

DIGITALE INTELLIGENTE VIDEOANALYSE

A.Dammeyer, W.Jürgensen, C.Krüwel, E.Poliak, S.Rutkowski, Th.Schäfer, M.Sirava und
Th.Hermes

Technologie-Zentrum Informatik
Universität Bremen (FB 3)

E-mail: [mrm|wibke|krue52|poliak|svenr|tstheo|zaris|hermes}@tzi.org](mailto:{mrm|wibke|krue52|poliak|svenr|tstheo|zaris|hermes}@tzi.org)

Zusammenfassung

Die Bedeutung von Videos nimmt für multimediale Systeme stetig zu. Dabei existiert eine Vielzahl von Produkten zur Betrachtung von Videos, allerdings nur wenige Ansätze, den Inhalt eines Videos zu erschließen. Das DiVA-System dient der automatischen Analyse von MPEG-I Videofilmen. Der dabei verfolgte Ansatz läßt sich in vier Phasen gliedern. Zunächst wird der Videofilm durch eine Shotanalyse in seine einzelnen Kameraeinstellungen (Shots) unterteilt. Darauf aufbauend findet eine Kamerabewegungsanalyse sowie die Erstellung von Mosaicbildern statt. Mit Methoden der künstlichen Intelligenz und der digitalen Bildverarbeitung wird das analysierte Material nach Bild- und Toninformationen ausgewertet. Das Resultat ist eine textuelle Beschreibung eines Videofilms, auf der mit Hilfe von Text-Retrieval-Systemen recherchiert werden kann.

Einleitung

Für multimediale Systeme nehmen Videos eine immer größere Rolle ein. Es existieren jedoch bis zum heutigen Zeitpunkt nur sehr wenige Ansätze, um ein Video inhaltlich zu erschließen und so eine Recherche auf umfangreichem Videomaterial zu ermöglichen. Bestehende Video Retrieval Systeme verwenden unterschiedliche Techniken bei der Videoanalyse und der Rechercheunterstützung – IRIS [KRA+97], Virage Video Engine [GHS+97] [GJ97] und Euromedia [TEC98].

Alle Systeme analysieren die Bildinformationen des Videos anhand von Textur-, Farb- und Kantenmerkmalen. Die Analyse von Audioinformationen wird nur von der Virage Video Engine unterstützt. Bewegungsinformationen werden von allen Systemen, zumindest für die Generierung von Einzelbildern (Mosaicing), verwendet. Die Erkennung von Objekten und damit die Möglichkeit einer späteren Suche nach Inhalten anstatt nach Merkmalen wird vom IRIS-System in Ansätzen, von den anderen Systemen gar nicht unterstützt.

Das DiVA-System analysiert Videos mit Hilfe einer Kombination verschiedener Verfahren zur Bild- und Audioanalyse sowie Methoden der künstlichen Intelligenz. Die Bildanalyse beinhaltet zusätzlich eine Optical Character Recognition (OCR) während unter Audioanalyse sowohl Sprach- als auch Geräuschanalyse verstanden wird. In Zusammenarbeit mit der Fernsehanstalt Radio Bremen befaßt sich das Projekt damit, den Dokumentar bei der inhaltlichen Erschließung (Annotation) der regionalen Nachrichtensendung *Buten & Binnen* zu unterstützen.

Die nach Bild- und Sachinhalten manuell durchgeführten Annotationen werden heutzutage im allgemeinen in einer Datenbank abgelegt, um z.B. Redakteuren oder Fernsehanstalten eine Recherchemöglichkeit zu bieten.

