
From Data Compression to Virtual Goods Technical Perspectives for the Usage of Digital Music

Karlheinz Brandenburg

Technische Universität Ilmenau, Institut für Medientechnik

Fraunhofer Institut für Digitale Medientechnologie

October, 2007



Overview

- Digital compressed music
- Some historic remarks
- To DRM or not to DRM ?
- Scenarios of media use
 - The “home stereo” of tomorrow
- How to find my favourite music ?
 - Demo of the “Music Finder”
- Conclusions

What is going on in consumer electronics ?

- Digital TV
 - Still not everywhere
- Internet Audio
 - On all computers, 40% of people use it
 - For Christmas, mp3 players are on the top spot of childrens wish lists 😊
- Music from all around
 - There is serious competition for radio
- DVD Audio / SACD
 - Where are they ?

What is happening to the „old media“ ?

- For the first time in 50 years: the average time young people spent in front of the TV screen is decreasing
- Radio loses a substantial amount of listeners, especially among young people
- People spend more money on computer games
- Record companies and Hollywood studios complain about less business

The history of „new media“

- Radio plays concerts
 - Cinema moves theatre everywhere
 - Movie theater goes to our homes: TV
 - All my records go with me on my .mp3-Player
 - Analog TV goes as is on digital TV
- > it always takes time for new art forms to develop

Global trends: Quantity leads to paradigm changes

- The number of TV programs:
 - 1
 - 5
 - 35
 - 300
 - 3000
- How do I find my favorite program ?
- EPGs, web sites, communities, Podcasts
- Example: MusicMatch Radio: Radio or not ?

Global trends: Global reach for special interests

- Special interests find their
 - Web sites
 - Music
 - Video
 - Discussion place on the Web
- Just 100 people can establish a well organised global group
- Blogs, fora, chat rooms, cooperative computer games create virtual worlds

New equipment (this was never sold)



Historic Remarks

A true fairy tale: how it all started

Basic Research:

- **End of the 70s:**
 - Prof. Seitzer: transmit music over ISDN
 - First ideas of perceptual coding:
 - MIT Lincoln Labs
 - AT&T Bell Laboratories
- **Early 80s:**
 - In Erlangen and Duisburg Ph.D. work starts
 - First implementation based on speech coding
 - First results and papers:
 - “this won’t work”
 - “who needs this anyway ?”
- **1986: First break through (OCF)**

The team in 1987



Conquering the market: MP3 and the Internet



Marketing the Internet way:

- 1995:
 - Winplay3 Demo, Name mp3 (14.7.95)
 - Registration codes are spread
 - l3enc / l3dec not meant for consumers
 - IUMA has music using Layer-2
 - Macromedia licenses Layer-3
- 1996:
 - Students write other decoders
 - First license to Microsoft
- 1997:
 - A stolen mp3 encoder can be found widely
 - mp3.com starts

State of the Art in Audio Compression

- Everywhere: mp3
- New systems: AAC, HeAAC
 - Apple iPod
 - iTunes music shop
 - XM Radio, portable phones etc.
- High end: AAC or lossless
- “good enough”: mp3, HeAAC, WMA etc.
- Newest: surround backwards compatible
 - mp3surround, MPEG surround













What quality can be reached today ?

Define the quality to reach for first:

- **High end:**
 - don't call it „transparent“
 - best listening conditions
 - listeners need years to be trained
 - large number of samples for statistics
- **„near CD“ - quality:**
 - defined as „good enough“, no formal definition
 - much more important for practical purposes
 - example: mp3 at 128 kbit/s for stereo

Demo: Can you hear it (Version 4, 2000) ?

Each "?" corresponds to either
O (Original, 1536 kbit/s for two channels) or
C (Coded, 48 kbp/s for two channels)
(HeAAC, demo provided by Coding Technologies)

Trumpet solo		O		?		?		?
Speech		O		?		?		?
Abba		O		?		?		?

