

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4332988号
(P4332988)

(45) 発行日 平成21年9月16日 (2009. 9. 16)

(24) 登録日 平成21年7月3日 (2009. 7. 3)

(51) Int. Cl.

F I

H O 4 N 5/91 (2006. 01)

H O 4 N 5/91 N

H O 4 N 5/76 (2006. 01)

H O 4 N 5/76 B

請求項の数 12 (全 44 頁)

(21) 出願番号 特願2000-132918 (P2000-132918)
 (22) 出願日 平成12年4月27日 (2000. 4. 27)
 (65) 公開番号 特開2001-313898 (P2001-313898A)
 (43) 公開日 平成13年11月9日 (2001. 11. 9)
 審査請求日 平成19年1月25日 (2007. 1. 25)

(73) 特許権者 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (74) 代理人 100067736
 弁理士 小池 晃
 (74) 代理人 100086335
 弁理士 田村 榮一
 (74) 代理人 100096677
 弁理士 伊賀 誠司
 (72) 発明者 安部 素嗣
 東京都品川区北品川6丁目7番35号 ソ
 ニー株式会社内
 (72) 発明者 西口 正之
 東京都品川区北品川6丁目7番35号 ソ
 ニー株式会社内

最終頁に続く

(54) 【発明の名称】 信号処理装置及び方法

(57) 【特許請求の範囲】

【請求項 1】

少なくとも第1の信号の区間とそれ以外の信号の区間とが時分割的に存在する入力信号から、所定の時間間隔を持つ信号の特徴的パターンに基づいて、第1の信号の候補区間を検出する候補区間検出手段と、

上記候補区間内又はその前後の入力信号から、上記第1の信号らしさを表わす特徴量を抽出する特徴量抽出手段と、

上記特徴量に基づき、上記第1の信号の区間を検出する検出手段と、

上記検出手段により検出された上記第1の信号の区間について、ルール判定を行い第1の信号の区間を決定する判定手段と、

上記第1の信号区間の判定結果に基づき、上記入力信号から第1の信号区間の信号を抽出する第1の信号抽出手段と、

上記抽出された第1の信号を記録する記録手段とを有し、

上記入力信号は、映像及び音声信号からなり、上記第1の信号の区間はコマーシャルメッセージ (CM) 区間であり、上記候補区間は、音量条件と映像条件を満たす所定の長さのフレーム区間であり、

上記判定手段で利用するルール判定は、CM候補区間中に他のCM候補が存在している場合は最小長さ優先ルールを利用し、前記最小長さ優先ルール適用後に他のCM候補が存在している場合には隣接優先ルールを利用し、前記隣接優先ルール適用後に他のCM候補が存在している場合にはスコア優先ルールを利用する判定である

10

20

信号処理装置。

【請求項 2】

上記記録手段は、更に当該特徴量を記録する請求項 1 記載の信号処理装置。

【請求項 3】

上記記録された第 1 の信号を代表して索引となる情報を抽出する索引情報抽出手段と、
上記抽出された索引情報を表示する表示手段とを有する請求項 1 記載の信号処理装置。

【請求項 4】

上記表示手段は、上記索引情報に基づく複数の CM の CM 索引画像又は映像を、ユーザの指示を入力するためのアイコンと共に 1 画面内に配置して表示する請求項 3 記載の信号処理装置。

【請求項 5】

上記アイコンとしては、上記複数の CM の CM 索引画像又は映像の中で既に選択されている CM の映像及び音声を再生させる CM 再生アイコン、既に選択されている CM の開始点音声又は終了点音声又は全音声を再生させる音声再生アイコン、現在表示されている画面の前ページ又は次ページを表示させる前ページアイコン又は次ページアイコンの少なくとも 1 つを含む請求項 4 記載の信号処理装置。

【請求項 6】

上記記録手段に記録された特徴量に基づき、上記第 1 の信号区間を検索する検索手段を有する請求項 1 記載の信号処理装置。

【請求項 7】

少なくとも第 1 の信号の区間とそれ以外の信号の区間とが時分割的に存在する入力信号から、所定の時間間隔を持つ信号の特徴的パターンに基づいて、第 1 の信号の候補区間を検出する候補区間検出工程と、

上記候補区間内又はその前後の入力信号から、上記第 1 の信号らしさを表わす特徴量を抽出する特徴量抽出工程と、

上記特徴量に基づき、上記第 1 の信号の区間を検出する検出工程と、

上記検出工程により検出された上記第 1 の信号の区間について、ルール判定を行い第 1 の信号の区間を決定する判定工程と、

上記第 1 の信号区間の判定結果に基づき、上記入力信号から第 1 の信号区間の信号を抽出する第 1 の信号抽出工程と、

上記抽出された第 1 の信号を記録する記録工程とを有し、

上記入力信号は、映像及び音声信号からなり、上記第 1 の信号の区間はコマーシャルメッセージ (CM) 区間であり、上記候補区間は、音量条件と映像条件を満たす所定の長さのフレーム区間であり、

上記判定工程で利用するルール判定は、CM 候補区間中に他の CM 候補が存在している場合は最小長さ優先ルールを利用し、前記最小長さ優先ルール適用後に他の CM 候補が存在している場合には隣接優先ルールを利用し、前記隣接優先ルール適用後に他の CM 候補が存在している場合にはスコア優先ルールを利用する判定である

信号処理方法。

【請求項 8】

上記記録工程では、更に当該特徴量を記録する請求項 7 記載の信号処理方法。

【請求項 9】

上記記録された第 1 の信号を代表して索引となる情報を抽出する索引情報抽出工程と、
上記抽出された索引情報を表示する表示工程とを有する請求項 7 記載の信号処理方法。

【請求項 10】

上記表示工程では、上記索引情報に基づく複数の CM の CM 索引画像又は映像を、ユーザの指示を入力するためのアイコンと共に 1 画面内に配置して表示する請求項 9 記載の信号処理装置。

【請求項 11】

上記アイコンとしては、上記複数の CM の CM 索引画像又は映像の中で既に選択されて

10

20

30

40

50

いるＣＭの映像及び音声を再生させるＣＭ再生アイコン、既に選択されているＣＭの開始点音声又は終了点音声又は全音声を再生させる音声再生アイコン、現在表示されている画面の前ページ又は次ページを表示させる前ページアイコン又は次ページアイコンの少なくとも１つを含む請求項１０記載の信号処理方法。

【請求項１２】

上記記録された特徴量に基づき、上記第１の信号区間を検索する検索工程を有する請求項７記載の信号処理方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、例えば映像音声記録再生装置に適用されるものであり、特に、テレビジョン放送に付加されるコマーシャルメッセージを蓄積、閲覧、視聴、検索等する場合に好適な信号処理装置及び方法に関するものである。

【０００２】

【従来の技術】

テレビジョン放送やラジオ放送等の公共放送は、一部の有料放送などを除き、企業や団体などを提供者とするコマーシャルメッセージ（以下、単にＣＭとする）を、番組本編中に挿入して放送されることが多い。

【０００３】

視聴者は番組本編の視聴中にはＣＭを不要と感ずることもある反面、ＣＭは未知の商品などに対する直接的な情報源であることや、有名俳優の登場するＣＭやストーリー性のあるＣＭなどに代表されるように、作品としてＣＭ自体が視聴対象となる場合も多い。

【０００４】

しかしながら、番組に関しては、例えば番組表などにより事前に放送時間を知ることができるのに対し、ＣＭに関しては、一般には何時放送されるのかを事前には知り得ないため、特定のＣＭを必要に応じて視聴することは難しい。そこで、例えば所定の時間内に放送されるＣＭがすべて蓄積され、必要に応じた閲覧、視聴、検索など可能となれば、上記のような多様な要求に対し、利便性の高い対応が図れると考えられる。

【０００５】

【発明が解決しようとする課題】

ところで、従来より、放送信号からＣＭを検出に関する技術としては、例えば特開平８－３１７３４２号公報、特開平３－１５８０８６号公報、特開平３－２６２２８７２号公報などに、ＣＭをスキップして視聴する技術が公開されている。しかしながら、これら技術ではＣＭの除去はできても、ＣＭの蓄積、閲覧、視聴、検索などを行うことは出来ない。

【０００６】

また、特開平１０－２２４７２４号公報には、ＣＭに付随する情報を検出、蓄積し、同一コマーシャルを１回だけ再生する技術が公開されている。しかしながら、この技術をしててもＣＭの蓄積、閲覧、視聴、検索などの要求を満たすものではなく、また、ステレオモードで放送されることの多いＣＭ区間とステレオ番組区間とが区別されないため、使用できる範囲は非常に限定されている。

【０００７】

以上のようなことから、放送信号からＣＭ部分を高精度に検出できると共に、そのＣＭの蓄積、閲覧、視聴、検索等を可能とすることが望まれている。

【０００８】

そこで、本発明は、以上のような状況を鑑みてなされたものであり、例えば、放送信号からＣＭ部分を高精度に検出できると共に、そのＣＭの蓄積、閲覧、視聴、検索等を可能とする、信号処理装置及び方法を提供することを目的とする。

【０００９】

【課題を解決するための手段】

本発明の信号処理装置は、少なくとも第１の信号の区間とそれ以外の信号の区間とが時

10

20

30

40

50

分割的に存在する入力信号から、所定の時間間隔を持つ信号の特徴的パターンに基づいて、第1の信号の候補区間を検出する候補区間検出手段と、上記候補区間内又はその前後の入力信号から、上記第1の信号らしさを表わす特徴量を抽出する特徴量抽出手段と、上記特徴量に基づき、上記第1の信号の区間を検出する検出手段と、上記検出手段により検出された上記第1の信号の区間について、ルール判定を行い第1の信号の区間を決定する判定手段と、上記第1の信号区間の判定結果に基づき、上記入力信号から第1の信号区間の信号を抽出する第1の信号抽出手段と、上記抽出された第1の信号を記録する記録手段とを有し、上記入力信号は、映像及び音声信号からなり、上記第1の信号の区間はコマーシャルメッセージ（CM）区間であり、上記候補区間は、音量条件と映像条件を満たす所定の長さのフレーム区間であり、上記判定手段で利用するルール判定は、CM候補区間中に他のCM候補が存在している場合は最小長さ優先ルールを利用し、前記最小長さ優先ルール適用後に他のCM候補が存在している場合には隣接優先ルールを利用し、前記隣接優先ルール適用後に他のCM候補が存在している場合にはスコア優先ルールを利用する判定である。

10

【0010】

また、本発明の信号処理方法は、少なくとも第1の信号の区間とそれ以外の信号の区間とが時分割的に存在する入力信号から、所定の時間間隔を持つ信号の特徴的パターンに基づいて、第1の信号の候補区間を検出する候補区間検出工程と、上記候補区間内又はその前後の入力信号から、上記第1の信号らしさを表わす特徴量を抽出する特徴量抽出工程と、上記特徴量に基づき、上記第1の信号の区間を検出する検出工程と、上記検出工程により検出された上記第1の信号の区間について、ルール判定を行い第1の信号の区間を決定する判定工程と、上記第1の信号区間の判定結果に基づき、上記入力信号から第1の信号区間の信号を抽出する第1の信号抽出工程と、上記抽出された第1の信号を記録する記録工程とを有し、上記入力信号は、映像及び音声信号からなり、上記第1の信号の区間はコマーシャルメッセージ（CM）区間であり、上記候補区間は、音量条件と映像条件を満たす所定の長さのフレーム区間であり、上記判定工程で利用するルール判定は、CM候補区間中に他のCM候補が存在している場合は最小長さ優先ルールを利用し、前記最小長さ優先ルール適用後に他のCM候補が存在している場合には隣接優先ルールを利用し、前記隣接優先ルール適用後に他のCM候補が存在している場合にはスコア優先ルールを利用する判定である。

20

30

【0011】

すなわち本発明によれば、入力信号から第1の信号を検出して記録することにより、入力信号から第1の信号を分離し、整理することが可能となる。また、本発明によれば、検出した第1の信号を特徴づける特徴量を抽出し、第1の信号と合わせて記録することにより、例えば第1の信号の閲覧、除去、検索などが容易に実現可能となる。また本発明によれば、第1の信号からそれを代表する索引情報を抽出し、表示することにより、第1の信号の閲覧、検索などが容易に実現可能となる。さらに、本発明によれば、記録された第1の信号から同一のものを例えば除去することで、記録効率を高め、また、閲覧時には同一の信号を複数回閲覧したり検索したり無駄を省くことができる。さらに、本発明によれば、第1の信号が部分的に同一若しくは類似したものを検出することで、様々な検索が可能となっている。

40

【0012】

【発明の実施の形態】

以下、本発明の好ましい実施の形態について、図面を参照しながら説明する。

【0013】

図1には、本発明の第1の実施の形態の放送信号処理装置の概略構成を示す。

【0014】

この図1に示す放送信号処理装置には、映像信号、音声信号、制御信号、番組ガイド信号等からなる信号200aが入力される。この入力信号200aは、地上波放送、衛星放送、ケーブル放送、電話回線等を通じて受信されたもの、若しくはそれらが図示しない別の

50

記録装置に記録された後、再生されたものである。ここで、映像信号、音声信号は、番組本編及びCM部分を含む、放送信号の主要部を構成する信号である。また、制御信号には、放送モード、放送時刻、放送周波数もしくはチャンネル等の情報が含まれ、番組ガイド信号には、デジタル放送や電話回線等により受信される、上記映像信号や音声信号に関連のあるデータを含む。以下の説明では、これらが単数の放送番組の信号であることを想定して説明するが、同時に複数のチャンネルの信号が入力されてもよい。

【0015】

この図1に示す放送信号処理装置において、CM検出部202は、上記映像信号/音声信号/制御信号/番組ガイド信号等よりなる入力信号200aから、少なくともCM区間の開始時刻(CM開始時刻と呼ぶ)及びCM区間の長さ(以下、CM長さと呼ぶ)を検出し、それらCM開始時刻及びCM長さを含むCM検出信号202aを出力する。なお、このCM検出部202の詳細な動作とCM開始時刻及びCM長さについては後述する。当該CM検出部202にて検出されたCM検出信号202aは、CM抽出部201とCM特徴抽出部203に送られる。

10

【0016】

CM抽出部201は、CM検出部202より供給されたCM検出信号202aに基づいて、入力信号202aからCM区間に相当する部分の信号201aを抽出する。すなわち、CM抽出部201は、CM検出部202より供給されたCM検出信号202aに含まれるCM開始時刻からCM長さで指定される時間までのCM区間に相当する信号部分(以下、適宜、CM部分信号201aと呼ぶ)を、上記入力信号200aから抽出し、そのCM区間に相当する映像信号/音声信号/制御信号/番組ガイド信号からなるCM部分信号201aを出力する。このCM部分信号201aは、CM記録部205に送られる。なお、上記CM検出部202からのCM検出信号202aは、入力信号200aに対して、その検出処理時間分の遅延が生じているので、CM抽出部201は、内設或いは外付けされた例えば磁気記録装置などからなる一時記憶装置を用いて、上記遅延分を吸収するようにしている。

20

【0017】

CM記録部205は、例えば磁気テープや磁気ディスク、光磁気ディスク、記録可能な光ディスク、半導体メモリ等の記録媒体の何れか一つ若しくはそれらの組み合わせを用いて、信号を記録及び再生する装置からなる。当該CM記録部205は、上記CM抽出部201からCM部分信号(CM区間に相当する映像信号/音声信号/制御信号/番組ガイド信号)201aが供給されたとき、そのCM部分信号201aを記録する。

30

【0018】

ここで、図2には、上記図1の構成のうち、上述したCM抽出部201、CM検出部202、CM記録部205の部分のみを抜き出したサブセットを示しており、図3にはこれら図2に抜き出した部分の動作の流れをフローチャートとして示す。

【0019】

図3において、まず、図2のサブセットでは、ステップS220として、上記入力信号200aである映像信号/音声信号/制御信号/番組ガイド信号が順次入力されると、ステップS221として、その入力信号200aがCM抽出部201に前記内設或いは外付けされた一時記憶装置に保存される。

40

【0020】

同時に、図2のサブセットでは、前述のようにCM区間の開始時刻とCM区間の長さをCM検出部202により検出する。ここで、ステップS222として、入力信号200aについて、信号がCM区間に相当するか否かの検出を行い、CM区間であると判定した場合にはステップS223の処理に進み、CM区間でないと判定した場合にはステップS225の処理に進む。

【0021】

ステップS223の処理に進むと、CM抽出部201により入力信号200aからCM区間に相当する映像信号/音声信号/制御信号/番組ガイド信号の信号201aを抽出し、

50

その後、ステップ S 2 2 4 において当該抽出された信号を C M 記録部 2 0 5 により保存する。

【 0 0 2 2 】

その後、図 2 のサブセットでは、ステップ S 2 2 5 の処理に進むと、C M 抽出部 2 0 1 に一時的に記憶した信号を破棄し、ステップ S 2 2 0 に戻って上述した処理を繰り返す。

