

(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2007-25900
(P2007-25900A)

(43) 公開日 平成19年2月1日(2007.2.1)

(51) Int. Cl.	F I	テーマコード (参考)
G06T 1/00 (2006.01)	G06T 1/00 340A	5B057
G06T 7/00 (2006.01)	G06T 7/00 350C	5C122
H04N 5/232 (2006.01)	H04N 5/232 Z	5L096
H04N 101/00 (2006.01)	H04N 101:00	

審査請求 未請求 請求項の数 8 O L (全 15 頁)

(21) 出願番号	特願2005-204735 (P2005-204735)	(71) 出願人	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成17年7月13日 (2005.7.13)	(74) 代理人	100076428 弁理士 大塚 康德
		(74) 代理人	100112508 弁理士 高柳 司郎
		(74) 代理人	100115071 弁理士 大塚 康弘
		(74) 代理人	100116894 弁理士 木村 秀二
		(72) 発明者	矢野 光太郎 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

最終頁に続く

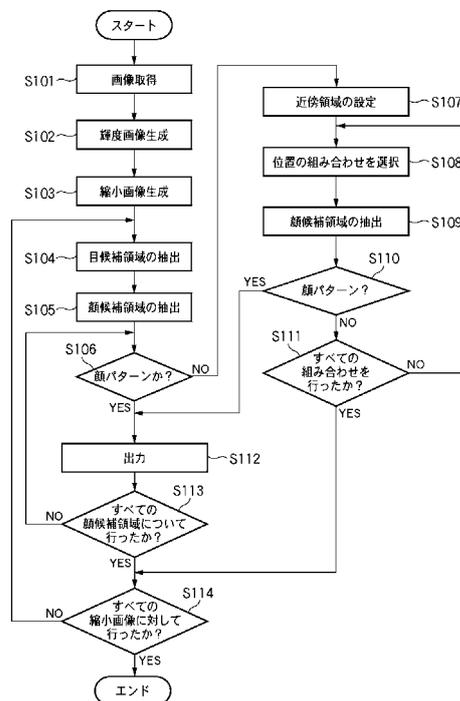
(54) 【発明の名称】 画像処理装置、画像処理方法

(57) 【要約】

【課題】 画像中の被写体を高速、且つ高精度に検出するための技術を提供すること。

【解決手段】 輝度画像の縮小画像内における目候補領域の位置を検出し (S104)、検出した位置のうち2つを包含し、且つこの2つの位置で規定される顔候補領域が顔パターンではない場合には、この顔候補領域内のそれぞれの目候補領域の位置近傍に近傍領域を設定し (S107)、一方の近傍領域内で取りうる第1の位置と、他方の近傍領域内で取りうる第2の位置とを包含し、且つ第1, 2の位置で規定される顔候補領域が顔パターンであるのかを判別する (S110)。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

画像から所定の被写体を検出する画像処理方法であって、
前記画像の輝度成分で構成される輝度画像を生成する生成工程と、
前記輝度画像を再帰的に縮小することで、複数枚の縮小画像を生成する縮小工程と、
縮小画像内における前記所定の被写体の部分的な特徴の候補としての特徴候補領域の位置を検出する第 1 の検出工程と、

前記第 1 の検出工程で検出した位置のうち 2 つを包含し、且つ当該 2 つの位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する第 1 の判別工程と

10

、
前記第 1 の判別工程で前記所定の被写体の領域ではないと判別された被写体候補領域内のそれぞれの特徴候補領域の位置近傍に近傍領域を設定する設定工程と、

一方の近傍領域内で取りうる第 1 の位置と、他方の近傍領域内で取りうる第 2 の位置とを包含し、且つ当該第 1 , 2 の位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する第 2 の判別工程と、

前記第 1 の判別工程、前記第 2 の判別工程のそれぞれで、前記所定の被写体の領域であると判断された被写体候補領域のうち一部若しくは全部のデータを出力する出力工程とを備えることを特徴とする画像処理方法。

【請求項 2】

前記第 2 の判別工程では、

20

それぞれの近傍領域を複数の矩形に分割する工程と、

一方の近傍領域内の矩形位置と、他方の近傍領域内の矩形位置との組み合わせを順次選択する工程と、

選択された組み合わせにおける一方の矩形位置と、他方の矩形位置とを包含し、且つ当該 2 つの矩形位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する工程と

を備えることを特徴とする請求項 1 に記載の画像処理方法。

【請求項 3】

前記出力工程では、前記第 1 の判別工程、前記第 2 の判別工程のそれぞれで、前記所定の被写体の領域であると判別された被写体候補領域のうち、当該判別の際に求めた値が最も大きい被写体候補領域のデータを出力することを特徴とする請求項 1 又は 2 に記載の画像処理方法。

30

【請求項 4】

前記所定の被写体は人の顔であることを特徴とする請求項 1 乃至 3 の何れか 1 項に記載の画像処理方法。

【請求項 5】

前記所定の被写体の部分的な特徴は目であることを特徴とする請求項 1 乃至 4 の何れか 1 項に記載の画像処理方法。

【請求項 6】

画像から所定の被写体を検出する画像処理装置であって、
前記画像の輝度成分で構成される輝度画像を生成する生成手段と、
前記輝度画像を再帰的に縮小することで、複数枚の縮小画像を生成する縮小手段と、
縮小画像内における前記所定の被写体の部分的な特徴の候補としての特徴候補領域の位置を検出する第 1 の検出手段と、

