

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2008年2月14日 (14.02.2008)

PCT

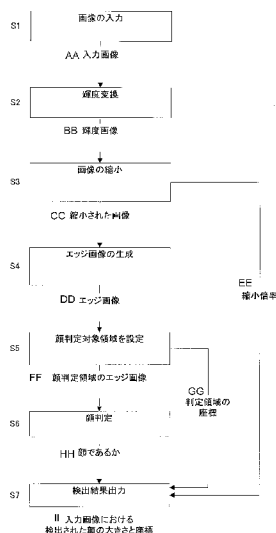
(10) 国際公開番号
WO 2008/018459 A1

- (51) 国際特許分類:
G06T 7/60 (2006.01) H04N 5/232 (2006.01)
- (21) 国際出願番号: PCT/JP2007/065446
- (22) 国際出願日: 2007年8月7日 (07.08.2007)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2006-215943 2006年8月8日 (08.08.2006) JP
- (71) 出願人 (米国を除く全ての指定国について): 株式会社ニコン (NIKON CORPORATION) [JP/JP]; 〒1008331 東京都千代田区丸の内三丁目2番3号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 宇津木 暁彦 (UTSUGI, Akihiko) [JP/JP]; 〒1008331 東京都千代田区丸の内三丁目2番3号 株式会社ニコン知的財産部内 Tokyo (JP).
- (74) 代理人: 永井 冬紀 (NAGAI, Fuyuki); 〒1000011 東京都千代田区幸町二丁目1番1号 飯野ビル Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE,

[続葉有]

(54) Title: IMAGE PROCESSING METHOD, IMAGE PROCESSING APPARATUS, IMAGE PROCESSING PROGRAM, AND IMAGE PICKUP APPARATUS

(54) 発明の名称: 画像処理方法、画像処理装置、画像処理プログラム、撮像装置



S1 ENTER IMAGE
AA INPUT IMAGE
S2 CONVERT BRIGHTNESS
BB BRIGHTNESS IMAGE
S3 REDUCE SCALE OF IMAGE
CC SCALE-REDUCED IMAGE
S4 GENERATE EDGE IMAGE
DD EDGE IMAGE
EE SCALE-DOWN RATE
S5 ESTABLISH AREA TO BE USED FOR FACE DETERMINATION
FF EDGE IMAGE OF FACE DETERMINATION AREA
GG COORDINATES OF DETERMINATION AREA
S6 DETERMINE FACE
HH FACE?
S7 OUTPUT DETERMINATION RESULT
II SIZE AND COORDINATES OF FACE DETERMINED IN INPUT IMAGE

(57) Abstract: An image processing method for determining whether an image is a particular type of image. The image processing method comprises acquiring an image consisting of a plurality of pixels; storing a look-up table indicative of a likelihood of a particular type of image for each pixel value and for each pixel position; generating, based on the acquired image, an image to be determined; obtaining, based on the pixel values and pixel positions of the pixels of the image to be determined, likelihoods of the particular type of image on those pixels by use of the look-up table; totalizing the obtained likelihoods of the particular type of image as to the pixels of the image to be determined; and determining, based on the totalization result, whether the input image is the particular type of image.

(57) 要約: 特定種類の画像であるかどうかを判定する画像処理方法は、複数の画素からなる画像を取得し、特定種類の画像らしさの度合いを画素値および画素位置ごとに示すルックアップテーブルを格納し、取得した画像に基づいて判定用画像を生成し、判定用画像の画素の画素値および画素位置に基づき、ルックアップテーブルを用いて、その画素における特定種類の画像らしさの度合いを求め、求めた判定用画像の画素の画像らしさの度合いを積算し、積算した結果に基づき、入力画像が特定種類の画像であるかどうかを判定する。



IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告書

明 細 書

画像処理方法、画像処理装置、画像処理プログラム、撮像装置 技術分野

[0001] 本発明は、取得した画像において特定の画像があるかどうか判定する画像処理方法、画像処理装置、画像処理プログラム、撮像装置に関する。

背景技術

[0002] デジタル画像処理において、撮影された画像の中から顔画像を検出する処理の需要は高い。例えば、デジタルカメラにおいて、検出された顔領域を好ましい色や階調に変換する処理や、ビデオ画像において、特定の人物の登場場面を抽出する処理や、監視カメラにおいて、不審者の画像を抽出する処理などがある。

[0003] 顔を検出するための画像処理としては、入力画像を様々な倍率で縮小し、縮小した画像の様々な位置に顔判定領域を設定し、その顔判定領域が顔であるか所定の顔判定方法により判定する方法が一般的である。従来の顔判定方法としては、テンプレートマッチング、SVM(サポートベクターマシン)、ニューラルネットワーク、AdaBoostなどが提案されている。

[0004] テンプレートマッチングによる顔判定では、テンプレート顔画像と判定対象画像との一致度を算出することにより、顔判定を行う。SVMによる顔判定では、多くのサポートベクター(学習サンプルから選ばれたテンプレート顔画像と非顔画像)と判定対象画像との一致度を算出することにより、判定対象顔画像の様々なバリエーションに柔軟に対応することができる。

[0005] ニューラルネットワーク(3層パーセプトロン)による顔判定の動作はSVMに少し似ているが、サポートベクターの代わりに、学習によって求められた加重係数マップを複数用いる。AdaBoostによる顔判定の動作は、ニューラルネットワークに少し似ているが、学習によって求めた加重係数マップの代わりに、学習によって選ばれた矩形フィルタを複数用いる。

[0006] 特許文献1:特開2005-44330号公報

発明の開示

発明が解決しようとする課題

[0007] しかしながら、テンプレートマッチングによる顔判定では、判定対象顔画像の様々なバリエーションに柔軟に対応することができない。特に、テンプレート顔画像と判別対象顔画像との照明条件が異なる場合などには、正しく判定できない場合が多い。また、SVMによる顔判定では、多くのサポートベクターと判定対象画像との一致度を算出するために、膨大な処理時間が必要であるという問題がある。

[0008] また、ニューラルネットワークによる顔検出はSVMに比べれば高速であるものの、AdaBoostなどに比べると判定に必要な処理時間がやや長いという問題がある。また、矩形フィルターを用いたAdaBoostによる顔判定は、SVMやニューラルネットワークよりも高速であるが、判定対象領域を変更するたびに矩形フィルターを演算する必要があるため、処理時間が必要である。

課題を解決するための手段

[0009] 本発明の第1の態様によると、特定種類の画像であるかどうかを判定する画像処理方法は、複数の画素からなる画像を取得し、特定種類の画像らしさの度合いを画素値および画素位置ごとに示すルックアップテーブルを格納し、取得した画像に基づいて判定用画像を生成し、判定用画像の画素の画素値および画素位置に基づき、ルックアップテーブルを用いて、その画素における特定種類の画像らしさの度合いを求め、求めた判定用画像の画素の画像らしさの度合いを積算し、積算した結果に基づき、入力画像が特定種類の画像であるかどうかを判定する。

本発明の第2の態様によると、第1の態様の画像処理方法において、特定種類の画像は、顔の画像であるのが好ましい。

本発明の第3の態様によると、第1または第2の態様の画像処理方法において、判定用画像は、取得した画像のエッジ成分を抽出して生成されるのが好ましい。

本発明の第4の態様によると、第1または第2の態様の画像処理方法において、判定用画像は、取得した画像の局所的に周辺より画素値がへこんでいる凹構造のエッジ成分を抽出して生成されるのが好ましい。

本発明の第5の態様によると、第1の態様の画像処理方法において、ルックアップテーブルは、特定種類の画像の特徴的な要素に対応する画素位置では、その画素

のエッジ成分が大きい場合の特定種類の画像らしさの度合いを、エッジ成分が小さい場合の特定種類の画像らしさの度合いに比べて大きな値とし、特定種類の画像の特徴的な要素以外に対応する画素位置では、その画素のエッジ成分が大きい場合の特定種類の画像らしさの度合いを、エッジ成分が小さい場合の特定種類の画像らしさの度合いに比べて小さな値とするのが好ましい。

本発明の第6の態様によると、第2の態様の画像処理方法において、判定用画像は、取得した画像のエッジ成分を抽出して生成され、目鼻口のいずれかの領域に対応する画素位置では、その画素のエッジ成分が大きい場合の顔の画像らしさの度合いを、エッジ成分が小さい場合の顔の画像らしさの度合いに比べて大きな値とし、目鼻口以外の領域に対応する画素位置では、その画素のエッジ成分が大きい場合の顔の画像らしさの度合いを、エッジ成分が小さい場合の顔の画像らしさの度合いに比べて小さな値とするのが好ましい。

本発明の第7の態様によると、第1から第6のいずれかの態様の画像処理方法において、ルックアップテーブルは、特定種類の画像に属する判定対象画像サンプル群と特定種類の画像に属さない非判定対象画像サンプル群とに基づく統計処理により生成されるのが好ましい。

本発明の第8の態様によると、第7の態様の画像処理方法において、統計処理において、判定用画像を生成するときと等価な処理により、判定対象画像サンプル群に基づいて第1の画像サンプル群を生成し、非判定対象画像サンプル群に基づいて第2の画像サンプル群を生成し、第1の画像サンプル群の画素位置 (x,y) における画素値が E となる頻度 $P_1(x,y)(E)$ と、第2の画像サンプル群の画素位置 (x,y) における画素値が E となる頻度 $P_2(x,y)(E)$ とを求め、判定用画像の画素位置 (x,y) における画素値 E に対してその画素における特定種類の画像らしさの度合い $V(x,y)$ を、 $V(x,y) = L(x,y)(E)$ で与える画素位置 (x,y) におけるルックアップテーブル $L(x,y)(E)$ を、 $L(x,y)(E) = f(P_1(x,y)(E), P_2(x,y)(E))$ により生成し、関数 $f(P_1(x,y)(E), P_2(x,y)(E))$ は、 $P_1(x,y)(E)$ について実質的に広義の単調増加関数であり、 $P_2(x,y)(E)$ について実質的に広義の単調減少関数であるのが好ましい。

本発明の第9の態様によると、第8の態様の画像処理方法において、関数 $f(P_1(x,y)(E), P_2(x,y)(E))$ は、 $P_1(x,y)(E)$ について実質的に広義の単調増加関数であり、 $P_2(x,y)(E)$ について実質的に広義の単調減少関数であるのが好ましい。

E), $P_2(x,y)(E)$ は、 $f(P_1(x,y)(E), P_2(x,y)(E)) = \log\{(P_1(x,y)(E) + \varepsilon_1) / (P_2(x,y)(E) + \varepsilon_2)\}$ であり、 ε_1 と ε_2 は所定の定数であるのが好ましい。

本発明の第10の態様によると、第1の態様の画像処理方法において、コントラストの程度に応じた複数のルックアップテーブルを格納し、取得した画像のコントラストを算出し、複数のルックアップテーブルからコントラストに応じたルックアップテーブルを選択するのが好ましい。

本発明の第11の態様によると、特定種類の画像であるかどうかを判定する画像処理方法は、複数の画素からなる画像を取得し、特定種類の画像らしさの度合いを画素値および画素位置ごとに示すルックアップテーブルを格納し、複数の異なる縮小倍率により取得した画像の複数の縮小画像を生成し、複数の縮小画像に基づいて判定用画像を生成し、複数の縮小画像の1つである第1の縮小画像に対して判定対象領域を設定し、判定対象領域の画素の画素値および判定対象領域内の画素位置に基づき、ルックアップテーブルを用いて、その画素における特定種類の画像らしさの度合いを求め、求めた判定対象領域の画素の特定種類の画像らしさの度合いを積算し、積算した結果に基づき、取得した画像内の判定対象領域に対応する画像が特定種類の画像であるかどうかを判定する。

本発明の第12の態様によると、第11の態様の画像処理方法において、第1の縮小画像よりもさらに縮小された第2の縮小画像に対して、判定対象領域に対応する第2の判定対象領域をさらに設定し、特定種類の画像らしさの度合いを画素値および第2の判定対象領域に対応した画素位置ごとに示す第2のルックアップテーブルをさらに格納し、第2の判定対象領域の画素の画素値および第2の判定対象領域内の画素位置に基づき、第2のルックアップテーブルを用いて、その画素における特定種類の画像らしさの度合いを求め、求めた第2の判定対象領域の画素の特定種類の画像らしさの度合いを積算し、判定対象領域の画素の特定種類の画像らしさの度合いの積算結果および第2の判定対象領域の画素の特定種類の画像らしさの度合いの積算結果に基づき、取得した画像内の判定対象領域に対応する画像が特定種類の画像であるかどうかを判定するのが好ましい。

