



Project no. FP6-507752

MUSCLE

Network of Excellence
Multimedia Understanding through Semantics, Computation and
Learning

DN 4.2.1: Feature Extraction Tools for Audio

Due date of deliverable: 30.11.2006
Actual submission date: 19.02.2007

Start date of project: 1 March 2004

Duration: 48 months

Deliverable Type: PU
Number: DN4.2.1
Nature: P
Task: WP4

Name of responsible: Andreas Rauber, TU Vienna-IFS

Revision 1.0

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Feature Extraction Tools for Audio

Within the [MUSCLE Network of Excellence](#) on multimedia understanding, datamining and machine learning researchers have developed a range of tools for audio analysis, speech recognition, sound description and music retrieval. This deliverable (DN 4.2) of WP4 represents an inventory of current audio feature extraction tools:

- [WinSnoori Speech Analysis Software](#)
- [RPextract Music Feature Extractor](#)
- [Sound Description Toolbox](#)

WinSnoori Speech Analysis Software

INRIA-Parole, Yves Laprie

Using tools for investigating speech signals is an invaluable help to teach phonetics and more generally speech sciences. For several years we have undertaken the development of the software WinSnoori which is for both speech scientists as a research tool and teachers in phonetics as an illustration tool. It consists of five types of tools:

- to edit speech signals,
- to annotate phonetically or orthographically speech signals. WinSnoori offers tools to explore annotated corpora automatically,
- to analyse speech with several spectral analyses and monitor spectral peaks along time,
- to study prosody. Besides pitch calculation it is possible to synthesise new signals by modifying the F0 curve and/or the speech rate,
- to generate parameters for the Klatt synthesiser. A user friendly graphic interface together with copy synthesis tools (automatic formant tracking, automatic amplitude adjustment) allows the user to generate files for the Klatt synthesiser easily.

In the context of speech sciences WinSnoori can therefore be exploited for many purposes, among them, illustrating speech phenomena and investigating acoustic cues of speech sounds and prosody.

Download (V 1.34-03): [WinSnoori_1.34_setup.exe](#)

Details and Guide: <http://www.loria.fr/~laprie/WinSnoori/index.html>

RPextract Music Feature Extractor

TU Vienna - IFS, [Thomas Lidy](#)

Content-based access to audio files, particularly music, requires the development of feature extraction techniques that capture the acoustic characteristics of the signal, and that allow the computation of similarity between pieces of music. At TU Vienna - IFS three different sets of descriptors were developed:

- Statistical Spectrum Descriptors: describe fluctuations by statistical measures on critical frequency bands of a psycho-acoustically transformed Sonogram
- Rhythm Patterns: reflect the rhythmical structure in musical pieces by a matrix describing the amplitude of modulation on critical frequency bands for several modulation frequencies
- Rhythm Histograms: aggregate the energy of modulation for 60 different modulation frequencies and thus indicate general rhythmic in music

The algorithm considers psycho-acoustics in order to resemble the human auditory system. The feature extractor is currently implemented in Matlab and processes au, wav, mp3 and ogg files. Feature vectors are output in SOMLib format.

Download (V 0.58) : [RP_extract_0.58.zip](#)

Usage Guide: http://www.ifs.tuwien.ac.at/mir/howto_matlab_fe.html

Sound Description Toolbox

AUTH, Emmanouil Benetos

The Sound Description Toolbox extracts a number of MPEG-7 standard descriptors as and other feature sets from WAV audio files. Features covered are:

- Energy: AudioPower
- Harmonic: AudioFundamentalFrequency
- Perceptual: Specific Loudness Sensation Coefficients
- Spectral: AudioSpectrumCentroid, Audio Spectrum Rolloff, AudioSpectrumSpread, MFCCs
- Temporal: Autocorrelation Coefficients, Log-attack Time, TemporalCentroid, Zero-crossing rate
- Various: AudioSpectrumFlatness

Instructions: Double-click on "ComputeFeatureMatrix.exe". A file-opener should appear, input the respective .wav file, say "file1.wav". After a short

period of time, a file named "file1.wav.fm" will appear on the path, containing the 1st and 2nd moments of 14 sound description features.

Download (V 0.1): [SoundDescriptionToolbox0.1.zip](#)