

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7477168号
(P7477168)

(45)発行日 令和6年5月1日(2024. 5. 1)

(24)登録日 令和6年4月22日(2024. 4. 22)

(51)Int. Cl.	F I
G 0 6 T 7/00 (2017. 01)	G 0 6 T 7/00 3 5 0 B
G 0 6 T 7/246 (2017. 01)	G 0 6 T 7/00 3 0 0 D
	G 0 6 T 7/00 3 0 0 B
	G 0 6 T 7/246

請求項の数 13 (全 19 頁)

(21)出願番号 特願2020-553348(P2020-553348)	(73)特許権者 503360115 国立研究開発法人科学技術振興機構 埼玉県川口市本町四丁目1番8号
(86)(22)出願日 令和1年10月18日(2019. 10. 18)	(74)代理人 100105924 弁理士 森下 賢樹
(86)国際出願番号 PCT/JP2019/041143	(72)発明者 石井 抱 広島県東広島市鏡山1丁目4番1号 広島 大学大学院 工学研究科内
(87)国際公開番号 W02020/080535	(72)発明者 岸 則政 東京都千代田区五番町7 国立研究開発法 人 科学技術振興機構内
(87)国際公開日 令和2年4月23日(2020. 4. 23)	(72)発明者 姜 明俊 広島県東広島市鏡山1丁目4番1号 広島 大学大学院 工学研究科内
審査請求日 令和4年9月15日(2022. 9. 15)	
(31)優先権主張番号 特願2018-196575(P2018-196575)	
(32)優先日 平成30年10月18日(2018. 10. 18)	
(33)優先権主張国・地域又は機関 日本国(JP)	

最終頁に続く

(54)【発明の名称】対象追跡方法、対象追跡システムおよび対象追跡プログラム

(57)【特許請求の範囲】

【請求項1】

第1のフレームレートで、学習的認識手法により、動画像内の少なくとも1つの対象を認識する認識ステップと、

前記第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、前記動画像内の対象を追跡する追跡ステップと、

前記動画像の実画像をバッファに記憶する記憶ステップと、

前記学習的認識手法による対象の認識が開始されたフレームで認識されたROIと、前記第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームのROIとの間の差異を補正する補正ステップであって、前記学習的認識手法による対象の認識が開始されたフレームのROIに基づき、前記バッファに記憶された実画像にテンプレートマッチングを継続した結果現在に至ったフレームのROIでカレントフレームのROIを置き換える補正ステップと、

前記置き換えられたROIに基づいて、前記カレントフレームのROI位置を初期化する初期化ステップと、を備える、

対象追跡方法。

【請求項2】

前記置き換えるステップは、前記対象の認識が開始されたフレームから前記カレントフレームまでの実画像を再生するステップを備える、

請求項1に記載の対象追跡方法。

【請求項 3】

前記実画像を再生するステップは、前記対象の認識が開始されたフレームから前記カレントフレームまでの実画像の一部を抽出したものを再生する、
請求項 2 に記載の対象追跡方法。

【請求項 4】

前記カレントフレームの R O I 位置は、前記第 1 のフレームレートで初期化されることを特徴とする、
請求項 1 乃至 3 のいずれか一項に記載の対象追跡方法。

【請求項 5】

前記学習的認識手法により認識された対象に基づいて、前記第 1 のフレームレートで、
前記テンプレートを初期化するステップをさらに備える、
請求項 1 乃至 4 のいずれか一項に記載の対象追跡方法。

10

【請求項 6】

対象の認識が開始されたフレームにおいて、前記学習的認識手法により認識された対象と、前記テンプレートマッチングにより追跡された対象とを比較するステップをさらに備え、

前記比較するステップでの比較結果が所定の値以下のときは、前記 R O I 位置は、前記テンプレートマッチングにより追跡された対象に基づいて初期化され、

前記比較するステップでの比較結果が所定の値より大きいときは、前記 R O I 位置は、
前記置き換えられた R O I に基づいて初期化されることを特徴とする、

20

請求項 1 乃至 3 のいずれか 1 項に記載の対象追跡方法。

【請求項 7】

対象の認識が開始されたフレームにおいて、前記学習的認識手法により認識された対象と、前記テンプレートマッチングにより追跡された対象とを比較するステップをさらに備え、

前記比較するステップでの比較結果が所定の値以下のときは、前記テンプレート及び R O I 位置は、前記テンプレートマッチングにより追跡された対象に基づいて初期化され、

前記比較するステップでの比較結果が所定の値より大きいときは、前記テンプレート及び R O I 位置は、前記学習的認識手法により認識された対象に基づいて初期化されることを特徴とする、

30

請求項 1 乃至 3 のいずれか一項に記載の対象追跡方法。

【請求項 8】

前記学習的認識手法による対象の認識が開始されたフレームからカレントフレームまでのフレームを記憶するステップをさらに備え、

前記置き換えるステップは、前記学習的認識手法による対象の認識が開始されたフレームの R O I 位置とテンプレートとを用いて、前記記憶したフレームに対して、テンプレートマッチングによる追跡を実行することによって行われることを特徴とする、

請求項 1 に記載の対象追跡方法。

【請求項 9】

前記学習的認識手法は、ディープラーニングを含むことを特徴とする、

請求項 1 乃至 8 のいずれか一項に記載の対象追跡方法。

40

【請求項 10】

前記動画は、高速カメラにより撮影された動画であることを特徴とする、

請求項 1 乃至 9 のいずれか一項に記載の対象追跡方法。

【請求項 11】

前記第 2 のフレームレートは、前記第 1 のフレームレートの 8 倍以上であることを特徴とする、

請求項 1 乃至 10 のいずれか一項に記載の対象追跡方法。

【請求項 12】

第 1 のフレームレートで、学習的認識手法により、動画内の少なくとも 1 つの対象を

50

認識する対象認識部と、

前記第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、前記動画像内の対象を追跡し、前記動画像の実画像をバッファに記憶する対象追跡部と、

前記学習的認識手法による対象の認識が開始されたフレームで認識されたROIと、前記第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームのROIとの間の差異を補正する補正部であって、前記学習的認識手法による対象の認識が開始されたフレームのROIに基づき、前記バッファに記憶された実画像にテンプレートマッチングを継続した結果現在に至ったフレームのROIでカレントフレームのROIを置き換える補正部と、

10

前記置き換えられたROIに基づいて、前記カレントフレームのROI位置を初期化するROI位置初期化部と、を備える、
対象追跡システム。

【請求項13】

第1のフレームレートで、学習的認識手法により、動画像内の少なくとも1つの対象を認識する認識ステップと、

前記第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、前記動画像内の対象を追跡する追跡ステップと、

前記動画像の実画像をバッファに記憶する記憶ステップと、

前記学習的認識手法による対象の認識が開始されたフレームで認識されたROIと、前記第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームのROIとの間の差異を補正する補正ステップであって、前記学習的認識手法による対象の認識が開始されたフレームのROIに基づき、前記バッファに記憶された実画像にテンプレートマッチングを継続した結果現在に至ったフレームのROIでカレントフレームのROIを置き換える補正ステップと、

20

前記置き換えられたROIに基づいて、前記カレントフレームのROI位置を初期化する初期化ステップと、をコンピュータに実行させる
対象追跡プログラム。

【発明の詳細な説明】

【技術分野】

30

【0001】

本発明は、対象追跡方法、対象追跡システムおよび対象追跡プログラムに関する。

【背景技術】

【0002】

対象追跡技術は、カメラで撮影した動画像中の特定の対象を認識し、追跡する技術である。近年の代表的な対象追跡技術には、テンプレートマッチングを用いて対象を局所的に追跡する手法（以下、「局所対象追跡手法」という）や、ディープラーニングなどの学習理論に基づく処理により対象を認識する手法（以下、「学習的認識手法」という）などがある（例えば非特許文献1）。

