

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4459795号
(P4459795)

(45) 発行日 平成22年4月28日(2010.4.28)

(24) 登録日 平成22年2月19日(2010.2.19)

(51) Int. Cl. F I
G06T 7/00 (2006.01) G06T 7/00 300F
G06T 1/00 (2006.01) G06T 1/00 340A

請求項の数 7 (全 29 頁)

(21) 出願番号	特願2004-350291 (P2004-350291)	(73) 特許権者	000002945
(22) 出願日	平成16年12月2日(2004.12.2)		オムロン株式会社
(65) 公開番号	特開2005-327242 (P2005-327242A)		京都市下京区塩小路通堀川東入南不動堂町
(43) 公開日	平成17年11月24日(2005.11.24)		801番地
審査請求日	平成18年6月16日(2006.6.16)	(73) 特許権者	598098331
(31) 優先権主張番号	200410038193.6		ツインファ ユニバーシティ
(32) 優先日	平成16年5月14日(2004.5.14)		中華人民共和国 ベイジン 100084
(33) 優先権主張国	中国 (CN)		, ハイダン ディストリクト, ツイン ファ ユニバーシティ
		(74) 代理人	100085006
			弁理士 世良 和信
		(74) 代理人	100106622
			弁理士 和久田 純一

最終頁に続く

(54) 【発明の名称】 特定被写体検出装置

(57) 【特許請求の範囲】

【請求項1】

画像の注目領域から2つの特徴量を取得しそれらの相対値を算出する算出処理を、複数種類実行可能な算出手段と、

前記相対値の取り得る範囲が複数の区間に区分されており、且つ、区間毎に判定値が対応付けられているルックアップテーブルを、前記算出処理毎に記憶する記憶手段と、

前記複数種類の算出処理で得られる複数の相対値の中から選択された1つ以上の相対値を用いて、前記注目領域内に特定の被写体が含まれるか否かを判定する第一の判定手段と

、
選択する相対値の組み合わせを変えて前記第一の判定手段に判定処理を行わせることにより、複数の判定結果を前記第一の判定手段から取得する制御手段と、

前記制御手段によって取得される複数の判定結果に基づいて、前記注目領域内に特定の被写体が含まれるか否かについての最終的な判定を行う第二の判定手段と、
を有し、

前記判定値は、予め用意された前記特定の被写体を含む画像である複数の正解画像から得られた相対値の度数分布と、予め用意された前記特定の被写体を含まない画像である複数の不正解画像から得られた相対値の度数分布とに基づいて決定される値であって、当該判定値に対応する区間内の相対値をもつ画像が前記特定の被写体を含む画像である確率が高いか否かを示す値であり、

前記第一の判定手段は、

10

20

前記選択された1つ以上の相対値毎に、その相対値を含む区間に対応付けられている判定値を、対応するルックアップテーブルから取得し、

該取得した判定値を元に得点を算出し、

該得点に基づいて前記注目領域内に特定の被写体が含まれるか否かを判定する特定被写体検出装置。

【請求項2】

前記2つの特徴量は所定のパターンに基づいて得られる値であって、

前記複数種類の算出処理は、複数の異なるパターンのそれぞれに基づいた算出処理である

請求項1に記載の特定被写体検出装置。

10

【請求項3】

前記第一の判定手段は、新たな組み合わせの複数の相対値から新たな得点を算出する際に、当該新たな組み合わせの複数の相対値それぞれに対する複数の判定値、及び既に第一の判定手段による判定処理が完了している他の組み合わせの相対値から算出された得点を用いる請求項1または2に記載の特定被写体検出装置。

【請求項4】

情報処理装置に、

画像の注目領域から2つの特徴量を取得しそれらの相対値を算出する算出処理を行うステップであって、複数種類の前記算出処理を実行可能なステップと、

前記複数種類の算出処理で得られる複数の相対値の中から選択された1つ以上の相対値を用いて、前記注目領域内に特定の被写体が含まれるか否かを判定する第一の判定ステップと、

20

選択する相対値の組み合わせを変えて前記第一の判定ステップでの判定処理を行うことにより、複数の判定結果を取得するステップと、

前記取得される複数の判定結果に基づいて、前記注目領域内に特定の被写体が含まれるか否かについての最終的な判定を行う第二の判定ステップと、
を実行させるプログラムであって、

前記情報処理装置は、前記相対値の取り得る範囲が複数の区間に区分されており、且つ、区間毎に判定値が対応付けられているルックアップテーブルを、前記算出処理毎に記憶する記憶手段を有しており、

30

前記判定値は、予め用意された前記特定の被写体を含む画像である複数の正解画像から得られた相対値の度数分布と、予め用意された前記特定の被写体を含まない画像である複数の不正解画像から得られた相対値の度数分布とに基づいて決定される値であって、当該判定値に対応する区間内の相対値をもつ画像が前記特定の被写体を含む画像である確率が高いか否かを示す値であり、

前記第一の判定ステップでは、

前記選択された1つ以上の相対値毎に、その相対値を含む区間に対応付けられている判定値を、対応するルックアップテーブルから取得し、

該取得した判定値を元に得点を算出し、

該得点に基づいて前記注目領域内に特定の被写体が含まれるか否かを判定するプログラム。

40

【請求項5】

前記2つの特徴量は所定のパターンに基づいて得られる値であって、

前記複数種類の算出処理は、複数の異なるパターンのそれぞれに基づいた算出処理である

請求項4に記載のプログラム。

【請求項6】

画像に対して部分領域を定義する領域パターン情報を参照する手段と、

前記領域パターン情報にしたがい前記部分領域から2つの特徴量を取得しそれらの相対値を算出する算出処理を、複数種類実行可能な演算手段と、

50

前記相対値の取り得る範囲が複数の区間に区分されており、且つ、区間毎に判定値が対応付けられているルックアップテーブルを、前記算出処理毎に記憶する記憶手段と、

前記複数種類の算出処理で得られる複数の相対値の中から選択された1つ以上の相対値を用いて、前記画像が所定の属性を有するか否かを判定する第一の判定手段と、

選択する相対値の組み合わせを変えて前記第一の判定手段に判定処理を行わせることにより、複数の判定結果を前記第一の判定手段から取得する制御手段と、

前記制御手段によって取得される複数の判定結果に基づいて、前記画像が所定の属性を有するか否かについての最終的な判定を行う第二の判定手段と、
を有し、

前記判定値は、予め用意された前記所定の属性を有する画像である複数の正解画像から得られた相対値の度数分布と、予め用意された前記所定の属性を有さない画像である複数の不正解画像から得られた相対値の度数分布とに基づいて決定される値であって、当該判定値に対応する区間内の相対値をもつ画像が前記所定の属性を有する画像である確率が高いか否かを示す値であり、

前記第一の判定手段は、

前記選択された1つ以上の相対値毎に、その相対値を含む区間に対応付けられている判定値を、対応するルックアップテーブルから取得し、

該取得した判定値を元に得点を算出し、

該得点に基づいて前記画像が所定の属性を有するか否かを判定する
特定被写体検出装置。

【請求項7】

コンピュータに、

画像に対して部分領域を定義する領域パターン情報を参照するステップと、

前記領域パターン情報にしたがい前記部分領域から2つの特徴量を取得しそれらの相対値を算出する算出処理を実行する演算ステップであって、複数種類の前記算出処理を実行可能な演算ステップと、

前記相対値の取り得る範囲が複数の区間に区分されており、且つ、区間毎に判定値が対応付けられているルックアップテーブルを、前記算出処理毎に記憶する記憶手段を参照するステップと、

前記複数種類の算出処理で得られる複数の相対値の中から選択された1つ以上の相対値を用いて、前記画像が所定の属性を有するか否かを判定する第一の判定ステップと、

選択する相対値の組み合わせを変えて前記第一の判定ステップでの判定処理を行うことにより、複数の判定結果を取得するステップと、

前記取得される複数の判定結果に基づいて、前記画像が所定の属性を有するか否かについての最終的な判定を行う第二の判定ステップと、

を実行させるプログラムであって、

前記判定値は、予め用意された前記特定の被写体を含む画像である複数の正解画像から得られた相対値の度数分布と、予め用意された前記特定の被写体を含まない画像である複数の不正解画像から得られた相対値の度数分布とに基づいて決定される値であって、当該判定値に対応する区間内の相対値をもつ画像が前記特定の被写体を含む画像である確率が高いか否かを示す値であり、

前記第一の判定ステップでは、

前記選択された1つ以上の相対値毎に、その相対値を含む区間に対応付けられている判定値を、対応するルックアップテーブルから取得し、

該取得した判定値を元に得点を算出し、

該得点に基づいて前記画像が所定の属性を有するか否かを判定する
プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

10

20

30

40

50

本発明は、撮像された画像から、この画像中に含まれる人物や動物や物体などの特定の被写体又は被写体の一部を検出する装置や方法などに適用されて有効な技術に関する。

【背景技術】

【0002】

従来技術として、撮像された画像から、この画像中に含まれる人物や動物や物体などの特定の被写体又は被写体の一部を検出する技術がある。このような従来技術の一例として、撮像された画像から人物の顔を検出する技術がある（非特許文献1参照）。

【0003】

非特許文献1では、特定の矩形（以下「顔判定矩形」と呼ぶ）を処理対象となる画像内で移動させ、各移動先の顔判定矩形内（以下「注目領域」と呼ぶ）に人物の顔が含まれるか否かが判定される。図14は、顔判定矩形の例（顔判定矩形P1）を示す図である。図14を用いて、顔判定矩形P1を用いた人物の顔の検出処理について説明する。

10

【0004】

顔判定矩形P1は、その矩形内に複数の他の矩形（以下「第一矩形」、「第二矩形」と呼ぶ）P2、P3を含む。第一矩形P2と第二矩形P3とは、それぞれ顔判定矩形P1内の定位置に配置される。第一矩形P2と第二矩形P3とは、一つの顔判定矩形P1内にそれぞれ1以上配置される。

【0005】

顔の検出処理では、各注目領域における第一矩形P2及び第二矩形P3によって囲まれるそれぞれの領域（以下、それぞれ「第一特徴領域」、「第二特徴領域」と呼ぶ）の特徴量が算出される。第一特徴領域及び第二特徴領域の特徴量とは、例えば各領域内の画素値の平均値を示す。

20

【0006】

次に、第一特徴領域の特徴量 L_a と、第二特徴領域の特徴量 L_b との差分が算出される。そして、この差分値が、あらかじめ設定されている閾値よりも大きいか否かに応じて、注目領域に人物の顔が含まれるか否かが判定される。閾値は、サンプル画像を用いた学習によって求められる。

【0007】

実際の処理では、このような顔判定矩形が複数パターン用意され、それぞれについて判定が実行される。各パターンでは、第一矩形P2や第二矩形P3の数や位置がそれぞれ異なる。そして、それぞれの判定結果を元に、注目領域に人物の顔が含まれるか否かについて最終的な判定が行われる。

30

【0008】

また、このように、第一特徴領域と第二特徴領域との特徴量について差分を算出することにより人物の顔を検出する技術は、特許文献1でも用いられている。

【特許文献1】特開2000-123148

【非特許文献1】Paul Viola, Michael Jones, "Robust Real-time Object Detection", SECOND INTERNATIONAL WORKSHOP ON STATISTICAL AND COMPUTATIONAL THEORIES OF VISION - MODELING, LEARNING, COMPUTING, AND SAMPLING VANCOUVER CANADA, JULY 13, 2001.

