

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5467346号
(P5467346)

(45) 発行日 平成26年4月9日(2014.4.9)

(24) 登録日 平成26年2月7日(2014.2.7)

(51) Int.Cl.		F I			
G06T	7/20	(2006.01)	G06T	7/20	B
H04N	19/50	(2014.01)	H04N	7/137	Z
G06T	7/40	(2006.01)	G06T	7/40	B

請求項の数 3 (全 26 頁)

(21) 出願番号	特願2009-298796 (P2009-298796)	(73) 特許権者	305060567
(22) 出願日	平成21年12月28日(2009.12.28)		国立大学法人富山大学
(65) 公開番号	特開2011-81761 (P2011-81761A)		富山県富山市五福3190
(43) 公開日	平成23年4月21日(2011.4.21)	(74) 代理人	100105809
審査請求日	平成24年12月14日(2012.12.14)		弁理士 木森 有平
(31) 優先権主張番号	特願2009-210991 (P2009-210991)	(74) 代理人	100126398
(32) 優先日	平成21年9月11日(2009.9.11)		弁理士 浅野 典子
(33) 優先権主張国	日本国(JP)	(72) 発明者	広林 茂樹
			富山県富山市五福3190 国立大学法人富山大学内
		(72) 発明者	上田 貴晃
			富山県富山市五福3190 国立大学法人富山大学内
		審査官	佐藤 実

最終頁に続く

(54) 【発明の名称】 動き推定方法、動き推定装置、及び動き推定プログラム

(57) 【特許請求の範囲】

【請求項1】

解析対象となる動画データを、前記動画データの3次元信号と、周波数及び初期位相を用いた位相と振幅とによって表される正弦波モデル信号との差の二乗和が最小値になるような前記周波数、前記振幅及び前記初期位相を、非周期信号のフーリエ変換式のパラメータとして求め、時空間スペクトルを抽出し、抽出された前記時空間スペクトル分布を、平面をなすスペクトル群毎に切り分けして1つ又は複数の平面群を推定し、推定された前記平面群の傾きを求めることで、前記動画内の1つ又は複数の移動物体の動きを推定する動き推定方法であって、

前記時空間スペクトルに対して最小二乗法を用いたクラスタリングを行い、1つ又は複数の平面群を推定するに際し、前記時空間スペクトルが所定の平面に属するものとした場合の存在確率変数を1又は小数に設定し、最小二乗法を用いたクラスタリングを行うことを特徴とする動き推定方法。

【請求項2】

解析対象となる動画データを入力する動画入力手段と、入力された前記動画データの3次元信号と、周波数及び初期位相を用いた位相と振幅とによって表される正弦波モデル信号との差の二乗和が最小値になるような前記周波数、前記振幅及び前記初期位相を、非周期信号のフーリエ変換式のパラメータとして求め、時空間スペクトルを抽出する時空間スペクトル抽出手段と、抽出された前記時空間スペクトル分布を、平面をなすスペクトル群毎に切り分けして1つ又は複数の平面群を推定する平面推定手段と、推定された前

記平面群の傾きを求める傾き算出手段とが備わっており、前記動画像内の1つ又は複数の移動物体の動きを推定する動き推定装置であって、

前記時空間スペクトルに対して最小二乗法を用いたクラスタリングを行い、1つ又は複数の平面群を推定するに際し、前記時空間スペクトルが所定の平面に属するものとした場合の存在確率変数を1又は小数に設定し、最小二乗法を用いたクラスタリングを行う構成であることを特徴とする動き推定装置。

【請求項3】

解析対象となる動画像データを入力する動画像入力手段と、入力された前記動画像データの3次元信号と、周波数及び初期位相を用いた位相と振幅とによって表される正弦波モデル信号との差の二乗和が最小値になるような前記周波数、前記振幅及び前記初期位相を、非周期信号のフーリエ変換式のパラメータとして求め、時空間スペクトルを抽出する時空間スペクトル抽出手段、前記時空間スペクトル分布を、平面をなすスペクトル群毎に切り分けして1つ又は複数の平面群を推定する平面推定手段、及び、推定された前記平面群の傾きを求める傾き算出手段としてコンピュータを機能させ、前記動画像内の1つ又は複数の移動物体の動きを推定するコンピュータ実行可能な動き推定プログラムであって、

10

前記時空間スペクトルに対して最小二乗法を用いたクラスタリングを行い、1つ又は複数の平面群を推定するに際し、前記時空間スペクトルが所定の平面に属するものとした場合の存在確率変数を1又は小数に設定し、最小二乗法を用いたクラスタリングを行うことを特徴とする動き推定プログラム。

【発明の詳細な説明】

20

【技術分野】

【0001】

本発明は、動画像内の移動物体の動き解析や動画像の圧縮符号化の分野に関し、特に、動画像内の1つ又は複数の移動物体の動きを推定する動き推定方法、動き推定装置、及び動き推定プログラムに関する。

【背景技術】

【0002】

動画像の圧縮符号化において、参照フレームからの移動物体の動き量を予測する動き予測技術は、動画像データの記述量を大幅に削減する上で極めて重要な技術である。

【0003】

30

動き予測に広く使われている方法としては、MPEG-4 AVC/H.264として規格化されている方法が知られている。MPEG-4 AVC/H.264の動画像符号化技術で用いられる動き予測技術においては、動画像から得られるデータを時間領域又は空間領域において解析することによって動き予測を行っている。すなわち、この動き予測技術は、2次元の画素ブロックに対して整数精度の直交変換を行う解析手法であり、また、その解析は、動画像データを構成する前後のフレームを利用し、これら前後のフレームの画素情報から予測した予測誤差画像に対して行っている。

【0004】

しかしながら、このような動き予測は、画素ブロック単位でのパターン認識に基づくものであることから、異なる方向へ移動する複数の物体や、複雑に形状を変えながら移動する物体については対応することができず、さらに別の手法を必要とする。

40

【0005】

このような別の手法としては、動画像を、高さ、幅、時間からなる3次元信号として捉えて3次元解析し、得られた時空間スペクトルを利用して動画像内の移動物体の動きを予測する手法がある。

【0006】

一般的に、移動物体を含む動画像を3次元解析して得られた時空間スペクトルのうち、高エネルギーを有するスペクトルに注目すると、これらのスペクトルは、3次元周波数領域空間において平面状に分布する平面群となる。この平面群の構成は、動画像内の物体の動きによって決まり、平面群の傾きを求めることによって動き予測が可能となる。したが

50

って、複数のそれぞれ異なる動きをする移動物体を含む動画像の場合には、移動物体の数と平面の数とが等しくなるため、各平面の傾きをそれぞれ求めることにより、個々の物体の動きを予測することが可能となる。

【 0 0 0 7 】

しかしながら、周波数解析手法として通常用いられる高速フーリエ変換 (Fast Fourier Transform ; F F T) や離散フーリエ変換 (Discrete Fourier Transform ; D F T) に代表される、等間隔な周波数分解幅を有する解析手法においては、分解能がフレーム長に依存し、動画像の解析において十分な分解能が得られないことから、分解能不足に起因して解析できない動きに対応するためには、さらに別の手法を組み合わせる必要がある。

10

【 0 0 0 8 】

例えば、非特許文献 1 には、F F T を用いて動画像を 3 次元解析して時空間スペクトルを得た後、時空間スペクトルで表される平面群の傾きを決定する方法として直線群フィルタを用いる手法が開示されている。

【 0 0 0 9 】

また、非特許文献 2 に記載された技術においては、平面推定を前提とした F F T アルゴリズムを利用して時空間周波数スペクトルを得ることによって動画像を 3 次元解析し、平面推定時にファジー推論を使用してスペクトルデータの誤差を補正することによって解析精度の向上を図っている。

【 0 0 1 0 】

20

一方、特許文献 1 には、本願発明者の一部が考案した周波数解析手法として、非周期信号の解析手法である Non-Harmonic Analysis (N H A) が開示されている。この N H A は、解析対象信号と、周波数 f' 及び初期位相 ϕ' を用いた位相と振幅 A' とによって表される正弦波モデル信号との差の二乗和が最小値になるような周波数 f' 、振幅 A' 、及び初期位相 ϕ' を、非周期信号のフーリエ変換式のパラメータとして算出するものである。

【 先行技術文献 】

【 特許文献 】

【 0 0 1 1 】

【 特許文献 1 】 国際公開第 2 0 0 9 / 0 3 8 0 5 6 号

【 非特許文献 】

30

【 0 0 1 2 】

【 非特許文献 1 】 Akira KOJIMA and Jun-ichi HISHIGAMI, "Motion Detection using 3D-FFT Spectrum", ITEC '92, 1992年

【 非特許文献 2 】 C. E. Erdem, G. Karabulut, E. Yanmaz and E. Anarim, "Motion Estimation in the frequency domain using fuzzy c-planes clustering", vol.10, pp. 1873-1879, 2001年12月

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 1 3 】

しかしながら、非特許文献 1 に記載された技術においては、得られる時空間スペクトルの精度が向上するわけではなく、移動物体の予測精度にも限界がある。また、非特許文献 2 に記載された技術においては、ファジー推論のために複雑な処理を必要とし、計算コストが高くなるという問題があった。

40

【 0 0 1 4 】

本発明は、このような実情に鑑みてなされたものであり、高精度に且つ少ない演算量で、動画像を 3 次元解析して得られた時空間スペクトルを利用して動画像内の移動物体の動きを推定することができる動き推定方法、動き推定装置、及び動き推定プログラムを提供することを目的とする。

【 課題を解決するための手段 】

【 0 0 1 5 】

50

本発明の動き推定方法は、解析対象となる動画像データを、前記動画像データの3次元信号と、周波数及び初期位相を用いた位相と振幅とによって表される正弦波モデル信号との差の二乗和が最小値になるような前記周波数、前記振幅及び前記初期位相を、非周期信号のフーリエ変換式のパラメータとして求め、時空間スペクトルを抽出し、抽出された前記時空間スペクトル分布を、平面をなすスペクトル群毎に切り分けして1つ又は複数の平面群を推定し、推定された前記平面群の傾きを求めることで、前記動画像内の1つ又は複数の移動物体の動きを推定する動き推定方法であって、

