

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4065961号
(P4065961)

(45) 発行日 平成20年3月26日(2008.3.26)

(24) 登録日 平成20年1月18日(2008.1.18)

(51) Int.Cl.		F I			
HO4N 5/60	(2006.01)	HO4N 5/60		Z	
HO4N 17/04	(2006.01)	HO4N 17/04		E	
HO4N 7/173	(2006.01)	HO4N 7/173	630		

請求項の数 13 (全 18 頁)

(21) 出願番号 特願2007-513550 (P2007-513550)
 (86) (22) 出願日 平成18年10月6日(2006.10.6)
 (86) 国際出願番号 PCT/JP2006/320098
 (87) 国際公開番号 W02007/049451
 (87) 国際公開日 平成19年5月3日(2007.5.3)
 審査請求日 平成19年3月30日(2007.3.30)
 (31) 優先権主張番号 特願2005-312111 (P2005-312111)
 (32) 優先日 平成17年10月27日(2005.10.27)
 (33) 優先権主張国 日本国(JP)

早期審査対象出願

(73) 特許権者 304021831
 国立大学法人 千葉大学
 千葉県千葉市稲毛区弥生町1番33号
 (72) 発明者 池田 宏明
 千葉県千葉市稲毛区弥生町1番33号 国立
 立大学法人千葉大学
 工学部内
 (72) 発明者 岩井 玲子
 千葉県千葉市稲毛区弥生町1番33号 国立
 立大学法人千葉大学
 工学部内

審査官 西谷 憲人

最終頁に続く

(54) 【発明の名称】映像と音声のずれ時間を、高精度かつ簡便に計測する方法および計測装置

(57) 【特許請求の範囲】

【請求項1】

デジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法において、

映像フレームを読み込むステップと、

該映像フレームの指定部分の画像を抽出するステップと、

該デジタル映像信号を構成するフレームについて、

フレーム間補間により、計測に必要とする時間分解能に相当する時刻毎の内挿フレームを求め基本映像信号とし、該基本映像信号を構成する各フレームの振幅領域差分を求め、各差分フレームの二次元フーリエ変換により空間周波数スペクトラムを求めて、該空間周波数スペクトラムの低域成分を取り出し、該低域成分を基本映像信号の時系列周波数とみなすステップと、

デジタル映像信号に伴うデジタル音声信号について、

低域フィルタを通し高周波成分を除去し、計測に必要とする時間分解能に相当する時刻毎のサンプルとなし得るように、短時間窓一次元フーリエ変換により、周波数スペクトラムを求めるステップと、

該周波数スペクトラムから、低域周波数スペクトラムを取り出し、該低域成分を基本音声信号の時系列周波数とみなすステップと、

以上により得られた基本映像信号の時系列周波数と基本音声信号の時系列周波数とに対して、

10

20

相互相関係数を計算して得られた基本映像信号の周波数領域の時系列データと基本音声信号の周波数領域の時系列データとの該相互相関係数の最大値を求めるステップと、該最大値を与える時系列サンプル数から、該デジタル映像信号と該デジタル音声信号のずれ時間を統計的に同定することを特徴とするデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法。

【請求項 2】

請求項 1 のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法において、

前記デジタル映像信号を、映像フレーム間補間によりサンプリング間隔を計測時間分解能と等しくするように映像フレームを増加し、

その短時間二次元フーリエ変換周波数スペクトラムから得られる情報を、基本映像信号の周波数領域の時系列データとすることを特徴とする請求項 1 記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法。

10

【請求項 3】

請求項 1 記載のデジタル音声信号を、デジタルフィルターにより高周波成分を除去し、請求項 2 記載の計測時間分解能と等しい時間窓毎の短時間一次元フーリエ変換を低周波領域デジタル音声信号に施し、

該デジタル音声信号に比べてサンプリング間隔を長くした周波数領域の音声信号時系列を得て、

該音声信号時系列を、基本音声信号の周波数領域の時系列データとすることを特徴とする請求項 2 に記載のデジタル映像信号とデジタル音声信号のずれ時間を計測する方法。

20

【請求項 4】

請求項 1 記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法において、

基本映像信号を、フレーム内で二次元フーリエ変換して、空間周波数成分ごとのフレーム内平均値を計算し、周波数領域における映像時系列データを生成し、

基本音声信号を、周波数成分ごとの時間窓内平均値を計算し、周波数領域における音声時系列データを生成し、

計測時間分解能と等しい時間窓毎の 2 つの情報を作成することを特徴とする請求項 1 から 3 いずれかに記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法。

30

【請求項 5】

基本音声信号を、

計測要求時間分解能に等しい短時間窓フーリエ変換により、周波数領域における音声時系列データを抽出することを特徴とする請求項 3 または 4 に記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法。

【請求項 6】

上記映像時系列データと上記音声時系列データの相互相関係数を計算して該相互相関係数の最大点を検出し、

該最大点に対応したデータ位置を求めることを特徴とする請求項 1 から 5 いずれかに記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法。

40

【請求項 7】

請求項 6 により検出された相互相関係数の最大点とするサンプル点を、サンプリング間隔に相当する時間から、デジタル映像信号とデジタル音声信号との間の時間的なずれと同定することを特徴とする請求項 6 に記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法。

【請求項 8】

請求項 7 により同定されたデジタル映像信号とデジタル音声信号との間の時間的なずれに相当するサンプル数を、デジタル音声に対してサンプルシフトを施し、デジタル音声を遅延して補正することによりデジタル映像信号とデジタル音声信号との間の時

50

間的なずれを補正する方法。

【請求項 9】

請求項 1 から 8 いずれかに記載の方法をコンピュータに実行させるためのプログラム。

【請求項 10】

請求項 9 のプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項 11】

請求項 9 のプログラムによりデジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する装置。

【請求項 12】

請求項 11 に記載の計測または補正装置を有することを特徴とするデジタル映像信号およびデジタル音声信号処理装置。

10

【請求項 13】

請求項 12 に記載のデジタル映像信号およびデジタル音声信号処理装置は、ネットワークによりデジタル映像信号およびデジタル音声信号が入力されることを特徴とするデジタル映像信号およびデジタル音声信号処理装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、時間窓周波数領域で信号処理することにより、デジタル映像とそれに付随するデジタル音声のずれ時間を、高精度かつ簡便に計測する方法および計測装置に関する。

20

【背景技術】

【0002】

近年、映像や音声をデジタル信号に置き変えて送信するデジタル方式が注目されており、実用・普及化が進みつつある。従来のアナログ方式と比べて、データ処理によりプログラム編集、ポストプロダクションが容易にでき、必要に応じてデータ圧縮が可能になり、この結果、各種品質のプログラムの制作が可能となる。記録媒体はアナログ式ビデオテープから、パーソナルコンピュータ内蔵ハードディスク装置やDVDになる。また、伝送媒体も従来のアナログ方式と比べて、インターネットによる配送になるので、全国どこの地域でも、テレビ画像が二重写しになる「ゴースト」もなく、欲しい情報を簡単に入手でき、視聴者参加の放送番組も試行されている。

