

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2008-176538
(P2008-176538A)

(43) 公開日 平成20年7月31日(2008.7.31)

| (51) Int.Cl. | F I | テーマコード (参考) |
|------------------------------|-----------------|-------------|
| G06F 17/30 (2006.01) | G06F 17/30 210A | 5B075 |
| H04N 5/91 (2006.01) | H04N 5/91 N | 5C053 |
| G11B 27/034 (2006.01) | H04N 5/91 Z | 5D110 |
| H04N 5/93 (2006.01) | G11B 27/034 | |
| | G06F 17/30 220A | |

審査請求 未請求 請求項の数 14 O L (全 22 頁) 最終頁に続く

(21) 出願番号 特願2007-8909 (P2007-8909)
(22) 出願日 平成19年1月18日 (2007.1.18)

(71) 出願人 000003078
株式会社東芝
東京都港区芝浦一丁目1番1号
(74) 代理人 100089118
弁理士 酒井 宏明
(72) 発明者 山本 晃司
東京都港区芝浦一丁目1番1号 株式会社東芝内
Fターム(参考) 5B075 ND12 NK06 NS01
5C053 FA06 FA14 FA20 FA23 FA27
GB06 GB08 GB09 HA30 JA16
LA06 LA07 LA14
5D110 AA13 AA27 AA29 CA10 CB06
CD16 CD26 DA15

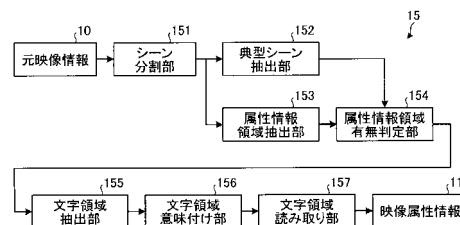
(54) 【発明の名称】 映像属性情報出力装置、映像要約装置、プログラムおよび映像属性情報出力方法

(57) 【要約】

【課題】映像コンテンツに関するメタデータを容易に生成することができる映像属性情報出力装置を提供する。

【解決手段】映像コンテンツ10をシーン分割部151で分割した隣接する各シーンにおける所定のフレーム間で、変化がない領域であって属性情報が表示される領域である属性情報領域を属性情報領域抽出部153で抽出し、属性情報領域内の個々の文字が存在する領域である文字領域を文字領域抽出部155で抽出するとともに当該文字領域の意味付けを文字領域意味付け部156で行い、文字領域読み取り部157で意味付けされた文字領域から映像コンテンツに関するメタデータである映像属性情報を読み取って出力する。これにより、例えば野球中継の場合における得点やボールカウントなどの映像コンテンツに関するメタデータである映像属性情報を容易に生成することができる。

【選択図】 図3



【特許請求の範囲】**【請求項 1】**

映像コンテンツのフレーム間の類似性が低くなるシーンチェンジを検出して、前記映像コンテンツを複数のシーンに分割するシーン分割手段と、

前記シーン分割手段で分割した隣接する前記各シーンにおける所定のフレーム間で、変化がない領域であって属性情報が表示される領域である属性情報領域を抽出する属性情報領域抽出手段と、

前記属性情報領域抽出手段により抽出された前記属性情報領域内の個々の文字が存在する領域である文字領域を抽出する文字領域抽出手段と、

前記文字領域抽出手段で抽出した前記文字領域について意味付けを記したルール情報を照合することにより当該文字領域の意味付けを行う文字領域意味付け手段と、

意味付けされた前記文字領域から前記映像コンテンツに関するメタデータである映像属性情報を読み取って出力する映像属性情報出力手段と、
を備えることを特徴とする映像属性情報出力装置。

10

【請求項 2】

前記シーン分割手段で分割した隣接する前記各シーンの中から前記映像コンテンツ中に何度も反復して出現する同じ構図で撮影した画像である典型シーンを抽出する典型シーン抽出手段と、

前記典型シーン抽出手段により抽出された前記典型シーンについて、前記属性情報領域抽出手段により抽出された前記属性情報領域が有るか否かを判定する領域判定手段と、
を更に備え、

20

前記文字領域抽出手段は、前記領域判定手段により前記典型シーンに有ると判定された前記属性情報領域抽出手段により抽出された前記属性情報領域内の個々の文字が存在する領域である文字領域を抽出する、

ことを特徴とする請求項 1 記載の映像属性情報出力装置。

【請求項 3】

前記属性情報領域抽出手段は、前記典型シーンとその直前のシーンの境界で前記属性情報領域を抽出する、

ことを特徴とする請求項 2 記載の映像属性情報出力装置。

【請求項 4】

前記典型シーン抽出手段は、

前記シーン分割手段で分割したシーン毎の特徴量に基づいて類似するシーンをグループに分類する分類手段と、

同じ構図で撮影した画像が前記グループに分類されたシーンの中に所定基準以上出現する場合に、当該グループに分類された全シーンを前記典型シーンとする判定手段と、
を備えることを特徴とする請求項 2 記載の映像属性情報出力装置。

30

【請求項 5】

前記判定手段の判定基準は、前記グループに含まれる同じ構図で撮影した画像の総数が定められた閾値を超えているか否かである、

ことを特徴とする請求項 4 記載の映像属性情報出力装置。

40

【請求項 6】

前記判定手段の判定基準は、前記グループに含まれる同じ構図で撮影した画像の総時間が前記映像コンテンツの時間に対して一定以上の割合を占めるか否かである、

ことを特徴とする請求項 4 記載の映像属性情報出力装置。

【請求項 7】

前記ルール情報は、意味付けの内容と、前記意味付けの内容に対応付けられており前記各文字領域を相対的な位置関係と配色の関係との少なくとも何れか一方によりグルーピングする第 1 のルールと、前記各文字領域の位置関係の第 2 のルールとを含み、

前記文字領域意味付け手段は、前記第 1 のルールグルーピングした前記文字領域から前記第 2 のルールを満たす前記意味付けの内容を選択し、前記文字領域の意味付けを行う、

50

ことを特徴とする請求項 1 記載の映像属性情報出力装置。

【請求項 8】

前記映像属性情報出力手段は、意味付けがなされた前記文字領域から読み取った前記映像属性情報に矛盾が生じた場合に認識結果を訂正する、
ことを特徴とする請求項 1 記載の映像属性情報出力装置。

【請求項 9】

前記映像属性情報出力手段から出力される前記映像属性情報について、前記映像属性情報の変化パターンとイベントの内容との対応関係を含むイベントルール情報と照合し、特定のパターンが現れたときにイベント情報として検出するイベント検出手段を更に備える、

10

ことを特徴とする請求項 1 ないし 8 の何れか一記載の映像属性情報出力装置。

【請求項 10】

請求項 1 ないし 9 の何れか一記載の映像属性情報出力装置と、
前記映像属性情報出力装置から出力された前記映像属性情報に含まれるイベントごとにその重要度を算出する重要度算出部と、

算出された重要度に従って、要約映像に含める映像コンテンツの映像区間を選択する映像区間選択部と、

前記イベント列から要約映像に用いる説明文の生成に用いる説明文用イベントを選択する説明文用イベント選択部と、

20

選択された前記説明文用イベントから説明文を生成する説明文作成部と、

選択された前記映像区間と前記説明文とを統合し、両方の情報を含んだ要約映像情報を生成する統合部と、

を備えることを特徴とする映像要約装置。

【請求項 11】

前記統合部は、前記説明文用イベント選択部で選択された前記説明文用イベントを含む前記要約映像情報を順次出力してスキップ再生する、
ことを特徴とする請求項 10 記載の映像要約装置。

【請求項 12】

前記統合部は、スキップ再生で変化する文字領域に対して強調表示を行う、
ことを特徴とする請求項 10 記載の映像要約装置。

30

【請求項 13】

映像コンテンツのフレーム間の類似性が低くなるシーンチェンジを検出して、前記映像コンテンツを複数のシーンに分割するシーン分割機能と、

前記シーン分割機能で分割した隣接する前記各シーンにおける所定のフレーム間で、変化がない領域であって属性情報が表示される領域である属性情報領域を抽出する属性情報領域抽出機能と、

前記属性情報領域抽出機能により抽出された前記属性情報領域内の個々の文字が存在する領域である文字領域を抽出する文字領域抽出機能と、

前記文字領域抽出機能で抽出した前記文字領域について意味付けを記したルール情報を照合することにより当該文字領域の意味付けを行う文字領域意味付け機能と、

40

意味付けされた前記文字領域から前記映像コンテンツに関するメタデータである映像属性情報を読み取って出力する映像属性情報出力機能と、
をコンピュータに実行させることを特徴とするプログラム。

