

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5229744号
(P5229744)

(45) 発行日 平成25年7月3日(2013.7.3)

(24) 登録日 平成25年3月29日(2013.3.29)

(51) Int.Cl.		F I			
G06F 17/30	(2006.01)	G06F 17/30	210D		
G06T 1/00	(2006.01)	G06T 1/00	200A		
		G06F 17/30	350C		
		G06F 17/30	170B		

請求項の数 8 (全 37 頁)

(21) 出願番号	特願2009-544662 (P2009-544662)	(73) 特許権者	504173471 国立大学法人北海道大学 北海道札幌市北区北8条西5丁目
(86) (22) 出願日	平成20年12月1日(2008.12.1)	(74) 代理人	100083806 弁理士 三好 秀和
(86) 国際出願番号	PCT/JP2008/071803	(74) 代理人	100101247 弁理士 高橋 俊一
(87) 国際公開番号	W02009/072466	(72) 発明者	長谷山 美紀 北海道札幌市北区北14条西9丁目 国立 大学法人 北海道大学 大学院情報科学研究科内
(87) 国際公開日	平成21年6月11日(2009.6.11)		
審査請求日	平成23年11月24日(2011.11.24)		
(31) 優先権主張番号	特願2007-312371 (P2007-312371)	審査官	田中 幸雄
(32) 優先日	平成19年12月3日(2007.12.3)		
(33) 優先権主張国	日本国(JP)		

最終頁に続く

(54) 【発明の名称】 画像分類装置および画像分類プログラム

(57) 【特許請求の範囲】

【請求項1】

複数の画像データを、類似した画像毎に分類する画像分類装置であって、
 複数の画像データが記憶された画像データベースと、
 前記複数の画像データのそれぞれについて、画像全体の特徴量を算出するとともに、前記画像データのエッジを検出し、検出されたエッジ部分の特徴量を算出する特徴量算出部と、
 前記画像全体における特徴量に基づいて、前記複数の画像データを複数のクラスタに分類する第1のクラスタリング部と、
 前記エッジ部分の前記特徴量に基づいて、前記第1のクラスタリング部によって分類された複数のクラスタを、更に複数のクラスタに分類する第2のクラスタリング部と、
 前記複数の画像データのそれぞれについて、被写体を構成する画素を決定し、前記第2のクラスタリング部によって分類された複数のクラスタを、前記被写体を構成する画素に基づいて統合するクラスタ統合部と、
 前記複数の画像データについてそれぞれのサムネイルをランダムに配置し、前記画像データの前記画像全体の特徴量、前記エッジ部分の特徴量、前記被写体を構成する画素の特徴量および前記クラスタ統合部により生成された前記画像データのクラスタに基づいて、前記各サムネイルの座標を更新して表示するとともに、前記各サムネイルの移動量を計算し、前記各サムネイルの移動量が0に収束するまで、座標を更新して表示する処理を繰り返す表示部

10

20

とを備える画像分類装置。

【請求項 2】

クエリ画像データの特徴ベクトルを算出するクエリ画像特徴量算出部と、
前記クエリ画像データの前記特徴ベクトルに基づいて、前記クラスタ統合部により生成されたクラスタから、前記クエリ画像データの所属するクラスタを決定するクラスタ決定部

とを備える請求項 1 に記載の画像分類装置。

【請求項 3】

前記特徴ベクトルは、画像全体における色コリログラムと、エッジ部分の色コリログラムと、被写体領域の色ヒストグラムと色コリログラムをパラメータとして有し、

前記クエリ画像特徴量算出部は、前記クエリ画像データの前記画像全体における色コリログラムと、前記クエリ画像データの前記エッジ部分の色コリログラムと、前記クエリ画像データの前記被写体領域の色ヒストグラムと色コリログラムを算出して、前記クエリ画像データの特徴ベクトルを算出する

請求項 2 に記載の画像分類装置。

【請求項 4】

前記クラスタ決定部は、前記クラスタ統合部により生成された各クラスタに属する画像データの特徴ベクトルの平均を算出し、前記クエリ画像データの特徴ベクトルとの距離を最小とするクラスタをクエリ画像の所属クラスタとする

請求項 3 に記載の画像分類装置。

【請求項 5】

複数の画像データを、類似した画像毎に分類する画像分類プログラムであって、コンピュータを、

画像データベースに記憶された複数の画像データのそれぞれについて、画像全体の特徴量を算出するとともに、前記画像データのエッジを検出し、検出されたエッジ部分の特徴量を算出する特徴量算出手段と、

前記画像全体における特徴量に基づいて、前記複数の画像データを複数のクラスタに分類する第 1 のクラスタリング手段と、

前記エッジ部分の前記特徴量に基づいて、前記第 1 のクラスタリング手段によって分類された複数のクラスタを、更に複数のクラスタに分類する第 2 のクラスタリング手段と、

前記複数の画像データのそれぞれについて、被写体を構成する画素を決定し、前記第 2 のクラスタリング手段によって分類された複数のクラスタを、前記被写体を構成する画素に基づいて統合するクラスタ統合手段と、

前記複数の画像データについてそれぞれのサムネイルをランダムに配置し、前記画像データの前記画像全体の特徴量、前記エッジ部分の特徴量、前記被写体を構成する画素の特徴量および前記クラスタ統合手段により生成された前記画像データのクラスタに基づいて、前記各サムネイルの座標を更新して表示するとともに、前記各サムネイルの移動量を計算し、前記各サムネイルの移動量が 0 に収束するまで、座標を更新して表示する処理を繰り返す表示手段

として機能させる画像分類プログラム。

【請求項 6】

クエリ画像データの特徴ベクトルを算出するクエリ画像特徴量算出手段と、

前記クエリ画像データの前記特徴ベクトルに基づいて、前記クラスタ統合手段により生成されたクラスタから、前記クエリ画像データの所属するクラスタを決定するクラスタ決定手段

として、更に前記コンピュータを機能させる請求項 5 に記載の画像分類プログラム。

【請求項 7】

前記特徴ベクトルは、画像全体における色コリログラムと、エッジ部分の色コリログラムと、被写体領域の色ヒストグラムと色コリログラムをパラメータとして有し、

前記クエリ画像特徴量算出手段は、前記クエリ画像データの前記画像全体における色コ

10

20

30

40

50

リログラムと、前記クエリ画像データの前記エッジ部分の色コリログラムと、前記クエリ画像データの前記被写体領域の色ヒストグラムと色コリログラムを算出して、前記クエリ画像データの特徴ベクトルを算出する

請求項 6 に記載の画像分類プログラム。

【請求項 8】

前記クラスタ決定手段は、前記クラスタ統合手段により生成された各クラスタに属する画像データの特徴ベクトルの平均を算出し、前記クエリ画像データの特徴ベクトルとの距離を最小とするクラスタをクエリ画像の所属クラスタとする

請求項 7 に記載の画像分類プログラム。

【発明の詳細な説明】

10

【技術分野】

【0001】

本発明は、複数の画像データを、分類する画像分類装置および画像分類プログラムに関する。

【背景技術】

【0002】

近年、デジタルカメラやイメージスキャナの普及、及び記録媒体の大容量化に伴い、ユーザが保持するデジタル画像の数は、急速に増加している。このような状況において、蓄積された大量の画像群から所望の画像を効率的に閲覧するために、画像検索技術の必要性が高まっている。代表的な検索手法として、ユーザの保持する検索要求画像（クエリ画像）とデータベース中に存在する画像との類似度に基づく手法がある（例えば、非特許文献 1 および非特許文献 2 参照）。これらの方法を用いることで、ユーザは、所望の画像をデータベースから取得することができる。

20

【0003】

また、クエリ画像に類似する画像を検索する方法として、クエリ画像と他の画像データとの画像間距離を計算して、画像間距離の小さいものから順に類似画像として抽出する画像検索装置がある（例えば、特許文献 1 参照）。上記特許文献 1 においては、クエリ画像を保持していなくとも、ユーザがクエリ画像を描画することにより、ユーザは、所望の画像を検索することができる。

【0004】

30

【特許文献 1】特開 2000 - 148795 号公報

【非特許文献 1】M.J. Swain and D.H. Ballard, "Color indexing," Int. J. Comput. Vision, vol.7, no.1, pp.11 - 32, 1991.

【非特許文献 1】中川俊昭、原武史、藤田広志、"局所的なパターンマッチングによる画像検索法、" 信学論(D-II), vol.J85-D-II, no.1, pp.149 - 152, Jan. 2002.

【発明の開示】

【0005】

しかしながら、上記に記載した方法は、ユーザがクエリ画像を保持していることを前提としている。従って、クエリ画像を保持していない場合や、検索要求が不明確である場合、ユーザは、検索することが困難となる。特許文献 1 においては、ユーザがクエリ画像を描画するので、ユーザが上手く描画できない場合には、ユーザが所望の画像を検索することが困難になってしまう。また、ユーザが所望の画像が明確でない場合には、特許文献 1 に記載の方法を適用することはできない。

40

【0006】

このような場合、サムネイル表示を用いてユーザにデータベース中に存在する全ての画像を提示することで、ユーザの画像検索作業が補助されていた。しかしながら、近年のデータベースの大規模化に伴い、データベース中に存在する全ての画像を一画面に表示することができなくなっている。その結果、サムネイル表示により画像検索の補助することは困難となっている。

【0007】

50

したがって、データベース中に存在する画像を自動で分類し、ユーザに分類結果を効果的に提示する技術が必要となる。すなわち、データベース中の画像を事前に類似した画像群に分類し、その分類結果を提示することで、ユーザはデータベースに含まれる画像を直観的に把握することができる。これにより、クエリ画像を保持していない場合や明確な検索意図が存在しない場合にも、ユーザが効率的に画像の検索することのできる技術が期待されている。

【0008】

従って本発明の目的は、複数の画像データを類似した画像毎に分類する画像分類装置および画像分類プログラムを提供することである。

【0009】

上記課題を解決するために、本発明の第1の特徴は、複数の画像データを、類似した画像毎に分類する画像分類装置に関する。即ち本発明の第1の特徴に係る画像分類装置は、複数の画像データが記憶された画像データベースと、複数の画像データのそれぞれについて、画像全体の特徴量を算出するとともに、画像データのエッジを検出し、検出されたエッジ部分の特徴量を算出する特徴量算出部と、画像全体における特徴量に基づいて、複数の画像データを複数のクラスタに分類する第1のクラスタリング部と、エッジ部分の特徴量に基づいて、第1のクラスタリング部によって分類された複数のクラスタを、更に複数のクラスタに分類する第2のクラスタリング部と、複数の画像データのそれぞれについて、被写体を構成する画素を決定し、第2のクラスタリング部によって分類された複数のクラスタを、被写体を構成する画素に基づいて統合するクラスタ統合部と、複数の画像データについてそれぞれのサムネイルをランダムに配置し、画像データの画像全体の特徴量、エッジ部分の特徴量、被写体を構成する画素の特徴量およびクラスタ統合部により生成された画像データのクラスタに基づいて、各サムネイルの座標を更新して表示するとともに、各サムネイルの移動量を計算し、各サムネイルの移動量が0に収束するまで、座標を更新して表示する処理を繰り返す表示部とを備える。

【0010】

このような第1の特徴に係る画像分類装置によれば、画像中の色とその分布に着目し、多段階に処理することで、高精度な画像の分類を実現することができる。具体的には、画像全体の特徴量に基づいて第1のクラスタリング部によるクラスタリングをした後、精度を上げるために、画像中の重要な特徴であるエッジに着目して、第2のクラスタリング部によるクラスタリング処理する。さらに、過分割による影響を抑制するため、画像データ中の被写体から得られた特徴量に基づいて、クラスタを統合する。本発明の第1の特徴に係る画像分類装置は、まず画像全体の特徴量に基づいて大まかに分類する。次に、画像を特徴付けるエッジ部分に注目して、更に階層的に詳細に分類する。更に、各分類により過分割されたクラスタを、被写体の特徴に基づいて、被写体ごとに統合する。最後に被写体に基づいてクラスタを統合することにより、ユーザの検索キーワードに合致しやすいクラスタを生成できる。

【0011】

ここで、複数の画像データについてそれぞれのサムネイルを任意に配置し、画像データの画像全体の特徴量、エッジ部分の特徴量、被写体を構成する画素の特徴量およびクラスタ統合部14により決定された画像データのクラスタに基づいて、各サムネイルの座標を更新して表示するとともに、各サムネイルの移動量を計算し、各サムネイルの移動量が0に収束するまで、座標を更新して表示する処理を繰り返す表示部を備えても良い。

【0012】

このような画像分類装置によれば、画像データベース中の画像データを直観的に把握することができるので、クエリ画像を保持していない場合や、明確な検索意図が存在しない場合においても、所望する画像データを検索することが可能となる。

【0013】

ここで、特徴量算出部は、複数の画像データのそれぞれについて、画像データにおける輝度値に基づいて、画像全体における色コリログラムを算出し、他の画像データの画像全

10

20

30

40

50

体との色コリログラムの距離を、画像全体の特徴量として算出するとともに、エッジ部分における色コリログラムを算出し、他の画像データのエッジ部分との色コリログラムの距離を、エッジ部分の特徴量として算出する。

【0014】

これによれば、画像データ中の色の量だけではなく、空間的な分布も考慮して、画像データの特徴量を算出することができる。

【0015】

また、第1のクラスタリング部は、画像全体における特徴量に基づいて、全てのクラスタによるクラスタリング誤差が閾値内になるように、クラスタの数と、クラスタに属する画像データを決定する。第2のクラスタリング部は、エッジ部分の特徴量に基づいて、第1のクラスタリング部によって分類された各クラスタのクラスタ内誤差が閾値内になるように、クラスタの数と、クラスタに属する画像データを決定する。

