

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第5795724号  
(P5795724)

(45) 発行日 平成27年10月14日(2015.10.14)

(24) 登録日 平成27年8月21日(2015.8.21)

(51) Int.Cl.

G06F 17/22 (2006.01)

F 1

G06F 17/22 611

請求項の数 5 (全 21 頁)

(21) 出願番号 特願2011-127703 (P2011-127703)  
 (22) 出願日 平成23年6月7日 (2011.6.7)  
 (65) 公開番号 特開2012-3753 (P2012-3753A)  
 (43) 公開日 平成24年1月5日 (2012.1.5)  
 審査請求日 平成26年6月5日 (2014.6.5)  
 (31) 優先権主張番号 12/819,656  
 (32) 優先日 平成22年6月21日 (2010.6.21)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 596170170  
 ゼロックス コーポレイション  
 XEROX CORPORATION  
 アメリカ合衆国、コネチカット州 068  
 56、ノーウォーク、ピーオーボックス  
 4505、グローバー・アヴェニュー 4  
 5  
 (73) 特許権者 502096543  
 パロ・アルト・リサーチ・センター・イン  
 コーポレーテッド  
 Palo Alto Research  
 Center Incorporated  
 アメリカ合衆国、カリフォルニア州 94  
 304、パロ・アルト、コヨーテ・ヒル・  
 ロード 3333

最終頁に続く

(54) 【発明の名称】注釈付き文書の画像からクリーンな文書を再構築するためのシステムと方法

## (57) 【特許請求の範囲】

## 【請求項 1】

注釈付き文書画像から文書を再構築するためのコンピュータ実装方法であって、少なくとも2つの注釈付き文書画像より成るセットをコンピュータメモリ内へ受信することであって、前記注釈付き文書画像は各々静止エレメントと注釈付きデータとを備えることと、

前記セットの前記少なくとも2つの注釈付き文書画像の各々のスキュームの測度、  
前記セットの前記少なくとも2つの注釈付き文書画像の各々の鮮明さの測度、及び  
前記セットの前記少なくとも2つの注釈付き文書画像の各々のライン品質の測度

に基づいて前記注釈付き文書画像のセットから代表画像を選択することと、

前記セットの前記少なくとも2つの注釈付き文書画像の各々について、前記選択された代表画像に対するアライメントを実行することと、

前記アライメントが実行された前記セットの前記少なくとも2つの注釈付き文書画像の各々に基づいて、前記アライメントが実行された前記セットの前記少なくとも2つの注釈付き文書画像の各々から導出された前記注釈付きデータのうちの少なくとも幾つかの前記注釈付きデータが曖昧にされる合意文書画像を形成することと、

前記合意文書画像を出力することと、

を含む方法。

## 【請求項 2】

注釈付き文書画像から文書を再構築するためのコンピュータ実装方法であって、

10

20

少なくとも 2 つの注釈付き文書画像より成るセットをコンピュータメモリ内へ受信することであって、前記注釈付き文書画像は各々静止エレメントと注釈付きデータとを備えることと、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々について、前記注釈付き文書画像のセットにおける少なくとも 1 つの他の注釈付き文書画像に対するアラインメントを実行することと、

前記アラインメントが実行された前記セットの前記少なくとも 2 つの注釈付き文書画像の各々に基づいて、前記アラインメントが実行された前記セットの前記少なくとも 2 つの注釈付き文書画像の各々から導出された前記注釈付きデータのうちの少なくとも幾つかの前記注釈付きデータが曖昧にされる合意文書画像を形成することと、

前記合意文書画像に基づいて、クリーンな文書を形成することと、

前記注釈付き文書画像のセットから注釈情報データを、前記セットの前記少なくとも 2 つの注釈付き文書画像の各々を前記クリーンな文書と比較することによって抽出することと、

を備え、

前記注釈情報データを抽出することは、

前記クリーンな文書内のバウンディングボックス及び前記セットの前記少なくとも 2 つの注釈付き文書画像の各々内の対応するバウンディングボックスを識別することと、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々の各バウンディングボックス内に含まれるが、前記クリーンな文書の前記対応するバウンディングボックスには存在していない注釈ピクセルを抽出するために異なる演算を実行することと、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々及び前記クリーンな文書の各バウンディングボックスについて、サイズ及びロケーション、前記抽出された注釈ピクセル、及び前記注釈付きデータが導出された前記注釈付き文書画像を識別する識別子を出力することと、

を含む、

方法。

### 【請求項 3】

注釈付き文書画像から文書を再構築するためのコンピュータ実装方法であって、

少なくとも 2 つの注釈付き文書画像より成るセットをコンピュータメモリ内へ受信することであって、前記注釈付き文書画像は各々静止エレメントと注釈付きデータとを備えることと、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々のスキュームの測度、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々の鮮明さの測度、及び

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々のライン品質の測度

に基づいて前記注釈付き文書画像のセットから代表画像を選択することと、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々及び前記選択された代表画像において、前記セットの前記少なくとも 2 つの注釈付き文書画像の各々内の幾つかのブロックを選択し、前記注釈付き文書画像内の前記ブロックとサイズ及びロケーションが一致する、前記代表画像における同数のブロックを識別し、かつ前記セットの前記少なくとも 2 つの注釈付き文書画像の各々の前記ブロック内の少なくとも 1 点を前記代表画像内の対応する前記ブロック内の少なくとも 1 点へマップする変換マトリクスを計算することによって、前記セットの前記少なくとも 2 つの注釈付き文書画像の各々についてグローバルアラインメントを実行することと、

前記選択された代表画像と、前記注釈付き文書画像のセットと、に基づいて、前記グローバルアラインメントが実行された前記セットの前記少なくとも 2 つの注釈付き文書画像の各々から導出された前記注釈付きデータのうちの少なくとも幾つかの前記注釈付きデータが曖昧にされる合意文書画像を形成することと、

前記合意文書画像を出力することと、

を含む方法。

10

20

40

50

**【請求項 4】**

注釈付き文書画像からクリーンな文書を再構築するためのコンピュータベースシステムであって、

コンピュータプロセッサと、

代表スキャン選択モジュール、グローバルアラインメントモジュール、合意画像形成モジュール、及びクリーンアップモジュールを格納するコンピュータメモリと、  
を備え、

前記代表スキャン選択モジュールは、

少なくとも 2 つの注釈付き文書画像のセットを受信し、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々のスキーの測度、

10

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々の鮮明さの測度、及び

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々のライン品質の測度

に基づいて前記注釈付き文書画像のセットから代表画像を選択するように動作可能であり

、前記グローバルアラインメントモジュールは、

前記注釈付き文書画像のセットについて、前記選択された代表画像に対して、グローバルアラインメントを実行するように動作可能であり、

前記グローバルアラインメントを実行することは、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々内で幾つかのブロックを選択することと、

20

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々内の前記ブロックとサイズ及びロケーションが一致する前記代表画像における同数のブロックを選択することと、

前記セットの前記少なくとも 2 つの注釈付き文書画像の各々の各ブロック内の少なくとも 1 点を前記代表画像内の対応するブロック内の少なくとも 1 点へマップする変換マトリクスを計算することと、

を含み、

前記合意画像形成モジュールは、

少なくとも前記グローバルアラインメントが実行された前記セットの前記少なくとも 2 つの注釈付き文書画像の各々に基づいて、合意文書画像を形成するように動作可能であり、かつ、

30

前記クリーンアップモジュールは、

少なくとも前記合意文書画像に基づいて、クリーンな文書を形成し、かつ、

前記クリーンな文書を出力するように動作可能であり、

前記コンピュータプロセッサは、前記代表スキャン選択モジュール、前記グローバルアラインメントモジュール、前記合意画像形成モジュール、及び前記クリーンアップモジュールを実装する

コンピュータベースシステム。

**【請求項 5】**

前記合意文書画像に基づいて、クリーンな文書を形成することと、

前記クリーンな文書を出力することと、

40

を更に含む請求項 1 に記載の方法。

**【発明の詳細な説明】****【技術分野】****【0001】**

例示的な実施形態は、注釈付き文書の画像セットからクリーンな文書を再構築するためのシステムと方法に関する。

**【背景技術】****【0002】**

所定の形式が幾名かのユーザによって完成され、かつ概して形式のスキャン及びユーザによる追加の識別を含む処理に提出されることは一般的である。記入済み（注釈付き）書

50

式の固定したテキストから可変テキスト（例えば、氏名、住所、日付け、金額、他）を分離することは、文書走査サービスの分野では困難である可能性がある。ユーザが情報を入力した書式フィールドのロケーションが既知であれば、分離は比較的容易である。しかしながら、過去の何れかの時点で生成されたものと思われる書式などは特に、この情報を入手できない場合が多い。

【発明の概要】

【発明が解決しようとする課題】

【0003】

従って、過去に全く目にしたことがない新しいタイプの書式に関しても、注釈付き文書の画像から注釈付きデータを分離するための、ほぼ完全に自動化された方法を保有することが望ましい。10

