

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3962517号
(P3962517)

(45) 発行日 平成19年8月22日(2007.8.22)

(24) 登録日 平成19年5月25日(2007.5.25)

(51) Int. Cl.

G06T 7/00 (2006.01)

F I

G06T 7/00 I O O C

請求項の数 22 外国語出願 (全 16 頁)

(21) 出願番号	特願平11-349498	(73) 特許権者	000001007
(22) 出願日	平成11年12月8日(1999.12.8)		キヤノン株式会社
(65) 公開番号	特開2001-167273(P2001-167273A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成13年6月22日(2001.6.22)	(74) 代理人	100076428
審査請求日	平成16年12月14日(2004.12.14)		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
		(74) 代理人	100101306
			弁理士 丸山 幸雄

最終頁に続く

(54) 【発明の名称】 顔面検出方法及びその装置、コンピュータ可読媒体

(57) 【特許請求の範囲】

【請求項 1】

複数の画素から形成されるカラーデジタル画像に含まれる顔面を検出する顔面検出装置の判定手段及び解析手段の各手段が実行する顔面検出方法であって、

前記判定手段が、前記画像の撮影時の照明条件を利用して前記画素の色の試験を行い、該画素から主に肌色を有する画素を判定する判定工程と、

前記解析手段が、前記判定工程で主に肌色を持つと判定された画素に基づいて、顔面特徴解析を実行する解析工程とを備え、

前記判定工程は、前記画像の撮影時の照明条件に依存するしきい値を使用し、前記照明条件はメタデータとして表現される

ことを特徴とする顔面検出方法。

【請求項 2】

前記照明条件は、該画像の撮影時に取得される

ことを特徴とする請求項 1 に記載の顔面検出方法。

【請求項 3】

前記画像は、予め定められたフォーマットに従って符号化され、前記照明条件は、前記予め定められたフォーマットに関連するメタデータとして表現される

ことを特徴とする請求項 1 に記載の顔面検出方法。

【請求項 4】

前記判定工程は、前記試験の前に、

10

20

前記画像を各々複数の前記画素を有する領域に分割する分割工程とを備え、
前記試験は、主に肌色である領域の1つを判定するために、前記領域各々内の画素に対し実行される

ことを特徴とする請求項1に記載の顔面検出方法。

【請求項5】

前記判定工程は、色分布モデルを利用する
ことを特徴とする請求項1に記載の顔面検出方法。

【請求項6】

前記色分布モデルは、特定画像撮影装置用に生成される
ことを特徴する請求項5に記載の顔面検出方法。

10

【請求項7】

前記色分布モデルは、異なる前記照明条件用に生成される
ことを特徴とする請求項5に記載の顔面検出方法。

【請求項8】

前記色分布モデルは、先にサンプリングされた顔面画像データの撮影時の異なる前記照明条件用に生成される
ことを特徴とする請求項7に記載の顔面検出方法。

【請求項9】

前記色分布モデルは、フラッシュを使用して撮影された画像群及びフラッシュを使用せずに撮影された画像群用に生成される
ことを特徴とする請求項8に記載の顔面検出方法。

20

【請求項10】

前記色分布モデルは、屋内で撮影された画像群及び屋外で撮影された画像群用に生成される
ことを特徴とする請求項8に記載の顔面検出方法。

【請求項11】

前記色分布モデルの各々は、色表現ベクトルの周波数ヒストグラムとして表現される
ことを特徴とする請求項1に記載の顔面検出方法。

【請求項12】

前記色分布モデルの各々は、色表現ベクトルの確率分布として表現される
ことを特徴とする請求項1に記載の顔面検出方法。

30

【請求項13】

前記色分布モデルの各々は、色表現ベクトルのバイナリマップとして表現される
ことを特徴とする請求項1に記載の顔面検出方法。

【請求項14】

前記色表現ベクトルは、先にサンプリングされた顔面画像データにおける予め定められた肌色画素の知覚色空間から抽出される
ことを特徴とする請求項1乃至請求項13のいずれか1項に記載の顔面検出方法。

【請求項15】

前記色表現ベクトルは、先にサンプリングされた顔面画像データにおける予め定められた肌色画素のRGB値から抽出される有彩色値を含む
ことを特徴とする請求項1乃至請求項13のいずれか1項に記載の顔面検出方法。

40

【請求項16】

前記バイナリマップは、前記先にサンプリングされた画像データにおいて識別される肌色画素の割合を有する
ことを特徴とする請求項13に記載の顔面検出方法。

【請求項17】

前記画素の1つは、該画素に対応する前記色表現ベクトルが前記バイナリマップ内で発生する場合に、肌色として区分される
ことを特徴とする請求項16に記載の顔面検出方法。

50

【請求項 18】

前記画素の各々は、該画素の各々に対応する前記色表現ベクトルの周波数が閾値周波数を超える場合に、肌色として区分される

ことを特徴とする請求項 11 に記載の顔面検出方法。

【請求項 19】

前記画素の各々は、該画素の各々に対応する前記色表現ベクトルの確率が確率閾値を超える場合に、肌色として区分される

ことを特徴とする請求項 12 に記載の顔面検出方法。

【請求項 20】

前記領域は、該領域内の前記画素の総数の予め定められた割合以上が肌色として区分される場合、主に肌色であると判定される

ことを特徴とする請求項 17 に記載の顔面検出方法。

【請求項 21】

複数の画素から形成されるカラーデジタル画像に含まれる顔面を検出する顔面検出装置であって、

前記画像の撮影時の照明条件を利用して前記画素の色の試験を行い、該画素から主に肌色を有する画素を判定する判定手段と、

前記判定手段で主に肌色を持つと判定された画素に基づいて、顔面特徴解析を実行する解析手段とを備え、

前記判定手段は、前記画像の撮影時の照明条件に依存するしきい値を使用し、前記照明条件はメタデータとして表現される

ことを特徴とする顔面検出装置。

【請求項 22】

複数の画素から形成されるカラーデジタル画像に含まれる顔面を検出する顔面検出装置の判定手段及び解析手段の各手段が実行する顔面検出処理をコンピュータに実行させるためのプログラムを記憶したコンピュータ可読媒体であって、

