(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3832879号
(P3832879)
(45)発行日　平成18年10月11日(2006.10.11)　(24)登録日　平成18年7月28日(2006.7.28)

(51)Int.Cl.　F I
G06K 9/20 (2006.01)　G06K 9/20 340J

請求項の数 2 (全 32 頁)

(21)出願番号　特願平7-243213
(22)出願日　平成7年9月21日(1995.9.21)
(65)公開番号　特開平8-185474
(43)公開日　平成8年7月16日(1996.7.16)
審査請求日　平成14年8月15日(2002.8.15)
(31)優先権主張番号　315875
(32)優先日　平成6年9月30日(1994.9.30)
(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　スタンフォード、ロング・リッジ・ロード　８００
(73)特許権者　000005496
富士ゼロックス株式会社
東京都港区赤坂二丁目１７番２２号
(74)代理人　100079049
弁理士　中島　淳
(74)代理人　100084995
弁理士　加藤　和詳

最終頁に続く

(54)【発明の名称】文書画像分割装置

(57)【特許請求の範囲】
【請求項１】
文書画像を少なくとも一つの文書エレメントへ分割する文書画像分割装置であって、
前記画像信号を入力する入力手段と、
前記画像信号を格納するメモリ手段と、
前記文書画像を形成する前記複数の画素から、隣接したオン画素から成る少なくとも一つの連結成分を識別する連結成分識別手段と、
前記少なくとも一つの連結成分のそれぞれの外周に、該連結成分を完全に取り囲む最小サイズの境界ボックスを発生する境界ボックス手段と、
前記境界ボックスを取り囲む背景エリアを閉ループ状の領域として認識すると共に該閉ループ状の領域を複数の閉じられた領域に分割し、該複数の閉じられた領域の各々を識別する背景エリア識別手段と、
前記背景エリアを構成する分割された複数の閉じられた領域のうち、水平方向および垂直方向の長さが閾値より小さいものを除去し、残った領域をグループ化することにより、前記背景エリアから、前記境界ボックスを取り囲む閉ループ状の複数の前記閉じられた領域である主要背景領域を抽出する主要背景領域抽出手段と、
抽出された前記主要背景領域に基づいて、前記少なくとも一つの文書エレメントを抽出する手段と、
前記入力手段、前記メモリ手段、前記連結成分識別手段、前記境界ボックス発生手段、前記背景エリア識別手段、前記主要背景領域抽出手段及び前記文書エレメント抽出手段を

制御する制御手段と、を備え、
前記主要背景領域抽出手段が、前記除去された領域を再検討し、該再検討された領域のうち前記主要背景領域に併合可能なものがあるかどうかを決定し、併合可能なものがある場合には併合を行う手段を含む、ことを特徴とする、
文書画像分割装置。
【請求項２】
各々画像信号の部分によって示される複数の画素によって形成される文書画像を少なくとも一つの文書エレメントへ分割する文書画像分割方法であって、
前記画像信号を入力するステップと、
前記画像信号を格納するステップと、
前記文書画像を形成する前記複数の画素から、隣接したオン画素から成る少なくとも一つの連結成分を識別するステップと、
前記少なくとも一つの連結成分のそれぞれの外周に、該連結成分を完全に取り囲む最小サイズの境界ボックスを発生するステップと、
前記境界ボックスを取り囲む背景エリアを閉ループ状の領域として認識すると共に該閉ループ状の領域を複数の閉じられた領域に分割し、該複数の閉じられた領域の各々を識別するステップと、
前記背景エリアを構成する分割された複数の閉じられた領域のうち、水平方向および垂直方向の長さが閾値より小さいものを除去し、残った領域をグループ化することにより、前記背景エリアから、前記境界ボックスを取り囲む閉ループ状の複数の前記閉じられた領域である主要背景領域を抽出するステップと、
前記抽出された主要背景領域に基づいて、前記少なくとも一つの文書エレメントを抽出するステップと、を含み、
前記主要背景領域を抽出するステップが、前記除去された領域を再検討し、該再検討された領域のうち前記主要背景領域に併合可能なものがあるかどうかを決定し、併合可能なものがある場合には併合を行うステップを含む、ことを特徴とする、
文書画像分割方法。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
本発明は、文書画像から主要ホワイト領域及び文書エレメントを抽出するための方法及び装置に関する。
【０００２】
【従来の技術】
テキスト領域を分割する方法は、１９９０年６月１６乃至２１日開催の第１０回「パターン認識」国際会議議事録、８２０乃至８２５頁の"Image Segmentation By Shape-Directed Covers（「形状指定されたカバーによる画像分割」"）においてBaird （ベアード）等によって提示されている。ベアードにおいて開示された方法は、文書画像内のホワイトスペースを解析しているが、テキスト領域分割に対する終了ルールを明確に開示していない。
【０００３】
文書画像内の文書エレメントを分割するための他の方法も開示されているが、これらの方法はどれも、文書画像内のホワイトエリアを解析していないし、また主要ホワイト領域を抽出、解析することにより文書エレメントを分割していない。ホワイト領域は連結成分を全く含まない文書のエリアである。
【０００４】
【発明が解決しようとする課題】
本発明は、文書画像上のホワイト領域のみを解析することによって文書画像内のエレメントを識別するためのシステムを提供する。
【０００５】

本発明は、必ずしも矩形に形付けられてない文書画像から文書画像を抽出することができる文書エレメント分割システムも提供する。

【０００６】

【課題を解決するための手段】

本発明を用いれば、どの連結成分がコヒーレントグループ即ち文書エレメントを形成するかを決定するのに文書エレメントを含む文書画像の部分を解析する必要はない。文書上の画像は、画像の電子的又はディジタル表示を生成するように走査される。主要ホワイト領域は、所定の最小サイズを有するホワイトスペースの矩形エリアである。文書エレメントは、テキストや図形などの情報を含むエリアであり、主要ホワイト領域によって互いから分離される。文書エレメントを含むと共に所定サイズのホワイトスペースによって他のエリアから分離されたエリアは、分離文書エレメントであると仮定される。文書画像内のこのような文書エレメントの例は、文書のタイトル（題名）、著者、脚注、テキストの活字体等である。一旦抽出されると、これらの文書エレメントは、光学的文字認識ユニット等によって更に処理される。

【０００７】

請求項１記載の発明は、文書画像を少なくとも一つの文書エレメントへ分割する文書画像分割装置であって、前記画像信号を入力する入力手段と、前記画像信号を格納するメモリ手段と、前記文書画像を形成する前記複数の画素から、隣接したオン画素から成る少なくとも一つの連結成分を識別する連結成分識別手段と、前記少なくとも一つの連結成分のそれぞれの外周に、該連結成分を完全に取り囲む最小サイズの境界ボックスを発生する境界ボックス手段と、前記境界ボックスを取り囲む背景エリアを閉ループ状の領域として認識すると共に該閉ループ状の領域を複数の閉じられた領域に分割し、該複数の閉じられた領域の各々を識別する背景エリア識別手段と、前記背景エリアを構成する分割された複数の閉じられた領域のうち、水平方向および垂直方向の長さが閾値より小さいものを除去し、残った領域をグループ化することにより、前記背景エリアから、前記境界ボックスを取り囲む閉ループ状の複数の前記閉じられた領域である主要背景領域を抽出する主要背景領域抽出手段と、抽出された前記主要背景領域に基づいて、前記少なくとも一つの文書エレメントを抽出する手段と、前記入力手段、前記メモリ手段、前記連結成分識別手段、前記境界ボックス発生手段、前記背景エリア識別手段、前記主要背景領域抽出手段及び前記文書エレメント抽出手段を制御する制御手段と、を備え、前記主要背景領域抽出手段が、前記除去された領域を再検討し、該再検討された領域のうち前記主要背景領域に併合可能なものがあるかどうかを決定し、併合可能なものがある場合には併合を行う手段を含む、ことを特徴とする。

また、請求項２記載の発明は、各々画像信号の部分によって示される複数の画素によって形成される文書画像を少なくとも一つの文書エレメントへ分割する文書画像分割方法であって、前記画像信号を入力するステップと、前記画像信号を格納するステップと、前記文書画像を形成する前記複数の画素から、隣接したオン画素から成る少なくとも一つの連結成分を識別するステップと、前記少なくとも一つの連結成分のそれぞれの外周に、該連結成分を完全に取り囲む最小サイズの境界ボックスを発生するステップと、前記境界ボックスを取り囲む背景エリアを閉ループ状の領域として認識すると共に該閉ループ状の領域を複数の閉じられた領域に分割し、該複数の閉じられた領域の各々を識別するステップと、前記背景エリアを構成する分割された複数の閉じられた領域のうち、水平方向および垂直方向の長さが閾値より小さいものを除去し、残った領域をグループ化することにより、前記背景エリアから、前記境界ボックスを取り囲む閉ループ状の複数の前記閉じられた領域である主要背景領域を抽出するステップと、前記抽出された主要背景領域に基づいて、前記少なくとも一つの文書エレメントを抽出するステップと、を含み、前記主要背景領域を抽出するステップが、前記除去された領域を再検討し、該再検討された領域のうち前記主要背景領域に併合可能なものがあるかどうかを決定し、併合可能なものがある場合には併合を行うステップを含む、ことを特徴とする。

【０００８】

【発明の実施の形態】
図１は、文書エレメント分割システム１００のブロック図の第１の好ましい実施の形態を示す。記述されている文書エレメント分割システム１００は、以下に説明される文書エレメント分割方法を実行するための多くの可能な実施の形態の内の一つにすぎない。従って、開示されている文書エレメント分割システム１００は、例として示されているだけであり、文書エレメント分割方法を実行するために使用され得るシステムの種々のタイプ及び構成を制限するものではない。
【０００９】
図１は、文書画像を分割するための種々の手段を有する汎用コンピュータ４００を示している。特に、文書エレメント分割システム１００は、メモリ４１０、連結成分識別手段４２０、境界ボックス発生手段４３０、主要ホワイト領域抽出手段４４０、文書エレメント抽出手段４５０、及びプロセッサ４６０を含み、これらの全てはバス手段４０５によって接続されている。スキャナー２００及びプリンタ３００も当該バス手段４０５へ接続されている。文書画像データはスキャナー２００によって又はメモリ４１０から連結成分識別手段４２０へ最初に入力される。メモリ４１０は、汎用コンピュータ内に配置されてもよいし、従来の技術において周知のように、ディスクドライブ、ＣＤ－ＲＯＭ、ＥＰＲＯＭ等の形態で汎用コンピュータの外部に配置されてもよい。同様に、スキャナ２００からの文書画像データは、連結成分識別手段４２０へ入力される前に、最初に、メモリ４１０へ格納されてもよい。文書画像データは、バイナリ（２値）画像又は複数ビットのディジタル信号の形態で、連結成分識別手段４２０へ入力され、各ビットは、文書画像の特定画素がオンであるか又はオフであるかを示す。
【００１０】
連結成分識別手段４２０が文書画像データを受け取った後、この連結成分識別手段４２０は、文書画像内の全ての連結成分を識別する。図３は、代表的な文書画像６００を示す。連結成分６１０は文書画像６００内で見つけられる。各連結成分６１０は、「オフ」（ホワイト）画素によって取り囲まれた、一連の隣接している「オン」（ブラック）画素を備える。文書画像６００内の連結成分６１０を識別するためのシステムは、従来の技術において周知である。
【００１１】
文書画像６００の連結成分６１０が識別されると、境界ボックス発生手段４３０が各連結成分６１０毎に境界ボックス６２０を発生する。図３に示されているように、境界ボックス６２０は、従来の技術において周知であるように、対応する連結成分６１０を完全に取り囲む最小サイズの矩形ボックスである。連結成分６１０から境界ボックス６２０を発生するためのシステムも従来の技術において周知である。
【００１２】
境界ボックス情報を有する文書画像データは主要ホワイト領域抽出手段４４０へ送られ、この主要ホワイト領域抽出手段４４０は、図４及び図５に示されているように、文書画像６００の垂直及び水平方向の主要ホワイト領域６６０を抽出する。
【００１３】
文書ホワイト領域抽出システム１１０の好ましい実施の形態においては、図２に示されているように、主要ホワイト領域抽出手段４４０は、二つの部分、即ち、垂直抽出部分４４１と水平抽出部分４４２に分かれている。垂直抽出部分４４１と水平抽出部分４４２は、それぞれ、プリミティブ（基本的エレメント）ホワイトエリア抽出手段４４３、比較手段４４４、除去手段４４５、及びグループ化手段４４６を備え、これらの各々がバス手段４０５に接続されている。垂直抽出部分４４１と水平抽出部分４４２は、同一成分を含んでおり、同様の方法で動作する。図４に示されているように、水平抽出部分４４２は、最初に、プリミティブホワイトエリア６３０－１乃至６３０－１０を抽出し、水平方向に主要ホワイト領域６６０を集める。同様に、図５に示されているように、垂直抽出部分４４１は、最初に、プリミティブホワイトエリア６３０－１１乃至６３０－１９を抽出し、垂直方向に主要ホワイト領域６６０を集める。

【００１４】
水平主要ホワイト領域６６０のアセンブリは、水平プリミティブホワイトエリア６３０－１乃至６３０－１０の隣接しているエリアを、特定のルールに従って、単一の水平のグループ化されたプリミティブホワイトエリア６３０’へグループ化し、併合することによって行われる。同様に、垂直主要ホワイト領域６６０のアセンブリは、垂直プリミティブホワイトエリア６３０－１１乃至６３０－１９の隣接しているエリアを、特定のルールに従って、単一の垂直のグループ化されたプリミティブホワイトエリア６３０’へグループ化し、併合することによって行われる。垂直及び水平のプリミティブホワイトエリアのグループ化及び併合が完了した後、幅閾値（しきい値）６４０より大きな幅と高さ閾値６５０’より大きな高さとを有する水平プリミティブホワイトエリア６３０と６３０’及び高さ閾値６５０より大きな高さと幅閾値６４０’より大きな幅とを有する垂直プリミティブホワイトエリア６３０と６３０’が主要ホワイト領域６６０として識別される。図４及び図５に示されている高さ及び幅の閾値６４０、６５０、６４０’及び６５０’が拡大縮小するように図示されていないことは理解されるべきである。

