

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4059631号
(P4059631)

(45) 発行日 平成20年3月12日 (2008. 3. 12)

(24) 登録日 平成19年12月28日 (2007. 12. 28)

(51) Int. Cl.

F I

HO 4 N	7/173	(2006. 01)	HO 4 N	7/173	6 4 O Z
HO 4 H	20/28	(2008. 01)	HO 4 H	1/00	2 3 6
HO 4 H	20/74	(2008. 01)	HO 4 H	1/00	2 8 6
HO 4 H	60/02	(2008. 01)	HO 4 H	1/00	6 O 2
HO 4 H	60/37	(2008. 01)	HO 4 H	1/00	6 3 4

請求項の数 25 (全 23 頁) 最終頁に続く

(21) 出願番号 特願2000-596737 (P2000-596737)
 (86) (22) 出願日 平成12年1月28日 (2000. 1. 28)
 (65) 公表番号 特表2002-541684 (P2002-541684A)
 (43) 公表日 平成14年12月3日 (2002. 12. 3)
 (86) 国際出願番号 PCT/IB2000/000135
 (87) 国際公開番号 W02000/045599
 (87) 国際公開日 平成12年8月3日 (2000. 8. 3)
 審査請求日 平成14年6月3日 (2002. 6. 3)
 (31) 優先権主張番号 9902235. 2
 (32) 優先日 平成11年2月1日 (1999. 2. 1)
 (33) 優先権主張国 英国 (GB)

前置審査

(73) 特許権者 501306483
 エミューズ メディア リミテッド
 アイルランド国、ダブリン 4、サンディ
 マウント、パーク アベニュー 80
 (74) 代理人 110000198
 特許業務法人湘洋内外特許事務所
 (72) 発明者 レインスフォード、パトリック、ジェイ.
 アイルランド国、ダブリン 4、サンディ
 マウント、パーク アベニュー 80

審査官 川崎 優

最終頁に続く

(54) 【発明の名称】 対話型システム

(57) 【特許請求の範囲】

【請求項 1】

オブジェクトを表わすデータをデジタルビデオ番組に対応付ける装置であって、
 少なくとも一部がオブジェクトを取り入れている複数の個別のフレームを有するデジ
 タルビデオ番組を提供する手段と、

編集リストを作成するために前記デジタルビデオ番組内の個別のショットを識別するこ
 とで前記デジタルビデオ番組をパースする手段と、

関連コンテンツを含む一連のショットを形成するために、関連コンテンツを含むショッ
 トを前記編集リストから識別する手段と、

前記一連のショットそれぞれの内から、前記オブジェクトを含む少なくとも1つのキー
 フレームを選択する手段と、

データが対応付けられるキーフレーム内で前記オブジェクトを選択する手段と、

前記キーフレームから前記オブジェクトの属性を抽出する手段と、

対話型データを、前記キーフレーム内の前記オブジェクトに対応付ける手段と、

前記一連のショットを通して前記オブジェクトを追跡するために前記オブジェクトの属
 性を利用する追跡手段とを含み、

前記対話型データは、前記一連のショット内の前記オブジェクトに対応付けられ、デー
 タ・シーケンス内の前記オブジェクトを表すデータとともに、前記対話型コンテンツデー
 タは埋め込まれ、

前記オブジェクトを追跡するために前記オブジェクトの属性を利用する前記追跡手段が

10

20

、異なるタイプの視覚環境において前記オブジェクトを自動的に追跡するために、或るシーケンスの視覚的複雑さに応じて、オブジェクトを追跡するための異なるアルゴリズムにより追跡を行うこと

を特徴とする装置。

【請求項 2】

関連コンテンツを含む一連のショットを形成するために、関連コンテンツを含むショットを識別する手段が、ショット群の階層を生成する手段を含むこと

を特徴とする請求項 1 記載の装置。

【請求項 3】

前記パースする手段が、ショットの変化を認識するのに用いられる基準を入力する手段を含むこと

10

を特徴とする請求項 1 または 2 に記載の装置。

【請求項 4】

前記オブジェクトの属性を抽出する手段が、フレーム上に形成された境界線内にオブジェクトを隔離する手段と、前記オブジェクトのエッジを識別し突き止めるために、前記境界線内でエッジ検出を行う手段と、および、前記オブジェクトの形状モデルを蓄積する記憶手段とを含むこと

を特徴とする請求項 1 乃至 3 のいずれかに記載の装置。

【請求項 5】

前記オブジェクトの属性を抽出する前記手段が、前記オブジェクトの形状、サイズ、位置、色、テクスチャ、輝度勾配の属性の少なくとも 1 つ、および前記属性に基づく時系列統計量を記録する手段も含むこと

20

を特徴とする請求項 1 乃至 4 のいずれかに記載の装置。

【請求項 6】

前記オブジェクトの属性を抽出する前記手段が、前記オブジェクトが以前に蓄積されたオブジェクトと区別できるかどうか判定するために、双方のオブジェクトの属性を比較する手段と、

前記オブジェクトが区別できないと判定されるときに、当該オブジェクトに関する追加の情報を与えるようにユーザに指示する手段と、

ユーザからの当該オブジェクトを定義し直す指示を受け付けた場合、受け付けた指示に従って定義し直す手段と、を提供することを特徴とする請求項 1 乃至 5 のいずれかに記載の装置。

30

【請求項 7】

前記オブジェクトの属性を抽出する前記手段が、前記オブジェクトがそのフレームにすでに蓄積されたオブジェクトと区別できるかどうか判定するために、双方のオブジェクトのフレーム内の位置を比較する手段と、

前記オブジェクトの位置が、別のオブジェクトの位置とは区別できない場合、どのオブジェクトが、そのような位置に対応づけられるか判定するために、前記オブジェクトにリンクを割当てようユーザに指示を行う手段と、

を提供することを特徴とする請求項 1 乃至 6 のいずれかに記載の装置。

40

【請求項 8】

前記オブジェクトを追跡するために前記オブジェクトの属性を利用する手段が、前記オブジェクトの属性がフレームからフレームへと変わると、前記オブジェクトのこれらの蓄積された属性を更新する手段を含むこと

を特徴とする請求項 1 乃至 7 のいずれかに記載の装置。

【請求項 9】

前記オブジェクトを追跡するために前記オブジェクトの属性を利用する前記追跡手段が、追跡されるすべてのフレームを下位の表現に変換する手段と、

各フレーム内に各オブジェクトを突き止めるために、距離測度を最小限に抑えることで

50

フレーム内の各オブジェクトの位置を決定する手段と、

各オブジェクトの隠蔽された部分、出口および入口を、平滑化する後処理を行う手段と

、
追跡されたシーケンス内の前記オブジェクトの位置に関し、ユーザから訂正の入力を受け付け、当該訂正に従って前記オブジェクトの位置の属性を訂正する手段と、を含むことを特徴とする請求項 1 乃至 8 のいずれかに記載の装置。

【請求項 10】

対話型データを、前記キーフレーム内の前記オブジェクトに対応付ける手段が、

URL、HTML ページ、ビデオクリップ、オーディオクリップ、テキストファイル、マルチメディア・カタログの 1 つまたは複数を含め、異なるタイプのデータのデータベースを提供する手段と、

前記オブジェクトに対応付けるために、前記データベースから前記対話型コンテンツデータを選択する手段と、を含むこと

を特徴とする請求項 1 乃至 9 のいずれかに記載の装置。

【請求項 11】

対話型データを、前記キーフレーム内の前記オブジェクトに対応付ける手段は、前記埋め込まれた対話型コンテンツデータが或るフレーム内のオブジェクトの位置、形状、ランク、ポイントに対応付けられるフレーム同期データであるか、または、或るグループ内のすべてのオブジェクトに対応付けられる群同期データであるか、あるいは、ジャストインタイム式にストリーム再生されるデータであるかどうか判定する手段を用いて前記データ・シーケンスを生成し、

前記フレーム同期データを該当するフレームに対応付ける手段と、

前記群同期データを、グループが変わるときのフレームに対応付ける手段と、

該当するオブジェクトに対応付けられる必要がある前に、ユーザにデータをジャストインタイム式にストリーム再生する手段と、を備えること

を特徴とする請求項 1 乃至 10 のいずれかに記載の装置。

【請求項 12】

異なる対話型コンテンツデータを、それぞれ異なるオブジェクトに対応付ける手段を備えること

を特徴とする請求項 1 乃至 11 のいずれかに記載の装置。

【請求項 13】

或るオブジェクトを表わす対話型データを、デジタルビデオ番組に対応付ける方法であって、

a) 受信したデジタルビデオ番組内の個別のショットを予め定めた基準に従って識別し編集リストを作成する工程と、

b) 前記作成した編集リストから、関連コンテンツを含むショットを識別し、関連コンテンツを含む一連のショットを形成する工程と、

c) 前記一連のショットそれぞれの内から、前記オブジェクトを含む少なくとも 1 のキーフレームを選択する工程と、

d) キーフレーム内のオブジェクトの選択を受け付ける工程と、

e) 前記選択を受け付けたオブジェクトの属性を抽出して記録する工程と、

f) 前記選択されたオブジェクトに対応付ける対話型データを受け付ける工程と、

g) 前記選択されたオブジェクトに前記受け付けた対話型データを埋め込む工程と、

h) 前記オブジェクトの属性を利用して、前記一連のショットを通して前記オブジェクトを追跡する工程と、

i) 前記追跡により突き止められたオブジェクトが含まれるフレームに前記オブジェクトに対応付ける対話型データを埋め込む工程と、を備え、

工程 h) が、異なるタイプの視覚環境において前記オブジェクトを自動的に追跡するために、或るシーケンスの視覚的複雑さに応じて、オブジェクトを追跡するための異なるアルゴリズムにより追跡を行う工程を含むこと

