

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-9599
(P2010-9599A)

(43) 公開日 平成22年1月14日(2010.1.14)

(51) Int.Cl. F I テーマコード(参考)
G06T 5/20 (2006.01) G06T 5/20 A 5B057

審査請求 未請求 請求項の数 4 O L (全 31 頁)

(21) 出願番号 特願2009-150284 (P2009-150284)
(22) 出願日 平成21年6月24日(2009.6.24)
(31) 優先権主張番号 12/147,867
(32) 優先日 平成20年6月27日(2008.6.27)
(33) 優先権主張国 米国(US)

(71) 出願人 502096543
パロ・アルト・リサーチ・センター・イン
コーポレーテッド
Palo Alto Research
Center Incorporated
アメリカ合衆国、カリフォルニア州 94
304、パロ・アルト、コヨーテ・ヒル・
ロード 3333
(74) 代理人 100075258
弁理士 吉田 研二
(74) 代理人 100096976
弁理士 石田 純
(72) 発明者 ドロン クレッター
アメリカ合衆国 カリフォルニア サン
マテオ タートル ベイ プレイス 20
最終頁に続く

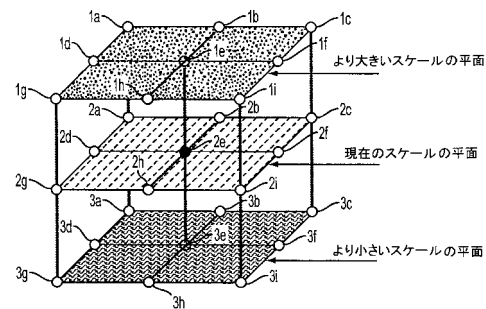
(54) 【発明の名称】 局所化されたスケール空間特性を使用してピクチャイメージ内で安定したキーポイントを検出するシステムおよび方法

(57) 【要約】

【課題】ピクチャイメージ内で安定したキーポイントを検出する。

【解決手段】キーポイントを検出する方法は、入力イメージのインテグラルイメージを計算するステップと、複数のスケールで前記入力イメージのスケール空間ピラミッド層表現を構成するステップであって、各スケールで、特定のフィルタのセットが、前記入力イメージの少なくとも一部の近似を作るために前記入力イメージに適用される、ステップと、スケールおよび空間の単一の関数を形成するために前記フィルタの出力を一緒に組み合わせるステップと、前記単一の関数が局所ピーク値を達成するピクセル位置として各スケールでの安定したキーポイント位置を識別するステップと、前記安定したキーポイント位置をメモリストレージに格納するステップと、を含む。

【選択図】図10



【特許請求の範囲】**【請求項 1】**

入力画像のインテグラルイメージを計算するステップと、
複数のスケールで前記入力画像のスケール空間ピラミッド層表現を構成するステップであって、各スケールで、特定のフィルタのセットが、前記入力画像の少なくとも一部の近似を作るために前記入力画像に適用される、ステップと、
スケールおよび空間の単一の関数を形成するために前記フィルタの出力を一緒に組み合わせるステップと、
前記単一の関数が局所ピーク値を達成するピクセル位置として各スケールでの安定したキーポイント位置を識別するステップと、
前記安定したキーポイント位置をメモリストレージに格納するステップと、
を含むことを特徴とする、局所化されたスケール空間特性を使用してピクチャイメージ内の安定したキーポイントを見つける方法。

10

【請求項 2】

請求項 1 に記載の方法であって、前記フィルタは、異なるサイズの長方形フィルタであることを特徴とする方法。

【請求項 3】

請求項 1 に記載の方法であって、前記構成するステップは、前記フィルタサイズに関わりなく所定の回数の演算を含むことを特徴とする方法。

20

【請求項 4】

電子データ処理システムであって、前記電子データ処理システムのメモリに格納されたソフトウェア動作を処理できる少なくとも 1 つのプロセッサを含み、動作の前記処理が、
入力画像のインテグラルイメージを計算することと、
複数のスケールで前記入力画像のスケール空間ピラミッド層表現を構成することであって、各スケールで、特定のフィルタのセットが、前記入力画像の少なくとも一部の近似を作るために前記入力画像に適用される、構成することと、
スケールおよび空間の単一の関数を形成するために前記フィルタの出力を一緒に組み合わせることと、
前記単一の関数が局所ピーク値を達成するピクセル位置として各スケールでの安定したキーポイント位置を識別することと、
前記安定したキーポイント位置をメモリストレージに格納することと、
を含むことを特徴とする、電子データ処理システム。

30

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、局所化されたスケール空間特性を使用してピクチャイメージ内で安定したキーポイントを検出するシステムおよび方法に関する。

【背景技術】**【0002】**

イメージ（画像）内のキーポイントの検出の技術の一例として、非特許文献 1 及び非特許文献 2 に記載された SIFT (Scalable Invariant Feature Transform) がある。処理速度を改善するために、SIFT では、ディファレンス・オブ・ガウシアン (Difference of Gaussian, DoG, ガウス関数の差分) のイメージピラミッドにより、ラプラシアン・オブ・ガウシアン (Laplacian of Gaussian, LoG) をさらに近似する。入力画像は、連続的にガウシアンカーネルによって平滑化され、ダウンサンプリングされる。ディファレンス・オブ・ガウシアンは、2 つの連続する平滑化されたイメージの差をとることで取得される。したがって、すべての DoG レベルは、平滑化およびサブサンプリングの組み合わせにより構成される。空間次元およびスケールのイメージピラミッド表現における局所的な 3D (3次元) の最大値は、キーポイントの位置およびスケールを決定する。

40

50

【先行技術文献】

【非特許文献】

【0003】

【非特許文献1】ロウエ (D. G. Lowe) , 「Distinctive Image Features from Scale-Invariant Keypoints」, International Journal of Computer Vision, 2004年, 第60巻 (Vol. 60), 第2号 (No. 2), p. 91 - 110

【非特許文献2】ブラウン (Matthew Brown) およびロウエ (D.G. Lowe) , 「Invariant Features from Interest Point Groups」, British Machine Vision Conference, BMVC 2002, カーディフ (Cardiff) , ウェールズ (Wales) , 2002年9月, p. 656 - 665

10

【発明の概要】

【発明が解決しようとする課題】

【0004】

キーポイントを検出する技術において特別な問題は、ノイズに対する頑健性 (robustness) の程度と、回転、スケール、および、キーポイントの検出の実行が要求されるときに生じる他の一般的なイメージ劣化による変化の程度である。

【0005】

従来技術では、かなり多くの計算量を必要とし、これは全体の性能を制限する。

【課題を解決するための手段】

【0006】

本発明の一態様の方法は、入力イメージのインテグラルイメージを計算するステップと、複数のスケールで前記入力イメージのスケール空間ピラミッド層表現を構成するステップであって、各スケールで、特定のフィルタのセットが、前記入力イメージの少なくとも一部の近似を作るために前記入力イメージに適用される、ステップと、スケールおよび空間の単一の関数を形成するために前記フィルタの出力を一緒に組み合わせるステップと、前記単一の関数が局所ピーク値を達成するピクセル位置として各スケールでの安定したキーポイント位置を識別するステップと、前記安定したキーポイント位置をメモリストレージに格納するステップと、を含むことを特徴とする、局所化されたスケール空間特性を使用してピクチャイメージ内の安定したキーポイントを見つける方法である。前記フィルタは、異なるサイズの長方形フィルタであってよい。前記構成するステップは、前記フィルタサイズに関わりなく所定の回数の演算を含んでいてよい。

20

30

【0007】

また、本発明の1つの態様のシステムは、電子データ処理システムであって、前記電子データ処理システムのメモリに格納されたソフトウェア動作を処理できる少なくとも1つのプロセッサを含み、動作の前記処理が、入力イメージのインテグラルイメージを計算することと、複数のスケールで前記入力イメージのスケール空間ピラミッド層表現を構成することと、各スケールで、特定のフィルタのセットが、前記入力イメージの少なくとも一部の近似を作るために前記入力イメージに適用される、構成することと、スケールおよび空間の単一の関数を形成するために前記フィルタの出力を一緒に組み合わせることと、前記単一の関数が局所ピーク値を達成するピクセル位置として各スケールでの安定したキーポイント位置を識別することと、前記安定したキーポイント位置をメモリストレージに格納することと、を含むことを特徴とする、電子データ処理システムである。

40

【図面の簡単な説明】

【0008】

【図1】本発明の実施形態の概念を実施できる環境を示す図である。

【図2A】安定したキーポイント識別動作の応用を表し、スクリーンに表示されたクエリカメライメージを示す図である。

【図2B】安定したキーポイント識別動作の応用を表し、スクリーンに表示されたクエリカメライメージにおける識別されたキーポイントの位置を示す図である。

【図2C】安定したキーポイント識別動作の応用を示し、カメライメージの印刷された版

50

(バージョン)を表す図である。

【図 2 D】安定したキーポイント識別動作の応用を示し、カメライメージの印刷された版において定義されたキーポイントを示す図である。

【図 2 E】安定したキーポイント動作の応用のさらなる例を示し、ターゲットイメージを示す図である。

【図 2 F】安定したキーポイント動作の応用のさらなる例を示し、ターゲットイメージのキーポイントを示す図である。

【図 3】インテグラルイメージ (integral image) を計算する効率的な方法を示す図である。

【図 4 A】3つの基本的なフィルタタイプのうちの水平フィルタを示す図である。

10

【図 4 B】3つの基本的なフィルタタイプのうちの垂直フィルタを示す図である。

【図 4 C】3つの基本的なフィルタタイプのうちの対角フィルタを示す図である。

【図 5 A】単位分散ガウス関数を示す図である。

【図 5 B】水平軸 X に関する単位分散ガウス関数の二次偏導関数を示す図である。

【図 5 C】垂直軸 Y に関する単位分散ガウス関数の二次偏導関数を示す図である。

【図 5 D】X と Y の両方のカスケードされた一次偏導関数に関する、対角線方向に関する単位分散ガウス関数二次偏導関数を示す図である。

【図 6 A】図 4 A のフィルタタイプの単純化された版を示す図である。

【図 6 B】図 4 B のフィルタタイプの単純化された版を示す図である。

【図 6 C】図 4 C のフィルタタイプの単純化された版を示す図である。

20

【図 7 A】スケール空間ピラミッドを作成する、SIFT のDOG法を示す図である。

【図 7 B】スケール空間ピラミッドを作成する、SIFT のDOG法を示す図である。

【図 8 A】水平フィルタのスケール空間ピラミッドを構成する方法を示す図である。

【図 8 B】垂直フィルタのスケール空間ピラミッドを構成する方法を示す図である。

【図 8 C】対角フィルタのスケール空間ピラミッドを構成する方法を示す図である。

【図 9】スケール空間ピラミッド層を作成するのに使用されるフィルタのセットを示す図である。

【図 10】スケール空間領域で $D(x, y, s)$ の極値点を検索する方法を示す図である。

【図 11】現在の層スケールでの対でのピクセルのテストを示す図である。

30

【図 12】3つの連続する層スケールの中の候補ピークに関する対でのピクセルテストを示す図である。

【図 13 A】スケール空間手法を使用するキーポイント識別プロセスを示し、入力イメージを示す図である。

【図 13 B】スケール空間手法を使用するキーポイント識別プロセスを示し、水平フィルタの出力を示す図である。

【図 13 C】スケール空間手法を使用するキーポイント識別プロセスを示し、垂直フィルタの出力を示す図である。

【図 13 D】スケール空間手法を使用するキーポイント識別プロセスを示し、対角フィルタの出力を示す図である。

40

【図 13 E】スケール空間手法を使用するキーポイント識別プロセスを示し、水平フィルタ、垂直フィルタ、及び対角フィルタの組み合わせの出力を示す図である。

【図 14】スケール空間ピラミッド層から安定したキーポイントの位置を識別するプロセスを示す図である。

【図 15】図の最下部に入力イメージがある、スケール空間イメージピラミッドを示す図である。

【発明を実施するための形態】

【0009】

本明細書で説明されるシステムおよび方法は、図 1 に示されたコンピュータネットワークのパラメータ内で働くことができる。コンピュータネットワーク 100 は、一連のワイ