前記時空間スペクトルに対して最小二乗法を用いたクラスタリングを行い、1つ又は複数の平面群を推定するに際し、前記時空間スペクトルが所定の平面に属するものとした場合の存在確率変数を1又は小数に設定し、最小二乗法を用いたクラスタリングを行うこと

10

【0016】

本発明の動き推定装置は、解析対象となる動画像データを入力する動画像入力手段と、入力された前記動画像データの3次元信号と、周波数及び初期位相を用いた位相と振幅とによって表される正弦波モデル信号との差の二乗和が最小値になるような前記周波数、前記振幅及び前記初期位相を、非周期信号のフーリエ変換式のパラメータとして求め、時空間スペクトルを抽出する時空間スペクトル抽出手段と、抽出された前記時空間スペクトル分布を、平面をなすスペクトル群毎に切り分けして1つ又は複数の平面群を推定する平面推定手段と、推定された前記平面群の傾きを求める傾き算出手段とが備わっており、前記動画像内の1つ又は複数の移動物体の動きを推定する動き推定装置であって、

20

前記時空間スペクトルに対して最小二乗法を用いたクラスタリングを行い、1つ又は複数の平面群を推定するに際し、前記時空間スペクトルが所定の平面に属するものとした場合の存在確率変数を1又は小数に設定し、最小二乗法を用いたクラスタリングを行う構成

であることを特徴とする。

【0017】

本発明の動き推定プログラムは、解析対象となる動画像データを入力する動画像入力手段と、入力された前記動画像データの3次元信号と、周波数及び初期位相を用いた位相と振幅とによって表される正弦波モデル信号との差の二乗和が最小値になるような前記周波数、前記振幅及び前記初期位相を、非周期信号のフーリエ変換式のパラメータとして求め、時空間スペクトルを抽出する時空間スペクトル抽出手段、前記時空間スペクトル分布を、平面をなすスペクトル群毎に切り分けして1つ又は複数の平面群を推定する平面推定手段、及び、推定された前記平面群の傾きを求める傾き算出手段としてコンピュータを機能させ、前記動画像内の1つ又は複数の移動物体の動きを推定するコンピュータ実行可能な動き推定プログラムであって、

30

前記時空間スペクトルに対して最小二乗法を用いたクラスタリングを行い、1つ又は複数の平面群を推定するに際し、前記時空間スペクトルが所定の平面に属するものとした場合の存在確率変数を1又は小数に設定し、最小二乗法を用いたクラスタリングを行うこと

を特徴とする動き推定プログラム。

【0018】

このような本発明にかかる動き推定方法、動き推定装置、及び動き推定プログラムが実装された装置においては、周波数分解能が解析窓長に依存しない周波数解析手法を動画像の3次元信号解析に用いることにより、従来の周波数解析手法を用いた場合に比べて高精度に時空間スペクトルを抽出することができ、動画像内の移動物体の動きが周波数領域空間に精度良く反映される。そのため、本発明にかかる動き推定方法、動き推定装置、及び動き推定プログラムが実装された装置においては、得られた時空間スペクトルのうち、高エネルギーを有するスペクトルの理論平面からの誤差を極めて少なくすることができる。

40

【発明の効果】

【0019】

本発明においては、周波数解析によって得られた時空間スペクトル分布を高精度に求めることができるため、動きの平面を推定する工程では複雑な処理を必要とせず、高精度に

50

且つ少ない演算量で動画像内の移動物体の動きを推定することができる。

【図面の簡単な説明】

【0020】

【図1】本発明の実施の形態として示す動き推定装置の構成を示すブロック図である。

【図2】本周波数解析手法とDFTとGHAとの違いを説明するための図であり、各手法の誤差を求めた結果を示す図である。

【図3】本発明の実施の形態として示す動き推定装置において、動画像内の1つ又は複数の移動物体の動きを推定する際の一連の処理を示すフローチャートである。

【図4】第1の検証にて用いた動画像を示す図である。

【図5】第1の検証において物体の移動速度 V_x を1.0ピクセル/フレームとした場合に、3次元FFTを用いて求めた時空間スペクトル分布を示す図である。

【図6】第1の検証において物体の移動速度 V_x を3.3ピクセル/フレームとした場合に、3次元FFTを用いて求めた時空間スペクトル分布を示す図である。

【図7】第1の検証において物体の移動速度 V_x を3.3ピクセル/フレームとした場合に、本周波数解析手法を用いて求めた時空間スペクトル分布を示す図である。

【図8】第2の検証にて用いた動画像を示す図である。

【図9】第2の検証において3次元FFTを用いて求めた時空間スペクトル分布を示す図である。

【図10】第2の検証において本周波数解析手法を用いて求めた時空間スペクトル分布を示す図である。

【図11】第3の検証にて用いた動画像を示す図である。

【図12】第3の検証において3次元FFTを用いて求めた時空間スペクトル分布を示す図である。

【図13】第3の検証において本周波数解析手法を用いて求めた時空間スペクトル分布を示す図である。

【図14】図13の時空間スペクトル分布の視点を変えて平面を描画した様子を示す図である。

【図15】類似する複数の動きを含む動画像について3次元FFTを用いて求めた時空間スペクトル分布を示す図である。

【図16】類似する複数の動きを含む動画像について本周波数解析手法を用いて求めた時空間スペクトル分布を示す図である。

【図17】図15の場合とは異なる複数の動きを含む動画像について3次元FFTを用いて求めた時空間スペクトル分布を示す図である。

【図18】図16の場合とは異なる複数の動きを含む動画像について3次元FFTを用いて求めた時空間スペクトル分布を示す図である。

【図19】第4の検証にて用いた動画像を示す図である。

【図20】第4の検証において本周波数解析手法を用いて求めた100本の時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

【図21】第4の検証において本周波数解析手法を用いて求めた125本の時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

【図22】第4の検証において本周波数解析手法を用いて求めた160本の時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

【図23】第4の検証において本周波数解析手法を用いて求めた180本の時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

【図24】第4の検証において本周波数解析手法を用いて求めた200本の時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

【図25】第4の検証において本周波数解析手法を用いて求めた225本の時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

【図26】第4の検証において本周波数解析手法を用いて求めた250本の時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

10

20

30

40

50

【図 27】第 5 の検証にて用いた動画像を示す図である。

【図 28】第 5 の検証において 3 次元 F F T を用いて求めた時空間スペクトル分布を示す図である。

【図 29】第 5 の検証において本周波数解析手法を用いて求めた時空間スペクトル分布を示す図である。

【図 30】第 6 の検証にて用いた動画像を示す図である。

【図 31】第 6 の検証において 3 次元 F F T を用いて求めた時空間スペクトル分布を示す図である。

【図 32】第 6 の検証において本周波数解析手法を用いて求めた時空間スペクトル分布を示す図である。

10

【図 33】第 7 の検証にて用いた動画像を示す図である。

【図 34】第 7 の検証において 3 次元 F F T を用いて求めた時空間スペクトル分布を示す図である。

【図 35】第 7 の検証において本周波数解析手法を用いて求めた時空間スペクトル分布を示す図である。

【図 36】図 34 に示す時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

【図 37】図 35 に示す時空間スペクトルに基づいて推定した平面を描画した様子を示す図である。

【発明を実施するための形態】

20

【0021】

以下、本発明を適用した具体的な実施の形態について図面を参照しながら詳細に説明する。

【0022】

この実施の形態は、動画像内の 1 つ又は複数の移動物体の動きを推定する動き推定装置である。特に、この動き推定装置は、非線形方程式を解くことでフーリエ係数を推定することによって周波数分解能が解析窓長に依存しない新たな周波数解析手法を利用して動画像の 3 次元信号解析を行うものである。

【0023】

[動き推定装置の構成]

30

動き推定装置は、例えばコンピュータ等から構成され、図 1 に示すように、各部を統括的に制御する CPU (Central Processing Unit) 11 と、各種プログラムを含む各種情報を格納する読み取り専用の ROM (Read Only Memory) 12 と、ワークエリアとして機能する RAM (Random Access Memory) 13 と、各種情報を読み出し及び / 又は書き込み可能に記憶する記憶部 14 と、ユーザインターフェースとしての図示しない所定の操作デバイスを介した入力操作の処理及び制御を行う入力操作制御部 15 と、各種情報を表示する表示部 16 とを備える。

【0024】

CPU 11 は、記憶部 14 等に格納されている各種アプリケーションプログラムをはじめとする各種プログラムを実行し、各部を統括的に制御する。

40

【0025】

ROM 12 は、各種プログラムをはじめとする各種情報を格納している。この ROM 12 に格納されている情報は、CPU 11 の制御のもとに読み出される。

【0026】

RAM 13 は、CPU 11 が各種プログラムを実行する際のワークエリアとして機能し、CPU 11 の制御のもとに、各種情報を一時記憶するとともに、記憶している各種情報を読み出す。

【0027】

記憶部 14 は、本発明にかかる動き推定プログラム等のアプリケーションプログラムの

50

他、解析対象となる動画像データをはじめとする各種情報を記憶する。この記憶部 14 としては、例えば、ハードディスクや不揮発性メモリ等を用いることができる。また、記憶部 14 には、本体に対して着脱可能とされるフレキシブルディスクやメモリカード等の記憶媒体に対して、各種情報の読み出し及び/又は書き込みを行うドライブ装置も含まれる。この記憶部 14 に記憶されている各種情報は、CPU 11 の制御のもとに読み出される。

【0028】

入力操作制御部 15 は、例えば、キーボード、マウス、キーパッド、赤外線リモートコントローラ、スティックキー、又はプッシュボタンといった、ユーザインターフェースとしての図示しない所定の操作デバイスを介した入力操作を受け付け、操作内容を示す制御信号を CPU 11 に対して供給する。

