

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4351994号
(P4351994)

(45) 発行日 平成21年10月28日(2009.10.28)

(24) 登録日 平成21年7月31日(2009.7.31)

(51) Int. Cl. F I
G06F 17/30 (2006.01) G06F 17/30 170D
 G06F 17/30 220A

請求項の数 8 (全 13 頁)

(21) 出願番号	特願2004-504237 (P2004-504237)	(73) 特許権者	503003854
(86) (22) 出願日	平成15年5月7日(2003.5.7)		ヒューレット-パカード デベロップメント カンパニー エル. ピー.
(65) 公表番号	特表2005-524916 (P2005-524916A)		アメリカ合衆国 テキサス州 77070
(43) 公表日	平成17年8月18日(2005.8.18)		ヒューストン コンパック センタ ド
(86) 国際出願番号	PCT/US2003/014528		ライブ ウェスト 11445
(87) 国際公開番号	W02003/096350	(74) 代理人	110000039
(87) 国際公開日	平成15年11月20日(2003.11.20)		特許業務法人アイ・ピー・エス
審査請求日	平成17年8月24日(2005.8.24)	(72) 発明者	ペン・ウー
(31) 優先権主張番号	10/140,794		アメリカ合衆国カリフォルニア州マウンテンビュー ギルモアストリート1540
(32) 優先日	平成14年5月7日(2002.5.7)	(72) 発明者	イーン・リ
(33) 優先権主張国	米国 (US)		アメリカ合衆国カリフォルニア州ロサンジエルス エイピーティ214 エレンデイルブレース2677

最終頁に続く

(54) 【発明の名称】 スケーラブルなビデオの要約化

(57) 【特許請求の範囲】

【請求項1】

コンピュータにビデオシーケンス(10)を要約させる方法であって、
 ビデオシーケンス(10)を一連の要素シーン(12)に分解するステップと、
 前記一連の要素シーン(12)における要素シーンのそれぞれに、前記要素シーンのそれぞれが有する特徴に基づいて重要値を割り当てるステップと、

前記要素シーン(12)をシーンセットとして扱い、前記シーンセットが決められた数より多いときのみ、2つの隣接するシーンセットの対全てにおいて、前記重要値を合算して重要値和を算出し、前記重要値和が最小である2つの隣接するシーンセットの対を併合して1つシーンセットとして扱い、前記併合を1回以上実行して予め決められた数以下のシーンセットを得るステップと、

前記シーンセットを得るステップにより得られたシーンセットのうち、1つの要素シーン(12)を含むシーンセットそれぞれにおいて、前記要素シーン(12)に含まれる複数のフレームからキーフレーム(18)を抽出するステップと、

前記シーンセットを得るステップにより得られたシーンセットのうち、複数の要素シーン(12)を含むシーンセットそれぞれに、新しいキーフレームを構築するステップであって、

それぞれ複数のフレームを有する要素シーン(12)を複数含むシーンセットから、予め決められた数のキーフレーム(18)を抽出するステップと、

前記抽出されたキーフレーム(18)の重要値に基づいて、前記抽出されたキーフレー

ム(18)を、割り当てられた表示エリア(30)に配置してサイズ変更することにより、前記割り当てられた表示エリア(30)に隙間なくパッキングするステップとを含むステップとを前記コンピュータに実行させる方法。

【請求項2】

前記重要値和を算出することを、繰り返し行うステップと、予め決められた数のシーンセットが残るまで、前記併合を繰り返し行うステップとをさらに含む請求項1に記載の方法。

【請求項3】

前記パッキングするステップは、前記割り当てられた表示エリア(30)に配置してサイズ変更するために、前記キーフレーム(18)を、前記割り当てられた表示エリア(30)をカバーする格子空間に割り当てるステップと、前記割り当てられた前記キーフレーム(18)の前記重要値に基づいて、前記格子空間を所定のサイズに拡張するステップとを含む請求項1に記載の方法。

【請求項4】

前記格子空間を所定のサイズに拡張するステップは、拡張された格子空間を避けるように、拡張されない格子空間を押しやるステップと、前記割り当てられた表示エリア(30)外に押しやられた格子空間を破棄するステップとを含む請求項3に記載の方法。

