

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6204199号
(P6204199)

(45) 発行日 平成29年9月27日 (2017.9.27)

(24) 登録日 平成29年9月8日 (2017.9.8)

(51) Int. Cl.		F I			
G06T	7/00	(2017.01)	G06T	7/00	Q
G06T	1/00	(2006.01)	G06T	7/00	350B
			G06T	1/00	340A

請求項の数 19 (全 17 頁)

(21) 出願番号	特願2013-553746 (P2013-553746)	(73) 特許権者	304043073
(86) (22) 出願日	平成24年2月20日 (2012.2.20)		イオムニサイエント ビーティーワイ リ
(65) 公表番号	特表2014-505952 (P2014-505952A)		ミテッド
(43) 公表日	平成26年3月6日 (2014.3.6)		オーストラリア国 ニューサウスウェール
(86) 国際出願番号	PCT/AU2012/000158		ズ州 ゴードン エルジン ストリート
(87) 国際公開番号	W02012/109712		19
(87) 国際公開日	平成24年8月23日 (2012.8.23)	(74) 代理人	110002000
審査請求日	平成27年2月17日 (2015.2.17)		特許業務法人栄光特許事務所
(31) 優先権主張番号	2011900557	(74) 代理人	100090343
(32) 優先日	平成23年2月18日 (2011.2.18)		弁理士 濱田 百合子
(33) 優先権主張国	オーストラリア (AU)	(74) 代理人	100192474
(31) 優先権主張番号	2011905341		弁理士 北島 健次
(32) 優先日	平成23年12月21日 (2011.12.21)	(74) 代理人	100105474
(33) 優先権主張国	オーストラリア (AU)		弁理士 本多 弘徳

最終頁に続く

(54) 【発明の名称】 画像品質の評価

(57) 【特許請求の範囲】

【請求項 1】

被写体の画像品質を評価するための、コンピュータで実施される方法であって、
前記画像の2つ以上の部分画像の特徴と前記被写体の包括的なモデルとの類似度を判定するステップにおいて、前記包括的なモデルが同じタイプの複数の異なる被写体の複数の訓練画像の組み合わせに基づいて決められ、前記モデルが部分モデルを有し、前記部分モデルの各々が複数の訓練画像からの複数の部分画像からモデル化され、前記部分モデルの各々が前記画像の前記部分画像に対応し、類似度の判定が、各部分画像の特徴とそれに対応する部分モデルによってモデル化された特徴との類似度に基づいて決められるステップと、

前記画像の品質を、前記2つ以上の部分画像の前記判定された類似度に基づいて判定するステップと、を含む方法。

【請求項 2】

前記画像のための前処理ステップとして使用され、追加処理が、前記画像の前記判定された品質に基づいて決められる、請求項1に記載のコンピュータで実施される方法。

【請求項 3】

各訓練画像が、品質を決める複数の好ましい特性を有し、好ましい特性は、所定の前記被写体内移動、所定の回転、所定のスケール、所定の解像度、所定の姿勢、所定の照度のうちの1つまたは複数を含み、前記被写体が顔である場合所定の表情を含む、請求項1または2に記載のコンピュータで実施される方法。

【請求項 4】

各部分モデルが、平均ベクトルおよび分散行列によって特徴をモデル化し、前記平均ベクトルおよび分散行列は、前記部分モデルが基にしている各訓練画像の部分画像の特徴に基づいて決められる、請求項 1 ~ 3 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 5】

部分画像の特徴が、前記部分画像のより低い周波数成分のみに略基づいて決められ、前記モデルが、訓練画像のより低い周波数成分のみに略基づいて決められる、請求項 1 ~ 4 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 6】

部分モデルによってモデル化された前記特徴が、前記部分モデルが基にしている各訓練画像の部分画像の前記特徴の平均値である、請求項 1 ~ 5 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 7】

部分モデルによってモデル化された前記特徴が、前記部分モデルが基にしている各訓練画像の部分画像の前記特徴の平均ベクトルおよび分散行列に特徴付けられるガウス確率密度関数に基づいて決められる、請求項 1 ~ 6 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 8】

類似度の判定が、前記ガウス確率密度関数に基づいて前記画像の前記部分画像の前記特徴の確率を判定するステップを含む、請求項 7 に記載のコンピュータで実施される方法。

【請求項 9】

前記画像の前記部分画像が、それに対応する部分モデルが基にしている各訓練画像の前記部分画像と位置合わせされる、請求項 1 ~ 8 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 10】

前記類似度の判定が、部分画像ごとに独立して実行される、請求項 1 ~ 9 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 11】

類似度の判定が、部分画像ごとに 1 つの類似度基準判定するステップを含み、前記画像品質の判定が、前記 2 つ以上の部分画像の前記判定された類似度の結合に基づいて決められる、請求項 1 ~ 10 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 12】

前記 2 つ以上の部分画像の前記判定された類似度の結合が、部分画像ごとに独立して確率を判定するステップと、前記画像の部分画像ごとの前記確率に基づいて結合確率を判定するステップとを含む、請求項 11 に記載のコンピュータで実施される方法。

【請求項 13】

前記画像の前記判定された品質の指標を出力として提供するステップをさらに含む、請求項 1 ~ 12 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 14】

前記画像の前記判定された品質の指標を不揮発性メモリに記憶するステップをさらに含む、請求項 1 ~ 13 のいずれか一項に記載のコンピュータで実施される方法。

【請求項 15】

請求項 1 ~ 14 のいずれか一項に記載のコンピュータで実施される方法であって、複数の画像のそれぞれの品質を判定するために、上記複数の異なる画像ごとに前記方法を繰り返すステップと、

前記判定された品質に基づいて前記画像の追加処理を実行するステップと、をさらに含む方法。

【請求項 16】

前記被写体が顔である、請求項 1 ~ 15 のいずれか一項に記載のコンピュータで実施さ

10

20

30

40

50

れる方法。

【請求項 17】

前記被写体が顔であり、

前記追加処理が、より良好な品質を有すると判定された画像のサブセットに対して顔認識または顔照合を実行する、請求項 2 または 15 に記載のコンピュータで実施される方法。

【請求項 18】

コンピュータによって実行されたときに請求項 1 ~ 17 のいずれか一項に記載の方法を前記コンピュータに実行させる、不揮発性のコンピュータ読み取り可能な媒体上に記録されたソフトウェア、すなわちコンピュータ命令。

10

【請求項 19】

被写体の画像の品質を評価するためのコンピュータであって、

同じタイプの複数の異なる被写体の複数の訓練画像の組み合わせに基づいて決められた前記被写体の包括的なモデルを記憶するためのコンピュータ記憶装置において、前記モデルが部分モデルを有し、前記部分モデルの各々が複数の訓練画像からの複数の部分画像からモデル化され、前記部分モデルの各々が前記画像の前記部分画像に対応する、コンピュータ記憶装置と、