【課題を解決するための手段】

【0004】

例示的な実施形態の一態様では、注釈付き文書の画像から文書を再構築するためのコンピュータ実装方法が提供される。本方法は、少なくとも2つの注釈付き文書画像のセットをコンピュータメモリ内へ受信することであって、注釈付き文書画像は各々静止エレメントと注釈付きデータとを備えることと、注釈付き文書画像の各々について、上記注釈付き文書画像セットにおける少なくとも1つの他の注釈付き文書画像に対するアライメントを実行することを含む。次に、アラインされた注釈付き文書画像を基礎として、上記アラインされた注釈付き文書画像から導出された少なくとも幾つかの注釈付きデータが曖昧にされた合意文書画像が形成される。クリーンな文書は、この合意文書画像を基礎として形成され、かつクリーンな文書は出力デバイスを介するビューリング用に表示される。20

【0005】

別の態様では、注釈付き文書の画像からクリーンな文書を再構築するためのコンピュータベースシステムが提供される。本システムは、コンピュータプロセッサと、グローバルアライメントモジュール、合意画像形成モジュール及びクリーンアップモジュールを格納するコンピュータメモリとを含む。グローバルアライメントモジュールは、注釈付き文書画像の各々について、選択された代表画像に対するグローバルアライメントを実行するように構成される。合意画像形成モジュールは、少なくともアラインされた注釈付き文書画像を基礎として、上記アラインされた注釈付き文書画像から導出された少なくとも幾つかの注釈付きデータが曖昧にされている合意文書画像を形成するように構成される。クリーンアップモジュールは、少なくともこの合意文書画像を基礎としてクリーンな文書を形成しつつクリーンな文書を出力デバイス上へビューリング用に表示するように構成される。コンピュータプロセッサは、これらのモジュールを実装する。30

【0006】

さらに別の態様において、コンピュータプログラムプロダクトは有形媒体符号化命令を含み、上記命令は、コンピュータ上で実行されると、注釈付き文書画像のセットについて、上記注釈付き文書画像セットから代表画像を選択することと、注釈付き文書画像セットにおける複数の他の画像の各々に対して、選択された代表画像に関するグローバルアライメントを実行してアラインされた注釈付き文書画像のセットを形成することと、アラインされた注釈付き文書画像のセットを基礎として合意文書画像を形成することと、合意文書画像を基礎としてクリーンな文書を形成することを含む方法を実行する。注釈付き文書画像の各々について、注釈付き画像がクリーンな画像と比較されて注釈が識別され、注釈が出力される。40

【0007】

本方法及びシステムは、注釈付き文書画像のセットから代表画像を選択し、かつ注釈付き文書画像セット内の各注釈付き文書画像について、選択された代表画像に対するグローバルアライメントを実行するように適合化されてもよい。

【0008】

本方法及びシステムは、選択された代表画像及び注釈付き文書画像のセットを基礎とし50

て合意文書画像を形成するように適合化されてもよい。

【0009】

本方法及びシステムのクリーンな文書は、全ての類似文書上に複写されるための静止エレメントを含む固定されたページレイアウトを有する文書であってもよい。

【0010】

本方法及びシステムは、少なくとも合意文書画像を基礎とし、ごま塩ノイズ及び隔絶された白黒ピクセルを除去しつつ破断したラインを修復することによってクリーンな文書を形成するように適合化されてもよい。

【0011】

本方法及びシステムは、注釈付き文書画像のセットから注釈情報データを、各注釈付き文書画像をクリーンな文書と比較することによって抽出するように適合化されてもよい。 10

【0012】

本方法及びシステムは、クリーンな文書内のバウンディングボックス及び注釈付き文書画像内の対応するバウンディングボックスを識別し、注釈付き文書画像の各ボックス内に含まれる、クリーンな文書の対応するボックスには存在していない注釈ピクセルを抽出するためにdiff演算を実行し、かつ各バウンディングボックスについて、サイズ及びロケーション、抽出された注釈ピクセル及び注釈データが抽出された注釈付き文書画像を識別する識別子を出力することによって、注釈情報を抽出するように適合化されてもよい。

【0013】

本方法及びシステムによって選択される代表画像は、注釈付き文書画像のスキュームの測度、注釈付き文書画像の鮮明さの測度及び注釈付き文書画像のライン品質の測度のうちの一つを基礎とするものであってもよい。 20

【0014】

本方法及びシステムの選択される代表画像及び注釈付き文書画像セットは、合意文書画像の形成前にぼかされてもよい。

【0015】

本方法及びシステムは、各注釈付き文書画像について、この注釈付き文書画像内の幾つかのブロックを選択し、代表画像において、上記注釈付き文書画像内のブロックとサイズ及びロケーションが一致する同数のブロックを識別し、かつ注釈付き文書画像の各ブロック内の少なくとも1点を代表画像内の対応するブロック内の少なくとも1点へマップする変換マトリクスを計算することによって、グローバルアラインメントを実行するように適合化されてもよい。 30

【0016】

本方法及びシステムの変換マトリクスは、注釈付き文書画像内の各ブロックに関する変換の平均であってもよい。

【0017】

本方法及びシステムは、注釈付き文書のブロック内の黒ピクセルと代表文書のブロック内の黒ピクセルとの重なりの測度を計算することによって、注釈付き文書画像の各ブロック内の少なくとも1点を代表画像内の対応するブロック内の少なくとも1点へマップするように適合化されてもよい。 40

【0018】

本方法及びシステムは、注釈付き文書画像のブロックと代表画像のブロックとの間の重なり合った黒ピクセルの数を、注釈付き文書画像のブロック及び代表画像の対応するブロックにおける黒ピクセルの平均数で除したものとして定義される品質スコアを算出することによって、黒ピクセルの重なりの測度を計算するように適合化されてもよい。

【0019】

本方法及びシステムは、変換マトリクスを計算するために粗い探索から細かい探索までの階層的探索を用いてもよい。

【0020】

本方法及びシステムの変換マトリクスは、回転、平行移動、均一なスケーリング、ペー 50

ジカールの補正及びキーストーニングの補正を考慮に入れてもよい。

**【0021】**

本方法及びシステムは、注釈付き文書画像のセットにおける各注釈付き文書画像について、上記注釈付き文書画像を二次ブロックのセットに分割し、注釈付き文書画像内の各二次ブロックと代表画像内の個々の二次ブロックとの一致を決定し、かつ一致が決定された複数の二次ブロックについて、グローバルアラインメント情報を基礎として、代表画像内の対応するブロックを取得するために注釈付き文書画像内の各二次ブロックへ平行移動ベクトルを割り当てるベクトルフィールドを生成することによって、注釈付き文書画像のセットについてローカルアラインメントを実行するように適合化されてもよい。

**【0022】**

本方法及びシステムのベクトルフィールドは、少なくとも1つの二次ブロックの平行移動ベクトルを推定するためにラプラス平滑化処理を用いてもよい。ラプラス平滑化処理は、少なくとも1つの二次ブロックに近接する二次ブロックからのベクトルを平均することによってベクトルを外挿してもよい。

**【0023】**

本方法及びシステムは、各々が合意文書画像内に一意の位置を有する複数の合意ピクセルを決定し、割り当てられた注釈付き文書画像の各々における対応するピクセルの特性を合計することを含み、かつ上記合計を所定のしきい値と比較することによって、合意文書画像を形成するように適合化されてもよい。さらに、上記合計は、代表画像内の対応するピクセルの加重値を含んでもよい。

**【図面の簡単な説明】**

**【0024】**

【図1】注釈データなしのクリーンな文書を示すグラフである。

【図2】注釈データが追加された図1のクリーンな文書を示すグラフである。

【図3】注釈データが追加された図1のクリーンな文書を示すグラフである。

【図4】注釈付き文書画像のセットからクリーンな文書を再構築するための例示的なシステムを示すハイレベル概観図である。

【図5】図4の例示的なシステムによって実行されるオペレーションを示すフロー図である。

**【図6】**図5に示すグローバルアラインメント方法をさらに詳しく示すフロー図である。

30

【図7】図5及び図6のグローバルアラインメント方法に使用される粗い探索から細かい探索までの階層的探索をさらに詳しく示すフロー図である。

【図8】図7に示す粗い探索から細かい探索までの階層的探索によって使用される探索ウインドウを示すグラフである。

【図9】図5によって説明されているローカルアラインメント方法をさらに詳しく示すフロー図である。

【図10】注釈付き文書画像へ印加されるべき平行移動ベクトルセットによる歪みマップを示すグラフである。

**【図11】**ノイズを含む合意画像を示す。

**【図12】**図2の注釈付き文書画像から抽出された注釈データを示す。

40

【図13】図3の注釈付き文書画像に関連して図6のグローバルアラインメント方法により選択された最初のブロックを示す。

【図14】図3の注釈付き文書画像に関連して図9のローカルアラインメント方法により使用された第2のブロックを示す。

**【発明を実施するための形態】**

**【0025】**

本明細書では、注釈付き文書画像のセットからクリーンな文書を再構築するための方法及びシステムを開示する。

**【0026】**

本明細書で使用しているように、「クリーンな文書」は、固定されたテキスト、固定さ

50

れた境界及び全ての類似文書において複写されるための他のページエレメント等の静止エレメントを含む固定されたページレイアウトを有する文書である。例えば、図1はクリーンな文書例2を示している。クリーンな文書2は、ユーザによって情報を記入される（注釈を付けられる）べき未記入の書式と見なすことができる。従って、注釈付き文書は、ユーザまたは他のエージェント（コンピュータシステム等）によって追加された可変データ（注釈）を含む修正されたクリーン文書である。注釈付き文書画像は、走査され、ダウンロードされかつ別途コンピュータメモリへ入力されている注釈付き文書からの1つまたは複数の画像である。図2及び図3は、注釈付き文書画像の例4、6を示す。