Eine unzureichende technische Unterstützung führt dabei zu folgenden Nachteilen: a) Lange Annotationszeit - das Verhältnis der benötigten Annotationszeit zur Sendezeit des Videos ist unangemessen hoch. Man geht von einer achtmal so langen maximalen Annotationszeit aus; b) Subjektivität - aus der subjektiven Wahrnehmung des Archivars resultieren unterschiedliche Annotationen. Einerseits ist die Auswahl von Sach- und Bildinhalten, die annotiert werden, verschieden, andererseits verwendet jeder Archivar selbst bei gleicher Auswahl der Inhalte ein unterschiedliches Vokabular zur Beschreibung. Auch Annotationsregeln schaffen nur eingeschränkt Abhilfe; c) Informationsverlust - aufgrund des historisch bedingten Wissenstandes eines Archivars, kann „älteres“ Videomaterial eventuell nicht korrekt annotiert werden; d) Routinetätigkeit - Teile der Auswer-

tung eines Videos sind Routinetätigkeiten und können somit als Prozeß technisch gut unterstützt werden.

Ziel des DiVA-Systems ist sowohl die möglichst weitgehende Unterstützung des Dokumentars bei der inhaltlichen Erschließung des Videomaterials durch eine automatische Analyse sowie die Unterstützung von Recherchen nach Sach- und Bildinhalten mit der Möglichkeit einer Videovorschau. Im folgenden werden einige der in dem Projekt DiVA verwendeten Ansätze vorgestellt. Der Abschnitt *Shotanalyse* beschäftigt sich mit der Aufbereitung des MPEG-I Videos für die Bild- und Audioanalyse. Hierfür wird das Video durch eine Shotanalyse in einzelne Kameraschnitte (Shots) unterteilt. Eine Kamerabewegungsanalyse liefert dann für jeden einzelnen Shot Informationen über die Kamerabewegung.

Nachfolgend wird dargestellt, wie die einzelnen Ergebnisse von Bild-, Sprach- und Geräuschanalysen zu der Gesamtklassifikation einer Videosequenz nach Bild- und Sachinformation zusammengeführt werden können. In dem Projekt DiVA wurde ein regelbasierter Ansatz sowie die Repräsentation von Wissen mit Hilfe von Datenbanken verfolgt [BBMR89].

Im Abschnitt *Architektur des DiVA-Systems* wird schließlich die grundlegende Architektur des DiVA-Systems vorgestellt.

Shotanalyse

Aufgabe der Shotanalyse im Rahmen des Projektes DiVA ist die Unterteilung des regionalen Nachrichten-Magazins *Buten & Binnen* in einzelne Kameraeinstellungen (Shots), um die weiteren Analysen des Bildmaterials pro Kameraeinstellung durchführen zu können.

Da die Unterteilung in Shots für die Qualität der nachfolgenden Analysen von entscheidender Wichtigkeit ist, wurde in der Entwicklung Wert auf ein möglichst genaues Ergebnis gelegt, d.h. eine Anforderung an die Shotanalyse war die framegenaue Detektion von Shotgrenzen. Die Shotanalyse muß somit jedes Frame betrachten.

Algorithmus der Shotanalyse

Die Detektion der einzelnen Shots basiert auf einer Analyse von Differenzen der Luminanz- und Chrominanz-Werte zweier direkt aufeinanderfolgender Frames. Hierbei werden die Differenzen zweier in der Anzeigereihenfolge aufeinanderfolgender Frames untersucht. Die Analyse der Differenzen zweier in der Dekodierreihenfolge aufeinanderfolgender Frames würde ein falsches Ergebnis liefern. Im weiteren wird davon ausgegangen, daß jedes dekodierte Frame im YUV 4:1:1 Format vorliegt (vergleiche [KÜS95] [ISO93]).