前記判定手段が、前記画像の撮影時の照明条件を利用して前記画素の色の試験を行い、該画素から主に肌色を有する画素を判定する判定工程と、

前記解析手段が、前記判定工程で主に肌色を持つと判定された画素に基づいて、顔面特徴解析を実行する解析工程とを備え、

前記判定工程は、前記画像の撮影時の照明条件に依存するしきい値を使用し、前記照明条件はメタデータとして表現される

ことをコンピュータに実行させるためのプログラムを記録したことを特徴とするコンピュータ可読媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、デジタルカラー画像における画像処理に関し、特に、カラーデジタル画像における顔面の検出に関するものである。

【0002】

カラーデジタル画像は、マルチメディアデータベースに格納され、様々なコンピュータアプリケーションに利用されることが多くなった。多くのコンピュータアプリケーションにおいては、多段階処理における 1 ステップとして、可視画像中の顔面の位置を検出することが望ましい。多段階処理は、内容ベースの画像検索や、自動金銭預け払い機 (ATM) 又は防犯カメラと共に使用される個人の識別又は照合や、あるいは人間と計算装置との自動対話などを含むと考えられる。

【従来の技術】

従来、固有顔面 (eigenfaces) や、ニューラルネットワークや、クラスタリングや、特徴識別及び肌色技法などを含む様々な顔面検出方法が知られている。これらの技法はいずれも長所と短所を有しているが、それらが共通して備えている 1 つの特徴は非常に大量の計

10

20

30

40

50

算を必要とし、そのために動作速度が大変に遅いこと、あるいは、計算は速いが、顔面を検出するのに十分な強固さを持っていないことである。

【 0 0 0 3 】

固有顔面 (eigenfaces) 又は固有ベクトル方法は顔面の認識に特に適しており、照明のばらつきをある程度許容できるが、顔面の様々に異なる視点には対処できず、また、様々な顔面特徴が隠された状態 (例えば、サングラスをかけている場合に起こる状態) も取り扱えない。また、スケールの相違に対し不変ではない。

【 0 0 0 4 】

ニューラルネットワーク方式は、大量の顔面画像及び非顔面画像に基づく訓練を利用するもので、実現するのが比較的簡単であり、顔面特徴が隠された状態でもある程度まで対応でき、また、照明のばらつきにもある程度まで対応できるという利点を有する。また、誤検出を使用してニューラルネットワークを再学習させることにより、検出速度を改善することも比較的容易である。しかしながら、この方式はスケールの相違に対し不変ではなく、様々に異なる視点又は方向に対応せず、顔面を画像上に位置決めするために集中的なプロセスを必要とする。

10

【 0 0 0 5 】

クラスタリング技法は、固有顔面方式に幾分類似している。通常、画像に沿って画素ウィンドウ (例えば、 20×20) を移動させ、その結果として得られるテストパターンと、プロトタイプ顔面画像やプロトタイプ非顔面画像との距離をベクトルによって表現する。ベクトルはテストパターンと顔面モデルとの類似度及び差を捕捉する。そこで、ベクトルが顔面又は非顔面を表現しているかに関して分類を行うために、ニューラルネットワークを訓練することができる。この方法は強固ではあるが、異なるスケール、異なる視点又は方向には対応しない。また、この方法は顔面を位置決めするために集中的なアプローチを必要とし、仮定上の変数に依存する。

20

【 0 0 0 6 】

特徴識別方法は、まゆ、目、鼻及び口などの顔面特徴群又は顔面特徴群と思われるものをサーチすることに基づいている。検出プロセスは、顔面特徴を識別し、それらの特徴を特徴対、部分顔面群又は顔面になりうるものにグループ分けすることから成る。このプロセスは、スケールの相違に対し相対的に不変であり、集中的なサーチが行われず、一部の顔面特徴が隠れた状態でも処理でき且つ様々に異なる視点及び方向も処理できるという利点を有する。主な欠点は、誤検出が多く起こりやすく、また、使用される顔面特徴検出アルゴリズムによって性能が大きく左右されてしまうことである。

30

【 0 0 0 7 】

人間の顔面を検出するために肌の色を使用することは、Yang J及びWaibel Aによる論文 (1995年) 「Tracking Human Faces in Real-Time」 (CMU-CS-95-210, School of Computer Science Carnegie Mellon University) の中に記載されている。この提案は、人間の視覚システムは様々な明るさや照明源に適応し、それは、人間による色の知覚が広範囲にわたる環境照明条件の中で首尾一貫していることを示唆するものであるという概念に基づいていた。従って、正確ではあるが、低次元の色情報を保持しつつ、肌色表現から明るさを除去することは可能であると考えられた。その結果、この従来技法においては、有彩色空間を使用した。有彩色 (例えば、 r 及び g) は、RGB値から次のように導出することができる。

40

【 0 0 0 8 】

$r = R / (R + G + B)$ 及び $g = G / (R + G + B)$

これらの有彩色は、明るさが無いときの「純」色として知られている。

【 0 0 0 9 】

この色空間を利用して、Yang及びWaibelは、人の別と人種の別の双方を含めて、様々な異なる人々の肌の色の分布をまとめてクラスタリングすることを発見した。すなわち、これは、様々な異なる人々の肌の色は、非常に近いものであり、その主な差は輝度の差であることを意味している。

50

【0010】

従来のこの方法は、何よりもまず、一連の顔面の画像例を使用し、それらの例から手操作により肌色領域を選択することにより、肌色分布モデルを生成していた。次に、試験画像を有彩色空間に変換した。更に、試験画像（変換後）中の各画像を肌色モデルの分布と比較した。最後に、そのようにして検出された全ての肌色画素を識別すると、隣接する肌色画素の領域を顔面になりうる領域として考えることができた。