【 0 0 2 3 】

図 1 に戻り、C M 特徴抽出部 2 0 3 では、上記 C M 検出部 2 0 2 より供給された C M 検出信号 2 0 2 a に基づいて、入力信号 2 0 0 a から C M の特徴を抽出する。すなわち、C M 特徴抽出部 2 0 3 は、C M 検出部 2 0 2 より供給された C M 検出信号 2 0 2 a に含まれる C M 開始時刻から C M 長さで指定される時間までの C M 区間に相当する信号部分について、その信号部分が有する特徴を抽出（後述する C M としての特徴を表すことになる特徴量を抽出）し、その特徴量 2 0 3 a を C M 記録部 2 0 5 へ出力する。なお、上記 C M 検出部 2 0 2 からの C M 検出信号 2 0 2 a は、入力信号 2 0 0 a に対して、その検出処理時間分の遅延が生じているので、C M 特徴抽出部 2 0 3 は、内設或いは外付けされた例えば磁気記録装置などからなる一時記憶装置を用いて、上記遅延分を吸収するようにしている。当該一時記憶装置は、C M 抽出部 2 0 1 の一時記憶装置と共有することが可能である。

【 0 0 2 4 】

ここで、上記 C M を特徴付ける特徴量としては、映像信号及び音声信号から抽出される、以下の振幅特徴量、スペクトル特徴量、線形予測係数、輝度ヒストグラム及び色ヒストグラム、平均輝度、輝度差分エネルギー、カットチェンジ回数及び時刻などのような物理量の一部又は全部を用いることができる。これらの特徴量は、同一の C M では同一のパターンとなり、異なる C M では異なるパターンとなるものであるから、C M を特徴付けるものと言うことができる。

【 0 0 2 5 】

以下の説明では、入力音声信号及び映像信号は離散化されているとし、図 4 の (b) に示す $S[m]$ により入力音声信号を表し、 $m = 0, 1, \dots, M - 1$ により離散化された時間を表し、 M により処理フレームに対応する離散時間を表わすことにする。また、図 4 の (a) に示す $I[x, y; l]$ により入力映像信号を表し、 $l = 0, 1, \dots, L - 1$ により入力映像の各映像フレームを表し、 L により処理フレームに対応する映像フレーム数を、 $x = 0, \dots, X - 1$ は横方向の画素番号を、 X は横方向の映像サイズを、 $y = 0, \dots, Y - 1$ は縦方向の画素番号を、 Y は縦方向の映像サイズを表すことにする。なお、処理フレームは、所定の時間長の処理単位であり、例えば 2 5 0 m s とする。さらに、図 4 の (c) に示すように、C M 区間長に対応するフレーム数を N で表わす。

【 0 0 2 6 】

以下、各特徴量について説明する。

【 0 0 2 7 】

先ず、処理フレーム n の振幅特徴量 $A[n]$ は、フレーム n 内の音声信号の平均二乗振幅値であり、式 (1) ように得られる。

【 0 0 2 8 】

【 数 1 】

$$A[n] = \frac{1}{M} \sum_{m=0}^{M-1} S^2[m + nM] \quad (1)$$

【 0 0 2 9 】

これをCM区間の全フレームについてまとめたベクトルを $A = (A[0], A[1], \dots, A[N-1])$ で表す。

【0030】

次に、処理フレーム n のスペクトル特徴量 $F[k;n]$ は、フレーム n 内の平均スペクトルであり、式(2)のようにして得られる。

【0031】

【数2】

$$F[k;n] = \left| \sum_{m=0}^{M-1} S[m] e^{-2\pi j m k / M} \right|^2 \quad (2)$$

10

【0032】

ただし、 $k = 0, \dots, K$ は離散化周波数を表わす番号であり、 K は対象とする最高周波数を表わす。この演算は、FFTや線形予測分析などを用いて実装される。周波数離散化のステップは、式(2)の k を再離散化することで、例えば1kHz毎といった線形なものにすることも、例えば1オクターブ毎といった非線形なものにすることもできる。また、これをCM区間の全フレームについてまとめたベクトルを $F = (F[0;0], F[1;0], \dots, F[K-1;N-1])$ で表わす。

20

【0033】

次に、処理フレーム n における線形予測係数 $P[k;n]$ は、例えばLinear Prediction of Speech (Markel 他者、Springer Verlag, 1978) 等 に示される、LPCアルゴリズムを用いて計算される。 $k = 0, \dots, K-1$ は、線形予測係数を表わす番号、 K は予測次数である。また、これをCM区間の全フレームについてまとめたベクトルを $P = (P[0;0], P[1;0], \dots, P[K-1;N-1])$ で表わす。

【0034】

次に、処理フレーム n の輝度ヒストグラム $H_1[q;n]$ は、フレーム n 内の映像信号の輝度のヒストグラムである。ただし、 $q = 0, \dots, Q-1$ は、輝度に対する升目を表わすインデックス番号であり、 Q はピストグラムの升目の数である。

30

【0035】

処理フレーム n の色ヒストグラム $H_C[q;n]$ は、フレーム n 内の信号の各色毎の強度のヒストグラムである。ただし、 $q = 0, \dots, Q-1$ は、色及び強度の升目を表わすインデックス番号であり、 Q はヒストグラムの升目の数である。

【0036】

これら輝度ヒストグラム、色ヒストグラムをそれぞれCM区間の全フレームについてまとめたベクトルを、 $H_1 = (H_1[0;0], H_1[1;0], \dots, H_1[Q-1;N-1])$ 、及び、 $H_C = (H_C[0;0], H_C[1;0], \dots, H_C[Q-1;N-1])$ により表す。

40

【0037】

次に、処理フレーム n の平均輝度 $B[n]$ は、フレーム n 内の映像信号の平均輝度であり、式(3)に示すように求められる。

【0038】

【数3】

$$B[n] = \frac{1}{XYL} \sum_{l=0}^{L-1} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} I[x, y; l] \quad (3)$$

【 0 0 3 9 】

10

これをCM区間の全フレームについてまとめたベクトルを $B = (B[0], B[1], \dots, B[N-1])$ で表す。

【 0 0 4 0 】

次に、処理フレーム n の輝度差分エネルギー $D[n]$ は、隣接する映像フレームの画素間差分エネルギーであり、例えば式(4)により求められる。

【 0 0 4 1 】

【数4】

20

$$D[n] = \frac{1}{XY(L-1)} \sum_{l=1}^{L-1} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \left(I[x, y; l] - I[x, y; l-1] \right)^2 \quad (4)$$

【 0 0 4 2 】

これをCM区間の全フレームについてまとめたベクトルを $D = (D[0], D[1], \dots, D[N-1])$ で表わす。なおこの演算は、画素間差分に代えて、全画面の平均輝度の差分や、水平×垂直方向に 8×8 画素や 16×16 画素といったブロック間の差分を用いることもできる。

30

【 0 0 4 3 】

次に、処理フレーム n のカットチェンジ回数 $C[n]$ は、処理フレーム n 内で、隣接映像フレーム間の画素差分エネルギーが所定の閾値を越えるフレーム数であり、式(5)により求められる。

【 0 0 4 4 】

【数5】

40

$$C[n] = \text{Count}_{l=1}^{L-1} \left(\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \left(I[x, y; l] - I[x, y; l-1] \right)^2 \geq D_{thsd} \right) \quad (5)$$

【 0 0 4 5 】

ただし、 $\text{Count}_a^b(f)$ は、 a から b までの区間内で評価式 f が満足される回数を表わす関数であり、 D_{thsd} は、輝度差分エネルギーに関する所定の閾値である。これをCM区間の全処理フレームについてまとめたベクトルを $C = (C[0], C[1], \dots, C[N-1])$ で表わす。

【 0 0 4 6 】

50

また、カットチェンジ時刻は、 $C[n] > 0$ となるフレーム番号 n として求めることができ、 CM 区間全体でのカットチェンジ回数は、 $C[n]$ の総和として求めることができる。

【 0 0 4 7 】

上記 CM 特徴抽出部 2 0 3 では、以上に説明した特徴量の一部又は全部を、各 CM 区間毎に検出する。すなわち、 CM 区間を特徴付ける特徴量は、式 (6) に示されるベクトル V として表わされることになる。

【 0 0 4 8 】

$$V = \{ A, F, P, H_I, H_C, B, D, C \} \quad (6)$$

また、特徴ベクトル V は、式 (7) のようにも表すことができる。

【 0 0 4 9 】

$V = \{ V[0], V[1], \dots, V[N-1] \}$ (7) 但し、 $V[n]$ は式 (8) に示されるように、処理フレーム n における各々の特徴量を、ベクトルとしてまとめたものである。

【 0 0 5 0 】

$V = \{ A[n], F[k;n], P[k;n], H_I[n], H_C[n], B[n], D[n], C[n] \}$ (8) CM 特徴抽出部 2 0 3 により抽出された上記の特徴量は、先に説明した CM 抽出部 2 0 1 で抽出された CM 区間に相当する CM 部分信号 (映像信号 / 音声信号 / 制御信号 / 番組ガイド信号) 2 0 1 a と合わせて、 CM 記録部 2 0 5 に蓄積される。

【 0 0 5 1 】

ここで、図 5 には、上記図 1 の構成のうち、上述した CM 抽出部 2 0 1、 CM 検出部 2 0 2、 CM 特徴抽出部 2 0 3、 CM 記録部 2 0 5 の部分のみを抜き出したサブセットを示しており、図 6 にはこれら図 3 に抜き出したサブセット部分の動作の流れをフローチャートとして示す。なお、図 6 のフローチャートにおいて、図 3 と重複する部分については図 3 と同じ指示符号を付してその説明は省略し、 CM 特徴抽出部 2 0 3 に係る部分に限って説明する。

【 0 0 5 2 】

図 6 において、図 5 のサブセットでは、前記ステップ $S 2 2 0$ により、の処理後、ステップ $S 2 3 3$ の処理に進むと、上記入力信号 2 0 0 a である映像信号 / 音声信号 / 制御信号 / 番組ガイド信号が順次入力されると、ステップ $S 2 3 1$ として、その入力信号 2 0 0 a が CM 特徴抽出部 2 0 3 に前記内設或いは外付けされた一時記憶装置に保存される。

【 0 0 5 3 】

同時に、 CM 検出部 2 0 2 により CM 区間の開始時刻と CM 区間の長さが検出され、前記ステップ $S 2 2 2$ において CM 区間であると判定した場合にはステップ $S 2 3 3$ の処理に進み、 CM 区間でないと判定した場合にはステップ $S 2 3 5$ の処理に進む。

【 0 0 5 4 】

図 5 のサブセットでは、ステップ $S 2 3 2$ の処理に進むと、 CM 特徴抽出部 2 0 3 により、 CM 区間に相当する映像信号 / 音声信号 / 制御信号 / 番組ガイド信号から、 CM の特徴量を抽出し、ステップ $S 2 3 4$ において当該抽出された特徴量を CM 記録部 2 0 5 により保存する。

【 0 0 5 5 】

その後、図 5 のサブセットでは、ステップ $S 2 3 5$ の処理に進むと、 CM 特徴抽出部 2 0 3 に一時的に記憶した信号を破棄し、ステップ $S 2 2 0$ に戻って上述した処理を繰り返す。

【 0 0 5 6 】

図 1 に戻り、上述のようにして CM 記録部 2 0 5 に記録された CM 区間に相当する CM 部分信号 (映像信号 / 音声信号 / 制御信号 / 番組ガイド信号) 2 0 1 a と、 CM 区間の特徴量 2 0 3 a は、当該 CM 記録部 2 0 5 から読み出され、 CM 索引生成部 2 0 6 と、 CM 閲覧部 2 0 8 に送られる。

【 0 0 5 7 】

CM 索引生成部 2 0 6 は、 CM 記録部 2 0 5 から供給された CM 部分信号 2 0 1 a と特徴

10

20

30

40

50

量 2 0 3 a に基づいて、C M を代表して索引となる情報を生成し、その情報（以下、C M 索引情報と呼ぶ）2 0 6 a を C M 閲覧部 2 0 8 に送る。

【 0 0 5 8 】

ここで、C M 索引情報 2 0 6 a としては、例えば以下の開始点画像、カット点画像、カット点映像、開始部音声、終了部音声等を示す情報を用いる。

【 0 0 5 9 】

以下、これらの C M 索引情報 2 0 6 a について説明する。

【 0 0 6 0 】

まず、開始点画像とは、C M 区間の開始点の画像であり、この開始点画像を索引情報の一つとする。

【 0 0 6 1 】

カット点画像とは、C M 区間の各カット点の画像であり、このカット点画像を索引情報の一つとする。なお、カット点は、特徴量 C [n] が 0 より大きい処理フレームであるので、その時刻の画像を用いる。通例、一つの C M 内にカット点は複数あるので、例えば最初又は最後のカット点といった、所定の基準により一枚の画像を選択する。

【 0 0 6 2 】

カット点映像とは、C M 区間の各カット点の画像をつないで時系列化し、映像としたものであり、このカット点映像を索引情報の一つとする。通例、一つの C M 内にカット点は複数あるので、カット点の画像を全てつなぐことで新たな短い映像が生成される。

【 0 0 6 3 】

開始部音声とは、C M 区間の最初の部分の所定の時間、例えば 2 秒間の音声であり、この開始部音声を索引情報の一つとする。特に、音声の開始部では、短い C M のイントロダクションとして特徴的な音声が存在する場合があるため、これを C M 索引とすることは有用である。

【 0 0 6 4 】

終了部音声とは、C M 区間の最初の部分の所定の時間、例えば 2 秒間の音声であり、この終了部音声を索引情報の一つとする。特に、音声の終了部では、商品名また会社や団体が共通に使用する映像や音声などが存在することが多いため、これを C M の索引とすることは有用である。

【 0 0 6 5 】

上記 C M 記録部 2 0 5 からの C M 部分信号 2 0 1 a 及び C M 区間の特徴量 2 0 3 a（以下これらを纏めて、適宜、記録部再生信号 2 0 5 a と呼ぶ）と、上記 C M 索引生成部 2 0 6 からの C M 索引情報 2 0 6 a とが供給される C M 閲覧部 2 0 8 は、表示プロセッサ 8 0 1 と表示部 8 0 2 とからなる。

【 0 0 6 6 】

表示部 8 0 2 は、例えば C R T（陰極線管）や液晶モニタ等の表示デバイスとスピーカなどで構成され、映像及び音声をユーザに提示する。

【 0 0 6 7 】

また、C M 閲覧部 2 0 8 の表示プロセッサ 8 0 1 には、ユーザ 2 0 9 の選択指示 2 0 9 a に応じて C M 選択部 2 0 7 が発生したユーザ選択情報 2 0 7 a も供給される。すなわち、後述するように、表示部 8 0 2 に表示された C M 索引画像又は映像やアイコン等をユーザ 2 0 9 が閲覧 2 0 8 a し、その表示部 8 0 2 に表示された C M 索引画像又は映像やアイコン等に対して、例えばマウスやリモートコマンド、タッチパネル等のポインティングデバイスを通じてユーザ 2 0 9 が選択指示 2 0 9 a を与えると、C M 選択部 2 0 7 はその選択指示 2 0 9 a に応じたユーザ選択情報 2 0 7 a を発生し、このユーザ選択情報 2 0 7 a が C M 閲覧部 2 0 8 の表示プロセッサ 8 0 1 に送られる。

【 0 0 6 8 】

表示プロセッサ 8 0 1 は、上記 C M 索引情報 2 0 6 a 及び記録部再生信号 2 0 5 a（特に映像 / 音声信号部分）と、上記 C M 選択部 2 0 7 からのユーザ選択情報 0 2 7 b とを受け、ユーザに提示する表示の操作を行う。なおこれは、プロセッサとソフトウェアなどによ

10

20

30

40

50

り実装される。

【 0 0 6 9 】

以下、図 7 に示す表示部 8 0 2 の表示画面例を参照しながら、上記表示プロセッサ 8 0 1 の動作を説明する。

【 0 0 7 0 】

表示プロセッサ 8 0 1 は、先ず、図 7 に示すように、上記 C M 索引情報 2 0 6 a に基づく複数の C M についての C M 索引画像又は映像 8 1 0 を、1 画面内に配置して、表示部 8 0 2 に表示する。なお、C M 索引画像又は映像 8 1 0 としては、各 C M の開始点画像又はカット点映像が表示される。C M 数が多い場合（C M 索引画像又は映像 8 1 0 の数が多い場合）には、複数のページに分けた表示等が行われる。また、表示プロセッサ 8 0 1 は、C M 選択部 2 0 7 においてユーザの指示を入力するためのアイコン 8 1 1 ~ 8 1 4 等を、上記 C M 索引画像又は映像 8 0 1 と共に 1 画面内に配置して、表示部 8 0 2 に表示する。図 7 の表示例では、1 画面内に 1 2 個の C M 索引画像又は映像 8 1 0 と、C M 再生アイコン 8 1 1、音声再生アイコン 8 1 2、前ページアイコン 8 1 3、次ページアイコン 8 1 4 が表示されている。

10

【 0 0 7 1 】

この図 7 のような表示がなされている状態で、C M 選択部 2 0 7 からのユーザ選択情報 2 0 7 a を受けると、表示プロセッサ 8 0 1 は、そのユーザ選択情報 2 0 7 a に従い、表示を変更する。すなわち、表示プロセッサ 8 0 1 は、例えば C M 選択部 2 0 7 を介してユーザ 2 0 9 が各 C M 索引画像又は映像 8 1 0 のうち何れか或いは複数を選択すると、その選択指示に応じた C M 索引画像又は映像 8 1 0 を、例えばハイライト表示等する。