30

【0003】

しかし、記録媒体内のテレビ情報を再現する装置内では、デジタル映像とデジタル音声とが別々に処理されるため、再現映像とそれに伴った再現音声の時間的なずれがより問題になる。また、インターネットプロトコルによるパケット伝送においては、多数のルータやネットワーク運用センターを経由して配送されるため、記録媒体内のデジタルテレビ情報に加わる不確実性が存在する。このため、国際連合電気通信諮問委員会 (ITU) や国際電気標準会議 (IEC) では、デジタルテレビに対応した、より高精度のずれ時間測定法の必要性が認識されている。一方、従来、ハイビジョン、NTSC、PAL などアナログ形態の信号処理では、映像と音声のずれの検出・補正のため、例えば、以下のような方法が提案されている。

40

【0004】

(1) 従来技術 1 (特許文献 1)

動画像とそれに付随する音声との間のずれを調整するために、動画像とそれに付随する音声のそれぞれに、レファレンス信号を挿入した、特殊な信号形態を対象にした技術が開示されている。

(2) 従来技術 2 (特許文献 2)

動画像とそれに付随する音声との間のずれを補正するために、ハイビジョン、NTSC、PAL などアナログ形態の信号を対象にした技術が開示されている。

(3) 従来技術 3 (特許文献 3)

50

動画像とそれに付随する音声との間のずれを補正するために、ハイビジョン、NTSC、PALなどアナログ形態の信号を対象にした従来技術1を改良する技術が開示されている。

(4) 従来技術4 (特許文献4)

動画像とそれに伴った音声との間のずれを補正するために、ハイビジョン、NTSC、PALなどアナログ形態の信号を対象にした従来技術1及び従来技術2と異なる技術が開示されている。

【特許文献1】特開平7-38771号

【特許文献2】特開2000-196917号

【特許文献3】特開2002-165153号

【特許文献4】特開2003-158643号

【発明の開示】

【発明が解決しようとする課題】

【0005】

しかしながら、これらの従来の方法では、いずれも時間領域を用いており、デジタル映像とそれに伴ったデジタル音声の時間的ずれを測定・補正するには以下の課題があり、簡易な構成で、高精度なデジタル映像とそれに付随するデジタル音声のずれ時間を計測することは困難である。

(1) 高精度の測定

ずれ時間の測定精度は、映像フレーム時間間隔の制約を受けるが、この制約を受けないより高精度の実現が求められている。

(2) 相互相関などの統計処理

映像とそれに伴った音声の時間的なずれには、統計的な揺らぎが有り得るが、ずれ時間を高精度に計測するための両タイミングに関する相互相関、相互共分散などの統計処理の適用が求められる。

(3) 周波数領域での同定

時間的に変化する映像とそれに伴った音声の短時間フーリエ変換による、周波数領域でのずれ時間最確値の高精度な同定をする必要がある。

(4) デジタルデータ対応

対象とするビデオ信号は、従来方式のハイビジョン、NTSC、PALばかりではなく、特に、インターネットなどデジタルネットワーク対応のビデオ形式、コンピュータ内蔵ハードディスク、DVDなど記録媒体に記録されたビデオ形式など、デジタルデータに対応する必要がある。

(5) 簡易性

デジタルビデオは、PDP、LCDなどの固定解像度大型表示装置、パーソナルコンピュータのモニタなどに再現されるが、デジタル映像とそれに伴ったデジタル音声の時間的ずれの計測結果に基づいた補正を簡易に実現する必要がある。

【0006】

本発明は、かかる事情に鑑みなされたものであって、上記課題を解決して、映像と音声のずれ時間を、高精度かつ簡便に計測する方法、特に、デジタル映像信号とそれに伴ったデジタル音声信号の時間的なずれを測定・補正するのに最適な映像と音声のずれ時間を、高精度かつ簡便に計測する方法、プログラム、プログラムを記憶した記録媒体、プログラムによりデジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する装置、デジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する装置を備えたデジタル映像信号およびデジタル音声信号処理装置を提供する。

【課題を解決するための手段】

【0007】

そこで、本発明者は、上記目的を達成するため、鋭意研究した結果、次の点に着目して、以下のように、本発明を構成する。

10

20

30

40

50

(着目点)

従来、時間領域であったものを、本発明においては、時間窓周波数領域において、時系列周波数スペクトラムを対象に、統計的にずれ時間を同定する。即ち、以下のように本発明を構成する。

【0008】

(発明の構成)

(1) 請求項1に係る発明

請求項1の発明においては、デジタル映像信号を構成するフレームについて、フレーム間補間により、計測に必要とする時間分解能に相当する時刻毎の内挿フレームを求め基本映像信号とし、基本映像信号を構成する各フレームの振幅領域差分を求め、各差分フレームの二次元フーリエ変換により空間周波数スペクトラムを求め、該空間周波数スペクトラムの低域成分を取り出し、該低域成分を基本映像信号の時系列周波数とみなし、デジタル映像信号に伴うデジタル音声信号について、信号処理により高周波成分を除去する低域フィルタを通し、計測に必要とする時間分解能に相当する時刻毎のサンプルとなし得るように、短時間窓一次元フーリエ変換により、周波数スペクトラムを求め、該周波数スペクトラムから、低域周波数スペクトラムを取り出し、該低域成分を基本音声信号の時系列周波数とみなし、

基本映像信号の時系列周波数と基本音声信号の時系列周波数とに対して、相互相関係数を計算し、

該相互相関係数の最大値を求め、該最大値を与える時系列サンプル数から、該デジタル映像信号と該デジタル音声信号のずれ時間を統計的に同定し、計測する方法に関する。

【0009】

(2) 請求項2に係る発明

請求項2の発明は、請求項1の発明において、前記デジタル映像信号を、映像フレーム間補間によりサンプリング間隔を計測時間分解能と等しくするように映像フレームを増加し、その短時間二次元フーリエ変換周波数スペクトラムから得られる情報を周波数領域の時系列信号とすることを特徴とする請求項1記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法に関する。

【0010】

(3) 請求項3に係る発明

請求項3の発明は、請求項1の発明において、前記デジタル音声信号を、デジタルフィルタにより高周波成分を除去し、請求項2記載の計測時間分解能と等しい時間窓毎の短時間一次元フーリエ変換を低周波領域デジタル音声信号に施し、よって、現デジタル音声信号に比べてサンプリング間隔を長くした、周波数領域の音声信号時系列を得る。それを、基本音声信号に関する情報とすることを特徴とする請求項1または2に記載のデジタル映像信号とデジタル音声信号のずれ時間を計測する方法に関する。

【0011】

(4) 請求項4に係る発明

請求項4の発明は、請求項1の発明において、請求項2記載の基本映像信号を、フレーム内で二次元フーリエ変換して、空間周波数成分ごとのフレーム内平均値を計算し、周波数領域における映像時系列データを生成し、請求項3記載の基本音声信号を、周波数成分ごとの時間窓内平均値を計算し、周波数領域における映像時系列データを生成し、計測時間分解能と等しい時間窓毎の2つの情報を生成し、請求項1から3いずれかに記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法に関する。

