

(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2008-117391
(P2008-117391A)

(43) 公開日 平成20年5月22日 (2008.5.22)

(51) Int.Cl.	F I	テーマコード (参考)
GO6T 1/00 (2006.01)	GO6T 1/00 340A	5B057
GO6T 7/00 (2006.01)	GO6T 7/00 200C	5L096

審査請求 未請求 請求項の数 31 O L (全 24 頁)

(21) 出願番号	特願2007-278534 (P2007-278534)	(71) 出願人	000002369
(22) 出願日	平成19年10月26日 (2007.10.26)		セイコーエプソン株式会社
(31) 優先権主張番号	11/556082		東京都新宿区西新宿2丁目4番1号
(32) 優先日	平成18年11月2日 (2006.11.2)	(74) 代理人	100095728
(33) 優先権主張国	米国 (US)		弁理士 上柳 雅誉
		(74) 代理人	100107261
			弁理士 須澤 修
		(74) 代理人	100127661
			弁理士 宮坂 一彦
		(72) 発明者	ジュウウィ ルー
			カナダ オンタリオ トロント クウイー
			ンズ クウェイ ダブリュ 600-10
			23
		最終頁に続く	

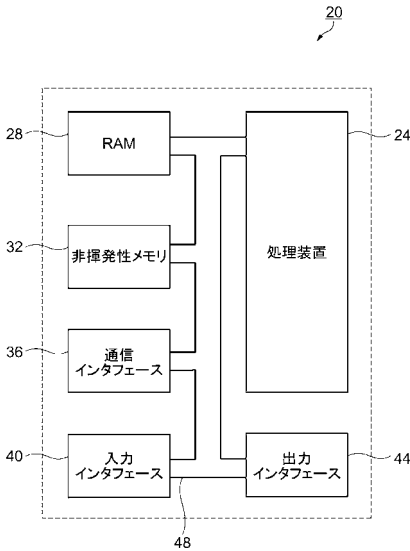
(54) 【発明の名称】 デジタル画像における顔を検出する方法、およびデジタル画像における顔を検出する装置

(57) 【要約】

【課題】 デジタル画像の顔を検出する方法および装置。

【解決手段】 デジタル画像における顔を検出する方法はデジタル画像の部分ウィンドウを選択することを含む。次に部分ウィンドウのサンプル領域が選択される。サンプル領域を分析し、部分ウィンドウが顔を表し得るか判定される。

【選択図】 図 1



【特許請求の範囲】**【請求項 1】**

デジタル画像における顔を検出する方法で、
前記デジタル画像の部分ウィンドウを選択するステップと、
前記部分ウィンドウにおいてサンプル領域を選択するステップと、
前記サンプル領域を分析して前記部分ウィンドウが顔を表し得るか判定するステップと

、
を含むことを特徴とするデジタル画像における顔を検出する方法。

【請求項 2】

前記サンプル領域は矩形であることを特徴とする請求項 1 に記載のデジタル画像における顔を検出する方法。 10

【請求項 3】

前記サンプル領域を選択する前に、前記部分ウィンドウをフレームに分割し、前記サンプル領域各々は前記フレームの異なったフレームに位置するステップを含むことを特徴とする請求項 1 に記載のデジタル画像における顔を検出する方法。

【請求項 4】

前記サンプル領域は前記フレームの前記境界からオフセットされていることを特徴とする請求項 3 に記載のデジタル画像における顔を検出する方法。

【請求項 5】

前記部分ウィンドウは少なくとも 4 つのフレームに分割されることを特徴とする請求項 3 に記載のデジタル画像における顔を検出する方法。 20

【請求項 6】

前記サンプル領域は前記部分ウィンドウの所定位置にあることを特徴とする請求項 1 に記載のデジタル画像における顔を検出する方法。

【請求項 7】

前記サンプル領域はパターンを形成することを特徴とする請求項 1 に記載のデジタル画像における顔を検出する方法。

【請求項 8】

前記パターンは格子パターンであることを特徴とする請求項 7 に記載のデジタル画像における顔を検出する方法。 30

【請求項 9】

前記分析はアダプスト方法を用いて前記サンプル領域を処理するステップを含むことを特徴とする請求項 1 に記載のデジタル画像における顔を検出する方法。

【請求項 10】

前記画像にわたり前記部分ウィンドウをパンするステップと、前記画像内における前記部分ウィンドウの各位置に対し前記選択および分析を再実施するステップと、を含むことを特徴とする請求項 1 に記載のデジタル画像における顔を検出する方法。

【請求項 11】

前記画像にわたり前記部分ウィンドウがパンされた後に、前記部分ウィンドウのスケールを調節し、前記パンおよび前記再実施のステップを繰り返すステップとを含むことを特徴とする請求項 10 に記載のデジタル画像における顔を検出する方法。 40

【請求項 12】

前記調節は前記部分ウィンドウが最小閾値サイズに縮小されるまで再実施されることを特徴とする請求項 11 に記載のデジタル画像における顔を検出する方法。

【請求項 13】

前記分析の際、前記サンプル領域は一連の処理段階にかけられ、前記部分ウィンドウにおける顔の存在を検出し確認することを特徴とする請求項 11 に記載のデジタル画像における顔を検出する方法。

【請求項 14】

前記処理段階は、肌色による分類、エッジ強度による分類、およびアダプストによる 50

分類の少なくとも2つを含むことを特徴とする請求項13に記載のデジタル画像における顔を検出する方法。

【請求項15】

前記部分ウィンドウにおける顔の存在を検出するために肌色による分類が用いられることを特徴とする請求項14に記載のデジタル画像における顔を検出する方法。

【請求項16】

前記部分ウィンドウにおける顔の存在を確認するためにエッジ強度による分類が用いられることを特徴とする請求項15に記載のデジタル画像における顔を検出する方法。

【請求項17】

前記部分ウィンドウにおける顔の存在を再確認するためにアダプストによる分類が用いられることを特徴とする請求項16に記載のデジタル画像における顔を検出する方法。

【請求項18】

デジタル画像における顔を検出する装置で、
前記デジタル画像の部分ウィンドウを選択する部分ウィンドウ選択器と、
前記部分ウィンドウ内のサンプル領域を選択するサンプル領域選択器と、
前記サンプル領域を分析して前記部分ウィンドウが顔を表し得るか判定するサンプル領域分析器と、

を含むことを特徴とするデジタル画像における顔を検出する装置。

【請求項19】

前記サンプル領域選択器は前記部分ウィンドウをフレームに分割し、前記サンプル領域の各々を前記フレームの異なったフレームから選択することを特徴とする請求項18に記載のデジタル画像における顔を検出する装置。

【請求項20】

前記サンプル領域は前記フレームの前記境界からオフセットされていることを特徴とする請求項19に記載のデジタル画像における顔を検出する装置。

【請求項21】

前記サンプル領域は前記部分ウィンドウにおける所定位置にあることを特徴とする請求項18に記載のデジタル画像における顔を検出する装置。

【請求項22】

前記サンプル領域はパターンとして選択されることを特徴とする請求項19に記載のデジタル画像における顔を検出する装置。

【請求項23】

前記パターンは格子パターンであることを特徴とする請求項22に記載のデジタル画像における顔を検出する装置。

【請求項24】

前記サンプル領域分析器は前記サンプル領域を一連の処理段階にかけ、前記部分ウィンドウにおける顔の存在を検出し確認することを特徴とする請求項19に記載のデジタル画像における顔を検出する装置。

【請求項25】

前記一連の処理段階は肌色による分類、エッジ強度による分類、およびアダプストによる分類の少なくとも2つを含むことを特徴とする請求項24に記載のデジタル画像における顔を検出する装置。

【請求項26】

デジタル画像における顔を検出する方法で、
前記デジタル画像の部分ウィンドウを選択するステップと、
前記部分ウィンドウの領域を選択するステップと、
前記部分ウィンドウの前記領域をフレームの二次元アレーに分割するステップと、
フレームの前記二次元アレーを分析し前記領域の各々に対し特徴を生成するステップと、
前記特徴を用い、前記部分ウィンドウは顔を表し得るか判定するステップと、

10

20

30

40

50

を含むデジタル画像における顔を検出する方法。

【請求項 27】

前記区域は少なくとも 4 つのフレームに分割されることを特徴とする請求項 26 に記載のデジタル画像における顔を検出する方法。

【請求項 28】

前記分析は、

前記フレームの各々のピクセル特性を閾値化して各フレームに対しバイナリ・マップを生成するステップと、

各フレームの前記バイナリ・マップの総和に対し関数を実施することにより前記特徴を生成するステップと、

を含むことを特徴とする請求項 27 に記載のデジタル画像における顔を検出する方法。

【請求項 29】

前記特性はピクセル輝度であることを特徴とする請求項 28 に記載のデジタル画像における顔を検出する方法。

【請求項 30】

前記特性は前記ピクセルの色値であることを特徴とする請求項 28 に記載のデジタル画像における顔を検出する方法。

【請求項 31】

前記特性は前記ピクセルのエッジ強度であることを特徴とする請求項 31 に記載のデジタル画像における顔を検出する方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は一般的に画像処理に関し、特にデジタル画像において顔を検出する方法および装置に関する。