前記画像の 2 つ以上の部分画像の特徴と前記モデルとの類似度を判定するためのプロセッサにおいて、類似度の判定が、各部分画像の特徴とそれに対応する部分モデルによってモデル化された特徴との類似度に基づいて決められ、前記判定された類似度に基づいて前記画像の品質を判定するためのプロセッサと、を備える、コンピュータ。

20

【発明の詳細な説明】

【技術分野】

【0001】

(参照による援用)

2011年2月18日に出願された「顔選択のためのパッチに基づく確率論的画像品質の評価および改良型のビデオに基づく顔認識 (Patch-based Probabilistic Image Quality Assessment for Face Selection and Improved Video Based Face Recognition)」という名称であるオーストラリア仮特許出願第 2011900557 号の開示を、参照としてここに援用する。2011年12月21日に出願された「画像品質の評価 (Image quality assessment)」という名称であるオーストラリア仮特許出願第 2011905341 号の開示もまた参照としてここに援用する。

30

【0002】

この開示は画像品質の評価に関する。詳細には、画像品質を評価するための、コンピュータで実施される方法、ソフトウェア、およびコンピュータについて記述されている。非限定的な例として、顔の画像品質は、顔認識に使用するための画像の適合性を示す。

【背景技術】

【0003】

画像処理は、例えば画像の強調、被写体の分類、識別などができるように、画像内の被写体を検出することを含む。

40

【0004】

顔認識は、画像処理の重要な領域となっている。カメラによって取り込まれた画像内の顔を自動検出するための技術は、バイオメトリック・アクセス・コントロールや、セキュリティ監視、画像位置合わせ、画像のタグ付けなど、多数の適用先があ。

【0005】

顔認識分野における主要な課題の 1 つは、理想的かつ制御可能な状態でセットアップされていないカメラによって元画像が取り込まれることである。得られる画像には、しばしば、低い基礎解像度、ブレ、姿勢の大きなばらつき、低いコントラストおよび頻繁な照明変化が存在する。その結果、同じ人物の複数の画像品質が大きく変動する可能性がある。この品質の変動は、(1) 画像固有の品質 (解像度や鮮鋭度、コントラスト、圧縮アーチ

50

ファクトなど)のばらつきと、(2)顔固有の品質(顔幾何形状や姿勢、検出可能な視野角/照明角度など)のばらつきを含む。

【発明の概要】

【発明が解決しようとする課題】

【0006】

現在の顔認識のための手法の大部分は、既知の顔画像を含んだ参照ギャラリーを使用するものであり、その目的は、調査画像の顔と整合するギャラリーの顔を確実に見つけ出すことである。当然ながら、ギャラリーの顔と調査画像の顔が異なる品質(解像度や鮮鋭度、コントラスト、圧縮アーチファクト、顔幾何形状、姿勢、検出可能な視野角/照明角度など)で取り込まれても、この整合を見出す必要がある。

10

【課題を解決するための手段】

【0007】

第1の態様では、被写体の画像品質を評価するための、コンピュータで実施される方法が提供され、上記方法は、

上記画像の2つ以上の部分画像の特徴と上記被写体の包括的なモデルとの類似度を判定するステップにおいて、上記包括的なモデルが同じタイプの複数の異なる被写体の複数の訓練画像(training images)の組み合わせに基づいて決められ、前記モデルが部分モデルを有し、前記部分モデルの各々が複数の訓練画像からの複数の部分画像からモデル化され、前記部分モデルの各々が前記画像の前記部分画像に対応し、類似度の判定が、各部分画像の特徴とそれに対応する部分モデルによってモデル化された特徴との類似度に基づいて決められるステップと、

20

上記画像の品質を、上記2つ以上の部分画像の上記判定された類似度に基づいて判定するステップと、を含む。

【0008】

1つの利点は、画像と、同一したがって単一の包括的モデルとの類似度のみから画像品質を判定できるので、品質評価に入力パラメータが必要ない点である。他の利点は、この方法では、被写体の実物の要素(顔の場合は目や鼻など)を検出する必要がない点である。その結果、他の利点は、この方法が、各要素の検出が不可能である低い解像度画像など、低い品質画像の使用に適している点である。

【0009】

30

1つの利点は、同じタイプの異なる被写体の複数の画像から1つのモデルが構築される点である。その結果、類似度の判定が、実物の参照画像との一連の比較に基づいて行われない。これにより、計算時間が短縮され、リアルタイムの品質判定が可能になる。例えば、この計算時間は、ストリーミング・ビデオの次のフレームとの間の時間よりも短い。計算時間は、この方法のスケーリング(scaling)にとっても重要になる。

【0010】

1つの利点は、部分画像を評価することによって、類似度の判定ごとに考慮する必要があるピクセル数が低減される点である。その結果、メモリの必要量と計算時間が低減される。他の利点は、個々の部分画像の計算を並列化して、現行のマルチコア・コンピュータ・アーキテクチャを最大限に利用できる点である。

40

【0011】

場合により、上記方法は画像の前処理ステップとして使用され、場合により、判定された画像品質に基づいて追加処理が行われる。1つの利点は、より優れた品質の追加処理用画像の選択に、上記の品質判定を使用することができ、それにより追加処理の精度が向上する点である。

【0012】

場合により、各訓練画像が、品質を決める複数の好ましい特性を有する。他の利点は、新たな好ましい特性を有する訓練画像の異なるセットを使用するだけで、同じ方法を使用して、異なる品質(つまり異なる前処理の目的に適合するように異なる特性セットに基づいて決められた品質)を判定できる点である。

50

【 0 0 1 3 】

場合により、上記好ましい特性は、所定の被写体内移動、所定の回転、所定のスケール、所定の解像度、所定の姿勢、所定の照度であり、上記被写体が顔である場合である。

【 0 0 1 4 】

場合により、各部分モデルは平均ベクトルおよび分散行列に基づいて決められ、この平均ベクトルおよび分散行列は、上記部分モデルが基にしている各訓練画像の部分画像の特徴に基づいて決められる。1つの利点は、モデルが2つのパラメータだけに基づいて決められる点である。その結果、類似度を判定する複雑さが低減される。