本発明の第13の態様によると、画像処理プログラムは、第1から12のいずれかの態

様の画像処理方法をコンピュータに実行させる画像処理プログラムとするものである。

本発明の第14の態様によると、画像処理装置は、第13の態様の画像処理プログラムを搭載する画像処理装置とするものである。

本発明の第15の態様によると、撮像装置は、第13の態様の画像処理プログラムを搭載する撮像装置とするものである。

発明の効果

[0010] 本発明は以上のように構成しているので、特定種類の画像を、照明条件の様々なバリエーションに影響を受けずに高速に判定することができる。

図面の簡単な説明

[0011] [図1]本発明の一実施の形態である画像処理装置を示す図である。

[図2]パーソナルコンピュータ1が実行する画像処理プログラムのフローチャートを示す図である。

[図3]エッジ抽出対象画素と周辺画素とを座標 xy で表した図である。

[図4]輝度の様々な構造に対して輝度凹部画像 $E_1(x,y)$ を作成した結果を示す図である。

[図5]具体的な顔の輝度画像について4種類のエッジ画像 $E_1(x,y) \sim E_4(x,y)$ を生成した例を示す図である。

[図6]具体的なエッジ画像について、顔らしさ $V_1(x,y)$ を生成し、顔らしさ V_{SUM1} を算出する処理を行った例を示す図である。

[図7]ルックアップテーブル $L_{1(x,y)}(E)$ の具体的な値をエッジの大きさ毎に表した図である。

[図8]図2のステップS6の顔判定の処理において、部分画像の顔らしさ $V_{sum1} \sim V_{sum4}$ を求めた後の処理のフローチャートを示す図である。

[図9]顔らしさ $L_{1(x,y)}(E)$ を求める処理のフローチャートを示す図である。

[図10]パーソナルコンピュータ1が実行する第2の実施の形態の画像処理プログラムのフローチャートを示す図である。

[図11]通常サイズエッジ画像および縮小サイズエッジ画像に対して、各部分画像の

顔らしさを求めた後の処理のフローチャートを示す図である。

[図12]撮像装置であるデジタルカメラ100の構成を示す図である。

発明を実施するための最良の形態

[0012] ー第1の実施の形態ー

図1は、本発明の一実施の形態である画像処理装置を示す図である。画像処理装置は、パーソナルコンピュータ1で実現される。パーソナルコンピュータ1は、デジタルカメラ2、CD-ROMなどの記録媒体3、他のコンピュータ4などと接続され、各種の画像(画像データ)の提供を受ける。パーソナルコンピュータ1は、提供された画像に対して、以下に説明する画像処理を行う。コンピュータ4は、インターネットやその他の電気通信回線5を経由して接続される。

[0013] パーソナルコンピュータ1が画像処理のために実行するプログラムは、図1の構成と同様に、CD-ROMなどの記録媒体や、インターネットやその他の電気通信回線を経由した他のコンピュータから提供され、パーソナルコンピュータ1内にインストールされる。パーソナルコンピュータ1は、CPU(不図示)およびその周辺回路(不図示)から構成され、CPUがインストールされたプログラムを実行する。

[0014] プログラムがインターネットやその他の電気通信回線を経由して提供される場合は、プログラムは、電気通信回線、すなわち、伝送媒体を搬送する搬送波上の信号に変換して送信される。このように、プログラムは、記録媒体や搬送波などの種々の形態のコンピュータ読み込み可能なコンピュータプログラム製品として供給される。

[0015] 本実施の形態のパーソナルコンピュータ1は、撮影された画像の中から顔画像を検出する画像処理を行う。具体的には、入力した画像に基づきエッジ成分を抽出してエッジ画像を生成し、生成したエッジ画像に基づき顔の画像があるかどうかを判定する。本実施の形態における処理では、このエッジ成分の抽出方法およびエッジ画像に基づく顔の判定方法に特徴を有する。

[0016] なお、以下では、画像に対して画像処理を行うという表現をするが、実際には入力した画像データに対して画像処理を行うことを意味する。また、本実施の形態で言うエッジとは、輝度値や画素値が周囲より小さい値を示しへこんでいる箇所(領域、画素)、周囲より大きい値を示し出っ張っている(突出している)箇所(領域、画素)、段

差になっている箇所(領域、画素)のことを言う。特に、周囲よりへこんでいる箇所(領域、画素)を凹構造のエッジ、周囲より出っ張っている箇所(領域、画素)を凸構造のエッジと言う。

[0017] 以下、本実施の形態のパーソナルコンピュータ1が撮影された画像の中から顔画像を検出する画像処理について詳細に説明する。図2は、パーソナルコンピュータ1が実行する画像処理プログラムのフローチャートを示す図である。

[0018] ステップS1では、デジタルカメラなどで撮影(撮像)した顔を検出する対象の画像(画像データ)を入力(取得)する。入力画像の各画素はR, G, Bの各色成分を含み、各色成分の範囲は0～255とする。ステップS2では、入力画像のR, G, Bに基づき、輝度画像Yを次の式で生成する。すなわち、輝度画像Y面を生成する。

$$Y = (R + 2G + B) / 4$$

[0019] ステップS3では、生成した輝度画像を階層的に縮小して出力する。例えば、0～31までの整数nに対して縮小倍率 κ を 0.9^n で与え、その32通りの縮小倍率 κ で縮小された輝度画像を出力する。なお、縮小方法は、例えばCubic変倍や線形変倍を用いればよい。このように複数通りの縮小画像を生成するのは、入力した画像にはどのようなサイズの顔画像があるかどうか不明であり、あらゆるサイズの顔の画像に対応できるようにするためである。

[0020] ステップS4では、縮小されたそれぞれの輝度画像 $Y(x, y)$ から4種類のエッジ画像 $E_1(x, y) \sim E_4(x, y)$ を以下の手順で生成する。以下では、x方向を画像の横方向あるいは水平方向、y方向を縦方向あるいは鉛直方向とする。

[0021] まず、以下の式より、縦方向に平滑化した画像 $Y_{LV}(x, y)$ と横方向に平滑化した画像 $Y_{LH}(x, y)$ を生成する。縦方向のエッジ成分を抽出するためには、横方向を平滑化した画像データを使用し、横方向のエッジ成分を抽出するためには、縦方向を平滑化した画像データを使用するのが好ましいからである。

$$Y_{LV}(x, y) = \{Y(x, y-1) + 2 \times Y(x, y) + Y(x, y+1)\} / 4$$

$$Y_{LH}(x, y) = \{Y(x-1, y) + 2 \times Y(x, y) + Y(x+1, y)\} / 4$$

[0022] 次に、横方向を平滑化した画像 $Y_{LH}(x, y)$ を使用して、以下の式より、縦方向のエッジ画像 $E_1(x, y)$ を生成する。なお、エッジ画像の各画素はエッジ画素と言う。

$$E_1^{\wedge}(x,y) = \text{Min}(Y_{LH}(x,y-1), Y_{LH}(x,y+2)) \\ - \text{Min}(Y_{LH}(x,y), Y_{LH}(x,y+1))$$

$$E_1(x,y) = \gamma(E_1^{\wedge}(x,y))$$

[0023] 次に、以下の式より、縦方向のエッジ画像 $E_2(x,y)$ を生成する。

$$E_2^{\wedge}(x,y) = |Y_{LH}(x,y-1) - Y_{LH}(x,y)| \\ + |Y_{LH}(x,y+1) - Y_{LH}(x,y)|$$

$$E_2(x,y) = \gamma(E_2^{\wedge}(x,y))$$

[0024] 次に、縦方向を平滑化した画像 $Y_{LV}(x,y)$ を使用して、以下の式より、横方向のエッジ画像 $E_3(x,y)$ を生成する。

$$E_3^{\wedge}(x,y) = \text{Min}(Y_{LV}(x-1,y), Y_{LV}(x+2,y)) \\ - \text{Min}(Y_{LV}(x,y), Y_{LV}(x+1,y))$$

$$E_3(x,y) = \gamma(E_3^{\wedge}(x,y))$$

[0025] 次に、以下の式より、横方向のエッジ画像 $E_4(x,y)$ を生成する。

$$E_4^{\wedge}(x,y) = |Y_{LV}(x-1,y) - Y_{LV}(x,y)| \\ + |Y_{LV}(x+1,y) - Y_{LV}(x,y)|$$

$$E_4(x,y) = \gamma(E_4^{\wedge}(x,y))$$

[0026] ここで、 $\text{Min}()$ は、 $()$ の中の最小の値を戻す関数である。また、 $\gamma(E)$ は、 γ 変換とクリッピングを行う関数であり、以下の演算を行い、0～31の整数を出力する。この $\text{MIN}()$ 処理は、非線形フィルタ処理である。また、 γ 変換やクリッピング処理を含めて非線形フィルタ処理と言ってもよい。

$$E < 0 \text{ の場合 } \gamma(E) = 0$$

$$E > 63 \text{ の場合 } \gamma(E) = 31$$

$$0 \leq E \leq 63 \text{ の場合 } \gamma(E) = (\text{int})(4 \times \sqrt{E})$$

[0027] 上記エッジ画像の生成について、図3を参照してさらに詳しく説明する。図3は、エッジ抽出対象画素と周辺画素とを座標 xy で表した図である。上記 $E_1^{\wedge}(x,y)$ は、輝度画像 $Y_{LH}(x,y)$ 面において、縦方向4画素 $Y_{LH}(x,y-1)$ 、 $Y_{LH}(x,y)$ 、 $Y_{LH}(x,y+1)$ 、 $Y_{LH}(x,y+2)$ のうち、対象画素 (x,y) を基準に、外側2画素 $Y_{LH}(x,y-1)$ 、 $Y_{LH}(x,y+2)$ の最小値と内側2画素 $Y_{LH}(x,y)$ 、 $Y_{LH}(x,y+1)$ の最小値の差を求めている。

- [0028] $E_1^-(x,y)$ の値が正の値を示すことは、対象画素(x, y)近辺の値が、縦方向周辺画素の値より小さい、すなわち画素値が縦方向の周辺よりへこんでいることを示す。従って、このようにして生成した $E_1^-(x,y)$ の値を画素値として取り扱い、生成された画像を縦方向輝度凹部画像と言う。
- [0029] 上記 $E_2^-(x,y)$ は、輝度画像 $Y_{LH}(x,y)$ 面において、対象画素(x, y)と縦方向に隣接する画素との輝度値の差分を足し込んだ値を示す。すなわち、縦方向隣接画素との間で輝度値の変化が大きい場合に大きな値が生成される。従って、このようにして生成した $E_2^-(x,y)$ の値を画素値として取り扱い、生成された画像を縦方向隣接画素差分画像と言う。縦方向隣接画素差分画像は、凹部構造のエッジ、凸部構造のエッジ、段差のエッジを区別なく検出する。
- [0030] 上記 $E_3^-(x,y)$ および $E_3^-(x,y)$ 、 $E_4^-(x,y)$ および $E_4^-(x,y)$ は、横方向のエッジ画像を生成するためのものである。上記 $E_1^-(x,y)$ および $E_1^-(x,y)$ 、 $E_2^-(x,y)$ および $E_2^-(x,y)$ に対して、縦と横をひっくり返して考え、後は同様に演算するものである。従って、このようにして生成された $E_3^-(x,y)$ を横方向輝度凹部画像、 $E_4^-(x,y)$ を横方向隣接画素差分画像と言う。
- [0031] 図4は、輝度の様々な構造に対して輝度凹部画像 $E_1^-(x,y)$ を作成した結果を示す図である。図4(a)は輝度が凹んでいる場合であり、図4(b)は輝度が突出している場合であり、図4(c)は輝度が段差になっている場合である。図4を見ると、輝度が凹んでいる場合のみ輝度凹部画像が正の値を持つことがわかる。従って、輝度凹部画像 E_1^- の負の値を0にクリッピングすれば、輝度の凹みだけに反応するエッジ画像 $E_1^-(x,y)$ が生成される。
- [0032] この輝度凹部画像によると、目鼻口などの局所的に暗い箇所特に良く反応する。図5は、具体的な顔の輝度画像について上記4種類のエッジ画像 $E_1^-(x,y) \sim E_4^-(x,y)$ を生成した例を示す図である。実際、輝度凹部画像は、目鼻口の位置に鋭いピークを持つ。特に、図5の縦方向輝度凹部画像 E_1^- では、目、鼻の穴、口などに反応し、その中でも目、鼻の穴などには強く反応し白くなっている。すなわち、その位置の E_1^- の値が大きな値となっている。従って、このような輝度凹部画像を解析することにより、顔を高精度に検出することができる。ただし、輝度凹部画像だけを用いるのではなく

