

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第4579759号  
(P4579759)

(45) 発行日 平成22年11月10日(2010.11.10)

(24) 登録日 平成22年9月3日(2010.9.3)

(51) Int.Cl.	F 1
HO4N 1/413 (2006.01)	HO4N 1/413 D
GO6T 1/00 (2006.01)	GO6T 1/00 310Z
GO6T 7/40 (2006.01)	GO6T 7/40 100A
HO4N 1/40 (2006.01)	HO4N 1/40 F

請求項の数 13 (全 19 頁)

(21) 出願番号	特願2005-124984 (P2005-124984)	(73) 特許権者	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成17年4月22日 (2005.4.22)	(74) 代理人	100126240 弁理士 阿部 琢磨
(65) 公開番号	特開2006-304062 (P2006-304062A)	(74) 代理人	100124442 弁理士 黒岩 創吾
(43) 公開日	平成18年11月2日 (2006.11.2)	(72) 発明者	三沢 玲司 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
審査請求日	平成20年4月11日 (2008.4.11)	審査官	富永 達朗

最終頁に続く

(54) 【発明の名称】 画像処理装置、画像処理方法、コンピュータプログラム

## (57) 【特許請求の範囲】

## 【請求項 1】

文書画像から判定対象領域を抽出する領域抽出手段と、  
前記領域抽出手段で抽出された判定対象領域内の閉ループ数を算出する閉ループ数算出手段と、

前記判定対象領域から前記閉ループ数と異なる特徴量を抽出する特徴量抽出手段と、  
前記閉ループ数算出手段により算出された閉ループ数が第1の閾値より小さい場合、当該判定対象領域を文字領域と判定し、

前記閉ループ数算出手段により算出された閉ループ数が第2の閾値(ここで第2の閾値  
>第1の閾値とする)より大きい場合、当該判定対象領域を写真領域と判定し、

前記閉ループ数算出手段により算出された閉ループ数が前記第1の閾値以上で且つ前記  
第2の閾値以下である場合、前記特徴量抽出手段で抽出された特徴量に基づいて当該判定  
対象領域が文字領域か否かを判定する領域判定手段と、

を備えることを特徴とする画像処理装置。

## 【請求項 2】

前記文書画像内に含まれる画素塊に基づいて、前記画像内から文字領域候補を判定する  
プレ判定手段を更に有し、

前記領域抽出手段では、前記プレ判定手段で判定された文字領域候補に含まれる文字矩  
形を、前記判定対象領域として抽出することを特徴とする請求項1に記載の画像処理装置  
。

**【請求項 3】**

多値画像から 2 値画像を生成する 2 値化手段を更に有する画像処理装置であって、

前記領域抽出手段での抽出対象となる文書画像は、前記 2 値化手段により生成された 2 値画像であることを特徴とする請求項 1 または 2 に記載の画像処理装置。

**【請求項 4】**

前記閉ループ数算出手段で算出される閉ループ数は、前記判定対象領域内の黒画素、または、白画素の輪郭線の数であることを特徴とする請求項 3 に記載の画像処理装置。

**【請求項 5】**

前記特徴量抽出手段で抽出される特徴量は、当該判定対象領域の色分散値であることを特徴とする請求項 1 乃至 4 のいずれかに記載の画像処理装置。

10

**【請求項 6】**

前記領域判定手段により文字領域と判定された領域に対して第 1 の圧縮処理を施し、前記文字領域と判定された領域以外の画像に対して第 2 の圧縮処理を施すことにより、前記文書画像の圧縮データを得る圧縮手段を、更に有することを特徴とする請求項 1 乃至 5 のいずれかに記載の画像処理装置。

**【請求項 7】**

前記第 1 の圧縮処理は 2 値画像に適した圧縮処理であり、前記第 2 の圧縮処理は多値画像に適した圧縮処理であることを特徴とする請求項 6 に記載の画像処理装置。

**【請求項 8】**

前記領域判定手段により文字領域と判定された領域から代表色を算出する代表色算出手段を更に有する画像処理装置であって、

20

前記圧縮データには、前記第 1 の圧縮処理により得られる第 1 圧縮コードと、前記第 2 の圧縮処理により得られる第 2 圧縮コードと、前記代表色算出手段で得られる代表色の情報とが含まれることを特徴とする請求項 6 に記載の画像処理装置。

**【請求項 9】**

前記第 2 の圧縮処理を施す対象の画像は、前記文書画像において前記文字領域と判定された領域の部分を周囲の色で埋めることにより生成された文書画像であることを特徴とする請求項 6 に記載の画像処理装置。

**【請求項 10】**

前記領域判定手段で文字領域と判定された領域に対して、文字認識処理を実行する文字認識処理手段を更に備えることを特徴とする請求項 1 乃至 5 のいずれかに記載の画像処理装置。

30

**【請求項 11】**

領域抽出手段が、文書画像から判定対象領域を抽出する領域抽出ステップと、

閉ループ数算出手段が、前記領域抽出ステップで抽出された判定対象領域内の閉ループ数を算出する閉ループ数算出手段と、

特徴量抽出手段が、前記判定対象領域から前記閉ループ数と異なる特徴量を抽出する特徴量抽出ステップと、

領域判定手段が、前記閉ループ数算出手段で算出された閉ループ数が第 1 の閾値より小さい場合、当該判定対象領域を文字領域と判定し、

40

前記閉ループ数算出手段で算出された閉ループ数が第 2 の閾値（ここで第 2 の閾値 > 第 1 の閾値とする）より大きい場合、当該判定対象領域を写真領域と判断し、

前記閉ループ数算出手段で算出された閉ループ数が前記第 1 の閾値以上で且つ前記第 2 の閾値以下である場合、前記特徴量抽出ステップで抽出された特徴量に基づいて当該判定対象領域が文字領域か否かを判定する領域判定ステップと、

を備えることを特徴とする、画像処理装置による画像処理方法。

**【請求項 12】**

コンピュータを、

文書画像から判定対象領域を抽出する領域抽出手段、

前記領域抽出手段で抽出された判定対象領域内の閉ループ数を算出する閉ループ数算出

50

手段、

前記判定対象領域から前記閉ループ数と異なる特徴量を抽出する特徴量抽出手段、

前記閉ループ数算出手段により算出された閉ループ数が第1の閾値より小さい場合、当該判定対象領域を文字領域と判定し、

前記閉ループ数算出手段により算出された閉ループ数が第2の閾値（ここで第2の閾値 > 第1の閾値とする）より大きい場合、当該判定対象領域を写真領域と判断し、

前記閉ループ数算出手段により算出された閉ループ数が前記第1の閾値以上で且つ前記第2の閾値以下である場合、前記特徴量抽出手段で抽出された特徴量に基づいて当該判定対象領域が文字領域か否かを判定する領域判定手段、

として機能させるためのコンピュータプログラム。

10

【請求項1 3】

請求項1 2に記載のコンピュータ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】

【技術分野】

【0 0 0 1】

本発明は、文書画像の領域判定を好適に行うことができる画像処理装置、画像処理方法、コンピュータプログラムに関する。

【背景技術】

【0 0 0 2】

20

近年、カラープリンタやカラースキャナ等の普及により、カラー化された文書が増え、これをスキャンにより取り込んで電子ファイルとして保存したり、インターネット等を介して第三者等に送付する機会が増えてきている。しかし、フルカラーデータのままでは記憶装置や回線への負荷が大きいため、圧縮処理等の方法で取り扱うデータ量を小さくする必要がある。