【 0 0 1 5 】

場合により、部分画像の特徴が、上記部分画像のより低い周波数成分のみにほぼ基づいて決められる。場合により、上記モデルが、訓練画像のより低い周波数成分のみにほぼ基づいて決められる。画像の部分画像とモデルとの類似度の判定は、画像の部分画像のより低い周波数成分と上記モデルとの類似度のみに基づいて決めることができる。1つの利点は、より低い周波数成分のみを使用することにより、画像の全成分の使用と比較して複雑さが低減される点である。より低い周波数成分のみを使用することによりまた、より認識に有用な識別的成分ではなく、記述のための生成的 (generative) 成分に上記方法が絞られる。品質の確実な判定に、少数の成分しか (3つなど) 必要ない。さらに、低い周波数成分を使用すると、主として高い周波数帯域にある顔の表情の変化に対する感度が低減される。

10

【 0 0 1 6 】

場合により、特徴が Haar-like 特徴、LBP または SIFT に基づいて決められる。1つの利点は、上記方法が、所与の状況にいつそう適し得る他の特徴生成方法との併用に柔軟に対応できる点である。

20

【 0 0 1 7 】

場合により、部分画像のより低い周波数成分が、離散コサイン変換に基づいて決められる。

【 0 0 1 8 】

場合により、被写体が顔である。

【 0 0 1 9 】

場合により、部分モデルによってモデル化された上記特徴が、上記部分モデルが基にしている各訓練画像の部分画像の上記特徴の平均値である。

30

【 0 0 2 0 】

場合により、類似度の判定が、部分モデルが基にしている各訓練画像の部分画像の特徴に基づいて決められる平均値および分散行列に特徴付けられるガウス確率密度関数に基づいて決められる。

【 0 0 2 1 】

場合により、類似度の判定が、上記ガウス確率密度関数に基づいて上記画像の上記部分画像の上記特徴の確率を判定するステップを含む。1つの利点は、ガウス確率密度関数が、部分画像の特徴の確率を正確に推定する簡単な方法である点である。他の利点は、ガウス確率密度関数が、異なる訓練画像または他の入力パラメータの変更に合わせて変換する必要がない値を返す点である。したがって、品質の閾値は固定したままにでき、異なる状況に適合させる必要がない。

40

【 0 0 2 2 】

場合により、画像の2つ以上の部分画像とモデルとの類似度を判定するステップが、部分画像ごとに独立して実行される。

【 0 0 2 3 】

類似度の判定が、部分画像ごとに1つの類似度基準判定するステップを含み、上記画像品質の判定が、上記2つ以上の部分画像の上記判定された類似度の結合に基づいて決められる。他の利点は、単一の品質スコアが直接判定される点である。その結果、融合法 (fusion method) は必要ない。融合法では、例えばコントラストや鮮鋭度、解像度、幾何形

50

状、姿勢、照射角などの異なる画像品質面が異なるアルゴリズムによって別々に測定され、1つの品質スコア生成するために融合される。

【0024】

場合により、全ての部分画像の類似度を結合するステップが、部分画像ごとに独立して確率を判定するステップと、画像の部分画像ごとの確率に基づいて結合確率を判定するステップを含む。1つの利点は、部分画像ごとの確率を独立して判定することによって、これらの確率を単に掛けるだけで結合確率を判定できる点である。この乗算は対数の合計に都合よく変換することができ、その結果、計算の複雑さが低減する。

【0025】

画像の部分画像は、部分モデルが基にしている各訓練画像の部分画像に対応する。典型的には、画像の部分画像は位置合わせされるはず、つまり、その画像内の、対応する部分モデルが基にしている訓練画像の部分画像と同じ位置に、位置合わせされるはずである。画像の部分画像間の空間的關係は、訓練画像の部分画像間の空間的關係と同一である。

10

【0026】

場合により、上記方法がさらに、対数正規化関数に基づいて画像を正規化するステップを含む。1つの利点は、対数正規化関数によって、より小さい輝度値は増幅され、より大きな輝度値は圧縮され、それによって上記方法の効率が高められる点である。

【0027】

場合により、上記方法がさらに、上記画像の上記判定された品質の指標を出力として提供するステップを含む。場合により、上記方法がさらに、上記画像の上記判定された品質の指標を揮発性メモリに記憶するステップを含む。

20

【0028】

場合により、上記方法がさらに、複数の画像のそれぞれの品質を判定するために、上記複数の異なる画像ごとに上記方法を繰り返すステップと、

上記判定された品質に基づいて上記画像の追加処理を実行するステップと、を含む。

【0029】

場合により、追加処理が、より高い品質を有する画像に、より大きな重みを与えるステップを含む。例えば、上記方法は、より高い品質を有する複数の画像のサブセットを識別するステップと、複数の顔画像のサブセットに対して追加処理を実行するステップとを含むことができる。別の例では、追加処理により、より良い品質を有する画像をより強調することができる。

30

【0030】

他の例では、場合により、所定の数のより良好な品質を有する画像が追加処理のために識別されるように、上記画像がランキングに基づいて決められる。1つの利点は、追加処理が最も機能する画像数を事前に決めることができる点である。

【0031】

場合により、追加処理が、上記の画像のサブセットなどに対して顔認識または顔照合を実行する。1つの利点は、顔認識のために複数の画像の全てではなく、判定された品質に基づいて決められたより少ない数しか考慮する必要がない点である。その結果、低い品質の顔画像によって顔認識の結果が損なわれることはない。さらなる結果として、多数の低い品質の顔画像とほんの少数の高い品質の顔画像からなるセットにおいて、高い品質の顔画像が外れ値として廃棄されることがなくなる。

40

【0032】

場合により、上記の顔画像のサブセットの選択が、閾値よりも良い品質を有する顔画像のみが選択されるように品質閾値に基づいて決められる。1つの利点は、顔画像を受け取りながら、その画像を追加処理のために使用するか廃棄するかどうか顔画像ごとに独立して決定することができる点である。

【0033】

場合により、画像は、ビデオのフレームなどの一続きの画像からのものである。

【0034】

50

場合により、上記方法がさらに、フレーム内の被写体を検出し、上述の方法で使用するための画像を形成するために、検出された被写体に基づいてフレームを切り出す(crop)最初のステップを含む。

【0035】

場合により、上記方法がさらに、画像からの2つ以上の部分画像を識別するステップを含む。

【0036】

第2の態様では、コンピュータによって実行されたときに上述の方法を上記コンピュータに実行させる、不揮発性のコンピュータ読み取り可能な媒体上に記録されたソフトウェア、すなわちコンピュータ命令も提供される。

10

【0037】

第3の態様では、被写体の画像の品質を評価するためのコンピュータが存在し、上記コンピュータは、

同じタイプの複数の異なる被写体の複数の訓練画像の組み合わせに基づいて決められた上記被写体の包括的なモデルを記憶するためのコンピュータ記憶装置において、前記モデルが部分モデルを有し、前記部分モデルの各々が複数の訓練画像からの複数の部分画像からモデル化され、前記部分モデルの各々が前記画像の前記部分画像に対応する、コンピュータ記憶装置と、

