

(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2014-16824
(P2014-16824A)

(43) 公開日 平成26年1月30日(2014.1.30)

(51) Int.Cl.	F 1	テーマコード (参考)
G06T 1/00 (2006.01)	G06T 1/00 340A	5B043
G06T 7/00 (2006.01)	G06T 7/00 510B	5B057

審査請求 未請求 請求項の数 6 O L (全 29 頁)

(21) 出願番号	特願2012-154011 (P2012-154011)	(71) 出願人	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成24年7月9日(2012.7.9)	(74) 代理人	100076428 弁理士 大塚 康德
		(74) 代理人	100112508 弁理士 高柳 司郎
		(74) 代理人	100115071 弁理士 大塚 康弘
		(74) 代理人	100116894 弁理士 木村 秀二
		(74) 代理人	100130409 弁理士 下山 治
		(74) 代理人	100134175 弁理士 永川 行光

最終頁に続く

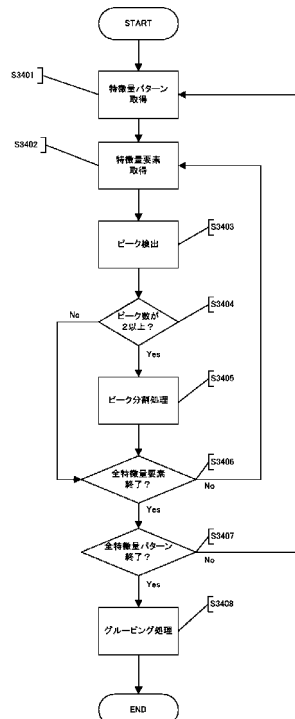
(54) 【発明の名称】 情報処理装置及びその制御方法及びプログラム

(57) 【要約】

【課題】 画像中の顔の個人認識処理において、個人認識精度を向上させた画像処理装置、画像処理方法、及びプログラムを提供する。

【解決手段】 画像データから抽出した被写体の顔領域の複数の部位の特徴量を含む特徴パターンを被写体毎に分類して、顔領域の認識に用いる辞書において管理する辞書管理手段と、被写体の顔領域の各部位の特徴量を、小分類化できるか否かを判定し、小分類化できると判定した場合にはその着目部位の特徴量を複数に小分類化して新たな特徴量とする小分類化手段と、小分類化手段により小分類化された場合は、着目部位の新たな特徴量と、管理手段において管理される着目部位以外の部位の特徴量と、を組み合わせた特徴パターンを、被写体を識別するための新たな特徴パターンとして辞書を更新する更新手段と、を備える。

【選択図】 図 2 8



【特許請求の範囲】

【請求項 1】

画像データから抽出した被写体の顔領域の複数の部位の特徴量を含む特徴パターンを被写体毎に分類して、顔領域の認識に用いる辞書において管理する辞書管理手段と、

前記被写体の顔領域の各部位の特徴量を、小分類化できるか否かを判定し、小分類化できると判定した場合にはその着目部位の特徴量を複数に小分類化して新たな特徴量とする小分類化手段と、

前記小分類化手段により小分類化された場合は、前記着目部位の新たな特徴量と、前記管理手段において管理される前記着目部位以外の部位の特徴量と、を組み合わせた特徴パターンを、前記被写体を識別するための新たな特徴パターンとして前記辞書を更新する更新手段と、

を備えることを特徴とする情報処理装置。

【請求項 2】

前記辞書管理手段において管理される特徴量及び前記新たな特徴量を組み合わせた新たな特徴パターンの複数の画像データにおける出現頻度が所定基準を満たすか否かを判定する判定手段をさらに備え、

前記更新手段は、前記判定手段で所定基準を満たさないと判定された特徴パターンを除外し、前記判定手段で所定基準を満たすと判定された特徴パターンを、前記被写体を識別するための新たな特徴パターンとして前記辞書を更新することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】

前記小分類化手段は、前記辞書において被写体の特徴パターンが予め設定された値を超えた場合に、前記各部位の分散値を求め、当該分散値が所定値以上になった場合に実行されることを特徴とする請求項 1 又は 2 に記載の情報処理装置。

【請求項 4】

各画像データの画像の撮影日時を示す情報を取得する取得手段をさらに備え、

前記小分類化手段は、前記被写体の撮影日時の時間軸に沿って特徴量をさらに分割し、

前記判定手段は、更に前記撮影日時も前記パターンの 1 要素として判定を行う

ことを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載の情報処理装置。

【請求項 5】

画像データから抽出した被写体の顔領域の複数の部位の特徴量を含む特徴パターンを被写体毎に分類して、顔領域の認識に用いる辞書において管理する辞書管理工程と、

前記被写体の顔領域の各部位の特徴量を、小分類化できるか否かを判定し、小分類化できると判定した場合にはその着目部位の特徴量を複数に小分類化して新たな特徴量とする小分類化工程と、

前記小分類化工程において前記着目部位の特徴量が複数に小分類化された場合は、前記着目部位の新たな特徴量と、前記管理工程において管理される前記着目部位以外の部位の特徴量と、を組み合わせた特徴パターンを、前記被写体を識別するための新たな特徴パターンとして前記辞書を更新する更新工程と、

を備えることを特徴とする情報処理装置の制御方法。

【請求項 6】

請求項 1 乃至 4 のいずれか一項に記載の情報処理装置の各手段としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は画像中の人物の顔を識別技術に関するものである。

【背景技術】

【0002】

デジタルスチールカメラ（以下「DSC」とも記載）の普及に伴い、昨今では時に数千枚

10

20

30

40

50

から数万枚といった多くの画像データをハンドリングする必要がある。ハンドリングする方法として、特に、重要な技術の1つとして、人物の顔を手掛かりに画像をハンドリングすることによって人物抽出を実現するものがある。

【0003】

人物抽出の方法は、大きく顔検出技術と個人認識技術に分けられる。前者は、画像内に存在する顔を見つけ出すものであり、顔の存在有無の判断に利用される。後者は、検出された顔に対し、誰の顔かを特定するものである。人物抽出する際には、各個人を識別するため、個人毎の顔の特徴パターンを記憶するためのデータベース（顔辞書）を参照する。そして、顔辞書に登録された画像をもとに、新たに追加された画像の解析処理を実施して人物抽出を行う。

10

【0004】

特許文献1では、人物の認識を実施する際、複数の異なる表情の特徴量と確率分布パラメータを利用することが開示されている。

【0005】

また、特許文献2では、認識する際の変動する照明や顔の向き、表情の変化についての精度低下防止が記述されている。データ属性ごとに用意された複数の辞書データの各々について、入力データと辞書データとの間に対応する局所パターンの相関値を算出する。データ属性ごとの相関値の集合を統合し、入力データの類似度を算出し、それに基づいて識別している。

20

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2005-208850号公報

【特許文献2】特開2011-134114号公報

【特許文献3】特開平06-095685号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

上記の特許文献1および特許文献2においては、個人認識の精度を向上させるために、表情や照明の変化に対応する記載がなされているものの、十分な個人認識の精度とは言えなかった。

30

【0008】

本発明は係る問題点に鑑みなされたものであり、画像中の顔の個人認識処理において、個人認識精度を向上させた画像処理装置、画像処理方法、及びプログラムを提供しようとするものである。

【課題を解決するための手段】

【0009】

この課題を解決するため、例えば本発明の情報処理装置は以下の構成を備える。すなわち、

画像データから抽出した被写体の顔領域の複数の部位の特徴量を含む特徴パターンを被写体毎に分類して、顔領域の認識に用いる辞書において管理する辞書管理手段と、

40

前記被写体の顔領域の各部位の特徴量を、小分類化できるか否かを判定し、小分類化できると判定した場合にはその着目部位の特徴量を複数に小分類化して新たな特徴量とする小分類化手段と、

前記小分類化手段により小分類化された場合は、前記着目部位の新たな特徴量と、前記管理手段において管理される前記着目部位以外の部位の特徴量と、を組み合わせた特徴パターンを、前記被写体を識別するための新たな特徴パターンとして前記辞書を更新する更新手段と、を備える。

【発明の効果】

【0010】

50

本発明によれば、個人認識精度を向上させることができる。例えば、デジタル画像中の顔の個人認識処理において、表情の異なる画像について、十分な精度を実現することができる。

【図面の簡単な説明】

【0011】

【図1】本発明のソフトウェアを実行可能なハードウェア構成を示す図。

【図2】本発明の処理のソフトウェアブロック図。

【図3】画像解析処理のフローチャート。

【図4】画像解析処理のフローチャート。

【図5】人物グループ生成処理のフローチャート。

10

【図6】自動レイアウト提案処理のフローチャート。

【図7】人物グループの表示例を示す図。

【図8】サムネイル形式による画像群の表示例を示す図。

【図9】カレンダー形式による画像群の表示例を示す図。

【図10】画像解析した結果得られる属性情報の例を示す図。

【図11】画像解析結果の保存形式の例を示す図。

【図12】ユーザが手動入力可能な属性情報の例を示す図。

【図13】手動でお気に入り度を入力するためのUI例を示す図。

【図14】手動でイベント情報を入力するためのUI例を示す図。

【図15】手動で人物属性情報を入力するためのUI例を示す図。

20

【図16】レイアウトテンプレートの一例を示す図。

【図17】レイアウトテンプレートの一例を示す図。

【図18】実施形態1における自動レイアウト生成処理のフローチャート。

【図19】実施形態1における不要画像フィルタリング処理のフローチャート。

【図20】自動トリミング処理の一例を示す図。

【図21】自動レイアウトを行う際のレイアウト評価値の例を示す図。

【図22】明るさ適正度の算出方法の説明図。

【図23】彩度適正度の算出方法の説明図。

【図24】トリミング欠け判定処理の説明図。

【図25】画像類似性判定処理の説明図。

30

【図26】自動レイアウト生成結果の表示例を示す図。

【図27】実施形態1における顔辞書作成処理のフローチャート。

【図28】実施形態1における細分化処理のフローチャート。

【図29】実施形態1における特徴量要素の細分化処理の一例を示す図。

【図30】実施形態1における個人認識処理の一例を示す図。

【図31】実施形態1における個人認識および顔辞書更新のフローチャート。

【図32】人間の特徴量の変化のイメージ図。

【図33】実施形態2における特徴量要素の細分化処理の一例を示す図。

【図34】実施形態2における個人認識処理の一例を示す図。

【図35】実施形態2における顔辞書作成処理のフローチャート。

40

【図36】実施形態2における顔辞書削減のフローチャート。

【図37】実施形態1及び2における特徴量パターン削減のフローチャート。

【発明を実施するための形態】

【0012】

以下、添付図面に従って本発明に係る実施形態を詳細に説明する。

【0013】

[実施形態1]

以下では、画像群から顔辞書を生成し、それに基づいた人物（被写体）抽出、レイアウトしたコラージュ画像生成、提案についての実施形態を説明する。これはあくまで実施の1つの形態を例として示したものであり、本発明は以下の実施形態に限定されるものでは

50

ない。

【 0 0 1 4 】

< ハードウェア構成の説明 >

図 1 は、第 1 の実施形態に係わる情報処理装置のハードウェア構成例を説明するブロック図である。図 1 において、情報処理装置 1 1 5 は、例えば、コンピュータである。この情報処理装置 1 1 5 は、画像処理装置として機能する。同図において、100はCPU（中央演算装置）であり、本実施形態で説明する情報処理方法をプログラムに従って実行する。101はROMであり、CPU100により実行されるBIOSプログラムが記憶されている。102はRAMで、CPU100により実行されるOSやアプリケーションを記憶したり、CPU100による各種情報の一時的に記憶のためのワークメモリとしても機能する。103はハードディスク等の2次記憶装置であり、OS、各種アプリケーションを記憶保持機能、並びに、記憶管理対象の画像ファイルを格納する画像記憶機能、更には、画像解析結果を保存するデータベース機能として作用する記憶媒体である。104は表示装置であり、本実施形態の処理結果など以下に示す種々のUI（User Interface）をユーザに提示する装置である。該表示装置 1 0 4 としては、例えば、ディスプレイが挙げられ、タッチパネル機能を備えても良い。また、1 1 0 は制御バス/データバスであり、上述の各部とCPU100とを接続する。このほかにユーザが画像補正の処理の指示等を入力するためのマウスやキーボードといった入力装置 1 0 5 も備える。