、従来の方法で作成したエッジ画像も合わせて用いることが望ましい。

[0033] なお、上記エッジ画像 E' をガンマ変換した理由は、エッジ量 E' を適切な特徴量 E に変換するためである。画像解析において、ほとんどエッジがない箇所での微妙なエッジ量の違いは、大きなエッジがある箇所での多少のエッジ量の違いよりも大きな意味を持つ。エッジ量 E' に対してガンマ変換を施すことにより上記の効果が実現され、ほとんどエッジがない箇所でのエッジ量の違いは特徴量 E の大きな違いに変換され、大きなエッジがある箇所でのエッジ量の違いは特徴量 E の小さな違いに変換される。

[0034] 次に、図2に戻って、ステップS5では、縮小した画像の1画素おきに 19×19 画素の顔判定対象領域を設定し、その領域におけるエッジ画像の部分画像を出力する。これをすべての縮小画像において行う。 19×19 画素の顔判定対象領域は、その領域が顔である場合に目や鼻や口などが2画素程度で検出できるのに適したサイズである。

[0035] ステップS6では、ステップ5で出力したエッジ画像の各部分画像に対して、この領域が顔の画像であるかどうか判定する。本実施の形態では、この顔の画像の判定を以下に説明する手法により行う。

[0036] まず、エッジ画像 $E_1(x,y)$ の部分画像の各画素位置 (x,y) ($0 \leq x \leq 18$, $0 \leq y \leq 18$)について、次の式に基づいてその位置の顔らしさ $V_1(x,y)$ を生成する。顔らしさ $V_1(x,y)$ は、各画素位置で顔らしさを数値化したもので、顔らしさの度合いや程度を示すものである。 $V_1(x,y)$ は、顔として尤もらしい度合いを表す尤度といってもよい。

$$V_1(x,y) = L_{1(x,y)}(E_1(x,y))$$

ここで、 $L_{1(x,y)}(E)$ は、各画素位置 (x,y) ($0 \leq x \leq 18$, $0 \leq y \leq 18$)について、後述する統計処理によりあらかじめ作成されているルックアップテーブルであり、画素位置 (x,y) のエッジ $E_1(x,y)$ が E である時のその箇所の顔らしさを表す。

[0037] そして、生成した顔らしさ $V_1(x,y)$ を全画素 (x,y) ($0 \leq x \leq 18$, $0 \leq y \leq 18$)について積算し、顔らしさ V_{SUM1} を算出する。

[0038] 図6は、具体的なエッジ画像について上記の処理を行った例を示す図である。図6の顔らしさ画像では、顔らしい箇所が白く表示され、顔らしくない箇所が黒く表示されている。図6(a)に示す顔のエッジ画像から生成した顔らしさ画像は、全体的に大きな

値を持つ。すなわち、全体的に白っぽい画像となる。しかし、図6(b)に示す非顔のエッジ画像から生成した顔らしさ画像は所々小さな値を持つ。すなわち、所々黒っぽくなった画像となる。

[0039] 図6(b)の非顔の例では、目の間、鼻、口の両横に対応する領域が顔らしくないとされて、顔らしさ画像ではその領域の画素値は小さな値となり黒い画像となっている。従って、非顔画像の顔らしさ画像を全画素積算した値 V_{SUM1} は小さな値になる。

[0040] 図7は、ルックアップテーブル $L_{1(x,y)}(E)$ の具体的な値をエッジの大きさ毎に表した図である。図7では、顔らしさの値が大きいほど白く表示されている。図7において、左側はエッジが小さい時の顔らしさであり、右側はエッジが大きい時の顔らしさである。なお、ルックアップテーブル $L_{1(x,y)}(E)$ の全ての値を図示するなら、前述の通りエッジは0~31の値で生成されているので、 $L_{1(x,y)}(0) \sim L_{1(x,y)}(31)$ の32通りの図ができる。しかし、図7では、図示の便宜上そのうちの8通りのみ表示している。

[0041] なお、図7のルックアップテーブル $L_{1(x,y)}(E)$ は、具体的な値をエッジの大きさ毎に視覚的に表した図である。実際には、画素位置(x, y)を引数とした画素値のテーブルが、エッジの値毎にメモリに格納されている。すなわち、32個の画素位置(x, y)を引数とした画素値のテーブルがメモリに格納されている。

[0042] 図7において、左側の図はエッジが小さい時の顔らしさを表す。左側の図を見ると、目、鼻、口の箇所の顔らしさが小さな値になっている。これは、目、鼻、口の箇所のエッジが小さい場合には、その箇所は顔らしくないということを表している。例えば、図6(a)の非顔の例では、鼻に対応する箇所のエッジが小さいので、その箇所は顔らしくないとされる。

[0043] また、図7の右側の図はエッジが大きい時の顔らしさを表す。右側の図を見ると、目、鼻、口以外の箇所の顔らしさが小さな値になっている。これは、目、鼻、口以外の箇所のエッジが大きい場合には、その箇所は顔らしくないということを表している。例えば、図6(a)の非顔の例では、目の間と口の両横に対応する箇所のエッジが大きいので、その箇所は顔らしくないとされる。

[0044] すなわち、顔の画像を特定種類の画像とし、目、鼻、口などを特定種類の画像の特徴的な要素であると考え、特定種類の画像の特徴的な要素に対応する画素位置

では、その画素のエッジ成分が大きい場合の特定種類の画像らしさの度合いを、エッジ成分が小さい場合の特定種類の画像らしさの度合いに比べて大きな値としている。また、特定種類の画像の特徴的な要素以外に対応する画素位置では、その画素のエッジ成分が大きい場合の特定種類の画像らしさの度合いを、エッジ成分が小さい場合の特定種類の画像らしさの度合いに比べて小さな値としている。

[0045] 上記ルックアップテーブルを参照する処理を整理すると、まず、エッジ画像 $E_1(x,y)$ の部分画像において、 $x=0$ 、 $y=0$ のエッジ E_1 の値を得る。次に、このエッジ E_1 の値に相当するルックアップテーブル $L_{1(x,y)}(E_1)$ を32個のルックアップテーブルの中から決める。ルックアップテーブル $L_{1(x,y)}(E_1)$ が決まると、このルックアップテーブル $L_{1(x,y)}(E_1)$ の画素位置(0, 0)の値を得る。これが、エッジ画像 $E_1(x,y)$ の画素位置(0, 0)の顔らしさの値である。この処理を、 $x=0$ 、 $y=0$ の画素から $x=18$ 、 $y=18$ の画素まで順次行い、顔らしさ画像 $V_1(x,y)$ を得る。そして、 $V_1(x,y)$ をすべて積算して $Vsum_1$ を得る。

[0046] 以上の処理により、エッジ画像 $E_1(x,y)$ に基づいて部分画像の顔らしさ $Vsum_1$ が生成される。そして、エッジ画像 $E_2(x,y) \sim E_4(x,y)$ に基づいて部分画像の顔らしさ $Vsum_2 \sim Vsum_4$ を生成する処理も同様に行う。

[0047] 図8は、図2のステップS6の顔判定の処理において、部分画像の顔らしさ $Vsum_1 \sim Vsum_4$ を求めた後の処理のフローチャートを示す図である。ステップS6の顔判定処理では、上記に説明したように、顔らしさ $Vsum_1 \sim Vsum_4$ を段階的に生成し、それらを積算した評価値が閾値よりも大きければ顔とする。ただし、評価値を閾値と比較する処理を図8に示すように各段階において行うことにより、明らかに顔ではない画像を早い段階で除外して、効率的な処理を行えるようにしている。

[0048] まず、ステップS11では、部分画像が顔の画像であるかどうかを判定する評価値を、エッジ画像 $E_1(x,y)$ の顔らしさ $Vsum_1$ とする。ステップS12では、評価値が所定の閾値 $th1$ より大きいかどうかを判定し、この評価値が閾値 $th1$ より大きければステップS13に進み、この評価値が閾値 $th1$ より小さくなければ、部分画像は顔の画像でないとし、対象の部分画像の顔判定の処理を終了する。

[0049] ステップS13では、評価値をステップS11の評価値にエッジ画像 $E_2(x,y)$ の顔らし

さ V_{sum_2} を足した値とする。ステップS14では、この評価値が所定の閾値th2より大きいかどうかを判定し、評価値が閾値th2より大きければステップS15に進み、この評価値が閾値th2より大きくなければ、部分画像は顔の画像でないとして、対象の部分画像の顔判定の処理を終了する。

[0050] ステップS15では、評価値をステップS13の評価値にエッジ画像 $E_3(x,y)$ の顔らしさ V_{sum_3} を足した値とする。ステップS16では、この評価値が所定の閾値th3より大きいかどうかを判定し、評価値が閾値th3より大きければステップS17に進み、この評価値が閾値th3より大きくなければ、部分画像は顔の画像でないとして、対象の部分画像の顔判定の処理を終了する。

[0051] ステップS17では、評価値をステップS15の評価値にエッジ画像 $E_4(x,y)$ の顔らしさ V_{sum_4} を足した値とする。ステップS18では、この評価値が所定の閾値th4より大きいかどうかを判定する。ステップS18において、評価値が閾値th4より大きければ、最終的にこの部分画像は顔の画像であると判定する。この評価値が閾値th4より大きくなければ、部分画像は顔の画像でないとして、対象の部分画像の顔判定の処理を終了する。

[0052] 以上説明した部分画像の顔判定処理を、各縮小画像において、1ビットずつずらした各部分画像についてすべて行い、顔の画像と判定できる部分画像をすべて抽出し、ステップS7に進む。

[0053] ステップS7では、ステップ6によりある部分画像が顔であると判定された場合には、その部分画像の入力画像に対する顔の大きさ S と座標 (X, Y) を出力する。 S, X, Y は、縮小画像における顔のサイズ $S' = 19$ と、顔とされた領域の座標 (X', Y') と縮小倍率 κ とを用いて、次の式で与えられる。

$$S = S' / \kappa$$

$$X = X' / \kappa$$

$$Y = Y' / \kappa$$

[0054] 以上のようにして、入力画像に顔の画像がある場合は、その顔の画像の位置と大きさが検出されて出力される。

[0055] <統計処理>

次に、前述した統計処理について説明する。すなわち、画素位置 (x,y) のエッジ $E_1(x,y)$ がEであるときのその画素の顔らしさ $L_{1(x,y)}(E)$ を求める方法を説明する。図9は、この顔らしさ $L_{1(x,y)}(E)$ を求める処理のフローチャートを示す図である。この処理は、パーソナルコンピュータ1において実行される。

[0056] ステップS21では、数百人以上の顔の画像を取得する。すなわち、数百人以上の顔をデジタルカメラ等で撮影(撮像)し、その画像(画像データ)を取得する。取得する画像は、図2のステップS1で入力する画像と同様な色成分で構成された画像である。ステップS22では、顔が撮影されている画像を、顔領域の大きさが 19×19 画素になるように変倍して、顔領域を切り出した部分画像を顔画像サンプル群とする。

[0057] ステップS23では、 19×19 画素の非顔画像サンプル群を、数百パターン以上取得する。これは、デジタルカメラで撮影した顔以外の画像から適宜抽出して非顔画像サンプル群とする。顔が写っている画像から、顔の領域を避けて抽出するようにしてもよい。この場合は、モニタに写された画像から、ユーザが適宜非顔画像の領域を指定すればよい。

[0058] ステップS24では、顔画像サンプル群からエッジ成分を抽出して、顔エッジ画像サンプル群を生成する。この処理は、顔検出処理においてエッジ画像 $E_1(x,y)$ を生成する処理と同様に行う。ステップS25では、非顔画像サンプル群からエッジ成分を抽出して、非顔エッジ画像サンプル群を生成する。この処理も、顔検出処理においてエッジ画像 $E_1(x,y)$ を生成する処理と同様に行う。

[0059] ステップS26では、顔エッジ画像サンプル群について、 (x,y) のエッジがEとなる頻度 $P_{\text{顔}}(x,y,E)$ を求める。すなわち、画素 (x,y) の値がEとなる画像がいくつあるかをカウントする。ステップS27では、非顔エッジ画像サンプル群について、同様に、 (x,y) のエッジがEとなる頻度 $P_{\text{非顔}}(x,y,E)$ を求める。

