

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6559415号  
(P6559415)

(45) 発行日 令和1年8月14日(2019.8.14)

(24) 登録日 令和1年7月26日(2019.7.26)

(51) Int.Cl.	F I				
<b>G06F</b>	<b>17/21</b>	<b>(2006.01)</b>	G06F	17/21	610
<b>G06T</b>	<b>11/60</b>	<b>(2006.01)</b>	G06T	11/60	100A
<b>G06K</b>	<b>9/20</b>	<b>(2006.01)</b>	G06K	9/20	340L
<b>H04N</b>	<b>1/387</b>	<b>(2006.01)</b>	G06K	9/20	340K
			H04N	1/387	

請求項の数 8 (全 40 頁)

(21) 出願番号	特願2014-235989 (P2014-235989)	(73) 特許権者	000005049
(22) 出願日	平成26年11月20日(2014.11.20)		シャープ株式会社
(65) 公開番号	特開2016-99793 (P2016-99793A)		大阪府堺市堺区匠町1番地
(43) 公開日	平成28年5月30日(2016.5.30)	(74) 代理人	110000338
審査請求日	平成29年9月25日(2017.9.25)		特許業務法人HARAKENZO WORLD PATENT & TRADEMARK
前置審査		(72) 発明者	松岡 輝彦
			大阪府大阪市阿倍野区長池町22番22号
			シャープ株式会社内
		(72) 発明者	高島 真彦
			大阪府大阪市阿倍野区長池町22番22号
			シャープ株式会社内

最終頁に続く

(54) 【発明の名称】 文書画像処理装置、それを備えた情報処理装置、プログラム、及び記録媒体

(57) 【特許請求の範囲】

【請求項1】

文書画像処理装置において、  
 文書を電子化した文書画像の構造解析を行う構造解析部と、  
 前記構造解析により前記文書画像から抽出した文字列の記述方向、数、1文字あたりの大きさ、および位置の少なくともいずれかを含む前記文字列の特徴量が所定の条件を満たすか否かを判定し、満たすと判定した場合は、前記文書画像に含まれる各文字、図、及び/又は表である各要素を表示領域に合わせて該文書画像中の文字列を折り返し表示することが可能な形式に再構成すると判定する変換判定部と、  
 前記構造解析部による解析結果に基づいて、前記文書画像の各要素の順序を記述した、再構成を行うための参照リストを生成する参照リスト生成部と、を備えており、  
 前記変換判定部が前記文書画像の各要素を再構成すると判定すると、前記参照リスト生成部を用いて前記参照リストを生成し、  
 前記特徴量は前記文字列の記述方向であり、前記変換判定部は、縦書きの行および横書きの行の比率が所定閾値より大きい場合に前記所定の条件を満たさないと判定し、前記文書画像を再構成しないと判定し、  
 前記変換判定部により各要素を再構成すると判定された文書画像については、当該文書画像の前記参照リスト及び当該文書画像に含まれる各要素の画像データを送信装置に出力することを特徴とする文書画像処理装置。

【請求項2】

文書画像処理装置において、

文書を電子化した文書画像の構造解析を行う構造解析部と、

前記構造解析により前記文書画像から抽出した文字列の記述方向、数、1文字あたりの大きさ、および位置の少なくともいずれかを含む前記文字列の特徴量が所定の条件を満たすか否かを判定し、満たすと判定した場合は、前記文書画像に含まれる各文字、図、及び/又は表である各要素を表示領域に合わせて该文書画像中の文字列を折り返し表示することが可能な形式に再構成すると判定する変換判定部と、

前記構造解析部による解析結果に基づいて、前記文書画像の各要素の順序を記述した、再構成を行うための参照リストを生成する参照リスト生成部と、を備えており、

前記変換判定部が前記文書画像の各要素を再構成すると判定すると、前記参照リスト生成部を用いて前記参照リストを生成し、

前記特徴量は横書きの文字列の行の1文字あたりの高さあるいは縦書きの文字列の行の1文字あたりの幅であり、前記変換判定部は、前記高さあるいは前記幅が予め定められた第2閾値以上である場合に前記所定の条件を満たさないと判定し、前記文書画像を再構成しないと判定し、

前記変換判定部により各要素を再構成すると判定された文書画像については、当該文書画像の前記参照リスト及び当該文書画像に含まれる各要素の画像データを送信装置に出力することを特徴とする文書画像処理装置。

#### 【請求項3】

文書画像処理装置において、

文書を電子化した文書画像の構造解析を行う構造解析部と、

前記構造解析により前記文書画像から抽出した文字列の記述方向、数、1文字あたりの大きさ、および位置の少なくともいずれかを含む前記文字列の特徴量が所定の条件を満たすか否かを判定し、満たすと判定した場合は、前記文書画像に含まれる各文字、図、及び/又は表である各要素を表示領域に合わせて该文書画像中の文字列を折り返し表示することが可能な形式に再構成すると判定する変換判定部と、

前記構造解析部による解析結果に基づいて、前記文書画像の各要素の順序を記述した、再構成を行うための参照リストを生成する参照リスト生成部と、を備えており、

前記変換判定部が前記文書画像の各要素を再構成すると判定すると、前記参照リスト生成部を用いて前記参照リストを生成し、

前記特徴量は複数の文字列あるいは図又は表よりなる複数のブロック間の位置のずれの大きさであり、前記変換判定部は、前記ずれの大きさが予め定められた第3閾値より大きい場合に前記所定の条件を満たさないと判定し、前記文書画像を再構成しないと判定し、

前記変換判定部により各要素を再構成すると判定された文書画像については、当該文書画像の前記参照リスト及び当該文書画像に含まれる各要素の画像データを送信装置に出力することを特徴とする文書画像処理装置。

#### 【請求項4】

請求項1から3の何れか1項に記載の文書画像処理装置と、

前記変換判定部により各要素を再構成しないと判定された文書画像をフィックス型の表示が可能にフォーマット変換するフォーマット変換処理部と、

前記変換判定部により各要素を再構成すると判定された文書画像については、当該文書画像の前記参照リスト及び当該文書画像に含まれる各要素の画像データを送信し、かつ、前記変換判定部により各要素を再構成しないと判定された文書画像については、前記フォーマット変換された文書画像を送信する送信装置と、を備えることを特徴とする情報処理装置。

#### 【請求項5】

請求項4に記載の情報処理装置から受信した文書画像を表示する表示装置であって、

前記変換判定部により各要素を再構成すると判定された文書画像の前記参照リスト及び当該文書画像に含まれる各要素の画像データに基づき、当該文書画像をリフロー型にて表示し、かつ、前記フォーマット変換された文書画像をフィックス型にて表示することを特

10

20

30

40

50

徴とする表示装置。

【請求項 6】

請求項 4 に記載の情報処理装置から、さらに前記変換判定部により各要素を再構成すると判定された文書画像をフィックス型の表示が可能なフォーマット変換した文書画像を受信し、

前記変換判定部により各要素を再構成すると判定された文書画像の表示を、リフロー型とフィックス型とで切り替える切替部を備えたことを特徴とする請求項 5 に記載の表示装置。

【請求項 7】

請求項 1 から 3 のいずれか 1 項に記載の文書画像処理装置を動作させるためのプログラムであって、コンピュータを上記の各部として機能させるためのプログラム。

【請求項 8】

請求項 7 に記載のプログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、文書画像の再構成を行う文書画像処理装置、それを備えた情報処理装置、コンピュータプログラム、及び記録媒体に関する。

【背景技術】

【0002】

近年、複写機又は複合機等の画像形成装置は高機能化が進められており、スキャナにより読み取った文書を文書画像データ（以下、文書画像）として保存し、保存した文書画像を管理する機能等が求められている。スキャナにより読み取った文書は画像データとして保存されるが、この画像データの形式は、フィックス型と呼ばれる固定の幅及び高さを持つファイル形式となっている。

【0003】

代表的なファイル形式として、例えば、PDF（Portable Document Format）ファイルやTIFF（Tagged Image File Format）ファイル等が挙げられる。これらのファイル形式の文書画像を、携帯電話、スマートフォン、タブレット等表示領域の小さな画像表示装置で表示すると、その表示領域に収まりきらず、垂直方向のスクロール操作と水平方向のスクロール操作の両方が必要となる場合がある。この場合、操作が非常に煩雑になる。

【0004】

そこで、例えば、表示領域の画素数に合わせて表示倍率を調整することにより、文書画像の幅を表示領域の幅に合わせて縮小表示することで、行方向のスクロール操作の省略が可能となる。しかし、縮小処理を実施することで文書画像中の文字の可読性が低下してしまう。そのため、読み取られた文書画像を、フィックス型のファイルではなく、表示領域に合わせて折り返し表示することが可能なリフロー型のファイルとして提供することが望ましい。

【0005】

リフロー型のファイルは固定の幅及び高さを持たず、画像表示装置の表示領域の範囲で行を自動的に折り返すことで、1行の文を表示領域からはみ出させることなく表示できる。よって、行方向にスクロールすることなく、行方向と直交する方向のスクロール操作のみで文書を読むことが可能となる。HTML（Hyper Text Markup Language）や、スマートフォンやタブレット向けに展開されている電子書籍機能が提供するファイル形式は、リフロー型の表示が可能なファイル形式の例である。前記読み取られた文書画像をリフロー型のファイル形式に変換することにより、上記の可読性の問題は解決される。

【0006】

ファイル形式を変換する装置として、例えば、特許文献 1 の文書ファイル表示装置がある。特許文献 1 の表示装置は、構造化された文書ファイル（doc、txt、odf、xls等）を文

10

20

30

40

50

書画像のファイル(jpeg、tiff、bmp等)に変換し、変換した文書画像から、文書を構成する個々の要素の存在領域及び要素の並び方向を含むレイアウト情報を検出する。そして、検出したレイアウト情報に基づき、個々の要素の存在領域に相当する要素画像(部分画像)を文書画像から抽出し、抽出した各要素画像の要素の並び方向に沿ったサイズに基づき、行情報を作成する。次に、作成した行情報のスクロール方向を決定し、決定したスクロール方向に沿って複数の行情報を配列することで、段落情報を作成する。そして、作成した段落情報を、表示部の表示範囲内でスクロール表示する。よって、特許文献1の文書ファイル表示装置では、表示領域と同じ横幅の文書画像を表示するため、一方向のスクロール操作のみで文書画像を閲覧することが可能となる。

【先行技術文献】

10

【特許文献】

【0007】

【特許文献1】特開2012-230623号公報(2012年11月22日公開)

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかしながら、特許文献1に記載の文書ファイル表示装置は、ユーザが指定した文書ファイルであれば、リフロー型に不向きな文書でも関係なくリフロー型の文書に変換してしまう。そのため、表やインデントを多用して文章を構成している文書など、元々意味を持っていたレイアウトの文書では、レイアウトが崩れてしまい、リフロー型の文書に変換したために内容が理解できなくなるといったことが起こり得る。

20

【0009】

そこで、本発明は、前述した問題に鑑みなされたものであり、文書画像をリフロー型に変換するか否かを判断し、常に最適なフォーマットでの表示が可能なように文書画像を処理する文書画像処理装置等を提供することを目的とする。

【課題を解決するための手段】

【0010】

上記の課題を解決するために、本発明の一態様に係る文書画像処理装置は、文書を電子化した文書画像の再構成を行う文書画像処理装置において、前記文書画像の構造解析を行う構造解析部と、前記構造解析により前記文書画像から抽出した文字列あるいは図又は表の特徴量に基づいて、前記文書画像に含まれる各文字、図、及び/又は表である各要素を再構成するか否かの判定を行う変換判定部と、前記変換判定部が前記文書画像の各要素を再構成すると判定すると、前記構造解析部による解析結果に基づいて、前記文書画像を再構成した際の上記各要素の順序を記述した参照リストを生成する参照リスト生成部と、を備えることを特徴とする。

30

【発明の効果】

【0011】

上記構成によると、文書画像をリフロー型に変換するか否かを判断し、常に最適なフォーマットでの表示が可能なように文書画像を処理することができる。

【図面の簡単な説明】

40

【0012】

【図1】本発明の一実施形態に係る画像形成装置の機能的構成を示すブロック図である。

【図2】上記画像形成装置の有する画像処理装置が備える変換処理部の構成を示すブロック図である。

【図3】行頭禁則の文字の例と行末禁則の文字の例を示す図である。

【図4】横書きの文字列の例を示す図である。

【図5】文書画像の例を示す図である。

【図6】行に分類した上記文書画像の例を示す図である。

【図7】(a)は、2段組の横書きの文書の例、(b)は、2段組の縦書きの文書の例を示す図である。

50

【図 8】(a) ~ (f) は、2つの行の行間距離を説明する図である。

【図 9】(a) は、行に分類した文書画像の例、(b) は、(a) をさらに行ブロックに分類した文書画像の例を示す図である。

【図 10】行と行ブロックとに分類した文書画像の例を示す図である。

【図 11】上記変換処理部の有するレイアウト解析処理部の構成を示すブロック図である。

【図 12】(a) は、行ブロックに分類した文書画像の例、(b) は、(a) をさらに段組に分類した文書画像の例、(c) は、(b) をさらにカラムに分類した文書画像の例を示す図である。

【図 13】2段組構成の文書画像の例を示す図である。

10

【図 14】行、行ブロック、段(カラム)、及び段組に分類した文書画像の例を示す図である。

【図 15】行順序リストの例を示す図である。

【図 16】行ブロック、カラム、及び段組についての情報を示す図である。

【図 17】文書構造ツリーの構造の例を示す図である。

【図 18】上記レイアウト解析処理部の有する段落解析処理部における改行判定処理の概要を示すイメージ図である。

【図 19】上記段落解析処理部の処理手順を示すフローチャートである。

【図 20】行IDバッファの更新処理の概要を示すイメージ図である。

【図 21】初期化された文書構造ツリーの例を示す図である。

20

【図 22】更新された行順序リストの例を示す図である。

【図 23】行順序リストに従って生成された文書構造ツリーの構造の例を示す図である。

【図 24】段落に分類した文書画像の例を示す図である。

【図 25】上記変換処理部の有する再配置処理部の処理手順を示すフローチャートである。

【図 26】HTML 言語で記述されたファイルの例を示す図である。

【図 27】ファイル記述処理の処理手順を示すフローチャートである。

【図 28】CSS 形式で記述したスタイルシートの外部ファイルの例を示す図である。

【図 29】(a) は、CSS 形式で記述したスタイルシートの外部ファイルの例を示す図であり、(b) は、HTML 言語で記述された参照リストの例を示す図である。

30

【図 30】本発明の別の実施形態に係る画像読取装置の機能的構成を示すブロック図である。

【図 31】ブロック位置が揃っていない文書画像の例を示す図である。

【図 32】(a) ~ (c) は、文書画像から線ベースのグラフを抽出する処理を説明する図である。

【発明を実施するための形態】

【0013】

以下に、本発明の実施の形態を図面に基づき詳述する。

【0014】

〔実施の形態 1：画像形成装置〕

40

以下の説明では、本発明に係る文書画像処理装置が変換処理部として画像処理装置の一部を成し、また、その画像処理装置が画像形成装置の一部を成す形態を例示する。

【0015】

〔1. 画像形成装置〕

図 1 は、実施の形態 1 に係る画像形成装置(情報処理装置) 100 の機能的構成を示すブロック図である。画像形成装置 100 は、コピー機能及びスキャナ機能等を有するデジタル複合機である。画像形成装置 100 は、画像処理装置 1、画像入力装置 2、画像出力装置 3、及び送信装置 4 を備えている。

【0016】

画像入力装置 2、画像処理装置 1、画像出力装置 3 及び送信装置 4 には、操作パネル 6

50

が接続されている。操作パネル6は、ユーザが画像形成装置100の動作モードを設定するための設定ボタン及びテンキー等の操作部(図示せず)と、液晶ディスプレイ等で構成される表示部(図示せず)とを備える。

【0017】

画像形成装置100で実行される各種処理は、図示しない制御部(CPU(Central Processing Unit)あるいはDSP(Digital Signal Processor)等のプロセッサを含むコンピュータ)が制御する。画像形成装置100の制御部は、図示しないネットワークカード及びLANケーブルを介して、ネットワークに接続されたコンピュータ及び他のデジタル複合機等とデータ通信を行う。

【0018】

以下、画像形成装置100の各部について詳述する。

【0019】

画像入力装置2は、原稿から画像を光学的に読み取る。画像入力装置2は、例えばCCD(Charge Coupled Device)を有するカラーキャナよりなり、原稿からの反射光像を、CCDを用いてRGB(R:赤, G:緑, B:青)のアナログ信号として読み取り、画像処理装置1へ出力する。画像入力装置2は、スキャナでなくてもよく、例えばデジタルカメラ等であってもよい。

【0020】

画像処理装置1は、画像入力装置2が読み取った画像データに処理を施し、処理を施した画像データを保存、あるいは、送信するために圧縮ファイルを生成する。

【0021】

画像処理装置1は、画像入力装置2から入力されたRGBのアナログ信号に対して、A/D変換部11、シェーディング補正部12、原稿種別判別部13、入力階調補正部14、及び領域分離処理部15にて各後述する画像処理を実行することによって、RGBのデジタル信号(以下、RGB信号という)からなる画像データを生成する。

【0022】

また、画像処理装置1は、領域分離処理部15が出力したRGB信号に対して色補正部16、黒色生成下色除去部17、空間フィルタ処理部18、出力階調補正部19、及び階調再現処理部20にて各後述する画像処理を実行することによって、CMYK(C:シアン, M:マゼンタ, Y:イエロー, K:ブラック)のデジタル信号からなる画像データを生成して、ストリームとして画像出力装置3へ出力する。なお、画像出力装置3へ出力される前に、画像データが記憶部5に一旦記憶されてもよい。記憶部5は、不揮発性の記憶装置(例えばハードディスク)である。

【0023】

画像出力装置3は、画像処理装置1が生成した画像データに基づいて画像を出力する。画像出力装置3は、画像処理装置1から入力された画像データに基づいて、熱転写、電子写真、又はインクジェット等の方式により、記録シート(例えば記録用紙等)上にカラー画像を形成(印刷)して出力する。

【0024】

本実施形態では、画像出力装置3はカラー画像を出力する構成とするが、記録シート上にモノクローム(白黒)画像を形成して出力する構成であってもよい。この場合、画像処理装置1にて、カラー画像の画像データがモノクローム画像の画像データに変換されてから画像出力装置3へ出力される。

【0025】

更にまた、画像処理装置1は、領域分離処理部15が出力したRGB信号に対して圧縮処理部21にて画像圧縮処理を実行することによって、圧縮されたカラー画像の画像データを有する圧縮ファイルを生成し、送信装置4へ出力する。なお、送信装置4へ出力される前に、圧縮ファイルが記憶部5に一旦記憶されてもよい。

