

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第5121599号  
(P5121599)

(45) 発行日 平成25年1月16日(2013.1.16)

(24) 登録日 平成24年11月2日(2012.11.2)

(51) Int.Cl.

G 06 T 11/60 (2006.01)  
H 04 N 1/387 (2006.01)

F 1

G 06 T 11/60 100 A  
H 04 N 1/387

請求項の数 10 (全 20 頁)

(21) 出願番号 特願2008-171178 (P2008-171178)  
 (22) 出願日 平成20年6月30日 (2008.6.30)  
 (65) 公開番号 特開2010-9509 (P2010-9509A)  
 (43) 公開日 平成22年1月14日 (2010.1.14)  
 審査請求日 平成23年6月28日 (2011.6.28)

(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  
 (74) 代理人 110001243  
 特許業務法人 谷・阿部特許事務所  
 (74) 代理人 100077481  
 弁理士 谷 義一  
 (74) 代理人 100088915  
 弁理士 阿部 和夫  
 (72) 発明者 山▲崎▼ 妙子  
 東京都大田区下丸子3丁目30番2号 キ  
 ャノン株式会社内  
 (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  
 ャノン株式会社内

最終頁に続く

(54) 【発明の名称】画像処理装置、画像処理方法およびそのプログラムならびに記憶媒体

## (57) 【特許請求の範囲】

## 【請求項 1】

画像から所定フォーマットの電子ドキュメントを生成する画像処理装置であって、  
前記画像を解析し文字領域を抽出する解析手段と、  
 前記解析手段で得られた文字領域について文字認識を行い認識結果の文字コードデータを得る文字認識手段と、  
 前記解析手段で得られた文字領域についてベクトル化を行い、ベクトル描画データを得るベクトル化手段と、

前記電子ドキュメントを所定プログラムで表示したときに前記ベクトル描画データが前記画像と同じレイアウトで前記所定プログラムのレイアウト編集領域に描画されるように、前記ベクトル描画データを前記電子ドキュメント内のレイアウト編集用データとして格納し、かつ、前記電子ドキュメントを前記所定プログラムで表示したときに前記文字コードデータが前記レイアウト編集領域と異なる前記所定プログラムのテキスト編集領域に描画されるように、前記文字コードデータを前記電子ドキュメント内のテキスト編集用データとして格納した電子ドキュメントを生成する電子ドキュメント生成手段と  
 を備えることを特徴とする画像処理装置。

## 【請求項 2】

画像から所定フォーマットの電子ドキュメントを生成する画像処理装置であって、  
前記画像を解析し文字領域を抽出する解析手段と、  
 前記解析手段で得られた文字領域について文字認識を行い認識結果の文字コードデータ

を得る文字認識手段と、

前記解析手段で得られた文字領域について、更に文字レイアウト情報を抽出する文字レイアウト抽出手段と、

前記解析手段で得られた文字領域についてベクトル化を行い、ベクトル描画データを得るベクトル化手段と、

背景画像を作成する背景画像作成手段と、

前記電子ドキュメントを所定プログラムで表示したときに前記文字レイアウト情報と前記文字コードデータとを組み合わせたレイアウト付き文字描画データと前記背景画像と前記ベクトル描画データとが前記所定プログラムのレイアウト編集領域に描画されるように、前記レイアウト付き文字描画データと前記背景画像と前記ベクトル描画データとを前記電子ドキュメント内のレイアウト編集用データとして格納し、かつ、前記電子ドキュメントを前記所定プログラムで表示したときに前記文字コードデータが前記レイアウト編集領域と異なる前記所定プログラムのテキスト編集領域に描画されるように、前記文字コードデータを前記電子ドキュメント内のテキスト編集用データとして格納した、電子ドキュメントを生成する電子ドキュメント生成手段と

を備えることを特徴とする画像処理装置。

**【請求項 3】**

前記電子ドキュメント生成手段は、前記レイアウト編集用データとして格納される前記レイアウト付き文字描画データのうち、前記文字認識手段による文字認識率が所定の閾値以上である前記レイアウト付き文字描画データに関しては、前記電子ドキュメントを前記所定プログラムで表示したときに前記所定プログラムの前記レイアウト編集領域および前記所定プログラムの要約情報領域の両方に描画されるように、前記レイアウト付き文字描画データを前記電子ドキュメントに格納し、前記文字認識率が所定の閾値未満である前記レイアウト付き文字描画データに関しては、前記電子ドキュメントを前記所定プログラムで表示したときに前記所定プログラムの前記レイアウト編集領域に描画され且つ前記所定プログラムの要約情報領域に描画されないように、前記レイアウト付き文字描画データを前記電子ドキュメントに格納することを特徴とする請求項 2 に記載の画像処理装置。

**【請求項 4】**

前記電子ドキュメント生成手段は、前記レイアウト編集用データとして格納される前記レイアウト付き文字描画データのうち、前記文字認識手段による文字認識率が所定の閾値未満であるレイアウト付き文字描画データに関しては、文字コードデータを削除した文字レイアウト情報のみのデータとしたうえで、前記電子ドキュメントを前記所定プログラムで表示したときに前記所定プログラムの前記レイアウト編集領域および前記所定プログラムの要約情報領域の両方に描画されるように前記電子ドキュメントに格納することを特徴とする請求項 3 記載の画像処理装置。

**【請求項 5】**

前記電子ドキュメント生成手段は、前記レイアウト付き文字描画データと前記背景画像と前記ベクトル描画データとを前記所定プログラムのレイアウト編集領域に描画する際、下から順に、前記レイアウト付き文字描画データ、前記背景画像、前記ベクトル描画データが描画されるように前記電子ドキュメントを生成することを特徴とする請求項 2 乃至 4 のいずれか 1 項に記載の画像処理装置。

**【請求項 6】**

前記電子ドキュメントの前記所定フォーマットは、XML フォーマットであ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】**

画像から所定フォーマットの電子ドキュメントを生成するために、画像処理装置によって実行される画像処理方法であって、

解析手段が、前記画像を解析し文字領域を抽出する解析ステップと、

文字認識手段が、前記解析ステップで得られた文字領域について文字認識を行い認識結果の文字コードデータを得る文字認識ステップと、

10

20

30

40

50

ベクトル化手段が、前記解析ステップで得られた文字領域についてベクトル化を行い、ベクトル描画データを得るベクトル化ステップと、

電子ドキュメント生成手段が、前記電子ドキュメントを所定プログラムで表示したときに前記ベクトル描画データが前記画像と同じレイアウトで前記所定プログラムのレイアウト編集領域に描画されるように、前記ベクトル描画データを前記電子ドキュメント内のレイアウト編集用データとして格納し、かつ、前記電子ドキュメントを前記所定プログラムで表示したときに前記文字コードデータが前記レイアウト編集領域と異なる前記所定プログラムのテキスト編集領域に描画されるように、前記文字コードデータを前記電子ドキュメント内のテキスト編集用データとして格納した電子ドキュメントを生成する電子ドキュメント生成ステップと