50

ヤ 102 からなるものとすることができ、ワイヤ 102 の多くは、分岐するかワイヤジャンクション 104 で第 3 のワイヤ 106 と接続されてもよく、独立型周辺デバイスに接続されるか、コンピュータ 108 , 109 などの他のデバイスに接続するために周辺機器を通過してもよく、ここで、コンピュータは、周辺デバイスと考えることができる。ネットワークは、カラープリンタ 110 またはカラー以外のプリンタ 112、ならびに少なくともカラーレーザプリンタ 120 , 122 またはカラー以外のレーザプリンタ 124 を組み込んでもよい。ネットワークは、スキャナ 130、またはファクシミリ機 140、写真複写機 150、カラー写真複写機 152、あるいは組合せカラープリンタ/スキャナ/ファクシミリ機 154 を組み込むこともできる。ネットワークは、パーソナルコンピュータおよび/または独立型コンピュータ端末 160、あるいは独立型ハードドライブデータストレージ媒体 164 を含むこともできる。ネットワークは、無線ネットワーク送受信器 170 を含み、少なくとも 1 つのラップトップコンピュータ 172 または複数のラップトップコンピュータ 174 とインターフェースすることもできる。ネットワークは、インターネット、イントラネット、または他の通信ネットワークを含むがこれらに限定されない任意の形のネットワーク 180 に相互接続されてもよい。別の形のネットワークとのインターフェースの使用を介して、本システムおよび方法は、デジタルスチルカメラ 191、デジタルビデオカメラ 192、セル電話機 193、スキャナ 194、携帯情報端末 195、または文書インデクシングシステム 196 を含むがこれらに限定はされない複数の周辺データ取り込みデバイス 190 とインターフェースすることができる。本概念を、単一のデバイスを有するネットワークから数千個以上の接続されたデバイスを含むネットワークにいたる、上記のコンポーネントのさまざまな組合せを有するネットワーク内で実施できることを理解されたい。さらに、上のコンポーネントのうちのさまざまなコンポーネントは、説明される概念を実施するのに有用である可能性がある複数の既知の構成のいずれかに配置されたメモリストレージアリアを有することができる。このストレージエリアは、RAM (Random Access Memory)、ROM (Read Only Memory)、フラッシュメモリ、または本発明の実施形態の概念を組み込んだソフトウェアを保持できる他のメモリタイプとすることができる。他のメモリストレージアリアを、複数のデータベースフォーマットのいずれかのさまざまなデジタルイメージを保持するように構成してもよい。

10

20

30

【0010】

コンピュータなどであるがこれに限定はされない図 1 のコンポーネントのうちのさらなるさまざまなコンポーネントは、そのコンポーネントにロードされるか他の形でそのコンポーネントによってアクセス可能なソフトウェアからの命令を処理するプロセッサを含む。プロセッサを有するコンポーネントのうちのさまざまなコンポーネントが、複数のプロセッサを有してもよく、これによって、命令の処理を複数のプロセッサの間で分割できることを理解されたい。代替案では、単一のプロセッサが、命令を分割するように動作でき、これによって、処理をマルチスレッド環境で行うことができる。

【0011】

図 2 A ~ 図 2 F に、本発明の実施形態のシステムおよび方法を使用する安定したキーポイント識別の成功した適用の例を示す。図 2 A に、スクリーンに表示されたカメライメージを示すが、このカメライメージは、多数のイメージのデータベース内にあるものとする
ことができる、図 2 E のイメージなどのターゲットイメージを識別する試みに使用される
劣化したクエリイメージである。同様に、図 2 C のイメージは、やはり図 2 E に示された
ターゲットイメージから劣化した（たとえば、イメージの一部の上に妨害を有する）印刷
された版のクエリカメライメージのもう 1 つの例である。図 2 A、図 2 C、および図 2 E
のイメージに対して本発明の実施形態のシステムおよび方法を動作させて得られるキー
ポイント位置を、それぞれ図 2 B、図 2 D、および図 2 F に示す。3 つのスケールで得られ
たキーポイント位置が示され、最小スケール位置は「x」でマークされ、中間スケール位
置は「」でマークされ、最大スケール位置は「o」でマークされている。最小スケール
位置は、最小サイズのフィルタの使用によって識別され、中間スケール位置は、最小サイ

40

50

ズのフィルタのスケールアップ版（スケールをより大きくした版）である中間サイズのフィルタの使用によって識別され、最大スケール位置は、中間サイズのフィルタのスケールアップ版である最大サイズのフィルタの使用によって識別される。したがって、図 2 A ~ 図 2 F は、以下でより詳細に開示する本発明の実施形態のシステムおよび方法の動作の結果を示す。

【0012】

イメージの間のイメージ品質、解像度、および変化する照明のかなりの差にもかかわらず、図 2 E のターゲットイメージに対応する図 2 F のキーポイントの多くは、それぞれ図 2 A および図 2 C の劣化したイメージに対応する図 2 B および図 2 D にも見られる。

【0013】

本発明の実施形態のシステムおよび方法の概要は、次の概念を含む。

【0014】

(a) フィルタサイズに関わりなく少数の加算だけで入力イメージ上の任意のサイズの長方形フィルタ形状の出力をすばやく計算できるようにするために入力イメージのインテグラルイメージを計算する。

【0015】

(b) 複数のスケールで入力イメージのスケール空間ピラミッド層表現を構成する。各スケールでは、特定のフィルタのセットを入力イメージに適用して、入力イメージの少なくとも一部の近似を作る。一実施形態で、近似は、ラプラシアン・オブ・ガウシアン (LoG) 関数の形とすることができる。各連続するフィルタは、前のフィルタのスケールアップ版である。フィルタの出力を一緒に組み合わせて、スケールおよび空間の単一の関数 $D(x, y, s)$ を形成する。

【0016】

(c) 安定したキーポイント位置は、各スケールで、関数 $D(x, y, s)$ がスケール空間近傍内で局所的なピーク値（最大値または最小値のいずれか）を達成するピクセル位置として識別される。

【0017】

上のステップを採用することによって、SIFTを含む既存の方法より計算において高速であり、なおかつイメージ内の安定したキーポイントの識別において同等に効率的と思われる手順がもたらされる。

【0018】

本発明の実施形態は、SIFTによって提案される入力イメージの事前の補間を主張しない。

【0019】

本発明の実施形態のシステムおよび方法の詳細を、以下で提供する。

【0020】

I. 1 インテグラルイメージ

インテグラルイメージを計算する効率的な方法を図 3 に示す。処理される入力イメージを、図 3 の上側部分 300 に示し、対応するインテグラルイメージの出力を、図 3 の下側部分 301 に示す。

【0021】

現在のピクセル 302 でのインテグラルイメージの値は、300 の左上角の原点 305 を含み、右下角としての現在のピクセル位置 302 までで現在のピクセル位置 302 を含む長方形区域内のすべてのピクセル値の合計として定義される。効率の一態様は、前に計算された結果を可能な範囲まで活用することである。入力イメージは、一般に、ラスタスキャン順で一時に 1 ラインずつ処理され、1 ライン内では、たとえば左から右へ一時に 1 ピクセルずつ処理される。

【0022】

任意の現在のピクセル位置 302 でのインテグラルイメージの値は、次の 2 つの量すなわち、(1) 前のサブ区域 304 (現在のピクセル位置の正確に 1 つ前のラインのピクセ

10

20

30

40

50

ル位置のインテグラルイメージ値として既に計算済み)と、(2)現在のライン上の最初のピクセルから始まり現在のピクセル値302までで現在のピクセル値302を含む、現在のサブライン303上のピクセル値の和と、を一緒に合計することによって構成することができる。

【0023】

インテグラルイメージは、長方形区域のピクセル値の合計に基づく。この方法は、ここで、入力イメージが共通のラスタスキャン順で得られると仮定する。しかし、行および列の役割は、イメージが通常のラスタスキャン行順ではなく列順で得られる場合に、簡単に逆転することができる。

【0024】

インテグラルイメージを計算する効率的な方法は、ラインアレイ306および307を使用して、正確に1イメージライン前のインテグラルイメージの結果を保持することによる。ラインアレイ内の要素の総数は、入力イメージの単一行内のピクセル数と正確に同一である。このアレイは、一次ストレージ用の環状バッファとして使用されている。

【0025】

ラインの始めから現在のピクセル位置302までのラインアレイピクセル値306は、前に処理されたインテグラルイメージ値および現在のライン上のピクセルの更新されたインテグラルイメージ値を含む。これらの値は、次のイメージラインの後続処理のために格納されている。それと同時に、現在のピクセル位置302から順方向にラインアレイの終りまでのラインアレイピクセル値307は、まだ更新されていない、前のライン上の前のインテグラルイメージ値を含む。現在のピクセル位置302のそれぞれで、前のラインのインテグラルイメージ値が、現在のアレイピクセル位置302から読み取られ、現在のインテグラルイメージ値を計算するために適用され、その後、更新された結果が、ラインアレイ307に書き戻されて、この位置の以前の値を上書きする。次に、このプロセスは、現在のライン上の次のピクセルに進む。このリード(読出し)・モディファイ(更新)・ライト(書込み)サイクルは、ラインの終りに達するまで、ピクセル位置ごとに継続する。

【0026】

各ピクセル位置で、現在のピクセル値302が、まず、図3の最下部に示された加算器309を使用して、前の行の和308に加算される。行の和は、現在の行の始めから現在のピクセル位置302までのピクセル値のランニングサム(running sum)を累算するのに使用される内部レジスタである。行和レジスタは、各ラインの始めに0にリセットされる。加算器309の出力は、行和レジスタに書き戻されて、各ピクセル位置でのその値を更新する。更新された行和値は、現在のサブライン303と現在のピクセル値302との合計と同等である。

【0027】

更新された行和値308は、次に、第2の加算器310を使用して、正確に1ライン戻った前のラインのインテグラルイメージ値307に加算される。というのは、その内容が、まだ現在のピクセル位置302によって更新されていないからである。加算器310の出力は、結果として得られる現在のピクセル位置302でのインテグラルイメージ値である。この結果は、次のラインの後続処理のためにその値を保持するために、現在のピクセル位置302でのラインアレイに書き戻される。したがって、各ラインアレイ要素は、一時に1ピクセルずつ単一のリード・モディファイ・ライトサイクルを受け、これによって、後に続くラインでの連続した再利用のためにインテグラルイメージ値を一時的に保持するためにラインアレイに連続的な1ラインランニングバッファウィンドウを同時に提供する。