In einem ersten Schritt werden die Luminanz- und die Chrominanz-Werte aufsummiert. Dabei werden die zwei Farbanteile U und V der Chrominanz getrennt betrachtet. Eine Normierung durch die Anzahl der Makroblöcke pro Frame (396 in der Standard MPEG-I Auflösung $N=352*288$) liefert die Unabhängigkeit von der Auflösung des MPEG-I Bitstroms:

$$Y_{Sum_{norm}} = \frac{1}{396} \left(\sum_{i=0}^N Y_i \right), U_{Sum_{norm}} = \frac{1}{396} \left(\sum_{i=0}^{\frac{N}{4}} U_i \right), V_{Sum_{norm}} = \frac{1}{396} \left(\sum_{i=0}^{\frac{N}{4}} V_i \right)$$

Sei nun Frame B ein in der Anzeigereihenfolge auf Frame A direkt folgendes Frame. Dann werden die Differenzen

$$YDiff = YSum_{norm}(B) - YSum_{norm}(A), UDiff = USum_{norm}(B) - USum_{norm}(A), VDiff = VSum_{norm}(B) - VSum_{norm}(A)$$

mit den jeweiligen Schwellwerten ThY , ThU und ThV für die Luminanz-Werte, den U-Anteil der Chrominanz-Werte und den V-Anteil der Chrominanz-Werte verglichen.

Eine Shotgrenze zwischen Frame B und Frame A wird bei einer der folgenden Bedingungen detektiert:

$$(YDiff > ThY) \text{ und } (UDiff > ThU) \text{ und } (VDiff > ThV), (YDiff > (2 * ThY)) \text{ und } (UDiff > (2 * ThU)), \\ (YDiff > (2 * ThY)) \text{ und } (VDiff > (2 * ThV)) \text{ oder } (UDiff > (2 * ThU)) \text{ und } (VDiff > (2 * ThV)).$$

Experimentelle Ergebnisse

Die Qualität der Shotanalyse wurde mittels Beispielfilmen aus 5 unterschiedlichen Genres getestet. Dabei wurden jeweils eine Vielzahl typischer Beispiele aus den Genres Werbung, Spielfilme, Cartoons, Sport und Nachrichten verwendet. Tabelle 1 stellt die Erkennungsraten der Shotanalyse aufgeschlüsselt nach den betrachteten Genres dar.

Genre	Erkennungsrate
Sport	93 %
Nachrichten	97 %
Spielfilme	99 %
Cartoons	89 %
Werbung	96 %
Mittlere Genauigkeit	94,8 %

Tabelle 1 Genauigkeit der Shotanalyse

In den durchgeführten Tests hat sich gezeigt, daß die im Projekt DiVA verwendete Shotanalyse vor allem auf extrem schnelle Bewegungen (schnelle Kameraschwenks) sowie auf Blitzlichter fehlerhaft reagiert. Da durch ein Blitzlicht jedoch die gesamte Bildhelligkeit ansteigt, ist es möglich die Shotanalyse dahingehend weiterzuentwickeln, daß auch Blitzlichter erkannt werden und falsch erkannte Shotgrenzen ausgeschlossen werden.

Kamerabewegungsanalyse

Die Kamerabewegungsanalyse des DiVA-Systems liefert die Bewegungsrichtung der Kamera bzw. die dazu komplementäre Bewegungsrichtung des Bildinhalts pro Shot. Hierzu werden die bei der Motion Estimation entstandenen Informationen eines MPEG-I Bitstroms genutzt [DJK+98]. Da jeweils die Kamerabewegung pro Shot ermittelt wird, und sowohl bei der Shotanalyse als auch bei der Kamerabewegungsanalyse der MPEG-I Bitstrom von Anfang bis Ende analysiert werden muß, bietet es sich an, beide Analyseverfahren in einem Durchgang auszuführen.

Vorgehensweise der Kamerabewegungsanalyse

Die in [DJK+98] beschriebene Motion Estimation wird in MPEG-I für die Makroblöcke der P- und B-Frames durchgeführt, wobei aber nicht zwangsläufig jeder Makroblock dieser Frames mit einem Vektor kodiert sein muß. Das heißt die Makroblöcke, zu denen der Kodierer keinen Vektor ermitteln konnte, werden als Intra-Makroblöcke kodiert. In der Praxis zeigt sich, daß mindestens die Hälfte der Makroblöcke pro P- und B-Frame mit einem Vektor kodiert sind.