【0026】

画像処理装置1は、操作パネル6においてフォーマット変換モードが選択されている場

10

20

30

40

50

合、領域分離処理部 15 が出力した RGB 信号に対して、変換処理部 22 にてフォーマット変換処理を実行する。後述のように、変換処理部 22 の処理によって、画像が有する文書レイアウトを解析して文書構造ツリーを生成する。

【0027】

変換処理部 22 は、本発明に係る文書画像処理装置として機能する。また、送信装置 4 へ出力する前に、変換されたファイルを記憶部 5 に一旦記憶してもよい。また、画像入力装置 2 から入力される文書画像が複数ページにわたる場合、操作パネル 6 において指定したページのみ、後述のように、文書レイアウトを解析して文書の再構成を行うようにすることができる。例えば、表紙ページは再構成の対象とせず、そのままページ全体を画像として出力するといった方法も可能とする。

10

【0028】

送信装置 4 は、画像処理装置が生成した圧縮ファイルを外部へ送信する。送信装置 4 は、図示しない公衆回線網、LAN (Local Area Network) 又はインターネット等の通信ネットワークに接続可能であり、ファクシミリ又は電子メール等の通信方法により、通信ネットワークを介して外部へ圧縮ファイルを送信する。例えば、操作パネル 6 において scan to e-mail モードが選択されている場合、ネットワークカード、モデム等を用いてなる送信装置 4 は、圧縮ファイルを e-mail に添付し、設定された送信先へ送信する。

【0029】

なお、ファクシミリ送信を行う場合は、画像形成装置 100 の制御部が、モデムを用いてなる送信装置 4 にて、相手先との通信手続きを行い、送信可能な状態が確保されたときに、圧縮ファイルに対して圧縮形式の変更等の必要な処理を施してから、相手先に通信回線を介して順次送信する。

20

【0030】

また、ファクシミリを受信する場合、画像形成装置 100 の制御部は、送信装置 4 にて通信手続きを行いながら、相手先から送信されてくる圧縮ファイルを受信して、画像処理装置 1 に入力する。画像処理装置 1 では、受信した圧縮ファイルに対し、不図示の圧縮/伸張処理部で伸張処理が施される。圧縮ファイルを伸張することによって得られた画像データには、必要に応じて、不図示の処理部で回転処理及び/又は解像度変換処理等が施され、また、出力階調補正部 19 で出力階調補正が施され、階調再現処理部 20 で階調再現処理が施される。各種画像処理が施された画像データは、画像出力装置 3 へ出力され、画像出力装置 3 にて、記録シート上に画像が形成される。

30

【0031】

[ 2 . 画像処理装置 ]

以下では、画像処理装置 1 の構成について、画像処理装置 1 における画像処理及びフォーマット変換処理を説明しながら詳述する。

【0032】

A/D 変換部 11 は、画像入力装置 2 から画像処理装置 1 へ入力された RGB のアナログ信号を RGB のデジタル信号 (即ち RGB 信号) に変換する。

【0033】

シェーディング補正部 12 は、A/D 変換部 11 から入力された RGB 信号に対して、画像入力装置 2 の照明系、結像系及び撮像系で生じる各種の歪みを取り除く。

40

【0034】

原稿種別判別部 13 は、シェーディング補正部 12 から入力された RGB 信号を RGB 各色の濃度を示す濃度信号に変換し、文字、写真、又は印画紙等の原稿のモードを判別する原稿種別判別処理を実行する。原稿種別をユーザが操作パネル 6 を用いてマニュアル設定する場合、原稿種別判別部 13 はシェーディング補正部 12 から入力された RGB 信号をそのまま後段の入力階調補正部 14 に出力する。原稿種別判別処理の処理結果 (原稿種別) は、後段の画像処理に反映される。

【0035】

入力階調補正部 14 は、原稿種別判別部 13 から入力された RGB 信号に対して、カラ

50

ーバランスの調整、下地濃度の除去、及びコントラストの調整等の画質調整処理を行う。

【 0 0 3 6 】

領域分離処理部 1 5 は、入力階調補正部 1 4 から入力された R G B 信号が表す画像中の各画素を、文字領域、網点領域、又は写真領域のいずれかに分離する。また、領域分離処理部 1 5 は、分離結果に基づき、各画素がいずれの領域に属しているかを示す領域識別信号を、黒色生成下色除去部 1 7、空間フィルタ処理部 1 8、階調再現処理部 2 0、及び圧縮処理部 2 1 へ出力する。更に、領域分離処理部 1 5 は、入力階調補正部 1 4 から入力された R G B 信号を、そのまま後段の色補正部 1 6、圧縮処理部 2 1 及び変換処理部 2 2 へ出力する。

【 0 0 3 7 】

色補正部 1 6 は、領域分離処理部 1 5 から入力された R G B 信号を C M Y のデジタル信号（以下、C M Y 信号という）へ変換し、色再現の忠実化実現のために、不要吸収成分を含む C M Y 色材の分光特性に基づいた色濁りを C M Y 信号から取り除く。

【 0 0 3 8 】

黒色生成下色除去部 1 7 は、色補正部 1 6 から入力された C M Y 信号に基づき、C M Y 信号から黒色（K）信号を生成する黒色生成処理と、C M Y 信号から黒色生成処理で得た K 信号を差し引いて新たな C M Y 信号を生成する処理とを行う。この結果、C M Y 3 色のデジタル信号は、C M Y K 4 色のデジタル信号（以下、C M Y K 信号という）に変換される。

【 0 0 3 9 】

黒色生成処理の一例としては、一般に、スケルトン・ブラックによる黒色生成を行う方法が用いられる。この方法では、スケルトン・カーブの入出力特性を  $y = f(x)$ 、入力されるデータを C、M、Y、出力されるデータを C'、M'、Y'、K'、UCR（Under Color Removal）率を  $(0 < \alpha < 1)$  とすると、黒色生成下色除去処理は、下記の式（1）～式（4）で表わされる。

$$K' = f(\min(C, M, Y)) \quad \dots (1)$$

$$C' = C - \alpha K' \quad \dots (2)$$

$$M' = M - \alpha K' \quad \dots (3)$$

$$Y' = Y - \alpha K' \quad \dots (4)$$

ここで、UCR 率  $(0 < \alpha < 1)$  とは、C M Y が重なっている部分を K に置き換えて C M Y をどの程度削減するかを示すものである。式（1）は、C M Y の各信号強度の内の最も小さい信号強度に応じて K 信号が生成されることを示している。

【 0 0 4 0 】

空間フィルタ処理部 1 8 は、黒色生成下色除去部 1 7 から入力された C M Y K 信号の画像データに対して、領域分離処理部 1 5 から入力された領域識別信号に基づいてデジタルフィルタによる空間フィルタ処理を行い、空間周波数特性を補正することによって、画像のぼやけ又は粒状性劣化を改善する。例えば、領域分離処理部 1 5 にて文字に分離された領域に対しては、空間フィルタ処理部 1 8 は、文字の再現性を高めるために、高周波成分の強調量が大きいフィルタを用いて空間フィルタ処理を行う。また、領域分離処理部 1 5 にて網点に分離された領域に対しては、空間フィルタ処理部 1 8 は、入力網点成分を除去するためのローパス・フィルタ処理を行う。

【 0 0 4 1 】

出力階調補正部 1 9 は、空間フィルタ処理部 1 8 から入力された C M Y K 信号に対して、画像出力装置 3 の特性である網点面積率に基づく出力階調補正処理を行う。

【 0 0 4 2 】

階調再現処理部 2 0 は、出力階調補正部 1 9 から入力された C M Y K 信号に対して、領域分離処理部 1 5 から入力された領域識別信号に基づいて、領域に応じた中間調処理を行う。例えば、領域分離処理部 1 5 にて文字に分離された領域に対しては、階調再現処理部 2 0 は、高域周波成分の再現に適した高解像度のスクリーンによる二値化又は多値化の処理を行う。また、領域分離処理部 1 5 にて網点に分離された領域に対しては、階調再現処

10

20

30

40

50



理部 20 は、階調再現性を重視したスクリーンでの二値化又は多値化の処理を行う。次いで、階調再現処理部 20 は、処理後の画像データを画像出力装置 3 へ出力する。

【 0043 】

圧縮処理部 21 は、領域分離処理部 15 から入力された領域識別信号と RGB 信号からなる画像データとに基づき、圧縮ファイルを生成する。圧縮処理部 21 に入力される画像データは、マトリクス状に配置されている複数の画素で構成されている。圧縮処理部 21 は、この画像データを、前景レイヤと背景レイヤとに分離する。そして、前景レイヤを更に 2 値画像に変換し、各 2 値画像を例えば MMR (Modified Modified READ) で可逆圧縮する。他方、背景レイヤを例えば JPEG で非可逆圧縮する。最後に、可逆圧縮された 2 値画像及び非可逆圧縮された背景レイヤと、これらを伸張してカラー画像の画像データと成すための伸張情報とを一つのファイルにまとめる。このファイルが圧縮ファイルである。また、この伸張情報としては、圧縮形式を示す情報、及びインデックス・カラー・テーブル (以下、IC テーブルという) 等が用いられる。画素毎に生成された領域識別信号の圧縮は、例えば、可逆圧縮方法である MMR 方式、MR (Modified READ) 方式に基づいて行われる。

10

【 0044 】

変換処理部 22 は、入力された文書画像に対してフォーマット変換処理を実行する。変換処理部 22 の詳細について、以下で説明する。

【 0045 】

以上の処理は、画像形成装置 100 に備えられる図示しない制御部により制御される。

20

【 0046 】

[ 3 . 変換処理部 ]

図 2 は、変換処理部 (文書画像処理装置) 22 の構成を示すブロック図である。変換処理部 22 は、行解析処理部 (構造解析部) 31 と、行ブロック解析処理部 (構造解析部) 32 と、レイアウト解析処理部 (構造解析部) 33 と、変換可否判定処理部 (変換判定部) 34、再配置処理部 (参照リスト生成部) 35 と、を備える。以下では、文書を構成する個々の文字、図、表を要素 (要素画像) と称する。図はグラフを含むものとする。

【 0047 】

行解析処理部 31 は、文書画像から各要素を抽出して、文字 (文字画像) から構成される文字列の行 (文字列行) と、図 (図画像) 又は表 (表画像) から成る行 (図表行) に分類する。更に、文書の横書き、縦書きといった記述方向を示す文書第 1 方向を解析する。

30

【 0048 】

行ブロック解析処理部 32 は、行解析処理部 31 で抽出された行を、少なくとも 1 つ以上有する行ブロックに統合する処理を行う。

【 0049 】

レイアウト解析処理部 33 は、行ブロック解析処理部 32 で分類された行ブロック同士の位置関係から、段組構成を解析して文書全体の行の順序付けを行い、行の前後関係から改行位置を検出することで、文書を 1 つ以上の段落に分類し、段落毎に行の情報を格納した文書構造ツリーを生成する。詳細は後述する。文書構造ツリーの各段落は、文字列の行の順序の情報と、図表の順序の情報をそれぞれ分けて格納することで、図表の配置を段落内で修正できるようにする。

40

【 0050 】

変換可否判定処理部 34 は、行解析処理部 31 からレイアウト解析処理部 33 までの処理にて得られた情報から、文書画像をリフロー型に変換するか否かを判定する。変換可否判定処理部 34 は、リフロー型に変換しないと判定した場合は、圧縮処理部 21 に判定信号を出力する。圧縮処理部 21 では、上記判定信号を受信すると、RGB の画像データを例えば、JPEG ファイルフォーマットに変換して出力する。すなわち、送信装置 4 の送信先の表示装置にてフィックス型で表示されるように送信装置 4 の送信先の表示装置での表示に適したフォーマットに変換して出力する。なお、圧縮処理部 21 は、JPEG ファイルフォーマット以外に、例えば、PNG ファイルフォーマットあるいは GIF ファイル

50

フォーマットに変換してもよい。

【 0 0 5 1 】

再配置処理部 3 5 は、変換可否判定処理部 3 4 にてリフロー型に変換すると判定された文書画像に対して、レイアウト解析処理部 3 3 で生成された文書構造ツリーに従って、文書画像の文字、図、表の各要素を順序通り参照するための命令と、段落の開始及び終了を宣言するための命令を列記した参照リストとを生成する。参照リストのフォーマットは特に固定されておらず、例えば HTML 等のマークアップ言語で記述した文書の形式として生成したものをファイル出力してもよい。

【 0 0 5 2 】

なお、本実施形態では、変換処理部 2 2 の処理について、画像入力装置 2 が読み取った画像データ（文書画像）を処理する場合を用いて説明するが、ネットワークを介して受信した、あるいは、USBメモリ等のメモリに格納されているデータ（PDFファイル又は構造化された文書ファイル（doc、txt、odf、xls等））については、次のように処理を行う。受信した、あるいは、メモリに格納されているデータを、不図示のソフトウェア処理部において、文書画像ファイル（jpeg、tiff、bmp等）に変換し、文書画像ファイルに変換されたデータを変換処理部 2 2 に入力する。受信した、あるいは、メモリに格納されたデータが文書画像ファイルである場合は、ソフトウェア処理部において、何ら処理は行わない。

【 0 0 5 3 】

文書画像ファイルに変換されたデータについては、変換可否判定処理部 3 4 において、リフロー型に変換するか否かの判定を行う。リフロー型に変換しないと判定された場合、圧縮処理部 2 1 に文書画像ファイルに変換されたデータが出力される。圧縮処理部 2 1 は、文書画像ファイルに変換されたデータが、JPEGファイルフォーマットである場合は、何も処理を行わずにそのまま出力する。文書画像ファイルに変換されたデータが、tiff又はbmpデータである場合は、文書画像ファイルに変換されたデータを、例えば、JPEGファイルフォーマットに変換して出力する。つまり、送信装置 4 の送信先の表示装置での表示に適したフォーマットに変換して出力する。

【 0 0 5 4 】

以下、変換処理部 2 2 の各処理部について詳述する。

【 0 0 5 5 】

[ 4 . 行解析処理部 ]

< 4 - 1 . 行解析処理部の構成 >

行解析処理部 3 1 は、文書画像から各要素画像を抽出し、文字列行と図表行とに分類する。なおグラフは図表行に含まれるものとする。行解析処理部 3 1 は、文字列抽出処理部 3 1 a 及び図表抽出処理部 3 1 b を備えて構成される。行解析処理部 3 1 は、更に、文書の横書き又は縦書きといった記述方向を示す文書第 1 方向を解析する。

【 0 0 5 6 】

< 4 - 2 . 文字列抽出処理部 >

文字列抽出処理部 3 1 a は、文書画像から個々の文字を抽出（検出して切り出す）すると共に、文字が複数並べられて構成される文字列を抽出する。文字及び文字列の抽出は、次の方法により行う。文書画像から文字領域の画素を抽出し、その中から 1 つの文字を構成していると思われる画素の集合を囲む最小外接矩形を文字構成要素として抽出する。更に、上下左右の各方向における近隣の各文字構成要素の矩形同士の距離から文字列として連続する文字構成要素の関係にあるかを判定し、その連続する矩形の連続数から、文字列領域を特定する。このとき、左右方向における連続数が上下方向における連続数を上回る場合は横書きの文字列領域として、上下方向における連続数が左右方向における連続数を上回る場合は縦書きの文字列領域として、文字列の持つ方向（文字列の方向、記述方向）を同時に取得する。

【 0 0 5 7 】

なお、文字及び文字列の抽出方法は、上記に記載の方法に限らず他の方法を用いること

10

20

30

40

50

ができる。例えば、光学式文字読取装置 (Optical Character Recognition; 以下OCR) で個々の文字及び文字列を抽出してもよい。

【0058】

< 4 - 3 . 図表抽出処理部 >

図表抽出処理部 3 1 b は、文書画像から図 (図領域) 及び表 (表領域) を抽出する。

【0059】

図領域の抽出は次の方法により行う。文書画像の所定領域毎に画素値の出現頻度 (即ち、ヒストグラム) を求めた場合に、図領域の一つである写真領域上の各画素では濃度変化が広範囲に及びヒストグラムが得られることを利用して、ヒストグラムのエントロピー (平均情報量) を算出する。このようにエントロピーが高い領域を抽出することで精度よく写真領域を抽出することが可能となる。図領域の抽出方法は、上記に記載の方法に限らず他の方法を用いてもよい。

10

【0060】

また、表領域の抽出は、次の方法により行う。文書画像からラインとなる可能性のある候補画素を抽出し、前記候補画素が水平方向もしくは垂直方向に所定画素数以上連続する場合に前記連続する候補画素の集合をラインとして抽出し、前記抽出された水平方向及び垂直方向のラインの位置関係から、各ラインが表を構成する罫線であるか単一のラインであるかを判定し、同一の表を構成するラインの集合について、それら全てを囲む最小外接矩形を表領域として抽出する。このように抽出することで、精度よく表領域を検出することが可能である。なお、表領域の抽出方法は、上記に記載の方法に限らず他の方法を用いてもよい。

20

【0061】

グラフは図の一種であるが、以下のようにグラフ領域を抽出することができる。円グラフの様にその形状だけでグラフが構成されているグラフや、棒グラフ等の様にグラフの軸の線と繋がっており、ベタや網掛け等のあるグラフの場合は、前述したエントロピー値を用いてグラフ領域を抽出することが可能である。また、折れ線グラフのような線ベースのグラフでグラフの軸から離れているようなグラフの場合は、次のように、抽出可能である。図32の(a)~(c)に示すように、前記エントロピー値と並行して表を構成する罫線ほどの水平又は垂直ラインではないが、単一の直線とは異なり、矩形やL字型やU字型のような水平又は垂直ラインの繋がりのある領域に対して、その領域の最外郭を矩形化处理し、その矩形の大きさが、予め定められた閾値以上の面積を持つ矩形であり、かつ、その矩形領域内にエントロピー値がある程度高い領域が存在するかを判定することでグラフ領域を抽出することが可能である。なお、グラフ領域の抽出方法は、上記に記載の方法に限らず他の方法を用いてもよい。

30

【0062】

なお、文字列抽出処理部 3 1 a で抽出した文字が、図表抽出処理部 3 1 b で抽出した図又は表として抽出した範囲と重複する場合、該抽出した文字をキャンセルする。特に、抽出された表には文字が含まれる可能性が高いが、表のサイズを表示領域の幅に合わせるためには表を構成する各列の幅を調整する必要がある。結果として、調整後の列幅に合わせて、表内の文字列は折り返し表示され、かえって可読性を低下させる原因となる。そのため、本実施の形態では、表として抽出された領域については、文字も含めたまま図表として抽出して表示する。