10

【0016】

このように、多段階にクラスタリングすることにより、高精度のクラスタリングを実行することができる。

【0017】

クラスタ統合部は、画像データに任意の境界線を設け、境界線によって得られる第1の領域と第2の領域における色ヒストグラム間の距離を算出し、境界線を動かすことによって色ヒストグラム間の距離が大きく変化する境界線を、構図を決定するための境界線とし、該境界線外を構成する画素の代表色との色差が閾値よりも大きな画素を、被写体を構成している画素とする。ここで、クラスタ統合部は更に、被写体を構成する画素の色コリログラムと色ヒストグラムを算出し、算出された色コリログラムと色ヒストグラムに基づいて、第2のクラスタリング部によって分類された複数のクラスタのうち、任意の2つのクラスタの非類似度が、閾値よりも高くなるまで、クラスタの統合を繰り返す。

20

【0018】

このように、被写体における特徴量が類似しているクラスタを統合することにより、過分割を抑制し、ユーザに把握させやすい画像分類を実現することができる。

【0019】

クエリ画像データの特徴ベクトルを算出するクエリ画像特徴量算出部と、クエリ画像データの特徴ベクトルに基づいて、クラスタ統合部により生成されたクラスタから、クエリ画像データの所属するクラスタを決定するクラスタ決定部とを備えても良い。

30

【0020】

このような画像分類装置によれば、複数の画像データから、類似する画像を抽出することができる。

【0021】

具体的には、特徴ベクトルは、画像全体における色コリログラムと、エッジ部分の色コリログラムと、被写体領域の色ヒストグラムと色コリログラムをパラメータとして有する。ここで、クエリ画像特徴量算出部は、クエリ画像データの画像全体における色コリログラムと、クエリ画像データのエッジ部分の色コリログラムと、クエリ画像データの被写体領域の色ヒストグラムと色コリログラムを算出して、クエリ画像データの特徴ベクトルを算出する。また、クラスタ決定部は、クラスタ統合部により生成された各クラスタに属する画像データの特徴ベクトルの平均を算出し、クエリ画像データの特徴ベクトルとの距離を最小とするクラスタをクエリ画像の所属クラスタとする。

40

【0022】

本発明の第2の特徴は、複数の画像データを、類似した画像毎に分類する画像分類プログラムに関する。即ち本発明の第2の特徴に係る画像分類プログラムは、コンピュータを、画像データベースに記憶された複数の画像データのそれぞれについて、画像全体の特徴量を算出するとともに、画像データのエッジを検出し、検出されたエッジ部分の特徴量を算出する特徴量算出手段と、画像全体における特徴量に基づいて、複数の画像データを複数のクラスタに分類する第1のクラスタリング手段と、エッジ部分の特徴量に基づいて、

50

第1のクラスタリング手段によって分類された複数のクラスタを、更に複数のクラスタに分類する第2のクラスタリング手段と、複数の画像データのそれぞれについて、被写体を構成する画素を決定し、第2のクラスタリング手段によって分類された複数のクラスタを、被写体を構成する画素に基づいて統合するクラスタ統合手段と、複数の画像データについてそれぞれのサムネイルをランダムに配置し、画像データの前記画像全体の特徴量、前記エッジ部分の特徴量、被写体を構成する画素の特徴量およびクラスタ統合手段により生成された画像データのクラスタに基づいて、各サムネイルの座標を更新して表示するとともに、各サムネイルの移動量を計算し、各サムネイルの移動量が0に収束するまで、座標を更新して表示する処理を繰り返す表示手段として機能させる。

【0023】

ここで、複数の画像データについてそれぞれのサムネイルを任意に配置し、画像データの画像全体の特徴量、エッジ部分の特徴量および被写体を構成する画素の特徴量に基づいて、各サムネイルの座標を更新して表示するとともに、各サムネイルの移動量を計算し、各サムネイルの移動量が0に収束するまで、座標を更新して表示する処理を繰り返す表示手段として、更にコンピュータを機能させても良い。

【0024】

特徴量算出手段は、複数の画像データのそれぞれについて、画像データにおける輝度値に基づいて、画像全体における色コリログラムを算出し、他の画像データの画像全体との色コリログラムの距離を、画像全体の特徴量として算出するとともに、エッジ部分における色コリログラムを算出し、他の画像データのエッジ部分との色コリログラムの距離を、エッジ部分の特徴量として算出しても良い。ここでは、色コリログラムに基づいて特徴量を算出する場合について説明するが、色コリログラムと色ヒストグラムを用いて特徴量を算出しても良い。

【0025】

第1のクラスタリング手段は、画像全体における特徴量に基づいて、全てのクラスタによるクラスタリング誤差が閾値内になるように、クラスタの数と、クラスタに属する画像データを決定しても良い。第2のクラスタリング手段は、エッジ部分の特徴量に基づいて、第1のクラスタリング手段によって分類された各クラスタのクラスタ内誤差が閾値内になるように、クラスタの数と、クラスタに属する画像データを決定しても良い。

【0026】

クラスタ統合手段は、画像データに任意の境界線を設け、境界線によって得られる第1の領域と第2の領域における色ヒストグラム間の距離を算出し、境界線を動かすことによって色ヒストグラム間の距離が大きく変化する境界線を、構図を決定するための境界線とし、該境界線外を構成する画素の代表色との色差が閾値よりも大きな画素を、被写体を構成している画素としても良い。ここで、クラスタ統合手段は更に、被写体を構成する画素の色コリログラムと色ヒストグラムを算出し、算出された色コリログラムと色ヒストグラムに基づいて、第2のクラスタリング手段によって分類された複数のクラスタのうち、任意の2つのクラスタの非類似度が、閾値よりも高くなるまで、クラスタの統合を繰り返しても良い。

【0027】

また、クエリ画像データの特徴ベクトルを算出するクエリ画像特徴量算出手段と、クエリ画像データの特徴ベクトルに基づいて、クラスタ統合手段により生成されたクラスタから、クエリ画像データの所属するクラスタを決定するクラスタ決定手段として、更にコンピュータを機能させても良い。

【0028】

ここで、特徴ベクトルは、画像全体における色コリログラムと、エッジ部分の色コリログラムと、被写体領域の色ヒストグラムと色コリログラムをパラメータとして有し、クエリ画像特徴量算出手段は、クエリ画像データの画像全体における色コリログラムと、クエリ画像データのエッジ部分の色コリログラムと、クエリ画像データの被写体領域の色ヒストグラムと色コリログラムを算出して、クエリ画像データの特徴ベクトルを算出しても良

10

20

30

40

50

い。

【0029】

クラスタ決定手段は、クラスタ統合手段により生成された各クラスタに属する画像データの特徴ベクトルの平均を算出し、クエリ画像データの特徴ベクトルとの距離を最小とするクラスタをクエリ画像の所属クラスタとしても良い。

【0030】

本発明によれば、複数の画像データを類似した画像毎に分類する画像分類装置および画像分類プログラムを提供することができる。

【図面の簡単な説明】

【0031】

【図1】図1は、本発明の最良の実施の形態に係る画像分類装置の機能ブロック図である。

【図2】図2は、本発明の最良の実施の形態に係る画像分類方法の概略を説明する図である。

【図3】図3は、本発明の最良の実施の形態に係る画像分類装置のハードウェア構成図である。

【図4】図4は、本発明の最良の実施の形態に係る画像分類装置の特徴量算出処理を説明するフローチャートである。

【図5A】図5Aは、本発明の最良の実施の形態に係る画像分類装置に用いられる色ヒストグラムと色コリログラムを説明する図である。(その1)

【図5B】図5Bは、本発明の最良の実施の形態に係る画像分類装置に用いられる色ヒストグラムと色コリログラムを説明する図である。(その2)

【図6】図6は、本発明の最良の実施の形態に係る画像分類装置において、第1のクラスタリング部と第2のクラスタリング部によるクラスタ分割を説明する図である。

【図7】図7は、本発明の最良の実施の形態に係る画像分類装置の第1のクラスタリング処理を説明するフローチャートである。

【図8】図8は、本発明の最良の実施の形態に係る画像分類装置の第2のクラスタリング処理を説明するフローチャートである。

【図9】図9は、本発明の最良の実施の形態に係る画像分類装置の第2のクラスタリング部において、Sobelフィルタによるエッジ強度の定義に用いられる画素の位置を説明する図である。

【図10】図10は、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合処理を説明するフローチャートである。

【図11A】図11Aは、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合部における構図を説明する図である。(その1)

【図11B】図11Bは、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合部における構図を説明する図である。(その2)

【図11C】図11Cは、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合部における構図を説明する図である。(その3)

【図11D】図11Dは、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合部における構図を説明する図である。(その4)

【図12A】図12Aは、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合部における境界線L1を説明する図である。

【図12B】図12Bは、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合部における境界線L2を説明する図である。

【図12C】図12Cは、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合部における境界線L3を説明する図である。

【図12D】図12Dは、本発明の最良の実施の形態に係る画像分類装置のクラスタ統合部における境界線L4を説明する図である。

【図13】図13は、本発明の最良の実施の形態に係る画像分類装置の表示処理を説明す

10

20

30

40

50

るフローチャートである。

【図14】図14は、本発明の最良の実施の形態に係る画像分類装置の表示部における透視変換を説明する図である。

【図15】図15は、本発明の最良の実施の形態に係る画像分類装置の表示部によって表示される初期画面の一例である。

【図16】図16は、本発明の最良の実施の形態に係る画像分類装置の表示部によって表示される結果画面の一例である。

【図17A】図17Aは、本発明の最良の実施の形態に係る画像分類装置の表示部におけるサムネイルの移動を説明する図である。(その1)

【図17B】図17Bは、本発明の最良の実施の形態に係る画像分類装置の表示部におけるサムネイルの移動を説明する図である。(その2)

【図17C】図17Cは、本発明の最良の実施の形態に係る画像分類装置の表示部におけるサムネイルの移動を説明する図である。(その3)

【図18】図18は、本発明の変形例に係る画像分類装置の機能ブロック図である。

【図19】図19は、本発明の変形例に係る画像分類方法の概略を説明する図である。

【図20】図20は、本発明の変形例に係る画像分類装置のクエリ画像特徴量算出処理を説明するフローチャートである。

【図21】図21は、本発明の変形例に係る画像分類装置の所属クラスタ決定処理を説明するフローチャートである。

【図22】図22は、本発明の変形例に係る画像分類装置の表示部によって表示される結果画面の一例である。

【発明を実施するための最良の形態】

【0032】

次に、図面を参照して、本発明の実施の形態を説明する。以下の図面の記載において、同一又は類似の部分には同一又は類似の符号を付している。

【0033】

(最良の実施の形態)

本発明の最良の実施の形態に係る画像分類装置1は、複数の画像データを、類似した画像毎に分類する。本発明の最良の実施の形態に係る画像分類装置1は、画像データベース51中に存在する画像データを類似した画像毎に分類し、得られる分類結果に基づいて画像を検索する。さらに、画像分類装置1は、分類、検索の結果を3次元空間上で可視化を行い、提示する。分類・検索結果を3次元空間上に配置することで、画像分類装置1は、空間的な距離により画像間の類似性を理解することが可能なユーザインターフェースを実現することができる。

【0034】

(画像分類装置の概要)

図2を参照して、本発明の最良の実施の形態に係る画像分類装置1の処理の概要を説明する。

【0035】

まず、ステップS1において、記憶装置107から、画像データベース51の画像データが読み出される。次に、ステップS2において、ステップS1で読み出された画像データベースの各画像データについて、特徴量が算出される。この特徴量は、画像全体の画素から算出した特徴量と、画像におけるエッジ部分の画素から算出した特徴量を含む。

【0036】

次にステップS3ないしステップS5において、画像分類装置1は、抽出された特徴量に基づいて多段階にクラスタリングする。このとき画像分類装置1は、色や構造などに基づいて、画像を分類する。具体的には、ステップS3において画像分類装置1は、画像全体の色分布に係る特徴量に基づいて、複数の画像データを、複数のクラスタに分割する。更にステップS4において画像分類装置1は、画像のエッジ部分の色分布に係る特徴量に基づいて、ステップS3で分類されたクラスタを、更に分割する。ステップS5において

10

20

30

40

50

画像分類装置 1 は、被写体部の色分布と、色の出現頻度に基づいて、ステップ S 4 において過剰に分割されたクラスタを統合する。

【 0 0 3 7 】

ステップ S 6 において画像分類装置 1 は、画像データベース 5 1 の各画像が分類された結果を、表示装置 1 0 5 に表示する。この際、各画像の特徴量と、ステップ S 3 ないしステップ S 5 の処理によるクラスタリング結果を用いて、画像分類装置 1 は、3次元空間上で分類結果を可視化する。

【 0 0 3 8 】

本発明の最良の実施の形態に係る画像分類装置 1 の分類部 1 0 は、画像中の色の分布に着目し、色コリログラムを用いて K-means 法によるクラスタリングをする (ステップ S 3)。このとき得られる分類結果について画像分類装置 1 は、画像中のさらに詳細な色分布を考慮するため、エッジ画素を対象として算出した色コリログラムを用い、各クラスタを分割する (ステップ S 4)。ここで、エッジ画素とは、Sobel フィルタで取得したエッジであって、被写体のエッジとは限らない。このとき、過分割の影響を抑制するため、画像分類装置 1 は、被写体から得られた色ヒストグラムと色コリログラムを用いてクラスタの再統合処理を加える (ステップ S 5)。これらの各処理において特徴量間の距離を算出する際には、色コリログラム間の 2 次形式距離が用いられる。これにより得られる各画像の特徴量と所属するクラスタ情報を用いて、画像分類装置 1 は、表示部 2 0 において、分類結果を 3 次元空間上で可視化する (ステップ S 6)。

ここで、ステップ S 3 およびステップ S 4 においては、色コリログラムのみを用いて処理する場合について説明したが、色ヒストグラムだけを用いても良いし、色コリログラムと色ヒストグラムを用いても良い

【 0 0 3 9 】

(画像分類装置のハードウェア構成)

