

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3950777号

(P3950777)

(45) 発行日 平成19年8月1日(2007.8.1)

(24) 登録日 平成19年4月27日(2007.4.27)

(51) Int. Cl.

G06T 7/40 (2006.01)

F I

G06T 7/40 100A

G06T 7/40 100Z

請求項の数 13 (全 18 頁)

(21) 出願番号	特願2002-285580 (P2002-285580)	(73) 特許権者	000001007
(22) 出願日	平成14年9月30日(2002.9.30)		キヤノン株式会社
(65) 公開番号	特開2004-126648 (P2004-126648A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成16年4月22日(2004.4.22)	(74) 代理人	100090538
審査請求日	平成17年6月15日(2005.6.15)		弁理士 西山 恵三
		(74) 代理人	100096965
			弁理士 内尾 裕一
		(72) 発明者	金津 知俊
			東京都大田区下丸子3丁目30番2号キヤ ノン株式会社内
		審査官	松尾 俊介

最終頁に続く

(54) 【発明の名称】 画像処理方法、画像処理装置および画像処理プログラム

(57) 【特許請求の範囲】

【請求項1】

第1の画素塊抽出手段が、二値画像から黒画素塊と白画素塊とを再帰的に抽出して記憶手段に記録するステップと、

ツリー構造作成手段が、前記第1の画素塊抽出手段で抽出されて前記記憶手段に記録された黒画素塊と白画素塊との位置関係を示すツリー構造データを作成するステップと、

反転画像作成手段が、前記ツリー構造データに含まれる黒画素塊のうち反転文字を含む黒画素塊の内部を白黒反転させて、反転画像を作成するステップと、

第2の画素塊抽出手段が、前記反転画像作成手段で作成された反転画像から、白画素塊と黒画素塊とを抽出するステップと、

前記ツリー構造追加手段が、前記第2の画素塊抽出手段で抽出された白画素塊と黒画素塊とに関するデータを、前記ツリー構造データの対応するノードに追加するステップとを有し、

前記反転画像作成手段では、前記ツリー構造データに含まれる黒画素塊のうち、黒画素密度が低い黒画素塊についての反転画像は作成しないことを特徴とする画像処理方法。

【請求項2】

前記ツリー構造追加手段で追加される白画素塊と黒画素塊とに関するデータは、当該抽出された黒画素塊に関するデータを子ノードとして有する白画素塊に関するデータを、前記ツリー構造データの当該白黒反転された黒画素塊の子ノードとして追加することを特徴とする請求項1に記載の画像処理方法。

## 【請求項 3】

前記反転画像作成手段では、前記第 1 の画素塊抽出手段で抽出した黒画素塊の輪郭は黒画素にしたまま、その内部の画素を白黒反転することによって前記反転画像を作成することを特徴とする請求項 1 に記載の画像処理方法。

## 【請求項 4】

更に、領域分割手段が、前記ツリー構造追加手段で白画素塊と黒画素塊とに関するデータが追加されたツリー構造データに基づいて、黒画素塊をグループ化することにより、前記二値画像を複数の領域に分割するステップを有することを特徴とする請求項 1 に記載の画像処理方法。

## 【請求項 5】

更に、領域ツリー構造作成手段が、前記ツリー構造追加手段で白画素塊と黒画素塊とに関するデータが追加されたツリー構造データに基づいて、黒画素塊をグループ化することにより、複数の領域の位置関係を示す領域ツリー構造データを作成するステップを有することを特徴とする請求項 1 に記載の画像処理方法。

## 【請求項 6】

前記黒画素塊とは黒画素の 8 方向輪郭追跡することによって得た黒画素輪郭で囲まれる領域であり、前記白画素塊とは白画素の 4 方向輪郭追跡することによって得た白画素輪郭で囲まれる領域であることを特徴とする請求項 1 に記載の画像処理方法。

## 【請求項 7】

前記二値画像とは、二値の原画像を OR 縮小することによって得た画像であり、  
前記反転画像とは、当該白黒反転させる対象の黒画素塊の領域と前記原画像を白黒反転させてから OR 縮小することによって得た画像とに基づいて作成されることを特徴とする請求項 1 に記載の画像処理方法。

## 【請求項 8】

前記第 2 の画素塊抽出手段で抽出される黒画素塊は、前記反転画像から抽出される複数の黒画素塊の中から、前記二値画像から抽出した白画素塊のうち所定以上の大きさを有する白画素塊に対応する位置にある前記反転画像から抽出した黒画素塊を除外し、該除外の後に残った黒画素塊であることを特徴とする請求項 1 に記載の画像処理方法。

## 【請求項 9】

前記第 2 の画素塊抽出手段では、前記反転画像から抽出される複数の黒画素塊のうち、ノイズに由来する黒画素塊の比率が所定値以上である場合、当該反転画像の作成元となった黒画素塊は反転文字を含まないと判断して処理を終了することを特徴とする請求項 1 に記載の画像処理方法。

## 【請求項 10】

前記二値の原画像は、多値の原画像を 2 値化することにより変換された画像であることを特徴とする請求項 7 に記載の画像処理方法。

## 【請求項 11】

二値画像から黒画素塊と白画素塊とを再帰的に抽出する第 1 の画素塊抽出手段と、  
前記第 1 の画素塊抽出手段で抽出された黒画素塊と白画素塊との位置関係を示すツリー構造データを作成するツリー構造作成手段と、  
前記ツリー構造データに含まれる黒画素塊のうち反転文字を含みうる黒画素塊の内部を白黒反転させて、反転画像を作成する反転画像作成手段と、  
前記反転画像作成手段で作成された反転画像から、白画素塊と黒画素塊とを抽出する第 2 の画素塊抽出手段と、  
前記第 2 の画素塊抽出手段で抽出された白画素塊と黒画素塊とに関するデータを、前記ツリー構造データの対応するノードに追加するツリー構造追加手段とを有し、  
前記反転画像作成手段では、前記ツリー構造データに含まれる黒画素塊のうち、黒画素密度が低い黒画素塊についての反転画像は作成しないことを特徴とする画像処理装置。

## 【請求項 12】

コンピュータを、

10

20

30

40

50

二値画像から黒画素塊と白画素塊とを再帰的に抽出して記憶手段に記録する第1の画素塊抽出手段、

前記第1の画素塊抽出手段で抽出されて前記記憶手段に記録された黒画素塊と白画素塊との位置関係を示すツリー構造データを作成するツリー構造作成手段、

前記ツリー構造作成手段で作成されたツリー構造データに含まれる黒画素塊のうち反転文字を含みうる黒画素塊の内部を白黒反転させて、反転画像を作成する反転画像作成手段

、

前記反転画像作成手段で作成された反転画像から、白画素塊と黒画素塊とを抽出する第2の画素塊抽出手段、

前記第2の画素塊抽出手段で抽出された白画素塊と黒画素塊とに関するデータを、前記ツリー構造データの対応するノードに追加するツリー構造追加手段、

として機能させるための画像処理プログラムであり、

前記反転画像作成手段では、前記ツリー構造データに含まれる黒画素塊のうち、黒画素密度が低い黒画素塊についての反転画像は作成しないことを特徴とする画像処理プログラム。