40

【0063】

< 4 - 4 . 行IDの設定 >

行解析処理部 3 1 は、以上のようにして抽出された文字列行及び図表行に対して、各行を識別する重複しない識別記号として行ID (Identification) を設定する。1つの行IDについて、その行IDを有する行が2つ以上存在しなければ、必ずしも文書の順序に従って行IDを割り振る必要は無い。行IDの設定方法として、行IDが「0」の場合を存在しない行である無効行とし、例えば、各ページの文書画像において、読み取った原稿の左上を原点 (0, 0) とし、原点に対して、右方向をX座標、下方向をY座標となる座標

50

系を採用し、行の範囲を表わす最も左上の Y 座標が小さい順に行 ID を連番で割り振る方法が挙げられる。なお、Y 座標が同じ行同士は X 座標が小さい方の行を割り振りにおいて優先する。この方法を用いる場合、段組構成により必ずしも文書の読み順序通りに行 ID が割り振られる訳ではないが、横書き文書であれば行が上にあるほど順序が先である可能性が高いため、比較的文書の順序を反映した行 ID の割り振り方になると言える。行 ID の設定方法はこれに限らず、自由に選択することができる。

#### 【 0 0 6 4 】

##### < 4 - 5 . 文書第 1 方向及び文書第 2 方向の決定 >

さらに、行解析処理部 3 1 は、1 ページの文書画像における全ての行について、文字列行か図表行かの分類が終わると、文字列の方向から、文書全体の方向を示す文書第 1 方向を決定する。文書第 1 方向は、横書きのとき水平となり、縦書きのとき垂直となる。文書第 1 方向は、取得した全ての文字列の持つ方向を分類し、その比率により決定する。文書第 1 方向を決定するための比率の算出方法の簡単な例として、単純に横書きもしくは縦書きの文字列の数をカウントして、その数の比率を算出する方法が挙げられる。この方法の場合、例えば横書きの行数と縦書きの行数を比較して、多い方の方向を文書第 1 方向として決定（設定）する。文書第 1 方向の決定は上記の方法に限らず様々な方法を採用することができる。

10

#### 【 0 0 6 5 】

ここで、算出した比率が所定閾値（例えば、0 . 7）以下である場合、文書には縦書きの行と横書きの行とが無視できない比率で混在しており文書全体の方向を一意に判別できないとして、行ブロック解析処理部 3 2 及びレイアウト解析処理部 3 3 での処理を行わず、変換可否判定処理部 3 4 にて入力文書画像をリフロー型に変換しないと判定する。

20

#### 【 0 0 6 6 】

さらに、上記の方法によって文書第 1 方向を決定すると、文書第 1 方向に直交する方向として文書第 2 方向を決定（設定）する。すなわち、文書第 1 方向が水平（横書き）の場合、文書第 2 方向は垂直、文書第 1 方向が垂直（縦書き）の場合、文書第 2 方向は水平となる。

#### 【 0 0 6 7 】

##### < 4 - 6 . 記号（約物）の統合処理 >

個々の文字の切り出しにおいて、以下に示すような記号（約物）の統合処理を追加することができる。記号には、例えば行頭に来ることが禁止とされる（行頭禁則）ものや、行末に来ることが禁止とされる（行末禁則）ものがあり、図 3 に示すような文字がその一部として挙げられる。切り出した個々の文字を表示した際、行の折り返しによりこれらのルールが守られず可読性が低下する場合がある。そこで、行頭禁則の記号については、1 つ前の文字と統合し、行末禁則の記号については、1 つ後の文字と統合することで、単独で行頭もしくは行末に来ることがなくなる。

30

#### 【 0 0 6 8 】

各文字が、前述したルールを持つ記号であるかの判定方法は公知の方法を使用することができる。例えば、OCR 処理を利用して文字種を照合してもよいし、文字の大きさや、文字を構成する画素の特徴から判別してもよい。例えば句読点の場合、図 4 のように横書きである場合に、行の下半分のみで構成され、また行の高さ（矢印で示された範囲）に比べて半分程度の幅を持つ場合、その文字が句読点である可能性が高いとして、1 つ前の文字と統合してもよい。

40

#### 【 0 0 6 9 】

図 4 では横書きの例を示したが、縦書きの場合も同様に統合を行うことができる。半角英小文字と区別するため、行を構成する他の文字の高さや幅の傾向から和文、英文の判定を加え、和文の場合のみ句読点と判定するようにする等の処理を追加してもよい。例えば、和文ではひらがな、カタカナ及び漢字等の全角文字が文章の大半を占めており、行の上半分もしくは下半分のみで構成される文字が少なくなる傾向がある。また、半角文字に比べて、全角文字では行の高さに対して文字の横幅が半分より大きい文字の種類が多い。従

50

って、(1)行を上下に分割する水平方向の直線をまたぎ、(2)文字の横幅が行の高さに所定係数(例えば0.6)を乗算した値以上である、文字数をカウントし、行を構成する文字数に対して前記(1)及び(2)を満たす文字数の割合が所定閾値(例えば0.5)以上である場合に、その行が和文であるとして判定する処理を適用することができる。和文、英文の判定方法はこの方法に限らず、他の方法により判定してもよい。また、縦書きの文書である場合は自動的に和文とみなしてもよい。

#### 【0070】

ここで、句読点と、「ァ」等小さい和字との区別がつかない可能性もあるが、これらの小さい和字も行頭禁則であるため句読点と同様に統合しても問題無い。そのため、厳密に句読点専用の処理とする必要はない。

#### 【0071】

##### <4-7. 処理例>

行解析処理部31が実行する処理の具体例として、図5に示す構造の文書画像(1ページ)に対して行解析処理を適用する場合について説明する。行解析処理部31は、図6に示すように、その行の要素を全て含んだ最小サイズの外接矩形の範囲を各行の領域として分離し、それぞれの行に、行IDを、外接矩形の左上の垂直座標(Y座標)位置の順で割り当てる。図6に示す文書画像では、行IDが105の行が図表の行であることを除いては、残りの行はいずれも横書きの文字列の行であり、縦書きの文字列の行は1つも含まれていない。そのため、行解析処理部31は、この文書画像における文書第1方向は水平方向であると決定する。

#### 【0072】

##### [5. 行ブロック解析処理部]

##### <5-1. 行ブロック解析処理部の処理>

行ブロック解析処理部32は、行解析処理部31で分類された文字列行を、少なくとも1つ以上の文字列行から成る文字列の行ブロックに統合し、重複しない行ブロックIDを持つ新規行ブロックとして記憶部5に記憶(登録)する。行ブロック解析処理部32は、図表行については、単一行で1つの行ブロックを構成するものとし、それぞれ重複しない行ブロックIDを持つ新規行ブロックとして登録する。

#### 【0073】

行ブロック解析処理部32による文字列の行ブロック統合処理について以下で詳細に説明する。初めに、行解析処理部31で分類された行のうち文字列行のグループから、注目行L1を選択する。続いて、注目行L1に関して、前方及び後方(定義は、後述の<5-5>章を参照)に連続する文字列行を探索する。具体的には、文字列行のグループの、注目行L1とは異なる文字列行から、注目行L1の前方もしくは後方に連続する文字列行を最大1つずつ選択する。注目行L1の連続行の候補となる文字列行は注目行L1を除く全ての文字列行であり、連続行の候補となる条件については後述する。

#### 【0074】

全ての文字列行について、前方及び後方に連続する文字列行を選択すると、連続する文字列行同士の繋がりから、前方及び後方の両方において連続する文字列行がなくなるまで1つの行ブロックとして分類、統合し、未割り当ての行ブロックIDを持つ新規行ブロックとして記憶部5に登録する。全ての文字列行がいずれかの行ブロックに登録されるまで処理を繰り返し、全ての文字列行についての登録が完了すると、行ブロック解析処理部32は処理を終了する。

#### 【0075】

##### <5-2. 連続行の候補の判定>

ここで、注目行L1とは別に選択された文字列行L2が、注目行L1の連続行の候補であるか否かを判定する方法について説明する。行L2が行L1の連続行の候補である条件として、少なくとも下記2つの条件を満たすものとする。

条件1：一方の行の先頭から末尾までの範囲において、もう一方の行の先頭もしくは末尾のうち少なくとも一方が存在する。

10

20

30

40

50

条件 2 : 2 つの行の行間変位量  $l i n e s p a c e ( L 1 , L 2 )$  が下記の式 ( 5 ) を満たす。

$$T H \_ M I N \_ L S \leq l i n e s p a c e ( L 1 , L 2 ) \leq T H \_ M A X \_ L S \quad \dots \quad ( 5 )$$

(  $T H \_ M I N \_ L S$ 、 $T H \_ M A X \_ L S$  は、予め設定される閾値 )

なお、条件 1 は、異なる段に属する行を連続行の候補として判定しないために用いる。条件 2 は、行間が広過ぎる又は狭過ぎる行を連続行の候補として判定しないために用いる。

#### 【 0 0 7 6 】

図 7 の ( a ) は、2 段組の横書きの文書の例、図 7 の ( b ) は、2 段組の縦書きの文書の例である。条件 1 を満たすために、行  $L 1$  と行  $L 2$  とは、文書第 1 方向で一部もしくは全部が重複している必要がある。例えば、図 7 の例の場合、行 a と行 b、行 c と行 d、行 e と行 f、行 g と行 h は、条件 1 を満たすため、これらの組合せは互いに連続行の候補となる。しかし、行 a と行 d、行 e と行 h 等の組合せでは条件 1 を満たさないため、これらの組合せは互いに連続行の候補とならない。

#### 【 0 0 7 7 】

条件 2 で示す行  $L 1$  と行  $L 2$  との行間変位量  $l i n e s p a c e ( L 1 , L 2 )$  は、図 8 に示すように、横書きであれば行  $L 1$  と行  $L 2$  とのうち下側にある方の行の上端座標と、もう一方の行の下端座標との差分値 ( 図 8 の ( a ) ~ ( c ) 参照)、縦書きであれば行  $L 1$  と行  $L 2$  とのうち左にある方の行の右端座標ともう一方の行の左端座標との差分値 ( 図 8 の ( d ) ~ ( f ) 参照) である。行  $L 1$  と行  $L 2$  とが重複しないとき、行間変位量  $l i n e s p a c e ( L 1 , L 2 )$  は 2 つの行の行間距離を示す。また、 $T H \_ M I N \_ L S$  及び  $T H \_ M A X \_ L S$  は、連続行同士の行間変位量として許容される差分値の最小値及び最大値を示す所定係数である。例えば、行  $L 1$  の文字サイズに所定係数  $r 1$  ( 例えば  $r 1 = 0 . 1$  ) を乗算したものを閾値  $T H \_ M I N \_ L S$  と設定し、所定係数  $r 2$  (  $r 2$  は正の数とする、例えば  $r 2 = 1 . 5$  ) を乗算したものを閾値  $T H \_ M A X \_ L S$  として設定する。

#### 【 0 0 7 8 】

閾値  $T H \_ M I N \_ L S$  及び閾値  $T H \_ M A X \_ L S$  は、他の方法により設定されてもよく、例えば行  $L 1$  と行  $L 2$  の文字サイズの平均値に所定係数を乗算したのものとしてもよい。また閾値  $T H \_ M I N \_ L S$  を正值に設定することで、重複のある 2 つの行同士を連続行として認めないようにすることができる。逆に閾値  $T H \_ M I N \_ L S$  を負値に設定することで、図 8 の ( c ) 及び ( f ) のように、行  $L 1$  と行  $L 2$  とが多少重複する場合も許容することができる。

#### 【 0 0 7 9 】

< 5 - 3 . 条件の強化 : インデントの範囲指定 >

また、連続行の候補を判定する条件を強化するために、上記条件 1 , 2 に加えて、別の条件を設定してもよい。例えば、次式 ( 6 ) を満たすことを条件として追加することができる。

$$i n d e n t ( L 1 , L 2 ) \leq T H \_ I N D E N T \quad \dots \quad ( 6 )$$

ここで、 $i n d e n t ( L 1 , L 2 )$  は行  $L 1$  の開始位置の文書第 1 方向成分と行  $L 2$  の開始位置の文書第 1 方向成分の差の大きさであり、すなわちインデントの大きさを意味する。また、閾値  $T H \_ I N D E N T$  は行の先頭のインデントとして許容される距離を示す所定係数である。閾値  $T H \_ I N D E N T$  は、例えば行  $L 1$  の文字サイズに所定係数 (  $r$  は正の数とする、例えば  $r = 1 . 5$  ) を乗算した値を与え、文字以内のインデントを許容することができる。閾値  $T H \_ I N D E N T$  は他の方法により設定してもよく、例えば行  $L 1$  と行  $L 2$  との文字サイズの平均値に所定係数  $r$  を乗算したのものとしてもよい。

#### 【 0 0 8 0 】

< 5 - 4 . 条件の強化 : 行終了位置の差異の許容範囲指定 >

連続行の候補を判定する条件を強化する他の条件として、例えば次式 ( 7 ) を満たすこ

10

20

30

40

50

とを条件として追加することで、行の終了位置がある程度近い行同士を連続行の候補とすることができる。

$|L1MAX1 - L2MAX1| \leq TH\_DIFF\_ENDPOS \cdot \cdot \cdot (7)$

ここで、 $L1MAX1$ は行 $L1$ の文書第1方向成分の最大値、 $L2MAX1$ は行 $L2$ の文書第1方向成分の最大値である。例えば、文書第1方向が水平方向（横書き）である場合、 $L1MAX1$ 及び $L2MAX1$ は、行 $L1$ 及び行 $L2$ の右端のX座標を指す。また、閾値 $TH\_DIFF\_ENDPOS$ は行の終了位置の差として許容される距離を示す所定係数である。例えば行 $L1$ の文字サイズの平均値に所定係数（ $\leq$ は正の数とする、例えば $= 0.5$ ）を乗算したものを閾値 $TH\_DIFF\_ENDPOS$ とすることで、文字以内のインデントを許容することになる。

10

#### 【0081】

< 5 - 5 . 前方又は後方の連続行の選択 >

行 $L1$ の連続行の候補として抽出された行から、行 $L1$ の前方で最も近い位置にある行、及び、後方で最も近い位置にある行を、それぞれ最大1つずつ選択する。なお、文書第1方向が水平方向（横書き）である場合、行 $L1$ より上にある行を前方の行、行 $L1$ より下にある行を後方の行とし、文書第1方向が垂直方向（縦書き）である場合、行 $L1$ より右にある行を前方の行、行 $L1$ より左にある行を後方の行とする。また、行の近さを表わす値として、例えば、前述の行間変位量 $linespace(L1, L2)$ を使用し、 $linespace(L1, L2)$ が小さい程、行が近いとみなすことができる。なお前方、後方とも、連続行は最大で1つずつであり、必ずしも連続行が存在する必要はない。

20

#### 【0082】

< 5 - 6 . 行ブロックへの分類及び統合 >

行ブロック解析処理部32は、以上のようにして、全ての文字列行について前方及び後方の連続行を選択すると、行ブロックへの分類、統合を行う。但し、複数の行から連続行として選択されるケースもあり得るため、相互に連続行であるとされていない行のペアについては、その間の連続関係を事前に解消しておく。例えば、文書画像が図9の(a)である場合、前方の連続行として行 $L3$ を選択する行は、行 $L4$ と行 $L5$ との2つ存在するが、行 $L5$ は行 $L3$ の後方の連続行として選択されていない。そのため、行 $L3$ と行 $L5$ との間の連続関係は解消される。同様にして行 $L4$ と行 $L6$ との間の連続関係も解消される。このことにより、図9の(a)に示すような例では、行 $L3$ 及び行 $L4$ において注目行 $L1$ からの連続関係が断たれるため、注目行 $L1$ と同一の行ブロックとして分類できなくなるケースも起こり得る。しかし、図9の(b)に示すように複数の行ブロックとして分類することができ、後段のレイアウト解析処理部33における段組解析処理部37で、同一の段組、及びその段組を構成する同一の段（カラム）として統合できるため、この時点でブロックが分かれてしまっても、問題とはならない。

30

#### 【0083】

行ブロックへの分類及び統合処理は、次のように行う。まず、行ブロックとして分類されていない文字列行のうち任意の行 $L1$ （注目行 $L1$ ）について、まず、行 $L1$ を新規の行ブロックとして設定する。続いて、行 $L1$ から前後の連続行をたどり、行ブロックの範囲を拡大する。前方及び後方とも、連続行が無くなると、行ブロックの拡大を終了し、その行ブロックに含まれる先頭の行から順に行IDを取得する。また、行ブロックの情報として、行ブロックに含まれる全ての行に外接する最小矩形の左上座標、幅及び高さ、並びに含まれる行数を取得する。以上のようにして得られた行の順序と各種情報とを持つ行ブロックを、既に登録済みの行ブロックと重複しないIDを持つ新規の行ブロックとして登録を行い、またその行ブロックに含まれる各行の所属行ブロックIDを更新する。行ブロック解析処理部32は、このようにして行われる行ブロックへの分類及び統合処理を、全ての文字列行がいずれかの行ブロックに分類されるまで繰り返す。

40

#### 【0084】

< 5 - 7 . 同一の行ブロックに分類できる（連続行の候補とできる）行の条件 >

なお、文書第1方向の文字列行は文書第1方向の文字列行とのみ、文書第2方向の文字

50

列行は文書第2方向の文字列行とのみ、行ブロックを構成する。すなわち、1つの行ブロックに、文書第1方向の文字列行と文書第2方向の文字列行とが混在することは無い。従って、注目行L1の連続行の候補を探索する際、注目行L1の文字列方向と異なる方向の文字列行は連続行の候補としない。

#### 【0085】

##### < 5 - 8 . 処理例 >

行ブロック解析処理部32が実行する行ブロック解析処理を、具体例を用いて説明する。既に示した図6のように文書画像から検出された複数の行に対して行ブロック解析処理を適用すると、文書画像は、図10のように行ブロックとして分類される。図10に示す例では、行ブロックB3は、章の見出しの行であり、行ブロックB4に比べて文字が大きい。このように文字のサイズが大きく異なる2つの行同士を連続行の候補として選択しないような、連続行の候補の判定の条件を追加することも有効な手段である。

10

#### 【0086】

##### [ 6 . レイアウト解析処理部 ]

##### < 6 - 1 . レイアウト解析処理部の構成 >

図11は、レイアウト解析処理部33の詳細構成を示すブロック図である。レイアウト解析処理部33は、前段の行ブロック解析処理部32で分類された行ブロック同士の上下左右の位置関係から、行ブロック構成を解析し、文書画像中の文章(本文)の読み順を推定する処理を行うものである。レイアウト解析処理部33は、段組解析処理部37、行順序付け処理部38、段落解析処理部39を備えて構成される。

20

#### 【0087】

##### < 6 - 2 . 段組解析処理部 >

段組解析処理部37は、複数の行ブロックの上下及び左右の位置関係から、段組及び段組を構成する各段(カラム)を分類する段組解析処理を実行する。文書は文書第2方向に段組が配置され、各段組構成内で文書第1方向にカラムが配置されているものとして、ページ内の行ブロックの集合を、適切に境界線を設定して行ブロックをまたぐことなく分割して初期段組とする。そして、同一の初期段組に含まれる行ブロックの集合を、適切に境界線を設定して行ブロックをまたぐことなく分割して、該初期段組を構成する初期カラムとする。