【0012】

(5) 請求項5に係る発明

請求項5の発明は、請求項1の発明において、請求項3に記載の基本音声信号を、計測要求時間分解能に等しい短時間窓フーリエ変換により、周波数領域における音声時系列データを抽出することを特徴とする請求項1から4いずれかに記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法に関する。

【 0 0 1 3 】

(6) 請求項 6 に係る発明

請求項 6 の発明は、上記映像時系列データと上記音声時系列データの相互相関係数を計算し、

該相互相関係数の最大点を検出し、該最大点に対応したデータ位置を求めることを特徴とする請求項 1 から 5 いずれかに記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する。

【 0 0 1 4 】

(7) 請求項 7 に係る発明

請求項 7 の発明は、上記により検出された相互相関係数の最大点とするサンプル点を、サンプリング間隔に相当する時間から、デジタル映像信号とデジタル音声信号との間の時間的なずれと同定することにより、請求項 1 から 6 いずれかに記載のデジタル映像信号とデジタル音声信号との間の時間的なずれを計測する方法に関する。

【 0 0 1 5 】

(8) 請求項 8 に係る発明

請求項 8 の発明は、前記により同定されたデジタル映像信号とデジタル音声信号との間の時間的なずれに相当するサンプル数を、デジタル音声に対してサンプルシフトを施し、デジタル音声を遅延して補正することによりデジタル映像信号とデジタル音声信号との間の時間的なずれを補正する方法に関する。

【 0 0 1 6 】

(9) 請求項 9 に係る発明

請求項 9 の発明においては、請求項 1 ~ 8 の方法を備えたことを特徴とするプログラムである。

(1 0) 請求項 1 0 に係る発明

請求項 1 0 の発明においては、請求項 9 のプログラムを記憶したことを特徴とする記録媒体である。

【 0 0 1 7 】

(1 1) 請求項 1 1 に係る発明

請求項 1 1 の発明においては、請求項 9 または 1 0 のプログラムによりデジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する装置である。

(1 2) 請求項 1 2 に係る発明

請求項 1 2 の発明においては、前記デジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する装置を備えたことを特徴とするデジタル映像信号およびデジタル音声信号処理装置である。

(1 3) 請求項 1 3 に係る発明

請求項 1 3 の発明においては、前記デジタル映像信号およびデジタル音声信号処理装置は、ネットワークによりデジタル映像信号およびデジタル音声が入力されることを特徴とするデジタル映像信号およびデジタル音声信号処理装置である。

【 発明の効果 】

【 0 0 1 8 】

以上のように構成された本発明は、以下に述べるように、上記課題を解決することができる。

(1) 高精度の測定

従来、ずれ時間の測定精度は、映像フレーム時間間隔の制約を受けるが、この制約を受けないより高精度で実現するため、本発明は、必要な時間分解能に相当するように映像フレーム時間間隔をフレーム間内挿により実現するので、本課題 (1) を解決することができる。

(2) 相互相関などの統計処理

本発明の構成により、基本映像信号と基本音声信号とから得られる、時系列データの相互相関などの統計処理をするので、本課題 (2) を解決することができる。

10

20

30

40

50

(3) 周波数領域での同定

本発明の構成により、周波数領域での同定が可能になるので、本課題(3)を解決することができる。

(4) デジタルデータ対応

本発明の構成により、デジタルデータ対応とするので、本課題(4)を解決することができる。

(5) 簡易性

本発明の構成により、特に、特許請求の範囲10に記載の発明は、当該プログラムの主要部分をハードウェア化して、デジタルテレビジョン受像機、ホームシアター機器、パーソナルコンピュータに実装するので、本課題(5)を解決することができる。

10

【発明を実施するための最良の形態】

【0019】

以下に、本発明の実施形態(以下、単に本発明という)を、図面により説明する。

【実施例1】

【0020】

本発明による実施例を、以下に示す。以下に使用する用語を説明する。発明の技術的範囲を限定するものではない。

【0021】

デジタルビデオは、デジタル映像信号とそれに伴うデジタル音声信号を統合した信号である。

20

デジタル映像信号は、デジタルフレームの時間系列である。デジタルフレームは、A/D変換によりカラー画像の画素をデータとしたものである。フレームと略記する。デジタル音声信号は、A/D変換により音声信号をデータの時間系列としたものである。

【0022】

次に、本発明による映像と音声のずれ時間を、高精度かつ簡便に計測する方法および計測装置を、以下に説明する。

(1) デジタル映像信号とそれに伴ったデジタル音声信号について

例えば、デジタル映像信号とそれに伴ったデジタル音声信号が統合されたデジタルビデオから、または、あらかじめコンピュータ可読形式のファイルとして記録されている媒体(例えば、AVI形式、DVD形式のCD等)から、デジタル映像データとデジタル音声データを、それぞれ分離して読み出す。実施例1では、ニュースを読むアナウンサのバストショットを想定する。デジタル映像信号とそれに伴ったデジタル音声信号の処理と時間遅れの同定による計測処理は、本発明に係る方法を示す全体図の図1による。

30

【0023】

(2) 映像データ前処理について

映像データ前処理は、(V1)

映像フレームの読み込み、(V2)

アナウンサの口元部分の抽出、(V3)

口元部分のモノクロ画像変換、(V4)

画像の強調処理、(V5) 2次元高速フーリエ変換、(V6) フーリエ係数のフレーム内統計処理、(V7) スペクトル時系列内挿処理、(V8) 周波数領域時系列データ抽出とする。映像データ前処理の手順を図2に示す。

40

【0024】

さらに、本発明に係る映像フレームの読み込み処理を、図3から7により、以下に、各ステップ毎に説明する。

(V1) 映像データ入力

映像フレームの読み込みでは、デジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する対象部分のフレームのみを読み込む。

(V2) アナウンサの口元部分の抽出

アナウンサの口元部分の抽出は、従来技術に基づく指定部分の画像抽出により実現する

50

。

(V3) 口元部分のモノクロ画像変換

上記抽出した口元部分を、モノクロ画像変換して、画素ごとの3次元カラーデータ(赤・緑・青)の線形結合により図3のように実現する。

(V4) フレーム内コントラスト強調処理

画像の強調処理は、コントラストを強調することにより図4のように行う。

(V5) フレーム内2次元フーリエ変換

二次元高速フーリエ変換は、上記(V4)による強調処理後の口元画像データに、一次元高速フーリエ変換を2回繰り返すことにより実現する。例を図5に示す。

(V6) フーリエ係数のフレーム内統計処理

周波数領域フレームデータに対して、フレーム内の統計処理(例えば、標準偏差の計算)を施し、フレームの特徴量を抽出する。

10

(V7) スペクトル時系列内挿処理

フーリエ係数の時間変化の内挿による補間処理は、空間周波数成分の低域平均のフレーム時間ごとのサンプルを元データとして、設定した時間分解能(実施例では2ms)に相当するサンプルを得るように内挿することにより実現する。内挿後の時系列データ数を $NV + 1$ とする。図6にその例を示す。点は内挿前データである。