【００１５】
文書画像６００において垂直及び水平の主要ホワイト領域６６０が一旦抽出されると、図６及び図７に示されているように、文書エレメント抽出手段４５０は、文書画像６００内の文書エレメント６７０を分割するために主要ホワイト領域６６０を使用する。文書エレメント６７０は、主要ホワイト領域６６０をトラックして閉ループを見つけることによって抽出される。その内部に他の閉ループを持たない閉ループは、図７に示されているように、文書エレメント６７０を文書画像６００の残りの部分から分割する。

【００１６】
図７に示されているように、他の閉ループを含まない閉ループが三つある。これは文書画像６００内に三つの分割された文書エレメント６７０が含まれていることを意味する。文書エレメント６７０は矩形形状である必要はなく、または四つ以上の辺を有していてもよい。

【００１７】
分割された文書エレメント６７０はメモリ４１０、プリンタ３００、及びプロセッサ４６０を含む種々のシステムコンポーネントへ出力される。分割された文書エレメント６７０がいかにして出力されるかは、当該分割された文書エレメント６７０を用いて何をなすべきかによる。文書エレメント６７０はメモリ４１０に格納されてもよいし又はプリンタ３００でプリントされてもよい。分割された文書エレメント６７０は、例えば、各分割された文書エレメント６７０内に含まれる連結成分に対して光学的文字認識処理を実行するように、プロセッサ４６０によって更に処理されてもよい。

【００１８】
文書エレメント抽出手段４５０は、分割された文書エレメント６７０又は当該分割された文書エレメント６７０に関する情報を、プリンタ３００、メモリ４１０、又は更なる処理のためにプロセッサ４６０へ出力する。

【００１９】
図８は、主要ホワイト領域６６０を抽出するための方法の好ましい実施の形態を大まかに示している。まず、ステップＳ１００でスタートし、このステップＳ１００において、文書画像６００が入力される。入力された文書画像６００は、多数の文書エレメント６７０を有している。ステップＳ１２０において、文書画像６００の連結成分６１０が識別される。次に、ステップＳ１３０において、境界ボックス６２０が、ステップＳ１２０において識別された各連結成分６１０毎に発生される。ステップＳ１４０において、主要ホワイト領域（背景）６６０が抽出される。ステップＳ１５０において、主要ホワイト領域６６０によって分割された文書エレメント６７０が抽出される。ステップＳ１６０においてプロセスが終了する。

【００２０】
図９は、ステップＳ１４０の主要ホワイト領域抽出プロセスの第１の好ましい実施の形態

を大まかに示している。ステップＳ２００においては、プリミティブホワイトエリア６３０が抽出される。図５に示されているように、プリミティブホワイトエリア６３０は、境界ボックス６２０同士の間のホワイトスペースの矩形エリアである。次に、ステップＳ３００において、各水平プリミティブホワイトエリア６３０の高さ及び幅が、幅閾値６４０と比較され、各垂直プリミティブホワイトエリア６６０の高さが、高さ閾値６５０’と比較される。水平方向の幅閾値６４０は、好ましくは、水平方向の文書画像６００の長さの３分の１にセットされる。水平方向の高さ閾値６５０は、文書画像におけるテキストの行間隔より大きな値にセットされる。垂直方向における高さ閾値６５０’は、好ましくは、垂直方向にある文書画像６００の長さの３分の１にセットされる。垂直方向における幅閾値６４０’は、文書画像内のテキスト内の行間隔より大きな値にセットされる。

【００２１】

ステップＳ４００において、幅閾値６４０より小さな幅を有する水平プリミティブホワイトエリア６３０と、高さ閾値６５０’より小さな高さを有する垂直ホワイトエリア６３０と、が除去される。ステップＳ５００において、残っているプリミティブホワイトエリア６３０がグループ化され、主要ホワイト領域６６０が設定される。最後に、ステップＳ６００において、対応している垂直又は水平の閾値より小さい垂直又は水平の凸部の内の少なくとも一つを有している主要ホワイト領域６６０が除去される。或いは、対応している閾値より小さな水平及び垂直の両凸部を有している主要ホワイト領域６６０のみが除去される。次いで、ステップＳ７００において、コントロールはステップＳ１５０へ戻る。

【００２２】

図１０乃至図１１は、ステップＳ２００の水平プリミティブホワイトエリア抽出プロセスの第１の好ましい実施の形態を大まかに示している。垂直プリミティブホワイトエリア抽出プロセスは、左右エッジが上下エッジの代わりに使用されていることを除いて、水平ホワイトエリア抽出プロセスと同じである。ステップＳ２１０において、ギャップリストが設定される。ステップＳ２１２において、ギャップリストは、最初に、第１のプリミティブホワイトギャップ６８０を含むようにセットされる。プリミティブホワイトギャップ６８０は、潜在的プリミティブホワイトエリア６３０であり、この潜在的プリミティブホワイトエリア６３０は、トップサイド（上部側）とレフトサイド（左側）及びライトサイド（右側）を設定させているが、ボトムサイド（下部側）はまだ設定させていない。ギャップリストに最初に表記された第１のプリミティブホワイトギャップ６８０は、文書画像６００の上部にトップを有し、文書画像６００の左側遠方にレフトサイドを有し、文書画像６００の右側遠方にライトサイドを有している。ステップＳ２１４において、プリミティブホワイトエリアリストは、最初は、空として設定される。

【００２３】

ステップＳ２１６において、現行走査線１０００が、文書画像６００の上部に配置されている。ステップＳ２１７において、現行走査線１０００に沿って左から右へ走査が開始される。ステップＳ２１８において、走査線１０００は、境界ボックスのエッジ又は文書画像６００のエッジが出会ったか否かを決定するためにチェックされる。エッジに出会った場合、コントロールはステップＳ２２０へ進む。エッジに出会わなかった場合、コントロールは、現行走査線１０００の更なる走査のためにステップＳ２１７へ戻る。

【００２４】

ステップＳ２２０において、出会ったエッジが、境界ボックス６２０の内の一つのボックスの以前に出会ってないトップである場合、コントロールはステップＳ２２２へ飛び越す。このステップ２２２において、図１２に示されているように、境界ボックス６２０のトップエッジに接触するギャップリスト内のプリミティブホワイトギャップ６８０Ａが識別される。一つ又はそれより多くのこのようなプリミティブホワイトギャップ６８０Ａがあってよい。ステップＳ２２３において、境界ボックス６２０のトップエッジに接触するプリミティブホワイトギャップ６８０Ａ毎にボトムが設定される。ステップＳ２２４において、プリミティブホワイトギャップ６８０Ａがギャップリストから除去される。ステップＳ２２６においては、ギャップリストから除去されたプリミティブホワイトギャップ６８

０Ａが、プリミティブホワイトエリア６３０としてプリミティブホワイトギャップリストに表記される。

【００２５】

次いで、図１３に示されているように、ステップＳ２２８において、境界ボックス６２０のトップエッジと同じラインに沿って配置されたトップと、境界ボックス６２０のサイドに隣接しているサイドと、を有する新しいプリミティブホワイトギャップ６８０Ｂが、設定される。図１４及び図１５に示されているように、前に出会っていない境界ボックスのトップエッジは、プリミティブホワイトギャップ６８０Ａの垂直エッジの片方又は両方と交差してもよい。この場合、ギャップリストへの新しいプリミティブホワイトギャップ６８０Ｂの追加は全く無いか又はあるとしても一つだけにすぎない。ステップＳ２２９において、新しいプリミティブホワイトギャップ６８０Ｂがギャップリストに表記される。次いで、コントロールは現行走査線１０００に沿って更なる走査を行うためにステップＳ２１７へ戻る。

【００２６】

現行走査線１０００の走査はステップＳ２１７において続行される。境界ボックス６２０の内の一つのボックスの他の以前に出会っていないトップエッジに出会った場合、上記のステップＳ２２２乃至Ｓ２２９が繰り返される。

【００２７】

ステップＳ２２０において、出会ったエッジが、境界ボックス６２０の内の一つの他の以前に出会っていないトップエッジでなかった場合、コントロールはステップＳ２３０へ進む。ステップＳ２３０においては、出会ったエッジが、境界ボックス６２０の内の一つのボックスの他の以前に出会っていないボトムエッジである場合、コントロールはステップＳ２３２へ進む。ステップＳ２３２においては、ボトムエッジが検出されている境界ボックス６２０のサイドに接触しているプリミティブホワイトギャップ６８０Ｂが識別される。一般に、もっと少なくなる可能性はあるが、二つ以下のこのようなギャップ６８０Ｂが存在する。稀ではあるが、境界ボックス６２０の片側が、文書画像６６０又は他の境界ボックス６２０のエッジと同じ高さである場合、ギャップ６８０Ｂだけが存在する。ステップＳ２３３において、境界ボックス６２０のサイドに接触しているプリミティブホワイトギャップ６８０Ｂ毎にボトムが設定される。ステップ２３４において、新しく設定されたボトムエッジを有するプリミティブホワイトギャップ６８０Ｂがギャップリストから除去される。ステップＳ２３６において、プリミティブホワイトギャップ６８０Ｂがプリミティブホワイトエリア６３０としてプリミティブホワイトエリアリスト上で検出される。ステップＳ２３８において、新しく出会った境界ボックス６２０のボトムエッジと同じラインに沿ったトップと、他の境界ボックス６２０のサイド又は文書画像６００の左側遠方に設定されたレフトサイド及び／又は右側遠方に設定されたライトサイドと、を有する新しいプリミティブホワイトギャップ６８０Ｃが設定される。ステップＳ２３９において、新しく設定されたプリミティブホワイトギャップ６８０Ｃは、ギャップリストに表記される。次いで、コントロールは、現行走査線１０００に沿って更なる走査を行うためにステップＳ２１７へ戻る。

【００２８】

図１４乃至図１８は、走査線１０００内で境界ボックス６２０のトップ又はボトムに出会った時に発生し得るいくつかの可能性を示す。図１４は、境界ボックス６２０のトップに接触しており、境界ボックス６２０のライトサイドの右へのみ拡張しているプリミティブホワイトギャップ６８０Ａを示す。プリミティブホワイトギャップ６８０Ａに対するボトムは、境界ボックス６２０のトップに設定され、新しいプリミティブホワイトギャップ６８０Ｂは、境界ボックス６２０のトップと同じラインに沿ったトップを有する境界ボックス６２０の右へ設定される。図１５は、境界ボックス６２０のトップエッジに接触しているが、この境界ボックス６２０のレフトサイド又はライトサイドのどちらにも拡張していないプリミティブホワイトギャップ６８０Ａを示す。このプリミティブホワイトギャップ６８０Ａに対するボトムは、境界ボックス６２０のトップに設定されるが、新しいプリミ

ティブホワイトギャップ６８０Ｂは全く設定されない。

【００２９】

図１６は、走査線１０００内で出会った境界ボックス６２０のボトムを示している。プリミティブホワイトギャップ６８０Ｂは、境界ボックス６２０のいずれかのサイドに配置され、境界ボックス６２０のサイドに接触する。従って、境界ボックス６２０のサイドのいずれかにあるプリミティブホワイトギャップ６８０Ｂのためにボトムが設定され、新しく設定されたボトムを有するプリミティブホワイトギャップ６８０Ｂがギャップリストから除去され、プリミティブホワイトエリアとしてプリミティブホワイトエリアリストに表記される。境界ボックス６２０のボトムと同じラインに沿って配置されたトップを有する新しいプリミティブホワイトギャップ６８０Ｃは、境界ボックス６２０のボトムの下に設定される。この新しいプリミティブホワイトギャップ６８０Ｃがギャップリストに表記される。

【００３０】

図１７は、走査線１０００内で境界ボックス６２０のボトムに出会っており、プリミティブホワイトギャップ６８０Ｂが境界ボックス６２０の右側に配置されているのを示している。境界ボックス６２０の右側に配置されたプリミティブホワイトギャップ６８０Ｂに対するボトムが、境界ボックス６２０のボトムと同じラインに沿って設定され、新しいプリミティブホワイトエリアを設定する。境界ボックス６２０のボトムと同じラインに沿って配置されたトップと第２の境界ボックス６２０のライトサイドに沿って配置されたレフトサイドと文書画像６００の右端遠方に沿って配置されたライトサイドとを有する新しいプリミティブホワイトギャップ６８０Ｃが設定される。

【００３１】

図１８は、走査線１０００内で境界ボックス６２０のボトムに出会っており、プリミティブホワイトギャップ６８０Ｂがこの境界ボックス６２０のサイドに全く接触していないことを示している。境界ボックス６２０のボトムにおいてトップを有し、第２の境界ボックス６２０のライトサイドに沿って配置されたレフトサイドを有し、第３の境界ボックス６２０のレフトサイド沿って配置されたライトサイドを有している新しいプリミティブホワイトギャップ６８０Ｃが設定される。

【００３２】

走査線１０００における走査は、文書画像の右エッジへ達するまで続けられる。ステップＳ２１８において検出されたエッジがトップエッジ又はボトムエッジでない場合、コントロールはステップＳ２２０及びステップＳ２３０を飛び越してステップＳ２４０へ進む。このステップにおいて、文書画像の下の次の走査線が現行走査線１０００になる。コントロールはステップＳ２４２へ進み、このステップにおいて、現行走査線１０００が文書画像の最終の走査線に一致しているか否かを見るためにチェックされる。一致している場合、コントロールはステップＳ２４４へ進み、このステップにおいて、ギャップリストに残っている全てのプリミティブホワイトギャップ６８０Ａ、６８０Ｂ、及び６８０Ｃに対するボトムが文書画像６００のボトムにおいて設定される。ステップＳ２４６において、新しいプリミティブホワイトエリア６３０が、文書画像６００のボトムにおいてボトムを有するプリミティブホワイトギャップ６８０に対応するプリミティブホワイトエリアリストに表記される。次いで、ステップＳ２４８において、コントロールはステップＳ３００へ戻る。

【００３３】

図１９乃至図２０は、ステップＳ５００のプリミティブホワイトエリアを主要ホワイトエリアへグループ化するためのプロセスに対する一つの好ましい実施の形態を大まかに示している。

【００３４】

まず、ステップＳ５１０においては、ステップＳ４００において、小さすぎるプリミティブホワイトエリアを除去した後に残ったプリミティブホワイトエリア６３０の各々が調べられる。図２１は、小さすぎるプリミティブホワイトエリアの全てが除去された後に残っ

