

(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2014-157409

(P2014-157409A)

(43) 公開日 平成26年8月28日(2014.8.28)

(51) Int.Cl.	F I	テーマコード (参考)
G06K 9/62 (2006.01)	G06K 9/62 620B	5B064
G06K 9/72 (2006.01)	G06K 9/72 B	
G06F 17/30 (2006.01)	G06F 17/30 419B	
	G06F 17/30 210A	

審査請求 未請求 請求項の数 9 O L (全 28 頁)

(21) 出願番号	特願2013-26808 (P2013-26808)	(71) 出願人	000005496
(22) 出願日	平成25年2月14日 (2013.2.14)		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
		(74) 代理人	100115129
			弁理士 清水 昇
		(74) 代理人	100102716
			弁理士 在原 元司
		(74) 代理人	100122275
			弁理士 竹居 信利
		(72) 発明者	木村 俊一
			神奈川県足柄上郡中井町境430 グリー
			ンテクなかい 富士ゼロックス株式会社内
		(72) 発明者	田中 瑛一
			神奈川県足柄上郡中井町境430 グリー
			ンテクなかい 富士ゼロックス株式会社内
			最終頁に続く

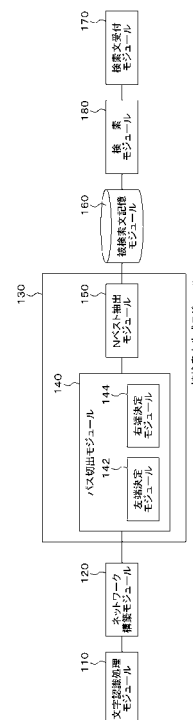
(54) 【発明の名称】 情報処理装置及び情報処理プログラム

(57) 【要約】

【課題】 確定していない複数の文字認識結果がある場合に、その文字認識結果で構成される文字列を被検索対象とするようにした情報処理装置を提供する。

【解決手段】 情報処理装置のネットワーク構築手段は、複数の文字認識結果の各々の文字をノードとし、隣接する文字画像のノード間をリンクで結ぶネットワークを構築し、第1の決定手段は、前記ネットワーク内の第1の切れ目候補を決定し、第2の決定手段は、前記ネットワーク内の前記第1の切れ目候補とは異なる第2の切れ目候補を決定し、抽出手段は、前記第1の切れ目候補と前記第2の切れ目候補の間にあるノードによって構成される文字列候補の組み合わせの中から複数の文字列候補を被検索対象として抽出する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

複数の文字認識結果の各々の文字をノードとし、隣接する文字画像のノード間をリンクで結ぶネットワークを構築するネットワーク構築手段と、

前記ネットワーク内の第 1 の切れ目候補を決定する第 1 の決定手段と、

前記ネットワーク内の前記第 1 の切れ目候補とは異なる第 2 の切れ目候補を決定する第 2 の決定手段と、

前記第 1 の切れ目候補と前記第 2 の切れ目候補の間にあるノードによって構成される文字列候補の組み合わせの中から複数の文字列候補を被検索対象として抽出する抽出手段を具備することを特徴とする情報処理装置。

10

【請求項 2】

前記第 1 の決定手段は、前記第 1 の切れ目候補を、前記ネットワーク内の先頭から最後尾まで移動させる

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】

前記第 1 の決定手段は、前記第 1 の切れ目候補を、該第 1 の切れ目候補の文字画像を含む文字列内における文字画像間の距離が閾値よりも大又は以上である箇所とする

ことを特徴とする請求項 1 又は 2 に記載の情報処理装置。

【請求項 4】

前記第 2 の決定手段は、前記第 1 の切れ目候補から予め定められたノード数離れた位置を前記第 2 の切れ目候補として決定する

ことを特徴とする請求項 1 から 3 のいずれか 1 項に記載の情報処理装置。

20

【請求項 5】

前記第 2 の決定手段は、前記第 1 の切れ目候補の文字画像から予め定められた距離だけ離れた位置を前記第 2 の切れ目候補として決定する

ことを特徴とする請求項 1 から 3 のいずれか 1 項に記載の情報処理装置。

【請求項 6】

前記第 2 の決定手段は、前記第 1 の切れ目候補の文字画像の高さ又は該第 1 の切れ目候補の文字画像を含む文字列の高さに基づいた距離だけ離れた位置を前記第 2 の切れ目候補として決定する

ことを特徴とする請求項 1 から 3 のいずれか 1 項に記載の情報処理装置。

30

【請求項 7】

前記抽出手段は、前記第 1 の切れ目候補と前記第 2 の切れ目候補の間にあるノードによって構成される文字列候補の組み合わせの評価値を算出し、該評価値に基づいて複数の文字列候補を抽出する

ことを特徴とする請求項 1 から 6 のいずれか一項に記載の情報処理装置。

【請求項 8】

検索文字列を受け付ける受付手段と、

前記抽出手段によって抽出された文字列候補を被検索文字列として、前記検索文字列を検索する検索手段

をさらに具備することを特徴とする請求項 1 から 7 のいずれか一項に記載の情報処理装置。

40

【請求項 9】

コンピュータを、

複数の文字認識結果の各々の文字をノードとし、隣接する文字画像のノード間をリンクで結ぶネットワークを構築するネットワーク構築手段と、

前記ネットワーク内の第 1 の切れ目候補を決定する第 1 の決定手段と、

前記ネットワーク内の前記第 1 の切れ目候補とは異なる第 2 の切れ目候補を決定する第 2 の決定手段と、

前記第 1 の切れ目候補と前記第 2 の切れ目候補の間にあるノードによって構成される文

50

字列候補の組み合わせの中から複数個の文字列候補を被検索対象として抽出する抽出手段として機能させるための情報処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理装置及び情報処理プログラムに関する。

【背景技術】

【0002】

特許文献1には、認識誤りによって検索できない場合を少なくすると共に、一般的なシステムでの検索が可能な、情報処理技術を提供することを課題とし、入力された画像が文字認識されると、この認識された文字データ候補からテキストデータを生成し、このとき、認識された文字データから複数のテキストデータが生成され、生成された複数のテキストデータを画像と共に電子データとして保存することが開示されている。

【0003】

特許文献2には、文字認識を用いた文書画像の検索において、誤認識があっても高精度な検索を行うことを課題とし、文字認識部は、文書画像を文字認識してテキストを生成し、再処理範囲決定部は、再処理範囲を決定し、ラティス生成部は、再処理範囲の文字認識をやり直した結果から候補文字ラティスを生成し、文字列生成部は、候補文字ラティスから文字列を生成し、テキスト追加部は、テキストに文字列を追加して、検索用テキストを生成し、インデックス生成部は、テキストを用いて文書画像にインデックスをばり、検索部は、このインデックスを用いた検索を実行することが開示されている。

【0004】

特許文献3には、従来の検索技術では、依然として無関係なヒットが発生してしまうことを課題とし、検索クエリと当該検索クエリの展開クエリとを基に、検索対象の文書データを検索して前記検索クエリ及び前記展開クエリに一致する文字列を抽出し、その抽出された文字列が、未知語領域を含むか否かを判定し、未知語領域を含まないと判定された場合に、抽出された文字列の類似度を低下させるように調整し、その調整された類似度に応じた順番で文字列を検索結果として出力することが開示されている。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2009-230671号公報

【特許文献2】特開2002-312365号公報

【特許文献3】特開2006-343870号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

本発明は、確定していない複数の文字認識結果がある場合に、その文字認識結果で構成される文字列を被検索対象とするようにした情報処理装置及び情報処理プログラムを提供することを目的としている。

【課題を解決するための手段】

【0007】

かかる目的を達成するための本発明の要旨とするところは、次の各項の発明に存する。

請求項1の発明は、複数の文字認識結果の各々の文字をノードとし、隣接する文字画像のノード間をリンクで結ぶネットワークを構築するネットワーク構築手段と、前記ネットワーク内の第1の切れ目候補を決定する第1の決定手段と、前記ネットワーク内の前記第1の切れ目候補とは異なる第2の切れ目候補を決定する第2の決定手段と、前記第1の切れ目候補と前記第2の切れ目候補の間にあるノードによって構成される文字列候補の組み合わせの中から複数個の文字列候補を被検索対象として抽出する抽出手段を具備することを特徴とする情報処理装置である。

【0008】

請求項2の発明は、前記第1の決定手段は、前記第1の切れ目候補を、前記ネットワーク内の先頭から最後尾まで移動させることを特徴とする請求項1に記載の情報処理装置である。

【0009】

請求項3の発明は、前記第1の決定手段は、前記第1の切れ目候補を、該第1の切れ目候補の文字画像を含む文字列内における文字画像間の距離が閾値よりも大又は以上である箇所とすることを特徴とする請求項1又は2に記載の情報処理装置である。

【0010】

請求項4の発明は、前記第2の決定手段は、前記第1の切れ目候補から予め定められたノード数離れた位置を前記第2の切れ目候補として決定することを特徴とする請求項1から3のいずれか1項に記載の情報処理装置である。

10

【0011】

請求項5の発明は、前記第2の決定手段は、前記第1の切れ目候補の文字画像から予め定められた距離だけ離れた位置を前記第2の切れ目候補として決定することを特徴とする請求項1から3のいずれか1項に記載の情報処理装置である。

【0012】

請求項6の発明は、前記第2の決定手段は、前記第1の切れ目候補の文字画像の高さ又は該第1の切れ目候補の文字画像を含む文字列の高さに基づいた距離だけ離れた位置を前記第2の切れ目候補として決定することを特徴とする請求項1から3のいずれか1項に記載の情報処理装置である。

20

【0013】

請求項7の発明は、前記抽出手段は、前記第1の切れ目候補と前記第2の切れ目候補の間にあるノードによって構成される文字列候補の組み合わせの評価値を算出し、該評価値に基づいて複数の文字列候補を抽出することを特徴とする請求項1から6のいずれか1項に記載の情報処理装置である。

【0014】

請求項8の発明は、検索文字列を受け付ける受付手段と、前記抽出手段によって抽出された文字列候補を被検索文字列として、前記検索文字列を検索する検索手段をさらに具備することを特徴とする請求項1から7のいずれか1項に記載の情報処理装置である。

30

【0015】

請求項9の発明は、コンピュータを、複数の文字認識結果の各々の文字をノードとし、隣接する文字画像のノード間をリンクで結ぶネットワークを構築するネットワーク構築手段と、前記ネットワーク内の第1の切れ目候補を決定する第1の決定手段と、前記ネットワーク内の前記第1の切れ目候補とは異なる第2の切れ目候補を決定する第2の決定手段と、前記第1の切れ目候補と前記第2の切れ目候補の間にあるノードによって構成される文字列候補の組み合わせの中から複数の文字列候補を被検索対象として抽出する抽出手段として機能させるための情報処理プログラムである。