40

前記第 1 の検出手段が検出した位置のうち 2 つを包含し、且つ当該 2 つの位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する第 1 の判別手段と

、
前記第 1 の判別手段によって前記所定の被写体の領域ではないと判別された被写体候補領域内のそれぞれの特徴候補領域の位置近傍に近傍領域を設定する設定手段と、

一方の近傍領域内で取りうる第 1 の位置と、他方の近傍領域内で取りうる第 2 の位置と

50

を包含し、且つ当該第 1, 2 の位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する第 2 の判別手段と、

前記第 1 の判別手段、前記第 2 の判別手段のそれぞれで、前記所定の被写体の領域であると判断された被写体候補領域のうち一部若しくは全部のデータを出力する出力手段とを備えることを特徴とする画像処理装置。

【請求項 7】

コンピュータに請求項 1 乃至 5 の何れか 1 項に記載の画像処理方法を実行させることを特徴とするプログラム。

【請求項 8】

請求項 7 に記載のプログラムを格納したことを特徴とする、コンピュータ読み取り可能な記憶媒体。 10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像中における顔を検出するための技術に関するものである。

【背景技術】

【0002】

画像から特定の被写体パターンを自動的に検出する画像処理方法は非常に有用であり、例えば人間の顔の判定に利用することができる。このような方法は、通信会議、マン・マシン・インタフェース、セキュリティ、人間の顔を追跡するためのモニタ・システム、画像圧縮などの多くの分野で使用することができる。このような画像中から顔を検出する技術としては、例えば、非特許文献 1 に各種方式が挙げられている。その中では、いくつかの顕著な特徴（2つの目、口、鼻など）とその特徴間の固有の幾何学的位置関係とを利用するか、又は人間の顔の対称的特徴、人間の顔色の特徴、テンプレート・マッチング、ニューラル・ネットワークなどを利用することによって、人間の顔を検出する方式が示されている。 20

【0003】

前者の例としては、非特許文献 2 で、blobおよびstreakの検出をもとに画像中から顔を検出する方式が提案されている。すなわち、目は暗い特徴（dark blob）、頬や鼻は明るい特徴（light blob）、眉、口、顔の輪郭は線状のエッジ（streak）であることを利用して、それらの特徴を画像中から検出し、幾何学的な位置関係を考慮して顔の検出を行っている。また、比較的簡単な特徴であるblobを低解像度の画像から検出して、その検出結果にもとづいて高解像度の画像からblobおよびstreakを精度よく検出するようにして、処理の高速化を図っている。しかしながら、このような特徴の検出にもとづく方法では、被写体の姿勢、照明条件、画像のノイズなどの変化に検出結果が影響されやすく、このような条件下において高精度に検出を行うことは困難である。 30

【0004】

一方、非特許文献 3 で提案されている方式は、ニューラル・ネットワークにより画像中の顔パターンを検出する方法である。以下、非特許文献 3 による顔検出の方法について簡単に説明する。 40

【0005】

まず、顔を含む画像データをメモリに読み込み、顔と照合する所定の領域を画像中から切り出す。そして、切り出した領域の画素値の分布を入力としてニューラル・ネットワークによる演算で一つの出力を得る。このとき、ニューラル・ネットワークの重み、閾値が膨大な顔画像パターンと非顔画像パターンによりあらかじめ学習されており、例えば、ニューラル・ネットワークの出力が 0 以上なら顔、それ以外は非顔であると判別する。

【0006】

そして、ニューラル・ネットワークの入力である顔と照合する画像パターンの切り出し位置を、例えば、図 3 に示すように画像全域から縦横順次に走査していくことにより、画像中から顔を検出する。また、様々な大きさの顔の検出に対応するため、図 3 に示すよう 50

に読み込んだ画像を所定の割合で順次縮小し、それに対して前述した顔検出の走査を行うようにしている。このような被写体パターンの照合にもとづく方法は、学習サンプルや前処理に工夫を行うことで被写体の姿勢、照明条件、画像のノイズなどの変化に比較的強い検出を行うことができるものの、画像中から切り出したパターンの照合を網羅的に行うため、膨大な計算時間が必要であった。

【0007】

以上のことから、両者の長所を組み合わせることで被写体の検出を行う方法が考えられる。例えば、検出の最初のステップで非特許文献2に係る方法でもって目の候補を画像中から抽出し、目の候補から選択した2つの目を両眼として顔の候補を設定する。そして、設定した顔候補を画像パターンとして切り出し、非特許文献3に係る方法でもってニューラル・ネットワークにより顔の判定を行う。

10

【0008】

この処理方法による顔の検出を図6に示す。(a)は入力画像の一部である顔画像であり、ここからまず目の位置a1、a2が抽出される。そして、a1、a2と顔全体の位置関係から顔領域bが設定され、この顔領域bが(b)に示すような画像パターンとして切り出され、照合が行われる。しかしながら、例えば、(c)に示すように被写体の視線がずれているような場合には、目の位置としてc1、c2が抽出されるので、(d)に示す如く顔領域が設定され、設定された顔領域が(d)に示すような画像パターンとして切り出されることになるので、切り出したこのような画像パターンでもって照合が行われることになる。

20

【0009】

この場合、照合パターンでの顔の位置がずれてしまうので、顔パターンと認識されず、正しく顔が検出されない可能性がある。なお、このような状況は、被写体の視線がずれている場合に限らず、例えば、照明の影響によって暗い部分の検出位置がずれたり、眉毛、めがねのフレームを目と誤ったりする場合にも発生する可能性がある。