#### 【0027】

この例示的な実施形態では、文書画像に対して、回転、平行移動及びx、y座標の分離スケーリングを考慮に入れる多変数変換が実行される。これは、2つ以上の注釈付き文書画像間のグローバルアラインメントを達成する。複数文書のアラインメントの後、本例示的方法及びシステムは、複数の注釈付き文書画像に共通するピクセルを含む新しい画像を計算する。これは、最初のグローバル変換が単に1ページ画像の他のページ画像に対する平行移動であった早期のアラインメント方法とは対照的である。

10

#### 【0028】

図4を参照すると、注釈付き文書画像のセットからクリーンな文書2を再構築するための例示的なシステム100が示されている。システム100は、注釈付き文書画像のセット104を受信するための入力デバイス102を含む。本明細書に提示している本例示的システム及び方法の説明を単純にするために、注釈付き文書画像のセット104は単一のクリーンな文書2に対応する2つ以上の注釈付き文書画像から成ることを想定するが、本方法及びシステムは、複数のクリーン文書に対応する注釈付き文書画像の集合体の処理に使用されることも可能である。入力に先立って、注釈付き文書画像のセット104は、スキヤナによって生成されるROMまたはRAMドライブ等の任意の適切な有形媒体に格納されてもよく、かつ／または、例えばインターネットを介してシステム100へ搬送波の形式で入力されてもよい。或いは、注釈付き文書画像104は、システム100自体の内部で生成されてもよい。入力デバイス102は、モデムリンク、有線または無線接続、USBポート、フロッピーまたはハードディスクレシーバまたはこれらに類似するものを含んでもよく、かつシステム100の他のコンポーネントと結合されても、分離されてもよい。

20

#### 【0029】

システム100は、処理の間に注釈付き文書画像104を格納するためのデータメモリ106を含む。システム100のメインメモリ108は、代表スキャン選択モジュール110と、グローバルアラインメントモジュール112と、任意選択のローカルアラインメントモジュール114と、合意画像形成モジュール116と、クリーンアップモジュール118と、注釈抽出モジュール119とを格納する。モジュール110、112、114、116、118及び119からの出力はメモリ106、108に格納されてもよく、または出力デバイス120を介してクライアント端末130またはプリンタ132の一方またはそれ以上へ、任意選択としてインターネット等のネットワーク134を通じて出力されてもよい。出力（クリーンな文書2または抽出された注釈157等）は、ビューア用にクライアント端末130またはプリンタ132上で表示されてもよい。

30

#### 【0030】

任意選択の代表スキャン選択モジュール110は、入力として入力デバイス102を介して注釈付き文書画像のセット104を受信し、この注釈付き文書画像のセット104から注釈付き文書画像のセット104を代表する（「最良の」）ものと考えられる注釈付き文書画像150を選択する。選択された代表の注釈付き文書画像150は、次にスキャン選択モジュール110からグローバルアラインメントモジュール112へ送られる。グローバルアラインメントモジュール112は、他の注釈付き文書画像104の各々に対して剛体変換（回転及び／または平行移動等）を実行し、これらを、選択された代表の注釈付き文書画像150へさらに密にアラインする。（任意選択のステップであることから）代

40

50

表の注釈付き文書画像が選択されていなければ、グローバルアラインメントモジュール 1  
 12 は、注釈付き文書画像 104 の各々について、他のロケーション情報に関する剛体及び / または相対変換を実行する。例えば、グローバルアラインメントモジュール 112 は、各注釈付き文書画像が、画像の端に平行する、かつ上記端から定距離を有するテキストの一一番上の行（または文書内の所定の行）を有するように、各注釈付き文書画像をアラインしてもよい。グローバルアラインメントモジュール 112 の出力は、任意選択のローカルアラインメントモジュール 114 へ送られる変換マトリクス 152 等の変換セットである。ローカルアラインメントモジュール 114 はさらに、とりわけページカール、キーストーニング（斜角から撮られた写真等）、他を補正する技術を用いて、選択された代表文書画像 150 に関して注釈付き文書画像のセット 104 をアラインする。10  
 ローカルアラインメントモジュール 114 の出力は、合意画像形成モジュール 116 へ送られるアラインされた文書画像 154 のセットである。合意画像形成モジュール 116 はアラインされた各画像を処理し、かつ少なくとも複数のアラインされた画像に幾何学的に共通するピクセルを識別して合意画像 156 を形成する。共通するピクセルを有することをそのピクセルが合意画像 156 に属するとして識別される前に必要とされる注釈付き文書画像のしきい値数はシステム実装に依存するが、これについては後に詳述する。合意画像形成モジュール 116 が合意画像 156 を形成すると、合意画像 156 は任意選択としてクリーンアップモジュール 118 へ送られる。クリーンアップモジュール 118 は、合意画像 156 内のごま塩ノイズ等の望ましくないアーチファクトを識別し、かつクリーンな文書 2 をメモリ 106、108、クライアント端末 130 及びプリンタ 132 のうちの 1つまたはそれ以上へ出力する。20  
 注釈抽出モジュール 119 は、注釈付き文書画像をクリーンな文書 2 と比較することによって注釈付き文書画像 114 の各々から注釈情報を抽出する。

### 【0031】

この例示的なシステムでは、コンポーネント 110、112、114、116、118 及び 119 はメインメモリ 108 に格納されたソフトウェア命令を備え、これはコンピュータプロセッサ 122 によって実行される。コンピュータの CPU 等のプロセッサ 122 は、メモリ 108 に格納された処理命令を実行することによってコンピュータシステム 100 の全体的なオペレーションを制御してもよい。コンポーネント 102、106、108、110、112、114、116、118、119、120 及び 122 は、データ制御バス 124 によって接続されてもよい。認識されるであろうが、システム 100 は、同じ機能性を維持しながらより少ない、またはより多いコンポーネントを含んでもよい。例えば、コンポーネント 110、112、114、116、118、119 は、より少ないコンポーネントを形成するために結合されても、より多い個々のコンポーネントを形成するために機能を分離されてもよい。30

### 【0032】

認識されるであろうが、クリーン文書再構築システム 100 は、パーソナルコンピュータ、PDA、ラップトップコンピュータ、サーバコンピュータ、文書処理ケイパビリティを有するスキャンデバイスまたはこれらの組合せ等の 1つまたは複数のコンピューティングデバイスを備えてもよい。メモリ 106、108 は一体型であっても分離型であってもよく、ランダムアクセスメモリ (RAM)、読み取り専用メモリ (ROM)、磁気ディスクまたはテープ、光ディスク、フラッシュメモリまたはホログラフィックメモリ等の任意タイプのコンピュータ読み取り可能媒体を表してもよい。ある実施形態では、メモリ 106、108 は、ランダムアクセスメモリと読み取り専用メモリとの組合せを備える。実施形態によつては、プロセッサ 122 とメモリ 106 及び / または 108 とが 1つのチップ内に組み合わされてもよい。40

### 【0033】

図 5 を参照すると、注釈付き文書画像のセットからクリーンな文書を再構築するための例示的な方法が示されている。本方法は、図 4 に示すシステム 100 を使用してもよい。本例示的な方法が図示されているステップより少ない、多い、または異なるステップを含んでもよく、かつ図示されている順序で進行する必要もないことは認識されるべきである。50

図5に示されている方法は、コンピュータ上で実行され得るコンピュータプログラムプロダクトにおいて実装されてもよい。コンピュータプログラムプロダクトは、コンピュータプログラムが記録される（ディスクまたは他のメモリ格納デバイス等の）有形コンピュータ読み取り可能記録媒体であっても、制御プログラムがデータ信号として具現される伝送可能な搬送波であってもよい。図示されている方法は、完全自動式であってもよく、本明細書に記載しているような何らかのユーザ入力を含んでもよい。

