

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2008年9月12日 (12.09.2008)

PCT

(10) 国際公開番号
WO 2008/107985 A1

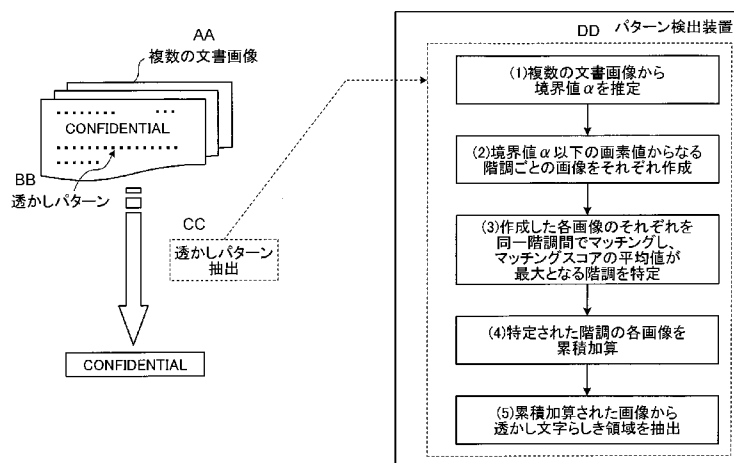
- (51) 国際特許分類:
H04N 1/387 (2006.01) H04N 1/40 (2006.01)
G06T 1/00 (2006.01)
- (21) 国際出願番号: PCT/JP2007/054465
- (22) 国際出願日: 2007年3月7日 (07.03.2007)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人 (米国を除く全ての指定国について): 富士通株式会社 (FUJITSU LIMITED) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 Kanagawa (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 藤井 勇作 (FUJII, Yusaku) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP).
- (74) 代理人: 酒井 宏明 (SAKAI, Hiroaki); 〒1006020 東京都千代田区霞が関三丁目2番5号 霞が関ビルディング 酒井国際特許事務所 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,

[続葉有]

(54) Title: PATTERN DETECTION PROGRAM, PATTERN DETECTING METHOD, AND PATTERN DETECTOR

(54) 発明の名称: パターン検出プログラム、パターン検出方法およびパターン検出装置

[図1]



AA DOCUMENT IMAGES

BB WATERMARK PATTERN

CC EXTRACT WATERMARK PATTERN

DD PATTERN DETECTOR

(1) ESTIMATE BOUNDARY VALUE α FROM DOCUMENT IMAGES(2) CREATE IMAGE OF EACH GRAY SCALE LEVEL COMPOSED OF PIXEL VALUES EQUAL TO OR LESS THAN BOUNDARY VALUE α

(3) MATCH CREATED IMAGES OF THE SAME GRAY SCALE LEVEL WITH ONE ANOTHER AND IDENTIFY GRAY SCALE LEVEL AT WHICH AVERAGE VALUE OF MATCHING SCORE TAKES ON MAXIMUM VALUE

(4) CUMULATIVELY ADD IMAGES OF IDENTIFIED GRAY SCALE LEVEL

(5) EXTRACT REGION WHICH IS PROBABLY WATERMARK CHARACTER FROM CUMULATIVELY ADDED IMAGE

(57) Abstract: A pattern detector estimates a boundary value α from document images inputted. The boundary value α is a value for distinguishing between the pixel values forming characters, diagrams and watermark patterns in document images and the pixel values forming the backgrounds. The pattern detector creates an image of each gray scale level composed of pixel values equal to or less than the boundary value α for each document image, matches the created images with the same gray scale level, identifies the gray scale level at which an average value of the matching scores is a maximum value, cumulatively adds the images of the identified gray scale level, and extracts the region which is probably a watermark character (the region where the probability of an embedded watermark character is high) from the cumulatively added image.

(57) 要約: パターン検出装置は、入力を受け付けた複数の文書画像から境界値 α を推定する。ここで、境界値 α とは、文書画像内で文字、図表および透かしパターンを構成する画素値と、背景を構成する画素値とを切り分けるための値である。次に、パターン検出装置は、各文書画像について、境界値 α 以下の画素値からなる階調ごとの画像をそれぞれ作成する。続いて、パター

ン検出装置は、作成された各画像のそれぞれを同一階調間でマッチングし、マッチングスコアの平均値が最大となる階調を特定する。そして、パター

[続葉有]



PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL,
SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, ZA, ZM, ZW.

KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE,
IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

(84) 指定国 (表示のない限り、全ての種類の広域保護が可
能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD,
SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY,

添付公開書類:
— 国際調査報告書

明 細 書

パターン検出プログラム、パターン検出方法およびパターン検出装置 技術分野

[0001] この発明は、複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パターンの検出処理をコンピュータに実行させるパターン検出プログラム、当該プログラムに対応するパターン検出方法およびパターン検出装置に関する。

背景技術

[0002] 従来より、機密情報や個人情報が入洩されることを防止するため、機密情報等であることを示すバーコード等のパターンを機密情報等が記録された紙文書内に埋め込み、閲覧時や複写時等に、そのパターンが紙文書内から検出された場合には閲覧や複写等を制限する技術が存在する。例えば、特許文献1では、機密文書の背景全体に機密文書であることを示す細かなパターンを印刷し、コピー機が複写時にそのパターンを検出した場合には複写を停止する方法が開示されている。

[0003] また、紙文書等に印刷された情報が機密情報等であることを示すパターンとして、「CONFIDENTIAL」や「複写禁止」、「社外秘」などの透かしパターン(ウォーターマーク)を紙文書の背景に印刷する技術が利用されている。

[0004] 特許文献1:特開2006-121200号公報

発明の開示

発明が解決しようとする課題

[0005] ところで、上記した従来の技術は、文書背景に埋め込まれた透かしパターンを検出することができないという問題点があった。すなわち、透かしパターンは文書背景として印刷されるので、前景文字や図表等によって部分的に隠された不完全な文字として存在することが多く、従来の技術では文書背景に埋め込まれた透かしパターンを検出することができないという問題点があった。

[0006] そこで、この発明は、上述した従来技術の課題を解決するためになされたものであり、文書背景に埋め込まれた透かしパターンを検出することが可能なパターン検出プログラム、パターン検出方法およびパターン検出装置を提供することを目的とする。

課題を解決するための手段

- [0007] 上述した課題を解決し、目的を達成するため、請求項1に係る発明は、複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パターンの検出処理をコンピュータに実行させるパターン検出プログラムであって、前記複数の文書画像の入力を受け付けて、当該複数の文書画像間で同一座標に位置する各画素の画素値の変動量から、文書画像内において文字、図表および前記共通パターンを構成する画素値を特定するための境界値を推定する境界値推定手順と、前記複数の文書画像について、前記境界値推定手順により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成する画像作成手順と、前記画像作成手順により作成された各画像について、同一の画素値で構成される各画像のそれぞれを照合して各同一画素値画像間の照合値を取得し、当該取得された各照合値の同一画素値画像集合における平均値が最大となる画素値を特定する画素値特定手順と、前記画素値特定手順により特定された画素値に対応する各画像を累積加算する画像累積手順と、前記画像累積手順により累積加算された画像から前記共通パターンが埋め込まれている画像内の領域を抽出する領域抽出手順と、をコンピュータに実行させることを特徴とする。
- [0008] また、請求項2に係る発明は、上記の発明において、前記画像作成手順は、前記複数の文書画像について、前記境界値推定手順により推定された前記境界値以下の各画素値をいくつかの階調に分割して、当該分割された階調ごとの画像をそれぞれ作成し、前記画素値特定手順は、前記画像作成手順により作成された各画像のそれぞれを同一階調間で照合して各同一階調画像間の照合値をそれぞれ取得し、当該取得された各照合値の同一階調における平均値が最大となる階調を特定し、前記画像累積手順は、前記画素値特定手順により特定された階調に対応する各画像を累積加算することを特徴とする。
- [0009] また、請求項3に係る発明は、上記の発明において、前記共通パターンを認識するための参照パターンを記憶部に記憶する参照パターン記憶手順と、前記参照パターン記憶手順により記憶部に記憶されている前記参照パターンを用いて、前記領域抽出手順により抽出された領域に埋め込まれている前記共通パターンを認識するパタ

ーン認識手順と、をさらにコンピュータに実行させることを特徴とする。

[0010] また、請求項4に係る発明は、上記の発明において、前記複数の文書画像間で同一座標に位置する各画素の画素値について、前記境界値推定手順により推定された前記境界値以上となる文書画像数に対する割合を算出し、当該境界値以上となる割合が所定の範囲内となる場合には、当該複数の文書画像間の同一座標における画素を文字、図表および前記共通パターンを構成する画素以外の背景画素として推定する背景画素推定手順をさらにコンピュータに実行させ、前記画像作成手順は、前記背景画素推定手順により前記背景画素と推定された画素を背景とみなして、前記境界値推定手順により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成することを特徴とする。

[0011] また、請求項5に係る発明は、複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パターンを検出するパターン検出方法であって、前記複数の文書画像の入力を受け付けて、当該複数の文書画像間で同一座標に位置する各画素の画素値の変動量から、文書画像内において文字、図表および前記共通パターンを構成する画素値を特定するための境界値を推定する境界値推定工程と、前記複数の文書画像について、前記境界値推定工程により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成する画像作成工程と、前記画像作成工程により作成された各画像について、同一の画素値で構成される各画像のそれぞれを照合して各同一画素値画像間の照合値を取得し、当該取得された各照合値の同一画素値画像集合における平均値が最大となる画素値を特定する画素値特定工程と、前記画素値特定工程により特定された画素値に対応する各画像を累積加算する画像累積工程と、前記画像累積工程により累積加算された画像から前記共通パターンが埋め込まれている画像内の領域を抽出する領域抽出工程と、を含んだことを特徴とする。