【請求項13】

請求項12に記載の画像処理プログラムを格納した、コンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は画像処理方法、装置、及びプログラムに関するものである。

【0002】

【従来の技術】

近年、情報の電子化が進み、紙文書と電子化文書を相互に変換する需要が高まっている。紙文書を電子化する際には、単にスキャナなどにより紙面を光電変換し、画像データ化することのみにとどまらず、記載されている内容に応じて、文書画像をテキスト、記号、図、写真、表などそれぞれ性質の異なる領域へと分割し、文字部は文字コード情報、図や線、表枠はベクトルデータ、写真は画像データ、表の内容は構造データ、といったように各領域に対し最適な形でデータ化することが望ましい。

【0003】

このように、紙文書の電子化処理においては、文書画像に書かれた内容を分析し、文字や図、写真、表など異なる性質の部分領域ごとに分割する処理、すなわち領域分割処理の重要性は高い。

【0004】

【発明が解決しようとする課題】

この領域分割処理の手法については、例えば、図21のような多値（グレースケール又はカラー）で読み取った文書画像を、輝度差による2値の画像に変換し、その中に存在する黒画素を輪郭に持つ画素塊をすべて抽出して、大きさなどから文字と非文字に分類し、さらに非文字の大きな黒画素塊内に存在する白画素領域の内部から再帰的に画素を探索することにより、画素塊の状況を図16のような階層的ツリー構造で表現することが提案されている。この画素塊のツリー構造に対して、同階層にある文字画素塊をグループ化して文字領域を得て、また、非文字画素塊の形状や周辺条件から図画や写真領域を得て、階層をなす画素の組として表領域を得るなどの処理で種々の属性を持つ領域へと分割することで、図22のような領域分割結果を得ることができる。また、このとき、各領域が図23のようなツリー構造を持つことで、文書の論理的構造の判断に適した情報となる。

【0005】

しかし、このような領域分割処理では処理の構成上、図21中に含まれる輝度反転文字部分、すなわち二値画像上で白地に黒ではなく、黒地上に白画素で構成されるような文字（反転文字、白抜き文字）の領域を抽出することは容易でなかった。また、黒白画素数を比

10

20

30

40

50

較して黒画素が多いと判断すれば画素を反転することにより反転文字を認識できるようにすることも考えられているが、通常文字と反転文字の関連性を得ることは難しく、その両方を含む文書から、その文書の通常文字と反転文字とを統合的に扱ったツリー構造を得ることはできなかった。

【 0 0 0 6 】

【課題を解決するための手段】

上記課題を解決するため、本発明の画像処理方法は、第1の画素塊抽出手段が、二値画像から黒画素塊と白画素塊とを再帰的に抽出して記憶手段に記録するステップと、ツリー構造作成手段が、前記第1の画素塊抽出手段で抽出されて前記記憶手段に記録された黒画素塊と白画素塊との位置関係を示すツリー構造データを作成するステップと、反転画像作成手段が、前記ツリー構造データに含まれる黒画素塊のうち反転文字を含みうる黒画素塊の内部を白黒反転させて、反転画像を作成するステップと、第2の画素塊抽出手段が、前記反転画像作成手段で作成された反転画像から、白画素塊と黒画素塊とを抽出するステップと、前記ツリー構造追加手段が、前記第2の画素塊抽出手段で抽出された白画素塊と黒画素塊とに関するデータを、前記ツリー構造データの対応するノードに追加するステップとを有し、前記反転画像作成手段では、前記ツリー構造データに含まれる黒画素塊のうち、黒画素密度が低い黒画素塊についての反転画像は作成しないことを特徴とする。

10

上記課題を解決するため、本発明の画像処理装置は、二値画像から黒画素塊と白画素塊とを再帰的に抽出する第1の画素塊抽出手段と、前記第1の画素塊抽出手段で抽出された黒画素塊と白画素塊との位置関係を示すツリー構造データを作成するツリー構造作成手段と、前記ツリー構造データに含まれる黒画素塊のうち反転文字を含みうる黒画素塊の内部を白黒反転させて、反転画像を作成する反転画像作成手段と、前記反転画像作成手段で作成された反転画像から、白画素塊と黒画素塊とを抽出する第2の画素塊抽出手段と、前記第2の画素塊抽出手段で抽出された白画素塊と黒画素塊とに関するデータを、前記ツリー構造データの対応するノードに追加するツリー構造追加手段とを有し、前記反転画像作成手段では、前記ツリー構造データに含まれる黒画素塊のうち、黒画素密度が低い黒画素塊についての反転画像は作成しないことを特徴とする。

20

上記課題を解決するため、本発明の画像処理プログラムは、コンピュータを、二値画像から黒画素塊と白画素塊とを再帰的に抽出して記憶手段に記録する第1の画素塊抽出手段、前記第1の画素塊抽出手段で抽出されて前記記憶手段に記録された黒画素塊と白画素塊との位置関係を示すツリー構造データを作成するツリー構造作成手段、前記ツリー構造作成手段で作成されたツリー構造データに含まれる黒画素塊のうち反転文字を含みうる黒画素塊の内部を白黒反転させて、反転画像を作成する反転画像作成手段、前記反転画像作成手段で作成された反転画像から、白画素塊と黒画素塊とを抽出する第2の画素塊抽出手段、前記第2の画素塊抽出手段で抽出された白画素塊と黒画素塊とに関するデータを、前記ツリー構造データの対応するノードに追加するツリー構造追加手段、して機能させるための画像処理プログラムであり、前記反転画像作成手段では、前記ツリー構造データに含まれる黒画素塊のうち、黒画素密度が低い黒画素塊についての反転画像は作成しないことを特徴とする。

30

【 0 0 0 7 】

40

【発明の実施の形態】

図1に本実施形態のブロック図を示す。

【 0 0 0 8 】

101は紙の文書が光電変換されて生成された画像データを入力する入力部、102は入力された画像データに対して二値化や縮小、ノイズ除去などの前処理を行う前処理部、103は画像データを文字、線、図、表などの属性毎の領域に分割する領域分割部である。なお、領域分割部103は、黒画素塊と白画素塊を抽出して画素塊のツリー構造データ(階層構造データ)を作成する画素塊抽出部1031、画素塊のツリー構造データに反転文字情報を追加する反転画素塊抽出部1032、画素塊のツリー構造データにおいて属性別に領域を分類する領域決定部1033からなる。104は画像から得られた領域分割を行

50

った結果の情報（領域のツリー構造データ）を出力する出力部である。

【0009】

図2に本実施形態を実現する装置構成の概観図を示す。スキャナ装置201は、入力部101の光電変換動作を実行して画像データを入力する。コンピュータ装置202は前処理102および領域分割処理103を実行し、キーボードやマウスなどの指示手段203によって、ユーザーからの動作制御を受ける。領域分割処理によって生成されたデータは、コンピュータに内蔵されたハードディスクなどの記憶媒体や、ディスプレイ204、プリンタ205、ネットワークを介して他の装置などへと出力される。