40

【発明の開示】

【発明が解決しようとする課題】

【0009】

画像における人物の顔の検出は、上記のような顔判定矩形P1を用いた方法が適用されることにより、その精度が向上した。しかしながら、現在、画像中の人物の顔を検出する処理は、携帯電話機などのマシンスペックの比較的高くない装置上でリアルタイムで動作することが求められているため、処理速度の向上も要求されている。また、画像における人物の顔検出の精度の向上についても、引き続き要求がなされている。

【0010】

本発明では、このような問題を解決し、画像中のある注目領域内に人物の顔が含まれる

50

か否かの判定処理に関して、高速化や高精度化を実現する装置など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】

【0011】

〔第一の態様〕

上記問題を解決するため、本発明は以下のような構成をとる。本発明の第一の態様は、特定被写体検出装置であって、記憶手段、算出手段及び判定手段を備える。

【0012】

記憶手段は、複数の各特徴量に応じて用意された各判定値を記憶する。以下、判定値とは、判定手段における判定処理に用いられる値である。判定値とは、例えば、対応する特徴量が算出手段によって算出された場合に、注目領域に特定の被写体が含まれる可能性が高いか否かを示す値である。

10

【0013】

算出手段は、注目領域内の特徴量を算出する。以下、注目領域とは、特定の被写体が含まれるか否かが判定される領域である。即ち、注目領域とは、特定被写体検出装置の処理対象となる領域を示す。また、特徴量とは、注目領域に含まれる画素の状態から一意に決まる値である。例えば、注目領域内の全画素又は一部の画素の画素値の平均値、合計値、分散などである。

【0014】

判定手段は、記憶手段に記憶される判定値のうち、算出手段によって算出された特徴量に対応する判定値を元に、注目領域内に特定の被写体が含まれるか否かを判定する。例えば、算出された特徴量に対応する判定値が、注目領域に特定の被写体が含まれる可能性が高いことを示す値である場合、判定手段は注目領域内に特定の被写体が含まれると判定する。

20

【0015】

このように構成された本発明の第一の態様では、判定手段による判定処理に用いられる判定値は、特徴量毎に対応する値として記憶手段に記憶される。このため、従来のように一つの閾値によって特徴量と判定値との対応付けがなされていた場合に比べて、より正確に判定値と特徴量とを対応させることが可能となる。従って、判定手段は、与えられた各特徴量について、注目領域内に特定の被写体が含まれるか否かをより正確に判定することが可能となる。

30

【0016】

この第1の態様は、以下のように変形することもできる。すなわち、変形された第1の態様は、画像に対して部分領域を定義する領域パターン情報を参照する手段と、上記領域パターン情報にしたがい所定の演算を実行することにより上記画像の特徴量を算出する演算手段と、複数のサンプル画像に対して算出された特徴量とその特徴量が算出された画像が有する属性に係る判定値とを組み合わせる記憶手段と、上記画像に対して算出される特徴量により当該画像が上記属性を有するか否かを判定する判定手段とを備える。

【0017】

40

この変形された第1の態様は、複数のサンプル画像に対して算出された特徴量とその特徴量が算出された画像が有する属性に係る判定値とを組み合わせる記憶手段と、ある属性（特定被写体の存在等）を有するサンプル画像に対して算出された特徴量とその属性有りとの判定値が記憶される。一方、その属性（特定被写体の存在等）を有しないサンプル画像に対して算出された特徴量とその属性なしとの判定値が記憶される。

【0018】

複数のサンプル画像に対して、このような特徴量と判定値の組み合わせを事前に上記判定値記憶手段に記憶しておけばよい。複数のサンプル画像に対して、上記特徴量の度数分布を求めてもよい。そして、例えば、ある属性（特定被写体の存在等）を有するサンプル画像に対する度数が所定値以上となるような特徴量の範囲に対して、その属性ありとの判

50

定値を記憶してもよい。また、その属性（特定被写体の存在等）を有しないサンプル画像に対する度数が所定値以上となるような特徴量の範囲に対して、その属性なしとの判定値を記憶してもよい。

【0019】

〔第二の態様〕

本発明の第二の態様は、特定被写体検出装置であって、記憶手段、算出手段、第一の判定手段、制御手段及び第二の判定手段を備える。

【0020】

記憶手段は、複数の各特徴量に応じて用意された各判定値を記憶する。記憶手段に記憶される判定値は、二値（例えば“0”又は“1”）であっても良いし、実数として与えられても良い。判定値が実数として与えられた場合、二値として与えられる場合に比べて、第一の判定手段や第二の判定手段における判定処理の精度が向上する。

10

【0021】

算出手段は、同一注目領域内の特徴量を、複数の異なる算出処理によって算出する。この複数の異なる算出処理は、例えば算出される値の種類が異なる（例：平均、合計、分散）ような処理であっても良いし、算出処理を実行する上での入力異なる（例：注目領域内の異なる部分領域に関するデータが入力される）ような処理であっても良い。また、算出される値の種類及び算出処理の入力が異なるような処理であっても良い。

【0022】

第一の判定手段は、算出手段によって算出された特徴量に対応する判定値を元に得点を算出し、この得点に基づいて注目領域内に特定の被写体が含まれるか否かを判定する。例えば、算出された複数の特徴量に対応する複数の判定値を累計することにより得点を算出し、この得点が、注目領域内に特定の被写体が含まれる可能性が高いことを示す値である場合（例えば所定の閾値を超えている場合）、判定手段は注目領域内に特定の被写体が含まれると判定する。

20

【0023】

制御手段は、算出手段による異なる算出処理によって得られた複数の特徴量をセットとして第一の判定手段に対し与えることにより、最終的な判定を得るに十分な数の判定結果を第一の判定手段から取得する。従って、第一の判定手段は、制御手段から与えられるセットに含まれる複数の特徴量それぞれについて、対応する判定値を記憶手段から読み出し、これらの判定値から得点を算出し、注目領域内に特定の被写体が含まれるか否かを判定する。このため、第一の判定手段は、複数の判定結果を導く。各判定結果は同一である必要は無い。

30

【0024】

また、得られた判定結果の数が、最終的な判定を得るに十分な数であるか否かは、制御手段によって動的に判断されても良いし、あらかじめこの数が決められていても良い。例えば、学習アルゴリズム等の実行や管理者の経験の感覚などによって、最終的な判定を得るに十分な判定結果の数があらかじめ設定されていても良い。このとき、数のみならず、算出手段によって実行される算出処理もあらかじめ設定されていても良い。

【0025】

なお、最終的な判定、即ち第二の判定手段による判定は、第一の判定手段による判定結果の数が多ほど精度が向上する。このため、最終的な判定を得るに十分な数とは、最終的な判定が一定の精度を確保するために必要な数を示す。

40

【0026】

第二の判定手段は、制御手段によって取得される第一の判定手段による複数の判定結果に基づいて、注目領域内に特定の被写体が含まれるか否かについて最終的な判定を行う。

【0027】

このように構成された本発明の第二の態様では、第一の判定手段による判定処理に用いられる判定値は、特徴量毎に対応して記憶手段に記憶される。このため、従来のように一つの閾値によって特徴量と判定値との対応付けがなされていた場合に比べて、より正確に

50

判定値と特徴量とを対応付けることが可能となる。従って、判定値の精度が向上し、この判定値を用いた処理を行う第一の判定手段における処理結果の精度も向上する。言い換えれば、第一の判定手段は、与えられた各特徴量を元に、注目領域内に特定の被写体が含まれるか否かをより正確に判定することが可能となる。

【0028】

また、各判定値の精度が向上するため、第一の判定手段に対してセットとして与えられる特徴量の数を削減しても、即ち第一の判定手段において用いる判定値の数を削減しても、処理結果の精度を維持することが可能となる。その結果、第一の判定手段における処理結果の精度を維持しつつ、セットとして与えられる特徴量の数を削減することにより高速化を図ることが可能となる。

10

【0029】

また、第一の判定手段における判定結果の精度が向上することにより、第一の判定手段による判定結果の数を削減しても、最終的な判定の精度（第二の判定手段による判定の精度）を確保することが可能となる。即ち、最終的な判定を得るに十分な数が減少し、第一の判定手段による判定処理の回数が削減される。このため、注目領域内に特定の被写体が含まれるか否かの最終的な判定までに要する時間を短くすることが可能となる。即ち、特定の被写体を検出する処理を、精度を落とすことなく高速化することが可能となる。

【0030】

また、第二の態様における第一の判定手段は、新たなセットとして複数の特徴量を制御手段から受け取り新たな得点を算出する際に、当該新たなセットとしての複数の特徴量それぞれにおける複数の判定値、及び既に第一の判定手段による判定処理が完了しているセットにおいて算出された得点を用いるように構成されても良い。

20

【0031】

このように構成される場合、第一の判定手段において、セットに含まれる特徴量における判定値のみならず、他のセット（既に第一の判定手段による判定処理が完了しているセット）における判定値も、判定処理に影響する。このため、第一の判定手段における処理の精度が向上する。第一の判定手段における判定結果の精度が向上することにより、第一の判定手段による判定結果の数を削減しても、最終的な判定の精度を確保することが可能となる。即ち、最終的な判定を得るに十分な数が減少し、第一の判定手段による判定処理の回数が削減される。このため、注目領域内に特定の被写体が含まれるか否かの最終的な判定までに要する時間を短くすることが可能となる。即ち、特定の被写体を検出する処理を、精度を落とすことなく高速化することが可能となる。

30

【0032】

〔第三の態様〕

本発明の第三の態様は、特定被写体検出装置であって、記憶手段、算出手段、第一の判定手段、制御手段及び第二の判定手段を備える。

【0033】

記憶手段は、複数の異なるパターンそれぞれについて、複数の各特徴量に応じて用意された各判定値を記憶する。従って、パターンと特徴量とが得られれば、判定値は一意に決まる。

40

【0034】

算出手段は、画像から同一注目領域内の特徴量を、複数の異なるパターンそれぞれに基づいて算出する。

【0035】

第一の判定手段は、算出手段が用いたパターンと算出手段によって算出された特徴量とに対応する判定値を得る。そして、得られた判定値を元に得点を算出し、この得点に基づいて注目領域内に特定の被写体が含まれるか否かを判定する。

【0036】

制御手段は、第一の判定手段に対し複数の特徴量をセットとして与えることにより、最終的な判定を得るに十分な数の判定結果を第一の判定手段から取得する。複数の特徴量と

50

は、複数の異なるパターンに基づいた算出処理によって得られる。即ち、パターン毎の特徴量が算出手段によって算出される。そして、パターン毎に得られた特徴量のセットが第一の判定手段に与えられ、第一の判定手段による判定結果が取得される。

【0037】

第二の判定手段は、制御手段によって取得される第一の判定手段による複数の判定結果に基づいて、注目領域内に特定の被写体が含まれるか否かについて最終的な判定を行う。

【0038】

このように構成された本発明の第三の態様では、第一の判定手段による判定処理に用いられる判定値は、各パターンについて、特徴量毎に対応する値として記憶手段に記憶される。このため、従来のように、各パターンについて一つの閾値によって特徴量と判定値との対応付けがなされていた場合に比べて、より正確に判定値と特徴量とを対応付けることが可能となる。従って、各判定値の精度が向上し、この判定値を用いた処理を行う第一の判定手段における処理結果の精度も向上する。言い換えれば、第一の判定手段は、各パターン毎に与えられる特徴量を元に、注目領域内に特定の被写体が含まれるか否かをより正確に判定することが可能となる。

10

【0039】

また、各判定値の精度が向上するため、第一の判定手段に対してセットとして与えられる特徴量の数を削減しても、即ち第一の判定手段において用いる判定値の数を削減しても、処理結果の精度を維持することが可能となる。その結果、第一の判定手段における処理結果の精度を維持しつつ、セットとして与えられる特徴量の数を削減することにより高速化を図ることが可能となる。

