



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 600 29 746 T2** 2007.10.18

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 138 151 B1**

(21) Deutsches Aktenzeichen: **600 29 746.2**

(86) PCT-Aktenzeichen: **PCT/EP00/09652**

(96) Europäisches Aktenzeichen: **00 971 311.6**

(87) PCT-Veröffentlichungs-Nr.: **WO 2001/028240**

(86) PCT-Anmeldetag: **02.10.2000**

(87) Veröffentlichungstag
der PCT-Anmeldung: **19.04.2001**

(97) Erstveröffentlichung durch das EPA: **04.10.2001**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **02.08.2006**

(47) Veröffentlichungstag im Patentblatt: **18.10.2007**

(51) Int Cl.⁸: **H04N 7/00** (2006.01)

G06F 17/30 (2006.01)

H04N 7/24 (2006.01)

H04H 9/00 (2006.01)

(30) Unionspriorität:

417288 13.10.1999 US

(73) Patentinhaber:

**Koninklijke Philips Electronics N.V., Eindhoven,
NL**

(74) Vertreter:

Volmer, G., Dipl.-Ing., Pat.-Anw., 52066 Aachen

(84) Benannte Vertragsstaaten:

DE, FR, GB

(72) Erfinder:

**DIMITROVA, Nevenka, NL-5656 AA Eindhoven,
NL; MCGEE, Thomas, NL-5656 AA Eindhoven, NL;
AGNIHOTRI, Lalitha, NL-5656 AA Eindhoven, NL**

(54) Bezeichnung: **AUTOMATISCHE SIGNATURBASIERTE ERKENNUNG, ERLERNUNG UND EXTRAHIERUNG VON
WERBUNG UND ANDEREN VIDEOINHALTEN**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung**ERFINDUNGSGEBIET**

[0001] Die vorliegende Erfindung betrifft allgemein die Videosignalbearbeitung und insbesondere Verfahren für das Bearbeiten von Videosignalen, um Werbesendungen oder andere Arten von Videoinhalt, der spezielle Kenngrößen aufweist, zu identifizieren und zu extrahieren.

HINTERGRUND DER ERFINDUNG

[0002] Es wurden viele unterschiedliche Systeme für das Nachweisen und Extrahieren von Werbesendungen aus gesendeten oder aufgezeichneten Videosignalen entwickelt. Zum Beispiel wird in der US-Patentschrift Nr. 4,782,401 mit dem Titel „Editing Method and Apparatus for Commercials During Video Recording“ eine hardwareorientierte Lösung für das Herausschneiden von Werbesendungen im Analogbereich beschrieben, die auf dem Vorliegen von Schwarz- oder Leer-Frames beruht, welche für das Abgrenzen des Werbesendungen verwendet werden.

[0003] Ein ähnliches System wird in der PCT-Anmeldung Nr. WO 83/00971 mit dem Titel „Reciprocating Recording Method and Apparatus for Editing Commercial Messages from Television Signals“ beschrieben. Dieses System schneidet Werbesendungen basierend auf dem Einblenden und Ausblenden zu Beginn bzw. am Ende einer Werbepause heraus.

[0004] Bei einem anderen Ansatz, der in der US-Patentschrift Nr. 4,750,052 mit dem Titel „Apparatus and Method for Deleting Selected Program Intervals from Recorded Television Broadcasts“ beschrieben ist, wird ein Schwunddetektor verwendet, um Werbesendungen in einem aufgezeichneten Sendeprogramm zu bearbeiten.

[0005] In der PCT-Anmeldung Nr. WO 94/27404 mit dem Titel „Method and Apparatus for Classifying Patterns of Television Programs and Commercials“ werden die Merkmalsextraktion und ein neuronales Netz zur Klassifizierung von Videosignalen verwendet. Das System weist Veränderungen in den Merkmalen, wie z.B. Leistungsamplitude über dem Frequenzspektrum, Farbe und Helligkeit, Vertikalintervall-Zeitcode, Untertitelsignal und Farbträger-Bildsprungsignal, nach.

[0006] In einem System, das in der PCT-Anmeldung Nr. WO 95/06985 mit dem Titel „Process and Device for Detecting Undesirable Video Scenes“ beschrieben ist, wird ein Bild aus einem gesendeten Programm, das einer Werbepause vorausgeht, gespeichert, so dass das Ende der Werbepause mit Hilfe des gespeicherten Bildes aufgefunden werden kann.

Bei diesem Ansatz wird von der Tatsache Gebrauch gemacht, dass Sender oft einen kleinen Teil des Programms nach dem Ende der Werbepause wiederholen.

[0007] In der Europäischen Patentanmeldung Nr. EP 735754 mit dem Titel „Method and Apparatus for the Classification of Television Signals“ wird ein Satz von Merkmalen und zugeordneten Regeln verwendet, um festzustellen, ob die aktuellen Werbesendungen mit einem bestimmtem Maß an „Unschärfe“ dieselben Kriterien erfüllen. Der Satz von Merkmalen, die zur Kennzeichnung von Werbesendungen verwendet werden, enthält z.B. Stereo gegen Mono, Zweikanalton, Schallpegel, Bildhelligkeit und -farbe sowie Logos. Es ist ein umfangreicher Satz von Regeln erforderlich, um die Grenzwerte und Parameterveränderungen für diese Merkmale unterzubringen.

[0008] In der US-Patentschrift Nr. 5,708,477 mit dem Titel „Video Signal Identifier for Controlling a VCR and Television Based on the Occurrence of Commercials“ wird ein Videosignalidentifikator verwendet, um zuvor identifiziertes Werbematerial wiederzuerkennen und es auszusondern, indem entweder das Fernsehtonsignal stumm geschaltet und/oder der Videorekorder angehalten wird, wenn er im Aufzeichnungsbetriebszustand ist. Ein erhebliches Problem bei diesem Ansatz besteht darin, dass er für einen automatischen Nachweis nicht geeignet ist, das heißt, er erfordert gewissermaßen eine Identifizierung des Materials vor seinem Nachweis.

[0009] In einem System, das in der US-Patentschrift Nr. 5,668,917 mit dem Titel „Apparatus and Method for Detection of Unwanted Broadcast Information“ beschrieben ist, werden die ständigen Wiederholungen der Werbesendungen genutzt, um das Werbematerial zu identifizieren. Dieses System speichert Video-Frames in einem komprimierten Format und vergleicht Frames im ursprünglichen „Roh“-Format Pixel für Pixel. Passen die Pixel innerhalb eines bestimmten Grenzwertes zueinander, dann werden die Frames als gleich angesehen. Ein ernsthafter Nachteil dieses Ansatzes ist, dass übermäßig viel Speicher und Rechnerressourcen benötigt werden. Insbesondere wird beim Speichern eines Videos selbst in einem komprimierten Format eine nicht vertretbar große Speicherkapazität, z.B. ungefähr 200 GByte pro Tag für einen Kanal mit hochauflösendem (HDTV) Fernsehinhalt, benötigt. Außerdem ist der Vergleich des Rohvideos sehr zeitaufwändig. Selbst unter der Annahme, dass das Komprimieren und Dekomprimieren von Video ohne zusätzlichen Rechenaufwand ausgeführt werden kann, ist das Vergleichen der Frames ein sehr langsamer Vorgang. Ein gegebenes eintreffendes Frame muss mit der oben erwähnten großen Menge an gespeichertem Videomaterial verglichen und der Vergleich abgeschlossen werden, bevor das nächste Frame eintrifft.