【請求項 14】

映像コンテンツのフレーム間の類似性が低くなるシーンチェンジを検出して、前記映像コンテンツを複数のシーンに分割するシーン分割工程と、

前記シーン分割工程で分割した隣接する前記各シーンにおける所定のフレーム間で、変化がない領域であって属性情報が表示される領域である属性情報領域を抽出する属性情報領域抽出工程と、

前記属性情報領域抽出工程により抽出された前記属性情報領域内の個々の文字が存在す

50

る領域である文字領域を抽出する文字領域抽出工程と、

前記文字領域抽出工程で抽出した前記文字領域について意味付けを記したルール情報を照合することにより当該文字領域の意味付けを行う文字領域意味付け工程と、

意味付けされた前記文字領域から前記映像コンテンツに関するメタデータである映像属性情報を読み取って出力する映像属性情報出力工程と、

を含むことを特徴とする映像属性情報出力方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、映像属性情報出力装置、映像要約装置、プログラムおよび映像属性情報出力方法に関する。

10

【背景技術】

【0002】

近年、放送の多チャンネル化等の情報インフラの発展により、多くの映像コンテンツが流通するようになってきている。一方、録画装置の側においては、ハードディスク・レコーダやチューナー搭載のパーソナルコンピュータの普及により映像コンテンツをデジタルデータとして保存することが出来るようになっており、様々な視聴方法が可能となっている。

【0003】

録画装置を用いた視聴方法の1つとしては、長時間の映像を短時間に要約して視聴する映像要約技術がある。例えば、特許文献1には、スポーツ中継番組の要約手法として、メタデータが付与された映像からユーザの好みに基づいてシーンを抽出し、ダイジェストを作成する手法が記載されている。ここで、上述したメタデータは、試合中の個々のプレイなどをイベントとして時刻情報と共に記述したものであり、映像を見ながら人手で入力されるものである。

20

【0004】

また、特許文献2には、メタデータを入力するための補助となる手法が記載されている。具体的には、野球中継におけるインニングや得点などの情報の表示される画像領域（スコア情報領域）を指定してキー画像として登録しておき、インニングが変わったなどのイベントが発生した際にキー画像に対する変化率が一定の基準以上になったと判定した場合に、その時点の画像を表示するものである。

30

【0005】

【特許文献1】特開2005-109566号公報

【特許文献2】特開2000-132563号公報

【発明の開示】

【発明が解決しようとする課題】

【0006】

しかしながら、特許文献1に記載されている手法によれば、ダイジェスト生成に必要なメタデータを用いることで高精度な要約を生成できる半面、メタデータをすべて人手で入力することは入力の手間やコストが負担であるという課題がある。

【0007】

そこで、特許文献2に記載されている手法を用いるようにすれば、野球中継におけるメタデータの入力について、ある程度の負担の軽減は望めるようになると考えられる。

40

【0008】

ところが、特許文献2の手法によれば、インニングや得点などの情報が表示される画像領域の指定やイベントとして検出された画像の読み取りは人手によるため、番組ごとに画像領域を指定する手間や読み取ったインニングや得点の入力の手間やコストなどの課題は依然として解決されない。

【0009】

本発明は、上記に鑑みてなされたものであって、映像コンテンツに関するメタデータを容易に生成することができる映像属性情報出力装置、映像要約装置、プログラムおよび映

50

像属性情報出力方法を提供することを目的とする。

【課題を解決するための手段】

【0010】

上述した課題を解決し、目的を達成するために、本発明の映像属性情報出力装置は、映像コンテンツのフレーム間の類似性が低くなるシーンチェンジを検出して、前記映像コンテンツを複数のシーンに分割するシーン分割手段と、前記シーン分割手段で分割した隣接する前記各シーンにおける所定のフレーム間で、変化がない領域であって属性情報が表示される領域である属性情報領域を抽出する属性情報領域抽出手段と、前記属性情報領域抽出手段により抽出された前記属性情報領域内の個々の文字が存在する領域である文字領域を抽出する文字領域抽出手段と、前記文字領域抽出手段で抽出した前記文字領域について意味付けを記したルール情報を照合することにより当該文字領域の意味付けを行う文字領域意味付け手段と、意味付けされた前記文字領域から前記映像コンテンツに関するメタデータである映像属性情報を読み取って出力する映像属性情報出力手段と、を備える。

10

【0011】

また、本発明の映像要約装置は、請求項1ないし9の何れか一記載の映像属性情報出力装置と、前記映像属性情報出力装置から出力された前記映像属性情報に含まれるイベントごとにその重要度を算出する重要度算出部と、算出された重要度に従って、要約映像に含める映像コンテンツの映像区間を選択する映像区間選択部と、前記イベント列から要約映像に用いる説明文の生成に用いる説明文用イベントを選択する説明文用イベント選択部と、選択された前記説明文用イベントから説明文を生成する説明文作成部と、選択された前記映像区間と前記説明文とを統合し、両方の情報を含んだ要約映像情報を生成する統合部と、を備える。

20

【0012】

また、本発明のプログラムは、映像コンテンツのフレーム間の類似性が低くなるシーンチェンジを検出して、前記映像コンテンツを複数のシーンに分割するシーン分割機能と、前記シーン分割機能で分割した隣接する前記各シーンにおける所定のフレーム間で、変化がない領域であって属性情報が表示される領域である属性情報領域を抽出する属性情報領域抽出機能と、前記属性情報領域抽出機能により抽出された前記属性情報領域内の個々の文字が存在する領域である文字領域を抽出する文字領域抽出機能と、前記文字領域抽出機能で抽出した前記文字領域について意味付けを記したルール情報を照合することにより当該文字領域の意味付けを行う文字領域意味付け機能と、意味付けされた前記文字領域から前記映像コンテンツに関するメタデータである映像属性情報を読み取って出力する映像属性情報出力機能と、をコンピュータに実行させる。

30

【0013】

また、本発明の映像属性情報出力方法は、映像コンテンツのフレーム間の類似性が低くなるシーンチェンジを検出して、前記映像コンテンツを複数のシーンに分割するシーン分割工程と、前記シーン分割工程で分割した隣接する前記各シーンにおける所定のフレーム間で、変化がない領域であって属性情報が表示される領域である属性情報領域を抽出する属性情報領域抽出工程と、前記属性情報領域抽出工程により抽出された前記属性情報領域内の個々の文字が存在する領域である文字領域を抽出する文字領域抽出工程と、前記文字領域抽出工程で抽出した前記文字領域について意味付けを記したルール情報を照合することにより当該文字領域の意味付けを行う文字領域意味付け工程と、意味付けされた前記文字領域から前記映像コンテンツに関するメタデータである映像属性情報を読み取って出力する映像属性情報出力工程と、を含む。

40

【発明の効果】

【0014】

本発明によれば、例えば野球中継の場合における得点やボールカウントなどの映像コンテンツに関するメタデータである映像属性情報が表示される画像領域の指定などの人手を介した作業を省略し、映像コンテンツから映像属性情報を自動的に認識して取得することができるので、映像コンテンツに関するメタデータを容易に生成することができる、とい

50

う効果を奏する。

【発明を実施するための最良の形態】

【0015】

以下に添付図面を参照して、この発明にかかる映像属性情報出力装置、映像要約装置、プログラムおよび映像属性情報出力方法の最良な実施の形態を詳細に説明する。

【0016】

[第1の実施の形態]

本発明の第1の実施の形態を図1ないし図16に基づいて説明する。本実施の形態は映像要約装置としてパーソナルコンピュータを適用した例である。

【0017】

図1は、本発明の第1の実施の形態にかかる映像要約装置1の構成を示すブロック図である。映像要約装置1は、情報処理を行うCPU (Central Processing Unit) 101、BIOSなどを記憶した読出し専用メモリであるROM (Read Only Memory) 102、各種データを書換え可能に記憶するRAM (Random Access Memory) 103、各種データベースとして機能するとともに各種のプログラムを格納するHDD (Hard Disk Drive) 104、記憶媒体110を用いて情報を保管したり外部に情報を配布したり外部から情報を入手するためのDVDドライブ等の媒体駆動装置105、ネットワーク2を介して外部の他のコンピュータと通信により情報を伝達するための通信制御装置106、処理経過や結果等を操作者に表示するLCD (Liquid Crystal Display) 等の表示部107、並びに操作者がCPU 101に命令や情報等を入力するためのキーボードやマウス等の入力部108等から構成されており、これらの各部間で送受信されるデータをバスコントローラ109が調停して動作する。

【0018】

このような映像要約装置1では、ユーザが電源を投入するとCPU 101がROM 102内のローダーというプログラムを起動させ、HDD 104よりOS (Operating System) というコンピュータのハードウェアとソフトウェアとを管理するプログラムをRAM 103に読み込み、このOSを起動させる。このようなOSは、ユーザの操作に応じてプログラムを起動したり、情報を読み込んだり、保存を行ったりする。OSのうち代表的なものとしては、Windows (登録商標) 等が知られている。これらのOS上で走る動作プログラムをアプリケーションプログラムと呼んでいる。なお、アプリケーションプログラムは、所定のOS上で動作するものに限らず、後述の各種処理の一部の実行をOSに肩代わりさせるものであってもよいし、所定のアプリケーションソフトやOSなどを構成する一群のプログラムファイルの一部として含まれているものであってもよい。

