

(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2007-172077

(P2007-172077A)

(43) 公開日 平成19年7月5日(2007.7.5)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 17/30 (2006.01)	G06F 17/30 350C	5B050
G06T 1/00 (2006.01)	G06F 17/30 170B	5B075
	G06F 17/30 310C	
	G06T 1/00 200E	

審査請求 未請求 請求項の数 12 O L (全 23 頁)

(21) 出願番号	特願2005-365409 (P2005-365409)	(71) 出願人	000005496 富士ゼロックス株式会社 東京都港区赤坂九丁目7番3号
(22) 出願日	平成17年12月19日 (2005.12.19)	(74) 代理人	100075258 弁理士 吉田 研二
		(74) 代理人	100096976 弁理士 石田 純
		(72) 発明者	小山 剛弘 東京都港区赤坂二丁目17番22号 富士 ゼロックス株式会社内
		Fターム(参考)	5B050 BA10 BA16 BA20 CA07 DA06 EA06 EA18 FA02 GA07 GA08 5B075 ND06 PP04 PP28 QM05

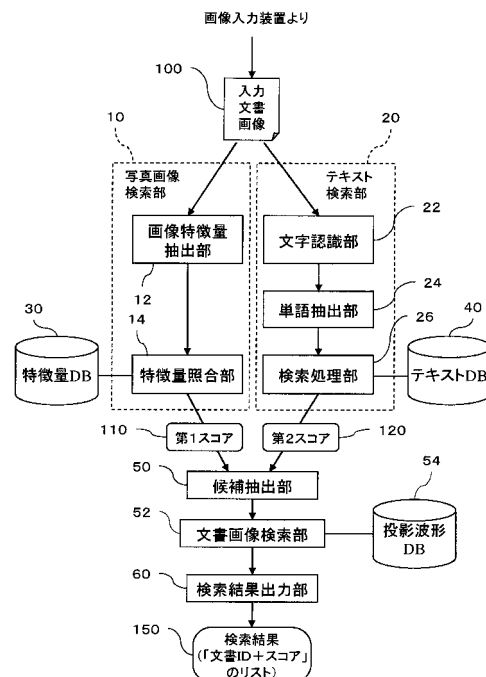
(54) 【発明の名称】 画像検索システム及び方法及びプログラム

(57) 【要約】

【課題】 多様な種類の画像に対して適切な画像検索方法を提供する。

【解決手段】 写真画像検索部10は、入力文書画像100をブロックに分割し、各ブロックごとにエッジ量などの画像特徴量を求める。そして、その画像特徴量の分布と各登録画像の画像特徴量の分布とを比較することで各登録画像の入力文書画像100に対する類似度スコア(第1スコア110)を計算する。テキスト検索部20は、画像を文字認識した結果としてのテキストの類似性に基づき、各登録画像の入力文書画像100に対する類似度スコア(第2スコア120)を計算する。候補抽出部50は、各登録画像の第1スコア110と第2スコア120とを統合し、得られた統合スコアが上位となる登録画像を候補として抽出する。文書画像検索部52は、それら候補と入力文書画像100とを、投影波形同士で比較することで、最終的な類似度を計算する。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

複数の登録画像の中から入力画像に類似する 1 以上の登録画像を検索する画像検索システムであって、

入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第 1 の類似度スコアを計算する第 1 スコア計算手段と、

入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と各登録画像のテキスト特徴量とに基づき、各登録画像の入力画像に対する第 2 の類似度スコアを計算する第 2 スコア計算手段と、

各登録画像の第 1 の類似度スコアと第 2 の類似度スコアを統合して統合スコアを求め、統合スコアが良い順に候補となる登録画像を抽出する候補抽出手段と、

入力画像の 1 以上の所定方向についてそれぞれ投影波形を求め、求められた 1 以上の投影波形と、候補抽出手段で候補として抽出された各登録画像の前記 1 以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第 3 の類似度スコアを計算する第 3 スコア計算手段と、

第 3 スコア計算手段が求めた第 3 の類似度スコアに基づき、入力画像に類似する 1 以上の登録画像を求める検索手段と、

を備える画像検索システム。

【請求項 2】

請求項 1 記載の画像検索システムにおいて、

前記候補抽出手段は、各登録画像の第 1 の類似度スコアに対して統計的な標準化処理を施すことにより各登録画像の第 1 の標準化スコアを求め、各登録画像の第 2 の類似度スコアに対して統計的な標準化処理を施すことにより各登録画像の第 2 の標準化スコアを求め、登録画像ごとにその登録画像の第 1 の標準化スコアと第 2 の標準化スコアとを統合することによりその登録画像の統合スコアを求める、

ことを特徴とする画像検索システム。

【請求項 3】

請求項 2 記載の画像検索システムにおいて、

前記候補抽出手段は、登録画像の前記第 1 の標準化スコアとして、前記各登録画像の第 1 の類似度スコア全体の中での当該登録画像の第 1 の類似度スコアの偏差値を計算し、登録画像の前記第 2 の標準化スコアとして、前記各登録画像の第 2 の類似度スコア全体の中での当該登録画像の第 2 の類似度スコアの偏差値を計算することを特徴とする画像検索システム。

【請求項 4】

複数の登録画像の中から入力画像に類似する 1 以上の登録画像を検索する画像検索システムであって、

入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第 1 の類似度スコアを計算する第 1 スコア計算手段と、

入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と各登録画像のテキスト特徴量とに基づき、各登録画像の入力画像に対する第 2 の類似度スコアを計算する第 2 スコア計算手段と、

入力画像に含まれるテキストの量を求め、テキストの量に基づき入力画像において連続調画像及びテキストのうちのいずれが優越的であるかを所定の判定条件に基づき判定し、連続調画像が優越的であると判定された場合は第 1 スコア計算手段に入力画像を供給し、テキストが優越的であると判定された場合は第 2 スコア計算手段に入力画像を供給し、その結果求められる各登録画像の第 1 の類似度スコア又は第 2 の類似度スコアが良い順に候補となる登録画像を抽出する候補抽出手段と、

入力画像の 1 以上の所定方向についてそれぞれ投影波形を求め、求められた 1 以上の投

10

20

30

40

50

影波形と、候補抽出手段で候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算する第3スコア計算手段と、

第3スコア計算手段が求めた第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める検索手段と、
を備える画像検索システム。

【請求項5】

請求項4記載の画像検索システムであって、

前記候補抽出手段は、前記入力画像がテキストが優越的であるとも連続調画像が優越的であるとも判定されなかった場合は、前記第1スコア計算手段と前記第2スコア計算手段の双方に前記入力画像を供給し、その結果求められる各登録画像の第1の類似度スコアと第2の類似度スコアを統合して統合スコアを求め、統合スコアが良い順に候補となる登録画像を抽出する、

10

ことを特徴とする画像検索システム。

【請求項6】

複数の登録画像の中から入力画像に類似する1以上の登録画像を検索する画像検索システムであって、

入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第1の類似度スコアを計算し、第1の類似度スコアが良い順に一次候補となる登録画像を抽出する第1候補抽出手段と、

20

入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と第1候補抽出手段で一次候補として抽出された各登録画像のテキスト特徴量とに基づき、それら各登録画像の入力画像に対する第2の類似度スコアを計算し、第2の類似度スコアが良い順に二次候補となる登録画像を抽出する第2候補抽出手段と、

入力画像の1以上の所定方向についてそれぞれ投影波形を求め、求められた1以上の投影波形と、第2候補抽出手段で二次候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算し、第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める検索手段と、

30

を備える画像検索システム。

【請求項7】

複数の登録画像の中から入力画像に類似する1以上の登録画像を検索する画像検索方法であって、

入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第1の類似度スコアを計算し、

入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と各登録画像のテキスト特徴量とに基づき、各登録画像の入力画像に対する第2の類似度スコアを計算し、

40

各登録画像の第1の類似度スコアと第2の類似度スコアを統合して統合スコアを求め、統合スコアが良い順に候補となる登録画像を抽出し、

入力画像の1以上の所定方向についてそれぞれ投影波形を求め、求められた1以上の投影波形と、候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算し、

第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める、
画像検索方法。

【請求項8】

複数の登録画像の中から入力画像に類似する1以上の登録画像を検索する画像検索方法

50

であって、

入力画像に含まれるテキストの量を求め、テキストの量に基づき入力画像において連続調画像及びテキストのうちのいずれが優越的であるかを所定の判定条件に基づき判定し、

連続調画像が優越的であると判定された場合は、入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第1の類似度スコアを計算し、

テキストが優越的であると判定された場合は、入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と各登録画像のテキスト特徴量とに基づき、各登録画像の入力画像に対する第2の類似度スコアを計算し、

10

計算された第1の類似度スコア又は第2の類似度スコアが良い順に、候補となる登録画像を抽出し、

入力画像の1以上の所定方向についてそれぞれ投影波形を求め、求められた1以上の投影波形と、候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算し、

第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める、
画像検索方法。

【請求項9】

複数の登録画像の中から入力画像に類似する1以上の登録画像を検索する画像検索方法であって、

20

入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第1の類似度スコアを計算し、第1の類似度スコアが良い順に一次候補となる登録画像を抽出し、