[0010] Der Konferenzbeitrag „On the detection and recognition of television commercials“ von R. Lienhart u.a. (Proceedings IEEE International Conference on Multimedia Computing and Systems '97, Ottawa, Ontario, Kanada, 3.– 6. Juni 1997, S. 509–516, ISBN 0-8186-7819-4) offenbart ein System zum Erkennen von Werbesendungen, das einen auf Merkmalen basierenden Ansatz mit einem Ansatz kombiniert, der Signaturen verwendet, um bekannte Werbesendungen zu erkennen. In einer Ausführungsform des Systems wird dann, wenn ein neuer Spot entdeckt wird, dieser neue Spot in einer ersten Liste gespeichert. Der Clip wird nur vorläufig in die Liste eingetragen. Er wird permanent gespeichert, wenn er auch in anderen Werbeblöcken aufgefunden wird. Dieser Ansatz entscheidet noch nicht genau genug zwischen Werbesendungen und anderen aufgefundenen Spots.

[0011] Aus den obigen Darlegungen wird ersichtlich, dass es einen Bedarf an verbesserten Verfahren zum Erkennen und Extrahieren von Werbesendungen und anderen Typen von Videoinhalt gibt, welche die Probleme umgehen, die mit den oben beschriebenen herkömmlichen Systemen verbunden sind.

KURZDARSTELLUNG DER ERFINDUNG

[0012] Die Erfindung verschafft verbesserte Verfahren für das Auffinden, Erkennen und Extrahieren von Werbesendungen oder anderen speziellen Typen von Videoinhalt in einem Videosignal. Erfindungsgemäß wird ein Videosignal bearbeitet, um Segmente zu erkennen, die wahrscheinlich mit einer Werbesendung oder einem anderen speziellen Typ von Videoinhalt zu verbinden sind. Aus jedem der so identifizierten Segmente wird eine Signatur extrahiert und die extrahierten Signaturen werden – möglicherweise zusammen mit zusätzlichen zeitlichen und kontextualen Informationen – verwendet, um zu bestimmen, welches der identifizierten Segmente tatsächlich mit dem speziellen Typ von Videoinhalt verbunden ist. Die zeitlichen Informationen können z.B. eine Angabe über die Zeitspanne einschließen, die zwischen einer gegebenen Signatur und einer dazu passenden Signatur aus einem früheren Segment des Videosignals verstrichen ist. Die kontextualen Informationen können z.B. Programminformationen, wie z.B. Programmnamen, Kanal, Sendeplatz und Einstufung, einschließen, wie sie von einem elektronischen Programmführer oder einer anderen Informationsquelle erhalten werden.

[0013] Eine oder mehrere der extrahierten Signaturen kann z.B. eine visuelle Frame-Signatur sein, die mindestens teilweise auf einer visuellen Kenngröße eines Frames oder eines Videosegments beruht, wie sie unter Verwendung von Informationen bestimmt wird, die auf DC- und Bewegungskoeffizienten des Frames oder auf DC- und AC-Koeffizienten des Fra-

mes basieren. Andere Verfahren zum Extrahieren einer visuellen Frame-Signatur können zumindest teilweise auf Farbhistogrammen basieren. Als ein weiteres Beispiel kann eine gegebene extrahierte Signatur eine Audiosignatur sein, die zumindest teilweise auf einer Kenngröße eines Audiosignals basiert, das mindestens einem Teil des Videosegments zugeordnet ist. Andere erfindungsgemäße Signaturen enthalten z.B. einen Untertiteltext, der ein in der Werbung auftretendes Produkt oder eine Dienstleistung bezeichnet, eine Frame-Nummer plus Informationen aus einem Unterbild von identifiziertem, dem Frame zugeordneten Text, wie z.B. eine 800-er Nummer, ein Name eines Unternehmens, ein Name eines Produkts oder einer Dienstleistung, eine Internetadresse (URL) usw., oder eine Frame-Nummer sowie eine Position und Größe eines Gesichts oder anderen Objekts im Bild, wie es durch eine geeignete Bounding-Box identifiziert wird, wie auch verschiedene Kombinationen von diesen und anderen Signaturtypen.

[0014] Nach einer anderen Ausbildung der Erfindung hält ein Videobearbeitungssystem verschiedene Sätze von Signaturlisten bereit, wobei die Sätze von Listen eine oder mehrere aus einem Satz wahrscheinlicher Listen, einem Satz von Kandidatenlisten und einem Satz von Fundlisten einschließt, wobei jeder Eintrag in eine gegebene Liste einer Signatur entspricht, die mit einem speziellen Videosegment verbunden ist. Die Sätze von Listen werden aktualisiert, wenn die verschiedenen extrahierten Signaturen bearbeitet werden. Zum Beispiel wird eine gegebene Signatur, von der erkannt wurde, dass sie wahrscheinlich mit dem speziellen Videoinhalt verbunden ist, zunächst auf eine der wahrscheinlichen Listen gesetzt, wenn sie nicht bereits zu einer beliebigen Signatur in einer der wahrscheinlichen Listen passt. Wenn die gegebene Signatur zu einer Signatur passt, die bereits auf einer der wahrscheinlichen Listen ist, dann wird die gegebene Signatur auf eine der Kandidatenlisten gesetzt. Eine gegebene Signatur auf einer Kandidatenliste wird auf eine Fundliste umgesetzt, wenn sie zu einer Signatur passt, die bereits auf einer der Kandidatenlisten ist. Eine gegebene Signatur kann auch in dem Falle, dass die Signatur in einem vorgesehenen Zeitraum nicht wiederholt wird, aus einer oder aus mehreren Listen entfernt werden.

[0015] Nach einer anderen Ausbildung der Erfindung kann das System eingerichtet sein, einen Benutzer in den Prozess des Auffindens, Erkennens und Extrahierens von Werbesendungen einzubeziehen. Zum Beispiel kann eine Benutzer-Fernbedienung zum Einsatz beim Fernsehen, an einer Set-Top-Box oder an einem anderen Videobearbeitungssystem so eingerichtet sein, dass sie eine „Nie-wieder“-Schaltfläche derart enthält, dass beim Drücken der Schaltfläche durch den Benutzer die Werbesendungssignatur automatisch extrahiert und

direkt in einer speziellen Fundliste gespeichert wird, ohne dass erst die oben erwähnten wahrscheinlichen und Kandidatenlisten durchlaufen werden.

[0016] Nach noch einer weiteren Ausbildung der Erfindung können spezielle Aktivitäten des Benutzers festgestellt und genutzt werden, um das automatische Extrahieren einer Signatur aus einem gegebenen Segment eines Videosignals zu veranlassen. Zum Beispiel kann das System eingerichtet sein, automatisch eine Signatur aus einem Teilbereich eines Videosignals zu extrahieren, den ein Benutzer im Schnellvorlauf durchläuft, wenn er eine Wiedergabe einer zuvor aufgenommenen Sendung anschaut.

[0017] Vorteilhaft ist, dass es durch die Erfindung möglich wird, Werbesendungen und andere Typen von Videoinhalt mit einem bedeutend geringeren Aufwand im Vergleich zu den oben erwähnten herkömmlichen Systemen zu identifizieren, durch das System zu erkennen und zu extrahieren. Insbesondere verringert die Erfindung durch die Verwendung von extrahierten Signaturen den Umfang des Speichers und der Rechnerressourcen, die für das Ausführen einer Identifizierung und Extraktion von Videoinhalt erforderlich sind. Diese und andere Merkmale und Vorzüge der vorliegenden Erfindung werden aus den beigefügten Zeichnungen und der folgenden ausführlichen Beschreibung klarer ersichtlich.