図 3 に示すように、本発明の最良の実施の形態に係る画像分類装置 1 は、中央処理制御装置 1 0 1、ROM (Read Only Memory) 1 0 2、RAM (Random Access Memory) 1 0 3 及び入出力インタフェース 1 0 9 を備える。これらは、バス 1 1 0 を介して接続されている。入出力インタフェース 1 0 9 には、入力装置 1 0 4、表示装置 1 0 5、通信制御装置 1 0 6、記憶装置 1 0 7 及びリムーバブルディスク 1 0 8 が接続されている。

【 0 0 4 0 】

中央処理制御装置 1 0 1 は、入力装置 1 0 4 からの入力信号に基づいて ROM 1 0 2 から画像分類装置 1 を起動するためのブートプログラムを読み出して実行する。中央処理制御装置 1 0 1 は、更に記憶装置 1 0 7 に記憶されたオペレーティングシステムを読み出す。更に中央処理制御装置 1 0 1 は、入力装置 1 0 4 や通信制御装置 1 0 6 などの入力信号に基づいて、各種装置の制御を行ったり、RAM 1 0 3 や記憶装置 1 0 7 などに記憶されたプログラム及びデータを読み出して RAM 1 0 3 にロードする。また、中央処理制御装置 1 0 1 は、RAM 1 0 3 から読み出されたプログラムのコマンドに基づいて、データの計算又は加工など、後述する一連の処理を実現する処理装置である。

【 0 0 4 1 】

入力装置 1 0 4 は、操作者が各種の操作を入力するキーボード、マウスなどの入力デバイスにより構成されている。入力装置 1 0 4 は、操作者の操作に基づいて入力信号を作成し、入出力インタフェース 1 0 9 及びバス 1 1 0 を介して中央処理制御装置 1 0 1 に送信する。表示装置 1 0 5 は、CRT (Cathode Ray Tube) ディスプレイや液晶ディスプレイなどである。表示装置 1 0 5 は、中央処理制御装置 1 0 1 からバス 1 1 0 及び入出力インタフェース 1 0 9 を介して表示装置 1 0 5 において表示させる出力信号を受信し、例えば中央処理制御装置 1 0 1 の処理結果などを表示する装置である。通信制御装置 1 0 6 は、LAN カードやモデムなどの装置である。通信制御装置 1 0 6 は、画像分類装置 1 をインターネットや LAN などの通信ネットワークに接続する装置である。通信制御装置 1 0 6 を介して通信ネットワークと送受信したデータは入力信号又は出力信号として、入出力インタフェース 1 0 9 及びバス 1 1 0 を介して中央処理制御装置 1 0 1 に送受信される。

【 0 0 4 2 】

記憶装置 1 0 7 は半導体記憶装置や磁気ディスク装置である。記憶装置 1 0 7 には、中央処理制御装置 1 0 1 で実行されるプログラムやデータが記憶されている。リムーバブルディスク 1 0 8 は、光ディスクやフレキシブルディスクのことである。ディスクドライブによって読み書きされた信号は、入出力インタフェース 1 0 9 及びバス 1 1 0 を介して中央処理制御装置 1 0 1 に送受信される。

【 0 0 4 3 】

本発明の最良の実施の形態に係る画像分類装置 1 の記憶装置 1 0 7 には、画像分類プログラムが記憶されるとともに、画像データベース 5 1 が記憶される。又、画像分類プログラムが画像分類装置 1 の中央処理制御装置 1 0 1 に読み込まれ実行されることによって、分類部 1 0 および表示部 2 0 が、画像分類装置 1 に実装される。

10

【 0 0 4 4 】

(画像分類装置の機能ブロック)

図 1 に示すように、本発明の最良の実施の形態に係る画像分類装置 1 は、画像データベース 5 1、分類部 1 0 および表示部 2 0 を備える。分類部 1 0 は、特徴量算出部 1 1、第 1 のクラスタリング部 1 2、第 2 のクラスタリング部 1 3 およびクラスタ統合部 1 4 を備える。

【 0 0 4 5 】

画像データベース 5 1 には、複数の画像データが記憶される。この画像データベースに記憶される画像データは、本発明の最良の実施の形態に係る画像分類装置 1 によって分類される対象となる。画像データベース 5 1 には、各画像データについて、後述する処理によって算出される特徴量や、クラスタの識別子を含むクラスタの情報などが関連づけられても良い。

20

【 0 0 4 6 】

特徴量算出部 1 1 は、画像データベース 5 1 に記憶された複数の画像データのそれぞれについて、画像全体の特徴量を算出する。さらに特徴量算出部 1 1 は、画像データのエッジを検出し、検出されたエッジ部分の特徴量を算出する。ここで、特徴量算出部 1 1 は、複数の画像データのそれぞれについて、画像データにおける輝度値に基づいて、色コリログラムを算出する。特徴量算出部 1 1 は、他の画像データとの色コリログラムの距離を、特徴量として算出する。画像全体の画素から算出した特徴量は、画像におけるエッジ部分の画素から算出した特徴量と、画像全体の画素のうち、エッジ部分と排他的画素から算出した特徴量と、を加算することによって、算出されても良い。

30

ここで、本発明の特徴量算出部 1 1 は、各画像データの色コリログラムに基づいて算出する場合について説明するが、色ヒストグラムを用いても良いし、色ヒストグラムと色コリログラムの二つを用いても良い。

【 0 0 4 7 】

第 1 のクラスタリング部 1 2 は、画像全体における特徴量に基づいて、複数の画像データを複数のクラスタに分類する。第 1 のクラスタリング部 1 2 は、画像全体における特徴量に基づいて、全てのクラスタによるクラスタリング誤差が閾値内になるように、クラスタの数と、クラスタに属する画像データを決定する。

40

【 0 0 4 8 】

第 2 のクラスタリング部 1 3 は、エッジ部分の特徴量に基づいて、第 1 のクラスタリング部 1 2 によって分類された複数のクラスタを、更に複数のクラスタに分類する。第 2 のクラスタリング部 1 3 は、エッジ部分の特徴量に基づいて、第 1 のクラスタリング部 1 2 によって分類された各クラスタのクラスタ内誤差が閾値内になるように、クラスタの数と、クラスタに属する画像データを決定する。

【 0 0 4 9 】

クラスタ統合部 1 4 は、複数の画像データのそれぞれについて画像の構図から被写体を構成する画素を決定し、第 2 のクラスタリング部 1 3 によって分類された複数のクラスタを、被写体を構成する画素に基づいて統合する。クラスタ統合部 1 4 は、画像データに任

50

意の境界線を設け、境界線によって得られる第1の領域と第2の領域における色ヒストグラム間の距離を算出する。クラスタ統合部14は、境界線を動かすことによって色ヒストグラム間の距離が大きく変化する境界線を、構図を決定するための境界線とする。ここでクラスタ統合部14は、この境界線外を構成する画素の代表色との色差が閾値よりも大きな画素を、被写体を構成している画素とする。クラスタ統合部14は更に、被写体を構成する画素の色コリログラムと色ヒストグラムを算出する。クラスタ統合部14は、算出された色コリログラムと色ヒストグラムに基づいて、第2のクラスタリング部13によって分類された複数のクラスタのうち、任意の2つのクラスタの非類似度が、閾値よりも高くなるまで、クラスタの統合を繰り返す。

【0050】

表示部20は、分類部10に依って分類された画像データを、表示装置105に可視的に表示する。表示部20は、複数の画像データについてそれぞれのサムネイルを任意に配置する。更に表示部20は、画像全体の特徴量、エッジ部分の特徴量、被写体を構成する画素の特徴量およびクラスタ統合部14により決定された画像データのクラスタに基づいて、各サムネイルの座標を更新して表示するとともに、各サムネイルの移動量を計算する。表示部20は、各サムネイルの移動量が0に収束するまで、座標を更新して表示する処理を繰り返す。表示部20は、画像全体の色コリログラム、エッジ部分の色コリログラムおよび被写体領域における色コリログラムから、各画像データ間の距離を定義する。表示部20は、この各画像データ間の距離に基づいて、サムネイルを移動させながら、表示装置105に表示する。本発明の最良の実施の形態においては、色コリログラムに基づいて各画像データ間の距離を算出する場合について説明するが、色ヒストグラムを用いても良いし、色ヒストグラムと色コリログラムの二つを用いても良い。

【0051】

(特徴量算出部)

本発明の最良の実施の形態に係る画像分類装置1は、カラー画像を分類の対象とし、画像中の色に基づく特徴を用いて画像を自動的に分類する。画像中の色を表現する特徴として、一般に色ヒストグラムが用いられる。しかしながら、色ヒストグラムを用いた場合、画像中の色の空間的な分布については考慮することができない。そこで、本発明の最良の実施の形態に係る画像分類装置1は、画像中の色の分布について考慮することが可能な特徴として、色コリログラムに着目する。色ヒストグラムに代えて、この色コリログラムを用いることで、色の空間的な分布の差異が考慮され、高精度な画像の自動分類が実現される。本発明の最良の実施の形態においては、色コリログラムに基づいて分類する場合について説明するが、色ヒストグラムを用いても良いし、色ヒストグラムと色コリログラムの二つを用いても良い。

【0052】

図4を参照して、特徴量算出部11の処理について詳述する。図4に示す例では、色ヒストグラムおよび色コリログラムに基づいて特徴量を算出する場合について説明する。色コリログラムのみ用いて特徴量を算出する場合は、ステップS102ないしステップS104およびステップS108は割愛されても良い。

図4において、ステップS102ないしステップS104の処理は、各画像データの色ヒストグラムを算出する処理である。ステップS105ないしステップS107の処理は、各画像データの色コリログラムを算出する処理である。ステップS108およびステップS109は、ステップS104およびステップS107において取得された各画像データの色ヒストグラムおよび色コリログラムに基づいて、各画像データ間の色コリログラムおよび色ヒストグラムの距離を算出する。

【0053】

まず、特徴量算出部11は、ステップS101ないしステップS107において、画像データベース51に格納された各画像データについて、色ヒストグラムおよび色コリログラムを算出する。

【0054】

10

20

30

40

50

具体的には、ステップS 1 0 1において、特徴量算出部 1 1は、画像データの各画素の輝度値を量子化する。ここでは、輝度値を基準に以下の処理を実行する場合について説明するが、RGBの各画素の値に基づいて処理されても良い。次にステップS 1 0 2において、特徴量算出部 1 1は、ヒストグラムのビンに対して、画像データの各画素に対応する画素値の値に投票する。ステップS 1 0 3において、特徴量算出部 1 1は、ステップS 1 0 2で取得したヒストグラムを正規化し、総和を1とする。ステップS 1 0 4において、特徴量算出部 1 1は、ステップS 1 0 3により算出された値を、この画像データの色ヒストグラムとして取得する。

【0055】

次にステップS 1 0 5において、特徴量算出部 1 1は、コリログラムのビンに対して、画像データの各画素に対応する画素値の値に投票する。ステップS 1 0 6において、特徴量算出部 1 1は、ステップS 1 0 5で取得したコリログラムを正規化し、総和を1とする。ステップS 1 0 7において、特徴量算出部 1 1は、ステップS 1 0 6により算出された値を、この画像データの色コリログラムとして取得する。

【0056】

次に、ステップS 1 0 8において、特徴量算出部 1 1は、各画像データの色ヒストグラム間の2次形式距離を算出し、画像データベース5 1に記憶する。更に、ステップS 1 0 9において、特徴量算出部 1 1各画像データの色コリログラム間の2次形式距離を算出し、画像データベース5 1に記憶する。例えば、画像データベース5 1に、50の画像データが格納されている場合、色ヒストグラム間の2次元形式距離および色コリログラム間の2次元形式距離は、それぞれ50 × 49 / 2個である。

【0057】

以下で、色ヒストグラム、及び色コリログラムの定義について説明する。

1. 色ヒストグラム

色ヒストグラムは画像中に特定の色が出現する確率の分布により定義される。画像Iにおける各画素の色は c_1, \dots, c_m のm階調に量子化されているものとし、各画素 $p = (x, y)$ Iに対して $I(p)$ をその画素の色とする。また、 $I_c = \{p | I(p) = c\}$ とする。このとき、画像Iの色 c_i に対する色ヒストグラム $h_{c_i}(I)$ は、次式により定義される。

【数1】

$$h_{c_i}(I) = Pr[p \in I_{c_i}] \quad \dots (式1)$$

このとき、 $Pr[p \in I_{c_i}]$ は、画像Iにおいて色が c_i となる画素pの確率を示す。このようにして定義される色ヒストグラムを特徴として用いることで、画像中の色に着目した画像分類が可能となる。

【0058】

2. 色コリログラム

色ヒストグラムが画像中に特定の色が出現する確率の分布として定義されるのに対し、色コリログラムは画像中の一定距離離れた画素間における特定の色の共起確率の分布として定義される。このことから、図5 Aおよび図5 Bに示すように、色ヒストグラムと色コリログラムは異なる特徴を表す。具体的には、図5 Aに示すように、大きい円が一つだけある画像についても、図5 Bに示すように小さい円が多数含まれている画像についても、同じ色ヒストグラムを有する。しかし、色コリログラムで表すことにより、図5 Aおよび図5 Bの画像それぞれの特徴を示すことができる。

【0059】

以降、色コリログラムの定義について説明する。

画像Iにおける各画素の色は c_1, \dots, c_m のm階調に量子化されているものとし、各画素 $p = (x, y)$ Iに対して $I(p)$ をその画素の色とする。また、 $I_c = \{p | I(p) = c\}$ とする。このとき、画像Iの色 c_i, c_j 、距離kに対する色コリログラム

【数 2】

$$\gamma_{c_i, c_j}^{(k)}$$

は、次式により定義される。

【数 3】

$$\gamma_{c_i, c_j}^{(k)} = Pr [p_2 \in I_{c_j} \mid |p_1 - p_2| = k, p_1 \in I_{c_i}] \quad \dots (式 2)$$

