

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2017 年 3 月 9 日(09.03.2017)



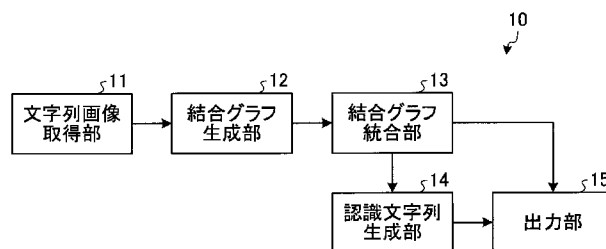
(10) 国際公開番号
WO 2017/038952 A1

- (51) 国際特許分類:
G06K 9/62 (2006.01) *G06K 9/03* (2006.01)
- (21) 国際出願番号: PCT/JP2016/075721
- (22) 国際出願日: 2016 年 9 月 1 日(01.09.2016)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2015-174414 2015 年 9 月 4 日(04.09.2015) JP
- (71) 出願人: 株式会社東芝 (KABUSHIKI KAISHA TOSHIBA) [JP/JP]; 〒1058001 東京都港区芝浦一丁目 1 番 1 号 Tokyo (JP). 東芝ソリューション株式会社 (TOSHIBA SOLUTIONS CORPORATION) [JP/JP]; 〒2128585 神奈川県川崎市幸区堀川町 7 2 番地 3 4 Kanagawa (JP).
- (72) 発明者: 吉田 篤弘 (YOSHIDA, Atsuhiko); 〒2128585 神奈川県川崎市幸区堀川町 7 2 番地 3 4 東芝ソリューション株式会社内 Kanagawa (JP). 黒沢 由明 (KUROSAWA, Yoshiaki); 〒2128585 神奈川県川崎市幸区堀川町 7 2 番地 3 4 東芝ソリューション株式会社内 Kanagawa (JP).
- (74) 代理人: 特許業務法人酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関 3 丁目 8 番 1 号 虎の門三井ビルディング 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, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (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, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ユーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

[続葉有]

(54) Title: CHARACTER RECOGNITION DEVICE, CHARACTER RECOGNITION METHOD, AND PROGRAM

(54) 発明の名称: 文字認識装置、文字認識方法およびプログラム



- 11 Character string image acquisition unit
12 Association graph generation unit
13 Association graph integration unit
14 Recognized character string generation unit
15 Output unit

(57) Abstract: A character recognition device (10) of an embodiment of the present invention is provided with: a character string image acquisition unit (11) that acquires a character string image; an association graph generation unit (12) that performs a character recognition process on the character string image and generates an association graph obtained by connecting a plurality of sets of character candidate information, which each represent recognition results for each character region regarded as one character and each include at least one character candidate, according to the arrangement order of the character regions in the character string image; an association graph integration unit (13) that integrates a plurality of association graphs generated from a plurality of character string images including an identical character string, or a plurality of association graphs generated by performing a plurality of different character recognition processes on one character string image; and an output unit (15) for outputting the integrated association graph or a recognized character string obtained on the basis of the integrated association graph.

(57) 要約:

[続葉有]



WO 2017/038952 A1



添付公開書類:

— 国際調査報告（条約第 21 条(3)）

実施形態の文字認識装置（10）は、文字列画像を取得する文字列画像取得部（11）と、文字列画像に対して文字認識処理を行い、1つの文字とみなされる文字領域ごとの認識結果を表す文字候補情報であって各々が1以上の候補文字を含む複数の文字候補情報を、文字列画像における各文字領域の並び順に従って接続した結合グラフを生成する結合グラフ生成部（12）と、同一文字列を含む複数の文字列画像から生成された複数の結合グラフ、または、1つの文字列画像に対して複数の異なる文字認識処理を行うことで生成された複数の結合グラフを統合する結合グラフ統合部（13）と、統合された結合グラフ、または、統合された結合グラフに基づいて得られる認識文字列を出力する出力部（15）と、を備える。

明 細 書

発明の名称：文字認識装置、文字認識方法およびプログラム

技術分野

[0001] 本発明の実施形態は、文字認識装置、文字認識方法およびプログラムに関する。

背景技術

[0002] OCR (Optical Character Recognition/Reader) に代表される文字認識の分野では、認識精度の向上を図るための様々な取り組みがなされている。例えば、同一の文字列を含む複数の文字列画像に対してそれぞれ文字認識処理を行い、対応する文字について信頼度の高い認識結果を選択して最終的な認識文字列を得るといった技術が知られている。

[0003] しかし、信頼度の高い認識結果を選択する従来の方法では、例えば、信頼度の高い認識結果が必ずしも正解ではないこと、また、文字列画像における文字の区切りが正しくない場合もあることから、正しい認識文字列が得られない場合も多く、さらなる改善が求められている。

先行技術文献

特許文献

[0004] 特許文献1：特開2003-331217号公報

発明の概要

発明が解決しようとする課題

[0005] 本発明が解決しようとする課題は、高精度な文字認識を行うことが可能な文字認識装置、文字認識方法およびプログラムを提供することである。

課題を解決するための手段

[0006] 実施形態の文字認識装置は、文字列画像取得部と、結合グラフ生成部と、結合グラフ統合部と、出力部と、を備える。文字列画像取得部は、文字列画像を取得する。結合グラフ生成部は、前記文字列画像に対して文字認識処理を行い、1つの文字とみなされる文字領域ごとの認識結果を表す文字候補情

報であって各々が1以上の候補文字を含む複数の前記文字候補情報を、前記文字列画像における各文字領域の並び順に従って接続した結合グラフを生成する。結合グラフ統合部は、同一文字列を含む複数の前記文字列画像から生成された複数の前記結合グラフ、または、1つの前記文字列画像に対して複数の異なる文字認識処理を行うことで生成された複数の前記結合グラフを統合する。出力部は、統合された前記結合グラフ、または、統合された前記結合グラフに基づいて得られる認識文字列を出力する。

図面の簡単な説明

- [0007] [図1]図1は、文字認識装置のハードウェア構成例を示すブロック図である。
- [図2]図2は、文字認識装置の機能的な構成例を示すブロック図である。
- [図3]図3は、結合グラフの一例を示す図である。
- [図4]図4は、結合グラフのデータ構造の一例を説明する図である。
- [図5]図5は、累積結合グラフおよび新規結合グラフの一例を示す図である。
- [図6]図6は、図5に示す新規結合グラフを累積結合グラフに統合することによって得られる新たな累積結合グラフを示す図である。
- [図7]図7は、文字認識装置による処理手順の一例を示すフローチャートである。
- [図8]図8は、図7のステップS105の統合処理の概要を説明するフローチャートである。
- [図9]図9は、図8のステップS205の処理手順を示すフローチャートである。
- [図10]図10は、図5に例示した累積結合グラフおよび新規結合グラフの一部の文字候補情報を抜き出して示す図である。
- [図11]図11は、結合グラフを単一の接続パスに分離する様子を示す図である。

発明を実施するための形態

- [0008] 以下、実施形態の文字認識装置、文字認識方法およびプログラムを、図面を参照して詳細に説明する。

[0009] 図1は、実施形態の文字認識装置10のハードウェア構成例を示すブロック図である。文字認識装置10は、例えば一般的なコンピュータとしてのハードウェア構成を採用することができる。この場合、文字認識装置10は、図1に示すように、CPU (Central Processing Unit) 101、ROM (Read Only Memory) 102、RAM (Random Access Memory) 103、HDD (Hard Disk Drive) 104、デバイスI/F 105、ネットワークI/F 106、これらを接続するバス107などを備える。そして、文字認識装置10は、例えば、CPU 101がRAM 103をワークエリアとして利用し、ROM 102やHDD 104などに格納されたプログラムを実行することによって、文字認識に関わる様々な機能を実現することができる。

[0010] デバイスI/F 105は、液晶ディスプレイなどの表示装置108、キーボードやマウスなどの操作入力装置109、カメラやスキャナなどの画像入力装置110といった周辺機器を文字認識装置10に接続するためのインタフェースである。ネットワークI/F 106は、文字認識装置10をインターネットやLAN (Local Area Network) などのネットワークに接続するための通信インタフェースである。

[0011] 図2は、実施形態の文字認識装置10の機能的な構成例を示すブロック図である。文字認識装置10は、例えば、上記のハードウェアとソフトウェア（プログラム）との協働により実現される機能的な構成要素として、図2に示すように、文字列画像取得部11と、結合グラフ生成部12と、結合グラフ統合部13と、認識文字列生成部14と、出力部15とを備える。

[0012] 文字列画像取得部11は、文字認識処理の対象となる文字列画像を取得する。文字列画像取得部11は、例えば、カメラやスキャナなどの画像入力装置110から入力される文字列画像をデバイスI/F 105を介して取得する構成であってもよいし、ネットワークに接続された外部装置から送信される文字列画像をネットワークI/F 106を介して取得する構成であってもよい。また、文字列画像取得部11は、予め取得した文字列画像をHDD 104などに格納しておき、文字認識処理を実行する際にHDD 104などが

ら文字列画像を読み出す構成であってもよい。

[0013] 文字列画像取得部 11 は、取得した文字列画像に対して例えば二値化処理などの文字認識処理を行う上で必要な前処理を施し、前処理を施した文字列画像を結合グラフ生成部 12 に渡す。なお、文字認識処理を行う上で必要な前処理は既存技術をそのまま利用できるため、詳細な説明は省略する。

[0014] 結合グラフ生成部 12 は、文字列画像取得部 11 から受け取った文字列画像に対して文字認識処理を行い、この文字列画像に対する文字認識処理の結果を纏めたグラフである結合グラフを生成する。文字認識処理は、例えば、文字列画像から 1 つの文字とみなされる文字領域をすべて抽出し、それぞれの文字領域から特徴量を求めて、その特徴量に基づいて文字領域ごとに 1 以上の候補文字とその確からしさを表す認識スコアを取得する処理である。また、文字認識処理は、文字列画像に対する文字領域の区切りと文字領域に対する文字認識とを同時に行うものであってもよい。結合グラフ生成部 12 は、文字列画像取得部 11 から受け取った文字列画像に対し、以上のような文字認識処理を行って、個々の文字領域の文字列画像 I S における位置やサイズ、個々の文字領域から各々取得された候補文字や認識スコアなどを纏めることにより、結合グラフを生成する。なお、文字列画像 I S に対する具体的な文字認識処理の手法、例えば文字領域の抽出方法や文字認識に用いる特徴量などは既存技術をそのまま利用できるため、詳細な説明は省略する。

[0015] 図 3 は、結合グラフ生成部 12 により生成される結合グラフ G の一例を示す図である。結合グラフ G は、図 3 に示すように、文字列画像 I S において 1 つの文字とみなされる文字領域ごとの認識結果を表す文字候補情報 210 を、文字列画像 I S における各文字領域の並び順に従って接続したグラフである。結合グラフ G は、文字列画像 I S における文字領域の区切りが異なる複数のパターンに対応して、複数の接続パスを含みうる。接続パスは、文字列画像 I S における文字候補情報 210 の繋がりを表す。図 3 の例では、「糸」と「色」が 2 つの文字とみなされる場合と、1 つの「絶」という文字とみなされる場合とで、接続パスが異なる。また、「女」と「子」が 2 つの文

字とみなされる場合と、1つの「好」という文字とみなされる場合とで、接続パスが異なる。このため、図3に示す結合グラフGには、「糸」→「色」→「女」→「子」と繋がる接続パスと、「糸」→「色」→「好」と繋がる接続パスと、「絶」→「女」→「子」と繋がる接続パスと、「絶」→「好」と繋がる接続パスとの4種類の接続パスが含まれる。なお、文字列画像I Sにおける文字領域の区切りが一意に特定される場合は、結合グラフGに含まれる接続パスは1つである。

[0016] 結合グラフGでは、隣り合う文字候補情報210同士の接続関係が接続情報220によって表される。ここでの接続とは、2つの文字候補情報210に各々対応する2つの文字が隣り合っていることを意味する。結合グラフGを図3のようにグラフィカルに表すと、接続情報220は、隣り合う2つの文字候補情報210の間に配置される。なお、特殊な接続情報220として、文字列の先頭には開始位置221が配置され、文字列の末尾には終了位置222が配置される。