10

【0029】

表示部 16 は、例えば、液晶ディスプレイ (Liquid Crystal Display ; LCD)、プラズマ・ディスプレイ・パネル (Plasma Display Panel ; PDP)、有機エレクトロルミネッセンス (Organic ElectroLuminescent) ディスプレイ、又は CRT (Cathode Ray Tube) といった、各種表示デバイスであり、CPU 11 の制御のもとに各種情報を表示する。例えば、表示部 16 は、CPU 11 によって動き推定プログラムが起動されると、その画面を表示し、入力された解析対象としての動画像データや動き推定結果等を表示する。

【0030】

20

このような各部を備える動き推定装置は、CPU 11 の制御のもとに、動き推定プログラムを実行すると、CPU 11 の制御のもとに、入力された動画像データの周波数解析を行うことによって時空間スペクトルを抽出し、得られた時空間スペクトルについて最小二乗法を用いたクラスタリングを行うことにより、時空間スペクトル分布を、平面をなすスペクトル群毎に切り分けして 1 つ又は複数の平面群を推定する。なお、周波数解析の対象となる信号、すなわち、解析対象となる動画像データは、図示しない動画像入力部を介して CPU 11 に入力される。例えば、動き推定装置は、テレビ映像を録画して得られた動画像データの圧縮符号化を行う場合には、当該動き推定装置を含む録画機とテレビとを接続する所定のインターフェースを介して解析対象としての動画像データを入力する。すなわち、動画像入力部は、解析対象としての動画像データを CPU 11 に入力させる機能を有する部位である。なお、動画像入力部は、アナログ信号を入力した場合には、A/D 変換を行ってデジタル信号に変換する機能をあわせ持つことはいうまでもない。このとき、動画像入力部は、必要に応じてアンチエイリアシングフィルタを含む A/D 変換器としてもよい。動き推定装置は、CPU 11 の制御のもとに、このようにして入力された解析対象としての動画像データの周波数解析を行うことによって動き推定を行い、圧縮符号化したデータ等を、図示しない出力部を介して記憶部 14 に記憶させたり、その他の機器に出力したりする。

30

【0031】

[周波数解析アルゴリズム]

まず、動き推定装置における一連の動き推定アルゴリズムの説明に先立って、動き推定を行う際に利用する周波数解析アルゴリズムについて詳述する。なお、動画像データは 3 次元信号であることはいうまでもないが、ここでは、説明の便宜上、1 次元の解析対象信号に対する周波数解析アルゴリズムについて説明するものとする。

40

【0032】

動き推定装置に適用する周波数解析手法 (以下、本周波数解析手法という。) においては、次式 (1) に示す非周期信号のフーリエ変換式の周波数パラメータを求める問題を非線形方程式の最適解を求める問題に置き換えている。

【0033】

【数 1】

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \quad \dots (1)$$

【0034】

具体的には、本周波数解析手法においては、次式(2)に示すように、解析対象信号 $x(n)$ と正弦波モデル信号との差の二乗和で表される非線形方程式の最適解として、この非線形方程式の右辺が最小値になるような周波数 f' 、振幅 A' 、及び初期位相 ϕ' を求める。なお、次式(2)において、 L はフレーム長(解析窓長)であり、 f_s はサンプリング周波数 [Hz] である。本周波数解析手法においては、このような最小二乗法によって非線形方程式の最適解を求める問題に帰着させることにより、解析窓の影響やエイリアシングの影響がなくなり、解析窓長が、1周期未満であってもよく、周期の整数倍でなくてもよく、さらには、不等間隔であってもよい等、柔軟な周波数解析処理を実現することが可能となる。

10

【0035】

【数 2】

$$F(A', f', \phi') = \frac{1}{L} \sum_{n=0}^{L-1} \left\{ x(n) - A' \cos(2\pi f' \frac{n}{f_s} + \phi') \right\}^2 \quad \dots (2)$$

【0036】

さて、上式(2)に示す非線形方程式の最適解を実際に求めるにあたっては、以下のような方法をとることができる。

20

【0037】

本周波数解析手法においては、振幅 A' 、周波数 f' 、及び初期位相 ϕ' のそれぞれについて適切な初期値を求め、これら初期値から非線形方程式の解法を用いて最適解に収束させる。この非線形問題では、上式(2)をコスト関数とする最小化問題とする。なお、適切な初期値は、離散フーリエ変換(Discrete Fourier Transform; DFT)やウェーブレット変換等の任意の周波数変換を行ったり、フィルタリングを行うことによっておおよその見当をつけたりする等、既存の任意の方法を適用して求めることができる。

【0038】

まず、本周波数解析手法においては、上式(2)における正弦波モデル信号の位相を構成する周波数パラメータ f' 、 ϕ' について、いわゆる最急降下法を適用し、周波数パラメータ f'_m 、 ϕ'_m を次式(3)及び次式(4)によって求める。

30

【0039】

【数 3】

$$f'_{m+1} \leftarrow f'_m - \mu_m \frac{\partial F}{\partial f} \quad \dots (3)$$

【数 4】

$$\phi'_{m+1} \leftarrow \phi'_m - \mu_m \frac{\partial F}{\partial \phi} \quad \dots (4)$$

40

【0040】

なお、上式(3)及び上式(4)においては、次式(5)と略している。また、 μ_m は、いわゆる減速法に基づく重み係数であり、各漸化式によって求められるコスト関数を単調減少数列にするために、適時 0 ~ 1 の値をとる。

【0041】

【数 5】

$$\partial F = \partial F(A'_m, f'_m, \phi'_m) \quad \dots (5)$$

【0042】

50

周波数パラメータ f'_m , ϕ'_m を求めることができれば、上式(2)における正弦波モデル信号の係数としての周波数パラメータ A'_m を一意に求めることができるため、本周波数解析手法においては、次式(6)によって周波数パラメータ A'_m を収束させる。

【0043】

【数6】

$$A'_{m+1} \leftarrow A'_m - \mu_m \frac{\partial F}{\partial A} \quad \dots (6)$$

【0044】

本周波数解析手法においては、これら一連の計算を反復して行うことにより、振幅 A' 、周波数 f' 、及び初期位相 ϕ' を高精度に収束させることができる。特に、本周波数解析手法においては、上式(2)における正弦波モデル信号の位相を構成する周波数パラメータ f' , ϕ' と、係数としての周波数パラメータ A' とを別個に求めることにより、計算を簡便に行うことができる。

10

【0045】

しかしながら、最急降下法は、比較的広い範囲から収束するものの、1回の反復では精度が低く、収束するまでに時間を要する。

【0046】

そこで、本周波数解析手法においては、最急降下法を適用して周波数パラメータ f'_m , ϕ'_m をある程度まで収束させた後、さらに、いわゆるニュートン法を適用して高精度に収束させるのが望ましい。具体的には、本周波数解析手法においては、ニュートン法として、次式(7)及び次式(8)に示す漸化式によって周波数パラメータ f'_m , ϕ'_m を求める。

20

【0047】

【数7】

$$f'_{m+1} = f'_m - \frac{\nu_m}{J} \begin{vmatrix} \frac{\partial F}{\partial f} & \frac{\partial^2 F}{\partial f^2} \\ \frac{\partial f \partial \phi}{\partial \phi} & \frac{\partial^2 F}{\partial \phi^2} \end{vmatrix} \quad \dots (7)$$

30

【数8】

$$\phi'_{m+1} = \phi'_m - \frac{\nu_m}{J} \begin{vmatrix} \frac{\partial^2 F}{\partial f^2} & \frac{\partial F}{\partial f} \\ \frac{\partial^2 F}{\partial f \partial \phi} & \frac{\partial F}{\partial \phi} \end{vmatrix} \quad \dots (8)$$

【0048】

ただし、上式(7)及び上式(8)において、 J は次式(9)とし、次式(10)と略している。また、 ν_m も μ_m と同様に減速法に基づく重み係数であり、適時0~1の値をとる。

40

【0049】

【数9】

$$J = \begin{vmatrix} \frac{\partial^2 F}{\partial f^2} & \frac{\partial^2 F}{\partial f \partial \phi} \\ \frac{\partial^2 F}{\partial f \partial \phi} & \frac{\partial^2 F}{\partial \phi^2} \end{vmatrix} \quad \dots (9)$$

【数10】

$$\partial^2 F = \partial^2 F(A'_m, f'_m, \phi'_m) \quad \dots (10)$$

50

【 0 0 5 0 】

本周波数解析手法においては、上式(7)及び上式(8)によって周波数パラメータ f_m' , ϕ_m' を求めた後、最急降下法と同様に、上式(6)によって周波数パラメータ A_m' を収束させ、この一連の計算をさらに反復して行う。

【 0 0 5 1 】

このように、本周波数解析手法においては、最急降下法とニュートン法とを組み合わせたハイブリッド型の解法を用いることにより、高速に且つ高精度に周波数パラメータ A' , f' , ϕ' を推定することができる。

【 0 0 5 2 】

また、本周波数解析手法においては、解析対象信号 $x(n)$ が複合正弦波の場合であっても、逐次減算処理することにより、近似的にスペクトルパラメータを導出することができる。ここで、解析対象信号 $x(n)$ が複数の正弦波の和であり、次式(11)のように表されているとする。

【 0 0 5 3 】

【数11】

$$x(n) = \sum_{k=1}^N A_k \cos(2\pi f_k \frac{n}{f_s} + \phi_k) \quad \dots (11)$$

【 0 0 5 4 】

パーセヴァル(Parseval)の定理より、解析対象信号 $x(n)$ の周波数 f_k と正弦波モデル信号の周波数パラメータ f' とが全く一致しない場合、すなわち、次式(12)である場合には、上式(2)に示す非線形方程式は次式(13)となる。また、周波数パラメータ f' , ϕ' の組が、周波数 f_k 及び初期位相 ϕ_k の組のいずれかに一致する場合には、上式(2)に示す非線形方程式は次式(14)となる。さらに、振幅 A_j が周波数パラメータ A' と一致した場合には、解析対象信号から推定スペクトルに関する周波数成分を完全に消去することができる。そのため、最適解を求める問題は、周波数に対して独立であり、解析対象信号から順次個別に推定すれば、複数の正弦波で表される信号にも応用することができる。

