

(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2010-3244

(P2010-3244A)

(43) 公開日 平成22年1月7日(2010.1.7)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 1/00 (2006.01)	G06T 1/00 200D	5B050
G06T 7/00 (2006.01)	G06T 7/00 300G	5B075
G06F 17/30 (2006.01)	G06F 17/30 350C	5L096
	G06F 17/30 170B	

審査請求 未請求 請求項の数 9 O L (全 19 頁)

(21) 出願番号	特願2008-163570 (P2008-163570)	(71) 出願人	000006747
(22) 出願日	平成20年6月23日 (2008.6.23)		株式会社リコー
			東京都大田区中馬込1丁目3番6号
		(74) 代理人	100110607
			弁理士 間山 進也
		(72) 発明者	大黒 慶久
			東京都大田区中馬込1丁目3番6号 株式
			会社リコー内
		Fターム(参考)	5B050 DA04 EA08 GA08
			5B075 ND06 QM08
			5L096 AA02 AA03 AA06 BA08 FA36
			FA37 HA08

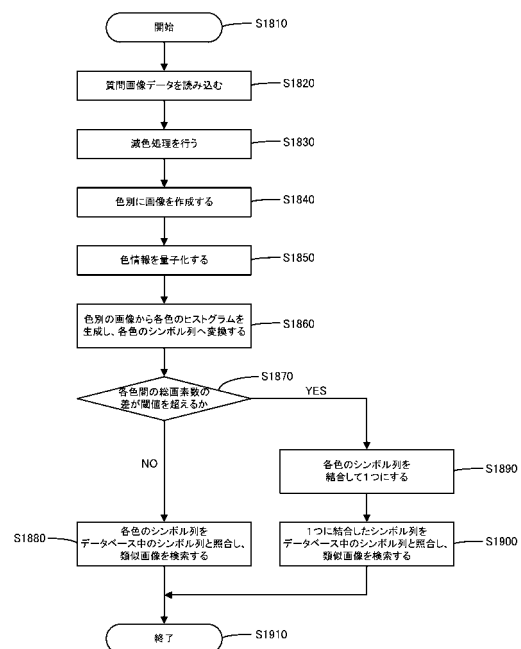
(54) 【発明の名称】 画像処理装置、画像処理方法およびプログラム

(57) 【要約】

【課題】 照合する画像間の分布形状を考慮した類似性を色別に求めることを可能とし、高精度で、かつ画像変形に対しても安定的に画像照合することができる画像処理装置および方法を提供する。

【解決手段】 この画像処理装置および方法では、行または列ごとに同じ色の画素を計数し、各色につきヒストグラムを生成し、ヒストグラムにおいて計数された各画素数を予め設定された画素数範囲に応じて量子化し、量子化した各値に対してシンボルを割り当て、ヒストグラムを複数のシンボルからなるシンボル列へ変換し、各色につき得られたシンボル列を結合し、各画像につき得られた、結合されたシンボル列を照合し、共通するシンボルが指定数以上ある場合に、両画像が類似すると評価する。

【選択図】 図18



【特許請求の範囲】**【請求項 1】**

行および列方向に複数の画素が配列してなる画像間の類似評価を行う画像処理装置であって、

前記行または列ごとに同じ色または同じ階調の画素を計数し、各色または各階調につきヒストグラムを生成する生成部と、

前記ヒストグラムにおいて計数された各画素数を予め設定された画素数範囲に応じて量子化し、量子化した各値に対してシンボルを割り当てて、前記ヒストグラムを複数の前記シンボルからなるシンボル列へ変換し、前記各色または各階調につき得られた前記シンボル列を結合する変換部と、

10

各画像につき得られた、結合された前記シンボル列を照合し、共通するシンボルが指定数以上ある場合に、両画像が類似すると評価する評価部とを備える、画像処理装置。

【請求項 2】

前記変換部は、総画素数が多い色順または階調順に、前記シンボル列を結合する、請求項 1 に記載の画像処理装置。

【請求項 3】

前記生成部は、前記行ごとに同じ色または同じ階調の画素を計数し、前記各色または各階調につき行方向に係るヒストグラムを生成するとともに、前記列ごとに同じ色または同じ階調の画素を計数し、前記各色または各階調につき列方向に係るヒストグラムを生成する、請求項 1 または 2 に記載の画像処理装置。

20

【請求項 4】

前記各色または各階調のヒストグラムから各色または各階調の総画素数を計算し、各色または各階調間の総画素数の差が閾値を超えるか否かを判定する判定部をさらに備え、

前記変換部は、前記差が前記閾値を超えないと判定された場合、前記各色または各階調のシンボル列を結合せず、前記評価部は、前記各色または各階調のシンボル列をそれぞれ照合する、請求項 1 ～ 3 のいずれか 1 項に記載の画像処理装置。

【請求項 5】

行および列方向に複数の画素が配列してなる画像間の類似評価を、画像処理装置により行う方法であって、前記画像処理装置は、生成部と、変換部と、評価部とを備えており、

前記生成部により、前記行または列ごとに同じ色または同じ階調の画素を計数し、各色または各階調につきヒストグラムを生成するステップと、

30

前記変換部により、前記ヒストグラムにおいて計数された各画素数を予め設定された画素数範囲に応じて量子化し、量子化した各値に対してシンボルを割り当てて、前記ヒストグラムを複数の前記シンボルからなるシンボル列へ変換し、前記各色または各階調につき得られた前記シンボル列を結合するステップと、

前記評価部により、各画像につき得られた、結合された前記シンボル列を照合し、共通するシンボルが指定数以上ある場合に、両画像が類似すると評価するステップとを含む、画像処理方法。

【請求項 6】

前記結合するステップでは、総画素数が多い色順または階調順に、前記シンボル列を結合する、請求項 5 に記載の画像処理方法。

40

【請求項 7】

前記生成するステップでは、前記行ごとに同じ色または同じ階調の画素を計数し、前記各色または各階調につき行方向に係るヒストグラムを生成するとともに、前記列ごとに同じ色または同じ階調の画素を計数し、前記各色または各階調につき列方向に係るヒストグラムを生成する、請求項 5 または 6 のいずれか 1 項に記載の画像処理方法。

【請求項 8】

前記各色または各階調のヒストグラムから各色または各階調の総画素数を計算し、各色または各階調間の総画素数の差が閾値を超えるか否かを判定するステップをさらに含み、

前記結合するステップでは、前記差が前記閾値を超えないと判定された場合、前記各色

50

または各階調のシンボル列を結合せず、前記評価するステップでは、前記各色または各階調のシンボル列をそれぞれ照合する、請求項 5～7 のいずれか 1 項に記載の画像処理方法。

【請求項 9】

請求項 5～8 のいずれか 1 項に記載の画像処理方法を実現するための画像処理装置により読み取り可能なプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、類似画像検索の高速化を図ることができ、画像変形に対しても安定的に画像照合を行うことができる画像処理装置、画像処理方法およびその方法を実現するための装置読み取り可能なプログラムに関する。

【背景技術】

【0002】

画像データベースからの類似画像の検索は、一般に、画像データに関連付けられたキーワードに基づき行われている。このキーワードに基づく検索は、高い精度で類似画像を検索することができるものの、画像データの記憶の際、その画像データにキーワードを関連付けて記憶しなければならず、キーワードが関連付けられていない場合には、1 つずつ画像を確認しなければならず、ユーザにとって利便性が悪いという問題があった。

