

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4733328号
(P4733328)

(45) 発行日 平成23年7月27日(2011.7.27)

(24) 登録日 平成23年4月28日(2011.4.28)

(51) Int.Cl.

F I

G 0 6 F 17/30 (2006.01)

G 0 6 F 17/30 4 1 4 A

G 0 6 F 17/30 2 2 O A

G 0 6 F 17/30 1 7 O D

請求項の数 6 (全 13 頁)

(21) 出願番号	特願2001-530817 (P2001-530817)	(73) 特許権者	596180076
(86) (22) 出願日	平成12年9月29日 (2000.9.29)		韓国電子通信研究院
(65) 公表番号	特表2003-511801 (P2003-511801A)		Electronics and Tel
(43) 公表日	平成15年3月25日 (2003.3.25)		ecommunications Res
(86) 国際出願番号	PCT/KR2000/001084		earch Institute
(87) 国際公開番号	W02001/027876		大韓民国大田廣域市儒城區柯亭洞 161
(87) 国際公開日	平成13年4月19日 (2001.4.19)		161 Kajong-dong, Yu
審査請求日	平成19年9月27日 (2007.9.27)		song-gu, Taejon kor
(31) 優先権主張番号	1999/43712		ea
(32) 優先日	平成11年10月11日 (1999.10.11)	(74) 代理人	110001243
(33) 優先権主張国	韓国 (KR)		特許業務法人 谷・阿部特許事務所
前置審査		(74) 復代理人	100115624
			弁理士 濱中 淳宏
		(74) 復代理人	100145388
			弁理士 藤原 弘和

最終頁に続く

(54) 【発明の名称】 効率的な概観及びブラウジングのためのビデオ要約記述構造と、ビデオ要約記述データ生成の方法及びシステム

(57) 【特許請求の範囲】

【請求項 1】

ビデオ要約記述データをブラウジングする装置において、
 前記ビデオ要約記述データは階層的な要約記述構造を有し、
 前記階層的な要約記述構造は、
 要約ビデオ区間にそれぞれ対応し、各要約ビデオ区間に該当する情報を記述する複数の
 ハイライトセグメント記述構造を有するハイライトレベル記述構造と；
 要約を構成する主題または事件情報を列挙して事件中心の要約およびブラウジング機能
 を提供する要約主題リスト記述構造とを含み、
 前記ビデオ要約記述データをブラウジングする装置は、
 原ビデオや要約ビデオを再生するビデオ再生部と；
 原ビデオの代表フレームを再生する原ビデオ代表フレーム部と；
 第1要約レベルの各ビデオ区間を代表する代表フレームと代表音響情報を再生する第1
 要約ビデオ代表フレーム部および代表音響部と；
 前記第1要約レベルより精巧に要約された第2要約レベルの各ビデオ区間を代表する代
 表フレームと代表音響情報を再生する第2要約ビデオ代表フレーム部および代表音響部と
 ；

前記第1要約レベルまたは前記第2要約レベルを選択して、前記ビデオ再生部が該当す
 る要約レベルの要約ビデオを再生するようにするレベル選択部と；

前記要約主題リスト記述構造によって提供される事件および主題を列挙して、使用者が

所望の事件をブラウジングするようにする事件選択部とを含むことを特徴とする、ビデオ要約記述データのブラウジング装置。

【請求項 2】

前記ハイライトセグメント記述構造は、

前記該当ハイライトセグメントの要約ビデオ区間の時間情報を記述するビデオセグメント位置指定記述構造と、

前記該当ハイライトセグメントの代表フレームを記述する映像位置指定記述構造とを含むことを特徴とする、請求項 1 に記載のビデオ要約記述データのブラウジング装置。

【請求項 3】

ビデオ要約記述データをブラウジングする装置が実施する方法において、

前記ビデオ要約記述データは階層的な要約記述構造を有し、

前記階層的な要約記述構造は、要約ビデオ区間にそれぞれ対応し、各要約ビデオ区間に該当する情報を記述する一つ以上のハイライトセグメント記述構造を有するハイライトレベル記述構造を含み、

前記ハイライトセグメント記述構造は、前記該当ハイライトセグメントの要約ビデオ区間の時間情報を記述するビデオセグメント位置指定記述構造、および前記該当ハイライトセグメントの代表フレームを記述する映像位置指定記述構造を含み、

前記ビデオ要約記述データをブラウジングする方法は、

ビデオ再生部が前記ハイライトセグメント記述構造と関連する第 1 要約レベルの要約ビデオを再生する第 1 段階と；

第 1 要約ビデオ代表フレーム部および代表音響部によって再生される要約ビデオ代表フレームを通じて前記第 1 段階で探そうとする場面が確認されると、前記ビデオ再生部が前記代表フレームに連結された原ビデオのビデオ区間を再生する第 2 段階と；

を含むことを特徴とする、ビデオ要約記述データのブラウジング方法。

【請求項 4】

前記ビデオ要約記述データをブラウジングする方法は、前記第 1 段階で探そうとする場面が確認されない場合、前記ビデオ再生部が前記第 1 要約レベルより精巧に要約された第 2 要約レベルの要約ビデオを再生する第 3 段階をさらに含むことを特徴とする、請求項 3 に記載のビデオ要約記述データのブラウジング方法。