【0011】

従来のこの方法は、色の処理が個別に顔面特徴を処理する方法よりはるかに高速で、色が方向によってほとんど左右されず、顔面特徴の一部が隠された場合でも影響を受けないという利点を有する。また、このシステムは視点やスケールの相違による影響もほぼ受けない。しかしながら、この方法には、照明条件の相違によって顔面の色表現が影響を受けるおそれがあること、及びカメラ（例えば、デジタルカメラ又はフィルムカメラ）が違えば、同じ環境で同じ人を撮影したとしても、色値が異なってしまうことなどを含むいくつかの欠点もある。

10

【0012】

しかしながら、従来の方法の重大な欠点は、肌色モデルがそれほど高い識別能力を持たないことである（すなわち、画素が肌色分布に含まれるか否かに基づいて画素を選択する場合、多数の肌色でない画素も誤って肌色分布に取り込まれる結果となる）。また、顔面になりうると思うことができる肌色画素のクラスタ又は領域の場所を確定することも難しい。

20

【発明が解決しようとする課題】

本発明の目的は、デジタルカラー画像中の1つ又は複数の顔面を検出する改良された方法を提供することである。

【課題を解決するための手段】

本発明の1つの目的に従えば、複数の画素から形成されるカラーデジタル画像中の顔面を検出する顔面検出方法を開示し、前記顔面検出方法は、前記画像から得られる少なくとも1つの画像撮影条件を利用して前記画素の色の試験を行い、該画素から主に肌色を有する画素を判定する判定工程と、前記判定工程で判定された画素のみに対し更に顔面特徴解析を実行し、主に肌色を有していない前記画素に対しては、前記顔面特徴解析を実行しない解析工程とを備える。

30

【0013】

好ましくは、各画像撮影条件は、画像の撮影時に取得される。有利な点としては、画像は、所定フォーマットに従って符号化され、少なくとも1つの画像撮影条件は、所定フォーマットに関連するメタデータとして表現される。より好ましくは、少なくとも1つの画像撮影条件は、画像の撮影時の照明条件を有する。

【0014】

特に、判定工程は、前記試験の前に、前記画像を各々複数の前記画素を有する領域に分割する分割工程とを備え、前記試験は、主に肌色である領域の1つを判定するために、前記領域各々内の画素に対し実行され、前記解析工程は、主に肌色であると判定される領域のみに顔面特定解析を実行する。

40

【0015】

本発明の別の目的に従えば、カラーデジタル画像中の顔面を検出する顔面検出方法を開示し、前記顔面検出方法は、前記画像を各々がほぼ同色を有する複数の領域に分割する分割工程と、上記分割工程で生成した前記領域の各々の色を試験して、主に肌色を有する領域を判定する判定工程と、上記判定工程で判定した領域のみに対し更に顔面特徴解析を実行し、前記分割工程で生成された、主に肌色を有していない前記領域に対しては、前記顔面特徴解析を実行しない解

50

析工程と
を備える。

【0016】

本発明を実行するための装置及びコンピュータ可読媒体もまた、開示される。

【発明の実施の形態】

図1は、各々がRGB値を有する 832×624 個の画素5を有する 832×624 サイズの典型的なカラーデジタル画像1を示している。

【0017】

本発明の実施形態1に従えば、先にYang及びWaibelによる従来の技術に関連して説明したように画素ごとにこの画像の肌色を考慮するのではなく、画像1を複数の領域に分割する。この分割の例が、各領域2の全ての画素がほぼ同色を有するということに基づく分割が、図2に概略的に示される。

10

【0018】

実施形態1は、図3のフローチャートで示されるプロセス30を実行し、この中で、画像の領域分割はステップ31で実行される。次に、ステップ32で、画像のそれぞれの領域を（先に説明した）有彩色空間に変換する。次のステップ33では、ステップ31で確定した領域のうち、肌色を有する画素が指定の割合（典型的には90から95%）を占める領域を選択する。これらの選択した領域は、境界ボックス又はその他の境界指示要素によって表現すると好都合である。最後に、ステップ34で、重複領域の組み合わせも含めた選択領域を更に解析（好ましくは、肌色に基づかない解析）して、それらの選択領域が1つ以上の顔面を表現するか否かを判定する。

20

【0019】

この初期色分けは、領域ベースの任意のカラー画像分割技術を使用できる。好ましくは、画像は、1つの局所領域にわたって類似の色を有する連結画素グループを探索することにより、複数の色領域に区分される。主要な色領域を検出し、雑音の影響を低減するため、非常に小さく、独立した空間グルーピングを無視しても良い。各初期空間領域の代表色は、その領域の平均色値によって決定される。

【0020】

1つの色領域は任意に選択された画素から始まり、その画素は隣接画素と比較される。色類似度閾値Tを使用して、色が類似している隣接画素を追加することにより、領域は広げられて行く。隣接画素のR、G、B値をそれぞれ R_p 、 G_p 、 B_p とし、且つ領域の平均R、G、B値を R_m 、 G_m 、 B_m とすると、 $|R_p - R_m| < T$ 及び $|G_p - G_m| < T$ 及び $|B_p - B_m| < T$ であれば、隣接画素を領域に追加する。

30

【0021】

領域が類似する色の隣接画素をそれ以上持たないとき、領域は成長を停止し、初期空間グルーピングの1つを表現することになる。この領域の大きさが所定閾値より小さければ、そのグルーピングは無視される。所定閾値以上の画素数を有する領域は、その平均色によって表現される。

【0022】

次に、まだ、どの領域にも属していない新たな画素を選択し、新たな色領域を始める。このプロセスは、画像中のあらゆる画素が初期空間グルーピングに属するか、または小領域の一部であるとして無視されるまで続く。

40

【0023】

初期空間グルーピングにより、各領域がその平均色で表現されるような画像の色領域分割が完了する。

【0024】

このようにして、画像の大部分が顔面又は顔面の一部ではない多くの画像については、大半の画素は明らかに顔面ではない領域又はオブジェクト（前景又は背景などである）にグルーピングされる。従って、これらの非顔面オブジェクトはその色に基づいて迅速に排除される。