【先行技術文献】

40

【非特許文献】

【0003】

【非特許文献1】姜明俊、高木健、石井抱 「ディープラーニングを用いた高速対象追跡」 日本機械学会ロボティクス・メカトロニクス講演会2018 講演論文集 2A1-K17(2018)

【発明の概要】

【発明が解決しようとする課題】

【0004】

局所対象追跡手法は、対象近傍のみで局所探索を行うため、実行速度を高速化できるというメリットがある反面、追跡開始時に初期テンプレートの人為的指定が必要である。

50

このため、背景画面の複雑さや、対象のオクルージョン・姿勢変化等に起因して追跡の失敗が多発すると、追跡が困難になるという問題がある。

【0005】

一方、畳み込みニューラルネットワーク（CNN）をベースとする学習的認識手法では、近年ディープラーニングの性能が著しく向上している。高性能GPUへの並列実装と相まって、学習的認識手法は、数十fps（frame per second）といったフレームレートでのリアルタイム対象認識を実現している。

【0006】

学習的認識手法は、フレームごとに独立した形で対象認識を行うため、一度対象を見失っても瞬時に対象を再認識できるというメリットがある。しかしながら、必ずしもすべてのフレームで認識が安定するとは限らないため、認識結果が間欠的になるという問題がある。またアクティブビジョン等の機械的なフィードバックでは、数百fpsといった高速リアルタイム対象検出が要求される。しかしながら、これは現状のコンピュータ技術では困難である。このように学習的認識手法には、実行速度上の限界があるという問題がある。

10

【0007】

これに対し、近年、局所的対象追跡手法と学習的認識手法の双方のメリット活かしたハイブリッド対象追跡手法が提案されている（例えば非特許文献1）。非特許文献1に記載の技術では、あるフレームレートで実行される学習的認識手法による対象認識と、これより高いフレームレートで実行される局所的対象追跡手法による対象追跡とが並列動作している。

20

【0008】

この技術では、追跡が正常に機能している限り、関連フィルタによってROI（Region Of Interest）の位置とテンプレートとが更新され、追跡が継続される。一方、追跡が外れて対象が見失われた場合は、学習的認識手法で認識された対象を基にROI位置とテンプレートとが初期化され、これらを基に新たな追跡が行われる。これにより、オクルージョンなど、対象のアピアランス変化に対しても、ロバストな対象追跡が実現される。

【0009】

しかし実際には、学習的認識手法による対象認識は、一定の計算時間を必要とする。このため、あるフレームを基に対象認識処理を開始してから、この処理が完了して対象が認識されるまでの間に、上記の計算時間が経過する。ROI位置やテンプレート初期化の対象となるフレームは、テンプレートマッチングを継続した結果現在に至ったフレーム（以下、「カレントフレーム」という）である。従って、対象認識処理を開始したフレームは、カレントフレームに対して遅延している。例えば、動画像のサンプリングレートが500fpsで、対象認識処理を開始したフレームがカレントフレームに対して30フレーム前のフレームだった場合、この遅延は60msとなる。

30

【0010】

この遅延の間にフレーム内で対象が動くと、対象認識処理を開始したフレームと、カレントフレームとの間で、ROI位置のずれが生じる。特に、高速カメラで撮影された動画像や、フレーム内の対象の動きが大きい動画像の場合、撮影中にカメラを左右にパンしたりすることにより発生するこのずれは大きいものとなる。このような場合、ハイブリッド対象追跡手法であっても、連続的で安定した追跡が困難となってしまう。

40

【0011】

本発明はこうした状況に鑑みてなされたものであり、その目的は、高速な動画像に対しても、連続的で安定した対象追跡を実現する技術を提供することにある。

【課題を解決するための手段】

【0012】

上記課題を解決するために、本発明のある態様の対象追跡方法は、第1のフレームレートで、学習的認識手法により、動画像内の少なくとも1つの対象を認識するステップと、

50

第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、動画像内の対象を追跡するステップと、学習的認識手法による対象の認識が開始されたフレームのROIと、第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームのROIとの間の差異を補正するステップと、補正されたROIに基づいて、カレントフレームのROI位置を初期化するステップと、を備える。

【0013】

差異を補正するステップは、対象の認識が開始されたフレームからカレントフレームまでの実画像を再生するステップを備えてよい。

【0014】

実画像を再生するステップは、対象の認識が開始されたフレームからカレントフレームまでの実画像の一部を抽出したものを再生してもよい。

【0015】

カレントフレームのROI位置は、第1のフレームレートで初期化されてよい。

【0016】

実施形態は、学習的認識手法により認識された対象に基づいて、第1のフレームレートで、テンプレートを初期化するステップをさらに備えてよい。

【0017】

実施形態は、学習的認識手法により認識された対象と、テンプレートマッチングにより追跡された対象とを比較するステップをさらに備え、比較するステップでの比較結果が所定の値以下のときは、ROI位置は、テンプレートマッチングにより追跡された対象に基づいて初期化され、比較するステップでの比較結果が所定の値より大きいときは、ROI位置は、補正されたROIに基づいて初期化されてよい。

【0018】

実施形態は、学習的認識手法により認識された対象と、テンプレートマッチングにより追跡された対象とを比較するステップをさらに備え、比較するステップでの比較結果が所定の値以下のときは、テンプレートは、テンプレートマッチングにより追跡された対象に基づいて初期化され、比較するステップでの比較結果が所定の値より大きいときは、テンプレートは、学習的認識手法により認識された対象に基づいて初期化されてよい。

【0019】

実施形態は、学習的認識手法による対象の認識が開始されたフレームからカレントフレームまでのフレームを記憶するステップをさらに備え、差異を補正するステップは、学習的認識手法による対象の認識が開始されたフレームのROI位置とテンプレートとを用いて、記憶したフレームに対して、テンプレートマッチングによる追跡を実行することによって行われてよい。

【0020】

本発明の別の対象追跡方法は、第1のフレームレートで、学習的認識手法により、動画像内の対象を認識するステップと、第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、動画像内の対象を追跡するステップと、学習的認識手法による対象の認識が開始されたフレームの画像から、第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームの画像を予測するステップと、学習的認識手法による対象の認識が開始されたフレームのROIと、予測されたカレントフレームのROIとの間の差異を補正するステップと、補正されたROIに基づいて、カレントフレームのROI位置を初期化するステップと、を備える。

【0021】

学習的認識手法は、ディープラーニングを含んでよい。

【0022】

動画像は、高速カメラにより撮影された動画像であってよい。

【0023】

10

20

30

40

50

第2のフレームレートは、第1のフレームレートの8倍以上であってよい。

【0024】

本発明の別の態様は、対象追跡システムである。この対象追跡システムは、第1のフレームレートで、学習的認識手法により、動画像内の対象を認識する対象認識部と、第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、動画像内の少なくとも1つの対象を追跡する対象追跡部と、学習的認識手法による対象の認識が開始されたフレームのROIと、第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームのROIとの間の差異を補正する補正部と、補正されたROIに基づいて、カレントフレームのROI位置を初期化するROI位置初期化部と、を備える。

10

【0025】

本発明のさらに別の態様は、対象追跡プログラムである。この対象追跡プログラムは、第1のフレームレートで、学習的認識手法により、動画像内の少なくとも1つの対象を認識するステップと、第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、動画像内の対象を追跡するステップと、学習的認識手法による対象の認識が開始されたフレームのROIと、第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームのROIとの間の差異を補正するステップと、補正されたROIに基づいて、カレントフレームのROI位置を初期化するステップと、をコンピュータに実行させる。

20

【0026】

なお、以上の構成要素の任意の組合せ、本発明の表現を装置、方法、システム、記録媒体、コンピュータプログラムなどの間で変換したものもまた、本発明の態様として有効である。

