

【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7部門第3区分

【発行日】平成21年3月19日(2009.3.19)

【公開番号】特開2004-282701(P2004-282701A)

【公開日】平成16年10月7日(2004.10.7)

【年通号数】公開・登録公報2004-039

【出願番号】特願2003-364142(P2003-364142)

【国際特許分類】

H 04 N 1/387 (2006.01)

G 06 T 7/40 (2006.01)

【F I】

H 04 N 1/387

G 06 T 7/40 100Z

【誤訳訂正書】

【提出日】平成21年1月29日(2009.1.29)

【誤訳訂正1】

【訂正対象書類名】特許請求の範囲

【訂正対象項目名】全文

【訂正方法】変更

【訂正の内容】

【特許請求の範囲】

【請求項1】

ピクセルから成る複合画像を、画像のレイアウト要素に対応する複数のフィールドに分割する方法であって、前記ピクセルが、画素の強度および／または色を示す値を有し、

画像の隣り合うピクセルの領域に対応し且つ画像のバックグラウンドを示す所定の特性を有するフィールドセパレータを見出すステップを含み、前記分割する方法がさらに、

少なくとも1つの分離方向に沿って画像の外側の境界まで前記フィールドセパレータを延長するステップと、

前記延長されたフィールドセパレータに対応するラインのテセレーショングリッドを構成するステップと、

前記テセレーショングリッドのラインによって囲まれる領域である基本長方形のセットを構成するステップと、

互いに隣り合い且つ1つのフィールドセパレータによって分離されない基本長方形同士を連結することによりフィールドを構成するステップとを含むことを特徴とする方法。

【請求項2】

基本長方形のセットを構成する前記ステップが、要素の2次元配列によって前記テセレーショングリッドを表わすマトリクスマップを構成することを含み、各要素は、テセレーショングリッドの基本長方形またはラインセグメントのいずれかを表わし、1つの要素が、フィールドセパレータに対応するラインを表わすための所定の第1の値、または基本長方形あるいは延長されたフィールドセパレータに対応するラインを表わすための更なる異なる値を有している、請求項1に記載の方法。

【請求項3】

フィールドを構成する前記ステップが、前記更なる異なる値を有するマトリクスマップ内の要素同士を連結することを含む、請求項2に記載の方法。

【請求項4】

前記フィールドセパレータが連結する点でノードが規定され、フィールドを構成する前記ステップが、前記テセレーショングリッドに対応し且つ前記テセレーショングリッド内のノードを示す要素を含むノードマトリクスを構成することを含む、請求項1から3のい

ずれか一項に記載の方法。

【請求項 5】

フィールドを構成する前記ステップが、マトリクスマップ内の連結された要素の各領域毎に外形を構成し、該外形を前記ノードマトリクス上に描出することによりフィールドを規定するノードを見出すことを含む、請求項 4 に記載の方法。

【請求項 6】

前記外形が、前記領域を拡張して、拡張された領域の領域を取り去ることにより構成される、請求項 5 に記載の方法。

【請求項 7】

分割が、画像のバックグラウンドを示す所定の特性を有する隣り合う領域に対応する辺と、辺が接続する頂点とを有するグラフを構成するとともに、前記フィールドセパレータを前記グラフの辺に関連付け、かつ

前記フィールドセパレータを画像の外側の境界まで延長することにより、前記テセレーショングリッドを形成することを含む、請求項 1 から 6 のいずれか一項に記載の方法。

【請求項 8】

グラフを構成するステップが、2つ未満の辺に連結される頂点を除去することにより、かつ／または2つ未満の辺に連結される頂点に連結する任意の辺を除去することにより、前記グラフをクリーニングすることを含む、請求項 7 に記載の方法。

【請求項 9】

前記テセレーショングリッド内のラインを2つの直交する分離方向でのズレを無くす、すなわちスナッピングすることを含む、請求項 1 から 8 のいずれか一項に記載の方法。

【請求項 10】

ピクセルから成る画像を複数のフィールドに分割するためのコンピュータプログラム製品であって、請求項 1 から 9 のいずれか一項に記載の方法をプロセッサに実行させるようにプログラムが動作する、コンピュータプログラム製品。

【請求項 11】

ピクセルから成る複合画像を、画像のレイアウト要素に対応する複数のフィールドに分割する装置であって、前記ピクセルが、画素の強度および／または色を示す値を有し、前記分割する装置が、

画像を入力するための入力ユニット(91)と、

隣り合うピクセルの領域に対応し且つ画像のバックグラウンドを示す所定の特性を有するフィールドセパレータを見出す処理ユニット(94)とを備え、

前記処理ユニット(94)が、

少なくとも1つの分離方向に沿って画像の外側の境界まで前記フィールドセパレータを延長し、

前記延長されたフィールドセパレータに対応するラインのテセレーショングリッドを構成し、

前記テセレーショングリッドのラインによって囲まれる領域である基本長方形のセットを構成し、

互いに隣り合い且つ1つのフィールドセパレータによって分離されない基本長方形同士を連結することによりフィールドを構成するように、構成されていることを特徴とする装置。

【請求項 12】

前記処理ユニット(94)が、

要素の2次元配列によって前記テセレーショングリッドを表わすマトリクスマップを構成し、各要素が、テセレーショングリッドの基本長方形またはラインセグメントのいずれかを表わし、1つの要素が、フィールドセパレータに対応するラインを表わすための所定の第1の値、または基本長方形あるいは延長されたフィールドセパレータに対応するラインを表わすための更なる異なる値を有するように構成されている、請求項 11 に記載の装置。

【請求項 1 3】

前記処理ユニット(94)が、

前記テセレーショングリッドに対応し且つ前記テセレーショングリッド内のノードを示す要素を含むノードマトリクスを構成するように構成されている、請求項11または12に記載の装置。

【請求項 1 4】

分割後に画像のフィールドを表示するディスプレイユニット(93)を備えている、請求項11から13のいずれか一項に記載の装置。

【誤訳訂正 2】

【訂正対象書類名】明細書

【訂正対象項目名】全文

【訂正方法】変更

【訂正の内容】

【発明の詳細な説明】

【発明の名称】最小領域による合成画像の分割

【技術分野】

【0001】

本発明は、ピクセルから成る複合画像を、画像のレイアウト要素に対応する複数のフィールドに分割する方法に関し、前記ピクセルが、画素の強度および／または色を示す値を有し、この方法は、画像の隣り合うピクセルの領域に対応し且つ画像のバックグラウンドを示す所定の特性を有するフィールドセパレータを見出すステップを含む。

【0002】

また、本発明は、ピクセルから成る複合画像を、画像のレイアウト要素に対応する複数のフィールドに分割する装置に関し、前記ピクセルが、画素の強度および／または色を示す値を有し、この装置は、画像を入力するための入力ユニットと、隣り合うピクセルの領域に対応し且つ画像のバックグラウンドを示す所定の特性を有するフィールドセパレータを見出す処理ユニットとを備える。

【0003】

更に、本発明はコンピュータプログラム製品に関する。

【背景技術】

【0004】

ページ分割方法は、1994年10月9日から12日にイスラエルのエルサレムで行なわれたパターン認識に関する第12回国際会議の議事録、IEEE-CSプレス、第2刊、339頁から344頁における、A. AntonacopoulousおよびR. T. Ritchingsによる文献、「Flexible page segmentation using the background (バックグラウンドを使用するフレキシブルなページ分割)」によって知られている。画像は、画素の強度および／または色を示す値を有するピクセルによって表わされる。この値は、バックグラウンド（通常、白）またはフォアグラウンド（通常、印刷スペースである黒）として分類される。ページ上の印刷領域を取り囲む白色のバックグラウンドスペースが、解析される。バックグラウンド白色スペースは、タイル、すなわちバックグラウンドピクセルの重なり合わない領域によって覆われる。