を特徴とする方法。

【請求項 1 4】

工程 a) の前記基準は、ショットの変化を認識するのに用いられる基準であること
を特徴とする請求項 1 3 記載の方法。

【請求項 1 5】

工程 b) が、ショットのシーケンス群の階層を生成する工程を含むこと
を特徴とする請求項 1 3 または 1 4 に記載の方法。

【請求項 1 6】

工程 d) が、
前記選択を受け付けたオブジェクトを隔離するために、前記キーフレーム上に境界線を
形成する工程と、
前記境界線で囲まれたオブジェクトを区分するためにエッジ検出を行う工程と、を含み

、
工程 e) が、前記オブジェクトの形状モデルを蓄積する工程を含むこと
を特徴とする請求項 1 3 乃至 1 5 のいずれかに記載の方法。

【請求項 1 7】

工程 e) が、前記オブジェクトの形状、サイズ、位置、色、テクスチャ、輝度勾配の属
性の少なくとも 1 つ、および前記属性に基づく時系列統計量を記録する工程を含むこと
を特徴とする請求項 1 3 乃至 1 6 のいずれかに記載の方法。

【請求項 1 8】

工程 e) が、
前記オブジェクトが以前に蓄積されたオブジェクトと区別できるかどうか判定するた
めに、双方のオブジェクトの属性を比較する工程と、
前記オブジェクトが区別できないと判定されるときに、前記オブジェクトに関するつ
いの情報を与えるようユーザに指示する工程と、
ユーザから前記オブジェクトを定義し直す指示を受け付けた場合、受け付けた指示に従
って定義し直す工程と、を含むこと
と特徴とする請求項 1 3 乃至 1 7 のいずれかに記載の方法。

【請求項 1 9】

工程 e) が、
前記オブジェクトがすでにそのフレーム内に蓄積されたオブジェクトと区別できるかど
うか判定するために、双方のオブジェクトのフレーム内の位置を比較する工程と、
前記オブジェクトの位置が、別のオブジェクトの位置とは区別できない場合には、どの
オブジェクトが、そのような位置に対応付けられるか判定するために、前記オブジェクト
にランクを割当てようユーザに指示を行う工程とを含むこと
を特徴とする請求項 1 3 乃至 1 8 のいずれかに記載の方法。

【請求項 2 0】

工程 h) が、前記オブジェクトの属性がフレームからフレームへと変わると、前記オブ
ジェクトのこれらの蓄積された属性を更新する工程を含むことを特徴とする請求項 1 3 乃
至 1 9 のいずれかに記載の方法。

【請求項 2 1】

工程 h) が、
追跡されるすべてのフレームを下位の表現に変換する工程と、
各フレーム内に各オブジェクトを突き止めるために、距離測度を最小限に抑えることで
フレーム内の各オブジェクトの位置を決定する工程と、
各オブジェクトの隠蔽された部分、出口および入口を、平滑化する後処理を行う工程と

、
追跡されたシーケンス内の前記オブジェクトの位置に関し、ユーザから訂正の入力を受け
付け、当該訂正に従って前記オブジェクトの位置の属性を訂正する工程と、を含むこと
を特徴とする請求項 1 3 乃至 2 0 のいずれかに記載の方法。

【請求項 2 2】

工程 f) の対応型データは、URL、HTML ページ、ビデオクリップ、オーディオクリップ、テキストファイル、マルチメディア・カタログの 1 つまたは複数を含む、異なるタイプのデータのデータベースとして提供されるデータの中から選択されること

を特徴とする請求項 1 3 乃至 2 1 のいずれかに記載の方法。

【請求項 2 3】

工程 i) が、

前記埋め込まれた対話型データが或るフレーム内のオブジェクトの位置、形状、ランク、ポイントに対応付けられるフレーム同期データであるか、または、或るグループ内のすべてのオブジェクトに対応付けられる群同期データであるか、あるいは、ジャストインタイム式にストリーム再生されるデータであるかどうか判定する工程と、

フレーム同期データを該当するフレームに対応付ける工程と、

群同期データを、グループが変わるときのフレームに対応付ける工程と、

該当するオブジェクトに対応付けられる必要がある前に、ユーザにデータをジャストインタイム式にストリーム再生する工程と、をそれぞれ含むこと

を特徴とする請求項 1 3 乃至 2 2 のいずれかに記載の方法。

【請求項 2 4】

工程 c) ~ 工程 i) において、異なる対話型データが、それぞれ異なるオブジェクトに対応付けられること

を特徴とする請求項 1 3 乃至 2 3 のいずれかに記載の方法。

【請求項 2 5】

或るオブジェクトを表わす対話型データを、デジタルビデオ番組に対応付けるためにコンピュータに、

a) 受信したデジタルビデオ番組内の個別のショットを予め定めた基準に従って識別し編集リストを作成する手順と、

b) 前記作成した編集リストから、関連コンテンツを含むショットを識別し、関連コンテンツを含む一連のショットを形成する手順と、

c) 前記一連のショットそれぞれの内から、前記オブジェクトを含む少なくとも 1 のキーフレームを選択する手順と、

d) キーフレーム内のオブジェクトの選択を受け付ける手順と、

e) 前記選択を受け付けたオブジェクトの属性を抽出して記録する手順と、

f) 前記選択されたオブジェクトに対応付ける対話型データを受け付ける手順と、

g) 前記選択されたオブジェクトに前記受け付けた対話型データを埋め込む手順と、

h) 前記オブジェクトの属性を利用して、前記一連のショットを通して前記オブジェクトを追跡する手順と、

i) 前記追跡により突き止められたオブジェクトが含まれるフレームに前記オブジェクトに対応付ける対話型データを埋め込む手順と、を実行させ、

手順 h) が、異なるタイプの視覚環境において前記オブジェクトを自動的に追跡するために、或るシーケンスの視覚的複雑さに応じて、オブジェクトを追跡するための異なるアルゴリズムにより追跡を行う手順を含めて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取可能な記録媒体。

【発明の詳細な説明】**【0001】**

本発明は、対話型システム、特にデジタルビデオ信号内にデータを多重化するシステムに関するものである。

【0002】

ブロードキャストされるか、あるいは、デジタルビデオディスク (DVD) またはビデオテープ上に与えられる場合があるデジタル信号形式のビデオ番組を提供することが知られているが、本発明は、放送番組用のビデオ信号を与える形式に限定されることはない。

【0003】

テレビ放送チャンネルの数が増すにつれて、広告収入が減少する。すなわち、商業上の理由で、広告主は、そのマーケティング努力を、限られた数の放送チャンネルに限定しているからである。さらに、望ましくない広告を受信できないようにするための、テレビ視聴者が利用できる装置（例えば、Vチップ）の稼働率も増してはいるが、ただし、現在のところ、選択的に広告を阻止する方法はなく、その結果、テレビ視聴者に興味のあるような広告までも、阻止されてしまう。

【0004】

インターネットの利用が高まるにつれて、ユーザは、パーソナルコンピュータ（PC）、あるいは、例えばテレビリモコンまたはマウスといったものに使用されるデジタル・セットトップ・ボックスを用いて、広範囲にわたる様々なデータソースや情報源にアクセスすることに慣れてきている。

10

【0005】

本発明は、インターネット上で行われることがある情報転送および／またはトランザクションを容易にするために、テレビ視聴者が、ブロードキャストされ得るビデオ信号と対話できるようにするシステムを提供しようとしている。

【0006】

本発明の一面により、ビデオ番組信号を提供する手段、少なくとも1つのオブジェクトに対応付けられる対話形コンテンツデータであって、前記オブジェクトが現れる前記ビデオ番組信号のフレームに対応付けられる前記データを生成する手段、前記ビデオ番組信号で前記データを多重化する手段、ビデオ番組信号を見る手段、前記データを検索する手段、および前記オブジェクトの詳細を得るために前記データを使用する手段を含む対話型システムが提供される。

20

【0007】

好ましくは、前記使用する手段は、前記オブジェクトの前記詳細を得るために対話型Webサイトにアクセスする手段を含む。好都合なことに、前記使用する手段は、前記オブジェクトの詳細のリストを生成する手段と、前記リストから選択する手段をさらに含む。

【0008】

好都合なことに、前記対話型Webサイトにアクセスする手段は、前記オブジェクト用の購入トランザクション、または広告カタログのブラウジングを含むことがある前記オブジェクトの詳細を入手するようにしている。

30

【0009】

好ましくは、前記生成する手段は、前記オブジェクトが現れる前記ビデオ番組信号のフレームごとに前記オブジェクトを追跡する手段と、それぞれの各フレーム内の前記オブジェクトの位置を識別する手段を含む。

【0010】

好ましくは、前記ビデオ番組の各フレームは、前記対話型コンテンツデータを含む。

【0011】

好都合なことに、前記追跡手段は、シーンブレイク（中断）を決定する手段と、前記オブジェクトが現れる次のフレーム内で前記オブジェクトを探す手段を含む。

【0012】

40

好都合なことに、前記多重化手段は、トランスポート・ストリーム（例えば、MPEG-2 / DV Bトランスポート・ストリーム）を生成するために、前記データを、前記番組信号のオーディオデータおよびビデオデータと同期化する手段を含む。

【0013】

好都合なことに、前記システムは、前記トランスポート・ストリームを、例えば衛星ネットワーク、地上ネットワーク、ケーブルネットワークの少なくとも1つを通じてブロードキャスト（放送）する手段を含む。