10

を含むことを特徴とする画像処理方法。

**【請求項 8】**

画像から所定フォーマットの電子ドキュメントを生成するために、画像処理装置によって実行される画像処理方法であって、

解析手段が、前記画像を解析し文字領域を抽出する解析ステップと、

文字認識手段が、前記解析ステップで得られた文字領域について文字認識を行い認識結果の文字コードデータを得る文字認識ステップと、

文字レイアウト抽出手段が、前記解析ステップで得られた文字領域について、更に文字レイアウト情報を抽出する文字レイアウト抽出ステップと、

ベクトル化手段が、前記解析ステップで得られた文字領域についてベクトル化を行い、ベクトル描画データを得るベクトル化ステップと、

20

背景画像作成手段が、背景画像を作成する背景画像作成ステップと、

電子ドキュメント生成手段が、前記電子ドキュメントを所定プログラムで表示したときに前記文字レイアウト情報と前記文字コードデータとを組み合わせたレイアウト付き文字描画データと前記背景画像と前記ベクトル描画データとが前記所定プログラムのレイアウト編集領域に描画されるように、前記レイアウト付き文字描画データと前記背景画像と前記ベクトル描画データとを前記電子ドキュメント内のレイアウト編集用データとして格納し、かつ、前記電子ドキュメントを前記所定プログラムで表示したときに前記文字コードデータが前記レイアウト編集領域と異なる前記所定プログラムのテキスト編集領域に描画されるように、前記文字コードデータを前記電子ドキュメント内のテキスト編集用データとして格納した、電子ドキュメントを生成する電子ドキュメント生成ステップと

30

を備えることを特徴とする画像処理方法。

**【請求項 9】**

請求項 7 または 8 に記載の方法をコンピュータに実行させるためのプログラム。

**【請求項 10】**

請求項 9 に記載のプログラムを記憶させ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】**

**【0001】**

本発明は、文書画像に対して画像処理を実施する画像処理装置に関する。

40

**【背景技術】**

**【0002】**

近年、インターネットに代表されるようなネットワークの広がりにより、文書が電子的に配布される機会も増えているが、電子文書が紙に印刷された状態で配布されことが多い。そこで、紙文書のみが手元に存在する場合でも、その内容を紙から再利用可能な電子データとして得られるようにするための技術が考えられている。

**【0003】**

たとえば、端末から、紙文書をスキャンして得た文書画像をサーバに送信し、サーバ上でこれを文字認識し、再利用可能なフォーマットに変換して端末に戻す技術がある（特許文献 1 参照）。

50

**【0004】**

また、文書画像を種類に対応した領域に分けて、個別に出力できるようにする技術もある（特許文献1参照）。

**【0005】**

- 【特許文献1】特開平11-167532号公報
- 【特許文献2】特開平09-091450号公報
- 【特許文献3】特許第3026592号公報
- 【特許文献4】特開2005-346137号公報
- 【特許文献5】特開2007-272601号公報

**【発明の開示】**

10

**【発明が解決しようとする課題】****【0006】**

ユーザが再利用したいデータの形式は場合によって異なるが、それぞれがユーザにとって抽出しやすい形態とされることが望まれる。また、文字認識技術にも限界があり、ある程度の誤判定もでてくる。再利用するコンテンツの認識精度が低ければ、逆にユーザにとっては使いにくいデータになる。特許文献1では再利用できるデータ形式は文字情報のみであり、認識精度を加味せずに変換している。しかし、紙文書においては、文書中のコンテンツそのもののほかに、紙面上のレイアウトや位置関係が重要な意味を持つことが多い。

**【0007】**

20

また、特許文献2では、文書を分割しそれぞれを出力しているため、コンテンツ同志の関係が欠落してしまう。

**【0008】**

また、画像内に含まれている文字画像について、ユーザが、ベクトルデータとして再利用したり、文字認識結果の文字コードとして再利用したりするのが、容易に行えるようになりたいという課題もある。

**【課題を解決するための手段】****【0009】**

前記課題を解決するために、本発明の画像処理装置は、画像から所定フォーマットの電子ドキュメントを生成する画像処理装置であって、前記画像を解析し文字領域を抽出する解析手段と、前記解析手段で得られた文字領域について文字認識を行い、認識結果の文字コードデータを得る文字認識手段と、前記解析手段で得られた文字領域についてベクトル化を行い、ベクトル描画データを得るベクトル化手段と、前記電子ドキュメントを所定プログラムで表示したときに前記ベクトル描画データが前記画像と同じレイアウトで前記所定プログラムのレイアウト編集領域に描画されるように、前記ベクトル描画データを前記電子ドキュメント内のレイアウト編集用データとして格納し、かつ、前記電子ドキュメントを前記所定プログラムで表示したときに前記文字コードデータが前記レイアウト編集領域と異なる前記所定プログラムのテキスト編集領域に描画されるように、前記文字コードデータを前記電子ドキュメント内のテキスト編集用データとして格納した電子ドキュメントを生成する電子ドキュメント生成手段とを備えることを特徴とする。

30

**【発明の効果】****【0010】**

文字領域の認識結果である文字コードデータと、ベクトル描画データの両方が提供されるので、ユーザが所望するデータを容易に利用できるようになる。また、再編集可能な状態での認識結果の提供が可能となる。更に、各ベクトル描画データが、入力画像と同じレイアウトで配置されるので、ユーザにとって、各文字領域のベクトル描画データ間の位置関係がわかり易く表示されることになり、閲覧や編集を行いややすくなる。

**【発明を実施するための最良の形態】****【0011】**

&lt;実施形態1&gt;

40

50

以下、図面を参照して本発明の好適な実施形態を詳細に説明する。

【0012】

【システム構成】

はじめに、本発明の第1の実施形態である画像処理装置を含むシステムの全体構成について、図1を用いて説明する。

【0013】

図1は、本実施形態1におけるシステム構成の一例を示す図である。

【0014】

同図において、100は、画像処理装置であり、以下の構成からなる。101は、文書を読み取って画像データに変換するスキャナである。102は、スキャナ101から得られる画像データに、後述の画像処理を施すための電子ドキュメント生成プログラムやその他のプログラムを実行するCPUである。103は、CPU102がこれらのプログラムを実行する際のワークメモリやデータの一時保存などに利用されるメモリであり、104はこれらのプログラムやデータを格納するハードディスクである。また、105は、外部装置とデータの入出力を行うためのネットワークIFである。なお、画像処理装置100は、スキャナ部101に加えて、ファックスやコピー、プリンタなどの機能を備える複合機であってもよい。

10

【0015】

パーソナルコンピュータ(PC)120は、画像処理装置100とLAN110などのネットワークで接続され、画像処理装置100から送信されるデータを受信する。PC120では、受信した電子ドキュメントを画面に表示し編集することができる表示・編集プログラム121を実行したり、またその電子ドキュメントの一部を別の再利用プログラム122によって利用することが可能である。