10

【 0 0 1 5 】

また、情報処理装置115は、内部撮像デバイス 1 0 6 を備えてもよい。該内部撮像デバイス 1 0 6 で撮像された画像は、所定の画像処理を経た後、2次記憶装置103に保存される。また、画像データはインターフェース（IF108）を介して接続された外部撮像デバイス111から読み込んで構わない。さらに、情報処理装置115は無線LAN（Local Area Network）109を備え、該LANはインターネット113に接続されている。該インターネットに接続された外部サーバ114より画像を取得することもできる。

20

【 0 0 1 6 】

最後に、画像等を出力するためのプリンタ112が、IF107を介して接続されている。なお、該プリンタはさらにインターネット上に接続されており、上記無線LAN109経由でプリントデータのやり取りをすることもできる。

【 0 0 1 7 】

< ソフトウェアのブロック図 >

図 2 は本実施形態における上記アプリケーションを含むソフトウェア構成のブロック図になっている。

30

【 0 0 1 8 】

まずハードウェア115が取得する、デジタルカメラ等で撮像された画像データは、通常JPEG（Joint Photography Expert Group）等の圧縮形式になっている。そのため、画像コーデック部200は、該圧縮形式を解凍していわゆるRGB点順次のビットマップデータ形式に変換する。変換されたビットマップデータは、表示・UI制御部201に伝達され、ディスプレイ等の表示装置104上に表示される。

【 0 0 1 9 】

上記ビットマップデータは、さらに画像センシング部203（アプリケーション）に入力され、同部において、詳細は後述するが、画像の様々な解析処理が行われる。上記解析処理の結果得られた画像の様々な属性情報は、所定の形式に従ってデータベース部202（アプリケーション）において、上述した2次記憶装置103に保存される。なお、以降においては、画像解析処理とセンシング処理は同義で扱う。

40

【 0 0 2 0 】

シナリオ生成部204（アプリケーション）では、ユーザが入力した様々な条件に応じて、詳細は後述するが、自動で生成すべきレイアウトの条件を生成する。レイアウト生成部205では上記シナリオに従って、自動でレイアウトを生成する処理を行う。

【 0 0 2 1 】

50

生成したレイアウトは、レンダリング部206で表示用のビットマップデータを生成し、該ビットマップデータは表示・UI制御部201に送られ、結果がディスプレイに表示される。また、レンダリング結果はさらにプリントデータ生成部207に送られ、同部でプリンタ用コマンドデータに変換され、プリンタに送出される。

【0022】

図3～図6を用いて、本実施形態の基本的な画像処理のフローについて説明する。

【0023】

図3および図4は、画像センシング部203のフローを示しており、複数の画像データ群を取得して、それぞれについて解析処理を施し、その結果をデータベースに格納するまでの処理の流れを示している。図5は、検出した顔位置情報に基づいて、同じ人物と思われる顔情報をグループ化する人物グループ生成処理の流れを示している。図6は、画像の解析情報およびユーザが入力した様々な情報に基づいて、レイアウト作成のためのシナリオを決定し、該シナリオに基づいて、自動でレイアウトを生成するための処理の流れを示している。

10

【0024】

図3を用いて、画像センシング部203が実行する処理のフローについて説明する。まず図3のS301では、画像データ群の取得を行う。画像データ群は、例えば、ユーザが、撮影画像が格納された撮像装置やメモリカードを情報処理装置115に接続し、これらから撮影画像を読み込むことで取得する。また、画像データ群として、内部撮像装置で撮影され、2次記憶装置103に保存されていた画像を取得してもよい。あるいは、画像は無線LANを介して、インターネット上に接続された外部サーバー114等、情報処理装置115以外の装置から取得をしてもよい。

20

【0025】

ここで、図8及び図9を用いて、画像データ群を取得した際の表示装置104の表示について説明する。画像データ群を取得すると、そのサムネイル群が図8や図9に示すようにUIに表示される。表示装置104上のUIの表示方法は、画像を確認できるものであれば、特に限定されるものではない。図8の801に示すように2次記憶装置103内のフォルダ単位で画像のサムネイル802を表示してもよいし、図9に示すようにカレンダーのようなUI901で日付ごとに画像データが管理されていてもよい。図9のUI901では、日付の部分902をクリックすれば、同日に撮影された画像が、図8のようなサムネイル一覧で表示される。

30

【0026】

次に、S302～305において、取得した画像データ群それぞれについて、解析処理および解析結果のデータベース登録を行う。

【0027】

具体的には、S302において、各画像のデコードを行う。まず、アプリケーションは、新規で保存され未だセンシング処理が行われていない画像をサーチし、抽出された各画像について、コーデック部が圧縮データからビットマップデータに変換する。

【0028】

次に、S303において、上記ビットマップデータに対して、各種センシング処理を実行させる。ここでいうセンシング処理には、図10に示されるような様々な処理を想定する。本実施形態では、センシング処理の例として、顔検出および顔領域の特徴量解析、画像の特徴量解析、シーン解析を挙げており、それぞれ同図に示すようなデータ型の結果を算出する。本実施形態では、画像基本特徴量としては、平均輝度(int:値0～255)、平均彩度(int:0～255)、平均色相(int:値0～359)を解析した。また、顔検出としては、人物顔の個数(int:値0以上(0～MAXFACE))、人物顔の位置情報である座標位置(int*8:値0以上(Width及びHeightそれぞれ同様))を解析した。さらに、顔領域内の平均Y(int:値0～255)、顔領域内の平均Cb(int:値-128～127)、顔領域内の平均Cr(int:値-128～127)を解析した。以下、それぞれのセンシング処理について説明する。

40

50

【 0 0 2 9 】

まず画像の基本的な特徴量である全体の平均輝度、平均彩度については、公知の方法で求めればよいため、簡単に説明する。平均輝度は、画像の各画素について、RGB成分を公知の輝度色差成分（例えばYCbCr成分）に変換し（変換式省略）、Y成分の平均値を求めればよい。また、平均彩度は、上記CbCr成分について画素毎に次式Sを算出し、その平均値を求めればよい。

$$S = (C b^2 + C r^2)^{1/2}$$

【 0 0 3 0 】

また、画像内の平均色相（AveH）は、画像の色合いを評価するための特徴量である。各画素毎の色相は、公知のHIS変換式を用いて求めることができ、それらを画像全体で平均化することにより、AveHを求めることができる。

10

【 0 0 3 1 】

また、上記特徴量は、画像全体で算出してもよいし、例えば、画像を所定サイズの領域に分割し、各領域毎に算出してもよい。

【 0 0 3 2 】

次に、人物の顔検出処理について説明する。本実施形態で使用する人物の顔検出手法としては、すでに様々な手法を用いることができる。特開2002-183731では、入力画像から目領域を検出し、目領域周辺を顔候補領域とする。

【 0 0 3 3 】

そして、該顔候補領域に対して、画素毎の輝度勾配、および輝度勾配の重みを算出し、これらの値を、あらかじめ設定されている理想的な顔基準画像の勾配、および勾配の重みと比較する。その時に、各勾配間の平均角度が所定の閾値以下であった場合、入力画像は顔領域を有すると判定する方法が記載されている。

20

【 0 0 3 4 】

また、特開2003-30667によれば、まず画像中から肌色領域を検出し、同領域内において、人間の虹彩色画素を検出することにより、目の位置を検出することが可能である。

【 0 0 3 5 】

特開平8-63597によれば、複数の顔の形状をしたテンプレートと画像とのマッチング度を計算する。該マッチング度が最も高いテンプレートを選択し、最も高かったマッチング度があらかじめ定められた閾値以上であれば、選択されたテンプレート内の領域を顔候補領域とする。同テンプレートを用いることにより、目の位置を検出することが可能である。

30

【 0 0 3 6 】

特開2000-105829によれば、まず、鼻画像パターンをテンプレートとし、画像全体、あるいは画像中の指定された領域を走査し最もマッチする位置を鼻の位置として出力する。次に、画像の鼻の位置よりも上の領域を目が存在する領域と考え、目画像パターンをテンプレートとして目存在領域を走査してマッチングをとり、ある閾値よりもマッチ度が度置きの画素の集合である目存在候補位置集合を求める。さらに、目存在候補位置集合に含まれる連続した領域をクラスタとして分割し、各クラスタと鼻位置との距離を算出する。その距離が最も短くなるクラスタを目が存在するクラスタと決定することで、顔の部位位置の検出が可能である。

40

【 0 0 3 7 】

その他、顔および部位位置を検出する方法としては、特開平8-77334、特開2001-216515、特開平5-197793、特開平11-53525、特開2000-132688、特開2000-235648、特開平11-250267などが挙げられる。さらには特登録2541688など、数多くの手法が挙げられる。本実施形態では特に手法を限定はしない。

【 0 0 3 8 】

上記処理の結果、各入力画像について、人物顔の個数と各顔毎の座標位置を取得することができる。

【 0 0 3 9 】

また、画像中の顔座標位置が分かれば、該顔領域毎に顔領域内に含まれる画素値の平均

50

YCbCr値を求めることにより、該顔領域の平均輝度および平均色差を得ることができる。

【0040】

また、画像の特徴量を用いてシーン解析処理を行うことができる。シーン解析処理については、様々な手法を用いることができ、例えば、出願人が開示している特開2010-251999や特開2010-273144等で開示されている技術を利用することができる。上記シーン解析の結果、風景(Landscape)、夜景(Nightscape)、人物(Portrait)、露出不足(Underexposure)、その他(Others)、という撮影シーンを区別するためのIDを取得することができる。

【0041】

なお、センシング情報は、上記のセンシング処理により取得されるものに限定されるものではなく、例えば、その他のセンシング情報を利用してもよい。

10

【0042】

上記のようにして取得したセンシング情報を、データベース202に保存する。データベースへの保存形式は、特に限定されないが、例えば、図11に示すような汎用的なフォーマット(例えば、XML: eXtensible Markup Language)で記述し、格納すればよい。

【0043】

同図においては、各画像毎の属性情報を、3つのカテゴリに分けて記述する例を示している。1番目のBaseInfoタグは、画像サイズや撮影時情報として、あらかじめ取得した画像ファイルに付加されている情報である。ここには、画像毎の識別子IDや、画像ファイルが格納されている保存場所、画像サイズ、撮影日時などが含まれる。

20

【0044】

次に2番目のSensInfoタグは、上述した画像解析処理の結果を格納するためのタグである。画像全体の平均輝度、平均彩度、平均色相やシーン解析結果が格納され、さらに、画像中に存在する人物の顔位置や顔色に関する情報が記述可能である。

【0045】

次に3番目のUserInfoタグは、ユーザが画像毎に入力した情報を格納することができるが、詳細については後述する。

【0046】

なお、画像属性情報のデータベース格納方法については、上記に限定されるものではない。その他どのような形式であっても構わない。

30

【0047】

<個人認識処理を用いた人物グルーピング>

次に、図3のS306において、S303で検出された顔位置情報を用いて、人物毎のグループを生成する処理を行う。あらかじめ人物の顔を自動でグループ化しておくことにより、その後ユーザが各人物に対して名前を付けていく作業を効率化することが可能となる。

【0048】

ここでの人物グループ生成処理は、公知の個人認識技術を用いて、図5の処理フローにより実行する。

【0049】

図5は人物グループ生成処理S306の基本的なフローチャートである。

40

【0050】

まず、S501で、2次記憶装置に保存されている画像を順次読みだしてデコード処理を行う。さらにS502でデータベースにアクセスし、該画像中に含まれる顔の個数と顔の位置情報を取得する。次に、S504において、個人認識処理を行うための正規化顔画像を生成する。

【0051】

ここで正規化顔画像とは、画像内に様々な大きさ、向き、解像度で存在する顔を切り出して、すべて所定の大きさと向きになるよう、変換して切り出した顔画像のことである。個人認識を行うためには、眼や口といった部位の位置が重要となるため、正規化顔画像のサイズは、上記部位が確実に認識できる程度であることが望ましい。このように正規化顔

50

画像を用意することにより、特徴量検出処理において、様々な解像度の顔に対応する必要がなくなるという利点がある。

【0052】

次に、S512で被写体毎の顔を分類するための顔辞書作成処理を実施する。対象となる画像から顔位置取得S504、正規化生成S504された顔画像をもとに、顔辞書を作成する。本処理について、詳細は後述する。