上記画像の2つ以上の部分画像の特徴と上記モデルとの類似度を判定するためのプロセッサにおいて、類似度の判定が、各部分画像の特徴とそれに対応する部分モデルによってモデル化された特徴との類似度に基づいて決められ、上記判定された類似度に基づいて上記画像の品質を判定するためのプロセッサと、を備える。

20

【0038】

上述の第1の態様の任意選択的な特徴は、等しく第2および第3の態様の任意選択の特徴になる。

【0039】

次に、本発明の例について説明する。

【図面の簡単な説明】

【0040】

【図1】ストリーミング・ビデオまたはビデオ・シーケンスから顔認識する方法を図示する。

30

【図2】パッチへの顔画像の分割を図示する。

【図3】2D-DCT基底関数の行列を図示する。

【図4】参照ギャラリー、3次元特徴ベクトルのセットおよび包括的な確率論的顔モデルを図示する。

【図5】上記方法を実行するためのコンピュータ・システムを図示する。

【図6】顔認識のためのワークフローを図示する。

【発明を実施するための形態】

【0041】

この例は、ストリーミング・ビデオから顔の認識を行う(つまり被写体が顔である)ための画像品質の評価に関する。図1にこの例の方法100のフローチャートが示されている。

40

【0042】

図5に、方法100を実行できるコンピュータ・システムが示されている。コンピュータ500は、上記コンピュータ・システムを方法100に従って動作させるようにインストールされたアプリケーション・ソフトウェアを有する。コンピュータ500はまた、キーボードやマウスおよび/またはタッチスクリーンなどのパーソナル・コンピュータの典型的な入力手段を有する。コンピュータ500は、異なる人々の顔の訓練画像のライブラリを含んだ内部および/または外部記憶装置506を含む。この例では、訓練画像は、中立的な表情を有する、適切な照明下で適切に位置合わせされた中心に位置する正面の顔画

50

像である。1つのライブラリの例として、F E R E Tデータセットが挙げられる。

【 0 0 4 3 】

方法100は、包括的顔モデルを構築すること(102)によって初期化される。この初期化ステップは一般に、上記方法の残り104~124が、ステップ102で構築された同一の包括的顔モデルに基づいて複数回繰り返されている間に1回実行される。

【 0 0 4 4 】

上記モデルは、同じタイプの異なる被写体をやはり取り込んでいる複数の訓練画像からなるライブラリを参照して構築される。この事例では、被写体のタイプは顔であり、訓練画像はそれぞれやはり単一の顔を取り込んでいる。複数の異なる人々の顔が訓練画像に取り込まれる。

10

【 0 0 4 5 】

品質評価の状況での包括的とは、単一のモデルが、ライブラリ内の複数の訓練顔(典型的には全ての顔)に基づいて構築されること、したがってそのモデルが特定の個人に特有のものではないことを意味する。これは、顔認識の状況での既存の方法とは対照的である。顔認識では、参照ギャラリーが別々の個人の顔を含み、その参照ギャラリーに独立した顔モデルが個人ごとに生成される。その結果、参照ギャラリーが1124個の異なる人の顔を含む場合、1124個の作成された顔モデルが存在し、取り込まれた顔画像が1124個の異なる顔モデルのそれぞれと比較され、1124個の異なる類似度スコアが判定される。最高のスコアを有する顔画像が最良の整合として採用される。品質評価では、1124個の異なるモデルの代わりに、唯一の単一の包括的モデルが訓練画像から構築され、取り込まれた顔画像に対して唯一のスコアが判定される。この例では、このスコアは顔の画像品質を直接反映する。

20

【 0 0 4 6 】

顔の画像は、移動(例えば垂直または水平方向の中心からの変位など)、顔面の回転、スケール、解像度、頭部の姿勢および照度を含んだ特性を有する。この例では、訓練画像は、顔認識解析に使用されるどの画像においても好ましい特性を共有する。つまり、そういった好ましい特性を有する画像は、その好ましい特性を示さない画像よりも品質が良いとみなされる。この例では、訓練画像は正面の顔画像であり、そのそれぞれがほぼ中心に配置され、ほぼ中立的な姿勢および中程度の照度で適切に位置合わせされている。その結果、訓練画像の内容により、何が良好な品質画像なのかを表すものとして、上述の特性が定義される。つまり、上記方法の異なる適用先では、その適用先に特有の品質画像を規定する、異なる特性を有する異なる訓練画像のセットを使用できるということである。

30

【 0 0 4 7 】

図2に、顔201を含んだ画像200が示されている。部分画像(パッチとも呼ばれる)202が、事前定義された経路203に沿ってフレーム200の全体にわたって段階的に移動する。パッチのサイズおよび段階のサイズは、次の段階のパッチが重なり合うようなものにすることができる。1つの例では、画像のサイズは64×64ピクセル、パッチのサイズは8×8ピクセルであり、次の段のパッチは7ピクセルだけ重なり合う。その結果、画像200の各列に64-8+1=59個のパッチが存在し、画像200全体で59×59=3481個のパッチが存在する。ここで注目されるのは、後で取り込まれる調査画像だけでなく訓練画像も、同じサイズに切り出し、サイズ変更されるということである。次いで、それらはさらに、訓練画像のパッチが、取り込まれる画像のパッチと位置合わせされるように、等しいサイズの小さなパッチに分割される。つまり、取り込まれたフレームにおける顔が適切に位置合わせされていない場合でも、切り出し、サイズ変更後のパッチは適切に位置合わせされる。

40

【 0 0 4 8 】

モデルは部分モデルからなる。つまり、包括的パッチモデルがパッチごとに生成される。包括的パッチモデルは、訓練画像のパッチから抽出された特徴に対して統計的なモデル化(多変量ガウスモデルなど)を実行することによって生成される。各パッチの特徴は、離散コサイン変換(discrete cosine transformation、

50

DCT) や主成分分析 (Principal Component Analysis、PCA)、ガボール・ウェーブレット変換 (Gabor Wavelet Transformation) などの任意の適した方法を使用して抽出することができる。この例では DCT が特徴抽出に適用される。

【0049】

図3に、2次元離散コサイン変換 (two-dimensional discrete cosine transformation、2D-DCT) の基底関数である行列300が示されている。この例では、行列300は 8×8 個の異なる基底関数を含む。各訓練画像の各パッチが、行列300を使用してスペクトル解析によって得られるベクトルによって表される。換言すれば、パッチごとに 8×8 個の係数、つまり行列300における 8×8 個の周波数のそれぞれに対して1つの係数が判定される。これらのDCT係数は、64次元の特徴ベクトルに配列される。したがって各パッチは、DCT係数を使用してベクトルに変換される。この明細書全体にわたって、顔面の要素と特徴の混同を回避するために、目や鼻などの顔の部品を顔面要素と呼ぶ。顔面の特徴という呼び方は、顔画像から抽出された、DCT係数などの特徴に用いる。