【0003】

この問題を解決するべく、画像から画像特徴量を抽出し、その画像特徴量を用いて類似画像を検索する技術が提案されている（例えば、特許文献 1 および非特許文献 1 参照）。これらは、画像の構図と色とを同時に考慮して類似画像検索を行う技術を開示するものであり、画像特徴量としては、画像内での全体的な色分布の状況を示すヒストグラムを用いている。ヒストグラムは、固定色ごとの画素数（度数）を求め、画像の全画素数で正規化したものである。

【0004】

特許文献 1 には、異なる画像データをそれぞれ複数のブロックに分割し、ブロックごとに取得された画像データの画像特徴量を所定の順序で並べて特徴量行列を作り、その特徴量行列を構成する各特徴量列間の距離を、DP マッチングを用いて算出し、算出した距離を 2 次元に配列してラインペナルティマトリクスを生成して、異なる画像データの類似度を、このマトリクスを用いて DP マッチングにより算出する装置および方法が記載されている。このように、各特徴量列間の距離を配列したマトリクスを用いて類似度を算出することで、画像特徴量の位置情報を考慮しつつ、その位置情報にある程度の違いがあった場合でも、高精度かつ高速に検索することができる。

【0005】

非特許文献 1 には、提示した例示画像から得られる特徴量とデータベース中の画像の特徴量との比較を行い、特徴量間の類似性の大きいデータベース中の画像を求めることにより、類似検索を行うことが記載されている。この技術は、最初に、30 枚の画像をデータベース中からランダムに選んで表示し、ユーザがこの中から必要な画像のイメージに近いものを何点か選択する。次に、これらの画像から得られる特徴量に基づいて検索のための重み付けを決定し、これらを用いて類似判定基準を作成し検索を行うものである。このように、ユーザが選択した複数枚の画像の特徴量から検索の際の類似判定に用いる重み付けを自動的に設定し、その重み付けを用いて検索を行うことができるため、ユーザの意図を反映した検索を可能にし、また、簡単で、かつ良好な結果を得ることができる。

【特許文献 1】特開 2006-146715 号公報

【非特許文献 1】南部藤太朗、八村広三郎、「類似判定基準の自動設定を行う画像検索システム」、[online]、平成 14 年 9 月、人文科学とコンピュータシンポジウム、[平成 20 年 6 月 12 日検索]、インターネット<URL:<http://www.arc.ritsumeai.ac.jp/img/nambu2002.pdf>>

10

20

30

40

50

【発明の開示】

【発明が解決しようとする課題】

【0006】

上記特許文献1では、画像データを複数のブロックに分割し、ブロック単位の照合結果を求めることにより、空間分布も考慮した照合を実現しようとしているが、各ブロック間の照合を多数回実施しなければならない、さらには、照合する画像間において分割単位が一致しない場合にはブロック単位の照合が失敗するおそれもある。これを回避するため、分割単位を小さくすることができるが、分割単位を小さくすると照合すべきブロック数が増加し、さらなる計算量の増加を招くという問題がある。

【0007】

また、上記非特許文献1では、各固定色の画像平面上のマクロな分布状況を示すものとして、固定色ごとに画像の縦・横両方に射影ヒストグラムを求め、その射影ヒストグラムの分布の中央値（メディアン）を与える座標位置を求める方法が記載されているが、この方法では、色の分布の中央値を与える座標位置を求めるのみで、分布形状そのものを特徴付けるものではなく、高い精度で照合を行うには不十分である。

【0008】

したがって、照合する画像間の分布形状を考慮した類似性を色別に求めることを可能とし、高精度で、かつ画像変形に対しても安定的に画像照合することができる画像処理装置や画像処理方法等の提供が望まれている。

【課題を解決するための手段】

【0009】

本発明の発明者は、鋭意検討の結果、色別に複数のヒストグラムを求め、それら複数のヒストグラムをシンボルが時系列に並ぶシンボル列によって表現し、複数のシンボル列を結合して1つの識別情報を生成し、色別に形成されたシンボル列を結合してなる識別情報を照合することで、高速で、かつ正確に画像間の比較処理を実行することができ、画像変形に対しても安定的に画像照合を行うことができることを見出した。

【0010】

本発明は、上記のことを見出すことによりなされたものであり、上記課題は、本発明の画像処理装置および画像処理方法、さらにはその方法を実現するための装置読み取り可能なプログラムを提供することにより解決することができる。

【0011】

すなわち、本発明によれば、行および列方向に複数の画素が配列してなる画像間の類似評価を行う画像処理装置であって、行または列ごとに同じ色または同じ階調の画素を計数し、各色につきヒストグラムを生成する生成部と、ヒストグラムにおいて計数された各画素数を予め設定された画素数範囲に応じて量子化し、量子化した各値に対してシンボルを割り当てて、ヒストグラムを複数のシンボルからなるシンボル列へ変換し、各色につき得られたシンボル列を結合する変換部と、各画像につき得られた、結合されたシンボル列を照合し、共通するシンボルが指定数以上ある場合に、両画像が類似すると評価する評価部とを備える、画像処理装置が提供される。この生成部および変換部では、各色につきヒストグラムを生成し、各色につきシンボル列を得ることに限られるのではなく、グレー画像の各階調につきヒストグラムを生成し、各階調につきシンボル列を得ることもできる。

【0012】

上記の変換部は、総画素数が多い色順に、シンボル列を結合することができる。総画素数が多い色のシンボル列は、総画素数の少ない色のシンボル列に比較して、画像データの特徴を有している。同じ画像なら、行または列方向ごとに同じようなシンボル列になるため、変化が激しい総画素数が多い色のシンボル列や、変化に乏しい総画素数が少ない色のシンボル列も、同じ傾向を示す。仮に、比較する原稿同士で原稿方向が異なっている、シンボル列の長さ順に従って結合しておけば、照合すべき方向同士が並ぶ順になる可能性が高い。このことから、原稿方向を揃える処理が不要となり、都合がよい。その結果、類似画像検索の高速化を図ることができる。

10

20

30

40

50

【0013】

上記の生成部は、行ごとに同じ色の画素を計数し、各色につき行方向に係るヒストグラムを生成するとともに、列ごとに同じ色の画素を計数し、各色につき列方向に係るヒストグラムを生成することが好ましい。このように複数の方向に係るヒストグラムを生成することで、より特徴が表れやすく、比較処理を高速に実行することができ、より高精度で、かつ画像変形に対しても類似評価を行うことができる。

【0014】

また、画像処理装置は、各色のヒストグラムから各色の総画素数を計算し、各色間の総画素数の差が閾値を超えるか否かを判定する判定部をさらに備えることができる。この場合、変換部は、その差が閾値を超えないと判定された場合、各色のシンボル列を結合せず、評価部は、各色のシンボル列をそれぞれ照合する。

10

【0015】

本発明によれば、行および列方向に複数の画素が配列してなる画像間の類似評価を、画像処理装置により行う方法も提供することができる。この画像処理方法は、上記の生成部と変換部と評価部とを備える画像処理装置により行うことができ、行または列ごとに同じ色または同じ階調の画素を計数し、各色につきヒストグラムを生成するステップと、ヒストグラムにおいて計数された各画素数を予め設定された画素数範囲に応じて量子化し、量子化した各値に対してシンボルを割り当てて、ヒストグラムを複数のシンボルからなるシンボル列へ変換し、各色につき得られたシンボル列を結合するステップと、各画像につき得られた、結合されたシンボル列を照合し、共通するシンボルが指定数以上ある場合に、両画像が類似すると評価するステップとを含む。

