

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4780564号
(P4780564)

(45) 発行日 平成23年9月28日(2011.9.28)

(24) 登録日 平成23年7月15日(2011.7.15)

(51) Int.Cl.		F I			
G06T	7/20	(2006.01)	G06T	7/20	200B
G06T	3/00	(2006.01)	G06T	3/00	400A
H04N	5/272	(2006.01)	H04N	5/272	

請求項の数 7 (全 26 頁)

(21) 出願番号	特願2006-314485 (P2006-314485)	(73) 特許権者	393031586
(22) 出願日	平成18年11月21日(2006.11.21)		株式会社国際電気通信基礎技術研究所
(65) 公開番号	特開2008-129864 (P2008-129864A)		京都府相楽郡精華町光台二丁目2番地2
(43) 公開日	平成20年6月5日(2008.6.5)	(74) 代理人	100067828
審査請求日	平成21年10月23日(2009.10.23)		弁理士 小谷 悦司
(出願人による申告) 平成18年度独立行政法人情報通信研究機構、研究テーマ「日常行動・状況理解に基づく知識共有システムの研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願		(74) 代理人	100096150
			弁理士 伊藤 孝夫
		(74) 代理人	100109438
			弁理士 大月 伸介
		(72) 発明者	キム ハンソン
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
		(72) 発明者	坂本 電基
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
			最終頁に続く

(54) 【発明の名称】 画像処理装置、画像処理方法及び画像処理プログラム

(57) 【特許請求の範囲】

【請求項1】

背景及び前景を含む動画像から前景を構成する対象物の領域である前景領域を抽出する画像処理装置であって、

処理対象画像となる動画像を取得する取得手段と、

前記取得手段により取得された処理対象画像から、前景を含まず、背景を含む背景モデル画像の輝度成分からなる輝度背景画像を差分して差分画像を作成する背景差分手段と、

前記差分画像の輝度成分を基に、前記処理対象画像を、当該領域が前景でなく背景である信頼性が高い第1の領域と、当該領域が前景でなく背景である信頼性が前記第1の領域より低い第2の領域と、当該領域が背景でなく前景である信頼性が低い第3の領域と、当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域とに区分する領域区分手段と、

前記第3領域の色成分と前記背景モデル画像の色成分との差分に基づいて、前記第3領域から対象物の陰の領域を抽出し、抽出した陰の領域を前記第2の領域に変更することにより、影の領域を除去する影除去手段と、

前記影除去手段により影の領域が除去された第1乃至4の領域を基に、前記処理対象画像から前記前景領域を抽出する抽出手段とを備えることを特徴とする画像処理装置。

【請求項2】

前記抽出手段は、

前記影除去手段により変更されていない第3の領域と、前記第4の領域とに対して、各領

域により表される対象物を特定するためのラベリングを行うラベリング手段と、

前記ラベリング手段によりラベリングされた各領域の輪郭を抽出する輪郭抽出手段と、

前記輪郭抽出手段により輪郭が抽出された領域において、前記1の領域、前記第2の領域、前記第3の領域、前記第4の領域の順に領域成長法を適用して、各領域を成長させる領域成長手段と、

前記領域成長手段により成長された各領域のうち前記第3及び第4の領域を前記前景領域として抽出する前景抽出手段とを含むことを特徴とする請求項1記載の画像処理装置。

【請求項3】

前記ラベリング手段は、前記影除去手段により変更されていない第3の領域と、前記4の領域との中から所定の大きさより小さい領域を除去し、残りの領域に対してラベリングを行うことを特徴とする請求項2記載の画像処理装置。

10

【請求項4】

前記領域区分手段は、前記背景モデル画像の輝度成分の標準偏差を基に、前記1乃至前記第4の領域を区分することを特徴とする請求項1～3のいずれかに記載の画像処理装置。

【請求項5】

前記影除去手段は、前記背景モデル画像の色成分の標準偏差を基に、前記第3領域から対象物の陰の領域を抽出することを特徴とする請求項1～4のいずれかに記載の画像処理装置。

【請求項6】

取得手段と、背景差分手段と、領域区分手段と、影除去手段と、抽出手段とを備える画像処理装置を用いて、背景及び前景を含む動画像から前景を構成する対象物の領域である前景領域を抽出する画像処理方法であって、

20

前記取得手段が、処理対象画像となる動画像を取得する第1のステップと、

前記背景差分手段が、前記第1のステップにおいて取得された処理対象画像から、前景を含まず、背景を含む背景モデル画像の輝度成分からなる輝度背景画像を差分して差分画像を作成する第2のステップと、

前記領域区分手段が、前記差分画像の輝度成分を基に、前記処理対象画像を、当該領域が前景でなく背景である信頼性が高い第1の領域と、当該領域が前景でなく背景である信頼性が前記第1の領域より低い第2の領域と、当該領域が背景でなく前景である信頼性が低い第3の領域と、当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域とに区分する第3のステップと、

30

前記影除去手段が、前記第3領域の色成分と、前記背景モデル画像の色成分との差分に基づいて、前記第3領域から対象物の陰の領域を抽出し、抽出した陰の領域を前記第2の領域に変更することにより、影の領域を除去する第4のステップと、

前記抽出手段が、前記第4のステップにおいて影の領域が除去された第1乃至4の領域を基に、前記処理対象画像から前記前景領域を抽出する第5のステップとを含むことを特徴とする画像処理方法。

【請求項7】

背景及び前景を含む動画像から前景を構成する対象物の領域である前景領域を抽出するための画像処理プログラムであって、

40

処理対象画像となる動画像を取得する取得手段と、

前記取得手段により取得された処理対象画像から、前景を含まず、背景を含む背景モデル画像の輝度成分からなる輝度背景画像を差分して差分画像を作成する背景差分手段と、

前記差分画像の輝度成分を基に、前記処理対象画像を、当該領域が前景でなく背景である信頼性が高い第1の領域と、当該領域が前景でなく背景である信頼性が前記第1の領域より低い第2の領域と、当該領域が背景でなく前景である信頼性が低い第3の領域と、当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域とに区分する領域区分手段と、

前記第3領域の色成分と前記背景モデル画像の色成分との差分に基づいて、前記第3領

50

域から対象物の陰の領域を抽出し、抽出した陰の領域を前記第2の領域に変更することにより、影の領域を除去する影除去手段と、

前記影除去手段により影の領域が除去された第1乃至4の領域を基に、前記処理対象画像から前記前景領域を抽出する抽出手段としてコンピュータを機能させることを特徴とする画像処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、背景及び前景を含む動画画像から前景を構成する対象物の領域である前景領域を抽出する画像処理装置、画像処理方法及び画像処理プログラムに関するものである。

10

【背景技術】

【0002】

動画画像から所望する対象物を抽出することは、監視カメラ、遠隔会議、画像編集、人とコンピュータとのインタフェース等の応用を含む画像処理分野の一つの重要な課題である。従来の対象物抽出アルゴリズムは、主要な区分基準に基づいて、以下の二つのアプローチに大別することができる。

【0003】

第1のアプローチは、基準として、空間的均一性を用いるものであり、形態的フィルタが画像を単純化するために使用され、Watershedアルゴリズムが領域境界決定に適用される。このアルゴリズムによる抽出結果は、Watershedアルゴリズムを用いているため、対象物の境界を他の方法より正確に追跡することができる。

20

【0004】

第2のアプローチは、動画画像における変化の検出を利用するものである。このアプローチに基づくいくつかのアルゴリズムでは、フレーム差分が使用され、ある期間の複数の画像から先行して取得された静的な背景画像を現在の画像から減算する背景差分が、最も一般的に使用される。このアルゴリズムは非常に高速に動作し、静的な背景から意味のある対象物の領域を抽出することができるので、多くの視覚システムに使用されている。

【0005】

例えば、非特許文献1には、オリジナルな背景の領域、影のある背景及び影の領域、ハイライトされた背景の領域、前景の動いている対象物の領域の4つの領域に画像を区分して対象物を抽出する頑強な背景差分及び影抽出方法が開示されている。