【請求項 5】

前記ハイライトセグメント記述構造は、前記該当ハイライトセグメントの代表音響情報を記述する音響位置指定記述構造をさらに含み、

前記第 2 段階は、前記第 1 要約ビデオ代表フレーム部および代表音響部によって再生される要約ビデオ代表音響情報を通じて前記第 1 段階で探そうとする場面を確認する段階を含むことを特徴とする、請求項 3 または 4 に記載のビデオ要約記述データのブラウジング方法。

【請求項 6】

前記階層的な要約記述構造は、要約を構成する主題または事件情報を列挙して事件中心の要約およびブラウジング機能を提供する要約主題リスト記述構造をさらに含むことを特徴とする、請求項 3 乃至 5 のうちのいずれか一項に記載のビデオ要約記述データのブラウジング方法。

【発明の詳細な説明】

【0001】

(技術分野)

本発明は効率的なビデオ概観 (overview) 及びブラウジングのためのビデオ要約記述構造に関する。また、ビデオ要約記述構造によってビデオ要約を記述するためのビデオ要約記述生成の方法及びシステムに関する。本発明の属する技術分野は、内容を基にしたビデオ索引 (indexing) 及びブラウジング/検索の分野で、ビデオを内容に基づいて要約し、これを記述する分野である。

10

20

30

40

50

【 0 0 0 2 】

(発明の背景)

ビデオ要約形態は大きく分けて、動的要約と静的要約になる。本発明によるビデオ記述構造は動的要約と静的要約とを統一的記述構造で効率的に記述するためのものである。

【 0 0 0 3 】

一般に、既存のビデオ要約及び記述構造はビデオ要約に含まれたビデオ区間に関する情報を提供するだけなので、現状のビデオ要約と記述構造は要約ビデオの再現を通じて全体ビデオの内容を伝達するのみに限定される。しかし、多くの場合、要約ビデオを通じて全体内容を概観するだけよりは、全体内容の概観を通じて関心のある部分を識別し再び呼び出すためのブラウジングが必要である。

10

【 0 0 0 4 】

また、既存のビデオ要約はビデオ要約提供者が定めた基準によって重要であると判断されるビデオ区間だけを使用者に提供する。したがって、使用者とビデオ提供者の基準が異なる場合、あるいは使用者が特別な基準を持つ場合、使用者は所望のビデオ要約を得ることができない。つまり、既存の要約ビデオは、いくつかのレベルの要約ビデオを提供して使用者が所望のレベルの要約ビデオを選択できるが、要約ビデオの内容による選択ができないので使用者の選択範囲が制限的である。

【 0 0 0 5 】

発明の名称が “ method and apparatus for video browsing based on content and structure ” で、登録番号が US 5 8 2 1 9 4 5 である特許では、ビデオを集約的に再現し、その再現を通じて所望の内容のビデオに接近する (access) ブラウジング機能を提供する。しかし、この特許では、代表フレームに基づいた静的な要約であって、既存の静的要約はビデオショットの代表フレームを利用して要約するが、この特許の代表フレームは単にそのショットを代表する映像情報だけを提供するため、要約を利用した情報伝達に限界がある。この特許に比べて、前記ビデオ記述構造とブラウジング方法はビデオセグメントに基づいた動的要約を利用する。

20

【 0 0 0 6 】

1 9 9 9 年 7 月に I S O / I E C J T C 1 / S C 2 9 / W G 1 1 M P E G - 7 O u t p u t D o c u m e n t N o . N 2 8 4 4 として発表された “ M P E G - 7 D e s c r i p t i o n S c h e m e (V 0 . 5) ” で提案されたビデオ要約記述構造は、動的要約ビデオの各ビデオセグメントの区間情報だけを記述する。これは動的要約を記述する基本的な機能は提供するが、次の側面で問題点を持っている。まず、従来は要約ビデオを構成する要約セグメントから原ビデオへの接近 (access) を提供できないという短所がある。つまり、使用者は要約ビデオを通じた概観と要約内容 (summary contents) に基づいてより詳細な内容把握のために原ビデオへ接近しようとするが、従来はこれが提供できなかった。また、オーディオ要約記述機能が十分に提供できない。最後に事件基盤の要約 (event-based summary) を表現しようとする時、重複記述と探索の複雑性が回避となる短所を持っている。

30

【 0 0 0 7 】

(発明の概要)

したがって、本発明は上記の問題点を改善するために、要約ビデオと共に要約ビデオに含まれた各ビデオ区間ごとに代表フレーム情報、代表音響情報を含み、要約ビデオの内容に対する使用者の選択を提供する使用者注文形 (user customized) の事件基盤要約 (event based summary) と効果的なブラウジングを可能にする階層的ビデオ要約記述構造と、その記述構造を利用したビデオ要約記述データ生成方法及びシステムを提供するのにその目的がある。