【 0 0 5 5 】

【数12】

$$f_k \neq f' \quad \dots (12)$$

【数13】

$$F(A', f', \phi') = A'^2 + \sum_{k=1}^N A_k^2 \quad \dots (13)$$

【数14】

$$F(A', f', \phi') = (A_j - A')^2 + \sum_{k=1, k \neq j}^N A_k^2 \quad \dots (14)$$

【 0 0 5 6 】

すなわち、本周波数解析手法においては、解析対象信号 $x(n)$ が複合正弦波の場合であっても、逐次残差信号に対して同様に処理を行い、複数の正弦波を抽出することができる。

【 0 0 5 7 】

音声信号や音響信号等の信号を複合正弦波によって表現するためには、これまで多くのスペクトル数(正弦波の数)が必要であったが、本周波数解析手法においては、そのような信号であっても僅かなスペクトル数で誤差なく表現することができる。すなわち、信号をより少ないスペクトル数で表現可能であることは、情報圧縮の用途に有効であることを示している。

10

20

30

40

50

【 0 0 5 8 】

[本周波数解析手法の有効性]

以下、本周波数解析手法の有効性について具体的に説明する。

【 0 0 5 9 】

本周波数解析手法は、非線形方程式の最適解を求めることにより、正弦波モデル信号の周波数 f' 、振幅 A' 、及び初期位相 ϕ' を高速に且つ高精度に求めることができる。具体的な精度を立証するために、本願発明者は、DFTと、DFTの発展型のうち最も解析精度が高いといわれているGHA (Generalized Harmonic Analysis) とを比較対象として精度の検証を行った。

【 0 0 6 0 】

なお、DFTやGHAは、1つの解析窓長に見かけ上複数の窓長を持たせていることから、周波数分解能が解析窓長に依存するが、その分解周波数が有限長であり、解析対象信号の周波数が分解周波数以外の周波数となった場合には解析することができず、解析対象信号が正確に解析できる周波数と異なる場合には、最も近い分解周波数の他に、その周辺に小さなスペクトルの周波数(側帯波成分)が現れ、複数の周波数が出現してしまう。

【 0 0 6 1 】

このような現象が本周波数解析手法においても生じるか否かについて、すなわち、本周波数解析手法の周波数分解能を検証するために、解析窓長を1秒(1024サンプル)とした1次元の非常に短い単一正弦波を解析し、各手法によって正弦波を1本抽出して元の信号との二乗誤差を調べた。その結果を図2に示す。

【 0 0 6 2 】

図2に示すように、DFTにおいては、基本周波数の整数倍以外の周波数における解析精度の悪化がみられた。また、GHAにおいては、1Hz以上の周波数ではDFTと比べて2~5桁程度の精度向上がみられた。これに対して、本周波数解析手法においては、1Hz以上の周波数ではDFTと比べて10桁以上、GHAと比べて5桁以上の精度向上がみられた。すなわち、本周波数解析手法は、既存の周波数解析手法と比べて10万~100億倍以上の精度向上がみられた。特に、1Hz以下の周波数を正確に推定することができるということは、解析窓長を超えた長い周期信号であっても解析可能であることを示している。

【 0 0 6 3 】

このように、本周波数解析手法は、最も解析精度が高いといわれているGHAと比べても驚くべき高精度に解析を行うことができるものである。動き推定装置は、このような周波数解析手法を利用して、図3に示すような一連の処理を行うことにより、入力された動画像内の1つ又は複数の移動物体の動きを推定する。

【 0 0 6 4 】

[動画像からの領域切り出し]

まず、動き推定装置は、図3に示すように、ステップS1において、CPU11の制御のもとに、図示しない動画像入力部を介して入力されてRAM13等のメモリに記憶されているオリジナルの動画像データから移動物体が含まれる領域を切り出す。ここで、切り出した領域の横方向のピクセル数、縦方向のピクセル数、フレーム数を、それぞれ、 N_x 、 N_y 、 N_z とし、各要素を i 、 j 、 k とする。なお、 i 、 j 、 k は、画像を再現するためのインデックスであり、実数である。また、切り出した3次元信号からなる動画像データを、 $o_{b_j}(i, j, k)$ とする。この切り出された動画像データは、RAM13等に記憶される。

【 0 0 6 5 】

[時空間スペクトルの抽出]

続いて、動き推定装置は、ステップS2において、CPU11の制御のもとに、次式(15)に示すように、上述した周波数解析手法を利用して3次元信号からなる動画像データ $o_{b_j}(i, j, k)$ からL本の時空間スペクトルを抽出する。次式(15)において、 f_{x_s} 、 f_{y_s} 、 f_{z_s} は、それぞれ、動画像の横軸方向、縦軸方向、時間軸方向の

10

20

30

40

50

サンプリング周波数 [Hz] であり、 A' 、 f_x' 、 f_y' 、 f_z' 、 ϕ' は、それぞれ、実際に抽出するスペクトルの振幅、各軸に対応する周波数、初期位相である。動き推定装置は、3次元信号である動画データを次式(15)に示すような正弦波モデル関数を用いて表し、実際の信号と、次式(15)に示す正弦波モデル信号との差が最小となるように、次式(16)を用いてパラメータを変化させ、各周波数を求める。

【0066】

【数15】

$$\xi(i, j, k) = A' \cos \left(2\pi \left(\frac{f_x'}{f_{xs}} i + \frac{f_y'}{f_{ys}} j + \frac{f_z'}{f_{zs}} k \right) + \phi' \right) \quad \dots (15)$$

10

【数16】

$$F(A', f_x', f_y', f_z', \phi') = \frac{1}{N_x N_y N_z} \sum_{l=0}^{N_x-1} \sum_{j=0}^{N_y-1} \sum_{k=0}^{N_z-1} \{ \xi_{obj}(i, j, k) - \xi(i, j, k) \}^2 \quad \dots (16)$$

【0067】

そして、動画データは、次式(17)に示すように、時空間スペクトルの合成によって表現することができるため、動き推定装置は、CPU11の制御のもとに、動画データからL本の時空間スペクトルを抽出する。このとき、1番目の時空間スペクトルを次式(18)に示すように表す。

20

【0068】

【数17】

$$\xi_{obj}(i, j, k) = \sum_{l=1}^L A(l) \cos \left(2\pi \left(\frac{f_x(l)}{f_{xs}} i + \frac{f_y(l)}{f_{ys}} j + \frac{f_z(l)}{f_{zs}} k \right) + \phi(l) \right) \quad \dots (17)$$

【数18】

$$S_l(f_x(l), f_y(l), f_z(l)) \quad l=1, \dots, L \quad \dots (18)$$

【0069】

[時空間スペクトルのクラスタリング]

そして、動き推定装置は、ステップS3において、CPU11の制御のもとに、時空間スペクトルのクラスタリングを行う。ここでは、I個の平面が存在すると仮定し、1番目の時空間スペクトルがi番目の平面に属するものとした場合の存在確率変数を $P_i(1)$ とする。なお、ここでは、確率変数Pは、0又は1の値をとるものとし、当初は0で初期化されているものとする。

30

【0070】

まず、動き推定装置は、CPU11の制御のもとに、時空間スペクトル S_1 から2本の時空間スペクトル S_u 、 S_v を選択する。なお、u、vは、次式(19)に示すように集合Lに含まれ、且つ、互いに異なる値である。

40

【0071】

【数19】

$$u, v \in L, \quad u \neq v \quad \dots (19)$$

【0072】

続いて、動き推定装置は、CPU11の制御のもとに、次式(20)に示すように、選択した時空間スペクトル S_u 、 S_v と原点とによって構成される平面の法線ベクトル n_u 、 n_v を求める。

【0073】

【数 2 0】

$$n_{u,v} = \begin{bmatrix} \alpha_{u,v} \\ \beta_{u,v} \\ \gamma_{u,v} \end{bmatrix} = \begin{bmatrix} f_x(u) \\ f_y(u) \\ f_z(u) \end{bmatrix} \times \begin{bmatrix} f_x(v) \\ f_y(v) \\ f_z(v) \end{bmatrix} \quad \dots (20)$$

【0074】

続いて、動き推定装置は、CPU 11の制御のもとに、次式(21)に示すように、求めた法線ベクトル $n_{u,v}$ を用いて平面の方程式を仮定する

【0075】

【数 2 1】

$$\alpha_{u,v} f_x + \beta_{u,v} f_y + \gamma_{u,v} f_z = 0 \quad \dots (21)$$

10

【0076】

続いて、動き推定装置は、CPU 11の制御のもとに、次式(22)に示すように、上式(21)によって求めた平面と、時空間スペクトル S_u, S_v を除く時空間スペクトル S_m との距離 $D_{u,v}(m)$ をそれぞれ求める。なお、 m は、集合 L に含まれ、且つ、 u, v とは異なる値である。

【0077】

【数 2 2】

$$D_{u,v}(m) = \frac{|\alpha_{u,v} f_x(m) + \beta_{u,v} f_y(m) + \gamma_{u,v} f_z(m)|}{\sqrt{\alpha_{u,v}^2 + \beta_{u,v}^2 + \gamma_{u,v}^2}}$$

20

$$m \in L, \quad m \neq u \neq v \quad \dots (22)$$

【0078】

そして、動き推定装置は、CPU 11の制御のもとに、次式(23)に示すように、求めた距離 $D_{u,v}(m)$ が所定の閾値以下となる時空間スペクトルを探索し、その確率変数 $P_i(m)$ を 1 とする。