【0053】

S510において、処理対象の画像データ中に、他の顔領域があるか否かを判定する。他の顔領域がある場合(S510でYES)、S502に戻る。一方、他の顔領域がない場合(S510でNO)、S511に進む。S511において、全ての画像についてS502～509の処理が終了したか否かを判定する。全ての画像について処理が終了した場合は、処理を終了し、全ての画像について処理が終了していない場合は、S502へ戻る。つまり、S510およびS511の判定処理により全ての顔および画像により、顔辞書が生成された後、S513により、顔辞書の細分化グループ生成が実施される。これは対象の画像をもとに生成された顔辞書を解析し、さらに細分化されたグループを生成するものであり、本処理についても詳細は後述する。

10

【0054】

グループ化の結果は、XMLなどのフォーマットで、上述したデータベースに保存しておく。

【0055】

なお、上記実施形態においては、図3に示したように、全ての画像のセンシング処理が終了した後に人物グループ生成処理を実行したが、これに限定されるものではない。例えば、図4に示すように、各画像に対してS403でセンシング処理を実行した後に、顔検出位置情報を利用してグループ化処理S405を行い、これを繰り返して全ての画像データの処理を行うようにしたとしても、同様の結果を生成することができる。

20

【0056】

また、上記の処理によって得られた各人物グループは、図7のようなUI701にて表示されることになる。同図において、702は人物グループの代表顔画像を表しており、その横には、該人物グループの名前を表示する領域703が存在する。自動の人物グループ化処理を終了した直後は、同図に示すように人物名は「No name」と表示される。また、704は該人物グループに含まれる複数の顔画像が表示される。後述するが、同図のUIにおいては、「No name」の領域703を指定して人物名を入力したり、人物毎に誕生日や続柄等の情報を入力することができる。

30

【0057】

また、上記センシング処理は、オペレーティングシステムのバックグラウンドタスクを利用して実行しても良い。この場合、ユーザはコンピュータ上で別の作業を行っていたとしても、画像群のセンシング処理を継続させることができる。

【0058】

<ユーザ情報の入力(人物名、誕生日、お気に入り度、など)>

本実施形態においては、ユーザが手動で画像に関する様々な属性情報を入力してもよく、自動で画像に関する様々な属性情報を設定するようにしてもよい。

40

【0059】

その属性情報の例の一覧を、図12に示す。手動登録情報は大きく分けると、画像毎に設定する画像の属性情報と、上記処理によりグループ処理した人物毎に設定する人物の属性情報に分かれる。

【0060】

まず、画像毎に設定する属性情報として、ユーザのお気に入り度がある。お気に入り度は、その画像を気に入っているかどうかを、例えば、ユーザが手動で入力することができる。例えば図13示すように、UI1301上で、所望のサムネイル画像1302をマウスポインタ1303で選択し、右クリックをすることでお気に入り度を入力できるダイアログを表示する

50

。ユーザはメニューの中で自分の好みに応じて、 の数を選択することができる。本実施形態では、お気に入り度が高いほど の数が多くなるよう設定した。

【 0 0 6 1 】

また、上記お気に入り度は、ユーザが手動で設定せずに、自動で設定するようにしてもよい。例えば、ユーザが図 8 に示す画像サムネイル一覧表示の状態から、所望の画像ファイルをクリックし、1 画像表示画面に遷移したとする。その遷移した回数を計測して、回数に応じてお気に入り度を自動設定してもよい。すなわち、閲覧した回数が多いほど、ユーザが該画像を気に入っていると判断する。

【 0 0 6 2 】

また他の例として、プリント回数に応じてお気に入り度を自動設定してもよい。具体的には、例えば、プリント行為を行った場合、当然その画像を気に入っていると判断し、プリント回数が多いほど、よりお気に入り度が高いと判断する。

10

【 0 0 6 3 】

以上で説明したように、お気に入り度については、ユーザが手動で設定する方法、閲覧回数やプリント回数に応じてお気に入り度を自動設定する方法など様々考えられる。上記設定および計測した情報については、それぞれ個別に、図 1 1 で示すようなXMLフォーマットで、データベース202のUserInfoタグ内に格納される。例えば、お気に入り度はFavoriteRateタグで、閲覧回数はViewingTimesタグで、プリント回数はPrintingTimesタグに格納される。

【 0 0 6 4 】

また、画像毎に設定する別の属性情報として、イベント情報が考えられる。イベント情報としては、例えば、家族旅行 “ travel ”、卒業式 “ graduation ”、結婚式 “ wedding ” が挙げられる。 イベントの指定は、例えば、図 1 4 で示すように、カレンダー上で所望の日付をマウスポインタ1402などで指定して、その日のイベント名を入力することで設定できるようにしてもよい。指定されたイベント名は、画像の属性情報の一部として、図 1 1 に示すXMLフォーマットに含まれることになる。同図のフォーマットでは、UserInfoタグ内のEventタグを使って、イベント名と画像を関連付けている（紐付けている）。

20

【 0 0 6 5 】

次に、人物の属性情報について説明する。

【 0 0 6 6 】

図 1 5 は、人物の属性情報を入力するためのUIを示している。同図において、1502は所定人物（この場合は “ father ”）の代表顔画像を示している。また、1504には、他の画像の中から検出し、S506で顔特徴量が類似していると判断された画像の一覧が表示される。

30

【 0 0 6 7 】

センシング処理が終了した直後は、図 7 に示したように各人物グループには名前が入力されていないが、「No name」の部分703をマウスポインタで指示することにより、任意の人物名を入力することができる。

【 0 0 6 8 】

また、人物毎の属性として、それぞれの人物の誕生日やアプリを操作しているユーザから見た続柄を設定することもできる。本実施形態では、図 1 5 の人物の代表顔1502をクリックすると、画面下部に図示するように、クリックした人物の誕生日を第 1 入力部1505で、続柄情報を第 2 入力部1506で入力することができるようにした。

40

【 0 0 6 9 】

以上、入力された人物の属性情報は、これまでの画像に紐付けられた画像の属性情報とは異なり、画像属性情報とは別にデータベース202内で管理される。

【 0 0 7 0 】

本実施形態においては、あらかじめ様々なレイアウトテンプレートが用意されている。レイアウトテンプレートとは図 1 6 および図 1 7 に示すようなもので、レイアウトする用紙サイズ上に、複数の画像配置枠1702,1902（以降、スロットと同義）で構成されている。

50

【 0 0 7 1 】

このようなテンプレートは多数用意されており、あらかじめ本実施形態を実行するためのソフトウェアが情報処理装置115にインストールされた時点で、2次記憶装置103に保存しておけばよい。また、その他の方法として、IF 107や無線LAN109を介して接続されたインターネット上に存在するサーバ114から、任意のテンプレートを取得してもよい。

【 0 0 7 2 】

これらのテンプレートは汎用性の高いページ記述言語、例えば上述したセンシング結果の格納と同様にXMLなどで記載されているものとする。本実施形態においては、このようなテンプレートを多数保持する。

【 0 0 7 3 】

上記のように、本実施形態で提示したアプリケーションは、入力された画像群に対して解析処理を実行し、人物を自動的にグループ化してUIで表示することができる。また、ユーザはその結果を見て、人物グループ毎に名前や誕生日などの属性情報を入力したり、画像毎にお気に入り度などを設定することができる。

【 0 0 7 4 】

さらにはテーマごとに分類された多数のレイアウトテンプレートを保持することができる。

【 0 0 7 5 】

以上の条件を満たすと、本実施形態のアプリケーションは、ある所定のタイミングで、自動的にユーザに好まれそうなコラージュ画像を生成し、ユーザに提示する処理を行う。これを、レイアウトの提案処理と呼ぶこととする。

【 0 0 7 6 】

図6は、レイアウトの提案処理を行うための基本的なフローチャートを示している。

【 0 0 7 7 】

同図において、まずS601において、提案処理のシナリオを決定する。シナリオには、提案するレイアウトのテーマおよびテンプレートの決定、レイアウト内で重視する人物（主人公）の設定、レイアウト生成に用いる画像群の選定情報が含まれる。

【 0 0 7 8 】

以下では簡単のため、2つのシナリオについて例を記載する。

【 0 0 7 9 】

例えば、各人物の誕生日の2週間前になった場合、自動的にコラージュレイアウトを生成し提示するように予め設定されているとする。そして、図15で自動グループ化されている人物“son”の1歳の誕生日が2週間後となったとする。この場合には、提案するレイアウトのテーマは成長記録“growth”と決定する。次にテンプレートの選択を行うが、この場合には成長記録に適した図17のようなものを選択する。次に、レイアウトに利用するための画像群を選定する。この例の場合には、データベースを参照し、上記人物“son”の誕生日からこれまでに撮影した画像群のうち、“son”を含む画像群を大量に抽出してリスト化する。以上が、成長記録レイアウトのためのシナリオ決定である。

【 0 0 8 0 】

上記とは異なる例として、1カ月以内に撮影された旅行の写真がある場合、自動的にコラージュレイアウトを生成して提示するように予め設定されているとする。図14で登録したイベント情報から、例えば数日前に家族旅行に行きその画像が大量に2次記憶装置に保存されていることがわかると、シナリオ決定部は、家族旅行のレイアウトを提案するためのシナリオを決定する。この場合には、提案するレイアウトのテーマは旅行“travel”と決定し、図16のようなレイアウトを選択する。次に、レイアウトに利用するための画像群を選定する。この例の場合には、データベースS602を参照し、上記旅行イベントに紐付けられた画像群を大量に抽出してリスト化する。以上が、家族旅行レイアウトのためのシナリオ決定である。

【 0 0 8 1 】

次に、図6のS603において、上述したシナリオに基づくレイアウトの自動生成処理を実

10

20

30

40

50

行する。ここで、図18を用いて、シナリオに基づくレイアウトの自動生成処理について説明する。図18はレイアウト処理部の詳細な処理フローを示している。

【0082】

同図では、まずS2101で、上述したシナリオ生成処理で決定され、レイアウトテーマと人物グループ情報が設定された後のレイアウトテンプレート情報を、より取得する。

【0083】

次に、S2103においては、上記シナリオで決定した画像群リストに基づいて、各画像毎に該画像の特徴量をデータベースから取得し、画像群属性情報リストを生成する。ここでいう画像群情報リストとは、図11に示したIMAGEINFOタグが画像リスト分だけ並んだ構成となっている。以降ではこの画像属性情報リストに基づいて、S2105～S2109における自動レイアウト生成処理を行う。

10

【0084】

このように、本実施形態の自動レイアウト生成処理では、このように画像データそのものを直接扱うのではなく、あらかじめ画像毎にセンシング処理を行ってデータベース保存しておいた属性情報を利用する。この理由は、レイアウト生成処理を行う際に、画像データそのものを対象としてしまうと、画像群を記憶するために非常に巨大なメモリ領域を必要としてしまうためである。すなわち、本実施形態のようにデータベース部に記憶しておいた属性情報を利用することにより、メモリ領域の使用量を低減することができる。

【0085】

次に、S2105において、入力された画像群の属性情報を用いて、入力された画像群の中から不要画像のフィルタリングを行う。フィルタリング処理は、図19のフローにて行う。同図では、各画像毎に、まずS2201で全体の平均輝度がある閾値(ThY_LowとThY_High)内に含まれているかの判定を行う。否の場合にはS2206に進み、注目画像はレイアウト対象から除去する。具体的には、S2202では、ID=Nである顔領域のAveYが所定閾値(ThfY_LowとThfY_High)の範囲に含まれているか否かの判定を行う。S2203では、ID=Nである顔領域のAveChが所定閾値(ThfY_LowとThfY_High)の範囲に含まれているか否かの判定を行う。S2204では、ID=Nである顔領域のAveCrが所定閾値(ThfY_LowとThfY_High)の範囲に含まれているか否かの判定を行う。S2205では、最後の顔であるか否かを判定する。最後の顔ではない場合は、S2202へ戻り、最後の顔である場合は、処理を終了する。

20

30

【0086】

同様に、S2202～S2205では、注目画像に含まれる顔領域それぞれについて、平均輝度、平均色差成分が、良好な肌色領域を示す所定閾値に含まれているかの判定を行う。S2202～S2205のすべての判定がYesとなる画像のみ、以降のレイアウト生成処理に適用される。

【0087】

なお、このフィルタリング処理では、以降の一時レイアウト作成処理に明らかに不要と判断できる画像の除去を目的としているため、上記閾値は比較的湯緩やかに設定することが望ましい。例えばS2201の画像全体輝度の判定において、ThY_HighとThY_Lowの差が画像ダイナミックレンジに比して極端に小さいと、それだけYesと判定される画像が少なくなってしまう。したがって、本実施形態のフィルタリング処理では、両者の差をできる限り広く設定し、かつ明らかに異常画像と判断されるものは除去できるような閾値を設定する。

