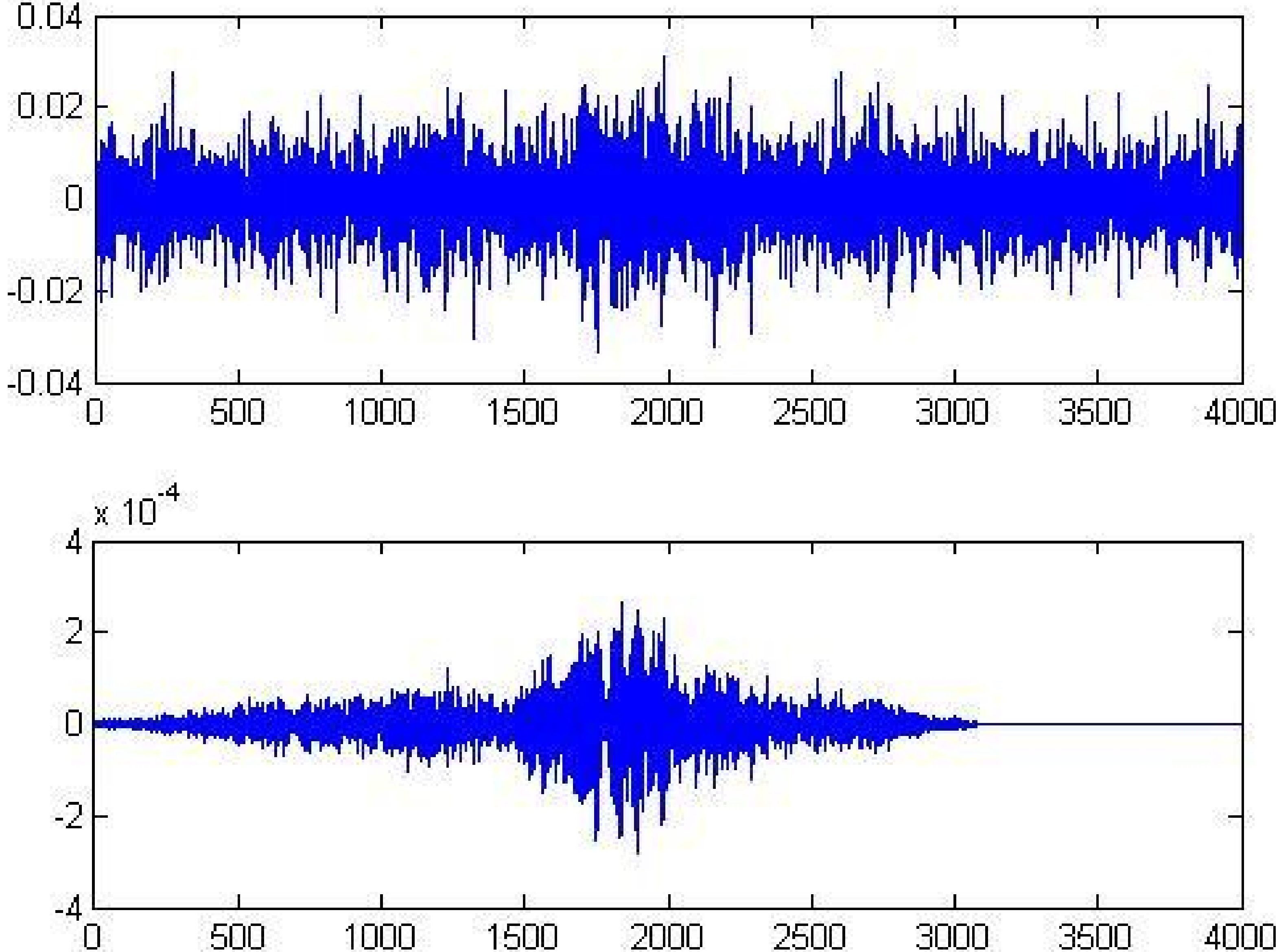




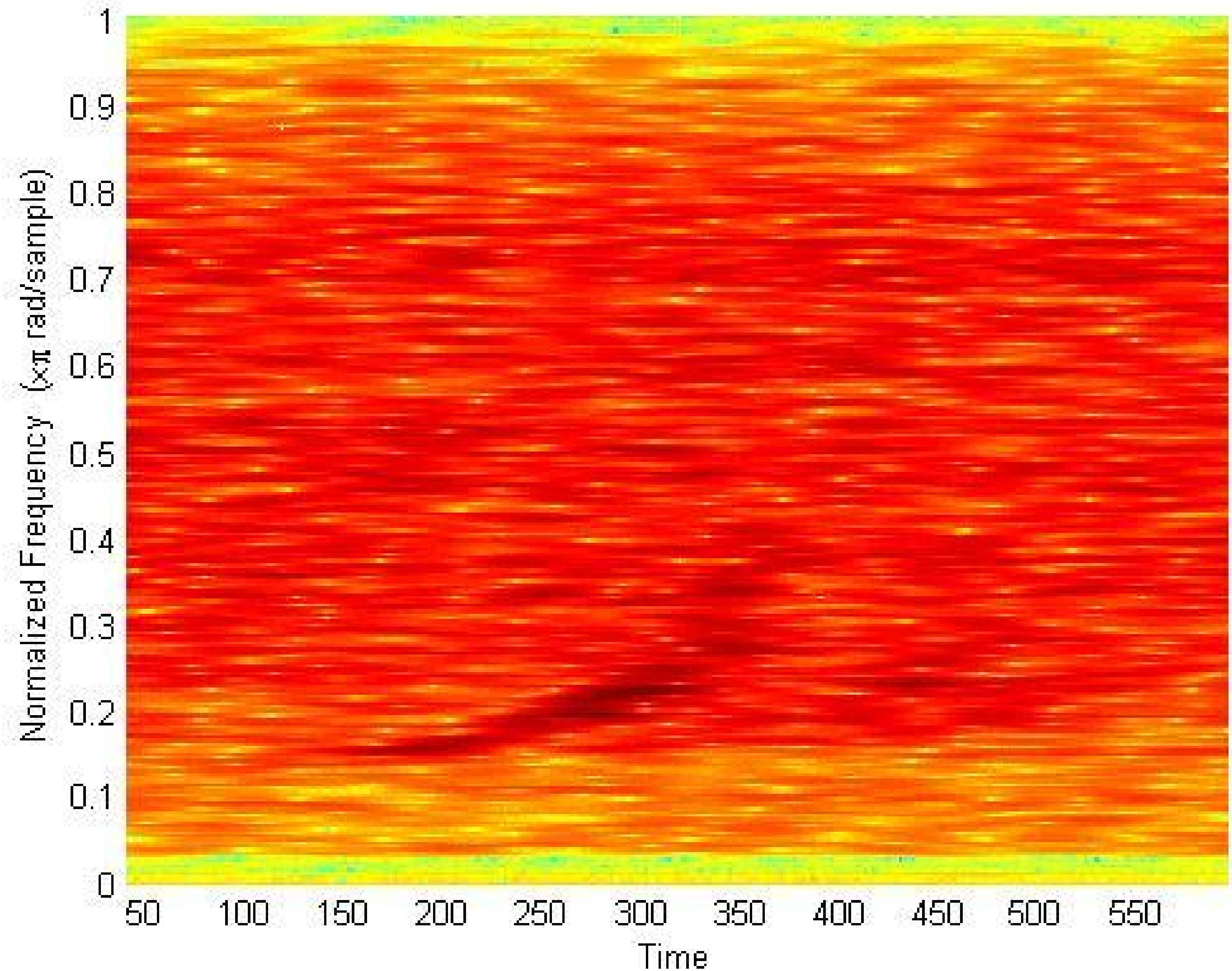
Noisy spectrogram

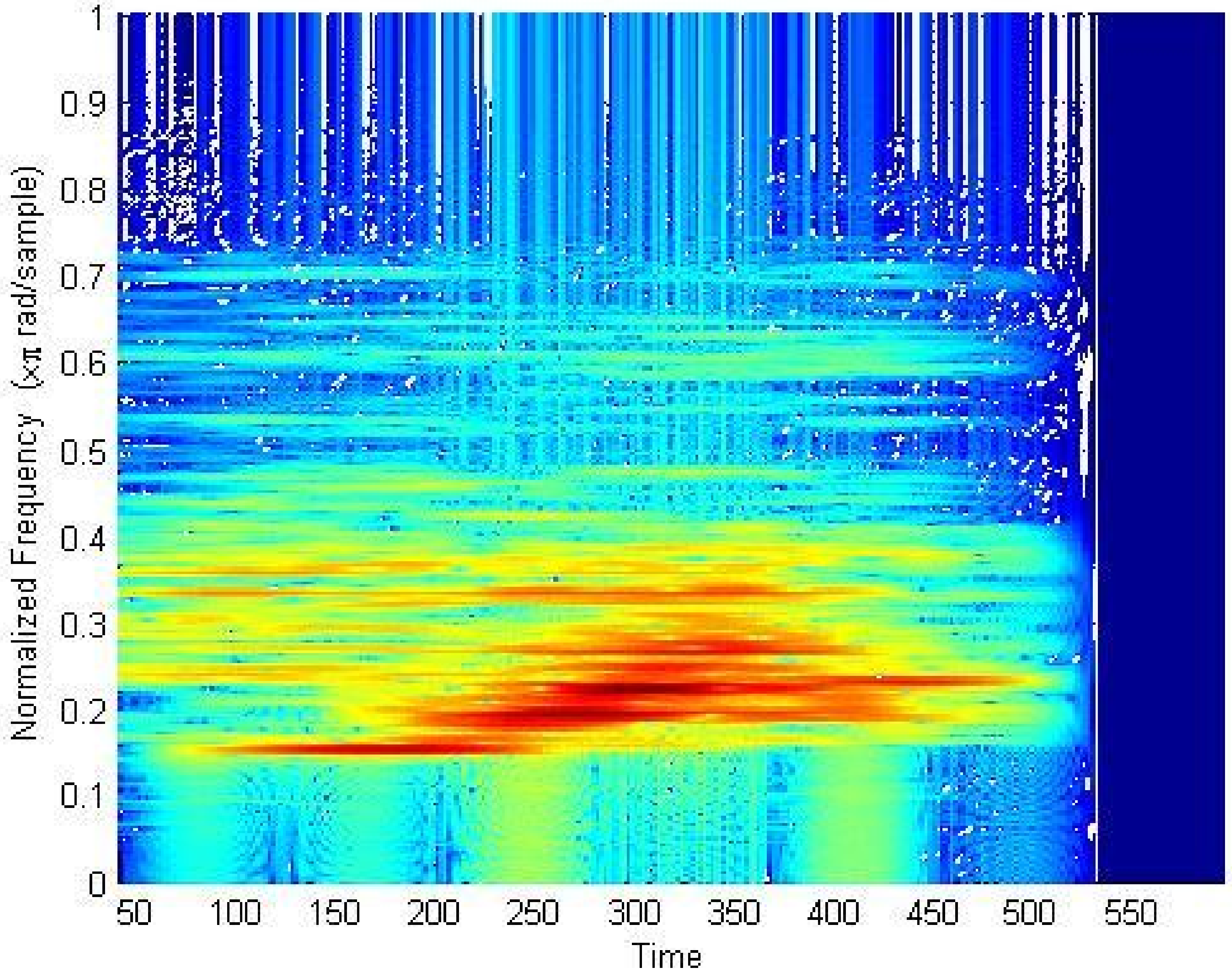
Attenuation factors of block thresholding. Spectrogram coefficients are grouped in blocks whose size is adaptive to signal property. There is no isolated coefficients responsible for musical noise.

Spectrogram of denoised signal.