【0019】

ここで、映像要約装置1は、アプリケーションプログラムとして、映像処理プログラムをHDD 104に記憶している。この意味で、HDD 104は、映像処理プログラムを記憶する記憶媒体として機能する。

【0020】

また、一般的には、映像要約装置1のHDD 104にインストールされるアプリケーションプログラムは、DVDなどの各種の光ディスク、各種光磁気ディスク、フレキシブルディスクなどの各種磁気ディスク、半導体メモリ等の各種方式のメディア等の記憶媒体110に記録され、この記憶媒体110に記録された動作プログラムがHDD 104にインストールされる。このため、DVD等の光情報記録メディアやFD等の磁気メディア等の可搬性を有する記憶媒体110も、アプリケーションプログラムを記憶する記憶媒体となり得る。さらには、アプリケーションプログラムは、例えば通信制御装置106を介して外部のネットワーク2から取り込まれ、HDD 104にインストールされても良い。

【0021】

映像要約装置1は、OS上で動作する映像処理プログラムが起動すると、この映像処理プログラムに従い、CPU 101が各種の演算処理を実行して各部を集中的に制御する。映像要約装置1のCPU 101が実行する各種の演算処理のうち、本実施の形態の特長的

10

20

30

40

50

な処理について以下に説明する。

【0022】

図2は、映像要約装置1の概略構成を示すブロック図である。図2に示すように、映像要約装置1は、映像処理プログラムに従うことにより、映像属性情報出力装置である映像属性情報作成部15、重要度算出部16、映像区間選択部17、説明文用イベント選択部18、説明文作成部19、統合部20を備える。また、10は元映像情報、11は映像属性情報、12は要約作成情報、13は説明文作成情報、14は要約映像情報である。

【0023】

元映像情報10は、例えばテレビの番組やユーザが撮影した映像など、様々な映像コンテンツがこれに該当する。本実施の形態では、元映像情報10は、デジタルデータとする。デジタルデータの形式は、どのような形式でもよい(例えば、MPEG-1やMPEG-2などの圧縮データ形式を用いるのが一般的である)。もちろん、ソース映像がアナログデータでもよいが、この場合には、予め外部でデジタルデータに変換するか、または、映像要約装置1がアナログ・デジタル変換を備えるものとする。また、映像コンテンツは、1つでも複数でもかまわなく、タイトルやIDなど映像コンテンツを一意に特定できる情報と、再生を開始する時間またはフレーム番号とを与えれば、任意の映像コンテンツを任意の位置から再生可能である。

【0024】

映像属性情報11は、映像コンテンツに関する様々な属性情報(メタデータ)である。映像属性情報11には、映像コンテンツ内で起きた事象をイベントとして列挙したイベント列が含まれる。イベントは、人物や物体の名称や動き(例えば、野球の場合、「選手のホームラン」など)や、時間的な区切り(「試合開始」、「試合終了」など)などの情報とその発生時間と共に記述される。発生時刻の記述は、シーン情報のように区間の始点・終点を記述する形式でもよいし、イベントが発生した瞬間の時刻を記述する形式でもよい。その他、属性情報には、スポーツの種目、対戦チーム名、日時、場所、出場選手、試合結果など、映像コンテンツ中の特定の時間に関連しない情報も含んでよい。また、ここでは具体例としてスポーツを挙げたが、属性情報はコンテンツに応じて適宜記述することができる。例えば、ドラマや情報番組についても登場人物やトピックのタイトルを用いることで、同様に記述することができる。

【0025】

要約作成情報12は、要約を作成するときのユーザの嗜好や、要約映像の時間、要約アルゴリズムのパラメータなどを含む情報である。ユーザの嗜好とは、ユーザが優先的に視聴したい部分を求めるためのキーワードなどの情報である。スポーツの試合であれば、好みのチームや好みの選手などの情報がこれに該当する。キーワードは、見たい対象だけ記述してもよいし、見たい対象だけではなく、見たくない対象についての記述をも可能にしてもよい。これらの情報は、直接入力してもよいし、毎回同じような条件を入力する手間を省くために予め内部あるいは外部の記憶装置上に保持しておいてもよい。

【0026】

説明文作成情報13は、映像属性情報11に含まれるイベントから、説明文を作成するとき用いるテンプレート(雛形)を含む情報である。ここで用いるテンプレートとは、例えば、「(時間)、(チーム)のチャンス。(選手)の(プレイ)！」のように、固定された文字列と可変の文字列の組み合わせによって与えられる文字列である。この例では、「(時間)」など()によって示された文字列が可変の文字列を表している。

【0027】

要約映像情報14は、映像要約装置1によって生成される要約映像である。要約映像情報14は、元映像コンテンツの一部と、生成された文字等の可視的情報や音声等の可聴的情報による説明文(例えばテロップやナレーションなど)を含む。また、元映像コンテンツとは別に用意されたコンテンツ(例えば、タイトル画面など)を含めてもかまわない。また、要約映像情報14の形態は、元映像とは別の映像コンテンツとして、それ単体で再生できる情報であってもよいし、元映像コンテンツを参照して再生しながら、文字や音声

10

20

30

40

50

による説明文を表示・再生するものであってもよい。後者には、SMILなど複数のメディアを同期して表示することのできる記述言語を用いることができる。音声の再生には、例えば、音声合成を用いることができる。

【0028】

以下において、映像要約装置1の各部について説明する。

【0029】

映像属性情報作成部15は、元映像情報10に基づいて映像属性情報11を自動的に作成する。なお、映像属性情報作成部15における映像属性情報11の自動作成処理の詳細は後述する。

【0030】

重要度算出部16は、入力された映像属性情報11および要約作成情報12に基づいて、映像属性情報11に含まれるイベントごとにその重要度を算出する。イベントの重要度の算出方法は、当該イベントに含まれる文字列と要約作成情報12に含まれるキーワードとの一致の具合の評価値を求めることによつて行う。例えば、重要度を w とし、キーワードの総数を N として、次に例示する式により重要度 w を求めることができる。ただし、次式において a_k は、要約作成情報12に含まれる k 番目のキーワードの重み値を示すパラメータである。

$$w = \sum_{k=1}^N a_k M(x)$$

ここで、 $\sum_{k=1}^N$ は $k=1 \sim N$ の範囲を総和の対象とする。また、 x 番目のキーワードが一致する場合には、 $M(x) = 1$ とし、 x 番目のキーワードが一致しない場合には、 $M(x) = 0$ とする。

【0031】

なお、キーワードとして見たくない対象についての記述をも可能にする構成においては、 $M(x)$ について、 x 番目のキーワードが一致し、 x 番目のキーワードが見たい映像に係るキーワードである場合には、 $M(x) = 1$ とし、 x 番目のキーワードが一致し、 x 番目のキーワードが見たくない映像に係るキーワードである場合には、 $M(x) = -1$ とし、 x 番目のキーワードが一致しない場合には、 $M(x) = 0$ とする。

【0032】

映像区間選択部17は、算出された重要度に従つて、要約映像に含める元映像コンテンツの区間を選択する。

【0033】

説明文用イベント選択部18は、要約映像に用いる文字や音声による説明文(テロップやナレーションなど)の生成に用いるイベントを、イベント列から選択する。選択は予め決められたイベントを自動的に選択してもよいし、ユーザが明示的に選択してもよい。例えば、野球中継の場合には、ヒットなどの安打のイベントを選択する。基本的には、要約映像の全部に対してのみ説明文を生成する処理と、要約映像の一部に対してのみ説明文を生成する処理と、要約映像の全部又は一部及び要約映像以外の所定の部分に対して説明文を生成する処理と、要約映像以外の所定の部分に対してのみ説明文を生成する処理が考えられる。これらのイベント選択処理は、映像区間選択部17とは独立して処理してもよいし、映像区間選択部17のデータを利用してよい。

【0034】

説明文作成部19は、説明文用イベント選択部18で選択された説明文用イベントと説明文作成情報13とから、(文字等の可視的情報や音声等の可聴的情報による)説明文を生成する。前述したように、説明文作成情報13は、説明文を生成するためのテンプレートを含む情報である。テンプレートとは、例えば、「(時間)、(チーム)のチャンス。(選手)の(プレイ)！」のように、固定された文字列と可変の文字列の組み合わせによつて、与えられる文字列である。この例では、「(時間)」など()によつて示された文字列が可変の文字列を表している。可変の文字列には、映像属性情報11や要約作成情報12から、直接的あるいは間接的に得ることができる文字列が入る。このようなテンプレートがイベントの種類などによつて複数用意される。