【背景技術】

【0002】

分類および認識システムは日常的にデジタル画像を処理し、顔などその中の特徴を検出している。デジタル画像における顔認識はパターン認識が係わるニクラス（顔または非顔）の分類問題である。しかし顔を表すパターンを認識することは、顔を表すパターン同士がしばしばかなり異なっており、容貌、照明、表情、および他の要因により通常極めて複雑であるため難問となる。その結果、画像における顔を検出するために用いられる方法は精度を向上させるために非常に複雑になっている。

【0003】

例えば、画像における顔検出に顔 / 非顔分類器のカスケードを用いる、学習に基づく方法が提案されている。これらの学習に基づく方法は画像訓練プロセスを通じて弱分類器（weak classifier）を学習し、学習した弱分類器を用いてより強分類器（stronger classifiers）を構築する。学習に基づく方法の一つは P. Viola および M. Jones 共著の「非対称 AdaBoost および検出カスケード」（Proc. Of Neural Information Processing Systems、カナダ・バンクーバー市、2001 年）と題される出版で提案されたアダブースト（AdaBoost）手法である。アダブースト手法は非常に複雑な分類器においてパターンを認識するのに特に適している。アダブースト手法は弱分類器のシーケンスを学習し、弱分類器を直線的に組み合わせ、強分類器を構築することにより弱分類器が指標として機能する能力を高める。弱分類器を直線的に組み合わせるために、処理される各画像における隣接矩形領域の小セットからハール（Haar）または類似した特徴が抽出される。各領域においてすべてのピクセルが分析されるので、この手法はプロセッサおよび時間集約的である。

【0004】

学習に基づく顔検出方法の明らかな利点にも拘わらず、カスケードの後期になると弱分類器が弱くなり過ぎるため、より高い性能の達成が制限される。現在の学習に基づく方法

10

20

30

40

50

は分類器を再訓練するために使用される非顔例（誤り検出）を収集するのにブートストラップを用いる。しかし強分類器の力がある点に達すると、ブートストラップで得られた非顔例は顔パターンに非常に類似するようになり、従って分類器を再訓練できなくなる。ハール状の、特徴に基づく弱分類器の分類エラーは50%に近いことが経験的に示され、従ってブートストラップは実際面で有効でなくなる。

【0005】

デジタル画像における顔検出の他の手法も提案されている。例えば、Zhang他の米国特許出願番号第2004/026474号明細書は画像内に複数の初期候補ウィンドウが確立される顔検出方法を開示している。初期候補ウィンドウ各々に対し色空間情報を用いて各ピクセルを肌色ピクセルまたは非肌色ピクセルのいずれかに分類する。肌色ピクセルおよび非肌色ピクセルの数に基づき、少なくとも1つの候補ウィンドウが非顔ウィンドウと分類される。分類された候補ウィンドウに対し信頼得点が判定され、信頼得点に基づき少なくとも1つの空間的に近接する候補ウィンドウの分類を選択的に飛ばすことができる。

10

【0006】

Chen他の米国特許出願番号第2004/017971号明細書は一連のカスケードされたテストが用いられる顔検出方法およびシステムを開示している。各テストは高い信頼性で非顔オブジェクトを切り捨て、殆どの顔を保持する。第1の色度テストは飽和した緑色および青色ピクセルなど、非肌色ピクセルを切り捨てる。続く幾何学テストで、ピクセルは色度に基づきグループ化され、幾何学的形状、サイズ、および位置が調べられる。続く格子パターン・テストで、平均格子パターン要素画像が幾何学テストから得られた残りの領域と比較される。格子パターン・テストに合格した部分画像は顔候補として印が付けられる。顔候補は続いて位置テストで確認され、密接した顔候補が組み合わせられて1つの顔候補になる。

20

【0007】

Xiao他の米国特許出願番号第2005/0013479号明細書は多段階顔検出方法を開示している。第1段階の際、線形フィルタリングを用いて画像の非顔的部分を取り除く。特に、画像の事前フィルタにアダプスト学習方法が用いられる。第2段階において、推進チェーンを導入して階層「チェーン」構造内の推進分類器が組み合わせられる。第3段階で、画像の前処理、SVMフィルタリング、および色フィルタリングを用いて事後フィルタリングが行われる。

30

【0008】

Morisadaの米国特許出願番号第2005/0069208号明細書はテンプレート照合を用いて顔候補を選ぶ、画像内の顔を検出する方法を開示している。各顔候補は次にサポート・ベクター・マシンによるパターン認識を用いて判定される。画像の肌色領域を特定の上、顔候補と照合して肌色の望ましいレベル未満のものを除外する。非顔を表すとみなされた候補は次に取り除かれる。

【0009】

Smith他の米国特許第5870138号明細書は顔内容に対応する画像部分をみつけ、置換する方法およびシステムを開示している。画像は比色分析され、顔の輪郭、唇、眼、等々の顔の特徴を特定する。特定された特徴の位置に対応する特徴抽出データを含む追跡信号が発信される。望まれる場合、特徴抽出データを用いて画像に存在する元の顔を代わりの顔で置換することができる。

40

【0010】

Kreschの米国特許第6463163号明細書は候補選択器が、候補選択器により選択された候補領域が実際に顔内容を含むか確認する顔検出器と連動して動作する顔検出システムおよび方法を開示している。候補選択器は入力画像から候補領域を選択するための直列で動作する線形照合フィルタおよび非線形フィルタを含む。線形照合フィルタはフィルタリング・カーネルを用いて入力画像に線形相関を行い、相関画像を導き出す。極大を有し、相関閾値を超えた相関値を有する領域が選択される。非線形フィルタが次に線形照合フ

50

イルタにより選択された画像領域の各種部分領域からのコントラスト値を調べ適切な候補領域を選別する。顔検出器は選択された領域に顔が含まれるか否かの判定にニューラル・ネットワークを用いる。

【 0 0 1 1 】

Abdel-Mottaleb他の米国特許第 6 5 7 4 3 5 4 号明細書は画像における顔を検出するシステムおよび方法を開示しており、肌色のピクセルがグループ化され、ピクセル・グループのエッジが取り除かれる。残りのグループ化された肌色のピクセルは分析され、顔を含むか判定される。分析はピクセル・グループのバウンディング・ボックス範囲、アスペクト比、検出された肌色のバウンディング・ボックスに対する比率、引き延ばしたオブジェクトの方位、およびバウンディング・ボックスの中心とバウンディング・ボックス範囲の重心との距離の判定を含む。

10

【 0 0 1 2 】

Ho他の米国特許第 6 6 6 1 9 0 7 号明細書は画像における顔を検出する方法を開示しており、画像は類似色の領域に分割される。顔検出分析は肌色領域に対してのみ行われる。

【 0 0 1 3 】

Tian他の米国特許第 6 8 7 9 7 0 9 号明細書は画像において中立的で無表情の顔を検出するシステムおよび方法を開示している。画像における顔のポーズと位置を検出し、顔の構成要素を見つけるために顔検出器が用いられる。画像で顔が検出されると、検出された顔は標準サイズの顔に正規化される。次に幾何学的な目鼻立ちのセットおよび口の区域におけるヒストグラムが 3 つ抽出される。目鼻立ちは分類器に供給され、検出された顔が中立的で無表情であるか判定される。

20

【 0 0 1 4 】

Matsuo他の米国特許出願番号第 2 0 0 2 / 0 1 9 1 8 1 8 号明細書は顔検出方法およびシステムを開示しており、画像に対しエッジ抽出が行われ、エッジ画像が作り出される。顔画像を含む候補である部分画像がエッジ画像から抽出される。顔検出は部分画像に対し学習辞書を用い部分画像が顔画像を含むか否かを検出することにより行われる。

【 0 0 1 5 】

Lestideauの米国特許出願番号第 2 0 0 3 / 0 0 5 3 6 8 5 号明細書は画像における顔を検出する方法を開示しており、人の肌である可能性が高い画像の部分が特定される。次に特定された部分に対しバウンディング・ボックスが判定される。バウンディング・ボックス内における区域の特徴を分析して高レベルのテクスチャが存在するか判定される。バウンディング・ボックス内で高レベルのテクスチャを有する区域が検出された場合、その区域は人の顔を表していないとみなされる。

30

【 0 0 1 6 】

Huang他の米国特許出願番号第 2 0 0 5 / 0 1 4 7 2 9 2 号明細書は画像に描かれる人とその顔のポーズを特定する顔検出システムおよび方法を開示している。顔領域が抽出され、画像の正規化に始まって前処理される。画像で顔が見つかった場合、顔はクロップされる。次に顔は顔のポーズに従い分類され、固有顔の方法を用いて顔が抽出される。

