

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4217969号
(P4217969)

(45) 発行日 平成21年2月4日(2009.2.4)

(24) 登録日 平成20年11月21日(2008.11.21)

(51) Int.Cl.

H04N 1/413 (2006.01)
H04N 1/40 (2006.01)

F 1

H04N 1/413
H04N 1/40D
F

請求項の数 6 (全 30 頁)

(21) 出願番号 特願2003-409680 (P2003-409680)
 (22) 出願日 平成15年12月8日 (2003.12.8)
 (65) 公開番号 特開2005-175641 (P2005-175641A)
 (43) 公開日 平成17年6月30日 (2005.6.30)
 審査請求日 平成18年11月28日 (2006.11.28)

(73) 特許権者 000005496
 富士ゼロックス株式会社
 東京都港区赤坂九丁目7番3号
 (74) 代理人 100122275
 弁理士 竹居 信利
 (72) 発明者 熊澤 幸夫
 神奈川県海老名市本郷2274番地 富士
 ゼロックス株式会社内
 審査官 松永 隆志

最終頁に続く

(54) 【発明の名称】 画像処理装置及びプログラム

(57) 【特許請求の範囲】

【請求項1】

処理対象となった画像データから、画素が連続する部分からなる複数の画素群を特定する画素群特定手段と、

前記画素群の各々に関する矩形を、基本矩形として画定する基本矩形画定手段と、

前記基本矩形同士を所定の連結条件に基づいて水平又は垂直方向に連づけ、互いに連づけられた基本矩形に關係する矩形を行矩形として、少なくとも一つ画定する行矩形画定手段と、

前記行矩形の数を調べる手段と、

前記行矩形の数が複数である場合に、各行矩形の幅や高さ、または位置のばらつきが所定のしきい値より大きいか否かにより、各行矩形に文字が含まれているか否かを判断する第1判断手段と、

前記行矩形の数が単数である場合に、前記行矩形に關係する基本矩形の数が「1」のときに、基本矩形の面積と、予め定めた面積しきい値とに基づいて当該行矩形に文字が含まれているか否かを判断する第2判断手段と、

を含み、

文字が含まれていると判断された行矩形内の前記画素群を文字部分として特定することを特徴とする画像処理装置。

【請求項2】

請求項1記載の画像処理装置であって、

前記行矩形に関係する基本矩形の数が「2」のときには、前記第2判断手段は、各基本矩形の面積の比と、各基本矩形間の距離とに基づいて当該行矩形に文字が含まれているか否かを判断することを特徴とする画像処理装置。

【請求項3】

請求項1または2記載の画像処理装置であって、

前記行矩形に関係する基本矩形の数が「3」以上のときには、前記第2判断手段は、基本矩形の数が予め定めた最大数を超えておらず、基本矩形の面積の最大値が予め定めた最大面積値を超えず、かつ、基本矩形の各組み合わせに係る面積の比が予め定めた面積比しきい値より大きくなない場合に、当該行矩形に文字が含まれていると判断する画像処理装置

。

10

【請求項4】

コンピュータを、

処理対象となった画像データから、画素が連続する部分からなる複数の画素群を特定する画素群特定手段と、

前記画素群の各々に関する矩形を、基本矩形として画定する基本矩形画定手段と、

前記基本矩形同士を所定の連結条件に基づいて水平又は垂直方向に関連づけ、互いに関連づけられた基本矩形に関係する矩形を行矩形として、少なくとも一つ画定する行矩形画定手段と、

前記行矩形の数を調べる手段と、

前記行矩形の数が複数である場合に、各行矩形の幅や高さ、または位置のばらつきが所定のしきい値より大きいか否かにより、各行矩形に文字が含まれているか否かを判断する第1判断手段と、

前記行矩形の数が単数である場合に、前記行矩形に関係する基本矩形の数が「1」のときに、基本矩形の面積と、予め定めた面積しきい値とに基づいて当該行矩形に文字が含まれているか否かを判断する第2判断手段と、

文字が含まれていると判断された行矩形内の前記画素群を文字部分として特定する手段と、

として機能させることを特徴とするプログラム。

【請求項5】

請求項4記載のプログラムであって、

コンピュータに、前記行矩形に関係する基本矩形の数が「2」のときには、各基本矩形の面積の比と、各基本矩形間の距離とに基づいて、当該行矩形に文字が含まれているか否かを判断させることを特徴とするプログラム。

【請求項6】

請求項4または5記載のプログラムであって、

コンピュータに、前記行矩形に関係する基本矩形の数が「3」以上のときには、基本矩形の数が予め定めた最大数を超えておらず、基本矩形の面積の最大値が予め定めた最大面積値を超えず、かつ、基本矩形の各組み合わせに係る面積の比が予め定めた面積比しきい値より大きくなない場合に、当該行矩形に文字が含まれていると判断させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像データから文字部分を特定する画像処理装置及びプログラムに関する。

【0002】

ラスタ画像のデータ（以下、区別するべき場合を除いて単に「画像データ」と呼ぶ）には、文字（テキスト）部分や、自然画の部分（絵柄部分）など、互いに性状の異なる多くの画像要素が含まれ得る。こうした画像要素は、その性状の相違から、例えば圧縮処理において異なる方式での圧縮が適していたりするなど、画像要素ごとに異なる画像処理を行うことが好ましい場合が多い。

40

50

20

【0003】

そこで従来から、いわゆるT/I分離と呼ばれる画像処理が研究・開発されている。従来、T/I分離の方法としては、例えば処理対象画像を二値化し、黒画素の連続する領域を画定し、当該画定した領域のサイズが予め定めたしきい値を下回る場合に当該領域に含まれる黒画素が文字を表すものと判定する方法等がある（特許文献1）。

【特許文献1】特開2003-8909号公報（段落番号0026を参照）

【特許文献2】特開2002-175532号公報

【発明の開示】**【発明が解決しようとする課題】****【0004】**

10

しかしながら、上記従来の画像処理方法では、例えば絵柄と判定された部分の内部にある文字などについては、それ以上の分離は為されず、絵柄と同じ画像処理が適用される結果となっていた。

【0005】

本発明は上記実情に鑑みて為されたもので、文字部分の抽出精度をより高めて、より適切な画像処理に供することのできる画像処理装置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】**【0006】**

20

請求項1記載の発明は、画像処理装置であって、処理対象となった画像データから、画素が連続する部分からなる複数の画素群を特定する画素群特定手段と、前記画素群の各々に関する矩形を、基本矩形として画定する基本矩形画定手段と、前記基本矩形同士を所定の連結条件に基づいて水平又は垂直方向に関連づけ、互いに関連づけられた基本矩形に関する矩形を行矩形として、少なくとも一つ画定する行矩形画定手段と、前記行矩形の数を調べる手段と、前記行矩形の数が複数である場合に、各行矩形の幅や高さ、または位置のばらつきが所定のしきい値より大きいか否かにより、各行矩形に文字が含まれているか否かを判断する第1判断手段と、前記行矩形の数が単数である場合に、前記行矩形に関する基本矩形の数が「1」のときに、基本矩形の面積と、予め定めた面積しきい値に基づいて当該行矩形に文字が含まれているか否かを判断する第2判断手段と、を含み、文字が含まれていると判断された行矩形内の前記画素群を文字部分として特定することとしたものである。

30

【0007】

請求項2記載の発明は、請求項1記載の画像処理装置であって、前記行矩形に関する基本矩形の数が「2」のときには、前記第2判断手段は、各基本矩形の面積の比と、各基本矩形間の距離とに基づいて当該行矩形に文字が含まれているか否かを判断することとしたものである。

【0008】

40

請求項3記載の発明は、請求項1または2記載の画像処理装置であって、前記行矩形に関する基本矩形の数が「3」以上のときには、前記第2判断手段は、基本矩形の数が予め定めた最大数を超えておらず、基本矩形の面積の最大値が予め定めた最大面積値を超える、かつ、基本矩形の各組み合わせに係る面積の比が予め定めた面積比しきい値より大きくなる場合に、当該行矩形に文字が含まれていると判断することとしたものである。

【0009】

請求項4記載の発明は、プログラムであって、コンピュータを、処理対象となった画像データから、画素が連続する部分からなる複数の画素群を特定する画素群特定手段と、前記画素群の各々に関する矩形を、基本矩形として画定する基本矩形画定手段と、前記基本矩形同士を所定の連結条件に基づいて水平又は垂直方向に関連づけ、互いに関連づけられた基本矩形に関する矩形を行矩形として、少なくとも一つ画定する行矩形画定手段と、前記行矩形の数を調べる手段と、前記行矩形の数が複数である場合に、各行矩形の幅や高さ、または位置のばらつきが所定のしきい値より大きいか否かにより、各行矩形に文字が含まれているか否かを判断する第1判断手段と、前記行矩形の数が単数である場合に、前記行矩形の数が

50

記行矩形に關係する基本矩形の数が「1」のときに、基本矩形の面積と、予め定めた面積しきい値とに基づいて当該行矩形に文字が含まれているか否かを判断する第2判断手段と、文字が含まれていると判断された行矩形内の前記画素群を文字部分として特定する手段と、として機能させることとしたものである。

【0010】

また請求項5記載の発明は、請求項4記載のプログラムであって、コンピュータに、前記行矩形に關係する基本矩形の数が「2」のときには、各基本矩形の面積の比と、各基本矩形間の距離とに基づいて、当該行矩形に文字が含まれているか否かを判断させることとしたものである。さらに請求項6記載の発明は、請求項4または5記載のプログラムであって、コンピュータに、前記行矩形に關係する基本矩形の数が「3」以上のときには、基本矩形の数が予め定めた最大数を超えておらず、基本矩形の面積の最大値が予め定めた最大面積値を超えず、かつ、基本矩形の各組み合わせに係る面積の比が予め定めた面積比しきい値より大きくない場合に、当該行矩形に文字が含まれていると判断させることとしたものである。

10

【発明を実施するための最良の形態】

【0011】

本発明の第1の実施の形態に係る画像処理装置は、図1に示すように、制御部11と記憶部12と画像入力部13と画像出力部14とを含んで構成されている。制御部11は、記憶部12に格納されているプログラムに従って動作しており、後に説明する各画像処理を遂行する。この画像処理の内容については、後に詳しく述べる。

20

【0012】

記憶部12は、制御部11によって実行されるプログラムを保持している。またこの記憶部12は、制御部11の処理の過程で生成される各種データ等を格納するワークメモリとしても動作する。具体的にこの記憶部12は、コンピュータ可読な記録媒体と当該記録媒体に対してデータを書き込み、又は当該記録媒体からデータを読み出す装置（例えばハードディスク装置やメモリ装置）として実装できる。

【0013】

画像入力部13は、例えばスキャナであり、原稿を光学的に読み取って得られた画像データを制御部11に出力する。ここではこの画像入力部13が出力する画像データにおいて、各画素の値がRGB（赤、緑、青）の色空間で表現されているとする。画像出力部14は、制御部11から入力される指示に従って画像データを出力するもので、例えば画像を記憶部12に格納し、又は、画像形成部（プリンタ等）に出力し、又はネットワークを介して外部の装置に送信する等の処理を行うものである。

30

【0014】

次に制御部11の処理の内容について説明する。本実施の形態の制御部11は、図2に機能的に示すように、画像入力部13から入力される画像データを処理対象として、この処理対象となった画像データに対して、所定前処理を行う前処理部21と、絵柄部分の候補となる部分（絵柄候補部分）を特定する絵柄候補部分特定処理部22と、文字線画を抽出する文字線画抽出処理部23と、レイアウト処理を遂行するレイアウト処理部24と、同一色領域分離部25と、穴埋処理部26と、圧縮処理部27とを含んで構成されている。

40

【0015】

以下、これら各部について具体的に説明する。

【0016】

[1. 前処理部]

この前処理部21では、画像入力部13から入力される画像データ（処理対象画像データ）の各画素の値をRGBからYCbCr（輝度と色差とからなる値）に変換する。具体的には、次の（1）式を用いて変換を行うことができる。なお、ここではRGBの各成分の値は0x00（「0x」は16進数であることを示す）から0xFFまでの値であるとしている。また、この前処理部21では、下地領域の輝度・彩度に基づいて各画素値を階調補正して

50

もよい。尤も、この階調補正の処理は、必ずしも必要なものではない。

【0017】

【数1】

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.16874 & -0.33126 & 0.5 \\ 0.5 & -0.41869 & -0.08131 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 0x80 \\ 0x80 \end{pmatrix} \quad (1)$$