入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と一次候補として抽出された各登録画像のテキスト特徴量とに基づき、それら各登録画像の入力画像に対する第2の類似度スコアを計算し、第2の類似度スコアが良い順に二次候補となる登録画像を抽出し、

入力画像の1以上の所定方向についてそれぞれ投影波形を求め、求められた1以上の投影波形と、二次候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算し、第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める、

30

画像検索方法。

【請求項10】

複数の登録画像の中から入力画像に類似する1以上の登録画像を検索する画像検索システムとしてコンピュータシステムを動作させるためのプログラムであって、前記コンピュータシステムに、

入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第1の類似度スコアを計算し、

40

入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と各登録画像のテキスト特徴量とに基づき、各登録画像の入力画像に対する第2の類似度スコアを計算し、

各登録画像の第1の類似度スコアと第2の類似度スコアを統合して統合スコアを求め、統合スコアが良い順に候補となる登録画像を抽出し、

入力画像の1以上の所定方向についてそれぞれ投影波形を求め、求められた1以上の投影波形と、候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算し、

第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める、

50

処理を実行させるためのプログラム。

【請求項 1 1】

複数の登録画像の中から入力画像に類似する 1 以上の登録画像を検索する画像検索システムとしてコンピュータシステムを動作させるためのプログラムであって、前記コンピュータシステムに、

入力画像に含まれるテキストの量を求め、テキストの量に基づき入力画像において連続調画像及びテキストのうちのいずれが優越的であるかを所定の判定条件に基づき判定し、

連続調画像が優越的であると判定された場合は、入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第 1 の類似度スコアを計算し、

10

テキストが優越的であると判定された場合は、入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と各登録画像のテキスト特徴量とに基づき、各登録画像の入力画像に対する第 2 の類似度スコアを計算し、

計算された第 1 の類似度スコア又は第 2 の類似度スコアが良い順に、候補となる登録画像を抽出し、

入力画像の 1 以上の所定方向についてそれぞれ投影波形を求め、求められた 1 以上の投影波形と、候補として抽出された各登録画像の前記 1 以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第 3 の類似度スコアを計算し、

20

第 3 の類似度スコアに基づき、入力画像に類似する 1 以上の登録画像を求める、処理を実行させるためのプログラム。

【請求項 1 2】

複数の登録画像の中から入力画像に類似する 1 以上の登録画像を検索する画像検索システムとしてコンピュータシステムを動作させるためのプログラムであって、前記コンピュータシステムに、

入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第 1 の類似度スコアを計算し、第 1 の類似度スコアが良い順に一次候補となる登録画像を抽出し、

30

入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と一次候補として抽出された各登録画像のテキスト特徴量とに基づき、それら各登録画像の入力画像に対する第 2 の類似度スコアを計算し、第 2 の類似度スコアが良い順に二次候補となる登録画像を抽出し、

入力画像の 1 以上の所定方向についてそれぞれ投影波形を求め、求められた 1 以上の投影波形と、二次候補として抽出された各登録画像の前記 1 以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第 3 の類似度スコアを計算し、第 3 の類似度スコアに基づき、入力画像に類似する 1 以上の登録画像を求める、

処理を実行させるためのプログラム。

【発明の詳細な説明】

40

【技術分野】

【0001】

本発明は、データベース等に登録された登録画像の中から入力画像に類似する画像を検索する技術に関する。

【背景技術】

【0002】

近年、企業における情報のセキュリティ強化、コンプライアンスあるいは個人情報の保護など情報の取扱いが重要視されてきている。例えば、企業側は、監査などに応じて業務がどのような情報に基づいて遂行されたかを公開する必要がある。そのためには、情報に対するアクセス状況をログ管理して、いつ誰がどのような情報をどこからどう処理したの

50

かなどを特定できるようにしておく必要がある。

【0003】

このような流れの中、例えば複写や印刷出力、ファクシミリ送信などによる文書の出力や、スキャンによる紙文書の電子化入力などの出来事があった場合、その時に出力又は入力された画像データを日付や操作者名などと共にログとして残しておき、後で漏洩の可能性がある文書が発見された場合、その文書と同じ文書をそのログ(画像ログ)から検索することで、漏洩源などを特定する等というシステムが提案されている。

【0004】

このようなシステムのためには、注目する文書に該当する画像をログの中から検索する必要がある。検索キーワードなどを細々と入力するのではなく、注目する文書をスキャンして得た画像そのものを検索条件として検索ができれば、操作性が非常によくなる。上記のようなセキュリティ目的に限らなくても、スキャンした画像に近い画像を画像データベースの中から検索できると便利である。

10

【0005】

このような画像検索に関する従来技術として、特許文献1~4に示すものがある。これら従来技術は、スキャンした文書の画像データから画像特徴量を算出し、その画像特徴量に基づいてデータベースから類似の画像データを検索している。

【0006】

また特許文献5に示される装置は、文書画像とその付属情報(検索要求文)とが検索条件として入力されると、その文書画像と画像特徴量が類似する登録文書をデータベースから検索して中間結果を求め、その中間結果の中から検索条件の付属情報と一致するものを最終的な検索結果としている。

20

【0007】

また特許文献6には、入力画像に対して横方向又は縦方向の投影波形を求め、登録画像の投影波形と比較することにより、入力画像に類似する登録画像を求める方式が示される。

【0008】

また特許文献7には、入力画像を複数のブロックに分割し、ブロックごとに等高線特徴や周波数分布特徴などといった複数の画像特徴量を求め、画像特徴量の分布が入力画像と近い画像を検索する方式が示される。

30

【0009】

【特許文献1】特開2004-139210号公報

【特許文献2】特開平9-270902号公報

【特許文献3】特開2003-281176号公報

【特許文献4】特開平10-49659号公報

【特許文献5】特開2005-149071号公報

【特許文献6】特開2003-91730号公報

【特許文献7】特開2001-319232号公報

【発明の開示】

【発明が解決しようとする課題】

40

【0010】

特許文献1~4の従来技術は、いずれも基本的に単一の画像特徴の値に基づき画像の類似性を判定している。ところが、このような単一の画像特徴による類似性の判定には、それぞれ得意な画像種類、不得意な画像種類があるなどして、必ずしも見た目の感覚に近い類似度が得られない。例えば風景写真画像同士の類似度を算出したい場合、OCR(光学文字認識)で求めたテキスト文字列を元にした特徴量を使用する類似度算出手段は無力である。逆に構造などが非常に似ているが書かれているテキスト内容が全く違う文書画像が複数あった場合、それらを画像の濃度や濃度分布を元にした特徴量を使用する類似度算出手段で処理するとスコアに差が出にくいいため、それらの中から所望の文書を識別することが困難になる。

50

【0011】

従来技術は、検索の対象が例えば写真のみというように特定の種類に限定されており、その種類に適合した特徴を用いている場合なら、かなり精度のよい判定ができる。しかしながら、企業活動では写真や図面、テキスト文書、或いはそれらが組み合わされた文書などといった多様な種類の文書を取り扱っており、それら多様な文書がログ乃至データベースに蓄積され、検索の対象となる。写真の識別に有用な画像特徴はテキスト文書の識別には向かない場合が多く、この逆も言える。したがって、多様な種類の文書画像を検索しようとした場合、単一の画像特徴に基づいていた上記従来技術はいずれも不十分なものであった。

【0012】

特許文献5の技術は、単一の画像特徴量に付属情報を組み合わせることで検索の精度の向上をもくろんでいるが、検索において着目する画像特徴量が検索条件の画像の種類に合っていないと、中間結果として適切なものが得られないため、最終的な検索結果の精度も望めない。また、この方式のためには、登録画像や検索条件の画像に付属情報を関連付ける必要がある。例えば、紙原稿を単にスキャンした画像の類似画像を検索する用途では付属情報は存在しないのでこの方式は利用できない。この方式を利用しようとするれば、ユーザに検索要求文のような付属情報を入力してもらう必要が出てくるため、ユーザの操作負担が増大する。

10

【0013】

また、特許文献6及び7の技術も、特許文献1～4と同様、基本的に単一の画像特徴の値に基づき画像の類似性を判定しているため、類似性の判定には、それぞれ得意な画像種類、不得意な画像種類があるなどして、必ずしも見た目の感覚に近い類似度が得られない可能性がある。

20

【0014】

本発明の一つの側面では、多様な種類の文書画像に適用できる検索技術を提供する。

【課題を解決するための手段】

【0015】

本発明の一つの側面では、複数の登録画像の中から入力画像に類似する1以上の登録画像を検索する画像検索システムであって、入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第1の類似度スコアを計算する第1スコア計算手段と、入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と各登録画像のテキスト特徴量とに基づき、各登録画像の入力画像に対する第2の類似度スコアを計算する第2スコア計算手段と、各登録画像の第1の類似度スコアと第2の類似度スコアを統合して統合スコアを求め、統合スコアが良い順に候補となる登録画像を抽出する候補抽出手段と、入力画像の1以上の所定方向についてそれぞれ投影波形を求め、求められた1以上の投影波形と、候補抽出手段で候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算する第3スコア計算手段と、第3スコア計算手段が求めた第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める検索手段と、を備える画像検索システム、を提供する。

