

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号  
特許第4250301号  
(P4250301)

(45) 発行日 平成21年4月8日(2009.4.8)

(24) 登録日 平成21年1月23日(2009.1.23)

(51) Int.Cl.

F I

HO4N 5/91 (2006.01) HO4N 5/91 N

HO4N 5/262 (2006.01) HO4N 5/262

HO4N 5/92 (2006.01) HO4N 5/92 H

請求項の数 5 (全 27 頁)

(21) 出願番号	特願2000-111125 (P2000-111125)	(73) 特許権者	000001007
(22) 出願日	平成12年4月12日 (2000.4.12)		キヤノン株式会社
(65) 公開番号	特開2000-350159 (P2000-350159A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成12年12月15日 (2000.12.15)	(74) 代理人	100076428
審査請求日	平成19年4月12日 (2007.4.12)		弁理士 大塚 康徳
(31) 優先権主張番号	PP9704	(74) 代理人	100112508
(32) 優先日	平成11年4月12日 (1999.4.12)		弁理士 高柳 司郎
(33) 優先権主張国	オーストラリア (AU)	(74) 代理人	100115071
(31) 優先権主張番号	PP9705		弁理士 大塚 康弘
(32) 優先日	平成11年4月12日 (1999.4.12)	(74) 代理人	100116894
(33) 優先権主張国	オーストラリア (AU)		弁理士 木村 秀二
(31) 優先権主張番号	PP9706		
(32) 優先日	平成11年4月12日 (1999.4.12)		
(33) 優先権主張国	オーストラリア (AU)		

最終頁に続く

(54) 【発明の名称】 映像シーケンスを編集する方法及びシステム

(57) 【特許請求の範囲】

【請求項 1】

少なくとも一つのクリップを有する映像シーケンスを編集する方法であって、  
それぞれ予め決められたセグメントの持続時間と前後の前記セグメントの間隔とを示す、複数の編集テンプレートの中から適用すべき編集テンプレートを選択し、前記選択した編集テンプレートと前記映像シーケンスの持続時間とに基づき、前記映像シーケンスを、前記間隔ごとに複数の前記セグメントに分割する適用工程と、  
前記映像シーケンスから前記分割された複数のセグメントを取り出して編集済みシーケンスを形成する編集工程とを有することを特徴とする方法。

【請求項 2】

前記編集テンプレートは、持続時間が異なる複数種類の前記セグメントの持続時間と、複数種類の前記セグメントのパターンとを示しており、前記適用工程は、前記編集済みシーケンスにおける前記複数種類のセグメントのパターンが前記定義されたパターンになるように、分割する前記セグメントの持続時間を前記複数種類のセグメントの中から選択することを特徴とする請求項 1 記載の方法。

【請求項 3】

前記適用工程は前記映像シーケンスのうち前記分割されたセグメントを示す編集表示リストを形成し、前記編集工程は前記編集表示リストに基づいて前記映像シーケンスから前記複数のセグメントを取り出して編集済みシーケンスを形成することを特徴とする請求項 1 又は 2 記載の方法。

## 【請求項 4】

少なくとも一つのクリップを有する映像シーケンスを編集するシステムであって、それぞれ予め決められたセグメントの持続時間と前後の前記セグメントの間隔とを示す、複数の編集テンプレートの中から適用すべき編集テンプレートを選択し、前記選択した編集テンプレートと前記映像シーケンスの持続時間とに基づき、前記映像シーケンスを、前記間隔ごとに複数の前記セグメントに分割する適用手段と、

前記映像シーケンスから前記分割された複数のセグメントを取り出して編集済みシーケンスを形成する編集手段とを有することを特徴とするシステム。

## 【請求項 5】

請求項 1 に記載の方法をコンピュータに実行させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

## 【発明の詳細な説明】

## 【0001】

## 【発明が属する技術分野】

本発明は生の動画場面の編集に関し、特に、所望の結果を提供するために生の場面の編集を容易にするようにフィルム又は映像情報から得られる画像クリップのシーケンスから情報を抽出することに関する。特定の実現形態は、殆ど関心のない場面部分の再生を回避するように再生時間を短縮しつつ生の場面の重要部分を捕捉するクリップのシーケンスを提供するための供給元画像資料の自動編集に関し、また、場面における重大イベントの識別、タイトルの配置及びプリントのための、オリジナルの場面を代表するフレームである一連の個別フレームの抽出に関する。

## 【0002】

## 【従来の技術】

生のビデオ又は映画フィルムからよどみない、リズムカルな編集結果を作成するに当たり、興味を引く面白い成果を得るためには、特殊な技術が必要である。フィルムを処理する場合、代表的には、コンピュータの助けを借りて画像のシーケンスを容易に操作できるように、フィルムを映像フォーマットに変換する。映像編集を利用して特定のシーケンスが完成したならば、オリジナルのフィルムを従来の方式に従ってカットし、繋ぎ合わせて、高品質の再生作品を確保できる。従って、このようなプロセスは（アナログであれ、デジタルであれ）映像の操作に関連しており、デジタルフィルム効果、編集及び音響デザインを含めたいくつかの分野における熟練した技術を必要とする。このような技術を一人の人間が身に付けていることはまれであり、それぞれの技術は高いレベルの訓練を要し、何年間もフィルム製作業界で働いて初めて身に付けられるものである。

## 【0003】

アマチュアがビデオを製作する場合、同一の材料を与えられたときにプロが達成できると思われる成果をそのまま実現するのに必要な時間、専門知識、精巧な機器を備えているケースはまれである。多くの場合、アマチュアの作品はビデオの関係者にのみ個人的に面白く、関係のない視聴者の興味は早い段階で画面から失われてしまう。そのような無関心さは、ある意味で「ごく普通の」オリジナル場面を面白味に富んだ最終編集バージョンに変えることのできる編集技法を十分に適用していないことによって引き起こされることが多い。アマチュアのビデオに欠けているプロの一般的な編集・製作の基本的技法には、興味を引くタイトルの挿入、リズム感のある編集方式、場面転換及びカットの適切な使用、サウンドトラック及びバックグランドトラック、色修正などのデジタル効果の適用、パーティクルアニメーション、異なる種類のショットの適用などがある。

## 【0004】

オリジナル場面の編集では、クリップが当初取り出された順序に対応するシーケンスでクリップを配置しなければならない。アマチュアとプロが同様に現時点で利用できるツールは、ビデオカードを伴って又はビデオカードなしでパーソナルコンピュータ（PC）で稼動し、編集を目的として線的な時刻表を管理するように構成されているソフトウェアを含む。オリジナル情報源テープから新たなテープへシーケンス付けを行うために、デュアル

10

20

30

40

50

ビデオカセットレコーダ（VCR）などのハードウェアを使用しても良い。いずれの方法をとっても、それぞれのクリップをシーケンス中の所定の場所に手作業で割り付けて行くという「手間のかかる」方式を必要とするため、編集は長い時間を要する作業である。ディゾルブ又はクロスフェードなどの場面転換も手作業で挿入しなければならない、多くの場合、コンピュータ補助製作装置に処理上の重い負荷を課すことになる。また、アマチュアのビデオ製作者は場面転換と、それらをどこで使用するべきかということを正しく理解する能力に欠けている場合が多く、そのため、製作システムからの資源の不適切な使用や、過剰な使用、即ち、垂れ流しになってしまい、とてもプロとは言えない仕上がりになる。現在のデュアルVCR方式には問題がある。例えば、編集が完了した後に、アマチュアがビデオの一部を手直したい場合、プロセス全体をもう一度実行しなければならない。

10

#### 【0005】

編集済みビデオにタイトルを挿入する場合にも、まず、新たなシーンの場所を判定するために、場面を解析しなければならない。この作業では、in-outポイントを記録しながら場面を慎重に検討しなければならないため、ビデオ製作者が利用可能である場面の量に比べて多少の時間を要し、更に、タイトルマットを挿入するのに、また時間がかかる。最良の結果を得るためには、新たなシーンを挿入するときに、映像の残り部分に至る交互の場面転換を挿入しなければならない。

#### 【0006】

「インタータイトル」としても知られている挿入タイトルは、以前、無声映画の製作に当たり、音声がなかったので視聴者に登場人物やストーリーに関する情報を伝えるのを助けるために使用されていた。今日の作品においても、アクションに対するコメントを容易にし、ユーモアを生み出し、時間や場所を設定し、また、全く異なるシーンの間につながりを生み出すために、挿入タイトルは使用されている。この挿入タイトルを作成する現在の方法は、植字工として知られている人によって実行されている。植字工は、映画又は作品の著者から内容を記した文書を渡される。植字工はそのテキストを写真製版や手書き或いはデスクトップパブリッシングシステムを使用して活字にする熟練工である。多くの場合、その言葉はストーリーの内容を知っている著者が考えたもので、ウィットに富んだ言い回しで書かれている。或いは、場所や時間のセッティングを伝える場合には、一般に直接的で、情報を提供する言葉となる。挿入タイトルは編集者の短いリストに組み込まれ、その後、映画の中に順次挿入される。挿入タイトルの持続時間は、主に、視聴者が理解しなければならない語と音節の数に従って設定される。作品のジャンルやスタイルによってもタイトルの時間は変わり、編集者が映画の中における連続性をどの程度まで維持できるかによっても異なる。

20

30

#### 【0007】

##### 【発明が解決しようとする課題】

従って、従来の方式により挿入タイトルを作成するには、それぞれが特殊技能を備えた多くの人が必要である。挿入タイトル向けのテキストを書くためには、映画のストーリー、ジャンルに関する知識と、視聴者の教養の程度に対する理解が必要である。映画のジャンルを良く表すようにテキストを活字組みするには、特殊なデザイン技術が必要であり、挿入タイトルを次々に流れて行く映画の適切な場所に挿入するためには、編集者の特殊な技能が要求される。そのため、挿入タイトルの作成は複雑で、費用もかかり、長い時間を要するプロセスである。

40

#### 【0008】

現在の音声編集方法はきわめて特殊な方法であり、最終編集後のリズムにビート同期で装飾音を付加するという概念は大半のアマチュアビデオ製作者の範囲をはるかに超えている。選択されたサウンドトラックの音声波形を解析し、映像カットを同期させるのに要する時間は論外のものであり、機器は多くのアマチュアにとってとても正当であるとは言えないほど高価であり、その技法はデュアルVCRによる編集で管理するには困難すぎる。

#### 【0009】

本発明の目的は、アマチュアであっても、簡単に最適な編集シーケンスの形成を可能と

50

することである。

【 0 0 1 0 】

【課題を解決するための手段】

本開示の一態様によれば、少なくとも一つのクリップを有する映像シーケンスを編集する方法であって、それぞれ予め決められたセグメントの持続時間と前後の前記セグメントの間隔とを示す、複数の編集テンプレートの中から適用すべき編集テンプレートを選択し、前記選択した編集テンプレートと前記映像シーケンスの持続時間とに基づき、前記映像シーケンスを、前記間隔ごとに複数の前記セグメントに分割する適用工程と、前記映像シーケンスから前記分割された複数のセグメントを取り出して編集済みシーケンスを形成する編集工程とを有することを特徴とする方法が提供される。

10

【 0 0 1 3 】

【発明の実施の形態】

以下、図面を参照しながら本発明に係る実施の形態を詳細に説明する。

【 0 0 1 4 】

本実施形態は、満足な再生を可能とするために生の映像場面の自動編集を補助することを全て目的とする多くの態様を含む。一つの態様においては、自動編集ツールは視聴者に面白くないと受け取られるおそれのある、又は視聴者の注目の限界を超える、過剰に長い映像カットを取り込むことを回避しつつ、オリジナルの生の場面の本質を捉えた編集済みのバージョンを提供するように、生の場面のリズム感のある編集を実行する。別の態様においては、生の場面のスチルショットによる概要を形成できるように、生の場面を代表する複数の個別フレームを映像カットの中から取り出す構成が提供される。更に別の態様においては、生の場面の様々に異なる段階を区別するために、編集済みのバージョンに挿入タイトルを配置する方法が開示される。