【0018】

【2. 絵柄候補領域特定処理部】

絵柄候補部分特定処理部22は、前処理部21が出力するYCbCr色空間で表現された画像データから絵柄の領域と推定される領域を、絵柄候補領域として特定する処理を行う。具体的に、絵柄候補部分特定処理部22の処理方法は、種々の方法があるので、以下そのうちの代表的な2つの方法の例について説明する。10

【0019】

【2a. 小領域除去による方法】

絵柄候補部分特定処理部22は、図3(a)に示すように、二値化処理部31と、連結画素抽出部32と、特徴量算出部33と、属性判定部34と、非絵柄領域処理部35と、背景領域塗潰部36と、絵柄候補領域作成部37とを含んで構成される。

【0020】

二値化処理部31は、前処理部21が出力するYCbCr色空間で表現された画像データを記憶部12にコピーして格納する。そして、当該コピーされた画像データ(絵柄領域特定画像データ)の各画素の値を参照しながら、当該画素値の輝度成分と、予め定めた二値化しきい値とを比較して、輝度成分が二値化しきい値以上の画素を黒画素(値「1」)、輝度成分が二値化しきい値未満の画素を白画素(値「0」)としていき、当該絵柄領域特定画像データを二値画像データに変換する。20

【0021】

連結画素抽出部32は、二値画像データ中から、黒画素が連結している領域(連結領域)を抽出する。この連結領域の抽出は、ラベリング処理として広く知られた処理を用いることができる。ラベリング処理を用いる場合、各連結領域には互いに異なるラベル識別子を付して、記憶部12に格納しておく。30

【0022】

特徴量算出部33は、連結画素抽出部32によって抽出された連結領域の各々について、その所定のスケール特徴量を演算する。ここでスケール特徴量には、連結領域に関係して定められる矩形の面積や、当該矩形内での黒画素の密度(矩形内の黒画素の数を矩形内の全画素の数で除したもの)などを含む。例えば特徴量算出部33は次のように処理を行う。すなわち特徴量算出部33は、各連結領域ごとの関係領域として、各連結領域に外接する矩形(外接矩形)を画定する。具体的に外接矩形は、連結領域に含まれる画素のうち、そのX座標が最小のもの(絵柄領域特定画像データ中で最も左にあるもの)の当該X座標をXminとし、同じくY座標が最小のもの(絵柄領域特定画像データ中で最も上にあるもの)の当該Y座標をYminとし、以下同様にX座標が最大のものと、Y座標が最大のものとにおける、X座標とY座標とのそれぞれをXmax、Ymaxとして、(Xmin, Ymin)から(Xmax, Ymax)までの線を対角線とする矩形を外接矩形と定める。そして当該外接矩形の面積Aを、(Xmax - Xmin) × (Ymax - Ymin)として演算する。なお、この面積Aは、当該外接矩形に含まれる全画素の数にも等しい。一方、特徴量算出部33は、当該外接矩形内の黒画素の数をカウントし、その数を上記全画素の数(すなわち面積Aの値)で除して密度Dを求める。40

【0023】

こうして特徴量算出部33は、各連結領域について、これら外接矩形の面積A(面積特徴量に相当する)と、黒画素密度D(密度特徴量に相当する)とを、当該連結領域のスケール特徴量として算出する。そして、各連結領域の識別子(例えばラベル識別子)とその50

スケール特徴量とを関連づけて、記憶部12にスケール特徴量データベースとして格納しておく。

【0024】

属性判定部34は、記憶部12に格納されたスケール特徴量データベースを参照しながら、各連結領域が絵柄候補領域であるか否かを判定する。具体的な例を挙げると、この属性判定部34は、文字であれば一定以上のサイズがあり、その内部の黒画素密度は比較的小さいはずであるとの前提に基づく処理を行う。つまり、一般的なドキュメントで利用されるフォントサイズの最小値（例えば6ポイント）に相当する面積（いわば最小面積）未満の領域（小領域）は、文字を含まない、絵柄の領域と判定できるので、こうした小領域を除去することで文字部分を特定できる。

10

【0025】

また、文字は基本的に線からなるので、複雑な文字であっても領域内の画素のうち一定割合以上が黒画素となることはない。そこで、最小面積より大きい面積を有する矩形であっても、その内部で黒画素の数の占める割合（つまり黒画素密度）が比較的大きい場合に、絵柄であると判定できる。

【0026】

さらに、本実施の形態では、例えば一般的なドキュメントで利用されるフォントサイズの最大値（例えば24ポイント）に相当する面積（いわば最大面積）以上の場合、文字は含まれないと基本的に判断できる。しかしながら、こうした領域は絵柄だけでなく線図（グラフなど）である場合もある。そこで、ここでも黒画素密度を利用して、最大面積より大きい面積を有する矩形のうち、黒画素密度が比較的大きい場合に限り、絵柄であると判定する。なぜならば、グラフなどの線図は、線を用いて構成されるので黒画素の密度は比較的低くなると推定されるからである。

20

【0027】

つまり、本実施の形態では、各連結領域ごとに、その面積特徴量が予め定めた複数の等級のいずれに属するかを判断し、各等級について予め定められた密度しきい値のうち、当該面積特徴量が属する等級に係る密度しきい値と密度特徴量とを比較して、絵柄であるか否かを表す情報（本発明にいう分類情報）を生成する。

【0028】

具体的に本実施の形態の属性判定部34は、上記最小面積を表す第1の面積しきい値1と、最大面積を表す第2の面積しきい値2（ただし、 $1 < 2$ ）並びに、第1、第2の密度しきい値（ $1, 2$ ）を予め定めておき、これらのしきい値を用いて、次のように判定する。

30

【0029】

すなわち属性判定部34は、図4に示すように、スケール特徴量データベースに含まれる連結領域のうち、この処理で未だ選択されていない連結領域の一つを注目領域として選択する（S1）。そしてこの注目領域に関連づけられたスケール特徴量を読み出して、当該スケール特徴量に含まれる面積Aが、第1の面積しきい値1未満となっているか否か（ $A < 1$ か否か）を調べる（S2）。ここで、 $A < 1$ であれば、注目領域を非絵柄領域として判定し、注目領域の識別子（例えばラベル識別子）と、その判定結果（非絵柄領域である旨を表す情報）とを関連づけて記憶部12に格納し（S3）、スケール特徴量データベース内にこの処理で未だ選択されていない連結領域があるか否かを調べて（S4）、未選択の連結領域があれば（Yesならば）、処理S1に戻って処理を続ける。また、処理S4において、未選択の連結領域がなければ（Noならば）、つまりすべての連結領域について判定を終了したならば、処理を終了する。

40

【0030】

また、処理S2において、 $A < 1$ でなければ、続いて注目領域の面積Aが第2の面積しきい値2を越えているか否か（ $A > 2$ であるか否か）を調べ（S5）、 $A > 2$ であれば、注目領域のスケール特徴量の一つである黒画素密度Dが、第1の密度しきい値1未満であるか否か（ $D < 1$ か否か）を調べ（S6）、 $D < 1$ であれば、注目領域が

50

絵柄領域である判定して、注目領域の識別子（例えばラベル識別子）と、その判定結果（絵柄領域である旨を表す情報）とを関連づけて記憶部12に格納し（S7）、処理S4に移行する。さらに処理S6において、D < 1でなければ、処理S3に移行して（つまり非絵柄領域と判定して）処理を続ける。

【0031】

また、処理S5において、A > 2でなければ、注目領域のスケール特徴量の一つである黒画素密度Dが、第2の密度しきい値2未満であるか否か（D < 2か否か）を調べ（S8）、D < 2であれば、処理S3に移行して（つまり非絵柄領域と判定して）処理を続け、D < 2でなければ、処理S7に移行して（つまり絵柄領域と判定して）処理を続ける。なお、第1，第2の密度しきい値1，2は、それぞれ実験によって線画等を抽出するのに適した値と、文字等を抽出するのに適した値とに定める。10

【0032】

また、面積特徴量又は面積しきい値の少なくとも一方を、処理対象となっている画像データの解像度に基づいて補正してもよい。例えば面積しきい値1，2が解像度r0で設定された値である場合において、処理対象となった画像データの解像度がrであったとき、第1の面積しきい値1を $(r / r_0) \times 1$ に補正し、第2の面積しきい値2を $(r / r_0) \times 2$ に補正する。

【0033】

非絵柄領域処理部35は、属性判定部34によって生成されて記憶部12に格納されている、各連結領域ごとの判定結果を参照して、絵柄領域特定画像データのうち、非絵柄領域と判定された連結領域に含まれる黒画素を白画素に変換して、非絵柄領域と背景とが白画素となっている絵柄領域特定画像データを生成する。20

【0034】

背景領域塗潰部36は、非絵柄領域と背景とが白画素となっている絵柄領域特定画像データに対して、絵柄部分の境界と連結している白画素を、所定の値（白、黒以外の任意の色）で塗りつぶす。ここでは、広く知られているような、閉じた領域（閉曲線で囲まれた内部又は外部）を塗つぶす処理を用いる。絵柄候補領域作成部37は、背景領域塗潰部36によって所定の値で塗潰されなかった画素の色を黒に設定する。この絵柄候補領域作成部37によって黒画素に設定された部分を絵柄候補領域として扱うことになる。

【0035】

この、小領域除去の方法による絵柄候補部分特定処理部22の動作について説明する。以下、図5(a)に示す画像データを処理対象とする場合を例として説明する。なお、図示の都合上、図5(a)においては、全体を線図で示しているが、図5(a)には、写真部分Pと、テキスト部分Tと、線図部分Gと、地図などの図版部分Mとが含まれているものとする。また、地図M内には、地図を構成する線分と一部重なり合う文字等が記述されている場合がある(X)。例えば道路名を示す文字などは、他の道路と重なり合って示されていることが多い。

【0036】

絵柄候補部分特定処理部22がこれを二値化処理して(図5(b))、連結画素を抽出し、各連結画素についての特徴量に基づき小領域除去の処理を行う。すると絵柄領域特定画像データは、図5(c)に示すように、テキスト部分(T)が除かれた状態となる。なお図5においては線図部分Gの黒画素密度が比較的高いので、当該線図部分Gは除去されずに残存した状態となっている。また、地図部分Mにおいて線分と重なり合っている文字は、当該重なり合う線分と一緒に連結画素として抽出されてしまう。このためこの連結画素を囲む矩形の面積が大きくなり、小領域として除去されず、残存することとなる(図5(c)のX)。40

【0037】

絵柄候補部分特定処理部22は、背景となっている白画素部分Bを所定の色で塗潰し(図5(d))、さらにこの所定の色で塗潰された部分以外の部分を黒画素とする。ここで背景となっている部分を塗潰す方法としては、画像データの四隅(左上、右上、左下、右50

下の四隅)から線分によって仕切られない領域を塗潰す、一般的なペイント処理を利用する方法を探ることができる。そして絵柄候補部分特定処理部22は、上記所定の色部分を白画素に戻す。すると、図5(e)に示すように、絵柄候補となる部分が黒画素となり、それ以外の部分が白画素となっているような絵柄領域特定画像データが得られることとなる。

【0038】

また、絵柄候補部分特定処理部22は、非絵柄領域処理部35によって絵柄領域特定画像データのうち、非絵柄領域と判定された連結領域に含まれる黒画素を白画素に変換した後の絵柄領域特定画像データに対して、膨張収縮処理を行ってもよい。すなわち、図3(b)に示すように、絵柄候補部分特定処理部22による処理は機能的に、二値化処理部31と、連結画素抽出部32と、特徴量算出部33と、属性判定部34と、非絵柄領域処理部35と、膨張収縮部38と、背景領域塗潰部36と、絵柄候補領域作成部37とを含んで構成されてもよい。10

【0039】

ここで膨張収縮部38は、非絵柄領域処理部35によって絵柄領域特定画像データのうち、非絵柄領域と判定された連結領域に含まれる黒画素を白画素に変換した後の絵柄領域特定画像データに含まれる画素を順次注目画素として選択する。

【0040】

そして膨張収縮部38は、注目画素の近傍(上下左右の画素からなる4近傍または、周辺8画素の8近傍)に黒画素が一つでもあれば、当該注目画素を黒画素とする処理(膨張処理)を行って、黒画素部分を膨張させて、次の注目画素を選択する。この選択順序は、上から下へとラインを走査し、ライン内は左から右へと走査する順序(いわゆるラスタスキャン順)とすればよい。20

【0041】

膨張収縮部38は、すべての画素に対する膨張処理を完了すると、再び、各画素を順次注目画素として選択しながら収縮処理を開始する。収縮処理では、注目画素の近傍(4近傍または8近傍)に白画素が一つでもあれば、注目画素を白画素とする。そして、全画素について収縮処理が完了すると、膨張収縮部38の処理を終了する。