【0079】

【数 2 3】

$$D_{u,v}(m) \leq \varepsilon, \quad P_i(m) = 1 \quad \dots (23)$$

30

【0080】

なお、ここでは、確率変数を 1 とすることにより、平面に完全に属するものとして演算を行っているが、動き推定装置においては、小数の確率変数を設定し、平面に近い点とみなす閾値を設け、この閾値を用いた演算を行うようにしてもよく、この閾値を次回の演算の際の算出パラメータの 1 つとしてもよい。

【0081】

そして、動き推定装置は、CPU 11の制御のもとに、上式(23)を満たす時空間スペクトル S_m に対して最小二乗法を適用し、次式(24)に示す平面の方程式を計算し、推定した平面の傾きから移動物体の横方向及び縦方向の速度 V_x, V_y を推定する。なお、次式(4)における d は、 f_z 軸の切片である。

40

【0082】

【数 2 4】

$$V_x \frac{N_z}{N_x} f_x + V_y \frac{N_z}{N_y} f_y + f_z - d = 0 \quad \dots (24)$$

【0083】

[時空間スペクトルからの平面推定]

50

動き推定装置は、ステップS4において、CPU11の制御のもとに、以上のようにして時空間スペクトルから平面を推定し、平面iに属していない時空間スペクトルが存在する場合には、上式(23)を満たしていない時空間スペクトルを用いて、ステップS3における時空間スペクトルのクラスタリングを繰り返す。

【0084】

動き推定装置は、このような一連の処理を行うことにより、入力された動画像内の1つ又は複数の移動物体の動きを推定することができる。特に、動き推定装置は、時空間スペクトルの抽出に上述した本周波数解析手法を利用することにより、理論平面からの誤差が少ない平面の推定を行うことができる。そのため、動き推定装置は、直線群フィルタやファジー推論を用いたクラスタリング法ではなく、平面推定の簡易的な手法である最小二乗法を用いたクラスタリング法を適用することが可能となり、高精度に且つ少ない演算量で動き推定を行うことができる。

10

【0085】

[動き推定装置の有効性]

以下、動き推定装置の有効性について具体的に説明する。

[第1の検証]

まず、図4に示すように、横方向のピクセル数、縦方向のピクセル数、フレーム数を、それぞれ、 N_x 、 N_y 、 N_z とし、初期フレームの左端にある黒い縦棒状の物体が、時間の進行にともない右方に移動するような動画像を考える。移動速度 V_x は、連続するフレーム間で物体が移動するピクセル数として定義し、フレームの右方又は上方に物体が移動する場合を正とする。速度がフレームあたりのピクセル数であり、 $N_z > (N_x / V_x)$ である場合には、物体は、 n 番目のフレームまで存在する。縦棒状の物体を用いていることから、縦方向の周波数 f_y はゼロである。したがって、動画像は、周波数 f_x 、 f_z によって特徴付けられることになる。本願発明者は、3次元の平面 $f_x - f_z$ 上のスペクトル構造を求め、振幅及び周波数 f_x 、 f_z の変化を観察した。この場合、時空間スペクトルは、平面 $f_x - f_z$ 上で直線状に分布し、次式(25)に示すように定義される。なお、次式(25)において、 d は、 f_z 軸の切片であり、速度 V_x は、理論移動の一次方程式における周波数 f_x の係数である。この直線は、理論直線と称される。

20

【0086】

【数25】

$$V_x \frac{N_z}{N_x} f_x + f_z - d = 0 \quad \dots (25)$$

30

【0087】

ここで、3次元FFTによる周波数解析を実行した場合、移動物体の速度及び解析窓は、スペクトルの分布が理論直線の形態をとるかどうかに影響する。また、3次元FFTの分解能も解析窓に依存する。 $N_z = N_x / V_x$ が一定の場合には、移動物体は、動画像を構成する全てのフレームにおいて存在し、動画像データは略周期的なものとなる。また、分解能及びフレーム数が整数であることから、速度もまた整数となる。したがって、この場合には、3次元FFTにおいて高エネルギーを有するスペクトルは理論直線上にあることになる。しかしながら、 $N_z < N_x / V_x$ 又は $N_z > N_x / V_x$ である場合には、解析データは、横方向 x 及び時間軸方向 z について値がゼロであるデータと等しくなる。このような場合には、解析窓内でデータが周期的でないことから、スペクトルは、理論直線上の複数個所で隆起したものとなる。

40

【0088】

実際に、図4に示した動画像として、サイズを $(N_x \times N_y \times N_z) = (64 \times 64 \times 64)$ とし、物体の大きさを64ピクセル×8ピクセルとした場合において、物体の移動速度 V_x を1.0ピクセル/フレームとした場合と、3.3ピクセル/フレームとした場合とについて時空間スペクトルの分布を求めた。なお、スペクトルの振幅は、全てのスペクトルが0dBよりも大きくなるように最大値を40dBに設定した。また、初期位相を

50

1 に、周波数 f_x , f_y , f_z をそれぞれ 64 Hz に正規化した。

【0089】

図5に、物体の移動速度 V_x を 1.0 ピクセル/フレームとした場合における3次元FFTの平面 $f_x - f_z$ 上の時空間スペクトルを示す。図5における破線は、理論直線であり、この直線を正確に求めることができれば正確な移動速度を得ることができる。図5においては、3次元FFTの分解能が1であり、物体の移動速度に一致していることから、高エネルギーを有するスペクトルが直線を形成している。しかしながら、物体の移動速度 V_x が、3次元FFTの分解能に等しくない3.3ピクセル/フレームである場合には、図6に示すように、スペクトルが理論直線上のみに存在するのではなく、その周囲にも存在するものとなる。したがって、正確な速度を得るために、理論直線上にないスペクトルのエネルギーを考慮することが重要である。

10

【0090】

これに対して、物体の移動速度 V_x を 3.3 ピクセル/フレームとした場合において、上述した本周波数解析手法を用いて求めた時空間スペクトルは、図7に示すように、理論直線上に存在し、その周囲には存在しないものとなった。なお、図示しないが、物体の移動速度 V_x が 1.0 ピクセル/フレームである場合にも、同様に、時空間スペクトルが理論直線上に存在する分布が得られた。

【0091】

これらの結果を定量的に評価するために、本願発明者は、3次元FFT及び本周波数解析手法のそれぞれについての時空間スペクトルと理論直線との分散を求めた。分散²は、次式(26)によって求めた。なお、次式(26)において、 f_x' , f_z' は、それぞれ、3次元FFT又は本周波数解析手法のスペクトルパラメータであり、 L はスペクトルの本数である。

20

【0092】

【数26】

$$\sigma^2 = \frac{1}{L} \sum_{l=1}^L \left\{ f_{zl}' - \left(V_x \frac{N_z}{N_x} f_{xl}' + d \right) \right\}^2 \quad \dots (26)$$

【0093】

この分散²が小さいほど、理論直線に対する誤差が小さくなる。3次元FFTの場合には、分散²は560.0となり、本周波数解析手法の場合には、0.0268と極めて小さい値となった。本願発明者は、振幅が約40dBから60dBのスペクトルについて、3次元FFTの場合には分散²が約570.0に収束する一方で、本周波数解析手法の場合には分散²が約0.03に収束することを確認している。

30

【0094】

したがって、本周波数解析手法によるスペクトルは、ほぼ理論直線上に存在することがわかる。移動物体の周期が解析窓の長さと同じ場合には、3次元FFTによるスペクトルは直線を形成するが、逆に、そうでない場合には、スペクトルは直線の形態をとらない。本周波数解析手法を適用した動き推定装置は、解析窓の大きさにかかわらず常に正確なスペクトルを得ることができ、高精度に動き推定を行うことができることがわかる。

40

【0095】

[第2の検証]

第1の検証においては、物体が1方向にのみ移動する場合を示したが、2方向に移動する場合の例として、図8に示すように、初期フレームの左下にある黒い正方形の物体が、時間の進行にともない縦横方向に均一な速度で右上に移動するような動画像を考えた。このような物体の移動は、図4に示したような物体の移動において、フレームの下方から上方に移動する横棒状物体の重ね合わせによって表現される。横方向の移動速度を V_x とし、縦方向の移動速度を V_y とした場合、3次元FFTにおいて高エネルギーを有する時空間スペクトルは、3次元周波数領域空間における平面上の共役スペクトルを除いてはこ

50

の平面上にあることになる。この平面は、次式(27)のような理論平面として定義される。したがって、2つの速度 V_x, V_y は、理論平面方程式における周波数 f_x, f_z の係数を計算することによって求められる。

【0096】

【数27】

$$V_x \frac{N_z}{N_x} f_x + V_y \frac{N_z}{N_y} f_y + f_z - d = 0 \quad \dots (27)$$

【0097】

ここで、 N_z が $N_x / V_x, N_y / V_y$ の最小公倍数と等しい場合には、動画データは略周期的なものとなり、3次元FFTによるスペクトルの大部分は、理論平面の近くに存在することになる。しかしながら、 N_z が $N_x / V_x, N_y / V_y$ の最小公倍数と等しくない場合には、時空間スペクトルは、理論平面に限らずどこにでも存在すると考えられる。したがって、3次元FFTによる時空間スペクトルは、解析窓の影響を受けると推測される。本願発明者は、上述した本周波数解析手法を用い、本周波数解析手法の分解能をテストするために速度を整数に設定することにより、この問題の解決を試みた。

10

【0098】

実際に、図8に示した動画データとして、サイズを $(N_x \times N_y \times N_z) = (64 \times 64 \times 64)$ とし、物体の大きさを8ピクセル×8ピクセルとした場合において、物体の移動速度を $V_x = V_y = 3.3$ ピクセル/フレームとした場合について時空間スペクトルの分布を求めた。なお、3次元周波数領域におけるスペクトル分布の観察を明確にするために、スペクトルの振幅は、全てのスペクトルが0dBよりも大きくなるように最大値を20dBに設定した。