20

【 0 0 1 5 】

図 1 A 及び図 1 B を参照すると、代表的な映像場面は、デジタルビデオカメラ 1 0 又はアナログビデオカメラ 2 0 の何れか一方から得られる。デジタルビデオカメラの場合、レコードボタン 1 2 を押すと、磁気テープ、磁気ディスク及び / 又は半導体メモリに代表される記録媒体に映像信号がデジタル記録される。デジタルビデオカメラの特有の利点の 1 つは、デジタル映像が記録されることを含め、またそれに加えてメタデータ 1 4 をカメラ 1 0 により自動的に生成でき、及び / 又はカメラ 1 0 のユーザの手動操作により生成できる構成を内蔵していることである。デジタルビデオカメラ 1 0 からはデジタル映像場面 1 6 が出力されるが、これは、図 1 A では番号 1、2、3、4、... により表されている複数の個別のクリップにより代表される。通常、映像と共に記録されるメタデータは個々のクリップの開始と中断に対する参照ポイントを含み、それらの参照ポイントは記録が実行された特定の実時間と関連している場合が多い。それらの時間、及び日付は自動的に記録されても良い。例えば、ユーザによって入力されるか、又は他のメタデータツールによって生成されるその他の詳細な情報として、場所（例えば GPS ロケータ装置により提供される）に関するデータ及び / 又はその時間に記録されるべきイベント情報及びその他の所望の詳細な情報が考えられる。自動的に生成されたメタデータは、典型的にはレコードボタン 1 2 を押す及び / 又は離す時点と同時に、クリップシーケンス 1 6 に挿入されるか、クリップシーケンス 1 6 と関連付けられる。このように、メタデータはクリップ及び / 又はその内容の特徴を示す情報の格納庫となる。

30

40

【 0 0 1 6 】

図 1 B を参照すると、アナログビデオカメラ 2 0 は、典型的には磁気テープ記録媒体などへの映像場面の記録を可能にするレコードボタン 2 2 を含む。記録された場面の再生及び / 又は編集のためにカメラ 2 0 から信号 2 4 が出力される。この信号 2 4 は、従来、記録された場面全体にわたり個々のクリップの開始又は中断に関する標識を全く伴わずに供給されていた。これは、異なるクリップを認識するための特定のメカニズムを通常含まない従来のセルロイドフィルムと同じ効果を持つ。この点に関して、従来のフィルム撮影の開始時に撮影された従来の「クリップボード」は、そのフィルムの編集担当者によって手作

50

業で識別される１つの標識となっており、これは、いずれか１つのクリップの単なる識別用ではなく、フィルムと音声を同期させるために特別に与えられたものである。

【００１７】

カメラ２０から取り出されたアナログ映像又はフィルム映像２６のいずれかをデジタルビデオカメラ１６と同様の方式で処理するためには、それぞれの信号をデジタル画像信号に変換するディジタイザ２８に信号２４又はフィルム映像２６のそれぞれを適宜入力しなければならない。ディジタイザ２８の出力はクリップ検出器３０に供給され、クリップ検出器３０はクリップ間の場面転換を検出し、メタデータを形成する。そのメタデータは加算器３２でディジタイザ２８の出力と組み合わせられて、デジタルビデオカメラ１０から取り出される信号１６に匹敵する有効性を備えたデジタルビデオ信号３４が形成される。

10

【００１８】

以上説明した構成は、Microsoft Corporationにより開発されたWindows（登録商標）オペレーティングシステムの環境をホストとするコンピュータアプリケーションプログラムとして実現されても良い。しかし、その他のオペレーティングシステムをホストとするコンピュータシステムにおいても上述の実施形態を実現し得ることは当業者には了承されるであろう。例えば、UNIX（登録商標）、OS/2（登録商標）、DOS（登録商標）を実行するコンピュータシステムで好適な実施形態を実行することができる。アプリケーションプログラムは、マウスやキーボードの操作に応答するメニュー項目及び制御要素を含むユーザインタフェースを有する。また、アプリケーションプログラムはホストコンピュータに直接に接続されるか又はネットワークを介してアクセスされる１台又は複数台の表示装置、プリンタ又は記憶装置に処理済データを送信する能力を有する。更に、アプリケーションプログラムは接続されるデジタル通信ネットワーク（例えば「インターネット」）との間でデータを送受信する能力を有する。

20

【００１９】

上述した構成は、図１０に示すコンピュータシステム４０のような従来の汎用（ホスト）コンピュータシステムを使用して実施できる。他の図面を参照して、更に説明されるべき上述のアプリケーションプログラムはコンピュータシステム４０で実行されるソフトウェアとして実現される。コンピュータシステム４０はコンピュータモジュール４１と、キーボード４２及びマウス４３などの入力装置と、プリンタ５７及び音声視覚出力装置（ビデオ表示装置）５６を含む出力装置とを備える。変復調器トランシーバ装置（モデム装置）５２は、例えば電話回線又は他の機能媒体を介して接続可能である通信ネットワーク５９との間で通信するために、コンピュータモジュール４１によって使用される。このモデム５２は、インターネット及び他のネットワークシステムに対するアクセスを得るために使用できる。

30

【００２０】

代表的な構成として、コンピュータモジュール４１は、少なくとも１つのプロセッサユニット４５と、例えば半導体ランダムアクセスメモリ（RAM）及び読み取り専用メモリ（ROM）から形成されるメモリユニット４６と、出力インタフェース４７を含む入出力（I/O）と、キーボード４２、マウス４３及びオプションであるジョイスティックに対応するI/Oインタフェース４８とを含む。出力インタフェース４７は、代表的にはビデオ表示装置及びスピーカ構成を組み込んだ音声視覚出力装置５６に結合している。記憶装置４９が設けられており、代表的な構成としては、ハードディスクドライブ５３と、フロッピーディスクドライブ６４とを含む。CD-ROMドライブ５５は、代表的には不揮発性データ供給源として設けられる。コンピュータモジュール４１の構成要素４５から４９及び５３から５５は、典型的には接続バス５０を介して、当業者に知られているコンピュータシステム４０の従来通りの動作モードをもたらすように通信する。この実施形態を実施できるコンピュータの例は、IBM-PC及びコンパチブル、Sun Sparcstations又はそこから発展した類似のコンピュータシステムを含む。好適な実施形態の代表的なアプリケーションプログラムはハードディスクドライブ５３に常駐しており、プロセッサ４５を使用して読み取られ、制御される。プログラムや、取り出されるデータの間格納は、可能

40

50

であればハードディスクドライブ 5 3 と協調して、半導体メモリ 4 6 を使用して実行しても良い。場合によっては、アプリケーションプログラムを C D - R O M 又はフロッピーディスクに符号化した形態でユーザに供給しても良いが、ユーザがモデム装置 5 2 を介してネットワークからアプリケーションプログラムを読み取ることもできる。

#### 【 0 0 2 1 】

特に、デジタル音声ストリーム 1 6 又は生の場面 2 4 は、コンピュータネットワーク及びモデム 5 2 を介することを含めた何らかの適切な方法によるか、又は C D - R O M ドライブ 5 5 などの可搬形メモリ装置によってコンピュータに提供されるか、或いは、例えば I / O インタフェース 4 8 の「映像」入力端子に直接提供される。このように、各々のクリップを含む生の映像場面の全体を利用して、コンピュータ 4 1 内でコンピュータ処理を実行することができる。

10

#### 【 0 0 2 2 】

図 1 0 に示すように、モデム装置 5 2 により、映像とそれに付随するサウンドトラックの双方を含むデジタル映像情報の供給源として機能し得るネットワーク 5 9 への接続が可能になる。或いは、例えばデジタルビデオカメラ 1 0 からデジタル映像情報を受信するように構成されたデジタル入力端子 9 1 と、アナログビデオカセットレコーダ 9 5 又はアナログビデオカメラ 2 0 のような装置から各々アナログ形態をとる映像情報 9 3 及び音声情報 9 4 を受信するように構成されたアナログ入力端子 9 2 とを含む映像入力インタフェース 9 0 が提供されても良い。これらの信号 9 3 及び 9 4 は対応するアナログ / デジタル変換器 9 6 及び 9 7 に入力され、変換器の出力は、デジタル入力 9 1 と同様に、分離バッファ 7 8 を介してシステムバス 5 0 に印加される。図 1 B に示すようなクリップ検出をコンピュータモジュール 4 1 内部のソフトウェアにより実行しても良く、図 1 A 及び図 1 B のシーケンス 1 6 及び 3 4 に匹敵する画像とサウンドトラックから構成されるメタデータ補強デジタルビデオシーケンスを編集及びその他の操作や、出力インタフェース 4 7 及び音声映像出力装置 5 6 を介する再生のために、システム 4 0 内に格納しても良い。

20

#### 【 0 0 2 3 】

##### 〔 リズム感のあるシーケンス編集 〕

図 2 に、場面の特定の 1 項目、この場合、海軍博物館への遠足の場面の特定の 1 項目から得られた様々な映像クリップ ( 0 1 番から 1 6 番まで ) のシーケンスを表すヒストグラム 3 0 0 を示す。図 2 から明らかなように、それぞれが約 4 秒 ( クリップ 1 5 ) の最短時間から 2 9 秒 ( クリップ 0 9 ) まで様々な異なる持続時間を有する個別のクリップが合わせて 1 6 個存在している。個々のクリップの間の描写は、各クリップに関連して格納されている上述したメタデータによって与えられる。

30

#### 【 0 0 2 4 】

直線的な ( 時系列 ) 順序で見て行くと、オリジナルの場面のいくつかの部分は退屈で、面白くないなどということに気づく。リズム感のあるシーケンス編集は、見ている側に対するアピールをより高めるように生の映像場面を編集することをいう。プロによる編集作品の入念な見直しを通して、視聴者の関心はいくつかの短い期間の後に失われる、特に、焦点の合った画像が殆ど、或いは全くないところで失われると判定された。これは、特に、家庭で制作された ( アマチュアの手になる ) ビデオ作品に見られるケースであると理解された。この場合、記録されている内容は、通常多くの場合に家族、友人、同僚などからなる視聴者よりも、実際にフィルムを製作した人に対して多くの関連性を持っている。このビデオ作品は、何十分もの長すぎるクリップであると考えられても、視聴者が興味を保ち続けることができる登場人物及び / 又はアクションを含む長編映画フィルム、テレビ映画などのプロの作品から区別されるべきである。

40

#### 【 0 0 2 5 】

そこで、どのような個別クリップであっても、その興味ある内容を最良の機会を捉えて再生できるように、このフィルムに自動的に適用できるいくつかの規則が判定された。本発明者により判定された規則は二倍の効果をもたらす。第一に、発明者はアマチュアが撮影する状況では、レコードボタン 1 2 ( 場合によっては 2 2 ) を押した直後に得られるクリ

50

ップの第1の部分は、通常、ほとんど関心と呼ばないか、又は低画質であると判定した。これは、ユーザがカメラの焦点をクリップの対象物に合わせるまでに、普通、時間を要するためである。この時間は、通常、クリップの約1秒間を占め、このため、好適な実現形態で編集を行う際に使用する第1の規則は、どのクリップについても、最初の1秒間を無視するということである。尚、1秒という期間は相対的なものであり、問題のクリップ又はオリジナルの場面を形成しているクリップの持続時間に応じて変わる。

