

MPEG-7 Overview

Gabriel Vigliensoni

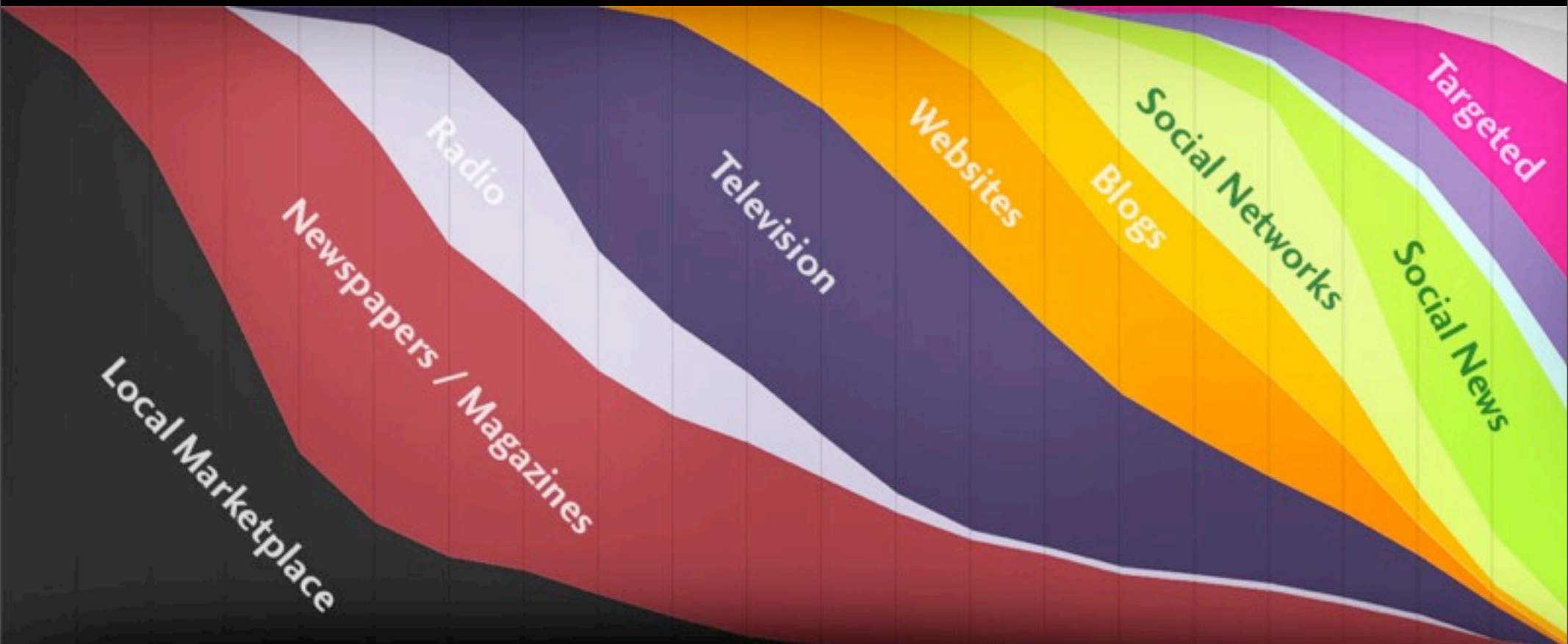
Music Technology Area
Schulich School of Music
McGill University

Background Information

Main Elements

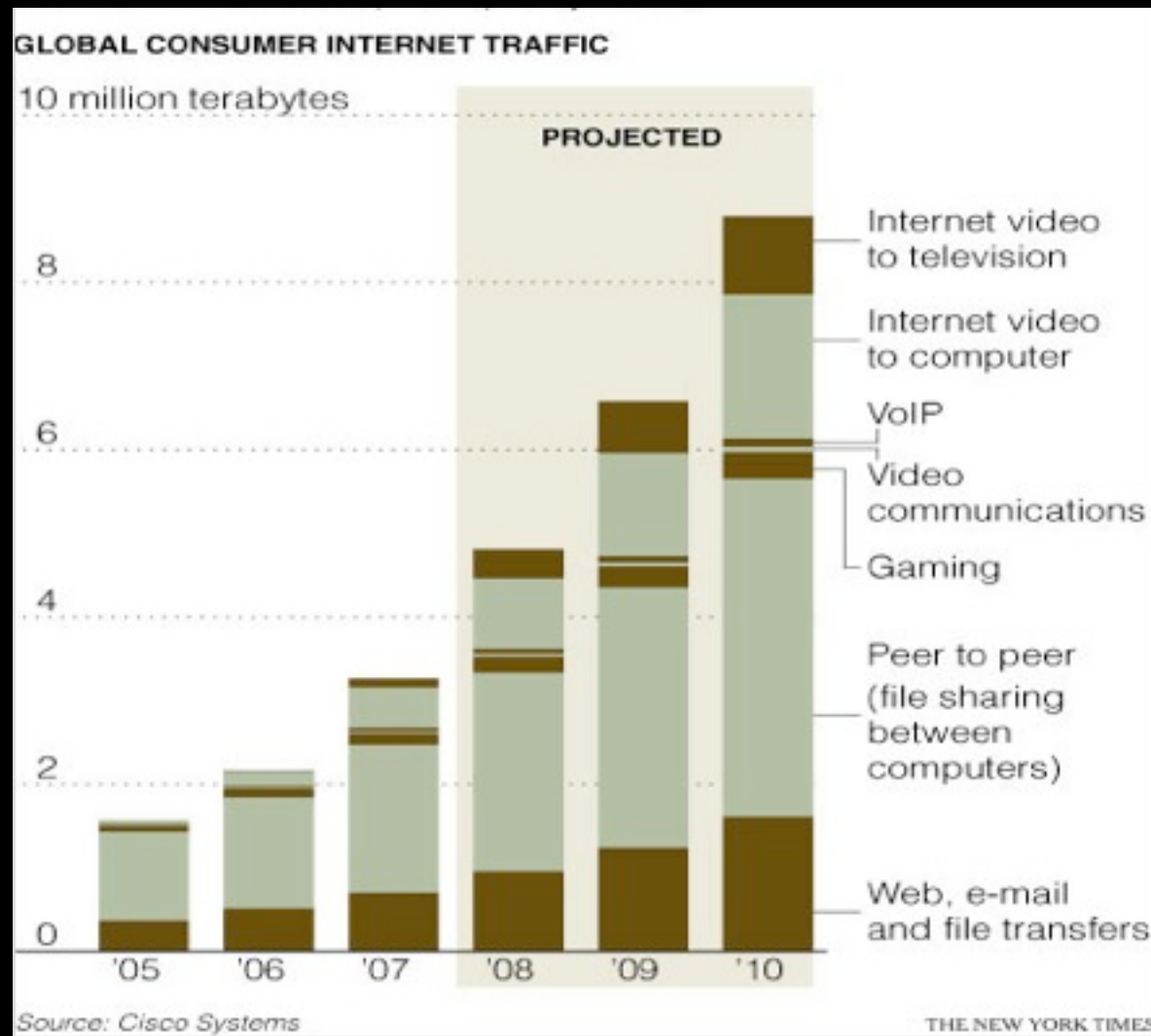
Applications

Information Explosion



Where is everyone? Baekdal, T. April 27, 2009
<http://www.baekdal.com/articles/Management/market-of-information/>