【0050】

行列300の左上の方のより低い周波数成分302が、粗い顔の外観および記述的な属性 (目や鼻の位置など) の特性を示すことが図3から明らかである。行列300の右下の方のより高い周波数成分は、より精細な顔の識別的 (discriminative) な属性の特性を示す。ここで提案する方法は、取り込まれた顔が包括的な平均顔と類似しているかどうか判定することを目的とするので、顔の識別的な属性すなわちより高い周波数成分に対する関心はあまりない。したがって、少数のより低い周波数成分のみが使用される。1つの例では、行列300の左上における 2×2 のマスの周波数のみが考慮される。DC成分を無視すると、行列300の左上要素により、次元が3の特徴ベクトルが得られる。当然ながら、他の実施形態では、特徴ベクトルはより少ない、またはより多くの次元を有することができる。

【0051】

各訓練画像のパッチごとに、DCTによって特徴ベクトルが判定される。次いで、全ての訓練画像の特徴ベクトルの平均および共分散が算出される。

【0052】

別の例では、特徴を、DCTによって判定せず、Haar-like特徴やローカル・バイナリ・パターン (LBP)、スケール不変特徴変換 (scale-invariant feature transform、SIFT) などの周波数成分に基づいて決めなくてもよい。

【0053】

図4に、記憶装置506に記憶された訓練画像400のライブラリが示されている。訓練画像400は、図2を参照して説明したようにパッチに分割される。例示的パッチ位置における1つのパッチ例402が、各訓練画像上に示されている。以下のステップについてパッチ402の位置を参照して説明するが、そういったステップはパッチ位置ごとに独立して実行されることを理解されたい。

【0054】

訓練画像400のパッチ位置402に対してDCTによって抽出された3次元の特徴ベクトル404のセットと、パッチ位置402のための包括的な確率論的パッチモデル406も示されている。特徴ベクトル404のセットは、各訓練画像400のパッチ位置402ごとに1つの特徴ベクトルを備える。1つの例では、ライブラリは、1124個の異なる訓練画像を備える。包括的な確率論的パッチモデル406は、パッチ位置402に対する特徴ベクトル404のセットの、3次元平均ベクトル408および 3×3 対角分散行列410を含む。残り全てのパッチ位置に対して、同様の包括的な確率論的パッチモデルが作成される。全てのパッチ位置に対する平均ベクトルおよび分散行列が、包括的な確率論的顔モデル412を形成する行列に入力される。パッチ位置402に関連する行列の入力

10

20

30

40

50

が黒ベタで示されている。

【 0 0 5 5 】

各パッチモデルからなる上記の包括的顔モデル 4 1 2 は、ステップ 1 0 4 ~ 1 2 4 によって規定される方法で使用できる状態にあるコンピュータ記憶装置 5 0 6 上に記憶される。1つの例では、包括的顔モデル 4 1 2 が算出されると、訓練画像が削除される。別の例では、高性能コンピュータで包括的顔モデルが 1 回算出され、より性能が低い複数のクライアントデバイスにエクスポートされる。

【 0 0 5 6 】

この例の図 1 における次のステップは、人の顔 5 0 4 を取り込むストリーミング・ビデオのフレームにアクセスするステップ 1 0 4 である。訓練画像と異なり、1 0 4 でアクセスされるフレームは、姿勢のばらつき、影状態の変動、ブレ、位置合わせ誤差（例えば平面内の回転、水平および垂直の移動など）を有する。

【 0 0 5 7 】

この取り出しステップは、ストリーミング・ビデオのフレームが、監視ビデオカメラ 5 0 2 などの取り込みデバイスから利用可能なすぐ次のフレームであるように、リアルタイムで実行することができる。あるいは、一連の画像をスチールカメラによって取り込み、コンピュータ記憶装置 5 0 6 に記憶し、その後以下のステップに従って処理することができる。

【 0 0 5 8 】

フレームの受け取りは、入力ポートへの直接接続によって、またはローカル・コンピュータ・ネットワーク（図示せず）を経由して、またはインターネットなどの広域ネットワーク（図示せず）を経由してなど、カメラ 5 0 2 からコンピュータ 5 0 0 への入力として行われる。1つの代替方法では、コンピュータ 5 0 0 がカメラ 5 0 2 に組み込まれる。あるいは、フレームを事前に記録し、記憶装置 5 0 6 から取り出すことができる。別の例では、同一の人物すなわち被写体をターゲットにした複数のカメラにコンピュータが接続され、少なくとも1つのカメラはその人物の顔の良好な品質の画像を取り込む可能性がより高まるようになっている。

【 0 0 5 9 】

顔検出アルゴリズムが上記フレームに適用される（1 0 6）。1つの例では、OpenCVのViola-Jones Haarに基づく顔検出器が使用されるが、別の例では、任意の適当な顔検出アルゴリズムを使用することができる。ここで注目されるのは、顔検出はフレームにおける顔の位置のみを判定し、その顔と関連付けられた人物の認識は行わないということである。顔検出アルゴリズムはまた、フレームに含まれる顔の位置および大きさをフレームを切り出して顔画像をもたらす。別の例では、顔はフレーム全体に広がり、顔検出アルゴリズムはフレームの切り出しを行わない。その例では、顔画像はフレームと同一である。別の例では、フレームは複数の顔を含み、顔検出アルゴリズムはフレームを別々の顔画像に切り出す。この例では、異なる顔の画像をその後の顔認識時に同一人物に属すると見なすことなく顔ごとの良好な品質の画像を別々に見つけることができるように顔を追跡する必要がある。

【 0 0 6 0 】

次のステップは、顔画像の大きさが訓練画像の大きさと同じであるように、顔画像の大きさを正規化することである。1つの例では、大きさは 6 4 × 6 4 ピクセルである。

【 0 0 6 1 】

次いで上記フレームは、以下の対数正規化関数に従って低輝度のピクセルを増幅し、高輝度のピクセルを圧縮するようにピクセルごとに輝度が正規化される（1 0 8）。

【 0 0 6 2 】

【 数 1 】

$$I_{\log}(r, c) = \ln[I(r, c) + 1]$$

10

20

30

40

50

【 0 0 6 3 】

ここで、 $I(r, c)$ は位置 (r, c) におけるピクセル輝度である。小さな輝度値の勾配は、より大きな輝度値の勾配よりも大きい。その結果、より小さな輝度値は増幅され、より大きな輝度値は圧縮される。この性質は、皮膚の色と背景の照明との輝度差を低減させるのに役立つ。