#### 【0026】

第2の本質的規則は、クリップの残り部分を複数のセグメントに分割することである。各セグメントは、それぞれが特定の時間周期を有する所定の数の間隔の中の1つである。この点に関して、発明者は、クリップをそれぞれが所定の時間周期を持つ複数のセグメントに分割し、所定の時間周期に整合しないクリップの他の部分を編集によりカットすることにより、再生すべき場面の量を有効に圧縮でき、しかも、クリップの本質と場面全体の連続性はそのまま維持されると判定した。好適な実現形態においては、発明者は、家庭で（アマチュアにより）撮影されたビデオ作品の編集には、約4秒から10秒の持続時間を有するクリップセグメントを使用するのが最良であると判定した。ユーザの特定の条件、提供される情報源資料の種類、又は使用される場所、選択した編集テンプレートの種類（後述する）によってそれらの時間周期が異なることは明白であろう。

#### 【0027】

図3に、図2に示す海軍博物館への遠足の場面について、上記の規則に従って行ったクリップ解析を示す。図3から明らかなように、図3でも、生の場面は同様に、16の生のクリップ（09から16）から構成されるものとして示されており、各クリップは何らかの方法により再生のために希望される1つのクリップ又は編集済みクリップ（001から026）に分割されている。編集済みクリップ（001から026）は、それぞれ、各生のクリップ（01から16）の少なくとも1秒後に始まっていることが図3から明白であろう。更に、約7秒の長さであることがわかる第1の生のクリップ（クリップ01）は、4秒の持続時間を有する第1のクリップセグメント（クリップ001）を形成するように編集されている。生の場面のクリップ01のその他の部分は別の編集セグメントに対応するには不十分であり、次の編集済みセグメントは次の生のクリップ02から取り出されている。この特定の例においては、編集は4秒のクリップと10秒のクリップを交互に使用することにより行われており、第2の生のクリップ02で見ると、10秒の編集済みセグメント002はそのクリップから取り出されている。更に、第2の生のクリップ02は20秒の持続時間を有するので、次の4秒クリップ003をこの第2の生のクリップ02から取り出せるというメカニズムが働く。図3から明らかであるように、何れか1つの生のクリップから取り出された編集済みクリップを分離するために、この実施形態では2秒である所定の時間周期が規定されている。

#### 【0028】

このようにして、海軍博物館への遠足の場面の生のクリップ01から16の各々は4秒セグメントと、10秒セグメントの交互配列を必要に応じて使用して編集される。図3から明白であるように、個々の生のクリップから取り出される編集済みセグメントの数はオリジナルの生のクリップの持続時間によって決まる。更に、生のクリップ15から明らかであるように、このクリップは5秒未満の持続時間であるので、上記の規則によれば、この生の場面からはクリップを編集できない結果となる。これは、クリップ15の最初の1秒間を取り除いてしまうと、4秒未満となり、シーケンスで望まれる10秒のセグメントが得られないからである。この例では、リズム感のあるシーケンス設定方法は、所望のセグメント間隔のシーケンス持続時間の70%に満たないと考えられる編集クリップは拒絶するように働く。好適な実施形態では、ある1つのクリップの一部が所望のセグメント持続時間の70%から200%である場合、1つの編集済みクリップを形成する修正後のその部分の再生時間が適切な10秒と4秒のセグメントの再生時間に整合するように、その部分は（時間圧縮又は時間伸張により）修正可能である。例えば、QUICKTIMEの商品名で販売されているソフトウェア製品を使用すると、約25%から400%の範囲で映像時間フ

10

20

30

40

50

レームの圧縮／伸張を行うことができる。

【 0 0 2 9 】

図 3 の例からわかるように、生の場面の総持続時間は 1 6 の生のクリップにわたり 3 2 7 秒であるが、図示されている通り、これを編集すると、編集済みクリップの数は 2 6 となり、再生可能な画像の時間は合わせて 1 7 6 秒である。その結果、編集済みバージョン全体の再生時間はオリジナルの場面と比較してほぼ半減したことになる。更に、編集済み場面はリズム感のある 4 秒と 1 0 秒の変化を示すので、視聴者の関心は薄れない。

【 0 0 3 0 】

以上のことに基づいて、クリップの集合体を提示するためのシステムは、ここでは「テンプレート」という用語で呼ばれる選択された一組の規則を適用するために、クリップの持続時間のプロファイルの作成と、その他の時間関連メタデータを基礎とすることができる。様々に異なる持続時間を有するクリップに対応するために、テンプレートの中に規則の階層を埋め込めれば良い。例えば、数秒又は数フレーム分の持続時間しかないクリップは、何時間又は何十分にもわたる持続時間のクリップとは異なる方法で管理することができる。

10

【 0 0 3 1 】

更に、オリジナルの場面から個々のセグメントをどのように編集するかは、その場面の実際の内容によって異なるであろう。例えば、図 3 は生のクリップから編集済みクリップを選択するために特定のタイミング規則を利用していたが、別の選択を行うことも可能である。一例を挙げると、図 4 に示すように、オリジナルの生の映像に付随するサウンドトラックの解析を利用して、関心領域を識別することができる。そのような関心領域の例として、スポーツの場面における観衆の歓声や、会議の場面での発言者の声などが挙げられる。このようにして、サウンドトラックを解析し、音声レベルが上昇した発言を識別することにより、その箇所では始めるか、又はその箇所にまたがって場面を設定するかなど、クリップの選択を行い、音声のピークの前、又はそれを含む期間、或いはそれに続けて関連性の高い、おそらくは興味を引くであろう内容を得ることができるのである。

20

【 0 0 3 2 】

興味あるクリップセグメントを識別するための音声検出を行うときには、所定の閾値と比較してピーク値を検査するだけで良いのであるが、多くの場合、その閾値を可変とし、総雑音レベルより背景雑音レベルを閾値に反映させる方が有利である。これにより、システムは、閾値ピーク検査及び識別のために、クリップ集合体の提示毎又はクリップ毎のプロファイルを生成することができる。

30

【 0 0 3 3 】

図 6 A に、きわめて可変性の高いと思われる大きな背景雑音が存在するところで音声ピークを判別することを可能にする構成 6 0 を示す。音声信号 6 2 を比較器 6 4 の非反転入力端子と、低域フィルタ 6 6 とに入力する。この低域フィルタ 6 6 の時定数は、メタデータなどの提供をトリガするのには望まれない低レベルの背景雑音を十分に除去しうる周期に設定されている。低域フィルタ 6 6 の出力は比較器 6 4 の反転入力端子に供給され、実際には低域フィルタ 6 6 の時定数に沿って平均された音声信号である信号を形成する。比較器 6 4 は平均音声信号と瞬時音声信号とを比較して、瞬時信号が平均信号を越えた時点を示すトリガ信号 6 8 を発生する。トリガ信号 6 8 は ( 別の ) メタデータとして映像シーケンスと共に取り込まれても良い。

40

【 0 0 3 4 】

図 6 B に、より洗練された音声検出構成 7 0 を示す。音声入力信号 8 4 を全波整流器 8 2 に入力し、全波整流器 8 2 は一对の低域フィルタ 7 4 及び 7 6 に全波整流信号 7 2 を供給する。低域フィルタ 7 4 及び 7 6 はそれぞれ対応する時定数 1 と、時定数 2 を有する。低域フィルタ 7 4 及び 7 6 の出力端子は、同様にトリガ信号 8 0 を供給するように構成されている比較器 7 8 の対応する入力端子に出力を供給する。この構成の場合、時定数 1 及び 2 は、背景雑音と所望の音声とを適切に判別するように設定されていても良い。例えば、2 は背景雑音を判別できるようにするために、背景雑音を公正に表現するよう

50



な相対的に長い周期（例えば、５秒）に設定されていても良い。１は、瞬間的な場面転換音声との判別は可能でありながら、所望の雑音内容（例えば、観衆の歓声又は所望の発言者の声）の検出が十分に行えるように、より短い周期に設定されても良い。この点に関して、１は約０．５秒の時間周期に設定されれば良い。その結果、回路７０はいつ所望の音声ピークが判定されたかに関するマーカー又はメタデータを十分に提供し得る２つの論理レベルの間で変化するトリガ信号８０を発生するように動作する。このメタデータは生の映像場面と組み合わせられて、最終的な編集済みバージョンでおそらく選択されるであろう関心セグメントを識別するためのクリップ解析手続きに使用されても良い。

#### 【００３５】

図６Ａ及び図６Ｂの構成は、例えば、音声－ＡＤＣ９７の入力端子の側で、アナログ回路を使用して実現されても良い。或いは、デジタルフィルタリング及びレベル比較を実行するハードウェア（コンピュータモジュール４１内に構成されるＤＳＰ装置）又はソフトウェア（コンピュータモジュール４１内で稼動する）の何れかによるデジタル構成を使用した実現も可能である。

#### 【００３６】

更に、生の場面の編集を実質的に、又は少なくとも知覚の上で、サウンドトラックと同期させて、サウンドトラックを編集済み場面にダビングすることも可能であろう。これは、適切なビートを識別するためにサウンドトラックを検査し、知覚の上での同期性を実現するために音声又は映像の一方、或いはその双方の再生速度を調整することを含む。例えば、毎分１２０回のビートを有する音楽は毎秒２ビートとなり、これを先に説明した１０秒、４秒のシーケンスのような０．５秒の整数倍である持続時間を有する編集済みクリップを含むリズム感のあるシーケンスの中に均等に分割するのである。

#### 【００３７】

以上説明した自動検出方法及び後述するその他の方法により、１つ又は複数のクリップから構成される生の映像場面を処理し、視聴者の関心をより引き易く、且つ保ち易いリズム感のある画像のシーケンスを形成する編集作品において複数のクリップセグメントを形成する複数の関心部分を識別することが可能になる。特定の実現形態によれば、生の映像信号の解釈時に適用される実際の規則は、所定の映像作成スタイルに基づいて編集済み映像シーケンスの作成を行うと共に、様々に異なる種類の生の映像に適合し得るテンプレート構成によって判定される。１つの編集済みシーケンスを作成するために生の映像を編集すべく使用できる所定の編集規則の組み合わせをそれぞれ組み込んだテンプレートの例は、

・標準１０－４フォーマット

・音楽ビデオ

・音楽トレーラ

・クイックルックサマリ

・ロマンス

・アクション

などである。

#### 【００３８】

それぞれ異なるテンプレートはそれぞれ異なるスタイル、構造の結果をもたらすように構成されており、生の映像場面が選択された特定のテンプレートに適合しているか否かに関わらず、生の映像場面の解釈を行うべく、ユーザにより選択される。

#### 【００３９】

##### [例１－標準テンプレート]

標準テンプレートは、多様な情報源場面の基本的編集を行うために適用できるテンプレートである。このテンプレートの様々な属性は次の通りである。

##### (１) シーケンス：

シーケンスは、最終編集結果を得るために場面をカットするための時間ベースである。即ち、ラインシーケンスは編集済みクリップの実際の持続時間を指定する。先の例では、１０－４秒フォーマットに符合する。１０－４フォーマットの代わりに、１２－４又は１２

10

20

30

40

50

- 6などの他のフォーマットを使用しても良い。

(2) 持続時間：

持続時間は、一般的には、生の場面中のクリップの数の持続時間によって決まる。編集済みシーケンスの総持続時間は、必然的に編集済み映像にダビングすべき付随するサウンドトラックの持続時間にも反映される。しかし、これは7分を越える長いサウンドトラックには不適切である。

