“IS IT TRUE ?” AUDIO RECOGNITION AND TAMPERING DETECTION AS MEANS FOR AUTHENTICATING COMMUNICATIONS

Karlheinz Brandenburg, Christian Dittmar

Fraunhofer Institute for Digital Media Technology IDMT
Technische Universität Ilmenau

Virtual Goods, Koblenz, September 2013

ILMENAU UNIVERSITY OF TECHNOLOGY

Fraunhofer IDMT

OVERVIEW

- Current challenges
  - Security
  - Authentication
- Technical means to proof tampering
  - Robust hashes (audio identification)
  - Recognize melodies
  - Identify editing of audio recordings
- Conclusions

BACKGROUND

- Digital audio everywhere
  - 20 million tracks of music
  - Every phone conversation
  - Billions of devices record / play back audio of all kinds

- But is it true ?
  - We all know that pictures can be modified
  - Audio has the same possibilities
    - Delete parts to change meaning
    - Re-use the artistic work of others

SOME SCENARIOS (SOME OF THEM HAVE PROBABLY NOT YET HAPPENED)

- Plagiarism:
  - A short piece of music is re-used as is
  - A melody is used in a different context
- Editing the original source
  - Some words are deleted from a sentence to change the meaning
  - An original source is used as a material to create a new sentence
- Resynthesizing speech:
  - Analysis of speech specifics, then synthesis with new meaning

SECURITY / AUTHENTICATION

- Use cryptography to secure the transmission
  - Not the topic of this talk
- Use hash functions or similar to authenticate the source of the transmission
  - Probably the only solution against the most advanced attacks
  - Again not the topic of this talk, but:
  - Can we produce robust hash functions for audio ?
    - Yes, see the next slides

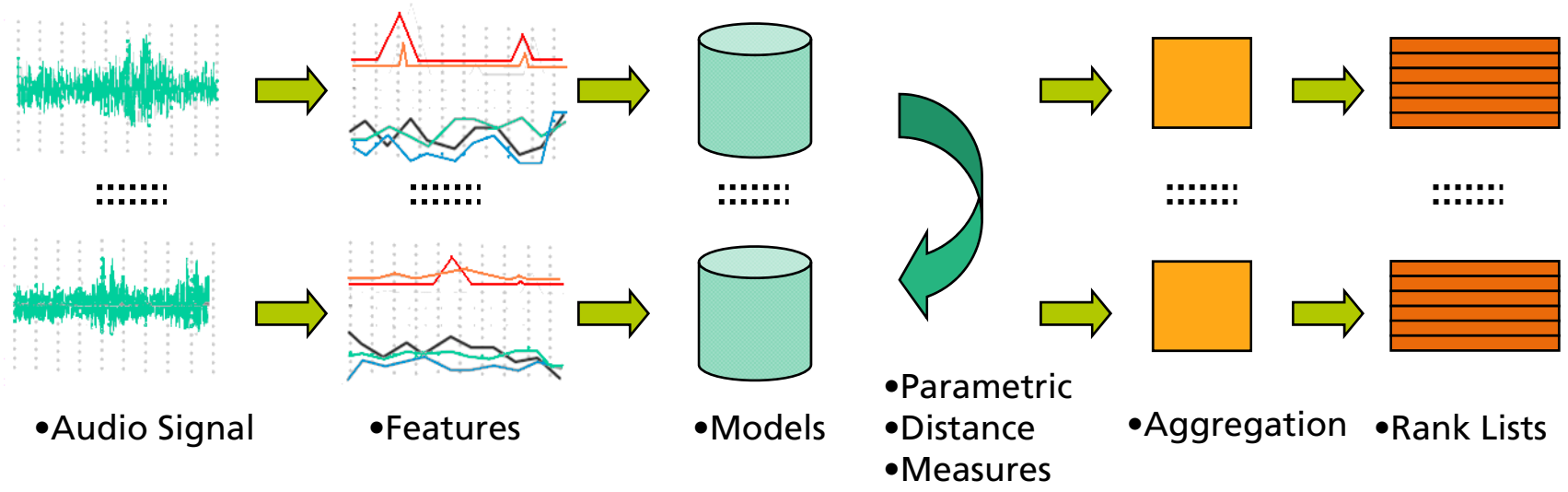
AUDIO IDENTIFICATION, AUDIO SIMILARITY: BASICS

- Audio identification is used in Apps like Shazam, SoundHound
  - Technically mature field
  - Use of machine learning
  - Accuracies approach 100 % even in difficult conditions