20

【0099】

図9及び図10に、それぞれ、3次元周波数領域において3次元FFT及び本周波数解析手法を用いて求めた時空間スペクトルを示す。図9及び図10において、3つの破線によって形成される平面が理論平面であり、上式(27)におけるパラメータ V_x, V_y 、すなわち、物体の移動速度は3.3である。これら図9及び図10は、求めた時空間スペクトルと理論平面との関係を示している。

30

【0100】

3次元FFTによる時空間スペクトルの一部は、理論平面上に存在するが、大部分のスペクトルは、理論平面の近くに存在している。一方、本周波数解析手法による時空間スペクトルは、その大部分が理論平面上に存在している。分散²は、次式(28)によって求めることができる。

【0101】

【数28】

$$\sigma^2 = \frac{1}{L} \sum_{l=1}^L \left\{ f_{zl}' - \left(-V_x \frac{N_z}{N_x} f_{xl}' - V_y \frac{N_z}{N_y} f_{yl}' + d \right) \right\}^2 \quad \dots (28)$$

40

【0102】

3次元FFTの場合には、分散²は11.66となり、本周波数解析手法の場合には、0.01と極めて小さい値となった。また、特に図示しないが、物体の大きさが32ピクセル×32ピクセルと大きくなった場合には、分散²は、対数的に増加する結果が得られた。

【0103】

このように、本周波数解析手法を適用した動き推定装置は、2方向の速度を有する物体の移動についての正確な時空間スペクトルを得るためにも有益であることがわかる。

【0104】

[第3の検証]

50

つぎに、同時に移動する2つの物体を含む動画像の例として、図11に示すように、初期フレームの左下及び右下にある黒い正方形の2つの物体a, bが、それぞれ、時間の進行にともない、分解能に一致していない異なる速度で右上及び左上に移動するような動画像を考えた。なお、動画像のサイズは、 $(N_x \times N_y \times N_z) = (64 \times 64 \times 64)$ であり、物体の大きさは、8ピクセル×8ピクセルである。

【0105】

図12及び図13に、それぞれ、3次元周波数領域において3次元FFT及び本周波数解析手法を用いて求めた時空間スペクトルを示す。

【0106】

図12から、3次元FFTによる時空間スペクトルは、分解能が不十分であり、スペクトルを正確に抽出することができないため、分解能に一致していない速度で移動している物体を含む動画像の場合には、スペクトルのエネルギーが理想とする場所の周囲に漏れてしまい、平面が積層して厚みを有する層のようなスペクトル分布となることがわかる。このような層状のスペクトル分布が得られた場合には、これらのスペクトルを分離して平面を推定することが困難となる。すなわち、DFTやFFTに代表される等間隔の周波数分解能を有する解析手法においては、その等間隔の周波数に一致する周波数成分であれば正確なスペクトルを得ることができるものの、そうでない場合には正確なスペクトルを得ることができず、理論平面の周囲の周波数に群れたスペクトルが出現し、層状のスペクトル分布となってしまう。

【0107】

一方、本周波数解析手法による時空間スペクトルは、図13に示すように、2つの物体に対応する2つの平面状に分布したものとなり、ばらつきもほとんどみられない。このような2つの平面は、上述したように、最小二乗法を用いて容易に推定することができる。図14に、2つの平面がよく観察できるように視点を変えて平面を描画した様子を示している。この図から、2つの平面が高精度に推定されていることがわかる。動き推定装置においては、推定した平面の f_x 軸及び f_y 軸の傾きをそれぞれ求めることにより、物体の横方向及び縦方向の速度を推定することができる。

【0108】

なお、上述した3次元FFTによる時空間スペクトルにおける層状のスペクトル分布は、特に類似する複数の動きを含む動画像の場合に顕著にみられる。本願発明者は、このような類似する複数の動きを含む動画像における本周波数解析手法の有効性を確認する実験も行っている。

【0109】

動画像のサイズを $(N_x \times N_y \times N_z) = (64 \times 64 \times 64)$ とし、物体の大きさを8ピクセル×8ピクセルとした場合において、先に図8に示したように、初期フレームにおいて左下にある2つの黒い正方形の物体が、時間の進行にともない縦横方向に均一な速度で右上に移動するような動画像を対象とした。2つの物体をa, bとすると、物体aの移動速度は、 $V_x = 1.0$ ピクセル/フレーム、 $V_y = 1.7$ ピクセル/フレームであり、物体bの移動速度は、 $V_x = 1.0$ ピクセル/フレーム、 $V_y = 2.3$ ピクセル/フレームである。すなわち、これら2つの物体a, bは、同じ場所を始点とし、時間の進行とともに、右上方向に移動しながら徐々に高さ方向に離間していく動きをするものである。なお、スペクトルの振幅は、全てのスペクトルが0dBよりも大きくなるように最大値を28dBに設定した。

【0110】

3次元FFT及び本周波数解析手法の双方とも200本のスペクトルを用いて分布を求めたところ、それぞれ、図15及び図16に示すような結果が得られた。そして、この結果から推定した平面の傾き、すなわち、物体a, bのそれぞれの移動速度を求めると、3次元FFTの場合には、

物体aについて、 $V_x = 1.0$ 、 $V_y = 1.8545$

物体bについて、 $V_x = 1.0$ 、 $V_y = 1.5484$

が得られたのに対して、本周波数解析手法の場合には、
 物体 a について、 $V_x = 0.99994$ 、 $V_y = 1.7211$
 物体 b について、 $V_x = 0.98712$ 、 $V_y = 2.3108$
 が得られた。すなわち、3次元FFTを用いた場合には、縦方向の移動速度 V_y の精度が悪く、本周波数解析手法を用いた場合には極めて高精度に動きを推定していることがわかる。

【0111】

同様に、物体 a の移動速度を、 $V_x = 1.0$ ピクセル/フレーム、 $V_y = 2.3$ ピクセル/フレームに変化させ、物体 b の移動速度を、 $V_x = 1.7$ ピクセル/フレーム、 $V_y = 3.2$ ピクセル/フレームに変化させて同様にスペクトル分布を求めた。ここでも、スペクトルの振幅は、全てのスペクトルが0dBよりも大きくなるように最大値を28dBに設定した。

【0112】

3次元FFTの場合には300本のスペクトル、本周波数解析手法の場合には200本のスペクトルを用いて分布を求めたところ、それぞれ、図17及び図18に示すような結果が得られた。そして、この結果から推定した平面の傾き、すなわち、物体 a , b のそれぞれの移動速度を求めると、3次元FFTの場合には、

物体 a について、 $V_x = 1.0034$ 、 $V_y = 2.1525$

物体 b について、 $V_x = 1.8053$ 、 $V_y = 3.5693$

が得られたのに対して、本周波数解析手法の場合には、

物体 a について、 $V_x = 1.0106$ 、 $V_y = 2.3116$

物体 b について、 $V_x = 1.7502$ 、 $V_y = 3.2304$

が得られ、この場合にも、3次元FFTを用いた場合に比べ、本周波数解析手法を用いた場合の方が良好な結果が得られた。

【0113】

このように、本周波数解析手法を適用した動き推定装置は、類似する複数の動きを含む動画像における動き推定を行う場合にも極めて有効である。

【0114】

[第4の検証]

つぎに、同時に移動する3つの物体を含む動画像の例として、図19に示すように、初期フレームの右上、左下及び右下にある黒い正方形の3つの物体 a , b , c が、それぞれ、時間の進行にともない、異なる速度で左下、右上及び左上に移動するような動画像を考えた。なお、動画像のサイズは、 $(N_x \times N_y \times N_z) = (64 \times 64 \times 64)$ であり、物体の大きさは、8ピクセル×8ピクセルである。また、物体 a の移動速度は、 $V_x = -1.0$ ピクセル/フレーム、 $V_y = -1.0$ ピクセル/フレームであり、物体 b の移動速度は、 $V_x = 1.3$ ピクセル/フレーム、 $V_y = 2.7$ ピクセル/フレームであり、物体 c の移動速度は、 $V_x = -3.3$ ピクセル/フレーム、 $V_y = 2.5$ ピクセル/フレームである。このような動画像について、本周波数解析手法を用いて時空間スペクトルを求め、使用するスペクトルの本数を変えながら平面を推定した。

【0115】

まず、100本のスペクトルを用いた場合には、図20に示すような結果が得られ、推定した3つの平面の傾き、すなわち、物体 a , b , c のそれぞれの移動速度は、

物体 a について、 $V_x = -1.0038$ 、 $V_y = -1.0016$

物体 b について、 $V_x = 1.2353$ 、 $V_y = 2.6933$

が得られたが、物体 c の移動速度については、スペクトルの本数が少なく求めることができなかった。

【0116】

また、125本のスペクトルを用いた場合には、図21に示すような結果が得られ、物体 a , b , c のそれぞれの移動速度は、

物体 a について、 $V_x = -1.0038$ 、 $V_y = -1.0024$

物体 b について、 $V_x = 1.2593$ 、 $V_y = 2.7166$
 物体 c について、 $V_x = -3.2332$ 、 $V_y = 2.452$
 が得られた。

【0117】

さらに、160本のスペクトルを用いた場合には、図22に示すような結果が得られ、
 物体 a, b, c のそれぞれの移動速度は、

物体 a について、 $V_x = -1.0053$ 、 $V_y = -1.0015$
 物体 b について、 $V_x = 1.2948$ 、 $V_y = 2.7012$
 物体 c について、 $V_x = -3.2897$ 、 $V_y = 2.4594$
 が得られた。

10

【0118】

さらにまた、180本のスペクトルを用いた場合には、図23に示すような結果が得られ、
 物体 a, b, c のそれぞれの移動速度は、

物体 a について、 $V_x = -1.003$ 、 $V_y = -1.0013$
 物体 b について、 $V_x = 1.3085$ 、 $V_y = 2.6974$
 物体 c について、 $V_x = -3.2906$ 、 $V_y = 2.4561$
 が得られた。