40

【 0 0 0 8 】

このような目的を達成するための本発明の一つの実施例による階層的要約記述構造は、ハイライトレベルを記述している一つ以上のハイライトレベル記述構造を含み、前記ハイラ

50

イトレベル記述構造は、そのハイライトレベルの要約ビデオを構成するハイライトセグメント情報を記述している最少限ハイライトセグメント記述構造を含むことを特徴とする。

【 0 0 0 9 】

好ましくは、前記ハイライトレベル記述構造は一つ以上の下位レベルのハイライトレベル D S (D S = 記述構造) で構成されることを特徴とする。

【 0 0 1 0 】

更に好ましくは、前記ハイライトセグメント記述構造は、前記該当ハイライトセグメントの時間情報又はビデオ自身を記述しているビデオセグメント位置指定記述構造を含むことを特徴とする。

【 0 0 1 1 】

前記ハイライトセグメント D S が、前記該当ハイライトセグメントの代表フレームを記述している映像位置指定 D S を、更に含むことが望ましい。

【 0 0 1 2 】

前記ハイライトセグメント D S が、前記該当ハイライトセグメントの代表音響情報を記述している音響位置指定 D S を、更に含むことが一層望ましい。

【 0 0 1 3 】

前記ハイライトセグメント D S が、前記該当ハイライトセグメントの代表フレームを記述している映像位置指定 D S 及び前記該当ハイライトセグメントの代表音響情報を記述している音響位置指定 D S を、更に含むことが望ましい。

【 0 0 1 4 】

前記映像位置指定 D S が、前記該当ハイライトセグメントに対応するビデオ区間の代表フレームの時間情報又は映像データを、記述することが一層望ましい。

【 0 0 1 5 】

前記ハイライトセグメント D S が、前記該当ハイライトセグメントの音響要約を構成している音響セグメント情報を記述している音響セグメント位置指定 D S を、更に含むことが望ましい。

【 0 0 1 6 】

前記音響セグメント位置指定 D S が、前記該当ハイライトセグメントの音響区間の時間情報又は音響データを記述することが一層望ましい。

【 0 0 1 7 】

好ましくは、前記階層的要約記述構造は、前記階層的要約記述構造に含まれる全てのサマリーコンポーネントタイプを列挙し、記述しているサマリーコンポーネントリストを含むことを特徴とする。

【 0 0 1 8 】

好ましくは、前記階層的要約記述構造は、要約に含まれた事件または主題を列挙し、その I D を記述している要約主題リスト記述構造 (Summary Theme List D S) を含み、事件中心の要約を記述し、使用者が要約ビデオを前記要約主題リストに記述された主題または事件別にブラウジングできるようにすることを特徴とする。

【 0 0 1 9 】

前記要約主題リスト D S が、要素としての要約主題を任意件数含み、前記要約主題が、該当する事件又は主題を表わす i d (識別記号) の属性を含み、この要約主題が、上位レベルの事件又は主題の i d を記述する親 I D の属性を更に含むことが一層望ましい。

【 0 0 2 0 】

共通の事件又は主題の前記 i d 属性を記述している主題 I d の属性を、前記ハイライトレベル D S が含むと好ましい場合は、該当ハイライトレベルを構成している全てのハイライトセグメント及びハイライトレベルが共通の事件と主題を有する場合である。

【 0 0 2 1 】

前記ハイライトセグメント D S が、前記 i d 属性を記述している主題 I d の属性を含み、該当ハイライトセグメントの事件又は主題を記述することが望ましい。

【 0 0 2 2 】

10

20

30

40

50

また、本発明によると、階層的要約記述構造が保存されたコンピュータで読むことができる記録媒体が提供される。階層的要約記述構造が、ハイトレベルDSを一つ以上含み、ハイトレベルDSが、ハイトレベルの要約ビデオを構成しているハイトセグメント情報を記述しているハイトセグメントDSを一つ以上含み、ハイトセグメントDSが、前記該当ハイトセグメントの時間情報又はビデオ自身を記述しているビデオセグメント位置指定DSを含むことが望ましい。

【0023】

また、本発明によると、原ビデオを入力してビデオ要約記述構造に従ってビデオ要約記述データを生成するビデオ要約記述データ生成方法が提供される。これは原ビデオを入力して分析し、ビデオ分析結果を出力するビデオ分析段階と；要約ビデオ区間を選択するための要約規則を定義する要約規則定義段階；前記原ビデオ分析結果と前記要約規則とを入力して原ビデオからビデオ内容を要約することができるビデオ区間を選択して、要約ビデオ区間情報を構成する要約ビデオ区間選択段階；及び前記要約ビデオ区間選択段階から出力された要約ビデオ区間情報の入力を受けて、階層的要約記述構造によってビデオ要約記述データを生成するビデオ要約記述段階を含んでなることを特徴とする。

10

【0024】

また、本発明によると、原ビデオを入力してビデオ要約記述構造に従ってビデオ要約記述データを生成するビデオ要約記述データ生成システムが提供される。これは原ビデオを入力して分析し、ビデオ分析結果を出力するビデオ分析手段と；要約ビデオ区間を選択するための要約規則を定義する要約規則定義手段；前記原ビデオ分析結果と前記要約規則とを入力して原ビデオからビデオ内容を要約することができるビデオ区間を選択して、要約ビデオ区間情報を構成する要約ビデオ区間選択手段；及び前記要約ビデオ区間選択手段で定義された要約ビデオ区間情報の入力を受けて、階層的要約記述構造を有するビデオ要約記述データを生成するビデオ要約記述手段を含んでなることを特徴とする。