20

【0016】

また、上記の画像形成方法を実現するための画像形成装置により読み取り可能なプログラムとして提供することもでき、そのプログラムが記録された画像処理装置により読み取り可能な記録媒体としても提供することができる。

【発明を実施するための最良の形態】

【0017】

図1は、画像処理装置の1つの実施形態を示した図である。図1では、パーソナルコンピュータ（PC）10、11、12がネットワーク13を介して互いに接続されている。PC10、11、12は、画像処理装置で行うべき画像間の類似評価を行う処理の一部をそれぞれが行うことができるように構成されている。

30

【0018】

PC10、11、12のそれぞれは、図2に示すように、画像間の類似評価を行うために構成されたプログラムを実行するCPU20と、そのプログラムを格納するメモリ21と、画像データ等を格納する外部記憶装置としてのHDD22、CD-ROMドライブ23、FDドライブ24と、ネットワーク13に接続された他のPCとの通信を行うための通信装置25と、ユーザに対し、処理の内容や処理結果を表示するための表示装置26と、ユーザが情報等を入力するための入力装置27とを備え、バス28を介して互いに接続されている。

【0019】

40

CPU20は、上記プログラムのほか、各種制御処理を実行し、メモリ21は、そのための制御プログラムや、ブートプログラム、その他の各種データを格納する。なお、PC10、11、12が備えるCPUは、上記プログラムを実行することにより、後述するヒストグラムを生成する生成部、シンボル列へ変換する変換部および両画像が類似すると評価する評価部として機能する。この実施形態の場合、例えば、PC10のCPUが、生成部として機能し、PC11のCPUが、変換部として機能し、PC12のCPUが、評価部として機能するように構成することができる。このように各PC10、11、12に各機能を分担させることに限られるものではなく、1つのPCですべての機能を実現することもできる。

【0020】

50

この画像処理装置は、行および列方向に複数の画素が配列してなる画像間の類似評価を行う装置であり、上述した生成部、変換部、評価部を備える。生成部は、行または列ごとに同じ色または同じ階調の画素を計数し、各色または各階調につきヒストグラムを生成する。変換部は、ヒストグラムにおいて計数された各画素数を予め設定された画素数範囲に応じて量子化し、量子化した各値に対してシンボルを割り当てて、ヒストグラムを複数のシンボルからなるシンボル列へ変換し、各色または各階調につき得られたシンボル列を結合する。そして、評価部は、各画像につき得られた、結合されたシンボル列を照合し、共通するシンボルが指定数以上ある場合に、両画像が類似すると評価する。

【0021】

ここで、画像間の類似評価を含む処理の一例として画像検索処理について説明する。図3(a)は、原稿画像を例示した図である。この原稿画像に対し、特定の行画像を検索する場合を考える。特定の行画像とは、原稿画像中と同一である必要はなく、解像度が異なっている場合でも、部分的な行として形状が同じであればよく、画像として完全一致することには限定されない。ここでは、1行全体を文字領域の単位として使用しているが、これに限られるものではなく、その1行のうちのある一部分や、2行以上であってもよい。

【0022】

図3(b)は、原稿画像中の黒画素で示される部分に外接する外接矩形を表示した図である。黒画素が連続して形成される部分を1つの字体と認識し、その字体を4本の線分により取り囲んで外接矩形を構成する。「文」や「も」は、すべての線がつながっているため、1つの字体として認識される。これに対し、「は」、「字」、「紙」は、これらの字を構成する線と線の間に隙間があり、黒画素がつながっていないため、黒画素がつながっている部分を1つの字体として認識し、その字体を4本の線分により取り囲んで外接矩形を構成する。

【0023】

このようにして構成された外接矩形を、近隣同士連結し、これを行に成長させると、図3(c)に示すような1行全体を取り囲む矩形となる。この矩形で囲まれた文字列を抽出する処理を、行切り出し処理という。

【0024】

図4は、上記のようにして切り出した文字列画像と、その射影ヒストグラムとを例示した図である。行方向を矢線Aに示す方向とし、黒画素を計数する方向、すなわち列方向を矢線Bに示す方向とする。切り出した文字列は、図4(a)に示すように「通信管理レポ」であり、1列ごとに矢線Bに示す方向へ黒画素の数が計数される。図4(b)に示す射影ヒストグラムは、縦軸を黒画素数、横軸を行方向位置として、計数された黒画素の数を、行方向位置に対応させて図示したものである。文字に縦線があると、黒画素が矢線Bに示す方向に連続して並んでいるため黒画素数が多く、文字と文字との間の隙間部分では、黒画素が矢線Bに示す方向に1つも存在しないので、黒画素数は0となっている。なお、図4(b)には、射影ヒストグラム中の各点がつなぎ合わされ、波形が形成されているのが示されている。

【0025】

図4(b)に示す射影ヒストグラムを参照すると、「通信管理レポ」の各文字に対応して波形が形成されるが、その波形は、文字によって異なっている。「通」と「レ」ではその波形は大きく異なっている。このことは、この射影ヒストグラムが、文字画像の特徴を抽出するために利用できることを意味する。なお、この射影ヒストグラムは、その波形を特徴とするものであるから、文字画像として文字間の分割を前提とするものではなく、一行に並ぶ文字列であれば、手書き文字、筆記体、アラビア文字等であってもよい。

【0026】

また、射影ヒストグラムは、矢線Bに示す方向に黒画素数を計数するため、同じ文字であっても画像サイズが異なれば黒画素数が異なることとなるが、隣り合う黒画素数の相対的な関係は同じであるから、画像サイズに依存しないことがわかる。

【0027】

射影ヒストグラムは、黒画素数を計数するものであるから、1画素単位で変化する。黒画素数は、画像を読み取る際の読み取りノイズやかすれ等によって容易に変化する。このため、精度的に問題が生じるように思われるが、射影ヒストグラムの波形形状は概ね安定している。よって、射影ヒストグラムの正確な値を用いて、波形を照合することに意味はなく、およその値を利用しても波形形状の特徴は十分に表現可能である。

【0028】

このことから、射影ヒストグラムの各黒画素数を、予め設定された画素数の範囲に割り当てられた固定段階の値へ量子化し、各黒画素数を、それらより少ない固定段階の値へと変換することができる。量子化は、行方向の座標 i の位置の射影ヒストグラムの値（黒画素数）を y_i とし、例えば、以下の基準に基づき5段階の値に変換することにより行うことができる。

10

【0029】

- (1) y_i が 0 以上 10 未満のとき、0 段階。
- (2) y_i が 10 以上 20 未満のとき、1 段階。
- (3) y_i が 20 以上 30 未満のとき、2 段階。
- (4) y_i が 30 以上 40 未満のとき、3 段階。
- (5) y_i が 40 以上 50 未満のとき、4 段階。

【0030】

図5に、5段階の値へ変換した後の図を示す。Cで示された波形は、黒画素数の値を示し、Dで示された波形は、量子化した後の波形を示す。図5を参照すると、量子化しても、波形形状の特徴を十分に表現できていることがわかる。

20

【0031】

このようにして量子化を行うと、各射影ヒストグラムの値は、5段階の値に制限され、その結果、各段階をシンボル化することが可能となる。各段階の値は、以下の規則に基づきシンボル化することができる。

【0032】

- (i) 0 段階のとき、a へ変換。
- (ii) 1 段階のとき、b へ変換。
- (iii) 2 段階のとき、c へ変換。
- (iv) 3 段階のとき、d へ変換。
- (v) 4 段階のとき、e へ変換。