(V8) 周波数領域時系列データ抽出

隣接するデータの差の絶対値を計算することにより、近似導関数絶対値を求める。時系列データ数は NV となる。これを映像データの周波数領域特徴量の時系列データとする。その例を図7に示す。

20

【0025】

次に、本発明に係る音声データ前処理について、図8から11により、以下に、各ステップ毎に説明する。

(3) 音声データ前処理

音声データ前処理は、(A1)

音声サンプルの読み込み、(A2)

音声サンプルの短時間窓1次元高速フーリエ変換、(A3) フーリエ係数の時間窓相当の時間ずれ修正、(A4) 時間差分による近似導関数の計算、(A5) 二乗平均の計算により、音声データの周波数領域特徴量の時系列データとする。音声データ前処理の手順を図8に示す

30

。

(A1) 音声データの入力

音声サンプルの読み込みは、(2)映像データ前処理の(V1)に伴ったデジタル音声データサンプルを読み込む。その例を図9に示す。通常、ステレオからモノラルへ変換する。

(A2) 短時間窓フーリエ変換

音声サンプルの短時間窓1次元高速フーリエ変換は、短時間窓を(2)映像データ前処理(V7)におけるフーリエ係数の時間変化の補間処理における設定した時間分解能に等しくして、逐次時間窓を移動させながらフーリエ変換することにより、図10に例示したように実現する。

なお、デジタル音声信号を、デジタルフィルターにより高周波成分を除去し、請求項2記載の計測時間分解能と等しい時間窓毎の短時間一次元フーリエ変換を低周波領域デジタル音声信号に施し、該デジタル音声信号に比べてサンプリング間隔を長くした周波数領域の音声信号時系列を得て、該音声信号時系列を、基本音声信号の周波数領域の時系列データとすることもできる。

40

(A3) 時間窓ずれ補正処理

フーリエ係数の時間窓相当の時間ずれ修正は、時間窓の1/2だけ時間がずれることの補正で、時間方向における相当量のデータのシフトにより実現する。基本周波数の時系列データ数を $NA + 1$ とする。

(A4) 差分による近似導関数処理

時間差分による近似導関数の計算は、隣接するデータの差とする。

50

(A5) 周波数領域時系列データ抽出

二乗平均の計算は、差の二乗の平方根を計算することにより実現する。時系列データ数はNAとなる。これを音声データの周波数領域特徴量の時系列データとする。NA=NVである。その例を図11に示す。

なお、基本映像信号は、フレーム内で二次元フーリエ変換して、空間周波数成分ごとのフレーム内平均値を計算し、周波数領域における映像時系列データを生成し、基本音声信号を、周波数成分ごとの時間窓内平均値を計算し、周波数領域における音声時系列データを生成し、計測時間分解能と等しい時間窓毎の2つの情報を生成することもできる。

【0026】

本発明に係る上記前処理されたデジタル映像信号とそれに伴ったデジタル音声信号におけるずれ時間の同定処理について、図12および13により以下に説明する。

(4) 同定処理について

ずれ時間の同定は、上記(2)(V8)による映像データの周波数領域特徴量の時系列データと、上記(3)(A5)による音声データの周波数領域特徴量の時系列データに対して、図12に示した手順で行う。同一の長さNA=NVの映像データの周波数領域特徴量の時系列データと音声データの周波数領域特徴量の時系列データとの間の相互相関係数を計算する。以下にさらに説明する。

まず、上記映像時系列データと上記音声時系列データの相互相関係数を計算して、該相互相関係数の最大点を検出し、該最大点に対応したデータ位置を求める。次に、検出された相互相関係数の最大点とする該サンプル点を、サンプリング間隔に相当する時間から、デジタル映像信号とデジタル音声信号との間の時間的なずれと同定する。相互相関係数の計算は、例えば、二つのランダム系列 x_i, y_j の相互相関係数 $R_{xy}(m)$ の計算を式(1)のように適用する。

【数1】

$$R_{xy}(m) = E\{x_{n+m}y_n^*\} = E\{x_n y_{n-m}^*\} \quad (1)$$

ここで、*は複素共役、 $E\{\cdot\}$ は期待値で、 x_n, y_n は同時生起定常ランダム系列であり、 $- < n < +$ が想定される、しかし、実際の問題では一定時間のサンプリング(サンプリング数、 $N > 1$)を使って相互相関係数を式(2)のように計算することになる。

【数2】

$$\begin{aligned} \hat{R}_{xy}(m) &= \sum_{n=0}^{N-m-1} x_{n+m} y_n^* \quad m \geq 0 \\ \hat{R}_{xy}(m) &= \hat{R}_{yx}^*(-m) \quad m < 0 \end{aligned} \quad (2)$$

【0027】

(測定結果、補正)

本実施例による測定結果を図13に示す。相互相関係数が最大となるサンプルデータ移動量に相当する時間をデジタル映像信号とデジタル音声信号との間の時間的なずれと同定することによりずれ時間を計測する。相関係数0.6681(最大値)の時、音声データに対して、映像データが、-0.0227秒遅延していることが判明して、高精度に測定できることを実証した。従って、このずれ時間に相当する音声信号サンプルを遅延することにより、デジタル映像信号とデジタル音声信号との間の時間的なずれを補正することができる。即ち、上記(4)により同定されたデジタル映像信号とデジタル音声信号との間の時間的なずれに相当するサンプル数を、デジタル音声に対してサンプルシフトを施し、デジタル音声を遅延して補正することによりデジタル映像信号とデジタル音声信号との間の時間的なずれを補正する。

【 0 0 2 8 】

上記前処理されたデジタル映像信号とそれに伴ったデジタル音声信号におけるずれ時間のずれを計測・補正する方法を実現するための本発明に係る装置について、以下に説明する。

(5) ハードウェア

全ての信号をあらかじめA/D変換回路でデジタル信号に変換し、記録してあるので、高性能・高速マイクロプロセッサで実現できる。データ処理量が多い高速フーリエ変換については、DSPなどの専用プロセッサを併用することができる。簡素な構成で小型化が実現できるので、放送局ばかりか取材現場や中継局あるいは受信装置においてこれを搭載することができる。

10

【 0 0 2 9 】

(6) 効果 (処理結果)

以上のように構成された本発明の方法・装置による効果について、さらに以下に説明する。

(測定精度)

測定精度に関して、公共放送局 (NHK) が放送したテレビニュースをデジタルビデオ (非圧縮AVI形式) として録画して、本発明を実施した。測定精度の効果を確認するために、音声データを分離し、意図的に時間ずれを導入した音声データと映像データのずれ時間を本発明方式により同定した。この結果、ずれ時間の高精度測定を実現した。従来方法・装置では、映像フレーム時間間隔の制約を受けるが、