【0042】

この場合、背景領域塗潰部36は、膨張収縮部38によって処理された後の絵柄領域特定画像データに対して、絵柄部分の境界と連結している背景部分の白画素を、所定の値(白、黒以外の任意の色)で塗りつぶすことになる。30

【0043】

このように膨張収縮処理を行うことによって、網点の領域が背景領域と区別されて絵柄領域として認識され易くなる。

【0044】

本実施の形態において特徴的なことの一つは、この絵柄領域候補を特定する処理において文字部分を絵柄と誤判定しても、後に行われるレイアウト処理によって文字部分がさらに抽出されるので、絵柄候補領域の抽出精度が比較的低くても構わないことである。

【0045】

なお、ここまで説明では、背景領域塗潰部36は、背景部分を所定の色で塗り潰すこととしているが、これに代えて次のようにしてもよい。すなわち、連結領域の抽出にラベリング処理を用いている場合に、当該抽出の際に得られた最大のラベルの値より大きい値(例えば最大のラベル値に1を加えた値)のラベル(背景特定ラベル)を背景部分の画素のラベルに置換えて設定していくこととしてもよい。この場合、絵柄候補領域作成部37は、当該背景特定ラベルが付されていない画素を黒画素に設定することで、図5(e)に示した絵柄領域特定画像データを生成してもよい。40

【0046】

制御部11は、特定された絵柄候補領域の各々を規定する座標情報(以下、絵柄候補領域画定情報と呼ぶ)を記憶部12に格納する。50

【0047】**[3. 文字線画抽出処理部]**

文字線画抽出処理部23は、前処理部21が出力するYCbCr色空間で表現された画像データから文字線画部分を抽出する処理を行う。この処理は、例えば特許文献2に開示されている通の処理を用いることができる。

【0048】

そして文字線画抽出部処理部23は、個々の文字や、線画部分についてそれらを取囲む矩形（文字線画外接矩形）の座標情報を記憶部12に格納する。

【0049】**[4. レイアウト処理部]**

10

レイアウト処理部24は、絵柄候補部分特定処理部22が生成した絵柄候補領域画定情報と、文字線画抽出処理部23が生成した文字線画外接矩形の座標情報を記憶部12から読み出す。

【0050】

レイアウト処理部24は、この絵柄候補領域画定情報によって画定される絵柄候補領域と、文字線画外接矩形の座標情報で画定される文字線画の領域とのそれぞれに対して互いに異なる判定条件を用いてレイアウト解析を行う。

【0051】

具体的に本実施の形態においては、記憶部12から読み出した絵柄候補領域画定情報によって画定される絵柄候補領域においてレイアウト解析処理を行い、当該絵柄候補領域の中からさらに文字部分の抽出を試みる。そして、抽出された文字部分を除く部分を絵柄領域として画定し、後段の処理に供する。

20

【0052】

ここでレイアウト解析処理の処理内容について説明する。レイアウト処理部24は図6に機能的に示すように、二値化処理部41と、連結画素抽出部42と、基本矩形画定部43と、第1セパレータ検出部44と、行矩形画定部45と、第2セパレータ検出部46と、文字領域画定部47と、ノイズ判定部48と、文字部分特定部49とを含んで構成される。

【0053】

二値化処理部41は、前処理部21が出力するYCbCr色空間で表現された画像データ（元の画像データ）のうち、絵柄候補領域画定情報で画定される領域内の部分的な画像データ（絵柄候補部分データ）を処理対象として、この処理対象となった絵柄候補部分データを二値化して、二値化絵柄候補部分データを生成する。

30

【0054】

連結画素抽出部42は、二値化絵柄候補部分データに対してラベリング処理を行い、所定の条件（例えば黒画素である等の条件）を満足する画素値の画素が連続する部分からなる複数の画素群（連結画素群）を特定する。

【0055】

なお、これらの処理は、絵柄候補部分特定処理部22の二値化処理部31と連結画素抽出部32とによる処理結果をそのまま用いてよい。その場合、レイアウト解析処理における二値化処理部41と連結画素抽出部42とは必ずしも必要ではない。

40

【0056】

基本矩形画定部43は、連結画素抽出部42が特定した連結画素群に関する矩形（例えば連結画素群に外接する矩形）を基本矩形として画定し、各連結画素群についての基本矩形の座標情報（当該矩形を画定するための座標情報）を生成する。そして、各基本矩形ごとに固有の識別子を発行し、当該識別子と基本矩形の座標情報とを関連づけて基本矩形データベースとして記憶部12に格納する。

【0057】

第1セパレータ検出部44は、処理対象となっている絵柄候補部分データの左上端の画素を初期位置として、各画素を左から右へと1ライン走査し、一つ下のラインについて同

50

様に（すなわちラスタスキャン順に）走査していき、上記ラベリング処理における所定の条件を満足しない画素値の画素（例えば白画素）が、予め定めた水平方向閾値より多く連続している場合に、当該連続画素部分を（水平方向の）第1セパレータとして検出し、当該第1セパレータを特定する情報（連続画素部分の左端画素の座標と右端画素の座標など、ここで座標は、元の画像データ上の座標であってもよいし、絵柄候補部分データ上のローカルな座標であってもよい）を生成して記憶部12に格納する。

【0058】

また、この第1セパレータ検出部44は、処理対象となっている絵柄候補部分データの左上端の画素を初期位置として、各画素をの画素を上から下へと1ライン走査し、一つ右のラインについて同様に走査していき、上記ラベリング処理における所定の条件を満足しない画素値の画素（例えば白画素）が、予め定めた垂直方向閾値より多く連続している場合に、当該連続画素部分を（垂直方向の）第1セパレータとして検出し、当該第1セパレータを特定する情報（連続画素部分の上端画素の座標と下端画素の座標など、ここで座標は、元の画像データ上の座標であってもよいし、絵柄候補部分データ上のローカルな座標であってもよい）を生成して記憶部12に格納する。10

【0059】

これらの処理において、水平方向閾値や垂直方向閾値は、ユーザが任意に定め得る。水平方向閾値は多段組のレイアウトにおいて各段を分かつための閾値であり、垂直方向閾値は2行以上の文字列を含む文書から、各行を分かつための閾値である。またユーザの設定によるだけでなく、水平方向閾値は基本矩形の幅の統計値（平均）などに基づく所定閾値として、また垂直方向閾値は基本矩形の高さの統計値（平均）などに基づく所定閾値としてそれぞれ定めてもよい。20

【0060】

具体的に第1セパレータは、図7（a）に示すような状態で検出されることになる。なお、図7（a）では各第1セパレータが相互に隣接して検出された結果として、一つの第1セパレータ領域のように示されている。

【0061】

行矩形画定部45は、記憶部12に格納されている基本矩形の一つを注目基本矩形として選択する。そして、記憶部12に格納されている基本矩形であって、今までに注目基本矩形として選択されていない基本矩形を処理対象基本矩形として順次選択しながら、次の処理を行う。30

【0062】

すなわち、注目基本矩形の中心座標（座標情報が対角位置にある各頂点の座標を表している場合、その中点座標）から、処理対象基本矩形の中心座標へのベクトルを算出する。さらに行矩形画定部45は、ベクトルの大きさ（各成分の二乗和の平方根）から注目基本矩形と処理対象基本矩形との距離を算出する。そして、この算出した距離が予め定めた距離閾値以下となっている場合は、上記算出したベクトルが、検出された第1セパレータのいずれかと交差するか否かを調べる。この処理は2つの線分が交差するか否かを調べる処理として広く知られたものを用いることができる。ここで、上記算出したベクトルが、検出された第1セパレータのいずれとも交差しない場合、注目基本矩形の識別子に、当該処理対象基本矩形の識別子を関連づけて基本矩形関係データベースとして記憶部12に格納する。40

【0063】

行矩形画定部45は記憶部12に格納されている基本矩形について順次注目基本矩形として選択しながら上記処理を行う。そしてこの処理の結果として得られた基本矩形関係データベースを参照しながら連鎖的に互いに関連する一連の基本矩形群（複数あってもよい）を特定し、特定された基本矩形群に含まれる基本矩形に外接する矩形を行矩形として画定する（例えば図7（b））。

【0064】

例えば基本矩形関係データベース内において、識別子が「1」の基本矩形と識別子が「50

2」の基本矩形とが関連づけられ、また識別子が「2」の基本矩形と識別子が「3」の基本矩形とが関連づけられている場合、行矩形画定部45は、これらの結果を統合して識別子「1」と「2」と「3」との各基本矩形を一連の基本矩形群として特定する。そして基本矩形群に含まれる基本矩形のうち、その座標情報のx(水平方向)の値の最大値と最小値とを抽出し、同じようにy(垂直方向)の値の最大値と最小値とを抽出する。そして、抽出されたxの最小値とyの最小値とを組とした第一座標と、xの最大値とyの最大値とを組とした第二座標とを、それぞれ左上座標、右下座標とする行矩形を画定する。つまり、行矩形は、この2つの座標値を含む座標情報によって画定される。

【0065】

行矩形画定部45は、こうして画定した各行矩形についてそれぞれ固有の識別子を発行し、各識別子とその行矩形の座標情報と当該行矩形に含まれる基本矩形群を特定する情報(各基本矩形の識別子のリストなど)とを関連づけて行矩形データベースとして記憶部12に格納する。10

【0066】

第2セパレータ検出部46は、処理対象となっている絵柄候補部分データの左上端の画素を初期位置として、各画素を左から右へと1ライン走査し、一つ下のラインについて同様に(すなわちラスタスキャン順に)走査していく、連結画素抽出部42でのラベリング処理における所定の条件を満足する画素値の画素(例えば黒画素)が、予め定めた水平方向閾値より多く連続している場合に、当該連続画素部分を(水平方向の)第2セパレータとして検出し、当該第2セパレータを特定する情報(連続画素部分の左端画素の座標と右端画素の座標など、ここで座標は、元の画像データ上の座標であってもよいし、絵柄候補部分データ上のローカルな座標であってもよい)を生成して記憶部12に格納する。20

【0067】

また、この第2セパレータ検出部46は、処理対象となっている絵柄候補部分データの左上端の画素を初期位置として、各画素の画素を上から下へと1ライン走査し、一つ右のラインについて同様に走査していく、連結画素抽出部42でのラベリング処理における所定の条件を満足する画素値の画素(例えば黒画素)が、予め定めた垂直方向閾値より多く連続している場合に、当該連続画素部分を(垂直方向の)第2セパレータとして検出し、当該第2セパレータを特定する情報(連続画素部分の上端画素の座標と下端画素の座標など、ここで座標は、元の画像データ上の座標であってもよいし、絵柄候補部分データ上のローカルな座標であってもよい)を生成して記憶部12に格納する。30

【0068】

これらの処理において、水平方向閾値や垂直方向閾値は、ユーザが任意に定め得る。水平方向閾値は多段組のレイアウトにおいて各段を分かつための閾値であり、垂直方向閾値は2行以上の文字列を含む文書から、各行を分かつための閾値である。またユーザの設定によるだけでなく、水平方向閾値は基本矩形の幅の統計値(平均)などに基づく所定閾数値として、また垂直方向閾値は基本矩形の高さの統計値(平均)などに基づく所定閾数値としてそれぞれ定めてよい。

【0069】

文字領域画定部47は、行矩形画定部45が画定した行矩形の一つを注目行矩形として選択する。そして文字領域画定部47は、記憶部12に格納されている行矩形であって、今までに注目行矩形として選択されていない行矩形を処理対象行矩形として順次選択しながら、次の処理を行う。40

【0070】

すなわち、注目行矩形の各頂点の座標と処理対象行矩形の対応する頂点の座標とを結ぶ線分、及び注目行矩形と処理対象行矩形の各辺とによって画定される多角形領域を生成し、この多角形領域と第2セパレータ(の領域)とが交差(領域同士が少なくとも一部で重なり合う)するか否かを調べる。この処理は2つの領域が交差するか否かを調べる処理として広く知られたものを用いることができる。

【0071】

50

20

30

40

50

ここで多角形領域と第2セパレータとが交差していない場合、注目行矩形の識別子と処理対象行矩形の識別子とを関連づけて行矩形関係データベースとして記憶部12に格納する。

【0072】

文字領域画定部47は記憶部12に格納されている行矩形について順次注目行矩形として選択しながら上記処理を行う。そしてこの処理の結果として得られた行矩形関係データベースを参照しながら連鎖的に互いに関連する一連の行矩形群（複数あってもよい）を特定し、特定された行矩形群に含まれる行矩形に外接する矩形を文字領域として画定する。

