

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2009-509218

(P2009-509218A)

(43) 公表日 平成21年3月5日(2009.3.5)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G06F 17/30 (2006.01)</b>	G06F 17/30 220A	5B075
<b>H04N 7/30 (2006.01)</b>	H04N 7/133 Z	5C052
<b>H04N 5/76 (2006.01)</b>	H04N 5/76 B	5C053
<b>H04N 5/92 (2006.01)</b>	H04N 5/92 H	5C059
	G06F 17/30 170D	
審査請求 未請求 予備審査請求 未請求 (全 76 頁)		

(21) 出願番号 特願2008-528577 (P2008-528577)  
 (86) (22) 出願日 平成18年9月1日 (2006.9.1)  
 (85) 翻訳文提出日 平成20年4月22日 (2008.4.22)  
 (86) 国際出願番号 PCT/GB2006/003243  
 (87) 国際公開番号 W02007/026162  
 (87) 国際公開日 平成19年3月8日 (2007.3.8)  
 (31) 優先権主張番号 60/712,810  
 (32) 優先日 平成17年9月1日 (2005.9.1)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 508059731  
 アストラグループ エイエス (ア ノルウ  
 ェギアン カンパニー)  
 ノルウェー王国 エヌオー・O250 オ  
 スロ、ベッディングゲン8  
 (74) 代理人 100096024  
 弁理士 柏原 三枝子  
 (74) 代理人 100125520  
 弁理士 高橋 剛一  
 (74) 代理人 100155310  
 弁理士 柴田 雅仁  
 (74) 代理人 100156339  
 弁理士 米村 道子

最終頁に続く

(54) 【発明の名称】 ポストレコーディング分析

## (57) 【要約】

## 【課題】

【解決手段】いくつかの形式のコンピュータあるいは計算機を利用してデジタルデータレコーディングを生成する場合、データは様々な方法で入力され、いくつかの形式の電子媒体に格納される。この処理では、データが計算および変換され、格納するためにデータが最適化される。本発明は、データ圧縮、アクティビティ検出およびオブジェクト認識などの多くの異なる処理それぞれに必要なものを含む方法により計算を設計することに関する。到着したデータがこのように計算および格納されると、各処理の情報が同時に抽出される。異なる処理の計算は、単一のプロセッサで連続的に実行され、あるいは複数の分散型プロセッサで並列に実行される。抽出処理を「要約分解」と称し、抽出された情報を「要約データ」と称する。通常、「要約データ」の語は、オリジナルデータの主要部を含まない。要約データは、生成される特定の問い合わせに対する事前のバイアスなしで生成され、このため、レコーディングを生成する前に検索基準を入力する必要がない。また、これは、要約分解をするのに利用されるアルゴリズム/計算の特性にも依存しない。(処理された)要約データを有する(処理された)オリジナルデータを含む得られたデータは、関連するデータベースに格納される。代替的に、簡単な形式の要約データを主要なデータの一部として格納してもよい。レコーディングが実施された後、要約データは、データの主要部を調べる必要なく分析できる。必要な計算の大部分がオリジナルデータのレコーディングのときに既に行われているため、この分析は、非常に高速で実施される。要約データを分析することにより、必要に応じて主要なデータレコーディングから関連するデータ

**【特許請求の範囲】****【請求項 1】**

多数の連続したデジタル化されたデータを問い合わせまたは検索する方法であって、

- (a) ピラミッド分解を利用してデータを分解するステップと、
- (b) 選別処理を実施してデータ属性(要約データ)の情報を分離するステップと、
- (c) 前記データおよび前記要約データをインデックスとともに格納するステップと、
- (d) 問い合わせあるいは検索基準を設定するステップと、
- (e) 要約データを取り出すステップと、
- (f) 前記問い合わせあるいは検索基準を前記取り出された要約データに適用するステップとを利用することを特徴とする方法。

10

**【請求項 2】**

請求項 1 に記載の方法において、前記インデックスが、対応するメインデータを取り出すのに利用されることを特徴とする方法。

**【請求項 3】**

請求項 1 に記載の方法において、前記分解が、ウェーブレットを利用して実施されることを特徴とする方法。

**【請求項 4】**

請求項 2 に記載の方法において、前記分解が、適応ウェーブレット階層を利用して実施されることを特徴とする方法。

**【請求項 5】**

20

請求項 1 に記載の方法において、前記選別処理が、ノイズ属性を抽出するのに利用されることを特徴とする方法。

**【請求項 6】**

請求項 1 に記載の方法において、前記選別処理が、静止した背景の情報を抽出するのに利用されることを特徴とする方法。

**【請求項 7】**

請求項 1 に記載の方法において、前記選別処理が、静的な背景の情報を抽出するのに利用されることを特徴とする方法。

**【請求項 8】**

請求項 1 に記載の方法において、前記選別処理が、動的な動作の情報を抽出するのに利用されることを特徴とする方法。

30

**【請求項 9】**

請求項 1 に記載の方法において、前記選別処理が、オブジェクトの情報を抽出するのに利用されることを特徴とする方法。

**【請求項 10】**

請求項 4 に記載の方法において、要約データが、1 または複数のマスク形式であることを特徴とする方法。

**【請求項 11】**

ピラミッド分解を利用して応用例のためのウェーブレット計算を促進する方法であって、

40

(a) 連続的な変数を利用して偶数ポイントのウェーブレット群をパラメータ化するステップと、

(b) 前記変数を利用してウェーブレット係数のセットを生成するステップとによりウェーブレット計算を促進する方法。

**【請求項 12】**

請求項 11 に記載の方法において、8 ビット表現で正確に表現可能な係数セットが生成されることを特徴とする方法。

**【請求項 13】**

請求項 11 に記載の方法において、前記ウェーブレット係数が、所定のスケールに「調整」されることを特徴とする方法。

50

**【請求項 14】**

ウェーブレットによるピラミッド分解を利用して、デジタル化されたデータシーケンスを処理する方法において、前記シーケンスの各データセットが、適応ウェーブレットを利用してウェーブレット表現に変換されることを特徴とする方法。

**【請求項 15】**

請求項 14 に記載の方法において、適応圧縮が、ノイズと、クリーンデータと、静止したデータと、静的なデータと、動的なデータとを含む異なる要素に前記データを分けることにより、ビデオシーケンスの分析の第 1 段階を行うために、いくつかの処理ノードで構成される繰返しループを用いることを特徴とする方法。

**【請求項 16】**

請求項 14 に記載の方法において、初めの繰返しで、前記シーケンス内の画像のウェーブレット変換と参照画像のウェーブレット変換との違いが、データポイント毎に演算されることを特徴とする方法。

**【請求項 17】**

請求項 14 に記載の方法において、後の繰返しで、ウェーブレットカーネル置換処理を用いて、照度の変化によるフレームの違いを除去することを特徴とする方法。

**【請求項 18】**

請求項 17 に記載の方法において、前記ウェーブレットカーネル置換が、照度の変化を調整するために、現在の画像の低解像度の特徴を前の画像の同じ特徴と置き換えることを特徴とする方法。

**【請求項 19】**

請求項 18 に記載の方法において、現在のテンプレートのカーネル要素が、現在の画像のカーネル要素に置換され、前記現在の画像およびそのウェーブレット変換の新たなバージョンを生成することを特徴とする方法。

**【請求項 20】**

請求項 19 に記載の方法において、前記新たなデータ J は、ノイズを評価し、様々なマスクを演算するために、オリジナル画像 I の代わりに使用できることを特徴とする方法。

**【請求項 21】**

請求項 14 に記載の方法において、フレームの違いの第 1 レベルのウェーブレット変換の本質的な特徴が関連付けられ、体系的なカメラモーションを計算するのに使用されることを特徴とする方法。演算されたシフトが、外挿処理により後続のカメラモーションを予測するために記録されることを特徴とする方法。

**【請求項 22】**

請求項 21 に記載の方法において、デジタルマスクが演算され、現在の画像と前の画像の重複部分と、計算および格納された重複した領域間の変換とを記録することを特徴とする方法。

**【請求項 23】**

請求項 22 に記載の方法において、体系的なカメラモーションの誤差が、カメラのぶれとして処理され、画像の静止した構成要素が、前記カメラのぶれを調整するための背景のテンプレートを形成するのに利用されることを特徴とする方法。

**【請求項 24】**

請求項 14 に記載の方法において、前記マスクは、画像ノイズ分布の統計パラメータをリファインし、且つこれらのパラメータを利用して前記画像をノイズ成分とクリーンな成分とに分けることにより生成されることを特徴とする方法。

**【請求項 25】**

請求項 24 に記載の方法において、所定の閾値に満たない差を有する画像部分が、前の画像に対して画像が変化しない領域を規定するマスクを生成するのに利用されることを特徴とする方法。前記マスクは、シーンについての更なる情報が取得されると、各繰返しにおいて調整されることを特徴とする方法。

**【請求項 26】**

請求項 1 4 に記載の方法において、現在のクリーンな画像が、ピラミッドの各レベルでウェーブレット特性が画像特性に適応した異なるウェーブレットを利用する新規な適応ウェーブレット変換を利用して、ピラミッド分解されることを特徴とする方法。

【請求項 2 7】

請求項 1 4 に記載の方法において、カーネルが変更された現在の画像が、前のテンプレートと比較され、違いがシーン内のモーションとして保存されることを特徴とする方法。

【請求項 2 8】

請求項 2 7 に記載の方法において、前記シーン内のモーションのマスクが生成され、後に参照するために格納されることを特徴とする方法。

【請求項 2 9】

請求項 1 4 に記載の方法において、テンプレートが、等式  $T_j = (1 - \alpha) T_{j-1} + \alpha I_j$  を利用して生成され、静的な背景を平準化し、移動する前面の存在を削除あるいは低減することを特徴とする方法。

【請求項 3 0】

請求項 1 4 に記載の方法において、複数のテンプレートが、複数の値のために保存され、 $\alpha$  は請求項 1 4 で定義されるようなメモリパラメータであることを特徴とする方法。

【請求項 3 1】

請求項 1 4 に記載の方法において、現在の画像およびそのピラミッド表現が、将来のデータと比較可能なテンプレートとして保存されることを特徴とする方法。

【請求項 3 2】

請求項 1 4 に記載の方法において、判断閾値が、背景が動く領域を鈍感にすべく、動的に設定されることを特徴とする方法。

【請求項 3 3】

請求項 3 2 に記載の方法において、動的な判断閾値の使用による背景の感度の喪失は、局所的な動作をぼかすべく、一定期間に亘り統合されたテンプレートを利用することにより補正されることを特徴とする方法。

【請求項 3 4】

請求項 1 4 に記載の方法において、動作が検出された画像部分が、画像領域の過去を示す検出の空間的な相関関係と時間的な相関関係を考慮して、再評価されることを特徴とする方法。

【請求項 3 5】

請求項 1 4 に記載の方法において、動的な背景データが、空間的および時間的に分析されることを特徴とする方法。

【請求項 3 6】

請求項 3 5 に記載の方法において、前記空間的な分析は、前記動的な背景の各要素が、当該要素のうち近接する要素との近接性によってスコアが付けられる相関分析であることを特徴とする方法。

【請求項 3 7】

請求項 3 5 に記載の方法において、前記時間的な分析は、前記動的な背景の要素を、先行するフレームの対応する要素と、先行するフレーム用に既に生成された要約データと比較することにより行われることを特徴とする方法。

【請求項 3 8】

請求項 3 5 に記載の方法において、前記空間的および時間的な相関関係のスコアリングが、事前に割り当てられた空間的および時間的パターンテーブルにより判断されることを特徴とする方法。

【請求項 3 9】

請求項 1 4 に記載の方法において、画像マスクが、データストリームの属性毎に生成され、前記属性が示す画像の位置を表わすことを特徴とする方法。

【請求項 4 0】

請求項 1 4 に記載の方法において、適切に符号化されたウェーブレットデータが、初め

10

20

30

40

50

に、ビットレートを低減する局所的な特徴に依存する適正な閾値および量子化の処理により、次に、効率的な記憶のために得られた係数の符号化により、圧縮されることを特徴とする方法。

【請求項 4 1】

請求項 4 0 に記載の方法において、静的な背景が存在するが静止した背景が存在しないウェーブレット表現内の位置が、マスクにより符号化され、前記位置に特有の閾値および量子化が与えられることを特徴とする方法。

【請求項 4 2】

請求項 1 4 に記載の方法において、画像データ G が、明確な時間制約を有する時間に依存する多数の構成要素の合計として表わされることを特徴とする方法。

10

【請求項 4 3】

請求項 1 4 に記載の方法において、ノイズフィルタが生成され、マスキング技術を利用して適用されることを特徴とする方法。

【請求項 4 4】

請求項 1 4 に記載の方法において、異なるレベルで使用されるウェーブレットが、このパラメータの異なる値を選択することにより、あるレベルから次のレベルに変化することを特徴とする方法。

【請求項 4 5】

請求項 1 4 に記載の方法において、テンプレートが参照画像として利用され、当該参照画面に対して、現在の画像の内容あるいは現在の画像上の変化を評価することを特徴とする方法。

20

【請求項 4 6】

請求項 1 4 に記載の方法において、時間  $I_j$  における画像ストリームの第 1 および第 2 の時間導関数の推定が利用されることを特徴とする方法。

【請求項 4 7】

請求項 1 4 に記載の方法において、レベルがノイズ分布の既知の確率密度に応じて調整可能であり、このため、ピクセルが異常であると誤認する既知の確率が存在し、動く背景が補正できることを特徴とする方法。

【請求項 4 8】

請求項 4 7 に記載の方法において、前記異常なピクセルをトラッキングする場合に、改良された基準が、前記ピクセルの位置あるいは空間的に隣接するピクセルが何をしているかを考慮することなく、各ピクセルにおける変化の時系列の歴史を利用することを特徴とする方法。

30

【請求項 4 9】

請求項 4 8 に記載の方法において、前記ノイズ分布の確率密度が不明な場合に、パラメータを利用せずに決定がなされることを特徴とする方法。

【請求項 5 0】

請求項 1 4 に記載の方法において、パラメータが初期値で設定され、フレームの短いシーケンスを検査した後に自動的に調整できることを特徴とする方法。

【請求項 5 1】

請求項 1 4 に記載の方法において、ブロックスコアリングを用いて、異常な隣接するピクセル自体の数に応じて異常なピクセル毎にスコアを割り当てることにより、異常なピクセルの密集度合を評価すること特徴とする方法。

40

【請求項 5 2】

請求項 1 4 に記載の方法において、シーン内のオブジェクトを特定およびトラックするモーションベクタを利用する方法に基づいたウェーブレットカーネル置換が利用されることを特徴とする方法。

【請求項 5 3】

請求項 1 4 に記載の方法において、速度フィールドが、ウェーブレット変換の  $SS$  成分の適切な対数の全スケールにおける空間的なグラディエントを利用して計算されることを

50

特徴とする方法。

【請求項 5 4】

請求項 1 4 に記載の方法において、適応圧縮が、閾値を決定して当該閾値よりも小さい係数を適切な方法により 0 に設定するステップと、残りの係数を量子化するステップと、これらの係数を効率的に表わすあるいは符号化するステップとで構成されていることを特徴とする方法。

【請求項 5 5】

請求項 1 4 に記載の方法において、適応圧縮が、適応閾値化および適応量子化の方法を利用して画像圧縮のタスクを実行するとともに、フレーム内の特別な関心のある領域の画像の質を維持することを特徴とする方法。

10

【請求項 5 6】

請求項 1 4 に記載の方法において、「ビットボローイング」処理が利用され、特徴に依存する方法により、あるデータポイントの量子化によるエラーを隣接するデータポイントに分散でき、これにより、局所的な領域の情報の内容全体をできる限り保存することを特徴とする方法。

【請求項 5 7】

請求項 1 4 に記載の方法において、フレームグループ内の画像に関連する総ての要約データが、要約画像グループにパッケージされ、これらのグループが、ウェーブレット圧縮されたデータチャンクに正確に対応するチャンクに包含されることを特徴とする方法。

【請求項 5 8】

請求項 1 4 に記載の方法において、圧縮された画像データが保存され、データベースおよび要約データにより参照されることを特徴とする方法。

20

【請求項 5 9】

請求項 1 4 に記載の方法において、時間および日付で検索する場合に、ユーザが、選択されたビデオストリームから、ある瞬間に取得したデータを要求し、特定の時間近くで発生した要約データから生成されたイベントが、前記ユーザに返却されることを特徴とする方法。

【請求項 6 0】

請求項 1 4 に記載の方法において、イベントあるいはオブジェクトを検索する場合に、ユーザが、選択されたビデオストリーム内のシーンの領域と、特定のイベントが発生する検索時間間隔を指定し、前記領域および時間間隔の要約データが検索され、対応するイベントが生成され、前記ユーザに返却されることを特徴とする方法。

30

【請求項 6 1】

請求項 1 4 に記載の方法において、要約データを検索する場合に、当該データは、オンにされたビットを対象にユーザが指定した領域のみが検索されるべきであることを意味する単一のビットプレーンであることを特徴とする方法。

【請求項 6 2】

請求項 1 4 に記載の方法において、成功したクエリーが要約データから生成されている場合に、対応するイベントが生成され、ユーザに返却されるイベントリストに追加されることを特徴とする方法。

40

【請求項 6 3】

請求項 6 2 に記載の方法において、イベントが、キーフレームの前後の複数のデータフレームが当該キーフレームの基準を満たさない場合でも、前記複数のデータフレームを備えることを特徴とする方法。

【請求項 6 4】

請求項 1 4 に記載の方法において、要約データヒットが得られた場合に、何らかの理由によりオブジェクトがサブセットに分類されないとき、これらのストリームに利用可能であり保存された画像から入手可能ないずれの要約データを組み合わせることにより分類が実施されることを特徴とする方法。

50

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

## ポストレコーディング分析

本発明は、データが記録された後に、非常に高速にデジタルデータの分析を実行できる処理に関する。

## 【0002】

## ウェーブレットのパラメータ化

本発明は、連続的なパラメータ化された一群のウェーブレットを生成する処理に関する。多くのウェーブレットは、8ビットあるいは16ビット表現で正確に表現できる。

10

## 【0003】

## ウェーブレットを利用した情報抽出、データ圧縮、およびポストレコード分析

本発明は、周囲の状況の変化に強い情報を抽出すべく適応ウェーブレットを使用する処理と、局所的な適応量子化および閾値化スキームを用いたデータ圧縮を実行する処理と、ポストレコーディング分析の実行する処理に関する。

## 【0004】

大量のデジタルデータが、監視、気象学、地質学、医学、および他の多くの分野で応用するために現在記録されている。

## 【0005】

関連する情報を抽出すべくこのデータを検索することは、単調で時間を浪費する作業である。

20

## 【0006】

レコードを生成する前に特定のマーカーが設定されない限り、データのインタロゲーションは、所望の情報を検索するために、記録された総てのデータを調べる必要がある。

## 【0007】

インタロゲーション処理は自動化できるが、総てのオリジナルデータを分析するため、インタロゲーションが行われる速度が制限される。例えば、デジタルビデオのレコーディングは、レコーディングするのと同じくらい再生に時間がかかるため、これらを分析は、非常に長い処理である。

## 【0008】

危機的な状況が起こり、情報が即座に必要な場合、レコーディングの全体の大きさおよび数により、あり得ない情報を迅速に抽出することができる。

30

## 【0009】

特定のマーカーが事前に設定されている場合、記録されたデータの後続のインタロゲーションが素速く行われるが、これらのマーカーにより決定される情報に限定される。何を探すかについての決定は、レコーディングが開始する前に行うべきであり、レコーディングごとに個別に実施すべき複雑なセットアップ処理が必要であろう。

## 【0010】

本発明の重要な特徴は、インタロゲーションの正確な要求は、レコーディングが完了するまで特定する必要がない。標準的な簡単なデータレコーディングは、データ分析の今後の必要を考慮することなく行うことができる。

40

## 【0011】

したがって、後者の分析が必要な場合、この処理は、インタロゲーションを非常に高速に実施し、短い時間で多くのデータ量を分析できる。

## 【0012】

これは、労力と費用を非常に節約するだけでなく、実際面で従前では不可能であった規模で、非常に多くのデジタル情報を分析できる。

## 【0013】

この処理は、限定ではなく、画像、音および振動データを含む任意の種類のストリームデジタルデータに応用できる。

50

## 【 0 0 1 4 】

この分析は、限定ではなく、データの動的な動作の変化、データの空間的な構造および分配の変化を含む多くの種類により構成してもよい。

## 【 0 0 1 5 】

この分析は、一般的でもよく（例えば、いずれの非反復的な動作あるいはいずれの大きさのオブジェクト）、あるいは細かくてもよい（例えば、特定の出入口あるいは同じような特定のフェースを介した動作）。

## 【 0 0 1 6 】

一般に分析されるデータ種類の例は、

（特定の種類のアクティビティを検出するための）デジタルビデオレコーディング、  
（フェースあるいはナンバープレートなどの特定の種類のオブジェクトを認識するための）デジタルビデオレコーディング、  
（鉱物等の存在を検出するための）振動のレコーディング、  
（骨、考古学的な遺物等の存在を検出するための）振動のレコーディング、  
（キーワード、特定の音、音声パターン等を検出するための）音のレコーディング、  
（カルジオグラムの特有の特徴を検出するための）医学的データのレコーディング、  
（交通流量、消費者の購買傾向等を監視するための）統計データ、  
（気象パターン、海流、温度等进行分析するための）環境データである。

## 【 0 0 1 7 】

ビデオシーケンスを分析する場合、ウェーブレットは、画像を分解するのによく用いられる。この目的のためにウェーブレットを使用することは多くの利点があり、多くの応用例で使用されている。

## 【 0 0 1 8 】

いくつかの応用例に特によく適しているウェーブレットのいくつかのクラスが、定義されている。例えば、DaubechieおよびCoifletウェーブレットである。本発明は、連続型変数を用いて、パラメータ化された方法により、これらのウェーブレットおよび他の総ての偶数ポイント（even-point）ウェーブレットを表わす方法を提供する。本発明は、適切な規模のために自動的に選択でき、データの内容に適応ウェーブレットを簡単に演算する方法を提供する。

## 【 0 0 1 9 】

DaubechieおよびCoifletウェーブレットを含む多くのウェーブレットは、無理数の演算が必要であり、浮動小数点計算を利用して演算しなければならない。本発明は、整数計算を利用して、任意に選択されたウェーブレットに近似するウェーブレットを算出する方法を提供する。整数演算は、正確であり丸めによるエラーなく元に戻すことができ、少ない電力によりマイクロプロセッサで実行でき、浮動小数点計算で必要とされる熱よりも少ない熱を発生させる。これは、多くの状況において有利である。

## 【 0 0 2 0 】

ノイズをフィルタリングし、背景動作と割り込み動作を区別する方法の改善は、要約的なデータの情報内容を最適化するのに有用である。本発明は、背景を決定するための複数のテンプレートの使用、背景の決定における「カーネル置換（kernel substitution）」の使用、ピクセルの違いの重要性を予測するための「ブロックスコアリング」の方法を含むこのような多くの改善を行う方法を提供する。

## 【 0 0 2 1 】

ウェーブレットを利用したビデオ画像の圧縮では、局所的な適応ウェーブレットの使用により、高圧縮から重要な画像の詳細を保護するメカニズムを提供する。特別に関心のありそうな画像内の領域を特定し、ノイズをフィルタリングし背景を決定する様々な方法を使用することにより、高圧縮アルゴリズムの適用からこれらの領域を除外すべく、マスクが作成できる。この方法では、特別に関心のある領域は、画像の残りの部分よりも高いレベルの詳細を維持し、画質を落とすことなく高圧縮方法を利用できる。

## 【 0 0 2 2 】



ウェーブレットの圧縮は、気象データの生成に関する多くの処理のために、普通のコンピュータ環境を提供する。特別に関心のある領域を特定するために生成されるマスクは、気象データとして利用可能な一組のデータを選択的に生成する。

#### 【 0 0 2 3 】

本発明は、画像処理分野における多くの特殊化の結果を利用および統合する。特に、本発明は、多くの新規性のある分析技術に基づいて、画像データの複数のミラミッド分解 (pyramidal decomposition) を利用する。複数のデータ表示を使用することにより、組み合わせたときにデータレベルで何が起きているかについての健全かつ信頼性のある表示を提供する複数の異なるデータビューが可能になる。この情報は、画像データと同時に記録可能な気象データを生成するために組み合わせられた一組の属性マスクとして符号化され、膨大な量のデータの高速なインタロゲーションおよび相関が可能になる。