[0017] 図3は、横方向に文字が並ぶ横文字列を含む文字列画像I Sを文字認識処理の対象とした場合に生成される結合グラフGをグラフィカルに表した例であり、横方向に並ぶ文字候補情報210の各々が、文字列画像I Sにおいて1つの文字とみなされる文字領域ごとの認識結果を表している。なお、図3に示す各文字候補情報210の文字は、対応する文字領域に対する文字認識により取得される候補文字のうち、認識スコアが最も高い候補文字を示している。以下では、このような横文字列を含む文字列画像I Sを文字認識処理の対象とする場合について説明するが、縦方向に文字が並ぶ縦文字列を含む文字列画像I Sを文字認識処理の対象とする場合も、文字候補情報210の並びが横方向から縦方向に変わるだけで、結合グラフGの基本的な構成は同様である。

[0018] ここで、結合グラフGのデータ構造の具体例について説明する。図4は、結合グラフGのデータ構造の一例を説明する図である。図4では、結合グラフGの中から1つの接続情報220および当該接続情報220に関連する複

数の文字候補情報 210 を部分的に抜き出したものを模式的に示している。

[0019] 文字候補情報 210 は、上述したように、1つの文字とみなされる文字領域に対する文字認識により得られる情報であり、例えば、フラッグ、候補数、文字コード、スコア、サイズ、位置、右ポインタ、左ポインタなどを含む。フラッグは、当該文字候補情報 210 の属性などを表す。候補数は、当該文字候補情報 210 に含まれる文字候補の数を表す。文字コードは、当該文字認識情報 210 に含まれる 1 以上の候補文字それぞれの文字コードである。スコアは、それぞれの候補文字に対応する認識スコアである。サイズは、当該文字候補情報 210 に対応する文字領域（文字の外接矩形）のサイズである。位置は、当該文字候補情報 210 に対応する文字領域の文字列画像 I S における位置（本実施形態では文字領域の左端位置や右端位置）を表す位置情報である。右ポインタは、当該文字候補情報 210 の右端位置に対応する接続情報 220 を指し示すポインタである。左ポインタは、当該文字候補領域 210 の左端位置に対応する接続情報 220 を指し示すポインタである。なお、ポインタは対象となる情報が格納されたメモリ上の領域を特定できればよく、例えばメモリ上のアドレスやインデックスなどを用いることができる。

[0020] 接続情報 220 は、隣り合う文字候補情報 210 を接続するための情報であり、フラッグ、複数の左ポインタ、複数の左接続位置、複数の右ポインタ、複数の右接続位置を含む。フラッグは、当該接続情報 220 の属性などを表す。左ポインタは、当該接続情報 220 を介して隣り合う文字候補情報 210 のうちの左側の文字候補情報 210 を指し示すポインタである。左接続位置は、左ポインタが指し示す文字候補情報 210 の位置を知るための情報であり、例えば当該文字候補情報 210 の位置情報である右端位置が登録される。右ポインタは、当該接続情報 220 を介して隣り合う文字候補情報 210 のうちの右側の文字候補情報 210 を指し示すポインタである。右接続位置は、右ポインタが指し示す文字候補情報 210 の位置を知るための情報であり、例えば当該文字候補情報 210 の位置情報である左端位置が登録さ

れる。

[0021] 上述したように結合グラフ G は複数の接続パスを含む場合があるので、文字候補情報 210 同士の接続関係は複数通り存在する。このため、接続情報 220 には、複数の左ポイントおよび左接続位置と、複数の右ポイントおよび右接続位置とが設けられている。各ポイントは有効／無効を切り替えられるようになっており、それぞれのポイントが有効であるか無効であるかは、例えばフラッグに記載される。

[0022] なお、図 3 に示す例のように、隣り合う文字候補情報 210 同士の接続関係を 2 つの接続情報 220 で表すことも可能である。この場合、2 つの接続情報 220 のうちの左側の接続情報 220 は、右ポイントの 1 つで右側の接続情報 220 を指し示し、その右ポイントに対応する右接続位置に、右側の接続情報 220 の右接続位置と同じ位置が登録される。また、2 つの接続情報 220 のうちの右側の接続情報 220 は、左ポイントの 1 つで左側の接続情報 220 を指し示し、その左ポイントに対応する左接続位置に、左側の接続情報 220 の左接続位置と同じ位置が登録される。

[0023] 図 3 に示す開始位置 221 は、右ポイントおよび右接続位置のみ登録された特殊な接続情報 220 であり、図 3 に示す終了位置 222 は、左ポイントおよび左接続位置のみ登録された特殊な接続情報 220 である。このような接続情報 220 の属性は、上述したフラッグに記載されている。なお、1 つの結合グラフ G には開始位置 221 と終了位置 222 が通常 1 つずつ設けられるが、結合グラフ G 内に開始位置 221 や終了位置 222 が複数存在していてもよい。

[0024] 本実施形態では、隣り合う文字候補情報 210 同士の接続関係を接続情報 220 によって表す構成の結合グラフ G を例示するが、これに限らない。例えば文字候補情報 210 が隣り合う他の文字候補情報 210 を直接指し示すように設定し、接続情報 220 を含まない構成の結合グラフ G としてもよい。この場合、文字候補情報 210 には 1 つの接続情報 220 を指し示す左ポイントや右ポイントの代わりに、隣り合う他の文字候補情報 210 を指し示

す複数の左ポインタや複数の右ポインタを設定すればよい。

[0025] 結合グラフ生成部 12 は、文字列画像取得部 11 から文字列画像 I S を受け取るたびに、以上のような結合グラフ G を生成して結合グラフ統合部 13 に渡す。特に本実施形態では、結合グラフ生成部 12 が 1 つの文字列に対して複数の結合グラフ G を生成して結合グラフ統合部 13 に渡す。例えば、結合グラフ生成部 12 は、同一文字列を含む複数の文字列画像 I S に対してそれぞれ文字認識処理を行うことで複数の結合グラフ G を生成し、これら複数の結合グラフ G を結合グラフ統合部 13 に渡す。また、結合グラフ生成部 12 は、1 つの文字列画像 I S に対して複数の異なる文字認識処理を行うことで複数の結合グラフ G を生成し、これら複数の結合グラフ G を結合グラフ統合部 13 に渡すようにしてもよい。なお、同一文字列を含む複数の文字列画像 I S は、例えば画像ファイルのファイル名などで識別可能な構成とすることができる。

[0026] 結合グラフ統合部 13 は、1 つの文字列に対して結合グラフ生成部 12 により生成された複数の結合グラフ G、すなわち、同一文字列を含む複数の文字列画像 I S から生成された複数の結合グラフ G、または、1 つの文字列画像 I S に対して複数の異なる文字認識処理を行うことで生成された複数の結合グラフ G を統合する。本実施形態では、結合グラフ G を 1 つずつ順次統合していく方法を採用する。以下では、それまでに統合された結合グラフ G を累積結合グラフ G_{acc}（第 1 の結合グラフ）と呼び、新たに統合する結合グラフ G を新規結合グラフ G_{new}（第 2 の結合グラフ）と呼ぶ。

[0027] 結合グラフ統合部 13 は、1 つの文字列に対して結合グラフ生成部 12 により生成された複数の結合グラフ G のうちの最初の結合グラフ G を受け取ると、これを初期の累積結合グラフ G_{acc} として保存する。そして、結合グラフ統合部 13 は、2 番目の結合グラフ G を受け取るとこれを新規結合グラフ G_{new} とし、この新規結合グラフ G_{new} を累積結合グラフ G_{acc} に統合して、統合された結合グラフ G を新たな累積結合グラフ G_{acc} として保存する。結合グラフ統合部 13 は、3 番目以降の結合グラフ G

に対しても同様の処理を繰り返し、1つの文字列に対して結合グラフ生成部12により生成されたすべての結合グラフGの統合が終了すると、最終的に得られた累積結合グラフG_{acc}を認識文字列生成部14または出力部15に渡す。

[0028] 累積結合グラフG_{acc}に対する新規結合グラフG_{new}の統合は、以下のように行われる。すなわち、結合グラフ統合部13は、累積結合グラフG_{acc}に含まれる各文字候補情報210と新規結合グラフG_{new}に含まれる各文字候補情報210との対応関係を特定し、対応する文字候補情報210同士をマージ（1つに併合）するとともに、累積結合グラフG_{acc}側の文字候補情報210のいずれにも対応しない新規結合グラフG_{new}側の文字候補情報210を累積結合グラフG_{acc}に追加することで、新規結合グラフG_{new}を累積結合グラフG_{acc}に統合する。

[0029] 以下では、このような統合処理の具体例について、図5および図6を参照しながら説明する。図5（a）は累積結合グラフG_{acc}の一例、図5（b）は新規結合グラフG_{new}の一例をそれぞれ示し、図6は、図5（b）の新規結合グラフG_{new}を図5（a）の累積結合グラフG_{acc}に統合することによって得られる新たな累積結合グラフG_{acc}を示している。図5では、累積結合グラフG_{acc}や新規結合グラフG_{new}に含まれる各文字候補情報210を区別するために、累積結合グラフG_{acc}側の文字候補情報210にA1，A2，A3，A4，A5，A6の符号を付し、新規結合グラフG_{new}側の文字候補情報210にB1，B2，B3，B4，B5の符号を付している。

[0030] 本実施形態では、文字候補情報210に含まれる位置情報（文字列画像ISにおける文字領域の左端位置や右端位置）を手掛かりとして、累積結合グラフG_{acc}に含まれる各文字候補情報210と新規結合グラフG_{new}に含まれる各文字候補情報210との対応関係を特定する。

[0031] 結合グラフ統合部13は、新規結合グラフG_{new}に含まれる文字候補情報210のそれぞれについて、位置情報として登録されている文字領域の

左端位置にほぼ一致する右接続位置を有する接続情報 220 と、位置情報として登録されている文字領域の右端位置にほぼ一致する左接続位置を有する接続情報 220 との組を、累積結合グラフ G_acc の中から検索する。ほぼ一致するとは、双方の位置の違いが予め定めた誤差範囲内にあることをいう。これにより、新規結合グラフ G_new 側の文字候補情報 210 の左右の接続情報 220 に対応する累積結合グラフ G_acc 側の 2 つの接続情報 220 が特定される。

[0032] 次に、結合グラフ統合部 13 は、特定した累積結合グラフ G_acc 側の 2 つの接続情報 220 に挟まれる 1 つの文字候補情報 210 が累積結合グラフ G_acc 内に存在するか否かを判断し、そのような文字候補情報 210 が累積結合グラフ G_acc 内にあれば、その文字候補情報 210 が新規結合グラフ G_new 側の文字候補情報 210 に対応するものと判断する。このとき、結合グラフ統合部 13 は、累積結合グラフ G_acc 側の文字候補情報 210 と新規結合グラフ G_new 側の文字候補情報 210 とが対応しているか否かを、さらに、双方の文字候補情報 210 に含まれる文字候補の一致度などを考慮して判断することが望ましい。例えば、双方の文字候補情報 210 に同じ文字候補が所定数以上含まれている場合に、双方の文字候補情報 210 が対応していると判断する。

[0033] 結合グラフ統合部 13 は、新規結合グラフ G_new に含まれる文字候補情報 210 のうち、累積結合グラフ G_acc 内で対応する文字候補情報 210 が見つかった文字候補情報 210 については、その新規結合グラフ G_new 側の文字候補情報 210 を、対応する累積結合グラフ G_acc 側の文字候補情報 210 にマージ（1 つに併合）する。具体的には、文字認識により得られた候補文字の文字コードと認識スコアをマージする。文字候補情報 210 をマージする際、候補文字の文字コードを認識スコア順にソートするが、同じ文字コードで認識スコアが異なる場合には、認識スコアが高い方を採用する。また、マージによって候補文字の数が予め定めた上限値を超える場合は、認識スコアが低い文字コードは登録しないようにする。

[0034] 図5に示す例では、新規結合グラフG__new側のB1, B2, B3, B4が、累積結合グラフG__acc側のA1, A2, A3, A4に各々対応するので、B1がA1、B2がA2、B3がA3、B4がA4にそれぞれマージされる。

[0035] また、結合グラフ統合部13は、新規結合グラフG__newに含まれる文字候補情報210のうち、累積結合グラフG__acc内で対応する文字候補情報210が見つからない文字候補情報210については、その新規結合グラフG__new側の文字候補情報210を、累積結合グラフG__accに新たな文字候補情報210として追加する。具体的には、結合グラフ統合部13は、追加すべき文字候補情報210の右ポイントが、その文字候補情報210の右側の接続情報220に対応する累積結合グラフG__acc側の接続情報220を指し示し、追加すべき文字候補情報210の左ポイントが、その文字候補情報210の左側の接続情報220に対応する累積結合グラフG__acc側の接続情報220を指し示すように、これらのポイントを変更する。また、結合グラフ統合部13は、追加すべき文字候補情報210の右側の接続情報220に対応する累積結合グラフG__acc側の接続情報220に、その文字候補情報210を指し示す左ポイントおよび左接続位置を追加登録するとともに、追加すべき文字候補情報210の左側の接続情報220に対応する累積結合グラフG__acc側の接続情報220に、その文字候補情報210を指し示す右ポイントおよび右接続位置を追加登録する。これにより、累積結合グラフG__acc側の文字候補情報210のいずれにも対応しない新規結合グラフG__new側の文字候補情報210が、累積結合グラフG__accに追加される。