10

20

30

40

50

【0035】

統合部20は、映像区間情報と文字テロップや音声ナレーションによる説明文とを統合し、両方の情報を含んだ要約映像情報14を生成する。文字の表示や音声の発話のタイミングは、イベントの発生時刻や映像区間に合わせて行う。発生時刻に合わせる場合、発生時刻の前に適当な間隔をおいて表示、発話したり、発話区間の特定の位置（例えば中間点）がイベントの発生時刻となるように表示、発話したりするようにタイミングを調整すればよい。また、映像区間に合わせる場合は、その区間の開始点や終了点から適当な間隔をとって表示、発話すればよい。もちろん、これらの複数の手法を組み合わせ、例えば、文字テロップは映像区間に合わせ、音声ナレーションはイベントの発生時刻に合わせて用いるなどしてもよい。

10

【0036】

統合時には、テロップのオーバーレイ表示や、音声合成による音声ナレーションを作成して、要約映像に含めてもよいし、テキストの状態でも要約映像情報14として記述して、再生装置側で、オーバーレイ表示音声合成を行ってもよい。

【0037】

なお、上記では、統合部20は、選択された映像区間情報と生成された説明文情報とを単純に重ね合わせて、1つの要約映像情報14として出力する例を示したが、映像区間や音量の調整を含むようにしてもよい。

【0038】

以上のように本実施の形態の映像要約装置1によれば、映像情報の中から、ユーザの入力したキーワードや時間などの条件やユーザの好みに基づいて、ユーザの興味のある映像区間のみを選択し、それにテロップや音声ナレーションを付加した要約映像を作成することができる。また、省略されてしまった重要なシーンを説明文で補うことで、より内容を理解しやすい要約映像を作成することができる。

20

【0039】

したがって、本実施の形態の映像要約装置1によれば、ユーザは1つの番組をまるごと再生することなく、自分の興味のある部分だけを含む要約を作成して、再生することが可能である。その際に、テロップや音声ナレーションによる説明を映像に現れない部分を含めて、ユーザの嗜好に合った文章や口調で受けることができる。

【0040】

次に、映像属性情報作成部15における映像属性情報11の自動作成処理について詳述する。ここで、図3は映像属性情報作成部15の概略構成を示すブロック図、図4は映像属性情報の自動作成処理の流れを概略的に示すフローチャートである。

30

【0041】

図3に示すように、映像属性情報作成部15は、元映像情報10を入力とし、映像属性情報11を出力する。このような映像属性情報作成部15は、元映像情報10をシーンに分割するシーン分割手段として機能するシーン分割部151と、同じ構図で撮影した画像（典型シーン）を抽出する典型シーン抽出手段として機能する典型シーン抽出部152と、属性情報領域を抽出する属性情報領域抽出手段として機能する属性情報領域抽出部153と、典型シーンに属性情報領域が含まれるか否かを判定する領域判定手段として機能する属性情報領域有無判定部154と、属性情報領域が含まれる文字領域を抽出する文字領域抽出手段として機能する文字領域抽出部155と、文字領域の意味付けを行う文字領域意味付け手段として機能する文字領域意味付け部156と、文字領域から映像属性情報を読み取って出力する映像属性情報出力手段として機能する文字領域読み取り部157と、を備えている。

40

【0042】

映像属性情報の自動作成処理の流れを概略的に説明する。各ステップの詳細については、後述する。図4に示すように、まず、元映像情報10のフレームごとの特徴量を算出し、特徴量が類似する連続フレームの集合である複数のシーンに元映像情報10を分割する（ステップS1：シーン分割部151）。すなわち、元映像情報10のフレーム間の類似

50

性が低くなるシーンチェンジを検出して、元映像情報 10 を複数のシーンに分割する。

【0043】

次いで、各シーンの中から典型シーンを抽出し（ステップ S 2：典型シーン抽出部 152）、画面の中から属性情報が表示される属性情報領域を抽出し（ステップ S 3：属性情報領域抽出部 153）、ステップ S 2 で抽出された各典型シーンについて、属性情報が表示されているか否かを判定する（ステップ S 4：属性情報領域有無判定部 154）。

【0044】

その後、属性情報が表示された典型シーンから文字領域を抽出し（ステップ S 5：文字領域抽出部 155）、抽出された文字領域とルール情報を照合して文字領域の意味付けを行う（ステップ S 6：文字領域意味付け部 156）。

【0045】

最後に、文字領域の読み取りを行って映像属性情報を生成し（ステップ S 7：文字領域読み取り部 157）、処理を終了する。

【0046】

ここで、上述の各用語について図面を参照しつつ説明する。

【0047】

図 5 は、野球中継番組の画面例を示す模式図である。図 5 に示すように、201 に示す領域には“回”、“得点”、“ボールカウント”、“出塁状況”などが表示されている。これらのスコア情報を属性情報、201 の領域を属性情報領域と呼ぶ。さらに、属性情報領域内の個々の文字が存在する領域（例えば、202 に示すボールカウントが存在するスコア領域）を文字領域と呼ぶ。

【0048】

図 6 は、典型シーンとなる同じ構図で撮影した画像について示す説明図である。301 は元映像情報 10 を示し、紙面を右に向かって時間が流れているとする。斜線部（例えば、302）は投球シーンを示す。投球シーンは、図 5 に示したようにピッチャーの後方からバッター方向を撮影したシーンである。野球中継では、図 5 に示すような画面に投球の度に切り替わることが多く、カメラの位置や向きは毎回、ほぼ同じである。そのため、303 に示すように、野球中継番組では投球シーンが繰り返し出現する。このように、映像コンテンツ中に何度も反復して出現する同じ構図で撮影した画像を典型シーンと定義する。

【0049】

次に、各ステップにおける処理の詳細について説明する。

【0050】

まず、シーン分割部 151 によるシーン分割（ステップ S 1）について説明する。シーン分割部 151 によるシーン分割は、元映像情報 10 の映像フレームの特徴量を比較することで行う。ここで、図 7 は入力映像からの特徴量抽出方法について模式的に示す説明図である。図 7 に示す 401 は、元映像情報 10 の映像フレームを順次並べたものである。これらの映像フレームから特徴量を直接抽出することもできるが、本実施の形態においては、時間的・空間的にサンプリングを行うことで処理量を削減しつつ、特徴量を抽出する方法について述べる。

【0051】

時間的なサンプリングは、図 7 に示すように、映像フレーム群 401 から映像の一部のフレーム 402 を抜き出して処理する。抜き出す方法としては、一定間隔で抜き出したり、MPEG の I ピクチャだけ抜き出す方法などがある。403 は抜き出した 1 フレームである。次いで、抜き出したフレーム 403 を縮小してサムネイル画像 404 を作成し、空間的にサンプリングする。サムネイル画像 404 の作成方法としては、複数の画素の平均値を求めて縮小する方法や MPEG の I ピクチャの DCT 係数の DC 成分だけを復号して算出する方法などがある。そして、サムネイル画像 404 をブロックに分割し、ブロックごとに色のヒストグラム分布 405 を求める。この色ヒストグラム分布をこのフレームの特徴量とする。

10

20

30

40

50

【 0 0 5 2 】

そして、隣接するフレーム 4 0 3 の間の特徴量の距離を求め、特徴量の距離が一定の基準を上回るフレーム 4 0 3 の間、すなわち類似度が低下するフレーム 4 0 3 の間でシーンを分割する。特徴量の距離には、例えばユークリッド距離を用いることができる。フレーム i の a 番目ブロックのヒストグラムにおける b 番目の頻度を $h_i(a, b)$ とすると、フレーム i とフレーム $i + 1$ 間のユークリッド距離 d は、下記式のようになる。

【数 1】

$$d^2 = \sum_a \sum_b (h_i(a, b) - h_{i+1}(a, b))^2$$

10

【 0 0 5 3 】

次に、典型シーン抽出部 1 5 2 による典型シーンの抽出（ステップ S 2）について説明する。ここで、図 8 は典型シーンの抽出処理の流れを示すフローチャートである。図 8 に示すように、まず初めに、シーン分割部 1 5 1 で分割したシーン毎の特徴量に基づいて類似するシーンをグループに分類する（ステップ S 1 1：分類手段）。すなわち、1つのグループには、類似するシーンのみが含まれることになる。シーンの特徴量としては、例えば、各シーンの先頭のフレームの特徴量を用いることができる。