30

(例えば、非特許文献1参照)。

【非特許文献1】ティホーブラサート(T Horprasert)他2名、「頑強な背景差分及び影抽出」(A Robust Background Subtraction and Shadow Detection)、Proc. ACCV、2000年

【発明の開示】

【発明が解決しようとする課題】

【0006】

しかしながら、上記の従来背景差分及び影抽出方法では、予め固定された4つの領域に画像を区分しているため、背景領域における対象物の陰及び照明のハイライトにより間違った領域を前景領域として抽出する場合があります、前景領域を正確に抽出することができない。

40

【0007】

本発明の目的は、背景及び前景を含む動画画像から前景領域を高精度に且つ高速に抽出することができる画像処理装置、画像処理方法及び画像処理プログラムを提供することである。

【課題を解決するための手段】

【0008】

本発明に係る画像処理装置は、背景及び前景を含む動画画像から前景を構成する対象物の領域である前景領域を抽出する画像処理装置であって、処理対象画像となる動画画像を取得

50

する取得手段と、前記取得手段により取得された処理対象画像から、前景を含まず、背景を含む背景モデル画像の輝度成分からなる輝度背景画像を差分して差分画像を作成する背景差分手段と、前記差分画像の輝度成分を基に、前記処理対象画像を、当該領域が前景でなく背景である信頼性が高い第1の領域と、当該領域が前景でなく背景である信頼性が前記第1の領域より低い第2の領域と、当該領域が背景でなく前景である信頼性が低い第3の領域と、当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域とに区分する領域区分手段と、前記第3領域の色成分と前記背景モデル画像の色成分との差分に基づいて、前記第3領域から対象物の陰の領域を抽出し、抽出した陰の領域を前記第2の領域に変更することにより、影の領域を除去する影除去手段と、前記影除去手段により影の領域が除去された第1乃至4の領域を基に、前記処理対象画像から前記前景領域を抽出する抽出手段とを備えるものである。

10

【0009】

本発明に係る画像処理装置において、処理対象画像となる動画画像が取得され、取得された処理対象画像から背景モデル画像の輝度成分が差分されて差分画像が作成され、この差分画像の輝度成分を基に、当該領域が前景でなく背景である信頼性が高い第1の領域と、当該領域が前景でなく背景である信頼性が前記第1の領域より低い第2の領域と、当該領域が背景でなく前景である信頼性が低い第3の領域と、当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域とに区分され、区分された第3領域の色成分と背景モデル画像の色成分との差に基づいて、第3領域から対象物の陰の領域が抽出され、抽出された陰の領域が第2の領域に変更されて影の領域が除去される。ここで、色成分は、輝度成分と異なり、対象物の陰及び照明のハイライトの影響を受けにくいいため、対象物の陰の領域を正しく検出し、当該領域を第2の領域に変更して前景領域から確実に排除することができる。このように、影の領域が前景領域から確実に除去された第1乃至4の領域を基に、処理対象画像から前景領域を抽出しているため、背景及び前景を含む動画画像から前景領域を高精度に抽出できるとともに、処理が高速な差分処理を基本にして上記の各処理を構成しているため、動画画像から前景領域を高速に抽出することができる。

20

【0010】

前記抽出手段は、前記影除去手段により変更されていない第3の領域と、前記4の領域とに対して、各領域により表される対象物を特定するためのラベリングを行うラベリング手段と、前記ラベリング手段によりラベリングされた各領域の輪郭を抽出する輪郭抽出手段と、前記輪郭抽出手段により輪郭が抽出された領域において、前記1の領域、前記第2の領域、前記第3の領域、前記第4の領域の順に領域成長法を適用して、各領域を成長させる領域成長手段と、前記領域成長手段により成長された各領域のうち前記第3及び第4の領域を前記前景領域として抽出する前景抽出手段とを含むことが好ましい。

30

【0011】

この場合、変更されていない第3の領域と4の領域とに対して、各領域により表される対象物を特定するためのラベリングが行われ、ラベリングされた各領域の輪郭が抽出され、輪郭が抽出された領域において、1の領域、第2の領域、第3の領域、第4の領域の順に領域成長法を適用して各領域が成長されるので、前景領域である第3及び4の領域の内部に本来あるべきホール、例えば、二人の人間が手を繋ぐことにより形成されるホールを背景領域として確実に復元することができる。このように、本来あるべきホールが確実に復元された後の第3及び第4の領域を前景領域として抽出しているため、前景に位置する対象物の形態を高精度に抽出することができる。

40

【0012】

前記ラベリング手段は、前記影除去手段により変更されていない第3の領域と、前記4の領域との中から所定の大きさより小さい領域を除去し、残りの領域に対してラベリングを行うことが好ましい。

【0013】

この場合、変更されていない第3の領域と4の領域との中から所定の大きさより小さい

50

領域を除去し、残りの領域に対してラベリングを行っているので、前景領域から小さなノイズ領域を確実に排除することができる。

【0014】

前記領域区分手段は、前記背景モデル画像の輝度成分の標準偏差を基に、前記1乃至前記第4の領域を区分することが好ましい。

【0015】

ここで、背景差分の分布は、ガウス分布ではなく、ゼロで非常にシャープなピークを有する零平均ラプラシアン分布であるため、背景モデル画像の輝度成分の標準偏差を基準に1乃至前記第4の領域を区分することにより、各領域を高精度に区分することができる。

【0016】

前記影除去手段は、前記背景モデル画像の色成分の標準偏差を基に、前記第3領域から対象物の陰の領域を抽出することが好ましい。

【0017】

ここで、背景差分の分布は、ガウス分布ではなく、ゼロで非常にシャープなピークを有する零平均ラプラシアン分布であるため、背景モデル画像の色成分の標準偏差を基準に第3領域から対象物の陰の領域を抽出することにより、対象物の陰を前景から確実に除去することができる。

【0018】

本発明に係る画像処理方法は、取得手段と、背景差分手段と、領域区分手段と、影除去手段と、抽出手段とを備える画像処理装置を用いて、背景及び前景を含む動画画像から前景を構成する対象物の領域である前景領域を抽出する画像処理方法であって、前記取得手段が、処理対象画像となる動画画像を取得する第1のステップと、前記背景差分手段が、前記第1のステップにおいて取得された処理対象画像から、前景を含まず、背景を含む背景モデル画像の輝度成分からなる輝度背景画像を差分して差分画像を作成する第2のステップと、前記領域区分手段が、前記差分画像の輝度成分を基に、前記処理対象画像を、当該領域が前景でなく背景である信頼性が高い第1の領域と、当該領域が前景でなく背景である信頼性が前記第1の領域より低い第2の領域と、当該領域が背景でなく前景である信頼性が低い第3の領域と、当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域とに区分する第3のステップと、前記影除去手段が、前記第3領域の色成分と、前記背景モデル画像の色成分との差分に基づいて、前記第3領域から対象物の陰の領域を抽出し、抽出した陰の領域を前記第2の領域に変更することにより、影の領域を除去する第4のステップと、前記抽出手段が、前記第4のステップにおいて影の領域が除去された第1乃至4の領域を基に、前記処理対象画像から前記前景領域を抽出する第5のステップとを含むものである。

【0019】

本発明に係る画像処理プログラムは、背景及び前景を含む動画画像から前景を構成する対象物の領域である前景領域を抽出するための画像処理プログラムであって、処理対象画像となる動画画像を取得する取得手段と、前記取得手段により取得された処理対象画像から、前景を含まず、背景を含む背景モデル画像の輝度成分からなる輝度背景画像を差分して差分画像を作成する背景差分手段と、前記差分画像の輝度成分を基に、前記処理対象画像を、当該領域が前景でなく背景である信頼性が高い第1の領域と、当該領域が前景でなく背景である信頼性が前記第1の領域より低い第2の領域と、当該領域が背景でなく前景である信頼性が低い第3の領域と、当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域とに区分する領域区分手段と、前記第3領域の色成分と前記背景モデル画像の色成分との差分に基づいて、前記第3領域から対象物の陰の領域を抽出し、抽出した陰の領域を前記第2の領域に変更することにより、影の領域を除去する影除去手段と、前記影除去手段により影の領域が除去された第1乃至4の領域を基に、前記処理対象画像から前記前景領域を抽出する抽出手段としてコンピュータを機能させるものである。

