Analysis of Audio-visual Content

Tools to bridge the semantic gaps

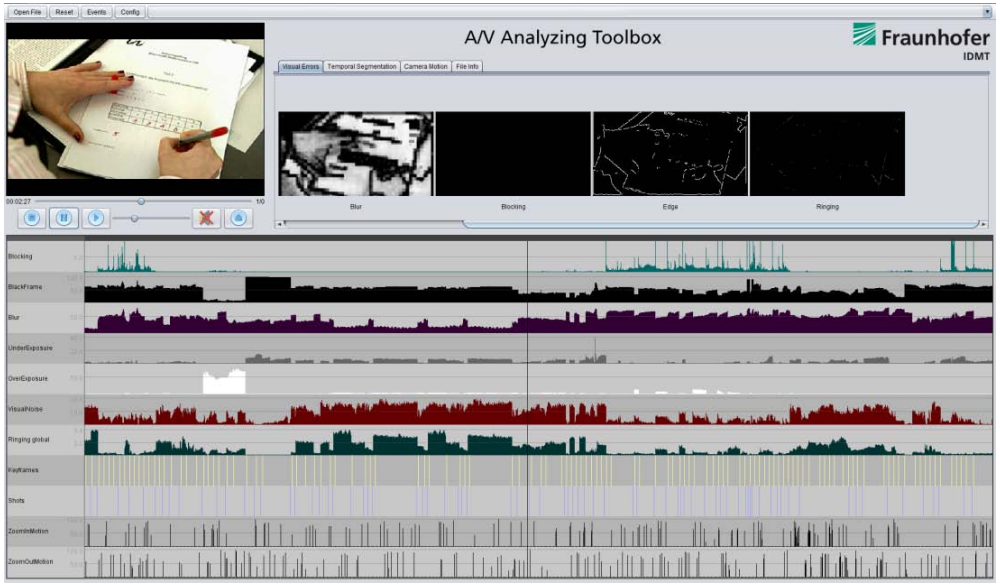
Fraunhofer Institute for Digital Media Technologie IDMT
Metadata Department

Uwe Kühhirt

Head of Audio-visual Systems

EMWRT IX, Amsterdam

13. September 2013



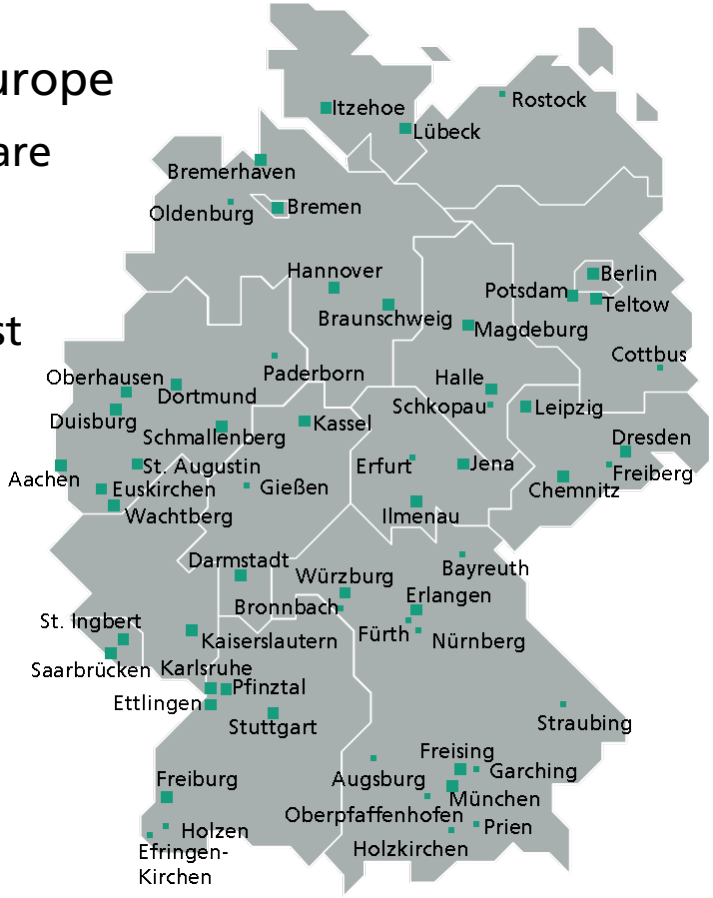
Fraunhofer
IDMT

FRAUNHOFER-GESELLSCHAFT
Overview

- Largest applied research organization in Europe
  - More than 80 research facilities, of which 66 are Fraunhofer institutes
  - 22.000 employees at more than 40 locations
  - Branches in Europe, USA, Asia and Middle East
- Non-profit:

"applied research of direct utility to private and public enterprise and of wide benefit to society"

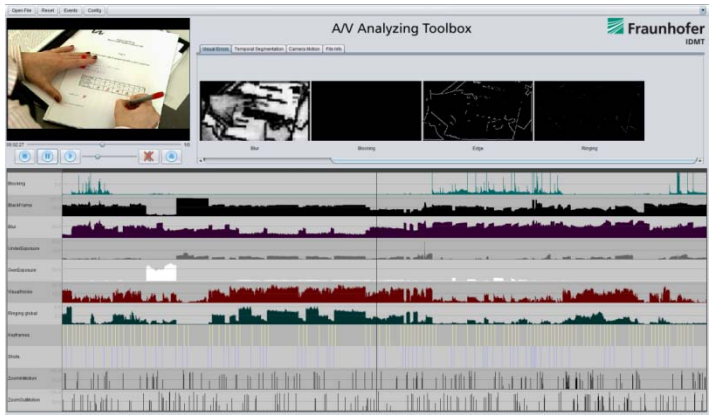
- Financial revenues
  - 30% German federal government and states (basic funding)
  - 70% industry and public research projects