【0028】

上で提案した方法を使用すると、インテグラルイメージ全体の計算が、1ピクセルあたり2つの加算だけを必要とする。動作全体を、もちろん、並列に動作する2つの別々の加算器を使用してパイプライン化して、インテグラルイメージ計算を1ピクセルあたり単一

10

20

30

40

50

の二重加算に減らすことができる。したがって、インテグラルイメージは、計算が速く、所定のイメージについて1回だけ計算すればよい。インテグラルイメージを計算した後に、その結果を繰り返して効率的に使用して、以下で説明する形で、さまざまな長方形フィルタ形状の計算を劇的に加速することができる。カスケードされたフィルタリングの伝統的な手法とは異なって、インテグラルイメージは、区域サイズに関わりなく固定された一定の時間で任意の長方形フィルタ区域の出力を計算する能力を提供する。

【0029】

I. 2 個々のスケール空間層の構成

スケール空間表現を作成する伝統的な手法は、イメージピラミッドによるものである。既存の方法の短所は、イメージピラミッドが、通常はガウシアンフィルタを繰り返して適用することと、その後、より大きいスケールに対応するより上のピラミッドレベルを入手するために入力イメージをダウンサンプリングすることによって構成されることである。既存の方法によるイメージピラミッド構成プロセスを、次のように示すことができる。入力イメージを、まず補間して、たとえばSIFT法と同様に、各方向でイメージ寸法を2倍にしてよい。次に、0平均および分散のガウシアンフィルタを、サイズを増やされたイメージに適用して、第1の最小スケールのピラミッド層である出力イメージを作る。

次に、結果の第1イメージ層を、2倍だけダウンサンプリングし、0平均および(2)の同等の分散の第2ガウシアンフィルタを適用して、第2の1つ大きいスケールの第2イメージピラミッド層を作る。各後続層は、前のピラミッド層出力をダウンサンプリングし、徐々により大きくなるガウシアンフィルタを結果のイメージに効果的に適用して、次のピラミッド層を作ることによって作られる。したがって、イメージピラミッドを作るプロセス全体は、順次的性質を有し、各後続空間スケールピラミッド層は、より小さいスケールの前のすべてのピラミッド層に完全に依存し、したがって、すべての他の層と同時に計算することはできない。さらに、各方向でサイズNピクセルの一般的な分離不能正方形ガウシアンフィルタは、計算にオーダー $O(N^2)$ の乗算および加算を必要とし、Oは、演算の回数である。演算の回数は、フィルタ面積サイズに比例し、したがって、各後続層は、増加したフィルタサイズに起因して前の層に対して相対的に計算により長い時間を要し、やはり、後続する層の結果を後のステージでキーポイント抽出に同時に使用可能にするために、後続する層の結果をメモリに格納することが必要になる。フィルタ動作の数は、フィルタが分離可能である場合には $O(N)$ 回の乗算および $O(2N)$ 回程程度の加算に減らすことができるが、それでも、複数のスケールについて計算するのにかなりの時間を要する。

【0030】

さらに、再帰的にフィルタリングし、ダウンサンプリングする必要のゆえに、このプロセス全体が、直列の性質を有し、より高い次数のレベルの計算を、すべてのピラミッドレベルについて同時に得ることはできない。しかし、キーポイント識別プロセスは、ある個数のレベル(通常、少なくとも3つから4つ)の同時使用可能性に依存する。したがって、中間ピラミッドレベルを、後の使用のためにメモリ内に格納しなければならない。

【0031】

既存の手法のように複数スケールイメージピラミッドを作成するために入力イメージを繰り返してフィルタリングし、ダウンサンプリングするのではなく、本システムおよび方法は、水平方向、垂直方向、および対角線方向での入力イメージの局所化可能な特性の直接識別に基づく異なる手法を使用する。さまざまなフィルタ出力を一緒に組み合わせ、入力イメージ内の局所化可能な「しみ様(blob-like)」の特徴および「接合点様(junction-like)」の特徴を識別する。

【0032】

さらに、以前の方法のように、予めフィルタリングされダウンサンプリングされた版の入力イメージの(徐々に小さくなる)出力にガウシアンフィルタを反復的に適用するのではなく、本発明の実施形態のシステムおよび方法は、フィルタをアップサンプリングし、より大きいフィルタ版をオリジナルの入力イメージに直接に適用する。したがって、より

大きいスケールのそれぞれで、入力イメージは一定のままにしながら、以前のフィルタ形状を、たとえばサイズにおいて2倍にすることができる。

【0033】

一見すると、この見た目には良好なスケーリング透視の変化は反直観的に見える。しかし、本システムおよび方法において、固定されたフィルタおよび順次小さくなるイメージサイズを固定されたイメージおよび順次大きくなるフィルタサイズに置換することが、性能の改善をもたらす得ることがわかった。

【0034】

この手法の利益は、より大きいフィルタを、既存の方法のように以前のフィルタ出力を使用して反復的にカスケードする必要なしに、オリジナルの入力イメージに直接に適用できることである。さらに、インテグラルイメージおよび特定のフィルタ形状の使用に起因して、任意のサイズのフィルタを処理するのに、正確に同一の長さの時間を要する。したがって、すべてのフィルタを、そのサイズと独立に、正確に同一の速度で計算することができる。さらに、各長方形フィルタ領域は、そのサイズに関わりなく、インテグラルイメージを使用して計算するのに少数の加算だけを必要とする。

10

【0035】

その結果、提案されるシステムおよび方法は、既存の伝統的な反復手法に対する相対的な性能改善を入手しながら、なおかつイメージ内の堅牢なスケール不変キーポイントの位置の識別において同等に効果的である。

【0036】

図4A～図4Cに、本発明の実施形態によるスケール空間ピラミッドの個々の層を構成するのに使用される3つの基本的なフィルタタイプを示す。

20

【0037】

各スケールで、3つの別々のフィルタすなわち、(a)図4Aの水平フィルタ、(b)図4Bの垂直フィルタ、および(c)図4Cの対角フィルタが入力イメージに適用される。「X」を用いてマークされた各フィルタの中央位置は、その回りにフィルタウィンドウが位置決めされる、入力イメージの現在のピクセル位置である。

【0038】

各フィルタのサイズは、それに適用されるスケールに比例する。各フィルタが、長方形区域のセットからなることに留意されたい。というのは、その意図が、高速計算のためにインテグラルイメージを適用することであるからである。また、図4Bの垂直フィルタが、図4Aの水平フィルタ形状を90度回転した版であることに留意されたい。

30

【0039】

図4Aの水平フィルタ形状は、すべてのピクセル値が所定の極性、たとえば+pをとる中央領域(斜めの線のパターンを用いて図示)と、すべてのピクセル値が反対の極性、たとえば-mをとる2つの外側領域(水平線のパターンを用いて図示)とからなり、ここで、「p」および「m」は、2つのゼロでない整数値である。

【0040】

中央領域が正であり、外側領域が負であるか、または、その反対であるかという2つの領域タイプの間の極性割り当ての意味は、一貫性をもって適用される限り、任意である。

40

【0041】

図4Aで白い領域として図示されたオプションの「ガードバンド」は、領域をお互いから分離するために、反対の領域タイプの間の境界に沿って導入される。この「ガードバンド」の目的は、フィルタ安定性を改善し、極性において変動する可能性がある、領域境界に沿ったノイズのあるピクセルを除外することによって、ノイズに対する感度を最小化することである。「ガードバンド」領域の内側のすべてのピクセル値に、0の値が割り当てられる。

【0042】

したがって、各フィルタピクセルの値は、3つの可能な値{-m, 0, +p}のうちの1つをとる。「p」および「m」の値は、一実施形態では整数値であり、相対領域面積サ

50

イズに基づいて一定の入力イメージの場合の偏りのないフィルタ出力をもたらすように選択される。たとえば、水平フィルタまたは垂直フィルタの場合に、式

$$p / m = b_1 / (c - b_2)$$

を満足するように「 p 」および「 m 」の値の対を選択する。

【0043】

b_1 、 b_2 、および c のすべてが、整数ピクセル番号値なので、「 p 」および「 m 」の対を見つけることが可能である。

【0044】

各フィルタの形状は、4つのパラメータ値{ a 、 b_1 、 b_2 、および c }のセットによってパラメータ化される。これらのパラメータの値は、スケールに比例する。フィルタサイズがスケールアップされると、すべてのパラメータ値が比例して増やされる。ただし、フィルタ形状は比例して正規化されたままに保たれる。また、4つのパラメータ値{ a 、 b_1 、 b_2 、および c }が、必ずしも、水平フィルタまたは垂直フィルタについて対角フィルタと同一である必要がないことに留意されたい。

【0045】

水平フィルタは、水平方向でよく局所化可能な、入力イメージ内の候補の高く狭い「細長い」特徴を検出するように設計される。外側領域に対して相対的により暗いまたはより明るい中央領域などであるがこれに限定はされない特徴極性は、後述するように、フィルタ出力の大きさを使用することによって無視される。フィルタ出力の大きさは、現在のピクセル位置ごとの中央フィルタ領域と外側フィルタ領域との間のコントラスト差に効果的に比例する。

【0046】

図4Aの水平フィルタは、一般に、幅 $2 \times b_1$ および高さ $2 \times c$ の狭く高い中央領域を有し、 $b_1 < c$ である。同様に、この水平フィルタは、組み合わされた幅 $2 \times (a - b_2)$ および高さ $2 \times c$ の2つの外側領域をも有する。幅 $(b_2 - b_1)$ および高さ $2 \times c$ を有するオプションの狭く高い垂直「ガードバンド」を、中央領域と外側領域との間に挿入してもよい。あるいは、パラメータ値 $b_2 = b_1$ をセットすることによって、「ガードバンド」を完全に除去することが可能である。

【0047】

図4Bの垂直フィルタは、垂直方向でよく局所化可能である入力イメージ内の候補の幅広く短い「細長い」特徴を検出するように設計される。垂直フィルタの形状は、図4Aの水平フィルタを90度回転した版である。

【0048】

図4Cの対角フィルタは、対角線方向でよく局所化可能である入力イメージ内の候補の「細長い」特徴を検出するように設計される。対角フィルタの形状は、4つの長方形領域からなり、そのうちの2つは(対角線パターンを用いて図示)、一方の対角線方向に向けられ、他方の2つ(水平線パターンで図示)は、他方の対角線方向に向けられる。すべての領域サイズが、対角フィルタの場合には同一なので、フィルタ応答は、設計によって既に偏りをなくされており、上述の「 p 」および「 m 」ではなく+1または-1の値を、上記で示したように領域タイプに応じてこれらの領域のそれぞれのすべてのピクセル値に割り当てることで十分であることに留意されたい。

