Musical Instruments Signal Analysis and Recognition Using Fractal Features

Athanasia Zlatintsi and Petros Maragos

School of ECE, National Technical University of Athens 15773, Greece

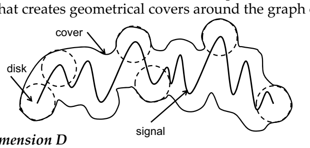
http://cvsp.cs.ntua.gr

1. Outline - Contributions

- Analyze musical signals using the theory of fractals: model the geometrical complexity of the waveform via the fractal dimension
- Novel approach for the analysis of the structure of musical instruments signals at multiple levels
- Timbre analysis for the distinction of different instrument classes and exploration of the differences between the attack and steady state of the tones
- Goal: Feature configurations focusing at multiple scales and the statistical self-similarity of the signals that offer a modest improvement to the performance of a recognition task

2. Multiscale Fractal Dimension(MFD)

- *Short-time fractal dimension* D: approximately quantifies the degree of turbulence.
- *Multiscale fractal dimension profile* MFD[s]: quantifies the multiscale complexity of the waveform, thus the degree of fragmentation.
- Algorithm: based on multiscale nonlinear operators of morphological filtering that creates geometrical covers around the graph of the signal

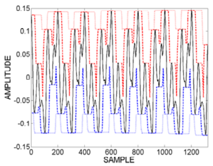


Fractal Dimension D

$$D = \lim_{s \to 0} \frac{\log[(\text{Area of dilated graph by disks of radius s}) / s^2]}{\log(1/s)} \quad (1)$$

Algorithm

- Measure the area of the difference between dilations and erosions
  - Structuring Element: symmetric 3-sample flat line segment
- Least-squares fit a straight line to the log-log data of (1) over a small scale window $\{s, s+1, \ldots s+w-1\}$ of w scales
- Move along the s axis and create a profile of local multiscale fractal dimensions $D(s,t)$ at each time location t of the short time analysis frame.
- D ∈ [1,2] for 1-D signals
- The larger D, the larger the amount of geometrical fragmentation.

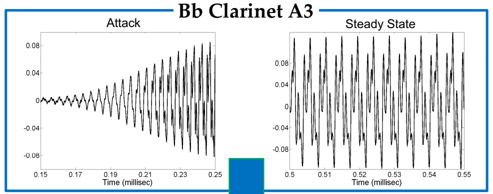


Bb Clarinet A3 Steady state (solid line) and its dilation (red lines) & erosions (blue lines) at scales s = 25, 75, with D = 1.1.

3. MFD on attack & steady state

1st Step: Processing of the input signal

Bb Clarinet A3

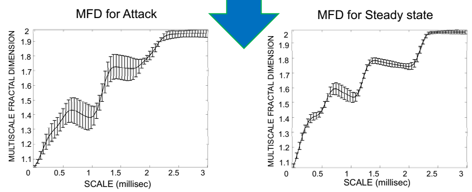


2nd Step: Analysis Using Multiscale Fractal Exponents

MFD

Calculate the short time MFDs of the tones, 30-ms segments, updated every 15-ms

- Sampling Frequency: 44.1KHz
- MFD[s] are analyzed for discrete scales $s = 1, \ldots, 133$
- Time scales s_t from 1/44.1 to 3 ms

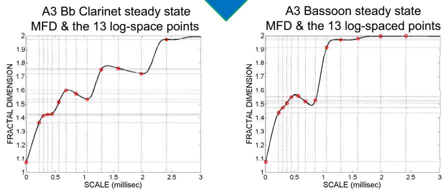


3rd Step: Feature set configuration

Feature sets

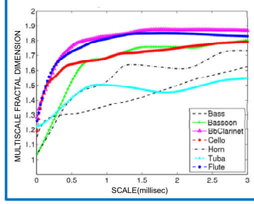
- 13 MFD[log-spaced] features (MFDLG)
- 6 PCA components (MFDPC)

examples



4. Analysis cases

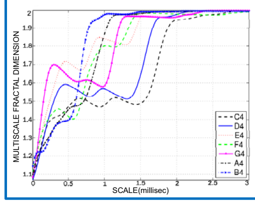
Average MFDs on Attack



Estimated for 7 instruments

- Similar as for the steady state
- Higher D for small scales s_t and more fragmentation.
- Increased value of $D(s = 1)$.
- Clear distinction of D among some of the analyzed instruments.