- Audio recognition
  - Much more difficult
  - We can recognize melodies etc.
  - Examples follow

AUDIO SIGNAL ANALYSIS: BASIC TASKS

- Audio Similarity Search: Query-by-Example → Common approach
  - Audio model given by distribution of low-level audio features
  - Distance between models → indicates similarity

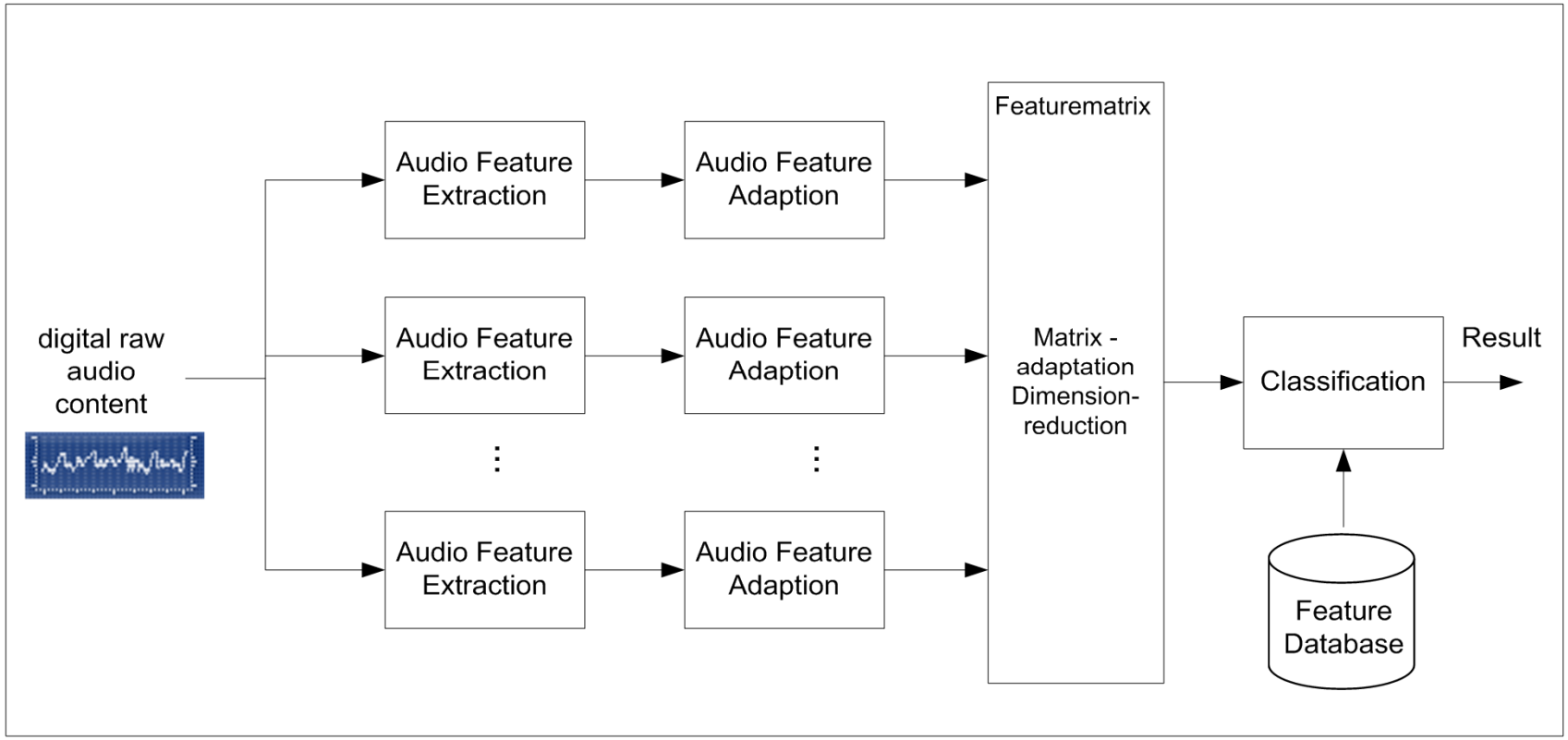


ILMENAU UNIVERSITY OF TECHNOLOGY

Fraunhofer IDMT

AUDIO SIGNAL ANALYSIS: BASIC TASKS

- Audio Pattern Recognition → Machine Learning



ILMENAU UNIVERSITY OF TECHNOLOGY

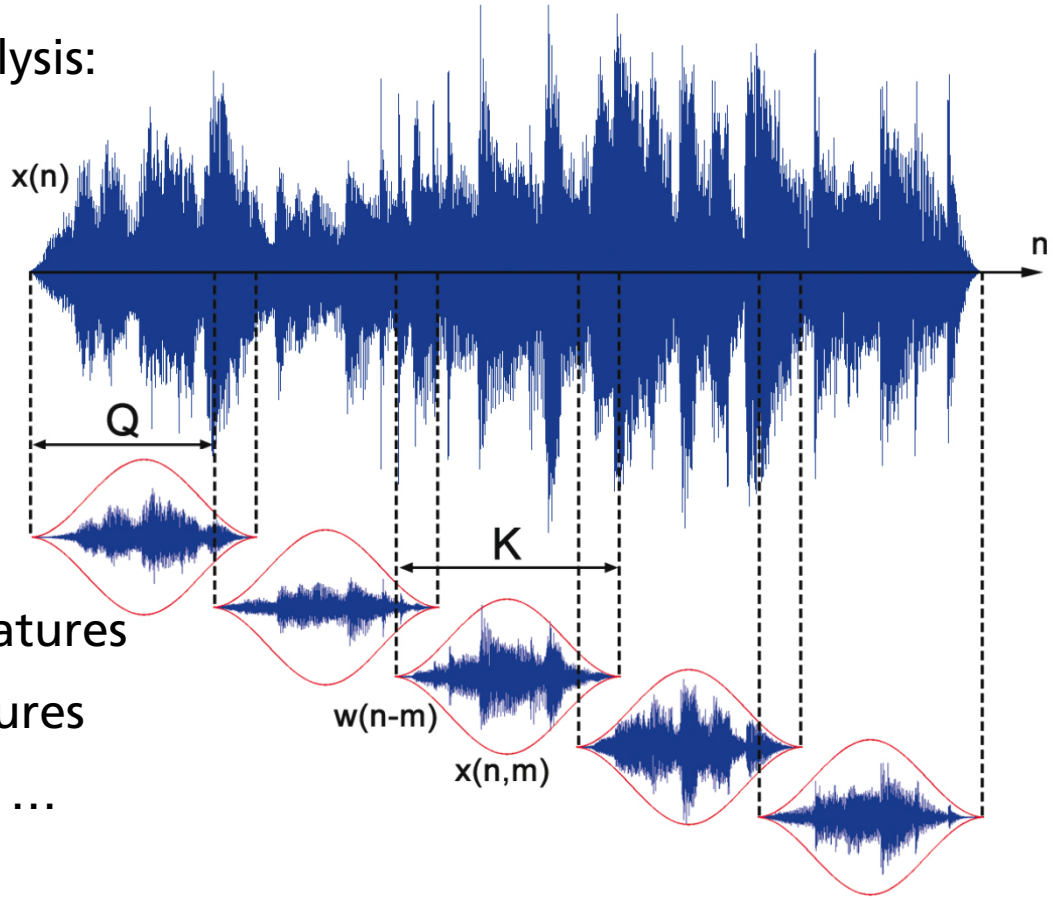