30

【0033】

この規則に従うことで、言語を問わず、一般的な文字領域画像はすべて、射影ヒストグラムへ変換し、量子化し、シンボル化するという過程を経て、文字領域画像は、例えば「a b a a b c c e e a a . . .」という文字画像の特徴を保持したまま、「a」、「b」、「c」、「d」、「e」といった文字や記号等からなるシンボルが並んだシンボル列へと変換することができる。

【0034】

このことから、2つの文字領域画像を比較する作業、例えば、画像検索においては被検索（データベース）領域画像と、検索（クエリー）領域画像の両方をシンボル化し、シンボル列へと変換することで、変換された2つのシンボル列を照合するのみで容易に、かつ高い精度で画像間の類似評価を行うことができる。

40

【0035】

なお、シンボル列の照合は、例えば、従来から開発されているテキスト処理の技術によって実現可能である。その1つとして、動的計画法に基づく文字列照合法を挙げることができる。この方法は、2つの文字列間の距離を算出し、その距離を利用して2つの文字列の類似性を判断するものである。距離尺度としては、編集距離等を挙げることができる。また、N-gram法を用いて2つの文字列間の類似性を判断することもできる。N-gram法は、検索対象を単語単位ではなく、文字数がN個からなるN文字単位で分解し、その文字単位の出現頻度を求める方法である。

50

【0036】

図6を参照して、N-gram法を利用した文字列照合処理について説明する。図6(a)は、「a」、「b」、「c」からなる文字列を例示した図である。固定文字数(ここでは、3文字)の部分文字列(trigram)を1単位とし、文字列中にそのtrigramが出現した回数を計数し、図6(b)に示すN-gramリストを作成している。その計数結果は、trigramが「abc」については2回、「bca」については2回、「caa」については2回、「aab」については1回となっている。

【0037】

このN-gramリストは、2つの文字列が類似しているほど、その結果が類似する。なお、類似するか否かの評価は、例えば、2つのN-gramリストA、Bにおいて出現回数が一致する部分文字列の数を計数し、その数が、指定した数以上か否かを基準に行うことができる。上記の例で言えば、指定数が3で、N-gramリスト中の「abc」、「bca」、「caa」、「aab」が、一方の画像については2、2、2、1回で、他方の画像については2、1、2、1回であった場合、「abc」、「caa」、「aab」については2、2、1回で共通しているため、共通する部分文字列の数は3であり、指定数3以上の条件を満たしている。よって、この場合、両画像は類似すると評価する。

【0038】

図6(c)には、文書Aと文書Bとが示されており、これらの文書に含まれる文字領域をシンボル化してシンボル列で表現されているのが示されている。文書から文字領域を抽出する方法としては、これまで知られたいかなる方法でも用いることができるが、領域識別やレイアウト分析等の文字認識分野において一般的な技術を採用することができる。これらのシンボル列をtrigramが出現した回数を計数し、N-gramリストを作成し、共通する部分文字列の数が指定数以上であるか否かを判定することにより、それら2つの文字列の一致度を評価することができる。

【0039】

N-gramリストの照合方法としては、順位相関係数を用いる方法がある。順位相関係数は、データが順位(ランク)という特殊な場合の相関係数で、相関係数は、2つの確率変数間の類似性の度合いを示す指標である。この方法については詳述しないが、照合方法としては、以下に例示する方法を用いることができる。

【0040】

図7を参照して、文書画像照合処理について詳細に説明する。ステップ700において、処理を開始し、ステップ710において、画像の入力を受け付け、ノイズ除去や2値化処理等の必要な処理を行う。次に、ステップ720において、レイアウト分析(領域識別)を行い、文字領域を抽出する。文字領域を抽出すると、図3(b)に示す状態となる。そして、ステップ730において、文字領域から行画像を抽出する。行画像を抽出すると、図3(c)に示す状態となる。

【0041】

ステップ740において、各行画像中の文字をシンボル化し、それをつなげてシンボル列へ変換する。この変換処理の詳細は、図8を参照して詳細に説明する。ステップ750において、変換したシンボル列からN-gramの出現頻度を集計する。そして、ステップ760において、比較対象の文書画像から同様に作成したN-gramの出現頻度の集計結果と照合する。なお、ステップ750および760における処理は、従来のDPマッチングによる文字列間距離測定によっても実現することができる。

【0042】

ステップ740における処理を、図8を参照して説明する。ステップ800において、抽出された行画像から行高さを推定する。具体的には、行画像の高さや、行内の矩形の最大高さとして行うことができる。ステップ810において、行方向に対して垂直な列方向へ黒画素数を計数し、射影ヒストグラムを作成する。ステップ820において、射影ヒストグラムの計数された各黒画素数を、ステップ800において推定した行高さを使用して量子化する。例えば、矩形の最大高さが50画素であるとき、その50画素を上記(1)～(

10

20

30

40

50

5) のように 5 つの段階に分け、量子化することができる。

【0043】

ステップ 830 において、量子化された各値を各シンボルへ変換し、シンボルが並んだシンボル列を作成する。ステップ 840 において、作成されたシンボル列をデータベースに記録する。ここでは行高さの最大高さを使用して量子化したが、行高さを使用することなく、射影ヒストグラム of 最大値を基準に量子化することもできる。

【0044】

これまで、文字行画像と文字行画像との類似度を計算する一例を示し、説明してきた。なお、文字行画像と文字行画像とを照合することができるのであれば、上記の方法に限定されるものではない。以下、文字行画像と文字行画像とをシンボル列に変換した後、類似度を判定するための照合処理を施すことが可能であるとして、文字行画像だけではなく、2次元画像についても適用可能な方法について説明する。

10

【0045】

まず、図 9 を参照して、従来の照合処理を簡単に説明しておく。図 9 (a) は、フルカラーの 2 次元画像を例示した図である。従来の処理では、図 9 (a) に示すカラー画像を減色した後、同色の画素のみを抽出した画素において、図 9 (b) に示すように、行方向および列方向に射影ヒストグラムを求め、その中央値 (メディアン) の座標を求め、画像の幅を高さで正規化したものを分布中心とし、それを画像特徴としている。

【0046】

この従来の処理では、図 9 (b) を参照してわかるように、画素の分布形状に関係なく中央値の座標のみを画像特徴として使用することから、分布中心は画像の大まかな特徴しか表現できない。このため、2つの画像の色と形状が同じであれば分布中心も同一であることは保証されるが、形状が大きく異なる場合でも分布中心がほぼ同一になるおそれがある。そうすると、形状が大きく異なるにもかかわらず、両画像は類似すると評価されることになる。

20

【0047】

そこで、図 10 (a) に示すように、行方向に求めた射影ヒストグラムと列方向に求めた射影ヒストグラムとを連結し、図 10 (b) に示すように、1つの射影ヒストグラムにした後、上述した方法によりシンボル化し、シンボル列へ変換する。このようにすることで、上述した文字行画像同士の照合処理と同様の方法で、類似判断を行うことができる。

30

【0048】

画像を構成する各画素は、色情報 (例えば、RGB 値) を有するため、その RGB 値により画素の色を判定することができ、その RGB 値に基づき同じ色を有する画素を抽出することができる。つまり、注目する単一色からなる画像を作成し、その画像から射影ヒストグラムを求める。比較対象画像に対しても同様に、その注目する色と同じ色からなる画像を作成し、射影ヒストグラムを作成する。これら 2つの射影ヒストグラムを比較し、形状が類似していれば元画像も類似していると判断することができる。