[0060] ステップS28では、画素位置 (x,y) のエッジ $E_1(x,y)$ がEであるときのその画素の顔らしさ $L_{1(x,y)}(E)$ を、次の式によって算出する。

$$L_{1(x,y)}(E) = \log \{ (P_{\text{顔}}(x,y,E) + \varepsilon_1) / (P_{\text{非顔}}(x,y,E) + \varepsilon_2) \}$$

ここで、 ε_1 と ε_2 は所定の定数であり、対数の発散や過学習を抑制するために導入している。 ε_1 の値は $P_{\text{顔}}(x,y,E)$ の平均的な値の1000分の1程度に設定すればよく

、 ε_2 の値は ε_1 の値の数十倍に設定すればよい。

[0061] 上記 $L_{1(x,y)}(E)$ を求める式において、 $\log\{(P_{顔}(x,y,E) + \varepsilon_1)\}$ は、単調増加関数であり、 $\log\{1/(P_{非顔}(x,y,E) + \varepsilon_2)\}$ は、単調減少関数である。すなわち、顔らしさ $L_{1(x,y)}(E)$ は、画素位置 (x,y) のエッジ $E_1(x,y)$ が E である顔画像サンプルの分布が増加していく方向にその値は単調増加し、画素位置 (x,y) のエッジ $E_1(x,y)$ が E である非顔画像サンプルの分布が増加していく方向にその値は単調減少していく関数である。なお、画素位置 (x,y) のエッジ $E_1(x,y)$ が E である顔画像サンプルの分布、および、画素位置 (x,y) のエッジ $E_1(x,y)$ が E である非顔画像サンプルの分布は、通常正規分布している。

[0062] エッジ画像 $E_2(x,y) \sim E_4(x,y)$ を顔らしさに変換するルックアップテーブル $L_{2(x,y)}(E) \sim L_{4(x,y)}(E)$ を生成するには、上記ステップS24、ステップS25のエッジ成分抽出処理を、顔検出処理におけるエッジ画像 $E_2(x,y) \sim E_4(x,y)$ を生成する処理と同様にすればよい。

[0063] 以上説明した第1の実施の形態の処理を行うと、次のような効果を奏する。

(1) 顔画像の目、鼻、口などの位置は周辺に比べて局所的に暗い。従来のエッジ抽出方法では、エッジ構造が局所的に暗い構造であるか、局所的に明るい構造であるか、あるいはそれ以外の構造であるかを識別することはできなかった。しかし、上記のように凹部構造のエッジを検出し、エッジ画像である凹部画像を生成することにより顔画像の局所的に暗い構造である目、鼻、口などを適切に抽出することができる。その結果、顔の画像を正確に判定することができる。

[0064] (2) 輝度凹部画像によると、目鼻口などの局所的に暗い箇所特に良く反応する。実際、輝度凹部画像は、目鼻口の位置に鋭いピークを持つ。従って、このような輝度凹部画像を解析することにより、顔を高精度に検出することができる。本実施の形態では、輝度凹部画像だけを用いるのではなく、従来の方法で作成したエッジ画像も合わせて用いるようにしているので、さらにより精度の高い顔の判定を可能としている。

[0065] (3) 上記エッジ E' エッジをガンマ変換した理由は、エッジ量 E' を適切な特徴量 E に変換するためである。画像解析において、ほとんどエッジがない箇所での微妙なエッ

ジ量の違いは、大きなエッジがある箇所での多少のエッジ量の違いよりも大きな意味を持つ。エッジ量 E' に対してガンマ変換を施すことにより、ほとんどエッジがない箇所でのエッジ量の違いは特徴量 E の大きな違いに変換され、大きなエッジがある箇所でのエッジ量の違いは特徴量 E の小さな違いに変換される。これにより、エッジ量の違いが画像の構造の違いに一致するようになる。この結果、顔判定の精度も高くなる。

[0066] (4) 上記実施の形態の図4から明らかなように、輝度が凹んでいる場合のみ輝度凹部画像が正の値を持つことがわかる。従って、本実施の形態では、輝度凹部画像 E' の負の値を0にクリッピングするようにした。これにより、輝度の凹みだけに反応するエッジ画像 $E_1(x,y)$ が生成され、エッジ画像 E_1 を使用する処理がしやすくなる。

[0067] (5) エッジ画像の画素値をルックアップテーブルを用いて顔らしさに変換して積算するという単純で高速な処理により、顔の画像を検出することができる。また、エッジ画像を判定することにより、画像を撮影する際の照明条件の影響を抑制する効果がある。

[0068] ー第2の実施の形態ー

第2の実施の形態では、判定対象画像のコントラストの変動に強い顔判定方法を説明する。第2の実施の形態は、第1の実施の形態と同様に、パーソナルコンピュータ1で実現される。従って、第2の実施の形態の画像処理装置の構成は、第1の実施の形態の図1を参照することとする。

[0069] <統計処理>

まず、以下で説明する統計処理を行い、第2の実施の形態の顔判定用のルックアップテーブル(LUT)を作成する。第2の実施の形態のルックアップテーブルの作成について、第1の実施の形態の図9を参照しながら以下説明をする。

[0070] まず、図9のステップS21からステップS23までと同様にして、数百以上の 19×19 画素の顔画像サンプル群と、数百以上の 19×19 画素の非顔画像サンプル群とを取得する。

[0071] 次に、顔画像サンプル群の画素値にゲインをかけ、画素値の分散が100程度になるように調整する。または、顔画像サンプル群の中から、画素値の分散が200未満であるものを抽出する。このようにして調整または抽出した顔画像サンプル群と、先に求

めた非顔画像サンプル群を使って、図9のステップS24からステップS28までと同様にして、顔判定用のルックアップテーブルを作成する。このようにして求めたルックアップテーブルを、低コントラスト顔判定用ルックアップテーブルと言う。

[0072] 次に、顔画像サンプル群の画素値に上記とは異なるゲインをかけ、画素値の分散が400程度になるように調整する。または、顔画像サンプル群の中から、画素値の分散が200以上であるものを抽出する。このようにして調整または抽出した顔画像サンプル群と、先に求めた非顔画像サンプル群を使って、図9のステップS24からステップS28までと同様にして、顔判定用のルックアップテーブルを作成する。このようにして求めたルックアップテーブルを、高コントラスト顔判定用ルックアップテーブルと言う。

[0073] 次に、上記のようにして求めた低コントラスト顔判定用ルックアップテーブルと高コントラスト顔判定用ルックアップテーブルを使用して、撮影された画像の中から顔画像を検出する画像処理について説明する。図10は、パーソナルコンピュータ1が実行する第2の実施の形態の画像処理プログラムのフローチャートを示す図である。

[0074] ステップS31からステップS34は、第1の実施の形態の図2のステップS1からステップS4と同様である。ステップS38では、輝度画像の積分画像 $I(x,y)$ と輝度画像の画素値の二乗の積分画像 $I_2(x,y)$ を次の式に基づいて作成する。

[数1]

【数 1】

$$I(x,y) = \sum_{i=0}^x \sum_{j=0}^y Y(i,j)$$

$$I_2(x,y) = \sum_{i=0}^x \sum_{j=0}^y Y(i,j)^2$$

[0075] ステップS35では、第1の実施の形態の図2のステップS5と同様に顔判定対象領域を設定する。ステップS39では、顔判定対象領域内の輝度画像 $Y(x,y)$ の画素値の分散 σ^2 を計算する。顔判定対象領域を、4点 $(x,y), (x+w,y), (x,y+h), (x+w,y+h)$ を頂点とす

る長方形領域とすると、その領域内で輝度画像 $Y(x,y)$ を積分した値 Y_{sum} と輝度画像の二乗を積分した値 Y_{sum2} 、は、次の式で算出される。

[数2]

【数 2】

$$Y_{sum} = I(x+w, y+h) - I(x-1, y+h) - I(x+w, y-1) + I(x-1, y-1)$$

$$Y_{sum2} = I_2(x+w, y+h) - I_2(x-1, y+h) - I_2(x+w, y-1) + I_2(x-1, y-1)$$

[0076] 上記演算によると、4点の画素値の加減算を行うだけで積分が求まるので、高速な演算が可能である。そして、顔判定対象領域内の輝度画像 $Y(x,y)$ の画素値の分散 σ^2 は次の式で与えられる。

[数3]

【数 3】

$$\sigma^2 = \frac{Y_{sum2}}{w \times h} - \left(\frac{Y_{sum}}{w \times h} \right)^2$$

[0077] ステップS40では、上記分散 σ^2 が200未満の場合には、低コントラスト顔検出用ルックアップテーブルを選択する。また、上記分散 σ^2 が200以上の場合には、高コントラスト顔検出用ルックアップテーブルを選択する。なお、分散 σ^2 が大きい場合は高コントラストの画像であることを示し、分散 σ^2 が小さい場合は低コントラストの画像であることを示している。

[0078] ステップS36では、ステップS40で選択された顔検出用ルックアップテーブルを用いて、第1の実施の形態のステップS6と同様の方法で顔判定処理を行う。ステップS37では、第1の実施の形態のステップS7と同様に検出結果を出力する。

[0079] 本実施の形態によると、顔判定対象領域のコントラストを高速に測定し、そのコントラストに応じて顔検出用ルックアップテーブルを選択することにより、判定処理時間を抑えながら、様々なコントラストに対して精度が高い判定をすることができる。

[0080] なお、コントラストによって異なるルックアップテーブルを使用するのは、コントラストが高い画像であると、エッジが大きめに出すぎてしまうためである。すなわち、コントラストが高いものは高いなりのルックアップテーブルを使用することにより、精度の高い顔判定が可能となるからである。

[0081] ー第3の実施の形態ー

第3の実施の形態では、異なる複数の解像度の画像を用いることにより、より高精度な顔判定を行う方法を説明する。第3の実施の形態は、第1の実施の形態と同様に、パーソナルコンピュータ1で実現される。従って、第3の実施の形態の画像処理装置の構成は、第1の実施の形態の図1を参照することとする。

[0082] <統計処理>

まず、以下で説明する統計処理を行い、第3の実施の形態の顔判定用のルックアップテーブル(LUT)を作成する。第3の実施の形態のルックアップテーブルの作成について、第1の実施の形態の図9を参照しながら以下説明をする。

[0083] まず、図9のステップS21からステップS28までと同様にして、第1の実施の形態と同様な顔判定用ルックアップテーブルを作成する。以後、このルックアップテーブルを、通常サイズ顔判定用ルックアップテーブルと言う。

[0084] 次に、図9のステップS22の段階で取得した顔画像サンプル群を縮小し、12×12画素程度の大きさにする。同様にして、図9のステップS23の段階で取得した非顔画像サンプル群を縮小し、12×12画素程度の大きさにする。このようにして作成した顔画像サンプル群と非顔画像サンプル群を使って、図9のステップS24からステップS28までと同様にして、顔判定用のルックアップテーブルを作成する。このようにして求めたルックアップテーブルを、縮小サイズ顔判定用ルックアップテーブルと呼ぶ。

[0085] 次に、上記のようにして求めた通常サイズ顔判定用ルックアップテーブルと縮小サイズ顔判定用ルックアップテーブルを使用して、撮影された画像の中から顔画像を検出する画像処理について説明する。パーソナルコンピュータ1が実行する第3の実施の形態の画像処理プログラムは、第1の実施の形態の図2のフローチャートと、処理の流れとしては同様であるので、図2を参照しながら以下説明をする。

