

Whale detection - Brainstorming session

Jiří Dutkevič

Lenka Kovářová

Milan Le

Signal processing, Sampling theorem

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

Původní spojité signál může být zrekonstruován ze získaných vzorků.

Signal processing, Aliasing

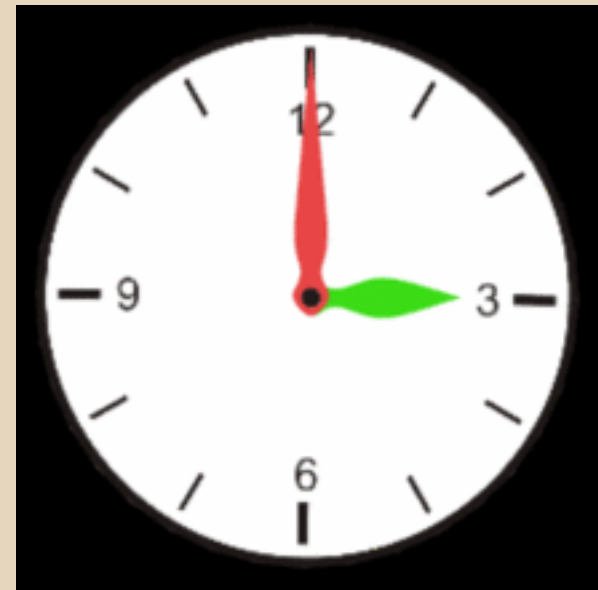
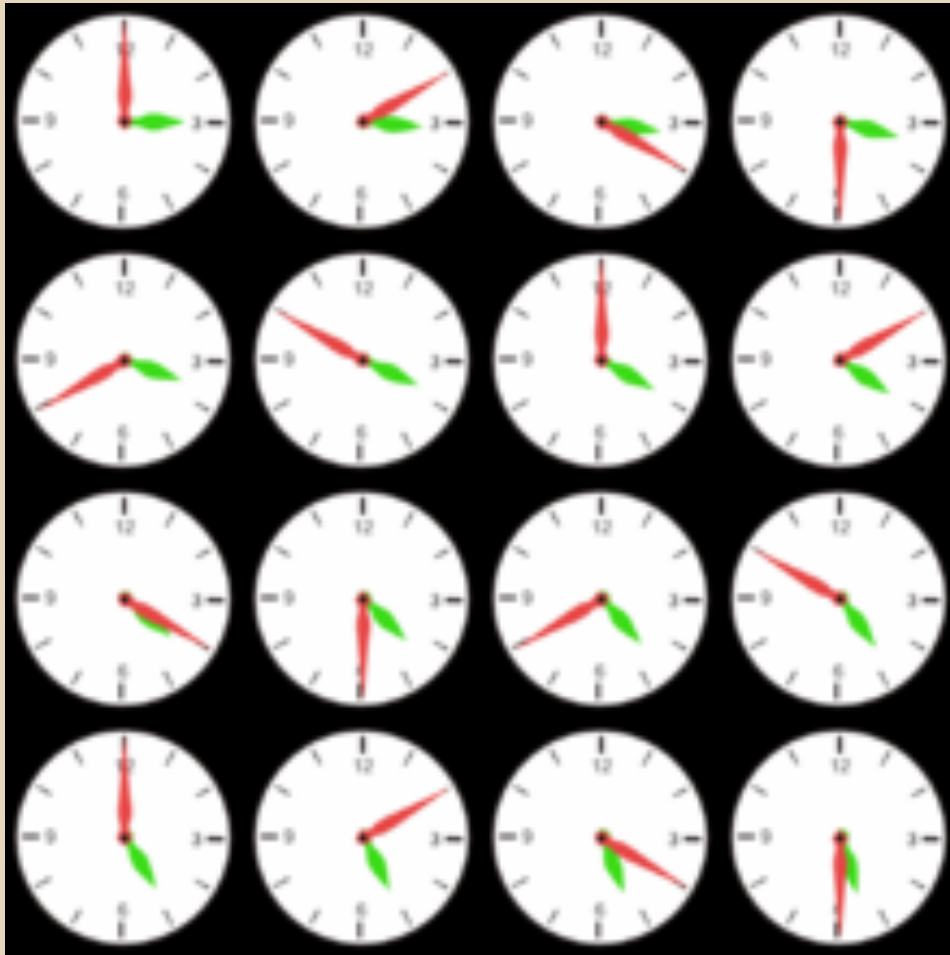
Jinými slovy: $F_s/2$ (*Nyquist frekvence*) je nejvyšší frekvence, které lze docílit rozkladem.