20

【0016】

【画像処理装置100の概要】

続いて、本実施形態1における画像処理装置100の概要を、図2を用いて説明する。

【0017】

図2は、本実施形態1における画像処理装置100の電子ドキュメント生成に係る構成と入出力されるデータを説明する図である。

【0018】

30

図2に示す各処理は、CPU102が電子ドキュメント生成プログラムを実行することによって実現されるものとするが、その一部又は全部を電気回路で構成するようにしても構わない。

【0019】

同図において、200はスキャナ101から入力された入力画像である。なお、本実施形態において入力画像はスキャナから入力された画像としたが、これに限るものではなく、例えば、ファックスなどを介して受信した画像を入力画像として入力するようにしてもよい。

【0020】

210は、この入力画像から電子ドキュメントを生成するための電子ドキュメント生成部(電子ドキュメント生成プログラム)である。

40

【0021】

220は、電子ドキュメント生成部210により生成された電子ドキュメントである。

【0022】

なお、ブロック211～215は、電子ドキュメント生成部210で実行される各処理を模式的に示したものであり、以下のものから構成される。

【0023】

211は、入力された文書画像を解析し文字領域を抽出する解析部である。

【0024】

212は、解析部211によって抽出された文字領域について文字認識を行い、認識さ

50

れた文字を文字コードデータに変換する文字認識部である。

【0025】

213は、解析部211によって抽出された文字領域について、それに含まれる各文字の輪郭情報を抽出し、抽出した輪郭情報をベクトル描画データに変換するベクトル化部である。

【0026】

214は、文字認識部212が出力する文字コードデータと、ベクトル化部213が出力するベクトル描画データのそれぞれが格納される、220の電子ドキュメント内の格納先を決定する格納先決定部である。

【0027】

215は、格納先決定部214で決定した格納先に応じて、文字コードデータとベクトル描画データとを電子ドキュメント220のフォーマットへ変換し、電子ドキュメント220を生成するフォーマット変換部である。

【0028】

電子ドキュメント生成部210によって生成された電子ドキュメント220は、以下の221～224の各データによって構成され、PC120が備える表示・編集プログラム121によって表示・編集が可能である。

【0029】

221は、表示・編集プログラム121においてレイアウト編集用に用いられるデータであり、ベクトル化部213で生成されたベクトル描画データ223が格納される。また、222は表示・編集プログラム121においてテキスト編集用に用いられるデータであり、文字認識部212による認識結果の文字コードデータ224が格納される。

【0030】

[電子ドキュメント生成部による処理例]

以下では、入力画像200として図3の入力画像300を電子ドキュメント生成部210が処理する場合を説明する。

【0031】

解析部211で公知の画像解析処理を用いて入力画像200中の文字領域を抽出する。具体的には、画像の中から各文字を構成する画素の集合を抽出し、さらにそれらのうち文字の大きさを有する画素集合で、かつ、ほぼ同じ大きさの画素集合が少なくとも縦または横に並んでいる領域を文字領域として抽出する。文字を構成する画素の集合を抽出する手法としては、文字を構成する画素の集合として、入力多値画像から近似色画素塊を抽出する手法や、多値画像を二値化して得た黒画素塊に基づいて抽出する手法がある。あるいは多値画像から微分エッジ情報を生成してエッジ内の画素として抽出する手法などもあるが、そのいずれを用いてもよい。

【0032】

図3に示す300の例では、点線で囲まれる領域301と302がそれぞれ文字領域として抽出されたものとする。文字認識部212では公知の文字認識技術を用いて各文字領域内の文字認識を行い、文字コードデータを生成する。

【0033】

ここで文字認識処理の一例について簡単に説明する。なお、ここで説明するのはあくまで一例であって別の手法を用いてもよい。

【0034】

文字認識処理ではまず文字領域の縦書き横書きすなわち行方向を判断する。これは画像を二値化し、縦と横の射影をとって、この射影の分散の低いほうを行方向と判定する方法があり、これを用いることができる。

【0035】

次に画像を個々の文字画像へと分割する。これは二値画像の行方向への射影を用いて切断すべき行間を決定して行画像に分割し、さらに行画像を行と垂直方向への射影を用いて切断すべき文字間を決定して文字画像へと分割すればよい。

10

20

30

40

50

**【0036】**

そしてそれら文字画像ひとつひとつに対して所定の特徴量を取り、あらかじめ全字種分の特徴量を保存した辞書から一番類似するものを探し、辞書が示す対応する文字コードを各文字の認識結果とする。

**【0037】**

図3の例では、文字領域301の認識結果として“テストテスト”に対応する文字コードデータと、文字領域302の認識結果として“これはテストです。本発明の為のみに使用されるものです。”に対応する文字コードデータが生成されるものとする。

**【0038】**

ベクトル化部213では公知のベクトル化技術を用いて各文字領域の画像データから文字のベクトル描画データを生成する。ベクトル化技術の例としては、特許文献3や特許文献4に開示された方法がある。

10

**【0039】**

たとえば特許文献3では、画像をラスタ走査しながら注目画素とその近傍画素の状態に基づいて、水平方向及び垂直方向の画素間ベクトルを検出する。次に、これら画素間ベクトル同士の接続状態をもとに、画像データの輪郭を抽出することで、アウトラインベクトルと呼ばれる連結画素データの周囲を画素間ベクトルの集合で記述する情報を生成する技術を開示している。また、特許文献4では、アウトラインベクトルを、直線や、2次や3次のベジェ曲線で近似することで、大きく変倍しても高画質なベクトル記述データを生成する技術を開示している。

20

**【0040】**

214の格納先決定部は、例えば図4に示すような、データに対し格納先を定める決定規則に基づいて、文字コードデータ、およびベクトル描画データの格納先を決定する。すなわち文字コードデータは電子ドキュメント220における224のテキスト編集用データとして、ベクトル描画データは電子ドキュメント220における221のレイアウト編集用データとしてそれぞれ格納先が決定される。

**【0041】**

フォーマット変換部215では、文字コードデータおよびベクトル描画データを、格納先決定部214で決定された格納先に従ったフォーマットに変換して電子ドキュメント220を生成する。

30

**【0042】****[電子ドキュメントの具体例]**

ここで、本実施形態1における電子ドキュメントの具体例を、図5を用いて説明する。

**【0043】**

図5の電子ドキュメント500は、本実施形態の説明のために、XMLフォーマットに従って作られた電子ドキュメント220の具体例である。

**【0044】**

同図の501は、電子ドキュメントのレイアウト編集用データを格納する部分で要素名<Layout>と</Layout>で挟まれた箇所に該当する。501には、ベクトル描画データ503が格納されている。なお、ベクトル描画データの各オブジェクト(各文字)は、入力画像と同じレイアウトで描画されるように位置情報も指定されているものとする。