【0 0 0 3】

従来、カラー画像を圧縮する方法として、例えば、誤差拡散等で擬似階調を持った2値画像にして圧縮する方法、JPEG形式で圧縮する方法、8ビットのパレットカラーに変換を行ってZIP圧縮やLZW圧縮をする方法等があった。また、領域判定とMMRによる2値圧縮とZIPによる可逆圧縮と、JPEGによる非可逆圧縮との組み合わせにより、通常の文字領域については高い品位が得られる圧縮方法等があった（例えば、特許文献1及び特許文献2参照）。

30

【0 0 0 4】

また、従来、文書画像処理に関する技術としては、文書を光学的に入力し、文字を認識してテキストコードを出力する光学的文書認識装置（OCR）の技術が存在する（例えば、特許文献3参照）。

【0 0 0 5】

OCRでは、濃度射影（ヒストグラム）により文字行を切り出し（抽出）、さらに1文字単位の文字ブロック切り出し（抽出）を行う。文字ブロックの切り出しに際しては、文字行方向に濃度射影を取り、濃度射影値の変化に基づいて文字行を分離し、さらに、各文字行について、文字行と垂直方向に濃度射影を取ることで個々の文字ブロックを抽出する。また必要に応じて、標準的な文字サイズや文字ピッチの推定値、および行と垂直方向に濃度射影値等の情報を用いて、1文字単位の文字画像となる、最終的な文字ブロックを切り出す。切り出された文字ブロックは、縦横寸法の正規化を行った後に、所定の特徴データ抽出の処理が施される。特徴データが抽出された個々の文字ブロックは、求められている標準パターンとの類似度が計算され、最も類似度の高い文字が認識結果とされる。標準パターンの集合は認識辞書と呼ばれる。

40

【特許文献1】特開2002-077633号公報

【特許文献2】特開2004-128880号公報

【特許文献3】特開2003-346083号公報

50

**【発明の開示】****【発明が解決しようとする課題】****【0006】**

特許文献1や特許文献2に記載の方法によれば、領域判定とMMRによる2値圧縮とZIPによる可逆圧縮と、JPEGによる非可逆圧縮との組み合わせにより、通常の文字領域については高い品位が得られる。しかし、領域判定の結果において、文字ではない領域（写真領域など、以下、非文字）を文字であると誤判断する場合もあり、その際は逆に大きな画質劣化を生じるという問題があった。

**【0007】**

また、OCR処理においては、文字ブロックとして切り出した領域が非文字であった場合、非文字に対して文字認識を行うことになる。非文字に対して文字認識を行ってしまうと、全体の処理速度を低下させてしまう他、意味のないテキストコードが認識結果の出力データに含まれてしまう場合もあり好ましくないという問題があった。

10

**【0008】**

本発明は、このような事情を考慮してなされたものであり、抽出した領域に対して、文字と非文字の属性判断を良好に行うことができる画像処理装置、画像処理方法、コンピュータプログラム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】****【0009】**

上記課題を解決する為に、本発明の画像処理装置は、文書画像から判定対象領域を抽出する領域抽出手段と、前記領域抽出手段で抽出された判定対象領域内の閉ループ数を算出する閉ループ数算出手段と、前記判定対象領域から前記閉ループ数と異なる特徴量を抽出する特徴量抽出手段と、前記閉ループ数算出手段により算出された閉ループ数が第1の閾値より小さい場合、当該判定対象領域を文字領域と判定し、前記閉ループ数算出手段により算出された閉ループ数が第2の閾値（ここで第2の閾値 > 第1の閾値とする）より大きい場合、当該判定対象領域を写真領域と判断し、前記閉ループ数算出手段により算出された閉ループ数が前記第1の閾値以上で且つ前記第2の閾値以下である場合、前記特徴量抽出手段で抽出された特徴量に基づいて当該判定対象領域が文字領域か否かを判定する領域判定手段と、を備えることを特徴とする。

20

**【0010】**

上記課題を解決する為に、本発明の画像処理方法は、領域抽出手段が、文書画像から判定対象領域を抽出する領域抽出ステップと、閉ループ数算出手段が、前記領域抽出ステップで抽出された判定対象領域内の閉ループ数を算出する閉ループ数算出ステップと、特徴量抽出手段が、前記判定対象領域から前記閉ループ数と異なる特徴量を抽出する特徴量抽出ステップと、領域判定手段が、前記閉ループ数算出ステップで算出された閉ループ数が第1の閾値より小さい場合、当該判定対象領域を文字領域と判定し、前記閉ループ数算出ステップで算出された閉ループ数が第2の閾値（ここで第2の閾値 > 第1の閾値とする）より大きい場合、当該判定対象領域を写真領域と判断し、前記閉ループ数算出ステップで算出された閉ループ数が前記第1の閾値以上で且つ前記第2の閾値以下である場合、前記特徴量抽出ステップで抽出された特徴量に基づいて当該判定対象領域が文字領域か否かを判定する領域判定ステップと、を備えることを特徴とする。

30

**【発明の効果】****【0011】**

本発明によれば、文字と非文字の領域判定を精度良く実行することが可能となる。したがって、この領域判定結果を圧縮技術に適用すると、良好な画質が得られると共に、圧縮効率を向上させることができるとなる。また、OCR技術に対して適用すると、処理速度の向上と共に、意味のないテキストコードを出力してしまうことを抑え、認識率を向上させることができる。

40

**【発明を実施するための最良の形態】****【0012】**

50

## (実施例 1 )

以下の実施の形態では、例えばカラー複写機に搭載可能な、カラー画像の圧縮技術において本発明の領域判定技術を適用する例を説明する。尚、カラー複写機の機能としては、例えば、カラーコピー機能、カラープリント機能及びカラースキャナ機能があるが、本実施形態で説明する領域判定技術は、このうちカラーコピー機能及びカラースキャナ機能で適用可能である。具体的には、カラー原稿を読み取ることにより得られたカラー画像データを圧縮する際に用いられる圧縮技術に適用できる。また、カラースキャナ機能としては、例えば、カラー原稿を読み取ることにより得られたカラー画像データを圧縮して外部へ送信するデータ送信機能及び同カラー画像データを圧縮して複写機内部の記憶手段に記憶する保存機能がある。

10

## 【 0 0 1 3 】

以下、図面を参照して、本発明を好適な実施形態に従って詳細に説明する。

## 【 0 0 1 4 】

図1は本発明の実施例によるシステム構成を示す概略図であり、ネットワーク通信機能を備えた複合機(MFP)101とホストコンピュータ(以下、PC)102が、ネットワーク103等の伝送媒体で接続された環境を示す図である。

## 【 0 0 1 5 】

また、点線104～105は、処理／制御の流れを示すものとし、以下順に説明を行う。104は、ユーザがMFP101のスキャナより紙文書を読み込ませる処理を示す。その際、ユーザは、後述するMFP101のユーザーインターフェース(図2の203)より、送信する宛先(例えば、PC102)、各種送信設定(例えば、解像度、圧縮率)、データ書式(例えば、JPEG、TIFF、PDF、PDF高压縮、PDF(OCR結果付き))を予め指定する。本実施例では、カラー画像の圧縮技術において本発明の領域判定方法を用いる例を説明するため、データ書式としてPDF高压縮を指定した場合について説明を行う。PDF高压縮の技術詳細については後述する。105は、指定された送信設定及びデータ書式に基づいて、MFP101の後述するソフトウェアあるいはハードウェア機能を利用してデータを生成し、指定された宛先に送信する処理を示す。ここで、PC102へ送信された画像は、PDFなどのファイルフォーマットで送信されることになるので、PC102の有する汎用的なビューアで閲覧可能である。