30

40

【0016】

本発明の別の側面では、複数の登録画像の中から入力画像に類似する1以上の登録画像を検索する画像検索システムであって、入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第1の類似度スコアを計算する第1スコア計算手段と、入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と各登録画像のテキスト特徴量とに基づき、各登録画像の入力画像に対する第2の類似度スコアを計算する第2スコア計算手

50

段と、入力画像に含まれるテキストの量を求め、テキストの量に基づき入力画像において連続調画像及びテキストのうちのいずれが優越的であるかを所定の判定条件に基づき判定し、連続調画像が優越的であると判定された場合は第1スコア計算手段に入力画像を供給し、テキストが優越的であると判定された場合は第2スコア計算手段に入力画像を供給し、その結果求められる各登録画像の第1の類似度スコア又は第2の類似度スコアが良い順に候補となる登録画像を抽出する候補抽出手段と、入力画像の1以上の所定方向についてそれぞれ投影波形を求め、求められた1以上の投影波形と、候補抽出手段で候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算する第3スコア計算手段と、第3スコア計算手段が求めた第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める検索手段と、を備える画像検索システム、を提供する。

10

【0017】

本発明の更に別の側面では、複数の登録画像の中から入力画像に類似する1以上の登録画像を検索する画像検索システムであって、入力画像を複数の領域に分割して領域ごとに画像特徴量を求め、求めた領域ごとの画像特徴量の分布と、各登録画像の領域ごとの画像特徴量の分布とに基づき、各登録画像の入力画像に対する第1の類似度スコアを計算し、第1の類似度スコアが良い順に一次候補となる登録画像を抽出する第1候補抽出手段と、入力画像に対する文字認識処理により得られたテキスト文字列に関して所定のテキスト特徴量を求め、このテキスト特徴量と第1候補抽出手段で一次候補として抽出された各登録画像のテキスト特徴量とに基づき、それら各登録画像の入力画像に対する第2の類似度スコアを計算し、第2の類似度スコアが良い順に二次候補となる登録画像を抽出する第2候補抽出手段と、入力画像の1以上の所定方向についてそれぞれ投影波形を求め、求められた1以上の投影波形と、第2候補抽出手段で二次候補として抽出された各登録画像の前記1以上の所定方向についての投影波形とに基づき、それら各登録画像の入力画像に対する第3の類似度スコアを計算し、第3の類似度スコアに基づき、入力画像に類似する1以上の登録画像を求める検索手段と、を備える画像検索システム、を提供する。

20

【発明を実施するための最良の形態】**【0018】**

以下、図面を参照して、本発明の好適な実施の形態（以下「実施形態」と呼ぶ）について説明する。

30

【0019】

図1を参照して、本実施形態の画像検索装置の構成を説明する。この画像検索装置は、既存の画像データベース又は画像ログ蓄積装置（図示省略）に登録された登録画像の中から、入力文書画像100に類似する度合いの高い登録画像を検索する。入力文書画像100は、スキャナで紙原稿を読み取ることで得た画像データや、各種のアプリケーションプログラムで作成された画像ファイルをビットマップ画像に変換したものなどである。入力文書画像100が与えられると、画像検索装置は、その入力文書画像を写真画像検索部10及びテキスト検索部20にそれぞれ入力する。

【0020】

写真画像検索部10は、写真等の連続調画像に適した検索処理を行う手段であり、入力された画像を複数のブロックに分割し、ブロックごとの画像特徴量の類似度合いに基づき画像検索を行う。写真画像検索部10は、典型的にはコンピュータでプログラムを実行することにより実現される。

40

【0021】

写真画像検索部10において、画像特徴量抽出部12は、入力文書画像100の画像特徴量を求める手段である。画像特徴量としては、例えばエッジ量の2次元分布を画像特徴量として用いてもよい。これは、図2に示すように、一定サイズの画像200を一定分割数（例えば 8×8 、 16×16 、 32×32 など）で一定サイズのブロック210に分割し、画像200に対してエッジ抽出フィルタを適用し、この結果求められるエッジの量をブロック210ごとに集計し、この結果求められる各ブロック210のエッジ量の組合せ

50

すなわち分布を画像特徴量とするものである。このエッジ量分布を用いる方式は、デジタルカメラ等で撮影された写真画像に対して好適な方式である。ブロック210ごとに、エッジ量の代わりに当該ブロックの平均色を求め、各ブロックの平均色の組合せ(分布)を画像特徴量として用いてもよい。また、画像を二値化し、ブロック210ごとの黒画素の割合の分布を画像特徴量とするなど、従来より様々な画像特徴量とそれを用いたスコア付け方式が利用されており、写真画像検索部10はそのいずれの画像特徴量及びスコア付け方式を用いてもよい。また、特許文献7に示される特徴量、例えばブロックごとに求めた等高線特徴の分布や、ブロックごとに求めた周波数分布特徴の分布などを、照合のための画像特徴量として用いることもできる。また、以上に例示した画像特徴量のうちの2以上の組合せを用いて照合を行うこともできる。

10

【0022】

なお、写真画像検索部10で用いるそれら画像特徴量は、比較的アルゴリズムが単純なのでソフトウェア処理でも高速に演算することができる。また、そのような画像特徴量の計算や画像特徴量を用いた照合処理は、ハードウェア回路による実装も可能であり、高速処理に向いている。

【0023】

特徴量照合部14は、画像特徴量抽出部12が求めた入力文書画像100の画像特徴量(又は複数の画像特徴量の組合せ)と、画像データベース又は画像ログ(共に図示省略)中の各登録画像の画像特徴量(又は複数の画像特徴量の組合せ)との類似度を計算する。各登録画像の画像特徴量は、それら登録画像を画像データベース又は画像ログに登録する際に画像特徴量抽出部12と同じアルゴリズムで計算し、特徴量DB(データベース)30に登録しておく。すなわち、特徴量DB30には、例えば、登録画像の文書ID(識別情報)に対応づけて当該登録画像の画像特徴量(例えば前述のエッジ量の分布など)が登録されている。特徴量照合部14は、各登録画像ごとに、当該登録画像の画像特徴量と入力文書画像100の画像特徴量との類似性の強さを示す類似度スコアを計算する。類似度スコアの計算は、従来公知の手法を用いて行えばよい。

20

【0024】

写真画像検索部10は、以上のようにして登録画像ごとに入力文書画像100に対するその登録画像の類似度スコア(画像特徴量に基づくスコア)を計算し、出力する。図1では、この類似度スコアを第1スコア110として示している。写真画像検索部10の出力は、例えば、登録画像ごとにその登録画像の文書IDと第1スコア110とをペアにしたものとなる。この出力は候補抽出部50に入力される。

30

【0025】

以上写真画像検索部10について説明した。次にテキスト検索部20について説明する。テキスト検索部20は、入力文書画像100中に存在するテキスト文字列の特徴を用いて類似の登録画像を検索する手段であり、典型的にはコンピュータでプログラムを実行することにより実現される。テキスト検索部20は文字認識部22、単語抽出部24及び検索処理部26を備える。

【0026】

文字認識部22は、入力文書画像100に含まれる文字をOCR(光学文字認識)アルゴリズム又はOCR回路により認識する。OCRアルゴリズム又はOCR回路としては、従来からあるものを用いればよい。

40

【0027】

単語抽出部24は、文字認識部22が出力する文字列に対して形態素解析等の公知の自然言語解析を実行することで、入力文書画像100中に現れる単語を抽出する。単語抽出部24の出力は、例えば、入力文書画像100中に含まれる単語の集合のデータとなる。この集合のデータには、単語ごとに、入力文書画像100でのその単語の出現回数の情報を含めることが好適である。また、抽出する単語を品詞(例えば、名詞のみなど)によって制限してもよい。

【0028】

50

検索処理部 26 は、単語抽出部 24 が求めた出現単語の集合を用いてテキスト DB 40 を検索し、各登録画像の入力文書画像 100 に対する類似度スコアを計算する。

【0029】

ここで、テキスト DB 40 には、単語を見出しとして、その単語が出現する登録画像の文書 ID のリストが登録されている。テキスト DB 40 は、例えば、登録画像の画像を画像データベースやログ等に登録する際に、文字認識及び単語抽出を行い、テキスト DB 40 の抽出した各単語を見出しとする欄にそれぞれ当該登録画像の文書 ID を登録することで作成できる。ここで、登録画像に同じ単語が複数回出現することもよくあることなので、好適には、単に見出し単語に対応づけて文書 ID を登録するだけでなく、当該単語の登録画像中での出現回数も対応づけて登録する。

10

【0030】

このテキスト DB 40 を用いた類似度スコアの計算は、例えば次のように行う。すなわちテキスト検索部 20 は、単語抽出部 24 で抽出された各単語をキーとしてテキスト DB 40 を検索し、単語ごとに、それが出現した登録文書の文書 ID と出現回数のペアを得る。1つの単語に対し複数の文書 ID と出現回数のペアが求められることも当然起こり得る。そして、得られた情報に基づき、単語ごとに、その単語が出現する登録画像の各文書 ID に対してスコアを加算する。