(3) 場面転換：

編集済みクリップ間の場面転換は、各クリップ間の4フレームクロスフェードを使用して行われるのが好ましい。

(4) カット規則：

好適な実現形態においては、次のようないくつかのカット規則が適用される。

(a) クリップは古い順にカットされる。

(b) 新たなクリップのカット長さを判定する前に、オリジナルの各クリップの始めと終わりの1秒間を取り除く。

(c) 同じオリジナルの生のクリップから取り出された2つの編集済みクリップの間には、12フレームのクロスフェードを加える。

(d) 可能であれば、10 - 4のリズム感のあるカットシーケンスを適用する。

(e) クリップの持続時間によって、2つ以上のクリップをカットできる場合、常に、残りの持続時間の中で、終わりから1秒間を省略し、且つ2つのクリップの間で4秒間省略できるだけの時間を確保する。

【0040】

[ カット規則の例 - 標準テンプレート ]

第1の生のクリップが7秒未満であれば、4秒にカットする。生のクリップが7秒以上、10秒未満であれば、オリジナルの生のクリップの時間を12秒に伸張した後、伸張クリップをカットして10秒（幾分スローモーションになった）クリップを作成する。次のオリジナルの生のクリップが14秒以上、20秒未満である場合には、初めの1秒を省略し、次の4秒をカットし、次の4秒を省略し、次の4秒をカットし、生のクリップの終わりに至るまで残り部分を省略する。次の生のクリップが20秒以上であれば、初めの1秒を省略し、4秒をカットし、次の4秒を飛び越し、残る10秒をカットし、27秒に至るまで残り部分をカットする。次のクリップが28秒以上であれば、初めの1秒を省略し、4秒をカットし、次の4秒を飛び越し、次に10秒をカットし、次の4秒を省略し、4秒をカットし、38秒に至るまで残り部分を省略する。

(5) 効果：

これは、映像場面に適用し得る視覚効果に関する。標準テンプレートでは効果は適用されない。

(6) 時間伸張：

編集済み映像の最終クリップを200%まで時間伸張し、新たに12秒の持続時間を作る。これを10秒にカットすることにより、クリップの始めと終わりの1秒ずつを省略する。終わりの3秒間は黒色又はテンプレートデフォルトにフェードアウトする。

(7) 音声：

音声はシーケンスに適合するようにビート伸張される（起こりうる最良の整合を得るために増減する）。

(8) マット：

(a) シーケンス持続時間中の第1のクリップへのフェードが起こる最初の10秒間に編集自在のタイトルマットを配置する。シーケンスの編集済みクリップの終結時に編集自在の「終り」マットを設ける。

(b) 編集自在のシーンマットやキャストマットを規定しても良いが、シーケンス中に配置する必要はない。

【0041】

[ 例2 - ロマンスマンタージュ ]

## ( 1 ) シーケンス : 1 2 - 4 秒

この点に関して、ロマンス形の場面は典型的には落ち着いたものであるので、シーケンスの持続時間を 1 0 - 4 シーケンスと比較してわずかに延長して、よりリラックスし、ゆっくりしたペースを与える。

## ( 2 ) 持続時間 :

持続時間は、一般的には、ディレクトリ中のクリップの数の持続時間によって決まる。持続時間シーケンスは必然的にサウンドトラックの持続時間に反映されるが、これは 7 分を越える長さのトラックには勧められない。

## ( 3 ) 場面転換 :

1 2 秒クリップの場合、次のクリップへのフェードインは現在のクリップが終了する前の最後の 2 秒と共に、0 から 1 0 0 % の不透明状態である。各クリップの間に 4 フレームのクロスフェードを使用する。

## ( 4 ) 時間伸張 :

( a ) 持続時間を 1 5 0 % に伸張することによりクリップをスピードダウンさせて、よりリラックスした、ロマンチックなフィーリングを与える。

( b ) 最終クリップの速度を 2 0 0 % まで伸張して、新たに 1 2 秒の持続時間を作り ( スローモーションの効果を発生させ )、そのクリップを 1 0 秒にカットすることによりクリップの始めと終わりの 1 秒を省略し、その 1 0 秒のうち終わりの 3 秒間にわたり黒色テンプレートへのフェードアウトを適用する。

## ( 5 ) カット規則 :

( a ) 古い順にカットする。

( b ) 新たなクリップカット長さを判定するために、はじめから 1 秒間を取り除く。

( c ) 同じショットから取り出された 2 つのクリップの間に 2 秒間のクロスフェードを追加する。

( d ) クリップの持続時間によって、2 つ以上のクリップをカットできる場合、常に、残る持続時間の中で、終わりから 1 秒間を省略し、且つ 2 つのクリップの間から 4 秒間を省略できるだけの間を確保する。

## 【 0 0 4 2 】

## [ カット規則の例 - ロマンズモンタージュ ]

第 1 の生のクリップが 8 秒未満であれば、4 秒にカットする。クリップが 8 秒以上、1 2 秒未満であれば、1 4 秒に時間伸張し、1 2 秒までカットする。次の生のクリップが 1 4 秒以上、2 0 秒未満であれば、初めの 1 秒を省略し、次を 4 秒にカットし、次の 4 秒を省略し、次のクリップを 4 秒にカットし、2 0 秒に至るまで残る時間を省略する。次の生のクリップが 2 0 秒以上である場合には、初めの 1 秒を省略し、4 秒をカットし、次の 4 秒を飛び越し、2 7 秒に至るまで残る時間を省略しつつ残る 1 2 秒をカットする。次の生のクリップが 2 8 秒以上であれば、初めの 1 秒を省略し、4 秒をカットし、次の 4 秒を飛び越し、続く 1 2 秒をカットし、次の 4 秒を省略し、3 8 秒に至るまで残り時間を省略しつつ 4 秒をカットする。

## ( 6 ) 効果 :

アニメーションのフォグフィルタを利用し、ぼんやりとした「ロマンチック」な雰囲気を出す。

## ( 7 ) 音声 :

起こりうる最良の整合を得るために増減するように映像シーケンスに適合させるべく音声をビート伸張 / 圧縮する。

## ( 8 ) マット :

( a ) 第 1 のクリップへのフェードを伴い、1 0 秒の持続時間のシーケンス中に編集自在のタイトルマットを配置する。

( b ) シーケンス中に編集自在の「終り」マットを設ける。

( c ) 編集自在のシーンマット及びキャストマットを設けても良いが、特定のシーケンス中に配置する必要はない。

10

20

30

40

50

## 【 0 0 4 3 】

## [ 例 3 - 音楽ビデオクリップ ]

## ( 1 ) シーケンス :

この例では、映像が音声を補足し、音声が映像を補足するのではない（時には逆に見えることもあるだろうが）ため、シーケンスは音声ビートによって決まる。例えば、毎分 100 ビート未満の音楽の場合、基礎として 10 - 4 シーケンスを使用する。毎分 100 ビート以上の場合には、8 - 3 ベースのシーケンスを使用できる。何れの場合にも、実際のクリップ間隔はほぼビート同期が成立するように調整される。例えば、96 ビート / 分の音楽では 1.6 ビート / 秒となるので、場面を 10 秒と 3.76 秒のシーケンスにカットすることにより、それぞれ 16 秒と 6 秒に近くなり、知覚の上での同期性が得られることになる。

10

## ( 2 ) 場面転換 :

各クリップの間に一般的には 4 フレームのクロスフェードを入れる。

## ( 3 ) 持続時間 :

カットのシーケンスの持続時間は必然的にサウンドトラックの持続時間に反映される。これは 6 分より長いトラックについては勧められない。

## ( 4 ) カット規則 :

( a ) 古い順にカットする。

( b ) 新たなクリップカット長さを判定する前に、各オリジナルクリップの始めと終わりから 1 秒を取り除く。

20

( c ) 同じショットから取り出されたクリップの間に 12 フレームのクロスフェードを追加する。

( d ) クリップの持続時間によって、2 つ以上のクリップをカットできる場合、常に、残りの持続時間の中で、終わりから 1 秒を省略し、且つ 2 つのクリップの間から 4 秒を省略するだけの時間を確保する。

## 【 0 0 4 4 】

## [ カット規則の例 - 音楽ビデオ ( 10 - 4 シーケンスの場合 ) ]

第 1 の生のクリップが 7 秒未満であれば、4 秒にカットし、クリップが 7 秒以上、10 秒未満であれば、12 秒まで時間伸張し、10 秒にカットする。次の生のクリップが 14 秒以上、20 秒未満であれば、初めの 1 秒を省略し、次の 4 秒をカットし、次の 4 秒を省略し、次の 4 秒をカットし、20 秒に至るまで残り時間を省略する。次の生のクリップが 20 秒以上であれば、初めの 1 秒を省略し、4 秒をカットし、次の 4 秒を飛び越し、27 秒に至るまで残り時間を省略しつつ残る 10 秒をカットする。次の生のクリップが 28 秒以上である場合には、初めの 1 秒を省略し、4 秒をカットし、次の 4 秒を飛び越し、次に 10 秒をカットし、次の 4 秒を省略し、4 秒をカットし、38 秒に至るまで残り時間を省略する。

30

## ( 5 ) 効果 : なし

## ( 6 ) 時間伸張 :

10 - 4 シーケンスの場合、最終クリップを 200 % まで時間伸張し、12 秒の新たな持続時間を作り、クリップの始めと終わりの 1 秒ずつを省略し、それを 10 秒にカットする。最後の 3 秒については黒色又はテンプレートデフォルトにフェードアウトする。

40

## ( 7 ) 音声 :

音声の一体性を確保するためには好ましくないが、シーケンスを適合させ、起こりうる最良の整合を得るために、ビートを伸張又は圧縮しても良い。

## ( 8 ) マット :

( a ) 第 1 のクリップについては 10 秒のシーケンス持続時間の中に編集自在のタイトルマットを配置し、第 1 のクリップへフェードインする。

( b ) シーケンス中に編集自在の「終り」マットを設ける。編集自在のシーンマット及びキャストマットを設けても良いが、シーケンス中に配置する必要はない。

## 【 0 0 4 5 】

50

#### 〔例４－クイックルックテンプレート〕

クイックルックテンプレートにより、ユーザは生の映像内容の中で提示し得る場面の全てを短時間のうちに流しながらプレビューすることができる。このクイックルックテンプレートは、例えば３０秒の指定デフォルト時間の中で、又はユーザにより指定された時間でプレビューを提示する。所定の時間（３０秒）の何倍も長い持続時間を有するオリジナルの生の場面に対応するために、クリップを長さがフレームのみの短い持続時間の集合にカットすることにより、リズム感のある編集シーケンスを適用する。クリップ速度は、例えば１００から６００％の範囲で変更自在であり、カット持続時間は数フレームから数秒の範囲である。希望に応じて多様なクイックルックテンプレートを形成できる。

#### 【００４６】

##### <クイックルックの例１：>

クリップを先に述べた１０－４規則に対応する方式で１０フレームと、４フレームの複数のセグメントに分割する。このような短い持続時間の中により多くの場面を盛りこむために、場面をオリジナルの再生速度の３００％まで伸張することもあり、場合によっては、オリジナル場面のフレーム速度を遅くする。例えば、図３に示す構成を使用する場合、１０－４秒規則を使用して１７６秒にわたる標準編集済みクリップを取り出していたが、それらと同じクリップを処理して、クリップ毎に１０フレームセグメントと、４フレームセグメントとを取り出して、再生のために１７６フレームを得る。例えば、PAL再生システムで使用されているような毎秒２５フレームのフレーム速度では、これは約７秒の再生時間に相当する。クイックルック方式によれば、４フレームセグメント及び１０フレームセグメントから選択されたセグメント、或いはそれらセグメントの全てを伸張して３０秒のプレビュー時間を実現する。ユーザはより長い生の場面の「プレビュー」を選択でき、フレーム速度や時間伸張などのシステムパラメータを調整することができる。