#### 【0034】

本方法は、ステップS100で開始される。ステップS200において、クリーン文書再構築システム100は、入力デバイス102を介してコンピュータメモリ106へ少なくとも2つの注釈付き文書画像のセット104を受信する。セット104内の各注釈付き文書画像は同じ形式またはクリーンな文書画像に対応しているべきであって、唯一異なる注釈を有する。例えば、図2及び図3の注釈付き文書画像4及び6は、画像4及び6が各々10a及び10b等の注釈を含むクリーン文書2の上位セットであることから、共に図1のクリーンな文書画像（即ち、未記入の書式）2に対応している。ある実施形態では、2つの注釈付き文書画像は、これらの2つの注釈付き文書画像が少なくとも所定の比率（例えば、50%）のピクセルを共有していれば同じクリーンな文書画像に対応している。例えば、2つの注釈付き文書画像は、画像の大きさに関して同じ位置に存在する同じ改行及び同じフォントを有する所定量のテキストを共有してもよい。出力されるクリーンな文書画像の品質は、少なくとも部分的に、入力される注釈付き文書画像の数に依存する。本例示的方法及びシステムでは少なくとも10、かつある実施形態では約20またはそれ以上の白黒、グレースケール及び/またはカラー画像がシステム100へ入力される。注釈付き文書画像は、TIFF、JPEG、GIF、PDF、他等の処理ソフトウェアによる認識が可能な任意のフォーマットであってもよい。任意選択として、グレースケールまたはカラー画像は、処理の前に2値（白黒）画像に変換されてもよい。

#### 【0035】

ステップS300では、代表スキャン選択モジュール110が入力された注釈付き文書画像のセット104から1つの注釈付き文書画像を選択し、選択された注釈付き文書画像を代表の注釈付き文書画像150として設定する。代表の注釈付き文書画像150は、入力された注釈付き文書画像のセット104から外されてもよいが、この限りではない。ある代替実施形態では、代表の注釈付き文書画像150はユーザによって手動で選択されてもよい。ある注釈付き文書画像が代表の注釈付き文書画像として選択されることに適するかどうかを決定し得る要素としては、上記注釈付き文書画像のスクュー、鮮明さ及びライン品質が含まれるが、この限りではない。例えば、ある注釈付き文書画像は、それが文書の余白に対して歪められておらず、鮮明でくっきりとしたエッジを有しあつ濃くて容易に区別可能な行を有していれば、代表の注釈付き文書画像と見なされてもよい。代表の注釈付き文書画像150の決定に際しては、これらの要素の任意のもの、または全てが考慮されてもよい。その目的は、最良の注釈付き画像（生成元であったクリーンな画像と構造が最もよく類似していると予想されたもの）を代表の注釈付き文書画像として選択することにあるが、処理ソフトウェアは必ずしも人が最良であると考える注釈付き画像を選択しない場合がある。本例示的実施形態では、代表の注釈付き文書画像の決定に際して、文書画像の鮮明さ、または品質よりも文書画像のスクューの方に重きがおかれる。説明を目的として、ここでは図2を最良の注釈付き画像として選択する。

#### 【0036】

ステップS350において、代表スキャン選択モジュール110（または他の任意の適切なモジュール）は、任意選択として入力された注釈付き文書画像のセット104及び選択された代表の注釈付き文書画像150をぼかす。使用されるぼかしアルゴリズム及び/またはコンポーネントは、任意の適切なぼかしアルゴリズムを実装してもよい。例えば、所定の実施形態は、周知のガウスぼかしアルゴリズムを実装して注釈付き文書画像104、150をぼかす。他の実施形態では、文書画像の黒ピクセルは膨張される。膨張は、出力されるピクセルの値が入力されるピクセルの近傍における全ピクセルの最大値である才

10

20

30

40

50

ペレーションを示す。例えば、2値画像において、入力されるピクセルの近傍における任意のピクセルが黒であれば、出力されるピクセルは黒に設定される。ばかしの目的は、1つまたは2つのピクセルによるピリオドまたはライン等の比較的小さいオブジェクトであっても幅はアライン可能であることの保証にある。

#### 【0037】

ステップS400において、グローバルアラインメントモジュール112は、入力された注釈付き文書画像のセット104内の各注釈付き文書画像 $A_1, A_2, \dots, A_n$ を、選択された代表の注釈付き文書画像B150に対してアラインするグローバルアラインメントを実行する。簡単に言えば、グローバルアラインメントモジュール112は、 $A_i$ 内の幾つかの初プロック478(図13)(例えば、本例示的実施形態では約20個-50個のプロック)を選択し、相互相關及び粗い探索から細かい探索までの階層的探索を用いてB内のその対応する(即ち、一致する)初プロック478を識別する。本例示的実施形態では、文書画像 $A_i$ 内の選択された初プロック478は、多くの黒(またはカラーの場合は暗い)ピクセルを含む複合プロックである。例えば、ある初プロック478は、それが10%を超える黒ピクセルを含んでいれば選択されてもよい。初プロック478は、少なくとも50ピクセルまたは少なくとも100ピクセルを含んでもよく、かつ実施形態によっては、約100,000ピクセルまで含んでもよい。選択された初プロックは全て同じサイズ及び形状であってもよく、異なる形状及び/またはサイズであってもよい。ある実施形態では、初プロック478は長方形であるが、他の形状も企図されている。概して、初プロック478は画像面積の約5%以下の面積を占める。

10

#### 【0038】

初プロック478(図13)の一一致する各ペアは、1対の点、即ち $A_i$ 内の( $x, y$ )及びB内の( $x', y'$ )を与える。一般的な最小二乗法(不良一致を除去する異常値排除を有する)を用いて、グローバルアラインメントモジュール112は、 $A_i$ 内の点をB内の点に近似的にマップする変換マトリクスを計算する。変換マトリクスは、回転及び平行移動等の剛体運動のみを許容してもよい。所定の実施形態では、変換マトリクスは(ファックスマシンで発生することが多い) $x$ 及び $y$ 方向に一様なスケーリングまたは異なるスケーリングも許容してもよく、かつ/またはページカール、キーストーニング、他を含む、但しこれらに限定されない変形部を補正するための追加的な変換も提供してもよい。ステップS400は、図6にさらに詳述されている。グローバルアラインメント法はステップS350から開始されかつS402で始まり、各注釈付き文書画像 $A_i$ について、選択された代表の注釈付き文書画像B150に関して実行される。

20

30

#### 【0039】

ステップS402において、グローバルアラインメントモジュール112は、画像 $A_i$ 内の複数の初プロックを選択する。例えば、図13は、選択された複数の初プロック478を有する図3の注釈付き文書画像6を示している。本例示的実施形態では、30個以上の初プロック478が選択されている。さらに所定の実施形態では、初プロック478は、文書画像 $A_i$ に渡って一様のサイズ、非重複及び一様分布のうちの少なくとも何れかである。初プロック478は互いに隣接していても、任意サイズのスペースによって分離されてもよい。300DPI文書画像の場合、初プロック478は、典型的には $128 \times 128$ ピクセルのサイズを有する。初プロックは、ランダムに選択されても、擬似ランダムに選択されても、ユーザによって選択されてもよいが、この限りではない。初プロックが擬似ランダムに選択される実施形態では、初プロックはユーザの手助けによって選択されてもよく、かつ/または所定の条件を満たすように選択されてもよい。

40

#### 【0040】

ステップS404において、グローバルアラインメントモジュール112は、文書画像 $A_i$ の初プロック478(図13)の各々について、文書画像B内の対応する初プロック478に対する最近一致を決定する。まず、 $A_i$ の初プロックの同じ $x$ 及び $y$ 座標を用いてB内の初プロックを位置決めすることによって、 $A_i$ の初プロック478が暫定的にBの初プロック478に一致される。文書画像の初プロックの最近一致は、有効変換の下で

50

文書画像  $A_i$  の初ブロック 478 に最も近接して似ている選択された文書画像  $B$  の初ブロックを発見することによって決定される。有効変換には、文書画像の初ブロックの平行移動、文書画像の初ブロックの回転、文書画像の初ブロックのスケーリング及び文書画像の初ブロックのアフィン変換が含まれるが、この限りではない。文書画像の初ブロック 478 の最近一致は、上記一致の品質に対応する品質スコアにも関連づけられてもよい。所定の実施形態では、選択された文書画像の初ブロックの品質はゼロから 1 までのスケールで評価される。但し、ゼロは一致が発見されないことを示し、1 は完全一致を示す。完全一致は、選択された文書画像の何らかの初ブロックが、回転等の有効変換下で文書画像の初ブロックに正確に一致する場合にのみ存在する。このような実施形態下では、品質スコアは、(適用された任意の変換に続く) 文書画像の初ブロックと選択された文書画像の初ブロックとの間の重なり合う黒ピクセル数を、文書画像の初ブロック及びこれに対応する選択された文書画像の初ブロックにおける黒ピクセルの平均数で除したものとして定義されてもよい。さらに、所定の実施形態において、最近一致の品質スコアが信頼度しきい値より低ければ、方法 S400 の残りのステップに関して文書画像の対応する初ブロックは無視されてもよい(一致なし)。文書画像の対応する初ブロックが注釈データのみ(または主として注釈データ)を含んでいれば、品質スコアは信頼度しきい値より低くなると思われる。10