【0049】

対角フィルタが、図4Cでその主軸に沿った白い領域として示されたオプションの「ガードバンド」をも含むことに留意されたい。「ガードバンド」領域は、フィルタ安定性およびノイズに対する免疫性を改善するために、対向する領域タイプの境界に沿って導入される。「ガードバンド」領域内のすべてのピクセル値には、0の値が割り当てられる。長方形対角フィルタの一般的なケースについて、「ガードバンド」は、水平軸の回りで $2 \times b_2$ の幅および垂直軸の回りで $2 \times b_1$ の幅を有する。「ガードバンド」の幅は、 $b_1 = b_2$ を選択することによって対称にすることができ、または、 $b_1 = b_2 = 0$ を選択することによって完全に除去することができる。「ガードバンド」区域を含まない各領域タイプの総

10

20

30

40

50

面積は、対角フィルタについて $2 \times (a - b_1) (c - b_2)$ である。

【0050】

パラメータ値 $\{a, b_1, b_2, \text{および } c\}$ の適当な選択を用いると、フィルタの各1つを、それぞれ水平方向、垂直方向、および対角線方向のガウシアンカーネルの二次偏導関数の近似とみなすことができる。より具体的には、図5Aに、単位分散ガウシアンフィルタ関数を示す。これは、より一般的なケースの分散ではなく正規化された分散1を有することを除いて、最小スケールで図2Aの入力イメージに適用される平滑化フィルタである。

【0051】

単位分散ガウス関数の水平、垂直、および対角の二次偏導関数を、それぞれ図5B、図5C、および図5Dに示す。この文脈での対角二次導関数は、水平一次偏導関数と垂直一次偏導関数との積である。平滑化ガウシアン二次導関数は、各ケースで長方形のトライナリ(すなわち、3つの)フィルタ形状を用いて近似される。

10

【0052】

本システムおよび方法は、入力イメージのダウンサンプリングを利用しないので、平滑化および事前フィルタリングの必要が大幅に減らされる。したがって、本手法によれば、図4A~図4Cのフィルタのそれぞれは、スケールが増えるにつれてますます大きくなる区域サイズにまたがる平均化フィルタである。しかし、ある平滑化を、ノイズに対する高められた頑健性を提供し、ブロッキングアーティファクト(blocking artifact)を最小化するために、後続フィルタ位置をわずかにオーバーラップさせることによって実現することができる。1/4から1/2までのフィルタオーバーラップ方式を各方向で使用することが、優れた結果をもたらす。これは、所定のスケールでの次のフィルタ位置を、 $a/2$ と a の間の値だけ水平方向で前進させ、これによって、前のフィルタ位置に部分的にオーバーラップさせることを意味する。ラインの終りに達した時に、次のフィルタ位置も、 $c/2$ と c の間の値だけ垂直方向で下に移動され、前のフィルタ位置との部分的垂直オーバーラップにもつながる。

20

【0053】

この手法の第1の態様は、インテグラルイメージの長方形のフィルタ形状に起因して、フィルタ結果を計算するのに少量の加算だけを要することである。各長方形フィルタ領域は、インテグラルイメージを使用して計算するのに、そのサイズに関わりなく4回の加算だけを必要とする。追加の節約を、隣接領域区域が共通の境界を共有する時に行うことができる。たとえば、図4Aの水平フィルタを計算するには、通常は $3 \times 4 = 12$ 回の加算が必要になる。というのは、このフィルタが1つの中央領域および2つの外側領域を含むからである。しかし、「ガードバンド」領域がない場合には、外側領域をフィルタ区域全体 $[-a, -c]$ から $[a, c]$ に延長することと、余分な外側区域の加算を無にするためにそれ相応に「p」および「m」の値を調整することによって、水平フィルタ全体を12ではなく8クロックだけで計算することができる。「ノイズバンド」領域なしの単純化された水平フィルタ、垂直フィルタ、および対角フィルタの形状を、それぞれ図6A~図6Cに示す。図6A~図6Cの単純化されたフィルタ形状は、図4A~図4Cのオリジナル形状より計算が高速であるが、「ノイズバンド」領域の欠如に起因して、ノイズに対して図4A~図4Cのオリジナル形状ほど頑健(robust)ではない。ノイズの多い入力イメージの状況では、図6A~図6Cの単純化されたフィルタ形状ではなく、図4A~図4Cのオリジナルフィルタ形状を適用することが好ましい場合がある。

30

40

【0054】

本手法の第2の態様は、フィルタを計算するのに要する動作および時間の量が、フィルタサイズに関わりなく正確に同一のままになる固定された定数であることである。たとえば、図4Aの水平フィルタが、サイズにおいて2倍にされる場合に、このフィルタは、計算に同一回数分の加算を要する。というのは、より大きいフィルタ形状が、それでも、比例して調整された角の点を有する3つの長方形領域からなるからである。これは、ガウシアンフィルタを繰り返して使用し、前のフィルタ出力をダウンサンプリングする場合と対照

50

的であり、この場合に、各より高いピラミッドレベルは、追加の計算を必要とし、前のレベルに加えて余分な時間を要する。

【0055】

本手法の第3の態様は、任意の順次ピラミッドレベル処理の除去である。これは、本システムおよび方法が、必ず、どのスケールレベルでもフィルタタイプごとに正確に同一の入力イメージを適用するという事実の直接の結果である。入力イメージをフィルタリングする必要は全くなく、したがって、すべての前のピラミッドレベルの完了時に、後続レベル用の入力を得るために前のレベルの出力をフィルタリングし、ダウンサンプリングする必要がある、1つのピラミッドレベルの順次依存性は、除去される。したがって、あるピラミッドレベルの計算は、他のすべてのピラミッドレベルから完全に分離される。

10

【0056】

本手法の第4の態様は、並列処理および/またはマルチスレッド環境での実行に理想的に適することである。これは、本システムおよび方法が、未知の入力イメージではなくフィルタをスケールリングすることを選択するからである。フィルタ形状は前もってすべてわかっているので、すべてのタイプ、サイズ、およびスケールのスケールされたフィルタ形状のセット全体を、前もって生成し、複数のプロセッサが使用可能な場合には並列に、または別々のスレッドで、入力イメージに適用することができる。ピラミッドレベルの間の任意の依存性の欠如に起因して、スケール空間ピラミッドを作成する処理方式全体を、並列に実行することができる。

20

【0057】

本手法の第5の態様は、複数のプロセッサまたはスレッドの間で計算負荷を平衡化することのたやすさである。これは、ほぼすべてのフィルタが、本手法に従って計算される時に、フィルタサイズに関わりなく計算に正確に同一の時間を要するという特性の結果である。したがって、スケール空間ピラミッドを作成する作業は、それぞれが実行にほぼ同一の時間を要する複数のプロセッサまたはスレッドの間で均等に分割することができ、このプロセスの終わりに、必要な個数のピラミッドレベルが、結果をメモリに格納する必要なしに、および/または不均一な計算負荷に起因するアイドル状態のプロセッサを有することなく、後続のキーポイント識別に使用可能にされる。

【0058】

この手法の第6の態様は、より大きいスケールのさらにより大きいフィルタを有するより高いピラミッドオクターブについて、後続のフィルタの位置の間の距離を、計算負荷をさらに減らすために比例して増やすこともできることである。より高いオクターブについて、対応する「しみ様」および/または「接合点様」の特徴は、さまざまなフィルタ出力を一緒に組み合わせた後に、やはり非常に大きくなり、したがって、詳細なサブピクセル精度でその位置を正確に指摘する必要なしに処理することができる。しかし、望まれる場合に、より高いピラミッドオクターブについて追加の補間方式を導入することができる。というのは、結果として得られる後続のフィルタの位置の間の距離が、大きくなる可能性があるからである。

30

【0059】

I.3 スケール空間ピラミッドの作成

40

スケール変化に対して不変のキーポイントの位置を識別する伝統的手法は、スケール空間として知られるスケールと空間との両方の連続関数を構成し、複数のスケールにまたがって安定した特徴を検索することである。

【0060】

前述のように、スケール空間ピラミッドを生成する1つの既存の方法は、SIFT法に見られる。SIFT法は、代替のディファレンス・オブ・ガウシアン(DoG)関数を用いるラプラシアン・オブ・ガウシアン(LoG)法の近似(最初は、リンドバーグ、「Scale Space Theory: A Basic Tool for Analyzing Structures at Different Scales」、Journal of Applied Statistics, Vol. 21, No. 2,

50

224 ~ 270 ページ、1994 年で提案された) に基づく。

【0061】

スケール空間ピラミッドを作成する S I F T 法の概念を、図 7 A ~ 図 7 B に示す。スケール空間のオクターブごとに (図 7 A は、1 つのスケールを示し、図 7 B は、追加のフィルタリングおよびダウンサンプリングの後のより高いスケールを示す)、初期入力イメージが、ガウシアンフィルタのセットを用いて繰り返して畳み込まれて、図 7 A ~ 7 B の左側に示されたスケール空間イメージのセットを作る。隣接するガウシアンフィルタリングされたイメージは、減算されて、図 7 A ~ 7 B の右側のディファレンス・オブ・ガウシアン $D_1 \sim D_3$ イメージを作る。各オクターブの後に、前のオクターブからのガウシアンイメージが、2 倍ダウンサンプリングされ、この処理が、繰り返される。スケール空間ピラミッドを作成する反復処理の概念を、図 15 に示す。入力イメージは、図 15 の最下部に示され、その上に、特定のスケールのガウシアンフィルタを用いるガウシアンフィルタリングおよびダウンサンプリングの 3 つの連続する反復の後のカスケードされた結果がある。隣接するガウシアンイメージを減算した後に、結果のディファレンス・オブ・ガウシアン $D_1 \sim D_3$ イメージが、識別されたキーポイントを円によって強調表示された状態で図 14 に示されている。

10

【0062】

次では、既存概念からの進歩を開示するが、オリジナルラプラシアンオブガウシアン (L o G) 関数は、上で説明したインテグラルイメージの特性を活用する機会を可能にする S I F T 法の (D o G) と異なる新しい近似に置換される。

20

【0063】

上のセクション I . 2 で説明したように、図 4 A ~ 図 4 C のフィルタは、ガウス関数の二次偏導関数のライナリ近似を提供する。したがって、フィルタ出力を一緒に組み合わせ、ラプラシアン・オブ・ガウシアン (L o G) の近似を提供することができる。ラプラシアン・オブ・ガウシアン (L o G) は、ある形では、(D o G) の近似でもある。本手法が、S I F T とは異なり、近似の精度に応じて異なる結果につながる可能性が高いことに留意されたい。

【0064】

ここで、本概念によるスケール空間差ピラミッドの構成に注意を向ける。

【0065】

図 8 A ~ 図 8 C では、少なくとも 3 つのスケール空間差ピラミッド層が、キーポイント位置を識別するのに必要である。これは、近傍スケールにまたがる安定した特徴を識別するために、現在のスケールと、現在のスケールの下の少なくとも 1 つのスケールおよび現在のスケールの上の少なくとも 1 つのスケールを含む。図 8 A は、水平フィルタリング (たとえば、 $F_{1h} \sim F_{3h}$) を表し、図 8 B は、垂直フィルタリング (たとえば、 $F_{1v} \sim F_{3v}$) を表し、図 8 C は、対角フィルタリング (たとえば、 $F_{1d} \sim F_{3d}$) を表す。図 8 A ~ 8 C に示されているように、入力イメージは、連続してより大きくなるフィルタ F_1 から F_3 までのセットを用いて処理される。その後、水平フィルタ、垂直フィルタ、および対角フィルタからの結果の出力が、一緒に組み合わせられて、キーポイント位置が識別される単一の尺度が形成される。

