

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4356371号
(P4356371)

(45) 発行日 平成21年11月4日(2009.11.4)

(24) 登録日 平成21年8月14日(2009.8.14)

(51) Int.Cl. F1
G06T 7/20 (2006.01) G06T 7/20 A

請求項の数 6 (全 22 頁)

<p>(21) 出願番号 特願2003-173524 (P2003-173524) (22) 出願日 平成15年6月18日(2003.6.18) (65) 公開番号 特開2005-11005 (P2005-11005A) (43) 公開日 平成17年1月13日(2005.1.13) 審査請求日 平成18年5月22日(2006.5.22)</p>	<p>(73) 特許権者 000005496 富士ゼロックス株式会社 東京都港区赤坂九丁目7番3号 (74) 代理人 100114937 弁理士 松本 裕幸 (72) 発明者 シュエ ジェンル 神奈川県足柄上郡中井町境430 グリー ンテクなかい 富士ゼロックス株式会社内 (72) 発明者 瀬古 保次 神奈川県足柄上郡中井町境430 グリー ンテクなかい 富士ゼロックス株式会社内 審査官 松尾 俊介</p>
---	--

最終頁に続く

(54) 【発明の名称】 情報処理装置

(57) 【特許請求の範囲】

【請求項1】

動画像から時間的に変化する対象を検出する情報処理装置において、
動画像の対象フレームと前記対象フレームと時間的に隣接する他のフレームとの対応する画素の画素値の差と予め定められた閾値との比較結果に基づき、画素値が時間的に変化する画素である変化画素を検出する変化画素検出手段と、

前記変化画素検出手段により検出された変化画素と他の画素を区別して表す動作画像について、分解能低減前の画像における予め定められた画素数を単位とした画素群に含まれる変化画素の数に基づいて、分解能を1段階低減させた画像の画素が変化画素であるか否かを決定する処理を、予め定められた分解能の画像が得られるまで繰り返し行い、これにより得られた分解能の異なる複数の画像を用いて、前記予め定められた分解能の画像における画素値が時間的に変化する画素領域の検出から開始して、当該検出した画素領域を処理対象領域として、分解能が1段階以上高い画像における画素値が時間的に変化する画素領域を検出することを複数段階繰り返す変化画素領域検出手段と、

前記対象フレームより時間的に前のフレームから検出された変化画素領域のエッジから前記対象フレームにおける変化画素領域のエッジを予測し、当該予測される変化画素領域のエッジに基づいて複数の追跡線を配置し、各追跡線についての勾配及びモード値に基づいてその追跡線のエッジを見つけることにより、前記変化画素領域検出手段により検出される画素領域のエッジを検出する変化画素領域エッジ検出手段と、

ランダムハフ変換法を用いて、前記変化画素領域エッジ検出手段による検出結果に基づ

いて時間的に変化する対象を検出する変化対象検出手段と、
を備えたことを特徴とする情報処理装置。

【請求項 2】

請求項 1 又は請求項 2 に記載の情報処理装置において、
前記変化画素領域検出手段は、ラベリング法を用いて、画素値が時間的に変化する複数の画素領域を識別して検出し、
前記変化画素領域エッジ検出手段は、前記変化画素領域検出手段により検出されるそれぞれの画素領域のエッジを検出し、
前記変化対象検出手段は、それぞれの画素領域毎に、前記変化画素領域エッジ検出手段による検出結果に基づいて時間的に変化する対象を検出する、
ことを特徴とする情報処理装置。

10

【請求項 3】

請求項 1 乃至請求項 3 のいずれか 1 項に記載の情報処理装置において、
前記変化対象検出手段により検出する対象は、パラメータを用いて表現される、
ことを特徴とする情報処理装置。

【請求項 4】

動画像から時間的に変化する対象を情報処理装置により検出する情報処理方法において、
前記情報処理装置は、変化画素検出手段と、変化画素領域検出手段と、変化画素領域エッジ検出手段と、変化対象検出手段と、を備えており、

20

前記変化画素検出手段が、動画像の対象フレームと前記対象フレームと時間的に隣接する他のフレームとの対応する画素の画素値の差と予め定められた閾値との比較結果に基づき、画素値が時間的に変化する画素である変化画素を検出し、

前記変化画素領域検出手段が、前記変化画素検出手段により検出された変化画素と他の画素を区別して表す動作画像について、分解能低減前の画像における予め定められた画素数を単位とした画素群に含まれる変化画素の数に基づいて、分解能を 1 段階低減させた画像の画素が変化画素であるか否かを決定する処理を、予め定められた分解能の画像が得られるまで繰り返し行い、これにより得られた分解能の異なる複数の画像を用いて、前記予め定められた分解能の画像における画素値が時間的に変化する画素領域の検出から開始して、当該検出した画素領域を処理対象領域として、分解能が 1 段階以上高い画像における画素値が時間的に変化する画素領域を検出することを複数段階繰り返し、

30

前記変化画素領域エッジ検出手段が、前記対象フレームより時間的に前のフレームから検出された変化画素領域のエッジから前記対象フレームにおける変化画素領域のエッジを予測し、当該予測される変化画素領域のエッジに基づいて複数の追跡線を配置し、各追跡線についての勾配及びモード値に基づいてその追跡線のエッジを見つけることにより、前記変化画素領域検出手段により検出される画素領域のエッジを検出し、

前記変化対象検出手段が、ランダムハフ変換法を用いて、前記変化画素領域エッジ検出手段による検出結果に基づいて時間的に変化する対象を検出する、
ことを特徴とする情報処理方法。

【請求項 5】

動画像の対象フレームと前記対象フレームと時間的に隣接する他のフレームとの対応する画素の画素値の差と予め定められた閾値との比較結果に基づき、画素値が時間的に変化する画素である変化画素を検出する変化画素検出機能と、

40

前記変化画素検出機能により検出された変化画素と他の画素を区別して表す動作画像について、分解能低減前の画像における予め定められた画素数を単位とした画素群に含まれる変化画素の数に基づいて、分解能を 1 段階低減させた画像の画素が変化画素であるか否かを決定する処理を、予め定められた分解能の画像が得られるまで繰り返し行い、これにより得られた分解能の異なる複数の画像を用いて、前記予め定められた分解能の画像における画素値が時間的に変化する画素領域の検出から開始して、当該検出した画素領域を処理対象領域として、分解能が 1 段階以上高い画像における画素値が時間的に変化する画素

50

領域を検出することを複数段階繰り返す変化画素領域検出機能と、

前記対象フレームより時間的に前のフレームから検出された変化画素領域のエッジから前記対象フレームにおける変化画素領域のエッジを予測し、当該予測される変化画素領域のエッジに基づいて複数の追跡線を配置し、各追跡線についての勾配及びモード値に基づいてその追跡線のエッジを見つけることにより、前記変化画素領域検出機能により検出される画素領域のエッジを検出する変化画素領域エッジ検出機能と、

ランダムハフ変換法を用いて、前記変化画素領域エッジ検出機能による検出結果に基づいて時間的に変化する対象を検出する変化対象検出機能と、

をコンピュータに実現させるためのプログラム。

【請求項6】

コンピュータに実行させるプログラムを当該コンピュータの入力手段により読み取り可能に記憶した記憶媒体において、

当該プログラムは、動画像の対象フレームと前記対象フレームと時間的に隣接する他のフレームとの対応する画素の画素値の差と予め定められた閾値との比較結果に基づき、画素値が時間的に変化する画素である変化画素を検出する変化画素検出機能と、

前記変化画素検出機能により検出された変化画素と他の画素を区別して表す動作画像について、分解能低減前の画像における予め定められた画素数を単位とした画素群に含まれる変化画素の数に基づいて、分解能を1段階低減させた画像の画素が変化画素であるか否かを決定する処理を、予め定められた分解能の画像が得られるまで繰り返し行い、これにより得られた分解能の異なる複数の画像を用いて、前記予め定められた分解能の画像における画素値が時間的に変化する画素領域の検出から開始して、当該検出した画素領域を処理対象領域として、分解能が1段階以上高い画像における画素値が時間的に変化する画素領域を検出することを複数段階繰り返す変化画素領域検出機能と、

前記対象フレームより時間的に前のフレームから検出された変化画素領域のエッジから前記対象フレームにおける変化画素領域のエッジを予測し、当該予測される変化画素領域のエッジに基づいて複数の追跡線を配置し、各追跡線についての勾配及びモード値に基づいてその追跡線のエッジを見つけることにより、前記変化画素領域検出機能により検出される画素領域のエッジを検出する変化画素領域エッジ検出機能と、