【0014】

好都合なことに、前記検索する手段は、マウス、キーボード、リモコン装置の1つを含む。

50

【 0 0 1 5 】

本発明の第2の面により、少なくとも一部がオブジェクトを取り入れている複数の個別のフレームを有するデジタルビデオ番組を提供する手段、前記オブジェクトがキーフレームを与えられるビデオ番組の1フレームを選択する手段、データが対応付けられるキーフレーム内で前記オブジェクトを選択する手段、キーフレームからオブジェクトの属性を抽出する手段、対話型データを、キーフレーム内の前記オブジェクトに対応付ける手段、ビデオ番組の後続フレームを通して前記オブジェクトを追跡するために前記オブジェクトの属性を利用する手段を含む、前記オブジェクトを表わすデータをデジタルビデオ番組に対応付ける装置であって、前記オブジェクトが追跡されているビデオ番組の後続フレーム内で、前記オブジェクトに前記対話型データに対応付けさせ、また或るデータ・シーケンス内で、前記オブジェクトを表わすデータとともに前記対話型コンテンツデータを埋め込ませる装置が提供される。

10

【 0 0 1 6 】

好都合なことに、前記データ・シーケンスを、標準データ・シーケンス（例えば、MPEG-2/DVB適合のデータ・シーケンス）に変換する手段が提供される。

【 0 0 1 7 】

ビデオ番組が、アナログ形式である場合には、好ましくは、そのビデオ番組を、デジタル化形式に変換する手段が提供される。

【 0 0 1 8 】

好ましくは、ビデオ番組の1フレームを選択する手段は、そのデジタル化ビデオ番組を、複数の関連ショットのシーケンスに分けるために編集リストを作成する手段と、各シーケンス内から少なくとも1つのキーフレームを選択する手段を含む。

20

【 0 0 1 9 】

好都合なことに、編集リストを作成する手段は、編集リストを作成するためにビデオ番組内の個別のショットを識別することでビデオ番組をパースする手段、関連コンテンツを含む一連のショットを形成するために、関連コンテンツを含むショットを識別する手段、および、ショット群の階層を生成する手段をさらに含む。

【 0 0 2 0 】

好都合なことに、前記パースする手段は、ショットの変化を認識するのに用いられる基準を入力する手段を含む。

30

【 0 0 2 1 】

好ましくは、前記オブジェクトの属性を抽出する手段は、フレーム上に形成された境界線内にオブジェクトを隔離する手段、前記オブジェクトのエッジを識別し突き止めるために、この境界線内でエッジ検出を行う手段、および、前記オブジェクトの形状モデルを蓄積する記憶手段を含む。

【 0 0 2 2 】

好都合なことに、前記オブジェクトの属性を抽出する前記手段はまた、前記オブジェクトの形状、サイズ、位置、色、テクスチャ、輝度勾配の属性の少なくとも1つ、および前記属性に基づく時系列統計量を記録する手段も含む。

【 0 0 2 3 】

好都合なことに、前記オブジェクトの属性を抽出する前記手段は、前記オブジェクトが、以前に蓄積されたオブジェクトと区別できるかどうか判定するために、双方のオブジェクトの属性を比較する手段を含み、また前記オブジェクトが区別できないと判定されるときには、例えば前記境界線を定義し直すことで、前記オブジェクトを定義し直す手段を提供する。

40

【 0 0 2 4 】

好ましくは、前記属性を抽出する前記手段は、前記オブジェクトが、すでに蓄積されたオブジェクトと区別できるかどうか判定するために、双方のオブジェクトのフレーム内の位置を比較する手段を含み、また前記オブジェクトの位置が、別のオブジェクトの位置とは区別できない場合には、どのオブジェクトが、そのような位置に対応付けられるか判定す

50

るために、オブジェクトにランクを割当てて手段を提供する。

【0025】

好ましくは、前記オブジェクトを追跡する手段は、オブジェクトが、異なるフレーム内に位置を移すと、オブジェクトのこれらの蓄積された属性を更新する手段を含む。

【0026】

好都合なことに、前記追跡する手段は、異なるタイプの視覚環境においてオブジェクトを自動的に追跡するために、或るシーケンスの視覚的複雑さに応じて使用する複数のアルゴリズム手段を含む。

【0027】

好都合なことに、前記追跡手段は、追跡されるすべてのフレームを下位の表現に変換する手段、各フレーム内に各オブジェクトを突き止めるために、距離測度を最小限に抑えることでフレーム内の各オブジェクトの位置を決定する手段、閉塞の障害を除く前記オブジェクトの位置と、オブジェクトが前記フレームに出入りする入口や出口を処理する手段、および、追跡されたシーケンス内のオブジェクトを再検討し、不適切な位置に置かれたオブジェクトがあれば、その位置の属性を訂正する手段を含む。

10

【0028】

好ましくは、前記対応付ける手段は、URL、HTMLページ、ビデオクリップ、オーディオクリップ、テキストファイル、マルチメディア・カタログの1つまたは複数を含め、異なるタイプのデータのデータベースを提供する手段と、前記オブジェクトに対応付けるために、前記データベースから前記対話型コンテンツデータを選択する手段を含む。

20

【0029】

好ましくは、前記対応付ける手段は、埋め込まれた対話型コンテンツデータが、或るフレーム内のオブジェクトの位置、形状、ランク、ポイントに対応付けられるフレーム同期データであるか、または、或るグループ内のすべてのオブジェクトに対応付けられる群同期データであるか、あるいは、ジャストインタイム式にストリーム再生されるデータであるかどうか判定する手段を用いて前記データ・シーケンスを生成し、そこでは、フレーム同期データを該当するフレームに対応付ける手段が提供され、また群同期データを、グループが変わるときのフレームに対応付ける手段が提供され、さらに、該当するオブジェクトに対応付けられる必要がある前に、ユーザにデータをジャストインタイム式にストリーム再生する手段が提供される。

30

【0030】

対話型コンテンツデータを或るオブジェクトに対応付けることに関して、上記のものを定義してきたが、異なる対話型コンテンツデータは、それぞれ異なるオブジェクトに対応付けられるはずである。

【0031】

本発明の第3の面により、デジタル化ビデオ信号内の或るオブジェクトに対応付けられる対話型コンテンツデータのデータ・シーケンスを受け取る手段、さらなるトランスポート・ストリームを生成するために、そのデータ・シーケンスを、そのデジタル化ビデオ信号のビデオおよびオーディオと同期化する手段、および、パケット識別子を、そのさらなるトランスポート・ストリームに対応付ける手段を含み、総称デジタル・トランスポート・ストリーム（例えば、DVB/MPEG-2またはATSC/MPEG-2）中にデータ・シーケンスを埋め込む装置が提供される。

40

【0032】

好ましい実施例では、さらなるトランスポート・ストリームをテレビ視聴者にブロードキャスト（放送）する手段が提供される。

【0033】

好ましくは、データ・シーケンスを受け取る手段は、デジタルビデオ信号ストリーム、デジタルオーディオ信号ストリーム、デジタルデータ・シーケンス・ストリーム、デジタル制御データストリームから成るエレメンタリーストリームを受け取る手段、これらのデータストリームのそれぞれをパケット化して固定サイズ・ブロックにし、プロトコルヘッダ

50

を付加して、パケット化エレメンタリーストリームを生成する手段、および、これらのデータストリームの相互の関係を確立するために、そのパケット化エレメンタリーストリームをタイムスタンプと同期化する手段を含む。

【 0 0 3 4 】

好ましくは、このデータ・シーケンスを同期化する手段は、パケット化エレメンタリーストリームを多重化して、同期バイトのヘッダが付けられたトランスポート・パケットにする手段と、異なるパケット識別子を各パケット化エレメンタリーストリームに割当てて手段を含む。

【 0 0 3 5 】

好都合なことに、前記パケット化エレメンタリーストリームをタイムスタンプと同期化する手段は、現在時間を示すために基準タイムスタンプでスタンプする手段と、データ・シーケンス・ストリームを、いつビデオストリームおよびオーディオストリームと同期化する必要があるか示すために、デコーディング・タイムスタンプでスタンプする手段を含む。

10

【 0 0 3 6 】

好都合なことに、さらなるトランスポート・ストリームをユーザにブロードキャスト（放送）する手段は、放送で使えるように、すべてのチャンネルをリストした番組対応付けテーブルを提供する手段、放送チャンネル内のすべてのエレメンタリーストリームを識別する番組マップ・テーブルを提供する手段、および、さらなるトランスポート・ストリーム中の個別のパケットとして、この番組対応付けテーブルと番組マップ・テーブルを送る手段を含む。

20

【 0 0 3 7 】

本発明の第4の面により、総称デジタル・トランスポート・ストリームに埋め込まれたデータが、総称デジタル・トランスポート・ストリームで表わされるオブジェクトに対応付けられるデータのデータ・シーケンスを含む、前記埋込みデータを検索する装置が提供され、また、前記装置には、ビデオ信号内のパケット識別子を認識する手段、総称デジタル・トランスポート・ストリームから、そのデータ・シーケンスを抽出する手段、対応付けられるデータを検索するビデオ・シーケンス内のオブジェクトを識別する手段、前記データシーケンスを、前記識別されたオブジェクトと同期させる手段、および、前記対応付けられるデータを対話形式で利用する手段が含まれる。

30

【 0 0 3 8 】

好ましくは、前記オブジェクトを識別する手段は、或るフレーム内の或るオブジェクトを選択する手段、前記オブジェクトに対応付けられるデータを表示する手段、表示されたデータのリストからデータを選択する手段、および、前記オブジェクトに関するデータに対応付けられる埋込みデータを抽出する手段を含む。

【 0 0 3 9 】

好都合なことに、埋め込まれた対応付けデータを有するオブジェクトを表示するために1フレームを選択する手段、前記オブジェクトに対応付けられるデータのリストを表示するために、これらの表示されたオブジェクトの1つを選択する手段、および、前記リストから選択する手段が提供される。