【 0 0 5 4 】

続くステップ S 1 2 では、グループ i ($i = 1, 2, 3, \dots, N$ (初期値 $i = 1$)) に分類されたシーンの中に基準以上の「反復があるシーン」が含まれるか否かを判定する。N は、グループの総数である。「反復があるシーン」は、映像の時系列中に何度も反復して出現する同じ構図で撮影した画像である。判定の基準としては、例えば、グループに含まれる「反復があるシーン」の総数が定められた閾値を超えているか否かであったり、グループに含まれる「反復があるシーン」の総時間が元映像の時間に対して一定以上の割合を占めるか否かである。

20

【 0 0 5 5 】

基準以上の「反復があるシーン」が含まれる、すなわち基準を満たすと判定された場合は（ステップ S 1 2 の Yes）、ステップ S 1 3 へ進み、グループ i に分類された全シーンを典型シーンとする。ここに、判定手段の機能が実行される。その後、ステップ S 1 4 へ進む。

30

【 0 0 5 6 】

一方、基準以上の「反復があるシーン」が含まれない場合は（ステップ S 1 2 の No）、そのままステップ S 1 4 に進む。

【 0 0 5 7 】

ステップ S 1 4 では全グループについての処理が終了したか否か、すなわち $i = N$ が否かを判定する。全グループについての処理が終了していなければ（ステップ S 1 4 の No）、 i を更新し（ステップ S 1 5）、ステップ S 1 2 へ戻る。全グループについての処理が終了していれば（ステップ S 1 4 の Yes）、選択した典型シーンを時間順に並び替え（ステップ S 1 6）、処理を終了する。

40

【 0 0 5 8 】

次に、属性情報領域抽出部 1 5 3 による属性情報領域の抽出（ステップ S 3）について説明する。ここで、図 9 は属性情報領域の抽出処理の流れを示すフローチャートである。図 9 に示すように、まず初めに、シーン分割部 1 5 1 で分割した各シーンの境界について、境界の前と境界の後のフレームを選択し、フレーム間の重複部分を抽出する（ステップ S 2 1）。なお、選択するフレームはシーン境界の直前、直後である必要はなく、数フレームあるいは数秒離れた位置のフレームを選択してもよい。また、映像中で処理対象とするシーン境界は映像中の全てのシーン境界としてもよいが、典型シーンとその直前のシーンの境界だけを対象としてもよい。これは、典型シーン（例えば、ピッチングシーン）は 1 つのプレイの開始であるため、属性情報（スコア情報）が表示されることが多く、属性情報領域の検出に適しているからである。

50

【 0 0 5 9 】

図 1 0 は、典型シーンをを用いてフレーム間の重複部分を抽出する処理について模式的に示す説明図である。図 1 0 中、5 0 1 は典型シーンの中の 1 シーンの先頭付近のフレーム、5 0 2 はその直前のシーンの終端付近のフレームを示す。すなわち、ステップ S 2 1 においては、典型シーンの中の 1 シーンの先頭付近のフレーム 5 0 1 と、その直前のシーンの終端付近のフレーム 5 0 2 との重複部分、すなわち変化がない領域が抽出される。図 1 0 においては、斜線で示す重複部分 5 0 3 が、変化がない領域として抽出される。

【 0 0 6 0 】

次いで、抽出された重複部分中において閾値を下回る画素を 2 値化によって求める（ステップ S 2 2）。

【 0 0 6 1 】

そして、抽出した全てのシーン境界の重複部分を合算し（ステップ S 2 3）、合算した重複部分中（図 1 0 中の 5 0 4）において閾値を下回る画素を 2 値化によって求める（ステップ S 2 4）。

【 0 0 6 2 】

最後に、これらの画素の分布密度が高い部分を属性情報領域として抽出する（ステップ S 2 5）。ここで、図 1 1 は属性情報領域の抽出方法を模式的に示す説明図である。図 1 1 に示すように、画面 6 0 1 の中で属性情報領域の候補を 6 0 2 とし、この領域 6 0 2 の位置を求める。画面 6 0 1 を x 軸方向、y 軸方向に射影すると、6 0 3、6 0 4 のような分布をもつヒストグラムが生成される。これらのヒストグラム 6 0 3、6 0 4 において閾値以上の頻度を持つ区間 6 0 5、6 0 6 を求め、その領域 6 0 2 を属性情報領域 2 0 1 の位置とする。

【 0 0 6 3 】

次に、属性情報領域有無判定部 1 5 4 による属性情報領域の有無の判定（ステップ S 4）について説明する。ステップ S 2 で抽出された各典型シーンに属性情報（スコア情報）が表示されているか否かの判定は、ステップ S 3 で合算された重複部分 5 0 4 と、各典型シーンとその直前のシーンとの重複部分（例えば、5 0 3）を比較することにより行なう。すなわち、重複部分 5 0 4 と、各典型シーンとその直前のシーンとの重複部分（例えば、5 0 3）を比較し、基準以上の画素が一致する場合に、その典型シーンには属性情報（スコア情報）が表示されていると判定する。

【 0 0 6 4 】

次に、文字領域抽出部 1 5 5 による文字領域抽出（ステップ S 5）について説明する。ここで、図 1 2 は文字領域の抽出処理の流れを示すフローチャート、図 1 3 は文字領域を抽出する処理について模式的に示す説明図である。図 1 2 に示すように、まず初めに、隣接する 2 つの典型シーン間の属性情報領域 2 0 1 における差分を抽出する（ステップ S 3 1）。例えば、図 1 3 においては、典型シーン 7 0 1 と典型シーン 7 0 2 の属性情報領域 2 0 1 における差分 7 0 3 を求める。

【 0 0 6 5 】

次に、閾値に基づいて差分を 2 値化する（ステップ S 3 2）。このようにすることで、2 つの典型シーンの属性情報領域 2 0 1 間で、変化のあった箇所だけが抽出される。野球の投球では 1 プレイごとに典型シーンが繰り返されるため、1 球ごとの変化が抽出される。この差分を他の隣接典型シーンの組み合わせについても求める。例えば、図 1 3 においては、典型シーン 7 0 2 の属性情報領域 2 0 1 と典型シーン 7 0 4 の属性情報領域 2 0 1 の差分 7 0 5 を求める。

【 0 0 6 6 】

続いて、これらの属性情報領域 2 0 1 内の差分を領域に分割する（ステップ S 3 3）。差分は 2 値化されているので、隣接画素が同じ値であれば 1 つの領域とする領域成長法によって領域に分割することができる。これらの領域が文字領域となる。

【 0 0 6 7 】

最後に、全ての典型シーンについて求めた文字領域から、映像中で変化のあった領域の

10

20

30

40

50

一覧706を作成する(ステップS34)。なお、一覧には重複する領域は一度だけ含まれるようにする。

【0068】

ここで、図14は重複する領域の処理例を模式的に示す説明図である。図14は、重複する領域の処理についてストライクカウントを例に説明したものである。なお、図14に示す破線部は説明のためのもので、実際には何もない。図14に示す801は、打者が交代することにより、2ストライクからノーストライクに変化したときの差分を示す領域である。図14に示す802は、ノーストライクから1ストライクに変化したときの差分を示す領域である。同様に、図14に示す803は、1ストライクから2ストライクに変化したときの差分を示す領域である。このとき、801の領域は802の領域と803の領域を合わせたものとなり、それぞれと重複する。この場合、801の領域は領域一覧には含めず、重複しない802と803の領域を含める。また、映像中に802や803と同等の領域は何度も出現するが、1回だけ採用する。

10

【0069】

次に、文字領域意味付け部156による文字領域の意味付け(ステップS6)について説明する。文字領域意味付け部156は、ステップS5で抽出された文字領域一覧の個々の領域の意味付けを行う。

【0070】

例えば、野球の場合、ボールカウント(ストライク, ボール, アウト)や得点を表すスコア文字領域(文字領域)に意味を割り当てる。意味付けにはルール情報を用い、スコア文字領域(文字領域)の特徴と照合して適合する内容を割り当てる。ルール情報は、スコア文字領域(文字領域)のグループ、他のスコア文字領域(文字領域)との相対的な位置関係やスコア文字領域(文字領域)内での大まかな位置に関する内容を含む。グループとは複数のスコア文字領域で1つの意味合いを持つものを指し、グループを構成するスコア文字領域(文字領域)の数、配色、位置関係をルール情報とする。例えば、ストライクカウントは2つのスコア文字領域(文字領域)によって構成されるので、これらの2つの領域を1グループとして扱う。そのため、意味付け処理の手順は初めに位置や色で領域をグルーピングし、次に、他のグループとの相対的な関係を比較するという2段階の処理で行う。