30

40

【0066】

フィルタ F_1 から F_3 は、前のセクションで説明した形でインテグラルイメージを使用して効率的に計算される。

【0067】

本願の一実施形態で、スケール空間の 1 つのオクターブは、図 8 A ~ 図 8 C に示された 3 つのフィルタの最小限のセットを使用することによって形成される。

【0068】

もう 1 つの実施形態で、追加のますます大きくなるフィルタが、スケール範囲を拡張するために F_3 を超えて適用される。たとえば、 F_4 フィルタ (図示せず) を、図 8 A ~ 8 C に追加することができ、その水平フィルタ、垂直フィルタ、および対角フィルタの出力

50

を一緒に組み合わせて、キーポイント位置を識別するもう1つの尺度層を形成することができる。しかし、1オクターブ内のフィルタの最大個数は、イメージサイズによって制限される。というのは、これが、イメージサイズより大きい特徴を探すために利益をもたらさないからである。したがって、最大のフィルタサイズは、イメージサイズによって決定される。

【0069】

1つのオクターブ内に3つを超える異なる層がある時には、層は、近傍スケールにまたがる特徴安定性を保証するために隣接する3つの層を含む組で処理される。すなわち、安定したキーポイント位置は、まず、最初の3つの層{F1, F2, F3}の内容に基づいて識別され、3つの層の後続セット{F2, F3, F4}からの安定したキーポイント位置が続き、新しい最も高い層{F4}は、最下位の層{F1}を置換し、層{F2, F3}は、両方のセットに共通する。同様に、3つの層の次のセットは、{F3, F4, F5}を含むことができ、以下同様である。しかし、ほとんどの場合に、通常、最初の3つを超える多数の層を使用することは必要ではない。というのは、イメージがサイズにおいて1メガピクセルから2メガピクセルを超えない限り、フィルタサイズおよび対応するイメージ特徴のスケールが、オリジナルイメージサイズによってすばやく制限されるようになるからである。1メガピクセルから2メガピクセルを超える場合には、より多くの層を追加することが有益である可能性がある。

10

【0070】

スケール空間ピラミッドの4層オクターブを作成するのに使用される水平フィルタのセットを、図9に示す。所定のオクターブ内で、各後続フィルタ F_i は、さらに、前のフィルタ $F(i-1)$ に対して相対的に係数「s」だけスケールアップされ、または、最高スケールの最初のフィルタ F_1 に対して相対的に合計($s \times i$)だけスケールアップされる。水平フィルタだけが、図9に示されていることに留意されたい。垂直フィルタおよび対角フィルタは、正確に同じ態様でスケールアップされる。

20

【0071】

SIFT法のカスケードリングフィルタの手法では、キーポイント位置を第1オクターブについて識別した後に、結果のイメージが、もう一度ダウンサンプリングされ、フィルタリングされ、隣接フィルタ出力が減算され、その結果が、より高いスケールでの追加のキーポイント位置の検出に使用される。このプロセスは、追加のスケールごとに繰り返される。

30

【0072】

しかし、本システムおよび方法では、フィルタリングされたイメージをダウンサンプリングするのではなく、基本的な $F_1 \sim F_3$ のフィルタサイズのそれぞれが、2倍にされ、スケール空間層が再度計算され、その後、より高いスケールでの任意の新しい安定したキーポイントが識別される。フィルタ形状 $F_1 \sim F_3$ は、事前に十分にわかっているので、 $F_1 \sim F_3$ のスケールアップされた版のすべてを、事前に計算でき、図1のコンポーネントに含まれるメモリなどであるがこれに限定はされないメモリに格納することができる。したがって、本システムおよび方法は、同時に並列にキーポイント位置を別々に識別することを可能にするという点で、並列処理および/またはマルチスレッド処理に理想的に適する。

40

【0073】

オクターブごとにフィルタサイズを2倍にするプロセスは、必要なだけ何度でも繰り返される。やはり、SIFT法とは異なって、任意のスケールでのフィルタへの入力イメージが、必ず同一すなわちオリジナル入力イメージであることに留意されたい。

【0074】

上の説明では、フィルタを2倍にすることによってフィルタのサイズを変更する。しかし、フィルタのセットを任意の比率について設計できることを理解されたい。これは、連続するライン上のピクセルの間の補間の問題にすぎない。

【0075】

50

たとえば、提案される 1.5 倍の比率をとり、 3×3 最小サイズ水平フィルタから開始されたい (1つの中央線と2つの反対の極性の外側領域、「ガードバンド」なし)。次に、連続して 1.5 を乗算することによって、3 は 4.5、6.75、10.125 などになる。したがって、第 2 フィルタは、サイズにおいて 4.5×4.5 であり、それぞれ 1.5 ラインの高さ (スケーリングによって) の 3 つの水平領域からなる。

【0076】

4.5 フィルタ高さを半分に分割し、中央領域は、この分割線のそれぞれの側で 0.75 ラインを含む。したがって、中央領域は、分割の両側の両方のラインのピクセル値の合計の 0.75 倍から計算される。

【0077】

同様に、外側領域は、

(a) 分割の両側の両方の周辺のラインのピクセル値の合計の 0.25 倍と、

(b) 分割の両側の 1 ライン離れたラインのピクセル値の合計と、

(c) 分割の両側の 2 ライン離れたラインのピクセル値の合計の 0.25 倍と、

の合計によって計算される。

【0078】

後続のより大きいフィルタは、類似する形で得られる。

【0079】

分かる通り、境界上の共有されるイメージピクセルは、単純に、領域の間の境界線までの距離によって線形に重みを付けられる (すなわち、補間される)。

【0080】

したがって、この説明は、非整数比が多少の追加の補間および丸めを必要とすることを除いて、任意の比率に拡張することができる。

【0081】

I.4 スケール空間での安定したキーポイント位置の識別

本手法の目的は、ターゲットイメージのターゲットのノイズのある (すなわち、ダウングレードされた) 版内で高い確率で信頼できる形で見つけることができるイメージ内の安定したキーポイントを識別することである。

【0082】

DOG 関数に基づく SIFT 法とは異なって、本願で説明される手法は、「しみ様」のイメージ特徴および「接合点様」のイメージ特徴の検出に基づく。要件は、これらの「しみ」特徴および「接合点」特徴を、ターゲットイメージのノイズのある版で信頼できる形で見つけることができることと、これらの特徴が、シーン照明および類似物の変動に堅牢であり、また、イメージに対する相対的なある分散特性を有する (すなわち、これらの特徴を、イメージがスケーリングされ回転された版においても高い確率で見つけることができる) ことである。

【0083】

上のセクション I.2 で示したように、本システムおよび方法の水平フィルタは、現在のスケールでのフィルタサイズに対応するサイズの水平の「細長い」イメージ特徴について大きい大きさの信号を作る傾向がある。そのような特徴は、水平方向でよく局所化可能であるが、必ずしも垂直方向でよく局所化可能ではない。同様に、垂直フィルタは、垂直方向に局所化可能な特徴について強い大きさの信号を作る。この 2 つの応答を一緒に組み合わせることによって、水平方向と垂直方向との両方で局所化できるイメージ内の「しみ様」特徴位置を識別することが可能である。

【0084】

同様に、対角フィルタは、現在のスケールでの対角フィルタサイズに対応するサイズの斜めの「角」イメージ特徴または「接合点様」イメージ特徴について出力で大きい大きさの信号を作るように働く。そのような特徴は、対角線方向でよく局所化可能でもある。このタイプの例は、交番するチェッカーボードの黒白/白黒パターンまたはその逆の 4 つの隣接する正方形の間の接合点あるいは任意の 90 度の暗い角または明るい角を含む。もち

10

20

30

40

50

ろん、イメージ内の他の位置を、接合点として判定することができる。

【0085】

2つの特徴タイプすなわち「しみ様」イメージ特徴および「接合点様」イメージ特徴を異なるフィルタ出力で別々に見つけることを試みるのではなく、この実施形態では、2つの特徴タイプが、次の単一の特徴尺度に組み合わせられる。

【0086】

$$D(x, y, s) = F_H(x, y, s) \times F_V(x, y, s) - [k \times F_D(x, y, s)]^2$$

ここで、 $F_H(x, y, s)$ 、 $F_V(x, y, s)$ 、および $F_D(x, y, s)$ は、それぞれ図4A～図4Cに示された水平フィルタ出力、垂直フィルタ出力、および対角フィルタ出力であり、 k^2 は、異なるフィルタ形状を有する可能性および整数ピクセル境界での離散値への変換に起因する2つの特徴タイプ間の正規化パラメータである。

【0087】

「しみ様」イメージ特徴について、対角フィルタの出力での信号の大きさは、小さい可能性が高いが、水平フィルタおよび垂直フィルタの出力での大きさは、大きい。同様に、「接合点様」イメージ特徴について、水平フィルタまたは垂直フィルタの出力は小さい可能性が高いが、対角フィルタの出力の大きさは大きい。したがって、この2つの特徴タイプを、 $D(x, y, s)$ の出力の大きさを最大にするすなわち、 $D(x, y, s)$ の最大点または最小点のいずれかの極値を探すという単一の目標と一緒に組み合わせることができる。したがって、キーポイント位置は、関数 $D(x, y, s)$ がスケール s と空間 (x, y) との両方でその極値点を達成する、イメージ内の位置として識別される。

【0088】

スケール空間領域での極値点を検索する方法を図10に示す。差関数 $D(x, y, s)$ は、スケール空間全体にまたがって計算される。 (x, y, s) 内のピクセル位置 $(2e)$ は、関数 $D(x, y, s)$ の値が、図10に示されているようにその直接近傍に対して最大値または最小値のいずれかのピークである場合に、候補キーポイント位置であると判定される。

【0089】

あるイメージ内で識別できる安定したスケール不変キーポイント位置の個数は、イメージ内容および以下で定義する所望のキーポイント強さに依存して、かなり変化する。キーポイントの最終的な個数は、所定の通常的な自然シーンイメージ内で、数百個から数千個まで変化する可能性がある。したがって、イメージピクセルの大多数は、候補キーポイント位置ではない。したがって、性能の理由から、特定のピクセルが有効なキーポイント位置ではない時に候補キーポイント位置としてのそのピクセル位置をすばやく除外できる方法を探ることが望ましい。明らかに、各ピクセル位置の周囲の近傍ピクセルのすべてをチェックする必要があるには伴わずに、できる限り高速にそれを行うことが望ましい。

【0090】

この問題を解決する本願の実施形態の手法は、3重ピクセル比較テストの順序付きセットに基づく。この方法は、ルールベースである。各テストでは、中央ピクセルの両側の対向するピクセルの対を調べる。この方法は、現在のフィルタスケールで空間ピクセル近傍 (x, y) 平面を調べることによって開始される。近傍は、図10の中央の平面で9個のピクセル(すべて数字2から始まる)を含む。この近傍の拡大図を、図11に示す。この方法の目的は、中央ピクセル位置 $(2e)$ を候補キーポイント位置にすることができるかどうかを判定することである。

