

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4300697号
(P4300697)

(45) 発行日 平成21年7月22日(2009.7.22)

(24) 登録日 平成21年5月1日(2009.5.1)

(51) Int.Cl.

F I

H O 4 N 5/91 (2006.01)

H O 4 N 5/91 Z

G O 6 N 3/00 (2006.01)

G O 6 N 3/00 5 6 O Z

G O 6 T 7/00 (2006.01)

G O 6 T 7/00 3 O O E

G 1 O L 11/02 (2006.01)

G 1 O L 3/00 5 1 3 A

G 1 O L 15/04 (2006.01)

G 1 O L 3/00 5 1 3 D

請求項の数 23 (全 37 頁) 最終頁に続く

(21) 出願番号 特願2000-307907 (P2000-307907)

(22) 出願日 平成12年10月6日(2000.10.6)

(65) 公開番号 特開2002-16873 (P2002-16873A)

(43) 公開日 平成14年1月18日(2002.1.18)

審査請求日 平成19年3月2日(2007.3.2)

(31) 優先権主張番号 特願2000-127658 (P2000-127658)

(32) 優先日 平成12年4月24日(2000.4.24)

(33) 優先権主張国 日本国(JP)

前置審査

(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 100082131

弁理士 稲本 義雄

(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

審査官 竹中 辰利

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、

前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

を有し、

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、C M候補区間中に他のC M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C M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C M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

信号処理装置。

【請求項2】

上記検出手段は、上記特徴量に基づいて上記候補区間が上記第1の信号である可能性を評価する特徴量評価手段を有する

請求項 1 記載の信号処理装置。

【請求項 3】

上記検出手段は、上記特徴量に基づいて、上記候補区間の信号と、別途指定した第 1 の信号との一致を判定する一致判定手段を有する

請求項 1 記載の信号処理装置。

【請求項 4】

上記入力信号の振幅を検出する振幅検出手段を有し、

上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号の振幅が所定の値より小さくなるパターンを検出する

請求項 1 記載の信号処理装置。

10

【請求項 5】

上記入力信号の変化を検出する変化検出手段を有し、

上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号が急激に所定の変化量を超えて変化するパターンを検出する

請求項 1 記載の信号処理装置。

【請求項 6】

上記入力信号の所定の信号成分が、所定の範囲内に収まる単位区間を検出する一様成分検出手段を有し、

上記候補区間を検出する際の特徴的パターンとして、所定の時間間隔に対応して上記入力信号の単位区間についての所定の信号成分が一様になるパターンを検出する

請求項 1 記載の信号処理装置。

20

【請求項 7】

上記特徴量抽出手段は、上記入力信号の振幅を検出する振幅検出手段を有し、上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間の前及び / 又は後の入力信号の振幅の大きさを抽出する

請求項 1 記載の信号処理装置。

【請求項 8】

上記特徴量抽出手段は、上記入力信号の振幅を検出する振幅検出手段を有し、上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間の前及び / 又は後の入力信号の振幅が、所定の閾値より小さい区間の時間長を抽出する

請求項 1 記載の信号処理装置。

30

【請求項 9】

上記特徴量抽出手段は、上記入力信号の相関を検出する相関検出手段を有し、上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間における入力信号の相互相関を抽出する

請求項 1 記載の信号処理装置。

【請求項 10】

上記特徴量抽出手段は、上記入力信号の振幅を検出する振幅検出手段を有し、上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間における入力信号の平均振幅を抽出する

請求項 1 記載の信号処理装置。

40

【請求項 11】

上記特徴量抽出手段は、上記入力信号の変化を検出する変化検出手段を有し、上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間において入力信号が急激に変化する回数又は頻度を抽出する

請求項 1 記載の信号処理装置。

【請求項 12】

上記特徴量抽出手段は、複数のモードを取り得る入力信号の当該モードを検出するモード検出手段を有し、上記第 1 の信号らしさを表わす特徴量として、上記第 1 の信号の候補区間における上記モードを抽出する

50

請求項 1 記載の信号処理装置。

【請求項 1 3】

上記特徴量抽出手段は、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前又は後に続く第 1 の信号の有無を抽出する

請求項 1 記載の信号処理装置。

【請求項 1 4】

上記特徴量抽出手段は、上記入力信号のスペクトルを検出するスペクトル検出手段を有し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前又は後の境界における上記スペクトルの変化を抽出する

請求項 1 記載の信号処理装置。

10

【請求項 1 5】

上記特徴量抽出手段は、上記入力信号の信号源を識別する信号源識別手段を有し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間における信号源の種類を抽出する

請求項 1 記載の信号処理装置。

【請求項 1 6】

上記特徴量抽出手段は、複数のジャンルに分け得る上記入力信号のジャンルを識別するジャンル識別手段を有し、上記第 1 の信号らしさを表す特徴量として、上記第 1 の信号の候補区間の前後の信号のジャンルを抽出する

請求項 1 記載の信号処理装置。

20

【請求項 1 7】

上記特徴量抽出手段は、上記第 1 の信号らしさを表す特徴量として、上記入力信号の振幅が基準値より小さいときの回数、その長さ、または上記入力信号の振幅の分散を抽出する

請求項 1 記載の信号処理装置。

【請求項 1 8】

上記特徴量評価手段は、上記特徴量を荷重加算し、当該加重加算後の特徴量に基づき、上記候補区間が上記第 1 の信号である可能性を評価する

請求項 2 記載の信号処理装置。

【請求項 1 9】

上記入力信号を記録及び / 又は再生する記録及び / 又は再生手段を有する

請求項 1 記載の信号処理装置。

30

【請求項 2 0】

上記第 1 の信号の区間をスキップするスキップ手段を有する

請求項 1 記載の信号処理装置。

【請求項 2 1】

上記第 1 の信号の区間のみを取り出す取り出し手段を有する

請求項 1 記載の信号処理装置。

【請求項 2 2】

上記入力信号は映像及び / 又は音声信号からなり、上記第 1 の信号の区間はコマースャルメッセージ区間である

請求項 1 記載の信号処理装置。

40

【請求項 2 3】

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し、

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し、

上記特徴量に基づき、上記第 1 の信号の区間を検出し、

検出した上記第 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

50

る

ステップを含み、

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。

【発明の詳細な説明】

【０００１】

10

【発明の属する技術分野】

本発明は、信号処理装置及び方法に関し、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自動的に検出或いは検索可能とする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】

一般に、放送番組についてスポンサーが付いているテレビジョン（以下、適宜ＴＶとする）放送では、本編（番組そのもの）の合間に、その番組を提供しているスポンサーのコマーシャルメッセージ（以下、単にＣＭとする。）が挿入される。

【０００３】

20

しかしながら、視聴者の興味は、主として本編にあるため、ＣＭを省略して視聴したいという要望は多い。

【０００４】

その一方で、関心のある製品のＣＭや、ストーリー性のあるＣＭ、有名俳優が登場するＣＭなど、ＣＭ自体を視聴対象とする場合も少なくない。

【０００５】

このようなことから、特に、磁気テープや磁気ディスク等の媒体にＴＶ放送番組等を録画する場合において、放送番組から本編とＣＭとを分離／区別して、必要なときに必要な部分を視聴することが可能となれば、上述したような多様な要望に対する有効な解決手段を与えることになると考えられる。

30

【０００６】

ところで、ＴＶ放送からＣＭのみを省略して試聴可能とする技術としては、従来より、いわゆるＣＭスキップ機能（広義のＣＭスキップ機能）が存在し、例えば家庭用ビデオ録画再生装置には当該ＣＭスキップ機能が搭載されているものが多い。

【０００７】

このＣＭスキップ機能（広義のＣＭスキップ機能）は、その仕組みに応じて４つに大別され、それぞれ一般には、ＣＭ早送り機能、ＣＭカット機能、放送モードに基づくＣＭスキップ機能（狭義のＣＭスキップ機能）、放送モードによらないＣＭスキップ機能（狭義のＣＭスキップ機能）と呼ばれている。なお、これら４つに大別されたＣＭスキップ機能のうち、上記ＣＭ早送り機能とは、日本国内においては通例としてＣＭが１５秒の倍数で製作されていることを利用し、家庭用ビデオ録画再生装置でのビデオ再生中に、例えばリモコンボタン操作によって（３０秒）＊（ボタンを押した回数）の時間だけ早送りを行うことにより、ＣＭ部分を飛ばす機能である。このＣＭ早送り機能に関する技術については、例えば特開平１０－２６９６５１号公報などに公開されている。また、特開平９－３０７８４１号公報には、上記ＣＭ早送り機能における早送り終了時刻を時間で決定せず、映像信号のブラックフレームと音声信号の無音部が同時に所定長さ以上にわたり生じる部分を、早送り終了時刻とする技術が公開されている。

40

【０００８】

ＣＭカット機能とは、日本国内においてはＣＭがステレオ放送されることが多く、また、ＴＶ放送信号には音声モノラルモード／ステレオモード／音声多重モードの何れである

50

かを示すパイロット信号が多重されていることを利用し、例えば音声モノラルモード又は音声多重モードとなされている本編を録画するような場合に、ステレオモード区間（すなわちCM区間）のみ録画を停止することで、CM部分をカットすることを可能にする機能である。このCMカット機能に関する技術については、例えば特開平3-158086号、特開平3-2622872号などの公報に公開されている。

【0009】

放送モードに基づくCMスキップ機能（狭義のCMスキップ機能）とは、CMカット機能と同様に放送モードの違いを利用する仕組みであるが、上記CMカット機能が録画時にCMをカットしているのに対し、当該放送モードに基づくCMスキップ機能では、録画時には全ての映像及び音声を記録すると同時にその放送モードも記録しておき、再生時にステレオ放送区間（すなわちCM区間）のみ自動的に早送りする機能である。この放送モードに基づくCMスキップ機能に関する技術については、例えば特開平5-250762号公報などに公開されている。

10

【0010】

放送モードによらないCMスキップ機能（狭義のCMスキップ機能）とは、放送信号中に含まれる音声信号の無音区間間隔や、映像信号の映像変換点（映像が急激に変化する点）の出現間隔、映像信号の黒レベル／白レベルの出現間隔などを利用して、これらが15秒の倍数間隔に合致する部分をCMとして早送りする機能である。当該放送モードによらないCMスキップ機能に関する技術については、特開平8-317342号公報や、文献「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、VIR97-22、19/23（1997））などに公開されている。

20

【0011】

【発明が解決しようとする課題】

しかしながら、上記CM早送り機能自体は、家庭用ビデオ録画再生装置にて録画したTV放送を再生して視聴する視聴者に対して、CMかどうかの判断を任せるものであり、TV放送から自動的にCM部分を検出するようなCM検出機能は備えていない。