KURZBESCHREIBUNG DER ZEICHNUNGEN

[0018] [Fig. 1](#) ist ein Blockdiagramm eines Videobearbeitungssystems, in dem die Erfindung ausgeführt werden kann.

[0019] Die [Fig. 2](#) und [Fig. 3](#) sind Flussdiagramme der Erkennungs- bzw. Abstimmungsprozesse gemäß einer veranschaulichenden Ausführungsform der Erfindung, die in dem Videobearbeitungssystem von [Fig. 1](#) ausgeführt werden kann.

AUSFÜHRLICHE BESCHREIBUNG DER ERFINDUNG

[0020] [Fig. 1](#) zeigt ein Videobearbeitungssystem **10**, in dem Verfahren für das Auffinden, Erkennen und Extrahieren von Video gemäß der vorliegenden Erfindung ausgeführt werden können. Das System **10** kann ein Fernsehgerät, eine Set-Top-Box, ein Desktop, ein Laptop- oder Palmtop-Rechner, ein Personaldigitalassistent (PDA), ein Videospeichergerät, wie z.B. ein Kassettenvideorekorder (VCR), ein Digitalvideorekorder (DVR), ein TiVo-Gerät usw., sowie auch Teile oder Kombinationen dieser und anderer Geräte sein. Das System **10** enthält eine oder mehrere Videoquellen **12**, ein oder mehrere Eingabe/Ausgabegeräte **14**, einen Prozessor **15** und einen Speicher **16**. Die Videoquelle(n) **12** kann (können) z.B. ein Fernsehempfänger, ein Videorekorder oder ein ande-

res Videospeichergerät sein, oder es kann ein beliebiger andere Typ einer Videoquelle sein. Alternativ kann (können) die Quelle(n) **12** eine oder mehrere Netzverbindungen für den Videoempfang aus einem Server oder aus Servern z.B. eines globalen Rechnerkommunikationsnetzes, wie z.B. des Internets, eines Fernnetzes, eines innerstädtischen Netzes, eines lokalen Netzes, eines terrestrischen Rundfunksystems, eines Kabelnetzes, eines Satellitennetzes, eines Funknetzes oder eines Telefonnetzes, sowie auch Teile oder Kombinationen von diesen oder anderen Typen von Netzen sein.

[0021] Die Eingabe/Ausgabegeräte **14**, der Prozessor **15** und der Speicher **16** kommunizieren über ein Kommunikationsmedium **17**. Das Kommunikationsmedium **17** kann z.B. ein Bus, ein Kommunikationsnetz, eine oder mehrere innere Verbindungen einer Schaltung, eine Chipkarte oder ein anderes Gerät sowie auch Teile und Kombinationen dieser und anderer Kommunikationsmedien sein. Das Eingangsvideo aus der (den) Quelle(n) **12** wird entsprechend einem oder mehreren Softwareprogrammen bearbeitet, die im Speicher **16** gespeichert sind, und durch den Prozessor **15** ausgeführt, um ein Ausgangsvideo zu erzeugen, das einem Anzeigegerät **18** zugeführt wird, welches z.B. ein Fernsehbildschirm, ein Rechnermonitor usw. sein kann.

[0022] Es sollte verständlich sein, dass die spezielle Konfiguration des Systems **10**, wie sie in [Fig. 1](#) dargestellt ist, lediglich als ein Beispiel dient. Fachleute werden einsehen, dass die Erfindung mit einer großen Vielfalt alternativer Systemkonfigurationen ausgeführt werden kann.

[0023] Die vorliegende Erfindung schafft verbesserte Verfahren für das Auffinden, Erkennen und Extrahieren von Werbesendungen oder anderen Typen von Videosegmenten.

[0024] Es ist von Vorteil, dass die erfindungsgemäßen Verfahren ohne übermäßig hohe Anforderungen an Speicher und Rechner ausgeführt werden können.

[0025] Es folgt ein Beispiel eines erfindungsgemäßen Prozesses zum Auffinden, Erkennen und Extrahieren von Werbesendungen aus einem Rundfunk-Videosignal. Für dieses Beispiel wird angenommen, dass das Eingangsvideo ein Rundfunk-Videosignal einschließlich mindestens eines Programms und mehrerer Werbesendungen enthält.

[0026] Schritt 0. Bei Vorliegen eines Eingangsvideos Ausführen der nachfolgenden Schritte 1 bis 5:

[0027] Schritt 1. Auffinden von ungewöhnlich ablaufenden Segmenten in dem Rundfunk-Videosignal. Das kann z.B. das Auffinden eines Bereichs im Rund-

funk-Videosignal mit einer hohen Schnittrate oder das Auffinden eines Bereiches mit einer hohen Textaktivität einschließen. Andere Beispiele schließen den Nachweis einer schnellen Veränderung im visuellen Bereich durch das Ausbilden von Farbhistogrammen, den Nachweis eines Anstieges im Tonpegel oder den Nachweis schneller Änderungen im Tonbereich von Musik auf Sprache, von einem Rhythmus auf einen anderen usw. ein.

[0028] Schritt 2. Die Segmente, bei denen im Schritt 2 erkannt wurde, dass sie eine ungewöhnlichen Aktivität enthalten, werden weiter bearbeitet, um festzustellen, ob sie wahrscheinlich mit einer Werbesendung verbunden sind. Die so bestimmten Segmente werden dann markiert. Beispiele für Merkmale, die verwendet werden, um diese Feststellung zu treffen, schließen ein:

- (a) angezeigten Text, der zu Eintragungen in eine gespeicherte Textdatei mit Namen bekannter Unternehmen, Namen von Produkten oder Dienstleistungen, 800-er Nummern oder anderen Telefonnummern, Internetadressen (URL) usw. gehört, die mit Werbesendungen verbunden sind;
- (b) Sprache. In diesem Falle kann die Sprache extrahiert, in Text umgewandelt und der sich ergebende Text mit Bezugnahme auf die oben erwähnte gespeicherte Textdatei analysiert werden, um Namen bekannter Unternehmen, Namen von Produkten oder Dienstleistungen, 800-er Nummern und andere Telefonnummern, Internetadressen usw. aufzufinden;
- (c) Fehlen von Untertitelinformationen verbunden mit einer hohen Schnittrate;
- (d) Untertitelinformationen, die mehrere Leerzeilen enthalten;
- (e) Abschluss eines Abspanns für einen Film, eine Show oder ein anderes Programm;
- (f) mittlerer Keyframe-Abstand oder Trend des mittleren Schnitt-Frame-Abstandes, z.B. ein zunehmender oder ein abnehmender Trend;
- (g) Fehlen von Logos, z.B. eingeblendeten Videologos, welche den Fernsehsender kennzeichnen;
- (h) Unterschiedliche Schrifttypen, Größen und Farben für eingeblendeten Text;
- (i) schnelle Veränderungen in der Farbpalette oder in anderen Farbkenngrößen.

[0029] Die Signaturen werden dann aus den Keyframes in den markierten Segmenten extrahiert und in einer speziellen „wahrscheinlichen“ Signaturliste angeordnet. Der Begriff „Keyframe“, wie er hier verwendet wird, bezeichnet allgemein ein oder mehrere Frames, die mit einer gegebenen Kameraeinstellung oder einem anderen Teil eines Videosignals, einem ersten Frame in einer speziellen Kameraeinstellung, verbunden sind. Beispiele für wahrscheinliche Signaturlisten schließen die Listen L1, Li, Ln usw. ein, auf die im Flussdiagramm von [Fig. 2](#) verwiesen wird. Bei einem ersten Durchlauf durch den Schritt 2 wird eine

gegebene Liste von den wahrscheinlichen Listen im Allgemeinen Signaturen für mehrere Werbesendungen sowie auch für Teile des Programms enthalten.