#### 【0088】

境界線の設定方法は特に指定はなく、最も簡単な例として、初期段組の分類には文書第2方向と平行な直線を使用し、初期カラムの分類には文書第1方向と平行な直線を使用することが挙げられる。例えば、図12の(a)のように横書きの文書画像から行ブロックの構造が解析された場合、図12の(b)のように行ブロックを初期段組に分類され、さらに初期段組は図12の(c)のようにそれぞれ初期カラムとして分類される。なお、図12の(b)及び(c)では、段組間の境界線は実線で、カラム間の境界線は一点鎖線で示されている。

30

#### 【0089】

##### < 6 - 3 . 段組の分割禁止(同一段組として許容される行ブロック間距離の算出) >

本来は同一段組であるが、偶然、行ブロックを分割することができるために複数の段組に分かれてしまうようなケースもまれに存在する。こうしたケースに対応するため、例えば連続する2つの行ブロック間の距離を算出し、その距離が所定値(例えば行ブロックの平均行間距離の2倍)以下の2つのブロック間には境界線を引くことを禁止する条件を追加することができる。

40

#### 【0090】

図13は、行ブロックB10, B11, B12を左側のカラム、行ブロックB20, B21を右側のカラムとした2段組構成の例を示す。行ブロックB10と行ブロックB11との間、行ブロックB20と行ブロックB21との間が空いているため、行ブロックB10と行ブロックB20から成る2段組構成、及び、行ブロックB11とB12と行ブロックB21とから成る2段組構成として分割してしまう恐れもある。しかし、行ブロックB

50



20の平均行間距離(20)に対して、行ブロックB20と行ブロックB21とのブロック間距離(30)が所定値(20×2=40)以下であるとして、行ブロックB20と行ブロックB21との間に境界線を引くことを禁止することで、これらの行ブロックが2つの異なる段組に分かれることを防ぐことができる。

【0091】

<6-4. 位置関係以外の情報の活用>

また、行ブロックの位置関係に加えて、行ブロックが持つ各種情報を利用して、段組及びカラムの分類(すなわち境界線の設定)を行うことができる。行ブロックが持つ各種情報の例として、行の長さや主要な文字のサイズ等が挙げられる。隣り合う行ブロック同士でこれらの情報が大きく異なる場合は、同一の段組に分類することを避けるようにすることができる。逆に位置が大きく離れた行ブロック同士でも、例えば同じカラム境界線を共有することができ、かつ類似する情報を持つ場合、同一段組として分類してもよい。

10

【0092】

<6-5. 行ブロックが0個もしくは1個しかない場合>

なお、段組解析処理部37に入力されたページ画像が、ただ1つの行ブロックを持つ場合、そのページ画像は1段構成の文書であるとして、境界線の設定は行わない。また、該ページ画像が、1つも行ブロックを持たない場合(すなわち白紙ページの場合)も境界線の設定は行わない。

【0093】

また、文書画像に文書第1方向の文字列行と文書第2方向の文字列行とが混在する場合は、文書第2方向の文字列行の行ブロックを図表行の行ブロックに置き換える。このことにより、文書第1方向に記述された文章の最中に、文書第2方向に記述された文章が混在することを防ぐことができる。

20

【0094】

<6-6. 処理例>

段組解析処理部37が実行する処理の具体例として、例えば既に示した図10に示す文書画像から検出された複数の行ブロックに対して段組解析処理を適用する場合について説明する。段組解析処理部37は、図10に示す文書画像を、図14に示す段組及びカラム(淡いグレー地)に分類する。カラムC1及びカラムC2は、それぞれ1段構成の段組G1及び段組G2を成し、カラムC3及びカラムC4は2段組構成の段組G3における左右のカラムを成している。なお、図14では、行(文字列行及び図表行)を直線、行ブロックを点線、カラムを一点鎖線で囲んでいる。

30

【0095】

<6-7. 行順序付け処理部>

行順序付け処理部38は、段組、カラム、行ブロック、及び行の位置関係から文書全体における行の順序を解析し、行順序リストを生成する処理を、以下のルール(1)~(9)に従って行う。

(1) 同じ行ブロックに属する行同士については、横書き文書であれば上から下、縦書き文書であれば右から左の順に優先順位を設定する。ここでは、上記のように既に、行ブロックに分類する際に、その行ブロックに含まれる行についての順序の情報も取得しているため、この情報を利用する。

40

(2) 同じカラムに属する行ブロック同士については、横書き文書であれば上から下、縦書き文書であれば右から左の順に優先順位を設定する。

(3) 連続する2つの行ブロック間では、優先順位の高い方の行ブロックの末尾の行の次に、優先順位の低い方の行ブロックの先頭の行が優先されるように設定する。

(4) 同じ段組に属するカラム同士については、横書き文書であれば左から右、縦書き文書であれば上から下の順に優先順位を設定する。

(5) 連続する2つのカラム間では、優先順位の高い方のカラムの末尾の行ブロックの次に、優先順位の低い方のカラムの先頭の行ブロックが優先されるように設定する。

(6) 同じページに属する段組については、横書き文書であれば上から下、縦書き文書で

50

あれば右から左の順に優先順位を設定する。

(7) 連続する2つの段組間では、優先順位の高い方の段組の末尾のカラムの次に、優先順位の低い方の段組の先頭のカラムが優先されるように設定する。

(8) 同じ文書画像ファイルに属するページ同士については、ページ番号の小さい順に優先順位が高くなるよう設定する。

(9) 連続する2つのページ間では、優先順位の高いページの末尾の段組の次に、優先順位の低い方のページの先頭の段組が優先されるように設定する。

#### 【0096】

行順序付け処理部38は、上記のルール(1)~(9)に従って、ページの順序、段組の順序、カラムの順序、行ブロックの順序を決定し、それらにより行の順序付けを行う。順序付けされた行は、各行が属する行ブロック、カラム、段組及びページの順序を示す番号を保有すると共に、先頭から順に各行の行IDを行順序リストに格納する。

行順序リストは、下記の規定(a)~(c)に従う形式であれば特に構造は問わない。

(a) 上記順序付けルールに従って決定された順序通りに行を呼び出すことができる。

(b) 呼び出した行について、その座標情報や種類(文字列行か図表行か)等の各種情報を参照することができる。

(c) 呼び出した行について、段落情報(後述)を格納することができる。

#### 【0097】

##### <6-8. 処理例>

行順序付け処理部38が、上記の規定(a)~(c)に従って、図14の構成の文書画像について行順序リストを生成した例を、図15に示す。行順序リストは、決定された順序の先頭から順に、行のIDと、行の情報として、所属する行ブロックID、行の種別(文字列行であるか図表行であるか)、及び行の範囲を示す外接矩形の左上座標及び右下座標、の情報を格納し、さらに行毎に段落情報を格納している。なお、図15の例では、改行が発生するときに、その行から新たな段落が始まるとして、段落情報を改行の有無を有り(Yes)か無し(No)かの2通りで示しており、事前に「No」で初期化している。図16のように行ブロックの情報やカラムの情報、段組の情報を別途作成し、相互参照により各行及び各行ブロックが所属するカラム、段組、ページを参照できるようにしておくことで、冗長の少ない行順序リストを構成することができる。もちろん、行順序リスト単独で各行に関する情報を全て抽出できるようにしてもよい。

#### 【0098】

##### <6-9. 段落解析処理部>

段落解析処理部39は、各行の前後の位置関係等の情報から、その行の位置で改行が発生しているかどうかを判定し、文書画像中の各行を1つ以上の段落に分類する処理を行う。具体的には、行順序リストから複数の行を参照して段落の切れ目、すなわち改行位置を判定し(改行判定処理)、段落毎に行の順序を記述した文書構造ツリーを生成する(文書構造ツリー生成処理)。

#### 【0099】

ところで、文書中の図表は、必ずしも段落の切れ目に配置されとは限らず、例えばページの端に挿入される場合が多く、それにより文章が図表を挟んで前後に分かれることがある。この順序のまま行を呼び出し、行を構成する要素(文字、図表)を挿入していくと、図表の挿入によって不自然に途切れた文章が出力されてしまう。そこで、本実施の形態では、段落毎に、文字列行と図表行とが混在した順序ではなく、文字列行の順序と図表行の順序をそれぞれ別に保有する文書構造ツリーを生成する。図17は、文書構造ツリーの構造を示す図である。それにより、文書画像を構成する文字列のみの順序を把握しながら、その段落に係る図表を、段落の先頭や末尾等にまとめて配置できるようにする。

#### 【0100】

##### <6-10. 改行判定処理>

図18は、段落解析処理部39における、改行判定処理の概要を示すイメージ図である。改行判定処理は、判定の対象となる注目行と、注目行より前に順序づけられるM個の行

と、注目行より後に順序づけられる $N$ 個の行と、の $M + N + 1$ 個の行によって判定される。なお、図18に示す例では、 $M = N = 2$ としている。なお、改行判定の対象となる行、及びその前後の行は、いずれも文字列行である。本実施の形態では、 $M + N + 1$ 個の行IDバッファ $L[0]$ 、 $L[1]$ 、 $\dots$ 、 $L[M + N]$ を記憶部5に備え、行順序リストで参照される $M + N + 1$ 個の行IDをそれぞれ格納することで、注目行と、注目行の前後の行との比較を行う。

#### 【0101】

以下、段落解析処理部39の処理内容について詳細に説明する。図19は、段落解析処理部39の処理手順を示すフローチャートである。段落解析処理の開始にあたり、事前に初期化を済ませておく。具体的には、リスト参照番号を $LNOW = 1$ とし、行IDバッファには全て無効行(0)を格納しておく。初期化が終わると、注目行にあたる行IDバッファ $L[M]$ に、行IDを選択して格納する(ステップS1、以下ではS1のように略す)。任意の行IDバッファ $L[k]$ ( $k = 0, 1, \dots, M + N$ )の選択方法は以下の(1)~(3)の通り行われる。

(1) 行順序リストにおいて、第 $LNOW$ 番目から順に、文字列行を探索する。

(2) 最初に見つかった文字列行の行IDを行IDバッファ $L[k]$ に格納し、そのときの行順序リストの位置(リスト番号)に1を加えた番号を新たなリスト参照番号 $LNOW$ として更新する。

(3) 文字列行が見つからないまま行順序リストの末尾まで探索が終了した場合、行IDバッファ $L[k]$ には無効行(0)を格納する。

#### 【0102】

行 $L[M]$ の更新後、行 $L[M]$ が有効行(ゼロでない行IDを持つ行)であるかどうかを判定し(S2)、有効行である場合(S2の判定がYES)、S3に移る。一方、 $L[M]$ が無効行である場合(S2の判定がNO)、入力した文書画像には文字列行が存在しない図表行のみの文書画像であるとして、文書構造ツリーの生成処理(後述)を実行する(S7)。

#### 【0103】

次に、注目行より後の行に当たる行IDバッファ $L[M + 1]$ 、 $\dots$ 、 $L[M + N]$ に、行IDを選択して格納する(S3)。各バッファにおける行IDの選択方法は上記と同様であるため省略する。続いて、注目行 $L[M]$ の改行判定を実行する(S4)。改行判定は、改行判定対象となる注目行 $L[M]$ と、注目行より前に位置する行 $L[0]$ 、 $\dots$ 、 $L[M - 1]$ 及び注目行より後に位置する行 $L[M + 1]$ 、 $\dots$ 、 $L[M + N]$ からなる複数の行を用いた公知の方法で行うことができる。簡単な例として、改行判定対象のインデントの有無を確認する方法がある。行 $L[M]$ の開始位置が、他の行に比べて文書第1方向に正値のずれが生じている場合に、行 $L[M]$ はインデントを持ち、行 $L[M]$ の位置で改行がなされているとみなすことができる。また、改行判定対象行 $L[M]$ の1つ前の行 $L[M - 1]$ が、他の行に比べて短い場合、行 $L[M]$ の位置で改行がなされているとみなすことができる。例えば、図18( $M = N = 2$ )のような横書きの文字列行が存在している場合、注目行 $L[2]$ の開始位置が、他の行に比べて右側(横書きの場合の正方向)にずれており、また1行前の行 $L[1]$ が他の行に比べて短いという特徴を持っており、これらの結果から、注目行 $L[2]$ は総合的に改行位置であるとして判定され易くなる。なお、文書編集者の好み等により段落の先頭行でインデントが付加されない場合もあり、また1つ前の段落の最終行が必ずしも短くなるとは限らないため注意する。

#### 【0104】

改行位置であるかどうかの判定の他の例としては次のものがある。注目行の文字サイズが周辺の行に比べて大きく異なるかどうかを判定することで、見出し行のように文字サイズが大きくなっている行、また補足コメント等のように逆に文字サイズが小さくなっている行等で改行位置と判定することができる。また、周辺の連続する2つの行の行間距離に比べて、注目行とその1つ前の行との行間距離が大きくなっている場合に、1つ前の行で段落が終了している可能性が高くなる。上記挙げられた条件を例として、様々な条件を複

合的に判定して、注目行における改行の有無を設定するのが好ましい。

【 0 1 0 5 】

行 L [ M ] の改行判定の結果は、行順序リストの L [ M ] に該当する行情報に段落情報として反映させる。段落情報は段落の切れ目が判るものであれば何でもよく、最も簡単な例として、改行の有無を Yes あるいは No の 2 通りで示すだけでもよい。

【 0 1 0 6 】

改行判定対象行 L [ M ] の改行判定が終了すると、注目行の次の行 L [ M + 1 ] が有効行であるかどうかを判定し ( S 5 )、L [ M + 1 ] が有効行である場合 ( S 5 の判定が Y E S )、行 I D バッファの更新を行い、次の行についての改行判定を行う準備をする ( S 6 )。行 I D バッファの更新は、具体的には、図 2 0 に示すように、L [ 0 ] = L [ 1 ] , . . . , L [ M + N - 1 ] = L [ M + N ] としてバッファを 1 つずつずらすとともに、バッファ L [ M + N ] を新たに行順序リストから選択する。バッファ L [ M + N ] の選択方法は前述の方法と同様であるため省略する。バッファの更新後、S 4 に戻り、更新された注目行 L [ M ] について、改行判定を行う。これを、S 5 で判定が NO となるまで繰り返す。S 5 の判定が NO となると、全ての文字列行について改行判定が終了したことになります。文書構造ツリーの生成処理を実行する ( S 7 )。

【 0 1 0 7 】

< 6 - 1 1 . 文書構造ツリー生成処理 >

段落解析処理部 3 9 による文書構造ツリー生成処理は、段落情報を考慮した行順序リストに格納された段落情報に従って実行される。但し、文書構造ツリーは事前に、図 2 1 に示すような 1 つの空の段落 ( 第 0 段落 ) を持つ状態に初期化されているものとする。初期段落番号を P N O W = 0 として、行順序リストの先頭から順に行の情報を参照し、段落情報から該行が段落開始行と判定された場合 ( 段落情報が Y E S ) のみ、P N O W = P N O W + 1 として段落番号の更新を行い、また文書構造ツリーに空の段落 ( 第 P N O W 段落 ) を新たに追加する。そして、該行が文字列行である場合、文書構造ツリーの第 P N O W 段落を持つ文字列ツリーの末尾に該行の行 I D を追加する。一方、該行が図表行である場合、文書構造ツリーの第 P N O W 段落を持つ図表ツリーの末尾に該行の行 I D を追加する。これを行順序リストの各行について反復し、末尾まで探索が終了すると、文書構造ツリーの生成処理を終了する。なお、文字列行の改行位置から次の改行位置までが 1 つの段落となるので、図表行が段落の分類から漏れるということはない。そのため、ひと固まりの文字列行の後に図表行があり、その後新しい段落が始まる ( 改行が発生する ) 場合は、その図表行は 1 つ前の段落に含まれることになる。

【 0 1 0 8 】

< 6 - 1 2 . 処理例 >

段落解析処理部 3 9 が実行する処理の具体例として、図 6 の構成の文書画像 ( 図 1 5 に示す初期の行順序リストを持つ ) に段落解析処理部 3 9 での処理を適用する場合について説明する。見出しに当たる行 I D : 1 0 1、行 I D 1 0 4、行 I D : 1 2 9 の各行は、行の文字サイズや 1 つ前の行間距離等の条件から、改行位置と判定される。また行 I D : 1 0 6、行 I D : 1 1 2、行 I D : 1 1 9 及び行 I D : 1 3 1 の各行は、前後の複数の行に比べ、行の開始位置が文書第 1 方向において正方向にシフトしている、よってインデントが存在するとして、改行位置と判定される。行 I D : 1 0 2 の行も、見出し行の行 I D : 1 0 1 の次の行であることから、改行位置と判定することができる。従って、行 I D : 1 0 1、1 0 2、1 0 4、1 0 6、1 1 2、1 2 9、1 3 1、1 1 9 の各行が改行位置として設定され、行順序リストの段落情報は、図 2 2 のように更新される。

【 0 1 0 9 】

例えば、行順序リストが図 2 2 のように、段落情報として改行の有無 ( Y e s もしくは N o ) が表されている場合、改行がある行から新しい段落が開始するとみなせるため、改行がある ( Y e s ) 場合に該行が段落開始行として判定することができる。また、段落が開始してから、次の改行位置が見つかるまでの行を同一の段落の範囲とみなす。更新された行順序リストの段落情報に従い、図 6 の文書画像を段落毎に分類すると、図 2 4 に示す

ように分類することができる。図 24 において、段落 R7 は、左下の 3 行の文字列行（行 ID：131，133，135）から、右上の図表行（行 ID：105）及び 4 行の文字列行（行 ID：111，113，115，117）までを同一の段落の範囲としており、行 ID：135 と行 ID：111 との文字列行の間に改行位置は存在しないため、一続きの文章を構成することが可能となる。更に、この行順序リストに従って文書構造ツリーを生成すると、図 23 に示すような文書構造ツリーを得ることができる。

#### 【0110】

##### [7. 変換可否判定処理部]

##### <7-1. 変換可否判定処理部の処理>

変換可否判定処理部 34 は、行解析処理部 31 からレイアウト解析処理部 33 までの処理にて得られた情報（文書画像から抽出した文字列あるいは図又は表の特徴量）から、文書画像を次に示す条件によってリフロー型に変換するか否か、言い換えれば、文書画像に含まれる各要素を再構成するか否か、を判定する。