【0012】

また、上記放送モードに基づくCMカット機能やCMスキップ機能（狭義のCMスキップ機能）では、放送モードの違い、すなわち例えばステレオモードによるCM検出を行うようになされているが、例えば本編自体がステレオモードとなされている場合や、CMがモノラルモードや音声多重モードとなっている場合には、全く効果がない（すなわちCMカットやCMスキップができない）。

30

【0013】

一方、放送モードによらないCMスキップ機能（狭義のCMスキップ機能）では、視聴者や放送モードに依存することなく、放送内容に基づいたCM検出が可能となっている。ただし、当該機能の場合は、音声信号の無音区間間隔や、映像信号の映像変換点（映像が急激に変化する点）の出現間隔などが、予め想定されている設定値と一致することなど、非常に狭い条件に基づいた決定論的な判定手続きによりCM区間を検出するようになされている。これに対し、実際の放送番組では、放送時間の調整の目的や人為的なスイッチングなどの要因によって無音区間間隔が例えば短縮されていたりする場合が少なからずあり、逆に、番組本編の中にも上記の条件を満たす区間が多く存在している。このため、無音区間間隔が予め想定されている長さより短いようなCMの場合は全く検出ができなくなる問題が発生し、逆に、上記設定値の条件を満たす本編の場合は当該本編をCM区間として誤って検出してしまう等の問題点が生じてしまう。

40

【0014】

さらに、上記4つの機能では、例えば複数のCMが連続されて放送されている場合には、CM全体の区間を検出することはできても、個々のCMの検出はできない。このため、CMを個別に抽出して視聴したいような場合（要望）には適さない。

【0015】

以上のようなことから、TV放送信号からCM部分を高精度に検出又は検索できることが

50

望まれている。

【 0 0 1 6 】

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ＴＶ放送信号に含まれるコマーシャルメッセージ部分を高精度に検出又は検索可能とする、信号処理装置及び方法を提供することを目的とする。

【 0 0 1 7 】

【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、その候補区間内又はその前後の入力信号から第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、その特徴量に基づき第１の信号の区間を検出する検出手段と、検出手段により検出された第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を有し、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。

【 0 0 1 8 】

ここで、検出手段は、特徴量に基づいて候補区間が第１の信号である可能性を評価する特徴量評価手段を有する。或いは、検出手段は、特徴量に基づいて、候補区間の信号と、別途指定した第１の信号との一致を判定する一致判定手段を有する。

【 0 0 1 9 】

次に、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し、その候補区間内又はその前後の入力信号から第１の信号らしさを表わす特徴量を抽出し、その特徴量に基づき第１の信号の区間を検出し、検出した第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ステップを含み、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。

【 0 0 2 2 】

【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。

【 0 0 2 3 】

先ず、本発明実施の形態の具体的構成について述べる前に、本発明に基づき、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。

【 0 0 2 4 】

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。

【 0 0 2 5 】

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。

【 0 0 2 6 】

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が示すパターンである。本発明では、このようにほぼ全てのＣＭが示す信号のパターンを「特徴的パターン」と称し、それを検出するための条件を「必須条件」と称することにする。

【 0 0 2 7 】

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う虞が残る。

10

【 0 0 2 8 】

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 0 2 9 】

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 0 3 0 】

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。

20

【 0 0 3 1 】

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。

【 0 0 3 2 】

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 0 3 3 】

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 0 3 4 】

６）ＣＭの放送モードは、ステレオモードであることが多い。

30

【 0 0 3 5 】

７）ＣＭ区間では、複数のＣＭが連続して放送されることが多い。

【 0 0 3 6 】

８）ＣＭ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 0 3 7 】

９）逆に、ＣＭ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ＣＭなど）。

【 0 0 3 8 】

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

40

【 0 0 3 9 】

１１）ＣＭは、音声と音楽を同時に含むことが多い。

【 0 0 4 0 】

１２）番組編成上、毎時丁度の時刻近辺では、ＣＭが放送される確率が高い。

【 0 0 4 1 】

１３）同様に、毎時３０分付近でもＣＭが放送される確率が高い。

【 0 0 4 2 】

１４）番組のジャンルによってＣＭ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）。

50

【0043】

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、CM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、TV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、CMである可能性（CMらしさ）を評価する際の有効な条件となる。

【0044】

さらに、TV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、TV放送信号からCMであろうと思われる信号部分（CM候補）を検出しようとする場合において、例えば、TV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CM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CM区間ではあり得ない。本発明では、TV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。

10

【0045】

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、TV放送信号から高精度でCM部分を検出可能としている。

【0046】

20

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TV放送信号からCM候補（CM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CMらしさ（CMである可能性）の統計論的な評価によってCM候補を選択し、「論理条件」によりCM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CM検出を実現するものである。

【0047】

図1には、上述のようにTV放送信号からCMを検出可能とした、本発明の第1の実施の形態としての映像音声記録装置の概略構成を示す。

【0048】

図1において、アンテナ6により受信されたTV放送信号は、チューナ1により同調され、放送信号1aとして復調器2に入力する。

30

【0049】

復調器2は、入力された放送信号1aに含まれるパイロット信号の変調周波数から、その入力時点での放送モード（モノラルモード/ステレオモード/音声多重モード）を判別し、また、放送信号1aに含まれる映像及び音声の変調信号を復調する。このとき、音声変調信号は、上記判別した放送モードに応じて復調される。当該復調器2により復調された映像信号2a、音声信号2b、及び放送モード信号2cは、CM検出部4に入力される。また、映像信号2a、音声信号2bは、映像音声記録部5にも入力される。

【0050】

映像音声記録部5は、磁気テープ、磁気ディスク、光ディスク若しくは光磁気ディスク等を記録媒体として用い、上記映像信号2a及び音声信号2bを記録再生可能な装置である。なお、映像音声記録部5は、映像信号及び音声信号に対して符号化や圧縮などを施して記録し、再生時にはその符号化や圧縮された映像信号及び音声信号を復号化、伸張する装置であってもよい。

40

【0051】

時計部3は、時刻を表す時間信号3aを発生し、当該時間信号3aをCM検出部4に供給する。

【0052】

CM検出部4は、本発明の主要部であり、詳細については後述するが、上記復調器2からの映像信号2a、音声信号2b、放送モード信号2cと、時計部3からの時間情報3aと

50

を入力とし、放送信号 1 a の映像及び音声信号に含まれる C M 区間を検出し、その開始時刻と長さを C M 検出力 4 a として出力する。当該 C M 検出力 4 による C M 検出力 4 a は、映像音声記録部 5 に送られ、上記映像信号 2 a 及び音声信号 2 b と共に記録される。

【 0 0 5 3 】

また、この C M 検出力 4 には、チューナ 1 から、上記同調された放送チャンネルを示すチャンネル情報 1 b も供給される。当該チャンネル情報 1 b は、C M が含まれない放送チャンネルが指定されたか否かを、C M 検出力 4 において判断する為に用いられる。すなわち、C M 検出力 4 は、上記チャンネル情報 1 b により、上記チューナ 1 において明らかに C M が放送されない放送チャンネルの指定がなされていると判断した場合、上記 C M 検出力 4 自身が行わないようにする。なお、ここでは、C M 検出力 4 自身がチャンネル情報 1 b に基づいて C M 検出力 4 自身が行うか否か判断することとしているが、チューナ 1 がチャンネル情報 1 b により C M 検出力 4 での C M 検出力 4 自身を制御するようにしても良い。

10

【 0 0 5 4 】

以上、チューナ 1、復調器 2、時計部 3、映像音声記録部 5 に示した各構成要素は、現在広く知られているビデオ装置などの磁気記録再生装置のものと同様であるため、以下では、C M 検出力 4 に限って詳細に説明する。

【 0 0 5 5 】

図 2 には、上記 C M 検出力 4 の第 1 の具体例の詳細な構成を示す。なお、この図 2 中の各信号のうち、図 1 と共通の信号については、図 1 と同じ指示符号を付している。また、当該 C M 検出力 4 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 2 3 は、チューナ 1 から供給された上記チャンネル情報 1 b に基づいて、上記チューナ 1 において明らかに C M が放送されない放送チャンネルの指定がなされているか否か判断し、その判断結果に応じて、当該図 2 の各部における C M 検出力 4 自身が行わないように制御するものである。

20

【 0 0 5 6 】

先ず、図 2 のフロントエンド部から説明する。

【 0 0 5 7 】

この図 2 において、図 1 の復調器 2 より供給された映像信号 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 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。

30

【 0 0 5 8 】

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、輝度が一定となるフレーム（以下、一定輝度フレームと呼ぶ）を検出する。

【 0 0 5 9 】

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、そのフレームを一定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 30 ms）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。

40

【 0 0 6 0 】

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定輝度フレームの検出について、より具体的に説明する。

【 0 0 6 1 】

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

50

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自乗和 $D[n]$ は、式 (1) により得られ、また、第 n フレームの輝度分散値 $V[n]$ は、式 (2) により得られる。

【 0 0 6 2 】

【 数 1 】

$$D[n] = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \left(I_n(x, y) - I_{n-1}(x, y) \right)^2 \quad (1) \quad 10$$

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

【 0 0 6 3 】

また、このときのカットチェンジ検出器 12 の検出出力 $C[n]$ は、式 (3) により表わされる。

【 0 0 6 4 】

【 数 2 】

$$C[n] = \begin{cases} 1 & \left(D[n] \geq D_{thsd} \text{ or } V[n] \leq V_{thsd} \right) \\ 0 & \left(D[n] < D_{thsd} \text{ and } V[n] > V_{thsd} \right) \end{cases} \quad (3) \quad 20$$

【 0 0 6 5 】

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 0 6 6 】

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 0 6 7 】

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。

【 0 0 6 8 】

次に、図 1 の復調器 2 より供給された音声信号 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 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 (

30

40

50

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。

【0069】

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15ms、以下、これを1フレーム周期とする)毎の短時間平均自乗振幅を検出する。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 (15ms、1フレーム周期)毎に、短時間平均自乗振幅を検出する。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。

10

【0070】

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自乗振幅 $A[n]$ を式(4)により計算する。すなわち、平均自乗振幅は15ms(1/2フレーム)毎に計算され、その15ms毎の平均自乗振幅の30ms(1フレーム)の期間における平均値がさらに演算され、最終的な、1フレームの平均自乗振幅とされる。

【0071】

【数3】

20

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

【0072】

当該振幅検出器15の検出出力である平均自乗振幅 $A[n]$ 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18へ送られる。

【0073】

30

相関検出器16は、音声信号バッファ14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16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、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。