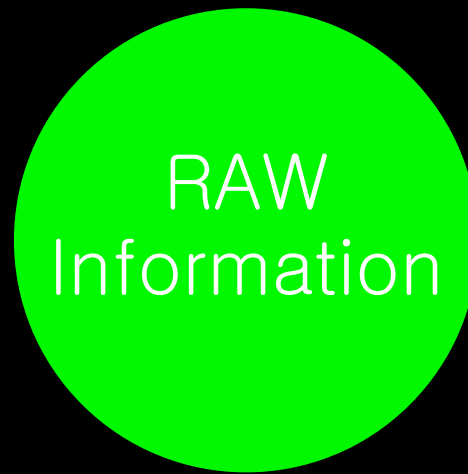
Information Explosion



The Internet Traffic Challenge: The Policy Dimension. The New York Times. March 13, 2008 <http://bits.blogs.nytimes.com/2008/03/13/the-internet-traffic-challenge-the-policy-dimension/>

found

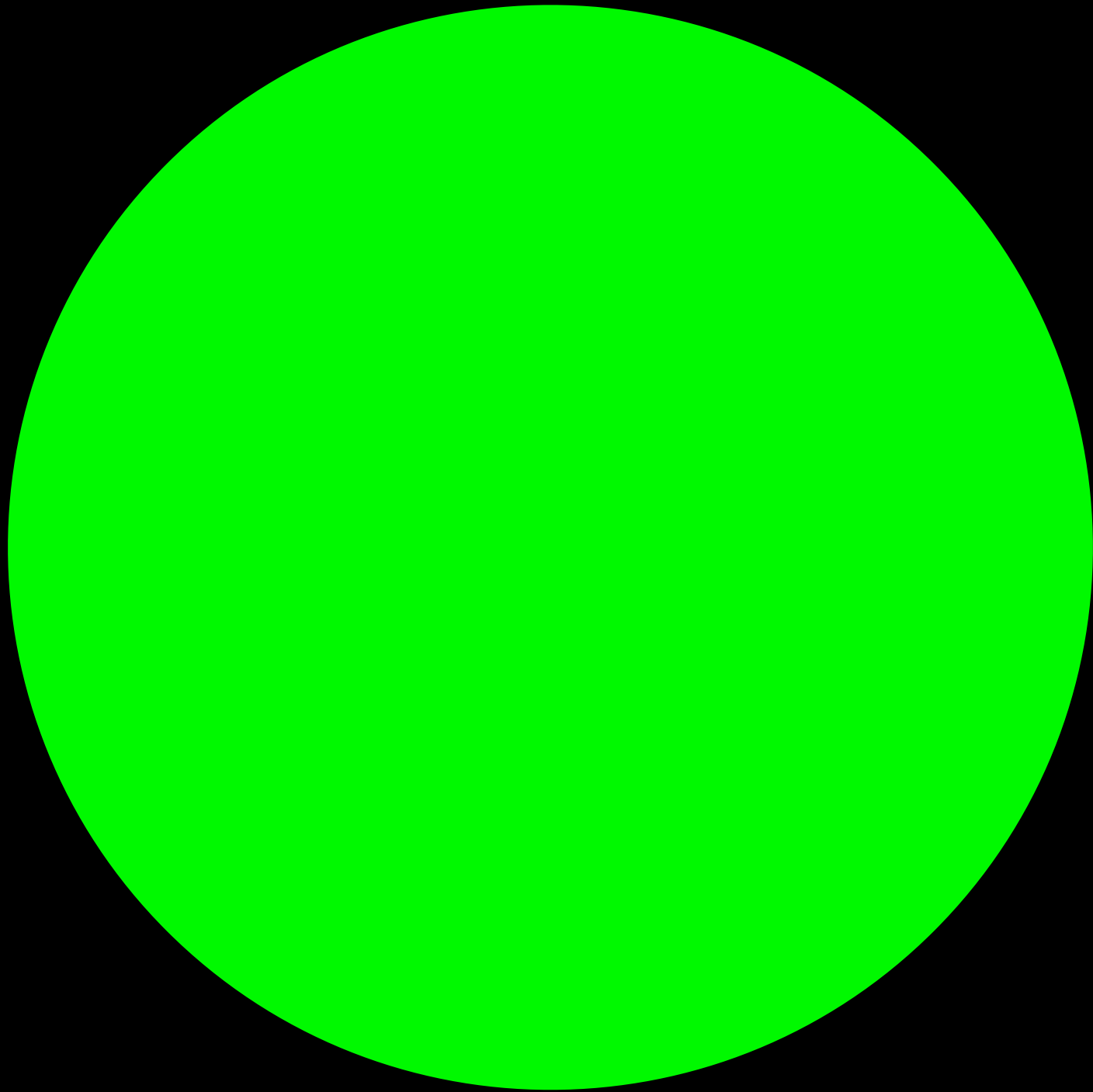
retrieve



RAW
Information

access

filter





INFORMATION

MPEG-7

(ISO/ICE 15938)