【0005】

画像内のフォアグラウンドフィールドの外形は、それを取り囲む白色タイルに沿ってトレースすることにより識別され、これにより、タイルの内側の境界が、更なる解析のためのフィールドの境界を構成する。この方法の問題は、効率的な更なる解析を妨げる複雑な記述によってフィールドの境界が表わされているという点にある。

【0006】

【特許文献1】米国特許第5,856,877号明細書

【非特許文献1】A. AntonacopoulousおよびR. T. Ritchings

、「Flexible page segmentation using the background (バックグラウンドを使用するフレキシブルなページ分割」、パターン認識に関する第12回国際会議の議事録、IEEE-CSプレス、1994年10月9日-12日、イスラエル、エルサレム、第2刊、p339-344

【非特許文献2】K.Y.Wong、R.G.Casey、およびF.M.Wahl、「Document analysis system (文書解析システム)」、IBM J.Res.、Dev 26、(1982)、p647-656

【発明の開示】

【発明が解決しようとする課題】

【0007】

本発明の目的は、信頼性が高く複雑でない画像分割方法および装置を提供することである。

【課題を解決するための手段】

【0008】

本発明の第1の態様によれば、前記目的は、冒頭の段落に規定された方法であって、少なくとも1つの分離方向に沿って画像の外側の境界までフィールドセパレータを延長するステップと、延長されたフィールドセパレータに対応するラインのテセレーショングリッドを構成するステップと、テセレーショングリッドのラインによって囲まれる領域である基本長方形のセットを構成するステップと、互いに隣り合い且つ1つのフィールドセパレータによって分離されない基本長方形同士を連結することによりフィールドを構成するステップとをさらに備えていることを特徴とする方法によって達成される。

【0009】

本発明の第2の態様によれば、前記目的は、冒頭の段落に規定された装置であって、少なくとも1つの分離方向に沿って画像の外側の境界までフィールドセパレータを延長し、延長されたフィールドセパレータに対応するラインのテセレーショングリッドを構成し、テセレーショングリッドのラインによって囲まれる領域である基本長方形のセットを構成し、互いに隣り合い且つ1つのフィールドセパレータによって分離されない基本長方形同士を連結することによりフィールドを構成するように、処理ユニットが構成されていることを特徴とする装置によって達成される。

【0010】

本発明の第3の態様によれば、前記目的は、前記方法を実行するためのコンピュータプログラム製品を用いて達成される。

【0011】

通常、画像は、フィールドセパレータを含む、このフィールドセパレータは、少なくとも2つの分離方向、一般には水平方向および垂直方向のうちの1つを有し、且つテキストフィールド等のレイアウト要素を連結しかつ／または横断しどもに取り囲む。本方法の結果、外側の境界へのフィールドセパレータの延長に基づくラインにより、テセレーショングリッド(tesselation grid)が形成される。グリッドによって囲まれるが小分割されない全ての領域が、基本長方形と称され、これらの基本長方形に関して更なる解析が行なわれる。基本長方形のセットの利点は、基本長方形を連結することによりフィールドを簡単に構成することができるという点である。なお、基本長方形のレベルに関する演算は、コンピュータで行なうと、個々のピクセル同士または小さなピクセルを基本とするオブジェクト同士を連結する場合よりも、実質的に効率的である。

【0012】

本発明は、以下の認識に基づいている。分割は、関連する階層レベルで画像内のオブジェクトを識別する処理である。例えば、新聞の一面において、階層レベルが最も低いものはピクセルであり、その次にレベルが低いものは、連結されたピクセルから成るオブジェクト(例えば、文字またはセパレータ)であり、その次がテキストラインであり、その次がテキストフィールドであり、その次がコラムであり、最後が記事である。発明者らは、構造化された画像内でフィールドを見出す場合には、フィールドセパレータの低いレベル

からビルディングブロックレベルへの変換により、フィールドの所要レベルを下回るビルディングブロックを構成できることが分かった。基本長方形は、セレーショングリッドによって効率的に構成することができるビルディングブロックである。基本長方形を1つの領域に連結するステップは、ビルディングブロックレベルで行なわれる。最後に、ビルディングブロックレベルからフィールドレベルへの変換は、画像のノードまたはフィールドセパレータの当初の連結点に基づいて、複数の基本長方形をフィールドに統合することにより行なわれる。したがって、基本長方形の構成は、主に多角形のフィールドを有するデジタル画像を分割する最中に、フィールドのビルディングブロックを決定するという便利な方法を与える。

【0013】

この方法の一実施形態において、基本長方形のセットを構成するステップは、要素の2次元配列によってセレーショングリッドを表わすマトリクスマップを構成することを含み、各要素は、セレーショングリッドの基本長方形またはラインセグメントのいずれかを表わし、1つの要素は、フィールドセパレータに対応するラインを表わすための所定の第1の値、または、基本長方形あるいは延長されたフィールドセパレータに対応するラインを表わすための更なる異なる値を有している。利点は、マトリクスマップが、基本長方形と基本長方形同士の間の境界とを含んでいるという点である。マトリクスマップは、簡単に処理することができる。これは、そうでなければ、演算を複雑にしかねない幾何学的な内容を伴うことなく、マトリクスマップが、フィールドのビルディングブロックのレベルで画像を表わしているためである。

【0014】

この方法の一実施形態において、ノードは、フィールドセパレータが連結する位置、およびセレーショングリッド内の対応する位置にあるオリジナル画像の点で規定される。また、フィールドを構成するステップは、セレーショングリッドに対応し且つセレーショングリッド内のノードを示す要素を含むノードマトリクスを構成することを含む。

【0015】

利点は、ノードマトリクスが、幾何学的な表示でノードへの参照を含んでいるという点である。ノードマトリクスにより、フィールドのビルディングブロックすなわち基本長方形のレベルを、ノードによってフィールドの表示へと簡単に変換することができる。

【0016】

本発明に係る装置の更に好ましい実施形態が、別の請求項に記載されている。

【0017】

本発明のこれらの態様および他の態様は、以下の説明に一例として記載された実施形態および添付図面を更に参照することにより、明らかとなり、解明される。

【0018】

図面は、概略図であり、一定の縮尺で描かれたものではない。図面において、既に説明した要素に対応する要素が、同じ参照符号で示されている。

【発明を実施するための最良の形態】

【0019】

図1は、知られている分割システムによる3つの基本的なステップを有する典型的な分割方法の概略を示している。入力画像11は、連結成分解析を使用して画像のピクセルを解析するCCA(Connected Component Analysis)モジュール14で処理される。最初に、白黒文書、グレースケール文書、またはカラー文書、例えば新聞の一面であっても良いオリジナル画像が、好ましくはグレースケールで走査される。このグレースケール走査された画像は、フォアグラウンド値(例えば黒)またはバックグラウンド値(例えば白)を各ピクセルに割り当てるため、ハーフトーン化される。CCAモジュール14は、同様の特性を有する隣り合うピクセルの連結成分(Connected Component, CC)を検出することにより、画像中のフォアグラウンド要素を見つける。分割処理におけるこの第1のステップの例は、例えばUSS, 856, 877に記載されている。CCAモジュールは、連結されたフォアグラウンドピクセルの