Ve filmu se vrtule točí:

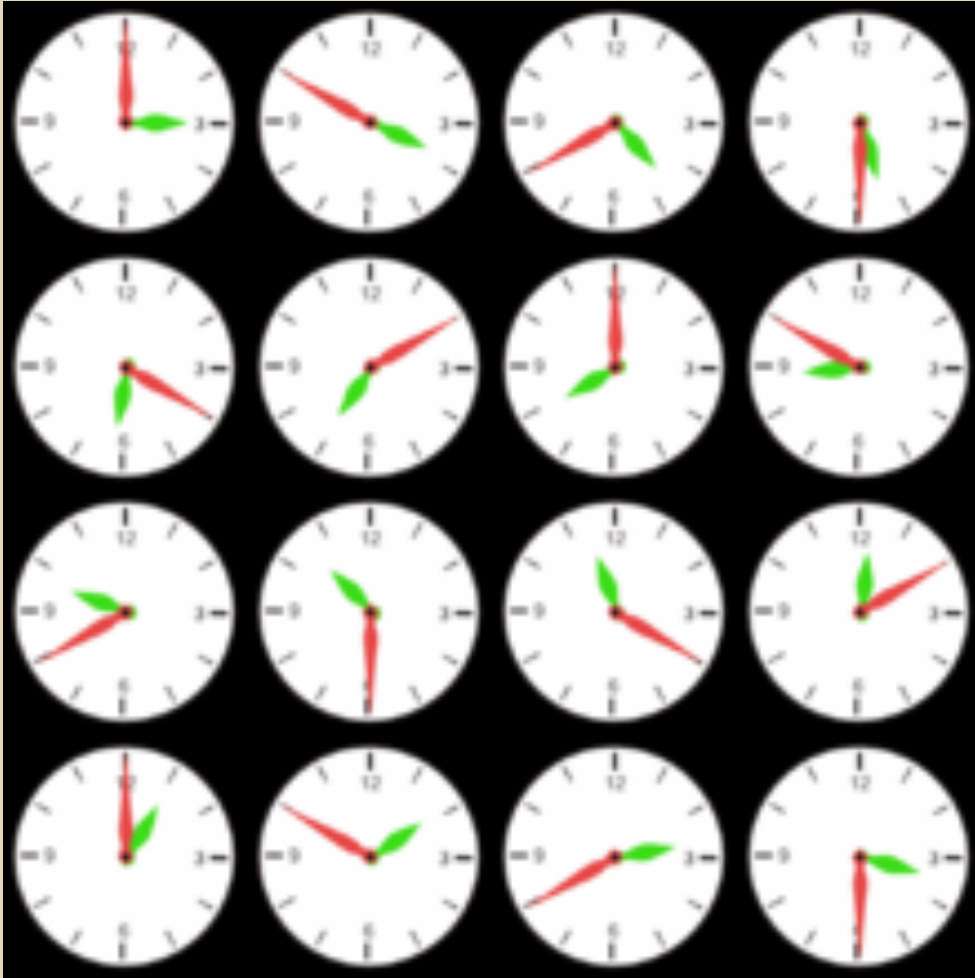
- V opačném směru než ve skutečnosti
- Nepřirozeně pomalu

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

Signal processing, Aliasing



Signal processing, Aliasing



Antialiasing

Spojité signál -> antialiasing filter -> diskretní 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

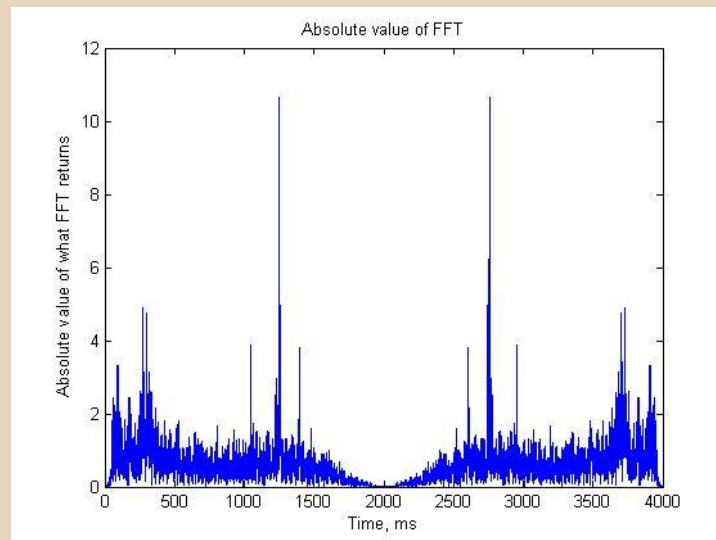
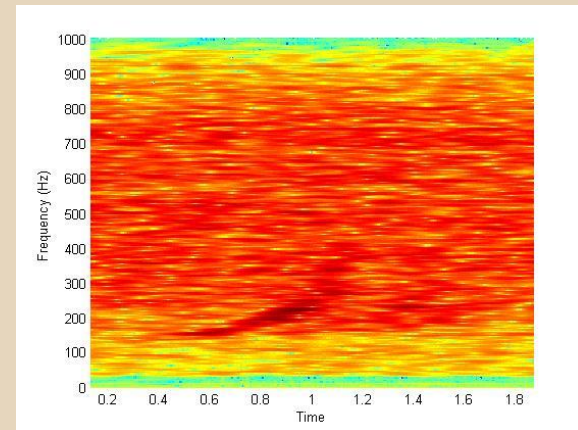
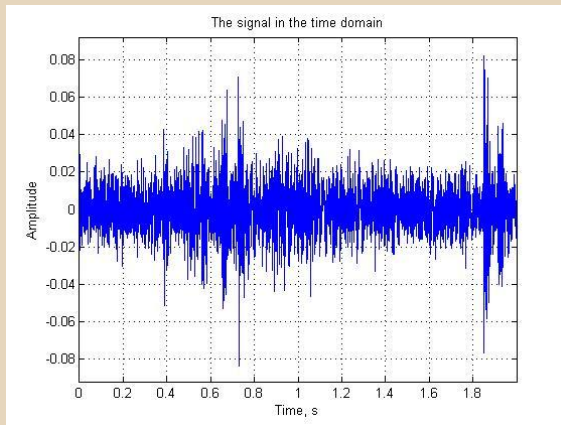
Vážení signálu s okénkovou funkcí