50

【 0 0 2 5 】

領域が確定されたならば、 r 値及び g 値を求めるために、先に挙げた式を利用してそれらの領域を「純」有彩色空間に変換する。この色変換後、所定領域の中にある画素の少なくとも 85% が顔面の色であるという規則のような許容範囲の大きい規則を使用して、更に検討するに値する領域を選択することができる。好ましくは、顔面の色の試験は、原画像の性質を、例えば、原画像はフラッシュありでまたはなしで撮影されたかを考慮に入れる。この情報は、通常、画像源（例えば、カメラ）から判定できる。

【 0 0 2 6 】

その後、それらの選択領域のみを更に試験して、顔面特徴の有無を判定する。この試験は、領域が顔面を構成しているか否かに関して最終的な判定を下す試験である。この点に関して、最終試験は計算速度の上では遅くなる傾向にあり、従って、上述の領域の排除によって、この低速度の方法が画像全体のうち相対的に小さな部分にのみ適用される結果となる。これにより、処理時間は全体として短縮される。従って、上記の方法は、すべての画素がない場合でも、ほとんどの部分をコンピュータ単純処理で実行し、肌色領域のみ複雑な処理を実行する。

【 0 0 2 7 】

領域が顔面を表現しているか否かを検証する好ましい方法は、顔面特徴を検出する手段としてのエッジ検出技法に依存する。特に、目、まゆ及び口などの顔面特徴は顔面の上に黒い棒として現われ、そのため、黒いエッジを描く場合が多い。

【 0 0 2 8 】

好ましい形態のエッジ検出は、エッジ検出フィルタの使用である。これは直交する方向に働く 2 つの関数を利用する。水平の棒を検出する場合、垂直方向には二次微分ガウス関数を使用し、水平方向にはガウス関数を使用する。

【 0 0 2 9 】

このようにしてエッジを確定したならば、検出したそれぞれのエッジを検査する。検出された一対のエッジは、いずれも、その相対的位置と大きさに従って、一対の目、一対のまゆ、又は目及びそれと関連するまゆのいずれかから派生しており、従って、それらのものを示すものとわかる。同様に、1 つのエッジが既に検出された目及び / 又はまゆに対して適切な位置に配置されていれば、口から派生し、従って、口を示すものであるとわかる。

【 0 0 3 0 】

このように処理を進行することにより、所定領域が肌色から、まゆ / 目を経て、口に至るまで顔面特徴を累積し始める。顔面でありそうな所定領域に関して顔面特徴が多く見出されるほど、その領域が実際に顔面である確率は高くなる。

【 0 0 3 1 】

更に、以上説明した方法は、ほぼ肌色と同色である背景領域を背に顔面が位置しているような状況にも対処できる。このような状況においては、Yang 及び Waibel の方法の場合、顔面と背景との境界を検出できないであろう。従って、その領域全体をそれ以降の試験のために選択することになる。しかしながら、本発明に係る上記実施形態の方法では、どの画素が肌色であるかに関する決定を実行する前に、色空間全体を利用して画像を分割する。そのため、顔面は背景から分離されやすい。加えて、本発明の方法が顔面の方向又は顔面の一部が隠された状態によって影響を受けないことは言うまでもない。

【 0 0 3 2 】

更に、本発明の方法によれば、後の段階で偽の肯定を検査することができるので、後の試験において、最終的には顔面領域であるとみなされるかもしれない領域が除外されることはない。

【 0 0 3 3 】

上述した実施形態 1 は、原画像の性質が初期顔面検出処理を実行する場合に考慮されても良いことに注意されたい。また、各実施形態は、この特徴に基づいて構成される。

【 0 0 3 4 】

カメラを使用して画像が撮影される場合で、人物を撮影するためには手動でカメラ設定（

10

20

30

40

50

例えば、シャッタースピード、絞り、焦点距離等）を確立すること、あるいはカメラに対し上記の設定を自動的に実行させることのいずれか一方が必要である。どちらの場合にしても、カメラ設定は、直接、撮影される画像の状況及び品質に影響を与える。特に、画像内の対象物の知覚明度、色、シャープネスは、カメラ設定がどのように行われたかにすべて依存する。例えば、異なるカメラ設定で同一シーンの写真を２枚とると、異なる色及び明度で表現される同一対象物の２つの画像を取得することができる。それゆえ、（デジタル）画像に含まれる（特に）色情報を校正する機能は、広範囲の様々な対象物検出及び、色が強い識別特徴となる区分作業を可能にする。

【 0 0 3 5 】

顔面検出は、１つの適用例であり、本発明は、それぞれが特定照明条件で適応された顔面
色識別モデル（ＣＤＭ）の生成が、顔面検出の精度及び信頼性の両方を改善できることを
規定している。照明条件の変化は、実施形態１の顔面検出方法で与えられる特徴であるよ
うなフラッシュの使用で起こりうる。明度は、輝度及びクロミナンスのような色特徴表現
であるので、このような色特徴は、顔面検出を定量化するために使用しても良い。

【 0 0 3 6 】

顔面色分布モデルを使用して画像が処理される前に、顔面色分布モデルは構築されなけれ
ばならない。これは、図５に示される方法５０に従って実行される。方法５０は、まず、
ステップ５２で顔面を含む代表画像である画像サンプルを取得し、この代表画像は、照明
条件の変化、つまり、輝度及びクロミナンスの変化を示す照明条件の下で取得される。次
に、ステップ５４で、これらの画像は手動で試験され、モデル形成における処理用に肌色
領域を抽出する。ステップ５４は、顔面彩色画素のサンプルの周囲の境界ボックスを手で
描くことによって実行されても良い。ステップ５６で、以下のように、抽出された画素に
対する色表現値を抽出する。これは、各画素が少なくとも２次元ベクトルによって表現さ
れるように、抽出された画素を、例えば、 $CIE\ L^*u^*v$ あるいは $CIE\ L^*a^*b$ の
ような知覚色空間に変換することによって実行されても良い。また、例えば、 HSI 及び
 HSV のような他の色空間を使用しても良い。好ましくは、各画素が、輝度値及びクロミ
ナンス値の両方を含む３次元ベクトルとして表現される。