【発明の効果】

【0016】

請求項1の情報処理装置によれば、確定していない複数の文字認識結果がある場合に、その文字認識結果で構成される文字列を被検索対象とすることができる。

40

【0017】

請求項2の情報処理装置によれば、被検索対象として、ネットワーク内の文字列が欠如してしまうことを防止することができる。

【0018】

請求項3の情報処理装置によれば、第1の切れ目候補を、文字列内における文字画像間の距離が閾値よりも大又は以上である箇所とすることができる。

【0019】

請求項4の情報処理装置によれば、予め定められた文字数の文字列を被検索対象とする

50

ことができる。

【0020】

請求項5の情報処理装置によれば、第1の切れ目候補の文字画像から予め定められた距離の間にある文字画像からの文字認識結果で構成される文字列を被検索対象とすることができる。

【0021】

請求項6の情報処理装置によれば、第1の切れ目候補の文字画像の高さ又は第1の切れ目候補の文字画像を含む文字列の高さに基づいた距離だけ離れた位置の間にある文字画像からの文字認識結果で構成される文字列を被検索対象とすることができる。

【0022】

請求項7の情報処理装置によれば、評価値に基づいて複数の文字列候補を抽出することができる。

【0023】

請求項8の情報処理装置によれば、確定していない複数の文字認識結果がある場合に、その文字認識結果に対して検索を行うことができる。

【0024】

請求項9の情報処理プログラムによれば、確定していない複数の文字認識結果がある場合に、その文字認識結果で構成される文字列を被検索対象とすることができる。

【図面の簡単な説明】

【0025】

【図1】本実施の形態の構成例についての概念的なモジュール構成図である。

【図2】本実施の形態による被検索文の生成処理例を示すフローチャートである。

【図3】本実施の形態による検索処理例を示すフローチャートである。

【図4】文字認識処理モジュール内の構成例についての概念的なモジュール構成図である。

。

【図5】文字認識処理モジュールによる処理例を示すフローチャートである。

【図6】記号例を示す説明図である。

【図7】記号例を示す説明図である。

【図8】記号例を示す説明図である。

【図9】記号例を示す説明図である。

【図10】記号例を示す説明図である。

【図11】文字切り出し位置が複数ある場合の処理例を示す説明図である。

【図12】記号例を示す説明図である。

【図13】文字切り出し位置が複数ある場合の処理例を示す説明図である。

【図14】文字列を構成するネットワークの例を示す説明図である。

【図15】文字列を構成するネットワーク内のノードが接続されている例を示す説明図である。

【図16】被検索文生成モジュールによる処理例を示すフローチャートである。

【図17】本実施の形態による処理例を示す説明図である。

【図18】本実施の形態による処理例を示す説明図である。

【図19】本実施の形態による処理例を示す説明図である。

【図20】本実施の形態による処理例を示す説明図である。

【図21】本実施の形態による処理例を示す説明図である。

【図22】本実施の形態による処理例を示す説明図である。

【図23】本実施の形態による処理例を示す説明図である。

【図24】本実施の形態による処理例を示す説明図である。

【図25】本実施の形態による処理例を示す説明図である。

【図26】本実施の形態による処理例を示す説明図である。

【図27】検索モジュールによる処理例を示すフローチャートである。

【図28】被検索文生成モジュールによる処理例を示すフローチャートである。

10

20

30

40

50

【図 29】本実施の形態を実現するコンピュータのハードウェア構成例を示すブロック図である。

【図 30】文字列画像の例を示す説明図である。

【図 31】文字境界候補の例を示す説明図である。

【図 32】外接矩形の例を示す説明図である。

【図 33】文字切り出し結果の例を示す説明図である。

【図 34】文字切り出し位置を示すネットワーク表現の例を示す説明図である。

【図 35】ネットワーク表現内のパターンの例を示す説明図である。

【図 36】ネットワークの例を示す説明図である。

【図 37】複数の文字候補がある場合のネットワーク例を示す説明図である。

10

【発明を実施するための形態】

【0026】

まず、本実施の形態を説明する前に、その前提又は本実施の形態を利用する情報処理装置について説明する。なお、この説明は、本実施の形態の理解を容易にすることを目的とするものである。

【0027】

例えば、図 30 に示すような文字列画像を対象として説明する。まず、この文字列画像を準文字（文字画像、文字セグメントともいわれる）に分割する。準文字とは、文字そのもの、あるいは文字の一部となる可能性があるものである。文字部分としては、例えば、漢字の偏、旁が該当する。ここでは、図 30 の例に示すような横書きの文字列画像を例にとる。横書きの画像では、垂直な線（あるいは垂直に近い線）で分割を行うことにより、準文字に分割する。例えば図 31 に示した縦線（切れ目候補 3110、切れ目候補 3120）で、文字列画像を分割して、3つの準文字「イ」、「ヒ」、及び、「学」を得ることができる。図 31 の例に示した縦線を切れ目候補と呼ぶこととする。切れ目候補 3110 が「イ」と「ヒ」を分け、切れ目候補 3120 が「ヒ」と「学」を分けている。

20

【0028】

次に、図 32 の例に示すように、各準文字に対し、その外接矩形（外接矩形 3210、外接矩形 3220、外接矩形 3230）を抽出する。

以下、既に知られている技術内容を例にして説明する。

前述の準文字を統合して、文字画像を決定する。複数の準文字を統合して1つの文字画像を形成する場合もあれば、1つの準文字が1つの文字となる場合もある。文字画像を決定するとは、文字の切り出し位置を決定することと同値であるから、以下では文字切り出し位置の決定という場合もある。

30

準文字の統合のパターンは複数存在する。複数存在するパターンの中で、最も文字画像として評価の高いものを選択することによって、最終的な文字切り出し位置を決定する。

図 32 の例に対しては、全ての文字切り出しパターンは、図 33 に示す例のようになる。つまり、図 33（a）の例では、パターン 1 として3つの文字画像（外接矩形 3210、外接矩形 3220、外接矩形 3230）、図 33（b）の例では、パターン 2 として2つの文字画像（外接矩形 3210 と 3220、外接矩形 3230）、図 33（c）の例では、パターン 3 として1つの文字画像（外接矩形 3210 と 3220 と 3230）、図 33（d）の例では、パターン 4 として2つの文字画像（外接矩形 3210、外接矩形 3220 と 3230）を示している。

40

【0029】

文字切り出し位置を示すネットワーク表現（グラフ表現ともいわれる）として、図 33 の例に示した複数の切り出しパターンを表すことができる。図 34 の例において、ネットワークは、始点ノード 3400、終点ノード 3490、中間ノード 3410（ノード 1）、中間ノード 3420（ノード 2）の4つのノードと、ノード間を接続するアークで構成されている（ノード間の接続線をアークと呼ぶこととする）。始点は、文字列画像の左端、終点は文字列画像の右端にあたる。中間ノード 3410（ノード 1）、中間ノード 3420（ノード 2）は、それぞれ、文字の切れ目候補位置（すなわち、図 31 の例に示した

50

切れ目候補 3 1 1 0、切れ目候補 3 1 2 0) を示す。中間ノード 3 4 1 0 (ノード 1) は、切れ目候補 3 1 1 0 に対応している。また、中間ノード 3 4 2 0 (ノード 2) は、切れ目候補 3 1 2 0 に対応している。

【0030】

始点から、各ノードを通して、終点に至る経路を以下、「パス」と呼ぶ。パスは、1 又は複数のアークから構成される。通常、複数のパスが存在する。図 3 3 の例に示した文字切り出しパターンは、これらの複数のパスに対応している。例えば、図 3 3 (b) の例に示したパターン 2 は、図 3 5 の太線で示したパス (文字切り出しパターン 3 4 0 4、文字切り出しパターン 3 4 2 2) に対応している。

ここで、どれか 1 つのアークには、1 つの文字画像の候補が対応している。例えば、始点ノード 3 4 0 0 と中間ノード 3 4 2 0 (ノード 2) を結ぶアークには、「化」という文字画像 (文字切り出しパターン 3 4 0 4) が対応している。1 つのアークに対応する文字に対して、その文字の評価値を決定することができる。これを「アーク評価値」と呼ぶこととする。

アーク評価値は、文字の形状情報や、文字認識における認識確度 (認識対象の準文字が文字認識結果である文字コードであることの確信度) などから算出する。アーク評価値の算出方法はさまざまある。例えば、(1) 特開平 9-1 8 5 6 8 1 号公報、(2) 特開平 8-1 6 1 4 3 2 号公報、(3) 特開平 1 0-1 5 4 2 0 7 号公報、(4) 特開昭 6 1-1 7 5 8 7 8 号公報、(5) 特開平 3-3 7 7 8 2 号公報、(6) 特開平 1 1-2 0 3 4 0 6 号公報等に記載の従来技術に示されている手法等がある。

【0031】

1 つのパスは、複数のアークから構成されている。複数のアーク評価値を用いて、そのアークから構成されるパスの評価値を計算することができる。これを「パス評価値」と呼ぶこととする。

文字切り出し位置を決定するため、複数のパスの中で、最もパス評価値の高いパスを選択する。パスが選択できれば、文字切り出し位置が確定して、文字を切り出すことができる。そして、切り出した文字 (文字画像) を文字認識した結果も確定することになる。

例えば、図 3 5 の例で、太線のパスが選択されたとする。この場合、文字切り出し位置は、始点 3 4 0 0 と、中間ノード 3 4 2 0 (ノード 2) と、終点 3 4 9 0 の 3 点となる。そして、確定した文字認識結果は、「化」 (文字切り出しパターン 3 4 0 4)、「学」 (文字切り出しパターン 3 4 2 2) となる。

【0032】

ここで、パス評価値の算出方法について説明する。基本的には、パス評価値は、アーク評価値の重み付け和が用いられる。i 番目のアークのアーク評価値を V_i とし、i 番目のアーク評価値に対する重みを w_i とし、アークの数を N 、パス評価値を P とすると、 P は次の (1) 式で与えられる。

【数 1】

$$P = \sum_{i=1}^N w_i V_i \quad (1)$$

パスの数が複数存在すると前述したが、実際の文字列には、多くの準文字が存在しているため、存在するパスの数は膨大となる。

そこで、特開平 3-2 2 5 5 7 9 号公報に記載の技術では、図 3 4 の例に示されたようなネットワーク内の複数のパスから最も評価値の高いパスを探索する方法として、ダイナミックプログラミング手法を用いることが述べられている。

ここでは、ダイナミックプログラミング手法の中で、このようなネットワークの最良パスを探索するのに適したビタビ法の説明を行う。

【0033】

図 3 6 に示す例において、始点ノード 3 6 0 0 から終点ノード 3 6 9 0 までのネットワ

ークを説明する。なお、ノード間のリンク（アークともいわれる）は、図 3 6 に示したものに限らず、さまざまな場合がある。もちろんのことながら、図 3 6 のように対称の結線でなくてもかまわない。

このネットワークにおいて、途中に複数のノード（中間ノード 3 6 1 1、中間ノード 3 6 1 2、中間ノード 3 6 1 3 等）を介して始点ノードから終点ノードに達するとする。途中のノードを中間ノードと呼ぶこととする。

各ノードとノードの間にはリンクが張られている。このリンクにはそのリンク固有の評価値（リンク値）が割り当てられている。始点ノード 3 6 0 0 から終点ノード 3 6 9 0 に向かうパスは複数存在する。パスは、複数のリンクから成り立つことになる。パスが含む複数のリンクのリンク値の和が、パスの評価値となる。