【0010】

したがって、従来例では、画像から被写体パターンを検出する処理において精度と計算時間を両立できないといった問題点があった。また、従来例を単純に組み合わせた場合においても、被写体パターンの検出精度が悪くなるといった問題点があった。

【非特許文献1】IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL.24, NO.1, JANUARY 2002, "Detecting Faces in Images: A Survey"

30

【非特許文献2】Theoretical Foundations of Computer Vision, vol.69, 1993, "Multiresolution Face Detection"

【非特許文献3】IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL.20, NO.1, JANUARY 1998, "Neural network-based face detection"

【発明の開示】

【発明が解決しようとする課題】

【0011】

本発明は以上の問題に鑑みて成されたものであり、画像中の被写体を高速、且つ高精度に検出するための技術を提供することを目的とする。

40

【課題を解決するための手段】

【0012】

本発明の目的を達成するために、例えば、本発明の画像処理方法は以下の構成を備える。

【0013】

即ち、画像から所定の被写体を検出する画像処理方法であって、
前記画像の輝度成分で構成される輝度画像を生成する生成工程と、
前記輝度画像を再帰的に縮小することで、複数枚の縮小画像を生成する縮小工程と、
縮小画像内における前記所定の被写体の部分的な特徴の候補としての特徴候補領域の位置を検出する第1の検出工程と、

50

前記第1の検出工程で検出した位置のうち2つを包含し、且つ当該2つの位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する第1の判別工程と

、
前記第1の判別工程で前記所定の被写体の領域ではないと判別された被写体候補領域内のそれぞれの特徴候補領域の位置近傍に近傍領域を設定する設定工程と、

一方の近傍領域内で取りうる第1の位置と、他方の近傍領域内で取りうる第2の位置とを包含し、且つ当該第1, 2の位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する第2の判別工程と、

前記第1の判別工程、前記第2の判別工程のそれぞれで、前記所定の被写体の領域であると判断された被写体候補領域のうち一部若しくは全部のデータを出力する出力工程と
を備えることを特徴とする。

10

【0014】

本発明の目的を達成するために、例えば、本発明の画像処理装置は以下の構成を備える。

【0015】

即ち、画像から所定の被写体を検出する画像処理装置であって、

前記画像の輝度成分で構成される輝度画像を生成する生成手段と、

前記輝度画像を再帰的に縮小することで、複数枚の縮小画像を生成する縮小手段と、

縮小画像内における前記所定の被写体の部分的な特徴の候補としての特徴候補領域の位置を検出する第1の検出手段と、

20

前記第1の検出手段が検出した位置のうち2つを包含し、且つ当該2つの位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する第1の判別手段と

、
前記第1の判別手段によって前記所定の被写体の領域ではないと判別された被写体候補領域内のそれぞれの特徴候補領域の位置近傍に近傍領域を設定する設定手段と、

一方の近傍領域内で取りうる第1の位置と、他方の近傍領域内で取りうる第2の位置とを包含し、且つ当該第1, 2の位置で規定される被写体候補領域が、前記所定の被写体の領域であるのかを判別する第2の判別手段と、

前記第1の判別手段、前記第2の判別手段のそれぞれで、前記所定の被写体の領域であると判断された被写体候補領域のうち一部若しくは全部のデータを出力する出力手段と
を備えることを特徴とする。

30

【発明の効果】

【0016】

本発明の構成により、画像中の被写体を高速、且つ高精度に検出することができる。

【発明を実施するための最良の形態】

【0017】

以下添付図面を参照して、本発明を好適な実施形態に従って詳細に説明する。

【0018】

[第1の実施形態]

本実施形態に係る画像処理装置は、PC(パーソナルコンピュータ)やWS(ワークステーション)等のコンピュータにより構成されており、デジタルカメラなどの撮像装置から入力した画像、インターネットなどのネットワークを介して外部機器からダウンロードした画像、CD-ROMやDVD-ROMなどの記憶媒体からの読み出しにより入力した画像など、様々な入力形態で入力した画像中に含まれている所定の被写体を検出する。なお、本実施形態では被写体として人間の顔を用いるが、その他の被写体を用いても良い。

40

【0019】

先ず、このような処理を行う本実施形態に係る画像処理装置について説明する。図10は、本実施形態に係る画像処理装置に適用可能なコンピュータのハードウェア構成を示す図である。

50

【0020】

201はCPUで、RAM202やROM203に格納されているプログラムやデータを用いてコンピュータ全体の制御を行うと共に、コンピュータが行う後述の各処理を実行する。

【0021】

202はRAMで、外部記憶装置207や記憶媒体ドライブ装置208から読み出したプログラムやデータを一時的に記憶するためのエリア、I/F209を介して外部から受信したデータを一時的に記憶する為のエリア、CPU201が各種の処理を実行する為に用いるワークエリア等、各種のエリアを適宜提供することができる。

【0022】

203はROMで、ここにブートプログラムや本コンピュータの設定データなどを格納する。

【0023】

204、205は夫々キーボード、マウスで、コンピュータの操作者が操作することで各種の指示をCPU201に対して入力することができる。

【0024】

206は表示部で、CRTや液晶画面などにより構成されており、CPU201による処理結果を文字や画像等でもって表示する。

【0025】

207は外部記憶装置で、ハードディスクドライブ装置等の大容量情報記憶装置であって、ここにOS（オペレーティングシステム）や、コンピュータが行う後述の各処理をCPU201に実行させるためのプログラムやデータが格納されており、これらはCPU201による制御に従って適宜RAM202に読み出される。