40

【0088】

次に図18のS2107において、上記処理でレイアウト対象となった画像群を用いて、大量(L個)の一時レイアウトを生成する。一時レイアウトの生成は、取得したテンプレートの画像配置枠に対して、入力画像を任意に当てはめる処理を繰り返すことにより実行する。この時に、以下のパラメータ(画像選択・配置・トリミング基準)をランダムで決定する。画像選択基準としては、レイアウト中の画像配置枠がN個の時、画像群の中からどの画像を選択するかが挙げられ、配置としては、選択した複数の画像を、どの配置枠に配

50

置するかが挙げられ、トリミング基準としては、・配置した際に、どの程度のトリミング処理を行うかというトリミング率が挙げられる。ここで、トリミング率は例えば0~100%で表わされ、画像の中心を基準として、図20のように所定の領域がトリミングされる。同図において、2301は画像全体を示し、2302はトリミング率50%でトリミングした際の切り取り枠を示している。

【0089】

上記の画像選択・配置・トリミング基準に基づいて、可能な限り数多くの一時レイアウトを生成する。生成した各一時レイアウトについても、XMLのように形式で表わすことができる。

【0090】

なお、ここで生成する一時レイアウトの数Lについては、後述するレイアウト評価ステップでの評価処理の処理量と、それを処理する情報処理装置115の性能に応じて決定されるが、例えば数十万通り以上の一時レイアウトを用意する。

【0091】

次に、図18のS2108において、上記で作成したL個の一時レイアウトに対して、それぞれ所定のレイアウト評価量を用いて評価を行う。本実施形態におけるレイアウト評価量の一覧を、図21に示す。同図に示すように、本実施形態で用いるレイアウト評価量は、主に3つのカテゴリに分けることができる。

【0092】

一つ目の評価カテゴリは、画像個別の評価量である。これは画像の明るさや彩度、ブレぼけ量等の状態を判断し、スコア化するものである。スコア化の一例を如何に示す。まず明るさの適正度については、図22に示すように、平均輝度がある所定レンジ範囲内においてはスコア値100とし、所定レンジ範囲外となるとスコア値が下げるよう設定している。また、彩度の適正度については、図23に示すように、画像全体の平均彩度がある所定の彩度値より大きい場合にはスコア値100とし、所定値より小さい場合にはスコア値を除所に下げていくよう設定している。

【0093】

二つ目の評価カテゴリは、画像とスロットの適合度の評価である。画像・スロットの適合度の評価としては、例えば、人物適合度、トリミング欠け判定が挙げられる。人物適合度とは、スロットに指定されている人物と、実際に該スロットに配置された画像内に存在する人物の適合率を表したものである。例を挙げると、あるスロットが、“father”、“son”が指定されているものとする。この時、該スロットに割り当てられた画像に該2人の人物が写っていたとすると、該スロットの人物適合度はスコア値100とする。もし片方の人物しか写っていなかったとすると、適合度はスコア値50とし、両者とも写っていなかった場合は、スコア値0とする。ページ内の適合度は、各スロット毎に算出した適合度の平均値とする。

【0094】

また別の画像・スロット適合度評価値として、トリミング領域2702の欠け判定が存在する。例えば、図24に示すように、画像中に存在する顔の位置2703が判明している場合、欠けた部分の面積に応じて、0から100までのスコア値を算出する。欠けた面積が0の場合、スコアは100とし、逆にすべての顔領域が欠けた場合、スコア値は0とする。

【0095】

三つ目の評価カテゴリは、レイアウトページ内のバランスの評価である。図21ではバランスを評価するための評価値として画像類似性、色合いばらつき、顔サイズばらつきを挙げている。

【0096】

まず、画像類似性について説明する。画像の類似性は、大量に生成した一時レイアウト毎に、それぞれの画像の類似性を算出する。例えば旅行テーマのレイアウトを作成したい場合、あまりに似通った類似度の高い画像ばかりが並んでいたとすると、それは良いレイアウトとは言えないことがある。例えば、類似性は、撮影日時によって評価することがで

10

20

30

40

50

きる。撮影日時が近い画像は、同じような場所で撮影された可能性が高いが、撮影日時が離れていれば、その分、場所もシーンも異なる可能性が高いからである。撮影日時は、図 11 で示したように、画像属性情報として、予めデータベース202に保存されている、画像毎の属性情報から取得することができる。撮影日時から類似度を求めるには以下のような計算を行う。例えば、今注目している一時レイアウトに図 25 で示すような4つの画像がレイアウトされているものとする。なお、図 25 において、画像IDで特定される画像には、それぞれ撮影日時情報が付加されている。具体的には、撮影日時として、年月日及び時間（西暦：YYYY、月：MM、日：DD、時：HH、分：MM、秒：SS）が付加されている。この時、この4つの画像間で、撮影時間間隔が最も短くなる値を算出する。この場合は、画像ID102と108間の30分が最も短い間隔である。この間隔をMinIntervalとし、秒単位で格納する。すなわち30分 = 1800秒である。このMinIntervalをL個の各一時レイアウト毎に算出して配列stMinInterval[l]に格納する。次に、該stMinInterval[l]の中で最大値MaxMinInterval値を求める。すると、l番目の一時レイアウトの類似度評価値Similarity[l]は以下のようにして求めることができる。

$$\text{Similarity}[l] = 100 \times \text{stMinInterval}[l] / \text{MaxMinInterval}$$

すなわち、上記Similarity[l]は、最小撮影時間間隔が大きいほど100に近づき、小さいほど0に近づく値となっているため、画像類似度評価値として有効である。

【0097】

次に、レイアウトページ内のバランスを評価するための評価量として、色合いのバラつきについて説明する。例えば旅行テーマのレイアウトを作成したい場合、あまりに似通った色（例えば、青空の青、山の緑）の画像ばかりが並んでいたとすると、それは良いレイアウトとは言えないことがある。この場合は、色合いのバラつきが大きくなるように設定する。ここでは注目しているl番目の一時レイアウト内に存在する画像の平均色相AveHの分散を算出して、それを色合いのバラつき度tmpColorVariance[l]として格納する。次に、該tmpColorVariance[l]の中での最大値MaxColorVariance値を求める。すると、l番目の一時レイアウトの色合いバラつき度の評価値ColorVariance[l]は以下のようにして求めることができる。

$$\text{ColorVariance}[l] = 100 \times \text{tmpColorVariance}[l] / \text{MaxColorVariance}$$

すなわち、上記ColorVariance[l]は、ページ内に配置された画像の平均色相のバラつきが大きいほど100に近づき、小さいほど0に近づく値となっているため、色合いのバラつき度評価値として有効である。

【0098】

次に、レイアウトページ内のバランスを評価するための評価量として、顔の大きさのバラつき度について説明する。例えば旅行テーマのレイアウトを作成したい場合、レイアウト結果を見て、あまりに似通った顔のサイズの画像ばかりが並んでいたとすると、それは良いレイアウトとは言えないことがある。レイアウト後の紙面上における顔の大きさが、小さいものもあれば大きいものもあり、それらがバランスよく配置されていることが、良いレイアウトとする場合、顔サイズバラつきが大きくなるように設定する。その場合、注目しているl番目の一時レイアウト内に配置された後の顔の大きさ（顔位置の左上から右下までの対角線の距離）の分散値を、tmpFaceVariance[l]として格納する。次に、該tmpFaceVariance[l]の中での最大値MaxFaceVariance値を求める。すると、l番目の一時レイアウトの顔サイズバラつき度の評価値FaceVariance[l]は、以下のようにして求めることができる。

$$\text{FaceVariance}[l] = 100 \times \text{tmpFaceVariance}[l] / \text{MaxFaceVariance}$$

すなわち、上記FaceVariance[l]は、紙面上に配置された顔サイズのバラつきが大きいほど100に近づき、小さいほど0に近づく値となっているため、顔サイズのバラつき度評価値として有効である。

【0099】

またその他カテゴリとして、ユーザの嗜好性評価が考えられる。

【0100】

10

20

30

40

50

以上説明したような、各一時レイアウト毎に算出した複数の評価値を、以下では統合化して、各一時レイアウト毎のレイアウト評価値とする。今、 l 番目の一時レイアウトの統合評価値を、 $EvalLayout[l]$ とし、上記で算出した N 個の評価値（図2-1の評価値それぞれを含む）の値を、 $EvalValue[n]$ とする。この時、統合評価値は以下で求めることができる。 $EvalLayout[l] = \{EvalValue[n] \times W[n]\}$ ここで、 $n=0,1,2,\dots,N$ の合算演算記号を示す。また、 $W[n]$ は、図2-1で示したシーン毎の各評価値の重みである。該重みはレイアウトのテーマ毎に異なる重みを設定していることが特徴となっている。例えば、図2-1に示すようにテーマを成長記録“growth”と旅行“travel”で比較した場合、旅行テーマの方は、できるだけ良質の写真をいろいろな場面で数多くレイアウトすることが望ましいとし、画像の個別評価値やページ内のバランス評価値を重視する傾向に設定する。一方、成長記録“growth”の場合、画像のパリエーションよりは、成長記録の対象となる主人公が確実にスロットに適合しているか否かが重要である。そのため、ページ内バランスや画像個別評価よりも、画像・スロット適合度評価を重視する傾向に設定する。

10

20

30

40

50

【0101】

このようにして算出した $EvalLayout[l]$ を用いて、S2109では、レイアウト結果表示のためのレイアウトリスト $LayoutList[k]$ を生成する。レイアウトリストは、予め定められた個数（例えば5個）に対して、 $EvalLayout[l]$ のうち、評価値が高いもの順に識別子 l を記憶しておく。例えば最も良いスコアを出したものが、 $l=50$ 番目に作成した一時レイアウトであった場合、 $LayoutList[0]=50$ となる。同様に、 $LayoutList[1]$ 以降は、スコア値が2番目以降の識別子 l を記憶しておく。

【0102】

次に、図1-9を用いて説明したレイアウト生成処理によって得られたレイアウト結果を表示する。本実施形態では、レンダリングした結果を、図6のS605でレンダリングして結果を図2-6のように表示し、提案する。S605では、まず $LayoutList[0]$ に格納されているレイアウト識別子を読み出し、該識別子に相当する一時レイアウト結果を、2次記憶装置あるいはRAM上から読み出す。レイアウト結果には、上述したようにテンプレート情報と該テンプレート内に存在するスロット毎に、割り当てられた画像名が設定されている。S605ではこれらの情報に基づいて、情報処理装置115上で動作するOSの描画関数を用いて、該レイアウト結果をレンダリングし、図2-6の2902のようにポップアップし、表示する。

【0103】

同図の他の画像ボタン2903を押下することにより、次点スコアである $LayoutList[1]$ の識別子を読み出し、上記と同様にレンダリング後、表示を行うことで、ユーザは様々なパリエーションの提案レイアウトを閲覧することができる。さらに、表示されたレイアウトが気に入った場合には、印刷ボタン2905を押下することで、情報処理装置115に接続されたプリンタ112からレイアウト結果をプリントアウトすることもできる。

【0104】

上述した実施形態のうち、図5のS512にあたる顔辞書作成処理について説明する。

【0105】

図2-7に顔辞書作成処理S512のフローチャートを示す。以下、本実施形態では、説明を簡単にするために、顔を構成する部位による判定として、眉の角度、目の大きさ、口の角度を例に挙げて説明するが、これに限定されるものではない。例えば、さらに多くの部位を用いて判定ようにしてもよく、他の部位を用いて判定してもよい。

【0106】

S3301～S3303は、それぞれ、眉の角度 B 、目の大きさ ES 、口の角度 M （以下、説明のために、これらを特徴量要素と呼ぶ）の算出を行う。

【0107】

S3301では、眉角度を算出する。眉の角度は、例えば、両端部を直線で結んだ際の水平からの傾きから算出することができる。眉の角度を左右別々に取得するようにしてもよいが、本実施形態では、左右の眉の角度の平均を求めて、これを眉の角度とした。S3302では、目の大きさを算出する。目の大きさは、目の横（水平）方向、縦（垂直）方向の

高さを、左右別々に算出してもよいが、本実施形態では縦方向の左右の平均の値とした。S3303では、口の角度は、眉毛と同様に、水平からの傾きから算出した。口の角度を算出する際には、口は顔の中心で変局点を取ることも多いので、最大角度を利用したり、近似曲線を使って求めたりしてもよい。