ている多数のプリミティブホワイトエリア６３０Ａ乃至６３０Ｎを示す。ステップＳ５１２においては、プリミティブホワイトエリア６３０Ａ乃至６３０Ｎの内の一つ又はそれより多くの各グループへグループラベルがラベル付けされる。プリミティブホワイトエリア６３０Ａ乃至６３０Ｎは、互いに唯一に接触するプリミティブホワイトエリアが同じグループ内にあるようにグループ化される。図２２に示されているように、プリミティブホワイトエリア６３０Ａ乃至６３０Ｎは７個の異なるグループ４０乃至４６へグループ化される。

【００３５】

ステップＳ５１４において、グループラベルがプリミティブホワイトエリア６３０Ａ乃至６３０Ｎに割り当てられた後で、図２２に示されているように、同じグループラベルを有するプリミティブホワイトエリア６３０のグループが、単一グループ４１乃至４６へ集められる。次に、ステップＳ５１６において、水平グループに対して、各水平グループ４０乃至４６の各々の高さ６５０が高さ閾値６５０と比較される。グループの高さ４７は、図２３に示されているように、各プリミティブホワイトエリア６３０の幅の方向に垂直な方向におけるグループのサイズである。従って、垂直グループに対しては、グループの「高さ」４７は、水平方向におけるグループのサイズであり、垂直グループの高さ４７は、幅閾値６４０’と比較される。本発明の好ましい実施の形態において、高さ閾値６５０又は６４０’は、文書画像６００内のテキストの行間隔より大きな値にセットされる。

【００３６】

ステップＳ５１８において、高さ閾値６５０又は６４０’より小さな高さ４７を有しており、他のグループに接触しないグループは、除去される。図２２のグループ４０がこの例である。ステップＳ５２０において、高さ閾値６５０又は６４０’より小さな高さ４７を有しており、一つの面上だけ（グループのボトム上又はグループのトップ上だけ）で他のグループと接触している分岐されたグループは、除去される。図２３のグループ４５がこの例である。ステップＳ５２２において、高さ閾値６５０又は６４０’より小さく分割され、分岐されたグループを除去した後で、残っているグループの数は、グループ４１乃至４４、及び４６のうちの二つ又はそれより多くを結合することによって減少されることができる。

【００３７】

ステップＳ５２４においては、高さ閾値より小さな高さを有するプリミティブホワイトエリアの終端がトリミングされる。これは、図３１乃至図３３に示されており、以下に詳細に記述されている。

【００３８】

ステップＳ５２６においては、（他のグループに）接触しており、共通の左エッジと右エッジを有するグループが単一のプリミティブホワイトエリア６３０へ併合される。図２５及び図２６に示されているように、プリミティブホワイトエリア６３０Ｏ及び６３０Ｐ、並びに６３０Ｑ、６３０Ｒ、６３０Ｓ、及び６３０Ｔはそれぞれ共通の左右エッジを有している。三つのグループ内の７個のプリミティブホワイトエリアは、図２６に示されているように、三つのプリミティブホワイトエリア６３０Ｏ、６３０Ｑ、及び６３０Ｕへ併合される。

【００３９】

ステップＳ５２８において、高さ４７が高さ閾値６５０より小さい、除去されたプリミティブホワイトエリア６３０が再検討されて、これらの再検討されたプリミティブホワイトエリア６３０の内のいくつかが、ステップＳ５３０において、残っているプリミティブホワイトエリア６３０の内の一つへ戻り、併合されることができるかを決定する。

【００４０】

ステップＳ５３２において、各残っているプリミティブホワイトエリア６３０の高さ４７が、高さ閾値６５０又は６４０’と比較される。ステップＳ５３４において、高さ閾値６５０又は６４０’より大きな高さ４７を有するプリミティブホワイトエリア６３０が、主要ホワイト領域６６０であると決定される。

【００４１】
図２７は、ステップＳ５１２のプリミティブホワイトエリアへラベルを割り当てるプロセスの一つの好ましい実施の形態を大まかに示している。ステップＳ５１２１において、文書画像の上部に配置された図２１のプリミティブホワイトエリア６３０Ａが現行ホワイトエリアとして選択される。現行検討されているプリミティブホワイトエリアは現行プリミティブホワイトエリアと呼ばれる。ステップＳ５１２２において、現行プリミティブホワイトエリアは、ラベルがこの現行プリミティブホワイトエリアへ既に割り当てられたか否かを決定するためにチェックされる。既に割り当てられていた場合、コンロトールは飛び越してステップＳ５１２７へ進む。割り当てられていなかった場合、コントロールはステップＳ５１２３へ進む。

【００４２】
ステップＳ５１２３において、現行プリミティブホワイトエリアのトップに接触している他のプリミティブホワイトエリアの数が決定される。当然、最上部のプリミティブホワイトエリア６３０Ａに対してその数はゼロである。これ以外のプリミティブホワイトエリア６３０に対しては、現行プリミティブホワイトエリアのトップに接触している他のプリミティブホワイトエリア６３０の数は、ゼロ、１又はそれ以上であってもよい。次に、ステップＳ５１２４においては、ステップＳ５１２３において決定された現行プリミティブホワイトエリアのトップに接触している他のプリミティブホワイトエリア６３０の数が正確に１に等しいか否かがチェックされる。１に等しい場合、コントロールはステップＳ５１２５へ進み、このステップにおいて、現行プリミティブホワイトエリアのトップに接触している単一の他のプリミティブホワイトエリア６３０に割り当てられたラベルは現行プリミティブホワイトエリアへも割り当てられる。１に等しくないとき、現行プリミティブホワイトエリアのトップに接触しているゼロ又は一つより多くの他のプリミティブホワイトエリアが存在していれば、コントロールはステップＳ５１２６へ飛び越し、このステップにおいて、新しいラベルが現行プリミティブホワイトエリアへ割り当てられる。これは、最上部のプリミティブホワイトエリア６３０Ａに対しても当然の状況である。次いで、ステップＳ５１２５及びステップＳ５１２６は、共に、ステップＳ５１２７へ進む。

【００４３】
ステップＳ５１２７において、現行プリミティブホワイトエリアのボトムに接触している他のプリミティブホワイトエリア６３０の数が決定される。次いで、ステップＳ５１２８において、現行プリミティブホワイトエリアのボトムに接触している数が、１以下か否かを決定するためにチェックされる。１以下の場合、コントロールはステップＳ５１３０へ飛び越す。１以下でない場合、コントロールはステップＳ５１２９へ進み、このステップにおいて、現行プリミティブホワイトエリアのボトムに接触している他のプリミティブホワイトエリア６３０の各々には新しくて異なるラベルが割り当てられる。

【００４４】
コントロールはステップＳ５１３０へ進み、このステップにおいて、文書画像は、ラベルが文書６００のプリミティブホワイトエリア６３０のどれかに割り当てられるべきかを決定するためにチェックされる。ラベルを割り当てるように決定された場合、コントロールはステップＳ５１３１へ戻り、このステップにおいて、最上部の割当てられていないプリミティブホワイトエリア６３０が現行プリミティブホワイトエリアとして選択される。次いで、コントロールは、ステップＳ５１３０からステップＳ５１２２へ戻る。ラベルが割り当てられない場合、割当てられていないプリミティブホワイトエリア６３０が残っていなければ、コントロールはステップＳ５１３２へ進み、このステップにおいて、コントロールはＳ５１４へ戻る。

【００４５】
図２８は、ステップＳ５２２のプリミティブホワイトエリアグループをグループ化する方法のための一つの好ましい実施の形態を大まかに示している。ステップＳ５２２１において、最上部及び最左部の検討されていないホワイトエリアグループが現行プリミティブホワイトエリアグループとして選択される。ステップＳ５２２２において、現行プリミティ

ブホワイトエリアグループのボトムに接触している他のプリミティブホワイトエリアグループの数が決定される。ステップＳ５２２３において、現行プリミティブホワイトエリアグループのボトムに接触している他のプリミティブホワイトエリアグループの数が１に等しくない場合、コントロールはステップＳ５２２７へ飛び越す。或いは、その数が１に等しい場合、コントロールはステップＳ５２２４へ進む。このステップにおいて、現行グループのボトムに接触している単一グループのトップに接触しているグループの合計が決定される。次いで、ステップＳ５２２５において、接触しているグループの合計がぴったり１に等しい場合、コントロールはステップＳ５２２６へ進む。１に等しくない場合、コントロールは再びステップＳ５２２７へ進む。ステップＳ５２２６において、現行プリミティブホワイトエリアグループのボトムに接触している単一プリミティブホワイトエリアグループに割り当てられたラベルは、現行プリミティブホワイトエリアに割り当てられたラベルと置き換えられ、これによって、二つのグループが併合される。次いで、コントロールはステップＳ５２２２へ戻る。

【００４６】

ステップＳ５２２７において、文書画像６００は、検討されていないプリミティブホワイトエリアグループが当該文書画像６００内に残っているか否かを決定するためにチェックされる。検討されていないプリミティブホワイトエリアグループが当該文書画像６００内に残っている場合、コントロールは、次のグループが現行グループとして選択されるステップＳ５２２８へ進む。次いで、コントロールはステップＳ５２２２へ飛び越して戻る。検討されていないプリミティブホワイトエリアグループが当該文書画像６００内に残っていない場合、コントロールはステップＳ５２２９へ飛び越し、このステップはコントロールをステップＳ５２４へ戻す。

【００４７】

図２９及び図３０は、ステップＳ５２４１のグループ内のプリミティブホワイトエリアの終端をトリミングするための手順の一つの好ましい実施の形態を大まかに示している。まず、ステップＳ５２４において、プリミティブホワイトエリアグループの第１のグループが、現行プリミティブホワイトエリアグループとして選択される。次に、ステップＳ５２４２において、現行プリミティブホワイトエリアグループの最左側（レフトサイド）が現行最左側として識別され、当該最左側を有する現行プリミティブホワイトエリアグループのプリミティブホワイトエリアが現行エリア又は現行サブグループとして識別される。次いで、ステップＳ５２４３において、プリミティブホワイトエリアの現行グループの現行プリミティブホワイトエリア又は現行サブグループの現行最左側の高さ４７が、高さ閾値６５０又は６４０’と比較される。高さ閾値６５０又は６４０’が越えられない場合、コントロールはステップＳ５２４４へ進み、このステップにおいて、現行グループが、現行エリア又はサブグループに接触しているプリミティブホワイトエリアが残っているか否かを見るためにチェックされる。

【００４８】

接触しているプリミティブホワイトエリアが残っている場合、コントロールはステップＳ５２４５へ進む。接触しているプリミティブホワイトエリアが残っていない場合、コントロールはステップＳ５２４６へ飛び越す。ステップＳ５２４５において、現行エリア又はサブグループに接触している現行プリミティブホワイトエリアの次の最左側が現行最左側として識別される。このプリミティブホワイトエリアは現行サブグループ又は現行エリアへ追加されて、現行サブグループを形成する。次いで、コントロールはステップＳ５２４３へ戻る。このループは、現行エリア又はサブグループの現行最左側の高さが高さ閾値６５０又は６４０’を越えなくなる迄、又は接触しているプリミティブホワイトエリアが無くなる迄、適所に残る。次いで、コントロールはステップＳ５２４６へ進む。

【００４９】

ステップＳ５２４６において、現行プリミティブホワイトエリア又はサブグループの現行最左側の左へ延びる現行プリミティブホワイトエリアグループ内のプリミティブホワイトエリアの終端が除去される。次いで、コントロールはステップＳ５２４７へ進み、このス

テップは、左端が未だトリミングされていないと共に当該左端の高さが高さ閾値に等しいか又はそれより小さい現行グループの他のプリミティブホワイトエリアがあるか否かを決定する。左端がトリミングを必要とする他のエリアが残っている場合、コントロールはステップＳ５２４７からステップＳ５２４２へ戻る。

【００５０】

左端がトリミングを必要とする他のエリアが無い場合、コントロールはステップＳ５２４８へ進み、このステップは、現行プリミティブホワイトエリアグループのプリミティブホワイトエリアの最右側（ライトサイド）を現行最右側として識別する。コントロールは次いでステップＳ５２４９へ進み、このステップにおいて、現行最右側の高さ４７は、高さ閾値６５０又は６４０’と比較される。高さ閾値６５０又は６４０’を越えない場合、コントロールはステップＳ５２５０へ進み、このステップにおいて、現行グループは、現行エリア又はサブグループに接触しているプリミティブホワイトエリアが残っているか否かを見るためにチェックされる。

【００５１】

現行エリア又はサブグループに接触しているプリミティブホワイトエリアが残っている場合、コントロールはステップＳ５２５１へ進む。或いは、現行エリア又はサブグループに接触しているプリミティブホワイトエリアが残っていない場合、コントロールはステップＳ５２５２へ飛び越す。ステップＳ５２５１において、現行プリミティブホワイトエリアグループの現行ホワイトエリア又はサブグループに接触しているプリミティブホワイトエリアの次の最右側が、現行最右側として識別される。コントロールは次いでステップＳ５２４９へ戻る。このループは、高さ閾値が越えられなくなる迄、又は接触しているプリミティブホワイトエリアが無くなる迄、続けられる。次いで、コントロールはステップＳ５２５２へ進み、このステップにおいて、現行最右側の右へ延びる現行プリミティブホワイトエリア又はサブグループのプリミティブホワイトエリアの終端が除去される。

【００５２】

コントロールはステップＳ５２５３へ進み、このステップは、右端が未だトリミングされていないと共に当該右端の高さが高さ閾値に等しいか又はそれより小さい現行グループの他のプリミティブホワイトエリアがあるか否かを決定する。右端がトリミングを必要とする他のエリアがある場合、コントロールはステップＳ５２４８へ飛び越して戻る。或いは、トリミングを必要とする他のエリアが無い場合、コントロールはステップＳ５２５４へ進み、このステップは、トリミングされていないプリミティブホワイトエリアグループが文書画像６００内に残っているか否かを決定する。トリミングされていないプリミティブホワイトエリアグループが文書画像６００内に残っている場合、コントロールはステップＳ５２５５へ進む、このステップにおいて、次のプリミティブホワイトエリアグループが現行プリミティブホワイトエリアグループとして選択される。コントロールは次いでステップＳ５２４２へ戻る。或いは、文書画像６００のプリミティブホワイトエリアグループの全てがトリミングされている場合、コントロールはステップＳ５２５６へ進み、このステップはコントロールをステップＳ５２６へ戻す。