Fraunhofer IDMT

Business Unit
Media Management & Delivery – M2D

- Research groups
  - Audio-visual Systems
  - Semantic Music Technologies
  - Media Distribution & Security



- Working fields
  - Analysis of audio-visual content with digital signal processing
  - Security and privacy in the media context

→ *Development of Technologies: methods, components, and systems*

→ *Evaluation & Testing of Technologies*

→ http://www.idmt.fraunhofer.de/m2d

Fraunhofer IDMT

Business Unit
Media Management & Delivery – M2D

- Topics
  - Automatic content annotation, metadata enrichment
  - Content-based recommendation
  - Robust content identification
  - Content provenance and authenticity
  - Copyright-awareness, security and privacy
  - Metadata modeling and interoperability
  - Media aggregation and distribution

- Tools to bridge the semantic gaps!

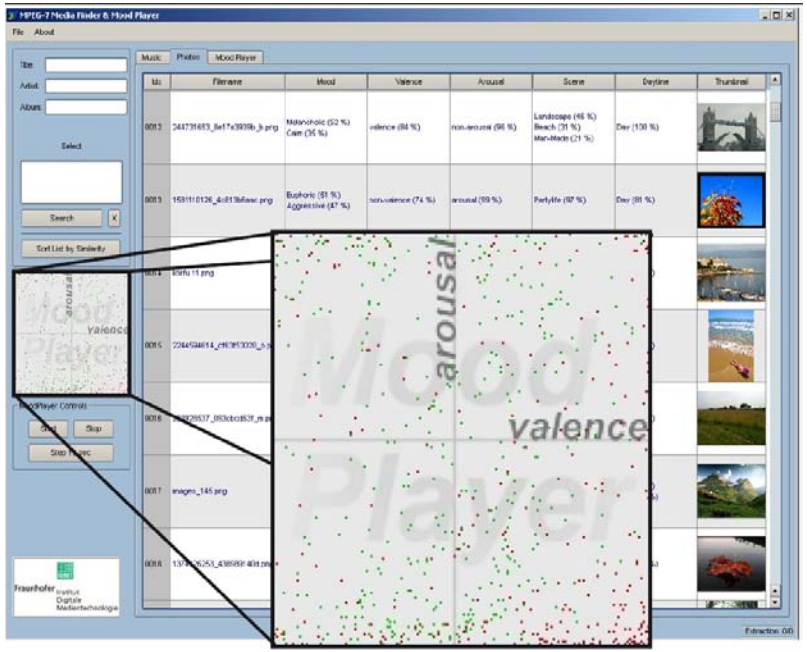
Tools to Bridge the Semantic Gaps 1:
Automatic Annotation

- Problem / motivation
  - No or few metadata (e.g. manual annotation too costly)
  - Erroneous metadata
  - Incomplete metadata

- Approach
  - Extract metadata directly from A/V content via signal analysis
  - Translate low-level content features to higher-level semantics
  - Combine metadata from different semantic levels and origin

Semantic Music Technologies
Core Technologies

- Audio Identification
  - TV consumption measurement
  - Social Media
- Automatic Music Annotation
  - Music segmentation
  - Genre- and Mood-classification
  - Timbre / rhythmic / tonal properties
- Music Recommendation
  - Content-based music similarity
  - MusicDNA, mufin, soundslike
- Music Transcription
  - Melody, Bass, Drums, Chords
  - Query by Humming System

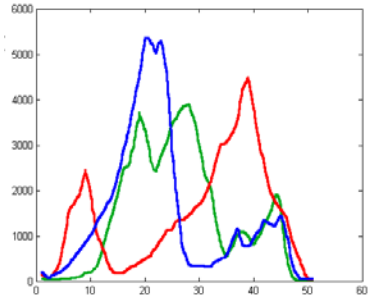


Fraunhofer IDMT

Feature Extraction

Texture / Corners / Blur

Color

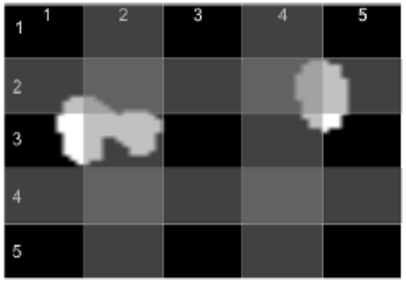


Layout/Form

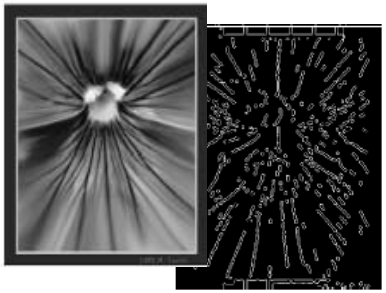
Photo metadata (EXIF)

F-Number
Shutter speed
ISO
...

Gestalt (1/3 rule, ...)