ランダムハフ変換法を用いて、前記変化画素領域エッジ検出機能による検出結果に基づいて時間的に変化する対象を検出する変化対象検出機能と、を当該コンピュータに実現させるためのプログラムである、

ことを特徴とする記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、動画像から時間的に変化する対象を検出する情報処理装置などに関し、例えば、背景が複雑であるような場合や、対象が突然出現或いは消滅したり対象の一部が隠れたりするような場合においても、確実に対象を検出することが可能な技術に関する。

【0002】

【従来の技術】

例えば、マシンビジョンにおいて、対象物を見つけて追従することは、1つの共通の目標となっている。このような画像処理を行うためのアルゴリズムとしては、一般に、2種類ある。1つは、領域分割（セグメンテーション）とマッチングの前処理として全フレーム範囲の画像処理を行うアルゴリズムであり、もう1つは、部分的な画像の特徴検出を行うアルゴリズムである。

【0003】

多くの追跡の問題はいずれのアルゴリズムによっても解決され得るが、解を得るためのデータ処理に対する要求は大きく異なる。オプティカルフローの計算或いは領域の分割のような全フレーム範囲に関するアルゴリズムは、大量のデータ処理になりやすく、オフラインで実行され、或いは、特別なハードウェアを使用して加速させられ得る。一方、特徴に

10

20

30

40

50

基づくアルゴリズムは、通常、画像の空間的な一部の領域に集中し、そして、処理が必要なデータの量が比較的少なく、標準的な計算を行うハードウェアで動作する一連のアルゴリズムにより取り扱われることが可能である（例えば、非特許文献2参照。）。

【0004】

部分的な特徴検出の手法の中で、実時間の映像（動画像）において目標物を追跡する基本的な方法として、一時的差分（TD：Temporal Differencing）法や（例えば、非特許文献7参照。）、テンプレート相関マッチング法がある。一時的差分法では、一定の時間 t だけ離隔した映像のフレームが比較されて、変化した領域を見つける。テンプレート相関マッチング法では、それぞれの映像のフレームが走査（スキャン）されて、画像のテンプレートと最も相関する領域を見つける。

10

【0005】

これらの方法は、それぞれ、長所と短所を有している。一時的差分法は、目標物が隠れてしまったりその動きをやめてしまうような場合には、うまく機能しない。テンプレート相関マッチング法は、一般に、目標物の外観が一定であることを要求し、対象物の大きさや方向が変化し或いは光の当たり具合が変化するような場合においても確実性が低下してしまう。

【0006】

しかしながら、これら2つの方法による追跡の特性は相補的なものである。つまり、目標物が静止している場合には、テンプレート相関マッチング法の確実性が最大となる一方、一時的差分法はうまく機能しない。また、目標物が動いている場合には、一時的差分法が

20

【0007】

多くの人工の物体は、円形や楕円形などの図形であるといった視覚的な特徴或いはこのような図形の部分を含むといった視覚的な特徴を有している。このため、抽出される幾何学的な特徴を用いて対象となる物体の見かけを決定することも、活発に研究されている論題である。

【0008】

人工の物体を追跡するために使用される点や線や輪郭などのような幾何学的な特徴の中で、特に注目されている特徴は円である。この理由は、円は、円という1つの特徴の中に、3個の自由度（DOF：degree of freedom）を有しているからである。なお、比較例として、線の場合には自由度は2であり、点の場合には自由度は1である（例えば、非特許文献3～5参照。）。

30

【0009】

また、ランダムハフ変換（RHT：Randomized Hough Transform）は、このような幾何学的な特徴を抽出するために使用されるのに非常に役立つ（例えば、非特許文献6参照。）。

ランダムハフ変換は、検索空間の中で、パラメータで表示されるテンプレートモデルと画像との相関を用いて、相関のピークを見つける。

【0010】

【非特許文献1】

Ayromlou, M., Vincze, M., Kubinger, W., Zillich, M, Robust, “Tracking of ellipses at Frame Rates”, OAGM workshop on Pattern recognition, 155-164, Steyr, Austria, May 27-28, 1999.

40

【非特許文献2】

Gregory D. Hager and Kentaro Toyama, “X Vision: A portable substrate for real-time vision applications.” Computer Vision and Image Understanding, vol.69, no.1 Jan, 1998, pp.23-37

【非特許文献3】

M. Ferri, F.Mangili, G.Viano, “Projective pose estimation of linear and quadratic primitives in monocular computer vision”, CVGIP:Image understanding 58(1)(1993)66-84.

50

【非特許文献 4】

D. Forsyth, J.L. Mundy, A.Zisserman, C.Coelho, A.Heller, C.Rothwell, “Invariant descriptors for 3-D objects recognition and pose”, IEEE PAMI 13 (10) (1991) 971-991.

【非特許文献 5】

P.Wunsch, G. Hirzinger, “Real-time visual tracking of 3-D object with dynamic handling of occlusion”, IEEE International Conference on Robotics and Automation, 1997.

【非特許文献 6】

Lei XU, Erkki OJA, and Pekka Kultanena. “A new curve detection method: Randomized hough transform.” Pattern Recognition Letters, (11):331-338, 1990 10

【非特許文献 7】

C.Anderson, P.Burt, G.van der Wal, “Change detection and tracking using pyramid transformation techniques”, SPIE-Intelligent Robots and Computer Vision, vol. 579, pp. 72-78, 1985.

【非特許文献 8】

JOrg, S., Langwald, J.,el.al, “Flexible Robot-Assembly using a multi-sensory approach”, IEEE ICRA, pp.3687-3694, 2000.

【非特許文献 9】

Martin A. Fischler and Robert C. Bolles, “Random Sample Consensus: A Paradigm for model fitting with applications to image analysis and automated cartography”, Communications of ACM: Graphics and Image Processing, Vol.24,no.6, June 1981. p. 381-395. 20

【非特許文献 1 0】

M. Isard and A. Blake, “Contour tracking by stochastic propagation of conditional density”, Proceedings of European Conference on Computer vision 96, pp. 343-356, 1996.

【非特許文献 1 1】

R.K.K. Yip, P.K.S, Tam,D.N.K, Leung, “Modification of Hough Transform for circles and ellipse detection using a 2-dimensional array”, Pattern Recognition 25 (9) (1992),pp.1007-1022. 30

【非特許文献 1 2】

Y.Lei, K.-C.Wong, “Ellipse detection based on the symmetry”, Pattern Recognition Lett. 20(1999)41-47.

【非特許文献 1 3】

M. Kass, A. Witkin, D. Terzopoulos, “Snakes: active contour model”, Int. J. Comput. Vision 1(1987)321-331.

【非特許文献 1 4】

D. Terzopoulos, R. Szeliski, “Tracking with Kalman snakes”, in: A.Blake, A.Yuille, Active vision, MIT Press, Cambridge, MA, 1992. 40

【非特許文献 1 5】

G. Hager, K.yoyama, “The Xvision-system: a portable substrate for real-time vision applications”, Comput.Vision Image Understanding 69(1) (1998)23-37.

【非特許文献 1 6】

Xue Jianru, “Real time circle detection in Video”, IML-TR-02-006 (2002), Information Media Lab, Fuji Xerox Co.,Ltd.

【非特許文献 1 7】

Xue Jianru, “Robust tracking of multiple non-rigid circles in realtime”, Technical Report, (2003), Information Media Lab, Fuji Xerox Co.,Ltd.