[0012] また、請求項6に係る発明は、上記の発明において、前記画像作成工程は、前記複数の文書画像について、前記境界値推定工程により推定された前記境界値以下の各画素値をいくつかの階調に分割して、当該分割された階調ごとの画像をそれぞれ作成し、前記画素値特定工程は、前記画像作成工程により作成された各画像のそ

れぞれを同一階調間で照合して各同一階調画像間の照合値をそれぞれ取得し、当該取得された各照合値の同一階調における平均値が最大となる階調を特定し、前記画像累積工程は、前記画素値特定工程により特定された階調に対応する各画像を累積加算することを特徴とする。

[0013] また、請求項7に係る発明は、上記の発明において、前記共通パターンを認識するための参照パターンを記憶部に記憶する参照パターン記憶工程と、前記参照パターン記憶工程により記憶部に記憶されている前記参照パターンを用いて、前記領域抽出工程により抽出された領域に埋め込まれている前記共通パターンを認識するパターン認識工程と、をさらに含んだことを特徴とする。

[0014] また、請求項8に係る発明は、上記の発明において、前記複数の文書画像間で同一座標に位置する各画素の画素値について、前記境界値推定工程により推定された前記境界値以上となる文書画像数に対する割合を算出し、当該境界値以上となる割合が所定の範囲内となる場合には、当該複数の文書画像間の同一座標における画素を文字、図表および前記共通パターンを構成する画素以外の背景画素として推定する背景画素推定工程をさらに含み、前記画像作成工程は、前記背景画素推定工程により前記背景画素と推定された画素を背景とみなして、前記境界値推定工程により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成することを特徴とする。

[0015] また、請求項9に係る発明は、複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パターンを検出するパターン検出装置であって、前記複数の文書画像の入力を受け付けて、当該複数の文書画像間で同一座標に位置する各画素の画素値の変動量から、文書画像内において文字、図表および前記共通パターンを構成する画素値を特定するための境界値を推定する境界値推定手段と、前記複数の文書画像について、前記境界値推定手段により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成する画像作成手段と、前記画像作成手段により作成された各画像について、同一の画素値で構成される各画像のそれぞれを照合して各同一画素値画像間の照合値を取得し、当該取得された各照合値の同一画素値画像集合における平均値が最大となる画素値を特定する画素値特定手段と、

前記画素値特定手段により特定された画素値に対応する各画像を累積加算する画像累積手段と、前記画像累積手段により累積加算された画像から前記共通パターンが埋め込まれている画像内の領域を抽出する領域抽出手段と、を備えたことを特徴とする。

[0016] また、請求項10に係る発明は、上記の発明において、前記画像作成手段は、前記複数の文書画像について、前記境界値推定手段により推定された前記境界値以下の各画素値をいくつかの階調に分割して、当該分割された階調ごとの画像をそれぞれ作成し、前記画素値特定手段は、前記画像作成手段により作成された各画像のそれぞれを同一階調間で照合して各同一階調画像間の照合値をそれぞれ取得し、当該取得された各照合値の同一階調における平均値が最大となる階調を特定し、前記画像累積手段は、前記画素値特定手段により特定された階調に対応する各画像を累積加算することを特徴とする。

[0017] また、請求項11の発明は、上記の発明において、前記共通パターンを認識するための参照パターンを記憶する参照パターン記憶手段と、前記参照パターン記憶手段により記憶されている前記参照パターンを用いて、前記領域抽出手段により抽出された領域に埋め込まれている前記共通パターンを認識するパターン認識手段と、をさらに備えたことを特徴とする。

[0018] また、請求項12の発明は、上記の発明において、前記複数の文書画像間で同一座標に位置する各画素の画素値について、前記境界値推定手段により推定された前記境界値以上となる文書画像数に対する割合を算出し、当該境界値以上となる割合が所定の範囲内となる場合には、当該複数の文書画像間の同一座標における画素を文字、図表および前記共通パターンを構成する画素以外の背景画素として推定する背景画素推定手段をさらに備え、前記画像作成手段は、前記背景画素推定手段により前記背景画素と推定された画素を背景とみなして、前記境界値推定手段により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成することを特徴とする。

発明の効果

[0019] 本発明によれば、複数の文書画像の入力を受け付けて、複数の文書画像間で同

一座標に位置する各画素の画素値の変動量(例えば、分散値)から、文書画像内において文字、図表および前記共通パターンを構成する画素値を特定するための境界値を推定し(例えば、各分散値の中で頻度が最大となる分散値を目安として境界値を推定し)、複数の文書画像について、推定された境界値以下(すなわち、画素濃度が濃くなる)の画素値ごとの画像をそれぞれ作成し、作成された各画像について、同一の画素値で構成される各画像のそれぞれを照合して各同一画素値画像間の照合値を取得し、取得された各照合値の同一画素値画像集合における平均値が最大となる画素値を特定し、特定された画素値に対応する各画像を累積加算し、累積加算された画像から共通パターンが埋め込まれている画像内の領域を抽出するので、複数の文書画像の背景に埋め込まれた共通パターン(例えば、機密情報等であることを示すウォーターマークなどの透かしパターン)を文書画像内から検出することが可能である。

[0020] また、本発明によれば、複数の文書画像について、推定された境界値以下の各画素値をいくつかの階調に分割して、分割された階調ごとの画像をそれぞれ作成し、作成された各画像のそれぞれを同一階調間で照合して各同一階調画像間の照合値を階調ごとにそれぞれ取得し、取得された各照合値の同一階調における平均値が最大となる階調を特定し、特定された階調に対応する各画像を累積加算するので、同一の画素値で構成される各画像のそれぞれを照合して各画像間の照合値を画素値ごとにそれぞれ算出するよりも効率的に処理を行うことが可能である。

[0021] また、本発明によれば、共通パターンを認識するための参照パターンを記憶部に記憶しておき、記憶部に記憶されている参照パターンを用いて、累積加算された文書画像から抽出された領域に埋め込まれている共通パターンを認識するので、例えば、認識の結果、文書画像に埋め込まれていた共通パターンがウォーターマークであった場合には、その文書画像の印刷やデータコピーなどの外部出力を停止したり、その文書画像を外部に出力しようとしたユーザの情報をログに保存したり、文書画像の出力先(特定のプリンタやフラッシュメモリ、CD-R)を限定したり、文書画像の出力時には責任者の承認を必要としたりする等、機密情報の漏洩防止に利用することが可能である。

[0022] また、本発明によれば、複数の文書画像間で同一座標に位置する各画素の画素値が、推定された境界値以上となる文書画像数に対する割合を座標ごとに算出し、境界値以上となる割合が所定の範囲内となる場合には、複数の文書画像間で同一座標に位置する各画素を文字、図表および共通パターンを構成する画素以外の背景画素として推定し、背景画素と推定された画素を背景とみなして、境界値以下の画素値ごとの画像をそれぞれ作成するので、文書画像内で文字、図表および共通パターンを構成する画素のみからなる画像をより鮮明に作成することが可能である。

図面の簡単な説明

[0023] [図1]図1は、実施例1に係るパターン検出装置の概要および特徴を説明するための図である。

[図2]図2は、境界値を説明するための図である。

[図3]図3は、各文書画像間で同一座標に位置する各画素値の分散値とその度数との関係を示す図である。

[図4]図4は、階調ごとに作成された画像例を示す図である。

[図5]図5は、階調ごとの画像作成からマッチングスコアの平均値算出までを模式的に示す図である。

[図6]図6は、平均マッチングスコアと各階調との関係を示す図である。

[図7]図7は、累積加算された画像例を示す図である。

[図8]図8は、累積加算された画像から抽出された領域に埋め込まれた透かしパターン例を示す図である。

[図9]図9は、実施例1に係るパターン検出装置の構成を示すブロック図である。

[図10]図10は、実施例1に係るパターン検出装置による処理の流れを示すフローチャートである。

[図11]図11は、パターン検出プログラムを実行するコンピュータを示す図である。

符号の説明

[0024] 10 パターン検出装置

11 記憶部

11a 文書画像記憶部

- 11b 参照パターン記憶部
- 12 制御部
 - 12a 画像入力受付部
 - 12b 境界値推定部
 - 12c 画像作成部
 - 12d 画像マッチング部
 - 12e 領域抽出部
 - 12f パターン認識部
- 20 コンピュータ
 - 21 HDD(Hard Disk Drive)
 - 21a 文書画像データテーブル
 - 21b 参照パターンデータテーブル
 - 22 RAM(Random Access Memory)
 - 22a 文書画像データ
 - 22b 参照パターンデータ
 - 23 ROM(Read Only Memory)
 - 23a 画像入力受付プログラム
 - 23b 境界値推定プログラム
 - 23c 画像作成プログラム
 - 23d 画像マッチングプログラム
 - 23e 領域抽出プログラム
 - 23f パターン認識プログラム
 - 24 CPU(Central Processing Unit)
 - 24a 画像入力受付プロセス
 - 24b 境界値推定プロセス
 - 24c 画像作成プロセス
 - 24d 画像マッチングプロセス
 - 24e 領域抽出プロセス

24f パターン認識プロセス

30 バス

発明を実施するための最良の形態

[0025] 以下に添付図面を参照して、この発明に係るパターン検出プログラム、パターン検出方法およびパターン検出装置の実施例を詳細に説明する。なお、以下では、本発明に係るパターン検出装置を実施例1として説明した後に、本発明に含まれる他の実施例を説明する。