【0074】

より具体的に説明すると、相関検出器16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(5)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(6)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(7)により計算する。

40

【0075】

【数4】

$$A_{LR}[n] = \sum_{m=0}^{M-1} S_L[m+nT_2] S_R[m+nT_2] \quad (5)$$

$$A_{LL}[n] = \sum_{m=0}^{M-1} S_L^2[m+nT_2] \quad (6)$$

$$A_{RR}[n] = \sum_{m=0}^{M-1} S_R^2[m+nT_2] \quad (7)$$

【 0 0 7 6 】

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ひとつとして特徴量バッファ 1 8 へ送られる。

10

【 0 0 7 7 】

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 0 7 8 】

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 (8) により表わされる。

20

【 0 0 7 9 】

【 数 5 】

$$F[k;n] = \left| \sum_{m=0}^{M-1} (S_L[m] + S_R[m]) e^{-2\pi j m k / M} \right|^2 \quad (8)$$

30

【 0 0 8 0 】

この式 (8) 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 0 8 1 】

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ひとつとして特徴量バッファ 1 8 へ送られる。

【 0 0 8 2 】

次に、図 1 の復調器 2 より供給された放送モード信号 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。

【 0 0 8 3 】

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 (9) のような数値 $B[n]$ となされる。

40

【 0 0 8 4 】

【 数 6 】

$$B[n] = \begin{cases} 0 & (\text{モノラルモード}) \\ 1 & (\text{ステレオモード}) \\ 2 & (\text{音声多重モード}) \end{cases} \quad (9)$$

【 0 0 8 5 】

50

この放送モード信号 2 c を離散化した数値 $B[n]$ は、TV 放送信号の特徴量の一つとして特徴量バッファ 18 へ送られる。

【0086】

同様に、図 1 の時計部 3 より供給された時間信号 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 18 へ送られる。

【0087】

特徴量バッファ 18 は、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 からの平均自乗振幅 $A[n]$ 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 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 (10) 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 CM 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 80 秒などとする。

$$G[n] = \{ C[n], A[n], A_{LR}[n], A_{LL}[n], A_{RR}[n], F[k;n], B[n], T[n] \} \quad (10)$$

【0088】

以上の A/D 変換器 10 から特徴量バッファ 18 までが、図 2 に示した CM 検出部 4 のフロントエンド部の構成であり、以下、図 3、図 4 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 3 のステップ S30 乃至 S32 までは映像信号 2 a についての処理の流れを表しており、図 4 のステップ S33 乃至 S40 までは音声信号 2 b 及び放送モード信号 2 c、時間信号 3 a についての処理の流れを表している。

【0089】

まず、映像信号 2 a についての処理の流れを表す図 3 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 2 a 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 2 a 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 2 a 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 2 a 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 2 a が押し出されて出力されるようになっている。

【0090】

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 2 a 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。

【0091】

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。

【0092】

一方、音声信号 2 b についての処理の流れを表す図 4 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 2 b 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 2 b 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 2 b 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 2 b 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 2 b 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 2 b が押し出されて出力されるようになっている。

【0093】

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 2 b 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 2 b を読み出して振幅検出器 1 5 に送り、前述のようにして、平均自乗振幅 $A[n]$ を求める。

【 0 0 9 4 】

同時に、フロントエンド部は、ステップ S 3 6 の処理として、音声信号バッファ 1 4 に蓄積された音声信号 2 b を相関検出器 1 6 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。

【 0 0 9 5 】

また同時に、フロントエンド部は、ステップ S 3 7 の処理として、音声信号バッファ 1 4 に蓄積された音声信号 2 b をスペクトル検出器 1 7 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。

【 0 0 9 6 】

さらに、フロントエンド部は、ステップ S 3 8 の処理として、図 1 の復調器 2 より供給された放送モード信号 2 c から、前述のように離散化した数値 $B[n]$ を求めると共に、図 1 の時計部 3 より供給された時間信号 3 a から、前述のように離散化された数値 $T[n]$ を求める。

【 0 0 9 7 】

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 1 2 からの検出出力 $C[n]$ と、振幅検出器 1 5 からの平均自乗振幅 $A[n]$ 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 $T[n]$ とからなる特徴量 $G[n]$ を、特徴量バッファ 1 8 に蓄積する。

【 0 0 9 8 】

図 2 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 1 8 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。

【 0 0 9 9 】

図 2 において、特徴量バッファ 1 8 に蓄積された特徴量は、フレーム毎に CM 候補検出器 1 9 に送られる。

【 0 1 0 0 】

当該 CM 候補検出器 1 9 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（11）で且つ式（12）で且つ式（13）の条件として書き下すことができる。

【 0 1 0 1 】

$$A[0] < A_{thsd} \quad (11)$$

$$C[0] = 1 \quad (12)$$

$$A[n_1] < A_{thsd}, C[n_1] = 1 \text{ 又は } A[n_2] < A_{thsd}, C[n_2] = 1 \text{ 又は } A[n_3] < A_{thsd}, C[n_3] = 1 \quad (13)$$

ただし、 A_{thsd} は所定の自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

10

20

30

40

50

多少の幅を持たせる。

【 0 1 0 2 】

ここで、図 5 を用いて、上記 C M 候補検出器 1 9 の動作の流れを説明する。

【 0 1 0 3 】

図 5 において、特徴量バッファ 1 8 では、ステップ S 5 0 のバッファシフト処理とステップ S 5 1 の特徴量入力処理として、図 3 のステップ S 3 2 で説明したフレームメモリと図 4 のステップ S 4 0 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 1 8 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 1 8 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。

10

【 0 1 0 4 】

上記ステップ S 5 0 及びステップ S 5 1 の処理により、特徴量バッファ 1 8 から 1 フレーム (1 サンプル) 分の特徴量が入力されると、C M 候補検出器 1 9 は、ステップ S 5 2 及びステップ S 5 3 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、C M 候補検出器 1 9 は、ステップ S 5 2 において、先ず最初のフレームの平均自乗振幅 $A[0]$ と所定の自乗振幅の閾値 A_{thsd} を比較し、次に、ステップ S 5 3 の処理として、前記検出出力 $C[0]$ 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。C M 候補検出器 1 9 では、これらステップ S 5 2 , S 5 3 の判定処理の結果、上記平均自乗振幅 $A[0]$ が所定の自乗振幅の閾値 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 C M 候補としてステップ S 5 7 以降 (ステップ S 5 4 乃至 S 5 6 については後述する) の処理に進み、逆に、上記平均自乗振幅 $A[0]$ が所定の自乗振幅の閾値 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 C M 候補にはならないとしてステップ S 5 0 の処理に戻る。

20

【 0 1 0 5 】

上記ステップ S 5 2 , S 5 3 の各判定処理の結果、上記平均自乗振幅 $A[0]$ が所定の自乗振幅の閾値 A_{thsd} を超えず、且つ、上記必須条件を満たしていると判定された場合、C M 候補検出器 1 9 は、ステップ S 5 7 の処理として C M 開始フレーム n_s を検索し、次に、ステップ S 5 8 の処理として C M 終了フレーム n_e の検索を行い、更に、ステップ S 5 9 の処理として C M 開始時刻 T_s を計算し、ステップ S 6 0 として C M 長さ W を計算する。

30

【 0 1 0 6 】

C M 候補検出器 1 9 は、以上のステップ S 5 7 乃至 S 6 0 の検索及び計算を行った後、ステップ S 6 1 において後述する C M 候補テーブルを参照し、もし、C M 開始時刻 T_s 及び C M 長さ T_w の一致する候補がすでに当該 C M 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 S 5 4 乃至 S 5 6 の処理に戻り、逆に存在しない場合には、新たな C M 候補として C M 候補テーブルに追加した後、再びステップ S 5 4 乃至 S 5 6 の処理に戻る。

40

【 0 1 0 7 】

ステップ S 5 4 乃至 S 5 6 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 S 5 0 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。

【 0 1 0 8 】

なお、上記 C M 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、平均自乗振幅 $A[n]$ が自乗振幅の閾値 A_{thsd} を越える最初のフレーム番号である。また、C M 終了フレーム n_e とは、0 番目のフレームより過去の方角に向かって、平均自乗振幅 $A[n]$ が自乗振幅の閾値 A_{thsd} を越えない最後のフレーム番号である。さらに C M 開始時刻 T_s は、C M 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 C M 長さ T_w は、 $T_w = T[n_e] - T[n_s]$ として求められる。

50

【 0 1 0 9 】

ここで、図 6 に、上記必須条件の算出例を示す。この図 6 に示す $A[n]$ の項において、「 o 」は自乗振幅の閾値 A_{thsd} 未満の平均自乗振幅を持つフレームを示し、「 x 」は自乗振幅の閾値 A_{thsd} 以上の平均自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。

【 0 1 1 0 】

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。

10

【 0 1 1 1 】

図 7 には、CM 候補テーブルの構成例を示す。この図 7 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 乃至 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 19 a 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 乃至 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

20

【 0 1 1 2 】

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 19 a は、付加条件算出器 20 に送られる。

【 0 1 1 3 】

付加条件算出器 20 では、CM 候補テーブル 19 a 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 乃至 Q_{11} を抽出し、それを CM 候補テーブル 19 a に追加記述し、CM 候補テーブル 20 a として付加条件判定器 21 に出力する。

30

【 0 1 1 4 】

図 8 には、当該付加条件算出器 20 における特徴量 Q_1 乃至 Q_{11} の算出例を示す。

【 0 1 1 5 】

この図 8 において、横軸はフレーム番号 (離散時間に相当) を表し、図 8 (A) はカットチェンジ検出力 $C[n]$ 、図 8 (B) は放送モード信号 2 c の離散化数値 $B[n]$ 、図 8 (C) は音声信号の短時間離散スペクトル $S[k, n]$ 、図 8 (D) は音声信号の平均自乗振幅 $A[n]$ 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 8 (A) において、図中 CT で示す位置はカットチェンジ検出力 $C[n]$ 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 8 (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 8 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 8 (D) において、図中 AM は自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 乃至 Q_{11} が計算される場所を示している。

40

【 0 1 1 6 】

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

【 0 1 1 7 】

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 8 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。

50

【 0 1 1 8 】

特徴量 Q_2 は後ブレイク長である。当該後ブレイク長とは、C M 候補区間直後の小音量区間（後ブレイク区間と称する）、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 8 中の一点鎖線で挟まれた区間長 $A B$ が後ブレイク長 Q_2 である。