連結成分であるCCオブジェクト12を出力として形成する。LA(Layout Analysis)モジュール15は、CCオブジェクト12を入力として受け、レイアウトオブジェクト13を形成する。この場合、CCオブジェクトを統合してグループ化することにより、テキストラインおよびテキストブロック等の大きなレイアウトオブジェクトが形成される。この段階中においては、ヒューリスティックス(経験則)を使用して、レイアウト要素をグループ化し、大きなレイアウト要素を形成する。これは、通常のボトムアップ処理における論理ステップである。AF(Article Formatation)モジュール16は、レイアウトオブジェクト13を入力として受けて、記事形成(article formatation)により出力として記事17を形成する。このモジュールにおいては、大きなエンティティを構成する幾つかのレイアウトオブジェクトが、一緒になってグループ化される。大きなエンティティは、オリジナル画像に適用されるレイアウトルールを使用してアセンブルされる。例えば、新聞の一面において、AFモジュールは、その特定の新聞様式のレイアウトルールにしたがって、テキストブロックおよび画像のようなグラフィック要素をグループ化し、個々の記事を形成する。例えば西洋タイプの雑誌のレイアウト、科学のテキストのレイアウト、または日本の記事のレイアウトといった画像のレイアウトタイプの知識は、ルールに基づいた記事形成方法において使用でき、これにより、テキストブロックのグループ化を向上させることができる。

【0020】

本発明によれば、以下に説明するように、分割に対して複数の更なるステップが加えられる。これらのステップは、画像を複数のフィールドに分割した後、1つのフィールド内で要素を検出すること、すなわち、更に小さい相互に関連する別個の項目によって構成される、レイアウトオブジェクトを形成することに関するものである。図2は、サンプルとしての日本の新聞を示している。このような新聞は、水平方向読み方向22および垂直方向読み方向21の両方を有するテキストラインを含む、特定のレイアウトを有している。検出された連結成分の従来のボトムアップグループ化処理における問題点は、グループ化をどの方向で進めるべきかが分からぬという点である。そのため、バックグラウンドを処理してその頁(一面)のフィールドを検出する別個のステップにより、分割が増大される。その後、文字のグループ化を行なう前に、日本の新聞の各フィールドにおける読み方向が検出される。

【0021】

この方法の一実施形態においては、例えば個々のコラムにおける黒ライン23といったセパレータ要素が検出されて、複数のバックグラウンド要素に変換される。このような選択肢を用いれば、実際に連結される垂直および水平ラインを含む黒ライン23の大きな要素を、様々なセパレータ要素に分離することができる。日本の新聞において、ラインは、レイアウトにおいてフィールドを分割するための非常に重要なオブジェクトである。これらのオブジェクトが、分離方向に沿うラインとして認識されることが求められる。この選択肢が無いと、これらのオブジェクトはグラフィックスとして分類される。この選択肢を使用すると、ラインを、様々な方向のセパレータ要素として、各分離方向毎に個別に扱うことができる。

【0022】

図3は、オブジェクトを1つの方向に統合する基本的な方法を示している。この図は、知られている方向に向けられたレイアウトオブジェクト、例えば読む順番が分かっている状況におけるテキストブロックを見つけるためのLAモジュール15の基本的な機能を示している。連結成分12は、統計的な解析により最初の解析ステップ31で処理され、これにより、算定閾値32が得られる。2番目の分類ステップ33においては、CC分類が補正され、これにより、補正された連結成分34が得られる。この補正された連結成分が、3番目の統合ステップ35で処理されることにより、文字がテキストラインに加えられ、その結果、テキストラインおよび他のオブジェクト36が得られる。4番目のテキスト統合ステップ37においては、テキストラインが、テキストブロック38(および、可能であれば他のグラフィックオブジェクト)に加えられる。日本の新聞に関する要求事項に

基づいて、従来のオブジェクトの統合は、少なくとも 2 つの読み方向に沿っていなければならず、そのため、前述した基本的な方法を改良しなければならない。

【0023】

図 4 は、オブジェクトの分割および 2 方向統合を示している。図 3 の 1 方向処理に対して、新たな別個のステップが加えられている。最初の（前）処理ステップにおいては、画像のグラフ 4-1 が構成される。フィールドセパレータを見つけることによりグラフを構成することについて以下に述べる。グラフにおいて、フィールドは、フィールド検出ステップ 4-2 で、グラフの辺によって囲まれる領域を見出すことにより検出される。当該領域は、テキストブロック 4-7 を含むフィールドとして分類される。テキストブロック 4-7 において（テキストブロック領域内にある補正されたみ連結成分 3-4 または連結成分 4-3 を使用して）、ステップ 4-4 で、読む順番 4-5 が決定される。読み方向検出は、文書スペクトルに基づいている。テキストブロック 4-7 のフィールド、含まれている連結成分 4-3、および読む順番 4-5 を入力として使用して、ライン形成ステップ 4-6 は、必要に応じて、見出された方向に沿って文字をラインに加える。

【0024】

ここで、グラフ 4-1 の構成について説明する。文書のグラフ表示は、走査のバックグラウンドを使用して形成される。走査におけるピクセルは、バックグラウンドとして（通常、白）或はフォアグラウンド（通常、黒）として分類される。白の大きな領域だけが、フィールドに関する情報を与えるため、例えば画像をダウンサンプリングすることにより、小さなノイズオブジェクトが除去される。1 つのフォアグラウンド（黒）ピクセルを除去するため、更に、ダウンサンプリングされた画像から斑点が除去されても良い。

【0025】

次の作業は、重要な白領域を抽出することである。最初のステップは、隣り合うバックグラウンドピクセルの 1 ピクセル高領域、所謂ホワイトランを検出することである。所定の最小長よりも短いホワイトランは、処理から除外される。

【0026】

図 5 は、一例として、垂直方向で隣り合う白ピクセルの 4 つの水平方向のホワイトラン 5-1 を示している。フォアグラウンド領域 5-3 は、ホワイトラン 5-1 を直接に取り囲むフォアグラウンドピクセルを有していると仮定される。「最大白長方形」は、隣り合うホワイトラン 5-1 によって構成することができる最も大きな長方形領域、したがって、黒（フォアグラウンド）ピクセルを含んでいると延長することができない長方形白領域として規定される。最大白長方形 5-2 は、垂直方向の破線によって示される長さ、及び 4 ピクセル分の幅を有する 4 つのホワイトラン 5-1 に基づいて示されている。白長方形は、これを延長することができない場合に、いわゆる最大分離力を有する。そのような長方形は、より重要な白領域の更に小さい部分ではない。したがって、長方形 5-2 は、4 ピクセル分の幅を有する考えられる唯一の最大長方形である。3 ピクセル分または 2 ピクセル分の幅を持つ更なる長方形を構成することができる。更なる例が図 6 に示されている。

【0027】

白長方形の構成は、例えば水平方向および垂直方向の白長方形といったように、異なる分離方向で別個に行なわれる。垂直方向の白長方形は、画像を回転させ、且つ回転された画像における水平方向のホワイトランを検出することにより検出される。なお、画像のタイプまたは用途に応じて、斜め方向等の他の分離方向が選択されても良いこと留意されたい。

【0028】

最大白長方形を構成するためのアルゴリズムは、以下の通りである。アルゴリズムの入力は、所定の画像から検出された全ての水平方向の 1 ピクセル高ホワイトラン（White Run、WR）から成る。各ホワイトランは、一組の座標 $((x_1, y_1), (x_2, y_2))$ によって特徴付けられる長方形として表わされる。ここで、 x_1 および y_1 は、その左上角部の座標であり、 x_2 および y_2 は右下角部の座標である。順序付けられたアクティブなオブジェクト INPUT LIST に存在する各ホワイトランは、延長の可

能性について検査される。延長の可能性は、 p のラベルが付された所定の W R が、最大白長方形 (M a x i m a l W h i t e R e c t a n g l e , M W R) を形成できるか否かといった様で表わされる。延長の可能性が偽である場合には、 p が既に最大のものであり、 p は、アクティブ INPUT LIST から削除されるとともに、アクティブ RESULT LIST に書き込まれる。延長の可能性が真である場合には、 p で始まる全ての M W R が構成されるまで、延長のための検査が繰り返される。その後、 p が INPUT LIST から削除され、 p から得られる全ての M W R が、アクティブ RESULT LIST に書き込まれる。INPUT LIST からの全ての白長方形が処理されると、RESULT LIST に全ての M W R が含まれるようになる。アルゴリズムの効率を高めるため、INPUT LIST において y の値がソートされる。まず最初に、水平方向の W R 、すなわち高さよりも幅が大きいホワイトランに関してアルゴリズムが適用される。そして、画像を 90° 回転させた後、垂直方向の W R に対してアルゴリズムを適用することができる。