実施例 1

[0026] 以下の実施例1では、実施例1に係るパターン検出装置の概要および特徴、パターン検出装置の構成および処理を順に説明し、最後に実施例1による効果を説明する。なお、以下の実施例1では、複数の文書画像の背景に埋め込まれた複数の文書画像に共通する共通パターンとして、文書画像の内容が機密情報であることを示す透かしパターンを例に挙げた場合を説明する。

[0027] [パターン検出装置の概要および特徴(実施例1)]

まず、図1～図8を用いて、実施例1に係るパターン検出装置の概要および特徴を説明する。図1は、実施例1に係るパターン検出装置の概要および特徴を説明するための図である。図2は、境界値を説明するための図である。図3は、各文書画像間で同一座標に位置する各画素値の分散値とその度数との関係を示す図である。図4は、階調ごとに作成された画像例を示す図である。図5は、階調ごとの画像作成からマッチングスコアの平均値算出までを模式的に示す図である。図6は、平均マッチングスコアと各階調との関係を示す図である。図7は、累積加算された画像例を示す図である。図8は、累積加算された画像から抽出された領域に埋め込まれた透かしパターン例を示す図である。

[0028] 実施例1に係るパターン検出装置は、複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パターンを検出することを概要とするが、文書背景に埋め込まれた透かしパターン(例えば、「CONFIDENTIAL」や「複写禁止」など)を検出することが可能である点に主たる特徴がある。

[0029] この主たる特徴について具体的に説明すると、実施例1に係るパターン検出装置は

、図1に示すように、入力を受け付けた複数の文書画像から境界値 α を推定する(図1の(1)参照)。ここで、境界値 α とは、図2に例示するように、文書画像内で文字、図表および透かしパターンを構成する画素値と、背景を構成する画素値とを切り分けるための値である。

[0030] また、境界値 α の推定の仕方について具体的に説明する。実施例1に係るパターン検出装置は、複数の文書画像間で同一座標に位置する各画素値の変動に関する分散値をそれぞれ算出し、図3に例示するように、座標ごとに算出された各分散値の中で度数が最大となる分散値を求め、その分散値を持つ各画素の画素値の平均を境界値 α と推定する。

[0031] 次に、実施例1に係るパターン検出装置は、各文書画像について、境界値 α 以下の画素値からなる階調ごとの画像をそれぞれ作成する(図1の(2)参照)。具体的に説明すると、推定された境界値を仮に“255”とすると、例えば、境界値 α 以下の画素値“254～0”までを16階調に分割して、各文書画像について階調ごとの画像をそれぞれ作成する(例えば、図4参照)。

[0032] 階調ごとの画像をそれぞれ作成した後、実施例1に係るパターン検出装置は、作成された各画像のそれぞれを同一階調間でマッチングし、マッチングスコアの平均値が最大となる階調を特定する(図1の(3)参照)。

[0033] 具体的に説明すると、実施例1に係るパターン検出装置は、図5に例示するように、Nページからなる各文書画像について、0階調～15階調からなる16階調ごとの画像をそれぞれ作成し、作成された各画像のそれぞれを同一階調間(例えば、0階調や15階調などの同一階調間)でマッチングして、取得された各マッチングスコアの平均値を算出する。例えば、10ページの文書画像である場合には、同一階調間でマッチングする一組の文書画像の組み合わせ数は45通りとなり、取得されるマッチングスコアの総数は45となる。そして、図6に例示するように、算出したマッチングスコアの平均値が最大となる階調(例えば、14階調)を特定する。すなわち、透かし文字は、通常、各文書画像の同じ位置に埋め込まれているので、透かし文字が文書画像内に鮮明に現れている階調のマッチングスコアの平均値は高くなることが予想されることを利用する趣旨である。

- [0034] マッチングスコアの平均値が最大となる階調を特定した後、実施例1に係るパターン検出装置は特定された階調の各画像を累積加算する(図1の(4)参照)。例えば、特定された階調が14階調である場合には、マッチングスコアが最大となる位置で14階調の各画像を累積加算する。すなわち、文書画像に埋め込まれた透かし文字を強調する趣旨である(図7参照)。また、累積加算により、文書画像内の前景文字等で欠けの発生した透かしパターンの欠けの部分が補間される。
- [0035] そして、実施例1に係るパターン検出装置は、累積加算された画像から透かし文字らしき領域(透かし文字が埋め込まれている蓋然性の高い領域)を抽出する(図1の(5)参照)。例えば、累積加算された画像を透かし文字とそれ以外とに2値化して、透かし文字が埋め込まれた最小矩形領域を画像から抽出する(図8参照)。
- [0036] このようなことから、実施例1に係るパターン検出装置は、上述した主たる特徴のごとく、文書背景として印刷された透かしパターンを検出することが可能である。
- [0037] [パターン検出装置の構成(実施例1)]
- 次に、図9を用いて、実施例1に係るパターン検出装置の構成を説明する。図9は、実施例1に係るパターン検出装置の構成を示すブロック図である。なお、同図では、実施例1に係るパターン検出装置を説明する上で必要となる処理部のみを記載し、その他の処理部については記載を省略する。
- [0038] 同図に示すように、実施例1に係るパターン検出装置10は、記憶部11と、制御部12とから構成される。
- [0039] 記憶部11は、制御部12による各種処理に必要なデータおよびプログラムを記憶する記憶部であり、特に本発明に密接に関連するものとしては、文書画像記憶部11aおよび参照パターン記憶部11bをそれぞれ備える。
- [0040] このうち、文書画像記憶部11aは、文書画像に関する各種の情報を記憶する記憶部であり、後述する画像入力受付部12aから格納された各文書画像をそれぞれ記憶して構成される。ここで、文書画像とは、例えば、紙文書をスキャナなどで取り込んで得られるJPEG画像や、画像表示イメージをそのまま画像化した電子データのことをいう。
- [0041] 参照パターン記憶部11bは、文書画像から抽出された透かし文字らしき領域を認

識するための参照パターンの情報を記憶する記憶部である。

- [0042] 制御部12は、所定の制御プログラム、各種の処理手順などを規定したプログラムおよび所要データを格納するための内部メモリを有し、これらによって種々の処理を実行する処理部であり、特に本発明に密接に関連するものとしては、画像入力受付部12aと、境界値推定部12bと、画像作成部12cと、画像マッチング部12dと、画像累積部12eと、領域抽出部12fと、パターン認識部12gとを備える。
- [0043] このうち、画像入力受付部12aは、複数の文書画像に関する情報の入力を受け付ける処理部であり、例えば、図には示していない画像読取部などにより読取られた文書画像の入力を受け付けると、受け付けた文書画像に関する情報を文書画像記憶部11aに格納する。
- [0044] 境界値推定部12bは、文書画像記憶部11aから読み出した複数の文書画像から境界値 α を推定する処理部である。ここで、境界値 α とは、文書画像内で文字、図表および透かしパターンを構成する画素値と、背景を構成する画素値とを切り分けるための値である(図2参照)。
- [0045] 境界値 α の推定の仕方について具体的に説明すると、境界値推定部12bは、複数の文書画像間で同一座標に位置する各画素の画素値の変動に関する分散値をそれぞれ算出し、算出された各分散値の中で度数が最大となる分散値を求め、その分散値を持つ各画素の画素値の平均を境界値 α と推定する(図3参照)。
- [0046] 画像作成部12cは、各文書画像について、境界値 α 以下の画素値からなる階調ごとの画像をそれぞれ作成する処理部である。具体的に説明すると、境界値推定部12bにより推定された境界値を仮に“255”とすると、例えば、境界値 α 以下の画素値“254～0”までを16階調に分割して、各文書画像について階調ごとの画像をそれぞれ作成する(図4参照)。
- [0047] なお、画像作成部12cは、各文書画像について、境界値 α 以下の画素値からなる階調ごとの画像をそれぞれ作成する場合に限られるものではなく、例えば、境界値 α 以下の画素値が“254～0”である場合には、各文書画像について画素値ごとの255枚の画像をそれぞれ作成するようにしてもよい。
- [0048] 画像マッチング部12dは、画像作成部12cにより作成された各画像のそれぞれを同

一階調間でマッチングし、マッチングスコアの平均値が最大となる階調を特定する処理部である。

[0049] 具体的に説明すると、画像マッチング部12dは、Nページからなる各文書画像について、0階調～15階調からなる16階調ごとの画像が画像作成部12cによりそれぞれ作成された場合には、作成された各画像のそれぞれを同一階調間（例えば、0階調同士や15階調同士などの同一階調間）でマッチングして、階調ごとに取得された各マッチングスコアの平均値を算出する（図5参照）。例えば、10ページの文書画像である場合には、同一階調間でマッチングする一組の文書画像の組み合わせ数は45通りとなり、取得されるマッチングスコアの総数は45となる。そして、算出したマッチングスコアの平均値が最大となる階調（例えば、14階調）を特定する（図6参照）。すなわち、透かし文字は、通常、各文書画像の同じ位置に埋め込まれているので、透かし文字が文書画像内に鮮明に現れている階調のマッチングスコアの平均値は高くなることが予想されることを利用する趣旨である。

[0050] また、画像マッチング部12dによるパターンマッチングの方法としては、マッチトフィルタや位相相関などの従来技術を採用する。

[0051] なお、画像マッチング部12dは、画像作成部12cにより各文書画像について画素値ごとの画像がそれぞれ作成された場合には、同一の画素値で構成される各画像（例えば、各文書画像について作成された画素値“254”の各画像等）のそれぞれを照合して各画像間のマッチングスコアを画素値ごとにそれぞれ取得し、取得された各マッチングスコアの平均値が最大となる画素値を特定するようにしてもよい。

