

(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2004-355665

(P2004-355665A)

(43) 公開日 平成16年12月16日(2004. 12. 16)

(51) Int.Cl.⁷

G06T 7/20

F I

G06T 7/20

A

テーマコード (参考)

5 L O 9 6

審査請求 有 請求項の数 2 O L (全 51 頁)

(21) 出願番号	特願2004-273817 (P2004-273817)	(71) 出願人	000003078
(22) 出願日	平成16年9月21日 (2004. 9. 21)		株式会社東芝
(62) 分割の表示	特願平11-1891の分割		東京都港区芝浦一丁目1番1号
原出願日	平成11年1月7日 (1999. 1. 7)	(74) 代理人	100058479
(31) 優先権主張番号	特願平10-1847		弁理士 鈴江 武彦
(32) 優先日	平成10年1月7日 (1998. 1. 7)	(74) 代理人	100091351
(33) 優先権主張国	日本国 (JP)		弁理士 河野 哲
(31) 優先権主張番号	特願平10-192061	(74) 代理人	100088683
(32) 優先日	平成10年7月7日 (1998. 7. 7)		弁理士 中村 誠
(33) 優先権主張国	日本国 (JP)	(74) 代理人	100108855
			弁理士 蔵田 昌俊
		(74) 代理人	100075672
			弁理士 峰 隆司
		(74) 代理人	100109830
			弁理士 福原 淑弘

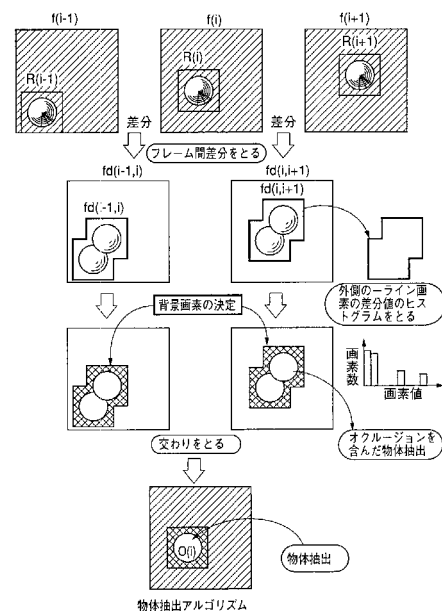
最終頁に続く

(54) 【発明の名称】 物体抽出装置

(57) 【要約】

【課題】 目的の物体以外の周囲の余分な動きに影響されずに、動画像から目的の物体を精度良く抽出／追跡する。

【解決手段】 時間的に連続な3つのフレーム $f(i-1)$ 、 $f(i)$ 、 $f(i+1)$ それぞれには、物体を囲む長方形 $R(i-1)$ 、 $R(i)$ 、 $R(i+1)$ が設定される。現フレーム $f(i)$ と第1参照フレーム $f(i-1)$ との間のフレーム間差分、および現フレーム $f(i)$ と第2参照フレーム $f(i+1)$ との間のフレーム間差分により、差分画像 $fd(i-1, i)$ 、 $fd(i, i+1)$ が得られる。多角形 $Rd(i-1, i) = R(i-1)$ or $R(i)$ と多角形 $Rd(i, i+1) = R(i)$ or $R(i+1)$ について、それぞれ背景領域が決定され、残りの領域が物体領域候補として選ばれる。これら物体領域候補の交わりを取ることで、現フレーム $f(i)$ の物体領域 $O(i)$ が抽出できる。



【選択図】 図8

【特許請求の範囲】

【請求項 1】

動画像データが記録されている記憶装置から前記動画像データを読み出し、前記動画像データを構成する複数フレーム内の所定フレーム上における物体領域を表すシェイプデータを動き補償することにより、前記読み出した動画像データを構成する各フレーム毎にシェイプデータを生成する手段と、

前記生成されたシェイプデータによって決定される各フレームの背景領域の画像データを背景メモリ上に逐次上書きすることによって、前記動画像データの背景画像を生成する手段と、

前記動画像データが記録されている記憶装置から前記動画像データを再度読み出し、前記読み出した動画像データを構成する各フレーム毎に前記背景メモリ上に蓄積されている背景画像の対応する画素との差分を求め、差分の絶対値が所定のしきい値よりも大きい画素を物体領域として決定する手段と、

により構成される物体抽出装置。

【請求項 2】

前記各フレームのシェイプデータによって決定される物体領域と、前記背景画像との差分の絶対値に基づいて決定される物体領域のいずれかを物体抽出結果として選択的に出力する手段をさらに有する請求項 1 に記載の物体抽出装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は画像の物体抽出装置に関し、特に入力動画像から目的とする物体の位置を検出して動物体の追跡／抽出を行う物体抽出装置に関する。

【背景技術】

【0002】

従来より、動画像中の物体を追跡／抽出するためのアルゴリズムが考えられている。これは、様々な物体と背景が混在する画像からある物体だけを抽出するための技術である。この技術は、動画像の加工や編集に有用であり、例えば、動画像から抽出した人物を別の背景に合成することなどができる。

【0003】

物体抽出に使用される方法としては、時空間画像領域分割（越後、飯作、「ビデオモザイクのための時空間画像領域分割」、1997年電子情報通信学会情報・システムソサイエティ大会、D-12-81、p. 273、1997年9月）を利用した領域分割技術が知られている。

【0004】

この時空間画像領域分割を用いた領域分割方法では、動画像の1フレーム内のカラーテクスチャによる小領域の分割を行ない、フレーム間の動きの関係を使ってその領域を併合する。フレーム内の画像を分割する際には、初期分割を与える必要があり、それによって分割結果が大きく左右されるという問題がある。そこで、これを逆に利用して、この時空間画像領域分割を用いた領域分割法では、別のフレームで初期分割を変えて、結果的に異なる分割結果を得て、フレーム間の動きで矛盾する分割を併合するという手法をとっている。

【0005】

しかし、この手法を動画像中の物体の追跡および抽出にそのまま適用すると、動きベクトルが、目的とする動物体以外の余分な動きに影響されてしまい、信頼度が十分でないことが多く、誤った併合を行なう点が問題となる。

【0006】

また、特開平8-241414号公報には、複数の動物体検出装置を併用した動物体検出・追跡装置が開示されている。この従来の動物体検出・追跡装置は、例えば監視カメラを用いた監視システムなどに用いられるものであり、入力動画像から動物体を検出してそ

の追跡を行う。この動物体検出・追跡装置においては、入力動画像は、画像分割部、フレーム間差分型動物体検出部、背景差分型動物体検出部、動物体追跡部にそれぞれ入力される。画像分割部では、入力動画像が予め定められた大きさのブロックに分割される。分割結果は、フレーム間差分型動物体検出部、および背景差分型動物体検出部にそれぞれ送られる。フレーム間差分型動物体検出部では、分割結果毎に、フレーム間差分を用いて入力動画像中の動物体が検出される。この場合、フレーム間差分を取る際のフレーム間隔は、動物体の移動速度に影響されずにその動物体を検出できるようにするために、背景差分型動物体検出部の検出結果に基づいて設定される。背景差分型動物体検出部では、これまでに入力された動画像を用いて分割結果毎に作成した背景画像と動物体との差分をとることにより、動物体が検出される。統合処理部では、フレーム間差分型動物体検出部および背景差分型動物体検出部それぞれの検出結果が統合されて、動物体の動き情報が抽出される。各フレームで物体を抽出した後、動物体追跡部では、フレーム間において対応する動物体同士に対応付けが行われる。

10

【0007】

この構成においては、フレーム間差分のみならず、背景差分をも用いて動物体の検出を行っているため、フレーム間差分だけを用いる場合に比べ検出精度は高くなる。しかし、入力動画像全体を対象としてその画像の中から動きのある物体をフレーム間差分および背景差分によって検出する仕組みであるため、フレーム間差分および背景差分それぞれの検出結果は、目的とする動物体以外の余分な動きに影響されてしまい、背景に複雑な動きがある画像ではうまく目的とする動物体を抽出・追跡できないという問題がある。

20

【0008】

また、別の物体抽出技術としては、複数のフレームを用いてまず背景画像を生成し、その背景画像と入力画像の画素値の差分が大きい領域を物体として抽出する方法も知られている。

【0009】

この背景画像を用いる物体抽出の既存の技術一例が、例えば、特開平8-55222号公報の「移動物体検出装置および背景抽出装置ならびに放置物体検出装置」に開示されている。

【0010】

現処理フレームの画像信号は、1フレーム分の画像を蓄えるフレームメモリ、第1の動き検出手段、第2の動き検出手段、スイッチに入力される。フレームメモリからは、1フレーム前の画像信号が読み出され、第1の動き検出手段に入力される。一方、背景画像を保持するために用意されたフレームメモリからは、その時点までに生成されている背景画像信号が読み出され、第2の動き検出手段と、スイッチに入力される。第1の動き検出手段と第2の動き検出手段では、各々入力される2つの画像信号の差分値などを用いて物体領域および物体領域が抽出され、いずれも論理演算回路に送られる。論理演算回路では、入力される2つの画像の論理積がとられ、それが最終的な物体領域として出力される。また、物体領域は、スイッチにも送られる。スイッチでは、物体領域によって、物体領域に属する画素については、背景画素信号が選択され、逆に、物体領域に属さない画素については、現処理フレームの画像信号が選択され、上書き信号としてフレームメモリに送られ、フレームメモリの画素値が上書きされる。

30

40

【0011】

この手法では、特開平8-55222号公報に示されている様に、処理が進行するにつれて、次第に背景画像が正しくなっていく、やがては、物体が正しく抽出されるようになる。しかし、動画像シーケンスの初めの部分においては背景画像に物体が混入しているために、物体の抽出精度が悪い。また、物体の動きが小さい場合には、いつまでたっても、その物体の画像が背景画像の中に残り、抽出精度は高くない。

【発明の開示】

【発明が解決しようとする課題】

【0012】

50

上述したように、従来の物体抽出／追跡方法では、入力動画像全体を対象としてその画像の中から動きのある物体を検出する仕組みであるため、目的とする動物体以外の余分な動きに影響を受けてしまい、目的の動物体を精度良く抽出・追跡することができないという問題があった。

【0013】

また、背景画像を用いる物体抽出法では、動画像シーケンスの初めの部分において抽出精度が悪く、また、物体の動きが小さい場合には、いつまでも背景画像が完成しないために抽出精度が良くならないという問題点があった。

【0014】

本発明は、目的の物体以外の周囲の余分な動きに影響を受けずにその物体の抽出／追跡を精度良く行うことが可能な動画像の物体抽出装置を提供することを目的とする。 10

【0015】

また、本発明は、背景画像を精度良く決定できるようにして、物体の動きの大小によらずに、且つ動画像シーケンスの初めの部分も最後の部分と同様に高い抽出精度が得られる物体抽出装置を提供することを目的とする。

【課題を解決するための手段】

【0016】

本発明は、物体抽出対象となる現フレームと、この現フレームに対し時間的に異なる第1の参照フレームとの差分に基づいて、前記現フレームと前記第1の参照フレームに共通の第1の背景領域を決定し、前記現フレームと、この現フレームに対し時間的に異なる第2の参照フレームとの差分に基づいて、前記現フレームと前記第2の参照フレームに共通の第2の背景領域を決定する背景領域決定手段と、前記現フレームの図形内画像の中で、前記第1の背景領域と前記第2の背景領域のどちらにも属さない領域を、物体領域として抽出する手段と、静止している物体領域を検出する物体静止検出手段とを具備する物体抽出装置を提供する。 20

【0017】

この物体抽出装置においては、物体抽出対象の現フレーム毎に二つの参照フレームが用意され、その現フレームと第1の参照フレームとの間の第1の差分画像により、現フレームと第1の参照フレームとで共通に用いられている第1の共通背景領域が決定され、また現フレームと第2の参照フレームとの間の第2の差分画像により、現フレームと第2の参照フレームとで共通に用いられている第2の共通背景領域が決定される。第1および第2のどちらの差分画像にも現フレーム上の物体領域が共通に含まれているため、第1の共通背景領域と第2の共通背景領域のどちらにも属さない領域の中で、現フレームの図形内画像に含まれる領域を検出することにより、現フレーム上の物体領域が抽出される。この物体領域が静止物体に相当する場合には、前の物体領域と現物体領域とに差分が存在しないとき静止物体領域が検出される。 30

【0018】

このようにして、時間的に異なる参照フレームに基づいて決定された複数の共通背景領域にいずれにも属さない領域を抽出対象物体と決定して物体の追跡を行うことにより、目的の物体以外の周囲の余分な動きに影響を受けずに、目的の物体を精度良く抽出・追跡することが可能となる。 40

【0019】

また、前記第1および第2の各参照フレームと前記現フレームとの間で背景の動きが相対的に零となるように、前記各参照フレームまたは現フレームの背景の動きを補正する背景補正手段をさらに具備することが好ましい。この背景補正手段を図形設定手段の入力段、あるいは背景領域決定手段の入力段のいずれかに設けることにより、例えばカメラをパンした時などのように背景映像が連続するフレーム間で徐々に変化するような場合であってもそれらフレーム間で背景映像を擬似的に一定にすることができる。よって、現フレームと第1または第2の参照フレームとの差分を取ることによって、それらフレーム間で背景を相殺することが可能となり、背景変化に影響されない共通背景領域検出処理および物 50

体領域抽出処理を行うことができる。背景補正手段は動き補償処理によって実現できる。

【0020】

また、前記背景領域決定手段は、前記現フレームと前記第1または第2の参照フレームとの差分画像の中で、前記現フレームの図形内画像または前記第1または第2の参照フレームの図形内画像に属する領域の輪郭線近傍における各画素の差分値を検出する手段と、前記輪郭線近傍の各画素の差分値を用いて、前記共通の背景領域と判定すべき差分値を決定する手段とを具備し、この決定された差分値を背景／物体領域判定のためのしきい値として使用して、前記差分画像から前記共通の背景領域を決定するように構成することが好ましい。このように輪郭線近傍における各画素の差分値に着目することにより、差分画像全体を調べることなく、容易にしきい値を決定することが可能となる。

10

【0021】

また、前記図形設定手段は、前記参照フレームの図形内画像を複数のブロックに分割する手段と、複数のブロックそれぞれについて、前記入力フレームとの誤差が最小となる前記入力フレーム上の領域を探索する手段と、探索された複数の領域を囲む図形を前記入力フレームに設定する手段とから構成することが好ましい。これにより、初期設定された図形の形状や大きさによらず、対象となる入力フレームに最適な新たな図形を設定することが可能となる。

【0022】

また、本発明は、既に物体領域を抽出したフレームから、物体抽出対象となる現フレーム上の物体の位置または形状を予測する予測手段と、この予測手段によって予測された現フレーム上の物体の位置または形状に基づいて、前記背景領域決定手段によって使用すべき前記第1および第2の参照フレームを選択する手段とをさらに具備する。

20

【0023】

このように、使用すべき参照フレームとして適切なフレームを選択することにより、常に良好な抽出結果を得ることが可能となる。

【0024】

ここで、 O_i 、 O_j 、 O_{curr} をそれぞれ参照フレーム f_i 、 f_j 及び抽出対象の現フレーム f_{curr} の物体とすると、正しく物体の形状を抽出するための最適な参照フレーム f_i 、 f_j とは、

$$(O_i \cap O_j) \subseteq O_{curr}$$

30

を満たすフレーム、つまり、 O_i 、 O_j の交わり部分が O_{curr} 内に属するようなフレーム f_i 、 f_j である。

【0025】

また、本発明は、互いに異なる方法によって物体抽出を行う複数の物体抽出手段を設け、これら物体抽出手段を選択的に切替ながら物体抽出を行うことを特徴とする。この場合、現フレームと、この現フレームとは時間的にずれた少なくとも2つの参照フレームそれぞれの差分を用いて物体抽出を行う第1の物体抽出手段と、フレーム間予測を使用して既に物体抽出が行われたフレームから現フレームの物体領域を予測することにより物体抽出を行う第2の物体抽出手段とを組み合わせ使用することが望ましい。これにより、物体が部分的に静止していて参照フレームとの差分が検出できないときでも、フレーム間予測を用いた物体抽出手段によってそれを補うことが可能となる。

40

【0026】

また、複数の物体抽出手段を設けた場合には、物体抽出対象となる現フレームから、その少なくとも一部の領域についての画像の特徴量を抽出する手段をさらに具備し、前記抽出された特徴量に基づいて、前記複数の物体抽出手段を切り替えることが好ましい。

【0027】

例えば、背景の動きがあるかどうか予めわかるならば、その性質を使った方がよい。背景の動きがある場合は、背景動き補償を行なうが、完全に補償できるとは限らない。複雑な動きをするフレームではほとんど補償できないこともある。このようなフレームは、背景動き補償の補償誤差量によって予め選別できるので、参照フレーム候補にしないなどの

50

工夫が可能である。しかし、背景の動きがない場合は、この処理は不必要である。別の物体が動いていると、誤った背景動き補償を行ったり、参照フレーム候補から外れたりして参照フレーム選択条件に最適なフレームであっても選ばれず、抽出精度が落ちることがあるからである。また、一つの画像中にも多様な性質が混在していることがある。物体の動きやテクスチャも部分的に異なり、同じ追跡・抽出方法及び装置やパラメータではうまく抽出できないことがある。従って、ユーザが画像中の特殊な性質を持つ一部を指定したり、画像中の違いを自動的に特徴量として検出し、その特徴量に基づいて、例えばフレーム内のブロック単位などで部分的に追跡・抽出方法を切替えて抽出したり、パラメータを変更した方がよい。

【0028】

10

このようにして、画像の特徴量に基づいて、複数の物体抽出手段を切替えれば、様々な画像中の物体の形状を精度良く抽出することが可能になる。

【0029】

また、現フレームと、この現フレームと時間的にずれた少なくとも2つの参照フレームとの差分を用いた第1の物体抽出手段と、フレーム間予測を使用した第2の物体抽出手段とを組み合わせる場合には、第2の物体抽出手段による予測誤差が所定の範囲内であるときは、第2の物体抽出手段による抽出結果が物体領域として使用され、予測誤差が所定の範囲を越えたときは第1の物体抽出手段による抽出結果が物体領域として使用されるように、予測誤差量に基づいて、フレーム内のブロック単位で第1および第2の物体抽出手段を選択的に切り替えて使用することが望ましい。

20

【0030】

また、第2の物体抽出手段は、参照フレームと物体抽出対象となる現フレームとの間のフレーム間隔が所定フレーム以上あくように入力フレーム順とは異なる順序でフレーム間予測を行うことを特徴とする。これにより、入力フレーム順でフレーム間予測を順次行う場合に比べフレーム間の動き量が大きくなるため、予測精度を向上でき、結果的に抽出精度を高めることが可能となる。

【0031】

すなわち、フレームの間隔によっては動きが小さ過ぎたり、複雑過ぎて、フレーム間予測による形状予測手法では対応できないことがある。従って、例えば形状予測の誤差が閾値以下にならない場合は、予測に用いる抽出済みフレームとの間隔をあけることにより、予測精度が上がり、結果的に抽出精度が向上する。また、背景に動きがある場合は、参照フレーム候補は抽出フレームとの背景動きを求め補償するが、背景の動きがフレームの間隔によっては小さ過ぎたり複雑過ぎたりして、背景動き補償が精度良くできない場合がある。この場合もフレーム間隔をあけることによって動き補償精度を上げることができる。このようにして抽出フレームの順序を適応的に制御すれば、より確実に物体の形状を抽出することが可能になる。

30

【0032】

また、本発明は、動画像データと、その動画像データを構成する複数フレーム内の所定フレーム上における物体領域を表すシェイプデータとを入力し、そのシェイプデータを用いて前記動画像データから物体領域を抽出する物体抽出装置において、前記動画像データが記録されている記憶装置から前記動画像データを読み出し、前記シェイプデータを動き補償することにより、前記読み出した動画像データを構成する各フレーム毎にシェイプデータを生成する手段と、前記生成されたシェイプデータによって決定される各フレームの背景領域の画像データを背景メモリ上に逐次上書きすることによって、前記動画像データの背景画像を生成する手段と、前記動画像データが記録されている記憶装置から前記動画像データを再度読み出し、前記読み出した動画像データを構成する各フレーム毎に前記背景メモリ上に蓄積されている背景画像の対応する画素との差分を求め、差分の絶対値が所定のしきい値よりも大きい画素を物体領域として決定する手段とを具備する。

40

【0033】

この物体抽出装置においては、記憶装置からの動画像データを読み出す1回目のスキャ

50

ン処理にて、背景画像が背景メモリ上に生成される。次いで、2回目のスキャン処理が行われ、1回目のスキャンで完成された背景画像を用いた物体領域の抽出が行われる。このようにして、動画像データが記憶装置に蓄積されていることを利用して動画像データを2回スキャンすることにより、動画像シーケンスの最初から十分に高い精度で物体領域を抽出することが可能となる。

【0034】

また、本発明は、前記各フレームのシェイプデータによって決定される物体領域と、前記背景画像との差分の絶対値に基づいて決定される物体領域のいずれかを物体抽出結果として選択的に出力する手段をさらに具備する。画像によっては、1スキャン目で得たシェイプデータによって決定される物体領域の方が、2スキャン目で、背景画像との差分を利用して得た物体領域よりも抽出精度が高い場合がある。したがって、1スキャン目で得た物体領域と2スキャン目で得た物体領域とを選択的に出力できるようにすることにより、さらに抽出精度の向上を図ることが可能となる。