【0073】

例えば行矩形関係データベース内において、識別子が「1」の行矩形と識別子が「2」の行矩形とが関連づけられ、また識別子が「2」の行矩形と識別子が「3」の行矩形とが関連づけられている場合、行矩形画定部45は、これらの結果を統合して識別子「1」と「2」と「3」との各行矩形を一連の行矩形群として特定する。そして行矩形群に含まれる行矩形のうち、その座標情報のx（水平方向）の値の最大値と最小値とを抽出し、同じようにy（垂直方向）の値の最大値と最小値とを抽出する。そして、抽出されたxの最小値とyの最小値とを組とした第一座標と、xの最大値とyの最大値とを組とした第二座標とを、それぞれ左上座標、右下座標とする文字領域の矩形を画定する。つまり、文字領域は、この2つの座標値を含む座標情報によって画定される。

【0074】

文字領域画定部47は、こうして画定した文字領域についてそれぞれ固有の識別子を発行し、各識別子とその文字領域の座標情報と当該文字領域に含まれる行矩形群を特定する情報（各行矩形の識別子のリストなど）とを関連づけて文字領域データベースとして記憶部12に格納する。

【0075】

ノイズ判定部48は、文字領域画定部47によって画定された文字領域のそれぞれについて文字が含まれているか否かを確認する、ノイズ判定処理を行う。ここでノイズ判定処理は、行矩形の数、又は各行矩形の性状を表す情報に基づいて、各行矩形に文字が含まれているか否かを判断する第1ノイズ判定処理と、行矩形に関係する基本矩形に関する情報に基づいて、当該行矩形に文字が含まれているか否かを判断する第2ノイズ処理とを含む。

【0076】

具体的にノイズ判定部48の処理は、図8に示すような処理として行われる。まずノイズ判定部48は記憶部12に格納された文字領域のうち、未だ注目文字領域として選択されていないものを注目文字領域として選択する（S11）。そして注目文字領域に含まれる行矩形の数を調べ、これが2以上か（すなわち、当該文字領域が複数行からなるか）否かを判断する（S12）。ここで、行矩形の数が2以上であれば（Yesならば）、注目文字領域に含まれる各行矩形の幅と高さ、並びにそれらの平均や標準偏差など、ばらつきを検定するための統計量を演算する（S13）。そして、これらの統計量に基づいて各行矩形の幅や高さのばらつきが所定のしきい値より大きいか否かを比較する（S14）。この比較は例えば標準偏差が、予め定めたしきい値を超えるか否かの比較とができる。そしてこの処理S14によって、ばらつきが大きいと判断される場合（Yesの場合）、注目文字領域には文字は含まれないと判断して、記憶部12の文字領域データベースから、注目文字領域を削除して（S15）、処理S18に移行する。これら処理S12からS15の処理が、第1ノイズ判定処理に相当する。すなわちここでは行矩形の性状を表す情報として各行矩形の幅や高さ、並びにそれらの統計量が用いられている。

【0077】

一方、処理S14において、ばらつきが小さいと判断される場合（Noの場合）には、注目文字領域に含まれる各行矩形に対して行内判定処理（第2ノイズ判定処理）を行う（S16）。この処理S16の具体的な内容については後述する。そしてこの処理S16においてノイズ（文字が含まれていない）と判定された行の数と注目文字領域に含まれる行矩

10

20

30

40

50

形の数との比に基づいて、ノイズと判定された行の数が、注目文字領域に含まれる行矩形の数に比して所定比率以上となっているか否かを判断し(S17)、所定比率以上であるときに、注目文字領域には文字は含まれないと判断して、処理S15に移行する。

【0078】

また、処理S17において、所定比率未満である場合は、未だ注目文字領域となっていない文字領域が記憶部12の文字領域データベースにあるか否かを調べ(S18)、未選択の文字領域があれば、処理S11に戻って処理を続ける。さらに処理S18において、未選択の文字領域がないならば(すべての文字領域について処理を行ったならば)、ノイズ判定の処理を終了する。

【0079】

さらに処理S12において行矩形の数が1つであれば(Noならば)、処理S16に移行して処理を続ける。この場合、当該単一の行矩形について文字が含まれているか否かを判断し、文字が含まれていれば(この場合は処理S17の比率は「0」となる)、注目文字領域には文字が含まれると判断され、当該単一の行矩形内に文字が含まれていないならば(この場合は処理S17の比率は「1」となる)、注目文字領域には文字が含まれないと判断される。

【0080】

ここで、処理S16における具体的な処理(第2ノイズ判定処理)の内容について説明する。この処理ではノイズ判定部48は、図9に示すように、処理の対象となった各行矩形の一つを注目行矩形として選択し(S21)、記憶部12に格納されている行矩形データベースを参照して、当該注目行矩形に含まれる基本矩形の数をカウントする(S22)。そしてカウントの結果、基本矩形の数が「1」である場合と、「2」である場合と、「3」以上である場合とに分岐して(S23)、基本矩形の数が「1」である場合は、注目行矩形に含まれる基本矩形の識別子のリストを取得し、このリストに含まれる基本矩形の座標情報を記憶部12の基本矩形データベースから読み出して、このリストに含まれる基本矩形の幅及び高さとその積(つまり面積)を演算する。そしてこの面積が予め定められた面積しきい値以下であるか否かを判断し(S25)、面積しきい値以下である場合は、注目行矩形には文字が含まれないと判断して、当該結果を記憶部12に格納する(S26)。そして未だ注目行矩形として選択されていない行矩形があるか否かを調べ(S27)、未選択の行矩形があれば、当該未選択の行矩形の一つを注目行矩形として選択するべく処理S21に戻って処理を続ける。一方、処理S27において未選択の行矩形がなければ、処理を終了して図8の処理に戻る。

【0081】

さらに処理S22におけるカウント値が「2」である場合、各基本矩形の面積を演算し、また、これらの基本矩形間の距離を演算する。基本矩形間の距離は、例えば基本矩形の中心同士の距離として演算できる。そして、距離が予め定めた距離しきい値より大きいか、または2つの基本矩形の面積の比が予め定めた面積比しきい値より大きいかを判断し(S31)、距離が予め定めた距離しきい値より大きいか、または2つの基本矩形の面積の比が予め定めた面積比しきい値より大きい場合は、処理S26に移行して(X)処理を続ける。

【0082】

さらに、処理S22におけるカウント値が「3」以上である場合には、カウント値(基本矩形の数)が所定最大数を超えているか否かを判断し(S32)、この所定最大数を超える場合は、注目行矩形には文字が含まれないと判断して処理S26に移行する(X)。これは、一行内に例えば100字を超える文字を含めることは通常あり得ないことなどに配慮したものであり、固定値として定めておいてもよいし、注目行矩形の幅に基づいて調整してもよい。また処理S32において基本矩形の数が所定最大数を超えていない場合は、さらに各基本矩形の面積を演算して、演算された面積の最大値が、所定最大面積値を超えているか否かを判断する(S34)。ここで所定最大面積値を超えていると判断される場合は、注目行矩形には文字が含まれないと判断して処理S26に移行する(X)。この

10

20

30

40

50

最大面積値も、固定値として定めてもよいし、注目行矩形の幅や高さの少なくとも一方（例えばそれらのうち小さい方）に基づいて調整してもよい。

【0083】

さらに処理S34において所定最大面積値を超えていないと判断される場合、さらに2つの基本矩形の組み合せ（任意に取り出された少なくとも一つの組み合せ）について、各組み合せに係る2つの基本矩形の面積比が予め定めた面積比しきい値より大きいかを判断し（S35）、2つの基本矩形の面積の比が予め定めた面積比しきい値より大きい場合は、処理S26に移行して処理を続ける（X）。

【0084】

この処理S35において2つの基本矩形の面積の比が予め定めた面積比しきい値より大きくない場合は、注目行矩形内に文字が含まれると判断して、その判断結果を記憶部12に格納し、処理S27に移行する。 10

【0085】

なお、処理S25において、面積が面積しきい値を超える場合、並びに、処理S31において、距離が予め定めた距離しきい値以下であり、かつ2つの基本矩形の面積の比が予め定めた面積比しきい値以下である場合には、処理S32（または処理S34）以下に移行して処理を続けることとする。

【0086】

また、処理S35においては各組み合せについて処理を行っているが、処理負荷を軽減するためには、例えば各基本矩形の面積の平均値（平均面積）や、最小値（最小面積）・最大値（最大面積）を演算し、平均面積と最小面積、平均面積と最大面積との比、あるいは最小面積と最大面積との比と、上記面積比しきい値との比較を行ってもよい。 20

【0087】

このように、ノイズ判定部48は、各行矩形について、そこに含まれる基本矩形の性状（面積、面積比、距離など）に基づき、各行矩形に真に文字が含まれているか否かを再確認する。

【0088】

なお、ノイズ判定部48の第1ノイズ判定処理は、ここで述べた例に限られない。例えばここでは行矩形の性状として各行矩形の幅や高さを用いていたが、これらとともに、またはこれらに代えて、行矩形の座標情報（の平均値や標準偏差などの統計量）を用いてもよい。これによると、文字領域内に含まれる行矩形の位置がばらついている場合などに、当該文字領域には文字が含まれていない（ノイズである）と判断して、記憶部12の文字領域データベースから、注目文字領域を削除することとなる。 30

【0089】

文字部分特定部49は、ノイズ判定部48の処理を経た文字領域データベースを記憶部12から読出して、当該文字領域データベースに含まれる文字領域（文字領域の座標情報そのもの）、または当該文字領域内の黒画素部分（文字領域の座標情報と、黒画素部分からなるピットマップ情報）を文字部分として特定し、当該文字部分を特定する情報（文字部分特定情報）を記憶部12に格納する。制御部11は、この時点で記憶部12に格納されている、基本矩形関係データベースや行矩形関係データベースを削除してもよい。 40

【0090】

このように、本実施の形態におけるレイアウト処理部24は、文字から行、行から領域へと段階的に文字領域を画定し、当該画定した文字領域内の行の状態に基づいて文字列が含まれているかを判断し、文字列が含まれていないと判断される場合には、さらに行内（文字単位）の状態に基づいて文字が含まれているかを判断することとしている。尤も、本実施の形態におけるレイアウト処理はこれに限られるものではなく、その他広く知られたレイアウト処理を用いても構わない。

【0091】

本実施形態において特徴的なことの一つは、いわゆるT/I分離処理においてレイアウト解析を用いて文字部分を抽出することとしていることである。これによって絵柄候補領 50

域に含まれる文字部分もレイアウト解析処理によって抽出されることとなり、文字部分の抽出精度を向上させることができる。

【0092】

一方、レイアウト処理部24は、文字線画として画定された部分（文字線画外接矩形）内についても、レイアウト解析処理を行う。そしてレイアウト処理部24は、当該レイアウト解析処理の結果として得られたレイアウト枠（少なくとも各文字に外接する矩形（基本矩形）を含む）を定め、これらレイアウト枠の情報（座標情報など）を記憶部12に格納する。

【0093】

なお、文字線画外接矩形内においても、上記基本矩形を水平または垂直方向に関係づけて得られる行矩形を画定し、この行矩形の座標情報などを記憶部12に併せて格納してもよい。

10

【0094】

制御部11のレイアウト処理部24は、記憶部12に格納された文字部分特定情報と、文字線画外接矩形（またはそれに対するレイアウト処理結果としてのレイアウト枠）とを参照し、処理対象となった画像データ全体について、画像データ内で文字を含む矩形領域の座標情報を生成する。具体的にレイアウト処理部24は、文字部分特定情報の座標情報で画定される矩形と、文字線画外接矩形情報（またはそれに対するレイアウト処理結果としてのレイアウト枠）で画定される矩形とを結合して、文字線画領域を生成する。すなわち、本実施の形態では絵柄候補領域の画定と、文字線画領域の画定とを別々に行っているため、絵柄候補領域内にも文字線画領域として特定された領域が存在し得る。このためここでこれらの領域を結合して、重複領域を一つの文字線画領域とする。

20

【0095】

そしてレイアウト処理部24は、結合後の各文字線画部分について固有の領域識別子（以下、ラベルデータと呼ぶ）を生成し、このラベルデータと、対応する文字線画領域を画定するための座標情報（頂点座標の情報等）とを関連づけて記憶部12に文字線画領域データベースとして格納する。

【0096】

[5. 同一色領域分離部]

同一色領域分離部25は、記憶部12の文字線画領域データベースに格納されている文字線画領域の各々について、同一の色の文字線画部分のみからなる領域に分割する処理を行う。同一色領域分離部25は、レイアウト処理の過程で記憶部12に格納された基本矩形（絵柄候補領域と文字線画領域との双方から画定された矩形）の座標情報を読み出す。そして、処理対象となっている画像データ（元画像データ）のうち、この座標情報で画定される各基本矩形内に含まれる画素値のうち代表値（代表色）の候補を決定する。