第 1 の条件：行解析処理部 31 により文書内に縦書きの行と横書きの行とが無視できない比率で混在しており文書全体の方向を一意に判別できないと判定された。

第 2 の条件：行解析処理部 31 の処理では図表行しか抽出されなかった。

第 3 の条件：行解析処理部 31 にて抽出された文字列の行数が、所定の閾値（THLn）（第 1 閾値）以下である。

第 4 の条件：行解析処理部 31 にて抽出された全ての文字列の行に対する、横書きの行の高さ（縦書きの場合は行の幅）が、所定の閾値（THCs）（第 2 閾値）以上である。

第 5 の条件：レイアウト解析処理部 33 にて順序付けられた横書き又は縦書きの行ブロックの位置が、一定の範囲（THrg）内には揃っていない。

#### 【0111】

変換可否判定処理部 34 は、入力文書画像が以上の 5 つの条件の何れかに当てはまる場合には、リフロー型に変換しないと判定する。そして、変換可否判定処理部 34 は、リフロー型に変換しないと判定した場合は、圧縮処理部 21 に判定信号を出力する。圧縮処理部 21 では、上記判定信号を受信すると、RGB の画像データを例えば、JPEG ファイルフォーマットに変換して出力する。すなわち、フィックス型で表示されるように画像全体をそのまま出力する。このように、本実施形態では、圧縮処理部 21 が、リフロー型に変換しないと判定された文書画像をフィックス型の表示が可能にフォーマット変換するフォーマット変換処理部として機能する。

#### 【0112】

また、変換可否判定処理部 34 は、入力文書画像が以上の 5 つの条件の何れかも当てはまらない場合には、リフロー型に変換すると判定する。リフロー型に変換すると判定された文書画像については、次の再配置処理部 35 にて、参照リストの生成が行われる。

#### 【0113】

##### <7-2. 判定条件の詳細>

上記 5 つの条件についてそれぞれ詳細に説明する。

#### 【0114】

まず、第 1 の条件について説明する。上記したように、行解析処理部 31 は、取得した全ての文字列の持つ方向を分類し、その比率により文書第 1 方向を決定する際に、例えば、単純に横書きもしくは縦書きの文字列の数をカウントする。そして、カウントした数の比率を算出することにより、算出した比率が所定閾値（例えば、0.7）以下である場合、文書画像には縦書きの行と横書きの行とが無視できない比率で混在していると判定する。そのため、文書全体の方向を一意に判別できず、また、レイアウト解析する際にも、縦書きと横書きの行をどう接続していけばよいかの判定が困難になる。よって、リフロー型に変換しないと判定する。

#### 【0115】

次に、第 2 の条件について説明する。行解析処理部 31 にて図表行しか抽出されなかった場合、図や表などはその中に文字があったとしても、レイアウトを崩さずにそのまま表

10

20

30

40

50

示しないと内容がわからなくなってしまう可能性がある。よって、リフロー型に変換しないと判定する。

【0116】

次に、第3の条件について説明する。行解析処理部31にて抽出した文字列の行数が例えば5行以下（ $THLn = 5$ ）であった場合には、わざわざリフロー型にして読むほどの行数でもないと考えられる。そのため、リフロー型に変換しないと判定する。この閾値（ $THLn$ ）をいくつにするかは、例えば、表示する画面のサイズと行の文字サイズに応じて1画面で表示できる行数分を閾値とすることが考えられる。

【0117】

次に、第4の条件について説明する。行解析処理部31にて抽出した全ての文字列の行に対し、横書きの行の高さ（縦書きの場合は幅）が例えば40画素以上（ $THcs = 40$ ）であった場合には、文字が十分大きいので、縮小表示しても十分可読性があると考えられる。よって、リフロー型に変換しないと判定する。この閾値（ $THcs$ ）をいくつにするかは、例えば、表示する画面のサイズと画像全体のサイズによる縮小率と、縮小した場合に読める最小のフォントサイズの画素数とから、元の画像での1文字の大きさの画素数を算出し、それを閾値とすることが考えられる。

【0118】

次に、第5の条件について説明する。レイアウト解析処理部33にて順序付けられた横書き又は縦書きの行ブロックの位置が、一定の範囲内に揃っていない場合には、整理されて段組みされたレイアウトでない可能性が高いため、行ブロックの接続が失敗し、誤った行ブロック同士をつないでしまう可能性が高い。そのため、リフロー型に変換しないと判定する。例えば、図31に示すように縦書きの行ブロックが複数存在し、また、図表も複数存在している文書画像について考える。この文書画像について、行ブロックのブロックサイズと位置情報からレイアウト解析処理部33で順序付けられた行ブロックの位置のずれを求める。ここで、一定の範囲内で揃っているかどうかの閾値（ $THrg$ ）として $THrg = 40$ とすると、どの行ブロックも閾値（ $THrg$ ）を超えており、一定範囲内に揃っていないということになる。このような場合、ブロック単位で上から順に行を接続していくと、縦書き1 縦書き2 縦書き4 縦書き3の順に接続されてしまい、正しい順序にならない。また、ブロック単位で右から順に行を接続した場合には、縦書き1 縦書き3 縦書き2 縦書き4の順に接続されてしまい、こちらも正しい順序にはならない。このように行の接続順序が失敗する可能性が高くなる。そのため、このようにブロック位置が揃っていない文書画像の場合には、リフロー型に変換しないと判定する。

【0119】

変換可否判定処理部34がリフロー型に変換しないと判定した文書画像（ページ）とリフロー型に変換すると判定した文書画像とを混在して表示させる処理については、後述の<9-3>の章にて説明する。

【0120】

[ 8 . 再配置処理部 ]

< 8 - 1 . 再配置処理部の処理 >

再配置処理部35は、変換可否判定処理部34にてリフロー型に変換すると判定された文書画像に対して、以下の様に参照リストを生成する。

【0121】

再配置処理部35は、レイアウト解析処理部33で定義した順序に従って行を呼び出す。そして、呼び出した行が文字列行である場合は対応する文字列行の先頭から順に文字を呼び出し、呼び出した行が図表行である場合は対応する図又は表を呼び出して、各要素の参照情報（要素を呼び出すための情報）を順に記述する。更に、要素が改行位置が含まれる行の末尾の要素である場合は、改行命令を挿入して参照リストを生成する。

【0122】

つまり、参照リストは、文字列行及び図表行の順序付けに従って、文書画像に含まれる各要素の順序が、文字列中の文字においては文書第1方向に沿って、記述されている。こ

10

20

30

40

50

ここで、参照リストは、HTMLに代表されるようなマークアップ言語形式で記述することができる。以下では、HTMLファイルとして記述する場合を例として、再配置処理部35の詳細について説明する。

【0123】

図25は、再配置処理部35の処理手順を示すフローチャートである。以下、図25に従って再配置処理部35の処理内容を説明する。

【0124】

図25に示すように、まず、ファイルのヘッダの記述を行う(S11)。ファイルのヘッダは、そのファイルの各種情報を記述するものである。例えばHTMLファイルの場合、図26に示すように、そのファイルがHTML言語で記述されていることの宣言(<HTML>タグ)や、本文には記載しないファイルの情報、例えばスタイルの定義やページタイトル等の情報、コメント、本文の記述が開始することの宣言(<BODY>タグ)等が含まれる。

10

【0125】

続いて、本文の記述を行う。ここでは、レイアウト解析処理部33で生成した文書構造ツリーの第0段落を初期呼び出し位置として、呼び出し位置において文書構造ツリーから段落の情報が呼び出し可能であるかどうかを判定する(S12)。呼び出し可能である場合(S12の判定がYES)、文書構造ツリーから段落を呼び出し、該段落が少なくとも1行以上の行を持つかどうか判定する(S13)。該段落が少なくとも1行以上の行を持つ場合(S13の判定がYES)、該段落を構成する全ての行の情報を順次呼び出し、該行に含まれる要素を参照してファイルで表示するための記述を行うファイル記述処理を実行する(S14)。他方、該段落が1つも行を持たない場合(S13の判定がNO)、呼び出し位置を次の段落に移し、S12に戻る。

20

【0126】

図27は、S14でのファイル記述処理の処理手順を示すフローチャートである。呼び出した段落について、初めに、段落を開始することを宣言する段落開始宣言命令(後述の<8-2>の章を参照)を実行する(S21)。その後、該段落が持つ文字列行を、文書構造ツリーの先頭から順次呼び出し、該行に含まれる文字要素を参照するための要素参照処理を実行する(S22)。つまり、該行の先頭から順に要素(文字)を呼び出し、該要素の参照命令を実行する。具体的には、該行の先頭から順に要素を呼び出して、全ての要素に対して同様の処理を繰り返す。

30

【0127】

全ての要素に対して参照命令を終了すると、該行についての参照処理を終了し、次の文字列行を呼び出して同様の処理を繰り返す。以上の処理を、文書構造ツリーにおける該段落が持つ全ての文字列行(該段落の文字列行ツリーに含まれる全ての文字列行)に対して実行する。該段落の文字列行ツリーに含まれる全ての文字列行に対する要素参照処理を終え、次に、図表行の要素参照処理を実行する(S23)。図表行は1つの要素しか持たないため、各図表行について1回の参照命令を行うと、該段落の持つ次の図表行(該段落の図表行ツリーに含まれる次の図表行)を呼び出し、同様の参照処理を実行する。該段落の図表行ツリーに含まれる全ての図表行について参照命令を実行すると、最後に、該段落についてのファイル記述処理が終了することを宣言する段落終了宣言(後述の<8-2>の章を参照)を実行し(S24)、ファイル記述処理(S14)を終了する。

40

【0128】

図25に戻り説明を続ける。S14の後、呼び出し位置を次の段落に更新し、S12に戻る。

【0129】

以上の処理を、S12の判定がNOになる、すなわち、文書構造ツリーの全ての段落に対するファイル記述処理を終了するまで反復する。S12の判定がNOになると、フッタの記述を行う(S15)。図26に示すように、フッタには、ヘッダの記述において開始を宣言した事項(例えば、本文の記述や、HTML言語の記述の開始等)の終了の宣言等

50

が含まれる。フッタ部分の記述を終えるとファイルを保存し ( S 1 6 )、終了する。なお、当該ファイルは、再配置処理部 3 5 から出力される。

【 0 1 3 0 】

< 8 - 2 . 各命令の例 >

段落開始宣言命令の例として、例えば、段落タグ < p > を挿入する方法が挙げられる。この場合、段落終了宣言命令は必ず段落タグ < / p > を挿入する方法とする。また、別の方法として、段落開始宣言命令は特に何も行わず、段落終了宣言命令として改行タグ < b r > を挿入する方法も挙げられる。要素の参照命令は、例えば、出力ファイルに画像表示タグ < i m g > を挿入して、該要素の切り取り画像ファイルのパスを指定して表示させる方法が挙げられる。また、図表要素の参照命令は文字要素の参照命令と同様としてもよいし、図表要素の場合のみ、改行タグ < b r > や表タグ < t a b l e > 及び < / t a b l e > 等の挿入処理と組み合わせることで、より視認性の高いレイアウトを構成することができる。

10

【 0 1 3 1 】

< 8 - 3 . 図表を段落の先頭に配置修正したい場合 >

上記では、各段落において図表を本文の後に配置する場合の処理手順について説明した。反対に、図表を本文の前に配置したい場合は、図 2 7 において、 S 2 2 と S 2 3 とを入れ替えることで容易に実現できる。また、図表を、段落の末尾に配置、段落の先頭に配置、もしくは図表を表示しない ( 文字列のみ表示する ) 等の複数のモードから操作パネル 6 を通じてユーザが指定できるようにしてもよい。

20

【 0 1 3 2 】

< 8 - 4 . 文書のスタイルの定義 >

上記命令のほか、スタイルシートを組み込むことで、文書のファイル書式を変更しても構わない。スタイルシートは、 HTML 言語等で構造化された文書の見栄え、表示形式を効率的に制御する公知の手段である。これらの見栄え等の情報を「スタイル」と呼ぶ。スタイルシートは専用のコンピュータ言語で実現され、その例として、 CSS ( C a s c a d i n g S t y l e S h e e t ) 等が挙げられる。また、スタイルシートを定義する場所は大きく分けて 3 つある。それは、ヘッダ要素内 ( < h e a d > ~ < / h e a d > 間 ) にスタイル要素 ( < s t y l e > ~ < / s t y l e > ) を追加してスタイルを記述する方法、スタイルを記述した外部ファイルを用意して、参照リストのヘッダ要素内で前記外部ファイルを呼び出す方法、及び本文部分 ( < b o d y > ~ < / b o d y > 間 ) における各種タグにおいて、そのタグにのみ有効となるスタイルを記述する方法、である。これらの方法はただ 1 つに選択されるものではなく、複数の方法を組み合わせてスタイルシートを定義することもできる。本実施の形態では、公知の方法を利用するものとして以下では詳述をせず、簡単な記述例の紹介に留める。

30

【 0 1 3 3 】

図 2 8 は、 CSS 形式で記述したスタイルシートの外部ファイルの例である。図中の範囲 A では、段落を定義するタグ < p > のスタイルを定義しており、この例では各段落の先頭行に、1文字分のインデントを付加することが定義されている。これにより、段落開始宣言タグが呼び出された場合は常に1文字分のインデントを付加することができる。範囲 B と範囲 C では、ともに画像を参照するタグ < i m g > のスタイルを定義しているが、範囲 B では g a i j i というクラスに属する場合に限定したスタイルであり、範囲 C では f i g というクラスに属する場合に限定したスタイルである。

40

【 0 1 3 4 】

このように、同じタグに対しても、それぞれ固有のスタイルを持つ複数のクラスを定義して外部ファイルとして保存し ( ここでは s t y l e . c s s というファイル名をつけている )、図 2 9 に示す例 ( HTML 言語で記述、一部のみ抜粋 ) のように、ヘッダ内で図 2 9 の ( a ) に示すような前記外部ファイル ( 図 2 8 のものと同じ ) を読み込み ( 図 2 9 の ( b ) の 3 行目 )、各タグにおいて所望のスタイルを持つクラスを指定することで、局所的なスタイルの指定を行うことが可能となる。図 2 8 及び図 2 9 の例では、画像の参照

50



時において、文字を表わす画像である場合に、imgタグ内において、図29の(b)に示すようにclass="gaiji"と記述することで、gaiji(外字)クラスを指定し、図表を表わす画像である場合にはclass="fig"と記述することで、fig(図)クラスを指定しており、このように要素毎に適切なスタイルを定義することで、より見栄えのよい文書を表示することが可能となる。

#### 【0135】

また、操作パネル6でこれらのスタイルを選択して指定できるようにしてもよい。例えば、文書画像データから決定した文書第1方向とは別に、再配置する際の文書の方向を、操作パネル6を通じてユーザが指定できるようにしてもよい。具体的には、HTML言語形式のフォーマットにおいてCSS形式によりhtml{writing-mode: 10  
t b - r l ; }とスタイルシートを定義することで、本文全体の文書の行方向を上から下、更に行が右から始まり左に進む、すなわち縦書き表記が可能となる。

#### 【0136】

従って、縦書きでのリフロー型の表示が選択された場合に、前記のようなスタイルシートの定義を追加するようにすることで、縦書き表示を実現できるようになる。なお、縦書き表記の実現方法は前記の方法以外の方法をとることができる。また、出力時の文書の方向として「自動モード」を準備し、行解析処理部31で得られた文書第1方向と同じ方向で出力するように自動的に選択するようにすることもできる。

#### 【0137】

##### < 8 - 5 . 他のファイル形式 >

本実施の形態の再配置処理部35においては、入力された文書画像からHTMLファイル 20  
を出力する場合について述べたが、出力するファイルは、HTMLファイルに限らず、リフロー型表示を実現するあらゆるファイル形式(例えば、XMLファイル、XPDFファイル等)から選択できる。

#### 【0138】

本実施の形態では、以上のように、文書画像からその構造を解析し、行の順序を理解することで、行を順に参照し、さらにその行を構成する文字や図表を先頭から順に参照して 30  
いくことで、フィックス型の画像ファイルとして生成された文書画像であっても、リフロー型のファイルとして変換するための情報(参照ファイル)を得ることができる。また、改行の有無を判定し、段落の範囲を定義することで、文書画像に行の折り返しがあっても、1つの段落内では改行をしないようにするほか、各段落に従属する図又は表の配置を、その段落の先頭や末尾にまとめて表示する等の修正をすることで、文字列と文字列の間に 30  
図又は表が挟まれていても、同一段落内とみなされていれば、そのまま図又は表を挟まずに配置するよう修正することにより、文章の連続性及び可読性を向上することができる。

#### 【0139】

以上のように、参照リストは、文字列行及び図表行の順序付けと文書の記述方向とに従ったものであり、リフロー型のファイルとして変換するための情報である。参照リストとして、文書画像をリフロー型表示が可能なファイル形式に変換したものを生成することで、処理量を減らして、あらゆる表示装置の表示領域(表示画面)にそれぞれ適したレイアウト配置が可能となる。参照リストを用いることで、表示装置では、文書画像の文書の記述方向に垂直な方向のスクロール操作のみでの表示を実現することが可能となる。 40

#### 【0140】

また、参照リストを、文書構造を列記したテキストデータとして参照リストを生成することで、閲覧用途でなく文書構造の解析用途とすることができる。加えて、出力されたテキスト形式の参照リストから更に所望のファイルフォーマットに変換することで、別のファイルフォーマットで出力したい場合に最初から処理をやり直さなくてもよくなる。

#### 【0141】

##### [ 9 . 表示装置 ]

##### < 9 - 1 . 表示装置での参照リストの使用 >

前記生成された参照リスト及び切り出した各要素の画像データは、本実施形態の画像形 50

成装置 100 が備える送信装置 4 から送信され、図示しない受信側の装置である表示装置（例えば、スマートフォンやタブレット等であってもよい）が備えるアプリケーションである閲覧プログラム（ビューア）を通じて閲覧することができる。参照リストが取るファイル形式によって最適な閲覧プログラムは異なり、例えば、参照リストとして HTML ファイルの形式をとる場合、Internet Explorer（登録商標）等、良く知られた HTML5 をサポートしたウェブブラウザを用いて開くことで、容易にリフロー型の表示を実現することが可能となる。

【0142】

なお、特定のマークアップ言語形式のファイルに変換せず、各要素（文字、図、表）を参照する順序と段落の開始宣言及び終了宣言等、再配置処理部 35 で得られた文書構造を列記したテキストデータとして参照リストを生成することもできる。この参照リストは、例えば文書構造の解析結果として利用することも可能であるし、受信側の装置が備えるコンピュータプログラム（変換プログラム）等によって、所望のファイル形式に変換することも可能である。また、このようにファイル形式の変換を二段階とすることで、処理ステップ数の増加はあるものの、同じ文書画像データから複数のファイル形式への変換を行いたい場合に参照リストを共通で使用することが可能となる。さらに、前記変換プログラムにおいてレイアウト解析結果の手動による修正も可能となる。