10

【0035】

また、本発明は、動画像データと、この動画像データを構成する複数のフレーム内の所定フレーム上の物体領域を表すシェイプデータとを入力し、前記シェイプデータが与えられたフレームあるいは既にシェイプデータを求めたフレームを参照フレームとして使用することにより、前記各フレームのシェイプデータを逐次求めていく物体抽出装置であって、現処理フレームをブロックに分割する手段と、前記ブロック毎に、画像データの図柄が相似であり、且つ面積が現処理ブロックよりも大きい相似ブロックを前記参照フレームから探索する手段と、前記参照フレームから相似ブロックのシェイプデータを切り出して縮小したものを、前記現処理フレームの各ブロックに貼り付ける手段と、前記貼り付けられたシェイプデータを現処理フレームのシェイプデータとして出力する手段とを具備する。

20

【0036】

この物体抽出装置においては、物体抽出対象の現フレームの各ブロック毎に、画像データ（テクスチャ）の図柄が相似であり、且つ面積が現処理ブロックよりも大きい相似ブロックの探索処理と、探索された相似ブロックのシェイプデータを切り出して縮小したものを現処理フレームのブロックに貼り付ける処理とが行われる。このように現処理ブロックよりも大きい相似ブロックのシェイプデータを縮小して張り付けることにより、シェイプデータで与えられる物体領域の輪郭線がずれていてもそれを正しい位置に補正することが可能となる。したがって、例えばユーザがマウスなどで最初のフレーム上の物体領域の輪郭を大まかになぞったものをシェイプデータとして与えるだけで、以降の入力フレーム全てにおいて物体領域を高い精度で抽出することが可能となる。

30

【0037】

また、本発明は、画像データと、その画像の物体領域を表すシェイプデータを入力し、そのシェイプデータを用いて前記画像データから物体領域を抽出する物体抽出装置において、前記シェイプデータの輪郭部分にブロックを設定し、各ブロック毎に、前記画像データの図柄が相似であり、且つ前記ブロックよりも大きい相似ブロックを同じ画像の中から探索する手段と、前記各ブロックのシェイプデータを各々の前記相似ブロックのシェイプデータを縮小したもので置き換える手段と、前記置き換えを所定の回数だけ繰り返す手段と、前記置き換えを繰り返されたシェイプデータを補正されたシェイプデータとして出力する手段とを具備する。

40

【0038】

このようにフレーム内のブロックマッチングによって、相似ブロックを用いた置き換え処理を行うことにより、シェイプデータによって与えられる輪郭線を正しい位置に補正することが可能となる。また、フレーム内のブロックマッチングであるので相似ブロックの探索および置き換えを同一ブロックについて繰り返し行うことができ、これにより補正精度をさらに高めることが可能となる。

【発明の効果】

【0039】

50

本発明によれば、目的とする物体を囲む図形を用いてその物体を追跡することにより、目的の物体以外の周囲の余分な動きに影響を受けずに、目的の物体を精度良く抽出・追跡することが可能となる。

【0040】

また、入力画像によらずに高い抽出精度を得ることが可能となる。さらに、動画シーケンスの初めの部分も最後の部分と同様に高い抽出精度が得られる。また、物体の動きが小さかったり、全く動かない場合でも正しく抽出される。

【発明を実施するための最良の形態】

【0041】

図1には、本発明の第1実施形態に係る動画像の物体追跡／抽出装置の全体の構成が示されている。この物体追跡／抽出装置は、入力動画像信号から目的とする物体の動きを追跡するためのものであり、初期図形設定部1と、物体追跡・抽出部2とから構成されている。初期図形設定部1は、外部からの初期図形設定指示信号a0に基づいて、追跡／抽出対象となる目的の物体を囲むような図形を入力動画像信号a1に対して初期設定するために使用され、初期図形設定指示信号a0によって例えば長方形、円、楕円などの任意の形状の図形が目的の物体を囲むように入力動画像信号a1の初期フレーム上に設定される。初期図形設定指示信号a0の入力方法としては、例えば入力動画像信号a1を表示する画面上に利用者がペンやマウスなどのポインティングデバイスを用いて図形そのものを直接書き込んだり、あるいはそれらポインティングデバイスを用いて入力する図形の位置や大きさを指定するなどの手法を用いることができる。これにより、目的の物体が現れる初期フレーム画像上において、追跡／抽出対象となる物体を外部から容易に指示することが可能となる。

【0042】

また、ユーザによる図形入力ではなく、図形の初期設定は、通常のフレーム画像を解析する処理によって例えば人や動物の顔、体の輪郭などを検出し、それを囲むように図形を自動設定することによっても実現できる。

【0043】

物体追跡・抽出部2は、初期図形設定部1で設定された図形内に含まれる図形内画像を基準として物体の追跡および抽出を行う。この場合、動物体の追跡・抽出処理では、図形で指定された物体に着目してその物体の動きが追跡される。従って、目的とする動物体以外の周囲の余分な動きに影響を受けずに目的とする動物体の抽出／追跡を行える。

【0044】

図2には、物体追跡・抽出部2の好ましい構成の一例が示されている。

【0045】

この物体追跡・抽出部は、図示のように、メモリ(M)11、14、図形設定部11、背景領域決定部12、および物体抽出部13から構成されている。

【0046】

図形設定部11は、これまでに入力および図形設定した任意のフレームを参照フレームとして使用しながら入力フレームに対して順次図形を設定するために使用される。この図形設定部11は、現フレーム画像101、参照フレームの図形内画像およびその位置103と、参照フレームの物体抽出結果106を入力し、現フレームの任意の図形で囲まれた領域内を表す画像102を出力する。すなわち、図形設定部11による図形設定処理では、参照フレームの図形内画像103と現フレーム画像101との相関に基づいて、参照フレームの図形内画像103との誤差が最小となる現フレーム画像上の領域が探索され、その領域を囲むような図形が現フレーム画像101に対して設定される。設定する図形は、長方形、円、楕円、エッジで囲まれた領域、など、何でも良い。以下では、簡単のために長方形の場合について述べる。また、図形設定部11の具体的な構成については、図5を参照して後述する。なお、物体を囲む図形を用いない場合は、図形内画像は全画像とし、位置を入出力する必要がない。

【0047】

メモリ 10 には、これまでに入力および図形設定されたフレームが少なくとも 3 つ程度保持される。保持される情報は、図形設定されたフレームの画像、設定された図形の位置や形状、図形内の画像などである。また、入力フレームの画像全体ではなく、その図形内画像だけを保持するようにしても良い。

【0048】

背景領域決定部 12 は、物体抽出対象の現フレーム毎にその現フレームとは時間的に異なるフレームの中の少なくとも二つの任意のフレームを参照フレームとして使用し、各参照フレーム毎に現フレームとの差分を取ることによってそれら各参照フレームと現フレーム間の共通の背景領域を決定する。この背景領域決定部 12 は、メモリ 10 で保持された、現フレームの任意の図形内画像およびその位置 102 と、少なくとも 2 つのフレームの任意の図形内画像およびその位置 103 と、該少なくとも 2 つのフレームの物体抽出結果 106 を入力し、現フレームと該少なくとも 2 つのフレームそれぞれの図形内画像との共通の背景領域 104 を出力する。すなわち、参照フレームとして第 1 および第 2 の 2 つのフレームを使用する場合には、現フレームと第 1 の参照フレームとの間のフレーム間差分を取るなどによって得られた第 1 の差分画像により、現フレームと第 1 の参照フレームのどちらにおいても背景領域として用いられている共通の第 1 の背景領域が決定されると共に、現フレームと第 2 の参照フレームとの間のフレーム間差分を取るなどによって得られた第 2 の差分画像により、現フレームと第 2 の参照フレームのどちらにおいても背景領域として用いられている共通の第 2 の背景領域が決定されることになる。この背景領域決定部 12 の具体的な構成は、図 4 を参照して後述する。また、背景メモリを使って、共通の背景を得る手法もある。

【0049】

なお、物体を囲む図形を用いない場合は、図形内画像は全画像とし、位置を入出力する必要がない。

【0050】

物体抽出部 13 は、背景領域決定部 12 にて決定された共通の背景領域を用いて現フレームの図形内画像から物体領域のみを抽出するために使用され、該現フレームと少なくとも 2 つのフレームそれぞれとの共通の背景領域 104 を入力し、現フレームの物体抽出結果 106 を出力する。第 1 および第 2 の差分画像のどちらにも現フレーム上の物体領域が共通に含まれているため、第 1 の共通背景領域と第 2 の共通背景領域のどちらにも属さない領域の中で、現フレームの図形内画像に含まれる領域を検出することにより、現フレーム上の物体領域が抽出される。これは、共通の背景領域以外の領域が物体領域の候補となることを利用している。つまり、第 1 の差分画像上においては第 1 の共通背景領域以外の領域が物体領域候補となり、第 2 の差分画像上においては第 2 の共通背景領域以外の領域が物体領域候補となるので、2 つの物体領域候補の重複する領域を、現フレームの物体領域と判定することができる。物体抽出結果 106 としては、物体領域の位置や形状を示す情報を使用することができる。また、その情報を用いて実際に現フレームから物体領域の画像を取り出すようにしても良い。

【0051】

メモリ 14 は、少なくとも 2 つの物体抽出結果を保持し、既に抽出されている結果をフ

【0052】

ここで、図 8 を参照して、本実施形態で用いられる物体抽出・追跡処理の方法について説明する。

【0053】

ここでは、時間的に連続する 3 つのフレーム $f(i-1)$ 、 $f(i)$ 、 $f(i+1)$ を用いて、現フレーム $f(i)$ から物体を抽出する場合を例示して説明する。

【0054】

まず、前述の図形設定部 11 によって図形設定処理が行われる。3 つのフレーム $f(i-1)$ 、 $f(i)$ 、 $f(i+1)$ についてもそれぞれ任意の参照フレームを使用すること

により図形設定処理が行われ、そのフレーム上の物体を囲むように長方形 $R(i-1)$, $R(i)$, $R(i+1)$ が設定される。なお、長方形の図形 $R(i-1)$, $R(i)$, $R(i+1)$ は位置および形状の情報であり、画像として存在するものではない。

【0055】

次に、背景領域決定部 12 にて共通背景領域が決定される。

【0056】

この場合、まず、現フレーム $f(i)$ と第 1 の参照フレーム $f(i-1)$ との間のフレーム間差分が取られ、第 1 の差分画像 $fd(i-1, i)$ が求められる。同様に、現フレーム $f(i)$ と第 2 の参照フレーム $f(i+1)$ との間のフレーム間差分も取られ、第 2 の差分画像 $fd(i, i+1)$ が求められる。

10

【0057】

第 1 の差分画像 $fd(i-1, i)$ を得ることにより、現フレーム $f(i)$ と第 1 の参照フレーム $f(i-1)$ とで共通の画素値を持つ部分については画素値が相殺されるためその画素の差分値は零となる。したがって、フレーム $f(i-1)$ と $f(i)$ の背景がほぼ同様であれば、基本的には、第 1 の差分画像 $fd(i-1, i)$ には、長方形 $R(i-1)$ の図形内画像と、長方形 $R(i)$ の図形内画像との OR に相当する画像が残ることになる。この残存画像を囲む図形は、図示のように、多角形 $Rd(i-1, i) = R(i-1) \cup R(i)$ となる。現フレーム $f(i)$ と第 1 の参照フレーム $f(i-1)$ の共通の背景領域は、多角形 $Rd(i-1, i)$ 内の実際の物体領域（ここでは、2 つの丸を一部重ねた結果得られる 8 の字の形状をした領域）以外の全領域となる。

20

【0058】

また、第 2 の差分画像 $fd(i, i+1)$ についても、長方形 $R(i)$ の図形内画像と、長方形 $R(i+1)$ の図形内画像との OR に相当する画像が残ることになる。この残存画像を囲む図形は、図示のように、多角形 $Rd(i, i+1) = R(i) \cup R(i+1)$ となる。現フレーム $f(i)$ と第 2 の参照フレーム $f(i+1)$ の共通の背景領域は、多角形 $Rd(i, i+1)$ 内の実際の物体領域（ここでは、2 つの丸を一部重ねた結果得られる 8 の字の形状をしたもの）以外の全領域となる。

【0059】

この後、第 1 の差分画像 $fd(i-1, i)$ から現フレーム $f(i)$ と第 1 の参照フレーム $f(i-1)$ の共通の背景領域を決定する処理が行われる。

30

【0060】

共通背景領域・物体領域の判定のためのしきい値となる差分値が必要になる。これは、ユーザーが与えてもよいし、画像のノイズや性質を検出して自動設定してもよい。その場合、一画面で一つのしきい値でなくとも、画像中の部分的な性質に応じて部分的に決定してもよい。画像の性質は、エッジの強さや差分画素の分散などが考えられる。また、物体を追跡する図形を用いて求めることもできる。

【0061】

この場合、共通背景領域／物体領域の判定のためのしきい値となる差分値が求められ、しきい値以下の差分値を持つ画素の領域が共通背景領域として決定される。このしきい値は、第 1 の差分画像 $fd(i-1, i)$ の多角形 $Rd(i-1, i)$ の外側のライン、つまり多角形 $Rd(i-1, i)$ の輪郭線上に沿った各画素の差分値のヒストグラムを用いて決定することができる。ヒストグラムの横軸は画素値（差分値）、縦軸はその差分値を持つ画素数である。たとえば、画素数が多角形 $Rd(i-1, i)$ からなる枠線に存在する全画素数の半分となるような差分値が、前述のしきい値として決定される。このようにしきい値を決定することにより、第 1 の差分画像 $fd(i-1, i)$ 全体にわたって画素値の分布を調べることなく、容易にしきい値を決定することが可能となる。

40

【0062】

次に、このしきい値を用いて、第 1 の差分画像 $fd(i-1, i)$ の多角形 $Rd(i-1, i)$ 内における共通背景領域が決定される。共通背景領域以外の領域はオクルージョンを含んだ物体領域となる。これにより、多角形 $Rd(i-1, i)$ 内の領域は背景領域

50

と物体領域に2分され、背景領域の画素値が“0”、物体領域の画素値が“1”の2値画像に変換される。

【0063】

第2の差分画像 $f_d(i, i+1)$ についても、同様にして、現フレーム $f(i)$ と第2の参照フレーム $f(i+1)$ の共通の背景領域を決定する処理が行われ、多角形 $R_d(i, i+1)$ 内の領域が画素値“0”の背景領域と、画素値“1”の物体領域に変換される。

【0064】

この後、物体抽出部13による物体抽出力が行われる。

【0065】

ここでは、第1および第2の差分画像との間で、多角形 $R_d(i-1, i)$ 内の2値画像と多角形 $R_d(i, i+1)$ 内の2値画像とのAND処理を画素毎に行う演算処理が行われ、これによってオクルージョン入りの物体の交わりが求められ、現フレーム $f(i)$ 上の物体 $O(i)$ が抽出される。

【0066】

なお、ここでは、フレーム差分画像内の物体領域以外の他の全ての領域を共通背景領域として求める場合について説明したが、各フレームから図形内画像だけを取り出し、フレーム上での各図形内画像の位置を考慮してそれら図形内画像同士の差分演算を行うようにしてもよく、この場合には、図形外の背景領域を意識することなく、多角形 $R_d(i-1, i)$ 内、および多角形 $R_d(i, i+1)$ 内の共通背景領域だけが決定されることになる。

【0067】

このように、本実施形態では、

1) 現フレームとこの現フレームに対し時間的に異なる第1および第2の少なくとも2つの参照フレームそれぞれとの差分画像を求めることにより、現フレームと第1参照フレーム間の図形内画像のORと、現フレームと第2参照フレーム間の図形内画像のORとを求め、

2) それら図形内画像のOR処理により得られた差分画像をAND処理し、これによって、現フレームの図形内画像から目的の物体領域を抽出するという、図形内画像に着目したORAND法による物体抽出が行われる。

【0068】

また、現フレームと2つの参照フレームとの時間関係は前述の例に限らず、例えば、現フレーム $f(i)$ に対して時間的に先行する2つのフレーム $f(i-m)$ 、 $f(i-n)$ を参照フレームとして使用したり、時間的に連続する2つのフレーム $f(i+m)$ 、 $f(i+n)$ を参照フレームとして使用することも可能である。

【0069】

例えば、図8において、フレーム $f(i-1)$ 、 $f(i)$ を参照フレームとして使用し、これら参照フレームそれぞれとフレーム $f(i+1)$ との差分を取って、それら差分画像に対して同様の処理を行えば、フレーム $f(i+1)$ から物体を抽出することができる。

【0070】

図3には、物体追跡・抽出部2の第2の構成例が示されている。

【0071】

図2の構成との主な違いは、背景動き削除部21が設けられている点である。この背景動き削除部21は、各参照フレームと現フレームとの間で背景の動きが相対的に零となるように背景の動きを補正するために使用される。

【0072】

以下、図3の装置について、具体的に説明する。

【0073】

背景動き削除部21は、現フレーム201と時間的にずれた少なくとも2つのフレーム

10

20

30

40

50

の任意の図形内画像およびその位置 206 を入力し、時間的にずれた少なくとも 2 つのフレームの背景の動きを削除した画像 202 を出力する。この背景動き削除部 21 の具体的な構成例については、図 6 で後述する。

【0074】

図形設定部 22 は、図 2 の図形設定部 11 に対応し、現フレーム 201 と、該背景の動きを削除した少なくとも 2 つの画像 202 と、画像 202 の物体抽出結果 206 を入力し、現フレームおよび該少なくとも 2 つの画像 202 の、任意の図形に囲まれた領域内を表す画像 203 を出力する。

【0075】

メモリ 26 は、任意の図形内画像とその位置を保持する。

10

【0076】

背景領域決定部 23 は、図 2 の背景領域決定部 12 に対応し、該任意の図形内画像およびその位置 203 と、画像 202 の物体抽出結果 206 を入力し、現フレームと該少なくとも 2 つの画像 202 との共通の背景領域 204 を出力する。物体抽出部 24 は、図 2 の物体抽出部 13 に対応し、該現フレームと少なくとも 2 つの画像との共通の背景領域 204 を入力し、現フレームの物体抽出結果 205 を出力する。メモリ 25 は、少なくとも 2 つの物体抽出結果を保持する。これは、図 2 のメモリ 14 に相当する。

【0077】

このように背景動き削除部 21 を設けることにより、例えばカメラをパンした時などのように背景映像が連続するフレーム間で徐々に変化する場合であってもそれらフレーム間で背景映像を擬似的に一定にすることができる。よって、現フレームと参照フレームとの差分を取った時に、それらフレーム間で背景を相殺することが可能となり、背景変化に影響されない共通背景領域検出処理および物体領域抽出処理を行うことができる。

20

【0078】

なお、背景動き削除部 21 を背景領域決定部 23 の入力段に設け、これにより参照フレームの背景の動きを現フレームに合わせて削除するようにしても良い。

【0079】

図 4 (a) には、背景領域決定部 12 (または 23) の具体的な構成例一例が示されている。

【0080】

変化量検出部 31 は、現フレームと前述の第 1 および第 2 の参照フレームとの差分を取るために使用され、現フレームと、時間的にずれたフレームの任意の図形内画像およびその位置 302 と、時間的にずれたフレームの物体抽出結果 301 を入力し、現フレームと時間的にずれたフレームの任意の図形内画像間の変化量 303 を出力する。変化量は、例えば、フレーム間の輝度差分や色の変化、オプティカルフロー、などを用いることができる。時間的にずれたフレームの物体抽出結果を使えば、フレーム間で物体が変化しない場合でも物体は抽出できる。例えば、変化量をフレーム間差分とすると、物体に属するフレーム間差分ゼロの部分は、物体が静止しているということなので、時間的にずれたフレームの物体抽出結果と同じになる。

30

【0081】

代表領域決定部 32 では、現フレームの任意の図形内画像およびその位置 302 を入力し、任意の図形内画像の背景を代表領域 304 として出力する。代表領域は、任意の図形内で最も背景が多いと予想される領域を選ぶ。例えば、図 8 で説明した差分画像上の図形の輪郭線などのように、図形内の最も外側に帯状の領域を設定する。図形は物体を囲むように設定されるので、背景となる可能性が高い。

40

【0082】

背景変化量決定部 33 では、代表領域 304 と、該変化量 303 を入力し、背景を判定する変化量 305 を出力する。背景変化量の決定は、図 8 で説明したように代表領域の差分値の変化量のヒストグラムをとり、例えば、全体の画素数の半分 (過半数) 以上の画素数に相当する変化量、つまり差分値をもつ領域を背景領域と決定する。

50

【0083】

代表領域の背景決定部34では、背景の変化量305を入力し、代表領域の背景306を判定し、出力する。代表領域の背景領域の決定は、先に決定した背景変化量かどうかで判定する。背景領域決定部35では、変化量303と、背景判定のしきい値305と、代表領域の背景306を入力し、代表領域以外の領域の背景307を出力する。代表領域以外の背景領域の決定は、代表領域から成長法で行う。例えば、決定済みの画素と図形の内部方向に隣接する未決定画素が、背景変化量と一致すれば、背景と決定する。背景と隣接しない画素や、背景変化量と一致しない画素は、背景以外と判定される。また、単純に先に決定した背景変化量かどうかで判定してもよい。このようにして、差分画像上の図形の輪郭線から内周に向かって絞り込みを行うことにより、図形内画像の中でどこまでが背景領域であるかを決定することができる。