【0108】

S3304では、顔辞書に登録された特徴量パターンの中に、類似のものがあるか否かを判定する。すなわち、算出した B 、 ES 、 M の組み合わせ（以下、これを特徴量パターンと呼ぶ）が顔辞書内に既に存在するか否かの判断を行う。このとき、既に顔辞書に格納されている特徴量パターンのうち、各特徴量要素の平均値 B_{Ave} 、 ES_{Ave} 、 M_{Ave} に対し、算出した眉の角度 B 、目の大きさ ES 、口の角度 M が予め定められ所定範囲内にあるか否かで判定する。この範囲を B_{Width} 、 ES_{Width} 、 M_{Width} と定義した場合には、次の評価を行えば良い。

$$(B_{Ave} - B_{Width}) < B < (B_{Ave} + B_{Width}) \quad \text{かつ}$$

$$(ES_{Ave} - ES_{Width}) < ES < (ES_{Ave} + ES_{Width}) \quad \text{かつ}$$

$$(M_{Ave} - M_{Width}) < M < (M_{Ave} + M_{Width})$$

【0109】

類似のものが存在しない場合は、S3305で新たな特徴量パターンとして、顔辞書に新規登録する。類似のものが存在する場合は、S3306において、該特徴量パターンと同一と判断され、追加登録される。

【0110】

次のS3307では、新規作成又は追加された特徴量パターンにおいて、各特徴量要素の統計値を作成又は更新を行う。統計値の一例としては、上記S3304で使用した平均値のほか、分散値などが挙げられる。

【0111】

以上により、顔辞書の追加、更新が繰り返され、S511において、全画像が終了と判断された場合には、既存画像をもとにした顔辞書が作成される。

【0112】

次にS513の細分化グループ生成について説明する。ここでは、顔辞書内の1つのグループ（例えば人物「太郎」）に属する特徴量をさらに細分化処理することで、よりきめ細かい判定が可能な辞書に修正する。図28に細分化グループ生成処理のフローチャートを示す。なお、細分化グループ生成処理は、該当する人物であると判定された顔の個数が、予め設定された個数以上になった場合に実行されるものとする。

【0113】

S3401では対象となる人物「太郎」と判定された画像群の特徴量の特徴量パターンを取得する。尚、上記と同様、特徴量パターンは、眉の角度、目の大きさ、口の角度の特徴量要素で構成されているものとして説明する。

【0114】

S3402では、特徴量要素の一つを取得する。そして、S3403にて、対象とする特徴量要素のピーク検出を行う。図29(a)に、ある特徴量パターンの一例を示す。横軸が特徴量を示し、矢印がその特徴量の存在範囲を示している。縦軸が度数（頻度）である。

【0115】

S3402において対象とする特徴量要素を眉の角度とした場合には、図29(a)では、2つのピークが検出される。ピークの検出は、幾つかの方法があるが、微分処理を用いて変局点（ここでは上に凸の変局点）を検出するものとする。なお、局所的な凹凸により誤検出してしまう場合は、スムージング処理や近似曲線を併用する方法も有用である。図29(a)の眉の角度を表わす特徴量範囲には、ピークが2つ検出される。また、同様の処理を適用すると、図29(a)では、目の大きさ、口の角度はそれぞれ1および2となる。

【0116】

S3404において、着目部位についてピークが複数あるか否かの判定、すなわち、小分類化できるか否かの判定を実施する。

10

20

30

40

50

【0117】

2つ以上のピークが存在する場合は、S3405へと進み、特徴量存在範囲を2以上の小範囲に分割する処理（ピーク分割処理）が実施される。その後、S3406に処理を進める。分割方法は、ピーク値を中心とした正規分布を考慮してもよいし、実際の半値幅をもとに推定する方法を利用してもよい。本実施形態では、図29(b)に示すように、分割をする。

【0118】

一方、S3404でピークが1以下であれば、ピーク分割はせず、現在の特徴量範囲を維持し、S3406に処理を進める。

【0119】

S3406において、全特徴量要素についてS3403からS3405が終了したか否かを判定する。終了した場合は、S3407に進み、終了していない場合は、S3402へ戻る。また、S3407において、全特徴量パターンについてS3402からS3406が終了したか否かを判定する。終了した場合は、S3408に進み、終了していない場合は、S3401へ戻る。

10

【0120】

S3408では、S3405において分割されたピークをもとに、再度グルーピングをやり直す。図29(a)では同一特徴量要素となっていたものを本処理後では、図29(b)のように分割し、それに基づき顔辞書を変更する。図29(b)の例では、図29(a)であった特徴量パターンが分割され、図29(b)の眉の角度は2種類（矢印の色：黒、白）、目は1種類のまま（同：黒）、口の角度は2種類（同：黒、白）に分割される。従って、分割処理後は下記2×1×2=4通りの特徴量パターンになる。

20

特徴量パターン1：眉（黒）×目（黒）×口（黒）

特徴量パターン2：眉（黒）×目（黒）×口（白）

特徴量パターン3：眉（白）×目（黒）×口（黒）

特徴量パターン4：眉（白）×目（黒）×口（白）

【0121】

尚、図29の眉の角度のように、正規分布などで推定した場合、重複する領域が存在することになる（同図(b)斜線部）。この場合、該領域に存在する画像をどちらに振り分けるのが問題となる。本実施形態では、両ピークの交点Aを求め、交点Aに相当する角度 M_A を境界に細分化を実施する。こうすることで、画像の細かい変化に感度をもつ辞書が作成されることになる。例えば、特徴量パターン1は、大きく笑った顔への検出、特徴量パターン2では少し笑った顔への検出感度をもつ辞書となる。

30

【0122】

さらに、本実施形態では、上記のようにして得られた各特徴量パターン毎（パターン1乃至4）の出現頻度を計数し、出現頻度が少ないパターンを除外する。図37にそのフローチャートを示す。S4301で対象とする特徴量パターンを取得し、S4302でその出現頻度（ここでは画像の中の出現回数）を計測する。S4303において、予め決められた所定数と比較し、所定数未満の場合はS4304へ進み、所定数以上の場合は、S4305へ進む。S4304では、該当するパターンは所定基準を満たさないとして削除する。すなわち、所定数以下であるパターンを個人識別のためのパターンから除外する。このとき、取りえない組み合わせのパターンは、個人識別のためのパターンから除外される。これを細分化した全特徴量パターンに対し実施する。S4305では、全特徴量パターンが終了したか判定し、終了していない場合は、S4301へ戻り、終了している場合は処理を終了する。これにより、各部位の取り得る組み合わせを規定し、目は笑っているものの口は怒っているなどの、実際に取り得ない組み合わせを除外し、出現頻度の高いパターンで辞書を更新することで、辞書の精度、すなわち、人物識別の精度の向上を図ることが出来る（精度の高い辞書管理が可能になる）。

40

【0123】

また、こうした処理は表情だけでなく、眉と口の角度は似ているものの、目の大きさが少し違うといった、似た顔をもつ異なる人物（例えば、兄弟や親子など）を区別するのにも有効となる。

50

【 0 1 2 4 】

尚、顔の変化の中には、メガネなどの装飾品を付けた場合も含めることが出来る。レンズの影響で目の大きさなどが変化することも起こり、この場合も独立した特徴量パターンとして辞書に登録することにより、検出性能を持つことが出来る。

【 0 1 2 5 】

上記の処理の結果、例えば人物「太郎」を識別するための顔辞書は、図30に示すように構築される。図示の如く、特徴量パターン1、2、4を用いて「太郎」を識別することと決定し、特徴パターン3は「太郎」を識別するためには不使用とする。不使用とするとは、識別対象の顔画像の特徴が特徴量パターン3であったとしても、それを「太郎」とは識別しないことを意味する。この結果、たまたま他人が特徴量パターン3を持つ顔を持っていたとしても、それを誤って太郎とは識別しなくなる。すなわち、他人が誤って「太郎」として判定される可能性を低くする事ができ、顔識別の精度を向上させることが可能になる。

10

【 0 1 2 6 】

上述した顔辞書を利用し、新たに入力された画像の個人認識を行う。その際、顔辞書についても、適宜修正を行うことにより、個人認識の精度を向上させることができる。当初の顔辞書を作成した画像群が、必ずしも十分なサンプル数であったとは限らないからである。一方で、十分最適な顔辞書になっているにも関わらず、修正を行い精度の低下を引き起こす恐れもある。それらを考慮し、下記処理により、辞書の更新を実施する。

【 0 1 2 7 】

図31に、新たに入力された画像内に存在する顔の個人認識および顔辞書の更新のフローチャートを示す。

20

【 0 1 2 8 】

S3701では、対象とする顔の画像を取得する。この顔の画像は、S504で記載されたものと同様に正規化処理実施済みのものとする。

【 0 1 2 9 】

S3702では、S3701で取得した画像より、特徴量要素である、眉の角度B、目の大きさES、口の角度Mを算出する。

【 0 1 3 0 】

S3703では、顔辞書内に登録された特徴量パターンを参照し、類似のパターンがあるか否かの判定を行う。図27のS3304で示した処理と同様の判定でも構わないが、予め設定する所定の範囲 B_{Width2} 、 ES_{Width2} 、 M_{Width2} （いずれも正の値としている）は、S3304のそれらよりも小さい値にすることが望ましい。既に細分化グループ生成が実施されているものであり、各特徴量パターンのもつ特徴量要素の幅が小さいためである。

30

$$(B_{Ave} - B_{Width2}) < B < (B_{Ave} + B_{Width2}) \text{ かつ}$$

$$(ES_{Ave} - ES_{Width2}) < ES < (ES_{Ave} + ES_{Width2}) \text{ かつ}$$

$$(M_{Ave} - M_{Width2}) < M < (M_{Ave} + M_{Width2})$$

【 0 1 3 1 】

上記条件による判定が実行され、類似パターンが存在する場合はS3704へ進み、存在しない場合はS3705に進む。S3704では、該特徴量パターンとひも付けられた（関連付けられた）名前が特定され、対象の顔の個人認識が実施される。

40

【 0 1 3 2 】

S3705では、該当する顔（人物）が存在しなかったため、処理は終了になる。

【 0 1 3 3 】

S3704では、個人認識が実施された後、辞書の更新処理に移行する。S3706では、対象となった特徴量パターンのもつ統計値を取得する。このうち、ばらつきを示す分散値 OLD に着目する。

【 0 1 3 4 】

S3707において、既存の分散値 OLD が、予め設定した所定値以上であるか否かの判定を実施する。既存の分散値 OLD が所定値以上ではない場合、すなわち、十分に小さい値に

50

なっている場合は、特徴量パターンが既に十分な精度が出ていると判断し、S3708に進む。この場合は、特に顔辞書は更新されない。

【0135】

一方、既存の分散値 σ_{OLD} が所定値以上と判定された場合は、まだ特徴量パターンを形成するのに十分なサンプリング数でないと判断し、S3709に進む。

【0136】

S3709では対象となる顔の特徴量パターンを追加した場合の統計量を算出する。新たに計算された分散値を σ_{NEW} とする。

【0137】

S3710では、 σ_{OLD} と σ_{NEW} の比較を行い、 σ_{NEW} が小さくなる場合、つまり追加した方が、ばらつきが小さくなる場合には、S3711に進み、顔辞書に特徴量パターンの追加が行われる。ばらつきが小さくならない場合は、S3708に進み、特に顔辞書は更新されない。

10

【0138】

本実施形態では、顔辞書を用いた個人認識精度を向上させることができる。例えば、デジタル画像中の顔の個人認識処理において、表情の異なる画像について、十分な精度を実現することができる。尚、異なる表情の中には、メガネなどの装飾品による変化も含まれる。以上のように、個人認識を実施すると同時に、必要と判断した場合には顔辞書の更新を行うことで、顔辞書の学習機能を有することが出来る。また、本実施形態に係る辞書を用いた上記処理により、提案型でコラージュ画像を作成する際に、既存の画像に類似する表情の画像であっても正確に個人認識することが可能となる。このように、誤った個人認識が抑制されることにより、適切なコラージュ画像を提案することができる。なお、本実施形態では、ピークを分割することにより、特徴量を細分化したが、特徴量を細分化できるものであれば、これに限定されない。

20

【0139】

[実施形態2]

上記実施形態1では、特徴量パターンを構成する特徴量要素として、眉の角度、目の大きさ、口の角度を例に挙げた。

これに対し、本実施形態では、人物識別のためのパターンの1要素として時間を取り入れた例を説明する。本実施形態は、顔辞書以外は、実施形態1と同様であるので、実施形態1と同様の部分については説明を省略する。