20

【0025】

また、本発明によると、上述したようなビデオ要約記述データ生成方法でビデオを階層的に要約するビデオ要約記述データ生成システムを機能させるためのプログラムを記録したコンピュータで読むことができる記録媒体が提供される。

【0026】

また、本発明によるサーバー/クライアント環境でのビデオブラウジングシステムは、原ビデオの入力を受けて階層的に要約記述構造に基づいてビデオ要約記述データを生成し、前記原ビデオとビデオ要約記述データとをリンクするビデオ要約記述データ生成システムを備えたサーバーと；

30

前記ビデオ要約記述データを利用して前記原ビデオを概観し、前記サーバーの原ビデオに接近してビデオをブラウジング及びナビゲーションするクライアントを備えることを特徴とする。

【0027】

以下、添付した図面を参照して本発明の好ましい一実施例を詳細に説明する。図中、参照番号は同一部分または同様部分を識別するために用いる。

【0028】

図1は本発明による記述構造(description scheme)によってビデオ要約記述データを生成するためのシステムを示したブロック図である。図1に示すように、本発明によるビデオ要約記述データ生成装置は特徴抽出部101、事件検出部102、エピソード検出部103、要約ビデオ区間選択部104、要約規則定義部105、代表フレーム抽出部106、代表音響抽出部107、及びビデオ要約記述部108で構成される。

40

【0029】

特徴抽出部101は原ビデオを入力して要約ビデオを生成するために必要な特徴を抽出する。一般的な特徴としてはショット境界、カメラの動き、字幕領域、顔領域などがある。特徴抽出段階ではこれら特徴を抽出して特徴の種類とこれら特徴が検出されるビデオ時間区間を(特徴種類、特徴一連番号、時間区間)の形態で事件検出段階に出力する。例えば

50

、カメラ動きの場合（カメラズーム、1、100～150）には第1カメラズームが100～150番目のフレームで検出されたという情報を表現する。

【0030】

事件検出部102は原ビデオに含まれたキーになる事件を検出する。これら事件は原ビデオの内容を代表的によく表現しなければならず、要約ビデオを生成するのに基準となるものであるため、一般に原ビデオのジャンルに従って区別するように定義される。事件は上位の意味レベルを示すこともあり、上位の意味を直接類推することができるビジュアル特徴であることもある。例えば、サッカービデオの場合、ゴール、シュート、字幕、リプレー（replay）などを事件として定義することができる。

【0031】

事件検出部102は検出した事件の種類とその時間区間を（事件種類、事件番号、時間区間）形態で出力する。例えば、最初のゴールが200～300フレームの間で発生したという事件情報は（ゴール、1、200～300）の形態で出力する。

【0032】

エピソード検出部103は検出された事件に基づいてビデオを話の筋に基づいた一つの事件よりさらに大きい単位のエピソードに分割する。キー事件を検出した後、そのキー事件を中心にその事件による付属事件を含んで一つのエピソードを検出する。一例として、サッカービデオの場合、ゴールとシュートはキー事件になり、その事件の付属事件としてゴールやシュートが発生した時のベンチ場面、観衆席場面、ゴールセレモニー場面、ゴール場面のリプレーなどがその事件の付属事件を構成する。つまり、ゴールとシュートとを中心

【0033】

エピソード検出情報は（エピソード番号、時間区間、優先順位、特徴ショット、関連事件情報）の形態で出力する。ここでエピソード番号はエピソードの一連番号であり、時間区間はそのエピソードの時間区間をショット単位で示す。優先順位はそのエピソードの重要度を示す。特徴ショットはそのエピソードを構成するショットの中で最も重要な情報を含んだショット番号を示し、関連事件情報はそのエピソードと関係する事件の事件番号を示す。例えば、（エピソード1、4～6、1、5、ゴール1、字幕3）のようなエピソード情報を表示する場合、その情報は、第1エピソードが4～6番目ショットを含み、優先順位が最高（1）、特徴ショットが5番ショットであり、関連事件が1番ゴールと3番字幕であることを示す。

【0034】

要約ビデオ区間選択部104は、検出されたエピソードに基づいて原ビデオ内容をよく要約することができるビデオ区間を選択する。この区間選択基準は予め決められた要約規則定義部105の要約規則によって行う。

【0035】

要約規則定義部105は要約区間を選択するための規則を定義して要約区間を選択するための制御信号を出力する。また、要約規則定義部105は要約ビデオ区間選択の基盤となる要約事件種類をビデオ要約記述部108に出力する。

【0036】

要約ビデオ区間選択部104は選択された要約ビデオの区間の時間情報をフレーム単位に出力し、ビデオ区間に対応する事件種類を出力する。つまり、（100～200、ゴール）、（500～700、シュート）等の形態は要約ビデオ区間として選択されたビデオセグメントが、100～200フレーム、500～700フレーム等であって、各セグメントの事件はゴールとシュートであることを示す。または、要約ビデオ区間だけで構成された追加的ビデオに接近できるようにファイル名などの情報を出力することも可能である。

