
Fraunhofer Institute Media Communication



iFinder: Medienanalyse und -Retrieval Media-Asset-Managementsystem mit MPEG-7

Dr. Jobst Löffler
Competence Center NetMedia
<http://www.imk.fraunhofer.de>
Jobst.Loeffler@imk.fraunhofer.de



Fraunhofer
Institut
Medienkommunikation the media innovation factory

Inhalt

- **Motivation: Medienanalyse und -Retrieval**
- **Automatische Erzeugung von Metadaten**
 - Internationaler Metadatenstandard MPEG-7
- **Media-Asset-Management mit MPEG-7**
 - Medienanalyse: iFinderSDK
 - Multimedia-Retrieval: iFinder Suchmaschine
- **Demo**
- **Zusammenfassung**



Fraunhofer
Institut
Medienkommunikation the media innovation factory

Motivation: Medienanalyse und -Retrieval

Wachsende Bedeutung Medienanalyse- und Retrieval

- > Enormer Zuwachs an verfügbarer Information in Form digitaler multimedialer Dokumente (Radio, TV, Online-News-Angebote)
- > Fehlende Konzepte und Werkzeuge zur Strukturierung, Anreicherung (Wissen) und aktiven Interaktion
- > Archive (offline, analog) entwickeln sich zur wichtigen Basis aller Unternehmensprozesse (online, offen, verteilt, Schnittstellen zu allen Bereichen) **Wissensmanagement**



Wissensgewinnung (Metadatenerzeugung)

- Wer ist zu sehen und was wird gesprochen?
- Wo enden Audio- bzw. Videosequenzen?

Suchanfragen (Metadatenabfrage)

- Gib mir alle Sequenzen mit Schröder.
- Wo wurde über „Elfmeterschießen“ gesprochen?
- Gib mir alle Interviews aus der Sportschau.

Bedeutung und Anwendungsfelder

Motivation für Medienanalyse und -Retrieval

- > Werkzeuge zur effizienten Verarbeitung und zum intuitiven Umgang mit Audio-/Video-/Textmaterial
- > Automatische Wissenserzeugung: Videoschnitte, Gesichtserkennung, Logoerkennung, Szeneninterpretation, Sprechererkennung, Schrifterkennung etc.
- > Möglichkeit der gezielten Suche in großen Beständen von Audio- und Videomaterial (Essenzen) und der Materialvorschau (Browsen)

Digitale Archive

- Broadcaster, Printmedien
- Wissensmanagement Industrie
- Öffentliche Einrichtungen

Sicherheitstechnik

- Gebäudeüberwachung
- Personenidentifizierung

Audio-/Video-Retrieval

Privatanwender

- Videosammlungen etc.

Medienbeobachtung

- Radio, TV, Zeitungen
- Präsenz von Personen im TV
- Programmanalyse: Häufigkeit von Werbung und Wiederholungen

Internet / WEB

- Multimediale Suchmaschine

Medienbeobachtung

Image	NAME	FID#	First
<input type="checkbox"/>	Hofer-Jan	187	12133
<input type="checkbox"/>	UNKNOWN	192	12883
<input type="checkbox"/>	UNKNOWN	195	13335
<input type="checkbox"/>	UNKNOWN	201	13486
<input checked="" type="checkbox"/>	Steinbrueck-Peer	213	14393
<input checked="" type="checkbox"/>	Steinbrueck-Peer	215	14444

Broadcaster

- Bestand: Film ~10 Mio. h, Audio/Video ~20 Mio. h
- Wachstum: Film 0.4%, Video 10%, Audio 4%

Nutzung TV-Bereich

- stark: 25% zirkulieren jährlich in Form von VHS-Kassetten
- Retrieval: manuelle Indizierung, Suchergebnisse als VHS-Kassetten

Zustand

- Medium veraltet (66 %), Zerfall (30%), eingeschränkte Nutzung (25%)

Strategische Ziele (wachsender Markt)

