

【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30年3月8日(2018.3.8)

【公表番号】特表2017-507444(P2017-507444A)

【公表日】平成29年3月16日(2017.3.16)

【年通号数】公開・登録公報2017-011

【出願番号】特願2016-567451(P2016-567451)

【国際特許分類】

G 06 F 17/27 (2006.01)

G 06 F 17/30 (2006.01)

G 06 F 17/21 (2006.01)

【F I】

G 06 F 17/27 6 4 5

G 06 F 17/30 3 8 0 C

G 06 F 17/21 6 7 0

【手続補正書】

【提出日】平成30年1月22日(2018.1.22)

【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】

コンピュータ上に記憶され、複数のページを有し、各ページがページコンテンツを有する電子文書に、自動的にインデックスを付ける方法であって、

前記文書の前記複数のページのうちの第1のページ内の第1の領域の選択をユーザから受け取る工程であって、前記第1の領域は、前記複数のページのうちの前記第1のページに対して第1の境界セットによって画定され、第1のベース選択ページコンテンツを含む、受け取る工程と、

前記複数のページのうちの前記第1のページから前記第1のベース選択ページコンテンツのテキスト文字列をユーザ介入なく取り出す工程と、

前記第1のベース選択ページコンテンツの取り出された前記テキスト文字列を、前記複数のページのうちの前記第1のページに関連付けられているページ・ロケーション・インデックスに、ユーザ介入なく割り当てる工程と、

前記複数のページのうちの第2のページから第1の複製された選択ページコンテンツのテキスト文字列をユーザ介入なく取り出す工程であって、前記第1の複製された選択ページコンテンツは、前記複数のページのうちの前記第2のページに対して前記第1の境界セットによって画定された同じ前記第1の領域内に含まれる、取り出す工程と、

前記第1の複製された選択ページコンテンツの取り出された前記テキスト文字列を、前記複数のページのうちの前記第2のページのページ・ロケーション・インデックスに、ユーザ介入なく割り当てる工程と、を備える、方法。

【請求項2】

前記ページ・ロケーション・インデックスはページ・ラベルである、請求項1に記載の方法。

【請求項3】

前記ページ・ロケーション・インデックスの1つは、付加されたページ・ブックマークである、請求項1に記載の方法。

【請求項 4】

第1の領域の選択を受け取る前記工程は、前記第1の境界セットに対応する第1の点および第2の点のカーソル入力で指定された前記文書上のグラフィックで選択された区域を含む、請求項1に記載の方法。

【請求項 5】

第1のベース選択ページコンテンツのテキスト文字列を取り出す前記工程は、

前記第1のベース選択ページコンテンツ内に位置決めされた基礎的なテキスト・データを抽出する工程をさらに含む、請求項1に記載の方法。

【請求項 6】

第1のベース選択ページコンテンツのテキスト文字列を取り出す前記工程は、

前記第1のベース選択ページコンテンツを光学式文字認識(OCR)モジュールにより構文解析し、前記第1のベース選択ページコンテンツの前記テキスト文字列を生成する工程をさらに含む、請求項1に記載の方法。

【請求項 7】

第1の複製された選択ページコンテンツのテキスト文字列を取り出す前記工程は、

前記第1の複製された選択ページコンテンツを前記OCRモジュールにより構文解析し、前記第1の複製された選択ページコンテンツの前記テキスト文字列を生成する工程をさらに含む、請求項6に記載の方法。

【請求項 8】

前記第1の領域に関連付けられている第1のフィールド・コードを含む書式構造を受け取る工程をさらに備え、

第1のベース選択ページコンテンツのテキスト文字列を取り出す前記工程は、前記書式構造に従って前記第1のベース選択ページコンテンツの前記テキスト文字列を配置する工程を含み、前記第1のベース選択ページコンテンツは、前記第1のフィールド・コードに置き換えられる、請求項1に記載の方法。

【請求項 9】

第1の複製された選択ページコンテンツのテキスト文字列を取り出す前記工程は、前記書式構造に従って第2のテキスト文字列を配置する工程を含み、前記第1の複製された選択ページコンテンツは、第1のフィールド・コードに置き換えられる、請求項7に記載の方法。

【請求項 10】

前記書式構造に従って配置されるように前記第1のベース選択ページコンテンツの前記テキスト文字列のプレビューを表示する工程をさらに備える、請求項7に記載の方法。

【請求項 11】

前記第1の領域が方形である、請求項1に記載の方法。

【請求項 12】

前記文書の前記複数のページのうちの第1のページ内の第2の領域の選択を受け取る工程であって、前記第2の領域は、前記複数のページのうちの前記第1のページに対して第2の境界セットによって画定され、第2のベース選択ページコンテンツを含む、受け取る工程と、

前記複数のページのうちの前記第1のページから前記第2のベース選択ページコンテンツのテキスト文字列を取り出す工程と、

前記第2のベース選択ページコンテンツの取り出された前記テキスト文字列を前記複数のページのうちの前記第1のページのページ・ロケーション・インデックスに割り当てる工程と、