[0086] ステップS1からステップS4は、第1の実施の形態のステップS1からステップS4と同

様である。

- [0087] ステップS5では、縮小画像の1画素おきに 19×19 画素の顔判定対象領域を設定し、その領域におけるエッジ画像 $E_1 \sim E_4$ の部分画像を出力する。ここで出力したエッジ画像を通常サイズエッジ画像と呼ぶ。さらに、前記縮小画像に対して、 0.9^4 の縮小倍率で縮小された第2の縮小画像に対して、上記 19×19 画素の顔判定対象領域と同一の被写体に対応する 12×12 画素の縮小サイズ顔判定対象領域を設定し、その領域における、前記第2の縮小画像に対して作成したエッジ画像 $E_1 \sim E_4$ の部分画像を出力する。ここで出力したエッジ画像を縮小サイズエッジ画像と呼ぶ。
- [0088] ステップS6では、通常サイズエッジ画像に対して、通常サイズ顔判定用ルックアップテーブルを用いて、第1の実施の形態と同様にして顔らしさを算出する。さらに、縮小サイズエッジ画像に対して、縮小サイズ顔判定用ルックアップテーブルを用いて、第1の実施の形態と同様にして顔らしさを算出する。
- [0089] 図11は、上記のようにして、通常サイズエッジ画像および縮小サイズエッジ画像に対して、各部分画像の顔らしさを求めた後の処理のフローチャートを示す図である。ステップS6の顔判定処理では、第1の実施の形態と同様に、顔らしさを段階的に生成し、それらを積算した評価値が閾値よりも大きければ顔とする。ただし、評価値を閾値と比較する処理を図11に示すように各段階において行うことにより、明らかに顔ではない画像を早い段階で除外して、効率的な処理を行えるようにしている。
- [0090] まず、ステップS51では、部分画像が顔の画像であるかどうかを判定する評価値を、縮小サイズエッジ画像 $E_1(x,y)$ の顔らしさ V_{sum}_1 とする。ステップS52では、評価値が所定の閾値 $th1$ より大きいかどうかを判定し、この評価値が閾値 $th1$ より大きければステップS53に進み、この評価値が閾値 $th1$ より大きくなければ、部分画像は顔の画像でないとして、対象の部分画像の顔判定の処理を終了する。
- [0091] ステップS53では、評価値をステップS51の評価値に縮小サイズエッジ画像 $E_2(x,y)$ の顔らしさ V_{sum}_2 を足した値とする。ステップS54では、この評価値が所定の閾値 $th2$ より大きいかどうかを判定し、評価値が閾値 $th2$ より大きければステップS55に進み、この評価値が閾値 $th2$ より大きくなければ、部分画像は顔の画像でないとして、対象の部分画像の顔判定の処理を終了する。

[0092] ステップS55では、評価値をステップS53の評価値に縮小サイズエッジ画像 $E_3(x,y)$ の顔らしさ V_{sum_3} を足した値とする。ステップS56では、この評価値が所定の閾値 th_3 より大きいかどうかを判定し、評価値が閾値 th_3 より大きければステップS57に進み、この評価値が閾値 th_3 より大きくなければ、部分画像は顔の画像でないとして、対象の部分画像の顔判定の処理を終了する。

[0093] ステップS57では、評価値をステップS55の評価値に縮小サイズエッジ画像 $E_4(x,y)$ の顔らしさ V_{sum_4} を足した値とする。ステップS58では、この評価値が所定の閾値 th_4 より大きいかどうかを判定し、評価値が閾値 th_4 より大きければステップS59に進み、この評価値が閾値 th_4 より大きくなければ、部分画像は顔の画像でないとして、対象の部分画像の顔判定の処理を終了する。

[0094] ステップS59からステップS66は、通常サイズエッジ画像について同様の処理を行う。その結果、ステップS66において、評価値が閾値 th_8 より大きければ、最終的にこの部分画像は顔の画像であると判定する。この評価値が閾値 th_8 より大きくなければ、部分画像は顔の画像でないとして、対象の部分画像の顔判定の処理を終了する。

[0095] ステップ7では、第1の実施の形態と同様に、顔検出結果を出力する。

[0096] 第3の実施形態によると、第1の実施の形態に比べて、さらに縮小サイズエッジ画像の顔らしさの評価を加えることにより、さらに高精度な顔判定処理を行うことができる。例えば、顔判定対象領域が 19×19 画素であると、目は2画素程度になって検出しやすいが、口は4画素程度になって検出しにくくなる。ところが、同じ顔判定対象領域が 12×12 画素になると、口は2画素程度になって検出しやすくなる。従って、このような縮小サイズエッジ画像の顔らしさの評価を加えることにより、口などの凹構造が検出しやすくなり、さらに高精度な顔判定処理を行うことができるようになる。

[0097] ー変形例ー

上記実施の形態では、顔の画像を判定する例を説明した。しかし、顔の画像以外の画像にも本発明は適用できる。すなわち、特定種類の画像が取得した画像にあるかどうかを判定する場合にも適用できる。このような場合は、その特定種類の画像らしさの度合いを画素値および画素位置ごとに示すルックアップテーブルを統計処理により準備し、このルックアップテーブルを用いて、判定用画像の各画素における特定

種類の画像らしさの度合いを求めるようにすればよい。

- [0098] 上記実施の形態では、エッジ画像として輝度凹部画像を生成し、顔の目鼻口などの局所的に暗い箇所を適切に判断する例を説明した。しかし、歯を見せて笑っている口や、光が当たって光っている頬や鼻では、輝度が周囲に比べて局所的に明るくなっている。このような顔の局所的に明るい箇所も適切に検出するために、次のような式により、エッジ画像として輝度凸部画像を生成し、同様にして顔らしさを求めるようにしてもよい。

$$E_5^{\wedge}(x,y) = \text{Max}(Y_{LH}(x,y-1), Y_{LH}(x,y+2)) \\ - \text{Max}(Y_{LH}(x,y), Y_{LH}(x,y+1)) \\ E_5(x,y) = \gamma(E_5^{\wedge}(x,y))$$

- [0099] 上記実施の形態では、顔らしさ $L_{1(x,y)}(E)$ を、次の式によって算出する例を説明した。

$$L_{1(x,y)}(E) = \log\{(P_{\text{顔}}(x,y,E) + \varepsilon_1) / (P_{\text{非顔}}(x,y,E) + \varepsilon_2)\}$$

しかし、次のような式を用いてもよい。

$$L_{1(x,y)}(E) = \sqrt{\{P_{\text{顔}}(x,y,E)\}} - \sqrt{\{P_{\text{非顔}}(x,y,E)\}}$$

第1項の $\sqrt{\{P_{\text{顔}}(x,y,E)\}}$ も単調増加関数と言え、第2項の $\sqrt{\{P_{\text{非顔}}(x,y,E)\}}$ も単調減少関数と言える。

- [0100] 上記実施の形態では、パーソナルコンピュータ1が、撮影された画像の中から顔画像を検出する画像処理を行う例を説明した。しかし、デジタルスチルカメラなどの撮像装置内で、撮像した画像に対し上記説明した処理を行うようにしてもよい。

- [0101] 図12は、このような撮像装置であるデジタルカメラ100の構成を示す図である。デジタルカメラ100は、撮影レンズ102、CCDなどからなる撮像素子103、CPUおよび周辺回路からなる制御装置104、メモリ105などから構成される。

- [0102] 撮像素子103は、被写体101を撮影レンズ102を介して撮影(撮像)し、撮影した画像データを制御装置104へ出力する。制御装置104は、撮像素子103で撮影された画像(画像データ)に対して、上記で説明した顔画像を検出する画像処理を行う。そして、制御装置104は、顔画像の検出結果に基づき撮影した画像に対し、ホワイトバランスの調整やその他の各種の画像処理を行い、画像処理後の画像データを適

宜メモリ105に格納する。また、制御装置104は、顔画像の検出結果を、オートフォーカス処理などにも利用することができる。なお、制御装置104が実行する画像処理プログラムは、不図示のROMに格納されている。

[0103] また、上記説明した処理をビデオカメラにも適用できる。さらに、不審者を監視する監視カメラや、撮影された顔画像に基づいて個人を識別したり、性別や年齢や表情を推定するような装置にも適用できる。すなわち、顔の画像など特定種類の画像を抽出して処理する画像処理装置や撮像装置などのあらゆる装置に本発明を適用することができる。

[0104] 上記では、種々の実施の形態および変形例を説明したが、本発明はこれらの内容に限定されるものではない。本発明の技術的思想の範囲内で考えられるその他の態様も本発明の範囲内に含まれる。

[0105] 次の優先権基礎出願の開示内容は引用文としてここに組み込まれる。

日本国特許出願2006年第215943号(2006年8月8日出願)

請求の範囲

- [1] 特定種類の画像であるかどうかを判定する画像処理方法であつて、
複数の画素からなる画像を取得し、
特定種類の画像らしさの度合いを画素値および画素位置ごとに示すルックアップテーブルを格納し、
前記取得した画像に基づいて判定用画像を生成し、
前記判定用画像の画素の画素値および画素位置に基づき、前記ルックアップテーブルを用いて、その画素における前記特定種類の画像らしさの度合いを求め、
前記求めた判定用画像の画素の画像らしさの度合いを積算し、
前記積算した結果に基づき、前記入力画像が前記特定種類の画像であるかどうかを判定する画像処理方法。
- [2] 請求項1に記載の画像処理方法において、
前記特定種類の画像は、顔の画像である画像処理方法。
- [3] 請求項1または2に記載の画像処理方法において、
前記判定用画像は、前記取得した画像のエッジ成分を抽出して生成される画像処理方法。
- [4] 請求項1または2に記載の画像処理方法において、
前記判定用画像は、前記取得した画像の局所的に周辺より画素値がへこんでいる凹構造のエッジ成分を抽出して生成される画像処理方法。
- [5] 請求項1に記載の画像処理方法において、
前記ルックアップテーブルは、
前記特定種類の画像の特徴的な要素に対応する画素位置では、その画素のエッジ成分が大きい場合の前記特定種類の画像らしさの度合いを、エッジ成分が小さい場合の前記特定種類の画像らしさの度合いに比べて大きな値とし、
前記特定種類の画像の特徴的な要素以外に対応する画素位置では、その画素のエッジ成分が大きい場合の前記特定種類の画像らしさの度合いを、エッジ成分が小さい場合の前記特定種類の画像らしさの度合いに比べて小さな値とする画像処理方法。

- [6] 請求項2に記載の画像処理方法において、
 前記判定用画像は、前記取得した画像のエッジ成分を抽出して生成され、
 目鼻口のいずれかの領域に対応する画素位置では、その画素のエッジ成分が大きい場合の前記顔の画像らしさの度合いを、エッジ成分が小さい場合の前記顔の画像らしさの度合いに比べて大きな値とし、
 目鼻口以外の領域に対応する画素位置では、その画素のエッジ成分が大きい場合の前記顔の画像らしさの度合いを、エッジ成分が小さい場合の前記顔の画像らしさの度合いに比べて小さな値とする画像処理方法。
- [7] 請求項1から6のいずれかに記載の画像処理方法において、
 前記ルックアップテーブルは、前記特定種類の画像に属する判定対象画像サンプル群と前記特定種類の画像に属さない非判定対象画像サンプル群とに基づく統計処理により生成される画像処理方法。
- [8] 請求項7に記載の画像処理方法において、
 前記統計処理において、
 前記判定用画像を生成するときと等価な処理により、前記判定対象画像サンプル群に基づいて第1の画像サンプル群を生成し、前記非判定対象画像サンプル群に基づいて第2の画像サンプル群を生成し、
 前記第1の画像サンプル群の画素位置 (x,y) における画素値が E となる頻度 $P_1(x,y)(E)$ と、前記第2の画像サンプル群の画素位置 (x,y) における画素値が E となる頻度 $P_2(x,y)(E)$ とを求め、
 前記判定用画像の画素位置 (x,y) における画素値 E に対してその画素における前記特定種類の画像らしさの度合い $V(x,y)$ を、 $V(x,y) = L(x,y)(E)$ で与える画素位置 (x,y) における前記ルックアップテーブル $L(x,y)(E)$ を、 $L(x,y)(E) = f(P_1(x,y)(E), P_2(x,y)(E))$ により生成し、
 前記関数 $f(P_1(x,y)(E), P_2(x,y)(E))$ は、 $P_1(x,y)(E)$ について実質的に広義の単調増加関数であり、 $P_2(x,y)(E)$ について実質的に広義の単調減少関数である画像処理方法。
- [9] 請求項8に記載の画像処理方法において、

前記関数 $f(P_1(x,y)(E), P_2(x,y)(E))$ は、
 $f(P_1(x,y)(E), P_2(x,y)(E)) = \log\{(P_1(x,y)(E) + \varepsilon_1) / (P_2(x,y)(E) + \varepsilon_2)\}$ であり、
 前記 ε_1 と ε_2 は所定の定数である画像処理方法。