【0084】

また、逆に、図形の輪郭線から外側に向かって図形からはみ出した物体領域を検出する。例えば、背景以外と決定された画素と図形の外部方向に隣接する未決定画素が、背景変化量と一致しなければ、背景以外と決定する。背景以外の画素と隣接しない画素や背景変化量と一致する画素は背景と判定される。このようにして差分画像上の図形の輪郭線から外側に向かって広げることにより、図形外の画像のどこまでが背景外領域であるかを決定することができる。この場合は、図形外でも変化量を求める必要があるので、任意の図形を数画素太くして充分物体がはみ出さない図形を新たに設定して、その内部で変化量を求めるか、単純にフレーム全体で変化量を求めてもよい。また、図形内部だけ変化量を求めておき、図形外部の判定の時に随時変化量を求めながら上記処理を行ってもよい。当然、物体が図形をはみ出さない場合は、例えば、輪郭線上に背景以外の画素がない場合は、図形外部の処理を行う必要がない。

【0085】

ところで、現フレームと参照フレームとの間で物体または物体の一部が静止している場合、現フレームと参照フレームとの差分が検出されず、物体の形状が正しく抽出されない場合がある。そこで、既に抽出されている参照フレームを使って現フレームの物体を検出する方法を図4(b)を参照して説明する。

【0086】

図4(b)は、静止物体領域検出部37を備えた背景領域決定部12(または23)を示している。これによると、変化量検出部31は現フレームの図形内画像及びその位置と、時間的にずれた少なくとも2つのフレームの任意の図形内画像及びその位置311を入力とし、現フレームと参照フレームの図形内画像の変化量313を検出する。

【0087】

形状予測部36は、現フレームの図形内画像及びその位置と、時間的にずれた少なくとも2つのフレームの任意の図形内画像及びその位置311と、既に抽出されたフレームの画像及び物体形状317とを入力とし、現フレームと時間的にずれたフレームのうち未だ物体が抽出されていないフレームについては物体の形状312を予測し、出力する。

【0088】

静止物体領域決定部37は予測した物体形状312と、参照フレームと現フレームとの変化量313と、既に抽出されたフレームの物体形状317を入力とし、少なくとも2つのフレームから現フレームに対して静止している物体領域314を決定する。

【0089】

背景領域決定部35は少なくとも2つのフレームに関する、現フレームに対して静止している物体領域314と、参照フレームと現フレームとの変化量313とを入力とし、少なくとも2つのフレームと現フレームとのそれぞれの共通の背景領域316を決定し、出力する。

【0090】

まず、参照フレームの物体が抽出されている場合は、現フレームで参照フレームとのフレーム間差分がゼロの領域について、参照フレームでは同じ位置の領域が物体の一部であ

れば、現フレームでのその領域は静止した物体の一部として抽出できる。逆に、その領域が参照フレームでは、背景の一部であれば、現フレームでのその領域は背景となる。

【0091】

しかし、参照フレームの物体が未だ抽出されていない場合、静止した物体または物体の一部は上記方法では抽出できない。その場合、既に物体が抽出されている他のフレームを用いて、未だ抽出されていない参照フレームの物体形状を予測して、物体の一部であるかどうかを判定することができる。予測の方法は、画像の符号化でよく用いられるブロックマッチング法やアフエイン変換法などを用いる。

【0092】

一例として、図13で示すようなブロックマッチング法が考えられる。このようにして物体の形状が予測されれば、フレーム間差分が検出されない領域について静止した物体の一部か、背景かを判定することが可能となる。 10

【0093】

物体を囲む図形を用いない場合は、図形内画像は全画像とし、位置を入出力する必要がない。この形状予測は、参照フレームを選択する場合の形状予測と同じものを使うことができる。また、別の物体抽出法と切り換える実施例では、別の物体抽出法で得た物体形状を用いることができる。

【0094】

図5には、図形設定部11（または22）の具体的な構成の一例が示されている。

【0095】

分割部41では、現フレームと時間的にずれたフレームの任意の図形内画像とその位置402を入力し、分割された該画像403を出力する。任意の図形内画像の分割は、2等分、4等分、でもよいし、エッジを検出し、エッジにそった分割をおこなってもよい。以下、簡単に、2等分とし、分割された図形は、ブロックと呼ぶことにする。動き検出部42では、該分割された任意の図形内画像とその位置403と、現フレームの任意の図形内画像とその位置401を入力し、該分割された画像の動きと誤差404を出力する。ここでは、ブロックが現フレームに対応する位置を、誤差が最小になるよう探索し、動きと誤差を求める。分割判定部43では、該動きと誤差404と、時間的にずれたフレームの物体抽出結果407を入力し、時間的にずれたフレームの任意の図形内画像を分割するか否かの判定結果406と、分割しない場合、動き405を出力する。ここでは、時間的にずれたフレームの物体抽出結果が分割されたブロック内に含まれていなければ、そのブロックは図形内から削除する。そうでなければ、求めた誤差から、誤差が閾値以上であれば、更に分割し、動きを求め直す。そうでなければ、ブロックの動きを決定する。図形決定部44は、動き405を入力とし、現フレームの図形内画像とその位置407を出力する。ここでは、図形決定部44は、各ブロックの、現フレームへの位置対応を求め、対応した位置のブロックを全て含むように、新しい図形を決定する。新しい図形は、全ブロックの結びつきも良く、全てのブロックを含むような長方形や円でもよい。 30

【0096】

このようにして、参照フレームの図形内画像を複数のブロックに分割し、複数のブロックそれぞれについて現フレームとの誤差が最小となる領域を探索し、そして探索された複数の領域を囲む図形を現フレームに設定することにより、初期設定された図形の形状や大きさによらず、図形設定対象となる入力フレームに対して最適な形状の新たな図形を設定することが可能となる。 40

【0097】

なお、図形設定に用いる参照フレームは既に図形が設定されていて且つ現フレームと時間的にずれたフレームであればよく、通常の符号化技術で前方向予測と後方向予測とが使用されていることと同様、現フレームよりも時間的に後のフレームを図形設定のための参照フレームとして用いることも可能である。

【0098】

図6には、背景動き削除部21の具体的な構成の一例が示されている。 50

【0099】

代表背景領域設定部51は、時間的にずれた任意の図形内画像とその位置501を入力とし、代表背景領域503を出力する。代表背景領域とは、任意の図形内のグローバルな動き、つまり図形内における背景の動きを代表して表す領域で、例えば、任意の図形を長方形にした場合、図7に示すような、長方形を囲む数画素幅の帯状の枠領域を設定する。また、図形内の外側の数画素を使っても良い。動き検出部52では、現フレーム502と、該代表背景領域503を入力し、動き504を出力する。先の例を用いると、長方形周囲の帯状の枠領域の現フレームに対する動きを検出する。枠領域を一つの領域として検出しても良い。また、図7のように、複数のブロックに分割して動きを求め、各々の平均動きを出力しても良いし、最も多い動きを出力しても良い。

10

【0100】

動き補償部53では、該時間的にずれたフレーム501と、該動き504を入力とし、動き補償画像505を出力する。先に求めた動きを使って、時間的にずれたフレームの動きを現フレームに合わせて削除する。動き補償は、ブロックマッチング動き補償またはアフィン変換を使った動き補償でもよい。

【0101】

なお、動き削除は、図形内の背景に対してのみならず、フレーム全体に対して行うようにしても良い。

【0102】

以上のように、本実施形態においては、(1)物体の輪郭ではなく、その物体を大まかに囲む図形を用いて物体を追跡していくこと、(2)現フレームに任意の図形を設定し、現フレームと少なくとも2つのフレームそれぞれの図形内画像との共通の背景領域を決定し、現フレームの物体を抽出すること、(3)該時間的にずれた少なくとも2つのフレームの背景の動きを削除すること、(4)任意の図形内画像間の変化量を検出し、代表領域を決定し、現フレームと少なくとも2つのフレームの図形内画像とその位置の、背景に対応する変化量を決定し、変化量と代表領域との関係から背景かどうかを判定すること、(5)図形内画像を分割し、任意の図形内画像又は分割された図形内画像の一部の動きを検出し、任意の図形内画像又は分割された図形内画像の一部を分割するか否かを判定し、現フレームの任意の図形内画像と位置を決定すること、(6)背景を代表する領域を設定し、背景の動きを検出し、該時間的にずれたフレームの背景の動きを削除した画像を作ることにより、目的の物体以外の周囲の余分な動きに影響を受けずに、且つ比較的簡単な処理で、目的の物体を精度良く抽出・追跡することが可能となる。

20

30

【0103】

また、本実施形態の物体抽出・追跡処理の手順はソフトウェア制御によって実現することもできる。この場合でも、基本的な手順は全く同じであり、図形の初期設定を行った後に、入力フレームに対して順次図形設定処理を行い、この図形設定処理と並行してあるいは図形設定処理完了後に、背景領域決定処理、物体抽出処理を行えばよい。

【0104】

次に、本発明の第2実施形態を説明する。

【0105】

前述の第1実施形態では、ORAND法による一つの物体抽出手段のみを備えていたが、入力画像によってはその手段だけでは十分な抽出性能が得られないこともある。また、第1実施形態のORAND法では、物体抽出対象となる現フレームと、この現フレームに対し時間的に異なる第1の参照フレームとの差分に基づいて共通の背景を設定し、また別の現フレームと時間的に異なる第2の参照フレームとの差分に基づいて共通の背景を設定している。しかし、この第1及び第2の参照フレームの選択方法は特に与えられていない。第1及び第2の参照フレームの選択によっては、物体の抽出結果に大きな差が生まれ、良好な結果が得られないことがあり得る。

40

【0106】

そこで、本第2実施形態では、入力画像によらずに物体を高精度に抽出できるように第

50

1 実施形態を改良している。

【0107】

まず、図9のブロック図を用いて、第2実施形態に係る物体追跡／抽出装置の第1の構成例について説明をする。

【0108】

以下では、第1実施形態の物体追跡／抽出部2に対応する構成のみについて説明する。

【0109】

図形設定部60は、図2で説明した第1実施形態の図形設定部11と同じものであり、フレーム画像601と、初期フレームまたは既に他の入力フレームに対して設定した図形602とを入力とし、フレーム画像601に図形を設定して出力する。スイッチ部SW61は、既に行われた物体抽出の結果605を入力とし、それに基づいて、使用すべき物体抽出部を切り替えるための信号604を出力する。 10

【0110】

物体追跡・抽出部62は、図示のように第1乃至第Kの複数の物体追跡・抽出部から構成されている。これら物体追跡・抽出部はそれぞれ異なる手法で物体抽出を行う。物体追跡・抽出部62の中には、第1実施形態で説明したORAND法を用いるものが少なくとも含まれている。また、別の方法による物体追跡・抽出部としては、例えばブロックマッチングによる形状予測法を用いたものや、アフィン変換による物体形状予測などを使用することができる。これら形状予測では、既に物体抽出されたフレームと現フレームとのフレーム間予測により現フレーム上の物体領域の位置または形状が予測され、その予測結果に基づいて現フレームの図形内画像603から物体領域が抽出される。 20

【0111】

ブロックマッチングによる形状予測の一例を図13に示す。現フレームの図形内画像は図示のように同じ大きさのブロックに分割される。各ブロック毎に絵柄（テクスチャ）が最も類似したブロックが、既に物体の形状及び位置が抽出されている参照フレームから探索される。この参照フレームについては、物体領域を表すシェイプデータが既に生成されている。シェイプデータは、物体領域に属する画素の画素値を“255”、それ以外の画素値を“0”で表したものである。この探索されたブロックに対応するシェイプデータが、現フレームの対応するブロック位置に張り付けられる。このようなテクスチャの探索およびシェイプデータの張り付け処理を現フレームの図形内画像を構成する全ブロックについて行うことにより、現フレームの図形内画像は、物体領域と背景領域を区別するシェイプデータによって埋められる。よって、このシェイプデータを用いることにより、物体領域に対応する画像（テクスチャ）を抽出することができる。 30

【0112】

スイッチ部SW61は、例えば第1の物体追跡抽出部と同様の操作を行ない、抽出精度が良い場合は第1の物体追跡抽出部を選ぶよう切り替え、そうでない場合は別の物体追跡抽出部を選ぶよう切り替える。例えば、第1の物体追跡抽出部が、ブロックマッチングによる物体形状予測手段であるとすれば、マッチング誤差の大きさによって物体追跡抽出部の切り替えを制御すればよい。また、アフィン変換による物体形状予測であれば、アフィン変換係数の推定誤差の大きさによって物体追跡抽出部を切り替えることができる。スイッチ部SW61での切替の単位は、フレーム単位ではなく、フレーム内の小領域、例えばブロック毎や、輝度や色に基づいて分割した領域毎である。これにより、使用する物体抽出法をよりきめ細かに選択することが出来、抽出精度を高めることができる。 40

【0113】

図10には、第2実施形態に係る動画像物体追跡／抽出装置の第2の例が示されている。

【0114】

図形設定部70は図2で説明した第1実施形態の図形設定部11と同じものであり、画像701と、初期フレームまたは既に他の入力フレームに対して設定した図形702とを入力し、フレーム画像701に図形を設定して出力する。 50

【0115】

第2の物体追跡・抽出部71は、ブロックマッチング法やアフィン変換などの形状予測によって物体領域を抽出するために使用され、図形設定部70から出力される現フレームの図形内画像703と、既に抽出されている別の参照フレーム上の物体の形状及び位置707を入力とし、現フレームの図形内画像703から物体の形状及び位置を予測する。

【0116】

参照フレーム選択部72は、第2の物体追跡・抽出部71によって予測された現フレームの物体の予測形状及び位置704と、既に抽出されている物体の形状及び位置707とを入力し、少なくとも2つの参照フレームを選択する。ここで、参照フレームの選択方法について説明する。

10

【0117】

O_i , O_j , O_{curr} は各々フレーム i , j 及び抽出中のフレーム $curr$ の物体とする。2つの時間的に異なる参照フレーム f_i , f_j との差分 d_i , d_j を取って、これら差分をAND処理して現フレーム f_{curr} の物体を抽出すると、抽出したい物体 O_{curr} 以外に、物体 O_i , O_j の重なり部分が時間的に異なるフレームのAND処理により抽出される。勿論、 $O_i \cap O_j = \phi$ 、つまり物体 O_i , O_j の重なり部分が存在せず、物体 O_i , O_j の重なりが空集合となる場合には問題ない。

【0118】

しかし物体 O_i , O_j の重なり部分が存在し ($O_i \cap O_j \neq \phi$)、かつ、この重なり部分が抽出したい物体の外部に存在する場合は、 O_{curr} と、 $O_i \cap O_j$ の二つが抽出結果として残る。

20

【0119】

この場合、図14(a)のように、 O_{curr} の背景領域 (O_{curr}^-) と、物体 O_i と、 O_j との全ての共通領域が存在しない場合 $\{O_{curr}^- \cap (O_i \cap O_j) = \phi\}$ であれば、問題はない。しかし、図14(b)のように、 O_{curr} の背景領域 (O_{curr}^-) と、物体 O_i と、 O_j との全ての共通領域が存在する場合 $\{O_{curr}^- \cap (O_i \cap O_j) \neq \phi\}$ は、 O_{curr} が斜線で示すような誤った形状で抽出される。

【0120】

従って、正しく物体の形状を抽出する最適な参照フレーム f_i , f_j とは、

$$(O_i \cap O_j) \cap O_{curr} \quad \dots (1)$$

30

を満たすフレーム、つまり、 O_i , O_j の重なり部分が O_{curr} 内に属するようなフレーム f_i , f_j である (図14(a))。

【0121】

また、2つ以上の参照フレームを選ぶ場合は、

$$(O_i \cap O_j \cap \dots \cap O_k) \cap O_{curr} \quad \dots (2)$$

となる。

【0122】

したがって、物体抽出対象となる現フレーム上の物体の位置または形状の予測結果に基づいて、(1)式または(2)式を満足するような参照フレームを選択することにより、確実に物体の形状を抽出することが可能になる。

40

【0123】

第1の物体追跡・抽出部73では、参照フレーム選択部72で選択された少なくとも2つの参照フレーム705と、現画像701を入力し、ORAND法により物体を抽出してその形状及び位置706を出力する。

【0124】

メモリ74には、抽出された物体形状及び位置706が保持されている。

【0125】

図11には、第2実施形態に係る物体追跡／抽出装置の第3の構成例が示されている。

【0126】

この物体追跡／抽出装置は、図示のように、図形設定部80、第2の物体追跡・抽出部

50

8 1、スイッチ部 S W 8 2、および第 1 の物体抽出部 8 3 から構成されている。図形設定部 8 0、第 2 の物体追跡・抽出部 8 1、および第 1 の物体抽出部 8 3 は、それぞれ図 1 0 の図形設定部 7 0、第 2 の物体追跡・抽出部 7 1、および第 1 の物体抽出部 7 3 に対応している。本例では、スイッチ部 S W 8 2 によって、第 2 の物体追跡・抽出部 8 1 の抽出結果と第 1 の物体抽出部 8 3 の抽出結果が選択的に使用される。

【0 1 2 7】

すなわち、図形設定部 8 0 では、画像 8 0 1 と初期図形の形状及び位置 8 0 2 を入力し、図形の形状及び位置 8 0 3 を出力する。第 2 の物体追跡・抽出部 8 1 では、図形の形状及び位置 8 0 3 と、既に抽出されている物体の形状及び位置 8 0 6 を入力し、未だ抽出されていない物体の予測形状及び位置 8 0 4 を予測し、出力する。スイッチ部 S W 8 2 では、第 2 の物体抽出部で予測された物体の形状及び位置 8 0 4 を入力し、第 1 の物体追跡・抽出部を行なうかどうかを切り替える信号 8 0 5 を出力する。物体追跡・抽出部 8 3 では、既に抽出されている物体の形状及び位置 8 0 6 と、未だ抽出されていない物体の予測形状及び位置 8 0 4 を入力し、物体の形状及び位置 8 0 5 を決定し、出力する。

【0 1 2 8】

スイッチ部 S W 8 2 での切替の単位は、上記で述べた例と同様にブロック毎に切り替えても良いし、輝度や色に基づいて分割した領域毎に切り替えても良い。切替を判断する方法として、例えば物体予測したときの予測誤差を用いることができる。すなわち、フレーム間予測を用いて物体抽出を行う第 2 の物体追跡・抽出部 8 1 における予測誤差が所定のしきい値以下の場合には、第 2 の物体追跡・抽出部 8 1 によって得られた予測形状が抽出結果として使用されるようにスイッチ部 S W 8 2 による切替が行われ、第 2 の物体追跡・抽出部 8 1 における予測誤差が所定のしきい値を越えた場合には、第 1 の物体追跡・抽出部 8 3 によって O R A N D 法にて物体抽出が行われるようにスイッチ部 S W 8 2 による切替が行われ、その抽出結果が外部に出力される。

【0 1 2 9】

図 1 5 は、予測の単位となるブロック毎にマッチング誤差に基づいて、使用する抽出部を切り替えた場合の抽出結果の例を示している。

【0 1 3 0】

ここで、網目で示した領域は第 2 の物体追跡・抽出部 8 1 による予測で得られた物体形状であり、斜線で示した領域は第 1 の物体追跡・抽出部 8 3 によって得られた物体形状である。

【0 1 3 1】

図 1 2 には、本第 2 実施形態に係る動画像物体追跡／抽出装置の第 4 の構成例が示されている。

【0 1 3 2】

この物体追跡／抽出装置は、図 1 1 の構成に加え、図 1 0 の参照フレーム選択部を追加したものである。

【0 1 3 3】

図形設定部 9 0 では、画像 9 0 1 と初期図形の形状及び位置 9 0 2 を入力し、図形の形状及び位置 9 0 3 を出力する。第 2 の物体追跡・抽出部 9 1 では、図形の形状及び位置 9 0 3 と、既に抽出されている物体の形状及び位置 9 0 8 を入力し、未だ抽出されていない物体の予測形状及び位置 9 0 4 を予測し、出力する。スイッチ部 S W 9 2 は、その予測した物体の形状及び位置 9 0 4 を入力とし、予測物体の精度が良いか否かを判断し、第 2 の物体抽出部で抽出された物体の出力を切り替える信号 9 0 5 を出力する。参照フレーム選択部 9 3 は、未だ抽出されていない物体の予測形状及び位置 9 0 4 と既に抽出されている物体の形状及び位置 9 0 8 を入力とし、少なくとも 2 つの参照フレームの物体または予測物体の形状及び位置 9 0 6 を選択し、出力する。物体追跡・抽出部 9 4 は、現画像 9 0 1 と、少なくとも 2 つの参照フレームの物体又は予測物体の形状及び位置 9 0 6 を入力とし、物体を抽出して、その形状及び位置 9 0 7 を出力する。メモリ 9 5 は、抽出した物体の形状及び位置 9 0 7 と、その予測した物体の形状及び位置 9 0 4 いずれかを保持する。