40

【 0 0 4 0 】

好都合なことに、前記選択する手段は、そのフレームを蓄積して、後で表示し、また後でフレームを呼び戻す手段を含む。

【 0 0 4 1 】

好ましい実施例では、その抽出された埋め込まれたデータは、eコマースなどの対話型通信を容易に行うために、インターネットwebサイトにアクセスする手段に用いられる。

【 0 0 4 2 】

本発明を用いることで、広告主により作成される広告は目立たなくなる。すなわち、テレビ視聴者は、もし要求があれば、対話なしで、その番組を見ることができる。別法として、テレビ視聴者は、番組を見て、その番組のフレームをこま止めし、マウス、キーボード

50

、またはＴＶのリモコンを用いてオブジェクトをクリックして、インターネット上で、ｅコマースのトランザクションを容易に実行できる。このような機能を果たすとき、テレビ視聴者は、ＶＤＵ画面を分割して、一方の部分が、その流れている番組を引き続き表示し、また他方の部分が静止画像を表示し、またインターネット情報が転送されるようにする場合がある。

【 ０ ０ ４ ３ 】

本発明は、デジタルビデオの娯楽の多数の面（特に、以下に示される放送）に使用できる。すなわち、

- １．正規のテレビ番組または映画における対話型製品配置。
- ２．ファッションＴＶ。
- ３．音楽ＴＶ。
- ４．教育番組。

10

【 ０ ０ ４ ４ 】

ｅコマースは、例えば、チケット販売を促進する場合がある。

【 ０ ０ ４ ５ 】

本発明は、テレビ視聴者が、関連のない情報でまごつくことなく、ビデオ信号番組内の興味のある品目について、さらなる情報を選択できるという利点がある。これは、その情報が広告の形式を取る場合に特に役立つものであって、ビデオ番組において見られるオブジェクトに、対応付けられる多重化（埋込み）データを持たせて、これらのオブジェクトに関係のあるさらなる情報（ビデオ信号内の情報か、または、データベースに蓄積された情報のいずれか）にリンクを与えることで、あるいは、インターネットwebサイトにアクセスすることで、得られる。

20

【 ０ ０ ４ ６ 】

広告主に関する限り、本発明は、広告のターゲットを正確に関連視聴者にしぼることができる、また広告を阻止する装置（例えば、Ｖチップ）では、広告がユーザに届くのを阻止できないという利点がある。複数の広告主は、その広告を、ビデオ番組シーケンスの各フレームに対応付けるから、本発明は、番組制作者や番組提供者に対して広告収入を維持または増加させながら、個々の広告主には、広告費用を減らす可能性がある。このようなやり方で、ビデオ番組信号の各フレームのデータ搬送潜在力が最大限に活用され、また放送チャンネルのデータ搬送容量が最大限利用される場合がある。本発明は、例えば、ゴールデンアワーのときに、放送用ＴＶに２分の割当て時間を与えるだけの余裕のない場合がある広告主に対して、新たな民主主義を生み出す方法を導くものと思われる。これは、本発明により、１オブジェクト当たりの複数の広告主、および／または、１フレーム当たりの複数のオブジェクトが、広告収入モデルの柔軟性を高めることができるからである。

30

【 ０ ０ ４ ７ 】

例えば、音楽ビデオの分野では、そのコンテンツは、レコード会社のバンドの音楽を売り込む目的で使用される場合があり、またミュージシャンとの対話により、ユーザは、音楽を買い付けて、じかに音楽をダウンロードする場合がある。

【 ０ ０ ４ ８ 】

さらに、複数の広告主は、同一の時間枠を買っている場合があり、言い換えれば、広告主のコンテンツは、番組コンテンツ内に完全に融合されていて、この広告コンテンツがユーザによりダウンロードされるまで、読まれることはない。したがって、デジタルＴＶ番組のあらゆるフレームは、広告収入として使用される場合がある。ｅコマースのデータベースは、WebサイトのURLアドレスからカタログ、パンフレット、ビデオプロモーションまで、広告主に関するあらゆる関連データを、ｅコマース・トランザクション設備に蓄積する場合がある。

40

【 ０ ０ ４ ９ 】

テレビ視聴者が、例えばマウスを用いてオブジェクトをクリックして、そのオブジェクトを選択するときには、そのオブジェクトは、多数の広告主を表わす場合がある。例えば、ミュージシャンは、衣類、腕時計、化粧品、楽器を宣伝して、テレビ視聴者が、このオブ

50

ジェクトに対応付けられる宣伝販売品のリストから選択することがある。したがって、衛星放送の伝送速度を最大にするプッシュ技術の手法が提示される。ユーザは、実際にトランザクションを行いたいと思う場合には、インターネットを介するリターンパスしか必要としない。

【 0 0 5 0 】

本発明は、ここでは、例示として、添付図面を参照して説明される。添付図面において、図 1 は、本発明の対話型システムのブロック模式図を示す。

【 0 0 5 1 】

図 2 は、或る番組の関連フレーム内の 1 オブジェクトに対応付けられる対話型コンテンツデータを生成するビデオ番組処理のブロック模式図を示す。

10

【 0 0 5 2 】

図 3 は、関連カメラ・ショット群によって得られた番組シーケンスを示す模式図を示す。

【 0 0 5 3 】

図 4 は、ショット群を生成させる、図 2 に示されるパーサのブロック模式図を示す。

【 0 0 5 4 】

図 5 は、ビデオ番組のキーフレームを示す。

【 0 0 5 5 】

図 6 は、図 5 のキーフレーム内で選択された 1 オブジェクトを示す。

【 0 0 5 6 】

図 7 は、ビデオ番組内のオブジェクトを 1 フレームずつ識別する流れ図を示す。

20

【 0 0 5 7 】

図 8 は、このオブジェクトを 1 フレームずつ追跡する、図 2 に示されるオブジェクト・トラッカの流れ図を示す。

【 0 0 5 8 】

図 9 は、図 2 に示されるストリーマの流れ図を示す。

【 0 0 5 9 】

図 1 0 は、対話型コンテンツデータをビデオ番組信号と組み合わせるブロック模式図を示す。

【 0 0 6 0 】

図 1 1 は、本発明に用いられるデータパケットの構造を示す。

30

【 0 0 6 1 】

図 1 2 は、ビデオ番組信号から対話型コンテンツデータを抽出するやり方を、ブロック模式図で示す。

【 0 0 6 2 】

以上の図では、同じ参照番号は、同じ部分を示す。

【 0 0 6 3 】

図 1 に示される対話型システムは、少なくとも 1 つのオブジェクトに対応付けられる対話型コンテンツデータを表わすデータ・シーケンスを生成する装置 2 0 0 を持ち、また、この対話型コンテンツデータは、デジタルビデオ番組を表わすビデオデータとオーディオデータで多重化 1 0 8 0 されている。この説明された実施例では、データトランスポート・ストリーム 1 0 0 1 が、衛星放送装置 2 0 のヘッドエンド装置 1 0 に与えられ、衛星放送装置 2 0 が衛星 2 5 に送信し、次に、それぞれが衛星放送受信用のパラボラアンテナ 3 1 を備えている複数のユーザ/テレビ視聴者 3 0 に、衛星 2 5 が、その放送信号を再送する。この受信信号は、TVカードを備えたPC 4 0 に加えられて、テレビ視聴者が対話を行う場合がある。この受信放送信号はまた(あるいは、別法として)、デジタルテレビジョン 5 5 のセットトップ・ボックス 5 0、またはセットトップ・ボックスの電子回路を組み入れたテレビジョンに与えられる場合がある。テレビ視聴者がTV上のアイコンを操作して、オブジェクトを選択し、かつ提供されることのあるメニューや操作と対話するために、このセットトップ・ボックスには、キーボード(図には示されてない)またはマウス 5 6 が備えられる場合がある。同様に、PC 4 0 は、キーボードを備えるが、慣習になって

40

50

いる通り、使用方法がセットトップ・ボックスと同じであるようにマウスも備えることがあり、したがって、テレビ視聴者/ユーザは、オブジェクトを選択して、対話型通信を行うことができる。PCへの入力やPCからの出力は、モデム45を通じて、公衆電話網60（例えば、PSTN、ISDN、xDSL、または衛星の場合がある）に対して行われ、同様に、セットトップ・ボックス50も公衆電話網60に接続される。公衆電話網60により、複数のテレビ視聴者が、eコマース管理システム70（専用管理システム、または、インターネット・サービス・プロバイダと相互にリンクされたシステムの場合がある）と相互接続される。ビデオ番組が放送されるシステムでは、eコマース管理システム70は、放送提供システムに接続されて、eコマース管理システム70が、その放送番組と結びついて、テレビ視聴者に送られるオブジェクト間の関係を持続させることができるようにしている。

10

【0064】

本発明のシステムでは、例えば人であることもあり、衣類、腕時計、化粧品、楽器などの物理的なオブジェクト、あるいは、例えば商標などであることもあるオブジェクトは、そのオブジェクトを搬送する該当番組のビデオ番組信号に多重化された（埋め込まれた）オブジェクトと対応付けられるデータを持っている。これを達成するためには、ビデオ番組全体にわたって、1フレームずつオブジェクトを識別し、追跡する必要がある。この説明された実施例では、ビデオ番組が放送されるが、このビデオ番組は、ビデオ番組を蓄積するデジタルビデオディスク（DVD）、テープ、または任意の公知手段上にあるはずである。次に、1オブジェクトを選択するとき、テレビ視聴者は、そのオブジェクトに関する細部と対話することができる。例えば、このオブジェクトが、ポップ音楽ビデオのミュージシャンである場合には、その音楽レコード、このミュージシャンが着て宣伝する衣類が、インターネット上で入手される場所に関して、情報が得られることもある。