【請求項5】

前記パッキングするステップは、前記キーフレーム(18)の構築のために、前記キーフレーム(18)をそれぞれの重要値により順序付けるステップと、前記キーフレーム(18)それぞれを、予め定められ、前記キーフレーム(18)の鮮明さを示す複数の可視レベルのうちの1つに量子化するステップとを含む、前記割り当てられた表示エリア(30)に配置してサイズ変更するために、前記抽出されたキーフレーム(18)を、表示エリア(30)をカバーする格子空間に割り当てるステップと、前記割り当てられたキーフレーム(18)の前記可視レベルに基づいて、前記格子空間を所定のサイズに順次拡張するステップとを含む、前記パッキングするステップは、拡張された格子空間を避けるように、拡張されない格子空間を移すステップと、前記表示エリア(30)外に移された格子空間を破棄するステップとをさらに含む請求項1に記載の方法。

【請求項6】

9個のキーフレーム(18)が、前記シーンセットから抽出される請求項1に記載の方法。

【請求項7】

前記キーフレーム(18)は、それぞれ3つの可視レベルのうちの1つに量子化される請求項5に記載の方法。

【請求項8】

前記複数の可視レベルは、

10

20

30

40

50

それぞれ関連する表示サイズ

を有し、

最大表示サイズは、最小表示サイズの少なくとも2倍である

請求項5に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

[発明の分野]

本発明は、包括的にはビデオ材料の要約化に関し、詳細にはビデオ要約化プロセスの自動化およびカスタマイズに関する。

10

【背景技術】

【0002】

[発明の背景]

デジタルビデオは、コンピュータ業界および電気通信業界で急成長している要素である。

多くの企業、大学、さらには家庭がすでに、アナログフォーマットのビデオおよびデジタルフォーマットのビデオの両方を大量に収納している。

例としては、放送ニュースに使用されるビデオ、訓練・教育ビデオ、セキュリティ監視ビデオ、およびホームビデオが挙げられる。

デジタルビデオの急速な発達により、多くの人々がマルチメディアを取り込みマルチメディアと対話する方法が変化しつつあり、その過程の中で多くの新しいニーズおよび用途がもたらされた。

20

【0003】

このような一用途はビデオ要約である。

ビデオ要約は、名称が示すようにより長いビデオシーケンスの内容を短くまとめたものであり、オリジナルの本質的なメッセージを良い状態で保持しながら、ビデオシーケンスの内容についての簡潔な情報をユーザに提供する。

理論上、ビデオ要約は手動で、または自動的に生成することができる。

しかし、大量のビデオデータがすでに存在し、作成される新しいビデオデータの量が今後も増え続けることから、ビデオ要約を手動で生成することはますます困難になっている。

30

したがって、ビデオ要約プロセスにおいて人間の介入を減らすように、完全に自動化されたビデオ解析・処理ツールを開発することがますます重要になりつつある。

【0004】

2つの基本的に異なる種類のビデオ要約、すなわち静止画像要約および動画像要約がある。

静止画像要約は、ビデオサマリとも呼ばれ、基礎となるビデオソースから抽出または生成された目立った画像（キーフレームとして知られる）の小さな集まりである。

動画像要約は、ビデオスキミングとも呼ばれ、画像シーケンスの集まり、ならびにオリジナルシーケンスから抽出された対応するオーディオ要約からなり、したがってそれ自体、長さのかなり短いビデオクリップである。

40

一般にビデオサマリは、視覚的情報のみが利用され、オーディオ情報またはテキスト情報を処理する必要がないため、スキミングよりもはるかに素早く構築することができる。

したがってビデオサマリは、タイミングすなわち同期の問題がないことからより容易に表示することができる。

さらに、抽出されたすべての代表フレームの時間順を空間順で表示することができ、したがって、ユーザはビデオの内容をより素早く把握することができる。

最後に、必要なときに、ビデオサマリ中の抽出静止画像をすべて非常に容易にプリントアウトすることができる。

【0005】

50

ビデオ要約化は、いずれの記憶媒体（テープ、ディスク等）中のビデオシーケンスにも適用して使用することが可能であるが、対象となる一般的な記憶媒体はDVDビデオディスクである。

DVDビデオは、人々がマルチメディア情報を利用する方法を劇的に変えつつある。

DVDビデオディスクの大きな記憶容量は、静止画像、テキスト、ビデオ、およびオーディオに対して理想的な記憶場所を提供する。

DVDビデオフォーマットによりサポートされるナビゲーション機能により、媒体の内容に対話的にアクセスすることができる。

DVDディスクに記憶することができる様々な媒体タイプに適應するために、DVDビデオフォーマット仕様に従って媒体を編成（organize）し、このように編成された媒体の内容をDVDディスクにエクスポートする技術に対するニーズが増している。