Multimedia
Content
Description
Interface

Scope



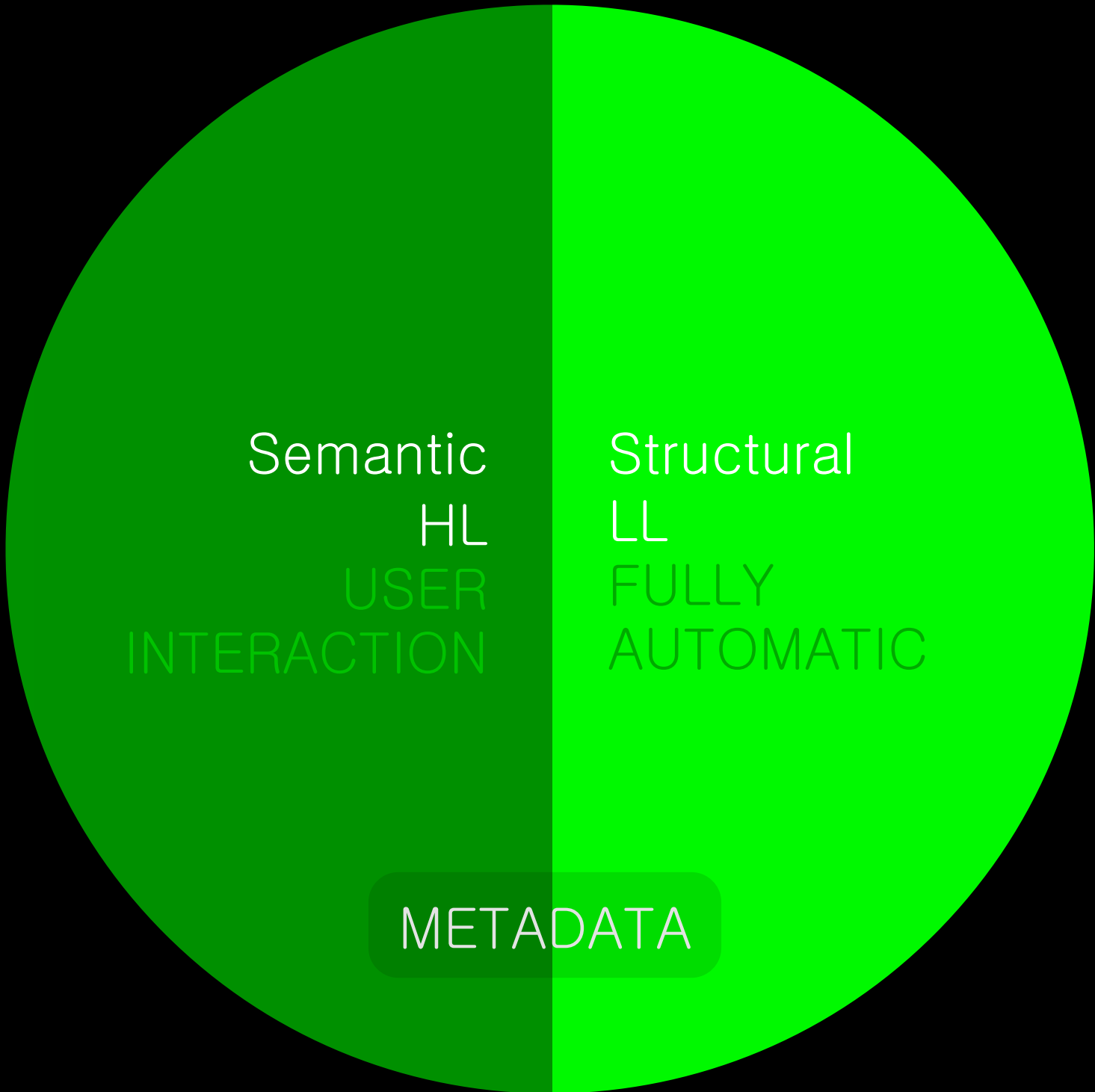
A/V Content
Description

Scope

Feature
Extraction

A/V Content
Description

Search
Engine



Interoperability

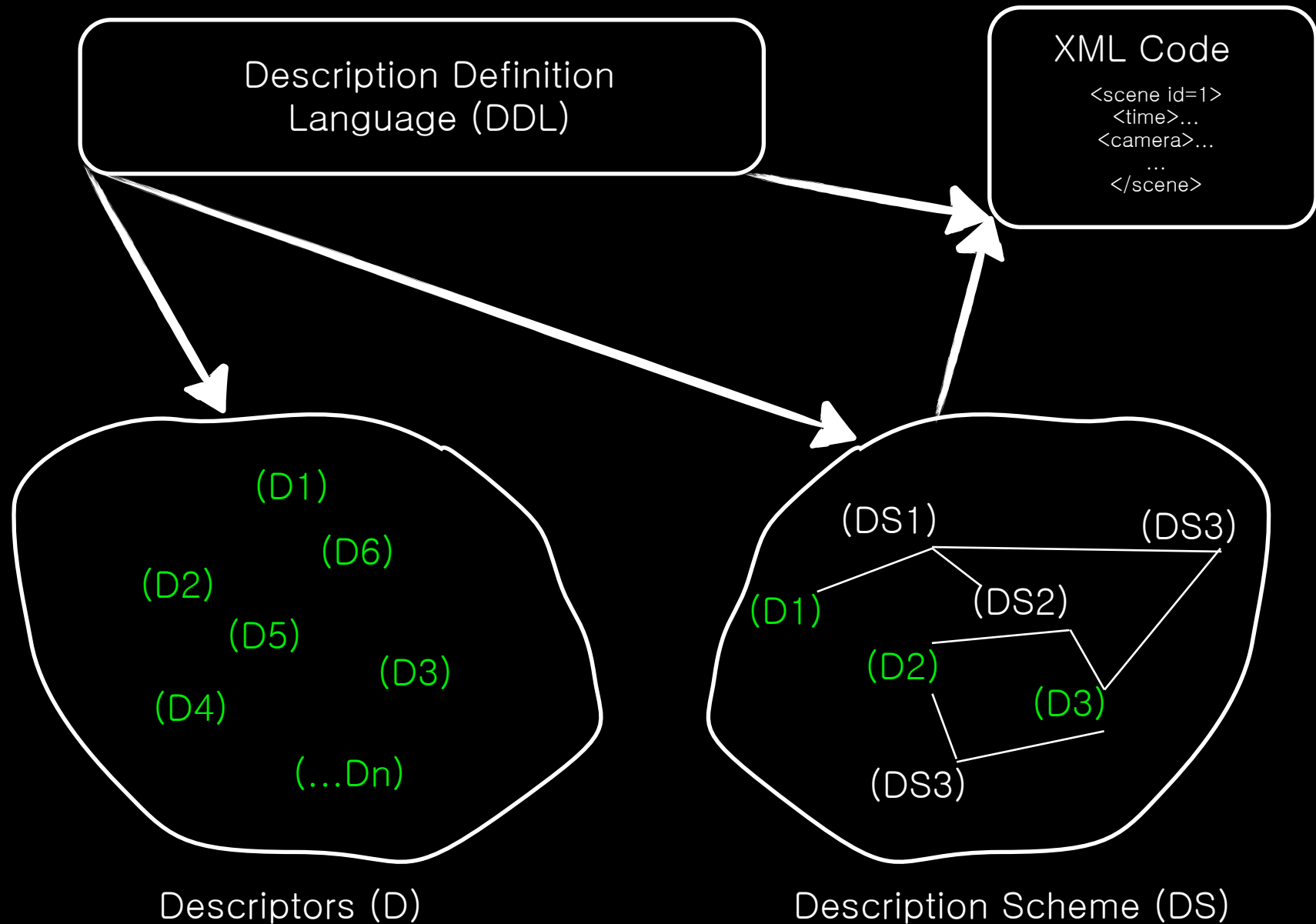
Feature
Extraction

A/V Content
Description

Search
Engine

XML Schema

Main Elements of MPEG-7



Descriptors

Visual

Color
Texture
Shape
Motion
Face

Audio

Basic
Basic Spectral
Signal Parameters
Timbral Temporal
Timbral Spectral
Spectral Basis

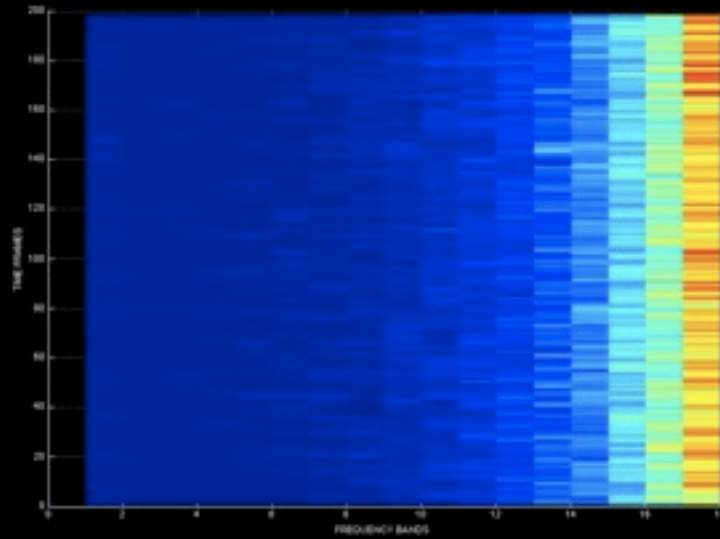
AudioDescriptors

Basic	AudioWaveform AudioPower
Basic Spectral	AudioSpectrumEnvelope AudioSpectrumCentroid AudioSpectrumSpread AudioSpectrumFlatness
Signal Parameters	AudioFundamentalFrequency AudioHarmonicity
Timbral Temporal	LogAttackTime TemporalCentroid
Timbral Spectral	HarmonicSpectralCentroid HarmonicSpectralDeviation HarmonicSpectralSpread HarmonicSpectralVariation & SpectralCentroid
Spectral Basis	AudioSpectrumBasis AudioSpectrumProjection