20

## 【 0 0 1 6 】

30

次に、図1におけるMFP101のハードウェアの詳細構成について図2を用いて説明する。

## 【 0 0 1 7 】

MFP101は、画像入力デバイスであるスキャナ部201、画像出力デバイスであるプリンタ部202、CPUやメモリ等で構成される制御ユニット(Controller Unit)204、ユーザーインターフェースである操作部203等を有する。制御ユニット204は、スキャナ部201、プリンタ部202、操作部203と接続し、一方では、LAN219や一般的の電話回線網である公衆回線(WAN)220と接続することで、画像情報やデバイス情報の入出力を行うコントローラである。CPU205はシステム全体を制御するコントローラである。RAM206はCPU205が動作するためのシステムワークメモリであり、画像データを一時記憶するための画像メモリでもある。ROM210はブートROMであり、システムのブートプログラムが格納されている。HDD211はハードディスクドライブで、システム制御ソフトウェア、画像データを格納する。操作部I/F207は操作部(UI)203とのインターフェース部で、操作部203に表示するための画像データを操作部203に対して出力する。また、操作部203から本画像処理装置の使用者が入力した情報を、CPU205に伝える役割をする。ネットワーク(Network)208は本画像処理装置をLAN219に接続し、パケット形式の情報の入出力を行う。モデム(MODEM)209は本画像処理装置を公衆回線220に接続し、情報の復調・変調を行い入出力を行う。以上のデバイスがシステムバス221上に配置される。

40

50

## 【0018】

イメージバスインターフェース( Image Bus I/F ) 212はシステムバス221と画像データを高速で転送する画像バス222とを接続し、データ構造を変換するバスブリッジである。画像バス222は、例えば、PCIバスやIEEE1394で構成される。

## 【0019】

画像バス222上には以下のデバイスが配置される。ラスターイメージプロセッサ( RIP) 213はPDLコードを解析し、ピットマップイメージに展開する。デバイスI/F部214は、信号線223を介して画像入出力デバイスであるスキャナ部201、信号線224を介してプリンタ部202、をそれぞれ制御ユニット204に接続し、画像データの同期系／非同期系の変換を行う。スキャナ画像処理部215は、入力画像データに対し補正、加工、編集を行う。プリンタ画像処理部216は、プリンタ部202に出力すべきプリント出力画像データに対して、プリンタ部202に応じた補正、解像度変換等を行う。画像回転部217は入力された画像データの回転を行い出力する。画像圧縮部218は、多値画像データに対してはJPEG圧縮伸長処理、または、デバイス固有の圧縮伸長処理を行い、2値画像データに対してはJBIG、MMR、MHの圧縮伸長処理を行う。以上が図1におけるMFP101のハードウェアの詳細構成である。

## 【0020】

次に、図2における制御ユニット204に実装されるソフトウェア構成について図3を用いて説明する。301はユーザーインターフェース(以下、UI)であり、オペレータが操作部203を用いてMFPに対する各種操作・設定を行う際の、機器とユーザ操作との仲介を行うモジュールである。本モジュールは、オペレータの操作に従い、後述の各種モジュールに入力情報を転送して処理の依頼、或いはデータの設定等を行う。

## 【0021】

302はアドレスブック(Address-Book)、即ちデータの送付先、通信先等を管理するデータベースモジュールである。アドレスブック302の内容は操作部203からの操作を、UI301で検知し、データの追加、削除、取得が行われ、オペレータの操作により後述の各モジュールにデータの送付・通信先情報を与えるものとして使用されるものである。

## 【0022】

303はWebサーバモジュール(Web-Serverモジュール)であり、Webクライアント(例えば、PC102)からの要求により、本MFPの管理情報を通知するために使用される。この管理情報は、後述の統合送信部(Universal-Sendモジュール)304、後述のリモートコピースキャンモジュール(Remote-Copy-Scanモジュール)309、後述のリモートコピープリントモジュール(Remote-Copy-Printモジュール)310、後述の制御API(Control-API)318を介して読み取られ、後述のHTTPモジュール312、TCP/IP通信モジュール316、ネットワークドライバ(Network-Driver)317を介してWebクライアントに通知される。Webサーバモジュール303はWebクライアントに渡すべき情報を、HTML形式等のいわゆるWebページ(ホームページ)形式のデータとして作成する。必要に応じてJava(登録商標)やCGIプログラム等が用いられる。

## 【0023】

304は統合送信部(Universal-Sendモジュール)、即ちデータの配信を司るモジュールであり、UI301を介してオペレータによって指定されたデータを、指示された通信(出力)先に配布するものである。また、オペレータにより、本MFPのスキャナ機能を使用して配布データの生成が指示された場合は、後述の制御API318を介して本MFPのスキャナ201を動作させ、データの生成を行う。

## 【0024】

305は統合送信部304内で出力先にプリンタが指定された際に実行されるモジュ

10

20

30

40

50

ルである。306は統合送信部304内で通信先にE-mailアドレスが指定された際に実行されるモジュールである。307は統合送信部304内で出力先にデータベースが指定された際に実行されるモジュールである。308は統合送信部304内で出力先に本MFPと同様のMFPが指定された際に実行されるモジュールである。

#### 【0025】

309はリモートコピースキャン( Remote - Copy - Scan )モジュールであり、MFP101のスキャナ機能を使用してスキャナ201で読み取った画像情報の出力先をネットワーク等で接続された他のMFPのプリンタで出力し、本MFP101単体で実現しているコピー機能と同等の処理を行うモジュールである。310はリモートコピー・プリント( Remote - Copy - Print )モジュールであり、ネットワーク等で接続された他のMFPのスキャナで読み取った画像情報を入力元として得られた画像情報をMFP101のプリンタ機能を使用して出力することにより、同様にMFP101単体で実現しているコピー機能と同等の処理を行うモジュールである。ボックスモジュール311はスキャン画像もしくはPDLプリント画像をHDDに格納し、格納した画像のプリンタ機能による印刷、統合送信( Universal - Send )機能による送信、HDDに格納した文書の削除、グルーピング(個別BOXへの格納)、BOX間移動、BOX間コピーなどの管理機能を提供する。なお、ボックスモジュール311は、HTTPモジュール312及びTCP/IPモジュール316によって通信機能が提供される。

#### 【0026】

312はHTTPモジュールであり、本MFPがHTTPにより通信する際に使用され、後述のTCP/IP通信モジュール316により前述のWebサーバモジュール303、Webブループリントモジュール311に通信機能を提供する。313はlprモジュールであり、後述のTCP/IP通信モジュール316により前述の統合送信部304内のプリンタモジュール305に通信機能を提供するものである。314はSMTPモジュールであり、後述のTCP/IP通信モジュール316により統合送信部304内のE-mailモジュール306に通信機能を提供する。315はSLM、即ちSalutation - Managerモジュールであり、後述のTCP/IP通信316モジュールにより前述の統合送信部304内のデータベースモジュール317、DPMモジュール318、及びリモートコピースキャンモジュール309、リモートコピー・プリントモジュール310に通信機能を提供する。

#### 【0027】

316はTCP/IP通信モジュールであり、後述のネットワークドライバ316を用いて、前述の各種モジュールにネットワーク通信機能を提供する。317はネットワークドライバであり、ネットワークに物理的に接続される部分を制御するものである。

#### 【0028】

318は制御APIであり、統合送信部304等の上流モジュールに対し、後述のジョブマネージャ( Job - Manager )319等の下流モジュールとのインターフェイスを提供するものであり、上流及び下流のモジュール間の依存関係を軽減し、それぞれの流用性を高めるものである。319はジョブマネージャであり、前述の各種モジュールより制御API318を介して指示される処理を解釈し、後述の各モジュール( 320、324、326 )に指示を与えるものである。また、ジョブマネージャ319は、FAXジョブの制御も含め本MFP内で実行される種々のジョブを一元管理するものである。