30

【0140】

人物識別のためのパターンの1要素として、時間要素を取り入れる場合、各画像とその撮影日時がひも付けされることが必要になる。本実施形態では、各画像は図11に示す属性情報を用いて記憶管理されるので、図11の記述におけるタグ「CaptureDateTime」からその撮影時刻を得ることができる。

【0141】

図32は、人間の顔のある変化の度合いを示したイメージ図である。縦軸はある特徴量を示す軸、横軸は時間軸である。図32に示すように、人間の顔は、誕生から変化を始め、死亡するまで常に変わり続ける。したがって、本実施形態では、加齢による変化という要素を加えて顔辞書を細分化する。

40

【0142】

実施形態1で細分化した特徴量パターンを用いることによる多様な表情検出について説明したが、本実施形態では、さらに、その中で加齢による変化という要素を加えてさらに細分化し、個人認識精度を向上させた。

【0143】

顔辞書の作成に係る処理手順を図35に示す。尚、S4101~S4107については、図27のS3301~S3307と同様であるため説明は省略する。

【0144】

S4108では、S4106で画像の特徴量要素を辞書に追加した後、該画像の時間情報を取得する。時間の情報は、画像ファイルに付加されているメタデータ、例えばExifの撮影日時を

50

参照すればよい。メタデータが含まれない場合は、ファイルの作成日時を参照することでも代用可能である。

【0145】

画像ファイルのExif情報は、撮影日時として秒単位の詳細なものが記載されているが、本処理ではそこまでの情報は必要ないため、日、月、年などの単位でまとめた時間の情報を取得するようにしてもよい。また、図32の顔の変化に示したように、幼少の頃の方が大きく、年齢を重ねるにつれ、変化率が小さくなることを考慮し、対象とする人物の年齢に応じ、取得する時間の情報の単位を切り替えるようにしても良い。

【0146】

最初の検出時で、まだ年齢や名前などのプロフィールが取得出来ていない場合は、予め決められた期間を使用すればよい。

10

【0147】

顔辞書の細分化グループ生成処理のフローは実施形態1の図28に示したものと同様であるが、S3408のグルーピング処理において、取得した時間の情報を利用し、実施を行う点で異なる。その部分について説明を加える。

【0148】

図33は、本実施形態の特徴量要素を示す図である。実施形態1の図29で示した特徴量要素、眉の角度、目の大きさ、口の角度に加え、時間の特徴量要素が加わったものになっている(図33)。時間の特徴量要素の場合、一定期間撮影した後、しばらく経って撮影というケースもあり、正規分布になるとは限らない。そうしたことを考慮し、上述した日、月、年など決められた期間の最初と最後の日時を決定すればよい。

20

【0149】

ピークで分離したものに対し、さらに時間の要素を考慮したグルーピング処理を実施する。図33において、特徴量要素としての時間を、黒、斜線、白の3つの時間範囲で表している。

【0150】

この時間の要素を、他の特徴量要素に適用したものが、それぞれのピーク部分である。例えば、眉の角度であれば、ピーク検出による細分化により、2つに分けられるが、さらに時間の要素により、5つに細分化される(黒:2、斜線:2、白:1)。同様に、目の大きさ、口の角度は、それぞれ、3(黒:1、斜線:1、白:1)、4(黒:2、斜線:1、白:1)に細分化される。顔辞書としては、同一時間の要素を掛け合わせて、登録することから、

30

$$\text{黒} : 2 \times 1 \times 2 = 4$$

$$\text{斜線} : 2 \times 1 \times 1 = 2$$

$$\text{白} : 1 \times 1 \times 1 = 1$$

の合計7つの特徴量パターンに分類される。

【0151】

この後、第1実施形態で図37を用いて説明した、出現頻度による、取り得ない組み合わせの除去を実施する。

【0152】

40

こうした処理を実施することにより、図34に示すような顔辞書が構成される。すなわち、7つの特徴量パターンのうち、特徴量パターン1、2、4、5、7は「一郎」を識別するために利用し、残りの特徴量パターン3、6は「一郎」を識別するためには使用しないことにする。

【0153】

上記で作成した時間を考慮した顔辞書を利用し、新たに入力された画像の個人認識を行っていく方法について説明する。実施形態1と同様に、個人認識すると同時に、その結果をもとに辞書の更新を行う。処理フローは、図31のフローと同様であるため、説明は省略する。

【0154】

50

ここでは、辞書の更新方法の一つとして、顔辞書の容量最適化処理について説明する。本実施形態のシステムでは、顔辞書のデータは、2次記憶装置、図1の103に記憶されることになる。2次記憶装置103は、一般的に有限な領域のため、顔辞書データを蓄積する際も限られた容量に抑える必要がある。つまり、予め決められた容量をオーバーした場合は、データベースに登録してある何かしらのデータを削除し、容量を押さえる処理が必要となる。

【0155】

図36は、容量を抑える処理の一例の処理フローを示したものである。S4201では、顔辞書データベースの容量が一定以上に到達しているか否かの判定処理を実施する。本判定処理で、一定以上に達していない、すなわち、まだ容量に十分な空きがあると判定された場合は、何も行われず、処理を終了する。容量が一定以上であり、データの削減が必要だと判断された場合は、S4202へと進む。S4202では、顔辞書を生成した際に利用した時間の特微量要素を参考に最も古い時間で構成される特微量パターンを検出し、取得する。S4203では、S4202で取得した特微量パターンの削除を実行する。実行後、再度、容量の確認を行い、一定以下になるまで繰り返し処理が実施される。

10

【0156】

図33に示した例では、時間の特微量要素のうち、左側が古い、つまり最も古いものを黒とした場合に、黒の特微量要素を含むデータを削除する。つまり、本例であれば、黒の特微量パターン4つが削除されることになる。

【0157】

顔が刻々と変化することは上述したが、新たに入力され画像に対し、個人認識処理を実施するため参照する顔辞書は、新しいデータで構成することにより、より個人認識精度を向上させることができる。すなわち、古い時間の含むものを取り除いて行くことにより、より個人認識を向上させることができる。なお、時間を利用し、細分化した特微量要素で構成されている顔辞書では、この処理を実現するのが容易である。

20

【0158】

尚、図36で記載した処理フローでは、容量を参照し、その実施可否を判断したが、これ以外にもユーザが明示的に容量を削減する処理を実施してもよい。

【0159】

[その他の実施形態]

以上説明した実施形態は本発明の効果を得るための一手段であり、類似の別手法を用いたり、異なるパラメータを用いたとしても、本発明と同等の効果が得られる場合は、本発明の範疇に含まれる。

30

【0160】

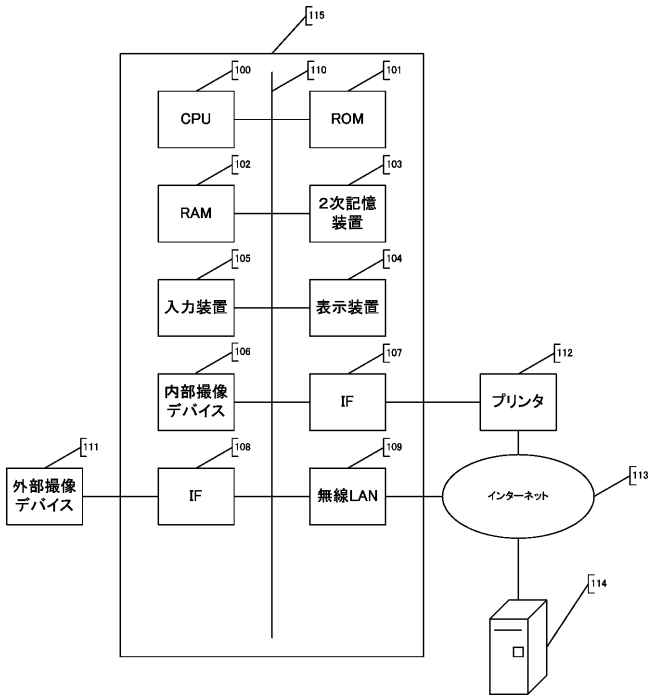
尚、本発明は、以下の処理を実行することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア(プログラム)を、ネットワーク又は各種記憶媒体を介してシステムまたは装置に供給し、そのシステムまたは装置のコンピュータ(またはCPUやMPU等)がプログラムを読み出して実行する処理である。また、プログラムは、1つのコンピュータで実行させても、複数のコンピュータを連動させて実行させるようにしてもよい。また、上記した処理の全てをソフトウェアで実現する必要はなく、一部または全部をハードウェアによって実現するようにしてもよい。

