(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第3840928号
(P3840928)

(45)発行日　平成18年11月1日(2006.11.1)　(24)登録日　平成18年8月18日(2006.8.18)

(51)Int.Cl.		F I		
G10L 11/00	(2006.01)	G10L	11/00	402C
G10L 15/00	(2006.01)	G10L	11/00	403
		G10L	11/00	101E
		G10L	15/00	200G

請求項の数 15　(全 20 頁)

(21)出願番号　特願2001-216716(P2001-216716)
(22)出願日　平成13年7月17日(2001.7.17)
(65)公開番号　特開2003-29772(P2003-29772A)
(43)公開日　平成15年1月31日(2003.1.31)
審査請求日　平成15年2月20日(2003.2.20)

前置審査

(73)特許権者　000002185
ソニー株式会社
東京都品川区北品川６丁目７番３５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　山下　剛史

最終頁に続く

(54)【発明の名称】信号処理装置および方法、記録媒体、並びにプログラム

(57)【特許請求の範囲】

【請求項１】

入力音響信号に対して、前記入力音響信号の振幅を抽出する振幅抽出手段と、

前記振幅抽出手段により抽出された振幅の時間的変化に基づいて、前記入力音響信号の内容の変化点を検出する変化点検出手段と

を備え、

前記振幅抽出手段は、前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、

前記変化点検出手段は、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、

前記スペクトル変化量が所定の閾値より大きい場合に変化点とする

ことを特徴とする信号処理装置。

【請求項２】

前記振幅抽出手段は、周波数分解されてスケールファクタとともに符号化された入力音響信号に対しては、前記スケールファクタを前記振幅として抽出する

ことを特徴とする請求項１に記載の信号処理装置。
【請求項３】
入力音響信号に対して、前記入力音響信号の振幅を抽出する振幅抽出ステップと、
前記振幅抽出ステップの処理で抽出された振幅の時間的変化に基づいて、前記入力音響信号の内容の変化点を検出する変化点検出ステップと
を含み、
前記振幅抽出ステップは、前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、
前記変化点検出ステップは、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、
前記スペクトル変化量が所定の閾値より大きい場合に変化点とする
ことを特徴とする信号処理方法。
【請求項４】
入力音響信号に対して、前記入力音響信号の振幅を抽出する振幅抽出ステップと、
前記振幅抽出ステップの処理で抽出された振幅の時間的変化に基づいて、前記入力音響信号の内容の変化点を検出する変化点検出ステップと
を含み、
前記振幅抽出ステップは、前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、
前記変化点検出ステップは、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、
前記スペクトル変化量が所定の閾値より大きい場合に変化点とする
ことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。
【請求項５】
入力音響信号に対して、前記入力音響信号の振幅を抽出する振幅抽出ステップと、
前記振幅抽出ステップの処理で抽出された振幅の時間的変化に基づいて、前記入力音響信号の内容の変化点を検出する変化点検出ステップと
を含み、
前記振幅抽出ステップは、前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、
前記変化点検出ステップは、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、
前記スペクトル変化量が所定の閾値より大きい場合に変化点とする
ことをコンピュータに実行させることを特徴とするプログラム。

【請求項６】
映像音響信号を入力とする信号処理装置において、
入力音響信号の振幅を抽出し、抽出した前記振幅の時間的変化に基づいて、前記入力音響信号の内容の変化点を検出する音響信号変化点検出手段を備え、
前記変化点検出手段は、
前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、
前記変化量が所定の閾値より大きい場合に変化点とし、
検出された前記音響信号変化点近傍の映像および音声を再生または記録する
ことを特徴とする信号処理装置。
【請求項７】
前記映像音響信号は、テレビジョン放送信号であり、前記所定の内容の信号は、コマーシャルメッセージである
ことを特徴とする請求項６に記載の信号処理装置。
【請求項８】
映像音響信号を入力とする信号処理装置の信号処理方法において、
入力音響信号の振幅を抽出し、抽出した前記振幅の時間的変化に基づいて、前記入力音響信号の内容の変化点を検出する音響信号変化点検出ステップを含み、
前記変化点検出ステップは、
前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、
前記変化量が所定の閾値より大きい場合に変化点とし、
検出された前記音響信号変化点近傍の映像および音声を再生または記録する
ことを特徴とする信号処理方法。
【請求項９】
映像音響信号を入力とする信号処理装置を制御するプロセッサに実行させるプログラムにおいて、
入力音響信号の振幅を抽出し、抽出した前記振幅の時間的変化に基づいて、前記入力音響信号の内容の変化点を検出する音響信号変化点検出ステップを含み、
前記変化点検出ステップは、
前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、

前記変化量が所定の閾値より大きい場合に変化点とし、
検出された前記音響信号変化点近傍の映像および音声を再生または記録する
ことを特徴とするプログラムが記録されている記録媒体。
【請求項１０】
映像音響信号を入力とする信号処理装置を制御するプロセッサに実行させるプログラムにおいて、
入力音響信号の振幅を抽出し、抽出した前記振幅の時間的変化に基づいて、前記入力音響信号の内容の変化点を検出する音響信号変化点検出ステップを含み、
前記変化点検出ステップは、
前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、
前記変化量が所定の閾値より大きい場合に変化点とし、
検出された前記音響信号変化点近傍の映像および音声を再生または記録する
ことを特徴とするプログラム。
【請求項１１】
映像音響信号を入力とする信号処理装置において、
入力映像信号の内容の変化点を検出する映像信号変化点検出手段と、
前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、前記変化量が所定の閾値より大きい場合に変化点として検出する音響信号変化点検出手段と、
前記検出した入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容を検出する検出手段と
を備えることを特徴とする信号処理装置。
【請求項１２】
前記映像音響信号は、テレビジョン放送信号であり、前記所定の内容の信号はコマーシャルメッセージである
ことを特徴とする請求項１１に記載の信号処理装置。
【請求項１３】
映像音響信号を入力とする信号処理装置の信号処理方法において、
入力映像信号の内容の変化点を検出する映像信号変化点検出ステップと、
前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、前記変化量が所定の閾値より大きい場合に変化点として検出する音響