【0049】

フルカラー (32ビット) 画像で表現可能な色は、16777216色と非常に多く、非常に細かい色の違いも区別することができる。特に、写真等の自然画像の場合、高い再現力を伴って描写された画像中に含まれる色数は、非常に多く、同一色を見つけることが難しい場合もある。同一画像ではなく、類似画像を求めることが目的であれば、多少の色の差は無視した方が類似画像を見つけやすい。このことから、類似する色を1つの代表色で表現するために、減色処理を施す。この減色処理により、画像中に含まれる色数を所望の数に制限し、色の差を考慮する程度を調整することができる。なお、色数を少なく制限すれば制限するほど、色の差が無視される傾向が強くなる。減色アルゴリズムとしては、例えば、メディアンカット減色法や、このメディアンカット減色法に誤差拡散処理を追加した方法や、K-Means法等を挙げることができる。

40

【0050】

メディアンカット減色法について簡単に説明すると、RGBで表される色空間を、まず

50

、赤成分の画素値を大きい順に並べ、全画素を画素数が半分になる箇所を分割面として分割する。同様にして、緑成分の分割面、青成分の分割面を求め、各分割面で分割する。このようにして、分割された色空間が選択すべき色数以下になるまで分割し、選択すべき色数に達したときの分割面を選択されたインデックスとする。インデックス作成後、各画素の色数に対応するインデックスカラーで埋めることにより減色画像を作成することができる。なお、このようにして減色処理を行うと、歪みが生じ、その歪みは輪郭を目立たせる。目立たないようにするために、減色処理後、ディザ処理を施すことができる。

【0051】

図11に減色処理が施される画像例を示す。図11(a)は、32ビットのフルカラー画像である。すなわち、16777216色で表現された画像である。図11(b)は、8色に減色された画像である。この図11(b)を参照すると、フルカラーから8色に減色されたとしても、印象が大きくことならないことがわかる。フルカラー画像と減色処理された画像では、色数の差が非常に大きい、人間の感覚ではほぼ類似した画像に見え、このことは、減色処理された画像を照合対象としても、フルカラー画像を照合対象とした場合とほぼ同様な類似性判断結果を得ることができることを意味する。

【0052】

図11(c)は、減色された画像のうち、最も画素数が多かった色のみを抽出し、その色を黒色とし、その他の色を白色とした場合の画像である。この最も画素数が多かった色は、この画像を特徴付けるにあたって最も支配的な色となる。この画像からは、周辺が同じ黒色であり、中心に別の白色が略円状にあるという大まかな特徴を認識することができる。

【0053】

図11(d)は、2番目に画素数が多かった色を抽出し、その色を黒色とし、その他の色を白色とした場合の画像を示し、図11(e)は、3番目に画素数が多かった色を抽出し、その色を黒色とし、その他の色を白色とした場合の画像を示す。このように色別に分離することにより画像内に描写されている物体の概形が特徴的に抽出されている。

【0054】

これは、カラー画像に限られるものではなく、グレイ画像についても階調の程度により同様に扱うことができる。図12は減色処理が施される画像例を示し、図12(a)はオリジナルのグレイ画像を、図12(b)はそのグレイ画像の階調数を減じた画像を、図12(c)は最も画素数が多かった階調のみを抽出し、そのときの色を黒色とし、その他の階調のものを白色とした場合の画像を、図12(d)は2番目に多かった階調のみを抽出し、そのときの色を黒色とし、その他の階調のものを白色とした場合の画像を、図12(e)は3番目に多かった階調のみを抽出し、そのときの色を黒色とし、その他の階調のものを白色とした場合の画像を示す。減色処理では、例えば256階調のグレイ画像を8階調へ減色することができる。

【0055】

図11に示した減色処理の詳細を、図13に示すフローチャートを参照して説明する。ステップ1300において処理を開始し、まず、ステップ1310において、記録すべき画像データがなくなったかを判定する。なくなった場合は、ステップ1380へ進み、この処理を終了する。記録すべき画像データがまだ存在する場合、ステップ1320へ進み、まだ記録されていない画像データを読み込む。次に、ステップ1330へ進み、読み込んだ画像データに対し、減色アルゴリズムを適用し、減色処理を行う。

【0056】

ステップ1340において、同じ画素値、すなわち同じ色から構成される画像を色別に作成する。これは、最も画素数が多かった色のみを抽出し、その色を黒色とし、その他の色を白色とした場合の画像や、2番目に多かった色のみを抽出し、そのときの色を黒色とし、その他の階調のものを白色とした場合の画像といったように、色別の画像を作成する。

【0057】

ステップ1350において、異なり色数のリストを作成するために、色の種類を色情報として記録する。画像によっては、赤味がかかったものや、青味がかかったものがあり、赤味がかかったものの場合、減色に際し、代表色として赤系の色が多く選択され、緑や青系の色の選択は少なくなる。このため、色情報として記録されていない色も存在し、このような色は新色として追加される。ステップ1360において、単色画像の画像特徴を抽出する。具体的には、まず、各画像の各列につき、黒画素数を計数して、各色のヒストグラムを生成する。この画像ヒストグラムを生成する処理は生成部が行う。そのヒストグラムの各値を段階に分け、量子化し、シンボル化する。シンボル化したものをつなぎ合わせ、シンボル列へ変換する。このシンボル列への変換処理は変換部が行う。ステップ1370において、そのシンボル列を色情報とともにデータベースに記録する。データベースは、図1に示すいずれかのPCが備える記憶装置、ネットワーク13に接続された別途設けられる記憶装置、図2に示すHDD等とすることができる。その後、再びステップ1310へ戻り、記録すべき画像データが存在するかを判定する。

10

【0058】

次に、図14を参照して、データベースに記録された画像の色情報を量子化し、色コードブックを作成する処理について説明する。この処理は、図13に示す減色処理によりデータベースに色情報を記録した後に行うことができる。

【0059】

ステップ1400において処理を開始し、ステップ1410で、図13のステップ1350において記録した色情報を読み込む。ステップ1420において、ベクトル量子化法（例えば、L B G法）を用いて、色情報から代表色リストを作成する。このベクトル量子化法は、サンプルをN個ずつ取り込み、N次元のベクトルとし、全てのサンプルを取り込んだ後、符号化する数であるM個に各ベクトルをクラスタリングし、M個に分けられたクラスタの各々から1つずつ代表ベクトルを決め、そのクラスタ内の他のベクトルをその代表ベクトルへ置き換え、その代表ベクトルを符号化することにより量子化する。したがって、ステップ1420では、読み込んだ色情報をN次元のベクトルとし、M個のクラスタに分け、各クラスタにつき代表ベクトルを決定し、その代表ベクトルを符号化することにより量子化して代表色を決定し、それをリスト化することで、代表色リストを作成する。クラスタは、近似した色の集合で、代表色は、そのクラスタを代表する色である。

20

【0060】

ステップ1430において、代表色リストを色コードブックとして記録し、ステップ1440で処理を終了する。

30

【0061】

図15を参照して、データベースに記録された画像の色情報を量子化する処理について説明する。ステップ1500において処理を開始し、ステップ1510で、未処理の画像特徴データがあるかを判定する。未処理の画像特徴データがない場合には、ステップ1550へ進み、この処理を終了する。

【0062】

ステップ1510で未処理の画像特徴データが存在する場合、ステップ1520へ進み、データベースから画像特徴データと色情報とを読み込む。ステップ1530で、図14のステップ1430にて作成した色コードブックを用いて、色情報を量子化する。量子化された色情報は、代表色リストのどの色で代替するかを決定する。