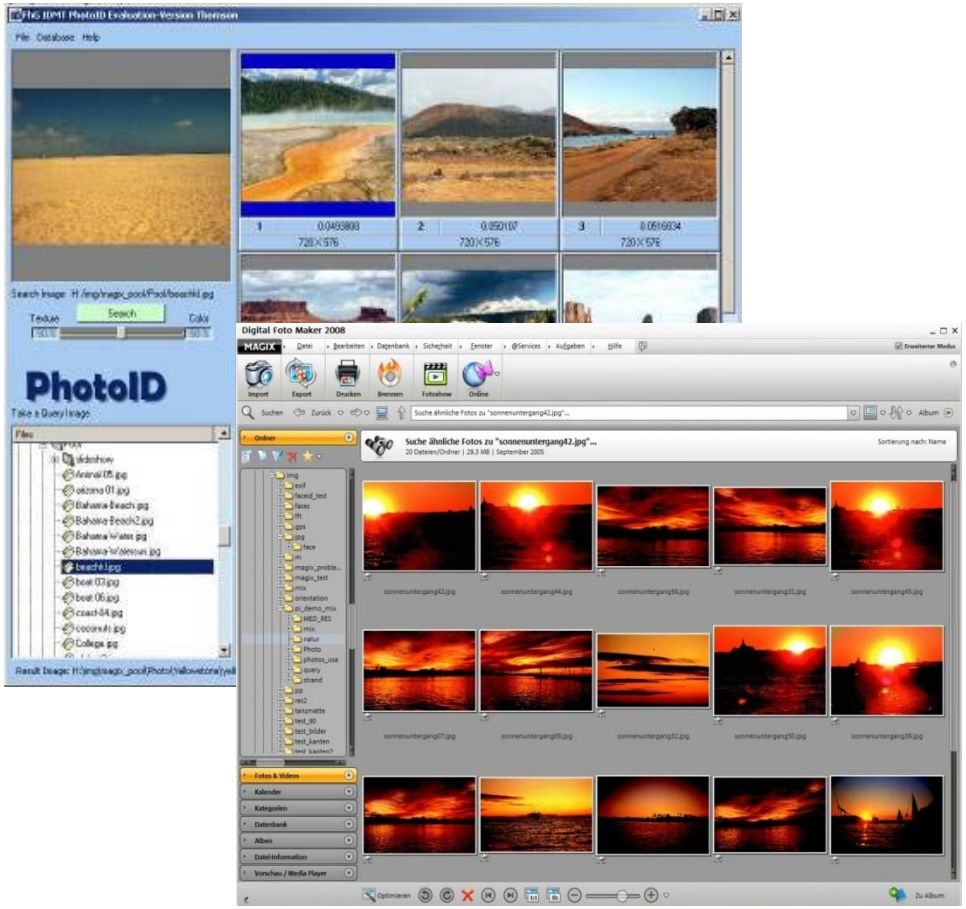


Symmetric



Fraunhofer IDMT

Photo ID / Similarity

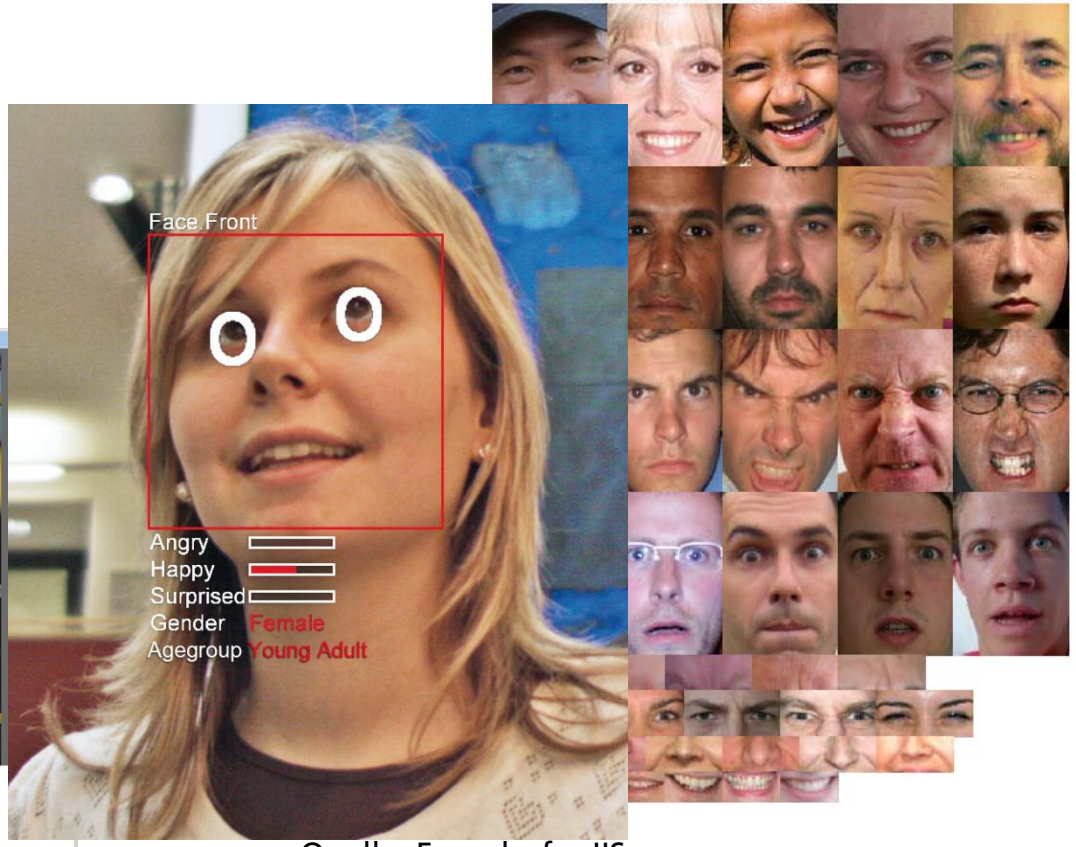


- Content-based Search
- Identical vs. Similar Photos
- Duplicate detection
- Parametrized Search:
  Color vs. Structure