30

【0097】

本実施の形態における同一色領域分離部25は、図10に示すように、色数判定部51と、第1代表色決定部52と、第2代表色決定部53と、限定色化部54とを含んで構成されている。

【0098】

色数判定部51は、記憶部12に格納された基本矩形（本発明の基本領域に相当する）の座標情報を参照し、それらを順次注目基本矩形として選択しながら、元画像データのうち注目基本矩形内の画素値に基づいて色数をカウントする。具体的にこの色数判定部51は、元画像データのうち注目基本矩形内の画素値のヒストグラム（発生頻度）を生成し、このヒストグラムにおいて所定しきい値（例えば注目基本矩形内の画素数の1/3など）を超える頻度で出現する画素値の数をカウントする。

40

【0099】

第1代表色決定部52は、色数判定部51によって判定された色数が所定の整数N（例えばN=1）以下であるか否かを調べ、色数が所定数N以下である場合は、注目基本矩形を処理対象として、当該注目基本矩形の画素値に基づいて少なくとも一つの代表画素値を

50

決定する。

【0100】

またこの第1代表色決定部52は、色数が所定数N以下でない場合は、注目基本矩形を処理対象としないこととなる。この場合、同一色領域分離部25は、当該注目基本矩形の部分を文字線画領域データベースから除去する処理を行って、当該注目基本矩形内の画素を絵柄として扱うこととしてもよい。

【0101】

なお、代表画素値の決定方法は、色数判定部51における色数判定の処理と同様に、元画像データのうち注目基本矩形内の画素値のヒストグラム（発生頻度）を生成し、このヒストグラムにおいて所定しきい値（例えば注目基本矩形内の画素数の1/3など）を超える頻度で出現する画素値を代表画素値として決定し、注目基本矩形を特定する情報（注目基本矩形に固有に発行された識別子）に関連づけて決定した代表画素値を代表画素値データベースとして記憶部12に格納する。10

【0102】

このように、各基本矩形内の色数に基づいて処理対象とするか否かが定められるので、グラデーション文字などが限定色化されることがない。

【0103】

第2代表色決定部53は、記憶部12に格納されている行矩形データベースを参照して、各行矩形を順次、注目行矩形として選択する。そして注目行矩形に含まれる少なくとも一つの基本矩形の識別子を取り出し、取り出した各識別子に関連づけられた代表画素値を代表画素値データベースから読出す。20

【0104】

以下、i番目の基本矩形についての代表画素値Pをa, b, cの3値で張られる色成分（a, b, cはL*a*b*やRGB等の色成分のどれであってもよい）を用いてPi (Pi a, Pi b, Pi c)と表現する。また、元の画像データにおいて当該基本矩形内の画素値の各色成分の標準偏差 i (ia, ib, ic)を演算する。

【0105】

第2代表色決定部53は、i番目の基本矩形とj（iとjとは互いに等しくないものとする）番目の基本矩形についてこれらの値、Pi, i, Pj, jに基づいて色空間内に形成される2つの所定幾何形状を生成し、この2つの幾何形状が色空間上で交差しているか（重なり合う部分があるか）否かによってi番目の基本矩形についての代表画素値と、j番目の基本矩形についての代表画素値が同じ色であるか否かを判断する。この幾何形状の交差の有無が本発明の所定均一色条件に相当する。30

【0106】

具体的には、幾何形状として直方体や橯円球を用いることができる。すなわち、直方体 (Pi a ± × ia, Pi b ± × ib, Pi c ± × ic)と、(Pj a ± × ja, Pj b ± × jb, Pj c ± × jc)とが重なり合う場合に、i番目の基本矩形についての代表画素値と、j番目の基本矩形についての代表画素値が同じ色であると判断してもよい。また、Piを中心としてiの各値を径とする橯円球と、Pjを中心としてjの各値を径とする橯円球とが重なり合う場合に、i番目の基本矩形についての代表画素値と、j番目の基本矩形についての代表画素値が同じ色であると判断してもよい。40

【0107】

さらに、標準偏差を演算する際に、色成分(a, b, c)に代えて、広く知られた主成分分析によって得られる3つの軸(a, b, c)を各成分として、これらの軸方向の標準偏差を演算し、直方体 (Pi a ± × ia, Pi b ± × ib, Pi c ± × ic)と、(Pj a ± × ja, Pj b ± × jb, Pj c ± × jc)とが重なり合う場合に、i番目の基本矩形についての代表画素値と、j番目の基本矩形についての代表画素値が同じ色であると判断してもよい。

【0108】

50

これらにおいて は、均一色との判定のされ易さを表すパラメータであり、例えば比較の対照となる基本矩形間の距離に応じて変化させてもよい（例えば隣接する基本矩形においては の値を大きくして、均一色と判定のされ易くするなど）し、予め定めた一定値としてもよい。

【 0 1 0 9 】

第2代表色決定部53は、注目行矩形内から取り出した2つの基本矩形の組み合せについて、上記処理によって互いに均一の色であると判断される基本矩形群を少なくとも一つ特定する。そして、特定した基本矩形群ごとに、それらの基本矩形群に含まれる画素を抽出して当該抽出した画素からなる色別行領域を生成する。これにより、行矩形内の画素であって、一つの代表値に関連づけられる画素ごとに、当該代表値と当該画素を再現する情報とを関連づけた情報を含む色別行領域情報が生成され、注目行矩形の識別子に関連づけて記憶部12に格納される。10

【 0 1 1 0 】

限定色化部54は、記憶部12に格納されている色別行領域情報ごとに、各色別行領域内の画素値をN色以下に限定色化して各画素値を設定する。

【 0 1 1 1 】

次に、図11を参照しながら、同一色領域分離部25の動作を説明する。ここでは具体的に図11(a)に示すような2行に配列された文字領域であって、各行の先頭にグラデーションがかけられた円が含まれ、文字の一部に赤色文字（破線で仮想的に囲んだ部分）が含まれるものとする。図11(a)では、各行について行矩形が画定されている状態が示されている。20

【 0 1 1 2 】

色数判定部51は、各文字についての色数を判定する。グラデーションがかけられた円については、複数色が含まれると判定され、他の文字部分については、黒又は赤色の一色と判定される。第1代表色決定部52は、グラデーションがかけられた円については処理対象から除外し、黒色文字については黒色の代表色を決定し、赤色文字については赤色の代表色を決定する。なお、この第1代表色決定部52の決定する代表色は、例えば同じ黒色と決定される色であっても、画素値自体は異なっていてよい。

【 0 1 1 3 】

第2代表色決定部53は、各行矩形内で、黒色と決定された文字についての基本矩形について、これらの基本矩形内の画素（黒画素）を含んでなる色別行矩形（図11(b)の(イ), (ハ)）と、赤色と決定された文字についての基本矩形について、これらの基本矩形内の画素（赤画素）を含んでなる色別行矩形（図11(b)の(ロ), (ニ)）を生成する。30

【 0 1 1 4 】

そして限定色化部54が、各色別行矩形について当該色別行矩形内の画素値に基づいてN色（第1代表色決定部52が用いる色数Nと同一の値）以下の限定色を定める。例えば黒画素を含んでなる色別行矩形については黒色に相当する一つの画素値を限定色として定め、赤画素を含んでなる色別行矩形については赤色に相当する一つの画素値を限定色として定める。また、各色別行矩形内の画素のうち限定色として定められる画素を黒色画素とし、それ以外の画素を白色画素として二値化しておく。これにより、画像データのうち、基本領域内の画素値が限定色化された状態となる。40

【 0 1 1 5 】

また、色別行矩形内の二値化画像をラスタスキャン順に走査しながら、この二値化画像に対してMMR（Modified Modified Read）などのランレンジス圧縮を行って、文字線画圧縮データを生成する。そして同一色領域分離部25は、行矩形の識別子と、当該行矩形に含まれる画素によって得られた色別行矩形とを関連づけ、さらに各色別行矩形とそれについて定めた限定色の情報を関連づけて、文字線画プレーンデータとして記憶部12に格納する。また、この限定色の情報は、画素値そのものではなく、各画素値を表す識別子の情報（以下、カラータグと呼ぶ）として保持してもよい。

【0116】

なお、基本矩形内の画素値がばらついている場合に配慮して、第1代表色決定部52は、平滑化処理を行ってから代表画素値を決定してもよい。ここで平滑化処理としては、各基本矩形内の各画素を順次注目画素として特定し、注目画素の値とそれに隣接する画素の値との平均値を注目画素の値とする処理などがある。

【0117】

さらにこの平滑化処理の際に、基本矩形内で文字を構成する画素（例えば二値化処理により黒画素となる部分）のみを注目画素として選択してもよい。また平滑化の処理において、平均値を演算する際は当該文字を構成する画素の値のみを参照して平均値を演算することとしてもよい。これにより、文字以外の部分の画素値を参照することにより、文字の代表色が背景色に影響されることが防止される。10

【0118】

ここで平滑化処理してから決定した代表値について補正を行ってもよい。すなわち、本実施の形態の同一色領域分離部25は、決定した代表値の候補についてその輝度を補正して、補正後の値を代表値として決定する。ここで輝度の補正是、例えば図12に示すようなトーンカーブ（補正関数）を用いて補正することができる。この図12に示すトーンカーブは、入力値（補正前の代表値候補の輝度）が最小値MINから第1しきい値TH1までに対する出力値（補正後の値、つまり代表値として決定される値の輝度）が最小値MINであり、第2しきい値TH2（ただしTH2 > TH1）から、最大値MAXまでに対する出力値が最大値MAXであるように設定されている。また、このトーンカーブは、入力値が最大値MAXと最小値MINとの間の中央の値MID（例えば最大値が「255」であり最小値が「0」であるときにはMIDは「128」となる）であるときに、これに対する出力値が略MIDとなるように設定されてもよい。20

【0119】

つまり同一色領域分離部25は、代表値の候補（本実施の形態ではYCbCrで表されることとしている）の輝度成分（Y）について、図12のトーンカーブによる補正を行って、代表値の輝度（Y）を決定し、このYと、代表値の候補の色差成分Cb,Crによって特定される値を代表値として決定する。

【0120】

なお、代表値候補がRGBなど、輝度成分を含まない色空間で表現されている場合は、L*a*b*や、YCbCrなど、輝度成分を含む色空間の値に変換してから上記処理を行うこととすればよい。30

【0121】

さらに、ここでは輝度のみを補正したが、色差成分についても補正を行ってもよい。具体的に同一色領域分離部25は、代表値候補の各色差成分が所定の条件を満足しているときに、当該代表色候補の輝度成分値に関する階調数を低減する補正を行い、当該補正後の値を代表値として決定してもよい。

【0122】

具体的には、図13に示すように、L*a*bの色空間で表現された代表値候補の色差成分（a*、b*）が、それぞれ対応する色差成分の値域の中心値からの所定範囲内（図13のTHa, THbで画定される円の内部）にあるとの条件を満足している場合に、例えば256階調で表現された輝度成分Lを4階調または8階調など所定階調に低減する。この場合、色差成分の値を上記中心値に設定してもよい。ここで、各成分ごとの所定範囲THa, THbは、同じ値であってもよいし、異なる値であってもよい。40

【0123】

この処理により、特に文字色がグレー（黒を含む）である場合に、その文字色の本来の色を再現した代表値が設定される。例えば文字色が黒であるときに色差成分と輝度成分とは本来「0」であるが、スキャナの特性や、元の画像データのエンコード形式（例えばJPEGなど）の特性によっては、色差成分が「0」でなくなってしまったり、輝度成分が「0」でなくなってしまう場合がある。そこでここで示した色差成分に関する処理を行う50

ことで、代表値を本来の黒色とすることができるようになる。

【0124】

このように本実施の形態によれば、処理対象となった画像データ内に画定される注目画像領域としての基本矩形や行矩形内の画素値に基づいて当該注目画像領域の代表色候補を決定し、その輝度を補正して、代表色を決定することとしている。

【0125】

なお、ここでは平滑化処理を行った後で、補正処理を行って代表値を決定しているが、この処理順序を逆にして各画素について上記補正処理を行った後で、各画素値の平滑化処理を行ってヒストグラムを演算し、代表値を決定してもよい。なお、ここでは文字の場合を説明したが、線画についても同様の処理が行われることとなる。

10

【0126】

こうした平滑化と補正の処理によって、本実施の形態においては、文字や線画を構成する画素値にはばらつきがある場合、当該ばらつきの影響を軽減して、元の画像データの色と違和感のない代表色を決定することができる。