Die DiVA-Kamerabewegungsanalyse beschränkt sich auf die Untersuchung der P-Frames. Durch diese Einschränkung ergibt sich ein Geschwindigkeitsgewinn, und die Anzahl der betrachteten Frames ist weiterhin ausreichend, wenn man davon ausgeht, daß eine Kamerabewegung von einer gewissen Dauer ist.

Kamerabewegungen sind beliebige Kombinationen von Kameraschwenks nach *oben*, *unten*, *rechts* oder *links*, die ihre Hauptrichtung über die Dauer eines Shots nicht verändern.

Eine verwackelte Kameraführung würde zwar dazu führen, daß von einem Frame A zu dem Folgeframe B Vektoren, die einheitlich in eine Richtung zeigen, vom Kodierer und somit von der Bewegungsanalyse erkannt werden, aber diese Bewegung besitzt im allgemeinen keine einheitliche Richtung über einen längeren Zeitraum, also über die gesamte Dauer eines Shots. Da die Kamerabewegungsanalyse aber die Richtung der Kamerabewegung pro Shot ermittelt, führen verwackelte Kameraführungen nicht zu Fehlern.

Zwischen einer Kamerabewegung und einer Bewegung des Bildinhalts muß unterschieden werden, d.h. es ist ein Unterschied, ob z.B. ein Kameraschwenk an einem Zug vorbei durchgeführt wird oder ob sich der Zug an der Kamera vorbei bewegt. Die Konsequenz aus dieser Überlegung ist, daß eine Mindestanzahl von Makroblöcken mit einem Vektor kodiert sein muß, damit die Bewegungsanalyse ausgeführt wird.

Algorithmus der Kamerabewegungsanalyse

Die Bewegungsanalyse wird ausgeführt, wenn das betrachtete Frame ein P-Frame ist, und die Anzahl der mit einem Vektor kodierten Makroblöcke größer als die Hälfte der Gesamtanzahl der Makroblöcke pro Frame ist. Die jeweiligen Vektoren liegen unterteilt nach horizontaler und vertikaler Länge kodiert im MPEG-I Bitstrom vor. Zunächst werden jeweils die Summen der horizontalen und vertikalen Längen aller Makroblöcke pro Frame errechnet. Dabei geht man von einem MPEG-I Bitstrom in Standard-Auflösung 352*288 aus, d.h. mit 396 Makroblöcken pro Frame. Sei \bar{x} die horizontale und \bar{y} die vertikale Länge des Vektors eines Makroblocks, dann werden mit folgenden Formeln die Summen errechnet:

$$X_{Sum} = \sum_{i=0}^{396} \bar{x}(i), Y_{Sum} = \sum_{i=0}^{396} \bar{y}(i)$$

Um aus diesen beiden Summen die Gesamtrichtung (MD) aller Vektoren eines Frames zu machen, wird mit folgender Formel die Richtung in Grad umgerechnet:

$$MD = \tan^{-1} \left(\frac{Y_{Sum}}{X_{Sum}} \right) * (180 * \pi) + 180$$

MD wird als Richtung einer Bewegung für ein Frame in Grad betrachtet. Es werden die Richtungen der letzten drei Frames berücksichtigt, um an dieser Stelle schon eine eventuell verwackelte Kameraführung herauszufiltern. Jetzt kann jedem Frame eine der Kombinationen der Richtungen *oben*, *unten*, *links* und *rechts* zugeordnet werden.

Nachdem für jedes P-Frame eine Kamerabewegung ermittelt wurde, muß für jeden detektierten Shot entschieden werden, ob es sich um einen Kameraschwenk handelt. Ist dies der Fall, so muß ermittelt werden, in welche Richtung die Kamera schwenkt. Dazu wird die Anzahl der P-Frames pro Shot und die Anzahl der gleichgerichteten Frames benötigt. Ein Kameraschwenk wird detektiert, wenn mindestens 50% der betrachteten Frames eines Shots die gleiche Richtung aufweist. In einem Shot kommt normalerweise nicht mehr als ein Kameraschwenk vor.

