

Automatische Medienanalyse im digitalen Archiv durch Einsatz von MPEG-7 und MXF

Felix Zielke und Jobst Löffler

Fraunhofer Institut Medienkommunikation, Schloss Birlinghoven
53754 Sankt Augustin, Germany
{Felix.Zielke, Jobst.Loeffler}@imk.fraunhofer.de

Abstract: Die Rolle des Archivs in den Arbeitsabläufen der Rundfunkanstalten wird mit der Konvergenz von IT-, Kommunikations- und A/V-Technik immer bedeutender. Die Wiederverwertbarkeit von Beiträgen ist zu einem wichtigen Ziel der Medienindustrie geworden. Digitale Archive erlauben den direkten und gleichzeitigen Zugriff auf gespeicherten Content. Damit Inhalte schnell und zielgenau gefunden werden können, werden qualitativ hochwertige Metadaten benötigt. Im vorliegenden Beitrag wird das Medienanalyzesystem iFinderSDK zur automatischen Erzeugung von MPEG-7-Metadaten vorgestellt. MPEG-7 bietet ein sehr detailliertes XML-basiertes Metadatenmodell zur effektiven Suche in großen multimedialen Archiven. Zur Integration der Medienanalyse in den digitalen Produktions- und Archivierungsworkflow der Rundfunkunternehmen ist jedoch auch ein Standardformat zum Austausch von A/V-Material nötig. MXF bietet sich für diesen Zweck als eine gute Lösung an. Ein Konzept zur Integration einer MXF-Schnittstelle in das bestehende Medienanalyzesystem iFinderSDK wird im folgenden genauer beschrieben.

1 Einleitung

In modernen Rundfunksystemen gerät der vernetzte Produktionspeicher in die Mitte eines sternförmigen Szenarios (Abb. 1). Die einzelnen Arbeitstationen sind in der Lage, parallel und unabhängig voneinander auf Content zuzugreifen. Content und Arbeitsprozesse werden von *Content Management Systemen* (CMS) verwaltet.

Das Archiv bildet einen Baustein innerhalb dieses Aufbaus und steht nicht länger am (vernachlässigten) Ende der Produktionskette. Daraus ergeben sich zwei wesentliche Vorteile: Das CMS kann damit auch den Zugriff auf das Archiv kontrollieren, so dass ein Redakteur am selben Arbeitsplatz aktuelles und archiviertes Material sichten kann. Zweitens kann recherchiertes Material aus dem Archiv über das vernetzte System in den aktuellen Produktionsprozess aufgenommen werden, anstatt in einem komplizierten Vorgang ausgeliefert und erneut digitalisiert zu werden.

Um die Recherchemöglichkeiten in der Archivdatenbank zu verbessern, kann das A/V-Material durch die Integration des Medienanalyzesystems iFinder automatisch mit Metadaten angereichert werden. Zur Codierung der Metadaten nutzt das iFinder-System den

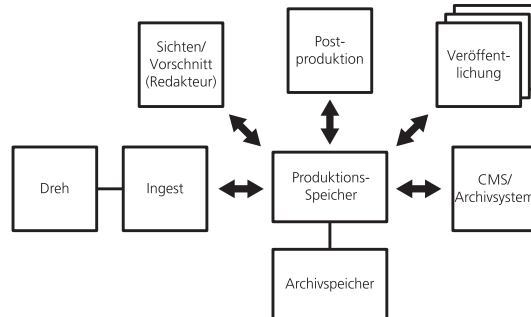


Abbildung 1: Sternförmiger Aufbau des digitalen Produktionsnetzes

internationalen Standard MPEG-7. Durch den Einsatz des standardisierten *Material Exchange Formats* (MXF) ist eine nahtlose Integration des iFinders in den modernen Rundfunk-Workflow möglich.

2 Der internationale Metadatenstandard MPEG-7

MPEG-7 [BSS02] ist ein erweiterbares standardisiertes Framework zur Beschreibung multimedialer Daten (Essenz). Der Standard verwendet W3C XML Schema zur Festlegung von Beschreibungselementen und deren Zusammenhang. Im Gegensatz zum MXF-Standard wird nicht zwischen strukturellen und beschreibenden Metadaten unterschieden. Beschreibungselemente wie zum Beispiel *<MultimediaContent>*, *<TemporalDecomposition>* und *<VideoSegment>* werden verwendet, um Metadaten für den zeitlichen Aufbau von Videoessenzen zu erzeugen.

Innerhalb der MPEG-7 Metadaten befindet sich ein Verweis auf das beschriebene multimediale Material (*<MediaURI>*). Durch die Trennung von Metadaten und Essenzen wird ein flexibler Umgang mit Formaten erreicht. Außerdem wird so eine sehr effektive Suche innerhalb des Archivs durch direkte Abfragen in den XML-Dokumenten ermöglicht.