この技術は一般に「DVDオーサリング」と呼ばれ、DVDオーサリングの本質的な一タスクは、ビデオソースからDVDビデオタイトルおよびナビゲーション構造を作成することである。

【0006】

DVDビデオタイトル構造は主に、対話型閲覧用にビデオの内容を編成するために使用される2つのエンティティ、すなわちタイトルおよびチャプタからなる。

DVDディスクのフォーマットにより、DVDディスクは最大で99個のタイトルを収容することが可能であり、1タイトルは最大で99個のチャプタを収容することが可能である。

したがってタイトルおよびチャプタは、全体のビデオシーケンスを意味のある部分に分け、各タイトルおよび/またはチャプタが1つの特定のビデオ部分の入口点である。

【0007】

DVDオーサリングでは、ビデオシーケンスからタイトルおよびチャプタの構造を自動的に作成することが非常に重要である。

たとえば、Hewlett PackardのMyDVDアプリケーションでは、ユーザがビデオからDVDを自動的に作成することを選択すると、シーン検出アルゴリズムに基づいてシーンが検出されたときに新しいチャプタが作成される。

次に、検出された各シーンからキーフレームが抽出される。

キーフレームは、基礎となるシーンを表し、DVDナビゲーションボタンにリンクされ、したがってユーザは、キーフレームを閲覧してビデオシーケンスの内容を素早く把握し、関連するボタンをクリックして対応するシーンを見ることができる。

【特許文献1】米国特許第5995095号

【特許文献2】欧州特許出願公開第1045316号

【特許文献3】国際公開第03/096229号

【非特許文献1】Uchihashi S 他, 「Summarizing Video Using a Shot Importance Measure and a Frame-Packing Algorithm」, ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 1999年会報, 1999 IEEE INTERNATIONAL CONFERENCE ON PHOENIX, AZ USA, 1999年3月15-19日, Piscatway NJ USA, IEEE, US, 1999年3月15日, p3041-3044

【非特許文献2】M. M. YeungおよびB. L. Yeo, 「Video Visualization For Compact Presentation and Fast Browsing of Pictorial Content」, 1997年10月, IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, no. 5

【発明の開示】

【発明が解決しようとする課題】

【0008】

時には、検出されたシーンの数が好ましい、または許可されているチャプタの数を上回る場合がある。

したがって、検出されたシーンを知的に併合する方法が必要である。

さらに、併合されたいくつかのオリジナルシーンがチャプタに含まれる場合、有意かつ有益なキーフレームを構築して基礎となる併合されたビデオの内容を表す方法が必要であ

10

20

30

40

50

る。

【課題を解決するための手段】

【0009】

[発明の概要]

ビデオシーケンスにおいて検出されたシーンを知的に併合し、キーフレームを構築して基礎となる併合されたビデオの内容を表す方法およびシステムを本明細書に記載する。

本発明による一実施形態では、ビデオシーケンスが一連の構成要素シーンに分解され、次に、構成要素シーン対が、所定数のシーンセットが残るまで併合される。

単一の構成要素シーンを含む各シーンセットからキーフレームが抽出され、複数の構成要素シーンを含む各シーンセットに新しいキーフレームが構築される。

10

【発明を実施するための最良の形態】

【0010】

[好ましい実施形態の説明]

以下の好ましい実施形態の詳細な説明では、本明細書の一部を成し、本発明を実施し得る特定の実施形態を例として示す添付図面を参照する。

他の実施形態も利用することができ、構造的または論理的な変更を本発明の範囲から逸脱することなく行い得ることを理解されたい。

したがって以下の詳細な説明は限定の意味で解釈されるべきではなく、本発明の範囲は添付の特許請求の範囲によって規定される。

【0011】

20

本明細書における本発明の説明では、ビデオシーケンスからDVDタイトル・チャプタ構造（本明細書ではビデオサマリと呼ぶこともある）をどのように自動的に作成するかという問題に対処する。

本明細書での説明は、各ビデオシーケンスが1つのタイトルの下にあると想定すると共に、そのタイトルの下にあるシーンを表すチャプタをどのように作成するかに焦点をあてる。

同様の手法を用いて、ビデオシーケンス中のシーンを表す複数のタイトルを作成することができ、この場合、チャプタを使用して各シーンをさらに細分することができることを意図する。