20

【0040】

また、第一の判定手段における判定結果の精度が向上することにより、第一の判定手段による判定結果の数を削減しても、最終的な判定の精度（第二の判定手段による判定の精度）を確保することが可能となる。即ち、算出手段及び第一の判定手段において用いられるパターンの数を削減しても、最終的な判定の精度を確保することが可能となる。言い換えれば、最終的な判定を得るに十分な判断結果の数が減少し、第一の判定手段による判定処理の回数が削減される。このため、注目領域内に特定の被写体が含まれるか否かの最終的な判定までに要する時間を短くすることが可能となる。即ち、特定の被写体を検出する処理を、精度を落とすことなく高速化することが可能となる。

30

【0041】

また、本発明の第三の態様における記憶手段は、複数の区間に分割された特徴量と、各区間における判定値とを対応付けて記憶しても良い。

【0042】

また、本発明の第三の態様において、各区間における判定値は以下のような判断基準生成装置によって求められた値であっても良い。判断基準生成装置とは、サンプル画像特徴量算出手段、度数獲得手段及び判定値決定手段を備える。

【0043】

サンプル画像特徴量算出手段は、任意のパターンに基づいて複数のサンプル画像各々の特徴量を算出する。度数獲得手段は、複数の特徴量の各区間について、サンプル画像特徴量算出手段によって算出された特徴量が区間内に含まれるサンプル画像の度数を求める。判定値決定手段は、複数の各区間について、この区間における度数に基づいて、このパターンに基づいて算出された特徴量がこの区間に含まれる注目領域には特定の被写体が含まれると判定すべきか否かを判断することにより判定値を決定する。

40

【0044】

また、本発明の第三の態様において、サンプル画像は、第一の判定手段が判定を行う対象である特定の被写体を含む正解画像と、特定の被写体を含まない不正解画像とを含むように構成されても良い。

【0045】

また、本発明の第三の態様において、各区間における判定値は、正解画像及び不正解画

50

像それぞれの指標の相対的な値に基づいて設定されても良い。

【0046】

また、本発明の第三の態様において、各区間における判定値は、正解画像及び不正解画像それぞれの度数の相対的な値に基づいて設定されても良い。相対的な値とは、例えば比や差分などである。

【0047】

また、本発明の第三の態様におけるパターンは第一の特徴領域と第二の特徴領域とを有し、各特徴領域は、特定の領域内に位置及び大きさがパターン毎にそれぞれ固定されるように構成されても良い。

【0048】

また、本発明の第三の態様の算出手段は、注目領域内の第一の特徴領域における第一の特徴量と第二の特徴領域における第二の特徴量との相対的な値を算出することによって注目領域内の特徴量を算出するように構成されても良い。相対的な値とは、例えば比や差分などである。

【0049】

また、第三の態様における第一の判定手段は、新たなセットとして複数の特徴量を制御手段から受け取り新たな得点を算出する際に、当該新たなセットとしての複数の特徴量それぞれにおける複数の判定値、及び既に第一の判定手段による判定処理が完了しているセットにおいて算出された得点を用いるように構成されても良い。

【0050】

このように構成される場合、第二の態様がこのように構成される場合と同様の効果が奏される。

【0051】

〔第四の形態〕

本発明の第四の形態は、判断基準生成装置であって、算出手段、度数獲得手段、判断手段及び判断基準生成手段を含む。

【0052】

算出手段は、任意のパターンに基づいて複数のサンプル画像各々の特徴量を算出する。

【0053】

度数獲得手段は、特徴量が複数の区間に分割された各区間について、算出手段によって算出された特徴量が区間内に含まれるサンプル画像の度数を求める。度数とは、例えばサンプル画像の数や、この数にサンプル画像毎に設定された重みを乗じた値などを示す。

【0054】

判断手段は、特徴量の各区間における度数に基づいて、各区間における判定値を決定する。判断手段は、パターンに基づいて算出された特徴量が任意の区間に含まれる注目領域には、特定の被写体が含まれると判定すべきか否かを判断することにより、判定値を決定する。例えば、あるパターンに基づいて算出された特徴量がいずれの区間に該当する場合に、特定の被写体が注目領域に含まれる可能性が高いかについて判断し、判定値を決定する。

【0055】

判断基準生成手段は、判断手段の判断結果に基づいて、各区間と判定値とが対応付けられた判断基準を生成する。このような判断基準の具体的な例として、各区間と判定値とを対応付けたテーブルなどがある。

【0056】

このように構成された本発明の第四の形態では、判定値が特徴量の各区間毎に対応付けられたテーブルが生成される。このため、従来のように一つの閾値によって特徴量と判定値との対応付けがなされていた場合に比べて、より正確に判定値と特徴量とが対応付けられたテーブルを生成することが可能となる。従って、このテーブルを利用して特定の被写体を検出する処理が実行される場合、注目領域内に特定の被写体が含まれるか否かをより正確に判定することが可能となる。

10

20

30

40

50

【 0 0 5 7 】

〔その他〕

第一乃至第四の態様は、プログラムが情報処理装置によって実行されることによって実現されても良い。即ち、上記した第一乃至第四の態様における各手段が実行する処理を情報処理装置に対して実行させるためのプログラム、或いは当該プログラムを記録した記録媒体をもって上述した作用及び効果が得られても良い。また、第一乃至第四の態様における各手段が実行する処理を情報処理装置が実行する方法をもって上述した作用及び効果が得られても良い。

【 発明の効果 】

【 0 0 5 8 】

本発明によれば、判定手段による判定処理に用いられる判定値は、特徴量毎に対応付けて記憶手段に記憶される。このため、従来のように一つの閾値によって特徴量と判定値との対応付けがなされていた場合に比べて、より正確に判定値と特徴量とを対応付けることが可能となる。従って、判定手段は、与えられた各特徴量について、注目領域内に特定の被写体が含まれるか否かをより正確に判定することが可能となる。

【 0 0 5 9 】

また、最終的な判定の精度を確保するために複数の判定結果に基づいてさらに判定が行われる場合、判定値を用いた判定結果の精度が向上することにより、このような判定値を用いた判定結果の数を削減しても、最終的な判定の精度を確保することが可能となる。このため、注目領域内に特定の被写体が含まれるか否かの最終的な判定までに要する時間を短くすることが可能となる。即ち、特定の被写体を検出する処理を精度を落とすことなく高速化することが可能となる。

【 発明を実施するための最良の形態 】

【 0 0 6 0 】

次に、図を用いて本発明による特定被写体検出装置などの実施形態について説明する。以下の説明では、特定被写体検出装置の具体例として、人物画像から人物の顔を検出する顔検出装置 4 (4 a , 4 b を含む : 図 1 1 参照) について説明する。

【 0 0 6 1 】

この説明において、人物画像とは、少なくとも人物の顔の一部または全部の画像が含まれる画像である。従って、人物画像は、人物全体の画像を含んでも良いし、人物の顔だけや上半身だけの画像を含んでも良い。また、人物画像は、複数の人物についての画像を含んでも良い。さらに、人物画像は、背景に人物以外の風景 (被写体として注目された物も含む) や模様などのいかなるパターンを含んでも良い。

【 0 0 6 2 】

なお、顔検出装置 4 についての以下の説明は例示であり、その構成は以下の説明に限定されない。

【 0 0 6 3 】

〔顔検出の原理〕

まず、顔検出装置 4 において適用される顔検出技術の原理について説明する。顔検出装置 4 に適用される顔検出技術は、従来の顔検出技術に比べていくつかの変更点を有するが、ここでは従来の顔検出技術の原理についてはじめに説明する。この従来の顔検出技術では、あらかじめ、サンプル画像を用いた学習が実行され (以下「学習処理」と呼ぶ) 、この学習結果に基づいて顔の検出が実行される (以下「顔検出処理」と呼ぶ) 。

【 0 0 6 4 】

〔学習処理〕

まず、サンプル画像を用いた従来の学習処理について説明する。サンプル画像として、大きさの等しい複数の顔画像 (正解画像) 及び非顔画像 (不正解画像) が用意される。ここでは、サンプル画像として、縦横のそれぞれの画素数が等しい複数の矩形の画像が用いられる。顔画像とは、人間の顔を含む画像であり、人の顔の大きさに対応してフレーミング又はトリミングなどされた画像である。また、非顔画像とは、人間の顔を含まない画像

10

20

30

40

50

であり、例えば風景の画像や動物の画像を用いて構成される。顔検出装置 4 では、特定の被写体として人物の顔を検出するため、正解画像としてこのような顔画像が用意される。同様に、不正解画像としてこのような非顔画像が用意される。他の具体的な特定被写体検出装置においては、それぞれの装置で検出される特定の被写体を含む画像が正解画像として、特定の被写体を含まない画像が不正解画像として用意される。

【 0 0 6 5 】

学習処理では、サンプル画像と同じ大きさの領域を囲む矩形（以下「顔矩形」と呼ぶ）が用いられる。図 1 は顔矩形の例を示す図である。顔矩形 1 は、第一矩形 2 と第二矩形 3 とを含む。顔矩形 1 は、第一矩形 2 及び第二矩形 3 の数や配置される位置に応じて複数のパターンを有する（図 1 では（a）～（l）に相当）。即ち、各顔矩形 1 は、第一矩形 2 及び第二矩形 3 について固有の数や配置をパターンとして有する。以下、顔矩形 1 とサンプル画像とを用いた学習について説明する。

10

【 0 0 6 6 】

まず、あるパターンの顔矩形 1 について全てのサンプル画像を用いたデータ収集が実行される。このデータ収集では、まず、サンプル画像において、第一矩形 2 及び第二矩形 3 に対応する領域（以下、それぞれ「第一特徴領域」、「第二特徴領域」と呼ぶ）の特徴量（例えば領域内の画素値の平均値）が算出される。第一特徴領域及び / 又は第二特徴領域が一つの顔矩形 1 に複数含まれる場合、それぞれの特徴量として各領域の特徴量の合計値が算出される。例えば、図 1（j）の場合、第一特徴領域の特徴量は、二つの第一特徴領域の特徴量の和として算出される。そして、第一特徴領域の特徴量と第二特徴領域の特徴量との相対値（例えば比や差分値。ここでは相対値として差分値を想定する）として差分値が算出される。この差分値が、注目領域の特徴量を示す。

20

【 0 0 6 7 】

次に、算出された差分値（注目領域の特徴量）に基づいて、各パターンの顔矩形 1 に対応する閾値が求められる。この閾値は、確率的な手法を使って求められる。通常、このような確率的な手法は、単純な数学モデル（例えばガウス分布）を想定して設計される。例えば、差分値がゼロから各値までにおける顔画像と非顔画像についてそれぞれのサンプル数の合計値（積分値）を求め、その合計値の差が最大となる値が閾値として設定される。

【 0 0 6 8 】

上記の処理が、用意された全てのパターンの顔矩形 1 について実行され、全てのパターンの顔矩形 1 それぞれに対応する閾値が設定される。

30

【 0 0 6 9 】

次に、閾値が設定された複数の顔矩形 1 のうち、いずれのパターンの顔矩形 1 を顔検出処理において使用すべきかが判断される。なお、本特定被写体検出装置は、レイヤという処理単位ごとに顔の有無について判定を実行する。例えば、レイヤ 1 の判定で、大まかに顔の存在の可能性を判定し、可能性のない場合には処理を停止する。一方、顔の存在の可能性がある場合に、次のレイヤ 2 でさらに詳細な判定を実行する。

【 0 0 7 0 】

上記判断の際に、顔検出処理において顔の有無の判定を実行する複数のレイヤ（レイヤの具体例については顔検出処理の欄で説明する）それぞれに対し、各レイヤにおいて使用されるパターンの顔矩形 1 の割り当てが実行される。この処理は、Ada Boost 等のブースティング学習アルゴリズムによって実行される。