【 0 0 1 1】

【発明が解決しようとする課題】

しかしながら、従来においては、動画像から時間的に変化する対象を検出する技術については、未だに不十分な点があり、更なる開発が要求されていた。具体的には、例えば、背景が複雑であるような場合や、対象が突然出現或いは消滅したり対象の一部が隠れたりするような場合においても、確実に対象を検出することが可能な技術の開発が望まれていた。

【0012】

本発明は、このような従来の課題を解決するために為されたもので、動画像から時間的に変化する対象を確実に検出することができる情報処理装置などを提供することを目的とする。更に具体的には、本発明は、例えば、背景が複雑であるような場合や、対象が突然出現或いは消滅したり対象の一部が隠れたりするような場合においても、確実に対象を検出することが可能な情報処理装置などを提供することを目的とする。

10

【0013】**【課題を解決するための手段】**

上記目的を達成するため、本発明に係る情報処理装置では、次のようにして、動画像から時間的に変化する対象を検出する。

すなわち、変化画素検出手段が、動画像のフレームを構成する複数の画素の中から、画素値が時間的に変化する画素を検出する。次に、変化画素領域検出手段が、変化画素検出手段による検出結果に基づいて、画素値が時間的に変化する画素領域を検出する。次に、変化画素領域エッジ検出手段が、変化画素領域検出手段により検出される画素領域のエッジを検出する。そして、変化対象検出手段が、変化画素領域エッジ検出手段による検出結果に基づいて、時間的に変化する対象を検出する。

20

【0014】

従って、上記のような処理の組み合わせにより、動画像から時間的に変化する対象を確実に検出することができる。

また、後述する本発明の実施例に示されるように、本発明に係る情報処理装置では、例えば、背景が複雑であるような場合や、対象が突然出現或いは消滅したり対象の一部が隠れたりするような場合においても、確実に対象を検出することを可能とすることができる。

【0015】

ここで、動画像としては、種々なものが用いられてもよく、例えば、時間的に連続した複数の画像フレームから構成されるものが用いられる。また、それぞれの画像フレームは、例えば、所定の複数の画素から構成される。それぞれの画素は画素値を有する。

30

【0016】

また、検出する対象としては、種々なものが用いられてもよい。

また、対象の時間的な変化としては、種々な変化であってもよく、例えば、対象が動くという変化や、対象自体の大きさや形や濃度が変わるという変化などであってもよい。

【0017】

また、画素値が時間的に変化する画素領域としては、例えば、画素値が時間的に変化する1又は複数の画素から構成される。この場合、変化画素検出手段により検出される画素から、1個の画素領域が検出されてもよく、或いは、複数個の画素領域が検出されてもよい。また、変化画素検出手段により検出される画素以外の画素を含む画素領域が検出されてもよい。

40

また、エッジとしては、例えば、対象と背景との境界を示すものが検出される。

【0018】

以下で、更に、本発明に係る構成例を示す。

本発明に係る情報処理装置では、一構成例として、変化画素検出手段は、時間的に隣接する動画像のフレームに対して一時的差分(TD: Temporal Differencing)法による処理を行って、画素値が時間的に変化する画素を検出する。

従って、一時的差分法により、画素値が時間的に変化する画素を適切に検出することができる。

50

【0019】

ここで、時間的に隣接する動画像のフレームとしては、例えば、連続した2つの画像フレームが用いられる。

また、一時的差分法では、例えば、2つの画像フレームにおいて、画素値が変化している画素を検出する。

【0020】

本発明に係る情報処理装置では、一構成例として、変化画素領域検出手段は、画像の分解能を低くして、画素値が時間的に変化する画素領域を検出する。

従って、分解能が低い画像を用いて、画素値が時間的に変化する画素領域を検出する処理に要する負担や時間を低減することができる。

10

ここで、画像の分解能を低くする態様としては、種々な態様が用いられてもよい。

【0021】

本発明に係る情報処理装置では、一構成例として、変化画素領域検出手段は、画像の分解能を低くして画素値が時間的に変化する画素領域を検出した後に、当該検出した画素領域を処理対象領域として、高い分解能の画像に基づいて画素値が時間的に変化する画素領域を検出する。

【0022】

従って、分解能が低い画像を用いて、画素値が時間的に変化する画素領域を検出する処理に要する負担や時間を低減することができ、その後、例えば、検出した画素領域のみを処理対象領域として、高い分解能の画像を用いて、画素値が時間的に変化する画素領域を検出することにより、処理の効率化を図ることができる。

20

ここで、画像の分解能を低くする態様や、その後高い分解能の画像を用いる態様としては、種々な態様が用いられてもよい。

【0023】

本発明に係る情報処理装置では、一構成例として、変化画素領域検出手段は、画素値が1値又は0値で表される画像（バイナリ画像）を処理対象とする場合に、次のようにして、画像の分解能を低くする。

すなわち、まず、画像フレームを構成する複数の画素に関して、所定の数の画素のまとまり毎にブロック化する。次に、画素値が1値である画素が画素値が0値である画素と比べて多く含まれるブロックに1値を割り当てる一方、画素値が0値である画素が画素値が1値である画素と比べて多く含まれるブロックに0値を割り当てる。そして、各ブロックを各画素とみなす。これにより、画像の分解能を低くする。

30

【0024】

従って、ブロック化を用いた簡易な処理により、画像の分解能を低下させることができる。具体的には、分解能低下前の画像（元の画像）における所定の数の画素のまとまりが、1個のブロックに相当し、分解能低下後の画像における1個の画素に相当する。

【0025】

ここで、1個のブロックを構成する所定の数の画素のまとまりとしては、種々なものが用いられてもよい。また、当該所定の数としては、種々な数が用いられてもよい。

また、例えば、1個のブロックに含まれる画素の数が偶数であり、画素値が1値である画素の数と画素値が0値である画素の数とが等しい場合には、当該ブロックに割り当てられる値（画素値）としては、1値が用いられてもよく、或いは、0値が用いられてもよい。

40

【0026】

本発明に係る情報処理装置では、一構成例として、変化画素領域エッジ検出手段は、E P I C (edge-projected integration of visual cues) 法を用いて、変化画素領域検出手段により検出される画素領域のエッジを検出する。

従って、E P I C 法により、画素領域のエッジを適切に検出することができる。

【0027】

本発明に係る情報処理装置では、一構成例として、変化対象検出手段は、ランダムハフ変換 (R H T : Randomized Hough Transform) 法を用いて、時間的に変化する対象を検出す

50

る。

従って、ランダムハフ変換法により、時間的に変化する対象を適切に検出することができる。

【0028】

本発明に係る情報処理装置では、例えば、動画像に時間的に変化する複数の対象が含まれるような場合に、一構成例として、次のような処理を行う。

すなわち、変化画素領域検出手段は、ラベリング法を用いて、画素値が時間的に変化する複数の画素領域を識別して検出する。また、変化画素領域エッジ検出手段は、変化画素領域検出手段により検出されるそれぞれの画素領域のエッジを検出する。また、変化対象検出手段は、それぞれの画素領域毎に、変化画素領域エッジ検出手段による検出結果に基づいて、時間的に変化する対象を検出する。

10

【0029】

従って、時間的に変化する複数の対象が動画像中に存在するような場合においても、これら複数の対象を確実に検出することができる。

ここで、複数の対象の数としては、種々な数が用いられてもよい。

また、複数の画素領域としては、種々な数が用いられてもよい。

また、それぞれの画素領域を識別する仕方としては、種々な仕方が用いられてもよく、例えば、番号などの識別情報を付する仕方などを用いることができる。

【0030】

本発明に係る情報処理装置では、例えば、変化対象検出手段により検出する対象は、パラメータを用いて表現される。

20

ここで、パラメータを用いて表現される対象としては、例えば、円や楕円などの曲線や、或いは、直線や、正方形や、長方形や、台形などの種々な図形が用いられてもよい。

【0031】

本発明では、以上に示したような種々な処理を行う方法を提供することも可能である。

一例として、本発明に係る情報処理方法では、次のようにして、動画像から時間的に変化する対象を検出する。

すなわち、動画像のフレームを構成する複数の画素の中から画素値が時間的に変化する画素を検出し、当該検出結果に基づいて画素値が時間的に変化する画素領域を検出し、当該検出される画素領域のエッジを検出し、当該検出結果に基づいて時間的に変化する対象を検出する。

30

【0032】

本発明では、以上に示したような種々な処理を行うためのプログラムを提供することも可能である。

一例として、本発明に係るプログラムでは、動画像のフレームを構成する複数の画素の中から画素値が時間的に変化する画素を検出する機能と、当該検出結果に基づいて画素値が時間的に変化する画素領域を検出する機能と、当該検出される画素領域のエッジを検出する機能と、当該検出結果に基づいて時間的に変化する対象を検出する機能をコンピュータにより実現する。

ここで、プログラムとしては、種々なものが用いられてもよい。

40

【0033】

本発明では、以上に示したような種々な処理を行うためのプログラムを記憶した記憶媒体を提供することも可能である。

一例として、本発明に係る記憶媒体では、コンピュータに実行させるプログラムを当該コンピュータの入力手段により読み取り可能に記憶している。

そして、当該プログラムは、動画像のフレームを構成する複数の画素の中から画素値が時間的に変化する画素を検出する処理と、当該検出結果に基づいて画素値が時間的に変化する画素領域を検出する処理と、当該検出される画素領域のエッジを検出する処理と、当該検出結果に基づいて時間的に変化する対象を検出する処理を当該コンピュータに実行させる。