【発明の効果】

【0027】

本発明によれば、高速な動画像に対しても、連続的で安定した対象追跡を実現することができる。

【図面の簡単な説明】

【0028】

【図1】実施形態に係る動画像のフレームを示す図である。

30

【図2】第1実施形態に係る対象追跡方法のフロー図である。

【図3】第2実施形態に係る対象追跡方法のフロー図である。

【図4】第3実施形態に係る対象追跡方法のフロー図である。

【図5】第4実施形態に係る対象追跡方法のフロー図である。

【図6】第5実施形態に係る対象追跡方法のフロー図である。

【図7】第6実施形態に係る対象追跡方法のフロー図である。

【図8】第7実施形態に係る対象追跡方法のフロー図である。

【図9】第8実施形態に係る対象追跡システムのブロック図である。

【図10】第9実施形態に係る対象追跡システムのブロック図である。

【図11】比較例に係る方法による対象追跡実験の結果を示す図である。図11(a)、図11(b)は、それぞれ追跡された対象のx座標、y座標を示す。

40

【図12】第6実施形態による対象追跡実験の結果を示す図である。図11(a)、図11(b)は、それぞれ追跡された対象のx座標、y座標を示す。

【図13】第9実施形態による対象追跡実験の結果を示す図である。

【発明を実施するための形態】

【0029】

実施形態を具体的に説明する前に、概要を説明する。

図1は、本発明の実施形態に係る動画像のフレームを示す。フレーム F_0 、 F_1 、 F_2 、 F_3 、...、 F_{N-1} は、追跡の対象となる動画像のフレームである。フレーム f_{1b} 、 f_{1a} 、 f_{2b} 、 f_{2a} は、ディープラーニングなどの学習的認識手法により、対象認識

50

が実行されるフレームである。フレーム $f I_{1b}$ 、 $f I_{1a}$ 、 $f I_{2b}$ 、 $f I_{2a}$ は、バッファ等に記憶された実画像である。

【0030】

図1の例では、第1の周期 t_1 で、学習的認識手法により、動画内の対象が周期的に認識される。以下、第1の周期 t_1 の逆数を第1のフレームレートと呼ぶ。学習的認識手法により対象が認識されると、この認識された対象に基づいて、ROI位置が初期化される。この初期化されたROI位置に基づいて、第2の周期 t_2 で、テンプレートを用いたテンプレートマッチングにより、動画内の対象が追跡される。以下、第2の周期 t_2 の逆数を第2のフレームレートと呼ぶ。また、学習的手法による対象認識に要する計算時間を t_3 とする。なお動画内の追跡対象は1つだけとは限らず、複数あってもよい。

10

【0031】

第1の周期 t_1 は、第2の周期 t_2 より長い。ここでは、 $t_1 = N \cdot t_2$ であるとする。 t_1 は t_2 の8倍以上であること、すなわち $N \geq 8$ であることが望ましい。換言すれば、第1のフレームレートは第2のフレームレートの8分の1以下であることが望ましい。

【0032】

f_{1b} は、ある時点で、学習的認識手法により対象の認識が開始されるフレームを示す。このときの時刻 t を $-t_3$ とする。前述の通り対象認識に要する計算時間が t_3 であるので、 $t = 0$ でこの計算が完了し、フレーム f_{1a} で対象が認識される。

【0033】

f_{1b} の時点から t_1 の時間が経過して $t = t_1 - t_3$ になると、フレーム f_{2b} に対して、次の周期の対象認識が開始される。そして、 $t = t_1$ でこの対象認識が完了する。以下同様である。

20

【0034】

$t = -t_3$ で対象の認識が開始されてから、 $t = 0$ で対象の認識が完了するまでの実画像は、バッファ等に記憶される。本例では、 $t = -t_3$ における実画像 $f I_{1b}$ から $t = 0$ における実画像 $f I_{1a}$ が、バッファ B_1 に記憶される。

【0035】

次の対象認識の周期では、 $t = t_1 - t_3$ における実画像 $f I_{2b}$ から $t = t_1$ における実画像 $f I_{2a}$ が、バッファ B_2 に記憶される。以下同様である。

【0036】

F_0 、 F_1 、 F_2 、 F_3 、... は、それぞれ $t = 0$ 、 t_2 、 $2 \cdot t_2$... で順次対象が追跡されたフレームを示す。前述の通り $t = 0$ で学習的認識手法により対象が認識されるので、この時点のフレーム F_0 のROI位置が更新される。このフレーム F_0 をカレントフレーム CF (Current Frame) と呼ぶ。

30

【0037】

時刻 $t = t_1 = N \cdot t_2$ で、次の周期の学習的認識手法による対象認識に基づいて、フレーム F_{N-1} のROI位置が新たな位置に初期化される。従って、もし $F_1 \sim F_{N-2}$ のいずれかのフレームで追跡が外れて対象が見失われた場合であっても、フレーム F_{N-1} における対象認識により、 F_N 以降のフレームで追跡を再開することができる。

【0038】

図1の下段に示されるように、学習的認識手法による対象認識が開始されたフレーム f_{1b} は、カレントフレーム CF に対し、時間にして t_3 だけ遅延している。もしこの t_3 の間に動画内で対象が大きく動くと、フレーム f_{1b} 内の対象位置 (ROI位置) と、カレントフレーム CF 内の対象位置との間に大きな差異が生じる。

40

【0039】

実施形態は、この差異を補正する。実施形態は、バッファ B_1 内に記憶したフレーム $f I_{1b}$ から $f I_{1a}$ までの実映像を再生して対象を追跡し、ROI位置を補正する。この実映像の再生は、記憶したすべてのフレームを用いた高速再生であってもよく、記憶したフレームから抽出したフレームのサブセットのコマ送りの再生であってもよい。

【0040】

50

図1の中段に示されるように、 $t = -t_3$ ではフレーム $f_{I_{1b}}$ の下左隅にあったROI I_{1b} は、 $t = 0$ ではフレーム $f_{I_{1a}}$ の下右隅の I_{1a} に移動している。一方、図1の下段に示されるように、 $t = 0$ で対象認識されたフレーム f_{1a} のROI I_1 は、 $t = -t_3$ におけるフレーム f_{1b} に基づいて計算されているので、フレーム f_{1a} の下左隅にある。このROI I_1 は、実画像に I_{1a} によって補正される。そしてこの補正されたROIに基づいて、カレントフレームCFのROI位置が更新される。

【0041】

このようにして、更新されるROI位置は、カレントフレームCF内の対象位置から遅延したものではなく、常にカレントフレームCFに追跡したものとなる。従って、滑らかに途切れのない対象追跡を実現することができる。

10

【0042】

以下、図2～図8を参照して、本発明の種々の実施形態を説明する。実施形態および変形例では、同一または同等の構成要素には同一の符号を付すものとし、適宜重複した説明は省略する。図2～図8のフロー図は、第1のフレームレートで、学習的手法により対象が認識される周期、すなわち第1の周期についての1周期分のフローを示すものである。

【0043】

(第1実施形態)

図2は、第1実施形態に係る対象追跡方法のフロー図である。

「開始」で、ある第1の周期についての学習的手法による対象の追跡が開始される。

【0044】

ステップS1で本方法は、第1のフレームレートで、学習的認識手法により、動画像内の少なくとも1つの対象を認識する。学習的認識手法には、例えばCNNをベースとするディープラーニングが使われてよい。対象の認識は例えば、(1)認識候補の検出、(2)ターゲットの決定、の2段階で行われてよい。これによりターゲットROIの画像、位置、サイズなどが認識される。

20

【0045】

ステップS2で本方法は、第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、動画像内の対象を追跡する。動画像内の対象の追跡には、相関フィルタによる局所領域追跡が使われてよい。相関フィルタによる局所領域追跡は、(1)相関フィルタに用いるROI画像選択、(2)周波数空間での相関計算、(3)相関ピーク検出に基づく追跡ROI領域更新、などの処理を含んでよい。