20

該周波数領域における時系列データの相互相関を統計的に同定する方法によるので、この制約を受けることがない。

【 0 0 3 0 】

(相互相関の時計処理)

相互相関などの統計処理の効果としては、映像信号の特徴抽出及び音声信号の特徴抽出に基礎統計量を使い、それにより抽出した映像信号の時系列データと音声信号の時系列データ間の統計的な相互相関量を計算することにより、両時系列データ間のずれに相当するデータサンプル量を同定できるという効果を確認した。

【 0 0 3 1 】

(周波数領域における同定)

周波数領域における同定については、時間領域におけるデジタル映像信号及びデジタル音声信号を短時間窓高速フーリエ変換することにより、映像の時間的な変化とそれに伴う音声の時間的な変化を関連した周波数の時間的な変化として捕らえことができ、その効果を確認した。

30

【 0 0 3 2 】

(デジタルデータ対応)

デジタルデータ対応については、本実施例では対象をデジタルビデオとしているものの、アナログビデオについても、(効果 1) に示したように、公知の技術であるA/D変換により容易にデジタルビデオを得ることができるので、本発明の効果を減ずるものではないことを確認した。

40

【 0 0 3 3 】

(簡易性)

簡易性については、本発明のソフトウェアを、磁気記録、光学記録、光磁気記録あるいはメモリ等の記録媒体に記録して、汎用のパーソナルコンピュータに実装することにより、簡便に実施でき、その効果を確認した。なお、パーソナルコンピュータによる計算はデータの積和が主要なものであるから、公知の技術であるDSPなどのハードウェアで置き換えることができ、さらに、小型化、高速化できる。

【 0 0 3 4 】

(まとめ)

以上から、上記課題 1、2、3、4 および 5 を解決して、映像と音声のずれ時間を、特

50

に、従来測定ができなかった高精度なデジタル映像とそれに伴ったデジタル音声の時間的なずれを測定できることが判明した。

【実施例 2】

【0035】

さらに、本発明で用いられる他の実施例を示す。

インターネットで広範に活用されている圧縮デジタルオーディオ・ビデオ（例えば、リアルビデオやウインドウメディア）についても、ネットワーク packets を対象にしてこれを AVI ファイル形式に変換する以外は、実施例 1 と同様の構成である。これにより、あらかじめコンピュータ可読形式のファイルとして記録されている媒体以外の実時間のデジタル映像信号とそれに伴ったデジタル音声信号についても、同等の効果を奏する。

10

【実施例 3】

【0036】

加えて、本発明で用いられる他の実施例を示す。

デジタル映像信号とそれに伴ったデジタル音声信号を再現する装置（デジタルシネマシステム、ホームシアターシステム、パーソナルコンピュータ、ゲーム機等）では、例外なくデジタル映像信号処理再現に要する時間は、デジタル音声信号処理再現に要する時間に比べて長くなる。このため、デジタル音声信号を遅延させるが、遅延時間を計測結果に基づいて加減することにより、高精度でデジタル映像信号とそれに伴ったデジタル音声信号を同期させることができる。この機構を再現装置に組み込む以外は、実施例 1 と同様の構成である。これにより、再現対象のデジタル映像信号とそれに伴ったデジタル音声信号の同期のみならず、システム全体としての正しい同期を実現するという効果を奏する。

20

【0037】

（他の実施例に関する付記）

本発明は、上記の実施例に限定されるものではなく、発明の技術的範囲内において、様々な変更をすることができる。また、極めて高精度かつ簡易な構成で実現できるため、信号処理装置と組み合わせ、リアルタイムでの動画と音声のずれ補正、記録、再生システムが可能であるため、放送局設備ばかりか小型中継局へ設置できる。

【0038】

また、デジタルビデオカメラへ搭載できる。あるいは、高精細のプラズマディスプレイ装置に応用できる。さらには、本発明に係るデジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する装置を備えたプロジェクターにより、デジタルシネマシステム、インターネットを利用したホームシアターシステム、テレビ機能を有するパーソナルコンピュータなどに応用することができる。

30

【産業上の利用可能性】

【0039】

以上説明したように、本発明は、映像と音声のずれ時間を、高精度かつ簡便に計測する方法および計測装置を提供する。簡易な構成にも関わらず、極めて高精度に、デジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する装置を実現できるため、上記のように、リアルタイムでデジタル映像信号とデジタル音声信号との間の時間的なずれを計測または補正する記録、通信、再生システムが可能であるため、極めて、有用である。