20

【0071】

図15は、ルール情報の一例を示す模式図である。図15に示すように、ルール情報は、それぞれの内容901と、対応するグループとしてのルール902と、位置関係のルール903とにより構成される。例えば、「ストライク」カウントに対応して、「水平方向に隣接する2つの同サイズ同色の領域である」というグループのルール902が定義されている。そこで、各領域を位置、色によってグルーピングする。なお、隣接と近接の違いは、領域間の距離の閾値の違いであり、隣接の方がより近い位置にある。次に、「ストライク」カウントに対応して、「ボールおよびアウトカウントと領域の左端のx座標が同じ(縦並びの場合)あるいはy座標が同じ(横並びの場合)」という他の要素との相対的な位置関係についてのルール903が定義されている。そこで、このような位置関係のルール903を用いて照合を行い、ルール902に基づいてグルーピングした領域からルール903を満たすものを選択し、意味付けを完了する。

30

40

【0072】

例えば野球中継においては、このように抽象度の高いルールを用いることで、放送局ごとのデザインの違いに影響を受けづらくなるので、汎用性が高まる。なお、ルール情報は1つである必要はなく複数のパターンを用意しておいてもよい。例えば、ボールカウントが数字で表示されるパターンを用意する。ルールで数字と定義されているものは属性情報領域201が表示された典型シーンを1つ以上選択し、数字として読み取ってみて、読み取ることができれば数字とする。スポーツ中継に用いられるスコア情報はスポーツごとにある程度決まった形があり、放送局が変わっても大きくは変わらないため、スポーツごとに数パターンを作成しておけば十分である。もちろん、汎用性が不要がない場合はこのよ

50

うな抽象的なルール情報ではなく、特定の番組のスコア情報の配置をテンプレート（雛形）とし、スコア文字領域（文字領域）の位置や大きさを細かく指定しても構わない。

【0073】

次に、文字領域読み取り部157による映像属性情報の生成（ステップS7）について説明する。文字領域読み取り部157は、ステップS6で意味づけられた各文字領域から映像コンテンツに関するメタデータである映像属性情報11を読み取って出力する。ここで、図16は野球のスコア情報の一例を示す模式図である。図16に示す野球のスコア情報の例では、回、得点、ボールカウント、出塁状況を含んでいるが、一部の情報（例えば、ボールカウントだけを含む情報）を生成してもよい。その場合、当然、ステップS6における処理も該当する情報だけを意味づけることができればよい。

10

【0074】

なお、スコア文字領域（文字領域）からの映像属性情報11の読み取りは、図形なのか文字なのかで処理が異なる。例えば、図形は、丸の数で表現されるボールカウントや出塁状況に用いられる。この場合は、各スコア文字情報が表示状態なのか、非表示状態なのかを判定する。一般的に表示状態は輝度や彩度が高い図形で表示されるので、閾値処理によって判定することができる。あるいは複数の典型シーンのスコア文字領域（文字領域）を2つのクラス（グループ）に分類し、輝度や彩度が高い方を選択する。一方、文字の読み取りの場合、野球中継で用いられる文字は数字と「表、裏」の文字であり、種類が限られている。そこで、これらの文字のテンプレートを予め作成しておき、比較することで読み取ることができる。もちろん、一般のOCRを用いたより汎用性が高い文字認識技術を用いても構わない。スコア文字領域（文字領域）から読み取ったスコア文字情報（映像属性情報11）はスポーツのルールと照らし合わせて矛盾がないように訂正される。例えば、回などに変化がないにも係わらず得点の数字が減少したときは誤認識である可能性が高いので、増加あるいは変化しない数字で読み取りの信頼性が高いものがあればそちらの数字を採用する。また、図形の読み取りに関しても、ボールカウントの1つ目が点灯せずに2つめ以降が点灯するなどの通常ありえない認識結果の場合には、前後の読み取り結果を考慮して訂正する。

20

【0075】

このように本実施の形態によれば、映像コンテンツを分割した隣接する各シーンにおける所定のフレーム間で、変化がない領域であって属性情報が表示される領域である属性情報領域を抽出し、属性情報領域内の個々の文字が存在する領域である文字領域を抽出するとともに当該文字領域の意味付けを行い、意味付けされた文字領域から映像コンテンツに関するメタデータである映像属性情報を読み取って出力する。これにより、例えば野球中継の場合における得点やボールカウントなどの映像コンテンツに関するメタデータである映像属性情報が表示される画像領域の指定などの人手を介した作業を省略し、映像コンテンツから映像属性情報を自動的に認識して取得することができるので、映像コンテンツに関するメタデータを容易に生成することができる。

30

【0076】

また、野球を初めとするスポーツ中継番組では投球時に投手の背後から打者方向を撮影したシーンのように、同じ構図で撮影した画像（典型的なシーン）が繰り返し現れる。このように番組に典型的なシーンはスポーツの1つのプレイごとに繰り返されることが多く、その度に得点やボールカウントなどのスコア情報に変化が生じる。そこで、本実施の形態では、対象となる映像コンテンツにおいて、野球の投球のように試合中の繰り返しの単位となる典型的なシーンを自動的に検出し、典型シーンの間での画面の変化によって映像属性情報を検出・認識することで、得点やボールカウントなどのメタデータを自動生成することにより、映像コンテンツに関するメタデータの生成を更に容易にすることができる。

40

【0077】

[第2の実施の形態]

次に、本発明の第2の実施の形態を図17および図18に基づいて説明する。なお、前

50

述した第1の実施の形態と同じ部分は同じ符号で示し説明も省略する。

【0078】

図17は、本発明の第2の実施の形態の映像属性情報作成部15の概略構成を示すブロック図である。図17に示すように、本実施の形態の映像属性情報作成部15は、第1の実施の形態の映像属性情報作成部15の構成にイベント検出部160を追加し、イベント検出部160で検出したイベント情報30を説明文用イベント選択部18(図2参照)に出力するようにしたものである。

【0079】

イベントとは試合の中で発生した何らかの事象である。例えば、野球ではヒット、ホームラン、打者の交代などが挙げられる。これらのイベントが発生すると、スコア情報(映像属性情報11)に特定の変化が起こる。そこで、イベント検出部160は、スコア情報(映像属性情報11)の変化パターンを予め用意されたイベントルール情報と照合し、特定のパターンが現れたときにイベント情報30として検出する。イベントルール情報としては、例えば、「ヒットの場合、ストライク、ボールのカウントが0に戻る」「アウトで打者が交代するときは、ストライク、ボールのカウントが0に戻り、アウトのカウントが1つ増える」「ホームランの場合は、出塁なしの状態になり、得点が増える」などが考えられる。このようなイベントルール情報を検出したいイベントの種類ごとに用意しておき、映像属性情報11と照合してイベント情報30を出力する。

10

【0080】

ここで、図16に示したスコア情報(映像属性情報11)を例に説明する。図16に示す1001から1002のスコア情報に変化したときに、ストライク、ボールのカウントがクリアされ、アウトカウントには変化がないため、1001の投球の結果がヒットであったことが分かる。また、1003から1004への変化では、ストライク、ボールのカウント、出塁がクリアされ、得点が増えているので、ホームランと判断できる。

20

【0081】

したがって、統合部20は、説明文用イベント選択部18で選択されたイベント情報30が含まれる典型シーンを順次再生することが可能になる。より詳細には、統合部20は、典型シーン抽出部152で得られた典型シーンの時刻を受け取り、スキップの指示があったときに典型シーンの単位で再生中の映像をスキップする。なお、次のイベントを含む典型シーンへスキップするタイミングは、前の典型シーンの再生後に自動的にスキップしてもよいし、外部からの指示があったときにスキップしてもよい。スキップした後、再生を再開するときに、説明文作成部19はその典型シーンに対応したスコア情報(映像属性情報11)に基づいて説明文を生成する。前述したように、説明文の生成はスコア情報を文章のテンプレート(説明文作成情報13)に当てはめることで行う。テンプレートは、スコア情報(映像属性情報11)の部分を除いた文章が決まっており、空白の部分にスコア情報(映像属性情報11)を当てはめる。テンプレートは1種類である必要はなく、スコア情報(映像属性情報11)の内容や前後の説明文、試合展開、ユーザの好みなどに応じて切り替えてもよい。例えば、図14の1005のスコア情報を「(回)回の(表裏)、(得点)の場面で、カウントは(ストライク)ストライク、(ボール)ボール」というテンプレートに当てはめると、「5回の裏、2対3の場面で、カウントは1ストライク、1ボール」となる。このようにして生成された説明文を音声合成を用いて音声等の可聴的情報に変換し、再生される映像に同期して再生する。このような構成とすることで、ユーザは安打のシーンや打者に対する最終投球だけを飛ばして視聴することができるので、短時間で視聴することができる。また、視聴を再開した場面での試合状況を音声で教えてくれるため、試合の流れも理解しやすくなる。

30

40

【0082】

また、上述したようなスキップの際に、スコア文字領域(文字領域)の変化部分を分かり易くするように、強調表示を行うようにしても良い。図18に示すように、1401のスコア文字領域(文字領域)を含む典型シーンから1402のスコア文字領域(文字領域)を含む典型シーンにスキップで変化した際に、統合部20は、変化したスコア文字領域