【0026】

208は記憶媒体ドライブ装置で、CD-ROMやDVD-ROMなどの記憶媒体に記録されているプログラムやデータを読み出して、RAM202や外部記憶装置207等に出力する。なお、上記外部記憶装置207に記憶されているプログラムやデータの一部を上記記憶媒体に記録しておいても良く、その場合には、これら記憶されているプログラムやデータを使用する際に、記憶媒体ドライブ装置208がこの記憶媒体に記録されているプログラムやデータを読み出して、RAM202に出力する。

【0027】

209はI/F（インターフェース）で、ここにデジタルカメラやインターネットやLANのネットワーク回線等を接続することができる。

【0028】

210は上述の各部を繋ぐバスである。

【0029】

なお、コンピュータへの画像の入力形態については特に限定するものではなく、様々な形態が考えられる。

【0030】

図1は、本実施形態に係る画像処理装置に適用可能なコンピュータの機能構成を示すブロック図である。同図に示す如く、本実施形態に係る画像処理装置は、画像入力部10、画像メモリ20、画像縮小部30、目候補抽出部40、照合パターン抽出部50、顔判別部60、顔領域出力部70、再探索候補設定部80により構成されている。

【0031】

画像入力部10は、例えばデジタルスチルカメラ、フィルムスキャナーなどの装置から出力された画像データを受け、後段の画像メモリ20に出力するものである。なお、上述の通り、画像の入力形態については特に限定するものではない。

【0032】

画像メモリ20は、画像入力部10から出力される画像データを格納するためのメモリである。

10

20

30

40

50

【0033】

画像縮小部30は、まず、画像メモリ20から受けた画像データの輝度成分で構成される輝度画像を生成する。そして、生成した輝度画像を再帰的に縮小することで、複数枚の縮小画像を生成する。生成したそれぞれの縮小画像（画像メモリ20から受けた画像データに基づいて生成したオリジナルの輝度画像も1/1の縮小画像と解釈すれば、このオリジナルもまた、縮小画像に含めることができる）は順次後段の目候補抽出部40に出力する。

【0034】

目候補抽出部40は、画像縮小部30から受けた縮小画像内で目の候補としての領域（目候補領域）を抽出し、抽出した領域の位置を求める。

10

【0035】

照合パターン抽出部50は、目候補抽出部40が求めた縮小画像上における目候補領域の位置を受けると、2つの位置を用いて、この縮小画像上における所定領域を以下の処理で用いる照合対象パターンとして抽出する。このような処理は、画像縮小部20から受けたそれぞれの縮小画像について行う。

【0036】

顔判別部60は、照合パターン抽出部50から受けた照合対象パターンが顔パターンであるか、それとも非顔パターンであるのかを判別する処理を行う。

【0037】

顔領域出力部70は、顔判別部60が顔パターンであると判別した照合対象パターンを出力する。

20

【0038】

再探索候補設定部80は、顔判別部60が顔パターンではないと判別した場合に動作するものであり、照合パターン抽出部50が照合対象パターンを抽出する為に用いた2つの目候補領域の位置のそれぞれの近傍に近傍領域を設定し、設定したそれぞれの近傍領域内で再度、目候補領域の位置を探索する。探索の結果、取得した2つの目候補領域の位置は照合パターン抽出部50に出力する。

【0039】

以上の各部は、例えば、CPU201の一機能として動作するものである。

【0040】

次に、CPU201が図1に示した各部として動作することでなされる処理、即ち、画像中に含まれている被写体を検出するための処理について、同処理のフローチャートを示す図2を用いて以下説明する。なお、同図のフローチャートに従った処理をCPU201に実行させるためのプログラムやデータは外部記憶装置207（もしくは記憶媒体ドライブ装置208が読み取り可能な記憶媒体）に保存されており、これをCPU201の制御に従って適宜RAM202にロードし、CPU201がこれを用いて処理を実行することで、コンピュータは以下説明する各処理を実行することになる。

30

【0041】

外部記憶装置207や、I/F209を介して外部から画像データが入力されると、CPU201は、これをRAM202内の上記画像メモリ20に相当するエリアに一時的に格納する（ステップS101）。なお、本コンピュータに入力した画像が圧縮されている場合には、これを伸張してからRAM202に一時的に格納する。

40

【0042】

本実施形態では、入力された画像データを構成する各画素は、R、G、Bで表現されるものであるとする。従って、CPU201は、ステップS101でRAM202に格納した画像データに基づいて、この画像の輝度成分で構成される画像（輝度画像）、即ち、この画像を構成する各画素の値をこの画素の輝度値に変換した画像を生成する（ステップS102）。しかし、ステップS101でRAM202に格納した画像データを構成する各画素がYCrCbで表現されるものである場合には、ステップS102では、Y成分のみを用いて輝度画像を生成する。

50

【0043】

次に、CPU 201は、生成した輝度画像を再帰的に縮小することで、複数枚の縮小画像を生成する(ステップS103)。例えば、ステップS102で生成した輝度画像(以下の説明上、縮小画像1と呼称する)の縦横のサイズを1/1.2倍した縮小画像2を生成し、次に縮小画像2の縦横のサイズを1/1.2倍した縮小画像3を生成する、というように、複数枚の縮小画像を生成する。これは、以後の処理で目の候補を抽出する際に様々な大きさの目の検出に対応するため複数のサイズの画像データに対して順次検出を行うようにしたためである。なお、生成する縮小画像の枚数については特に限定するものではない。