[0052] 画像累積部12eは、画像マッチング部12dにより特定された階調の各画像を累積加算する処理部である。例えば、画像マッチング部12dにより特定された階調が14階調である場合には、14階調の各画像を累積加算する。すなわち、文書画像に埋め込まれた透かし文字を強調する趣旨である（図7参照）。また、累積加算により、文書画像内の前景文字等で欠けの発生した透かしパターンの欠けの部分が補間される。なお、画像を累積加算する場合には、画像マッチング部12dにおいてマッチング時に合わせて取得しておいた並進移動量に基づいて、各画像の位置あわせを行ってから累積加算する。なお、画像累積部12eは、画像マッチング部12dにより画素値

が特定された場合には、その画素値の各画像を累積加算するようにしてもよい。

[0053] 領域抽出部12fは、画像累積部12eにより累積加算された画像から透かし文字らしき領域(透かし文字が埋め込まれている蓋然性の高い領域)を抽出する処理部である。例えば、累積加算された画像を透かし文字とそれ以外とに2値化して、透かし文字が埋め込まれた最小矩形領域を画像から抽出する(図8参照)。

[0054] パターン認識部12gは、領域抽出部12fにより抽出された透かし文字らしき領域を認識する処理部である。具体的に説明すると、参照パターン記憶部11bに記憶されている参照パターンを用いて、領域抽出部12fにより抽出された透かし文字らしき領域に参照パターンで示される透かし文字が埋め込まれているか否か認識する。

[0055] [パターン検出装置による処理(実施例1)]

続いて、図10を用いて、実施例1に係るパターン検出装置による処理を説明する。図10は、実施例1に係るパターン検出装置による処理の流れを示すフローチャートである。なお、以下では、画像入力受付部12aにより、複数の文書画像の入力を受け付けられた後の処理の流れについて説明する。

[0056] 同図に示すように、境界値推定部12bは、文書画像記憶部11aから読み出した複数の文書画像から境界値 α を推定する(ステップS1001)。ここで、境界値 α とは、文書画像内で文字、図表および透かしパターンを構成する画素値と、背景を構成する画素値とを切り分けるための値である(図2参照)。

[0057] 境界値 α の推定の仕方について具体的に説明すると、境界値推定部12bは、複数の文書画像間で同一座標に位置する各画素の画素値の変動に関する分散値をそれぞれ算出し、算出された各分散値の中で度数が最大となる分散値を求め、その分散値を持つ各画素の画素値の平均を境界値 α と推定する(図3参照)。

[0058] 次に、画像作成部12cは、各文書画像について、境界値 α 以下の画素値からなる階調ごとの画像をそれぞれ作成する(ステップS1002)。具体的に説明すると、境界値推定部12bにより推定された境界値を仮に“255”とすると、例えば、境界値 α 以下の画素値“254～0”までを16階調に分割して、各文書画像について階調ごとの画像をそれぞれ作成する(図4参照)。

[0059] 続いて、画像マッチング部12dは、画像作成部12cにより作成された各画像のそれ

それを同一階調間でマッチングし、マッチングスコアの平均値が最大となる階調を特定する(ステップS1003)。

[0060] 具体的に説明すると、画像マッチング部12dは、Nページからなる各文書画像について、0階調～15階調からなる16階調ごとの画像が画像作成部12cによりそれぞれ作成された場合には、作成された各画像のそれぞれを同一階調間(例えば、0階調同士や15階調同士などの同一階調間)でマッチングして、階調ごとに取得された各マッチングスコアの平均値を算出する(図5参照)。例えば、10ページの文書画像である場合には、同一階調間でマッチングする一組の文書画像の組み合わせ数は45通りとなり、取得されるマッチングスコアの総数は45となる。そして、算出したマッチングスコアの平均値が最大となる階調(例えば、14階調)を特定する(図6参照)。すなわち、透かし文字は、通常、各文書画像の同じ位置に埋め込まれているので、透かし文字が文書画像内に鮮明に現れている階調のマッチングスコアの平均値は高くなることが予想されることを利用する趣旨である。

[0061] そして、画像累積部12eは、画像マッチング部12dにより特定された階調の各画像を累積加算する(ステップS1004)。例えば、画像マッチング部12dにより特定された階調が14階調である場合には、14階調の各画像を累積加算する。すなわち、文書画像に埋め込まれた透かし文字を強調する趣旨である(図7参照)。また、累積加算により、文書画像内の前景文字等で欠けの発生した透かしパターンの欠けの部分が補間される。なお、画像を累積加算する場合には、画像マッチング部12dにおいてマッチング時に合わせて取得しておいた並進移動量に基づいて、各画像の位置あわせを行ってから累積加算する。なお、画像累積部12eは、画像マッチング部12dにより画素値が特定された場合には、その画素値の各画像を累積加算するようにしてもよい。

[0062] 続いて、領域抽出部12fは、画像累積部12eにより累積加算された画像から透かし文字らしき領域(透かし文字が埋め込まれている蓋然性の高い領域)を抽出する(ステップS1005)。例えば、累積加算された画像を透かし文字とそれ以外とに2値化して、透かし文字が埋め込まれた最小矩形領域を画像から抽出する(図8参照)。

[0063] 最後に、パターン認識部12gは、領域抽出部12fにより抽出された透かし文字らし

き領域を認識する(ステップS1006)。具体的に説明すると、参照パターン記憶部11bに記憶されている参照パターンを用いて、領域抽出部12fにより抽出された透かし文字らしき領域に参照パターンで示される透かし文字が埋め込まれているか否か認識する。

[0064] [実施例1による効果]

上述してきたように、実施例1によれば、複数の文書画像の入力を受け付けて、複数の文書画像間で同一座標に位置する各画素の画素値の分散値をそれぞれ算出し、算出された各分散値の中で頻度が最大となる分散値を目安として、文書画像内で文字、図表および透かしパターンを構成する画素値を特定するための境界値 α を推定し、複数の文書画像について、推定された境界値 α 以下の画素値ごとの画像をそれぞれ作成し、作成された各画像について、同一の画素値で構成される各画像のそれぞれを照合して各画像間のマッチングスコアを画素値ごとにそれぞれ取得し、取得された各マッチングスコアの同一画素値画像集合における平均値が最大となる画素値を特定し、特定された画素値に対応する各画像を累積加算し、累積加算された画像から透かしパターンが埋め込まれている画像内の領域を抽出するので、複数の文書画像の背景に埋め込まれた透かしパターン(例えば、機密情報等であることを示すウォーターマークなどの透かしパターン)を文書画像内から検出することが可能である。

[0065] また、実施例1によれば、複数の文書画像について、推定された境界値 α 以下の各画素値をいくつかの階調に分割して、分割された階調ごとの画像をそれぞれ作成し、作成された各画像のそれぞれを同一階調間で照合してマッチングスコアを取得し、取得された各マッチングスコアの同一階調内での平均値を求め、それが最大となる階調を特定し、特定された階調に対応する各画像を累積加算するので、同一の画素値で構成される各画像のそれぞれを照合して各画像間のマッチングスコアを画素値ごとにそれぞれ算出するよりも効率的に処理を行うことが可能である。

[0066] また、実施例1によれば、透かしパターンを認識するための参照パターンを記憶部に記憶しておき、記憶部に記憶されている参照パターンを用いて、累積加算された文書画像から抽出された領域に埋め込まれている透かしパターンを認識するので、

例えば、認識の結果、文書画像に埋め込まれていた共通パターンがウォーターマークであった場合には、その文書画像の印刷やデータコピーなどの外部出力を停止したり、その文書画像を外部に出力しようとしたユーザの情報をログに保存したり、文書画像の出力先(特定のプリンタやフラッシュメモリ、CD-R)を限定したり、文書画像の出力時には責任者の承認を必要としたりする等、機密情報の漏洩防止に利用することが可能である。

実施例 2

[0067] さて、これまで本発明の実施例1について説明したが、本発明は上述した実施例以外にも、種々の異なる形態にて実施されてよいものである。そこで、以下では、本発明の含まれる他の実施例を説明する。

[0068] (1)背景画素の推定

上記の実施例1において、複数の文書画像間で同一座標に位置する各画素値が、推定された境界値以上となる割合を座標ごとに算出し、境界値以上となる割合が所定の範囲内となる場合(すなわち、各文書画像の各画素が背景もしくは文字、図表のいずれかに該当する場合)には、複数の文書画像間で同一座標に位置する各画素を文字、図表および透かしパターンを構成する画素以外の背景画素として推定し、背景画素と推定された画素を背景とみなして、境界値以下の画素値ごとの画像をそれぞれ作成するようにしてもよい。

[0069] このようにすれば、文書画像内で文字、図表および透かしパターンを構成する画素のみからなる画像をより鮮明に作成することが可能である。

[0070] (2)装置構成等

また、図9に示したパターン検出装置10の各構成要素は機能概念的なものであり、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、パターン検出装置10の分散・統合の具体的形態は図示のものに限られず、例えば、画像マッチング部12dと領域抽出部12fとを統合するなど、その全部または一部を、各種の負荷や使用状況などに応じて、任意の単位で機能的または物理的に分散・統合して構成することができる。さらに、パターン検出装置10にて行なわれる各処理機能(例えば、図10参照)は、その全部または任意の一部が、CPUおよび当該CPUにて解析実行

されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実現され得る。

[0071] (3)パターン検出プログラム

ところで、上記の実施例1で説明した各種の処理(例えば、図10参照)は、あらかじめ用意されたプログラムをパーソナルコンピュータやワークステーションなどのコンピュータシステムで実行することによって実現することができる。そこで、以下では、図11を用いて、上記の実施例1と同様の機能を有するパターン検出プログラムを実行するコンピュータの一例を説明する。図11は、パターン検出プログラムを実行するコンピュータを示す図である。

