

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第3区分

【発行日】平成17年9月8日(2005.9.8)

【公開番号】特開2001-318948(P2001-318948A)

【公開日】平成13年11月16日(2001.11.16)

【出願番号】特願2000-142232(P2000-142232)

【国際特許分類第7版】

G 06 F 17/30

【F I】

G 06 F 17/30 350 C

G 06 F 17/30 170 A

G 06 F 17/30 340 B

G 06 F 17/30 380 Z

【手続補正書】

【提出日】平成17年3月17日(2005.3.17)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

文書情報を文字コードデータとして蓄積したテキストデータベースを対象として、
登録対象テキストデータから所定の部分文字列と該部分文字列の該登録対象テキストデータにおける出現情報を抽出し、出現情報ファイルとして記憶する文書登録ステップと、
検索条件から所定の部分文字列を抽出し、該出現情報ファイルから取得した該部分文字列の出現情報を用いて、検索条件の内容とテキストデータベース中のテキストの内容が類似する度合いを類似度として所定の方法で算出する文書検索ステップを有する文書検索方法において、

該部分文字列から重要な部分文字列を選択し、該類似度を算出する際に用いた情報の中で、該重要な部分文字列に関する情報を表示する重要情報表示ステップを有することを特徴とした文書検索方法。

【請求項2】

請求項1載の文書検索方法において、前記重要情報表示ステップは、前記部分文字列の重要度を所定の方法で算出し、該重要度の降順に所定の数の部分文字列を抽出する高重要度部分文字列抽出ステップと、該高重要度部分文字列抽出ステップで抽出した部分文字列に関する前記出現情報あるいは前記類似度の算出に寄与した度合いを表示する高重要度情報表示ステップを有し、

前記重要な部分文字列とは、該高重要度部分文字列抽出ステップで、該重要度の降順に所定の数だけ抽出した部分文字列であり、

前記重要な部分文字列に関する情報とは、該重要な部分文字列に関する該出現情報あるいは該類似度の算出に寄与した度合いを含むことを特徴とした文書検索方法。

【請求項3】

請求項1記載の文書検索方法において、

前記文書検索ステップは、検索条件から所定の部分文字列を抽出し、情報を表示する部分文字列を該部分文字列の中からユーザに選択させる情報表示部分文字列選択ステップと、

前記出現情報ファイルから取得した該部分文字列の出現情報を用いて、検索条件の内容

とテキストデータベース中のテキストの内容が類似する度合いを類似度として所定の方法で算出する類似度算出ステップを有し、

前記重要な情報表示ステップは、該情報表示部分文字列選択ステップにおいて選択された部分文字列を取得する選択部分文字列取得ステップと、該選択部分文字列取得ステップで取得した部分文字列に関する前記出現情報あるいは前記類似度の算出に寄与した度合いを表示する選択情報表示ステップを有し、

前記重要な部分文字列とは、該情報表示部分文字列選択ステップにおいて、選択された部分文字列であり、

前記重要な部分文字列に関する情報とは、該重要な部分文字列に関する該出現情報あるいは該類似度の算出に寄与した度合いを含むことを特徴とした文書検索方法。

【請求項4】

請求項1記載の文書検索方法において、

前記文書検索ステップは、検索条件から所定の部分文字列を抽出し、該部分文字列に対して部分文字列の追加あるいは削除をユーザに実施させる部分文字列編集ステップと、

該部分文字列編集ステップで編集された該部分文字列の出現情報を、前記出現情報ファイルから取得し、該出現情報を用いて、検索条件の内容とテキストデータベース中のテキストの内容が類似する度合いを類似度として所定の方法で算出する編集後類似度算出ステップを有し、

前記重要な情報表示ステップは、該部分文字列編集ステップにおいて追加された部分文字列を取得する追加部分文字列取得ステップと、該追加部分文字列取得ステップで取得した部分文字列に関する前記出現情報あるいは前記類似度の算出に寄与した度合いを表示する追加情報表示ステップを有し、

前記重要な部分文字列とは、該部分文字列編集ステップにおいて、追加された部分文字列であり、

前記重要な部分文字列に関する情報とは、該重要な部分文字列に関する該出現情報あるいは該類似度の算出に寄与した度合いを含むことを特徴とした文書検索方法。

【請求項5】

請求項1記載の文書検索方法において、

前記重要な情報表示ステップは、前記部分文字列から前記類似度の算出に寄与した順に所定の数の部分文字列を抽出する寄与部分文字列抽出ステップと、該寄与部分文字列抽出ステップで抽出した部分文字列に関する前記出現情報あるいは該類似度の算出に寄与した度合いを表示する寄与情報表示ステップを有し、

前記重要な部分文字列とは、該寄与部分文字列抽出ステップで、該類似度の算出に寄与した順に所定の数だけ抽出した部分文字列であり、

前記重要な部分文字列に関する情報とは、該重要な部分文字列に関する該出現情報あるいは該類似度の算出に寄与した度合いを含むことを特徴とした文書検索方法。

【請求項6】

文書情報を文字コードデータとして蓄積したテキストデータベースを対象として、

登録対象テキストデータから所定の部分文字列と該部分文字列の該登録対象テキストデータにおける出現情報を抽出し、出現情報ファイルとして記憶する文書登録ステップと、

検索条件から所定の部分文字列を抽出し、該出現情報ファイルから取得した該部分文字列の出現情報を用いて、検索条件の内容とテキストデータベース中のテキストの内容が類似する度合いを類似度として所定の方法で算出する文書検索ステップと、