【0044】

そして、ステップS104以降では、生成したそれぞれの縮小画像について処理を行う。即ち、ステップS104以降の処理は、生成した縮小画像の数だけ繰り返し行うことになる。

【0045】

以降の説明では、生成した縮小画像はサイズの大きい順に縮小画像1、縮小画像2、、、、縮小画像Nと呼称するものとし、先ず、縮小画像1について以降の処理を行うものとする。なお、処理の対象として選択する順番については特に限定するものではない。

【0046】

先ず、CPU 201は、縮小画像1を所定のサイズの部分領域毎に参照し、参照した部分領域内で局所的に暗い領域を目の候補領域(目候補領域)として抽出し、抽出した目候補領域の位置を取得する(ステップS104)。

【0047】

より詳しくは、先ず、輝度画像1から所定のサイズ(例えば5×5画素)のブロックを縦横順次切り出していく。なお、各ブロックは隣接するブロックと多少重なり合うように切り出す方がよい。次にブロック内の輝度データを照合パターンとし、あらかじめ多数の目の輝度分布から得られた標準のテンプレートとの照合を行う。ここでは、照合パターンとテンプレートとの相関演算(入力輝度で正規化された相関演算がより望ましい)を行い、相関値が所定の閾値を越えた場合、そのブロックの中心位置を目候補領域の位置(目の特徴点)として抽出する。このような処理は照合に必要な部分領域が小さいので、直接顔の照合を行う場合に比べるとより少ない演算回数で処理を行うことができる。

【0048】

次に、CPU 201は、ステップS104で取得した目候補領域の位置を用いて、輝度画像1上における顔の候補領域を抽出する(ステップS105)。

【0049】

ステップS105における処理についてより詳細に説明すべく、図7を用いて説明する。図7は、顔候補領域を設定する処理を説明する図である。同図ではe1, e2が目候補領域の位置(特徴点の位置)である。

【0050】

先ず、ステップS104で取得した目候補領域の位置を、2つの位置毎に用いる。そして2つの位置を用いて1つの矩形を規定する。例えば、図7の場合、2つの位置e1, e2を通る直線方向にこの2つの位置間をA倍した長さを有する一辺(第1の辺)と、2つの位置e1, e2を通る直線に対して垂直な直線方向にこの2つの位置間をB倍した長さを有する一辺(第2の辺)とを有する矩形700を、2つの位置の中心から第2の辺方向に距離Cだけ移動した位置がこの矩形700の中心位置となるような矩形700を規定する。よって、2以上の目候補領域の位置が存在する場合には複数の矩形を設定することになる。

【0051】

この矩形内の領域が顔の候補領域(顔候補領域)である。よって、ステップS105では、この矩形を規定すると共に、この矩形内における顔候補領域を抽出する。なお、図7ではe1を右目、e2を左目として顔候補領域を設定したが、e1を左目、e2を右目と

10

20

30

40

50

して顔候補領域を設定してもよいし、2つの顔候補領域を設定してもよい。

【0052】

図2に戻って、次に、ステップS105で抽出した顔候補領域の輝度パターンが顔のパターンであるのかを判別する(ステップS106)。より詳しくは、先ず、ステップS105で抽出した顔候補領域内の各画素の輝度分布を正規化する処理を行う。例えば、ヒストグラム平滑化などの輝度補正を行う。これは、撮像される被写体パターンはその照明条件によって輝度分布が変わるので被写体照合の精度が劣化するのを抑制するためである。そして、正規化された輝度パターンが顔のパターンであるのかを判別する。

【0053】

図4は、所定領域内のパターンを識別する為のニューラルネットワークの動作について示した図である。同図においてRは、例えば画像上で識別する領域を示すものであり、本実施形態ではこの領域Rを同図に示す如く、3種類の方法にてさらに領域分割し、各ニューロン(Nで示す)への受容野とする。そして、分割された領域の輝度分布を各ニューロンに入力し、中間層での出力が得られる。そして、各ニューロンの出力を出力層のニューロンの入力として最終出力が得られる。

10

【0054】

ここで、各ニューロンでは予め学習によって得られた重みと輝度分布との積和演算およびその結果のシグモイド関数による演算が行われる。本実施形態では、出力層のニューロンの出力値と所定の閾値との比較を行い、閾値を越える出力のパターンを顔、それ以外を顔でない(非顔)とした(ニューラル・ネットワークの詳細および学習の方法については、上記非特許文献3を参照されたい)。なお、輝度分布が正規化されたパターンが顔パターンであるか否かを判別する処理(ステップS106における処理)についてはこれに限定するものではなく、所定の入力パターンに対して被写体の判別を行い、顔パターンと非顔パターンとの判別が可能な値を出力するようなものであればよい。例えば、入力である部分領域を図5に示す如く顔の輪郭、眼、鼻、口、頬の位置関係にもとづいて分割するように受容野を設定してもよい。図5は、顔の各部分を示す図である。

20

【0055】

また、ニューラル・ネットワークによる判別に限らず、例えば、Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2001に"Rapid Object Detection using a Boosted Cascade of Simple Features"と題するViolaとJonesによる報告で提案されているAdaBoostによる方式を用いてもよい。