【0134】

以下、図16を参照して、本例における物体追跡／抽出方法の手順を説明する。

【0135】

(ステップS1)

参照フレームの候補としては、現フレームと時間的にずれたフレームが予め設定される。これは現フレーム以外の全てのフレームでも良いし、現フレーム前後の数フレームと限定してもよい。例えば、初期フレームと、現フレームより前3フレーム、現フレームより後1フレームの合計5フレームに限定する。ただし、前フレームが3フレームない場合はその分後のフレームの候補を増やし、後フレームが1フレームない場合はその分前4フレームを候補とする。

10

【0136】

(ステップS2)

まず、ユーザが初期フレームに抽出したい物体を書こむ図形を例えば長方形で設定する。以降のフレームの図形は初期設定の図形をブロックに分割し、マッチングを取って対応する位置にブロックを張り付ける。全ての張り付けたブロックを含むように新たに長方形を設定することで物体を追跡する。全ての参照フレーム候補に物体追跡の図形を設定する。物体が抽出される度にそれを使って先のフレームの物体追跡図形を求め直すほうが抽出エラーを防ぐことができる。また、ユーザは初期フレームの物体形状を入力する。

【0137】

以下、抽出するフレームは現フレームとし、現フレームより前のフレームは物体が既に抽出されており、先のフレームは抽出されていないとする。

20

【0138】

(ステップS3)

参照フレーム候補の図形の周辺に適当な領域を設定し、現フレームとの背景の動きを検出して参照フレームの図形内の背景を削除する。背景の動きを検出する方法として、図形の周囲数画素の幅の領域を設定し、この領域を現フレームに対してマッチングを取り、マッチング誤差が最小となる動きベクトルを背景の動きとする。

【0139】

(ステップS4)

背景動き削除時の動きベクトル検出誤差が大きい参照フレームは候補から外すことにより、背景動き削除が適当でない場合の抽出エラーを防ぐことができる。また、参照フレーム候補が減った場合、新たに参照フレーム候補を選び直してもよい。新たに付け加えた参照フレーム候補の図形設定や背景動き削除が行なわれていない場合は、新たに図形設定および背景動き削除を行なう必要がある。

30

【0140】

(ステップS5)

次に、未だ物体が抽出されていない現フレームと、現フレームより先の参照フレームの候補の物体形状を予測する。現フレーム又は先の参照フレームの候補に設定された長方形を例えばブロックに分割して既に物体が抽出されているフレーム（前のフレーム）とマッチングを取り、対応する物体形状を張り付けて物体形状を予測する。物体が抽出される度にそれを使って先のフレームの物体予測をやり直すほうが抽出エラーを防ぐことができる。

40

【0141】

(ステップS6)

この時、予測誤差が小さいブロックは、予測した形状を抽出結果としてそのまま出力する。また物体形状の予測をブロック単位で処理を行なうとマッチング誤差によりブロック歪みが生じる場合があるので、それを消去するようなフィルターをかけ、全体の物体形状を滑らかにしてもよい。

【0142】

物体追跡及び物体形状予測の時に行なう長方形の分割は、固定ブロックサイズで行なっ

50

ても良いし、マッチング閾値によって階層的ブロックマッチングによって行なっても良い。

【0143】

予測誤差が大きいブロックについては、以下の処理を行う。

【0144】

(ステップS7)

参照フレームの候補から仮の参照フレームを設定し、各々の組合せについて、式(1)又は式(2)を満たす参照フレームのセットを選ぶ。全参照フレーム候補のどの組合せも式(1)又は式(2)を満たさなかった場合、 $O_i \cap O_j$ 内の画素数が最小のものを選ぶのがよい。また、背景動き削除時の動きベクトル検出誤差がなるべく小さいフレームを選ぶように、参照フレーム候補の組合せを考慮するほうがよい。具体的には、式(1)又は式(2)による条件が同じ参照フレームセットがあるばあい、背景動き削除時の動きベクトル検出誤差が小さい方を選ぶ、などの方法がある。以下、参照フレームは2フレーム選択されたとする。

【0145】

(ステップS8)

参照フレームが選択されると、現フレームとのフレーム間差分を求め、設定された図形内のフレーム間差分に注目する。設定された図形の外側1ライン画素の差分の絶対値のヒストグラムを求め、多く現れる差分の絶対値を背景領域の差分値とし、設定された図形外側の1ライン画素の背景画素を決定する。設定された図形外側の1ライン画素の背景画素から内側に向けて、隣接する背景領域の差分値をもつ画素を背景画素と決定し、背景画素でないと判定されるまで順次続ける。この背景画素は、現フレームと一つの参照フレームとの共通の背景領域となる。この時、ノイズの影響で背景領域とそれ以外の部分の境界が不自然であることがあるので、境界を滑らかにするフィルターや、余分やノイズ領域を削除するフィルターをかけてもよい。

【0146】

(ステップS9)

各々の参照フレームに対して共通の背景領域が求まると、二つの共通背景領域に含まれない領域を検出し、それを物体領域として抽出する。先に予測した物体の形状を用いない部分について、ここでの結果を出力し、物体全体の形状を出力する。

【0147】

共通の背景から求めた形状を用いる部分と先に予測した物体形状を用いる部分の整合性が取れない場合、最後にフィルターをかけて出力結果を見ために良いものにできる。

【0148】

以上説明したように、本第2実施形態によれば、入力画像によらずに物体を精度良く抽出できる。又は、物体抽出に適した参照フレームを選択することができる。

【0149】

次に、本発明の第3実施形態を説明する。

【0150】

まず、図17のブロック図を用いて、第3実施形態に係る物体追跡／抽出装置の第1の例を説明をする。

【0151】

ここでは、物体抽出対象となる現フレームから、その少なくとも一部の領域についての画像の特徴量を抽出し、その特徴量に基づいて複数の物体抽出手段を切り替える構成が採用されている。

【0152】

すなわち、本物体追跡／抽出装置は、図示のように、図形設定部110、特徴量抽出部111、スイッチ部SW112、複数の物体追跡・抽出部113、およびメモリ114から構成されている。図形設定部110、スイッチ部SW112、複数の物体追跡・抽出部113は、それぞれ第2実施形態で説明した図9の図形設定部60、スイッチ部SW61

、および複数の物体追跡・抽出部 6 2 と同じであり、特徴量抽出部 1 1 1 によって抽出された現フレームの画像の特徴量に基づいて、使用する物体追跡・抽出部の切替が行われる点が異なっている。

【0 1 5 3】

図形設定部 1 1 0 は、抽出フレーム 1 1 0 1 と、ユーザ設定による初期図形 1 1 0 2 と、既に抽出されたフレームの抽出結果 1 1 0 6 を入力とし、抽出フレームに図形を設定してその図形を出力する。図形は長方形、円、楕円、など幾何図形でもよいし、ユーザが物体形状を図形設定部 1 1 0 に入力してもよい。その場合、図形は精密な形状でなくても、大まかな形状でもよい。特徴量検出部 1 1 1 は、図形が設定された抽出フレーム 1 1 0 3 と、既に抽出されたフレームの抽出結果 1 1 0 6 とを入力とし、特徴量 1 1 0 4 を出力する。スイッチ部 S W 1 1 2 は、特徴量 1 1 0 4 と、既に抽出されたフレームの抽出結果 1 1 0 6 とを入力とし、図形が設定された抽出フレーム 1 1 0 3 の物体追跡・抽出部への入力を制御する。

10

【0 1 5 4】

スイッチ部 S W 1 1 2 は、画像全体に対して特徴量を得た場合は、画像の性質を検出し、画像に対して適当な物体追跡・抽出部への入力の制御に用いることができる。図形内部は適当な大きさに分割され、特徴量は各分割図形毎に与えても良い。特徴量は分散や輝度勾配、エッジ強度などであり、この場合は、これらを自動的に算出することができる。また、人間が視覚的に認知した物体の性質がユーザによってスイッチ部 W 1 1 2 に与えられてもよい。例えば、目的とする物体が人物であれば、エッジが不鮮明な髪の毛を指定して抽出時のパラメータが特別に選ばれ、前処理にエッジ補正してから抽出されてもよい。

20

【0 1 5 5】

特徴量は、設定された図形内部（物体及びその周辺）に関してだけでなく、図形外部（背景部）に関する特徴量でも良い。

【0 1 5 6】

複数（第 1 ～ k）の物体追跡・抽出部 1 1 3 の各々では、図形が設定された抽出フレーム 1 1 0 3 と、既に抽出されたフレームの抽出結果 1 1 0 6 とを入力とし、物体を追跡・抽出した結果 1 1 0 5 を出力する。

【0 1 5 7】

複数の物体追跡・抽出部 1 1 3 は、O R A N D 法を使用して物体を抽出するもの、クロマキーを使用して物体を抽出するもの、ブロックマッチングやアフィン変換によって物体を抽出するものなどを含む。

30

【0 1 5 8】

なお、実施形態 1 では、設定された図形の周囲の画素値のフレーム間差分のヒストグラムを用いて、背景画素が決定されているが、単純に、フレーム間差分が閾値以下の画素が背景画素と決定されても良い。また、実施形態 1 では、設定された図形から図形内部に向かって背景画素（差分値が一定値以下）が決定されているが、図形から図形外部へ向けて物体画素（差分値が一定値以上）も決定されても良いし、任意の操作順でもよい。

【0 1 5 9】

メモリ 1 1 4 は、物体を追跡・抽出した結果 1 1 0 5 を入力とし、それを保持する。

40

【0 1 6 0】

以下、画像の性質を示す特徴量によって、追跡／抽出方法を切替えると、よりよい抽出結果が得られる理由について説明する。

【0 1 6 1】

例えば、背景の動きがあるかどうか予め分かるならば、その性質を使った方がよい。背景の動きがある場合は、背景動き補償が行なわれるが、完全に補償できるかわからない。複雑な動きをするフレームではほとんど動き補償できない。このようなフレームは、背景動き補償の補償誤差で予め分かるので、参照フレーム候補にしないなど工夫が可能である。しかし、背景の動きがない場合は、この処理は不必要である。別の物体が動いていると、誤った背景動き補償が行なわれたり、そのフレームは参照フレーム候補から外れたりし

50

て参照フレーム選択条件に最適なフレームであっても選ばれず、抽出精度が落ちることがある。

【0162】

また、一つの画像中にも多様な性質が混在している。物体の動きやテクスチャも部分的に異なり、同じ追跡・抽出方法及び装置やパラメータではうまく物体が抽出できないことがある。従って、ユーザが画像中の特殊な性質を持つ一部を指定したり、画像中の違いを自動的に特徴量として検出して、部分的に追跡・抽出方法を切替えて物体を抽出したり、パラメータを変更した方がよい。

【0163】

このようにして、複数の物体の追跡・抽出手段を切替えれば、様々な画像中の物体の形状を精度良く抽出することが可能になる。 10

【0164】

次に、図18のブロック図を用いて、第3実施形態に係る動画像物体追跡／抽出装置の第2の構成例について説明する。

【0165】

図形設定部120は、抽出フレーム1201と、ユーザ設定による初期図形1202と、既に抽出されたフレームの抽出結果1207を入力とし、抽出フレームに図形を設定して出力する。第2の物体追跡・抽出部121は、ブロックマッチング法やアフィン変換などの形状予測によって物体領域を抽出するために使用され、図形が設定された抽出フレーム1203と、既に抽出されたフレームの抽出結果1207を入力とし、物体の追跡・抽出結果1204を出力する。 20

【0166】

特徴量抽出部122は、物体の追跡・抽出結果1204を入力とし、物体の特徴量1205をスイッチ部SW123に出力する。スイッチ部SW123は、物体の特徴量1205を入力として、第一の物体追跡・抽出部への物体の追跡・抽出結果1204の入力を制御する。例えば、第2の物体追跡・抽出部121でブロックマッチング法により物体形状が追跡・抽出された場合、特徴量をマッチング誤差として、このマッチング誤差が小さい部分は第2の物体追跡・抽出部121による予測形状の抽出結果として出力される。また他の特徴量として、ブロック毎に輝度勾配や分散、テクスチャの複雑さを表すパラメータ（フラクタル次元など）がある。輝度勾配を用いた場合、輝度勾配がほとんどないブロックに対しては、ORAND法による第1の物体追跡・抽出部124の結果が使用されるように第一の物体追跡・抽出部への入力が制御される。またエッジ検出をして、エッジの有無や強度を特徴量とした場合、エッジのない所、弱い所では第1の物体追跡・抽出部124の結果が使用されるように第一の物体追跡・抽出部への入力が制御される。このように、画像の一部分であるブロック単位や領域単位で、切替の制御が変えられる。切替の閾値を大きくしたり小さくしたりすることで、適応的な制御ができる。 30

【0167】

第1の物体追跡・抽出部124は、抽出フレーム1201と、物体の追跡・抽出結果1204と、既に抽出されたフレームの抽出結果1207を入力とし、抽出フレームの追跡・抽出結果1206をメモリ125に出力する。 40

【0168】

メモリ125は、抽出フレームの追跡・抽出結果1206を入力とし、保持する。

【0169】

次に、図19のブロック図を用いて、本第3実施形態に係る物体追跡／抽出装置の第3の構成例をする。

【0170】

この物体追跡／抽出装置は、図18の構成に加え、第2実施形態で説明した参照フレーム選択部を追加したものである。すなわち、この物体追跡／抽出装置は、図示のように、図形設定部130、第2の物体追跡・抽出部131、特徴量抽出部132、スイッチ部SW133、参照フレーム選択部134、第1の物体追跡・抽出部135、およびメモリ1 50

36から構成されている。

【0171】

図形設定部130では、抽出フレーム1301と、ユーザ設定による初期図形1302と、既に抽出されたフレームの抽出結果1308を入力とし、抽出フレームに図形を設定して出力する。第2の物体追跡・抽出部131は、ブロックマッチング法やアフィン変換などの形状予測によって物体領域を抽出するためのものであり、図形を設定された抽出フレーム1303と、既に抽出されたフレームの抽出結果1308を入力とし、物体の追跡・抽出結果1304を出力する。

【0172】

特徴量抽出部132は、物体の追跡・抽出結果1304を入力とし、物体の特徴量1305を出力する。スイッチ部SW133は、物体の特徴量1305を入力とし、第1の物体追跡・抽出部135への物体の追跡・抽出結果1304の入力を制御する。

【0173】

参照フレーム選択部134は、第1の物体追跡・抽出部135への物体の追跡・抽出結果1304と、既に抽出されたフレームの抽出結果1308を入力とし、参照フレーム1306を出力する。

【0174】

物体の特徴の一例として、動きの複雑さがある。第2の物体追跡・抽出部131でブロックマッチング法により物体を追跡・抽出する場合、マッチング誤差が大きい部分に対して、第1の物体抽出結果が出力される。部分的に複雑な動きがあると、その部分はマッチング誤差が大きくなり、第1の物体追跡・抽出部135で抽出されることになる。従って、このマッチング誤差を特徴量として第1の物体追跡・抽出部135で用いる参照フレームの選択方法が切替えられる。具体的には物体形状全体ではなく、第1の物体追跡・抽出部135で抽出する部分だけについて第2実施形態で説明した式(1)または(2)の選択条件を満たすように参照フレームの選択方法を選ぶのがよい。

【0175】

背景の特徴量の例は、1)背景が静止している画像である、2)ズームがある、3)パンがあるという情報等である。この特徴量はユーザが入力しても良いし、カメラから得たパラメータが特徴量として入力されても良い。背景の特徴量としては、背景の動きベクトル、背景動き補正画像の精度、背景の輝度分布、テクスチャ、エッジなどがある。例えば背景動き補正画像の精度を背景動き補正画像と補正前画像との差分平均を特徴量として、参照フレーム選択方法が制御できる。制御例としては、差分平均が非常に多い場合、そのフレームは参照フレームの候補にしなかったり、そのフレームの選択順位を下げてフレームを選んだりできる。背景が静止している場合や、背景動き補正がすべてのフレームについて完全であると差分がゼロとなる。参照フレーム選択法は、第2実施形態と同じ方法を用いることができる。

【0176】

第1の物体追跡・抽出部135は、抽出フレーム1301と、参照フレーム1306と、既に抽出されたフレームの抽出結果1308を入力とし、ORAND法により抽出フレームの追跡・抽出結果1307をメモリ135に出力する。メモリ135は、抽出フレームの追跡・抽出結果1307を入力し、保持する。

【0177】

次に、図22を用いて、先に挙げた例のうち、第2の物体追跡・抽出部からの出力から特徴量を得て、それによって複数の参照フレーム選択部を切替える例を第4の構成例として説明する。

【0178】

図形設定部160は、抽出フレーム1601と、ユーザが設定した初期図形1602と、既に物体が抽出されたフレーム1608を入力とし、設定図形1603を出力する。第2の物体追跡・抽出部161は、ブロックマッチング法やアフィン変換などの形状予測によって物体領域を抽出するために使用され、設定図形1603と、既に物体が抽出された

フレーム 1608 を入力とし、物体追跡・抽出結果 1604 を出力する。特徴量検出部 163 は、物体追跡・抽出結果 1604 を入力とし、特徴量 1605 をスイッチ部 SW164 に出力する。スイッチ部 SW164 は、特徴量 1605 を入力し、参照フレーム選択部への物体追跡・抽出結果 1604 の入力を制御する。

【0179】

複数の参照フレーム選択部 165 は、物体追跡・抽出結果 1604 と、既に物体が抽出されたフレーム 1608 を入力とし、少なくとも 2 つの参照フレーム 1606 を出力する。

【0180】

第 1 の物体追跡・抽出部 166 は、ORAND 法により物体抽出を行うために使用され、参照フレーム 1696 と、抽出フレーム 1601 を入力とし、物体の追跡・抽出結果 1607 をメモリ 167 に出力する。メモリ 167 は、物体の追跡・抽出結果 1607 を入力とし、保持する。 10

【0181】

次に、ブロック図 23 を用いて、先に述べた例のうち、背景の情報を得て、背景動き補正の誤差によって複数の参照フレーム選択部の入力を制御する例について説明する。

【0182】

図形設定部 170 は、抽出フレーム 1701 と、ユーザが設定した初期図形 1702 と、既に物体が抽出されたフレーム 1710 を入力とし、設定図形 1703 を出力する。第 2 の物体追跡・抽出部 171 は、設定図形 1703 と、既に物体が抽出されたフレーム 1710 を入力とし、物体の追跡・抽出結果 1704 を出力する。スイッチ部 SW172 では、ユーザが指定した背景の情報 1705 を入力し、背景動き補正部 173 への抽出フレーム 1701 の入力を制御する。 20

【0183】

背景動き補正部 173 は、抽出フレーム 1701 と、既に物体が抽出されたフレーム 1710 を入力とし、背景動きを補正したフレーム 1706 を出力する。

【0184】

背景特徴量検出部 174 は、抽出フレーム 1701 と、背景動きを補正したフレーム 1706 を入力とし、背景の特徴量 1707 をスイッチ部 SW175 へ出力する。このスイッチ部 SW175 は、背景の特徴量 1707 を受け、参照フレーム選択部 176 への物体の追跡・抽出結果 1704 の入力を制御する。参照フレーム選択部 176 は、物体の追跡・抽出結果 1704 と、既に物体が抽出されたフレーム 1710 を入力し、少なくとも 2 つの参照フレーム 1708 を出力する。 30

【0185】

第 1 の物体追跡・抽出部 177 は、少なくとも 2 つの参照フレーム 1708 と、抽出フレーム 1701 を入力とし、物体の追跡・抽出結果 1709 をメモリ 178 に出力する。メモリ 178 は、物体の追跡・抽出結果 1709 を受け、保持する。

【0186】

次に、図 20 のブロック図を用いて、本第 3 実施形態に係る物体追跡／抽出装置の第 5 の構成例を説明する。 40

【0187】

抽出フレーム出力制御部 140 は、画像 1401 と抽出するフレームの順序 1405 を入力とし、抽出フレーム 1402 を出力する。フレーム順序制御部 141 は、ユーザが与えたフレーム順序に関する情報 1405 を入力とし、フレーム順序 1406 を出力する。物体追跡・抽出装置 142 は、動画像信号から目的とする物体の抽出／追跡を行なう物体追跡／抽出方法及び装置であり、抽出フレーム 1402 を入力とし、追跡・抽出結果 1403 を追跡・抽出結果出力制御部 143 に出力する。追跡・抽出結果出力制御部 143 は、追跡・抽出結果 1403 と、フレーム順序 1406 を入力とし、フレーム順序を画像 1401 の順序に並び変えて出力する。

【0188】

フレームの順序は、ユーザが与えても良いし、物体の動きに応じて適応的に決定しても良い。物体の動きが検出しやすいフレーム間隔が決定され、物体が抽出される。すなわち、参照フレームと物体抽出対象となる現フレームとの間のフレーム間隔が少なくとも2フレーム以上となるように入力フレーム順とは異なる順序で物体抽出処理が行われるようにフレーム順の制御が行われる。これにより、入力フレーム順にフレーム間予測による形状予測や、ORAND演算を行う場合に比べ、予測精度を向上でき、結果的に抽出精度を高めることが可能となる。ORAND法の場合には適切な参照フレームを選択することによって抽出精度を高めることが可能となるため、ブロックマッチングなどによるフレーム間予測による形状予測法について特に効果がある。