Experimentelle Ergebnisse

Die Qualität der Kamerabewegungsanalyse wurde an den gleichen Videos getestet wie die Shotanalyse.

Genre	Erkennungsrate
Sport	95 %
Nachrichten	99 %
Spielfilme	100 %
Cartoons	92 %
Werbung	98 %
Mittlere Genauigkeit	96,8 %

Tabelle 2 Genauigkeit der Kamerabewegungsanalyse

Tabelle 2 gibt die Erkennungsraten der Kamerabewegungsanalyse gegliedert nach den getesteten Genres wieder. Die Qualität der Kamerabewegungsanalyse hängt ganz wesentlich von der Qualität der Shotanalyse ab. Somit sind die Erkennungsraten der Kamerabewegungsanalyse in den Videos besonders hoch, in denen auch die Shotanalyse gute Ergebnisse liefert.

In den Tests hat sich gezeigt, daß die Kamerabewegungsanalyse bei sehr schnellen Kameraschwenks teilweise fehlerhafte Ergebnisse liefert. Ansonsten sind die Ergebnisse auch dort gut, wo die Shotanalyse zu viel Shots erkannt hat.

Mosaicing

Es ist sinnvoll, vor einer Analyse eines Shots die Anzahl der zu analysierenden Informationen möglichst gering zu halten. Somit wird nur eine kleine Auswahl an Frames pro Shot analysiert. Hierfür stellt man den Shot möglichst nur durch einige wenige Bilder dar. Zwei gängige Verfahren für diesen Zweck sind: a) Auswahl von Keyframes - es werden die Frames eines Shots ausgewählt, die diesen möglichst vollständig repräsentieren. Diese Methode ist gut geeignet, wenn in dem Shot keine Kamerabewegungen enthalten sind; b) Erzeugung eines Mosaicbildes (Mosaicing)

- bei Shots mit Kamerabewegungen wird durch Mosaicing ein für diesen Shot repräsentatives Bild erzeugt. Dabei wird dann also der gesamte Shot durch ein einziges Bild repräsentiert.

Mit Hilfe der Kamerabewegungsanalyse wird die Entscheidung getroffen, welche der beiden Verfahren verwendet wird. Das Mosaicing bietet die Möglichkeit mehrere zusammengehörende Bilder, in der Regel die Frames eines Shots, zu einem einzigen Bild zusammenzufassen (Vergleiche [DJK+98] [MP95]).

Algorithmus des Mosaicing

Das DiVA-Mosaicing nutzt, wie auch die anderen beiden vorgestellten Verfahren, Shotanalyse und Kamerabewegungsanalyse, Informationen die im MPEG-I Bitstrom enthalten sind. Zur Bestimmung der Veränderungen zwischen zwei Frames werden die im MPEG-Bitstrom kodierten Motion Vektoren verwendet. Aus den Motion Vektoren lassen sich die Verhältnisse zwischen den Frames sehr gut ableiten, so daß sich die für die Berechnung der Koordinatentransformation notwendigen vier Punkte pro Framepaar [KRA+97] gut bestimmen lassen. Bei der Bestimmung dieser vier Punkte wird das Frame gefiltert, und aus jedem Viertel des Frames wird der am häufigsten auftretende Wert für die Motion Vektoren verwendet.