- Effizienter Zugriff: überall, jederzeit, gleichzeitig, präzise
- Langzeitsicherung und Rechtsmanagement

Fraunhofer Institut Medienkommunikation the media innovation factory

Automatische Erzeugung von Metadaten

Vorteile der automatischen Metadaterzeugung

- > Schnell
- > Geringe Kosten
- > Sehr detaillierte Beschreibungen

Anwendung von modernen Erkennungsmethoden

Audio

- > Sprechersegmentierung (Bayes-Klassifikatoren, BIC)
- > Sprecherunabhängige Spracherkennung (Hidden-Markov-Modelle)
- > Optimales Audio/Text-Alignment durch Methoden der Spracherkennung und Ausnutzung vorhandener Transkripte

Video

- > Videoschnitterkennung und intelligente Keyframe-Auswahl
- > Gesichtsfindung und -erkennung (Neuronale Netze und HMM)

```

    graph LR
      A[AV-Daten] --> B[Segmentierung]
      B --> C[Erkennung]
      C --> D[Metadaten]
    
```

Fraunhofer Institut Medienkommunikation the media innovation factory

Videoindizierung: Zeitliche Segmentierung

Auffinden zusammenhängender Segmente und Extraktion von Keyframes

Abrupte Schnitte (Cut)

- > Global Histogram Comparison

$$D(i, i + 1) = \sum_{j=1}^n |H_i(j) - H_{i+1}(j)|$$

Weiche Übergänge (Dissolve, Fade Out/In)

- > Twin Comparison Method

The figure shows two histograms, (a) and (b), plotting distance $D(i, i+1)$ against frame i . Histogram (a) shows a sharp peak at frame F_{cut} , indicating an abrupt cut. Histogram (b) shows a broader peak between frames F_s and F_e , indicating a soft transition. Vertical dashed lines mark F_{cut} , F_s , and F_e .

Three grids of video frames illustrate different transition types. The first grid shows an abrupt cut between two scenes. The second grid shows a soft transition (fade) between two scenes. The third grid shows a dissolve transition between two scenes.

Koprinska, Carrato (2000)

Fraunhofer Institut Medienkommunikation the media innovation factory

MPEG-7 International Standard ISO/IEC

Hintergrund: Bedarf nach effizienten und interoperablen Anwendungen

- > Retrieval in sehr umfangreichen Multimediaarchiven
- > Editing und Synchronisierte Medien, etc.

Standardisiertes Framework zur Beschreibung von multimedialen Inhalten (XML Schema)

- > Module: Systems, DDL, Video, Audio, MDS, Ref.-Software (XM)
- > Description Definition Language (DDL)
- > Description Scheme (DSs)
- > Descriptor (Ds): Merkmalsrepräsentation von AV Daten

The flowchart shows a process starting with 'Metadaten Produktion' (Metadata Production), which leads to 'MPEG-7 Beschreibung' (MPEG-7 Description). This description is noted as the 'Normativer Teil des MPEG-7 Standards' (Normative part of the MPEG-7 Standards). The final step is 'Metadaten Anwendung' (Metadata Application).

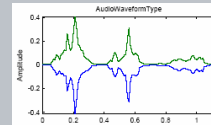
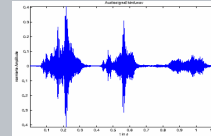
Fraunhofer Institut Medienkommunikation the media innovation factory

MPEG-7 Audio Low-Level-Deskriptoren

AudioWaveformType

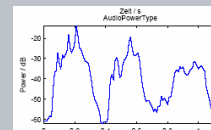
- > Minimaler und maximaler Amplitudenwert des Audiosignals in nicht-überlappenden Blöcken (HopSize 10ms)

```
<complexType name="AudioWaveformType">
  <complexContent>
    <extension base="mpeg7:AudioLLDScalarType">
      <attribute name="minRange" type="float" use="optional"/>
      <attribute name="maxRange" type="float" use="optional"/>
    </extension>
  </complexContent>
</complexType>
```