信号変化点検出ステップと、
前記検出した入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容を検出する検出ステップと
を含むことを特徴とする信号処理方法。
【請求項１４】
映像音響信号を入力とする信号処理装置を制御するプロセッサに実行させるプログラムにおいて、
入力映像信号の内容の変化点を検出する映像信号変化点検出ステップと、
前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、前記変化量が所定の閾値より大きい場合に変化点として検出する音響信号変化点検出ステップと、
前記検出した入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容を検出する検出ステップと
を含むことを特徴とするプログラムが記録されている記録媒体。
【請求項１５】
映像音響信号を入力とする信号処理装置を制御するプロセッサに実行させるプログラムにおいて、
入力映像信号の内容の変化点を検出する映像信号変化点検出ステップと、
前記入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、前記第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる前記係数の値を検出するとともに、帯域毎の、前記第１の信号の平滑されたスペクトル振幅の値と、前記第２の信号の平滑されたスペクトル振幅の値に検出した値の前記係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、前記変化量が所定の閾値より大きい場合に変化点として検出する音響信号変化点検出ステップと、
前記検出した入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容を検出する検出ステップと
を含むことを特徴とするプログラム。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
本発明は、信号処理装置および方法、記録媒体、並びにプログラムに関し、特に、音響の変換点を検出することができるようにした信号処理装置および方法、記録媒体、並びにプログラムに関する。
【０００２】
【従来の技術】
テレビジョン放送、ラジオ放送、ビデオテープ、またはビデオディスク等から得られる映像音声信号に対して、編集、インデクシング、検索、または要約を行う場合、コンテンツの内容や場面が切り替わる点（シーチェンジ）が利用される。
【０００３】
シーンチェンジの検出は、例えば、映像信号が大きく変化する場面（カットチェンジ）を検出し、それをシーンチェンジとして代用することで行われる。しかしながら、同一のシ

ーン内には、複数のカットチェンジが存在する場合もあり、このときこの方法では、必要以上にシーンチェンジが検出されてしまう。

【０００４】

なお、特開平5-37893号公報には、映像信号のフレーム間の差分を検出し、その時点の静止映像を順次記録、再生することでテレビジョン信号の要約の記録再生を行う方法が開示されている。

【０００５】

また、特開平2000-5111743号公報には、一定区間の音声信号の最小レベルを、背景音声レベルとし、その背景音声レベルの変化が大きい時点をシーンチェンジとする方法が開示されている。

【０００６】

しなしながら、背景音声とそうでない音声とが明確に区別されていない場合があること、背景音声レベルの変化を伴わないシーンチェンジが存在すること、そして同一シーン内で背景音声レベルが変化する場合があること等の理由から、この方法でも、シーンチェンジを正確に検出することができない。

【０００７】

さらに、無音区間を、シーンチェンジとして検出する方法も考えられているが、無音区間が存在せずに、急激に音響信号の切り換えが行われる場合もあるので、この方法でも、シーンチェンジを正確に検出することはできない。

【０００８】

【発明が解決しようとする課題】

すなわち、従来においては、シーンチェンジを正確に検出することができないという課題があった。

【０００９】

本発明はこのような状況に鑑みてなされたものであり、シーンチェンジをより正確に検出することができるようにするものである。

【００１０】

【課題を解決するための手段】

本発明の第１の信号処理装置は、入力音響信号に対して、入力音響信号の振幅を抽出する振幅抽出手段と、振幅抽出手段により抽出された振幅の時間的変化に基づいて、入力音響信号の内容の変化点を検出する変化点検出手段とを備え、振幅抽出手段は、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、変化点検出手段は、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、スペクトル変化量が所定の閾値より大きい場合に変化点とすることを特徴とする。

【００１１】

振幅抽出手段は、周波数分解されてスケールファクタとともに符号化された入力音響信号に対しては、スケールファクタを振幅として抽出することができる。

【００１２】

本発明の第１の信号処理方法は、入力音響信号に対して、入力音響信号の振幅を抽出する振幅抽出ステップと、振幅抽出ステップの処理で抽出された振幅の時間的変化に基づいて、入力音響信号の内容の変化点を検出する変化点検出ステップとを含み、振幅抽出ステップは、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、変化点検出ステップは、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も

小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、スペクトル変化量が所定の閾値より大きい場合に変化点とすることを特徴とする。

【００１３】

本発明の第１の記録媒体のプログラムは、入力音響信号に対して、入力音響信号の振幅を抽出する振幅抽出ステップと、振幅抽出ステップの処理で抽出された振幅の時間的変化に基づいて、入力音響信号の内容の変化点を検出する変化点検出ステップとを含み、振幅抽出ステップは、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、変化点検出ステップは、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、

スペクトル変化量が所定の閾値より大きい場合に変化点とすることを特徴とする。

【００１４】

本発明の第１のプログラムは、入力音響信号に対して、入力音響信号の振幅を抽出する振幅抽出ステップと、振幅抽出ステップの処理で抽出された振幅の時間的変化に基づいて、入力音響信号の内容の変化点を検出する変化点検出ステップとを含み、振幅抽出ステップは、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、変化点検出ステップは、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、スペクトル変化量が所定の閾値より大きい場合に変化点とすることをコンピュータに実行させることを特徴とする。