前記複数のページのうちの第2のページから第2の複製された選択ページコンテンツのテキスト文字列をユーザ介入なく取り出す工程であって、前記第2の複製された選択ページコンテンツは、前記複数のページのうちの前記第2のページに対して前記第2の境界セットによって画定された同じ前記第2の領域内に含まれる、取り出す工程と、

前記第2の複製された選択ページコンテンツの取り出された前記テキスト文字列を前記

複数のページのうちの前記第2のページのページ・ロケーション・インデックスに割り当てる工程と、をさらに備える、請求項1に記載の方法。

【請求項13】

前記第1の領域に関連付けられている第1のフィールド・コード、前記第2の領域に関連付けられている第2のフィールド・コード、および少なくとも1つのフィールド区切り文字を含む書式構造を受け取る工程をさらに備え、

第1のベース選択ページコンテンツのテキスト文字列を取り出す前記工程は、前記書式構造に従って、前記第1のベース選択ページコンテンツおよび前記第2のベース選択ページコンテンツの前記テキスト文字列ならびに前記少なくとも1つのフィールド区切り文字を連結する工程を含み、前記第1のベース選択ページコンテンツは、前記第1のフィールド・コードに置き換えられ、前記第2のベース選択ページコンテンツは、前記第2のフィールド・コードに置き換えられる、請求項12に記載の方法。

【請求項14】

前記書式構造に従って配置されるように前記第1のベース選択ページコンテンツおよび前記第2のベース選択ページコンテンツの前記テキスト文字列のプレビューを表示する工程をさらに備える、請求項13に記載の方法。

【請求項15】

前記フィールド区切り文字はスペース文字である、請求項13に記載の方法。

【請求項16】

第1のベース選択ページコンテンツのテキスト文字列を取り出す前記工程は、

前記第1のベース選択ページコンテンツを光学式文字認識(OCR)モジュールにより構文解析し、前記第1のベース選択ページコンテンツの前記テキスト文字列を生成する工程をさらに含み、

第1の複製された選択ページコンテンツのテキスト文字列を取り出す前記工程は、

前記第1の複製された選択ページコンテンツを前記OCRモジュールにより構文解析し、前記第1の複製された選択ページコンテンツの前記テキスト文字列を生成する工程をさらに含み、

第2のベース選択ページコンテンツのテキスト文字列を取り出す前記工程は、

前記第2のベース選択ページコンテンツを光学式文字認識(OCR)モジュールにより構文解析し、前記第2のベース選択ページコンテンツの前記テキスト文字列を生成する工程をさらに含み、

第2の複製された選択ページコンテンツのテキスト文字列を取り出す前記工程は、

前記第2の複製された選択ページコンテンツを前記OCRモジュールにより構文解析し、前記第2の複製された選択ページコンテンツの前記テキスト文字列を生成する工程をさらに含む、請求項12に記載の方法。

【請求項17】

文書内の複数のページにラベルを自動的に付ける方法であって、

前記複数のページのうちの第1のページ上の1以上のページ領域選択を受け取る工程であって、前記ページ領域選択の各々は、ラベルコンテンツを中心に含む、受け取る工程と、

前記1以上のページ領域選択の各々に基づいて、複数の選択されたページからの1以上のラベルコンテンツを、テキスト・データとしてユーザ介入なく抽出する工程と、

対応する前記ラベルコンテンツの前記テキスト・データを、特定のラベルコンテンツが抽出された前記ページに対するラベルとして割り当てる工程と、を備える、方法。

【請求項18】

前記1以上のラベルコンテンツが抽出された、選択された前記ページの範囲の指定を受け取る工程をさらに備える、請求項17に記載の方法。

【請求項19】

1以上のラベルコンテンツを抽出する前記工程は、前記ページ領域選択をOCRエンジンにより構文解析し、前記対応するテキスト・データを生成する工程を含む、請求項17に記載の方法。

【請求項 2 0】

文書内の複数のページに対するブックマークを自動的に生成する方法であって、

前記複数のページのうちの第1のページ上の1以上のページロケーション選択を受け取る工程であって、前記ページロケーション選択の各々は、ブックマークコンテンツを中に含む、受け取る工程と、

前記1以上のページロケーション選択の各々に基づいて、複数の選択されたページからの1以上のブックマークコンテンツを、テキスト・データとしてユーザ介入なく抽出する工程と、

対応する前記ブックマークコンテンツの前記テキスト・データで、特定のラベルコンテンツが抽出された前記ページに対するブックマークを生成する工程と、を備える、方法。

【請求項 2 1】

前記1以上のブックマークコンテンツが抽出された、選択された前記ページの範囲の指定を受け取る工程をさらに備える、請求項20に記載の方法。

【請求項 2 2】

1以上のブックマークコンテンツを抽出する前記工程は、前記ページロケーション選択をOCRエンジンにより構文解析し、前記対応するテキスト・データを生成する工程を含む、請求項20に記載の方法。