【 0 0 2 9 】

一実施形態において、最大長方形を構成するためのアルゴリズムは、以下の通りである。まず、長方形データがリンクリストとして記憶される。この場合、長方形データには、少なくとも長方形の頂点の座標が含まれている。INPUT LIST および RESULT LIST も、リンクリストとして記憶される。このリンクリストには、少なくとも 3 つの要素、すなわち、白長方形の数、リンクリスト内の最初および最後の要素のポインタが含まれている。次に、以下のステップが実行される。すなわち、INPUT LIST をアクティブにして、RESULT LIST を開始し、選択された長方形の一次的な座標のための BUFFER を開始する。順序付けられたアクティブな INPUT LIST にあるもののうち、 p_1 のラベルが付された最初の白長方形から始める。リスト中の次の白長方形に p_2 のラベルが付される。INPUT LIST 中の各白長方形毎に、 p_1 が延長の可能性を有しているか否かを検討する。アクティブな白長方形 p_1 に関し、順序付けられたアクティブな INPUT LIST 中で、以下の条件を満たす p_{nj} , $j = 1, \dots, l$ のラベルが付された最初のものを見つける。

【 数 1 】

$$y_2(p_1) = y_1(p_{nj})$$

【 数 2 】

$$x_1(p_{nj}) \leq x_2(p_1)$$

【 数 3 】

$$x_2(p_{nj}) \geq x_1(p_1)$$

【 0 0 3 0 】

この検索により、 $\{ p_{n1}, p_{n2}, \dots, p_{nl} \}$ のセットが得られる。このセット $\{ p_{n1}, p_{n2}, \dots, p_{nl} \}$ が空でない場合にだけ、 p_1 が延長の可能性を有していると言われる。

【 0 0 3 1 】

p_1 が延長の可能性を有していない場合には、 p_1 が最大白長方形である。 p_1 を RESULT LIST に書き込んで、 p_1 を INPUT LIST から削除し、 p_2 に関して処理を進める。

【 0 0 3 2 】

p_1 が延長の可能性を有している場合には、延長処理を p_1 に適用する。 p_2 に関して処理を進める。なお、 p_1 がそれ自体最大であっても延長の可能性を有し得る。延長処理

は以下の通りである。まず、 p_1 が延長の可能性を有していると仮定すると、セット { $p_{n_1}, p_{n_2}, \dots, p_{n_j}$ } が存在する。延長処理は、{ $p_{n_1}, p_{n_2}, \dots, p_{n_j}$ } の各要素に対して一貫して適用される。長方形 p_{n_j} , $j = 1, \dots, l$ を用いて延長可能な白長方形 p_1 に関して、以下の座標を有する新たな長方形 p_{1,n_j} を構成する。

【数4】

$$x_1(p_{1,n_j}) = \max\{x_1(p_1), x_1(p_{n_j})\}$$

【数5】

$$x_2(p_{1,n_j}) = \min\{x_2(p_1), x_2(p_{n_j})\}$$

【数6】

$$y_1(p_{1,n_j}) = y_1(p_1)$$

【数7】

$$y_2(p_{1,n_j}) = y_2(p_{n_j})$$

【0033】

$p_1, n_j, j = 1, \dots, l$ の座標を「座標」バッファに書き込む。ここで、 p_1, n_j に関して延長可能性の検査を繰り返す。検査が真である場合には、 p_1, n_j が最大である。 p_1, n_j を RESULT LIST に書き込み、さもなければ、 p_1, n_j を延長する。

【0034】

延長処理を p_1, n_j に適用する前に、吸収作用に関して p_1 および p_{n_j} をチェックする。 p_1, n_j を用いた吸収作用に関する p_1 および p_{n_j} の検査は、以下の通りである。吸収作用とは、 $p_1 (p_{n_j})$ 又はこれらの両方が、 p_1, n_j に完全に含まれている状態を意味する。座標において、このことは、以下の状態を意味する。

【数8】

$$x_1(p_{t,n_j}) \leq x_1(p_k)$$

【数9】

$$x_2(p_{1,n_j}) \geq x_2(p_k)$$

ここで、 $k = 1, n_j, j = 1, \dots, l$ である。

【0035】

状態が p_1 に関して真である場合には、 p_1 は p_1, n_j によって吸収される。 p_1 を INPUT LIST から除去する。状態が p_{n_j} に関して真である場合には、 p_{n_j} は p_1, n_j によって吸収される。 p_{n_j} を INPUT LIST から除去する。

【0036】

アルゴリズムでは、長方形の幅が高さよりも大きく、したがって、長方形が、主として水平方向であると仮定される。垂直方向の MWR を構成するため、オリジナルの 2 値画像が 90° 時計周りに回転される。回転された画像に関して前述したアルゴリズムが繰り返される。その結果、オリジナル画像において全ての垂直方向の MWR が構成される。

【0037】

図 6 は、最大白長方形の構成を示している。水平方向の x 軸および垂直方向の y 軸に沿って、ピクセル座標が表わされている。4 つのホワイトラン 6 1 が、図の左側に示されている。ホワイトラン (WR) は、その上角部および下角部の座標が、以下の座標に対応する長方形として描かれている。

【数 1 0】

WR₁ : ((10,1),(50,2))

【数 1 1】

WR₂ : ((10,2),(50,3))

【数 1 2】

WR₃ : ((5,3),(30,4))

【数 1 3】

WR₄ : ((40,3),(60,4))

【0038】

これらのホワイトランによって全ての最大白長方形が構成される。図の右側部分には、結果として得られる 5 つの最大白長方形 (MWR) が、62、63、64、65、および 66 で示されている。図示された 5 つの MWR は、図の左側部分に示された WR における MWR の完全なセットである。構成アルゴリズムは以下の通りである。

【0039】

INPUT LIST に 4 つのホワイトラン 6 1 を含ませる。INPUT LIST からの最初の要素は、WR₁ ((10, 1), (50, 2)) である。WR₁ に p₁ のラベルを付ける。前述したように、延長の可能性に関して p₁ を検査する。延長における第 1 の候補は、WR₂ ((10, 2), (50, 3)) である。WR₂ に p_{n1} のラベルを付ける。前述した延長に関する方式にしたがって p_{n1} を用いて p₁ を延長する。これにより、座標 ((10, 1), (50, 3)) を有する新たな長方形 p_{1, n1} が与えられる。p_{1, n1} を用いた吸収作用に関して p₁ および p_{n1} を検査する。以下の通り、吸収検査により、p₁ および p_{n1} の両方が、p_{1, n1} によって吸収される。したがって、p₁ および p_{n1} を INPUT LIST から削除する。p_{1, n1} に関して処理を進める。延長の可能性に関して p_{1, n1} を検査する。これにより、第 1 の候補 WR₃ ((5, 3), (30, 4)) が与えられる。WR₃ に p_{t1} のラベルを付ける。延長に関する方式にしたがって p_{t1} を用いて p_{1, n1} を延長する。その結果、座標 ((10, 1), (30, 4)) を有する新たな長方形 p_{(1, n1), t1} が得られる。p_{(1, n1), t1} を用いた吸収作用に関して p_{t1} を有する p_{1, n1} を検査する。検査は失敗する。