【0010】

なお、本実施形態を実行するコンピュータ装置は、実際の処理演算を行うCPU、プログラムを読み込んでワークエリアとして用いるRAM、後述するフローチャートに対応する処理を実行するためのプログラムや各種データを格納するための記憶媒体（ハードディスク、ROM、リムーバブルディスク（フロッピー（R）ディスク、CD-ROM等）など）、各種操作を行うためのキーボードやポインティングデバイス、処理対象の文書等を表示するためのディスプレイ、ネットワークと接続するためのネットワークインターフェースなどで構成される。CPUに実行させるための画像処理プログラムは、前記記憶媒体から供給されるものであってもよいし、ネットワークを介して外部装置から読み込むものであってもよい。なお、本実施形態ではCPUにおいてプログラムを実行することにより実現するものとするが、その一部又は全ての処理をハードウェア（電気回路）で構成するようにしても構わない。

【0011】

本実施形態で実行される画像処理手順を、図3を用いて説明する。

【0012】

ステップS301では、スキャナなどで紙文書を読み取って画像データを生成し、該画像データがコンピュータに入力される。

【0013】

ステップS302では、前処理部102において、画像データを後段の領域分割処理を実行するのに適した二値画像に変換する。具体的には、入力された画像データがカラーやグレースケールなどの多値画像の場合、適応的に閾値を設定して白黒の2値（なお、本実施形態では黒画素の画素値を1、白画素の画素値を0として扱う）に変換する二値化処理や、また孤立点などを除去するノイズ除去処理をおこなう。さらに、画像の領域分割処理を高速に行うため、入力された画像データのサイズを適正な画像サイズに変更する処理（文書画像の解像度変換）を行うようにしてもよい。例えば、画像解像度を1/2に縮小するなら2×2範囲（1/4の縮小なら4×4範囲）の画素がすべて0の場合は代表画素値0にし、それ以外の場合は画素値を1とするOR縮小によっておこなわれる。

【0014】

ステップS303では、領域分割処理103における画素塊抽出処理1031が実行され、二値画像から黒画素、白画素のかたまりを再帰的に抽出し、ツリー状の構造を作成する。この画素塊抽出処理について、図4のフローチャートを用いて詳細に説明する。

【0015】

ステップS401では、画像から黒画素の8連結輪郭塊の抽出を行う。黒画素塊を抽出できた場合はS402に進み、抽出できなかった場合はS408へ進む。なお、黒画素の8連結輪郭塊とは、縦横斜めのいずれかで接触した黒画素を検知して黒画素集合の輪郭を追跡することによって抽出した黒画素の集合（領域）のことである。以降、この集合を黒画素塊と呼ぶ。なお、この黒画素塊は、輪郭以外の画素の白黒は問わず、内部に白画素の空洞を有しても構わない。なお、この黒画素の輪郭追跡方法については、公知の方法を用いることが可能であり、以下に、図5を用いてこの黒画素塊抽出処理の概要を簡単に説明する。

【0016】

ステップS501では、（白画素値0、黒画素値1で構成される）二値画像を左上から順

10

20

30

40

50

にラインスキャンし、画素値が1となる点（黒画素）を探す。例えば、図8の801の矢印で示したような順番で、ライン順に走査して黒画素を探す。

【0017】

ステップS502では、黒画素が見つかったかどうか判断し、黒画素が見つかったときは、当該黒画素を開始点ならびに注目点QとしてステップS503に進んで輪郭抽出の処理を開始する。一方、S502で見つからなかった場合は終了する。例えば、図8では画素810が開始点で且つ最初の注目点Qとなる。

【0018】

ステップS503では、直前の追跡方向d（直前の画素から注目点Qに追跡してきた方向）と周辺画素状態とに基づき、図11の表に記載した順序で周囲の画素の検査を行うことにより、注目点Qから次の輪郭画素への追跡方向d'を定める。ここで追跡方向は、図10に示したように、N、NE、E、SE、S、SW、W、NWの8方向で表すものとする。また例外処理として、最初の注目点Q（開始点）では、直前の追跡方向dはSEであったものと定義する。図11に示した表は、直前の追跡方向dに対して注目点Qの周囲の画素を検査する順序nを示している。例えば、直前の追跡方向dがSWであった場合は、注目点Qから見て、NW方向の画素から順に反時計回りに周囲の画素を検査していき（NW方向画素、W方向画素、SW方向画素・・・の順に検査）、黒画素が見つかった時点で周囲の画素の検査を終了して、その黒画素が見つかった方向を次の追跡方向d'とする。図11に示したような順序で検査することにより、次の追跡方向d'を定める際には、直前の追跡方向dを定める際に検査した画素の検査を行わないようにしている。

【0019】

ステップS504では、注目点Qの画素に輪郭画素であることを示すラベル付けをおこなう。ラベルには、“A”、“B”、“C”、“D”の4種類があり、輪郭で囲まれる領域において、画素行の左端のエッジとなる画素にはラベルAが付与され、右端のエッジとなる画素にはラベルBが付与され、左右両方のエッジとなっている画素にはラベルCが付与され、左右のどちらのエッジでもない輪郭画素にはラベルDが付与されるものとする。このラベル値は、図11と図12を用いて、直前の追跡方向dと次の追跡方向d'と現在のQのラベル値から決定される。

【0020】

ステップS505では、次の追跡方向d'にある画素を、新たな注目点Qとし、追跡方向d'を直前の追跡方向dに代入する。

【0021】

ステップS506において、新たな注目点Qが開始点に等しいかどうか判断し、等しい場合はステップS507に進み、等しくなければステップS503へ戻って追跡処理を繰り返す。

【0022】

ステップS503～506で実行される処理について、図8の801を例にとって示す。まず、開始点810を最初の注目点Qとし、図11に記載された順に、方向SWから検査を始める。ここで最初の注目点Q（810）からは、次の追跡方向d' = SW（n = 1）において黒画素が見つかり、図12を参照して、注目点Q（画素810）にラベル“A”が付与された後、注目点Qは左下に移動する。以下同様に繰り返すと図8の802のようなラベル付けがされた輪郭が得られる。

【0023】

ステップS507では、注目点Q（開始点）に他の分岐が存在するかどうかを調べる。他の分岐が存在するかどうかの判断は直前の追跡方向dがNEの場合に行われ、直前の追跡方向がNE以外の場合は他の分岐は存在しないと判断する。直前の追跡方向dがNEの場合、当該注目点Qの周囲の画素について、Qから見てSE方向の画素、E方向の画素の順で黒画素が存在するか検査を行い、黒画素が存在した時点で他の分岐が存在すると判断し、その存在した方向d'に注目点Qを移動させるとともに追跡方向d'を直前の追跡方向dに代入してステップS503に戻る。一方、他の分岐が存在しない場合はステップS5

10

20

30

40

50

09に進む。

【0024】

ステップS509では、輪郭ラベル付けされた黒画素にかこまれる画素の塊を1個の黒画素塊として記録する。具体的には、図8の803のように、各y座標（各画素行）毎に、左エッジラベル“A”と右エッジラベル“B”の組に挟まれた連続画素、あるいは単独の“C”画素を1セグメントとし、それらセグメントの集合として画素塊を記録する。なお、図8の803では、説明を容易にするために、x座標とy座標は黒画素塊に外接する矩形の左上の点を原点にしているが、好適には、x座標とy座標はS302で縮小した画像の左上を原点にして表すものとする。