10

#### 【 背景技術 】

#### 【 0 0 2 4 】

##### 関連技術の説明

本発明は、多数の分野の方法および装置に関し、この分野は、ビデオデータマイニング、ビデオ動画検出および分類、画像セグメンテーション、ウェーブレット画像圧縮である。当業者であれば、これらの分野に関連する従来技術に精通しているであろう。本発明が取り組む主要な課題は、この種の画像処理をリアルタイムで実行する要求であり、この要求は、例えば、テレビおよびビデオレコーディングがHDTVに移行するように、アルゴリズムに大きな制約をかける。

20

#### 【 0 0 2 5 】

シーンの光の変化は、リアルタイムのビデオストリームのセグメント化における問題の主な原因である。このような状況におけるフレーム間の比較は、特に光の変化が急速で一時的な場合に、困難でありモデルに依存する。ここで、これをリアルタイムで処理する場合に、簡単で効果的なモデルに依存しない方法を紹介する。我々が採用する方法は、誤検出の割合が非常に低い状態で処理される画像背景 (揺れる木) 内の動く要素を可能にする。

#### 【 0 0 2 6 】

画像セグメンテーション。今では古典的な論文であるToyama,K.;Krumm,J.;Brumitt,B.;Meyers,B.1999.Wallflower:Principles and Practice of Background Maintenance"。In International Conference on Computer Vision, 255-261およびマイクロソフト社に関連するウェブページ (<http://research.microsoft.com/~jckrumm/WallFlower/TestImages.htm>) は、広い文献の主題である「Wallflower system」の情報源である。(1997, IEEE Trans Patt. Anal. Machine Intel., 19, 394でCaselles等により例示されている) 偏微分方程式に基づくセグメント化方法は興味深い、リアルタイムの応用に対し未だ現実的ではない。他の処理では、Kalman Filtering, Mixture of Gaussian Models and Hidden Markov modelsがある。

30

#### 【 0 0 2 7 】

画像からのノイズのフィルタリング。これは、長い歴史を有する課題である。手軽な一様な閾値からリソースを最大限に消費するエントロピ式の方法に至るまで、ノイズの構成要素を特定する多くの方法が存在する。ウェーブレットの世界は、Donohoおよび協力者の革新的な功績 (例えば、the pioneering D.L. Donoho and I.M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," Biometrika, vol.81, pp.425-455, 199) により支配されており、これらは以下に記されている。また、G.Ramponi, "Detail-preserving filter for noisy image", Electronics Letters, 1995, 31, 865などの初期研究により実証された非線形フィルタに基づく、特徴保存ノイズ除去 (feature-preserving noise removal) のための多くの解決策がある。加重メジアンフィルタおよび他の順序統計量に基づくフィルタは、ほぼ間違いなくJ.W. Tuley's "Nonlinear methods for smoothing data", Conf.Rec.Eascom(174)p673."を思い出す。

40

#### 【 0 0 2 8 】

50

分類および検索。最近の研究のいくつかの意図は、10以上前のプロジェクト、VISION（ネットワークを介した検索のためのビデオインデックス）プロジェクト、DVL S（デジタルビデオライブラリシステム）およびQBIC（画像およびビデオコンテンツによるクエリー）に由来する。例えば、M. Flicker, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, Query by Image and Video Content: The QBIC System, Computer, v.28 n.9, p.23-32, September 1995および"The VISION Digital Video Library and Information Science. Vol.68, Supplement 31, 2000, pp.366-381., 2000を参照。以前からビデオデータの自動検索の分野において開発が盛んに行われている。

#### 【0029】

画像の多重解像度（multi-resolution）表示およびウェーブレット。画像を処理するための階層的な（多重解像度）ウェーブレット変換の使用は、ノイズの除去、特徴検出およびデータ圧縮を含む幅広い話題を網羅する広範な研究を有する。応用例ごとに作られた特定目的のウェーブレットを有するいずれのウェーブレット研究が最もよいのか、またその理由についてに関する疑問について議論がよく交わされている。

#### 【0030】

他の画像処理タスク。安全および監視の分野の狭い領域内でさえ、手ぶれなどの画像収集の態様や、領域マッチング、動作検出および目標追跡などの画像シーケンス処理の態様を網羅するイメージングの応用例を目にする。この技術の多くは、商業製品に組み込まれている。ランダムなカメラモーションおよびトラッキングシステム動作は、多くの研究者により取り組まれている。ここでは、天文学界の適応制御光学（AO）プログラムからいくつかの研究を引用する。試験済の多くの方法の中で、Quad Correlation法が、リアルタイムの状況では非常に簡単で効果的である。Herriot等の(2000)Proc SPIE, 115, 4007がオリジナルの資料である。最近の見解である天文学の画像安定化コンテキストのThomas等の(2006)Mon. Not. R Asrt. Soc. 371, 323を参照。

#### 【発明の開示】

#### 【課題を解決するための手段】

#### 【0031】

何らかの形でコンピュータあるいは計算機を用いてデジタルデータレコーディングを作成する場合、データが様々な方法で入力され、何らかの形で電子媒体に記録される。この処理の間、データに対して計算および変換が実行され、記録のためにデータを最適化する。

#### 【0032】

本発明は、データ圧縮、アクティビティ検出および目標認識などの多くの異なる処理に必要なものが計算に含まれているような方法により、計算を設計することに関する。

#### 【0033】

到着するデータがこれらの計算で管理および蓄積されるため、各処理の情報が同じタイミングで抽出される。

#### 【0034】

異なる処理の計算は、単一のプロセッサで連続的に、あるいは複数の分散型プロセッサで並行して実行される。

#### 【0035】

我々は、抽出処理を「要約分解（synoptic decomposition）」と呼び、抽出された情報を「要約データ」と呼ぶ。「要約データ」の語は通常、オリジナルデータの主要部分を含まない。

#### 【0036】

要約データは、実行されるインタロゲーションを特定する事前のバイアスなしに生成され、レコーディングを作成する前に、検索基準を入力する必要がない。これは、要約的な分解を実行するのに利用されるアルゴリズム／計算の特性に依存しない。

#### 【0037】

10

20

30

40

50

(処理された)要約データと(処理された)オリジナルデータを含む得られたデータは、関連するデータベースに格納される。代替的に、簡単な形式の要約データは、メインデータの一部分として記録してもよい。

【0038】

レコーディングが実行された後、要約データは、データの主要部を調べる必要なく分析できる。

【0039】

この分析は、必要な計算の大部分が元のレコーディングと同時に既に行われているため、非常に速く行うことができる。

【0040】

要約データを分析は、必要に応じて、メインデータのレコーディングから関連するデータにアクセスするのに利用可能なマーカーを提供する。

【0041】

この方法による分析の実質的な効果は、従来の方法による分析では数日あるいは数週間要していた大量の記録されたデジタルデータを、数秒あるいは数分で分析できることである。

【0042】

この分析を行うのに必要なユーザインタフェースの形式には制限がない。

【0043】

ある実施例では、本発明はリアルタイムの画像処理に依存し、この処理により、必要な画像が、ストリームデータセットのサイズ、色、形状、位置、移動のパターン、あるいは他のこのような属性を予測することなく、シーンの中で動く総てのターゲットを確実に特定する方法で分析およびセグメント化される。この特定は、利用可能な資源内で可能な限り、システムあるいはランダムなカメラモーションから独立しており、シーンイルミネーションの変化から独立している。

【発明を実施するための最良の形態】

【0044】

セクション1：ポストレコーディング分析

図1は、通常形式の処理のブロック図である。ブロック1乃至8は、「レコーダ」を備え、ブロック9乃至15は、「アナライザ」を備えている。各ブロックは、新規あるいは周知の小さな処理あるいは一組の処理を示している。一連のデジタル化されたデータは、レコーダに入力され、1以上のピラミッド分解される(ブロック1)。このような分解の例は、ウェーブレット変換であるが、いずれのピラミッド分解を実施してもよい。分解されたデータは、異なる種類の情報の内容を分離する1以上の「ふるい(シーブ)」により「選別される」(ブロック2)。これらの例は、ノイズフィルタあるいは動作検出器である。このふるいは一度行われ、あるいは反復して何度も行ってもよい。選別処理の結果は、適用の目的に依存する3つのカテゴリ、

(a) 通常はノイズであるが、無損失処理あるいは無損失データ圧縮が必要な場合、このカテゴリはNULLでもよい「必要でない」データ(ブロック3)、

(b) (a)以外の総ての情報を含む「メイン」データ(ブロック4)、

(c) 適用の目的に応じて、選択された多数の選別処理の結果で構成される「要約」データ(ブロック5)に分けられる。

【0045】

要約データの重要な特徴は、選別されたデータであり、選別処理は、選別されたデータから一般的な特性を抽出し、このデータの特定の位置における特定の特徴あるいはイベントを単純に特定しない。

【0046】

任意のステップでは、分離されたメインデータが圧縮され(ブロック6)、また、分離された要約データも圧縮してもよい(ブロック7)。選別処理がピラミッド分解の頂点でデータに適用されたとすると、要約データは通常、メインデータの大きさよりもかなり小

10

20

30

40

50

さくなるであろう。

【 0 0 4 7 】

次に、メインデータおよび要約データは、データベースに格納され（ブロック 8）、連続的にインデックスが付される。このインデックスは、メインデータと対応する要約データを関連づける。これにより、処理のレコーディング段階が完了する。

【 0 0 4 8 】

分析段階は、例えば、特定のイベントの発生、特定の特性を有する特定のオブジェクトの存在、あるいはデータシーケンスのテクスチャトレンドの存在などのデータについて、特定のクエリーの形で行われるインタロゲーション処理のセットアップから始まる（ブロック 9）。この処理のユーザインタフェースはいずれの形式でもよいが、クエリーは、要約データのフォーマットおよび範囲に対応すべきである。

【 0 0 4 9 】

関連するデータの連続的なサブセットは、クエリーにより決定され、例えば、クエリーは、インタロゲーションを所定の時間間隔に制限してもよく、対応する要約データは、データベースから抽出され、必要に応じて解凍される（ブロック 10）。次に、抽出された要約データは問い合わせされる（ブロック 11）。インタロゲーション処理は、ブロック 2 で行われた選別処理の完了（completion）を含み、データ内の特定の位置における部分的あるいは一時的な特定の特徴あるいはイベントを特定する決定段階に処理を移行する。特定の情報を抽出するのに必要な詳細は、インタロゲーション段階、即ち、レコーディングが行われた後に提供される（ブロック 9）。インタロゲーションの結果は、クエリー条件が満たされたデータ内の一組の特定の位置である（ブロック 12）。結果は、要約データに含まれる情報の量により制限される。より詳細な結果が必要な場合、特定の位置に対応するメインデータのサブセットは、データベースから抽出すべきであり（ブロック 13）、必要に応じて解凍される。次に、より詳細な選別処理がこれらのサブセットに行われ、詳細なクエリーに応答する（ブロック 14）。

【 0 0 5 0 】

ブロック 13 あるいは 14 のいずれかから得られる対応するデータを表示すべく、好適なグラフィカルユーザインタフェースあるいは他の表示用プログラムを使用できる。これはいずれの形式でもよい。メインデータの解凍が、選別あるいは表示のいずれかを必要とする場合（ブロック 13 あるいは 14）、元のピラミッド分解は可逆である必要がある。

【 0 0 5 1 】

要約データから情報を抽出するのに必要な演算量は、情報を抽出し且つメインデータのサブセットの選別処理をさらに実行するのに必要な演算量よりも少ないが、これらの処置双方は、要約データから供給される情報を含まない記録されたメインデータの選別処理よりも少ない演算を必要とする。

【 0 0 5 2 】

この処理の詳細な実施例は、セクション 3 に示す。

【 0 0 5 3 】

セクション 2：ウェーブレットおよびウェーブレット分解

1 次元のウェーブレット

1 次元のデータセットのウェーブレット変換は、データの伸張の数学的な演算であり、データは、変換により 2 つパートに分割される。1 のパートは、単純にオリジナルデータの半分の大きさに圧縮されたデータである。これが、単に 2 つの要素により拡張された場合、これの元であるオリジナルデータを復元しないだろう、即ち、情報は圧縮処理で失われる。ウェーブレット変換にとってよいのは、圧縮したデータだけでなく、伸張によりオリジナルデータを復元するのに必要な多くのデータを生成することである。

【 0 0 5 4 】

和および差

図 2 を参照する。変換されたデータは、オリジナルと同じ大きさであるが、これは 2 つのパートで構成されており、一方のパートは、圧縮されたデータであり、他方のパートは

10

20

30

40

50

、拡張の際に再追加すべき総ての特徴である。ウェーブレット変換のパートであるこれらの和を  $S$ 、差を  $D$  と称する。

【 0 0 5 5 】

平凡な例

総合的な平凡な例は、2つの数字  $a$  および  $b$  で構成されるデータセットであると考えられる。和は、 $S = (a + b) / 2$  であり、差は、 $D = (a - b) / 2$  である。オリジナルデータは、単純に  $a = S + D$  ,  $b = S - D$  を行うことにより復元される。これは、総てのウェーブレットうち最も初歩的な基礎であるハールウェーブレットである。あらゆるポイントで機能すると同時に、これを実行を実行する様々なウェーブレットが存在する。これらは総て、いくらか異なる特性を有しており、データに対して異なる処理をする。したがって、未解決の問題は常に、いずれの状況下で、これらのうちのいずれを使用するのが最適であるのかということである。

【 0 0 5 6 】

レベル

ウェーブレットのパートの和は、オリジナルデータよりも4倍短いデータを生成すべく、変換されたウェーブレットとすることができる。これは、ウェーブレット変換の第2のレベルとみなされる。したがって、オリジナルデータはレベル0であり、1回目のウェーブレット変換はレベル1である。

圧縮されたデータが単純に1ポイントである限り、継続可能である（実際には、これは、オリジナルデータの長さが2の累乗であることを要する）。

【 0 0 5 7 】

4ポイントウェーブレットフィルタ

$N$ ポイントウェーブレットフィルタは10以上前に有名になっており（I. Daubechies, 1992, Ten Lectures on Wavelets, SIAM, Philadelphia, PAを参照）、ウェーブレット変換の歴史はさらに過去に遡る。このテーマについての多くの研究や多くの取り組みが存在し、これらは総て、多くの書籍および論文に記載されている。

ここで重要なのは、一群のウェーブレットであり、明確にするために、4ポイントフィルタに注目する。結論は、6ポイントおよびそれ以上のポイントである。

【 0 0 5 8 】

4ポイントフィルタ

4ポイントウェーブレットフィルタは、4つの係数を持ち、これらは  $\{ \quad_0, \quad_1, \quad_2, \quad_3 \}$  で示される。ある線上の4等分点における関数の値  $(h_0, h_1, h_2, h_3)$  とすると、我々は、2つの数字  $s_0$  および  $d_0$  を算出できる。

$$s_0 = \alpha_0 h_0 + \alpha_1 h_1 + \alpha_2 h_2 + \alpha_3 h_3$$

$$d_0 = \alpha_3 h_0 - \alpha_2 h_1 + \alpha_1 h_2 - \alpha_0 h_3$$

( [ 0 0 5 8 ] ) . 1

2  $N$  データポイントの線に沿ってフィルタ  $\{ \quad_0, \quad_1, \quad_2, \quad_3 \}$  をシフトする場合、2ポイントのステップでは、 $N$  対の数字  $(s_i, d_i)$  を算出できる。したがって、係数の再配列では、

$$\{h_0, h_1, h_2, \dots, h_{2N}\} \rightarrow \{s_0, \dots, s_N\} \{d_0, \dots, d_N\}$$

( [ 0 0 5 8 ] ) . 2

となる。

重要な要件は、この変換が可逆であることである。これは、以下の条件を課す。

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1$$

$$\alpha_0\alpha_2 + \alpha_1\alpha_3 = 0$$

( [ 0 0 5 8 ] ) . 3

また、

10

$$\alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 = 0$$

$$\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 = \sqrt{2}$$

( [ 0 0 5 8 ] ) . 4

である。変換されたデータが、消失モーメント ( vanishing moment ) などの特定の所望の特性を備えるために、さらに条件を課することができる。

【 0 0 5 9 】

幾何学的解釈

2つの関係 ( [ 0 0 5 8 ] ) . 3 は、これらの4つのウェーブレットを分類でき、正確な整数値を有する関心のある係数のセットを導出できる簡単で精密な幾何学的解釈を可能にする。

図2を参照する。原点Oを有する一組の直角軸 { O x , O y } が示されており、角度が 45 ° の直線OCが描かれている。点CはOからユニットの距離にあり、Cを中心とするユニットの直径の円を描く。円とO x が交わる点Lおよび円とO y が交わる点Mを明記するのが有用である。直線OCは、延在してIで円と交わり、したがって、OIは直径であり、ユニットの長さを有する。

ここで、円上の2点PおよびQを見ると、角POQは直角である。次に、PQは円の対角線である。次に、OPとO y 軸で形成される角を とする。構造的に、 は、OQとO x 軸で形成する時計周回り角である。最後に、PおよびQに座標を割り当て、

$$P = P(\alpha_0, \alpha_3)$$

$$Q = Q(\alpha_2, \alpha_1)$$

( [ 0 0 5 9 ] ) . 1

必要なものが総て揃う。

円がユニットの直径を有し、PQが直径であることから、 $OP^2 + OQ^2 = 1$  であると分かる。割り当てられた点座標において、これは以下について示している。

20

30

40

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1$$

( [ 0 0 5 9 ] ) . 2

ベクタ O P および O Q が直交していることから、

10

$$\alpha_1 \alpha_3 + \alpha_0 \alpha_2 = 0$$

( [ 0 0 5 9 ] ) . 3

が得られ、これらは ( [ 0 0 5 8 ] ) . 3 と全く同じである。また、O L = O M = 1 / 2 であるため、

20

$$\alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 = 0$$

$$\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 = \sqrt{2}$$

( [ 0 0 5 9 ] ) . 4

であると分かり、これは、( [ 0 0 5 8 ] ) . 4 である。

( [ 0 0 5 8 ] ) . 2、( [ 0 0 5 8 ] ) . 3、( [ 0 0 5 8 ] ) . 4 の関係を変更しないで入力置換するのは自由であることに留意すべきである。

30

$$\psi \rightarrow \psi' = -\left(\frac{\pi}{4} + \psi\right)$$

( [ 0 0 5 9 ] ) . 2

【 0 0 6 0 】

4 ポイントウェーブレット群

40

O P と O y 軸により形成される角 は、一群のウェーブレットを決定する。これは、等式 ( [ 0 0 5 8 ] ) . 3 が 4 ポイントウェーブレット係数の必要十分条件であるため、4 ポイントウェーブレットの完全群 ( complete family ) である。一般性を喪失することなく、 $-45^\circ < \quad < 45^\circ$  の範囲の を選択した。より有名なウェーブレット群が表に記されている。

名称	$\psi$
Daubechies 4	$-15^\circ$
Haar	$0^\circ$
Coiflet 4	$15^\circ$

10

Daubechies 4 および Coiflet 4 ウェーブレットの間には、従前では確認されなかったよい対称性が存在する。

角  $\psi$  により、2つのウェーブレット群がどれほど近似するかが示される。

【0061】

代替的なパラメータ化

我々は、2つの数字  $p$  および  $q$  を導入できる。

20

$$\tan \psi = \frac{\alpha_3}{\alpha_0} = -\frac{\alpha_1}{\alpha_2} = -\frac{p}{q}$$

( [ 0 0 6 1 ] . 1 )

30

$$\frac{IP}{OP} = \tan \left( \frac{\pi}{4} - \psi \right) = \frac{p-q}{p+q}$$

( [ 0 0 6 1 ] . 2 )

であるので、我々は

40

$$\alpha_0 = OP \sin \psi = -q \frac{p-q}{p+q},$$

$$\alpha_3 = OP \cos \psi = p \frac{p-q}{p+q}$$

( [ 0 0 6 1 ] . 3 ) を有する。

このため、ウェーブレット係数は、



$$\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\} = \frac{1}{p+q} \{-q(p-q), q(p+q), p(p+q), p(p-q)\}$$

( [ 0 0 6 1 ] . 4 ) である。

10

取得した正確な標準化要因を元に戻す。

$$\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\} = \frac{1}{\sqrt{2}(p^2 + q^2)} \{-q(p-q), q(p+q), p(p+q), p(p-q)\}$$

20

( [ 0 0 6 1 ] . 5 )

p および q が整数の場合、標準化の語句とは別に、我々は整数を有する。

【 0 0 6 2 】

整数近似

3 7 / 4 のとき、d a u b 4 ウェーブレットの周知の表現で表わされる無理数は、  
3 + 9 19 / 4 および 3 - 3 5 / 4 であり、p = 19 および q = 5 のとき、標準  
化されていない整数の近似が導出される。

$$W_{daub4} \approx \{-35, 60, 228, 133\}$$

30

( [ 0 0 6 2 ] . 1 )

これは、実際の値 d a u b 4 = - 1 5 ° に比べて、 = - 1 4 ° . 7 4 4 に対応する。  
標準化されていない係数を用いてこれを近似するのに有用な別の 4 ポイントウェーブレッ  
トが存在する。

$$W_A \approx \{-3, 5, 20, 12\}$$

40

( [ 0 0 6 2 ] . 2 )

これは、 = - 1 4 ° . 0 3 を有する。

別のウェーブレットを提供するために、同一の係数を置換できることに留意すべきである

。

$$W_B \approx \{-3, 12, 20, 5\}$$

( [ 0 0 6 2 ] . 3 )

これは、 $p = 5$  および  $q = 3$  であり、予想とおり、 $\theta = -30^\circ . 96$  である。 $W_A$  および  $W_B$  は、異なる有効帯域幅を有する。

10

最も簡単なウェーブレットは、

$$W_X \approx \{-1, 2, 6, 3\}$$

$$W_Y \approx \{-1, 3, 6, 2\}$$

( [ 0 0 6 2 ] . 3 ) である。

20

$W_X$  は、広範な有効帯域幅を有する 4 ポイントウェーブレットであることが分かっている。

【 0 0 6 3 】

整数近似の稠密集合

無理数を近似するのに、無理数をより近似する集合を形成する有理数の無理数が存在する。このため、所定のウェーブレットを任意に近似する整数の係数を有する標準化されていないウェーブレットが存在する。

【 0 0 6 4 】

6 ポイントウェーブレットおよび高次元

図 3 を参照すると、前述の処理が、6 ポイントおよびより高次元のポイントウェーブレットに一般化されるのかが示されている。図 3 の上側の図は、図 4 を更新したものであり、 $P$  の座標は、新たに  $P(A, B)$  とし、 $OP$  を直径とする新たな円が追加され、長方形  $ORPS$  は、この新たな円の内側に描かれている。このため、三角形  $OSP$  および  $ORP$  は、直角をなし、角  $SOR$  は直角であり、すなわち、 $OS$  および  $OR$  は直交する。図 3 の下側の図は、上側の図の長方形  $ARPS$  および三角形  $QOP$  を抜き出したものであり、これらは総て必要である。

30

ここで、以下の関係が満たされることを証明するのは容易である。

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2 + \alpha_5^2 = 1$$

40

( [ 0 0 6 4 ] . 2 )

$$\alpha_0\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_4 + \alpha_3\alpha_5 = 0$$

$$\alpha_0\alpha_3 + \alpha_1\alpha_4 + \alpha_2\alpha_5 = 0$$

$$\alpha_0\alpha_4 + \alpha_1\alpha_5 = 0$$

( [ 0 0 6 4 ] . 3 )

10

$$\alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 - \alpha_5 = 0$$

$$\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = \sqrt{2}$$

( [ 0 0 6 4 ] . 4 )

したがって、この構成をとまなう。

20

$$W\{\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_2, \alpha_3\}$$

( [ 0 0 6 4 ] . 4 )

6ポイントウェーブレットは、4ポイント $\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$ に基づいて作られる。実際、4ポイントおよび6ポイントウェーブレットは、 $Q = Q(\alpha_2, \alpha_1)$ に基づいて、4ポイントの作成を開始する（円は、自動的にPおよび所定のQに至る）。

一組の6ポイントウェーブレットを生成する次の段階は、OPを直径とする別の円を描いて、この円の内側に長方形ORPSを描くことから始まり、次いてOSを利用して処理を

30

【 0 0 6 5 】

ウェーブレット群

一組の6ポイントウェーブレットを生成する次の段階は、OPを直径とする別の円を描いて、この円の内側に長方形ORPSを描くことから始まり、次いてOSを利用して処理を継続する。これは、ウェーブレットのポイントの数を毎回2つ増加するメカニズムを提供する。完全群は、第1のポイントQに関連し、したがって角度に関連する。