30

【0046】

ステップS3で本方法は、ステップS1で学習的認識手法による対象の認識が開始されたフレームのROIと、ステップS2で第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームのROIとの間の差異を補正する。前述のようにこの補正は、例えば、学習的認識手法により対象の認識が開始されたフレームからカレントフレームまでのビデオ映像をバッファに記憶しておき、このバッファ内の実映像を再生して対象を追跡することにより行われてよい。この実映像の再生は、バッファに記憶したすべてのフレームを用いた高速再生であってもよく、バッファ内から抽出したフレームのサブセットのコマ送りの再生であってもよい。

40

【0047】

ステップS4で本方法は、ステップS3で補正されたROIに基づいて、カレントフレームのROI位置を初期化する。

【0048】

「終了」で、この周期についての学習的手法によるフロー内の処理が終了する。続けてフローは次の周期に遷移し、次の第1の周期についての学習的手法による対象の追跡が開始される。以下、第1の周期ごとに同様のフローが繰り返される。

【0049】

本実施例によれば、ROI位置がカレントフレーム内の対象位置から遅延することなく

50

、常にカレントフレームに追跡したものとなる。これにより、滑らかで途切れのない対象追跡を実現することができる。

【 0 0 5 0 】

(第2実施形態)

図3は、第2実施形態に係る対象追跡方法のフロー図である。

第2実施形態は、第1実施形態のステップS4に代えて、ステップS41を備える。その他の処理は、第1実施形態と共通である。

【 0 0 5 1 】

ステップS41で本方法は、ステップS3で補正されたROIに基づいて、カレントフレームのROI位置を、第1のフレームレートで初期化する。すなわち本実施形態では、第1のフレームレートで学習的認識手法により対象が認識されると、それに応じて、ROI位置は毎回初期化される。

10

【 0 0 5 2 】

本実施形態によれば、学習的認識手法で得た認識結果を、ROI位置の更新に最大限利用することができる。

【 0 0 5 3 】

(第3実施形態)

図4は、第3実施形態に係る対象追跡方法のフロー図である。

第3実施形態は、第1実施形態の各処理に加えて、ステップS5をさらに備える。その他の処理は、第1実施形態と共通である。

20

【 0 0 5 4 】

ステップS5で本方法は、ステップS1で学習的認識手法により認識された対象に基づいて、第1のフレームレートでテンプレート及びROI位置を初期化する。すなわち本実施形態では、第2のフレームレートで行われる対象追跡に用いられるテンプレート及びROI位置が、第1のフレームレートで学習的認識手法によって対象が認識される度に、毎回初期化される。

【 0 0 5 5 】

本実施形態によれば、学習的認識手法で得た認識結果を、テンプレート及びROI位置の更新に最大限利用することができる。

【 0 0 5 6 】

(第4実施形態)

図5は、第4実施形態に係る対象追跡方法のフロー図である。

第4実施形態は、第1実施形態の各処理に加えて、ステップS42およびステップS6をさらに備える。その他の処理は、第1実施形態と共通である。

30

【 0 0 5 7 】

ステップS6で本方法は、対象の認識が開始されたフレームにおいて、学習的認識手法により認識された対象と、テンプレートマッチングにより追跡された対象とを比較する。この比較は、対象のROIの位置ずれや、テンプレートの形状のパターンの相違などを含んでよい。ステップS6での比較結果が所定の値以下のときは、ステップS42で、テンプレートマッチングにより追跡された対象に基づいてROI位置を初期化する。一方、ステップS6での比較結果が所定の値より大きいときは、ステップS4で、補正されたROIに基づいてカレントフレームのROI位置を初期化する。すなわち本実施形態では、追跡が正常に機能していて、学習的認識手法により認識された対象と、テンプレートマッチングにより追跡された対象との間に所定の差異がない限り、ROI位置の初期化は、テンプレートマッチングにより追跡された対象に基づいて行われる。追跡が外れて対象が見失われた場合に限り、ROI位置の初期化は、学習的認識手法により認識された対象に基づいて行われる。

40

【 0 0 5 8 】

本実施形態によれば、追跡軌跡が連続するように、ROI位置を更新することができる。

50

【 0 0 5 9 】

(第 5 実施形態)

図 6 は、第 5 実施形態に係る対象追跡方法のフロー図である。

第 5 実施形態は、第 3 実施形態の各処理に加えてステップ S 5 2 およびステップ S 6 をさらに備え、ステップ S 5 に代えてステップ S 5 1 を備える。その他の処理は、第 3 実施形態と共通である。

【 0 0 6 0 】

ステップ S 6 で本方法は、学習的認識手法により認識された対象と、テンプレートマッチングにより追跡された対象とを比較する。この比較は、対象の R O I の位置ずれや、テンプレートの形状のパターンの相違などを含んでよい。ステップ S 6 での比較結果が所定の値以下のときは、ステップ S 5 2 で、テンプレートマッチングにより追跡された対象に基づいてテンプレート及び R O I 位置を初期化する。一方、ステップ S 6 での比較結果が所定の値より大きいときは、ステップ S 5 1 で、学習的認識手法により認識された対象に基づいてテンプレート及び R O I 位置を初期化する。すなわち本実施形態では、対象の認識が開始されたフレームにおいて、追跡が正常に機能していて、学習的認識手法により認識された対象と、テンプレートマッチングにより追跡された対象との間に所定の差異がない限り、テンプレート及び R O I 位置の初期化は、テンプレートマッチングにより追跡された対象に基づいて行われる。追跡 R O I が認識結果から逸脱した場合に限り、テンプレート及び R O I 位置の初期化は、学習的認識手法により認識された対象に基づいて行われる。

【 0 0 6 1 】

本実施形態によれば、必要な場合にのみ、学習的認識手法で得た認識結果を、テンプレートの更新に利用することができる。

【 0 0 6 2 】

(第 6 実施形態)

図 7 は、第 6 実施形態に係る対象追跡方法のフロー図である。

第 6 実施形態は、第 1 実施形態の各処理に加えてステップ S 7 をさらに備え、ステップ S 3 に代えてステップ S 3 1 を備える。その他の処理は、第 1 実施形態と共通である。

【 0 0 6 3 】

ステップ S 7 で本方法は、学習的認識手法による対象の認識が開始されたフレームからカレントフレームまでのフレームを記憶する。

【 0 0 6 4 】

ステップ S 3 1 で本方法は、学習的認識手法による対象の認識が開始されたフレームの R O I 位置とテンプレートとを用いて記憶したフレームに対して、テンプレートマッチングによる追跡を実行することによって、対象の認識が開始されたフレームの R O I と、カレントフレームの R O I との間の差異を補正する。このときのテンプレートマッチングは、例えば、第 2 のフレームレートにおける 1 フレーム分に相当する時間で実行される。

【 0 0 6 5 】

本実施形態によれば、学習的認識手法により対象の認識が開始されたフレームからカレントフレームまでのテンプレートマッチングによる追跡を実行することより、R O I 位置を補正することができる。

【 0 0 6 6 】

(第 7 実施形態)

図 8 は、第 7 実施形態に係る対象追跡方法のフロー図である。

第 7 実施形態は、第 1 実施形態の各処理に加えてステップ S 8 をさらに備え、ステップ S 3 に代えてステップ S 3 2 を備える。その他の処理は、第 1 実施形態と共通である。

【 0 0 6 7 】

ステップ S 8 で本方法は、S 1 で学習的認識手法により対象の認識が開始されたフレームから、カレントフレームの画像を予測する。この予測には、線形予測などの既知の手法が用いられてよい。