30

【0056】

そして、ステップS105で抽出した顔候補領域が顔パターンと判別された場合には(正確には、この顔候補領域の輝度分布を正規化したパターンが顔パターンと判別された場合)、処理をステップS112に進め、この顔候補領域内の画像データを、顔領域のデータとして出力する(ステップS112)。出力先については特に限定するものではないが、RAM202内の所定のエリアであっても良いし、外部記憶装置207やI/F209を介してデータ通信可能な外部の装置であっても良い。

【0057】

そして次に、全ての顔候補領域についてステップS106以降の処理を行ったのかを判別し(ステップS113)、行っていない場合には未だ処理対象となっていない顔候補領域についてステップS106以降の処理を行う。一方、全ての顔候補領域についてステップS106以降の処理を行ったのであれば、処理をステップS114に進め、全ての縮小画像について以上の処理を行ったのかを判断し(ステップS114)、まだ処理対象となっていない縮小画像がある場合には処理をステップS104に戻し、次の縮小画像についてステップS104以降の処理を行う。本実施形態では現在縮小画像1について処理を行ったので、次は縮小画像2について処理を行う。

40

【0058】

一方、ステップS105で抽出した顔候補領域が顔パターンと判別されなかった場合には(正確には、この顔候補領域の輝度分布を正規化したパターンが顔パターンと判別され

50

なかった場合)、処理をステップS107に進め、以降の処理を行う。

【0059】

ここで、ステップS105で抽出した顔候補領域が実際には顔パターンであるのに、顔パターン中の目の位置がずれているだけで、顔パターンと判別されない場合がある。従って、本実施形態では、ステップS105で抽出した顔候補領域が顔パターンではないと判別されても、この顔候補領域内の目候補領域の位置を再度検討し、検討した目候補領域の位置で規定される顔候補領域について再度、この顔候補領域が顔パターンであるのか否かを判別する。これにより、例えば被写体の右目、左目がどのような方向を向いていたとしても、その目の位置ずれを許容して顔検出を行うことができる。

【0060】

先ず、ステップS105で抽出した顔候補領域内における2つの目候補領域の位置のそれぞれについて、図8(a)に示す如く、その位置の近傍に近傍領域(図8(a)ではE1, E2)を設定する(ステップS107)。図8(a)は近傍領域の設定例を示す図である。

【0061】

即ち、一方の目候補領域の位置を包含し、この位置を中心とする所定サイズの領域を、一方の目候補領域に対する近傍領域として設定すると共に、他方の目候補領域の位置を包含し、この位置を中心とする所定サイズの領域を、他方の目候補領域に対する近傍領域として設定する。

【0062】

次に、ステップS108において以下に説明する処理を行う。先ず、それぞれの近傍領域を複数の矩形に分割する。図8(b)は近傍領域E1(E2)を3×3の矩形に分割した例を示す図である。

【0063】

ここで、以下の説明のために便宜上、近傍領域E1内の各矩形(9個の矩形)をそれぞれB11, B12, , B19と呼称する。一方、近傍領域E2内の各矩形(9個の矩形)をそれぞれB21, B22, , B29と呼称する。

【0064】

よって、ステップS108では、次に、近傍領域E1内の矩形位置(矩形の中心位置)と、近傍領域E2内の矩形位置との組み合わせを選択する。即ち、B11の位置とB21の位置との組み合わせ、B11の位置とB22の位置との組み合わせ、、、、というように、この場合81通りの組み合わせがあるが、それぞれの近傍領域の中心の矩形同士の組み合わせについては既に処理済みであるので、実際には80通りの組み合わせについて順次選択する。

【0065】

そして、ステップS108で選択された組み合わせにおける一方の位置と他方の位置とを用いてステップS105と同様の処理を行うことで、ステップS108で選択された組み合わせにおける一方の位置と他方の位置とを2つの目候補領域の位置とした場合の顔候補領域を規定し、これを抽出する(ステップS109)。図8(c)は抽出された顔パターンの一例を示す図である。

【0066】

そして、ステップS106と同様に、この顔候補領域が顔パターンであるのか否かを判別し、顔パターンである場合には処理をステップS112に進め、この顔パターンを出力するのであるが、顔パターンではないと判別された場合には処理をステップS111に進め、ステップS109で全ての組み合わせについて処理を行ったかを判別し、全ての組み合わせについて行っていない場合には処理をステップS108に戻し、次の組み合わせを選択し、選択した組み合わせについて以降の処理を行う。

【0067】

なお、以上の繰り返し処理において、出力された顔候補が画像上で重なり合う場合には、例えば、ステップS106やステップS111で用いた上記ニューラルネットワークの

10

20

30

40

50

出力層のニューロンの出力値が最も高い顔候補領域の顔パターンを出力するようにする。

【0068】

以上の説明により、本実施形態によって、被写体の目が様々な方向を向いていることにより目の位置がずれているような場合であっても、精度良く画像から顔の領域を抽出することができる。

【0069】

なお、本実施形態では、被写体の顔の部分的な特徴として目を用いたがその他のものを用いても良い。

【0070】

[第2の実施形態]

本実施形態では、顔候補領域が顔パターンであるのかを2段階で判別する。従って、本実施形態に係る画像処理装置に適用可能なコンピュータの機能構成は図9に示す如く、図1に示した機能構成において、顔判別部60が2つ(第1の顔判別部61、第2の顔判別部62)に分かれている。

【0071】

図9は、本実施形態に係る画像処理装置に適用可能なコンピュータの機能構成を示すブロック図である。なお、同図において図1と同じ部分については同じ番号を付けており、その説明は省略する。