#### 【００４７】

##### <クイックルックの例２：>

この例では、１０－４規則のようなリズム感のあるカット規則を適用しない。ユーザは、通常情報源場面の全長に関するクイックルックの持続時間を指定する。例えば、５分の生の場面を３０秒のクイックルック再生に圧縮することを希望しても良い。そこで、ユーザはカット長さをその持続時間全体を均等に分割したものとなるように調整することができる。３０秒の出力の場合、これは互いに継ぎ合わされた３０個の１秒セグメントから形成されていても良い。各セグメントは、各クリップを、それぞれが１秒の（余り又はカット）部分により分離された複数の３秒部分に分割することにより得られる。各３秒部分を３００％時間圧縮して所望の再生持続時間を得る。その後、それらの部分のうち３０個を使用してクイックルックプレビューを形成する。生のクリップの持続時間が様々に異なっている場合には、テンプレートにおいて、各生のクリップから１つの部分を取り出すように保証することが望ましい。

#### 【００４８】

##### <クイックルック－比較例>

この例は、以下の表１に示すように、多くのクイックルックシーケンス規則をアクションシーケンスと比較したものである。

#### 【００４９】

##### 【表１】

10

20

30

40

規則	アクション	クイックルック
クリップ開始から「IN」ポイント設定	1秒	2秒
クリップ終了から「OUT」ポイント設定	1秒	2秒
持続時間が許せば、同じクリップ内で次の「IN」ポイントを設定する前にある期間パス	2秒	1秒
クリップ速度	100%	200%
持続時間がカット持続時間よりx%短ければ、クリップを飛び越す	70%	70%
1つのリズムサイクルを形成するカット数	2	サイクルなし
リズムサイクル中のカットの持続時間	10秒	—
リズムサイクル中のカットの持続時間	4秒	—

10

## 【0050】

上記の例は、あるスタイルの編集済みビデオ作品を実現し得るいくつかの異なるテンプレート配列を表しているに過ぎない。自動編集に関わる上記の規則を修正して、別のテンプレート構成を設定できることは当業者には理解されるであろう。その一例は、操作すべき特定のジャンルに基づいて様々に異なる範囲の圧縮／伸張を使用できることである。その他の型のテンプレートの例として、「軍隊戦術」、「SF」、「ホラー」、「戦争」及び「ウェスタン」などの様々な既に確立されている映画のスタイルを反映したものが挙げられる。「スポーツ」や「アクション」などの他のスタイルを使用しても良い。希望に応じて生の場面に複数のテンプレートを適用しても良い。例えば、生の場面をロマンステンプレートに従って編集し、編集済みバージョンを今度はアクションテンプレートに従って編集しても良い。更に、複数のテンプレートを順次又は組み合わせて使用すれば、様々な規則が何らかの特定の効果を無効にすることなく、規則のその適用の優先順位を守るように、規則の階層を適用できるであろう。

20

## 【0051】

何れの場合も、特定のスタイルは、望まれる編集済み作品再生のスタイルに一般的には適用可能であるとわかっているいくつかの標準編集済みクリップの長さ（例えば、10 - 4, 12 - 4）の使用を基礎としている。上記のテンプレートは2つの基本クリップ長さしか与えていないが、特定の状況に従って、更に多くのクリップ長さを使用して良い。例えば、野生生物や風景の場面は、少なくとも持続時間の短いクリップと交互に、より長いクリップの編集に適しているであろう。更に、先に説明した様々なテンプレートにおいては1 - 4及び12 - 4フォーマットが好ましいが、その他の持続時間を使用しても良い。代表的には、「より短い」持続時間は1秒から8秒の周期であり、より長い持続時間は12秒から20秒である。また、各クリップの始めから1秒間のカットも変更できる。通常、0.5秒から1秒の任意の周期を使用して良い。更に、編集済みセグメントの選択の間の2秒の間隔を変更しても良い。1秒から5秒の持続時間が適切であろう。更に、図3の実施形態は4秒の編集済みクリップ長さ、10秒の編集済みクリップ長さを交互に示していたが、例えば、短い、短い - 短い - 長いといった別のパターンになるように様々なクリップ長さの選択を代えても良いし、或いは使用すべき特定の持続時間を擬似ランダム方式で選択しても良い。これは、基本持続時間が3つ以上ある場合に有用である。

30

40

## 【0052】

システムは、選択されたテンプレート構造に従って、そのリズム感のあるシーケンスを使用して適切な場面転換、ビート同期音楽で生の場面をカットし、デジタル効果を追加する。デジタル効果の例としては、オリジナルのカラーパレットの変更、フォグフィルタ効果、画像を歪ませることなどがある。このように、ここで説明したリズム感のあるシーケンス編集システムはフィルムエディタ、サウンドエディタ及び特殊効果エディタの技術をアマチュアが撮影した情報源映像に適用することにより、アマチュアが結果を修正、調整又は単純に鑑賞するために、映像の再配列を自由に指示できるようになる。これらの効果

50

を生の映像に適用するプロセスは高速であり、コンピュータシステム 40 内部におけるオフライン（即ち、リアルタイムでない）処理に適している。このプロセスは、アマチュアのフィルム製作者に、クリップ中に場面転換や in - out ポイントを配置するという繰り返しの作業を通して時間を浪費するのではなく、編集中の映像の内容に関してより高いレベルの判断を自由に行わせることができる。また、この構成はリアルタイム動作も可能にする。例えば、ある生の映像源について、単純なキーストローク又はマウス 43 のクリックにより所望のテンプレートを選択すると、その結果システム 40 により映像源の自動編集が実行され、編集済みシーケンスはビデオ表示装置 56 にレンダリングされてユーザがそれを直ちにすることができる。更に、複数のウィンドウを操作すれば、複数のテンプレートに従って情報源映像を同時にリアルタイムで編集、再生することができる。例えば、これにより、10 - 4 テンプレートと 12 - 4 テンプレートとをリアルタイムで比較して、ユーザがよりアピールすると思う方を出力させることが可能になる。

10

#### 【0053】

動作に関しては、上記のリズム感のあるシーケンス編集システムは、生の映像場面に関連するメタデータを検査して、上記のテンプレートから情報の組み合わせを表す編集判断リスト（EDL）を生成することにより編集結果を提供する。基本的なリズム感のあるシーケンス処理は、実際の映像内容をアクセスする必要も如何なる知識もなく、クリップ番号、持続時間、フレーム番号などを含むメタデータのみに基づき行われるので、編集判断リストの評価は、ビデオ製作者が in - out ポイントを適切に設定する（通常何時間もの）時間を費やすことなく迅速に実行できる。一度編集判断リストが作成されたならば、そのリストは生の場面に適用され、再生のために適切なビットシーケンスが選択される。これはリアルタイムに実行されても良いし、或いは編集済み出力を連続する再生ファイルにコピーすることにより実行されても良い。

20

#### 【0054】

図 9 は、生の映像場面をある特定の実現形態に従って編集する方法を示すデータの流れ図である。ステップ 100 では、ビデオ画像とそれに付随するサウンドトラックから構成される生のデジタル場面が、その場面と関連し、少なくとも場面の中の様々なクリップを識別するメタデータと共に入力される。入力された場面は抽出プロセス 102 に提供され、このプロセスは生のデジタル場面 106 からメタデータ 104 を取り出す。画像と付随する音声を含む生のデジタル場面 106 は適切な方法により、通常磁気ハードディスク記憶装置 108 に格納される。特定の構成が許すならば、入力される生の場면을線 154 により示すように記憶装置 108 から取り出しても良い。

30

#### 【0055】

生のデジタル場面から更なるメタデータを取り出すことが望まれる場合には、その場면을記憶装置 108 から取り出し、メタデータ判定プロセス 110 に提供する。このプロセスは、追加のメタデータ 112 を取り出すように生のデジタル場面 106 を処理すべく動作し、メタデータ 112 は加算装置 114 においてオリジナルのメタデータ 104 と組み合わせられる。メタデータ抽出プロセス 110 は、図 4、図 6 A 又は図 6 B に示すような音声抽出構成を含んでいても良い。その代わりに、又はそれに加えて、別の抽出プロセスを使用しても良い。これは、例えば生の場面の個々のフレームを比較して、画像の各部分の動きや、それらの部分の間の衝突を識別することなどである。これは、例えばキャプション付き図形及びノイズ又はサウンド効果を得るときに有用である。その他のメタデータ抽出プロセスには、面検出、線検出、動き検出を始めとして多くのものがある。図 9 に示すように、プロセス 110 は、ユーザに実行すべき特定のメタデータ抽出プロセスを選択させる入力 111 を含む。メタデータ 112 をプロセス 110 を使用して抽出した場合、それを生の映像と共に既存のメタデータに加えて記憶装置 108 にセーブしておいても良い。デフォルトでは、追加メタデータ抽出プロセスは全く実行されない。加算器 114 は組み合わせメタデータ 116 を適用モジュール 118 へ出力する。

40

#### 【0056】

最良の同期を実行するために、オーバーダブ音声源 136 をリズム感のあるシーケンス編

50

集で利用できる音声源 1 3 6 のビートを識別するビート抽出プロセス 1 3 8 により解析する。抽出されたビート 1 3 9 は適用モジュール 1 1 8 に入力される。

【 0 0 5 7 】

また、適用モジュール 1 1 8 には、ユーザによりテンプレートの格納庫 1 2 4 からマルチプレクサ 1 2 2 を介して選択された特定の編集テンプレート 1 2 0 が入力される。図 9 に示すように、ユーザに直接に提供される所定のテンプレートに加え、ユーザが独自に作成するか、1 つ又は 2 つ以上の既存のテンプレートの変形として作成するユーザ自身のテンプレートを定義できるようになっている。適用モジュール 1 1 8 は選択されたテンプレートをメタデータ 1 1 6 及び抽出されたビート 1 3 9 に適用して、最終編集済みバージョンにおいて再生するために生のデジタル場面から選択されるべきクリップの実際のセグメント及びそれらの対応する周期を表す編集表示リスト ( E D L ) 1 2 6 を形成する。編集表示リスト 1 2 6 は、ユーザに編集済みバージョンと関連するタイトルセグメントを編集させる入力 1 2 8 を更に含む。

【 0 0 5 8 】

組み合わせメタデータ 1 1 6 はリストとして表現され、編集表示リスト 1 2 6 と共に保持されても良く、また最終編集済みシーケンスにおいて重要な編集済みクリップをマーキングするために使用されても良い。

【 0 0 5 9 】

編集表示リスト 1 2 6 は別の適用モジュール 1 3 0 に入力され、このモジュールは編集表示リストを解釈して、記憶装置 1 0 8 に格納されている生のデジタル場面をカットし、適切な編集済みセグメントを取り出す。また、適用モジュール 1 3 0 はアニメーション及びキャプションを含めた図形を、適切なサウンド効果と共に記憶装置 1 3 2 から取り出し、これらは編集済み映像と組み合わせて編集済み映像出力 1 3 4 を形成する。編集表示リスト 1 2 6 は、適用モジュール 1 3 0 により形成されるリズム感のあるシーケンス編集に整合するようにオーバーダブ音声源 1 3 6 の再生速度を変更するように構成されたビート調整ユニット 1 5 8 へビート制御指令 1 5 6 を適宜出力することができる。この点に関して、場合によっては音声再生速度を特定の編集間隔 (例えば、1 0 - 4 ) にほぼ整合させるか、或いは音声源 1 3 6 のビートにほぼ整合するように編集間隔 (例えば、1 2 - 4 から 1 1 . 5 - 3 . 5 ) を調整しても良い。