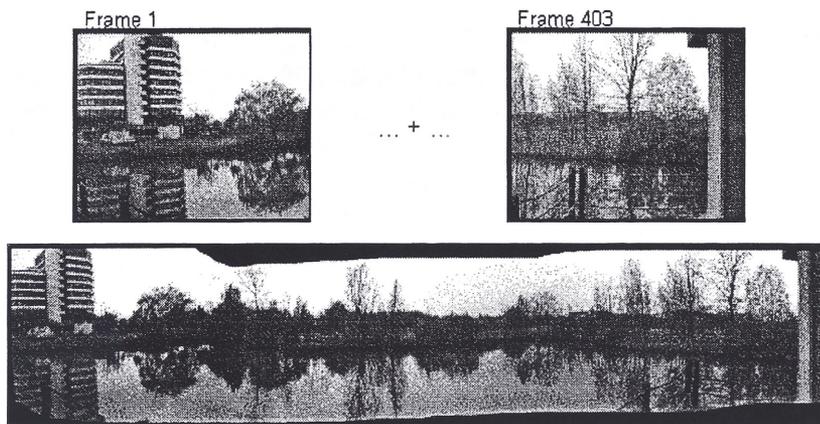


Abbildung 1 Mosaikbild mit den dazugehörigen Grenzframes

Danach kann die Formel der eigenschaftsbasierten Methode zur Berechnung der Koordinatentransformation benutzt werden. Da dabei nicht die speziellen Merkmale der Frames verwendet werden, kann das Korrespondenzproblem vernachlässigt werden. Für die Berechnung der Koordinatentransformation wird das projektive Modell verwendet [JÄH91] [MAN98]. Da nicht für alle Frames Motion Vektoren im MPEG-Bitstrom kodiert und die Motion Vektoren nicht für alle Richtungen definiert sind, werden im DiVA-Mosaicing die Parameter des projektiven Modells nicht für jedes Frame eines Shots berechnet. Entsprechend der Kodierung der Frames [DJK+98] im MPEG-Bitstrom können die Parameter für I-Frames nicht aus den Motion Vektoren berechnet werden. Für P-Frames werden nur die Parameter der Rückrichtung (aus Forward Motion Vektoren) und für B-Frames die Parameter der Hin- und Rückrichtung (Forward und Backward Motion Vektoren) berechnet. Im DiVA-Mosaicing wird zur Wiedergabe im allgemeinen das letzte Frame eines Shots als Referenzframe verwendet.

Aufgrund der oben genannten Gründe werden nicht alle Frames eines Shots für das Erstellen des Mosaikbildes verwendet. Es werden lediglich die I- und P-Frames benutzt. Bei den von uns verwendeten Einstellungen zur Erstellung des MPEG-I Bitstroms ist dies in der Regel jedes dritte Frame [DJK+98]. Als Filter wurden *replace first*, *replace last* und *median* implementiert. Die besten Ergebnisse wurden jedoch bei unseren Tests mit dem *replace last* Filter erzielt.

In der Abbildung 1 ist ein Ergebnis der Arbeit von dem Mosaicing-Tool des DiVA-Systems vorgestellt. Dabei wurde aus einem Videoshot mit 114 Frames (von Frame 1587 bis Frame 1700 eines MPEG-Streams mit insgesamt 6700 Frames) ein Mosaikbild erstellt. Das Mosaikbild wurde auf einem Pentium 133 mit 64 MB RAM in 150 sec. erstellt.

Wissensrepräsentation

In dem Projekt DiVA gliedert sich das zu repräsentierende Wissen zunächst in Informationen zum Bildinhalt einerseits und zum Sachinhalt andererseits. Beide Punkte untergliedern sich jeweils in Bild-, Ton- und Textinhalt eines Berichts.

Diese Vielzahl von Merkmalen müssen in einer Wissensbasis abgelegt und abrufbar sein. Im Idealfall soll eine solche Wissensbasis vom Benutzer wartbar und erweiterbar sein. Dies erreicht man durch die Verwendung bewährter Verfahren der Wissensrepräsentation (zum Beispiel Taxonomien, Baumstrukturen und Vererbung; vgl. [MÜL93] [PUP90]).

Ergebnis einer Sequenzinterpretation ist die Einordnung in ein Themengebiet. Für jedes Gebiet existiert in der Datenbank ein Datensatz, der alle zugehörigen Merkmale (für typische Begriffe, Bildobjekte, Geräusche) enthält, die eine Klassifizierung ermöglichen. Ein Gebiet wird durch eine Datenbankabfrage ermittelt, die alle Datensätze herausfiltert, die den in der Analysephase berechneten Merkmalen entsprechen.