例えば、リンク値は、ノード間の距離であるとする。この場合、パス評価値が最小のパスは、始点ノードから終点ノードに向かうパスの中で、最小距離のパスということになる。あるいは、パス評価値が最大のパスを求める問題とすることも可能である。

【0 0 3 4】

ここで、ビタビ法は、あるノードにある方向から入力されるリンクを 1 に限定することによって、最適ではないパスを削減する。これにより、演算処理量と所要メモリ量を削減する手法である。

例えば、今、ノード x（中間ノード 3 6 2 1）に左から入力されるリンクがすでに、1 に限定されているとする。同様に、ノード y（中間ノード 3 6 2 2）、ノード z（中間ノード 3 6 2 3）に関しても 1 に限定されているとする。このとき、ノード X（中間ノード 3 6 3 1）に左から入力されるリンクを限定する。ノード X（中間ノード 3 6 3 1）には、ノード x（中間ノード 3 6 2 1）、ノード y（中間ノード 3 6 2 2）、ノード z（中間ノード 3 6 2 3）の 3 つのノードからリンクが張られている。このとき、ノード X（中間ノード 3 6 3 1）を通るパスとして、最適な可能性があるのは、ノード x（中間ノード 3 6 2 1）、ノード y（中間ノード 3 6 2 2）、ノード z（中間ノード 3 6 2 3）からノード X（中間ノード 3 6 3 1）に向かうリンクのいずれかである。この 3 つのうちで最適なものだけを残し、残りの 2 つを削除する。このようにして、ノード X（中間ノード 3 6 3 1）に左から入力されるパス（リンク）を 1 に限定する。ノード Y（中間ノード 3 6 3 2）、ノード Z（中間ノード 3 6 3 3）に関しても同様に左から入力されるパスを 1 に限定する。

このような手順を左のノード A（中間ノード 3 6 1 1）、ノード B（中間ノード 3 6 1 2）、ノード C（中間ノード 3 6 1 3）から順に右の方向に行く。最終的にノード P（中間ノード 3 6 8 1）、ノード Q（中間ノード 3 6 8 2）、ノード R（中間ノード 3 6 8 3）に入る 3 つのパスに限定する。この 3 つのパスの中で最適なものを選択すればよい。

このようなビタビ法を用いた最適パス選定方法を、図 3 4 に例示のネットワークにも同様に適用し得る。文字切り出し位置をノードとする。また、アーク評価値を前述のリンク値とすればよい。

【0 0 3 5】

従来技術が扱っていた図 3 4 に例示するようなネットワークは、1 つのアークに対し、1 つの文字候補が存在しているものである。しかし、実際には 1 つのアークに対して、複数の文字候補が存在している。つまり、1 つの準文字に対して文字認識結果が複数ある場合である。例えば、図 3 7 のように、複数の文字候補を割り当てることになる。図 3 7 は、複数の文字候補がある場合のネットワーク例を示す説明図である。図 3 7 の例では、「化学」（文字画像 3 7 0 2）という画像を 1 文字として認識したときの文字認識結果が、文字候補 3 7 0 2 A の「梢」、文字候補 3 7 0 2 B の「際」、文字候補 3 7 0 2 C の「傾」の 3 種類ある。「化」（文字画像 3 7 0 4）という画像を 1 文字として認識したときの文字認識結果が、文字候補 3 7 0 4 A の「化」、文字候補 3 7 0 4 B の「比」、文字候補 3 7 0 4 C の「任」の 3 種類ある。同様に、他の文字画像に対しても文字認識結果としての文字候補が複数ある。なお、図 3 7 の例では、各文字画像に対して 3 種類の文字候補があるが、異なる数の文字候補を割り当ててもよい。例えば、認識確度が予め定められた値

10

20

30

40

50

以上である文字を文字候補とした場合は、各文字画像に対して文字候補の数が異なる場合がある。

【0036】

以下、図面に基づき本発明を実現するにあたっての好適な一実施の形態の例を説明する。

図1は、本実施の形態の構成例についての概念的なモジュール構成図を示している。

なお、モジュールとは、一般的に論理的に分離可能なソフトウェア（コンピュータ・プログラム）、ハードウェア等の部品を指す。したがって、本実施の形態におけるモジュールはコンピュータ・プログラムにおけるモジュールのことだけでなく、ハードウェア構成におけるモジュールも指す。それゆえ、本実施の形態は、それらのモジュールとして機能させるためのコンピュータ・プログラム（コンピュータにそれぞれの手順を実行させるためのプログラム、コンピュータをそれぞれの手段として機能させるためのプログラム、コンピュータにそれぞれの機能を実現させるためのプログラム）、システム及び方法の説明をも兼ねている。ただし、説明の都合上、「記憶する」、「記憶させる」、これらと同等の文言を用いるが、これらの文言は、実施の形態がコンピュータ・プログラムの場合は、記憶装置に記憶させる、又は記憶装置に記憶させるように制御するの意である。また、モジュールは機能に一对一に対応していてもよいが、実装においては、1モジュールを1プログラムで構成してもよいし、複数モジュールを1プログラムで構成してもよく、逆に1モジュールを複数プログラムで構成してもよい。また、複数モジュールは1コンピュータによって実行されてもよいし、分散又は並列環境におけるコンピュータによって1モジュールが複数コンピュータで実行されてもよい。なお、1つのモジュールに他のモジュールが含まれていてもよい。また、以下、「接続」とは物理的な接続の他、論理的な接続（データの授受、指示、データ間の参照関係等）の場合にも用いる。「予め定められた」とは、対象としている処理の前に定まっていることをいい、本実施の形態による処理が始まる前はもちろんのこと、本実施の形態による処理が始まった後であっても、対象としている処理の前であれば、そのときの状況・状態に応じて、又はそれまでの状況・状態に応じて定まることの意を含めて用いる。「予め定められた値」が複数ある場合は、それぞれ異なった値であってもよいし、2以上の値（もちろんのことながら、全ての値も含む）が同じであってもよい。また、「Aである場合、Bをする」という意味を有する記載は、「Aであるか否かを判断し、Aであると判断した場合はBをする」の意味で用いる。ただし、Aであるか否かの判断が不要である場合を除く。

また、システム又は装置とは、複数のコンピュータ、ハードウェア、装置等がネットワーク（一対一対応の通信接続を含む）等の通信手段で接続されて構成されるほか、1つのコンピュータ、ハードウェア、装置等によって実現される場合も含まれる。「装置」と「システム」とは、互いに同義の用語として用いる。もちろんのことながら、「システム」には、人為的な取り決めである社会的な「仕組み」（社会システム）にすぎないものは含まない。

また、各モジュールによる処理毎に又はモジュール内で複数の処理を行う場合はその処理毎に、対象となる情報を記憶装置から読み込み、その処理を行った後に、処理結果を記憶装置に書き出すものである。したがって、処理前の記憶装置からの読み込み、処理後の記憶装置への書き出しについては、説明を省略する場合がある。なお、ここでの記憶装置としては、ハードディスク、RAM（Random Access Memory）、外部記憶媒体、通信回線を介した記憶装置、CPU（Central Processing Unit）内のレジスタ等を含んでいてもよい。

【0037】

本実施の形態である情報処理装置は、画像から文字を認識し、その文字認識結果から検索の対象（被検索対象、以下、被検索文ともいう）を生成し、検索を行うものであって、図1の例に示すように、文字認識処理モジュール110、ネットワーク構築モジュール120、被検索文生成モジュール130、被検索文記憶モジュール160、検索文受付モジュール170、検索モジュール180を有している。被検索文生成モジュール130は、

パス切出モジュール140、Nベスト抽出モジュール150を有している。パス切出モジュール140は、左端決定モジュール142、右端決定モジュール144を有している。この情報処理装置を実現するのに、例えば、文字認識処理モジュール110、ネットワーク構築モジュール120、被検索文生成モジュール130を有している情報処理装置Aと、被検索文記憶モジュール160を有している情報処理装置Bと、検索文受付モジュール170、検索モジュール180を有している情報処理装置Cが通信回線を介して接続されているシステム構成としてもよいし、情報処理装置Aによる処理結果を情報処理装置Bが記憶し、その後に、情報処理装置Aと情報処理装置Bを切り離し、情報処理装置Bと情報処理装置Cを、通信回線を介して接続してもよい。

【0038】

文字認識処理モジュール110は、ネットワーク構築モジュール120と接続されている。文字認識処理モジュール110は、画像から文字を認識し、その文字認識結果をネットワーク構築モジュール120に渡す。文字認識結果には、1つの準文字に対しての複数の文字コードを少なくとも含み、その他に、その文字コードの認識確度、その準文字の位置（例えば、画像内のXY座標系における座標）、大きさ（例えば、準文字を囲む外接矩形の横幅、縦長）等を含んでいてもよい。文字認識処理モジュール110の処理の詳細については、図4の例を用いて後述する。

ネットワーク構築モジュール120は、文字認識処理モジュール110、パス切出モジュール140と接続されている。ネットワーク構築モジュール120は、文字認識処理モジュール110から文字認識結果を受け取り、その複数の文字認識結果の各々の文字をノードとし、隣接する準文字のノード間をリンクで結ぶネットワークを構築する。具体的には、前述したネットワークを生成する。

【0039】

パス切出モジュール140は、ネットワーク構築モジュール120、Nベスト抽出モジュール150と接続されている。左端決定モジュール142は、ネットワーク内の第1の切れ目候補を決定する。右端決定モジュール144は、ネットワーク内の第1の切れ目候補とは異なる第2の切れ目候補を決定する。左端決定モジュール142と右端決定モジュール144によって、第1の切れ目候補から第2の切れ目候補までの間にある文字列を特定することになる。

左端決定モジュール142は、第1の切れ目候補を、ネットワーク内の先頭から最後尾まで移動させるようにしてもよい。例えば、ネットワーク内の先頭ノードから最後尾ノードまで欠けることなく、第1の切れ目候補を選択した場合は、ネットワーク内で表される全ての文字列候補を抽出することができるようになる。

左端決定モジュール142は、第1の切れ目候補を、その第1の切れ目候補の文字画像を含む文字列内における文字画像間の距離が閾値よりも大又は以上である箇所とするようにしてもよい。対象とする画像が、英語のように分かち書きされているような場合（単語と単語の間に空白が存在する場合）は、単語に相当する文字列候補とするためである。

【0040】

右端決定モジュール144は、第1の切れ目候補から予め定められたノード数離れた位置を第2の切れ目候補として決定するようにしてもよい。ここで「予め定められたノード数」とは、予め定められた準文字数と同義である。つまり、第1の切れ目候補から数えて予め定められた準文字数（例えば、3文字分等）の文字列候補ということになる。

右端決定モジュール144は、第1の切れ目候補の文字画像から予め定められた距離だけ離れた位置を第2の切れ目候補として決定するようにしてもよい。ここで「距離」は、例えば、画素数、mm（ミリメートル）等で規定される。第1の切れ目候補からの距離は、前述の文字認識結果内の位置、大きさを用いて算出すればよい。予め定められた距離としては、例えば、予め知られている単語の長さの距離としてもよい。