50

(文字領域)が分かり易くなるように1403のように変化部分を囲むなどして強調表示を行う。文字領域抽出部155で隣接する2つの典型シーン間の属性情報領域201における差分が得られているので、その部分の領域を強調表示する。強調表示方法は何でもよく、1403のように囲むほか、例えば色を変化させたり、点滅させたりする。このような表示とすることで、画面を切り替えたときにどの部分に動きがあったのかがすぐに把握できる。なお、ここではスキップで変化した際に強調表示を行うようにしたが、継続的な再生時にも強調表示を行うようにしても良い。

【0083】

このように本実施の形態によれば、映像属性情報から詳細なメタデータであるイベント情報を生成することが可能であり、また、このようなイベント情報が発生したシーンだけを視聴することで短時間での視聴が可能となる。

【図面の簡単な説明】

【0084】

【図1】本発明の第1の実施の形態にかかる映像要約装置の構成を示すブロック図である。

【図2】映像要約装置の概略構成を示すブロック図である。

【図3】映像属性情報作成部の概略構成を示すブロック図である。

【図4】映像属性情報の自動作成処理の流れを概略的に示すフローチャートである。

【図5】野球中継番組の画面例を示す模式図である。

【図6】典型シーンについて示す説明図である。

【図7】入力映像からの特徴量抽出方法について模式的に示す説明図である。

【図8】典型シーンの抽出処理の流れを示すフローチャートである。

【図9】属性情報領域の抽出処理の流れを示すフローチャートである。

【図10】典型シーンを用いてフレーム間の重複部分を抽出する処理について模式的に示す説明図である。

【図11】属性情報領域の抽出方法を模式的に示す説明図である。

【図12】文字領域の抽出処理の流れを示すフローチャートである。

【図13】文字領域を抽出する処理について模式的に示す説明図である。

【図14】重複する領域の処理例を模式的に示す説明図である。

【図15】ルール情報の一例を示す模式図である。

【図16】野球のスコア情報の一例を示す模式図である。

【図17】本発明の第2の実施の形態の映像属性情報作成部の概略構成を示すブロック図である。

【図18】強調表示の一例を模式的に示す説明図である。

【符号の説明】

【0085】

- 1 映像要約装置
- 10 映像コンテンツ
- 11 映像属性情報
- 15 映像属性情報出力装置
- 16 重要度算出部
- 17 映像区間選択部
- 18 説明文用イベント選択部
- 19 説明文作成部
- 20 統合部
- 151 シーン分割手段
- 152 典型シーン抽出手段
- 153 属性情報領域抽出手段
- 154 領域判定手段
- 155 文字領域抽出手段