10

【0143】

受信側の表示装置では、画像形成装置 100 から送信される参照リストとして、表示装置の備えるビューアアプリケーションに適したファイル形式（HTML 等）に変換したものを受信すると、そのアプリケーションにおいてファイルを読み込むだけでリフロー型の文書に変換されたものを表示できるので、特別な処理はない。ただし、参照リストが特定のビューアと関連付けされていない、たとえば、最も単純な形式として座標情報等を列記しただけのようなものである場合、そのままではリフロー型の表示はできない。このような参照リストを受信した場合、再配置するための処理は必要となるが、公知の処理を利用することができる。

20

【0144】

また、表示装置の表示領域の表示幅に合わせて、参照リストが参照する文字は、折り返し部分に変更され、図や表については、表示幅に合わせて縮小又は拡大表示される。例えば、上記 < 8 - 4 > で説明したように、図又は表に fig クラスを割り当て、HTML ファイル形式で出力する場合、画像の幅（もしくは高さ）を表示装置の表示幅（もしくは高さ）に対する割合として設定することで、表示装置の表示幅や表示倍率に変更となった場合でも、そのときの表示幅（もしくは高さ）に合わせて自動調整させることが可能となる。

30

【0145】

上記のように表示装置の表示幅（横書きの場合）に対する割合として画像の幅を設定する方法の例として、たとえば、img タグの呼び出し時に fig クラスに属するもののみ img タグのサイズ属性として下記のようにパーセンテージを追加する、

```
<img class = "fig" src = " ~ ~ " width = " 90 % " / > . . . ( 8 )
```

40

もしくは CSS 形式で fig クラスのスタイルシートとして下記を追加する。

```
.fig { width : 90 % ; } . . . ( 9 )
```

等の方法がある。上記（8）又は（9）の方法はいずれも、表示幅に対して 90% のサイズが図又は表の幅として設定され、表示領域が変更された際には自動で変更後の表示幅の 90% を図又は表の幅として再設定する。（8）は、上記のサイズ属性の設定を追加した img タグのみに適用される、すなわち個別に適用されるのに対し、（9）は、fig クラスを割り当てた図表を一括で設定する点で異なる。もちろん、上記以外の公知の方法を利用することも可能である。

【0146】

< 9 - 2 . 図表の表示の別の例 >

50

上記< 8 - 1 >、< 8 - 3 >では、段落の最後又は最初に図表を表示する方法を開示したが、それらとは別に、図表を本文と同じファイル上に混在して表示せずに、画像へのリンク一覧（先頭の段落から順に図表を並べる）を別途作成し、そのリンク一覧からユーザが所望したときに選択した画像を個別に表示できるようにしてもよい。

【 0 1 4 7 】

例えば、画像形成装置 1 0 0 にて生成された参照リスト及び抽出した（切り出した）各要素の画像を受信する表示装置において、コンテンツを表示するコンテンツ表示領域とユーザ操作を受け付ける操作領域とを個別に備え、操作領域に、本文と画像リンク一覧との表示を切り替える操作機能を実行する手段を備える。このような構成とすることで、ユーザが所望するときに、コンテンツ表示領域に表示する内容を本文と図表とで切り替えるようにすることができる。

10

【 0 1 4 8 】

また、表示装置において、コンテンツ表示領域と操作領域とを個別に備え、操作領域に、リンク一覧の先頭から順に各画像へのリンクを表示し、選択できるように構成してもよい。リンクの表示の方法は、例えば、符号（図 1、図 2、・・・等順序を認識できるものが望ましい）や、画像のサムネイル等を利用することができる。リンクを選択すると、選択したリンクに対応する図表を表示する。図表の表示方法は、例えば、コンテンツ表示領域に表示する方法のほかに、コンテンツ表示領域に重畳して図表表示領域を生成し、図表表示領域に該当する図表を表示する、いわゆるポップアップ形式を採用してもよい。

【 0 1 4 9 】

20

上記のように表示装置がコンテンツ表示領域と操作領域とを備える場合、操作領域は、コンテンツ表示領域と必ずしも分離されている必要は無く、コンテンツ表示領域に重畳して表示してもよく、また、操作領域は常時表示せずに表示の命令が入力された場合にのみ画面上に表示するようにしてもよい。表示の命令の入力方法は、例えば、表示装置が備えるタッチパネルのうち表示領域に該当する範囲において、一定時間以上、タッチ開始時点でのタッチ座標から一定距離以上離れることなくタッチパネルをタッチした状態を継続する（いわゆる長押し）等が挙げられる。

【 0 1 5 0 】

なお、上記の方法自体は、図表の順序や対応する段落との関連付けの精度を向上するものではないが、図表の表示位置を固定せず、「（ほぼ）本文の順序通りに並んだ図表の一覧」からユーザ自身が選択して閲覧することで、対応する段落との関連付けが困難な複雑なレイアウトの文章を読み進める上での違和感を解消させることができる。

30

【 0 1 5 1 】

< 9 - 3 . リフロー型とフィックス型の文書の表示 >

変換可否判定処理部 3 4 がリフロー型に変換不可と判定したページとリフロー型に変換可能と判定したページとを混在して表示させる処理について以下で説明する。

【 0 1 5 2 】

フィックス型の場合、ページ全体を 1 つの画像として扱うことで、リフロー型と混在させて表示させることが可能である。ただし、画像として扱われるフィックス型のページをリフロー型と混在させて表示させると画像が全画面表示になってしまう。そのため、文字が読みづらいページもある。その場合には、表示装置をリフロー型とフィックス型の表示を切り替えられる構成にしておくと、各ページに合わせて読みやすいフォーマットで表示させることが可能となる。

40

【 0 1 5 3 】

例えば、HTML 5 などを用いてウェブブラウザで表示させる場合、ブラウザの表示領域内にリフロー型表示とフィックス型表示との切替ボタン（切替部）を表示させ、押されたボタンに合わせて表示領域にリフロー型での表示とフィックス型での表示を切り替えて表示できるようにしておけばよい。例えば、切替ボタンとしては、各型を選択するボタンをそれぞれ設ける、あるいは、トグル状のスイッチを設け、何れかの型を選択できるようにすればよい。

50

## 【 0 1 5 4 】

〔実施の形態 2：画像読取装置〕

実施の形態 1 では、本発明に係る文書画像処理装置を画像形成装置が有する画像処理装置に適用した構成について説明したが、これに限るものではない。そこで、本実施の形態では、本発明に係る文書画像処理装置を変換処理部として、フラットベッドスキャナ等の画像読取装置が有する画像処理装置に適用した例について説明する。

## 【 0 1 5 5 】

なお、実施の形態 1 の説明に用いた図面に記載されている部材と同じ機能を有する部材については、以下の説明においても同じ符号を付記する。また、それらの各部材の詳細な説明はここでは繰り返さない。

## 【 0 1 5 6 】

図 30 は、実施の形態 2 に係る画像処理装置 1 a を備える画像読取装置（情報処理装置）200 の構成を示すブロック図である。図 30 に示すように、画像読取装置 200 は、画像処理装置 1 a、画像入力装置 2、送信装置 4、記憶部 5、及び操作パネル 6 を備えている。画像処理装置 1 a は、A/D 変換部 11、シェーディング補正部 12、原稿種別判別部 13、入力階調補正部 14、領域分離処理部 15、圧縮処理部 21、及び変換処理部（文書画像処理装置）22 を備えている。当該変換処理部 22 にて、実施の形態 1 にて説明したのと同様に、参照リストが生成される。

## 【 0 1 5 7 】

画像読取装置 200 で実行される各種処理は、画像読取装置 200 に備えられる図示しない制御部（CPU あるいは DSP 等のプロセッサを含むコンピュータ）により制御される。

## 【 0 1 5 8 】

本実施の形態では、画像読取装置 200 は、スキャナに限定されることはなく、例えば、デジタルスチルカメラ、書画カメラ、あるいは、カメラを搭載した電子機器類（例えば、携帯電話、スマートフォン、タブレット端末等）であってもよい。これらカメラあるいはカメラを搭載した電子機器類においては、自装置にて文書画像の構造解析を行って、自装置の表示部にてリフロー型あるいはフィックス型で表示することが可能に構成されていてもよい。

## 【 0 1 5 9 】

〔実施の形態 3：ネットワークを経由した文書画像処理装置〕

上記では、本発明に係る文書画像処理装置を、画像形成装置 100 が有する画像処理装置 1 あるいは画像読取装置 200 が有する画像処理装置 1 a に適用する例を示したが、これに限るものではない。本発明に係る文書画像処理装置を、例えばサーバ装置に適用してもよい。この場合のサーバ装置の構成の一例は、画像形成装置あるいは画像読取装置により画像読取及び各種画像処理が施された文書画像をネットワークを介して受信する受信装置と、実施の形態 1 にて説明した変換処理部 22 での処理を実行する文書画像処理装置と、当該文書画像処理装置から出力されたファイル（文書画像及び参照リスト）をネットワークを介して送信する送信装置と、を備えたサーバ装置（情報処理装置）である。

## 【 0 1 6 0 】

このようにサーバ装置を構成することにより、画像形成装置あるいは画像読取装置にて画像読取及び各種画像処理が施された文書画像を、ネットワークを経由して受信して、上記の変換処理部 22 での処理を実行する文書画像処理装置により参照リストを生成し（フォーマット変換を適用し）、出力されたファイルをユーザの端末装置（例えば、スマートフォンやタブレット端末等）に送信する、という使い方が可能となる。また、このサーバ装置により、既に設置された画像形成装置あるいは画像読取装置を交換することなく、フォーマット変換機能を利用することが可能となる。また、フォーマット変換後のファイルをサーバ装置に記憶しておくことで、ユーザが望むときに変換後のファイルを受信して閲覧することも可能となる。

## 【 0 1 6 1 】

あるいは、本発明に係る文書画像処理装置を、例えば、携帯電話、スマートフォン、タブレット端末、電子書籍専用端末等の通信端末装置に適用してもよい。この場合の通信端末装置の構成の一例は、画像形成装置あるいは画像読取装置により画像読取及び各種画像処理が施された文書画像をネットワークを介して受信する受信部と、実施の形態1にて説明した変換処理部22と、実施の形態1にて説明した表示装置と、を備えた通信端末装置（情報処理装置）である。当該通信端末装置は、電子化された文書画像を受信し、受信した文書画像の構造解析を行って、リフロー型あるいはフィックス型にて表示することが可能である。

#### 【0162】

なお、上記サーバ装置の受信装置又は上記通信端末装置が受信する文書画像は、画像形成装置や画像読取装置等にて生成された文書画像である必要は無い。例えばWordファイルやPDFファイル等のように、構造化された文書ファイルを電子化したものを文書画像として受信してもよい。

#### 【0163】

〔実施の形態4：記録媒体・プログラム〕

上記で説明した画像処理装置1, 1a（特に、変換処理部22, 22a）、サーバ装置（特に、文書画像処理装置）、通信端末装置（特に、変換処理部）は、集積回路（ICチップ）等に形成された論理回路（ハードウェア）によって実現してもよいし、CPUを用いてソフトウェアによって実現してもよい。

#### 【0164】

後者の場合、画像処理装置1, 1a、サーバ装置、通信端末装置は、各機能を実現するソフトウェアであるプログラムの命令を実行するCPU、上記プログラム及び各種データがコンピュータ（又はCPU）で読み取り可能に記録されたROM（Read Only Memory）又は記憶装置（これらを「記録媒体」と称する）、上記プログラムを展開するRAM（Random Access Memory）等を備えている。そして、コンピュータ（又はCPU）が上記プログラムを上記記録媒体から読み取って実行することにより、本発明の目的が達成される。上記記録媒体としては、「一時的でない有形の媒体」、例えば、テープ、ディスク、カード、半導体メモリ、プログラマブルな論理回路等を用いることができる。また、上記プログラムは、該プログラムを伝送可能な任意の伝送媒体（通信ネットワークや放送波等）を介して上記コンピュータに供給されてもよい。なお、本発明は、上記プログラムが電子的な伝送によって具現化された、搬送波に埋め込まれたデータ信号の形態でも実現され得る。

#### 【0165】

なお、上記で説明した文書画像処理装置及び文書画像処理方法は、カラーの画像データを扱う構成としたが、これに限るものではなく、白黒の画像データを扱う構成であってもよい。

#### 【0166】

本発明は上述した各実施の形態に限定されるものではなく、種々の変更が可能である。すなわち、本発明の要旨を逸脱しない範囲内において適宜変更した技術的手段を組み合わせ得られる実施の形態についても本発明の技術的範囲に含まれる。

#### 【0167】

〔まとめ〕

本発明の態様1に係る文書画像処理装置（変換処理部22）は、文書を電子化した文書画像の再構成を行う文書画像処理装置において、前記文書画像の構造解析を行う構造解析部（行解析処理部31、行ブロック解析処理部32、レイアウト解析処理部33）と、前記構造解析により前記文書画像から抽出した文字列あるいは図又は表の特徴量に基づいて、前記文書画像に含まれる各文字、図、及び/又は表である各要素を再構成するか否かの判定を行う変換判定部（変換可否判定処理部34）と、前記変換判定部が前記文書画像の各要素を再構成すると判定すると、前記構造解析部による解析結果に基づいて、前記文書画像を再構成した際の前記各要素の順序を記述した参照リストを生成する参照リスト生成

10

20

30

40

50

部（再配置処理部 35）と、を備える。

【0168】

上記構成によると、構造解析により前記文書画像から抽出した文字列あるいは図又は表の特徴量に基づいて、前記文書画像に含まれる各要素を再構成するか否かの判定を行うことができる。このように、文書画像に含まれる各要素を再構成するか、つまり、文書画像をリフロー型のファイル形式にするか、文書画像に含まれる各要素を再構成しないか、つまり、フィックス型（固定レイアウト）のファイル形式にするか、を文書画像処理装置にて判断することで、常に最適なファイル形式での表示を行うことが可能となる。また、構造解析部による解析結果に基づいて、文書画像を再構成した際の各要素の順序を記述した参照リストの生成により、文書画像をリフロー型のファイル形式にするための情報を生成

10

【0169】

本発明の態様 2 に係る文書画像処理装置は、態様 1 において、前記構造解析により抽出された前記文字列からは前記文書画像での文書の記述方向を決定できない場合に、前記変換判定部は、前記文書画像を再構成しないと判定する。

【0170】

文書画像に縦書きの行と横書きの行とが無視できない比率で混在していると、文書画像全体の方向を一意に判別できず、また、レイアウト解析する際にも、縦書きと横書きの行をどう接続すればよいかの判定が困難になる。よって、正しくリフロー型に変換するのが困難となる。そこで、上記構成により変換判定部が上記判定を行うことで、正しくリフロー型に変換できない、というミスを防ぐことが可能となる。

20

【0171】

本発明の態様 3 に係る文書画像処理装置は、上記態様 1 において、前記構造解析により前記文書画像が図及び / 又は表のみから構成されていると判定される場合に、前記変換判定部は、前記文書画像を再構成しないと判定する。

【0172】

図及び / 又は表は、その中に文字があったとしても、レイアウトを崩さずにそのまま表示しないと内容がわからなくなってしまうことが多い。そこで、上記構成により変換判定部が上記判定を行うことにより、図及び / 又は表のみから構成されている文書画像を間違

30

【0173】

本発明の態様 4 に係る文書画像処理装置は、上記態様 1 において、前記構造解析により前記文書画像から抽出された文字列の数が、予め定められる第 1 閾値以下である場合に、前記変換判定部は、前記文書画像を再構成しないと判定する。

【0174】

文字列が少ない場合、わざわざリフロー型にして読むほどの文章の長さではないと考えられる。そのため、上記構成により変換判定部が上記判定を行うことにより、無駄にリフロー型に変換することを防ぐことが可能となる。

【0175】

本発明の態様 5 に係る文書画像処理装置は、上記態様 1 において、前記構造解析により前記文書画像から抽出された文字列の高さあるいは幅が、予め定められた第 2 閾値以上である場合に、前記変換判定部は、前記文書画像を再構成しないと判定する。

40

【0176】

文字が十分大きい場合には、縮小表示しても十分可読性があると考えられるため、リフロー型にする必要がない。そのため、上記構成により変換判定部が上記判定を行うことにより、無駄にリフロー型に変換することを防ぐことが可能となる。

【0177】

本発明の態様 6 に係る文書画像処理装置は、上記態様 1 において、前記構造解析により、前記文書画像から抽出された複数の文字列あるいは図又は表よりなるブロックのそれぞれの位置が不規則である場合に、前記変換判定部は、前記文書画像を再構成しないと判定

50

する。

【0178】

文字列あるいは図又は表よりなるブロックの位置が一定の範囲内で揃っていない場合には、整理された段組みされたレイアウトの文書画像ではない可能性が高い。そのため、行ブロックの接続が失敗し、誤った行ブロック同士をつないでしまう可能性が高い。よって、上記構成により変換判定部が上記判定を行うことにより、正しくリフロー型に変換できない、というミスを防ぐことが可能となる。

【0179】

本発明の態様7に係る情報処理装置は、上記態様1から6の何れか1つに記載の文書画像処理装置と、前記変換判定部により各要素を再構成しないと判定された文書画像をフィックス型の表示が可能にフォーマット変換するフォーマット変換処理部と、前記変換判定部により各要素を再構成すると判定された文書画像については、当該文書画像の前記参照リスト及び当該文書画像に含まれる各要素の画像データを送信し、かつ、前記変換判定部により各要素を再構成しないと判定された文書画像については、前記フォーマット変換された文書画像を送信する送信装置と、を備える。

10

【0180】

上記情報処理装置は、例えば、画像形成装置、画像読取装置、サーバ装置等であってもよい。また、画像読取装置は、スキャナ、デジタルスチルカメラ、書画カメラ、あるいは、カメラを搭載した電子機器類（例えば、携帯電話、スマートフォン、タブレット端末等）等であってもよい。上記情報処理装置が、例えば、画像形成装置である場合、読み込まれた画像データに変換処理を施して画像データを再構成し、アドレスを指定して再構成された画像データを受信側装置（例えば、タブレット端末等）に送信することにより、受信側装置にて、一方向（文書の記述方向と直交する方向）のスクロールのみで画像を閲覧することができる。

20

【0181】

本発明の態様8に係る表示装置は、上記態様7の情報処理装置から受信した文書画像を表示する表示装置であって、前記変換判定部により各要素を再構成すると判定された文書画像の前記参照リスト及び当該文書画像に含まれる各要素の画像データに基づき、当該文書画像をリフロー型にて表示し、かつ、前記フォーマット変換された文書画像をフィックス型にて表示する。

30

【0182】

上記構成によると、変換判定部により各要素を再構成すると判定された文書画像をリフロー型にて表示し、かつ、上記フォーマット変換された文書画像をフィックス型にて表示することができ、リフロー型のページとフィックス型のページとが混在した原稿の表示を行うことができる。