50

【 0 0 3 4 】

ここで、コンピュータとしては、種々なものが用いられてもよい。
また、プログラムとしては、種々なものが用いられてもよい。
また、コンピュータの入力手段としては、種々なものが用いられてもよい。
また、記憶媒体としては、種々なものが用いられてもよく、例えば、フロッピー（登録商標）ディスクや、C D（Compact Disc）- R O M（Read Only Memory）や、D V D（Digital Video Disc）などを用いることができる。

【 0 0 3 5 】

【 発明の実施の形態 】

本発明に係る一実施例を図面を参照して説明する。
本例では、動画像に含まれる時間的に変化する円を検出する処理に本発明を適用した場合を示す。

10

図 1 には、本発明の一実施例に係る情報処理装置が備える機能の構成例を示してある。

【 0 0 3 6 】

本例の情報処理装置には、一時的差分（T D : Temporal Differencing）法による処理や閾値処理を行う変化画素検出機能 1 と、ブロック化処理やフィルタ処理や多段階分解能の連結要素ラベリング処理を行う変化画素領域検出機能 2 と、E P I C（edge-projected integration of visual cues）法による処理を行う変化画素領域エッジ検出機能 3 と、ランダムハフ変換（R H T : Randomized Hough Transform）法による処理を行う変化対象検出機能 4 が備えられている。それぞれの機能 1 ~ 4 により行われる処理の具体例については、後述する。

20

【 0 0 3 7 】

なお、本例の情報処理装置は、例えば、C P U（Central Processing Unit）やメモリを有したコンピュータを用いて構成される。メモリとしては、例えば、R A M（Random Access Memory）や、R O M（Read Only Memory）や、外部のハードウェア装置などを用いることができる。

【 0 0 3 8 】

また、本例の情報処理装置の構成としては、種々な構成が用いられてもよく、例えば、1 つの装置から構成されてもよく、或いは、複数の装置を組み合わせで構成されてもよい。また、本例の情報処理装置は、例えば、汎用の装置として構成されてもよく、或いは、専用

30

【 0 0 3 9 】

本例の情報処理装置では、概略的には、（ 1 ）一時的差分法を用いて動画像に含まれる時間的に変化する画素を検出する処理と、（ 2 ）多段階分解能（multi-resolution）の連結要素（connected component）のラベリング（labeling）を用いて動画像に含まれる時間的に変化する画素領域を検出する処理と、（ 3 ）E P I C法を用いて動画像に含まれる時間的に変化する画素領域のエッジを検出する処理と、（ 4 ）ランダムハフ変換法を用いて動画像に含まれる時間的に変化する対象物を検出する処理を行う。

【 0 0 4 0 】

以下で、本例の情報処理装置により行われるそれぞれの処理（ 1 ）～（ 4 ）について説明する。

40

また、その後、（ 5 ）実験結果の一例と、（ 6 ）まとめ等を示す。

【 0 0 4 1 】

（ 1 ）一時的差分法を用いて動画像に含まれる時間的に変化する画素を検出する処理について説明する。この処理では、画素の光強度（例えば、明るさ）が 0 ~ 2 5 5 の値（画素値）で表される画像を処理対象として用いる。

目標物の追跡では、動作情報（動きに関する情報）は重要な役割を果たす。適切な動作情報を得ることができれば、例えば、後処理として行われるマッチングにおける検索空間を減少させるために非常に役立つ。

50

【 0 0 4 2 】

目標物を追跡する多くのシステムは、カルマンフィルタ (Kalman filter) に基づいている。しかしながら、これらは単一モードのガウシアン密度 (unimodal Gaussian densities) に基づいていることから、同時に他の動作の予測を為すことができないため、限られた範囲でしか使用することができない (例えば、非特許文献10参照。)。例えば、目標物が点であるような単純な場合には、カルマンフィルタを拡張して複数のデータの結合を扱うことがうまく機能するが、目標物自体が変動してしまう (“no-rigid” である) ような場合には単純に拡張することはできない。

【 0 0 4 3 】

また、従前のカルマンフィルタより優れていて確実な追跡を行うことができる “Condensation” というランダムなアルゴリズムが開発されている (例えば、非特許文献10参照。)。しかしながら、“Condensation” のアルゴリズムでは、観測式や状態式を立てることが必要であり、実時間への適用の点で広い使用が制限されてしまう。

【 0 0 4 4 】

そこで、本例では、一時的差分法を使用する。

一時的差分法は、目標物を追跡するための基本的な方法の1つである。一時的差分法には多くの変形があるが、最も簡単なものは、連続した映像 (動画像) フレームを取り出して絶対的な変化 (違い) を決定する方法である。このような変化を決定するためには、閾値に関する関数が使用される。

【 0 0 4 5 】

例えば、 n 番目のフレームにおける座標 (u, v) に位置する画素の光強度 (画素値) が $I[n](u, v)$ であるとすると、画素値に関する変化の関数 $\Delta[n](u, v)$ は式1のように表される。なお、座標系としては、直交座標系を用いている。また、式1で、“ $||$ ” は絶対値を表す。

【 0 0 4 6 】

【数1】

$$\Delta[n](u, v) = |I[n](u, v) - I[n-1](u, v)| \quad \dots (式1)$$

【 0 0 4 7 】

そして、閾値処理を行うことにより、動作画像 $M[n](u, v)$ が式2に示されるようにして得られる。なお、式2で、 T は画素値に関する閾値である。当該閾値 T としては、例えば処理の状況などに応じて、種々な値が設定されてもよい。

【 0 0 4 8 】

【数2】

$$\begin{aligned} \Delta[n](u, v) \geq T \text{ のとき、 } M[n](u, v) &= I[n](u, v) \\ \Delta[n](u, v) < T \text{ のとき、 } M[n](u, v) &= 0 \end{aligned} \quad \dots (式2)$$

【 0 0 4 9 】

(2) 多段階分解能の連結要素ラベリングを用いて動画像に含まれる時間的に変化する画素領域を検出する処理について説明する。この処理では、画素の光強度 (例えば、明るさ) が $0 \sim 255$ の値 (画素値) で表される画像に対して、例えば閾値 = 127 として、画素値が閾値以下である画素を “0” 値へ変換する一方、画素値が閾値以上である画素を “1” 値へ変換することにより、画素値が1値又は0値で表される画像 (バイナリ画像) を得て、当該バイナリ画像を処理対象として用いる。

【 0 0 5 0 】

10

20

30

40

50

上記した処理(1)により動作画像 $M[n](u, v)$ が決定された後に、連結要素の基準を用いて、動いている部分を動作領域 $R[n](i)$ としてまとめる。ここで、 i は、動作領域を識別する番号を示す。また、動作領域 $R[n](i)$ は、動いている部分の画素の集合に相当する。

【0051】

一般に、実時間に適用する場合には、連結要素の分析は処理時間を要するものである。そこで、本例では、連結要素の分析の能率及び確実性を改善するために、多段階の分解能を使用する連結要素の解析のアルゴリズムを新たに開発した。本例のアルゴリズムでは、まず、比較的分解能が低く粗い画像を用いて連結要素の解析を行い、その後、より分解能が高く細かい一連の画像を用いて、位置を正確にしていく。

10

【0052】

また、本例では、例えばサンプリングの数を減らして画像の画素数を減らすのではなく、“smearing”というブロック化技術を使用する。具体的には、分解能が1段階低減させられた画像と、これと比べて分解能が1段階上である画像を考えた場合に、分解能が1段階低減させられた画像のそれぞれの画素が、分解能が1段階上である画像における $(N \times M)$ 個の画素に相当するようにする。

【0053】

そして、分解能が1段階低減させられた画像のそれぞれの画素の画素値としては、例えば、当該画素に相当する前記 $(N \times M)$ 個の画素のうち半分以上が1値(例えば、白)である場合には1値であるとし、そうでない場合にはつまり半分以上が0値(例えば、黒)である場合には0値であるとする。

20

【0054】

本例の多段階分解能の連結要素ラベリングの解析処理では、次の(処理手順1)~(処理手順6)のような処理が行われる。

(処理手順1)原画像の分解能のまま得られるバイナリ画像 $B[0]$ のフレームを複数のサブブロックへ分割する。ここで、各サブブロックの大きさは $(N \times M)$ 画素である。そして、各サブブロックは、当該バイナリ画像 $B[0]$ の分解能を1段階低減させた画像 $B[1]$ の各画素に相当する。なお、各サブブロックの大きさである $(N \times M)$ 画素としては、種々な大きさが用いられてもよい。

【0055】

(処理手順2)また、原画像 $B[0]$ においてそれぞれのサブブロックに含まれる0値の画素の数をカウントし、0値の画素の数のカウント値が $(N \times M) / 2$ を超えたサブブロックについては分解能が1段階低減された画像 $B[1]$ において対応する画素の画素値を0値とし、他のサブブロックについては分解能が1段階低減された画像 $B[1]$ において対応する画素の画素値を1値とする。