- [10] 請求項1に記載の画像処理方法において、
 コントラストの程度に応じた複数のルックアップテーブルを格納し、
 前記取得した画像のコントラストを算出し、
 前記複数のルックアップテーブルから前記コントラストに応じたルックアップテーブルを選択する画像処理方法。
- [11] 特定種類の画像であるかどうかを判定する画像処理方法であって、
 複数の画素からなる画像を取得し、
 特定種類の画像らしさの度合いを画素値および画素位置ごとに示すルックアップテーブルを格納し、
 複数の異なる縮小倍率により前記取得した画像の複数の縮小画像を生成し、
 前記複数の縮小画像に基づいて判定用画像を生成し、
 前記複数の縮小画像の1つである第1の縮小画像に対して判定対象領域を設定し、
 、
 前記判定対象領域の画素の画素値および前記判定対象領域内の画素位置に基づき、前記ルックアップテーブルを用いて、その画素における前記特定種類の画像らしさの度合いを求め、
 前記求めた判定対象領域の画素の前記特定種類の画像らしさの度合いを積算し、
 前記積算した結果に基づき、前記取得した画像内の前記判定対象領域に対応する画像が前記特定種類の画像であるかどうかを判定する画像処理方法。
- [12] 請求項11に記載の画像処理方法において、
 前記第1の縮小画像よりもさらに縮小された第2の縮小画像に対して、前記判定対象領域に対応する第2の判定対象領域をさらに設定し、
 特定種類の画像らしさの度合いを画素値および前記第2の判定対象領域に対応した画素位置ごとに示す第2のルックアップテーブルをさらに格納し、
 前記第2の判定対象領域の画素の画素値および前記第2の判定対象領域内の画

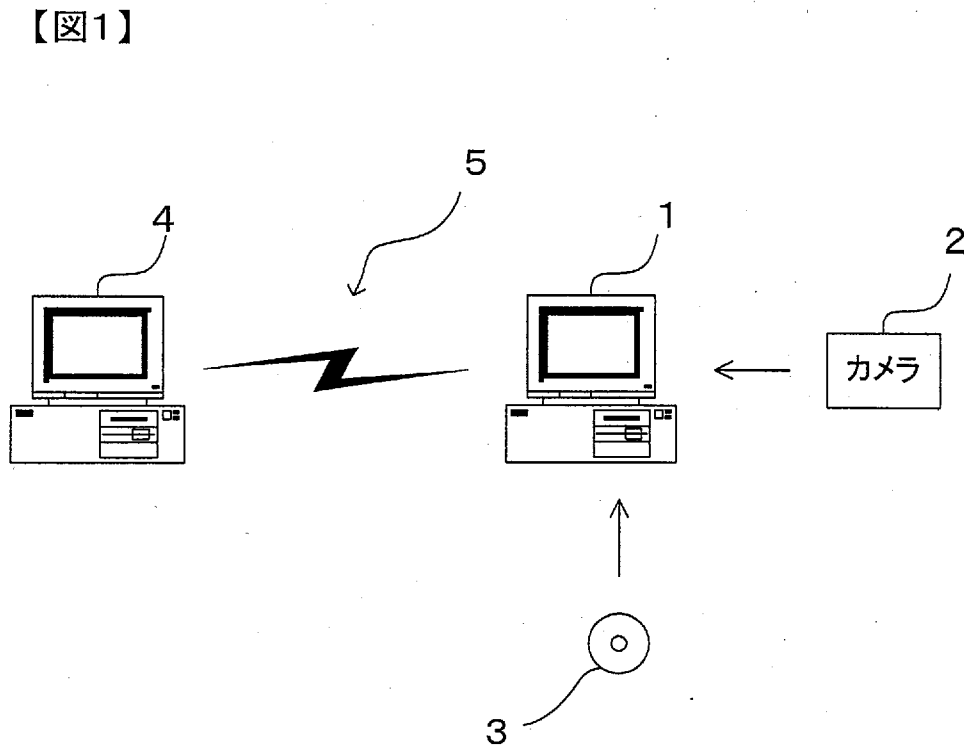
素位置に基づき、前記第2のルックアップテーブルを用いて、その画素における前記特定種類の画像らしさの度合いを求め、

前記求めた第2の判定対象領域の画素の前記特定種類の画像らしさの度合いを積算し、

前記判定対象領域の画素の前記特定種類の画像らしさの度合いの積算結果および前記第2の判定対象領域の画素の前記特定種類の画像らしさの度合いの積算結果に基づき、前記取得した画像内の前記判定対象領域に対応する画像が前記特定種類の画像であるかどうかを判定する画像処理方法。

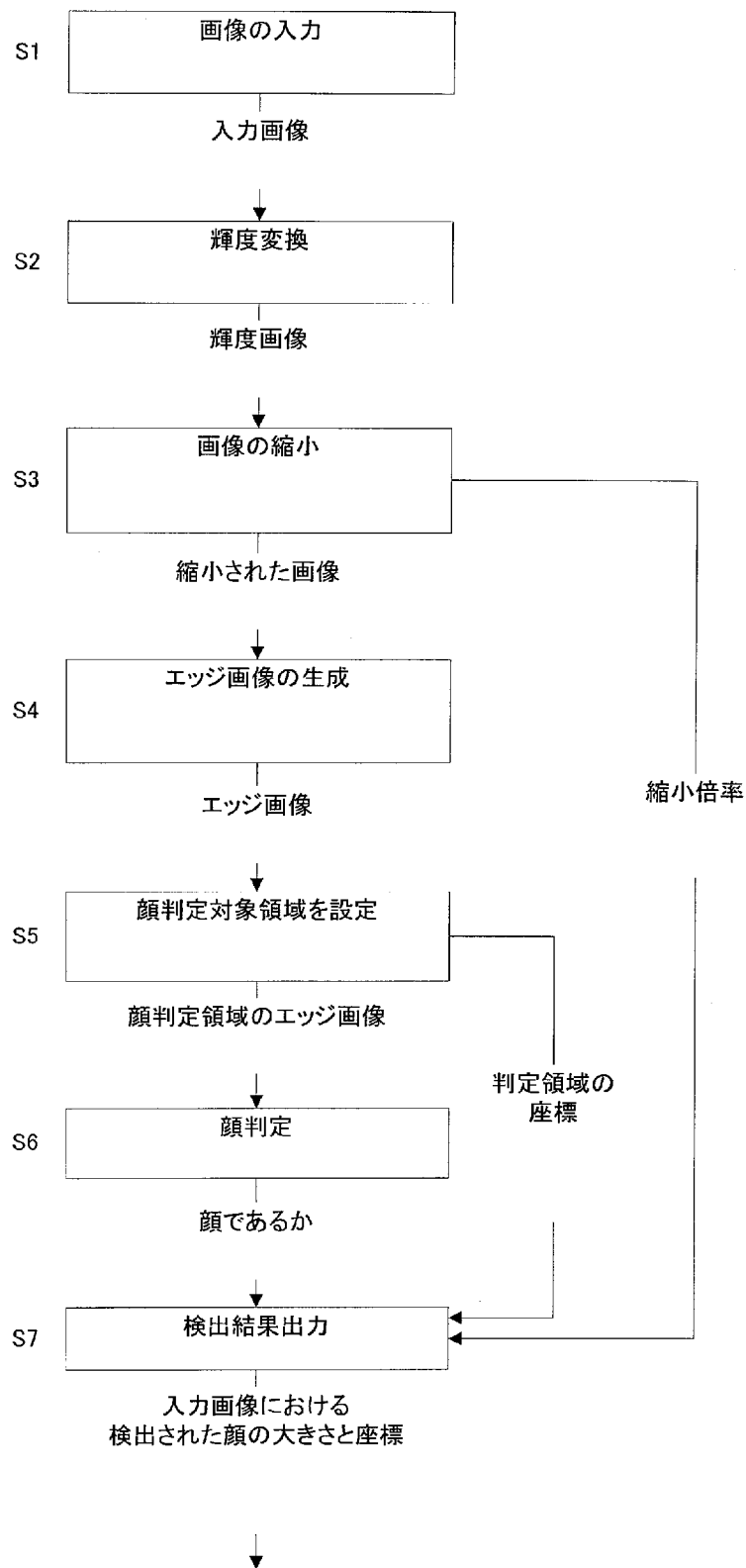
- [13] 請求項1から12のいずれかに記載の画像処理方法をコンピュータに実行させる画像処理プログラム。
- [14] 請求項13に記載の画像処理プログラムを搭載する画像処理装置。
- [15] 請求項13に記載の画像処理プログラムを搭載する撮像装置。

[図1]



[図2]

【図2】



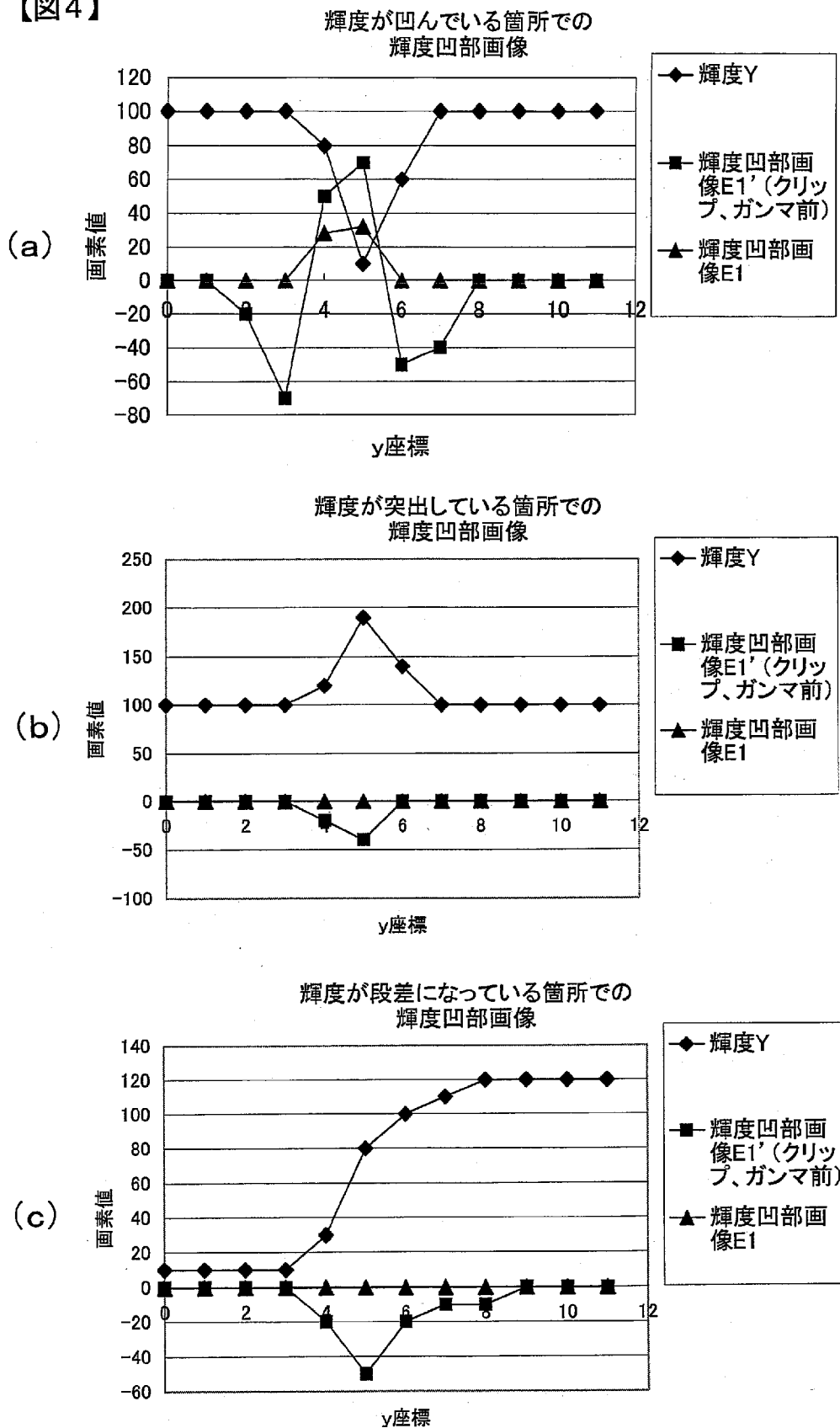
[図3]

【図3】

y+2	x-1,y+2	x,y+2	x+1,y+2	x+2,y+2
y+1	x-1,y+1	x,y+1	x+1,y+1	x+2,y+1
y	x-1,y	x,y	x+1,y	x+2,y
y-1	x-1,y-1	x,y-1	x+1,y-1	x+2,y-1
	x-1	x	x+1	x+2