10

20

30

40

50

【 0 0 6 8 】

ステップ S 4 2 で本方法は、ステップ S 1 で学習的認識手法により対象の認識が開始されたフレームの R O I と、ステップ S 8 で予測されたカレントフレームの R O I との間の差異を補正する。

【 0 0 6 9 】

本実施形態によれば、実映像のカレントフレームを用いる代わりに、カレントフレームを予測することにより、テンプレートマッチングの処理量を大きく削減することができる。

【 0 0 7 0 】

(第 8 実施形態)

図 9 は、第 8 実施形態に係る対象追跡システムのブロック図である。

対象追跡システム 1 は、対象認識部 1 0 と、対象追跡部 1 1 と、補正部 1 2 と、初期化部 1 3 と、を備える。

【 0 0 7 1 】

対象認識部 1 0 は、第 1 のフレームレートで、学習的認識手法により、動画像内の対象を認識する。対象追跡部 1 1 は、第 1 のフレームレートより高い第 2 のフレームレートで、テンプレートを用いたテンプレートマッチングにより、動画像内の対象を追跡する。補正部 1 2 は、対象認識部 1 0 で学習的認識手法による対象の認識が開始されたフレームの R O I と、対象追跡部 1 1 で第 2 のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームの R O I との間の差異を補正する。初期化部 1 3 は、補正部 1 2 で補正された R O I に基づいて、カレントフレームのテンプレート及び R O I 位置を初期化する。

【 0 0 7 2 】

図 1 0 は、第 9 実施形態に係る対象追跡システム 2 のブロック図である。

本実施形態に係る対象追跡システムは、対象認識部 1 0 と、対象追跡部 1 1 と、補正部 1 2 と、初期化部 1 3 と、を備える対象追跡装置 3 と、その他の周辺装置を含む。対象追跡装置 3 の構成および動作は、第 8 実施例に係る対象追跡システム 1 の構成及び動作と共通であるため、説明を省略する。

【 0 0 7 3 】

対象追跡システム 2 は、高速カメラ 1 4 を備えてよい。高速カメラ 1 4 は、例えば毎秒 5 0 0 フレーム以上のフレームレートで動画像を撮影し、撮影した動画像データを対象追跡装置 3 に送信する。

【 0 0 7 4 】

対象追跡システム 2 は、操作用 P C 1 5 を備えてよい。操作用 P C 1 5 は操作者によって操作され、操作者が設定した設定データや各種のコマンドを対象追跡装置 3 に送信する。

【 0 0 7 5 】

対象追跡装置 3 は、表示装置 1 6、スマートフォン 1 7 またはサーバ 1 8 を備えてよい。対象追跡装置 3 で追跡された追跡対象データは、表示装置 1 6 やスマートフォン 1 7 に送信され、これらの装置で動画像として表示される。あるいは対象追跡装置 3 で追跡された追跡対象データは、サーバ 1 8 に送信され、動画像データとして蓄積される。

【 0 0 7 6 】

本実施例によれば、R O I 位置がカレントフレーム内の対象位置から遅延することなく、常にカレントフレームに追跡したものとなるので、滑らかで途切れのない対象追跡を実現することができる。

【 0 0 7 7 】

(第 9 実施形態)

追跡の対象は複数あってもよい。特にこれらの対象は、異なる速度で運動していてもよい。以下、図 2 を参照して、第 9 実施形態に係る対象追跡方法のフローを説明する。「開始」で、ある第 1 の周期についての学習的手法による対象の追跡が開始される。

10

20

30

40

50

【0078】

ステップS1で本方法は、第1のフレームレートで、学習的認識手法により、動画像内の複数の対象を認識する。複数の対象は、例えば単一のカメラで撮影された動画像内で運動する複数の対象物であってよい。特にこれらの対象物の運動速度は異なってもよい。以下、例示として、第1の対象として「人」、第2の対象として「自動車」を認識するものとする。なおこの場合の学習的認識手法も、単一の対象を追跡する場合と同様、例えばCNNをベースとするディープラーニングが使われてよい。

【0079】

ステップS2で本方法は、第1のフレームレートより高い第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、動画像内の人と自動車をそれぞれ追跡する。追跡には、相関フィルタによる局所領域追跡が使われてよい。相関フィルタによる局所領域追跡は、(1)相関フィルタに用いるROI画像選択、(2)周波数空間での相関計算、(3)相関ピーク検出に基づく追跡ROI領域更新、などの処理を含んでよい。

10

【0080】

ステップS3で本方法は、人および自動車に関し、ステップS1で学習的認識手法による認識が開始されたフレームの人および自動車のROIと、ステップS2で第2のフレームレートでテンプレートマッチングを継続した結果現在に至ったカレントフレームのROIとの間の差異を補正する。前述のようにこの補正は、例えば、学習的認識手法により人および自動車の認識が開始されたフレームからカレントフレームまでのビデオ映像をバッファに記憶しておき、このバッファ内の実映像を再生して追跡することにより行われてよい。この実映像の再生は、バッファに記憶したすべてのフレームを用いた高速再生であってもよく、バッファ内から抽出したフレームのサブセットのコマ送りの再生であってもよい。

20

【0081】

ステップS4で本方法は、ステップS3で補正されたROIに基づいて、カレントフレームの人および自動車のそれぞれのROI位置を初期化する。

【0082】

「終了」で、この周期についての学習的手法によるフロー内の処理が終了する。続けてフローは次の周期に遷移し、次の第1の周期についての学習的手法による人および自動車の追跡が開始される。以下、第1の周期ごとに同様のフローが繰り返される。

30

【0083】

従来の対象追跡技術では、運動速度の異なる複数の対象物を追跡しようとした場合、対象により追跡のしやすさが異なる。すなわち、一般に高速で運動する対象は、低速で運動する対象より追跡が困難である。これは、追跡中に互いに対象物を取り違えてしまったりすることや、画面上で2つの画像が交錯することにより発生するオクルージョンに起因して対象の追跡が失敗したりすることの原因となる。本実施形態によれば、複数の対象のそれぞれのROI位置がカレントフレームに遅延することなく、常にカレントフレームに追従したものとなるため、滑らかで途切れのない追跡が実現され、継続した追跡が可能となる。

【0084】

以上の実施形態は、任意の動画像に適用することができる。特に実施形態は、近年性能の向上が著しく、産業上広く用いられている高速カメラにより撮影された動画像に適用してよい。画像取り込みのサンプリングレートが高い高速カメラの場合、対象追跡技術における前述の課題は特に深刻となる。従って、本発明の実施形態は、高速カメラにより撮影された動画像に対して特に効果的である。

40

【0085】

第2のフレームレートの方が第1のフレームレートより高い限り、これらのフレームレートは任意の値であってよい。一例として、第1のフレームレートは毎秒100フレーム以下であってよく、第2のフレームレートは毎秒500フレーム以上であってよい。ディープラーニング等の学習的手法は、毎秒100フレームを超えると、計算機コストや計算

50

時間上のデメリットが大きい。一方、現在および将来の高速カメラの性能や仕様における要求条件に照らせば、テンプレートマッチングによる追跡は、毎秒500フレーム以上で行われることが望ましい。

【0086】

(検証実験)

本発明者らは、本発明の効果を確認するために、実施形態による対象追跡実験を行った。

【0087】

図11は、比較例に係る方法による対象追跡実験の結果を示す図である。

ここでは比較例として、非特許文献1に記載の技術を用いた。実験は、サンプリングレート500fpsで取得した動画像に対し、これと同じフレームレート500fpsで相関フィルタを用いた追跡を行うとともに、フレームレート50fpsでディープラーニングによる対象認識を行うことにより行った。追跡が正常に機能している限り、ROI位置とテンプレートは相関フィルタによって更新され、追跡が継続される。追跡が外れて対象が見失われた場合は、ディープラーニングで認識された対象を基にテンプレートROIが更新され、これらを基に追跡が再開される。