【００５３】

図３１乃至図３３は、図２９及び図３０に示されたトリミングプロセスの動作をグラフィカルに示している。図３１に示されているように、グループ４８は、個々のプリミティブホワイトエリア６３１乃至６３７を備える。最初に、プリミティブホワイトエリア６３５及び６３６は、グループ４８の最左端を画定する。プリミティブホワイトエリア６３５及び６３６の共通左端の垂直高さ４７は、高さ閾値６５０と比較される。図３１に示されているように、プリミティブホワイトエリア６３６の最下端から測定される時、高さ閾値６５０は、プリミティブホワイトエリア６３４の上端を越えているがプリミティブホワイトエリア６３３の上端より下にある。従って、プリミティブホワイトエリア６３５及び６３６は、現行サブグループとして一緒にグループ化され、この現行サブグループに接触しているプリミティブホワイトエリアが識別される。図３１においては、プリミティブホワイトエリア６３４及び６３７を含む。

【００５４】
従って、図３２に示されているように、プリミティブホワイトエリア６３５及び６３６の左端は、プリミティブホワイトエリア６３４の左端へ戻るようにトリミングされる。この新しい左端の高さは、高さ閾値６５０と比較される。高さ閾値６５０が越えられていないので、プリミティブホワイトエリア６３４乃至６３６の左端は、プリミティブホワイトエリア６３３の左端に戻るようにトリミングされる。同様に、プリミティブホワイトエリア６３１の左端もプリミティブホワイトエリア６３２の左端に戻るようにトリミングされる。現行エリア又はグループの左端は、現行エリア又はグループに接触しているプリミティブホワイトエリアの次の最左端に合わせて調整されるか又はトリミングされる。

【００５５】
図３１乃至図３３において、プリミティブホワイトエリア６３２は、プリミティブホワイトエリア６３１に接触しているが、他のプリミティブホワイトエリア６３３乃至６３７は、プリミティブホワイトエリア６３１に接触していない。プリミティブホワイトエリア６３１の高さ４７は、高さ閾値６５０に比較するには及ばないが、当該プリミティブホワイトエリア６３１は、グループ４８の最左点が高さ閾値６５０を越えないプリミティブホワイトエリアのプリミティブホワイト領域のトリミングされたサブグループの最左端より更に左へ延びないように、戻るようにトリミングされる。図３３に示されているように、プリミティブホワイトエリア６３３乃至６３６によって画定されたグループ４８の最左端は、高さ閾値６５０を越えないので、グループ４８の最左端に対するトリミングプロセスが終了する。

【００５６】
次いで、このプロセスは、グループ４８の最右端に対して繰り返される。この場合、プリミティブホワイトエリア６３６及び６３７の最右端によって画定されたグループ４８の最右端は、まず、プリミティブホワイトエリア６３１及び６３２の最右端と同一線になるように戻るようにトリミングされる。これらのプリミティブホワイトエリアは隣接しておらず、どれも単独では高さ閾値６５０を越えないので、プリミティブホワイトエリア６３１、６３２、６３６、及び６３７の最右端は、プリミティブホワイトエリア６３１乃至６３４の最右端へ戻るように更にトリミングされる。プリミティブホワイトエリア６３１乃至６３４によって画定された最右端が高さ閾値４７を越えているので、トリミングプロセスは終了する。

【００５７】
図３４は、二つ又はそれより多くのプリミティブホワイトエリア６３０を単一のプリミティブホワイトエリア６３０へ併合するプロセスの一つの好ましい実施の形態を示している。ステップＳ５３０１において、第１のプリミティブホワイトエリア６３０が現行プリミティブホワイトエリアとして選択される。ステップＳ５３０１において、現行ホワイトエリアが、その垂直高さが垂直閾値より大きいか又はそれより小さいかを決定するためにチェックされる。現行ホワイトエリアの垂直高さが垂直閾値より大きい場合、コントロールはステップＳ５３０２を飛び越してステップＳ５３１１へ進む。或いは、現行ホワイトエリアの垂直高さが垂直閾値より小さい場合、コントロールはステップＳ５３０３へ進む。

【００５８】
ステップＳ５３０３において、現行プリミティブホワイトエリアに接触している他のプリミティブホワイトエリアのリストが発生される。次いで、ステップＳ５３０４において、接触エリアのリストは、二つの接触エリアが現行ホワイトエリアの同じ部分に接触しているか否かを決定するためにチェックされる。例えば、図３５に示されているように、接触しているプリミティブホワイトエリア６３１及び６３３は、現行ホワイトエリ６３２の同じ部分に接触している。同様に、接触しているプリミティブホワイトエリア６３４及び６３５も現行ホワイトエリア６３２の同じ部分に接触している。現行ホワイトエリアの同じ部分に接触している、リスト上のペアの接触エリアが存在しない場合、コントロールはステップＳ５３０８へ飛び越す。或いは、現行ホワイトエリアの同じ部分に接触している二つの接触エリアが存在している場合、コントロールはステップＳ５３０５へ進む。

【００５９】
ステップＳ５３０５において、ペアの接触エリアのエッジが充分に位置合わせされているか否かを決定するためにチェックされる。例えば、図３５に示されているように、接触しているホワイトエリア６３１及び６３３のエッジは充分に位置合わせされているが、接触エリア６３４及び６３５のエッジは充分に位置合わせされていない。ステップＳ５３０４の通過毎に二つの接触エリアの一つのペアが選択され、ステップＳ５３０５において、この選択されたペアがチェックされる。ステップＳ５３０５において、エッジが充分に位置合わせされていない場合、コントロールはステップＳ５３０４へ飛び越して戻り、このステップは、接触エリアのペアの内の他の一つをチェックする。ステップＳ５３０５を介してステップＳ５３０４へ戻るこのループは、各接触ペアがチェックされるまで続けられる。

【００６０】
ステップＳ５３０５において、ペアの接触エリアのエッジが充分に位置合わせされている場合、コントロールはステップＳ５３０６へ進み、このステップにおいて、ペアの接触エリアが接触エリアのリストから除去され、図３６に示されているように、現行ホワイトエリアは三つの部分へ分割される。当然、ペアの接触エリアの両左端又は両右端が、それぞれ、現行ホワイトエリアの左端又は右端と位置合わせされている場合、現行ホワイトエリアは二つの部分に分割されるにすぎない。ステップＳ５３０７において、これら二つの接触エリアと現行ホワイトエリアの接触された部分が単一ホワイトエリアへ併合される。このプロセスは、図３６及び図３７に示され、これらの図において、接触エリア６３２及び６３３並びに現行ホワイトエリア６３２ｃが単一ホワイトエリア６３１へ併合される。次いで、ステップＳ５３０４へ戻り、コントロールは接触エリアのペアの内の他の一つのペアをチェックする。

【００６１】
ステップＳ５３０４乃至Ｓ５３０７のループにおいて、全ての接触エリアのペアが一旦分析されると、コントロールはステップＳ５３０４からステップＳ５３０８へ飛び越す。ステップＳ５３０８において、接触エリアのリストが、当該リストにエリアが残っているか否かを決定するためにチェックされる。接触しているエリアがリスト上に残っていない場合、コントロールはステップＳ５３０８からステップＳ５３１１へ進む。そうでなく、接触しているエリアのリストにエリアが残っている場合、コントロールはステップＳ５３０９へ進む。

【００６２】
接触しているエリアのリストに残っている最も幅の広い接触エリアが選択され、当該選択された最も幅の広いエリアが接触している現行ホワイトエリアの部分が更に三つの部分に分割される。図３６及び図３７において、最も幅の広い残っている部分６３４が、現行ホワイトエリア６３２の部分６３２ｂに接触している。従って、部分６３２ｂは、更に、非接触部分６３２ｄ及び接触された部分６３２ｃへ分割される。図３８に示されているように、接触された部分６３２ｃは最も幅の広いエリア６３４へ併合される。

【００６３】
ステップＳ５３１０において、選択された最も幅の広いエリアがリストから除去され、当該選択された最も幅の広いエリアと同じ現行プリミティブホワイトエリアの部分に接触しているリスト上の他のエリアも、接触エリアのリストから除去される。コントロールは、ステップＳ５３１０からステップＳ５３０８へ戻って、接触しているエリアのリストにエリアが残っているか否かを再び決定する。リスト上に残っている全ての接触しているエリアが一旦選択され、これらの選択されたエリアが接触している現行ホワイトエリアの部分がそれらに併合されると、コントロールは、ステップＳ５３０８からステップＳ５３１１へ飛び越し、このステップＳ５３１１は、選択されていないプリミティブホワイトエリアが文書画像内に残っているか否かを決定する。選択されていないプリミティブホワイトエリアが文書画像内に残っている場合、コントロールはステップＳ５３１１からステップＳ５３１３へ飛び越し、このステップにおいて、次のプリミティブホワイトエリアが現行ホ

ワイトエリアとして選択される。ステップＳ５３１３において、次のプリミティブホワイトエリアが現行ホワイトエリアとして選択されると、コントロールは飛ばしてステップＳ５３０２へ戻り、このステップにおいて、現行ホワイトエリアが垂直閾値に対して再度チェックされる。そうではなく、文書画像のプリミティブホワイトエリアの全てが選択され、垂直閾値と比較された場合、コントロールはステップＳ５３１２へ進み、このステップはコントロールをステップＳ５３２へ戻す。

【００６４】

図３５に示されているように、現行ホワイトエリア６３２には、６個の他のプリミティブホワイトエリア６３１、及び６３３乃至６３７が接触している。プリミティブホワイトエリア６３１及び６３４は、垂直閾値と前もって比較され、これらのエリアが閾値より小さい場合、前もって解析される。この場合、接触しているプリミティブホワイトエリア６３１又は６３４のいづれも垂直閾値より小さい。上記のように、接触しているホワイトエリアのリストが一旦発生されると、現行プリミティブホワイトエリア６３２の共通部分に接触し且つ充分に位置合わせしているエッジを有するこれらの接触しているプリミティブホワイトエリアは、単一ホワイトエリアへ併合される。従って、図３６に示されているように、接触しているプリミティブホワイトエリア６３１及び６３３は、現行部分６３２の共通に接触している部分６３２ｃに沿って共に併合される。また、接触しているプリミティブホワイトエリア６３１及び６３３のエッジは、共通に接触している部分６３２ｃのエッジに位置合わせされるように、戻るようにトリミングされる。現行プリミティブホワイトエリアの共通部分にも接触している他の接触しているホワイトエリア６３４及び６３５のエッジは、充分に位置合わせされていないので、これらは併合されない。

【００６５】

図３７に示されているように、接触しているエリアのリストに最も広い残っている接触エリアが選択され、現行ホワイトエリア部分６３２の接触された部分６３２ｃが、選択された最も（幅の）広い残っているホワイトエリア６３４に併合される。同時に、当該接触されたエリア６３２ｃにも接触している接触しているエリアのリスト上のあらゆる他の接触エリアが接触エリアのリストから除去される。最も広い残っている接触エリアを検索するプロセスは、接触しているプリミティブホワイトエリアがプリミティブホワイトエリアのリストから無くなるまで続けられる。従って、図３８に示されているように、現行エリア６３２の接触された部分が接触ホワイトエリア６３６及び６３７へ併合され、現行ホワイトエリアは接触されていない部分６３２ｄ、６３２ｆ、及び６３２ｇへ更に分割される。

【００６６】

図３９は、ステップＳ１５０の文書エレメントを抽出する方法の一つの好ましい実施の形態を大まかに示している。図３９に示されているように、文書画像６７０を抽出するために、閉ループを形成しない主要ホワイト領域６６０は除去される。ステップＳ８００において、残っている主要ホワイト領域６６０がセグメントへ分割される。ステップＳ９００において、最も外側のセグメント、即ち、文書画像６００の境界に最も近接しているセグメントが第１の方向においてトラッキングされ、各トラッキングされたセグメントの第１の方向のフィールドがマーキングされる。ステップＳ１０００において、セグメントのリストから第１の最上部の水平セグメントが第１の及び現行セグメントとして選択される。

【００６７】

次に、ステップＳ１１００において、第１のセグメントを含む現行閉ループは、前記第１の方向と反対の方向に移動することによって発生され、この方向に対応している各セグメントの方向フィールドがマーキングされる。ステップＳ１２００において、マーキングされた両方向のフィールドを有するセグメントは除去される。ステップＳ１３００において、現行閉ループによって定義された現行文書エレメント６７０が出力される。ステップＳ１４００においては、セグメントのリストがセグメントが当該リストに残っているか否かを決定するためにチェックされる。セグメントが当該リストに残っている場合、コントロールはステップＳ１５００へ進み、このステップにおいて、当該リストに残っている最上部の水平セグメントが当該セグメントのリストから選択される。コントロールはステップ

Ｓ１１００へ進む。当該セグメントのリストにセグメントが残っていない場合、コントロールはステップＳ１４００からステップＳ１６００へ進み、このステップはコントロールをステップＳ１６０へ戻す。

【００６８】

図４０は、ステップＳ７００の閉ループを形成しない主要ホワイト領域を除去するプロセスの一つの好ましい実施の形態を大まかに示している。最初に、ステップＳ７１０においては、少なくとも二つの他の主要ホワイト領域６６０と接触していない主要ホワイト領域６６０が検出される。これは、図４４にグラフィカルに示されており、この図において、主要ホワイト領域６６０乃至６６７の内の主要ホワイト領域６６６と６６７が、それぞれ、一つだけの他の主要ホワイト領域と接触している。従って、これらの領域はステップＳ７２０において除去される。次いで、ステップＳ７３０において、文書画像６００は、少なくとも二つの他の領域に接触できない領域が残っているか否かを決定するために再度チェックされる。図４４に示されているように、主要ホワイト領域６６５が、二つの他の主要ホワイト領域６６１と６６７とに最初に接触した。しかしながら、主要ホワイト領域６６７が前もって除去されているので、主要ホワイト領域６６５は、この場合、一つだけの他の主要ホワイト領域と接触している。