#### 【0041】

文書画像の初ブロック 478 (図 13) の最近一致を発見するためには、粗い探索から細かい探索までの階層的探索手順が使用されてもよい。図 7 を参照すると、図 8 の例示的な  $15 \times 15$  ピクセルの探索ウィンドウ 453 へ適用されるような階層的探索手順 450 が示されている。例示を目的として、探索ウィンドウ 453 は、図 2 及び図 3 における初ブロック 478 内にも示されている。階層的探索手順は、ステップ S452 において、文書画像  $A_i$  の初ブロックを中心として探索ウィンドウ 453 を画定することによって始まる。探索ウィンドウは、 $A_i$  の選択された初ブロックに対応する選択された代表画像  $B$  内のブロックを発見するために使用される。近似ロケーションを決定するより高度な方法が使用されてもよいが、所定の実施形態では、文書画像の初ブロックのロケーションは単に、選択された文書内の対応するロケーションにマップされる。例えば、幅  $w_0$  及び高さ  $h_0$  を有する文書画像内のロケーション ( $x_0, y_0$ ) における初ブロックについて考察されたい。選択された文書画像が幅  $w_1$  及び高さ  $h_1$  を有していれば、選択された文書内の対応するロケーション ( $x_1, y_1$ ) は下記のように定義されることが可能である。2030

#### 【数 1】

$$(1) \quad x_1 = \frac{w_1}{w_0} * x_0$$

40

#### 【数 2】

$$(2) \quad y_1 = \frac{h_1}{h_0} * y_0$$

**【0042】**

ステップS454において、グローバルアラインメントモジュール112は、探索ウィンドウ453より小さくかつ探索ウィンドウ内に中心を置かれた第1のボックス（点457で画定されている）を画定する。

**【0043】**

ステップS456において、グローバルアラインメントモジュール112は、第1のボックス内に位置決めされた点457のセットの一致スコアを決定する。所定の実施形態では、一致スコアは、4つ角と、各辺の中点と、ボックスの中心との離隔された9点について決定されるが、離隔された点の数はこれより少なく、または多く選択されてもよい。ある点の一致スコアは、文書画像A<sub>i</sub>の初ブロックを上記点に近接する選択された文書画像Bの領域へ一致させるべく試行することによって決定される。オペレータは、オペレータによる近接の定義が広義であるほど階層的探索手順450は遅くなるという理解の下で、近接を自らが適当と思うように定義してもよい。よってステップS458では、最も高い一致スコア（この図では、右上の点457）を有する点が選択される。例えば、図2及び図3の探索ウィンドウ453は各々ドル記号（\$）を含むが、これは、対応する点に関して高い一致スコアを与える。

10

**【0044】**

ステップS460では、先のボックスより小さくかつ先のボックスからの最近一致の辺りに中心を置かれたボックスが（点463によって）画定される。

20

**【0045】**

ステップS462において、グローバルアラインメントモジュール112は、より小さいボックス内に位置決めされた点について、第1のボックスに類似する方法で一致スコアを決定する。ステップS464では、最も高い一致スコアを有する点（この場合は、右中央の点463）が選択される。

**【0046】**

ステップS466において、グローバルアラインメントモジュール112は、より小さいボックスの点が類似する一致スコアを有するかどうかを決定する。例えば、点463の一致スコアは互いに比較され、点間の差が所定のしきい値内（例えば、0.0から1.0まで）に含まれるかどうかが決定される。より小さいボックスの点が互いに類似していないければ、ステップS460、S462、S464及びS466が繰り返される。

30

**【0047】**

ステップS468において、グローバルアラインメントモジュール112は、最近一致が探索ウィンドウ453の境界においてあるかどうかを決定する。例えば、点471で表される右側の点のうちの何れかがそのボックスの最近一致であると決定されれば、選択される点は探索ウィンドウの境界に存在する。最近一致が境界上に存在しなければ、方法450はS472へ進む。存在すれば、方法450は、ステップS470において最近一致の辺りに中心が置かれる新しい探索ウィンドウを画定し、ステップS454からS468までを繰り返す。本方法におけるこの時点で、グローバルアラインメントモジュール112は、文書画像A<sub>i</sub>内に、文書画像B内の点セットに対応する9点までのセットを識別している。

40

**【0048】**

図6に戻って参照すると、ステップS406において、グローバルアラインメントモジュール112は、A<sub>i</sub>へ適用されると文書画像A<sub>i</sub>を文書画像Bへアラインする文書画像A<sub>i</sub>の変換マトリクスを決定する。この決定は、最近一致間の差を最小化するための最小二乗分析及び文書画像の対応する初ブロック478（図13）を用いて行われる。基本的には、その目的は、文書画像A<sub>i</sub>の初ブロック内の点を選択された文書画像Bの対応する初ブロック内の点へマップすることにある。これを実行する際には、最小二乗分析は1つまたは複数の有効変換を考慮してもよい。但し、有効変換には、回転、平行移動、スケーリング、ページカール、キーストーニング、他が含まれるが、この限りではない。所定の

50

実施形態では、最小二乗分析は、誤一致の点ペアに対するプロセス感度を下げるために異常値排除も含む場合がある。

#### 【0049】

例証のために、選択された文書画像 B 内の初ブロック  $r_s$  のロケーションを、文書画像  $A_i$  内の対応する初ブロック  $r_d$  のロケーションの関数  $f$  としてモデリングできるものとする。さらに、関数  $f$  は、関数  $f$  が 2 つの初ブロック間でどのようにマッピングするかを制御する可変パラメータのベクトル  $\theta$  をとるものとする。パラメータには、回転、スケーリング、平行移動、他の量が含まれてもよい。よって、誤差  $S$  の二乗和は、

#### 【数 3】

$$S = \sum_{i=1}^n d((x_i, y_i), f(x_i, y_i))^2, \quad (3)$$

10

と定義されることが可能である。但し、 $n$  はステップ S 4 0 2 で選択された初ブロックの数に一致し、 $f(\cdot)$  は新しい点となる点  $(x_i, y_i)$  への変換を適用する関数であり、 $d$  は点  $(x_i, y_i)$  と平行移動点  $f(x_i, y_i)$  との距離の測度であり、 $i$  は選択された文書画像 B (即ち、最近一致) 内の  $i$  番目の初ブロックに対応し、かつ  $f(x_i, y_i)$  は文書画像  $A_i$  内の  $i$  番目の初ブロックのロケーションに対応する。

20

#### 【0050】

二乗和  $S$  が算定されると、文書画像  $A_i$  の変換は、ベクトル  $\theta$  の可変パラメータを解くことによって決定される。即ち、ベクトル  $\theta$  は  $m$  個のパラメータを有し、かつ  $\theta_j$  は  $j$  番目のパラメータに対応するものとする。 $j$  番目のパラメータに対する解は、一次方程式等の従来技術を用いてパラメータ  $\theta_j$  を解くことにより決定されることが可能である。

#### 【0051】

30

認識されるべきものであるように、このようにして決定される可変パラメータは、選択された文書の初ブロック 478 (図 13) への文書画像の初ブロックのマッピングに付随する誤差の二乗和を最小化する。変換マトリクスが文書画像全体に渡って剛体的に適用される傾向があることを想定すれば、可変パラメータは、文書画像の変換をグローバルに算定するために用いられることが可能である。このグローバルに算定される変換は、文書画像  $A_i$  について計算される変換マトリクスの平均であってもよい。従って、ステップ S 4 0 6 の終わりでは、グローバルアラインメントモジュール 112 は文書画像  $A_i$  の各々について 1 つの変換マトリクス 152 を決定している。

#### 【0052】

40

再度図 5 を参照すると、ステップ S 5 0 0 において、ローカルアラインメントモジュール 114 は、任意選択として、入力された注釈付き文書画像のセット 104 内の各注釈付き文書画像  $A_1, A_2, \dots, A_n$  について、選択された代表の注釈付き文書画像 B 150 に関するさらなるローカルアラインメントを実行する。このステップは、グローバルなアラインメント変換は複数の初ブロック 478 (図 13) に渡る平均変換である、という事実を明らかにする手助けをする。文書画像  $A_i$  の各領域は、全て個々に計算されて文書画像  $A_i$  へ適用されるその固有の、僅かに異なる変換を有する。簡単に言えば、文書画像  $A_i$  は、 $A_i$  を非重複の (または部分的にのみ重なる) 二次ブロック 480 (図 14) のグリッドに分割することによって B へ局所的にアラインされる。例えば、 $128 \times 128$  ピクセルは、選択された代表の文書画像 B における対応ブロックが決定されている  $A_i$  内の各二次ブロックの妥当な二次ブロックサイズマッチである。これらのマッチングは、