【0025】

なお、1つの画素塊を抽出終了後（S509での処理終了後）、該抽出した画素塊にS402～S407による属性分類が行なわれた後、S501に戻り、前回の開始点の右隣りの画素からラインスキャンを再開して次の開始点を探索する。ただし、その時点までで得られている黒画素塊の輪郭内部の画素については探索をスキップする、すなわち既出の黒画素塊が持つセグメントと重なる画素を無視して、画素値＝“1”の画素を探す。

【0026】

図4の説明に戻ると、S401での処理（図5の処理）で抽出された全ての黒画素塊に対して、各黒画素塊の形状と各黒画素塊に外接する外接矩形の形状とを用いて、S402以降の処理において属性を分類する。

【0027】

ステップS402では、黒画素塊の外接矩形サイズが、予め予想される最大文字高さおよび幅に対し定められた閾値以下の大きさであった場合、文字要素と判定する。この黒画素塊には“CHAR”という属性を付加する。

【0028】

ステップS403では、黒画素塊の外接矩形サイズが、所定の比率以上で縦長あるいは横長であった場合、“LINE”の属性を与える。

【0029】

ステップS404では、黒画素塊中の黒画素のなす輪郭に注目し、その形状が細い斜めの線状であると判定した場合、“LINE”の属性を与える。

【0030】

ステップS405では、“CHAR”、“LINE”以外の黒画素塊の内部に存在する白画素の4連結輪郭塊を抽出する。白画素の4連結輪郭塊とは、縦横に連結した白画素の輪郭で囲まれる画素集合のことである。以降この集合を白画素塊と呼ぶ。

【0031】

白画素塊を抽出する方法は、図5で説明した黒画素の輪郭抽出処理において、“0”と“1”を逆転し、且つ、斜め方向の連結は考慮せずに、白画素は縦又は横（N、E、S、W）の4方向の連結のみを許可して輪郭を抽出する。さもないと白画素の輪郭追跡中に、注目点がベースとなる黒画素塊の外部に出てしまうからである。この制限のため、図5のS503に相当する処理においては、図11、12のテーブルにおいて、追跡方向が斜め方向の判断は行わずに同様の処理を行う。なお、図11、12のテーブルの代わりに白画素輪郭追跡用のテーブルを用意しても良く、この場合は追跡方向が4方向しか存在しないので非常に小さなテーブルとなる。

【0032】

ステップS406では、黒画素の輪郭形状がほぼ四角形かどうかを調べ、ほぼ四角形ならばステップS407に進む。四角形でなければ、黒画素塊を“NONCHAR”と判定する。四角形の黒画素塊および四角形以外の黒画素塊の例を図13に示す。

【0033】

ステップS407では、ほぼ四角形と判断された黒画素塊の内部から抽出される白画素塊の形状全てがほぼ長方形であり、且つそれらが黒画素塊の内部をほぼ隙間なしに占めている場合（白画素塊の外接矩形同士が重なることなく整然と並んでいる場合）、白画素塊の

10

20

30

40

50

整列が良、とする。そして、内部白画素塊の整列が良い黒画素塊には、“TABLE”の属性を与え、一方、整列が悪い黒画素塊は“NONCHAR”の属性を与える。図14に、内部白画素塊の整列の例を示す。図14(a)および(b)は内部白画素塊の整列が良いとして“TABLE”と分類される黒画素塊の例であり、図14(c)は内部白画素塊の整列が悪いとして“NONCHAR”と分類される黒画素塊の例である。

#### 【0034】

ステップS408では、“NONCHAR”あるいは“TABLE”と分類された黒画素塊内部にある白画素塊に注目し、さらにその白画素塊内部を対象にして、S401と同様に黒画素塊の抽出処理を行ってS402～S407と同様の分類処理をおこなう。

#### 【0035】

このS401～S408の処理によって、画像内の黒画素塊、ならびに黒画素塊中の白画素塊、さらに、“TABLE”や“NONCHAR”内部の白画素塊から再帰的に黒画素塊が抽出される。

#### 【0036】

図4のような処理を行って得た黒画素塊および白画素塊に対して、各画素塊内部に存在する画素塊を子ノードとして、ツリー構造を作成する。このときツリー構造において、黒画素塊についてはS402～S407で分類された属性をつけて表し、白画素塊については“WHITE”の属性で表すものとする。例えば、図15に対して画素塊抽出をおこなうと、図16のようなツリー構造を持ったデータが得られる。なお、図16では図を簡略化するため、画素塊の個数などは一部省略してあるが、実際には多数の画素塊によるノードが存在するものとする。図16のツリーにおいて、“WHITE”ノードに相当するのは前記白画素塊であり、また、便宜上、ツリーのルートには画像全体に相当する“WHITE”の属性を与えてある。つまり、図16のようなツリー構造において、“WHITE”属性を後景と考え、その他の“CHAR”“TABLE”“LINE”などの属性を前景と考えると、図16のツリー構造において、親ノードと子ノードは、後景と前景が交互に出現することになる。なお、ツリー構造の各ノードには、画素塊の領域情報(セグメント情報)と属性とが含まれているものとする。

#### 【0037】

以上説明したように画素塊のツリー構造を得た後、図3のステップS304において、領域分割部103の追加処理として、画素塊ツリー構造中の“NONCHAR”あるいは“TABLE”の属性が付与された黒画素塊に注目し、この黒画素塊内において黒下地上にある反転文字の抽出を目的とする画素塊抽出をおこなう。以下、具体的な処理例について、図6を用いて説明する。

#### 【0038】

ステップS601では、注目する黒画素塊(黒画素塊Aとする)の形状的な特徴から、黒画素塊A内に反転文字が存在する可能性を類推する。ここでは、黒画素密度が著しく低い場合、すなわち線の骨組のような黒画素塊については、反転文字を含まないと判断する。なお、黒画素密度とは、黒画素塊の全セグメント(図8の803で示されるような黒画素塊の領域)の総画素数をPとしたとき、{(黒画素塊の全セグメントにおいて画素値が1の数)/P}で求められる値である。このような黒画素密度などの容易に計算できる処理を用いて、反転文字が存在する可能性が著しく低いと判断できる場合は、S602～S608のような複雑な解析処理をスキップできるので、処理を速く行うことができる。

#### 【0039】

ステップS602では、黒画素塊A内部の画素の画素値(0と1)を反転した画像Rを作成する。このとき、黒画素塊Aの輪郭をなす画素については、反転せずに画素値を1のままとする。図17に画像R作成の例を示す。

#### 【0040】

なお、S302において縮小を施した画像を用いてS303の領域分割処理(画素塊抽出)を行っていた場合、ステップS602においては、縮小前の元画像から当該注目している黒画素塊の領域に対応する領域を抽出し、該抽出した領域に対して画素値(0と1)の

10

20

30

40

50



反転を行い（ただし輪郭の画素は反転しない）、その反転が行なわれた抽出領域に対して O R 縮小処理を施して得られる画像を、反転画像 R とする。このようにしないと、領域分割処理で O R 縮小が行なわれた画像においては反転文字部分がつぶれてしまっている可能性が高いためである。このように、縮小前の画像を反転してから縮小するので、反転文字部分がつぶれたり、かすれたりすることを防ぐことができる。