【0183】

本発明の態様9に係る表示装置は、上記態様8の表示装置において、上記態様7の情報処理装置から、さらに前記変換判定部により各要素を再構成すると判定された文書画像をフィックス型の表示が可能にフォーマット変換した文書画像を受信し、前記変換判定部により各要素を再構成すると判定された文書画像の表示を、リフロー型とフィックス型とで切り替える切替部を備える。

40

【0184】

上記構成によると、変換判定部により各要素を再構成すると判定された文書画像の表示を、ユーザがリフロー型とフィックス型とで切り替えることができる。

【0185】

なお、上記文書画像処理装置、上記情報処理装置、又は上記表示装置は、コンピュータによって実現してもよい。この場合には、コンピュータを上記各部として動作させることにより上記文書画像処理装置、上記情報処理装置、又は上記表示装置をコンピュータにて実現させるプログラム、及びそのプログラムを記録したコンピュータ読み取り可能な記録媒体も、本発明の範疇に入る。

50

【産業上の利用可能性】

【0186】

本発明は、文書画像の再構成を行う文書画像処理装置等に利用することができる。

【符号の説明】

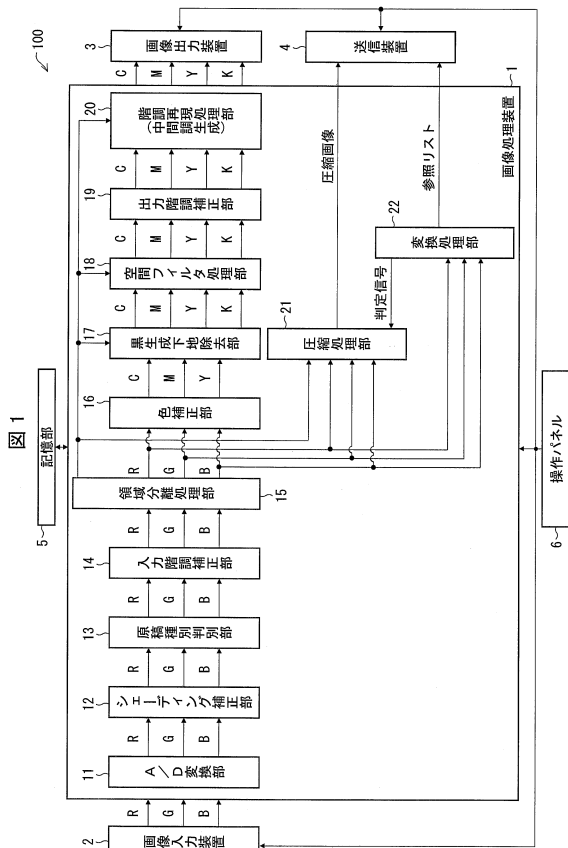
【0187】

- 1, 1a 画像処理装置
- 4 送信装置
- 22, 22a 変換処理部(文書画像処理装置)
- 31 行解析処理部(構造解析部)
- 31a 文字列抽出処理部
- 31b 図表抽出処理部
- 32 行ブロック解析処理部(構造解析部)
- 33 レイアウト解析処理部(構造解析部)
- 34 変換可否判定処理部(変換判定部)
- 35 再配置処理部(参照リスト生成部)
- 37 段組解析処理部
- 38 順序付け処理部
- 39 段落解析処理部(段落解析部)
- 100 画像形成装置(情報処理装置)
- 200 画像読取装置(情報処理装置)

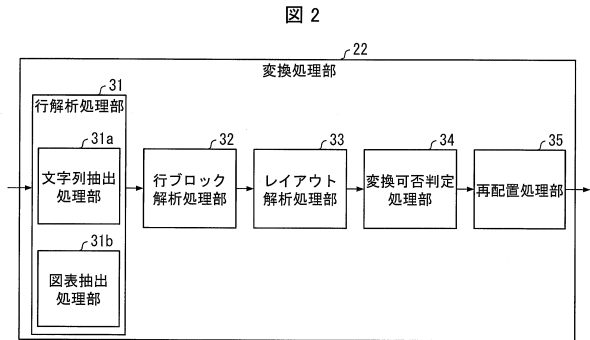
10

20

【図1】



【図2】



【図3】

図 3

種別	ルール	対象となる文字の例
行頭禁則	行頭に 来てはいけない	句読点 。(句点)、(読点)、(ピリオド)、(コンマ) など 終わり括弧類 ) ] ] ] } } ] ] " など
行末禁則	行末に 来てはいけない	始め括弧類 ( [ [ [ [ [ [ 《 『 『 “ など



【図4】

図4

これは、句読点を含んだ文です。これ

これは、句読点を含んだ文です。これ

【図5】

図5

### 文書タイトル

これは序文の文章です。これは序文の文章です。これは序文の文章です。

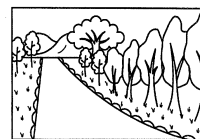
#### 1章タイトル

これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。

これは第1章の第2段落の文章です。これは第1章の第2段落の文章です。これは第1章の第2段落の文章です。これは第1章の第2段落の文章です。これは第1章の第2段落の文章です。これは第1章の第2段落の文章です。

#### 2章タイトル

これは第2章の第1段落の文章です。これは第2章の第1段落の文章です。これは第2章の第1段落の文章です。これは第2章の第1段落の文章です。これは第2章の第1段落の文章です。



これは第2章の第1段落の文章です。これは第2章の第1段落の文章です。これは第2章の第1段落の文章です。これは第2章の第1段落の文章です。これは第2章の第1段落の文章です。

これは第2章の第2段落の文章です。これは第2章の第2段落の文章です。これは第2章の第2段落の文章です。これは第2章の第2段落の文章です。これは第2章の第2段落の文章です。これは第2章の第2段落の文章です。

【図6】

図6

行ID:101 **文書タイトル**

行ID:102 これは序文の文章です。これは序文の文章です。これは序文の文章です。

行ID:103

行ID:104 **1章タイトル**

行ID:106 これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。これは第1章の第1段落の文章です。

行ID:107

行ID:108

行ID:109

行ID:110

行ID:112

行ID:114

行ID:116

行ID:118

行ID:120

行ID:122

行ID:124

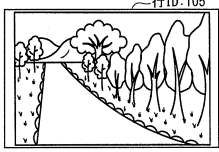
行ID:126

行ID:129 **2章タイトル**

行ID:131

行ID:133

行ID:135

行ID:105 

行ID:111

行ID:113

行ID:115

行ID:117

行ID:119

行ID:121

行ID:123

行ID:125

行ID:127

行ID:128

行ID:130

行ID:132

行ID:134

行ID:136

【図7】

行a

行b

行c

行d

行e

行f

行g

行h

行i

行j

行k

行l

行m

行n

行o

行p

行q

行r

行s

行t

行u

行v

行w

行x

行y

行z

行aa

行ab

行ac

行ad

行ae

行af

行ag

行ah

行ai

行aj

行ak

行al

行am

行an

行ao

行ap

行aq

行ar

行as

行at

行au

行av

行av

行aw

行ax

行ay

行az

行ba

行bb

行bc

行bd

行be

行bf

行bg

行bh

行bi

行bj

行bk

行bl

行bm

行bn

行bo

行bp

行bq

行br

行bs

行bt

行bu

行bv

行bv

行bw

行bx

行by

行bz

行ca

行cb

行cc

行cd

行ce

行cf

行cg

行ch

行ci

行cj

行ck

行cl

行cm

行cn

行co

行cp

行cq

行cr

行cs

行ct

行cu

行cv

行cv

行cw

行cx

行cy

行cz

行da

行db

行dc

行dd

行de

行df

行dg

行dh

行di

行dj

行dk

行dl

行dm

行dn

行do

行dp

行dq

行dr

行ds

行dt

行du

行dv

行dv

行dw

行dx

行dy

行dz

行ea

行eb

行ec

行ed

行ee

行ef

行eg

行eh

行ei

行ej

行ek

行el

行em

行en

行eo

行ep

行eq

行er

行es

行et

行eu

行ev

行ev

行ew

行ex

行ey

行ez

行fa

行fb

行fc

行fd

行fe

行ff

行fg

行fh

行fi

行fj

行fk

行fl

行fm

行fn

行fo

行fp

行fq

行fr

行fs

行ft

行fu

行fv

行fv

行fw

行fx

行fy

行fz

行ga

行gb

行gc

行gd

行ge

行gf

行gg

行gh

行gi

行gj

行gk

行gl

行gm

行gn

行go

行gp

行gq

行gr

行gs

行gt

行gu

行gv

行gv

行gw

行gx

行gy

行gz

行ha

行hb

行hc

行hd

行he

行hf

行hg

行hh

行hi

行hj

行hk

行hl

行hm

行hn

行ho

行hp

行hq

行hr

行hs

行ht

行hu

行hv

行hv

行hw

行hx

行hy

行hz

行ia

行ib

行ic

行id

行ie

行if

行ig

行ih

行ii

行ij

行ik

行il

行im

行in

行io

行ip

行iq

行ir

行is

行it

行iu

行iv

行iv

行iw

行ix

行iy

行iz

行ja

行jb

行jc

行jd

行je

行jf

行jg

行jh

行ji

行jj

行jk

行jl

行jm

行jn

行jo

行jp

行jq

行jr

行js

行jt

行ju

行jv

行jv

行jw

行jx

行jy

行jz

行ka

行kb

行kc

行kd

行ke

行kf

行kg

行kh

行ki

行kj

行kk

行kl

行km

行kn

行ko

行kp

行kq

行kr

行ks

行kt

行ku

行kv

行kv

行kw

行kx

行ky

行kz

行la

行lb

行lc

行ld

行le

行lf

行lg

行lh

行li

行lj

行lk

行ll

行lm

行ln

行lo

行lp

行lq

行lr

行ls

行lt

行lu

行lv

行lv

行lw

行lx

行ly

行lz

行ma

行mb

行mc

行md

行me

行mf

行mg

行mh

行mi

行mj

行mk

行ml

行mm

行mn

行mo

行mp

行mq

行mr

行ms

行mt

行mu

行mv

行mv

行mw

行mx

行my

行mz

行na

行nb

行nc

行nd

行ne

行nf

行ng

行nh

行ni

行nj

行nk

行nl

行nm

行nn

行no

行np

行nq

行nr

行ns

行nt

行nu

行nv

行nv

行nw

行nx

行ny

行nz

行oa

行ob

行oc

行od

行oe

行of

行og

行oh

行oi

行oj

行ok

行ol

行om

行on

行oo

行op

行oq

行or

行os

行ot

行ou

行ov

行ov

行ow

行ox

行oy

行oz

行pa

行pb

行pc

行pd

行pe

行pf

行pg

行ph

行pi

行pj

行pk

行pl

行pm

行pn

行po

行pp

行pq

行pr

行ps

行pt

行pu

行pv

行pv

行pw

行px

行py

行pz

行qa

行qb

行qc

行qd

行qe

行qf

行qg

行qh

行qi

行qj

行qk

行ql

行qm

行qn

行qo

行qp

行qq

行qr

行qs

行qt

行qu

行qv

行qv

行qw

行qx

行qy

行qz

行ra

行rb

行rc

行rd

行re

行rf

行rg

行rh

行ri

行rj

行rk

行rl

行rm

行rn

行ro

行rp

行rq

行rr

行rs

行rt

行ru

行rv

行rv

行rw

行rx

行ry

行rz

行sa

行sb

行sc

行sd

行se

行sf

行sg

行sh

行si

行sj

行sk

行sl

行sm

行sn

行so

行sp

行sq

行sr

行ss

行st

行su

行sv

行sv

行sw

行sx

行sy

行sz

行ta

行tb

行tc

行td

行te

行tf

行tg

行th

行ti

行tj

行tk

行tl

行tm

行tn

行to

行tp

行tq

行tr

行ts

行tt

行tu

行tv

行tv

行tw

行tx

行ty

行tz

行ua

行ub

行uc

行ud

行ue

行uf

行ug

行uh

行ui

行uj

行uk

行ul

行um

行un

行uo

行up

行uq

行ur

行us

行ut

行uu

行uv

行uv

行uw

行ux

行uy

行uz

行va

行vb

行vc

行vd

行ve

行vf

行vg

行vh

行vi

行vj

行vk

行vl

行vm

行vn

行vo

行vp

行vq

行vr

行vs

行vt

行vu

行vv

行vv

行vw

行vx

行vy

行vz

行wa

行wb

行wc

行wd

行we

行wf

行wg

行wh

行wi

行wj

行wk

行wl

行wm

行wn

行wo

行wp

行wq

行wr

行ws

行wt

行wu

行wv

行wv

行ww

行wx

行wy

行wz

行xa

行xb

行xc

行xd

行xe

行xf

行xg

行xh

行xi

行xj

行xk

行xl

行xm

行xn

行xo

行xp

行xq

行xr

行xs

行xt

行xu

行xv

行xv

行xw

行xx

行xy

行xz

行ya

行yb

行yc

行yd

行ye

行yf

行yg

行yh

行yi

行yj

行yk

行yl

行ym

行yn

行yo

行yp

行yq

行yr

行ys

行yt

行yu

行yv

行yv

行yw

行yx

行yy

行yz

行za

行zb

行zc

行zd

行ze

行zf

行zg

行zh

行zi

行zj

行zk

行zl

行zm

行zn

行zo

行zp

行zq

行zr

行zs

行zt

行zu

行zv

行zv

行zw

行zx

行zy

行zz

【 図 8 】



図 8

【 図 9 】

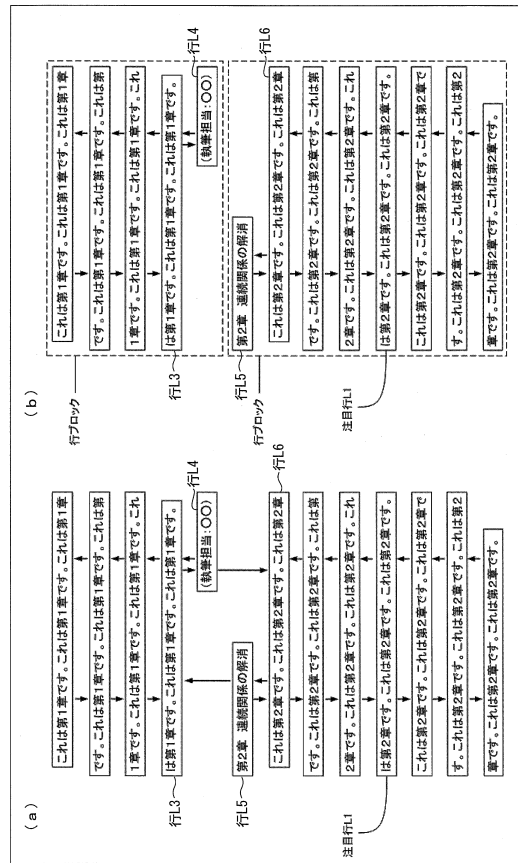
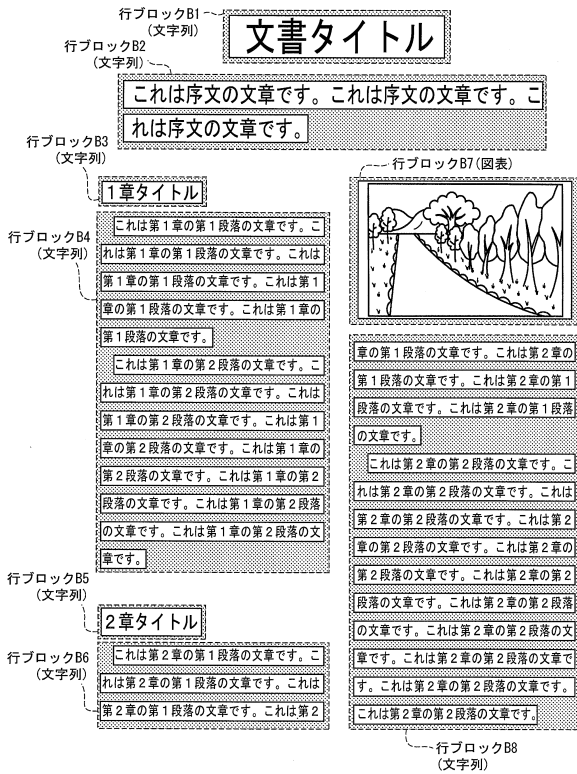