20

【0065】

第1の段階は、そのオブジェクトまたは各オブジェクトが現れる番組のあらゆるフレームに、そのようなオブジェクトに動的に対応付けられる対話型データを生成することである。例えば、5分のビデオ・シーケンスは、一般に、7,500のフレームから成るが、それに対して、90分の映画は、135,000のフレームであることもある。

【0066】

入力ビデオ番組がデジタル形式を取らない場合には、この番組は、まず最初に、それ自体公知の手段により、デジタル化されなければならない。

30

【0067】

図2を参照して、或る番組の関連フレーム内の1オブジェクトに対応付けられる対話型コンテンツデータを生成する装置200が示されている。デジタルビデオソース201からのデジタル化番組は、図4に詳細に示されるパーサ400により、関連ショット300（図3に示される）に分けられる。本発明に関連して、「ショット」は、或るシーンの一回のカメラ「撮影」である。5分のビデオ・シーケンスは、一般に、一連のフレーム F_n から成る100の上記ショットまたは編集を持つ場合がある。この場合、例えば、 $F_n = 25 \times 60 \times 5 = 7,500$ のフレームであるが、それに対して、90分のビデオは、何千ものショットを持つ場合がある。どのフレームにおいてショット300が変わるかを示すオブション編集リスト202といっしょに、デジタル化ビデオ番組を提供する場合には、これを利用して、その番組を別々のショット300に分けることがある。

40

【0068】

基本的には、パーサ400は、このビデオを分解して、一群のシーケンス321、322、323にする（図3）。これらのシーケンスは、意味的に関係のある一連のショットから成り、例えば、或るシーケンスは、ポップ音楽ビデオのリードシンガーを呼び物にするすべてのショットを含む場合がある。それゆえ、パーサ400の働きは、この番組を分解して、共通のスレッドにより統合されるシーケンスにすることである。後述されるトラック800が、オブジェクトが見つかりそうなシーケンス内で、オブジェクトだけを探すような操作が必要である。パーサ400は、ショットの変化、カメラアングルの変化、ワイ

50

プ、ディゾルブ、および、他の可能な任意の編集機能または光遷移効果を検出する。図4に示されるパーサ400は、デジタル番組を受け取り、また、例えば、そのビデオ番組のそれぞれの連続するフレームのエッジマップを比較して、所定のしきい値を超えるような前記エッジマップの位置変化が発生するときに、ショットの終わりが現れると定めることで、ショットの終わりが検出される410。ショットの終わりを判定するのに用いられる基準420は、或るオブジェクトに対応付けられるデータをビデオ番組シーケンスに埋め込んでいるユーザにより、カット/ショット検出プログラムに入力される。異なるショットの情報は、編集リスト430に入れられる。

【0069】

次に、いくつかのフレームは、各ショット300からのキーフレーム識別子440の中で選択されて、そのショット300を表わすキーフレーム500となる(図5も参照のこと)。ショットごとに、2つ以上のキーフレームを必要とする場合があり、そこでは、ショット300が、例えば、パンまたはズームなどの複雑なカメラ移動を含んで、1つのキーフレーム500が、全コンテンツを表わさないようにしている。さらに、ビデオ番組がポップ・グループのものであって、そのシーケンスが、すべてのバンドメンバのロングショットから始まって、すみやかにリードシンガーをズームし、最後に、そのリードシンガーの顔を画面いっぱいに映し出す場合には、いかなる単一のフレームも、ショット全体を表わさないであろうが、ただし、例えば3つのキーフレームの妥当な選択は、最初のフレーム311、ズームのほぼ半ばのフレーム312、および最後のフレーム313(図3に示される)であろう。したがって、ショット300のビデオコンテンツを表わすキーフレーム311、312、313が自動的に選択される。

【0070】

図3と図4に示されるように、ショット300は、シーン・グループ450により、シーケンスにまとめられて、シーン・グループ450が、各ショット300からのキーフレーム311~313を、他のそれぞれのショット304、307からのキーフレーム311~313と比較する。これは、色相関図表(colour correlelograms)、データマップ、テクスチャなどの下位の特徴を用いて、それらのショットからのキーフレームを比較することで行われる。同様なコンテンツを持つショットは、シーン・グループ450により階層構造にまとめられて、共通のテーマを持つショット群にする。例えば、ポップ音楽ビデオでは、いくつかの異なるセットが使用されることもあるが、ビデオ内の多数の個所に1セットが現れる場合がある。シーン・グループ450は、或る水準の同一セットと、別の水準での同一セットの同様なタイプのショット/シーケンスを用いて、ショット300、304、307のシーケンスをまとめる。このようなやり方で、シーケンスの階層構造(コンテンツ・ツリー460と呼ばれる)が構築される。このようにまとめる目的は、対話型コンテンツデータで識別されるオブジェクトを選択するのに役立て、さらに、この映画のすべてのショットを通じてではなく、関連ショット300、304、307内でのみ、特定のオブジェクトを確実に探せるようにすることで、後述されるように、後で、ビデオ番組を通じて、その選択されたオブジェクトを追跡する効率も高めることである。したがって、パーサ400は、或るオブジェクトの追跡を関連ショット(すなわち、シーケンス321、322、323内のショット)に限定することで、このトラックの効率を高めるだけでなく、強力なブラウジング・オブジェクト選択装置を提供することによっても、ビデオ番組の全構造と複雑さをユーザに理解させるのに役立てる。

【0071】

ショット300をシーケンス321、322、323にまとめた後で、シーケンス・キーフレームは、各ショットのキーフレーム311、312、313から選択されて、そのシーケンスを表わす。次に、或るオブジェクトを表わす対話型コンテンツデータをビデオ番組に入力したいと思っているユーザは、これらの上位のキーフレームを使用して、ユーザに興味のあるオブジェクトを含むショットのシーケンスを選択する場合がある。これらのキーフレームは、好ましくは、シーケンス321、322、323のコンテンツ・ツリー460の階層構造を表わす形式でユーザに提示される。シーン・グループ450の出力4

10

20

30

40

50

70は、単一ショットのシーケンスの数、これらのシーケンスを表わすキーフレーム311、312、313、および、キーフレームで反映される通り、これらのシーケンスの相互の階層関係を示すコンテンツ・ツリーである。

【0072】

対話型コンテンツデータをビデオ番組に挿入するつもりユーザは、図5に示されるように、キーフレームの階層構造を見て、第1のキーフレーム311を選択する。好ましい実施例では、すべてのキーフレームは、縮小化された形式で画面上でユーザに提示され、またユーザは、その縮小化キーフレーム上にカーソルを置いて、そのキーフレームを選択する場合がある。次に、キーフレーム311からオブジェクトを選択するために、そのキーフレームのフルサイズ版をユーザに提示する。次に、ユーザは、この番組ビデオに埋め込まれた対話型コンテンツデータに対応付けるつもりキーフレーム311内のオブジェクト600に、マウスなどのポインティング・デバイスを用いてマークを付ける(図6に示される通り)。オブジェクトの周りに境界線囲み610を描くことで、このオブジェクトにマークが付けられる場合がある。キーフレーム311内のオブジェクト600を選択するためには、ユーザは、カーソルが左上の角にあるときにマウスボタンをクリックし、そのマウスのカーソルを、オブジェクト600の右下の角までドラッグする。このようにして、その選択されたオブジェクト600の周りに境界線囲み610が表示される。

10

【0073】

例えば、ポップグループのツアー日程に関するデータ情報を埋め込むためには、キーフレーム全体を選択する場合がある。このキーフレームがキーボードを含む場合には、このキーボードを選択して、キーボードメーカーに代って、キーボードを宣伝し、かつ/またはキーボードを販売する場合がある。さらに、キーフレーム内に現れるリードシンガーも選択される場合がある。図6に示される境界線囲み610は長方形であり、これは、好ましいデフォルト形状であるが、ただし、平行四辺形またはユーザ定義の多角形などの他の形状を使用することもある。

20

【0074】

オブジェクトが選択され、図7に詳しく示される通りに、そのオブジェクトが識別される600。このように、ユーザは、オブジェクト710を識別し、オブジェクト600をポイントクリックして、初期オブジェクト選択を行う715。各オブジェクト600が、キーフレーム311内で選択されると、連続フレームを通してオブジェクトを追跡するのに用いられる属性を計算して、それを、すでに記録されたオブジェクト720の属性と比較し、新規オブジェクトが、そのフレームに関して、すでに記録された他のすべてのオブジェクトとはっきり異なることを保証する。これらの属性は、これらの属性に基づく時系列統計量だけでなく、このオブジェクトの形状、サイズ、位置、色、テクスチャ、輝度勾配の任意のものも含むことがある。新規オブジェクトが、以前に記録されたオブジェクトにあまりにも似ている場合には、この新規オブジェクトに関する追加情報を与えるようにユーザに指示する。そうでない場合には、このオブジェクトの属性が記録される。

30

【0075】

ブロック725における選択されたオブジェクトは、そのフレームの他の部分から隔離されていると考えられる。次に、ユーザは、境界線囲み610を変更して、他のオブジェクトに対して、さらに正確に区別する730ことで、オブジェクト600を定義するか、または、2つのオブジェクトが、画面上の同一場所を占有するように重なり合う場合には、ユーザは、その重なり合ったオブジェクトのそれぞれにランクを割り当てることで、どちらのオブジェクトが優先するか示すことがある。例えば、上に示された例では、ポップグループのツアー日程に関する情報は、フレーム全体に対応付けられものであって、その情報には低いランクが与えられて、例えば、このフレーム内のどこにでも現れる任意の他のオブジェクトが、さらに高いランクをつねに持ち、フレーム311全体に対応付けられるデータにより無効にされないようにしている場合がある。このような処理は、シーケンス321、322、323をそれぞれ表わすキーフレーム311のそれぞれに対して繰り返される。