#### 【0029】

320はコーデックマネージャ( CODEC - Manager )であり、ジョブマネージャ319が指示する処理の中でデータの各種圧縮・伸長を管理・制御するものである。321はFBEエンコーダモジュール( FBE - Encoder )であり、ジョブマネージャ319、後述のスキャンマネージャ( Scan - Manager )324により実行されるスキャン処理により読み込まれたデータをFBEフォーマットにより圧縮するものである。322はJPEGコーデックモジュール( JPEG - CODEC )であり、ジョブマネージャ319、スキャンマネージャ324により実行されるスキャン処理、及びブ

10

20

30

40

50

リントマネージャ (Print - Manager) 326 により実行される印刷処理において、読み込まれたデータの JPEG 圧縮及び印刷データの JPEG 展開処理を行うものである。323 は MMR コーデック (MMR - CODEC) であり、ジョブマネージャ 319、スキャンマネージャ 324 により実行されるスキャン処理、及びプリントマネージャ 326 により実行される印刷処理において、スキャナから読み込まれたデータの MMR 圧縮及びプリンタへ出力すべき印刷データの MMR 伸長処理を行うものである。

#### 【0030】

324 はスキャンマネージャ (Scan - Manager) であり、ジョブマネージャ 319 が指示するスキャン処理を管理・制御するものである。325 は SCSI ドライバであり、スキャンマネージャ 324 と本 MFP が内部的に接続しているスキャナ部 201 との通信を行うものである。326 はプリントマネージャ (Print - Manager) であり、ジョブマネージャ 319 が指示する印刷処理を管理・制御するものである。327 はエンジンインターフェース (Engine - I / F) であり、プリントマネージャ 326 とプリンタ部 202 との I / F を提供する。328 はパラレルポートドライバであり、Web ブルプリント 311 がパラレルポートを介して不図示の出力機器にデータを出力する際の I / F を提供する。

#### 【0031】

次に Address - Book 302 の詳細について説明する。この Address - Book 302 は、MFP 101 内の不揮発性の記憶装置 (不揮発性メモリやハードディスクなど) に保存されており、この中には、ネットワークに接続された他の機器の特徴が記載されている。例えば、以下に列挙するようなものが含まれている。

- ・機器の正式名やエイリアス名 ,
- ・機器のネットワークアドレス ,
- ・機器の処理可能なネットワークプロトコル ,
- ・機器の処理可能なドキュメントフォーマット ,
- ・機器の処理可能な圧縮タイプ ,
- ・機器の処理可能なイメージ解像度 ,
- ・プリンタ機器の場合の給紙可能な紙サイズや給紙段情報 ,
- ・サーバ (コンピュータ) 機器の場合のドキュメントを格納可能なフォルダ名

以下に説明する各アプリケーションは、上記 Address - Book 302 に記載された情報により配信先の特徴を判別することが可能となる。

#### 【0032】

この Address - Book 302 を参照して、MFP 101 はデータを送信することができる。例えば、リモートコピーアプリケーションは、配信先に指定された機器の処理可能な解像度情報を前記 Address - Book 302 より判別し、それに従い、スキャナにより読み込まれた画像 2 値画像を公知の MMR 圧縮を用いて圧縮し、それを公知の TIFF (Tagged Image File Format) 化し、SLM 303 に通して、ネットワーク上のプリンタ機器に送信する。SLM 303 については、詳細には説明しないが、公知の Salutation - Manager と呼ばれる機器制御情報などを含んだネットワークプロトコルの一種である。

#### 【0033】

次に、図 1 におけるホストコンピュータ 102 のハードウェア構成について図 4 を用いて説明する。ホストコンピュータ 102 については、一般的なパーソナルコンピュータの構成、機能を有しており、周辺機器であるモニタ 401、キーボード・マウス 402 と、ホストコンピュータ 102 全体を制御する中央処理装置 CPU 403、アプリケーションやデータを記憶するハードディスク 405、メモリ 406 等からなる。また、ネットワーク・インターフェース 406 を介してネットワーク 103 等の伝送媒体に接続されている。

#### 【0034】

次に、前述した PDF 高圧縮に関して、図 5 及び図 6 を用いて説明する。

#### 【0035】

10

20

30

40

50

ここでいうPDF高压縮とは、カラー画像の圧縮技術であり、領域判定を行い、各領域の属性に応じて、MMRによる2値圧縮とJPEGによる非可逆圧縮とを適応的に変えて圧縮することにより、圧縮率を高くできるとともに、文字領域については高い品位が得られる圧縮方法である。

#### 【0036】

多値画像である入力画像(501)は、2値化部(502)で2値化され、2値画像(503)が生成される。領域判定部(504)は、2値画像(503)を入力とし、例えば、所定値の画素(例えば、黒画素)の輪郭線追跡等により画素塊を取得し、当該画素塊の大きさや位置に基づいてグループ化していくことにより領域を形成し、当該形成された領域内の画素塊の大きさや並び方などに基づいて文字領域を判別して、文字領域情報を生成する。文字領域情報は、文字領域の位置や大きさを示す情報である。また、領域判定部(504)が文字領域を判定することで、それ以外の部分は、写真やイラストや背景等の自然(階調)画像を示す写真領域として判定する。文字切り出し部(505)は、領域判定部(504)により、文字領域と判定した領域に対して、文字領域内における各文字(単位文字領域)を文字切り矩形として切り出し、文字切り矩形情報を生成する。文字切り矩形情報は、文字きり矩形の位置や大きさを示す情報である。文字領域情報、及び文字切り矩形情報は、文字領域情報(506)として情報管理されるものとする。また、2値画像(503)を入力とし、領域判定部(504)により文字領域と判定された領域について、文字領域毎の2値画像である部分2値画像(507)を生成する。

#### 【0037】

一方で、入力画像(501)は、縮小部(512)により縮小(又は低解像度化)され、縮小多値画像(513)が生成される。代表色抽出部(510)は、部分2値画像(507)を入力とし、文字領域情報(506)及び縮小多値画像(513)を参照しながら、文字切り矩形の代表色を算出し、その結果を文字色情報(511)として情報管理する(尚、この処理の詳細については、特許文献2参照)。文字領域穴埋め部(514)は、縮小多値画像(513)を入力とし、文字領域情報(506)及び部分2値画像(507)を参照しながら、縮小多値画像(513)の各文字領域あるいは文字切り矩形を、その周辺色で塗り潰す処理を行う(尚、この処理の詳細については、特許文献1参照)。

#### 【0038】

以上の処理の後、部分2値画像(507)は各々、MMR圧縮部(508)により圧縮コード1(509)として圧縮される。また、文字領域穴埋め部(514)にて穴埋めされた穴埋め多値画像は、JPEG圧縮部(515)により圧縮コード2(516)として圧縮される。

#### 【0039】

このようにして、各構成要素から得られる文字領域情報(506)、圧縮コード1(509)、文字色情報(511)、圧縮コード2(516)のデータ群を結合した圧縮データ(517)が生成される。この圧縮データ(517)を、更に、PDFなどで可逆圧縮することにより、PDF高压縮データが生成される。

#### 【0040】

図6は、前述したように生成された圧縮データ(517)を伸長する概略構成を示す図である。MMR伸長部(601)は圧縮コード1(509)を入力とし、MMR伸長処理を行い、部分2値画像(602)を生成する。JPEG伸長部(605)は圧縮コード2(516)を入力し、JPEG伸長処理を行い、さらに拡大部(606)で拡大処理を行うことで、多値画像(607)を生成する。合成部(603)は文字領域情報(506)を参照しながら、文字色情報(511)を部分2値画像(602)の黒画素に割り当て、その文字色が割り当てられた部分2値画像を多値画像(607)の上に合成して表示する。この際、部分2値画像(602)の白画素は透明色が割り当てられており、多値画像(607)を透過する。