【 0 0 6 0 】

その後、編集済み映像 1 3 4 は、音声源 1 3 6 から直接取り出されるか、又は必要に応じてビート調整ユニット 1 5 8 から取り出されるオーバーダブサウンドトラックと加算ユニット 1 4 0 で組み合わせられる。加算ユニット 1 4 0 は音声・視覚場面 1 4 2 を出力し、それは記憶装置 1 4 4 に格納されるか、又は再生装置 1 4 6 へ直接出力され、ビデオ表示装置 1 5 0 及び音声スピーカ 1 5 2 を含む再生システム 1 4 8 において再生される。

【 0 0 6 1 】

[ 挿入タイトルの生成 ]

編集済みビデオ作品へタイトルを挿入することができ、図 9 に示す構成では、編集表示リスト 1 2 6 への入力 1 2 8 を介して実現される。これにより、映像を構成する一連のクリップの中に文脈上正しく、図形により活字組みされた挿入タイトルを自動的に挿入することができる。これは、ユーザの時間と労力をほとんど必要とせず実現可能である。挿入タイトル発生器が設けられており、これは、著者、植字工及び編集者の役割を果たし、語句データベース、規則ベース適用メタデータ解析及びユーザ側で入力されたメタデータに対応するインタフェースを使用する。そのような構成の一例を図 1 1 に示す。図 1 1 には、挿入タイトル発生器 ( I T G ) 2 0 0 が示されている。I T G 2 0 0 は、ユーザ側で入力されるメタデータ 2 1 4 に対応する入力モジュール 2 1 6 を含む。ユーザが入力するメタデータ 2 1 4 は、オリジナルの生の場面 1 0 0 を供給源とするメタデータ 1 0 4 から直接取り出されても良い。そのようなメタデータの例としては、ユーザが実際に生の場面を記録する時点で実際の画像捕捉時間と共に挿入したメタデータがある。ユーザ入力メタデータ 2 1 4 は、更にユーザが選択した生成プロセス 1 1 0 から編集システムへ取り出され



る追加メタデータ 1 1 2 も含む。更に、ユーザが編集段階で直接メタデータを入力できるように、全く別個のユーザ入力メタデータ入力部 2 0 2 を設けても良い。ユーザ入力メタデータ 2 1 4 はメタデータ解析ユニット 2 0 4 に供給され、解析ユニットはメタデータ 2 1 4 の特性を解析して、語句データベース 2 0 8、活字組みデータベース 2 1 0 及び図形データベース 2 1 2 の組み合わせから挿入タイトル出力 1 2 8 を形成するように機能する規則ベースアプリケーション 2 0 6 を実現する。

#### 【 0 0 6 2 】

従来の挿入タイトル作成方法に関して言えば、「著者」の役割は、クリップの持続時間や記録日時に関する情報ばかりでなく、場面の基礎を成すストーリー、重要なシーン及び映画のジャンルに関する情報も提供するユーザ入力メタデータ 2 1 4 により容易になっている。メタデータ解析ユニット 2 0 4 がメタデータ 2 1 4 の内容を解析して、時間コード及びクリップ持続時間に関する情報を獲得する。時間コードメタデータを使用して、複数の相対的時間に撮影されたクリップをクラスタリングすることができる。

#### 【 0 0 6 3 】

例えば、生の映像が早朝、ほぼ午前 7 時、その後の午前 9 時、ほぼ正午にそれぞれ撮影されたいくつかのショットから構成されているとする。メタデータ解析ユニット 2 0 4 はこのメタデータを使用して、どこで挿入タイトルを入れれば良いかを表す 3 つの明確に異なるシーンを判定する。ユーザは、ITG 2 0 0 でプリセットされたデフォルトを変更したいときに、クリップ間の時間クラスタリングの閾値を調整することができる。

#### 【 0 0 6 4 】

ITG 2 0 0 は、解析されたメタデータを語句データベース 2 0 8 から取り出された生活様式に適合するキャッチフレーズ、ことわざ及び面白い俗語と相互参照することにより、挿入タイトルの内容又は文面を供給する。相互参照の結果は、文脈上はユーザ入力メタデータ 2 1 4 が元になるウィットに富んだ挿入タイトルを形成する。この文面は、活字組みデータベース 2 1 0 から取り出されたプロ並みのデザインのテンプレートに挿入される。このテンプレートは植字工の役割を果たし、テキストを操作する際に他の図形ツールを必要としない。希望に応じて、挿入タイトルのマット背景幕を形成するために、適宜、図形データベース 2 1 2 を取り出しても良い。時間コードメタデータと、挿入タイトルの配置及び持続時間を定義する一連の規則は挿入タイトルのシーケンスを決定し、それにより、作品中により高いレベルの連続性と面白さを与える上で、編集者の役割を完璧に果たす。挿入タイトルの持続時間は、オリジナルの場面に適用されるリズム感のあるシーケンス編集及び表示される、即ち、視聴者によって読み取られるべきテキストの長さを含むいくつかの要素に従って判定される。それらの要素はテンプレートに組み込まれても良い。

#### 【 0 0 6 5 】

例えば、ユーザがスカウトのキャンプ遠足のビデオを撮影していた場合、ITG 2 0 0 を使用して、そのイベントを全体として記述すると共に、システムにより識別される各々の特定のシーンを記述する一連のキーワードを入力する。1 つのイベントとしての場面は、「キャンピング (camping)」、「スカウティング (scouting)」、「チームワーク (teamwork)」、「サバイバル (survival)」などと記述される。テントを張っている間に楽しんでいるスカウトたちを映した場面からなる第 1 のシーンは、ユーザにより「友達 (friends)」、「楽しい (fun)」、「少年 (boys)」、及び「助け (help)」と記述されている。そこで、システムはそれらのキーワードを使用して、イベントのコンテキストを判定し、次に 1 つのキャッチフレーズ又はキャッチフレーズのリストをユーザに提示し、そこから選択を行えば良い。この例におけるキャッチフレーズの例は、

「Boys will be boys」(男の子のいたずらはしょうがない)

「A friend in need is a friend indeed」(まさかの時の友こそ真の友)

「Survival of the fittest」(適者生存)

「Time flies when you're having fun」(光陰矢のごとし)

などとなるであろう。

#### 【 0 0 6 6 】

ユーザは入力 2 1 8 を介して選択を行い、I T G 2 0 0 はその選択された語句を受け入れ、記述されたシーンの始めに配置する。これは、選択されたタイトルを適切なシーンを識別するメタデータと関連付けることにより実行される。このように、ユーザはシーン毎にキーワードを提供することにより、又は単純に映画を全体として記述し、システムに時間コード又は他の種類のメタデータを相互参照させることにより、タイトルマットを生成することができる。特定の日時に行われる共通イベントを参照するタイトルマットを生成することができる。例えば、ほぼ正午に撮影されたクリップは「昼食時間 ( lunch time ) 」などの語句を示唆させるものになるであろう。クリップが午前 7 時に始まるのであれば、「夜明け ( a new day dawns ) 」などの語句が代わりに示唆されるであろう。グローバルポジショニングシステム ( G P S ) 構成から獲得されるメタデータを適宜使用して、イベントが行われる場所の位置関係を示す内容を有する挿入タイトルを示唆することも可能である。このような相互参照は、処理すべき特定のシーンに適合する正しいタイトルを集中的に探す上で有用であろう。

#### 【 0 0 6 7 】

最も基本的なシーンの識別は自動的に、生の場面と関連する記録時間メタデータを使用して実行される。これは図 8 に示されている。図 8 に、4 分 4 8 秒の範囲内のクリップクラスタリングに基づいてタイトルを配置する場合の I T G システム 2 0 0 の結果を表す。先行するクリップの終了時から 4 分 4 8 秒を越えてリアルタイムでずれているクリップは、何れも、システムにより新たなシーンとして考えられる。

#### 【 0 0 6 8 】

図 8 に示すように、クリップ番号 1 から 4 ( それぞれのクリップの持続時間は図 8 から不確定であり、ここでの説明には無関係である ) はそれぞれほぼ同じ時間 ( ほぼ午前 9 時 3 0 分 ) に撮影されており、各クリップの間には 4 分 4 8 秒を越える間隔は存在していない。その結果、クリップ 0 1 から 0 4 をオリジナル場面の第 1 のクリップのクラスタ ( クラスタ 0 1 ) として一纏めにする。図からわかるように、クリップ 0 5 はほぼ午前 1 0 時に撮影されており、クリップ 0 1 から 0 4 及びクリップ 0 6 以降の後続クリップから時間の上で離れている。従って、クリップ 0 5 は独自のクラスタ、この場合はクラスタ 0 2 を形成する。クリップ 0 6 はほぼ午前 1 0 時 4 0 分に撮影され、その後、クリップ 0 7 から 1 1 が続いているが、それぞれの撮影間隔は 4 分 4 8 秒を越えない。従って、クリップ 0 6 から 1 1 は第 3 のクリップのクラスタを形成する。このようにして、オリジナル場面の各クリップをクラスタに分割し、各クラスタの始まりがオリジナル場面における様々なシーンを記述するために挿入タイトルを配置できる場所を形成する。

#### 【 0 0 6 9 】

以上説明した時間コードの規則ベース解析は、類似した時間に撮影されたクリップのクラスタと、異なる日付で撮影されたクリップのクラスタとを定義する。この点に関して、クリップ 2 7 はほぼ午後 5 時 3 0 分に撮影され、クリップ 2 8 は明らかに次の日のほぼ午前 6 時 4 5 分に撮影されていることがわかる。この結果に従って、挿入タイトル発生器 2 0 0 を使用してシーンマットが自動的に挿入されるか、又は図形ユーザインタフェースを介して個々のシーンのユーザに指示がなされる。

#### 【 0 0 7 0 】

挿入タイトルテンプレート、並びに視聴者が読み取らなければならない音節、語及び文字の数の選択は、稼動するタイトルマットの持続時間を判定するように機能する。I T G 2 0 0 はこれらの要因を組み合わせるシステムを利用して、メッセージの理解の程度、編集リズム並びに最大限の読みやすさを確保するためにどの字体を選択するのが正しいかといった問題の妥協点を探る。更に、適切であれば、シーケンスのジャンルによって挿入タイトルの持続時間を判定しても良い。例えば、そのジャンルが緊張又は恐怖の一方 ( 例えば、ホラー映画 ) である場合、持続時間を短縮して、挿入タイトルを読んでいるときに視聴者にストレスをかけるようにしても良い。

#### 【 0 0 7 1 】

タイトルマットの持続時間は、先に説明したテンプレート構造の編集リズムに整合してい

10

20

30

40

50

るのが好ましい。10 - 4 規則から成るテンプレート構造は、4 秒の持続時間に整合するタイトルマットを有するのが好ましい。同様に、8 - 3 規則から成るテンプレート構造は3 秒の持続時間に整合するタイトルマットを有していることになるであろう。ITG200 は選択されたテンプレートの持続時間規則を解析して、挿入タイトルに適する持続時間を判定するように機能する。

#### 【0072】

状況によっては、テンプレートがタイトルマットの持続時間に都合よく対応する構造を有している場合もある。例えば、3 秒又は4 秒の持続時間で十分に視聴者はメッセージを理解できるであろう。このように、ITG200 は選択されたテンプレートを識別する、マルチプレクサ122 からの入力120 を含むことができる。このように、語句データベース208 は、規則ベースアプリケーション206 に理解に必要な利用可能な持続時間を超えるメッセージである場合に別の持続時間を指定させることができるように、タイトル持続時間に関する埋め込み規則を含んでいても良い。このようなことは、スタイルの整った有力テンプレートを使用する場合に起こり得る。そのようなテンプレートはニュース放送や音楽ビデオなどの特定の時間フォーマットの作品から構成されている。