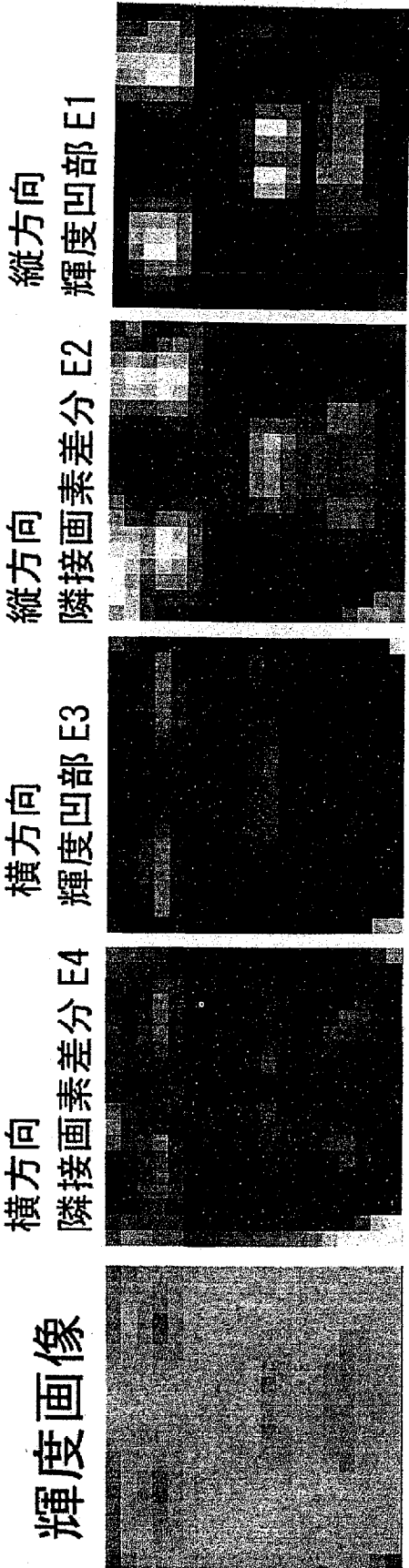
[図4]

【図4】



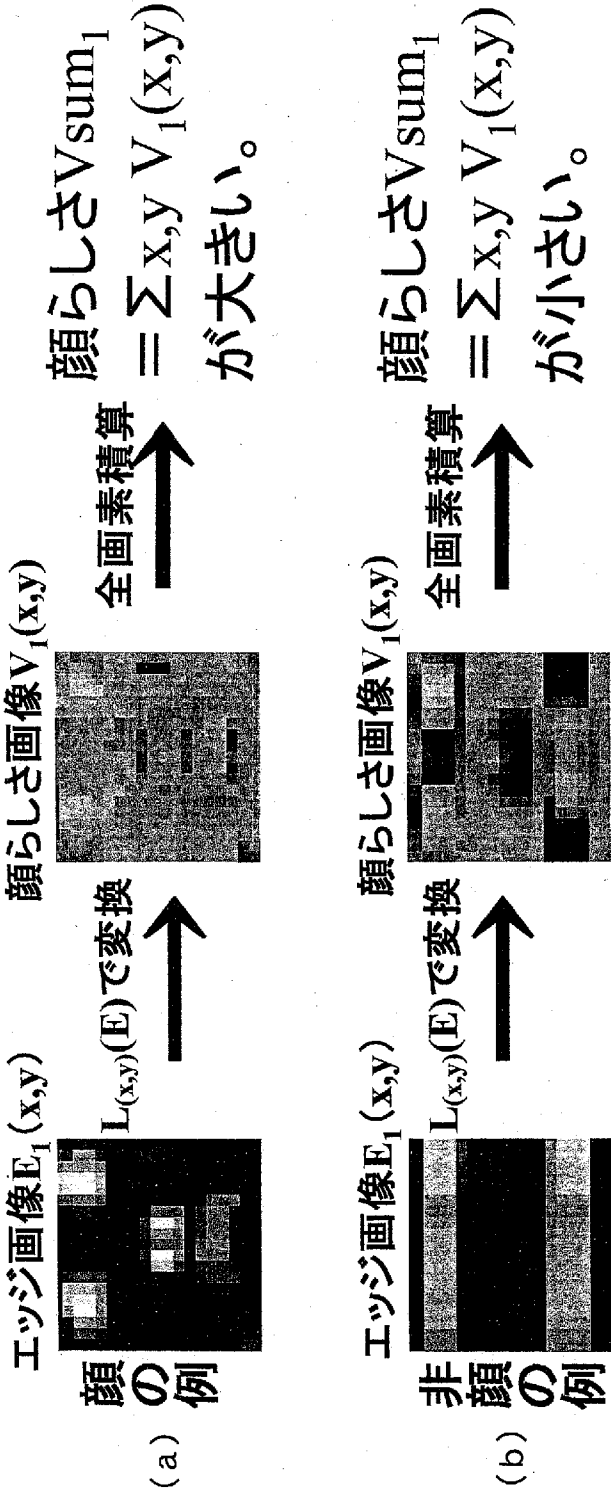
[図5]

【図5】

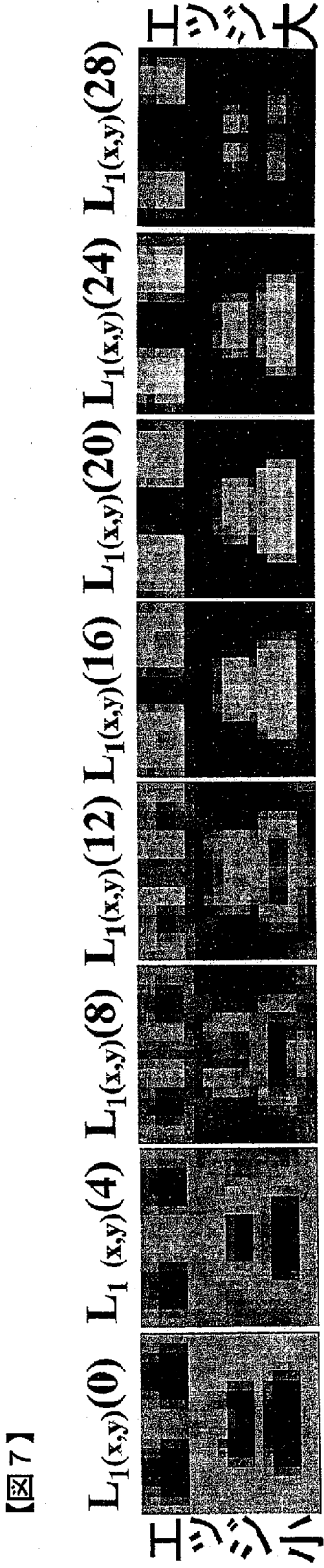


[図6]

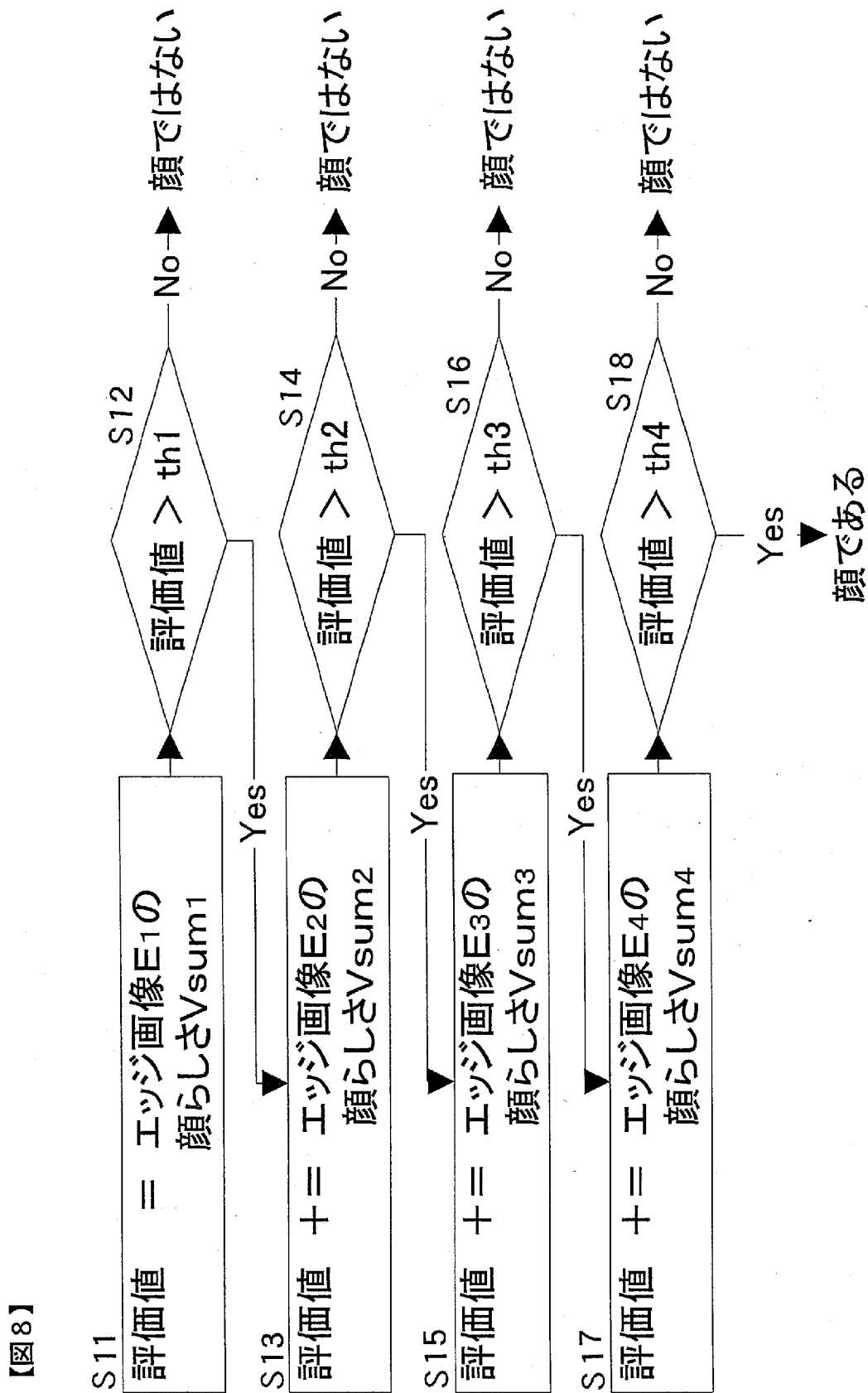
【図6】



[図7]