20

【 0 0 7 2 】

また、図 7 のような表示がなされている状態で、例えば C M 選択部 2 0 7 を介してユーザ 2 0 9 により各アイコンのうちの何れかが選択指示されると、表示プロセッサ 8 0 1 は、その選択指示がなされたアイコンに応じた処理を行う。すなわち、表示プロセッサ 8 0 1 は、例えば、C M 再生アイコン 8 1 1 がユーザ選択情報 2 0 7 a により指定された場合、上記 C M 索引画像又は映像 8 1 0 の中で既に選択されている C M の映像及び音声を再生させる。また、表示プロセッサ 8 0 1 は、例えば音声再生アイコン 8 1 2 がユーザ選択情報 2 0 7 a により指定されると、上記 C M 索引画像又は映像 8 1 0 の中で既に選択されている C M の開始点音声又は終了点音声又は全音声を再生させる。また、表示プロセッサ 8 0 1 は、例えば、前ページアイコン 8 1 3 がユーザ選択情報 2 0 7 a により指定されると、その時の表示画面の直前に表示されていたページの C M 索引画像又は映像 8 1 0 群を画面上に表示させ、一方、例えば次ページアイコン 8 1 4 がユーザ選択情報 2 0 7 a により指定されると、次のページの C M 索引画像又は映像 8 1 0 群（閲覧されていない他の C M 索引画像又は映像 8 1 0 群）を画面上に表示させる。

30

【 0 0 7 3 】

ここで、図 8 には、上記図 1 の構成のうち、上述した C M 記録部 2 0 5、C M 索引生成部 2 0 6、C M 閲覧部 2 0 8、C M 選択部 2 0 7 の部分のみを抜き出したサブセットを示しており、図 9 にはこれら図 8 に抜き出したサブセット部分の特に C M 閲覧部 2 0 8 の動作の流れをフローチャートとして示す。

40

【 0 0 7 4 】

図 9 において、C M 閲覧部 2 0 8 は、先ず、ステップ S 2 4 0 として、最初のページの C M 群を指定し、次に、ステップ S 2 4 1 として、上記指定された C M 群に対応する映像信号 / 音声信号を C M 記録部 2 0 5 より取得し、その後、ステップ S 2 4 2 として、C M 索引生成部 6 により C M 索引情報 2 0 6 a が生成される。

【 0 0 7 5 】

次に、ステップ S 2 4 3 として、C M 閲覧部 2 0 8 の表示プロセッサ 8 0 1 は、上記 C M 索引情報 2 0 6 a に基づいて、上記図 7 に示したような表示を行うための表示画面を生成して表示部 8 0 2 に表示させ、ステップ S 2 4 4 として、その表示状態でユーザ 2 0 9 の指示入力（ユーザ選択情報 2 0 7 a の入力）がなされるまで待機する。

50

【 0 0 7 6 】

ここで、ユーザ 2 0 9 の指示入力に応じて上記 C M 選択部 2 0 7 により生成されたユーザ選択情報 2 0 7 a が、C M 索引を指定するものであった場合、表示プロセッサ 8 0 1 は、ステップ S 2 4 5 として、ユーザ選択情報 2 0 7 a により指定された C M 索引画像又は映像 8 1 0 を選択し、ステップ S 2 4 3 として、当該選択した C M 索引画像又は映像 8 1 0 をハイライト表示した表示画面を再生成して表示部 8 0 2 に表示させた後、ステップ S 2 4 4 により、その表示状態でユーザ 2 0 9 の指示入力の待機状態となる。

【 0 0 7 7 】

また、上記ユーザ選択情報 2 0 7 a が、C M 再生アイコン 8 1 1 を指定するものであった場合、表示プロセッサ 8 0 1 は、ステップ S 2 4 6 として、ユーザ選択情報 2 0 7 a により既に指定されている C M 索引画像又は映像 8 1 0 に対応する C M の映像及び音声信号を C M 記録部 2 0 5 から再生させ、ステップ S 2 4 3 として、当該再生された映像信号に応じた表示画面を再生成して表示部 8 0 2 に表示させた後、ステップ S 2 4 4 により、その表示状態でユーザ 2 0 9 の指示入力の待機状態となる。

10

【 0 0 7 8 】

また、上記ユーザ選択情報 2 0 7 a が、音声再生アイコン 8 1 2 を指定するものであった場合、表示プロセッサ 8 0 1 は、ステップ S 2 4 7 として、ユーザ選択情報 2 0 7 a により既に指定されている C M 索引画像又は映像 8 1 0 に対応する C M の開始音声又は終了音声又は全音声信号を C M 記録部 2 0 5 から再生させ、ステップ S 2 4 3 として、当該再生された音声信号に応じた音声をスピーカから出力させた後、ステップ S 2 4 4 により、その状態でユーザ 2 0 9 の指示入力の待機状態となる。

20

【 0 0 7 9 】

また、上記ユーザ選択情報 2 0 7 a が、前ページアイコン 8 1 3 を指定するものであった場合、表示プロセッサ 8 0 1 は、ステップ S 2 4 8 として、その直前に表示されていた表示画面の C M 索引画像又は映像 8 1 0 に対応する C M の映像及び音声信号を C M 記録部 2 0 5 から再生させ、ステップ S 2 4 3 として、当該再生された映像信号に応じた表示画面を表示部 8 0 2 に表示させた後、ステップ S 2 4 4 により、その表示状態でユーザ 2 0 9 の指示入力の待機状態となる。

【 0 0 8 0 】

また、上記ユーザ選択情報 2 0 7 a が、次ページアイコン 8 1 4 を指定するものであった場合、表示プロセッサ 8 0 1 は、ステップ S 2 4 9 として、次に表示される表示画面の C M 索引画像又は映像 8 1 0 に対応する C M の映像及び音声信号を C M 記録部 2 0 5 から再生させ、ステップ S 2 4 3 として、当該再生された映像信号に応じた表示画面を表示部 8 0 2 に表示させた後、ステップ S 2 4 4 により、その表示状態でユーザ 2 0 9 の指示入力の待機状態となる。

30

【 0 0 8 1 】

以下、この繰り返しである。

【 0 0 8 2 】

図 1 に戻り、前記 C M 記録部 2 0 5 に記録されている各 C M の特徴ベクトル V は、信号 2 0 5 b として特徴比較部 2 0 4 に送られる。

40

【 0 0 8 3 】

特徴比較部 2 0 4 は、読み出し制御信号 2 0 4 a により C M 記録部 2 0 5 に記録されている各 C M の特徴ベクトル V を読み出す。当該特徴比較部 2 0 4 は、上記 C M 記録部 2 0 5 から読み出した各 C M の特徴ベクトル V_i を用いて、各 C M が同一の C M か否かの比較を行う。ここで、 i は各 C M を区別するためのインデックス（インデックスの変数）である。当該特徴比較部 2 0 4 における 2 つの C M の特徴ベクトル V_i と V_j の比較は次のようにして行う。

【 0 0 8 4 】

先ず、評価関数 $J(i, j)$ を、式 (9) のように計算する。

【 0 0 8 5 】

50

【数 6】

$$J(i, j) = \left| V_i - V_j \right|^2 = \sum_{n=0}^{N-1} \left| V_i[n] - V_j[n] \right|^2 \quad (9)$$

10

【0086】

続いて、 $J(i, j)$ を所定のスレッシュホールド J_{thsd} と比較し、 J_{thsd} 以下ならば同一のCMと判断し、 J_{thsd} より大きければ異なるCMと判断する。このような比較を、CM記録部205に記録されている全てのCMについて行い、同一と判断されたCMのうち、どちらかをCM記録部205から除去する。これにより、CM記録部205に記録されている信号のうち、同一のCMにかかる信号は除去されることとなる。

【0087】

ここで、図10には、上記図1の構成のうち、上述したCM記録部205、特徴量比較部204の部分のみを抜き出したサブセットを示しており、図11にはこれら図10に抜き出したサブセット部分の特に特徴比較部204の動作の流れをフローチャートとして示す。

20

【0088】

図11において、特徴比較部204は、先ず、ステップS250として、上記インデックスの変数 i を0にセットし、次に、ステップS251として、インデックスの変数 j を $i+1$ にセットする。

【0089】

続いて、特徴比較部204は、ステップS252として、特徴ベクトル V_i, V_j より評価関数 $J(i, j)$ を計算し、所定の閾値 J_{thsd} と比較する。ここで、当該比較の結果、閾値を下回っていると判定した場合(Yes)は、同一のCMと判断し、ステップS253として、特徴ベクトル V_j で表わされるCMをCM記憶部205より除去する。一方、閾値を下回っていないと判定した場合(No)は、異なるCMであると判断し次のステップS254に進む。

30

【0090】

ステップS254に進むと、特徴比較部204は、 j が対象とするCMの中で、最後のCM番号かどうかチェックする。ここで、最後のCMでないと判定した場合(No)は、ステップS255にて j をインクリメントし、続いてステップS252にて再び閾値との比較に戻る。一方、最後のCMであると判定した場合(Yes)は、ステップS256として、 i が対象とするCMの中で、最後のCMであるかどうかをチェックする。ここで、最後のCMではないと判定した場合(No)は、ステップS257として i をインクリメントし、続いてステップS251にて再び j のセットに戻る。最後のCMであった場合(Yes)は、処理を終了する。

40

【0091】

次に、本発明実施の形態の放送信号処理装置では、前述した図7の表示部802の表示画面例に更に検索アイコンを追加し、ユーザがCMの検索を行いたい場合に、当該検索アイコンを指示することで、ユーザ所望のCMを検索可能としている。図12には、検索アイコン815が追加表示された、表示部802の表示画面例を示す。以下、この図12の表示画面例を参照しながら前記CM記録部205、特徴量比較部204、CM閲覧部208、CM索引生成部206、CM選択部207等の動作を説明する。

【0092】

すなわち、上記検索アイコン815がユーザ209により指定されると、CM選択部20

50

7はそのユーザ選択に応じたユーザ選択情報207aを発生し、そのユーザ選択情報207をCM閲覧部208に送ると共に、特徴比較部204にも送る。

【0093】

特徴比較部204では、上記ユーザ選択情報207aが供給されると、前記CM索引画像又は映像810の中で既に選択されているCMとその特徴量が部分的に一致する特徴量を有するCMをCM記録部205内から検索する。

【0094】

すなわち、特徴比較部204では、CM選択部207からのユーザ選択情報207aにより、任意のCM*i*が指定された場合、そのCM*i*の特徴量とCM記録部205に記録されている他のCM*j*の特徴量との比較を行う。

10

【0095】

この比較の際には、まず、CMの部分区間の評価関数J*(i, j)*を、式(10)に従って計算することにより行う。

【0096】

【数7】

$$J'(i, j) = \sum_{n=N_s}^{N_e} |V_i[n] - V_j[n]|^2 \quad (10)$$

20

【0097】

但し、式中のN_sは、比較するCM部分区間の最初の処理フレーム番号、N_eは比較するCM部分区間の最後の処理フレーム番号である。

【0098】

続いて、特徴比較部204は、その評価関数J*(i, j)*を所定の閾値J_{thsd}と比較し、閾値より大きい場合には不一致と判断し、閾値より小さい場合には一致と判断する。

30

【0099】

なお、この比較において、例えば異なるCMであっても、同一提供者が提供するCMには、商品もしくは会社もしくは団体などに共通する映像/音声信号を挿入することで、視聴者に共通感をもたせるよう作成されている場合が多く存在する。また、そのための共通する映像/音声信号は、1秒程度であることが多い。従って、例えばN_e=N-1をCM区間の最後のフレーム番号とし、N_sをその1秒前の処理フレームとすることで、一つのCMを指定すると、その同一会社の提供するCMを全て検索することが可能となる。

【0100】

また、特徴量の比較の際の評価関数J*(i, j)*は、式(11)のように設定することもできる。

40

【0101】

【数8】

$$J'(i, j) = \text{Min}_s \sum_{n=s}^{(s+1)N_w} |V_i[n] - V_j[n]|^2 \quad (11)$$

【 0 1 0 2 】

10

但し、 N_w は一致検索を行う区間長であり、 s は部分一致検索区間をCM区間の最初から最後までシフトしてゆくインデックス番号であり、 $\text{Min}_s()$ は、全ての s のなかで最小となる値を表わす。またこの場合、ベクトル $V[n]$ のなかで、映像に関する特徴量 $H_l[n]$ 、 $H_c[n]$ 、 $B[n]$ 、 $D[n]$ 、 $C[n]$ を除き、音声に関する特徴量 $A[n]$ 、 $F[k;n]$ 、 $P[k;n]$ のみを用いて行うことも有効である。

【 0 1 0 3 】

このような関数を用いると、CM中の何れかの部分で一致するCMを検索することができる。これによると、例えば同じ音楽を背景音楽として用いるCMなどが検索できる。例えば同じ商品又は会社又は団体が、幾つかの継続性のあるCMを提供している場合、視聴者に共通性を感じさせるため同じ背景音楽が用いられることが多いことから、当該関数は、

20

【 0 1 0 4 】

なお、評価関数 $J(i, j)$ 及び $J'(i, j)$ は、2つのCMの類似度と称することもある。

【 0 1 0 5 】

上述のようにしてCM記録部205から検索されたCMの映像及び音声信号は、前記記録部再生信号205aとしてCM閲覧部208に送られる。

【 0 1 0 6 】

またこのとき、前記索引生成部206では、上記特徴量が一致するCMの索引情報を生成し、その索引情報206aをCM閲覧部208に送る。

【 0 1 0 7 】

30

これにより、CM閲覧部208では、上記検索されたCMの映像が表示され、また音声再生可能となる。

【 0 1 0 8 】

図13には、上記図1の構成のうち、上述したCM検索を実現するための、CM記録部205、特徴量比較部204、CM索引生成部206、CM閲覧部208、CM選択部207の部分のみを抜き出したサブセットを示しており、図14にはこれら図13に抜き出したサブセット部分の特に特徴比較部204の動作の流れをフローチャートとして示す。なお、図14のフローチャートは、前述した図9のフローチャートのステップS244の後段の処理となるものである。

【 0 1 0 9 】

40

この図14において、図9のフローチャートのステップS244のユーザ指示待機状態となっており、例えば上記ユーザ選択情報207aが、CM件サックアイコン815を指定するものであった場合、特徴比較部204は、ステップS260として、インデックスの変数 i を、上記選択されているCMのインデックスにセットし、次に、ステップS261として、インデックスの変数 j を0にセットする。

【 0 1 1 0 】

続いて、特徴比較部204は、ステップS262として、特徴ベクトル V_i 、 V_j より評価関数 $J(i, j)$ を計算し、所定の閾値 J_{thsd} と比較する。ここで、閾値 J_{thsd} を下回っていると判定した場合(Yes)には、類似するCMと判断され、ステップS263において、特徴ベクトル V_j で表されるCMをマークしておく。一方、閾値 J_{thsd} を下回

50

っていないと判定した場合（No）は、類似していないCMと判断して次のステップS294に進む。

【0111】

ステップS264に進むと、特徴比較部204は、jが対象とするCMの中で、最後のCM番号かどうかをチェックする。ここで、最後のCMではないと判定した場合（No）は、ステップS265にてjをインクリメントした後、ステップS262に戻り再び閾値との比較を行う。一方、最後のCMであった場合（Yes）には、ステップS266として、マークしたCM群をまとめて指定し、図9に示したステップS241、すなわち指定されたCM群の信号をCM記録部205より取得する処理へ戻る。

【0112】

なお、本実施の形態において、上述した同一CMの除去を行わず、全てのCMを蓄積し、かつ検索を全CM区間の一致にて行うことで、同一のCMが放送された回数及びその時刻を検索することも可能となる。

【0113】

以上説明したようなことから、本実施の形態の放送信号処理装置によれば、放送信号からCM部分だけを抽出し、蓄積することが可能となり、これによって、例えば、番組部分を記録せず、CM部分のみを蓄積するデータベース装置を提供することが可能となる。

【0114】

また、本実施の形態の放送信号処理装置によれば、放送信号からCM部分だけを抽出し、該CM部分の特徴付ける特徴量を抽出し、それらを蓄積することが可能となり、これによって、例えば、番組部分を記録せず、CM部分及びその特徴量のみを蓄積するデータベース装置を提供することが可能となる。これらの特徴量は、特に同一CM検索、類似CM検索などに有用なものである。

【0115】

また、本実施の形態の放送信号処理装置によれば、上述のように蓄積されたCMを、表示、閲覧することが可能となり、これによって、視聴者（ユーザ）は蓄積されたCMの一覧の表示、再生及び検索などを行うことができる。これは、CMを探して視聴する場合に有用である。

【0116】

また、本実施の形態の放送信号処理装置によれば、上述のように蓄積されたCMから、同一のCMを除去することが可能となり、これによって、蓄積容量が節約されると同時に、同一のCMを何度も閲覧してしまう手間を回避可能となる。

【0117】

また、本実施の形態の放送信号処理装置によれば、類似するCMの検索が可能となり、これによって、蓄積されたCMから、同一商品のCM、同一提供者のCMなどを、容易に検索、表示することが可能となる。