【0091】

この方法の第1のルールは、図11の前のピクセル位置 $(2d)$ がキーポイント位置として識別されたかどうかをチェックすることである。そうである場合には、位置 $(2d)$ での関数 $D(x, y, s)$ の値は、その近傍に対して最大値または最小値でなければならない。したがって、 $D(x, y, s)$ は、1ピクセル離れた位置 $(2e)$ でもピーク(すなわち、キーポイント)になることはできず、この場合に、 $(2e)$ のテストは、終了し

、このプロセスは、次のピクセルに進むことができる。

【0092】

しかし、前のピクセル位置(2d)がキーポイントではなかった場合に、この方法は、(2e)の両側の対向するピクセル対位置(2d)および(2f)での $D(x, y, s)$ の値を取り出す。中央ピクセル(2e)での $D(x, y, s)$ の値は、(2e)が図11の方向1に沿った $D(x, y, s)$ の局所極値点としての資格を有するために、最大値の場合には(2d)および(2f)のいずれよりも大きくなければならず、あるいは(2d)および(2f)より小さくなければならない(最小値の場合)。

【0093】

さらに、ノイズの存在および/またはあるピクセルから次のピクセルへの $D(x, y, s)$ の小さい丸め誤差に起因するランダムな弱いキーポイント位置の偶然の識別の尤度を除去するために、しきい値「T」を導入する。このしきい値は、次のように定義される。

すべての近傍 x, y, s について、

$$|D(x, y, s) - D_{(2e)}(x, y, s)| > T$$

すなわち、中央ピクセル位置での $D(x, y, s)$ の値は、そのピクセル近傍位置のいずれかの $D(x, y, s)$ のすべての他の値より少なくともTだけ大きいまたは小さいものでなければならない。しきい値Tの値は、識別できるキーポイントの個数を決定する。

【0094】

Tが大きい時には、最も強いピーク大きさ値 $D(x, y, s)$ を有する最も強いキーポイント位置だけが識別される。Tの値を減らすにつれて、より弱いピーク大きさの追加キーポイントを識別することができる。したがって、Tの値は、キーポイントの所望の個数および識別されるキーポイントの結果の強さを制御するのに使用することができる。

【0095】

本願の一実施形態で、Tの値は、イメージの期待されるタイプ、キーポイントの所望の個数、およびキーポイントの相対的な強さに基づいて事前に予め決定される定数値である。

【0096】

本願の他の1つの実施形態で、Tの値は、上記で示したように最小キーポイント強さを決定する固定された定数部分と、当該の現在のピクセル位置の周囲の局所領域内の $D(x, y, s)$ 大きさの平均変動に基づく動的に可変な部分と、の合計からなる。Tの動的な調整は、他のより静かなイメージ区域に対して騒がしいイメージ区域内で識別されるキーポイントの個数を制御するのに使用することができる。

【0097】

本願の第3の実施形態で、対向する対のピクセルの $D(x, y, s)$ 値と(2e)での中央 $D(x, y, s)$ 値との間の大きさの差が、キーポイント強さの尺度を提供するために、ピクセル対にまたがって累算され、正規化される。キーポイントの強さは、その後、より弱くより不安定なキーポイントを除去し、その強さに従ってキーポイントをソートするのに使用することができる。

【0098】

図11を継続して参照すると、(2e)での中央 $D(x, y, s)$ 値が、対向する近傍位置(2d)および(2f)に関してピーク値になるためのすべての必要条件を満足してはいない場合に、やはり、(2e)をキーポイント位置にすることができないと結論することができる。このプロセスは、次のピクセル位置に移ることができる。

【0099】

そうではなく、中央位置(2e)が、まだキーポイント位置であり得る場合に、この方法は、図11の方向2に沿った(2e)の両側のピクセル位置(2a)および(2i)の次の対向する対の $D(x, y, s)$ の値を取り出す。この方法は、中央位置(2e)がまだキーポイント位置として識別され得るかどうかを判定するために、上の分析を繰り返す。この対でのピクセル比較テストは、図11に示された残りのピクセル対方向3および4について、中央位置(2e)の周囲の円で継続される。

10

20

30

40

50

【0100】

この方法は、候補キーポイント位置として中央位置(2e)を識別することに失敗するピクセル対が最初に出現した時に終了するように設計される。検索は、しばしば、追加テストにより長い時間を費やすことなく、これらの対でのピクセルテストのうちの2~3回のみ後に、すばやく終了する。しかし、この実施形態では、ピクセル位置は、すべての可能な対でのテストおよびすべての他の追加のルールがすべて完全に満足された後に限って、安定したキーポイント位置として識別される。

【0101】

もう1つの実施形態では、すべての可能な対でのテストを行う必要はないものとしてとすることができる。たとえば、計算の回数を減らすために、対でのテストの一部を除去することができる(たとえば、図12の遠い角1a、1c、1g、1iから)。1ピクセルだけ離れた近傍のみを使用することによって、これは、テストの回数をこのケースで9回から5回に減らす(スケールの間で)。したがって、この実施形態では、すべての可能な対でのテストが使用されるわけではない。

10

【0102】

上で述べた追加ルールの例を、以降の段落で展開する。さらに、そのような追加ルールは、たとえば1ピクセルを超えて離れたさらなる対向するピクセル対を調べ、または中央ピクセルの両側の両方が同一の符号を有するかどうかをチェックすることによって、現在の同一スケールのケースにも適用可能である。たとえば、2aの左に1ピクセルおよび2iの右に1ピクセルは、図11の1と2との間の新しい方向での新しいテストをもたらすなどである。

20

【0103】

現在のスケール平面のテストにおいて、上で概要を示した図11の方向1~4に沿った4つの対でのピクセルテストを完了したならば、この方法は、中央ピクセル位置(2e)をそのスケールに関してキーポイント位置として識別できるかどうかのチェックに進む。

【0104】

9つの追加の対向する対でのテストを、図12に示された方向1~9に沿って適用する。各対でのピクセルテストは、本質的に、キーポイントの強さのしきい値Tを含めて、上で説明したものと同一である。しかし、すべてのピクセル位置が同一の現在のスケールにある、4つの対でのテストの前のケースとは異なって、今回は、対向するピクセル位置が、必ず異なるスケールから選択され、一方は現在のスケールより小さいスケール、他方はより大きいスケールから選択される。これは、中央位置(2e)が実際にすべての可能な方向でD(x, y, s)の真のスケール空間ピークであることを保証するために行われる。

30

【0105】

やはり、この方法は、前述の対でのピクセルテストのいずれかの最初の失敗で終了するようにセットされる。

【0106】

本願の他の実施形態では、追加の対でのルールを追加して、識別できる結果のキーポイントに対するさらなる制約を設けることができる。たとえば、D(x, y, s)の符号一貫性に関する追加の制約、または、1つ小さいスケールおよび1つ大きい平面でのさらに離れた対向するピクセルの対を追加して、より多くの角度での追加の方向テストを実現することによるものなどである。

40

【0107】

図13A~図13Eに、本発明の実施形態に従って実行されるキーポイント識別手順の結果を示す。より具体的には、入力イメージが、図13Aに示されている。図13Bは、水平フィルタによって生成された出力であり、図13Cは、垂直フィルタによって生成された出力であり、図13Dは、対角フィルタによって生成された出力であり、このすべてが同一の特定のスケールである。最後に、図13Eは、3つすべてのフィルタを組み合わせた時の出力を示す。

50

【0108】

図13Aの入力イメージは、グレイスケールイメージに変換されたオリジナルイメージとすることができる。次に、異なるサイズのフィルタのセットをグレイスケールイメージに適用して、図9に示されたスケール空間ピラミッド層を作成する。ある特定のスケールでの水平フィルタ、垂直フィルタ、および対角フィルタの通常の出力を、それぞれ図13B、図13C、および図13Dに示す。この3つのフィルタ出力を一緒に組み合わせて、少なくとも3つの連続するスケールで関数 $D(x, y, s)$ を形成する。最後に、3つのスケールのうちの1つすなわち現在のスケールの出力を、図13Eに示す。 $D(x, y, s)$ がそのすべての近傍に対してスケール空間内で極値を達成するピクセル位置が、現在のスケール空間ピラミッド層の安定したキーポイント位置として識別される。

10

【0109】

追加のキーポイントをより高いピラミッド層から抽出することができる。

【0110】

図14に、より高いピラミッド層内で追加のキーポイント(円によって示される)を識別するプロセスを示す。参照のために、図15に、各ピラミッド層でフィルタサイズを2倍にすることに起因するイメージサイズの対応する縮小を考慮に入れた後の「同等の」DOGイメージピラミッドを示す。図15が、参照のみのために提供され、提案される方法実施態様の一部と混同されてはならないことに留意されたい。

【0111】

本願の提案される方法は、SIFTなどの既存の方法のどれよりも計算がかなり高速でありながら、イメージ内の安定したキーポイントの識別において同様に効果的であると認められる。

20

【0112】

ガウシアンフィルタは、非常に高速に計算でき、既存の方法に対してはるかに少数の計算を必要とする、はるかにより単純なハールライク(Haar-like)特徴突出インジケータ長方形フィルタに置換される。既存の方式の多くとは異なって、ガウシアンフィルタリングは、現在説明されている複数スケール分析で直接には使用されない。ガウシアンフィルタリングは、その優れた平滑化特性に起因して、空間スケール分析に望ましいと判定された。しかし、実際には、ガウシアン形状は、離散値にデジタル化することと、さらに、切り捨てなければ無限の尾をある点を超えて切り捨てることとによって近似されなければならないという点で、理想的ではない。さらに、正確なガウシアンフィルタリングを用いても、エイリアシングアーティファクトが、イメージをダウンサンプリングする時に導入され得る。非常に重要視されている技法のいくつかは、ガウシアン形状のさらなる近似の下であっても優れた結果を実証した(たとえば、SIFT法の一部としてのディファレンス・オブ・ガウシアン(DOG)イメージピラミッドとしてのラプラシアン・オブ・ガウシアン(LOG)イメージピラミッドのD. Loweの近似)。

30

【0113】

本システムおよび方法は、再帰的フィルタリングを使用せずに、オリジナルイメージに直接に異なる長方形形状サイズを適用することによって、イメージピラミッド全体を便利に効率的に計算することができる。したがって、本システムおよび方法は、反復的にイメージをスケールダウンするのではなく、フィルタサイズをスケールアップすることによって複数スケール分析を実行することができる。

40

【0114】

本願の実施形態のシステムおよび方法は、非常に高速のレートを達成するだけでなく、その速度は、すべてのフィルタサイズについて正確に同一である。長方形フィルタ形状を用いると、これを達成するための面積合計積分の特性が活用される。まず、所定の左下のイメージ角に対するオリジナルイメージの各ピクセルの面の積合計を計算する。これは、行および列を記憶しながら前の面積の合計に現在のピクセル値を再帰的に加算することによって、1ピクセルあたり約1回の加算で効率的に行うことができる。面積合計積分をイメージ全体について計算し、メモリに格納したならば、任意の長方形フィルタ形状の下