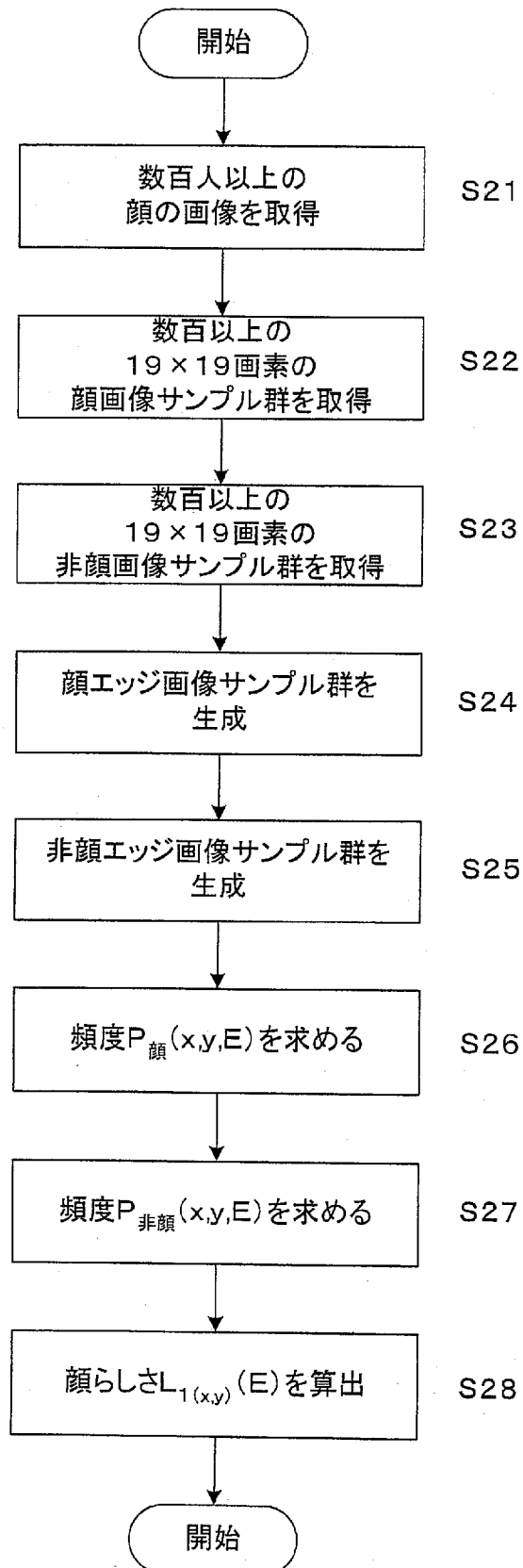


【図8】

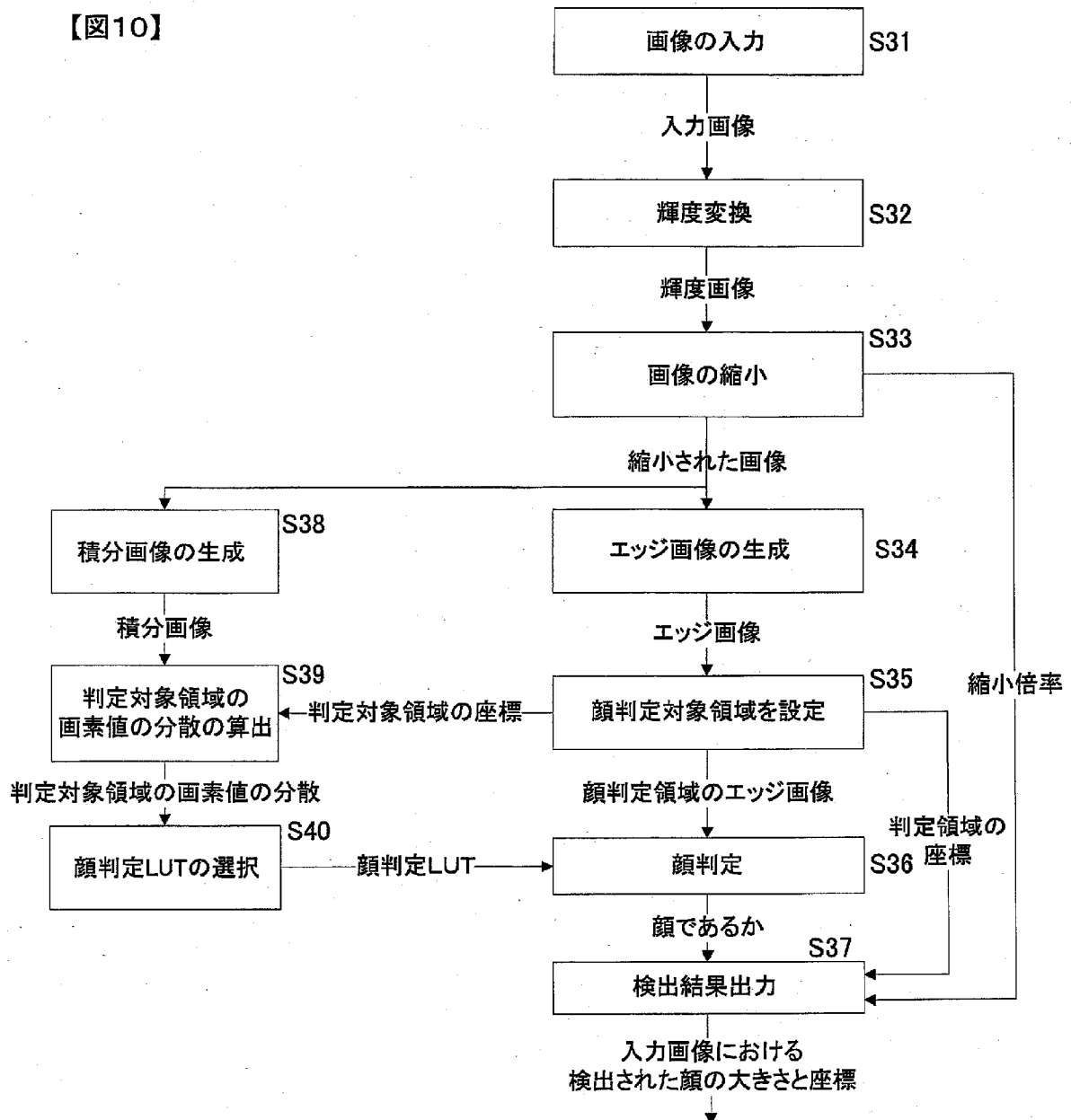


[図9]

【図9】

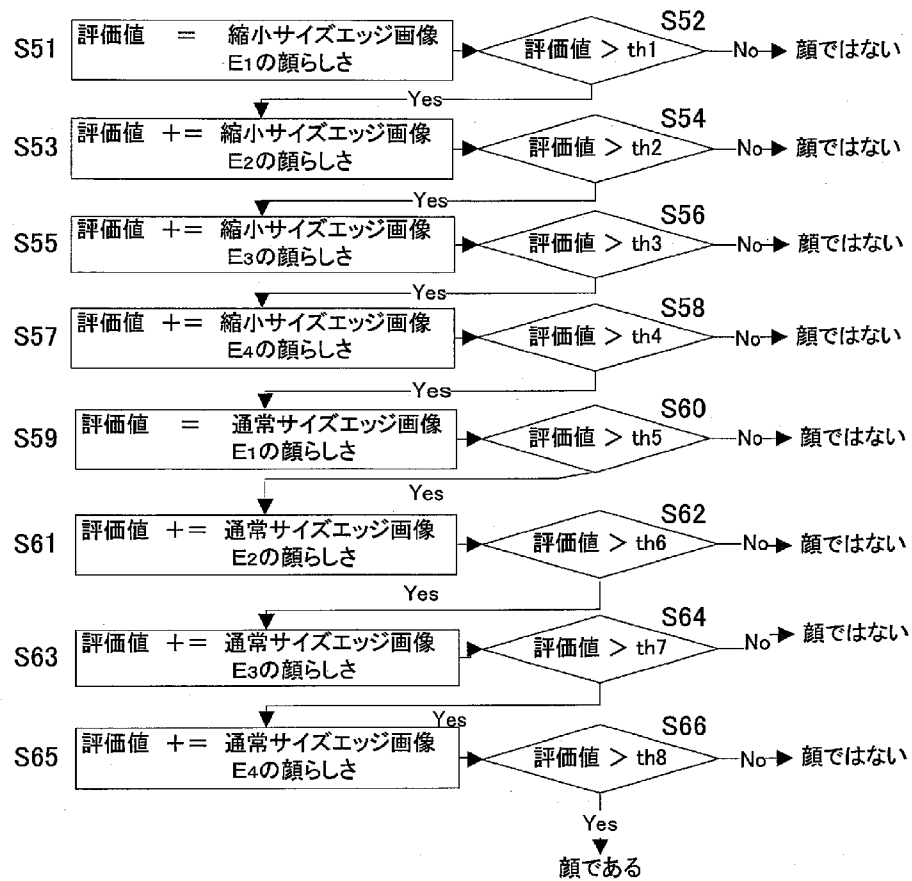


[図10]



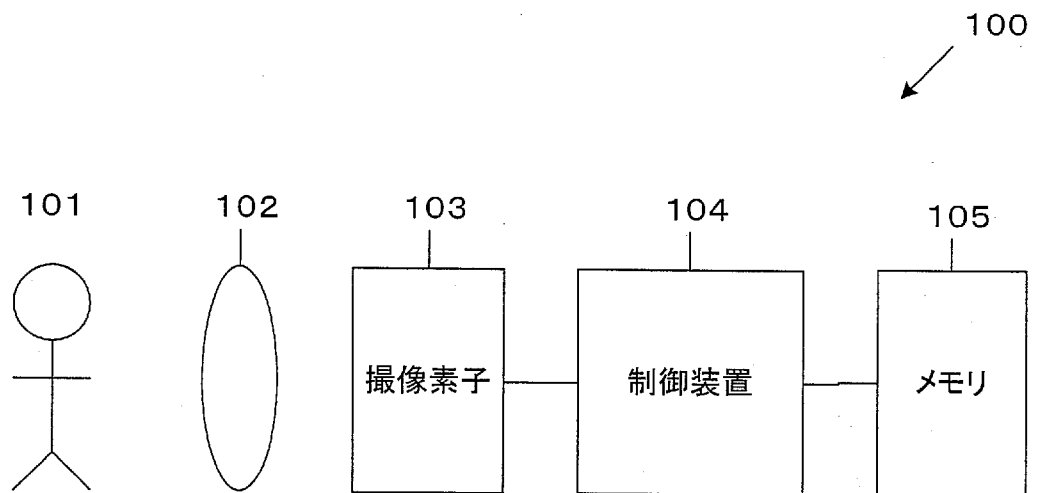
[図11]

【図11】



[図12]

【図12】



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2007/065446

A. CLASSIFICATION OF SUBJECT MATTER

G06T7/60(2006.01)i, H04N5/232(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T7/60, H04N5/232

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2007
Kokai Jitsuyo Shinan Koho	1971-2007	Toroku Jitsuyo Shinan Koho	1994-2007

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 6-348851 A (Sony Corp.), 22 December, 1994 (22.12.94), Par. Nos. [0018] to [0074]; Figs. 1 to 14 (Family: none)	1-15
Y	JP 8-83341 A (Nippon Telegraph And Telephone Corp.), 26 March, 1996 (26.03.96), Par. Nos. [0008] to [0012], [0017]; Fig. 4 (Family: none)	1-15
Y	JP 2006-92095 A (Sony Corp.), 06 April, 2006 (06.04.06), Par. No. [0061] & US 2000/62477 A1	3, 4

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
28 August, 2007 (28.08.07)

Date of mailing of the international search report
04 September, 2007 (04.09.07)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2007/065446

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2001-175869 A (Samsung Electronics Co., Ltd.), 29 June, 2001 (29.06.01), Par. Nos. [0032] to [0036]; Fig. 3 & US 6611613 B1 & NL 1016006 A	5,6
Y	JP 2000-259833 A (Toshiba Corp.), 22 September, 2000 (22.09.00), Par. No. [0033]; Fig. 3 (Family: none)	7-9
Y	JP 2001-192898 A (Fuji Photo Film Co., Ltd.), 11 May, 2001 (11.05.01), Par. Nos. [0108] to [0122]; Figs. 19 to 20 & US 7068328 B1	10

A. 発明の属する分野の分類 (国際特許分類 (I P C))

Int.Cl. G06T7/60(2006.01)i, H04N5/232(2006.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (I P C))

Int.Cl. G06T7/60, H04N5/232

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 0 7 年
日本国実用新案登録公報	1 9 9 6 - 2 0 0 7 年
日本国登録実用新案公報	1 9 9 4 - 2 0 0 7 年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 6-348851 A (ソニー株式会社) 1994. 12. 22, 段落【0018】乃至【0074】, 第 1-14 図 (ファミリーなし)	1-15
Y	JP 8-83341 A (日本電信電話株式会社) 1996. 03. 26, 段落【0008】乃至【0012】, 段落【0017】, 第 4 図 (ファミリーなし)	1-15
Y	JP 2006-92095 A (ソニー株式会社) 2006. 04. 06, 段落【0061】 & US 2000/62477 A1	3, 4

☒ C 欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の 1 以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

2 8 . 0 8 . 2 0 0 7

国際調査報告の発送日

0 4 . 0 9 . 2 0 0 7

国際調査機関の名称及びあて先

日本国特許庁 (I S A / J P)
郵便番号 1 0 0 - 8 9 1 5
東京都千代田区霞が関三丁目 4 番 3 号

特許庁審査官 (権限のある職員)

松尾 俊介

5 H

9 7 4 9

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 3 1

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 2001-175869 A (三星電子株式会社) 2001.06.29, 段落【0032】 乃至【0036】, 第3図 & US 6611613 B1 & NL 1016006 A	5, 6
Y	JP 2000-259833 A (株式会社東芝) 2000.09.22, 段落【0033】, 第3 図 (ファミリーなし)	7-9
Y	JP 2001-192898 A (富士写真フイルム株式会社) 2001.05.11, 段落 【0108】乃至【0122】, 第19-20図 & US 7068328 B1	10