#### 【0041】

このように、画像伸長装置は、画像圧縮装置により生成された圧縮データを伸長し、画

10

20

30

40

50

像を復元する。

【0042】

図7は、図5及び図6で使用、または生成される画像の概略図を示す。

【0043】

701は、入力画像(501)を示す。702は、2値画像(503)を示す。

【0044】

703は、領域判定部(504)により、文字領域、写真領域に領域判定された結果を示す。ここで、704及び706は文字領域と判定され、705は写真領域として判定されたものとする。

【0045】

707、708は、領域判定部(504)により文字領域と判定された領域の部分2値画像(507)を示す。

【0046】

709は、文字切り出し部(505)により切り出された文字切り矩形の概略図を示す。710は、文字領域704の文字切り矩形であり、711、712は、文字領域706の文字切り矩形である。ここで、711、712に示すように文字領域内の文字切り矩形の中に、文字と写真が混在することがある。例えば、特許文献1のように画素の集まりを位置の近さやサイズの一致に基づいてグループ化した場合、文字サイズに近い写真領域が文字領域内に混在する場合がある。これらの矩形全てを文字として扱うと、712のような本来写真として扱うべき矩形は、2値画像として処理が行われるため、情報の欠落が生じる。仮に文字領域内の文字切り矩形を全て文字として扱った場合に生成される圧縮データ(517)または、PDF高压縮データを713に示す。ここで、714に示すように本来階調や色を有する写真領域が文字領域として扱われて2値化されてしまい、情報欠落が生じることになる。

【0047】

これらの問題点を解決するために、本発明では、図8に示すように領域判定部2(801)を更に設け、文字切り矩形の領域判定を行う。その他の構成要素は図5と同様である。

【0048】

次に、図9のフローチャートを用いて、本発明のポイントである領域判定部2(801)の説明を行う。ここで、図9のフローチャートは、図8の処理の一部であるため、図8を適宜参照する。また、領域判定部2(801)は、図9の907～910の処理を示す。

【0049】

まず、ステップ901にて、入力画像(501)に対して2値化部(502)により2値化を行う。

【0050】

次に、ステップ902にて、2値画像(503)に対して領域判定部(504)により領域判定を行う。ステップ902での領域判定は、例えば、2値画像において輪郭線追跡を行うことによって画素塊を取得し、近くの画素塊同士をグループ化することにより分割されてしまっている文字や文字行が結合されることになる。このグループ化によって形成された領域に含まれる画素塊の大きさや位置関係などに基づいて、当該領域が1又は複数の文字を含む文字領域かどうかの判定が行われる。

【0051】

次に、ステップ903にて、領域数のカウンタであるnを初期化する。次に、ステップ904にて、注目領域が文字領域と判定された領域である場合は、ステップ905へ、非文字領域と判定された領域である場合は、ステップ912へ進む。

【0052】

ステップ905では、文字切り出し部(505)にて文字切り出しを行う。例えば、水平方向にヒストグラムを取って文字行を切り出し、各文字行の垂直方向のヒストグラムを

10

20

30

40

50

取って文字矩形を切り出すことができる。

【0053】

ステップ906にて、文字切り矩形数のカウンタであるmを初期化する。次に、領域判定部2(801)において、まず、ステップ907にて、文字切り矩形内の閉ループ数(NumLoop)を算出する。閉ループ数とは、文字切り矩形内における黒画素、または白画素の輪郭数であり、詳細については、後述する。次に、ステップ908にて、閉ループ数(NumLoop)と予め設定された閾値th(例えば閾値th=9)を比較する(ここで、閾値thは、固定値でもよいが、文字切り矩形の大きさに応じて動的に可変としてもよい)。閾値よりも小さい場合は、ステップ909にてTEXTと判断され、閾値以上の場合は、ステップ910にてIMAGEと判断される。

10

【0054】

ステップ911にて、文字切り矩形数のカウンタmと文字領域内の文字切り矩形数Mとの比較を行い、全ての文字切り矩形についてステップ907~910の処理が終了している( $m == M - 1$ )と判断するとステップ912へ進む。一方、未処理の文字切り矩形があると判断した場合は、ステップ913にてカウンタmを増やしてステップ907に戻る。

【0055】

ステップ912にて、領域数のカウンタnと領域数Nの比較を行い、全ての領域についての判定が終了したと判断すると本処理を終了し、未処理の領域があると判断した場合は、ステップ914にてカウンタnを増やしてステップ904に戻る。

20

【0056】

このように、領域判定部2(801)では、文字領域内に含まれる各文字切り矩形がTEXTであるかIMAGEであるかの判定を、閉ループ数に基づいて行う。ここで、入力画像(501)を701とした場合、領域判定部2(801)で各文字切り矩形の判定を行った結果を図10の1001に示す。また、図8の802は、この結果にもとづいて、領域が文字領域かつ領域判定部2(801)でTEXTと判定された文字切り矩形領域を用いて部分2値画像を生成することを示しており、この生成された部分2値画像の概略図を図10の1002に示す。領域判定部2での判定処理が無い場合は図7の714のように写真領域が2値化されてしまっていたが、領域判定部2での判定処理を行った場合に生成される圧縮データ(517)は、高精度に領域判定できるので、図10の1003のように写真領域が2値化されてしまわずに済む。

30

【0057】

次に、ステップ907の閉ループ数の算出方法について図11を用いて説明する。2値画像からの輪郭線の抽出は、まず2値画像における輪郭線追跡のため追跡開始点を見つけることから始められる。そして、この追跡開始点が発見されると、次にこの追跡開始点から順にその画像の輪郭を追跡していく、追跡の終わった輪郭点には次々に追跡済のマークを付けながら追跡を続行する。こうして、この追跡が一巡した時点での輪郭点列(輪郭線)を求めている。このような手順を繰り返し行なうことにより、その画像中のすべての輪郭線を抽出することができる。文字切り矩形内におけるこれらの輪郭線の総数を本実施形態では、閉ループ数(NumLoop)としている。

40

【0058】

図11は、文字切り矩形の一例であり、まず、ラスタスキャンを行って白画素から黒画素に変化する点を輪郭線追跡のため追跡開始点1101として発見する。次に、破線矢印1102に沿って輪郭を追跡する。1102が追跡開始点1101に戻ると閉ループ数は、1つカウントされる。そして、次の追跡開始点1103が発見され、以下、同様の処理が行われる。図11では、閉ループ数は2とカウントされる。

【0059】

また、図12の1205~1209はそれぞれ、文字切り矩形の例であり、夫々の閉ループ数(NumLoop)を示している。ここで、1205~1208のような文字の場合は、閉ループ数が1桁と少ないが、1209のような写真の場合は、閉ループ数が極端

50

に多いという特徴がある。

【0060】

以上のように、領域判定部2(801)において、文字切り矩形の閉ループ数にもとづいて、文字切り矩形が文字か写真かを判定することで、高精度に領域判定を行うことができるようになる。高精度に領域判定できるようになるので、良好な画質の圧縮データ(517)またはPDF高压縮データを得ることが可能となる。

【0061】

(実施例2)

実施例1では、文字切り矩形の閉ループ数にもとづいて、文字切り矩形が文字か写真かを判定する場合について説明した。しかし、例えば、複雑な漢字は、アルファベット、ひらがなを含む一般的な文字に対して、閉ループ数が比較的多くなる傾向があり、また、逆に、複雑な漢字の閉ループ数程度に少ない閉ループ数をもつ写真も存在する。

【0062】

実施例1では閉ループ数の閾値 $t_h$ を1つ用いて判断することにしていたが、実施例2では、閉ループ数の閾値を複数段階に分け、閉ループ数以外の特徴量を利用できるようにすることにより、文字切り矩形の判定をより良好に行うための領域判定方法について説明する。尚、実施例1と同様の処理部分に関しては、基本的に、同図面、同図番を使用することとし、説明を省略する。