【 0 0 6 6 】

セクション3：ウェーブレットを利用した情報抽出、データ圧縮、およびポストレコーディング分析

40

本発明は、多数の個別の処理を備えており、当該処理のいくつかあるいは総ては、ウェーブレットを利用して、多次元にデジタル化されたデータから情報を抽出し、このデータを圧縮する場合に適用できる。また、本発明は、セクション1に記載されているようなポストレコーディング分析を実行する自然なコンテキスト（natural context）を提供する。

【 0 0 6 7 】

データは、少なくとも2次元のデジタル化されたデータの形式を採用できる。通常、次元の一つは時間であり、連続したデータを生成する。この処理は、特に、デジタル化されたビデオイメージの処理に適しており、これは、2つの空間的な次元、追加の色、および情報のインテンシティプレーン（intensity plane）を有する連続した画像ピクセルを備

50

えている。

【 0 0 6 8 】

以下の説明では、好適な実施例が示されているが、この処理は、同様にいずれの多次元のデジタル化データセットに適用できる。

【 0 0 6 9 】

処理のうち、特に関連のあるものは以下である。

a . カーネル置換 ( 段落 [ 0 0 8 6 ] および [ 0 1 5 3 ] )

b . 画像の適応ウェーブレット表現 ( 段落 [ 0 0 8 6 ] および [ 0 1 5 3 ] )

c . 画像の違いの自動閾値化 ( 段落 [ 0 0 8 9 ] および [ 0 1 7 5 ] )

d . 複数の比較方法を可能にする特注のテンプレートの使用 ( 段落 [ 0 0 9 8 ] 、 [ 0 1 5 9 ] および [ 0 1 6 5 ] )

e . 調整可能な特別なウェーブレットのセット ( 段落 [ 0 0 6 0 ] )

f . 検出されたイベントの精度の高い識別と分類のためのブロックの計算方法 ( 段落 [ 0 1 8 6 ] )

g . 圧縮されたデータの知覚品質を改善するための、制御された誤差拡散を伴う局所的な閾値および量子化レベルの使用 ( 段落 [ 0 1 1 4 ] 、 [ 0 1 4 9 ] 、 [ 0 2 0 0 ] および [ 0 2 0 5 ] )

【 0 0 7 0 】

本発明の実施例の詳細が記されており、この実施例は、添付の図面に図示されている。この例は、システムを示しており、このシステムでは、連続したビデオ画像が取得され、要約データの形式の情報を抽出すべく処理され、圧縮され、記録され、検索され、問い合わせられ、結果が表示される。概要が図 5 に示されている。

【 0 0 7 1 】

可能な限り、同一あるいは同様のパートを参照するために、図面および説明全体を通して同一の参照番号が用いられる。

【 0 0 7 2 】

シーケンス内の各画像フレームは、ウェーブレット分解が行われる。好適な実施例では、使用は、セクション 2 で説明するようにパラメータ化されたウェーブレットで構成され、処理の演算を支援する。しかしながら、いずれの好適なウェーブレット表現を使用できる。

【 0 0 7 3 】

以下の説明では、特に明記しない限り、「画像」あるいは「フレーム」が処理されるという趣旨の説明は、オリジナル画像だけでなく、ウェーブレット階層の全体について言及している。

【 0 0 7 4 】

図 5 は、取得から ( ブロック 1 2 ) 、処理 ( ブロック 1 3 ) および分類 ( ブロック 1 4 ) を介して格納し ( ブロック 1 5 ) 、クエリーにより検索する ( ブロック 1 6 ) 処理全体を示している。

【 0 0 7 5 】

ブロック 1 2 では、ある実施例では、一時的な一連のビデオ画像 1 1 が、1 以上のビデオリソースから受信され、必要であれば、以下のステップに適したデジタル形式に変換される。いずれのビデオリソースからのデータは、所望のフレームレートで打ち切られる。多くのリソースからのデータは、後に複数のストリームにアクセスするために、並行して処理でき、相互参照できる。

【 0 0 7 6 】

ブロック 1 3 では、画像が取得されると、低レベル分析が実行される。この分析は、画像データの一連のピラミッド ( 多重解像度 ) 変換が実行され、画像圧縮の前の適応ウェーブレット変換が行われる。

この分析は、不必要なノイズおよびを特定および除去し、体系的なあるいはランダムなカメラモーションを特定する。低価格の C C T V カメラは最も弱いと、画像の色の構成要

素のノイズを処理することは重要である。説明される一連の処理は、画像のいずれの部分  
が静止あるいは静的な背景を構成するか、且ついずれの部分のシーンがシーンの動的な構成要素  
であるかを特定する。これは、カメラモーションおよび証明の変化から独立して行われる。  
詳細は、図 6 および段落 [ 0 0 8 4 ] 乃至 [ 0 1 0 4 ] に示されている。

【 0 0 7 7 】

デジタルマスクは、現在の処理の重要な部分である。マスクは符号化され、一時的に 1  
あるいは複数レベルのビットプレーンとして格納される。一組のデジタル画像マスクが生  
成され、異なる属性を有する画像領域を描画する。ある位置の 1 ビットマスクデータは、  
属性を備え、あるいは備えていない。多数のビットで符号化されたマスクは、属性値を格  
納できる。マスクは、画像がマスクされていない場合に画像を破壊する可能性のある処理  
から、画像の特定の部分を保護するのに利用され、あるいはデータの一部を選択的に修正  
するのに利用される。

【 0 0 7 8 】

ブロック 1 4 では、ブロック 1 3 の分析結果が量的に見積もられ、シーンの動的な部分  
に対する深い分析が行われる。この結果は、後の要約データになる一組のデジタルマスク  
として表わされる。詳細は、図 7 および段落 [ 0 1 0 5 ] 乃至 [ 0 1 1 1 ] に示されてお  
り、このようなマスクの例が、図 1 5 に表わされている。

【 0 0 7 9 】

ブロック 1 5 は、ブロック 1 4 で説明した処理の出力である。オリジナルシーンおよび  
これに関連する要約データの適応ウェーブレット表現は圧縮され、後の検索のためにディ  
スクに格納される。詳細は、図 8 および段落 [ 0 1 1 3 ] 乃至 [ 0 1 1 6 ] に示されてい  
る。

【 0 0 8 0 】

ブロック 1 6 では、ブロック 1 5 で格納された要約データが問い合わせされ、問い合わ  
せされ、クエリーからの肯定的な応答が、圧縮された一連の画像データから検索され、イ  
ベントとして表示される。この場合の「イベント」は、ビデオフレームの連続的なシーケ  
ンスであり、この間、問い合わせされた行動が、他のビデオリソースからの複数の関連す  
るフレームと共に継続する。詳細は、図 A E および A F 並びに段落 [ 0 1 1 8 ] 乃至 [ 0  
1 2 5 ] に示されている。

【 0 0 8 1 】

図 6 は、本発明に従ってビデオシーケンス 2 1 を構成要素に分解する第 1 のフェースを  
構成するいくつかの「処理ノード」で構成される長いループ（ブロック 2 2 乃至 3 1 ）を  
示している。

【 0 0 8 2 】

このループには多くの重要な特徴が存在する。（ 1 ）これは何度も実行可能であり、こ  
のため、リソースがこれを利用可能である。（ 2 ）いずれのノードにおける処理の実行は  
、時間、リソース、および全体のアルゴリズムストラテジーに応じて任意である。（ 3 ）  
処理はさらに、リソースの可用性に応じて、先の画像を考慮してもよい。この反復処理は  
、以下のように表わすことができる。

$$S_j = S_{j-1} + I_j, \quad S_{-1} = 0$$

（ [ 0 0 8 2 ] . 1 ）

ここで、 $S_{j-1}$  はループ  $j-1$  の終わりの知識（knowledge）の状態であり、 $I_j$  はル  
ープ  $j$  の新しい状態  $S_j$  を生成すべく追加する情報である。

【 0 0 8 3 】

このループの目的は、データを多数の構成要素、(1)ノイズ、(2)分析のために明瞭(クリーン)にされ、圧縮されるデータ、(3)データの静止した構成要素、データの静的な構成要素、およびデータの動的な構成要素に分割することである。これらの言葉の定義は辞書にあり、この分割されるデータの詳細な説明は、段落[0128]乃至[0131]にある。

【0084】

ブロック21では、一連のビデオフレームが受信される。

【0085】

ブロック22では、各フレーム21は、いくつかの適応ウェーブレットを利用して変換される。ある実施例では、演算効率のために、小さな整数係数を有する4タップ整数ウェーブレットが利用される。これにより、演算効率のよいデータの初回分析が可能になる。

10

【0086】

ブロック23では、ブロック22で演算された最新のビデオフレームのウェーブレット変換とこれの前の状態との違いが、算出および格納される。この処理の一の実施例では、データポイントとデータポイントの違いが算出される。これにより、演算効率のよいデータの初回分析が可能になる。この処理の別の実施例では、フレーム間のより複雑な違いが、段落[0153]に詳細に説明される「ウェーブレットカーネル置換」を用いて算出される。ウェーブレットカーネル置換の利点は、明確な背景モデルの必要なく証明が変化するため、違いを除去するのが効率的であることである。

【0087】

20

ブロック24では、連続的なフレームが、体系的なカメラモーションについてチェックされる。ある実施例では、これは、ブロック23で算出されるフレームの違いにおける第1レベルのウェーブレット変換の相関する本質的な特徴により実施される。段落[0134]は、この処理の他の実施例をさらに詳細に説明している。演算されたシフトは、外挿処理(extrapolation process)を介して連続的なカメラモーションを予測するために記録される。デジタルマスクはコンピュータ処理され、先の画像と重なる最新の画像部分を記録し、演算および格納された重複した領域間の変換を記録する。

【0088】

ブロック25では、体系的なカメラモーションの残りの部分が、不規則なカメラモーションによるカメラのぶれとして扱われる。カメラのぶれは、視覚的に画像を見辛くするだけでなく、連続的なフレームを相互に関連付けないため、オブジェクトの特定がより困難になる。通常、カメラのぶれの補正は反復的な処理であり、初めの近似は、画像フィールド内の静止した背景が分かると改良できる(段落を参照)。この特性により、画像の静止した構成要素は維持され、これにより、この本来の目的のための特別なバックグラウンドのテンプレートを素速く形成できる。このテンプレートの主要な特徴を分離することにより、相対的に容易にカメラのぶれの補正が可能になる。詳細については、段落[0134]を参照。

30

【0089】

ブロック26では、いくつかの(自動的に)決定された閾値の範囲内で異なる最新の画像データ部分が、先の画像に対して変化しない画像領域を決定するマスクを生成するのに利用される。ブロック26を通る最初のパスにおいて、この処理のある実施例では、閾値が、異なる画像の切りつめられたヒストグラムの最大値から算出され、別の実施例では、ピクセルの違いのメジアン統計から算出される。このマスクは、各パスで再調整される。詳細については、段落[0135]を参照。

40

【0090】

ブロック27では、ブロック26で算出されたマスクが、画像ノイズの分配の統計的なパラメータの精度を高めるために使用される。これらのパラメータは、画像をノイズ成分および明瞭(クリーン)な構成要素に分離するのに用いられる。

【0091】

ある反復的な実施例では、カメラモーションおよびノイズを評価するために、処理がブ

50

ロック 23 に戻る。

【0092】

低価格の C C T V を使用する場合、信号の色の構成要素内のノイズを適切に扱うことが重要であり、これは非常に重要である。画像のはっきりしたエッジは、特に色のノイズの影響を受けやすい。

【0093】

ブロック 28 では、ブロック 27 からの最新の明瞭な画像が、新規の適応ウェーブレット変換を利用してピラミッド分解される。データのこのようなピラミッド分解では、ピラミッドの各レベルは、ウェーブレットの特性が各レベルで画像特性に適応するウェーブレットを利用して形成される。ある実施例では、ピラミッドの解像度の高い（上位）レベルで使用されるウェーブレットは、高解像度ウェーブレットであり、ピラミッドの解像度の低いレベルで使用されるウェーブレットは、同じパラメータ化された群の低解像度ウェーブレットである。さらに、この処理は、段落 [0139] に示されており、段落 [0060] および [0065] で説明されており、好適な様々なウェーブレット群が表わされている。

【0094】

画像のこの適応ウェーブレット分解を表わす多数の係数は、検知、量子化および圧縮できる。分解のいずれのレベルでも、検知および量子化は、（a）ウェーブレット変換で発見された特徴がある位置、（b）モーションが（ブロック 26 のモーションマスクから、あるいは処理が繰り返される場合にはブロック 30 から）検出された位置に応じて変更できる。

【0095】

ブロック 29 では、最新の画像の新しい画像が、低解像度情報を利用して、先行する画像のウェーブレット変換から生成される。この新しい画像は、先の画像と同じ全体的な照度を有する。この新規な処理「ウェーブレットカーネル置換」は、フレーム間の照度の変化を補正するのに利用される。この処理は、段落 [0153] でより詳細に説明する。

【0096】

ブロック 30 では、ブロック 29 のカーネルが変更された最新の画像と先行する画像との違いは、シーン内のモーションによるものであり、カーネル置換は、照度の変化による大きな除去効果を有する。デジタルマスクが生成され、モーションが検出される領域を決定することができる。

【0097】

[段落 0096] と同じ原理が、既に格納された先行する多数の画像およびテンプレートに適用される。様々なテンプレートの格納戦略が利用可能である。この処理のある実施例では、等比数列の 1 つ古いデータフレーム（即ち、直前のデータフレーム）、2 つ古いフレーム、4 つ古いフレーム等の様々なテンプレートが格納される。この制限は、データストレージと、多くのテンプレートをチェックするのに必要な追加の演算リソースによるものであり、テンプレートのより詳細な説明は、段落 [0159] にある。

【0098】

テンプレートは、様々な方法により、データのウェーブレット変換から生成される。最も単純なテンプレートは、1 つ前の画像のウェーブレット変換である。ある実施例では、先行する  $m$  ウェーブレット画像の平均が、追加のテンプレートとして格納される。別の実施例では、過去のウェーブレット画像における時間により重み付けされた平均が格納される。これは、最近の画像  $I_j$  を利用してテンプレート  $T_{j-1}$  を  $T_j$  に更新するのに以下の式が利用される場合は演算効率がよい。

$$T_j = (1 - \alpha) T_{j-1} + \alpha I_j$$

( [ 0 0 9 8 ] . 1 )

ここで、 $\alpha$  は、テンプレートに対する最新の画像の僅かな貢献 (fractional contribution) である。この種の式では、 $\alpha^{-1}$  順序フレームのメモリを有し、最前面のオブジェクトの動作は不鮮明になり、最終的には消える。揺れる葉を有する木などの静的な背景は、この平滑作用により処理され、顕著な動きに対するモーション検出は必要ない (段落 [ 0 1 3 1 ] を参照)。このようなテンプレートを取得するには、少なくとも  $\alpha^{-1}$  フレームの「ウォームアップ」期間が必要である。

この処理の別の実施例では、複数のテンプレートが、複数の  $\alpha$  値のために格納される。いくつかの実施例では、 $\alpha$  は、画像  $I_j$  がどれ程直前の画像と異なるかに依存し、非常に類似しない画像は、 $\alpha$  がそのフレームのために小さくされない限りテンプレートを汚染する。

【 0 0 9 9 】

いくつかのテンプレートの過去のマスクが生成され、ノイズフィルタされた画像の過去の活動のレベルを示す。過去に蓄積された長さは、各マスクの各ピクセルに割り当てられたメモリの量に依存し、マスクを継続的に更新可能な計算能力に依存する。マスクは、ウェーブレット変換の総てのレベルで保持する必要はない。

ある実施例では、これらのマスクは 8 ビットである。「最新のマスク」は、先の 8 つのフレームにおける各ピクセルの活動を 0 ビットあるいは 1 ビットとして符号化する。2 つの「活動レベルマスク」は、「0」および「1」の状態と、過去の連続する「1」の数の連続的なランレングスとの間の平均的な変化の割合を符号化する。他の実施例では、他の状態分析が利用され、様々な可能性が存在する。これは、前面および背面のモーションを分割する前に、画像の総ての位置における活動のレベルを符号化する手段を提供する。1 以上の活動レベルマスクは、要約データの一部として格納してもよい。しかしながら、これらは通常、あまり圧縮されないため、ある実施例では、解像度の低いマスクのみが、テンプレートの更新の割合  $\alpha$  に応じて間隔をあけて格納される。

【 0 1 0 0 】

最新の画像およびそのピラミッド表現は、将来のデータと比較するためにテンプレートとして格納される。最も古いテンプレートは、ストレージに問題があるときに失ってもよい。テンプレートの詳細については、段落 [ 0 1 5 9 ] を参照。

【 0 1 0 1 】

ある反復的な実施例では、ノイズおよび照度の変化の影響の評価の精度を高めるために、処理がブロック 2 7 に戻る。このループには多くの重要な特徴が存在し、( 1 ) これは何度も実行可能であり、このため、リソースがこれを利用可能であり、( 2 ) いずれのノードにおける処理の実行は、時間、リソース、および全体のアルゴリズムストラテジーに応じて任意であり、( 3 ) さらに処理が、リソースの可用性に応じて、先の画像を考慮してもよい。反復する場合には、初めのループにおいて総てのステージを実行する必要はない。

【 0 1 0 2 】

ブロック 3 1 では、モーション分析は、( 動作が自由である静止した背景とは対照的に ) 動作が制限されている静的な背景を考慮する方法により行われる。判断閾値が、動的に設定され、背景の動作が存在する領域を効果的に鈍感 (desensitise) にし、複数の過去のテンプレートとの比較が行われる。これにより生じ得る感度の喪失は、複数の期間に亘って統合されたテンプレートを利用することにより補われ、これにより、局所的な動作をばかす (段落 [ 0 0 9 8 ] 並びに段落 [ 0 1 3 1 ] および段落 [ 0 1 5 9 ] の説明を参照)。



## 【 0 1 0 3 】

この結果は、前面の活動があるウェーブレット変換された画像の位置の暫定的な特定である。これは、空間的および時間的な相互関係が考慮される場合にリファインされる（次の段落および段落 [ 0 1 8 4 ] を参照）。

## 【 0 1 0 4 】

ブロック 3 2 では、ブロック 3 1 で動作が検出された画像の位置は、検出と画像の領域の変遷を表わす時間的な相関関係との空間的な相関関係を考慮して再評価される。この再評価は、多重解像度のウェーブレット階層の総てのレベルにおいて行われる。この詳細については、段落 [ 0 1 8 6 ] を参照。

## 【 0 1 0 5 】

図 7 は、動的な前面のピクセルを要約データになる一連の対象とするマスクに時間的および空間的にグループ化する処理を示している。ブロック 3 2 は、図 6 から本図に繋がっている。

## 【 0 1 0 6 】

ブロック 4 3 では、ブロック 3 1 で明らかになった動的な前面データが、空間的および時間的に分析される。この再評価は、多重解像度のウェーブレット階層の総てのレベルで行われる。

## 【 0 1 0 7 】

ある実施例では、空間的な分析は、実質的に相関関係分析であり、ブロック 3 1 で明らかになった前面の動的な各要素は、これらの間の近傍の要素との近接性に応じて記録される（ブロック 4 4）。これは、総てのスケールにおけるまとまりのあるピクセルグループであることが好適であり、分散および分離したピクセルは好ましくない。

ある実施例では、時間的な分析は、前面の動的な要素を、先行するフレームの対応する要素と比較し、先行するフレームのために既に生成された要約データと比較することにより行われる（ブロック 4 4）。この実施例では、格納された時間的な参照は、過去の 1, 2, 4, 8, ... フレームに保持されている。この変遷の唯一の制限は、最初のストレージの利用可能性である。

## 【 0 1 0 8 】

ブロック 4 5 では、空間的および時間的相関関係の記録の結果が解釈される。ある実施例では、これは、事前に割り当てられた空間的および時間的パターンテーブルに従って行われる。これらは、空間的および時間的シーブと称される（ブロック 4 6 および 4 7）。

## 【 0 1 0 9 】

ブロック 4 8 では、様々な空間的および時間的なパターンが、オブジェクトおよびシーン選別に分類される。オブジェクトの場合、動作ベクタは、様々な手段（[ 段落 0 1 9 9 ] を参照）により算出でき、必要であれば、ウェーブレット変換の解像度の低い構成要素を利用してサムネイルを格納できる。シーンの変化の場合、必要であれば、関連する過去の画像のシーケンスは、ウェーブレット変換の解像度の低い構成要素から収集でき、将来の参照のために検査可能なトレーラを生成する。ある実施例では、これらのマスクを生成する処理およびパラメータのオーディットも保持される。

## 【 0 1 1 0 】

ブロック 4 9 では、画像マスクが、ブロック 4 8 で発見されたデータストリームの各属性のために生成され、画像データ内で属性が配置される位置を描く。別の実施例は、別のカテゴリを示すマスクセットを提供する。これらのマスクは、要約データの基礎を形成する。図 1 5 は、シーンの主要な変化する構成要素を示すこれらのマスクを示している。

## 【 0 1 1 1 】

ブロック 5 0 では、最終的にノイズの無いウェーブレット符号化されたデータが、次の段階である圧縮のために利用できる。ウェーブレット係数の圧縮は、環境に依存する。

## 【 0 1 1 2 】

図 8 は、後のクエリーおよび検索のためにデータを圧縮、符号化および格納するのに必要な処理を示している。ブロック 4 9 および 5 0 は、図 7 から繋がっている。

10

20

30

40

50

## 【0113】

ブロック61では、ブロック49で生成された要約データが、データチェックサムにより可逆圧縮され、符号化が行われ、この符号化は必要である。

## 【0114】

ブロック62では、適切に符号化されたウェーブレットデータは、効率的に格納すべく、初めにビットレートを低減する局所的な適切な閾値および量子化により、次に得られた係数の符号化により圧縮される。ある実施例では、少なくとも2つの位置が決定され、単一のマスクで符号化され、ウェーブレット表現の位置は、動的な前面モーションが存在するのものと、何も存在しないものとがある。別の実施例では、静止しているが、静止した背景が存在しない（例えば、動く葉）ウェーブレット表現のこれらの位置は、マスクにより符号化され、それ自体の閾値および量子化が与えられる。マスクは、検索および再構成のために符号化および格納され、画像バリデーションコードが、法的な目的で生成される。ある実施例では、得られる圧縮データは、チェックサムと共に符号化され提供される。

## 【0115】

ブロック63では、ブロック61および62からのデータは、データベースフレームワークに挿入される。ある実施例では、これは、コンピュータファイルシステムの単純な使用であり、別の実施例では、これは関連するデータベースである。複数の入力データストリームの場合、時刻同期の情報は、特にデータが時刻帯境界を跨ぐときに、非常に重要である。

## 【0116】

ブロック64では、総てのデータが、ローカルあるいはネットワークストレージシステムに格納される。データは、同時に追加および検索される。ある実施例では、データが任意のストレージ媒体（例えば、DVD）に格納できる。認証されたオーディットトラレーが、データに沿って書き込まれる。

## 【0117】

図9はデータ検索の処理を示しており、この処理では、クエリーが要約データにより呼ばれ、このクエリーに応じて、このクエリーの条件を満たす記録されたイベントのリストが生成される。このクエリーは、イベントの最後の選択が終了するまでリファインできる。ブロック64は、明確にするために図8と重複している。

## 【0118】

ブロック71では、データがブロック72のクエリーに利用されるようになる。ブロック72のクエリーは、データベースを備えるローカルコンピュータに、あるいはコンピュータネットワーク上のリモートステーションを介して発行してもよい。クエリーは、要約データのための1以上のデータストリームと、このようなデータを持たない関連するストリームを必要とするであろう。クエリーは、複数の場所にある異なるデータベースに分配された要約データを呼び出してもよく、複数の異なる場所にある複数の異なるデータベースからデータにアクセスしてもよい。

## 【0119】

ブロック73では、クエリーに合致する要約データが検索される。クエリーに合致するフレームリストが生成される。これを「キーフレーム」と称する。ブロック74では、イベントリストが、発見されたキーフレームに基づいて形成される。イベントと、当該イベントが生成されるデータフレーム（キーフレーム）には大きな違いがある。イベントは、単一のフレーム、あるいは複数の入力データストリームから複数のフレームで構成してもよい。複数のデータストリームが関与する場合、異なるストリームで定義されるイベントは、時間的に一致している必要はなく、クエリーにより見つかったキーフレームと同じデータベースからのイベントである必要もない。これにより、データベースは広範な調査目的に利用できる。この広範なマッチングは、ブロック75で行われる。キーフレームの周囲のイベントの生成は、段落[0234]で説明される。

## 【0120】

ブロック76では、ブロック74および75で生成される複数のイベントに関連するデ

10

20

30

40

50

ータは、関連するウェーブレット符号化されたデータ（ブロック 77）から検索され、関連する利用可能な外部データ（ブロック 78）から検索され、必要に応じて圧縮される。ブロック 77 および 78 からのデータフレームは、イベントにグループ分けされ（ブロック 79）、表示される（ブロック 80）。