40

**【0045】**

一方、502は、テキスト編集用データを格納する部分で、要素名<Text>と</Text>で挟まれた箇所に該当する。502には文字コードデータ504が格納されている。

**【0046】**

なお、要素名およびデータ構造は表示・編集プログラム121の仕様に準じている。表示・編集プログラム121の種類によって図5に示したもの以外の要素名やデータ構造で出力されることもありうる。

50

## 【0047】

なお、表示・編集プログラム121がマイクロソフト社のパワーポイント（商標）である場合、ベクトル描画データ503はスライド領域（スライドペイン）に表示されるようになり、文字認識結果は、ノート領域（ノートペイン）に表示されるようにする。この場合、電子ドキュメント500のフォーマットは「pptx」であるとする。

## 【0048】

## [表示・編集プログラムによる表示例]

次に、図3の画像300から生成された電子ドキュメント220の例である図5の電子ドキュメント500を、図1のパーソナルコンピュータ120で実行される表示・編集プログラム121で処理する際の表示例を、図6を用いて説明する。

10

## 【0049】

図6は表示・編集プログラム121に電子ドキュメント500を入力した場合の表示例である。

## 【0050】

601は表示・編集プログラム121による表示ウィンドウの全体を示すものであり、その中にレイアウト編集ウィンドウ602、テキスト編集ウィンドウ603、要約情報表示ウィンドウ604をそれぞれ含む。表示・編集プログラム121がパワーポイントの場合、レイアウト編集ウィンドウ602はスライドペイン、テキスト編集ウィンドウ603はノートペイン、要約情報表示ウィンドウ604はアウトラインペインである。

## 【0051】

20

レイアウト編集ウィンドウ602には、電子ドキュメント220中のレイアウト編集用データ221の内容に従った表示がなされる。本実施形態で例示する図5の電子ドキュメント500に対しては、ベクトル描画データ503によって図3の301および302と外観が一致する画像情報が描画されている。

## 【0052】

このレイアウト編集ウィンドウ602内で、ユーザはベクトル描画データによって描画されたオブジェクトの拡大縮小、色情報の変更などといった編集作業が可能である。また編集後の電子ドキュメントを保存したり、602に表示されたとおりの外観で紙面を印刷することも可能である。

## 【0053】

30

テキスト編集ウィンドウ603には、電子ドキュメント220中のテキスト編集用データの内容がテキストデータとして表示される。本実施形態で例示する図5の電子ドキュメント500に対しては、文字コードデータ504の内容である文字列、すなわち図3の301の文字認識結果である文字列“テストテストテスト”、および302の文字認識結果である文字列“これはテストです。本発明の為のみに使用されるものです”が連続して表示されている。

## 【0054】

ユーザはこの603の内容を602に表示されているドキュメント外観の註釈・備考のような補助的な情報を格納するために用いることができ、さらに同ウィンドウの内容をテキストデータとして編集することが可能である。

40

## 【0055】

要約情報ウィンドウ604は、電子ドキュメント220中のレイアウト編集用データ221の内容のうち、要約情報として使用されるデータの文字情報が表示される。ただし本実施形態で例示する図5の電子ドキュメント500には要約情報に使用されるデータが存在しないので、図6のように空の内容が表示されている。

## 【0056】

以上説明したように、本実施形態の画像処理装置は、画像から、文字領域をベクトル化した情報と文字認識した情報を、それぞれの格納先に保存した電子データを生成する。このように生成した電子データは、表示・編集プログラム121の表示画面において、入力された画像から生成されたデータとして文字コードデータとベクトル描画データの両方

50

が提示されているため、ユーザはどちらのデータも直ちに利用することができる。すなわち本実施形態では、文字領域の認識結果である文字コードデータとベクトル描画データの両方を一覧し、再編集可能な状態での提供が可能である。更に、各ベクトル描画データが、入力画像と同じレイアウトで配置されるので、ユーザにとって、各文字領域のベクトル描画データ間の位置関係がわかり易く表示されることになり、閲覧や編集を行いやすくなる。

**【0057】**

＜実施形態2＞

次に、本発明の第2の実施形態における画像処理装置について、図7を用いて説明する。

10

**【0058】**

図7は、図1と同様のシステム構成で実施される、本発明の第2の実施形態における画像処理装置の電子ドキュメント生成に係る構成と入出力されるデータを説明する図である。

**【0059】**

図2と同様に、各処理はCPU102が電子ドキュメント生成プログラムを実行することによって実現されるものとするが、その一部又は全部を電気回路で構成するようにしても構わない。

**【0060】**

同図に示す700は、スキャナ101などから入力された入力画像である。

20

**【0061】**

710は、この入力画像700から電子ドキュメントを生成するための電子ドキュメント生成部である。

**【0062】**

720は、電子ドキュメント生成部710により生成された電子ドキュメントである。同図に示すブロック711～717は、電子ドキュメント生成部710で実行される各処理を模式的に示したものである。

**【0063】**

711は、入力された文書画像を解析し文字領域を抽出する解析部である。

**【0064】**

712は、解析部711から抽出された文字領域について文字認識を行い、文字コードデータに変換する文字認識部である。

30

**【0065】**

713は、解析部711から抽出された文字領域について、文字レイアウトの情報を抽出する文字レイアウト抽出部である。文字レイアウト情報にはたとえば、文字領域の座標、文字のフォント、大きさ、色などがあるが、ここでは少なくとも文字領域の座標と文字の大きさが抽出されるものとする。

**【0066】**

714は、解析部711から抽出された文字領域について各文字の輪郭情報を抽出し、ベクトル描画データに変換するベクトル化部である。

40

**【0067】**

715は、入力画像中の文字領域以外の部分、すなわち背景部分を描画するためのデータを作成する背景作成部である。

**【0068】**

716は、電子ドキュメント720内におけるデータの格納先を決定する格納先決定部である。ここで対象となるデータは、文字認識部712が出力する文字コードデータと、同文字コードデータと文字レイアウト情報とを組みあわせたレイアウト付き文字描画データと、ベクトル化部714が出力するベクトル描画データである。

**【0069】**

717は、文字コードデータ、レイアウト付き文字描画データ、ベクトル描画データ、

50

背景描画データをそれぞれ電子ドキュメント720のフォーマットへ変換するフォーマット変換部である。

【0070】

以上の電子ドキュメント生成部710によって生成された電子ドキュメント720は、以下の721～727の各データによって構成され、PC120にある表示・編集プログラム121によって表示・編集が可能である。

【0071】

721は表示・編集プログラム121においてレイアウト編集用に用いられるデータであり、723と724の2種類のレイアウト付き文字描画データ、725の背景描画データ、および726のベクトル描画データによって構成される。

10

【0072】

一方、722は、表示・編集プログラム121においてテキスト編集用に用いられるデータであり、727の文字コードデータによって構成される。

【0073】

[電子ドキュメント生成部による処理例]