30

【0056】

なお、本例では、上記した(処理手順2)において、0値の画素の数をカウントしており、これは、一般的な動画画像では、多くの場合において0値の画素である確率が1値の画素である確率と比べて大きいためである。これにより、処理時間を節約することが可能である。

40

【0057】

(処理手順3)上記した(処理手順1)及び上記した(処理手順2)と同様な処理を繰り返して行うことにより、要求される分解能の画像が得られるまで、分解能が j 段階低減された画像 $B[j]$ を順次生成していく。これにより、 $j = J$ (J は、例えば、1以上の整数)となるまで処理を行うと、原画像と比べて分解能が低減された一連の J 個の画像 $B[1]$ 、 $B[2]$ 、 \dots 、 $B[J]$ が得られる。

【0058】

(処理手順4)次に、分解能が最も低く粗い画像 $B[J]$ に対して、連結要素の解析を行い、そして、サイズフィルタを用いて雑音を除去しつつ、粗いラベリングの結果を取得する。ここで、ラベリングでは、時間的な変化があった画素を1以上の動作領域 $R[n]($

50

i) へまとめることが行われ、概略的には、隣接する画素或いは近くに位置する画素は同一の動作領域 $R[n](i)$ にまとめ、他の離隔した画素は異なる動作領域 $R[n](i)$ にまとめる。

【0059】

また、サイズフィルタによる処理では、例えば、ラベリングの後に、当該ラベリングによりまとめられた画素領域（動作領域 $R[n](i)$ ）に含まれる画素の数が所定の閾値と比べて小さい場合には、当該画素領域は雑音であるとみなして、当該画素領域を処理対象から除去することが行われる。当該閾値としては、例えば処理の状況などに応じて、種々な値が用いられてもよい。

【0060】

（処理手順5）また、必要に応じて、分解能が J 段階低減された画像 $B[J]$ でラベリングが為された画素領域（動作領域 $R[n](i)$ ）について、分解能が1段階上である画像 $B[J-1]$ 或いは更に分解能が高い画像に対して、上記した連結要素の解析や上記したラベリングなどを行うことも可能である。このように、比較的分解能が低く粗い画像で対象となる領域を絞り込んで、その後、比較的分解能が高く細かな画像で当該領域のみに対して処理を行うと、総じて、処理の量や時間を短くして、効率的に精密な処理を行うことが可能である。

（処理手順6）このようにして、要求される分解能の処理結果が得られるまで、処理を継続する。

【0061】

ここで、本例では、連結要素のアルゴリズムにおいて、6連結の隣接の構造が採用されている。本例の6連結の隣接構造では、上側（N）、右側（E）、右下側（SE）、下側（S）、左側（W）、左上側（NW）の6個の方向に位置するデータを隣接するものとみなす。

なお、本例では、6方向を隣接とする6連結を用いたが、例えば、上下左右の4方向を隣接とする4連結や、上下左右及び斜め4方向を隣接とする8連結などのように、他の種々な連結が用いられてもよい。

また、走査（スキャン）は、左から右へ、上から下へ、行われる。

【0062】

また、ブロック化処理では、画像の分解能を低減させることができるとともに、上記した処理（1）における一時的差分法により発生した雑音を除去することができる。

また、本例では、一時的差分法により欠けた輪郭が生じた場合には、当該欠けた部分を埋めるために、形態論的な輪郭フィリング処理を行う。

また、本例では、一時的差分法により得られた輪郭の内部の画素の画素値を0値から1値へ変更する内部フィリング処理を行う。

また、ラベリングは、例えば、このような輪郭フィリング処理や内部フィリング処理の後に行われ、また、このような処理の前や中間といった種々な時期に行われてもよい。

【0063】

ここで、図2には、領域Aが処理対象である場合における、6連結領域の一部の領域B、C、Dを示してあるとともに、走査処理の方向を矢印で示してある。なお、それぞれの領域A、B、C、Dは、処理対象としている分解能の画像におけるそれぞれの画素に相当する。

また、図3には、領域Aが処理対象である場合においてラベリングを行う処理を記述した“pseudo-Pascal”によるアルゴリズムの一例を示してある。

【0064】

図3に示したアルゴリズムでは、まず、領域Aの画素値が0値である場合には何もしない。

一方、領域Aの画素値が1値である場合には、次のような処理を行う。

すなわち、領域Dにラベルが付されている場合には、領域Dのラベルを領域Aにコピーして付する。また、領域Dにラベルが付されてなく、領域Bと領域Cの両方に同一のラベル

10

20

30

40

50

が付されている場合には、当該ラベルを領域 A にコピーして付する。また、領域 D にラベルが付されてなく、領域 B と領域 C に互いに異なるラベルが付されている場合には、領域 B のラベルを領域 C のラベルへ変更して当該ラベルを領域 A にコピーして付することにより、これら 3 つの領域 A、B、C のラベルを同一とする。また、領域 D にラベルが付されてなく、領域 B と領域 C の一方のみにラベルが付されている場合には、当該ラベルを領域 A にコピーして付する。また、領域 D にラベルが付されてなく、領域 B と領域 C のいずれにもラベルが付されていない場合には、新たなラベルを選択して領域 A に付する。

【 0 0 6 5 】

1 回目の走査が終了すると、同一の目標物に属するとみなされる画素には同一のラベルが付される。

10

また、必要に応じて、2 回目の走査を行って、例えば “ r e m a r g e ” により、離れてはいるが近くに位置する 1 値の領域（画素）の間に存在する 0 値の領域（画素）を 1 値の領域（画素）へ変更する補間処理が為され、これにより、同一の目標物には同一のラベルが付されることが保証される。

【 0 0 6 6 】

具体的には、ラベリングでは、例えば、画像フレーム中に P 個の動作領域 $R[n](1)$ 、 $R[n](2)$ 、 \dots 、 $R[n](P)$ が存在する場合には、それぞれに異なるラベルが付される。

なお、画像フレーム中に対象物として複数の円が存在する場合には、例えば、隣接する円が十分に離れている場合には各動作領域 $R[n](1)$ 、 $R[n](2)$ 、 \dots 、 $R[n](P)$ にそれぞれ 1 個の円が含まれるが、隣接する円が見かけ上繋がっていたり重なっているような場合には、1 個の動作領域 $R[n](i)$ に 2 以上の円が含まれることもあり得る。

20

【 0 0 6 7 】

また、2 回目の走査では、各動作領域 $R[n](i)$ について、画素の総数に相当する面積及び重心を容易に得ることができる。当該面積からは円の半径が得られ、当該重心からは X - Y 直交座標系を考えた場合における重心の X 座標の値及び Y 座標の値が得られ、これら 3 つのパラメータにより円の予測された状態が定義される。これら 3 つのパラメータ（半径或いは面積、重心の X 座標、重心の Y 座標）は、後処理であるエッジ検出の処理で非常に役に立つ。

30

【 0 0 6 8 】

なお、ブロック化処理や、輪郭フィリング処理や、内部フィリング処理や、連結要素の解析処理や、ラベリング処理などを行う対象となる画像の分解能としては、必ずしも本例で示したものに限定されず、それぞれ、種々な分解能が用いられてもよい。また、例えば、全ての処理が原画像のように分解能が高い画像に対して行われてもよい。

【 0 0 6 9 】

(3) E P I C 法を用いて動画像に含まれる時間的に変化する画素領域のエッジを検出する処理について説明する。この処理では、上記した処理 (2) による処理結果を原画像と同一の光強度（画素値）を有する画像へ反映させて、再び、画素の光強度（例えば、明るさ）が 0 ~ 2 5 5 の値（画素値）で表される画像を処理対象として用いる。

40

【 0 0 7 0 】

まず、エッジ検出に関して、関連する技術を説明する。

円を検出するための最も知られている技術は、ハフ変換（H T : Hough Transform）である（例えば、非特許文献 6 参照。）。他の方法と比べて、ハフ変換は非常に確実性がある。基本的な方法の変形では、計算時間を低減するために複数ステップのアプローチが使用され、例えば、ランダムハフ変換や（例えば、非特許文献 6 参照。）、2 次元の累算器のアレイを使用するモディファイド・ハフ変換（m o d i f i e d H T）があり（例えば、非特許文献 1 1 参照）、或いは、幾何学的な対称性を使用する（例えば、非特許文献 1 2 参照。）。