【0118】

また、本実施の形態の放送信号処理装置によれば、CMの最後の部分が一致するものを検索することができる。すなわち、CMの最後の部分には商品、会社、団体等に共通する映像や音声が含まれることが多いことから、これにより、同一商品や提供者のCMを容易に検索、表示可能となる。

【0119】

さらに、本実施の形態の放送信号処理装置によれば、背景音楽の一致するCMを検索することができる。すなわち、同一商品や提供者が、継続性をもって作成したCMには同一の背景音楽が用いられることが多いため、これにより、継続性をもって作成されたCMを、に検索、表示可能となる。

【0120】

最後に、本実施の形態の放送信号処理装置によれば、同一CMが放送された時刻及び回数を計測することができる。これは、例えばCMの作成者などが、放送予定と実際の放送を比較する等の目的に対して、有効な装置となる。

10

20

30

40

50

【 0 1 2 1 】

次に、本発明実施の形態において、放送信号からＣＭを検出し、ＣＭ開始時刻やＣＭ長さを検出する、図１のＣＭ検出部２０２の詳細について以下に説明する。

【 0 1 2 2 】

以下、ＣＭ検出部２０２において、例えばＴＶ放送信号からＣＭ部分を検出する際の原理について概説する。なお、以下の説明で、特徴量という記載は、前記ＣＭ特量抽出部２０３や特徴比較部２０４における特徴と同じ場合と異なる場合とがあるが、それらは共にＣＭの特徴を表すものであるため、ここではそれらを特に区別せずに使用している。

【 0 1 2 3 】

一般に、ＴＶ放送されるＣＭは、放送者の指定する規格に基づいて製作されるため、その「時間長（１つのＣＭの時間）はごく少数の種類に限定」される。例えば日本国内においては、特殊な例を除くほぼ全てのＣＭが、１５秒、３０秒、６０秒の長さで製作されている。

10

【 0 1 2 4 】

また、ＣＭの放送時には、番組本編や他のＣＭとは独立に製作されたものがＴＶ放送ストリーム中に挿入されるため、各ＣＭの前後では必然的に「音声レベルが下がる（すなわち小音量となる）」こと、及び、「映像信号が切り替わる」こと、という特徴を持つ。ここで、「音声レベルが下がる」とは、必ずしも無音（ここでは極微小なノイズしかない部分という意味）と同義ではない。すなわち、実際には、ＣＭと本編との切り替えのタイミングなどにより、必ずしも完全に無音とはならないまま切り替わることがあるからである。

20

【 0 1 2 5 】

上述したように、ＣＭについての「規定時間長（少数種類の時間長）」、「小音量」、「映像切り替わり」という３つの特徴は、ほぼ全てのＣＭが満たす条件である。本発明では、以下、これら３つの特徴に基づく条件を「必須条件」と称することにする。

【 0 1 2 6 】

したがって、ＴＶ放送信号から当該必須条件に対応する信号部分を検出するようにすれば、誤棄却がほとんどなく決定論的にＣＭの候補（すなわちＣＭであろうと思われる信号部分）を検出することが可能となる。但し、番組本編内にも、偶然そのような必須条件を満たしてしまう部分が多く存在するため、上記必須条件を用いただけでは、番組本編の一部をＣＭ候補として誤検出してしまいう虞が残る。

30

【 0 1 2 7 】

一方で、上記必須条件と比べて例外は多く存在するものの、ＣＭの性質上、多くのＣＭが満たすか若しくは一定の傾向を示す特徴としては、以下のようなものがある。

【 0 1 2 8 】

１）ＣＭの前後（ＣＭが開始される直前とＣＭが終了して本編番組が開始又は再開される直前）では、通常の番組本編内よりも音声レベルが低くなることが多い。

【 0 1 2 9 】

２）ＣＭと番組本編との間、及び、あるＣＭと他のＣＭとの間の、ほぼ無音となる区間長は、数百ミリ秒程度であることが多い。

【 0 1 3 0 】

３）ＴＶ放送内に含まれる有音区間は、ＣＭの規定時間長（１５秒、３０秒、６０秒等）より百ミリ秒程度以上短いことが多く、また１秒程度以上短いことは少ない。

40

【 0 1 3 1 】

４）ステレオ音声信号の左チャンネル（Ｌチャンネル）と右チャンネル（Ｒチャンネル）の相関値は、１より有意に小さいことが多い。

【 0 1 3 2 】

５）ＣＭ期間中は、番組本編より音量が大きめである傾向がある。

【 0 1 3 3 】

６）ＣＭの放送モードは、ステレオモードであることが多い。

【 0 1 3 4 】

50

7) C M 区間では、複数の C M が連続して放送されることが多い。

【 0 1 3 5 】

8) C M 期間中は、映像カットの切り替わり頻度が高いことが多い。

【 0 1 3 6 】

9) 逆に、C M 期間中であっても、カットの切り替わり頻度が極端に低いものがある（例えば静止画による C M など）。

【 0 1 3 7 】

10) 番組本編と C M との境界や、ある C M と他の C M の境界では、音質が大きく変化することが多い。

【 0 1 3 8 】

11) C M は、音声と音楽を同時に含むことが多い。

【 0 1 3 9 】

12) 番組編成上、毎時丁度の時刻近辺では、C M が放送される確率が高い。

【 0 1 4 0 】

13) 同様に、毎時 30 分付近でも C M が放送される確率が高い。

【 0 1 4 1 】

14) 番組のジャンルによって C M が放送される確率の高い時間帯がある（例えばサッカー中継のハーフタイムなど）。

【 0 1 4 2 】

本発明では、以下、これらの特徴に基づく条件を「付加条件」と称することにする。すなわち、当該付加条件は、C M が、規格に基づいて製作されるという制約、短い時間で宣伝効果を上げるためのものであるという制約、及び、番組構成上の都合などによる制約の元で製作された結果として、T V 放送信号上に現れてくることによる条件である。したがって、この付加条件は、決定論的な取り扱いができるほど確実な条件ではないものの、C M である可能性（C M らしさ）を評価する際の有効な条件となる。

【 0 1 4 3 】

さらに、T V 放送においては、同時に同じチャンネルで複数の映像及び音声が発送されることは物理的にありえないという特徴がある。すなわち、T V 放送信号から C M であろうと思われる信号部分（C M 候補）を検出しようとする場合において、例えば、T V 放送信号中に、上記付加条件を満たす複数の映像及び音声区間がオーバーラップして存在し、何らかの処理の結果、当該オーバーラップ区間で C M 候補が検出されたとしても、そのオーバーラップしている複数の映像及び音声内の少なくともどちらかの区間は、正しい C M 区間ではあり得ない。本発明では、T V 放送におけるこのような特徴に基づく条件を、「論理条件」と称することにする。

【 0 1 4 4 】

本発明では、以上説明した「必須条件」、「論理条件」、「付加条件」を合理的かつ効果的に利用することにより、T V 放送信号から高精度で C M 部分を検出可能としている。

【 0 1 4 5 】

より具体的に言うと、本発明では、「必須条件」に基づき、決定論的に T V 放送信号中から C M 候補（C M であろうと思われる信号部分）を抽出し、「付加条件」に基づく C M らしさ（C M である可能性）の統計論的な評価によって C M 候補を選択し、「論理条件」により C M 候補のオーバーラップ関係を解消することにより、精度の高い C M 検出を実現するものである。

【 0 1 4 6 】

図 15 には、上記 C M 検出部 202 の第 1 の具体例の詳細な構成を示す。当該 C M 検出部 202 は、大別して、フロントエンド部とバックエンド部とから構成されている。また、図中の動作制御部 23 は、放送チャンネルを示すチャンネル情報 1b に基づいて、明らかに C M が放送されない放送チャンネルであるか否か判断し、その判断結果に応じて、当該図 15 の各部における C M 検出動作を行わないように制御するものである。

【 0 1 4 7 】

10

20

30

40

50

先ず、図 15 のフロントエンド部から説明する。

【 0 1 4 8 】

この図 15 において、当該フロントエンド部には、前記入力信号 2 0 0 a に含まれる映像信号 2 a、音声信号 2 b、及び放送モード信号 2 c が入力される。また、フロントエンド部には、図示しない時計により発生した時間情報 3 a も入力される。

【 0 1 4 9 】

映像信号 2 a は、A / D 変換器 1 0 にてデジタル化され、フレームメモリ 1 1 に蓄えられる。なお、フレームメモリ 1 1 には、少なくとも 2 フレーム分の映像信号を蓄積可能なメモリである。当該フレームメモリ 1 1 からフレーム毎に読み出された映像信号は、カットチェンジ検出器 1 2 に送られる。

10

【 0 1 5 0 】

カットチェンジ検出器 1 2 は、フレームメモリ 1 1 より供給されたフレーム毎の映像信号に基づいて、映像が急激に変化するフレーム（以下、映像変化フレームと呼ぶ）と、輝度が一樣となるフレーム（以下、一樣輝度フレームと呼ぶ）を検出する。

【 0 1 5 1 】

すなわち、カットチェンジ検出器 1 2 は、フレームメモリ 1 1 に蓄えられた時間的に隣接する 2 つのフレーム映像間で、各画素毎に輝度の差分の自乗和を求め、当該自乗和が所定の閾値を越えた場合に、上記隣接する 2 つのフレームのうちの時間的に後のフレームを、上記映像が急激に変化する映像変化フレームとして検出する。また、カットチェンジ検出器 1 2 は、フレームメモリ 1 1 に蓄えられた各フレーム映像の輝度の分散を求め、その輝度の分散値が所定の閾値以下である場合に、そのフレームを一樣輝度フレームであるとして検出する。なお、フレームの間隔（NTSC 方式では約 3 0 m s）が、後述する音声信号処理において説明するフレーム周期と一致しない場合には、当該フレーム間隔を再離散化することによって、フレーム周期と一致させておくようにする。

20

【 0 1 5 2 】

以下、当該カットチェンジ検出器 1 2 における映像変化フレームと一樣輝度フレームの検出について、より具体的に説明する。

【 0 1 5 3 】

ここで、離散化された映像信号の横サイズを X、縦サイズを Y、縦横の画素番号を x, y とし、第 n フレームの映像を $I_n(x, y)$ 、当該第 n フレームに対して時間的に 1 フレーム前の第 n - 1 フレームの映像を $I_{n-1}(x, y)$ として表わすと、第 n フレームと第 n - 1 フレームの間の各画素毎の輝度差分の二乗和 D[n] は、式 (1 2) により得られ、また、第 n フレームの輝度分散値 V[n] は、式 (1 3) により得られる。

30

【 0 1 5 4 】

【 数 9 】

$$D[n] = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \left(I_n(x, y) - I_{n-1}(x, y) \right)^2 \quad (12)$$

$$V[n] = \frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} I_n^2(x, y) - \left(\frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} I_n(x, y) \right)^2 \quad (13)$$

40

【 0 1 5 5 】

また、このときのカットチェンジ検出器 1 2 の検出出力 C[n] は、式 (1 4) により表わされる。

【 0 1 5 6 】

【 数 1 0 】

$$C[n] = \begin{cases} 1 & \left(D[n] \geq D_{thsd} \text{ or } V[n] \leq V_{thsd} \right) \\ 0 & \left(D[n] < D_{thsd} \text{ and } V[n] > V_{thsd} \right) \end{cases} \quad (14)$$

【 0 1 5 7 】

10

ただし、式中の D_{thsd} は上記映像変化フレームを検出する際の前記自乗和に対する所定の閾値であり、 V_{thsd} は上記一様輝度フレームを検出する際の前記輝度の分散値に対する所定の閾値である。

【 0 1 5 8 】

当該カットチェンジ検出器 12 の検出出力 $C[n]$ は、映像信号についての特徴量として特徴量バッファ 18 へ送られる。

【 0 1 5 9 】

なお、上記の 2 つのフレーム映像間で輝度差分を求める際には、2 フレーム分の映像信号を蓄積可能なメモリが必要となり、また、2 フレーム分の映像信号に対する演算量も必要となる。そこで、例えばフレーム映像全面を同時に処理する代わりに、フレーム映像を適切な小ブロック毎に分け、その小ブロック毎に輝度差分を求めるようにしたり、或いは、フレーム映像間の各画素毎に輝度差分を求めるのではなく、各フレーム映像毎に輝度ヒストグラムを求めて、その輝度ヒストグラムのフレーム間差分を求めるようにしたり、又は、各フレーム映像毎に平均輝度を求めて、その平均輝度のフレーム間差分を求めるようにするで、メモリ容量や演算量を減らすことも可能である。逆に、メモリや演算量に余裕がある場合には、例えば、カラー映像における R (赤)、G (緑)、B (青) 成分のようなカラー成分毎に、上記輝度差分やカラーヒストグラム差分を求めることで、より検出精度を高めることも可能である。

20

【 0 1 6 0 】

次に、音声信号 2b は、A/D 変換器 13 にてデジタル化され、音声信号バッファ 14 に蓄えられる。なお、音声信号バッファ 14 には、少なくとも所定時間 T_1 (例えば 30 ms、以下、これを 1 フレーム長とする) 分の左 (L) 右 (R) 2 チャンネルのステレオ音声信号を蓄積可能なメモリである。当該音声信号バッファ 14 から読み出された音声信号は、振幅検出器 15、相関検出器 16、スペクトル検出器 17 に送られる。

30

【 0 1 6 1 】

振幅検出器 15 は、音声信号バッファ 14 に蓄えられた音声信号を用いて、所定の時間 T_2 (例えば 15 ms、以下、これを 1 フレーム周期とする) 毎の短時間平均二乗振幅を検出する。すなわち、振幅検出器 15 は、音声信号バッファ 14 に左右 2 チャンネルのステレオ音声信号が蓄積されている場合、当該音声信号バッファ 14 より読み出された左右 2 チャンネルのステレオ音声信号 $S_L[m]$ 、 $S_R[m]$ から、所定の時間 T_2 (15 ms、1 フレーム周期) 毎に、短時間平均二乗振幅を検出する。なお、上記 m ($m = 0, \dots, M-1$) は、離散化された時間を表わすバッファ内のサンプル番号であり、最大番号 M が 1 フレーム長 T_1 に対応する。

40

【 0 1 6 2 】

より具体的に説明すると、振幅検出器 15 は、第 n フレームにおける左右 2 チャンネルの音声信号の平均二乗振幅 $A[n]$ を式 (15) により計算する。

【 0 1 6 3 】

【 数 1 1 】

$$A[n] = \frac{1}{4M} \sum_{m=0}^{M-1} \left(S_L[m + nT_2] + S_R[m + nT_2] \right)^2 \quad (15)$$

【 0 1 6 4 】

10

当該振幅検出器 1 5 の検出出力である平均二乗振幅 $A[n]$ は、音声信号についての特徴量の一つとして特徴量バッファ 1 8 へ送られる。

【 0 1 6 5 】

相関検出器 1 6 は、音声信号バッファ 1 4 に蓄えられた音声信号を用いて、1 フレーム毎の音声信号について規格化前の相関係数を検出すると共に、後段にて行われる規格化のための短時間エネルギーも同時に検出する。すなわち、相関検出器 1 6 は、音声信号バッファ 1 4 に左右 2 チャンネルのステレオ音声信号が蓄積されている場合、当該音声信号バッファ 1 4 より読み出された左右 2 チャンネルのステレオ音声信号 $S_L[m]$ 、 $S_R[m]$ から、1 フレーム毎の左右 2 チャンネルの音声信号について規格化前の相関係数を検出すると共に、後段にて行われる規格化のための短時間エネルギーも同時に検出する。

20

【 0 1 6 6 】

より具体的に説明すると、相関検出器 1 6 は、第 n フレームにおける左右 2 チャンネルの音声信号の相関係数 $A_{LR}[n]$ を式 (1 6) により計算し、左チャンネルの音声信号エネルギー $A_{LL}[n]$ を式 (1 7) により計算し、右チャンネルの音声信号エネルギー $A_{RR}[n]$ を式 (1 8) により計算する。

【 0 1 6 7 】

【 数 1 2 】

$$A_{LR}[n] = \sum_{m=0}^{M-1} S_L[m + nT_2] S_R[m + nT_2] \quad (16)$$

30

$$A_{LL}[n] = \sum_{m=0}^{M-1} S_L^2[m + nT_2] \quad (17)$$

$$A_{RR}[n] = \sum_{m=0}^{M-1} S_R^2[m + nT_2] \quad (18)$$

【 0 1 6 8 】

当該相関検出器 1 6 の検出出力である相関係数 $A_{LR}[n]$ と音声信号エネルギー $A_{LL}[n]$ 、 $A_{RR}[n]$ は、それぞれが音声信号についての特徴量の一つとして特徴量バッファ 1 8 へ送られる。

40

【 0 1 6 9 】

スペクトル検出器 1 7 は、音声信号バッファ 1 4 に蓄えられた音声信号を用いて、短時間スペクトルを計算する。すなわち、スペクトル検出器 1 7 は、音声信号バッファ 1 4 に左右 2 チャンネルのステレオ音声信号が蓄積されている場合、当該音声信号バッファ 1 4 より読み出された左右 2 チャンネルのステレオ音声信号 $S_L[m]$ 、 $S_R[m]$ から、短時間スペクトルを計算する。

