

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-228302

(P2005-228302A)

(43) 公開日 平成17年8月25日(2005.8.25)

(51) Int.Cl.<sup>7</sup>

G06F 3/03

G06T 1/00

H04N 1/387

F I

G06F 3/03

G06F 3/03

G06T 1/00

H04N 1/387

テーマコード (参考)

5B057

5B068

5C076

審査請求 未請求 請求項の数 39 O L 外国語出願 (全 35 頁)

(21) 出願番号 特願2005-1434 (P2005-1434)

(22) 出願日 平成17年1月6日(2005.1.6)

(31) 優先権主張番号 10/753, 176

(32) 優先日 平成16年1月6日(2004.1.6)

(33) 優先権主張国 米国 (US)

(特許庁注: 以下のものは登録商標)

1. Bluetooth

(71) 出願人 500046438

マイクロソフト コーポレーション

アメリカ合衆国 ワシントン州 9805

2-6399 レッドモンド ワン マイ

クロソフト ウェイ

(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 ジャン ワン

アメリカ合衆国 98052 ワシントン

州 レッドモンド ワン マイクロソフト

ウェイ マイクロソフト コーポレーシ

ョン内

最終頁に続く

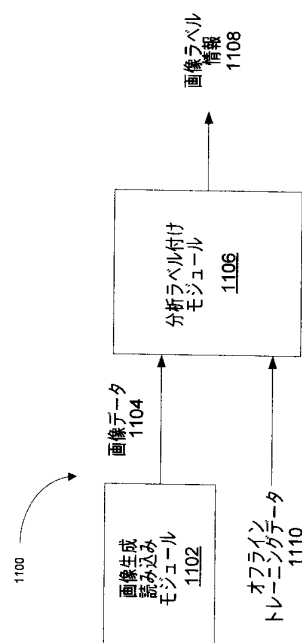
(54) 【発明の名称】 位置エンコードされた文書画像の分析およびラベル付け

(57) 【要約】

【課題】 位置エンコードされた文書画像の分析およびラベル付けを提供すること。

【解決手段】 迷路パターン透かしなどの位置エンコードされた情報を含む文書画像の分析と、テキストなどの文書の内容が位置エンコーディング情報を覆う程度に基づいて画像にラベル付けを行う。そのような覆いの程度に応じて、カメラによって取り込まれた画像の文書内の位置を判定するのに十分な位置エンコーディングビットを、文書のカメラによって取り込まれた画像から抽出することが不可能になる場合がある。分析ラベル付けモジュールが、画像生成取り込みモジュールによって出力される画像データおよびオフライントレーニングデータを入力として受け取り、分析ラベル付け処理を実行し、画像ラベル情報を出力する。文書分析ラベル付け処理の結果を、位置エンコードされた文書内のカメラによって取り込まれた画像の位置を効率的に判定するのに使用することができる。

【選択図】 図11



## 【特許請求の範囲】

## 【請求項 1】

文書画像から抽出できるエンコードされた位置情報の量を示す位置エンコードされた情報を含む前記文書画像にラベル付けを行う方法であって、

前記文書画像を得るステップと、

少なくとも、実質的に位置エンコードされた情報だけを含み、かつ実質的に文書内容を含まない第 1 の型、ならびに前記位置エンコードされた情報の少なくとも一部を覆う文書内容を含む第 2 の型から選択される型となるように前記文書画像にラベル付けを行うステップと

を備えたことを特徴とする方法。

10

## 【請求項 2】

前記文書画像を得るステップは、電子文書を印刷された文書に対応するビットマップ表現とするステップを含むことを特徴とする請求項 1 に記載の方法。

## 【請求項 3】

前記文書画像を得るステップは、読み取られた紙の文書进行处理するステップを含むことを特徴とする請求項 1 に記載の方法。

## 【請求項 4】

前記第 2 の型は、前記文書内容による前記位置エンコードされた情報を覆うそれぞれ異なる程度を表す複数の下位の型に分割されることを特徴とする請求項 1 に記載の方法。