40

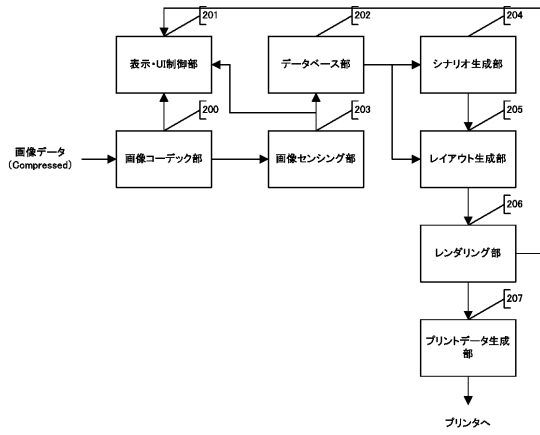
【0161】

上述した実施形態では、レイアウト出力物として、1ページに複数の画像を配置した出力物について説明したが、勿論、複数ページの出力にも適用することができる。

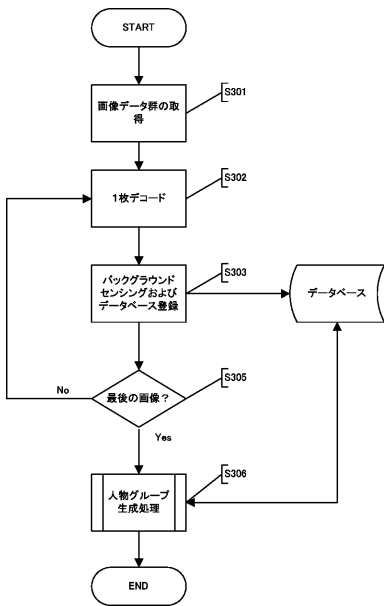
【 図 1 】



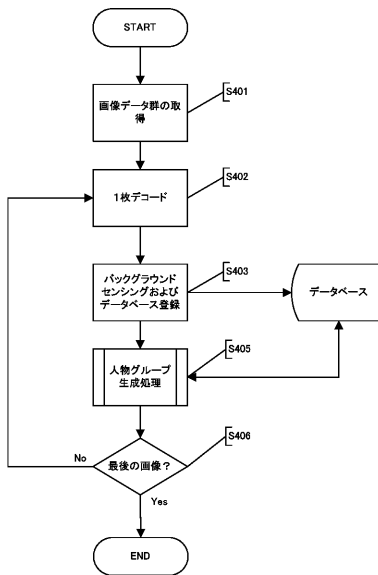
【 図 2 】



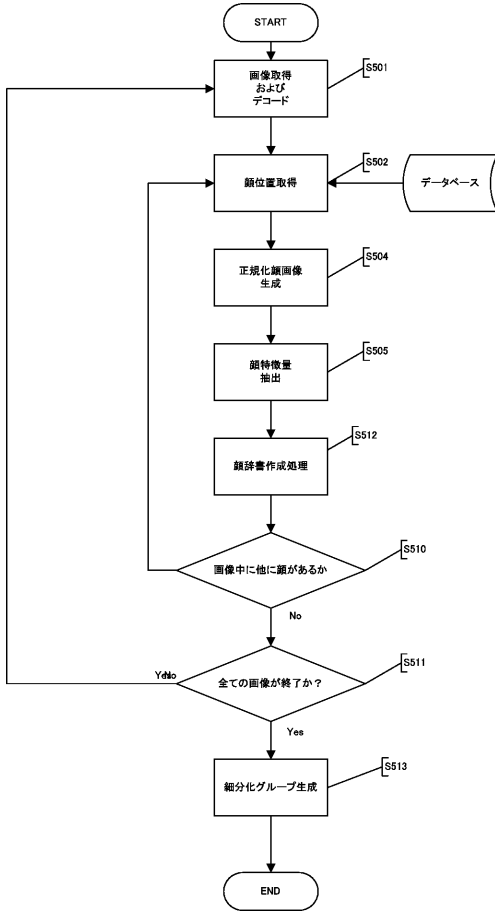
【 図 3 】



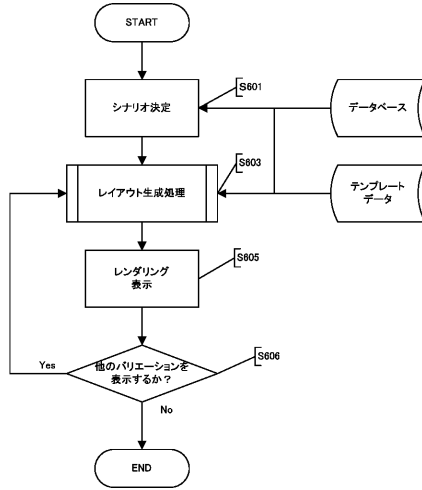
【 図 4 】



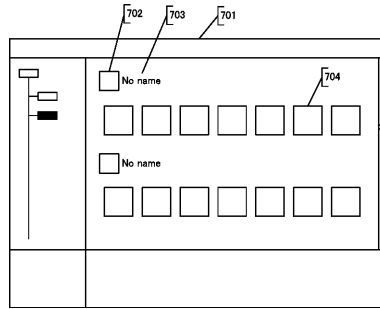
【 図 5 】



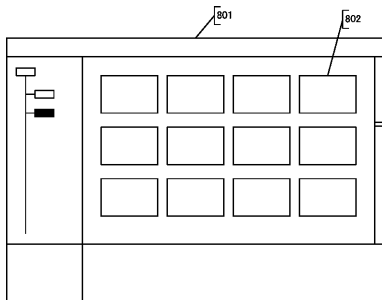
【 図 6 】



【 図 7 】



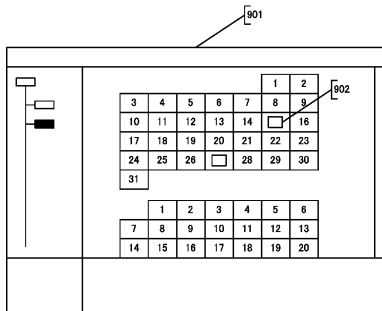
【 図 8 】



【 図 10 】

センシング大分類	センシング小分類	データの型	値	
画像基本特徴量	平均輝度	int	0~255	
	平均彩度	int	0~255	
	平均色相	int	0~359	
顔検出	人物顔の個数	int	0~MAXFACE	
	座標位置	int * 8	0~Width or Height	
	顔領域内の平均Y	int	0~255	
	顔領域内の平均Cb	int	-128~127	
	顔領域内の平均Cr	int	-128~127	
シーン解析	シーン結果		char	Landscape Nightscape Portrait Underexposure Others

【 図 9 】



【 図 11 】

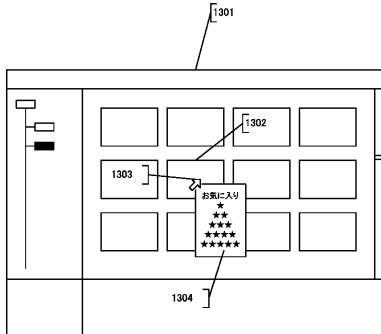
```

<?xml version="1.0" encoding="utf-8"?>
<IMAGEINFO>
  <BaseInfo>
    <ID>00000001</ID>
    <ImagePath>C:\My Picture\IMG0001.jpg</ImagePath>
    <ImageSize width=3000, height=2000 />
    <CaptureDateTime>20100101:120000</CaptureDateTime>
  </BaseInfo>
  <SensInfo>
    <AveY>128</AveY>
    <AveS>38</AveS>
    <AveH>50</AveH>
    <SceneType>Landscape</SceneType>
  </SensInfo>
  <Person>
    <ID>0</ID>
    <Position>
      <LeftTop x=420, y=200/>
      <LeftBottom x=420, y=300/>
      <RightTop x=520, y=200/>
      <RightBottom x=520, y=300/>
    </Position>
    <AveY>128</AveY>
    <AveCb>20</AveCb>
    <AveCr>20</AveCr>
  </Person>
  ...
</SensInfo>
<UserInfo>
  <FavoriteRate>3</FavoriteRate>
  <ViewingTimes>5</ViewingTimes>
  <ImageRate>70</ImageRate>
  <PrintingTimes>3</PrintingTimes>
  <Event>Travel</Event>
</UserInfo>
  ...
</IMAGEINFO>
    
```

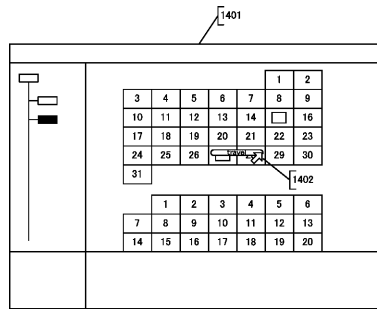
【図 1 2】

分類	内容	データの型	値
画像	自動お気に入り度	int	0~100
	自動お気に入り度	int	0~100
	イベント	char	"travel" "graduation" "wedding"
人物	名前	char	"NAME"
	誕生日	char	YYYYMMDD
	続柄	char	"family"

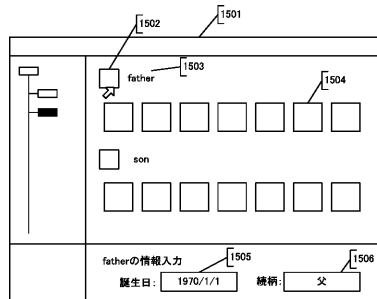
【図 1 3】



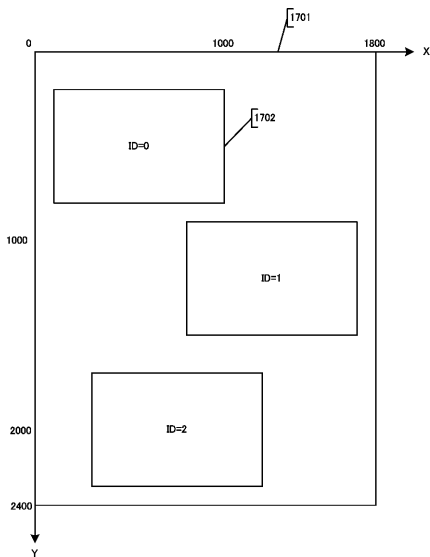
【図 1 4】



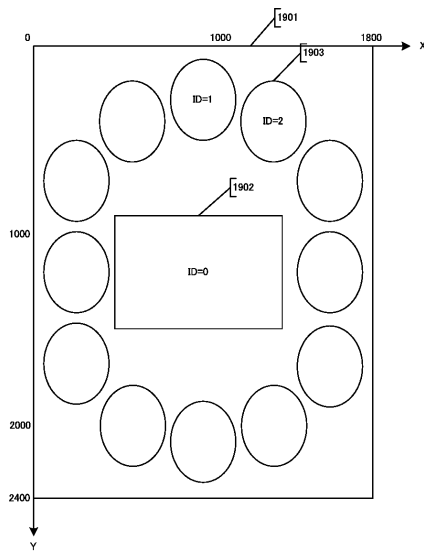
【図 1 5】



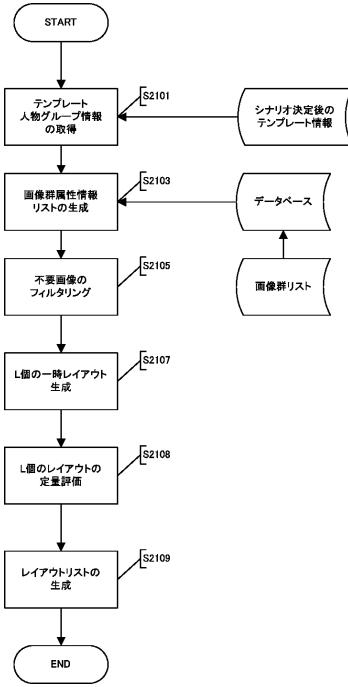
【図 1 6】



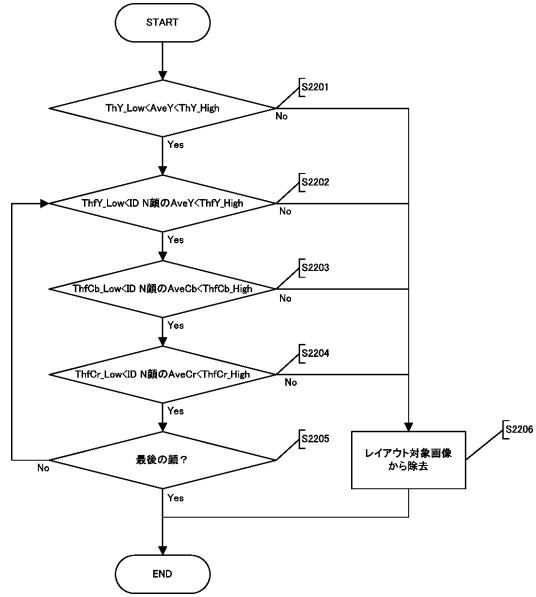
【図 1 7】



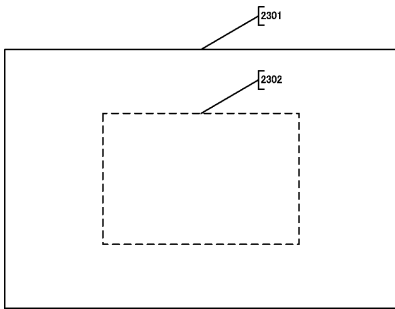
【図18】



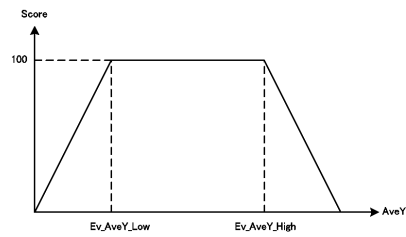
【図19】



【図20】



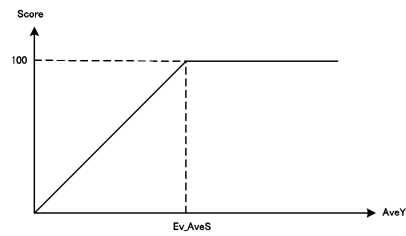
【図22】



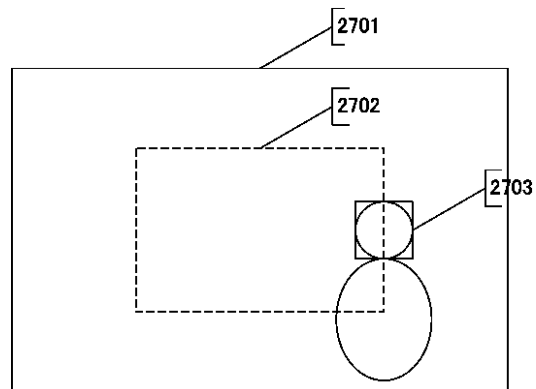
【図21】

カテゴリ	内容	スコアレンジ	テーマ毎重要度(重みW)			
			growth	travel	favorite	...
画像個別評価	明るさ適正度	0~100	0.5	1.0	0.0	
	彩度適正度	0~100	0.5	1.0	0.0	
画像・スロット適合度評価	人物適合度	0~100	1.0	0.5	0.0	
	トリミング欠け判定	0~100	1.0	0.5	0.0	
ページ内バランス評価	画像類似性	0~100	0.5	1.0	0.0	
	色合いばらつき	0~100	0.5	1.0	0.0	
	顔サイズばらつき	0~100	0.5	1.0	0.0	
その他	ユーザー嗜好性	0~100	0.8	0.8	1.0	