#### 【0121】

ブロック 81 では、検索をリファインする可能性により検索の結果が評価される（ブロック 82）。検索は、選択されたイベントのリストを生成して終了する（ブロック 83）。

#### 【0122】

図 10 は、イベントの選択の後に行われる処理を示している（明確にするためにブロック 81 が重複している）。

#### 【0123】

ブロック 91 では、イベントデータは、好適なフォーマットに変換される。ある実施例では、このフォーマットは、オリジナルデータを格納するときに用いられるのと同じ適応ウェーブレット圧縮である。別の実施例では、フォーマットは第三者のフォーマットでもよく、第三者のために利用可能なデータビューワが存在する（例えば、Ogg-Vorbis フォーマットのオーディオデータ）。

#### 【0124】

ブロック 92 では、データは、後の参照のため、あるいは検査目的のために注釈が付される。このような注釈は、単一のローカルデータベースに格納されるテキストでもよく、あるいはこのようなデータアクセスのために設計された第三者のツール（例えば、SGML に基づいたツール）でもよい。ブロック 93 では、このデータ検索がどのように計画され、実行されたかを示すオーディットトレースと、データの完全性を保証するバリデーションコードが、パッケージに付加される。

#### 【0125】

ブロック 94 では、クエリーから得られイベントデータを有するイベントリスト全体（ブロック 79）といずれかの注釈（ブロック 92）が、パッケージが検索されるデータベースあるいは位置に格納するためにパッケージされる。ブロック 95 では、検索の結果が他の媒体にエクスポートされ、ある実施例では、取り外し可能あるいは光学式記憶（例えば、リムーバブルディスク装置あるいは DVD）である。

#### 【0126】

データ構成要素

#### 【0127】

ノイズ（N）は、シーンのいずれかの部分を正確に表示しない画像データの一部である。ノイズは通常、機器の影響により発生し、画像データのクリアな評価を損なう。一般に、ノイズの構成要素は、画像データに関係しないもの（例えば、重なった映像「スノー」）であると考えられる。これは、ノイズが画像の局所的特性に直接的に依存するため、必ずしも問題であるとは限らない。

#### 【0128】

静止した背景（S）は、固定され且つカメラレスポンス、照度、あるいは移動するオブジェクトによるオクルージョンにおける変化によってのみ変化するシーン要素で構成される。静止した背景は、カメラがパン、ティルト、ズームしているときでも存在してもよい。あるシーンを異なる時間に再び閲覧することにより、同じ静止した背景要素が表示される。建物および道路は、静止した背景を構成する要素の例である。何日もかけて木から落ちる葉は、この分類に該当するであろうが、これは、単にタイムスケールの問題にすぎない。

#### 【0129】

静的な背景（M）は、シーン内で固定されたシーン要素で構成され、このシーンでは、異なる時間にシーンを再び閲覧することにより、僅かに置き換えられる方法により同じ要素が表示される。モーションは局所化され、制限され、そのタイムバリデーションは一時

10

20

30

40

50

的でもよい。窓の反射は、このカテゴリに属するであろう。静的な背景の構成要素は、制限された静的なランダム処理 (stationary random process) としてモデルができる。

【 0 1 3 0 】

動的な前面 (D) は、データを取得する間に、シーンを入力あるいは残す、あるいは多くの動作を実行するシーンの特徴である。このプロジェクトの一の目標は、誤った検出や検出漏れがなく、前面で行われるイベントを特定することである。

【 0 1 3 1 】

構成要素 ( [ 0 1 2 7 ] 乃至 [ 0 1 3 0 ] ) 間のこれらの区別は、処理を行う者が、構成要素の分離の様々な様子 (アスペクト) の取り扱いについて決定できる実用的な区別である。人がシーンに入り、椅子を動かして、シーンから立ち去る場合を想定する。この椅子は、移動される前および移動された後は、シーンの静止した部分である。移動中では、椅子は、椅子を動かしている人と同様にシーンの動的な部分である。これは、構成要素の分類が時間とともに変化することを強調しており、分類する手段は、このことを考慮すべきである。

【 0 1 3 2 】

これらの区別を設計する場合に、いくつかの注意点がある。「静止した」背景と「静的な」背景との区別は、なされる価値判断に関連するタイムスケールの選択の問題である。3本の枝は、何秒か風に揺られ、一方、この同じ木は、数週間かけて葉を失う。揺れ動く木の枝は、背景の「動く」構成要素を備え、一方、このような動作が無い場合、葉の喪失は、静止した背景の一部として正確に表示される (たとえゆっくり変化する構成要素であっても)。木の様子は、暗くなるにつれ変化するが、これは、分解の静止した態様として考えるのが最もよい。

【 0 1 3 3 】

これは、数学的に要約すると、画像データ  $G$  を時間に依存する多数の構成要素の合計として表わされる。

$$G(\underline{x}, t) = G^S(\underline{x}) + G^M(\underline{x}, \varepsilon t) + G^D(\underline{x}, t)$$

( [ 0 1 3 3 ] ) . 1

最初の構成要素は、実際に静止しており、2番目の構成要素は、前述したようにゆっくり動き、3番目の構成要素は、前面および背景の貢献 (contribution) に分類されるべき動的な構成要素である。当面の目的のために、体系的に移動するカメラは、 $G^S$  に纏められる。より正確な定義は、カメラモーションから得られる空間的座標  $\underline{x}$  における変換を明確に示す必要があるであろう。 $G^D$  を前面  $G^{DF}$  構成要素および背景  $G^{DB}$  構成要素に分類する原理は、 $G^{DB}$ 、動的な背景構成要素、が実質的に静的であることを示すべきであり、

$$\int_{-\infty}^T G^{DB}(\underline{x}, t) dt \rightarrow \hat{G}^S(\underline{x})$$

( [ 0 1 3 3 ] ) . 2

この場合、いくつかの静的な背景は、

10

20

30

40

$$\hat{G}^S(\underline{x})$$

である（これは、木が風に揺られていない場合には、木がある位置を表わす）。時間により重み付けされたテンプレートを使用することにより、これを実現でき、また動的な前面の構成要素の分類を可能にする（段落 [ 0 1 5 9 ] を参照）。

パラメータ は、変化の低速性により、何を意味するのかを決定する。理想的には、 は映像の取得レートより少なくとも一桁小さい。いくつかの移動する構成要素が存在してもよく、各要素がそれ自身のレート を有する。

10

$$G(\underline{x}, t) = G^S(\underline{x}) + \sum_i G_i^M(\underline{x}, \varepsilon_i t) + G^D(\underline{x}, t)$$

( [ 0 1 3 3 ] ) . 3

これらのうちの最も遅いものは、静止した構成要素に纏めてもよく、静止した構成要素の「断熱な (adiabatic)」変化を明らかにするものが提供される。

20

【 0 1 3 4 】

カメラモーション、特にカメラのぶれの補正は、長い歴史を有する技術であり、多くの解決手段が存在する。ある実施例では、Herriot等のQuad Correlation法、(2000)Proc SPIE, 115, 4007、が用いられている。最近の見解である天文学の画像安定化コンテキストのThomas等の(2006)Mon. Not. R Asrt. Soc. 371, 323を参照。

【 0 1 3 5 】

第 1 レベルのノイズフィルタ

【 0 1 3 6 】

ノイズ成分の最初の評価は、同一のシーンの 2 つの連続したフレームの違いを算出し、「静止した背景」に分類された画像部分の統計的な分配、即ちこれらの違いをマスクしたものを調べることにより取得される。ノイズの変化は、

30

$$\sigma_n = 1.483 \text{ Median}(M_n - M_{n-1})$$

( [ 0 1 3 6 ] ) . 1

によりおおよそ算出でき、ここで、

40

$$M_n = \mathbf{M}(F_n - F_{n-1})$$

( [ 0 1 3 6 ] ) . 2

は、処理されていないフレーム間の違いをマスクしたものである。

初めのパスでは、フレーム  $F_n$  について何も決められていないため、マスクは空である ( $M = I$ 、アイデンティティ)。

50

## 【 0 1 3 7 】

この違いのメジアンは、（フレーム間の知覚可能な違いにより生ずる）異常値に対してより安定しているため、分散（variance）を算出するのに用いられる。これは特に、演算速度の利点において、分散が画像ピクセルの無作為な副標本から予測される場合に有益である。

（１）シーン間の全体的な光量の変動の補正、（２）静止した背景部分以外の画像要素の補正の２つの補正は、ノイズ分布のこの予測を必要とする。これらの補正の第１の補正は、「ウェーブレットカーネル置換」処理（段落「 0 1 5 3 」）を介して行われる。これらの補正の第２の補正は、いずれの画像部分が大きく変化したかを調査する分析の「VMD」構成要素を介して行われる。

10

## 【 0 1 3 8 】

マスクが空（ $M = I$ ）の場合、クリーニングは、分散に要素数を掛けた値よりも小さい値を有する異なる画像の総てのピクセルをゼロに設定することにより行われ、違いのヒストグラムが再形成され、最初の違いがゼロになる（「ウェーブレットシュリンケージ（wavelet shrinkage）およびその変化」）。

マスクが空でない場合、分散の値は、映像が変化した領域と、フィルタにより画像の外観を傷付ける可能性のある位置（重要なエッジなど）とを考慮して、フレーム  $F_n$  を空間的にフィルタをかけるのに利用される。（１）位相依存Weiner-type Filteringおよび（２）非線形feature sensitive filter（例えば、Teager-style filter）などの空間的なフィルタリングに依存する機能に関する実行可能な技術が存在する。

20

ノイズの除去は、画像のウェーブレット変換が行われ直前に行われ、ノイズ除去は圧縮にとって有利である。

## 【 0 1 3 9 】

図 1 1 は、新たに取得した画像の分析の最初のループで実施される処理を統合した図である。この図は、一組のフレーム  $F_0, F_{-1}, F_{-2}, F_{-3} \dots$  を示しており、これは、既に取り得られ、一連のテンプレート  $T_0, T_{-1}, T_{-2}, T_{-3} \dots$  およびエッジ特徴画像  $E_0, E_{-1}, E_{-2}, E_{-3} \dots$  を生成するのに使用される。これらの画像  $E_i$  は、カメラのぶれの検出および監視に使用される。 $F_0$  および  $T_0$  は、新しい画像  $F_1$  の参照画像になる。

カメラのぶれが検出された場合、カメラのぶれは、この時点で補正される（段落 [ 0 1 3 4 ] を参照）。補正は、続く反復処理において、後の改良を必要としてもよい。ここで、（できる限りシェイクを補正した） $F_1$  は、先行するフレーム  $F_0$  と現在のテンプレート  $T_0$  と比較される。違いマップは計算され、VMDディテクタに送信されると、違いマップ双方の変化を検出するか、あるいは検出しないか２つの可能性がある。これは、段落 [ 0 1 3 7 ] に記載されている。変化が検出されなかった場合、ノイズの特性は、 $F_1 - F_0$  の違いから直接的に予測でき、この場合、いずれの違いもノイズによるものでなければならない。 $F_1 - F_0$  はクリーンにでき、以前の  $F_0$  のクリーンなバージョン  $f_0$  に追加できる。これは、 $F_1$  のクリーンなバージョン  $f_1$  を生成し、これは、次の反復処理で利用できる。

30

違いが存在した場合、ノイズの補正は、フレーム  $F_1$  で直接に行われる。 $F_1$  と  $F_0$  あるいは  $F_1$  と  $T_0$  との違いがどこに存在するかを示すマスクを利用して、このレベルで検出された変化が存在する  $F_1 - F_0$  および  $F_1 - T_0$  の部分を保護する。これらの違いをクリーニングすることにより、変化が検出された位置以外の位置をクリーンにされた  $F_1$  のバージョン  $f_1$  を可能にする。変化が検出されたマスク内のこれらの領域は、Teagerフィルタあるいはこれを一般化したフィルタなどの単純な非線形クリーニングエッジ保存ノイズフィルタ（nonlinear cleaning edge preserving noise filter）を利用して、クリーンにすることができる。

40

## 【 0 1 4 0 】

ピラミッド変換によるデータ表現

## 【 0 1 4 1 】

50

ウェーブレット変換および他のピラミッド変換は、多重解像度分析の例である。このような分析は、スケールの階層でデータを表示でき、科学および工学の分野で利用可能である。この処理は図 1 2 に示されている。ピラミッドの各レベルは、オリジナルデータを再生成するために追加される情報を示す一組のデータと共に、オリジナルデータよりも小さく解像度が低くデータを含んでいる。通常、常にではないが、ピラミッドのレベルは、各次元の要素によりデータを再構成する。

#### 【0142】

これを実施する方法は多く存在し、ここで使用する方法は、数学者がこれを発見した後は、Mallatの多重解像度表現と呼ばれている。図 1 3 の上図は、初めにウェーブレット $W_1$ を利用して、次にウェーブレット $W_2$ を利用して、階層がどのように生成されるのかを示している。下図は、データが格納される方法を示している。

1次元データセットのウェーブレット変換は、隣接データグループの合計および違いに関連する2つの部分で構成される処理である。合計は、これらの隣接データの利点を提供し、収縮(shrunk)を提供するのに利用される。解像度の低いデータ。違いは、変換の合計した部分により生成された平均からの偏差を示し、データの再生成に必要である。合計した部分はSで示され、違いはDで示される。2次元データは、初めに各行を水平に処理し、次に各列を垂直に処理する。これは、図 1 3 の{SS, SD, DS, DD}で示す4つのパートを生成する。

#### 【0143】

ウェーブレット階層。通常、所定の多くのウェーブレットから選択した単一の特定のウェーブレットにより生成されるデータ階層が利用される。したがって、図 1 3 では $W_1 = W_2$ である。この文脈では、ウェーブレットの共通の選択は、CDF群、実施が容易であるために特に広く人気があるCDF(2, 2)変化(variant)(これはまた、「5 - 3 ウェーブレット」として知られている)の様々な個体である。

#### 【0144】

適応ウェーブレット階層。ここで説明する処理では、ウェーブレット変換の特定の階層が利用され、階層のメンバが、1以上の値によりパラメータ化されたウェーブレットの連続的なセットから選択される。この群の4ポイントウェーブレットは、1のパラメータのみを必要とし、6ポイントのメンバは、2のパラメータ等を必要とする。パラメータ値の離散集合の場合、4ポイントのメンバは、有理数である係数を有し、これらは、演算効率がよく正確である。

#### 【0145】

異なるレベルで利用されるウェーブレットは、このパラメータの異なる値を選択することにより、あるレベルから次のレベルでは異なる。これを適応ウェーブレット変換という。この処理のある実施例では、解像度の高いウェーブレットは、高解像度レベルで利用され、次いで、解像度の低いウェーブレットは、移動した低解像度レベルで利用される。

#### 【0146】

別々のウェーブレットの場合、効果的なフィルタ帯域幅は、ウェーブレットフィルタのFourier変換により決めることができる。他よりも広い通過帯域を有するものが存在し、我々は、狭い通過帯域のウェーブレットを上位(高解像度)レベルで使用し、広い通過帯域のウェーブレットを下位(低解像度)レベルで使用する。この処理のある実施例では、帯域幅により順序付けられたパラメータセットに分けられたウェーブレットが使用される。

#### 【0147】

最低レベル(これらのレベルは、ほぼオリジナルデータの大きさである画像に変換が影響を与えるレベルである)では、我々は、これらのレベルの圧縮を適正にすべく詳細を保存し背景を取得することに関心がある。最高レベル(これらのレベルは最小の画像を有する)では、我々は、重要な特徴をもたない画像内で大きな構造を写像(map)する。さらに、ここでは精度は、エラーがブロックアーティファクトとして鮮明に表示される最低レベルに伝わるため重要である。

10

20

30

40

50

## 【 0 1 4 8 】

閾値。ウェーブレット変換のSD、DSおよびDD部分を閾値化は、画像データ圧縮の観点から無視してもよいピクセル値を除去する。閾値がより大きくなり得るこれらの位置を特定することは、高圧縮を実現する重要な方法である。不適切な位置を特定することは、認識される画像の劣化を最小限にするため重要である。特徴検出およびイベント検出は、高閾値化が避けられる位置（空間的および時間的）を指摘する。

## 【 0 1 4 9 】

量子化。量子化は、数値の範囲が少ない数値により表示される処理を参照し、これにより、データの表示を（おおよそではあるが）よりコンパクトにできる。量子化は閾値化の後に行われ、局所的な（空間的および時間的）画像コンテンツに依存する。閾値化が控えめな位置は、量子化も控えめであるべきである。

10

## 【 0 1 5 0 】

ビットBORROWING (Bit-borrowing)。データ値を表示すべく非常に僅かな数値を使用することは、多くの欠点があり、再形成された画像の質にとって非常に有害である。状態は、様々な周知な技術が非常により改善できる。この処理のある実施例では、あるデータポイントの量子化によるエラーは、近接するデータポイントに伝わり、これにより、局所的な領域の情報コンテンツ全体をできる限り抑えることができる。残りのもの (remainder) の均一な再分配は、均一な照度の領域内の輪郭の抑制を助ける。さらに、特徴を有するこの残りのものの賢明な配置は、画像の詳細に対する損傷を抑制し、非常に見た目がよくなる。これにより、アーティファクトなどの輪郭等を低減する。これを「ビットBORROWING」と称す。

20

## 【 0 1 5 1 】

ビットBORROWING技術における残りのものの再分配のメカニズムは、ウェーブレットのような分析が相対的に平準化されたデータから画像の特徴を容易に表わすため、ウェーブレット分析において単純化される。各レベルにおける変換のSDおよびDS部分は、残りのものの再分配に付される重み付けを決定する。これは、ビットBORROWING処理の演算効率をよくする。

## 【 0 1 5 2 】

ウェーブレットカーネル、テンプレートおよび閾値

## 【 0 1 5 3 】

ウェーブレットカーネル置換。これは処理であり、この処理では、先行する画像の大きな縮尺（低解像度）の特徴が、現在の画園内の同様の特徴を置き換えるために作られる。照度は一般的には大きな属性であるため、この処理は、別の画像上の一の画像から光を本質的に写生し、非常に強く速い変光面で実施される（他のものの）動作検出を可能にする効果がある。この技術は、ウェーブレットでは各レベルのSD、DSおよびDD構成要素が非常に小さいDC構成要素のみを有するため、非常に効果的である。

30

## 【 0 1 5 4 】

この処理のある実施例では、画像の事前処理サイクルの一部として実施される第1のレベルVMDを改善するためにカーネル置換が行われる。これは、照度の変化を取り除き、画像の前面における変化を発見するのを改善する。

40

## 【 0 1 5 5 】

ウェーブレットカーネル置換が図14に示されており、現在の画像のカーネル構成要素F3の代わりに配置される現在のテンプレートのカーネル構成要素T3が示されており、ウェーブレット構成要素が $J_i \{ J_0, J_1, J_2, T_3 \}$ である現在の画像の新しいものを生成する。この新しいデータは、ノイズを予測し様々なマスクを算出するために、オリジナル画像 $I_i \{ F_0, F_1, F_2, F_3 \}$ の代わりに使用する。

形式的に、この処理は以下のように説明できる。取得した画像を $\{ I_i \}$ と称する。我々は、 $\{ J_i \}$ と呼ばれるウェーブレット変換を介してこの一組の画像から取得でき、このウェーブレット変換では、照度の空間的に大きな変化が、先行する画像の変換のカーネルを利用することにより行われる。

50



S S 構成要素階層を有するウェーブレット変換を有する同じシーケンスから 2 つの画像 { $I_i$ } および { $I_j$ } がある場合、

$$\{I_i\} = \{^1SS(i), ^2SS(i), ^3SS(i), \dots, ^kSS(i)\}$$

( [ 0 1 5 5 ] ) . 1

10

$$\{I_j\} = \{^1SS(j), ^2SS(j), ^3SS(j), \dots, ^kSS(j)\}$$

( [ 0 1 5 5 ] ) . 2

我々は、画像 j のために画像 i のカーネルを用いて新しい画像を作る。

$$\{J_j\} = \{^1\overline{SS}(j), ^2\overline{SS}(j), ^3\overline{SS}(j), \dots, ^k\overline{SS}(j)\}$$

20

( [ 0 1 5 5 ] ) . 3

新しいウェーブレットの S S 部分の上線に注意すべきであり、これらは、i 番目のウェーブレットカーネルを用いて画像 j を再形成したという事実により修正される。また、我々は、変換の S D、D S または D D を修正しないことに留意すべきであり、これらは、 $^k S S(i)$  から { $J_i$ } を再形成するのに直接的に利用される。

したがって、我々は、画像  $i = j - m$  および j の間の違いが修正された周辺光を算出できる。

30

$$\delta_{j,(m)} = J_j - I_{j-m}$$

( [ 0 1 5 5 ] ) . 4

この別個の画像は、m フレーム前の画像が得られるので、周辺光による変化に加えて画像内の変化を表す。

40

j - m の画像を有する画像 j のカーネルを更新するかどうかについての問題が存在する。実際には、演算効率により、メモリに一次的に保存された現在の画像のウェーブレット変換全体が存在するため、前述したような置換が実行される。

【 0 1 5 6 】

相対的变化。実際には、ウェーブレット変換の単一のレベル p の変化のみを見ることができる。

$$\delta_{j,(m)}^p = {}^p\overline{SS}(j) - {}^pSS(i), \quad m = j - i$$

( [ 0 1 5 6 ] ) . 1

これは、画像  $i$  のウェーブレット変換を有する画像  $j$  のカーネル置換されたウェーブレット変換の  $p$  番目のレベルの  $SS$  部分の違いを示している。タイムラグの値  $m$  は、単純にフレームレートに依存し、実際には、動作の変化が知覚できる所定の時間の長さになる。しかしながら、これを実施することにより、多重解像度分析により得られるサイズ識別を失い、可能であれば、変換全体を利用することがよい。

10

【 0 1 5 7 】

現在の画像。現在の画像を先行する画像に対して評価される単一の画像であると単純にみなすことはよくあることである。これは通常の場合である。しかしながら、この処理の実施例が存在し、この処理では、先行する画像の選択の平均を有する単一の現在の画像を置換するのが有用である。

【 0 1 5 8 】

一時的な除去。周囲の監視の応用例では、動物、人々および車両などの一時的な事象により画像が汚染されないことが有用である。最近の画像における適切な時間加重平均であるデータを利用することにより、これらの一時的なものが除去する。これを「一時的なものが除去された現在の画像」と称する。

20

このような状況に適応されるこの処理のある実施例では、以下の式は、直前の画像  $I_j$  を利用して、「一時的なものが除去された現在の画像」 $C_{j-1}$  を定義および  $C_{j-1}$  を  $C_j$  に更新するのに用いられる。

$$C_j = (1 - \tau) C_{j-1} + \tau I_j$$

30

( [ 0 1 5 8 ] ) . 1

ここで、 $\tau$  はテンプレートに対する現在の画像の分数で表わされる貢献 (fractional contribution) である。この種の式の場合、画像は、 $j-1$  番目のフレームの情報を保持している。この応用例では、テンプレートは、 $j-1$  フレームよりも十分に長い時間 (秒に対して、日あるいは週) にわたって格納される。

【 0 1 5 9 】

テンプレートおよびマスク

【 0 1 6 0 】

40

テンプレート。ここに説明した処理を通して、「画像テンプレート」と呼ばれるものの変化が一時的に格納される。通常、テンプレートは、画像データ自体 (あるいはピラミッド変換) の階層的なレコードであり、単独あるいは組み合わせにより、現在の画像と先行する画像とを比較する基礎を提供する。このようなテンプレートは通常、常にではないが、好適な重み係数を有する統合された先行する画像群により構成される (段落 [ 0 1 6 5 ] 参照)。

【 0 1 6 1 】

また、テンプレートは、現在の画像の一種でもよく、現在の画像の平準化された画像は、例えば、アンシャープマスク処理あるいは他の単一の画像処理のために保持される。

50

## 【 0 1 6 2 】

マスク。テンプレートのようなマスクも画像であるが、これらは、画像の特定の外観を描写するために生成される。したがって、マスクは、画像内あるいはピラミッド変換内で、閾値よりも大きな動作がどこに存在するか、あるいは特定のテクスチャが発見された位置を示してもよい。したがって、このマスクは、マップの情報の内容を規定する属性および値のリストを有するマップである。属性値が、「trueあるいはfalse」あるいは「yesあるいはno」である場合、情報は、1のビットマップとして符号化できる。属性がテクスチャの場合、マップは、フラクタルローカル次元(fractal local dimension)を4ビット整数等として符号化する。