[0036] 図5に示す例では、新規結合グラフG__new側のB5の累積結合グラフG__acc側の接続位置の間にA2およびA3の2つの文字候補情報210があり、新規結合グラフG__new側のB5に対応する累積結合グラフG__acc側の1つの文字候補情報210が見つからないため、新規結合グラフG__new側のB5が、累積結合グラフG__acc側のA1とA4の間に、

新たな文字候補情報 210 として追加される。

[0037] 結合グラフ統合部 13 は、新規結合グラフ G_{new} 内のすべての文字候補情報 210 について、左からの接続順に以上のような統合処理を順次行っていく。また、新規結合グラフ G_{new} 側の文字候補情報 210 の左右に対応する累積結合グラフ G_{acc} 側の接続情報 220 の組が複数見つかる場合があるが、その場合には、それぞれに対して上述した文字候補情報 210 のマージあるいは追加を行う。この統合によって、図 5 に示す累積結合グラフ G_{acc} および新規結合グラフ G_{new} から、図 6 に示す新たな累積結合グラフ G_{acc} が生成される。

[0038] 次に、例外的な処理について説明する。新規結合グラフ G_{new} の文字候補情報 210 の左右に対応する累積結合グラフ G_{acc} 側の接続情報 220 が 1 つも見つからない場合は、その文字候補情報 210 は誤読である可能性が高いため、累積結合グラフ G_{acc} へのマージや追加は行わない。

[0039] また、新規結合グラフ G_{new} の文字候補情報 210 の左側に対応する累積結合グラフ G_{acc} 側の接続情報 220 が見つかったが、右側に対応する接続情報 220 がみつからない場合は、その文字候補情報 210 を累積結合グラフ G_{acc} に追加するとともに、その文字候補情報 210 の右側の接続情報 220 を、新たな終了位置 222 として累積結合グラフ G_{acc} に追加する。このとき、新たな終了位置 222 として追加する接続情報 220 が右ポインタと右接続位置を有する場合は、これらを削除する。また、新たな終了位置 222 として追加する接続情報 220 が、追加する文字候補情報 210 以外の文字候補情報 210 を指し示す左ポインタと左接続位置を有する場合は、これらも削除する。

[0040] また、新規結合グラフ G_{new} の文字候補情報 210 の右側に対応する累積結合グラフ G_{acc} 側の接続情報 220 が見つかったが、左側に対応する接続情報 220 がみつからない場合は、その文字候補情報 210 を累積結合グラフ G_{acc} に追加するとともに、その文字候補情報 210 の左側の接続情報 220 を、新たな開始位置 221 として累積結合グラフ G_{acc}

cに追加する。このとき、新たな開始位置221として追加する接続情報220が左ポインタと左接続位置を有する場合は、これらを削除する。また、新たな開始位置221として追加する接続情報220が、追加する文字候補情報210以外の文字候補情報210を指し示す右ポインタと右接続位置を有する場合は、これらも削除する。

[0041] また、新規結合グラフG_{new}の文字候補情報210の右側に対応する累積結合グラフG_{acc}側の接続情報220が開始位置221である場合、その文字候補情報210を開始位置221の左に接続される文字候補情報210として累積結合グラフG_{acc}に追加するとともに、累積結合グラフG_{acc}側の開始位置221にその文字候補情報210を指し示す左ポインタおよび左接続位置を追加し、フラグの属性を書き換えることによって、開始位置221を通常の接続情報220に変更する。また、その文字候補情報210の左側の接続情報220を、新たな開始位置221として累積結合グラフG_{acc}に追加する。このとき、新たな開始位置221として追加する接続情報220が左ポインタと左接続位置を有する場合は、これらを削除する。また、新たな開始位置221として追加する接続情報220が、追加する文字候補情報210以外の文字候補情報210を指し示す右ポインタと右接続位置を有する場合は、これらも削除する。

[0042] また、新規結合グラフG_{new}の文字候補情報210の左側に対応する累積結合グラフG_{acc}側の接続情報220が終了位置222である場合、その文字候補情報210を終了位置222の右に接続される文字候補情報210として累積結合グラフG_{acc}に追加するとともに、累積結合グラフG_{acc}側の終了位置222にその文字候補情報210を指し示す右ポインタおよび右接続位置を追加し、フラグの属性を書き換えることによって、終了位置222を通常の接続情報220に変更する。また、その文字候補情報210の右側の接続情報220を、新たな終了位置222として累積結合グラフG_{acc}に追加する。このとき、新たな終了位置222として追加する接続情報220が右ポインタと右接続位置を有する場合は、これら

を削除する。また、新たな終了位置 2 2 2 として追加する接続情報 2 2 0 が、追加する文字候補情報 2 1 0 以外の文字候補情報 2 1 0 を指し示す左ポイントと左接続位置を有する場合は、これらも削除する。

[0043] 累積結合グラフ G_{acc} は、複数の開始位置 2 2 1 や複数の終了位置 2 2 2 を持つ構成であってもよいが、これら開始位置 2 2 1 や終了位置 2 2 2 を 1 つに絞り込む必要がある場合は、以下のようにして絞り込みを行う。すなわち、複数の開始位置 2 2 1 のうち、最左端以外の開始位置 2 2 1 の右ポイントをすべて無効化する。同様に、複数の終了位置 2 2 2 のうち、最右端以外の終了位置 2 2 2 の左ポイントをすべて無効化する。文字候補情報 2 1 0 の右ポイントや左ポイントが指し示す接続情報 2 2 0 の対応するポイントが無効となっている場合には、その文字候補情報 2 1 0 の右ポイントや左ポイントも無効化する。この処理を無効化されるポイントがなくなるまで繰り返し実施する。そして、最後にポイントがすべて無効な接続情報 2 2 0 および文字候補情報 2 1 0 を削除する。

[0044] なお、以上は、隣り合う文字候補情報 2 1 0 同士の接続関係を接続情報 2 2 0 によって表す構成の結合グラフ G の統合処理について説明したが、文字候補情報 2 1 0 が隣り合う他の文字候補情報 2 1 0 を直接指し示す構成、つまり、文字候補情報 2 1 0 に接続情報 2 2 0 の機能も持たせた構成の結合グラフ G を用いる場合であっても、同様の統合処理を適用できる。この場合、上述の説明において、文字候補情報 2 1 0 の左右の接続情報 2 2 0 を、文字候補情報 2 1 0 内の接続情報に置き換えればよい。

[0045] 結合グラフ統合部 1 3 は、統合すべきすべての結合グラフ G について、以上説明したような統合処理を繰り返し、すべての結合グラフ G の統合が終了すると、統合された結合グラフ G を認識文字列生成部 1 4 または出力部 1 5 に渡す。

[0046] 認識文字列生成部 1 4 は、結合グラフ統合部 1 3 から統合された結合グラフ G を受け取り、この統合された結合グラフ G に対して、例えば知識処理などの所定の処理を施すことにより、最終的な文字認識結果である認識文字列

を生成する。そして、認識文字列生成部 14 は、生成した認識文字列を出力部 15 に渡す。なお、最終的な文字認識結果である認識文字列を生成するための知識処理などの処理は既存技術をそのまま利用できるように、詳細な説明は省略する。

[0047] 出力部 15 は、認識文字列生成部 14 により生成された認識文字列を出力する。また、出力部 15 は、認識文字列生成部 14 により生成された認識文字列に代えて、あるいは認識文字列とともに、結合グラフ統合部 13 によって統合された結合グラフ G を出力する構成としてもよい。出力部 15 が統合された結合グラフ G のみを出力する構成の場合、実施形態の文字認識装置 10 は、上述した認識文字列生成部 14 を備えない構成とすることができる。

[0048] 出力部 15 による認識文字列や統合された結合グラフ G の出力形態は、認識文字列や統合された結合グラフ G を表示装置 108 に表示する形態であってもよいし、ネットワークに接続された外部装置に、ネットワーク I/F 106 を介して認識文字列や統合された結合グラフ G を送信する形態であってもよい。

[0049] 次に、実施形態の文字認識装置 10 の動作を説明する。図 7 は、文字認識装置 10 による処理手順の一例を示すフローチャートである。文字認識装置 10 は、例えば、図 7 のフローチャートで示す一連の処理手順に従って動作する。

[0050] 文字認識装置 10 が動作を開始すると、まず、文字列画像取得部 11 が、文字認識処理の対象となる文字列画像 I S を取得し（ステップ S 101）、取得した文字列画像 I S に対して前処理を施して（ステップ S 102）、結合グラフ生成部 12 に渡す。

[0051] 次に、結合グラフ生成部 12 が、文字列画像取得部 11 から受け取った文字列画像 I S に対する文字認識処理を実行し（ステップ S 103）、文字列に対応する結合グラフ G を生成する（ステップ S 104）。本実施形態では、結合グラフ生成部 12 が、同一文字列を含む複数の文字列画像 I S に対してそれぞれ文字認識処理を行う、あるいは、1 つの文字列画像 I S に対して

複数の異なる文字認識処理を行うことで、1つの文字列に対応する複数の結合グラフGを生成する。結合グラフ生成部12により生成された複数の結合グラフGは、結合グラフ統合部13に順次渡される。

[0052] 次に、結合グラフ統合部13が、結合グラフ生成部12から受け取った複数の結合グラフG、つまり、1つの文字列に対応する複数の結合グラフGの統合処理を実行し（ステップS105）、統合した結合グラフGを認識文字列生成部14に渡す。なお、上述したように、出力部15が統合された結合グラフGを出力する構成の場合は、結合グラフ統合部13は、統合した結合グラフGを出力部15に渡す。

[0053] 次に、認識文字列生成部14が、結合グラフ統合部13から受け取った統合された結合グラフGに基づいて、最終的な文字認識結果である認識文字列を生成し（ステップS106）、この認識文字列を出力部15に渡す。なお、出力部15が統合された結合グラフGのみを出力する構成の場合は、このステップS106の処理は省略される。

[0054] 最後に、出力部15が、認識文字列生成部14から受け取った認識文字列を出力する（ステップS107）。なお、出力部15は、認識文字列に代えて、あるいは認識文字列とともに、結合グラフ生成部12から受け取った統合された結合グラフGを出力してもよい。

[0055] 図8は、図7のステップS105の統合処理の概要を説明するフローチャートであり、新規結合グラフG_{new}を累積結合グラフG_{acc}に順次統合する統合処理の手順を示している。なお、図中のiはカウンタ値を示し、nは統合すべき結合グラフGの数を示している。

[0056] 統合処理が開始されると、結合グラフ統合部13は、まずカウンタ値iを初期化（ $i = 0$ ）する（ステップS201）。その後、結合グラフ生成部12により結合グラフGが生成されると、結合グラフ統合部13は、結合グラフ生成部12から結合グラフGを受け取り（ステップS202）、カウンタ値iをインクリメント（ $i = i + 1$ ）する（ステップS203）。

[0057] 次に、結合グラフ統合部13は、カウンタ値iが1であるかどうかを確認

することにより、ステップS 2 0 2で受け取った結合グラフGが、統合すべき複数の結合グラフGのうちの最初の結合グラフGかどうかを判定する（ステップS 2 0 4）。

[0058] ここで、ステップS 2 0 2で受け取った結合グラフGが最初の結合グラフGである場合（ステップS 2 0 4 : Y e s）、結合グラフ統合部1 3は、その結合グラフGをそのまま累積結合グラフG__a c cとして保存する（ステップS 2 0 6）。一方、ステップS 2 0 2で受け取った結合グラフGが最初の結合グラフGではない場合（ステップS 2 0 4 : N o）、結合グラフ統合部1 3は、その結合グラフGを新規結合グラフG__n e wとして、保存している累積結合グラフG__a c cに統合する（ステップS 2 0 5）。そして、統合した結合グラフGを、新たな累積結合グラフG__a c cとして保存する（ステップS 2 0 6）。

[0059] その後、結合グラフ統合部1 3は、カウンタ値iがnに達したか否かを判断することにより、統合すべきすべての結合グラフGを統合したか否かを判定する（ステップS 2 0 7）。そして、統合していない結合グラフGがあれば（ステップS 2 0 7 : N o）、ステップS 2 0 2に戻って以降の処理を繰り返し、すべての結合グラフGを統合すると（ステップS 2 0 7 : Y e s）、保存している累積結合グラフG__a c cを認識文字列生成部1 4や出力部1 5に渡して、一連の処理を終了する。