さらに、本明細書に記載の本発明はDVDオーサリングとの併用に限定されないことが理解される。

30

むしろ本発明は、検出されたシーンの数が好ましい数または許可されている数を上回り、シーンの併合が望ましい、または必要ないずれのビデオ要約にも採用することが可能である。

【0012】

通常、図1に見られるように、ビデオシーケンス10は一連のシーン12、ショット14、およびフレーム16に基づいた階層ツリーとして表すことができる。

本明細書で使用される場合、ショット14は連続した撮影期間中に取り込まれたビデオセグメントとして定義され、シーン12は意味的に関連するショット14のセットからなる。

40

ショット14は一連の個々のフレーム16からなる。

ショットおよびシーンを検出する様々なアルゴリズムが存在し、そのうちのいずれもビデオシーケンスをシーン、ショット、およびフレーム構造に分解するにあたって使用するのに適している場合がある。

解析されるビデオシーケンス10のタイプに応じて、或るアルゴリズムが別のアルゴリズムよりも好ましい場合がある。

【0013】

DVDディスクの文脈の中では、シーン12は意味的な基本単位とみなされ、DVDディスクのチャプタにマッピングされる。

1つまたは複数のキーフレーム18が、様々な特徴に基づいた基礎となるビデオの内容

50

の複雑性に依存する均一サンプリング (uniform sampling、均等サンプリング) または適応的方式によって各シーン 12 から抽出される。

したがって、ビデオシーケンス 10 は、図 2 に示すように各シーン 12 (またはチャプタ) を 1 つのキーフレーム 18 を使用して表すことによって要約される。

シーン 12 (ひいては抽出されるキーフレーム 18) の数が予め設定された (present を preset として訳) 限度を超える場合は、シーン 12 の数が予め設定された限度内になるまでいくつかのシーン 12 を併合しなければならない。

【 0 0 1 4 】

異なるシーン 12 は閲覧者に対して異なる視覚的インパクトを有する場合があるため、各シーン 12 に重要値が割り当てられる。

シーン、ショット、およびフレームの重要値を計算する様々な方法およびアルゴリズムが知られている。

たとえば、本明細書と同じ日に出願され、同一の譲受人に譲渡された「Scalable Video Summarization and Navigation System and Method」という名称の同時係属中の米国特許出願第 10 / 140511 号 [HP ファイル番号 10019975 - 1] が参照により援用され、シーン、ショット、およびフレームの重要値を計算するアルゴリズムを記載している。

重要値の計算には、シーンおよびショットの長さ、シーンおよびショットのアクティビティレベル、シーンに含まれる構成要素ショットの数、検出されたカメラの動き、フレーム中の皮膚色のピクセルの割合、フレーム中の検出された人の顔の数、フレーム中のエッジエネルギーの分布、およびフレームに含まれる動きアクティビティ量等の要因を考慮することができる。

【 0 0 1 5 】

本明細書において本発明を説明するために、本説明全体を通して以下に定義するシーン構造を使用する。

m_startFrame はシーンの開始フレームであり、

m_endFrame はシーンの終了フレームであり、

m_containedFrames[] はシーンに含まれるフレームの配列であり、

m_keyframe はシーンから抽出された代表フレームであり、

m_import はシーンの重要値である。

【 0 0 1 6 】

上に定義したように、値 m_import はシーンの重要度である。

アルゴリズムの説明を容易にするために、キーフレームも同様にこの重要値を有するものと想定することができる。

しかし、ビデオシーケンスのフレームは、シーン 12 中のすべてのフレーム 16 からのキーフレーム 18 の選択に使用される、計算されたそれぞれの重要値を有することができることが認識されよう。

【 0 0 1 7 】

[対ごとの繰り返しシーン併合]

図 3 は、本発明による対ごとの繰り返しシーン併合方法を示す。

構成要素シーンに分解されたビデオシーケンス 10 を考えると、N 個のシーンが検出され、S [1], 1 = 1, . . . , N と表され、各シーン S [1] は上に定義したシーン構造を有するものと想定する。