【発明の効果】

【0020】

本発明によれば、差分画像の輝度成分を基に処理対象画像を信頼性に応じた２種類の背景領域（第１及び第２の領域）及び２種類の前景領域（第３及び第４の領域）に正確に区分し、この第３領域の色成分と背景モデル画像の色成分との差に基づいて、第３領域から対象物の陰の領域が抽出されて除去され、影の領域が確実に除去された第１乃至４の領域を基に処理対象画像から前景領域を抽出しているため、背景及び前景を含む動画像から前景領域を高精度且つ高速に抽出することができる。

【発明を実施するための最良の形態】

【００２１】

以下、本発明の一実施の形態による画像処理装置について図面を参照しながら説明する。図１は、本発明の一実施の形態による画像処理装置の構成を示すブロック図である。

10

【００２２】

図１に示す画像処理装置は、複数のビデオカメラ１１～１ｎ（ｎは任意の整数）、背景モデル記憶部１２、画像取得部１３、背景差分部１４、領域区分部１５、影除去部１６、ラベリング部１７、輪郭抽出部１８、領域成長部１９、及び前景領域抽出部２０を備える。

【００２３】

ｎ台（ｎは１以上の整数）のビデオカメラ１１～１ｎは、撮影空間の所定箇所に固定され、前景となる対象物、例えば、人物をその背後に位置する背景、例えば、室内の壁、調度品等を含めて撮影し、撮影したカラー動画像を処理対象画像として画像取得部１３へ出力する。例えば、ビデオカメラ１１～１ｎとして、７台のビデオカメラを用い、１０２４

20

× 768画素のRGBストリームが30フレーム/秒で撮影される。

【００２４】

背景モデル記憶部１２、画像取得部１３、背景差分部１４、領域区分部１５、影除去部１６、ラベリング部１７、輪郭抽出部１８、領域成長部１９、及び前景領域抽出部２０は、入力装置、表示装置、ROM（リードオンリメモリ）、CPU（中央演算処理装置）、RAM（ランダムアクセスメモリ）、画像I/F（インターフェース）部及び外部記憶装置等を備えるコンピュータを用いて、ROM又は外部記憶装置に予め記憶されている、後述する各処理を行うための画像処理プログラムをCPU等で実行することにより実現される。なお、背景モデル記憶部１２、画像取得部１３、背景差分部１４、領域区分部１５、影除去部１６、ラベリング部１７、輪郭抽出部１８、領域成長部１９、及び前景領域抽出部２０の構成例は、本例に特に限定されず、各ブロックを専用のハードウェアから構成したり、一部のブロック又はブロック内の一部の処理のみを専用のハードウェアで構成したりする等の種々の変更が可能である。

30

【００２５】

背景モデル記憶部１２は、前景を含まず、背景を含む背景モデル画像の、輝度成分からなる輝度背景画像データ及び色成分からなる色背景画像データと、領域区分及び陰除去の基準となる複数の閾値とを予め記憶している。すなわち、背景モデル画像は、２つの識別可能な部分、輝度モデル画像と色モデル画像とにモデル化され、輝度背景画像データが輝度モデル画像となり、色背景画像データが色モデル画像となる。

【００２６】

ビデオカメラ１１～１ｎから入力されるビデオストリームは、RGB成分の３チャンネルを有しているが、それらはノイズ及び照明条件に対して敏感であるため、本実施の形態では、初期的な対象物すなわち前景の区分にカラー画像の輝度成分を使用している。背景モデル画像の輝度すなわち輝度背景画像データは、下記式で計算される。

40

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad \dots (1)$$

【００２７】

しかしながら、輝度成分は、背景領域の影及び前景領域の照明の反射により急激に変化する。本実施の形態では、間違った前景の区分を防止するために、画像の色成分を用いた色背景画像データを用いる。背景モデル画像の色成分Hすなわち色背景画像データは、下記のようにしてHSIモデルを基に算出される。なお、本発明に使用可能なモデルは、こ

50

の例に特に限定されず、HSVモデルを用いても差し支えない。

【0028】

【数1】

$$\begin{aligned}
 I &= \max(R, G, B) \\
 S &= \begin{cases} (I - \min(R, G, B)) / I & \text{if } I \neq 0 \\ 0 & \text{otherwise} \end{cases} \\
 H &= \begin{cases} (G - B) \times 60 / S & \text{if } I = R \\ 180 + (B - R) \times 60 / S & \text{if } I = G \\ 240 + (R - G) \times 60 / S & \text{if } I = B \end{cases} \dots(2) \\
 &\text{if } H < 0 \text{ then } H = H + 360
 \end{aligned}$$

【0029】

ここで、ノイズ、背景の素材の特性及び照明条件の変化等により、静的なシーンにおいてさえ、フレーム間の変化が発生する。図2は、連続した150フレームから計算された各画素の平均からの偏差の分布（ノイズ分布）を示す図であり、(a)は輝度成分を示し、(b)は色成分を示している。

【0030】

ビデオカメラ11～1nが所定期間同一シーンを観測し続けている場合、ノイズは、キャンセル又は抑圧され、零平均ラプラシアン分布として表現することができる。したがって、本実施の形態では、ビデオカメラ11～1nにより連続的に、例えば、150フレーム間撮影された背景モデル画像の平均及び標準偏差を算出し、算出した各画素のRGB成分の平均値から、式(1)及び式(2)を用いて、輝度背景画像データ及び色背景画像データを予め計算し、これらの輝度背景画像データ及び色背景画像データと、輝度背景画像データの標準偏差に所定の重み係数を乗算した3つの閾値及び色背景画像データの標準偏差に所定の重み係数を乗算した1つの閾値とが背景モデル記憶部12に予め記憶されている。

【0031】

画像取得部13は、ビデオカメラ11～1nから処理対象画像となる動画像を取得して背景差分部14へ出力する。背景差分部14は、背景モデル記憶部12から輝度背景画像データを読み出し、処理対象画像データから輝度背景画像データを差分して差分画像データを作成して領域区分部15へ出力する。

【0032】

領域区分部15は、輝度背景画像データの標準偏差に所定の重み係数を乗算した3つの閾値を背景モデル記憶部12から読み出し、差分画像データの輝度成分と3つの閾値とを比較する。領域区分部15は、比較結果に応じて、処理対象画像を、当該領域が前景でなく背景である信頼性が高い第1の領域と、当該領域が前景でなく背景である信頼性が第1の領域より低い第2の領域と、当該領域が背景でなく前景である信頼性が低い第3の領域と、当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域とに区分し、影除去部16へ出力する。

【0033】

影除去部16は、背景モデル記憶部12から色背景画像データ及び色背景画像データの標準偏差に所定の重み係数を乗算した1つの閾値を読み出し、第3領域の色成分と背景モデル画像の色成分すなわち色背景画像データとの差分と閾値とを比較する。影除去部16は、比較結果に応じて、第3領域から対象物の陰の領域を抽出し、抽出した陰の領域を第

2の領域に変更することにより、影の領域を除去してラベリング部17へ出力する。

【0034】

ラベリング部17は、影除去部16により第2の領域に変更されていない第3の領域と、4の領域とに対して、各領域により表される対象物を特定するためのラベリングを行い、輪郭抽出部18へ出力する。

【0035】

輪郭抽出部18は、ラベリングされた各領域の輪郭を抽出して領域成長部19へ出力する。領域成長部19は、輪郭が抽出された領域において、1の領域、第2の領域、第3の領域、第4の領域の順に領域成長法を適用して各領域を成長させ、対象物の輪郭の内側に実際に存在するホールを復元し、前景領域抽出部20へ出力する。