Pozadujeme

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

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

DFT, Spectral analysis

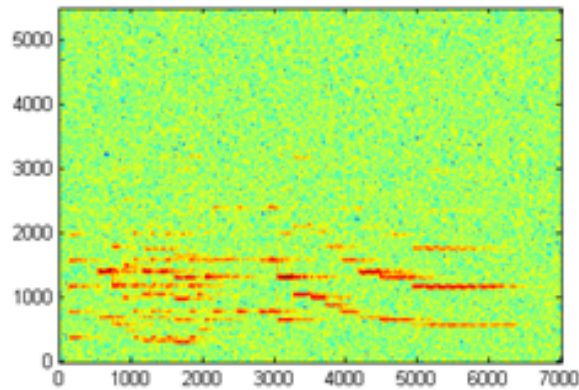


Denoising "odšumění"

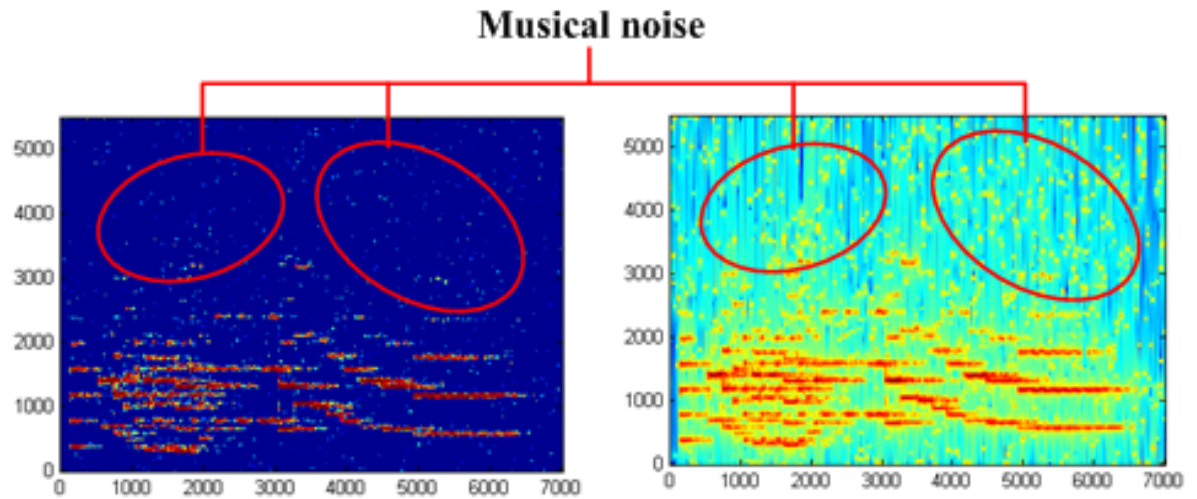
Audio Denoising by Time-Frequency Block Thresholding. Volba všech parametrů adaptivně na vlastnostech signálu minimalizací Steinovým odhadem risku.

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

Power subtraction



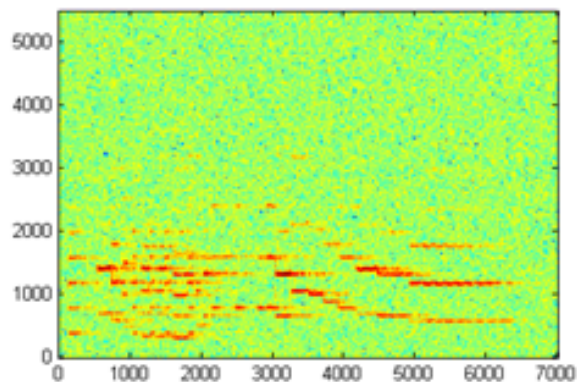
Noisy spectrogram



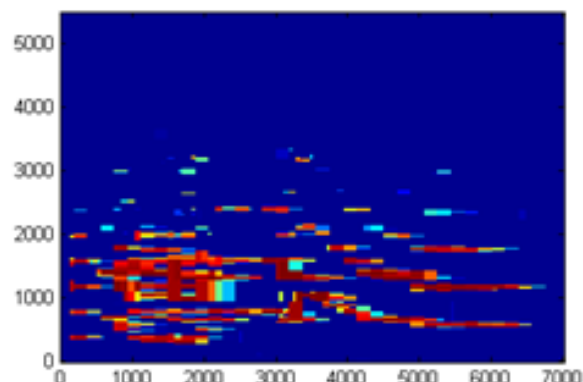
Attenuation factor of power subtraction. The isolated coefficients restore isolated short-time Fourier structures perceived as musical noise.