【 0 1 1 9 】

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

【 0 1 2 0 】

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

10

【 0 1 2 1 】

特徴量 Q_5 は左右相関値である。当該左右相関値 Q_5 は、C M 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 (5) 乃至式 (7) の $A_{LR}[n]$ 、 $A_{LL}[n]$ 、 $A_{RR}[n]$ を利用して、式 (1 4) に従って算出することができる。

【 0 1 2 2 】

【 数 7 】

$$Q_5 = \frac{\sum_{n=n_s}^{n_e-1} A_{LR}[n]}{\sum_{n=n_s}^{n_e-1} A_{LL}[n] \sum_{n=n_s}^{n_e-1} A_{RR}[n]} \quad (14)$$

20

【 0 1 2 3 】

この式 (1 4) 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 1 2 4 】

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（平均自乗振幅）である。これは、式 (1 5) により計算することができる。

30

【 0 1 2 5 】

【 数 8 】

$$Q_6 = \sqrt{\frac{1}{n_e - n_s} \sum_{n=n_s}^{n_e-1} A[n]^2} \quad (15)$$

【 0 1 2 6 】

この式 (1 5) 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。

40

【 0 1 2 7 】

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 1 2 8 】

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。

【 0 1 2 9 】

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

50

にある有音区間もＣＭ候補であるかどうかを表わし、両側ともＣＭ候補であれば「２」、片側のみＣＭ候補であれば「１」、どちらもＣＭ候補でなければ「０」の値をとる。この演算は、ＣＭ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 と後ブレーク長 Q_2 の和($T_s + T_w + Q_2$)が、他のＣＭ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 と前ブレーク長 Q_1 の差($T_s - Q_1$)が、他のＣＭ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。

【０１３０】

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(１６)乃至(２１)に従って計算される。

10

【０１３１】

【数９】

$$S_1[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n'_e - n] \quad (16)$$

$$S_2[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n_s + n] \quad (17)$$

$$S_3[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n_e - n] \quad (18)$$

20

$$S_4[k] = \frac{1}{N} \sum_{n=0}^{N-1} F[k; n'_s + n] \quad (19)$$

$$Q_{10} = \frac{1}{S_{norm}^2} \sum_k \left(S_2[k] - S_1[k] \right)^2 \quad (20)$$

$$Q_{11} = \frac{1}{S_{norm}^2} \sum_k \left(S_4[k] - S_3[k] \right)^2 \quad (21)$$

30

【０１３２】

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８参照)、 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。

【０１３３】

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。

40

【０１３４】

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。

【０１３５】

図９には、付加条件判定器２１の概略構成を示す。

【０１３６】

50

この図 9 において、C M 候補テーブル 2 1 a の各特徴量 Q_1 乃至 Q_L は、それぞれ対応する関数演算器 5 0₁ 乃至 5 0_L に送られ、それぞれ対応するパラメータ変換関数 $H_1()$ 乃至 $H_L()$ 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 5 1₁ 乃至 5 1_L により荷重 W_1 乃至 W_L との積がとられる。各重み付け器 5 1₁ 乃至 5 1_L により重み付けがなされた後の特徴量は、総和加算器 5 2 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 5 2 から出力されたスコア R は、スコア判定器 5 3 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 C M 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 5 3 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 C M 候補は、テーブルから消去される。

【 0 1 3 7 】

10

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 (2 2) に従って行われる。

【 0 1 3 8 】

【 数 1 0 】

$$R = \sum_{l=1}^L W_l H_l(Q_l) \quad (22)$$

【 0 1 3 9 】

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 乃至 1 1 のうちの任意の数である。

20

【 0 1 4 0 】

ここで、各関数演算器 5 0₁ 乃至 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 1 0 (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 (2 3) により例えば Q_1 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 1 4 1 】

$$H_l(Q_l) = \text{Rect}(Q_l; t_{1l}, t_{2l}) \quad (23)$$

なお、前記境界付近で滑らかに 0 から 1、1 から 0 へ推移させるようにする場合には、例えば式 (2 4) 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。

30

【 0 1 4 2 】

【 数 1 1 】

$$H_l(Q_l) = \text{Sigm}(Q_l; t_{1l}, t_{2l}) = \frac{1}{1 + \exp\left(-\frac{(Q_l - t_{1l})}{\sigma_{1l}}\right)} \cdot \frac{1}{1 + \exp\left(\frac{(Q_l - t_{2l})}{\sigma_{2l}}\right)} \quad (24)$$

【 0 1 4 3 】

図 1 0 (C) にシグモイド関数の概形を示す。ただし、 s_{1l} 、 s_{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に基づき決定しておく。

40

【 0 1 4 4 】

また、上記各重み付け器 5 1₁ 乃至 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 乃至 1 1 のうちの任意の数である。

【 0 1 4 5 】

さらに、上記判定器 5 3 におけるスコア判定は、式 (2 5) のように、スコア R の閾値処理により行う。

50

【 0 1 4 6 】

$$Z = \text{Unit}(R - t_r) \quad (25)$$

ただし、 $\text{Unit}(x)$ は、図 10 (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 1 4 7 】

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。

10

【 0 1 4 8 】

以下、ルール判定器 22 の動作を図 11 のフローチャートを用いて説明する。

【 0 1 4 9 】

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、判定すべき CM 候補を選択する。この選択は、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われる。 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。

【 0 1 5 0 】

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、他の CM 候補が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。

20

【 0 1 5 1 】

一方、ステップ S71 において、他の CM 候補が存在すると判定された場合（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。

30

【 0 1 5 2 】

図 12 を用いて、この最小長さ優先ルールの一例を説明する。

【 0 1 5 3 】

なおこの例には、図 12 (A) のように、実際には 4 つの CM 1 乃至 CM 4 が連続して放送されている区間に対し、図 12 (B) 中 A 乃至 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。

【 0 1 5 4 】

まず、図 12 (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 12 (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 12 (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 12 (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

【 0 1 5 5 】

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

50

。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。

【 0 1 5 6 】

図 1 1 に戻り、ステップ S 7 2 のルール判定の後、ルール判定器 2 2 の処理は、ステップ S 7 3 に進む。ステップ S 7 3 に進むと、ルール判定器 2 2 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 S 7 3 において、判定中のＣＭが棄却されと判断した場合（Ｙｅｓ）、ルール判定器 2 2 は、その候補をＣＭ候補テーブルから消去し、ステップ S 7 0 に戻る。一方、ステップ S 7 3 において、判定中のＣＭが棄却されていないと判断した場合（Ｎｏ）、ルール判定器 2 2 は、ステップ S 7 4 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。

10

【 0 1 5 7 】

このステップ S 7 4 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 2 2 は、ステップ S 8 0 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 S 7 4 にて他のＣＭ候補が存在すると判断した場合（Ｙｅｓ）、ルール判定器 2 2 の処理は、ステップ S 7 5 に進む。

【 0 1 5 8 】

ステップ S 7 5 に進むと、ルール判定器 2 2 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。

20

【 0 1 5 9 】

図 1 3 を用いて、当該隣接優先ルールについて説明する。

【 0 1 6 0 】

なおこの例には、図 1 3（Ａ）のように、実際には４つのＣＭ 1 1 乃至ＣＭ 1 4 が連続して放送されている区間に対し、図 1 3（Ｂ）中 I 乃至 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。

【 0 1 6 1 】

30

このような例において、まず、図 1 3（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 1 3（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 1 3（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。

【 0 1 6 2 】

以上により、この図 1 3 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。

40

【 0 1 6 3 】

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。

【 0 1 6 4 】

図 1 1 に戻り、ステップ S 7 5 の処理後、ルール判定器 2 2 の処理は、ステップ S 7 6 に進む。ステップ S 7 6 に進むと、ルール判定器 2 2 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 S 7 6 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 2 2 は、その候補をＣＭ候補テーブルから消去し、ステップ S 7 0 の処理に戻る。一方、ステップ S 7 6 において棄却されて

50

いないと判定された場合 (No)、ルール判定器 22 は、次のステップ S 77 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。

【0165】

このステップ S 77 において、他の CM 候補が存在しなと判定された場合 (No)、ルール判定器 22 は、ステップ S 80 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 S 77 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 S 78 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。

10

【0166】

図 14 を用いて、当該スコア優先ルールについて説明する。

【0167】

なおこの例には、図 14 (A) のように、実際には 4 つの CM 21 乃至 CM 24 が連続して放送されている区間に対し、図 14 (B) 中 P 乃至 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。

【0168】

この例において、まず、図 14 (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。

20

【0169】

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 14 (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。

30

【0170】

しかしながら、図 14 (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。

【0171】

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 14 (F) に示すように、Q と競合関係にある候補 U を棄却する。

【0172】

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。

40

【0173】

なお、仮に、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。

【0174】

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。

50

【 0 1 7 5 】

図 1 1 に戻り、ステップ S 7 8 の処理後、ルール判定器 2 2 の処理は、ステップ S 7 9 に進む。ステップ S 7 9 に進むと、ルール判定器 2 2 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 S 7 9 において、判定中の候補が棄却されたと判断した場合 (Y e s)、ルール判定器 2 2 は、その候補を C M 候補テーブルより消去し、ステップ S 7 0 に戻る。一方、ステップ S 7 9 において、棄却されなかった場合、ルール判定器 2 2 は、ステップ S 8 0 の C M 検出出力として、開始時刻とその長さを出力し、C M 候補テーブルから消去した後、ステップ S 7 0 に戻る。

【 0 1 7 6 】

以上説明したように、本実施の形態の第 1 の具体例の C M 検出部 4 においては、ほぼ全ての C M が満足する必須条件に基づき、決定論的に番組中から C M の候補を抽出し、C M 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 C M 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 T V 放送の映像音声記録装置を例にとったが、デジタル T V 放送等に適用される場合にも同様の C M 検出部 4 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 C M 検出部 4 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。

【 0 1 7 7 】

次に、本発明の第 2 の具体例としての C M 検出部 4 について以下に説明する。

【 0 1 7 8 】

図 1 5 には、本発明の第 2 の具体例としての C M 検出部 4 の詳細な構成を示す。

【 0 1 7 9 】

当該第 2 の具体例の C M 検出部 4 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) 乃至 1 4) の付加条件のうち、1 1) 乃至 1 4) 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) 。

【 0 1 8 0 】

この第 2 の具体例の C M 検出部 4 も、図 2 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 1 5 において、図 2 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。

【 0 1 8 1 】

以下、図 1 5 の構成において、新たに追加された各構成要素 (1 0 1 , 1 0 2 , 1 0 3) と、付加条件算出器 2 0 において新たに拡張された機能についてのみ説明する。