10

【0036】

前景領域抽出部20は、実際に存在するホールが復元された第3及び第4の領域を前景領域として抽出し、その他の領域をマスクする前景マスクを生成して動画像をマスクすることにより、動画像から前景領域すなわち対象物を抽出して表示する。

【0037】

上記の構成により、本実施の形態では、複数閾値を用いた背景差分により、処理対象画像がそれらの信頼性に基づいて4つの領域に区分され、対象物の影の領域が色成分を用いて除去される。また、各対象物が自身の識別番号でラベリングされ、シルエット抽出技術により前景領域の境界がスムージング処理されて前景領域内の不要なホールが除去され、領域成長技術により前景領域内に実際に存在するホールが復元される。このように、不要なホール及び影が除去されるとともに、実際に存在するホールが復元された前景領域を用いて前景マスクが生成され、動画像から対象物が確実に且つ正確に、さらに高速に抽出される。

20

【0038】

本実施の形態では、画像取得部13が取得手段の一例に相当し、背景差分部14が背景差分手段の一例に相当し、領域区分部15が領域区分手段の一例に相当し、影除去部16が影除去手段の一例に相当し、ラベリング部17、輪郭抽出部18、領域成長部19及び前景領域抽出部20が抽出手段の一例に相当する。また、ラベリング部17がラベリング手段の一例に相当し、輪郭抽出部18が輪郭抽出手段の一例に相当し、領域成長部19が領域成長手段の一例に相当し、前景領域抽出部20が前景抽出手段の一例に相当する。

30

【0039】

次に、上記のように構成された画像処理装置により、背景及び前景を含む動画像から前景を構成する対象物の領域である前景領域を抽出する画像処理である前景領域抽出処理について説明する。図3は、図1に示す画像処理装置による前景領域抽出処理を説明するためのフローチャートである。

【0040】

まず、ステップS11において、ビデオカメラ11～1nは、前景となる人物をその背後に位置する背景を含めて撮影し、画像取得部13は、撮影された動画像を処理対象画像として取得する。

【0041】

次に、ステップS12において、背景差分部14は、背景モデル記憶部12から輝度背景画像データを読み出し、画像取得部13から出力される処理対象画像データから輝度背景画像データを差分して差分画像データを作成して領域区分部15へ出力する。

40

【0042】

次に、ステップS13において、下記の式(3)に従い、領域区分部15は、輝度背景画像データの標準偏差に所定の重み係数を乗算した3つの閾値を背景モデル記憶部12から読み出し、差分画像データの輝度成分と3つの閾値とを比較して、処理対象画像を、(a)当該領域が前景でなく背景である信頼性が高い第1の領域(信頼できる背景)と、(b)当該領域が前景でなく背景である信頼性が第1の領域より低い第2の領域(疑わしい背景)と、(c)当該領域が背景でなく前景である信頼性が低い第3の領域(疑わしい前

50

景)と、(d)当該領域が背景でなく前景である信頼性が前記第3の領域より高い第4の領域(信頼できる前景)とに区分し、影除去部16へ出力する。なお、下式において、 L_I 及び L_B は、現在のフレーム及び背景モデル画像の輝度成分を示し、 σ は、背景モデル画像の輝度成分の標準偏差を示し、 $K_1 \sim K_3$ は、重み係数を示す。

【0043】

【数2】

$$BD(p) = |L_I(p) - L_B(p)|$$

$$\left\{ \begin{array}{ll} BD < K_1 \sigma(p) & \Rightarrow \text{(a) Reliable Background} \\ K_1 \sigma(p) \leq BD \leq K_2 \sigma(p) & \Rightarrow \text{(b) Suspicious Background} \\ K_2 \sigma(p) \leq BD \leq K_3 \sigma(p) & \Rightarrow \text{(c) Suspicious Foreground} \\ K_3 \sigma(p) \leq BD & \Rightarrow \text{(d) Reliable Foreground} \end{array} \right\} \dots(3)$$

【0044】

ここで、上記の背景差分処理及び領域区分処理について詳細に説明する。図4は、7台のビデオカメラのうちの2台のカメラを用いて生成した背景モデル画像の輝度成分の一例を示す図であり、図5は、図4に示す背景モデル画像の色成分の一例を示す図であり、図6は、本実施の形態に使用した処理対象画像の一例を示す図であり、図7は、図6に示す処理対象画像の差分後の輝度成分を示す図である。なお、図5において、 $0 \sim 360^\circ$ 範囲で記述される色成分は、 $0 \sim 255$ のグレースケール値にマッピングされている。

【0045】

上記の各画像を用いて、現在のフレームの輝度成分から背景モデル画像を減算することにより、初期の背景差分が実行される。このとき、背景と同様の明るさを有する影及び前景領域のような多くの曖昧な領域を、固定されたシンプルな閾値を用いて区分するために、上記の式(3)に示すように、それらの信頼性に基づく複数の閾値を用いて、処理対象領域が4つのカテゴリーに分類される。

【0046】

背景差分の分布がガウス分布であると仮定すると、所望の検出比を達成するように、背景差分の標準偏差によって閾値を決定することができるが、図2に示したように、背景差分の分布は、ガウス分布ではなく、ゼロで非常にシャープなピークを有する零平均ラプラス分布である。このため、本実施の形態では、背景モデル画像の標準偏差を用いて閾値を決定する。画像の各画素の偏差は、背景を構成する素材の特性及び照明条件に基づいて異なるため、背景を長時間撮影することにより、各画素の標準偏差を閾値として使用することができる。

【0047】

図8は、図6及び図7に示す画像を用いた分類結果を示す図である。図8に示す例では、式(3)に用いられる重み係数 $K_1 \sim K_3$ として、5、12、60を用いている。図8に示す黒、濃い灰色、薄い灰色、白の各領域が、第1の領域(信頼できる背景)、第2の領域(疑わしい背景)、第3の領域(疑わしい前景)、第4の領域(信頼できる前景)を示している。

【0048】

次に、ステップS14において、下記の式(4)に従い、影除去部16は、背景モデル記憶部12から色背景画像データ及び色背景画像データの標準偏差に所定の重み係数を乗算した1つの閾値を読み出し、第3領域の色成分と色背景画像データとの差分と閾値とを比較して第3領域から対象物の陰の領域を抽出し、抽出した陰の領域を第2の領域に変更することにより、影の領域を除去してラベリング部17へ出力する。なお、下式において、 H_I 及び H_B は、現在のフレーム及び背景モデル画像の色成分を示し、 H は、背景モ

10

20

30

40

50

デル画像の色成分の標準偏差を示す。

【 0 0 4 9 】

【 数 3 】

$$\left. \begin{array}{l} \text{if } (p \in \text{region } (c) \ \& \ |H_I(p) - H_B(p)| < K_1 \sigma_H(p)) \\ \text{then } p \Rightarrow \text{region } (b) \end{array} \right\} \dots(4)$$

【 0 0 5 0 】

ここで、上記の影除去処理について詳細に説明する。図 8 に示す結果から、対象物の影が背景の明るさを変化させ、フロア上の背景の多くの部分が、第 3 の領域（疑わしい前景）に組み込まれていることがわかる。このため、本実施の形態では、影が背景の色特性を変化させることなく、輝度のみを変化させることに着目し、式（4）に示すように、色成分を用いて、第 3 の領域（疑わしい前景）の影を第 2 の領域（疑わしい背景）にマージし、第 3 の領域から影を除去する。図 9 は、影除去処理後の結果の一例を示す図である。

10

【 0 0 5 1 】

次に、ステップ S 1 5 において、ラベリング部 1 7 は、影除去部 1 6 による影除去処理後の第 3 の領域と、4 の領域とに対して、各領域により表される対象物を特定するためのラベリングを行い、輪郭抽出部 1 8 へ出力する。

【 0 0 5 2 】

20

ここで、上記のラベリング処理について詳細に説明する。このラベリング処理においては、式（3）におけるすべての前景領域である第 3 の領域と 4 の領域が、自身の識別番号でラベリングされる。このとき、4 近傍規則を用いて連結されたすべての前景画素が、領域成長技術を用いて同じラベルを割り付けられる。しかしながら、初期対象物領域である第 3 の領域と 4 の領域には、小さなノイズ領域が存在する可能性がある。