右端決定モジュール144は、第1の切れ目候補の文字画像の高さ又はその第1の切れ目候補の文字画像を含む文字列の高さに基づいた距離だけ離れた位置を第2の切れ目候補として決定するようにしてもよい。「高さに基づいた距離」としては、高さに予め定めら

10

20

30

40

50

れた値を乗算等して算出した距離がある。つまり、文字の大きさ（文字の高さ）に応じて文字列候補の長さを決定しようとするものである。

【0041】

Nベスト抽出モジュール150は、パス切出モジュール140、被検索文記憶モジュール160と接続されている。Nベスト抽出モジュール150は、第1の切れ目候補と第2の切れ目候補の間にあるノードによって構成される文字列候補の組み合わせの中から複数の文字列候補を被検索対象として抽出する。そして、抽出した文字列候補を被検索対象として被検索文記憶モジュール160に格納する。なお、以下、複数の文字列候補をNベストともいう。

また、Nベスト抽出モジュール150は、第1の切れ目候補と第2の切れ目候補の間にあるノードによって構成される文字列候補の組み合わせの評価値を算出し、その評価値に基づいて複数の文字列候補を抽出するようにしてもよい。

被検索文記憶モジュール160は、Nベスト抽出モジュール150、検索モジュール180と接続されている。Nベスト抽出モジュール150によって抽出された文字列候補を被検索対象として記憶している。

【0042】

検索文受付モジュール170は、検索モジュール180と接続されている。検索文受付モジュール170は、検索文字列を受け付ける。

検索モジュール180は、被検索文記憶モジュール160、検索文受付モジュール170と接続されている。検索モジュール180は、被検索文記憶モジュール160に格納された文字列候補（つまり、Nベスト抽出モジュール150によって抽出された文字列候補）を被検索文字列として、検索文字列を検索する。

検索文受付モジュール170、検索モジュール180による処理内容については、図26～図28を用いて後述する。

【0043】

図2は、本実施の形態による被検索文の生成処理例を示すフローチャートである。

ステップS202では、文字認識処理モジュール110が、画像を受け付ける。

ステップS204では、文字認識処理モジュール110が、文字認識を行う。

ステップS206では、ネットワーク構築モジュール120が、文字認識結果を用いてネットワークを構築する。

ステップS208では、被検索文生成モジュール130が、被検索文を生成し、被検索文記憶モジュール160に格納する。

【0044】

図3は、本実施の形態による検索処理例を示すフローチャートである。確定していない複数の文字認識結果がある場合に、その文字認識結果に対して検索を行うものである。確定していない文字認識結果とは、人間による確認作業が完了していない文字認識結果であって、例えば、文字切り出しの誤りがあって1つの準文字が1つの文字に対応していなかったり、1つの準文字に対して複数の文字認識結果があるような状態をいう。

ステップS302では、検索文受付モジュール170が、検索文を受け付ける。

ステップS304では、検索モジュール180が、被検索文記憶モジュール160に記憶されている被検索文から検索文を検索する。

ステップS306では、検索モジュール180が、検索結果を出力する。

【0045】

図4は、文字認識処理モジュール110内の構成例についての概念的なモジュール構成図である。

文字認識処理モジュール110は、画像受付モジュール410、文字列抽出モジュール420、切出位置抽出モジュール430、文字候補抽出モジュール440を有している。

【0046】

画像受付モジュール410は、文字列抽出モジュール420と接続されている。画像受付モジュール410は、画像を受け付けて、その画像を文字列抽出モジュール420へ渡

10

20

30

40

50

す。画像を受け付けるとは、例えば、スキャナ、カメラ等で画像を読み込むこと、ファックス等で通信回線を介して外部機器から画像を受信すること、ハードディスク（コンピュータに内蔵されているものの他に、通信回線を介して接続されているもの等を含む）等に記憶されている画像を読み出すこと等が含まれる。画像は、2値画像、多値画像（カラー画像を含む）であってもよい。受け付ける画像は、1枚であってもよいし、複数枚であってもよい。また、画像の内容として、文字列が含まれているのであれば、ビジネスに用いられる文書、広告宣伝用のパンフレット等であってもよい。

【0047】

文字列抽出モジュール420は、画像受付モジュール410、切出位置抽出モジュール430と接続されている。文字列抽出モジュール420は、画像受付モジュール410が受け付けた画像から文字列を抽出する。

文字列抽出モジュール420が対象とする画像は、横書きあるいは縦書きの、1列のみの文字列画像である。なお、ここで、列とは、横書きの場合は横に並ぶ列であり、縦書きの場合は縦に並ぶ列である。

したがって、文字列抽出モジュール420は、画像受付モジュール410が受け付けた画像が1列のみの文字列画像であれば、そのまま用いればよい。画像受付モジュール410が受け付けた画像は、複数の文字列が存在するものである場合がある。このような複数文字列を単一の文字列になるように分離する手法としては、従来よりさまざまなものが提案されているため、それらを用いればよい。単一の文字列となるように分離する例としてもさまざまな方式があるため、そのうちのいずれかを用いればよい。例えば、（1）特開平4-311283号公報、（2）特開平3-233789号公報、（3）特開平5-73718号公報、（4）特開2000-90194号公報、等を用いればよい。これら以外の方法であってもよい。

【0048】

切出位置抽出モジュール430は、文字列抽出モジュール420、文字候補抽出モジュール440と接続されている。切出位置抽出モジュール430は、文字列抽出モジュール420が抽出した文字列画像から文字画像を切り出す位置を抽出する。つまり、文字列画像を複数の準文字に分割する。この文字切り出し位置の抽出方法としてもさまざまな方式があるため、そのうちのどれかを用いればよい。例えば、（1）特開平5-114047号公報、（2）特開平4-100189号公報、（3）特開平4-92992号公報、（4）特開平4-68481号公報、（5）特開平9-54814号公報、（6）特開平9-185681号公報の0021段落に記載の文字の境界候補抽出方式、（7）特開平5-128308号公報の0005段落に記載の文字切り出し位置決定方法等を用いればよい。これら以外の方法であってもよい。なお、ここでの文字画像とは、いわば文字候補画像であり、必ずしもその文字画像が1つの文字を表す画像でなくてもよい。

【0049】

また、切出位置抽出モジュール430は、複数の切り出し位置を抽出してもよい。複数の切り出し位置を抽出することによって、1つの文字列画像に対して、文字切出位置群が複数ある。文字切出位置群とは、1つの文字列画像に対して、1つ以上の文字切り出し位置があることをいう。例えば、2つの文字切り出し位置があれば、4つの文字画像に分けられることになる。また、文字切出位置群が複数あるとは、1つの文字列画像に対して、文字切出位置で分けられた文字画像列が複数あることをいう。例えば、2つの文字切り出し位置で3つの文字画像からなる文字画像列と、4つの文字切り出し位置で5つの文字画像からなる文字画像列となる。具体例で説明すると、「化学」という文字列画像に対して、「イ」「ヒ」「学」からなる文字画像列と「化」「学」からなる文字画像列を生成することである。

【0050】

文字候補抽出モジュール440は、切出位置抽出モジュール430と接続されている。文字候補抽出モジュール440は、切出位置抽出モジュール430によって抽出された位置によって分けられた文字画像に対して文字認識を行った結果である複数の文字候補を抽

出し、ネットワーク構築モジュール120に渡す。この抽出処理を行うために、文字認識処理を行う。そのために、文字候補抽出モジュール440内に文字認識モジュールを含む。文字認識処理による認識結果は、前述のように、1つの文字画像に対して複数の文字候補が対応する。つまり、その文字画像を文字認識して、第1位の認識確度を有している文字候補、第2位の認識確度を有している文字候補等のように、複数の文字候補を文字認識結果とする。そして、文字認識結果として、文字候補の他に、その文字候補の認識確度を有していてもよい。また、抽出するのに、1つの文字画像に対して予め定められた数の文字候補を抽出してもよいし、予め定められた値以上の認識確度を有している文字候補を抽出してもよい。なお、認識確度は、文字認識処理において、その認識結果としての信頼性を表す値であってもよいし、文字画像の外接矩形のサイズ、縦横比等によって定められる文字らしさを表す値であってもよい。

10

【0051】

図5は、文字認識処理モジュール110による処理例を示すフローチャートである。

ステップS502では、画像受付モジュール410が、対象となる画像を受け付ける。

ステップS504では、文字列抽出モジュール420が、画像から文字列画像を抽出する。

ステップS506では、切出位置抽出モジュール430が、文字列画像を対象として切り出し位置を抽出する。

ステップS508では、文字候補抽出モジュール440が、切り出された文字画像を文字認識する。

20

ステップS510では、文字候補抽出モジュール440が、複数の文字認識結果を文字画像の文字候補として抽出する。

【0052】

次に、図6から図15を用いて、ネットワーク構築モジュール120による処理を説明する。

本実施の形態は、パス評価値の高いパスを出力することによって、複数の文字列候補を抽出するものである。また、パスの探索にダイナミックプログラミングの手法を用いてもよい。

本実施の形態のネットワークにおいては、始点ノード、終点ノード、複数の中間ノードがある。また、各ノード間のリンクには、リンク値を与える。始点ノードから1あるいは複数の中間ノードを介して、終点ノードに至るパスは、介するノードに依存したリンクを通ることになる。始点ノードから終点ノードに至るパスのパス評価値は、そのパスが通ったリンクのリンク値の重み付け和として表すことができる。

30

本実施の形態のネットワーク構築モジュール120は、1つの文字列画像に対して、複数の文字認識結果が存在しているときに、前述のノード、リンク、パスの構成（ネットワーク構造）を生成するものである。ネットワーク構造が与えられれば、ネットワーク構築モジュール120によってビタビ法などの手法を用いて、最適パスを探索することも可能である。

【0053】

< A1. 文字切り出し位置が固定の場合 >

40

まず、切出位置抽出モジュール430の抽出する文字切り出し位置が固定（一種類）の場合について説明する。

図6は、記号例を示す説明図である。記号の種類として、長方形610、横棒である接続線620、622、624、627、628、円弧630、丸である文字候補642、644、646がある。

図6の例において、長方形610A、610B、610C、610D（図7に例示する長方形610）は、それぞれ準文字を表す。

また、横棒である接続線620、622、624、626、628は、文字切り出し位置を示す（図8に例示する接続線620、接続線622）。準文字は文字切り出し位置を介して、隣接する準文字につながっている。

50

さらに、丸で表されている文字候補 6 4 2 A、6 4 4 A 等は、1 つの準文字を 1 文字として認識したときの、複数の文字候補を示す。円弧 6 3 0 A、6 3 0 B、6 3 0 C、6 3 0 D は、下の 1 つの準文字だけを対象に文字認識を行っていることを示している。

【0054】

図 9 の例では、文字候補 6 4 2、6 4 4、6 4 6 は、その下の長方形 6 1 0 が示す 1 文字の準文字を認識したときの複数の文字候補であることを示している。また、円弧 6 3 0 は、その下の長方形 6 1 0 1 つだけを対象に文字認識を行うことを示している。