[0030] Eine gegebene Signatur kann z.B. auf einer visuellen Frame-Signatur oder einer Audiosignatur oder anderen geeigneten identifizierenden Kenngrößen beruhen. Eine visuelle Frame-Signatur kann z.B. unter Verwendung eines Extraktionsverfahrens, das auf DC- und AC-Koeffizienten (DC+AC) basiert, eines Extraktionsverfahrens, das auf DC- und Bewegungskoeffizienten (DC+M) basiert oder anderer geeigneter Extraktionsverfahren, z.B. von Verfahren, die auf Wavelets und anderen Transformationen beruhen, extrahiert werden.

[0031] Das oben erwähnte DC+AC-Verfahren wird ausführlicher z.B. von N. Dimitrova, T. McGee und H. Elenbaas in „Video Keyframe Extraction and Filtering: A Keyframe is not a Keyframe to Everyone“, Proceedings of ACM Conference on Information and Knowledge Management, Nov. 1997, beschrieben, und es kann zum Erzeugen einer visuellen Frame-Signatur verwendet werden, die z.B. einen DC-Koeffizienten und fünf AC-Koeffizienten umfasst.

[0032] Als ein weiteres Beispiel kann das oben erwähnte DC+M-Verfahren verwendet werden, um Signaturen der Form (Keyframe1, Signatur1, Keyframe2, Signatur2 usw.) zu erzeugen. Dieses DC+M-Extraktionsverfahren wird z.B. ausführlicher beschrieben in der US-Patentschrift Nr. 5,870,754 mit dem Titel „Video Retrieval of MPEG Compressed Sequences Using DC and Motion Signature“, die am 9. Februar 1999 im Namen der Erfinder N. Dimitrova und M. Abdel-Mottaleb ausgefertigt wurde, und in „Content-Based Video Retrieval By Example Video Clip“, Proceeding of Storage and Retrieval for Image and Video Databases V, SPIE Bd. 3022, S. 59–70, San Jose, Ca, 1997 von N. Dimitrova und M. Abdel-Mottaleb.

[0033] Andere visuelle Frame-Signatur-Extraktionsverfahren können zumindest teilweise auf Farbhistogrammen aufgebaut werden, wie es z.B. in „Color Super-histograms for Video Representation“, IEEE International Conference on Image Processing, Kobe, Japan, 1999, von N. Dimitrova, J. Martino, L. Agnihotri und H. Elenbaas beschrieben ist.

[0034] Eine Audiosignatur Ai kann solche Informationen, wie z.B. die Tonhöhe (z.B. Maximum, Minimum, Median, Mittelwert, Anzahl der Spitzenwerte usw.), mittlere Amplitude, mittlere Energie, Bandbreite und Spitzenwerte der Mel-Frequenz-Cepstrumkoeffizienten (MFCC), enthalten. Ein solche Signatur kann in der Form z.B. eines einzelnen Objekts A1 vorliegen, das in den ersten 5 Sekunden aus einer Werbesendung extrahiert wird. Als ein weiteres Beispiel könnte die Audiosignatur ein Satz von Audiosi-

gnaturen {A1, A2, ... An} sein, die z.B. in einer gekennzeichneten Zeitspanne, die auf jeden identifizierten Schnitt folgt, extrahiert werden.

[0035] Die Erfindung kann auch zahlreiche andere Typen von Signaturen verwenden. Zum Beispiel kann ein weiterer Signaturtyp in der Form eines Untertiteltextes vorliegen, der ein Produkt oder eine Dienstleistung, für die geworben wird, beschreibt. Als ein weiteres Beispiel könnte die Signatur in der Form einer Frame-Nummer plus Informationen aus einem Unterbild von identifiziertem, dem Frame zugeordneten Text vorliegen, wie z.B. eine 800-er Nummer, ein Name eines Unternehmens, ein Name eines Produkts oder einer Dienstleistung, eine Internetadresse usw. Als noch ein weiteres Beispiel könnte die Signatur eine Frame-Nummer und eine Lage sowie Größe eines Gesichts oder anderen Objekts in dem Bild sein, wie es durch eine geeignete Bounding-Box identifiziert wird. Es könnten auch verschiedene Kombinationen von diesen und anderen Signaturtypen verwendet werden.

[0036] Schritt 3. Beim Auffinden eines neuen möglichen Werbesendungssegmentes wird jedes Mal die Signatur dieses Segments mit den anderen Signaturen in den wahrscheinlichen Listen verglichen, wie es ausführlicher in Zusammenhang mit dem Flussdiagramm von [Fig. 2](#) beschrieben ist. Passt die neue Signatur nicht zu irgendeiner Signatur, die bereits auf einer der wahrscheinlichen Listen ist, dann wird die neue Signatur zu einer wahrscheinlichen Liste hinzugefügt. Passt die neue Signatur zu einer oder mehreren Signaturen auf einer der wahrscheinlichen Listen, dann werden die eine oder die mehreren Signaturen in einer speziellen „Kandidaten“-Liste von Signaturen angeordnet. Beispiele für Kandidatenlisten von Signaturen schließen die Listen C1, Cj, Cm usw. ein, auf die im Flussdiagramm von [Fig. 2](#) Bezug genommen wird.

[0037] Es sollte angemerkt werden, dass dann, wenn eine neue Signatur nicht ähnlich zu einer beliebigen Signatur für ein Segment ist, welches mehr als etwa 30 Sekunden oder weniger als etwa 10 Minuten in der Zeit zurückliegt, aber ähnlich zu einer Signatur für ein Segment ist, das etwa 10–13 Minuten in der Zeit zurückliegt, eine erhöhte Wahrscheinlichkeit dafür vorliegt, dass es Teil einer Werbesendung sein kann. Mit anderen Worten, diese Zeitrelation zwischen ähnlichen Signaturen spiegelt die Tatsache wieder, dass eine gegebene wahrscheinliche Liste Werbesendungssegmente enthalten kann, die eine vorgesehene ungefähre Zeitspanne, z.B. 10 Minuten, auseinander liegen. Diese Zeitabstandsrelation kann experimentell für verschiedene Typen von Programmen, Rundfunksendezeiten, Ländern usw. bestimmt werden.

[0038] Im Vergleichsprozess können andere Typen

von zeitlichen und kontextualen Informationen berücksichtigt werden. Tritt zum Beispiel eine spezielle Signatur an einem Tag in etwa dem gleichen Zeitabschnitt auf wie am vorhergehenden Tag, dann kann es wahrscheinlicher sein, dass sie mit einer Werbesendung verbunden ist. Die Listen könnten auch in verschiedene Gruppen für unterschiedliche Tage, Zeiten oder Kanalplätze unterteilt werden, um so den Vergleichsprozess zu erleichtern. Zum Beispiel sollten Shows für Kinder im Allgemeinen auf Sendeplätzen am frühen Morgen laufen, und sie würden mit einer größeren Wahrscheinlichkeit andere Werbesendungen aufweisen als ein Abendprogramm, wie z.B. Fußball am Montagabend. Zur Bereitstellung dieser und anderer Informationen könnte ein im System **10** vorgesehener elektronischer Programmführer (EPG) verwendet werden. Zum Beispiel könnte eine Signatur mit einem speziellen Shownamen und einer Einstufung verknüpft werden, was eine Anordnung wie z.B. (Showname, Einstufung, Kanal, Keyframe1, Signatur, Keyframe5, Signatur usw.) ergibt. Es könnten auch Programmkategorieinformationen aus dem EPG verwendet werden, um zur Identifizierung von Werbesendungen in den Listen beizutragen.