#### 【 0 0 4 1 】

ステップ S 6 0 3 では、反転画像 R の内部に対して、S 4 0 5 と同様の処理を行って白画素の 4 連結輪郭塊（白画素塊集合 B）を抽出する。

#### 【 0 0 4 2 】

ステップ S 6 0 4 では、ステップ S 6 0 3 で抽出された白画素塊集合 B の内部から、黒画素塊（黒画素の 8 連結輪郭塊）を抽出する。抽出された黒画素塊集合を C とする。

10

#### 【 0 0 4 3 】

ステップ S 6 0 5 では、反転前の黒画素塊 A 内部の白画素塊集合を抽出し、その反転前黒画素塊 A 内部の白画素塊集合のうち、所定以上の大きさを持つ白画素塊については反転文字（白抜き文字）ではないと判断し、ステップ S 6 0 4 で得た黒画素塊集合 C と比較して、該所定以上の大きさを持つ白画素塊と座標上で重なる黒画素塊を集合 C から取り除く。このように反転前の黒画素塊 A 内部から白画素塊を抽出して判断するので、反転文字ではなく背景部分であるとして予め容易に判断できる。また、元画像をそのまま O R 縮小した画像内の黒画素塊 A 内においては、白画素塊が潰れることはあったとしても、元画像で分離していた白画素塊同士が 1 つの白画素塊として結合してしまうことはないので、背景と判断される部分を正確に取り出すことができる（一方、元画像を反転してから縮小した画像において黒画素塊を抽出した場合、元画像で分離していた白画素塊同士が結合して黒画素塊となっている可能性があるため、元画像の白画素背景部分と反転文字部分が近いと結合してしまう可能性があり、その場合反転文字も除外されてしまう可能性があるため、S 6 0 5 では反転前の黒画素塊 A から白画素塊を抽出している）。

20

#### 【 0 0 4 4 】

ステップ S 6 0 2 ~ S 6 0 5 で行なわれる処理の例を図 9 に示す。元画像 9 0 1 に対し、S 3 0 2 で O R 縮小された画像から S 3 0 3 の領域分割処理で抽出された黒画素塊が 9 1 0 であった場合、この黒画素塊に対応する領域の元画像を反転・縮小して作成した反転画像が 9 2 0 である。黒画素塊 9 1 0 の内部にある白画素塊の領域 9 1 2（反転画像 9 2 0 上では白画素塊 9 2 1 内部の黒画素塊 9 2 2 となって抽出される）は、明らかに反転文字でないと判断される場合（例えば、予め定めておいた所定サイズより大きいとき反転文字ではないと判断される）、この黒画素塊 9 2 2 は白画素塊 9 1 2 と位置的に重なるので除外される処理が行われる。この結果、反転文字部分に相当する黒画素塊集合 9 2 3 は残る。

30

#### 【 0 0 4 5 】

ステップ S 6 0 6 では、集合 C 中の黒画素塊を S 4 0 2 と同等の判定規準（所定の閾値以下かどうか）で“CHAR”とそれ以外に分類する。

#### 【 0 0 4 6 】

ステップ S 6 0 7 では、“CHAR”黒画素塊を大きさが微少なものとそうでないものに分ける。個数をそれぞれ N, M とする。

40

#### 【 0 0 4 7 】

ステップでは、N をノイズに由来する画素塊の個数とみなし、それ以外の画素塊個数 M との比較によって、画素塊が文字集合であるか否かを判断する。ここでは、 $M = 0$  または  $N / M$  が所定の比率 T 以上の場合、画素塊は文字でないとして S 6 1 0 に進む。それ以外の場合は S 6 0 9 にすすむ。

#### 【 0 0 4 8 】

ステップ S 6 0 9 では、“CHAR”黒画素塊を反転文字の画素塊として、白画素塊を親ノードとするツリー構造を作成し、元の黒画素塊 A の直下に追加するよう画素塊ツリーを更新する。なお、親ノードとした白画素塊は、白画素塊 B をそのまま使ってもよいし、反

50

転文字の画素塊に外接するような範囲を白画素塊として定義してもよい。

【0049】

このように白画素塊を親ノードとすることにより、親ノードと子ノードの関係において後景と前景が交互に出現するというツリー構造の特性を保ったまま、反転文字をツリー構造の前景として追加することができる。

【0050】

図15の例において、1501, 1502, 1503の3つの黒画素塊を対象に反転文字抽出判断処理をおこなって、ノイズ比などから判断して1501と1503から反転文字の画素塊が抽出された場合、図16のツリーは図18のように更新される。このとき挿入される仮想的な白画素塊は、反転文字の下地であるという情報が含まれている以外は通常の“WHITE”と同等に扱われる。本実施形態では、“WHITE(R)”と書くことにする。

10

【0051】

図3に戻り、ステップS305では、領域決定部1033が、画素塊のツリー構造およびその分類結果を利用し、画像を文字/画像/表/線など矩形状の領域へと分割する。なお、ステップS303で抽出された画素塊とステップS304で追加された画素塊は、ツリー構造の特性を保っているので、区別なく同様の処理で分割することができる。この画素塊の分類、領域化の処理を、図7を用いて説明する。

【0052】

ステップS701では、“CHAR”と分類された画素塊に注目し、縦あるいは横に一定距離内にあるものをグループ化していく。このグループを囲む矩形を各々文字領域とする。さらに、文字領域内で文字列が横方向であるか縦方向であるかを調べる。たとえば、領域内の各画素塊から左右の最近傍の画素塊への水平距離、および上下最近傍の画素塊への垂直距離を求め、それらの平均値が小さいほうが文字列の方向とすればよい。

20

【0053】

ステップS702では、“NONCHAR”画素塊のうち、同程度の大きさで縦または横に連なっている集合を検出し、これをグループ化してタイトル文字領域とする。

【0054】

ステップS703では、“NONCHAR”画素塊のうち、輪郭内の黒画素と白画素の比が小さい、すなわち黒画素密度の低いものを抽出し、これを線画領域とする。

30

【0055】

ステップS704では、“NONCHAR”画素塊のうち、黒密度が高く大きな画素塊、あるいはある領域に群をなしている画素塊をグループ化したものをハーフトーン領域とする。ハーフトーン領域とは、写真など中間調の領域である。ハーフトーン領域に“CHAR”や“LINE”の画素塊が含まれる場合は、各々の元の領域を破棄してハーフトーン領域に統合する。

【0056】

ステップS705では、“LINE”画素塊を囲む矩形を、線領域とする。

【0057】

ステップS706では、“TABLE”領域を含む矩形を、表領域とする。

40

【0058】

以上の処理を、すべての黒画素塊に対しておこなう。ただしグループ化の対象は1つの“WHITE”画素塊の内部に存在する黒画素塊集合ごととする。

【0059】

例えば、図21の画像から図18のような画素塊のツリー構造が得られたとすると、更にS701～S706の処理を行って、図19のような属性毎の領域が決定される。このとき、画素塊ツリーの構成を元にして、領域間には図20の領域ツリー構造が作成される。この領域分割処理の出力結果においては、各領域の座標情報とともに、自分の親となる領域と、子供が存在すればそれらの領域へのリンク情報を保持する。