[0060] 図9は、図8のステップS 2 0 5の処理手順を示すフローチャートである。なお、図中のjはカウンタ値を示し、mは新規結合グラフG__n e wに含まれる文字候補情報2 1 0の数を示している。

[0061] 結合グラフ統合部1 3は、まずカウンタ値jを初期化（j = 0）する（ステップS 3 0 1）。その後、結合グラフ統合部1 3は、新規結合グラフG__n e wの左から順に文字候補情報2 1 0を1つ取り出し（ステップS 3 0 2）、カウンタ値jをインクリメント（j = j + 1）する（ステップS 3 0 3）。

[0062] 次に、結合グラフ統合部1 3は、ステップS 3 0 2で取り出した文字候補

情報 210、すなわち新規結合グラフ G_{new} 側の左から j 番目の文字候補情報 210 の左右に対応する累積結合グラフ G_{acc} 側の 2 つの接続情報 220 を特定する（ステップ S304）。そして、結合グラフ統合部 13 は、ステップ S304 で特定した 2 つの接続情報 220 に挟まれる 1 つの文字候補情報 210 が、累積結合グラフ G_{acc} 側に存在するか否かを判定する（ステップ S305）。

[0063] ここで、累積結合グラフ G_{acc} 側にそのような文字候補情報 210 があれば（ステップ S305：Yes）、結合グラフ統合部 13 は、その文字候補情報 210 を新規結合グラフ G_{new} 側の左から j 番目の文字候補情報 210 に対応する累積結合グラフ G_{acc} 側の文字候補情報 210 とみなし、新規結合グラフ G_{new} 側の左から j 番目の文字候補情報 210 を、その累積結合グラフ G_{acc} 側の文字候補情報 210 にマージする（ステップ S306）。一方、累積結合グラフ G_{acc} 側にそのような文字候補情報 210 がなければ（ステップ S305：No）、結合グラフ統合部 13 は、新規結合グラフ G_{new} 側の左から j 番目の文字候補情報 210 に対応する文字候補情報 210 が累積結合グラフ G_{acc} にはないと判断し、新規結合グラフ G_{new} 側の左から j 番目の文字候補情報 210 を累積結合グラフ G_{acc} に追加する（ステップ S307）。

[0064] その後、結合グラフ統合部 13 は、カウンタ値 j が m に達したか否かを判断することにより、新規結合グラフ G_{new} に含まれるすべての文字候補情報 210 に対する処理が終了したか否かを判定する（ステップ S308）。そして、処理が終了していない文字候補情報 210 があれば（ステップ S308：No）、ステップ S302 に戻って以降の処理を繰り返し、すべての文字候補情報 210 に対する処理が終了すると（ステップ S308：Yes）、一連の処理を終了する。

[0065] 以上、具体的な例を挙げながら詳細に説明したように、実施形態の文字認識装置 10 は、文字列画像 I_S に対する文字認識処理により各々が 1 以上の候補文字を含む文字候補情報 210 を接続した結合グラフ G を生成し、1 つ

の文字列に対して生成した複数の結合グラフGを統合して、統合した結合グラフG、または、統合した結合グラフGに基づいて生成した認識文字列を出力する。したがって、複数の文字認識結果から対応する文字について信頼度の高い認識結果を選択して最終的な認識文字列を得る従来の方法に比べて、誤読や文字の区切りの誤りに対して頑強な認識結果を出力することができ、高精度な文字認識を行うことが可能となる。

[0066] 以下では、上述した実施形態の変形例について説明する。

<変形例1>

上述した実施形態では、複数の結合グラフGにおける文字候補情報210の対応付けを文字候補情報210に含まれる位置情報に基づいて行っているが、複数の結合グラフGが異なる文字列画像ISから生成されたものである場合、対応する文字候補情報210の位置情報が必ずしも一致しているとは限らない。上述した実施形態では、位置情報の一致判定に誤差範囲を設けているが、同一の文字列を含む複数の文字列画像ISにおいて、同じ文字が存在する位置が大きく異なることも想定される。

[0067] そこで、同一の文字列を含む複数の文字列画像ISから生成された複数の結合グラフGを統合する場合には、複数の文字列画像ISの位置合わせ（レジストレーション）を行い、位置合わせの結果に応じて変換した位置情報に基づいて、複数の結合グラフGにおける文字候補情報210同士の対応付けを行うようにしてもよい。

[0068] この場合、結合グラフ統合部13は、結合グラフ生成部12から結合グラフGを受け取る際に、その結合グラフGの生成に用いた文字列画像ISも受け取る。そして、結合グラフGの統合を行う際には、まず、文字列画像ISの位置合わせを行い、その位置合わせの結果に応じて、統合する結合グラフGに含まれる各文字候補情報210の位置情報を変換する。そして、変換した位置情報を用いて、上述した実施形態と同様の方法により文字候補情報210の対応付けを行う。なお、画像の位置合わせ（レジストレーション）は既存技術をそのまま適用できるため、詳細な説明は省略する。

[0069] 本変形例では、文字列画像 I S の位置合わせの結果に応じて変換した位置情報に基づいて、複数の結合グラフ G における文字候補情報 210 の対応付けを行うことにより、複数の文字列画像 I S において同じ文字が存在する位置が大きく異なる場合であっても、文字候補情報 210 の対応付けを適切に行って、高精度な文字認識を行うことが可能となる。

[0070] <変形例 2>

複数の結合グラフ G における文字候補情報 210 の対応付けは、文字候補情報 210 の位置情報だけでなく、隣り合う文字候補情報 210 の接続性を手掛かりとして行うこともできる。以下では、隣り合う文字候補情報 210 の接続性を手掛かりとして文字候補情報 210 の対応付けの方法の一例について説明する。

[0071] 図 10 は、図 5 に例示した累積結合グラフ G_{acc} および新規結合グラフ G_{new} の一部の文字候補情報 210 を抜き出して示す図である。この図 10 において、累積結合グラフ G_{acc} 側の文字候補情報 210 (A1, A2, A5) と新規結合グラフ G_{new} の側の文字候補情報 210 (B1, B2, B5) とを結ぶ線が、それぞれの文字候補情報 210 の対応付けの候補を表している。図 10 に示すように、1 つの文字候補情報 210 が複数の対応付けの候補を有する。

[0072] 本変形例では、このような対応付けの候補の 1 つ 1 つにそれぞれスコアを準備する。スコアの初期値としては、文字列における各文字の相対的な位置関係から求まる位置ずれ量や認識結果の近さなどからスコアを設定する。例えば、文字列内の座標値を左上が 0、右下が 1 となるように正規化して表現し、その正規化座標値に基づいてスコアを計算する。具体的には、累積結合グラフ G_{acc} 側の文字候補情報 210 の正規化座標値と新規結合グラフ G_{new} 側の文字候補情報 210 の正規化座標値との差の絶対値の二乗を計算し、それら全部の和を取るなどの方法がある。また、累積結合グラフ G_{acc} 側の文字候補情報 210 と新規結合グラフ G_{new} 側の文字候補情報 210 とで同じ文字コードがある場合に、それに対応する認識スコアの

和を取り、最もよい認識スコアとなる文字コードを見つけ、その文字コードの認識スコアに基づいて、ここでの対応付け候補のスコアを決めるようにしてもよい。また、以上のような2つのスコアを合成することで、ここでの対応付け候補のスコアを決めるようにしてもよい。

[0073] 次に、新規結合グラフ G_{new} において隣り合う2つの文字候補情報210について、これらの文字候補情報210の対応付け候補となっている累積結合グラフ G_{acc} 側の隣り合う2つの文字候補情報210の組を探し出す。一般には、このような文字候補情報210の組は複数見つかる。

[0074] 次に、新規結合グラフ G_{new} 側の2つの文字候補情報210と累積結合グラフ G_{acc} 側の2つの文字候補情報210との対応付け候補のスコアに基づいて、それぞれのスコアを更新する。例えば、双方の対応付け候補のスコアが平均スコアを上回れば、それぞれのスコアに対して所定の定数を加算し、双方の対応付け候補のスコアが平均スコアを下回れば、それぞれのスコアから所定の定数を減算し、それ以外はスコアを加減算を行わない。この処理を繰り返すことによって、尤もらしい対応付け候補のスコアが上昇し、尤もらしくない対応付け候補のスコアが下降することになる。以上の処理を一定回数、あるいはスコアの変動が閾値以下になるまで行う。

[0075] 次に、対応付け候補のスコアが大きい順から、新規結合グラフ G_{new} 側の文字候補情報210と累積結合グラフ G_{acc} 側の文字候補情報210との対応付けを決定していくが、その過程では、すでに対応付けが決まった文字候補情報210を含む対応付けは採用しないようにする。また、対応付け候補のスコアが閾値を下回る場合は、その文字候補情報210同士の間に対応付けを採用しないようにする。これにより、最終的に妥当な文字候補情報210同士の間に対応付けを得ることができる。なお、ここでの対応付けは、新規結合グラフ G_{new} 側と累積結合グラフ G_{acc} 側のすべての文字候補情報210を1対1に対応付けるものではなく、1対1に対応しない文字候補情報210、つまり1対0や0対1の対応付けを含むものである。

[0076] 以上のような対応付けの方法は、緩和法として知られている方法である。

上述した実施形態の文字認識装置 10 は、結合グラフ統合部 13 における結合グラフ G の統合処理において、以上のような緩和法により文字候補情報 210 の対応付けを行うようにしてもよい。これにより、位置情報に基づく文字候補情報 210 の対応付けが困難な場合であっても、文字候補情報 210 の対応付けを適切に行って、高精度な文字認識を行うことが可能となる。

[0077] <変形例 3>

次に、複数の結合グラフ G の統合方法の他の例を説明する。本例の統合方法は、複数の接続パスを有する累積結合グラフ G_{acc} と、複数の接続パスを有する新規結合グラフ G_{new} のそれぞれを単一の接続パスに分離する。そして、累積結合グラフ G_{acc} 側と新規結合グラフ G_{new} 側との接続パスの対応関係を特定し、対応する接続パスに含まれる文字候補情報 210 同士をマージする。また、累積結合グラフ G_{acc} 側のいずれの接続パスにも対応しない新規結合グラフ G_{new} 側の接続パスについては、その接続パスに含まれる文字候補情報 210 を、累積結合グラフ G_{acc} 側のいずれかの接続パスに追加する。その後、累積結合グラフ G_{acc} 側のすべての接続パスを結合して、新たな累積結合グラフ G_{acc} とする。

[0078] 図 11 は、結合グラフ G を単一の接続パスに分離する様子を示す図である。結合グラフ G から分離された単一の接続パスの集合を、以下では多重単線パス MP という。多重単線パス MP は、結合グラフ G に含まれる文字候補情報 210 を左から順に辿り、分岐があるごとに個別の接続パスを生成することで構築することができる。また、このとき、生成した各接続パスに含まれる各文字候補情報 210 が元の結合グラフ G 内のどの文字候補情報 210 に由来するかのデータを付属させておく。また、例えば文字候補情報 210 に含まれる認識スコアなどから接続パスのスコアを計算し、上位 n 個の接続パスのみを残す、あるいは、スコアが閾値以上の接続パスのみを残すようにして、多重単線パス MP に含まれる接続パスの数に制限を設けるようにしてもよい。

[0079] 本例では、以上のような接続パスの分離を、累積結合グラフ G_{acc} と

新規結合グラフ G_new の双方に対して行う。そして、累積結合グラフ G_acc 側の接続パスと新規結合グラフ G_new 側の接続パスとの対応関係を、それぞれの接続パスに含まれる文字候補情報 210 同士のマッチングスコアを用いて特定する。具体的には、以下の方法により、累積結合グラフ G_acc 側の接続パスと新規結合グラフ G_new 側の接続パスとの対応関係を特定する。

[0080] 累積結合グラフ G_acc 側の接続パスにおいて連続する文字候補情報 210 を $A_0, A_1, \dots, A_{n-1}$ とし、新規結合グラフ G_new 側の接続パスにおいて連続する文字候補情報 210 を $B_0, B_1, \dots, B_{m-1}$ とする。文字候補情報 210 同士のマッチングスコアは、それぞれの文字候補情報 210 に含まれる認識スコア、文字領域の位置やサイズなどを用いて計算する。このような文字候補情報 210 同士のマッチングスコアを接続パスの先頭から所定数の文字候補情報 210 の組み合わせについて計算し、その中で最良のマッチングスコアが得られた文字候補情報 210 同士を特定する。そして、累積結合グラフ G_acc 側の接続パスと新規結合グラフ G_new 側の接続パスのそれぞれで、最良のマッチングスコアが得られた文字候補情報 210 の次の文字候補情報 210 から所定数の文字候補情報 210 の組み合わせについて、文字候補情報 210 同士のマッチングスコアを同様に計算する。そして、得られた最良のマッチングスコアをそれまでのマッチングスコアに加算していく。