【 0 0 1 7 】

【特許文献 1】米国特許出願公開第 2 0 0 4 / 0 2 6 4 7 4 4 号明細書

40

【発明の開示】

【発明が解決しようとする課題】

【 0 0 1 8 】

上記文献は顔検出の各種方法を開示しているが、改良が求められる。従って本発明の目的はデジタル画像における顔を検出する新規の方法および装置を提供することである。

【課題を解決するための手段】

【 0 0 1 9 】

よって、一態様においてデジタル画像における顔を検出する方法が提供され、前記デジタル画像の部分ウィンドウを選択するステップと、前記部分ウィンドウにおけるサンプル領域を選択するステップと、前記サンプル領域を分析し、前記部分ウィンドウが顔を表し

50

得るか判定するステップと、を含むことを特徴とする。

【0020】

一実施形態で、サンプル領域は矩形である。サンプル領域を選択する前に、部分ウィンドウはフレームに分割され、各サンプル領域は異なったフレームの1つに位置する。サンプル領域はフレームの境界からオフセットされ、パターンを形成する。

【0021】

部分ウィンドウは画像にわたりパンされ、部分ウィンドウの画像内における各位置に対し、選択および分析が再実施される。部分ウィンドウが画像にわたりパンされた後、部分ウィンドウのスケールが調節され、パンおよび再実施が繰り返される。調節は部分ウィンドウが最小サイズの閾値に達するまで行われる。

10

【0022】

分析の際、サンプル領域は部分ウィンドウに顔が存在することを検出し確認するために一連の処理段階を受ける。処理段階は少なくとも肌色分類、エッジ強度、およびアダプスト分類を含む。肌色分類は部分ウィンドウに顔が存在することを検出するために用いられる。エッジ強度およびアダプスト分類は部分ウィンドウに顔が存在することを確認するために用いられる。

【0023】

別の態様において、デジタル画像における顔を検出する装置が提供され、前記デジタル画像の部分ウィンドウを選択する部分ウィンドウ選択器と、前記部分ウィンドウ内のサンプル領域を選択するサンプル領域選択器と、前記サンプル領域と分析して前記部分ウィンドウが顔を表し得るかを判定するサンプル領域分析器と、を含むことを特徴とする。

20

【0024】

さらに別の態様において、デジタル画像における顔を検出するコンピュータ・プログラムを含むコンピュータ読み取り可能な媒体が提供され、前記コンピュータ・プログラムは、前記デジタル画像の部分ウィンドウを選択するコンピュータ・プログラム・コードと、前記部分ウィンドウにおけるサンプル領域のコンピュータ・プログラム・コードと、前記サンプル領域と分析して前記部分ウィンドウが顔を表し得るかを判定するコンピュータ・プログラム・コードと、を含むことを特徴とする。

【0025】

さらに別の態様において、デジタル画像における顔を検出する方法が提供され、前記デジタル画像の部分ウィンドウを選択するステップと、前記部分ウィンドウの区域を選択するステップと、前記部分ウィンドウの前記区域をフレームの二次元アレーに分割するステップと、前記フレームの二次元アレーを分析して前記区域各々に対し特徴を生成するステップと、前記特徴を用い、前記部分ウィンドウが顔を表し得るかを判定するステップと、を含むことを特徴とする。

30

【0026】

一実施形態で、区域は少なくとも4つのフレームに分割される。各フレームのピクセルの特性は閾値化され、各フレームのバイナリ・マップが生成され、バイナリ・マップの和に関数を実施することにより特徴が生成される。特性はピクセル輝度であることができる。あるいは特性はピクセルの色またはエッジ強度であることができる。

40

【0027】

さらに別の態様において、デジタル画像における顔を検出する装置が提供され、前記デジタル画像の部分ウィンドウの区域を選択し、前記部分ウィンドウの前記区域をフレームの二次元アレーに分割する部分ウィンドウ選択器と、前記フレームの二次元アレーを分析し、前記区域各々に対し特徴を生成し、前記特徴を用いて前記部分ウィンドウが顔を表し得るかを判定する部分ウィンドウ分析器と、を含むことを特徴とする。

【0028】

さらに別の態様において、デジタル画像における顔を検出するコンピュータ・プログラムを含むコンピュータ読み取り可能な媒体が提供され、前記コンピュータ・プログラムは、前記デジタル画像の部分ウィンドウを選択するコンピュータ・プログラム・コードと、

50

前記部分ウィンドウの区域を選択するコンピュータ・プログラム・コードと、前記部分ウィンドウの前記区域をフレームの二次元アレーに分割するコンピュータ・プログラム・コードと、前記フレームの二次元アレーを分析して前記区域各々に対し特徴を生成するコンピュータ・プログラム・コードと、前記特徴を用いて前記部分ウィンドウが顔を表し得るかを判定するコンピュータ・プログラム・コードと、を含むことを特徴とする。

【0029】

本発明の一態様である、デジタル画像における顔を検出する方法は、前記デジタル画像の部分ウィンドウを選択するステップと、前記部分ウィンドウにおいてサンプル領域を選択するステップと、前記サンプル領域を分析して前記部分ウィンドウが顔を表し得るか判定するステップとを含むことを特徴とする。

10

【0030】

また、前記サンプル領域は矩形であるものであってもよい。

【0031】

また、さらに、前記サンプル領域を選択する前に、前記部分ウィンドウをフレームに分割し、前記サンプル領域各々は前記フレームの異なったフレームに位置するステップを含むものであってもよい。

【0032】

また、前記サンプル領域は前記フレームの前記境界からオフセットされているものであってもよい。

【0033】

20

また、前記部分ウィンドウは少なくとも4つのフレームに分割されるものであってもよい。

【0034】

また、前記サンプル領域は前記部分ウィンドウの所定位置にあるものであってもよい。

【0035】

また、前記サンプル領域はパターンを形成するものであってもよい。

【0036】

また、前記パターンは格子パターンであるものであってもよい。

【0037】

また、前記分析はアダプスト方法を用いて前記サンプル領域を処理することを含むものであってもよい。

30

【0038】

また、さらに、前記画像にわたり前記部分ウィンドウをパンするステップと、前記画像内における前記部分ウィンドウの各位置に対し前記選択および分析を再実施するステップと、を含むものであってもよい。

【0039】

また、さらに、前記画像にわたり前記部分ウィンドウがパンされた後に、前記部分ウィンドウのスケールを調節し、前記パンおよび前記再実施のステップを繰り返すステップとを含むものであってもよい。

【0040】

40

また、前記調節は前記部分ウィンドウが最小閾値サイズに縮小されるまで再実施されるものであってもよい。

【0041】

また、前記分析の際、前記サンプル領域は一連の処理段階にかけられ、前記部分ウィンドウにおける顔の存在を検出し確認するものであってもよい。

【0042】

また、前記処理段階は肌色による分類、エッジ強度による分類、およびアダプストによる分類の少なくとも2つを含むものであってもよい。

【0043】

また、前記部分ウィンドウにおける顔の存在を検出するために肌色による分類が用いら

50

れるものであってもよい。

【0044】

また、前記部分ウィンドウにおける顔の存在を確認するためにエッジ強度による分類が
用いられるものであってもよい。

【0045】

また、前記部分ウィンドウにおける顔の存在を再確認するためにアダプーストによる分
類が用いられるものであってもよい。

【0046】

本発明の一態様である、デジタル画像における顔を検出する装置は、前記デジタル画像
の部分ウィンドウを選択する部分ウィンドウ選択器と、前記部分ウィンドウ内のサンプル
領域を選択するサンプル領域選択器と、前記サンプル領域を分析して前記部分ウィンドウ
が顔を表し得るか判定するサンプル領域分析器と、を含むことを特徴とする。

10

【0047】

また前記サンプル領域選択器は前記部分ウィンドウをフレームに分割し、前記サンプル
領域の各々を前記フレームの異なったフレームから選択するものであってもよい。

【0048】

また、前記サンプル領域は前記フレームの前記境界からオフセットされているものであ
ってもよい。

【0049】

また、前記サンプル領域は前記部分ウィンドウにおける所定位置にあるものであっても
よい。

20

【0050】

また、前記サンプル領域はパターンとして選択されるものであってもよい。

【0051】

また、前記パターンは格子パターンであるものであってもよい。

【0052】

また、前記サンプル領域分析器は前記サンプル領域を一連の処理段階にかけ、前記部分
ウィンドウにおける顔の存在を検出し確認するものであってもよい。

【0053】

また、前記一連の処理段階は肌色による分類、エッジ強度による分類、およびアダプ
ーストによる分類の少なくとも2つを含むものであってもよい。

30

【0054】

本発明の一態様である。デジタル画像における顔を検出する方法は、前記デジタル画像
の部分ウィンドウを選択するステップと、前記部分ウィンドウの領域を選択するステップ
と、前記部分ウィンドウの前記領域をフレームの二次元アレーに分割するステップと、フ
レームの前記二次元アレーを分析し前記領域の各々に対し特徴を生成するステップと、前
記特徴を用い、前記部分ウィンドウは顔を表し得るか判定するステップと、を含むこと
を特徴とする。