[0039] Schritt 4. Wenn ein neues mögliches Werbesendungssegment erkannt wird, dann wird jedes Mal die Signatur dieses Segments auch mit den Signaturen in den oben erwähnten Kandidatenlisten verglichen. Passt die neue Signatur zu einer Signatur in einer der Kandidatenlisten, dann wird die neue Signatur in eine spezielle Liste der „aufgefundenen Werbesendungen“ übertragen, die hier auch als eine permanente Liste bezeichnet wird. Beispiele für Listen aufgefundener Werbesendungen sind die Listen P1 und Pk, auf die im Flussdiagramm von [Fig. 2](#) Bezug genommen wird.

[0040] Schritt 5. Gibt es in einer gegebenen Liste aufgefundener Werbesendungen mindestens eine Signatur, dann wird die Signatur eines beliebigen neuen möglichen Werbesendungssegments zuerst mit der(den) Signatur(en) in dieser Liste verglichen. Wird eine Übereinstimmung gefunden, dann wird ein Werbesendungsfrequenzzähler, welcher der entsprechenden Signatur zugeordnet ist, um eins heraufgesetzt. Gibt es keine Übereinstimmung mit einer Signatur in einer Liste aufgefundener Werbesendungen, dann wird die neue Signatur mit den Signaturen in einer oder mehreren Kandidatenlisten verglichen. Wird für die neue Signatur eine Übereinstimmung in einer gegebenen Kandidatenliste gefunden, dann wird die neue Signatur wie beim Schritt 4 in die Liste der aufgefundenen Werbesendungen eingesetzt. Gibt es keine Übereinstimmung mit irgendeiner Signatur in der Kandidatenliste, dann wird die neue Signatur in eine der wahrscheinlichen Listen gesetzt. Eine ausführlichere Beschreibung dieses Prozesses wird im Flussdiagramm von [Fig. 2](#) gegeben.

[0041] Der oben erwähnte Zähler für die Signaturen in einer Liste aufgefundener Werbesendungen kann überwacht werden, um festzustellen, wie oft er heraufgesetzt wird, und die Ergebnisse können verwendet werden, um weitere Informationen zur Identifizierung von Werbesendungen zu erhalten. Wird der Zähler zum Beispiel in einer verhältnismäßig kurzen Zeitspanne, in der Größenordnung von etwa 1–5 Minuten, heraufgesetzt, dann ist es wahrscheinlich keine Werbesendung. Wird als ein anderes Beispiel der Zähler über einen sehr langen Zeitraum hinweg, z.B. in der Größenordnung von einer Woche oder mehr, nicht heraufgesetzt, dann kann der Zähler heruntergesetzt werden, so dass die Werbesendung schließlich vom System "vergessen" wird. Dieser Typ von Zeitrelationspolitik kann auch für die Signaturen in den oben erwähnten wahrscheinlichen Listen durchgeführt werden.

[0042] Der Signaturzähler kann auch in anderen Anwendungen eingesetzt werden; zum Beispiel kann er durch die Werbefirmen verwendet werden, um festzustellen, wie oft eine gegebene Werbesendung ausgestrahlt wurde.

[0043] [Fig. 2](#) und [Fig. 3](#) zeigen detailliertere Beispiele der Erkennungs- bzw. Anpassungsvorgänge, die in dem oben beschriebenen Prozess zum Auffinden, Erkennen und Extrahieren von Werbesendungen verwendet werden. Es sollte angemerkt werden, dass diese veranschaulichende Ausführungsform zwar auf der Verwendung visueller Signaturen aufbaut, die beschriebenen Verfahren aber auch auf andere Typen von Signaturen, einschließlich der Audio-Signaturen, anwendbar sind. Mit Bezugnahme auf [Fig. 2](#) initialisiert ein Initialisierungsschritt **100** eine erste wahrscheinliche Liste L1, eine erste Kandidatenliste C1 und eine erste permanente Liste P1 zu Leermengen und die Variablen i, j und k auf null. Der Initialisierungsschritt **100** setzt auch die Variablen „vorhergehende Signatur“ und „aktuelle Signatur“ auf null.

[0044] Im Schritt **102** wird das Eingangsvideo bearbeitet, um ein Keyframe zu erkennen, eine Schnittrate zu berechnen und ein Schwarz-Frame in Verbindung mit einer Tonpause zu erkennen. Im Schritt **103** wird die vorhergehende Signatur gleich der aktuellen Signatur gesetzt. Im Schritt **104** wird bestimmt, ob nach einem Schwarz-Frame mit Tonpause eine hohe Schnittrate vorliegt. Ist das so, dann wird eine visuelle Frame-Signatur extrahiert, die aktuelle Signatur wird gleich der extrahierten Signatur gesetzt, und der Index i wird um eins heraufgesetzt, wie in Schritt **106** dargestellt ist. Die aktuelle Signatur wird dann im Schritt **108** in der wahrscheinlichen Liste Li gespeichert. Liegt nach einem Schwarz-Frame mit Tonpause keine hohe Schnittrate vor, dann wird der Schritt **102** so lange wiederholt, bis eine solche Bedingung nachgewiesen wird. Die Bearbeitungsvorgänge von

Schritt **102** werden so lange fortgesetzt, wie Eingangsvideo vorliegt.

[0045] Wie zuvor erwähnt wurde, könnten im Schritt **102** zusammen mit oder anstelle des Nachweises von Schnittrate und Schwarz-Frames plus Tonpause viele andere Verfahren verwendet werden, um Bereiche des Eingangsvideos zu identifizieren, welche wahrscheinlich Werbesendungen zuzuordnen sind.

[0046] Ist der Index i größer als eins, dann wird im Schritt **110** versucht, die aktuelle Signatur mit den Signaturen, die in den wahrscheinlichen Listen L1, ... Li-1 gespeichert sind, abzugleichen. Wird keine Übereinstimmung gefunden, dann werden im Schritt **112** aus den wahrscheinlichen Listen L1 bis Li die alten Eintragungen gelöscht, und der Prozess kehrt zum Schritt **102** zurück. Die alten Eintragungen können als solche Eintragungen identifiziert werden, welche in den wahrscheinlichen Listen länger als eine bestimmte vorgegebene Zeitspanne vorgelegen haben. Wird im Schritt **110** eine Übereinstimmung zwischen der aktuellen Signatur, die in Li gespeichert ist, und einer weiteren Signatur, z.B. einer in Ln gespeicherten Signatur, gefunden, dann wird im Schritt **114** bestimmt, ob der Index j größer als null ist. Ist j nicht größer als null, dann wird im Schritt **116** der Index j heraufgesetzt, und dann wird die aktuelle Signatur in der Kandidatenliste Cj gespeichert, wie im Schritt **118** angezeigt ist. Aus den Kandidatenlisten C1 bis Cj werden dann im Schritt **120** die alten Eintragungen gelöscht, und der Prozess kehrt zum Schritt **102** zurück, um zu versuchen, eine weitere hohe Schnittrate nach einem Schwarz-Frame mit Tonpause nachzuweisen.