【００１５】

本発明の第１の信号処理装置および方法、並びにプログラムにおいては、入力音響信号の振幅が抽出され、抽出された振幅の時間的変化に基づいて、入力音響信号の内容の変化点が検出され、その際、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅が抽出され、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値が検出されるとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出された値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量が算出され、スペクトル変化量が所定の閾値より大きい場合に変化点とされる。

【００１６】

本発明の第２の信号処理装置は、入力音響信号の振幅を抽出し、抽出した振幅の時間的変化に基づいて、入力音響信号の内容の変化点を検出する音響信号変化点検出手段を備え、変化点検出手段は、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる

係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、変化量が所定の閾値より大きい場合に変化点とし、検出された音響信号変化点近傍の映像および音声を再生または記録することを特徴とする。

【００１７】

映像音響信号は、テレビジョン放送信号であり、所定の内容の信号は、コマーシャルメッセージであるようにすることができる。

【００１８】

本発明の第２の信号処理方法は、入力音響信号の振幅を抽出し、抽出した振幅の時間的変化に基づいて、入力音響信号の内容の変化点を検出する音響信号変化点検出ステップを含み、変化点検出ステップは、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、変化量が所定の閾値より大きい場合に変化点とし、検出された音響信号変化点近傍の映像および音声を再生または記録することを特徴とする。

【００１９】

本発明の第２の記録媒体のプログラムは、入力音響信号の振幅を抽出し、抽出した振幅の時間的変化に基づいて、入力音響信号の内容の変化点を検出する音響信号変化点検出ステップを含み、変化点検出ステップは、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、変化量が所定の閾値より大きい場合に変化点とし、検出された音響信号変化点近傍の映像および音声を再生または記録することを特徴とする。

【００２０】

本発明の第２のプログラムは、入力音響信号の振幅を抽出し、抽出した振幅の時間的変化に基づいて、入力音響信号の内容の変化点を検出する音響信号変化点検出ステップを含み、変化点検出ステップは、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、変化量が所定の閾値より大きい場合に変化点とし、検出された音響信号変化点近傍の映像および音声を再生または記録することを特徴とする。

【００２１】

本発明の第２の信号処理装置および方法、並びにプログラムにおいては、入力音響信号の振幅が抽出され、抽出された振幅の時間的変化に基づいて、入力音響信号の内容の変化点が検出され、その際、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅が抽出され、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペ

クトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値が検出されるとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量が算出され、変化量が所定の閾値より大きい場合に変化点とされ、検出された音響信号変化点近傍の映像および音声が再生または記録される。

【００２２】

本発明の第３の信号処理装置は、入力映像信号の内容の変化点を検出する映像信号変化点検出手段と、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、変化量が所定の閾値より大きい場合に変化点として検出する音響信号変化点検出手段と、検出した入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容を検出する検出手段とを備えることを特徴とする。

【００２３】

映像音響信号は、テレビジョン放送信号であり、所定の内容の信号はコマーシャルメッセージであるようにすることができる。

【００２４】

本発明の第３の信号処理方法は、入力映像信号の内容の変化点を検出する映像信号変化点検出ステップと、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、変化量が所定の閾値より大きい場合に変化点として検出する音響信号変化点検出ステップと、検出した入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容を検出する検出ステップとを含むことを特徴とする。

【００２５】

本発明の第３の記録媒体のプログラムは、入力映像信号の内容の変化点を検出する映像信号変化点検出ステップと、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、変化量が所定の閾値より大きい場合に変化点として検出する音響信号変化点検出ステップと、検出した入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容を検出する検出ステップとを含むことを特徴とする。

【００２６】

本発明の第３のプログラムは、入力映像信号の内容の変化点を検出する映像信号変化点

検出ステップと、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅を抽出し、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値を検出するとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量を算出し、変化量が所定の閾値より大きい場合に変化点として検出する音響信号変化点検出ステップと、検出した入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容を検出する検出ステップとを含むことを特徴とする。

【００２７】

本発明の第３の信号処理装置および方法、並びにプログラムにおいては、入力映像信号の内容の変化点が検出され、入力音響信号をスペクトル分解して各スペクトル信号のスペクトル振幅が抽出され、帯域毎の、サンプリングされた第１の信号の平滑化されたスペクトル振幅の値と、第１の信号より１つ前にサンプリングされた第２の信号の平滑化されたスペクトル振幅の値に所定の値の係数を乗算した結果得られた値との差の合計値が最も小さくなる係数の値が検出されるとともに、帯域毎の、第１の信号の平滑されたスペクトル振幅の値と、第２の信号の平滑されたスペクトル振幅の値に検出した値の係数を乗算した結果得られた値との差の合計値を、第１の信号のスペクトルエネルギーで正規化してスペクトル変化量が算出され、変化量が所定の閾値より大きい場合に変化点として検出され、検出された入力音響信号の内容の変化点と入力映像信号の内容の変化点がほぼ同じタイミングで検出されたと判定したときに所定の信号内容が検出される。

【００３１】

【発明の実施の形態】

図１は、本発明の音声信号処理装置１の構成例を示している。この場合、音声信号処理装置１には、コンテンツを構成する映像音声信号のうちの音声信号が入力されるものとする。

【００３２】

スペクトル分解部１１は、バンドパスフィルタ群や高速フーリエ変換器等により構成されており、入力される音声信号を、複数の周波数帯域の信号に分解する。複数の周波数帯域に分解された信号（以下、サブバンド信号と称する）は、振幅検出部１２に供給される。

【００３３】

振幅検出部１２は、スペクトル分解部１１からの各サブバンド信号の振幅を検出し、その検出結果を平滑部１３に供給する。

【００３４】

平滑部１３は、メジアンフィルタ、モードフィルタ、最小値フィルタなどの平滑フィルタにより構成され、振幅検出部１２から供給された各サブバンド信号の振幅の急激な変動を平滑化する。平滑部１３は、平滑したサブバンド信号の振幅（以下、平滑化バンドパス振幅と称する）を、平滑時間に併せてサンプリングし、そのサンプリング結果を、記憶部１４および算出部１５に出力する。