10

ここで、式 2 における 2 画素間の距離 $|p_1 - p_2|$ は、次式により定義される。

【数 4】

$$|p_1 - p_2| = \max(|x_1 - x_2|, |y_1 - y_2|) \quad \dots (式 3)$$

ただし、最良の実施の形態に係る画像分類装置 1 では、画素の周辺の局所領域における色の分布を考慮するため、式 2 を次式に置き換える。

【数 5】

$$\gamma_{c_i, c_j}^{(k)} = Pr [p_2 \in I_{c_j} \mid |p_1 - p_2| \leq k, p_1 \in I_{c_i}] \quad \dots (式 4)$$

20

【0060】

このように、色コリログラムを特徴として用いることで、本発明の最良の実施の形態において画像分類装置 1 は、画像中の色の空間的な分布の差異を考慮することができる。また、画像分類装置 1 は、高精度に画像を自動的に分類することができる。

【0061】

本発明の最良の実施の形態において画像分類装置 1 は、以上の 2 つの特徴を分類対象画像より抽出し、各画像データに対するそれらの距離を定義する。さらに画像分類装置 1 は、定義された距離に基づき画像間の類似度を算出し、得られる類似度に基づいて画像を分類する。本発明の最良の実施の形態に係る画像分類装置 1 は、画像全体、エッジ、被写体と着目する領域を限定しながら特徴を算出し、多段階に処理することで、より高精度に画像を自動的に分類することができる。

30

【0062】

次に、特徴間の距離に基づく画像間の類似度の定義を説明する。

本発明の最良の実施の形態に係る画像分類装置 1 は、複数の画像データにおける、画像から算出した色ヒストグラム間、および色コリログラム間の距離を用いて、画像間の類似度を評価する。色ヒストグラム間の距離尺度としては様々なものが考えられるが、その中でも人間の知覚に近い距離尺度であることが報告されている 2 次形式距離を用いることで、適切な分類結果を得ることが可能となる。ただし、2 次形式距離は色ヒストグラム間の距離として定義されており、そのままでは色コリログラム間の距離尺度としては用いることができない。そこで、本発明の最良の実施の形態に係る画像分類装置 1 は、色ヒストグラム間の 2 次形式距離の概念を拡張することで、色コリログラム間の距離を定義する。

40

【0063】

2 つの色ヒストグラム間の 2 次形式距離は、各ピンの値を要素とするベクトル h_i, h_j を用いて次式により定義される。

【数 6】

$$D_h(h_i, h_j) = (h_i - h_j)^T S (h_i - h_j) \quad \dots (式 5)$$

式 5 における $S = [s_{xy}]$ は色類似度行列と呼ばれ、 x 番目と y 番目のピンに対応する色の

50

類似度として

【数 7】

$$s_{xy} = e^{-\alpha d_{xy}^2}$$

・・・(式6)

で定義される。ただし、 α は、正定数である。 d_{xy} は、 x 番目のピンと y 番目のピンに対応する色の L*a*b 表色系における色差である。本発明の最良の実施の形態に係る画像分類装置 1 は、色ヒストグラム間の距離として式 5 の 2 次形式距離を用いる。

【0064】

本発明の最良の実施の形態に係る画像分類装置 1 は、さらにこの色ヒストグラム間の 2 次形式距離の概念を拡張することで、色コリログラム間の距離尺度を定義する。色ヒストグラムの各ピンが単一の色に対応するのに対し、色コリログラムの各ピンは 2 つの色の組み合わせに対応する。従って式 5 における色類似度行列 S をそのまま用いることはできない。そこで、本発明の最良の実施の形態に係る画像分類装置 1 は、 $S = [s_{xy}]$ を次式に示すように変更する。

【数 8】

$$\tilde{S} = [\tilde{s}_{xy}] \quad \dots (式7)$$

【数 9】

$$\tilde{s}_{xy} = \min \left(e^{-\alpha(d_{x_1y_1}^2 + d_{x_2y_2}^2)}, e^{-\alpha(d_{x_1y_2}^2 + d_{x_2y_1}^2)} \right) \quad \dots (式8)$$

ただし、色コリログラムの x 番目のピンと y 番目のピンは、それぞれ色 x_1, x_2 , および色 y_1, y_2 の組み合わせに対応する。本発明の最良の実施の形態においては、色コリログラム間の距離は、各ピンの値を要素とするベクトル c_i, c_j と

【数 10】

$$\tilde{S} = [\tilde{s}_{xy}]$$

を用いて次式により定義される。

【数 11】

$$D_c(c_i, c_j) = (c_i - c_j)^T \tilde{S} (c_i - c_j) \quad \dots (式9)$$

【0065】

本発明の最良の実施の形態は、2 次形式距離を算出する際、以下のように色ヒストグラムを変換することで計算量を削減する。まず、 S の固有値分解が $S = U \Lambda U^T$ で表される。ただし、 Λ は S の固有値を降順に並べた対角行列、 U は対応する固有ベクトルを並べた行列である。ここで、 Λ の上位 l 個の固有値を並べた対角行列 Λ_l と対応する固有ベクトルを並べた行列 U_l により、 n 次元の色ヒストグラム h_i, h_j は、次式を用いて l 次元のベクトル

【数 12】

$$h'_i, h'_j$$

に変換される。

10

20

30

40

【数 1 3】

$$\begin{aligned} h'_i &= \Lambda_l^{1/2} U_l^T h_i \\ h'_j &= \Lambda_l^{1/2} U_l^T h_j \\ &\dots \text{(式 1 0)} \end{aligned}$$

式 1 0 より算出される

【数 1 4】

$$h'_i, h'_j \quad 10$$

を用い、 h_i, h_j の l 次元空間上での距離は、次式により定義される。

【数 1 5】

$$D_h^l(h_i, h_j) = (\tilde{h}_i - \tilde{h}_j)^T (\tilde{h}_i - \tilde{h}_j) \quad \dots \text{(式 1 1)}$$

このとき、式 5 における $D_h(h_i, h_j)$ は、 l 次元空間上の距離

【数 1 6】

$$D_h^l(h_i, h_j) \quad 20$$

により近似することが可能である。本発明の最良の実施の形態では、この低次元での 2 次形式距離を用いて類似画像を分類する。色コリログラム間の 2 次形式距離の算出の際についても、色ヒストグラム間の 2 次形式距離算出と同様に、式 9 を次式で近似することにより、計算量が削減される。

【数 1 7】

$$D_c^l(c_i, c_j) = (c'_i - c'_j)^T (c'_i - c'_j) \quad \dots \text{(式 1 2)}$$

ただし、

【数 1 8】

$$\begin{aligned} c'_i &= \tilde{\Lambda}_l^{1/2} \tilde{U}_l^T c_i \\ c'_j &= \tilde{\Lambda}_l^{1/2} \tilde{U}_l^T c_j \\ &\dots \text{(式 1 3)} \end{aligned}$$

である。

【0 0 6 6】

本発明の最良の実施の形態においては、画像分類装置 1 は、

【数 1 9】

$$h'_i, h'_j \quad 40$$

を特徴ベクトルとし、距離

【数 2 0】

$$D_h^l(h_i, h_j), D_c^l(c_i, c_j)$$

を用いて画像間の類似度を評価し、画像を分類する。

【0 0 6 7】

50

(クラスタ処理部)

第1のクラスタリング部12、第2のクラスタリング部13およびクラスタ統合部14について説明する。

本発明の最良の実施の形態は、画像中の色とその分布に着目し、多段階に処理することで、高精度な画像の自動分類を実現する。まずは画像中の色の分布に着目し、色コリログラム間の2次形式距離に基づくK-means法により、画像が分類される。具体的には、画像分類装置1は、図6に示すように、画像データベース51に記憶された画像データ群を、K-means法により分類し、クラスタC11、C12、C13およびC14を生成する。

【0068】

しかしながら、K-means法により得られる分類は、1枚の画像の特徴を1つの色コリログラムのみを用いて表現しているため、画像中のより詳細な特徴については考慮されていない。そのため、この時点では十分な分類精度が得られていない可能性がある。そこで、本発明の最良の実施の形態は、画像中のより詳細な特徴を用いて分類結果の高精度化を図ることで、より高精度な画像の自動分類を実現する。

【0069】

本発明の最良の実施の形態は、より詳細な特徴として、画像中の重要な特徴であるエッジに着目する。画像分類装置1は、画像よりエッジ画素の検出を行い、エッジ画素とその周囲の画素のみから算出された色コリログラムを用いて、既に得られている各クラスタを分割する。ここで、エッジ画素とは、Sobelフィルタで取得したエッジであって、被写体のエッジとは限らない。例えば画像分類装置1は、K-means法により分割されたクラスタのうち、クラスタ内誤差が所定の閾値 T_h より大きいクラスタを更に分割する。図6に示す例では、クラスタ内誤差が所定の閾値 T_h より大きいクラスタは、クラスタC11およびC13であるので、画像分類装置1は、クラスタC11を、それぞれクラスタC21およびC22に分割する。さらに画像分類装置1は、クラスタC13を、クラスタC23およびC24に分割する。この結果得られたクラスタC24のクラスタ内誤差が所定の閾値 T_h より大きいので、画像分類装置1は、さらにクラスタ24を分割し、クラスタC31およびクラスタC32に分割する。このように画像分類装置1は、各クラスタのクラスタ内誤差が所定の閾値 T_h より小さくなるように、クラスタを分割する。これにより本発明の最良の実施の形態において、画像中のさらに詳細な色分布を考慮し、分類結果の高精度化が可能となる。

【0070】

ただし、このときクラスタを分割することにより過分割が発生してしまう危険性がある。そこで、画像分類装置1は、過分割の影響を抑制するため、被写体から得られた色ヒストグラムと色コリログラムを用いたクラスタの再統合処理を加える。これにより、画像分類装置1は、高精度に類似画像を分類することができる。具体的には、画像分類装置1は、分割結果のクラスタC21、C22、C12、C23、C31、C32およびC14の7つのクラスタについて、同一の被写体を含むと考えられるクラスタを6つ以下のクラスタに統合する。

以降、各処理について詳細に述べる。

【0071】

(第1のクラスタリング部)

本発明の最良の実施の形態に係る第1のクラスタリング部12は、N枚の分類対象画像 f_i ($i = 1, \dots, N$) を複数の集合に分類するため、K-means法を用いてクラスタリングを行う。ここで、K-means法によるクラスタリングは、非階層型クラスタリング手法の一種である。

【0072】

図7を参照して、本発明の最良の実施の形態に係る第1のクラスタリング部12による第1のクラスタリング処理を説明する。

【0073】

ステップS201において、第1のクラスタリング部12は、最初にクラスタ数Kを1

10

20

30

40

50

に設定する。次に、ステップS202において、第1のクラスタリング部12は、クラスタ数をKにするクラスタリングを行う。このとき、ステップS203において、第1のクラスタリング部12は、全てのクラスタに対するクラスタリング誤差E(k)を算出する。

【0074】

ステップS204において、第1のクラスタリング部12は、クラスタ数がK=1の場合、ステップS205においてKをインクリメントする。更に、ステップS202に戻り、第1のクラスタリング部12は、クラスタ数をKにするクラスタリングが行われる。

【0075】

一方ステップS204において、クラスタ数がK=1でない場合、ステップS206において、第1のクラスタリング部12は、現在のクラスタ数におけるクラスタリング誤差E(k)と、現在のクラスタ数より一つ少ないクラスタ数におけるクラスタリング誤差E(k-1)との差が、所定の閾値以上である場合、ステップS205においてKをインクリメントする。更に、ステップS202に戻り、第1のクラスタリング部12は、クラスタ数をKにするクラスタリングを行う。

【0076】

一方、ステップS206において、所定の閾値以上でない場合、現在のクラスタ数の状態で、クラスタリング誤差が収束したとして、第1のクラスタリング部12は、第1のクラスタリング処理を終了する。

【0077】

ここで、第1のクラスタリング部12の処理を詳述する。

本発明の最良の実施の形態に係る第1のクラスタリング部12は、K-means法に用いる特徴ベクトルを生成する。第1のクラスタリング部12は、各 f_i について、R、G、Bの輝度値をそれぞれm階調に量子化し、距離 k_1 に対する色コリログラムを算出する。ここで、 k_1 は予め与えられる値である。得られた色コリログラムに対し、式13を適用することで1次元ベクトル c'_i を算出し、特徴ベクトルとする。以上の処理により得られた特徴ベクトルに対し、K-means法を適用する。ただし、K-means法はクラスタリング結果が初期クラスタ、およびクラスタ数により変化する。そこで、本発明の最良の実施の形態においては、クラスタリング結果を評価するため、次式により表されるクラスタリング誤差 E_k を用いる。

【数21】

$$E_K = \frac{1}{N} \sum_{k=1}^K \sum_{h'_i \in C_k} (h'_i - v_k)^T (h'_i - v_k) \quad \dots (式14)$$

ただし、 C_k 、及び v_k ($k = 1, \dots, K$, K はクラスタ数)は k 番目のクラスタとその中心ベクトルを表す。本発明の最良の実施の形態においては、クラスタ数 K を $K = 1, 2, \dots$ と変化させる。第1のクラスタリング部12は、各クラスタについて初期値をランダムに変更しながらK-means法をM回適用する。第1のクラスタリング部12は、 E_k の値が最小となる結果をクラスタ数 K におけるクラスタリング結果とする。

【0078】

クラスタリングの終了判定には、スクリープロット概念を用いる。スクリープロットは、主成分分析において主成分数を決定するために用いられ、第一主成分から順に固有値をプロットし、2点間の差分が小さくなる主成分数を採用するものである。本発明の最良の実施の形態においては、主成分数をクラスタ数、固有値を E_k に対応づけ、 $|E_k - E_{k-1}| < TH_{c1}$ を満たす場合に現在の状態を、最終的なクラスタリング結果とする。

【0079】

(第2のクラスタリング部)

K-means法を適用することにより得られた分類結果は、画像全体の色の分布に着目したものであるため、画像の詳細な特徴については考慮されていない。よって、異なる被写体