以下、入力画像700として、図8の入力画像801を電子ドキュメント生成部710が処理する場合をさらに説明する。

【0074】

まず解析部711が入力画像700中の文字領域を抽出する。本処理は図2中の211の説明と同じであるため説明は省略する。

20

【0075】

800の入力画像に対しては、点線で囲まれる領域801、802および803がそれぞれ文字領域として抽出されたものとする。

【0076】

認識部712は各文字領域内の文字認識を行い、文字コードデータを生成する。本処理は図2中の212の処理と同じであるため説明は省略する。

【0077】

図8の入力例からは文字領域801、802、803の認識結果としてそれぞれ“ドキュメント”、“これはサンプル電子ド+メント”、“写真写真”という文字コードデータが得られたものとする。なお、この例では、文字領域802の認識結果には、誤認識した文字認識結果が含まれている。

30

【0078】

次いで、文字レイアウト抽出部713が各文字領域の文字レイアウト情報を取得する。ここで抽出されるのは文字領域の座標情報と、文字の大きさの情報であるとする。ここでの文字レイアウト情報抽出は公知の抽出処理を用いるものとする。たとえば文字領域の座標情報は解析部711で当該文字領域を抽出する際に得られ、文字の大きさは文字認識部712で各文字を認識する際に得られる文字の大きさを平均した値として得ることができるが、他の方法を用いてもよい。

【0079】

次に、ベクトル化部714が各文字領域の画像データから文字のベクトル描画データを生成する。本処理は図2中の213の処理と同じであるため説明は省略する。

40

【0080】

次いで、背景作成部715が入力画像から文字領域内の文字部にあたる画素を周囲の背景と同様の色に変更されることにより、背景描画データを作成する。すなわち、入力画像の文字の画素を周囲の色で埋める処理を行うことにより、入力画像から文字部分を消した背景画像を作成する。背景作成部715の処理は公知の処理を用いればよい。たとえば特許文献5には入力画像中の文字をベクトル化し、その部分の画素を背景色で補完する処理が記載されている。

【0081】

次いで、格納先決定部716が、例えば図9に示すような、データとその格納先を定め

50

る決定規則に基づいて、文字コードデータ、レイアウト付き文字描画データ、背景描画データおよびベクトル描画データの格納先を決定する。

【0082】

図9の決定規則に従えば、レイアウト付き文字描画データのうち、文字コードの文字認識率が高いものは723のレイアウト付き文字描画データ1(要約表示用)として、電子ドキュメント720におけるレイアウト編集用データ721内へと格納先が決定される。一方、文字認識率の低いものは724レイアウト付き文字描画データ2(非要約表示用)として、電子ドキュメント720におけるレイアウト編集用データ721内へと、文字認識率の高いものの格納先とは別個の格納先が決定される。背景描画データは725の背景描画データとして、またベクトル描画データは726のベクトル描画データとして、721のレイアウト編集用データ内へと格納先が決定される。一方、文字コードデータは727の文字コードデータとして、722のテキスト編集用データ内へと格納先が決定される。

【0083】

図9の決定規則に従う決定においては、文字認識部712が各文字の認識処理の際に辞書との比較に用いた特徴量間の距離を、領域内の全文字について平均化した値が予め定めた閾値より小さいものを文字認識率の高いデータとする。すなわち、文字領域内に含まれる各文字の類似度が高ければ、文字認識率(認識信頼度)の高いデータとする。一方、認識結果の距離を平均化した値があらかじめ定めた閾値より大きいものは文字認識率の低いデータとする。すなわち、文字領域内に含まれる文字の類似度が低ければ、文字認識率(認識信頼度)の低いデータとする。これはあくまで一例であって、別の方で各文字の認識信頼度を定めて閾値と比較してもよい。あるいは、認識後の文字列を単語辞書と照合して合致する単語が含まれるものと文字認識率の高いデータとし、含まれないものを文字認識率の低いデータとしてもよい。

【0084】

717では、各データを、716で決定された格納先に従ったフォーマットに変換して、電子ドキュメント720を構成する。

【0085】

〔電子ドキュメントの具体例〕

ここで、本実施形態2における電子ドキュメントの具体例を、図10を用いて説明する。

【0086】

図10に示す電子ドキュメント1000は、本実施形態の説明のために作られた仮想的なXMLフォーマットに従って作られた電子ドキュメント720の具体例である。

【0087】

1001は電子ドキュメントのレイアウト編集用データを格納する部分で要素名<Layout>と</Layout>で挟まれた箇所に該当する。1001には、1003のレイアウト付き文字描画データ1、1004のレイアウト付き文字描画データ2、1005の背景描画データ、1006のベクトル描画データが格納されている。

【0088】

1002は、テキスト編集用データを格納する部分で要素名<text>と</text>で挟まれた箇所に該当する。1002には1007の文字コードデータが格納されている。

【0089】

なお、実施形態1と同じく図10の要素名およびデータ構造は、表示・編集プログラム121の仕様に準じている。したがって、表示・編集プログラム121の種類によっては、電子ドキュメント1000は、図10に示したもの以外の要素名やデータ構造で出力されることもありうる。

【0090】

〔表示・編集プログラムによる表示例〕

10

20

30

40

50

次に、図8の画像800から生成された電子ドキュメント720の具体例である図10の電子ドキュメント1000を、図1のパーソナルコンピュータ120で実行される表示・編集プログラム121で処理する際の表示例を、図11を用いて説明する。

#### 【0091】

図11は表示・編集プログラム121に電子ドキュメント1000を入力した場合の表示例である。

#### 【0092】

1101は表示・編集プログラム121の表示ウィンドウの全体であり、その中に1102のレイアウト編集ウィンドウと、1103のテキスト編集ウィンドウと、1104の要約情報表示ウィンドウをそれぞれ含む。

10

#### 【0093】

レイアウト編集ウィンドウ1102には、電子ドキュメント720中のレイアウト編集用データ721の内容に従った表示がなされる。

#### 【0094】

図10の電子ドキュメント1000に対しては、1003のレイアウト付き文字描画データおよび1004のレイアウト付き文字描画データ、1005の背景描画データ、1006のベクトル描画データがこの順番で下から順に描画された結果が表示されている。

#### 【0095】

ただし1003および1004のレイアウト付き文字描画データの上に1005の背景描画データが描画される。つまりレイアウト付き文字描画データ1003とレイアウト付き文字描画データ1004は、背景描画データ1005に上書きされた状態で表示される。このため、実際にユーザが見えるのは背景描画データ1005に基づく背景の上に、1006のベクトル描画データが描画された状態となる。その結果、各文字から抽出された輪郭情報から生成されたベクトル描画データによって、図8の800と外観が一致する画像情報が描画されている。

20

#### 【0096】

前述の実施形態1と同様に、このレイアウト編集ウィンドウ1102内で、ユーザはベクトル描画データによって描画されたオブジェクトの拡大縮小、色情報の変更などといった編集作業が可能である。また編集後の電子ドキュメントを保存したり、レイアウト編集ウィンドウ1102に表示されたとおりの電子ドキュメントの外観で紙面を印刷することも可能である。