【００３５】

なお、平滑部１３は、例えば、所定の区間（数秒間）のサブバンド信号の振幅を平均することで、サブバンド信号の振幅を平滑化するようにすることもできる。

【００３６】

記憶部１４は、１サンプリング時間または数サンプリング時間に渡る平滑化サブバンド振幅を記憶する。

【００３７】

算出部１５は、平滑部１３および記憶部１４から供給される平滑化サブバンド振幅を利用して、式（１）を解き、スペクトル変化量Ｄを算出する。

【００３８】
【数１】

$$D = \min_{w} \frac{\sum_{k=1}^{K} \{F(n,k) - wF(n-1,k)\}^2}{\sum_{k=1}^{K} F^2(n,k)} \quad \cdots (1)$$

【００３９】
式中、ｎは、サンプリング時刻を表し、ｋは、サブバンド信号の周波数帯域を表し、Ｋは、サブバンド信号の数を表し、そしてＦ（ｎ，ｋ）は、ｋで表される周波数帯域のサンプリング時刻ｎの平滑化サブバンド振幅の値を表す。
【００４０】
すなわち、予め決められた複数の値がｗに順に設定されて、式（１）が演算され、その演算結果が最小となるｗが検出される。そしてそのときの演算結果が、スペクトル変化量Ｄとなる。従って、スペクトル変化量Ｄは、時刻ｎと、時刻ｎ－１のスペクトルの差分エネルギーを、現在のスペクトルエネルギーで正規化した値（０乃至１）である。
【００４１】
なお、スペクトル変化量Ｄは、実際は計算を簡単にするために、式（１）と等値の式（２）を演算することで求められる。式中、Ｓaaは式（３）、Ｓbbは式（４）、そしてＳabは式（５）で求められる。
【００４２】
【数２】

$$D = \frac{s_{aa}s_{bb} - s_{ab}^2}{s_{aa}s_{bb}} \quad \cdots (2)$$

【数３】

$$s_{aa} = \sum_{k=1}^{K} F(n,k)F(n,k) \quad \cdots (3)$$

$$s_{bb} = \sum_{k=1}^{K} F(n-1,k)F(n-1,k) \quad \cdots (4)$$

$$s_{ab} = \sum_{k=1}^{K} F(n,k)F(n-1,k) \quad \cdots (5)$$

【００４３】
算出部１５は、算出したスペクトル変化量Ｄを、適宜、シーンチェンジ判定部１６に出力する。
【００４４】
シーンチェンジ判定部１６は、算出部１５からのスペクトル変化量Ｄが予め設定された閾値より大きいか否かを判定し、閾値より大きいと判定した場合、例えば、スペクトル変化量Ｄが得られた音声信号の位置（時刻）をシーンチェンジの検出信号として出力する。
【００４５】
なお、検出されたシーンチェンジが隣接して複数個存在する場合、それらのうちの１つをシーンチェンジとするようにすることもできる。
【００４６】
次に、音声信号処理装置１の動作を、図２のフローチャートを参照して説明する。なお、ここでは、図３（Ａ）に示すテレビジョン放送の音声信号からシーンチェンジを検出するものとする。

【００４７】
ステップＳ１において、スペクトル分解部１１は、入力されたテレビジョン放送の音声信号を、８個の周波数帯域のサブバンド信号に分解する。ステップＳ２において、振幅検出部１２は、スペクトル分解部１１からの各サブバンド信号の振幅を検出し、検出結果を平滑部１３に出力する。
【００４８】
次に、ステップＳ３において、平滑部１３は、各サブバンド信号の振幅を、図３（Ｄ）に示すように平滑化するとともに、平滑時間に合わせてサンプリングを行い、そのサンプリング結果を、記憶部１４および算出部１５に出力する。図３（Ｄ）には、各サブバンド信号の平滑化サブバンド振幅が示されている。なお、図中、縦軸に示されている数値は、各周波数帯域の中心周波数である。
【００４９】
次に、ステップＳ４において、算出部１５は、平滑部１３からの平滑化サブバンド振幅の値Ｆ（ｎ，ｋ）と、記憶部１４からの、それより１つ前にサンプリングされた平滑化サブバンド振幅の値Ｆ（ｎ－１，ｋ）を利用して式（２）を解き、スペクトル変化量Ｄを算出する。
【００５０】
図３（Ａ）において、放送内容がＣＭから番組本編に切り替わる、時刻１０秒付近では、その時間帯の各平滑化サブバンド振幅を拡張した図４に示すように、変化の割合が各平滑化サブバンド振幅毎で異なるので（高域側での増加率は大きいが、低域側での増加率は小さいので）、図３（Ｃ）に示すように、このとき大きなスペクトル変化量Ｄが算出される。
【００５１】
また図３（Ａ）において、ガラスの破砕音が発生する、時刻１３秒付近、および手を叩した音が発生する時刻２９秒付近においても、大きなスペクトル変化量Ｄが算出される。
【００５２】
ステップＳ５において、シーンチェンジ判定部１６は、算出部１５により算出されたスペクトル変化量Ｄが、所定の閾値より大きいか否かを判定し、ステップＳ６において、その判定結果に基づく処理を行う。
【００５３】
例えば、スペクトル変化量Ｄが閾値より大きいと判定された場合、シーンチェンジ判定部１６は、シーンチェンジの検出信号を、外部の装置に出力する。
【００５４】
その後、処理は終了し、次に入力される音声信号に対してステップＳ１乃至ステップＳ６の処理が同様に行われる。
【００５５】
すなわち、本発明によれば、図３のテレビジョン放送からは、ＣＭから番組に切り替わったとき、ガラスの破砕音が発生したとき、また手を叩いた音が発生したときを、シーンチェンジとして検出することができる。
【００５６】
それに対して、映像信号を利用した従来の方法では、図３（Ｂ）の縦方向の線で示されるように、同一シーン（例えば、１つのＣＭ）から複数のカットチェンジ（シーンチェンジ）が検出されてしまう。
【００５７】
図５は、他のテレビジョン放送の音声信号からシーンチェンジを検出した場合の例を表している。
【００５８】
この場合においても、本発明によれば、図５（Ｃ）に示すように、図５（Ａ）中上向きの矢印で示されている、音声が切り替わるとき、ＣＭ１からＣＭ２に切り替わるとき、ＣＭ２から番組に切り換わるとき、大きなスペクトル変化量Ｄが算出されるので、その時点を