本実施の形態では、準文字の複数の文字候補をノードとして捉える。さらに、隣接する準文字の文字候補と、リンクを接続する。図 10 の例にリンクを太線で記入して示す。

ここでリンク値生成処理が生成するリンク値としては、リンク左右のノードの相互作用を示すものを使ってもよい。具体的には、リンクの左の文字候補とリンクの右の文字候補が連続して日本語の文章中に出現する確率（バイグラム）を用いる。

このようにノードとリンクを構成することによって、全てのネットワーク構造が規定できる。ネットワーク構造が規定できれば、ビタビ法等により、最適パスを選択できる。

【0055】

< A 2. ノード内情報も用いる場合 >

前述では、リンク値として、ノード間の相互作用を示すもの（文章中に出現する確率）だけを用いたが、さらにノード単独の評価値を用いるようにしてもよい。ここでは、ビタビ法を用いて最適パスを探索するものとする。左から順にノードの左から入るリンクをノード毎に 1 つだけに限定していく処理を行う。

今、図 10 の例で文字候補 6 4 2 B、6 4 4 B、6 4 6 B のリンクを限定する段階であるとする。

ここで、文字候補 6 4 2 B、6 4 4 B、6 4 6 B と、左側にある文字候補 6 4 2 A、6 4 4 A、6 4 6 A 間のリンク値を生成する。リンク値としては、各ノード間の相互作用を表すバイグラムなどの値と、各ノード内部の値の両方を用いる。ノード内部の値とは、例えば、文字候補 6 4 2 B の文字の認識確度などがある。

【0056】

ここで、リンクは、文字候補 6 4 2 B、6 4 4 B、6 4 6 B と文字候補 6 4 2 A、6 4 4 A、6 4 6 A 間にあるため、リンク値としては、文字候補 6 4 2 B、6 4 4 B、6 4 6 B と文字候補 6 4 2 A、6 4 4 A、6 4 6 A 間の評価値を算出するのが簡単である。しかし、ノード内部の値は、文字候補 6 4 2 B、6 4 4 B、6 4 6 B と文字候補 6 4 2 A、6 4 4 A、6 4 6 A 間にあるのではなく、この場合、文字候補 6 4 2 B、6 4 4 B、6 4 6 B に存在する。

つまり、ノード間情報はリンクの内部に存在していて、ノード内情報はリンクの端点に存在する。このような発生位置、又は概念が異なる値を一度に扱う。

【0057】

従来では、図 3 4 に示す始点ノード 3 4 0 0、中間ノード 3 4 1 0（ノード 1）、中間ノード 3 4 2 0（ノード 2）、終点ノード 3 4 9 0 など（つまり、文字切り出し位置）をノードとして、ノード間のアーク評価値を算出している。本実施の形態のように、複数の文字コードをノードとして、ノード間のリンク値を算出するわけではない。そのため、従来技術をそのまま用いることはできない。

本実施の形態では、リンクの評価値として、リンクの内部に存在する値（例えば、バイグラムの値）と、リンクの一方の端点のみに存在する値（例えば、文字候補 6 4 2 B の文字認識確度）を用いる。他方の端点に存在する値（例えば、文字候補 6 4 2 A の文字認識確度）は用いない。このようにすることで、リンクの内部の値と、リンクの端点の値をともに用いる評価が可能となる。

最終的には、文字列の評価値（パス評価値）として、（1）式で、全てのリンクの評価値を加算することになる。そのため、リンクの評価値の中に、リンクの内部の評価値と、リンクの一方の端点の評価値が含まれていれば、パス評価値の中に全てのリンク内部の評価値とリンク端点の評価値が 1 つずつ含まれることになる。

10

20

30

40

50

【0058】

リンク値生成処理では、前述のリンク内部の値と、リンク端点の値のように、複数の値（バイグラムや認識確度）などを特徴量として、リンク値を算出すればよい。このように複数の値から、1つのリンク値を算出する方法としては、（１）特開平９－１８５６８１号公報、（２）特開昭６１－１７５８７８号公報、（３）特開平３－３７７８２号公報、（４）特開平１１－２０３４０６号公報等に記載の技術を用いればよい。これ以外の方法であってもよい。

なお、複数の値を特徴量ベクトルとして把握し、リンク値は、特徴量ベクトルを対象として、リンク評価値（スカラー値）を出力する関数で実現できる。

【0059】

< A 3. リンク情報として２以上のノードを用いる場合 >

前述では、リンク左右のノードの相互情報としてバイグラムを用いていた。この場合、リンク情報として２つのノード間の関係情報を用いていることになる。

ビタビ法を用いる場合、例えば、ノードである文字候補６４２Ａ、６４４Ａ、６４６Ａの左側のリンク数はすでに１個に限定されていることになる。この場合には、２以上のノードの情報を用いてリンク情報を構築することが可能となる。

例えば、２つの連続する文字の生起確率であるバイグラムではなく、３つの連続する文字の生起確率であるトライグラムを用いることも可能となる。

今、リンク値生成処理では、ノードである文字候補６４２Ｂ、６４４Ｂ、６４６Ｂの左側のリンク値を生成しようとする。

例えば、文字候補６４２Ａ－文字候補６４２Ｂ間のリンク値を算出する。バイグラムであれば、文字候補６４２Ａと文字候補６４２Ｂが連続する生起確率を求めればよい。ここで、トライグラムを求める場合を説明する。文字候補６４２Ａの左側のリンクは１つに限定されているため、実は、文字候補６４２Ａの左の文字も確定していることになる。この文字を保持するノードをGとする。トライグラムとしては、ノードG－ノード（文字候補６４２Ａ）－ノード（文字候補６４２Ｂ）の３つの文字に関する生起確率を求めればよい。

以上のように求めたトライグラムをノード（文字候補６４２Ａ）－ノード（文字候補６４２Ｂ）間のリンク値として生成すればよい。同様に、N g r a mであっても、求めることが可能となる。

【0060】

< A 4. 文字切り出し位置が確定していない場合 >

文字切り出し位置が確定していない場合（つまり、切出位置抽出モジュール４３０が複数の文字切り出し位置を抽出した場合であり、具体的には、「化」のように、「イ」と「ヒ」、あるいは、「化」のどちらになるか分からない場合）、文字候補の選択と、文字切り出し位置の選択を行うようにしてもよい。文字切り出し位置が確定していない場合、文字候補の選択は、文字切り出し位置の選択となる。

図１１は、文字切り出し位置が複数ある場合の処理例を示す説明図である。ここでは、円弧の記号の意味が追加されている。円弧が下にある複数の準文字（長方形）を指し示す場合、その円弧はその複数の準文字を統合した画像を１文字とみなして認識することを示す。円弧１１１０Ａは、長方形６１０Ａと長方形６１０Ｂを統合した画像を１文字とみなして文字認識結果として文字候補１１２２Ａ、１１２４Ａ、１１２６Ａを有している。また、円弧１１１０Ｃは、長方形６１０Ａ、６１０Ｂ、６１０Ｃ、６１０Ｄを統合した画像を１文字とみなして文字認識結果として文字候補１１２２Ｃ、１１２４Ｃ、１１２６Ｃを有している。

図１２の例に示すように、円弧６３０Ａと円弧６３０Ｂの下に２つの準文字（長方形６１０Ａ、長方形６１０Ｂ）「イ」と「ヒ」がある場合に、その２つを含む円弧１１１０の上の文字候補（文字候補１１２２、１１２４、１１２６）は、「イ」と「ヒ」を統合した１つの準文字「化」を認識したときの複数の文字候補に相当する。

【0061】

文字切り出し位置が確定していない場合のリンクの接続は、図13の例に示すようになる。図13は、文字切り出し位置が複数ある場合の処理例を示す説明図である。

ここでは、文字切り出し位置に注目する。今、図13の矢印で示す文字切り出し位置に関連するノードのリンクを対象とする。この文字切り出し位置でリンクされるノードには、

(1) 左側ノード：矢印の文字切り出し位置に円弧の右側が存在するノード（斜線でハッチングしたノード、文字候補1342A、文字候補1344A、文字候補1362A、文字候補1364A、文字候補1372A、文字候補1374A等）と、

(2) 右側ノード：矢印の文字切り出し位置に円弧の左側が存在するノード（内部が白のノード、文字候補1342B、文字候補1344B、文字候補1362B、文字候補1364B、文字候補1372B、文字候補1374B等）

の2種類がある。このとき、左側ノードと、右側ノード間にリンクを形成することによって、グラフ構造を構築することができる。

例えば、全ての左側ノードが、全ての右側ノードに直接接続できるようにリンクを形成すればよい。さらに、全ての文字切り出し位置において、前述のように、左側ノードと右側ノードのリンクを形成し、さらに、左側が文字列の端点の場合には始点ノードに接続し、また、右側が文字列の端点の場合には終端ノードに接続すれば、全てのグラフ構造を構築することができる。

【0062】

この場合においても、リンク値は、リンク左右のノードの相互作用を示すものを使って

もよいし、ノード内部の評価値を用いてもよい。
特に、この場合には、文字切り出し位置が確定していないため、ノード内部の評価値として、文字の形状情報を用いることができる。文字形状情報の例として、文字の縦横比や、文字左右の空白量、等を用いることができる。

【0063】

このようにすることによって、文字列の左端から、文字列の右端に至る全てのパスを書き出すことができる。各パスは、認識出力文字列に該当する。つまり、文字列を構成するネットワークにおいて、左端ノードから、右端ノードへ至るパスを決定することによって、文字列を決定できることになる。

【0064】

以上をまとめると、文字列を構成するネットワークは、例えば、図14のようになる。図14の例において、○は、ノードを示す。○と○を接続する実線は、ノード間の接続を示すアークあるいはリンクである。ノードAは、文字列の左端を示し、ノードBは、文字列の右端を示す。このネットワークにおいて、左端ノードAから、右端ノードBへ至るパスを決定することによって、文字列を決定できることになる。

【0065】

図15の例を用いて説明する。図15は、文字列を構成するネットワーク内のノードが接続されている例を示す説明図である。

例えば、ノードD（文字候補「会」）に左から3本のリンクが接続されているとする。そして、そのリンクはそれぞれノードA（文字列候補「ヒ学」）、ノードB（文字列候補「化学」）、ノードC（文字列候補「ト学」）から接続されているとする。

ノードA、B、Cには、さらに左からノードが接続されているが、その部分は説明では省略している。ノードAの文字列候補は「ヒ学」、ノードBの文字列候補は「化学」、ノードCの文字列候補は「ト学」となっている。

ここで、ノードDに左から接続されるリンクを対象とする。ノードAからのパスであれば、文字列候補は「ヒ学会」であり、ノードBからのパスであれば、文字列候補は「化学会」であり、ノードCからのパスであれば、文字列候補は「ト学会」となる。被検索文生成モジュール130は、左端から右端に向かう複数のパスがあるので、そのパスの中から被検索対象である文字列候補を抽出すればよい。

【0066】

次に被検索文生成モジュール130による処理について説明する。

文字列画像の左端を始点、文字列画像の右端を終点として、始点から終点に向かう複数のパスを対象とする。なお、一つの準文字に対して、複数の候補が存在する場合があります、これら複数の候補も加味して、複数のパスを対象とする。前述したように、始点から終点までの各パスに評価値を与えることができる。この評価値の高いものから順にN個の文字認識出力（文字列候補）を得ることができる。このN個の文字認識出力をNベストと呼ぶこととする。