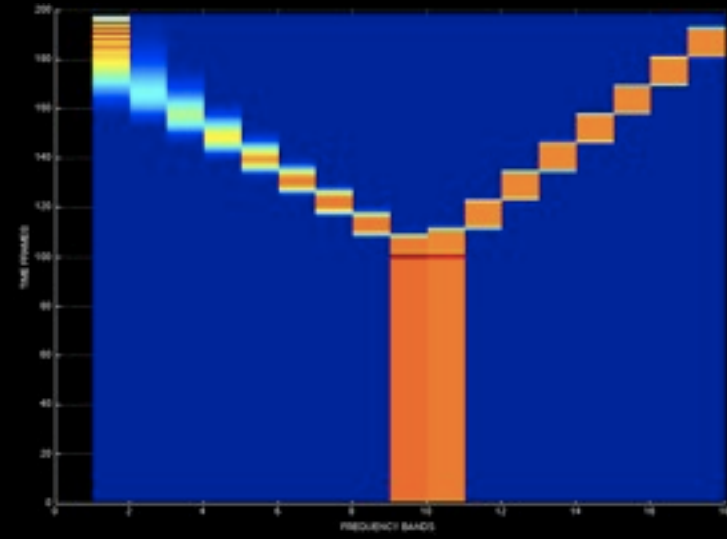
BasicSpectral Descriptors

AudioSpectrumEnvelope

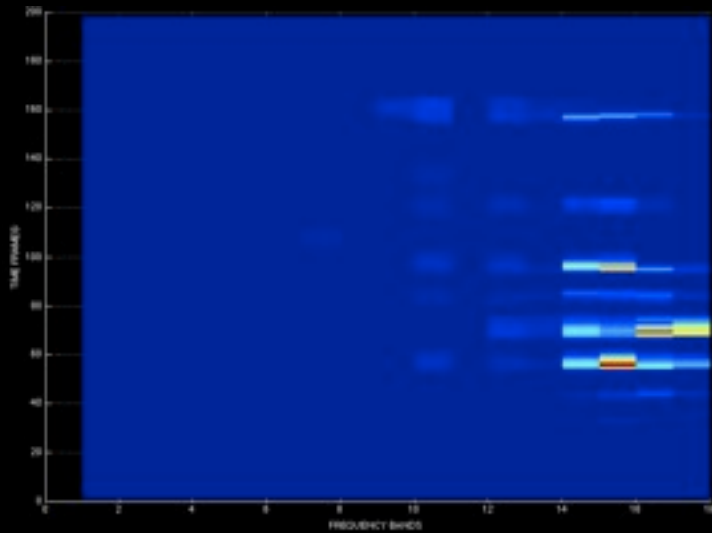
White Noise



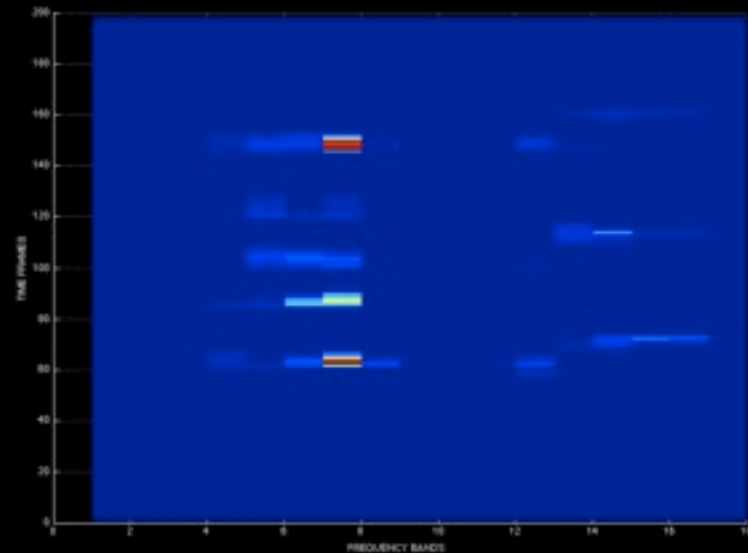
1Khz tone + double sweep



Trumpet



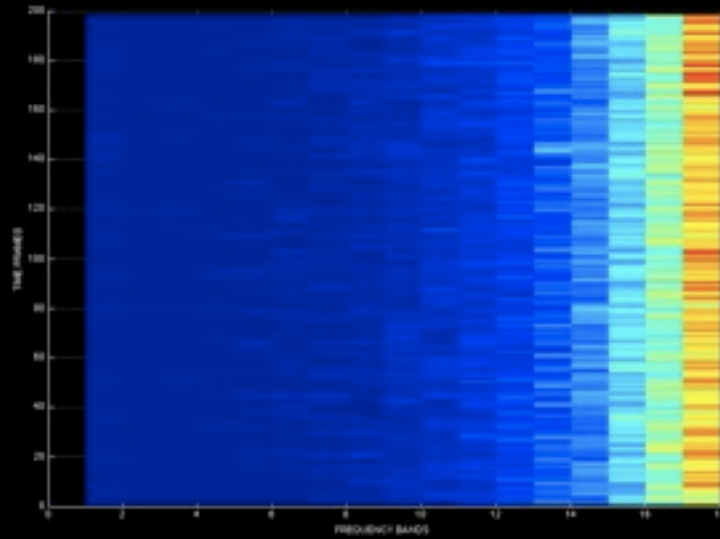
Speech



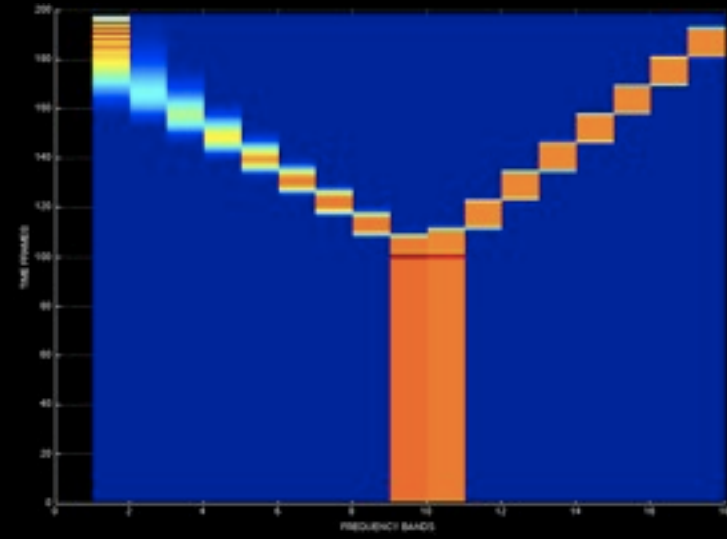
BasicSpectral Descriptors

AudioSpectrumEnvelope

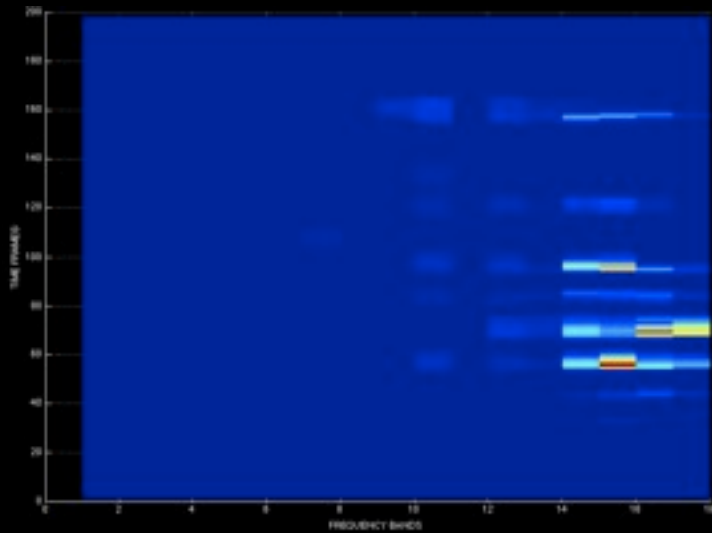
White Noise



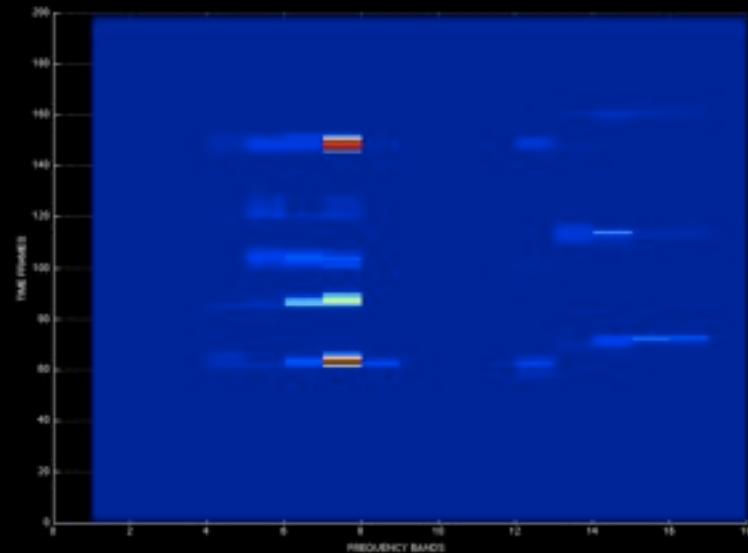
1Khz tone + double sweep



Trumpet



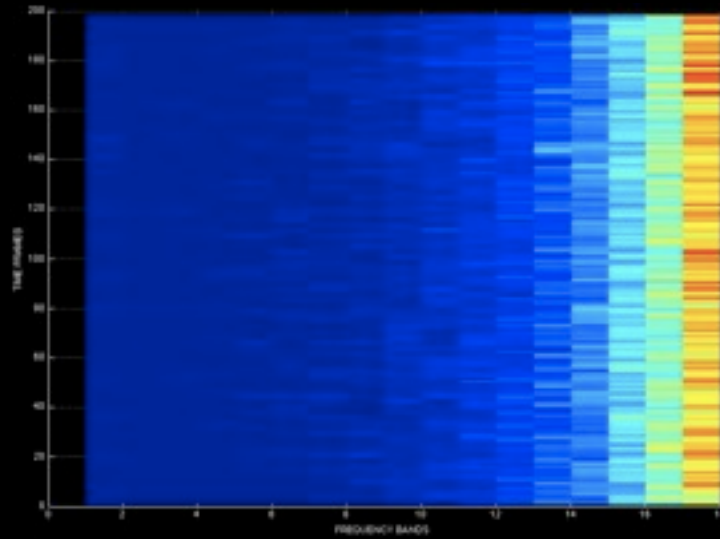
Speech



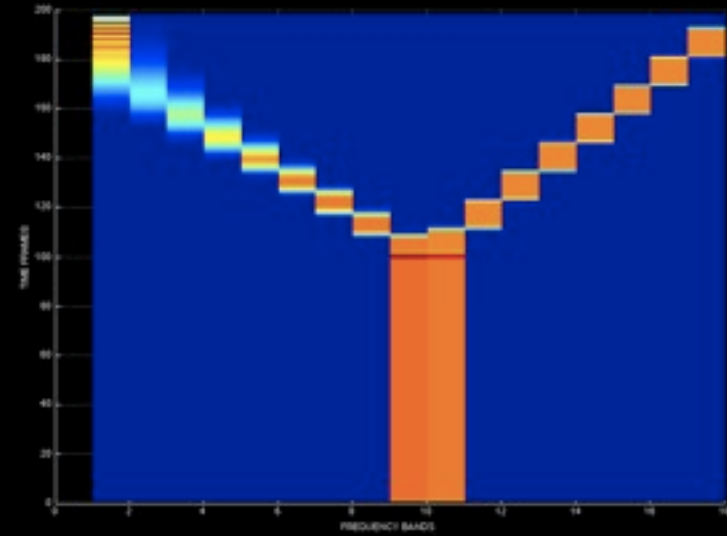
BasicSpectral Descriptors

AudioSpectrumEnvelope

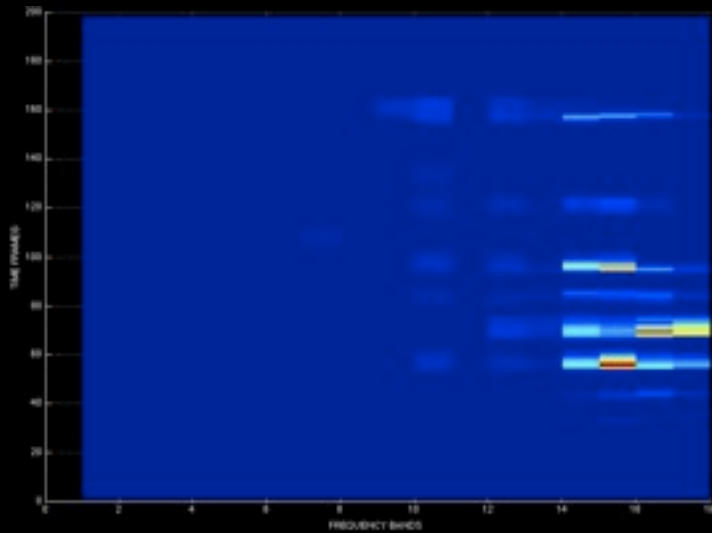
White Noise



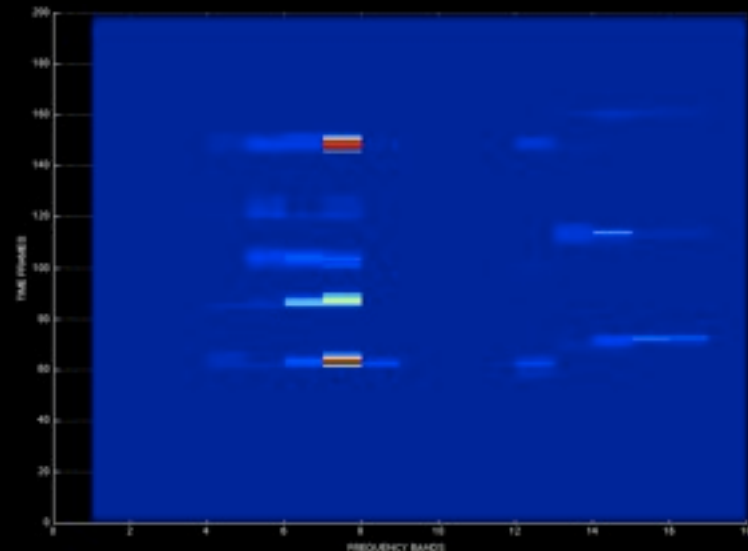
1KHz tone + double sweep



Trumpet



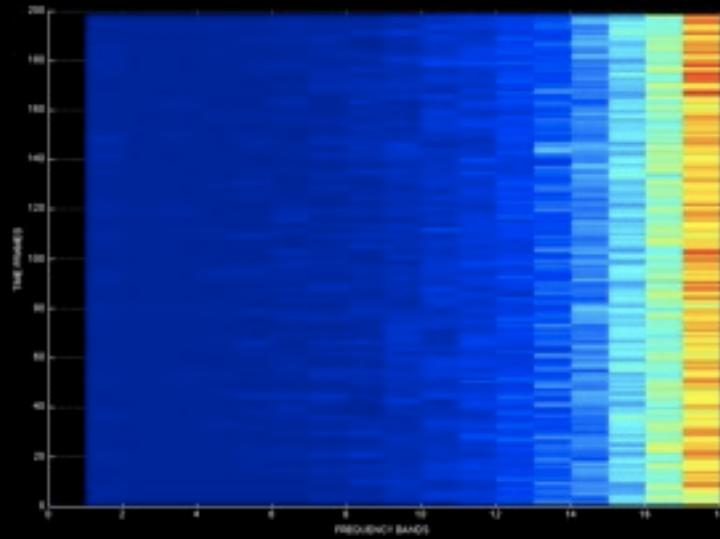
Speech



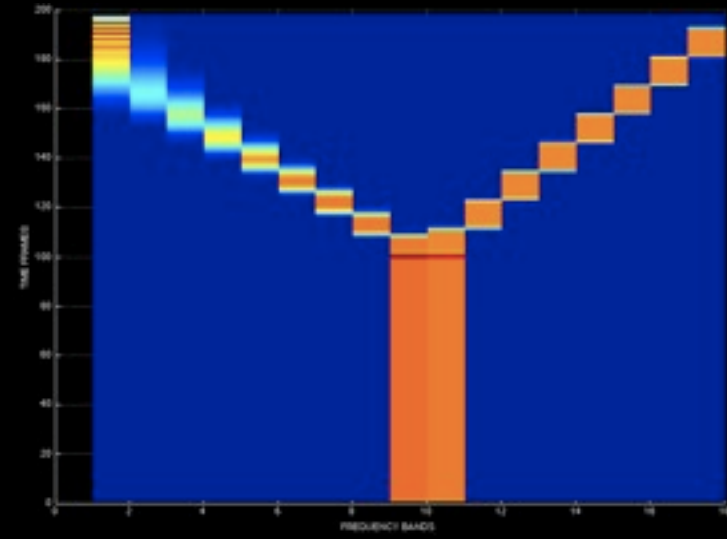
BasicSpectral Descriptors

AudioSpectrumEnvelope

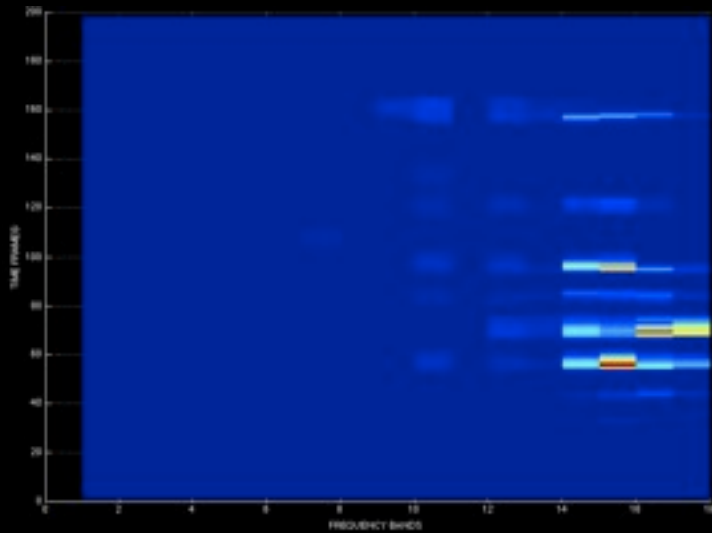
White Noise