シーンチェンジとして検出することができる。それに対して映像信号を利用した従来の方法では、図５（Ｂ）に示すように、同一シーンから複数のカットチェンジが検出される。なお、図５（Ｅ）は、音声信号の波形を表している。

【００５９】

図６は、音声信号処理装置１の他の構成例を表している。この音声信号処理装置には、図１のスペクトル分解部１１および振幅検出部１２に代えて、スケールファクタ抽出部２１が設けられている。他の部分は、図１における場合と同様である。

【００６０】

この例の場合、音声信号処理装置１には、ＭＰＥＧ１方式など、周波数分解を伴う符号化方法で符号化された音声信号が入力される。通常、ＭＰＥＧ１方式で符号化される場合、周波数分解された信号は、符号化効率を向上させるために正規化され、その正規化のために算出された拡大縮小率（スケールファクタ）が音声信号とともに符号化されるようになされている。

【００６１】

すなわち、スケールファクタ抽出部２１は、入力された符号化信号から各サブバンド信号のスケールファクタを抽出し、それを平滑部１３に供給する。

【００６２】

平滑部１３は、スケールファクタ抽出部２１から供給されたスケールファクタの急激な変動を平滑化する。平滑部１３は、平滑化したスケールファクタを、平滑時間に合わせてサンプリングし、そのサンプリング結果を、記憶部１４および算出部１５に出力する。

【００６３】

算出部１５は、図１における場合と同様に、平滑部１３および記憶部１４からのサンプリング値に基づいて式（２）を演算してスペクトル変化量Ｄを算出し、シーンチェンジ判定部１６は、スペクトル変化量Ｄに基づいてシーンチェンジを検出する。

【００６４】

スケールファクタは、各周波数帯域のフレーム内の平均振幅と概略比例関係にあるので、このようにスケールファクタを利用してシーンチェンジを検出することができる。すなわち、この場合、符号化信号を復号することなく、シーンチェンジを検出することができる。

【００６５】

図７は、音声信号処理装置１を利用した記録再生装置の構成例を示している。この装置は、コンテンツを構成する映像音声信号から代表画像や代表音声を記録再生することができる。すなわち、ユーザは、例えば、この装置に記録された代表画像または代表音声を視聴することで、コンテンツの内容を容易に把握することができる。またユーザは、この装置に記録された画像や音声をコンテンツの頭出しのインデックスとして利用し、迅速に頭出しを行うこともできる。

【００６６】

信号処理装置１は、上述したように、入力された音声信号のスペクトル変化量Ｄに基づいてシーンチェンジを検出し、その検出結果を、制御部３２に出力する。

【００６７】

映像信号処理部３１は、入力される映像信号が大きく変化する部分を検出し、すなわち、カットチェンジを検出し、その検出結果を、制御部３２に出力する。なお、映像信号処理部３１は、例えば、特開平5-37893号公報に開示されている技術を利用してカットチェンジを検出する。

【００６８】

制御部３２は、音声信号処理装置１からの検出結果および映像信号処理部３１からの検出結果に基づいて、代表映像および代表音声を決定する。

【００６９】

例えば、制御部３２は、音声信号処理装置１または映像信号処理部３１のいずれか一方で、シーンチェンジ（カットチェンジ）が検出されたとき、または両者においてほぼ同じタ

イミングで検出されたとき、そのシーンチェンジに対応する画像および音声を代表画像および代表音声とする。

【００７０】

制御部３２は、代表画像および代表音声を決定したとき、それらの画像音声の開始時刻とその長さを、音声抽出部３３、画像抽出部３４、およびタイムスタンプ生成部３５にそれぞれ出力する。

【００７１】

音声抽出部３３は、入力される音声信号から、制御部３２からの情報に対応する音声を抽出し、記録部３６に出力する。画像抽出部３４は、入力される映像信号から、制御部３２からの情報に対応する画像を抽出し、記録部３６に出力する。

【００７２】

タイムスタンプ生成部３５は、制御部３２からの情報に基づいて、音声抽出部３３により抽出された音声信号および画像抽出部３４により抽出された映像信号を適切に再生させるためのタイムスタンプを生成し、記録部３６に出力する。

【００７３】

記録部３６は、例えば、磁気テープ、磁気ディスク、光ディスク、または光磁気ディスク等を記録媒体として用い、音声抽出部３３および画像抽出部３４からの音声信号および映像信号を、タイムスタンプ生成部３５からのタイムスタンプと対応させて記録する。

【００７４】

再生部３７は、必要に応じて記録部３６に記録されている音声（代表音声）および画像（代表画像）を再生する。

【００７５】

図８は、音声信号処理装置１を利用したＣＭ検出装置の構成例を示している。この装置は、テレビジョン放送の番組本編に付加されるコマーシャルメッセージ（ＣＭ）を検出することができる。