Fraunhofer IDMT

AUDIO FEATURES

- Low level
  - Signal derived, simple math
  - Sufficient for certain applications
  - Building blocks for more complex tasks
- Mid level
  - May already have semantic meaning
  - Combined or derived from low level features
- High level
  - Could be called “output parameters”
  - Can be understood by a human listener

AUDIO FEATURES: LOW-LEVEL FEATURE EXTRACTION

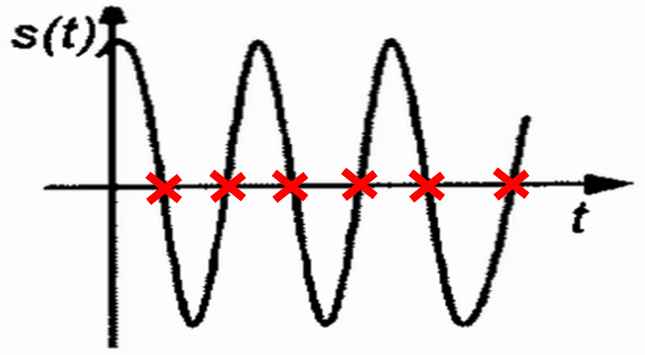
- Principle of short-term analysis:
  - Q: Hop-size
  - K: Window-/Block-size
  - w: Window-function
  - x: Signal frame