【0189】

10

すなわち、フレームの間隔によっては動きが小さ過ぎたり、複雑過ぎて、フレーム間予測による形状予測手法では対応できないことがある。従って、例えば形状予測の誤差が閾値以下にならない場合は、予測に用いる抽出済みフレームとの間隔をあけることにより、予測精度が上がり、結果的に抽出精度が向上する。また、背景に動きがある場合は、参照フレーム候補は抽出フレームとの背景動きを求め補償するが、背景の動きがフレームの間隔によっては小さ過ぎたり複雑過ぎたりして、背景動き補償が精度良くできない場合がある。この場合もフレーム間隔をあけることによって動き補償精度を上げることができる。このようにして抽出フレームの順序を適応的に制御すれば、より確実に物体の形状を抽出することが可能になる。

【0190】

20

次に、図21のブロック図を用いて、本第3実施形態に係る物体追跡／抽出装置の第6の例を説明をする。

【0191】

抽出フレーム出力制御部150は、画像1501と抽出するフレームの順序1505を入力とし、抽出フレーム1502を出力する。フレーム順序制御部151は、ユーザが与えたフレーム順序に関する情報1505を入力とし、フレーム順序1506を出力する。すなわち、フレーム順序制御部151は、フレーム間隔が与えられ、フレームの抽出順序を決定する。複数の物体追跡／抽出装置152は、動画像信号から目的とする物体の抽出／追跡を行なう物体追跡／抽出方法及び装置であり、フレーム順序1506にしたがって抽出フレーム1502の入力が制御され、追跡・抽出結果1503を出力する。追跡・抽出結果出力制御部153は、追跡・抽出結果1503と、フレーム順序1506を入力とし、それを画像1501の順序に並び変えて出力する。

30

【0192】

飛ばされた間のフレームは、既に抽出されたフレームから内挿しても良いし、参照フレーム候補の選び方を変えて同じアルゴリズムで抽出しても良い。

【0193】

ここで、図25を用いて、図21の物体の追跡／抽出装置の処理の例について説明する。

【0194】

図25で、斜線で示すフレームは2フレーム間隔開けて先に抽出するフレームである。飛ばされたフレームは第2の物体追跡・抽出装置によって抽出される。図25のように両脇のフレームが抽出された後に、両脇のフレームの抽出結果から内挿して物体形状を求めてもよい。また、閾値などのパラメータを変えたり、これら両脇のフレームを参照フレーム候補に加えて、両脇のフレームと同じ方法で抽出してもよい。

40

【0195】

次に、図24のブロック図を用いて、物体追跡・抽出装置の他の構成例を説明する。

【0196】

スイッチ部SW182は、ユーザが指定した背景の情報1805を入力とし、背景動き補正部183への抽出フレーム1801の入力を制御する。背景動き補正部183は、抽出フレーム1801と、既に物体が抽出されたフレーム1811を入

50

力とし、背景動きを補正したフレーム1806を出力する。背景特徴量検出部184は、抽出フレーム1801と、背景動きを補正したフレーム1806を入力とし、背景の特徴量1807を出力する。スイッチ部SW187は、背景の特徴量1807を入力とし、参照フレーム選択部188への物体の追跡・抽出結果1804の入力を制御する。図形設定部180は、抽出フレーム1801と既に物体が抽出されたフレーム1811並びにユーザが設定した初期図形1802を入力とし、図形を設定した抽出フレーム1803を出力する。第二の物体追跡・抽出部185は、図形を設定した抽出フレーム1803と既に物体が抽出されたフレーム1811を入力とし、物体追跡・抽出結果1804を出力する。特徴量検出部185は、物体追跡・抽出結果1804を入力とし、特徴量1808を出力する。スイッチ部SW186は、特徴量1808を入力とし、参照フレーム選択部への物体追跡・抽出結果1804の入力を制御する。参照フレーム選択部188は、物体追跡・抽出結果1804と、既に物体が抽出されたフレーム1811を入力とし、少なくとも2つの参照フレーム1809を出力する。

10

【0197】

第一の物体追跡・抽出部189は、少なくとも2つの参照フレーム1809と、抽出フレーム1801を入力とし、物体の追跡・抽出結果1810をメモリ190に出力する。メモリ190は、物体の追跡・抽出結果1810を保持する。

【0198】

処理の流れは以下になる。

【0199】

ユーザが初期フレームにおいて抽出したい物体を大まかに囲む。以降のフレームの長方形は既に抽出された物体を囲む長方形を上下左右に数画素広げて設定する。この長方形をブロックに分割し、抽出済みのフレームとマッチングを取って対応する位置に抽出済み物体の形状を張り付ける。この処理によって得られた物体形状（予測物体形状）が大まかな物体を表す。予測の精度がある閾値以下にならない場合、別のフレームから予測を直してより予測精度を上げるように処理しても良い。

20

【0200】

予測精度が良い場合、この予測形状全部又は一部をそのまま抽出結果として出力する。この方法は、物体を追跡しつつ、物体も抽出できる。

【0201】

物体追跡及び物体形状予測の時にこなうブロック化は、長方形を固定ブロックサイズで分割しても良いし、マッチング閾値によって階層的ブロックマッチングによって行なっても良い。フレームを固定のサイズで分割し、物体を含むブロックだけを用いても良い。

30

【0202】

予測が悪い場合を考えて、物体予測形状を数画素分拡張して、予測エラーによる凹凸や穴が修正される。この方法で全ての参照フレーム候補に予測物体形状が設定される。物体が抽出される度にその物体を使って先のフレームの物体追跡図形が求め直されそれにより抽出エラーが防がれる。この追跡図形が物体を囲むように設定された追跡図形とする。

【0203】

以下、抽出フレームより前のフレームについては物体が既に抽出されており、先のフレームについては物体が抽出されていないものとする。

40

【0204】

参照フレームの候補は、一定間隔おきに抽出するフレームに対して時間的に一定間隔毎にずれた前後5フレームとする。参照フレームの候補は、例えば、初期フレームと現フレームより前3フレーム、現フレームより後1フレーム、の合計5フレームのように限定する。ただし、前フレームが3フレームない場合はその分後のフレームの候補を増やし、後フレームが1フレームない場合はその分前4フレームを候補とする。

【0205】

参照フレーム候補の物体の周辺に適当な領域が設定され、この領域と現フレームとの背景の動きが検出され、これにより参照フレームの図形内の背景が削除される。背景の動き

50

を検出する方法として、物体を除いた全領域で現フレームに対してマッチングを取り、マッチング誤差が最小となる動きベクトルが背景の動きと判定される。

【0206】

背景動き削除時の動きベクトル検出誤差が大きい参照フレームは候補から外すことにより、背景動き削除が適当でない場合の抽出エラーを防ぐことができる。また、参照フレーム候補が減った場合、新たに参照フレーム候補を選び直してもよい。新たに付け加えた参照フレーム候補の図形設定や背景動き削除が行なわれていない場合は、新たに図形設定や背景動き削除を行なう必要がある。

【0207】

予め背景の動きがないと分かる場合は、この処理は行なわない。

10

【0208】

参照フレームの候補から仮の参照フレームを設定し、これらフレームの組合せについて、第2実施形態の式(1)又は式(2)を満たす参照フレームのセットを選ぶ。全参照フレーム候補のどの組合せも式(1)又は式(2)を満たさなかった場合、 $O_i \cap O_j$ 内の画素数が最小のフレームを選ぶのがよい。

【0209】

また、背景動き削除時の動きベクトル検出誤差がなるべく小さいフレームを選ぶように、参照フレーム候補の組合せを考慮するほうがよい。具体的には、式(1)又は式(2)による条件が同じ参照フレームセットがある場合、背景動き削除時の動きベクトル検出誤差が小さい方のフレームを選ぶなどの方法がある。背景の動きがない場合は、フレーム間差分が十分検出できるフレームを優先的に選ぶようにできる。

20

【0210】

また、物体予測の精度がよく、物体の一部をそのまま出力する場合、物体予測結果を抽出結果としない領域のみを対象に、式(1)又は式(2)による条件を満たすフレームを選ぶ。

【0211】

以下、2参照フレームが選択されたときの処理を説明する。

【0212】

参照フレームが選択されると、抽出フレームとのフレーム間差分を求め、設定された図形内のフレーム間差分に注目する。

30

【0213】

設定された閾値でフレーム間差分を2値化する。2値化に用いる閾値は画像に対して一定でもよいし、背景動き補償の精度に応じてフレーム毎に変えても良い。制御例としては、背景動き補償の精度が悪ければ、背景に余分な差分が多く発生しているので、2値化の閾値を大きくする例などがある。また、物体の部分的な輝度勾配やテクスチャ、エッジ強度に応じて変えても良い。この制御例として、輝度勾配が少ない領域や、エッジ強度が小さい領域のように比較的平坦な領域は2値化の閾値を小さくする。更に、ユーザが物体の性質から閾値を与えても良い。

【0214】

物体追跡図形の外側の画素について、隣接する背景領域の差分値を持つ画素を背景画素と決定する。また、同時に物体追跡図形の内側の画素についても、隣接する背景領域の差分値をもたない画素を背景画素でない、と決定する。

40

【0215】

フレーム間差分は、物体の静止領域では検出できない。従って、予測に用いたフレームとのフレーム間差分がゼロで、かつ予測に用いたフレームでは物体内部の画素である場合は、静止領域画素として背景画素に加えない。

【0216】

この背景画素は、現フレームと一つの参照フレームとの共通の背景領域となる。この時、ノイズの影響で背景領域とそれ以外の部分の境界が不自然であることがあるので、画像信号に境界を滑らかにするフィルターや、余分なノイズ領域を削除するフィルターをかけ

50

てもよい。

【0217】

各々の参照フレームに対して共通の背景領域が求まると、二つの共通背景領域に含まれない領域が検出され、それが物体領域として抽出される。先に予測した物体の形状を用いない部分に対しては抽出結果が出力され、物体全体の形状が抽出される。共通の背景から求めた形状を用いる部分と先に予測した物体形状を用いる部分の整合性が取れない場合、最後にフィルターをかけて得た出力結果は見ために良いものにできる。

【0218】

最後に抽出順が入力フレームの順序に置き換えて抽出された物体領域が出力する。

【0219】

本発明のような物体の形状を抽出する方法及び装置は、現在標準化が固まりつつあるMPEG-4のオブジェクト符号化の入力手段として用いることができる。このMPEG-4と物体抽出の応用例として、物体形状をウィンドウ形式とする表示システムがある。このような表示システムは、多地点会議システムに有効である。限られた大きさのディスプレイにテキスト資料と、各地点で会議に参加する人物を四角いウィンドウで表示するよりも、図26のように人物は人物の形状で表示することにより、省スペース化できる。MPEG-4の機能を使えば、発言中の人物だけを大きくしたり、発言していない人物を半透明にしたりでき、システムの使用感がよくなる。

【0220】

以上説明したように、本第3実施形態によれば、画像の性質に応じて方法及び装置で物体を選ぶことによって、不必要な処理を省き、安定な抽出精度が得られる。また、時間順という制約を外すことによって物体の動きによらずに十分な抽出精度が得ることができる。

【0221】

また、本第3実施形態は、第1実施形態及び第2実施形態の性能を改善するものであり、第1実施形態及び第2実施形態の各構成と第3実施形態で説明した構成とを適宜組み合わせ使用することもできる。

【0222】

図27には、本発明の第4実施形態に係る物体抽出装置の第1の構成例が示されている。

【0223】

外部のカメラで撮像されたり、ビデオテープ、ビデオディスクなどの蓄積媒体から読み出されたりした後に、本物体抽出装置に入力されるテクスチャ画像221は、記録装置222、スイッチ部223、動き補償による物体抽出回路224に入力される。記録装置222は、入力されたテクスチャ画像221を保持するものである。例えば、パソコンなどで用いられているハードディスク、光磁気ディスクなどである。記録装置222は後にテクスチャ画像221を再び用いるために必要であり、テクスチャ画像221が外部の蓄積媒体に記録されていた画像である場合は、記録装置222を別に用意する必要はなく、その蓄積媒体が記録装置222として用いられる。この際は、記録装置222にテクスチャ画像221を入力しなおす必要はない。テクスチャ画像は、例えば、各画素の輝度(Y)を0~255の値で表した画素をラスタ順序(画像の左上の画素から右方向へ、上のラインから下のラインへの順序)で並べて形成され、一般に画像信号と呼ばれている。後に述べるシェイプ画像と区別するために、ここではテクスチャ画像と呼ぶことにする。テクスチャ画像としては、輝度以外にも、色差(U, Vなど)、色(R, G, Bなど)が用いられても良い。

【0224】

一方、最初のフレームにおいて、操作者が抽出したい物体を別途抽出しておいたシェイプ画像225が、動き補償による物体抽出回路224に入力される。シェイプ画像は、例えば、物体に属する画素の画素値を“255”、それ以外の画素の画素値を“0”で表した画素をテクスチャ画像と同様にラスタ順序で並べて生成される。

10

20

30

40

50

【0225】

ここで、最初のフレームのシェイプ画像25を生成する実施例を図34などを用いて詳しく説明する。

【0226】

図34では、省略しているが、背景や前景にも図柄があり、そのうちで、家の形をした物体226を抽出したいとする。操作者は、モニタに表示された画像227に対して、物体226の輪郭をマウスやペンでなぞる。その輪郭線の内側の画素に“255”、外側の画素に“0”を代入して得た画像をシェイプ画像とする。操作者が細心の注意をはらって輪郭線を描けば、このシェイプ画像の精度は高いものになるが、ある程度精度が低い場合でも、以下の方法を用いれば、精度を上げることができる。

10

【0227】

図35には、操作者によって描かれた線228と、物体226の輪郭線229が示されている。この段階では、輪郭線229の正しい位置はもちろん抽出されていないが、線228との位置関係を表すために輪郭線229が示している。

【0228】

まず、輪郭線228を含むようにブロックが設定される。具体的には、画面をラスト順でスキャンし、輪郭線228があった時、つまり、輪郭線228のシェイプ画像において、隣接する画素値に差があった時、その画素を中心にして所定のサイズのブロックを設ける。この際、既に設定したブロックと今回のブロックが重なる場合には、今回のブロック設定は行わずに、スキャンを進めるようにすると、図36のように互いに重なりがなく、なおかつ接するようにブロックが設定できる。しかし、これだけでは、部分230、231、232がブロックに入っていない。そこで、もう一度スキャンを行い、ブロックに含まれない輪郭線があった時、やはり、その画素を中心にしてブロックが設けられる。但し、2度目のスキャンの時には、今回のブロックが既に設定したブロックと重なる部分があっても、中心とする画素が既に設定したブロックに含まれない限り、今回のブロックの設定を行う。図37において斜線で示すブロック233、234、235、236が2スキャン目で設定されたブロックである。ブロックサイズは、固定にしてもよいが、輪郭線228によって囲まれる画素数が多い場合には大きく、その画素数が少ない場合には小さく、輪郭線228の凸凹が少ない場合には大きく、凸凹が多い場合には小さく、あるいは、画像の図柄が平坦な場合には大きく、図柄が細かい場合には小さく、設定してもよい。

20

30

【0229】

画面の端では、普通にブロックを設定すると画面からはみだしてしまうことがある。そういう場合は、そのブロックだけ、画面からはみ出さないようにブロックの端を切って長方形のブロックにする。この場合は相似ブロックも長方形とする。

【0230】

以上がシェイプ画像におけるブロックの設定方法である。

【0231】

次に、ブロック毎に、その相似なブロックをテクスチャ画像を用いて探索する。ここで、相似とは、ブロックサイズが異なるブロック同士で、一方のブロックサイズを他方と同じになるように、拡大あるいは縮小した時に、対応する画素の画素値がほぼ等しくなることをいう。例えば、図38のブロック237に対しては、ブロック238が、テクスチャ画像の図柄が相似になる。同様に、ブロック239に対しては、ブロック240が、ブロック241に対してはブロック242が、相似である。本実施形態では、相似ブロックは、輪郭線上に設定したブロックよりも大きくする。また、相似ブロックは、画面全体を探索するのではなく、例えば、図39に示す様に、ブロック243の近くのブロック244、245、246、247を四隅とするある一定の範囲内で探索すれば十分である。図39は各ブロックの中心を起点におき、ブロック243の起点を用いて、ブロック244、245、246、247の起点を所定の画素幅だけ、上下方向と左右方向に動かした場合である。起点をブロックの左上角においた場合を図40に示す。

40

【0232】

50

探索範囲内でも、一部が画面からはみ出す相似ブロックは、探索の対象から外すのであるが、ブロックが画面の端にあると、探索範囲にある全ての相似ブロックが探索の対象から外れてしまうことがある。そういう場合には、画面の端のブロックについては、探索範囲を画面の内側にずらして対応する。

【0233】

相似ブロックの探索は、多段階探索を行うと、演算量を少なくできる。多段階探索とは、例えば1画素や半画素ずつ起点をずらしながら、探索範囲全体を探索するのではなく、初めに、とびとびの位置の起点で誤差を調べる。次に、その中で誤差が小さかった起点の周囲だけを少し細かく起点を動かして誤差を調べるということを繰り返しながら、相似ブロックの位置をしばらくこんでいく方法である。

10

【0234】

相似ブロックの探索において、相似ブロックの縮小処理を毎回行くと、処理時間が多く必要である。そこで、予め画像全体を縮小したものを生成し、別のメモリに保持しておけば、相似ブロックに対応する部分のデータをそのメモリから読み出すだけで済む。

【0235】

図38では3つのブロック237, 239, 241についてだけ、相似ブロックを示しているが、実際には、図37で示した全てのブロックに対して相似ブロックを求める。以上が相似ブロックの探索方法である。相似ブロックの探索はシェイプ画像ではなく、テクスチャ画像を用いることが肝要である。画面内で相似ブロックをブロックに写像する一次変換を考えた時に、テクスチャ画像の輪郭線は、この一次変換において不変である。

20

【0236】

次に、各ブロックとその相似ブロックの位置関係を用いて、シェイプ画像の輪郭がテクスチャ画像の輪郭に合うように補正する方法を説明する。

【0237】

図41において、輪郭線228が操作者によって描かれた線である。この線が、正しい輪郭線229に近づけばよい。そのために、シェイプ画像の相似ブロック238の部分を読み出し、それをブロック237と同じサイズに縮小したもので、シェイプ画像のブロック237の部分置き換える。この操作には、輪郭線を、相似ブロックからブロックへの一次変換の不動点を含む不変集合に近づける性質があるので、輪郭線228は輪郭線229に近づく。相似ブロックの一辺がブロックの一辺の2倍の長さの時、1回の置き換えで、輪郭線228と正しい輪郭線229の隔たりは概して、 $1/2$ になる。この置き換えを全てのブロックに対して1回行った結果が図42の輪郭線248である。このブロックの置き換えを繰り返せば輪郭線248は、正しい輪郭線にさらに近づき、やがて、図43に示すように、正しい輪郭線に一致する。実際には、2本の輪郭線のずれが画素間距離よりも小さい状態は意味がないので適当な回数で置き換えを終了する。本手法は、シェイプ画像で設定した $N \times N$ 画素のブロックにテクスチャ画像での輪郭線が含まれる時に有効なのであるが、その場合、シェイプ画像の輪郭線とテクスチャ画像の輪郭線の距離は最大でおよそ $N/2$ である。相似ブロックの一辺の長さが、ブロックの一辺の長さの A 倍とした時、1回の置き換えにつき、2本の輪郭線の距離は $1/A$ になるのであるから、この距離が1画素よりも短くなることを式で表すと、置き換え回数を x として、

30

40

$$(N/2) \times (1/A)^x < 1$$

となる。ここで $\wedge$ はべき乗を表し、上式では $(1/A)$ を x 回乗するという意味である。上式から、

$$x > 1 \log(2/N) / \log(1/A)$$

となる。例えば $N=8$, $A=2$ の時は、

$$x > 2$$

であり、置き換え回数は3回で十分である。

【0238】

この物体抽出装置のブロック図を図30に示す。まず、操作者によって入力されるシェイプ画像249が、シェイプメモリ250に記録される。シェイプメモリ250において

50

は、図 3 6, 3 7 を用いて前述した様にブロックが設定される。一方、テクスチャ画像 2 5 1 は、テクスチャメモリ 2 5 2 に記録される。テクスチャメモリ 2 5 2 からは、シェイプメモリ 2 5 0 から送られる、ブロックの位置情報 2 5 3 を参照して、ブロックのテクスチャ画像 2 5 4 が探索回路 2 5 5 に送られる。同時に、図 3 9 や図 4 0 を用いて説明した様に相似ブロックの候補もテクスチャメモリ 2 5 2 から、探索回路 2 5 5 に送られる。探索回路 2 5 5 では、相似ブロックの各候補を縮小した後に、ブロックとの誤差を計算し、その誤差が最小となったものを相似ブロックとして決定する。誤差としては輝度値の差分の絶対値和や、それに色差の差分の絶対値和を加えたものなどが考えられる。輝度だけに比べて、色差も用いると、演算量は多くなるが、物体の輪郭において輝度の段差が小さくても、色差の段差が大きい場合に正しく相似ブロックを決定できるので精度が向上する。相似ブロックの位置の情報 2 5 6 は縮小変換回路 2 5 7 に送られる。縮小変換回路 2 5 7 には、シェイプメモリ 2 5 0 から、相似ブロックのシェイプ画像 2 5 8 も送られる。縮小変換回路 2 5 7 では、相似ブロックのシェイプ画像が縮小され、その縮小された相似ブロックは輪郭線が補正されたシェイプ画像 2 5 9 として、シェイプメモリ 2 5 0 に返され、対応するブロックのシェイプ画像が上書きされる。このシェイプメモリ 2 5 0 の置き換えが所定の回数に達した時は、補正されたシェイプ画像 2 5 9 は外部に出力される。シェイプメモリ 2 5 0 の書き換えは、ブロック毎に逐次上書きしても良いし、メモリを 2 画面分用意して、一方から他方へ、初めに画面全体のシェイプ画像をコピーした後に、輪郭部分のブロックは相似ブロックを縮小したもので置き換えるようにしても良い。