40

【0063】

ステップ1540で、画像特徴データと色情報と量子化した色情報とを関連付けて記録する。その後、再びステップ1510へ戻り、未処理の画像特徴データが存在するかを判定する。

【0064】

図16を参照して、この量子化した結果を用いて行う検索処理について説明する。ステップ1600において処理を開始し、ステップ1610で、まず、質問画像データを読み込む。例えば、画像をスキャンすることにより、質問画像データを読み込むことができる

50

。ステップ1620で、その質問画像に対し、減色処理を施す。上述した減色アルゴリズムを適用することにより、減色処理を施すことができる。ステップ1630で、図11(c)～(e)に示すような色別の画像データを生成し、ステップ1640で、図14のステップ1430にて作成した色コードブックを用いて、色別に色情報を量子化する。そして、ステップ1650で、色別の画像データの画像特徴を抽出する。このとき、生成部は、色別の画像データからヒストグラムを生成し、変換部は、画像特徴であるシンボル列への変換を行う。

【0065】

次に、ステップ1660で、質問画像の画像特徴を、データベースに記録されている画像特徴と照合し、類似画像の検索を行う。すなわち、質問画像から変換されたシンボル列を、データベースに記録されているシンボル列と照合し、シンボル列中のシンボルがどの程度一致しているかを判断して、類似画像を検索する。その際、量子化した色情報が一致することが必要であるが、ヒストグラム形状の類似性を色別に判定すると、色別に類似度を算出する必要があるが、さらに、色別の類似度から総合的類似度を算出する必要がある。これでは高速検索を行うことができない。このため、変換部は、色別に作成されたヒストグラム形状から求めたシンボル列を、画素数の多い色順に連結することにより、1つのシンボル列にまとめる。変換部が1つにまとめた後、評価部は、1つにまとめたシンボル列を照合し、それらシンボル列間の類似性を評価する。

【0066】

画像を色別に分離してから、色別にヒストグラム形状を求め、シンボル列に変換しているので、シンボルを連結しても複数色からなる元画像から直接、ヒストグラム形状を求め、シンボル列に変換した場合とは異なり、色別の形状特徴は保存されたままである。したがって、形状特徴と色情報とを同時に用いて画像の類似度を評価することができる。このようにすれば、色別に類似度を評価する必要があるため、高速処理を実現することができる。

【0067】

そのための処理を、図17を参照して詳細に説明する。図17は、データベースに画像を記録するためのフローを示した図である。ステップ1700で処理を開始し、ステップ1710で、記録すべき画像データが存在するか否かを判定する。ここで、記録すべき画像データが存在しない場合は、ステップ1800へ進み、この処理を終了する。ステップ1710で、記録すべき画像データが存在する場合、ステップ1720へ進み、処理すべき画像データを読み込む。そして、ステップ1730で、読み込んだ画像データに対し、減色アルゴリズムを適用する。ステップ1740において、同じ画素値（同じ色）から構成される画像を色別に作成する。

【0068】

その後、ステップ1750において、異なり色数リストを作成するために、色の種類を色情報として記録する。新色であれば、色を追加する。ステップ1760で、単色画像の画像特徴を抽出する。このとき、生成部は、列ごとに黒画素数を計数し、画素ヒストグラムを作成し、変換部は、画素ヒストグラムの各値（各列における黒画素数）を段階に分け、それをシンボル化し、シンボルをつなぎ合わせてシンボル列へ変換する。

【0069】

ステップ1770で、画素ヒストグラムから色別に画素の総数を求める。ステップ1780で、画素数が多い色の順に、予め決定した方向から各色のシンボル列を連結し、1つのシンボル列を作成する。例えば、画素数が多い順に、左側からつなぎ合わせて1つのシンボル列を作成する。

【0070】

ステップ1790で、そのシンボル列（画像特徴）と色情報（減色後の色と量子化後の色の組）をデータベースに記録する。このようにして記録された被検索データを検索する場合は、図16に示した検索処理を実行することにより行うことができる。

【0071】

10

20

30

40

50

色別に求められたヒストグラムから変換されたシンボル列を1つのシンボル列へ連結するための順序は、類似画像であれば同じ順になるような、また異なる画像であれば違う順になるような、一定の規則に基づいて行う。これは、連結する順序が異なると、部分的には類似するシンボル列が見つかるものの、全体としては大きく異なってしまうシンボル列パターンが形成されてしまうので、上述した傾向を満たす規則に基づき連結順を決める必要がある。

【0072】

色別の画素数の多い順にパターンを連結する場合、色別の画素数の差が大きく、スキャンや減色化といった画像条件によって、画素数の多い順の順序が変化しない場合には問題は生じない。しかしながら、色別の画素数がわずかなものである場合、比較対象画像間の画像条件が異なると、色別の画素数に基づく順序が変化する場合があり、そういった場合、シンボル列パターンが部分的に入れ替わってしまうおそれがある。そうすると、色別の画素数の差が僅差である場合には、実際に両画像は類似しているものの、評価部による評価では類似していないと評価される場合がある。

10

【0073】

そこで、色別の画素数の差が大きいものである場合は、上述した方法で類似性を評価し、その差が小さい場合には、同じ色あるいは類似する色同士のパターンを比較し、類似性を評価するように判定部を設ける。この判定部は、各色のヒストグラムから各色の総画素数を計算し、各色間の総画素数の差が閾値を超えるか否かを判定する。この判定を受けて、変換部は、その差が閾値を超えないと判定された場合、各色のシンボル列を結合せず、評価部は、各色のシンボル列をそれぞれ照合する。

20

【0074】

このように、色別の画素数の差によって、色別のシンボル列を連結するか否かを切り替え、照合することで、色別の画素数が特徴的な画像に対しては高速検索を実現しつつ、色別の画素数が微差である画像に対しても、質問画像と被検索画像との間の画像条件の違いに起因して検索が失敗することがなくなる。この処理を、図18を参照して詳細に説明する。

【0075】

図18は、色別の画素数の差に応じ、画像照合の方法を切り替えて画像検索する処理を示したフローチャートである。ステップ1810で処理を開始し、ステップ1820で、質問画像データを読み込む。ステップ1830で、読み込んだ質問画像データに対し、減色処理を施す。ステップ1840で、色別に画像を作成する。ステップ1850で、図14のステップ1430で作成した色コードブックを用いて、色別に色情報を量子化する。

30

【0076】

次に、ステップ1860で、色別の画像データの画像特徴を抽出する。生成部は、列ごとに黒画素数を計数し、各色のヒストグラムを作成し、変換部は、ヒストグラムの各値（各列における黒画素数）を段階に分け、それに対応するシンボルへ変換し、変換されたシンボルをつなぎ合わせて各色のシンボル列を作成する。これまでは、図16に示した検索処理と同様である。ステップ1870では、判定部は、作成した色別のヒストグラムから求めた各色間の総画素数の差あるいは比が閾値を超えるか否かを判定する。この判定は、各色間につき行われ、いずれの色間についても閾値を超える場合に、閾値を超えると判定する。このようにして判定された結果が、閾値を超える場合にはステップ1890へ進み、閾値以下である場合にはステップ1880へ進む。

40

【0077】

ステップ1880で、画素毎に色別にシンボル化されているデータベース中の被検索画像と、ステップ1860で求めた各色のシンボル列とをそれぞれ比較し、類似度を算出する。類似度は、共通するシンボル数が多いほど高くなるが、予め指定した数以上のシンボルが共通する場合に両画像が類似すると評価することができる。この評価は、シンボルを1つずつ照合するのではなく、上述したN個の文字からなる部分文字列により照合することもできる。