40

【図面の簡単な説明】

【0040】

【図 1】本発明に係る方法を示す全体図

【図 2】本発明に係る映像データ前処理手順を示す図

【図 3】口元部分の抽出画像（モノクロ画像）の例示

【図 4】口元部分コントラスト強調処理の例示

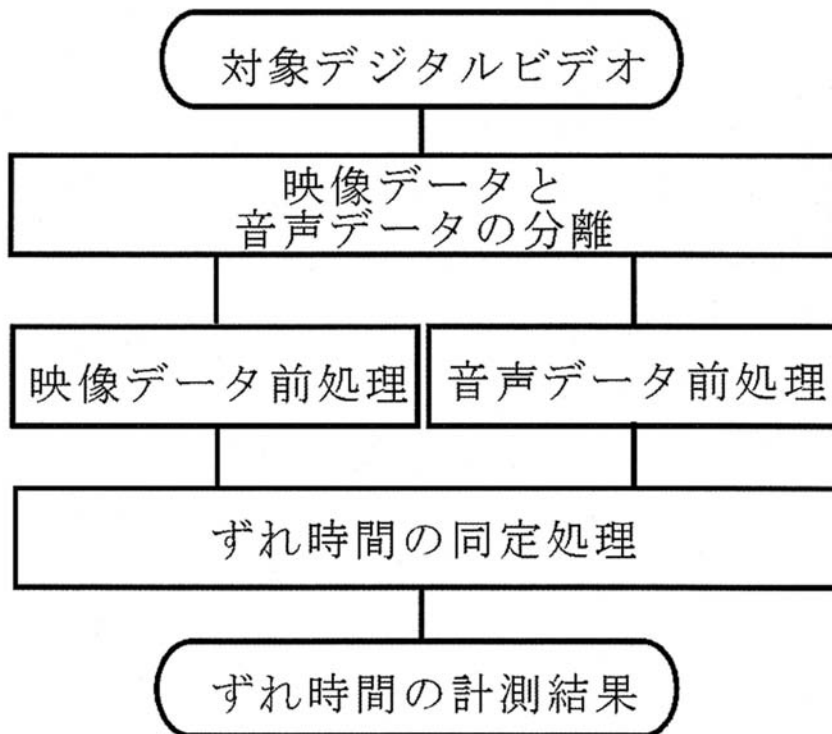
【図 5】口元画像の 2 次元フーリエ変換の例示

【図 6】空間周波数成分の低域平均の時間的な変化

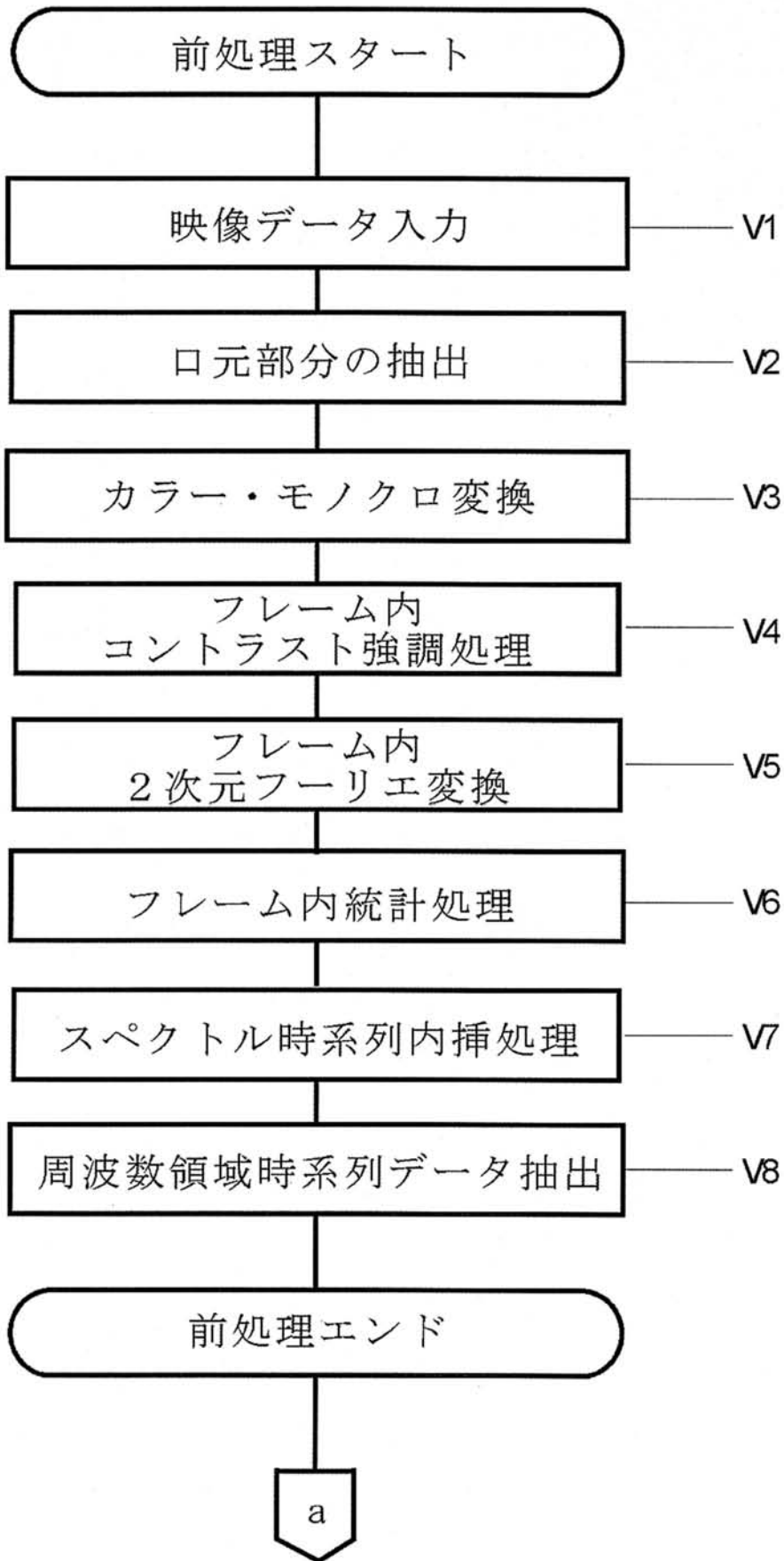
50

- 【図7】映像データの周波数領域特徴量の時系列データの例示
- 【図8】本発明に係る音声データ前処理手順を示す図
- 【図9】音声データサンプルの時間的な変化の例示
- 【図10】音声サンプルの短時間窓1次元高速フーリエ係数の時間的な変化の例示
- 【図11】音声データの周波数領域特徴量の時系列データの例示
- 【図12】本発明に係るずれ時間の同定処理手順の図
- 【図13】映像データの周波数領域特徴量の時系列データと音声データの周波数領域特徴量の時系列データとの間の相互相関係数の例示