50

B 内の対応するブロックを取得するために A<sub>i</sub> 内のブロックの中心へ平行移動ベクトルを割り当てるベクトルフィールドとを考えることができる。初期ベクトルは、グローバルアライメントステップ (S 4 0 0) において計算される変換マトリクスによって決定されるが、最終ベクトルは x 及び y の各々において初期ベクトルから約 5 ピクセルまで偏向されてもよい。確信的には一致され得ない二次ブロックのベクトルは、ラプラス平滑化処理を用いてその近傍から外挿されることが可能であり、この場合、上記ベクトルはその（通常は）4 つの近傍ベクトルの平均に設定される。画像 A<sub>i</sub> 内の任意の点 (x, y) における平行移動ベクトルを発見するために、(x, y) を包囲する 4 つのグリッド点における平行移動ベクトルからの双線形補間（即ち、2 点間の変換推定の発見）が用いられる。適用される変換ベクトルのセットにアラインされた画像は、A'<sub>i</sub> で示される。ステップ S 5 0 0 は、図 9 において詳述されている。任意選択であるローカルアライメント法 S 5 0 0 はステップ S 4 0 0 の後に開始され、S 5 5 2 で始まる。方法 S 5 0 0 は、注釈付き文書画像 A<sub>i</sub> の各々について、かつ二次ブロックの各々について、選択された代表の注釈付き文書画像 B 1 5 0 に関して実行される。10

#### 【0053】

ステップ S 5 5 2 において、ローカルアライメントモジュール 1 1 4 は文書画像 A<sub>i</sub> を、上記文書画像全体を包含する二次ブロック 4 8 0（図 14）のセットに分割する。但し、各二次ブロック 4 8 0 は上下左右（境界上の二次ブロックを除く）に近傍ブロックを有する。認識されるべきものであるように、二次ブロック 4 8 0 は文書画像の単なる一領域である。所定の実施形態では、二次ブロック 4 8 0 は、非重複的、二次ブロック 4 8 0 間の間隙を防止するために少し（例えば、10 ピクセル）だけ重複的、かつ一様サイズのうちの少なくとも 1 つである。ローカルアライメントモジュール 1 1 4 またはユーザは、二次ブロック 4 8 0 のサイズを、任意選択として少なくとも下記の考慮事項、即ち 1) 二次ブロックのサイズが小さいほど、必要な処理時間は長くなること、及び 2) 二次ブロックのサイズが大きいほど、ローカルアライメントの効果は下がること、を加重した後に選択してもよい。本例示的実施形態では、300 D P I（1 インチ当たりのドット数）文書画像のブロックサイズは 128 × 128 ピクセルであり、即ち、各二次ブロック 4 8 0 は少なくとも 1 cm<sup>2</sup> の面積を有してもよく、かつ画像は少なくとも 50 個または少なくとも 100 個の二次ブロックに、例えば二次ブロック 5 0 0 個にまで分割されてもよい。20

#### 【0054】

ステップ S 5 5 4 において、ローカルアライメントモジュール 1 1 4 は、A<sub>i</sub> 内の各ブロックについて文書 B 内のブロックとの最近一致を決定する。あるブロックの最近一致は、有効変換下の上記ブロックに最もよく似ている選択された文書画像 B の初ブロック 4 7 8（図 13）を発見することによって決定される。さらに、最近一致は、一致の品質に対応する品質スコアを含んでもよい。従って、認識されるべきであるように、あるブロックの最近一致は、図 7 に関連して論じた文書画像の初ブロック 4 7 8 の最近一致と同じ方法で決定されてもよく、よって前述の論考に注目されたい。30

#### 【0055】

ステップ S 5 5 6 において、ローカルアライメントモジュール 1 1 4 は、ステップ S 5 5 4 で決定された選択された文書画像 B 内の最近一致の二次ブロック 4 8 0（図 14）に対応する文書画像 A<sub>i</sub> 内の各ブロックの平行移動ベクトルを決定する。文書画像 A<sub>i</sub> 内の二次ブロック 4 8 0 の平行移動ベクトルは、単に、二次ブロック 4 8 0 が選択された文書画像 B 内のその最近一致にアラインするために必要とする平行移動の量である。所定の実施形態では、ある最近一致の品質スコアが信頼度しきい値より下であれば、二次ブロック 4 8 0 の平行移動ベクトルはその近傍の平行移動ベクトルから外挿されてもよい。即ち、このような二次ブロック 4 8 0 の平行移動ベクトルは、その近傍の平均平行移動ベクトルに設定されることが可能である。40

#### 【0056】

図 10 を参照すると、文書画像に適用された平行移動ベクトルのセットを示す文書画像50

の歪みマップが示されている。本例では、文書画像が $6 \times 8$ グリッドの二次ブロック480(図14)に分割されているが、これより少い、または多い二次ブロックが使用されてもよい。各ブロックは、1つの平行移動ベクトル(単に平行移動方向を示す矢印によって指示されている)に関連づけられている。ブロックの離散セットの平行移動ベクトルしか決定されていないが、双線形補間を用いれば、文書画像 $A_i$ 内の任意の点に関する平行移動ベクトルを、上記点を包囲する4つのグリッド点における平行移動ベクトルから決定することができる。

#### 【0057】

ステップS558において、ローカルアラインメントモジュール114は、ステップS556からの決定された平行移動を適用することによって文書画像 $A_i$ 内の各ブロックをシフトする。  
10

#### 【0058】

ステップS560において、ローカルアラインメントモジュール114は、ステップS558からのシフトされたブロックを組み立てる(纏めてステイッチする)ことによって変換された文書画像 $A'_i$ を生成する。従って、ステップS560が実行された後は、変換された文書画像154のセット $A'_1, A'_2, \dots, A'_n$ が生成される。但し、nは入力された文書画像のセット104内の文書画像の数である。制御は次に、S600へ進む。

#### 【0059】

ステップS600において、合意画像形成モジュール116は、アラインされた文書画像154 $A'_1, A'_2, \dots, A'_n$ を結合して合意画像156を形成する。合意画像156を形成するためにアラインされた文書画像154を結合する方法は幾つか存在する。任意選択として、任意カラーのアラインされた文書画像は、さらなる処理の前にグレースケールまたは2値(白黒)画像に変換されてもよい。アラインされた文書画像を結合する下記の例は何れも、別々に、または組み合わせて使用されてもよい。1つの方法は、単に、所定の特性(カラー値または2進値等)について全画像を合計し、次に適切な値で合計をスレッショルドして、図11に示すようにこれを白黒画像にすることである。合計はピクセル単位であり、よって位置( $x, y$ )における合意ピクセル $C(x, y)$ は合計からスレッショルドされる。  
20

$$(5) B(x, y) + \sum_{i=1}^n A'_i(x, y) = t$$

但し、 $B(x, y)$ 及び $A'_i(x, y)$ はピクセルのカラー値(但し、白は0、黒は1)であり、iは1からnまでの範囲の指標であり、nはシステム100へ入力される注釈付き文書画像の数であり、tはしきい値である。言い替えれば、位置( $x, y$ )におけるピクセルは、その計算された合計がしきい値以下であれば合意画像から曖昧にされる。例えば、50個の注釈付き文書画像の場合、しきい値は、複合画像内のピクセルが黒に設定されるために文書画像 $A'_i$ の少なくとも例えば30個がロケーション( $x, y$ )において黒ピクセルを有するように設定されてもよく、グレースケール及びカラー画像の場合、所定のしきい値を超える合計されたカラー特性(強度等)に設定されてもよい。また、選択された文書画像B内の( $x, y$ )におけるピクセルのカラー/強度も、式5に示すように、合計へと計数されてもよい。  
30

#### 【0060】

合意画像を形成する別 の方法は、文書画像B及び $A'_1, A'_2, \dots, A'_n$ 内のロケーション( $x, y$ )における各ピクセルが、もしそのピクセルが黒であれば、またはグレースケール及びカラー画像の場合、そのピクセルが所定のしきい値を超えるカラー特性(強度等)を有していれば計数されるプロセスを提案することによる。例えば、合意画像形成モジュール116は、 $C(x, y)$ が黒であるために、文書画像B及び $A'_1, A'_2, \dots, A'_n$ 内の( $x, y$ )におけるピクセルの3分の2が黒であることを要求してもよい。

#### 【0061】

合意画像を形成するためのさらに別の方法は、選択された代表の文書画像Bを特別に処  
40

理することである。例えば、 $C(x, y)$ は、 $B(x, y)$ が黒でありかつ $A'_{\cdot 1}(x, y)$ 、 $A'_{\cdot 2}(x, y)$ 、……、 $A'_{\cdot n}(x, y)$ の3分の2が黒である場合にのみ黒に設定される。