50

の面積を、任意の長方形フィルタサイズについて2回の加算だけを使用して計算でき、性能の改善につながる。

【0115】

たとえばヘッセ行列ベースの技法と同様にガウシアン二次導関数を使用するのではなく、本願の実施形態では、他の点では複数スケール分析として知られる、2つの空間イメージ次元空間内ならびにさまざまなスケール内で局所化できる独特なキーポイントを見つけることを試みる。この方法は、そのようなキーポイントの位置を識別するのに単純な長方形フィルタ形状を使用する。本願の実施形態では、イメージ内の最も特徴的なキーポイントのセットを探すので、フィルタの正確な形状は、決定的ではない。すなわち、鋭いビルディングの角または目の角などの明確に局所化されたイメージキーポイントは、より単純なフィルタを用いて同様によく見つけられる可能性が高い。本願の実施形態では、垂直方向、水平方向、および対角線方向でイメージ特徴を取り込むように設計されたフィルタ形状を利用する。特定のフィルタ形状は、ガウシアン形状の長方形近似と解釈することができる。入力イメージの品質および予想される取り込みデバイスに応じて、これらのフィルタは、さらに、ノイズおよびサンプリングアーティファクトを除去するのを助けるために分離「ガードバンド」を含むことができる。基本的なフィルタ形状は、複数スケールイメージ分析をもたらすためにスケールアップされる。しかし、フィルタ形状の実際のスケールリングは、必要ではない。長方形のフィルタ形状に起因して、スケールリングを、前もって新しい角の座標を記録し、その後、新しい異なるサイズのフィルタ結果を得るためにイメージ積分を使用することによって、効率的に行うことができる。

10

20

【0116】

本システムおよび方法は、各ピラミッドレベルを、上記で概要を示した面積合計を使用して入力イメージから直接に計算できるので、並列な実施態様に特に適する。各ピラミッドレベルの計算は、他のピラミッドレベルと完全に独立であり、フィルタサイズに関わりなく正確に同一の時間を要する。

【0117】

本願の実施形態では、イメージ内の独特なキーポイントの位置を判定するのにルールベースの方法を使用する。数百万個のピクセルを含む通常のイメージは、数百個から数千個の候補キーポイントを生じる場合がある。ほとんどのイメージピクセルはキーポイントではないので、任意の所定のイメージピクセルで費やされる時間の最小化が探される。本システムおよび方法は、現在のピクセルを強いキーポイント候補にすることができるかどうかをすばやく判定するために、さまざまなスケールおよび方位で当該の現在のピクセルの両側の近傍ピクセルの対などの間で一連の連続するテストを実行する。これらのテストは、尤度順によって編成される。各連続するテストで、現在のピクセルとその隣接物との間の関係が、要求される条件を満足しない場合には、この方法は、キーポイントとしての現在の位置を除外し、次のピクセルに移る。したがって、全体的な時間が最小化されると同時に、各最終的なキーポイントは、有効なキーポイントであるためのすべての条件を満足する。本システムおよび方法は、複数のそのようなルールを使用して、ありそうなキーポイント候補を判定する。

30

【0118】

本願の実施形態では、結果の滑らかさを改善し、フィルタ境界に沿った潜在的な不連続性の影響を減らすために、各後続フィルタ位置のサンプリングのオーバーラップという概念も導入される。長方形フィルタは、スケールリングの後に一定のままになるように正規化される。さらに、フィルタサイズが、各より大きいスケールの分析のために増やされる時に、サンプリング位置をも対応して増やして、キーポイントの識別に関するより高いピラミッド層の有効解像度を便利に下げることができる。

40

【0119】

キーポイントの候補それぞれのキーポイントの強さの尺度も導入される。結果のキーポイントを、イメージ内の最良の最も特徴的なキーポイントを識別し、イメージの劣化したノイズのある版で見つかる可能性が高くはないすべての他の弱いキーポイントを破棄する

50

ために、強さの順でソートすることができる。通常、最も強いキーポイントのうちの最上位の数百個が、イメージ照合のために最も有用である。したがって、性能の改善を、最も強いキーポイントの小さいサブセットに集中することと、照合結果が結論に達しない場合に限ってより弱いキーポイントを追加することとによって得ることができる。代替案では、キーポイントの所望の個数および強さを前もって決定することができ、キーポイント識別プロセスを、十分な個数の適切なキーポイントが識別されるや否や停止することができ、これによって、さらなるキーポイントを識別するために時間を浪費しなくなる。キーポイントの強さは、各テストの個々の寄与を使用して、テストのシーケンス中にオンザフライで計算される。キーポイントの強さの尺度は、さらなる下流処理でアンカキーポイントを選択するためなどであるがこれには限定されないある種の特定用途向け処理に特に有用である。

10

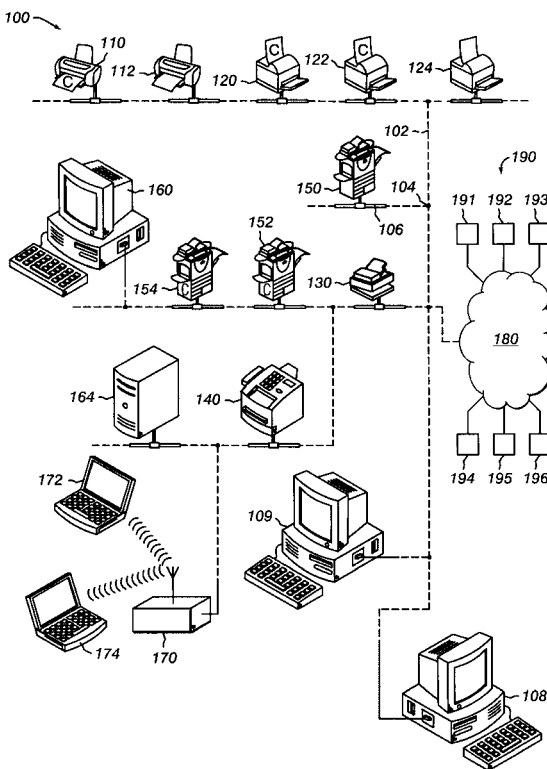
【符号の説明】

【0120】

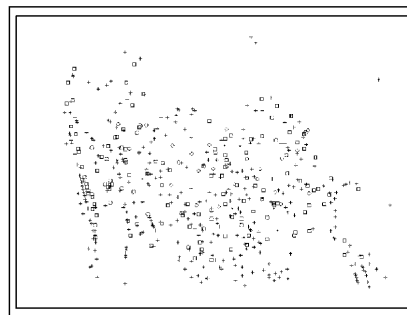
100 コンピュータネットワーク、102, 106 ワイヤ、104 ワイヤジャンクション、108, 109 コンピュータ、110 カラープリンタ、112 カラー以外のプリンタ、120, 122 カラーレーザプリンタ、124 カラー以外のレーザプリンタ、130 スキャナ、140 ファクシミリ機、150 写真複写機、152 カラー写真複写機、154 組合せカラープリンタ/スキャナ/ファクシミリ機、160 パーソナルコンピュータおよび/または独立型コンピュータ端末、164 独立型ハードドライブデータストレージ媒体、170 無線ネットワーク送受信器、172, 174