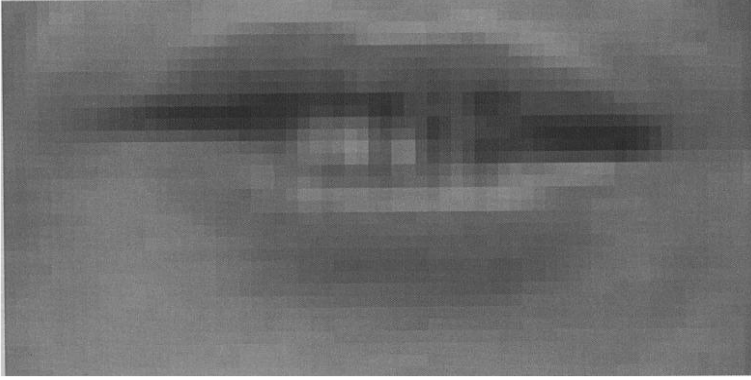
【図1】



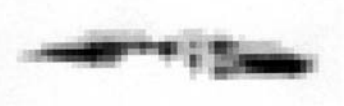
【図2】



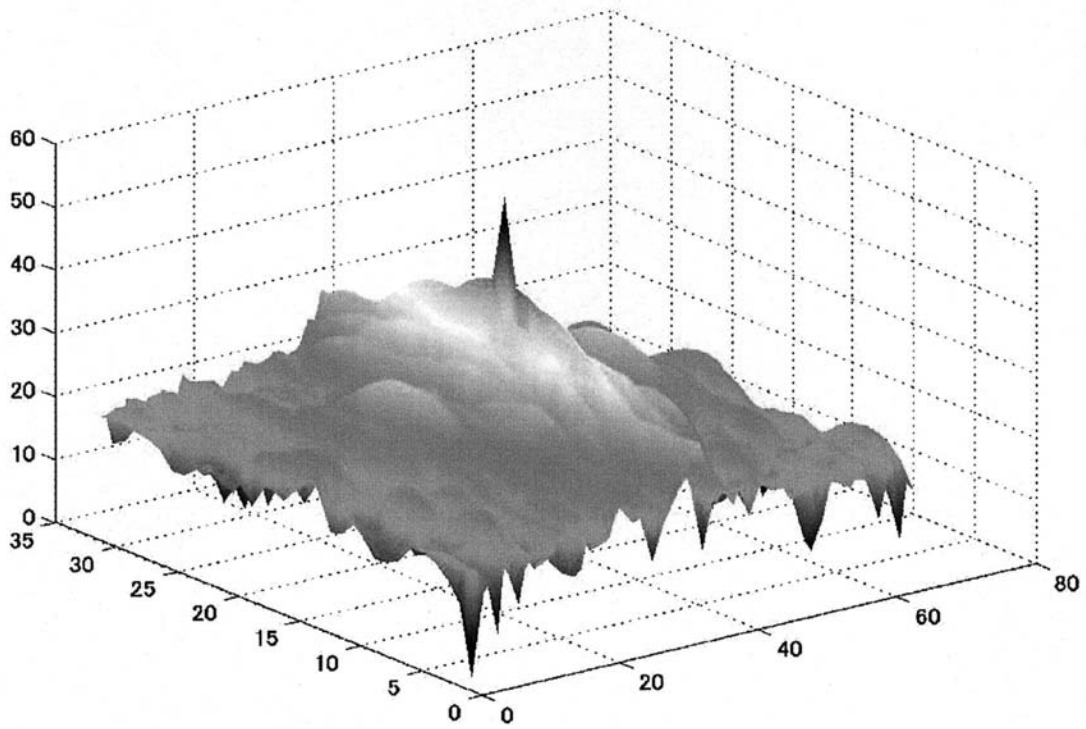
【 図 3 】



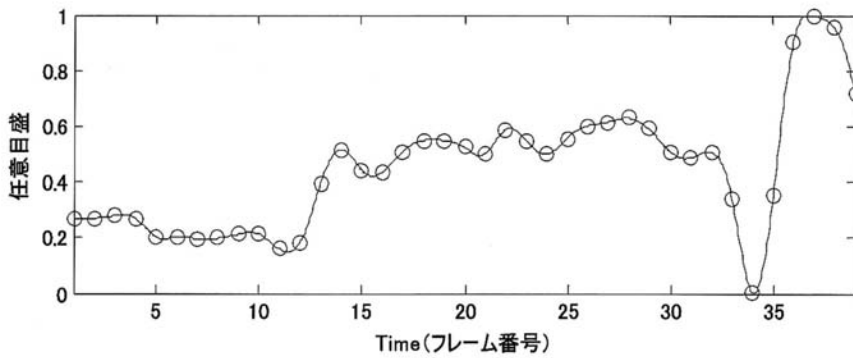
【 図 4 】



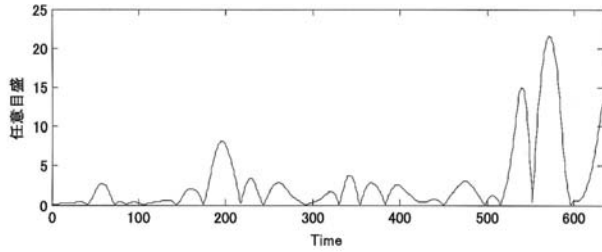
【 図 5 】



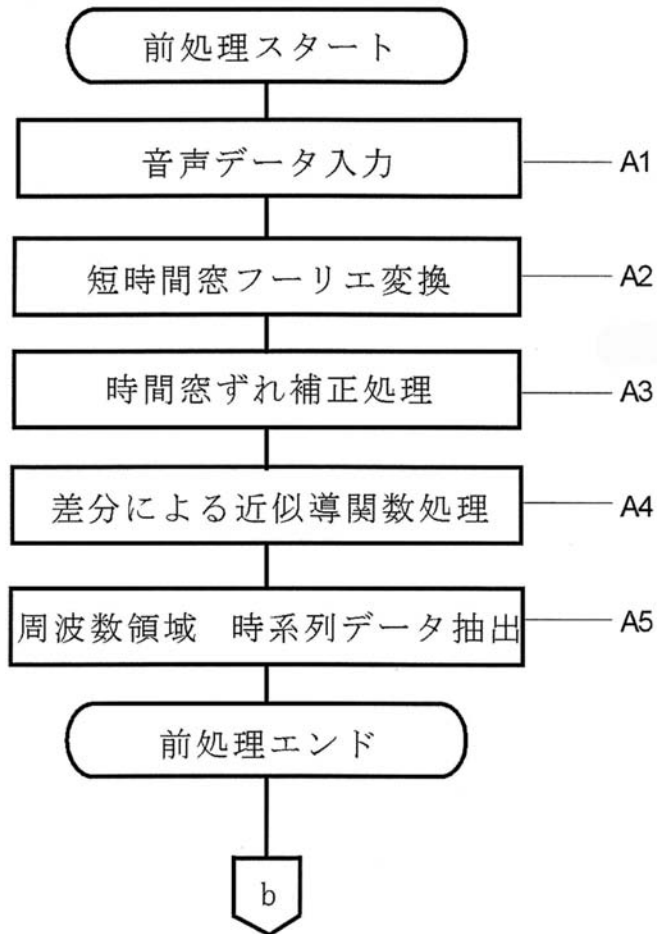
【 図 6 】



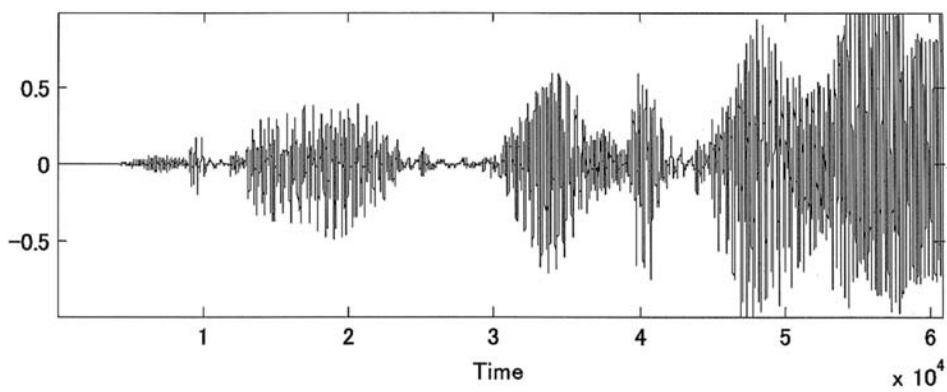
【 図 7 】



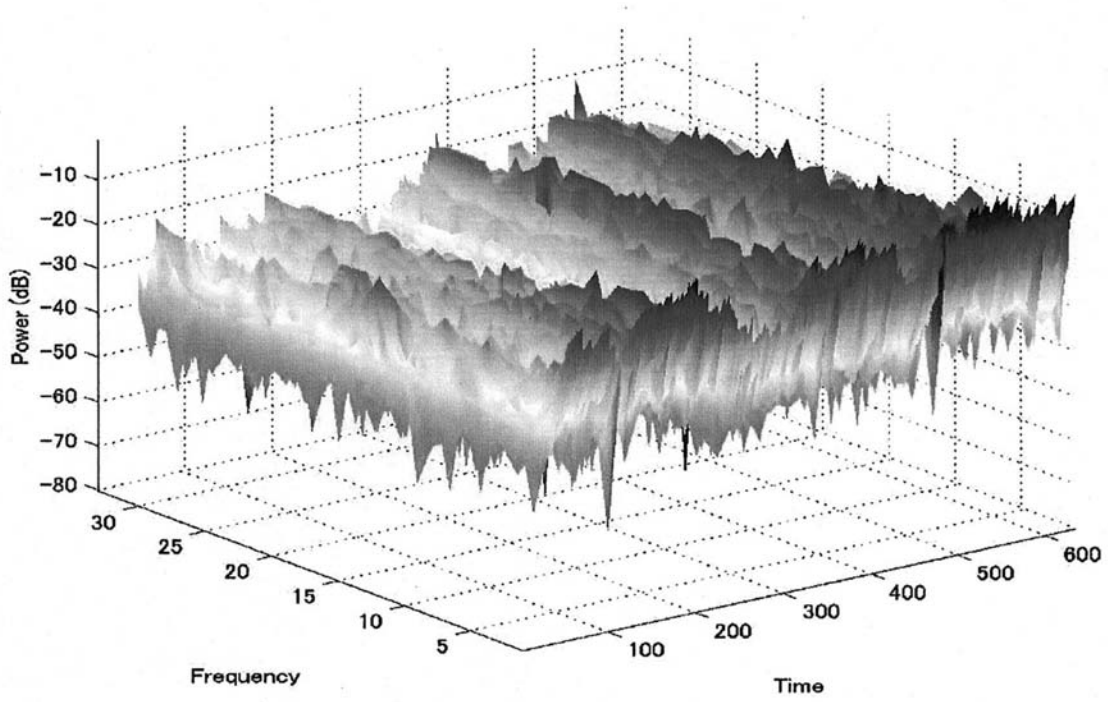
【 図 8 】