【 0 1 7 0 】

より具体的に説明すると、スペクトル検出器 1 7 は、第 n フレームにおける左右 2 チャンネルの音声信号の離散スペクトル $F[k;n]$ を求める。なお、 $k = 0, \dots, K-1$ を離散化された周波数を表わす番号とすると、離散スペクトル $F[k;n]$ は式 (1 9) により表

50

わされる。

【 0 1 7 1 】

【 数 1 3 】

$$F[k;n] = \left| \sum_{m=0}^{M-1} (S_L[m] + S_R[m]) e^{2\pi j m k / M} \right|^2 \quad (19)$$

10

【 0 1 7 2 】

この式 (1 9) の演算は、例えば、高速フーリエ変換 (F F T) 又は線形予測分析 (L P C) などを用いて実現される。

【 0 1 7 3 】

当該スペクトル検出器 1 7 の計算出力である短時間離散スペクトル $F[k;n]$ は、音声信号についての特徴量の一つとして特徴量バッファ 1 8 へ送られる。

【 0 1 7 4 】

次に、放送モード信号 2 c は、上述した音声信号処理のフレームに合わせて離散化された数値となされる。

20

【 0 1 7 5 】

より具体的に説明すると、第 n フレームにおける放送モード信号 2 c は、例えば式 (2 0) のような数値 $B[n]$ となされる。

【 0 1 7 6 】

【 数 1 4 】

$$B[n] = \begin{cases} 0 & (\text{モノラルモード}) \\ 1 & (\text{ステレオモード}) \\ 2 & (\text{音声多重モード}) \end{cases} \quad (20)$$

30

【 0 1 7 7 】

この放送モード信号 2 c を離散化した数値 $B[n]$ は、T V 放送信号の特徴量の一つとして特徴量バッファ 1 8 へ送られる。

【 0 1 7 8 】

同様に、時間信号 3 a も、音声信号処理のフレームに合わせて離散化された数値 $T[n]$ となされ、特徴量の一つとして特徴量バッファ 1 8 へ送られる。

40

【 0 1 7 9 】

特徴量バッファ 1 8 は、上記カットチェンジ検出器 1 2 からの検出出力 $C[n]$ と、振幅検出器 1 5 からの平均二乗振幅 $A[n]$ と、相関検出器 1 6 からの相関係数 $A_{LR}[n]$ 、音声信号エネルギー $A_{LL}[n]$ 、 $A_{RR}[n]$ と、スペクトル検出器 1 7 からの短時間離散スペクトル $F[k;n]$ と、放送モード信号 2 c の離散化数値 $B[n]$ と、時間信号 3 a の離散化数値 $T[n]$ とからなる、式 (2 1) に示される特徴量 $G[n]$ を、所定の時間 T_3 に渡って蓄積する。なお、時間 T_3 は C M 部分を最低でも 1 つ以上に渡って記憶できる時間であり、例えば 8 0 秒などとする。

【 0 1 8 0 】

50

$G[n] = \{C[n], A[n], A_{LR}[n], A_{LL}[n], A_{RR}[n], F[k;n], B[n], T[n]\}$ (21) 以上の A/D 変換器 10 から特徴量バッファ 18 までが、図 15 に示した CM 検出部 202 のフロントエンド部の構成であり、以下、図 16、図 17 のフローチャートを用いて当該フロントエンド部における処理の流れを説明する。なお、図 16 のステップ S30 ~ S32 までは映像信号 2a についての処理の流れを表しており、図 17 のステップ S33 ~ S40 までは音声信号 2b 及び放送モード信号 2c、時間信号 3a についての処理の流れを表している。

【0181】

まず、映像信号 2a についての処理の流れを表す図 16 において、フロントエンド部は、ステップ S30 の処理として、A/D 変換器 10 によりデジタル化された、少なくとも 1 フレーム分の映像信号 2a をフレームメモリ 11 に蓄える。このフレームメモリ 11 は、1 フレーム分の映像信号 2a を 1 サンプルとして扱うようになされており、1 フレーム分の映像信号 2a が入力されると、当該フレームメモリ 11 内に既に蓄積されている映像信号 2a が 1 フレーム分シフトし、最も時間的に過去に入力された 1 フレームの映像信号 2a が押し出されて出力されるようになっている。

10

【0182】

次に、フロントエンド部は、ステップ S31 の処理として、フレームメモリ 11 から映像信号 2a を読み出してカットチェンジ検出器 12 に送り、前述のようにして検出出力 $C[n]$ を求める。

【0183】

その後、フロントエンド部は、ステップ S32 の処理として、当該検出出力 $C[n]$ を特徴量バッファ 18 に蓄える。

20

【0184】

一方、音声信号 2b についての処理の流れを表す図 17 において、フロントエンド部は、ステップ S33 及びステップ S34 の処理として、A/D 変換器 13 によりデジタル化された、音声信号 2b を音声信号バッファ 14 に入力すると共に、当該音声信号バッファ 14 に少なくとも 1 フレーム周期 T_2 分の音声信号 2b を蓄積する。この音声バッファ 14 は、1 フレーム周期 T_2 分の音声信号 2b を 1 サンプルとして扱うようになされており、1 フレーム周期 T_2 分の音声信号 2b が入力されると、当該音声バッファ 14 内に既に蓄積されている音声信号 2b が 1 フレーム周期 T_2 分だけシフトし、最も時間的に過去に入力された 1 フレーム周期 T_2 分の音声信号 2b が押し出されて出力されるようになっている。

30

【0185】

上記音声信号バッファ 14 に少なくとも 1 フレーム周期 T_2 分の音声信号 2b が蓄積されると、フロントエンド部は、ステップ S35 の処理として、当該音声信号バッファ 14 に蓄積された音声信号 2b を読み出して振幅検出器 15 に送り、前述のようにして、平均二乗振幅 $A[n]$ を求める。

【0186】

同時に、フロントエンド部は、ステップ S36 の処理として、音声信号バッファ 14 に蓄積された音声信号 2b を相関検出器 16 に送り、前述のようにして、相関係数 $A_{LR}[n]$ と音声信号エネルギー $A_{LL}[n]$ 、 $A_{RR}[n]$ を求める。

40

【0187】

また同時に、フロントエンド部は、ステップ S37 の処理として、音声信号バッファ 14 に蓄積された音声信号 2b をスペクトル検出器 17 に送り、前述のようにして、短時間離散スペクトル $F[k;n]$ を求める。

【0188】

さらに、フロントエンド部は、ステップ S38 の処理として、放送モード信号 2c から離散化した数値 $B[n]$ を求めると共に、時間信号 3a から、離散化された数値 $T[n]$ を求める。

【0189】

50

フロントエンド部は、以上のようにして求められた、上記カットチェンジ検出器 12 からの検出出力 $C[n]$ と、振幅検出器 15 からの平均二乗振幅 $A[n]$ と、相関検出器 16 からの相関係数 $A_{LR}[n]$ 、音声信号エネルギー $A_{LL}[n]$ 、 $A_{RR}[n]$ と、スペクトル検出器 17 からの短時間離散スペクトル $F[k;n]$ と、放送モード信号 2c の離散化数値 $B[n]$ と、時間信号 3a の離散化数値 $T[n]$ とからなる特徴量 $G[n]$ を、特徴量バッファ 18 に蓄積する。

【0190】

図 15 に戻り、バックエンド部の説明を行う。なお、以下の説明において、番号 n は、特徴量バッファ 18 内にフレーム毎に蓄積される特徴量の、各フレーム番号を表わすものとする。また、最新のフレームの特徴量を $G[0]$ とし、過去のフレームの特徴量となるにしたがって n の値が増加し、新たなフレームの特徴量が入力された場合には、全てのデータが 1 ずつシフト（フレーム番号が 1 ずつシフト）するものとする。

10

【0191】

図 15 において、特徴量バッファ 18 に蓄積された特徴量は、フレーム毎に CM 候補検出器 19 に送られる。

【0192】

当該 CM 候補検出器 19 は、ほぼ全ての CM が満たす、前述した「必須条件」に基づき、フレーム毎に CM 区間の候補を算出する。ここで、必須条件とは、前述したように、CM の音声信号が「小音量」であること、すなわち音声信号の音量が所定の閾値以下となっているフレーム（以下、音量条件と呼ぶ）であり、且つ、CM の「映像切り替わり」があること、すなわち映像信号が急激に変換するフレーム又は一様な輝度となるフレーム（以下、映像条件と呼ぶ）であり、さらに、「規定時間長（少数種類の時間長）」であること、すなわち上記音量条件と映像条件を満たす 2 つのフレームの間隔が所定の CM 長と合致する区間（以下、時間条件と呼ぶ）となるような条件であり、具体的には、前述の特徴量を用いて、以下のような式（22）で且つ式（23）で且つ式（24）の条件として書き下すことができる。

20

【0193】

$$A[0] < A_{thsd} \quad (22)$$

$$C[0] = 1 \quad (23)$$

$$A[n_1] < A_{thsd}, C[n_1] = 1 \text{ 又は } A[n_2] < A_{thsd}, C[n_2] = 1 \text{ 又は}$$

$$A[n_3] < A_{thsd}, C[n_3] = 1 \quad (24)$$

30

ただし、 A_{thsd} は所定の二乗振幅の閾値であり、 n_1, n_2, n_3 はそれぞれ CM 長として規定されている時間長（本実施の形態では、一例として 15 秒、30 秒、60 秒の 3 種類の時間長がある場合を説明に用いている）を、フレーム周期単位に換算した数である。なお、CM の実際の放送時間には誤差があるため、実用上は、 n_1, n_2, n_3 にはそれぞれ多少の幅を持たせる。

【0194】

ここで、図 18 を用いて、上記 CM 候補検出器 19 の動作の流れを説明する。

【0195】

図 18 において、特徴量バッファ 18 では、ステップ S50 のバッファシフト処理とステップ S51 の特徴量入力処理として、図 16 のステップ S32 で説明したフレームメモリと図 17 のステップ S40 で説明した音声信号バッファと同様に、1 フレーム単位の入力、シフト及び出力の動作を行うようになされている。すなわち、特徴量バッファ 18 は、1 フレーム分の特徴量を 1 サンプルとして扱うようになされており、1 フレーム分の特徴量が入力されると、当該特徴量バッファ 18 内に既に蓄積されている特徴量が 1 フレーム分だけシフトし、最も時間的に過去に入力された 1 フレーム分の特徴量が押し出されて出力されるようになっている。

40

【0196】

上記ステップ S50 及びステップ S51 の処理により、特徴量バッファ 18 から 1 フレーム（1 サンプル）分の特徴量が入力されると、CM 候補検出器 19 は、ステップ S52 及びステップ S53 の処理として、1 フレーム（サンプル）に特徴量が上記必須条件の音量

50

条件、映像条件、時間条件を満たすか否かの評価を行う。すなわち。CM候補検索器19は、ステップS52において、先ず最初のフレームの平均二乗振幅 $A[0]$ と所定の二乗振幅の閾値 A_{thsd} を比較し、次に、ステップS53の処理として、前記検出力 $C[0]$ が1となるか否か調べることにより、当該フレームが上記必須条件である音量条件、映像条件、時間条件を満たすか否かの判定を行う。CM候補検索器19では、これらステップS52, S53の判定処理の結果、上記平均二乗振幅 $A[0]$ が所定の二乗振幅の閾値 A_{thsd} を超えず、且つ、上記必須条件を満たしていると判定した場合、当該フレームをCM候補としてステップS57以降(ステップS54~S56については後述する)の処理に進み、逆に、上記平均二乗振幅 $A[0]$ が所定の二乗振幅の閾値 A_{thsd} を超えたか、或いは上記必須条件を満たしていないと判定した場合、当該フレームがCM候補にはならないとしてステップS50の処理に戻る。

10

【0197】

上記ステップS52, S53の各判定処理の結果、上記平均二乗振幅 $A[0]$ が所定の二乗振幅の閾値 A_{thsd} を超えず、且つ、上記必須条件を満たしている場合、CM候補検索器19は、ステップS57の処理としてCM開始フレーム n_s を検索し、次に、ステップS58の処理としてCM終了フレーム n_e の検索を行い、更に、ステップS59の処理としてCM開始時刻 T_s を計算し、ステップS60としてCM長さ W を計算する。

【0198】

CM候補検索器19は、以上のステップS57~S60の検索及び計算を行った後、ステップS61において後述するCM候補テーブルを参照し、もし、CM開始時刻 T_s 及びCM長さ T_w の一致する候補がすでに当該CM候補テーブル中に存在するならば、そのまま再びステップS54~S56の処理に戻り、逆に存在しない場合には、新たなCM候補としてCM候補テーブルに追加した後、再びステップS54~S56の処理に戻る。

20

【0199】

ステップS54~S56では、全ての時間長に対して上述同様の処理を行った後、ステップS50に戻り、次の入力に対して同じ処理を繰り返すことを表している。

【0200】

なお、上記CM開始フレーム n_s とは、 n_1, n_2, n_3 で表される各フレームのうち時間条件に合致したフレームから、最新フレームの方向へ向かって、平均二乗振幅 $A[n]$ が二乗振幅の閾値 A_{thsd} を越える最初のフレーム番号である。また、CM終了フレーム n_e とは、0番目のフレームより過去の方

30

向に、平均二乗振幅 $A[n]$ が二乗振幅の閾値 A_{thsd} を越えない最後のフレーム番号である。さらにCM開始時刻 T_s は、CM開始フレーム番号 n_s を用いて $T_s = T[n_s]$ として求められる。同様にCM長さ T_w は、 $T_w = T[n_e] - T[n_s]$ として求められる。

【0201】

ここで、図19には、上記必須条件の算出例を示す。この図19に示す $A[n]$ の項において、「o」は二乗振幅の閾値 A_{thsd} 未達の平均二乗振幅を持つフレームを示し、「x」は二乗振幅の閾値 A_{thsd} 以上の平均二乗振幅を持つフレームを示している。この例では、 $A[0], C[0]$ 及び $A[n_1], C[n_1]$ が条件を満たし、 n_1 より左方で最初に $A[n] = x$ となるフレームが n_s 、0より右方に連続する最後の $A[n] = o$ となるフレームが n_e となる。

40

【0202】

以上の処理により、CM候補検出器19では、1フレーム(1サンプル)の特徴量が入力される毎にCM候補の検出を行い、CM候補が検出された場合にはCM候補テーブルにエントリーする。

【0203】

図20には、CM候補テーブルの構成例を示す。この図20において、CM候補テーブルの項目は、開始時刻 T_s 、長さ T_w 、及び後述する付加条件算出器20で算出する特徴量 $Q_1 \sim Q_{11}$ 、及び後述する付加条件判定器21で算出するスコア R とスコア判定結果 Z からなる。CM候補検出器19によるCM候補テーブル19aの段階では、開始時刻 T_s 、長さ T_w のみが記述される。このように、CM候補テーブルは、CM候補検出器19で得ら

50

れるCM開始時刻 T_s 、長さ T_w と、付加条件算出器20で算出される特徴量 $Q_1 \sim Q_{11}$ と、付加条件判定器21で算出されるスコア R 及びスコア判定結果 Z とを記述し、それら特徴量を管理するための表である。また、CM候補テーブルは、そのエントリーがCMであるかないかの判定を受けるまで保持され、CMであると判断された場合には、後述するルール判定器22からCM検出出力4aとして出力され、CMでないと判断された場合には破棄される。

【0204】

上記CM候補検出器19により開始時刻 T_s 、長さ T_w のみが記述されたCM候補テーブル19aは、付加条件算出器20に送られる。

【0205】

付加条件算出器20では、CM候補テーブル19aにエントリーされた候補区間より、特徴量バッファ18を参照しながら、以下に示すような特徴量 $Q_1 \sim Q_{11}$ を抽出し、それをCM候補テーブル19aに追加記述し、CM候補テーブル20aとして付加条件判定器21に出力する。

【0206】

図21には、当該付加条件算出器20における特徴量 $Q_1 \sim Q_{11}$ の算出例を示す。

【0207】

この図21において、横軸はフレーム番号（離散時間に相当）を表し、図21の（a）はカットチェンジ検出出力 $C[n]$ 、図21の（b）は放送モード信号2cの離散化数値 $B[n]$ 、図21の（c）は音声信号の短時間離散スペクトル $S[k,n]$ 、図21の（d）は音声信号の平均二乗振幅 $A[n]$ を表わし、 n_1 の間隔（図中点線で挟まれた区間）がCM候補である。なお、図21の（a）において、図中CTで示す位置はカットチェンジ検出出力 $C[n]$ が1となっている位置（すなわちカットチェンジが検出された位置）を示している。また、図21の（b）において、図中Mで示す区間はその区間が何らかの放送モードとなっていることを示している。図21の（c）において、図中 S_1, S_2, S_3, S_4 は何らかのスペクトル成分が存在することを示し、図21の（d）において、図中AMは二乗振幅の変化を表している。また、図中 Q_1 から Q_{11} は、上記付加条件算出器20にて特徴量 $Q_1 \sim Q_{11}$ が計算される場所を示している。