【 0 1 8 2 】

フロントエンド部に設けられた音源識別器 1 0 1 は、デジタル化及びフレーム化された音声信号 2 b を入力とし、この音声信号 2 b 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、「V Q 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-II、No.3、529/537(1998)に記載された技術、安部による、特願平 1 1 - 1 9 0 6 9 3 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。

【 0 1 8 3 】

この音源識別器 1 0 1 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 1 8 に入力される。

【 0 1 8 4 】

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 1 0 2 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、テレビ番組表などが

10

20

30

40

50

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 1 1 - 1 9 0 6 9 3 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。

【 0 1 8 5 】

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 1 8 6 】

フロントエンド部のその他の各構成要素は、図 2 の例と同様である。

【 0 1 8 7 】

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 (1 0) に示した特徴量 $G[n]$ が、式 (2 6) のように拡張されることになる。

【 0 1 8 8 】

$G[n] = \{ C[n], A[n], A_{LR}[n], A_{LL}[n], A_{RR}[n], F[k;n], B[n], T[n], U[n], W[n] \}$ (26)

バックエンド部の CM 検出器 1 9 は、前述の図 2 のものと同様のものであるが、当該第 2 の具体例の場合、CM 候補テーブル 1 9 a 乃至 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 乃至 2 1 a は、前述した Q_1 から Q_{11} までの特徴量に加え、図 1 6 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 1 6 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。

【 0 1 8 9 】

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 1 9 0 】

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 1 9 1 】

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 (2 7) に従って検出される。

【 0 1 9 2 】

【 数 1 2 】

$$Q_{12} = \begin{cases} 1 & \left(\text{if } {}^3U[n] = (\text{音声}) \text{ or } {}^3U[n] = (\text{音声} + \text{音楽}) \text{ for } n_s \leq n < n_e \right) \\ 0 & (\text{otherwise}) \end{cases} \quad (27)$$

【 0 1 9 3 】

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 (2 8) に従って検出される。

【 0 1 9 4 】

【 数 1 3 】

$$Q_{13} = \begin{cases} 1 & \left(\text{if } {}^3U[n] = (\text{音声}) \text{ or } {}^3U[n] = (\text{音声} + \text{音楽}) \text{ for } n_s \leq n < n_e \right) \\ 0 & (\text{otherwise}) \end{cases} \quad (28)$$

10

【 0 1 9 5 】

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 1 9 6 】

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 1 9 7 】

付加条件判定器 21 以降は、変数としての特徴量 Q_{12} 乃至 Q_{15} が拡張されるだけであり、
前述の図 2 の C M 検出部 4 の場合と同様であるため、説明を省略する。

20

【 0 1 9 8 】

この場合の C M 検出部 4 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 1 9 9 】

C M 検出部 4 の第 3 の具体例として、例えば、図 17 に示すように、小振幅回数、小振幅区間、および信号分散を、それぞれ特徴量 Q_{16} 乃至 Q_{18} として、付加条件算出器 20 により算出させるようにすることができる。

【 0 2 0 0 】

小振幅回数とは、音声信号の振幅が、予め設定されている所定の閾値を下回る回数を意味する。付加条件算出器 20 は、例えば、図 18 のフローチャートに示すような処理を行うことで、小振幅回数を計算する。

30

【 0 2 0 1 】

最初にステップ S90 において、カウンタ C とフラグ F がリセットされる。カウンタ C には、小振幅回数が保持され、フラグ F は、小振幅区間であることを表す。ステップ S90 においては、さらに、時刻 n が C M 候補の開始時刻に設定される。

【 0 2 0 2 】

ステップ S91 において、音声信号の振幅信号 $A[n]$ が取得される。 n は、離散化された時刻に対応する。

40

【 0 2 0 3 】

ステップ S92 において、いまフラグ F がセットされているか否かが判定され、セットされていない場合、ステップ S93 に進み、ステップ S91 で取得された信号振幅 $A[n]$ が、予め設定されている所定の閾値 A_1 より小さいか否かが判定される。音声信号の振幅 $A[n]$ の値が閾値 A_1 より等しいか、それより大きいと判定された場合、ステップ S97 に進み、時刻 n の値がインクリメントされる。そして、ステップ S98 において、時刻 n の値が C M 候補の終了時刻に達したか否かが判定され、終了時刻に達していない場合には、ステップ S91 に戻り、次の時刻のタイミングにおける振幅 $A[n]$ が取得される。

【 0 2 0 4 】

以上のような処理がステップ S93 において、振幅 $A[n]$ の値が閾値 A_1 より小さいと

50

判定されるまで繰り返し実行される。ステップ S 9 3 において、振幅 A [n] の値が閾値 A 1 より小さいと判定された場合、ステップ S 9 4 に進み、カウンタ C の値が 1 だけインクリメントされ、かつ、フラグ F がセットされる。

【 0 2 0 5 】

その後、ステップ S 9 7 に進み、時刻 n がインクリメントされ、ステップ S 9 8 において、インクリメントされた時刻 n の値が C M 候補の終了時刻に達しているか否かが判定され、まだ達していない場合には、ステップ S 9 1 に戻り、次のタイミングの振幅 A [n] が取得される。

【 0 2 0 6 】

そして、ステップ S 9 2 において、フラグ F がセットされているか否かが判定され、いまの場合、フラグ F がセットされているので、ステップ S 9 5 に進み、ステップ S 9 1 で取得された振幅 A [n] の値が、予め設定されている閾値 A 2 より大きいか否かが判定される。なお、この閾値 A 2 の値は、ステップ S 9 3 において比較される閾値 A 1 より大きい値 ($A 2 > A 1$) とされている。

10

【 0 2 0 7 】

ステップ S 9 5 において、振幅 A [n] の値が閾値 A 2 より大きくないと判定された場合、ステップ S 9 7 に進み、時刻 n の値がインクリメントされる。

【 0 2 0 8 】

ステップ S 9 8 において、時刻 n の値が C M 候補の終了時刻に達しているか否かが再び判定され、達していない場合には、ステップ S 9 1 に戻り、次のタイミングの振幅 A [n] が取得される。

20

【 0 2 0 9 】

ステップ S 9 2 において、フラグ F がセットされているか否かが再び判定され、いまの場合、まだセットされているので、ステップ S 9 5 に進み、取得された振幅 A [n] が閾値 A 2 より大きくないと判定された場合、ステップ S 9 7 に進み、上述した場合と同様の処理が繰り返し実行される。

【 0 2 1 0 】

以上のようにして、振幅 A [n] の値が、より小さい閾値 A 1 より小さいと判定された場合、カウンタ C の値が 1 だけインクリメントされた後、振幅 A [n] の値が、閾値 A 1 より若干大きい値の閾値 A 2 より大きくなるまで待機する。

30

【 0 2 1 1 】

ステップ S 9 5 において、振幅 A [n] の値が閾値 A 2 より大きいと判定された場合、ステップ S 9 6 に進み、フラグ F がリセットされる。その後ステップ S 9 7 に進み、時刻 n の値がインクリメントされる。ステップ S 9 8 において、時刻 n の値が C M 候補の終了時刻に達したか否かが判定され、達していない場合には、ステップ S 9 1 に戻り、それ以降の処理が繰り返し実行される。

【 0 2 1 2 】

ステップ S 9 8 において、時刻 n の値が C M 候補の終了時刻に達したと判定された場合、処理は終了される。

【 0 2 1 3 】

40

ステップ S 9 5 において基準とされる閾値 A 2 の値を、ステップ S 9 3 の処理において設定される閾値 A 1 より大きく設定することで、判定処理に、いわゆるヒステリシス特性を持たせることが可能となる。すなわち、振幅 A [n] の値が、より小さい閾値 A 1 より小さくなったとき、小振幅区間に入ったと判定されるが、振幅 A [n] の値が、閾値 A 1 より若干大きくなっても、閾値 A 2 より小さい場合には、まだ小振幅期間中であると判定され、閾値 A 1 より大きい閾値 A 2 よりさらに大きくなったとき、初めて、小振幅期間が終了したと判定される。これにより、小振幅期間中のわずかな振幅の変化に起因して、小振幅回数が必要以上に大きな値にカウントされることが防止される。

【 0 2 1 4 】

以上のようにして、C M 候補の期間における小振幅の回数がカウンタ C に設定され、この

50

カウンタCの値が特徴量 Q_{16} として出力される。

【0215】

図19は、小振幅回数の具体的な計測結果の例を表している。図19における横軸は、小振幅回数を表し、縦軸は、相対度数を表している。図19(A)のグラフは、実験データから得られた490個のCM候補のうち、実際にCMであった352個の度数分布を表しており、図19(B)は、そのうちのCMでなかった138個の度数分布を表している。すなわち、図19(B)は、本編中で偶然、音量条件や映像条件が満たされたためにCM候補として検出されたものである。

【0216】

これらの図を比較して明らかなように、CMである場合(図19(A))、小振幅回数は、0回から2回に集中するのに対して、CMでない場合には(図19(B))、小振幅回数は、7回乃至9回と多くなることがわかる。

10

【0217】

次に、図20のフローチャートを参照して、CM検出部4の付加条件算出器20が実行する小振幅区間長計算処理について説明する。最初に、ステップS110において、小振幅区間長を表すカウンタDの値がリセットされ、かつ、時刻nの値がCM候補の開始時刻にセットされる。

【0218】

次にステップS111において、音声信号の振幅 $A[n]$ が取得され、ステップS112において、ステップS111で取得された振幅 $A[n]$ の値が、予め設定されている所定の閾値 A_1 より小さいか否かが判定される。この閾値 A_1 は、図18のステップS93における閾値 A_1 と等しい値とされているが、異なる値とすることも可能である。

20

【0219】

ステップS112において、振幅 $A[n]$ の値が閾値 A_1 と等しいか、それより大きいと判定された場合、ステップS114に進み、時刻nの値がインクリメントされる。そして、ステップS115において、インクリメントされた時刻nの値が、CM候補の終了時刻に達したか否かが判定され、終了時刻に達していない場合には、ステップS111に戻り、次のタイミングの振幅 $A[n]$ が取得される。

【0220】

そして、その振幅 $A[n]$ の値が、ステップS112において、閾値 A_1 より小さいか否かが再び判定され、振幅 $A[n]$ の値が閾値 A_1 より小さくない場合には、ステップS114に進み、時刻nの値がさらにインクリメントされる。

30

【0221】

以下、同様の処理が繰り返し実行され、ステップS112において、振幅 $A[n]$ の値が、閾値 A_1 と等しいか、それより大きいと判定された場合、ステップS113に進み、カウンタDの値が1度だけインクリメントされる。その後、ステップS114に進み、時刻nの値がインクリメントされる。ステップS115において、時刻nの値が、CM候補の終了時刻に達したか否かが判定され、達していない場合には、ステップS111に戻り、次のタイミングの振幅 $A[n]$ が取得され、ステップS112において、その振幅 $A[n]$ の値が、閾値 A_1 より小さいか否かが再び判定される。振幅 $A[n]$ の値が閾値 A_1 より小さい場合には、ステップS113において、カウンタDの値が再び1度だけインクリメントされる。