Fraunhofer IDMT

Face Detection & Face Recognition

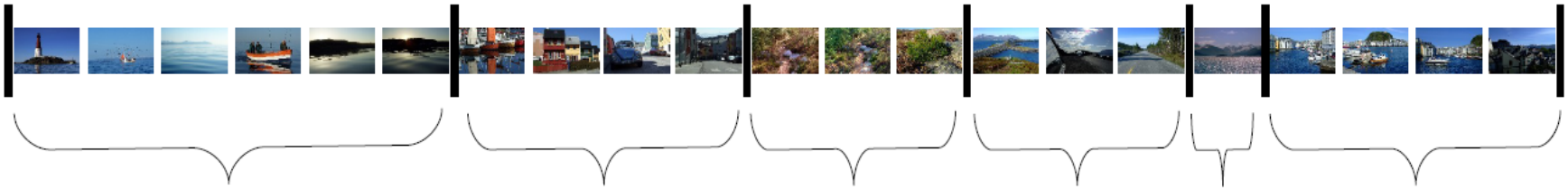
- Search for faces
- Recognize faces
- Facial expression
- Mood



Quelle: Fraunhofer IIS

Photo Summary

- Automatic generation of a representative „Photo Summary“
- Content-based fast selection: *n* out of *m*
- Usage of:
  - Quality Screening
  - Temporal Event Segmentation
  - Image Appeal
  - Near-Duplicate and Object-Duplicate Detection
  - ....