【 0 0 5 3 】

このため、従来のノイズ領域除去方法では、形態的操作を用いて小さいノイズ領域をフィルタリングしていたが、本実施の形態では、閉鎖及び開口処理による初期マスクを改良している。すなわち、ラベリング部 1 7 は、各領域の大きさに基づく降順ですべてのラベリングされた領域をソート及び再ラベリングし、再ラベリング処理において、所定の閾値 TH_{RG} より小さい領域を除去している。図 1 0 は、ラベリング処理の結果の一例を示す図である。図 1 0 に示すように、右側の画像では、二人の人間がシーン内に存在しているが、手を繋いでいるので、一つの対象物としてラベリングされている。

30

【 0 0 5 4 】

次に、ステップ S 1 6 において、輪郭抽出部 1 8 は、ラベリングされた各前景領域の輪郭を抽出して領域成長部 1 9 へ出力する。ここで、図 1 0 に示す状態では、背景として間違っていて区分されたり、対象物の内部に間違っていてラベリングされた多くの間違っただ領域が存在する。このため、本実施の形態では、前景領域の境界をスムージングし、領域内のホールを除去するために、カーマープロファイル抽出技術を改良したシルエット抽出技術を用いて、各領域の輪郭を抽出している。

40

【 0 0 5 5 】

図 1 1 は、初期対象物に適用される輪郭抽出処理を説明するための模式図である。図 1 1 の（a）～（d）に示すように、輪郭抽出部 1 8 は、重み付けされた一画素の厚いドレープ $D 1 \sim D 4$ を初期対象物 RA に対して順次一方側から反対側へ移動させ、4 つのドレープ $D 1 \sim D 4$ によって包まれた領域 SI が最終的に前景領域を示すこととなる。なお、ドレープ $D 1 \sim D 4$ の隣接画素は、その幅が所定の閾値 M より小さい浸潤ギャップのない対象物を覆う弾性パネによって連結されている。

【 0 0 5 6 】

次に、ステップ S 1 7 において、領域成長部 1 9 は、輪郭が抽出された領域において、1 の領域、第 2 の領域、第 3 の領域、第 4 の領域の順に領域成長法を適用して、各領域を

50

成長させ、対象物の輪郭の内側に実際に存在するホールを復元して前景領域抽出部 20 へ出力する。

【0057】

ここで、上記のシルエット抽出技術では、複数の対象物が撮影シーン中に存在するときに、対象物の内部に実際に存在するホールをも覆う可能性があり、対象物間の領域に重大なエラーを引き起こすという欠点を有している。このため、本実施の形態では、第1の問題を避けるために、各ラベリングされた領域に対してシルエット抽出技術を独立に適用し、第2の問題に対して、閾値 TH_{RG} より大きい領域に対して、シルエット内の信頼性が高い背景領域から領域成長技術を順次実行している。

【0058】

図12は、輪郭抽出処理の結果の一例を示す図であり、図13は、内部ホール復元処理の結果の一例を示す図である。図12の(a)及び図13の(a)に示すように、一人の人間が前景に位置する場合、輪郭抽出処理の結果と内部ホール復元処理の結果とで大きな差はないが、図12の(b)及び図13の(b)に示すように、二人の人間が手を繋いだ状態で前景に位置する場合、図12の(b)において喪失していた、二人の人間によって作られたホールが、図13の(b)では、領域成長処理によって第4の領域(信頼できる前景)から復元されていることがわかる。

【0059】

最後に、ステップ S18 において、前景領域抽出部 20 は、実際に存在するホールが復元された第3及び第4の領域を前景領域として抽出し、その他の領域をマスクする前景マスクを生成して動画像をマスクすることにより、動画像から対象物を抽出して表示する。その後、ステップ S11 に戻って、次の処理対象画像に対して上記の処理が繰り返される。

【0060】

次に、上記の前景領域抽出処理を実際に行なった結果について詳細に説明する。ビデオカメラ 11 ~ 1n は、7 台の同期型 IEEE-1394 カメラ(カメラ 1 ~ 7)から構成され、撮影空間の中心方向に向けられ、ほぼ同一シーンを撮影した。各カメラは、1024 × 768 画素の RGB ストリームを 30 フレーム/秒で撮影し、通常のパーソナルコンピュータを用いて上記の前景領域抽出処理を実行した。このとき、背景モデル画像のフレーム数として、 $N = 150$ 、領域区分に使用した標準偏差の各重み係数として、 $K_1 = 5$ 、 $K_2 = 12$ 、 $K_3 = 60$ 、最小有効領域サイズ(画素)として、 $TH_{RG} = 300$ 、輪郭抽出処理に用いる最大ギャップとして、 $M = 12$ を用いた。

【0061】

図14は、各カメラにより撮影された処理対象画像に対する前景領域抽出処理の抽出結果を示す図であり、図15~図21は、カメラ1~7の撮影画像及び区分された前景を示す図である。

【0062】

本例では、照明器具及び背景の操作のような特別な条件なしに、典型的な室内環境において、映像が撮影され、図15~図21の(a)及び(c)の各画像は、7台のカメラ1~7により撮影された同一フレームの画像を示しており、図15~図21の(b)及び(d)の画像の外側は、詳細な境界を示すためにカットされている。

【0063】

図15~図21から、本例では、対象物である人物が素早く動いたにもかかわらず、また、すべてのカメラが異なる方向に向けて異なる背景を撮影しているにもかかわらず、前景を非常に高精度に区分することができ、本画像処理装置が種々の条件下で十分に機能することがわかった。

【0064】

また、図14に示す評価結果は、手動区分によって生成されたグランド・トゥース区分マスクによる区分結果と、本実施の形態による区分結果とを比較し、各マスクにおけるエラー画素の割合を計算したものであり、FBは、背景領域として間違っ

10

20

30

40

50

領域を意味し、F F は、前景領域として間違っ て分類された背景領域を意味する。

【 0 0 6 5 】

一般に、F F のエラーは、F B のエラーより人間の目にとって不快であり、多くの視覚システムにとって受け入れがたい。しかしながら、図 1 4 に示す結果では、素早い動き及び対象物の境界からのボケのために、F F のエラーが F B のエラーより大きくなっているが、前景の対象物の速度及び大きさに確実に依存するため、すべてのシーンにおける平均エラーレートは、1 . 2 5 % より低かった。

【 0 0 6 6 】

さらに、上記の前景領域抽出処理のランタイム分析を行った結果、背景差分処理（ステップ S 1 1 ~ S 1 3 ）は 1 5 m s e c、影除去処理（ステップ S 1 4 ）は 4 7 m s e c、ラベリング処理（ステップ S 1 5 ）は 1 2 5 m s e c、輪郭抽出処理（ステップ S 1 6 ）は 2 3 5 m s e c、ホール復元処理（ステップ S 1 7、S 1 8 ）は 1 2 5 m s e c であり、合計は 5 4 7 m s e c であった。これらの時間は、一人の人間が撮影空間を動いたときの平均処理時間であり、本例の画像解像度では、リアルタイムに処理することはできなかったが、処理の複雑さは画像サイズの自乗に比例するため、V G A 又は Q V G A のような低解像度で撮影した画像の場合、リアルタイムに動作させることができることがわかった。また、G P U のようなハードウェアアクセラレータを用いたり、さらに処理を最適化することによって、X G A 動画に対してもリアルタイムに動作させることができることがわかった。

【 0 0 6 7 】

上記のように、本実施の形態では、差分画像の輝度成分を背景モデル画像の輝度成分の標準偏差に基づく複数の閾値と比較して処理対象画像を信頼性に応じた 2 種類の背景領域（第 1 及び第 2 の領域）及び 2 種類の前景領域（第 3 及び第 4 の領域）に正確に区分し、この第 3 領域の色成分と背景モデル画像の色成分との差を背景モデル画像の色成分の標準偏差に基づく閾値と比較して第 3 領域から対象物の陰の領域を抽出して除去し、影の領域が確実に除去された第 3 及び第 4 の領域を前景領域として抽出しているため、背景及び前景を含むカラー動画から前景領域を高精度且つ高速に抽出することができる。