10

【0088】

図11(a)、図11(b)は、それぞれ追跡された対象の、画面上におけるx座標、y座標を示す。ここでx座標は画像の水平方向を、y座標は画像の垂直方向をそれぞれ意味する(以下同様)。図示されるように、グラフは複数の時点で途切れており、追跡が失敗している。また追跡が成功しているところであっても、グラフは滑らかな曲線を描いていない。これらは、ディープラーニングによる対象認識時点のフレームがカレントフレームに対して遅延していることに起因して、正しい追跡が行われていないことを表すと考えられる。

20

【0089】

図12は、本発明に係る第6実施形態による対象追跡実験の結果を示す図である。

本実施形態と比較例との違いは、本実施形態が、ディープラーニングによる対象認識時点のフレームとカレントフレームとの間のROIの差異を補正するステップを備える点である。その他の実験条件は、比較例と共通である。

【0090】

図12(a)、図12(b)は、それぞれ追跡された対象の、画面上におけるx座標、y座標を示す。図11と異なり、測定時間の5秒間全体にわたって、グラフには途切れがない。またグラフ全体は、概ね滑らかな曲線を描いている。これは正しい対象追跡が実現されていることを表すと考えられ、本発明の有効性が確認された。

30

【0091】

図13は、本発明の第9実施形態による対象追跡実験の結果を示す図である。図13は、対象すなわち「人」と「自動車の」画面上におけるx座標を示す。追跡された自動車は、測定時間の5秒間全体にわたって、グラフにはほぼ途切れがない。一方、追跡された人は、時刻3秒から3.7秒の間にカメラの前を自動車が通過したためオクルージョンが発生し、この間グラフが途切れている。しかしながら自動車が通過した後に即座に認識がされ、その後の継続した追跡に成功している。このように、運動速度の異なる複数の対象に対して、途中でオクルージョンが発生した場合であっても、本発明の有効性が確認された。ただし移動体は反転や方向転換時などでは停止期間もあるため、速度の概念は速度=0を含むものとする。

40

【0092】

以上、本発明を実施例を基に説明した。これらの実施例は例示であり、それらの各構成要素や各処理プロセスの組合せにいろいろな変形例が可能なこと、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。

【0093】

例えば第2のフレームレートは、動画像を取り込むときのフレームレートと一致してい

50

てよい。あるいは第2のフレームレートは、取り込んだ動画像のフレームの一部のフレームに関するものであってもよい。すなわち、テンプレートマッチングによる対象の追跡は、フレームバイフレームで行ってもよいし、フレームのサブセットで行ってもよい。

【0094】

高速カメラは、肉眼で認識できる画像を撮影するための可視光撮影カメラに代えて、赤外線等の非可視光を可視化するための非可視光撮影カメラであってもよい。

【0095】

本発明による手法は、監視カメラ等の固定カメラに限らず、車載用カメラやドローン用カメラなどの移動体用カメラ、パンなどの大きな動きを伴うカメラで撮影された画像など、広い分野での動画像に適用が可能であり、産業上の利用性が高いものである。

10

【産業上の利用可能性】

【0096】

本発明は、対象追跡方法、対象追跡システムおよび対象追跡プログラムに関する。

【符号の説明】

【0097】

S1・・・第1のフレームレートで、学習的認識手法により、動画像内の対象を認識するステップ

S2・・・第2のフレームレートで、テンプレートを用いたテンプレートマッチングにより、動画像内の対象を追跡するステップ

S3・・・対象の認識が開始されたフレームのROIと、カレントフレームのROIとの間の差異を補正するステップ

20

S31・・・学習的認識手法による対象の認識が開始されたフレームのROI位置とテンプレートとを用いて記憶したフレームに対して、テンプレートマッチングによる追跡を実行することによって、対象の認識が開始されたフレームのROIと、カレントフレームのROIとの間の差異を補正するステップ