【0055】

また、前記領域は少なくとも4つのフレームに分割されるものであってもよい。

40

【0056】

また、前記フレームの各々のピクセル特性を閾値化して各フレームに対しバイナリ・マ
ップを生成するステップと、各フレームの前記バイナリ・マップの総和に対し関数を実施
することにより前記特徴を生成するステップと、を含むものであってもよい。

【0057】

また、前記特性はピクセル輝度であるものであってもよい。

【0058】

また、前記特性は前記ピクセルの色値であるものであってもよい。

【0059】

また、前記特性は前記ピクセルのエッジ強度であるものであってもよい。

50

【 0 0 6 0 】

方法および装置はデジタル画像における顔検出に対する高速な方法を提供する。デジタル画像の部分ウィンドウの区域を表すサンプル領域を分析することにより、精度を著しく低減させることなく演算コストを削減することができる。さらに、フレームの二次元アレーを分析して特徴を生成することにより、画像の部分ウィンドウを顔または非顔に分類する場合より少ない特徴を利用することができ、それにより処理条件および時間を削減できる。

【 発明を実施するための最良の形態 】

【 0 0 6 1 】

以下に添付図面を参照して実施形態をさらに詳しく説明する。以下の説明において、デジタル画像における顔を検出する方法、装置、およびコンピュータ・プログラムを具現するコンピュータ読み取り可能な媒体が提供される。方法の際、異なったサイズおよび位置のいくつかの部分ウィンドウが分析される。場合により、部分ウィンドウ内のサンプル領域のセットのみ分析され、それにより演算コスト（すなわち処理能力および時間）が削減される。各部分ウィンドウに対し、次の分類器 - 肌色に基づく分類器、エッジ強度に基づく分類器、およびジェントル・アダプストに基づく分類器 - がカスケードの段階のセットに決定され、部分ウィンドウが顔を含むか検出する。第 1 段階は演算上速く、または「安く」、続くテストの各段階で必要な処理能力が増える。いずれかの段階で、部分ウィンドウがおそらく顔を表していないと判定された場合、部分ウィンドウの分析は終了し、次の部分ウィンドウの分析を開始することができる。部分ウィンドウは 3 つの分類器各々に合格して初めて顔を含むと判定される。

【 0 0 6 2 】

図 1 を見ると、デジタル画像における顔を検出する装置が示され、全般的に装置 2 0 で識別される。本実施形態で、装置はパソコン等々で、処理装置 2 4、ランダムアクセス・メモリ（「RAM」）2 8、非揮発性メモリ 3 2、通信インタフェース 3 6、入力インタフェース 4 0、および出力インタフェース 4 4 を含み、すべてローカル・バス 4 8 を通じて通信している。処理装置 2 4 は非揮発性メモリ 3 2 に格納される顔検出アプリケーションを実行する。装置 2 0 は通信インタフェース 3 6 経由でネットワークまたは画像および顔検出結果を格納するサーバに接続されることができる。入力インタフェース 4 0 はキーパッド、マウス、および / あるいは他のユーザ入力装置を含み、ユーザが顔検出アプリケーションとやりとりできるようにする。入力インタフェースはさらに顔検出のために分析される画像を取り込むスキャナを含むことができる。出力インタフェース 4 4 はディスプレイを含み、望まれる場合は顔検出の結果を視覚的に表示することができ、調節用に顔検出アプリケーションの設定を表示することができる。

【 0 0 6 3 】

次に図 2 乃至図 4 を見ると、顔検出アプリケーションの実行の際、装置 2 0 が画像を処理してその中の顔を検出するために行う一般工程が示される。最初に特定スケールの部分ウィンドウが選択される（ステップ S 1 1 0）。部分ウィンドウは $x(m, n, s)$ の形を取り、 m および n は各々画像の左上の隅から水平および垂直のオフセットをピクセルで表し、 s は部分ウィンドウのスケールを表す。本実施形態で、部分ウィンドウは正方形で、正方形の高さおよび幅は両方 s ピクセルに等しい。図 3 は画像 2 0 0 に対する部分ウィンドウ 2 0 4 の例を示す。選択される最初の部分ウィンドウのスケール、 s 、は画像 2 0 0 にはまる最大サイズの正方形領域である。すなわち、 s は画像の高さおよび幅のより小さい方に設定され、 m および n は最初ゼロ（0）に設定される。部分ウィンドウ 2 0 4 は次にいくつかの等しいフレームに分割され、本例では図 4 に示すように 4 つのフレーム A ~ D に分割される（ステップ S 1 1 5）。部分ウィンドウ 2 0 4 は次に画像 2 0 0 の左上の隅に適用され、部分ウィンドウ 2 0 4 内のフレームは肌色に基づく分類器を用いて分析され、部分ウィンドウ内に顔が存在する判定される（ステップ S 1 2 0）。部分ウィンドウ 2 0 4 に顔が見つかった場合、部分ウィンドウ内のフレーム A ~ D が次にエッジ強度に基づく分類器を用いて分析され部分ウィンドウにおける顔の存在を確認する（ステップ S

130)。顔の存在が確認された場合、部分ウィンドウ204内のフレームA～Dはジェントル・アダプストに基づく分類器を用いて再度分析され、部分ウィンドウにおける顔の存在が確認される(ステップS140)。顔の存在が確認された場合、部分ウィンドウ204は顔を包含するものとして登録される(ステップS150)。

【0064】

次に部分ウィンドウ204が画像200全体にわたりパンされたか判定するために調査が行われる(ステップS160)。されていない場合、部分ウィンドウの位置は増分され(ステップS170)、ステップS120、S130、S140、S150、S160が新しい部分ウィンドウ位置に対し繰り返される。ステップS120、S130、S140、S150、S160は画像200に対しあり得る部分ウィンドウ位置各々に対し繰り返されることが理解されよう。

【0065】

ステップS120で顔が存在しないとみなされるか、または部分ウィンドウ204における顔の存在がステップS130またはステップS140で確認されない場合、プロセスはステップS160に進む。

【0066】

一旦、部分ウィンドウ204が画像全体にわたりパンされると、他のスケールの部分ウィンドウが選択できるか調査が行われる(すなわち部分ウィンドウが最小サイズになっているか)(ステップS180)。本実施形態で、部分ウィンドウの最小サイズは17×17ピクセルに等しい。部分ウィンドウが最小サイズになっていない場合、部分ウィンドウは14%縮小の上、最も近い整数に四捨五入され(ステップS190)、ステップS120、S130、S140、S150、S160がその部分ウィンドウに対し繰り返される。上記プロセスはより小さいスケールの追加部分ウィンドウが選択できなくなるまで続けられる。

【0067】

ステップS120において、部分ウィンドウ204内のフレームを分析する際、部分ウィンドウ204のフレーム内の各ピクセルの赤、緑、および青(「RGB」)値がバイナリ・ベイズ分類器に供給される。バイナリ・ベイズ分類器はピクセルのRGB色値に基づき各ピクセルが肌または非肌を表す可能性を判定する。

バイナリ・ベイズ分類器は訓練画像から取られた部分ウィンドウのサンプル・セットを用いて訓練される。訓練の際、各訓練画像の各部分ウィンドウは顔または非顔を表すと手動で分類され、部分ウィンドウのピクセルは各々肌および非肌のヒストグラムを生成するのに用いられる。ヒストグラムは三次元アレーで、各次元はRGB色空間の赤Rピクセル値、緑Gピクセル値、および青Bピクセル値に対応する。具体的に、ヒストグラムは32×32×32の寸法である。訓練画像は肌を表すか表さないか手動で特定されるので、しかるべき肌または非肌のヒストグラムに訓練画像からのピクセル値が入れられる。これらのヒストグラムは次に肌および非肌の対象からもたらされるピクセル色値のベイズの確率を計算するために用いられる。具体的に、特定のピクセル色値、 z 、が肌からもたらされる確率 $P(z|skin)$ は以下による。

【0068】

【数1】

$$P(z|skin) = \frac{H_s(z)}{N_s}$$

【0069】

ここで、

$H_s(z)$ は、分析されるピクセルと同じRGBビンにある訓練画像のセットから構築された肌ヒストグラムにおけるピクセル数、 N_s は、肌ヒストグラムに含まれるピクセルの合計数である。

10

20

30

40

50

同様に、特定のピクセル色値、 z 、が非肌からもたらされる確率 $P(z | non - skin)$ は以下による。

【 0 0 7 0 】

【 数 2 】

$$P(z | non - skin) = \frac{H_n(z)}{N_n}$$

【 0 0 7 1 】

ここで、

10

$H_n(z)$ は分析されるピクセルと同じ RGB ピンにある訓練画像のセットから構築された非肌ヒストグラムにおけるピクセル数、 N_n は非肌ヒストグラムに含まれるピクセルの合計数である。