【0127】

さらに、ここまで説明では行矩形ごと、かつ代表色ごとに限定色化処理を行っているが、色別行矩形ごとに定められた限定色について、さらに上記第2代表色決定部53におけるのと同様の処理によって互いに均一色と判断される限定色に関連づけられている色別行矩形内の画素を抽出し、当該抽出した画素を含み、文字線画部分を画定する矩形と同一サイズの矩形を、色別文字線画部分データとして生成して、これについてさらに限定色を定めてもよい。

20

【0128】

また、色別行矩形を生成する代りに、第2代表色決定部53が直接、一つの文字線画部分に含まれる基本矩形であって、互いに均一色と判断される代表画素値に関連づけられている基本矩形内の画素を抽出し、当該抽出した画素を含み、文字線画部分を画定する矩形と同一サイズの矩形を、色別文字線画部分データとして生成して、これについて限定色を定めてもよい。

【0129】

これら色別文字線画部分データを生成する場合は、当該色別文字線画部分データの元となつた文字線画部分の識別子と、それぞれ限定色を特定する情報に関連づけられた色別文字線画部分データ（内部の画素値は二値化し、MMR符号化してもよい）とを、文字線画プレーンデータとして生成することになる。

30

【0130】

[6. 穴埋処理部]

穴埋処理部26は、元の画像データのうち、絵柄候補領域に相当する領域を抽出し、この領域内からレイアウト処理部24の処理で検出された文字の画素（文字を構成する画素）を除去して、絵柄部分画像データを生成する。

【0131】

そして、この絵柄部分画像データの各画素をラスタスキヤン順に走査し、走査により選択される注目画素が除去された画素でなければ、当該注目画素の画素値をそのままとするとともに、当該注目画素の画素値を直前画素値として記憶部12のワークメモリに記憶する。なお、既に他の画素値が直前画素値として記憶されている場合は、その記憶内容に上書きする。

40

【0132】

また、走査により選択される注目画素が除去された画素である場合、当該注目画素の画素値を、記憶している直前画素値に設定する。これにより除去された部分の画素値が、ラスタスキヤン順に直前画素値と同一になり、多くの圧縮処理において圧縮効率を向上させることができるようになる。

【0133】

そしてこの処理を行った後の絵柄部分画像データを絵柄プレーンデータとして記憶部1

50

2に格納する。

【0134】

[7.圧縮処理部]

圧縮処理部27は、記憶部12に格納されている絵柄プレーンデータを、JPEG圧縮し、圧縮絵柄プレーンデータを生成する。また、この圧縮処理部27は、記憶部12に格納されている文字線画プレーンデータと、この圧縮絵柄プレーンデータとを連結して一連のデータを生成する。

【0135】

具体的にこの一連のデータとしては、PDF(Portable Document Format)データとすることができます。すなわち、圧縮絵柄プレーンデータを伸長して生成したビットマップ(絵柄プレーンデータのビットマップ)を生成させる指示と、当該絵柄プレーンデータのビットマップ上に文字線画プレーンデータに含まれる各文字や線画を描画させる指示とを含むPDFデータとする。

10

【0136】

ここで文字線画プレーンデータに含まれる各文字や線画を描画させる指示は、文字線画圧縮データと、これに関連する代表色の情報と基本矩形等の座標情報との組を一つずつ読み出し、それぞれの組について、文字線画圧縮データを伸長して二値化画像を生成し、この二値化画像の黒画素の色を代表色に設定し、絵柄プレーンデータのビットマップにおいて、上記基本矩形等の座標情報として設定された位置に透過合成する指示である。ここで透過合成とは、二値化画像のうち黒画素(代表色に設定された画素)以外の画素については上書きせず、代表色に設定された画素のみを上書きすることをいう。

20

【0137】

圧縮処理部27は、この生成したPDFデータを記憶部12に格納し、または画像出力部14にこのPDFデータを出力して、外部の装置に送出させる。

【0138】

[動作]

本実施の形態の画像処理装置は、上述の構成を有しているので、次のように動作する。ここでは図14(a)に示すような文字部分(T1, T2)と、写真部分(P)と、線画部分としての地図部分(M)とを含むドキュメントが画像入力部13から入力され、このドキュメントの画像データを処理対象とする場合を例として説明する。この図14(a)の例においては写真部分(P)内に文字部分の一部(T2)が重ね合わせられている。また地図部分(M)には、道路線図と文字とが入組んでいる。なお、ここでは便宜的に白黒で示しているが、実際には地図部分の道路線図と文字とは互いに異なる色で表され、写真是カラーで構わない。

30

【0139】

前処理部21は、この画像データの画素値を所定色空間(YCbCr)の値に変換する。絵柄候補部分特定処理部22は、この画像データ(元の画像データ)を二値化処理し、その小領域を除去して文字部分(T1)や線図(M)を除去した画像データを生成する(図14(b))。このとき、文字や道路線図のほとんどが除去されるが(部分的に残存しても構わない)、写真部分に重ね合わせられた文字は、そのまま絵柄候補部分として特定された状態となる。

40

【0140】

文字線画抽出処理部23は、元の画像データを二値化し、小領域部分を特定するなどの方法で文字線画部分を抽出する。この際ににおいて、元の画像データを複数の領域に分割し、分割して得られた各領域ごとに二値化のしきい値を適応的に定める方法(特許文献2に開示の方法など)によって二値化を行うことで、地図等着色された部分からも文字の抽出を可能としている(図14(c))。

【0141】

レイアウト処理部24は、絵柄候補部分内でレイアウト解析処理を行い、絵柄候補部分内に残存する文字部分(T2)を抽出する。同一色領域分離部25は、文字線画抽出処理

50

部23やレイアウト処理部24で抽出された文字部分や線画部分について、その代表色を定める。本実施の形態において特徴的なことの一つは、ここで元の画像データが画像入力部13での読み取り誤差や元の画像データのエンコードの特性（例えばJPEGエンコードにおけるモスキートノイズ）によって本来一色であった文字線画部分に複数の色が混ざり合う場合があることに配慮し、これらの画素を平滑化し、さらに輝度成分について暗いものをより暗く、明るい部分をより明るくする補正を行うことである。この補正を平滑化処理とともにを行うことで、平滑化処理によって輝度部分に生じる変動を抑制し、文字や線画の代表色と元の画像データにおける色と違和感のないように設定できるようになる。なお、色差成分も併せて補正することとしてもよい。

【0142】

10

こうして制御部11は、各文字・線画部分（T1, T2, M）のそれについてその領域を画定する座標情報と、その領域内部の画素の代表色を表す情報と、その領域内部で代表色とするべき部分を特定する二値ビットマップデータの圧縮画像とを組とした文字線画プレーンデータを生成する。

【0143】

穴埋処理部26は、絵柄候補部分から、それに内在していた文字部分（T2）を除去した画像データを生成する（図14（d））。なお、本来は文字部分T2の文字形状通りに白抜きとなるのであるが、図14（d）では画面を見やすくするため当該文字形状を含む矩形部分を白抜きにして図示している。また、図14（d）、（e）は絵柄画像部分のみを図示している。そして穴埋処理部26が当該除去された画素の値を、スキャンライン順で最近傍の画素値（除去されていない画素値）に設定して（図14（e））、絵柄プレーンデータを生成する。

20

【0144】

圧縮処理部27は、絵柄プレーンデータについてJPEG圧縮を行い、文字線画プレーンデータと組み合せてPDFデータを生成し、これを画像出力部14に出力する。画像出力部14は、このPDFデータを外部の装置に出力する。

【0145】

30

ここで圧縮処理部27は、絵柄プレーンデータについてJPEG圧縮の前に画像のサイズを縮小する処理（縮小処理）を行って圧縮率をより向上させることとしてもよい。

【0146】

[第2実施形態]

次に、本発明の第2の実施の形態に係る画像処理装置について説明する。本実施の形態の画像処理装置は、図1に示した第1の実施の形態に係る画像処理装置と同様の構成を採るものであるが、制御部11における処理の一部が異なる。

【0147】

40

すなわち、本実施の形態の制御部11が実行する処理は、図15に機能的に示すように、画像入力部13から入力される画像データを処理対象として、この処理対象となった画像データに対して、所定前処理を行う前処理部21と、グレイ変換部61と、しきい値算出部62と、単純二値化部63と、第1レイアウト処理部64と、浮動二値化部65と、処理領域設定部66と、第2レイアウト処理部67と、文字部分統合処理部68と、同一色領域分離部25と、穴埋処理部26と、圧縮処理部27とを含んで構成されている。なお、図2に示した処理と同様の構成となる部分については同一の符号を付して詳細な説明を省略する。

【0148】

グレイ変換部61は、前処理部21が入力する画像データから輝度成分を抽出して、色差成分のないグレースケールの画像データ（グレイ画像データ）を生成する。しきい値算出部62は、グレイ画像データの特定の場所の画素に基づいて二値化のしきい値を決定する。このしきい値算出部62は、例えば四隅の画素の画素値の平均を算出し、当該平均の値を二値化のしきい値として決定する。

【0149】

50

単純二値化部 6 3 は、グレイ画像データの各画素のうち、しきい値算出部 6 2 が決定した二値化のしきい値（第 1 の条件）を越える画素値の画素を白画素、それ以外の画素を黒画素とした第 1 の二値化画像データを生成する。

【 0 1 5 0 】

第 1 レイアウト処理部 6 4 は、第 1 の二値化画像データに対して所定のレイアウト解析処理（第 1 レイアウト解析処理）を行って、第 1 の文字部分特定情報を生成する。この第 1 レイアウト処理部 6 4 の具体的動作については、後に詳しく述べる。

【 0 1 5 1 】

浮動二値化部 6 5 は、グレイ画像データに対して、特許文献 2 に開示された方法によって二値化処理を行い、第 2 の二値化画像データを生成する。例えば、グレイ画像データの各画素を順次注目画素として選択し、注目画素近傍の画素値の平均値に基づいて局所的な二値化しきい値を定め、当該局所的な二値化しきい値よりも当該注目画素の画素値が大きい場合に注目画素を白画素、小さい場合に注目画素を黒画素と設定していく。これにより第 2 の二値化画像データが生成される。

【 0 1 5 2 】

処理領域設定部 6 6 は、第 1 文字部分データ上でオンとなっている画素に対応する、第 2 の二値化画像データ上の画素の値を白画素に変更する。これにより、第 1 レイアウト処理部 6 4 によって文字部分として既に抽出された画素については、以下の処理を行わないよう設定するのである。

【 0 1 5 3 】

第 2 レイアウト処理部 6 7 は、処理領域設定部 6 6 によって処理された第 2 の二値化画像データについて、所定のレイアウト解析処理（第 2 レイアウト解析処理）を行って第 2 の文字部分特定情報を生成する。この第 2 レイアウト処理部 6 7 の具体的動作については、後に詳しく述べる。

【 0 1 5 4 】

文字部分統合処理部 6 8 は、第 1 の文字部分特定情報と第 2 の文字部分特定情報を統合（透過合成）して、統合文字部分特定情報を生成し、記憶部 1 2 に格納する。従って同一色領域分離部 2 5 は、前処理部 2 1 が出力する画像データのうち、この統合文字部分特定情報によって特定される部分が文字部分であるとして処理を行うことになる。なお、ここでの透過合成とは、2つのデータの対応する画素について論理和（いずれか一方でも黒画素であれば、結果を黒画素とする）を演算し、対応する画素の値とする合成方法である。

【 0 1 5 5 】

また、穴埋処理部 2 6 は、前処理部 2 1 が出力する画像データのうち、この統合文字部分特定データによって特定されない部分（統合文字部分特定データの各画素値を反転したデータによって特定される部分）を絵柄部分であるとして処理を行うことになる。

【 0 1 5 6 】

ここで第 1 レイアウト処理部 6 4 と、第 2 レイアウト処理部 6 7 との動作について説明する。これらのレイアウト処理部が行うレイアウト解析処理は、具体的には第 1 の実施の形態に係る制御部 1 1 におけるレイアウト処理部 2 4 の動作とほぼ同様の処理である。

【 0 1 5 7 】

尤も、第 1 レイアウト処理部 6 4 については、絵柄領域を二値化処理によって除去した結果に対して処理が行われると考えられるため、ノイズ判定部 4 8 の処理は必ずしも必要でない。すなわち、第 1 レイアウト処理部 6 4 は、二値化処理部 4 1 と、連結画素抽出部 4 2 と、基本矩形画定部 4 3 と、第 1 セパレータ検出部 4 4 と、行矩形画定部 4 5 と、第 2 セパレータ検出部 4 6 と、文字領域画定部 4 7 と、文字部分特定部 4 9 とを含んで構成される。この場合、文字部分特定部 4 9 は、文字領域画定部 4 7 が生成した文字領域データベースを記憶部 1 2 から読出して、当該文字領域データベースに含まれる文字領域（文字領域の座標情報そのもの）、または当該文字領域内の黒画素部分（文字領域の座標情報と、黒画素部分からなるビットマップ情報）を文字部分として特定し、当該文字部分を特