【0031】

この計算では、例えば、単語ごとに得られた文書 ID と出現回数のペアの情報を、文書 ID ごとに整理することで、個々の登録画像について、その登録画像に現れる各単語とその出現回数のペアを求める。そして、入力文書画像 100 と登録画像との間で各単語の出現回数の差を計算し、その差の絶対値の合計（或いは二乗和や二乗平均）を計算する。この計算結果は、入力文書画像 100 と登録画像とが同一のものであれば「0」となり、両者の差が大きくなるほど値が大きくなる。そこで、その計算結果の符号を反転したものを当該登録文書の類似度スコアとすれば、登録画像が入力文書画像 100 にテキストとして似ているほどそのスコアは大きい値を持つようにできる。このような計算を、画像データベース又は画像ログに登録された全ての登録文書について行う。

20

【0032】

なお、以上では計算を画像データベース又は画像ログに登録された全ての登録文書について行うとしたが、これに限らず、所定の条件を満たすものを対象として行うようにしても良い。このような計算対象の絞り込みの条件は、例えば、画像データベース又は画像ログに対して登録画像が登録された日時の範囲や、登録画像を登録したユーザの属するグループなどとして指定できる。すなわち、指定された範囲に属する日時に登録され文書のみを計算対象としたり、指定されたグループに属するユーザが登録した文書のみを計算対象としたりするなどである。このような絞込の条件は、画像検索装置のユーザインタフェースを介してユーザが指定するようによい。

30

【0033】

検索処理部 26 は、このようにして求めた登録文書の文書 ID と類似度スコア（図 1 における第 2 スコア 120）とのペアを、全ての登録文書について出力する。出力されたデータは候補抽出部 50 に入力される。

40

【0034】

以上、テキスト検索部 20 における類似度スコア算出の例を説明したが、テキストベースでの類似度スコア算出方式はこれに限定されるものではない。従来よりテキスト検索の分野では、キーワード群やキーワードの論理式で示された検索条件に対して各登録文書の適合度を示す検索スコアを求める方式が様々に開発されており、これら従来技術を本実施形態のテキスト検索部 20 の方式として採用することももちろん可能である。

【0035】

なお、本実施形態には、一般的なキーワードベースでの文書検索と異なる検索目的の特徴がある。すなわち、キーワード検索は基本的にそのキーワードを含む文書を検索するものであって、キーワードと検索される文書はそもそも別物であるが、本実施形態の検索目

50

的の一つである漏洩検証では、1ページ乃至複数ページ分の画像（入力文書画像100）と全く同一乃至は極めて近い1ページ乃至複数ページの登録画像を検索するというように、基本的に同一に近いものを見つけようとする。したがって、本実施形態の検索では、単語の出現回数だけでなく、更に個々の単語の画像中での出現位置を考慮することで、更に検索の精度を向上させることができる。この検索方式は、例えば次のようなものとなる。

【0036】

すなわち、この方式では、図3に示すように画像300を複数のブロック310に分割する。図示例では縦8×横8の64個のブロック310に分割している。そして、その画像中から抽出された単語の先頭文字が属するブロックを、当該単語の出現位置とする。この方式では、テキストDB40には、見出しの単語ごとに、登録画像の文書IDとその画像内でのその単語の出現位置（ブロック番号など）のリストとが登録される。検索では、入力文書画像100から抽出した単語の出現位置と同じ位置に登録画像中でも同じ単語があれば類似度スコアを高くし、そうでなければ（入力文書画像100の単語が登録画像中の同じ位置にない場合、及び登録画像中の単語が入力文書画像100の同じ位置にない場合などでは）類似度スコアを低くする。具体例としては、例えば、入力文書画像100中の単語ごとに、当該単語の入力文書画像100での出現位置及び登録画像での出現位置の総数（両画像で同一位置が出現位置となっている場合は1つとして計数する）のうち、両画像で共通して出現位置となっているものの割合を求める。そして、この（割合+1）の逆数の2倍を係数として、前述した当該単語についての入力文書画像100と登録画像との間での出現回数の差に乗ずる。そして、その乗算結果の絶対値を、入力文書画像100及び登録画像における全出現単語にわたって合計（或いは二乗和等）した結果を符号反転したものを類似度スコアとする。なお、ここに例示したのは、単語の出現位置を類似度スコアに反映させる方式の一例に過ぎず、この他にも様々な方式を用いることができる。

10

20

【0037】

なお、以上の例では、文字認識結果から単語を抽出したが、単語レベルまで分析しなくても、文字認識結果に現れる個々の部分文字列を求め、その部分文字列に対して上記と同じ処理を施してもよい。

【0038】

以上に説明したテキスト検索部20は、検索速度は比較的高速であり、文書画像中に含まれるテキストの量が多いと高い検索精度が得られるというメリットがある。しかし、テキストが抽出できない文書画像や、抽出できても少量しか抽出できないような文書画像については、高い検索精度を得ることができない。このようにテキスト検索部20も苦手な文書画像を持っている。

30

【0039】

以上に説明した写真画像検索部10の処理とテキスト検索部20の処理とは、同時並列的に実行してもよいし、1つずつ順番に実行してもよい。

【0040】

候補抽出部50は、写真画像検索部10の出力した第1スコア110とテキスト検索部20の出力した第2スコア120とを統合することで、画像特徴とテキスト特徴とを総合評価した統合スコアを計算する。そして、統合スコアが上位の登録画像を後続の文書画像検索部52での検索対象の候補として抽出する。

40

【0041】

図1の装置では、画像特徴に基づく類似度スコアとテキスト特徴に基づく類似度スコアという特性の大きく異なるスコアを統合するために、統計的な標準化という考え方を採用する。すなわち、各類似度スコアは、それぞれ異なる尺度であるため、それらを単純に比較したり、それらに対してそのまま加算・乗算等の演算を加えたのでは、適切なスコアとならない可能性が高いので、本実施形態では、そのような生の類似度スコアを、登録画像の全体集合の中での位置を示す値として標準化（規格化）する。標準化の方式としては、登録文書の類似度スコアを、登録文書全体の中での偏差値に変換する方式を例示することができる。以下、この方式を説明する。処理の流れは図5を参照する。

50

【 0 0 4 2 】

ある画像群（登録画像群） $G = \{ G_1, G_2, G_3, \dots, G_n \}$ から、任意の画像 a （入力文書画像 100）に類似した画像群を選択し、それらを類似している順に並べて出力する類似画像検索システムを考える。入力画像 a に対する登録画像群 G の各画像 G_i （ $i = 1, 2, 3, \dots, n$ ）の各種の特徴量 F_j （ $j = 1, 2, 3, \dots, m$ ）を用いた類似度スコアを $S_{ij}(a)$ とする。特徴量 F_j は、例えばエッジ量分布などの画像特徴量、単語の出現回数に基づくテキスト特徴量などである。これら各特徴量 F_j に着目した時の各登録画像 G_i の類似度スコア $S_{ij}(a)$ は、上述した写真画像検索部 10 及びテキスト検索部 20 の処理結果として取得できる（S1）。

【 0 0 4 3 】

候補抽出部 50 は、各スコア $S_{ij}(a)$ を特徴量の種類 j ごとに標準化した偏差値 $Z_{ij}(a)$ を算出する。この計算は、以下のような式に基づき行えばよい。

【 0 0 4 4 】

【 数 1 】

$$Z_{ij}(a) = \frac{S_{ij}(a) - \overline{S_j(a)}}{D_j(a)} \quad : \text{偏差値}$$

$$\overline{S_j(a)} = \frac{\sum_{i=1}^n S_{ij}(a)}{n} \quad : \text{平均}$$

$$D_j(a) = \sqrt{\frac{\sum_{i=1}^n (S_{ij}(a) - \overline{S_j(a)})^2}{n-1}} \quad : \text{標準偏差}$$

【 0 0 4 5 】

ここで、

【 数 2 】

$$\overline{S_j(a)}, D_j(a), Z_{ij}(a)$$

は、それぞれ、特徴量 F_j を使用した時の入力画像 a に対する登録画像 G_i における類似度 $S_{ij}(a)$ の平均値、標準偏差、偏差値となっている。

【 0 0 4 6 】

図 1 の例でいえば、テキスト特徴に基づく類似度スコアの登録画像全体の中で偏差値を求めてこれを標準化テキストスコアとすると共に（S2a）、画像特徴に基づく類似度スコアの登録画像全体の中で偏差値を求めてこれを標準化画像特徴量スコアとする（S2b）。

【 0 0 4 7 】

このように偏差値 $Z_{ij}(a)$ をスコアとして用いることで、異なる特徴量を使用した類似度スコア間でも、値として比較可能となる。ただし、これは、登録画像の数が十分多く、同じ特徴量 F_j についての各登録画像の類似度スコアの分布が正規分布に近い分布を示すことを仮定している。登録画像が、テキスト文書、図面、写真、それらの混合画像などと様々な種類にわたり、しかもそれらが十分に多い数であれば、この仮定は妥当な場合