10

#### 【0073】

短い期間で流れる作品又は25 フレーム未満のクリップ持続時間を有する高速カットを含む作品では、タイトルマットにおけるメッセージについて、視聴者がその内容を理解するために利用できる時間に適合するようにレイアウト、語数 / 音節数及び活字組みを整えなければならない。例えば、タイトルマットの持続時間がわずかに14 フレームである場合、ITG200 は語句データベース208 から、1 語又は数語の短いメッセージのみを選択し、活字組みデータベース210 が許すならば、メッセージを読みやすい太字の活字面で組む。

20

#### 【0074】

図9の配列リズム感のあるシーケンス編集システムを参照してITG200を説明したが、その用途と動作はそのような特定の適用に限定されるものではない。メタデータの入力、語句や活字組みの選択に関する特定の規則を伴うシーンの変更などの識別を、オリジナル場面のリズム感のある編集に依存せずに実現しても良い。その結果、オリジナル場面の編集を必要とせずに、挿入タイトルをそのオリジナル場面に直接追加することができるであろう。

30

#### 【0075】

挿入タイトル発生器200に関連して以上説明したシーンの識別は、タイトルマットの場所ホルダを識別するためばかりでなく、正しい場面転換の使用などの他の要素にも使用されて良い。特に、リズム感のあるシーケンス編集に関して言えば、1つのシーンの始まりに使用される「フェーディング」場面転換及びそのシーンの終了時の「フェードアウト」は、作品全体を通して他のカット技法とはスタイルの上で異なっている。これは、各々がビデオ作品のストーリーの中で異なる意味を表しているからである。

#### 【0076】

シーン識別システムをタイトル及びリズム感のあるシーケンス編集のテンプレート構造と組み合わせると、サウンドトラック及びサウンド効果を自動的に配置することができる。システムはこの情報を処理し、その結果を使用して、どこでタイトル音楽を入れるべきか及びどこでサウンドトラックを始めるべきかを知る。サウンドトラックのビートは取り出された後、特定の映像テンプレート規則に応じて、編集シーケンスに整合させるために、又は編集シーケンスをビートに整合するように編集シーケンスを修正するために、又はその双方を修正するために修正されても良く、その結果、音楽と映像カットとが同期する。

40

#### 【0077】

このようにして、リズム感のあるシーケンス編集システムはオリジナルの生の場面を検討するために余分の時間を費やす必要なくビデオを見ることを望んでいるユーザに対して、編集済みの面白味に富んだ結果を生成することができるのである。生の場面は多くの場合に所期の時間より長いショットを含んでおり、場合によっては、全く記録するはずではな

50

かったショットをも含んでおり、それらを好都合で、ユーザにほぼ見えるような方式で編集できる。

【 0 0 7 8 】

[ プリントフレーム選択 ]

生の映像に適用し得るもう 1 つの編集プロセスは、プリント又はオリジナル場面を示すスライド表示を行うために望まれる個別のフレームの選択を実行する。図 5 を参照すると、海軍博物館への遠足の生の場面が再び、それぞれ持続時間の異なる 16 個の個別クリップの連続として示されている。更に、図 12 も併せて参照すると、クリップは場面全体を表す個別のサムネイルを形成するように分割される。

【 0 0 7 9 】

図 12 に示すように、プリントフレーム選択を実行する方法 250 はステップ 252 で、場面の個々のクリップを区別し、且つそれらのクリップの持続時間を識別するために必要なメタデータなどのメタデータを含む生のデジタル場面の入力をもって始まる。ステップ 254 では、ユーザは特定の場面について必要なプリントフレームの数を識別することを要求される。選択された数のプリントフレームが場面全体の内容を表示するように、実際の数はオリジナルの生の場面の全体の長さ及び / 又はその場面に含まれている個々のクリップの数に基づいて示唆されても良い。プリントフレームの数を選択したならば、生の場面、特に個々のクリップを分割して、複数のセグメントを形成する。それらのセグメントから個々のプリントフレームを選択することになる。図 5 に示すように、生の場面は 64 のプリントフレームを形成するように分割され、各々のクリップから 4 つのプリントフレームが選択されている。このように、各クリップはそのクリップから取り出す必要のあるプリントフレームの数により決まる等しい持続時間を有する 4 つのセグメントに分割されるのである。ステップ 258 では、各セグメントを処理して、プリントに最良のフレームを取り出す。ステップ 260 では、セグメントごとの最良フレームをサムネイル画像にフォーマットし、ステップ 262 では、選択されてフォーマットされた最良フレームをプリントに適する方式でフォーマットする。

【 0 0 8 0 】

最良フレームを取り出すために個々のセグメントを処理するステップ 258 は、最良フレームを判定するようにセグメント中の個々のフレームにいくつかの規則を適用するように機能する。生の場面の状況に応じて及び / 又はユーザが希望する特定のテンプレートに従って、何れか 1 つ又は複数の規則が適用される。発明者の研究によれば、何れか 1 つのクリップの最良又は最適の画質の画像はそのクリップの第 3 のセグメントにあることがわかつていいる。更に、何れか 1 つのセグメントの中では、発明者はそのセグメントの中央に最良フレームがあると判定している。セグメントの持続時間に応じて、これは他の規則に従って、更に検査されるべき個々のフレームの数を制限することがあり得る。

【 0 0 8 1 】

他の選択規則は、セグメント内に現れるフレームに適用できる様々なプロセスに基づいている。1 つの規則は、図 4 に示すような音声解析を含むことができ、この場合、生の場面と関連するサウンドトラック中のピークとほぼ一致するとして 1 つのフレームを選択する。そのようなピークは、スポーツイベントにおける観衆の歓声又は会議で発言している発言者の声を含むと考えられるであろう。

【 0 0 8 2 】

複数の一致が起こった場合、その問題を 1 つのプリントフレームに解決するためにユーザインタフェースを適用しても良い。例えば、整合するフレームの各々をビデオ表示装置 56 にレンダリングすることにより、プリントのためにユーザが 1 つのフレームを選び出しても良い。

【 0 0 8 3 】

別の規則は面解析のような画像解析を含むことができる。これによれば、背景である風景の情報などに対して人間の顔を示す特徴を検出することに基づいて 1 つのプリントフレームを選択する。面検出ソフトウェア及びそのプロセスは当業者には知られており、図 12

10

20

30

40

50

のステップ258を考慮しながら限られた数のフレームに容易に適用できる。その他の選択基準には、大きな相対運動を示すフレームはピント外れなどの画質の低い記録につながりやすいとして選択から排除する隣接フレームの間の動き検出がある。

【0084】

図5及び図12の構成から出力されるフォーマッティング済みプリント出力の一例を図7に示す。これは、グレートバリアリーフでのダイビングの場面を主題としている。24の個別のプリントフレームサムネイルが1つのページにフォーマッティングされており、それぞれのフレームがダイビングの異なる段階を示している。例えば、第1のフレームはダイバーをグレートバリアリーフへ運ぶボートの航跡を示し、次の3つのフレームはダイビングの準備をしているダイバーたちを示している。それ以降のフレームは、ダイビング中のクリップから取り出された様々なサンプルを示す。

10

【0085】

オリジナル場面を見ることを希望している人に負担をかけることなく、オリジナル場面の主題を要約したような生の映像場面の有用なバージョンを複数提供するために、生の映像場面の編集を可能にするいくつかの構成が提供されたことは以上の説明から明白であろう。リズム感のあるシーケンス編集方式は、編集後のバージョンの視聴者の関心を高め且つそれを維持するように考案されたスタイルの整った再生が行われるように、オリジナル場面のサイズを好都合に縮小するものである。また、リズム感のあるシーケンス編集はオリジナル場面とオーバーダブ音声との組み合わせを支援し、場面と音声をほぼ同期するように合わせることで、視聴者の関心は保たれる。リズム感のあるシーケンス生成を伴う場合も、伴わない場合もあるが、挿入タイトルを配置することにより、視聴者にあらすじなどを伝えるために再生すべき場面中の様々なシーンを識別することができる。プリントフレーム選択構成によれば、オリジナル映像場面のあらすじを静止画像によって捉え、都合よく再生することができる。それぞれがその生成に際して異なるプロセスの組み合わせを利用しているので、それらを別個に適用した場合、リズム感のあるシーケンス編集とプリントフレーム選択が各出力で同一の映像フレームを再生させるという保証はない。

20

【0086】

尚、本発明は複数の機器（例えば、ホストコンピュータ、インタフェイス機器、リーダー、プリンタなど）から構成されるシステムに適用しても、一つの機器からなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

30

【0087】

また、本発明の目的は前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システム或いは装置に供給し、そのシステム或いは装置のコンピュータ（CPU若しくはMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

【0088】

この場合、記憶媒体から読出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【0089】

プログラムコードを供給するための記憶媒体としては、例えばフロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

40

【0090】

また、コンピュータが読出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOS（オペレーティングシステム）などが実際の処理の一部又は全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0091】

50

更に、記憶媒体から読出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部又は全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【図面の簡単な説明】

【図1A】デジタル供給源からのデジタルビデオクリップの供給を示す図。

【図1B】アナログ供給源からのデジタルビデオクリップの供給を示す図。

【図2】一体となってオリジナルの生の場面を形成する複数のクリップの提示ヒストグラム。

10

【図3】本実施形態にて定義される「10-4」規則に従った図2のクリップの解析を表す図。

【図4】音声解析に基づくクリップの分割を示す図。

【図5】フレームプリントに使用するための図2の生の場面の分割を示す図。

【図6A】音声解析を実現するための様々な構成を示す図。

【図6B】音声解析を実現するための様々な構成を示す図。

【図7】図5のクリップ分割からサンプリングした映像フレームを示す図。

【図8】別のクリップ配列例に基づくタイトルの挿入を示す図。

【図9】好適な編集方法のデータ流れ図。

【図10】本実施形態における汎用コンピュータの構成を示す概略ブロック図。

20

【図11】挿入タイトル発生器の概略ブロック図。

【図12】プリントフレーム選択方法を示す流れ図。

【符号の説明】

- 10 デジタルビデオカメラ
- 20 アナログビデオカメラ
- 30 クリップ検出器
- 40 コンピュータシステム
- 41 コンピュータモジュール
- 42 キーボード
- 43 マウス
- 45 プロセッサユニット
- 46 メモリユニット
- 47 ビデオインタフェース
- 48 I/Oインタフェース
- 49 記憶装置
- 50 接続バス
- 52 モデム装置
- 53 ハードディスクドライブ
- 54 フロッピーディスクドライブ
- 55 CD-ROM
- 56 ビデオ表示装置
- 57 プリンタ
- 59 ネットワーク
- 102 抽出プロセス
- 108 記憶装置
- 110 メタデータ判定プロセス
- 114 加算装置
- 118 適用モジュール
- 122 マルチプレクサ
- 124 テンプレート格納庫