[0072] 同図に示すように、パターン検出装置としてコンピュータ20は、HDD21、RAM22、ROM23およびCPU24をバス30で接続して構成される。

[0073] そして、ROM23には、上記の実施例1に示したパターン検出装置と同様の機能を発揮するパターン検出プログラム、つまり、図11に示すように、画像入力受付プログラム23a、境界値推定プログラム23b、画像作成プログラム23c、画像マッチングプログラム23d、画像累積プログラム23e、領域抽出プログラム23fおよびパターン認識プログラム23gがあらかじめ記憶されている。なお、これらのプログラム23a、23b、23c、23d、23e、23fおよび23gについては、図9に示したパターン検出装置10の各構成要素と同様、適宜統合または分散してもよい。なお、ROM23は、不揮発性の「RAM」でもよい。

[0074] そして、CPU24が、これらのプログラム23a、23b、23c、23d、23e、23fおよび23gをROM23から読み出して実行することで、図11に示すように、各プログラム23a、23b、23c、23d、23e、23fおよび23gは、画像入力受付プロセス24a、境界値推定プロセス24b、画像作成プロセス24c、画像マッチングプロセス24d、画像累積プロセス24e、領域抽出プロセス24fおよびパターン認識プロセス24gとして機能するようになる。なお、各プロセス24a、24b、24c、24d、24e、24fおよび24gは、図9に示したパターン検出装置10の画像入力受付部12a、境界値推定部12b、画像作成部12c、画像マッチング部12d、画像累積部12e、領域抽出部12fおよびパターン認識部12gにそれぞれ対応する。

[0075] また、HDD21には、図11に示すように、文書画像データテーブル21aおよび参照パターンデータテーブル21bがそれぞれ設けられる。なお、文書画像データテーブル21aおよび参照パターンデータテーブル21bは、図9に示した文書画像記憶部11aおよび参照パターン記憶部11bにそれぞれ対応する。そして、CPU24は、文書画像データテーブル21aおよび参照パターンデータテーブル21bから、文書画像データ22aおよび参照パターンデータ22bをそれぞれ読み出してRAM22に格納し、RAM22に格納された文書画像データ22aおよび参照パターンデータ22bに基づいて処理を実行する。

[0076] なお、上記した各プログラム23a、23b、23c、23d、23e、23fおよび23gについては、必ずしも最初からROM23に記憶させておく必要はなく、例えば、コンピュータ20に挿入されるフレキシブルディスク(FD)、CD-ROM、DVDディスク、光磁気ディスク、ICカードなどの「可搬用の物理媒体」、または、コンピュータ20の内外に備えられるHDDなどの「固定用の物理媒体」、さらには、公衆回線、インターネット、LAN、WANなどを介してコンピュータ20に接続される「他のコンピュータ(またはサーバ)」などに各プログラムを記憶させておき、コンピュータ20がこれらから各プログラムを読み出して実行するようにしてもよい。

産業上の利用可能性

[0077] 以上のように、本発明に係るパターン検出プログラム、パターン検出方法およびパターン検出装置は、複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パターンの検出処理をコンピュータに実行させる場合等に有用であり、特に、文書背景に埋め込まれた透かしパターンを検出することに適する。

請求の範囲

- [1] 複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パターンの検出処理をコンピュータに実行させるパターン検出プログラムであって、
- 前記複数の文書画像の入力を受け付けて、当該複数の文書画像間で同一座標に位置する各画素の画素値の変動量から、文書画像内において文字、図表および前記共通パターンを構成する画素値を特定するための境界値を推定する境界値推定手順と、
- 前記複数の文書画像について、前記境界値推定手順により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成する画像作成手順と、
- 前記画像作成手順により作成された各画像について、同一の画素値で構成される各画像のそれぞれを照合して各同一画素値画像間の照合値を取得し、当該取得された各照合値の同一画素値画像集合における平均値が最大となる画素値を特定する画素値特定手順と、
- 前記画像特定手順により特定された画素値に対応する各画像を累積加算する画像累積手順と、
- 前記画素値累積手順により累積加算された画像から前記共通パターンが埋め込まれている画像内の領域を抽出する領域抽出手順と、
- をコンピュータに実行させることを特徴とするパターン検出プログラム。
- [2] 前記画像作成手順は、前記複数の文書画像について、前記境界値推定手順により推定された前記境界値以下の各画素値をいくつかの階調に分割して、当該分割された階調ごとの画像をそれぞれ作成し、
- 前記画素値特定手順は、前記画像作成手順により作成された各画像のそれぞれを同一階調間で照合して各同一階調画像間の照合値をそれぞれ取得し、当該取得された各照合値の同一階調における平均値が最大となる階調を特定し、
- 前記画像累積手順は、前記画素値特定手順により特定された階調に対応する各画像を累積加算することを特徴とする請求項1に記載のパターン検出プログラム。
- [3] 前記共通パターンを認識するための参照パターンを記憶部に記憶する参照パターン記憶手順と、

前記参照パターン記憶手順により記憶部に記憶されている前記参照パターンを用いて、前記領域抽出手順により抽出された領域に埋め込まれている前記共通パターンを認識するパターン認識手順と、

をさらにコンピュータに実行させることを特徴とする請求項1または2に記載のパターン検出プログラム。

- [4] 前記複数の文書画像間で同一座標に位置する各画素の画素値について、前記境界値推定手順により推定された前記境界値以上となる文書画像数に対する割合を算出し、当該境界値以上となる割合が所定の範囲内となる場合には、当該複数の文書画像間の同一座標における画素を文字、図表および前記共通パターンを構成する画素以外の背景画素として推定する背景画素推定手順をさらにコンピュータに実行させ、

前記画像作成手順は、前記背景画素推定手順により前記背景画素と推定された画素を背景とみなして、前記境界値推定手順により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成することを特徴とする請求項1に記載のパターン検出プログラム。

- [5] 複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パターンを検出するパターン検出方法であって、

前記複数の文書画像の入力を受け付けて、当該複数の文書画像間で同一座標に位置する各画素の画素値の変動量から、文書画像内において文字、図表および前記共通パターンを構成する画素値を特定するための境界値を推定する境界値推定工程と、

前記複数の文書画像について、前記境界値推定工程により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成する画像作成工程と、

前記画像作成工程により作成された各画像について、同一の画素値で構成される各画像のそれぞれを照合して各同一画素値画像間の照合値を取得し、当該取得された各照合値の同一画素値画像集合における平均値が最大となる画素値を特定する画素値特定工程と、

前記画素値特定工程により特定された画素値に対応する各画像を累積加算する

画像累積工程と、

前記画像累積工程により累積加算された画像から前記共通パターンが埋め込まれている画像内の領域を抽出する領域抽出工程と、

を含んだことを特徴とするパターン検出方法。

- [6] 前記画像作成工程は、前記複数の文書画像について、前記境界値推定工程により推定された前記境界値以下の各画素値をいくつかの階調に分割して、当該分割された階調ごとの画像をそれぞれ作成し、

前記画素値特定工程は、前記画像作成工程により作成された各画像のそれぞれを同一階調間で照合して各同一階調画像間の照合値をそれぞれ取得し、当該取得された各照合値の同一階調における平均値が最大となる階調を特定し、

前記画像累積工程は、前記画素値特定工程により特定された階調に対応する各画像を累積加算することを特徴とする請求項5に記載のパターン検出方法。

- [7] 前記共通パターンを認識するための参照パターンを記憶部に記憶する参照パターン記憶工程と、

前記参照パターン記憶工程により記憶部に記憶されている前記参照パターンを用いて、前記領域抽出工程により抽出された領域に埋め込まれている前記共通パターンを認識するパターン認識工程と、

をさらに含んだことを特徴とする請求項5または6に記載のパターン検出方法。

- [8] 前記複数の文書画像間で同一座標に位置する各画素の画素値について、前記境界値推定工程により推定された前記境界値以上となる文書画像数に対する割合を算出し、当該境界値以上となる割合が所定の範囲内となる場合には、当該複数の文書画像間の同一座標における画素を文字、図表および前記共通パターンを構成する画素以外の背景画素として推定する背景画素推定工程をさらに含み、

前記画像作成工程は、前記背景画素推定工程により前記背景画素と推定された画素を背景とみなして、前記境界値推定工程により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成することを特徴とする請求項5に記載のパターン検出方法。

- [9] 複数の文書画像の背景に埋め込まれた当該複数の文書画像に共通する共通パタ

ーンを検出するパターン検出装置であって、

前記複数の文書画像の入力を受け付けて、当該複数の文書画像間で同一座標に位置する各画素の画素値の変動量から、文書画像内において文字、図表および前記共通パターンを構成する画素値を特定するための境界値を推定する境界値推定手段と、

前記複数の文書画像について、前記境界値推定手段により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成する画像作成手段と、

前記画像作成手段により作成された各画像について、同一の画素値で構成される各画像のそれぞれを照合して各同一画素値画像間の照合値を取得し、当該取得された各照合値の同一画素値画像集合における平均値が最大となる画素値を特定する画素値特定手段と、

前記画素値特定手段により特定された画素値に対応する各画像を累積加算する画像累積手段と、

前記画像累積手段により累積加算された画像から前記共通パターンが埋め込まれている画像内の領域を抽出する領域抽出手段と、

を備えたことを特徴とするパターン検出装置。

- [10] 前記画像作成手段は、前記複数の文書画像について、前記境界値推定手段により推定された前記境界値以下の各画素値をいくつかの階調に分割して、当該分割された階調ごとの画像をそれぞれ作成し、

前記画素値特定手段は、前記画像作成手段により作成された各画像のそれぞれを同一階調間で照合して各同一階調画像間の照合値をそれぞれ取得し、当該取得された各照合値の同一階調における平均値が最大となる階調を特定し、