[0081] 今、 A_{k-1} と B_{h-1} のマッチングスコアが最良であったとする。この場合、次のステップでは、 $A_k \sim A_{k+d-1}$ の d 個の文字候補情報 210 と、 $B_h \sim B_{h+d-1}$ の d 個の文字候補情報 210 との間で、合計 $2d-1$ 組の文字候補情報 210 の組み合わせについて各々マッチングスコアを計算する。そして、得られたマッチングスコアのうちの最良のマッチングスコアを、 A_{k-1} と B_{h-1} までの処理で得られているマッチングスコアに加算する。このとき、累積結合グラフ G_acc 側の接続パスと新規結合グラフ G_new 側の接続パスにおいて、最良のマッチングスコアが得られた文

字候補情報 2 1 0 同士が連続しない場合、その間の文字候補情報 2 1 0 の数に応じて、マッチングスコアが低くなるように調整する。この処理を累積結合グラフ G_{acc} 側の接続パスの最後の文字候補情報 2 1 0 と新規結合グラフ G_{new} 側の接続パスの最後の文字候補情報 2 1 0 との組み合わせまで行い、累積結合グラフ G_{acc} 側の接続パスと新規結合グラフ G_{new} 側の接続パスとの最終的なマッチングスコアを得る。ここで使うスコア計算法はレーベンシュタイン距離と呼ばれるものの一種であり、マッチング方式はダイナミックプログラミング (DP) と呼ばれるものである。ただし、スコア計算の方法やマッチング方式は、以上の例に限定されるものではない。

[0082] 以上の説明では、2 d - 1 組の文字候補情報 2 1 0 の組み合わせのうち、最良のマッチングスコアが得られた文字候補情報 2 1 0 の組み合わせをマッチングがとれた文字候補情報 2 1 0 の組み合わせとみなして処理を進めているが、マッチングスコアが高い順から上位 T 個の組み合わせを候補として残り、それぞれの組み合わせに対して上述と同様の処理を行うようにしてもよい。このような上位 T 個の組み合わせを残す手法はビームサーチと呼ばれる。

[0083] 本例では、累積結合グラフ G_{acc} 側のすべての接続パスと、新規結合グラフ G_{new} 側のすべての接続パスとの組み合わせについて、以上の処理により接続パス同士のマッチングスコアを計算する。そして、マッチングスコアが最大となる累積結合グラフ G_{acc} 側の接続パスと新規結合グラフ G_{new} 側の接続パスとの組みを特定し、そのマッチングスコアが所定の閾値を超える場合に、これらの接続パスが対応しているものとして、これらの接続パスに含まれる文字候補情報 2 1 0 同士を上述した実施形態と同様の方法によりマージする。一方、マッチングスコアが閾値以下の接続パスの組については、新規結合グラフ G_{new} 側の接続パスに含まれる文字候補情報 2 1 0 を、上述した実施形態と同様の方法により、累積結合グラフ G_{acc} 側の接続パスに追加する。そして最後に、累積結合グラフ G_{acc}

側のすべての接続パスを結合して、新たな累積結合グラフ G_{acc} とする。

[0084] 上述した実施形態の文字認識装置 10 は、結合グラフ統合部 13 における結合グラフ G の統合処理を、以上説明した本例の方法により行うようにしてもよい。これにより、累積結合グラフ G_{acc} や新規結合グラフ G_{new} の接続パスの数が多い場合であっても、結合グラフ G の統合処理を適切に行って、高精度な文字認識を行うことが可能となる。

[0085] <補足説明>

以上説明した実施形態の文字認識装置 10 における各機能は、例えば、文字認識装置 10 のハードウェア構成としてコンピュータを用いる場合、このコンピュータで所定のプログラムを実行することにより実現できる。文字認識装置 10 として用いるコンピュータで実行されるプログラムは、例えば、インストール可能な形式又は実行可能な形式のファイルで CD-ROM (Compact Disk Read Only Memory)、フレキシブルディスク (FD)、CD-R (Compact Disk Recordable)、DVD (Digital Versatile Disc) などのコンピュータで読み取り可能な記録媒体に記録されてコンピュータプログラムプロダクトとして提供される。

[0086] また、文字認識装置 10 として用いるコンピュータで実行されるプログラムを、インターネットなどのネットワークに接続された他のコンピュータ上に格納し、ネットワーク経由でダウンロードさせることにより提供するように構成してもよい。また、文字認識装置 10 として用いるコンピュータで実行されるプログラムをインターネットなどのネットワーク経由で提供または配布するように構成してもよい。また、文字認識装置 10 として用いるコンピュータで実行されるプログラムを、コンピュータ内部の ROM 102 などに予め組み込んで提供するように構成してもよい。

[0087] 文字認識装置 10 として用いるコンピュータで実行されるプログラムは、文字認識装置 10 の上述した機能的な構成要素（文字列画像取得部 11、結合グラフ生成部 12、結合グラフ統合部 13、認識文字列生成部 14、およ

び出力部 15) を含むモジュール構成となっており、実際のハードウェアとしては、例えば、CPU 101 が上記記録媒体からプログラムを読み出して実行することにより、上記の各構成要素が RAM 103 などの主記憶部上にロードされ、上記の各構成要素が主記憶部上に生成されるようになっている。なお、文字認識装置 10 の機能的な構成要素は、その一部または全部を、ASIC (Application Specific Integrated Circuit) や FPGA (Field-Programmable Gate Array) などの専用のハードウェアを用いて実現することも可能である。

[0088] 以上、本発明の実施形態を説明したが、この実施形態は例として提示したものであり、発明の範囲を限定することは意図していない。この新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、請求の範囲に記載された発明とその均等の範囲に含まれる。

請求の範囲

[請求項1]

文字列画像を取得する文字列画像取得部と、

前記文字列画像に対して文字認識処理を行い、1つの文字とみなされる文字領域ごとの認識結果を表す文字候補情報であって各々が1以上の候補文字を含む複数の前記文字候補情報を、前記文字列画像における各文字領域の並び順に従って接続した結合グラフを生成する結合グラフ生成部と、

同一文字列を含む複数の前記文字列画像から生成された複数の前記結合グラフ、または、1つの前記文字列画像に対して複数の異なる文字認識処理を行うことで生成された複数の前記結合グラフを統合する結合グラフ統合部と、

統合された前記結合グラフ、または、統合された前記結合グラフに基づいて得られる認識文字列を出力する出力部と、を備える文字認識装置。

[請求項2]

前記結合グラフ統合部は、統合する複数の前記結合グラフを第1の結合グラフおよび第2の結合グラフとしたときに、前記第1の結合グラフに含まれる前記文字候補情報と前記第2の結合グラフに含まれる前記文字候補情報との対応関係を特定し、前記第1の結合グラフと前記第2の結合グラフとで対応する前記文字候補情報同士を1つの前記文字候補情報に併合するとともに、前記第2の結合グラフに含まれる前記文字候補情報であって、前記第1の結合グラフに含まれるいずれの前記文字候補情報にも対応しない前記文字候補情報を前記第1の結合グラフに追加することにより、前記第1の結合グラフと前記第2の結合グラフとを統合する、請求項1に記載の文字認識装置。

[請求項3]

前記文字候補情報は、前記文字列画像における文字領域の位置を表す位置情報を含み、

前記結合グラフ統合部は、前記位置情報に基づいて、前記第1の結合グラフに含まれる前記文字候補情報と前記第2の結合グラフに含ま

れる前記文字候補情報との対応関係を特定する、請求項 2 に記載の文字認識装置。

[請求項4] 前記結合グラフ統合部は、同一文字列を含む複数の前記文字列画像から生成された複数の前記結合グラフを統合する場合に、複数の前記文字列画像の位置合わせを行って、当該位置合わせの結果に応じて変換した前記位置情報に基づいて、前記第 1 の結合グラフに含まれる前記文字候補情報と前記第 2 の結合グラフに含まれる前記文字候補情報との対応関係を特定する、請求項 3 に記載の文字認識装置。

[請求項5] 前記結合グラフ統合部は、前記第 1 の結合グラフに含まれる前記文字候補情報と前記第 2 の結合グラフに含まれる前記文字候補情報との対応関係を緩和法により特定する、請求項 2 に記載の文字認識装置。

[請求項6] 前記結合グラフは、前記文字列画像における文字領域の区切りが異なる複数のパターンに対応して、各パターンにおける前記文字候補情報の繋がりを表す接続パスを複数含み、

前記結合グラフ統合部は、前記第 1 の結合グラフと前記第 2 の結合グラフをそれぞれ単一の前記接続パスに分離した後、前記第 1 の結合グラフの前記接続パスと前記第 2 の結合グラフの前記接続パスとの対応関係を特定し、前記第 1 の結合グラフと前記第 2 の結合グラフとで対応する前記接続パスに含まれる前記文字候補情報同士を 1 つの前記文字候補情報に併合するとともに、前記第 2 の結合グラフの前記接続パスであって、前記第 1 の結合グラフのいずれの前記接続パスにも対応しない前記接続パスに含まれる前記文字候補情報を、前記第 1 の結合グラフのいずれかの前記接続パスに追加した上で、前記第 1 の結合グラフの複数の前記接続パスを結合することにより、前記第 1 の結合グラフと前記第 2 の結合グラフとを統合する、請求項 2 に記載の文字認識装置。

[請求項7] 前記結合グラフは、隣り合う前記文字候補情報同士の接続関係を表す接続情報を含み、

前記結合グラフ統合部は、前記第1の結合グラフに含まれる前記接続情報に対し、前記第2の結合グラフに含まれる前記文字候補情報との接続関係を追加することにより、前記第2の結合グラフに含まれる前記文字候補情報を前記第1の結合グラフに追加する、請求項2乃至6のいずれか一項に記載の文字認識装置。

[請求項8]

文字列画像を取得するステップと、

前記文字列画像に対して文字認識処理を行い、1つの文字とみなされる文字領域ごとの認識結果を表す文字候補情報であって各々が1以上の候補文字を含む複数の前記文字候補情報を、前記文字列画像における各文字領域の並び順に従って接続した結合グラフを生成するステップと、

同一文字列を含む複数の前記文字列画像から生成された複数の前記結合グラフ、または、1つの前記文字列画像に対して複数の異なる文字認識処理を行うことで生成された複数の前記結合グラフを統合するステップと、

統合された前記結合グラフ、または、統合された前記結合グラフに基づいて得られる認識文字列を出力するステップと、を含む文字認識方法。

[請求項9]

コンピュータに、

文字列画像を取得する機能と、

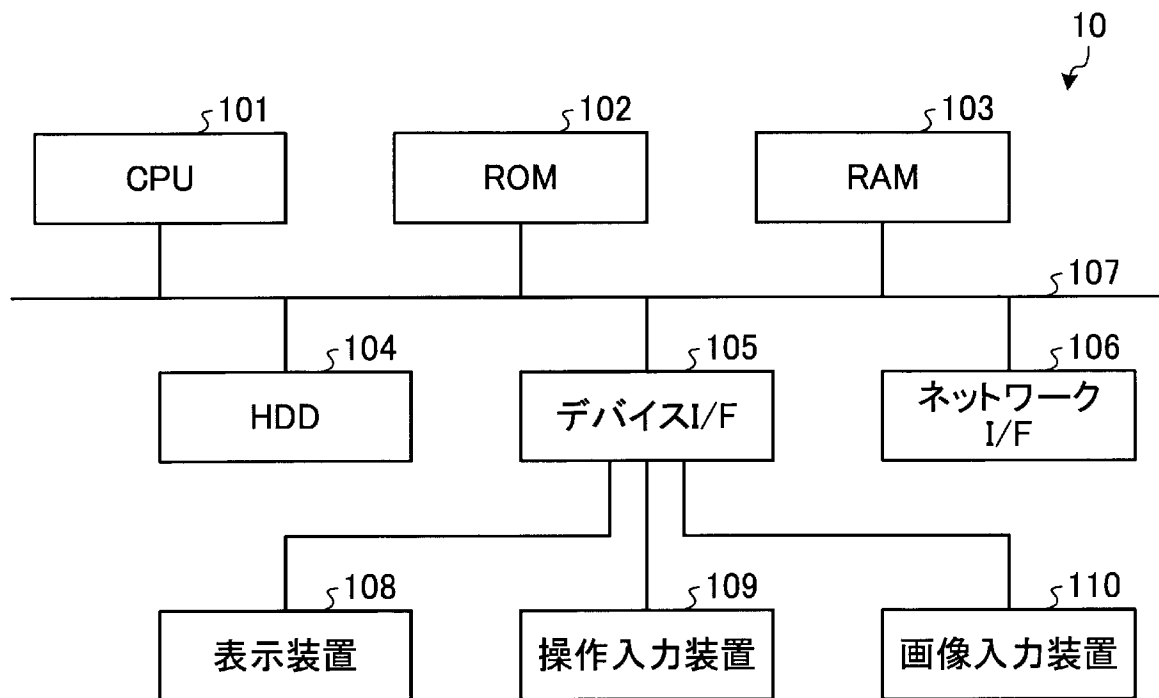
前記文字列画像に対して文字認識処理を行い、1つの文字とみなされる文字領域ごとの認識結果を表す文字候補情報であって各々が1以上の候補文字を含む複数の前記文字候補情報を、前記文字列画像における各文字領域の並び順に従って接続した結合グラフを生成する機能と、