【００７６】

音声信号処理装置１は、上述したように、入力されるテレビジョン放送の音声信号のスペクトル変化量Ｄに基づいてシーンチェンジを検出し、その検出結果を制御部４１に出力する。

【００７７】

映像信号処理部３１（図７の記録再生装置の映像信号処理部３１と同じもの）は、入力されたテレビジョン放送の映像信号からカットチェンジを検出し、その検出結果を制御部４１に出力する。

【００７８】

制御部４１は、信号処理装置１および映像処理部３１からの検出結果に基づいてＣＭの開始時刻と長さを検出し、記録部４２に出力する。

【００７９】

記録部４２は、入力される音声信号および映像信号とともに、制御部４１からのＣＭの開始時刻と長さを記録する。

【００８０】

次に、ＣＭ検出装置の制御部４１の動作を、図９のフローチャートを参照して説明する。

【００８１】

ステップＳ１１において、制御部４１は、音声信号処理装置１および映像信号処理部３１からの検出結果に基づいて、音響的なシーンチェンジと映像のカットチェンジの両方が同じタイミングで（ほぼ同一時刻に）検出されたか否かを判定し、同じタイミングで検出されたと判定した場合、ステップＳ１２に進み、そのタイミングに対応する時刻を内蔵するメモリに記憶させる。

【００８２】

次に、ステップＳ１３において、制御部４１は、ステップＳ１２で記憶された時刻から時間Ｔ1だけ前の時刻付近に、シーンチェンジとカットチェンジの両方が検出されていたか

否かを判定し、検出されていたと判定した場合、ステップＳ１４に進む。

【００８３】

なお、同じタイミングでシーンチェンジとカットチェンジの両方が検出されたときの時刻は、以前行われたステップＳ１２の処理でメモリに記録されているので、制御部４１は、メモリの内容を参照してステップＳ１３での判定を行う。

【００８４】

また、時間Ｔ1は、ＣＭの放送時間である。テレビジョン放送されるＣＭは、放送者の指定する規格に基づいて制作されるため、その時間長はいくつかに限定されている。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、または６０秒の長さで制作されている。すなわち、時間Ｔ1は、それらの時間のいずれかの時間とされる。

【００８５】

ステップＳ１４において、制御部４１は、ステップＳ１２で記憶された時刻から時間Ｔ1＋時間Ｔ2だけ前の時刻付近に、シーンチェンジとカットチェンジの両方が検出されたか否かを判定し、検出されたと判定した場合、ステップＳ１５に進む。

【００８６】

時間Ｔ2も、ＣＭ放送の時間であり、１５秒、３０秒、または６０秒のいずれかの時間である。

【００８７】

ステップＳ１５において、制御部５１は、ステップＳ１２で記録された時刻から時間Ｔ1だけ前の時刻、および時間Ｔ1＋時間Ｔ2だけ前の時刻を、ＣＭの開示時刻と判定する。そして制御部５１は、検出したＣＭの開始時刻とその長さ（時間Ｔ1、および時間Ｔ1＋時間Ｔ2）を記録部４２に出力する。記録部４２は、入力される音声信号および映像信号とともに、制御部４１からのＣＭの開始時刻とその長さを記録する。

【００８８】

ステップＳ１１，Ｓ１３，Ｓ１４で、シーンチェンジとカットチェンジが同じタイミングで検出されていないと判定された場合、またはステップＳ１５でＣＭ開始時刻とその長さが検出された場合、ステップＳ１１戻り、次に入力されるシーンチェンジおよびカットチェンジの検出結果に対して、それ以降の処理が実行される。

【００８９】

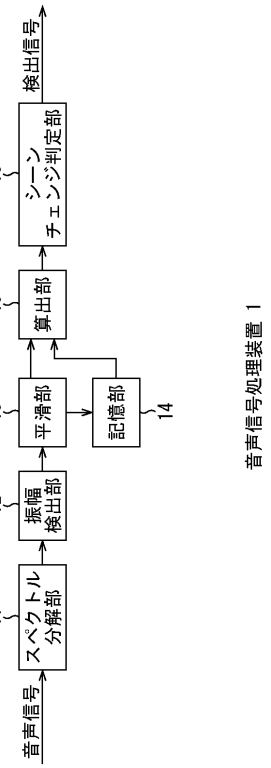
上述した一連の処理は、ハードウエアにより実現させることもできるが、ソフトウエアにより実現させることもできる。一連の処理をソフトウエアにより実現する場合には、そのソフトウエアを構成するプログラムがコンピュータにインストールされ、そのプログラムがコンピュータで実行されることより、上述した音声信号装置１が機能的に実現される。

【００９０】

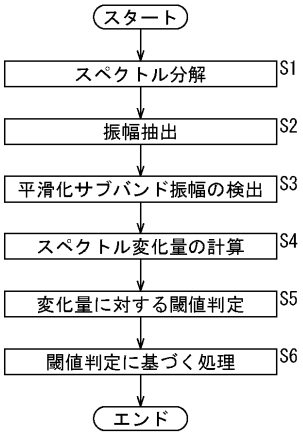
図１０は、上述のような音声信号装置１として機能するコンピュータ１０１の一実施の形態の構成を示すブロック図である。CPU（Central Processing Unit）１１１にはバス１１５を介して入出力インタフェース１１６が接続されており、CPU１１１は、入出力インタフェース１１６を介して、ユーザから、キーボード、マウスなどよりなる入力部１１８から指令が入力されると、例えば、ROM（Read Only Memory）１１２、ハードディスク１１４、またはドライブ１２０に装着される磁気ディスク１３１、光ディスク１３２、光磁気ディスク１３３、若しくは半導体メモリ１３４などの記録媒体に格納されているプログラムを、RAM（Random Access Memory）１１３にロードして実行する。これにより、上述した各種の処理が行われる。さらに、CPU１１１は、その処理結果を、例えば、入出力インタフェース１１６を介して、LCD（Liquid Crystal Display）などよりなる出力部１１７に必要に応じて出力する。なお、プログラムは、ハードディスク１１４やROM１１２に予め記憶しておき、コンピュータ１０１と一体的にユーザに提供したり、磁気ディスク１３１、光ディスク１３２、光磁気ディスク１３３，半導体メモリ１３４等のパッケージメディアとして提供したり、衛星、ネットワーク等から通信部１１９を介してハードディスク１１４に提供することができる。