前記画像累積手段は、前記画素値特定手段により特定された階調に対応する各画像を累積加算することを特徴とする請求項9に記載のパターン検出装置。

- [11] 前記共通パターンを認識するための参照パターンを記憶する参照パターン記憶手段と、

前記参照パターン記憶手段により記憶されている前記参照パターンを用いて、前記領域抽出手段により抽出された領域に埋め込まれている前記共通パターンを認識

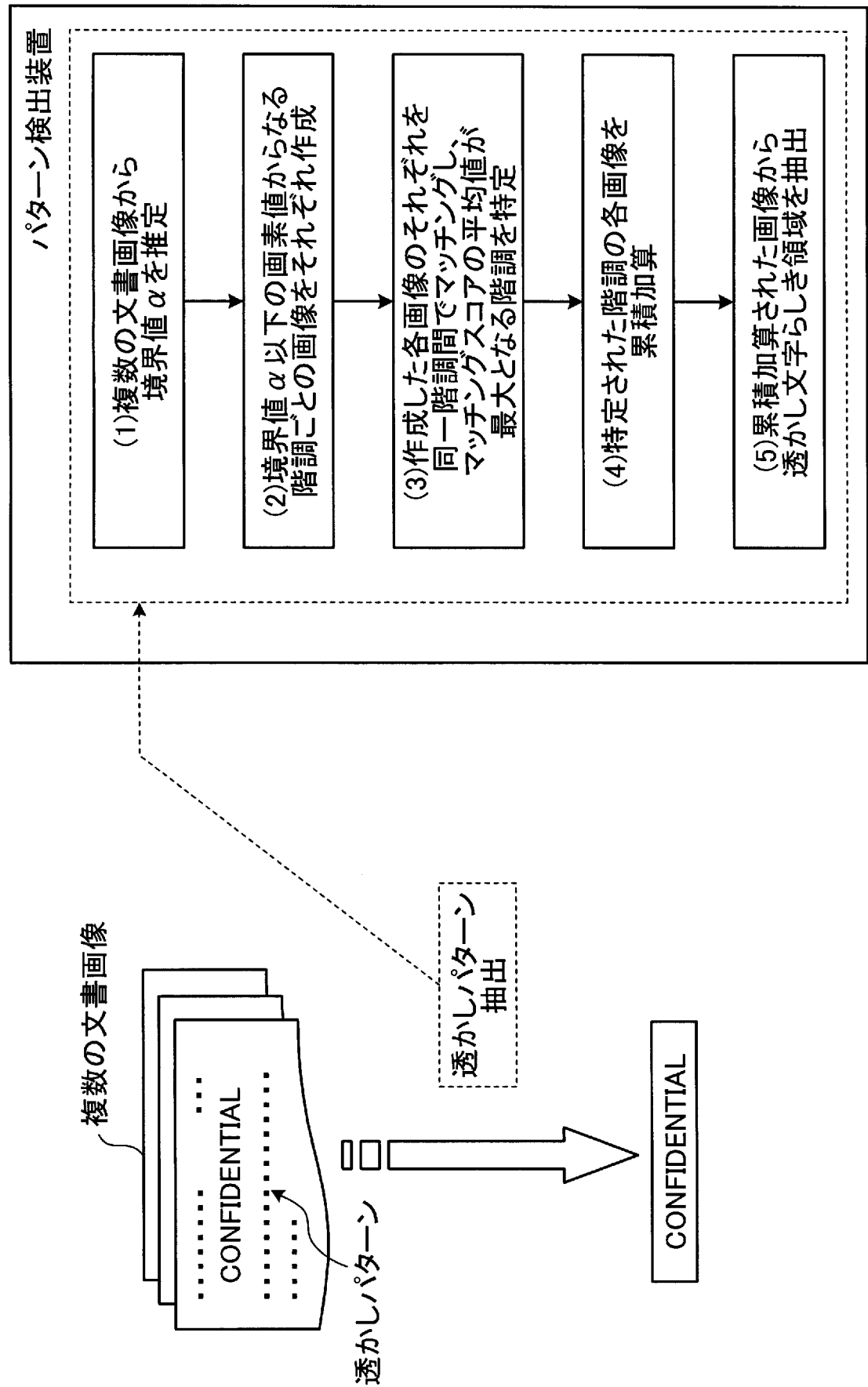
するパターン認識手段と、

をさらに備えたことを特徴とする請求項9または10に記載のパターン検出装置。

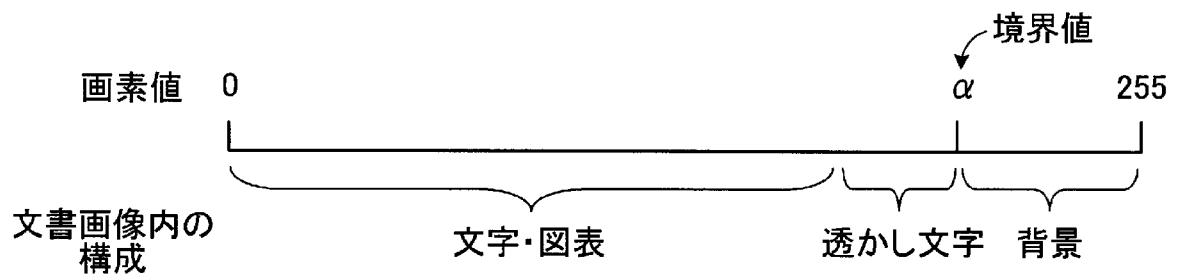
- [12] 前記複数の文書画像間で同一座標に位置する各画素の画素値について、前記境界値推定手段により推定された前記境界値以上となる文書画像数に対する割合を算出し、当該境界値以上となる割合が所定の範囲内となる場合には、当該複数の文書画像間の同一座標における画素を文字、図表および前記共通パターンを構成する画素以外の背景画素として推定する背景画素推定手段をさらに備え、

前記画像作成手段は、前記背景画素推定手段により前記背景画素と推定された画素を背景とみなして、前記境界値推定手段により推定された前記境界値以下の画素値ごとの画像をそれぞれ作成することを特徴とする請求項9に記載のパターン検出装置。

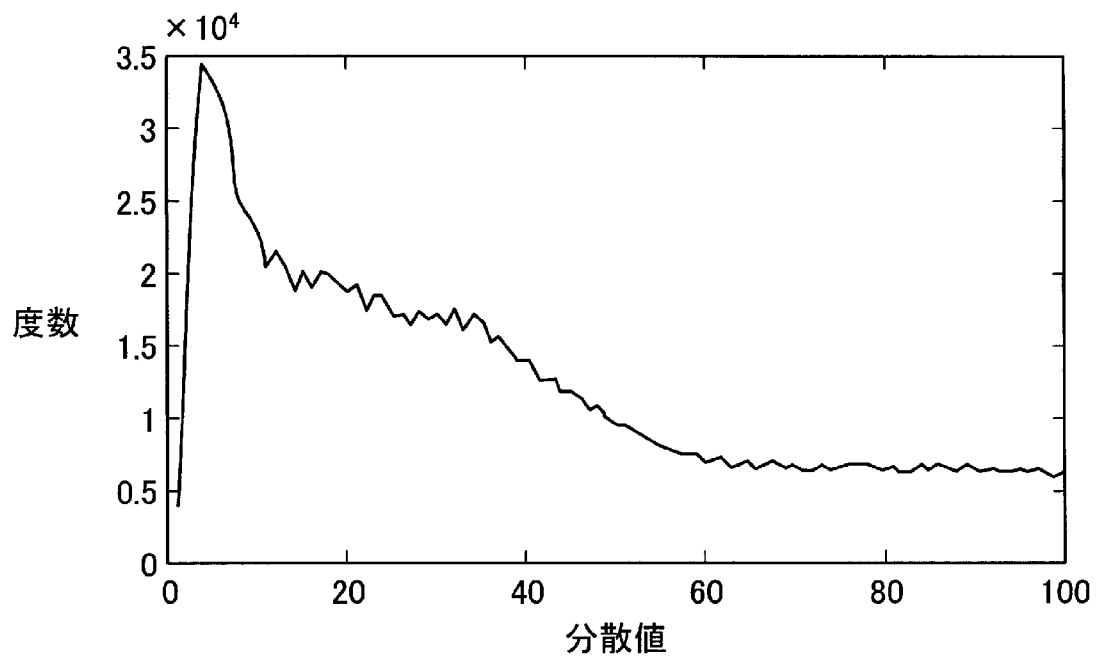
[図1]



[図2]

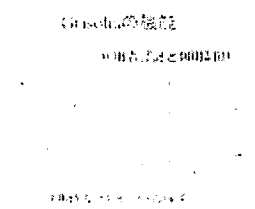
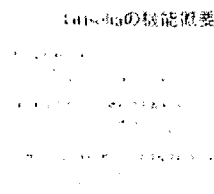
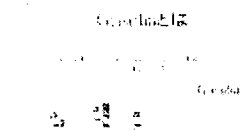


[図3]

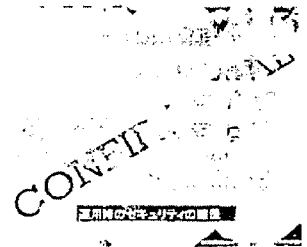
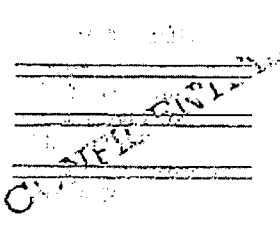
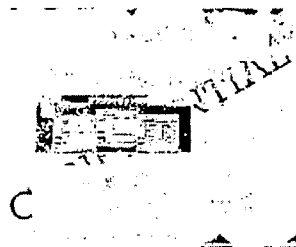


[図4]