【図23】



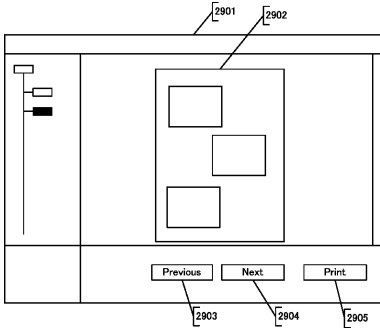
【図24】



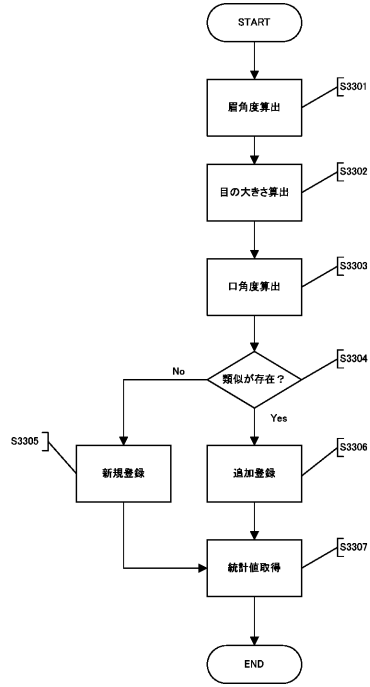
【図 25】

画像ID	撮影日時(YYYYMMDD:HHMMSS)
25	20100101:120000
86	20100101:150000
102	20100101:170000
108	20100101:173000

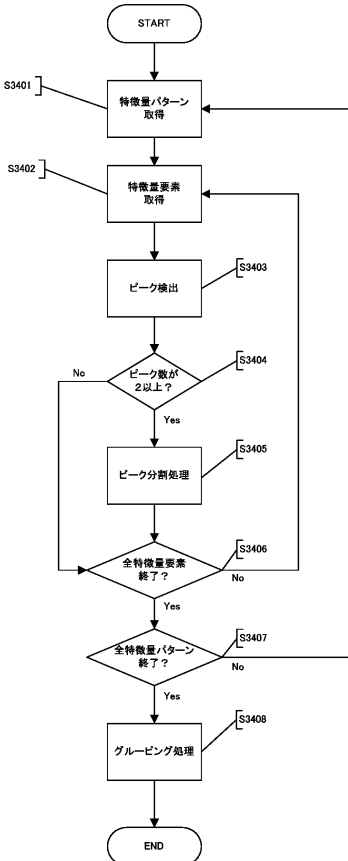
【図 26】



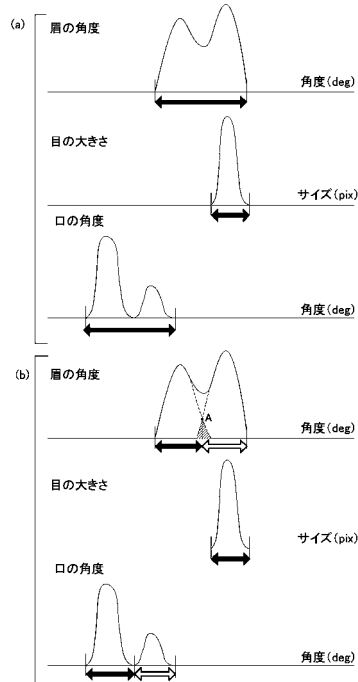
【図 27】



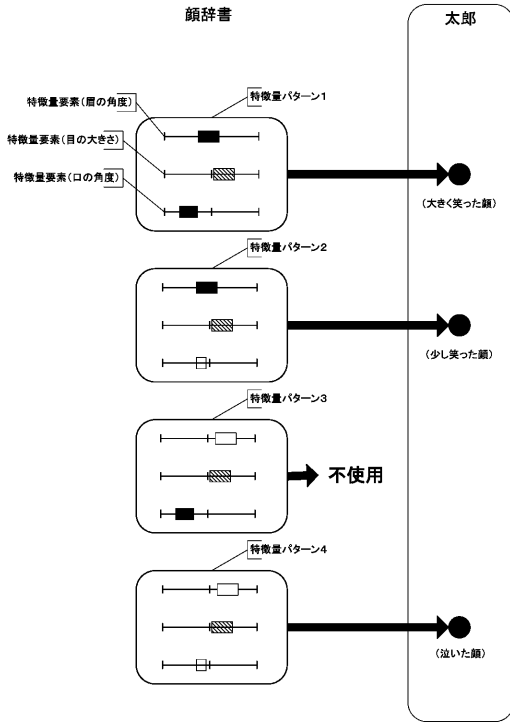
【図 28】



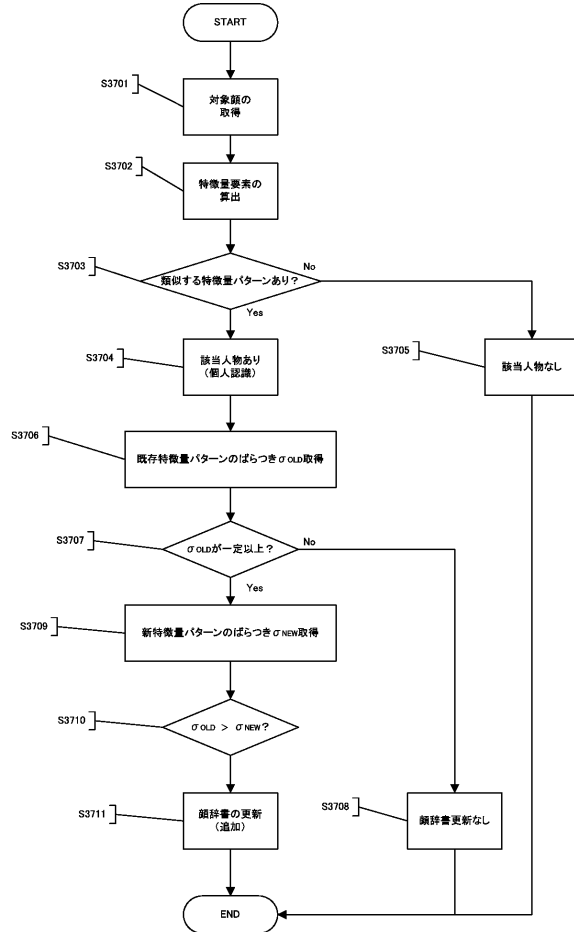
【図 29】



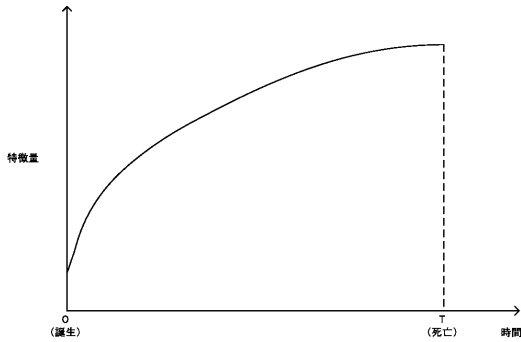
【図30】



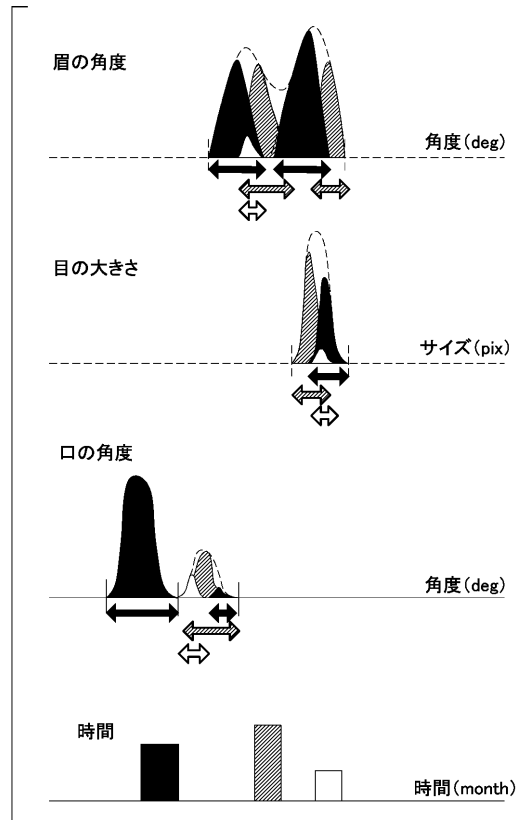
【図31】



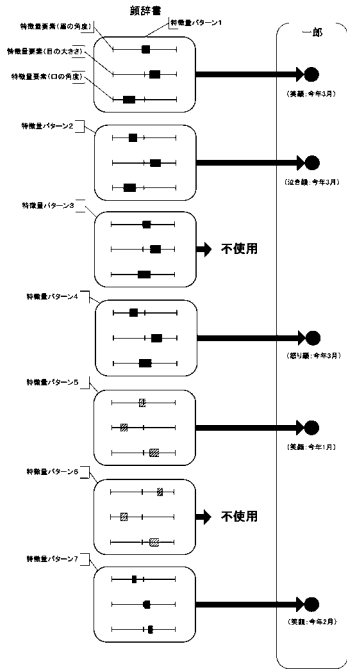
【図32】



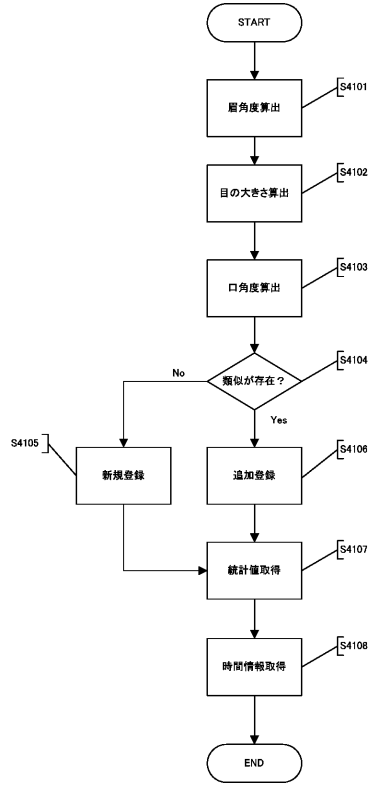
【図33】



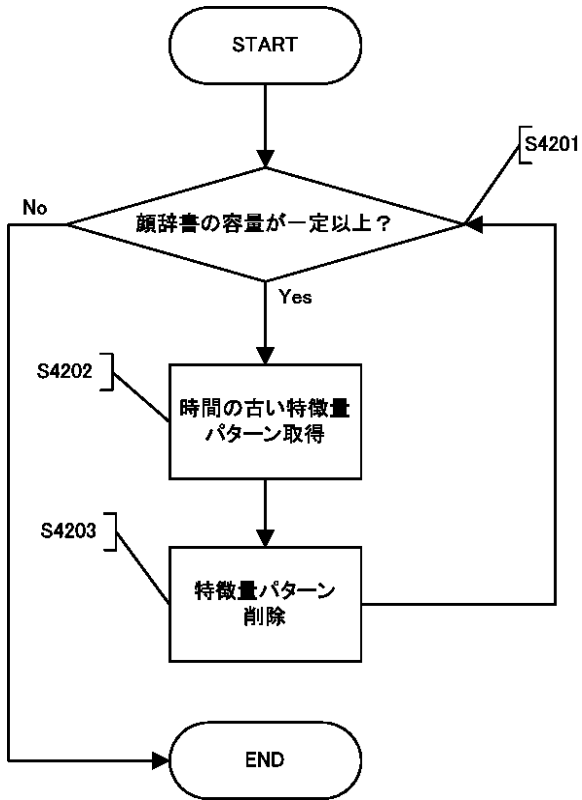
【図34】



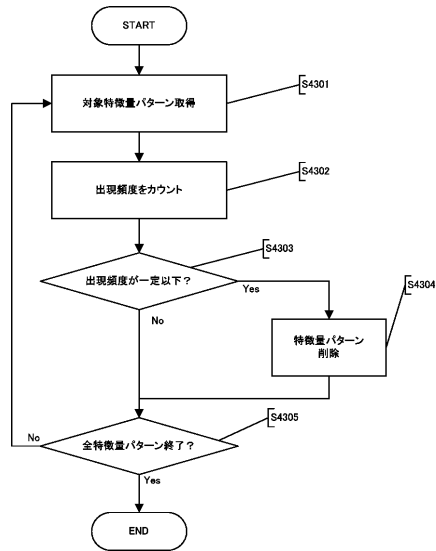
【図35】



【図36】



【図37】



フロントページの続き

- (72)発明者 諏訪 徹哉
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
- (72)発明者 國枝 寛康
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
- (72)発明者 橋井 雄介
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
- (72)発明者 鷺見 尚紀
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
- (72)発明者 中村 隆
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
- (72)発明者 藤田 貴志
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
- (72)発明者 梅田 清
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
- (72)発明者 酒井 洋行
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

Fターム(参考) 5B043 AA09 BA04 FA08 GA17

5B057 BA02 CA01 CA08 CA12 CA16 CE16 CH11 DA07 DA12 DA16
DC04 DC05 DC25 DC36 DC38 DC39