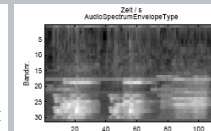
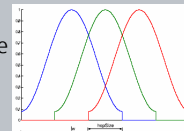
AudioPowerType

- > Zeitliches Verhalten der mittleren Leistung der Blöcke



AudioSpectrumEnvelope

- > Kurzzeit-Leistungsdichtespektrum mit logarithmischer Frequenzachse



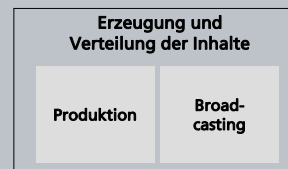
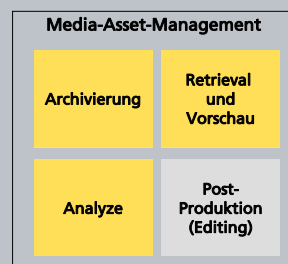
Media-Asset-Management mit MPEG-7

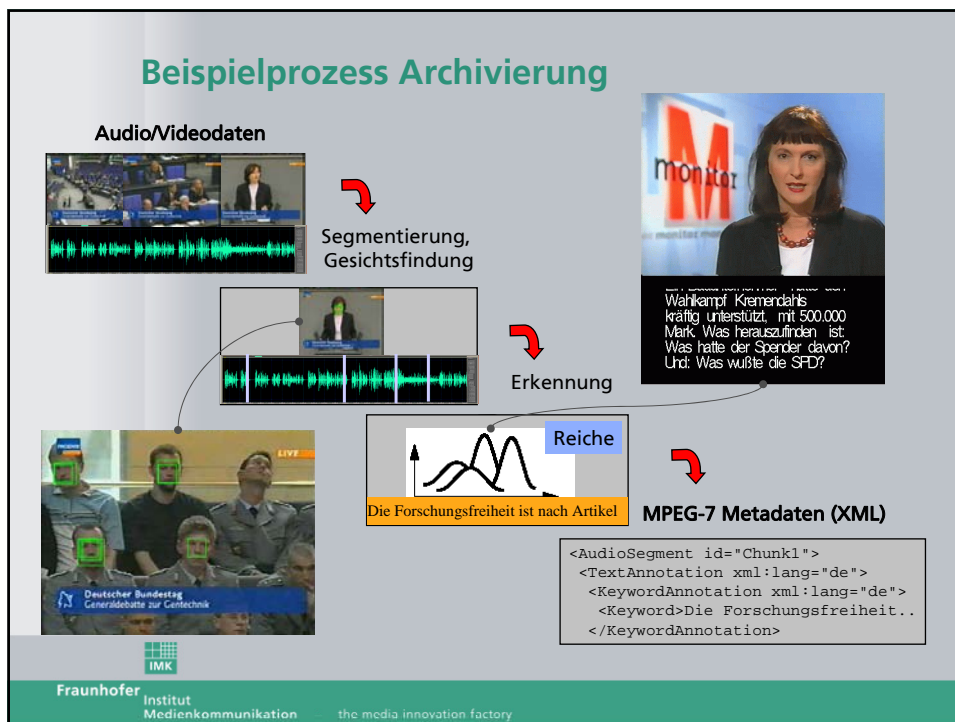
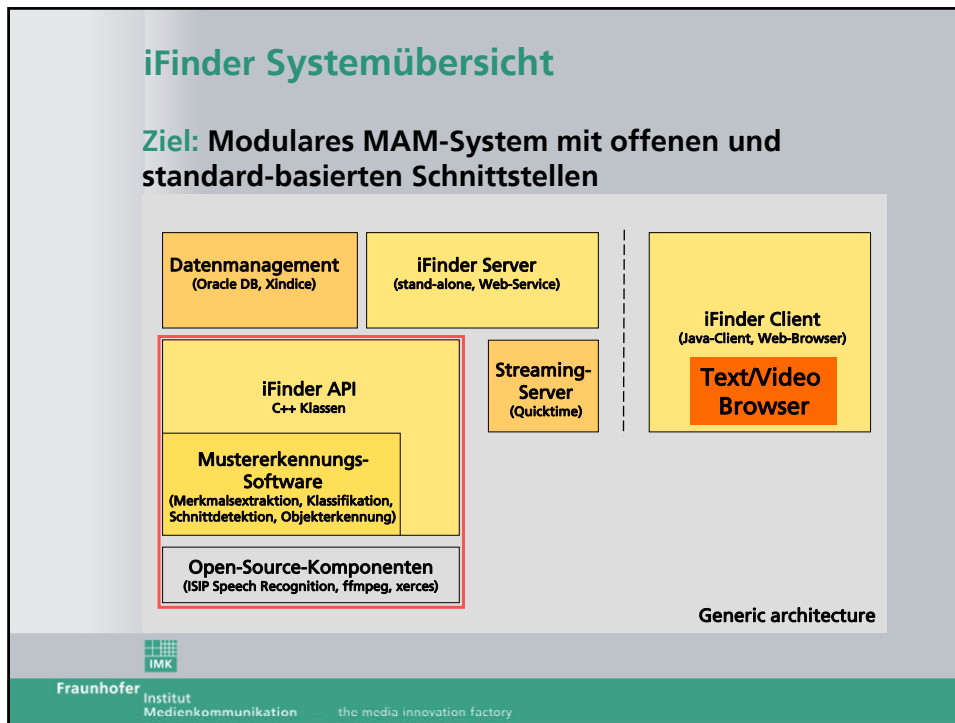
iFinder MAM-Produkt am Fraunhofer IMK entwickelt