S32・・・学習的認識手法による対象の認識が開始されたフレームのROIと、予測されたカレントフレームのROIとの間の差異を補正するステップ

S4・・・補正されたROIに基づいて、カレントフレームのROI位置を初期化するステップ

S41・・・補正されたROIに基づいて、第1のフレームレートで、カレントフレームのROI位置を初期化するステップ

30

S42・・・テンプレートマッチングにより追跡された対象に基づいて、カレントフレームのROI位置を初期化

S5・・・学習的認識手法により認識された対象に基づいて、第1のフレームレートで、テンプレート及びROI位置を初期化するステップ

S51・・・学習的認識手法により認識された対象に基づいて、カレントフレームのテンプレート及びROI位置を初期化するステップ

S52・・・テンプレートマッチングにより追跡された対象に基づいて、カレントフレームのテンプレート及びROI位置を初期化するステップ

S6・・・対象の認識が開始されたフレームにおいて、学習的認識手法により認識された対象と、テンプレートマッチングにより追跡された対象との差を比較するステップ

40

S7・・・学習的認識手法による対象の認識が開始されたフレームからカレントフレームまでのフレームを記憶するステップ

S8・・・学習的認識手法による対象の認識が開始されたフレームの画像から、カレントフレームの画像を予測するステップ

1・・・対象追跡システム

10・・・対象認識部

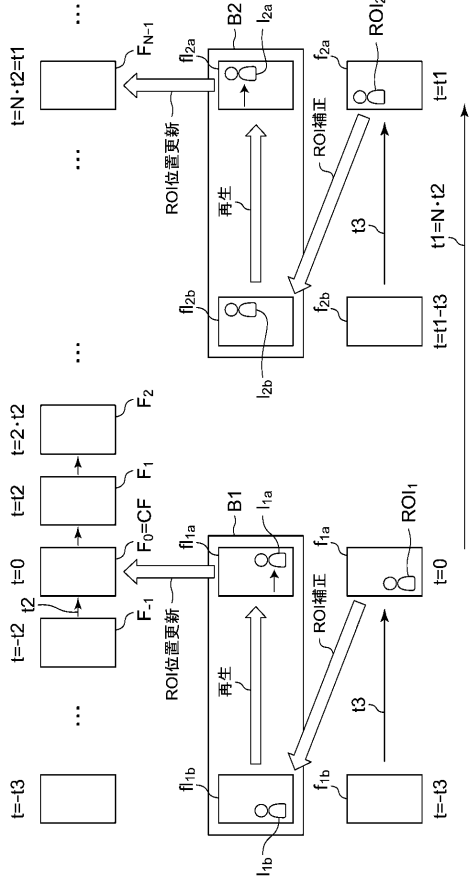
11・・・対象追跡部

12・・・補正部

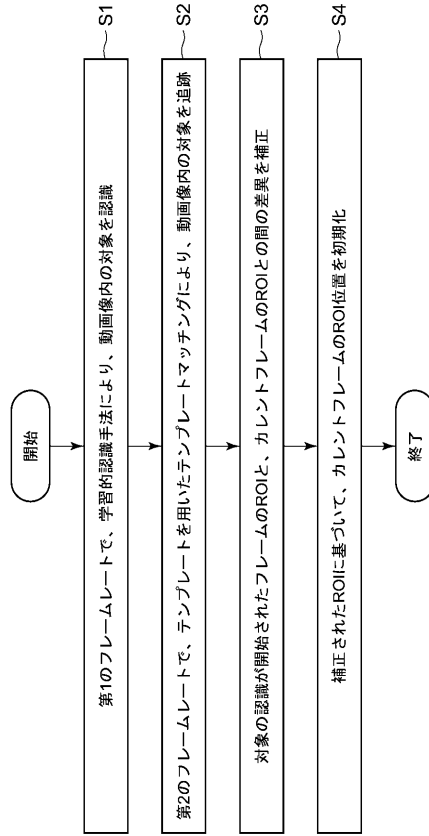
13・・・初期化部

50

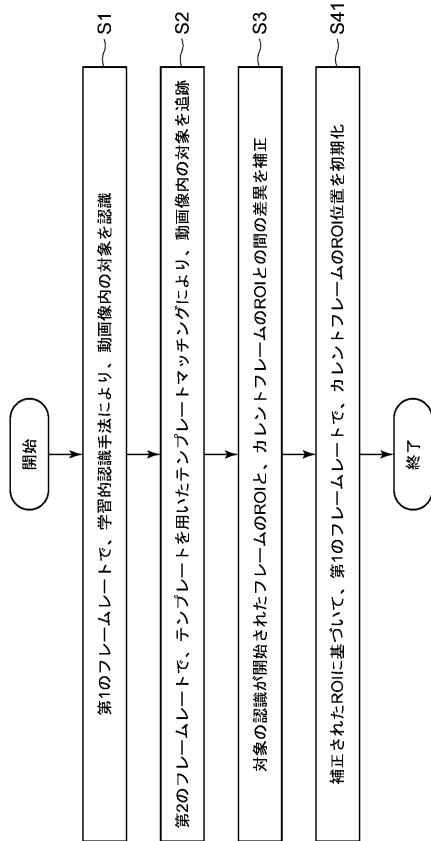
【図 1】



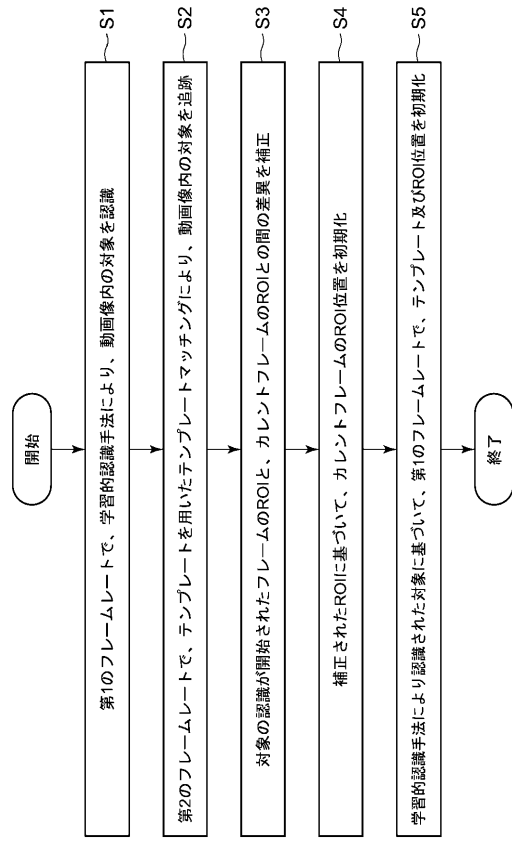
【図 2】



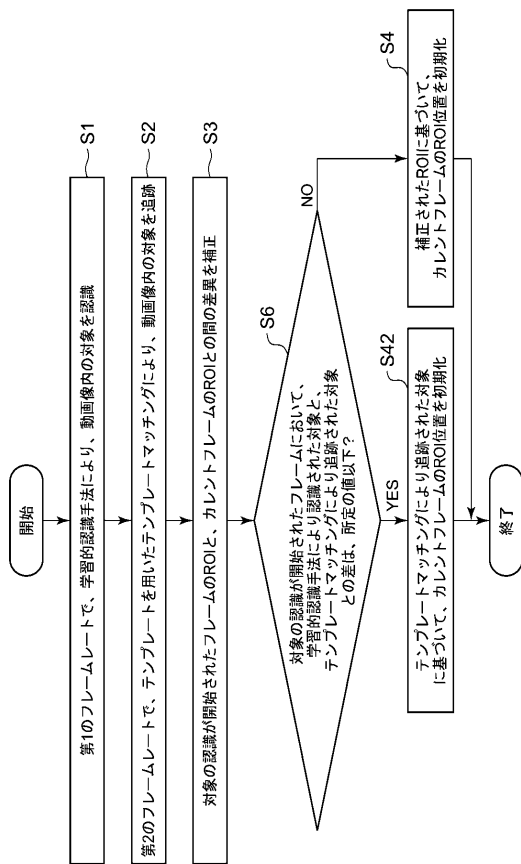
【図 3】



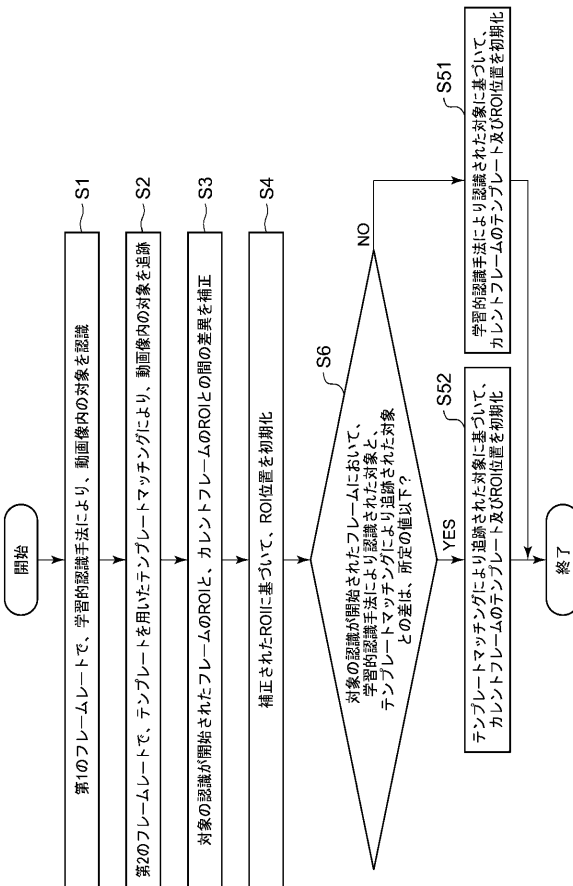
【図 4】