## 【請求項 5】

カメラによって取り込まれた画像から抽出できるエンコードされた位置情報の量を示す位置エンコードされた情報を含む前記画像にラベル付けを行う方法であって、

前記カメラによって取り込まれた画像を得るステップと、

少なくとも、実質的に位置エンコードされた情報だけを含み、かつ実質的に文書内容を含まない第 1 の型、および前記位置エンコードされた情報の少なくとも一部を覆う文書内容を含む第 2 の型から選択される型となるように前記画像にラベル付けを行うステップとを備えたことを特徴とする方法。

20

## 【請求項 6】

前記画像が前記第 1 の型または前記第 2 の型のいずれであるかを判定するために、グラディエント画像ヒストグラムをサポート間隔が使用されることを特徴とする請求項 5 に記載の方法。

30

## 【請求項 7】

グラディエント画像を得るために前記画像にグラディエント演算子を用いるステップをさらに備えたことを特徴とする請求項 6 に記載の方法。

## 【請求項 8】

前記グラディエント演算子は、S o b e l エッジ演算子であることを特徴とする請求項 7 に記載の方法。

## 【請求項 9】

前記グラディエント画像のヒストグラムを生成するステップをさらに備えたことを特徴とする請求項 8 に記載の方法。

40

## 【請求項 10】

前記グラディエント画像ヒストグラムをサポート間隔として、非 0 の値を有する前記ヒストグラムの x 軸上の最大値を使用するステップをさらに備えたことを特徴とする請求項 9 に記載の方法。

## 【請求項 11】

オフライントレーニングセッションおよびオンラインラベル付けセッションは、前記画像が前記第 1 の型および前記第 2 の型のいずれであるかを判定するのに使用されることを特徴とする請求項 6 に記載の方法。

## 【請求項 12】

前記第 1 の型の画像と前記第 2 の型の画像とを区別するために使用される閾値は、トレ

50

ーニングデータ画像に対して実行される前記オフライントレーニングセッションの結果に基づいて選択されることを特徴とする請求項 1 1 に記載の方法。

【請求項 1 3】

前記閾値は、前記オンラインラベル付けセッション中に、前記画像が前記第 1 の型または前記第 2 の型のいずれであるかを判定するため、前記画像のグラディエント画像ヒストグラムの前記サポート間隔と比較されることを特徴とする請求項 1 2 に記載の方法。

【請求項 1 4】

文書画像から抽出されるエンコードされた位置情報の量を示す位置エンコードされた情報を含む前記文書画像にラベル付けを行うシステムであって、

前記文書画像を得る画像生成取り込みモジュールと、

少なくとも、実質的に位置エンコードされた情報だけを含み、かつ実質的に文書内容を含まない第 1 の型、および前記位置エンコードされた情報の少なくとも一部を覆う文書内容を含む第 2 の型から選択される型となるように前記文書画像にラベル付けを行う分析ラベル付けモジュールと

を備えたことを特徴とするシステム。

【請求項 1 5】

前記画像生成取り込みモジュールは、電子文書を印刷された文書に対応するビットマップ表現とすることを特徴とする請求項 1 4 に記載のシステム。

【請求項 1 6】

前記画像生成取り込みモジュールは、読み取られた紙の文書进行处理することを特徴とする請求項 1 4 に記載のシステム。

【請求項 1 7】

前記第 2 の型は、前記文書内容による前記位置エンコードされた情報を覆うそれぞれ異なる程度を表す複数の下位の型に分割されることを特徴とする請求項 1 4 に記載のシステム。

【請求項 1 8】

カメラによって取り込まれた画像から抽出されるエンコードされた位置情報の量を示すために、位置エンコードされた情報を含む前記画像にラベル付けを行うシステムであって、

前記カメラによって取り込まれた画像を得る画像生成取り込みモジュールと、

少なくとも、実質的に位置エンコードされた情報だけを含み、かつ実質的に文書内容を含まない第 1 の型、および前記位置エンコードされた情報の少なくとも一部を覆う文書内容を含む第 2 の型から選択される型となるように前記画像にラベル付けを行う分析ラベル付けモジュールと