【 0 0 6 8 】

なお、上記の説明では、抽出される対象物として、人間を例に説明したが、他の動物、他の物体等にも同様に適用することができ、同様の効果を得ることができる。

【 図面の簡単な説明 】

【 0 0 6 9 】

【 図 1 】 本発明の一実施の形態による画像処理装置の構成を示すブロック図である。

【 図 2 】 連続した 1 5 0 フレームから計算された各画素の平均からの偏差の分布を示す図である。

【 図 3 】 図 1 に示す画像処理装置による前景領域抽出処理を説明するためのフローチャートである。

【 図 4 】 7 台のビデオカメラのうちの 2 台のカメラを用いて生成した背景モデル画像の輝度成分の一例を示す図である。

【 図 5 】 図 4 に示す背景モデル画像の色成分の一例を示す図である。

【 図 6 】 本実施の形態に使用した処理対象画像の一例を示す図である。

【 図 7 】 図 6 に示す処理対象画像の差分後の輝度成分を示す図である。

【 図 8 】 図 6 及び図 7 に示す画像を用いた分類結果を示す図である。

【 図 9 】 影除去処理後の結果の一例を示す図である。

【 図 1 0 】 ラベリング処理の結果の一例を示す図である。

【 図 1 1 】 初期対象物に適用される輪郭抽出処理を説明するための模式図である。

【 図 1 2 】 輪郭抽出処理の結果の一例を示す図である。

【 図 1 3 】 内部ホール復元処理の結果の一例を示す図である。

【 図 1 4 】 各カメラにより撮影された処理対象画像に対する前景領域抽出処理の抽出結果を示す図である。

10

20

30

40

50

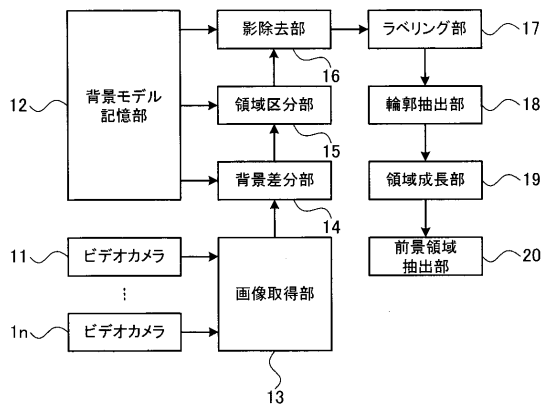
- 【図15】カメラ1の撮影画像及び区分された前景を示す図である。
- 【図16】カメラ2の撮影画像及び区分された前景を示す図である。
- 【図17】カメラ3の撮影画像及び区分された前景を示す図である。
- 【図18】カメラ4の撮影画像及び区分された前景を示す図である。
- 【図19】カメラ5の撮影画像及び区分された前景を示す図である。
- 【図20】カメラ6の撮影画像及び区分された前景を示す図である。
- 【図21】カメラ7の撮影画像及び区分された前景を示す図である。

【符号の説明】

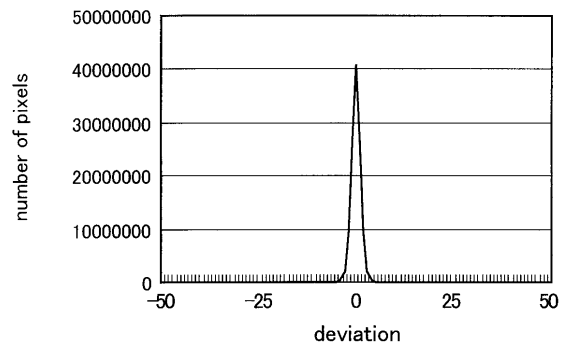
【0070】

- 11 ~ 1n ビデオカメラ
- 12 背景モデル記憶部
- 13 画像取得部
- 14 背景差分部
- 15 領域区分部
- 16 影除去部
- 17 ラベリング部
- 18 輪郭抽出部
- 19 領域成長部
- 20 前景領域抽出部

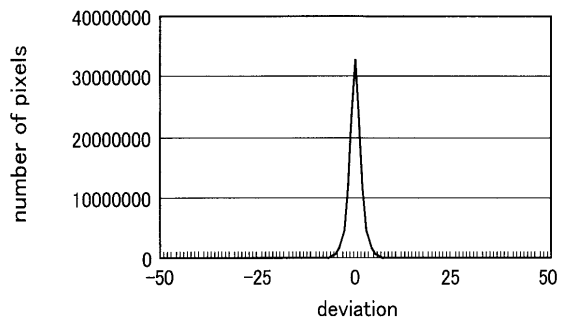
【図1】



【図2】

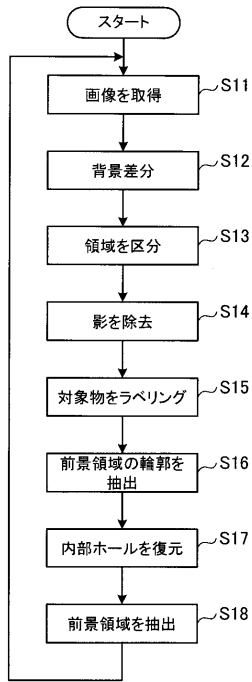


(a)



(b)

【 図 3 】



【 図 1 4 】

	Frame 25		Frame 1067	
	FB	FF	FB	FF
カメラ 1	0.08	0.42	0.16	1.19
カメラ 2	0.17	0.78	0.20	1.57
カメラ 3	0.04	0.57	0.08	0.49
カメラ 4	0.02	0.54	0.00	0.51
カメラ 5	0.09	0.56	0.05	1.67
カメラ 6	0.08	0.83	0.17	1.04
カメラ 7	0.05	0.60	0.14	1.47
Avg.	0.08	0.61	0.11	1.13

【 図 4 】

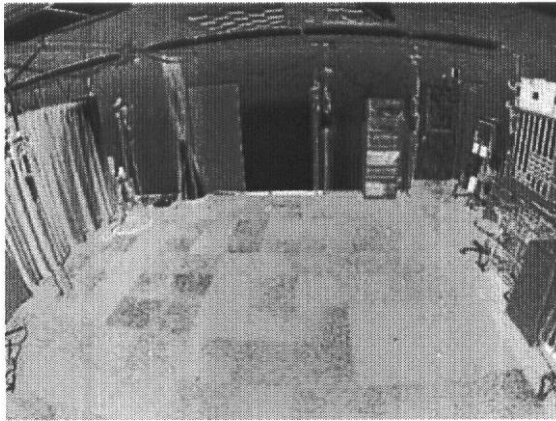


(a)

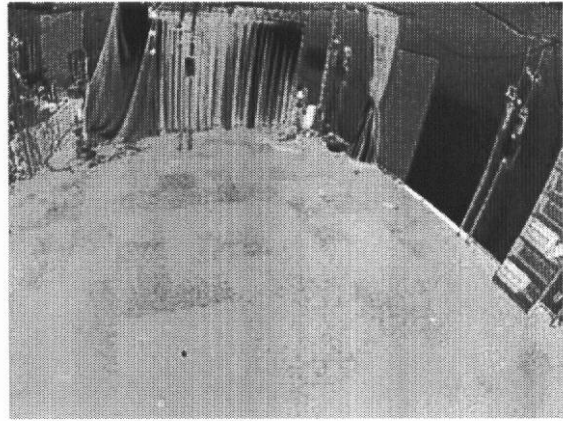


(b)

【 図 5 】



(a)



(b)

【 図 6 】



(a)