【００６９】

少なくとも二つの他の主要ホワイト領域に接触できない主要ホワイト領域が残っている場合、コントロールはステップＳ７３０からステップＳ７１０へ戻る。少なくとも二つの他の主要ホワイト領域に接触できない主要ホワイト領域が残っていない場合、コントロールはステップＳ７３０からステップＳ７４０へ進み、このステップにおいて、コントロールがステップＳ８００へ戻される。これにより、図４４における文書画像６００は、図４５に示されているように、主要ホワイト領域６６１乃至６６４を残して除去された主要ホワイト領域６６５乃至６６６を有する。

【００７０】

図４１は、ステップＳ８００の残っている主要ホワイト領域６６０をセグメントへ分割するためのプロセスの一つの好ましい実施の形態を示している。最初に、ステップＳ８１０においては、主要ホワイト領域６６０の内の一つが現行ホワイト領域として選択される。次に、ステップＳ８２０においては、現行主要ホワイト領域６６０上の交差点、即ち、当該現行主要ホワイト領域６６０が他の主要ホワイト領域６６０と接触する場所が検出される。ステップＳ８３０において、現行主要ホワイト領域が各交差点で分割されて、セグメントを形成する。このプロセスは、図４６及び図４７にグラフィカルに示されている。文書画像６００は主要ホワイト領域６６１乃至６６７を備える。図４７に示されているように、主要ホワイト領域６６１は、それが、主要ホワイト領域６６２、６６４、及び６６５と交差する三つの交差点を有する。従って、主要ホワイト領域６６１は、四つのセグメント６６１ａ乃至６６１ｄへ分割される。同様に、主要ホワイト領域６６２は、それが、主要ホワイト領域６６１、６６６、及び６６３と交差される三つの交差点を有する。従って、主要ホワイト領域６６２は、セグメント６６２ａ乃至６６２ｄへ分割される。同様に、主要ホワイト領域６６３乃至６６７が、セグメント６６３ａ乃至６６３ｃ、６６４ａ乃至６６４ｄ、６６５ａ乃至６６５ｃ、６６６ａ乃至６６６ｃ、及び６６７ａ乃至６６７ｃへ分割される。

【００７１】

次いで、ステップＳ８４０において、図４８に示されているように、単一交差点のみによって境界付けされた、文書画像６００のセグメントは、除去される。図４８に示されているように、セグメント６６１ａ、６６１ｄ、６６２ａ、６６２ｄ、６６３ａ、６６３ｃ、６６４ａ、６６４ｄ、６６５ａ、６６５ｄ、６６６ａ、６６６ｃ、６６７ａ、乃び６６７ｃは、交差点によって境界付けされない一端を有しているので、文書画像６００から除去される。従って、これは、文書画像６６０上に残っているセグメント６６１ｂ乃至ｃ、６６２ｂ乃至ｃ、６６３ｂ、６６４ｂ乃至ｃ、６６５ｂ乃至ｃ、６６６ｂ及び６６７ｂを残す。

【００７２】
次に、ステップＳ８５０において、現行主要ホワイト領域がセグメントされ、その境界付けされていないセグメントが除去された後で、文書画像６００は、まだ分割されていない主要ホワイト領域が残っているか否かを決定するためにチェックされる。まだ分割されていない主要ホワイト領域が残っている場合、コントロールはステップＳ８６０へ進み、このステップにおいて、次の分割されていない主要ホワイト領域が現行領域として選択される。コントロールは飛び越してステップＳ８２０へ戻る。文書画像６００内の主要な未分割のホワイト領域が残っていない場合、コントロールはステップＳ８５０からステップＳ８７０へ飛び越し、このステップはコントロールをステップＳ９００へ戻す。
【００７３】
図３８は、ステップＳ９００の最も外側のセグメントのリストを発生するための方法の一つの好ましい実施の形態を示す。各セグメントは、このセグメントが各方向（水平セグメントの場合、左から右、右から左、垂直セグメントの場合、上から下、下から上）においてトラックされていることを示すために、当該セグメントに対応している二つのフィールドを有している。最初に、ステップＳ９１０において、最上部及び最下部の水平セグメントが、第１のセグメント及び現行セグメントとして選択される。水平サポートを選択することは必要とされず、その代わりに垂直セグメントが選択されることが理解されよう。選択されたセグメントにおいて、左から右への方向を示すフィールドがマーキングされる。図４９に示された文書画像６００の場合、これはセグメント６６２ｂである。
【００７４】
次に、ステップＳ９２０において、第１のセグメントの右交差点が最終の交差点として選択される。図４９において、これは、セグメント６６２ｂと６６２ｃを結ぶ交差点である。
【００７５】
ステップＳ９４０において、現行セグメントは、現行セグメントの次の交差点に出会うまで、最終交差点から現行セグメントに沿って反時計回りに移動することによってトラックされる。ステップＳ９５０において、次のセグメントは、トラッキング（追跡）の反対方向から出発しながら、セグメントが見つかるまで、次の交差点の回りを反時計回りに探索することによって選択される。図４９において、これはセグメント６６１ｂであり、セグメント６６２ｂの左手交差点回りの動作がラインＡで示されている。トラックされた方向、右から左の方向に対応しているフィールドがマーキングされる。
【００７６】
ステップＳ９６０において、次のセグメントがそれが第１のセグメントであるか否かを見るためにチェックされる。第１のセグメントでなければ、コントロールはステップＳ９７０へ進み、このステップにおいて、次の交差点が最終の交差点として選択され、次のセグメントが現行セグメントとして選択される。コントロールは次いでステップＳ９３０へ進む。このプロセスは、セグメント６６１ｃ、６６４ｂ、６６４ｃ、６６３ｂ、及び６６２ｃをトラックする。各セグメント内のこれらのトラックされた方向に対応しているフィールドがマーキングされる。
【００７７】
しかしながら、ステップＳ９６０において、次のセグメントが第１のセグメントである場合、コントロールはステップＳ９８０へ飛び越し、このステップはコントロールをステップＳ１０００（図３９）へ戻す。これらのセグメントの内のどれかがステップＳ９１０の開始点として選択され得ることが理解されよう。
【００７８】
図４３はステップＳ１１００の現行の最も内側の閉ループを発生するための方法の一つの好ましい実施の形態を示している。ステップＳ１１１０において、セグメントのリストの最上部の水平セグメントは、第１のループセグメントとして及び現行ループセグメントとして選択され、現行の最も外側のセグメントの左交差点は最終セグメントとして選択される。ステップＳ１１２０において、現行ループセグメントは、現行ループセグメントの次

の交差点が見つかるまで、最終交差点から時計回りにトラックすることによって走査される。ステップＳ１１３０において、トラッキングの反対方向から開始しながら、セグメントが見つかるまで、次の交差点の回りを反時計回りに探索することによって次のセグメントが選択される。図５０に示されているように、ステップＳ１１３０において形成されたラインは、セグメント６６２ｂの右手交差点の回りのラインＢに沿って反時計回りに回転される。

【００７９】

次いで、ステップＳ１１４０においては、次のループセグメントが、それが第１のループセグメントであるか否かを見るためにチェックされる。次のループセグメントが第１のループセグメントでなければ、コントロールはステップＳ１１５０へ進み、このステップにおいて、次のループセグメントが現行ループセグメントとして選択され、次の交差点が最終交差点として選択される。このプロセッスは、図５１に示されており、この図において、セグメント６６２ｂ、６６６ｂ、６６５ｃ、６６５ｂ、及び６６１ｃが、第１の文書エレメント６７１を取り囲むインターループを形成する。各セグメント内のトラックされた方向に対応しているフィールドがマーキングされる。

【００８０】

ステップＳ１１４０において、次のセグメントが第１のセグメントである場合、コントロールはステップＳ１１４０へ飛び越し、このステップにおいて、現行閉ループは終了したものとして処理され、この閉ループの内部が次の文書エレメントとして識別される。次いで、コントロールはステップＳ１１７０へ進み、このステップはコントロールをステップＳ１２００（図３９）へ戻す。

【００８１】

図５２及び図５３は、第２及び第３の文書エレメント６７２及び６７３を識別するためのプロセスを示している。図５２に示されているように、第１の文書エレメント６７１が一旦検出されると、セグメント６６１ｂ及び６６２ｂが除去される。ステップＳ１２００に対して上記に示されているように、最も外側のループを見つけるにあたって、これらの各セグメントの両フィールドが既にマーキングされていたので、これらのセグメント６６１ｂ及び６６２ｂは除去される。閉ループが反時計回り方向に追跡されている場合、次の交差点からのラインが時計回りに回転されることが理解されよう。

【００８２】

【発明の効果】

本発明は、文書画像上のホワイト領域のみを解析することによって文書画像内のエレメントを識別するためのシステムを提供する。

【図面の簡単な説明】

【図１】本発明の文書分割システムの好ましい実施の形態を示すブロック図である。

【図２】主要ホワイト領域抽出手段の好ましい実施の形態を示すブロック図である。

【図３】サンプル文書画像を示す図である。

【図４】抽出された水平プリミティブホワイトエリアを有する文書画像を示す図である。

【図５】抽出された垂直プリミティブホワイトエリアを有する文書画像を示す図である。

【図６】抽出された主要ホワイト領域を有する文書画像を示す図である。

【図７】分割され、抽出された文書エレメントを示す図である。

【図８】主要ホワイト領域及び文書エレメントを抽出する方法の一つの実施の形態を示すフローチャートである。

【図９】主要ホワイト領域を抽出する方法の一つの実施の形態を示すフローチャートである。

【図１０】プリミティブホワイトエリアを抽出する方法の一つの実施の形態を示すフローチャートである。

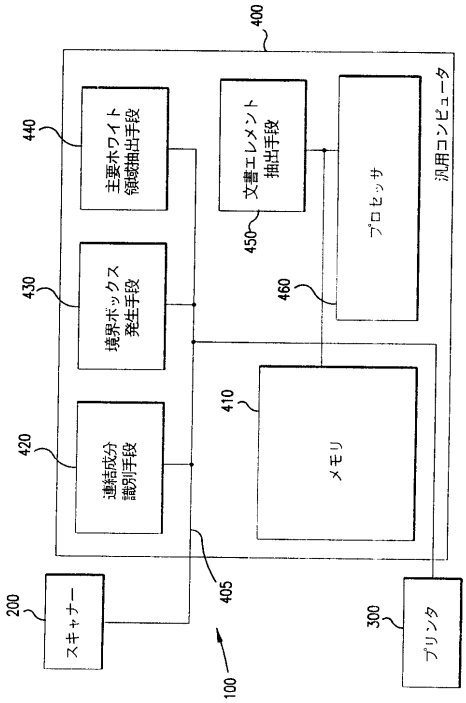
【図１１】プリミティブホワイトエリアを抽出する方法の一つの実施の形態を示すフローチャートである。

【図１２】境界ボックスのトップが走査線内で出会った時にプリミティブホワイトエリア

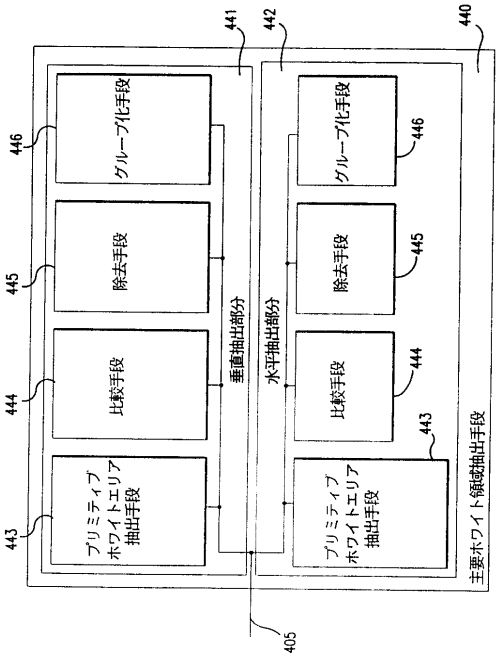
が形成され、プリミティブホワイトギャップが検出される種々の方法をグラフィカルに示す図である。
【図１３】境界ボックスのトップが走査線内で出会った時にプリミティブホワイトエリアが形成され、プリミティブホワイトギャップが検出される種々の方法をグラフィカルに示す図である。
【図１４】境界ボックスのトップが走査線内で出会った時にプリミティブホワイトエリアが形成され、プリミティブホワイトギャップが検出される種々の方法をグラフィカルに示す図である。
【図１５】境界ボックスのトップが走査線内で出会った時にプリミティブホワイトエリアが形成され、プリミティブホワイトギャップが検出される種々の方法をグラフィカルに示す図である。
【図１６】境界ボックスのボトムが走査線内で出会った時にプリミティブホワイトエリアが形成され、プリミティブホワイトギャップが検出される種々の方法をグラフィカルに示す図である。
【図１７】境界ボックスのボトムが走査線内で出会った時にプリミティブホワイトエリアが形成され、プリミティブホワイトギャップが検出される種々の方法をグラフィカルに示す図である。
【図１８】境界ボックスのボトムが走査線内で出会った時にプリミティブホワイトエリアが形成され、プリミティブホワイトギャップが検出される種々の方法をグラフィカルに示す図である。
【図１９】残っているプリミティブホワイトエリアを主要ホワイト領域へグループ化する方法の一つの実施の形態を示すフローチャートである。
【図２０】残っているプリミティブホワイトエリアを主要ホワイト領域へグループ化する方法の一つの実施の形態を示すフローチャートである。
【図２１】プリミティブホワイトエリアがラベル付けされ、除去され、グループ化される種々の方法をグラフィカルに示す図である。
【図２２】プリミティブホワイトエリアがラベル付けされ、除去され、グループ化される種々の方法をグラフィカルに示す図である。
【図２３】プリミティブホワイトエリアがラベル付けされ、除去され、グループ化される種々の方法をグラフィカルに示す図である。
【図２４】プリミティブホワイトエリアがラベル付けされ、除去され、グループ化される種々の方法をグラフィカルに示す図である。
【図２５】プリミティブホワイトエリアがラベル付けされ、除去され、グループ化される種々の方法をグラフィカルに示す図である。
【図２６】プリミティブホワイトエリアがラベル付けされ、除去され、グループ化される種々の方法をグラフィカルに示す図である。
【図２７】グループラベルをプリミティブホワイトエリアへ割り当てる方法の一つの実施の形態を示すフローチャートである。
【図２８】再グループ化のために残っているグループを検査する方法の一つの実施の形態を示すフローチャートである。
【図２９】グループ内のプリミティブホワイトエリアの終端をトリミングする方法の一つの実施の形態を示すフローチャートである。
【図３０】グループ内のプリミティブホワイトエリアの終端をトリミングする方法の一つの実施の形態を示すフローチャートである。
【図３１】プリミティブホワイトエリアの終端がどのようにトリミングされ、併合されるかをグラフィカルに示す図である。
【図３２】プリミティブホワイトエリアの終端がどのようにトリミングされ、併合されるかをグラフィカルに示す図である。
【図３３】プリミティブホワイトエリアの終端がどのようにトリミングされ、併合されるかをグラフィカルに示す図である。