さらに、好ましいシーン (またはチャプタ) の数は M であり、 $M < N$ であると想定する。

次いで、以下の対ごとの繰り返しシーン併合方式を使用して M 個のチャプタを作成することができる。

【 0 0 1 8 】

本明細書に述べる対ごとの繰り返しシーン併合はシーンの重要値に基づく。

N 個のシーンが、結果得られるシーン 12 の数が好ましい数である M 個のチャプタに見

10

20

30

40

50

合うまで繰り返し併合される。

各最終シーンは、1つまたは複数のオリジナルシーンを含み得るため、本明細書ではシーンセットと呼ぶ。

【0019】

各繰り返し中に、2つの隣接するシーン12が選択されて併合される。

したがって各繰り返し後、シーンの数は1だけ低減することになる。

併合するシーン12を選択する基準はそれぞれの重要値に基づく。

具体的には、重要値の和が最小になる隣接シーン12対が併合されることになる。

【0020】

たとえば、第1の繰り返しでは、隣接するすべてのシーン対は{(S[1], S[2]), (S[2], S[3]), ..., (S[N-1], S[N])}であり、総計でN-1個の対になる。

各対の重要値は、

$$I_pair[i] = S[i].m_import + S[i+1].m_import$$

として計算される。

但し、 $i = 1, \dots, N - 1$ である。

$I_pair[i]$ を最小にする2つの隣接するシーン12が併合され、次の繰り返しでは、重要度が、和である $I_pair[i]$ である単一のシーンとしてみなされる。

隣接するシーンの繰り返し併合は、シーンをN-M回繰り返し併合した後、M個の最終シーンセットが残るまで続けられる。

1つのみのシーンを含むシーンセットごとに、シーンはチャプタに直接マッピングされる。

しかし、複数のシーンを含むシーンセットの場合は、シーンセットからチャプタを作成する方法が必要である。

【0021】

キーフレームのサイズ変更およびパッキング

図4は、シーンセットS'が複数の構成要素シーン12を含む場合に使用される本発明によるキーフレームのサイズ変更およびパッキング方法を示す。

$S[1] \sim S[1+m-1]$ のm個のシーンを含むシーンセットS'を考えると、以下の公式を使用してシーン構造に記憶される情報を更新する。

【0022】

【数1】

$$S'.m_startFrame = S[l].startFrame$$

$$S'.m_endFrame = S[l+m-1].endFrame$$

$$S'.m_containedFrames[] = S[l].m_containedFrames[] \otimes \dots \otimes S[l+m-1].m_containedFrames[]$$

【0023】

但し、

【0024】

10

20

30

40

【数 2】



【0025】

は連結 $S'.m_import=S[l].m_import+\dots+S[l+m-1].m_import$ を表す。

【0026】

最終操作は、シーンセット S' に含まれるすべてのシーンを表すことが望ましいキーフレーム18の更新である。

一方策は、シーンセット中で最も重要なシーンのキーフレーム18を抽出するというものである。

しかしこの方策は、シーンセット S' 中の他の構成要素シーン12からの情報が失われることになる。

より良い方策は、シーンセット中のすべての構成要素シーンからできるだけ多くの情報を包含するというものである。

この目的を達成するために、キーフレームのサイズ変更およびパッキング方式を用いてシーンセット S' に新しいキーフレーム18を構築する。

【0027】

ここに述べるキーフレームのサイズ変更およびパッキング方法は、それぞれの重要値に基づいてシーン $S[l]$ 、 $S[l+1]$ 、 \dots 、 $S[l+m-1]$ のすべてのキーフレーム18をサイズ変更し、予め割り当てられた表示エリアにパッキングする。

重要値の大きなキーフレーム18ほど、重要値の低いキーフレームよりも基礎となるビデオの内容を良好に表しているものと想定される。

したがって、高い重要値を有するキーフレーム18ほど大きな表示サイズが割り当てられ、閲覧者はビデオシーケンスをより良く理解することができる。

【0028】

複数のフレームをサイズ変更してパッキングする方策が研究されてきた。

しかし、既存の方法はDVDチャプタの作成には不適切である。

特に、一手法は抽出されたすべてのキーフレームをサイズ変更して表示エリアにパッキングする自動サイズ変更およびパッキングアルゴリズムである(S. Uchihashi およびJ. Foote著「Summarizing Video Using a Shot Importance Measure and a Frame-Packing Algorithm」ICASSP'99, vol. 6, 1999参照のこと)。

この方法は、構成要素キーフレームのサイズをもはや明確に認識することができない程度まで縮小する場合がある。

一方、比較的重要度の低いキーフレームの場合、より重要なキーフレームを認識できないほど小さくするという事までしてこういったフレームのために空間を設けることに価値はない。