10

20

30

40

50

【0 2 3 9】

この物体抽出方法を図 4 8 のフローチャートを参照して説明する。

【0 2 4 0】

(フレーム内の縮小ブロックマッチングによる物体抽出方法)

ステップ S 3 1 では、シェイプデータの輪郭部分にブロックが設定される。ステップ S 3 2 では、現処理ブロックと画像データの図柄が相似である相似ブロックが同じ画像データから見つけられる。ステップ S 3 3 では、現処理ブロックのシェイプデータを相似ブロックのシェイプデータが縮小したデータで置き換えられる。

【0 2 4 1】

ステップ S 3 4 で処理済みブロック数が所定の数に達したらステップ 3 5 に進む、そうでない場合は次のブロックに処理対象を進めてステップ 3 2 に戻る。

【0 2 4 2】

ステップ S 3 5 では置き換えの繰り返し回数が所定の数に達したらステップ 3 6 に進む、そうでない場合は、置き換えられたシェイプデータを処理対象としてステップ 3 1 に戻る。ステップ S 3 6 では、置き換えを繰り返されたシェイプデータが物体領域として出力される。

【0 2 4 3】

この方法はブロックのエッジと相似ブロックのエッジが合った場合に効果がある。従って、ブロックに複数のエッジがある場合には、エッジが正しく合わないことがあるので、そういうブロックについては置き換えをせずに入力されたままのエッジを保持する。具体的には、ブロックのシェイプ画像を左右方向と上下方向に各ラインをスキャンし、1 つのラインで“0”から“2 5 5”へ、あるいは“2 5 5”から“0”に変化する点が 2 つ以上あるラインが所定の数以上あるブロックは置き換えをしない。また、物体と背景の境界であっても部分によっては、輝度などが平坦な場合がある。このような場合もエッジ補正の効果が期待できないのでテクスチャ画像の分散が所定値以下のブロックについても置き換えをせずに入力されたままのエッジを保持する。

【0 2 4 4】

相似ブロックの誤差が所定値よりも小さくならない場合、縮小をあきらめ、同じサイズで相似ブロックを求めても良い。この際、自分とはなるべく重ならないように相似ブロックを選ぶ。縮小を行わないブロックだけでは、エッジが補正される効果はないが、縮小を行うことによってエッジが補正されたブロックから、その補正されたエッジをコピーする

ことで、縮小を行わないブロックについても、間接的にエッジが補正される。

【0245】

図48に示したフローチャートは相似ブロックを見つけた直後にシェイプ画像の置き換えを行う例であったが、全ブロックの相似ブロックの位置情報を保持するようにすることで、初めに、相似ブロックの探索を全ブロックについて行い、次に、シェイプ画像の置き換えを全ブロックについて行う方法を図50のフローチャートを参照して説明する。

【0246】

この例では、1回の相似ブロックの探索に対してシェイプ画像の置き換えが複数回反復できる。

【0247】

ステップS41では、シェイプデータの輪郭部分にブロックが設定される。ステップS42では、現処理ブロックと画像データの図柄が相似である相似ブロックが同じ画像データ内から見つけられる。ステップS43では、全てのブロックについて相似ブロックを見つける処理が終わったとき、つまり、処理済みブロック数が所定の数に達したときにはステップS44に進む。そうでない場合はステップS42に戻る。ステップS44では、現処理ブロックのシェイプデータを相似ブロックのシェイプデータを縮小したもので置き換える。

【0248】

ステップS45では、全てのブロックについて置き換える処理が終わったとき、つまり、処理済みブロック数が所定の数に達したときにはステップS46に進む。そうでない場合はステップS44に戻る。ステップS46では全ブロックの置き換え回数が所定の回数に達した場合はS47に進む。そうでない場合はS44に戻る。ステップS47では、置き換え変換を繰り返されたシェイプデータが物体領域として出力される。

【0249】

次にエッジ補正の精度を上げることができるブロックの設定方法を説明する。

【0250】

前述したようにシェイプ画像の輪郭線の周囲にブロックを設定する方法では、図51(a)に示されるように、正しい輪郭線301の一部がブロックに含まなくなることがある。ここで、シェイプ画像の輪郭線302は太い線で示してある。仮に輪郭線の右下側が物体で左上側が背景だとすると、本当は背景である部分303は物体と誤って設定されているにもかかわらず、ブロックに含まれないために修正される可能性がない。このようにブロックと正しい輪郭線の間に隙間があると、正しく補正されない。

【0251】

ブロックと正しい輪郭線の隙間を小さくするには、図51(b)に示したようにブロックをある程度重なり合わせる方法がある。こうすると、ブロックの数が増えるので、演算量は増加するが隙間304は小さくなる。従って抽出の精度は向上する。しかし、この例では、まだ隙間は完全にはなくなる。

【0252】

隙間を小さくするには、図51(c)に示したようにブロックサイズを大きくすることも有効である。この例では、前述したブロックの重ねあわせを併用した。これにより、この例では隙間が完全になくなる。

【0253】

このように、輪郭線の補正可能な範囲を広げるにはブロックサイズを大きくすることが有効である。しかし、ブロックサイズが大きすぎると、ブロックに含まれる輪郭線の形状が複雑になり、相似ブロックが見つかりにくくなる。その例が図52に示されている。

【0254】

図52(a)では、斜線の部分305が物体領域、白色の部分306が背景領域を表す。与えられたシェイプ画像の輪郭線307は黒線で示されている。このように、シェイプ画像の輪郭線307は正しい輪郭線と大きく隔たっており、また、正しい輪郭線には凹凸がある。これに対して、前に説明した方法とは異なる方法でブロックを配置した結果が図

10

20

30

40

50

5 2 (b) に示されている。ここでは、まず、互いに重ならず、かつ、隙間がないような矩形ブロックで画像が分割される。ブロック毎にテクスチャ画像での分散が計算され、分散が所定値よりも小さいブロックはその設定を解消した。従って図 5 2 (b) では、分散が所定値以上のブロックだけが残っている。これらのブロック毎に相似ブロックを求めるのであるが、例えばブロック 3 0 8 の近くにこれを縦横 2 倍にした図柄は存在しないし、他の多くのブロックについても同様である。従って、誤差最小な部分を相似ブロックとして選択はするものの、その位置関係を用いてシェイプ画像の置き換え変換を反復しても図 5 2 (c) に示す通り正しい輪郭線には合致しない。ただ、図 5 2 (a) のシェイプ画像の輪郭線 3 0 7 と比較して、エッジ補正後の図 5 2 (c) のシェイプ画像の輪郭線 3 0 9 は、テクスチャ画像の輪郭線の大まかな凹凸（左と右に山があってその間に谷があるという程度のもの）は反映されている。この例で仮にブロックサイズを小さくすると、この大まかな補正さえされなくなってしまう。

【0 2 5 5】

このように、補正の範囲を広げるためにブロックサイズを大きくすると、ブロックに含まれる輪郭線の形状が複雑になり、相似ブロックが見つかりにくくなることもある。その結果エッジの補正が大まかにしかされなくなる。このような場合には、ブロックのサイズを初めは大きなサイズでエッジ補正を行い、その結果に対して、再度ブロックサイズを小さくしてエッジ補正を行うと、補正の精度が向上する。図 5 2 (c) に対して、ブロックサイズを縦横 1 / 2 にして補正を再度行い、さらに 1 / 4 にして補正を行くた結果が図 5 2 (d) に示される。このように、ブロックサイズを次第に小さくしながら補正を繰り返せば補正の精度を向上できる。

【0 2 5 6】

ブロックサイズを次第に小さくする方法を図 5 3 のフローチャートを参照して説明する。

【0 2 5 7】

ステップ S 5 1 ではブロックサイズ $b = A$ と設定する。ステップ S 5 2 では、図 4 8 または図 5 0 に示したエッジ補正と同様なエッジ補正を行う。ステップ S 5 3 では、 b を観察し、 b が Z ($< A$) より小さくなると、この処理は終了する。 b が Z 以上の場合にはステップ S 5 4 に進む。ステップ S 5 4 でブロックサイズ b を半分にして S 5 2 に進む。

【0 2 5 8】

以上、ブロックサイズを初めは大きめにし、次第に小さくしながら補正を繰り返すことで補正の精度を向上する例を示した。

【0 2 5 9】

図 5 4 (a) に、ブロックを 4 5 度傾けることで、ブロックと正しい輪郭線の間に隙間をできにくくする例が示されている。このように、輪郭線が斜めの場合にはブロックサイズを図 5 1 (c) ほど大きくしなくても、ブロックを傾ければ正しい輪郭線を覆うことができる。また、この例では、ブロックの重なりを無くしても図 5 4 (b) のように正しい輪郭線を覆える。このように、シェイプ画像の輪郭線と同じ向きにブロックの辺を傾けることで、ブロックと正しい輪郭線の間に隙間を生じにくくすることができる。具体的には、アルファ画像の輪郭線の傾きを検知しそれが水平か垂直に近い場合にはブロックの向きは図 5 1 (c) のようにし、そうでない場合にはブロックの傾きは図 5 4 (b) のようにする。水平や垂直に近いという判断はしきい値との比較で行う。

【0 2 6 0】

以上が、最初のフレームの物体抽出処理である。これは、必ずしも動画像の最初のフレームだけではなく、静止画像一般に用いることができる手法である。なお、置き換えを 1 回行ったシェイプ画像に対して、ブロックを設定しなおし、その相似ブロックを求めなおして、2 回目の置き換えを行うというように、置き換えの度にブロック設定と相似ブロックの探索を行えば、演算量は増えるが、より補正の効果が得られる。

【0 2 6 1】

また、相似ブロックはブロックのなるべく近い部分から選ばれるのが好ましいので、相

似ブロックを探す範囲をブロックサイズによって切り換えるとよい。即ち、ブロックサイズが大きい場合には、相似ブロックを探す範囲を広くし、ブロックサイズが小さい場合には、相似ブロックを探す範囲を狭くする。

【0262】

また、本手法では、シェイプデータの置き換えの過程で、シェイプデータに小さい穴や、孤立した小領域が誤差として出現することがある。そこで、ステップS34、S35、S36、S45、S46、S47、S53の前などで、シェイプデータから、小さい穴や、孤立した小領域を除くことにより、補正の精度を向上することができる。小さい穴や、孤立した小領域を除くには、例えば、画像解析ハンドブック（高木、下田監修、東京大学出版会、初版1991年1月）575～576頁に記載されている膨張と収縮を組み合わせた処理や、677頁に記載された多数決フィルタなどを用いる。 10

【0263】

また、ブロックは図49に示すように、より簡易的に設定しても良い。すなわち、画面を単純にブロック分割し、そのうちブロック2200など、輪郭線228を含むブロックについてのみ、相似ブロックの探索や置き換えの処理を行う。

【0264】

また、与えられるテクスチャ画像が予めフラクタル符号化（特公平08-329255号公報「画像の領域分割方法及び装置」）によって圧縮されているのであれば、その圧縮データに、各ブロックの相似ブロックの情報が含まれている。従って、輪郭線228を含むブロックの相似ブロックとしては、圧縮データを流用すれば、改めて相似ブロックを探索する必要はない。 20

【0265】

図27に戻り、動画から物体を抽出する物体抽出装置の説明を続ける。

【0266】

動き補償による物体抽出回路242では、テクスチャ画像221から検出される動きベクトルを用いながら、最初のフレームのシェイプ画像25を元にして、2フレーム目以降のフレームのシェイプ画像260を生成する。

【0267】

図29に動き補償による物体抽出回路224の例が示される。最初のフレームのシェイプ画像225が、シェイプメモリ261に記録される。シェイプメモリ261においては、図45のフレーム262に示した様に、画面全体にブロックが設定される。一方、テクスチャ画像221は、動き推定回路264に送られ、また、テクスチャメモリ263に記録される。テクスチャメモリ263からは、1フレーム前のテクスチャ画像265が動き推定回路264に送られる。動き推定回路264では、現処理フレームのブロック毎に、1フレーム前のフレーム内から誤差が最小となる参照ブロックを見つける。図45に、ブロック267と、1フレーム前のフレーム266から選ばれた参照ブロック268の例が示される。ここで、誤差が所定のしきい値よりも小さくなるのであれば、参照ブロックはブロックよりも大きくする。ブロック269と縦横2倍の大きさの参照ブロック70の例も図45に示す。 30

【0268】

図29に戻り、参照ブロックの位置の情報271は動き補償回路272に送られる。動き補償回路272には、シェイプメモリ261から、参照ブロックのシェイプ画像273も送られる。動き補償回路272では、参照ブロックの大きさがブロックと同じ場合は、そのまま、参照ブロックの大きさがブロックよりも大きい場合は、参照ブロックのシェイプ画像が縮小され、その参照ブロックのシェイプ画像は現処理フレームのシェイプ画像260として出力される。また、次のフレームに備えて、現処理フレームのシェイプ画像260はシェイプメモリ261に送られて、画面全体のシェイプ画像が上書きされる。 40

【0269】

参照ブロックがブロックよりも大きい場合、先に図41、42を用いて説明したのと同様に、輪郭線が正しい位置からずれていた場合に補正する効果がある。従って、与えられ 50

る最初のフレームのシェイプ画像に続く、動画シーケンスの全てのフレームにおいて、物体が高い精度で抽出される。従来手法のように、動画シーケンスの最初の方や、物体の動きが小さい時に、精度が悪いという不具合はない。

【0270】

フレーム間の動き補償による物体抽出を図47のフローチャートを参照して説明する。

【0271】

ステップS21で現処理フレームがブロックに分割される。ステップS22では現処理ブロックと画像データの図柄が相似であり、かつ、現処理ブロックよりも大きい参照ブロックを各フレームあるいは、既にシェイプデータを求めたフレーム内から見つける。ステップS23では参照ブロックのシェイプデータを切り出して縮小したサブブロックを現処理ブロックに貼り付ける。 10

【0272】

ステップS24では処理済みブロック数が所定の数に達したらステップ25に進む、そうでない場合は次のブロックに処理対象を進めてステップ22に戻る。ステップS25では貼り合わされたシェイプデータを物体領域として出力する。

【0273】

ここで、各フレームとは、本実施例では最初のフレームであり、予めシェイプ画像が与えられるフレームのことである。また、参照ブロックは必ずしも1フレーム前のフレームでなくても、ここで述べたように、既にシェイプ画像が求まっているフレームならよい。

【0274】

以上が動き補償を用いた物体抽出の説明である。物体抽出回路224としては、以上で説明した方法の他に、先に出願した、特開平10-001847「動画像の物体追跡／抽出装置」にあるフレーム間差分画像を用いる方法などもある。 20

【0275】

図27に戻り、動画から物体を抽出する物体抽出装置の実施例の説明を続ける。

【0276】

シェイプ画像260はスイッチ部223とスイッチ部281に送られる。スイッチ部223では、シェイプ画像260が“0”（背景）の時には、テクスチャ画像221が、背景メモリ274に送られ、記録される。シェイプ画像260が“255”（物体）の時には、テクスチャ画像221は、背景メモリ274には送られない。これをいくつかのフレームに対して行い、そのシェイプ画像260がある程度正確であれば、物体を含まない、背景部分だけの画像が、背景メモリ274に生成される。 30

【0277】

次に、記録装置222から、テクスチャ画像275が再度最初のフレームから順に読み出され、あるいは、操作者が指定する物体を抽出したいフレームだけが読み出され差分回路276に入力される。同時に、背景メモリ274から、背景画像277が読み出され、差分回路276に入力される。差分回路276では、テクスチャ画像275と背景画像277の、互いに画面内で同じ位置にある画素同士の差分値278が求められ、背景画像を用いた物体抽出回路279に入力される。物体抽出回路279では、シェイプ画像280が生成されるのであるが、これは、差分値278の絶対値が予め定めるしきい値よりも大きい画素は、物体に属するとして画素値を“255”とし、そうでない画素は、背景に属するとして画素値“0”とすることで生成される。テクスチャ画像として、輝度だけでなく、色差や色も用いる場合は、各信号の差分の絶対値の和をしきい値と比較して物体か背景かが決定される。あるいは、輝度や色差毎に別々にしきい値を定めて、輝度、色差のいずれかにおいて、差分の絶対値がそのしきい値よりも大きい場合に物体、そうでない場合に背景とが判定される。このようにして生成されたシェイプ画像280がスイッチ部281に送られる。また、操作者によって決定される選択信号282が外部からスイッチ部281に入力され、この選択信号282によって、シェイプ画像260とシェイプ画像280のうちのいずれかが選択され、シェイプ画像283として外部に出力される。操作者は、シェイプ画像260とシェイプ画像280を各々、ディスプレイなどに表示し、正確な 40 50

方を選択する。あるいは、シェイプ画像 260 が生成された段階でそれを表示し、その精度が満足するものでなかった場合に、シェイプ画像 280 を生成し、シェイプ画像 260 の精度が満足するものであった場合には、シェイプ画像 280 を生成せずに、シェイプ画像 260 をシェイプ画像 283 として外部に出力するようにすれば、処理時間を節約できる。選択は、フレーム毎に行ってもよいし、動画像シーケンス毎に行ってもよい。

【0278】

図 27 の物体抽出装置に対応する物体抽出方法を図 46 のフローチャートを参照して説明する。

【0279】

(背景画像を用いる物体抽出方法)

ステップ S11 では与えられる各フレームにおけるシェイプデータを動き補償することにより各フレームのシェイプデータが生成される。ステップ S12 ではシェイプデータによって決定される背景領域の画像データが背景画像としてメモリに記憶される。

【0280】

ステップ S13 では処理済みフレーム数が所定の数に達したらステップ 14 に進む、そうでない場合は次のフレームに処理対象を進めてステップ 11 に戻る。ステップ S14 では画像データと背景画像との差分の絶対値が大きい画素を物体領域とし、そうでない画素を背景領域とする。

【0281】

本実施形態においては、例えば撮像するカメラに動きがあると背景が動く。この場合は、前のフレームからの背景全体の動き（グローバル動きベクトル）を検出し、1 スキャン目ではグローバル動きベクトルの分だけ前のフレームからずらして背景メモリに記録し、2 スキャン目では、グローバル動きベクトルの分だけ前のフレームからずらした部分を背景メモリから読み出す。1 スキャン目で検出したグローバル動きベクトルをメモリに記録しておき、2 スキャン目では、それを読み出して用いれば、グローバル動きベクトルを求める時間を節約できる。また、カメラが固定していることなどから、背景が静止していることが既知の場合は、操作者がスイッチを切り替えることなどによって、グローバル動きベクトルの検出を行わないようにして、グローバル動きベクトルは常にゼロにするようにすれば、処理時間はさらに節約できる。グローバル動きベクトルを半画素精度で求める時は、背景メモリは、入力される画像の縦横とも 2 倍の画素密度とする。すなわち、入力画像の画素値は 1 画素おきに背景メモリに書き込まれる。例えば次のフレームでは背景が横方向に 0.5 画素動いていた場合には、先に書き込まれた画素の間にやはり 1 画素おきに画素値が書き込まれる。このようにすると、1 スキャン目が終了した時点で、背景画像に一度も書き込まれない画素ができることがある。その場合は、周囲の書き込まれた画素から内挿してその隙間を埋める。

【0282】

また、半画素の動きベクトルを用いる／用いないに関わらず、動画シーケンス全体を通じて、一度も背景領域にならない部分は、1 スキャン目を終わっても背景メモリに画素値が代入されない。このような未定義の部分は、2 スキャン目では、常に物体と判定する。これは、特に未定義の部分を記録するためのメモリを用意して、未定義か否かをいちいち判定しなくても、背景に希にしか出てこないと予想される画素値 (Y, U, V) = (0, 0, 0) などで、予め背景メモリを初期化してから 1 スキャン目を開始すればよい。未定義の画素にはこの初期画素値が残るので 2 スキャン目では、自動的に物体と判定される。

【0283】

これまでの説明では、背景メモリを生成する時に、既に背景の画素値が代入されている画素についても、背景領域であれば、画素値が上書きされている。この場合、動画シーケンスの最初の方でも最後の方にでも背景である部分には、動画シーケンスの最後の方の背景の画素値が背景メモリに記録される。動画シーケンスの最初と最後でそういった背景が全く同じ画素値ならば問題はないが、カメラが非常にゆっくりと動いたり、背景の明るさが少しずつ変化するなどして、画素値がフレーム間で微少に変動する場合には、動画シー