0.04
0.02
0
-0.02
-0.04
0
500
1000
1500
2000
2500
3000
3500
4000
x 10^-4
4
2
0
-2
-4
0
500
1000
1500
2000
2500
3000
3500
4000


Normalized Frequency (×π rad/sample)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
50
100
150
200
250
300
350
400
450
500
550
Time



Normalized Frequency (×π rad/sample)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
50
100
150
200
250
300
350
400
450
500
550
Time

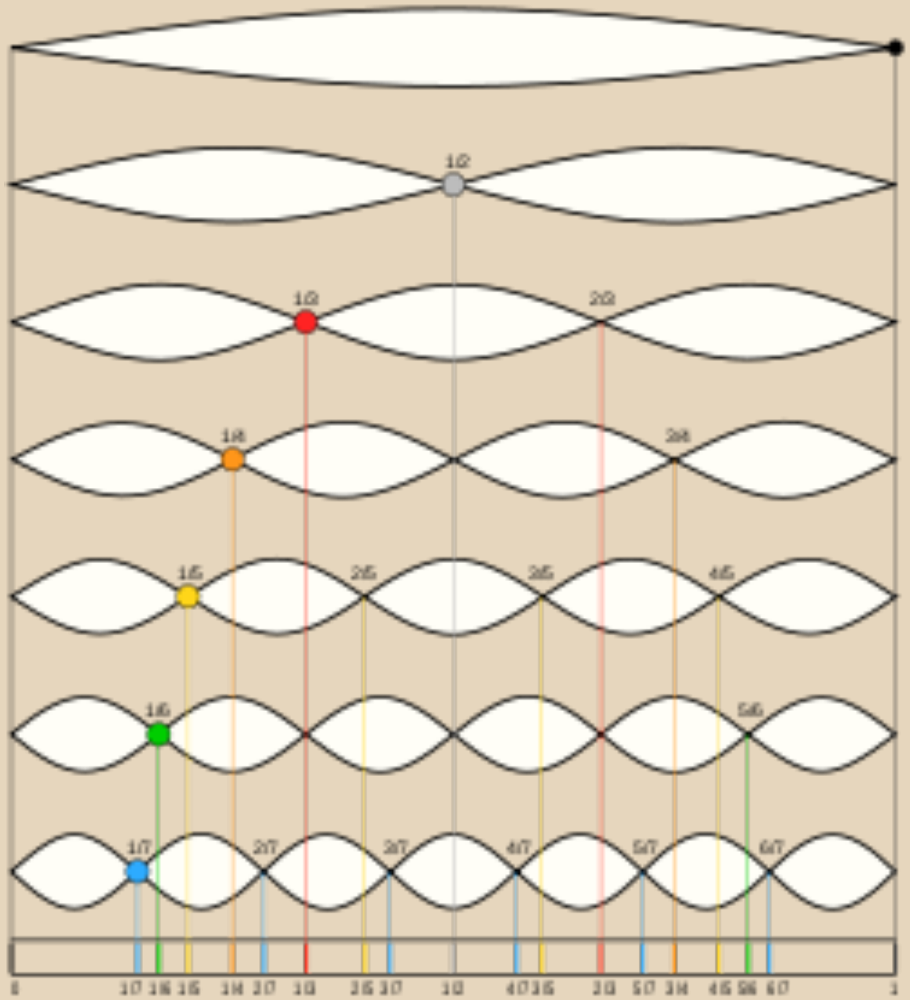
Audio Features

Spectralní features

- popisují spectrum, nejoblíbenější
- cepstral koefy, v krátkých frame

```
c=real(ifft(log(abs(fft(x)))))
```