40

【 0 0 7 1 】

顔検出処理の実行の際に必要となるレイヤ数や、各レイヤに割り当てられる顔矩形 1 の数も、この判断の際に設計者によって決定される。このとき、顔検出処理において用いられる顔矩形 1 の数が多いほど処理の精度が増すため、設計者は実験や経験等に基づいて顔検出処理における最終的な判定を得るに十分な顔矩形 1 の数を決定する。設計者は、この数に基づいて、レイヤ数や各レイヤに割り当てられる顔矩形 1 の数を決定する。この数は、顔検出処理に求められる処理の速度や精度に応じて適宜決定される。

【 0 0 7 2 】

50

〔顔検出処理〕

次に、従来の顔検出処理について説明する。図2は、顔検出処理の流れを示す図である。まず、図2を用いて、顔検出処理のおおまかな流れについて説明する。

【0073】

顔検出処理は、複数のレイヤによって実行される。各レイヤには、異なる組み合わせの顔矩形1が割り当てられる。図2では、各レイヤに割り当てられる顔矩形1の数も異なる。また、各レイヤには判定を実施する順番が割り当てられており、各レイヤはその順番に従って処理を行う。即ち、例えば図2において、レイヤ1(Layer 1)の次にレイヤ2(Layer 2)が判定を実施し、その次にレイヤ3(Layer 3)が判定を実施する。

10

【0074】

各レイヤは、自身に割り当てられた順番に、自身に割り当てられたパターンの顔矩形1を用いて、注目領域内に人物の顔が含まれるか否かを判定する。あるレイヤにおいて、注目領域内に人物の顔が含まれないと判定された場合、それ以降の順番のレイヤでは、この注目領域についての判定は実施されない。そして、最後の順番のレイヤ(図2では、レイヤn(Layer n))による判定で、注目領域内に人物の顔が含まれると判定された場合に、最終的に顔検出処理において注目領域内に人物の顔が含まれると判定される。

【0075】

図3は、顔検出処理の流れを示すフローチャートである。次に、図3を用いて、顔検出処理の具体的な流れについて説明する。

20

【0076】

顔検出処理では、まず処理の対象となる注目領域が人物画像から選択される(S01)。基本的には、この注目領域は、人物画像の端から一定間隔で縦方向又は横方向へずれることによって順に選択される。例えば、人物画像をラスタスキャンすることによって注目領域が選択される。このとき、任意の人物画像について、複数の大きさの注目領域について選択が実行される。この選択法には、注目領域の大きさを固定して人物画像の大きさを变化させる方法や、人物画像の大きさを固定して注目領域の大きさを変更する方法などがある。図4は、注目領域の大きさを固定する場合の方法を示す図である。図5は、人物画像の大きさを固定する場合の方法を示す図である。注目領域の大きさが変化する場合、この注目領域の大きさの変化に従って顔矩形1、第一矩形2、及び第二矩形3の大きさも変化する。即ち、各レイヤにおいて使用される顔矩形1の大きさは、注目領域の大きさが変化する場合、その注目領域の大きさと同一又は略同一になるように制御され、顔矩形1の大きさに従って第一矩形2及び第二矩形3の大きさも変化する。

30

【0077】

次に、選択された注目領域について、注目領域に人物の顔が含まれるか否かの判定が実施される。この判定は、複数のレイヤ毎に実施される。このため、まず、順番に従って判定を実施するレイヤが選択される(S02)。

【0078】

次に、選択されたレイヤによって判定処理が実施される(S03)。このレイヤの判定において、注目領域に人物の顔が含まれると判定されなかった場合(S04-NO)、S07以降の処理が実行される。S07以降の処理については後述する。一方、注目領域に人物の顔が含まれると判定された場合(S04-YES)、直前の判定処理(S03の判定処理)が最後のレイヤによる処理であったか否かが判断される。最後のレイヤでなかった場合(S05-NO)、S02の処理に戻り次のレイヤが選択され、新たに選択されたレイヤによって判定処理が実施される。一方、最後のレイヤであった場合(S05-YES)、現在の注目領域に人物の顔が含まれるという最終的な判定がなされる(S06)。この時点で、顔検出装置4がこの注目領域に人物の顔が含まれると判断したこととなる。即ち、この時点ではじめて顔検出装置4が人物の顔を検出したこととなる。

40

【0079】

次に、判定の対象となっていた注目領域が、人物画像の中で最後の注目領域であったか

50

否かを判断する。最後の注目領域ではなかった場合（S07-NO）、S01の処理に戻り、次の注目領域が選択され、S02以降の処理が実行される。一方、最後の注目領域であった場合（S07-YES）、この人物画像に対する顔検出処理は終了する。

【0080】

図6は、各レイヤにおける判定の処理例を示す図である。次に、図6を用いて、レイヤ及び各レイヤにおける判定の処理内容について説明する。

【0081】

各レイヤには、1以上のパターンの顔矩形1が割り当てられる。この割り当ては学習処理において、AdaBoost等のブースティング学習アルゴリズムによって実施される。各レイヤは、自身に割り当てられたパターンの顔矩形1に基づいて、注目領域内に顔が含まれるか否かの判定を行う。

10

【0082】

各レイヤでは、各レイヤに割り当てられた各パターンの顔矩形1に基づいて、注目領域内の第一特徴領域及び第二特徴領域における特徴量が算出される。このとき、特徴量が各領域内の画素値の合計や平均などである場合、即ち特徴量が画素値の合計を用いて算出される値である場合、積分画像を用いて特徴量が算出されると良い。図7は、積分画像の例を示す図である。図7を用いて積分画像を用いた特徴量の算出処理について説明する。

【0083】

積分画像では、各画素は、元画像における各画素から左上全ての画素の画素値の合計をその画素値として持つ。例えば図7の画素aは、元画像における領域Aに含まれる全画素の画素値の合計を画素値として持つ。このため、例えば元画像における領域Dに含まれる全画素の画素値の合計（即ち領域Dの特徴量）は、dの画素値からb及びcの画素値を減算し、aの画素値を加算することによって算出される。

20

【0084】

次に、算出された特徴量の相対値としての差分値が算出され、この差分値に基づいて注目領域に人物の顔が含まれるか否かの判定が実施される。具体的には、算出された差分値が、判定に用いられているパターンの顔矩形1に設定された閾値よりも大きいか否か又は小さいか否かが判定される。そして、この判定の結果に応じて、注目領域における人物の顔の存否が判定される。

【0085】

ただし、この時点での判定は、各パターンの顔矩形1それぞれに基づいた判定であり、レイヤによる判定ではない。このように、各レイヤでは、割り当てられた全てのパターンの顔矩形1それぞれに基づいて個別に判定が実施され、それぞれの判定結果が得られる（図6の「顔矩形個別の判定」に相当）。

30

【0086】

次に、レイヤにおける得点が算出される。各パターンの顔矩形1には個別の得点（例：Pt1, Pt2, …, Ptn）が割り当てられている。注目領域に人物の顔が含まれると判定されると、そのとき使用されたパターンの顔矩形1に割り当てられた得点が参照され、当該レイヤの得点に加算される。このようにして、加算された得点の総計が、レイヤにおける得点（以下、このレイヤにおける得点の総計を、各パターンにおける得点と区別するため、「総得点」と呼ぶ）として算出される。そして、このレイヤにおける総得点が特定の閾値を越えた場合に、このレイヤは注目領域に人物の顔が含まれると判定する。一方、このレイヤにおける総得点が特定の閾値を越えない場合、このレイヤは注目領域に人物の顔が含まれないと判定する。

40

【0087】

S02～S06（図3参照）において、各レイヤによる判定は、処理の軽いレイヤ（例えば割り当てられた顔矩形1の数が少ないレイヤ）から順に実施される（図2参照）。また、各レイヤによる判定の前に、注目領域内の明るさの分散を算出し、算出された値に基づいて各レイヤによる判定を行うか否かを判断するように構成されても良い。このような構成においては、各レイヤによる判定を行わないと判断された場合、図3におけるS07

50

の処理が実行される。例えば、ほとんど明るさの変化のない注目領域（例えば真っ暗な注目領域や真っ白な注目領域など）については、各レイヤによる判定を行うまでもなく顔が含まれていないと認められるためである。

【 0 0 8 8 】

[第一実施形態]

[原理]

上記では、顔検出装置 4 において適用される顔検出技術のうち、従来の顔検出技術について説明した。次に、この顔検出技術において、顔検出装置 4 の第一実施形態である顔検出装置 4 a に適用されるにあたり変更された処理について説明する。なお、以下の説明では、変更された処理について説明する。即ち、以下において説明の無い処理については、上記で説明した顔検出技術と同様の処理が実行される。

10

【 0 0 8 9 】

従来の顔検出技術では、各パターンの顔矩形 1 における閾値の算出の際に単純な数学モデルが想定されていたため、差分値と顔画像及び非顔画像のサンプル数とのヒストグラムが実際にはどのような形状になるかについて想定されていなかった。例えば、図 6 の最上に例示される顔矩形 1 の場合、左右の目周囲における特徴量が第一特徴領域の特徴量として算出され、鼻や左右の頬周囲における特徴量が第二特徴領域の特徴量として算出される。

【 0 0 9 0 】

従来は、このような特徴領域の特徴量はそれぞれ単純な数学モデルに基づいて分布すると想定され、閾値が算出されていた。しかし、例えば上記の第一特徴領域の具体例の場合、実際には、左右の目が閉じている場合、片方の目が閉じている場合、左右の目が開いている場合など、特徴量が大きく変化するのであろう三つのケースが想定できる。また、例えば上記の第二特徴領域の具体例の場合、頬や鼻は顔の中でも凸型の部位であるため、凸の程度や肌の状態によっては光の反射が顕著な場合やそうでない場合など、特徴量が大きく変化するのであろう二つのケースが想定できる。そこで、顔検出装置 4 a では、このような想定に基づき、二つの特徴領域の特徴量の差分値は、ガウス分布のような単純な分布ではなく、複数のピークを有する分布になることを想定した。

20

【 0 0 9 1 】

図 8 は、学習処理におけるデータ収集によって算出された差分値のヒストグラムの例を示す図である。このヒストグラムは、ある一つのパターンの顔矩形 1 に関するものであり、各パターンの顔矩形 1 について同様にヒストグラムが形成される。

30

【 0 0 9 2 】

ヒストグラムの横軸は、第一特徴領域の特徴量と第二特徴領域の特徴量との差分値を示す。ヒストグラムの縦軸は、対応する差分値が算出されたサンプル画像の数（度数）を示す。正解分布とは顔画像のサンプル画像に関する分布を示し、不正解分布とは非顔画像のサンプル画像に関する分布を示す。

【 0 0 9 3 】

第一実施形態における学習処理では、ヒストグラムが形成されると、横軸が特定の間隔で区切られる。この間隔は、一定の幅であっても良いし、差分値に応じて異なる幅であっても良い。次に、各区間において、正解分布の値（顔画像の度数）と不正解分布の値（非顔画像の度数）とに基づき、判定値が求められる。図 9 は、形成されたヒストグラムに基づいて各区間の判定値が求められる様子を示す図である。判定値は、対応する差分値の区間に分布する画像が顔画像である確率が高いか否かを示す値である。例えば、判定値は、顔画像である確率が高い区間（図 9 において色の淡い区間）では“ 1 ”を有し、顔画像である確率が低い区間（図 9 において色の濃い区間）では“ 0 ”を有する。また、例えば判定値は、ある区間における正解分布の度数が不正解分布の度数よりも高い場合などに“ 1 ”を有し、ある区間における正解分布の度数が不正解分布の度数よりも低い場合などに“ 0 ”を有する。

40

【 0 0 9 4 】

50

そして、このヒストグラムに基づいて、LUT（ルックアップテーブル（Look Up Table））19aが作成される。図10は、LUT19aの例を示す図である。LUT19aは、差分値の各区間に対応する判定値を有する。変更後の顔検出処理では、閾値ではなく、このような学習処理によって作成されたLUT19aに基づいて、画像中の人物の顔が検出される。