【００９１】
なお、本明細書において、記録媒体により提供されるプログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。
【００９２】
【発明の効果】
第１の本発明によれば、変化点を正確に検出することができる。
【００９３】
第２の本発明によれば、変化点に応じた所望の映像および音声を再生または記録することができる。
【００９４】
第３の本発明によれば、変化点を正確に検出することができる。
【図面の簡単な説明】
【図１】　本発明を適用した音声信号処理装置の構成例を示すブロック図である。
【図２】図１の音声信号処理装置の動作を説明するフローチャートである。
【図３】スペクトル変化量を説明する図である。
【図４】スペクトル変化量を説明する他の図である。
【図５】スペクトル変化量を説明する他の図である。
【図６】本発明を適用した音声信号処理装置の他の構成例を示すブロック図である。
【図７】本発明を適用した記録再生装置の構成例を示すブロック図である。
【図８】本発明を適用したＣＭ検出装置の構成例を示すブロック図である。
【図９】図８のＣＭ検出装置の動作を説明するフローチャートである。
【図１０】パーソナルコンピュータ１０１の構成例を示すブロック図である。
【符号の説明】
１１　スペクトル分解部，　１２　振幅検出部，　１３　平滑部，　１４　記憶部，　１５　算出部，　１６　シーンチェンジ判定部，　２１　スケールファクタ抽出部，　３１　映像信号処理部，　３２　制御部，　３３　音声抽出部，３４　画像抽出部，　３５　タイムスタンプ生成部，　３６　記録部，　３７
再生部，　４１　制御部，　４２　記録部