Der MPEG-7-Standard stellt weder Werkzeuge zur Extraktion von Merkmalen zur Verfügung, noch werden Konzepte zum Aufbau von Suchmaschinen oder Filteragenten beschrieben. Das Ziel von MPEG-7 ist die Definition einer standardisierten Schnittstelle zwischen Anwendungen zur Metadatenerzeugung und zur Metadatenutzung. Aus diesem Grund wird MPEG-7 auch als *Multimedia Content Description Interface* bezeichnet. Der Standard soll ermöglichen, dass Programme unterschiedlicher Hersteller zur Medienanalyse und zur Suche in großen verteilten Archiven (z.B. im Web) problemlos zusammenarbeiten können (Abb. 2).

MPEG-7 erlaubt unterschiedliche Abstraktionsebenen: Low-Level Metadaten, wie Farbhistogramminformation oder Audiospektrogramme, können meist automatisch erzeugt und dann durch die in MPEG-7 definierten Audio- bzw. Video-Descriptor detailliert re-

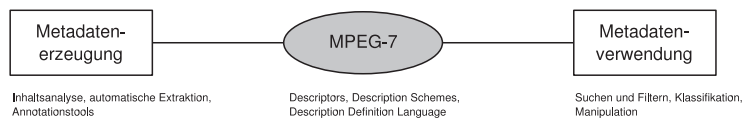


Abbildung 2: MPEG-7 als standardisierte Schnittstelle zwischen Erzeugern und Verbrauchern von Metadaten.

präsentiert werden. High-Level Metadaten, die semantische Zusammenhänge oder den Aufbau einer Videoszene wiedergeben, werden in vielen Fällen manuell annotiert.

3 Das Material Exchange Format MXF

Das Material Exchange Format [DW04] wird vom *Pro-MPEG-Forum* entwickelt und ist 2004 von der *Society of Motion Picture and Television Engineers* (SMPTE) als S377M [SMP04] standardisiert worden.

MXF ist ein *Wrapper*-Format, das A/V-Material zusammen mit Metadaten in einer Datei verpackt, um den Austausch zwischen Systemen über hersteller- und plattformspezifische Grenzen hinweg zu ermöglichen. Die innerhalb einer MXF-Datei obligatorisch enthaltenen strukturellen Metadaten geben dafür eindeutig Auskunft über Format, Länge, Kompression und Komplexität der enthaltenen Bild- und Tondaten. A/V-Daten werden nicht neu codiert sondern im *Generic Container* in die MXF-Datei eingebettet.

Der physikalische Aufbau einer MXF-Datei basiert auf dem *Data Encoding Protocol* [SMP01]. Es definiert die Codierung der Datenpakete im KLV-Format. Ein *Key* (K) ist ein bei der SMPTE registriertes Label, das das Datenpaket eindeutig identifiziert. Darauf folgt ein *Length*-Feld (L), welches die Länge des nun folgenden *Value*-Feldes (V) bestimmt, das wiederum die eigentlichen Daten transportiert. Diese Codierung ermöglicht es einem Decoder, unbekannte Datenpakete anhand der Längenangabe zu überspringen.

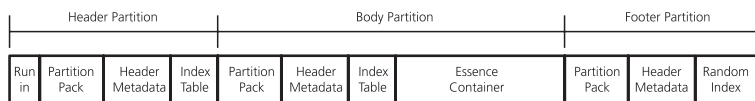


Abbildung 3: Die Dateistruktur einer MXF-Datei

Abbildung 3 zeigt die Dateistruktur der gesamten MXF-Datei. In der Header Partition befinden sich sowohl die strukturellen Metadaten als auch optionale Metadaten, die den Inhalt der Essenzen beschreiben. Im Body befinden sich alle A/V-Daten. Im optionalen Footer können die Metadaten aus dem Header aktualisiert werden, etwa wenn eine MXF-Datei aus einem A/V-Stream erzeugt wird. Optionale Indextabellen lokalisieren die Essenzen innerhalb der Datei und beschleunigen so den direkten Zugriff.

Der komplexere logische Aufbau einer MXF-Datei basiert auf einem objektorientierten Modell, das von den strukturellen Metadaten innerhalb der Datei umgesetzt wird. Alle Essenzen befinden sich demnach in sogenannten *File Packages*. Ein MXF-Player spielt nur einen definierten Teil der Essenzen ab, die im sog. *Material Package* liegen, das die *Playlist* der MXF-Datei darstellt. Jedes Package beinhaltet genau einen Timecode-Track, sowie beliebig viele dazu synchron laufende Audio- und Videotracks. Die Tracks können wiederum in Segmente eingeteilt sein, die etwa Schnitte im Videomaterial definieren.

Einer MXF-Datei können neben Essenzen und strukturellen Metadaten beliebig viele beschreibende Metadaten zugefügt werden. Sie werden in einem virtuellen DM (*descriptive metadata*)-Track mit dem Timecode-Track innerhalb eines Packages synchronisiert. Die Daten können einzelne Frames, Clips oder die gesamte Datei beschreiben.