- Mel-freq cepstral koefy
- spekrální centroid, bandwidth

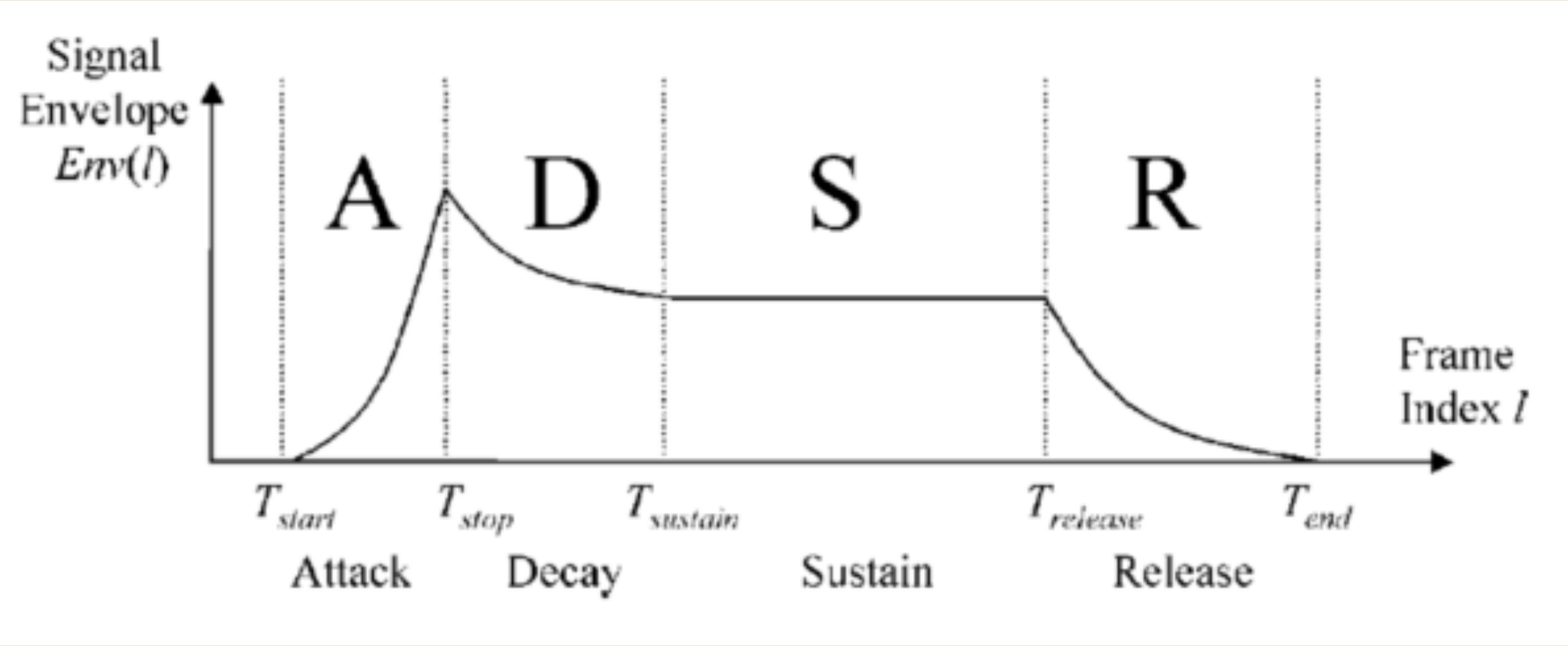


Harmonické vlny

Features

Temporální features

- pro rozpoznání pohlaví, mluvčího, nástrojů



Features

Zero-crossing rate

$$zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} \mathbb{I}\{s_t s_{t-1} < 0\}$$

ZCR koreluje se spectrálním centroidem

193 vytažených features

Energic

- AudioPower = čtverec waveform

Harmonic

- stupeň harmoničnosti v nahrávce

Perceptual

Spectral

Temporal

Jiné

Referencies

BlockThresholding

http://www.cmap.polytechnique.fr/~yu/research/ABT/samples.html

G. Yu, S. Mallat, E. Bacry, Audio Denoising by Time-Frequency Block Thresholding, IEEE Trans. on Signal Processing, vol 56, no. 5, pp. 1830-1839, May 2008.

Audio Descriptors

http://www-sipl.technion.ac.il/Info/Teaching_Projects_MPEG-7-Audio-Descriptors_e.shtml

Sound Description Toolbox

Assembled (and partially written) by Emmanouil Benetos with a little help by M. Kotti & C. Kotropoulos Inspired by: G. Peeters, "A large set of audio features for sound description in the CUIDADO project", IRCAM Technical Report, 2004.