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

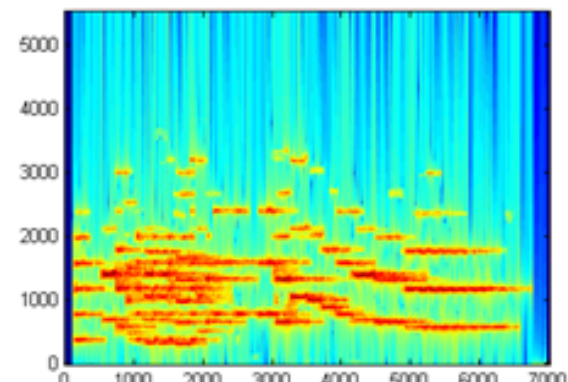
Block thresholding



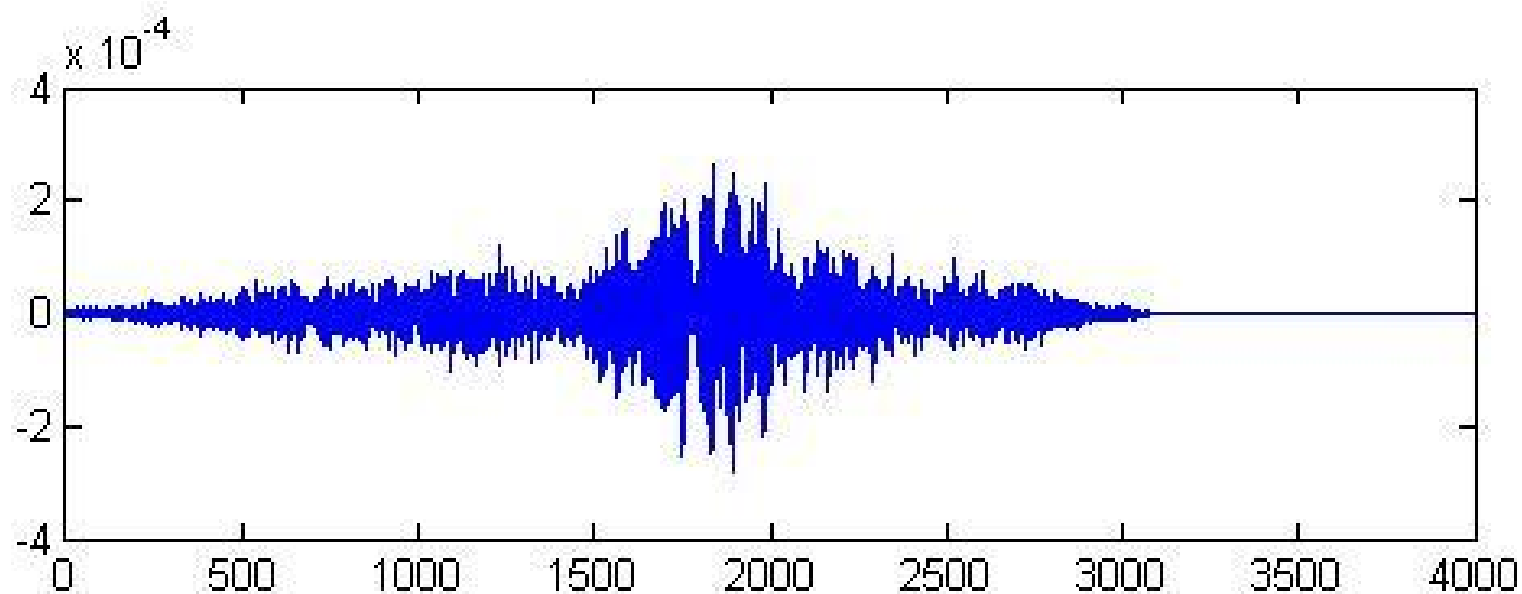
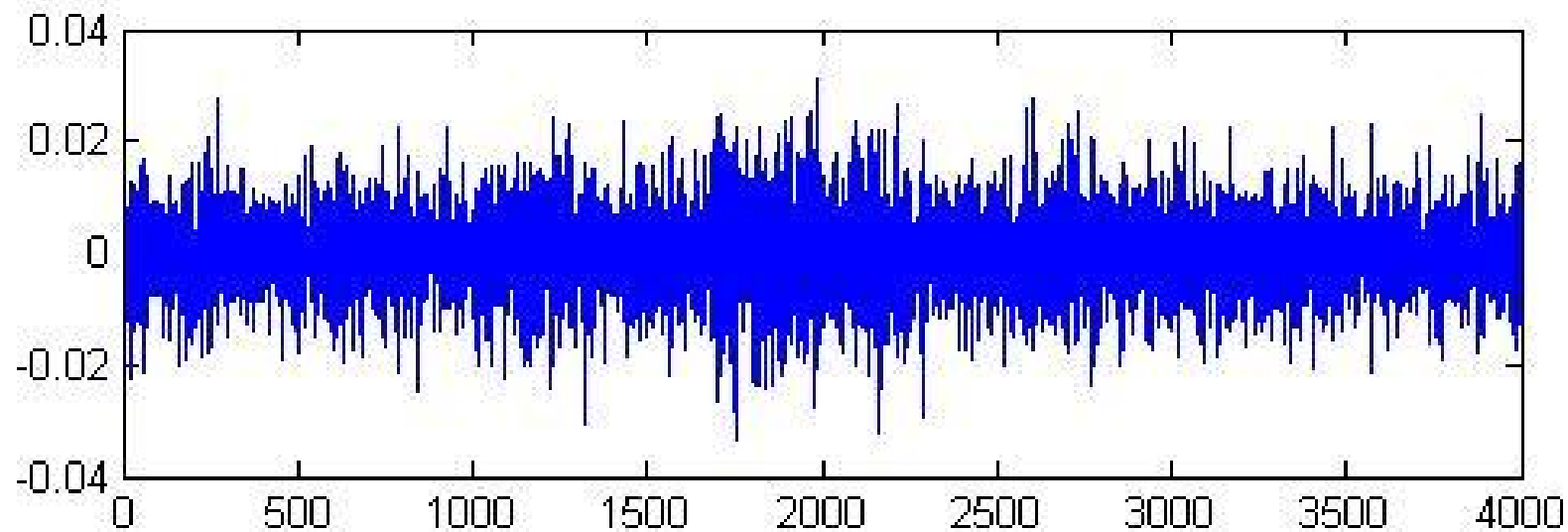
Noisy spectrogram