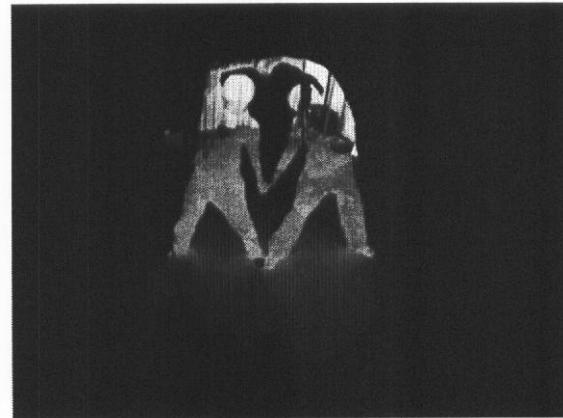


(b)

【 図 7 】

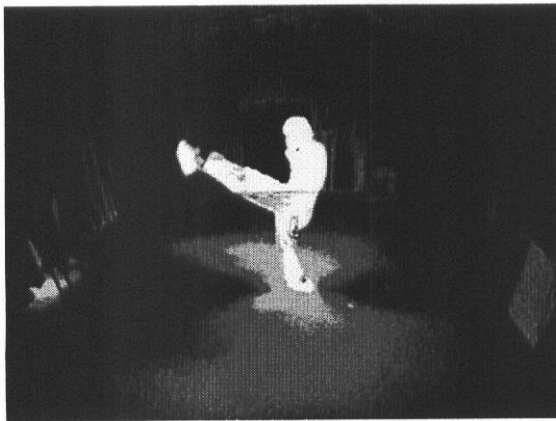


(a)



(b)

【 図 8 】



(a)

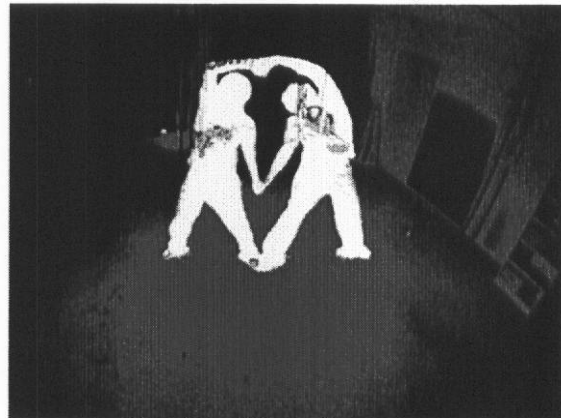


(b)

【 図 9 】



(a)

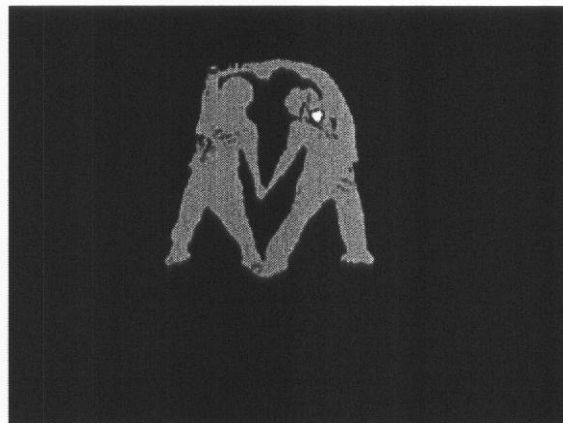


(b)

【 図 10 】

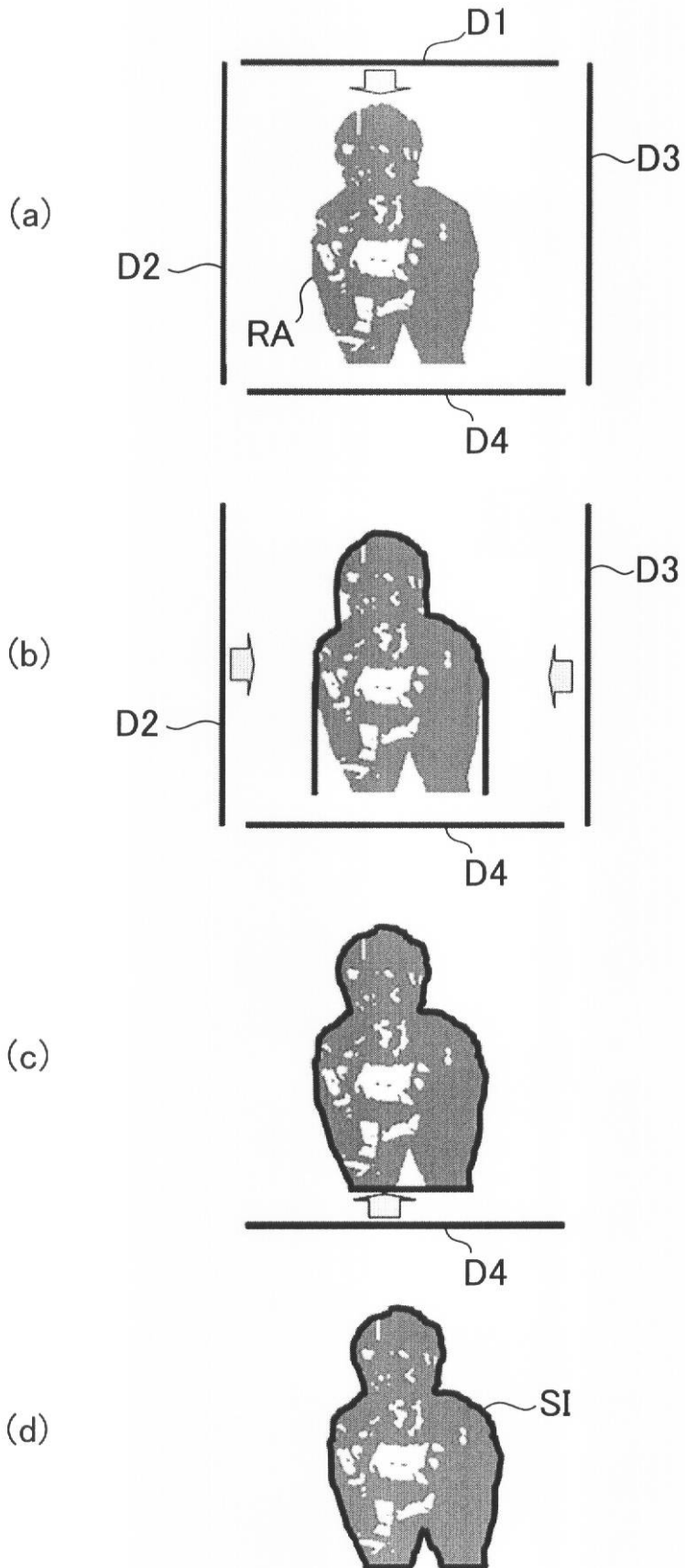


(a)

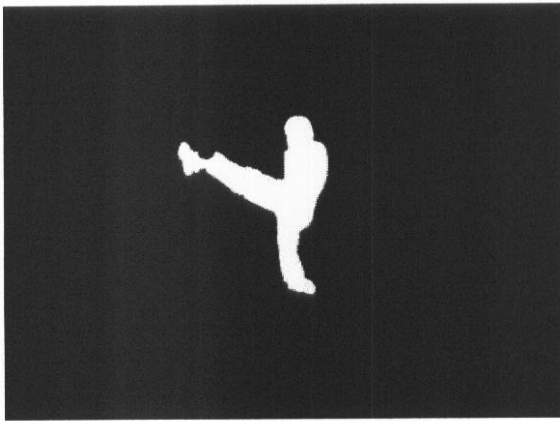


(b)