Semantic Classification

Visual Concepts

Familiy

Partylife

City

Beach

Snow

Indoor / Outdoor

Night

Sunset

Day

Appeal / Aesthetics

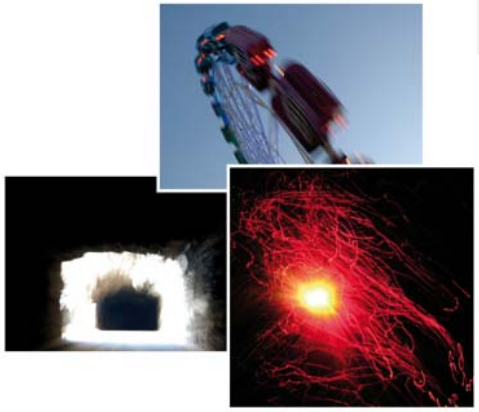
+

-

Mood: Calm, happy, aggressive, melancholic

Fraunhofer IDMT

Cross-modal Mood Model



aggressive
dramatic
agitated

arousal

euphoric
happy
playful

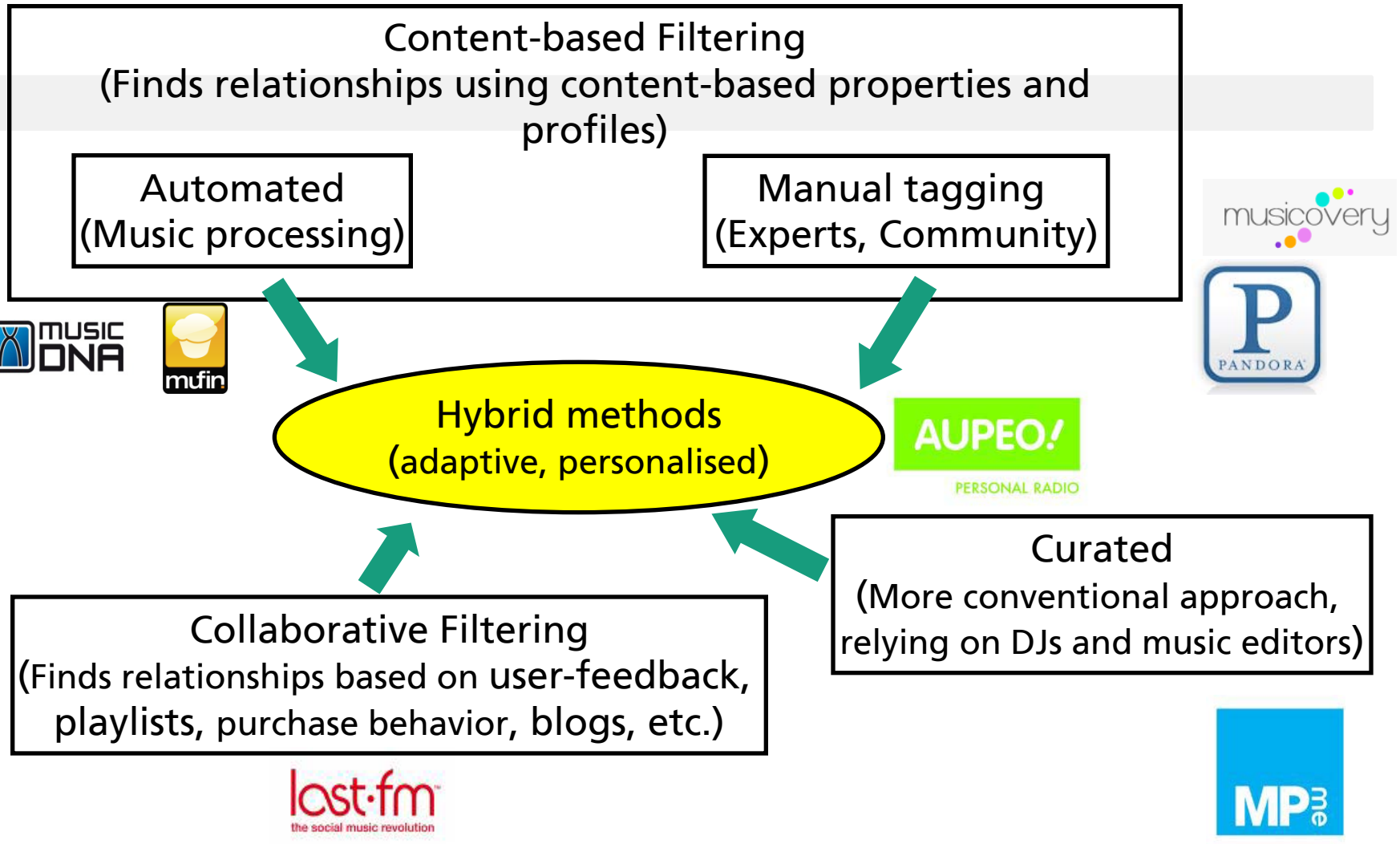
valence

melancholy
sad
depressing

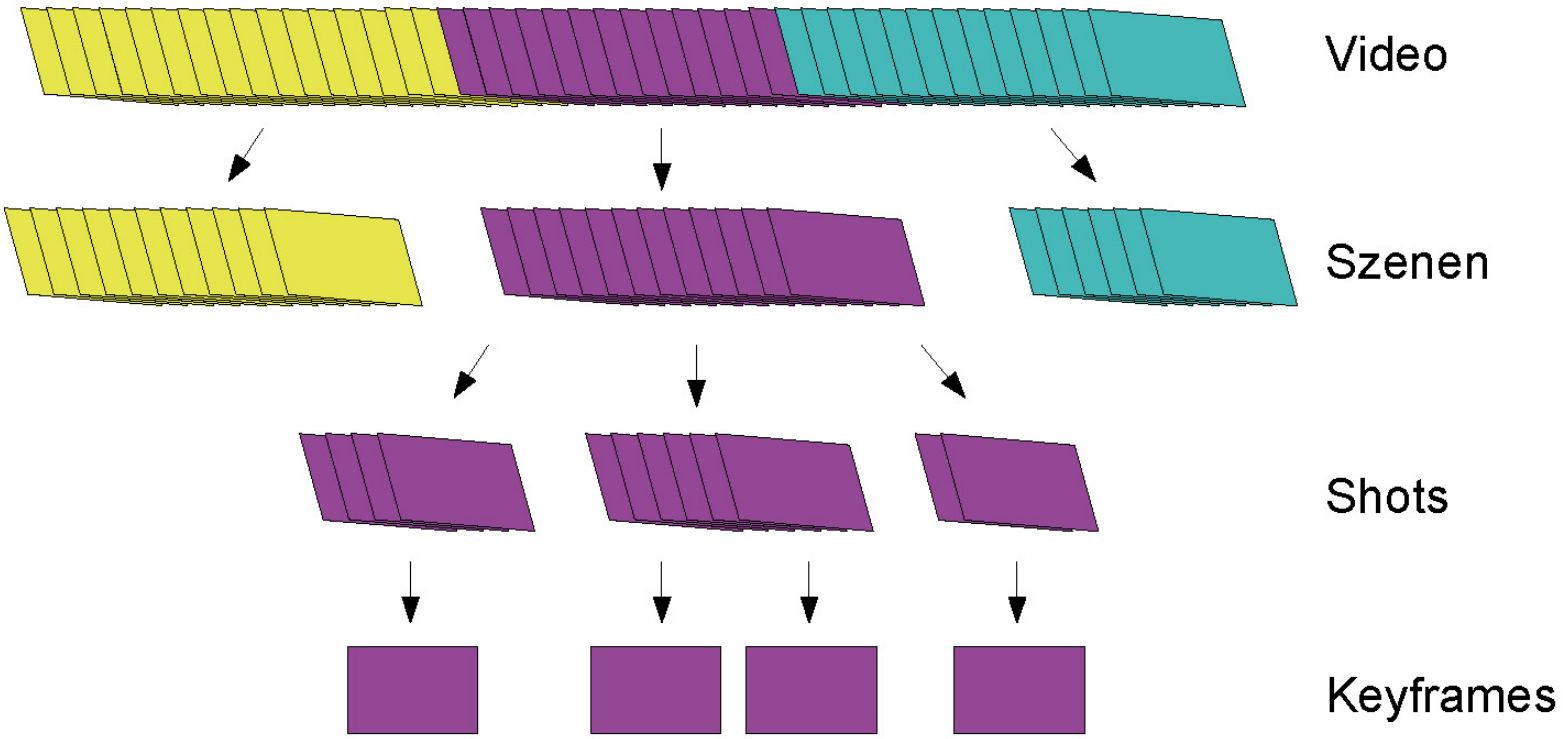
calm
soothing
dreamy

Fraunhofer IDMT