別の手法では、いくつかのテンプレートが提供されて、抽出されたキーフレームをそれぞれの重要値に基づいて予め割り当てられた表示エリアにパッキングする(M. M. Yeung およびB. L. Yeo著「Video Visualization For Compact Presentation and Fast Browsing of Pictorial Content」IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, no. 5, Oct. 1997を参照のこと)。

しかし、この手法におけるテンプレートレイアウトはすべて固定され、基礎となるシーンの構造に適合するように調整することができない。

【0029】

本発明によれば、最も重要なキーフレーム18の可視性を確実にしながらシーンセット S' のキーフレーム18を適応的にパッキングすることができるキーフレームのサイズ変

10

20

30

40

50

更およびパッキング方法について説明する。

【0030】

キーフレームのサイズ変更およびパッキング方法は以下の2つの前提に基づく。

第1の前提は、シーンセットS'中の構成要素キーフレームをすべて最終的に構築されるキーフレームに含める必要はないことである。

本明細書での説明では、9枚の最も重要な構成要素キーフレームに対応する最大で9枚のキーフレームが構築される、と想定される。

もちろん、本明細書に記載の方法を用いて他の任意の数の構成要素キーフレームを構築することができ、想定される9枚の構成要素キーフレームは例示のみを目的として使用され、本発明に対する限定と解釈されるべきではない。

10

【0031】

第2の前提は、選択された構成要素キーフレームがそれぞれの重要値に基づいて3つの可視性レベルのうちの1つに量子化され、各可視性レベルには予め定められた表示サイズが割り当てられることである。

第1のレベルにおけるあらゆる構成要素キーフレームの表示サイズは「非常によく見える」である一方で、第2のレベルの構成要素キーフレームの表示サイズは「見える」であり、第3のレベルの場合は「ようやく見える」である。

所望であればより多数または少数のレベルを使用してもよく、3レベルを使用することは単なる例示である。

「非常によく見える」、「見える」、および「ようやく見える」の表示エリア同士のサイズの区別は経験的に求めることができる。

20

たとえば、本明細書では、「見える」表示サイズには「非常によく見える」表示サイズの75%が割り当てられ、「ようやく見える」エリアには「非常によく見える」表示サイズの50%が割り当てられる。

予め定められた表示エリアのサイズは予期される表示装置に依存する。

たとえば、ラップトップコンピュータでは、「非常によく見える」表示エリアはおよそ1平方インチ(6.45平方cm)であるが、大画面テレビセットの場合、「非常によく見える」表示エリアはおよそ3平方インチ(19.35平方cm)以上であることができる。

【0032】

30

以下の3つのステップがサイズ変更およびパッキングプロセスに採用される。

第1に、選択された構成要素キーフレーム(この例では9枚の構成要素キーフレーム)が降順でそれぞれの重要値によりソートされ、 s_1, s_2, \dots, s_9 として表される。

第2に、構築されたキーフレームの表示エリア30が 3×3 の格子に分割され、9枚の構成要素キーフレームが図5に示す順序に従って格子空間に割り当てられる。

返されるキーフレームが9枚よりも少ない場合、空いた格子空間は空のまま残される。

第3に、構成要素キーフレームが、それぞれの重要度順でそれぞれの予期される表示サイズ(すなわち、「非常によく見える」、「見える」、または「ようやく見える」)に拡張される。

40

特に、各構成要素キーフレームの量子化レベル「非常によく見える」、「見える」、または「ようやく見える」に基づいて、格子空間のエリアが予め割り当てられたサイズに拡張される。

拡張された格子空間が隣接する格子空間の1つに重複する場合、重複した格子空間は脇に「押しやられ」、拡張された格子空間を避ける。

格子空間が構築されたキーフレームの表示エリア30外に最終的に押しやられる場合、その格子空間および格子空間が含む構成要素キーフレームは単に破棄される。

図3に示す番号を有する格子空間は結果として、最高重要度の構成要素キーフレーム18が構築されたキーフレーム18に残される一方で、最低重要度の構成要素キーフレーム18がまず構築されたキーフレーム18の表示エリア30から押し出されることになる。

50

さらに、拡張プロセス後に表示エリア 30 にいくつかのエリアが空白のままである場合、構成要素キーフレーム 18 のサイズがそれぞれの重要値に基づいて調整されて、表示エリア 30 全体を充填する。