[0047] Wenn im Schritt **114** festgestellt wird, dass der Index j größer als null ist, dann wird im Schritt **122** versucht, die aktuelle Signatur mit den Kandidatenlisten C1, ... Cj abzugleichen. Wird im Schritt **122** keine Übereinstimmung gefunden, dann wird im Schritt **124** versucht, die vorhergehende Signatur mit der letzten Signatur in Cj abzugleichen. Ergibt der Schritt **124** keine Übereinstimmung, dann werden die Schritte **116**, **118** und **120** ausgeführt, und der Prozess kehrt zum Schritt **102** zurück. Ergibt der Schritt **124** eine Übereinstimmung, dann werden die Schritte **118** und **120** ausgeführt, und der Prozess kehrt dann zum Schritt **102** zurück.

[0048] Wird im Schritt **122** eine Übereinstimmung zwischen der in Li gespeicherten visuellen Signatur und einer Signatur in einer Kandidatenliste, z.B. einer in Cm gespeicherten Signatur, gefunden, dann wird im Schritt **126** die Kandidateneinstufung für die in der Kandidatenliste Cm gespeicherte Signatur Si erhöht. Im Schritt **128** wird dann bestimmt, ob die Einstufung für irgendeine Signatur in Cm über einem festgelegten Grenzwert liegt. Wenn nicht, dann kehrt der Prozess zum Schritt **102** zurück. Liegt eine Einstufung

für eine oder mehrere Signaturen in Cm über dem Grenzwert, dann wird der Index k im Schritt **130** heraufgesetzt, die Signaturen) in Cm wird(werden) aus dieser Kandidatenliste entfernt und in der permanenten Liste Pk gespeichert, und die Kandidatenlisten werden dann bis hin zu Cj umgeordnet, wie in Schritt **132** dargestellt ist, wonach der Prozess zum Schritt **102** zurückkehrt.

[0049] [Fig. 3](#) zeigt ausführlicher einen veranschaulichenden Prozess zum Abgleich einer aktuellen Signatur mit Signaturen in den Kandidatenlisten C1, ... Cm, der weitgehend dem Prozessschritt **122** von [Fig. 2](#) entspricht. Der Abgleichvorgang von [Fig. 3](#) ist dazu bestimmt, eine Übereinstimmung zwischen der aktuellen Signatur und z.B. einer mittleren Signatur in einem Signaturesatz in einer gegebenen Kandidatenliste zu orten. Im Schritt **150** wird eine Übereinstimmung zwischen der aktuellen Signatur und einer zuvor gespeicherten Signatur Smp in der Liste Cm = {Sm1, Sm2, ..., Smp, ... Smt} gefunden. Im Schritt **152** wird dann bestimmt, ob die Variablen "begin" und "end" alle gleich null sind. Diese Variablen werden verwendet, um den Beginn bzw. das Ende einer einzelnen Werbesendung zu markieren, wenn mehrere Werbesendungen in einer gegebenen Kandidatenliste vorgefunden werden, und sie werden in dem oben erwähnten Initialisierungsschritt **100** zu null initialisiert. Wenn entweder begin oder end ungleich Null sind, dann werden beide auf einen Wert p gesetzt, wie in den Schritten **154** und **156** angegeben ist, wobei p der Indexwert ist, welcher der Signatur Smp zugeordnet ist. Wenn sowohl begin als auch end gleich null sind, dann behält begin seinen Wert null bei, und der Wert von end wird im Schritt **156** auf p gesetzt.

[0050] Im Schritt **158** wird dann die Differenz zwischen begin und end mit einem vorgegebenen Grenzwert verglichen. Ist die Differenz end-begin größer als der Grenzwert und wird im Schritt **160** festgelegt, dass begin gleich eins und end gleich einer vorgesehenen Zeit t ist, dann wird der Index j im Schritt **162** heraufgesetzt. Im Schritt **164** werden dann die Signaturen Smbegin, ... Smend in die Kandidatenliste Cj kopiert und diese Signaturen aus der Kandidatenliste Cj-1 entfernt. Im Schritt **166** werden dann sowohl end als auch begin auf null zurückgesetzt, und der Prozess von [Fig. 3](#) endet und kehrt zu dem entsprechenden Punkt, d.h. dem Schritt **122**, in dem Prozess von [Fig. 2](#) zurück. Wenn im Schritt **158** festgestellt wird, dass end-begin kleiner als der Grenzwert ist, oder wenn im Schritt **160** begin nicht gleich eins oder end nicht gleich t ist, dann endet der Prozess in [Fig. 3](#) und kehrt zum Schritt **122** von [Fig. 2](#) zurück.

[0051] Sobald die Werbesendungen oder andere Typen von Videoinhalt unter Verwendung der oben beschriebenen Verfahren identifiziert und durch das System erkannt wurden, können herkömmliche An-

sätze für das Bearbeiten oder anderenfalls für das Extrahieren der Werbesendungen, z.B. durch Stummschalten eines Fernsehgeräts, Schnellvorlauf durch eine aufgezeichnete Sendung usw., direkt verwendet werden.

[0052] Andere Ausführungsformen können den Benutzer in den Prozess des Auffindens, Erkennens und Extrahierens von Werbesendungen einzubeziehen. Zum Beispiel kann eine Benutzer-Fernbedienung für den Einsatz in dem oben beschriebenen System so eingerichtet sein, dass sie eine „Nie-wieder“-Schaltfläche enthält, die bei Betätigung durch den Benutzer die Werbesendungssignatur automatisch extrahiert und direkt in einer bestimmten P-Liste speichert, ohne dass erst die L- und die C-Listen durchlaufen werden. In anderen Ausführungsformen können spezielle Aktivitäten des Benutzers nachgewiesen und genutzt werden, um das automatische Extrahieren einer Signatur aus einem gegebenen Segment eines Videosignals zu veranlassen. Zum Beispiel kann das System eingerichtet sein, automatisch eine Signatur aus einem Teilbereich des Videosignals zu extrahieren, den ein Benutzer im Schnellvorlauf durchläuft, wenn er eine Wiedergabe einer zuvor aufgenommenen Sendung, einer TiVO-Ausgabe usw., anschaut.

[0053] In einem erfindungsgemäßen Verfahren zum Auffinden, Erkennen und Extrahieren von Werbesendungen können auch Informationen von einem Diensteanbieter, wie z.B. eine Liste von Werbesendungen, verwendet werden. Zum Beispiel kann ein Diensteanbieter Informationen, wie z.B. vorab extrahierte Signaturen, einem Benutzergerät zur Verfügung stellen, um das Identifizieren und Extrahieren des entsprechenden Videoinhalts zu erleichtern.

[0054] Wie oben erwähnt wurde, reduziert die Erfindung beträchtlich den Speicher und die Rechnerressourcen, die für das Identifizieren und Extrahieren von Werbesendungen oder anderen Typen von Videoinhalt benötigt werden. Zum Beispiel können die oben beschriebenen Signaturen so konfiguriert sein, dass sie pro Frame nicht mehr als einige wenige hundert Bytes des Speichers benötigen, und die Rechnerressourcen werden im Vergleich zu herkömmlichen Verfahren bis zu mehreren Größenordnungen reduziert.

[0055] Die oben beschriebenen Ausführungsformen der Erfindung sind nur zur Veranschaulichung gedacht. Die Erfindung kann zum Beispiel verwendet werden, um das Auffinden, Erkennen und Extrahieren von anderen Typen von Bildsignalen in einem beliebigen gewünschten Typ eines Videobearbeitungssystems oder -gerätes und in vielen anderen Anwendungen als den hier beschriebenen zu realisieren. Die Erfindung kann auch zumindest teilweise in der Form eines oder mehrerer Softwareprogramme reali-

siert werden, die auf einem ansonsten herkömmlichen elektronischen, magnetischen oder optischen Speichermedium oder einem anderen Speichertyp gespeichert und durch ein Verarbeitungsgerät, z.B. durch den Prozessor **15** des Systems **10** in [Fig. 1](#), ausgeführt werden. Diese und zahlreiche andere Ausführungsformen im Gültigkeitsbereich der nachfolgenden Ansprüche sind für Fachleute offensichtlich.