40

【0222】

以上のような処理が繰り返し実行されることで、カウンタDの値は、振幅 $A[n]$ の値が閾値 A_1 より小さい期間に対応する値となる。

【0223】

ステップS115において、時刻nの値がCM候補の終了時刻に達したと判定された場合、ステップS116に進み、カウンタDの値が規格化される。すなわち、カウンタDの値は、サンプリング周波数 f_s で割算されることで、規格化され、その値が特徴量 Q_{17} として出力される。

50

【 0 2 2 4 】

図 2 1 は、図 1 9 における場合と同様に、4 9 0 個の C M 候補のうちの 3 5 2 個の実際の C M と、1 3 8 個の C M でなかった場合の小振幅区間長の総和を表している。図 2 1 において、横軸は、小振幅区間長の総和（単位は秒）を表し、縦軸は、相対度数を表している。図 2 1 (A) と図 2 1 (B) を比較して明らかなように、C M の小振幅区間長の総和は、2 0 ms 程度以下に集中している（図 2 1 (A) ）のに対して、C M でない場合には、1 . 0 s 以上の長さに集中している（図 2 1 (B) ）。

【 0 2 2 5 】

さらに、付加条件算出器 2 0 は、音声信号の振幅の分散を式 (2 9) に基づいて演算する。この式 (2 9) において、s は、C M 候補の離散開始時刻を表し、e は、C M 候補の離散終了時刻を表し、v は、信号の分散を表す。付加条件算出器 2 0 は、この値 v をそのまま特徴量 Q_{18} として出力するか、あるいはその平方根を取って、標準偏差を特徴量 Q_{18} として出力する。あるいはまた、付加条件算出器 2 0 は、標準偏差を平均値で割算し、相対標準偏差を特徴量 Q_{18} として出力することができる。

10

【 0 2 2 6 】

【 数 1 4 】

$$v = \frac{1}{e-s} \sum_{n=s}^{e-1} A^2[n] - \left(\frac{1}{e-s} \sum_{n=s}^{e-1} A[n] \right)^2 \quad (29)$$

20

【 0 2 2 7 】

図 2 2 は、振幅の分散の例を表している。図 2 2 (A) は、4 9 0 個の C M 候補のうちの、3 5 2 個の実際に C M であった場合の振幅の分散を表しており、図 2 2 (B) は、1 3 8 個の C M でなかった場合の分散を表している。なお、図 2 2 において、横軸は、相対標準偏差を表しており、縦軸は、相対度数を表している。

【 0 2 2 8 】

これらの図を比較して明らかなように、C M の場合（図 2 2 (A) ）、振幅の相対標準偏差がほぼ 0 . 6 以下に集中しているのに対して、C M でない場合には（図 2 2 (B) ）、0 . 7 以上となることが多いことがわかる。

30

【 0 2 2 9 】

従って、小振幅回数、小振幅区間長、および振幅分散を特徴量として利用することで、より正確に C M を検出することが可能となる。

【 0 2 3 0 】

次に、本発明の第 2 の実施の形態としての映像音声記録装置について以下に説明する。

【 0 2 3 1 】

図 2 3 には、第 2 の実施の形態の映像音声記録装置の概略構成を示す。

【 0 2 3 2 】

なお、この図 2 3 において、図 1 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。また、第 2 の実施の形態の映像音声記録装置の場合の C M 検出部 4 は、前記第 1 の具体例、第 2 の具体例、および第 3 の具体例の何れも適用できる。

40

【 0 2 3 3 】

以下、図 2 3 の構成において、新たに追加された各構成要素 (1 1 0 , 1 1 1) と、C M 検出部 4 において新たに拡張された機能についてのみ説明する。

【 0 2 3 4 】

先ず、この第 2 の実施の形態の映像音声記録装置における C M 検出部 4 は、前述のように、式 (1 0) に示される各 C M 候補の特徴量 $G[n]$ を、内部で算出している。また、当該第 2 の実施の形態の場合、C M 検出部 4 は、最終的に C M として検出されたものに関して

50

、その開始時刻及び時間長と共に、CM開始フレーム $n = n_s$ から終了フレーム $n = n_e$ に渡って、 $G[n]$ をCMデータベース110に出力するように機能が拡張されている。

【0235】

CMデータベース110は、上記検出されたCMに関して、その開始時刻、時間長、特徴量 $G[n]$ を保存する。

【0236】

CM特徴量比較器111は、ユーザが入力する検索指令に基づき、データベース110に保存されている全て又は一部のCMから、ユーザが指定したCMと同じCMを抽出し、CM検索出力111aとして出力する。

【0237】

このCM特徴量比較器111の動作を、図24を用いて説明する。

【0238】

先ず例えば、ユーザは、映像音声記録部5による映像信号及び音声信号を視聴することにより、検索したいCMを選択したとする。このとき、CM特徴量比較器111には、ステップS120として、上記ユーザによる選択に応じた検索指令が入力されることになる。

【0239】

このとき、CM特徴量比較器111は、ステップS121の処理として、上記入力された検索指令に基づいて、CMデータベース110から、その検索指令に該当するCMの特徴量 $G[n]$ を取得する。

【0240】

続いて、CM特徴量比較器111は、ステップS122としてCMデータベース110より、検索される候補CMを一つ選択し、さらにステップS123として、その候補CMに対応する特徴量 $G'[n]$ を取得する。

【0241】

次に、CM特徴量比較器111は、ステップS124として、上記選択されたCMについて、式(30)の計算を行い、それを予め定める所定の閾値 J_{thsd} と比較する。

【0242】

【数15】

$$J(G, G') = \sum_R \left| G[n] - G[n'] \right|^2 \quad (30)$$

【0243】

ここで、このステップS124において、 $J(G, G') < J_{thsd}$ と判定したならば(Yes)、CM特徴量比較器111は、ステップS125に進み、特徴量が一致したとして検索結果を出力し、ステップS122に戻って再び他の候補CMについて同様の処理を行う。一方、ステップS124において、 $J(G, G') < J_{thsd}$ でないと判定した場合(No)、CM特徴量比較器111は、特徴量が一致しなかったとして、ステップS122に戻り、再び他の候補CMについて同様の処理を行う。

【0244】

以上により、CM特徴量比較器111では、ユーザにより指定されたCMと同じCMを、映像音声記録部5に記録されているデータの中から検索することができる。

【0245】

次に、図25には、上述した図2や図15に示したCM検出部4を実装する場合のハードウェア構成の一例を示す。

【0246】

この図25において、A/D変換器40は、前記図2や図15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。

【0247】

A/VプロセッサまたはDSP(デジタルシグナルプロセッサ)42は、前記カッチ

10

20

30

40

50

エンジン検出器 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。

【0248】

プロセッサ 44は、前記CM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、CM確率データベース 103等の機能を備えている。

【0249】

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DSP（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。

【0250】

以上説明したような本発明の各実施の形態によれば、TV放送信号からCM部分を正確に検出可能とすることにより、例えばCMを不要としている視聴者や、CMを必要としている視聴者の双方に対して利便を図ることが可能となる。すなわち例えばCM部分を不要とする場合、テレビ放送信号からCM部分をスキップして視聴可能とする装置を実現でき、これは例えば番組本編のみを連続視聴する要求に対して有用な装置となる。また例えば、CM部分のみを必要とする場合、TV放送信号からCM部分のみを視聴できる装置を実現でき、これは例えばCMのみを連続視聴する要求に対して有用な装置となる。

【0251】

さらに、TV放送信号からCM部分を正確に検出可能とすることにより、例えば特定のCMの放送状況を調査する場合などにも有用となる。

【0252】

【発明の効果】

本発明の信号処理装置及び方法によれば、例えば、TV放送信号に含まれるコマーシャルメッセージ部分を高精度に検出又は検索可能となる。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態の映像音声記録装置の概略構成図である。

【図2】第1の具体例のCM検出部の詳細な構成図である。

【図3】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4】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。

【図5】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6】必須条件の算出例の説明に用いる図である。

【図7】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8】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9】付加条件算出器の構成図である。

【図10】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。

【図11】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12】最小長さ優先ルールの説明に用いる図である。

【図13】隣接優先ルールの説明に用いる図である。

【図14】スコア優先ルールの説明に用いる図である。

【図15】第2の具体例のCM検出部の詳細な構成図である。

【図16】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17】第3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18】CM検出部の付加条件算出器における小振幅回数計算処理を説明するフローチャートである。

【図19】小振幅回数の計算の具体例を示す図である。

10

20

30

40

50

【図20】CM検出部の付加条件算出器における小振幅区間長計算処理を説明するフローチャートである。

【図21】小振幅区間長の具体例を示す図である。

【図22】振幅分散の具体例を示す図である。

【図23】本発明の第2の実施の形態の映像音声記録装置の概略構成図である。

【図24】CM特徴量比較器の動作の流れを示すフローチャートである。

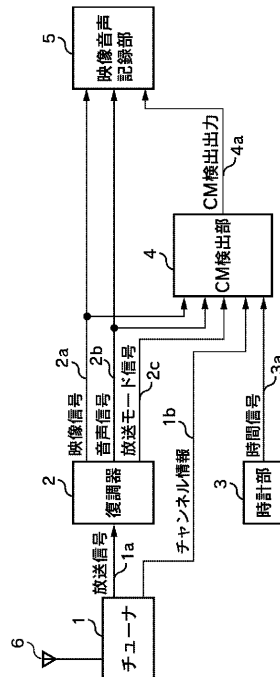
【図25】CM検出部を実装する場合の一例としてのハードウェア構成図である。

【符号の説明】

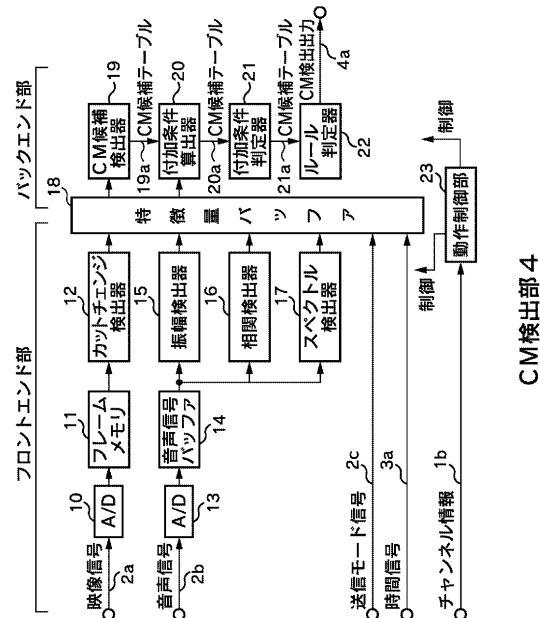
1 チューナ、 2 復調器、 3 時計部、 4 CM検出器、 5 映像音声記録部、 10、13 A/D変換器、 11 フレームメモリ、 14 音声信号バッファ、 12 カットチェンジ検出器、 15 振幅検出器、 16 相関検出器、 17 スペクトル検出器、 18 特徴量バッファ、 19 CM候補検出器、 20 付加条件算出器、 21 付加条件判定器、 22 ルール判定器、 23 動作制御部、 101 音源識別器、 102 番組ジャンルデータまたは番組ジャンル識別器、 103 CM確率データベース、 110 CMデータベース、 111 CM特徴量比較器