図 9

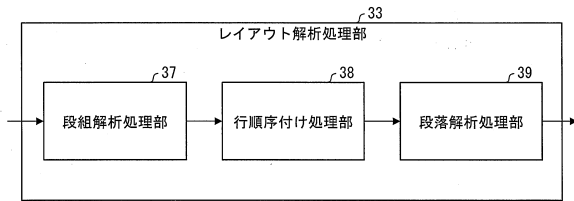
【 図 10 】

図 10

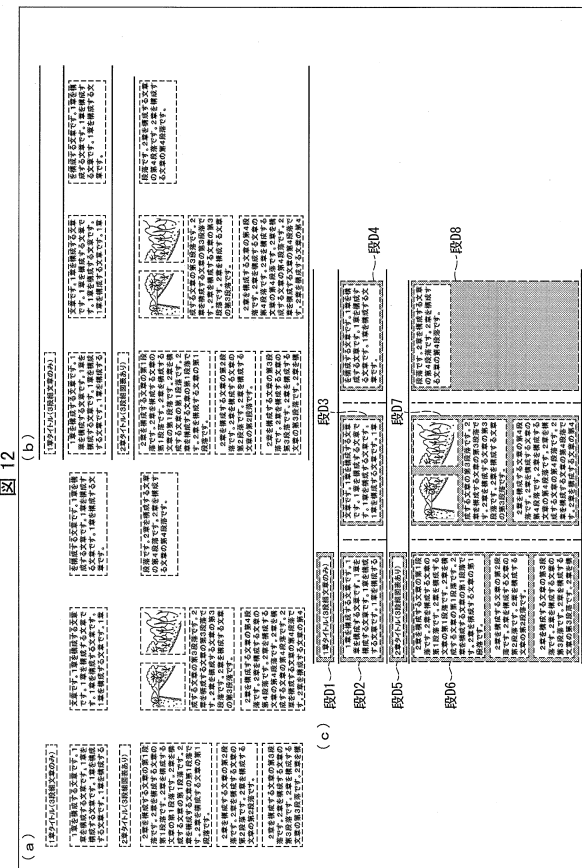


【 図 11 】

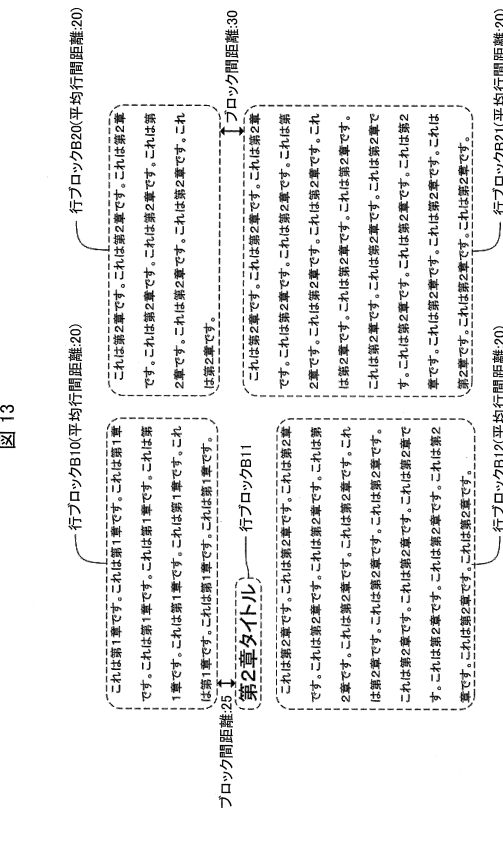
図 11



【図 1 2】

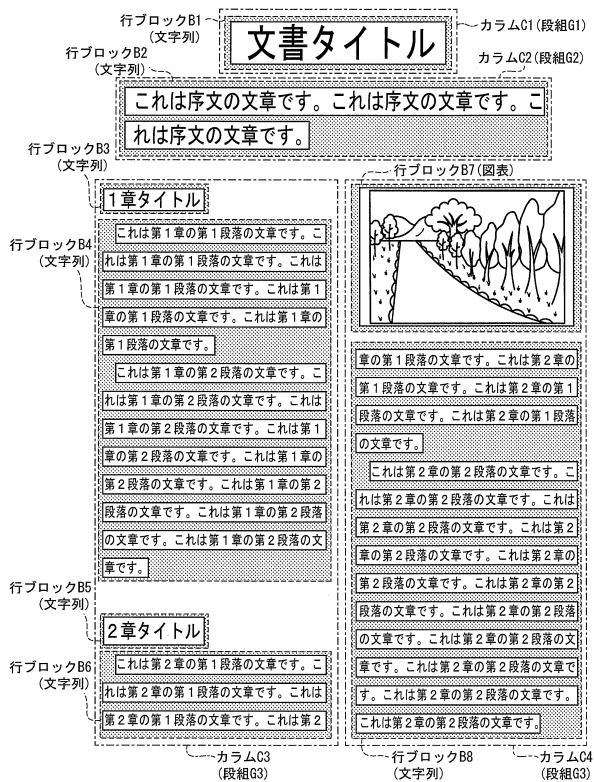


【図 1 3】



【図 1 4】

図 14



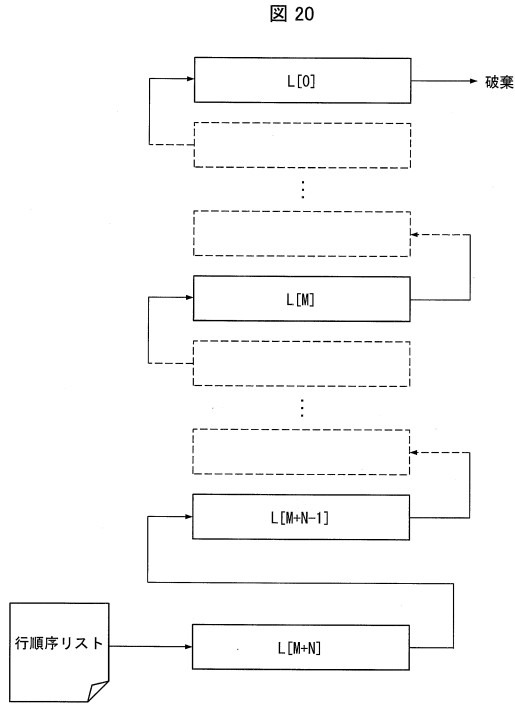
【図 1 5】

図 15

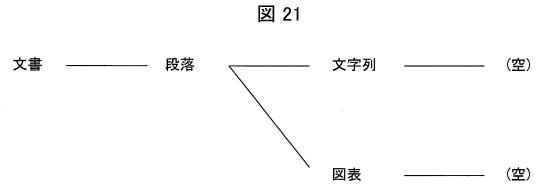
順位	行ID	ページ番号	行ブロックID	種別	左上座標, 右下座標	段落情報
1	101	1	B1	文字列	(800, 100), (1700, 240)	No
2	102	1	B2	文字列	(250, 350), (2260, 430)	No
3	103	1	B2	文字列	(250, 480), (1090, 560)	No
4	104	1	B3	文字列	.	No
5	106	1	B4	文字列	.	No
...	...	...	...	...	(以下省略)	...
9	110	1	B4	文字列	.	No
10	112	1	B4	文字列	.	No
11	114	1	B4	文字列	.	No
...	...	...	...	...	.	...
17	126	1	B4	文字列	.	No
18	129	1	B5	文字列	.	No
19	131	1	B6	文字列	.	No
20	133	1	B6	文字列	.	No
21	135	1	B6	文字列	.	No
22	105	1	B7	図表	.	No
23	111	1	B8	文字列	.	No
24	113	1	B8	文字列	.	No
25	115	1	B8	文字列	.	No
26	117	1	B8	文字列	.	No
27	119	1	B8	文字列	.	No
...	...	...	...	...	.	...
35	134	1	B8	文字列	.	No
36	136	1	B8	文字列	.	No



【 図 2 0 】



【 図 2 1 】

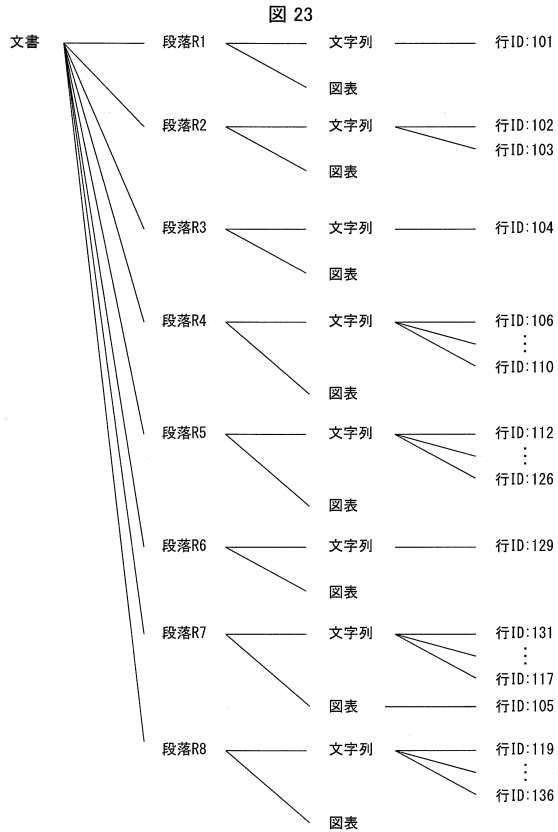


【 図 2 2 】

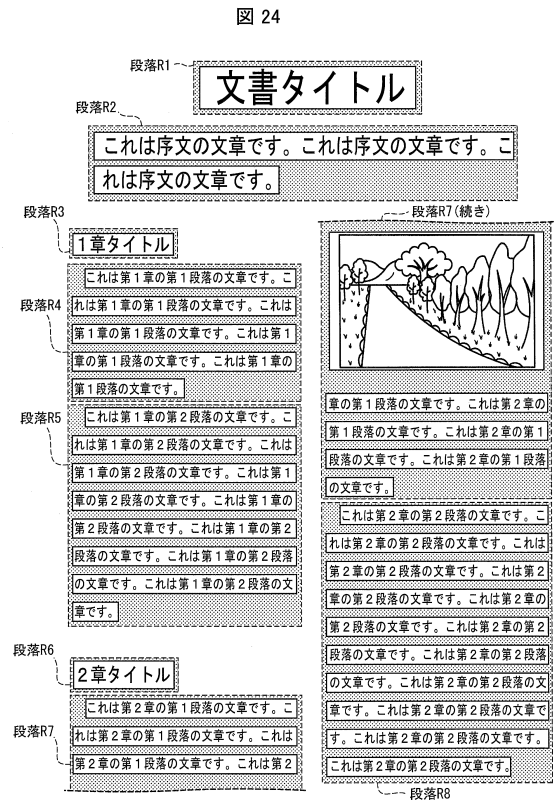
図 22

順位	行ID	ページ番号	行ブロック ID	種別	左上座標, 右下座標	段落情報
1	101	1	B1	文字列	(800, 100), (1700, 240)	Yes
2	102	1	B2	文字列	(250, 350), (2260, 430)	Yes
3	103	1	B2	文字列	(250, 480), (1090, 560)	No
4	104	1	B3	文字列	.	Yes
5	106	1	B4	文字列	.	Yes
...	...	...	...	...	(以下省略)	...
9	110	1	B4	文字列	.	No
10	112	1	B4	文字列	.	Yes
11	114	1	B4	文字列	.	No
...	...	...	...	...	.	...
17	126	1	B4	文字列	.	No
18	129	1	B5	文字列	.	Yes
19	131	1	B6	文字列	.	Yes
20	133	1	B6	文字列	.	No
21	135	1	B6	文字列	.	No
22	105	1	B7	図表	.	No
23	111	1	B8	文字列	.	No
24	113	1	B8	文字列	.	No
25	115	1	B8	文字列	.	No
26	117	1	B8	文字列	.	No
27	119	1	B8	文字列	.	Yes
...	...	...	...	...	.	...
35	134	1	B8	文字列	.	No
36	136	1	B8	文字列	.	No

【 図 2 3 】

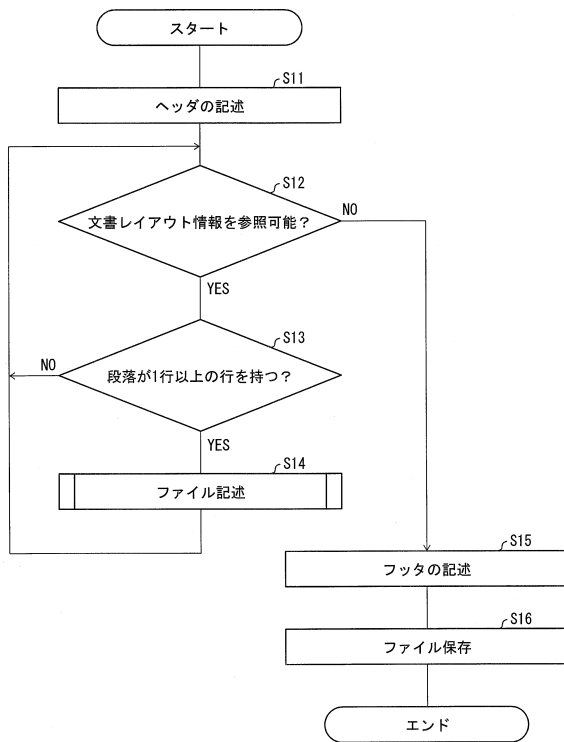


【 図 2 4 】



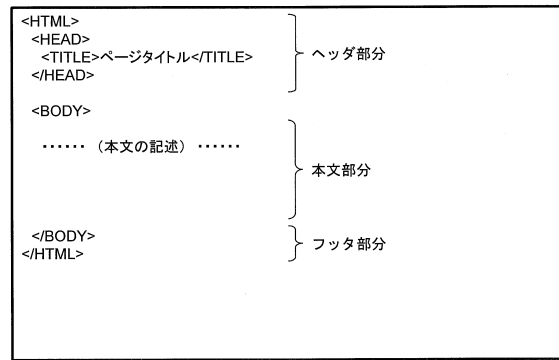
【図 25】

図 25



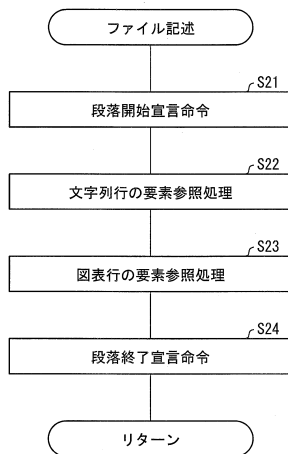
【図 26】

図 26



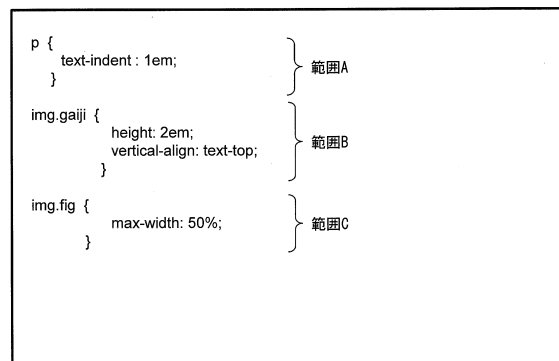
【図 27】

図 27

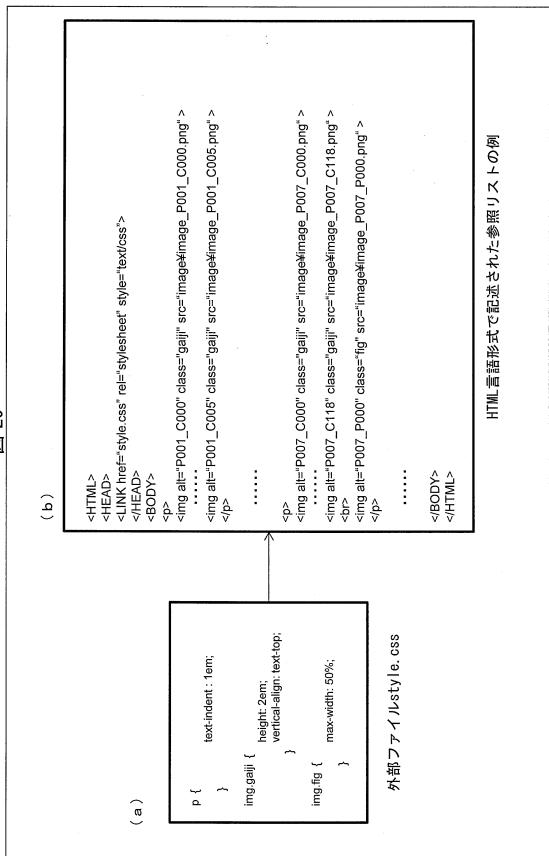


【図 28】

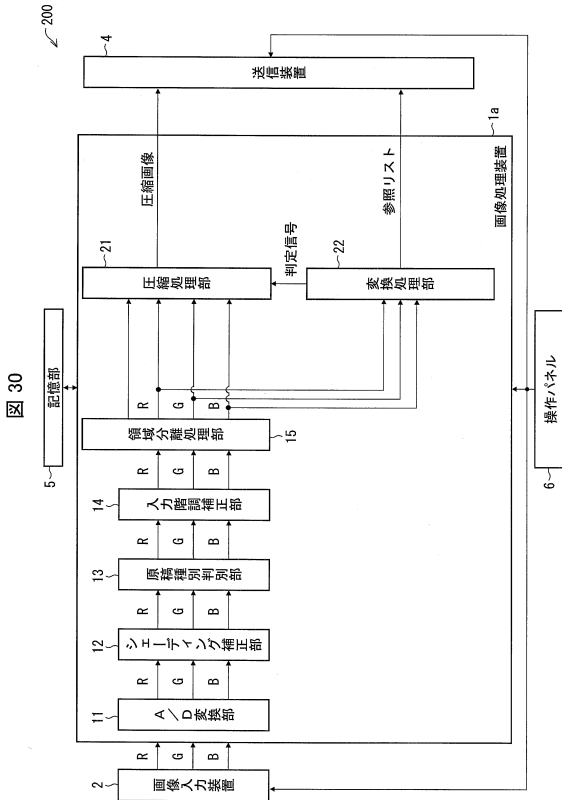
図 28



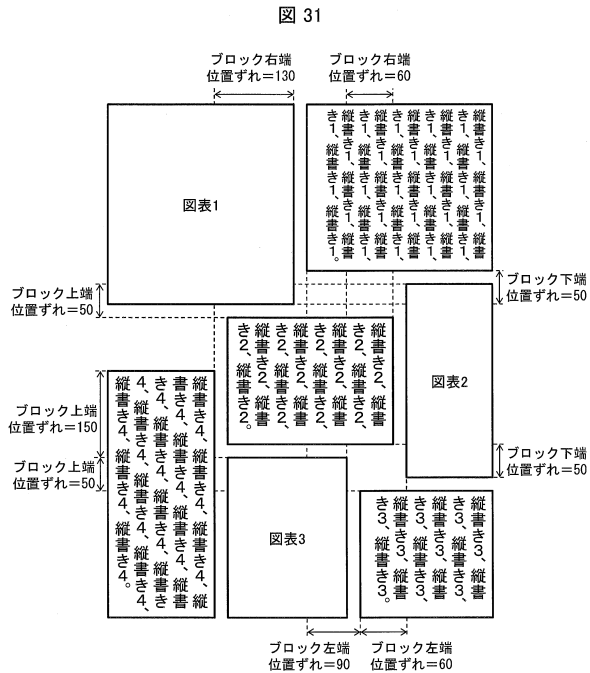
【図 29】



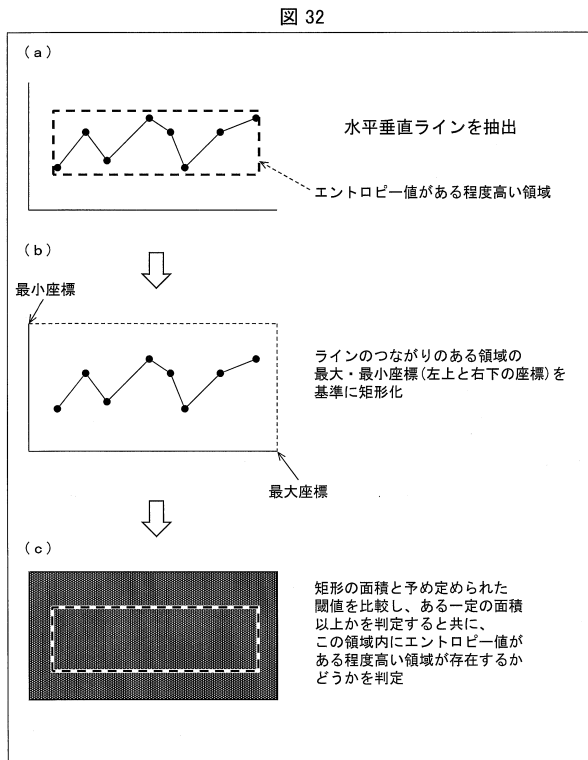
【図 30】



【図 31】



【図 32】



---

フロントページの続き

(72)発明者 濱田 和之  
大阪府大阪市阿倍野区长池町2番2号 シャープ株式会社内

審査官 成瀬 博之

(56)参考文献 米国特許出願公開第2013/0259377(US, A1)  
特開平05-328096(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 17/20 - 17/26

G06K 9/20

G06T 11/60

H04N 1/387