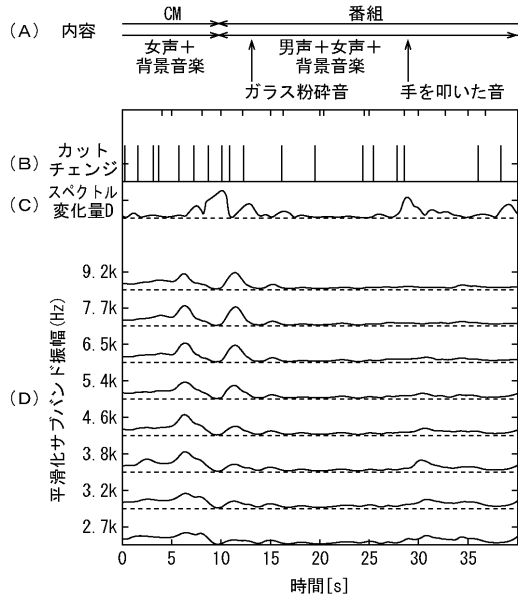
【図１】
図1



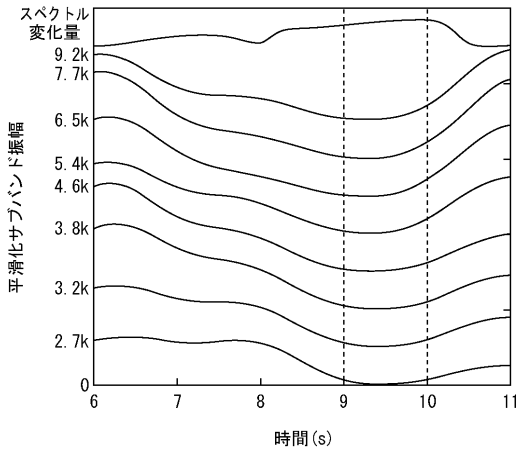
【図２】
図2



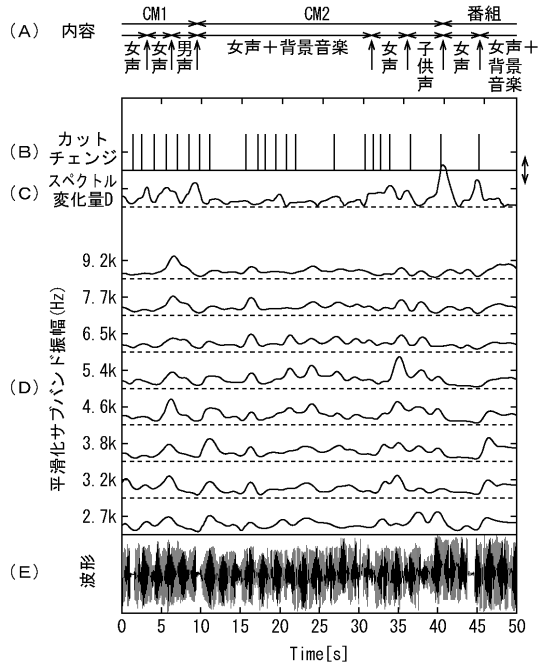
【図３】
図3



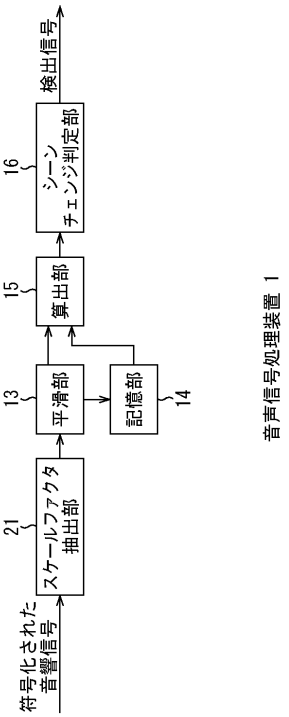
【図４】
図4



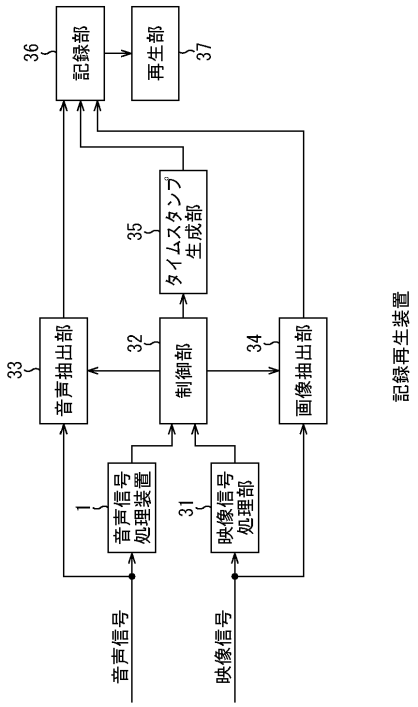
【図５】
図5



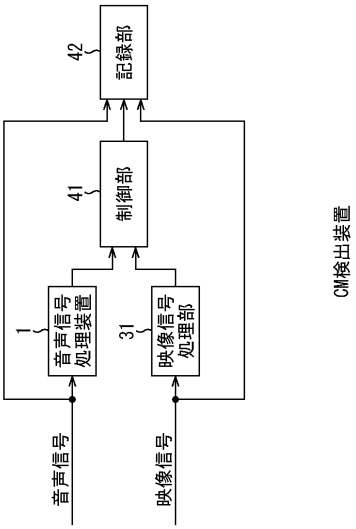
【図６】
図6



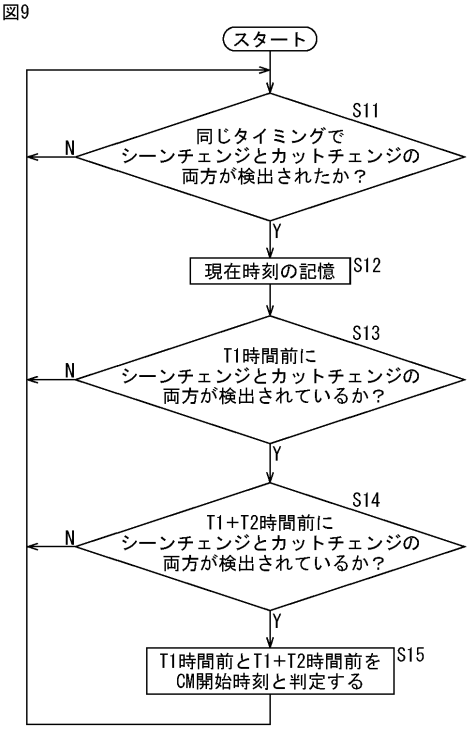
【図７】
図7



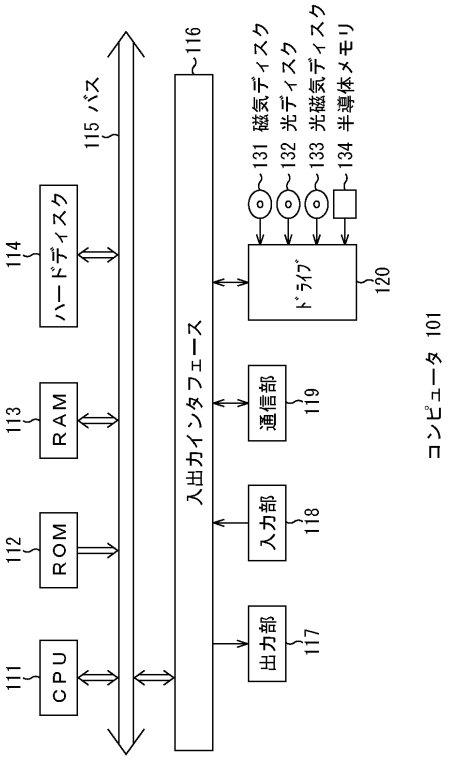
【図８】
図8



【図９】
図9


スタート
S11
同じタイミングで
シーンチェンジとカットチェンジの
両方が検出されたか？
N
Y
現在時刻の記憶
S12
S13
T1時間前に
シーンチェンジとカットチェンジの
両方が検出されているか？
N
Y
S14
T1+T2時間前に
シーンチェンジとカットチェンジの
両方が検出されているか？
N
Y
T1時間前とT1+T2時間前を
CM開始時刻と判定する
S15

【図１０】
図10


CPU
111
ROM
112
RAM
113
ハードディスク
114
115 バス
入出力インタフェース
116
出力部
117
入力部
118
通信部
119
ドライブ
120
131 磁気ディスク
132 光ディスク
133 光磁気ディスク
134 半導体メモリ
コンピュータ 101

フロントページの続き

(56)参考文献 特表２０００－５１１７４３（ＪＰ，Ａ）
特開２００１－１４２４８０（ＪＰ，Ａ）
特開２０００－１８１４７３（ＪＰ，Ａ）
特開２０００－３２２０７６（ＪＰ，Ａ）
特開昭６１－３４７６４（ＪＰ，Ａ）
特表２００３－５３００２７（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
G10L 11/00,15/00
H04N 5/91- 5/93