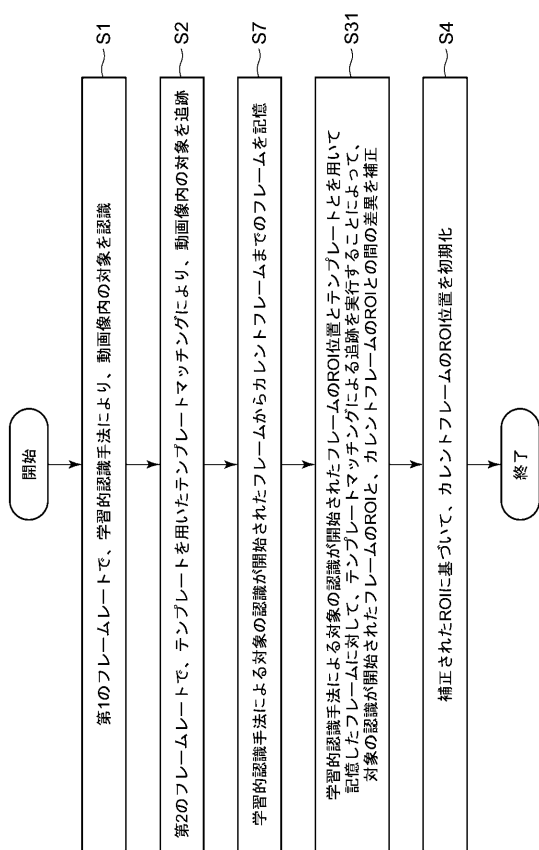
【 図 5 】



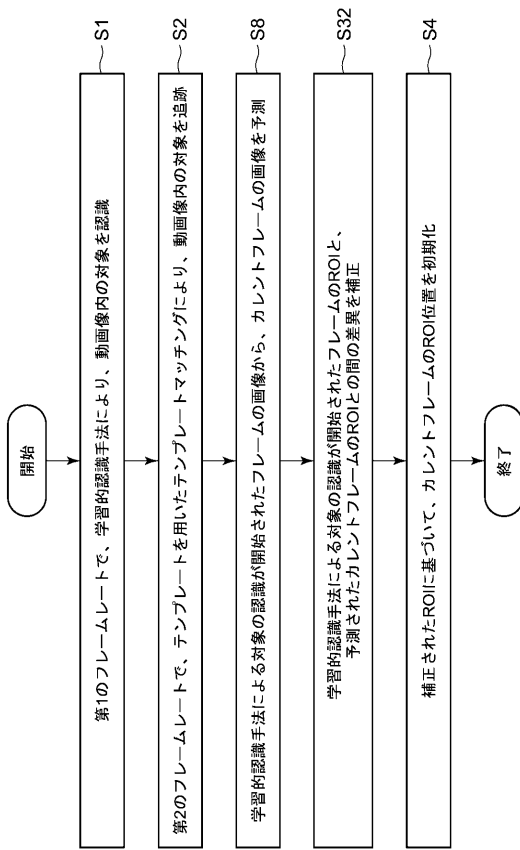
【 図 6 】



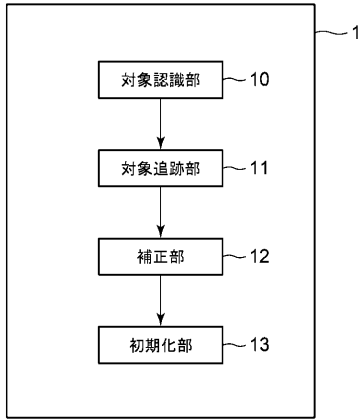
【 図 7 】



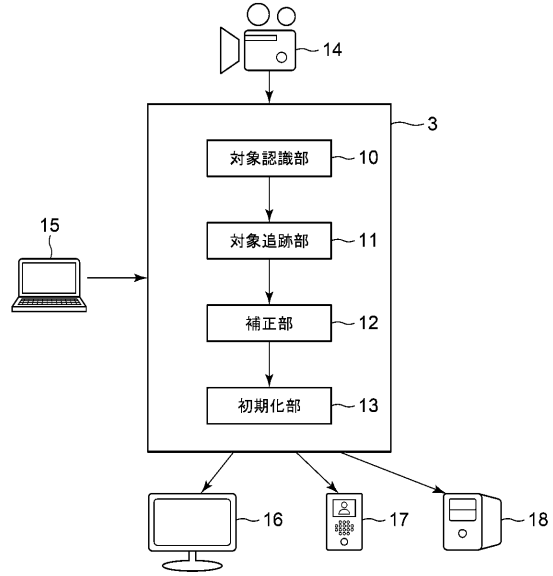
【 図 8 】



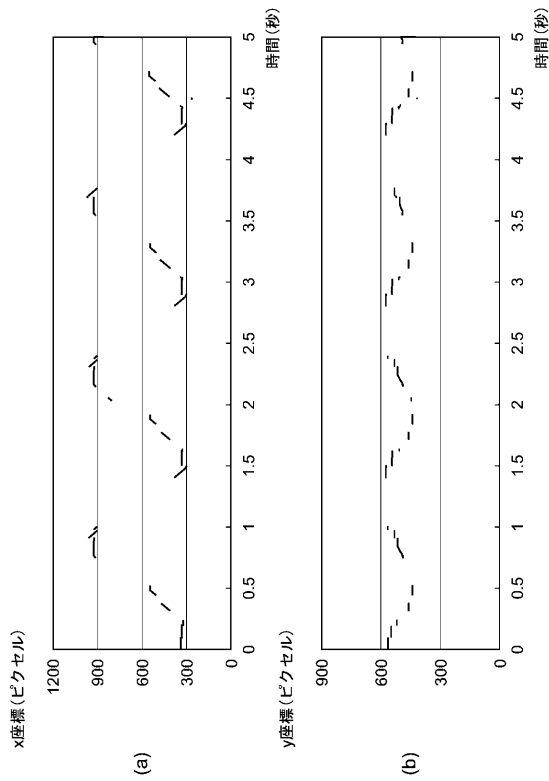
【図9】



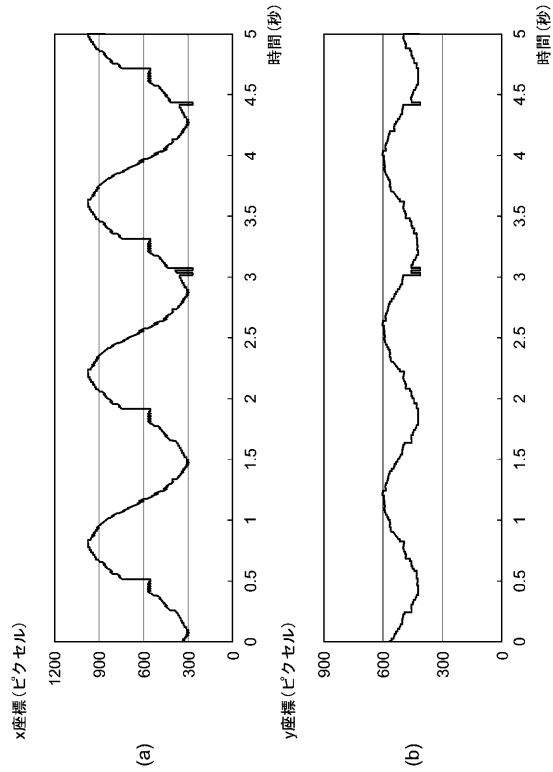
【図10】



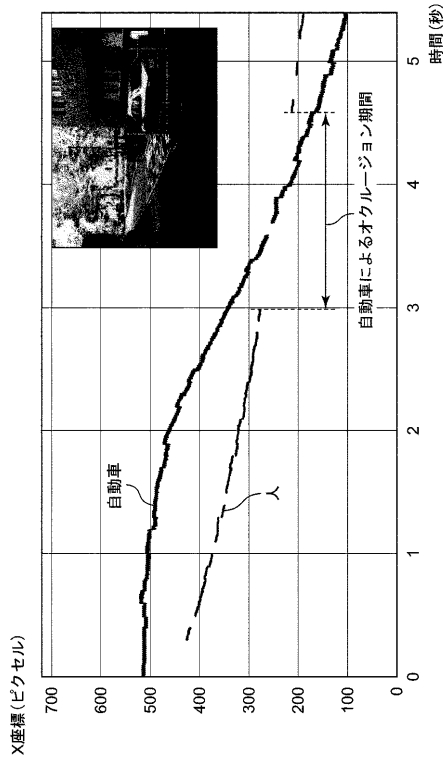
【図11】



【図12】



【 図 1 3 】



フロントページの続き

審査官 久保 光宏

- (56)参考文献 姜 明俊(外2名),「ディープラーニングを用いた高速対象追跡」,ロボティクス・メカトロニクス講演会2018講演論文集, Article number: 2A1-K17, 日本, [DVD-ROM], 一般社団法人日本機械学会, 2018年06月01日, 全4頁.
Mingjun Jiang, et al., "High-frame-rate Target Tracking with CNN-based Object Recognition", Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018年10月05日, Pages 599-606, ISBN: 978-1-5386-8094-0, <DOI: 10.1109/IROS.2018.8594300>.

(58)調査した分野(Int.Cl., DB名)

G 0 6 T 7 / 0 0 - 7 / 9 0
G 0 6 N 3 / 0 2 - 3 / 1 0
G 0 6 N 2 0 / 0 0 - 2 0 / 2 0
G 0 6 V 1 0 / 0 0 - 2 0 / 9 0
H 0 4 N 7 / 1 8
C S D B (日本国特許庁)
学術文献等データベース(日本国特許庁)
I E E E X p l o r e (I E E E)