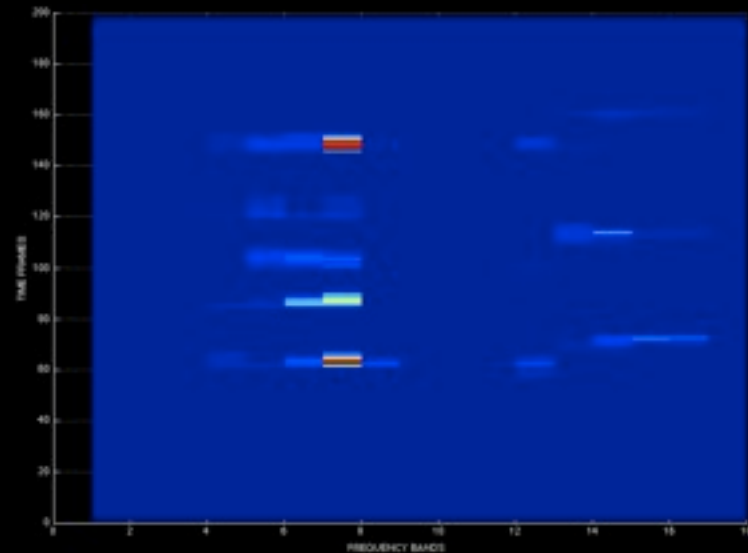
1Khz tone + double sweep



Trumpet



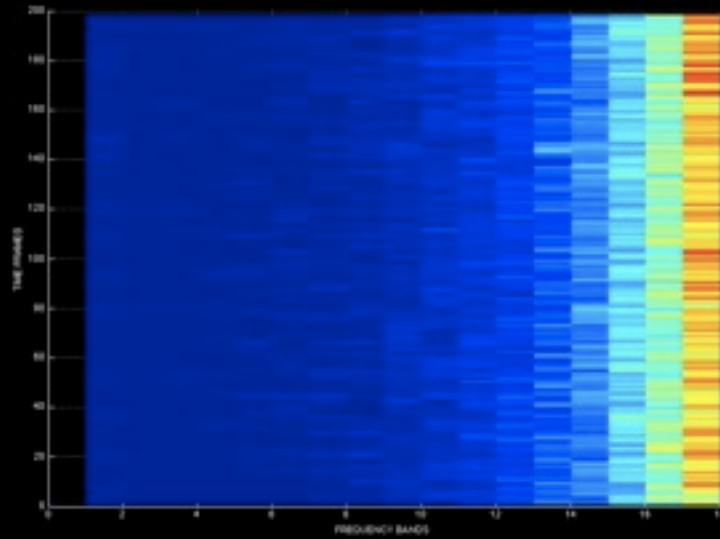
Speech



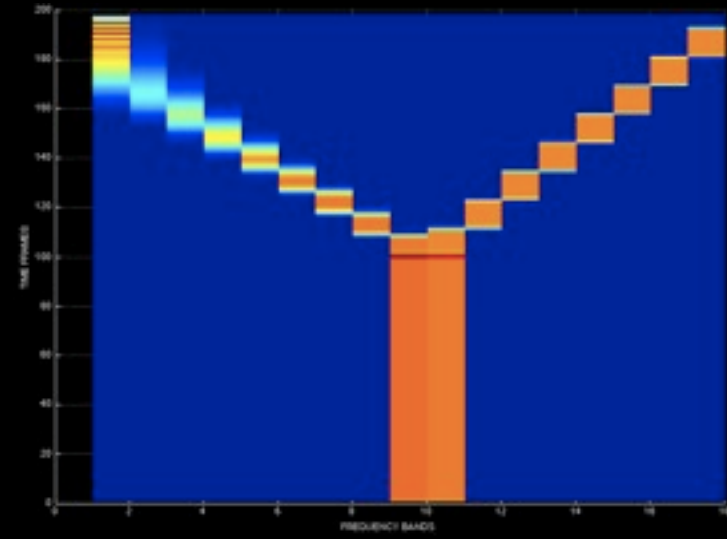
BasicSpectral Descriptors

AudioSpectrumEnvelope

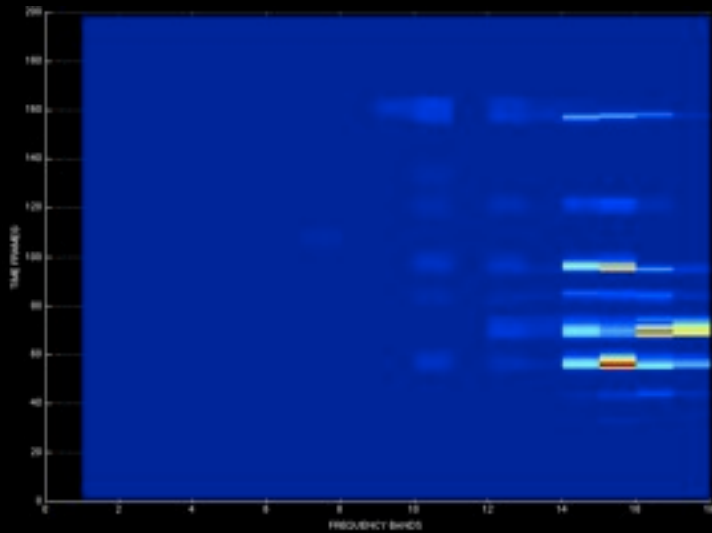
White Noise



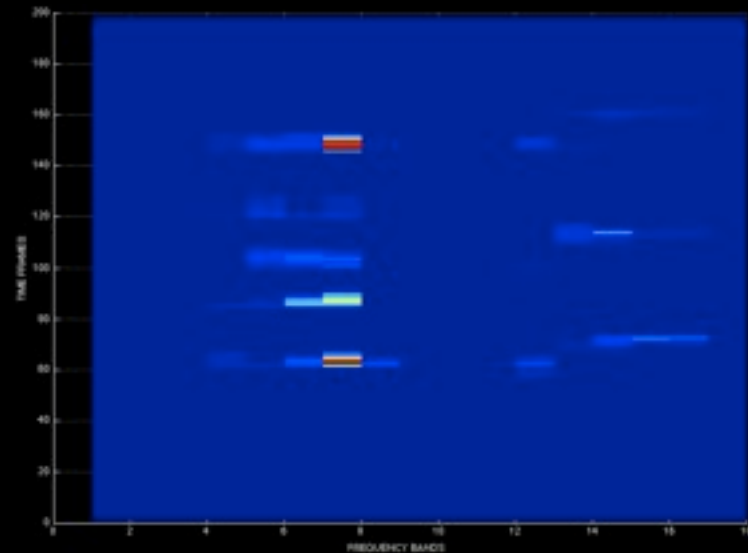
1KHz tone + double sweep



Trumpet



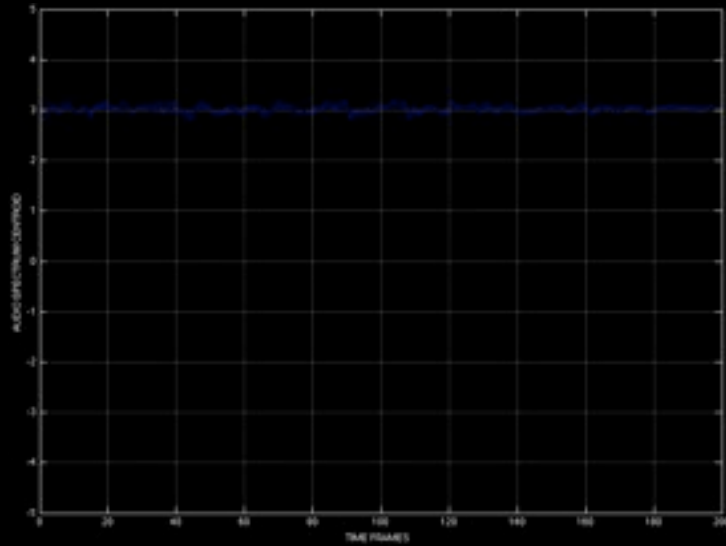
Speech



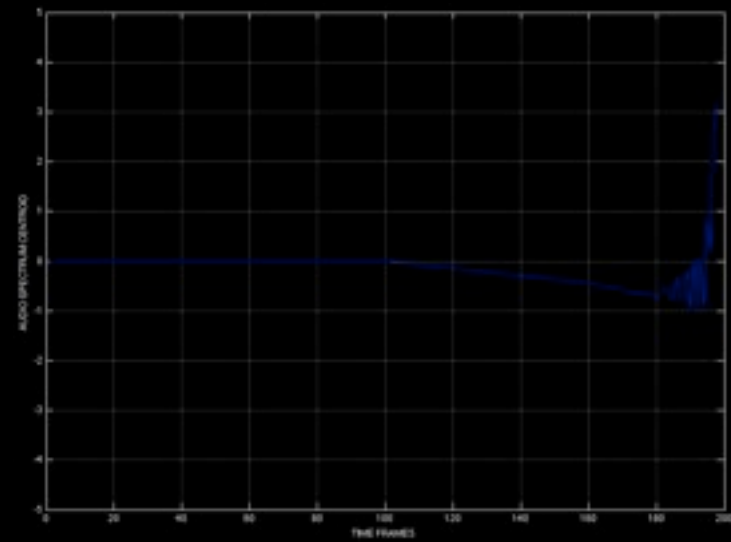
BasicSpectral Descriptors

AudioSpectrumCentroid

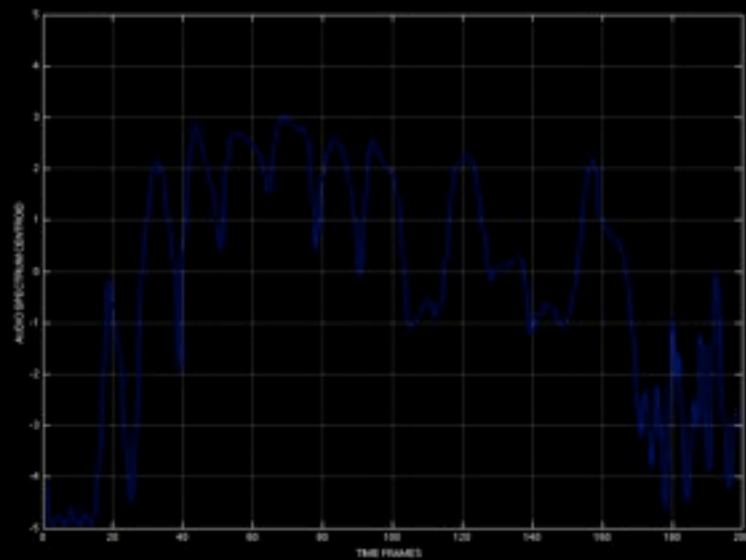
White Noise



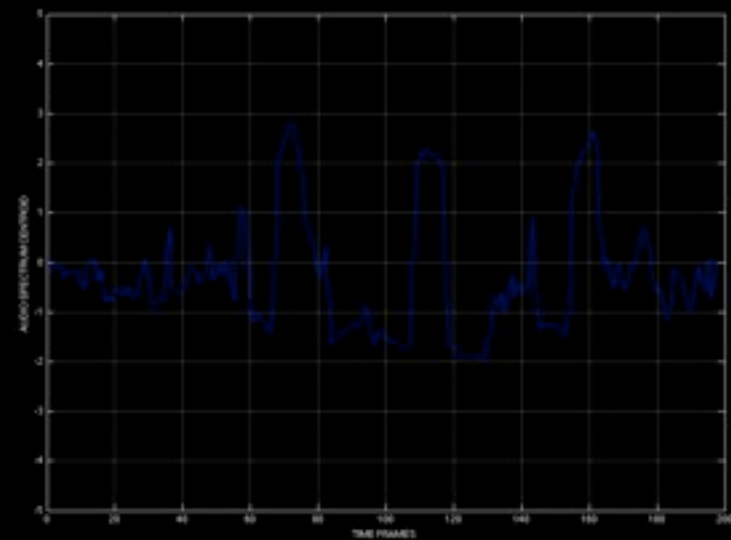
1KHz tone + double sweep



Trumpet



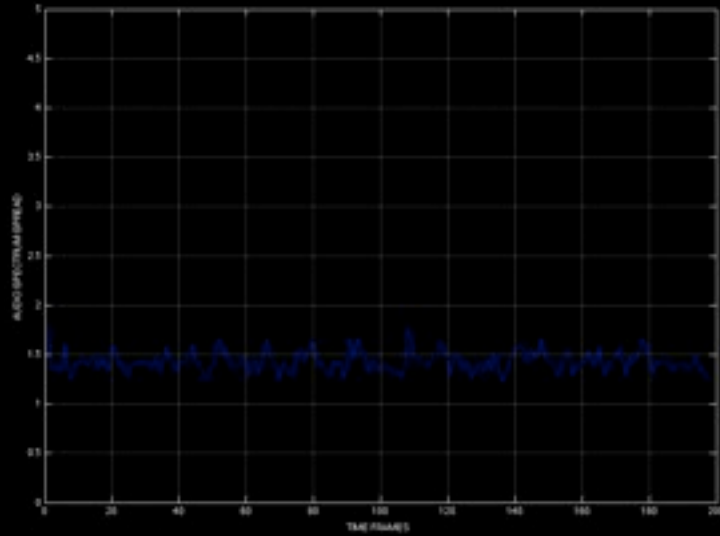
Speech



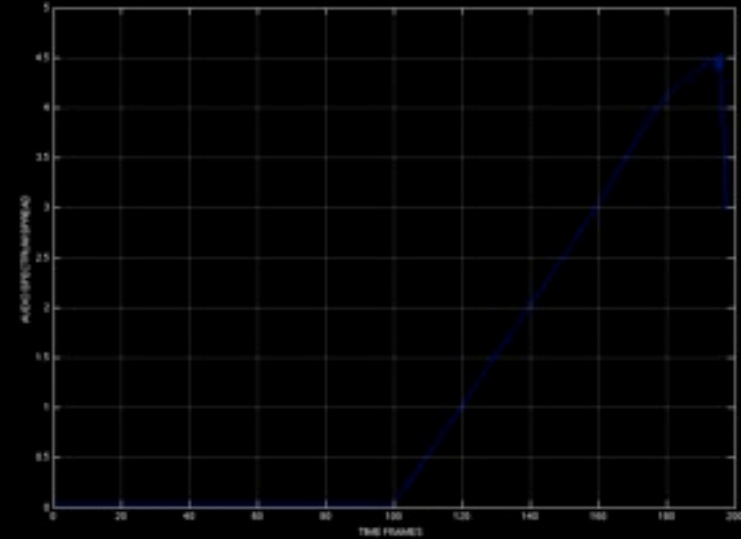
BasicSpectral Descriptors

AudioSpectrumSpread

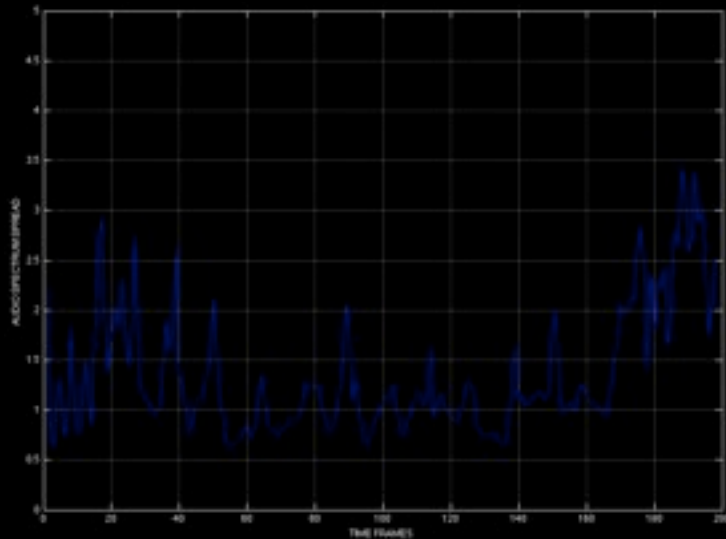
White Noise



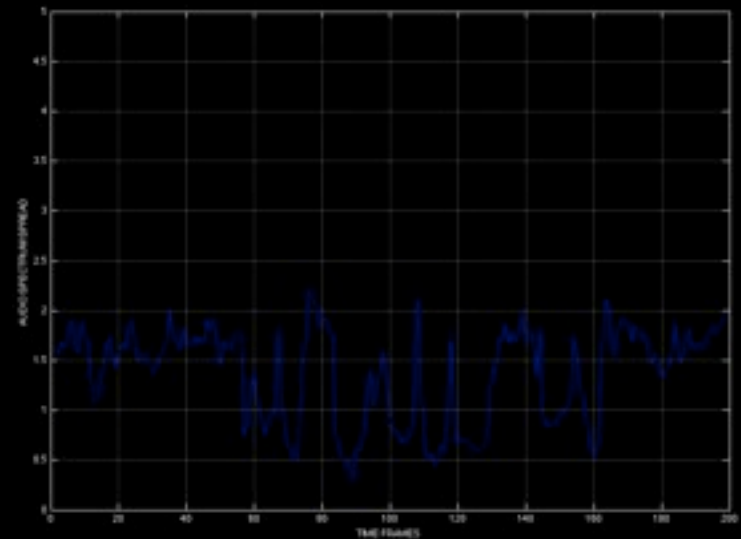
1Khz tone + double sweep



Trumpet



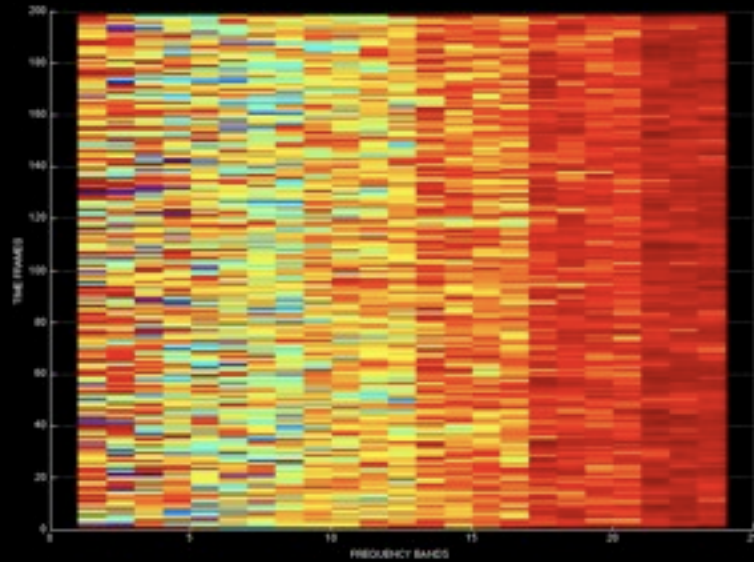
Speech



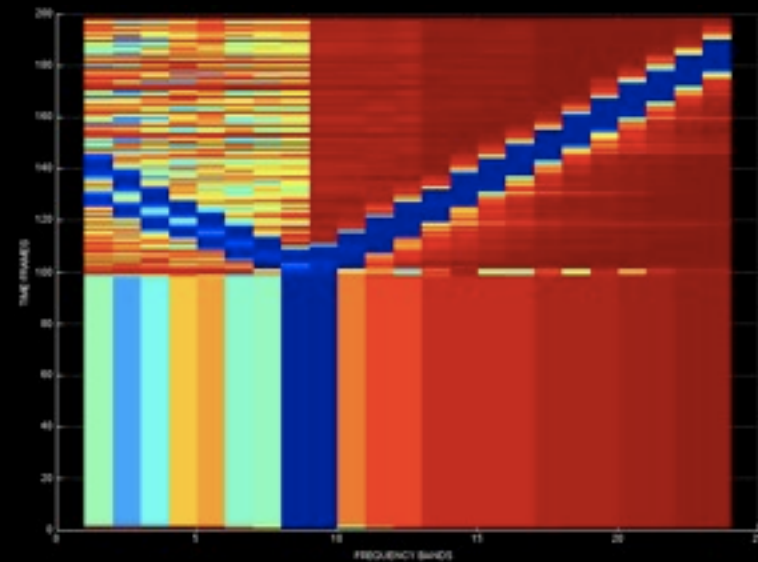