【0037】

要約ビデオ区間選択が完了すれば、その要約ビデオ区間情報を利用して、代表フレームと代表音響を代表フレーム抽出部106と代表音響抽出部107から各々抽出する。代表フレーム抽出部106はその要約ビデオ区間を代表する映像のフレーム番号またはその映像

10

20

30

40

50

データを出力する。代表音響抽出部 107 はその要約ビデオ区間を代表する音響データまたは音響時間区間を出力する。

【0038】

ビデオ要約記述部 108 は、図 2 に記述された本発明による階層的記述構造によって効果的な要約及びブラウジング機能を可能にする関連情報を記述する。階層的要約 DS の主要情報は、要約ビデオの要約事件種類、各要約ビデオ区間を記述する時間情報、代表フレーム、代表音響、及び各区間の事件種類を含む。

【0039】

ビデオ要約記述部 108 は図 2 に示された記述構造によるビデオ要約記述データを出力する。

10

【0040】

図 2 は、本発明によるビデオ要約記述データを記述する階層的要約記述構造 (Hierarchical Summary DS) のデータ構造を UML (Unified Modeling Language) で示した図面である。

【0041】

ビデオ要約を記述する階層的要約記述構造 201 は一つ以上のハイライトレベル記述構造 (Highlight Level DS) 202 と 1 個または 0 個の要約主題リスト記述構造 (Summary Theme List DS) 203 を含んでいる。要約主題リスト DS (Summary Theme List DS) は要約を構成する主題または事件の情報を網羅的に記述することで、事件中心の要約及びブラウジングの機能を提供する。

20

【0042】

ハイライトレベル記述構造 (Highlight Level DS) 202 は、そのレベルの要約ビデオを構成するビデオ区間数だけの個数のハイライトセグメント記述構造 (Highlight Segment DS) 204 と 0 個または数個のハイライトレベル記述構造 (Highlight Level DS) で構成される。ハイライトセグメント記述構造は各要約ビデオ区間に対応する情報を記述する。ハイライトセグメント記述構造は一つのビデオセグメント位置指定記述構造 (Video Segment Locator DS) 205、0 個または数個の映像位置指定記述構造 (Image Locator DS) 206、そして 0 個または数個の音響位置指定記述構造 (Sound Locator DS) 207 及びオーディオセグメント位置指定記述構造 (Audio Segment Locator DS) 208 を含んでいる。

30

【0043】

以下、この階層的要約記述構造についてより詳細に説明する。

【0044】

階層的要約記述構造 (Hierarchical Summary DS) は階層的要約 DS により包括される要約形態を明確に示すサマリーコンポーネントリスト (Summary Component List) という属性 (attribute) を有する。サマリーコンポーネントリスト (Summary Component List) は要約タイプ (Summary Component Type) に基づいて派生し、サマリーコンポーネントタイプを含む全てのものを列挙して記述する。

【0045】

サマリーコンポーネントリストにはキーフレーム、キービデオクリップ、キーオーディオクリップ、キーイベント及びアンコンストレイン (unconstrained) の 5 種類がある。キーフレームは代表フレームで構成されたキーフレーム要約を示す。キービデオクリップは主要ビデオ区間の集合で構成されたキービデオクリップ要約を示し、キーイベントは事件または主題に対応するビデオ区間で構成された要約を示し、キーオーディオクリップは代表オーディオ区間の集合で構成されたキーオーディオクリップ要約を示す。アンコンストレインは前記要約以外の、使用者が定義した要約形態を示す。

40

【0046】

また、事件中心の要約を記述するために階層的要約記述構造は要約に含まれた事件 (または主題) を列挙してその ID を記述する要約主題リスト記述構造 (Summary Theme List DS) を含むこともできる。

50

【 0 0 4 7 】

要約主題リストは任意の数の要約テーマを要素 (element) として持つ。要約テーマは I D 形の i d という属性を有して親 I D という属性を選択的に持つ。

【 0 0 4 8 】

要約主題リスト記述構造は、要約主題リストに記述された各事件又はいくつかの主題の観点から、使用者が要約ビデオをブラウジングできるようにする。つまり、記述データを入力する応用ツールは、要約主題リスト記述構造を解析し、この情報を使用者に提示して、使用者が望む主題を選択させる。この時、このような主題を単純な形態に列挙する場合、主題の数が多ければ使用者が望む主題を探すのが容易でないことがある。

【 0 0 4 9 】

したがって、目次 (ToC=Table of Content) と類似したツリー構造として主題を表現することによって、使用者は、所望の主題を発見した後、主題別ブラウジングが効率的にできる。このために本発明では、要約テーマに親 I D の属性を選択的に使用できるようにする。この親 I D とは、ツリー構造における上位の要素 (上位の主題) を意味する。

【 0 0 5 0 】

本発明の階層的な要約記述構造はハイライトレベル記述構造 (Highlight Level DSs) を含み、各ハイライトレベル記述構造は、要約ビデオを構成するビデオセグメント (または区間) に対応する一つ以上のハイライトセグメント記述構造を含む。