【図３４】隣接しているプリミティブホワイトエリアを単一プリミティブホワイトエリアへ併合する方法の一つの実施の形態を示すフローチャートである。
【図３５】プリミティブホワイトエリアが他の隣接しているプリミティブホワイトエリアにどのように併合されるかをグラフィカルに示す図である。
【図３６】プリミティブホワイトエリアが他の隣接しているプリミティブホワイトエリアにどのように併合されるかをグラフィカルに示す図である。
【図３７】プリミティブホワイトエリアが他の隣接しているプリミティブホワイトエリアにどのように併合されるかをグラフィカルに示す図である。
【図３８】プリミティブホワイトエリアが他の隣接しているプリミティブホワイトエリアにどのように併合されるかをグラフィカルに示す図である。
【図３９】文書画像を抽出するための方法の一つの実施の形態を示すフローチャートである。
【図４０】閉ループを形成しない主要ホワイト領域を除去する方法の一つの実施の形態を示すフローチャートである。
【図４１】残っている主要ホワイト領域をセグメントへ分割するための方法の一つの実施の形態を示すフローチャートである。
【図４２】最も外側のセグメントのリストを発生する方法の一つの実施の形態を示すフローチャートである。
【図４３】最も内側のセグメントのリストを発生する方法の一つの実施の形態を示すフローチャートである。
【図４４】主要ホワイト領域が除去される前に主要ホワイト領域へ分割された文書画像を示す図である。
【図４５】除去された閉ループを形成しない前記主要ホワイト領域を有する図４４の文書画像を示す図である。
【図４６】残っている主要ホワイト領域がセグメントへ分割される前に他の文書画像を示す図である。
【図４７】セグメントへ分割された図４６の文書画像を示す図である。
【図４８】除去された唯一の交差点によって境界付けされたセグメントを有する図４７のセグメントされた文書画像を示す図である。
【図４９】発生された最も外側のセグメントのリストを有する図４８の文書画像を示す図である。
【図５０】内部の閉ループを形成する方法をグラフィカルに示す図である。
【図５１】第１の閉ループが終了した後の図４８の文書画像を示す図である。
【図５２】第２の閉ループが終了した後の図４８の文書画像を示す図である。
【図５３】第３及び最終の閉ループが終了した後の図４８の文書画像を示す図である。
【符号の説明】
１００　文書エレメント分割システム
４１０　メモリ
４２０　連結成分識別手段
４３０　境界ボックス発生手段
４４０　主要ホワイト領域抽出手段
４５０　文書エレメント抽出手段
４６０　プロセッサ

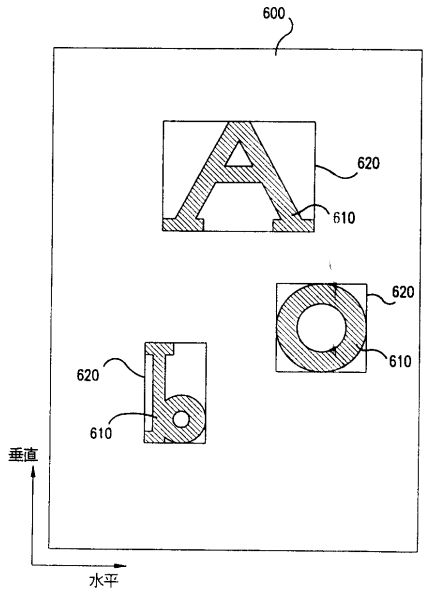
【図１】


400
汎用コンピュータ
440
主要ホワイト領域抽出手段
450
文書エレメント抽出手段
460
プロセッサ
430
境界ボックス発生手段
410
メモリ
420
連結成分識別手段
405
200
スキャナー
300
プリンタ
100

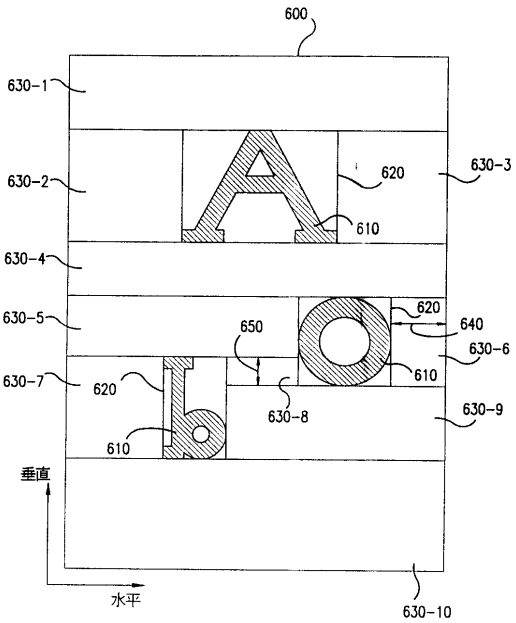
【図２】


441
442
440
446
グループ化手段
445
除去手段
444
比較手段
443
プリミティブホワイトエリア抽出手段
垂直抽出部分
水平抽出部分
主要ホワイト領域抽出手段
405

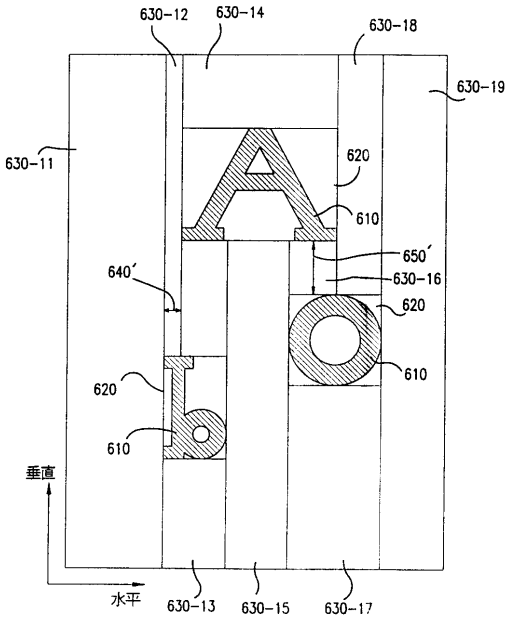
【図３】


600
620
610
620
610
620
610
垂直
水平

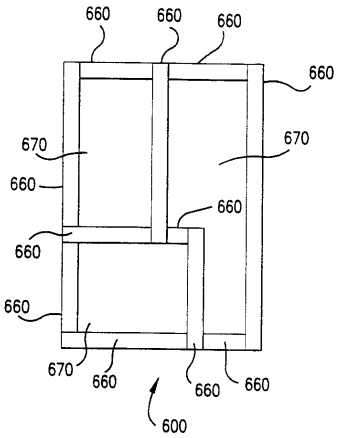
【図４】


600
630-1
630-2
620
610
630-3
630-4
630-5
620
650
640
630-6
630-7
620
610
630-8
630-9
610
垂直
水平
630-10

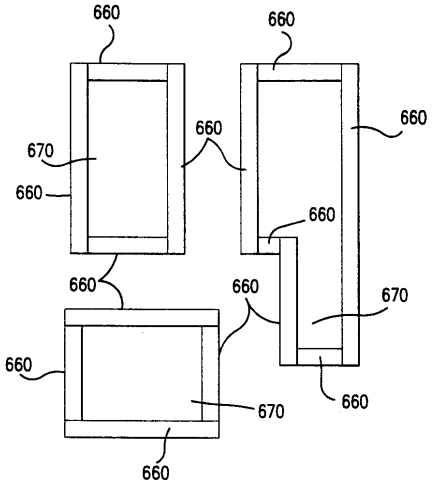
【図５】


630-12
630-14
630-18
630-19
630-11
620
610
650'
630-16
640'
620
610
620
610
垂直
水平
630-13
630-15
630-17

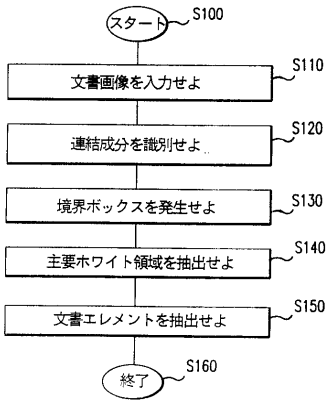
【図６】


660
660
660
660
670
670
660
660
660
660
670
660
660
660
600

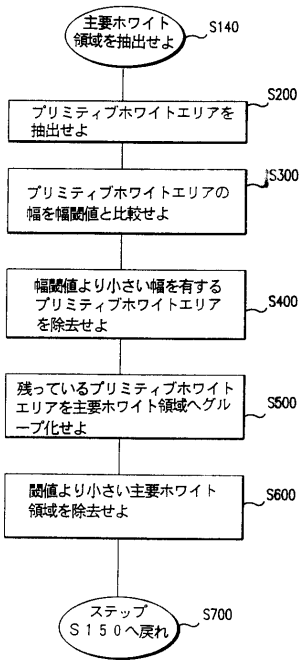
【図７】


660
660
660
670
660
660
660
660
670
660
660
670
660

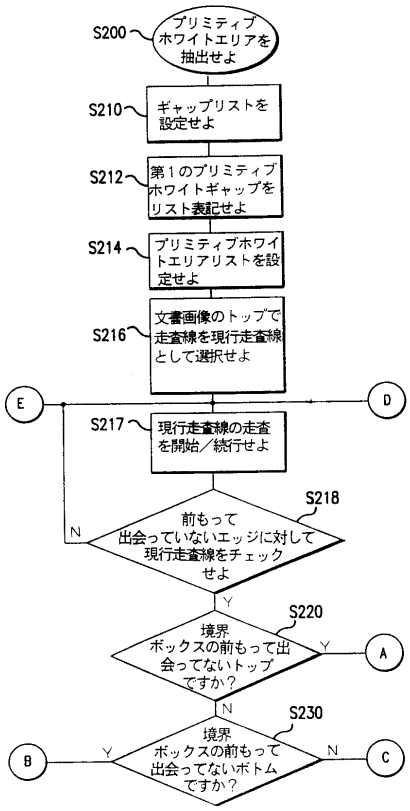
【図８】


スタート
S100
文書画像を入力せよ
S110
連結成分を識別せよ
S120
境界ボックスを発生せよ
S130
主要ホワイト領域を抽出せよ
S140
文書エレメントを抽出せよ
S150
終了
S160