【0060】

50

なお、図 19 において、斜線塗りのテキスト領域は、反転文字として抽出されたテキスト領域である。これは出力結果からも図 20 のツリー構造において、WHITE ( R ) の下にあるテキストであることから判別できる。しかしながら、S 7 0 1 ~ S 7 0 6 の処理は反転領域を区別することなく行われるので、判定アルゴリズムなどは非反転領域・反転領域のどちらにも共通にすることができる。

【 0 0 6 1 】

【発明の効果】

以上説明したように、反転文字（白抜き文字）も、通常の文字と同じ階層ツリー構造で管理することができる。

【 0 0 6 2 】

また、紙をスキャンして得た画像を、文字や図、写真、表など異なる性質の部分要素に分割する領域分割処理を行う際に、白地に黒の文字領域と同様の領域抽出手法を利用して、黒地に白の文字領域も抽出することができる。

【図面の簡単な説明】

【図 1】第 1 の実施形態の構成を示すブロック図である

【図 2】第 1 の実施形態を実現する装置構成例の図である

【図 3】第 1 の実施形態の領域分割処理を説明するためのフローチャートである

【図 4】画素塊抽出処理を説明するためのフローチャートである

【図 5】黒画素の輪郭追跡処理を説明するためのフローチャートである

【図 6】反転文字抽出処理を説明するためのフローチャートである

【図 7】領域分類処理を説明するためのフローチャートである

【図 8】輪郭追跡による画素塊抽出処理の例である

【図 9】反転文字抽出処理の例である

【図 10】8 つの追跡方向を示す図である

【図 11】追跡方向を決定するための表である

【図 12】注目画素に対し付与するラベルを得るための表である

【図 13】四角形の黒画素塊とそれ以外の黒画素塊の例を示す図である

【図 14】黒画素塊内部の白画素塊の整列状態例を示す図である

【図 15】領域分割処理の入力となる二値画像の例である

【図 16】画素塊のツリー構造の例を示す図である

【図 17】反転文字抽出の為に作成される画像の例である

【図 18】反転文字を付与した画素塊のツリー構造の例を示す図である

【図 19】反転文字部を含む領域分割処理結果の例である

【図 20】反転文字部を含む領域のツリー構造の例である

【図 21】領域分割をおこなう文書原稿の例である

【図 22】従来の領域分割結果の例である

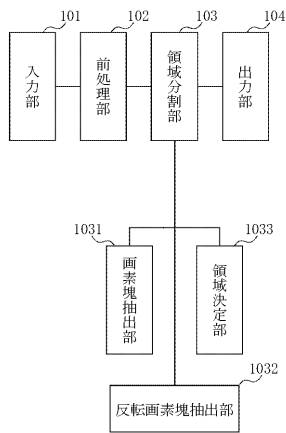
【図 23】従来の領域ツリー構造の例である

10

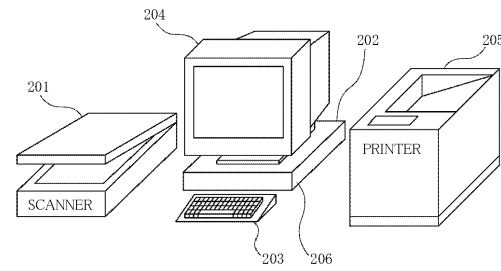
20

30

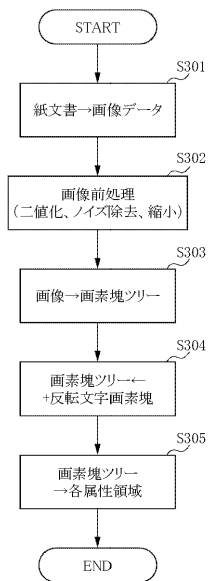
【図 1】



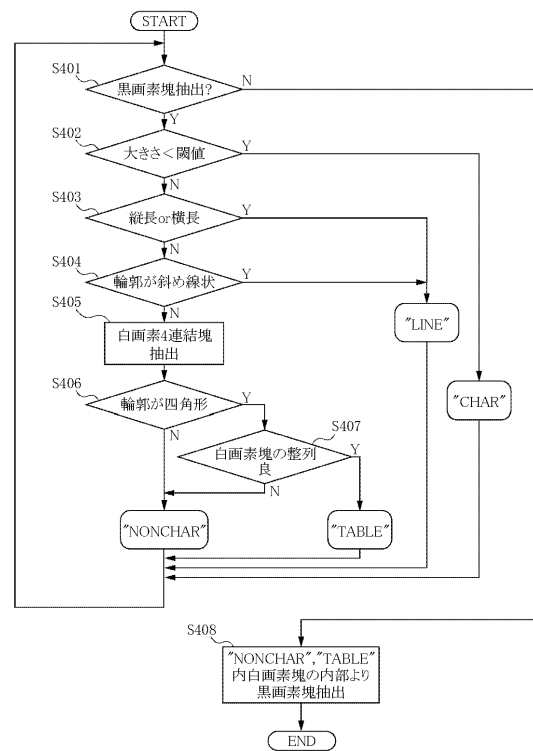
【図 2】



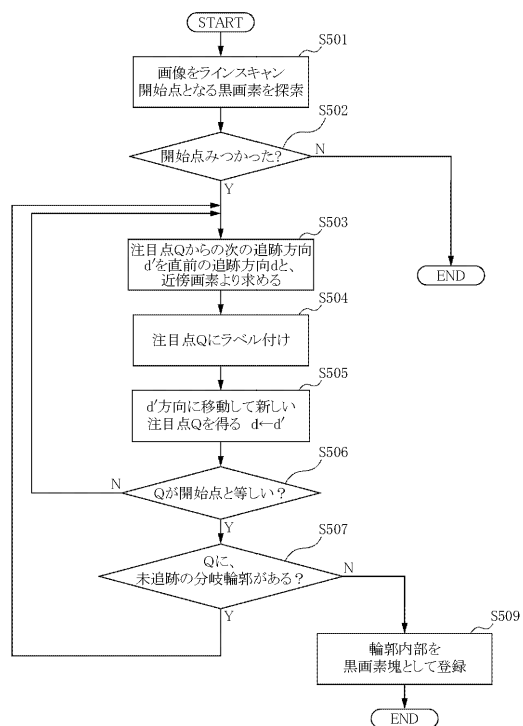
【図 3】



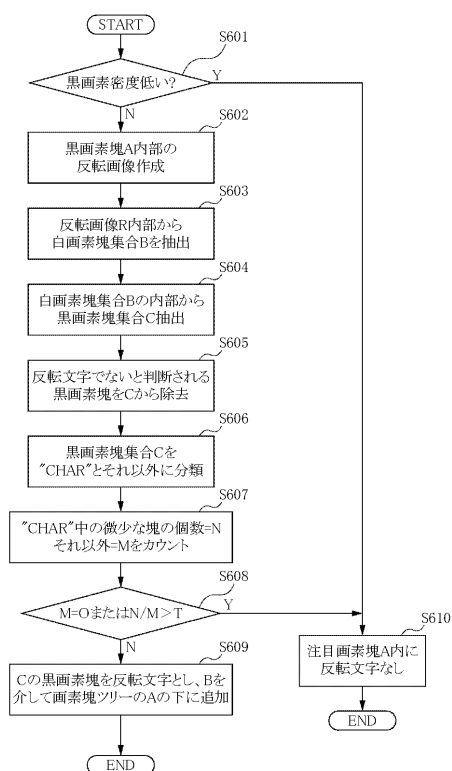
【図 4】



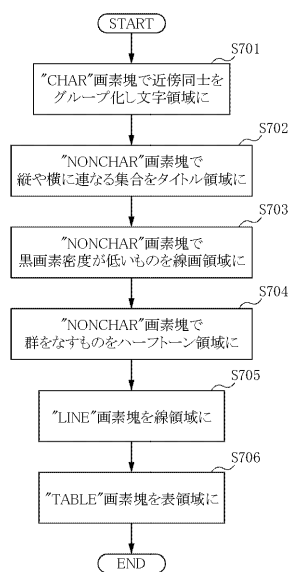
【圖 5】



【图 6】



【 図 7 】



【 図 8 】

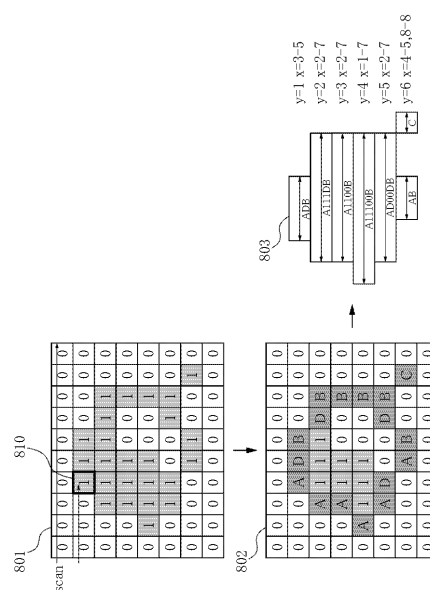
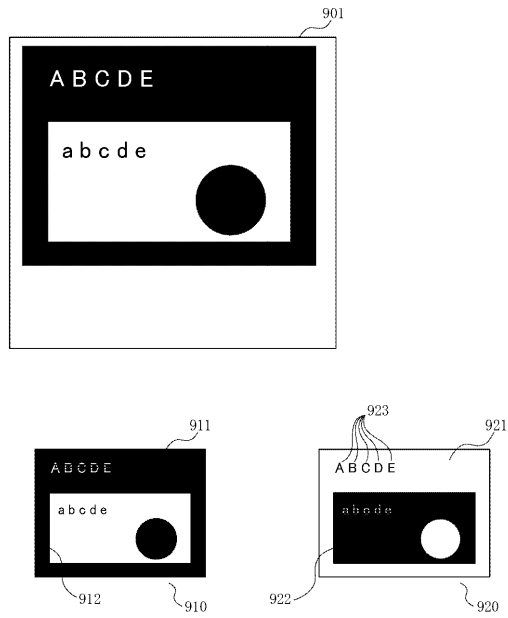