10

20

30

40

50

定する情報（第1の文字部分特定情報）を記憶部12に格納する。

【0158】

また、第2レイアウト処理部67は、第1の実施の形態に係る制御部11におけるレイアウト処理部24と同様に、二値化処理部41と、連結画素抽出部42と、基本矩形画定部43と、第1セパレータ検出部44と、行矩形画定部45と、第2セパレータ検出部46と、文字領域画定部47と、ノイズ判定部48と、文字部分特定部49とを含んで構成される。この第2レイアウト処理部67の文字部分特定部49が第2の文字部分特定情報を生成して記憶部12に格納する。

【0159】

これらの各部は、いずれも第1の実施の形態において同一の符号を付して説明したものとほぼ同様であるので、繰り返しての説明を省略する。10

【0160】

次に、本実施の形態の画像処理装置の動作について、第1の実施の形態の動作の説明と同様に、図14(a)に示したドキュメントが画像入力部13から入力され、このドキュメントの画像データを処理対象とする場合を例として説明する。

【0161】

前処理部21は、この画像データの画素値を所定色空間(YCbCr)の値に変換する。グレイ変換部61はこの画像データ(元の画像データ)の輝度成分(Y成分)を取り出して、グレイスケールのグレイ画像データを生成する。しきい値算出部62は、グレイ画像データの四隅の画素の画素値の平均を算出し、当該平均の値を二値化のしきい値として決定し、単純二値化部63は、グレイ画像データの各画素のうち、しきい値算出部62が決定した二値化のしきい値(第1の条件)を越える画素値の画素を白画素、それ以外の画素を黒画素とした第1の二値化画像データを生成する。この第1の二値化画像データは、図16(a)に示すようなものとなる。第1レイアウト処理部64は、第1の二値化画像データに対して第1レイアウト解析処理を行って、第1の文字部分特定情報を生成する。この第1レイアウト解析処理では、画素塊ごとの外接矩形のサイズや、各外接矩形間の位置関係やサイズの比較から文字部分が特定される。すなわち図14(a)の例では、T1を含む領域が第1の文字部分特定情報によって特定される(図16(b))。20

【0162】

浮動二値化部65は、グレイ画像データに対して、その各画素を順次注目画素として選択し、注目画素近傍の画素値の平均値に基づいて局所的な二値化しきい値を定め、当該局所的な二値化しきい値よりも当該注目画素の画素値が大きい場合に注目画素を白画素、小さい場合に注目画素を黒画素と設定していく。これにより第2の二値化画像データ(図16(c))が生成される。30

【0163】

図16(c)では、絵柄に含まれる文字部分(T2)が当該絵柄の背景から分離されて黒画素となっている。また、絵柄部分のうち、一部は黒画素となることがある。なお、絵柄以外の文字部分T1等は、単純二値化の結果と同様のものとなる。

【0164】

処理領域設定部66は、第1文字部分データ上でオンとなっている画素に対応する、第2の二値化画像データ上の画素の値、又は第1レイアウト処理部64によって文字部分として画定された領域に対応する領域内の画素を白画素に変更する(図16(d))。第2レイアウト処理部67は、図16(d)に示した、処理領域を設定した後の第2の二値化画像データについて、第2レイアウト解析処理を行って第2の文字部分特定情報を生成する。この処理においては、小さい画素塊などが除去されて絵柄内の文字部分T2が特定され、当該文字部分T2を含む領域が、第2の文字部分特定情報によって画定される(図17(a))。40

【0165】

文字部分統合処理部68は、第1の文字部分特定情報と第2の文字部分特定情報を統合(透過合成)して、統合文字部分特定情報(図17(b))を生成する。この統合文字50

部分特定情報には、従って文字部分 T₁ と T₂ を画定する情報が含まれる。この統合文字部分特定情報は、記憶部 12 に格納される。

【 0 1 6 6 】

同一色領域分離部 25 は、統合文字部分特定情報によって特定された文字部分について、その代表色を定める。そしてこれにより、各文字部分 (T₁, T₂) のそれぞれについてその領域を画定する座標情報と、その領域内部の画素の代表色を表す情報と、その領域内部で代表色とするべき部分を特定する二値ビットマップデータの圧縮画像とを組とした文字プレーンデータを生成する。

【 0 1 6 7 】

穴埋処理部 26 は、前処理部 21 が出力する画像データのうち、この統合文字部分特定データによって特定されない部分を取り出す。図 14 (a) に示した例においては、絵柄部分 P に内在していた文字部分 (T₂) を除去し、この除去処理後の絵柄部分と、線図部分 M とが処理対象となる。そしてこの処理対象部分について、上記除去された画素の値を、スキャンライン順で最近傍の画素値（除去されていない画素値）に設定して、絵柄プレーンデータを生成する。

10

【 0 1 6 8 】

圧縮処理部 27 は、絵柄プレーンデータについて JPEG 圧縮を行い、文字プレーンデータと組み合せて PDF データを生成し、これを画像出力部 14 に出力する。画像出力部 14 は、この PDF データを外部の装置に出力する。

20

【 0 1 6 9 】

ここで圧縮処理部 27 は、絵柄プレーンデータについて JPEG 圧縮の前に画像のサイズを縮小する処理（縮小処理）を行って圧縮率をより向上させることとしてもよい。

【 0 1 7 0 】

なお、ここでは線画を絵柄プレーンデータに含めるようにしているが、例えば単純二値化の結果であって、統合文字部分特定情報によって文字部分として特定されない領域 (M) を線画領域として、穴埋処理部 26 において、前処理部 21 が出力する画像データのうち、統合文字部分特定データ及び線画領域によって特定されない領域（図 14 (a) における領域 P から文字部分 T₂ を除去した領域）を絵柄領域として処理してもよい。

【 0 1 7 1 】

この場合は、当該線画領域を画定する座標情報と、その内部のビットマップデータとを線画プレーンデータとして生成し、圧縮処理部 27 において、ランレンゲス圧縮、又は LZ (Lempel-Ziv) 法等、辞書方式を基礎とした圧縮法（例えば z i p 等のフレート (flat e) 圧縮）を行い、文字プレーンデータと線画プレーンデータと絵柄プレーンデータとを含む PDF データを生成し、これを画像出力部 14 に出力することとしてもよい。

30

【 図面の簡単な説明 】

【 0 1 7 2 】

【図 1】本発明の第 1 の実施の形態に係る画像処理装置の一例を表す構成ブロック図である。

【図 2】本発明の第 1 の実施の形態に係る画像処理装置の制御部によって実行される処理内容を表す機能ブロック図である。

40

【図 3】絵柄候補部分特定処理部 22 の処理内容例を表す機能ブロック図である。

【図 4】属性判定部 34 の処理例を表すフローチャート図である。

【図 5】絵柄候補部分特定処理部 22 の処理例を表す説明図である。

【図 6】レイアウト処理部 24 の処理内容例を表す機能ブロック図である。

【図 7】レイアウト処理部 24 の処理例を表す説明図である。

【図 8】レイアウト処理部 24 の処理例を表すフローチャート図である。

【図 9】レイアウト処理部 24 の処理例を表すフローチャート図である。

【図 10】同一色領域分離部 25 の処理内容例を表す機能ブロック図である。

【図 11】同一色領域分離部 25 の処理例を表す説明図である。

【図 12】同一色領域分離部 25 において利用されるトーンカーブの例を表す説明図であ

50

る。

【図13】同一色領域分離部25における補正処理の処理条件を表す説明図である。

【図14】本発明の第1の実施の形態に係る画像処理装置の処理例を表す説明図である。

【図15】本発明の第2の実施の形態に係る画像処理装置の制御部によって実行される処理内容を表す機能ブロック図である。

【図16】本発明の第2の実施の形態に係る画像処理装置の処理例を表す説明図である。

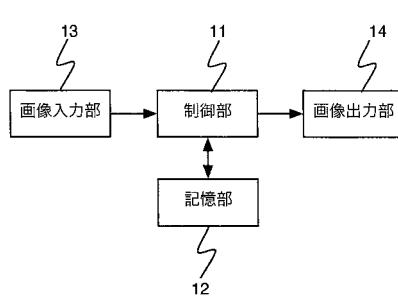
【図17】本発明の第2の実施の形態に係る画像処理装置の処理例を表す説明図である。

【符号の説明】

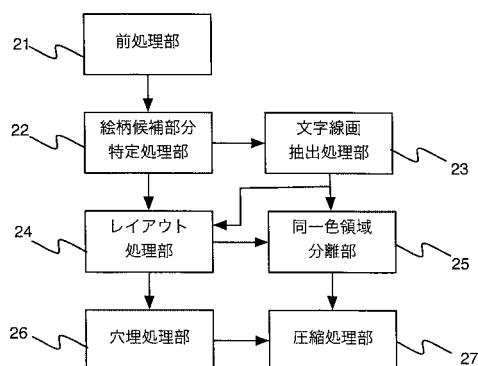
【0173】

11 制御部、12 記憶部、13 画像入力部、14 画像出力部、21 前処理部
 、22 絵柄候補部分特定処理部、23 文字線画抽出処理部、24 レイアウト処理部
 、25 同一色領域分離部、26 穴埋処理部、27 圧縮処理部、31, 41 二値化
 処理部、32, 42 連結画素抽出部、33 特徴量算出部、34 属性判定部、35
 非絵柄領域処理部、36 背景領域塗潰部、37 絵柄候補領域作成部、38 膨張収縮
 部、43 基本矩形画定部、44 第1セパレータ検出部、45 行矩形画定部、46
 第2セパレータ検出部、47 文字領域画定部、48 ノイズ判定部、49 文字部分特
 定部、51 色数判定部、52 第1代表色決定部、53 第2代表色決定部、54 限
 定色化部、61 グレイ変換部、62 しきい値算出部、63 単純二値化部、64 第
 1レイアウト処理部、65 浮動二値化部、66 処理領域設定部、67 第2レイアウ
 ツ処理部、68 文字部分統合処理部。 10
 20