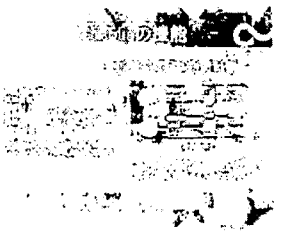
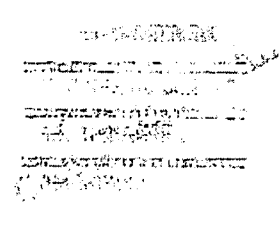
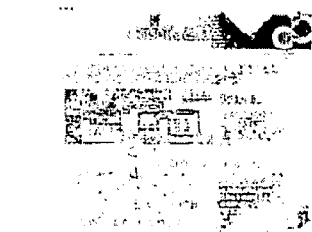
3
階調



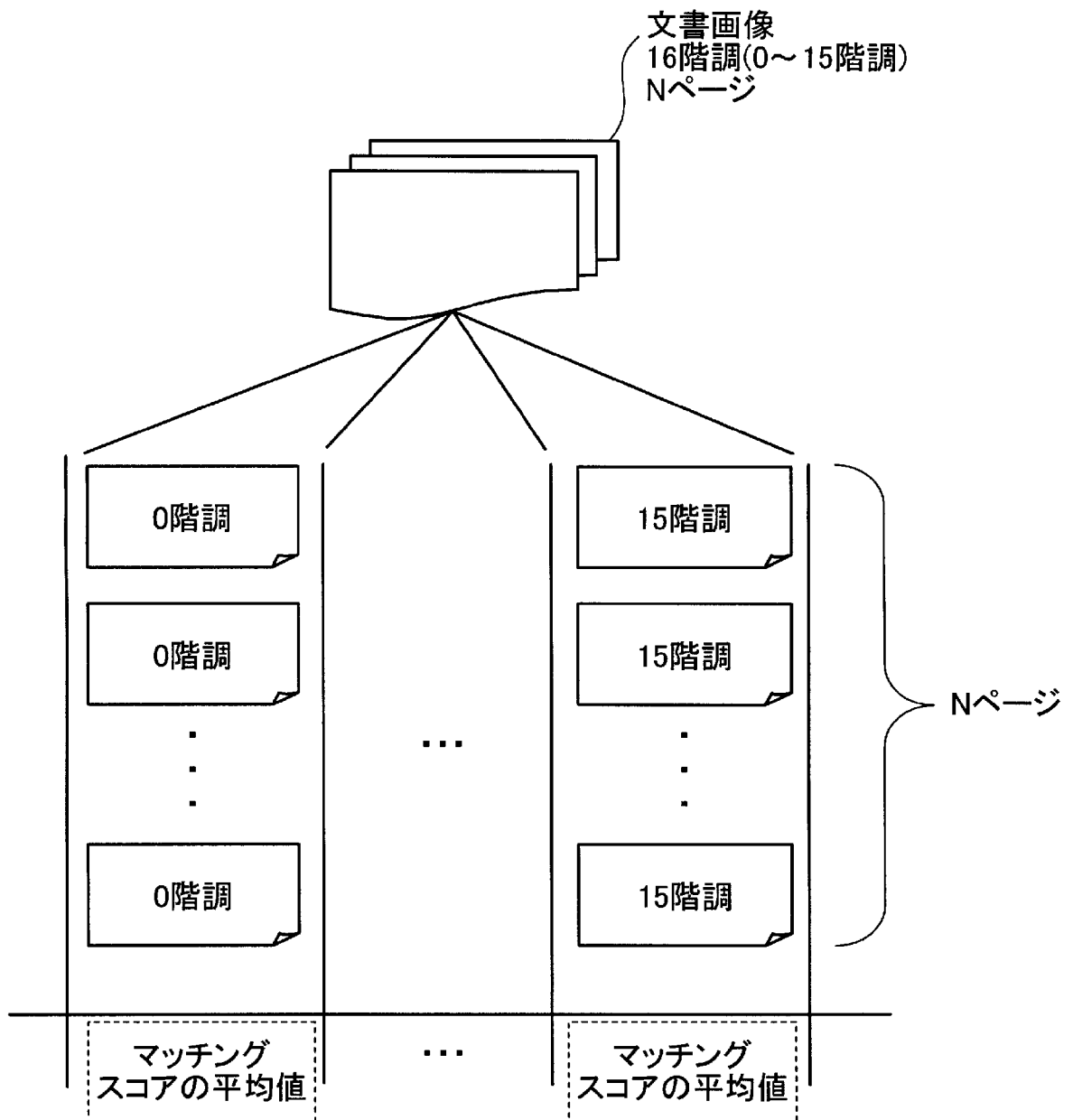
1
4
階調



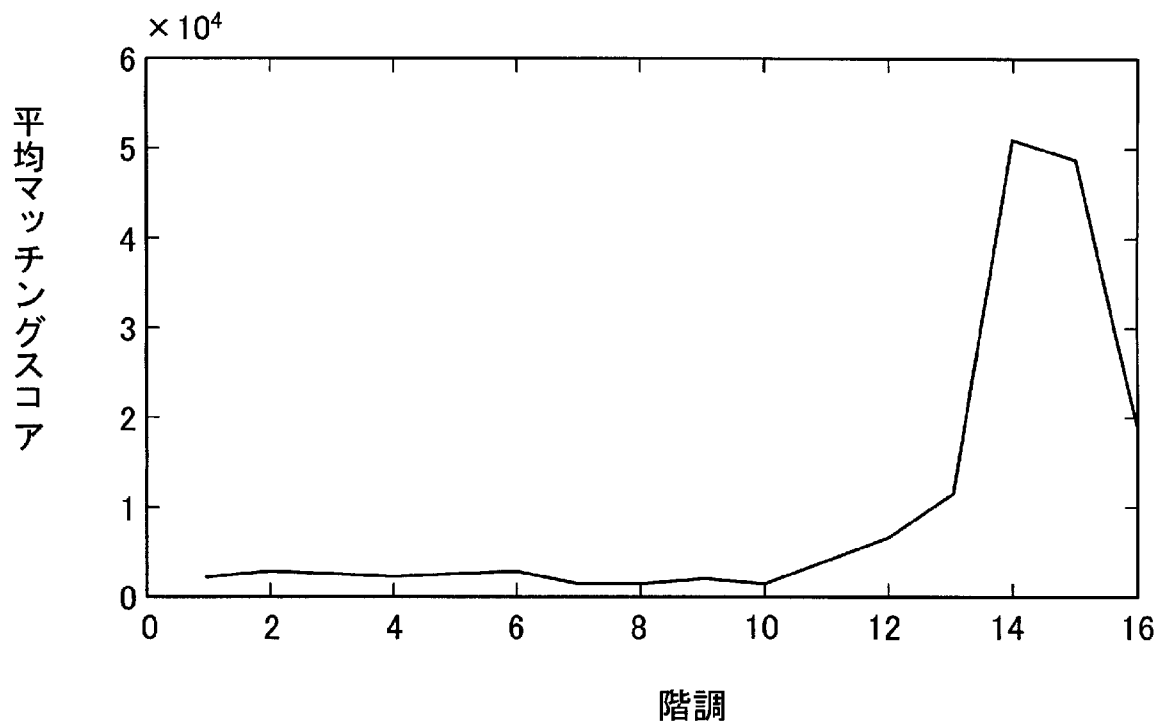
1
5
階調



[図5]



[図6]



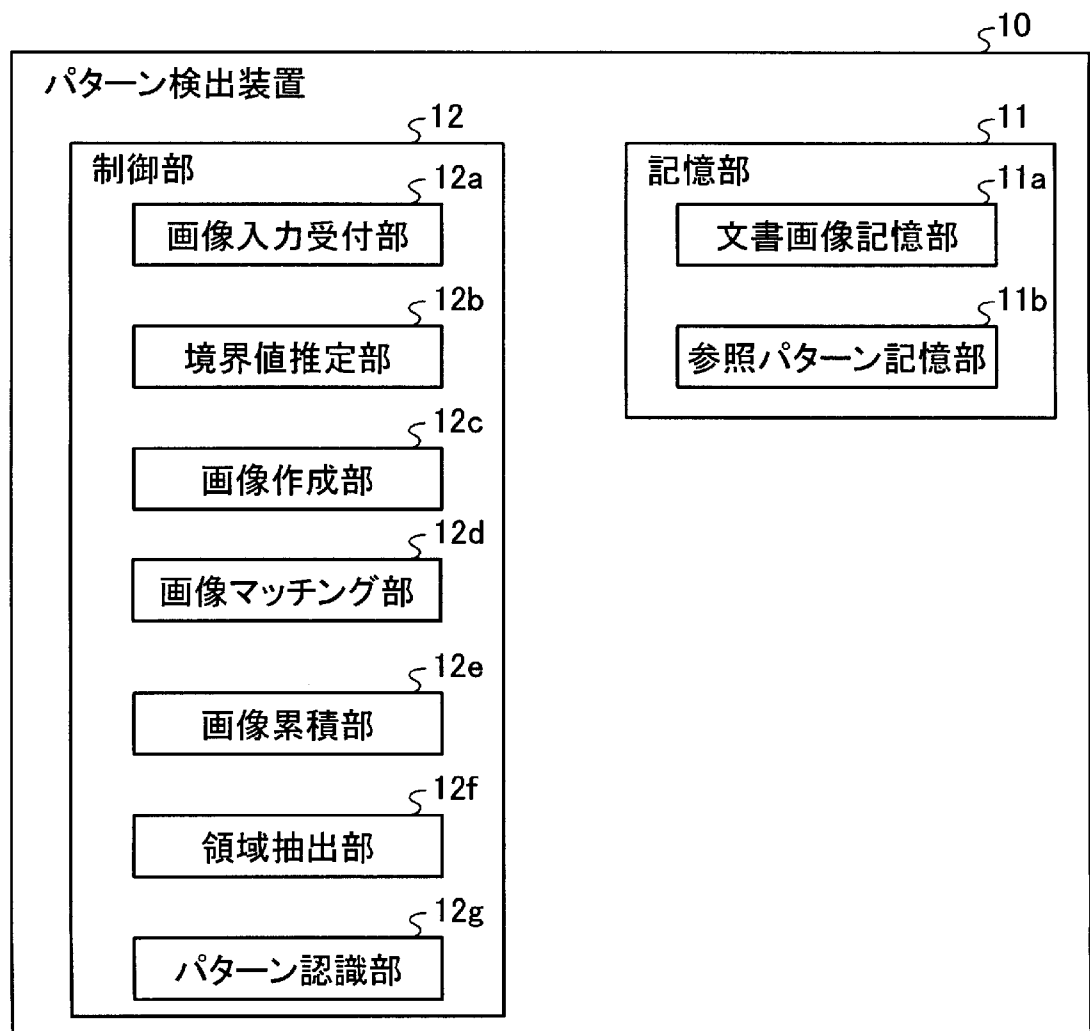
[図7]