【0208】

以下、付加条件算出器20で算出される各特徴量 $Q_1 \sim Q_{11}$ について個々に説明する。

【0209】

特徴量 Q_1 は前ブレーク長である。当該前ブレーク長とは、CM候補区間直前の小音量区間（前ブレーク区間と称する）、すなわち連続して $A[n]$ が所定の閾値 A_{thsd} 以下である時間長であり、図21中の一点鎖線で挟まれた区間長BBが前ブレーク長 Q_1 である。

【0210】

特徴量 Q_2 は後ブレーク長である。当該後ブレーク長とは、CM候補区間直後の小音量区間（後ブレーク区間と称する）、すなわち連続して $A[n]$ が所定の閾値 A_{thsd} 以下である時間長であり、図21中の一点鎖線で挟まれた区間長ABが後ブレーク長 Q_2 である。

【0211】

特徴量 Q_3 は前ブレーク最小振幅である。当該前ブレーク最小振幅 Q_3 は、前記の前ブレーク区間における $A[n]$ の最小値である。

【0212】

特徴量 Q_4 は後ブレーク最小振幅である。当該後ブレーク最小振幅 Q_4 は、前記の後ブレーク区間における $A[n]$ の最小値である。

【0213】

特徴量 Q_5 は左右相関値である。当該左右相関値 Q_5 は、CM候補区間の音声の左右2チャンネルの音声信号 $S_L[m]$ 、 $S_R[m]$ の相関値である。これは、式(16)～式(18)の $A_{LR}[n]$ 、 $A_{LL}[n]$ 、 $A_{RR}[n]$ を利用して、式(25)に従って算出することができる。

【0214】

【数15】

10

20

30

40

50

$$Q_5 = \frac{\sum_{n=n_s}^{n_e-1} A_{LR}[n]}{\sum_{n=n_s}^{n_e-1} A_{LL}[n] \sum_{n=n_s}^{n_e-1} A_{RR}[n]} \quad (25)$$

【 0 2 1 5 】

10

この式 (2 5) の演算では、フレームのオーバーラップにより原波形が部分的に複数回加算されることになるが、そのことはこのシステムに実質的な影響は及ぼさない。また、原波形をそのまま保持できるだけのメモリ容量及び処理速度がある場合には、この演算は原波形の相互相関と置き換えることもできる。

【 0 2 1 6 】

特徴量 Q_6 は平均振幅値である。当該平均振幅値 Q_6 は、C M 候補区間の音声信号の振幅の R M S 値 (平均二乗振幅) である。これは、式 (2 6) により計算することができる。

【 0 2 1 7 】

【 数 1 6 】

20

$$Q_6 = \sqrt{\frac{1}{n_e - n_s} \sum_{n=n_s}^{n_e-1} A[n]^2} \quad (26)$$

【 0 2 1 8 】

30

この式 (2 6) の演算では、上記左右相関演算の場合と同様に、フレームのオーバーラップ次第では原波形が部分的に複数回加算されることになるが、そのことは実質的な影響を及ぼさない。また、原波形をそのまま保持できるだけのメモリ容量及び処理速度がある場合には、この演算は原波形の R M S 演算と置き換えることもできる。

【 0 2 1 9 】

特徴量 Q_7 はカット数である。当該カット数 Q_7 は、C M 候補区間中に存在するカットチェンジの回数 (上記 C T の数) を数える演算となる。すなわちこれは、 $[n_s, n_e)$ の区間で $C[n] = 1$ となる回数を数える演算となる。

【 0 2 2 0 】

特徴量 Q_8 は放送モードである。ここでの放送モードは、C M 候補区間中で最も支配的な放送モードのことである。これは、 $[n_s, n_e)$ の区間の $B[n]$ 値の中で、最も頻発する放送モード Q_8 を選ぶ演算である。

40

【 0 2 2 1 】

特徴量 Q_9 は隣接候補数である。当該隣接候補数 Q_9 は、ある C M 候補に対して、その前後にある有音区間も C M 候補であるかどうかを表わし、両側とも C M 候補であれば「 2 」、片側のみ C M 候補であれば「 1 」、どちらも C M 候補でなければ「 0 」の値をとる。この演算は、C M 候補テーブルを検索することで行われ、開始時刻 T_s と長さ T_w と後ブレイク長 Q_2 の和 ($T_s + T_w + Q_2$) が、他の C M 候補の開始時刻 (T_s) と一致するかどうかで後側候補の判定が行われる。同様に、開始時刻 T_s と前ブレイク長 Q_1 の差 ($T_s - Q_1$) が、他の C M 候補の開始時刻 T_s と長さ T_w の和 ($T_s + T_w$) と一致するかどうかで、前側候補の判定が行われる。

50

【 0 2 2 2 】

特徴量 Q_{10} 、 Q_{11} はスペクトル差分エネルギーである。当該スペクトル差分エネルギー Q_{10} 、 Q_{11} は、CM と番組本編や CM と他の CM との境界での音質変化を定量化するために用いられる。これは、上記境界の両側における平均スペクトルの差の自乗和として定義され、式 (27) ~ (32) に従って計算される。

【 0 2 2 3 】

【 数 1 7 】

$$S_1[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n'_e - n] \quad (27)$$

10

$$S_2[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n_s + n] \quad (28)$$

$$S_3[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n_e - n] \quad (29)$$

$$S_4[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n'_s + n] \quad (30)$$

$$Q_{10} = \frac{1}{S_{norm}^2} \sum_k \left(S_2[k] - S_1[k] \right)^2 \quad (31)$$

$$Q_{10} = \frac{1}{S_{norm}^2} \sum_k \left(S_4[k] - S_3[k] \right)^2 \quad (32)$$

20

【 0 2 2 4 】

但し、式中の N はスペクトルの平均をとるフレーム数、 n_e は CM 候補区間の直前の有音区間の終了フレーム番号 (図 21 参照)、 n_s は CM 候補区間の直後の有音区間の開始フレーム番号、 $S_1[k]$ は CM 候補区間の直前の有音区間の終了直前の平均スペクトラム、 $S_2[k]$ は CM 候補区間開始直後の平均スペクトラム、 $S_3[k]$ は CM 候補区間終了直前の平均スペクトラム、 $S_4[k]$ は CM 候補区間の直後の有音区間の開始直後の平均スペクトラム、 S_{norm} は適切な規格化定数である。

【 0 2 2 5 】

30

上記付加条件算出器 20 は、以上により算出した Q_1 から Q_{11} までの特徴量を、CM 候補テーブル 19a に追加記述し、CM 候補テーブル 20a として出力する。当該 CM 候補テーブル 20a は、付加条件判定器 21 に送られる。

【 0 2 2 6 】

付加条件判定器 21 は、CM 候補テーブル 20a を入力とし、CM 候補の特徴量を、閾値関数などにより非線型にパラメータ変換した後、荷重加算することで CM 候補に対するスコア R を算出し、 R が所定の閾値以上である場合には有力 CM 候補と判定する。付加条件判定器 21 は、これらスコア R とスコア判定結果 Z を CM 候補テーブル 20a に追加記述し、CM 候補テーブル 21a として出力する。

【 0 2 2 7 】

40

図 22 には、付加条件判定器 21 の概略構成を示す。

【 0 2 2 8 】

この図 22 において、CM 候補テーブル 21a の各特徴量 $Q_1 \sim Q_L$ は、それぞれ対応する関数演算器 50₁ ~ 50_L に送られ、それぞれ対応するパラメータ変換関数 $H_1() \sim H_L()$ による変換演算が施された後、さらにそれぞれ対応する重み付け器 51₁ ~ 51_L により重み $W_1 \sim W_L$ との積がとられる。各重み付け器 51₁ ~ 51_L により重み付けがなされた後の特徴量は、総和加算器 52 での総和加算によりスコア R が算出される。この総和加算器 52 から出力されたスコア R は、スコア判定器 53 にて所定の閾値と比較され、スコア R が所定の閾値以上である場合には有力 CM 候補である旨を示す判定結果が出力される。なお、スコア判定器 53 によるスコア判定により所定の閾値未満であると判定された CM 候補は

50

、テーブルから消去される。

【 0 2 2 9 】

より具体的に説明すると、当該付加条件判定器 2 1 におけるスコア算出演算は、式 (3 3) に従って行われる。

【 0 2 3 0 】

【 数 1 8 】

$$R = \sum_{l=1}^L W_l H_l(Q_l) \quad (33)$$

10

【 0 2 3 1 】

ただし、 $H_l()$ は各特徴量に対して予め定めるパラメータ変換関数、 W_l は予め決定しておく荷重、 L は特徴量数 (= 1 1) である。なお、 L は 1 ~ 1 1 のうちの任意の数である。

【 0 2 3 2 】

ここで、各関数演算器 5 0 1 ~ 5 0 L におけるパラメータ変換関数 $H_l()$ は、最も簡単には矩形関数でよい。すなわち例えば、図 2 3 の (b) に示すような矩形関数 $\text{Rect}(x; t_1, t_2)$ を用い、予め各特徴量について決定しておく標準値の下上限値を t_{1l} , t_{2l} とし、式 (3 4) により例えば Q_l が標準値の範囲内ならば 1、範囲外ならば 0 となるようにする。

【 0 2 3 3 】

20

$$H_l(Q_l) = \text{Rect}(Q_l; t_{1l}, t_{2l}) \quad (34)$$

なお、前記境界付近で滑らかに 0 から 1、1 から 0 への推移させるようにする場合には、例えば式 (3 5) のような、シグモイド関数 $\text{Sigm}(x; t_1, t_2)$ を用いることもできる。

【 0 2 3 4 】

【 数 1 9 】

$$H_l(Q_l) = \text{Sigm}(Q_l; t_{1l}, t_{2l}) = \frac{1}{1 + \exp\left(-\left(x - t_{1l}\right) / \sigma_{1l}\right)} \cdot \frac{1}{1 + \exp\left(\left(x - t_{2l}\right) / \sigma_{2l}\right)}$$

(35)

30

【 0 2 3 5 】

図 2 3 の (c) にはその概形を示す。ただし、 t_{1l} , t_{2l} は推移の程度を表わす定数であり、予め特徴量の分布などに基づき決定しておく。

【 0 2 3 6 】

また、上記各重み付け器 5 1 1 ~ 5 1 L による加算荷重 W_l は、予め特徴量の統計的性質に基づき人為的に決定しておくこともできるが、既知の学習サンプルに対して、ニューラルネットワーク (例えば中川著「パターン情報処理」丸善 (1999) などに詳説) の要領で学習することで、自動的に荷重を決定することも可能である。なお、 L は 1 ~ 1 1 のうちの

40

【 0 2 3 7 】

さらに、上記判定器 5 3 におけるスコア判定は、式 (3 6) のように、スコア R の閾値処理により行う。

【 0 2 3 8 】

$$Z = \text{Unit}(R - t_r) \quad (36)$$

ただし、 $\text{Unit}(x)$ は、図 2 3 の (a) に示すように、 $x > 0$ で 1、 $x < 0$ で 0 となる単位ステップ関数であり、 t_r は予め定めるか或いは学習により自動的に決まる判定閾値である。

【 0 2 3 9 】

50

次に、ルール判定器 22 は、上記付加条件判定器 21 でのスコア判定により得られた CM 候補テーブル 21a を入力とし、後述するような所定のルール判定により m 最終的な CM 検出出力 4a として CM 開始時刻と長さを出力する。すなわち、当該ルール判定器 22 では、同一時刻に複数の CM 候補があった場合（以下、競合関係という）に、どちらがより CM として確からしいかをルール処理により判定する。

【0240】

以下、ルール判定器 22 の動作を図 24 のフローチャートを用いて説明する。

【0241】

まず、ルール判定器 22 は、ステップ S70 として、CM 候補テーブルより、判定するべく CM 候補を選択する。これは、CM 候補テーブル中で最古の候補であり、予め設定された時間 T_4 が経過したものから順に行う。 T_4 は、数個の CM が十分含まれる程度の時間長であり、例えば 150 秒間とする。

10

【0242】

続いて、ルール判定器 22 は、ステップ S71 として、選択した CM 候補の区間中（ T_s から $T_s + T_w$ までの間）に、他の CM 候補が存在するかどうか、CM 候補テーブル中を検索する。このステップ S71 において、存在しないと判定した場合（No）、この CM 候補は CM 検出出力として出力され、CM 候補テーブルより消去される。

【0243】

一方、ステップ S71 において、存在すると判定した場合（Yes）、それらは競合関係にあるとして、ステップ S72 にてまず最小長さ優先ルールが適用される。ここで、最小長さ優先ルールは、ある時区間が、複数の長さの異なる CM 候補の組み合わせにより構成され得る場合、より短い CM 候補で構成されている方を優先するというルールである。すなわち、例えば 30 秒という時区間に対して、1 つの 30 秒 CM という可能性と、2 つの 15 秒 CM の組み合わせという可能性の、両方が同時に候補として存在する場合には、15 秒 CM を選択し、30 秒 CM を棄却するというルールである。

20

【0244】

図 25 を用いて、この最小長さ優先ルールの一例を説明する。

【0245】

なおこの例には、図 25 の (a) のように、実際には 4 つの CM1 ~ CM4 が連続して放送されている区間に対し、図 25 の (b) 中 A ~ H で示すような 8 つの候補が CM 候補テーブルに存在する場合が示されている。

30

【0246】

まず、図 25 の (c) に示すように、A の CM 候補が判定中であるとする、この候補 A と競合する候補は E と H である。しかしながら、E の区間は A と B で記述でき、また、H の区間は A と B と C と D 等で記述できることから、それぞれ棄却され、A が採用される。続いて、図 25 の (d) に示すように、B が判定中となったときには、F が競合相手となる（このとき、E、H は A の判定により棄却済みとなっている）が、F の区間は B と C で記述できることから棄却され、B が採用される。同様に、図 25 の (e) に示すように、C が判定中の場合には、G が競合相手となるが、G の区間は C と D で記述されることから棄却され、C が採用される。最後に、図 25 の (f) に示すように、D が判定されるときには、すでに競合相手は存在しないので、そもそもこのルールを適用する必要はなく、当該 D がそのまま採用される。

40

【0247】

以上により、この時区間からは、CM 候補として A、B、C、D が選択されることとなる。このルールが適用できない競合関係については、そのまま CM 候補テーブルに残してこの処理を終了する。

【0248】

図 24 に戻り、ステップ S72 のルール判定の後、ルール判定器 22 の処理は、ステップ S73 に進む。ステップ S73 に進むと、ルール判定器 22 は、最小長さ優先ルールを適用した結果、判定中の CM が棄却されたか否か判断する。このステップ S73 において、

50

判定中のCMが棄却されと判断した場合（Yes）、ルール判定器22は、その候補をCM候補テーブルから消去し、ステップS70に戻る。一方、ステップS73において、棄却されていないと判断した場合（No）、ルール判定器22では、ステップS74において、再び判定中のCM候補の区間中に他のCM候補が存在するかどうか、テーブル中を検索する。

【0249】

このステップS74において他のCM候補が存在しないと判定した場合（No）、ルール判定器22では、ステップS80において判定中のCM候補をCM検出出力から出力し、CM候補テーブルから消去する。一方、ステップS74にて存在すると判断した場合（Yes）、ルール判定器22の処理は、ステップS75に進む。

10

【0250】

ステップS75に進むと、ルール判定器22は、隣接優先ルールを適用する。ここで、隣接優先ルールとは、複数のCM候補が競合関係にある場合、それぞれ直前又は直後に隣接するCM候補を検索し、それが存在する方を優先するというルールである。

【0251】

図26を用いて、当該隣接優先ルールについて説明する。

【0252】

なおこの例には、図26の（a）のように、実際には4つのCM11～CM14が連続して放送されている区間に対し、図26の（b）中I～Nで示すような6つの候補が存在する場合が示されている。また、この例の場合、候補M及びNは、偶然CM中にカットチェンジや小音量区間が存在したために候補となっているが、このような候補は、実際には誤った区間であるとはいえ、内容的にはCMを含んでいるため、CMらしさを判定する付加条件のスコア判定によっても、棄却されない場合があるものである。

20

【0253】

このような例において、先ず、図26の（c）に示すように、最古のIが判定される候補となる。当該Iと競合するものとしてMがあるが、Iには隣接する候補Jが存在するのに対し、Mには隣接する候補がないため、Iを採用し、Mを棄却する。次に、図26の（d）に示すように、Jが判定される候補となった場合、Jと競合する候補としてNがあるが、Jには隣接する候補I、Kが存在するのに対し、Nには存在しないため、Jが採用されNが棄却される。次に、図26の（e）、（f）に示すように、残りの候補K、Lには、既に競合する候補がなくなるため、このルールは適用されず、これらK、Lがそのまま採用される。

30

【0254】

以上により、この図26に例示した区間からは、I、J、K、LがCM候補として選択されることとなる。

【0255】

なお、競合関係の候補のいずれにも隣接候補が無い場合、及び複数の候補にそれぞれ隣接候補がある場合には、それらはどちらも棄却されず、CM候補テーブルに残される。

【0256】

図24に戻り、ステップS75の処理後、ルール判定器22の処理は、ステップS76に進む。ステップS76に進むと、ルール判定器22は、隣接優先ルールを適用の結果、判定中のCMが棄却されたか否か判断する。このステップS76において、判定中のCMが棄却されたと判断した場合（Yes）、ルール判定器22は、その候補をCM候補テーブルから消去し、ステップS70の処理に戻る。一方、ステップS76において棄却されていない場合（No）、ルール判定器22は、次のステップS77において、再び判定中のCM候補の区間中に、他のCM候補が存在するかどうか、CM候補テーブル中を検索する。