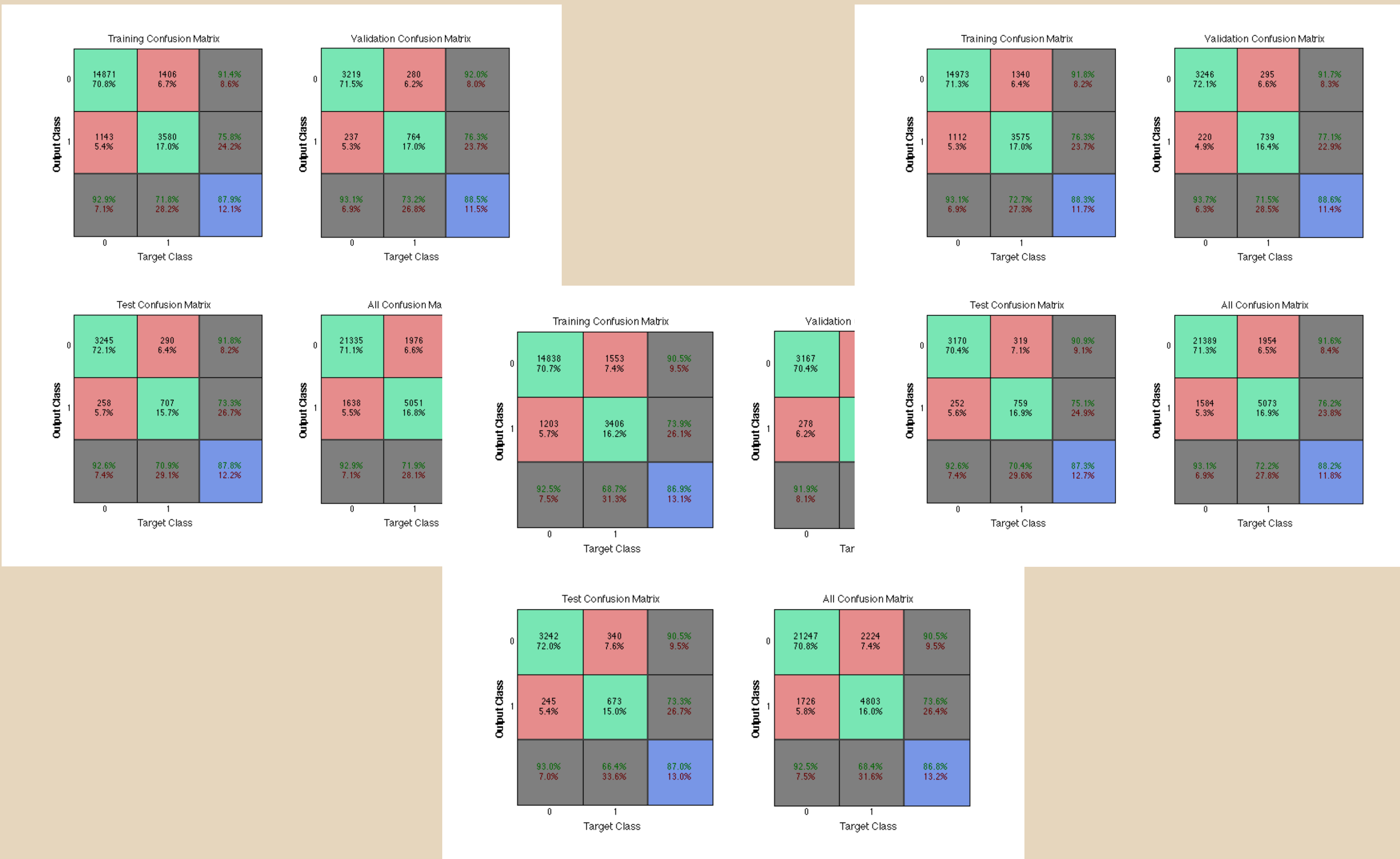
Referencies

Klapury, ISMIR Graduate School, October 4th-9th, 2004

Neuronové sítě s extrahovanými příznaky

- použita klasifikační neuronová síť s jednou skrytou vrstvou a jedním neuronem ve výstupní vrstvě
- náš nejlepší výsledek: úspěšnost 87%
- benchmark: 72.18% (Cornell University)
- vítěz: 98.384% (tým SluiceBox)