10

【図1】

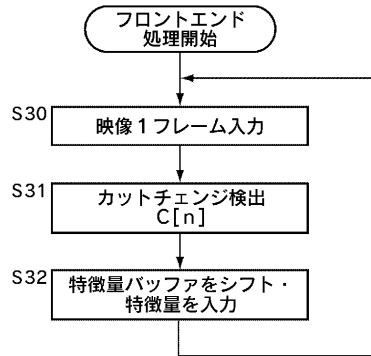


【図2】

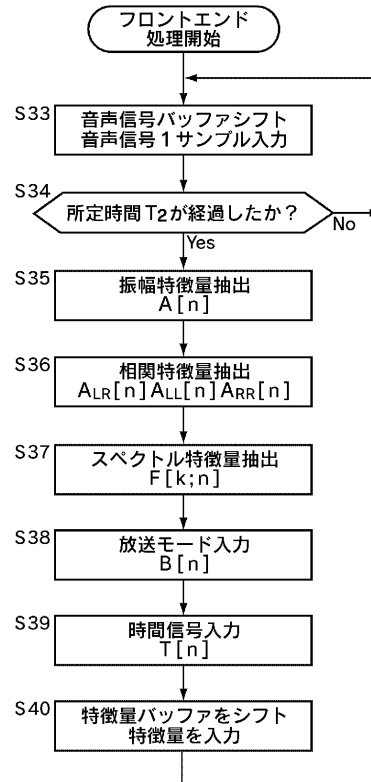


CM検出部4

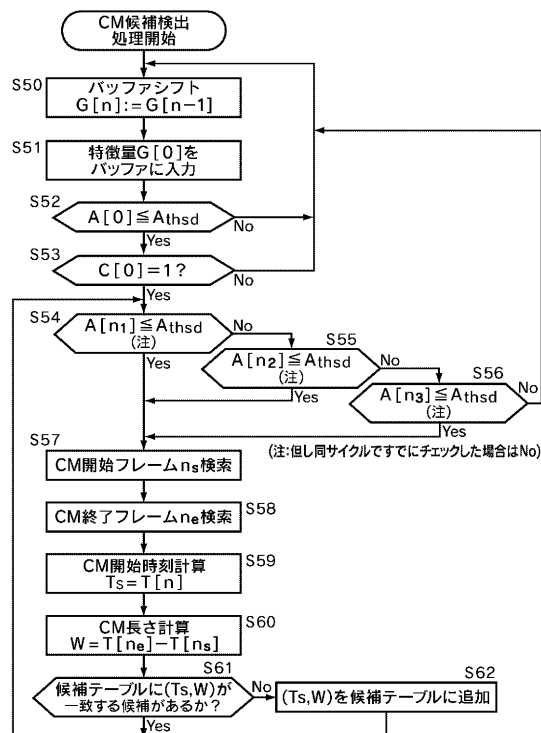
【 図 3 】



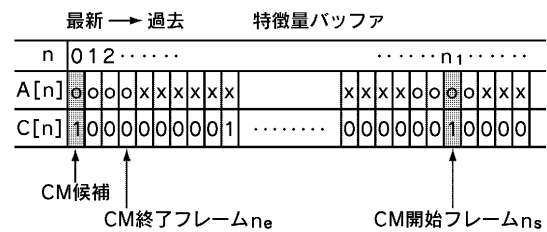
【 図 4 】



【 図 5 】



【 図 6 】

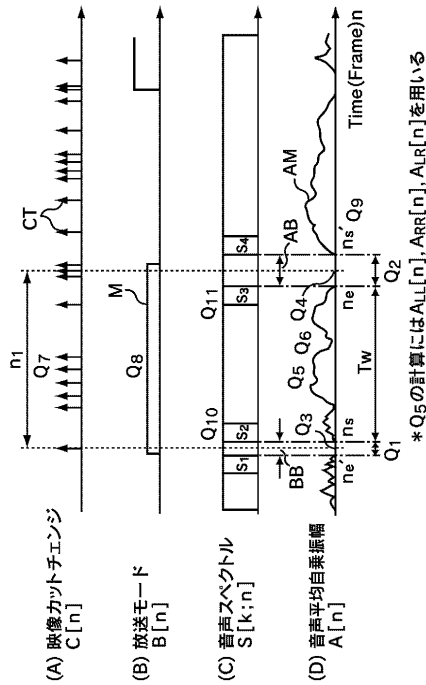


【圖 7】

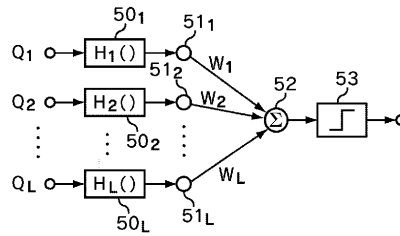
項 目	記号	単位	必須条件例 (19a)	必須条件例 (20a)	判定例 (21a)
開始時刻	Ts	時分秒	1:23'45	1:23'45	1:23'45
長さ(有音長)	Tw	秒	14.63	14.63	14.63
前ブレイク長	Q1	ms	—	300.0	300.0
前ブレイク長	Q2	ms	—	300.0	300.0
前ブレイク最小振幅	Q3	注	—	0.00015	0.00015
後ブレイク最小振幅	Q4	注	—	0.00020	0.00020
左右相関値	Q5	—	—	0.934	0.934
平均振幅値	Q6	注	—	0.010	0.010
カット数	Q7	個	—	9	9
放送モード	Q8	—	—	1	1
隣接候補数	Q9	個	—	2	2
前スペクトル差分エネルギー	Q10	—	—	0.41	0.41
後スペクトル差分エネルギー	Q11	—	—	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 8】