10

20

30

40

50

を撮像した画像が、同一のクラスタに含まれる可能性がある。そこで、画像分類装置 1 は、より詳細な特徴を用いて各クラスタを分割することで、分類結果の高精度化を図る。具体的には、画像分類装置 1 は、画像の詳細な特徴としてエッジに着目し、エッジ画素と、エッジ画素を中心とする周辺の画素との色の関係を考慮することで、より高精度な分類結果を得る。ここで、エッジ画素とは、Sobel フィルタで取得したエッジであって、被写体のエッジとは限らない。また、周辺の画素とは、エッジ画素の座標を (x, y) としたとき、エッジ画素からの距離 k_2 は、

【数 2 2】

$$\sqrt{(x-a)^2 + (y-b)^2} \leq k_2$$

10

を満たす座標 (a, b) の画素である。ここで k_2 は、予め与えられるパラメータである。

【0080】

図 8 を参照して、本発明の最良の実施の形態に係る第 2 のクラスタリング部 1 3 による第 2 のクラスタリング処理を説明する。

【0081】

まず、ステップ S 3 0 1 において、第 2 のクラスタリング部 1 3 は、画像データベース 5 1 に記憶された全ての画像データに対し、Sobel フィルタを適用し、各画像データにおけるエッジ画像を取得する。ステップ S 3 0 2 において、第 2 のクラスタリング部 1 3 は、各画像データについて、ステップ S 3 0 1 で取得したエッジ画像の色コリログラムを生成する。

20

【0082】

次に、ステップ S 3 0 3 において、第 2 のクラスタリング部 1 3 は、第 1 のクラスタリング部 1 2 によって生成された各クラスタにおいて、クラスタ内誤差 E を算出する。このクラスタ内誤差 E は、クラスタに属する画像データについて、ステップ S 3 0 2 で生成されたエッジ画像の色コリログラム間の 2 次形式距離に基づく。この各画像データのエッジ部分における色コリログラム間の 2 次元形式距離は、画像データベース 5 1 に記憶されることが好ましい。

【0083】

ステップ S 3 0 4 において、第 2 のクラスタリング部 1 3 は、全てのクラスタにおいて、クラスタ内誤差 E が、所定の閾値未満であるか否かを判定する。閾値未満でない場合、ステップ S 3 0 5 において、第 2 のクラスタリング部 1 3 は、クラスタ内誤差 E が閾値未満でないクラスタを抽出する。更に、第 2 のクラスタリング部 1 3 は、その抽出したクラスタを 2 つに分割する。この分割においては、エッジ画素の色コリログラム間の 2 次形式距離に基づいて、 $K = 2$ の K-means 法が用いられる。更に、ステップ S 3 0 3 に戻り、第 2 のクラスタリング部 1 3 は、ステップ S 3 0 5 において分割された後の各クラスタにおいて、クラスタ内誤差 E を算出する。第 2 のクラスタリング部 1 3 は、ステップ S 3 0 5 において、全てのクラスタのクラスタ内誤差 E が閾値未満であるかを判断する。

30

【0084】

全てのクラスタのクラスタ内誤差 E が、閾値未満になったと判断されるまで、ステップ S 3 0 3 ないしステップ S 3 0 5 の処理は繰り返される。全てのクラスタのクラスタ内誤差 E が、閾値未満になると、第 2 のクラスタリング処理は終了する。

40

【0085】

ここで、第 2 のクラスタリング部 1 3 の処理を詳述する。

まず、Sobel フィルタを用いて各画像におけるエッジ画素を検出する。このとき、画素の 3×3 近傍の画素値に図 9 に示す記号を与えると、Sobel フィルタによるエッジ強度 e_s は次式で定義される。

【数 2 3】

$$e_s = \sqrt{e_x^2 + e_y^2} \quad \dots (式 15)$$

$$e_x = |(a - c) + 2(d - f) + (g - i)| \quad \dots (式 16)$$

$$e_y = |(a - g) + 2(b - h) + (c - i)| \quad \dots (式 17)$$

10

しかしながら、本手法はカラー画像を分類対象としているため、式 16 および式 17 を次式に変更する。

【数 2 4】

$$e_x = ||a - c|| + 2||d - f|| + ||g - i|| \quad \dots (式 18)$$

$$e_y = ||a - g|| + 2||b - h|| + ||c - i|| \quad \dots (式 19)$$

ここで、例えば $||a - c||$ は a と c の L^*a^*b 色空間における色差を表す。

【0086】

20

本発明の最良の実施の形態においては、このように拡張した Sobel フィルタを用いてエッジ画素を検出する。これにより取得されるエッジに着目して K-means 法により得られた各クラスタを分割することで、分類結果の高精度化を図る。エッジに着目するため、前述の拡張された Sobel フィルタにより得られるエッジ画素から距離 k_2 ($k_2 < k_1$) に対する色コリログラムを作成する。その後、式 13 を用いて算出される特徴ベクトル

【数 2 5】

$$c'_i$$

を用いることで、エッジに着目した処理を可能とする。このようにして得られた特徴ベクトルを用い、各クラスタを K-means 法によりさらに 2 つのクラスタに分割する。

30

【0087】

以上の処理を、全てのクラスタについて、次式により得られるクラスタ内の誤差 $E(C_k)$ が $E(C_k) < TH_d$ を満たすまで行う。

【数 2 6】

$$E(C_k) = \frac{1}{n_k} \sum_{c'_i \in C_k} D_c(c'_i, v_k) \quad \dots (式 20)$$

40

ただし、 n_k は、クラスタ C_k に属する分類対象画像の数を表す。以上の処理により、各クラスタを画像の重要な特徴であるエッジに着目して分割することが可能となる。

【0088】

(クラスタ統合部)

以上の処理により、画像分類装置 1 は、K-means 法を用いて得た各クラスタを、画像の重要な特徴であるエッジに着目して分割することが可能となる。ただし、これに伴って過分割が行われ、同一の被写体を撮影した画像が複数のクラスタに分割される場合がある。そこで、本発明の最良の実施の形態は、分割された各クラスタを、被写体に着目して統合することで、最終的な分類結果を得る。

【0089】

50

図10を参照して、本発明の最良の実施の形態に係るクラスタ統合部14によるクラスタ統合処理を説明する。

まずクラスタ統合部14は、画像データベース51に記憶された全ての画像データについて、ステップS401ないしステップS404の処理を実行する。ステップS401ないしステップS404の処理では、クラスタ統合部14は、画像データの構図を決定する。

【0090】

画像データベース51に記憶された画像データの1つについて、ステップS401において、クラスタ統合部14は、任意の境界線を設定して、画像を2つの領域に分割する。次にステップS402において、クラスタ統合部14は、ステップS401で分割された2つの領域のそれぞれについて、各領域に位置する画素についての色ヒストグラムを算出する。

10

【0091】

次にステップS403において、クラスタ統合部14は、2つの領域における色ヒストグラム間の2次元形式距離に基づいて、ステップS401で設けた境界線が、画像の構図を示すかを判定する。

【0092】

また、ステップS401ないしステップS404の処理は、図12Aないし図12Dに示すように、境界線の変更を繰り返される。これによりクラスタ統合部14は、画像データを、境界線内の画素と、境界線外の画素とに分割することができる。

20

【0093】

全ての画像データについて、ステップS401ないしステップS404の処理が終了すると、クラスタ統合部14は、ステップS405を処理する。ステップS405において、クラスタ統合部14は、境界線内の画素と境界線外の画素との特性に基づいて、被写体を構成する画素を抽出する。

【0094】

更に、クラスタ統合部14は、ステップS405で抽出した被写体の画素に基づいて、類似する被写体の画素を有する画像が含まれるクラスタを、統合する。具体的にはクラスタ統合部14は、複数のクラスタのうち、被写体が類似するクラスタを、1つのクラスタに統合する。この結果得られた各画像データが所属するクラスタの情報は、画像データベース51に記憶されることが好ましい。

30

【0095】

ここで、クラスタ統合部14の処理を詳述する。

最良の実施の形態において、被写体を構成する画素を推定するにあたり、分類対象画像をそのおおまかな色の分布から図11Aおよび図11Bに示す2つの構図に大別する。

【0096】

構図1：画像中に被写体が存在する構図(図11A参照)。例えば、図11Cに示す、中央付近に被写体が撮像されている画像。

構図2：構図1以外の構図(図11B参照)。例えば、図11Dに示す、テクスチャ画像等、画像全体を被写体が占める画像。

40

【0097】

以下で、構図を推定するための特徴量を、色ヒストグラムを用いて定義する。本発明の最良の実施の形態においては、図12Aないし図12Dに示す境界線 L_i ($i = 1, \dots, 4$)を平行移動することにより、クラスタ統合部14は、画素数の比が $j : 1 - j$ である2つの領域 $A_{i,j,1}, A_{i,j,2}$ に、画像を分割する。このとき、 j を $1/d$ から $(d-1)/d$ まで $1/d$ ずつ変化させたときの、2つの領域における色ヒストグラム間の2次形式距離の変化は、画像の構図により固有の傾向を示す。そこで、本発明の最良の実施の形態においては、構図の決定に用いる特徴量は、 $h_{i,j,1}, h_{i,j,2}$ 間の2次形式距離を用いて次式により算出される

【数 27】

$$D_i^{max}, D_i^{min}$$

を定義する。

【数 28】

$$D_i^{max} = \max_j D_h^l(h_{i,j,1}, h_{i,j,2}) \quad \dots (式 21)$$

$$D_i^{min} = \min_j D_h^l(h_{i,j,1}, h_{i,j,2}) \quad \dots (式 22)$$

10

各構図について、2つの領域における色ヒストグラム間の2次形式距離の変化が示す傾向、及び画像の構図（構図1、あるいは構図2）の決定法が、以下に示される。

【0098】

構図1：構図1の画像を2つの領域に分割すると、それぞれの領域の色ヒストグラム間の2次形式距離は、境界線の位置により大きく変化する。そこで、

【数 29】

$$D_i^{max} - D_i^{min}$$

20

を最大とする*i*を求め、

【数 30】

$$\hat{i}$$

とする。このとき、

【数 31】

$$D_{\hat{i}}^{max} - D_{\hat{i}}^{min} > TH_{diff}$$

30

である場合にその画像の構図を構図1とする。

構図2：構図1以外の画像を構図2とする。

【0099】

本発明の最良の実施の形態は、構図1の画像から被写体を構成する画素を推定する。一般に、背景は画像の端に接するため、構図1を決定する際に得られた境界線

【数 32】

$$\hat{i}$$

40

を用い、画像の端に存在する領域である

【数 33】

$$A_{\hat{i}, \frac{1}{d}, 1}, \text{ 及び } A_{\hat{i}, \frac{d-1}{d}, 2}$$

内に含まれる画素のL*a*b表色系における輝度値のメディアン値を求め、その画像の背景を構成する画素の代表色とする。このとき得られた、背景を構成する画素の代表色との色差が、閾値 TH_{object} よりも大きな画素を、画像分類装置1は、被写体を構成している画素であると推定する。一方、構図2の画像については画像全体を1つの被写体とみなし、

50

画像分類装置 1 は、全画素を被写体を構成する画素であると推定する。

【 0 1 0 0 】

これにより得られる被写体を構成する画素に着目してクラスタを統合することで、クラスタの過分割を抑制し、クラスタ統合部 1 4 は、最終的な分類結果を得る。被写体に着目するため、前述のように被写体を構成すると推定された画素から距離 k_1 に対する色コリログラムを算出する。ところで、色コリログラムは、非常に類似した画像間で高い類似度を示すが、被写体の撮像方向等の影響を受けやすいことが知られている。そこで、最良の実施の形態は、被写体を構成する画素から色ヒストグラムを得、さらに式 1 0 より特徴ベクトル h'_i を算出し、色コリログラムによる特徴ベクトルと共に用いてクラスタを統合する。クラスタの統合には、階層的クラスタリングの一手法であるウォード法が用いられる。

10

【 0 1 0 1 】

ウォード法は、クラスタ C_k と C_l の非類似度 $S(C_k, C_l)$ を

【数 3 4】

$$S(C_k, C_l) = \frac{n_k n_l}{n_k + n_l} (v_k - v_l)^T (v_k - v_l) \quad \dots (式 2 3)$$

で定義し、最も非類似度の低い 2 つのクラスタを逐次統合する手法である。本発明の最良の実施の形態は、色ヒストグラムと色コリログラムを同時に考慮するため、式 2 3 を次式のように変更する。

20

【数 3 5】

$$S(C_k, C_l) = \frac{n_k n_l}{n_k + n_l} (v'_k - v'_l)^T A (v'_k - v'_l) \quad \dots (式 2 4)$$

ただし、

【数 3 6】

$$v'_k, v'_l$$

30

はそれぞれクラスタ C_k 、及びクラスタ C_l の色ヒストグラムにおけるクラスタ中心と色コリログラムにおけるクラスタ中心を縦に並べたベクトルであり、 $A = [a_{ij}]$ は

【数 3 7】

$$a_{ij} = \begin{cases} w1 & \text{if } i = j, i \leq l \\ w2 & \text{if } i = j, i > l \\ 0 & \text{if } i \neq j \end{cases} \quad \dots (式 2 5)$$

40

を満たす行列である。

本発明の最良の実施の形態においては、 $S(C_k, C_l)$ の最小値が、閾値 TH よりも高い値を示すまで統合処理を繰り返し行う。以上の処理を行うことで、被写体に着目した高精度な類似画像分類が実現可能となる。

【 0 1 0 2 】

クラスタ統合部 1 4 は、以上の処理により得られる各特徴量と、クラスタリング結果を画像ごとに保存しておき、可視化の際に使用する。ただし、保存しておく各特徴量は、可視化の際に特徴量間の距離を算出することを考慮し、式 1 0 および式 1 3 により変換を行った値とする。

【 0 1 0 3 】

50

(表示部)

分類部 10 の処理によって得られた画像の特徴量、及びその分類結果に基づき、本発明の最良の実施の形態に係る画像分類装置 1 は、画像の分類結果を可視化する。この画像分類結果の可視化は、画像を多次元空間上に配置することで実現可能となる。