#### 【0062】

ステップS600が実行された後は、図11に示されているもの等の合意画像8が生成される。合意画像8は、図2及び図3に示す画像4及び画像6等の注釈付き文書画像から形成されることに起因して、十分に画定されたライン（境界等）及びラベル「食料品」及び「娯楽」等の特徴を含む。さらに、重なり合う散乱ピクセルのみを有する注釈付き文書画像104に対する異形及び注釈のスキャンから形成されるノイズ12が存在する場合がある。このノイズ12は、クリーンな文書2を生成するために合意画像8から除去してもよい。10

#### 【0063】

例えば、ステップS700において、クリーンアップモジュール118は、クリーンな文書画像2を形成するために、ごま塩ノイズ、隔絶された白黒ピクセル及びピクセルの小さい接続されたコンポーネントを除去することによって合意画像8を清浄化する。クリーンアップモジュール118は、破断された水平及び垂直ラインも修復する場合がある。

#### 【0064】

ステップS750において、注釈抽出モジュール119は、入力された注釈付き文書画像104の各々から注釈情報データ157を、入力された文書画像104をクリーンな文書2と比較することによって抽出する。本例示的実施形態では、注釈抽出モジュール119は、画像データが存在している注釈付き文書画像における空のフィールド及び対応するバウンディングボックスまたはエリアに対応するクリーン文書内のバウンディングボックスまたはエリアを識別し、微分（d i f f）演算を実行して注釈付き文書画像から注釈情報データを抽出する。抽出される注釈情報データには、注釈自体、バウンディングボックスの座標及び注釈が抽出された注釈付き文書画像を識別する識別子が含まれてもよい。バウンディングボックスは、クリーン文書において、幾何学的配置を基礎として識別される。例えば、図1において、点P1、P2、P3、P4によって境界をつけられるボックスは、長方形（未記入の書式では一般的であると思われる）を形成するためにP1、P2、P3、P4を結ぶラインを有することから、クリーンな文書2の明確なエレメントとして識別される。対応するボックスは、次に、図2のアラインされた注釈付き文書画像4内に同じ相対的な点P1、P2、P3、P4を用いて決定される。バウンディングボックスが決定された後、注釈付きバウンディングボックス内に存在する、クリーンなバウンディングボックスには存在しないピクセルを抽出するために、d i f f演算が実行される。ある実施形態では、このステップは、アラインされた注釈付き画像ではなく原初の注釈付き画像に対して実行されてもよい。注釈付き画像4のバウンディングボックスP1、P2、P3、P4に関してd i f f演算が実行された後は、図12の注釈データ情報が入手される。本例において、注釈データ情報は、「リンゴ、バナナ、ピザ、クッキー、バター」と記入する表記法、注釈付き文書画像4を識別する識別子及びその情報が正確にはどこから抽出されたかを知るために他のコンポーネントによって用いられてもよい座標P1、P2、P3、P4を含む。注釈付き画像から抽出された画像データは、手書きの注釈の場合、テキストのコンテンツを決定するためにOCRまたはワード認識技術を用いて処理されてもよい。3040

#### 【0065】

さらに、注釈データ情報が抽出された後、注釈抽出モジュール119は、任意選択として、ワード検出アルゴリズムを用いて注釈データを処理してもよい。

#### 【0066】

ステップS800において、クリーン文書再構築システム100は、クリーンな文書画像2及び抽出された注釈データ157を、出力デバイス120を介してメモリ106へ、またはクライアント端末130またはプリンタ132等の別の出力デバイスへ出力する。任意選択として、クライアント端末130、プリンタ132及び/または他の任意の出力50

デバイスは、インターネット等のネットワーク 134 を介して接続されてもよい。

【0067】

本方法は、S900で終了する。

【図1】

2 ↘

予算ワークシート			
月／年 ___ / ___ 使える金額 \$ ___ - ___			
このワークシートは、現金の管理に使用してください。 月々の現金支出先の把握に役立ちます。1つまたはそれ以上のカテゴリで現金を使い過ぎていれば、必要に応じて他の分野の金額を切りつめるか、調整しましょう。			
支出カテゴリ	第1週	第2週	第3週
P1 食費：	P4 P3 \$ ___	\$ ___	\$ ___
P2 娯楽費：	\$ ___	\$ ___	\$ ___
衣料費：	\$ ___	\$ ___	\$ ___
家庭用品：	\$ ___	\$ ___	\$ ___
家賃：	\$ ___	\$ ___	\$ ___
交通費：	\$ ___	\$ ___	\$ ___
その他：	\$ ___	\$ ___	\$ ___
合計：	\$ ___	\$ ___	\$ ___

図1

【図2】

4 ↘

予算ワークシート			
月／年 11 / 2010 使える金額 \$ 1500			
このワークシートは、現金の管理に使用してください。 月々の現金支出先の把握に役立ちます。1つまたはそれ以上のカテゴリで現金を使い過ぎていれば、必要に応じて他の分野の金額を切りつめるか、調整しましょう。			
支出カテゴリ	第1週	第2週	第3週
P1 食費： P3 りんご、バナナ、ピザ、 クッキー、バター	P4 P3 \$ 89	\$ 73	\$ 109
P2 娯楽費： 映画、ビデオゲーム、本	\$ 52	\$ 97	\$ 65
衣料費： 帽子、ミトン、靴、スカート	\$ 20	\$ 37	\$ 16
家庭用品： タオル、石けん、タッパー/ウェア	\$ 108	\$ 42	\$ 38
家賃： エバーグリーンテラス 742	\$ 850	\$ 850	\$ 850
交通費： 車、トラック、ガス	\$ 237	\$ 250	\$ 224
その他： おむつ、塗料、公共料金	453 478 \$ 100	\$ 150	\$ 150
合計：	\$ 1456	\$ 1499	\$ 1452

図2

【図3】

予算ワークシート  
月／年 2 / 2009 使える金額 \$ 1600  
このワークシートは、現金の管理に使用してください。  
日々の現金支出先の把握に役立ちます。1つまたはそれ以上のカテゴリで現金を使い過ぎていれば、必要に応じて他の分野の金額を切り替えるか、調整しましょう。

支出カテゴリ	第1週	第2週	第3週
食費： 卵、ピーナツバター、サンドイッチ、小麦粉	\$ 46	\$ 140	\$ 104
娯楽費： 野球シーズンチケット、映画、ナイターゲーム	\$ 45	\$ 104	\$ 87
衣料費： コート、バンダナ、サングラス	\$ 100	\$ 137	\$ 160
家庭用品：ペーパータオル、銀食器、タイヤクリーナー	\$ 125	\$ 43	\$ 147
家賃： ヒルクレストドライブ 6 1 8	\$ 600	\$ 600	\$ 600
交通費： 定期券	\$ 125	\$ 125	\$ 125
その他： 公共料金、保険料、ジム会費	453	\$ 175	\$ 200
合計：	\$ 1216	\$ 1324	\$ 1423