【0063】

以下、閉ループ数の閾値を複数設けた場合の領域判定方法について、図13のフローチャートを用いて説明する。

【0064】

図13は、実施例1で使用した図9の907～910に対応する領域判定部2(801)である。まず、ステップ1301にて、実施例1で前述したように閉ループ数を算出する。次に、ステップ1302にて、閉ループ数(NumLoop)と予め設定された第1の閾値 $t_h1$ を比較する。ここで、閾値 $t_h1$ よりも小さい場合は、ステップ1303進み、閾値 $t_h1$ 以上の場合は、ステップ1307にてIMAGEと判断される。ステップ1303では、閉ループ数(NumLoop)と予め設定された第2の閾値 $t_h2$ と比較する。ここで、閾値 $t_h2$ よりも小さい場合は、ステップ1306にて、TEXTと判断される。閾値 $t_h2$ 以上の場合は、ステップ1304にて色分散値(ColorStd)を算出する。色分散値の算出方法については、後述する。次に、ステップ1305にて、色分散値と予め設定された色分散値の閾値 $t_h3$ を比較する。ここで、閾値 $t_h3$ よりも小さい場合は、ステップ1306にて、TEXTと判断される。閾値 $t_h3$ 以上の場合は、ステップ1307にてIMAGEと判断される。なお、各閾値は実験的に求めた値を用いればよく、例えば、 $t_h1 = 18$ 、 $t_h2 = 6$ というような値などが用いてもよい。

【0065】

ここで、前述したように閉ループ数が比較的多くなる複雑な漢字や閉ループ数が比較的少ない写真を第2の閾値 $t_h2$ より大きくなるように第2の閾値 $t_h2$ を設定し、一般的な単色(例えば、黒)の文字に対して、一般的な多階調表現をもつ写真は、色分散値が大きくなるという特徴を利用することで、ステップ1305にてそれらを良好に判定することが可能となる。

【0066】

次に、色分散値の算出方法について説明する。ここで、色情報は、縮小多値画像(513)の色(例RGB値の各8bit)を参照する必要がある。さらに、好ましくは、RGB値を輝度、色差情報に変換した値(例えは、YCbCr値の各8bit)を参照する。ここでは、例としてYCbCr値のCb値の色分散値を算出する。RGBからYCbCrへの変換方法については、公知であるため、説明を省略する。

【0067】

色分散値は、一般的に知られるvarianceを用い、下記の式で求められる。

【0068】

10

20

30

40

50

**v a r i a n c e ( 分散 ) :**  $( C b ( i ) - m )^2 / n$   
 ( n : データ数 ( 文字切り矩形内の黒画素数 ) 、

**C b ( i ) :** 文字切り矩形内の黒画素と位置的に対応する縮小多値画像の C b 値、  
**m ( 平均 ) :**  $C b ( i ) / n$  )

以上のように、閉ループ数の閾値を段階的に分け、各々の段階に応じて閉ループ数以外の特徴量を併用することで、より良好な領域判定が可能となる。

#### 【 0 0 6 9 】

尚、実施例 2 では、閉ループ数とは異なる別の特徴量として色分散値を利用して説明を行ったが、これに限るものではなく、例えば、文字切り矩形内に占める黒画素の密度情報、または孤立画素情報などを利用してもよい。

10

#### 【 0 0 7 0 】

##### ( 実施例 3 )

実施例 1 では、カラー複写機に搭載可能な、カラー画像の圧縮技術において本領域判定方法を用いる例について説明したが、これに限るものではなく、本発明は、モノクロ複写機におけるモノクロ画像の圧縮技術、あるいは、カラー複写機におけるモノクロ画像の圧縮技術においても利用可能である。モノクロ画像の圧縮技術の場合は、図 5 の入力画像 ( 5 0 1 ) は、例えばグレースケール ( 8 b i t ) であり、2 値画像 ( 5 0 3 ) 及び縮小多値画像 ( 5 1 3 ) を M M R による 2 値圧縮と J P E G による非可逆圧縮との組み合わせで実現する。

#### 【 0 0 7 1 】

20

##### ( 実施例 4 )

実施例 1 ~ 3 では、画像の圧縮技術において本領域判定方法を用いる例を説明した。実施例 4 では、光学的文書認識 ( O C R ) 技術を用いる際に、本領域判定方法を用いる例を説明する。

#### 【 0 0 7 2 】

30

前述したように、 O C R 処理では、文書画像に対して濃度射影 ( ヒストグラム ) を取ることにより文字行を切り出し ( 抽出 ) 、さらに 1 文字単位の文字ブロック切り出し ( 抽出 ) を行う。そして、個々の文字ブロックから特徴データを抽出して、標準パターンとの類似度が計算され、最も類似度の高い文字を認識結果として出力する。即ち、文字ブロック切り出し ( 抽出 ) 処理までは、実施例 1 の図 9 で前述したように、2 値化、領域判定、文字切り出しを行うことを意味する。また、前述したように、文字切り矩形が非文字である場合、非文字に対して文字認識を行ってしまうと、全体の処理速度を低下させてしまう他、意味のないテキストコードを出力してしまう場合もあり好ましくない。

#### 【 0 0 7 3 】

ここで、実施例 4 では、 O C R 処理を行う前に、図 8 の領域判定部 2 で、予め文字切り矩形の領域判定を行うことで、文字か非文字かを判定しておき、文字と判定された場合のみ O C R 処理を行うことでこれらの問題を解決する。この処理を図 14 のフローチャートに示す。図 14 において、 9 0 1 ~ 9 1 4 の処理部分は、実施例 1 で前述した図 9 と同様であるため、説明を省略する。ステップ 9 0 8 にて、閉ループ数 ( N u m L o o p ) と予め設定された閾値 t h を比較し、閾値よりも小さい場合は、ステップ 9 0 9 にて T E X T と判断されるため、ステップ 1 4 0 1 にて O C R 処理を行い、文字認識結果を出力する。また、閾値以上の場合は、ステップ 9 1 0 にて I M A G E と判断されるため、 O C R 処理は行わない。

40

#### 【 0 0 7 4 】

以上のように、 O C R 技術を用いる際、不要な O C R 処理を行わないため、処理速度の向上が図られ、また、意味のないテキストコードを出力してしまうことを抑えることが可能となる。

#### 【 0 0 7 5 】

##### ( 実施例 5 )

実施例 1 では、図 8 のように、領域判定部 ( 5 0 4 ) で複数の文字を含む文字領域を判

50

定し、その領域から文字切りだし部（505）で文字矩形領域を切り出して、領域判定部2（801）において、閉ループの数に基づいて文字か否か判断していたが、この処理の流れに限るものではない。例えば、領域判定部（504）で文字領域の判定を行わずに、文字切りだし部（505）において、画像503から直接文字矩形領域を切り出すことができるようにして、領域判定部2（801）において、該文字矩形領域の閉ループの数に基づいて文字か否か判断するようにしてもよい。

#### 【0076】

##### （実施例6）

本発明は、複数の機器（例えばホストコンピュータ、インターフェース機器、リーダ、プリンタなど）から構成されるシステムに適用しても、一つの機器からなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

10

#### 【0077】

また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記録媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記録媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記録媒体から読み出されたプログラムコード自体が本発明の新規な機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。プログラムコードを供給するための記憶媒体としては、例えば、フロッピー（登録商標）ディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

20

#### 【0078】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼動しているOSなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

#### 【0079】

さらに、記録媒体から読み出されたプログラムコードが、コンピュータに挿入された拡張機能ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードに指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