Recommendation Systems: Music

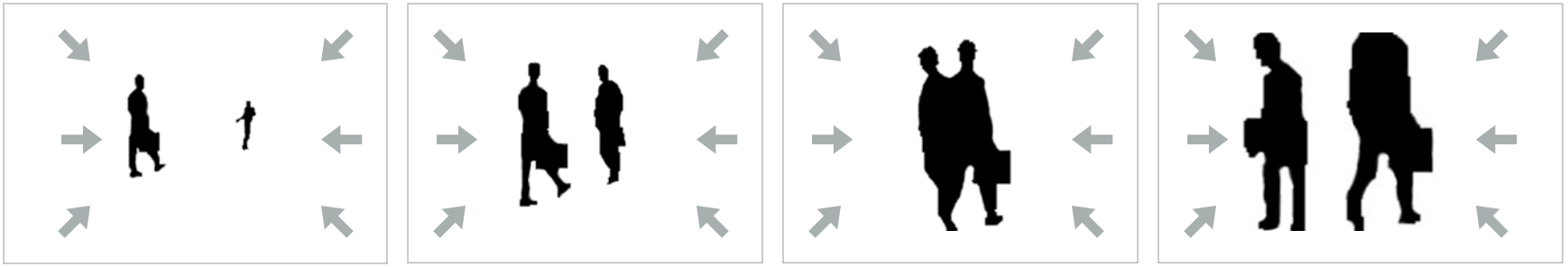


Temporal Video Segmentation

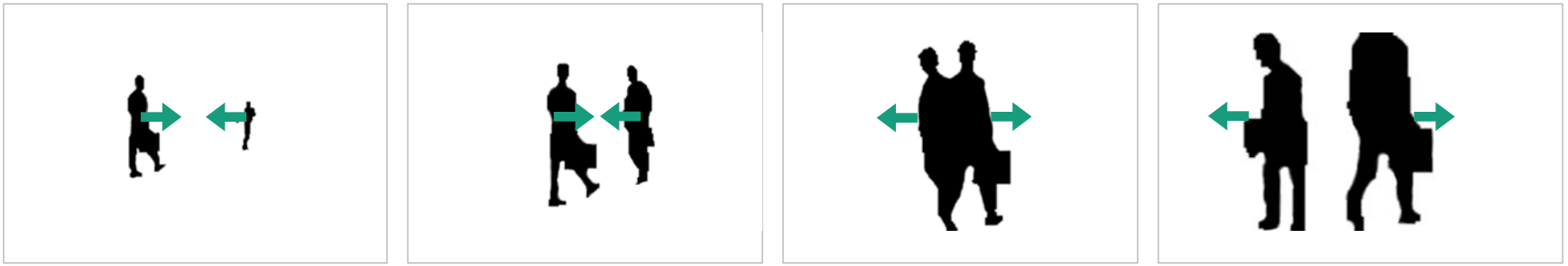


Motion Detection

Goal → **Detection and discrimination** of global and local motion



Global motion: Zoom



Local motion

Video Classification

- Classification of video scenes
  - Genre, mood, …
  - Visual concepts

Goal: automated selection and recommendation of films

- Determination of categories of film sequences
- automated classification of film sequences