【0095】

上記の処理が、用意された全てのパターンの顔矩形1について実行され、全てのパターンの顔矩形1それぞれに対応するLUT19aが作成される。

【0096】

次に、得られた複数のLUT19aのうち、いずれのLUT19aを顔検出処理において使用すべきかが判断される。即ち、いずれのパターンの顔矩形1を用いて顔検出処理を実行すべきかが判断される。この判断の際に、顔検出処理において顔の検出を実行する複数のレイヤそれぞれに対し、各レイヤにおいて使用されるパターンの顔矩形1の割り当てが実行される。この処理は、AdaBoost等のブースティング学習アルゴリズムによって実行される。

【0097】

顔検出装置4aの顔検出処理における各レイヤでは、各レイヤに割り当てられた各パターンの顔矩形1に基づいて、注目領域内の第一特徴領域及び第二特徴領域の特徴量が算出される。次に、算出された特徴量の差分値が算出され、この差分値に基づいて注目領域に人物の顔が含まれるか否かの判定が実施される。具体的には、算出された差分値に対応する判定値が、各パターンの顔矩形1に対応するLUT19aから求められ、この値によって判定される。例えば、図10に示されるLUT19aに対応するパターンの顔矩形1を用いた判定では、少なくとも、差分値が40以上60未満の場合、100以上120未満の場合、及び140以上160未満の場合には、注目領域内に人物の顔が含まれないと判定される。一方、少なくとも、差分値が60以上100未満の場合及び120以上140未満の場合には、注目領域内に人物の顔が含まれると判定される。

【0098】

このように、従来は単純な分布の想定により設定された閾値を用いて実施されていた判定処理が、顔検出装置4aでは複数のピークを有する分布の想定により設定されたLUT19aを用いて実施される。

【0099】

〔システム構成〕

顔検出装置

次に、従来に比べて上記のような変更点を有する顔判定技術が適用された顔検出装置4aの構成について説明する。顔検出装置4aは、ハードウェア的には、バスを介して接続されたCPU（中央演算処理装置）、主記憶装置（RAM）、補助記憶装置などを備える。補助記憶装置は、不揮発性記憶装置を用いて構成される。ここで言う不揮発性記憶装置とは、いわゆるROM（Read-Only Memory：EPROM（Erasable Programmable Read-Only Memory）、EEPROM（Electrically Erasable Programmable Read-Only Memory）、マスクROM等を含む）、FRAM（Ferroelectric RAM）、ハードディスク等を指す。

【0100】

図11は、顔検出装置4aの機能ブロックを示す図である。顔検出装置4aは、補助記憶装置に記憶された各種のプログラム（OS、アプリケーションプログラム等）が主記憶装置にロードされCPUにより実行されることによって、入力部5、出力部6、LUT記憶部7a、判定部8a、及び設定記憶部9等を含む装置として機能する。以下、図11を用いて、顔検出装置4aが備える各機能部について説明する。

【0101】

入力部

入力部5は、人物画像の原画像のデータ（以下、「原画像のデータ」と呼ぶ）を顔検出

10

20

30

40

50

装置 4 a へ入力するためのインタフェースとして機能する。原画像のデータは、静止画像のデータであっても良いし、動画のデータであっても良い。入力部 5 によって、顔検出装置 4 a の外部から、原画像のデータが顔検出装置 4 a へ入力される。入力部 5 は、顔検出装置 4 a へ原画像のデータを入力するためのどのような既存技術を用いて構成されても良い。

【 0 1 0 2 】

例えば、ネットワーク（例えばローカル・エリア・ネットワークやインターネット）を介して原画像のデータが顔検出装置 4 a へ入力されても良い。この場合、入力部 5 はネットワークインタフェースを用いて構成される。また、デジタルカメラやスキャナやパーソナルコンピュータや記録装置（例えばハードディスクドライブ）等から原画像のデータが顔検出装置 4 a へ入力されても良い。この場合、入力部 5 は、デジタルカメラやパーソナルコンピュータや記録装置などと顔検出装置 4 a とをデータ通信可能に接続する規格（例えば U S B (Universal Serial Bus) や S C S I (Small Computer System Interface) 等の有線接続や b l u e t o o t h 等の無線接続の規格）に応じて構成される。また、記録媒体（例えば各種フラッシュメモリやフロッピー（登録商標）ディスクや C D (Compact Disk) や D V D (Digital Versatile Disc, Digital Video Disc)）に記録された原画像のデータが顔検出装置 4 a へ入力されても良い。この場合、入力部 5 は、記録媒体からデータを読み出す装置（例えばフラッシュメモリリーダーやフロッピーディスクドライブや C D ドライブや D V D ドライブ）を用いて構成される。

【 0 1 0 3 】

また、顔検出装置 4 a がデジタルカメラ等の撮像装置又はデジタルカメラ等の撮像装置を備える各種装置（例えば P D A (Personal Digital Assistant) や携帯電話機）の内部に含まれ、撮像された人物画像が原画像のデータとして顔検出装置 4 a へ入力されても良い。この場合、入力部 5 は、C C D (Charge-Coupled Devices) や C M O S (Complementary Metal-Oxide Semiconductor) センサ等を用いて構成されても良いし、C C D や C M O S センサなどによって撮像された原画像のデータを顔検出装置 4 a に入力させるためのインタフェースとして構成されても良い。また、顔検出装置 4 a が、プリンタやディスプレイ等の画像出力装置の内部に含まれ、出力データとしてこの画像出力装置に入力された人物画像が原画像のデータとして顔検出装置 4 a へ入力されても良い。この場合、入力部 5 は、これらの画像出力装置に入力された原画像のデータを顔検出装置 4 a において取り扱い可能なデータに変換する装置などを用いて構成される。

【 0 1 0 4 】

また、入力部 5 は上記された複数の場合に応じることが可能となるように構成されても良い。

【 0 1 0 5 】

出力部

出力部 6 は、判定部 8 a によって人物の顔が検出されたか否かを示すデータ及び / 又は検出された顔の位置や大きさなどを示すデータを、顔検出装置 4 a の外部へ出力するためのインタフェースとして機能する。出力部 6 は、顔検出装置 4 a から、人物の顔の検出結果に関するデータを出力するためのどのような既存技術を用いて構成されても良い。

【 0 1 0 6 】

例えば、ネットワークを介して検出結果に関するデータが顔検出装置 4 a から出力されても良い。この場合、出力部 6 はネットワークインタフェースを用いて構成される。また、パーソナルコンピュータ等の他の情報処理装置や記録装置へ検出結果に関するデータが出力されても良い。この場合、出力部 6 は、パーソナルコンピュータ等の他の情報処理装置や記録装置と顔検出装置 4 a とをデータ通信可能に接続する規格に応じて構成される。また、記録媒体に対し検出結果に関するデータが出力（書き込み）されても良い。この場合、出力部 6 は、これらの記録装置又は記録媒体へデータを書き込む装置（例えばフラッシュメモリライターやフロッピーディスクドライブや C D - R ドライブや D V D - R ドライブ）を用いて構成される。

【 0 1 0 7 】

出力部 6 によって出力されるデータの具体的な用途例についても説明する。例えば、顔検出装置 4 a によって検出された顔の領域を示す図形をディスプレイ等の表示装置に出力するために、出力部 6 によって出力されるデータが使用されても良い。このような場合、出力部 6 は、例えばディスプレイ等の表示装置とデータ通信するためのインタフェースとして構成されても良いし、ディスプレイ等の表示装置に接続又は内蔵された情報処理装置にデータを渡すためのインタフェースとして構成されても良い。また、例えば、顔検出装置 4 a がデジタルカメラ又はデジタルカメラを備える各種装置の内部に含まれる場合、デジタルカメラは、出力部 6 によって出力されるデータを元に焦点制御や露出補正などの撮像に関する制御を行うように構成されても良い。このような場合、出力部 6 は、例えばデジタルカメラ内の情報処理装置とデータ通信可能なインタフェースとして構成されても良い。また、例えば顔検出装置 4 a が、画像補正処理を行う情報処理装置の内部に含まれる場合やこのような情報処理装置に接続される場合、情報処理装置は、出力部 6 によって出力されるデータを元に画像補正処理の処理領域や処理内容などを決定するように構成されても良い。このような場合、出力部 6 は、例えばこのような情報処理装置やその内部の装置とデータ通信可能なインタフェースとして構成されても良い。

10

【 0 1 0 8 】

また、出力部 6 は上記された複数の場合に応じることが可能となるように構成されても良い。

【 0 1 0 9 】

L U T 記憶部

L U T 記憶部 7 a は、不揮発性記憶装置を用いて構成される。L U T 記憶部 7 a は、判定部 8 a が顔検出処理を行う際に利用する L U T 1 9 a を記憶する。即ち、L U T 記憶部 7 a は、学習処理の結果として得られた各パターン顔矩形 1 に対する L U T 1 9 a を記憶する。このため、L U T 記憶部 7 a は、複数の L U T 1 9 a を記憶する可能性がある。

20

【 0 1 1 0 】

判定部

判定部 8 a は、L U T 記憶部 7 a に記憶された L U T 1 9 a を用いて、設定記憶部 9 に記憶された設定内容に基づいて顔検出処理を実行する。判定部 8 a は、顔検出処理の実行による結果を、出力部 6 へ渡す。なお、判定部 8 a は、不図示の入力部と出力部とによって、入力部 5 , 出力部 6 , L U T 記憶部 7 a 及び設定記憶部 9 へのデータの入出力を行う。

30

【 0 1 1 1 】

判定部 8 a は、顔検出プログラムが C P U によって実行されることにより実現される。また、判定部 8 a は、専用のチップとして構成されても良い。

【 0 1 1 2 】

図 1 2 は、判定部 8 a の内部の機能ブロックを示す図である。図 1 2 を用いて、判定部 8 a の機能ブロックについて説明する。判定部 8 a は、特徴量算出部 1 0 , 第一判定部 1 1 a , 制御部 1 2 , 及び第二判定部 1 3 a を含む。

【 0 1 1 3 】

特徴量算出部 1 0 は、各レイヤにおいて算出される、第一特徴領域及び第二特徴領域の特徴量を算出する。そして、特徴量算出部 1 0 は、算出された二つの特徴量の相対値（ここでは差分値）を算出する。

40

【 0 1 1 4 】

第一判定部 1 1 a は、特徴量算出部 1 0 によって算出された差分値に対応する判定値を取得し、1 以上の判定値に基づいて注目領域に人物の顔が含まれるか否かの判定を行う。具体的には、第一判定部 1 1 a は、特徴量算出部 1 0 によって特徴量の算出に利用されたパターンに対応する L U T 1 9 a を読み出す。第一判定部 1 1 a は、特徴量算出部 1 0 によって算出された差分値に対応する判定値を、読み出された L U T 1 9 a に基づいて取得する。第一判定部 1 1 a は、各レイヤに割り当てられた全パターンに対応する判定値を取

50

得し、それらの値に基づいて各レイヤにおける総得点を算出し、注目領域に人物の顔が含まれるか否かの判定を行う（図3のS03，S04に相当）。

【0115】

制御部12は、設定記憶部9に記憶される各種設定内容を読み出し、各パターンの第一矩形2及び第二矩形3の位置や大きさなどを特徴量算出部10に与える。また、制御部12は、特徴量算出部10によって算出された特徴量と、この特徴量の算出に利用されたパターンに対応するLUT19aとを第一判定部11aに与える。そして、第一判定部11aによる判定結果を第二判定部13aへ与える。また、制御部12は、その他にも注目領域の選択（図3のS01に相当）や、判定部8aの動作の制御（図3のS02，S05，S07などに相当）を行う。