【0040】

p_{(1, n1), t1} に関して延長の可能性の検査を繰り返す。検査は失敗する。すなわち、p_{(1, n1), t1} は延長の可能性を有していない。このことは、p_{(1, n1), t1} が最大であることを意味する。座標 ((10, 1), (30, 4)) を有する p_{(1, n1), t1} を RESULT LIST に書き込む。

【0041】

p_{1, n1} に関して再び処理を進め、延長の可能性に関して p_{1, n1} を検査する。第 2 の候補 WR₄ ((40, 3), (60, 4)) が見出される。WR₄ に p_{t2} のラベルを付ける。延長に関する方式にしたがって p_{t2} を用いて p_{1, n1} を延長する。その結

果、座標((40, 1), (50, 4))を有する新たな長方形 $p_{(1, n_1), t_2}$ が得られる。

【0042】

$p_{(1, n_1), t_2}$ を用いた吸収作用に関して p_{t_2} を有する p_{1, n_1} を検査する。検査は失敗する。すなわち、吸収がない。 $p_{(1, n_1), t_2}$ に関して延長の可能性の検査を繰り返す。検査は失敗する。すなわち、 $p_{(1, n_1), t_2}$ は延長の可能性を有していない。このことは、 $p_{(1, n_1), t_2}$ が最大であることを意味する。座標((40, 1), (50, 4))を有する $p_{(1, n_1), t_2}$ を RESULT LIST に書き込む。

【0043】

延長の可能性に関して p_{1, n_1} を再び検査する。検査は失敗し、 p_{1, n_1} が最大である。座標((10, 1), (50, 3))を有する p_{1, n_1} を RESULT LIST に書き込む。

【0044】

INPUT LIST に戻る。この段階における INPUT LIST は、2つのホワイトラン、すなわち、 $WR_3 : ((5, 3), (30, 4))$, $WR_4 : ((40, 3), (60, 4))$ を含んでいる。 WR_3 から開始して、これに p_2 のラベルを付ける。 p_2 に関して延長の可能性の検査を繰り返す。検査は失敗し、 p_2 が最大である。座標((5, 3), (30, 4))を有する p_2 を、RESULT LIST に書き込む。INPUT LIST から p_2 を除去する。 WR_4 に関して処理を進め、これに p_3 のラベルを付ける。 p_3 に関して延長の可能性の検査を行なうことにより、 p_3 が最大であることが分かる。座標((40, 3), (60, 4))を有する p_3 を RESULT LIST に書き込む。INPUT LIST から p_3 を除去する。最終的に、RESULT LIST は、5つの最大白長方形、すなわち、図6に64で示される $MWR_1 : ((10, 1), (50, 3))$ と、62で示される $MWR_2 : ((10, 1), (30, 4))$ と、63で示される $MWR_3 : ((40, 1), (50, 4))$ と、65で示される $MWR_4 : ((5, 3), (30, 4))$ と、66で示される $MWR_5 : ((40, 3), (60, 4))$ とを含んでいる。

【0045】

図7は、本発明にかかる方法における次のステップ、すなわち、最大白長方形をオーバーラップさせるクリーニングステップを示している。このクリーニングステップにおいて、オーバーラップする複数の最大白長方形は、後で詳述するように、オリジナルの最大白長方形の最も関連する特性を組み合わせる、1つの所謂「情報提供最大長方形」(Informative Maximum Rectangle, IWR)に統合される。

【0046】

クリーニングは、サイズおよび空間的な関係をチェックする等のステップを更に含んでいる。図7の上側の部分は、一例として、2つの最大白長方形 MWR_1 および MWR_2 を示している。これらの対は、図の下側の部分に示されるように、クリーニングステップにおいて、1つの情報提供白長方形 IWR に統合される。オーバーラップを検知する処理および統合する処理は、関連する対をもはや形成することができなくなるまで繰り返される。対を形成する基準は、オーバーラップ領域のサイズであっても良い。

【0047】

また、クリーニングステップは、薄い或は短い長方形、すなわち、アスペクト比が所定の値を下回る長方形を除去することを含んでいても良い。除去する基準は、画像のタイプに基づいていても良い。例えば、所定のピクセル数を下回る幅は、テキストラインのセパレータを示し、フィールドの分離に関係しない。特定の値を下回る長さは、フィールドの期待されるサイズに関連しない。

【0048】

クリーニングステップのためのアルゴリズムは、以下の通りである。クリーニング処理の開始は、図5および図6に関して前述したように構成される MWR のセット全体である

。クリーニング処理は、情報を提供しない M W R を廃棄するべく適用される。このため、情報を提供しないことに関する尺度が規定される。例えば、長い M W R は、短いものよりも多くの情報を与える。低いアスペクト比は、情報提供量が少ない正方形を多かれ少なかれ示す。また、例えば 2 つのテキストライインを分離する極めて薄い長方形は、除外しなければならない。最初に、全ての M W R は、その高さと幅との間の比が計算されることにより、水平方向、垂直方向、あるいは正方形であるとして分類される。正方形の M W R は、その情報提供性が無いことにより、削除される。残りの水平方向および垂直方向の M W R に関しては、以下の 3 つのステップから成るクリーニング技術が適用される。

【 0 0 4 9 】

長さまたは幅が所定の値を下回る各 M W R が削除される。

【 0 0 5 0 】

長い辺の長さを短い辺の長さで割った比として規定されるアスペクト比 (A R) が、所定の値を下回る各 M W R が削除される。

【 0 0 5 1 】

互いにオーバーラップする水平方向（または垂直方向）の M W R₁ ((x₁, y₁), (x₂, y₂)) および水平方向（または垂直方向）の M W R₂ ((a₁, b₁), (a₂, b₂)) から成る各対毎に、以下の座標を用いて、情報提供白長方形 I W R が構成される。

(a) 水平方向のオーバーラップ

【 数 1 4 】

$$x_1 = \min \{ x_1, a_1 \}$$

【 数 1 5 】

$$y_1 = \max \{ y_1, b_1 \}$$

【 数 1 6 】

$$x_2 = \max \{ x_2, a_2 \}$$

【 数 1 7 】

$$y_2 = \min \{ y_2, b_2 \}$$

(b) 垂直方向のオーバーラップ

【 数 1 8 】

$$x'_1 = \max \{ x_1, a_1 \}$$

【 数 1 9 】

$$y'_1 = \min \{ y_1, b_1 \}$$

【 数 2 0 】

$$x'_2 = \min \{ x_2, a_2 \}$$

【数21】

$$y'_2 = \max \{y_2, b_2\}$$

【0052】

この処理は、オーバーラップするMWRから成る全ての対に関して繰り返される。ここで、MWRのセットは、情報提供白長方形IWRを含んでいる。これらのIWRは、レイアウト要素に対応する複数のフィールドに、画像を分割するアルゴリズムのための開始点を形成する。IWRは、有力なフィールドセパレータであり、そのため、「分離要素」と呼ばれる。アルゴリズムは、IWRを使用して、画像の地理的記述へと更に処理するためにグラフを構成する。

【0053】

図8は、新聞の一面におけるそのようなグラフを示している。画像は、ダウンサンプリングされた新聞の一面のデジタル画像80を示している。オリジナルテキストは、図2に対応するダウンサンプリングされたバージョンにおいて黒で見ることができる。分離要素を構成する情報提供長方形IWRが、灰色で示されている。グラフの構成のため、水平方向および垂直方向の白IWRによって構成される分離要素の交差部が決定される。2つのIWRの交点は、頂点すなわちグラフの頂点81を示す小さな黒色正方形で表わされている。一面内でフィールドを分離するラインを示す辺82は、頂点81から成る対を「フィールドセパレータ」によって接続することにより構成される。グラフの辺82が白で示されている。辺の2つの頂点間の距離、すなわち長さは、更なる処理のために加重として辺に対して割り当てられる。他の実施形態においては、異なるパラメータが使用され、例えばピクセルの色等の加重が割り当てられる。グラフを構成するためのアルゴリズムは、以下の通りである。