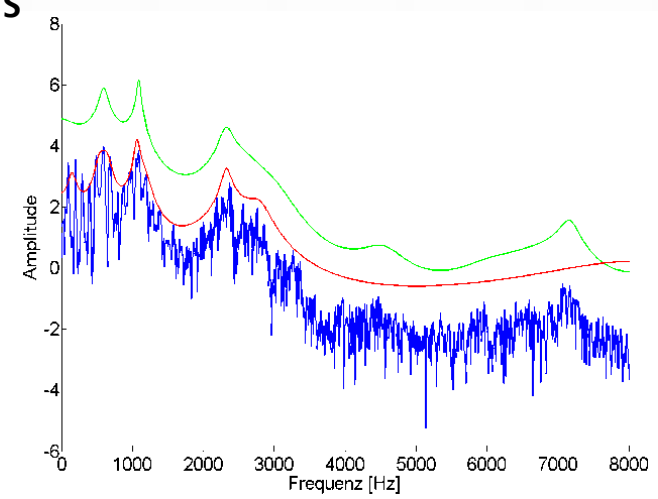


- In each analysis frame:
  - Time signal based LL features
  - Spectrum based LL features
  - Cepstrum based, others …

AUDIO FEATURES: LOW-LEVEL FEATURE EXTRACTION

- Time signal based LL features:
  - ZCR (Zero Crossing Rate): number of sign changes of the audio waveform per time frame → can be used to distinguish between low-pitched and high-pitched sounds, less suited for mixtures of multiple sounds
  - LPC (Linear Prediction Coefficients): compute filter coefficients, whose impulse response is as close to the spectral envelopes of the input signal as possible → originally used for speech coding