10

【0116】

第二判定部13aは、第一判定部11aの結果に基づいて、即ち各レイヤによる判定結果に基づいて、現在の注目領域に顔が含まれるか否かの最終的な判定を行う（図3のS06に相当）。

【0117】

設定記憶部

設定記憶部9は、不揮発性記憶装置を用いて構成される。設定記憶部9は、判定部8aが顔検出処理を行う際の各種設定内容を記憶する。例えば、設定記憶部9は、各パターンの顔矩形1について記憶する。具体的には、各パターンの第一矩形2及び第二矩形3の位置や大きさなどについて記憶する。また、例えば設定記憶部9は、各レイヤにいずれのパターンの顔矩形1が割り当てられているかを記憶する。また、設定記憶部9は、人物画像内から注目領域を選択する方法などを記憶する。

20

【0118】

テーブル生成装置

次に、顔検出装置4aによって利用されるLUT19aを生成するテーブル生成装置14aの構成について説明する。テーブル生成装置14aは、ハードウェア的には、バスを介して接続されたCPU（中央演算処理装置）、主記憶装置（RAM）、補助記憶装置などを備える。補助記憶装置は、不揮発性記憶装置を用いて構成される。

【0119】

図13は、テーブル生成装置14aの機能ブロックを示す図である。テーブル生成装置14aは、補助記憶装置に記憶された各種のプログラム（OS，アプリケーション等）が主記憶装置にロードされCPUによって実行されることによって、特徴量算出部15，度数獲得部16，判断部17a，及びテーブル生成部18aを含む装置として機能する。

30

【0120】

テーブル生成装置14aは、全体として、変更後の学習処理を実行する。以下、図13を用いて、テーブル生成装置14aが備える各機能部について説明する。

【0121】

特徴量算出部15は、各サンプル画像について、パターンに従って第一特徴領域及び第二特徴領域の特徴量を算出する。そして、特徴量算出部15は、各特徴量の相対値として差分値を算出する。各パターンの特徴（各特徴領域の大きさや位置など）は、特徴量算出部15に記憶されても良いし、不図示の他の機能部によって記憶されても良い。

40

【0122】

度数獲得部16は、特徴量算出部15の算出結果に基づき、正解分布及び不正解分布を得る。そして、度数獲得部16は、正解分布及び不正解分布について、各区間の度数を獲得する。

【0123】

判断部17aは、度数獲得部16によって獲得される正解分布及び不正解分布の各区間の度数に基づいて、各区間の判定値を求める。

【0124】

テーブル生成部18aは、判断部17aによって求められた判定値とその区間とを対応

50

付けたLUT19aを生成する。また、テーブル生成部18aは、ブースティング学習アルゴリズムの実行により、いずれのLUT19aを顔検出装置4aにおいて使用すべきかを判断し、各レイヤへの割り当てを行う。

【0125】

〔作用/効果〕

顔検出装置4aによれば、顔検出処理の各レイヤの判定処理において、各パターンの顔矩形1に基づいた判定の際に、閾値ではなくLUT19aが使用される。このLUT19aには、各特徴領域における特徴量の差分値の範囲と、差分値の各範囲に対応する判定値とが保持され、この判定値を元に各パターンにおける判定が実施される。

【0126】

このため、顔検出装置4aでは、ある閾値に基づいて注目領域内における顔の存否を判定する場合と比べて、各パターンの顔矩形1を用いた判定において、より正確な判定を行うことが可能となる。例えば、図9に示されるようなヒストグラムが学習によって得られた場合、従来の技術では単純な分布が想定されていたため、例えば左から4つ目の区間と5つ目の区間との境界が閾値として設定されていた。即ち、小さいピークを有する分布(図9では右から2つ目の区間及び3つ目の区間)については具体的に想定されることはなかった。しかし、顔検出装置4aではLUT19aを用いることにより、このような小さいピークを有する分布についても独立した判定が行われる。従って、顔検出装置4aでは、各パターンの顔矩形1を用いた個々の判定が従来に比べてより正確になった。

【0127】

また、顔検出装置4aでは、各レイヤに割り当てられる顔矩形1のパターンが削減される。及び/又は、顔検出装置4aでは、顔検出処理において実行されるレイヤの数が削減される。即ち、一つの注目領域に対する顔検出処理において、判定が実施される顔矩形1のパターンの総数が削減される。

【0128】

そもそも、顔検出装置4aに適用される顔検出技術の顔検出処理において複数のパターンの顔矩形1に基づいた判定が実施される理由は、各パターンの顔矩形1に基づいた判定が個々では非常に不正確であるためである。即ち、個々では不正確な判定が実施されるため、多くのパターンの顔矩形1に基づいた判定によって判定処理の精度の向上が図られる。しかし、顔検出装置4aによれば、個々のパターンの顔矩形1に基づいた判定の精度が向上される。このため、一つの注目領域に対する顔検出処理において利用される顔矩形1のパターンの総数を削減することが可能となり、全体として顔検出処理の精度を落とすことなく高速化を図ることが可能となる。

【0129】

〔変形例〕

上記では、特定被写体検出装置の具体例として、画像中から人物の顔を検出する顔検出装置4aについて説明した。この他、特定被写体検出装置の具体例として、画像中から自動車の車体を検出する装置や、猫や犬などの特定の動物を検出する装置や、特定の文字や記号や標章などを検出する装置などがある。これらの装置は、学習処理に用いられるサンプル画像が異なるのみであり、基本的な構成などは顔検出装置4aと同様に実装することが可能である。また、検出されるそれぞれの特定被写体に応じて、顔検出装置4aに対し適宜変更が施されても良い。

【0130】

また、上記では、第一特徴領域の特徴量と第二特徴領域の特徴量との差分値が用いられたが、それぞれの特徴量の比など相対的な値が用いられても良い。

【0131】

〔第二実施形態〕

〔原理〕

次に、顔検出装置4の第二実施形態である顔検出装置4bにおいて適用される顔検出技術の原理について説明する。なお、以下の説明では、第一実施形態に適用された顔検出技

10

20

30

40

50

術と異なる点について説明する。

【 0 1 3 2 】

第一実施形態では、ヒストグラム（図 9 参照）の横軸が特定の間隔で区切られた後、各区間に対し“ 0 ”又は“ 1 ”のいずれかの判定値が与えられた。これに対し、第二実施形態では、各区間に対し実数の判定値が与えられる。図 1 5 は、第二実施形態においてヒストグラムの各区間に対し与えられる判定値の例を示す図である。第二実施形態において、判定値は、対応する差分値の区間に分布する画像が顔画像である確率の高さ又は確率そのものを示す。即ち、第一実施形態における判定値は「注目領域の画像が顔である可能性が高いか否か」を示したのに対し、第二実施形態における判定値は「注目領域の画像が顔である可能性の程度」を示す。例えば、判定値は“ 0 ”から“ 1 ”までの実数を取り、その値が大きいほど顔画像である確率が高いことを示す。より具体的には、判定値は例えば以下の数式に基づいて算出されても良い。数 1 において、判定値は $h(x)$ として算出される。

10

【 数 1 】

$$\text{If } f_{\text{Haar}}(x) \in \text{bin}_j \text{ then } h(x) = \frac{1}{2} \ln \left(\frac{\overline{W}_{+1}^j + \varepsilon}{\overline{W}_{-1}^j + \varepsilon} \right)$$

$$\text{where } \overline{W}_l^j = P(f_{\text{Haar}}(x) \in \text{bin}_j, y = l), l = \pm 1, j = 1, \dots, n.$$

20

$$\text{bin}_j = \left[\frac{j-1}{n}, \frac{j}{n} \right), j = 1, \dots, n$$

f_{Haar} is the Haar feature

【 0 1 3 3 】

また、この判定値は、正解分布の度数と不正解分布の度数との差や比に基づいて求められても良い。この場合、正解分布の度数が不正解分布の度数に対して高いほど判定値は大きい値を取り、逆に正解分布の度数が不正解分布の度数に対して低いほど判定値は小さい値をとる。

【 0 1 3 4 】

30

そして、このヒストグラムの各区間における判定値に基づいて L U T 1 9 b が作成される。図 1 6 は、L U T 1 9 b の例を示す図である。L U T 1 9 b は、差分値の各区間に対応する判定値を有し、その判定値は実数として表される。第二実施形態では、第一実施形態における L U T 1 9 a と同様に、全てのパターンの顔矩形 1 それぞれに対応する L U T 1 9 b が作成される。そして、ブースティング学習アルゴリズムによって、L U T 1 9 b が複数のレイヤそれぞれに対し割り当てられる。

【 0 1 3 5 】

第二実施形態における顔検出処理では、各レイヤ（最初に処理を行うレイヤを除く）は、第一実施形態における各レイヤと異なる処理を行う。図 1 7 は、第二実施形態における各レイヤの処理の概要を示す図である。まず、最初のレイヤ（L a y e r 1）は、第一実施形態における各レイヤと同じように、このレイヤに割り当てられた各パターンの判定値を取得する。そして、L a y e r 1 は、各パターンの判定値を元にレイヤの総得点を算出し、注目領域内における顔の存否を判断する。一方、L a y e r 2 以降のレイヤは、各レイヤに割り当てられた各パターンの顔矩形 1 に基づいて得られる判定値、及び一つ前のレイヤにおいて算出された総得点に基づいて、注目領域内における顔の存否を判断する。即ち、第二実施形態における各レイヤは、一つ前のレイヤにおける総得点を考慮して総得点を算出する点で、第一実施形態における各レイヤと異なる。なお、第二実施形態における各レイヤは、各パターンにおける判定値をそのまま各パターンにおける得点として考慮する。ただし、各パターンにおける判定値を元に得られる他の値を各パターンの得点として扱うように構成されても良い。

40

50

【0136】

図18は、第二実施形態における各レイヤの処理の具体例を示す図である。レイヤ m （レイヤ m は最初に処理を行うレイヤではない）は、このレイヤ m に割り当てられた各パターンに基づいて特徴量を計算する。次に、レイヤ m は、各パターン毎に、LUT19b及び算出された特徴量に基づいて判定値（ $p_{t2} \sim p_{tn}$ ）を取得する。さらに、レイヤ m は、判定値 p_{t1} として、一つ前のレイヤ（レイヤ（ $m-1$ ））における総得点を取得する。

【0137】

また、第一実施形態では、各レイヤにおける総得点を算出する際に、判定値が“1”となったパターンにおける得点のみが考慮されていたが、第二実施形態における各レイヤは全てのパターンにおける実数の判定値を考慮し総得点を算出する。従って、レイヤ m は、各判定値（ $p_{t1} \sim p_{tn}$ ）の全てに基づいて総得点を算出しレイヤ m としての判定を実施する。レイヤ m は、注目領域内に顔が含まれると判定した場合、レイヤ m における総得点を次のレイヤ（レイヤ（ $m+1$ ））に渡す。そして、最後のレイヤにおいて、最終的に注目領域内における顔の存否の判定が下される。

【0138】

〔システム構成〕

顔検出装置

次に、第二実施形態における顔検出装置4bの構成について説明する。顔検出装置4bは、LUT記憶部7a、判定部8aに代えて、LUT記憶部7b、判定部8bを備える点で顔検出装置4aと異なる。以下、顔検出装置4bについて、顔検出装置4aと異なる点について説明する。

【0139】

LUT記憶部

LUT記憶部7bは、LUT19a（図10参照）に代えてLUT19b（図16参照）を記憶する点でLUT記憶部7aと異なる。LUT記憶部7bは、他の点については、LUT記憶部7aと同様の構成をとる。