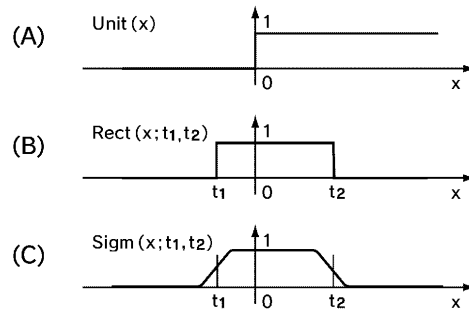


【図 9】

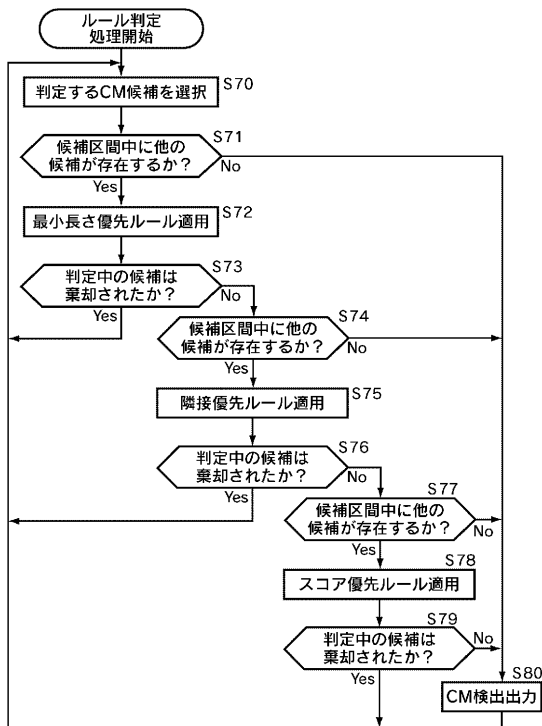


付加条件判定器 21

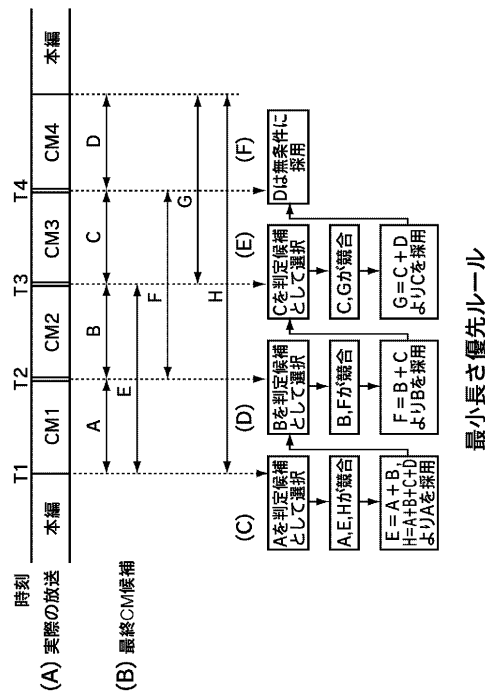
【図 10】

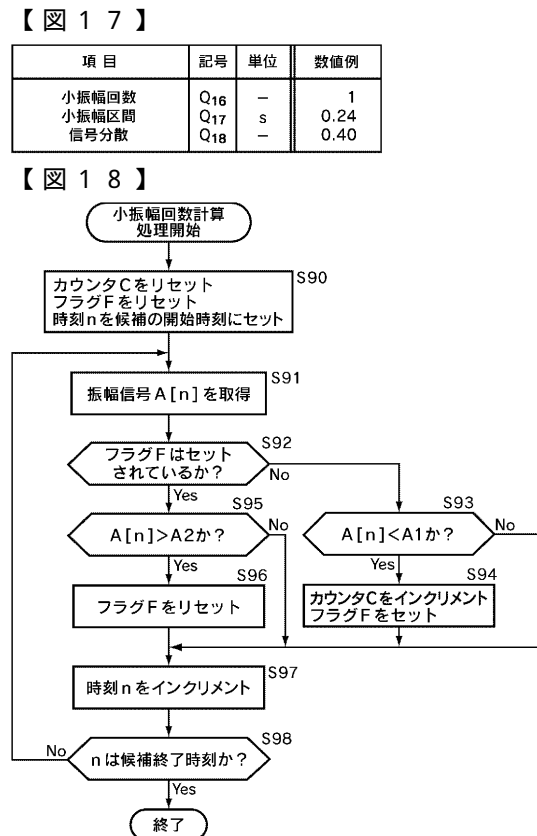
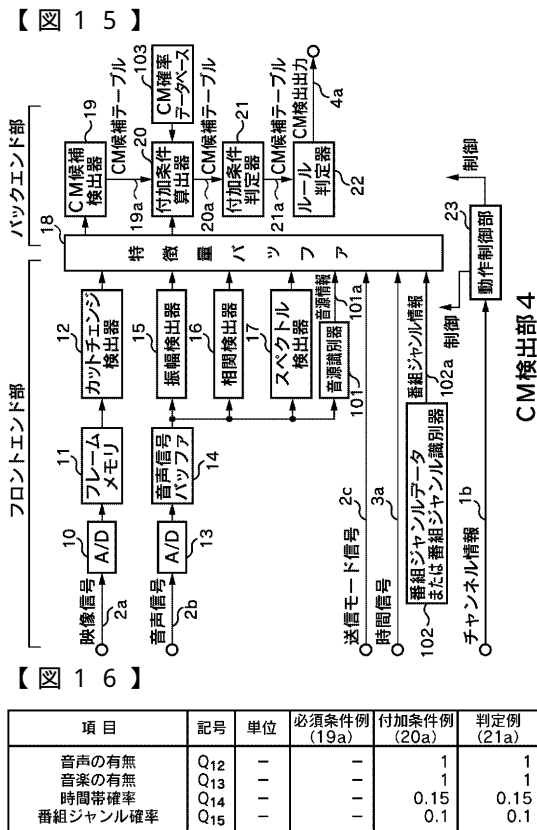
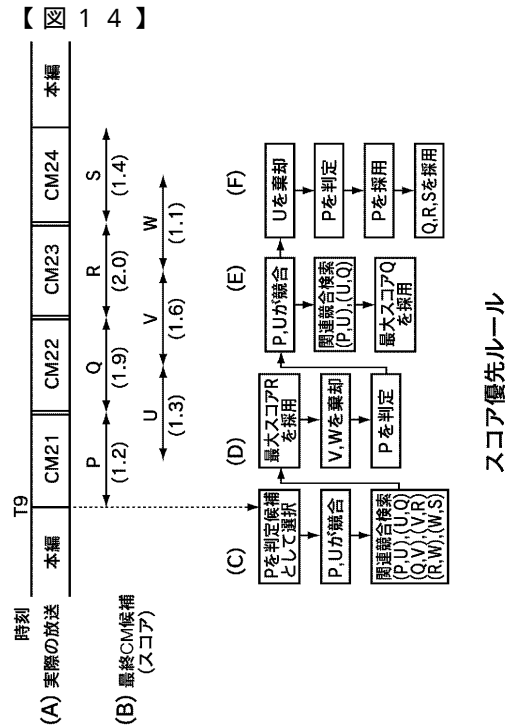
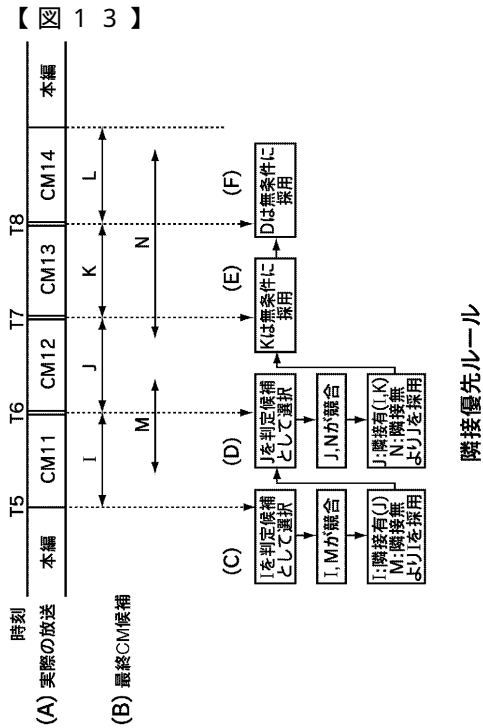


【図 11】

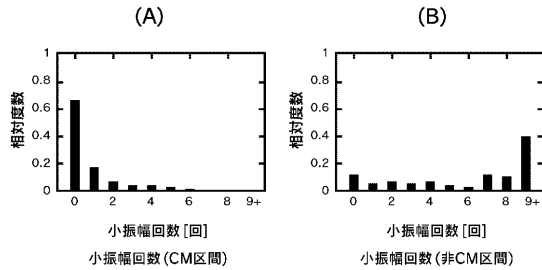


【図 12】

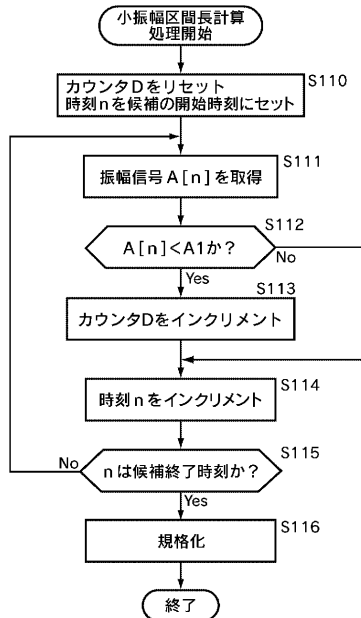




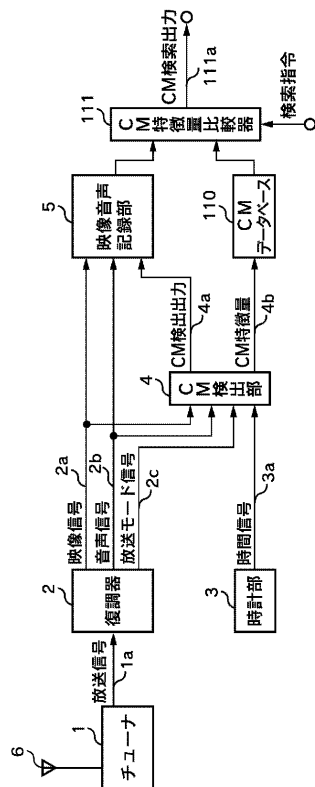
【図 19】



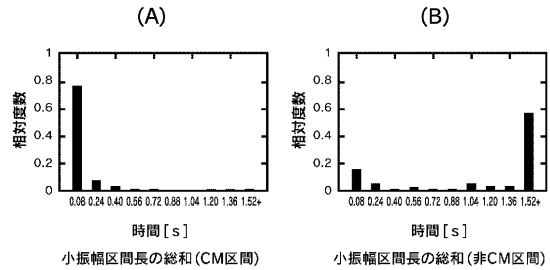
【図 20】



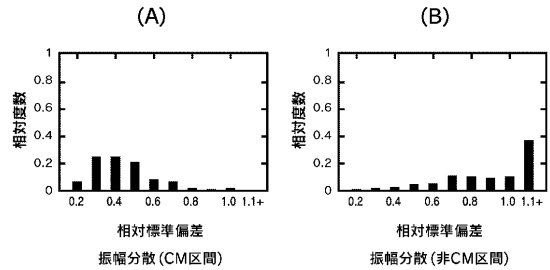
【図 23】



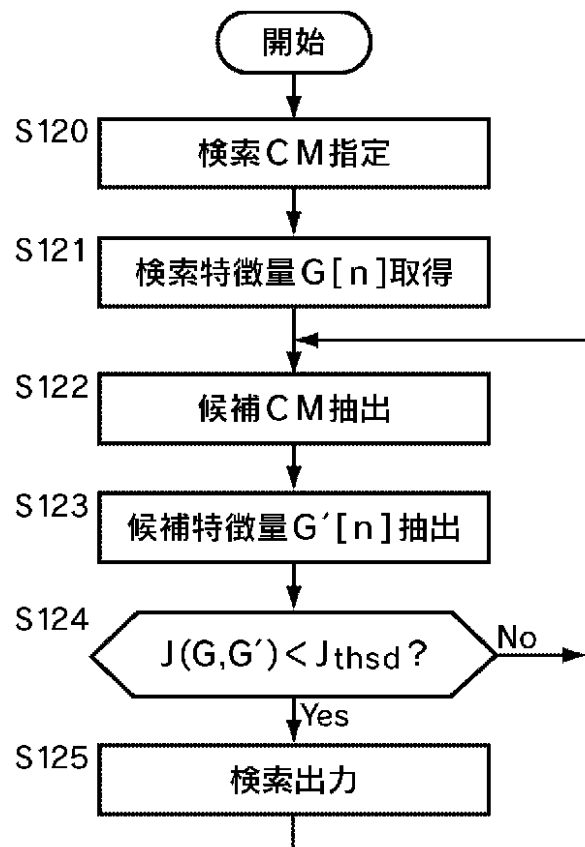
【図 21】



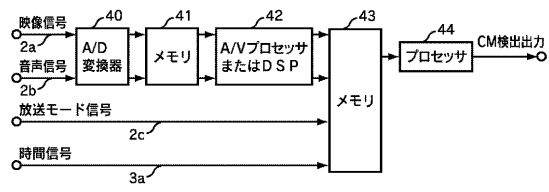
【図 22】



【図 24】



【図 25】



 フロントページの続き

(51)Int.Cl.		F I	
G 1 0 L 15/16	(2006.01)	G 1 0 L 3/00	5 3 9
G 1 0 L 15/02	(2006.01)	G 1 0 L 7/08	A
H 0 4 N 17/00	(2006.01)	H 0 4 N 17/00	N
		G 1 0 L 101:02	
		G 1 0 L 101:065	

(56)参考文献 特開平 1 1 - 0 5 3 7 8 3 (J P , A)
 特開平 0 9 - 3 0 7 8 4 1 (J P , A)
 特開平 1 0 - 2 1 5 4 3 6 (J P , A)
 特開平 1 0 - 2 2 4 7 2 2 (J P , A)
 特開平 0 9 - 3 1 2 8 2 7 (J P , A)
 特開平 1 1 - 0 6 9 3 2 2 (J P , A)
 特開平 1 0 - 2 2 2 9 3 3 (J P , A)
 特開平 0 5 - 0 2 8 7 1 7 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

H04N 5/76-5/956
 G06N 3/00
 G06T 7/00
 H04N 17/00