【 0 0 3 7 】

次に、ステップ５８で、画素の色表現値は、各画像の撮影時に得られる照明条件に従って
複数のセット（５８ａ、５８ｂ、...、５８ｎ）に分割される。セット例としては、フラッ
シュ、フラッシュなし、屋内、屋外、これらの組み合わせがある。一方、例えば、フラッ
シュ動作のようなカメラから直接取得される照明変数は、そのセットを識別し区別するた
めに使用されても良い。晴れあるいは曇り、夕暮れあるいは夜明け、あるいは蛍光灯、白
熱電球あるいはハロゲンのような人工照明の種類のような他の照明条件も、これらの目的
のために使用あるいは検出されても良い。これらの詳細は、画像撮影時に人間の手によっ
て与えられても良い。

【 0 0 3 8 】

次に、顔面サンプルのセット（５８ａ、５８ｂ、...、５８ｎ）のそれぞれに対し、ステッ
プ６０で、顔面色画素のサンプルに最適な対応色分布モデル（ＣＤＭ）（６０ａ、...、６
０ｎ）を構築する。ＣＤＭは、ヒストグラム、確率密度関数、あるいはバイナリビットマ
ップであり得る。１つの実施形態では、ガウスＰＤＦの混合は、当業者には周知の技術を
使用するサンプルデータに適し、この周知の技術は、例えば、相互確認、ジャックナイフ
のいずれか一方を有する期待値最大化（ＥＭ）アルゴリズムや、色分布モデルを最適に評
価するために使用されるブートストラップ技術がある。

【 0 0 3 9 】

各ＣＤＭ（６０ａ、...、６０ｎ）が構築されると、次に、ステップ６２に示されるように
、非顔面画に関連するように区分される色ベクトル以下で、かつ潜在的な顔面画素として
区分される色ベクトル以上の対応確率閾値（６２ａ、...、６２ｎ）を設定することが好ま
しい。加えて、顔面色確率は、以下に詳述される顔面画像解析ステップで直接使用される
。好ましい実施形態では、ＣＤＭは知覚色空間（例えば、 $CIE\ L^*u^*v$ あるいは CI

10

20

30

40

50

E L * a * b) を使用して抽出される色表現値から構築され、そして、入力画像の色フォーマット、即ち、R G BあるいはY U Vのいずれか一方に逆変換される。これは、入力画像の知覚色空間への変換を必要としない。

【 0 0 4 0 】

種類の異なる画像撮影装置は、光学要素（例えば、レンズ、ミラー、絞り等）の品質及びサイズによって異なるパフォーマンスを有するので、典型的なC D MあるいはC D Mのセットは、特定撮影装置に対し生成される。露出計を有する画像撮影装置（例えば、カメラ）の1操作においては、ある瞬間での露出計の読取値での画像が必要なC D Mを決定するために使用される。このような方法においては、色モデルにより広い範囲を持たせたり、可能なヒューマンインタフェースを介在しないで選択することもできる。このような抑制は、人間がフラッシュ操作を手動で選択する場合、そうでなければ、フラッシュの自動操作が必要とされない場合に発生する。また、フラッシュ/屋外での上述のセット例では、4つのC D Mセットが得られる。露出計を使用する、いわゆる、4ビット符号化では、16モデルを提供することができる。また、露出計の使用は、拡張再生出力を提供し、使用される顔面サンプルには、ラボラトリ条件及びカメラ製造時の設備が考慮された色分布モデルを生成することを可能にする。

10

【 0 0 4 1 】

実施形態2に従う画像の処理70は、図6に示される。ステップ72で入力画像が提供され、ステップ74で、画像が撮影された時の照明条件が決定される。この決定は、カメラ（例えば、フラッシュ+屋内、フラッシュなし+屋外、フラッシュなし+屋内、フラッシュ+屋外）から直接取得されるバイナリデータあるいは、画像から提供されるあるいは所定フォーマットに従って符号化あるいは通信された画像に伴う対応メタデータに基づいても良い。一旦、照明条件が決定されると、先に決定されたC D M（60a、...、60n）を保持するルックアップテーブル78の列から対応あるいは類似C D Mが選択される。ステップ80で、入力画像72の第1画素が選択され、ステップ82で、画素の（R G BあるいはY U V）色成分が選択されたC D M（60a、...、60n）内に含まれているかどうか試験される。

20

【 0 0 4 2 】

比較ステップ82に続く図6に示されるステップは、C D Mが記憶されている順に依存する。好ましい実施形態では、ステップ62（図5）で選択された閾値は、バイナリマップあるいはルックアップテーブルを構築するために使用され、このバイナリマップあるいはルックアップテーブルでは、色ベクトルが閾値化顔面色分布内に含まれる場合に代表色ベクトルは1で表現され、色ベクトルが閾値化色分布内に発生しない場合に代表色ベクトルは0で表現される。一方、C D Mは、閾値化色分布の代表色ベクトルの周波数で表現しても良い（即ち、C D Mは事実上の代表色ベクトルのヒストグラムである）。別の変形例は、サンプル化分布がガウスあるいはガウスの混合のような変数モデルによって近似される場合である。後者の場合、C D Mは変数モデルの変数（例えば、平均、共分散）で構成される。

30

【 0 0 4 3 】

図6に示されるように、好ましい実施形態に従って、ステップ82から得られる1あるいは0値は、ステップ84でマップに追加される。ステップ86で、処理対象の画像内に未処理の画素があるかどうかを判定し、ステップ88で、最適な試験のために次の画素を取得して、ステップ82に進む。選択されたC D Mに対し、すべての画素が試験された場合、ステップ90は、上述のステップの処理結果として、検出された肌色画素を使用して形成されたバイナリ顔面画像マップを生成する。

40

【 0 0 4 4 】

次に、ステップ92で、バイナリ顔面画像マップは肌色画素の解析が実行され、ステップ94で、画像に対する顔面検出マップを出力する。ステップ92の解析は、実施形態1と類似し、好ましくは、顔面色の考慮から独立している。