【 0 0 6 4 】

正規化ステップの後、顔画像は、ここでパッチと呼ばれる部分画像に分割される (110)。顔画像の分割は訓練画像の分割と同様である。同じ数の部分画像が入力画像と訓練画像の両方から定義され、その全てが位置合わせされる。その位置合わせは、ビデオフレームからの顔画像の各パッチが1つの部分モデルに対応し、したがって図4にある包括的顔モデル412の行列における1入力であるようになっている。言い換えれば、入力画像の各パッチ間の空間的關係が訓練画像で使用されるパッチ間の空間的關係と同じになっている。1つの例では、各パッチは、顔画像間のコントラストのばらつきを吸収するように、ゼロ平均および単位分散を有するように正規化される。

【 0 0 6 5 】

パッチごとに、特徴ベクトルつまりDC T係数は、上記でステップ102における訓練画像のパッチに関して説明したように判定される (112)。この例では、3つのDC T係数からなるベクトルがパッチごとに判定されるが、別の例ではそれより多い係数を使用しても、それより少ない係数を使用してもよい。

【 0 0 6 6 】

次のステップは、顔画像の各パッチとそれに対応する包括的顔モデルのパッチとの類似度を判定することである (114)。この類似度は、パッチごとの事後確率 (posterior probability) として表される。事後確率は、そのパッチに対して判定されたベクトルを、包括的顔モデルに基づいて決められる以下のガウス確率密度関数に送ることによって判定される。

【 0 0 6 7 】

【 数 2 】

$$p(\mathbf{x}_i | \lambda_i) = \frac{e^{-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_i)}}{(2\pi)^{\frac{d}{2}} |\boldsymbol{\Sigma}_i|^{\frac{1}{2}}}$$

【 0 0 6 8 】

ここで、 \mathbf{x}_i は位置固有の多変量ガウスモデル、 $\boldsymbol{\mu}_i$ および $\boldsymbol{\Sigma}_i$ は対応する包括的顔モデルのパッチの平均および分散行列である。ステップ102に従って、パッチごとの包括的モデル412は、適切な照明下の中立的な表情を有する正面画像のプールを使用して訓練された。訓練画像の全てと、ビデオフレームからの顔画像は、スケーリングや位置合わせなどによって 64×64 ピクセルという固定サイズに正規化される。

【 0 0 6 9 】

訓練画像などの訓練データから各パッチをモデル化するための他の例は、多変量ガウス分布、ラプラス分布、周波数分布のヒストグラム、ポアソン分布である。一般に、任意の分布モデル生成技術またはそれらの技術の組み合わせをパッチのモデル化に使用することができる。

【 0 0 7 0 】

この段階での目的は、個々の顔を認識することではなく、ビデオフレームからの顔のパッチが同じ位置にある包括的顔モデルで表される平均パッチにどのくらい類似しているかを判定することである。上記公式は、確率論的なモデルの平均 $\boldsymbol{\mu}_i$ からかけ離れたDC T係数を有するパッチには低い確率を割り当て、確率論的なモデルの平均 $\boldsymbol{\mu}_i$ に近いDC T

10

20

30

40

50

係数を有するパッチには高い確率を割り当てる。その結果、事後確率は、ビデオフレームからの顔画像パッチと包括的顔モデル412との類似度を反映する。

【0071】

他の例では、DCT特徴の置き換えに、ローカル・バイナリ・パターン(LBP)や主成分分析(PCA)、ガボール、SIFTなどの他の画像特徴およびそれらの組み合わせを使用することができる。測定に必要な画像の特性が何かに応じて、表情の変化などの異なる画像品質面の特性を示すためにこういった特徴を使用してもよい。

【0072】

他の例では、パッチサイズは様々な値をとることができる。つまり、パッチサイズは画像全体にわたって固定的ではない。パッチサイズは画像の領域別に変化してよい。例えば、目領域の周りのパッチサイズを、口または顎領域の周りのパッチサイズと異なるサイズにすることができる。

【0073】

画像特徴と可変パッチサイズの組み合わせも可能である。画像領域別に、様々な特徴とパッチサイズを使用することができる。例えば、目領域の周りで、パッチサイズ8×8を有するDCT特徴を使用する一方、顎領域など他の顔領域で、パッチサイズ12×10を有するガボール特徴を使用することもできる。

【0074】

次に、画像品質が判定される。つまり、顔画像全体の品質を判定する(116)ために、全てのパッチの事後確率が結合される。パッチの確率が統計的に互いに独立していると仮定すれば、フレーム全体の結合確率を判定するために、これらのパッチ確率を乗算することができる。いくつかの例では、パッチが統計的に独立していない可能性があるが、過度の不正確をもたらずに独立性が近似として仮定される。加算は乗算より計算コストが大きくないことから、確率の対数が使用され、その結果、顔画像Iの品質スコアQを判定するために対数が合計される。

【0075】

【数3】

$$Q(I) = \log[P(\mathbf{X}|\Lambda)] = \frac{1}{N} \sum_{i=1}^N \log[p(x_i|\lambda_i)]$$

【0076】

融合法に頼らずに、上記方法により顔画像ごとの品質スコアが生成される。このスコアは、位置合わせ誤差、姿勢のばらつき、スケール類似度、影および画像鮮鋭度(基礎解像度)の程度を同時に表す。顔面要素(すなわち目、鼻および口)の位置特定もまた、この方法では必要ない。1つの例では、ステップ106における顔検出アルゴリズムは、特に低い解像度画像では不正確であり、その結果、顔画像は数ピクセルだけ位置合わせがずれる。

【0077】

小さなパッチへの分割の性質上、2ピクセルを超える位置合わせずれを確実に検出することができ、その結果、顔検出アルゴリズムの不正確さを発見することができる。パッチは比較的小さいので、各パッチにおけるモデルは、2ピクセルの移動などの小さな変化を検出するのに十分に高い感度を有する。任意選択で、こういった低い品質の画像を、上記方法でその画像に追加解析を実行しないことにより廃棄することができる。同様にして、上記方法を、モデルに対して最も位置合わせされた顔画像の識別に役立てることができる。顔検出器が顔を検出するときに、検出された領域の周囲で数多くの様々な方法により顔画像を切り出すことができる。各方法で、切り出し済みの顔の画像品質が測定される。品質に基づいて最良の切り出し済みの顔画像を選択することによって、位置合わせずれを追加処理する際の誤差を最小限にするために、ここで提案する方法を役立てることができる

【 0 0 7 8 】

姿勢、照度および解像度の変化もここで提案する方法で検出される。

【 0 0 7 9 】

他の例では、品質スコアは、確率の加重和、エントロピー、または確率の対数の加重和に基づいて決められる。一般に、単一のスコアの生成に、確率に基づく任意の融合法またはそういった技術の組み合わせを使用することができる。