【0140】

判定部

判定部8bは、LUT記憶部7bに記憶されたLUT19bを用いて、設定記憶部9に記憶された設定内容に基づいて顔検出処理を実行する。以下、図12を用いて判定部8bの機能ブロックについて説明する。判定部8bは、第一判定部11aに代えて第一判定部11bを備える点、第二判定部13aに代えて第二判定部13bを備える点で判定部8aと異なる。以下、判定部8bについて、判定部8aと異なる点について説明する。

【0141】

第一判定部11bは、特徴量算出部10によって算出された差分値に対応する判定値を取得し、1以上の判定値に基づいて注目領域に人物の顔が含まれるか否かの判定を行う。具体的には、第一判定部11bは、特徴量算出部10によって特徴量の算出に利用されたパターンに対応するLUT19bを読み出す。第一判定部11bは、特徴量算出部10によって算出された差分値に対応する実数の判定値を、即ち各パターンにおける判定値を、読み出されたLUT19bに基づいて取得する。第一判定部11bは、これらの判定値に基づいて、各レイヤにおける総得点を算出し、注目領域に人物の顔が含まれるか否かの判定を行う。

【0142】

第一判定部11bは、二つ目以降のレイヤにおいては、判定値の一つとして、一つ前のレイヤにおける総得点に基づいた値を用いる。即ち、第一判定部11bは、二つ目以降のレイヤにおいては、一つ前のレイヤにおける総得点に基づいた値、及びこのレイヤに割り当てられた各パターンに対応する全ての判定値を用いて、このレイヤにおける総得点を算出する。そして、算出されたこの総得点に従って、現在の注目領域内に顔が含まれるか否かの判断について、このレイヤにおける判定を行う。

【0143】

第二判定部13bは、第一判定部11bの結果に基づいて、即ち各レイヤによる判定結果に基づいて、現在の注目領域に顔が含まれるか否かの最終的な判定を行う（図3のS06に相当）。

【0144】

テーブル生成装置

次に、顔検出装置4bによって利用されるLUT19bを生成するテーブル生成装置14bの構成について説明する。テーブル生成装置14bは、第二実施形態における学習処理を行う点でテーブル生成装置14aと異なる。即ち、テーブル生成装置14bは、判断部17aとテーブル生成部18aに代えて、判断部17bとテーブル生成部18bを備える点で、テーブル生成装置4aと異なる。以下、テーブル生成装置14bについて、テーブル生成装置14aと異なる点についてのみ説明する。

10

【0145】

判断部17bは、度数獲得部16によって獲得される正解分布及び不正解分布の各区間の度数に基づいて、数1の式に従い、各区間の実数の判定値を算出する。

【0146】

テーブル生成部18bは、判断部17bによって算出された実数の判定値とその区間とを対応付けたLUT19bを生成する。また、テーブル生成部18bは、ブースティング学習アルゴリズムの実行により、いずれのLUT19bを顔検出装置4bにおいて使用すべきかを判断し、各レイヤへの割り当てを行う。

20

【0147】

〔作用/効果〕

第二実施形態における顔検出装置4bによれば、顔検出処理の各レイヤの判定処理において、各パターンの顔矩形1に基づいた判定の際に、LUT19aではなくLUT19b（図16参照）が使用される。このLUT19bには、差分値の各範囲に対応する判定値として“0”又は“1”の二値ではなく、“0”～“1”の実数の値が保持される。

【0148】

このため、顔検出装置4bでは、LUT19aを用いた処理を行う顔検出装置4aに比べて、各レイヤにおける処理の精度が向上する。例えば、LUT19aの場合、正解分布と不正解分布との度数の違いがほとんどみられずわずかの差で判定値が“0”と判断された場合（以下、ケース1と呼ぶ）も、明らかに不正解分布の度数が圧倒的に多いために判定値が“0”と判断された場合（以下、ケース2と呼ぶ）も、全く同等に扱われてしまっていた。一方、LUT19bの場合、ケース1の場合は例えば判定値が“0.4”とされ、ケース2の場合は例えば判定値が“0.1”とされ、区別して考慮される。このため、ケース1が多い場合とケース2が多い場合とは、異なる状態（異なる得点）として考慮することが可能となり、顔を検出する際の精度向上を図ることができる。

30

【0149】

また、このように各パターンにおける判定値が実数となり精度が向上されるため、処理の精度を維持しつつ各レイヤに割り当てられるパターンの数を削減することが可能となる。即ち、これまでよりも少ないパターンによって判定処理をすることが可能となる。このため、処理の高速化を実現することが可能となる。同様の理由により、レイヤの数を削減し高速化を図ることも可能となる。

40

【0150】

また、第二実施形態における顔検出装置4bによれば、既に判定処理が完了したレイヤにおける総得点を、判定処理が完了していないレイヤにおいて用いることによりこのレイヤの判定処理が実施される。言い換えれば、既に判定処理が完了したレイヤにおける各パターンの判定値が、後のレイヤにおける判定処理に反映される。従って、後のレイヤにおいて、実際に用いられるパターンの数に比べて、判定処理に影響を及ぼすパターンの数が仮想的に増加したこととなり、このような処理を行わない顔検出装置4aに比べて各レイヤにおける判定処理の精度を向上させることができる。このため、後のレイヤにおいて、

50

判定処理の精度を維持しつつ、各レイヤに割り当てられるパターン数を削減し処理の高速化を実現することが可能となる。同様に、レイヤの数を削減し処理の高速化を実現することも可能となる。また、パターン数が削減されるため、顔検出装置 4 b において利用されるリソースを削減することが可能となる。

【 0 1 5 1 】

〔 変形例 〕

図 1 5 に示されたヒストグラム及び図 1 6 に示された L U T 1 9 b の例では、判定値は有効桁数 1 の小数点第一位までの小数として判定値が表されているが、判定値はこのような基準に拘束される必要はない。即ち、判定値の有効桁数や小数点第何位までの値として判定値が表されるかは、設計者が都合に応じて自由に設定して良い。

10

【 0 1 5 2 】

第二判定部 1 3 b は、各レイヤに割り当てられている各パターンに対応する判定値全てを用いて各レイヤにおける総得点を算出するのではなく、各レイヤに割り当てられている各パターンに対応する判定値のうち、閾値（例えば “ 0 . 2 ” , “ 0 . 5 ” ）を超えた判定値のみを用いて各レイヤにおける総得点を算出するように構成されても良い。

【 0 1 5 3 】

また、第二判定部 1 3 b は、各レイヤにおける総得点を算出する際に、一つ前のレイヤの総得点に限らず、そのレイヤよりも以前に処理が実行された 1 以上のレイヤにおける総得点に基づいた値を用いるように構成されても良い。

【 0 1 5 4 】

また、第二判定部 1 3 b は、以前に処理が実行された 1 以上のレイヤにおける総得点を用いて判定値を得る際に、総得点をそのまま判定値と使用しても良いし、総得点になんらかの重み付けを行うことにより判定値として使用しても良い。

20

【 図面の簡単な説明 】

【 0 1 5 5 】

【 図 1 】 顔矩形のパターンの例を示す図である。

【 図 2 】 顔検出処理の流れを示す図である。

【 図 3 】 顔検出処理のフローチャートである。

【 図 4 】 注目領域の大きさを固定した場合の注目領域の選択方法を示す図である。

【 図 5 】 人物画像の大きさを固定した場合の注目領域の選択方法を示す図である。

30

【 図 6 】 第一実施形態における各レイヤの処理の例を示す図である。

【 図 7 】 積分画像の例を示す図である。

【 図 8 】 差分値と画像数とのヒストグラムの例を示す図である。

【 図 9 】 第一実施形態においてヒストグラムの各区間に対し与えられる判定値の例を示す図である。

【 図 1 0 】 第一実施形態における L U T の例を示す図である。

【 図 1 1 】 顔検出装置の構成例を示す機能ブロック図である。

【 図 1 2 】 判定部の構成例を示す機能ブロック図である。

【 図 1 3 】 テーブル生成装置の構成例を示す機能ブロック図である。

【 図 1 4 】 顔判定矩形の例を示す図である。

40

【 図 1 5 】 第二実施形態においてヒストグラムの各区間に対し与えられる判定値の例を示す図である。

【 図 1 6 】 第二実施形態における L U T の例を示す図である。

【 図 1 7 】 第二実施形態における各レイヤの処理の概要を示す図である。

【 図 1 8 】 第二実施形態における各レイヤの処理の具体例を示す図である。

【 符号の説明 】

【 0 1 5 6 】

- 1 顔矩形
- 2 第一矩形
- 3 第二矩形

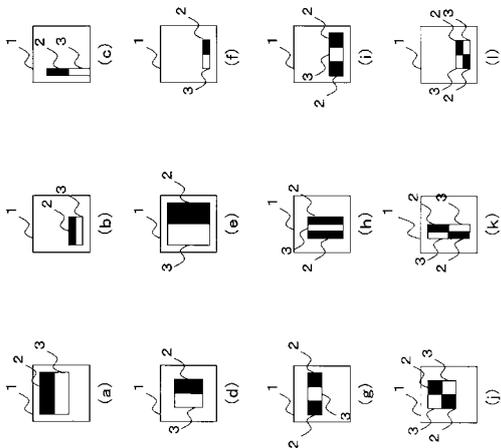
50

- 4 , 4 a , 4 b 顔検出装置
- 5 入力部
- 6 出力部
- 7 a , 7 b L U T 記憶部
- 8 a , 8 b 判定部
- 9 設定記憶部
- 10 特徴量算出部
- 11 a , 11 b 第一判定部
- 12 制御部
- 13 a , 13 b 第二判定部
- 14 a , 14 b テーブル生成装置
- 15 特徴量算出部
- 16 度数獲得部
- 17 a , 17 b 判断部
- 18 a , 18 b テーブル生成部
- 19 a , 19 b L U T
- P 1 顔判定矩形
- P 2 第一矩形
- P 3 第二矩形