30

#### 【0080】

以上のように、本発明によれば、領域判定精度を向上させることができることが可能となる。また文字と非文字の好適な領域判定を実行することができるので、良好な画質が得られると共に、圧縮効率を向上させることができとなる。また、OCR技術を用いる際、処理速度の向上と共に、意味のないテキストコードを出力してしまうことを抑え、認識率を向上させることができとなる。

#### 【図面の簡単な説明】

#### 【0081】

40

【図1】本実施例1におけるシステムの概略図

【図2】本実施例1におけるMFPのハードウェア構成

【図3】本実施例1におけるMFPのソフトウェア構成

【図4】本実施例1におけるPCのハードウェア構成

【図5】画像圧縮装置のブロック図1

【図6】画像伸長装置のブロック図

【図7】入力画像のサンプル～出力画像のサンプル

【図8】本発明における画像圧縮装置のブロック図2

【図9】本実施例1における領域判定のフローチャート

【図10】本実施例1における入力画像～出力画像のサンプル

50

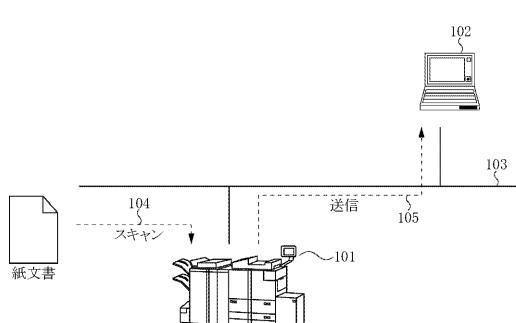
【図1】閉ループ算出の説明図1

【図1】閉ループ算出の説明図2

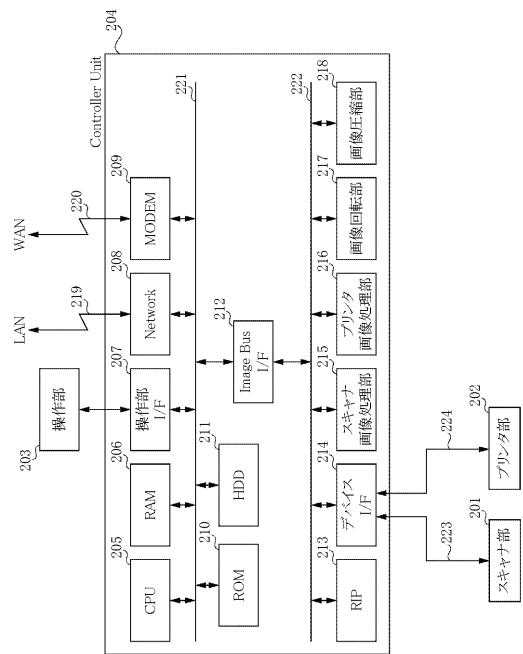
【図1】本実施例2における領域判定のフローチャート

【図1】本実施例4における領域判定のフローチャート

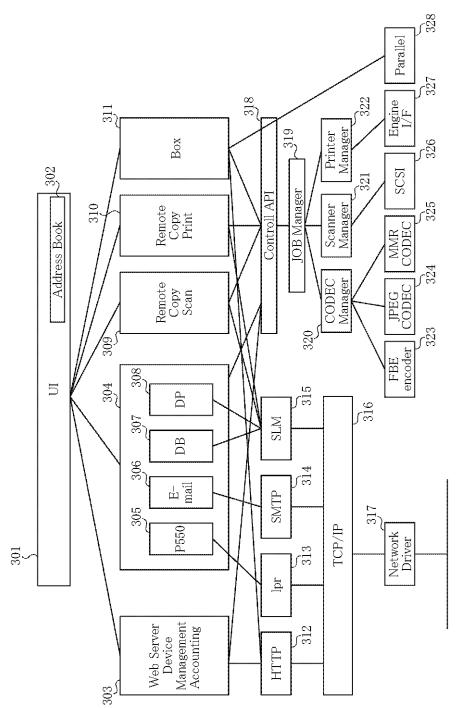
【図1】