【 0 0 7 1 】

50

しかしながら、これらの方法では、フレーム速度を高めることが難しい。例えば、処理時間は、エッジマップの画素の数に非常に強く依存する。

また、円を追跡する場合には、アルゴリズムでは前回の画像と比較して円の輪郭を見つけることが可能である。円を追跡するために使用することが可能な技術の1つとして、輪郭追跡によるものがある。精密で簡潔なアルゴリズムとして、“snake”という動的な輪郭のモデルがある（例えば、非特許文献13参照。）。

【0072】

この“snake”のアルゴリズムでは、画像から得られる内部の力と外部の力にスプラインを従わせ、当該スプラインを曲線（例えば、円）に合わせる。この方法は実時間で機能するが、例えば、対象物が複雑な背景の前を動いて、輪郭の勾配が大きく変化して符号が変化するような場合には、追跡において問題が発生する。

10

【0073】

近年、古典的な“snake”のアルゴリズムは、改良されてきている（例えば、非特許文献14参照。）。これらの輪郭追跡は、明白に円に使用されたことはない。しかしながら、これらの輪郭追跡は、検出される輪郭に円を合わせれば、実行可能な技術であると考えられる。

また、例えばSSD追跡（SSD-Tracking）のように（例えば、非特許文献15参照。）、領域を基礎とする追跡では、円が対象物のより大きな表面の一部である場合や、円自体が良好なテクスチャ及び妥当な大きさを有している場合にのみ、確実性がある。

20

【0074】

次に、本例で使用するEPI C法による処理について説明する。

EPI C法の基本的な考えは、巧みに配置された追跡線に対して低レベルな1次元の画像解析を行うことと、輪郭のエッジを速く信頼性高く見つけ直すための円の幾何学的なモデルとの組み合わせである（例えば、非特許文献1参照。）。EPI C法では、追跡線に沿ってエッジの画素を非常に速く見つけることができ、本例では、検出された動作領域 $R[i]$ の中から円の輪郭のエッジを確実に信頼性高く見つけることができる。

【0075】

EPI C法では、概略的には、それぞれのサイクルにおいて、次の（処理手順1）～（処理手順2）のような処理が行われる。

30

（処理手順1）前回における円の状態から、円の新しい状態を予測し、予想される輪郭に沿って複数である m 個の追跡線を置く。各追跡線は、予め設定された長さ L を有し、予め設定された一定の幅を有する。なお、当該幅としては、例えば、1画素の幅が用いられる。

また、第1回目のサイクルでは、前回における円の状態として、例えば、上記した処理（2）において得られる3つのパラメータにより予測された円の状態が用いられる。

【0076】

（処理手順2）次に、勾配とモード値を用いた処理により、それぞれの追跡線でエッジを見つける。なお、勾配は、画素の光強度（画素値）の勾配を示す。また、モードは、最も確率が高いものを示す。具体的には、例えば、或る追跡線上に複数の画素があって、これら複数の画素のそれぞれについての画素値の勾配において、“A”という値の勾配を有する画素が最も多い（最も確率が高い）場合には、モード値を当該“A”とする。

40

【0077】

以下で、更に、EPI C法による処理を説明する。

それぞれの追跡線の状態は、位置（position）と、方向（orientation）と、モデル値（model value）を含む。これらの値は、追跡線を置くときに、更新される。また、本処理では、円の状態は、上記した円の3つのパラメータと、これに加えて、全ての追跡線に関する平均的なモード値により表される。

【0078】

本例では、円（例えば予測された円）のパラメータから、追跡線を置く位置がつけられ

50

る。通常、追跡線の数は5以上であることが必要であり、追跡線の幅は1画素である（例えば、非特許文献1参照。）。

ここで、図4には、円と、追跡線を置く主な配置の一例を示してある。

同図の例では、追跡線の数は8である。これらの追跡線は、円の輪郭に対して垂直（直交）の方向に向いており、円の境界線の周囲に等しい角度間隔で配置されている。

【0079】

次に、追跡線が置かれた後に、それぞれの追跡線のエッジが見つけれられる。

ここで、それぞれの追跡線は、関連する状態ベクトルを有している。当該状態ベクトルは、追跡線の midpoint の座標値（ x 、 y ）と、方向 と、2つのモード値 $m[\text{left}]$ 、 $m[\text{right}]$ を有している。位置（ x 、 y ）と方向 は、配置処理（ $\text{placement procedure}$ ）により決定される。

10

【0080】

そして、正しいエッジが見つかった追跡線については、次回の追跡のサイクルにおいても今回の追跡のサイクルにおけるモード値 $m[\text{left}]$ 、 $m[\text{right}]$ が継続される。つまり、或る追跡線により有効なエッジを見つけた場合には、次回のサイクルのモード値 $m[\text{left}]$ 、 $m[\text{right}]$ としても、今回のサイクルの当該追跡線によるものと同じの値が使用される。他の場合には、円の実際の輪郭上にあるエッジが決定され、当該エッジに対するモード値 $m[\text{left}]$ 、 $m[\text{right}]$ が計算し直される。

【0081】

具体的には、それぞれのサイクルにおける上記した（処理手順2）には、次のような2つの操作（操作1）、（操作2）が含まれる。

20

（操作1）エッジを見つける。本例では、 (8×1) の大きさのプレウィットフィルタ（ Prewitt filter ）を用いて、追跡線に沿った光強度 $I(z)$ の一時微分 $\text{grad} I(z)$ を計算する。ここで、光強度 $I(z)$ は、追跡線に沿った方向の位置 z における画素値を表す。

【0082】

そして、それぞれの追跡線において、設定された閾値を用いることにより、全てのエッジ候補 $\text{DM}(k)$ をを見つける。ここで、勾配 $\text{grad} I(z)$ の大きさが閾値を超える場合に当該勾配 $\text{grad} I(z)$ の値をエッジ候補 $\text{DM}(k)$ とする。また、閾値としては、種々な値が用いられてもよく、例えば、4などが設定される。また、 k は、エッジ候補 $\text{DM}(k)$ の番号を示す。

30

また、エッジ候補 $\text{DM}(k)$ に対応する位置は、エッジが存在する位置の候補を定義し、追跡線に沿った位置（間隔）を定義する。追跡線が有する2つの端は、最も左端の位置と最も右端の位置として使用される。

【0083】

（操作2）ヒストグラムの技術を用いて、エッジ候補 $\text{DM}(k)$ の位置に対して左側と右側とのそれぞれの間隔について、モード値 $m[\text{left}]$ 、 $m[\text{right}]$ を計算する。本例のヒストグラムでは、1つのビン（ bin ）は5個分の画素値（本例では、5個分の画素値に相当する勾配の値）から構成されており、これにより確実な結果が得られる。

【0084】

40

ここで、エッジ候補 $\text{DM}(k)$ の位置に対して左側と右側のそれぞれの間隔におけるモード値 $m[\text{left}]$ 、 $m[\text{right}]$ としては、最大のカウンタ値（画素の数）に相当するビンの値（勾配の値）が設定される。つまり、追跡線上の左側の間隔において、或るビンに対応した勾配範囲に含まれる勾配を有する画素の数が最も多い場合には当該ビンに対応した勾配を $m[\text{left}]$ とし、同様に、追跡線上の右側の間隔において、或るビンに対応した勾配範囲に含まれる勾配を有する画素の数が最も多い場合には当該ビンに対応した勾配を $m[\text{right}]$ とする。

【0085】

このように、エッジ候補 $\text{DM}(k)$ の位置に対する間隔の左右の向きに基づいて、それぞれの番号 k のエッジ候補 $\text{DM}(k)$ について、左側のモード値 $m[\text{left}]$ と右側のモ

50

ード値 $m[\text{right}]$ が割り当てられる。本例では、常に、円の中心に近い方が追跡線の左側であるとし、円の外側の領域に向かっていている方が追跡線の右側であるとする。

これにより、それぞれのエッジ候補 $DM(k)$ は、2個のモード値 $m[\text{left}]$ 、 $m[\text{right}]$ を有する。

【0086】

次に、エッジ候補 $DM(k)$ に係るエッジが円の輪郭上にある実際のエッジである確からしさ $L(k)$ は、式3～式7のように表される。ここで、 w_g は、情報(キュー) C_g の重みである。また、式5において、 \max は、最大値を表す。また、式6及び式7において、 t 、 $(t-1)$ はサイクルの数を表す。

【0087】

【数3】

$$L(k) = \frac{1}{W} \sum_{g=1}^4 w_g C_g, \quad g=1,2,3,4$$

$$W = \sum_{g=1}^4 w_g$$

・・・(式3)