Quelle: http://www.netflix.com/BrowseGenres/New_Releases/gnr?lnktrk=GID_NR

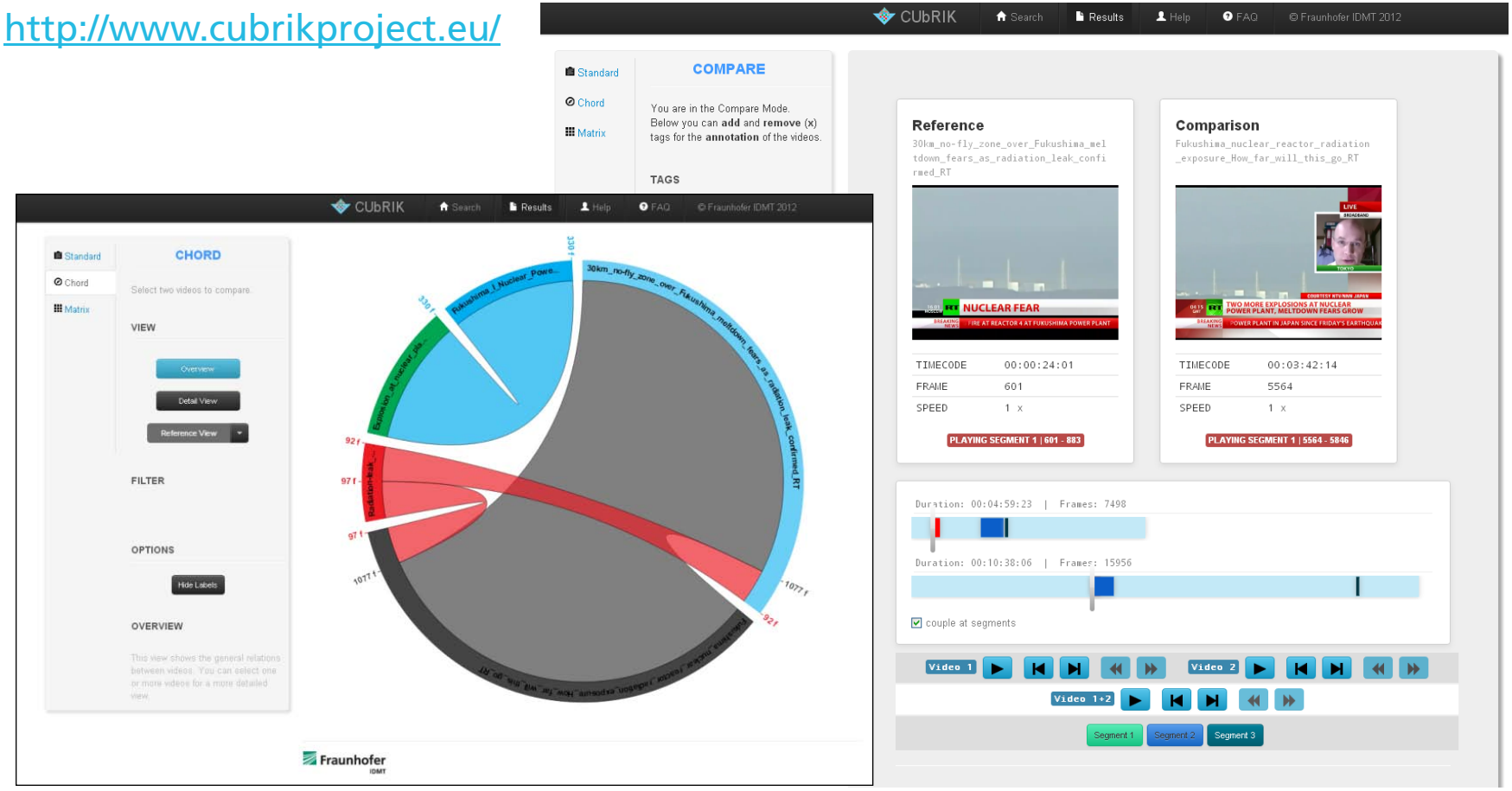
Video Identification

- Identification of videos and video segments
- based on video finger prints
- robust against distortions and manipulations
  - Grop, transcoding, compostiting, …
- Applications
  - Content tracking and repository cleanup
  - Content linking
  - Content history
  - Rights tracking
  - …
- News history @ CUbRIK

Video Identification

News Content History: GUI

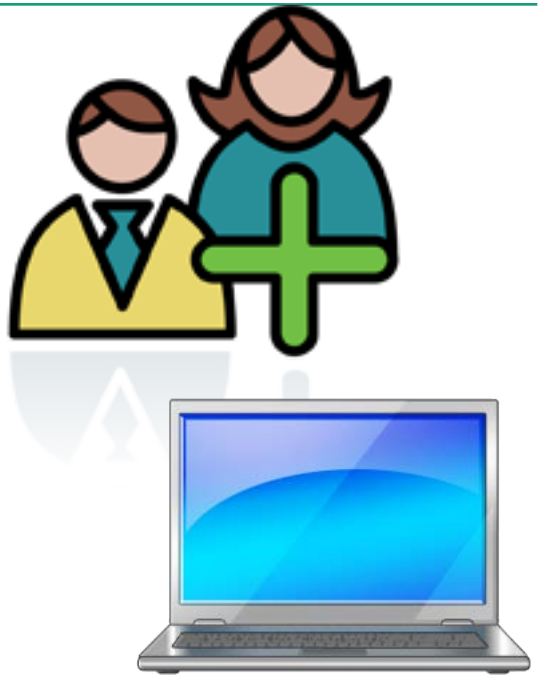
http://www.cubrikproject.eu/



Fraunhofer IDMT

Combining Automatic and Manual Annotation

- In many cases, it makes sense to combine automatic and manual annotation:
  - Some annotations tasks can be performed better by humans
  - Some annotations tasks can be performed better by machines
  - User feedback can improve automatic annotators
  - Validation by automatic annotators can improve manual annotations
  - Manual annotation can be interactively assisted by automatic annotation
- See e.g. CUbRIK
  human-enhanced multimedia search, IP, EU FP7



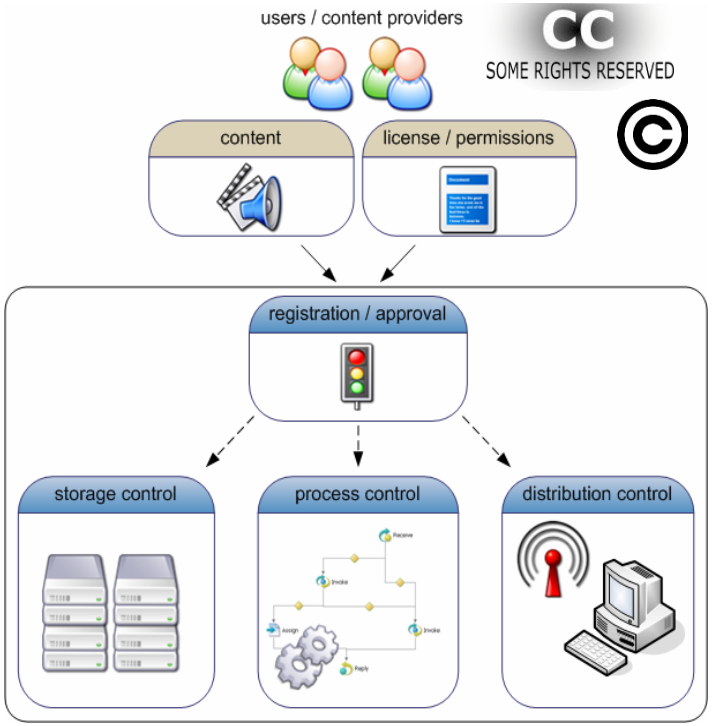
Tools to Bridge the Semantic Gaps 2:
Copyright-Awareness

- Problem / motivation
  - Huge amounts of content with heterogenuous license information
  - Content source / provenance and reliability often unclear
  - Concrete usage permissions for storage, processing, presentation and delivery often difficult to determine (copyright law too complex)
  - Huge costs for manual checking

- Approach
  - Determine trust into content / acceptance
  - Interpret license and rights metadata to fine-grain usage permissions
  - Enforce permissions within the system
  - Semi-automatic approach

Tools to Bridge the Semantic Gaps 2:
Copyright-awareness

- Content approval
  - Detect content provenance / source, to determine trustworthiness
  - Identify and authenticate content using crypto and signal analysis, to avoid forged authorship, and detect reuse
- Retrieve and interpret license and rights metadata
  (also: completion by users/crowds)
  - Derive fine-grain usage rights for storage, processing, presentation, and distribution
  - Enforce permissions throughout the system



Thank you!