この方式を使用すると、シーンセット S' 中の最も重要な構成要素キーフレーム 18 が、構築されたキーフレーム 18 の表示エリア 30 により高い可視性をもって常に保持されることになる。

【0033】

したがって、本明細書に記載の本発明は、ビデオシーケンスのダイナミクスを考慮すると共に、基礎となる内容を良好に表す有意な DVD エンティティ (タイトルおよびチャプタ) を有するビデオサマリを自動的かつ知的に作成する。

10

本発明は所望数のチャプタを提供し、各チャプタに有意な代表キーフレームを構築する。

【0034】

図 6 に概略的に示すように、本明細書に記載のシステムは、選択されたプラットフォームが十分な計算力を有するプロセッサ 102、ビデオシーケンス 10、アルゴリズム、およびビデオサマリを記憶するデータ記憶システム 104、ならびにユーザがビデオサマリを見ることができる表示装置 108 を有する限り、家庭用コンピュータ 100 等の様々なコンピューティングプラットフォームで実施することができる。

データ記憶システム 104 は DVD ディスク、ハードディスクドライブ、もしくは他の永久記憶装置、または選択されたプラットフォームのランダムアクセスメモリであること

20

ができる。表示装置 108 はたとえば、ビデオモニタやプリンタによって生成されたハードコピーであることができる。

【0035】

特定の実施形態を好ましい実施形態の説明を目的として本明細書に図示し説明したが、本発明の範囲から逸脱することなく、図示し説明した特定の実施形態に代えて広範な代替および/または等価の実施態様を用い得ることが当業者により理解されよう。

コンピュータ分野および電気分野の当業者は、本発明を非常に広い範囲の実施形態で実施し得ることを容易に理解するであろう。

本願は、本明細書において説明した好ましい実施形態の適合形態および変形形態をいず

30

れも包含することを意図する。したがって、本発明は請求項および請求項の等価物によってのみ限定されることを明らかに意図する。

【図面の簡単な説明】

【0036】

【図 1】ビデオ階層構造の概略図である。

【図 2】DVD タイトル・チャプタ構造の概略図である。

【図 3】本発明による対ごとに繰り返し行われるシーン併合を示すフローチャートである。

【図 4】本発明によるキーフレームのサイズ変更およびパッキングを示すフローチャートである。

40

【図 5】本発明による新しいキーフレームの構築に使用可能な 1 つの格子である。

【図 6】本発明によるビデオサマリの作成に使用可能なコンピュータシステムの概略図である。

【符号の説明】

【0037】

10・・・ビデオシーケンス,

12・・・シーン 1,

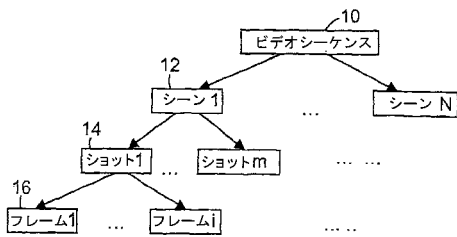
14・・・ショット 1,

16・・・フレーム 1,

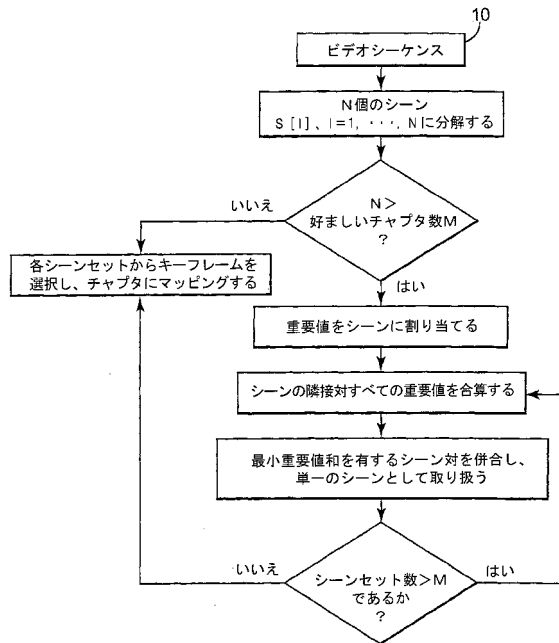
50

- 1 8 . . . キーフレーム 1 ,
- 1 0 . . . ビデオシーケンス ,
- 1 0 2 . . . データプロセッサ ,
- 1 0 4 . . . データ記憶装置 ,
- 1 0 8 . . . 表示装置