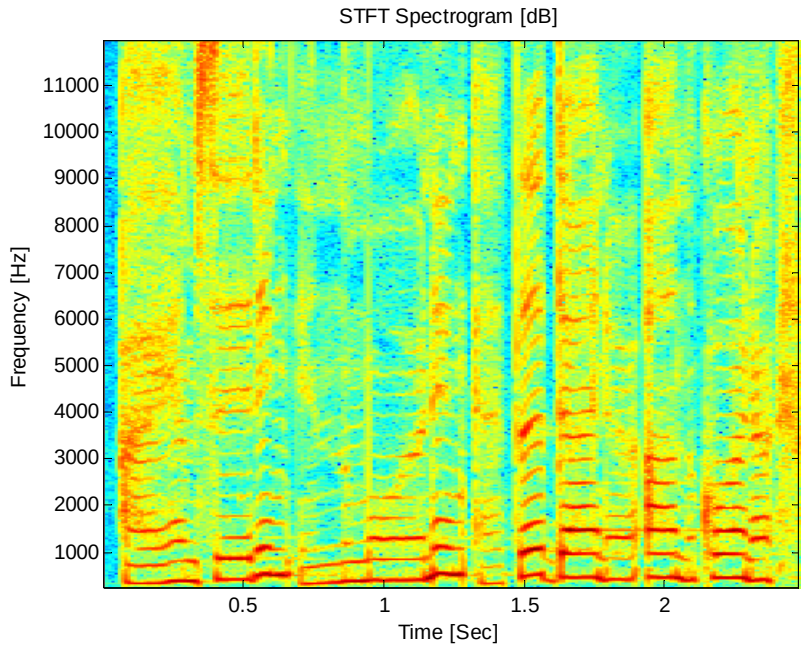


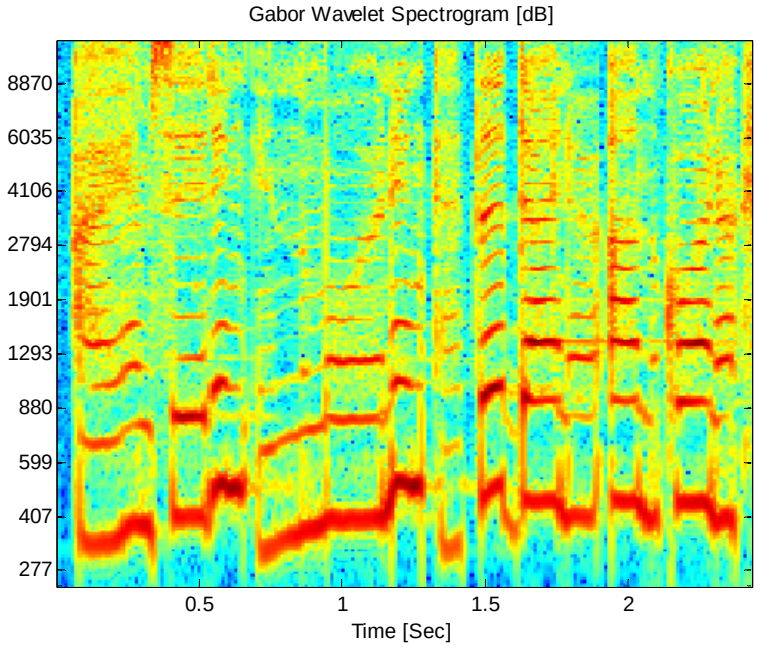
ILMENAU UNIVERSITY OF TECHNOLOGY

Fraunhofer IDMT

AUDIO FEATURES: LOW-LEVEL FEATURE EXTRACTION

- Spectrum-based features:
  - Spectrogram → Duality between time and frequency resolution
  - Linear vs. Logarithmic frequency axis





ILMENAU UNIVERSITY OF TECHNOLOGY

Fraunhofer IDMT

MUSIC PLAGIARISM ANALYSIS: MOTIVATION

- Plagiarism is know since ancient times:

  The word „plagiarius“ was used for somebody kidnapping poems.

- In legal terms, different types of music plagiarism are discerned:
  - Unconscious plagiarism → The Chiffons vs. George Harrison example
  - Parallel creation → two authors create a work independently
  - Adaption → editing extensive enough to create new work
  - Free usage → original material must not be recognizable in derived one

MUSIC PLAGIARISM ANALYSIS: MOTIVATION

- Music Plagiarism:
  - Melody sequences
  - Rhythm patterns
  - Chord sequences

  } Similarities on a semantic level

- Sampling Plagiarism:
  - Re-use of existing recordings into a new work
  - Timbre qualities → Similarity on a signal level

- There are web-communities that search & document such cases

  (www.whosampled.com; www.the-breaks.com; www.secondhandsongs.com)