To DRM or not to DRM 😊

- Ideas for protected media content have been around since the mid 90s
- Early systems lacked sophistication
- SDMI
 - Initiated by the music industry
 - Work on requirements for DRM systems
 - Success or failure ?
- Current systems:
 - Windows DRM
 - Fairplay
 - OMA 1.0, 2.0
 - Music distribution without DRM

Requirements for successful applications:

- Ease of use:
 - work everywhere, on all equipment
= **interoperability**
- Add value:
 - digital is more than just the same product
add-on text / pictures / video / special sale etc.
- Make illegal use more difficult
 - enough copy protection to make piracy a
conscious decision

Basic Building Blocks and Concepts

- Secure envelope
- Scrambling / encryption
- Robust embedding of data (watermarking)
- Authentication
 - User authentication
 - “Bind to”, authenticate content
- Secure Authenticated Channel
- Super Distribution

User expectations

- For personal use, music (on Redbook CD) can be
 - moved (from home stereo to the car)
 - copied (for example to tape)
 - shared between family members
 - lent to my best friend
- All compliant CDs used to play on all equipment with the same physical form factor (interoperability)
- Current technology (mp3) allows all of the above
- Secure Electronic Music Distribution (EMD) has to meet these expectations



Current Status

- **DRM systems**
 - Fairplay (Apple): limited usability, only on Apple hardware
 - Windows Media DRM: on many devices, together a small market share
 - Other: nearly non-existent
- **Non-DRM distribution**
 - EMusic and others: # 2 in U.S. market
 - Apple with AAC non-encrypted (EMI)
 - Amazon in beta test (EMI, Universal)
 - Super distribution: Potato et. al.

Szenarios for media use

old ideas, now reality

Where are my multimedia data ?

- Physically stored in my home (CD, DVD)
- On my home server
- On my PDA or mobile phone
- Detached on the net
- Just recreated from metadata whenever I need them

Examples: The future home stereo setup

(really ?? 😊)



Prof. Karlheinz Brandenburg, bdg@iis.fraunhofer.de

The home stereo of tommorow:

- **Remembers my music:**
 - Each CD gets ripped just when I play it
- **Helps to search for my favorite music**
 - “query by humming”
 - Look for certain genres
 - Searches the Internet
 - Where can I listen to / buy this music ?
- **Helps to assemble my favorite music playlist**
 - What is similar, but new ?
 - “please play my favorite music for this time of the day”

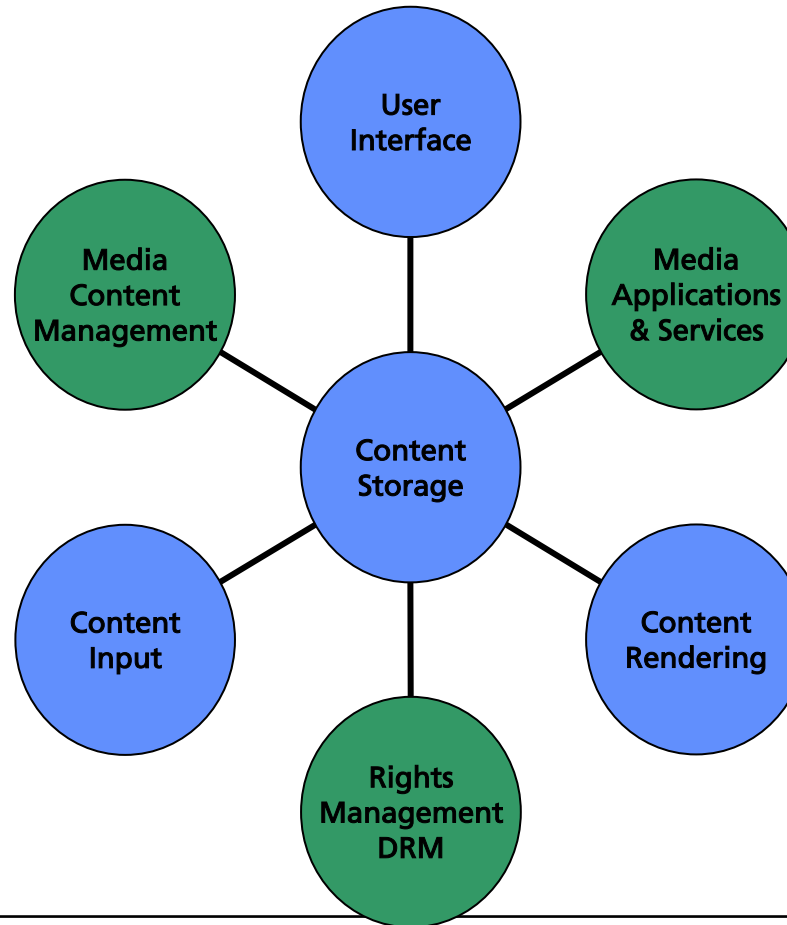


Music just like in the concert hall:

- Two channel stereo can sound good, surround can be better, but:
 - Only if equipment and room are of the highest quality
 - Only for very few or single listeners in the sweet spot
- IOSONO brings the concert hall to my home:
 - Improve room acoustics
 - Wave Field Synthesis (acoustical holography)
 - Invisible flat panel loudspeakers
 - “we are now in row 7, seat 18 of the Musikvereinssaals, Vienna”
 - The singer can virtually stand **in** the room

The future of consumer electronics

Components tomorrow



Rich Media Management

Semantic
Media
Analysis

Multimedia
Search and
Recommen-
dation

Services and
Applications

Rich Media Management



Multimodal metadata retrieval

- Context-based metadata creation (User, Usage)
- Content-based metadata creation (Rhythm, Genre, Mood, Harmony, ..)

Storage of feature vectors and textual metadata

Machine learning (Clustering, context adaptation,..)

Supports content-based/semantic queries

Allows for a lot of new services & applications

How to gather metadata ?

Sources

- Expert knowledge, deploying the public, metadata services (CDDDB, FreeDB)
- Context-derived, collaborative filtering

Content-derived

- Automated extraction of relevant features from the content data
- Find clusters in the high-dimensional feature space that represent semantic properties or similarities
- Efficient algorithms (→ »embedded«)

Multimedia Information Retrieval – available technologies

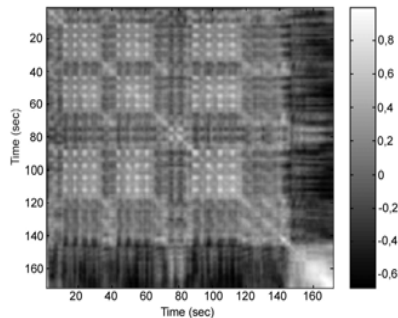
Recognition, Similarity

- AV-recognition (AudioID, PhotoID, VideoID)
- AV-similarity (Soundslike, PhotoID)
- Song-segmentation

Semantic Analysis

- Genre classification (GenreID), Speech-/Music discrimination
- Dominant melody transcription, Query by Humming (QbH)
- Bass transcription
- Tempo-, beat determination
- Drum pattern transcription
- Harmony transcription
- Face-, object detection & identification
- (Lead-) instrument recognition
- Singer detection and characterization

MIR-Technologies – Music Segmentation



- Inner-song similarities / dissimilarities based on a set of low-level audio features
- Measure of innovation
- Segment classification (verse, chorus)

Intelligent navigation within songs

- Retrieval of most significant part of a song (“audio-thumbnail”)
- Basis for a deeper semantic analysis of the segments (lead instrument, vocals, etc.)



MIR-Technologies – Rhythm

The Pulse Of Time...

- Important for beat-synchronized playlists
- Crucial for recommendation and genre determination