【0119】

また、200本のスペクトルを用いた場合には、図24に示すような結果が得られ、
 物体 a, b, c のそれぞれの移動速度は、

物体 a について、 $V_x = -1.0024$ 、 $V_y = -1.0008$
 物体 b について、 $V_x = 1.3104$ 、 $V_y = 2.7007$
 物体 c について、 $V_x = -3.2874$ 、 $V_y = 2.4595$
 が得られた。

20

【0120】

さらに、225本のスペクトルを用いた場合には、図25に示すような結果が得られ、
 物体 a, b, c のそれぞれの移動速度は、

物体 a について、 $V_x = -1.0023$ 、 $V_y = -1.0011$
 物体 b について、 $V_x = 1.3093$ 、 $V_y = 2.6994$
 物体 c について、 $V_x = -3.2998$ 、 $V_y = 2.4494$
 が得られた。

30

【0121】

そして、250本のスペクトルを用いた場合には、図26に示すような結果が得られ、
 物体 a, b, c のそれぞれの移動速度は、

物体 a について、 $V_x = -1.0019$ 、 $V_y = -1.001$
 物体 b について、 $V_x = 1.3091$ 、 $V_y = 2.683$
 物体 c について、 $V_x = -3.2756$ 、 $V_y = 2.4747$
 が得られた。

【0122】

すなわち、本周波数解析手法を適用した動き推定装置の精度は、時空間スペクトルの本数に依存し、図19に示した動画像の場合には、約180本程度の時空間スペクトルを用いれば、十分な精度を実現できることがわかる。なお、この十分な精度を実現可能な時空間スペクトルの本数は、移動物体の速度や移動方向等、動画像の内容によって異なる。換言すれば、動き推定装置においては、動画像に応じて適切な時空間スペクトルを用いることにより、演算量を削減しながらも高精度に動き推定を行うことが可能である。

40

【0123】

[第5の検証]

つぎに、移動しない背景がある動画像の前景としての物体が移動する場合の例として、図27に示すように、初期フレームの左下にある黒い正方形の物体が時間の進行にとまらぬ、右上に移動するような動画像を考えた。なお、動画像のサイズは、 $(N_x \times N_y \times$

50

N_z) = (64 × 64 × 64) であり、物体の大きさは、8 ピクセル × 8 ピクセルであり、物体の移動速度は、 $V_x = V_y = 3.3$ ピクセル/フレームである。また、背景は、所定の繰り返しパターンから構成されている。具体的には、背景は、振幅 $A = 128$ (0 ~ 255)、位相 = 0、周波数 $f_x = 8$ 、 $f_y = 0$ 、 $f_z = 1$ のパターンと、振幅 $A = 128$ (0 ~ 255)、位相 = 0、周波数 $f_x = 0$ 、 $f_y = 8$ 、 $f_z = 1$ のパターンとから構成されている。このような動画像について、3次元FFT及び本周波数解析手法のそれぞれを用いて時空間スペクトルを求め、平面を推定した。なお、スペクトルの振幅は、全てのスペクトルが0 dBよりも大きくなるように、3次元FFTの場合には最大値を25.07 dB、本周波数解析手法の場合には22.75 dBに設定した。

【0124】

10

まず、3次元FFTによる101本の時空間スペクトルの分布を求めると、図28に示すように、背景のパターンの影響を受けて断続的な分布が得られた。そして、この結果から、上述した閾値を変化させながら物体の移動速度を求めると、
 閾値 = 0.1の場合、 $V_x = 3.4697$ 、 $V_y = 3.4697$
 閾値 = 0.2の場合、 $V_x = 3.3824$ 、 $V_y = 3.3824$
 閾値 = 0.3 ~ 0.5の場合、 $V_x = 3.2857$ 、 $V_y = 3.2857$
 といったように、閾値にかかわらず誤差が大きい結果が得られた。

【0125】

一方、本周波数解析手法による100本の時空間スペクトルの分布を求めると、図29に示すように、背景のパターンにかかわらず連続的な分布が得られた。そして、この結果から、上述した閾値を変化させながら物体の移動速度を求めると、
 閾値 = 0.1の場合、 $V_x = 3.3093$ 、 $V_y = 3.3033$
 閾値 = 0.2の場合、 $V_x = 3.306$ 、 $V_y = 3.3053$
 閾値 = 0.3 ~ 0.5の場合、 $V_x = 3.3037$ 、 $V_y = 3.3031$
 といったように、誤差が極めて小さい結果が得られた。

20

【0126】

このように、本周波数解析手法を適用した動き推定装置は、背景の影響を受けにくく、高精度に動き推定を行うことが可能である。

【0127】

[第6の検証]

30

つぎに、カラーの動画像において前景が移動する場合の例として、図30に示すように、人物が左方から右方に移動するような動画像を考えた。なお、動画像のサイズは、($N_x \times N_y \times N_z$) = (64 × 64 × 64) であり、人物の移動速度は、 $V_x = 1.641$ ピクセル/フレームである。また、動画像は、所定の温度分布からなる環境において撮像した熱画像から構成されている。このような動画像について、3次元FFT及び本周波数解析手法のそれぞれを用いて時空間スペクトルを求め、平面を推定した。なお、スペクトルの振幅は、全てのスペクトルが0 dBよりも大きくなるように、3次元FFTの場合には最大値を19.4 dB、本周波数解析手法の場合には22.58 dBに設定した。

【0128】

まず、3次元FFTによる101本の時空間スペクトルの分布を求めると、図31に示すような結果が得られ、この結果から、上述した閾値を変化させながら人物の移動速度を求めると、

40

閾値 = 0.4の場合、 $V_x = 1.4838$

閾値 = 0.5の場合、 $V_x = 1.5732$

といったように、閾値にかかわらず誤差が大きい結果が得られた。

【0129】

一方、本周波数解析手法による100本の時空間スペクトルの分布を求めると、図32に示すような結果が得られ、この結果から、上述した閾値を変化させながら人物の移動速度を求めると、

閾値 = 0.4の場合、 $V_x = 1.6452$

50

閾値 = 0.5 の場合、 $V_x = 1.7027$
 といったように、誤差が極めて小さい結果が得られた。

【0130】

このように、本周波数解析手法を適用した動き推定装置は、カラー動画画像の場合にも高精度に動き推定を行うことが可能である。特に、動き推定装置においては、動画画像に応じて適切な閾値を用いることにより、動き推定の精度を向上させることができる。

【0131】

[第7の検証]

つぎに、自然動画画像に適用した場合の例として、図33に示すように、いわゆる「Coast

Guard」という標準動画画像を用いた。この動画画像は、前景であるボートはほとんど静止した状態であるが、カメラワークによって背景が左方に移動するように観察されるものである。この実験では、130フレーム目から161フレーム目までの32フレーム間におけるボートの移動量を推定した。

【0132】

まず、3次元FFTによる時空間スペクトルの分布を求めると、図34に示すように、周波数 f_x 、 f_y の低周波領域に誤差と思われるばらつきが顕著にみられた。一方、本周波数解析手法による100本の時空間スペクトルの分布を求めると、図35に示すように、ばらつきがほとんどみられない結果が得られた。そして、これらの結果から、最小二乗法を用いたクラスタリング法によって平面を推定すると、3次元FFT及び本周波数解析

手法の場合についてそれぞれ図36及び図37に示す結果が得られた。求めた平面の傾きに基

づいて速度を求めると、次表1に示す結果が得られた。

【0133】

【表1】

	オリジナル	3次元FFT	本周波数解析手法
背景(V_x, V_y)	(-0.8750, 0.0313)	(-0.6031, 0.0157)	(-0.8897, 0.0330)
ボート(V_x, V_y)	(-0.0156, 0.0469)	-	(-0.0199, 0.0586)
誤差	-	(-2.7306, 0.0463)	-

【0134】

3次元FFTの場合には、背景については一応推定ができていたが、その推定速度結果に誤差があり、また、ボートについては推定することができず、誤った速度を推定していることがわかる。一方、本周波数解析手法の場合には、誤差が小さいことが確認された。

【0135】

このように、本周波数解析手法を適用した動き推定装置は、自然動画画像の場合にも高精度に動き推定を行うことが可能である。

【0136】

[動き推定装置の効果]

以上説明したように、本周波数解析手法を適用した動き推定装置は、動画画像符号化において動き補償や動き予測、補間といった処理を行う場合に、演算量を大幅に削減し且つ高精度に処理を行うことができる、したがって、この動き推定装置は、動画画像内に含まれる物体の動きを推定・予測したり、フレーム補間等を行ったりする際に利用して好適である。

【0137】

また、この動き推定装置は、動画画像符号化に応用することにより、従来に比べて大幅な高圧縮化を図ることができる。さらに、この動き推定装置は、大規模な変更を加えることなく、処理の質的・量的変化に対応することができる、スケーラブルな符号化技術の開発に利用可能である。これにより、この動き推定装置は、昨今増加傾向にあるネットワークトラフィックの低減に貢献することができる。

【0138】

なお、本発明は、上述した実施の形態に限定されるものではない。例えば、上述した実施の形態では、動き推定装置によってソフトウェアによる周波数解析を行うものとして説明したが、本発明は、本周波数解析手法を含む動き推定処理のアルゴリズムを実装したDSP (Digital Signal Processor) 等、積和演算を行うことが可能であればハードウェアによっても実現することができる。

【0139】

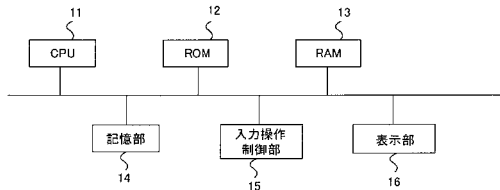
このように、本発明は、その趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