10

20

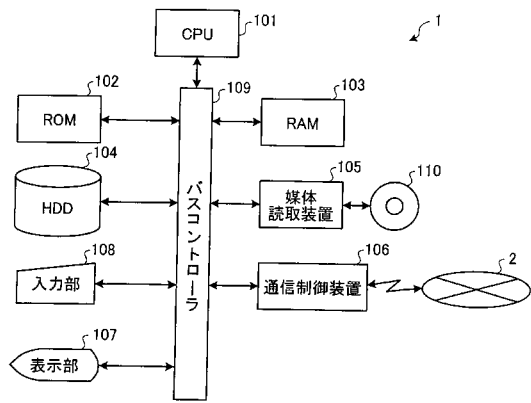
30

40

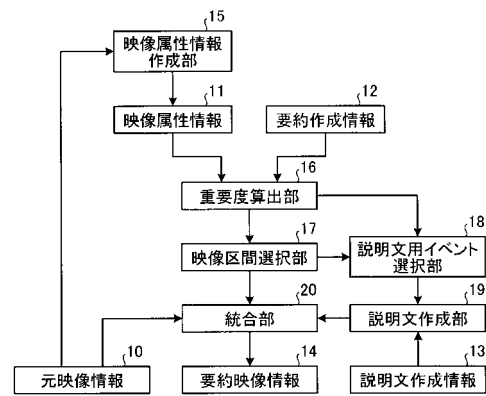
50

- 1 5 6 文字領域意味付け手段
- 1 5 7 映像属性情報出力手段

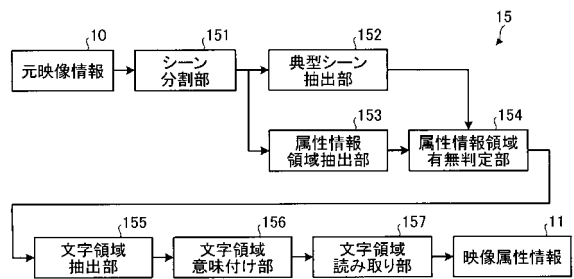
【 図 1 】



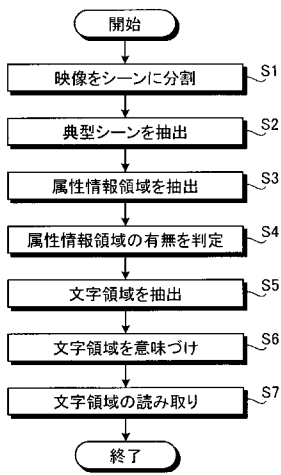
【 図 2 】



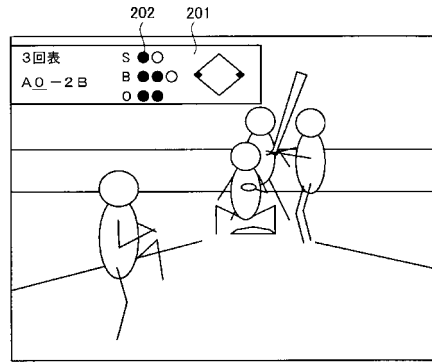
【 図 3 】



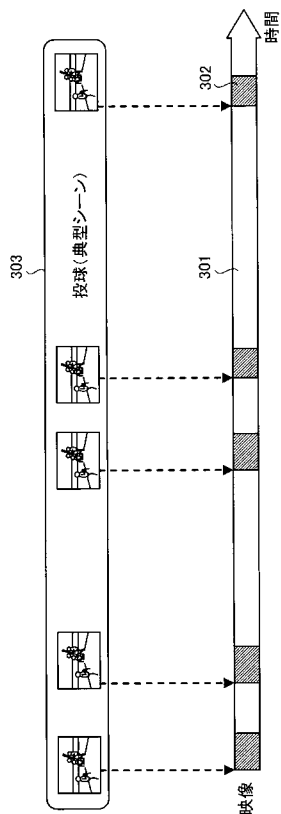
【 図 4 】



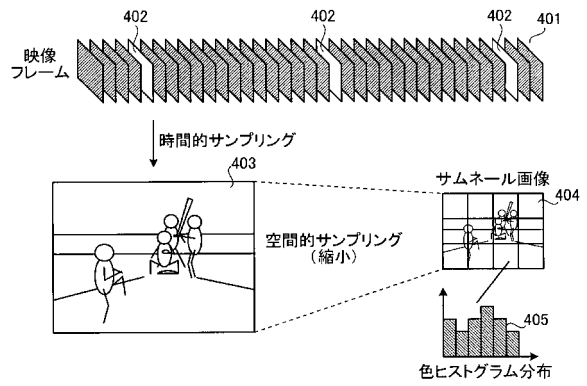
【 図 5 】



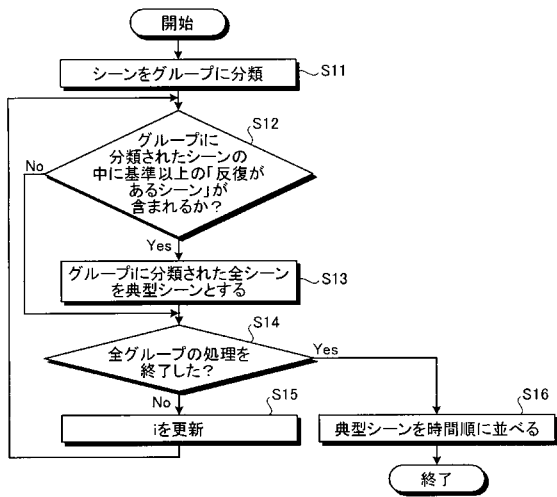
【 図 6 】



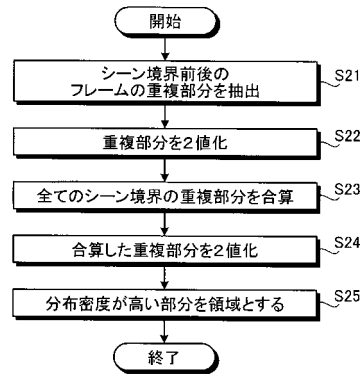
【 図 7 】



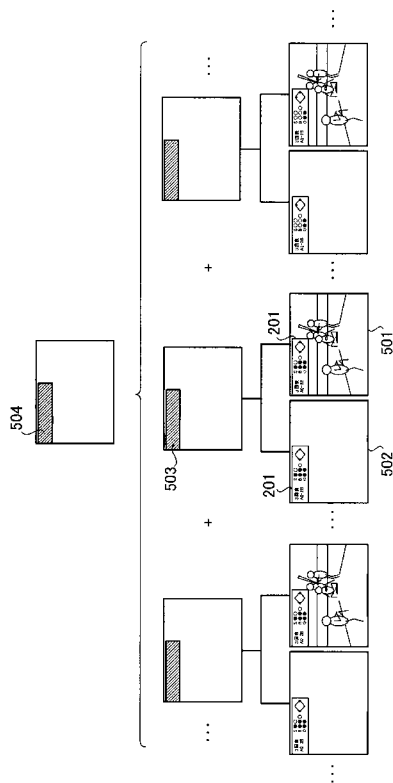
【 図 8 】



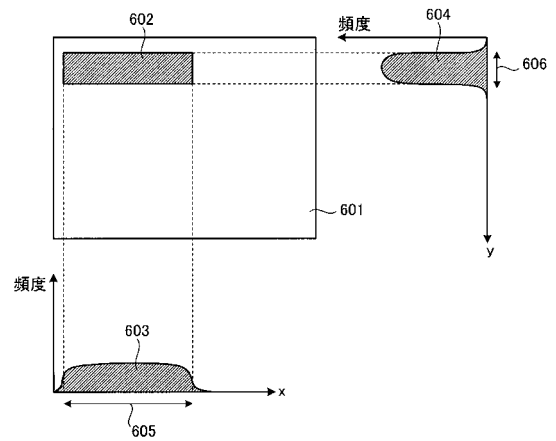
【 図 9 】



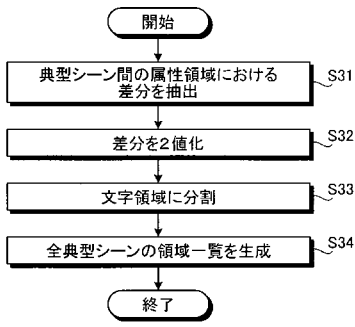
【 図 10 】



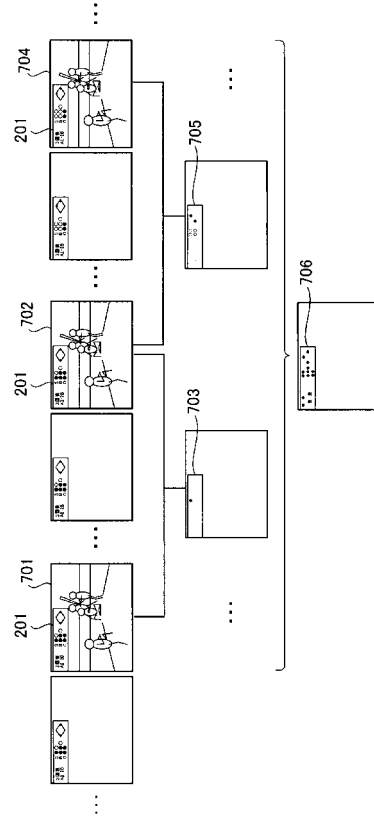
【 図 11 】



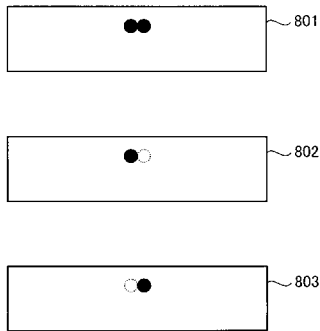
【 図 1 2 】



【 図 1 3 】



【 図 1 4 】



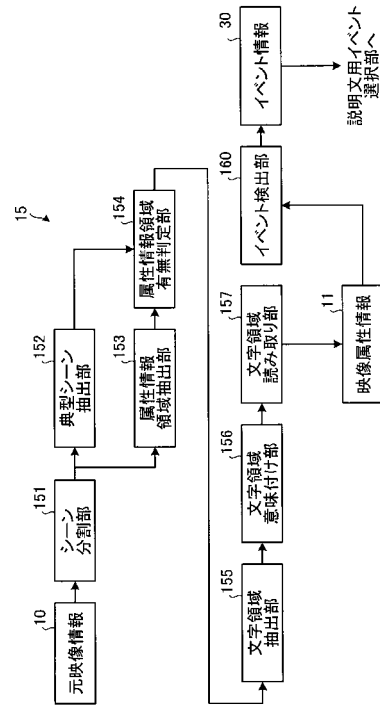
【 図 1 5 】

| 内容 | ルール (グループ) | | ルール (位置関係) |
|-------|--------------------------------|--------------------------------|--|
| | 901 | 902 | |
| ストライク | 水平並び2個領域 隣接, 同サイズ, 同色 | 水平並び2個領域 隣接, 同サイズ, 同色 | 縦並び時: ボール, アウトと同じ座標 横並び時: ボール, アウトと同じ座標 |
| ボール | 水平並び3個領域 隣接, 同サイズ, 同色 | 水平並び3個領域 隣接, 同サイズ, 同色 | 縦並び時: ストライク, アウトと同じ座標 横並び時: ストライク, アウトと同じ座標 |
| アウト | 水平並び2個領域 隣接, 同サイズ, 同色 | 水平並び2個領域 隣接, 同サイズ, 同色 | 縦並び時: ストライクより下の座標 横並び時: ストライクより下の座標 |
| 回+派/獲 | 水平並び2領域 近接, 同サイズ, 同色, 数字+文字 | 水平並び2領域 近接, 同サイズ, 同色, 数字+文字 | スコア領域のまままたは上部分 |
| 得点 | 水平並び2領域 近接, 同サイズ, 同色, 数字 | 水平並び2領域 近接, 同サイズ, 同色, 数字 | スコア領域のまままたは上部分 |
| 出塁状況 | ひし形の頂点3領域 近接, 同サイズ, 同色 | ひし形の頂点3領域 近接, 同サイズ, 同色 | スコア領域のまままたは上部分 |

【 図 1 6 】

| 時刻 | 回 | 表裏 | ストライク | ボール | アウト | 出塁 | 得点 |
|---------|---|----|-------|-----|-----|-----|-----|
| 0:41:11 | 5 | 表 | 0 | 2 | 1 | 1 | 2-3 |
| 0:42:12 | 5 | 表 | 1 | 2 | 1 | 1 | 2-3 |
| 0:43:15 | 5 | 表 | 1 | 3 | 1 | 1 | 2-3 |
| 0:44:47 | 5 | 表 | 0 | 0 | 1 | 1,2 | 2-3 |
| 0:45:20 | 5 | 表 | 1 | 0 | 1 | 1,2 | 2-3 |
| 0:46:07 | 5 | 表 | 1 | 1 | 1 | 1,2 | 2-3 |
| 0:46:52 | 5 | 表 | 2 | 1 | 1 | 1,2 | 2-3 |
| 0:47:28 | 5 | 表 | 2 | 2 | 1 | 1,2 | 2-3 |
| 0:48:15 | 5 | 表 | 0 | 0 | 2 | 1,2 | 2-3 |
| 0:49:05 | 5 | 表 | 1 | 0 | 2 | 1,2 | 2-3 |
| 0:50:32 | 5 | 表 | 1 | 1 | 2 | 1,2 | 2-3 |
| 0:53:19 | 5 | 裏 | 0 | 0 | 0 | - | 2-3 |
| 0:53:57 | 5 | 裏 | 1 | 0 | 0 | - | 2-3 |
| 0:54:29 | 5 | 裏 | 1 | 1 | 0 | - | 2-3 |
| 0:54:16 | 5 | 裏 | 1 | 2 | 0 | - | 2-3 |
| 0:56:23 | 5 | 裏 | 0 | 0 | 0 | 1 | 2-3 |
| 0:57:15 | 5 | 裏 | 0 | 1 | 0 | 1 | 2-3 |
| 0:58:01 | 5 | 裏 | 0 | 2 | 0 | 1 | 2-3 |
| 1:01:59 | 5 | 裏 | 0 | 0 | 0 | - | 2-5 |
| 1:02:17 | 5 | 裏 | 0 | 1 | 0 | - | 2-5 |
| 1:02:52 | 5 | 裏 | 1 | 1 | 0 | - | 2-5 |
| 1:02:28 | 5 | 裏 | 2 | 1 | 0 | - | 2-5 |
| 1:03:12 | 5 | 裏 | 0 | 0 | 1 | - | 2-5 |

【 図 1 7 】



【 図 1 8 】



フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

G 0 6 F 17/30 1 7 0 D
H 0 4 N 5/93 Z