【 0 0 5 1 】

ハイライトレベル記述構造は I D R E F S 形のテーマ I d s の属性を有する。このテーマ I d s は、該当ハイライトレベルに含まれた全てのハイライトセグメント記述構造または該当ハイライトレベル記述構造の子ハイライトレベル記述構造に共通した、主題及び事件 i d を記述するが、この i d は前記要約主題リスト記述構造に記述されている。テーマ I d s は数個の事件を意味することができ、事件中心の要約をする時、そのレベルを構成する全セグメント内で同一 i d が不必要に繰り返される問題点を解決するために、そのレベルを構成するハイライトセグメント内で共通した主題の形を示すテーマ I d s をおく。

【 0 0 5 2 】

ハイライトセグメント記述構造は一つのビデオセグメント位置指定記述構造 (Video Segment Locator DS) と、一つ以上の映像位置指定記述構造 (Image Locator DS) と、0 個または 1 個の音響位置指定記述構造 (Sound Locator DS) と、0 個または 1 個のオーディオセグメント位置指定記述構造 (Audio Segment Locator DS) を含む。

【 0 0 5 3 】

ここで、ビデオセグメント位置指定記述構造は、要約ビデオを構成するビデオセグメントのビデオ自身または時間情報を記述する。映像位置指定記述構造は、そのビデオセグメントの代表フレームの映像データ情報を記述する。音響位置指定記述構造は、該当ビデオセグメント区間を代表する音響情報を記述する。オーディオセグメント位置指定記述構造は、オーディオ要約を構成するオーディオセグメントの区間時間情報又はオーディオ情報自身を記述する。

【 0 0 5 4 】

ハイライトセグメント記述構造はテーマ I d s の属性を有する。このテーマ I d s は、該当ハイライトセグメントに関連する前記要約主題リスト記述構造内で記述された主題又は事件を、要約主題リスト記述構造内で定義された i d を利用して、記述する。テーマ I d s は複数の事件を意味することができ、一つのハイライトセグメントに複数の主題を含ませることができるようにして、事件基盤の要約をするための既存の方法を使う時、事件 (または主題) 毎にビデオセグメントを記述することにより生ずる不可避な重複記述の問題点を解決するという、本発明の効率的な記述方法である。

【 0 0 5 5 】

要約ビデオを構成するハイライトセグメントを記述する時、単にそのハイライトビデオ区間の時間情報だけを記述した既存の階層的な要約記述構造とは異なって、各ハイライトセグメントのビデオ区間情報、代表フレーム情報、代表音響情報を記述できるように、ビデオ

10

20

30

40

50

セグメント位置指定記述構造、映像セグメント位置指定記述構造、サウンド位置指定記述構造を設定して、本発明は、ハイライトセグメントビデオを通じた概観とそのセグメントの代表フレーム及び代表音響を活用したナビゲーション及びブラウジングを、要約ビデオを構成するハイライトセグメントを記述するためのハイライトセグメント記述構造を導入し、効率的に使えるようにする。

【 0 0 5 6 】

ビデオ区間に該当する代表音響を記述することができるサウンド位置指定記述構造を設定して、実際に、そのビデオ区間を代表できる特徴的な音響、例えば銃声、かん声、サッカーでアンカーのコメント（例、ゴール、シュート）、ドラマでの俳優の名前、特定単語などを通じて、そのビデオ区間を再生してみなくても短時間に、その区間が使用者が望む内容を含む重要な区間であるかどうか、どんな内容が含まれた区間であるかを大略的に把握することで効率的なブラウジングが可能である。

10

【 0 0 5 7 】

図 3 は、図 2 と同じ記述構造で記述されたビデオ要約記述データを入力する要約ビデオの再生及びブラウジングのためのツールの使用者インターフェースの構成図である。ビデオ再現部 3 0 1 は使用者の制御に従って原ビデオまたは要約ビデオを再生する。原ビデオ代表フレーム部 3 0 5 は原ビデオのショットの代表フレームを再現する。つまり、一連の縮小映像で構成される。原ビデオのショットの代表フレームは、本発明の階層的要約記述構造ではなく別途の記述構造で記述され、この記述データが本発明の階層的要約記述構造で記述された要約記述データと共に提供される時に、活用できる。使用者は、代表フレームをクリックすることにより代表フレームに対応する原ビデオのショットに、接近する。要約ビデオレベル 0 代表フレーム部及び代表音響部 3 0 7 と要約ビデオレベル 1 代表フレーム部及び代表音響部 3 0 6 は、要約ビデオレベル 0 と要約ビデオレベル 1 夫々の各ビデオ区間を代表するフレームと音響情報を与える。つまり、一連の縮小された映像及び音響を示すアイコン状映像で構成される。使用者が要約ビデオ代表フレーム部及び代表音響部の代表フレームをクリックすると、その代表フレームに対応する原ビデオ区間に接近する。この時、要約ビデオの代表フレームに対応する代表音響アイコンをクリックすると、そのビデオ区間の代表音響が再生される。

20

【 0 0 5 8 】

要約ビデオ制御部 3 0 2 は要約ビデオを再生するために使用者の選択のための制御を入力する。多階層の要約ビデオが提供される場合、使用者がレベル選択部 3 0 3 を通じて所望のレベルの要約を選択することにより、概観しブラウジングする。事件選択部 3 0 4 は要約主題リストによって提供される事件及び主題を列挙して、使用者は所望の事件を選択することにより概観し、ブラウジングする。結局、これが使用者注文形の要約を実現する。