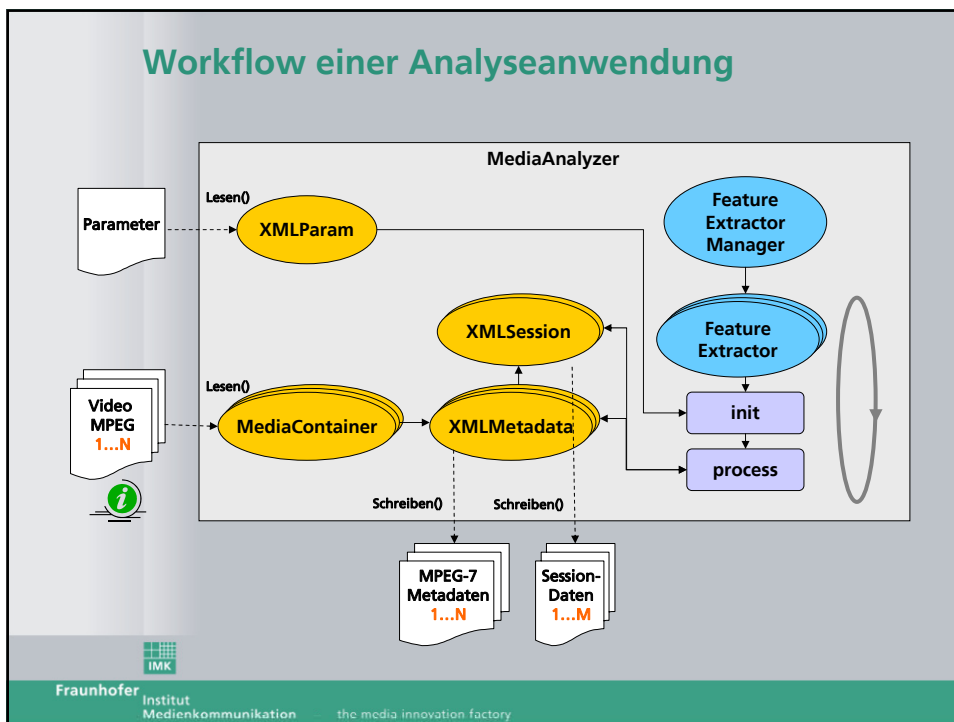
- > Automatische Medienanalyse
- > Asset-Management
- > Audio- und Videoretrieval