【0072】

以下、この2段階の顔判別処理について説明する。

【0073】

第1の顔判別部61は、顔判別部60と基本的には同様の処理を行うのであるが、第1の顔判別部61は、比較的幅広く顔パターンを検出するように(顔パターンが顔でないと判別しないように)あらかじめ学習されている。従って、学習の際に例えば図6(d)に示すようなパターンも顔として判別されるよう考慮されている。よって、少なくとも、第1の顔判別部61により顔パターンであると判別された顔候補領域しか、以降の処理では意味を成さない。そこで、第1の顔判別部61で顔パターンと判別された顔候補領域についてのみ、以降の処理対象となる。即ち、画像中に多数の顔候補領域が存在しても第1の顔判別部61により、少なくとも顔パターンとは見なされないのものは以降の処理対象とは成らないので、全体の処理効率を上げることができると共に、顔検出の精度を高めることができる。

【0074】

第2の顔判別部62は、比較的確度の高い顔パターンのみを検出するようにあらかじめ学習されている。第2の顔判別部62は、第1の顔判別部61によって顔パターンであると判別された顔候補領域について、顔判別部60と同様の処理を行い、この顔候補領域が顔パターンであるのかを判別する。以降の処理については第1の実施形態と同じである。

【0075】

このような処理により、図2のステップS107以降の処理を行う必要のある顔候補領域の数を軽減させることができ、より効率よく顔の探索を行うことができる。

【0076】

なお、本実施形態において、第1の顔判別部61による処理を処理時間を重視して比較的簡単な演算で行い、第2の顔判別部62による処理を判別精度を重視してより複雑な演算を行うように構成すれば、より効率よく顔の探索を行うことができる。

【0077】

以上、目の候補から選択した2つの目を両眼として顔の候補を設定し、設定した顔候補を画像パターンとして切り出して、顔の判定を行う方式において、一旦、非顔と判別した場合に目の近傍領域において再探索候補を設定して判別を行う上記方式は、被写体の視線がずれている場合や照明の影響によって暗い部分の検出位置がずれたり、眉毛、めがねのフレームを目と誤ったりするような目の抽出位置ずれに強いものであることを示した。

【0078】

10

20

30

40

50

[その他の実施形態]

また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記録媒体（または記憶媒体）を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記録媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記録媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記録した記録媒体は本発明を構成することになる。

【0079】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム（OS）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0080】