Bb Clarinet (C4-B4) Steady State MFD Variability



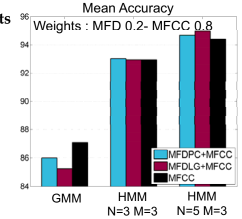
- Dependence on the acoustical frequency and the MFD profile that increases rapidly for higher frequency sounds.
- The instruments' specific MFDs beholds the shape observed for the specific octave.

5. Experimental Evaluation

4th Step: Recognition Experiments

Experimental Setup

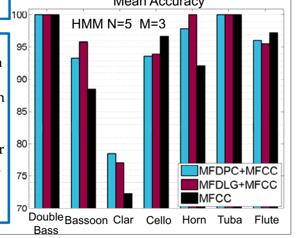
- 1333 notes, 7 instruments
- 5 feature sets
  - 6 MFDPC
  - 13 MFDLG
  - 13 MFCC
  - 6 MFDPC + 13 MFCC
  - 13 MFDLG + 13 MFCC
- 2 classification methods
  - GMM M[1-3] mixtures
  - Multi-stream HMM N[3-9] states and M[1-3] mixtures



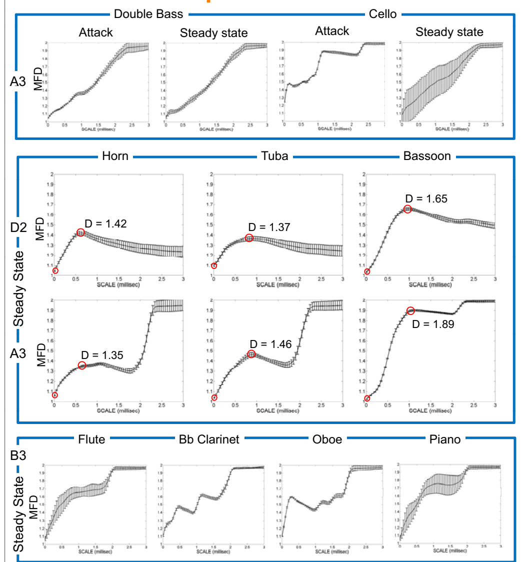
Results

MFD+MFCC — MFD

- Double Bass, Bassoon & Tuba best recognized
- Low discriminability between Bb Clarinet & Flute
- Enhanced discriminability for Bassoon, Bb Clarinet and Horn
- Decreased for Cello & Flute



6. More examples



References

[1] M. Bigerelle and A. Iost, Fractal dimension and classification of music, *Chaos, Solitons, & Fractals,* vol. 11, pp.2179-2192, 2000.

[2] A. Eronen, Comparison of features for musical instrument recognition, in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.,* pp. 19-22, 2001.

[3] K. J. Hsu and A. J. Hsu, Fractal geometry of music, in *Proc. Natl. Acad. Sci. USA Physics,* vol. 87, pp. 938-941, Feb. 1990.

[4] B. Mandelbrot, The Fractal Geometry of Nature, W.H. Freeman, San Francisco, 1982.

[5] P. Maragos, Fractal signal analysis using mathematical morphology, *Advances in electronics and electron physics, Academic Press,* vol. 88, pp. 199-246, 1994.

[6] P. Maragos and A. Potamianos, Fractal dimension of speech sounds: Computation and application to automatic speech recognition, *J. Acoust. Soc. Am.,* vol. 105, no. 3, pp. 1925-1932, 1999.

[7] Z.-Y. Su and T. Wu, Music walk, fractal geometry in music, *Physica A,* vol. 380, pp. 418-428, 2007.

[8] G. Tzanetakis and P. Cook, Musical genre classification of audio signals, *IEEE Transactions on Speech and Audio Processing,* vol. 10(5), pp. 292-302, July 2002.

[9] R. F. Voss and J. Clarke, '1/f noise' in music and speech, *Nature,* vol. 258, pp. 317-318, Nov. 1975.

[10] University of Iowa Musical Instrument Sample Database. [Online]. Available: http://theremin.music.uiowa.edu/index.html

Acknowledgments

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) – Research Funding Program: Heracleitus II. Investing in knowledge society through the European Social Fund.

More information : Please contact: *[nzlat,maragos]@cs.ntua.gr*