【符号の説明】

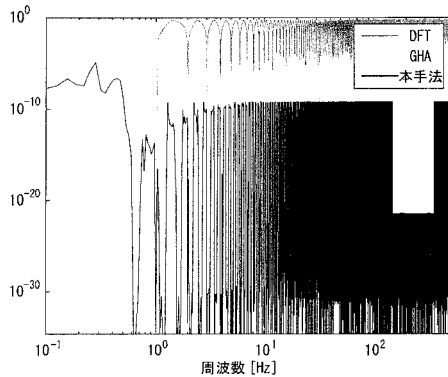
【0140】

- 11 CPU
- 12 ROM
- 13 RAM
- 14 記憶部
- 15 入力操作制御部
- 16 表示部

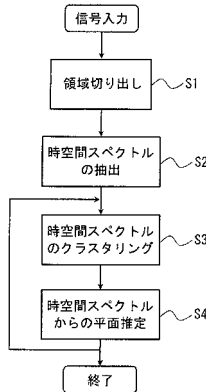
【図1】



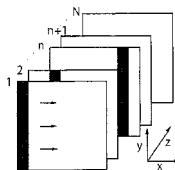
【図2】



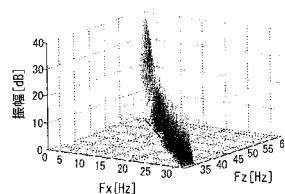
【図3】



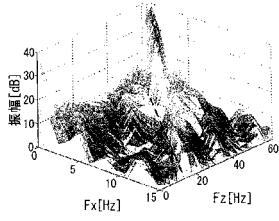
【図4】



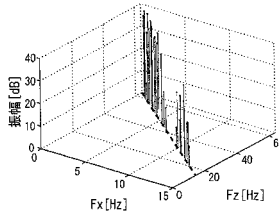
【図5】



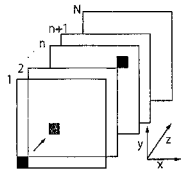
【 図 6 】



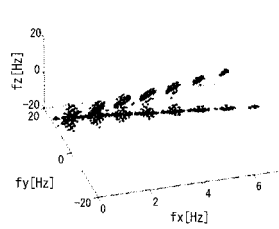
【 図 7 】



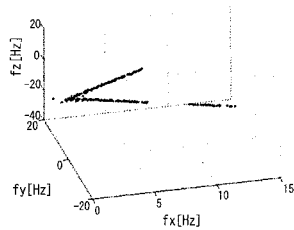
【 図 8 】



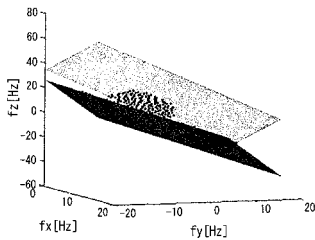
【 図 12 】



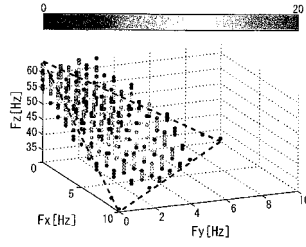
【 図 13 】



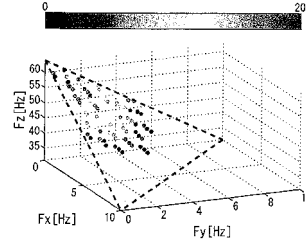
【 図 14 】



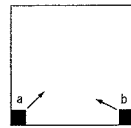
【 図 9 】



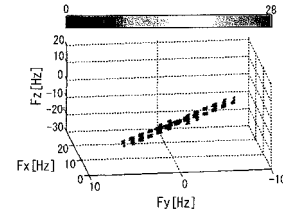
【 図 10 】



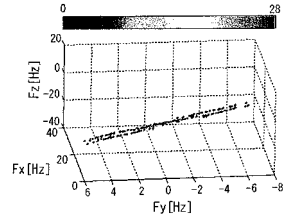
【 図 11 】



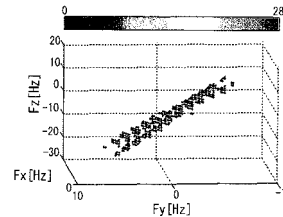
【 図 15 】



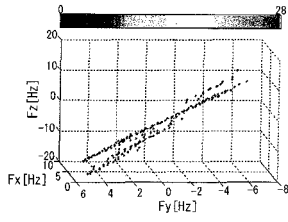
【 図 16 】



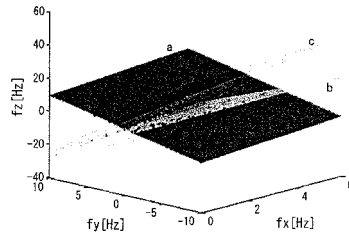
【 図 17 】



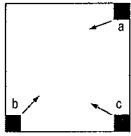
【 18 】



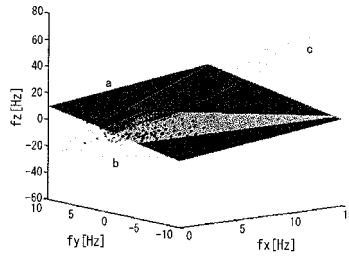
【 21 】



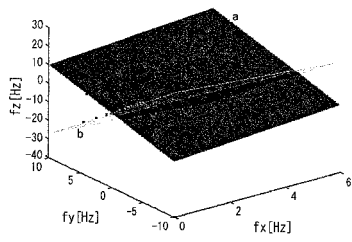
【 19 】



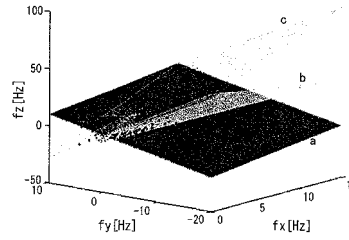
【 22 】



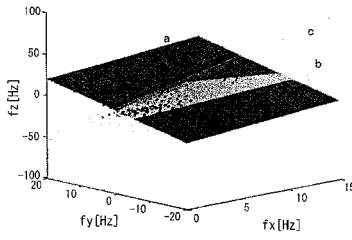
【 20 】



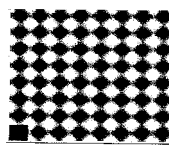
【 23 】



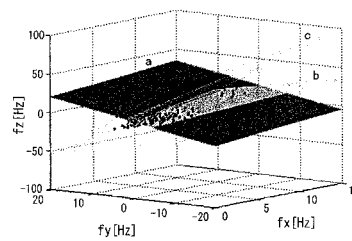
【 24 】



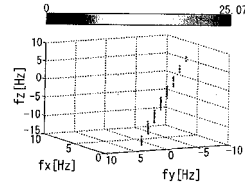
【 27 】



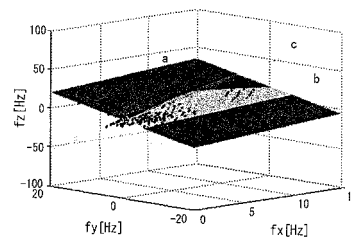
【 25 】



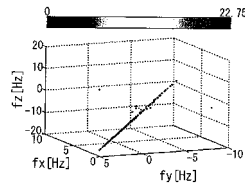
【 28 】



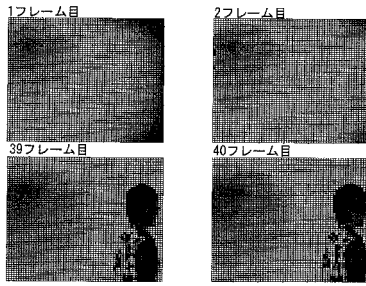
【 26 】



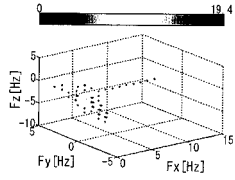
【 29 】



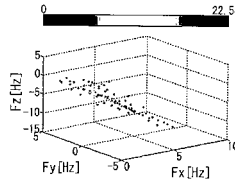
【図 30】



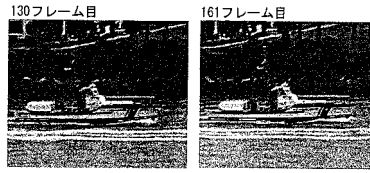
【図 31】



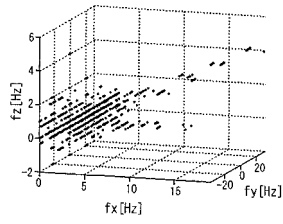
【図 32】



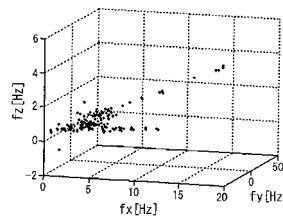
【図 33】



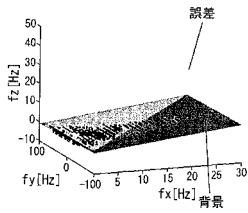
【図 34】



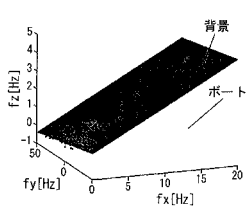
【図 35】



【図 36】



【図 37】



フロントページの続き

特許法第30条第1項適用 平成21年10月17日~19日 The Institute of Electrical and Electronic Engineers, Inc. (IEEE) 発行の「2009 2nd International Congress on Image and Signal Processing Volume 4」(2009第2回国際会議画像信号処理研究会抄録誌 第4巻)に発表。

- (56)参考文献 船田純一 外1名, 自然動画像の中に潜む非フーリエ運動, 電子情報通信学会技術研究報告 NC96-155~208 ニューロコンピューティング, 社団法人電子情報通信学会, 1997年 3月18日, 第96巻 第584号, 第399~406頁
東山三樹夫 外1名, 高い周波数分析精度の信号分析手法, 日本音響学会誌, 社団法人日本音響学会, 1998年 8月 1日, 第54巻 第8号, 第568~574頁
古川徹生, ボルツマンマシンによるモデル選択: 確率的動作と期待値動作, 電子情報通信学会技術研究報告 NC2001-22~36 ニューロコンピューティング, 社団法人電子情報通信学会, 2001年 6月22日, 第101巻 第154号, 第57~64頁

(58)調査した分野(Int.Cl., DB名)

G06T 7/00 ~ 7/60
H04N 7/32