

Intelligent Audio Systems: A review of the foundations and applications of semantic audio analysis and music information retrieval



Jay LeBoeuf
Imagine Research
jay@imagine-research.com

July 2008

These lecture notes contain hyperlinks to the CCRMA Wiki.

On these pages, you can find supplemental material for lectures - providing extra tutorials, support, references for further reading, or demonstration code snippets for those interested in a given topic .

Click on the  symbol on the lower-left corner of a slide to access additional resources.

WIKI REFERENCES...



Day 9

- Projects
 - High-level review: your designs and process (5-10 min / person)
 - Suggestions from the crowd. Code sharing of non-creative work.
- Mini-lecture on MPEG-7 and some Real-World Applications
- Lunch
- Projects
- Perry Cook (Princeton / CS, SoundLab)

Matlab Central

- <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=18596&objectType=FILE>
- Common sharing code to wiki site.
<http://ccrma.stanford.edu/~mw/>

FEATURES: ADVANCED

MPEG-7

MPEG-7 aims to standardize a core set of quantitative measures of audio-visual features

"First international standard providing content-derived audio metadata"

Content-based indexing features

Example Applications:

- Making AV as searchable as text

- Transcriptive, physical, perceptual, and medium-based descriptions



MPEG-7 Example Applications

- General Sound Recognition / Indexing
- Archiving and Restoration
- Instrument Timbre Search
- Melody Search
- Audio Identification / Fingerprinting
- Surveillance
- Film, video, and radio archives (journalism)

MPEG-7

- Low Level Descriptors (17)
 - Basic
 - AudioWaveform (stats)
 - AudioPower (stats)
 - Silence (no significant sound)
 - Basic Spectral
 - Centroid, spread, flatness, etc
 - Signal Parameters
 - AudioFundamentalFrequency
 - AudioHarmonicity

MPEG-7

- Low Level Descriptors (17) cont'd
 - Timbral Temporal
 - LogAttackTime
 - Temporal Centroid
 - Timbral Spectral (specific to harmony perception)
 - Harmonic Spectral Centroid
 - Spectral Spread
 - Spectral Basis

MPEG-7

- Additional (added in 2003)
 - Audio signal quality description:
 - Background Noise, channel cross-correlation, relative delay, balance, DC offset, bandwidth, transmission technology
 - Errors in recordings (clicks, clipping, drop outs, ...)
 - Musical tempo [bpm]

MPEG-7

- High-level descriptors
 - Audio signature
 - Musical instrument timbre
 - melody description