40

【0257】

このステップS77において、他のCM候補が存在しないと判定された場合（No）、ルール判定器22は、ステップS80において、判定中のCM候補をCM検出出力から出力し

50

、ＣＭ候補テーブルから消去する。一方、ステップＳ７７において、他のＣＭ候補が存在すると判定した場合（Ｙｅｓ）、ルール判定器２２は、ステップＳ７８において、スコア優先ルールを適用する。ここで、スコア優先ルールとは、上記の各ルールによっても競合関係が解消されない場合、付加条件判定器２１により得られた判定スコアＲの高い候補を優先するというルールである。このスコア優先ルールは、対象となる競合関係が解消するまで繰り返し適用する。

【０２５８】

図２７を用いて、当該スコア優先ルールについて説明する。

【０２５９】

なおこの例には、図２７の（ａ）のように、実際には４つのＣＭ２１～ＣＭ２４が連続して放送されている区間に対し、図２７の（ｂ）中Ｐ～Ｗで示すような７つの候補が存在する場合が示されている。

10

【０２６０】

この例において、まず、図２７の（ｃ）に示すように、最古のＰが判定される候補となるが、この候補ＰはＵと競合関係がある。但し、このときの競合関係は、前記最小長さ優先ルールによっても、また、隣接優先ルールによっても競合が解消されない。

【０２６１】

したがって、この場合には、これら競合関係にある候補と関連する全ての競合関係を、ＣＭ候補テーブル中から検索する。すなわち、この場合は、（Ｐ－Ｕ）、（Ｕ－Ｑ）、（Ｑ－Ｖ）、（Ｖ－Ｒ）、（Ｒ－Ｗ）、（Ｗ－Ｓ）という、７候補に対する６つの競合関係が全て関連しているので、スコア優先ルールでは、これら関連する候補の中で最もスコアの高い候補を採用する。この例の場合、判定スコアＲ（２．０）が最も高いスコアであるため、図２７の（ｄ）に示すように、このスコアが採用され、その結果、Ｒと競合関係にある候補Ｖ、Ｗは棄却される。

20

【０２６２】

しかしながら、図２７の（ｅ）に示すように、これによっても（Ｐ－Ｕ）の競合関係は解消されていない。したがって、再びこれらと関連する全ての競合関係を、ＣＭ候補テーブル中から検索する。今回は、Ｖが棄却されたことにより、（Ｐ－Ｕ）、（Ｕ－Ｑ）という、３つの候補が関係する２つの競合関係のみとなる。

【０２６３】

さらに、これらの候補の中で最もスコアの高い候補Ｑ（１．９）を採用し、図２７の（ｆ）に示すように、Ｑと競合関係にある候補Ｕを棄却する。

30

【０２６４】

以上によって、Ｐに関係する競合関係はなくなり、Ｐが採用される。また、Ｕ、Ｖ、Ｗは全て棄却され、Ｑ、Ｒ、Ｓが採用されることとなる。

【０２６５】

なお、もしも、関連する全ての競合関係を検索せず、対象となる競合関係（この例の場合、Ｐ，Ｕ）のみでスコア優先ルールを適用すると、まずＵが採用され、Ｐは棄却される。後にＵとＱとの競合関係により、一時採用されたＵもまた棄却されてしまう。このように、ルール判定器２２では、偶然の処理順序により候補Ｐが棄却されるようなことのないよう、関連競合関係の検索を行っている。

40

【０２６６】

以上のスコア優先ルールにより、選択された候補に関する競合関係は必ず解消されることになる。

【０２６７】

図２４に戻り、ステップＳ７８の処理後、ルール判定器２２の処理は、ステップＳ７９に進む。ステップＳ７９に進むと、ルール判定器２２は、スコア優先ルールを適用の結果、判定中の候補が棄却されたか否か判断する。このステップＳ７９において、判定中の候補が棄却されたと判断した場合（Ｙｅｓ）、ルール判定器２２は、その候補をＣＭ候補テーブルより消去し、ステップＳ７０に戻る。一方、ステップＳ７９において、棄却されな

50

った場合、ルール判定器 22 は、ステップ S80 の CM 検出力として、開始時刻とその長さを出力し、CM 候補テーブルから消去した後、ステップ S70 に戻る。

【0268】

以上説明したように、本実施の形態の第 1 の具体例の CM 検出部 202 においては、ほぼ全ての CM が満足する必須条件に基づき、決定論的に番組中から CM の候補を抽出し、CMらしさの指標である付加条件に基づく特徴量の統計論的な評価により候補を選択し、論理条件により候補のオーバーラップ関係を解消することで、精度の高い CM 検出を実現している。また、本実施の形態では、例えば現行のアナログ TV 放送の映像音声録画録音装置を例にとったが、デジタル TV 放送等に適用される場合にも同様の CM 検出部 202 が適用できることは明らかである。また、例えばラジオ放送に適用される場合には、上記 CM 検出部 202 から映像信号の処理を担当する部分を省略することで同様の機能が実現できる。

10

【0269】

次に、本発明の第 2 の具体例としての CM 検出部 202 について以下に説明する。

【0270】

図 28 には、本発明の第 2 の具体例としての CM 検出部 202 の詳細な構成を示す。

【0271】

当該第 2 の具体例の CM 検出部 202 は、前述した付加条件の中でも基本的なもののみを実装するようにしており、前述した 1) ~ 14) の付加条件のうち、11) ~ 14) に関しては導入しないことで、装置構成を簡略化している（複雑になることを防いでいる）。

20

【0272】

この第 2 の具体例の CM 検出部 202 も、図 15 の例と同様にフロントエンド部とバックエンド部とから構成されている。なお、この図 28 において、図 15 の各構成要素と同じ動作を行う部分については、同一の指示符号を付して、それらの説明は省略する。

【0273】

以下、図 28 の構成において、新たに追加された各構成要素（101, 102, 103）と、付加条件算出器 20 において新たに拡張された機能についてのみ説明する。

【0274】

フロントエンド部に設けられた音源識別器 101 は、デジタル化及びフレーム化された音声信号 2b を入力とし、この音声信号 2b の該当フレームに関する音源名を出力する。音源名としては、例えば、音声、音楽、音声と音楽、その他を挙げることができる。なお、入力された音声信号の音源識別を実現する技術としては、例えば、河地、他による、「VQ 歪みに基づく放送音の自動分類」信学技報、DSP97 95/SP97 50、43/48(1998)に記載された技術や、南、他による、「音情報を用いた映像インデクシングとその応用」信学論、Vo1.J81 D 11、No.3, 529/537(1998)に記載された技術、安部による、特願平 11 - 190693 号の明細書及び図面に記載された技術などがあり、これらを利用することができる。

30

【0275】

この音源識別器 101 により識別された各音源名は、例えば音声 = 1、音楽 = 2、などのように、各フレーム毎に適切に数値化され、特徴量 $U[n]$ として特徴量バッファ 18 に入力される。

40

【0276】

フロントエンド部に設けられた番組ジャンルデータ又は番組ジャンル識別器 102 は、現在処理している番組のジャンル名を出力するものである。番組ジャンルは、例えば、ニュース、ドラマ、野球、サッカーなどである。番組ジャンルデータは、テレビ番組表などから入力してもよく、また近年ではインターネット等を通じて自動的に取得することもできる。または、外部情報に頼らず音声及び映像信号から番組ジャンルを識別する装置を用いることも可能である。なお、音声及び映像信号から番組ジャンルを識別する技術としては、例えば安部による、特願平 11 - 190693 号の明細書及び図面に記載された技術などを利用することができる。

50

【 0 2 7 7 】

この番組ジャンルデータ又は番組ジャンル識別器 1 0 2 により分類された番組ジャンル名は、例えばニュース = 1、ドラマ = 2、などのように、各フレーム毎に適切に数値化され、特徴量 $W[n]$ として特徴量バッファ 1 8 に入力される。

【 0 2 7 8 】

フロントエンド部のその他の各構成要素は、図 1 5 の例と全く同じである。

【 0 2 7 9 】

この第 2 の具体例の場合、フロントエンド部に、上記音源識別器 1 0 1 と番組ジャンルデータ又は番組ジャンル識別器 1 0 2 を設け、これらにより得られた各特徴量 $U[n]$ と $W[n]$ を特徴量バッファ 1 8 に蓄積することで、当該特徴量バッファ 1 8 においては、式 (2 1) に示した特徴量 $G[n]$ が、式 (3 7) のように拡張されることになる。

10

【 0 2 8 0 】

$$G[n] \equiv \{C[n], A[n], A_{LR}[n], A_{LL}[n], A_{RR}[n], F[k;n], B[n], T[n], U[n], W[n]\} \quad (37)$$

バックエンド部の CM 検出器 1 9 は、前述の図 1 5 のものと全く同じであるが、当該第 2 の具体例の場合、CM 候補テーブル 1 9 a ~ 2 1 a は、次のように拡張される。すなわち、この第 2 の具体例の場合の CM 候補テーブル 1 9 a ~ 2 1 a は、前述した Q_1 から Q_{11} までの特徴量に加え、図 2 9 に示すように、後述する特徴量 Q_{12} から Q_{15} が拡張される。なお、図 2 9 は、 Q_1 から Q_{11} までの特徴量についての図示を省略している。

20

【 0 2 8 1 】

また、バックエンド部の CM 確率データベース 1 0 3 には、予め、時間帯に応じた CM の放送確率、及び、番組ジャンルと経過時間に応じた CM の放送確率をデータとして蓄積してある。この CM 確率データベース 1 0 3 からは、現在時刻に応じてそれらの確率が読み出され、付加条件算出器 2 0 に入力するようになされている。なお、これらの確率のデータは、実際の放送を元に統計をとることで作成することができる。

【 0 2 8 2 】

この第 2 の具体例の場合の付加条件算出器 2 0 は、前述の特徴量 Q_1 から Q_{11} に加え、次の特徴量 Q_{12} から Q_{15} の演算を行うよう拡張される。

【 0 2 8 3 】

ここで、特徴量 Q_{12} は、CM 候補区間中に、音声区間があったどうかを検出して求められるものである。音声の有無を表す特徴量 Q_{12} は、式 (3 8) に従って検出される。

30

【 0 2 8 4 】

【 数 2 0 】

$$Q_{12} = \begin{cases} 1 & \left(\text{if } \exists U[n] = (\text{音声}) \text{ or } \exists U[n] = (\text{音声} + \text{音楽}) \text{ for } n_s \leq n < n_e \right) \\ 0 & (\text{otherwise}) \end{cases} \quad (38)$$

40

【 0 2 8 5 】

特徴量 Q_{13} は、上記音声の有無と同様に、CM 候補区間中に、音楽区間があったどうかを検出して求められるものである。この音楽の有無を表す特徴量 Q_{13} は、式 (3 9) に従って検出される。

【 0 2 8 6 】

【 数 2 1 】

$$Q_{13} = \begin{cases} 1 & \left(\text{if } \exists U[n] = (\text{音声}) \text{ or } \exists U[n] = (\text{音声} + \text{音楽}) \text{ for } n_s \leq n < n_e \right) \\ 0 & (\text{otherwise}) \end{cases} \quad (39)$$

【 0 2 8 7 】

10

特徴量 Q_{14} は、現在時刻に応じた C M の発生確率（時間帯確率）である。付加条件算出器 20 では、C M 確率データベース 103 より提供される C M の放送確率を、そのまま特徴量 Q_{14} に代入する。

【 0 2 8 8 】

特徴量 Q_{15} は、番組ジャンル及びその番組の開始からの経過時間に従う C M の放送確率（番組ジャンル確率）である。付加条件算出器 20 では、C M 確率データベース 103 より提供される C M の放送確率を、そのまま特徴量 Q_{15} に代入する。

【 0 2 8 9 】

付加条件判定器 21 以降は、変数としての特徴量 $Q_{12} \sim Q_{15}$ が拡張されるだけであり、前述の図 15 の C M 検出部 202 の場合と同様であるため、説明を省略する。

20

【 0 2 9 0 】

この場合の C M 検出部 202 においては、以上の拡張により、放送信号の音源に応じた C M 検出を行うことができ、また、現在時間に応じた C M 検出を行うこと、さらに、番組ジャンルに応じた C M 検出を行うことが可能となる。

【 0 2 9 1 】

次に、図 30 には、上述した図 15 や図 28 に示した C M 検出部 202 を実装する場合のハードウェア構成の一例を示す。

【 0 2 9 2 】

この図 19 において、A / D 変換器 40 は、前記図 15 や図 28 の A / D 変換器 10 及び 13 の機能を備え、メモリ 41 は、前記フレームメモリ 11 及び音声信号バッファ 14 の機能を備えている。

30

【 0 2 9 3 】

A / V プロセッサまたは D S P（デジタルシグナルプロセッサ）42 は、前記カッチェンジ検出器 112、振幅検出器 15、相関検出器 16、スペクトル検出器 17、音源識別器 101 等の機能を備え、メモリ 43 は、前記特徴量バッファ 18 の機能を備えている。

【 0 2 9 4 】

プロセッサ 44 は、前記 C M 候補検出器 19、付加情報算出器 20、付加条件判定器 21、ルール判定器 22、C M 確率データベース 103 等の機能を備えている。

【 0 2 9 5 】

40

前記動作制御部 23 の機能については、A / V プロセッサまたは D S P（デジタルシグナルプロセッサ）42 か、或いは、プロセッサ 44 が備えることができる。

【 0 2 9 6 】

【 発明の効果 】

本発明においては、入力信号から第 1 の信号区間を検出し、その第 1 の信号区間の検出結果に基づいて入力信号から第 1 の信号部分を抽出して記録することにより、例えば、入力信号としての放送信号から、第 1 の信号としての C M 部分を高精度に検出できると共に、その C M の蓄積、閲覧、視聴、検索等も可能となる。

【 図面の簡単な説明 】

【 図 1 】 本発明実施の形態の放送信号処理装置の全体構成図である。

50

【図 2】CM 検出部、CM 抽出部、CM 記録部のサブセットの構成図である。

【図 3】図 2 のサブセットにおける動作の流れを示すフローチャートである。

【図 4】入力音声信号及び映像信号の離散化とフレームの説明に用いる図である。

【図 5】CM 検出部、CM 抽出部、CM 特徴抽出部、CM 記録部のサブセットの構成図である。

【図 6】図 5 のサブセットにおける動作の流れを示すフローチャートである。

【図 7】表示部の表示画面例を示す図である。

【図 8】CM 記録部、CM 索引生成部、CM 閲覧部、CM 選択部のサブセットの構成図である。

【図 9】図 8 のサブセットにおける動作の流れを示すフローチャートである。

10

【図 10】CM 記録部、特徴比較部のサブセットの構成図である。

【図 11】図 10 のサブセットにおける動作の流れを示すフローチャートである。

【図 12】表示部の他の表示画面例（検索アイコンを表示する例）を示す図である。

【図 13】CM 記録部、特徴比較部、CM 索引生成部、CM 閲覧部、CM 選択部のサブセットの構成図である。

【図 14】図 13 のサブセットにおける動作の流れを示すフローチャートである。

【図 15】第 1 の具体例の CM 検出部の詳細な構成図である。

【図 16】CM 検出部のフロントエンド部における映像信号処理の流れを示すフローチャートである。

【図 17】CM 検出部のフロントエンド部における音声信号処理の流れを示すフローチャートである。

20

【図 18】CM 検出部の CM 候補検出器における動作の流れを示すフローチャートである。

【図 19】必須条件の算出例の説明に用いる図である。

【図 20】第 1 の具体例の CM 検出部における CM 候補テーブルを示す図である。

【図 21】CM 検出部の付加条件算出器における特徴量の算出例の説明に用いる図である。

【図 22】付加条件算出器の構成図である。

【図 23】スコア算出演算の際の単位ステップ関数、矩形関数、シグモイド型関数の説明に用いる図である。

30

【図 24】ルール判定器の動作の流れを示すフローチャートである。

【図 25】最小長さ優先ルールの説明に用いる図である。

【図 26】隣接優先ルールの説明に用いる図である。

【図 27】スコア優先ルールの説明に用いる図である。

【図 28】第 2 の具体例の CM 検出部の詳細な構成図である。

【図 29】第 2 の具体例の CM 検出部における CM 候補テーブル（拡張部分のみ）を示す図である。

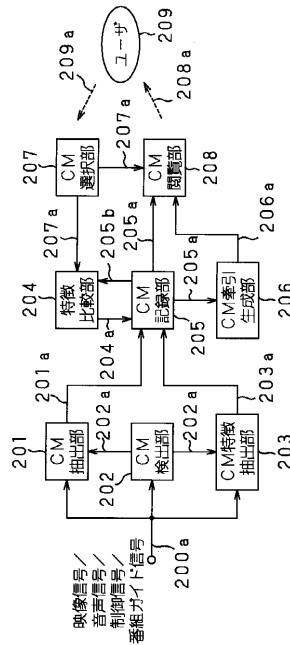
【図 30】CM 検出部を実装する場合の一例としてのハードウェア構成図である。

【符号の説明】

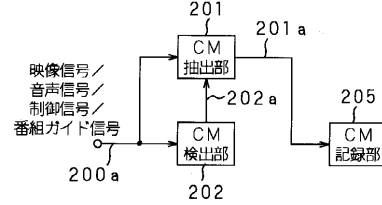
201 CM 抽出部、 202 CM 検出部、 203 CM 特徴抽出部、 204 特
 徴比較部、 205 CM 記録部、 206 CM 索引生成部、 207 CM 選択部、
 208 CM 閲覧部、 801 表示プロセッサ、 802 表示部、 810 CM
 索引画像又は映像、 811 CM 再生アイコン、 812 音声再生アイコン、 813
 前ページアイコン、 814 次ページアイコン、 815 検索アイコン