Analog dazu können diese Datensätze in einem regelbasierten System in Form von Einzelregeln abgelegt werden. Mit einer Abfrage, die als Parameter die errechneten Merkmale beinhaltet, kann durch Regelableitungen ebenfalls eine Einordnung in ein Themengebiet erfolgen [DJK+98].

Architektur des DiVA-Systems

Dieser Abschnitt beschäftigt sich mit der Architektur des DiVA-Systems, also der Aufteilung in Module, deren Schnittstellen, die Prozesse und Abhängigkeiten zwischen ihnen sowie die benötigten Ressourcen [MCD91]. Das DiVA-System ist als verteilte Anwendung konzipiert [BS95].

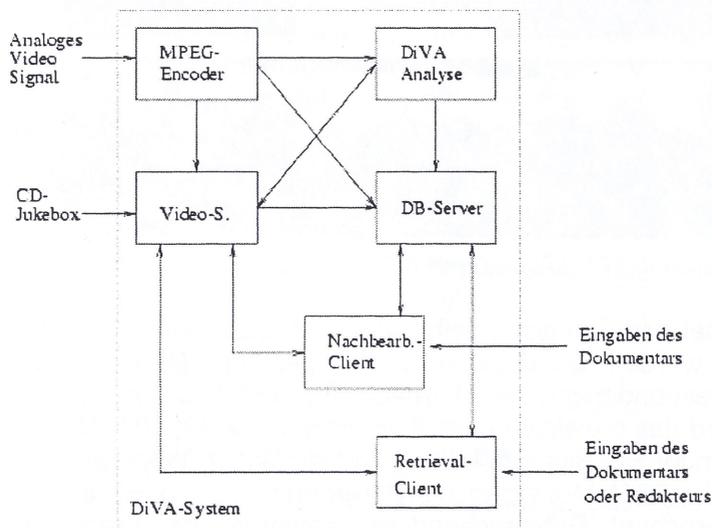


Abbildung 2 DiVA-Module und ihre Schnittstellen (Pfeile)

Eine Aufteilung der Funktionalitäten des DiVA-System auf Module, die sich auf verschiedenen Rechnern eines LANs befinden können, wird in Abbildung 2 illustriert.

Im folgenden werden die Aufgaben dieser Module beschrieben:

MPEG-Encoder: Das analoge Videosignal der zu analysierenden Sendung wird digitalisiert und in einen MPEG-Stream umgewandelt. Nach der Encodierung wird als Referenz auf den Stream eine Archivnummer vergeben, die dem Datenbank-Server und der daraufhin startenden Analyse mitgeteilt wird.

Video-Server: Die erzeugten MPEG-Streams werden auf Massenmedien archiviert. Die Verwaltung der Ablage und des Zugriffs auf die Streams erfolgt über den Video-Server.

Datenbank-Server: Anfragen bezüglich der Datenbank werden von diesem Modul ausgewertet und in eine datenbankspezifische Anfrage umgewandelt. Die Anfrage basiert auf SQL und wird über standardisierte Schnittstellen (ODBC, siehe [GE195]) an das Datenbanksystem abgesetzt.

Nachbearbeitungs-Client: über die grafische Benutzeroberfläche steht dem Dokumentar eine Schnittstelle zu den Informationen einer Sendung zur Verfügung. Die textuellen Beschreibungen

aus der Analyse des Diva-Systems bzw. des Dokumentars lassen sich aus der Datenbank abrufen, verändern sowie ergänzen. Die dazugehörigen Videosequenzen können angezeigt werden.

Retrieval-Client: Dokumentar und Redakteur können per Suchkriterien Anfragen an die Datenbank starten. Die Trefferliste aus einer Anfrage enthält Informationen über das Datum der Sendung sowie den Teil des Textes in der Annotation, in der der gesuchte Begriff vorkommt. Der Benutzer hat die Möglichkeit, sich die korrespondierenden Videosequenzen anzeigen zu lassen.