10

20

30

40

50

が多いと考えられる。

【0048】

このようにして各特徴量 F_j についての類似度スコアを標準化すると、次に候補抽出部 50 は、同じ登録画像についてのそれら各特徴量についての標準化スコアを統合することで、その登録画像についての統合スコアを計算する (S3)。入力画像 a に対する登録画像 G_i の統合スコアを $S_i(a)$ とすると、これは次式で求められる。

【0049】

【数3】

$$S_i(a) = f(Z_{i1}(a), Z_{i2}(a), Z_{i3}(a), \dots, Z_{im}(a))$$

10

【0050】

ここで、 f は各特徴量 F_j についての偏差値 $Z_{ij}(a)$ から統合スコアを求めるための関数である。このような関数 f としては、例えば、変数各項の中の最大値を採る関数を用いることができる。画像についての各種特徴量には画像の種類に識別力の高低があるのは既述の通りだが、入力画像 a の種類に対して識別力の高い特徴量を用いたときの類似度スコアは、入力画像 a に一致又は極めて近い登録画像に対しては高い値になり、入力画像 a と異なる登録画像に対してはかなり低い値となってくる。したがって、このような類似度スコアから求めた偏差値は、入力画像 a に一致又は極めて近い登録画像については他よりも特異的に大きな値となる。これに対し、入力画像 a の種類の画像に対して識別力が低い特徴量を用いた場合の類似度スコアは、入力画像 a に一致又は極めて近い登録画像のものと、他の登録画像とでそれほど顕著な差が出てこないため、偏差値に変換してもそれほど大きな値とはならない。したがって、関数 f として各特徴量 F_j についての標準化スコア (偏差値) $Z_{ij}(a)$ の中で最大値をとるものを用いれば、入力画像 a に一致又は極めて近い登録については非常に高い値となり、他の登録画像については (仮にどの特徴量に対する偏差値が最大値として採用されるとしても)、一致又は極めて近い登録画像よりもかなり低い値となる。したがって、このような方式で求められる統合スコアは、多様な種類の入力画像に対して一致又は極めて近い登録画像を見つけるという用途に非常に適したものである。

20

30

【0051】

また、関数 f としては、この他にも、変数各項の相加平均や相乗平均などを採る関数を用いることもできる。

【0052】

このスコア統合処理のポイントは、検索対象となる登録画像群 G_i をサンプル集合と仮定することで、単一の登録画像との比較だけでは不可能であった標準化を行うことができ、特徴量 F_j とそれを使用したときの類似度スコアに対する詳細な知見を用いなくても、ある程度妥当性の高い統合された類似度スコアを提供できるという点にある。単一の登録画像 (これは画像群 G_i の中の画像数が 1 であるケースと等価) しか考慮しない場合、テキストと画像濃度などといった評価尺度や特性の全く異なる特徴についての類似度を何の補正もなしに使用してしまうか、或いはそれらを組み合わせるための確からしいモデルを考案してパラメータ評価を実施するなどが考えられる。しかし、前者では妥当性の高いスコアは得難く、後者は少なくとも特徴量 F_j とそれを使用した場合の類似度に対する詳細な知見が必要となるため、多様な種類の画像を対象としようとする場合実現が難しい。

40

【0053】

例えば、企業活動で使用される文書の中には、特許明細書のようにレイアウトが決まっています。ほとんど文字ばかりの文書、表計算ソフトで作成した名簿その他の一覧表のように文字と罫線からなる定型文書、テンプレートを用いて作成されたプレゼンテーション資料のようにレイアウトの類似性が高い文書、図を多用したプレゼンテーション資料、特許図

50

面のように図形が主で文字が少ない文書、ほとんど全面写真のパンフレット、写真と文章が混在したパンフレット、などと多様な種類の文書がある。このような多様な文書に対する汎用のモデルを考案するには多大の困難がある。これに対し、このスコア統合方式を用いれば、テキスト文書に強い特徴量を用いたスコア算出部、写真に強い特徴量を用いたスコア算出部、図面に強い特徴量を用いたスコア算出部などといった多様な特徴量に基づくスコア算出部を併用し、それらを各々の中で標準化した上で統合することで、好適な統合スコアを得ることができる。

【 0 0 5 4 】

ここで、統合する類似度スコアは、相互の相関が低い性質に基づき求めたものであることが好適である。例えば、前述のテキストベースでのスコアと、エッジ量分布や平均色分布などを用いた写真に強いスコアとを併用すれば、テキスト文書、写真のいずれに対しても高い精度で検索を行うことができる。

10

【 0 0 5 5 】

以上では、写真画像検索部 1 0 及びテキスト検索部 2 0 は、すべての登録画像についての類似度スコアを候補抽出部 5 0 に供給したが、この代わりに類似度スコアが予め定めた閾値以上となった登録画像のみの類似度スコアを候補抽出部 5 0 に供給するようにしてもよい。この場合、候補抽出部 5 0 は、類似度スコアを受け取っていない登録画像については、その類似度スコアを「0」などといった予め定めた値とした上で、上述のスコア統合処理を行えばよい。

【 0 0 5 6 】

候補抽出部 5 0 は、このようにして、登録画像ごとに、入力画像 a (入力文書画像 1 0 0) に対する類似度合いを示す統合スコアを計算する。

20

【 0 0 5 7 】

そして、候補抽出部 5 0 は、求めた統合スコアが上位のものから順に、次の文書画像検索部 5 2 での検索対象となる候補を選択する。すなわち、候補抽出部 5 0 は、文書画像データベース又は文書画像ログ記憶装置 (図示省略) に格納された登録画像群から、第一次の候補の絞り込みを行うわけである。候補の抽出では、統合スコアが上位から順に所定数の登録画像を抽出してもよいし、また登録画像全体のうち統合スコアが上位となる所定割合のものを抽出してもよい。また、統合スコアが予め定めた閾値以上の登録画像を抽出してもよい。もちろん、これらは一例であり、統合スコアが上位から順に候補を抽出する条件としてはこのほかの条件を用いることもできる。候補抽出部 5 0 は、このようにして抽出した各候補の文書 ID を文書画像検索部 5 2 に渡す。なお、文書画像検索部 5 2 が統合スコアを考慮した検索を行う場合には、各候補の文書 ID と統合スコアとを文書画像検索部 5 2 に渡す。

30

【 0 0 5 8 】

なお、以上に例示した統合スコアの計算方式は一例に過ぎず、他の方式で統合スコアを計算してもよい。

【 0 0 5 9 】

文書画像検索部 5 2 は、受け取った候補群のなかから入力文書画像 1 0 0 に対する類似度の高いものを、画像の投影波形を用いた照合処理により検索する。この投影波形に基づく画像照合は、本出願人による特開 2 0 0 3 - 9 1 7 3 0 号公報 (特許文献 6) に示されたものである。概説すると、図 4 に示すように、横方向の投影波形 4 1 0 は、入力文書画像 1 0 0 の各画素の値を例えば横方向に投影したものであり、言い換えれば、横方向同一行に沿った各画素の値を加算していき、各行の加算結果の列 (縦) 方向に沿った分布をとったものである。このような横方向に投影した投影波形の代わりに縦方向に投影した投影波形を画像特徴量として用いてもよいし、横方向の投影波形及び縦方向の投影波形のペアを画像特徴量として用いてもよい。なお、投影の方向は、横方向及び縦方向に限定されるわけではない。

40

【 0 0 6 0 】

投影波形を用いた類似度スコアの計算では、図 4 に示したように、入力文書画像 1 0 0

50

から求めた投影波形410と、投影波形DB54に登録された登録画像420の投影波形430との間で、例えば、縮尺、位置を合わせた上で両者の波形や微分波形の相関の強さを示す相関係数を求め、この係数を類似度スコアとする。投影波形DB54には、全登録画像の投影波形情報が登録されており、文書画像検索部52はそのうち候補抽出部50が抽出した各候補の投影波形情報をそれぞれ読み出し、入力文書画像100の投影波形410と照合するのである。縦横両方向の投影波形を用いる場合は、例えば縦方向の波形についての類似度スコアと横方向の波形についての類似度スコアの和や平均などを最終的な類似度スコアとするなどすればよい。このほか、前述の特開2003-91730号公報に示された様々な投影波形に基づく類似度の判定手法を用いることができる。

【0061】

このような投影波形ベースの照合・検索は、全般的に精度が高く、また写真画像検索部10やテキスト検索部20とは異なった特徴に着目して解析しているため、それら両検索部10及び20と組み合わせることで多様な角度からの類似性評価ができるという利点がある。しかし、その反面、演算処理の負荷が高く、演算に時間を要するか、或いは非常に高い演算性能を持つ演算装置を用いる必要があるという欠点がある。ところが、本実施形態では、あらかじめ候補抽出部50で絞り込んだ候補のみにしかこの照合を行わないので、演算時間や演算能力についての問題はさほど大きな問題とはならない。

【0062】

文書画像検索部52は、以上のようにして、各候補につき投影波形に基づく類似度スコアを求め、その類似度スコアが高い順に、各候補の文書IDとその類似度スコアとのペアを並べたリストを出力する。なお出力するリストは、類似度スコアが上位の所定数の文書に限定してもよいし、類似度スコアがあらかじめ定められた値以上となる文書に限定してもよい。