【 0 0 8 0 】

次いで、フレームの品質が顔認識に十分かどうか判定するために、品質スコアが閾値と比較される(118)。1つの例では、品質閾値は1.4である。他の例では、品質閾値は、3シグマ閾値 $\mu_g - 3\sigma_g$ などの分布の値に基づいて与えることができる。品質が不十分の場合、フレームは廃棄される(120)。

【 0 0 8 1 】

次いで、十分な数の顔画像が顔認識に使用可能かどうか、またはビデオの終わりに達したかどうか、つまり使用可能な全てのフレームの処理が完了したかどうか判定される(122)。顔画像の必要数は10個など、事前に決められた数であってよい。あるいは、全てのフレームからの全ての顔の画像品質を判定することができ、10個など、最高スコアの顔画像数を選択することもできる。他の代替方法として、品質閾値を超える全ての顔画像を画像認識に使用してもよい。顔画像の使用可能な数が不十分である場合、上記方法は、より多くのフレームを収集するためにステップ102から再スタートする。他の例では、上記システムは、より高い認識品質を有する上位10個の画像を選択することができる。

【 0 0 8 2 】

顔画像が選択されると、選択されたフレームに対して顔認識が実行される(124)。顔認識は、確率論的なマルチ・リージョン・ヒストグラム(Multi Region Histogram、MRH)またはローカル・バイナリ・パターン(Local Binary Pattern、LBP)を使用して、相互部分空間法(Mutual Subspace Method、MSM)に基づいて実行することができる。これらの方法は、上述したように、より識別力のあるより高い周波数成分を使用することができる。

【 0 0 8 3 】

コンピュータ500のディスプレイ上に方法100の結果を表示することができる。例えば、品質表示がグラフで表示されたり、他のフレームと比較されたフレームの数値ランキングが表示されたりする。この品質表示は、簡易的に、品質基準を満たした(108)フレームの表示であってもよい。

【 0 0 8 4 】

別法または追加として、整合済みの顔画像およびその整合済みの画像に示されている人物の情報などの、顔認識の結果を表示することもできる。1つの例では、品質スコアが顔認識アルゴリズムに提供され、分類または顔認識のために特徴を結合するための重みとして使用される。より良好な品質を有する顔画像は、より大きな重みが割り当てられ、これは、良好な品質画像が認識結果に対してより大きく寄与すべきことを示している。

【 0 0 8 5 】

他の例では、上記方法を、所望の特性に合致する画像を選択するために用いることができる。その特性は、包括的モデルの生成に使用される訓練画像に取り込まれている。例えば、訓練画像は、笑顔などの一定の顔の表情を示すことができ、その場合、上記方法を良好な品質を有する、つまり表情が笑顔である画像の識別に使用することができる。このようにして、特性が、ひいては品質の定義が、この方法ではプラグアンドプレイ可能になっている。

【 0 0 8 6 】

図6に、顔認識システムのためのワークフロー600が示されている。このワークフローは、顔検出テストに適合済みの上述したように正規化されたビデオフレームなどの取り

10

20

30

40

50

込まれた画像のセット602を備える。次いで、この取り込まれた顔画像のセット602は、上述した方法に従って画像品質のための評価が行われる(604)。品質評価の結果に基づいて、顔画像が選択され(606)、その結果、顔認識に使用される顔画像のサブセットが得られる(608)。ここで注目されるのは、顔認識を単一の顔画像に対して実行することができるが、顔画像が低品質のものであると、その認識は正確にならないということである。顔認識アルゴリズムの正確さは、使用される顔画像の数と共に高まり、ビデオ監視の状況では、同じ顔の多数の画像が様々なビデオフレームで容易に利用可能である。しかしながら、不十分な品質の顔画像があると、顔認識の品質が損なわれる可能性がある。したがって、上記で提案した方法で、最良の画像品質を有する顔画像のみを選択する。

10

【0087】

顔画像のサブセット608が選択されると、最も適当な顔整合方法が選択され(610)、これはギャラリー画像612に基づいて決められる。顔整合方法を選択する主要な基準は、顔画像がどのくらい適切に位置合わせされているかということである。顔画像が適切に位置合わせされている場合(614)、PCAやLBPのような方法が最も適当であるが、顔画像が適切に位置合わせされていない場合(616)、MRH類の方法が選択されると好ましい。

【0088】

顔整合方法の結果が、ランキング表示などの後処理方法に提供され(618)、最後に顔の識別情報が判定される(620)。

20

【0089】

上記の実施形態は顔画像のための品質評価の適用例について特に論じたが、モデル化できる他のタイプの被写体に適用することもできる。被写体のタイプの例には、道路標識およびナンバープレートが含まれる。当然ながら、被写体は、物理的な3次元の被写体や、画像に現れる着目する被写体を定義するその他任意の特徴の集まりなど、任意の適当なタイプの被写体であってよい。例えば、画像に表示される道路標識や被写体に印刷された商標図案などがある。

【0090】

選択またはフィルタ処理のための使用とは別に、品質スコアを認識などのその後の処理のための画像重要度の重みとして使用することもできる。

30

【0091】

この方法は、テンプレートと位置合わせするための画像の位置合わせのために使用することもできる。これは一連の候補画像を生成することによって行うことができ、最大品質のスコアが最良の位置合わせ選択肢を示す。

【0092】

この方法はまた、顔照合と共に使用することもでき、この顔照合では、顔の画像品質に基づいて閾値を適合するように選択する。

【0093】

本発明で使用されるコンピュータは、カメラ自体の内部に組み込まれたプロセッサなどの任意の適当なコンピュータ、またはネットワークをなど介した分散処理システムにすることができる。一つの代替方法では、コンピュータは、組み込みシステム、つまり特定の制御機能またはタスクのために設計されたコンピュータ・システムにすることができる。これは大型システムの一部として組み込むことができる。組み込みシステムは、典型的にはマイクロ・コントローラまたはデジタル信号プロセッサ(DSP)である処理コアを含む。こういった組み込みシステムは、タスク固有であるので、サイズおよびコストを低減するように設計することができ、したがって一般に、携帯電話や様々なタイプのカメラなどの携帯デバイスの領域に見られる。

40

【0094】

ここで説明した技術は、種々の技術を使用して実施できることを理解されたい。例えば、ここで説明した方法は、適当なコンピュータ読み取り可能な媒体上にある一連のコンピ

50

ユーザ実行可能な命令によって実施することができる。適当なコンピュータ読み取り可能な媒体は、揮発性（例えばRAMなど）および/または不揮発性（例えばROM、ディスクなど）メモリ、搬送波および伝送媒体（例えば銅線、同軸ケーブル、光ファイバ媒体など）を含むことができる。搬送波の例には、ローカルネットワークまたはインターネットなどの公衆がアクセス可能なネットワークに沿ってデジタルデータ流を搬送する電気信号、電磁信号または光信号の形をとるものが挙げられる。