同一文字列を含む複数の前記文字列画像から生成された複数の前記結合グラフ、または、1つの前記文字列画像に対して複数の異なる文字認識処理を行うことで生成された複数の前記結合グラフを統合する

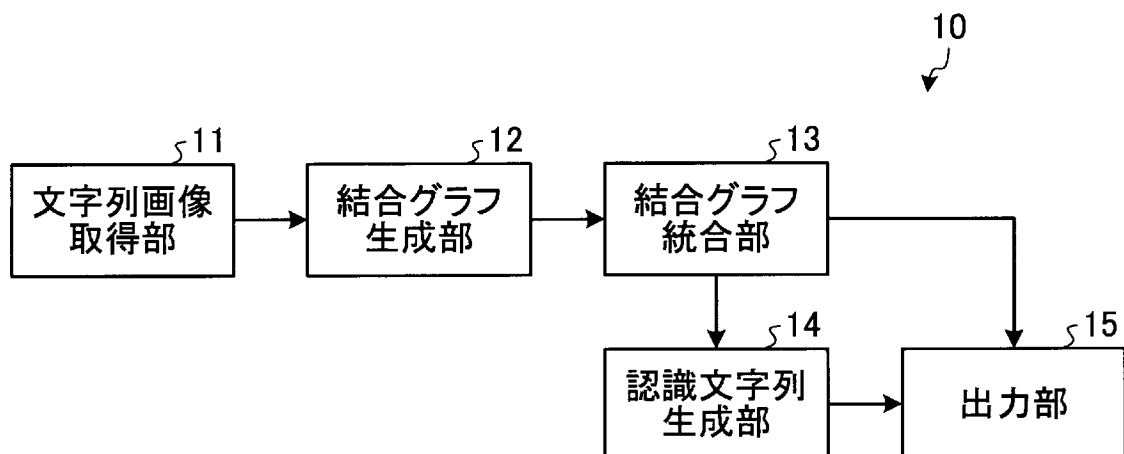
機能と、

統合された前記結合グラフ、または、統合された前記結合グラフに基づいて得られる認識文字列を出力する機能と、を実現させるためのプログラム。

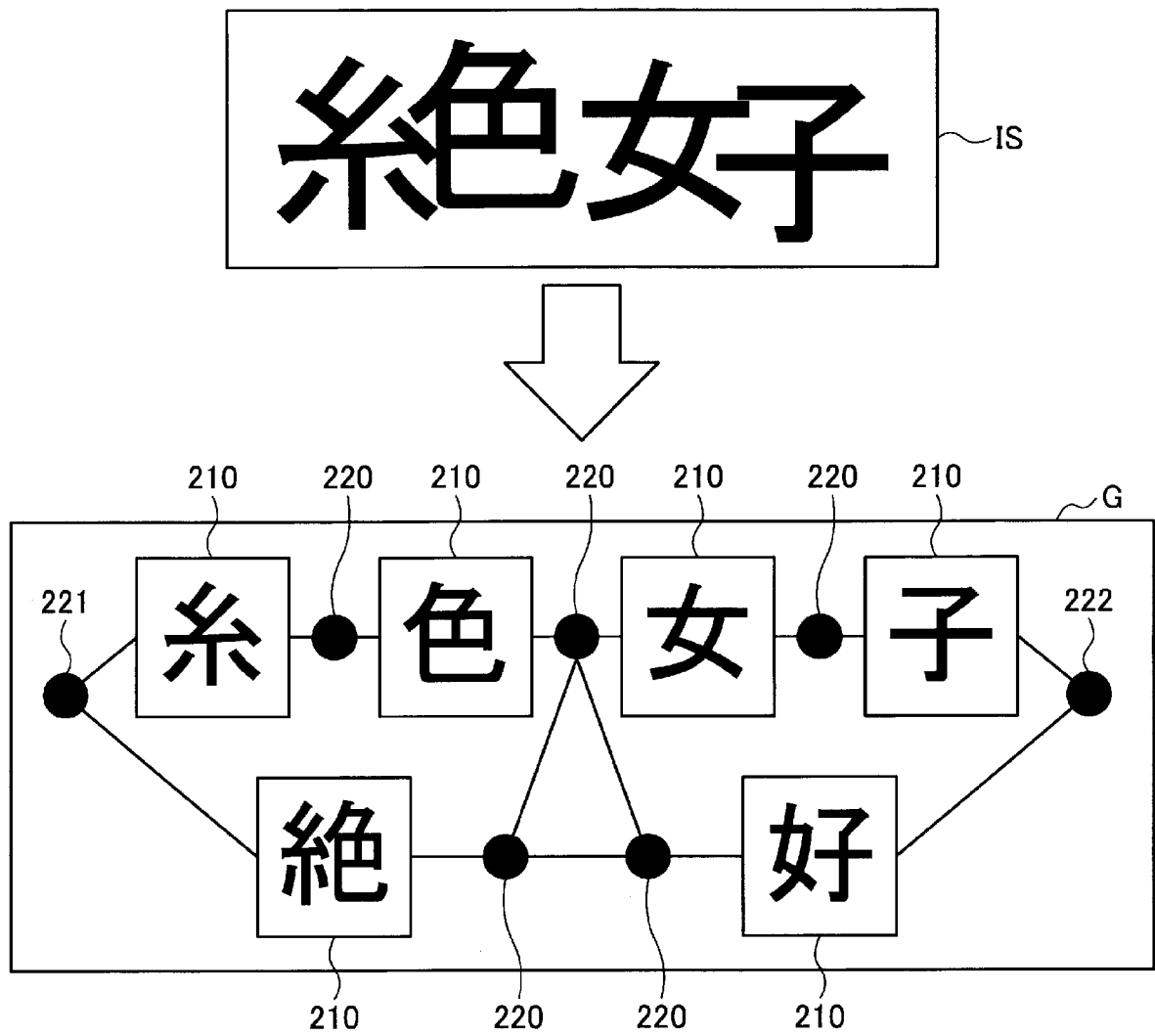
[図1]



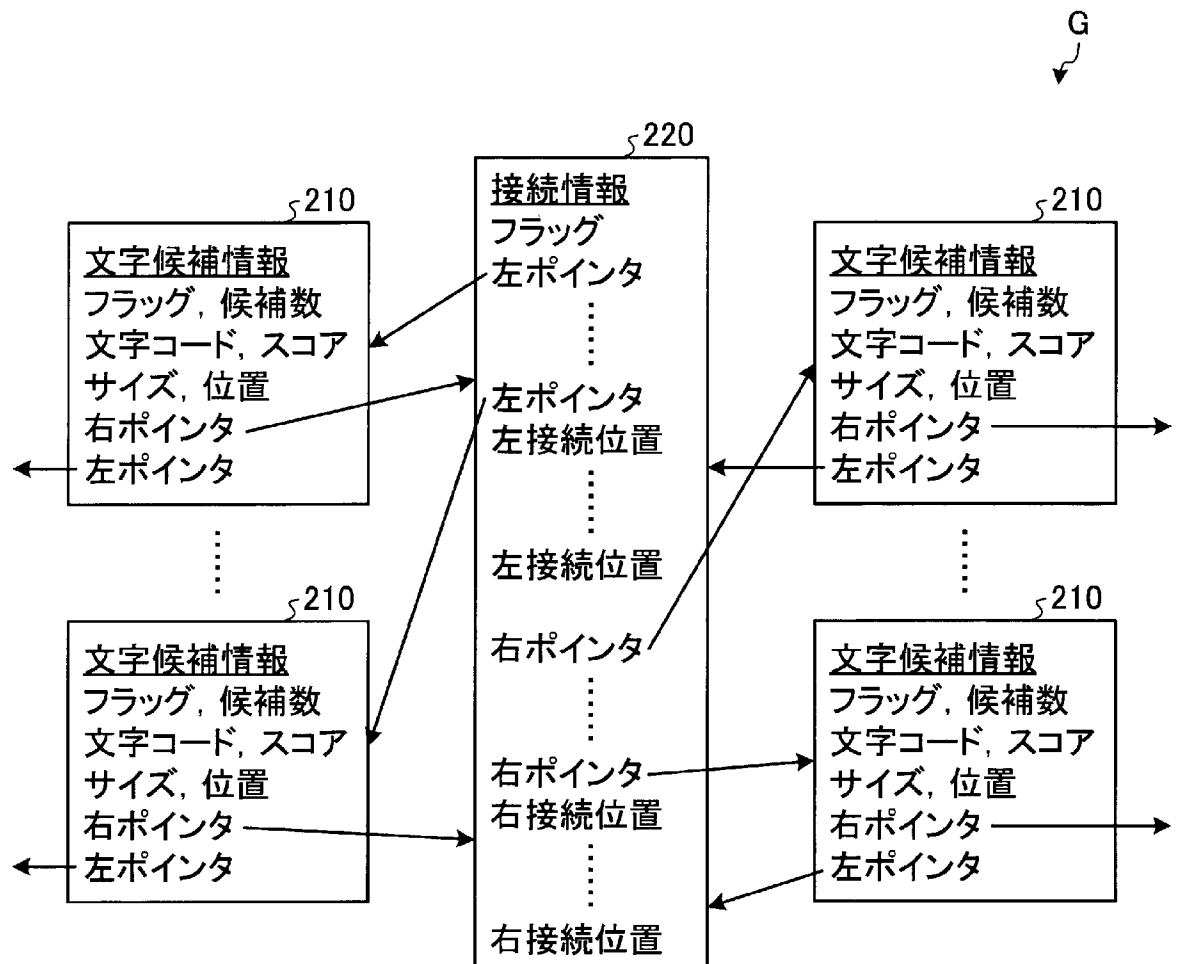
[図2]