【0104】

図 13 を参照して、本発明の最良の実施の形態に係る表示部 20 による表示処理が説明される。

まず、ステップ S501 において、表示部 20 は、画像データベース 51 に記憶された各画像データについて、画像全体の色コリログラム、エッジ領域における色コリログラムおよび被写体領域における色ヒストグラムと色コリログラムを算出する。この算出処理は、上述した分類部 10 における処理結果値が参照されることが好ましい。各特徴値が算出されると、表示部 20 は、ステップ S502 において、画像データ間の距離を定義する。

【0105】

次にステップ S503 において、表示部 20 は、画像データのサムネイルを、表示装置 105 の画面上に配置する。このとき表示部 20 は、ランダムな位置に画像データのサムネイルを表示する。次に、ステップ S504 において、表示部 20 は、ステップ S502 で定義された画像データ間の距離や、分類部 10 によって決定されたクラスタに基づいて、画像データのサムネイルの座標を更新する。

【0106】

ステップ S505 において、表示部 20 は、座標更新の収束を判定をする。収束していない場合、ステップ S504 に戻り、更に画像データのサムネイルの座標を更新する。

一方、ステップ S505 において、表示部 20 は、表示装置 105 の座標更新が収束したと判定する場合、表示処理を終了する。

【0107】

表示部 20 における、ステップ S502 の画像データ間の距離の定義を説明する。

まず、表示部 20 は、各画像データから特徴ベクトルを作成する。特徴ベクトルの要素としては、表示部 20 は、上述した画像全体の色コリログラム、エッジ領域における色コリログラム、被写体領域における色ヒストグラムと色コリログラムの特徴量を用いる。表示部 20 は、このようにして求められた画像 I_i の特徴ベクトル v_i から、式 26 に基づき画像間の距離 $D_I(i, j)$ を算出する。

【数 38】

$$D_I(i, j) = \begin{cases} (v_i - v_j)^T (v_i - v_j) \times w_{same} & \text{if } C_i = C_j \\ (v_i - v_j)^T (v_i - v_j) \times w_{other} & \text{otherwise} \end{cases}$$

・・・(式 26)

ただし、クラスタ C_i, C_j は I_i, I_j を満たすクラスタである。 w_{same}, w_{other} は画像が所属するクラスタに応じた重みである。

【0108】

このとき本発明の最良の実施の形態において表示部 20 は、重み w_{same}, w_{other} に対して、 $w_{same} < w_{other}$ となる値を割り当てることで同じクラスタに属する画像間の距離を短くする。これにより、分類結果において同じクラスタに所属するとされた画像同士は、近距離に配置される。本発明の最良の実施の形態において表示部 20 は、この画像間の距離に基づき画像のサムネイルの座標を更新するとともに、その収束過程をアニメーションで表示することにより、画像分類結果の可視化を実現する。

【0109】

画像データ間の距離に基づいて画像の座標を移動させていき、収束するまでの過程を可視化することで、ユーザが、分類結果を直感的に把握することが可能となる。本発明の最良の実施の形態において表示部 20 は、まず画像を空間上にランダムに配置し、それらが

10

20

30

40

50

互いに類似した画像毎に集まっていく様子をアニメーションにより可視化する。本発明の最良の実施の形態は、多次元空間として、人間が知覚可能な最大次元である3次元空間を選択する。具体的には、表示部20は、分類対象画像より算出された特徴の次元を3次元まで削減することで、画像の3次元空間での座標を決定し、3次元空間上に配置する。次に、これを2次元のスクリーン上で表現するため、表示部20は、図14に示す透視変換を行い、最終的な出力画像を得る。

【0110】

本発明の最良の実施の形態に係る表示部20は、以下のような手順により構成される。

手順1：初期配置

図15に示すように、表示部20は、各画像を3次元空間上にランダムに配置し、透視変換することで2次元のスクリーン上に表示する。この際、表示部20は、各画像を配置する座標には多面体を描画し、その各面に対象画像を描画する。これにより、表示部20が、画像の3次元上の位置をユーザに分かり易く提示することが可能となる。

10

【0111】

手順2：画像の移動

表示部20は、画像間の距離に基づき各画像の移動量を計算し、各画像の3次元空間上の座標を変更する。このとき、画像 l_i の3次元座標を p_i とすると、その移動量

【数39】

$$m_i^*$$

20

は、以下の式に基づき算出される。

【数40】

$$m_i^* = \arg \min_{|m_i| \leq \sqrt{3}} \sum_j \left| D_I(i, j) - D_I^{real}(i, j, m_i) \right| \quad \dots (式27)$$

$$D_I^{real}(i, j, m_i) = \|p_i + m_i - p_j\| \quad \dots (式28) \quad 30$$

ただし

【数41】

$$D_I^{real}(i, j)$$

は、画像 l_i の座標 p_i を m_i 移動させたときの、3次元空間上における画像 l_i 、 l_j 間の距離を示す。したがって、画像の移動量

【数42】

$$m_i^*$$

40

は、予め定義した画像間の距離 $D_I(i, j)$ と実際の3次元空間上の距離

【数43】

$$D_I^{real}(i, j)$$

の差が最小となるものとして求められる。このようにして求められた画像の移動量

【数 4 4】

$$m_i^*$$

に基づき、表示部 2 0 は、次式のように各画像の座標を更新する。

【数 4 5】

$$p_i \leftarrow p_i + m_i^* \quad \dots (式 2 9)$$

【0 1 1 2】

10

手順 3 : 繰り返し

表示部 2 0 は、手順 2 を、全ての画像の移動量

【数 4 6】

$$m_i^*$$

が 0 に収束するまで繰り返す。0 に収束した時点で、表示装置 1 0 5 には、図 1 6 に示す画面が表示される。

【0 1 1 3】

図 1 6 においては、もみじを示す絵や、建物を示す絵など、類似した画像のサムネイルが、近く配置されている。一方、類似していない画像のサムネイルが、遠く配置されている。

20

【0 1 1 4】

このような処理により、各画像データのサムネイル間の距離は、予め定義した距離 $D_i(i, j)$ に近づいていく。類似した画像データのサムネイルが、近距離に配置されることとなる。

【0 1 1 5】

ここで、上述した手順 2 および手順 3 の処理が詳述される。各画像データのサムネイルの座標は、以下に示す評価関数を最小とする方向に移動する。

【数 4 7】

30

$$f(\mathbf{m}) = \sum_j |D_i(i, j) - D_i^{real}(i, j, \mathbf{m})|$$

ここで、簡単のため各画像（画像データのサムネイル）の座標を 1 次元に限定した例が、図 1 7 A に示される。図 1 7 A に示すように、各画像間の実際の距離が “ 7 ”、“ 2 ” であり、予め定義された画像間の距離が “ 5 ”、“ 4 ” であることが確認される。これらの値を用いて、表示部 2 0 は、画像 I_2 の座標を移動させた時の評価関数を算出し、値が最小となる方向に画像を移動させる。

【0 1 1 6】

(1) 1 ステップ目

40

画像が左に 1 移動された場合

$$f(-1) = |5-6| + |4-3| = 2$$

画像が移動されない場合

$$f(0) = |5-7| + |4-2| = 4$$

画像が右に 1 移動された場合

$$f(1) = |5-8| + |4-1| = 6$$

以上より、画像を左に移動させたときに評価関数が、最小となることが分かる。したがって、この場合画像 I_2 は、左へ移動する。このときの各画像の位置が、図 1 7 B に示される。表示部 2 0 は、この処理を繰り返すことで、各画像を類似した画像の近傍に配置することができる。

50

【 0 1 1 7 】

(2) 2ステップ目

次のステップでは、評価関数は以下ようになる。

画像が左に1移動された場合

$$f(-1)=|5-5|+|4-4|=0$$

画像が移動されない場合

$$f(0)=|5-6|+|4-3|=2$$

画像が右に1移動された場合

$$f(1)=|5-7|+|4-2|=4$$

以上より、画像 I_2 は、左へ移動する。このときの各画像の位置が、図 1 7 C に示される

10

【 0 1 1 8 】

(3) 3ステップ目

次のステップにおける評価関数は、以下ようになる。

画像が左に1移動された場合

$$f(-1)=|5-4|+|4-5|=2$$

画像が移動されない場合

$$f(0)=|5-5|+|4-4|=0$$

画像が右に1移動された場合

$$f(1)=|5-6|+|4-3|=2$$

20

このとき、評価関数は画像を移動させないとき最小となる。この時点で評価関数は収束しているので、表示部 2 0 は、画像（画像データのサムネイル）の移動を終了する。以上の処理を全ての画像に関して、表示部 2 0 が評価関数が収束するまで行うことで、画像分類結果の可視化が実現する。

【 0 1 1 9 】

次に、図 1 5 および図 1 6 を説明する。

本発明の最良の実施の形態に係る表示部 2 0 は、まず、図 1 5 に示す初期画面 P 2 0 1 を、表示装置 1 0 5 に表示する。初期画面 P 2 0 1 は、分類過程表示部 2 0 1 を備えている。分類過程表示部 2 0 1 には、上述した手順 1 において、画像データのサムネイルをランダムに配置した状態が表示される。この初期画面 P 2 0 1 は、各画像データのサムネイルを 3 次元空間にランダムに配置した後に透視変換を施すことにより得られる画面である

30

【 0 1 2 0 】

さらに、手順 2 および手順 3 において、画像データのサムネイルが移動する様子が、表示画面 1 0 5 に表示される。この様に、自動分類の過程をアニメーションを用いて表示することで、ユーザがデータベースに含まれる画像を直観的に把握することができる。

手順 2 および手順 3 が繰り返された結果、移動量が収束すると、図 1 6 に示す結果画面 P 2 0 2 が、表示装置 1 0 5 に表示される。結果画面 P 2 0 2 は、分類結果表示部 2 0 2 を備える。分類結果表示部 2 0 2 において、類似した画像が近くに表示される。一方、類似しない画像が遠くに表示される。

40

【 0 1 2 1 】

図 1 5 および図 1 6 に示す画面には、画像データのサムネイルが移動する過程を可視化するための制御ボタンがいくつか設けられている。具体的には、下記のボタンが設けられる。

【 0 1 2 2 】

(1) 早送りボタン

このボタンがクリックされると、アニメーションの速度が速くなる。これにより、ユーザが最終的な収束結果を見たい場合には、表示部 2 0 は、それを迅速に提示することができる。

(2) 初期状態復元ボタン

50

このボタンがクリックされると、表示部 20 は、分類過程を初期状態に戻すことができる。

(3) マップ表示・非表示ボタン

分類画面右上部に現在位置と視点の方向を示すマップが用意されている。これにより、ユーザは画像を配置した 3 次元空間上を自由に移動することができる。このボタンがクリックされることで、マップの表示・非表示を切替えることが可能である。

(4) 視点変更ボタン

表示部 20 は、画像を配置した 3 次元空間上での移動、および視点の変更を、マウスのドラッグにより行うことができる。その他、ユーザが 4 つの視点変更ボタンをクリックすることによっても、表示部 20 は、予め定めた 4 つの位置・視点に変更することができる。

10

【0123】

図 15 に示す初期画面 P 201 において、ユーザは、分類開始・一時停止ボタンをクリックすることで、表示部 20 は、分類過程をアニメーションにより再現する。その後、最終的に図 16 に示す結果画面 P 202 のように分類対象画像群はそれぞれ類似した画像毎に集まり、表示部 20 は、分類結果を表示することができる。ユーザは、分類過程表示部 201 および分類結果表示部 202 において、マウスのドラッグにより自由に 3 次元空間上での移動、および視点の変更が可能である。ユーザは、自由な視点で分類過程・結果を閲覧することができる。これにより、ユーザは、データベース中の画像を直観的に把握することができる。また、クエリ画像を保持していない場合、あるいは明確な検索意図が存在しない場合においても、ユーザは、所望する画像を検索することが可能となる。

20

【0124】

また、表示部 20 は、特定の画像と同じクラスに所属する画像群を表示した結果を示しても良い。このとき、選択した画像データとともに、選択した画像データと同じクラスに所属する画像データも表示することが好ましい。ユーザは選択した画像データと同じクラスに所属する画像データから画像を選択できる。ユーザがこのうちの 1 枚をクリックすることで、ユーザは、新たに画像を選択することができる。

【0125】

このように、本発明の最良の実施の形態に係る画像分類装置 1 は、画像の自動分類における分類過程および分類結果をユーザに提示することで、データベース中に含まれる画像を直観的に把握することができる。これによりユーザがクエリ画像を保持していない場合、あるいは明確な検索意図が存在しない場合においても、画像分類装置 1 は、ユーザの所望する画像を提供することが可能となる。

30

【0126】

(効果)

本発明の最良の実施の形態に係る画像分類装置 1 は、分類結果を表示する際に 3 次元空間上において可視化する。これにより画像分類装置 1 は、分類結果を効率的に閲覧させることができる。具体的には、画像分類装置 1 は、画像間の非類似度に基づいて 3 次元空間上における距離を決定し、画像のサムネイルを表示装置 105 に配置する。これにより、データベース中の画像のサムネイルが類似画像ごとに集まった状態が、表示装置 105 に表示される。

40

【0127】

また、画像分類装置 1 は、非類似度に基づいて画像間の距離を決定している。従って、ある画像から距離が近ければ近いほど、類似した画像が、表示装置 105 上に配置される。一方、遠ければ遠いほど、似ていない画像が、表示装置 105 上に配置される。これにより分類結果が表示される際、ある画像の周辺においてごく狭い範囲を見た場合には、類似している画像の中でも、特に類似している画像が近傍に配置される。同様に、あまり類似していない画像が遠くに配置されることとなる。また、より多くの画像を含む範囲で見た場合にも、類似している画像が近傍に配置され、似ていない画像が遠くに配置されることとなる。データベース中のすべての画像についてこのような配置が行われるため、ユーザが分類結果を閲覧する際に、画像分類装置 1 は、直感的に理解しやすく、効率的に閲覧