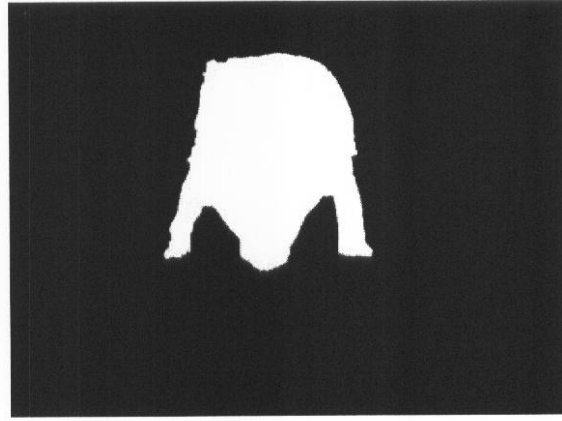
【図11】



【 図 1 2 】

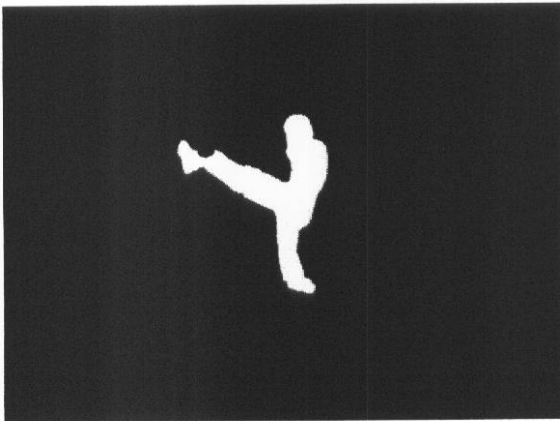


(a)

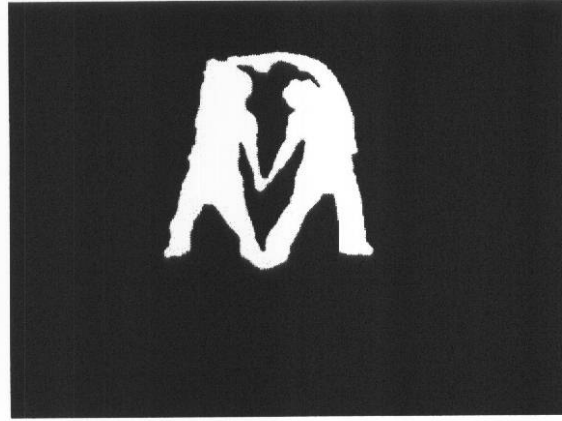


(b)

【 図 1 3 】



(a)

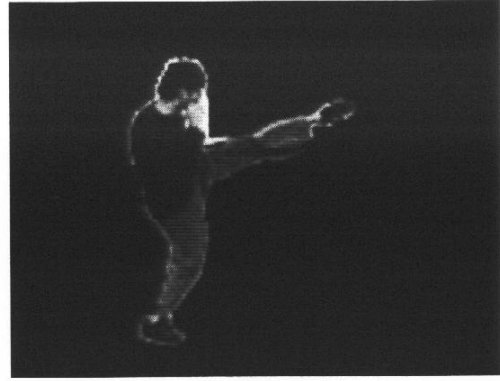


(b)

【 図 15 】



(a)



(b)



(c)



(d)

【 図 16 】



(a)



(b)



(c)



(d)

【 図 17 】



(a)



(b)

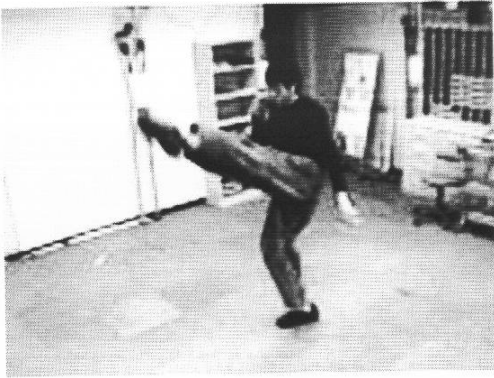


(c)

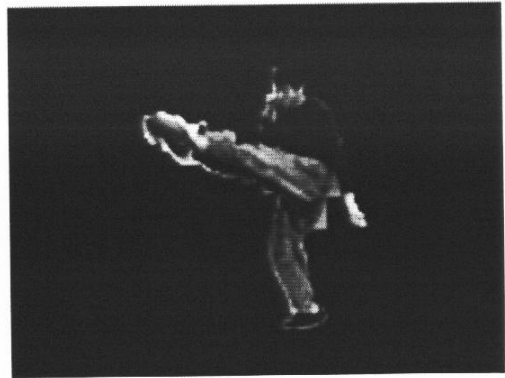


(d)

【 図 18 】



(a)



(b)

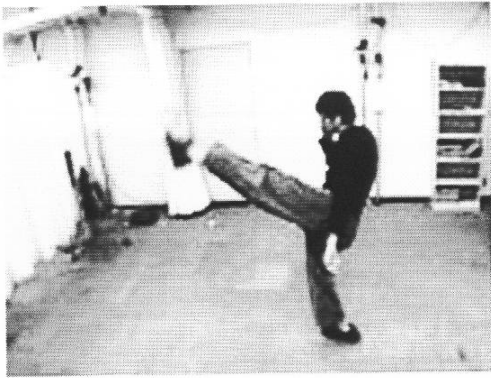


(c)

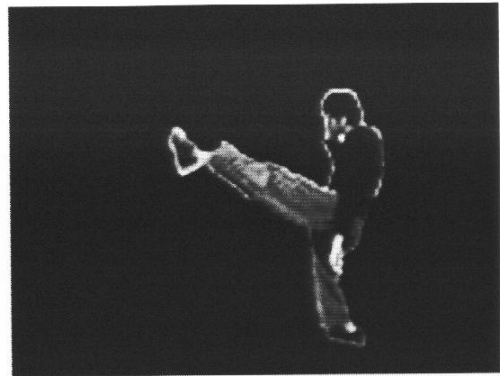


(d)

【 図 19 】



(a)



(b)

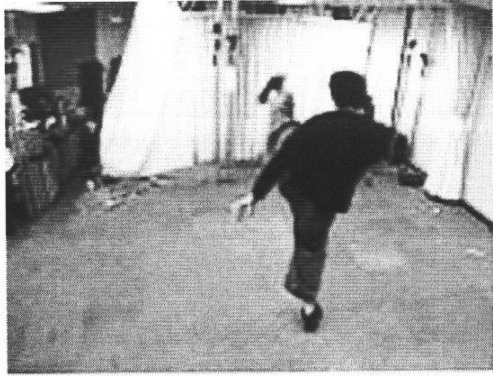


(c)



(d)

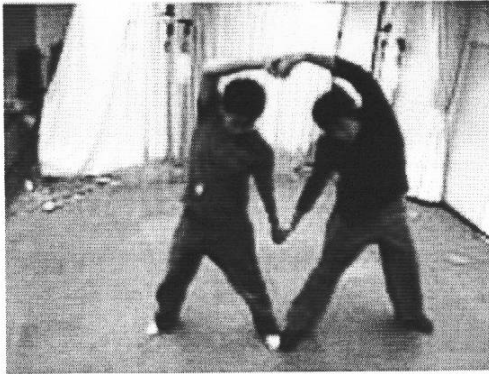
【 図 20 】



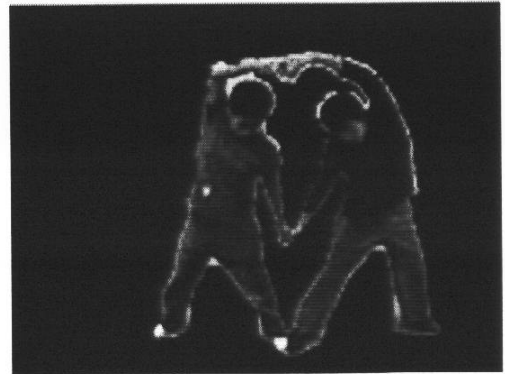
(a)



(b)



(c)



(d)

【 図 2 1 】



(a)



(b)



(c)



(d)

フロントページの続き

- (72)発明者 北原 格
京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
- (72)発明者 鳥山 朋二
京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内
- (72)発明者 小暮 潔
京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

審査官 松永 稔

- (56)参考文献 特開平10-247247(JP,A)
特開平10-021408(JP,A)

- (58)調査した分野(Int.Cl., DB名)
- | | |
|---------|-----------|
| G 0 6 T | 7 / 2 0 |
| G 0 6 T | 3 / 0 0 |
| H 0 4 N | 5 / 2 7 2 |