Contact:
Uwe Kühhirt
uwe.kuehhirt@idmt.fraunhofer.de
+49 3677 467-205

http://www.idmt.fraunhofer.de/m2d

IBC 2013:
Visit us at booth 8.B80

Components A/V Analysis
mature, continuously extended

- A/V quality control and quality comparison
- A/V alignment, linking and recommendation
- temporal video segmentation
- visual genre and mood analysis
- video motion analysis
- visual similarity and identification
- photo summary
- speech-music-discrimination
- music transcription: rhythm, tempo, main melody, bass and chord extraction
- music annotation: style/genre, subgenre, emotion and mood, color, texture, tempo, distortion, dynamic, density, percussiveness, naturalness
- music similarity
- music segmentation
- speech audibility enhancement

Components A/V Analysis

mature, but needs application-specific training / adaptation

- audio event detection
- speech recognition
- speaker discrimination
- sampling plagiarism detection
- semantic concept classification

early prototypes

- language detection
- face detection and recognition
- audio tampering detection
- audio device / microphone classification

Media Distribution & Security

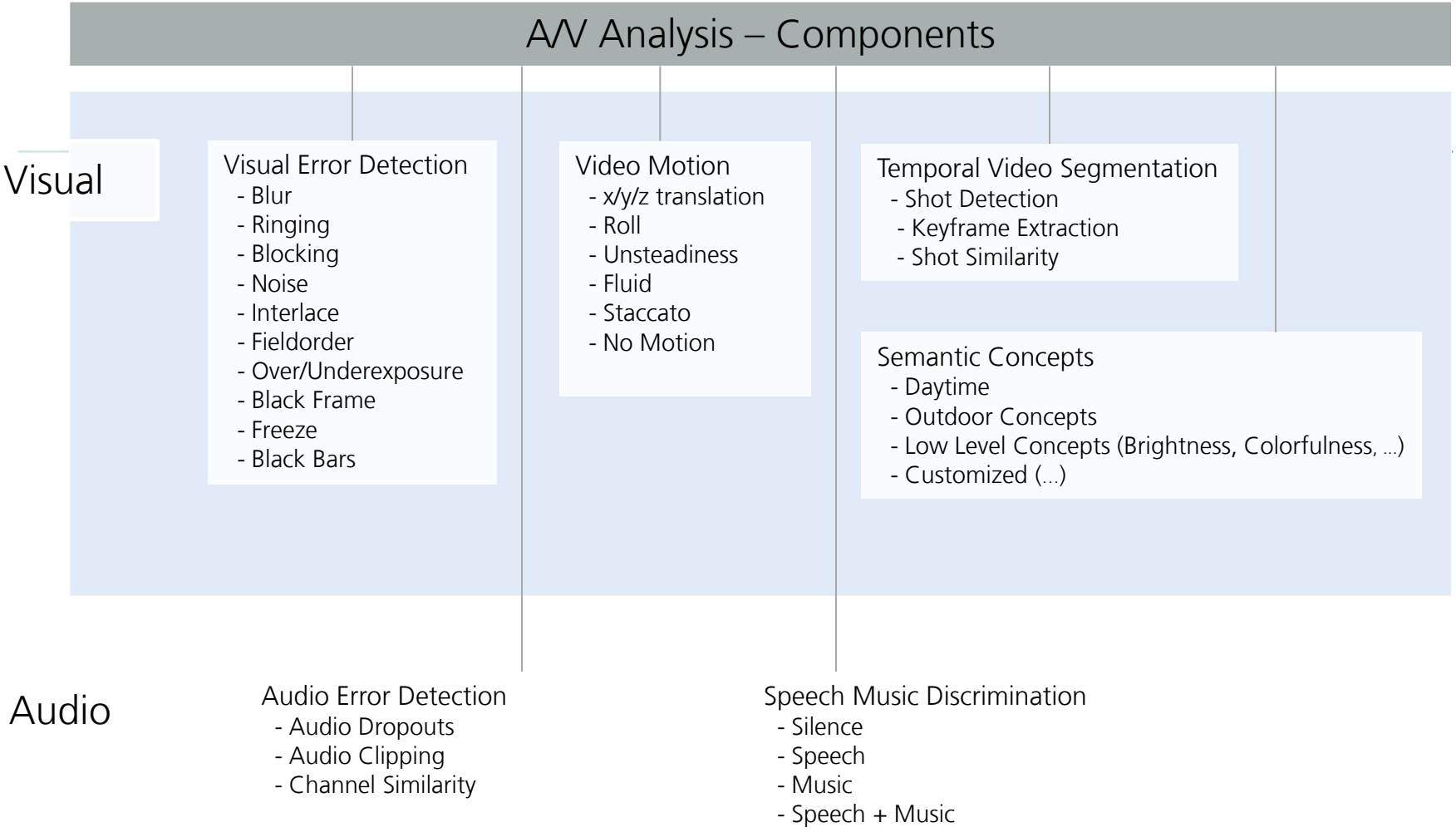
Media aggregation and distribution

- metadata modeling and interoperability, metadata cleanup
- copyright modeling and efficient, copyright-aware processing
- convergent production and distribution (broadcast, mobile, IP-based incl. P2P)
- Web Services, SOA for media processing

Privacy and security in the media context

- security, incl. could media security; data integrity, authenticity and confidentiality
- privacy-by-design and privacy enhancement technologies
- ID management and access control
- media-related cloud security aspects

A/V Analysis – Components



Possibilities for Cooperations with Fraunhofer IDMT

- Contractual R&D
  - direct commissioning, payed R&D
  - studies, consulting, development, evaluations, prototypes, …
  - if possible: long-term cooperation
- Licensing of technologies
- Collaboration within public projects
  - IDMT as a contractor or subcontractor
- Joint R&D, acquisition, and selling
- Informal cooperation
  - Exchange of ideas, requirements, example content, …
  - Preparation of further cooperations