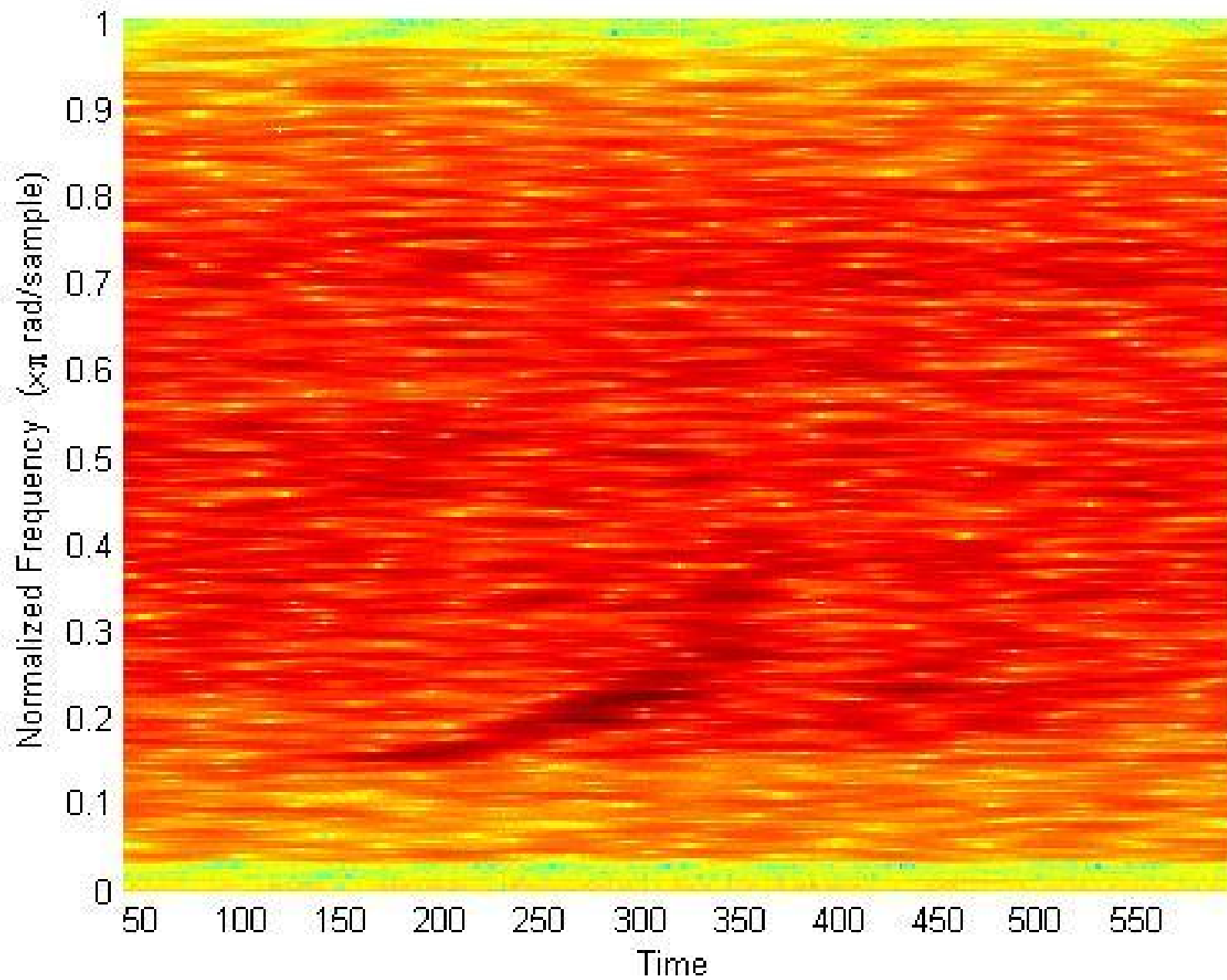


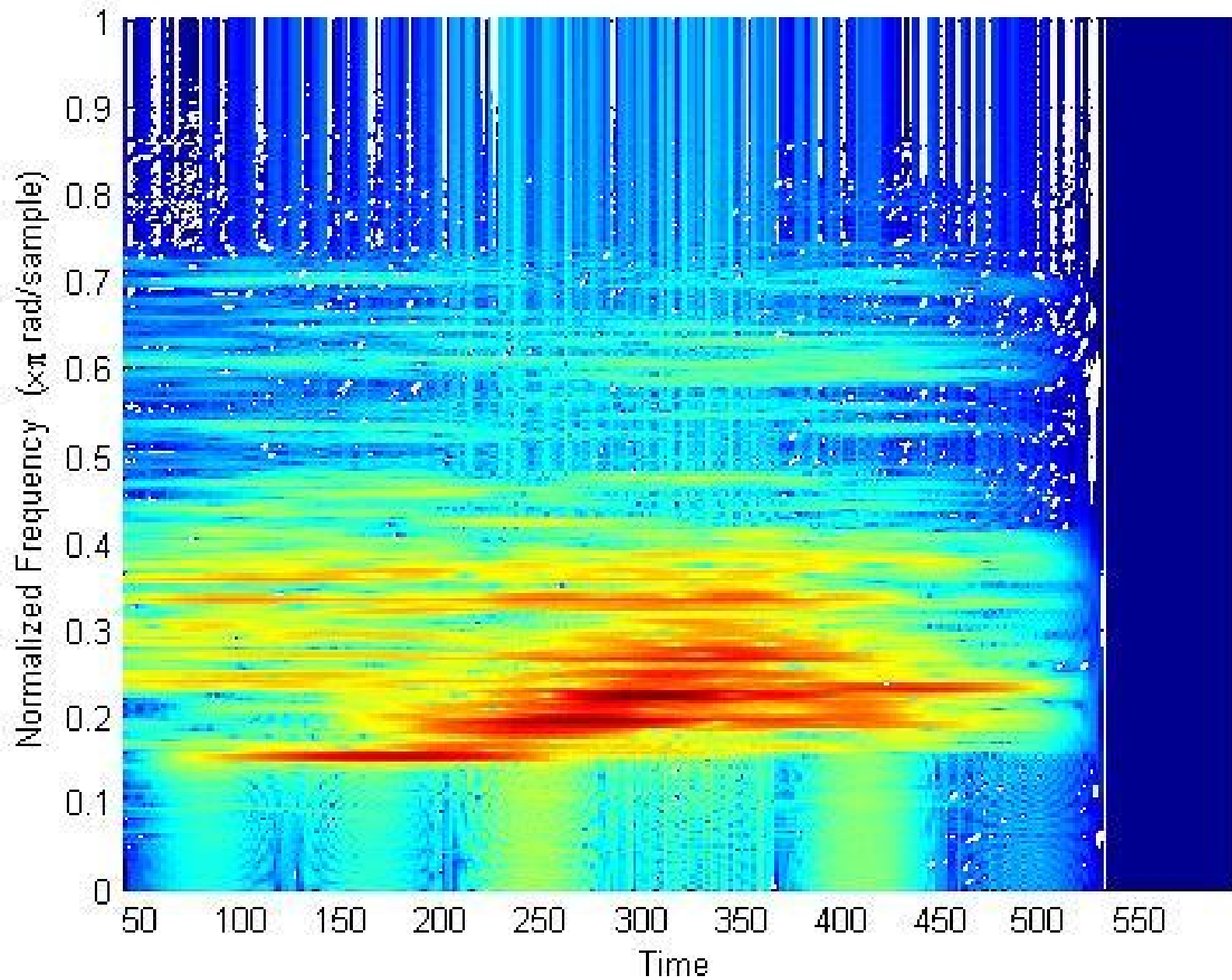
Attenuation factors of block thresholding.
Spectrogram coefficients are grouped in blocks whose size is adaptive to signal property. There is no isolated coefficients responsible for musical noise.



Spectrogram of denoised signal.