40

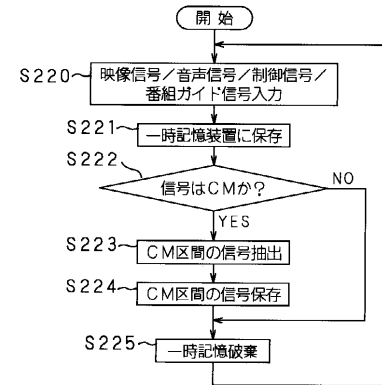
【図 1】



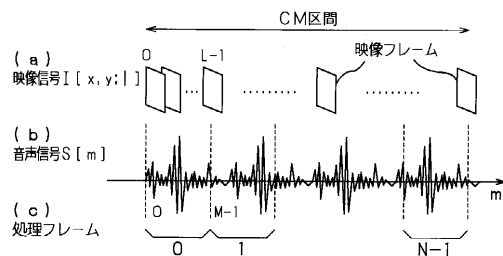
【図 2】



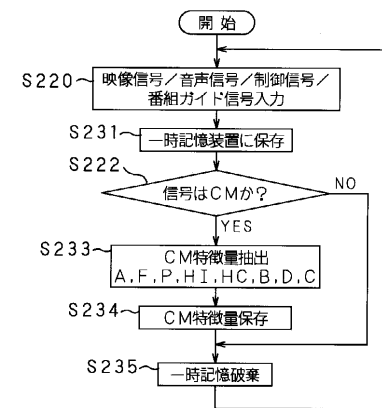
【図 3】



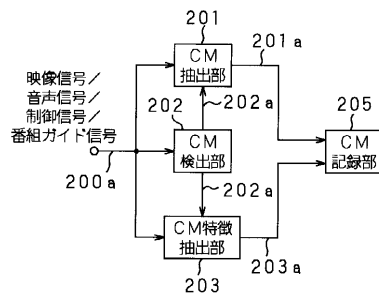
【図 4】



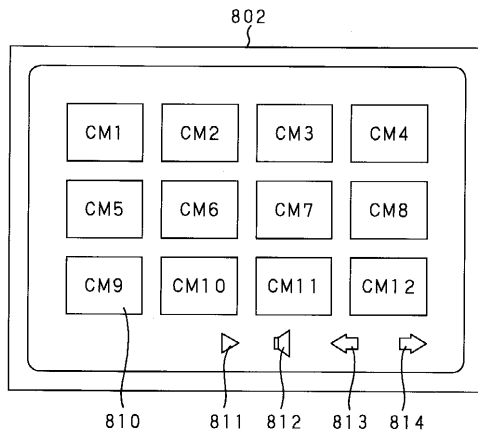
【図 6】



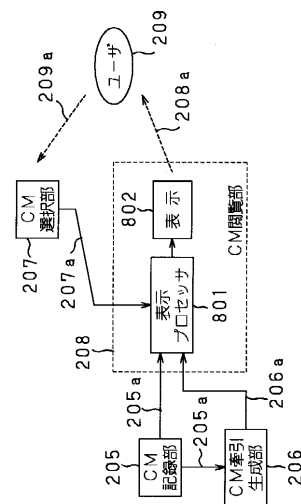
【図 5】



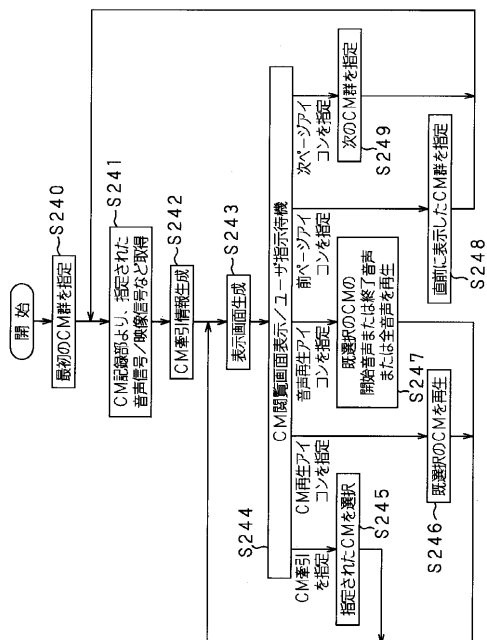
【図 7】



【図 8】



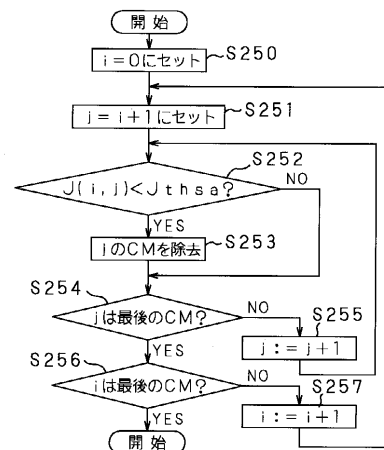
【図 9】



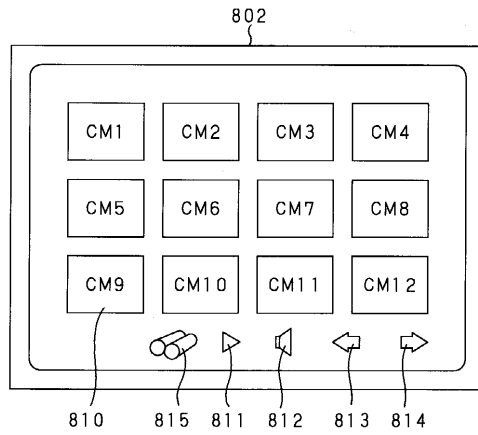
【図 10】



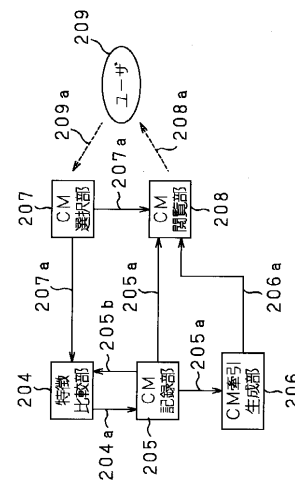
【図 11】



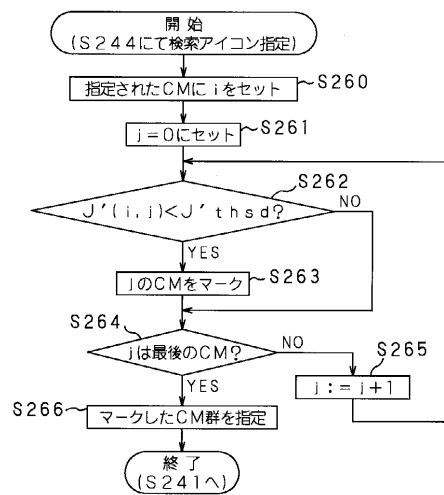
【図 12】



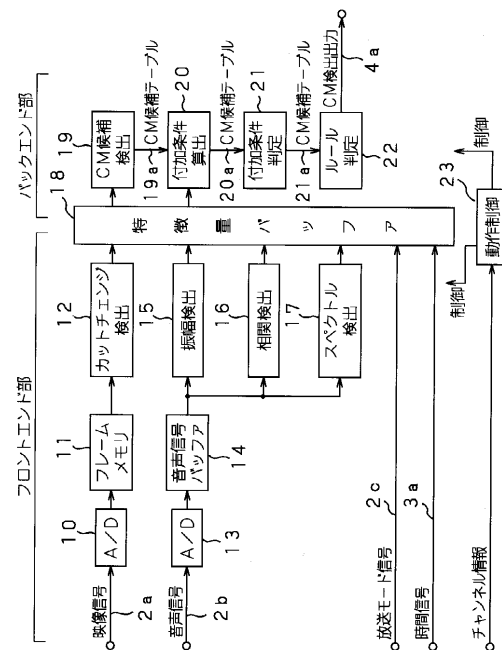
【図 13】



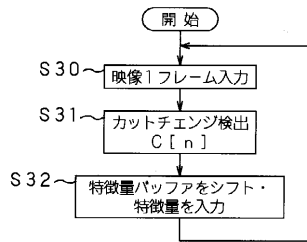
【図 14】



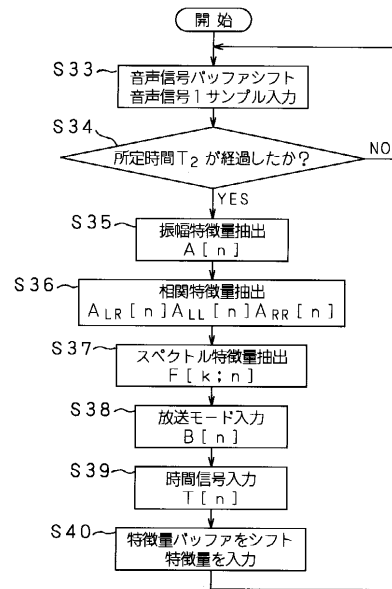
【図 15】



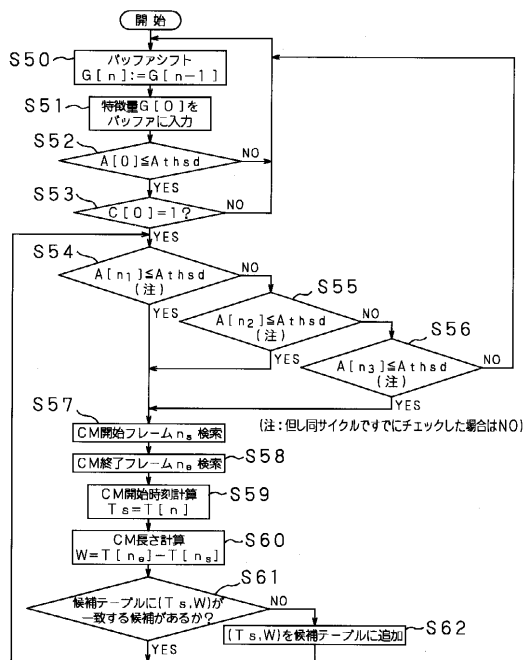
【図 16】



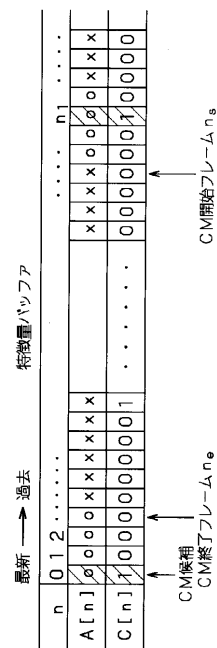
【図 17】



【図 18】



【図 19】

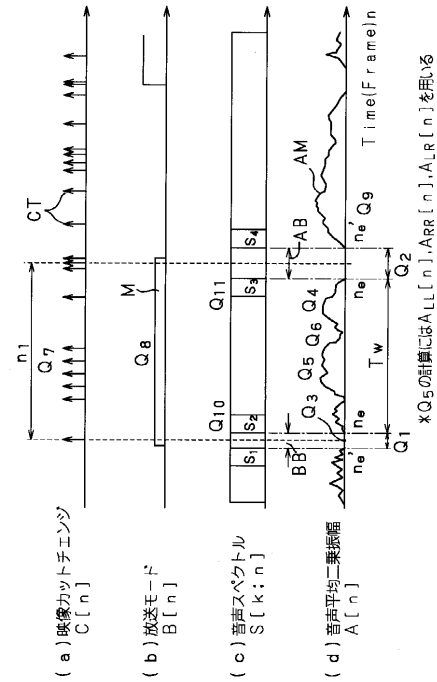


【図 20】

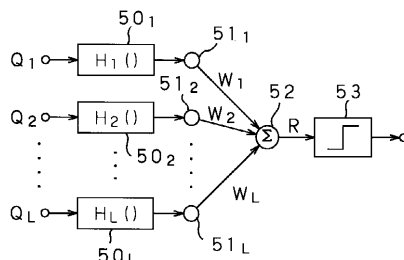
項目	記号	単位	必須条件例 (19a)	付加条件例 (20a)	準条件例 (21a)
開始時刻	T_s	時分秒	1:23:45	1:23:45	1:23:45
長さ(有音長)	T_w	秒	14.63	14.63	14.63
前ブレイク長	Q_1	ms	—	300.0	300.0
後ブレイク長	Q_2	ms	—	300.0	300.0
前ブレイク最小振幅	Q_3	注	—	0.00015	0.00015
後ブレイク最小振幅	Q_4	注	—	0.00020	0.00020
左右相関値	Q_5	—	—	0.934	0.934
平均振幅値	Q_6	注	—	0.010	0.010
カット数	Q_7	個	—	9	9
放送モード	Q_8	—	—	1	1
隣接候補数	Q_9	個	—	2	2
前スペクトル差分エネルギー	Q_{10}	—	—	0.41	0.41
後スペクトル差分エネルギー	Q_{11}	—	—	0.63	0.63
スコア	R	—	—	—	1.80
スコア判定結果	Z	—	—	—	1

注：ここでの音声信号の振幅に関する量は、最大振幅に対する比率として表している

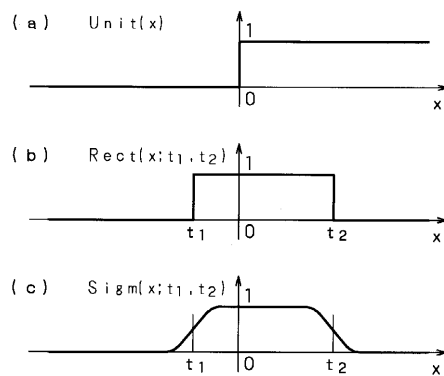
【図 21】



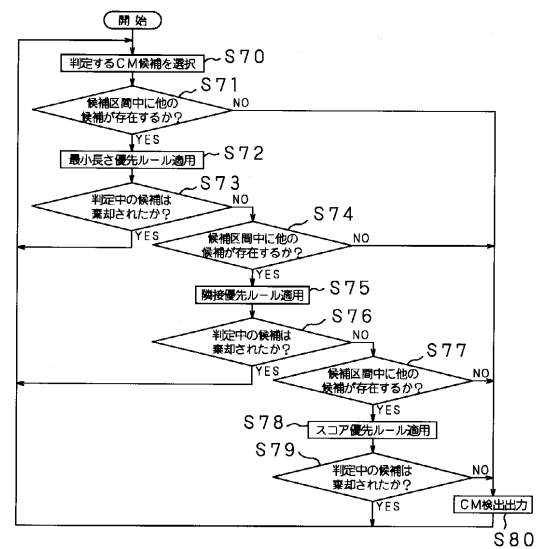
【図 22】



【図 23】



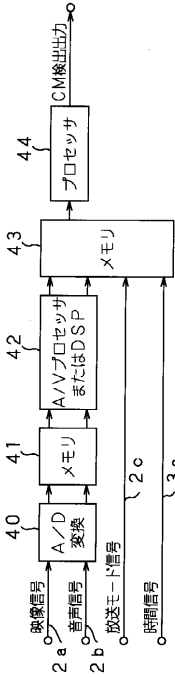
【図 24】



【図 29】

項目	記号	単位	必須条件例 (19 a)	付加条件例 (20 a)	判定例 (21 a)
音声の有無 音楽の有無 時間遅延率 番組ジャンル確率	Q12	—	—	—	1
	Q13	—	—	—	1
	Q14	—	—	0.15	0.15
	Q15	—	—	0.1	0.1

【図 30】



フロントページの続き

審査官 小林 大介

(56)参考文献 特開平11-102550(JP,A)

特開平09-247612(JP,A)

特開平05-205344(JP,A)

特開2000-354213(JP,A)

和田哲夫,CM放送チェッカーの開発,映像情報メディア学会技術報告,日本,社団法人映像情報メディア学会,1998年5月20日,Vol.22 No.25,p.55-59

高橋克直,画紋を用いた動画検索方式,情報処理学会研究報告,日本,社団法人情報処理学会,1998年12月4日,Vol.98 No.111,p.15-20

多田行伸,TV放送のCM検出方式についての一検討,映像情報メディア学会技術報告,日本,社団法人映像情報メディア学会,1997年3月14日,Vol.21 No.23,p.19-23

(58)調査した分野(Int.Cl.,DB名)

H04N 5/76-5/95