30

【 0 0 5 9 】

図 4 は、本発明の要約ビデオを利用した階層的ブラウジングのためのデータ及び制御の流れに関する構成図である。ブラウジングは図 3 の使用者インターフェースを利用して図 4 の方法でブラウジングのためのデータに接近して行う。ブラウジングのためのデータは要約ビデオと要約ビデオの代表フレーム、原ビデオ 4 0 6 と原ビデオ代表フレーム 4 0 5 である。要約ビデオは二つのレベルを有するものとする。もちろん要約ビデオが二つ以上のレベルを有することもある。要約ビデオレベル 0（記号 4 0 1）は要約ビデオレベル 1（記号 4 0 3）より短く要約されたものである。つまり、要約ビデオレベル 1 が要約ビデオレベル 0 より多くの内容を含んでいる。要約ビデオレベル 0 代表フレーム 4 0 2 は要約ビデオレベル 0 の代表フレームであり、要約ビデオレベル 1 代表フレーム 4 0 4 は要約ビデオレベル 1 の代表フレームである。

40

【 0 0 6 0 】

要約ビデオと原ビデオは、図 3 のビデオ再現部 3 0 1 を通じて再現される。要約ビデオレベル 0 代表フレームは要約ビデオレベル 0 代表フレーム部及び代表音響部 3 0 6 で表示され、要約ビデオレベル 1 代表フレームは要約ビデオレベル 1 代表フレーム部及び代表音響部 3 0 7 で表示される。原ビデオ代表フレームは原ビデオ代表フレーム部 3 0 5 に表示さ

50

れる。

【 0 0 6 1 】

図 4 に示された階層的ブラウジング方法は、次の例のように多様な形態の階層的経路を有することができる。

場合 1) (1) - (2)

場合 2) (1) - (3) - (5)

場合 3) (1) - (3) - (4) - (6)

場合 4) (7) - (5)

場合 5) (7) - (4) - (6)

【 0 0 6 2 】

全体的なブラウジング技法は次の通りである。まず、原ビデオの要約ビデオを見て原ビデオの全体内容を把握する。この時、要約ビデオは要約ビデオレベル 0 又は要約ビデオレベル 1 を再現できる。要約ビデオを見た後、要約ビデオでさらに詳細にブラウジングしようとする時、関心のあるビデオ区間を要約ビデオ代表フレームを通じて確認する。正確に探そうとする場面が要約ビデオ代表フレームで確認できると、その代表フレームを連結された原ビデオのビデオ区間に直ちに接近して再生する。より詳細な情報が必要な場合、次のレベルの代表フレームを把握したり、原ビデオの代表フレームの内容を階層的に把握して、所望の原ビデオに接近する。このような階層的ブラウジング技法は所望の内容に接近するために、原ビデオを再生しながらブラウジングすると多くの時間がかかるが、原ビデオの内容を階層化された代表フレームを通じて直接に接近するために、ブラウジング時間を相当減らすことができる。

【 0 0 6 3 】

既存の一般的なビデオ索引及びブラウジング技法は、原ビデオをショット単位に分割し、各ショットを代表する代表フレームを構成して代表フレームから所望のショットを認識してそのショットに接近する。この場合、原ビデオのショットの個数が非常に多いので多くの数の代表フレームから所望の内容をブラウジングするのに多くの時間と努力を要する。本発明では要約ビデオの代表フレームで階層的な代表フレームを構成して速く所望のビデオに接近することができる。

【 0 0 6 4 】

場合 1) は、要約ビデオレベル 0 を再現して要約ビデオレベル 0 代表フレームから直ちに原ビデオに接近する場合である。場合 2) は、要約ビデオレベル 0 を再現して要約ビデオレベル 0 代表フレームから最も関心のある代表フレームを選択して原ビデオに接近する前にさらに詳細な情報を把握するために、その代表フレームの近くに該当する要約ビデオレベル 1 の代表フレーム内に所望の場面を確認して、原ビデオに接近する場合である。場合 3) は、場合 2) で要約ビデオレベル 1 代表フレームから直ちに原ビデオに接近するのが難しい場合、さらに詳細な情報を得るために最も関心のある代表フレームを選択し、その代表フレーム近くの原ビデオ代表フレームによって所望の場面を確認し、原ビデオの代表フレームを利用して原ビデオに接近する場合である。場合 4) と場合 5) とは、要約ビデオレベル 1 の再現で開始して経路は上述した場合と類似している。

【 0 0 6 5 】

このような本発明をサーバー/クライアント環境に適用すると、複数のクライアントが一つのサーバーに接近してビデオを概観及びブラウジングできるシステムを提供することができる。サーバーに原ビデオを受信して階層的な要約記述構造に基づいてビデオ要約記述データを生成し、前記原ビデオとビデオ要約記述データをリンクするビデオ要約記述データ生成システムを設ける。クライアントは通信網を通じてサーバーに接近し、ビデオ要約記述データを利用してビデオを概観して原ビデオに接近してビデオをブラウジング及びナビゲーションする。