## 【 0 1 6 3 】

10

マスクが得られる画像にマスクが適用される場合、特定のマスクの属性値を共有する画像の領域が描写される。同じ属性を有する2つのマスクが一組の画像に適用される場合、マスクの違いは、属性における画像の違いを示している。

## 【 0 1 6 4 】

1以上のマスクからの情報は、データストリーム用の要約データの生成に利用される。要約は、様々なマップを決定する属性を示しており、要約はこのマップから生成される。図15は、要約データストリームに入れられる動的な前面と静止したおよび静的な背景構成要素に対応する3つのレベルのマスクを示している。

この図では、VMDマスクが、開いたドアとこのドアから出てくる人を現わしている。動く背景マスクは、動く葉と低木を示している。照度マスクは、動く木の影により光が変化する位置を示している(この最後の構成要素は、ウェーブレットカーネル置換によりかなり除去されているため、動く背景として現れていない)。

20

## 【 0 1 6 5 】

特有のテンプレート。テンプレートは、当該テンプレートに対して現在の画像の内容あるいは現在の画像の変化を評価するための参照用の画像である。最も簡単なテンプレートは、前の画像である。

$$T_j = I_{j-1}$$

30

## ( [ 0 1 6 5 ] ) . 1

過去のm個の画像の平均は、ややリファインされている。

$$\bar{T}_j = \frac{1}{m} \sum_{k=1}^m I_{j-k}$$

40

## ( [ 0 1 6 5 ] ) . 2

これは、ノイズを低減するテンプレートを生成する効果を有する。過去の画像の時間加重平均は、さらに有用である。

$$\tilde{T}_j = \alpha I_j + (1 - \alpha) \tilde{T}_{j-1}$$

( [ 0 1 6 5 ] ) . 3

ここで、 $\alpha$  は、テンプレートに対する現在の画像の分数で表わされる貢献である。この最後の等式は、代替的に以下のように説明できる。

10

$$\tilde{T}_n = \alpha \sum_{r=0}^n (1 - \alpha)^r I_{n-r}$$

( [ 0 1 6 5 ] ) . 4

これは、

20

$$\tilde{T}_n$$

を、予め重み係数  $(1 - \alpha)^r$  を有するフレーム  $r$  画像を有する過去のフレームの加重和として示している。この種の式の場合、テンプレートは、 $(1 - \alpha)^{-1}$  フレームのようなメモリを有し、このため、このテンプレートを取得するには、少なくとも  $(1 - \alpha)^{-1}$  フレームの期間の「ウォームアップ」が必要である。

実際には、 $\alpha$  は、画像  $I_j$  と先行する画像  $I_{j-1}$  との違いに依存し、全く類似しない画像は、 $\alpha$  がそのフレームのために小さくされない限り、画像を汚染するだろう。 $\alpha$  を選択するフレキシビリティは、動的な前面のオクルージョンがテンプレートを大きく変更する場合に用いられる（段落 [ 0 1 8 0 ] 参照）。

30

【 0 1 6 6 】

最近のマスク。「最近のマスク」は、先行する 8 フレームが 0 ビットあるいは 1 ビットの間、各ピクセルの活動を符号化する。

【 0 1 6 7 】

活動レベルマスク。2 の「活動レベルマスク」は、それまでの「1 (ones)」の連続的な数字の平均および変化を符号化し、第 3 の最近の活動マスクは、「1」の現在のランレングスを符号化する。

40

【 0 1 6 8 】

他のテンプレート。テンプレートを生成する場合は、 $I_j$  より前の画像に限定されないことに留意すべきである。以下のような将来の画像に基づいてテンプレートを考えることは特定の目的に有用である。

$$\dot{T}_j = I_{j+1} - I_{j-1}$$

50

( [ 0 1 6 8 ] ) . 1

あるいは

$$\ddot{T}_j = \frac{1}{2}(I_{j+1} - 2I_j + I_{j-1})$$

( [ 0 1 6 8 ] ) . 2

この表記が示唆するように、これらは、 $I_j$  が得られたときの画像ストリームの第 1 および第 2 の時間導関数を推定するものである。このようなテンプレートを使用することは、「将来の」画像が取得された場合に、ストリームの分析をバッファリングすることにより採用するタイムラグが必要である。

他の様々な可能性が存在する。平準化された画像テンプレートは

$$S_j = \text{Smooth}(I_j)$$

( [ 0 1 6 8 ] ) . 3

であり、ここで、「Smooth」は、画像  $I_j$  に適用するスムージングオペレータ (smoothing operator) の可能性のある数字を表わしている。マスクされた画像テンプレートは、

$$\hat{T}_j = \text{Mask}(T_j)$$

( [ 0 1 6 8 ] ) . 4

であり、ここで、「マスク」オペレータは、適切に定義された画像マスクをテンプレート画像  $T_j$  に適用する。このリストは明らかに完全ではなく、単に説明のためである。

【 0 1 6 9 】

最近のマスク。「最近のマスク」は、先行するフレームにおけるシーンの各ピクセルの活動の測定を符号化する。ある活動の測定は、2つの連続するフレームの間あるいは1のフレームとそのときのテンプレートとの間のピクセルの違いが、段落 [ 0 1 8 1 ] に定義する閾値よりも上か否かである。

ある実施例では、これは、画像データの大きさの 8 ビットマスクとして格納され、活動は、「0」あるいは「1」として過去の 8 フレーム分記録される。ピクセルの違いの数値が求められる度に、このマスクは、適切なビットプレーンを変更することにより更新される。

【 0 1 7 0 】

長期間の過去のマスク。最近のマスクと同様に、これらは、先行するシーンから過去のデータを符号化する。違いは、このようなマスクが、過去の基準時における活動データを格納できることである。均等に離れた位置は更新し易いが、更新するのが困難である幾何学的に配置された位置と同じくらい有用でない。このようなマスクは、シーン活動に関する長期間の動きの削除を容易にする。

【 0 1 7 1 】

活動レベルマスク。2つの「活動レベルマスク」は、最近のマスクで示したように、所定のピクセルの統計的概要を表わしている。これらの最初のマスクのエントリは、このピクセルで行われた状態の変化の数あるいはレートを記録する。これは、移動平均として容易に保持され、レートが  $R_{j-1}$  で、次の変化が  $e_j = 0$  あるいは  $1$  である場合、レート  $R$  の推定は、以下のように更新される。

$$R_j = \varepsilon R_{j-1} + (1-\varepsilon) e_j$$

10

( [ 0 1 7 1 ] ) . 1

数字 は、レートが平均化されたデータの範囲を示す。

2番目のマスクは、ランの平均長の計数を保持し、ここでは、 $e_j = 1$  : 「活動ランレングス」である。これは、レートの推定と同じように算出され、レートが前述したように平均である場合、活動ランレングスである。

【 0 1 7 2 】

これらの活動マスクは、維持するのに非常に費用がかかるため、いくつかの実施例では、データピラミッドの小さなレベルおよびこれよりも上の小さなレベルにマスクを制限するのが有用である。通常、主要な画像の解像度の半分の最大を維持することが、最も適切であると知られており、これは、図 1 2 のレベル 1 あるいはレベル 2 である。

20

【 0 1 7 3 】

背景変化マスク - 動作の検出がない。静止した背景（本質的に変化しない）に対する2つの質問がある。通常、何も存在しない静止した背景の一部として見なされるもののなかには何か存在するのか。反対に、以前は存在しなかった静止した背景の一部であるものが、今は存在するのか。明らかに、この種の変化は、変化するシーンに動作があることを必要とする。しかしながら、この質問は、単に変化を発見することについての質問よりも、非常に複雑である。この質問は、静止した背景が既に回復したのか、もし回復したのならそれはいつか。

30

【 0 1 7 4 】

背景の動作を記録するマスクは、これを処理することができないため、比較あるいは相関により静止した背景の特徴を識別可能な特別な背景変化マスクを使用する必要がある。このマスクは、道程な前面のオブジェクトにより隠される位置を除き、静止した背景の構成要素が変化しない場合は一定のままである。このため、静止した背景マスク間の違いは、理想的には0であり、格納するのに何も必要としない。

この目的のための理想的なマスクは、ウェーブレットピラミッド（図 1 2 を参照）のレベル 1 の SD および DS 部分の合計であり、これは、比較的高い解像度でシーンの特徴をマップする。カーネル置換されたウェーブレット表現から生成されたこのような2の連続的なマスクを差別化により比較がされ、対応する動的な構成要素のマスクにアクセスできる。後者の場合、我々は、シーンの移動する部分に対応する特徴を除去できる。得られた背景変化マスクは圧縮され、要約データの一部として格納される。

40

【 0 1 7 5 】

画像間の違い

【 0 1 7 6 】

違い画像。このセクションのために、我々は、「画像」の語を以下に示すものと見なす。(1) データストリームから取得された画像、(2) データストリームから取得された画像であって連続的に処理されるもの。この場合、縮小された画像などの画像の変換あるいはそのウェーブレット変換も含まれる。(3) 画像あるいはその一の変換の一部。

すなわち、このような配列のストリームから取得されるデータ配列を先行する配列と比較

50

することを考慮する。

【 0 1 7 7 】

我々は、ストリーム内の  $j$  番目のこのような配列を記号  $I_j$  で示し、我々が比較する関連するオブジェクト（「テンプレート」）を起動  $T_j$  で示す。 $T_j$  は、ストリーム  $I_j$  の他のメンバから定義される様々なテンプレートとなり得る（セクション 0 を参照）。

我々は、画像とこれらの様々なテンプレートとの違いを評価する方法を検討する。この違い画像を、

$$\delta_j = I_j - T_j$$

10

( [ 0 1 7 7 ] ) . 1

とする。

$j$  を生成するピクセルの平均は、テンプレート  $T_j$  および画像  $I_j$  を生成する総ての画像が一致しない限り、ゼロを必要としない。これは、 $j$  のピクセル値の分析を考慮する場合に重要な点である。

画像  $j$  内のピクセルの値は、周辺光の変化が、カーネル置換（段落 [ 0 1 5 3 ] 乃至 [ 0 1 5 5 ]）によるものである場合、平均して 0 である。ピクセルが 0 でない場合、我々は、これらが画像内の実際の変化に対応するか、これらが統計変動であるかを判断する必要がある。

20

【 0 1 7 8 】

偏差ピクセル。ここで、時間の関数、異なる画像のピクセルの値として、トラッキングに集中する。我々が開発した基準は、ピクセルの位置や、空間的に隣接するピクセルが何をしているかを考慮することなく、各ピクセルの時系列の変化を利用する。これは、ノイズの空間的な分配を想定することなく、不規則なノイズが処理できる利点を有する。この変化の空間的な分配は後述する（段落 [ 0 1 8 4 ] 参照）。

この処理の一実施例では、データ中のピクセルの時刻歴は、以下に記載されモデル化される。ピクセル閾値レベル  $L_i$  は、各ピクセルの時刻歴を示すランダム処理のための「ランニング識別レベル」 $M_i$  と称す量の観点により、この時刻歴から定義される。

30

【 0 1 7 9 】

違い画像  $j$  のために、我々が、（統計的な試験により）ピクセル値はノイズ「偏差ピクセル値」によるものではないと考えられる閾値レベル  $L_i$  以上を決定できると仮定する。安全計数を  $\lambda$  とした場合に、

$$|\Delta_j| > \lambda L_i$$

40

( [ 0 1 7 9 ] ) . 1

であるとき、我々は、違い画像  $j$  内で値  $j$  の偏差を有するピクセルであると判断する（我々は、 $j$  内のピクセル値の傾斜分布のために、 $j$  の肯定的および否定的な値のための異なる境界を選択するが、記載を簡潔にするために、これらは同じであると仮定する）。

ピクセル値の変化  $j$  は、静的でないランダムな処理であるため、 $L_i$  の値は、 $|j|$  の値の上側のエンベロープを示すべきである。上側のエンベロープは、このような処理にとって予測するのが困難であることはよく知られているため、我々は、単純化された推測を用いる。これは特に、これがピクセル毎に実施され、演算時間の制限がある場合に当て

50

はまる。

【 0 1 8 0 】

識別レベル。ピクセルごとにこれらの値の演算を利用した  $j$  の  $m$  個前の値、以下のよう  
な等式に基づく識別レベル  $M_j$  を検討する。

$$M_j = \max\{|\Delta_{j-1}|, |\Delta_{j-2}|, |\Delta_{j-3}|, \dots, |\Delta_{j-m}|\}$$

$$M_j = \text{mean}\{|\Delta_{j-1}|, |\Delta_{j-2}|, |\Delta_{j-3}|, \dots, |\Delta_{j-m}|\} + K$$

10

$$M_j = \beta |\Delta_{j-1}| + (1 - \beta) M_{j-1}$$

( [ 0 1 8 0 ] ) . 1

第 1 の等式は、 $m$  時間間隔枠の移動における信号の高さを計測することによりエンベロー  
プを直接的に取得する試みである。第 2 の等式は、安全マージン とともに最近の  $m$  信号  
の高さの係数の平均値を利用する。最後の等式は、先行する信号の高さの時間加重平均で  
あり、量 は、相対的な時間の重み付けを示す。これは、好適なメカニズムである。

20

【 0 1 8 1 】

ピクセル閾値レベル。前述したような識別レベルの場合 ( 段落 [ 0 1 8 0 ] )、我々は  
、以下に示すように各ピクセル毎にピクセル閾値レベル  $L_j$  を算出する。このピクセルの  
閾値は、 を「メモリパラメータ」とすると、以下のように設定される。

$$L_j = \alpha L_{j-1} + (1 - \alpha) M_j$$

30

( [ 0 1 8 1 ] ) . 1

は、識別レベル  $M_j$  ( [ 0 1 8 0 ] . 2 の第 3 の等式 ) の計算に代入される量 と同一  
ではないことに留意すべきである。我々は、ピクセルが偏差しているとして「マーク」す  
るか否かを判断するために比較を行い、ピクセルが偏差しているか否かにより、次のフレ  
ームの計算のために  $L_j$  の値をリセットする。

40

$$L_j = \begin{cases} L_j, & \text{if } \Delta_j > \lambda L_i, \\ M_j & \text{otherwise} \end{cases}$$

( [ 0 1 8 1 ] ) . 2

換言すれば、我々は、このピクセルがすれている場合、このピクセルの閾値を更新しない  
。これは、異常な状況により閾値を決定することによって導入されるバイアスを避ける。  
合否基準が 3 偏差に基づく場合、例えば、この処理は、単に閾値の計算における 3 拒

50



否 ( rejection ) である。

【 0 1 8 2 】

移動する背景の補正。この処理は、閾値がノイズピークを超えることを可能にする。所定のノイズの分散の確率密度の場合、レベルは、ピクセルが間違いなく偏差していると考えられる所定の確率が存在するように構成される。ピクセル違いの分配の所定の確率密度がない場合、パラメトリック的ではなく、リファイン度を変化させる標準的な試験を利用して決定がされる。移動する背景の実質的な効果は、制限的および反復的な方法でシーンが変化する領域の動作の検出を鈍感にすることである。これは、例えば太陽により木の影が風により動く場合に生じ、画像違いの局所的な変化が増加するため、閾値が引き上げられる。

10

これは、映像検出システムで続いて起こる誤認警報を避けるための重要なメカニズムである。この欠点は、鈍感にすることにより重要なイベントを見逃す危険性があるため、このような状況では、補足的な検出メカニズムが必要であることである。ある実施例では、これは、相対的に長いメモリを有するテンプレートがこのような動作を顕在化し吸い上げるため、このようなテンプレートを利用することにより解決される。画像比較は、素速く動く背景の特徴が比較的少ない背景を対象にする ( 段落 [ 0 1 3 1 ] および [ 0 1 5 9 ] 参照 ) 。

【 0 1 8 3 】

パラメータ。前述した実施例では、画像ストリーム内の大きな変化を検出するために設定されたいくつかのパラメータが存在する。これらのパラメータの一部は、最初は固定されており、他のパラメータは、周囲の状況とともに変化し「学習」する。我々は、前述した処理を利用する場合、設定あるいは決定されるべきいくつかのパラメータを特定できる。

20

m

これは、比較するためのフレームのタイムラグである。毎秒 2 5 フレームは、毎秒 3 フレームよりも明らかに長い。毎秒 2 5 フレームのサンプルを毎秒 3 フレームでサンプル化して、同一の値 m を利用して終了することは明らかである。このため、m は、フレームレートに正比例する。比例定数の値は、探索される動作が、フレーム探索速度 ( traversal speed ) の観点からどれほど速いかによる。

30

これは、所定のピクセルにおける検出の感度であり、先に観測された値に対する観測されたピクセルの値の変化の程度である。我々は、ピクセルの値を試験するために平均あるいは標準偏差よりも境界の上限を使用することを留意すべきである。 は、偏差していない値の標本の一次解析に関連する。

メモリ要因は、次のフレームの閾値の値を更新するときに考慮した過去の閾値がどのほどであるのかを示している。これは、初めの閾値の値を無意味にするのに十分に周辺環境が変化する期間を示しているため、フレーム取得レートに関連する。

これらのパラメータは、初期値で設定され、10 フレームほど調べた後に自動的に調整できる。学習方法はより複雑であるが ( 人が所定の期間ノイズの分析を予測し計算できるが、これは、実際にはする価値がない )、これは比較的短い「ティーチングサイクル」である。

40

【 0 1 8 4 】

偏差ピクセル分析。前述した実施例は、画像内で一組の偏差ピクセルを生成し、このピクセルは、データ値の変化が、自動的に割り当てられた閾値を超える。これまでは、シーン内のピクセルの位置は不適切であり、我々は、所定のピクセルの変化の値と、その位置の前歴を単に比較するだけであった。これは、空間的に一様でないノイズの分配を処理することができる。

この問題は、これらが、画像内の本当の変化を表示しているのか、あるいは単に画像ノイズおよび周囲条件の一連の統計変動であるかを判断することである。これを支援すべく、

50

偏差したピクセルの空間的な分配のコヒーレンスに着目する。

【0185】

偏差したピクセルの空間的なコヒーレント。我々が画像内に例えば10の偏差したピクセルを発見した場合、これらが画像内でランダムに分散しているときよりも、これらが密集していたときのほうがより印象的であろう。実際、我々がノイズの分散の詳細について知っている場合、我々がランダムに分散している10の偏差したピクセルを取得する可能性を計算する。

【0186】

ブロックスコアリング。ここで、我々は、下表のようにスコアを割り当てることにより、偏差したピクセルの密集の度合を評価する簡単な方法を示す。

10

異常ピクセル	スコア2
それぞれが、隣接するピクセルに水平あるいは垂直に接する	スコア2
それぞれが、隣接するピクセルに斜めに接する	スコア1

偏差したピクセルはそれぞれ、どれだけの隣接するピクセルが偏差しているかに依存する。スコアが中心のピクセルに割り当てられた多数の3×3のパターンが、図16の「ピクセルスコア」に示されている。

20

スコアは、僅かに偏差したピクセルの存在が一目で分かるため、隣接するピクセルの数が増加すると急激に増加しており、重要でないと考えられるパターンは、他のパターンよりもスコアが小さい。水平および垂直に交差する5つのピクセルのスコアは10であり、一方、斜めの6つのピクセルのスコアは9である（再度の行のパターン1およびパターン3）。

状態は、全体のパターンのスコア、即ち、所定の領域の偏差した総てのブロックのスコア全体を調べる場合に決まる。図16の「特別なパターンのスコア」は、3×3ブロックの偏差したピクセルのスコア全体を示しており、ここでは、3×3ブロックが独立しており、隣接する偏差したピクセルが無いと仮定する。ブロックスコアの直線的でない相互の補強が存在し、このため、スコアは、3×3の領域内のブロックパターンが密集している場合

30

【0187】

ある実施例では、ブロックは、画像内の水平、垂直あるいは斜めの構造のスコアリングを助長すべく、重み付けられている。これは、パターン分類の第1の段階である。この処理は、明らかに階層的に実行され、これに関する唯一の制限は、計算資源における要求を倍にすることである。

【0188】

最後の説明として、偏差したピクセルの要約画像は、ピクセルのスコアを格納する必要がないことに留意すべきであり、これらは、必要なときはいつでも再計算でき、偏差したピクセルの位置を知ることができる。したがって、偏差したピクセルを報告する要約画像は、単純な1ビットプレーンのビットマップであり、対応するピクセルが偏差している場合は1であり、その他の場合は0である。

40

これにより、映像の変化のための要約データの検索が速くなる。

【0189】

動作ベクタ。

【0190】

動作ベクタを計算することは、多くの圧縮アルゴリズムおよびオブジェクト認識アルゴリズムの本質的な部分である。しかしながら、極度の圧縮が求められない限り、圧縮用の動作ベクタは必ずしも使用する必要はない。我々は、シーン内のオブジェクトを特定し追跡するために動作ベクタを使用する。利用される方法は新規性があり、これは、ブロック

50

あるいは相関関係に基づくものではない。この方法は、ウェーブレットカーネル置換技術（段落〔0153〕乃至〔0155〕）の使用の利益を享受し、これは、背景の照度の体系的な変化を十分に除去する（背景の照度の問題は、オプティカルフロー計算の周知の問題である）。

【0191】

この説明は、カーネル置換されたウェーブレット変換の構成要素 $\{^jSS\}$ に該当する。ウェーブレットレベル毎に、 $\{^jSS\}$ の各構成要素のピクセル値の対数を生成する。0および負の値を避けるために（後者は、ウェーブレット変換の結果として現れる）、我々は、総ての値が完全に正となるように、レベルに依存する一定のオフセットをピクセル値に付加する。

10

$$^j\rho = \ln(^j\kappa + ^jg), \quad ^jg \in \{^jSS\}$$

（〔0191〕）．1

計算で使用する総ての画像は、同じオフセットを取得する。対数で表わされたピクセル値は、浮動小数点数として保持されるが、計算速度を高めるために、これらは、4あるいは5ビットの符号付整数に変更してもよい。

20

$^j$  の時間導関数の値を求めるために、我々は、3つの瞬間における $\{^jSS\}$ を必要とし、この3つの瞬間は、現在の時間、先行するフレームの時間、および次のフレームの時間である。この3つの瞬間のデータ値を、添字 - 1、0 および 1 とする。したがって、

$$\dot{\rho}_0 = (\rho_1 - \rho_{-1})$$

30

（〔0191〕）．2

および

$$\ddot{\rho}_0 = \frac{1}{2}(\rho_1 - 2\rho_0 + \rho_{-1})$$

（〔0191〕）．3

40

これらのフィールド毎に、新しい、より平準化されたフィールドを算出する。

$$\phi(x) = \sum_{\text{neighbors } i} w_i \dot{\rho}_0(x_i)$$

（〔0191〕）．4

および

50

$$\Phi(x) = \sum_{\text{neighbors } i} w_i \ddot{\rho}_0(x_i)$$

( [ 0 1 9 1 ] ) . 5

重み係数  $w_i$  は、双方の等式で同じである。重み付けは、潜在的なフィールドが、 の第 1 および第 2 の時間導関数、対数密度であるソースを有するラプラスの方程式の近似解になるように選択される。

10

速度フィールドは、ウェーブレット変換の総てのスケールにおけるこれらの潜在的なフィールドの空間的なグラディエントを利用して算出される。

【 0 1 9 2 】

低いフレームレートでは、第 1 のデリバティブフィールド は、侵入があった場合でもゼロという結果を生成してもよいことに留意すべきである。これは、侵入が現在のフレームにおいてのみ発生するときに、両側の画像フィールドが同一であるためである。しかしながら、これは、第 2 のデリバティブフィールド で確実に拾われるであろう。

反対に、一様にゆっくり動くターゲットは、第 2 のデリバティブフィールド にゼロを与えるが、これは、第 1 のデリバティブフィールド で確実に拾われるであろう。

20

双方のフィールドは、偏差したピクセルの分析が変化がないことを示した場合、ゼロあるいはゼロに近い値になることを留意すべきである。速度を測定するためには変化があるはずである。

【 0 1 9 3 】

圧縮および格納。

【 0 1 9 4 】

ウェーブレット符号化されたデータ。この段階では、データストリームは、ウェーブレットデータのストリームとして符号化され、オリジナルデータよりも多くのメモリを占有する。ウェーブレット表現の利点は、非常に圧縮できることである。しかしながら、高い質を維持する重要な圧縮への道は真っ直ぐではなく、多くの技術を組み合わせる必要がある。

30

【 0 1 9 5 】

データ構造。図 1 7 は、データ圧縮処理の要素を要約している。オリジナル画像データストリームは、一組の画像  $\{ F_i \}$  で構成されている。これらは、様々な比較が行われるテンプレート  $\{ T_i \}$  の実行中のシーケンスに組み込まれる。これらの 2 つのストリーム、画像およびテンプレートから、別のストリームが生成され、これは、違い映像  $\{ D_i \}$  のストリームである。