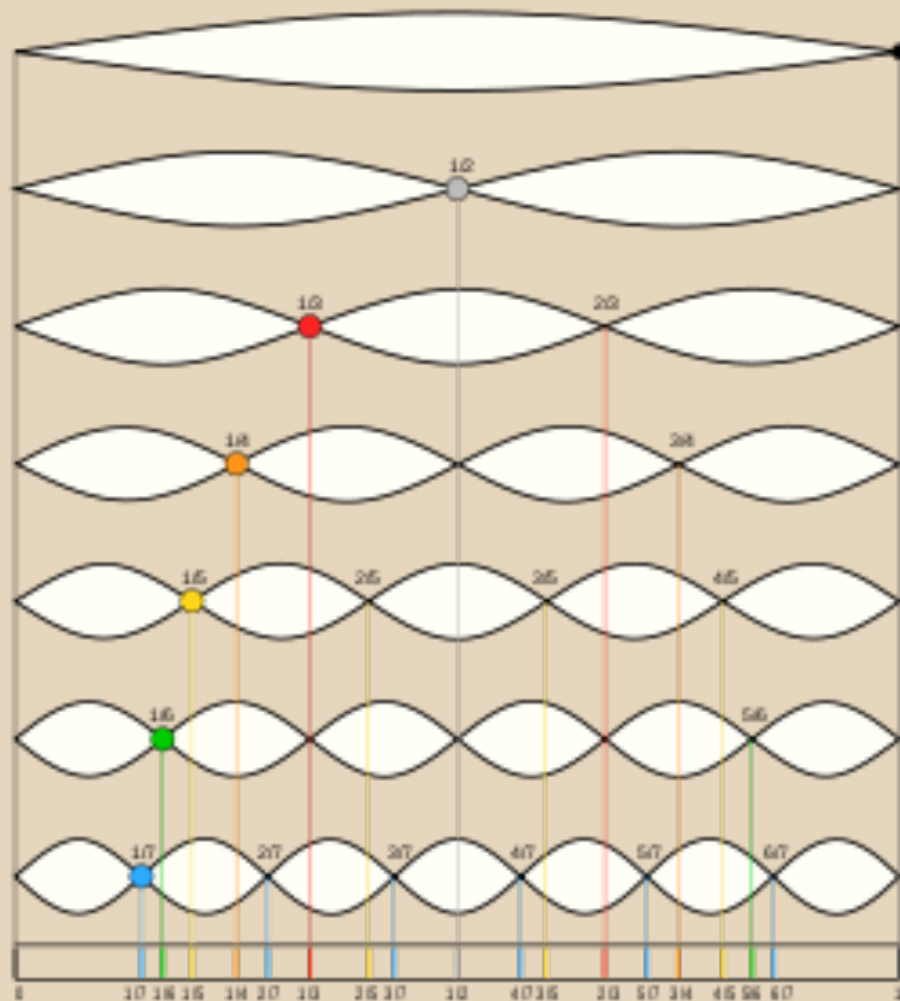
Audio Features

Spectralní features

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

```
c=real(iffcft(log(abs(ffcft(x))))))
```

- Mel-freq cepstral koefy
- spektrální centroid, bandwidth

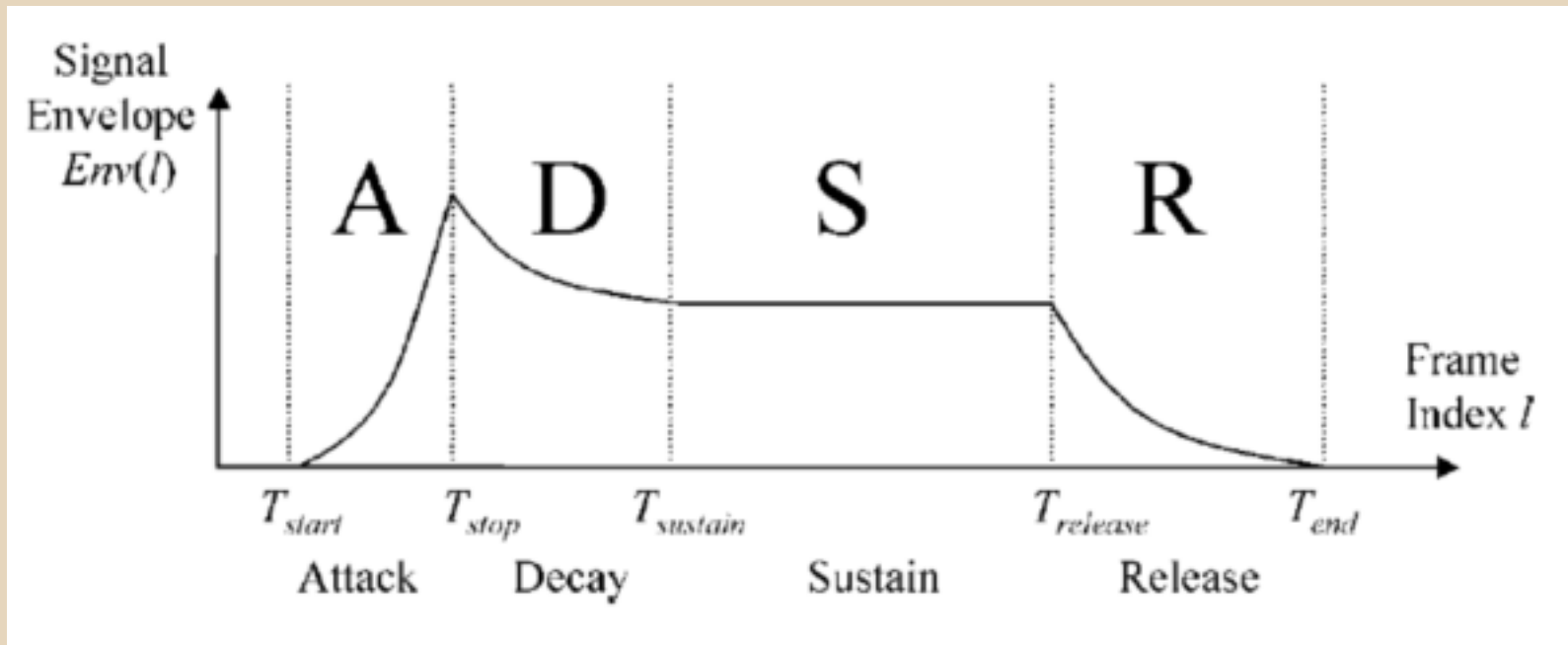


Harmonické vlny

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

Referencencies

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

Training Confusion Matrix

Output Class	0	1	
0	14871 70.8%	1406 6.7%	91.4% 8.6%
1	1143 5.4%	3580 17.0%	75.8% 24.2%
	92.9% 7.1%	71.8% 28.2%	87.8% 12.1%
	0	1	Target Class