50

させることができる。

【0128】

また、本発明の最良の実施の形態に係る画像分類装置1は、段階的にクラスタリング処理をする。これにより画像分類装置1は、画像全体の色の分布だけでなく、画像中のより詳細な特徴を考慮して画像を分類することができる。具体的には、画像全体の色分布に基づいてクラスタリングすることで、画像分類装置1は、画像を、全体的に類似しているものにクラスタリングすることができる。次に、エッジ領域における色分布に基づいてクラスタを分割することで、画像分類装置1は、より高精度にクラスタリングすることができる。さらに、画像中に存在する被写体を推定し、被写体の特徴に基づいてクラスタを再統合することで、画像分類装置1は、画像中のもっとも重要な情報である被写体を考慮して、画像を分類することができる。また、クラスタリングの各段階において、色ヒストグラム、色コリログラム間の距離を算出する際に、距離尺度として2次形式距離を用いることで、画像分類装置1は、より人間の視覚特性に近い分類を実現する。

10

このように、本発明の実施の形態に係る画像分類装置1は、第1のクラスタリング部12により、まず画像全体の特徴量に基づいて大まかに分類する。次に画像分類装置1は、第2のクラスタリング部13により、画像を特徴付けるエッジ部分に注目して、更に階層的に詳細に分類する。更に、第1のクラスタリング部12および第2のクラスタリング部13により過分割されたクラスタを、クラスタ統合部14は、被写体の特徴に基づいて、被写体ごとにクラスタを統合する。最後に被写体に基づいてクラスタを統合することにより、画像分類装置1は、ユーザの検索キーワードに合致しやすいクラスタを生成できる。

20

【0129】

(変形例)

本発明の変形例に係る画像分類装置1aは、最良の実施の形態に係る分類機能を利用することで、クエリ画像を与えた際に自動的に画像を検索することができる。

【0130】

本発明の変形例に係る画像分類装置1aは、まず入力されたクエリ画像から特徴量を抽出する。本発明の変形例に係る画像分類装置1aのクエリ画像処理部30は、抽出される特徴量を用いてあらかじめ特徴量の算出と分類が完了しているデータベース中の画像と比較し、クエリ画像が所属するクラスタを決定する。このクエリ画像の特徴量とクラスタリング結果をデータベース中の画像と合わせて用いることで、表示部20は、3次元空間上で検索結果を可視化する。

30

【0131】

本発明の変形例に係る画像分類装置1aは、入力されたクエリ画像から特徴量を抽出し、画像データベース51中の画像と比較することで、クエリ画像に類似した画像を検索する。本発明の変形例に係る画像分類装置1aは、入力されたクエリ画像から、画像を分類する際に使用したすべての特徴量を抽出する。そして、事前に特徴量の抽出および分類を完了しているデータベース中の画像と比較を行い、クエリ画像が所属するクラスタを決定する。これにより得られるクエリ画像の特徴量および所属するクラスタを、画像データベース51中の画像のデータと合わせて用いることで、表示部20は、検索結果を3次元空間上で可視化する。

40

【0132】

図18を参照して、本発明の変形例に係る画像分類装置1aを説明する。

本発明の変形例に係る画像分類装置1aは、図1に示す本発明の最良の実施の形態に係る画像分類装置1と比べて、クエリ画像分類部30を備えている点が異なる。クエリ画像分類部30は、クエリ画像特徴量算出部31と、所属クラスタ決定部32を備えている。

【0133】

クエリ画像特徴量算出部31は、クエリ画像データ52の特徴ベクトルを算出する。ここで、特徴ベクトルは、画像全体における色コリログラムと、エッジ部分の色コリログラムと、被写体領域の色ヒストグラムと色コリログラムをパラメータとして有するベクトルである。クエリ画像特徴量算出部31は、クエリ画像データ52の画像全体における色コ

50

リログラムと、クエリ画像データ52のエッジ部分の色コリログラムと、52クエリ画像データの被写体領域の色ヒストグラムと色コリログラムを算出して、クエリ画像データ52の特徴ベクトルを算出する。

【0134】

所属クラスタ決定部32は、クエリ画像データ52の特徴ベクトルに基づいて、クラスタ統合部14により生成されたクラスタから、クエリ画像データ52の所属するクラスタを決定する。所属クラスタ決定部32は、クラスタ統合部14により生成された各クラスタに属する画像データの特徴ベクトルの平均を算出し、クエリ画像データ52の特徴ベクトルとの距離を最小とするクラスタをクエリ画像の所属クラスタとする。

【0135】

図19を参照して、本発明の変形例に係る画像分類装置1aにおける画像分類処理の概略を説明する。

まずステップS21において、画像分類装置1aは、クエリ画像データ52の特徴ベクトルを算出する。更にステップS22において画像データベース51から各画像データを読み出し、各画像データにおける特徴ベクトルを取得する。

【0136】

更にステップS23において画像分類装置1aは、クエリ画像データ52と、各画像データとの特徴ベクトルの距離を算出する。ステップS24において画像分類装置1aは、クエリ画像データ52が所属するクラスタを決定する。このとき、画像分類装置1aは、ステップS23で算出した距離が最小となるクラスタを、クエリ画像データ52の所属クラスタとする。

【0137】

更にステップS25において画像分類装置1aは、クエリ画像データ52と、画像データベース51に記憶された各画像データと、が分類される過程を、表示装置105に表示する。

【0138】

(クエリ画像特徴量算出部)

図20を参照して、本発明の変形例に係るクエリ画像特徴量算出部31が説明される。

まず、ステップS601において、クエリ画像特徴量算出部31は、記憶装置107からクエリ画像データ52を読み出し、ステップS602において、クエリ画像データ52の画像全体の色コリログラムを算出する。次にステップS603において、クエリ画像特徴量算出部31は、クエリ画像データ52のエッジ領域における色コリログラムを算出する。さらにステップS604において、クエリ画像特徴量算出部31は、被写体領域における色コリログラムを算出する。

【0139】

ステップS605において、クエリ画像特徴量算出部31は、ステップS602ないしステップS604の出力から、特徴ベクトルを生成する。

【0140】

クエリ画像特徴量算出部31は、抽出する特徴量として、分類部10において画像を自動的に分類する際に使用した特徴量を用いる。具体的には、クエリ画像特徴量算出部31は、画像全体の色コリログラム、エッジ領域における色コリログラム、被写体領域における色ヒストグラムと色コリログラムを算出する。得られた色ヒストグラム、色コリログラムによる特徴量は、2次形式距離による類似度を算出するため、式10、式13により変換される。クエリ画像特徴量算出部31は、これにより得られる値を要素とするベクトルを作成し、これを特徴ベクトル v_{query} とする。

【0141】

(所属クラスタ決定部)

図21を参照して、本発明の変形例に係る所属クラスタ決定部32を説明する。

まず、ステップS701において、所属クラスタ決定部32は、画像データベース51の各画像データから、特徴ベクトルを生成する。更にステップS702において、所属ク

10

20

30

40

50

ラスタ決定部 3 2 は、クラスタ統合部 1 4 によって生成された各クラスタの特徴ベクトルの平均ベクトルを算出する。

【 0 1 4 2 】

次にステップ S 7 0 3 において、所属クラスタ決定部 3 2 は、クエリ画像データ 5 2 の特徴ベクトルと、ステップ S 7 0 2 で算出した各クラスタの特徴ベクトルの平均との距離を算出する。ステップ S 7 0 4 において、所属クラスタ決定部 3 2 は、ステップ S 7 0 3 において算出した特徴ベクトル間の距離を最小とするクラスタを、クエリ画像データ 5 2 の所属するクラスタとする。

【 0 1 4 3 】

所属クラスタ決定部 3 2 は、画像データベース 5 1 中の各画像データに対して、あらかじめ算出されている特徴量を用いて、クエリ画像データ 5 2 の特徴ベクトルと同様の特徴ベクトルを作成する。得られる特徴ベクトルを用いて、各クラスタの特徴ベクトルの平均ベクトルを算出し、得られるベクトルを v_i とする。クエリ画像の特徴ベクトル v_{query} と各クラスタの平均特徴ベクトル v_i 間において、次式を用いて距離を算出する。

10

【 0 1 4 4 】

以上の処理により得られるクエリ画像の特徴量と所属するクラスタが、画像データベース 5 1 中の画像の特徴量とクラスタと合わせて保存され、可視化の際に使用される。

【 0 1 4 5 】

(表示部)

本発明の変形例に係る表示部 2 0 a は、クエリ画像分類部 3 0 において得られたクエリ画像データ 5 2 の特徴量と所属するクラスタの情報を画像データベース 5 1 に追加する。さらに表示部 2 0 a は、本発明の最良の実施の形態で説明した画像分類装置 1 と同様に可視化する。

20

【 0 1 4 6 】

本発明の最良の実施の形態で説明した画像分類装置 1 の表示部 2 0 は、画像が互いに類似するクラスタごとに集まっていく過程がアニメーションにより表示された。一方、本発明の変形例に係る画像分類装置 1 a の表示部 2 0 a においては、クエリ画像データ 5 2 を入力するため、クエリ画像データ 5 2 に注目して可視化する。具体的には、表示部 2 0 a は、図 2 2 に示すように透視変換の際に常に画面中央にクエリ画像が表示されるようにカメラパラメータを更新する。これにより、画面上ではクエリ画像データ 5 2 に類似した画像が集まる過程が、可視化されることとなる。

30

【 0 1 4 7 】

表示部 2 0 a は、例えば、図 2 2 に示す結果画面 P 2 0 3 を表示する。この結果画面 P 2 0 3 は、本発明の変形例に係る画像分類装置 1 a において、クエリ画像データ 5 2 に類似する画像データを抽出した図である。クエリ画像表示部 2 0 4 には、クエリ画像データ 5 2 のサムネイルが表示されている。分類結果表示部 2 0 3 の中心には、クエリ画像データ 5 2 のサムネイルが表示され、分類結果表示部 2 0 3 においては、クエリ画像データ 5 2 に類似する画像のサムネイルが近くに、類似しない画像のサムネイルが遠くに表示されている。類似画像表示部 2 0 5 には、分類結果表示部 2 0 3 において表示された画像データのうち、クエリ画像データ 5 2 と同じクラスタに属する画像データが、類似画像として表示されている。このとき、同じクラスタに属する画像データのうち、特徴ベクトルの近い画像データのみを表示しても良い。

40

【 0 1 4 8 】

以上により、クエリ画像データに類似した画像を検索することが可能となる。本発明の変形例に係る画像分類装置 1 a は、3次元空間上においてクエリ画像データ 5 2 の周囲に類似した画像を配置することで、ユーザが所望の画像を直感的に検索することを実現した。

【 0 1 4 9 】

(効果)

本発明の変形例に係る画像分類装置 1 a は、検索結果を表示する際に3次元空間上にお

50

いて可視化することで、検索結果を効率的に閲覧させることができる。具体的には、画像間の非類似度を3次元空間上における距離とすることで、画像分類装置1 aは、ユーザに検索結果を直感的に理解させることができる。さらに、3次元空間上における距離を、画像の非類似度に基づいて設定することで、画像分類装置1 aは、検索結果として表示されるクエリ画像以外の画像においても類似した画像が近傍に表示されるため、ユーザに、効率的に検索結果を閲覧をさせることが可能となる。

【0150】

また、本発明の変形例に係る画像分類装置1 aは、検索する際に検索対象データベース中の画像におけるクラスタリング結果を利用することで、少ない計算量での効率的な検索を可能とする。具体的には、検索する際には、まず検索対象の画像データベース5 1中の画像において、最良の実施の形態において説明した方法で、特徴量の算出とクラスタリングが完了しているものとする。そして、クエリ画像データ5 2が入力された際に、画像分類装置1 aは、画像データベース4 1中のクラスタリング結果を利用してクエリ画像のクラスタリングをすることで、効率的な処理を可能とする。

10

【0151】

(その他の実施の形態)

上記のように、本発明の最良の実施の形態および変形例によって記載したが、この開示の一部をなす論述及び図面はこの発明を限定するものであると理解すべきではない。この開示から当業者には様々な代替実施の形態、実施例及び運用技術が明らかとなる。

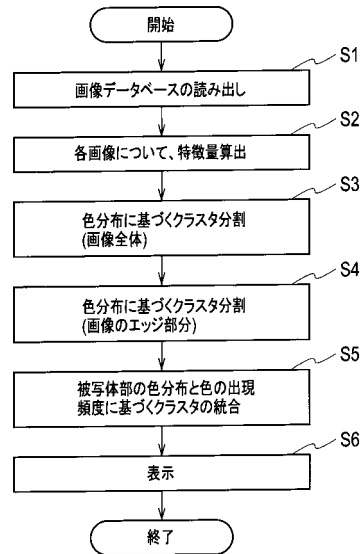
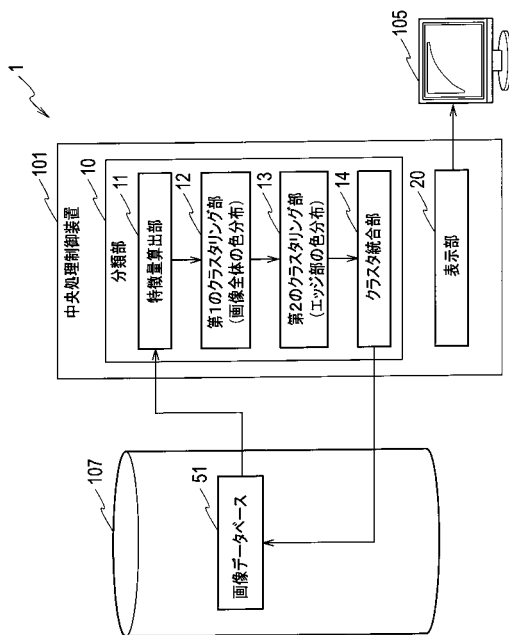
例えば、本発明の最良の実施の形態に記載した画像分類装置は、図1に示すように一つのハードウェア上に構成されても良いし、その機能や処理数に応じて複数のハードウェア上に構成されても良い。又、既存の情報システム上に実現されても良い。