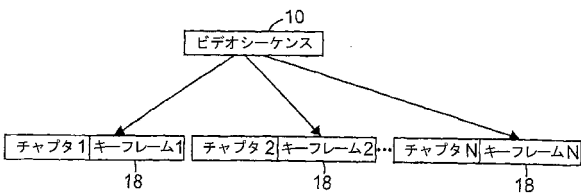
【 図 1 】



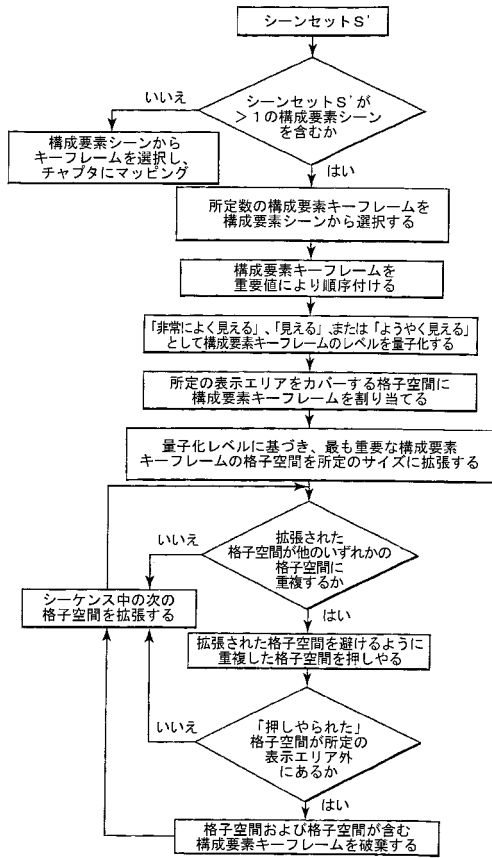
【 図 3 】



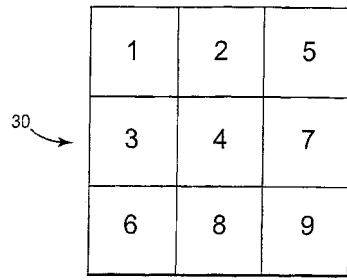
【 図 2 】



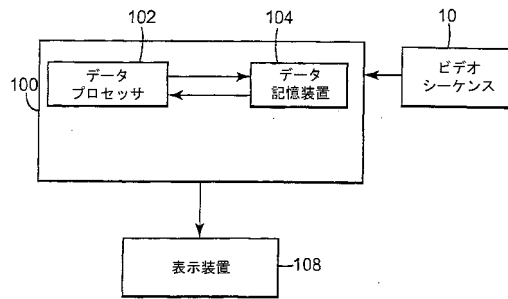
【図4】



【図5】



【図6】



フロントページの続き

(72)発明者 ダニエル・アール・トレッター
アメリカ合衆国カリフォルニア州サンノゼ カルデスチュアルダ 1 5 3 3

審査官 鈴木 和樹

(56)参考文献 特開平 1 0 - 2 3 2 8 8 4 (J P , A)
特開 2 0 0 0 - 3 0 8 0 0 8 (J P , A)
特開平 1 1 - 2 3 9 3 2 2 (J P , A)
特開 2 0 0 0 - 2 9 9 8 2 9 (J P , A)
特開 2 0 0 1 - 1 5 5 1 6 9 (J P , A)
特表 2 0 0 1 - 5 2 7 3 0 4 (J P , A)
特表 2 0 0 5 - 5 2 5 0 3 4 (J P , A)
是津耕司、外 2 名、時刻印付オーサリンググラフによるビデオ映像のシーン検索、情報処理学会
論文誌、日本、社団法人情報処理学会、1 9 9 8 年 4 月 1 5 日、第 3 9 巻、第 4 号、p . 9 2
3 - 9 3 2
Shingo Uchihashi、外 1 名、SUMMARIZING VIDEO USING A SHOT IMPORTANCE MEASURE AND A FRAM
E-PACKING ALGORITHM、Proc. of the IEEE Int'l Conf. on Acoustics, Speech, and Signal Pr
ocessing[online]、米国、IEEE、1 9 9 9 年 3 月 1 5 日、p . 3041~3044、[検索日：2008/8/
19]、U R L、<http://www.fxpai.com/publications/FXPAL-PR-99-134.pdf>

(58)調査した分野(Int.Cl.、D B 名)

G06F 17/30

NRIサイバーパテント