一般に、被検索文は長い。そのため、被検索文の先頭から終点を全て包含した文に対して、Nベストを取得しても、検索語にヒットする文字列を取得できる可能性は低くなる。

例えば、2ベストを用いて説明する。「E E G H I J K L < M O P Q R」が第1位文字列とする。第2位文字列は、例えば、この第1位文字列のうち1文字だけが置換されたものになり、「E E G H I J K L < M O P Q P」である。これは、最後の「R」が「P」に置換されたものである。正解は、「E F G H I J K L < M O P Q P」であったとする。つまり、「P」の認識確度が、2番目の認識文字「F」の認識確度よりも高い場合に、このような第2位文字列となる。ここで、検索語「E F G」で検索を行ったとしても、2ベスト（「E E G H I J K L < M O P Q R」、「E E G H I J K L < M O P Q P」）内には、その検索語はないので、一致しない（ヒットしない）という結果になる。

【0067】

これに対して、本実施の形態の被検索文生成モジュール130は、Nベストを部分文字列で抽出する。例えば、3文字からなる文字列で2ベストを生成した場合は、前述の例では先頭から3文字の「E E G」の第2位文字列は「E F G」となって、検索語「E F G」にヒットすることが可能となる。すなわち、始点から終点に向かうパスを対象としてNベストを生成するのではなく、それよりも短い文字列に対してNベストを生成する。ただし、極端な例で、1文字単位にNベストを生成する案がある。この場合、単語の検索語に対してヒットさせることができない。すなわち2文字以上であって、ネットワーク全体における始点から終点までのパスの文字数未満の長さの文字列に対して、Nベストを抽出すればよい。

そして、Nベストを抽出した結果、文字認識結果として、複数の結果をつなぎ合わせたテキストコードを出力すればよい。このテキストコードを検索することによって、Nベストのどれがヒットしても、文書を検索することができる。つまり、文字認識結果ファイルとして、文全体の第1候補に加えて、文の部分文字列のNベストを付与して、検索用のテキストファイルを作成する。

【0068】

図16は、被検索文生成モジュール130による処理例を示すフローチャートである。

ステップS1602では、左端決定モジュール142が、Nベスト左端位置を決定する。Nベストを抽出するための文字列の左端を「Nベスト左端」、Nベストを抽出するための文字列の右端を「Nベスト右端」と記す。

ステップS1604では、パス切出モジュール140が、文字認識結果の文字列が終了したか否かを判断し、終了した場合は処理を終了し（ステップS1699）、それ以外の場合はステップS1606へ進む。

ステップS1606では、右端決定モジュール144が、Nベスト右端位置を決定する。

ステップS1608では、Nベスト抽出モジュール150が、Nベスト左端位置からNベスト右端位置までの文字によって構成されるNベストを抽出する。

ステップS1610では、Nベスト抽出モジュール150が、抽出したNベストを被検索文記憶モジュール160に格納する。

【0069】

図17は、本実施の形態による処理例を示す説明図である。

準文字は、文字切れ目候補で区切られた文字、又は、文字部分となる。準文字は、画素として、連結領域の塊を一つの準文字としてもよいし、複数の連結領域をまとめて一つの

10

20

30

40

50

準文字としてもよい。あるいは、連結領域であっても、連結部分が細い場合には分割して複数の準文字としてもよい。

図17(a)の例に示す矩形は、準文字の外接矩形である。

以下では、図17(b)の例に示すように、一つの矩形が、一つの準文字であるとみなした説明を行う。

例えば、「化学」という文字列であれば、矩形の一つ一つに「イ」（「化」のへん）、「ヒ」（「化」の旁）、「学」が入る。

【0070】

図18は、本実施の形態による処理例を示す説明図である。

本実施の形態では、例えば、Nベストを抽出するための文字列の左端を、全ての文字切れ目候補で行う。そして、Nベストを抽出するための文字列の右端を、例えば、その左端からの準文字の個数（ノード数）で規定する。

図18は、準文字の個数を3として動作を示す例である。なお、前述したように、準文字の個数は3以外であってもよい。2以上であって、ネットワーク全体の始点から終点までの全ての準文字の個数未満であればよい。

まずNベスト左端を対象とする画像の左端とする（ステップS1602）。Nベスト左端から右方向へ3個の準文字分の位置をNベスト右端とする（ステップS1606）。

Nベスト左端を始点、Nベスト右端を終点とするパスの中で、Nベストを決定する。決定したNベスト（N個の文字列）を抽出する（ステップS1608、図18(a)）。

次に、1個分、準文字を右にシフトして（2巡目のステップS1602）、同様の処理を行う（図18(b)）。

対象とする文字列の終端に到達すれば（ステップS1604）、終了とする（ステップS1699、図18(c)）。

ここで求めたNベスト結果を、文字認識結果ファイルとして被検索文記憶モジュール160に記憶する（ステップS1610）ことで、検索が可能となる。

【0071】

図19は、本実施の形態による別の処理例を示す説明図である。

Nベスト左端とNベスト右端の距離は、準文字個数でなくともよい。

例えば、Nベスト左端に位置する準文字の高さ又はその準文字を含む文字列の高さをH（画素）とする。予め定められた数値（例えば、10（画素）等）をAとし、Nベスト左端とNベスト右端の距離を、 $X = H \times A$ 等として定める。

この値Xに最も近い位置の切れ目位置をNベスト右端とする。又は、値Xを最初に超える切れ目位置をNベスト右端とする。

なお、高さを「画素」単位で計測したが、「画素」ではなく、「mm（ミリメートル）」等の物理量であってもよい。

また、値XをHとは関係なく予め定められた値としてもよい。つまり、Nベスト左端から予め定められた距離（値X）だけ離れた位置をNベスト右端としてもよい。

【0072】

図20は、本実施の形態による処理例を示す説明図である。

Nベスト左端は、全ての文字切れ目候補を網羅する必要はない。例えば、Nベスト左端位置を、ネットワーク内の先頭から最後尾まで、予め定められた数毎に移動させてもよい。例えば、図20(a)に示すように、一つおきに移動させてもよい。

また、Nベスト左端を、準文字間の間隔が閾値よりも大きい又は以上である箇所としてもよい。例えば、図20(b)に示すように、準文字間が離れている箇所がNベスト左端となる。閾値としては、予め定められた値であってもよいし、その文字列における準文字間の距離の統計値（例えば、平均値、最頻値、中央値等又はそれらに定数を乗算したもの等）としてもよい。

【0073】

図21は、本実施の形態による処理例を示す説明図である。Nベスト左端とNベスト右端を決定することは、連続した準文字を選択することと同義である。つまり、「連続した

10

20

30

40

50

準文字の選択」と、「Nベスト左端、Nベスト右端の選択」に関し、原理的な違いは無い。

例えば、図21に示すように、準文字数を3とする場合、連続した3つの準文字を選択すればよい。左から順に選択してもよいし、ランダムに選択してもよい。

【0074】

図22は、本実施の形態による処理例を示す説明図である。前述したように、Nベスト右端を選択するための準文字の個数は固定であってもよい。また、変動する例として、単語と単語の間に空白が存在するような文書である場合には、空白で区切るようにしてもよい。

また、準文字の個数を複数用いてもよい。例えば、準文字個数を3個と4個として、Nベスト左端を固定したときに、2種のNベスト右端を抽出してもよい。図22(a)の例のように、まず3個でNベストを抽出し、次に図22(b)の例のように4個で抽出してもよい。

【0075】

図23は、より具体的な本実施の形態による処理例を示す説明図である。

文字認識処理モジュール110は、文字列画像全体を文字認識する。ここでは文字認識誤りが発生し得る。例えば、図23(a)に示すように「化学の組織論」という文字列画像を文字認識した場合、文字認識結果の第1候補は「イヒ学の組織論」のように文字切り出しの誤りが発生して、文字認識誤りが発生している。なお、「イヒ」の部分を「化」と文字認識しているものが、第2候補以下にある。

被検索文生成モジュール130は、文字認識結果を部分に分割して、複数の文字列候補を抽出する。その複数の文字列候補中に正解が存在する場合がある。例えば、図23(b)に示すように予め定められた長さの文字列画像を文字認識した結果は、「化学の」、「イヒ学の」となる。

そして、複数の文字列候補を抽出する箇所をずらして、さらに複数の文字列候補を抽出する。例えば、図23(c)に示すように準文字「イ」と準文字「ヒ」の間から始まる文字列画像を文字認識した結果は、「ヒ学の」、「ヒ字の」となる。

最終結果として、例えば、図24に示すように認識結果を抽出する。このように、第1候補の文字列「イヒ学の組織論」では、「化学」は検索できないが、図24に示すように認識結果のテキストファイルであれば、「化学」が検索可能となる。

【0076】

図25は、本実施の形態による処理例を示す説明図である。例えば、図25の左側に示す画像を文字認識して、図25の右側に示す文字認識結果ファイルを作成する。この文字認識結果ファイルは、図24の例に示すように、Nベスト左端とNベスト右端の間にある文字列画像を文字認識した結果であるNベスト（複数の文字列候補）が記憶されている。

ユーザは、図26の例に示すような「検索」窓に検索語であるテキストコードを打ち込み、データベース等に蓄積されている画像を検索する。なお、検索対象となる実体は、画像に対応する文字認識結果ファイルである。つまり、画像や、文字認識結果ファイルは、データベース等に蓄積されている。画像と文字認識結果ファイルの間にはリンクが張られているため、文字認識結果ファイルが検索できれば、画像ファイルも検索可能である。

図25の例に示すように、文字認識が正しければ、「ABC」の検索語に対し、文書Aが検索できる。文字認識が誤っている場合（「EFG」が正解であるが、第1候補は「EEG」となった場合）、第1候補のみからなる文字認識結果ファイルでは「EFG」の検索語を入れても、文書Bを検索できない。本実施の形態では、被検索文の拡張を行って、文書Bの文字認識結果ファイルでは、「EEG」、「HIJ」の文字認識結果の第1候補に加えて、3つの準文字からなる文字認識の第2候補以下である「EFG」、「HIJ」等も記憶しているので、検索を可能とする。

【0077】

図27は、検索モジュール180による処理例を示すフローチャートである。

図3の例に示したフローチャートのステップS304として、図27の例に示すような

処理を行ってもよい。

ステップS 2 7 0 2では、検索モジュール1 8 0が、検索文の文字数と一致する文字数のNベストを選択する。例えば、複数種の文字数のNベストが文字認識結果ファイル内に記憶されている場合は、検索文の文字数と一致するものを選択する。

ステップS 2 7 0 4では、検索モジュール1 8 0が、選択したNベストによって構成されている被検索文から検索する。

【0 0 7 8】

図2 8は、被検索文生成モジュール1 3 0による処理例を示すフローチャートである。

ステップS 2 8 0 2では、検索文の文字数とNベストの文字数の比較を行い、一致するNベストがある場合はステップS 2 8 0 4へ進み、検索文の文字数が短い場合はステップS 2 8 0 8へ進み、検索文の文字数が長い場合はステップS 2 8 1 2へ進む。