【0054】

初めに、IWRに関して以下の表記法および定義が与えられる。R = {r₁, ..., r_m} が空でなく、また全てのIWRの有限のセットが所定の画像1から得られると仮定する。ここで、各IWRは、それぞれ、その左上角部および右下角部のx座標およびy座標((x₁(), y₁()), (x₂(), y₂())), = 1, 2, ..., mによってによって特定される。各長方形r_iは、その高さと幅との比に基づいて、水平方向、垂直方向、あるいは正方形として分類される。H = {h₁, ..., h₁}, V = {v₁, ..., v_k}, S = {s_i, ..., s_d} は、以下の関係を成すように、水平方向、垂直方向、および正方形のIWRのサブセットを示す。

【数22】

$$H \cup V \cup S = R \text{ 及び } m = l + k + d$$

【数23】

$$H \cap V = \emptyset, V \cap S = \emptyset, H \cap S = \emptyset$$

ここで、以下が仮定される。

【数24】

$$H \neq \emptyset, V \neq \emptyset$$

【0055】

また、Sの内容は無視され、HおよびVのサブセットだけが使用される。これは、多くの場合、テキストブロックまたは非テキストブロックの境界を形成する白空間が、細長い垂直方向または水平方向領域であるといった考えに基づいている。hが座標((x₁, y

$(x_1, y_1), (x_2, y_2)$ を有する H の一部であるとし、また、v が座標 $((a_1, b_1), (a_2, b_2))$ を有する V の一部であるとする。この時、以下の条件の下、h および v はオーバーラップを有する。

【数 2 5】

$$\begin{cases} x_1 \leq a_2 \\ y_1 \leq b_2 \\ x_2 \geq a_1 \\ y_2 \geq b_1 \end{cases}$$

【0 0 5 6】

オーバーラップの場合の h および v の交点により、以下の座標によって規定される唯一の点 P が得られる。

【数 2 6】

$$\begin{cases} x_p = \frac{1}{2}(\max\{x_1, a_1\} + \min\{x_2, a_2\}), \\ y_p = \frac{1}{2}(\max\{y_1, b_1\} + \min\{y_2, b_2\}) \end{cases}$$

【0 0 5 7】

IWRにおいては、考えられる全てのタイプのオーバーラップから、2つのオーバーラップだけが生じる。すなわち、長方形となるオーバーラップと、点となるオーバーラップだけが生じる。ラインのオーバーラップは生じない。これは、ラインのオーバーラップが MWR の概念と矛盾するからである。

【0 0 5 8】

図 9 は、最大長方形の 2 つのタイプの交差部を示している。グラフを構成するため、垂直方向および水平方向の情報提供最大長方形の交点が決定され、これにより、グラフの頂点の位置が見出される。すなわち、頂点の正確な座標が決定される。図の左側部分は、垂直方向の IWR である v と、水平方向の IWR である h との交差部の第 1 のタイプを示している。このタイプの交差では、交差部の中心が点 P である長方形領域 8 8 が形成される。図の右側部分は、垂直方向の IWR である v と水平方向の IWR である h との交差部の第 2 のタイプを示している。このタイプの交差では、交差部の中心が P' である 1 つの交点 8 9 が生じる。

【0 0 5 9】

交点に基づいてグラフを構成するアルゴリズムは、以下の通りである。

【0 0 6 0】

$P = \{p_1, \dots, p_N\}$ は、垂直方向の IWR および水平方向の IWR の全ての交点のセットを示している。この場合、P の中の各 p は、その x 座標および y 座標 (x_p, y_p) によって特定される。ここで、 $p = 1, \dots, N$ である。セット P が見出され、 $G = (X, A)$ が P に対応する方向性が無いグラフであるとする。グラフ $G = (X, A)$ は、交点に対して直接に関係する有限の数の頂点 X と、交点間の関係を描く有限の数の辺 A とから成る。これは、数学的には、以下のように表わされる。

【数 2 7】

$$G(P) = (X(P), A(P \times P))$$

$$P : H \times V \rightarrow \{x_p, y_p\}$$

ここで、

【数28】

$$\begin{aligned} X &= \{1, \dots, N\} \text{及び} \\ A &= (\{1, \dots, N\} \times \{1, \dots, N\}) \end{aligned}$$

【数29】

$$A(i, j) = \begin{cases} \infty, i \text{と} j \text{が、4連鎖接続されていないとき} \\ d_{ij}, i \text{と} j \text{が、4連鎖接続されているとき} \end{cases}$$

ここで、 d_{ij} は、点 i と点 j の間のユークリッド距離を示している。また、4連鎖接続 (4-chain connected) は、4つの可能な移動方向で長方形ブロックの頂点が接続されることを意味している。前述した2つの点 i, j は、 $\min d_{ij}$ を有する4つの接続連鎖コードによって1方向に動き回ることにより、これらの点に達し得る場合には、4連鎖接続である。

【0061】

ここで、構成されたグラフを更に処理して、グラフの内部の領域をテキストブロックとして分類し、あるいは、画像のタイプに応じて同様の分類を行なっても良い。一実施形態において、グラフは、解析において、例えば黒ラインまたは破線 / 点線等のパターン化されたラインといったフォアグラウンドセパレータを含めることにより増大される。また、検出される写真またはグラフィックオブジェクトのエッジを、解析に含めることもできる。

【0062】

また、この分割方法は、フォアグラウンドセパレータを除去するステップを含んでいても良い。まず最初に、フォアグラウンドセパレータが認識されて、1つのオブジェクトとして再構成される。パターン化されたラインを構成する成分は、要素ヒューリスティックス、空間的関係ヒューリスティックス、およびラインヒューリスティックスを解析することにより、すなわち、1方向で組み合わされた要素を構築し且つそれがラインとして分類するか否かを検出することにより接続される。パターン化されたラインから実線を再構成するための更なる方法は、ダウンサンプリングおよび / または「Document analysis system (文書解析システム)」IBM J. Res. Dev. 26 (1982) 647 - 656において、K. Y. Wong, R. G. Casey, F. M. Wahにより説明されている、ラン・レンゲス・スマージング・アルゴリズム (RLSA) 使用することである。フォアグラウンドセパレータの検出後、フォアグラウンドセパレータは、バックグラウンドピクセルに取って代えられる。その結果、より大きな最大白長方形を構成することができ、あるいは、バックグラウンドピクセル特性を使用して、バックグラウンドセパレータを見出す任意の他の適した方法をサポートすることができる。

【0063】

図11は、フィールドセパレータに基づいてフィールドを規定する方法に関する図を示している。

【0064】

基本的に、この方法のタスクは、画像内でフィールドを規定することである。この場合、フィールドは、新聞画像内のテキストブロック等といった、相互に関連するフォアグラウンド要素を含む領域として規定される。画像内のフィールドは、所定の方向を有し且つ厚さが0である幾何学的なラインであると理解されるフィールドセパレータによって分離される。フィールドセパレータは、分離方向、一般には水平または垂直方向で、長方形状を有する連結されたバックグラウンドピクセルの領域に対応している。フィールドセパレータの交差点は、ノードと呼ばれている。本方法においては、最初に、画像内のフィールドセパレータが検出され、その後、フィールドセパレータの解析に基づいて、フィールド

が決定される。

【0065】

S E P A R ステップ95においては、画像が解析されることにより、フィールドセパレータが得られる。フィールドセパレータは、前述した最大白長方形を使用する解析に基づいていることが好ましい。最大白長方形を使用する解析では、辺と辺が接続する頂点とを有するグラフが得られる。本発明の方法において、フィールドセパレータおよびノードは、グラフの辺および頂点にそれぞれ対応している。また、フィールドセパレータを決定するために、他の適した方法を使用しても良い。なお、セパレータを得るための処理は、先立って既に終了してしまっていても良い。すなわち、画像は、既にセパレータを示す更に高いレベルでの構造の表示である。