50

【0078】

ステップ1890では、色別のシンボル列を連結し、1つのシンボル列を作成する。その後、ステップ1900で、画像毎に1つに連結されたシンボル列を記録したデータベース中の被検索画像と、ステップ1890で求めた連結されたシンボル列とを比較し、類似度を算出する。

【0079】

このように、被検索データについては、色別のシンボル列と、色別のシンボル列を連結した1つのシンボル列との両方を記録しておけば、質問画像に応じて検索シンボルを切り替える場合において、処理毎に、シンボル列を連結する必要がないため、高速処理を実現することが可能となる。

【0080】

なお、明細書中には特に示していないが、画像の読み込みは入力部、減色処理は減色処理部、画像データの作成は画像データ生成部、色情報の量子化は量子化部が行うことができ、画像処理装置は、これらの各部を備えることができる。これらの各部および上記判定部も、上記の生成部、変換部および評価部と同様、プログラムとして構成し、CPUがプログラムを実行することにより、CPUをこれらの各部として機能させることができる。したがって、本発明は、画像間の類似評価を行う画像処理装置および画像処理方法として提供することもできるが、その画像処理装置により読み取り可能なプログラムとしても提供することができるものである。

【0081】

これまで本発明を実施の形態をもって説明してきたが、本発明は上述した実施の形態に限定されるものではなく、他の実施の形態、追加、変更、削除など、当業者が想到することができる範囲内で変更することができ、いずれの態様においても本発明の作用・効果を奏する限り、本発明の範囲に含まれるものである。

【図面の簡単な説明】

【0082】

【図1】本発明の画像処理装置の1つの構成例を示した図。

【図2】図1に示すPCの構成例を示した図。

【図3】原稿、矩形抽出、行切り出しを例示した図。

【図4】射影ヒストグラムを例示した図。

【図5】量子化した結果を射影ヒストグラムとともに例示した図。

【図6】N-gramリストを用いて文書画像の一致度を評価しているところを示した図。

【図7】文書画像照合処理の流れを示した図。

【図8】シンボル列への変換処理を示した図。

【図9】分布中心を用いて2次元画像を表現したところを示した図。

【図10】2次元画像を1つのヒストグラムへ変換する1実施形態を示した図。

【図11】フルカラー画像を減色し、画素数が多い順に色別画像を作成しているところを示した図。

【図12】グレー画像の階調数を減じ、画素数が多い順に階調別画像を作成しているところを示した図。

【図13】データベース中に画像を記録する処理の流れを示した図。

【図14】データベース中に記録された画像の色情報の色別コードブックを作成する処理の流れを示した図。

【図15】色情報のコードブックを用いてデータベース中の画像の色情報を量子化する処理の流れを示した図。

【図16】質問画像の色情報を量子化し、データベース中の画像を検索する処理の流れを示した図。

【図17】データベース中に画像を蓄積する別の処理の流れを示した図。

【図18】質問画像の色情報を量子化し、データベース中の画像を検索する別の処理の流れを示した図。

10

20

30

40

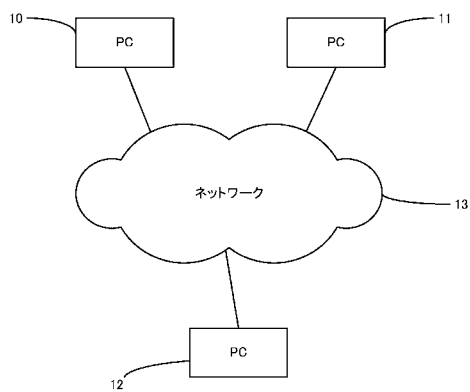
50

【符号の説明】

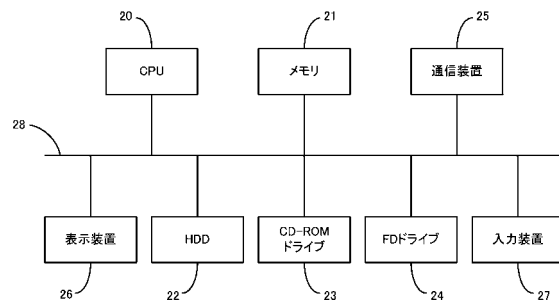
【0083】

10、11、12…PC、13…ネットワーク、20…CPU、21…メモリ、22…HDD、23…CD-ROMドライブ、24…FDドライブ、25…通信装置、26…表示装置、27…入力装置、28…バス

【図1】



【図2】



【図 3】

(a)

文字認識技術は、紙データを電子化する手段の一つです。オフィス業務のペーパーレス化が進んだ今日でも紙が使用される局面は多く、文字認識技術に対する期待は小さくなることはないようです。

(b)

文字認識技術は、紙データを電子化する手段の一つです。オフィス業務のペーパーレス化が進んだ今日でも紙が使用される局面は多く、文字認識技術に対する期待は小さくなることはないようです。

(c)

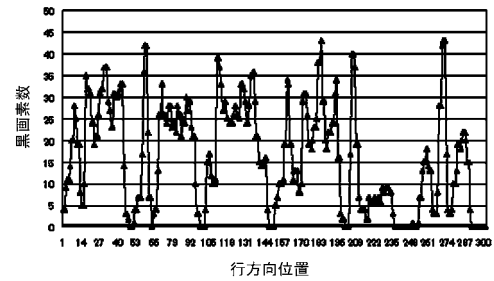
文字認識技術は、紙データを電子化する手段の一つです。オフィス業務のペーパーレス化が進んだ今日でも紙が使用される局面は多く、文字認識技術に対する期待は小さくなることはないようです。

【図 4】

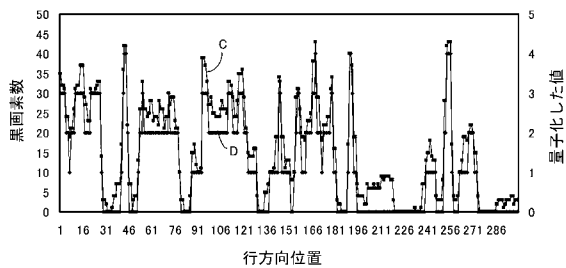
(a)



(b)

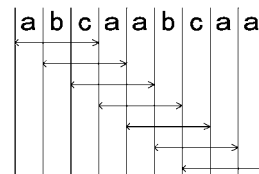


【図 5】



【図 6】

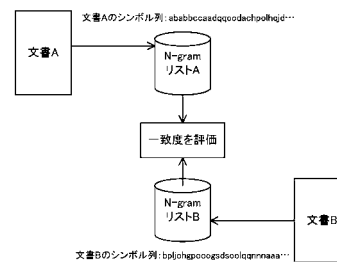
(a)



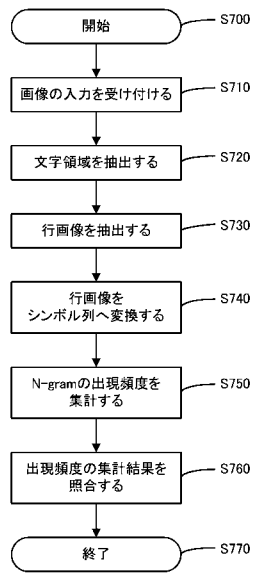
(b)

trigram	回数
abc	2
bca	2
caa	2
aab	1

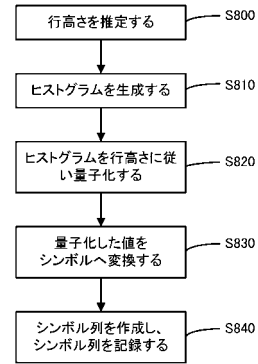
(c)