【0095】

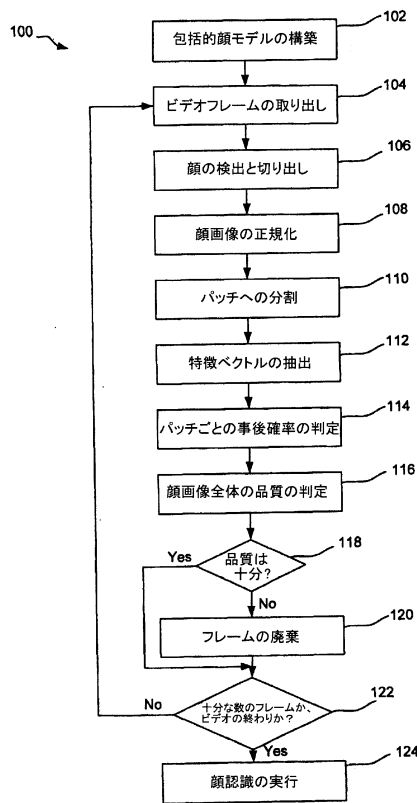
特に明記されていない限り、以下の説明から明らかなように、上記説明全体を通して、「処理する」、「構築する」、「算出する」、「計算する」、「判定する」、「表示する」などの用語を用いた記述は、コンピュータ・システムまたは同様の電子計算デバイスの動作および処理を示すと認識されていることも理解されたい。コンピュータ・システムは、そのコンピュータ・システムのレジスタおよびメモリ内部の物理的（電子的）な量として表されるデータを処理し、そのコンピュータ・システムのメモリまたはレジスタまたはその他のかかる情報記憶装置、伝送または表示デバイス内部の物理的な量として同様に表される他のデータに変換する。

【0096】

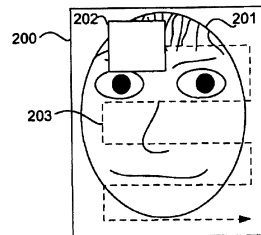
したがって、本明細書の各実施形態は、あらゆる点で、例示的かつ非限定的であるとみなされるべきである。

10

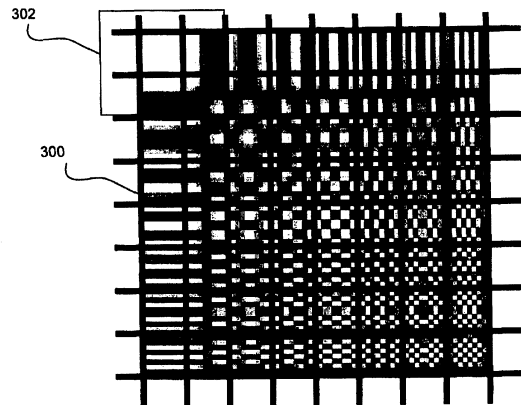
【図1】



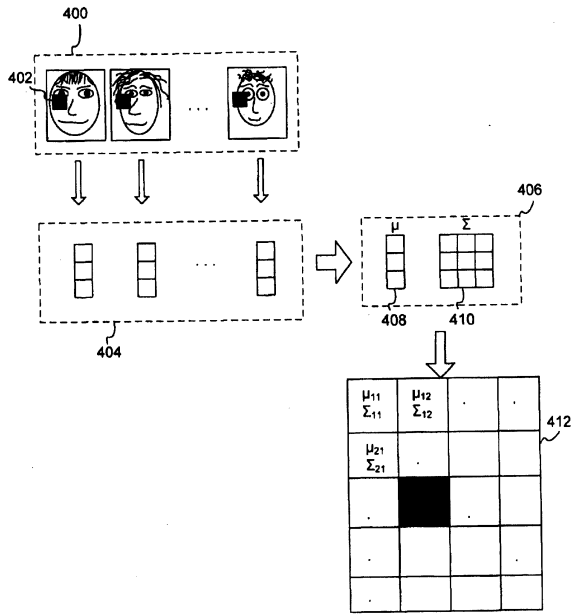
【図2】



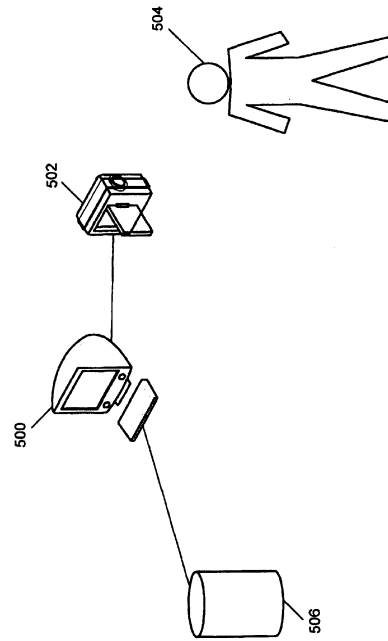
【図3】



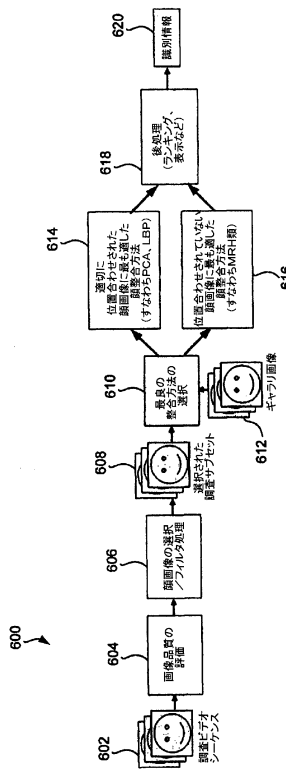
【 図 4 】



【 図 5 】



【 図 6 】



フロントページの続き

(72)発明者 チェン, シャオカン
オーストラリア国, 2015, ニュー サウス ウェールズ州, エヴェレイ, ガーデン ストリー
ト 13, レヴェル 5 気付

(72)発明者 ウォン, ヨンカン
オーストラリア国, 2015, ニュー サウス ウェールズ州, エヴェレイ, ガーデン ストリー
ト 13, レヴェル 5 気付

審査官 板垣 有紀

(56)参考文献 特開2002-032766(JP, A)
特開2001-167110(JP, A)
特開2003-296735(JP, A)
特開2011-019192(JP, A)

(58)調査した分野(Int.Cl., DB名)
G06T 7/00
G06T 1/00