【 0 0 7 2 】

上記の 2 つの確率を用い、各ピクセルに対するベイズ分類器は以下のとおりである。

【 0 0 7 3 】

【 数 3 】

$$g(z) = \begin{cases} 1, & \text{if } \frac{p(z | skin)}{p(z | non - skin)} > \theta_g \\ 0, & \text{otherwise} \end{cases}$$

20

【 0 0 7 4 】

ここで、

θ_g は正しい肌検出と偽陽性との間のトレードオフを調節するのに用いることのできる閾値である。

このように部分ウィンドウ 204 に対しバイナリの肌 / 非肌色マップが生成され、肌を表すとみなされるピクセルには 1 の値が割り当てられ、非肌を表すとみなされるピクセルにはゼロ (0) の値が割り当てられる。図 5 は部分ウィンドウ 204 がデジタル画像の肌色マップに適用されることを示す。

30

【 0 0 7 5 】

次に部分ウィンドウ 204 に対し以下のように 7 つの和に基づくスカラー特徴 f_1 から f_7 が計算される。

【 0 0 7 6 】

【 数 4 】

$$f_1 = \text{sum}(A) + \text{sum}(B) + \text{sum}(C) + \text{sum}(D) \text{ ---- (方程式1)}$$

$$f_2 = \text{sum}(A) + \text{sum}(C) \text{ ----- (方程式2)}$$

40

$$f_3 = \text{sum}(B) + \text{sum}(D) \text{ ----- (方程式3)}$$

$$f_4 = |\text{sum}(A) + \text{sum}(C) - \text{sum}(B) - \text{sum}(D)| \text{ ---- (方程式4)}$$

$$f_5 = \text{sum}(A) + \text{sum}(B) \text{ ----- (方程式5)}$$

$$f_6 = \text{sum}(C) + \text{sum}(D) \text{ ----- (方程式6)}$$

$$f_7 = |\text{sum}(A) + \text{sum}(B) - \text{sum}(C) - \text{sum}(D)| \text{ ---- (方程式7)}$$

【 0 0 7 7 】

50

ここで、

$sum(Z)$ はフレーム Z に対応する肌色マップのピクセル値 $g(z)$ の和を意味する。スカラー特徴 $f_1 \sim f_7$ は肌色マップの範囲総和表または積分画像を用いて効率的に計算される。スカラー特徴 $f_1 \sim f_7$ は各々に対応閾値と比較することにより部分ウィンドウ 204 を分類するために用いられる。すなわち以下のとおりである。

【0078】

【数5】

$$f_i(x) \begin{cases} \geq \theta_i \Rightarrow \text{facepattern} \\ < \theta_i \Rightarrow \text{non-facepattern} \end{cases}$$

10

【0079】

ここで、

θ_i は訓練データを評価することにより決定される閾値である。

【0080】

比較中スカラー特徴 $f_1 \sim f_7$ は、スカラー特徴 $f_1 \sim f_7$ の順序でカスケードされた形で各々対応する閾値と比較される。部分ウィンドウ 204 に対するいずれかのスカラー特徴がその対応する閾値を満足しない場合、部分ウィンドウ 204 は非顔を表すとみなされ、残りのスカラー特徴および他の分類器は打ち切られる。

【0081】

20

ステップ $S120$ において肌色に基づく分類器を用いて部分ウィンドウ 204 が顔を表すとみなされる場合、前述のとおりステップ $S130$ においてエッジ強度に基づく分類器を用いて顔の存在を確認する。ステップ $S130$ における顔の確認の際、各ピクセルのエッジ強度を用いて入力画像のエッジ強度マップが生成される。エッジ強度は次の一次導関数を用いて判定される。

【0082】

【数6】

$$s(I) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2}$$

30

【0083】

本実施形態で、ソーベル・エッジ検出手法が用いられる。ソーベル・エッジ検出手法は 3×3 ピクセル・カーネルを用い、部分ウィンドウ 204 の各ピクセルについて8つの隣接ピクセルの輝度値に対するそのピクセルの輝度値に基づきエッジ強度を判定する。結果はデジタル画像における各ピクセルのエッジ強度値を含むエッジ強度マップが得られる。

【0084】

図6は図5のデジタル画像のバイナリ・エッジ強度マップに部分ウィンドウ 204 を適用したところを示す。バイナリ・エッジ強度マップは各ピクセルについて次を判定することにより得られる。

40

【0085】

【数7】

$$e(I) = \begin{cases} 1, & \text{if } s(I) \geq \theta_e \\ 0, & \text{otherwise} \end{cases}$$

【0086】

ここで、

50

$s(I)$ はピクセルのエッジ強度、 θ は調節可能な閾値である。

【0087】

部分ウィンドウ 204 のエッジ強度マップが $e(x)$ として、スカラー特徴 $f_1 \sim f_7$ はエッジ強度マップのピクセル値を用いて再度直列的に計算される。各スカラー特徴が計算されるにつれ、対応する閾値と比較される。部分ウィンドウ 204 に対するいずれかの計算されたスカラー特徴が各々の閾値を満足しない場合、部分ウィンドウは非顔を表すとみなされ、残りのスカラー特徴の判定および分類器は打ち切られる。

【0088】

上述のとおり、ステップ S130 が顔を表すとみなされる部分ウィンドウ 204 をもたらした場合、ステップ S140 でジェントル・アダプーストに基づく分類器を用いて結果が確認される。訓練画像のセットが与えられた場合、ジェントル・アダプーストの発想は訓練段階中に訓練画像セットに対し弱いが多様な分類器を特定し、次に弱分類器を直線的に組み合わせて 1 つ以上の強分類器を形成することである。訓練エラーを直接最小限にするために撤回機構が導入される。これは非効率的な弱分類器を取り除き、強分類器を構築するために組み合わされる弱分類器の数を削減するのに役立つ。各弱分類器は訓練画像の部分ウィンドウにおける 1 つのスカラー特徴と結び付けられる。スカラー特徴はテンプレートに従い部分ウィンドウの特定区域におけるピクセル輝度の荷重和を計算することにより判定される。弱分類器の組み合わせに基づく部分ウィンドウの分類は実質的に人が顔を認識するために行うパターン認識のファジー論理に類似している。

【0089】

離散的ジェントル・アダプーストの基本形は 2 階級問題を扱う。N ラベルの訓練画像のセットが $(x_1, y_1), \dots, (x_N, y_N)$ として与えられ、 $y_i \in \{+1, -1\}$ は訓練画像セット $x_i \in \mathbb{R}^n$ に対する階級ラベルである。

【0090】

【数 8】

$w_j^{(m)}$; 訓練画像セットの分布-----式(A)

【0091】

ジェントル・アダプーストは訓練画像セットの分布（上記式（A）による）に関し訓練画像セットからの弱分類器 $h_m(x)$ ($m = 1, 2, \dots, M$) の学習シーケンスに対し手順があると仮定する。

【0092】

弱分類器はテンプレートを用いて生成された 1 つのスカラー特徴に対応する。テンプレートは 20×20 ピクセル・アレーを用いて訓練中に選択される。従って、部分ウィンドウ 204 に対するスカラー特徴を判定するために、部分ウィンドウは 20×20 ピクセルに再度縮小され、テンプレートが適用される。

【0093】

図 7 は部分ウィンドウ 204 ならびにスカラー特徴に結び付いた 3 つのテンプレート 212a, 212b, 212c を示す。各テンプレート 212a, 212b, 212c は 9 つの 3×3 のフレーム 216 に分割され、各フレーム内に矩形のサンプル領域 220 が位置される。このように矩形のサンプル領域 220 はテンプレート 212a, 212b, 212c において格子パターンを形成する。テンプレート 212b は $X_0 \sim X_3$ の水平範囲と $Y_0 \sim Y_3$ の垂直範囲にかかるよう示される。

【0094】

9 つの矩形のサンプル領域 220 は、下記の式（B）で示される。

【0095】

10

20

30

40

【数 9】

 $\{W_{ij}\}_{i,j=1}^3$; 矩形のサンプル領域-----式(B)

【0 0 9 6】

これらの9つの矩形のサンプル領域 2 2 0 (上記式 (B) による) の位置および寸法は以下の変数のセットに支配される。

【0 0 9 7】

【数 1 0】

 $\{X_k\}_{k=0}^3$ -----; フレーム216間の各分割線のx座標

10

 $\{Y_k\}_{k=0}^3$ -----; フレーム216間の各分割線のy座標

 $\{du_{ij}\}_{i,j=1}^3$ -----; サンプル領域220の各フレーム上端からの垂直オフセット

20

 $\{dv_{ij}\}_{i,j=1}^3$ -----; サンプル領域220の各フレーム上端からの水平オフセット

 $\{w_{ij}\}_{i,j=1}^3$ -----; 各サンプル領域の幅

 $\{h_{ij}\}_{i,j=1}^3$ -----; 各サンプル領域の高さ

30

【0 0 9 8】

例示の場合、各サンプル領域 2 2 0 は関連しているフレーム 2 1 6 の境界からオフセットされている。

【0 0 9 9】

スカラー加重係数、 B_{ij} R、が各サンプル領域 (上記式 (B) による) と結び付いている。部分ウィンドウ 2 0 4 のスカラー特徴はテンプレートで規定される9つのサンプル領域 2 2 0 内におけるピクセルの輝度値総和の線形加重組み合わせにより計算される、すなわち、