Patentansprüche

1. Verfahren für das Bearbeiten eines Videosignals, wobei das Verfahren die Schritte umfasst:

- Identifizieren (**102**, **104**) einer Anzahl von Segmenten des Videosignals, die wahrscheinlich einem speziellen Typ von Videoinhalt zuzuordnen sind,
- Extrahieren (**106**) einer Signatur aus jeder von mindestens einer Untermenge des einen oder der mehreren identifizierten Segmente, wobei die extrahierten Signaturen verwendet werden, um zu bestimmen, welches der identifizierten Segmente mit dem speziellen Typ von Videoinhalt verbunden ist,
- Speichern der extrahierten Signatur in einer ersten Liste von Signaturen des speziellen Videotyps (**108**), gekennzeichnet durch
- Vergleichen der extrahierten Signatur mit gespeicherten Signaturen in einer ersten Liste und in einer zweiten Liste (**110**, **122**),
- so dass im Falle einer Übereinstimmung mit einer gespeicherten Signatur in der ersten Liste (**110**) die passende Signatur in der zweiten Liste (**118**) als Kandidatensignatur gespeichert wird, und
- im Falle einer Übereinstimmung mit einer gespeicherten Signatur in der zweiten Liste (**122**) eine Kandidatenbewertung für die Kandidatensignatur in der zweiten Liste (**126**) erhöht wird.

2. Verfahren nach Anspruch 1, wobei der Identifizierungsschritt die Bestimmungsschritte umfasst, ob ein gegebenes der Segmente mindestens ein Schwarz-Frame verbunden mit einer Tonpause in seinem zugeordneten Audiosignal enthält, eine Schnittrate einen spezifizierten Wert überschreitet, eine spezifizierte Veränderung in einer Farbkenngröße vorliegt, eine spezifizierte Veränderung in einer Kenngröße des zugeordneten Audio vorliegt, ein spezifizierter Texttyp vorliegt, ein spezifizierter Sprachtyp vorliegt, Untertitelinformationen nicht oder in einem speziellen Typ vorliegen und ein eingeblendetes Videologo fehlt.

3. Verfahren nach Anspruch 1, wobei die Signatur, die aus einem gegebenen der Videosegmente extrahiert ist, eine visuelle Frame-Signatur umfasst, die zumindest teilweise auf einer visuellen Kenngröße eines Frame des Videosegments basiert.

4. Verfahren nach Anspruch 3, wobei die visuelle Frame-Signatur unter Verwendung von Informatio-

nen basierend auf DC- und Bewegungskoeffizienten des Frame extrahiert ist.

5. Verfahren nach Anspruch 3, wobei die visuelle Frame-Signatur unter Verwendung von Informationen basierend auf DC- und AC-Koeffizienten des Frame extrahiert ist.

6. Verfahren nach Anspruch 1, wobei die Signatur, die aus einem gegebenen der Videosegmente extrahiert ist, mindestens eines von den folgenden enthält: ein Histogramm, Untertiteltext, einen Name eines Unternehmens, einen Namen eines Produkts oder einer Dienstleistung, eine Telefonnummer, eine Internetadresse, eine Frame-Nummer in Verbindung mit der Platzierung eines Textfeldes, Farb- oder Schriftzeichencodeinformationen und eine Frame-Nummer in Verbindung mit Platzierungsinformationen eines Gesichts oder anderen Objekts.

7. Verfahren nach Anspruch 1, das ferner den Schritt zum Verwenden von Programmidentifizierungsinformationen aus einem elektronischen Programmführer in Verbindung mit den extrahierten Signaturen enthält, um zu bestimmen, welches der identifizierten Elemente mit dem speziellen Typ von Videoinhalt verbunden ist.

8. Verfahren nach Anspruch 1, das ferner den Schritt zum Bereithalten von Sätzen von Signaturlisten enthält, wobei die Sätze von Listen eine oder mehrere von mindestens einer wahrscheinlichen Liste, mindestens einer Kandidatenliste und mindestens einer Fundliste enthalten, wobei jede Eintragung in eine gegebene Liste einer Signatur entspricht, die mit einem speziellen Videosegment verbunden ist.

9. Verfahren nach Anspruch 1, das ferner den Schritt zum automatischen Extrahieren einer Signatur aus einem speziellen Segment eines Videosignals als Reaktion auf den Nachweis einer Aktion des Benutzers einschließt, die sich auf das spezielle Segment bezieht.

10. Gerät zum Bearbeiten eines Videosignals, wobei das Gerät umfasst:

- einen Prozessor (**15**), der funktionsfähig ist, mindestens ein Segment des Videosignals zu ermitteln, das mit einem speziellen Typ von Videoinhalt verbunden ist, und eine Signatur aus dem ermittelten Element zu extrahieren, wobei die extrahierte Signatur verwendet wird, um zusätzliche Segmente des Videosignals zu ermitteln, die auch mit dem speziellen Typ von Videoinhalt verbunden sind, und
- einen Speicher (**16**), der an den Prozessor gekoppelt ist, um mindestens einen Teil der extrahierten Signatur zu speichern, dadurch gekennzeichnet, dass
- der Prozessor ferner funktionsfähig ist
- die extrahierte Signatur mit gespeicherten Signatu-

ren in einer ersten Liste und in einer zweiten Liste (**110**, **122**) zu vergleichen,

- so dass im Falle einer Übereinstimmung mit einer gespeicherten Signatur in der ersten Liste (**110**) die passende Signatur in der zweiten Liste (**118**) als Kandidatensignatur gespeichert wird, und
- im Falle einer Übereinstimmung mit einer gespeicherten Signatur in der zweiten Liste (**122**) eine Kandidatenbewertung für die Kandidatensignatur in der zweiten Liste (**126**) erhöht wird.

11. Fertigungserzeugnis mit einem maschinenlesbaren Medium, das ein oder mehrere Softwareprogramme enthält, die beim Ausführen die Schritte realisieren:

- Identifizieren (**102**, **104**) einer Anzahl von Segmenten des Videosignals, die wahrscheinlich einem speziellen Typ von Videoinhalt zuzuordnen sind,
- Extrahieren (**106**) einer Signatur aus jeder der mindestens einen Untermenge des einen oder der mehreren identifizierten Segmente, wobei die extrahierten Signaturen verwendet werden, um zu bestimmen, welches der identifizierten Segmente mit dem speziellen Typ von Videoinhalt verbunden ist,
- Speichern der extrahierten Signatur in einer ersten Liste von Signaturen des speziellen Videotyps (**108**), gekennzeichnet durch
- Vergleichen der extrahierten Signatur mit gespeicherten Signaturen in einer ersten Liste und in einer zweiten Liste (**110**, **122**),
- so dass im Falle einer Übereinstimmung mit einer gespeicherten Signatur in der ersten Liste (**110**) die passende Signatur in der zweiten Liste (**118**) als Kandidatensignatur gespeichert wird, und
- im Falle einer Übereinstimmung mit einer gespeicherten Signatur in der zweiten Liste (**122**) eine Kandidatenbewertung für die Kandidatensignatur in der zweiten Liste (**126**) erhöht wird.