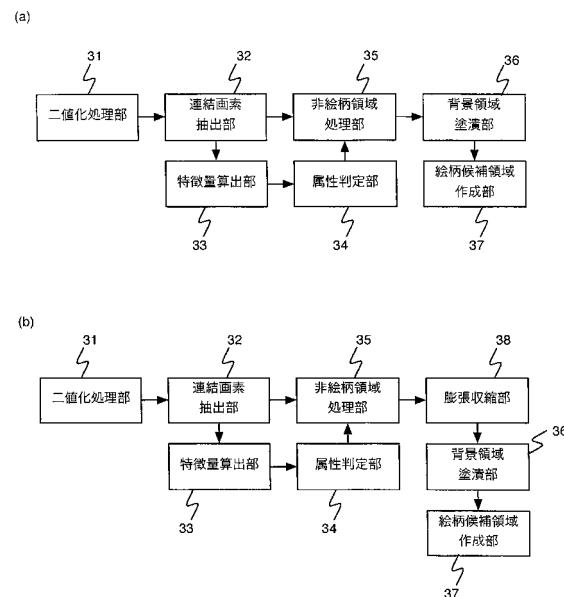
【図1】



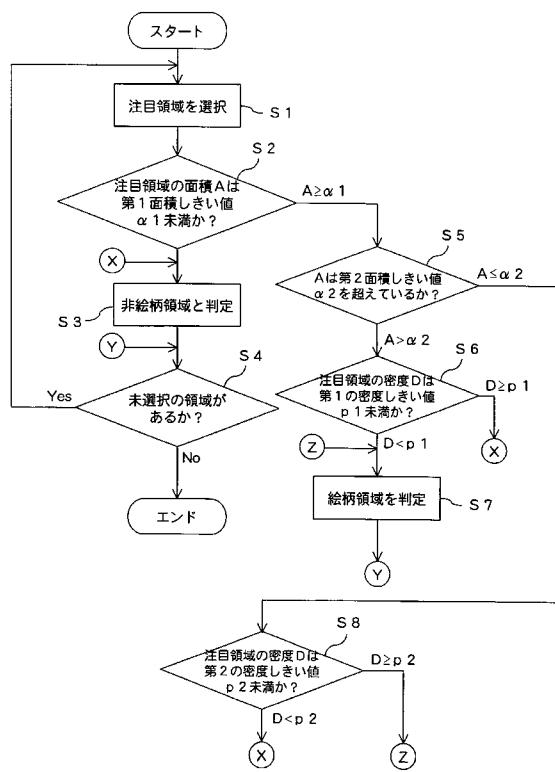
【図2】



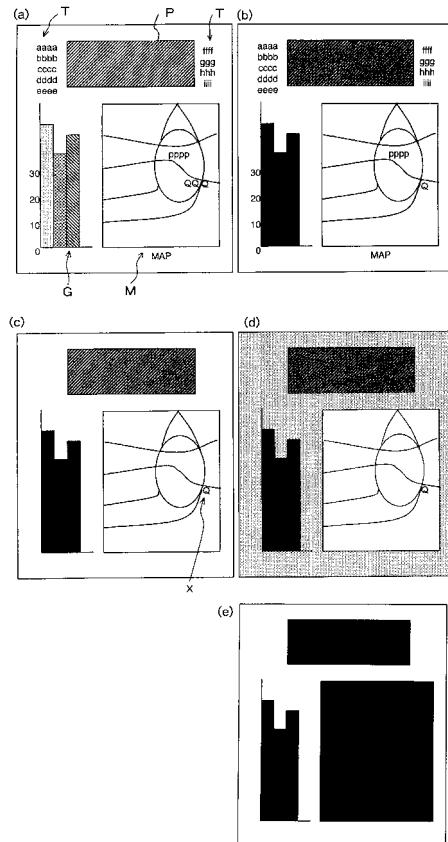
【図3】



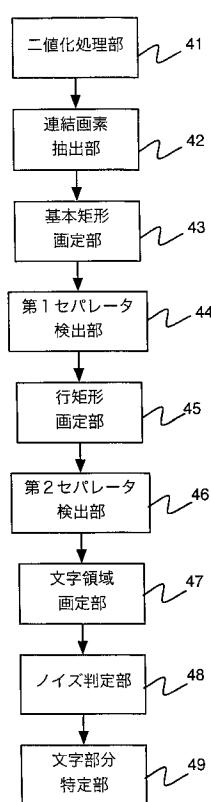
【図4】



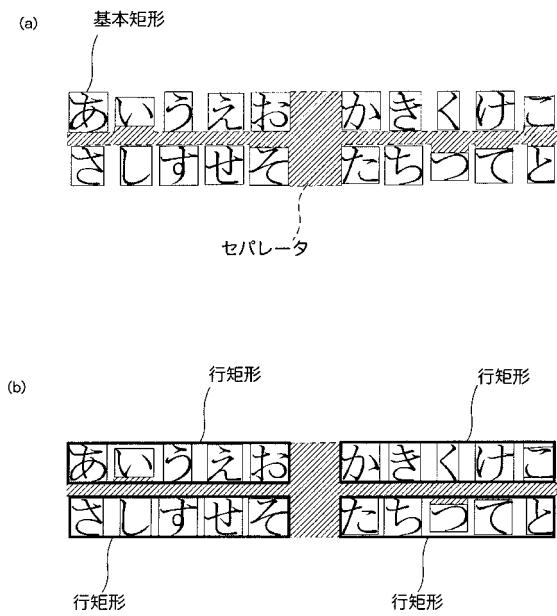
【図5】



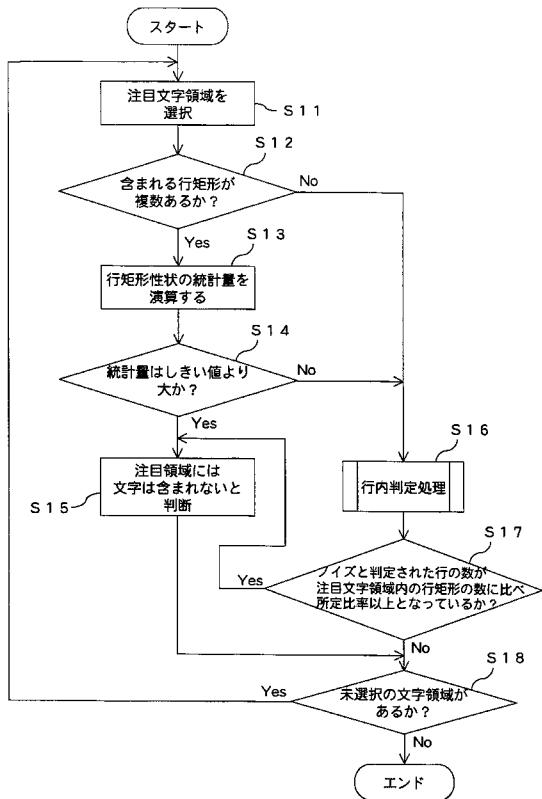
【図6】



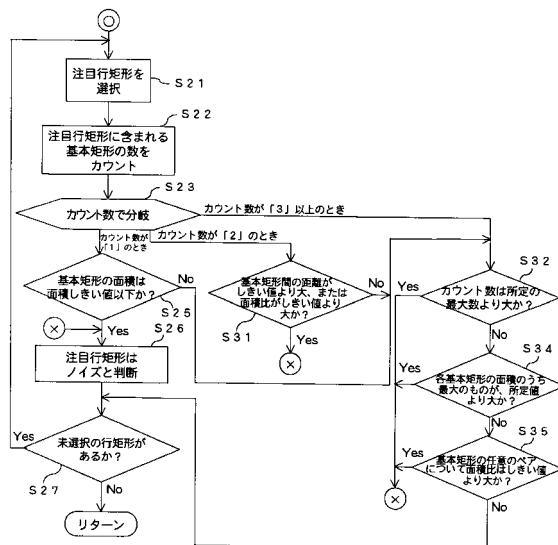
【図7】



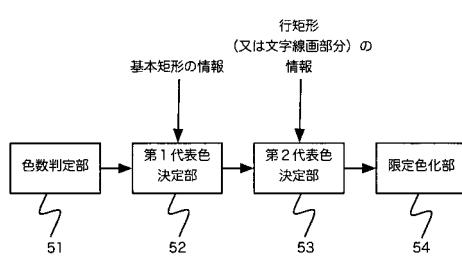
【図 8】



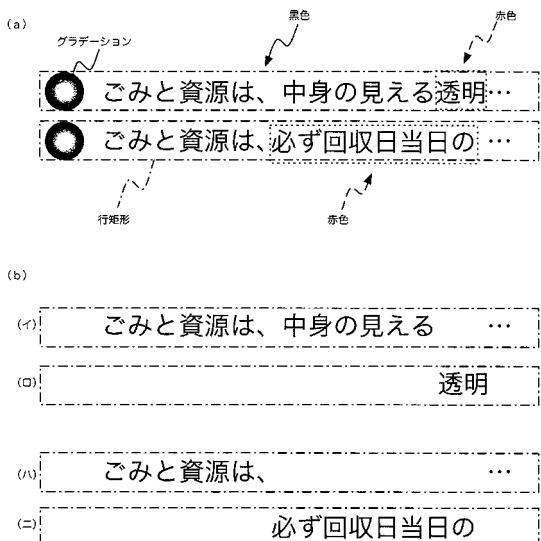
【図 9】



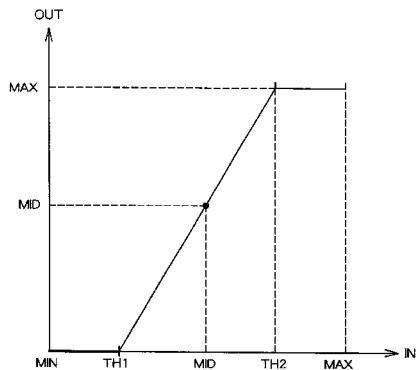
【図 10】



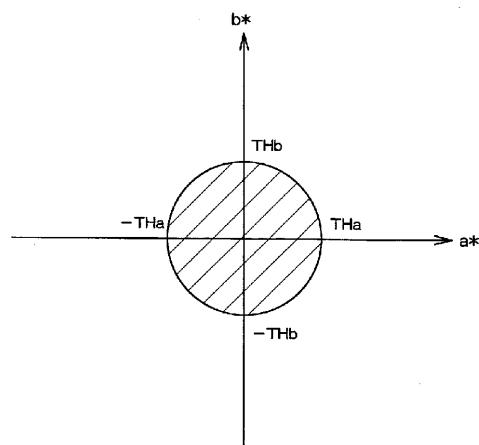
【図 11】



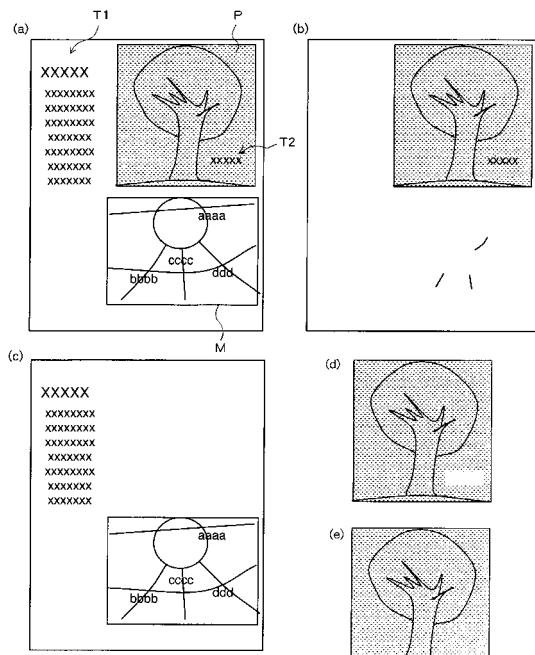
【図12】



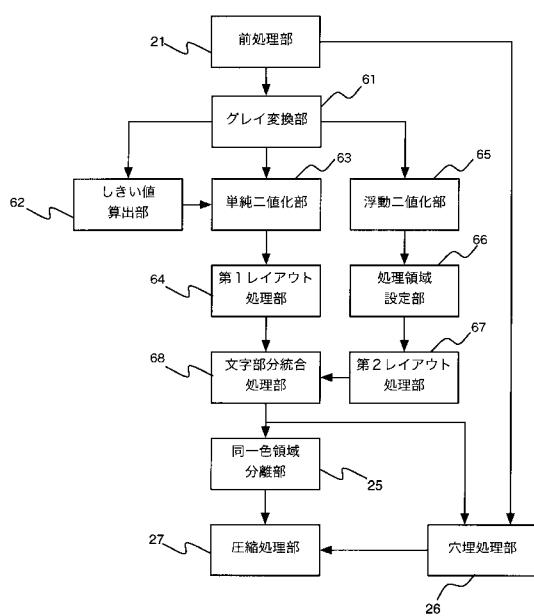
【図13】



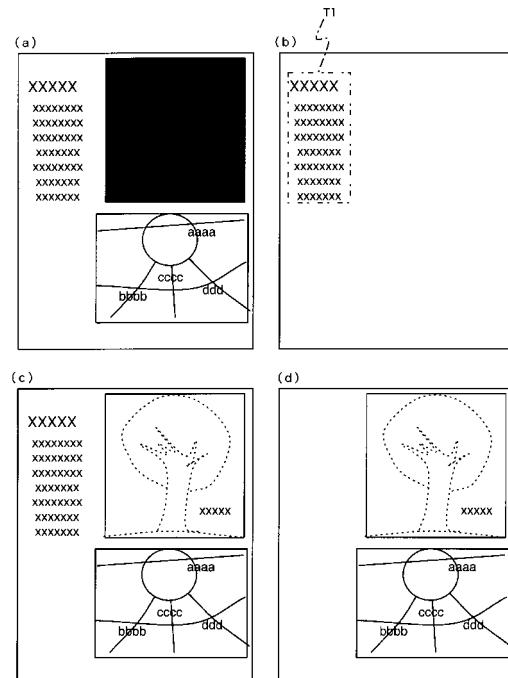
【図14】



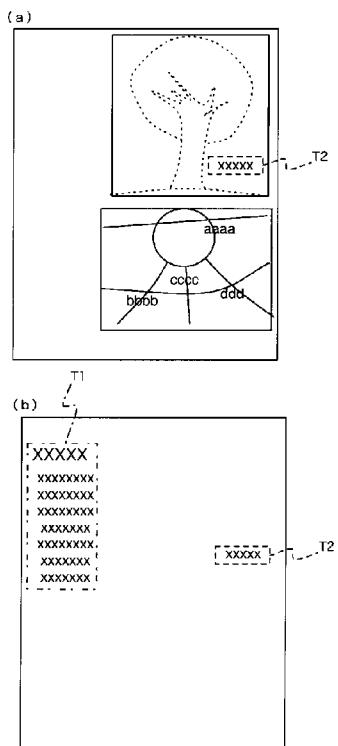
【図15】



【図16】



【図17】



フロントページの続き

(56)参考文献 特開2002-300408(JP,A)

特開平06-187490(JP,A)

特開平10-31716(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04N 1/413

H04N 1/40