40

【0 1 0 0】

【数 1 1】

$$fn(\delta) = \sum_{i=1}^3 \sum_{j=1}^3 B_{ij} \cdot sum(W_{ij}) \text{ ----(方程式8)}$$

【0 1 0 1】

ここで、

【0 1 0 2】

【数 1 2】

$$\delta = \left\{ \{X_k\}_{k=0}^3, \{Y_k\}_{k=0}^3, \{du_{ij}\}_{i,j=1}^3, \{dv_{ij}\}_{i,j=1}^3, \{w_{ij}\}_{i,j=1}^3, \{h_{ij}\}_{i,j=1}^3, \{B_{ij}\}_{i,j=1}^3 \right\}$$

【0 1 0 3】

はすべての変数のセット、 $\text{sum}(W_{ij})$ はサンプル領域 W_{ij} のすべてのピクセルにわたる総和を意味する。

関数 $\text{sum}(W_{ij})$ はSIGGRAPH、1984年、18(3) 巻、207-212頁記載のF.Crow著「テクスチャ・マッピング用の範囲総和表」(「Summed-area tables for texture mapping」)と題する資料で説明されるような範囲総和表から効率的に演算することができる。あるいは関数 $\text{sum}(W_{ij})$ はInternational Journal of Computer Vision、57巻、2004年5月、137~154頁記載のPaul Viola およびMichael J. Jones共著「頑強なりアルタイム顔検出」(「Robust real-time face detection」)と題する資料で説明されるような積分画像から効率的に演算される。可変セット の値を変化させることにより各種スカラー特徴が一般化される。

図 8 は選択されたスカラー特徴を判定するのに使用される加重係数 B_{ij} の各種スキームを用いるテンプレートを示す。装置 2 0 で用いられる 1 6 の加重スキームの内 6 つが示される。本実施形態で、- 1、0、および 1 といった 3 つの別個の加重係数が用いられる。テンプレートにおけるサンプル領域 W_{ij} に対する加重係数 B_{ij} は次の方程式を満足することが特記される。

【0 1 0 4】

【数 1 3】

$$\sum_{i=1}^3 \sum_{j=1}^3 B_{ij} \cdot \text{Area}(W_{ij}) = 0$$

【0 1 0 5】

1 つのスカラー特徴を演算するためにはサンプル領域 2 2 0 に対しピクセルの 9 つの総和が必要であることが方程式 8 から分かる。このように演算されるスカラー特徴の演算上の複雑度は単純なハールの特徴より 4 . 5 倍高い。その結果、演算されたスカラー特徴のセットは顔検出の目的に対しより多くの情報を提供する。

【0 1 0 6】

テンプレートの複雑さのため、エッジより複雑な画像特徴を検出することができる。単一のハール特徴は一次元的なエッジの検出しかできない。その結果、本件の顔検出方法はハール方法と同じ精度をかなり減らされたスカラー特徴を用いて達成することができる。

【0 1 0 7】

強分類器は図 9 に示すようにカスケードされた形で組織される。推進強分類器は顔を表す部分ウィンドウに対し高い検出率を維持しながら非顔の部分ウィンドウ 2 0 4 の大部分を有効に排除する。このカスケード構成において、初期の強分類器の方が弱分類器数が少なく偽陽性検出率が高い。部分ウィンドウ処理の際、強分類器に合格しない部分ウィンドウはそれ以上後続の強分類器により処理されない。部分ウィンドウ 2 0 4 は n 個の強分類器すべてに合格すると顔を表すとみなされる。段階毎に処理コストおよび識別レベルが高くなる。従って、部分ウィンドウ 2 0 4 がいずれかの段階でテストに不合格となると、部分ウィンドウをさらに分析してさらに処理能力および時間を無駄にしないということがない。このカスケード式強分類器の方法は検出プロセスの速度を相当高め、偽陽性を減らすことができる。

【0 1 0 8】

上述のように、スカラー特徴を判定するために用いられるテンプレートは訓練プロセス中に選択される。縮小し直した、サイズが 20×20 ピクセル部分ウィンドウの場合スカ

ラー特徴となり得るものは何万とある。これらのスカラー特徴は部分ウィンドウに対し過剰に揃ったスカラー特徴のセットを形成する。訓練段階の際、各種サイズ、加重方式および位置の二次元テンプレートを用いてスカラー特徴が生成され、人間のオペレータにより分類される各部分ウィンドウに適用される。

【 0 1 0 9 】

各弱分類器 $h_m(x)$ は 1 つのスカラー特徴 f_i と結び付いている。従って最良の新規弱分類器をみつけるという課題は最良の対応スカラー特徴を選択することに等しい。弱分類器は、訓練中に顔および非顔を表すと特定された部分ウィンドウに対し生成されたヒストグラムに基づき、肌色およびエッジ強度分類器に対しベイズ確率が判定されるのと同じように、部分ウィンドウ 204 が顔を表すというベイズの確率を判定することにより構築される。ベイズ確率は次に訓練セットを用いてその特定のスカラー特徴の結果を評価することにより判定される閾値と比較される。

10

【 0 1 1 0 】

【 数 1 4 】

$$h_m(f_i, x_j) = \begin{cases} v_i^{(m)}, & \text{if } f_i \geq \theta_i \\ \mu_i^{(m)}, & \text{otherwise} \end{cases}, \text{ with } v_i, \mu_i \in [-1, 1]$$

20

$$v_i^{(m)} = \frac{W_+^r - W_-^r}{W_+^r + W_-^r}$$

$$\mu_i^{(m)} = \frac{W_+^l - W_-^l}{W_+^l + W_-^l}$$

30

$$W_+^r = \sum_{j: (y_j=1) \& (f_i(x_j) \geq \theta_i)} w_j^{(m)}$$

$$W_-^r = \sum_{j: (y_j=-1) \& (f_i(x_j) \geq \theta_i)} w_j^{(m)}$$

40

$$W_+^l = \sum_{j: (y_j=1) \& (f_i(x_j) < \theta_i)} w_j^{(m)}$$

$$W_-^l = \sum_{j: (y_j=-1) \& (f_i(x_j) < \theta_i)} w_j^{(m)}$$

50

【 0 1 1 1 】

強分類器は次に従ってM個の弱分類器を組み合わせることにより構築される。

【 0 1 1 2 】

【 数 1 5 】

$$H_M(x) = \sum_{m=1}^M h_m(x) - \theta_b$$

【 0 1 1 3 】

ここで、

θ_b は検出率と偽陽性との間のトレードオフを支配する閾値である。

【 0 1 1 4 】

x の分類は $y(x) = \text{sign}[H_M(x)]$ として求められ、正規化された信頼得点は $|H_M(x)|$ である。 $h_m(x)$ の元の形は離散関数である。ジェントル・アダプーストは次の加重最小二乗エラーを最小限にすることを目標としている。

【 0 1 1 5 】

【 数 1 6 】

$$E(f_i) = \sum_{j=1}^N w_j^{(m)} (h_m(f_i, x_j) - y_j)^2$$

【 0 1 1 6 】

h_m のパラメータは最良のスカラー特徴 f_i と併せ、エラーを最小限にすることにより判定できる。

【 0 1 1 7 】

【 数 1 7 】

$$f_* = \arg \min_{f_i} E(f_i)$$

【 0 1 1 8 】

訓練段階の際、テンプレートのパラメータおよび位置はさまざまで、テンプレートの各サンプル領域に結び付いた加重係数も同様である。これは人間のオペレータにより分類された各部分ウィンドウに対し行われる。このような差異は多数あるため過剰に揃ったスカラー特徴のセットが提供されることが理解されよう。差異の数のため訓練は比較的長時間かかる。

【 0 1 1 9 】

スカラー特徴はすべて顔を表すと特定された部分ウィンドウとの結び付きに基づき優先順位が付けられる。強分類器を生成するために、弱分類器は顔識別の望ましいレベルが達成されるまで組み合わせられる。最初の強分類器の場合、弱分類器は加重された形で少なくとも98%の真陽性、最大50%の偽陽性が達成されるまで組み合わせられる。