図3

【図4】

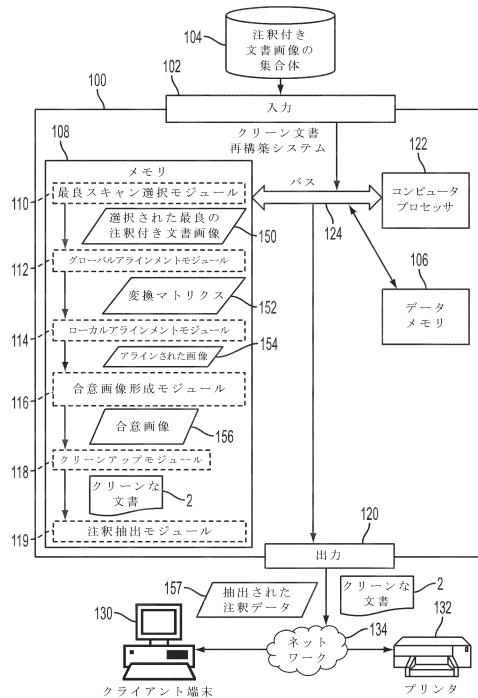


図4

【図5】

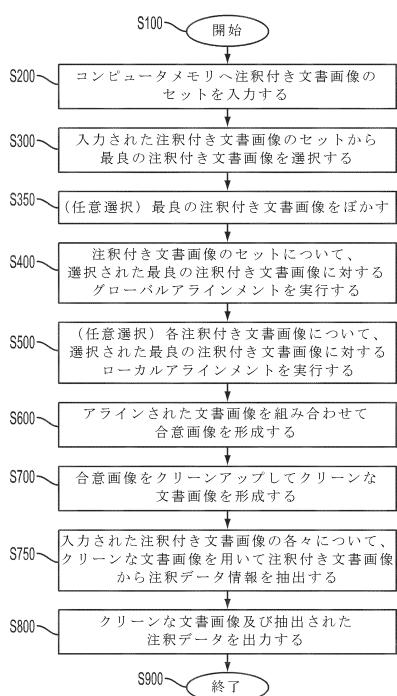


図5

【図6】

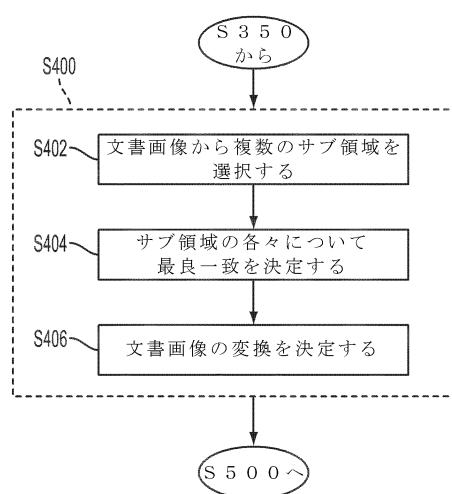


図6

【図7】

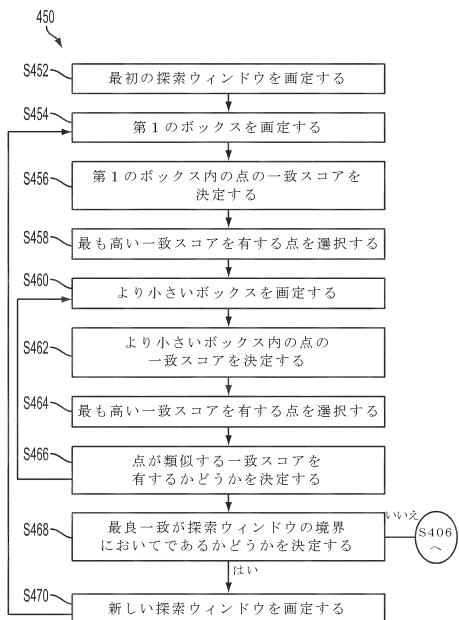


図7

【図8】

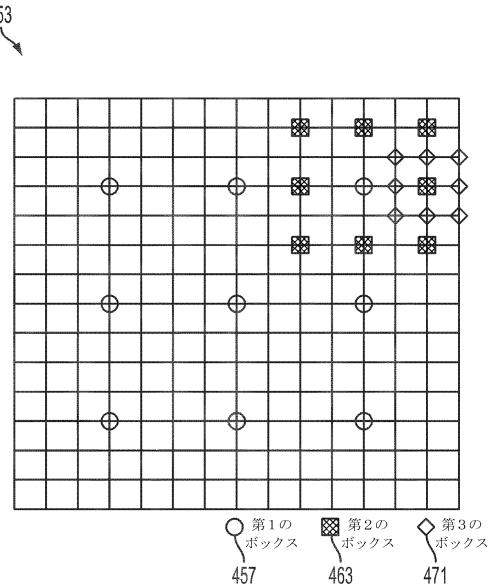


図8

【図9】

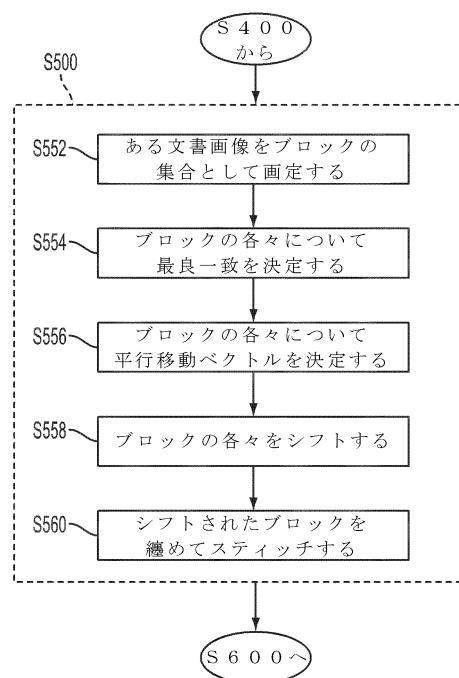


図9

【図10】

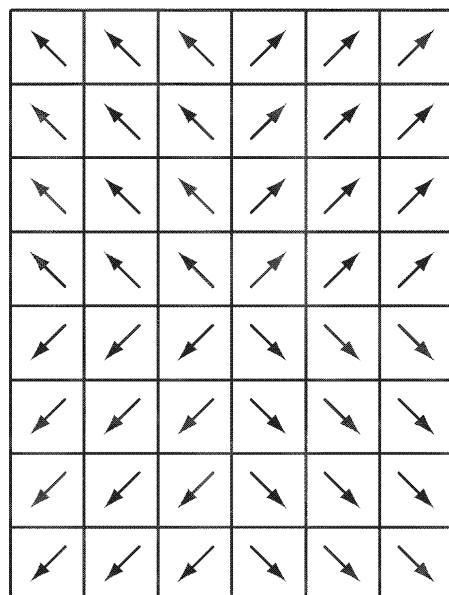


図10

【図 1 1】

8 ↘

予算ワークシート			
月／年 20 ___ / ___ 使える金額 \$ _____			
このワークシートは、現金の管理に使用してください。月々の現金支出先の把握に役立ちます。1つまたはそれ以上のカテゴリで現金を使い過ぎていれば、必要に応じて他の分野の金額を切りつめるか、調整しましょう。			
支出カテゴリ	第1週	第2週	第3週
食費： 12	\$ 12	\$ 12	\$ 12
娯楽費：	\$ 12	\$ 12	\$ 12
衣料費：	\$ 12	\$ 12	\$ 12
家庭用品：	\$ 12	\$ 12	\$ 12
家賃：	\$ 12	\$ 12	\$ 12
交通費：	\$ 12	\$ 12	\$ 12
その他： 12	\$ 12	\$ 12	\$ 12
合計：	\$ 12	\$ 12	\$ 12

図 1 1

【図 1 2】

りんご、バナナ、ピザ、  
クッキー、バター

図 1 2

【図 1 3】

予算ワークシート			
月／年 2 / 2009 使える金額 \$ 1 600			
このワークシートは、現金の管理に使用してください。月々の現金支出先の把握に役立ちます。1つまたはそれ以上のカテゴリで現金を使い過ぎていれば、必要に応じて他の分野の金額を切りつめるか、調整しましょう。			
支出カテゴリ	第1週	第2週	第3週
食費： 卵、ビーナツバター、サンドイッチ、小麦粉	\$ 46	\$ 140	\$ 104
娯楽費： 野球シーズンチケット、映画、ナイターゲーム	\$ 45	\$ 104	\$ 87
衣料費： コート、パンダナ、サングラス	\$ 100	\$ 137	\$ 160
家庭用品：ペーパータオル、銀食器、タイヤクリーナ	\$ 125	\$ 43	\$ 147
家賃： ヒルクレストドライブ618	\$ 600	\$ 600	\$ 600
交通費： 定期券	\$ 125	\$ 125	\$ 125
その他： 公共料金、保険料、ジム会費	\$ 175	\$ 175	\$ 200
合計：	\$ 1216	\$ 1324	\$ 1423

図 1 3

【図 1 4】

予算ワークシート			
月／年 2 / 2009 使える金額 \$ 1 600			
このワークシートは、現金の管理に使用してください。月々の現金支出先の把握に役立ちます。1つまたはそれ以上のカテゴリで現金を使い過ぎていれば、必要に応じて他の分野の金額を切りつめるか、調整しましょう。			
支出カテゴリ	第1週	第2週	第3週
食費： 卵、ビーナツバター、サンドイッチ、小麦粉	\$ 46	\$ 140	\$ 104
娯楽費： 野球シーズンチケット、映画、ナイターゲーム	\$ 45	\$ 104	\$ 87
衣料費： コート、パンダナ、サングラス	\$ 100	\$ 137	\$ 160
家庭用品：ペーパータオル、銀食器、タイヤクリーナ	\$ 125	\$ 43	\$ 147
家賃： ヒルクレストドライブ618	\$ 600	\$ 600	\$ 600
交通費： 定期券	\$ 125	\$ 125	\$ 125
その他： 公共料金、保険料、ジム会費	\$ 175	\$ 175	\$ 200
合計：	\$ 1216	\$ 1324	\$ 1423

図 1 4

---

フロントページの続き

(74)代理人 100079049

弁理士 中島 淳

(74)代理人 100084995

弁理士 加藤 和詳

(72)発明者 マーシャル・ダブリュ・バーン

アメリカ合衆国 カリフォルニア州 94070 サン・カーロス シーダー・ストリート 36  
8

(72)発明者 アレハンドロ・イー・ブリト

アメリカ合衆国 カリフォルニア州 94070 マウンテン・ビュー オルテガ・アヴェニュー  
163

(72)発明者 フランソワ・ラグネ

フランス国 ヴェノン 38610 プレ・ペロウ

審査官 成瀬 博之

(56)参考文献 特開2009-232085(JP,A)

特開平06-119488(JP,A)

特開2009-182870(JP,A)

特開平11-003431(JP,A)

特開2001-243423(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 17/20 - 17/26