を備えたことを特徴とするシステム。

【請求項 1 9】

前記画像が前記第 1 の型または前記第 2 の型のいずれであるかを判定するために、グラディエント画像ヒストグラムのサポート間隔が使用されることを特徴とする請求項 1 8 に記載のシステム。

【請求項 2 0】

前記分析ラベル付けモジュールは、グラディエント画像を得るために前記画像にグラディエント演算子を用いることを特徴とする請求項 1 9 に記載のシステム。

【請求項 2 1】

前記グラディエント演算子は、S o b e l エッジ演算子であることを特徴とする請求項 2 0 に記載のシステム。

【請求項 2 2】

前記分析ラベル付けモジュールは、前記グラディエント画像のヒストグラムを生成することを特徴とする請求項 1 9 に記載のシステム。

【請求項 2 3】

前記分析ラベル付けモジュールは、前記グラディエント画像ヒストグラムのサポート間

10

20

30

40

50

隔として、非 0 の値を有する前記ヒストグラム of x 軸上の最大値を使用することを特徴とする請求項 22 に記載のシステム。

【請求項 24】

前記分析ラベル付けモジュールは、前記画像が前記第 1 の型および前記第 2 の型のいずれであるかを判定するために、オフライントレーニングセッションおよびオンラインラベル付けセッションを実行することを特徴とする請求項 23 に記載のシステム。

【請求項 25】

前記第 1 の型の画像と前記第 2 の型の画像とを区別するために使用される閾値は、トレーニングデータ画像に対して実行される前記オフライントレーニングセッションの結果に基づいて選択されることを特徴とする請求項 24 に記載のシステム。

10

【請求項 26】

前記分析ラベル付けモジュールは、前記オンラインラベル付けセッション中に、前記画像が前記第 1 の型または前記第 2 の型のいずれであるかを判定するため、前記閾値を前記画像のグラディエント画像ヒストグラムのサポート間隔と比較することを特徴とする請求項 25 に記載のシステム。

【請求項 27】

文書画像から抽出されるエンコードされた位置情報の量を示す位置エンコードされた情報を含む前記文書画像にラベル付けを行うコンピュータ読み取り可能な命令を備えたコンピュータ読み取り可能な媒体であって、

前記文書画像を得るステップと、

20

少なくとも、実質的に位置エンコードされた情報だけを含み、かつ実質的に文書内容を含まない第 1 の型、および前記位置エンコードされた情報の少なくとも一部を覆う文書内容を含む第 2 の型から選択される型となるように前記文書画像にラベル付けを行うステップと

を実行することによってラベル付けすることを特徴とするコンピュータ読み取り可能な媒体。

【請求項 28】

前記文書画像を得るステップは、電子文書を印刷された文書に対応するビットマップ表現とするステップを含むことを特徴とする請求項 27 に記載のコンピュータ読み取り可能な媒体。

30

【請求項 29】

前記文書画像を得るステップは、読み取られた紙の文書を処理するステップをさらに含むことを特徴とする請求項 27 に記載のコンピュータ読み取り可能な媒体。

【請求項 30】

前記第 2 の型は、前記文書内容による前記位置エンコードされた情報を覆うそれぞれ異なる程度を表す複数の下位の型に分割されることを特徴とする請求項 27 に記載のコンピュータ読み取り可能な媒体。

【請求項 31】

カメラによって取り込まれた画像から抽出されるエンコードされた位置情報の量を示す位置エンコードされた情報を含む前記画像にラベル付けを行うコンピュータ読み取り可能な命令を備えたコンピュータ読み取り可能な媒体であって、