Es folgen 3 Blatt Zeichnungen

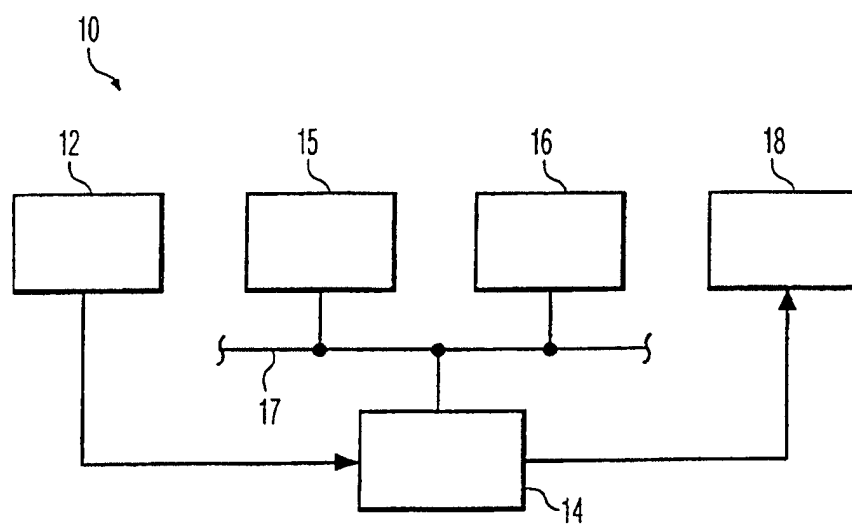
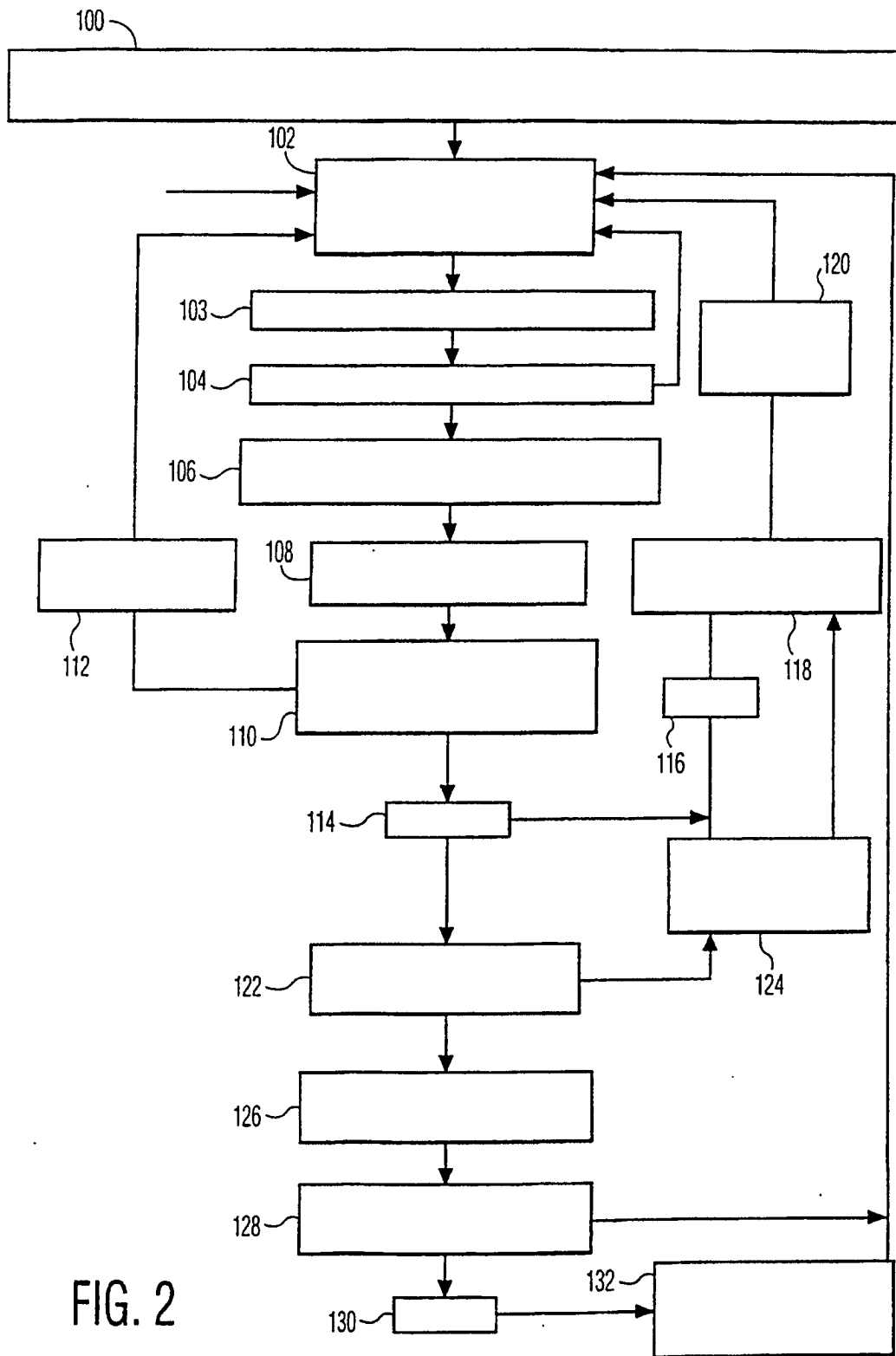


FIG. 1



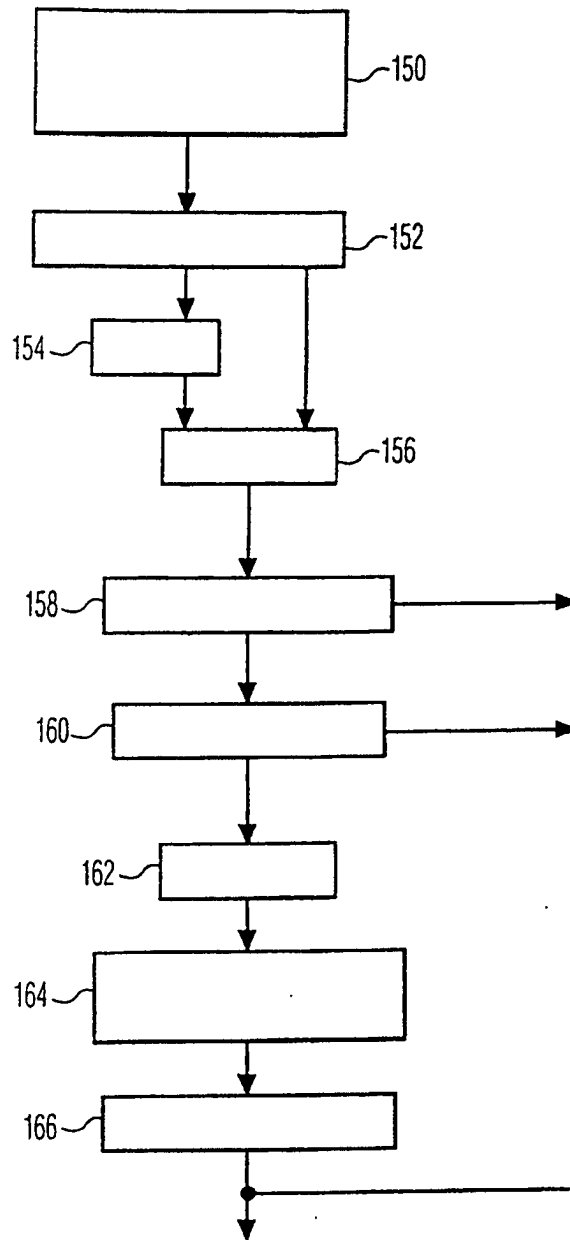


FIG. 3