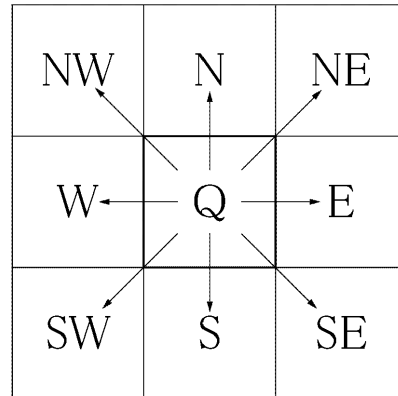
Diagram illustrating memory access patterns for different values of  $x$ :

- $y=1 \quad x=3-5$
- $y=2 \quad x=2-7$
- $y=3 \quad x=2-7$
- $y=4 \quad x=1-7$
- $y=5 \quad x=2-7$
- $y=6 \quad x=4-5, 8-8$

【図 9】



【図 10】



【図 11】

d \ n	1	2	3	4	5	6	7
NW	NE	N	NW	W	SW	S	SE
W	NW	W	SW	S	SE	E	
SW	NW	W	SW	S	SE	E	NE
S	SW	S	SE	E	NE	N	
SE	SW	S	SE	E	NE	N	NW
E	SE	E	NE	N	NW	W	
NE	SE	E	NE	N	NW	W	SW
N	NE	N	NW	W	SW	S	

【図 12】

d \ n	1	2	3	4	5	6	7
NW	*1	*1	*1	*1	*2	*2	*2
W	*4	*4	*3	*3	*3	*3	
SW	*4	*4	*3	*3	*3	*3	*2
S	*3	*3	*3	*3	*2	*2	
SE	*3	*3	*3	*3	*2	*2	*2
E	*4	*4	*1	*1	*1	*1	
NE	*4	*4	*1	*1	*1	*1	*2
N	*1	*1	*1	*1	*2	*2	

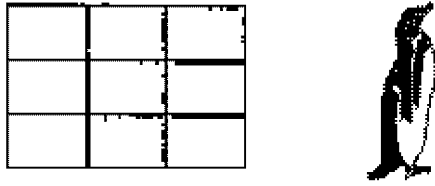
\*1:既にラベルAが付与されていた場合はラベルCにし、  
それ以外の場合はラベルBにする

\*2:ラベルCを付与する

\*3:既にラベルBが付与されていた場合はラベルCにし、  
それ以外の場合はラベルAにする

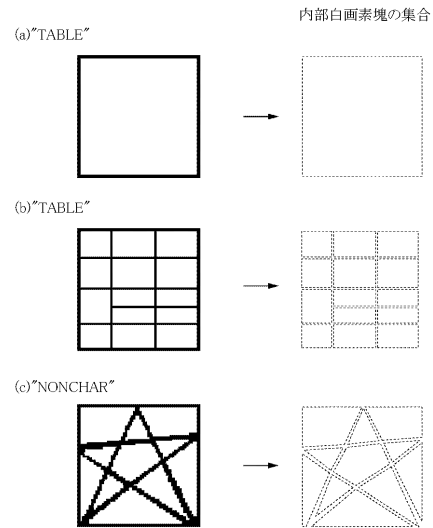
\*4:ラベルDを付与する

【 図 1 3 】



四角形黒画素塊(左)と非四角形黒画素塊(右)の例

【 図 1 4 】



【 図 1 5 】

**NAS 2001**

The Front Line of the Internet

**PREVIEW OF NAS 2001**

"NAS 2001" will be held on Japan convention Center in Chiba. The exhibition displays hardware, software and services of network business.

**IT Zone**  
Data warehouse, ERP, groupware, map information system, network administration tools. The highlight is B2B enterprise zone.

**Web Zone**  
Miscellaneous web servers, Application service providers, Content creation software.

**Network Zone**  
LAN/WAN router hub/network software, mobile computer, FAX server, Gateway, Firewall and virus scan are displayed.