【 0 1 2 0 】

後続の強分類器を判定するために、各々前の強分類器に合格した20,000個の顔および20,000個の非顔を有してなる訓練セットを用いて弱分類器のセットを判定し、これらは次に同様に優先順位が付けられる。すなわち、後続の強分類器を判定するために、前の強分類器を用いて顔と分類された部分ウィンドウに基づき新規訓練セットが選択される。次に望ましい合格率を有する強分類器が判定されるまで弱分類器が加重された形で組み合わせられる。このように、提案された過剰に揃ったスカラー特徴のセットが最も重要なスカラー特徴を選択するためにジェントル・アダプースト学習が用いられる。

【 0 1 2 1 】

10

20

30

40

50

後続の強分類器生成のために訓練画像のセットを生成することおよび可能なスカラー特徴すべての分析は非常に時間がかかり得る。しかし一旦、強分類器を生成するのに用いられる弱分類器が訓練段階で選択されると、その後の部分画像の分析は効率的かつ望ましい正確さのレベルで行われる。部分ウィンドウの各フレームのフレーム全体でなく、サンプル領域を分析することは必要な計算を減らし、従って結果の精度を著しく減少させることなく処理時間を減らす。

【0122】

本明細書で提案される顔検出装置および方法は色分析、演算し易いスカラー特徴およびジェントル・アダプスト学習など、画像処理、パターン認識、および機械学習におけるいくつかの強力な手法を組み合わせている。これらの手法を用い3つの個別的顔分類器が作り出され、これらはカスケードされた形で組織される。これにより装置は肌色、エッジ、およびグレースケールに基づくスカラー特徴など利用可能な各種の情報をを用い、デジタル画像における顔を望ましいレベルの精度を以って迅速に識別できる。

10

【0123】

図面を参考に一実施形態が説明されたが、当業者であれば各種の変形が可能であることを理解しよう。例えば、上述の総和に基づくスカラー特徴の代わりに、またはこれを補強するために偏差に基づくスカラー特徴を用いることができる。このように、顔検出に用いるスカラー特徴の数をかなり増やすことができる。また、部分ウィンドウの肌色およびエッジ分析の際、部分ウィンドウ全体が分析されるように説明されたが、当業者であれば部分ウィンドウの分類の精度を著しく低減させることなく部分ウィンドウのサンプル領域のみ分析することも可能であることを理解しよう。

20

【0124】

ジェントル・アダプストに基づく分析は色、エッジ方向、エッジ強度等々、部分ウィンドウの他の特性を用いて行うことができる。またアダプスト方法の他の形態をジェントル・アダプストの代わりに用いることができる。

【0125】

使用テンプレートはいくらでも異なった加重係数を含むことができる。テンプレートのフレームおよびサンプル領域の特性も変えることができる。

【0126】

提案された総和に基づくスカラー特徴は色画像、グレースケール画像、色画像の肌色マップ、および色またはグレースケール画像のエッジ・マップなど任意の二次元信号から抽出することができる。提案される顔検出システムにこれらのスカラー特徴は3つの要素分類器 - 肌色に基づく分類器、エッジ強度に基づく分類器、およびジェントル・アダプストに基づく分類器 - において用いられる。

30

【0127】

顔検出アプリケーションは独立したデジタル画像ツールとして実行でき、または他の利用可能なデジタル画像処理アプリケーションに組み入れ、これらのデジタル画像処理アプリケーションに拡大機能を提供できる。ソフトウェア・アプリケーションはルーチン、プログラム、オブジェクト・コンポーネント、データ構造等々を含むプログラム・モジュールを含むことができ、コンピュータ読み取り可能な媒体に格納されるコンピュータ読み取り可能なプログラム・コードとして具現されることができる。コンピュータ読み取り可能な媒体は後にコンピュータ・システムにより読み取られるデータを格納する任意のデータ記憶装置である。コンピュータ読み取り可能な媒体の例としては例えば読み取り専用メモリ、ランダムアクセス・メモリ、ハードディスク・ドライブ、磁気テープ、ならびにCD-ROMおよび他の光学データ記憶装置が含まれる。コンピュータ読み取り可能なプログラム・コードは、コンピュータ読み取り可能なコードが分散した形で格納され実行されるよう、連結コンピュータ・システムなどを含むネットワーク上で分散されることもできる。

40

【0128】

特定の実施形態が説明されたが、当業者であれば添付クレームに定義される精神および

50

範囲を逸脱することなく変形および修正をなすことが可能であることを理解しよう。

【図面の簡単な説明】

【0129】

【図1】 デジタル画像における顔を検出する装置の概略図である。

【図2】 図1の装置により用いられる顔検出方法のフローチャートである。

【図3】 顔検出の際、分析されるデジタル画像における部分ウィンドウのパラメータを示す。

【図4】 図3に示す部分ウィンドウのフレームを示す。

【図5】 デジタル画像の肌色マップに適用された図4の部分ウィンドウを示す。

【図6】 図5のデジタル画像におけるエッジ・マップに適用された図4の部分ウィンドウを示す。

【図7】 部分ウィンドウのフレームに適用されるテンプレートを示す。

【図8】 6つのテンプレートを示す。

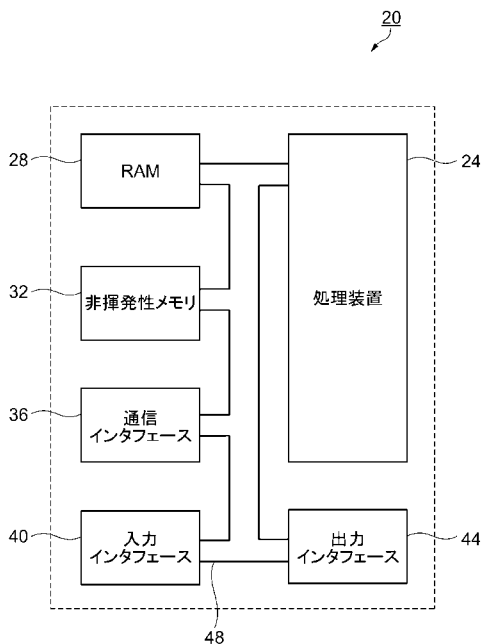
【図9】 図1の装置により用いられるジェントル・アダブースト (Gentle AdaBoost) に基づく分類方法のフローチャートである。

【符号の説明】

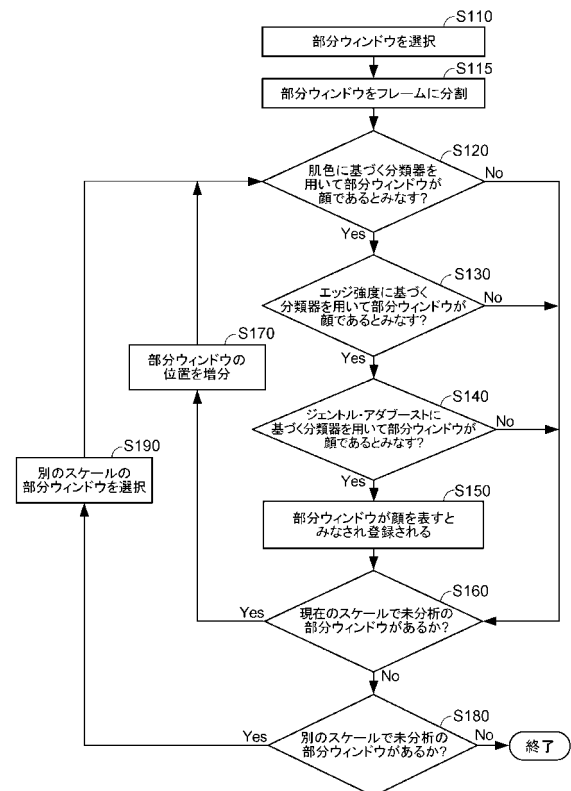
【0130】

20 ... 装置、24 ... 処理装置、32 ... 非揮発性メモリ、36 ... 通信インタフェース、40 ... 入力インタフェース、44 ... 出力インタフェース。