Anwendungsfeld Deutscher Bundestag

- > Forschungsprojekt AGMA (BMBF)
- > Setzt auf kombinierte Ansätze
- > Automatische Erzeugung von MPEG-7 Metadaten
- > Wortgenaue Suche und exakte Vorschau







Video: Gesichtsfindung auf Fernsehmaterial

Gesichtsfindung (Rowley/Kanade, 1998)

- > Vorverarbeitung: Bildpyramide, 20x20 Pixel Bildregionen
- > Merkmalsextraktion
- > Klassifikation (Neuronales Netz): Gesicht / Nicht-Gesicht
- > Nachbearbeitung: Zusammenfassung ähnlicher Hypothesen

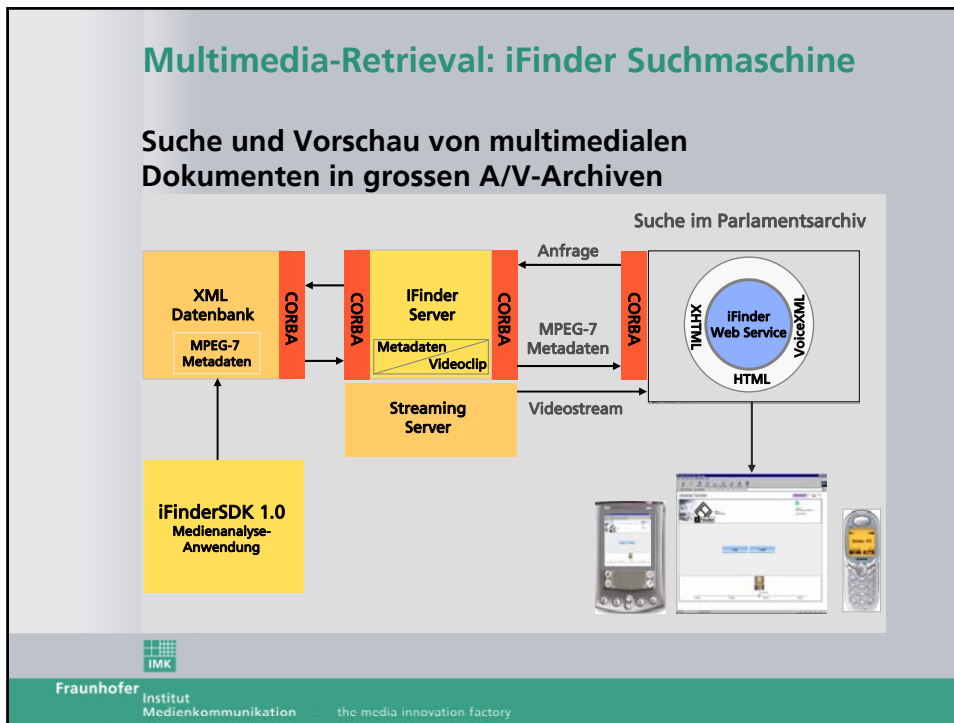
Preprocessing
Neural network

Fraunhofer Institut Medienkommunikation
the media innovation factory

Video: Gesichtsfindung auf Fernsehmaterial

Rowley-Verfahren

Fraunhofer Institut Medienkommunikation
the media innovation factory



Demo: Der Text/Video-Browser

Fraunhofer Institut Medienkommunikation the media innovation factory

Zusammenfassung

Wachsender Bedarf an Werkzeugen und Konzepten zum effizienten Umgang mit multimedialen Dokumenten

Automatische Verfahren gewinnen auch in der Praxis an Bedeutung

Der internationale ISO/IEC Standard MPEG-7 setzt sich als breit anwendbares Metadatenformat weiter durch

Das Media-Asset-Managementsystem iFinder bietet Lösungen für Analyse, Datenmanagement und Retrieval

iFinderSDK ist als Produkt verfügbar und wird in Zukunft kontinuierlich durch unsere Forschung weiterentwickelt