MUSIC PLAGIARISM ANALYSIS: MELODY PLAGIARISM

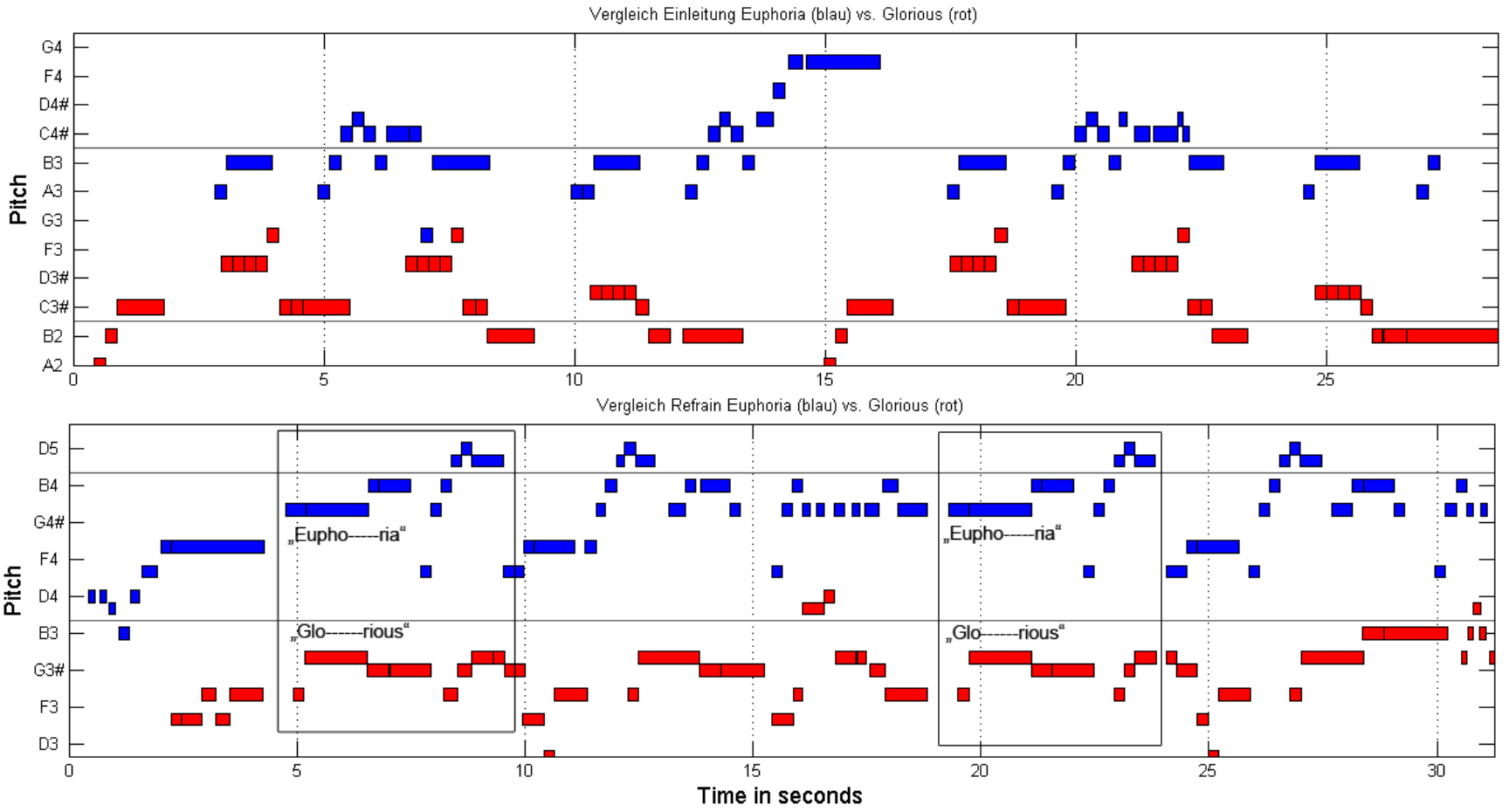
- Allegations of music plagiarism against the German entry to the European Song Contest

- Frontpage in biggest German newspaper Bild-Zeitung on 17.02.2013
  - Based on expertise by phonetician from University Kiel
  - Public broadcaster NDR commisioned musicologist expertise by Matthias Pogoda → result published 25.02.2013

- Sample „Loreen - Euphoria“
  - Tempo 131 BPM, Key F#-Minor

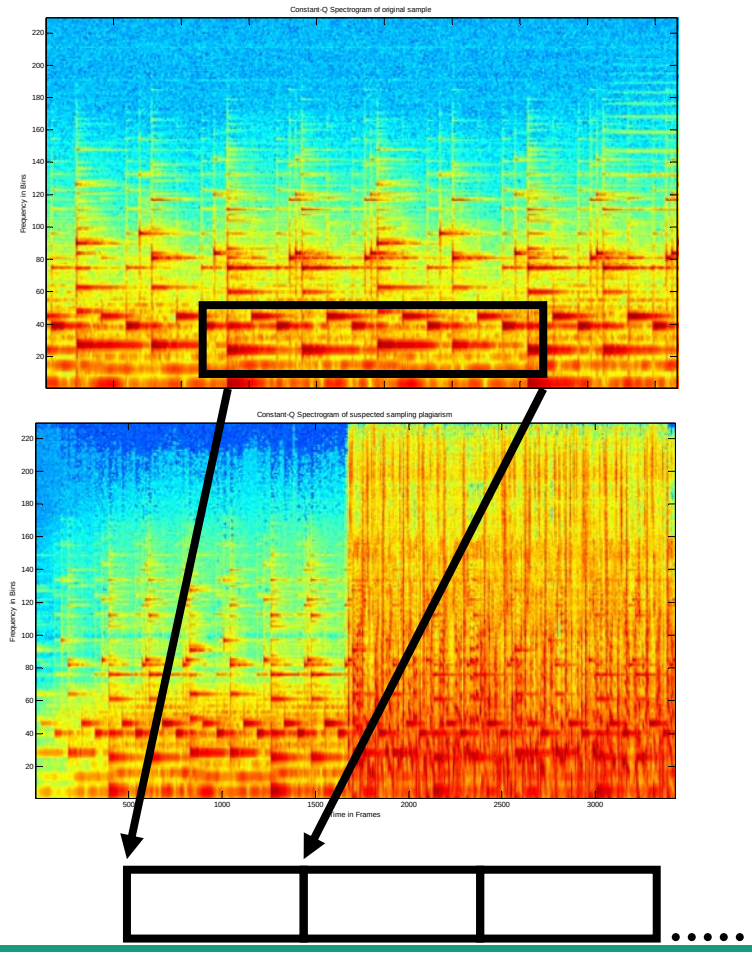
- Sample „Cascada - Glorious“
  - Tempo 128 BPM, Key G-Minor