【0066】

このようにして見つけられたフィールドセパレータは、例えば走査のアライメントミス等により、基本的な水平方向および垂直方向から僅かに外れるかもしれません、更なる処理ステップにおいてエラーを招く可能性がある。したがって、この時点で、フィールドセパレータのx座標またはy座標の小さなズレを0にする、「スナップ・トゥ・グリッド(snap to grid)」ステップが処理に加えられても良い。

【0067】

T E S S ステップ96においては、ビルディングブロックレベルへの変換が行なわれる。このステップにおいては、フィールドセパレータが画像の外側の境界に達するまでフィールドセパレータを延長することにより、画像内のフィールドのビルディングブロックを形成する基本的な複数の長方形に画像が分割される。このようにして、所謂テセレーショングリッドが形成され、(延長された)フィールドセパレータによって囲まれた領域が、基本長方形として規定される。

【0068】

テセレーショングリッドの形成については、図12および図13を参照しながら詳細に後述する。

【0069】

基本的に、ここで、方法は、フィールドセパレータによって分離されない基本長方形をフィールドに連結する。この処理を行なうための特に有効な方法は、以下のステップを含んでいる。

【0070】

M A T R I X ステップ97においては、テセレーション画像の新たな表示が、マトリクスマップの形態で形成される。マトリクスマップにおいて、基本長方形およびテセレーショングリッド要素は、マトリクス要素によって表わされる。このステップについては、図14を参照しながら更に後述する。

【0071】

C O N N ステップ98においては、基本長方形が連結され、これにより、連結基本長方形の領域が形成される。基本長方形は、これらがラインの延長部分によって分離される場合に、連結されていると見なされ、フィールドセパレータに関連付けられたライン部分によって分離される場合に、連結されていないと見なされる。連結成分アルゴリズムは、図14に関して後述するように、このステップで使用される。

【0072】

このステップでこの時に決定される連結基本長方形のセットは、オリジナル画像のフィールドに対応している。

【0073】

N O D E ステップ99においては、オリジナル画像内においてフィールドの位置を規定するために、C O N N ステップで見出されたフィールドを境界付けるオリジナルノードが検索される。

【0074】

最後に、F I E L D ステップ100において、先のステップで検索されたオリジナルノ

ードは、連結基本長方形の各領域毎にフィールドを規定するデータ構造に組み合わされる。これにより、マトリクス表示が変換されてピクセルドメインに戻される。このステップについては、図15から図17を参照しながら更に後述する。

【0075】

ここで、アルゴリズムのT E S Sステップについて更に詳しく説明する。

【0076】

図12は、画像の表示を示している。画像は、フィールド109を囲むフィールドセパレータ110に関連付けられたラインによって表わされている。フィールドセパレータ110は、バックグラウンド、通常は新聞の白を表わしており、黒ラインとして示される。例えばこの例ではフィールド109であるフィールドセパレータ間のフォアグラウンド領域は、フィールドとして規定される。実行されるタスクは、画像内のフィールドを識別することである。

【0077】

図13は、図12の入力画像に基づく画像上のテセレーショングリッドを示している。テセレーショングリッドを形成するため、全てのフィールドセパレータ（図13の連続ライン110）が画像の境界まで延長された。その結果、画像は、4つのXセグメント X_1 から X_4 の垂直ライン、および6つのYセグメント Y_1 から Y_6 の水平ラインによって小分割される。フィールドセパレータ110の延長が、破線111によって示されている。例えば、ノード2、6はフィールドセパレータの実際のノードであり、延長により、ノード2とノード6との間に仮想ノード116が形成される。ノード2とノード6との間のラインの直ぐ右側の領域には、2つの基本長方形が形成される。フィールドセパレータの延長に基づくラインによって形成されるテセレーショングリッド内の全ての長方形は、所謂基本長方形である。例えば、基本長方形113は、影が付けられた領域によって示される連結領域の一部である。前記連結領域は、フィールドセパレータによって基本長方形113から分離されない全ての基本長方形によって構成される。連結基本長方形の領域は、図14に関して後述するように、簡単に構成することができる。

【0078】

なお、このような方法は、実質的に長方形構造ではない領域にも拡大して適用することができることに留意されたい。平面グラフの区分的線形化および／または弾力的変形は、「湾曲した境界付けられた」領域を含む画像を処理するために適用できる。

【0079】

基本的なアルゴリズムのM A T R I Xステップにおいて、図13に示されるテセレーション画像は、マトリクス表示に変換される。このマトリクス表示においては、全ての基本長方形および全てのラインセグメントが、マトリクス要素に関連付けられる。テセレーション画像は、水平方向で横切る場合には、4つの基本長方形とフィールドセパレータに関連付けられた5つの垂直ラインとにまたがる。したがって、マトリクス表示は9個の縦列を有する。テセレーション画像は、垂直方向で横切る場合には、6つの基本長方形と7つの水平ラインとにまたがる。したがって、マトリクス表示は13個の横列を有する。

【0080】

最初に、全てのマトリクス要素に値1が与えられる。その後、全てのマトリクス要素は、オリジナル画像のフィールドセパレータに関連付けられているか否かが体系的にチェックされる。関連付けられている場合には、マトリクス要素の値1が値0に変更される。したがって、フォアグラウンド要素が1で表わされ、バックグラウンド要素が0で表わされる。

【0081】

また、フィールドセパレータのリストをチェックすることにより、マトリクス要素が0に変更されても良い。フィールドセパレータのリストをチェックすることは、通常、演算を殆ど必要としない。

【0082】

図14は、結果として得られる図13の画像のマトリクスマップ120を示している。

例えば、基本長方形 113 は、ここでは、マトリクスの 1 つの要素 123 に換算され、また、延長されたラインセグメント 111 は、ここでは、マトリクスの要素 121 となる。ノード 2、6 は、要素 124、125 で表わされている。また、仮想ノード 116 に対応するマトリクス要素も示されている。この要素は値 1 を有している。なぜなら、この要素は、フィールドセパレータの一部だからである。なお、地理的形状は保たれていないことに留意されたい。これは、ノード間のラインの長さが考慮されていないためである。画像およびテセレーショングリッドの表示におけるオリジナルノード間の関係は、図 17 を参照して後述するように個別に記憶される。

【0083】

領域 109（図 12）は、全て 1 となる要素から成る影付き領域 122 として図 14 に示されている。

【0084】

アルゴリズムの CONN ステップにおいて、形成されたマトリクスマップには、その後、1 の値を有する連結要素のセットをマトリクス内で見つけるための連結成分処理が施される。連結成分アルゴリズムは、文献で広く知られているため、本明細書では、これ以上説明しない。

【0085】

以下、アルゴリズムの NODE ステップについて更に詳細に説明する。一例として、図 15 は、図 14 のマトリクス内の 1 つの連結領域 130 を示している。図示のマトリクスは、前述したテセレーショングリッドに基づいているが、連結成分処理によって検出される連結領域 130 だけが、影付き領域で示されている。連結領域の構成要素は、値 1 を有するとともに、0 の値を有する要素によって囲まれている。以下のステップにおいて、フィールドは、連結領域の周囲の外形に基づいて規定される。

【0086】

図 16 は、連結領域の外形 140 を示している。この外形は、連結領域 130 に対応する値 0 を有する領域が、値 1 を有する影付き領域によって取り囲まれることにより示される。外形を見出すため、最初に、領域 130 が 1 ピクセル分だけ拡張される。その後、オリジナル領域が取り去られる。