【0063】

また、文書画像検索部52が求めた投影波形に基づく類似度スコアに対し、候補抽出部50で求めた統合スコアを加味することで、第2の統合スコアを求め、この第2の統合スコアに基づき各候補の順位付けを行い、この順位に従ったリストを作成してもよい。第2の統合スコアの計算は、投影波形に基づく類似度スコアに対して上述の標準化を行い、標準化結果の値と、候補抽出部50から得た統合スコアと、を上述の方法で統合することにより求めることができる。

【0064】

検索結果出力部60は、文書画像検索部52からリストを受けとり、そのリストの登録画像を統合スコアの値の降順にソートして検索結果として出力する。これにより、入力画像aに似ている順に登録画像を配列した検索結果150が提供できる。

【0065】

以上に説明した図1のシステムでは、まず、比較的高速に処理が可能な写真画像検索部10とテキスト検索部20により全登録画像の中から、入力画像に比較的類似したものを候補として抽出する。そして、それら絞り込まれた候補に対して、比較的高精度ではあるが演算に時間を要する文書画像検索部52での検索を行うことで最終的な検索結果を求める。このような構成をとることにより、全体的な演算に要する時間や演算能力を抑えつつも、高速で精度のよい検索を実現することができる。

【0066】

本実施形態において、写真画像検索部10は、高速で、全般的に検索精度が高く、苦手な文書が比較的少ないので、図1の構成ではこれを前段の候補絞り込みに用いることで、高速な候補絞り込を実現する。また、テキスト検索部20は、比較的高速であり、テキスト量が多いと検索精度が高いので、図1の構成ではこれを前段の候補絞り込みに用いることで、高速な候補絞り込を実現している。テキスト検索部20は、テキスト量が少ない画像の検索精度は良くないが、このようなテキスト量の少ない画像については並列した写真画像検索部10の検索で救うことができる。

【0067】

10

20

30

40

50

後段の文書画像検索部 5 2 が行う投影波形に基づく照合は、登録画像に対して書き込みやスタンプなどが追加された画像が入力された場合でも、高い精度でその登録画像を検索できるという強みがある。すなわち、それら書き込みやスタンプはページ全体に対しては一般に小さいものなので、投影波形にしてしまえば、全体に対するそれら書き込み等の影響は小さくなってしまふ。したがって、そのような入力画像と登録画像とを投影波形同士で照合すれば、正しく高い類似度スコアが得られる。これに対し、写真画像検索部 1 0 が行うエッジ量分布等に基づく照合では、書き込みやスタンプの影響により、類似度スコアがかなり下がってしまう場合がある。このように、文書画像検索部 5 2の方が写真画像検索部 1 0 よりも適切に類似性の高い文書画像を検索できる場合がある。特に、ビジネス分野での文書の流出監視などの目的では、原文書にスタンプや書き込みが施されることは多々あることであり、実際に流出する文書も原文書そのものではなくそのような書き込み等がなされたものである場合が少なくない。したがって、検索機構の一部として文書画像検索部 5 2 を用いることには大きな意義がある。

10

【 0 0 6 8 】

ただし、文書画像検索部 5 2 は前述したように処理が比較的低速であるが、本実施形態のように、高速な検索部 1 0 及び 2 0 により絞り込んだ候補についてのみ処理を行うことで、そのような低速という欠点を補うことができる。

【 0 0 6 9 】

次に、図 6 を参照して、実施形態の変形例について説明する。図 6 において、図 1 に示した構成要素と同一又は類似の構成要素については、同一符号を付してその説明を省略する。

20

【 0 0 7 0 】

この変形例の画像検索装置は、写真画像検索部 1 0 とテキスト検索部 2 0 との前段に振り分け部 5 を有する。振り分け部 5 は、入力文書画像 1 0 0 を解析することで、入力文書画像 1 0 0 が写真画像検索部 1 0 とテキスト検索部 2 0 のどちらの検索に適した性質を持つ画像であるかを判定し、適した方に入力画像 1 0 0 を選択的に振り分ける。

【 0 0 7 1 】

ここで、振り分け部 5 は、例えば、複写機やスキャナなどでよく知られている自動像域分離（テキスト・イメージ分離とも呼ばれる）処理を入力文書画像 1 0 0 に適用し、その画像 1 0 0 を文字領域と（写真）画像領域とに分離する。そして、文字領域の方が画像領域よりも大きい場合にはその画像 1 0 0 をテキスト検索部 2 0 に、その逆の場合は写真画像検索部 1 0 に、それぞれ選択的に供給する。そして、候補抽出部 5 0 a は、写真画像検索部 1 0 またはテキスト検索部 2 0 のうち選択された一方が求めた各登録画像のスコアのリストを受け取り、そのスコアが上位の登録文書群を文書画像検索部 5 2 に渡す候補として抽出する。文書画像検索部 5 2 は図 1 の装置の場合と同様の処理を行えばよい。

30

【 0 0 7 2 】

なお、文字領域と画像領域の大きさの差が小さい場合には、どちらか一方の検索部 1 0 又は 2 0 のみを選択的に用いたのでは、候補絞込の精度が十分に得られない可能性がある。そこで、文字領域と画像領域の大きさの差が、あらかじめ定めた閾値以下の場合には、図 1 の装置と同様、両方の検索部 1 0 及び 2 0 に入力画像を供給し、候補抽出部 5 0 a がそれら両検索部 1 0 の求めたスコアを統合し、その結果得られる統合スコアに基づき候補を抽出するようにしてもよい。

40

【 0 0 7 3 】

以上の例では、文字領域と画像領域の大きさの比較に基づき、使用する検索部 1 0 又は 2 0 を選択したが、これは一例に過ぎない。例えば、テキスト検索部 2 0 の検索精度は、画像に含まれる文字の数に依存する部分も大きいので、振り分け部 5 が入力文書画像 1 0 0 に含まれる文字数をカウントし、その文字数の情報に基づき振り分けを行ってもよい。一般に、文字数が多いほど、テキスト検索部 2 0 の検索精度は向上する。したがって、例えば、求めた文字数が、あらかじめ定めた閾値より多ければ、テキスト検索部 2 0 を選択し、そうでなければ写真画像検索部 1 0 を選択する構成をとればよい。例えばテキスト検

50

索部 20 はテキストが少ない画像では十分な精度が得られないのに対し、写真画像検索部 10 の検索方式はテキストのみからなる画像にもある程度の精度での検索が可能なので、文字数が少ない場合に写真画像検索部 10 を選べば、ある程度の精度を確保しつつ候補の抽出ができる。

【0074】

また、文字数について第 1 の閾値と、それより小さい第 2 の閾値とを設定しておき、入力文書画像 100 中の文字数が第 1 の閾値より大きければテキスト検索部 20 を選択し、第 2 の閾値よりも小さければ写真画像検索部 10 を選択するように制御してもよい。この場合、文字数が第 1 の閾値から第 2 の閾値までの間であれば、両方の検索部 10 及び 20 に対して入力文書画像 100 を供給し、両者が求めたスコアを候補抽出部 50a で統合す

10

【0075】

なお、入力文書画像 100 中の文字数は、振り分け部 5 で OCR 等の文字認識処理を行い、その結果求められた文字の数をカウントすることで求めることができる。この段階では文字数だけ分かればよく、個々の文字が何であるかまでは分かる必要はない。したがって、この段階では、完全な文字認識処理を実行する必要はなく、個々の文字の分離（セグメンテーション）までを行えばよい。この文字の分離結果のデータをテキスト検索部 20 で利用するようにすれば、テキスト検索部 20 の文字認識部 22 は、振り分け部 5 が行う文字認識処理の部分の機能を持たなくてよくなる。なお、振り分け部 5 で完全な文字認識を行うようにしてもよく、このようにすればテキスト検索部 20 は文字認識部 22 を持た

20

【0076】

また、入力文書画像 100 に含まれる文字数のみに基づき判定する代わりに、前述の像域分離などを用いて画像領域（連続調画像の部分）のページ全体に対する割合を求め、文字数と画像領域の割合とを総合することで、検索部 10 と 20 のどちらが適切かを判定してもよい。例えば、文字数が所定の第 1 の文字数閾値より多く画像領域の割合が所定の第 1 の領域閾値より小さければ、テキスト検索部 20 が適切であると判定し、文字数が所定の第 2 の文字数閾値（第 1 の文字数閾値より小さい値である）より少なく画像領域の割合が所定の第 2 の領域閾値（第 1 の領域位置より大きい値である）より大きければ、写真画像検索部 10 が適切であると判定すればよい。そして、そのいずれの場合にも該当しない

30

【0077】

なお、この変形例では、振り分け部 5 による振り分けにより、写真画像検索部 10 及びテキスト検索部 20 のうち適切な方に入力文書画像 100 が供給されるので、写真画像検索部 10 又はテキスト検索部 20 の検索において十分に高い類似度スコアが得られる場合がある。このように、写真画像検索部 10 又はテキスト検索部 20 の検索で、あらかじめ定めた閾値（この閾値は、写真画像検索部 10 とテキスト検索部 20 とで個別に定めておく）より高い類似度スコアを持つ登録画像が求められた場合は、後段の文書画像検索部 52 による検索を省略し、その閾値以上のスコアを持つ登録画像を検索結果として出力する