4 iFinderSDK - MPEG-7-basierte Medienanalyse

Das iFinderSDK (*Software-Development-Kit*) [LBK04] ist Teil des iFinder-Systems zur Analyse und zum Retrieval von multimedialen Daten. MPEG-7-Metadaten werden bei der Medienanalyse automatisch erzeugt (iFinderSDK) und dann zur Suche in verteilten Archiven verwendet (iFinder Suchmaschine). Mit dem iFinderSDK, das modulbasiert aufgebaut ist, können Medienanalyseanwendungen flexibel erstellt werden. Die automatische Analyse von A/V-Material unterstützt Archivare dabei, Fernseh- und Radiobeiträge mit sinnvollen Metadaten zu versehen. Ziel ist die Wiederverwertbarkeit des Materials, das im Archiv themengenaue gesucht werden kann.

iFinderSDK setzt automatische Verfahren zur Segmentierung, zum Clustering und zur Erkennung ein. Derzeit sind folgende Analysemodule im iFinderSDK enthalten: *Audiosegmentierung*, *Audioklassifizierung* (Sprache, Nicht-Sprache), *Sprecher-Clustering*, *Audio/Text-Alignment*, *Keyword-Spotting*, *Videoschnitterkennung* (weiche und harte Übergänge) sowie *Gesichtsfindung*. Module zur Audio-Jingle-Detektion, zum Gesichts-Clustering und zur Logoerkennung sind in Vorbereitung. Das iFinderSDK wird vom Fraunhofer IMK entwickelt und ist als Produkt für die Betriebssysteme Linux und Windows erhältlich.

5 Integration des iFinderSDK in den Archivworkflow mit MXF

Zur Integration der iFinder-Medienanalyse in den Archivierungs-Workflow von Rundfunkanstalten wird das iFinderSDK um eine Schnittstelle zur Eingabe und Ausgabe von MXF-Dateien erweitert. In der iFinderSDK-Analyse muss unterschieden werden, welche der erzeugten Metadaten die Struktur der MXF-Dateien verändern sollen, und welche optional als *Descriptive Metadata* (DM) in die MXF-Dateien integriert werden können. Die Module, die das AV-Material segmentieren, können strukturelle Metadaten erzeugen, die die MXF-Tracks in Segmente unterteilen. Die Ergebnisse der weiteren Module dienen der näheren Beschreibung der Essenzen und liegen zur Datenbankrecherche weiterhin im MPEG-7-Format vor. Sie können optional zusätzlich in die MXF-Dateien integriert wer-

den, abhängig davon, ob sie in anderen Systemen des Produktionsprozesses benötigt werden. Eine denkbare Anwendung wäre zum Beispiel eine zeitsynchrone Transkription zur Erstellung von Untertiteln mit dem *Text-Alignment-Modul* des iFinderSDK.

Abbildung 4 zeigt das Konzept der MXF-kompatiblen Archivilösung für Rundfunkanstalten in einem möglichen Einsatzszenario. Viele Rundfunkunternehmen setzen bereits Produktionssysteme ein, die Arbeitsplätze zum Browsen im Produktionsmaterial in Vorschauqualität integrieren. Die Redakteure können hier ihr Material sichten oder bereits vorschneiden. Durch die Integration des iFinder-Clients kann auch das Browsen im Archiv vom selben Arbeitsplatz aus ermöglicht werden.

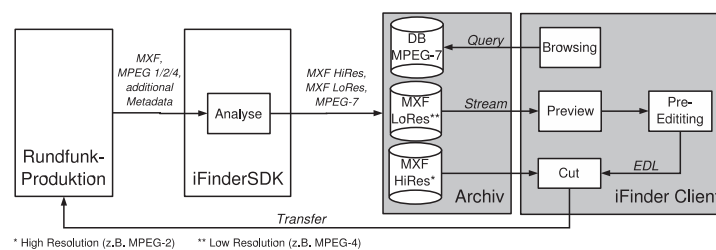


Abbildung 4: Archivworkfbw mit dem iFinder

6 Zusammenfassung

Der Beitrag hat gezeigt, wie A/V-Material aus der Rundfunkproduktion zur modernen Archivierung automatisiert mit Metadaten angereichert werden kann. Es wurden Möglichkeiten des Einsatzes von MPEG-7 zur Beschreibung der Ergebnisse des Medienanalyse-Systems iFinder erläutert. Außerdem wurde eine sinnvolle Kombination von MPEG-7-Metadaten mit dem standardisierten Austauschformat MXF diskutiert.

Literatur

- [BSS02] B.S.Manjunath, P. Salembier und T. Sikora, Hrsg. *Introduction to MPEG-7*. John Wiley & Sons, Ltd., 2002.
- [DW04] B. Devlin und J. Wilkinson. The Material Exchange Format. In Brad Gilmer, Hrsg., *File Interchange Handbook*, Seiten 123–176. Elsevier Inc., Focal Press, 2004.
- [LBK04] J. Löffler, K. Biatov und J. Köhler. Automatic Extraction of MPEG-7 Audio Metadata Using the Media Asset Management System iFinder. *Proceedings AES 25th International Conference: Metadata for Audio*, 2004.
- [SMP01] SMPTE. S336M: Data Encoding Protocol using Key-Length-Value, 2001.
- [SMP04] SMPTE. S377M: Material Exchange Format (MXF) - File Format Specification, 2004.