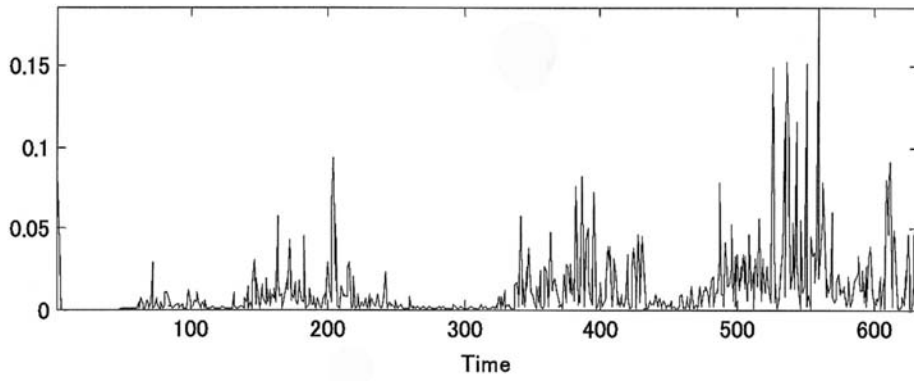
【 図 9 】



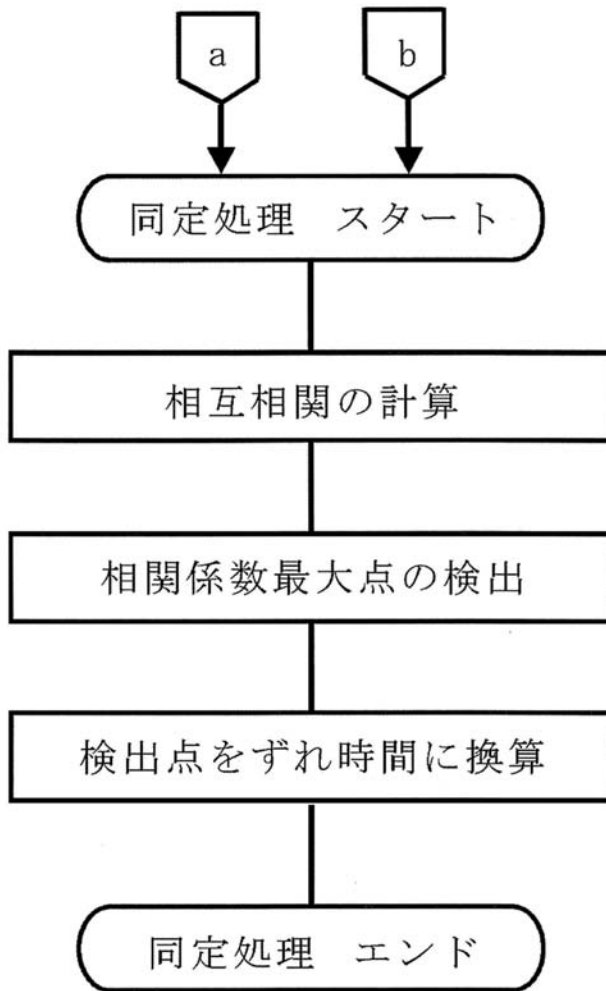
【 1 0 】



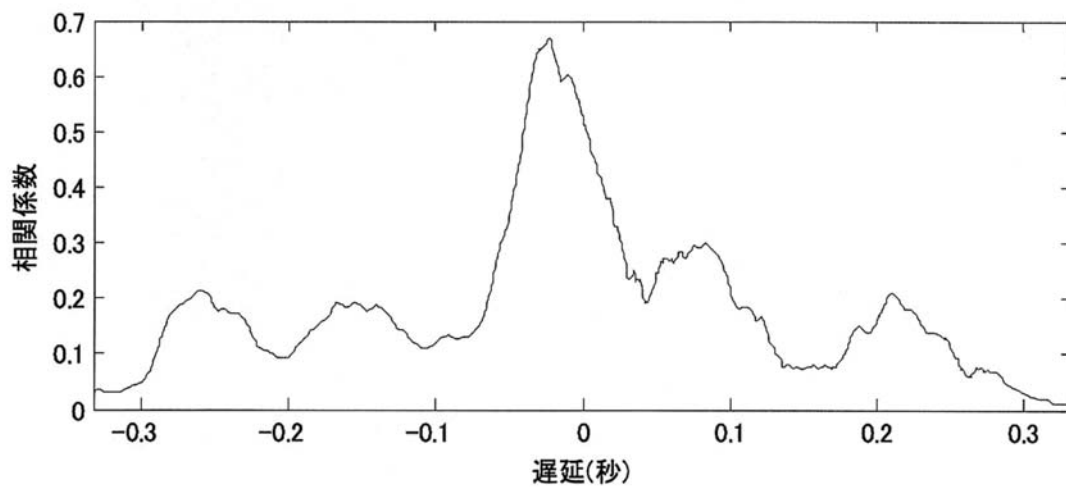
【 1 1 】



【 図 1 2 】



【 図 1 3 】



フロントページの続き

- (56)参考文献 特開2003-046901(JP,A)
特開平09-130789(JP,A)
特開2005-033314(JP,A)
特開平09-083469(JP,A)
特開平07-177479(JP,A)
特開平10-145645(JP,A)
国際公開第2006/100727(WO,A1)

(58)調査した分野(Int.Cl., DB名)

H04N 5/50-5/63
H04N 7/173
H04N 17/04