40

50

【 0 0 7 6 】

このキーフレーム内で、各オブジェクトが選択されると、次のステップは、このオブジェクトをデータを用いて識別して、このデータを、そのオブジェクトとともに埋め込むことである。好ましくは、データのレコード・アドレスは、データベースに保持され、そのデータが、特定のオブジェクトに対応付けられるか、あるいは、別法として、レコード・アドレスを使用しないで、データ自体が埋め込まれる場合がある。好ましくは、グラフィカル・ユーザ・インタフェース 7 5 0 を用いて、そのデータを表わすアイコンを、フレーム 3 1 1 内のオブジェクト 6 0 0 上にドラッグする。

【 0 0 7 7 】

それにより、ユーザは、「ドロップ・アンド・ドラッグ」技法を用いて、その区分化されたフレーム内の各オブジェクトに、広告コンテンツを追加して、例えば、広告主を表わすアイコンを、マウスを用いてオブジェクト上にドラッグして、関連データは、そのオブジェクトに自動的に埋め込まれるようになっている。この処理は、すべてのオブジェクトが、対話型データとともに埋め込まれるまで続けられる。それにより、或るオブジェクトを表わすデータが、ビデオ番組信号に埋め込まれて 7 6 0、オブジェクト 7 6 5 に対応付けられる対話型コンテンツデータと、出力 7 7 0 としてそれぞれの埋め込まれたコンテンツデータに対応付けられるキーフレームの数とを提供する。

【 0 0 7 8 】

したがって、識別子 7 0 0 は、これらのオブジェクトを識別して、各シーケンスからの少数のキーフレームにアクセスし、コンテンツを埋め込むことにより、そのコンテンツを、それらのオブジェクトに埋め込ませる。

【 0 0 7 9 】

オブジェクト・デスクリプタをキーフレームに埋め込んで、コンテンツを提供した後で、ビデオ番組の連続フレームを通して、オブジェクトを追跡することが必要である。

【 0 0 8 0 】

図 8 を参照して、ビデオ番組全体にわたって、或るオブジェクトを追跡することが必要であり、さらに、或るオブジェクトが、フレーム内で移動して、時折、他のオブジェクトにより覆い隠されるか、あるいは、見られているフレームから完全に去るときにも、その追跡の必要がある。基本的には、これらのオブジェクトは、一連の境界線形状と下位の特徴機能（例えば、形状、エッジ、色、テクスチャ、輝度勾配の情報）を組合わせたものとして定義される。このようなオブジェクトの表現を用いて、これらのオブジェクトが、ビデオ・シーケンスの残りのフレームを通して、対話形式で追跡される。複数のオブジェクトが追跡されて、それらのオブジェクトが現れるすべてのフレーム内に突き止められたときには、最初のキーフレーム 3 1 1 に埋め込まれた関連コンテンツが、すべてのシーケンスの残りのフレームに自動的に追加される。これは、特に図 8 に示されるオブジェクト・トラッカ 8 0 0 の働きである。

【 0 0 8 1 】

カットされていないシーケンスや選択されたオブジェクト 8 1 0 は、フレーム内のオブジェクトを比較するのに用いられる下位の表現 8 2 0 に変換される 8 1 5。すべてのフレームでは、距離測度を利用して、各フレーム内に各オブジェクトを突き止める。都合の良い距離測度は、それ自体公知のハウスドルフ測度であるが、ただし、この測度は、他の技法を用いると、大きくなる場合がある。順次フレームを通してのオブジェクトの追跡 8 2 5 が、繰り返して行われ、それにより、このオブジェクトは、当初、キーフレームでは、エッジ検出を行って、境界囲み 6 1 0 に囲まれたエッジを区分化することで得られた二次元幾何学形状として定義される。次に、オブジェクト 6 1 0 が、次のフレーム 3 1 2 内に突き止められて、そのオブジェクトの属性が、フレーム間に発生した位置と形状の変化を反映するために更新される。次に、これらの新たな属性を持つオブジェクトが、次のフレーム内に突き止められて、トラッカ 8 0 0 の処理が続く。

【 0 0 8 2 】

一連のショットのすべてのフレーム内の各オブジェクトの位置が決定されると、閉塞の障

10

20

30

40

50

害を除く位置と、オブジェクトの出口や入口の後処理が行われる。

【0083】

このシステムは、照明変化、閉塞、カメラ移動、ショット、ブレイク（中断）、光遷移効果（例えば、ワイプ、フェード、ディゾルブ）に耐えうる。このシステムは、様々な公知技法を使用して（例えば、選択が、このシーケンスの視覚的な複雑さによって決まるようなアルゴリズム群を用いて）、あらゆるタイプの視覚環境において自動的に追跡できるようにしている。これらのアルゴリズムは、それ自体公知のものであるが、ただし、当業者は、ヒューリスティックスを使用して、追跡用に性能を最適化する場合がある。次に、キーフレーム内のオブジェクトに追加されるデータは、ビデオ・シーケンス830全体にわたって、オブジェクトを追跡するときに、すべてのフレーム内のオブジェクトに自動的に追加される。

10

【0084】

ユーザは、生成されたトラックを再検討し、訂正があれば、入力する835。このような訂正は、その再検討されたシーケンスを、誤りのあるフレームの所で停止し、誤っているオブジェクトをクリックして、そのオブジェクトを、その正しい位置までドラッグして、行われる。このように、グラフィカル・ユーザ・インタフェースを用いて、オブジェクトの位置が誤って識別された位置でビデオを停止して、境界線囲み610を、その正しい位置にドラッグ・アンド・ドロップし、それにより、そのフレームに対して、このオブジェクトの属性を定義し直して、後続フレーム用のオブジェクトの定義を、その新規定義に基づかせ、それにより、検証済みのトラック845を生成する。

20

【0085】

最後に、このビデオのすべてのシーケンス内のすべてのフレームは、その関連オブジェクトが、識別され、かつ対話型コンテンツデータ850とともに埋め込まれる。

【0086】

トラック800からの出力は、図9に示されるストリーマ900に加えられ、ストリーマ900内で、その埋め込まれたコンテンツデータの妥当性がチェックされ、必要であれば、その埋め込まれた対話型コンテンツデータを出力する順序が、視聴覚フレームと同期化される。

【0087】

ストリーマは、すべてのフレーム内のすべてのオブジェクトが、埋め込まれたコンテンツデータ850を持つかどうか、また、そのコンテンツにラベルが付けられ、かつ、このコンテンツが妥当であって、エンコード設定920を用いて、エンコードとエラーチェック910に作用するかどうか確かめる。コンテンツに正しくラベルが付けられ、かつコンテンツが妥当であることの検証940が行われ、また出力930は、完全な放送適合のトランスポート・ストリーム（例えば、MPEG-2/DVBのオーディオ、ビデオ、および、埋め込まれたオブジェクトとコンテンツデータ）であるか、あるいは、埋め込まれたオブジェクトとコンテンツデータ単独としてのものである。

30

【0088】

ストリーマ900は、埋め込まれたコンテンツデータが、3つの種類（すなわち、フレーム同期データ、セグメント同期データ、または、ジャストインタイム式にストリーム再生されるデータ）のどれに属するか判定しなければならない。フレーム同期データは、オブジェクトの位置、形状、ランクから成り、またデータを指すポインタの1テーブルを指すポインタは、ソース201からのビデオ番組内の正確なフレーム番号に対応付けられる場合がある。セグメント同期データは、オブジェクトが変わると、埋め込まれたコンテンツデータも変わるように、その埋め込まれたコンテンツデータを指すポインタのテーブルを更新する目的で使用される。このデータは、コンテンツが変わるときのフレーム番号に対応付けられることもある。「ジャストインタイム」式にストリーム再生されるデータは、これらのオブジェクトのどれかにより求められる前に、エンドユーザにストリーム再生されなければならない。次に、このトランスポート・ストリームは、MPEG-2/DVB適合のパケットにパケット化される。

40

50

【 0 0 8 9 】

完全に埋め込まれた視聴覚番組が求められる場合には、図 1 0 に示されるように、パケット化トランスポート・ストリームとビデオ番組がいっしょに多重化される。

【 0 0 9 0 】

図 1 0 を参照して、この埋め込まれたビデオ番組を構成する異なる要素は、組合わされて単一のトランスポート・ストリーム 1 0 0 1 にして、ネットワーク・オペレータによる放送に備えている。この番組は、ビデオストリーム 1 0 1 0 とオーディオストリーム 1 0 2 0 から成り、それらのストリームは両方とも、圧縮されてない。ビデオデータ 1 0 1 0 もオーディオデータ 1 0 2 0 も、それぞれの M P E G - 2 基本エンコード 1 0 1 5 と 1 0 2 5 において符号化され、圧縮されて、それぞれ、データ 1 0 3 0、1 0 3 5 のエレメンタリーストリームを生成する。M P E G - 2 適合のデータ・シーケンス 9 3 0 は、エラー検査されて 1 0 3 7、データ 1 0 4 0 のエレメンタリーストリームを生成する。エレメンタリーストリーム 1 0 3 0、1 0 3 5、1 0 4 0 は、パケット化装置 1 0 5 0、1 0 5 5、1 0 6 0 に加えられ、それらのパケット化装置はそれぞれ、データを蓄積して、固定サイズ・ブロックにし、さらに、それらのブロックには、プロトコルヘッダが追加される。それらのパケット化装置からの出力は、パケット化エレメンタリーストリーム (P E S) 1 0 7 0 と呼ばれる。パケット化エレメンタリーストリーム 1 0 7 0 は、デジタル制御データ (P S I) 1 0 7 5 と組み合わせて、システムクロック 1 0 8 5 を有するシステム層多重化装置 1 0 8 0 に加えられる。P E S パケットは、情報 1 0 3 0、1 0 3 5 とデータ・シーケンス 9 3 0 の連続エレメンタリーストリームを、パケットのストリームに変換する機構である。このエレメンタリーストリームは、P E S パケットに埋め込まれると、タイムスタンプと同期化される場合がある。これは、受像機 (P C または T V) が、その埋め込まれたビデオ番組を構成するすべてのビデオストリーム、オーディオストリーム、データストリームの相互の関係を決定できるようにするために必要である。