MUSIC PLAGIARISM ANALYSIS: MELODY PLAGIARISM



SAMPLING PLAGIARISM

- Known data:
  - Original music excerpt
  - Suspected sampling plagiarism
- Edit operations:
  - Cropping
  - Looping
  - Time-stretching
  - Pitch-shifting
  - Mixing of new instruments



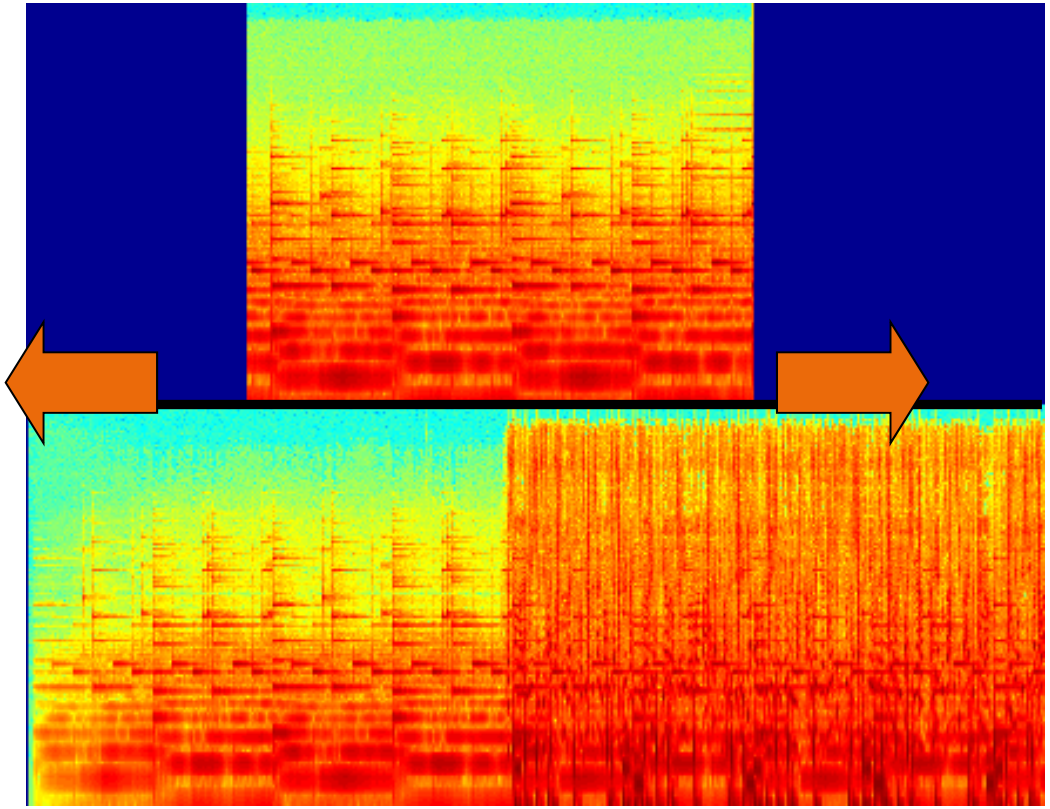
ILMENAU UNIVERSITY OF TECHNOLOGY

Fraunhofer IDMT

SAMPLING PLAGIARISM: BRUTE FORCE APPROACH

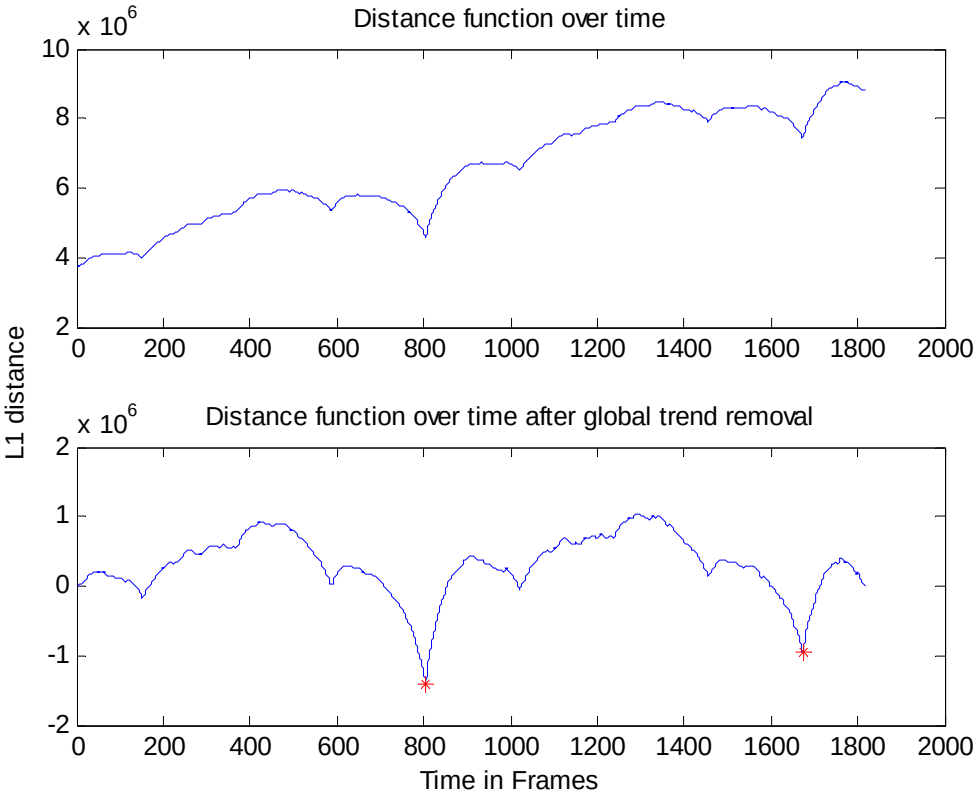
- Shift original along suspected plagiarism → find best match

- Distance measures: L1 distance, L2 distance, Correlation …



SAMPLING PLAGIARISM: BRUTE FORCE APPROACH

- Restrict search space by preliminary beat estimation
- only test timestretching factors at reasonable multiples of the beat
- only compare frame by frame around beat positions



ILMENAU UNIVERSITY OF TECHNOLOGY

Fraunhofer IDMT

DETECTING EDITING OR OTHER TAMPERING

- Watermarking:
  - Insert inaudible signals into the music / speech
  - Tradeoff between
    - Bitrate
    - Robustness
    - Inaudibility of the watermark
  - Can survive some modification of the signal (even transmission from loudspeaker to microphone) or can be fragile on purpose
  - Often does not survive heavier modifications
  - Clearly a forensic tool, it is often not known that a watermark has been applied

DETECTING EDITING OR OTHER TAMPERING