20

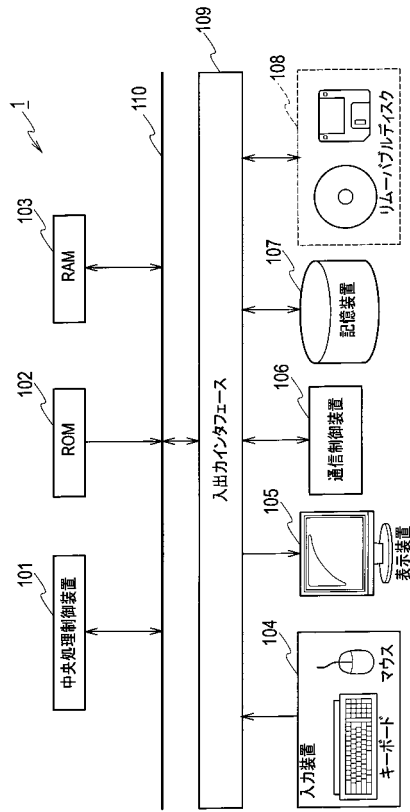
本発明はここでは記載していない様々な実施の形態等を含むことは勿論である。従って、本発明の技術的範囲は上記の説明から妥当な特許請求の範囲に係る発明特定事項によってのみ定められるものである。

【図1】

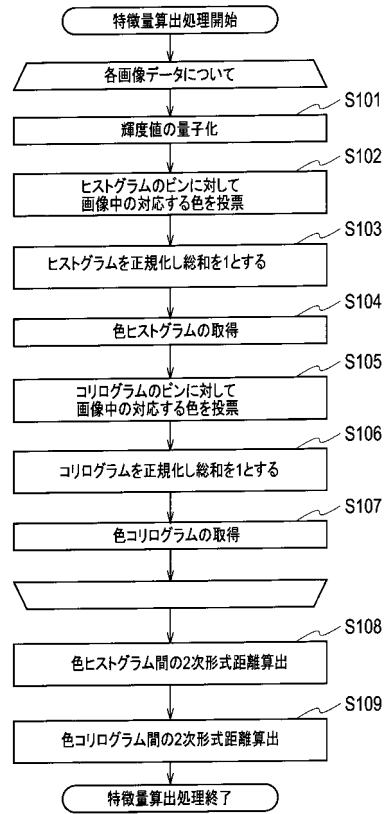
【図2】



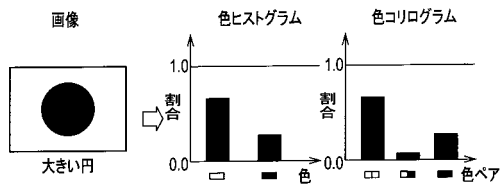
【図3】



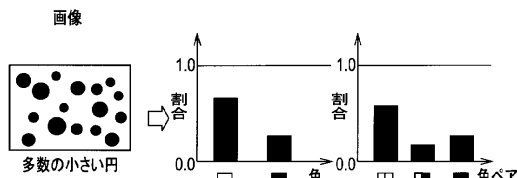
【図4】



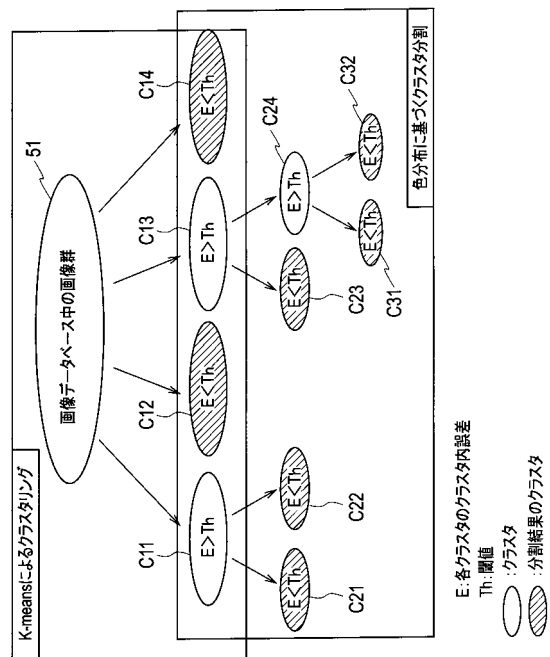
【図5A】



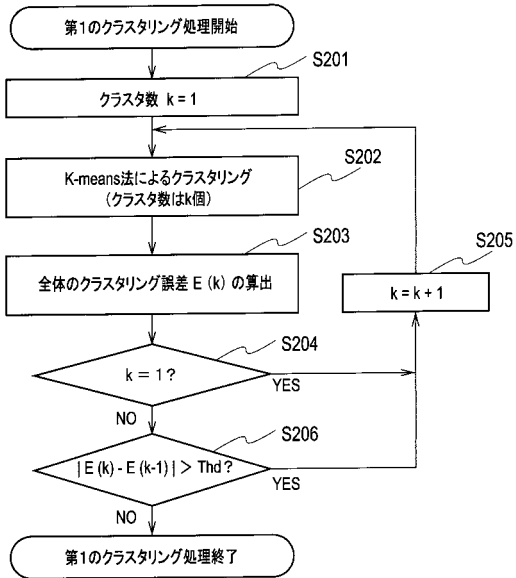
【図5B】



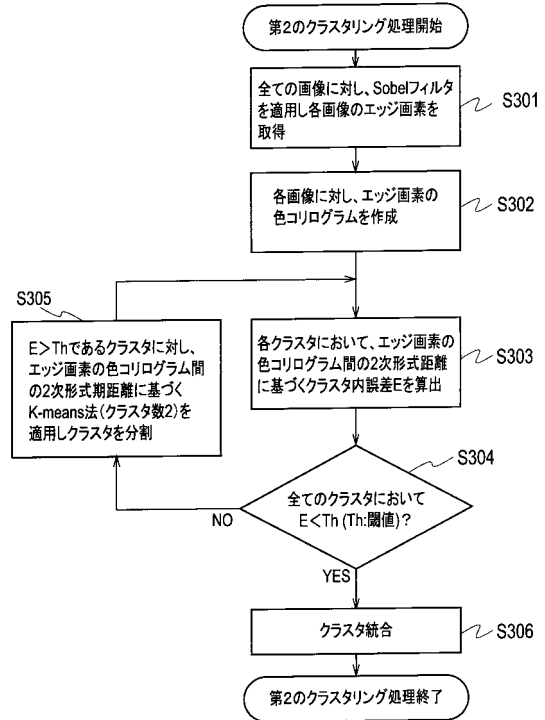
【図6】



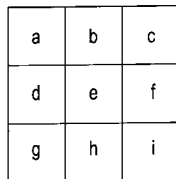
【 図 7 】



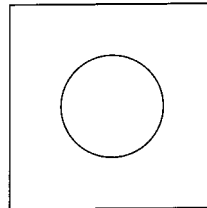
【 図 8 】



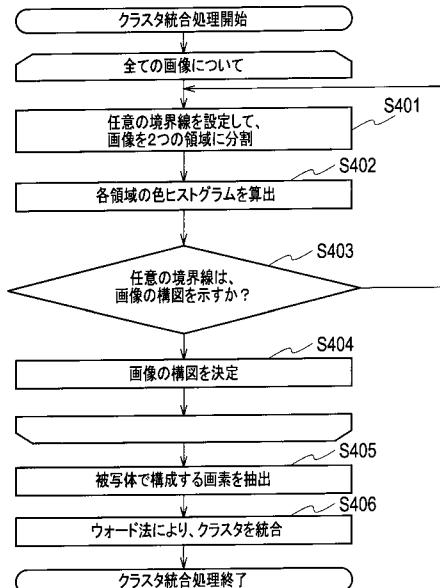
【 図 9 】



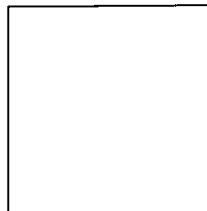
【 図 1 1 A 】



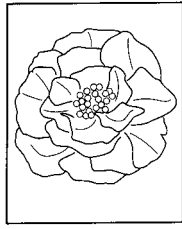
【 図 1 0 】



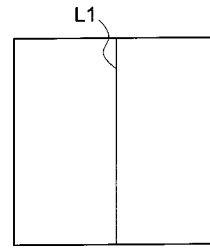
【 図 1 1 B 】



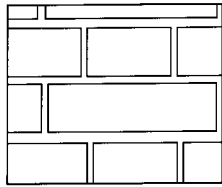
【図11C】



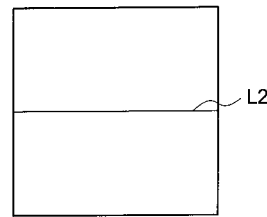
【図12A】



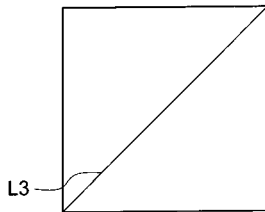
【図11D】



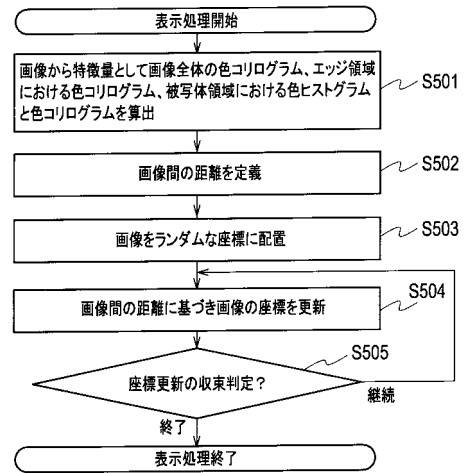
【図12B】



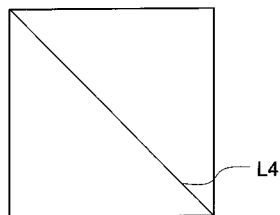
【図12C】



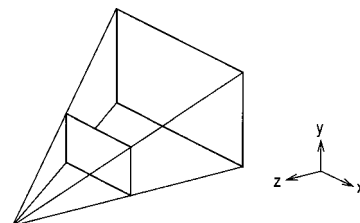
【図13】



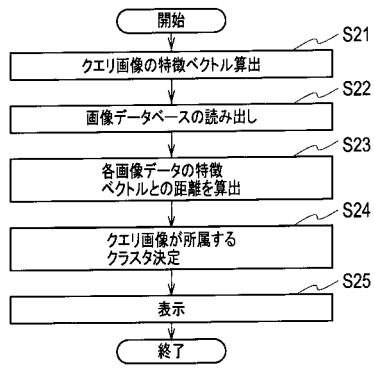
【図12D】



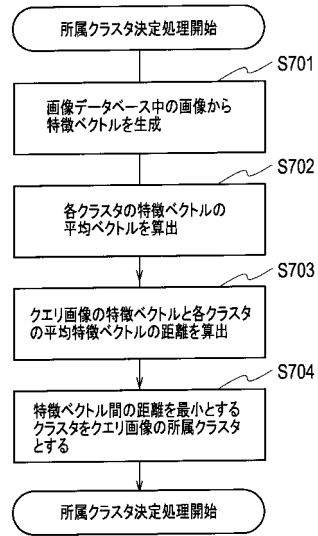
【図14】



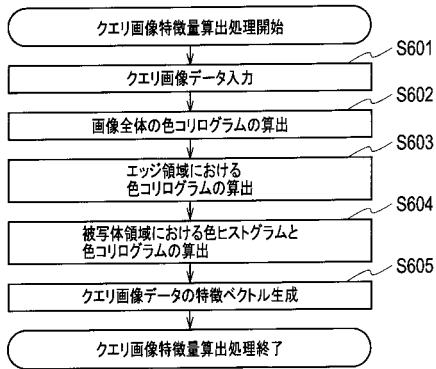
【図19】



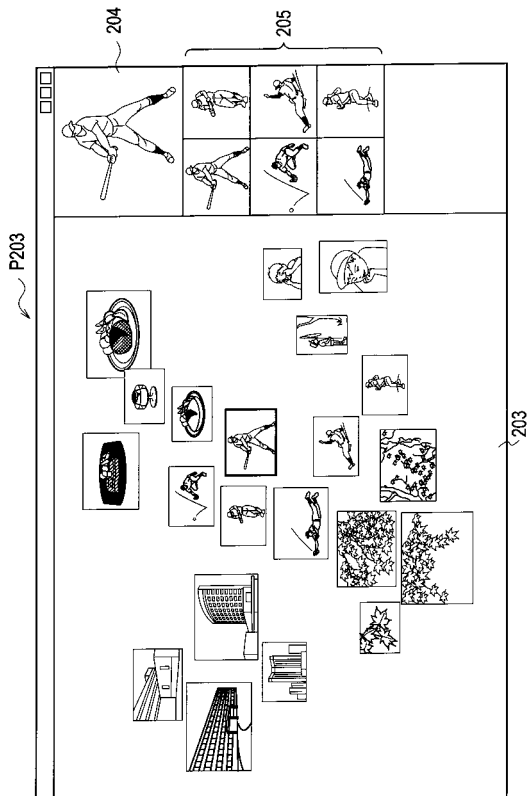
【図21】



【図20】



【図22】



フロントページの続き

(56)参考文献 特開2006-344155(JP,A)

渡辺隆志ほか, エッジを考慮した類似画像分類の高精度化に関する考察, 映像情報メディア学会技術報告, 日本, (社)映像情報メディア学会, 2007年 2月22日, 第31巻 第10号, 7-10頁

堀田政二ほか, 次元圧縮とクラスタリングに基づく画像の近似kNN検索, 映像情報メディア学会誌, 日本, (社)映像情報メディア学会, 2001年12月 1日, 第55巻 第12号, 1665-1668頁

(58)調査した分野(Int.Cl., DB名)

G06F 17/30

G06T 1/00