**PC Zone**  
New machines of saved space desktop computer NEC2861, hand-held optical character recognizer TOM81. Super server H10000 can proceed 300T transactions per second.

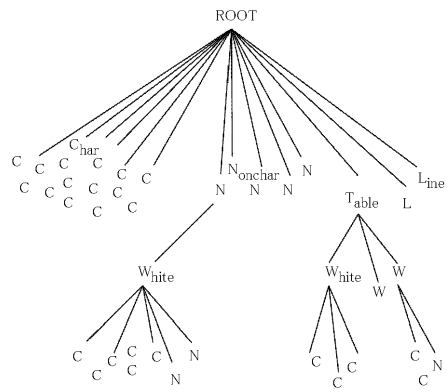
Tokyo-Nippori  
Keisei-Tsudanuma  
Makuhari-hongu  
Terijou-mae

Tokyo	
Shinjuku	
Babukuro	

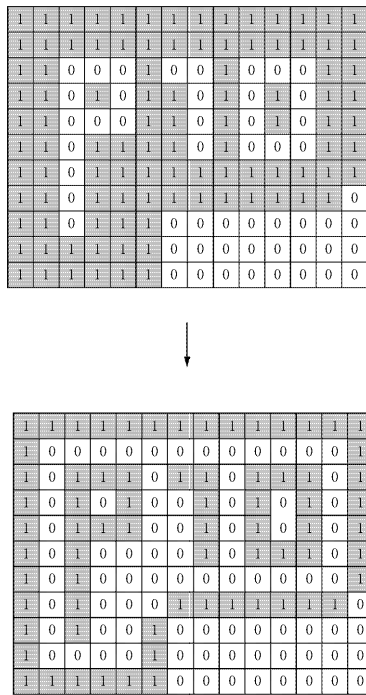
Printer can print at 60 pages per minute, the printer features a standard internal 2.1GB harddisk, paper capacity of 4,000 sheets and a remote display. With its high speed printing, high image quality, extensive document finishing capabilities and extensive integration, the printer will enable you to produce your everyday print applications with ease.

322 The CompNet magazine Jan. 2001

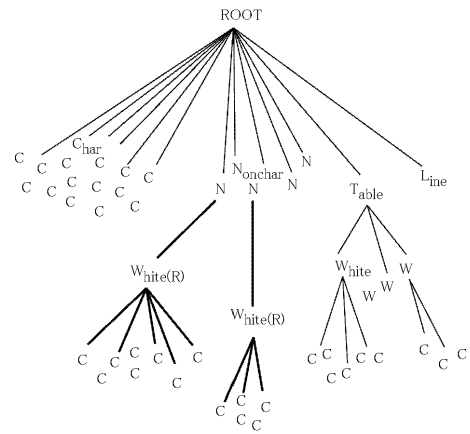
【 図 1 6 】



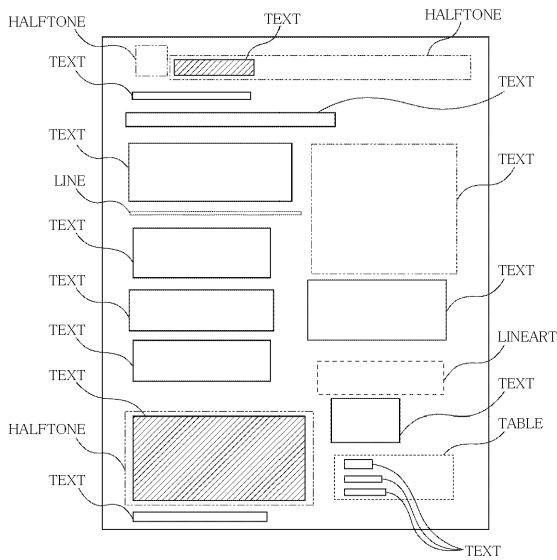
【 図 1 7 】



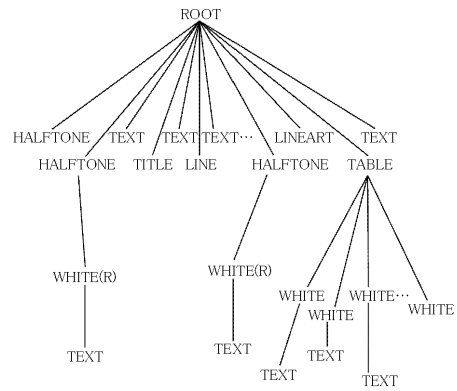
【 図 1 8 】



【 図 1 9 】

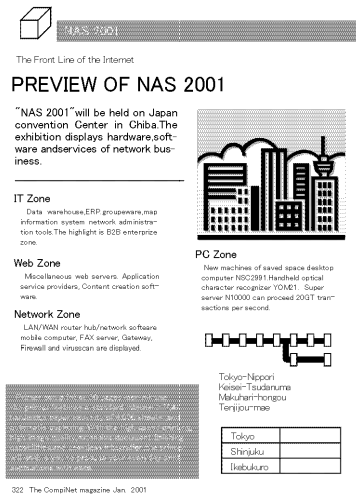


【 図 2 0 】

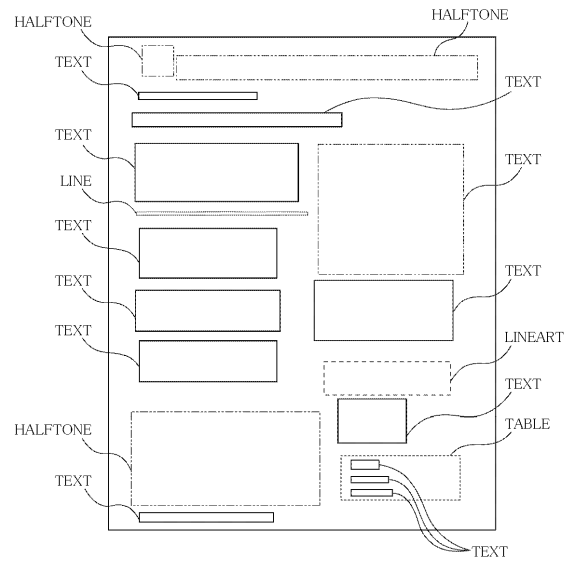




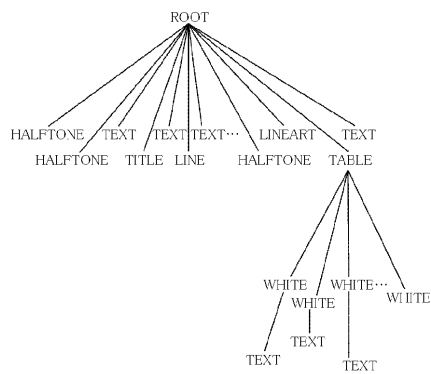
【 図 2 1 】



【 図 2 2 】



【 図 2 3 】



---

フロントページの続き

- (56)参考文献 特開昭62-271080(JP,A)  
特開平07-152920(JP,A)  
特開平08-235349(JP,A)  
特開平07-192084(JP,A)  
特開昭59-142678(JP,A)  
特開平09-269970(JP,A)

- (58)調査した分野(Int.Cl., DB名)

G06T 7/40