【 0 0 9 1 】

各 P E S パケットは、システム多重化装置 1 0 8 0 に送られる。そこでは、これらのパケットは、トランスポート・パケットにカプセル化されて、放送に使用されるトランスポート・ストリーム 1 0 0 1 を形成する。この点で、トランスポート・ストリーム 1 0 0 1 は、1 8 8 バイトのブロックでパケットを搬送し、またトランスポート・ストリーム 1 0 0 1 は、放送のために、ネットワーク・オペレータに送られる完全ないわゆる e M U S E チャンネルを構成する。本質的に、このトランスポート・ストリームは、固定長のパケットを用いて、複数のストリームを組み合わせる汎用方式である。

【 0 0 9 2 】

パケットの構造は、図 1 1 に示されている。図 1 1 に示されるパケット 1 1 0 0 は、1 バイトの同期バイト、1 3 ビットのパケット識別子 (P I D)、およびこのパケットを処理すべき方法を示す一組のフラグを有するヘッダ 1 1 1 0 を持っている。このトランスポート多重化装置は、異なるパケット識別子を各 P E S 1 0 7 0 に割り当てて、個々のストリームを一意的に識別する。このようなやり方で、このパケット化データ・シーケンス 9 3 0 が一意的に識別される。これらのエレメンタリーストリームの同期化は、タイムスタンプをトランスポート・ストリーム 1 0 0 1 で送ることで、容易となる。

【 0 0 9 3 】

以下の 2 つのタイプのタイムスタンプが用いられる場合がある：

- 1 . 現在時間を示す基準タイムスタンプ、すなわち、クロック 1 0 8 5 の情報。
- 2 . デコーディング・タイムスタンプ。

【 0 0 9 4 】

デコーディング・タイムスタンプは、P E S に挿入されて、データストリームが、ビデオストリームおよびオーディオストリームと同期化される必要のある厳密な時間を示す。デコーディング・タイムスタンプは、動作では、基準タイムスタンプに頼っている。トランスポート・ストリームがブロードキャスト (放送) された後で、P C または T V は、これらのタイムスタンプを使用して、ビデオストリームとオーディオストリームに関して、そ

のデータ・シーケンスを処理する。

【 0 0 9 5 】

受像機（PCまたはTV）が、このチャンネルをデコードする方法を知るためには、受像機は、パケット自体のPIDテーブルとともに、トランスポート・ストリーム1001内で別々のパケットとして送られる特定番組向け情報（PSI）ラベルとして知られた一組の信号テーブルにアクセスする必要がある。受像機が或るチャンネルをデコードできるようにするのに必要なテーブルが2つある。第1のテーブルは、番組対応付けテーブル（PAT）1130であって、これは、衛星放送の範囲内で使えるすべてのチャンネルをリストし、かつ、そのテーブルの識別を容易にする0のパケットID（PID）を持っている。この例では、eMUSEチャンネル、すなわち、ビデオ番組を搬送するチャンネルが、PID111として表わされている。

10

【 0 0 9 6 】

番組テーブル・マップ（PMT）1140は、埋め込まれたビデオ信号に入っているすべてのエレメンタリーストリームを識別する。各エレメンタリーストリームは、PID値により識別される。例えば、ビデオカメラ1からのビデオはPID71である。データ・シーケンス930は、図11の例においてPID値92を持っている。受像機のビデオデコーダとオーディオデコーダは、PMTテーブルを探索して、デコードするのに適したパケットを見つける。同様に、埋め込まれたデータを検索するプログラムが、PMTを探索して、図11の例ではPID92であるデータシーケンスを見つける。次に、このデータ検索プログラムは、これらのパケットをフィルタリングして選別し、選別されたパケットを、適切なビデオおよびオーディオと同期化させて、ユーザが様々なオブジェクトを選択できるようにしている。

20

【 0 0 9 7 】

対話型コンテンツデータをビデオ番組信号に埋め込んだ後で、このビデオ番組信号が、ブロードキャスト（放送）される。ここでは、このデータを受け取って、検索する方法を、図12を参照して説明する。

【 0 0 9 8 】

ユーザのPC40またはデジタル・セットトップ・ボックス50上にある衛星放送受信カード1210にハードウェアが設けられており、また、ソフトウェア（例えば、インターネットへのアクセスを容易に行うもの、Internet ExplorerやNetscapeなどのインターネットブラウザ、また、TV用として、Sun社のオープンTVオペレーティング・システムと互換性のあるもの）により、テレビ視聴者は、放送中の動的オブジェクトと対話することができる。

30

【 0 0 9 9 】

受け取られたMPEG-2/DVB信号は、MPEG-2ビデオ1215、MPEG-2オーディオ1220、データ・シーケンス930に分けられ、またデコードされたビデオ1225、オーディオ、データ・シーケンスは、シンクロナイザ1230に加えられる。対話型コンテンツデータが埋め込まれているビデオ番組を含め、シンクロナイザから出力されるものは、PCのVDUまたはTV画面により、表示される1240。

【 0 1 0 0 】

ユーザは、興味のある1オブジェクトを含む1フレームにおいて、マウス56をクリックするか、または、リモコンボタンを押し、それにより、画面上の表示が2つに分割される。例えば、左側の画面では、ビデオ番組が、常態として流れ続け、また右側の画面では、マウスをクリックしたときにアクティブであったフレーム内にあるオブジェクトが、カットされた部分として表示されて、それらの介在スペースが空白になる。次に、ユーザは、興味のあるオブジェクトをクリックして、そのオブジェクトがどの広告主を表わしているのか知る。例えば、ユーザがリードシンガーをクリックする場合には、画面は、そのリードシンガーだけを表示し、また文字の広告主リスト、あるいは、アイコンベースの広告主表示が見られる。ユーザが広告主の名前またはアイコンをクリックする場合には、ユーザは、宣伝された製品をじかに見ることになる。

40

50

【0101】

このサイトと対話した後で、ユーザは、eコマースのトランザクションを通じて、製品を購入することに決める場合がある。さらに、ユーザが、リードシンガーのスーツをクリックすると、ストリーム再生されたデジタル放送の一部として、スーツメーカーの全カタログが入手できるようになっていてもよい。このようなインターネットを介するリターンパスは、ユーザがフレーム内の多数のオブジェクトの中から選択すると、広告情報をストリーム再生するプッシュ技術の手法が、データ・シーケンス930により起動されるから、トランザクションをまったく容易にする。

【0102】

ユーザは、上に説明されるようなオンラインの手法で放送と対話できるが、別法として、データは、オフラインで見られることもある。すなわち、テレビ視聴者が、引き続き、或る番組を見ているときに、ユーザは、放送中に様々なフレームを選択し、それらのフレームを蓄積して、後で、その関連データを検索する場合がある。このデータを蓄積するのに十分なローカル・メモリがない場合には、ローカル・データベースまたはリモート・データベース（例えば、Webサイト）に前記データのアドレスを蓄積して、後で、エンドユーザは、それらのデータベースにアクセスして、このデータを検索することができる。つぎに、ユーザは、興味のあるオブジェクト600を、マウスまたはリモコンを用いて選択し、次に、オブジェクト600と、そのオブジェクトに対応付けられるデータ要素のメニューを示す別の画面を表示することがある。ユーザは、メニュー項目の1つをクリックし、その宣伝された製品上のデータを直接に見るか、あるいは、インターネット上のWebサイトにアクセスすることができる。別法として、ユーザがメニュー項目を選択するとすぐ、放送信号に埋め込まれているカタログが見られる場合がある。

【0103】

エンドユーザがアクセスするデータは、放送信号でストリーム再生される場合があるか、あるいは、ビデオ・シーケンスを見る前に、エンドユーザの装置に前もってインストールされるローカル・データベースに保持される場合がある。放送でストリーム再生された情報を見ているときに、特定の番組に対応付けられる情報は、その番組と同時にストリーム再生されて、ユーザのマシンに蓄積される。ユーザが或るオブジェクトを選択すると、このようなローカルデータが表示される。