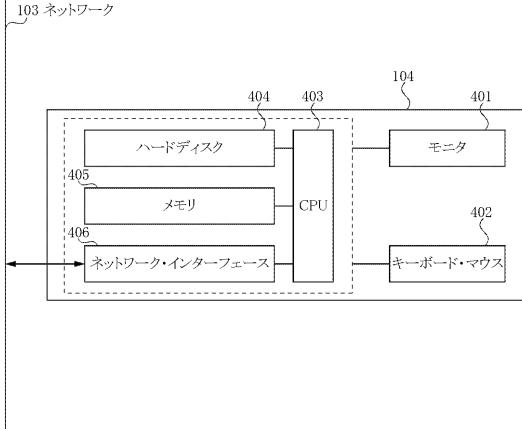
【図2】



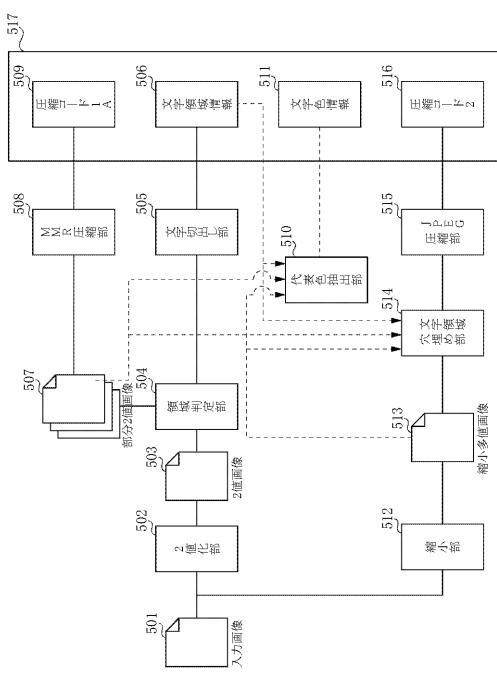
【図3】



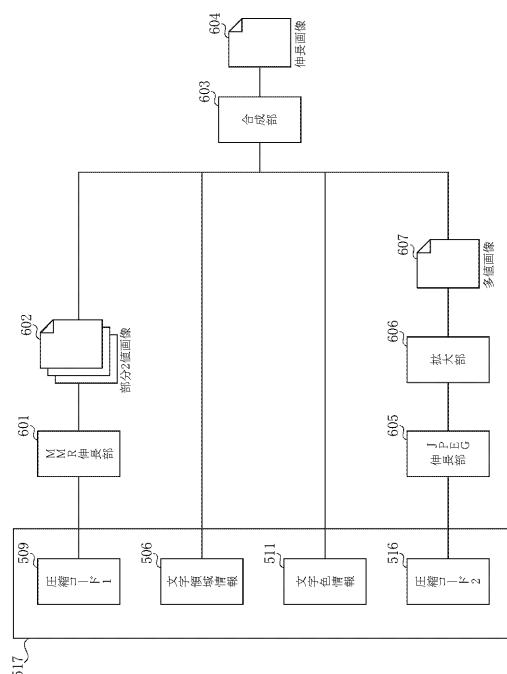
【 図 4 】



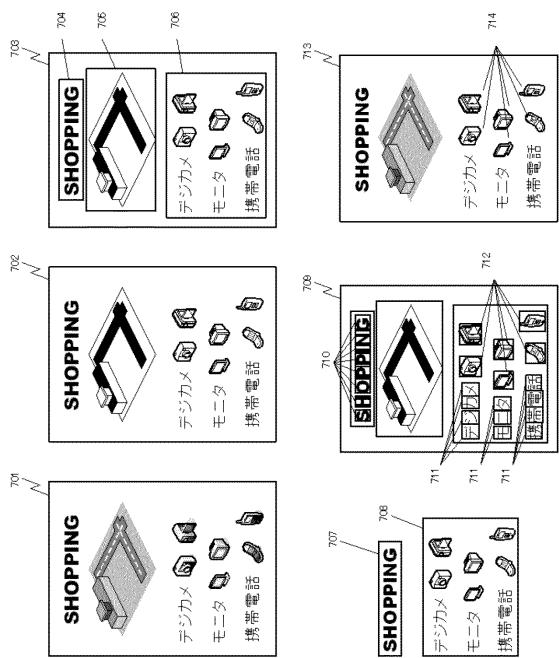
【 図 5 】



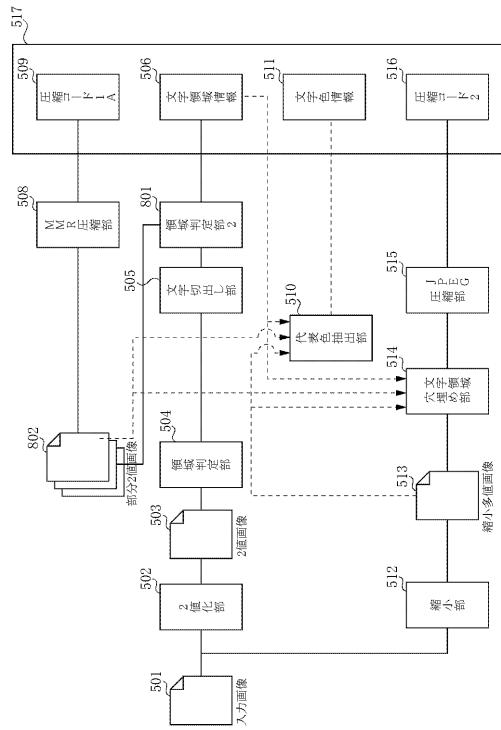
【図6】



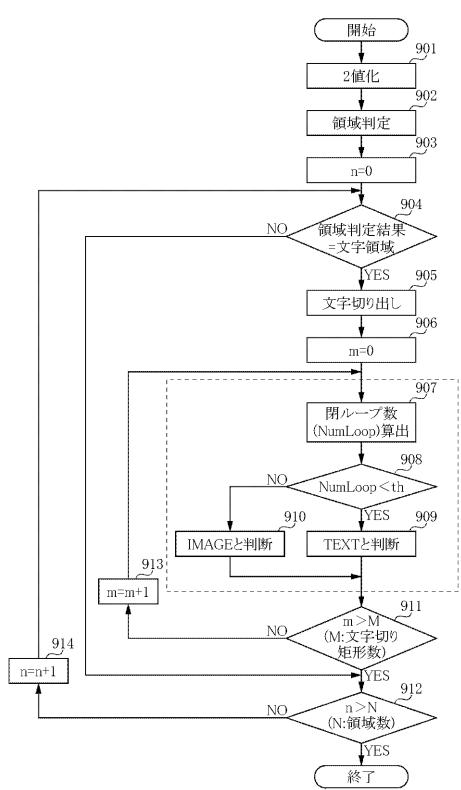
【図7】



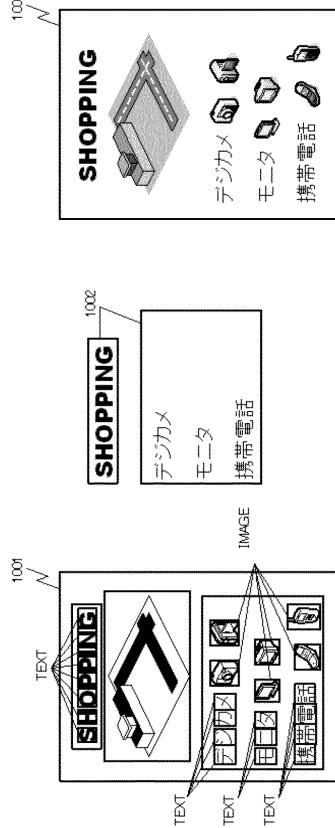
【図8】



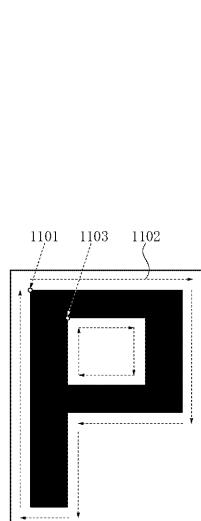
【図9】



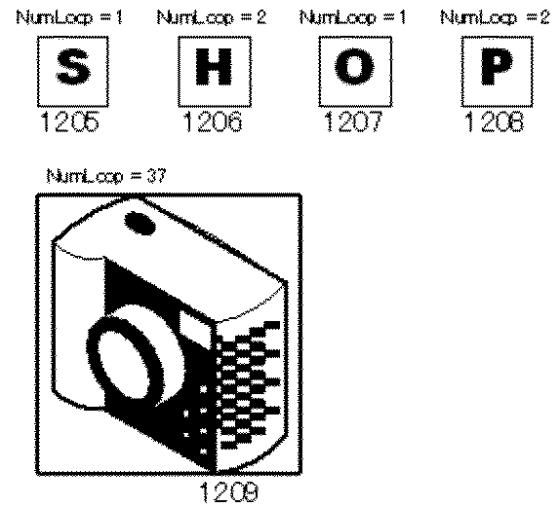
【図10】



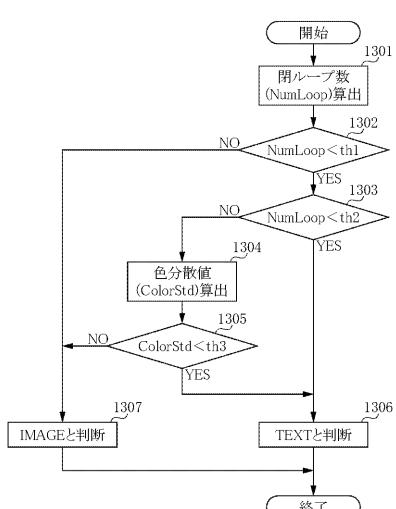
【図11】



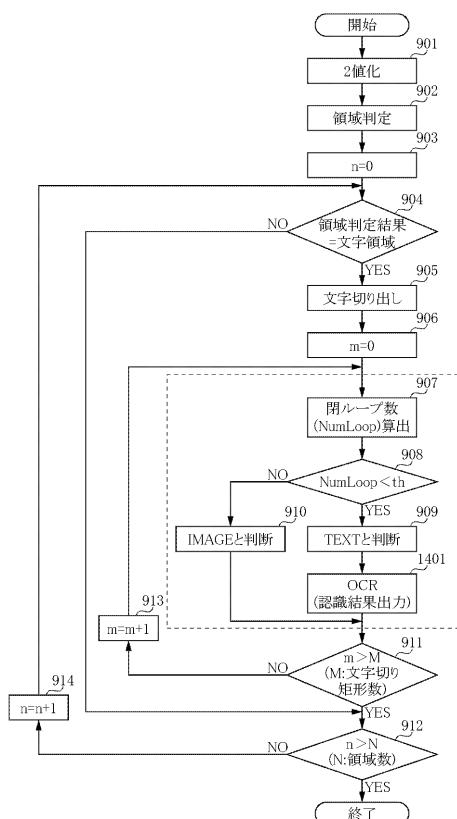
【図12】



【図13】



【図14】



---

フロントページの続き

(56)参考文献 特開2004-128880(JP, A)  
特開平10-260993(JP, A)

(58)調査した分野(Int.Cl., DB名)

H 04 N	1 / 413
G 06 T	1 / 00
G 06 T	7 / 40
H 04 N	1 / 40