30

#### 【0097】

テキスト編集ウィンドウ1103には、電子ドキュメント720中のテキスト編集用データの内容がテキストデータとして表示される。

#### 【0098】

本実施形態2で例示する図10の電子ドキュメント1000に対して、図11のテキスト編集ウィンドウ1103には、“ドキュメントこれはサンプル電子ド+ ユメント写真写真”が表示されている。このテキスト編集ウィンドウ1103は、レイアウト編集ウィンドウ1102に表示されているドキュメント外観のノート（註釈・備考）のような補助的な情報として用いられる。さらに同ウィンドウの内容をテキストデータとして編集することも可能である。

40

#### 【0099】

1104の要約情報ウィンドウは、電子ドキュメント720中のレイアウト編集用データ721の内容のうち、723のレイアウト付き文字描画データ1に含まれる文字コード情報が表示されるようになっている。

なお本実施形態の表示編集プログラムでは、レイアウト編集用データのうち<High RecognitionText></High RecognitionText>で挟まれた文字列を、要約情報ウィンドウに表示するレイアウト付き文字描画データ1とする。また、レイアウト編集用データのうち<Low RecognitionText></Low RecognitionText>で挟まれた文字列を、要約情報ウィンドウに表

50

示しないレイアウト付き文字描画データ2とする。したがって、本実施形態2で例示する図10の電子ドキュメント1000に対しては、図10の1003部分に含まれる文字列「ドキュメント」と「写真写真」が、要約情報に使用されるデータとして図11の1104のように表示されている。

#### 【0100】

以上説明したように、本実施形態の画像処理装置は、レイアウト付き文字描画データを文字認識率によって決定された格納先に保存した電子データを生成する。このように生成された電子データに対して、表示・編集プログラム121は、文字認識率の高いデータを要約情報としても提示するのでユーザは信頼性の高い情報をさらに有効に再利用できる。すなわち本実施形態では、文字認識率の高いデータを積極的に再利用できるとともに、閲覧しやすい状態での提供が可能となる。また、レイアウト編集ウィンドウ上にレイアウト付き文字描画データ（すなわち文字認識結果の文字コード）が描画されているので、テキスト検索を行うと、その文字の個所がハイライト表示される。したがって、ユーザ所望のテキストの表示位置を検索しやすくなる。

10

#### 【0101】

なお、表示・編集プログラム121がパワーポイントの場合、レイアウト編集ウィンドウ602はスライドペイン、テキスト編集ウィンドウ603はノートペイン、要約情報表示ウィンドウ604はアウトラインペインである。この場合、文字認識結果の類似度が高い文字列だけがアウトラインペインに表示されるように、類似度が高い文字認識結果をレイアウト編集用データ部分に記述した電子ドキュメントを生成する。

20

#### 【0102】

<実施形態3>

次に、本発明の第3の実施形態について説明する。

#### 【0103】

上述の実施形態2では、格納先決定部716は、図9の決定規則に従って格納先を決定した。この実施形態2と同様の構成による本実施形態3においては、図12に示す決定規則に基づいて格納先を決定するものとする。

#### 【0104】

本実施形態3では、レイアウト付き文字描画データは、文字コードの文字認識率が高いものは723のレイアウト付き文字描画データ1として、電子ドキュメント720における721のレイアウト編集用データ内へと格納先が決定される。一方、文字コードの文字認識率の低いものは、文字コードの情報を削除したうえで同じく723のレイアウト付き文字描画データ1として、電子ドキュメント720における721のレイアウト編集用データ内へと格納先が決定される。

30

#### 【0105】

[電子ドキュメントの具体例]

ここで、本実施形態3における電子ドキュメントの具体例を、図13を用いて説明する。

#### 【0106】

図13の電子ドキュメント1300は、図8の入力画像800に対し、本実施形態3において生成された電子ドキュメント720の具体例である。

40

#### 【0107】

電子ドキュメント1300のレイアウト編集用データを格納する部分である1301には、文字認識率の高い文字領域801および803のデータとして図13に示す1303と1305が、723のレイアウト付き文字描画データ1として格納されている。また、文字認識率の低い文字領域802のデータとして図13に示す1304が、文字コードデータが削除された状態で、レイアウト付き文字描画データ1として格納されている。

#### 【0108】

[表示・編集プログラムによる表示例]

次に、図8の画像800から生成された電子ドキュメントの具体例である図13の電子

50

ドキュメント 1300 を、図 1 のパーソナルコンピュータ 120 で実行される表示・編集プログラム 121 で処理する際の表示例を、図 14 を用いて説明する。

【0109】

図 14 は表示・編集プログラム 121 に上記電子ドキュメント 1300 を入力した場合の表示例である。

【0110】

レイアウト編集ウィンドウ 1402 およびテキスト編集ウィンドウ 1403 には図 11 と同様の表示がなされている。

【0111】

1404 の要約情報表示ウィンドウには、文字領域 801 および 803 の認識結果の間に、正確に認識できなかった文字領域 802 のデータ（“これはサンプル電子ドキュメント”）に対応する空行 1405 が表示されている。ユーザはこの空行 1405 に対して手動で、あるいはテキスト編集ウィンドウ 1403 の該当部分の認識結果のテキスト（“これはサンプル電子ドキュメント”）からコピーし編集することで、その文字列を挿入し、要約情報を完成することが可能である。

【0112】

またレイアウト編集ウィンドウ内の文字コードと要約情報表示ウィンドウ内の文字コードはリンクしているので、空行 1405 に文字を挿入すると、対応するレイアウト付き文字描画領域である 1406 部分にも文字が挿入される。なお、レイアウト編集ウィンドウ内の文字列は背景画像に隠れているのでそのままでは見えない。しかしながら、この文字の描画領域を背景やベクトル描画データの描画領域よりも前面になるような操作を加えることで、入力された文字列が、原画像から抽出された文字レイアウト情報、すなわち位置や大きさなどを備えた状態で利用可能になる。

【0113】

以上説明したように、本実施形態の画像処理装置は、レイアウト付き文字描画データを文字認識率によって出力形態を変えて保存した電子データを生成する。このように生成された電子データに対し、表示・編集プログラム 121 は、レイアウト編集ウィンドウ上では文字認識率の低い文字情報をあえてユーザに提供せず編集可能な空行を提示するため、ユーザは修正作業や画像の復元が容易になる。すなわち本実施形態では、文字認識率の低いデータを修正しやすい状態での提供が可能である。

【0114】

なお、表示・編集プログラム 121 がパワーポイントの場合、レイアウト編集ウィンドウ 602 はスライドペイン、テキスト編集ウィンドウ 603 はノートペイン、要約情報表示ウィンドウ 604 はアウトラインペインである。この場合、文字認識結果の類似度が高い文字列がアウトラインペインに表示され、類似度が低い文字領域が空行でアウトラインペインに表示されるように、レイアウト編集用データ部分にデータを記述した電子ドキュメントを生成する。