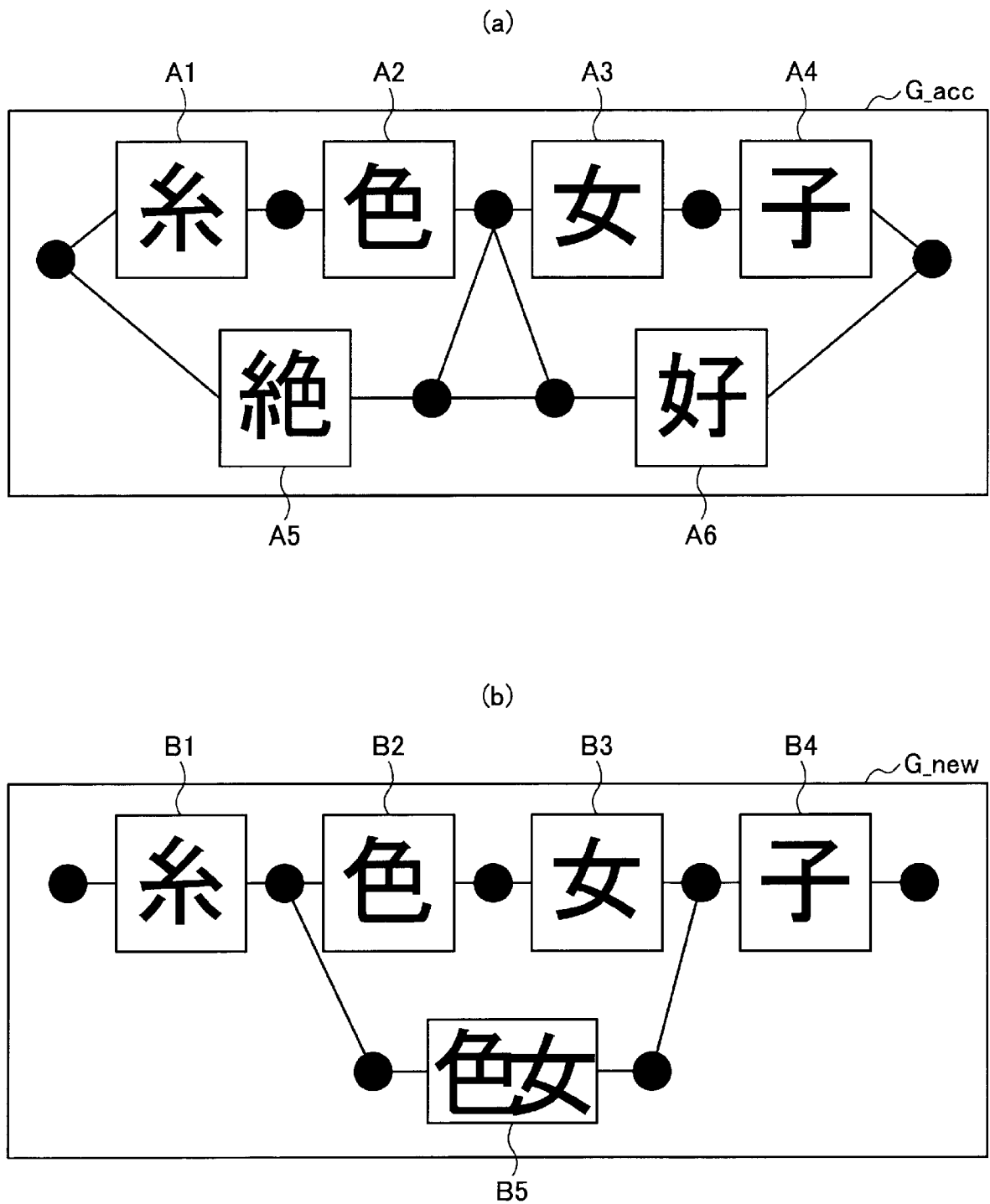
[図3]



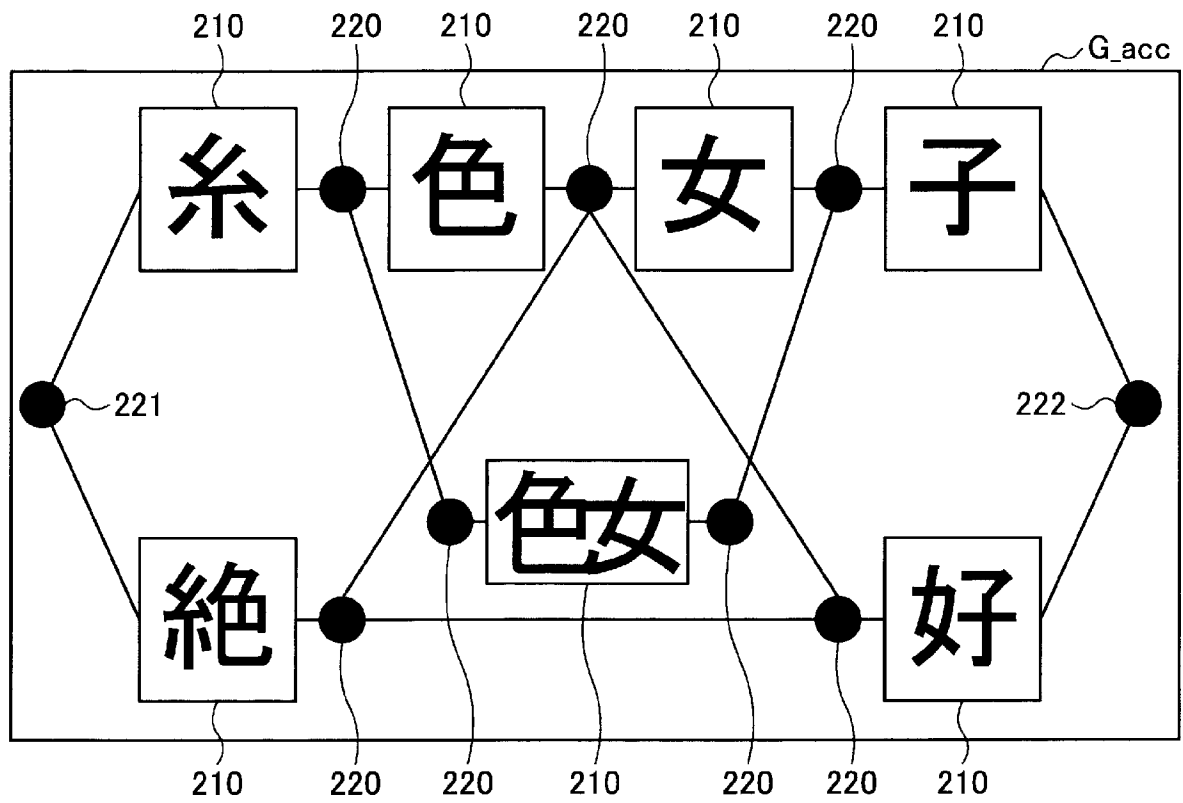
[図4]



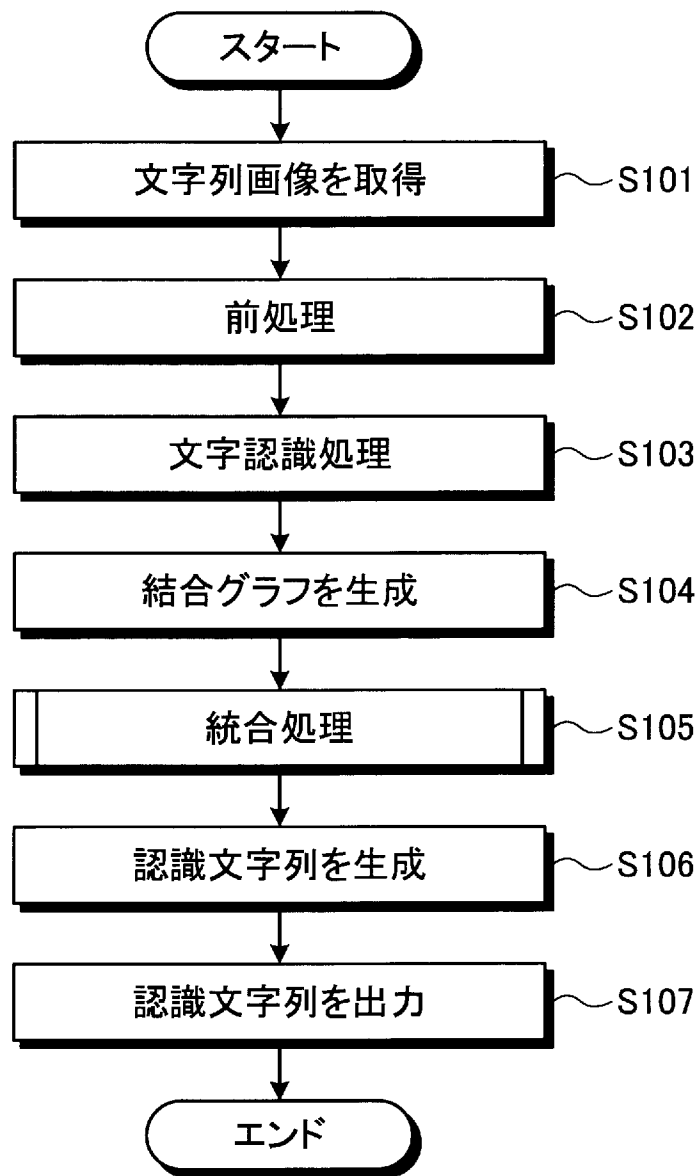
[図5]



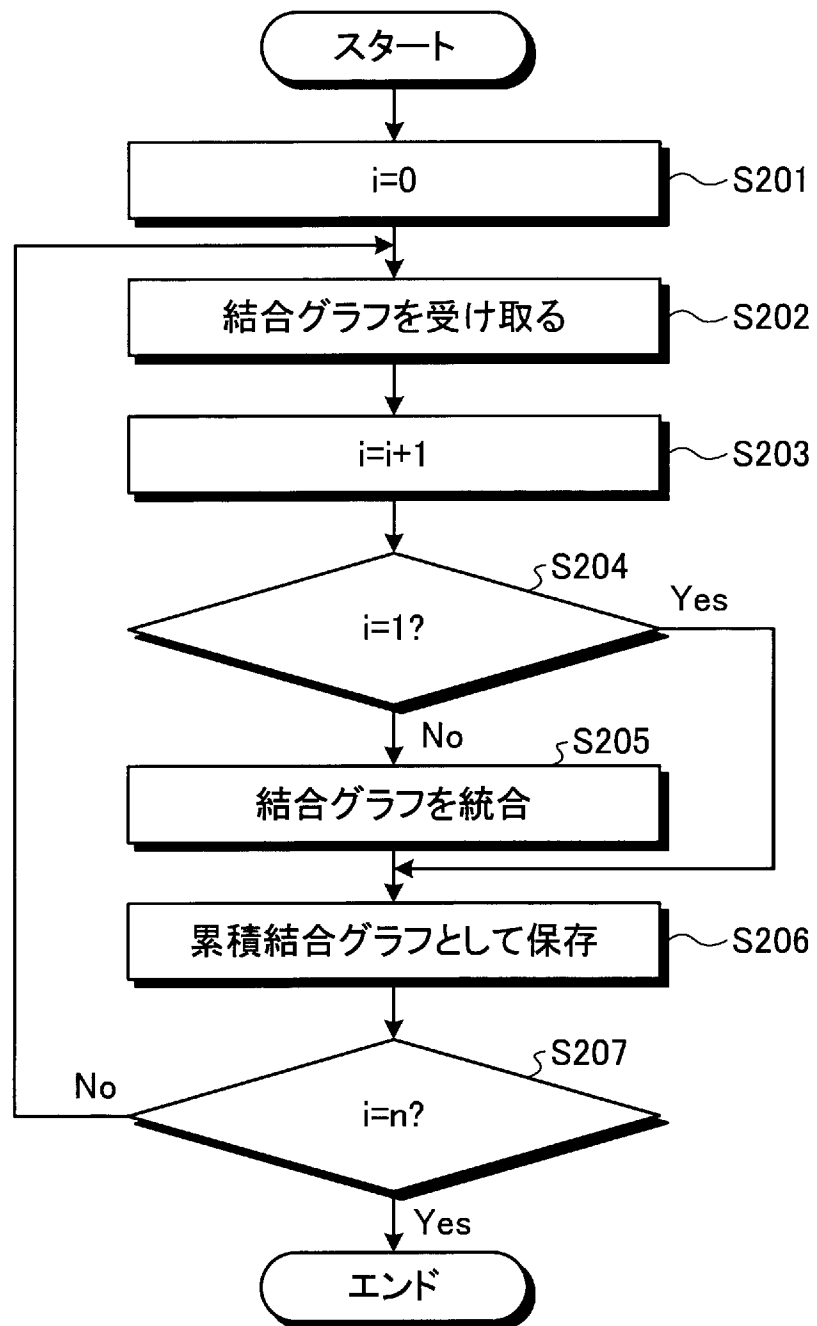
[図6]