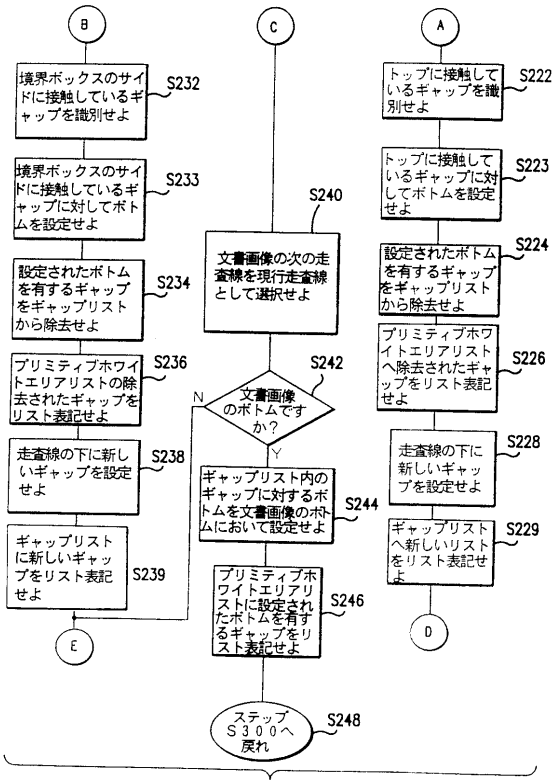
【図９】



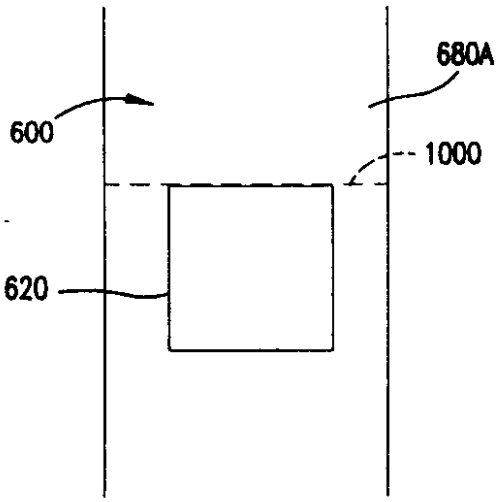
【図１０】



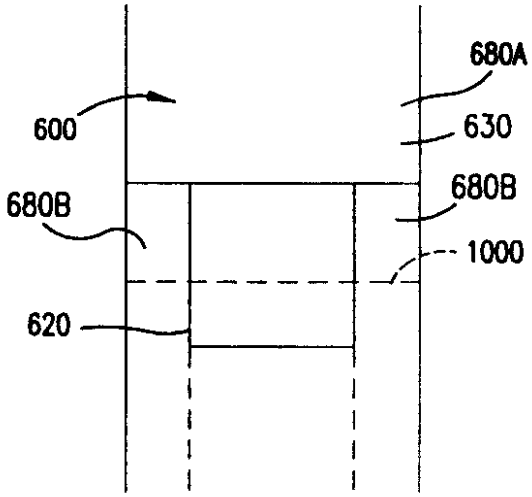
【図１１】



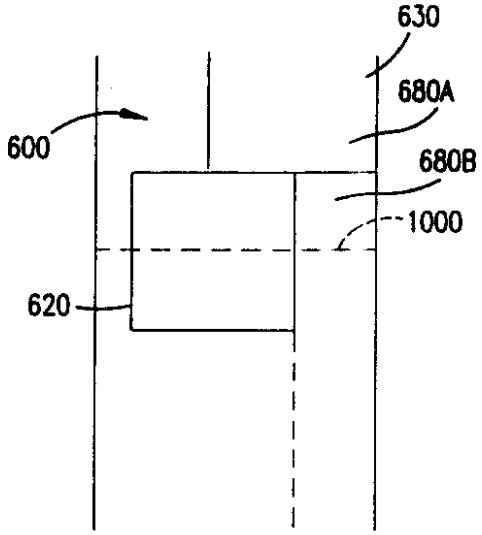
【図１２】



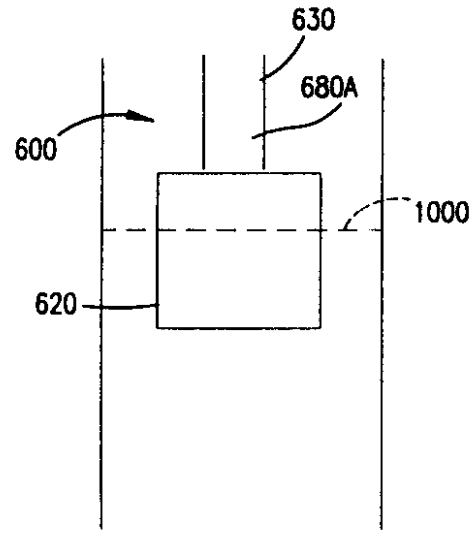
【図１３】



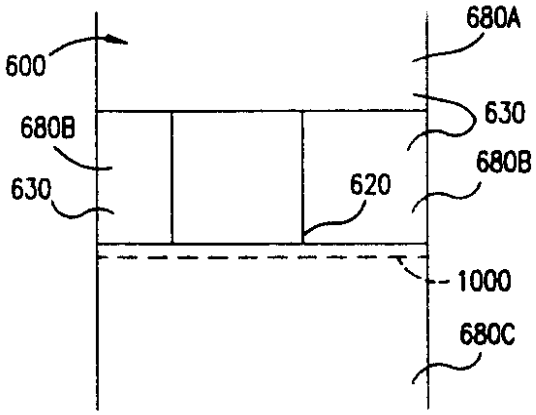
【図１４】



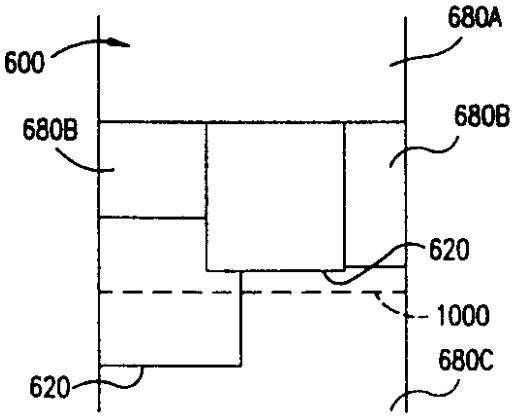
【図１５】



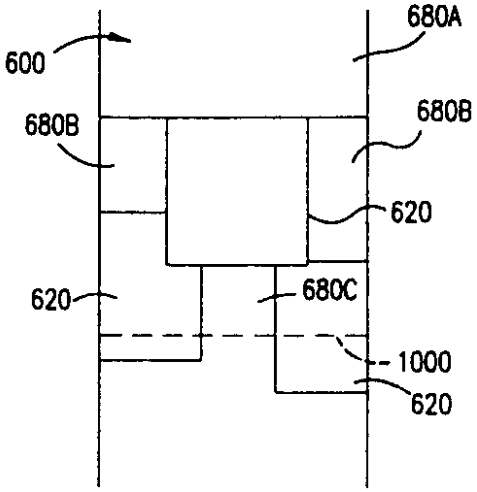
【図１６】



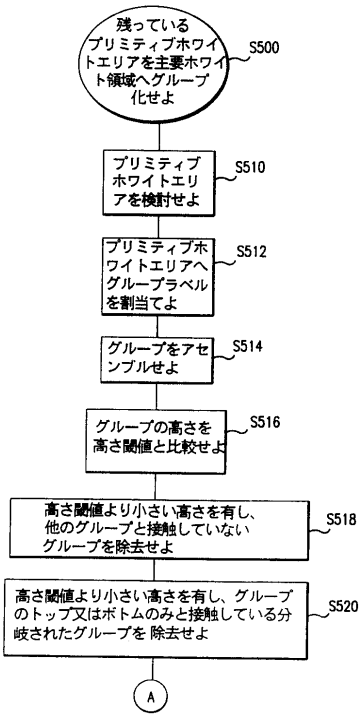
【図１７】



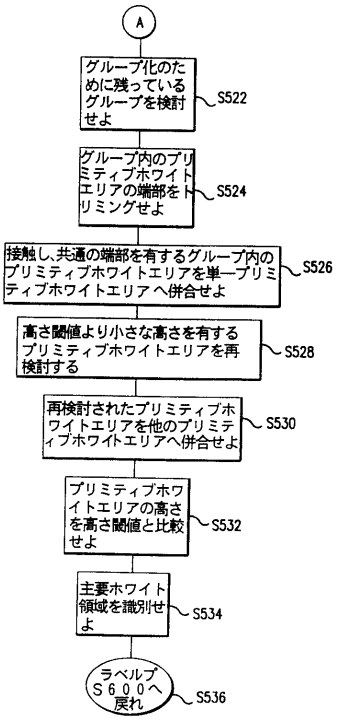
【図１８】



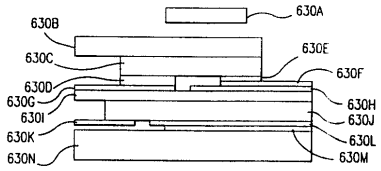
【図１９】



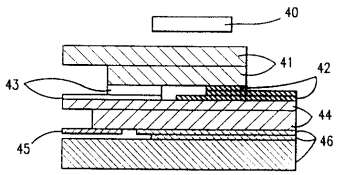
【図２０】



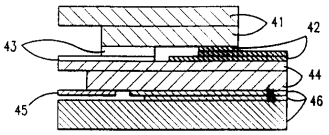
【図２１】



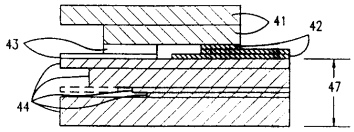
【図２２】



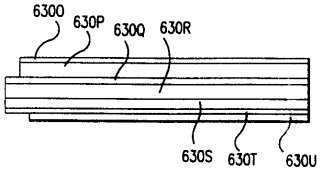
【図２３】



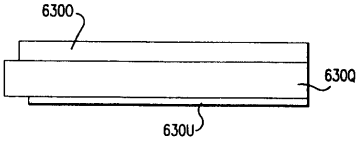
【図２４】



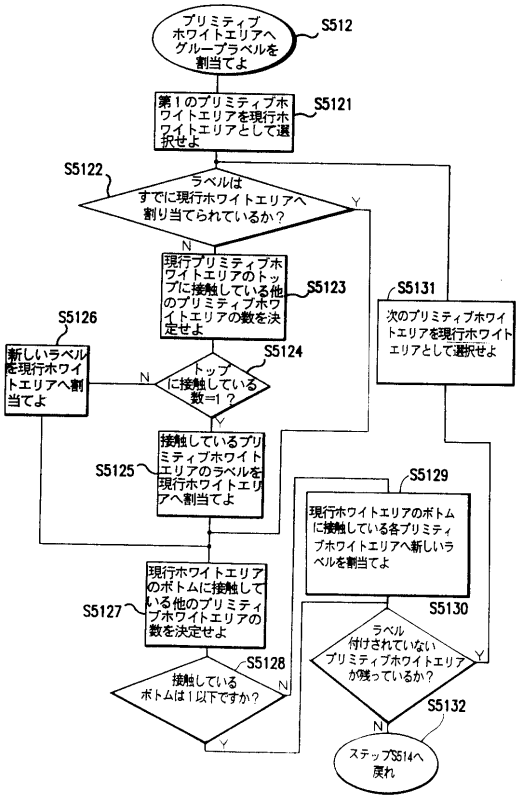
【図２５】



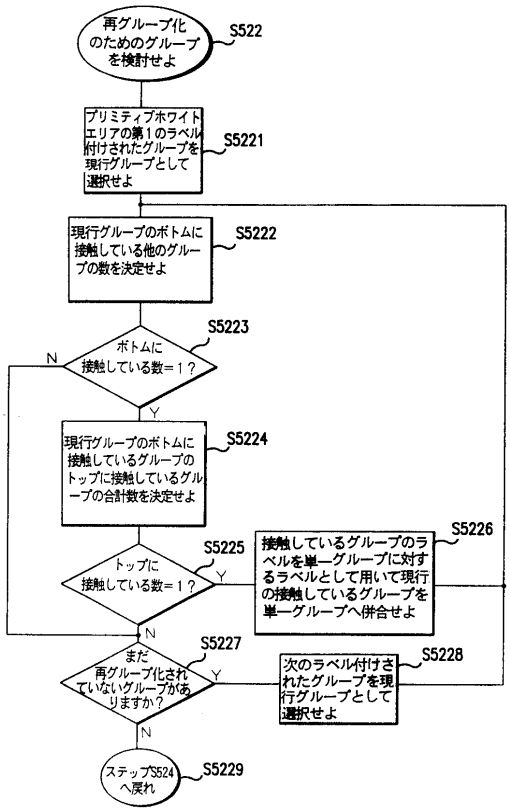
【図２６】



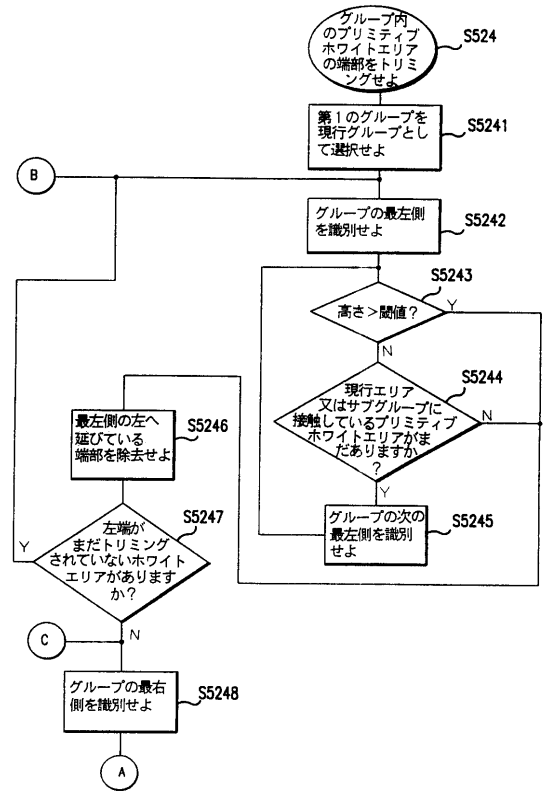
【図２７】



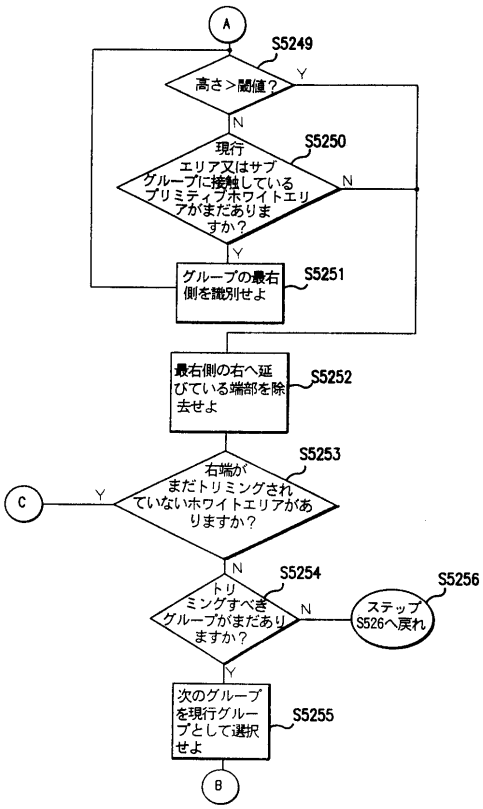
【図２８】



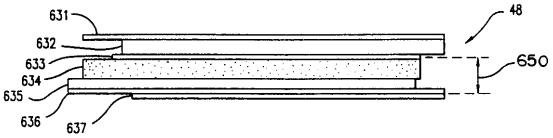
【図２９】



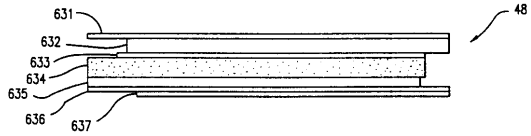
【図３０】



【図３１】



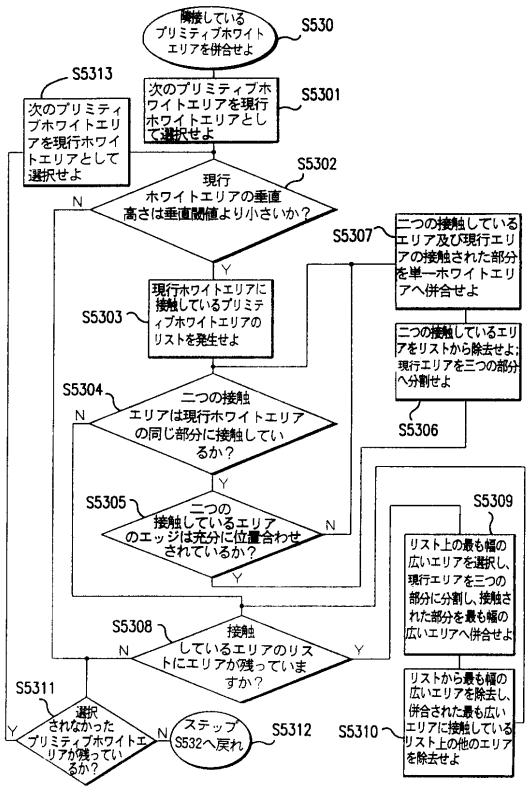
【図３２】



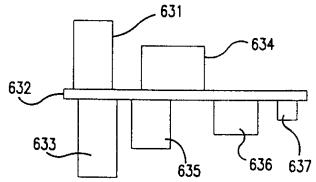
【図３３】