【 0 0 4 5 】

50

実際には、ステップ 890 で形成されるバイナリ顔面マップは、顔面画素（1 の画素）で周囲を囲まれた小さい非顔面画素（0 の画素）がある領域、あるいはその逆の場合の領域を含んでいても良い。ステップ 92 に従う解析に対する 1 つの処理は、バイナリ顔面画像の処理であり、この処理は、潜在的な顔面の最小サイズよりも小さい領域内に含まれる任意の画素位置に 0 を設定し、かつ、その画素が顔面色画素とおもわれる画素によって周囲を囲まれる場合に任意の画素位置を 0 から 1 に設定するためのものである。これは、最適に形成された構成要素での形態的開始及び終了動作のペアを使用して実行されても良い。

第 1 構成要素は、例えば、

【 0 0 4 6 】

【 数 1 】

10

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

【 0 0 4 7 】

が開始動作で使用され、このサイズ以下の潜在的な顔面候補画素位置を除去する。第 2 構成要素は、例えば、

20

【 0 0 4 8 】

【 数 2 】

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

30

【 0 0 4 9 】

が終了動作で使用され、潜在的な顔面候補画素位置で任意の欠落部分を埋める。

【 0 0 5 0 】

構成要素を使用する別の方法では、Hough 変換、あるいは肌色を有する領域内の画素数を計数し、所定割合値に対し閾値を設定することの使用を含んでいる。他の方法も、これらの処理を実行するために使用されても良い。

【 0 0 5 1 】

図 6 の方法 70 は、すべての可能性を見込んだ状態で顔面が検出され、顔面が存在する入力画像内の画素位置の顔面検出マップを出力する。

【 0 0 5 2 】

40

顔面が存在するかどうかを判定するために顔面画素とおもわれる画素の処理を行う上述のエッジ検出方法は、方法 70 から得られる顔面検出マップ 94 に対し実行されても良い。

【 0 0 5 3 】

好ましい実施形態では、顔面色分布モデルは、フラッシュ、フラッシュなし、屋内、屋外等の照明条件の識別数に対し構築される。しかしながら、この技術は、カメラから取得される変数に直接基づく任意の照明条件の通常の場合以上に拡張されても良い。このような状況内で使用されても良いカメラ変数のリストは以下ようになる。

【 0 0 5 4 】

i) ホワイトバランス

ii) ホワイトバランスモード

50

- iii) 絞り (アイリス)
- iv) シャッタースピード
- v) オートゲインコントロール (AGC)
- vi) 自動露出 (AE) モード
- vii) ガンマ
- viii) 基盤 (pedestal) レベル
- ix) 光斑補償

カメラから取得される変数は、各画像 (あるいはビデオシーケンス) の撮影に関連するメタデータストリームから取得されるのが好ましい。そのような転送プロトコルの例には、I E E E 1 3 9 4 (「ファイヤワイヤ」) が含まれる。また、I S O 基準は、M P E G - 7、M P E G - 4、J P E G の画像及びビデオへのメタデータの付加方法を定義している。

10

【 0 0 5 5 】

一方、図 1 から図 3 を参照して説明される実施形態 1 は、ほぼ同色の領域に従って画像を分割し、実施形態 2、実施形態 3 では、そのように処理されない。

【 0 0 5 6 】

実施形態 3 は、入力画像 1 5 2 を提供し、それぞれが実施形態 2 のステップ 7 4、ステップ 7 6、ステップ 7 8 に対応するステップ 1 5 4、ステップ 1 5 6、ステップ 1 5 8 に従って処理する方法 5 0 によって図 7 に示される。一旦、ステップ 1 5 6 で、最適な C D M が選択されると、ステップ 1 6 0 は、1 つ以上の領域として入力画像を処理し続ける。1 つの領域として、画像全体が画素単位で処理される。一方で、入力画像は、幾何学的に単純画素ブロック (例えば、2 5 × 2 5 画素、1 0 × 2 0 画素) に分割されても良く、この単純画素ブロックは、ラスタ順に形成され処理される。また、実施形態 1 のように、領域はほぼ同色に基づいて分割される。

20

【 0 0 5 7 】

ステップ 1 6 2 で、処理対象の第 1 領域を選択し、ステップ 1 6 4 で第 1 領域の第 1 画素を選択する、ステップ 1 6 6 で、選択画素と、実施形態 2 のステップ 8 2 に対応する方法で C D M とを比較する。色分布モデルと選択画素が合致する場合、ステップ 1 6 8 は、基準を満たす領域内の画素数をインクリメントする。ステップ 1 7 0 で、処理対象の領域となる任意の画素があるかどうかを判定し、任意の画素がある場合、ステップ 1 7 2 で、次の画素を取得し、最適な試験のためにステップ 1 6 6 に戻る。領域内のすべての画素が処理された場合、ステップ 1 7 4 で、所定割合閾値と、肌色とする領域に対して区分される画素の割合とを比較するために処理を続ける。この割合が所定割合閾値未満である場合、領域は非顔面領域と判定され、ステップ 1 7 6 で、処理対象の任意の領域があるかどうかを試験するために処理を続ける。任意の領域がある場合、ステップ 1 7 8 で、次の領域を選択し、ステップ 1 6 4 の処理に戻る。ここで、計数値はリセットされる。任意の領域がない場合、方法 1 5 0 は、ステップ 1 8 4 で終了する。

30

【 0 0 5 8 】

割合が所定割合閾値以上である場合、領域は顔面領域の可能性があると判定され、ステップ 1 8 0 で、顔面検出解析に従って領域を評価するために処理を続ける。この解析が顔面を検出しない場合、方法 1 5 0 は、ステップ 1 7 6 に進み、任意の別の領域を処理する。このステップ 1 8 0 の解析が顔面を検出する場合、ステップ 1 8 2 で、この領域を顔面領域として登録し、ステップ 1 7 6 に戻る。

40

【 0 0 5 9 】

ステップ 1 8 0 による最適な処理の結果として実行されても良い解析の一例は、上述の実施形態 1 に関連して説明されるエッジ検出方法である。

【 0 0 6 0 】

上記実施形態のそれぞれは、画像内の顔面検出が 2 段階処理、つまり、候補画素あるいは領域の可能性のある画素あるいは領域を取得するための画像の第 1 フィルタリングと同種の第 1 表現処理、第 1 段階を通過した画素あるいは領域に対し動的判定を行うための解析