CONFIDENTIAL

[図8]

CONFIDENTIAL

[図9]



[図10]

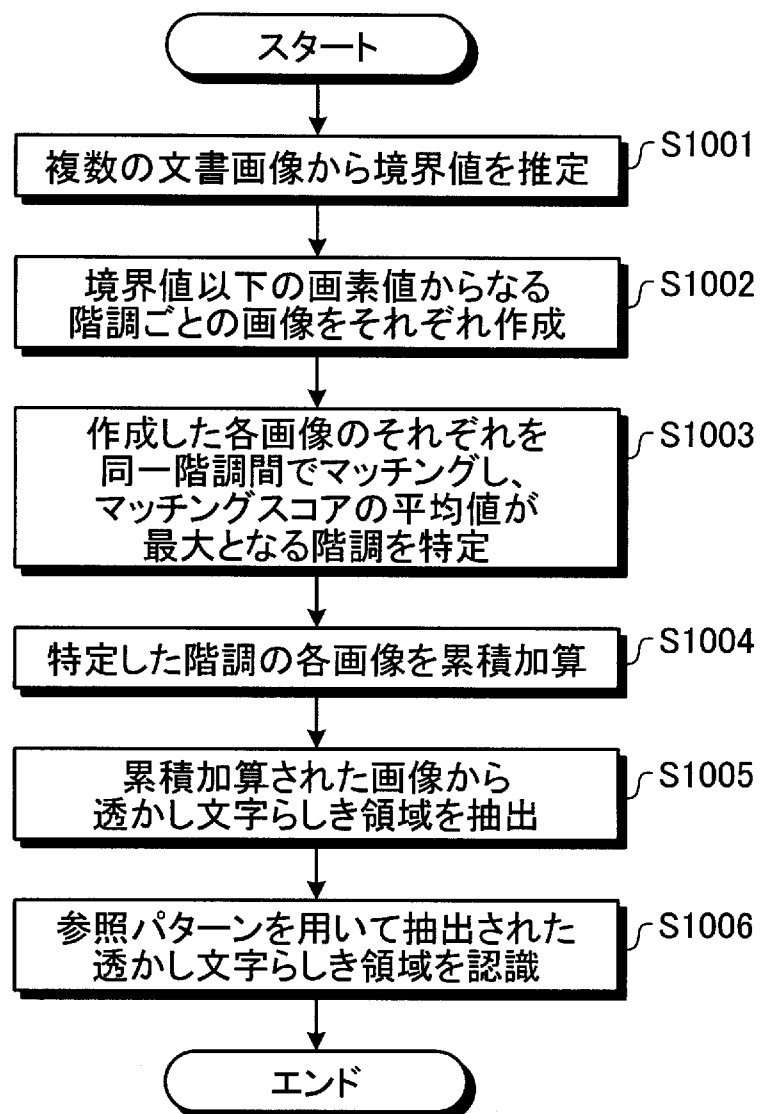
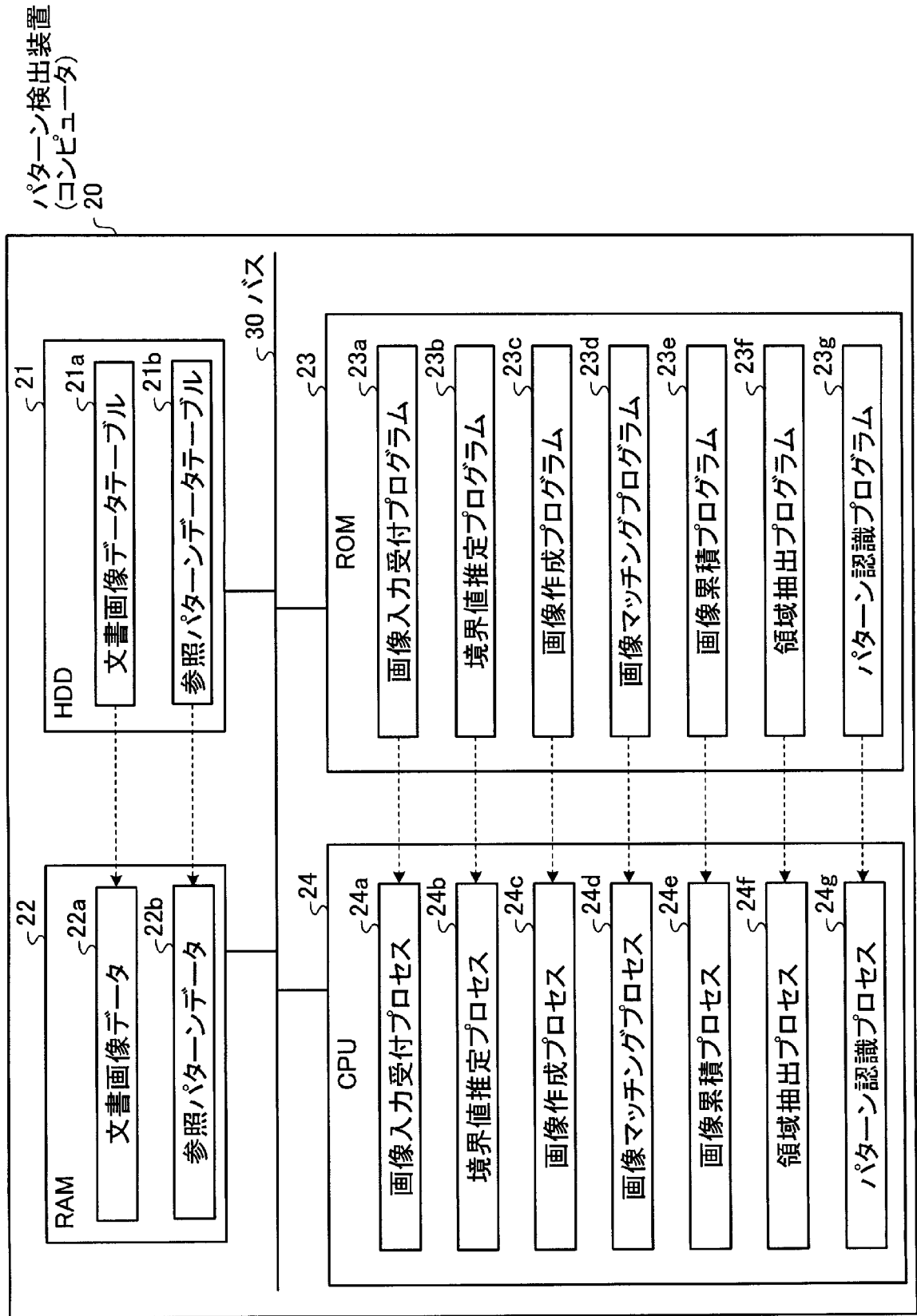


図11



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2007/054465

A. CLASSIFICATION OF SUBJECT MATTER

H04N1/387(2006.01) i, G06T1/00(2006.01) i, H04N1/40(2006.01) i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N1/387, G06T1/00, H04N1/40

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

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-96375 A (Fujitsu Ltd.), 25 March, 2004 (25.03.04), Full text; Figs. 1 to 22 (Family: none)	1-12
A	JP 2000-49985 A (Canon Inc.), 18 February, 2000 (18.02.00), Full text; Figs. 1 to 14 (Family: none)	1-12

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

☐ See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search
25 May, 2007 (25.05.07)

Date of mailing of the international search report
05 June, 2007 (05.06.07)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. H04N1/387(2006.01)i, G06T1/00(2006.01)i, H04N1/40(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. H04N1/387, G06T1/00, H04N1/40

最小限資料以外の資料で調査を行った分野に含まれるもの

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

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

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	J P 2 0 0 4－9 6 3 7 5 A（富士通株式会社） 2 0 0 4. 0 3. 2 5, 全文, 第1－2 2 図 (ファミリーなし)	1－1 2
A	J P 2 0 0 0－4 9 9 8 5 A（キヤノン株式会社） 2 0 0 0. 0 2. 1 8, 全文, 第1－1 4 図 (ファミリーなし)	1－1 2

☐ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

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

の日の後に公表された文献

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

国際調査を完了した日

2 5. 0 5. 2 0 0 7

国際調査報告の発送日

0 5. 0 6. 2 0 0 7

国際調査機関の名称及びあて先

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

特許庁審査官（権限のある職員）

白石 圭吾

5 V

9 8 5 6

電話番号 0 3－3 5 8 1－1 1 0 1 内線 3 5 7 1