BasicSpectral Descriptors

AudioSpectrumFlatness

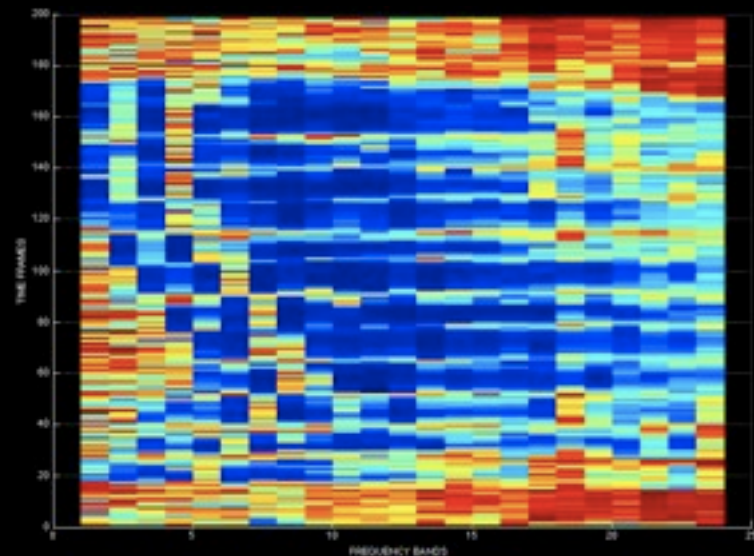
White Noise



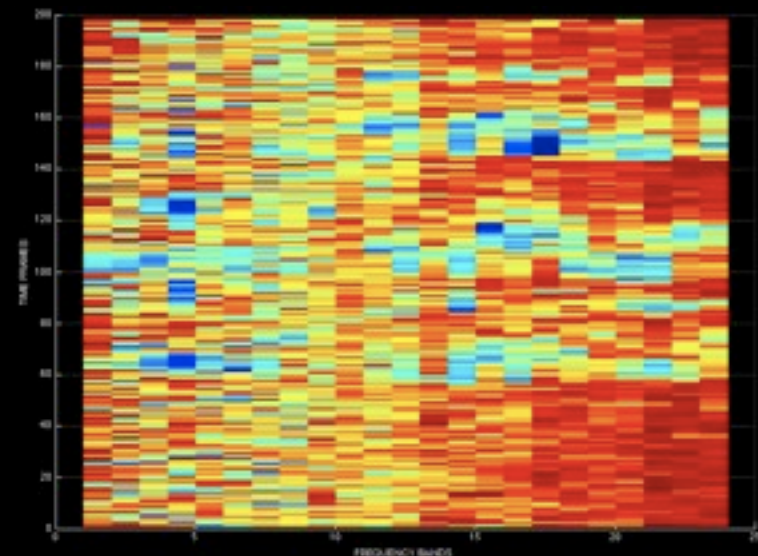
1Khz tone + double sweep



Trumpet



Speech



Some Audio Applications

Automatic Audio Segmentation

Spoken Document Retrieval

Automatic Classification for AudioFX

Highlights Extraction for Sport Using
Audio Event Detection

Sound Indexing and Browsing of
Home Video Using Spoken Annotations

Query by humming

Some Visual Applications

Automatic Video Segmentation
and Indexing

Video Personalization / TV AnyTime
Applications / SetTop Boxes

MPEG-7
Multimedia Database System (MPEG-7
MMDB)

Video Pattern Recognition

Query by example &
Semantic Retrieval

Face Analysis

Examples

www.eptascape.com : Image Understanding (surveillance, intelligent vision, smart cameras, etc.)

www.ist-live.org : Information retrieval and Filtering (quickly and efficiently searching for various types of multimedia documents of interest to the user)