ステップS 2 8 0 4では、検索モジュール1 8 0が、検索文の文字数と一致する文字数のNベストを選択する。

ステップS 2 8 0 6では、検索モジュール1 8 0が、選択したNベストによって構成されている被検索文から検索する。

【0 0 7 9】

ステップS 2 8 0 8では、検索文の文字数に近いNベストを選択する。

ステップS 2 8 1 0では、選択したNベストによって構成される被検索文に対して、前方一致検索を行う。

ステップS 2 8 1 2では、検索文の文字数に近いNベストを選択する。

ステップS 2 8 1 4では、検索文の文字数以上になるようにNベストを連結する。この連結は、対象のNベスト同士の全ての組み合わせを生成するように行う。

ステップS 2 8 1 6では、連結したNベストによって構成される被検索文に対して、前方一致検索を行う。

【0 0 8 0】

図2 9を参照して、本実施の形態の情報処理装置のハードウェア構成例について説明する。図2 9に示す構成は、例えばパーソナルコンピュータ（P C）などによって構成されるものであり、スキャナ等のデータ読み取り部2 9 1 7と、プリンタなどのデータ出力部2 9 1 8を備えたハードウェア構成例を示している。

【0 0 8 1】

C P U（C e n t r a l P r o c e s s i n g U n i t）2 9 0 1は、前述の実施の形態において説明した各種のモジュール、すなわち、文字認識処理モジュール1 1 0、ネットワーク構築モジュール1 2 0、被検索文生成モジュール1 3 0、パス切出モジュール1 4 0、左端決定モジュール1 4 2、右端決定モジュール1 4 4、Nベスト抽出モジュール1 5 0、検索文受付モジュール1 7 0、検索モジュール1 8 0、画像受付モジュール4 1 0、文字列抽出モジュール4 2 0、切出位置抽出モジュール4 3 0 1、文字候補抽出モジュール4 4 0等の各モジュールの実行シーケンスを記述したコンピュータ・プログラムに従った処理を実行する制御部である。

【0 0 8 2】

R O M（R e a d O n l y M e m o r y）2 9 0 2は、C P U 2 9 0 1が使用するプログラムや演算パラメータ等を格納する。R A M（R a n d o m A c c e s s M e m o r y）2 9 0 3は、C P U 2 9 0 1の実行において使用するプログラムや、その実行において適宜変化するパラメータ等を格納する。これらはC P Uバスなどから構成されるホストバス2 9 0 4により相互に接続されている。

【0 0 8 3】

ホストバス2 9 0 4は、ブリッジ2 9 0 5を介して、P C I（P e r i p h e r a l C o m p o n e n t I n t e r c o n n e c t / I n t e r f a c e）バスなどの外部バス2 9 0 6に接続されている。

【0 0 8 4】

キーボード2 9 0 8、マウス等のポインティングデバイス2 9 0 9は、操作者により操

10

20

30

40

50

作される入力デバイスである。ディスプレイ 2910 は、液晶表示装置又は CRT (Cathode Ray Tube) などがあり、各種情報をテキストやイメージ情報として表示する。

【0085】

HDD (Hard Disk Drive) 2911 は、ハードディスクを内蔵し、ハードディスクを駆動し、CPU 2901 によって実行するプログラムや情報を記録又は再生させる。ハードディスクには、受け付けた画像、文字認識結果、ネットワーク構造、対象文字列、被検索文、検索文などが格納される。さらに、その他の各種のデータ処理プログラム等、各種コンピュータ・プログラムが格納される。

【0086】

ドライブ 2912 は、装着されている磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリ等のリムーバブル記録媒体 2913 に記録されているデータ又はプログラムを読み出して、そのデータ又はプログラムを、インタフェース 2907、外部バス 2906、ブリッジ 2905、及びホストバス 2904 を介して接続されている RAM 2903 に供給する。リムーバブル記録媒体 2913 も、ハードディスクと同様のデータ記録領域として利用可能である。

【0087】

接続ポート 2914 は、外部接続機器 2915 を接続するポートであり、USB、IEEE 1394 等の接続部を持つ。接続ポート 2914 は、インタフェース 2907、及び外部バス 2906、ブリッジ 2905、ホストバス 2904 等を介して CPU 2901 等に接続されている。通信部 2916 は、通信回線に接続され、外部とのデータ通信処理を実行する。データ読み取り部 2917 は、例えばスキャナであり、ドキュメントの読み取り処理を実行する。データ出力部 2918 は、例えばプリンタであり、ドキュメントデータの出力処理を実行する。

【0088】

なお、図 29 に示す情報処理装置のハードウェア構成は、1つの構成例を示すものであり、本実施の形態は、図 29 に示す構成に限らず、本実施の形態において説明したモジュールを実行可能な構成であればよい。例えば、一部のモジュールを専用のハードウェア (例えば特定用途向け集積回路 (Application Specific Integrated Circuit: ASIC) 等) で構成してもよく、一部のモジュールは外部のシステム内にあり通信回線で接続しているような形態でもよく、さらに図 29 に示すシステムが複数互いに通信回線によって接続されていて互いに協調動作するようにしてもよい。また、複写機、ファックス、スキャナ、プリンタ、複合機 (スキャナ、プリンタ、複写機、ファックス等のいずれか 2 つ以上の機能を有している画像処理装置) などに組み込まれていてもよい。

【0089】

前述の実施の形態においては、対象として日本語の文字を示したが、中国語、英語等の文字を対象としてもよい。

前述の実施の形態においては、横書きの文字列を前提として、左が始点で右が終点であるような記述の仕方を行ってきた。しかし、前述の説明は、縦書きや、右から左に記述する文字列でも同様に成り立つ。例えば、縦書きの場合は、「左」を「上」、「右」を「下」とする変換を行えばよい。右から左に記述する文字列の場合は、「左」を「右」、「右」を「左」とする変換を行えばよい。

なお、数式を用いて説明したが、数式には、その数式と同等のものが含まれる。同等のものとは、その数式そのものの他に、最終的な結果に影響を及ぼさない程度の数式の変形、又は数式をアルゴリズム的な解法で解くこと等が含まれる。

【0090】

なお、説明したプログラムについては、記録媒体に格納して提供してもよく、また、そのプログラムを通信手段によって提供してもよい。その場合、例えば、前記説明したプログラムについて、「プログラムを記録したコンピュータ読み取り可能な記録媒体」の発明

10

20

30

40

50

として捉えてもよい。

「プログラムを記録したコンピュータ読み取り可能な記録媒体」とは、プログラムのインストール、実行、プログラムの流通などのために用いられる、プログラムが記録されたコンピュータで読み取り可能な記録媒体をいう。

なお、記録媒体としては、例えば、デジタル・バーサタイル・ディスク（DVD）であって、DVDフォーラムで策定された規格である「DVD-R、DVD-RW、DVD-RAM等」、DVD+RWで策定された規格である「DVD+R、DVD+RW等」、コンパクトディスク（CD）であって、読出し専用メモリ（CD-ROM）、CDレコーダブル（CD-R）、CDリライタブル（CD-RW）等、ブルーレイ・ディスク（Blu-ray Disc（登録商標））、光磁気ディスク（MO）、フレキシブルディスク（FD）、磁気テープ、ハードディスク、読出し専用メモリ（ROM）、電氣的消去及び書換可能な読出し専用メモリ（EEPROM（登録商標））、フラッシュ・メモリ、ランダム・アクセス・メモリ（RAM）、SD（Secure Digital）メモリーカード等が含まれる。

そして、前記のプログラム又はその一部は、前記記録媒体に記録して保存や流通等させてもよい。また、通信によって、例えば、ローカル・エリア・ネットワーク（LAN）、メトロポリタン・エリア・ネットワーク（MAN）、ワイド・エリア・ネットワーク（WAN）、インターネット、イントラネット、エクストラネット等に用いられる有線ネットワーク、あるいは無線通信ネットワーク、さらにこれらの組み合わせ等の伝送媒体を用いて伝送させてもよく、また、搬送波に乗せて搬送させてもよい。

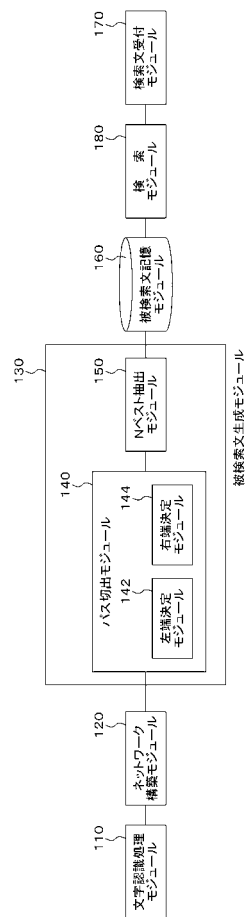
さらに、前記のプログラムは、他のプログラムの一部分であってもよく、あるいは別個のプログラムと共に記録媒体に記録されていてもよい。また、複数の記録媒体に分割して記録されていてもよい。また、圧縮や暗号化など、復元可能であればどのような態様で記録されていてもよい。

【符号の説明】

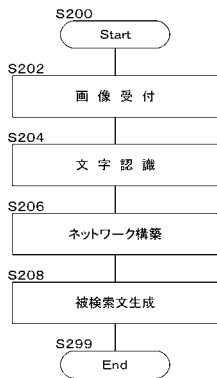
【0091】

- 110…文字認識処理モジュール
- 120…ネットワーク構築モジュール
- 130…被検索文生成モジュール
- 140…パス切出モジュール
- 142…左端決定モジュール
- 144…右端決定モジュール
- 150…Nベスト抽出モジュール
- 160…被検索文記憶モジュール
- 170…検索文受付モジュール
- 180…検索モジュール
- 410…画像受付モジュール
- 420…文字列抽出モジュール
- 430…切出位置抽出モジュール
- 440…文字候補抽出モジュール