違いは、隣接するフレーム間の違い、あるいはフレームと選択されたテンプレートとの違いである。我々は、「隣接する」により、隣接するフレームが直前のフレームであることを主張しているのではなく、比較は、フレームレートおよび画像ストリームの他のパラメータに依存するタイムラグにより実施してもよい。

40

様々な可能なテンプレートの説明については、段落 [ 0 0 9 8 ] および段落 [ 0 1 6 0 ] 以降を参照。また、「現在のフレーム」の利用の代替に関する段落 [ 0 0 9 8 ] および段落 [ 0 1 5 8 ] を参照。説明は、この原理の他の可能な実施例の存在を認めるとともに、一般性を失うことなくフレームおよびテンプレートについて続けて言及する。

違い処理のパートナーを参照画像  $\{ R_j \}$  と称する。すなわち、 $R_j$  は、 $T_i$  あるいは  $F_i$  のうちの一つでもよい。

【 0 1 9 6 】

圧縮のオブジェクトは、データ  $\{ D_i \}$  および  $\{ R_j \}$  で構成されるデータストリームである。これらのストリーム双方は、適応ウェーブレットを利用して、我々の事例では一

50

組のウェーブレットを利用して、ウェーブレット変換される。ウェーブレットは、浮動小数点あるいは整数でもよく、あるいはこれらの混合でもよい。象徴的に、我々は以下のように記すことができる。

$$F_k = R_i + D_k$$

10

( [ 0 1 9 6 ] ) . 1

どれだけの  $D_k$  を所定の  $R_j$  とともに使用すべきかということは重要な質問である。原則的に、我々は、1の参照画像  $R_0$  のみを必要とする。しかしながら、非常に長いシーケンスは、( a )  $D_k$  が、参照画像と大きく異なる将来のフレームのように大きくなり、( b ) 最近の  $D_k$  の圧縮が非常に長いデータシーケンスの処理を要するため不利である。

【 0 1 9 7 】

本質的に、個々の  $\{ D_i \}$  は、参照フレーム  $\{ R_j \}$  よりも圧縮する。この状況自体は、 $\{ R_i \}$  をこれ自体から区別し、シーケンス  $\{ R_j \}$  を新しいシーケンス  $\{ R_j, \{ k \} \}$  として表現することにより支援される。

20

$$R_k = R_i + \delta_k$$

( [ 0 1 9 7 ] ) . 2

シーケンス  $\{ R_j \}$  のメンバの事前の類似性により、 $\delta_k$  は  $R_k$  よりも少ないビットで表わすことができる。 $\{ R_j \}$  の圧縮は、格納された画像の質の決定における中心的な要素である。損失は、格納された  $R_k = R_i + \delta_k$  の質を低下と同じであるため、 $\{ \delta_k \}$  シーケンスの圧縮は、ほぼ損失無く行われるべきである。圧縮されるデータストリームは、

30

$$\{ \{ R_i, D_i, D_{i+1}, \dots, D_{i+m-1} \}, \{ \delta_k, D_k, D_{k+1}, \dots, D_{k+m-1} \}, \dots \}, \quad k = m + i$$

図 1 7 は、違いの体系化の様子を概略的に示している。

【 0 1 9 8 】

最終的な段階は、圧縮されたデータストリームをマスクするのに必要な総てのウェーブレット変換を取得することである。

$$R_k \rightarrow W_k$$

$$D_k \rightarrow v_k$$

40

( [ 0 1 9 8 ] ) . 3 a

また、参照フレームを再生成する場合、

$$\delta_k \rightarrow \omega_k$$

( [ 0 1 9 8 ] ) . 3 b

ウェーブレット変換ストリームは、

$$\{\{W_i, v_i, v_{i+1}, \dots, v_{i+m-1}\}, \{\omega_k, v_k, v_{k+1}, \dots, v_{k+m-1}\}, \dots\}, \quad k = m + i$$

10

であり、サイクル長はmである。圧縮が行われていないことに注意すべきである。

【 0 1 9 9 】

ウェーブレットデータストリーム内の各データブロックは、ウェーブレット係数の一連の配列で構成されている。

$$v_j = \{^1Q_j, ^2Q_j, \dots, ^KQ_j\},$$

20

( [ 0 1 9 9 ] ) . 4

ここで、

$$^N Q_j = \{^N SS, ^N DS, ^N SD, ^N DD\}$$

( [ 0 1 9 9 ] ) . 5

30

は、レベルN、同様に参照画像およびその違いの変換 $W_i$ および $W_k$ の場合のウェーブレット変換配列である。ウェーブレットKとして表わされるこれらの配列のうちの最も小さいものは、「ウェーブレットカーネル」と呼ばれる小さな画像を含む。現在の表記では、ウェーブレットカーネルは、

$$\text{データウェーブレットカーネル} = ^K SS$$

( [ 0 1 9 9 ] ) . 6

40

である。

【 0 2 0 0 】

圧縮。異なる種類のフレーム、参照フレーム $R_i$ 、違いフレーム $D_i$ 、あるいは異なる参照 $R_i$ は、画像の質を高く維持しつつ圧縮効率を最大限にすべく、これら自体の特別な処理を必要とする。

ここで、我々は本質的な原理のみを想起し、この処理は、閾値を決定し、好適な方法により当該閾値以下で係数を0に設定するステップと、残りの係数を量子化する方法と、最後に、これらの係数を効果的に表現あるいは符号化する方法とで構成されている。

【 0 2 0 1 】

適応コーディング。我々はまた、ウェーブレットプレーンの違い領域が、異なる閾値お

50

よび量子化を有することを想起し、閾値および量子化の特定の値を保持するデータの各領域が、マスクにより決定される。このマスクは、データの内容を示しており、データにより符号化される。

画像の一部が、その動作により、あるいは質のよい細部の存在により、特別に関心のあるものであると認定されると仮定する。特別に関心のあるこれらの領域は、低い閾値および質のよい量子化の程度（さらにレベル）を選択することができる。係数コードの異なるテーブルは、特別に関心のあるこれらの領域のために生成される。より数の多い値のために短いコードを使用することができ、こつは、2つのテーブルを保持することである。この2つのテーブルとともに、閾値の2つの値および量子化のスケーリング要因の2つの値を保持する必要がある。

10

#### 【0202】

閾値化。閾値化は、圧縮量を制御する主要な手段の一つである。あるレベルでは、閾値化は、ノイズと見なされるものを除去するが、閾値レベルが上がり、より多くの係数がゼロにされるに伴い、画像の特徴が失われていく。ウェーブレット変換マトリックスのSD、DSおよびDD構成要素は、画像データの曲率の状態を測定するため、これは、初めに損傷する画像のピクセルスケールの低曲率部分である。実際、ウェーブレット圧縮された画像は、閾値を厳しくすると「ガラスのような」外観になる。ウェーブレット変換マトリックスの $\hat{J}SD$ 、 $\hat{J}DS$ および $\hat{J}DD$ 構成要素を無効にすると、構成要素 $\hat{J}SS$ を単純に平準化し拡大した画像 $\hat{J}^{-1}SS$ が得られ、これを1以上のレベルについて実施することにより、特徴のない画像が生成される。ウェーブレットのより高いレベル（より小さい配列）は、注意深く保存されるべきであり、一方、低いレベル（大きな配列）は、閾値化が注意深く行われている場合には、知覚される損傷を画像に与えることなく破棄できることは経験則である。

20

#### 【0203】

量子化。ウェーブレット係数の量子化はまた、係数の数を低減し、これらを効率的に符号化することにより圧縮のレベルに貢献する。理想的には、量子化は、係数のヒストグラムに依存すべきであるが、実際には、これは、演算リソースに非常に高い要求する。最も簡単で一般的に効率的な量子化の方法は、係数の大きさを変更し、結果をビットプレーンに分割することである。これは、実質的に、対数間隔量子化（logarithmic interval quantization）である。係数のヒストグラムが急激に分散する場合、これは、理想的な方法であらう。

30

不適当な量子化による影響は特に、小さな輝度勾配（intensity gradient）を有する平らな復元領域に影響自体を印象付け、復元された輪郭が非常にきつくなる。幸い、例えば、誤差拡散を利用した賢明な復元は、画像の他の部分を損傷することなく、問題を解決することができる（段落[0150]および[0183]）。

ウェーブレットプレーンのスケーリング要因は、圧縮されたデータヘッダの一部として保持される。

#### 【0204】

符号化。ウェーブレット変換が閾値化および量子化されると、明確な係数の値の数が、非常に小さくなり（これは、量子化された値の数の数に依存する）、ハフマンのようなコードがこれらに割り当てられる。

40

コードテーブルは、各ウェーブレットプレーンにより保存される。通常、同一の映像ストリームからの数多くのフレーム用の同じテーブルを使用することは可能であり、好適なヘッダ圧縮技術はこれを効率的に処理し、これにより、フレーム毎にいくつかのテーブルを格納することによるオーバーヘッドを低減する。ストレージユニットは、圧縮されたウェーブレット群（以下を参照）であり、群全体に同じテーブルを使用させることは可能である。

#### 【0205】

ビットポローイング。非常に小さい数を利用してデータ値を表現することは、多くの欠点を有し、画像の質を復元するのに非常に悪影響である。この状況は、様々な周知の技術

50

により改善できる。この処理の一の実施例では、あるデータポイントの量子化によるエラーは、隣接するデータポイントに分散し、これにより、局所的な領域の情報の内容全体をできる限り保護することができる。残りのものの均等な再分配は、照度が均等な領域における輪郭削りを抑制する。さらに、特徴を有するこの残りのものの賢明な配置転換は、画像の詳細への損傷を抑制し、非常に見た目をよくする。これは、輪郭削りや他のアーティファクトを低減する。これを「ビットボローイング」と称す。

#### 【0206】

バリデーションおよびエンクリプション。我々は画像を見たときに、実際に画像が、取得、圧縮および格納された画像と同じであるか知りたいと願う。これは、画像のバリデーションの処理である。

10

我々はまた、画像データへのアクセスを制限し、復元係数を暗号化し、ユーザが有効な暗号キーを提供した場合に、これらを正しい値に変換することを所望する。

これらの問題双方は、量子化されたウェーブレット係数のテーブルを暗号化することにより同時に解決できる。アクセスが制限されていない場合、通常のキーは、ストリームデータ自体に基づいて使用される。データが本物である場合、データは正しく解凍される。第2のキーは、データアクセスが制限されている場合に使用される。

#### 【0207】

パッケージング。圧縮された画像データは、圧縮された参照フレームまたはテンプレートで構成された「パケット」に入り、当該参照フレームは、これらから抽出される一組のフレームが後に続く。これをフレームグループと称する。これは、他の圧縮スキームによる「画像のグループ」に似ているが、ここでは、参照フレームは全体的に人工的な構成であるため、我々は若干異なる名称を使用する。これは、有効に格納可能な最小のパケットである。

20

フレームグループを備える画像のウェーブレット変換のグループは、同様にウェーブレットグループと呼ぶことができる。

このようなフレームグループをより大きなパッケージに包含させることは有用であり、他に適切な用語がないために、この大きなパッケージを「データチャンク」と称し、これから派生する圧縮されたデータのパケットを「圧縮データチャンク」と称する。

通常、フレームグループはメガバイトあるいはそれより小さく、一方、有用なチャンクの大きさは数十メガバイトである。より大きな記憶素子を利用することにより、ディスクドライブからのデータアクセスをより効率的にすることができる。これは、DVD+RWなどの取り外し可能なメディアに書き込むときにも有利である。

30

#### 【0208】

要約データ。

#### 【0209】

圧縮およびエンクリプション。要約データは一組のデータ画像で構成されており、各画像は、オリジナル画像から取得されるオリジナル画像の特定の状態を要約している。要約された状態は通常、画像に含まれる情報の一部のみであり、要約データは、オリジナル画像よりも十分に小さいサイズに圧縮される。例えば、要約データの一部が、背景の動作が検出された画像領域を示す場合、各ピクセルのデータは、単一のビット（検出されたあるいは検出されていない）により表わすことができる。通常、背景で何も行われていない領域は、多くの0が存在するだろう。

40

#### 【0210】

要約データは不可逆圧縮される。

#### 【0211】

パッキング。要約データのサイズは、たとえオリジナルデータがクリーンであり圧縮されたとしても、オリジナルデータより遙かに小さい。アクセスを容易にすべく、要約データは、ウェーブレット圧縮されたデータと同一の方法によりパッケージされる。フレームグループ内の画像に関連する総ての要約データは、要約画像グループにパッケージされ、これらのグループは、ウェーブレット圧縮されたデータチャンクに正確に対応するチャン

50



クに包含される。

【0212】

データベース。

【0213】

タイムライン。オリジナルデータはストリームに入るため、フレーム識別子およびフレームが取得された時間の双方あるいは一方により、総ての形式のデータを扱うことが適切である。

圧縮されたデータは、多数のフレームグループを含むチャンクに格納される。データベースは、利用可能な総てのチャンクのリストと、各チャンクの内容のリスト（フレームグループ）と、各フレームグループの内容のリストとを保持する。

格納されたデータアイテムの最も簡単なデータベースリストは、格納されたデータアイテム、例えばチャンク、フレームグループ、あるいは単一のグループのID番号および開始終了時間からなる識別子で構成される。また、データ要素のバイトサイズに関する情報を保持することは、効率的な検索に有用である。

【0214】

図18は、要約画像データとウェーブレット圧縮データとの一対一の対応を示している。タイムラインを利用して、分析用の要約画像あるいは表示用のウェーブレット圧縮データのいずれかにアクセスできる。

【0215】

要約データおよびウェーブレット圧縮データを同じ位置に保持する必要はないことに留意すべきである。

【0216】

局所的な時間分割。この処理の主要な応用例は、ポストレコーディング分析機能によるデジタル画像レコーディングであるため、カレンダーに基づいてデータを保存するのは理にかなっている。

【0217】

要約画像。要約画像は通常、様々な解像度の1ビットプレーン画像である。これらを表示することは意味がないが、これらは、検索にとって非常に効率的である。

【0218】

圧縮画像データ。圧縮画像データは、ユーザがクエリーに応じて調べる最終的なデータである。

これは、要約データと同じレポジトリを必要としないが、データベースおよび要約データにより参照されるべきである。

【0219】

データストレージ。

【0220】

データベース。最終的にデータは、ある種類のストレージ媒体、例えばハードディスクあるいはDVD等に保存される。

最も簡単なレベルでは、データは、コンピュータ自身のファイリングシステムの一部として格納できる。この場合、論理的なカレンダー形式でデータを格納するのが有用である。毎日、その日のフォルダが生成され、データが時間毎にフォルダに格納される（UTC時間基準を使用することは、サマータイムによる時間の变化に関連する予測不能の变化を避けることができる）。

高いレベルでは、データベース自体が、記憶システムを備えて、独自の記憶方針の観点により、格納されるデータ要素を保存してもよい。

ストレージメカニズムは、使用されるクエリーシステムから独立しており、データベースインタフェースは、いずれのストレージメカニズムであっても、またデータがいつ格納されたかに拘わらず、要求されているデータへのアクセスを提供すべきである。

【0221】

媒体。コンピュータストレージ媒体は、非常に多種である。最も簡単な分類は、取り外

10

20

30

40

50

し可能な媒体および取り外し不可能なある。取り外し不可能な媒体の例は、ハードディスクであるが、いくつかのハードディスクは、取り外しができる。

実用的な違いは、取り外し可能な媒体は、独自のデータベースを保持しており、これにより、これらの媒体は、取り外し可能なだけでなく持ち運びができる。この方法により取り外し可能な媒体を管理することは、常に容易であるとは限らず、これは、使用されるデータベース次第であり、媒体がこのような容易性を備えているかによる。また、取り外し可能な媒体は、このデータがどうやって、いつ、どこで取得されたのかを示すオーディットのコピーを保持する必要がある。

【0222】

データ検索。

10

【0223】

図19は、データ検索および分析サイクルのステップを示している。ユーザのクエリーに応じて、クエリーに合致する要約データが検索される。成功したヒットにより、イベントが生成され、ユーザに返却されるイベントリストに追加される。主要な画像データは、ユーザがリスト内のイベントを調べるのを望むまで処理されない。図18は、主要な格納データが、どのように要約データに関連付けられるかを示している。

【0224】

存在しているものに基づいて、ユーザは、許容可能なイベントリストが見つかるまで検索の精度を高めることができる。選択されたイベントリストは、異なるストレージフォーマットに変換でき、注釈を付すことができ、パッケージでき、将来に使用するためにエク

20

【0225】

クエリー。

【0226】

検索基準。この種のストレージシステムは、ある実施例では、少なくとも2種類のデータ検索を可能にする。

時間および日付による検索：ユーザは、選択されたビデオストリームから、ある瞬間に取得されたデータを要求できる。要約データ内で特定の時間近くに発生したイベントが存在する場合、このイベントはユーザに通知される。

イベントあるいはオブジェクトの検索：ユーザは、選択されたビデオストリーム内のシーンの領域と、特定のイベントが発生する検索時間間隔を指定する。この時間間隔の要約データが検索され、発見されたイベントがユーザに通知される。検索は非常に高速であるため（数週間のデータは1分以内に検索できる）、ユーザは、非常に長い期間を効率的に検索できる。

30

要約データ内で発見されるイベントは、事前に記録した選択基準では予測されないことに留意すべきである。

【0227】

複数のストリームの検索。複数のストリームからの要約データリストは、ユーザが設定したロジックに基づいて形成し、組み合わせることができる。このロジックを可能にするメカニズムは、ユーザインタフェース次第であり、検索は単に、要求された総てのストリー

40

ムにおける総てのヒットのリストを提供し、ユーザが設定する論理的な基準に基づいてこれらを組み合わせる。

例えば、ユーザは、ユーザの検索ストリームの一つにおけるヒットに応じて、他のビデオストリーム何が発生しているのかを調べることができる。ユーザは、ヒットしたストリームだけを同時にあるいは所定に時間間隔内で確認できる。ユーザは、他のストリームで確認されたヒットに付随する一のストリームのヒットを調べることができる。

【0228】

イベント - 成功したクエリーの結果。成功したクエリーの結果は、ユーザが調査および評価可能なムービークリップ表示であるべきである。ムービークリップは、ユーザが評価できるように、十分な数のビデオフレームを表示すべきである。クエリーが複数のビデオ

50

ストリームを含む場合、ディスプレイは、これらのストリームの同期化されたビデオリプレイを含むべきである。

ここで利用される技術は、要約データにおける成功したヒットのリストを生成し、これらを他のフレームとともに小さなムービーあるいは「イベント」にパッケージすることである。ユーザは、イベントのみを確認し、ユーザが要求するまで個々のフレームを確認しない。

#### 【0229】

要約データ検索。

#### 【0230】

ヒット。要約データの検索は、特定の特徴の画像のシーケンスを検索することである。ここでの利点は、通常、データが単一のビットプレーンであり、我々は、オンにされたビットを探すために、ユーザが指定した領域だけを検索することである。これは、要約データマップが適切に符号化された場合に更なる高速化が可能な非常に高速な処理である。ヒットは、複数のビデオストリームから発生してもよく、クエリーにより設定されたロジックにより、複数のストリーム検索の結果が組み合わせられる。ヒットは、ブロックスコア全体、モーションの方向あるいはサイズなどの要約データから直接的あるいは間接的に利用可能な他の様々な属性の値に基づいて変更してもよい。

#### 【0231】

ディスプレイ。要約データセット内でヒットを発見した場合、要約データからのヒットは、表示可能なイベントに組み込まれるべきである。表示および評価に関する2つの選択肢がある。(1):トレイラが格納された場合に、トレイラを表示すること。(2):十国して完全なデータを取得すること。

#### 【0232】

スピード。要約データの検索は、分析が既の実施されているため非常に高速である。さらに、要約データセットのサイズは通常、オリジナルデータよりも数桁小さい。検索のうち最も遅いものは、実際に、ストレージ媒体のデータにアクセスすることである。これは特に、ストレージ媒体がDVD(アクセススピードはおおよそ1秒当たり10Mバイトである)の場合に当てはまり、この場合、要約データベース全体を頻繁にメモリにキャッシュするのが有用である。ユーザインタフェースのインテリジェントマスキングは、容易にこれを実行でき、最初の検索はデータをリードする時間であるが、以降の検索はほぼ一瞬である。

ネットワークを介した検索は、要約データが高速のローカルアクセスによりハードディスクに保持され、結果のみがクライアントに送信されるため非常に効率的である。

#### 【0233】

関連するデータの検索。

#### 【0234】

イベントの定義および生成。イベントは、1以上のデータ資源から連続したデータフレームを収集したものである。この収集を実施する少なくとも1つのフレームの一つ、キーフレームは、要約データに対するユーザのクエリーとして定式化された所定の基準を満たすであろう。このクエリーは、時間、位置、ある領域の色、動作スピード等に属性に関連してもよい。クエリーに対する成功した結果を「ヒット」を称する。

この処理の一の実施例を考えると、この処理で単一の「ヒット」があった場合、ユーザは、アクションを評価するためにヒットの数秒前および数秒後のビデオを調べたいであろう。2以上のヒットが数秒以内に発生した場合、これらは組み合わせられ、より長いイベントクリップが提供される。したがって、この実施例では、ユーザにより指定される前後のヒット時間の合計よりもヒットの間隔が短い場合、連続的なヒットは同一のクリップに組み込まれる。

あるデータストリームからの単一のキーフレームにより、複数のフレームをカバーするイベントを表わすことは可能であり、このように、キーフレームに関連する総てのデータストリームは相互参照できる。イベントは、(前後のアラーム画像シーケンスなどの)キー

フレームの基準を満たさないキーフレームの前後の複数のデータフレームを備えてもよい。

図20は、どのようにデータが取得され、処理され、格納され、および検索されるかを示している。クエリーに応じて、キーフレームが発見され、これらのキーフレームに関するイベントが生成される。

#### 【0235】

イベントクリップの生成。要約データの各フレームは親フレームに関連し、これは、(ウェーブレット圧縮された)オリジナルビデオデータ内の親フレームから取得される。要約データ内のヒットにより定義されるようにイベント内で参照されるフレームは、ウェーブレット圧縮されたデータストリームから取り出される。これらは、認証され、(必要に応じて)復号され、圧縮される。その後、これらは、表示に適した内的なデータフォーマットに変換される。

データフォーマットは、ユーザのコンピュータ上で表示され、あるいはTVモニタ上で表示するためのエンコーダチップあるいはグラフィックカードによりアナログCCTVビデオフォーマットに変換される場合、(DIBあるいはJPGなどの)コンピュータフォーマットでもよい。

#### 【0236】

イベント分析。要約データヒットのためのオリジナルビデオフレームが取得されると、これらは、要約データに含まれていなかった他の基準を満たすかどうか分析できる。したがって、要約データは、処理時間における演算資源の制限により、オブジェクトを人々、動物、あるいは車両に分類しない。この分類は、これらのストリームに利用可能な要約データを組み合わせることにより、また格納された画像から実施できる。

#### 【0237】

オーディオデータの追加。イベントが再生され、あるいはエクスポートされた場合、シーケンスに付随するオーディオチャネルへのアクセスが必要とされることがある。

オーディオチャネルは、この説明の観点によると、単に別のデータストリームであり、他のストリームと同じ方法により、アクセスおよび表示される。

#### 【0238】

ワークフロー。

#### 【0239】

データアクセスおよびバリデーション。データが符号化された場合、ユーザインタフェースは、データを表示する前に、データを復号化する権限を要求するだろう。同じコンピュータに記録された総てのデータは、同じユーザアクセスコードを備えている。異なるストリームは、異なるセキュリティレベルを有する場合、追加のストリームアクセスコードを備えてもよい。

データバリデーションは、データバリデーションコードが画像データ上で生成されたデータチェック式のほとんど固有の結果であるため、復号化と同時にされる(コードが、限られたビットを有するため、我々は「ほとんど固有」と記す。したがって、天文学的に起こりそうもないが、2つの画像が同じコードを持つこともあり得る。)

#### 【0240】

クエリーの繰返しあるいはリファイン。ユーザインタフェースは、問い合わせを繰り返す、問い合わせの精度を高める、あるいは完全に異なるデータストリームの別の問い合わせの結果と一の問い合わせの結果を組み合わせる選択肢を有する。

要約データ内の検索処理は非常に高速であるため、異なるパラメータあるいは異なるロジックによる問い合わせを繰り返すのに費用が抑えられる。これは単に、プログラムの効率の問題である。

#### 【0241】

データエクスポート - オーディット。ユーザが、クエリーを満たすイベントセットを有する場合、他のプログラムにより使用でき、あるいは表示および通知の目的で使用される方法で発見されたこれらのイベントを保存する必要がある。どのように結果が得られたか