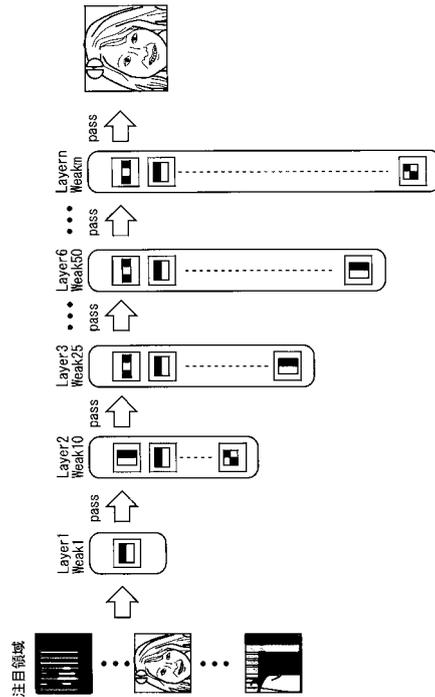
10

20

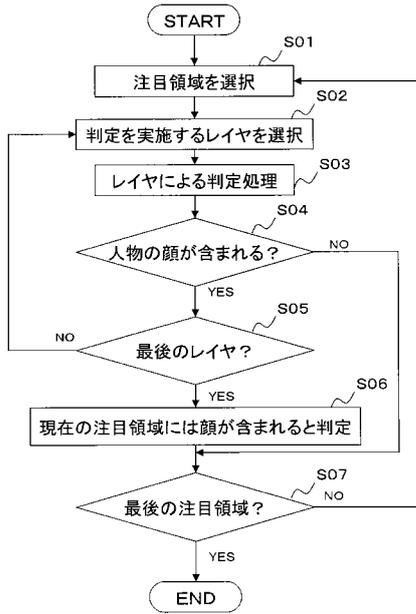
【 図 1 】



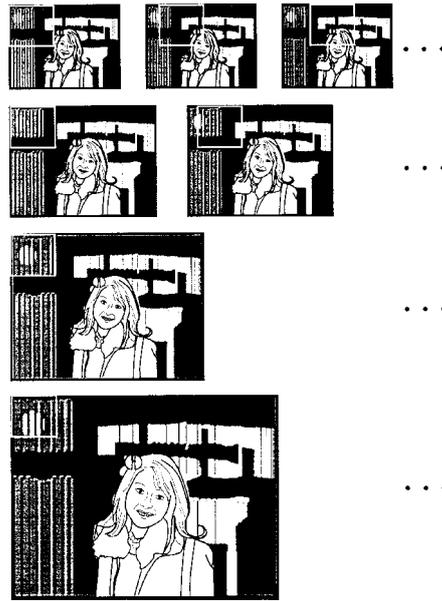
【 図 2 】



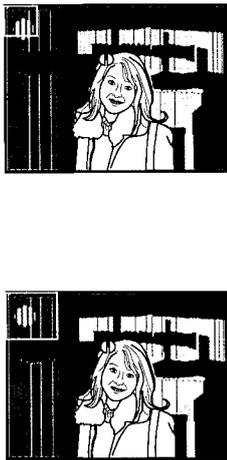
【図3】



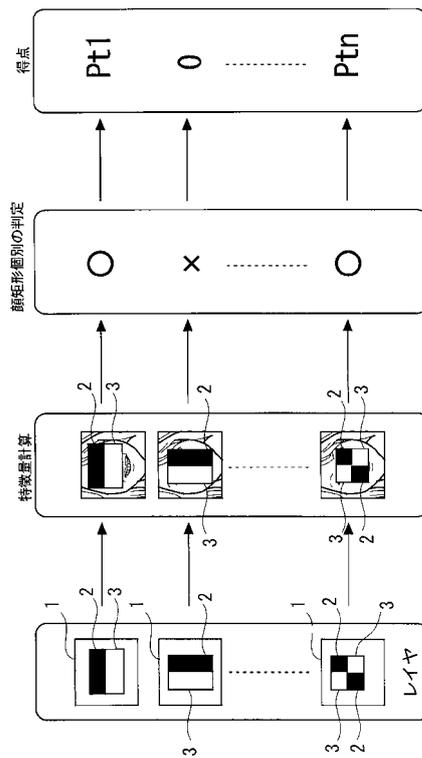
【図4】



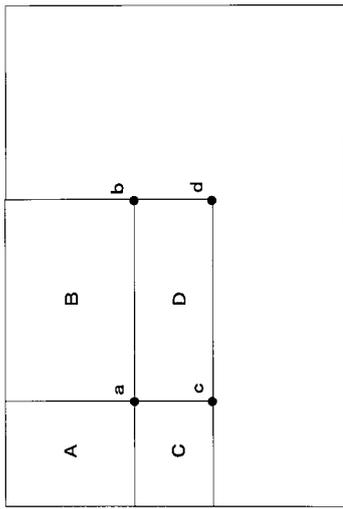
【図5】



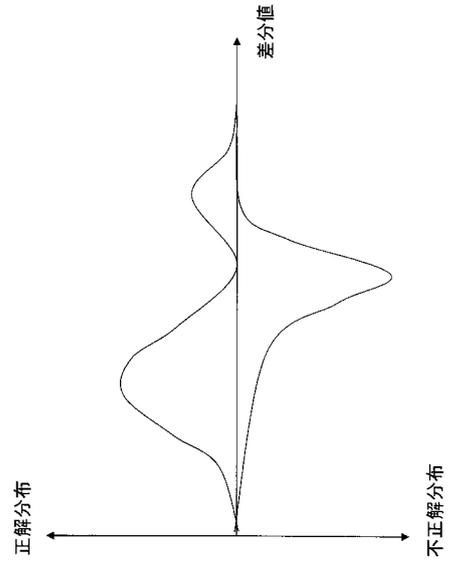
【図6】



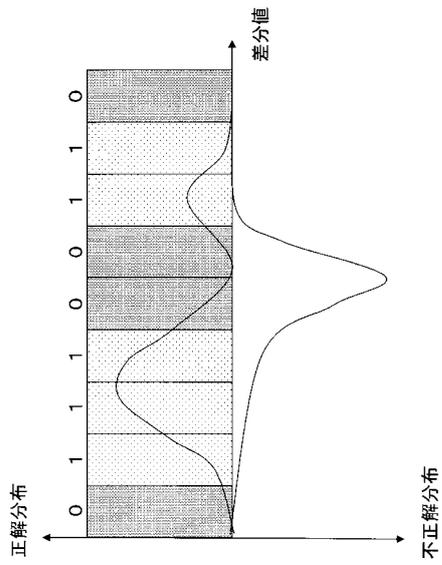
【 図 7 】



【 図 8 】



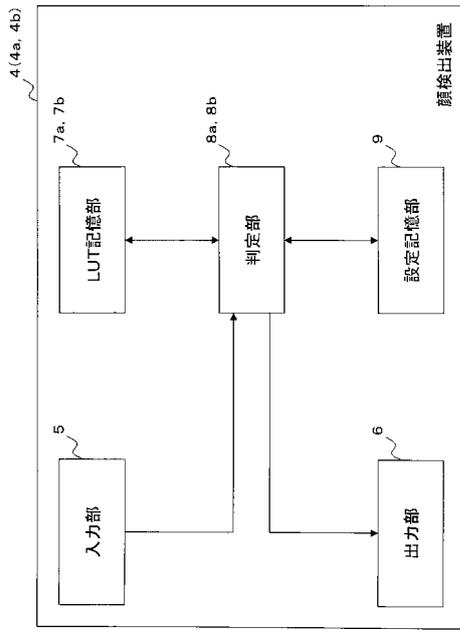
【 図 9 】



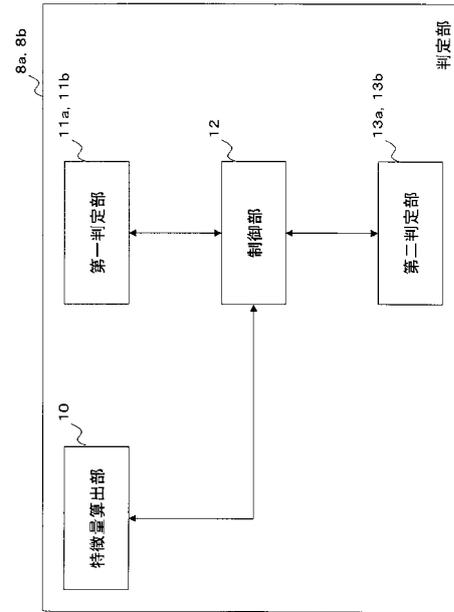
【 図 10 】

差分値	...	40 ~60	60 ~80	80 ~100	100 ~120	120 ~140	140 ~160	...
判定値	...	0	1	1	0	1	0	...

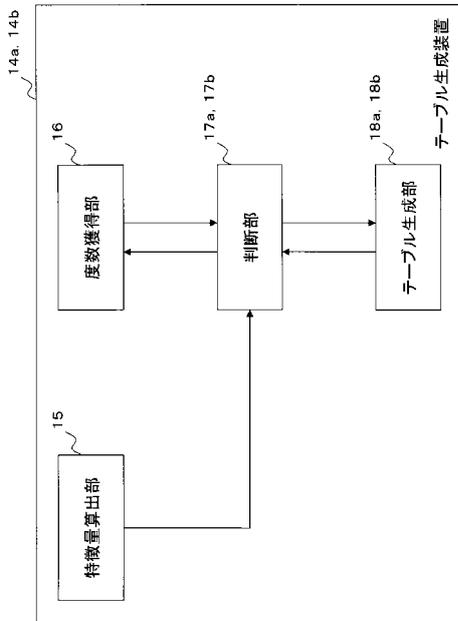
【図 1 1】



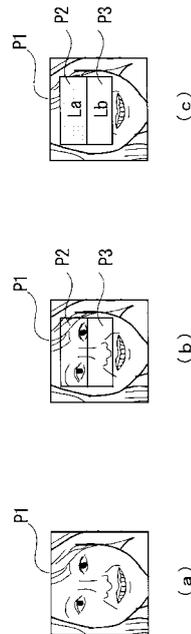
【図 1 2】



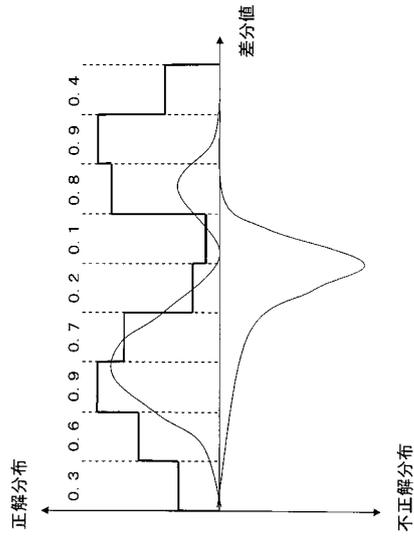
【図 1 3】



【図 1 4】



【 図 15 】

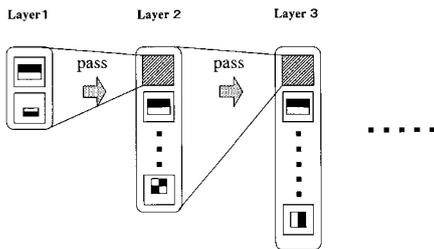


【 図 16 】

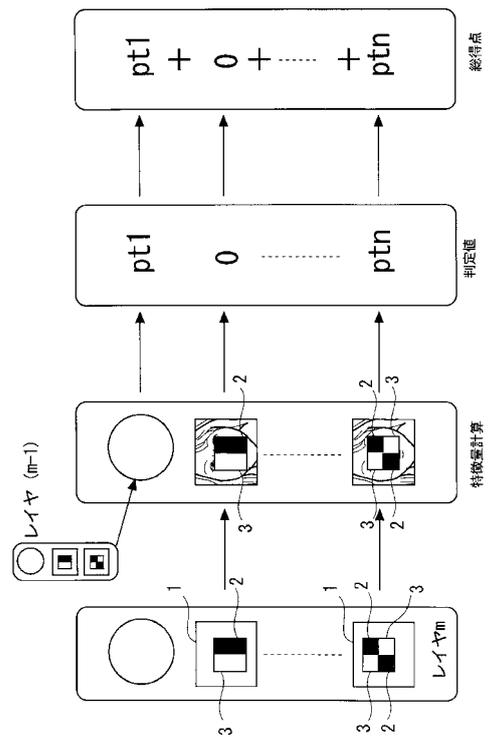
差分値	...	40	60	80	100	120	140	...
判定値	...	~60	~80	~100	~120	~140	~160	...
		0.6	0.9	0.7	0.2	0.1	0.8	

19b

【 図 17 】



【 図 18 】



フロントページの続き

- (72)発明者 艾 海舟
中華人民共和国北京市海淀區清華園 清華大学 内
- (72)発明者 黄 暢
中華人民共和国北京市海淀區清華園 清華大学 内
- (72)発明者 武 勃
中華人民共和国北京市海淀區清華園 清華大学 内
- (72)発明者 勞 世紅
京都府京都市下京区塩小路通堀川東入南不動堂町 8 0 1 番地 オムロン株式会社 内

審査官 佐藤 実

- (56)参考文献 特開 2 0 0 2 - 2 0 3 2 4 0 (J P , A)
Paul Viola, Michael Jones, Robust Real-time Object Detection, SECOND INTERNATIONAL WORKSHOP ON STATISTICAL AND COMPUTATIONAL THEORIES OF VISION - MODELING, LEARNING, カナダ,
2 0 0 1 年 7 月 1 3 日

- (58)調査した分野(Int.Cl., DB名)
G 0 6 T 7 / 0 0 ~ 7 / 6 0
G 0 6 T 1 / 0 0