10

20

30

40

50

ケンスの最初の方の背景と最後の方の背景とでは、画素値の差が大きくなるので、この背景メモリを用いると、動画シーケンスの最初の方のフレームで背景部分も物体と誤検出されてしまう。そこで、その前までのフレームでは、一度も背景領域にはならず、現処理フレームで初めて背景領域となった画素についてのみ背景メモリへの書き込みを行い、既に背景の画素値が代入されている画素の上書きはしないようにすれば、背景メモリには動画シーケンスの最初の方の背景が記録されるので、正しく物体が抽出される。そして、2スキャン目にも、その物体抽出結果に応じて、現処理フレームの背景領域を背景メモリに上書きするようにすれば、現処理フレームの直前のフレームの背景と現処理フレームの背景という相関の高い背景同士を比較することになり、その部分が物体と誤検出されにくくなる。2スキャン目の上書きは、背景に微少な変動がある場合に有効なので、操作者が背景の動きは無しという意味にスイッチを切り替えるなどした場合は、上書きは行わない。このスイッチは、先のグローバル動きベクトルを行うか行わないかを切り替えるスイッチと共通でも構わない。

【0284】

1スキャン目は背景画像を生成するのが目的であるから、必ずしも全てのフレームを用いる必要はない。1フレームおき、2フレームおきなどとフレームを間引いても、ほぼ同じ背景画像が得られ、処理時間は短くなる。

【0285】

背景領域のうち、フレーム間差分がしきい値以下の画素だけを背景メモリに記録するようにすれば、画面に入り込んでくる他の物体が背景メモリに記録されずに済む。また、1スキャン目の物体領域が実際よりも物体側に誤検出された場合、物体の画素値が背景メモリに記録されてしまう。そこで、背景領域でも物体領域に近い画素は背景メモリに入力しないようにする。

【0286】

観光地で撮影した画像などで、前景の人などを除いた背景画像だけが必要な場合は、背景メモリに記録された背景画像を外部に出力する。

【0287】

以上が本実施形態の第1の構成例の説明である。本例によれば、動画シーケンスの初めの部分も最後の部分と同様に高い抽出精度が得られる。また、物体の動きが小さかったり、全く動かない場合でも正しく抽出される。

【0288】

次に、図28を用いて、生成されたシェイプ画像280を修正する例を説明する。シェイプ画像280が生成されるまでは、図27と同じなので説明を省略する。

【0289】

シェイプ画像280は、背景パレットを用いるエッジ補正回路284に入力される。また、テクスチャ画像275が、背景パレットによるエッジ補正回路284と縮小ブロックマッチングによるエッジ補正回路285に入力される。エッジ補正回路284の詳細なブロック図を図31に示す。

【0290】

図31において、シェイプ画像280は、補正回路286に入力され、同じフレームのテクスチャ画像275は比較回路287に入力される。背景パレットを保持するメモリ288からは、背景色289が読み出され、比較回路287に入力される。ここで、背景パレットは、背景部分に存在する輝度(Y)と色差(U, V)の組すなわちベクトルの集まり、

(Y1, U1, V1)
 (Y2, U2, V2)
 (Y3, U3, V3)

のことで、予め用意される。具体的には、背景パレットは、最初のフレームにおいて、背景領域に属する画素のY, U, Vの組を集めたものである。ここで、例えばY, U, Vが

各々 2 5 6 通りの値をとるとすると、その組み合わせは膨大な数になり、背景の (Y, U, V) の組み合わせ数も多くなり、後に説明する処理の演算量が多くなってしまいうので、所定のステップサイズで Y, U, V の値を各々量子化することにより、組み合わせ数を抑制できる。これは、量子化をしない場合は異なるベクトル値だったもの同士が、量子化により同じベクトル値になる場合があるからである。

【0291】

比較回路 287 では、テクスチャ画像 275 の各画素の Y, U, V が量子化され、そのベクトルがメモリ 288 から順次送られてくる背景パレットに登録されているベクトル、すなわち背景色 289 のいずれかと一致するかどうか調べられる。画素毎に、その画素の色が背景色かどうかの比較結果 290 が、比較回路 287 から補正回路 286 へ送られる。補正回路 286 では、シェイプ画像 280 のある画素の画素値が “255” (物体) であるにもかかわらず、比較結果 290 が背景色であった場合に、その画素の画素値を “0” (背景) に置き換えて、補正されたシェイプ画像 291 として出力する。この処理により、シェイプ画像 280 において物体領域が背景領域にはみ出して誤抽出されていた場合に、その背景領域を正しく分離できる。ただ、背景と物体に共通の色があり、背景パレットに物体の色も混じって登録されていると、物体のその色の部分もが背景と判定されてしまう。そこで、最初のフレームでは、先に説明したパレットを背景の仮のパレットとしておき、同様の方法で最初のフレームの物体のパレットも作る。次に、背景の仮のパレットの中で物体のパレットにも含まれる色については、背景の仮のパレットから除き、残ったものを背景パレットとする。これにより、物体の一部が背景になってしまう不具合を回避できる。

【0292】

また、最初のフレームで与えられるシェイプ画像に誤差がある場合を考慮し、シェイプ画像のエッジの近傍の画素は、パレットの生成に用いないようにしても良い。また、各ベクトルの出現頻度を数え、頻度が所定値以下のベクトルはパレットに登録しないようにしても良い。量子化ステップサイズを小さくしすぎると、処理時間が多くなったり、背景色に非常に似た色でもベクトル値がわずかに異なるために背景と判定されなかったりし、逆に量子化ステップサイズを大きくしすぎると、背景と物体に共通するベクトルばかりになってしまう。そこで、最初のフレームに対して、いくつかの量子化ステップサイズを試し、与えられるシェイプ画像の様に背景色と物体色が分離される量子化ステップサイズが選ばれる。

【0293】

また、途中から新しい色が背景や物体に現れることがあるので、途中のフレームで背景パレットを作りなおしても良い。

【0294】

図 28 に戻り、シェイプ画像 291 は、エッジ補正回路 285 に入力される。エッジ補正回路 285 は、先に説明した図 30 の回路において、シェイプ画像 249 をシェイプ画像 291、テクスチャ画像 251 をテクスチャ画像 275 とする回路と同じであるので、説明は省略するが、シェイプ画像のエッジがテクスチャ画像のエッジに合うようにシェイプ画像の補正を行う。補正されたシェイプ画像 292 はスイッチ部 281 に送られる。スイッチ部 281 からは、シェイプ画像 292 とシェイプ画像 260 のうちから選択されたシェイプ画像 293 が出力される。

【0295】

本例では、エッジ補正回路を物体抽出回路 279 の後段に設けたが、物体抽出回路 224 の後段に設ければ、シェイプ画像 260 の精度を向上できる。

【0296】

また、エッジ補正によって、抽出精度がかえって悪化する場合も希にある。そういう時に、悪化したシェイプ画像 292 が出力されてしまわないように、図 28 において、シェイプ画像 280 やシェイプ画像 291 もスイッチ部 281 に入力すれば、エッジ補正を行わないシェイプ画像 280 や、背景パレットによるエッジ補正だけを施したシェイプ画像

291を選択することも可能となる。

【0297】

図44は、背景パレットに登録された背景色の画素をクロスハッチで示しており、先に図30や図29を用いて説明した相似ブロックの探索の時に、図44の情報を用いると、輪郭抽出の精度をさらに高めることができる。背景に図柄がある場合に、物体と背景のエッジではなく、背景の図柄のエッジに沿うように相似ブロックが選ばれてしまうことがある。このような場合、ブロックと相似ブロックを縮小したブロックとの誤差を求める時に、対応画素がいずれも背景色同士の時は、その画素の誤差は、計算に含めないようにすると、背景の図柄のエッジがずれていても、誤差が発生せず、従って、物体と背景のエッジが合うように相似ブロックが正しく選択される。

10

【0298】

図32は、本実施形態の物体抽出装置294を組み込んだ画像合成装置の例である。テクスチャ画像295はスイッチ部296と物体抽出装置294に入力され、最初のフレームのシェイプ画像2100は物体抽出装置294に入力される。物体抽出装置294は、図27や図28で構成されており、各フレームのシェイプ画像297が生成され、スイッチ部296に送られる。一方、記録回路298には、予め合成用背景画像299が保持されており、現処理フレームの背景画像299が記録回路298から読み出され、スイッチ部296に送られる。スイッチ部296では、シェイプ画像の画素値が“255”（物体）の画素ではテクスチャ画像295が選択されて合成画像2101として出力され、シェイプ画像の画素値が“0”（背景）の画素では、背景画像299が選択されて合成画像2101として出力される。これにより、背景画像299の前景にテクスチャ画像295内の物体を合成した画像が生成される。

20

【0299】

図33はエッジ補正を行う別の例を示す。図33のように設定されたブロックのうちの 하나가、図33のブロック2102であるとする。輪郭線を境界にして、ブロックは物体領域と背景領域に分けられている。この輪郭を左右方向にずらして得られたブロックが2103, 2104, 2105, 2106である。それぞれずらす幅と向きが異なる。文献：福井「領域間の分離度に基づく物体輪郭抽出」（電子情報通信学会論文誌、D-II、Vol. J80-D-II、No. 6、pp. 1406-1414、1997年6月）の1408ページに記述されている分離度を各々の輪郭線について求め、ブロック2102～2106のうちで分離度が最も高い輪郭線を採用する。これにより、シェイプ画像の輪郭がテクスチャ画像のエッジに合う。

30

【0300】

以上述べてきたように、本第4実施形態によれば、動画シーケンスの初めの部分も最後の部分と同様に高い抽出精度が得られる。また、物体の動きが小さかったり、全く動かない場合でも正しく抽出される。さらに、現処理ブロックよりも大きい相似ブロックのシェイプデータを縮小して張り付けることにより、シェイプデータで与えられる物体領域の輪郭線がずれていてもそれを正しい位置に補正することが可能となり、物体領域の輪郭を大まかになぞったものをシェイプデータとして与えるだけで、以降の入力フレーム全てにおいて物体領域を高い精度で抽出することが可能となる。

40

【0301】

なお、以上の第1乃至第4実施形態は適宜組み合わせることもできる。また、第1乃至第4実施形態の物体抽出方法の手順はすべてソフトウェアによって実現することもでき、この場合には、その手順を実行するコンピュータプログラムを記録媒体を介して通常のコンピュータに導入するだけで、第1乃至第4実施形態と同様の効果を得ることができる。

【図面の簡単な説明】

【0302】

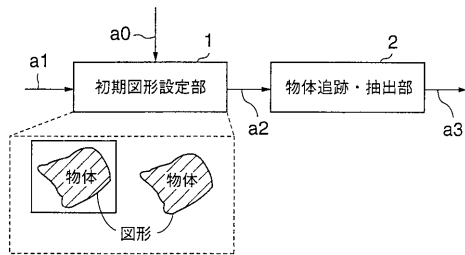
【図1】本発明の第1実施形態に係る動画像の物体追跡／抽出装置の基本構成を示すブロック図。

50

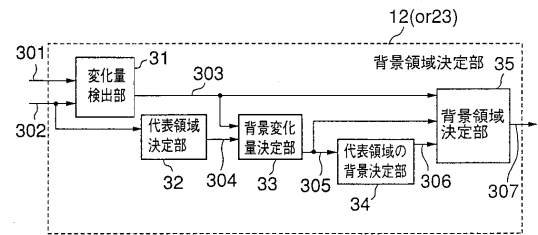
- 【図 2】同実施形態の物体追跡／抽出装置の第 1 の構成例を示すブロック図。
- 【図 3】同実施形態の物体追跡／抽出装置の第 2 の構成例を示すブロック図。
- 【図 4】同実施形態の物体追跡／抽出装置に設けられた背景領域決定部の具体的な構成の一例を示すブロック図。
- 【図 5】同実施形態の物体追跡／抽出装置に設けられた図形設定部の具体的な構成の一例を示すブロック図。
- 【図 6】同実施形態の物体追跡／抽出装置に設けられた背景動き削除部の具体的な構成の一例を示すブロック図。
- 【図 7】同実施形態の物体追跡／抽出装置に設けられた背景動き削除部で使用される代表背景領域の一例を示す図。 10
- 【図 8】同実施形態の物体追跡／抽出装置の動作を説明するための図。
- 【図 9】本発明の第 2 実施形態に係る第 1 の動画像物体追跡／抽出装置を表すブロック図。
- 【図 10】同第 2 実施形態に係る第 2 の動画像物体追跡／抽出装置を表すブロック図。
- 【図 11】同第 2 実施形態に係る第 3 の動画像の物体追跡／抽出装置を表すブロック図。
- 【図 12】同第 2 実施形態に係る第 4 の動画像の物体追跡／抽出装置を表すブロック図。
- 【図 13】同第 2 実施形態の物体追跡／抽出装置で用いられる物体予測の方法を説明するための図。
- 【図 14】同第 2 実施形態の物体追跡／抽出装置で用いられる参照フレーム選択方法を説明するための図。 20
- 【図 15】同第 2 実施形態の物体追跡／抽出装置において第 1 の物体追跡／抽出部と第 2 の物体抽出部を切り替えて物体を抽出した結果の例を表す図。
- 【図 16】同第 2 実施形態の物体追跡／抽出装置を用いた動画像の物体追跡／抽出処理の流れを説明する図。
- 【図 17】本発明の第 3 実施形態に係る第 1 の動画像物体追跡／抽出装置を表すブロック図。
- 【図 18】同第 3 実施形態に係る第 2 の動画像物体追跡／抽出装置を表すブロック図。
- 【図 19】同第 3 実施形態に係る第 3 の動画像物体追跡／抽出装置を表すブロック図。
- 【図 20】同第 3 実施形態に係る第 5 の動画像物体追跡／抽出装置を表すブロック図。
- 【図 21】同第 3 実施形態に係る第 6 の動画像物体追跡／抽出装置を表すブロック図。 30
- 【図 22】同第 3 実施形態に係る第 4 の動画像物体追跡／抽出装置を表すブロック図。
- 【図 23】同第 3 実施形態に係る動画像物体追跡／抽出装置の他の構成例を表すブロック図。
- 【図 24】同第 3 実施形態に係る動画像物体追跡／抽出装置のさらに他の構成例を表すブロック図。
- 【図 25】同第 3 実施形態に係る動画像物体追跡／抽出装置に適用されるフレーム順序制御による抽出フレーム順の例を説明する図。
- 【図 26】同第 3 実施形態に係る動画像物体追跡／抽出装置の応用例を表す図。
- 【図 27】本発明の第 4 実施形態に係る物体抽出装置を示すブロック図。
- 【図 28】同第 4 実施形態に係る物体抽出装置にエッジ補正処理を適用した場合の構成例 40
- を示すブロック図。
- 【図 29】同第 4 実施形態に係る物体抽出装置に適用される動き補償部の構成例を示すブロック図。
- 【図 30】同第 4 実施形態に係る物体抽出装置に適用される縮小ブロックマッチングによる物体抽出部の構成例を示すブロック図。
- 【図 31】同第 4 実施形態に係る物体抽出装置で使用される背景パレットによるエッジ補正回路を示す図。
- 【図 32】同第 4 実施形態に係る物体抽出装置に適用される画像合成装置を示す図。
- 【図 33】同第 4 実施形態に係る物体抽出装置で使用される分離度を用いたエッジ補正の原理を説明する図。 50

- 【図 3 4】同第 4 実施形態に係る物体抽出装置で処理される処理画像全体を示す図。
- 【図 3 5】同第 4 実施形態で用いられる操作者によって描かれた輪郭線を示す図。
- 【図 3 6】同第 4 実施形態で用いられるブロック設定（1 スキャン目）の様子を示す図。
- 【図 3 7】同第 4 実施形態で用いられるブロック設定（2 スキャン目）の様子を示す図。
- 【図 3 8】同第 4 実施形態で用いられる相似ブロックを説明するための図。
- 【図 3 9】同第 4 実施形態で用いられる相似ブロックの探索範囲を説明するための図。
- 【図 4 0】同第 4 実施形態で用いられる相似ブロックの探索範囲の別の例を説明するための図。
- 【図 4 1】同第 4 実施形態で用いられるシェイプ画像の置き換え変換前の様子を示す図。
- 【図 4 2】同第 4 実施形態で用いられるシェイプ画像の置き換え変換後の様子を示す図。 10
- 【図 4 3】同第 4 実施形態において抽出された輪郭線を示す図。
- 【図 4 4】同第 4 実施形態において抽出された背景色の部分を表す図。
- 【図 4 5】同第 4 実施形態で使用される動き補償を説明するための図。
- 【図 4 6】同第 4 実施形態で使用される背景画像を用いた物体抽出方法のフローチャート。
- 【図 4 7】同第 4 実施形態で使用される動き補償による物体抽出方法のフローチャート。
- 【図 4 8】同第 4 実施形態で使用されるフレーム内の縮小ブロックマッチングによる物体抽出方法のフローチャート。
- 【図 4 9】同第 4 実施形態で用いられるブロック設定の他の例を示す図。
- 【図 5 0】エッジ補正を説明するためのフローチャート図。 20
- 【図 5 1】ブロック設定の例を示す図。
- 【図 5 2】物体領域の輪郭線を探索する過程を示す図。
- 【図 5 3】ブロックサイズを次第に小さくする方法を説明するフローチャート図
- 【図 5 4】ブロック設定の他の例を示す図。
- 【符号の説明】
- 【0 3 0 3】
- 1 … 初期図形設定部、2 … 物体追跡・抽出部、1 1 … 図形設定部、1 2 … 背景領域決定部、1 3 … 物体抽出部、2 1 … 背景動き削除部、2 2 … 図形設定部、2 3 … 背景領域決定部、2 4 … 物体抽出部、3 1 … 変化量検出部、3 2 … 代表領域決定部、3 3 … 背景変化量決定部、3 4 … 代表領域の背景決定部、3 5 … 背景領域決定部、3 6 … 形状予測部、3 7 … 静止物体領域決定部、4 1 … 分離部、4 2 … 動き検出部、4 3 … 分割判定部、4 4 … 図形決定部、5 1 … 背景代表領域設定部、5 2 … 動き検出部、5 3 … 動き補償部、6 1 … 図形設定部、6 2 … 複数の物体追跡・抽出部、7 0 … 図形設定部、7 1 … 第 2 の物体追跡・抽出部、7 2 … 参照フレーム選択部、7 3 … 第 1 の物体追跡・抽出部、1 1 1 … 特徴量抽出部、1 4 1 … フレーム順序制御部、2 2 4 … 物体抽出部、2 7 9 … 物体抽出部、2 8 4 … エッジ補正部、2 8 5 … エッジ補正部 30