【0115】

以上、本発明の諸実施形態について説明した。

【0116】

なお、本発明の目的は、上述した実施形態で示したフローチャートの手順を実現するプログラムコードを記憶した記憶媒体から、システムあるいは装置のコンピュータ（または CPU や MPU）がその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。この場合、記憶媒体から読み出されたプログラムコード自体が、コンピュータに、上述した実施形態の機能を実現させることになる。そのため、このプログラムコード及びプログラムコードを記憶 / 記録したコンピュータ読み取り可能な記憶媒体も本発明の一つを構成することになる。

【0117】

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、

10

20

30

40

50

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。

【0118】

また、前述した実施形態の機能は、コンピュータが、読み出したプログラムを実行することによって実現される。また、このプログラムの実行とは、そのプログラムの指示に基づき、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う場合も含まれる。

【0119】

さらに、前述した実施形態の機能は、コンピュータに挿入された機能拡張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よっても実現することができる。この場合、まず、記憶媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張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る。その後、そのプログラムの指示に基づき、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う。こうした機能拡張ボードや機能拡張ユニットによる処理によっても前述した実施形態の機能が実現される。

【図面の簡単な説明】

【0120】

【図1】本発明の第1の実施形態におけるシステム構成の一例を示す図である。

【図2】同実施形態1における画像処理装置の電子ドキュメント生成に係る構成と入出力されるデータを説明する図である。

【図3】同実施形態1において説明に用いる入力画像を示す図である。

【図4】同実施形態1においてデータの格納先を決定するために用いる決定規則を示す図である。

【図5】同実施形態1において、図3の入力画像に対して生成される電子ドキュメントの具体例を示す図である。

【図6】同実施形態1における表示・編集プログラムに図5の電子ドキュメントを入力した場合の表示例を示す図である。

【図7】実施形態1と同様のシステム構成で実施される、本発明の第2の実施形態における画像処理装置の電子ドキュメント生成に係る構成と入出力されるデータを説明する図である。

【図8】同実施形態2において説明に用いる入力画像を示す図である。

【図9】同実施形態2においてデータの格納先を決定するために用いる決定規則を示す図である。

【図10】同実施形態2において、図8の入力画像に対して生成される電子ドキュメントの具体例を示す図である。

【図11】同実施形態2における表示・編集プログラムに図10の電子ドキュメントを入力した場合の表示例を示す図である。

【図12】本発明の第3の実施形態においてデータの格納先を決定するために用いる決定規則を示す図である。

【図13】同実施形態3において、図8の入力画像に対して生成された電子ドキュメントの具体例を示す図である。

【図14】同実施形態3における表示・編集プログラムに図13の電子ドキュメントを入力した場合の表示例を示す図である。

【符号の説明】

【0121】

100 画像処理装置

101 スキャナ

102 CPU

103 メモリ

104 ハードディスク

105 ネットワークI/F

10

20

30

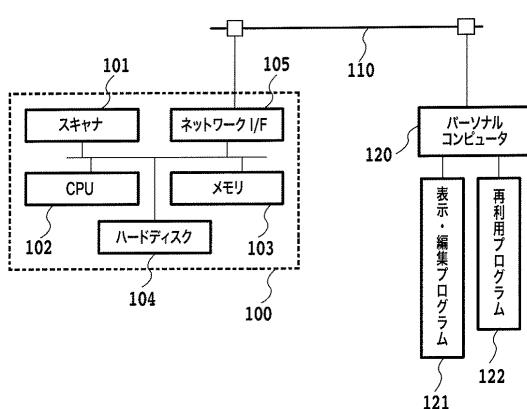
40

50

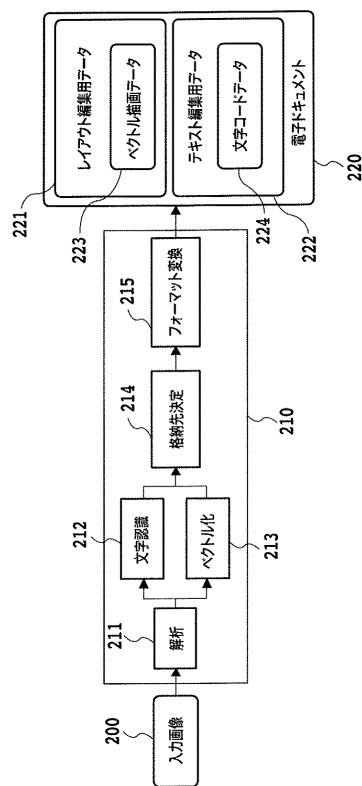
1 1 0 ネットワーク  
 1 2 0 パーソナルコンピュータ  
 1 2 1 表示・編集プログラム  
 1 2 2 再利用プログラム  
 2 0 0 入力画像  
 2 1 0 電子ドキュメント生成部（電子ドキュメント生成プログラム）  
 2 1 1 解析部  
 2 1 2 文字認識部  
 2 1 3 ベクトル化部  
 2 1 4 格納先決定部  
 2 1 5 フォーマット変換部  
 2 2 0 電子ドキュメント  
 2 2 1 レイアウト編集用データ  
 2 2 2 テキスト編集用データ  
 2 2 3 ベクトル描画データ  
 2 2 4 文字コードデータ