Analyse: Die Analyse führt eine formale und inhaltliche Klassifikation des MPEG-Streams durch. Die erzeugte textuelle Beschreibung wird in der Datenbank zur Nachbearbeitung durch den Dokumentar abgelegt.

Danksagungen

Diese Arbeit entstand im Rahmen des studentischen Projektes DiVA (Digitale intelligente Video Analyse) am Technologie-Zentrum Informatik der Universität Bremen. Das Projekt wurde in Kooperation mit der Fernsehanstalt Radio Bremen durchgeführt. Besonderer Dank geht an die Mitarbeiter des FS-Archiv und an den betreuenden Hochschullehrer Prof. Dr. O. Herzog.

Literatur

- [BBMR89] R.J. Brachman and A. Borgida and D.L. McGuinness and L.A. Resnick. The CLASSIC knowledge representation system, or KL-ONE: The next generation. In *Workshop on Formal Aspects of Semantic Networks*, Santa Catalina Island, California, February 1989.
- [BS95] U. Borghoff and J. Schlichter. *Rechnergestützte Gruppenarbeit – Eine Einführung in Verteilte Anwendungen*. Springer-Verlag, 1995.
- [GEI95] K. Geiger. *Inside ODBC*. Microsoft Press, 1995.
- [GHS+97] A. Gupta, B. Horowitz, C. Shu, C. Fuller, J. Bach, and M. Gorkani. Virage Video Engine. *SPIE*, 3022:188-198, 1997
- [GJ97] A. Gupta and R. Jain. Visual Information Retrieval. *Communications of the ACM*, 40(5), 1997.
- [ISO93] ISO. Coding of moving pictures and associated audio for digital storage media up to 1.5 mbit/s. MPEG-I Draft, 1993. ISO/IEC 11172-2.
- [JÄH91] B. Jähne. *Digitale Bildverarbeitung*. Springer-Verlag, 1991.
- [KRA+97] J. Kreyß, M. Röper, P. Alshuth und T. Hermes und O. Herzog. Video retrieval by still image analysis with imageminer. *Proc. Of SPIE – The Inter. Soc. For Optical Engineering, Storage and Retrieval for Image and Video Databases V*, 36-44, 1997.
- [KÜS95] H. Küsters. *Bilddatenkomprimierung mit JPEG und MPEG*. Franzis-Verlag GmbH, 1995.
- [MAN98] S. Mann. <http://www.wearcam.org/chirplet.html>. WorldWideWeb, 10.04.1998. Chirplets.
- [MCD91] J.A. McDermid. *Software engineer's reference book*. Butterworth-Heinemann, Oxford, 1991.
- [MP95] S. Mann und R.W. Picard. Video orbits of the projective group: A new perspective on image mosaicing. *M.I.T. Media Laboratory Perceptual Computing Section Technical Report*, 338:239-266, 1995.
- [MÜL93] J. Müller. *Verteilte Künstliche Intelligenz*. BI Wissenschaftsverlag, 1993.
- [PUP90] F. Puppe. *Problemlösung in Expertensystemen*. Studienreihe Informatik. Springer-Verlag, 1990.
- [SIE96] J. Siegel. *CORBA – Fundamentals and Programming*. John Wiley & Sons, Inc., 1996.
- [TEC98] GmbH TecMath. <http://www.foyer.de/euromedia>. WorldWideWeb, 14.05.1998. Synopsis zu dem ESPRIT-Projekt Euromedia.
- [DJK+98] A. Dammeyer, W. Jürgensen, C. Krüwel, E. Poliak, Th. Schäfer, M. Sirava und Th. Hermes. Videoanalyse mit DiVA. In *Beiträge zum Workshop „Inhaltsbezogene Suche von Bildern und Videosequenzen in Digitalen Multimedialen Archiven“ im Rahmen der 22. KI-Jahrestagung*, Bremen, 1998.