【0087】

図 17 はノードマトリクスを示している。このマトリクスは、マトリクスマップと同じ大きさを有している。要素の値は、ノード番号（0 から 19）または空のいずれかである。ノード番号は、図 12 に示されるオリジナル画像のノードを参照する。先に得られた連結領域 130 の外形 140 は、ノードマトリクス上に描出され、影付き領域 141 で示されている。

【0088】

ノードマトリクスは、以下のように構成される。最初に、要素の値が「空」に設定される。その後、例えばグラフの頂点リストに基づいて、フィールドセパレータの実際のノードがマトリクスに入力される。

【0089】

タスクは、領域 130 の外形 140 に属する全てのノードを抽出することである。外形内に存在するノードは、外形をトレースしてその中のノードを表示することにより検索される。

【0090】

外形をトレースした後、ノードは、アルゴリズムの FIELD ステップにおいて、オリジナル画像表示に組み合わせられる。必要に応じて、「スナップ・トゥ・グリッド（snap-to-grid）」処理の逆が適用され、ノード番号が、再びノードのオリジナルセットと一体になる。最後に、必要に応じて、ノードおよびノードの辺が、例えば時計周り方向で順序付けられる。この順序付けは、領域演算または表示のために必要とされても良い。

【0091】

無論、ノード抽出およびフィールド決定は、画像内の全てのフィールドに関して行なわれなければならない。

【0092】

なお、領域が互いに取り囲むことにより、分離した多角形、例えばテキスト包囲体を形成しても良いことに留意されたい。複数の分離した多角形によって境界付けられる領域に関して演算できるように、これらの多角形を連結する知られている技術が使用される。多角形の2つの外形は、所謂「0領域ブリッジ」により、実際には2つのラインセグメントにより連結される。ラインセグメントの一方は内側の外形に入り、ラインセグメントの他方は内側の外形から出る。

【0093】

図10は、画像を分割するための本発明に係る方法が実施される装置を示している。この装置は、デジタル画像を入力するための入力ユニット91を有している。入力ユニットは、紙からの画像を走査する電子光学スキャナ等の走査ユニット、またはインターネット等のネットワークからの画像を受けるデジタル通信ユニット、または光ディスクドライブ等の記録キャリアからデジタル情報を検索する再生ユニットを備えていても良い。入力ユニット91は、記憶ユニット92と協働する処理ユニット94に接続されている。処理ユニットは、汎用コンピュータ中央処理ユニット(CPU)および支援回路を備えていても良く、前述した分割を行なうためのソフトウェアを使用して動作する。特に、ソフトウェアは、テセレーショングリッドを構成するためのモジュール(図面には単独で示されていない)を有している。このモジュールは、フィールドセパレータを画像の外側の境界まで延長して基本長方形を構成し、フィールドセパレータによって分離されない隣り合う基本長方形同士を連結することによりフィールドを構成することによって、テセレーショングリッドを構成する。また、ソフトウェアは、テセレーショングリッドを表示するマトリクスマップを構成し、且つテセレーショングリッド内のノードに関連するノードマトリクスを構成するモジュールを有している。

【0094】

処理ユニットは、キーボード、マウス装置、またはオペレータボタン等の制御手段を備えたユーザインタフェース95を更に有していても良い。処理ユニットの出力部は、ディスプレイユニット93に接続されている。一実施形態において、ディスプレイユニットは、処理された画像を紙上に出力するための印刷ユニット、または磁気テープや光ディスク等の記録キャリア上に分割された画像を記憶する記録ユニットである。

【0095】

分割されるデジタル画像として、新聞の一面の実施形態によって本発明を主に説明してきたが、本発明は、例えばIC設計のためのレイアウト画像における電気回路や、シティーマップ上の通りや建物など、バックグラウンド上にフィールドを備える任意のデジタル表示にも適している。また、最短のサイクルによって分割を実行するための開始点としてのグラフは、MWRシステムに基づく前述したグラフとは異なる方法で構成されても良い。例えば、グラフは、前述したAntonacopoulosによる論文に記載されたタイルを使用して構成されても良い。また、グラフの辺に割り当てられる加重は、必ずしも距離でなくても良い。加重は、最短サイクルに寄与するように対応して選択されなければならない。例えば、加重は、タイルの表面であっても良い。なお、この明細書において、動詞「備える」、「含む」、「有する」及びその活用の用法は、記載された要素以外の他の要素またはステップの存在を排除しない。また、要素に先立つ单語「1つの」は、そのような要素が複数存在することを除外せず、任意の参照符号は、特許請求の範囲を限定しない。また、本発明、および全てのユニット、または前述した手段は、適当なハードウェアおよび/またはソフトウェアによって実施することができ、幾つかの「手段」または「ユニット」は、同じ要素によって表わすことができる。また、本発明の範囲は実施形態に限定されず、本発明は、前述した個々の新規な特徴、全ての新規な特徴、または特徴の組み合わせにある。

【図面の簡単な説明】

【0096】

- 【図1】典型的な分割方法の概略を示している。
- 【図2】サンプルとしての日本の新聞の一部を示している。
- 【図3】1方向に沿うオブジェクトの統合を示している。
- 【図4】オブジェクトの分割および2方向統合を示している。
- 【図5】ホワイトランからの最大長方形の構成を示している。
- 【図6】最大白長方形の構成を示している。
- 【図7】最大白長方形をオーバーラップさせるクリーニングを示している。
- 【図8】新聞の一面上のグラフを示している。
- 【図9】最大長方形の2つのタイプの交差部を示している。
- 【図10】画像を分割するための装置を示している。
- 【図11】フィールドセパレータに基づいてフィールドを規定するための方法に関する図を示している。
- 【図12】画像の表示を示している。
- 【図13】画像のテセレーショングリッドを示している。
- 【図14】テセレーショングリッドのマトリクスマップを示している。
- 【図15】マトリクス内の1つの連結領域を示している。
- 【図16】連結領域の外形を示している。
- 【図17】ノードマトリクスを示している。

【符号の説明】

【0097】

- 1 1 入力画像
- 1 2 C C オブジェクト
- 1 3 レイアウトオブジェクト
- 1 4 C C A モジュール
- 1 5 L A モジュール
- 1 6 A F モジュール
- 1 7 記事
- 2 1 垂直方向読み方向
- 2 2 水平方法読み方向
- 2 3 黒ライン
- 3 1 最初の解析ステップ
- 3 2 算定閾値
- 3 3 分類ステップ
- 3 4 補正された連結成分
- 3 5 統合ステップ
- 3 6 オブジェクト
- 3 7 テキスト統合ステップ
- 3 8 テキストブロック
- 4 1 画像のグラフ
- 4 2 フィールド検出ステップ
- 4 3 連結成分
- 4 4 ステップ
- 4 5 読む順番
- 4 6 ライン形成ステップ
- 4 7 テキストブロック
- 5 1 ホワイトラン
- 5 2 最大白長方形
- 5 3 フォアグラウンド領域
- 6 1 ホワイトラン

6 2 、 6 3 、 6 4 、 6 5 、 6 6 最大白長方形
8 0 デジタル画像
8 1 グラフの頂点
8 2 辺
8 8 長方形領域
8 9 交点
9 1 入力ユニット
9 2 記憶ユニット
9 3 ディスプレイユニット
9 4 処理ユニット
9 5 ユーザインターフェース
9 6 T E S S ステップ
9 7 M A T R I X ステップ
9 8 C O N N ステップ
9 9 N O D E ステップ
1 0 0 F I E L D ステップ
1 0 9 フィールド
1 1 0 フィールドセパレータ
1 1 1 ラインセグメント
1 1 3 基本長方形
1 1 6 仮想ノード
1 2 0 マトリクスマップ
1 2 1 、 1 2 3 、 1 2 4 、 1 2 5 要素
1 2 2 、 1 4 1 影付き領域
1 3 0 連結領域
1 4 0 外形