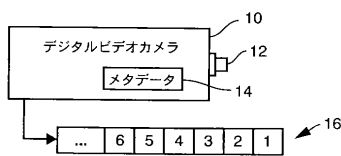
30

40

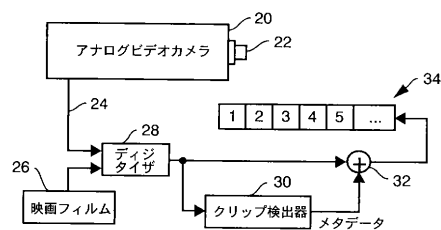
50

- |       |           |
|-------|-----------|
| 1 2 6 | 編集表示リスト   |
| 1 3 0 | 適用モジュール   |
| 1 3 2 | 記憶装置      |
| 1 3 6 | オーバーダブ音声源 |
| 1 3 8 | ビート抽出プロセス |
| 1 4 4 | 記憶装置      |
| 1 4 6 | 再生プロセス    |
| 1 5 0 | ビデオ表示装置   |
| 1 5 2 | スピーカ      |
| 1 5 8 | ビート調整ユニット |
| 2 0 0 | 挿入タイトル発生器 |

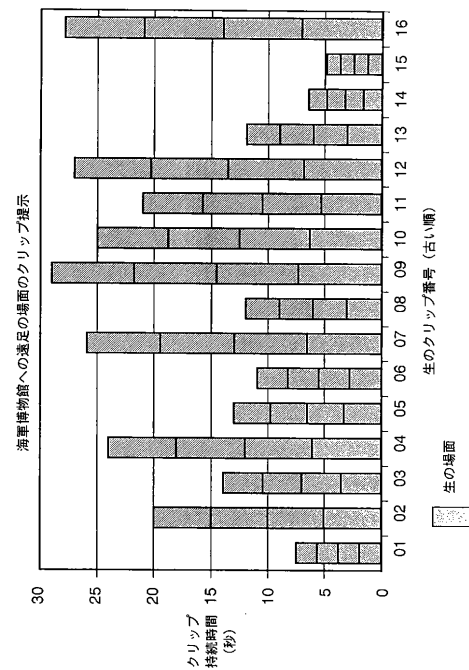
【 図 1 A 】



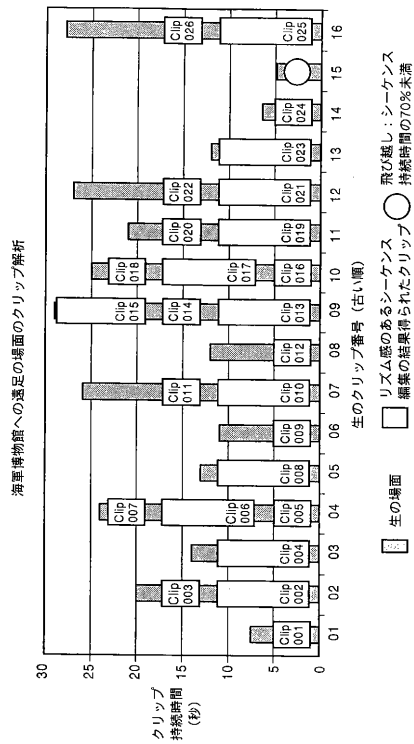
【 圖 1 B 】



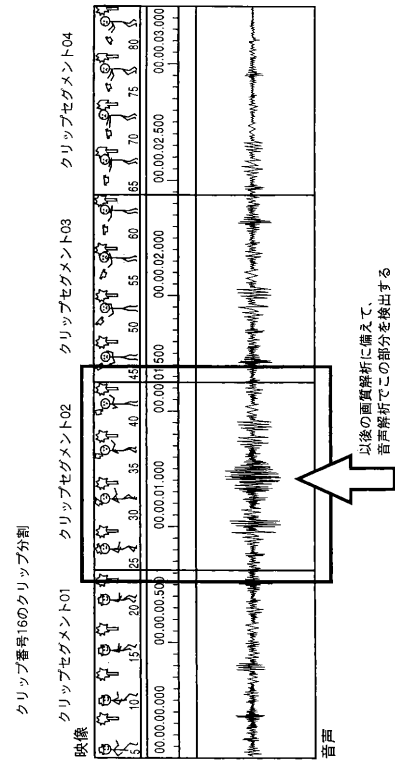
【圖 2】



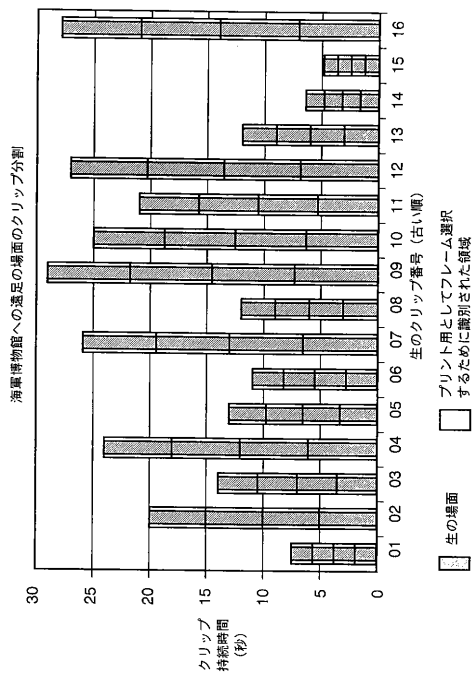
【図 3】



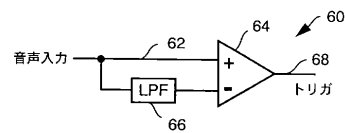
【図 4】



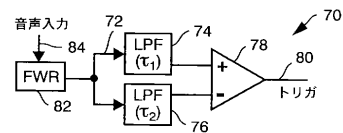
【図 5】



【図 6 A】

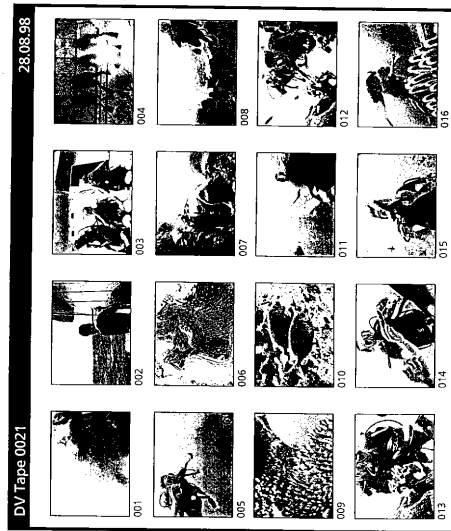


【図 6 B】

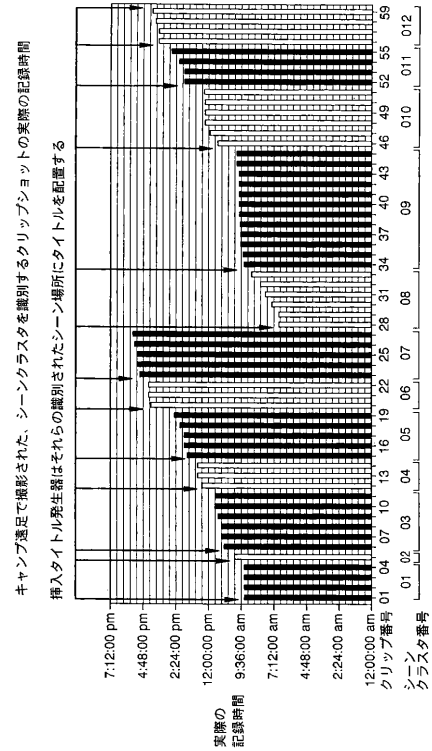




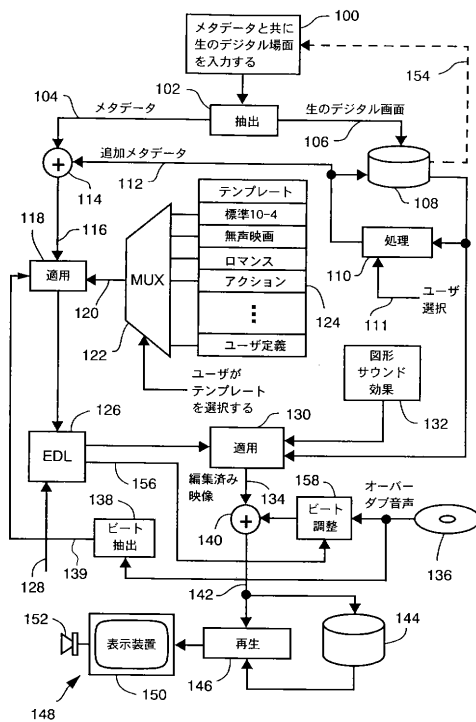
【図 7】



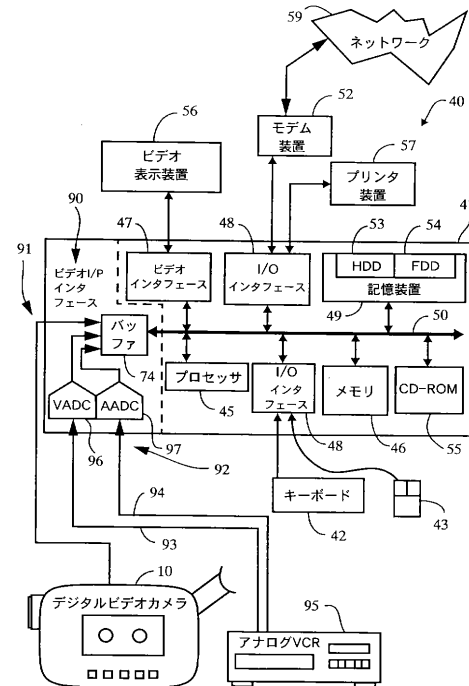
【図 8】



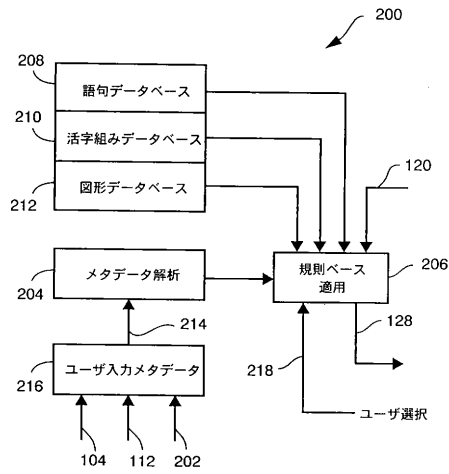
【図 9】



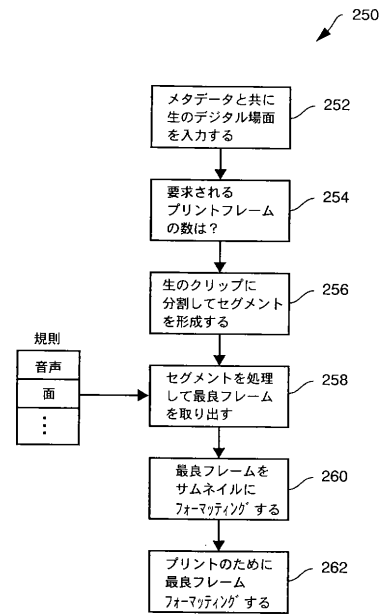
【図 10】



【図 11】



【図 12】



---

フロントページの続き

(72)発明者 ジュリー ラエ コウアルド  
オーストラリア国 2 1 1 3 ニュー サウス ウェールズ州, ノース ライド, トーマス ホル  
ト ドライブ 1 キヤノン インフォメーション システムズ リサーチ オーストラリア プ  
ロプライエタリー リミテッド 内

審査官 小林 大介

(56)参考文献 特開平 0 9 - 1 4 7 5 2 9 ( J P , A )  
特開平 0 7 - 1 5 3 2 3 9 ( J P , A )  
特開平 0 5 - 2 9 0 5 4 9 ( J P , A )

(58)調査した分野(Int.Cl. , D B 名)  
H04N 5/91 - 5/93  
H04N 5/262  
G11B 27/00 -27/038