10

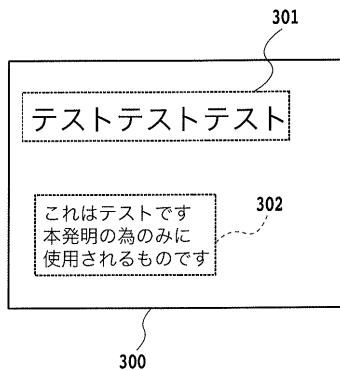
【図1】



【図2】



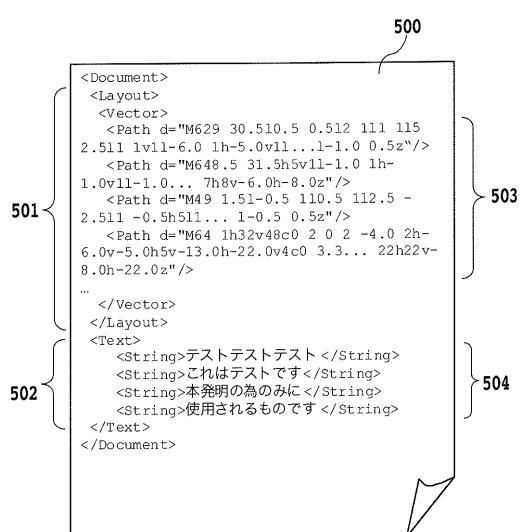
【図3】



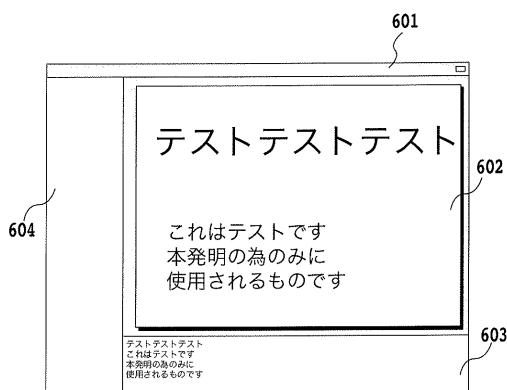
【図4】

データ	格納先
文字認識データ	テキスト編集用データ
ベクトルデータ	レイアウト編集用データ

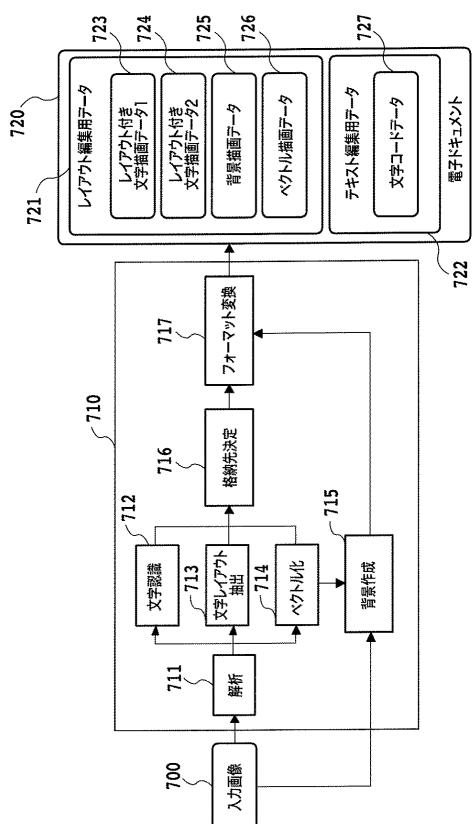
【図5】



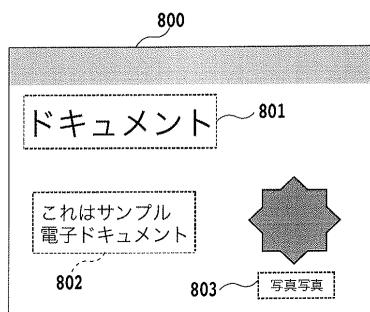
【図6】



【図7】



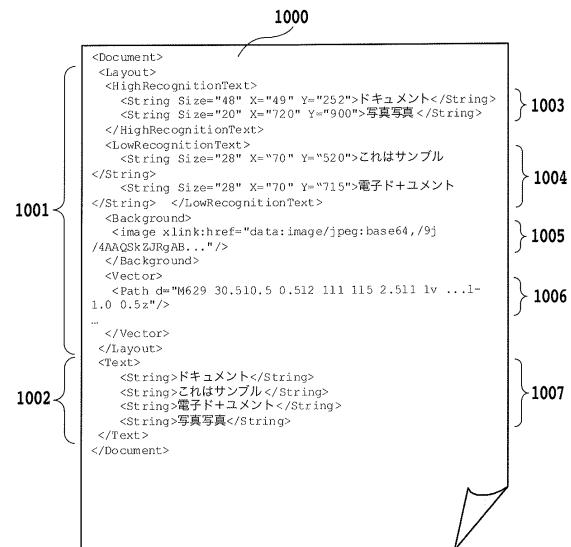
【図 8】



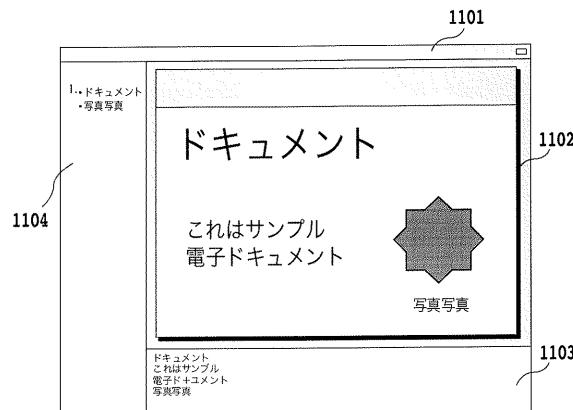
【図 9】

データ	格納先
文字認識データ	テキスト用編集データ
レイアウト付き文字データ	レイアウト編集用 (レイアウト付き文字データ1 として)
	レイアウト編集用 (レイアウト付き文字データ2 として)
ベクトルデータ	レイアウト編集用
背景データ	レイアウト編集用

【図 10】



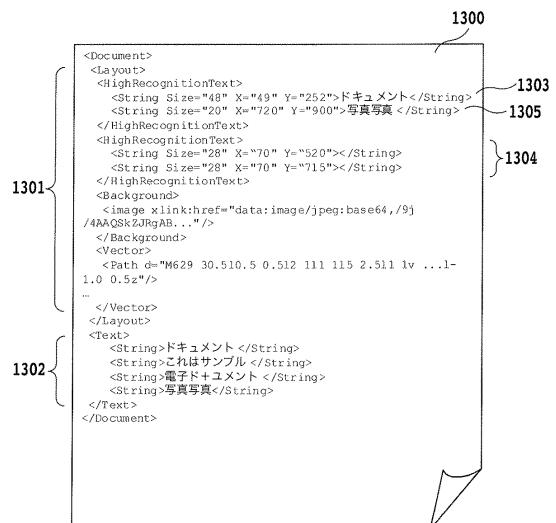
【図 11】



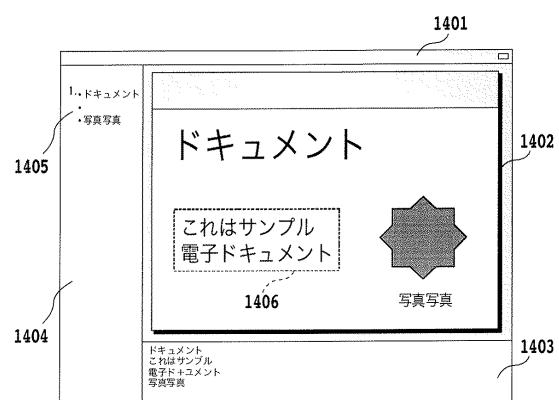
【図 12】

データ	格納先
文字認識データ	テキスト用編集データ
レイアウト付き文字データ	レイアウト編集用 (レイアウト付き文字データ1 として)
	レイアウト編集用 (文字情報を削除したうえで レイアウト付き文字データ1 として)
ベクトルデータ	レイアウト編集用
背景データ	レイアウト編集用

【図13】



【図14】



---

フロントページの続き

(72)発明者 榎本 誠

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

(72)発明者 金田 北洋

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 岡本 俊威

(56)参考文献 特開2008-092419(JP, A)

特開2006-155380(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06T 11/60 - 11/80

H04N 1/387