20

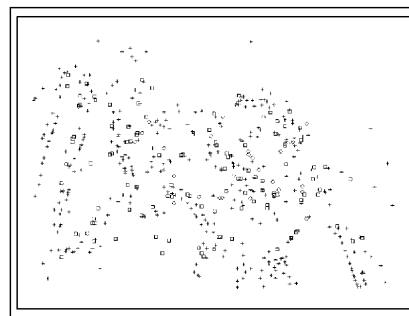
【図1】



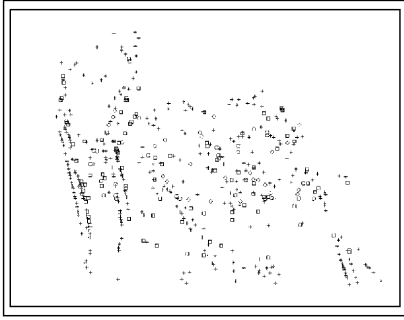
【図2B】



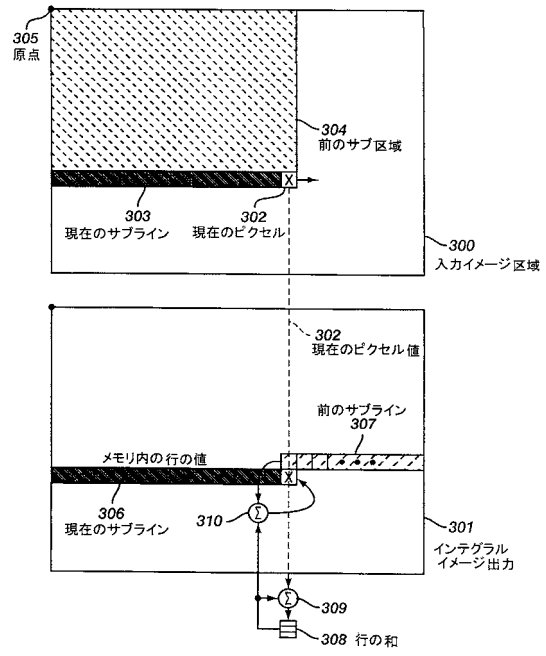
【図2D】



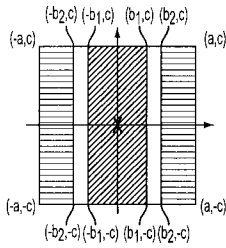
【 図 2 F 】



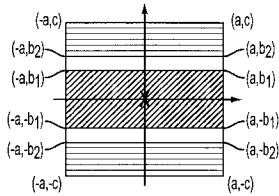
【 図 3 】



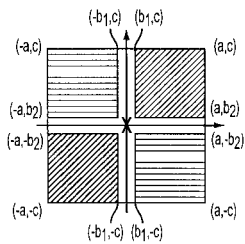
【 図 4 A 】



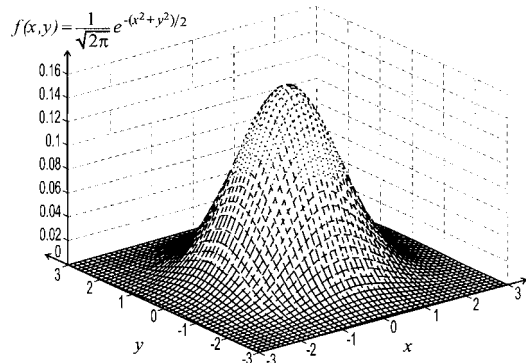
【 図 4 B 】



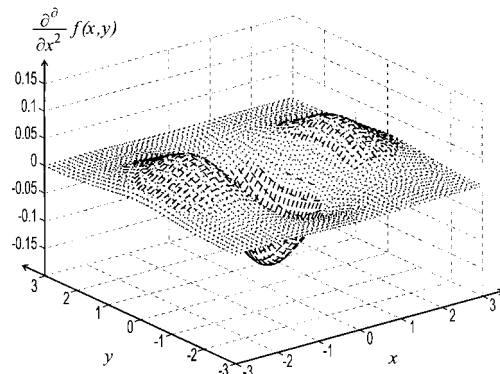
【 図 4 C 】



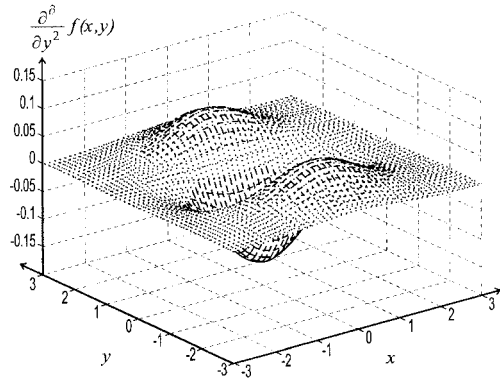
【 図 5 A 】



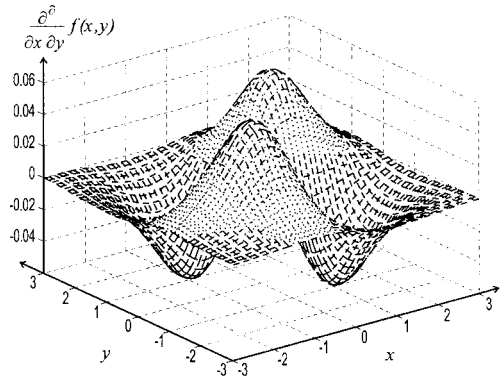
【 図 5 B 】



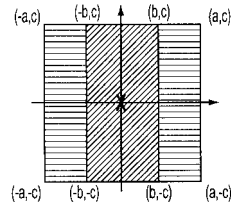
【 図 5 C 】



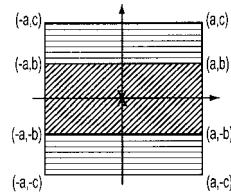
【 図 5 D 】



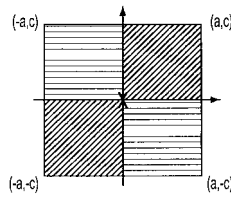
【 図 6 A 】



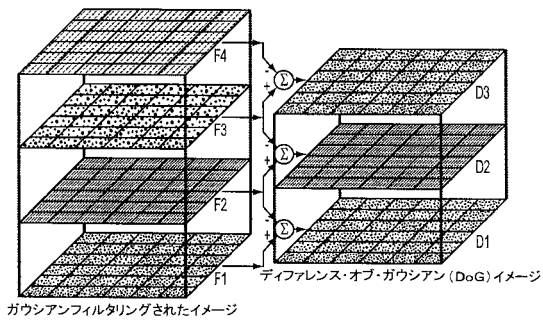
【 図 6 B 】



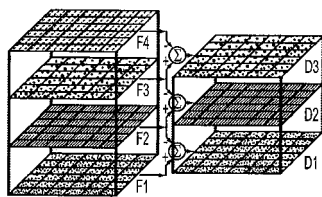
【 図 6 C 】



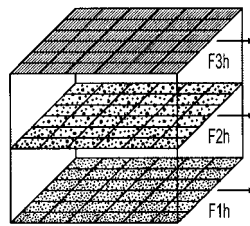
【 図 7 A 】



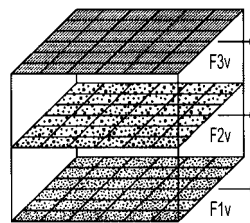
【 図 7 B 】



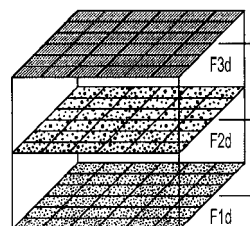
【 図 8 A 】



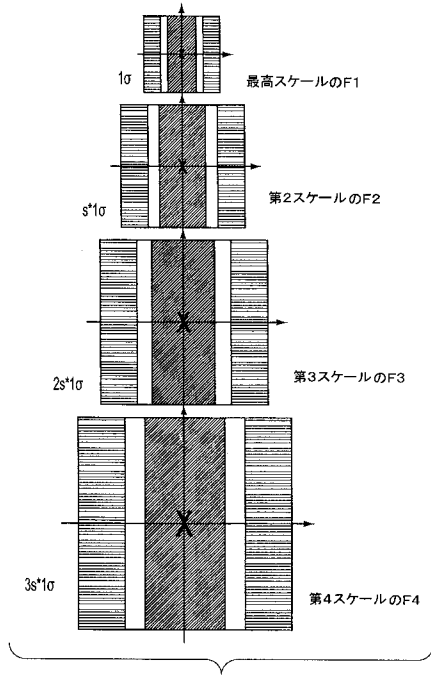
【 図 8 B 】



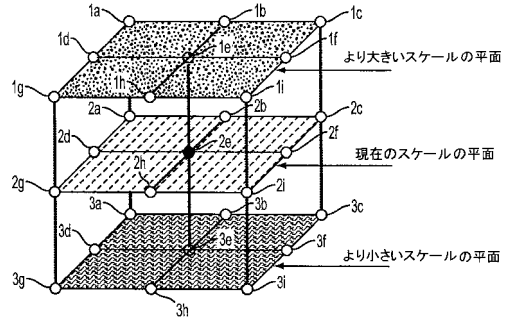
【 図 8 C 】



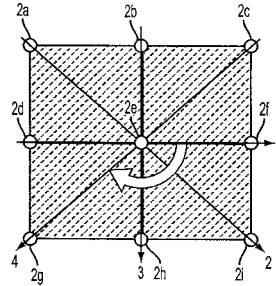
【 図 9 】



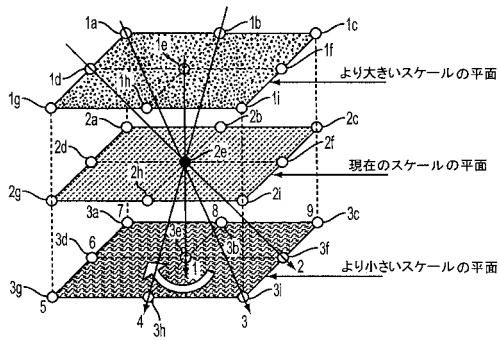
【 図 1 0 】



【 図 1 1 】



【 図 1 2 】



【 図 2 A 】



【 図 2 C 】



【 図 2 E 】



【 図 1 3 A 】



【図 13 B】



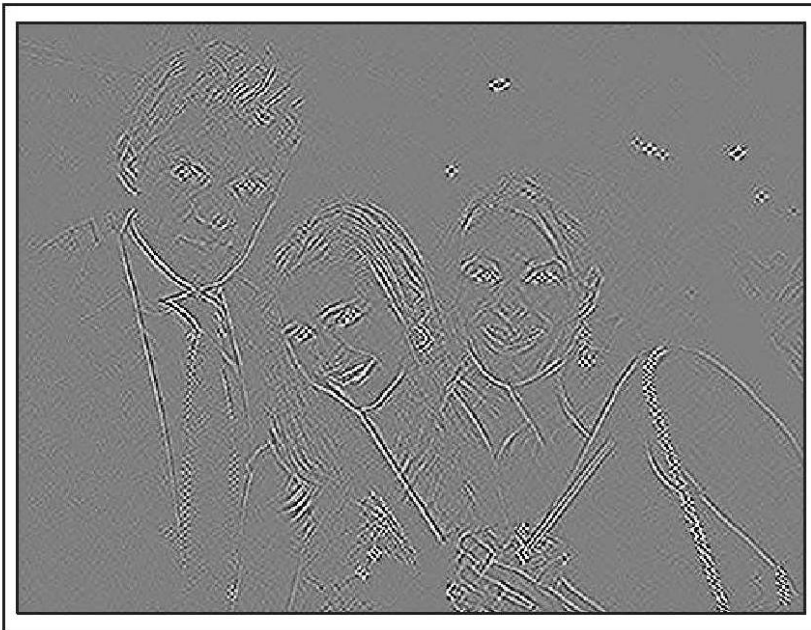
【図 13 C】



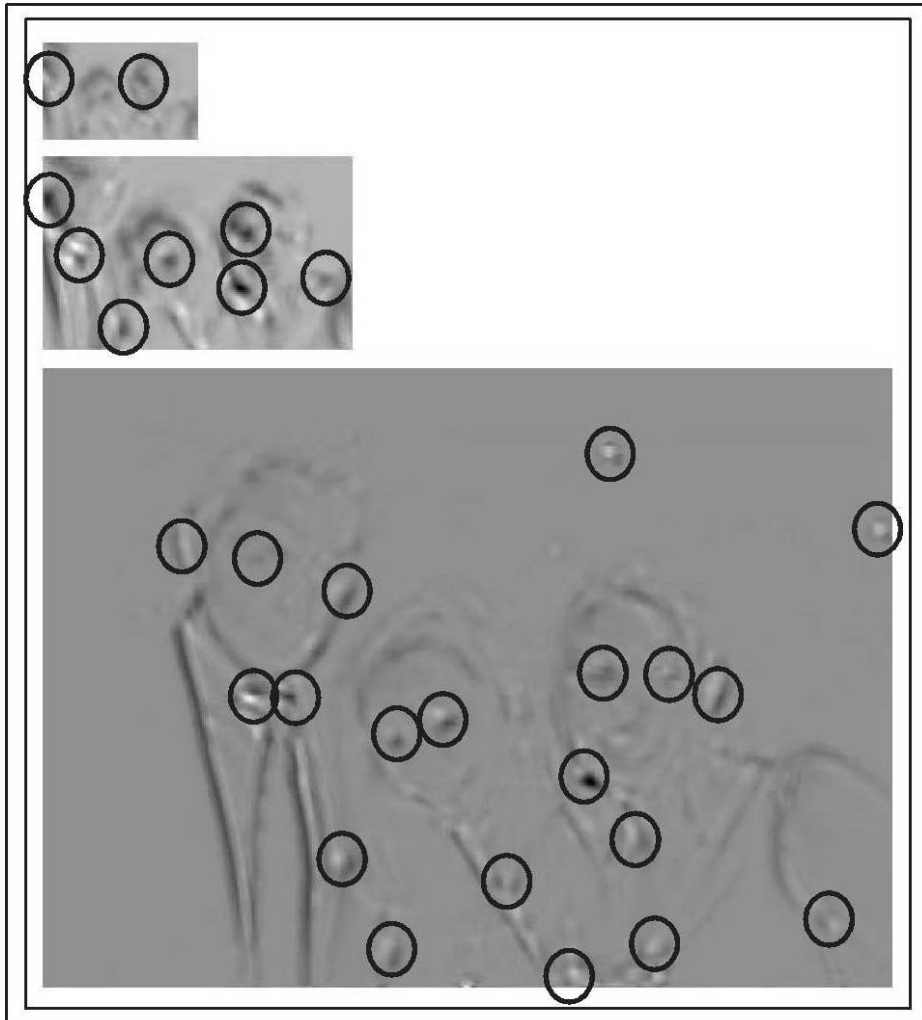
【 図 1 3 D 】



【 図 1 3 E 】



【 図 1 4 】



【 図 1 5 】



フロントページの続き

Fターム(参考) 5B057 CA08 CA12 CA16 CB08 CB12 CB16 CC02 CD05 CE03 CE05
CE06 CF05 CH09 CH11 DA07 DA16 DB02 DB09 DC30