【0088】

【数4】

$DM(k) > 8$ である場合、 $C_1 = 1$ 、 $w_1 = 1$

他の場合、 $C_1 = 0$

・・・(式4)

【0089】

【数5】

$\max(DM(k))$ である場合、 $C_2 = 1$ 、 $w_2 = 0.3$

他の場合、 $C_2 = 0$

・・・(式5)

【0090】

【数6】

$$C_3 = 1 - \{ (m[\text{left}]) [t] - (m[\text{left}]) [t-1] \} / 255$$

・・・(式6)

【0091】

【数7】

$$C_4 = 1 - \{ (m[\text{right}]) [t] - (m[\text{right}]) [t-1] \} / 255$$

・・・(式7)

【0092】

ここで、上記したそれぞれのキュー $C_1 \sim C_4$ の機能は、次の通りである。

すなわち、キュー C_1 は、閾値を超える全てのエッジ候補 $DM(k)$ を選択し、これは、雑音により発生するエッジ候補 $DM(k)$ を削除する目的のみを有する。また、キュー C_2 は、最大となるエッジ候補 $DM(k)$ を選択するものである。また、キュー C_3 及びキュー C_4 は、前回のエッジ候補 $DM(k)$ と類似するモード値 $m[\text{left}]$ 、 $m[\text{right}]$ を有するエッジ候補 $DM(k)$ を選択する。上記した式6及び上記した式7により表されるように、第 t 番目のサイクルと第 $(t-1)$ 番目のサイクルとでモード値 $m[\text{left}]$ 、 $m[\text{right}]$ が同一-或いは類似する場合には、確からしさ $L(k)$ が大

10

20

30

40

50

きくなる。

【0093】

そして、確からしさ $L(k)$ の値が最大となるエッジ候補 $DM(k)$ に係るエッジが、円の輪郭の新しい位置を示すものとし、後続して行われる円の検出のための決定機構において使用される。

本例の処理は全ての追跡線について行われ、これにより、それぞれの追跡線では1個(のみ)のエッジの位置(点)が見つけれられる。新たなエッジの位置及び新たなモード値 $m[left]$ 、 $m[right]$ の情報に基づく値が、それぞれの状態ベクトルに蓄積される。

【0094】

(4) ランダムハフ変換を用いて動画像に含まれる時間的に変化する対象物を検出する処理について説明する。この処理では、画素の光強度(例えば、明るさ)が0~255の値(画素値)で表される画像を処理対象として用いる。

本例では、図形のモデルに基づく決定アルゴリズムを用いて、追跡線のエッジから円を見つける。これは、ランダムハフ変換の考えに基づく。本例の決定アルゴリズムでは、図形のモデルとして、円を表す式が使用されており、これにより、円を表す式と最も適合する円の輪郭を見つける。

【0095】

本例の決定アルゴリズムの原理としては、全ての選択可能なエッジの中から Q 個のエッジをランダムに選択して、当該選択した Q 個のエッジを用いて円を決定する。ここで、円は3個のパラメータにより決定され得るため、 $Q=3$ と設定する。また、選択可能なエッジとしては、上記した処理(3)により見つけれられた複数の追跡線のエッジを用いる。

【0096】

なお、ランダムハフ変換を実施する詳細については、例えば、一般的に知られている方法を用いることができる(例えば、非特許文献16参照)。

また、ランダムハフ変換では、例えば、2個の異なる円の一部が重なっているような場合においても、当該重なりが小さければ、これらを2個の異なる円として検出することが可能である。

【0097】

(5) 実験結果の一例を示す。

本例の情報処理装置により動画の映像中に存在する円を追跡した結果の一例を示す。

図5には、直径が変化する動きのある円を含む、連続した動画像のフレームの一例を示してある。これは、原画像であり、同図中に示される円が追跡する対象となる実際の円である。なお、解像度は(2000×2000)[dpi]であり、フレーム速度は7.5フレーム/秒(fps:frame per second)である。

【0098】

図6には、上記した処理(1)により一時的差分法による処理を行った結果を示してある。この結果では、円の周囲の部分に、画素値が時間的に変化する画素の領域が見られる。

図7には、図6に示した画像をバイナリ画像へ変換した結果を示してある。なお、バイナリ画像へ変換するための画素値に関する閾値としては20を用いた。図7に示した画像では、図6に示した画像と比べて、白黒が明確化されている。

【0099】

図8には、図7に示した画像に対して上記した処理(2)によりブロック化処理を行って分解能を低減させた結果を示してある。ブロック化処理は、非常に有用であり、雑音を低減することができるばかりでなく、フィリング処理や連結要素ラベリングのような後の処理の速度を高めることができる。なお、本例では、ブロックの大きさとして、(8×8)画素の大きさを用いている。また、図8では、スケール調整して300%にしてある。

【0100】

図9には、図8に示した画像に対して、輪郭の欠けている0値(本例では、黒)の部分や輪郭の内部の0値(本例では、黒)の部分に相当するホールをフィリング処理により1値

10

20

30

40

50

(本例では、白)へ変換した結果を示してある。図9に示した画像では、図8に示した画像中に見られる円の輪郭の欠けた部分も1値(白)へ変更されており、当該円の内部も1値(白)へ変更されている。

また、図9に示した画像に対して、上記した処理(2)により連結要素ラベリングの処理が行われ、その後、上記した処理(3)によりE P I C法による処理が行われて複数の追跡線が引かれてエッジ検出が行われる。

【0101】

図10には、上記した処理(4)によりランダムハフ変換による処理を行った結果を示してある。具体的には、図5に示したのと同様な追跡対象となる円(左側の方の円)と、これと比べて少し右側にずれた円であってランダムハフ変換により見つけれられた円(右側の方の円)と、円の周囲に引かれた複数の追跡線が示されている。なお、図10では、追跡対象となる円とランダムハフ変換により見つけれられた円とが少しずれているが、これは誤差の範囲内であり、特に問題はない。追跡の精度については、例えば、処理に使用される画像の分解能などが調整されることにより、必要に応じて任意に設定されればよい。

10

【0102】

図11には、各処理に要した時間の一例を示してある。具体的には、一時的差分(TD)法による処理と閾値処理とブロック化処理には230msかかり、輪郭や内部のフィリング処理には80msかかり、連結要素ラベリング処理には10msかかり、E P I C法による処理には10msかかり、ランダムハフ変換(RHT)による処理には1msかかり、全体では331msかかった。なお、コンピュータの機能としては、P I I 733MHz及び128Mbを用いた。

20

【0103】

本例の実験結果を、例えばE P I C法とカルマンフィルタ(例えば、非特許文献1参照。)とを組み合わせた実験の結果と比較すると、本例の実験結果の方が、円が突然現れたり消えたりするような状況においても追跡に確実性があり、更に、円の一部が隠れるような状況においても追跡に確実性がある。このように、本例の情報処理装置では、円自体が変動する(non-rigidである)ような場合や、複雑な背景が存在するような場合においても、実時間で追跡を行うことができ、追跡の性能が高く、高度な正確さで円の検出や追跡を行うことができる。

【0104】

(6)まとめ等を示す。

以上のように、本例の情報処理装置では、一時的差分法とランダムハフ変換を組み合わせた方法を用いて、動的映像において、媒介変数(パラメータ)を有する曲線などを確実に追跡することができる。具体的には、本例では、(1)一時的差分法により動作領域を検出し、(2)多段階の分解能(解像度)の連結要素ラベリングを行い、(3)E P I C法によりエッジを検出し、(4)ランダムハフ変換によりマッチングを行い、これにより、曲線などを検出や追跡することができる。

30

【0105】

このように、本例の情報処理装置では、一時的差分法とランダムハフ変換を組み合わせることにより、連続した高分解能画像中における媒介変数(パラメータ)を有する曲線などを実時間で追跡することができる新しい枠組みを提供している。また、本例の情報処理装置では、簡易で効率的な多段階分解能の連結要素ラベリング技術を提供している。

40

【0106】

また、本例のように一時的差分法とランダムハフ変換を組み合わせると、動き検出により導かれて追跡が行われるため、例えば、追跡処理において予測的なフィルタリングが必要となることを回避することが可能である。ここで、一時的差分法では、曲線などが存在する大雑把な領域を検出する。

なお、実時間で完全な円を追跡することは検討されているが(例えば、非特許文献1、8参照。)、本例の処理とは異なっている。

【0107】

50

本例の処理では、例えば、動作領域を検出するために一時的差分法を用いており、静止した円についてはパラメータのテンプレート相関マッチング法を用いており、これにより、検索空間を非常に低減することができ、カルマンフィルタでは不可能である円の突然の出現や消滅を扱うことができる。