50

である第2表現として処理されても良いことを示している。それぞれの場合において、画像の撮影に関連する照明条件は、第1段階で実行される判定に寄与する。

【0061】

上述の方法は、好ましくは、一般的な汎用コンピュータシステム100を用いて実現され、例えば、図4に示したシステム100では、図3かつ/あるいは図5及び図6の処理は、例えば、コンピュータシステム100内で実行するアプリケーションプログラムのようなソフトウェアとして実現される。特に、上記方法のステップは、コンピュータによって実行されるソフトウェア内の命令によって実現される。このソフトウェアは、2つの部分に分割しても良く、1つは上述のステップを実行する部分、もう1つはユーザとのユーザインタフェースを実現する部分である。本ソフトウェアは、コンピュータ可読媒体、例えば、後述の記憶装置に格納すると良い。本ソフトウェアは、そのコンピュータ可読媒体からコンピュータ内にロードされ、コンピュータにより実行される。かかるソフトウェアやコンピュータプログラムが記録されたコンピュータ可読媒体は、コンピュータプログラム製品である。コンピュータでコンピュータプログラム製品を用いることは、本発明の実施形態に従って顔面候補領域を検出するための特長的な装置を実現する。

10

【0062】

コンピュータシステム100は、コンピュータモジュール101と、キーボード102やマウス103などの入力装置と、プリンタ115や表示装置114を含む出力装置を備える。変復調装置(モデム)116は、コンピュータモジュール101により用いられ、例えば、電話回線121や他の機能的な媒体を介して通信ネットワーク120と通信する。モデム116は、インターネットや他のネットワークシステム、例えば、LAN(Local Area Network)やWAN(Wide Area Network)等へのアクセスを取得するために使用され、これらは、入力画像源及び検出画像の出力先となり得る。

20

【0063】

コンピュータモジュール101は、典型的には、少なくとも1つのプロセッサユニット105と、例えば、半導体のメモリであるRAMやROMからなるメモリユニット106や、ビデオインタフェース107を含む入出力(I/O)インタフェースと、キーボード102やマウス103や(不図示)オプションとしてのジョイスティック用の入出力(I/O)インタフェースと、モデム116用のインタフェース108とを含んでいる。記憶装置109が設けられ、典型的なハードディスク110やフロッピーディスク装置111を含んでいる。磁気テープ装置(不図示)を用いてもよい。典型的には、CD-ROM装置112が不揮発性データとして用いられる。コンピュータモジュール101の構成要素105~113は、典型的には、内部接続バス104を介して、そして、当業者には周知のコンピュータシステム100の動作モードとなるような方法に従って、通信を行う。実施形態が実現可能なコンピュータの例として、IBM-PCとその互換機、SunのSPARCステーションやその発展型機である。

30

【0064】

典型的には、上記好適な実施形態のアプリケーションプログラムは、ハードディスク100上に配置され、読み取られ、プロセッサ105により実行されるに際して制御される。プログラムやネットワーク120から取り込まれるデータの間記憶は、半導体メモリ106を用いてハードディスク装置110と協同して実現されても良い。いくつかの例では、アプリケーションプログラムはユーザにCD-ROMやフロッピーディスク上に符号化された形で供給され、対応する装置112または111を介して読みとられ、または、ユーザがモデム116を介してネットワーク120から読みとってよい。更に、本ソフトウェアは、他のコンピュータ可読媒体、例えば、磁気テープやROMまたは集積回路、光磁気ディスク、コンピュータモジュール101と他の装置間での電波や赤外線の通信チャネル、PCMCIAカード等のコンピュータ可読カード、さらには、E-mail送信やWebサイト上に記録された情報を含むインターネットやイントラネット等、からロードすることができ。上述のものは、コンピュータ可読媒体の単に例示に過ぎない。他のコンピュータ可読媒体も、本発明の目的及び精神から逸脱することなく実施されても良い

40

50

。

【 0 0 6 5 】

候補顔面画像及び領域に対する後処理は、コンピュータシステム 1 0 0 あるいは 9 はそのような処理に対する周知の構成によって実行されても良い。

【 0 0 6 6 】

顔面候補領域を検出する本発明の方法は、別途、図 3 かつ / あるいは図 5 及び図 6 に示した種々の機能や副機能を実行する 1 つ以上の集積回路などの専用ハードウェアに組み込まれていてもよい。そのような専用ハードウェアは、グラフィックプロセッサ、デジタル信号プロセッサ、1 つ以上のマイクロプロセッサ並びに周辺メモリを含んでいてもよい。

【 0 0 6 7 】

10

[産業上の利用可能性]

上記のことから、本発明の実施形態は、自動銀行窓口装置や防犯カメラ、さらには、人間とコンピュータ機器との間の自動対話に用いるための、コンテンツベースの画像検索、個人識別若しくは認証などに適用可能であることが明らかである。

【 0 0 6 8 】

以上、本発明のいくつかの実施形態のみを説明したが、請求項で定義される本発明の目的及び精神から逸脱しない変形かつ / あるいは変更が可能である。

【 図面の簡単な説明 】

【 図 1 】 カラーデジタル画像の画素を概略的に表現する図である。

【 図 2 】 実施形態 1 に従って、図 1 の画像を各々がほぼ同色を有する複数の領域に分割した状態を示す図である。

20

【 図 3 】 実施形態 1 に従う顔面検出処理を示すフローチャートである。

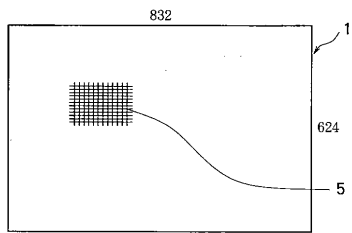
【 図 4 】 本発明の実施形態が実行可能な汎用コンピュータのブロック図である。