Neuronové sítě s extrahovanými příznaky



Random forests

Klasifikační metoda

Během trénovací fáze algoritmu se vytvoří několik rozpoznávacích stromů

Výsledek modelu je získán jako modus výsledků jednotlivých stromů

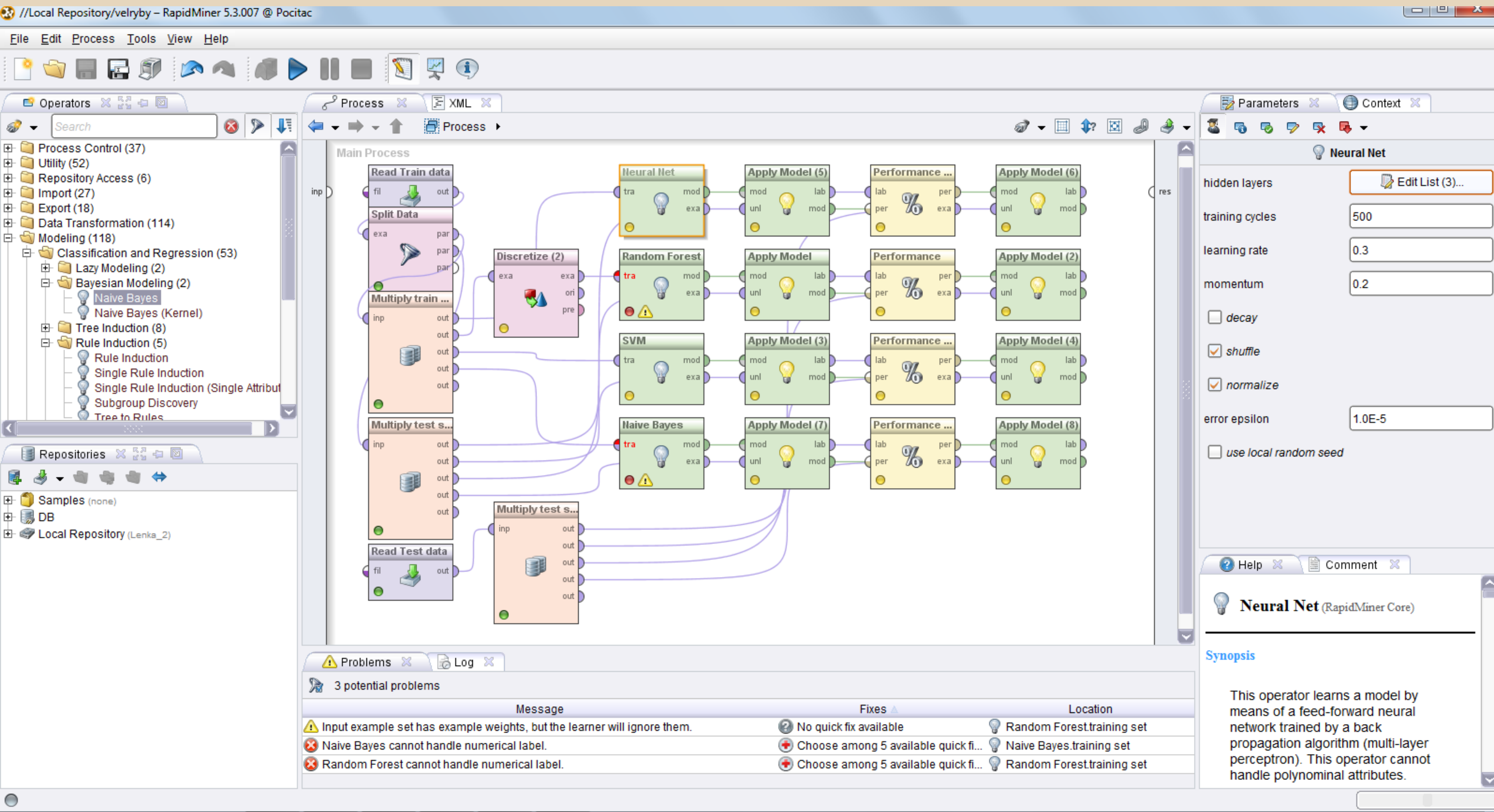
Support vector machines

Metoda strojového učení - klasifikátor

Hledá se nadrovina, která v prostoru atributů rozdělí trénovací data do dvou skupin

Rozdělovací nadrovina je lineární funkcí prostoru atributů

Data mining v programu Rapidminer



Evolutionary algorithms

- pokus o řešení problému s příliš mnoha příznaky
- nevedlo to k dobrým výsledkům

Další

- rozpoznávání obrazu
- jpeg / mp3
- deep learning