40

【0078】

このようにすることにより、次のような利点がある。すなわち、投影波形に基づく検索処理は、背景画像がある場合などのように投影波形に特徴的なピークが現れない画像については検索精度が劣化する。したがって、写真画像検索部 10 又はテキスト検索部 20 で非常に高い類似度スコアが得られた登録画像でも、その登録画像が投影波形による検索に適さない種類のものであれば、文書画像検索部 52 で得られる類似度スコアが低くなってしまう場合がある。ところが、上述のように写真画像検索部 10 又はテキスト検索部 20 で非常に高い類似度スコアが得られたら、後段の文書画像検索部 52 の検索をスキップする制御を採用すれば、無駄な演算処理を行わずに済むと共に、不適切な検索技術が検索結

50

果に悪影響を与えるリスクを低減できる。

【0079】

次に、図7を参照して、画像検索装置の更なる変形例を説明する。図7において、図1に示した構成要素と同一又は類似の構成要素については、同一符号を付してその説明を省略する。

【0080】

この変形例の装置は、上述した各検索技術の特徴を考慮して、1段目に写真画像検索部10を、2段目にテキスト検索部20を、3段目に文書画像検索部52を、直列に配した構成をとった。この構成では、まず写真画像検索部10の検索により、全登録画像の中から入力文書画像100に類似した一次候補を抽出し、次にテキスト検索部20がその一次候補の各登録画像と入力文書画像100とを照合することで、一次候補の中からテキスト検索でのスコアが高い二次候補を抽出する。

10

【0081】

なお、このテキスト検索部20での検索において、テキスト検索での類似度スコアだけで絞り込むのではなく、このスコアを前段から受け取った画像特徴に基づく類似度スコアと統合し、その結果得られる統合スコアに基づき二次候補を絞り込むようにすることも好適である。このようにすれば、入力文書画像100がテキスト量の多い画像であれば、テキスト検索部20の検索により入力文書画像100により近い登録画像を優先的に抽出することができると共に、入力文書画像100がテキスト量の少ない画像であっても、それは画像特徴に基づく検索での類似度スコアを加味することで検索精度の劣化を抑えることができる。そして、最後に文書画像検索部52が、それら二次候補と入力文書画像100との間で、投影波形に基づく照合を行って、最終的な検索結果を求める。

20

【0082】

この構成では、全般的に検索精度が高く、高速処理が可能であり、苦手な文書種類が比較的少ない写真画像検索部10を1段目に配することで、まず高速に精度良く候補を絞り込むことができる。そして、2段目のテキスト検索部20は、高速処理が可能なので、二次候補の絞込も高速に行うことができる。しかも、テキスト検索部20で、1段目の類似度スコアも統合して絞込を行うようにすれば、テキスト量が少ない入力文書画像100でも検索精度の劣化を抑えることができる。そして、3段目で、書き込みやスタンプなどの重畳に強い文書画像検索部52で最後の絞込を行うことで、そのような書き込み等を考慮した上で類似度の高い登録画像を求めることができる。文書画像検索部52の処理が多少遅くても、この処理は2段階の絞込を経て十分に数が少なくなった二次候補のみにしか行わないので、処理に要する時間は大きな問題とはならない。

30

【0083】

なお、この変形例でも、1段目の写真画像検索部10である閾値以上の類似度スコアを持つ登録画像が得られた場合は、2段目以降の検索処理をスキップして検索結果を求めてもよい。同様に、2段目のテキスト検索部20で閾値(1段目の閾値とは独立に定められる)以上の類似度スコアを持つ登録画像が得られた場合は、3段目の検索処理をスキップして検索結果を求めてもよい。

【0084】

以上、本発明の実施形態及びその変形例について説明した。以上に説明した画像検索装置は、典型的には、汎用のコンピュータシステムにてそれら各部の機能又は処理内容(詳細は後述)を記述したプログラムを実行することにより実現される。コンピュータシステムは、ハードウェアとして、図8に示すように、CPU(中央演算装置)80、メモリ(一次記憶)82、各種I/O(入出力)インタフェース84等がバス86を介して接続された回路構成を有する。また、そのバス86に対し、例えばI/Oインタフェース84経由で、ハードディスクドライブ88やCDやDVD、フラッシュメモリなどの各種規格の可搬型の不揮発性記録媒体を読み取るためのディスクドライブ90が接続される。このようなドライブ88又は90は、メモリに対する外部記憶装置として機能する。実施形態の処理内容が記述されたプログラムがCDやDVD等の記録媒体を経由して、又はネットワ

40

50

ーク経由で、ハードディスクドライブ 88 等の固定記憶装置に保存され、コンピュータシステムにインストールされる。固定記憶装置に記憶されたプログラムがメモリに読み出され CPU により実行されることにより、実施形態の処理が実現される。

【0085】

ここで、写真画像検索部 10 やテキスト検索部 20 としては、既存の検索アプリケーションをそのまま利用することもできるので、本実施形態特有のプログラムとしては、入力文書画像 100 を複数の検索部 10, 20 及び 52 に渡してスコア計算を行わせる機能と、候補抽出部 50 の機能を記述したプログラムを提供すればよい。写真画像検索部 10 やテキスト検索部 20 等の個々の検索部は、そのプログラムに対してプラグインなどの形で追加できるようにしてもよい。

10

【0086】

なお、以上では、画像検索装置が 1 台のコンピュータ上に構築される例を説明したが、これはあくまで一例に過ぎない。画像検索装置を構成する上述の各構成要素がインターネットや LAN 等のネットワーク上に分散したシステム構成も本発明の範囲内である。例えば、特徴量 DB 30 及びテキスト DB 40 のうちの一方又は両方が、他の構成要素群がインストールされたコンピュータとは別のデータベース装置としてネットワーク上に分散して設けられ、写真画像検索部 10 やテキスト検索部 20 からネットワークを介してそれらデータベースを利用するようなシステム構成が考えられる。また、ネットワーク上にある特徴量 DB 30 及びテキスト DB 40 は、一台とは限らず、複数あってもよい。

【0087】

以上に説明した実施形態は、あくまで一例に過ぎず、本発明の範囲内で様々な形態をとることができる。

20

【図面の簡単な説明】

【0088】

【図 1】実施形態の画像検索装置の構成を示す機能ブロック図である。

【図 2】写真画像検索を説明するための図である。

【図 3】単語の出現位置を考慮した類似テキスト検索の説明のための図である。

【図 4】画像の投影波形に基づく類似度の算出処理を説明するための図である。

【図 5】候補抽出部の処理手順を示すフローチャートである。

【図 6】変形例の画像検索装置の構成を示す機能ブロック図である。

30

【図 7】別の変形例の画像検索装置の構成を示す機能ブロック図である。

【図 8】画像検索システムが実装されるコンピュータシステムのハードウェア構成の一例を示す図である。

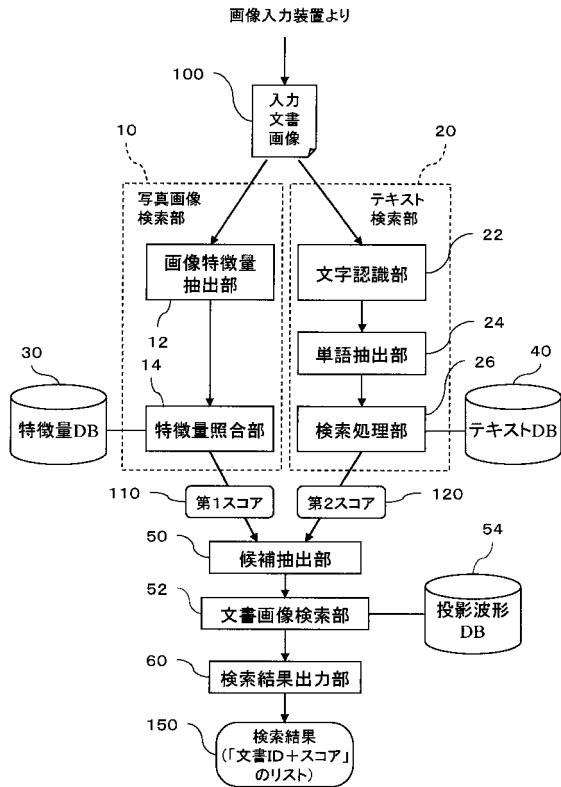
【符号の説明】

【0089】

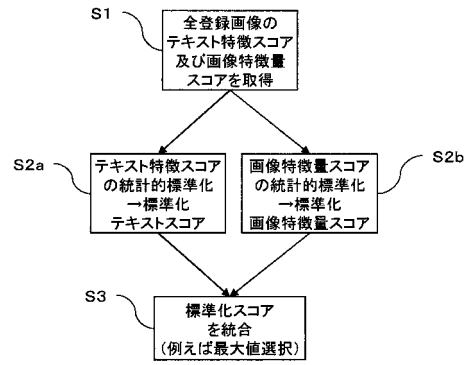
10 写真画像検索部、12 画像特徴量抽出部、14 特徴量照合部、20 テキスト検索部、22 文字認識部、24 単語抽出部、26 検索処理部、30 特徴量 DB、40 テキスト DB、50 候補抽出部、52 文書画像検索部、54 投影波形 DB、60 検索結果出力部、80 CPU、82 メモリ、84 I/O インタフェース、86 バス、88 ハードディスクドライブ、90 可搬型の不揮発性記録媒体を読み取るためのディスクドライブ、100 入力文書画像、110 第 1 スコア、120 第 2 スコア、150 検索結果。

