

【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 19 年 2 月 15 日 (2007.2.15)

【公開番号】特開 2004-348706 (P2004-348706A)

【公開日】平成 16 年 12 月 9 日 (2004.12.9)

【年通号数】公開・登録公報 2004-048

【出願番号】特願 2004-67779 (P2004-67779)

【国際特許分類】

G 0 6 F 17/30 (2006.01)

G 0 6 Q 50/00 (2006.01)

G 0 6 T 1/00 (2006.01)

G 0 6 T 7/00 (2006.01)

【F I】

G 0 6 F 17/30 1 7 0 B

G 0 6 F 17/30 3 5 0 C

G 0 6 F 17/60 1 4 2

G 0 6 T 1/00 2 0 0 C

G 0 6 T 7/00 3 0 0 F

【手続補正書】

【提出日】平成 18 年 12 月 20 日 (2006.12.20)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

入力される文書画像に類似する画像データを検索する情報処理装置であって、

前記入力される文書画像を属性ごとに複数の領域に分割する分割手段と、

前記分割された領域ごとに、前記属性に基づく検索手段を用いて、類似度を算出する類似度算出手段と、

前記分割された領域ごとに算出された類似度に重み付けをした総合類似度を算出する総合類似度算出手段と

を備えることを特徴とする情報処理装置。

【請求項 2】

前記検索手段は、

前記分割手段により分割された領域のうち、テキスト領域から文字認識により抽出されたテキストデータに基づいて全文検索する手段と、

前記テキストデータに基づいて概念検索する手段と、

前記分割手段により分割された領域のうち、イメージ領域から抽出された特徴量に基づいてイメージ検索する手段と

を備え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】

前記総合類似度算出手段は、

前記入力される文書画像における前記分割された各領域の占有率に基づいて、重み付けを行う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 4】

前記画像データごとの著作権ロイヤリティに関する情報を格納したサーバと通信する通信

手段を更に備え、

前記画像データについて、前記著作権ロイヤリティに関する情報を取得可能であ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 5】

前記著作権ロイヤリティに関する情報があつた場合には、前記サーバからの許可を受けない限り、前記画像データについての処理が行われないことを特徴とする請求項 4 に記載の情報処理装置。

【請求項 6】

ユーザに対して、前記著作権ロイヤリティに基づく課金処理を実行する課金手段を更に備えることを特徴とする請求項 4 に記載の情報処理装置。

【請求項 7】

前記入力される文書画像をベクトルデータに変換する際の料金をユーザに対して課金する課金手段と、

前記課金手段により課金された料金の支払い手続きが行われた場合に、前記入力される文書画像をベクトルデータに変換するベクトルデータ変換手段と、

前記ベクトルデータに変換された文書画像を格納する格納手段と
を更に有す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 8】

前記課金手段は、前記ベクトルデータに変換された文書画像の著作権ロイヤリティを管理する際の料金を課金することを特徴とする請求項 7 に記載の情報処理装置。

【請求項 9】

前記入力される文書画像をベクトルデータに変換するベクトルデータ変換手段と、

前記ベクトルデータに変換された文書画像の特徴に基づいて、格納先の候補カテゴリを出力する候補カテゴリ出力手段と

を更に備え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 10】

前記候補カテゴリ出力手段により出力された候補カテゴリの中から、格納先のカテゴリをユーザに選択させるカテゴリ選択手段と、

前記カテゴリ選択手段で選択されたカテゴリに前記ベクトルデータに変換された文書画像を格納する格納手段と

を更に備えることを特徴とする請求項 9 に記載の情報処理装置。

【請求項 11】

前記候補カテゴリ出力手段により出力された第 1 位の候補カテゴリに対し、前記ベクトルデータに変換された文書画像を格納する格納手段を更に備えることを特徴とする請求項 9 に記載の情報処理装置。

【請求項 12】

前記ベクトルデータ変換手段は、前記総合類似度算出手段の算出結果に基づいて前記文書画像に対応するオリジナルデータファイルが見つからなかった場合に、前記文書画像をベクトルデータに変換することを特徴とする請求項 9 に記載の情報処理装置。

【請求項 13】

前記ベクトルデータ変換手段は、前記文書画像に対応するオリジナルデータファイルがイメージファイルであつた場合に、前記文書画像をベクトルデータに変換することを特徴とする請求項 9 に記載の情報処理装置。

【請求項 14】

更に、前記入力される文書画像と比較対象の画像データとのレイアウト類似度を算出し、当該算出されたレイアウト類似度が予め定められた閾値以上の前記比較対象画像データについては前記類似度算出手段及び前記総合類似度算出手段による処理を実行するよう制御し、当該算出されたレイアウト類似度が予め定められた閾値より低い前記比較対象画像データについては前記類似度算出手段及び前記総合類似度算出手段による処理を実行しないように制御する、レイアウト類似度算出手段を有す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 15】

更に、前記分割手段で分割された複数の領域のうちから類似度算出に用いる領域を選択させる選択手段を有し、

前記類似度算出手段は、前記選択手段で選択された領域に関して、前記属性に適した検索手段を用いて、類似度を算出す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 16】