【 図 5 】 顔面色分布モデルの生成を示すフローチャートである。

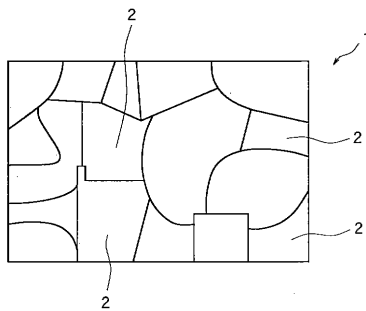
【 図 6 】 実施形態 2 に従う顔面検出処理を示すフローチャートである。

【 図 7 】 実施形態 3 に従う顔面検出処理を示すフローチャートである。

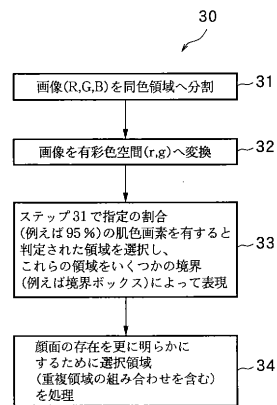
【図 1】



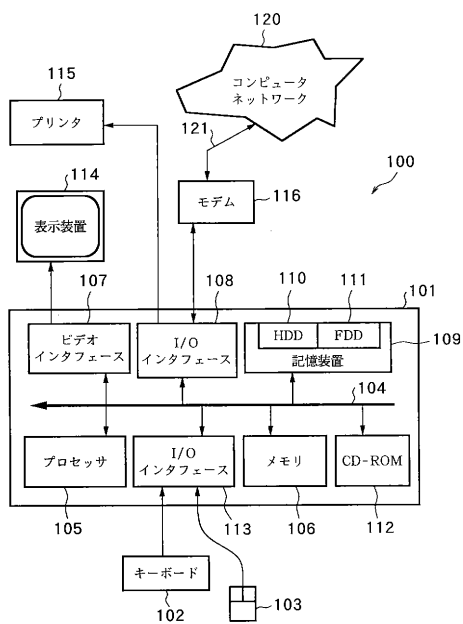
【図 2】



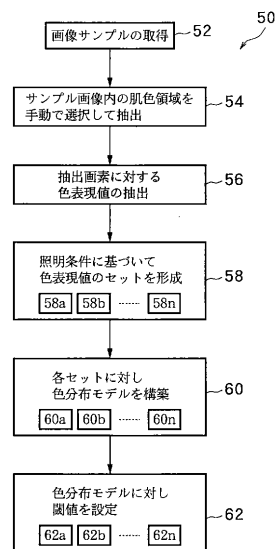
【図 3】



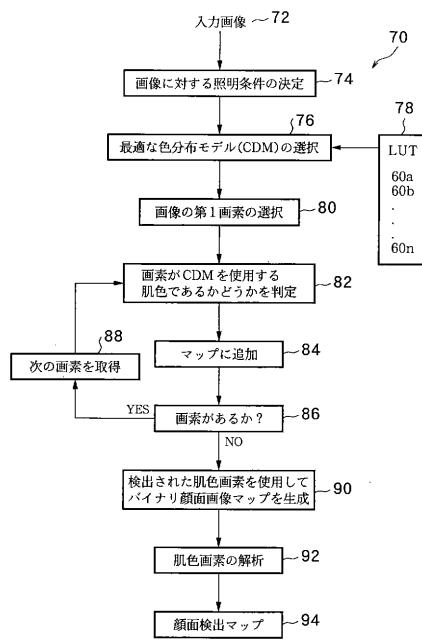
【図 4】



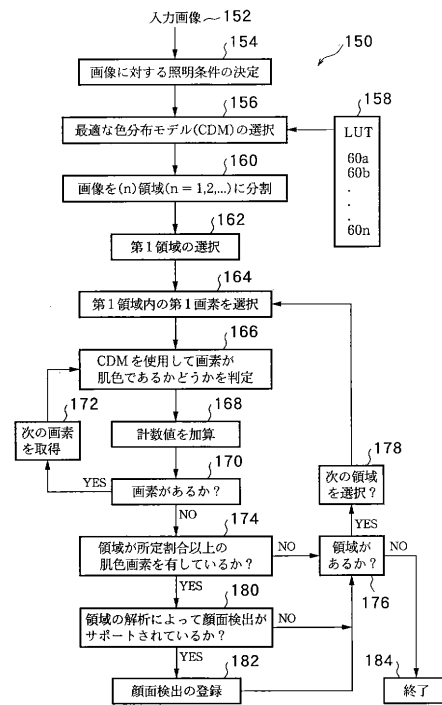
【図 5】



【図 6】



【図 7】



 フロントページの続き

- (72)発明者 エドウィン フー
オーストラリア国 2113 ニュー サウス ウェールズ州, ノース ライド, トーマス ホルト
ドライブ 1 キヤノン インフォメーション システムズ リサーチ オーストラリア プ
ロプライエタリー リミテッド内
- (72)発明者 アリソン レノン
オーストラリア国 2113 ニュー サウス ウェールズ州, ノース ライド, トーマス ホルト
ドライブ 1 キヤノン インフォメーション システムズ リサーチ オーストラリア プ
ロプライエタリー リミテッド内
- (72)発明者 アンドリュー ピーター ブラッドリー
オーストラリア国 2113 ニュー サウス ウェールズ州, ノース ライド, トーマス ホルト
ドライブ 1 キヤノン インフォメーション システムズ リサーチ オーストラリア プ
ロプライエタリー リミテッド内

審査官 松尾 俊介

- (56)参考文献 特開平11-155849(JP, A)
特開平10-191246(JP, A)
特開平08-063597(JP, A)
特開平06-160993(JP, A)
国際公開第99/035606(WO, A1)
特開平11-316846(JP, A)
特開平11-234560(JP, A)
特開平09-322178(JP, A)
木村朝子, 外2名, 肌色情報を用いた顔画像追跡, 電子情報通信学会技術報告, 日本, 社団法人
電子情報通信学会論文, 1996年 6月20日, Vol. 96, No. 115, pp. 65~
70

- (58)調査した分野(Int.Cl., DB名)

G06T 7/00