【図面の簡単な説明】

【図1】 図1は、本発明の対話型システムのブロック模式図を示す。

【図2】 図2は、或る番組の関連フレーム内の1オブジェクトに対応付けられる対話型コンテンツデータを生成するビデオ番組処理のブロック模式図を示す。

【図3】 図3は、関連カメラ・ショット群によって得られた番組シーケンスを示す模式図を示す。

【図4】 図4は、ショット群を生成させる、図2に示されるパーサのブロック模式図を示す。

【図5】 図5は、ビデオ番組のキーフレームを示す。

【図6】 図6は、図5のキーフレーム内で選択された1オブジェクトを示す。

【図7】 図7は、ビデオ番組内のオブジェクトを1フレームずつ識別する流れ図を示す。

【図8】 図8は、このオブジェクトを1フレームずつ追跡する、図2に示されるオブジェクト・トラックの流れ図を示す。

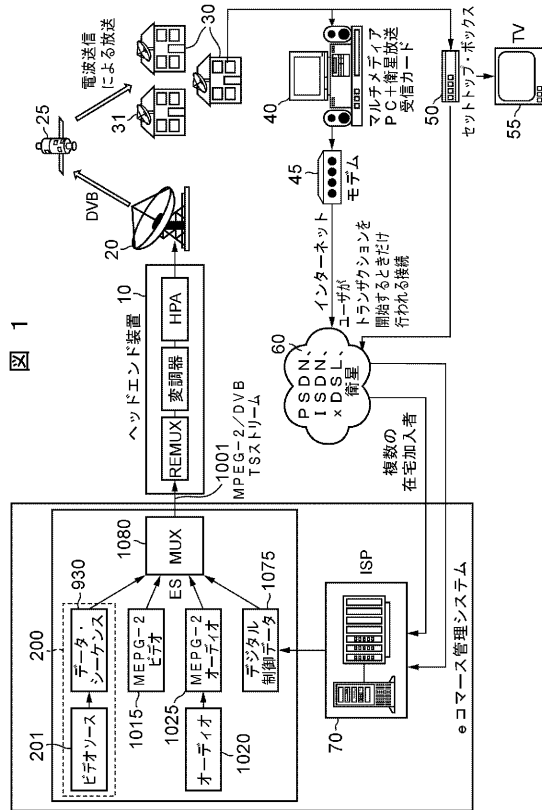
【図9】 図9は、図2に示されるストリーマの流れ図を示す。

【図10】 図10は、対話型コンテンツデータをビデオ番組信号と組み合わせるブロック模式図を示す。

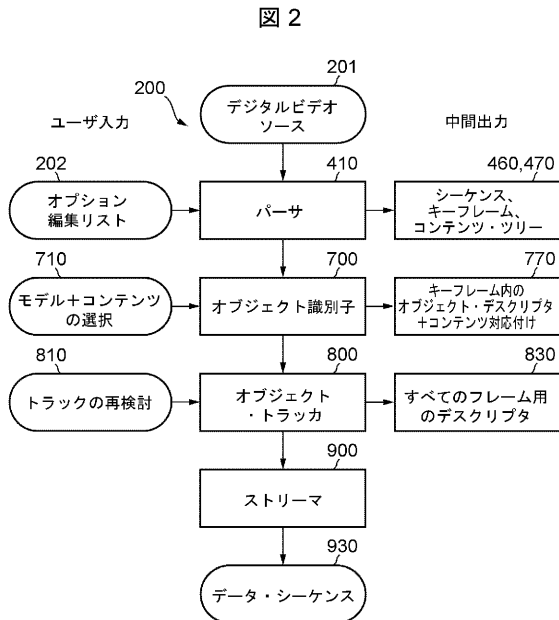
【図11】 図11は、本発明に用いられるデータパケットの構造を示す。

【図12】 図12は、ビデオ番組信号から対話型コンテンツデータを抽出するやり方を、ブロック模式図で示す。

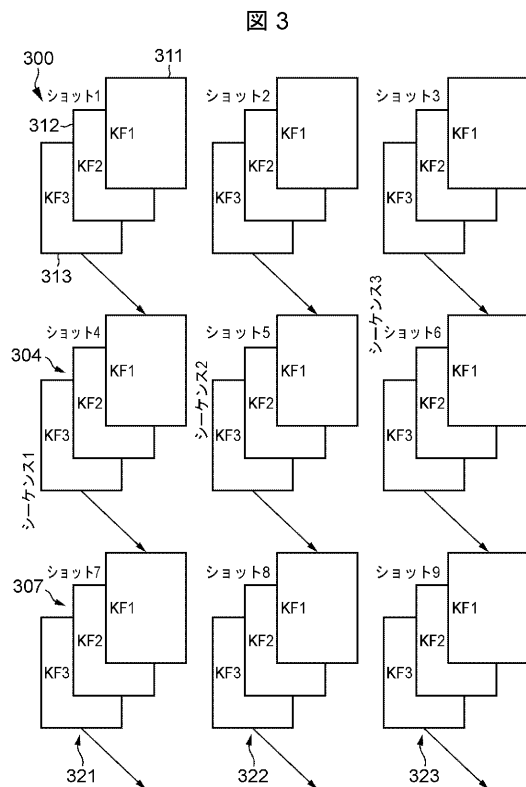
【 図 1 】



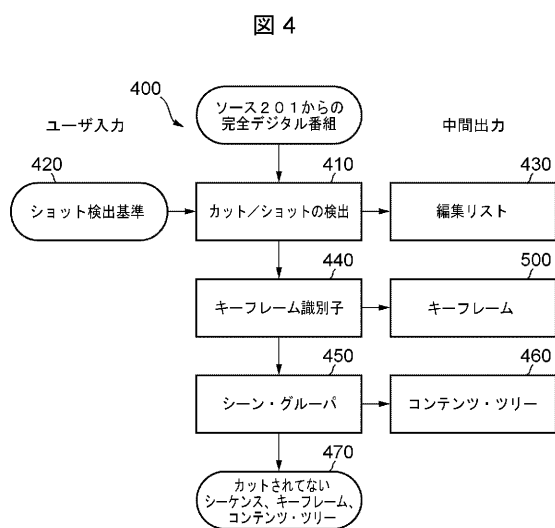
【 図 2 】



【圖 3】

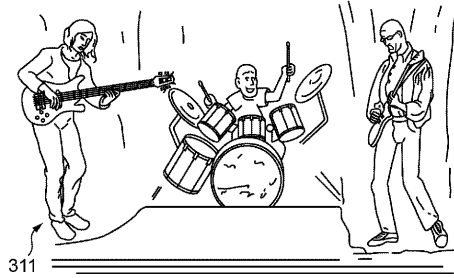


【圖 4】



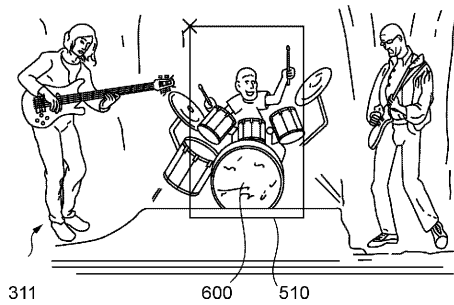
【図 5】

図 5



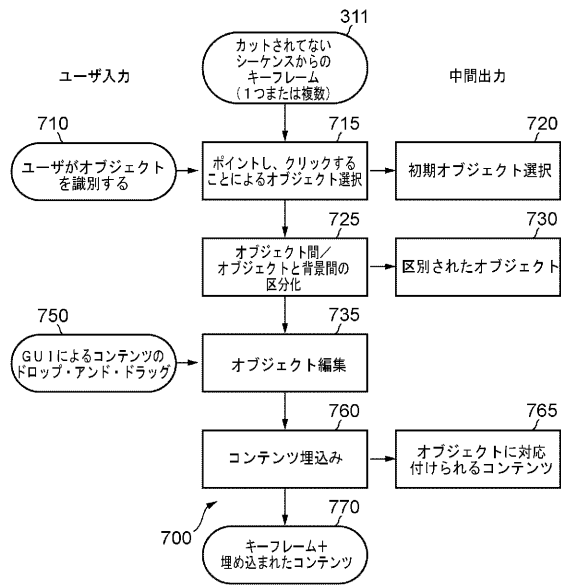
【図 6】

図 6



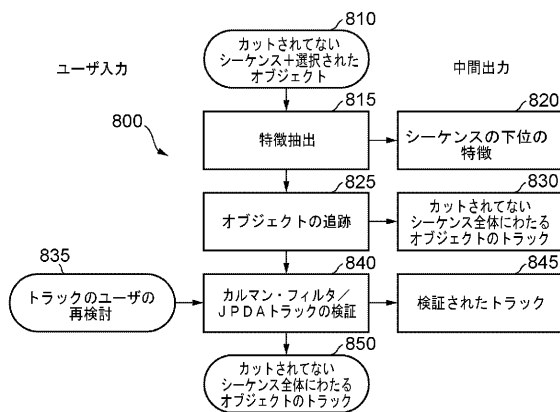
【図 7】

図 7



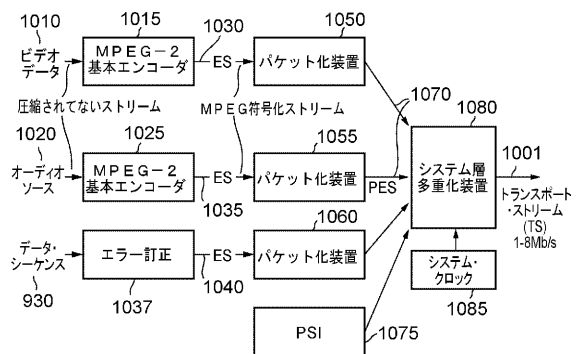
【図 8】

図 8



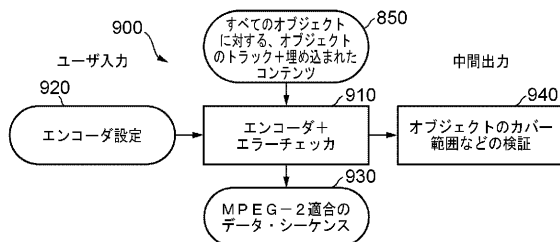
【図 10】

図 10



【図 9】

図 9



 フロントページの続き

(51)Int.Cl.		F I			
H 0 4 H 60/64	(2008.01)	H 0 4 H	1/00	6 5 9	
H 0 4 H 60/81	(2008.01)	H 0 4 H	1/00	6 7 6	

(56)参考文献 米国特許第05708845 (U S , A)
 美濃導彦, 知的メディア検索技術の動向, 人工知能学会誌, 1996年 1月 1日, 第11巻、第1号, P.3~9
 浅井光太郎ほか, デジタル放送を支える先端技術, 三菱電機技報, 第72巻 第8号, P.24~30(特にP.29)
 Eren,P.E.,et al, Interactive Object-based Analysis and Manipulation of Digital Video , Proc. of IEEE Workshop on Multimedia Signal Processing, 1998年12月 7日, P.335-340

(58)調査した分野(Int.Cl. , D B 名)
 H04N 7/16-173、24、26
 H04H 1/00
 G06F 17/30