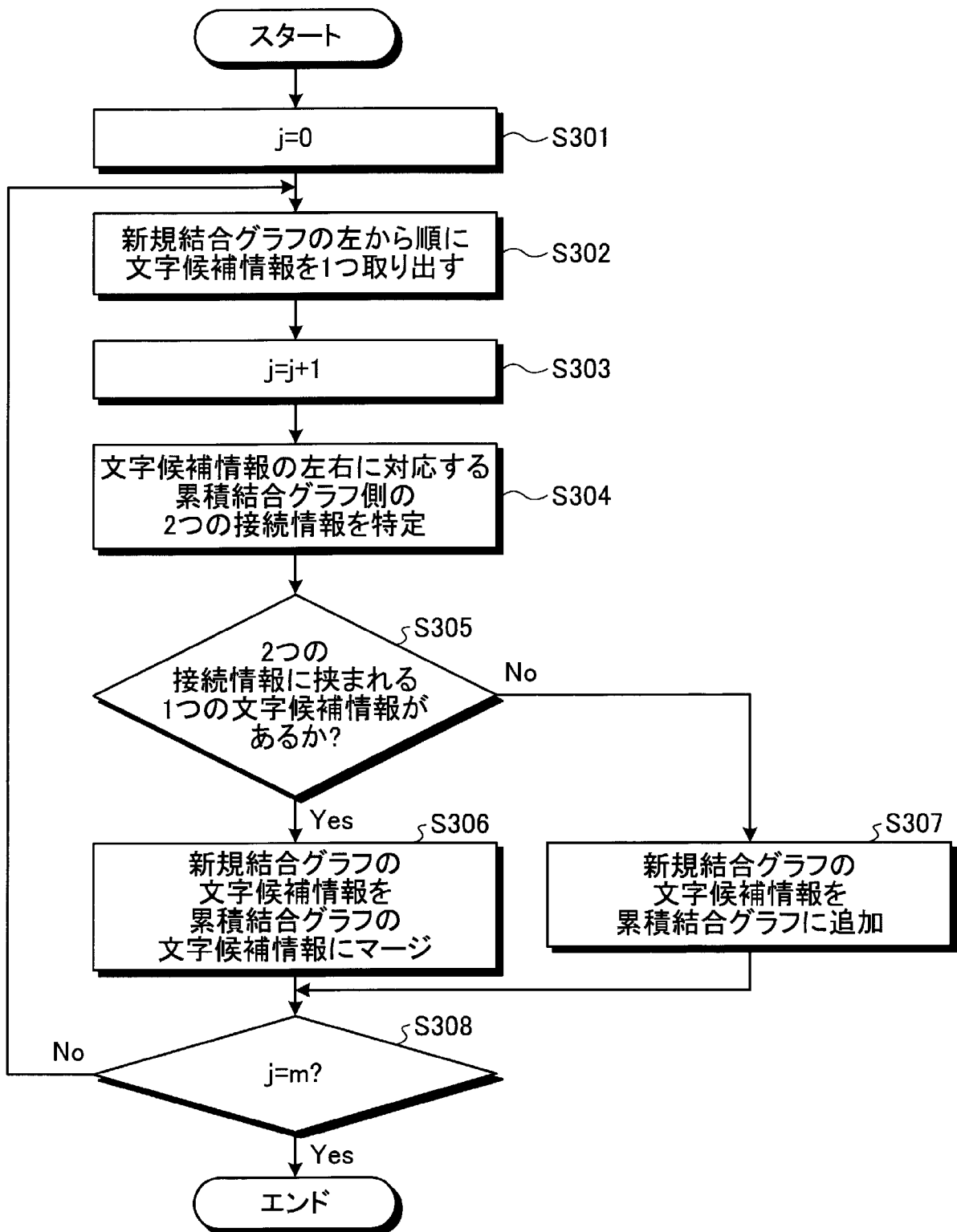
[図7]



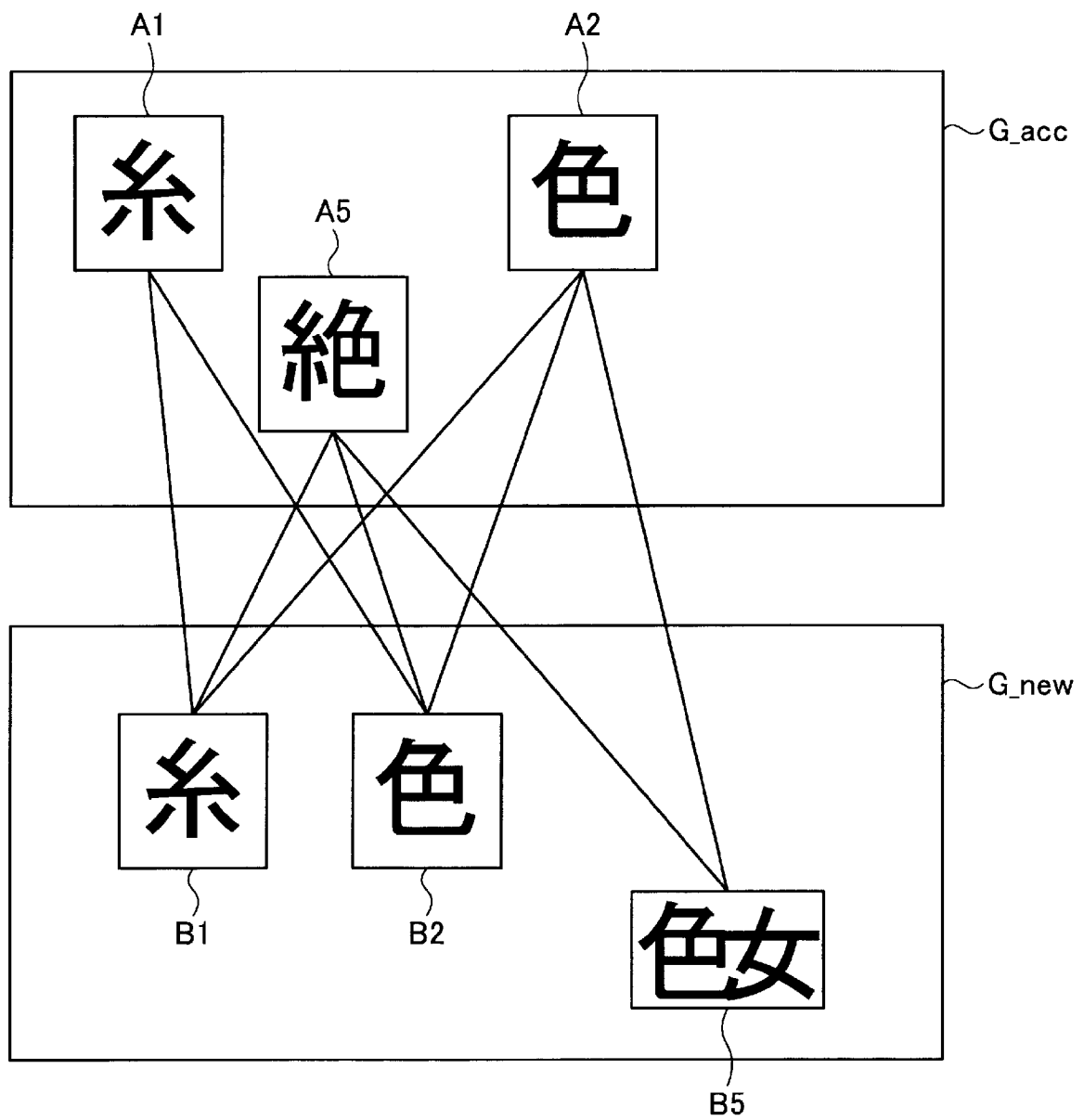
[図8]



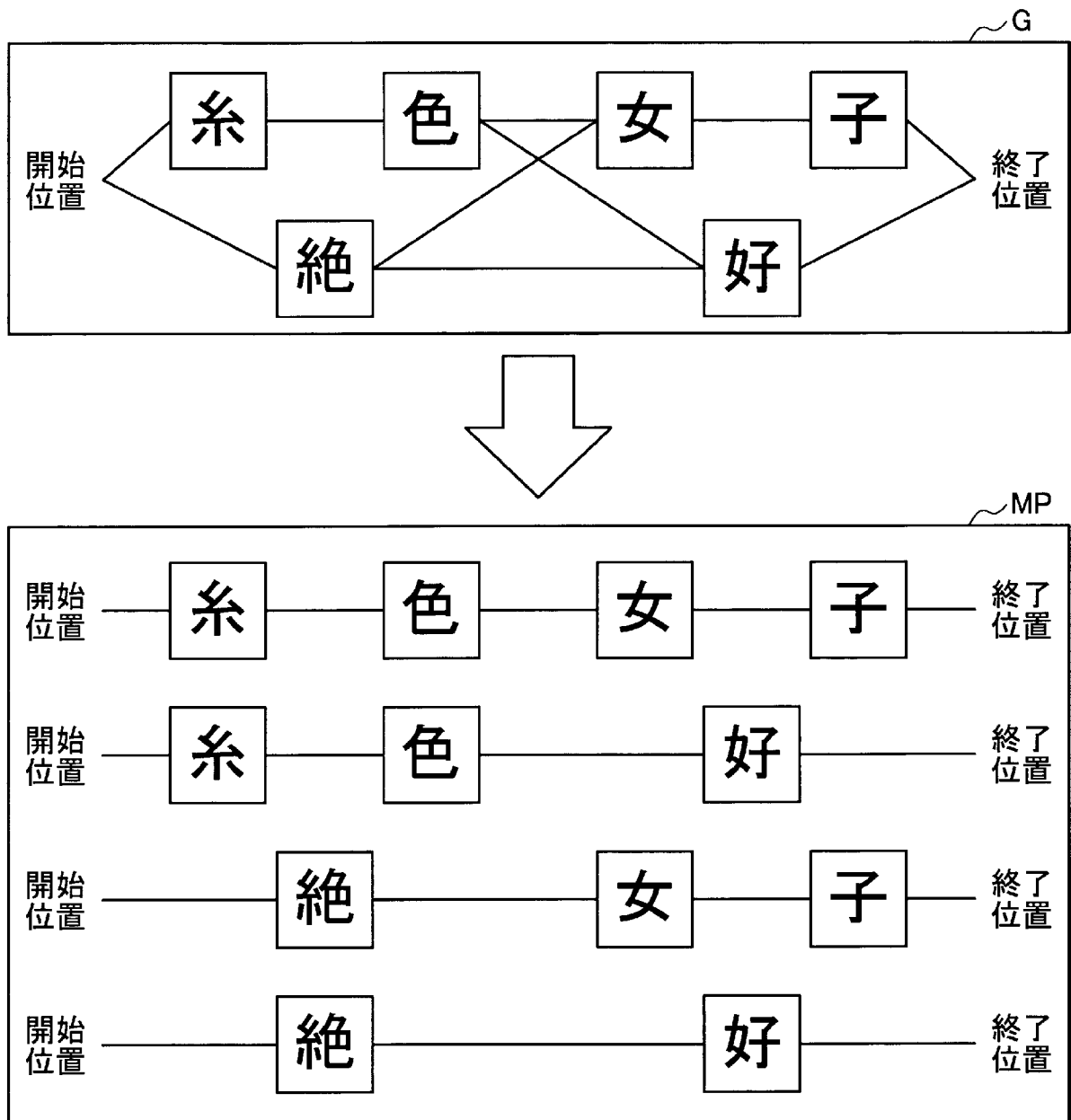
[図9]



[図10]



[図11]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/075721

A. CLASSIFICATION OF SUBJECT MATTER

G06K9/62(2006.01) i, G06K9/03(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)

G06K9/62, G06K9/03

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	1996-2016
Kokai Jitsuyo Shinan Koho	1971-2016	Toroku Jitsuyo Shinan Koho	1994-2016

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.
X Y	JP 2012-174163 A (Fujitsu Ltd.), 10 September 2012 (10.09.2012), claims 1 to 5; paragraphs [0011] to [0043], [0082] to [0153]; fig. 10 (Family: none)	1-4, 8-9 5-7
Y	Takahiko HORIUCHI et al., "Extended Relaxation Matching Method, Including Dynamic Programming Matching Method", The Transactions of the Institute of Electronics, Information and Communication Engineers, 25 April 1992 (25.04. 1992), vol.J75-D-II, no.4, pages 714 to 719, ISSN 0915-1923	5-7



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

11 November 2016 (11.11.16)

Date of mailing of the international search report

22 November 2016 (22.11.16)

Name and mailing address of the ISA/

Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/075721

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2000-207491 A (Hitachi, Ltd.), 28 July 2000 (28.07.2000), paragraphs [0015] to [0076]; fig. 7, 9 to 11 (Family: none)	6-7

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06K9/62(2006.01)i, G06K9/03(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. G06K9/62, G06K9/03

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

国際調査で利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 請求項の番号
X	JP 2012-174163 A（富士通株式会社） 2012.09.10, [請求項1] - [請求項5], 段落 [0011] - [0043], 段落 [0082] - [0153], [図10] (ファミリーなし)	1-4, 8-9
Y		5-7
Y	堀内 隆彦 外3名, DP整合法を含む拡張型弛緩整合法の提案, 電子情報通信学会論文誌 (J 7 5 - D - I I) 第4号, 1992.04.25, 第J75-D-II 巻 第4号, pp.714-719, ISSN 0915-1923	5-7

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

1 1 . 1 1 . 2 0 1 6

国際調査報告の発送日

2 2 . 1 1 . 2 0 1 6

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100-8915

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）

新井 則和

電話番号 03-3581-1101 内線 3531

5H

8937

C (続き) . 関連すると認められる文献		
引用文献の カテゴリ*	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 請求項の番号
Y	JP 2000-207491 A (株式会社日立製作所) 2000.07.28, 段落 [0015] – [0076], [図 7], [図 9] – [図 11] (ファミリーなし)	6-7