www.locatetv.com : TVAnytime (helps people to discover programmes on TV, DVD and online)

mpeg7audioenc.sourceforge.net : Audio Feature Extraction (to research similarities between different classes of sounds in a collection)

MPEG-7 Cons

XML Schema has several semantic ambiguities

Provided tools and descriptions are not associated with unique semantics

Description tools are not so much flexible

It does not have characteristics for the full description of a piece (e.g.: recording location, session date, engineers, etc), like other metadata schemas (Dublin Core)

Interoperability problems due to different interpretation of semantics

Bibliography

1. Manjunath, B. S., P. Salembier, et al. (2002). Introduction to MPEG-7 : multimedia content description interface. Chichester, England ; New York, Wiley.
2. Kim, H.-G., N. Moreau, et al. (2005). MPEG-7 audio and beyond : audio content indexing and retrieval. Chichester, England ; Hoboken, NJ, Wiley.
3. ISO/MPEG N4674, Overview of the MPEG-7 Standard, v 10.0, J.M. Martínez, ed., MPEG Requirements Group, Palma de Mallorca, España, Oct. 2004.
4. Casey, M. (2003). Music Processing with MPEG-7 Low Level Audio Descriptors (DAFx03 presentation) <http://mpeg7.doc.gold.ac.uk/DAFX03.ppt.zip>
5. Day, N (2001) MPEG-7 Applications: Multimedia Content Retrieval (IMA Workshop 6B Digital Libraries: Classification, Retrieval and Visualization) <http://www.ima.umn.edu/talks/workshops/2-26-3-2.2001/day/day.htm>
6. Lai, C., I. Fujinaga, et al. (2005). Metadata for phonograph records: facilitating new forms of use and access to analog sound recordings. Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries. Denver, CO, USA,

Thanks !

Some Visual Applications

Semantical Concepts



News Anchor

News

Sports News



Commercial

Weather Forecast

Program Header



Zoom Out

Close Up

Replay

Wei, Y., S. M. Bhandarkar, et al. (2009). Client-centered multimedia content adaptation. ACM Trans. Multimedia Comput. Commun. Appl. 5(3): 1-26.