10

20

30

40

50

についてのオーディットは、エクスポートに共に発行され、この処理は、必要に応じて再び実行できる（検索の結果を繰り返す可能性は、時折、訴訟事件で要求される）。

【0242】

エクスポート。

【0243】

イベントデータは、多くの標準的なフォーマットのいずれにもエクスポートできる。これらの多くは、Microsoft Windows（商標）ソフトウェア、あるいはLinuxのソフトウェアと互換性がある。多くがMPEG標準（これは、現在のWindowsメディアプレーヤによりサポートされていないが）を基にしている。

【0244】

本発明は図示された実施例により説明されているが、当業者であれば、この実施例の変更例が存在し、これらの変更例は、本発明の目的および範囲内であることは理解できるであろう。したがって、添付した請求項の意図および範囲から逸脱しない範囲で当業者により様々な変更がされるであろう。

【0245】

表記

記号による表記

以下では、明確にするために、様々な種類のデータおよび画像を意味する記号を使用する。

【0246】

データ、画像およびオペレータ

これらの画像あるいはこれらの組み合わせ上で動作するプロセスは、オペレータとして表わされる。したがって、Fが画像フレームを意味し、Nがノイズをフィルタするオペレータを意味する場合、NFは、プロセスの結果を表わし、F - NFは、Fのノイズ成分として認識される剰余を意味する。

連続して動作するオペレータは、右から左に動作すべく採用される。したがって、 $N_1$ および $N_2$ が、画像フレーム上で動作する2つのオペレータである場合、 $N_1 N_2 F$ は、初めに $N_1$ をFに適用し、次に $N_2$ を適用した結果である。

オペレータは線形である必要はなく、また交換する必要がある。すなわち、 $N_1$ および $N_2$ が、画像フレームF上で動作する2つのオペレータである場合、 $N_1 N_2 F$ および $N_2 N_1 F$ は、必ずしも同じであるとは限らない。

フレームFの一般的な時間および空間依存性は、記号 $F(x, t)$ により表わすことができ、ここでxは、時間tにおけるフレームの2次元画像データである。

我々はまた、どのようにこれらの様々な画像が生成され関連するのかを示すために擬似コードを使用する。詳細については別表に記載されている。

【0247】

表記

	説明
$F_n$	シーケンス内のn番目の加工されていないフレーム インデックスnは、過去の負の値から将来の正の値を取る。 n=0は、現在処理されているフレームに関連する。

10

20

30

40

$\overline{F}_n$	シーケンス内でn番目に処理されたフレーム。  上線は、異なる処理結果を表わすために、ドット、ハット、ティルデなどの様々な好適な装飾と置き換えてもよい。
$F_n^S, F_n^M, F_n^D$	時間解像度分析後の、静止した、移動するおよび動的なフレーム構成要素である。
$\Delta_n = F_n - F_{n-1}$	加工されていない2つの連続したフレーム間の違い
$\sigma(D)$	画像データDのヒストグラムの変化
メジアン(D)	画像Dのヒストグラムのメジアン  オペレータメジアンは、平準化(スモース)、フィルタ、マスク、変換、あるいはオペレーションを示す他の単語と等しい。  オペレータ名は、代替的に、S, F, M, Tのようなボールドフェイスシンボルで表わしてもよい。
$M_n = \mathbf{m}(F_n)$	n番目のフレームをマスクしたもの
$G(\underline{x}, t)$	位置 $\underline{x}$ および時間tに依存した通常のデータ
$G^C(\underline{x}, t)$	構成要素の特性を示すインデックスCが付されたGの特定の構成要素  構成要素の特定の例は、ノイズ(C='N')、背景データ(C='B')、及び動的なデータ(C='D')である。

10

20

30

40

表記は非常に多くなり、通常のデータが、我々が特に示したいサイズを有する値のマトリックスにより表示される場合を考える。我々は、内容から推定可能なものは残しつつ、必要な下付文字および上付文字のみを保持する通常の単純化のステップを採用する。

【 0 2 4 8 】

等式の番号付け

等式は2つの数字、等式が記載されている段落への直接的な参照と、その段落中の等式の番号の参照を有する。したがって、( [ 0 0 9 3 ] ) . 3の番号が付された等式は、段

50

落 [ 0 0 9 3 ] の 3 番目の等式である。

【 0 2 4 9 】

用語解説

ビットボローイング

高いレベルの忠実度が求められる画像の一部を、画像の他の部分よりも、高品質で圧縮することができる処理である。実際には、画像の一部のビットを借りて、他の部分をよく表す。これは、保存する前に、ウェーブレット係数を符号化する間に行われる。係数コードの特別なテーブルが、特別に関心のあるこれらの領域のために作られる。より数の多い値のためのより短いコードを使用することができ、この秘訣は、2つのテーブルを保持することである。2つのテーブルに加えて、閾値の2つの値および量子化倍率 ( scaling factor ) の値を保持することも必要である。 10

【 0 2 5 0 】

C D F ( 2 , 2 )

高いパスフィルタの5ポイントおよび低いパスフィルタの3ポイントを利用するために双直交5 - 3ウェーブレットとしても知られているCohen、DaubechiesおよびFeaveauによる双直交ウェーブレット ( bi-orthogonal wavelet ) 大きなクラスの単純なメンバである。

【 0 2 5 1 】

現在の画像

現在の注目されているシーケンス内の画像である。これは通常、ストリーム内で得られた最新の画像であるが、現在の画像の処理が多数の連続した画像に依存する場合は、最後のもの、最後から2番目、あるいは最後からn番目である可能性がある ( これは、我々が時間導関数を評価している場合に起こりうる ) 。 20

【 0 2 5 2 】

D C D

データチェンジディテクション：通常のVMD形式である。また、VMD、ビデオモーションディテクションも参照。

【 0 2 5 3 】

異常ピクセル

ピクセルの過去の歴史の時系列の分析により、ピクセルの過去が示すものに関して例外的な値を有すると思われる画像データ内のピクセルである。異常ピクセルは、各時点における時間的な動き ( time behaviour ) の観点により定義され、これらのピクセルの重要性は、異常なピクセルの空間的なパターンをスコアリングすることにより、別のピクセルに対するこれらの相対的な近接性の観点により評価される。 30

【 0 2 5 4 】

D i v X

相対的に高い画質を維持したまま、冗長なビデオセグメントを小さなサイズに圧縮するために人気のあるビデオファイル形式である。D i v Xは、不可逆のMPEG - 4パート2圧縮を利用し、D i v Xコーデックは、完全にMPEG - 4 - Advanced Simple Profileに準拠する。現在、D i v X形式は特許の対象であり、オープンソースではない。D i v Xは、MPEG - 4パート10としても知られている新しいH . 2 6 4 / MPEG - 4 AVCに劣るが、CPUの負担は非常に少ない。 40

一般的には、これは、X v i dとして知られているオープンソース形式により置換されている。

【 0 2 5 5 】

動的な前面

データ収集が動的な前面を有する期間中に、シーン内に入るもしくは残っている、または連続的な動作を実行するシーン内の特徴である ( 静止したおよび静的な背景 ) 。静止した背景、静的な背景も参照。

【 0 2 5 6 】

## イベント

イベントは、1以上データ資源からの連続したデータフレームの集めたものである。これを収集するフレームのうち少なくとも一つは、(時間、位置、ある領域の色、動作のスピードなどの)特定の基準を満たす。一のデータストリームの単一のキーフレームにより、複数のストリームを網羅するイベントを表すことができ、この方法により、キーフレームに関連する総てのデータが相互参照できる。一つのイベントは、(例えば、前後アラーム画像シーケンス内の)キーフレーム基準を満たさないキーフレームの前後の複数のデータフレームを備えてもよい。

ビデオイベントディテクションを参照。

【0257】

10

## GUI

グラフィカルインタフェース。これは、コンピュータ、個人用デジタル補助装置(PDA)、携帯電話等で動作するコンピュータプログラムであり、利用可能なプログラムおよびデータの「ウィンドウ」あるいは「グラフィカル」ビューによりユーザに表示する。ユーザは、マウスおよびキーボードなどのポインティングデバイスを用いてプログラムを制御しデータにアクセスする。GUIは容易性および機能を決め、これらにより、ユーザがプログラムを動作させ、データを処理する。

【0258】

## 画像マスク

画像データ上で特定の処理から保護される画像内の領域である。したがって、マスクは、画像中の特徴のエッジを網羅するように構成されており、平準化処理は、曖昧な特徴を生成しない。

【0259】

20

## 画像テンプレート

現在の画像と、場合によりそれに先行する多数の画像から生成される画像である。このような画像の目的は、画像およびその過去の画像の特定の様子を強調することである。テンプレートの一例は、現在の画像の単なるエッジで構成された画像でもよい。現在の画像を特別に設計されたテンプレートと比較することにより、我々は、画像の特定の様子における変化を分離することができる。

【0260】

30

## マスク

画像マスクは、ある画像の領域のマップであり、これの総てのポイントは、特定の性質を共有する。このマップは、これ自体の画像であるが、むしろ単純化された画像であり、これは、当該画像が通常、画像上のポイントがこの特定の性質を有するか否かを示すためである。二値(YesあるいはNo)のマップは、単一のビットマップで表される。マスクは、顕著な赤色がある画像、特定の方向への動きが存在する画像などの1以上の画像についての特定の情報を要約するのに利用される。したがって、マスクは、マップの情報の内容を決定する属性および値のリストを有するマップである。

1以上のマスクからの情報は、データストリームの要約データを生成するのに利用される。

40

またマスクは、画像の特定の部分がマスクされていない場合に、この画像の特定の部分を破壊する可能性のある処理から、この画像の特定の部分を保護するのに利用される。

要約データを参照。

【0261】

## MPEG

Moving Picture Experts Group: 1988に設立された組織である。この組織は、デジタルオーディオおよびビデオ信号の映像信号の符号化表示の規格を開発の責任がある。この規格により、MPEG-1、MPEG2、MPEG-4およびMP3などのデータファイル形式が作られた。この規格の文書は、自由に利用できるわけではなく、この規格を使用するには、ライセンス契約を必要である。MPEGは、実際にはオープンソース規格で

50



はない。

【0262】

ノイズ

ノイズ構成要素は、シーンの一部を正確に表示しない画像データの一部である。通常、ノイズは、機器の影響により発生し、画像データの明確な評価を阻害する。通常、ノイズ成分は、画像データと相関関係がなくあるいは直交すると考えられているが、ノイズは画像の局所的な特性に直接的に依存するため、必ずしもそうとは限らない。

【0263】

ピラミッド分解

Ma11a1の多重解像度分解の指針に従い $n$ 次元データを縮小された低解像度のデータに連続的に縮尺および分解する。また、高解像度のデータベースセットを前の低解像度のデータベースセットから生成するときに生じるエラーが格納される。この例はウェーブレット変換であるが、総てのピラミッド分解がウェーブレットに基づくのではなく、1次でないピラミッド型メジアン変換 (nonlinear pyramidal median transform) が重要な例である。

10

【0264】

ランダムなカメラモーション

ランダムで制限のあるカメラの動作により、認識される画像シーケンスがぶれ、誤った動作検出が発生する。ランダムなカメラモーションは、体系的なカメラモーションに重ね合わせることができ、この場合、これは、画像アスペクトにおける他の滑らかな変化からランダムにずれているように見える。

20

体系的なカメラモーションを参照。

【0265】

参照画像

画像、できれば人工画像であり、これに対して、現在のシーンで発生した重要なイベントが存在するか否かを決定する。人工画像は、過去に取得された他の画像 (この例は平均である) から形成できる。また、現在分析されている画像に続く画像を含むことも可能である (これが実施される場合は、むしろ好ましい)。テンプレートを参照。

【0266】

シーンシフト

シーンの静止した背景の景色が変化するときのビデオストリームの時間であり、シーンは、シフトの直前および直後に空間的に相関しない。

30

シーンマーカーを参照。

【0267】

シーンマーカー

シーンの重要な変化がある位置のしるしである。このような変化は通常、異なる光景を有するシーケンスを開始する体系的なカメラの動作によるものであり、あるいはシーケンスを提供するカメラの変化によるものである。

シーンシフトを参照。

【0268】

ふるい (シーブ)

「ふるいにかける」という動詞は、「選別する」と同義である。辞書では、「好適に試験するために調べること」、「注意深くチェックおよびソートすること」、および「区別および識別すること」と定義している。ふるい (シーブ) (名詞) は、ふるいを可能にする装置である。本明細書では、名詞は、エラトステネスのふるい法により例示される数学的な概念の意味で使用されており、これは、与えられた数 $N$ 以下の総ての素数を区別および識別するアルゴリズムである。したがって、我々は、データストリーム内の属性を区別および識別できる処理を提供する。

40

選別、空間的なシーブ、時間的なシーブを参照。

【0269】

50

## 選別 (sift)

辞書では、これは「調べること、特に有用あるいは価値のあるものを選び出すこと<選別された証拠> - - しばしば、<アレシボ望遠鏡により拾われた信号を選別する>のように使用される」とある。また、これは「ふるいを利用すること」、「ふるいにより識別するように区別すること」、および「注意深く調べること」を意味する。ふるい(シープ)を参照。

(辞書を使用することにより、「Scaleinvariant feature transform」の頭字語 S I F T と混同してはならず、この S I F T は、画像から特有の特徴を抽出するコンピュータビジョンアルゴリズムであり、オブジェクトあるいはシーンの異なるビュー(例えば、立体視)のマッチングやオブジェクト認識などのタスク用のアルゴリズムで使用される。)

ふるいを参照。

【0270】

## スナップショット

「スナップショット」は、アクションの一つのフレームの小さなサムネイルビューを提供するイベントから取得される単一の画像である。このようなフレームは、トレイラの一部とすることができ、あるいは要約データに保持された特別に生成されたフレームとすることができる。

【0271】

## 空間的なシープ

空間的に変化する信号あるいは一連の信号がある特徴を抽出および保存するアルゴリズムあるいは装置である。ハフ変換は、空間的なシープである。

シープ、時間的なシープを参照。

【0272】

## 静止した背景

固定したシーン要素であって、カメラレスポンス、照度、あるいは移動するオブジェクトによるオクルージョンの変化の値によってのみ変化するシーン要素で構成される。静止した背景は、カメラがパン、ティルト、ズームしているときでも存在してもよい。あるシーンを異なる時間に再び閲覧すると、同じ静的な背景要素が表示される。建物および道は、静止した背景を構成する要素の例である。

静的な背景、動的な前面を参照。

【0273】

## 静的な背景

あるシーンを異なる時間に再び閲覧すると、僅かに移動した様子の同じ静的な要素が表示されるという意味で固定されたシーン要素で構成される。動く木の枝および葉は、静的な背景構成要素の例である。動作は局所的および限定的であり、この時間変化は一時的でもよい。窓の反射はこのカテゴリに属するであろう。

静止した背景、動的な前面を参照。

【0274】

## 要約データ

要約データは、一組のデータ画像で構成されており、それぞれが、要約データが得られたオリジナルデータの特定の様子を要約する。

【0275】

## 体系的なカメラモーション

カメラは、オペレータあるいはプログラムの制御によりパン、ティルトおよびズームする機能を有する。このような場合、我々は、一連のアフィン変換によりモデル化できるシーン内の体系的なシフトを見る。動作が速い場合、連続的なシーンは、互いにあまり関係なくあるいは無関係である。

ランダムなカメラモーションを参照。

【0276】

## テンプレート

10

20

30

40

50

画像、できれば人工画像であり、これに対して、現在のシーンで発生した重要なイベントが存在するか否かを決定する。人工画像は、過去に取得された他の画像（この例は平均である）から形成できる。また、現在分析されている画像に続く画像を含むことも可能である（これが実施される場合は、むしろ好ましい）。参照画像を参照。

#### 【0277】

##### 時間的なシーブ

時間により変化する信号あるいは一連の信号がある特徴を抽出および保存するアルゴリズムあるいは装置である。フィルタの通過帯域は、信号の周波数成分の周波数シーブの選択である。

シーブ、空間的なシーブを参照。

#### 【0278】

##### サムネイル

アクティビティが検出されたシーンを示す小さな画像である。これらの小さな画像は、並列データのストリームあるいは要約データの一部として格納できる。これらは、ムービークリップの素早いブラウジングが必要な場合に、完全な画像の代わりに表示できる。

#### 【0279】

##### トレイラ

イベントを構成するフレームのサンプリングされた小さなバージョン。これらの小さなフレームは、並列データのストリームあるいは要約データの一部として格納できる。これらは、ムービークリップの素早いブラウジングが必要な場合に、完全なデータの代わりに再生できる。トレイラは、サムネイルを集めたものではなく、費用が高くて格納することができない。

#### 【0280】

##### ビデオイベント検出

ビデオイベントは、1以上のビデオデータ資源の連続的なビデオフレームのコレクションである。このコレクションを形成するフレームの少なくとも一つ、キーフレームは、イベントを決定する。連続的なフレームのコレクションは、キーフレームを含む総てのフレームを含むコレクションであり、どれほど大きなキーフレーム間のギャップが異なるイベントを表わしているかを示すための基準が存在する。コレクションは、初めのキーフレームの後で、最後のキーフレームの前の多数のフレームを含んでもよい。これは、単一のビデオフレームの領域のモーションを検出に関連するビデオモーション検出とは対照的である。ビデオモーション検出により検出されたモーションを含むビデオフレームは、多くの場合、ビデオイベントを決定するキーフレームである。

イベント、VMD、ビデオモーション検出も参照。

#### 【0281】

##### ビデオフレーム

ここで使用されているフレームは、単一の画像として表示されるビデオシーケンスの最も小さい時間的な単位として定義される。

#### 【0282】

##### ビデオシーケンス

ここで使用されているビデオシーケンスは、時間により整列した個別のデジタル画像のシーケンスとして定義されており、これは、デジタル電子カメラあるいはコンピュータのグラフィックアートアプリケーションなどのデジタル資源から直接的に生成してもよく、あるいは、テレビ放送あるいは記録媒体により提供される信号などのアナログ信号の視覚的な部分をデジタル変換（デジタル化）することにより生成してもよく、あるいは、映画フィルムをデジタル変換（デジタル化）することにより生成してもよい。

#### 【0283】

##### ビデオモーション検出

ビデオモーション検出：主要な目的の一つは、単に周囲の状況の変化によるものでないシーン内の変化を見つけることである。モーションにはいくつかの種類がある。我々は、

10

20

30

40

50

通常の変化（風の中で動く木など）を侵入による変化（車両など）と区別する。前者のモーションは、このようなモーションがシーン内に限定されており、明らかに反復するという事実により認識される。

【0284】

VMD

ビデオモーション検出を参照。

【0285】

ウェーブレット係数

ウェーブレット変換手段による画像の表示は、画像を正確に再形成するのに利用可能な数字の配列を提供する。変換は、ウェーブレット係数と称される一組の数字を用いて画像ピクセルのグループを処理することにより行われる。多くの種類のウェーブレットが存在し、それぞれが、特定の係数セットにより表わされる。画像圧縮の場合、最大限の圧縮を可能にするこれらの係数セットは有利である。しかしながら、これらの係数により生成されるデータは、より大きな圧縮レベルを得るために、検閲され、近似されるだろう。したがって、この検査および近似にも拘わらず、確実に正確な再形成を提供する係数セットは好適である。多くの議論は、これらの観点双方において、いずれの特定のウェーブレット係数セットが最も良いかに集中している。

10

【0286】

ウェーブレット圧縮

2つの要素が、ウェーブレットデータの大幅な圧縮を可能にする。画像のウェーブレット表現の階層構造は、値がほぼ0である多数の階層的に関連する係数が存在する状態になり易い。係数の閾値化処理は、この階層内の0値の数を増やし、量子化処理は、0値が効果的に表示されることを保証する。したがって、非常に少ない格納領域を使用するより効率的な方法でデータを表示することができる。

20

【0287】

ウェーブレット符号化

ウェーブレット係数が量子化される場合、参照テーブルに格納されたコードにより表現される相対的に少ない値が存在する（ウェーブレット量子化を参照）。コードの数字が、再形成のために検索される。しかしながら、格納の前にコード値を提供するテーブルを符号化することができ、この結果、暗号手段を利用しないプログラムは、画像を再形成できない。

30

【0288】

ウェーブレットカーネル

画像のウェーブレット変換は、常にサイズが減少する画像の階層で構成される。階層のレベル間の倍率は通常、必ずではないが、2の一次元ファクタ（linear factor）の場合、4ピクセルの2×2のブロックが1ピクセルになる。より高い（より大きな）画像は、逆のウェーブレット変換によりこれらから形成されるため、使用される最小レベルを「ウェーブレットカーネル」と称する。

【0289】

ウェーブレット量子化

データのウェーブレット変換は、オリジナルデータを再生成するのに利用可能な一組の数字で構成される。十分なレベルの圧縮を実現すべく、数字を単純化し、僅かな典型的な数字により実際の値を表わすことは有用である。典型的な値が選択される方法は、結果が、再形成されたデータの認識可能な変化を生じさせない。この処理は、本質的に連続的な値のセット（元のウェーブレット係数）を適切な不連続の値のセットに変更するため、量子化と呼ばれる。僅かな不連続な値が、各値を再形成処理の間に発見可能な特定のコードに置き換えてコード化される。このため、値29.6135は、文字「W」で表わすことができ、各「W」は、再生成で29.6135により置き換えられる。コード化は、データの符号化の可能性を開く。

40

【0290】

50

## ウェーブレット閾値化

データのウェーブレット変換は、オリジナルデータを再形成するのに利用可能な一組の数字で構成される。十分なレベルの圧縮を実現すべく、再形成されるデータの認識可能な変化を生じさせることないほど十分に小さな数を切り捨てることは有用である。閾値化は、数が安全に切り捨て可能かどうかの決定がなされる一つの方法である。閾値の選択的な値を決定し、閾値化が行われた場合にデータを利用して何をするのかを決定する多くの方法が存在する。このような方法の一つは、「SURE」(「Stein's Unbiased Risk Estimator」)と称される。

【0291】

## ウェーブレット変換

連続的なデータあるいは画像データの変換であり、これにより、変換されたデータは、オリジナルデータの線形スケールレングス(linear scale length)の半分である。低減されたデータセットは、削減されたデータからオリジナルデータを再生成するのに必要な情報を含む別のデータセットにより保持される。収縮されたデータからオリジナルデータを再生成する能力は、ウェーブレットの重要な特徴である。

【0292】

## X v i D

X v i Dは、無償のオープンソースによるビデオコーデックである。X v i Dは、2001年7月に中止されたOpenDivXソースの後、ボランティアのプログラマにより作られた。1.0.xリリースでは、GNU GPL v2ライセンスは、明示的な地理的な制約なしに利用されるが、X v i Dの適法な利用は、地域の法律により制限されている。X v i Dによりエンコードされたファイルは、CDあるいはDVDに書き込みでき、DivXに準拠したDVDプレーヤで再生できることに留意すべきである。

【0293】

## 関連出願のクロスリファレンス

本出願は、2005年9月1日出願された米国暫定特許出願第60/7,12,810の優先権を主張し、この出願の全体は、当該出願を参照することにより本明細書に組み込まれている。

【図面の簡単な説明】

【0294】

【図1】図1は、一般的な形態における処理のブロック図である。

【図2】図2は、ウェーブレット変換階層である。異なる変換は、異なるレベル間で生じる。

【図3】図3は、4つのポイントのウェーブレットのウェーブレットファミリーを生成する処理である。

【図4】図4は、ウェーブレットファミリーが6ポイントおよびこれより高次の偶数ポイントのウェーブレットに生成される処理である。

【図5】図5は、本発明の実現の個別のステージを示している。

【図6】図6は、データの取得ポイントから、詳細な分析および要約データの生成に十分なリファインされたデータポイントまでのステップを示している。このステップは、カメラモーションから生じるアーティファクトおよび画像ノイズを除去するステップと、静止および静的な背景および動的な前面に分解するステップとを含む。

【図7】図7は、要約データとなる一連のオブジェクトマスクに動的な前面を一時的および部分的にグループ分けする処理を示している。

【図8】図8は、画像データのウェーブレット表現および要約データが圧縮されるデータ格納処理を示している。

【図9】図9は、データの問い合わせおよび検索処理を示している。

【図10】図10は、イベント選択の後の処理を示している。

【図11】図11は、新たに取得した画像の分析の初めのループで実行される処理を示している。

10

20

30

40

50

【図 1 2】図 1 2 は、ピラミッド変換であり、ピラミッドの各層は、オリジナルデータよりも小さくて解像度の低いデータを含んでいる。

【図 1 3】図 1 3 は、初めにウェーブレット  $W_1$  を用いてどのように階層が生成されるか示しており、次に、ウェーブレット  $W_2$  を用いてどのように階層が生成されるか示している。下図は、データが格納される方法を示している。