【0108】

また、本例の処理では、多段階分解能の連結要素ラベリング法を用いることにより、雑音の除去と対象物のカウントを高速化している。

また、本例の処理では、円の決定処理において、RANSACのようなサンプリング（例えば、非特許文献9参照。）ではなく、ランダムハフ変換を用いた。一般に、ランダムハフ変換の方が、RANSACのようなサンプリングと比べて、計算上の複雑さが低い。

10

【0109】

なお、本例の情報処理装置では、上記した一連の処理（1）～（4）を動画像のそれぞれの画像フレーム毎に行う実施の形態を用いることが可能であるばかりでなく、例えば、或る画像フレームについての処理で得られた円の状態に関する情報を次の画像フレームにおいて予測値として使用するような実施の形態を用いることも可能である。

【0110】

具体的には、一例として、或る画像フレームについて上記した一連の処理（1）～（4）を行い、次の画像フレームについては前回の処理結果に基づいて上記したEPIC法による処理（3）と上記したランダムハフ変換による処理（4）のみを行うようなことが可能である（例えば、非特許文献17参照。）。つまり、隣接する前後の画像フレームの間で円の状態の変化（動き）が小さいような場合には、前の画像フレームにおける円の状態に関する情報を後の画像フレームにおいて利用しても、高い精度を保つことが可能である。

20

【0111】

また、本例では、検出や追跡する対象として円を用いたが、例えば、楕円や、四角や、線などのように種々な図形が検出や追跡する対象として用いられてもよく、このように、必ずしも曲線に限られない。また、検出や追跡する対象としては、長さや大きさや位置や濃度などの種々な要素が変化してもよく、本例の処理により検出や追跡することが可能である。なお、好ましくは、比較的簡易な数式で表現することが可能な図形が検出や追跡する対象とされるのがよいが、処理時間をかければ、複雑な図形を検出や追跡することも可能である。

30

【0112】

ここで、本発明に係る情報処理装置などの構成としては、必ずしも以上に示したものに限られず、種々な構成が用いられてもよい。なお、本発明は、例えば本発明に係る処理を実行する方法或いは方式や、このような方法や方式を実現するためのプログラムなどとして提供することも可能である。

また、本発明の適用分野としては、必ずしも以上に示したものに限られず、本発明は、種々な分野に適用することが可能なものである。

【0113】

また、本発明に係る情報処理装置などにおいて行われる各種の処理としては、例えばプロセッサやメモリ等を備えたハードウェア資源においてプロセッサがROM（Read Only Memory）に格納された制御プログラムを実行することにより制御される構成が用いられてもよく、また、例えば当該処理を実行するための各機能手段が独立したハードウェア回路として構成されてもよい。

40

また、本発明は上記の制御プログラムを格納したフロッピー（登録商標）ディスクやCD（Compact Disc）-ROM等のコンピュータにより読み取り可能な記憶媒体（記録媒体）や当該プログラム（自体）として把握することもでき、当該制御プログラムを記憶媒体からコンピュータに入力してプロセッサに実行させることにより、本発明に係る処理を遂行させることができる。

【0114】

【発明の効果】

50

以上説明したように、本発明に係る情報処理装置によると、例えば、一時的差分 (TD) 法を用いて動画像のフレームを構成する複数の画素の中から画素値が時間的に変化する画素を検出し、多段階分解能の連結要素ラベリング法を用いて当該検出結果に基づいて画素値が時間的に変化する画素領域を検出し、E P I C法を用いて当該検出される画素領域のエッジを検出し、ランダムハフ変換 (RHT) 法を用いて当該検出結果に基づいて時間的に変化する対象を検出するようにしたため、動画像から時間的に変化する対象を確実に検出することができ、例えば、背景が複雑であるような場合や、対象が突然出現或いは消滅したり対象の一部が隠れたりするような場合においても、確実に対象を検出することを可能とすることができる。

【図面の簡単な説明】

10

【図 1】 本発明の一実施例に係る情報処理装置の構成例を示す図である。

【図 2】 6 連結した領域 (一部) と走査処理の一例を示す図である。

【図 3】 p s e u d o - P a s c a l によるアルゴリズムの記述の一例を示す図である。

【図 4】 円の上に追跡線を配置する一例を示す図である。

【図 5】 原画像の一例を示す図である。

【図 6】 一時的差分法により得られる画像の一例を示す図である。

【図 7】 画素値のバイナリ化により得られる画像の一例を示す図である。

【図 8】 ブロック化処理により分解能が低減させられた画像の一例を示す図である。

【図 9】 フィリング処理後の画像の一例を示す図である。

20

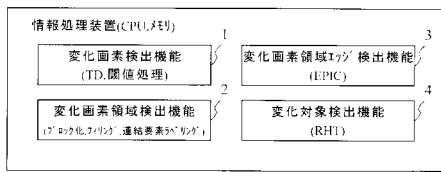
【図 10】 ランダムハフ変換法による円の検出結果の一例を示す図である。

【図 11】 各処理に要する時間の一例を示す図である。

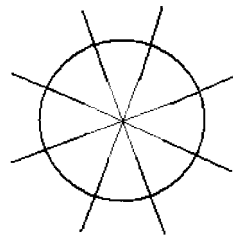
【符号の説明】

- 1・・・変化画素検出機能、 2・・・変化画素領域検出機能、
- 3・・・変化画素領域エッジ検出機能、 4・・・変化対象検出機能、

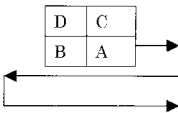
【図 1】



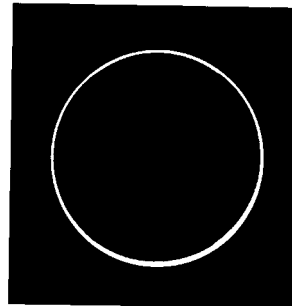
【図 4】



【図 2】



【図 5】

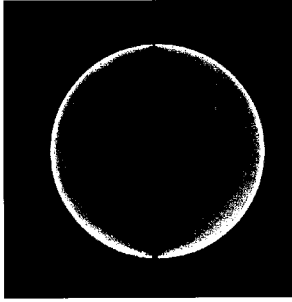


【図 3】

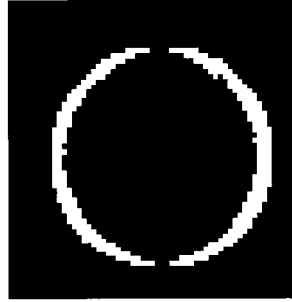
```

Begin
CASE A is Zero. NOP;
CASE D is labeled
  Copy the label to A;
  EXIT;
CASE B or C is labeled
  IF both B and C is labeled. THEN
    IF B's label equals C's label. THEN
      Copy the label to A;
      EXIT;
    ELSE IF
      Change B's label to C;
      Copy the label for A;
      EXIT;
    END IF
  END IF
  EXIT
CASE neither B nor C is labeled
  Choose a new label to A;
  EXIT
End
  
```

【図 6】



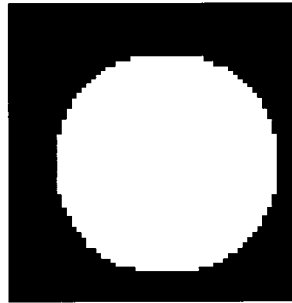
【図 8】



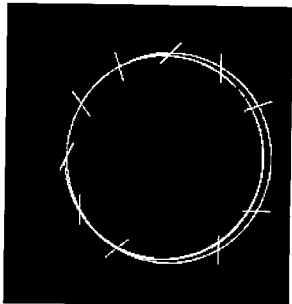
【図 7】



【図 9】



【図 10】



【図 11】

一時的差分法 閾値処理	フィルタ	連結要素 ラベル化	EPIC	RUII	全体
230ms	80ms	10ms	10ms	1ms	331ms

フロントページの続き

(56)参考文献 特開2000-069366(JP,A)

特開2001-167282(JP,A)

特開平08-185521(JP,A)

特開2003-070742(JP,A)

Markus Vincze 外2名, Fast Tracking of Ellipses using Edge-Projected Integration of Cues, Pattern Recognition, 2000. Proceedings. 15th International Conference on, IEEE Society, 2000年9月3日, Vol. 4, p. 72~75, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=902867

塩野, 黒点ランダム抽出とHough曲面の交点計算による円図形検出の一手法, 情報処理学会論文誌, 日本, 社団法人情報処理学会, 1991年2月15日, 第32巻, 第2号, p. 179~187

(58)調査した分野(Int.Cl., DB名)

G06T7/00~7/60