入力される文書画像に類似する画像データを検索する情報処理方法であって、

前記入力される文書画像を属性ごとに複数の領域に分割する分割工程と、

前記分割された領域ごとに、前記属性に基づく検索工程を用いて、類似度を算出する類似度算出工程と、

前記分割された領域ごとに算出された類似度に重み付けをした総合類似度を算出する総合類似度算出工程と

を備えることを特徴とする情報処理方法。

【請求項 17】

前記検索工程は、

前記分割工程により分割された領域のうち、テキスト領域から文字認識により抽出されたテキストデータに基づいて全文検索する工程と、

前記テキストデータに基づいて概念検索する工程と、

前記分割工程により分割された領域のうち、イメージ領域から抽出された特徴量に基づいてイメージ検索する工程と

を備えることを特徴とする請求項 16 に記載の情報処理方法。

【請求項 18】

前記総合類似度算出工程は、

前記入力される文書画像における前記分割された各領域の占有率に基づいて、重み付けを行うことを特徴とする請求項 16 に記載の情報処理方法。

【請求項 19】

前記画像データごとの著作権ロイヤリティに関する情報を格納したサーバと通信する通信工程を更に備え、

前記画像データについて、前記著作権ロイヤリティに関する情報を取得可能であることを特徴とする請求項 16 に記載の情報処理方法。

【請求項 20】

前記著作権ロイヤリティに関する情報があった場合には、前記サーバからの許可を受けない限り、前記画像データについての処理が行われないことを特徴とする請求項 19 に記載の情報処理方法。

【請求項 21】

ユーザに対して、前記著作権ロイヤリティに基づく課金処理を実行する課金工程を更に備えることを特徴とする請求項 19 に記載の情報処理方法。

【請求項 22】

前記入力される文書画像をベクトルデータに変換する際の料金をユーザに対して課金する課金工程と、

前記課金工程により課金された料金の支払い手続きが行われた場合に、前記入力される文書画像をベクトルデータに変換するベクトルデータ変換工程と、

前記ベクトルデータに変換された文書画像を格納する格納工程と

を更に有することを特徴とする請求項 16 に記載の情報処理方法。

【請求項 23】

前記課金工程は、前記ベクトルデータに変換された文書画像の著作権ロイヤリティを管理する際の料金を課金することを特徴とする請求項 22 に記載の情報処理方法。

【請求項 24】

前記入力される文書画像をベクトルデータに変換するベクトルデータ変換工程と、

前記ベクトルデータに変換された文書画像の特徴に基づいて、格納先の候補カテゴリを出力する候補カテゴリ出力工程と

を更に備えることを特徴とする請求項 16 に記載の情報処理方法。

【請求項 25】

前記候補カテゴリ出力工程により出力された候補カテゴリの中から、格納先のカテゴリをユーザに選択させるカテゴリ選択工程と、

前記カテゴリ選択工程で選択されたカテゴリに前記ベクトルデータに変換された文書画像を格納する格納工程と

を更に備えることを特徴とする請求項 24 に記載の情報処理方法。

【請求項 26】

前記候補カテゴリ出力工程により出力された第 1 位の候補カテゴリに対し、前記ベクトルデータに変換された文書画像を格納する格納工程を更に備えることを特徴とする請求項 24 に記載の情報処理方法。

【請求項 27】

前記ベクトルデータ変換工程は、前記総合類似度算出工程の算出結果に基づいて前記文書画像に対応するオリジナルデータファイルが見つからなかった場合に、前記文書画像をベクトルデータに変換することを特徴とする請求項 24 に記載の情報処理方法。

【請求項 28】

前記ベクトルデータ変換工程は、前記文書画像に対応するオリジナルデータファイルがイメージファイルであった場合に、前記文書画像をベクトルデータに変換することを特徴とする請求項 24 に記載の情報処理方法。

【請求項 29】

更に、前記入力される文書画像と比較対象の画像データとのレイアウト類似度を算出し、当該算出されたレイアウト類似度が予め定められた閾値以上の前記比較対象画像データについては前記類似度算出工程及び前記総合類似度算出工程による処理を実行するよう制御し、当該算出されたレイアウト類似度が予め定められた閾値より低い前記比較対象画像データについては前記類似度算出工程及び前記総合類似度算出工程による処理を実行しないように制御する、レイアウト類似度算出工程を有することを特徴とする請求項 16 に記載の情報処理方法。

【請求項 30】

更に、前記分割工程で分割された複数の領域のうちから類似度算出に用いる領域を選択させる選択工程を有し、

前記類似度算出工程は、前記選択工程で選択された領域に関して、前記属性に適した検索工程を用いて、類似度を算出することを特徴とする請求項 16 に記載の情報処理方法。

【請求項 31】

請求項 16 乃至 30 のいずれか 1 つに記載の情報処理方法をコンピュータによって実行させるための制御プログラムを格納した記憶媒体。

【請求項 32】

請求項 16 乃至 30 のいずれか 1 つに記載の情報処理方法をコンピュータによって実行させるための制御プログラム。