- Digital Signatures
  - Not part of the signal: may be deleted
  - Additional data necessary
  - Can implement a “bind identity to the content”
  - Can easily be stripped from the main data

- Tampering detection without any additional signal:
  - Find discontinuities in the speech signal
  - E.g. in the phase of Electric Network Frequency (ENF) signals

PHASE ANALYSIS

- Idea:
  - Modifications cause changes in the ENF phase
  - Using this changes to detect tamperings
  - Works without any reference data

- Approach
  - Extraction phase from ENF
  - Detection discontinuities
  - Segmentation of recording

WANT TO LEARN MORE?

Visit WASP workshop this Friday 20.09.2013 (RoomF 413)

- 11:00 - 12:30 Session 2 / 4

  Sebastian Mann: Combining ENF Phase Discontinuity Checking and Temporal Pattern Matching for Audio Tampering Detection

- 13:30 – 14:45 Session 3 / 4 (Posters)

  Christian Dittmar: Estimating MP3PRO Encoder Parameters From Decoded Audio

CONCLUSIONS:

- More and more people are concerned about privacy and security:
  - We need to do more about these topics
  - We do have technical means to help
- In the audio world:
  There are several methods to help against unwanted tampering:
  - Identification of plagiarism
  - Identification of changes to a signal
- Authentication is a topic which deserves more attention