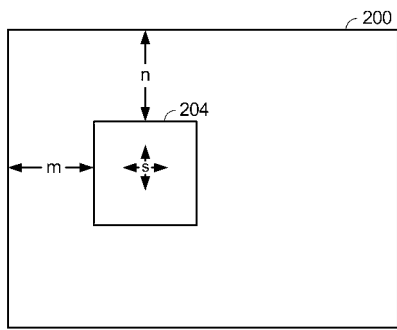
【図1】



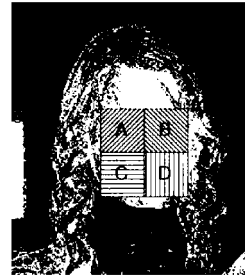
【図2】



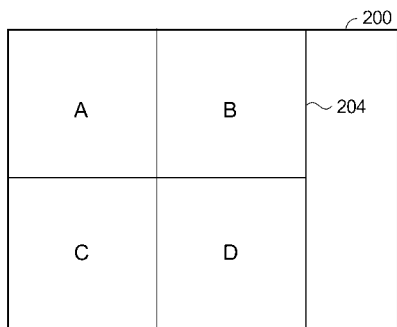
【図 3】



【図 5】



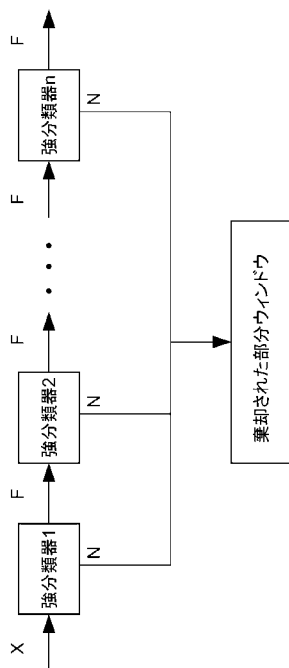
【図 4】



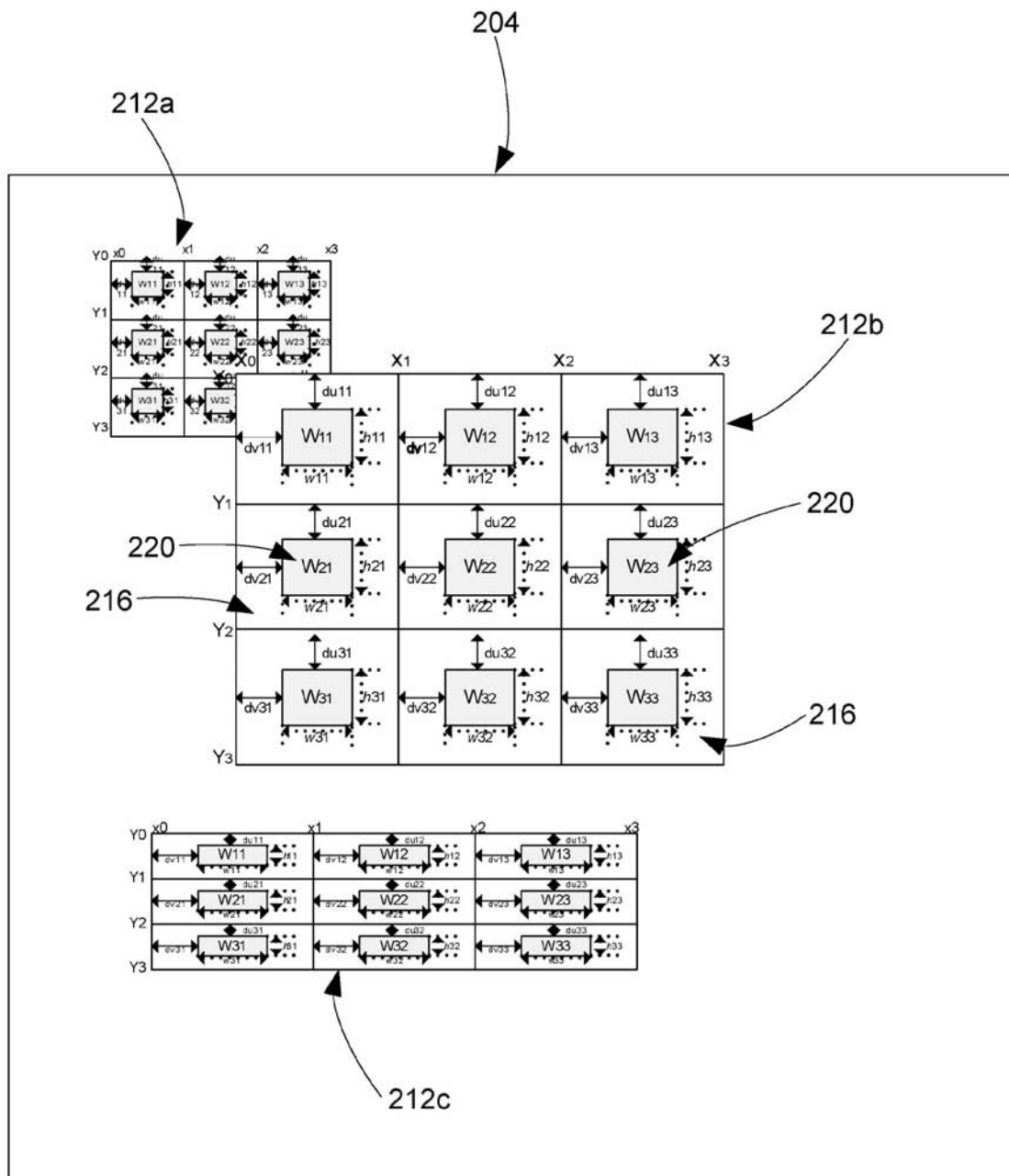
【図 6】



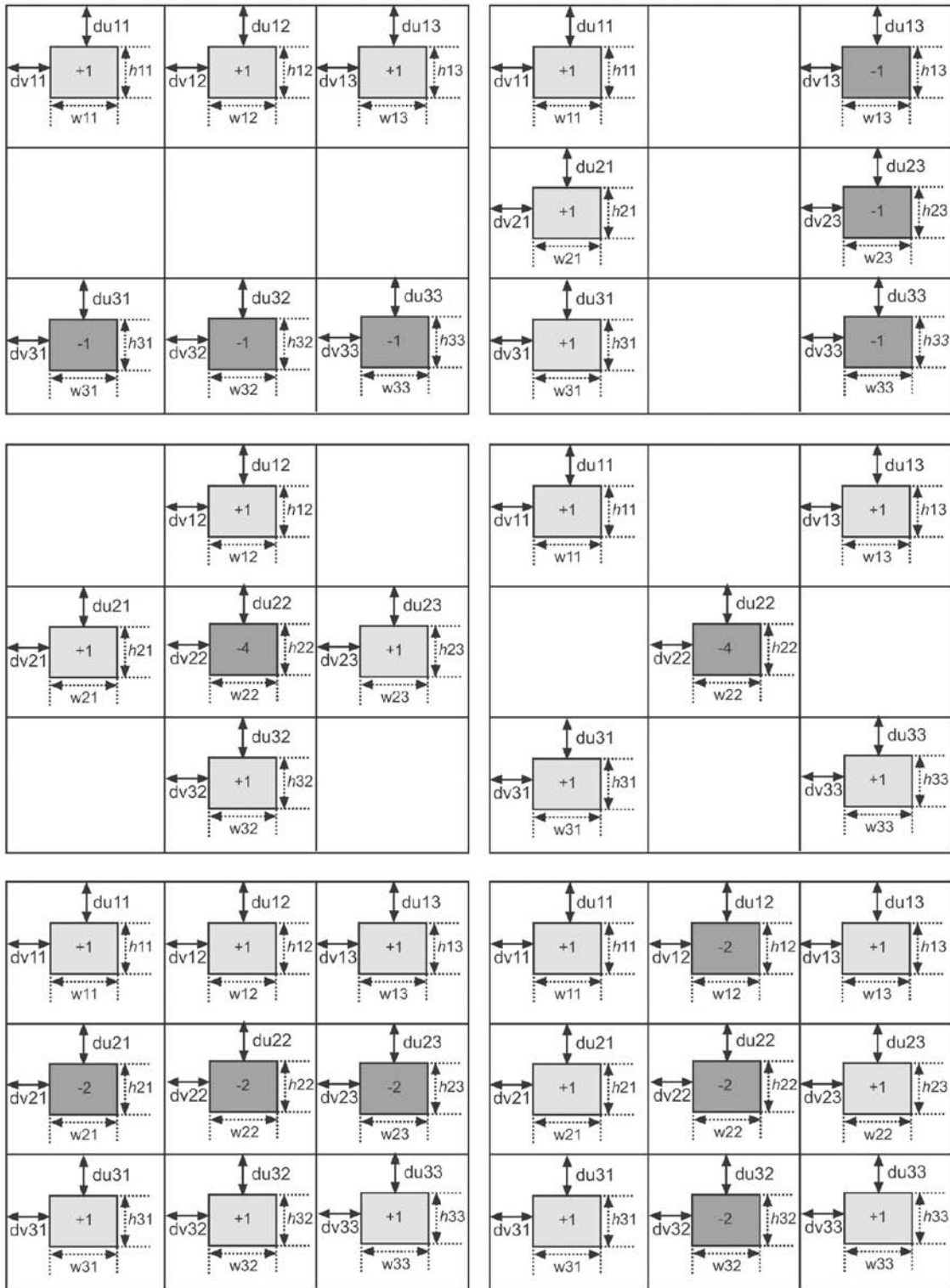
【図 9】



【 図 7 】



【 図 8 】



フロントページの続き

(72)発明者 ファイ ゾウ

 カナダ オンタリオ トロント スプリングデール ブルーバード 1 3 6

Fターム(参考) 5B057 CA08 CA12 CA16 DA08 DA12 DB02 DB06 DB09 DC17 DC25

 DC36 DC40

 5L096 AA02 AA06 CA02 FA06 GA17 GA19 GA38