【図 1 4】図 1 4 は、ウェーブレットのカーネル置換処理である。

【図 1 5】図 1 5 は、一連の画像から抽出された一組のデジタルマスクである。これらのマスクは、後に要約データの一部分になる。

【図 1 6】図 1 6 は、スコアが中央のピクセルに割り当てられた多くの  $3 \times 3$  の図形が示されており（上図）、特定の  $3 \times 3$  のブロックの総合的に逸脱したピクセルのスコアが示されている（下図）。

【図 1 7】図 1 7 は、データ圧縮処理の要素を要約している。

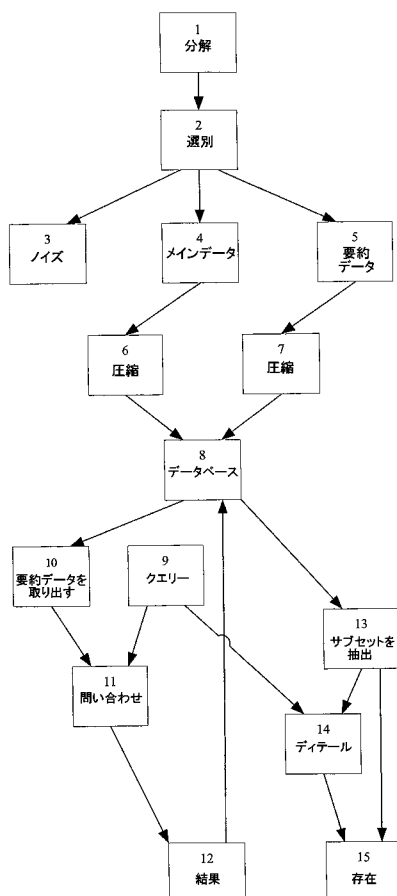
【図 1 8】図 1 8 は、どのように要約データとウェーブレット圧縮データが一对一に対応するのかを示している。

【図 1 9】図 1 9 は、データの検索および分析サイクル内のステップを示している。

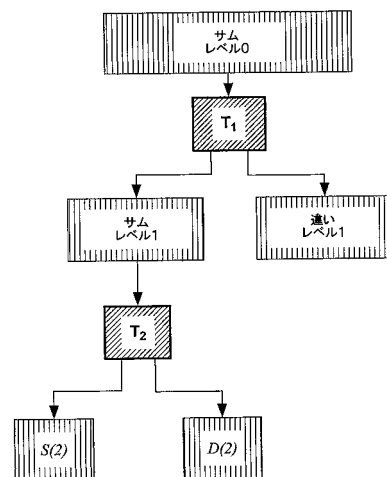
【図 2 0】図 2 0 は、データがどのように取得、処理、格納および検索されるかを示している。

10

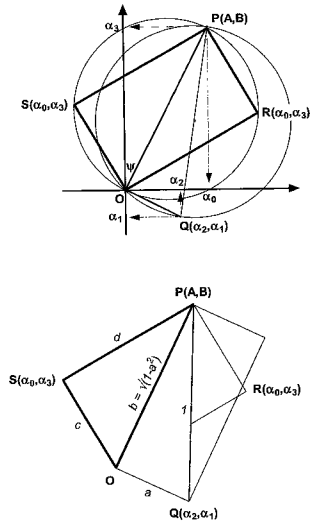
【図 1】



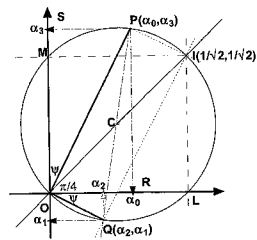
【図 2】



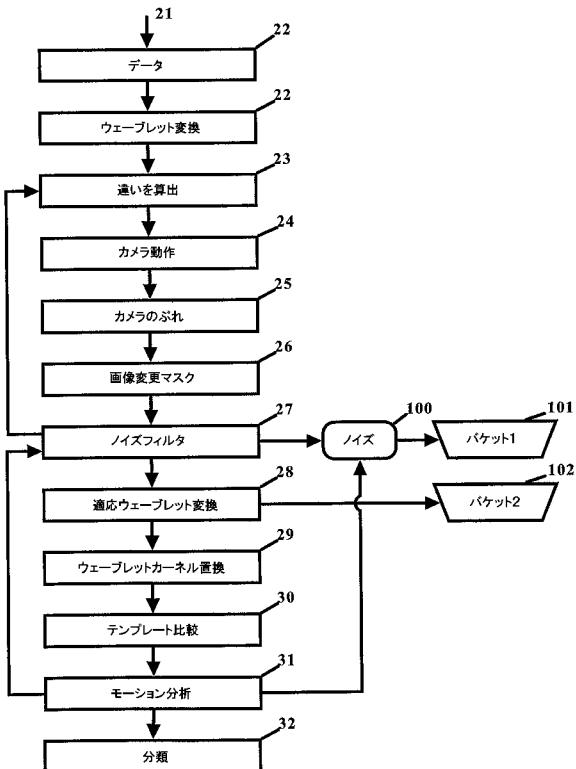
【図3】



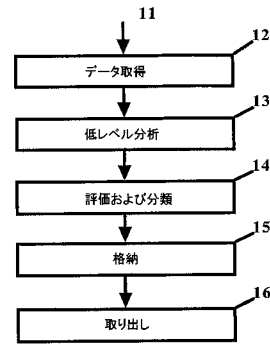
【図4】



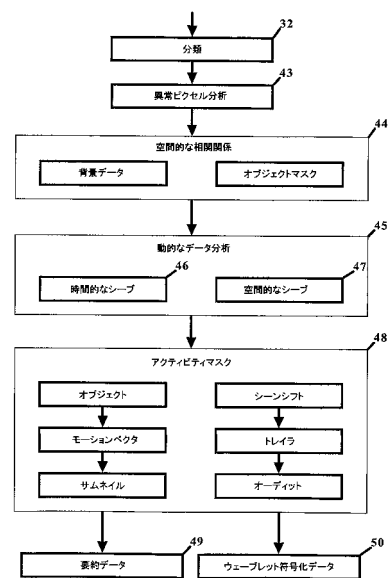
【図5】



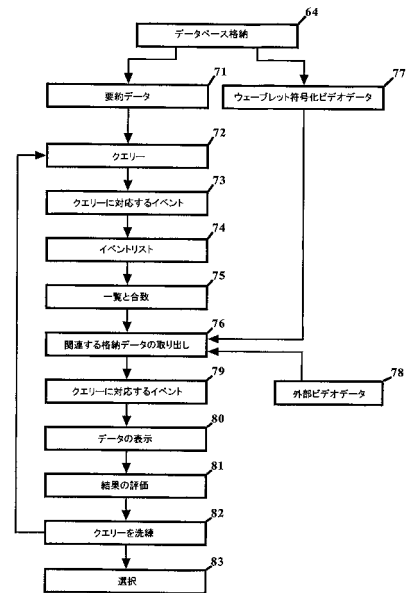
【図6】



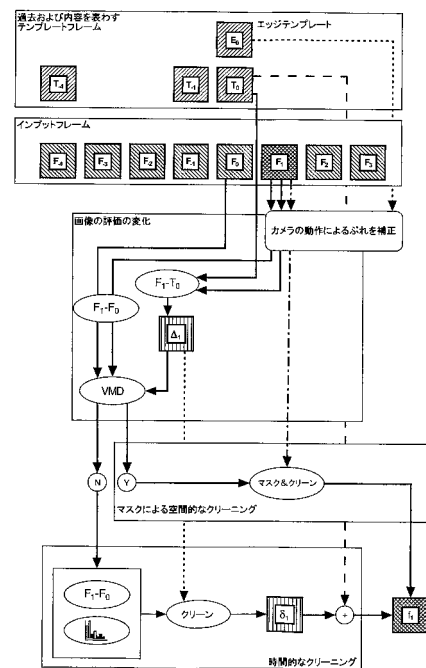
【図7】



【 図 9 】

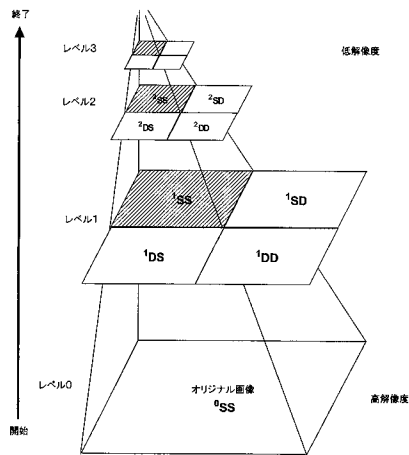


【 図 1 1 】

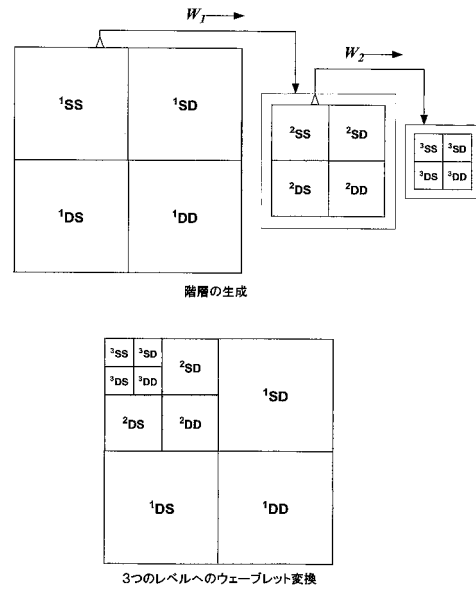




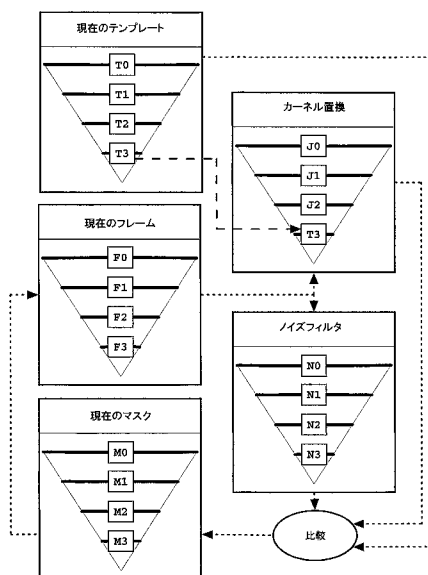
【図 12】



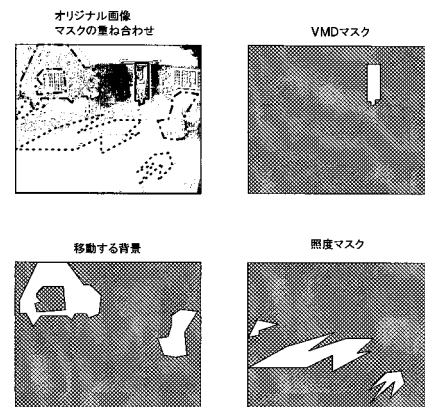
【図 13】



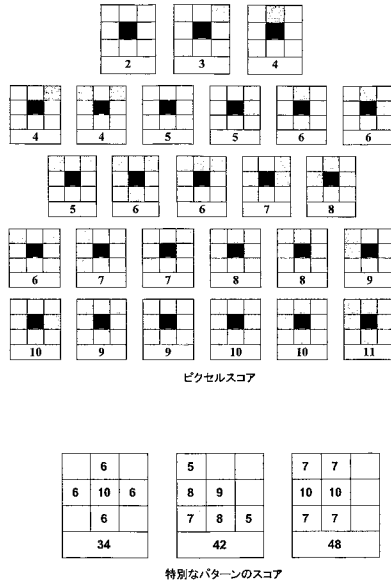
【図 14】



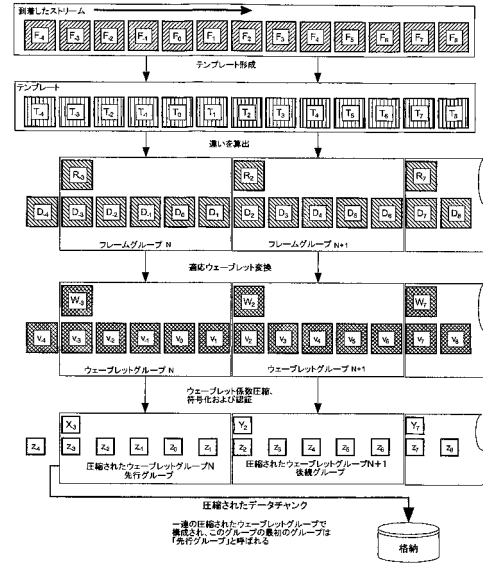
【図 15】



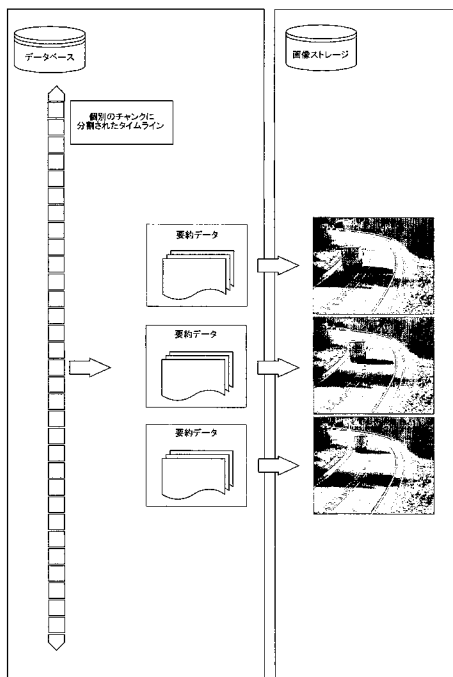
【図 16】



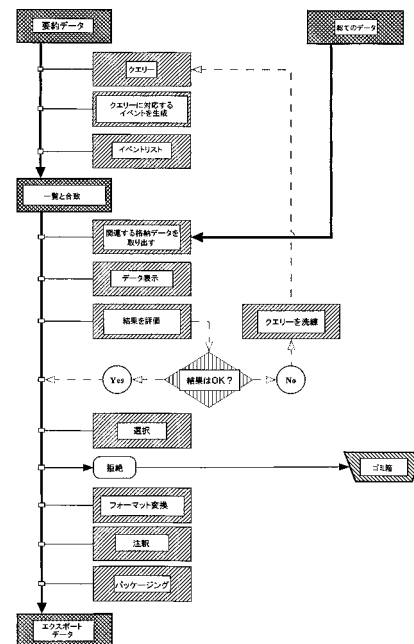
【図 17】



【図 18】



【図 19】



[illegible]

## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No PCT/GB2006/003243
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> INV. G06F17/30 G06T7/20		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) G06F G06T G06K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, INSPEC		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 802 361 A (WANG KATHERINE [US] ET AL) 1 September 1998 (1998-09-01) abstract; figures 2a,3	1-10, 59-64
Y	column 4, line 42 - column 5, line 34 column 7, line 36 - column 11, line 18	1-10, 59-64
X	LETANG J M ET AL: "Motion detection based on a temporal multiscale approach" PROCEEDINGS, 11TH IAPR INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 1992 THE HAGUE, NETHERLANDS 30 AUG.-3 SEPT. 1992, IEEE CS, 30 August 1992 (1992-08-30), pages 65-68, XP010029773 ISBN: 0-8186-2910-X	14,15, 39,42, 47-53
Y	the whole document	1-10, 59-64
----- -/-		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		<input checked="" type="checkbox"/> See patent family annex.
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "Z" document member of the same patent family		
Date of the actual completion of the international search  11 June 2007		Date of mailing of the international search report  15/06/2007
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo.nl, Fax: (+31-70) 340-3016		Authorized officer  LAURENTOWSKI, A

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/GB2006/003243

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SMITH J R ET AL: "AUTOMATED BINARY TEXTURE FEATURE SETS FOR IMAGE RETRIEVAL" 1996 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING - PROCEEDINGS; ATLANTA, MAY 7 - 10, 1996, IEEE, US, vol. CONF. 21, 7 May 1996 (1996-05-07), pages 2239-2242, XP000681679 ISBN: 0-7803-3193-1 the whole document	1-7,9,10
X	JAMES ZE WANG ET AL: "Wavelet-based image indexing techniques with partial sketch retrieval capability" PROCEEDINGS OF IEEE INTERNATIONAL FORUM ON RESEARCH AND TECHNOLOGY ADVANCES IN DIGITAL LIBRARIES, ADL '97, WASHINGTON, USA, 7-9 MAY 1997, IEEE CS, 7 May 1997 (1997-05-07), pages 13-24, XP010231204 ISBN: 0-8186-8010-5 abstract; figures 2,10,11 page 18, paragraph III.2 - page 23, left-hand column, paragraph IV	1-7,9,10
X	MANDAL M K ET AL: "EFFICIENT IMAGE INDEXING TECHNIQUES IN THE JPEG2000 DOMAIN" JOURNAL OF ELECTRONIC IMAGING, SPIE / IS & T, US, vol. 13, no. 1, January 2004 (2004-01), pages 182-190, XP001195298 ISSN: 1017-9909 abstract; figures 2,3,7,8 page 183, left-hand column, paragraph 3 - page 186, paragraph 3.4.1 page 187, left-hand column, paragraph 4	1-5,10
X	HABIBI A: "Introduction to wavelets" MILITARY COMMUNICATIONS CONFERENCE, 1995. MILCOM '95, CONFERENCE RECORD, SAN DIEGO, CA, USA 5-8 NOV. 1995, IEEE, vol. 2, 5 November 1995 (1995-11-05), pages 879-885, XP010154067 ISBN: 0-7803-2489-7 page 880, right-hand column; table 1	11-13
	----- -/-	

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/GB2006/003243

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BROOKS R R ET AL: "RECOGNITION IN THE WAVELET DOMAIN: A SURVEY" JOURNAL OF ELECTRONIC IMAGING, SPIE / IS&T, US, vol. 10, no. 3, July 2001 (2001-07), pages 757-784, XP001086913 ISSN: 1017-9909	14,15
A	pages 769-770, paragraph 3.3.1; figures 15,16	1-13, 27-39, 42, 45-53, 59-64
X	TOREYIN B U ET AL: "Moving object detection in wavelet compressed video" SIGNAL PROCESSING, IMAGE COMMUNICATION, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 20, no. 3, March 2005 (2005-03), pages 255-264, XP004740778 ISSN: 0923-5965	14-16, 27-39, 42,45-53
Y	abstract	1-10, 59-64
	pages 255-256, paragraph 1; figure 1 pages 257-259, paragraph 3	
A	JASPERS E G T ET AL: "Candela-Storage, Analysis and Retrieval of Video Content in Distributed Systems: Real-Time Video Surveillance and Retrieval" MULTIMEDIA AND EXPO, 2005. ICME 2005. IEEE INTERNATIONAL CONFERENCE ON AMSTERDAM, THE NETHERLANDS 06-06 JULY 2005, PISCATAWAY, NJ, USA, IEEE, 6 July 2005 (2005-07-06), pages 1553-1556, XP010843967 ISBN: 0-7803-9331-7 figures 1,6 page 1555, left-hand column, lines 9-15 page 1555 - page 1556, paragraph 3	1-10, 59-64
A	HAMPAPUR A ET AL: "Smart surveillance: applications, technologies and implications" PROCEEDINGS OF THE JOINT INTERNATIONAL CONFERENCE ON INFORMATION, COMMUNICATIONS AND SIGNAL PROCESSING AND THE PACIFIC RIM CONFERENCE ON MULTIMEDIA, SINGAPORE, IEEE, vol. 2, 15 December 2003 (2003-12-15), pages 1133-1138, XP010702808 ISBN: 0-7803-8185-8 page 1133, paragraph 2 - page 1137, paragraph 4.3	1-10, 14-16, 27-39, 42, 45-53, 59-64

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/GB2006/003243**Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:  
1-16, 27-39, 42, 45-53, 59-64
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.

International Application No. PCT/GB2006/003243

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

## 1. claims: 1-10, 59-64

The first group of claims appears to describe a method for interrogating or searching a body of sequential digitised data using the following steps:

- (a) decompose data using pyramidal decomposition;
- (b) apply a sifting process to separate information about data attributes (synoptic data);
- (c) store the data and the synoptic data with an index;
- (d) set up the interrogation or search criteria;
- (e) retrieve synoptic data;
- (f) apply the interrogation or search criteria to the retrieved synoptic data.

The problem to be solved by the subject matter of these claims seems to be how to provide an efficient retrieval method for digitised data.

---

## 2. claims: 11-13

The second group of claims appears to describe a method for aiding the computation of wavelets for applications using pyramidal decomposition comprising the following steps:

- (a) parameterising families of even-point wavelets using a continuous variable;
- (b) using the variable to generate sets of wavelet coefficients.

The problem to be solved by the subject matter of these claims seems to be how to provide a mathematical method to compute wavelet transformations.

---

## 3. claims: 14-16, 27-39, 42, 45-53

The third group of claims appears to describe a method for processing a sequence of digitised data using pyramidal decomposition by wavelets, wherein a data-point by data-point difference between the wavelet transforms of an image in the sequence and the wavelet transforms of a reference image is computed.

The problem to be solved by the subject matter of these claims seems to be how to provide a method for motion detection in image sequences/video.

---

## 4. claims: 17-20



International Application No. PCT/GB2006/003243

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

The fourth group of claims appears to describe a method for processing a sequence of digitised data using pyramidal decomposition by wavelets, wherein on a later iteration a specific process is used to eliminate the frame differences due to changes in illumination.

The problem to be solved by the subject matter of these claims seems to be how to provide a video processing method compensating for changes/fluctuations in illumination/light intensity.

---

**5. claims: 21-25, 43**

The fifth group of claims appears to describe a method for processing a sequence of digitised data using pyramidal decomposition by wavelets, used to calculate the systemic camera movement or to separate the image into a noise component and clean component.

The problem to be solved by the subject matter of these claims seems to be how to provide a method for camera movement compensation and/or video de-noising.

---

**6. claims: 26, 40-41, 44, 54-58**

International Application No. PCT/GB2006/003243

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

The sixth group of claims appears to describe a method for processing a sequence of digitised data using pyramidal decomposition by wavelets, wherein the adaptively coded wavelet data is compressed first by a process of locally feature dependent adaptive threshold and quantization to reduce the bit-rate, and then an encoding of the resulting coefficients for efficient storage.

The problem to be solved by the subject matter of these claims seems to be how to provide a method for efficient storage, compression and encoding of data.

\*\*\*\*\*

The application relates to a plurality of inventions, or groups of inventions, in the sense of Rule 13.1 PCT. They have initially been divided into three and further, 'a posteriori', into six inventions, as defined above, because the prior art gave evidence of a further lack of unity within the claims 14-64. No further invitation to pay further additional fees will be issued. This is because Article 17(3)(a) PCT stipulates that the ISA shall establish the International Search Report on those parts of the international application which relate to the invention first mentioned in the claims ('main invention') and for those parts which relate to inventions in respect of which the additional fees were paid. Neither the PCT nor the PCT guidelines provide a legal basis for further invitations to pay further additional search fees (W17/00, point 11 and W1/97, points 11-16). The applicant has paid three search fees altogether. Therefore, only claims 1-16, 27-39, 42, 45-53, and 59-64, corresponding to the first three inventions (groups of inventions) defined above, have been searched by ISA.

For a detailed reasoned statement concerning the non-unity finding see the Written Opinion of the ISA (WOISA).

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2006/003243

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5802361	A	01-09-1998	NONE

## フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(特許庁注：以下のものは登録商標)

1 . L i n u x

(72)発明者 ジョーンズ, バーナード

イギリス ケント ティーエヌ 1 3 3 アールエス, セブンオークス, ノールウェイ 2 3

Fターム(参考) 5B075 ND12 ND14 ND20 NS01 UU40

5C052 AA01 AB09 AC08 CC11 DD08

5C053 FA23 GB06 GB21 HA29

5C059 MA24 MA31

## 【要約の続き】

にアクセスするのに利用されるマーカーが提供される。この方法により分析を行う実質的な効果は、従来の方法で分析すると数日あるいは数週間を要する膨大な記録されたデジタルデータが、数秒あるいは数分で分析できることである。また、本発明は、連続したパラメータ化されたウェーブレット群の処理に関する。多くのウェーブレットは、8ビットあるいは16ビット表現内で表現できる。また、本発明は、適応ウェーブレットを利用して周囲の状況の変化に強い情報を抽出する処理、局所的な適応量子化および閾値化スキームを利用してデータ圧縮を実行する処理、およびポストレコーディング分析を実行する処理に関連する。

【選択図】なし