Retrievable Metadata

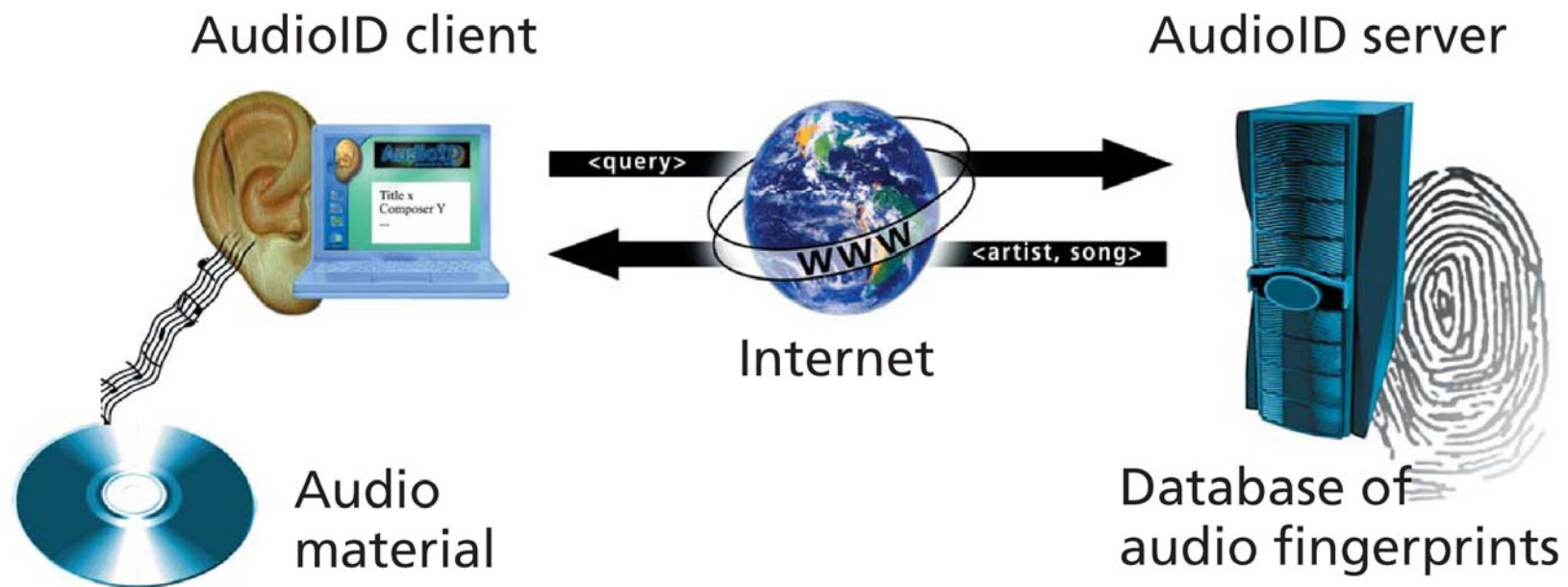


- Tempo (bpm)
- Bar measure
- Beat positions
- Tatum (grid)
-
- Transcribed rhythmic patterns
- → applicable for rhythmic similarity calculation

AudiID



Audiold – the basic idea



Query by Humming at Saturn-Markt, München



Music-controlled animations or -games



- Avatar:
 - Content-derived information retrieval from music, e.g. beat-, intensity, mood
 - Control animation, content-based (beat-synchronous and emotion-preserving)
 - Visualization of 3-dimensional virtual figures in realtime (OpenGL, Direct3D)

AudioID – Sonos: Modeling Perceptual Music Similarities

Holistic

- Each song represents just one single class
- Use distance metric from a holistic feature space (e.g. euclidian distance) → too simple!

vs.

Modular
classification
approach

- Each song exhibits certain semantic properties (aspects)
- Measure the aspect qualities and build an aspect-profile
- → Translate audio content into the semantic space

Select aspects according to their perceptual relevancy

subjective	objective
<ul style="list-style-type: none">• mood: happy vs. sad• aggressive vs. calmative• slow vs. fast	<ul style="list-style-type: none">• genre affiliation• drums vs. no drums• singer vs. no singer

Prof. Karlheinz Brandenburg, bdg@iis.fraunhofer.de

Megatrends

- **More is more than just more** 😊
 - More TV programs
 - More bandwidth
 - More accessibility
 - More individualized services

- **We will get to**
 - Virtual worlds
 - New methods of communication
 - New types of broadcasting
 - E.g. blogs

Conclusions

- The age of digital media has just started
- Signal processing brings us a lot of new possibilities for future home multimedia devices
- Searching and recommendation are crucial applications for tomorrow's media architectures
- Immersive media are one future killer applications, for audio we can already do it

To summarize:

- There are plenty of applications ahead for Audio and Music Computing for Multimedia