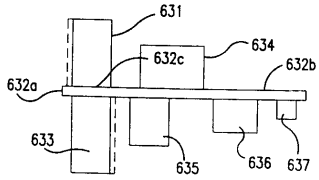
【図３４】



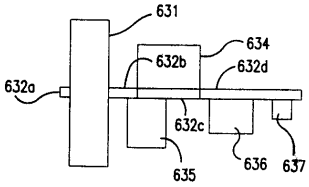
【図３５】



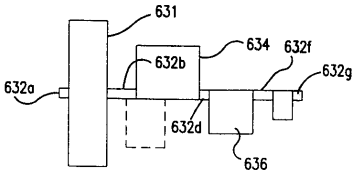
【図３６】



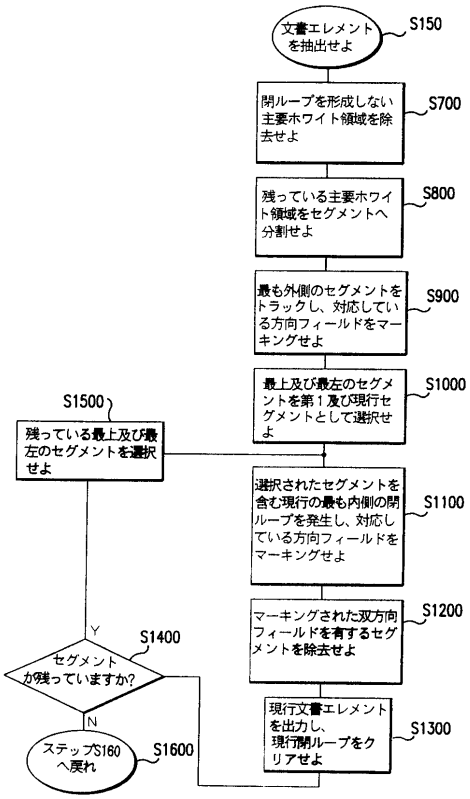
【図３７】



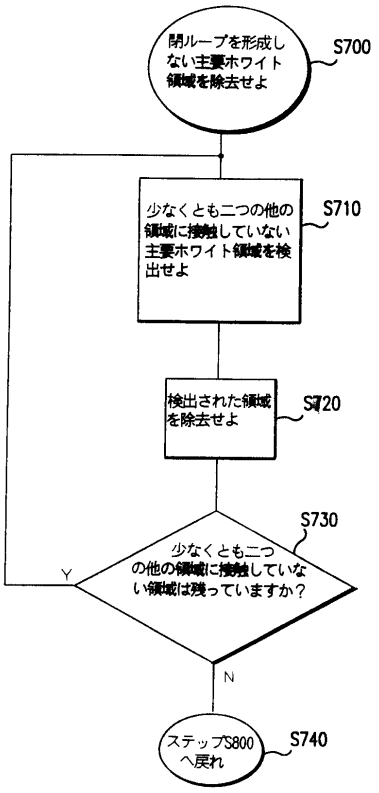
【図３８】



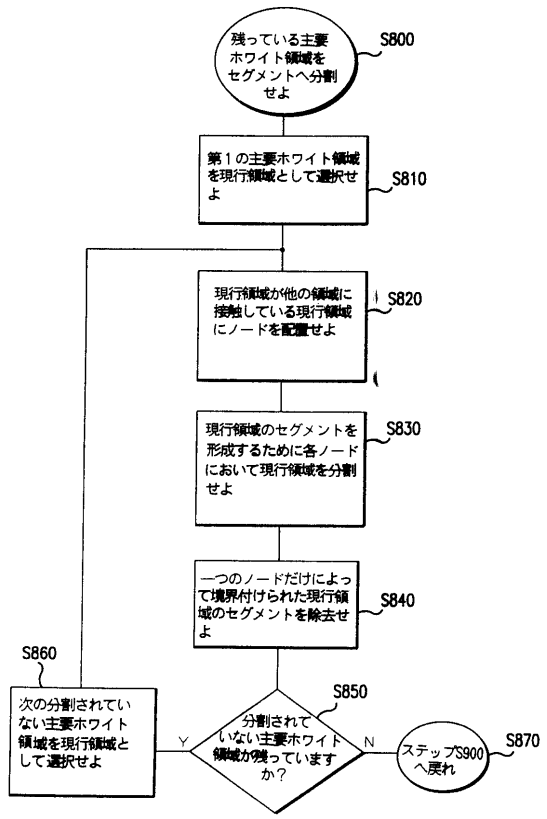
【図３９】



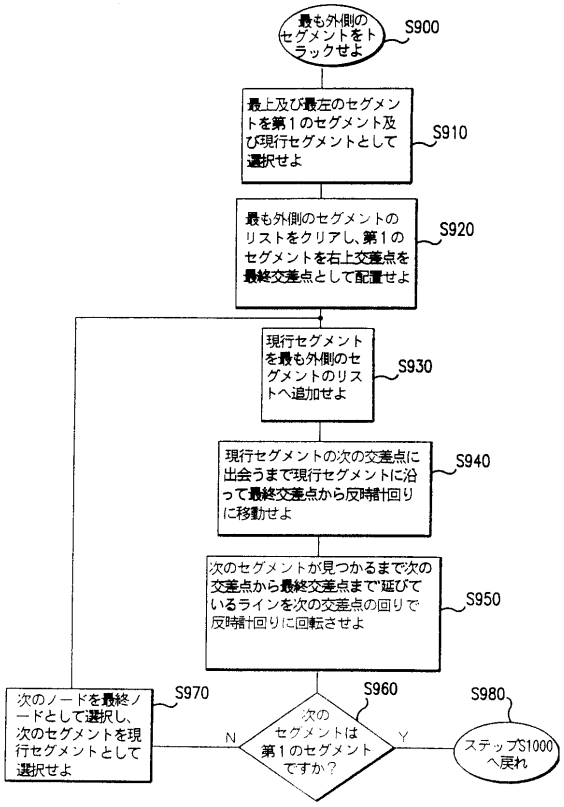
【図４０】



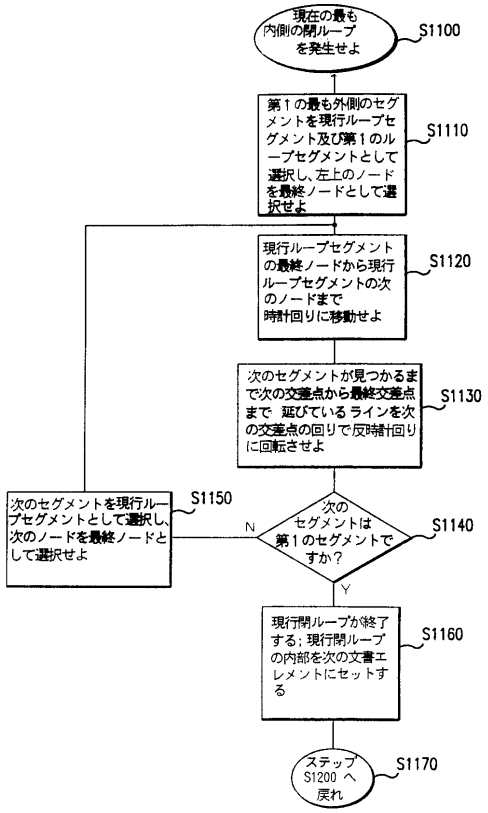
【図４１】



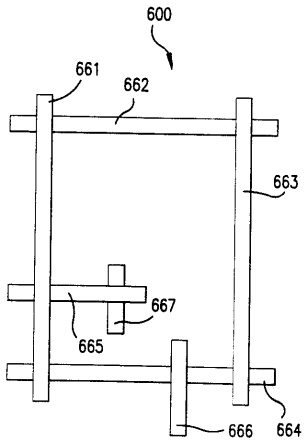
【図４２】



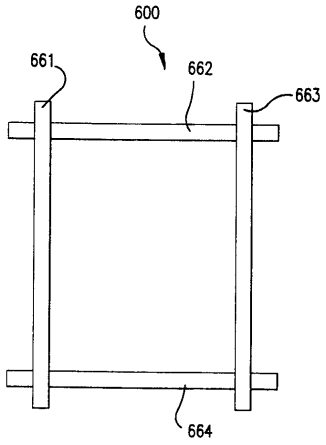
【図４３】



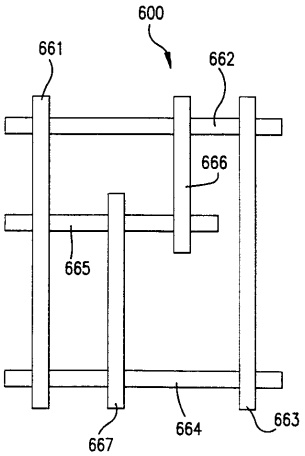
【図４４】



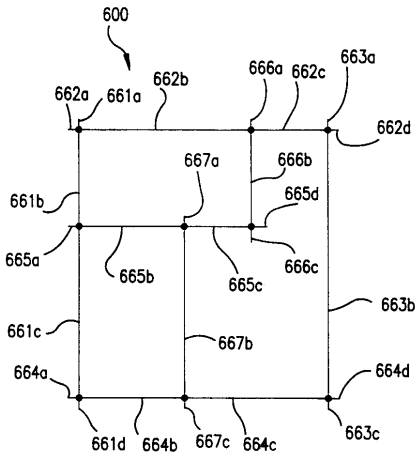
【図４５】



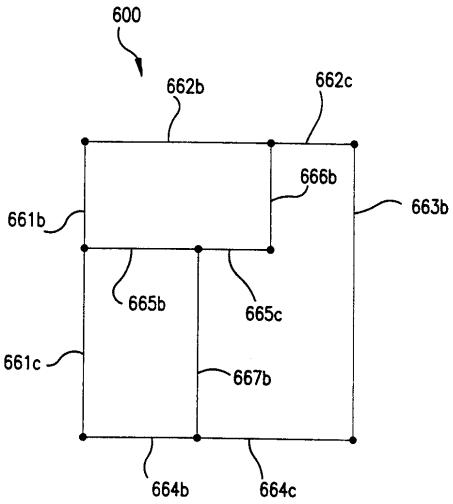
【図４６】



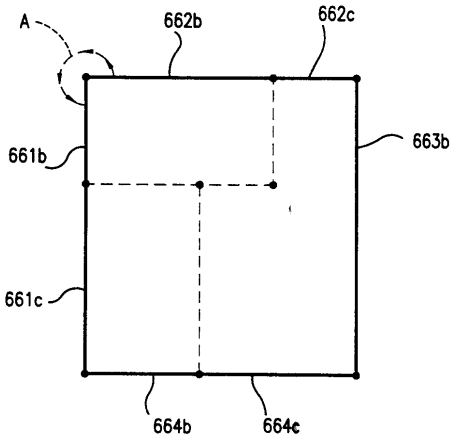
【図４７】



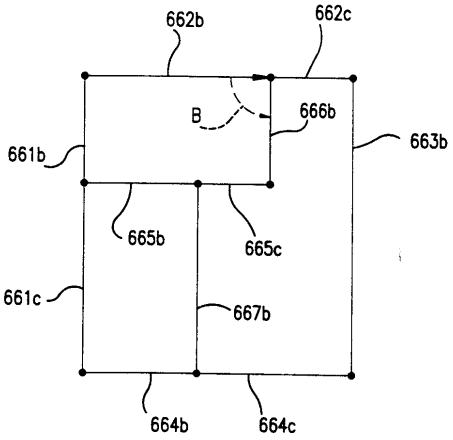
【図４８】



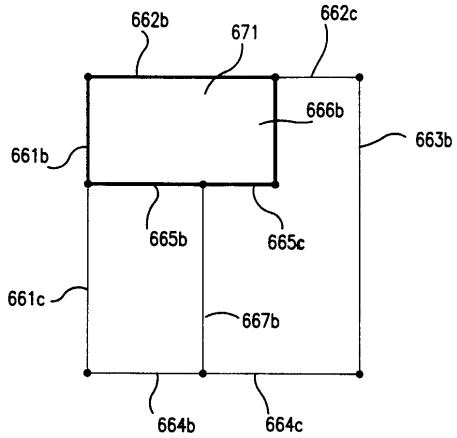
【図４９】



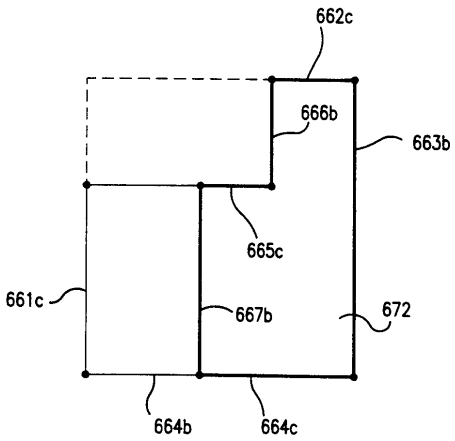
【図５０】



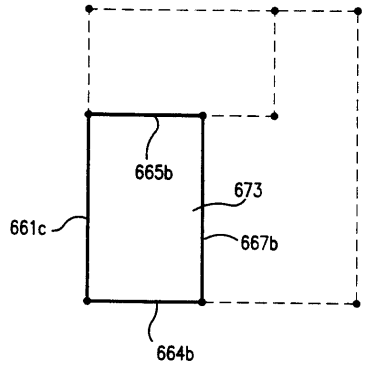
【図５１】



【図５２】



【図５３】



フロントページの続き

(72)発明者　尾▲崎▼　正治
アメリカ合衆国　カリフォルニア州　９４３０４　パロ　アルト　ピーエイエイチヴィー－４０３
ヒルビュー　アヴェニュー　３４００

審査官　新井　則和

(56)参考文献　特開平０５－１２８３０５（ＪＰ，Ａ）
特開平０４－１２０６７８（ＪＰ，Ａ）
特開平０６－０２００９２（ＪＰ，Ａ）
特開平０３－０７８８９２（ＪＰ，Ａ）
特開平０３－２５０２７９（ＪＰ，Ａ）
特開平０５－２６６２４９（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
G06K 9/20