【 0 0 6 6 】

本発明の技術思想は前記の好ましい実施例によって具体的に記述されたが、前記実施例はその説明のためのものであり、その制限のためのものでないことを注意すべきである。

また、本発明の技術分野の通常の専門家であれば本発明の技術思想の範囲内で様々な実施例が可能であるが理解できるであろう。

【 0 0 6 7 】

以上で説明したように本発明は、要約ビデオの生成と記述構造を通じてビデオ全体内容を速い時間に把握し、要約ビデオの各ビデオ区間の代表フレーム情報と代表音響情報を利用して効果的な階層的ブラウジングを可能にする。また、事件基盤の要約ビデオ記述を通じて事件及び主題による要約ビデオ及びブラウジング使用者に提供できる使用者注文形の機能も含む。

【図面の簡単な説明】

【図 1】 本発明による記述構造 (description scheme:DS) によってビデオ要約記述データを生成するためのシステムを示したブロック図である。

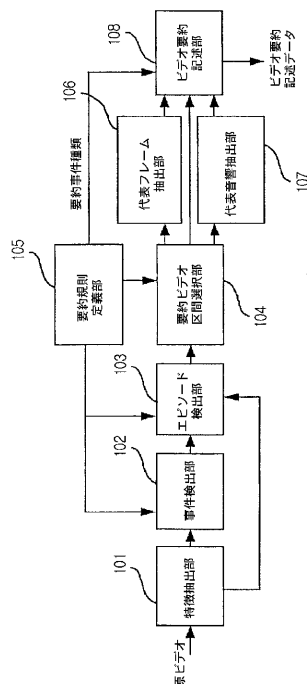
10

【図 2】 本発明による要約ビデオを記述するための階層的記述構造の資料構造を UML (Unified Modeling Language) で示した図面である。

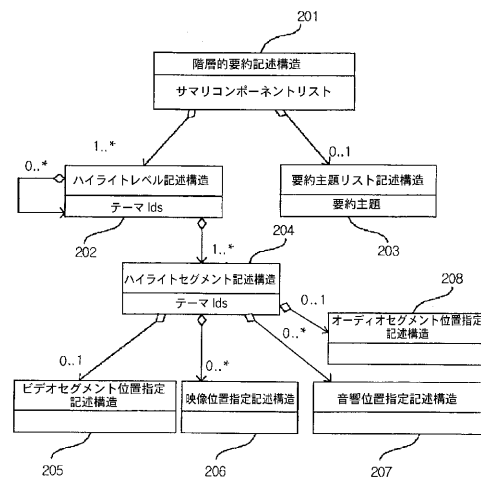
【図 3】 本発明による要約ビデオの再現及びブラウジングツールの使用者インターフェースの一実施例である。

【図 4】 本発明によるビデオ要約記述データを利用した階層的ブラウジングのためのデータ及び制御流れに関する構成図である。

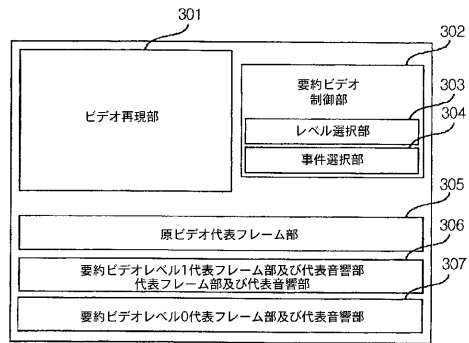
【図 1】



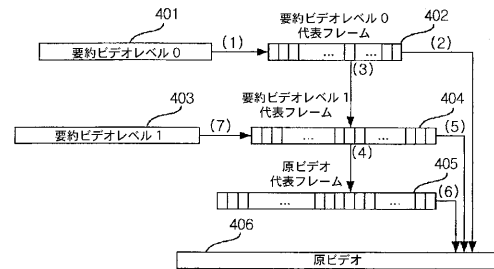
【図 2】



【図 3】



【図 4】



フロントページの続き

(72)発明者 キム ゼ ゴン

大韓民国 302-777 デゾンシ ソ-グ ドウンサン-ドン 908 サンムリ アパート
メント 203-402

(72)発明者 ジャン ヒョン ソン

大韓民国 302-741 デゾンシ ソ-グ マンヨン-ドン ガンピョン アpartment
112-506

(72)発明者 キム ムンチョル

大韓民国 302-734 デゾンシ ソ-グ ドウンサン 2-ドン 912 ドウンジ アパ
artment 110-405

(72)発明者 キム ジン ウン

大韓民国 305-761 デゾンシ ユソン-グ ジョンミン-ドン エクスポ アパートメン
ト 305-1603

審査官 鈴木 和樹

(56)参考文献 特開平11-088807(JP, A)

国際公開第99/032993(WO, A1)

橋本 隆子, 番組インデックスを利用したダイジェスト視聴方式の検討, 映像情報メディア学会
技術報告, 日本, 社団法人映像情報メディア学会, 1999年 3月24日, 第23巻 第28号

(58)調査した分野(Int.Cl., DB名)

G06F 17/30

H04N 5/91