該文書検索ステップで該類似度を算出したテキストの中でユーザに指定されたテキストを表示する文書表示ステップを有する文書検索方法において、

該文書表示ステップは、該部分文字列から重要な部分文字列を選択し、該指定されたテキストの中で、該重要な部分文字列が含まれる部分の表示形式を変更する表示形式変更ステップを有することを特徴とした文書検索方法。

【請求項7】

請求項6記載の文書検索方法において、

前記文書登録ステップは、前記登録対象テキストデータから所定の部分文字列と該部分文字列の該登録対象テキストデータにおける出現情報を抽出し、出現情報ファイルとして記憶する出現情報ファイル作成ステップと、

該部分文字列の重要度を所定の方法で算出し、重要度ファイルとして記憶する重要度ファイル作成ステップを有し、

前記文書表示ステップは、該部分文字列の重要度を該重要度ファイルから取得し、該取得した重要度の降順に所定の数の部分文字列を抽出する重要部分文字列抽出ステップと、

前記指定されたテキストの中で、該重要部分文字列抽出ステップで抽出した部分文字列が含まれる部分の表示形式を変更する重要部分文字列表示形式変更ステップを有し、

前記重要な部分文字列とは、該重要部分文字列抽出ステップで、該重要度の降順に所定の数だけ抽出した部分文字列であることを特徴とした文書検索方法。

【請求項 8】

請求項 6 記載の文書検索方法において、

前記文書検索ステップは、検索条件から所定の部分文字列を抽出し、表示形式を変更する部分文字列を該部分文字列の中からユーザに選択させる表示形式変更部分文字列選択ステップと、

前記出現情報ファイルから取得した該部分文字列の出現情報を用いて、検索条件の内容とテキストデータベース中のテキストの内容が類似する度合いを類似度として所定の方法で算出する類似度算出ステップを有し、

前記文書表示ステップは、該表示形式変更部分文字列選択ステップにおいて選択された部分文字列を取得する選択部分文字列取得ステップと、

前記指定されたテキストの中で、該選択部分文字列取得ステップで取得した部分文字列が含まれる部分の表示形式を変更する選択部分文字列表示形式変更ステップを有し、

前記重要な部分文字列とは、該表示形式変更部分文字列選択ステップにおいて選択された該表示形式を変更する部分文字列であることを特徴とした文書検索方法。

【請求項 9】

請求項 6 記載の文書検索方法において、前記文書検索ステップは、検索条件から所定の部分文字列を抽出し、該部分文字列に対しユーザに部分文字列の追加あるいは削除を実行させる部分文字列編集ステップと、

該部分文字列編集ステップで編集された該部分文字列の出現情報を、前記出現情報ファイルから取得し、該出現情報を用いて、検索条件の内容とテキストデータベース中のテキストの内容が類似する度合いを類似度として所定の方法で算出する編集後類似度算出ステップを有し、

前記文書表示ステップは、該部分文字列編集ステップにおいて追加された部分文字列を取得する追加部分文字列取得ステップと、

前記指定されたテキストの中で、該追加部分文字列取得ステップで取得した部分文字列が含まれる部分の表示形式を変更する追加部分文字列表示形式変更ステップを有し、

前記重要な部分文字列とは、該部分文字列編集ステップにおいて追加された該部分文字列であることを特徴とした文書検索方法。

【請求項 10】

請求項 6 記載の文書検索方法において、

前記文書表示ステップは、前記部分文字列から前記類似度の算出に寄与した順に所定の数の部分文字列を抽出する寄与部分文字列抽出ステップと、前記指定されたテキストの中で、該寄与部分文字列抽出ステップで抽出した部分文字列が含まれる部分の表示形式を変更する寄与部分文字列表示形式変更ステップを有し、

前記重要な部分文字列とは、該寄与部分文字列抽出ステップで、該類似度の算出に寄与した順に所定の数だけ抽出した部分文字列であることを特徴とした文書検索方法。

【請求項 11】

文書情報を文字コードデータとして蓄積したテキストデータベースを対象として、

登録対象テキストデータから所定の部分文字列と該部分文字列の該登録対象テキストデ

ータにおける出現情報を抽出し、出現情報ファイルとして記憶する文書登録手段と、

検索条件から所定の部分文字列を抽出し、該出現情報ファイルから取得した該部分文字列の出現情報を用いて、検索条件の内容とテキストデータベース中のテキストの内容が類似する度合いを類似度として所定の方法で算出する文書検索手段と、

該文書検索手段で該類似度を算出したテキストの中でユーザに指定されたテキストを表示する文書表示手段を備える文書検索装置において、

該文書表示手段は、該部分文字列から重要な部分文字列を選択し、該指定されたテキストの中で該重要な部分文字列が含まれる部分の表示形式を変更する表示形式変更手段を備えることを特徴とした文書検索装置。

【請求項 1 2】

文書情報を文字コードデータとして蓄積したテキストデータベースを対象として、
登録対象テキストデータから所定の部分文字列と該部分文字列の該登録対象テキストデータにおける出現情報を抽出し、出現情報ファイルとして記憶する文書登録モジュールと

、
検索条件から所定の部分文字列を抽出し、該出現情報ファイルから取得した該部分文字列の出現情報を用いて、検索条件の内容とテキストデータベース中のテキストの内容が類似する度合いを類似度として所定の方法で算出する文書検索モジュールと、

該文書検索モジュールで該類似度を算出したテキストの中でユーザに指定されたテキストを表示する文書表示モジュールを含む文書検索システムを構築するためのプログラムを格納した記憶媒体において、

前記プログラムにおいて、該文書表示モジュールは、該部分文字列から重要な部分文字列を選択し、該指定されたテキストの中で該重要な部分文字列が含まれる部分の表示形式を変更する表示形式変更モジュールを含むことを特徴とした記憶媒体。