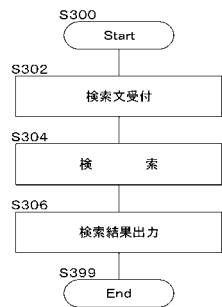
【図 1】



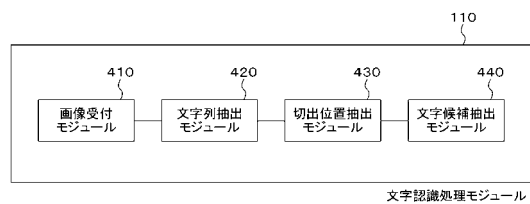
【図 2】



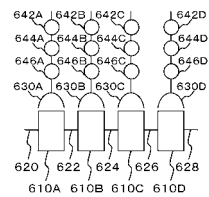
【図 3】



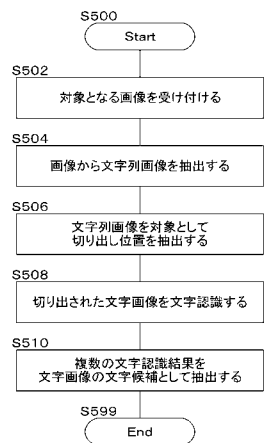
【図 4】



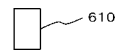
【図 6】



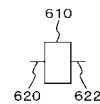
【図 5】



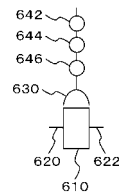
【図 7】



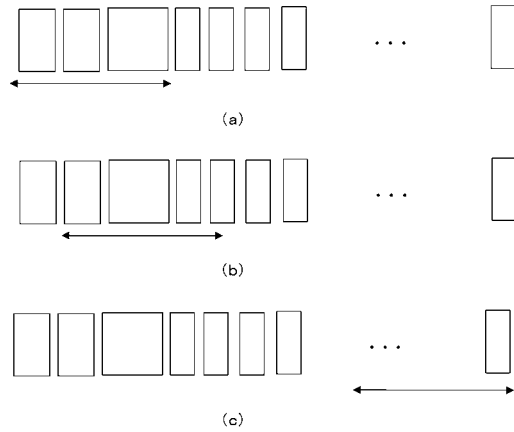
【図 8】



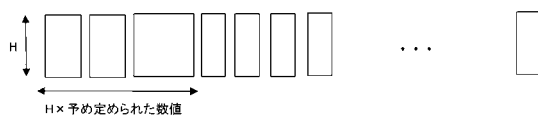
【図 9】



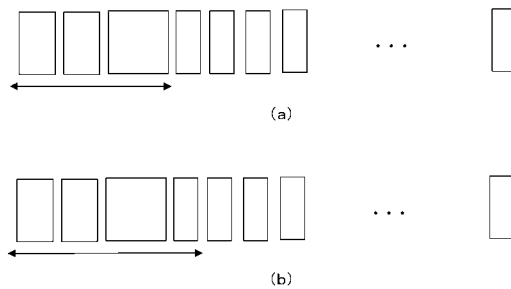
【図 18】



【図 19】



【図 22】



【図 23】

化学の組織論

(a)

認識結果
イヒ学の組織論

化学の組織論

(b)

認識結果
化学の
イヒ学の

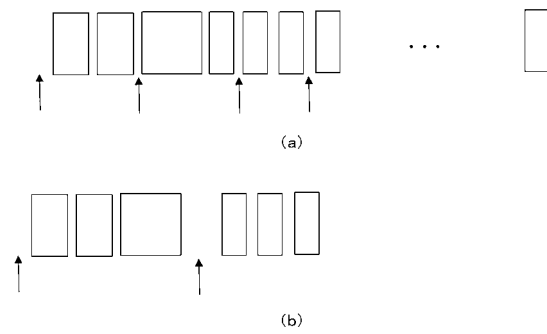
化学の組織論

(c)

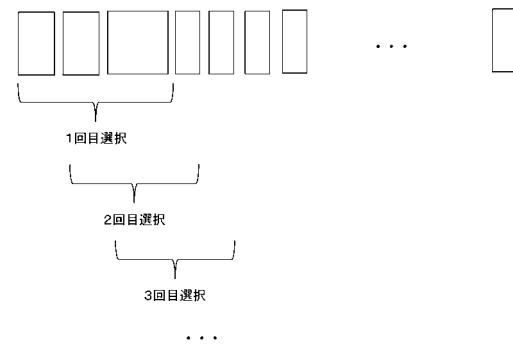
認識結果
ヒ学の
ヒ字の

...

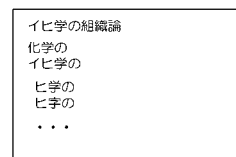
【図 20】



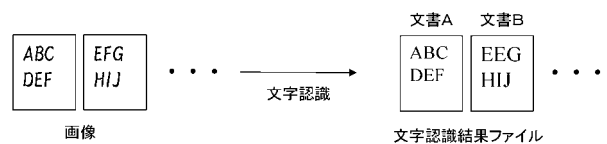
【図 21】



【図 24】



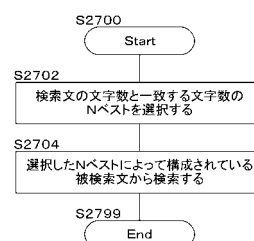
【図 25】



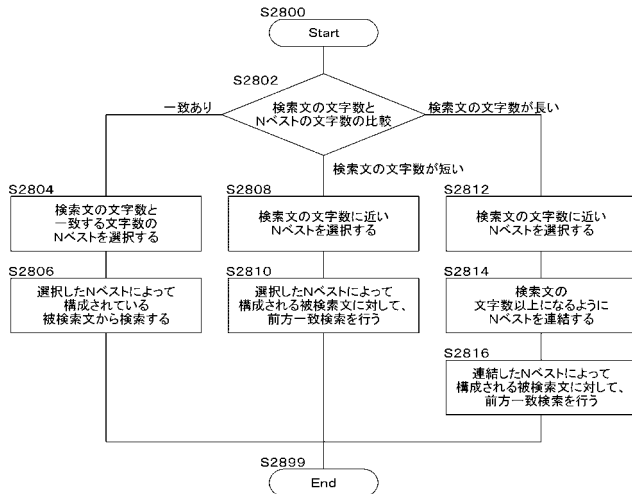
【図 26】



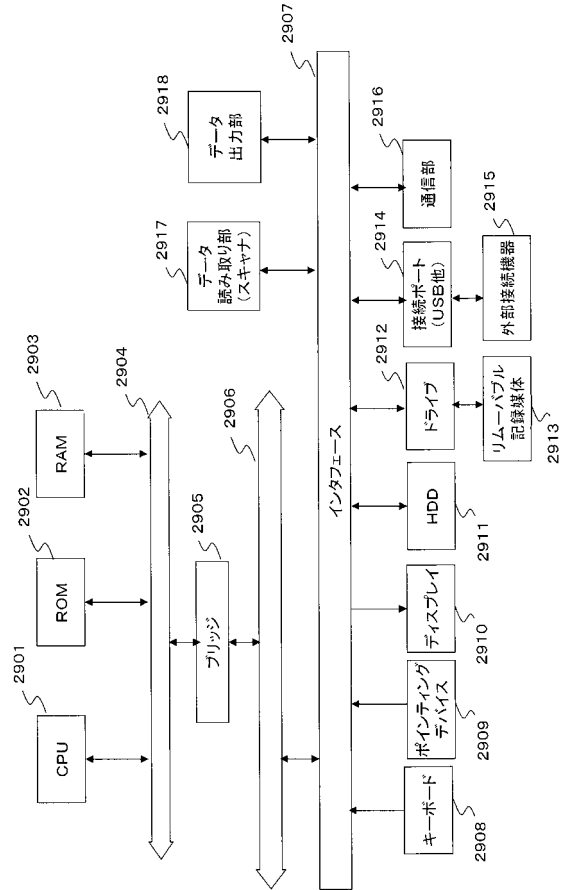
【図 27】



【図 28】



【図 29】



【図 30】

化学

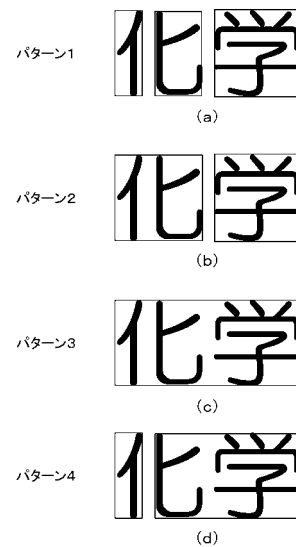
【図 31】

化学

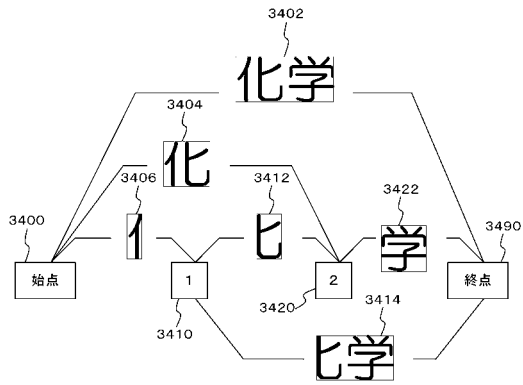
【図 32】

化学

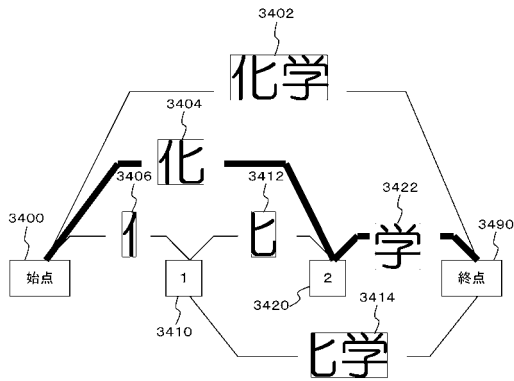
【図 33】



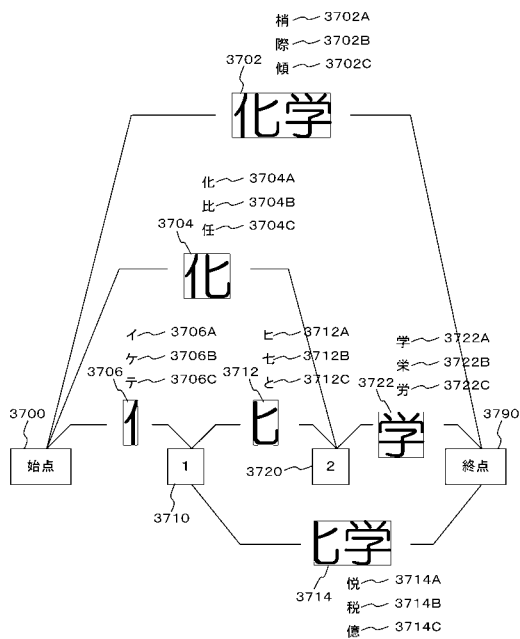
【図 3 4】



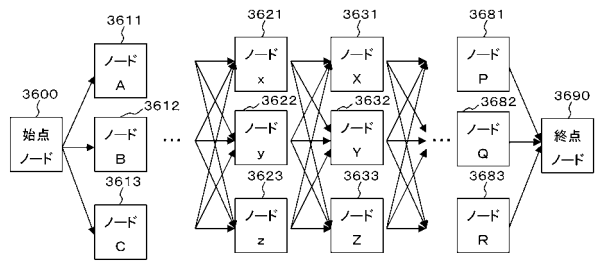
【図 3 5】



【図 3 7】



【図 3 6】



フロントページの続き

- (72)発明者 桜井 拓也
神奈川県足柄上郡中井町境430 グリーンテクなかい 富士ゼロックス株式会社内
- (72)発明者 鷹合 基行
神奈川県横浜市西区みなとみらい六丁目1番 富士ゼロックス株式会社内
- (72)発明者 外池 昌嗣
神奈川県横浜市西区みなとみらい六丁目1番 富士ゼロックス株式会社内
- (72)発明者 山根 洋平
神奈川県横浜市西区みなとみらい六丁目1番 富士ゼロックス株式会社内
- Fターム(参考) 5B064 AA01 AB02 AB13 AB14 AB16 BA01 EA18