Validation Confusion Matrix

Output Class	0	1	
0	3219 71.5%	280 6.2%	92.0% 8.0%
1	237 5.3%	764 17.0%	76.3% 23.7%
	93.1% 6.9%	73.2% 26.8%	88.5% 11.5%
	0	1	Target Class

Training Confusion Matrix

Output Class	0	1	
0	14973 71.3%	1340 6.4%	91.8% 8.2%
1	1112 5.3%	3575 17.0%	76.3% 23.7%
	93.1% 6.9%	72.7% 27.3%	88.3% 11.7%
	0	1	Target Class

Validation Confusion Matrix

Output Class	0	1	
0	3246 72.1%	295 6.6%	91.7% 8.3%
1	220 4.9%	739 16.4%	77.1% 22.9%
	93.7% 6.3%	71.5% 28.5%	88.6% 11.4%
	0	1	Target Class

Test Confusion Matrix

Output Class	0	1	
0	3245 72.1%	290 6.4%	91.8% 8.2%
1	258 5.7%	707 15.7%	73.3% 26.7%
	92.6% 7.4%	70.9% 29.1%	87.8% 12.2%
	0	1	Target Class

All Confusion Matrix

Output Class	0	1	
0	21335 71.1%	1976 6.6%	
1	1638 5.5%	5051 16.8%	
	92.9% 7.1%	71.9% 26.1%	
	0	1	Target Class

Training Confusion Matrix

Output Class	0	1	
0	14838 70.7%	1553 7.4%	90.5% 9.5%
1	1203 5.7%	3406 16.2%	73.9% 26.1%
	92.5% 7.5%	68.7% 31.3%	86.9% 13.1%
	0	1	Target Class

Validation

Output Class	0	1	
0	3167 70.4%		
1	278 6.2%		
	91.9% 8.1%		
	0	1	Target Class

Test Confusion Matrix

Output Class	0	1	
0	3170 70.4%	319 7.1%	90.9% 9.1%
1	252 5.6%	759 16.9%	75.1% 24.9%
	92.6% 7.4%	70.4% 29.6%	87.3% 12.7%
	0	1	Target Class

All Confusion Matrix

Output Class	0	1	
0	21389 71.3%	1954 6.5%	91.6% 8.4%
1	1584 5.3%	5073 16.9%	76.2% 23.8%
	93.1% 6.9%	72.2% 27.8%	88.2% 11.8%
	0	1	Target Class

Test Confusion Matrix

Output Class	0	1	
0	3242 72.0%	340 7.6%	90.5% 9.5%
1	245 5.4%	673 15.0%	73.3% 26.7%
	93.0% 7.0%	66.4% 33.6%	87.0% 13.0%
	0	1	Target Class

All Confusion Matrix

Output Class	0	1	
0	21247 70.8%	2224 7.4%	90.5% 9.5%
1	1726 5.8%	4803 16.0%	73.6% 26.4%
	92.5% 7.5%	68.4% 31.6%	86.8% 13.2%
	0	1	Target Class

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

Local Repository/velryby - RapidMiner 5.3.007 @ Pocatac

File Edit Process Tools View Help

Process

Operators

- Process Control (37)
- Utility (52)
- Repository Access (6)
- Import (27)
- Export (18)
- Data Transformation (114)
- Modeling (118)
 - Classification and Regression (53)
 - Lazy Modeling (2)
 - Bayesian Modeling (2)
 - Naive Bayes
 - Naive Bayes (Kernel)
 - Tree Induction (8)
 - Rule Induction
 - Single Rule Induction
 - Single Rule Induction (Single Attribut
 - Subgroup Discovery
 - Tree to Rules

Repositories

- Samples (none)
- DB
- Local Repository (Lenka_2)

Main Process

Read Train data

Split Data

Discretize (2)

Multiply train ...

Multiply test s...

Read Test data

Neural Net

Random Forest

SVM

Naive Bayes

Apply Model (5)

Apply Model

Apply Model (3)

Apply Model (7)

Performance ...

Performance

Performance ...

Performance ...

Apply Model (6)

Apply Model (2)

Apply Model (4)

Apply Model (8)

Parameters

Context

Neural Net

hidden layers

training cycles

learning rate

momentum

decay

shuffle

normalize

error epsilon

use local random seed

Help Comment

Neural Net (RapidMiner Core)

Synopsis

This operator learns a model by means of a feed-forward neural network trained by a back propagation algorithm (multi-layer perceptron). This operator cannot handle polynomial attributes.

Problems

Log

3 potential problems

Message	Fixes	Location
Input example set has example weights, but the learner will ignore them.	No quick fix available	Random Forest.training set
Naive Bayes cannot handle numerical label.	Choose among 5 available quick fi...	Naive Bayes.training set
Random Forest cannot handle numerical label.	Choose among 5 available quick fi...	Random Forest.training set

Evolutionary algorithms

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

Další

- rozpoznávání obrazu
- jpeg / mp3
- deep learning