40

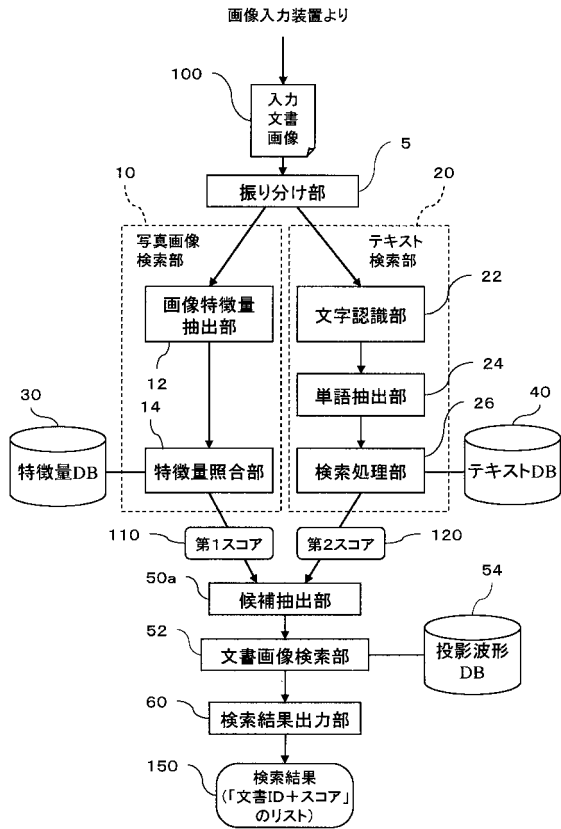
【 図 1 】



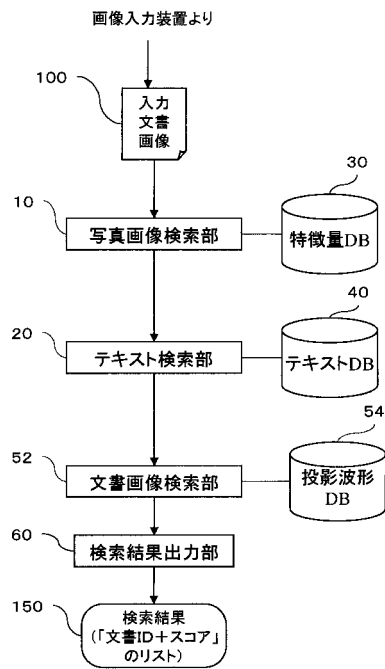
【 図 5 】



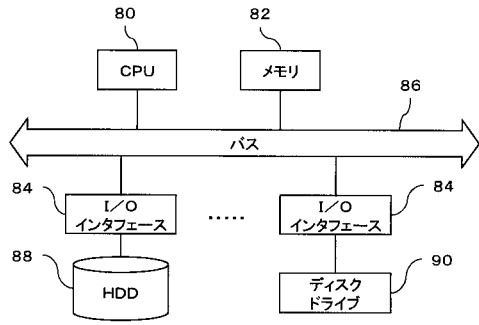
【 図 6 】



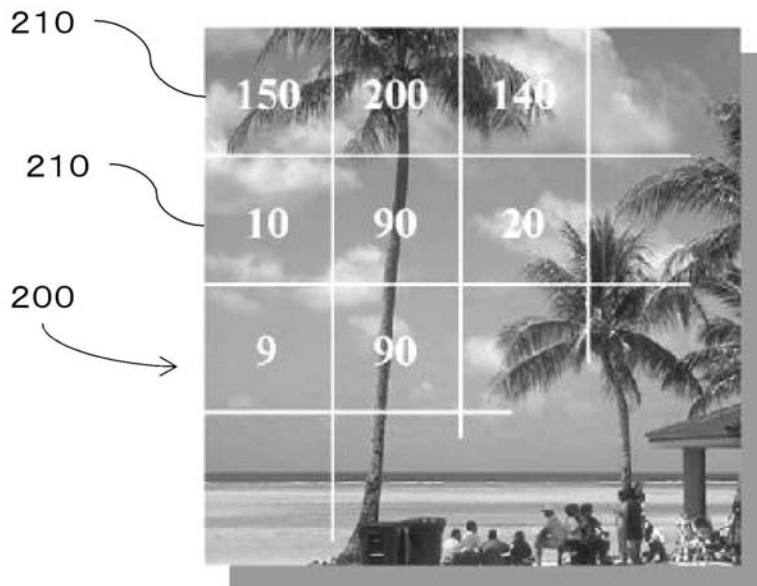
【 図 7 】



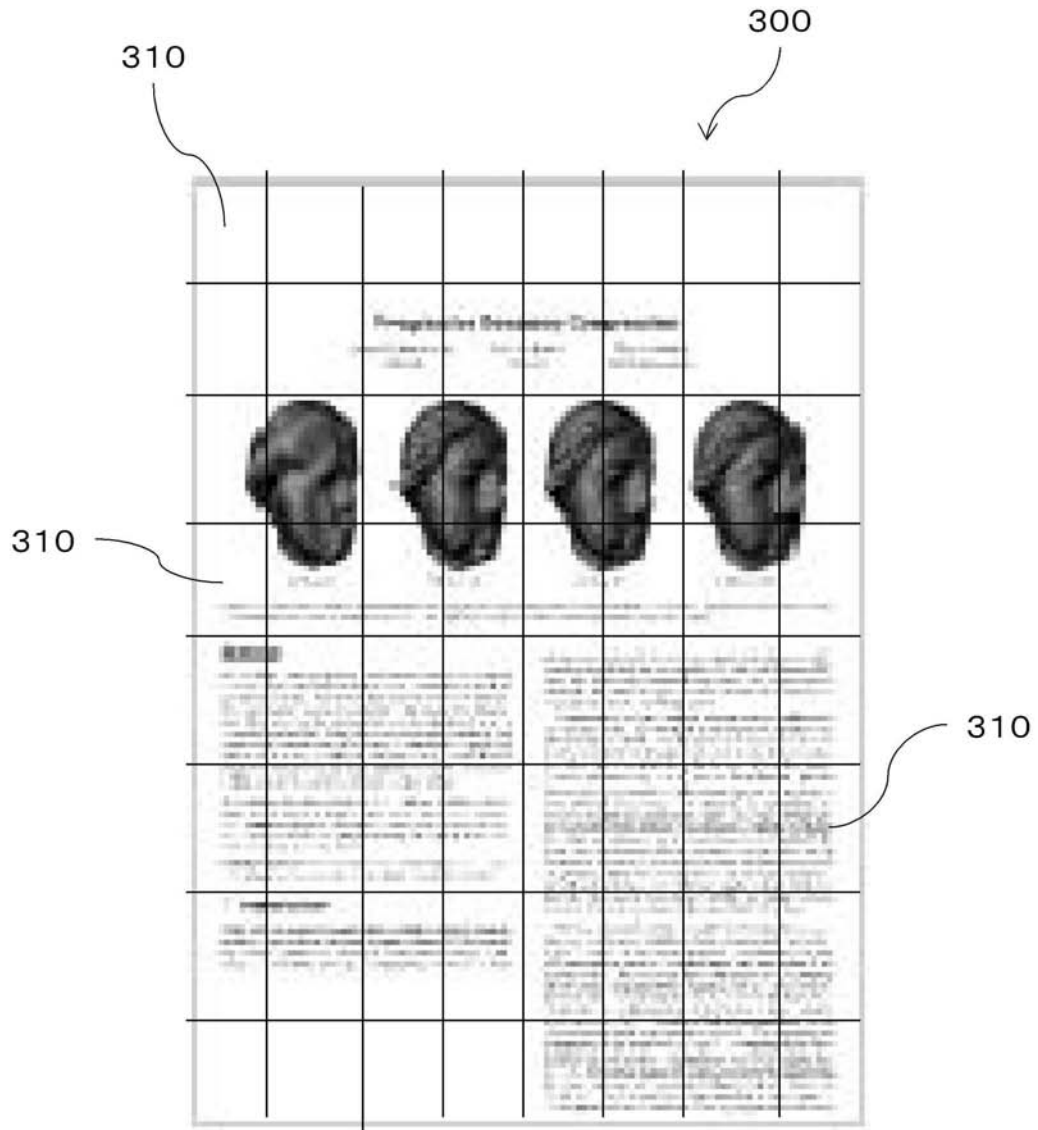
【 図 8 】



【 図 2 】



【 図 3 】



【 図 4 】