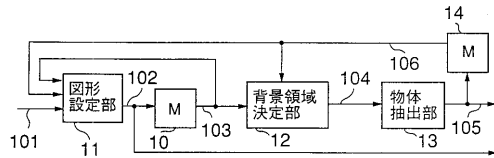
【図 1】



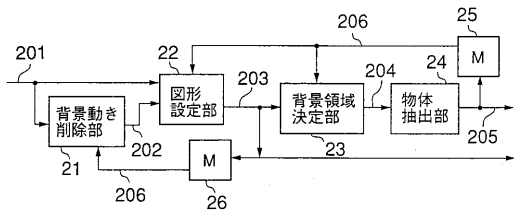
【図 4】



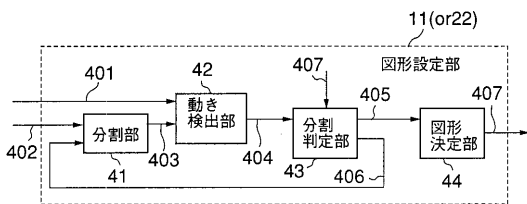
【図 2】



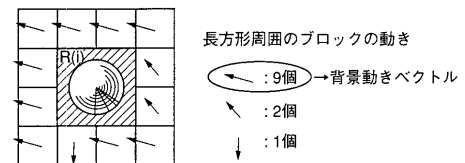
【図 3】



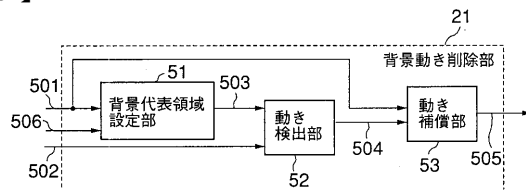
【図 5】



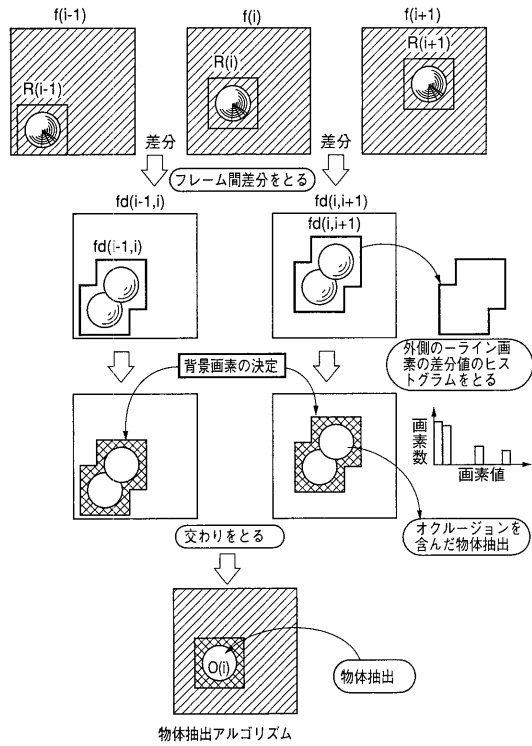
【図 7】



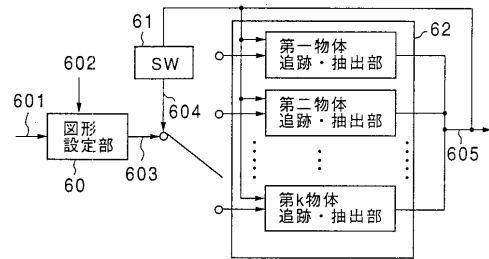
【図 6】



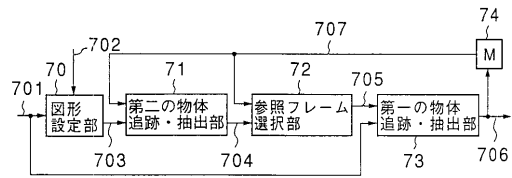
【図 8】



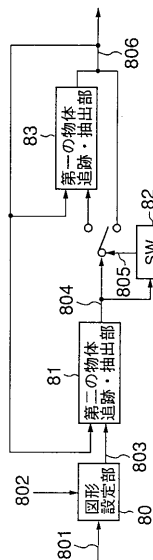
【図 9】



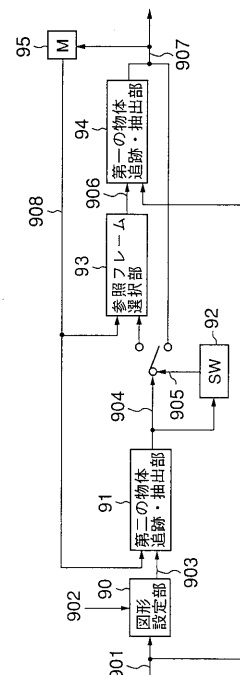
【図 10】



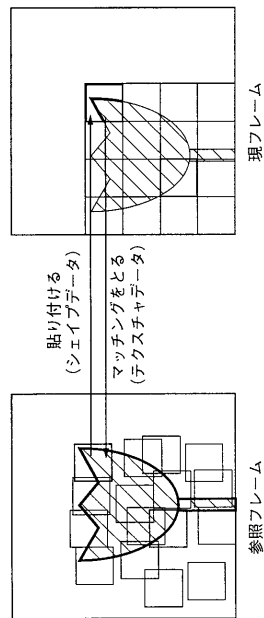
【図 11】



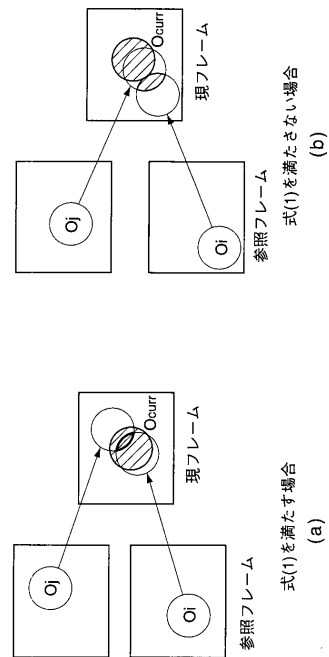
【図 12】



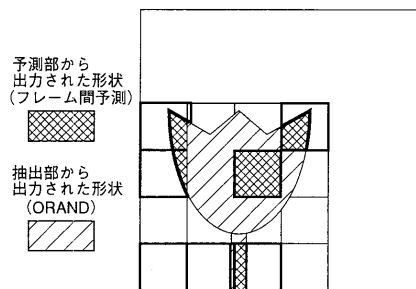
【図 1 3】



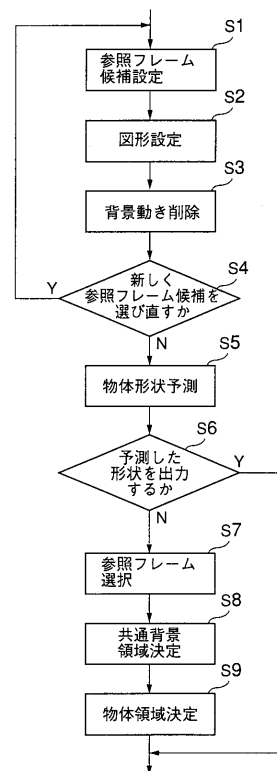
【図 1 4】



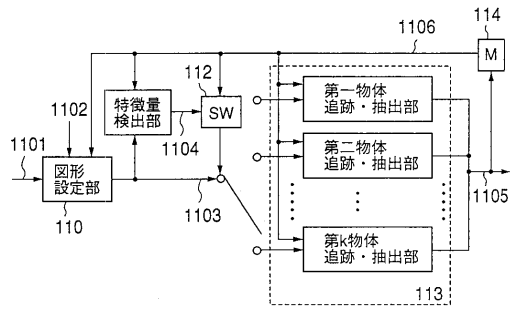
【図 1 5】



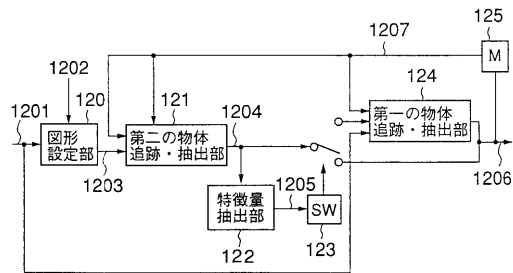
【図 1 6】



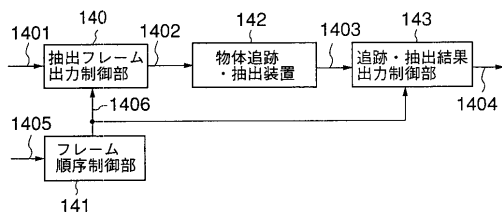
【図 17】



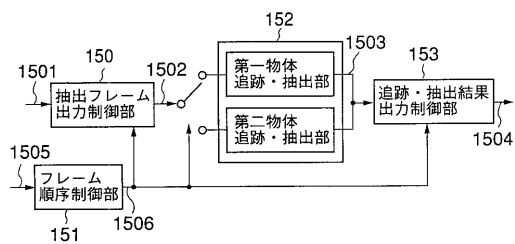
【図 18】



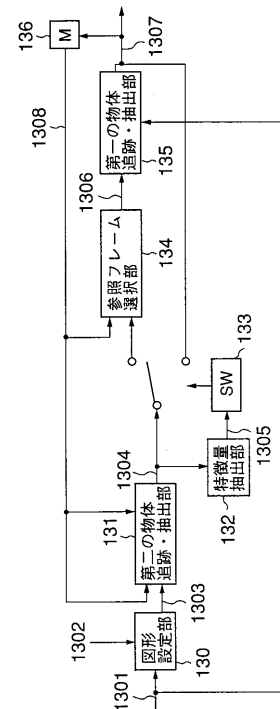
【図 20】



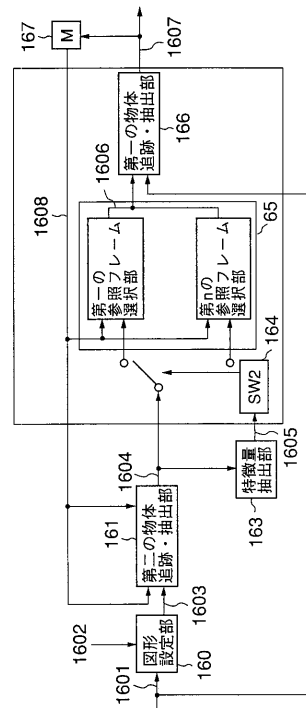
【図 21】



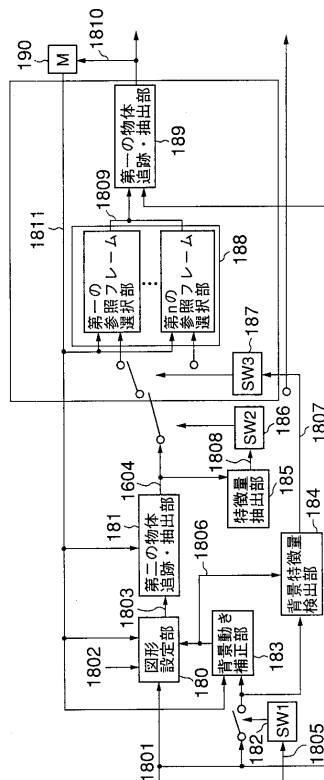
【図 19】



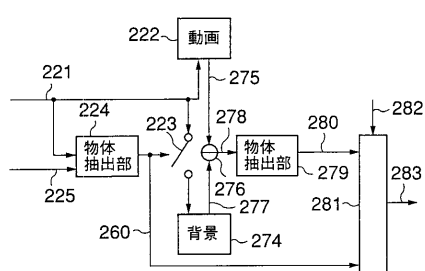
【図 22】



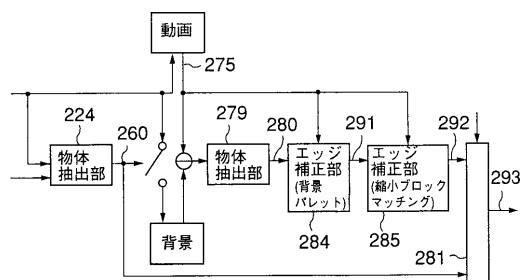
【图 24】



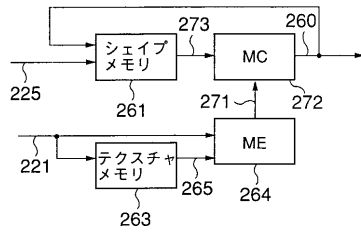
【図 27】



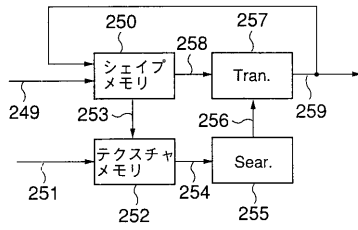
【图 28】



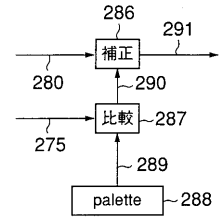
【図 29】



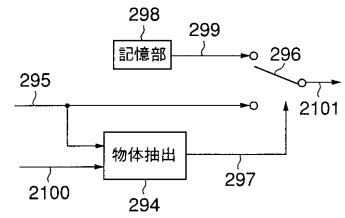
【図 30】



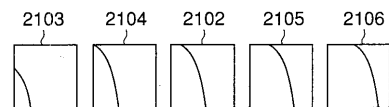
【図 31】



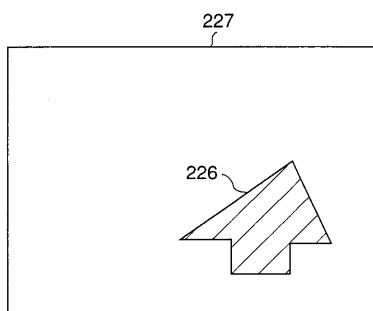
【図 32】



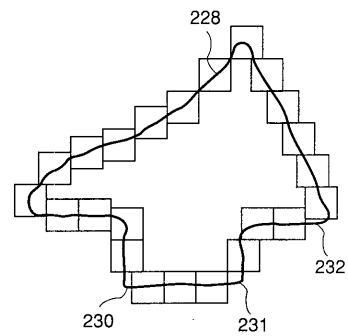
【図 33】



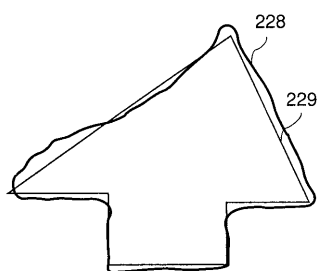
【図 34】



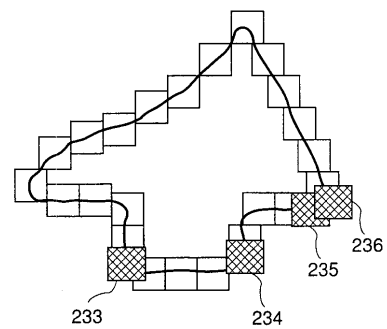
【図 36】



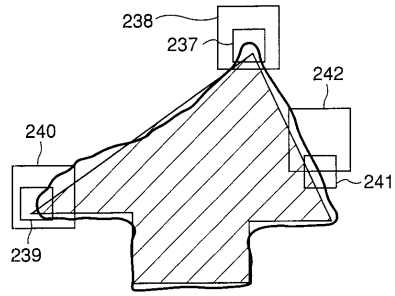
【図 35】



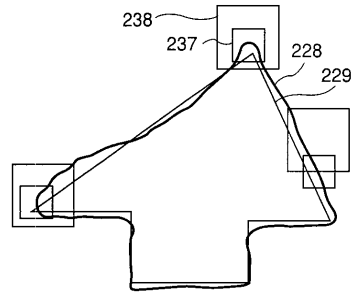
【図 37】



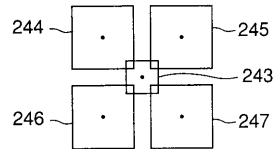
【図 3 8】



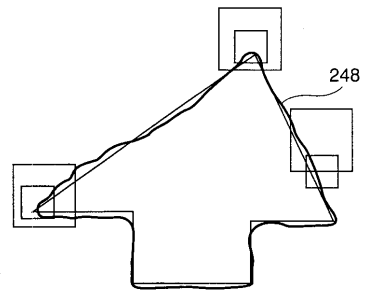
【図 4 1】



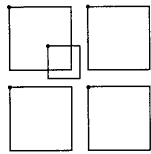
【図 3 9】



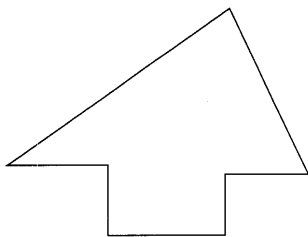
【図 4 2】



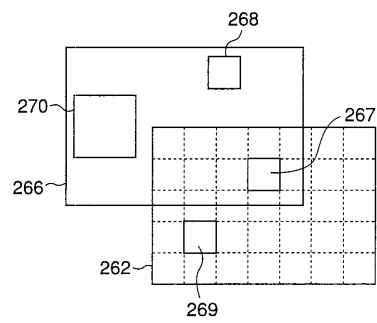
【図 4 0】



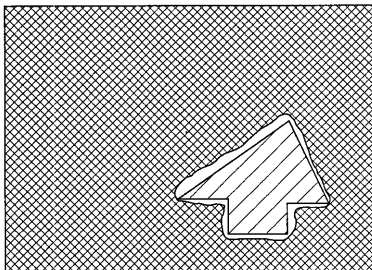
【図 4 3】



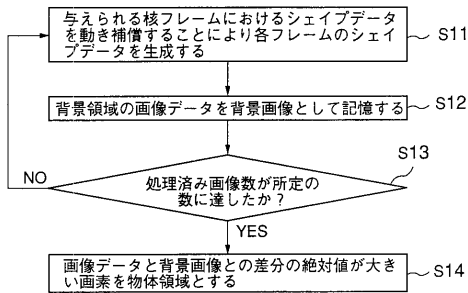
【図 4 5】



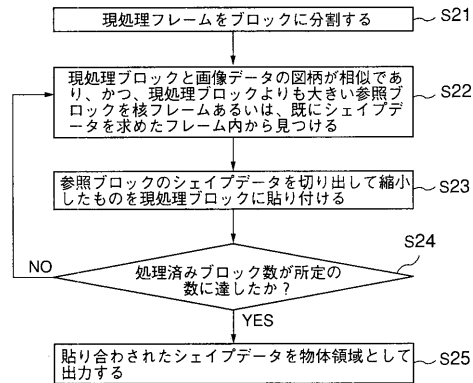
【図 4 4】



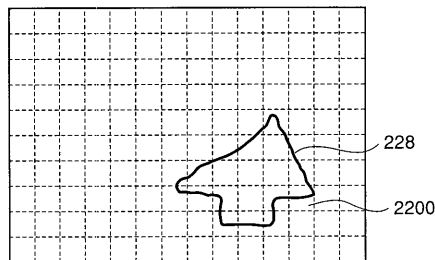
【図 46】



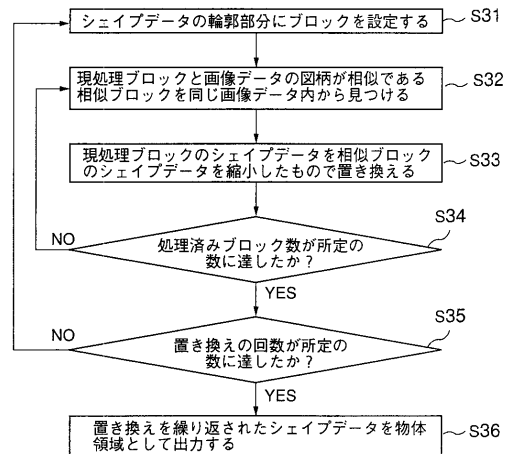
【図 47】



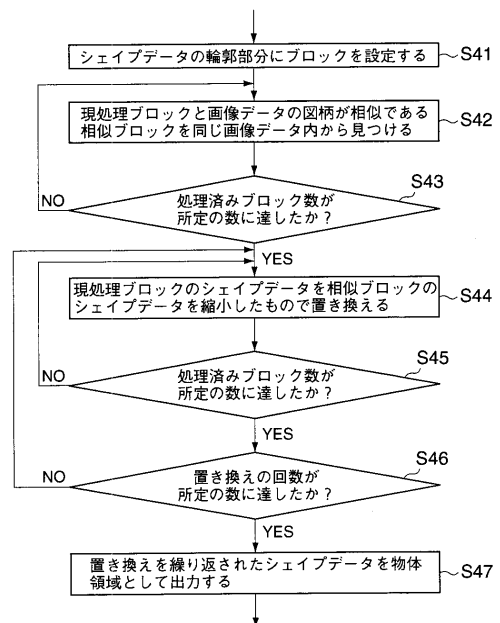
【図 49】



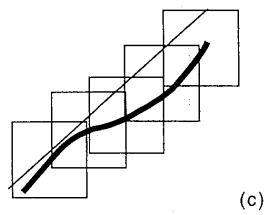
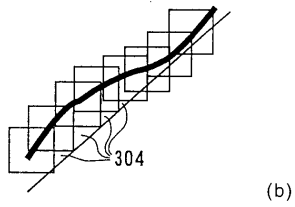
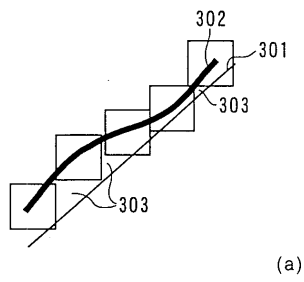
【図 48】



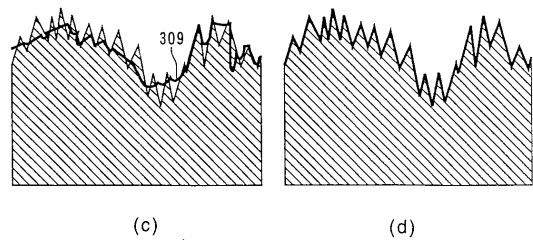
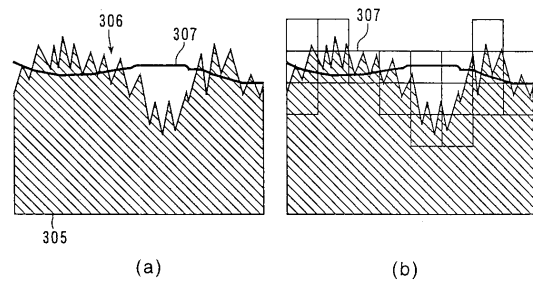
【図 50】



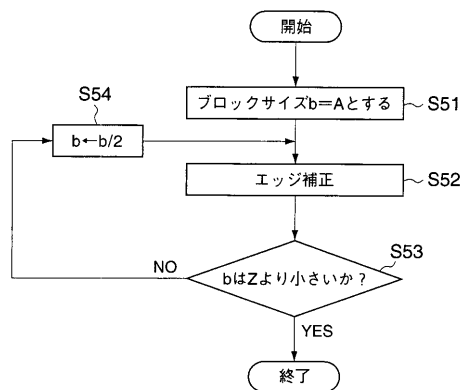
【図 5 1】



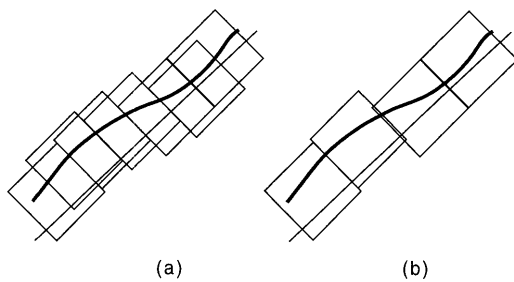
【図 5 2】



【図 5 3】



【図 5 4】



フロントページの続き

(74)代理人 100084618

弁理士 村松 貞男

(74)代理人 100092196

弁理士 橋本 良郎

(72)発明者 三本杉 陽子

神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

(72)発明者 渡邊 敏明

神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

(72)発明者 井田 孝

神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

Fターム(参考) 5L096 AA06 CA04 GA08 GA51 HA03