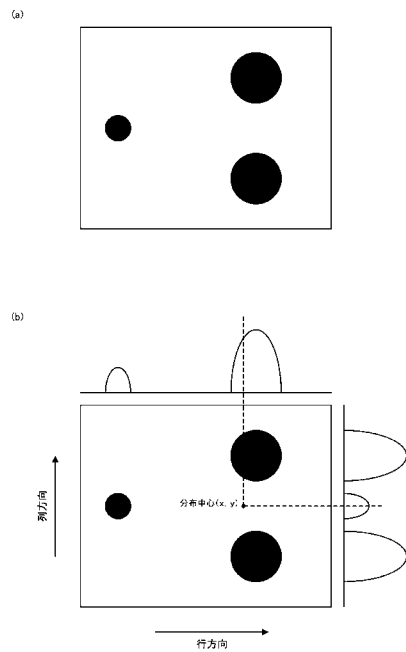
【図 7】



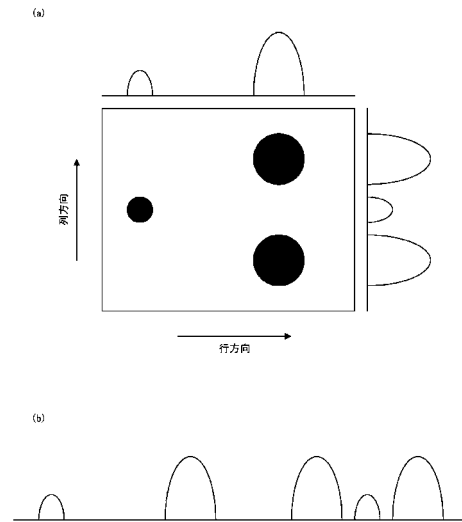
【図 8】



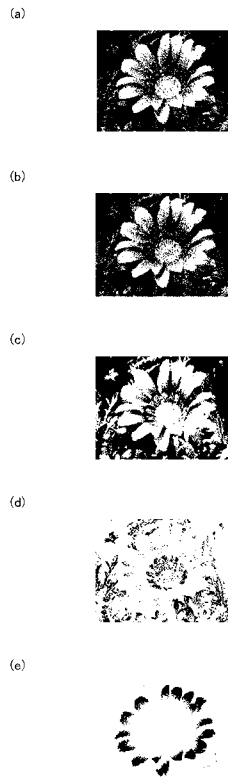
【図 9】



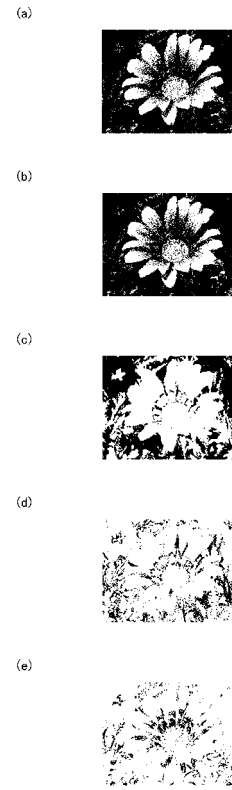
【図 10】



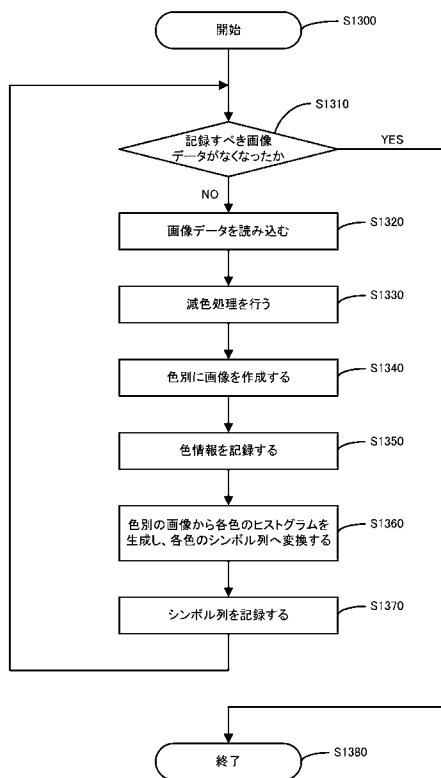
【図 1 1】



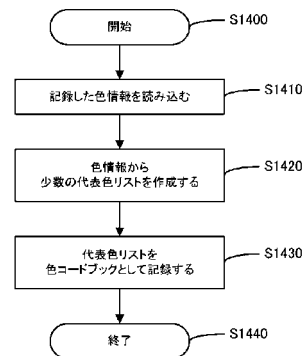
【図 1 2】



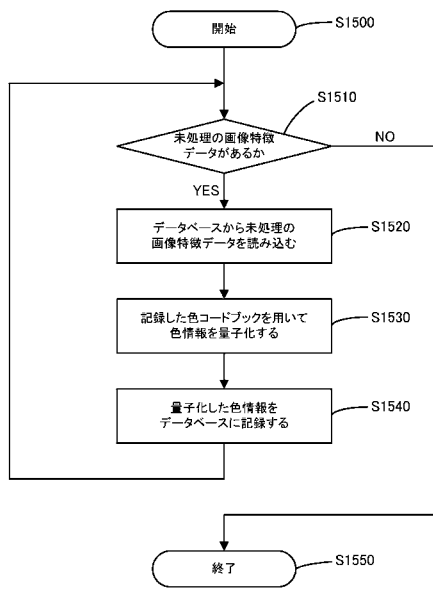
【図 1 3】



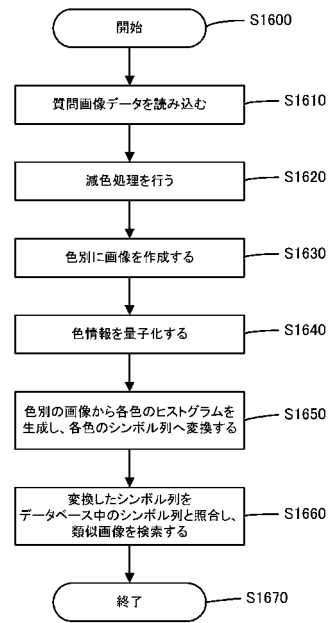
【図 1 4】



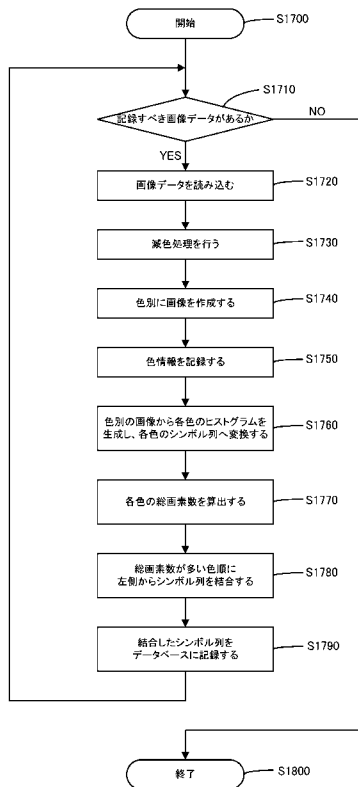
【図 15】



【図 16】



【図 17】



【図 18】