さらに、記録媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0081】

本発明を上記記録媒体に適用する場合、その記録媒体には、先に説明したフローチャートに対応するプログラムコードが格納されることになる。

【図面の簡単な説明】

【0082】

【図1】本発明の第1の実施形態に係る画像処理装置に適用可能なコンピュータの機能構成を示すブロック図である。

【図2】画像中に含まれている被写体を検出するための処理のフローチャートである。

【図3】様々なサイズの縮小画像について、被写体パターンの照合を行う処理を説明する図である。

【図4】所定領域内のパターンを識別する為のニューラルネットワークの動作について示した図である。

【図5】顔の各部分を示す図である。

【図6】従来顔検出について説明する図である。

【図7】顔候補領域を設定する処理を説明する図である。

【図8】ステップS107～ステップS109における処理を説明する図である。

【図9】本発明の第2の実施形態に係る画像処理装置に適用可能なコンピュータの機能構成を示すブロック図である。

【図10】本発明の第1の実施形態に係る画像処理装置に適用可能なコンピュータのハードウェア構成を示す図である。

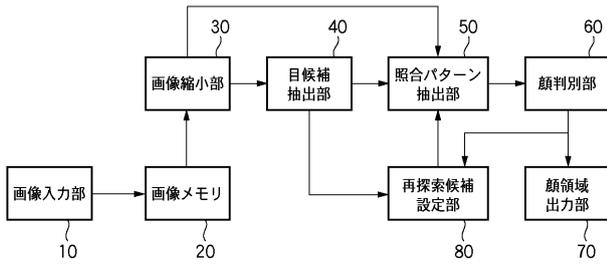
10

20

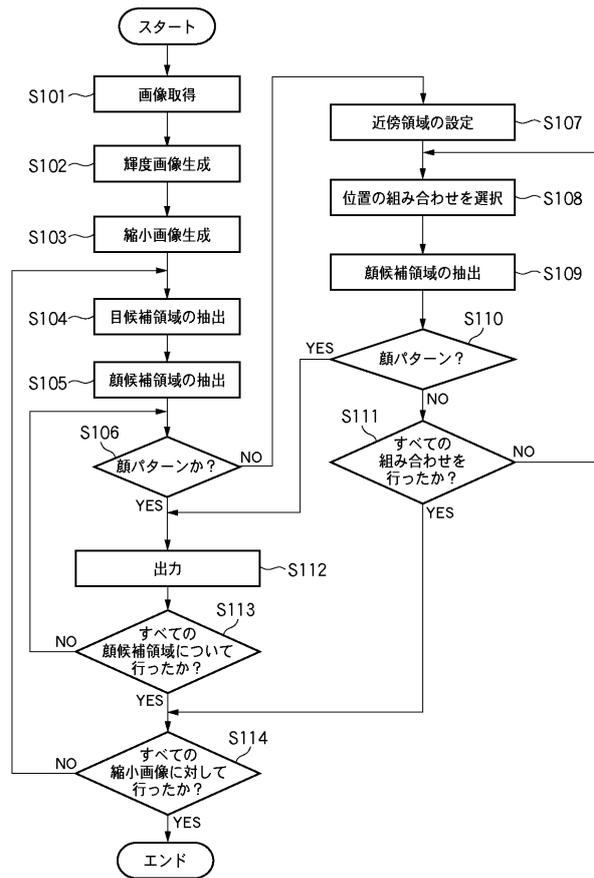
30

40

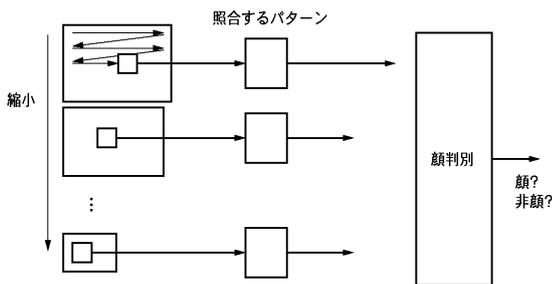
【図1】



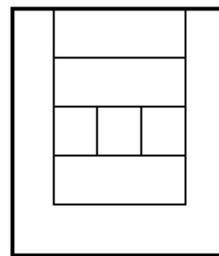
【図2】



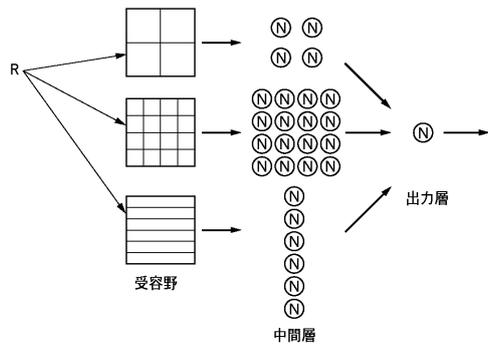
【図3】



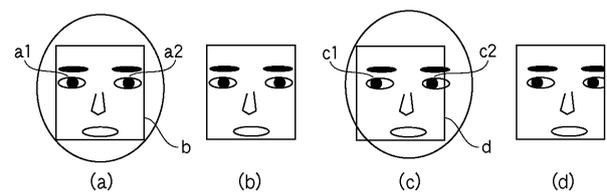
【図5】



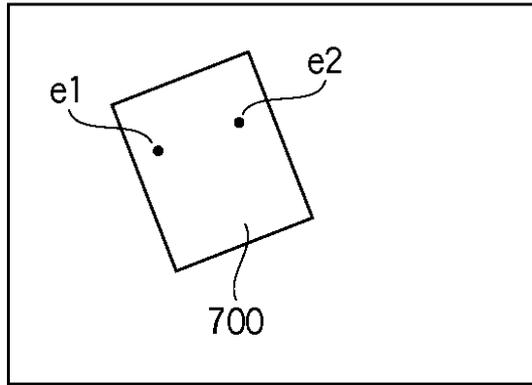
【図4】



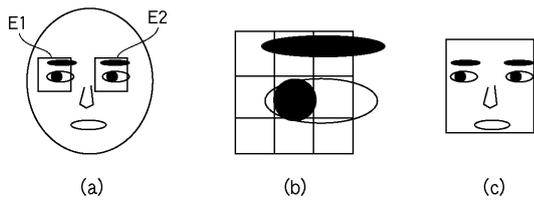
【図6】



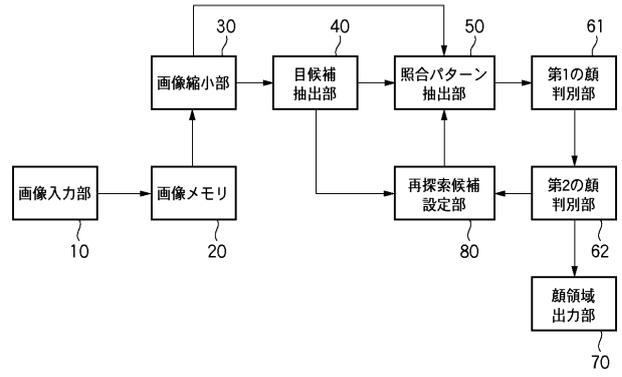
【図7】



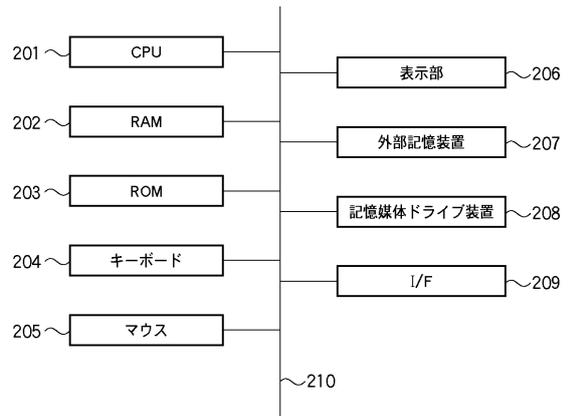
【図8】



【図9】



【図10】



フロントページの続き

F ターム(参考) 5B057 BA02 CA01 CA08 CA12 CA16 CB01 CB08 CB12 CB16 CD05
CE17 CF05 DA07 DC25 DC36
5C122 DA04 DA08 DA11 EA65 FH07 FH09 FH14 FH15 HA13 HA35
HB01 HB05
5L096 AA02 AA06 BA08 BA20 CA02 DA01 EA03 EA12 EA13 EA17
FA34 GA30 HA08 HA11 MA03