40

前記カメラによって取り込まれた画像を得るステップと、

少なくとも、実質的に位置エンコードされた情報だけを含み、かつ実質的に文書内容を含まない第 1 の型、および前記位置エンコードされた情報の少なくとも一部を覆う文書内容を含む第 2 の型から選択される型となるように前記画像にラベル付けを行うステップと

を実行することによってラベル付けすることを特徴とするコンピュータ読み取り可能な媒体。

【請求項 32】

前記画像が前記第 1 の型または前記第 2 の型のいずれであるかを判定するために、グラディエント画像ヒストグラムのサポート間隔が使用されることを特徴とする請求項 31 に

50

記載のコンピュータ読み取り可能な媒体。

【請求項 3 3】

グラディエント画像を得るために前記画像にグラディエント演算子を用いるステップを実行するコンピュータ実行可能命令をさらに備えたことを特徴とする請求項 3 2 に記載のコンピュータ読み取り可能な媒体。

【請求項 3 4】

前記グラディエント演算子は、S o b e l エッジ演算子であることを特徴とする請求項 3 3 に記載のコンピュータ読み取り可能な媒体。

【請求項 3 5】

前記グラディエント画像のヒストグラムを生成するステップを実行するコンピュータ実行可能命令をさらに備えたことを特徴とする請求項 3 4 に記載のコンピュータ読み取り可能な媒体。 10

【請求項 3 6】

前記グラディエント画像ヒストグラムのサポート間隔として、非 0 の値を有する前記ヒストグラムの x 軸上の最大値を使用するステップを実行するコンピュータ実行可能命令をさらに備えたことを特徴とする請求項 3 5 に記載のコンピュータ読み取り可能な媒体。

【請求項 3 7】

前記画像が前記第 1 の型および前記第 2 の型のいずれであるかを判定するために、オフライントレーニングセッションおよびオンラインラベル付けセッションが使用されることを特徴とする請求項 3 2 に記載のコンピュータ読み取り可能な媒体。 20

【請求項 3 8】

前記第 1 の型の画像と前記第 2 の型の画像とを区別するために使用される閾値は、トレーニングデータ画像に対して実行される前記オフライントレーニングセッションの結果に基づいて選択されることを特徴とする請求項 3 7 に記載のコンピュータ読み取り可能な媒体。

【請求項 3 9】

前記閾値は、前記オンラインラベル付けセッション中に、前記画像が前記第 1 の型または前記第 2 の型のいずれであるかを判定するため、前記画像のグラディエント画像ヒストグラムの前記サポート間隔と比較されることを特徴とする請求項 3 8 に記載のコンピュータ読み取り可能な媒体。 30

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明の実施形態は、位置エンコードされた ( p o s i t i o n a l l y e n c o d e d ) 媒体とデジタルペンの間の相互作用に関し、より詳細には、本発明の実施形態は、文書の内容が位置エンコーディング情報を覆う程度に基づいた、位置エンコードされた情報を含む文書画像のラベル付けに関する。

【背景技術】

【0 0 0 2】

コンピュータユーザは、パーソナルコンピュータと対話する形として、マウスおよびキーボードを使用することに慣れている。パーソナルコンピュータは、書かれる文書に対して多くの利点を有するが、ほとんどのユーザは、印刷される紙を使用する特定の機能を実行している。これらの機能の一部に、書かれた文書を読み、注釈を付ける機能が含まれる。注釈の場合に、ユーザによって書かれた注釈のゆえに、印刷された文書がより重要になる。しかし、注釈を有する印刷された文書を用いることの 1 つの問題点は、後に電子形式の文書に注釈を入力する必要があることである。これは、オリジナルユーザまたは別のユーザが、注釈全体を読み進め、それをパーソナルコンピュータに入力することを必要とする。ある場合には、ユーザは、注釈およびオリジナルテキストを読み取り、これによって新しい文書を作成する。これらの複数のステップのために、印刷された文書とその文書のデジタルバージョンとの間の相互作用を基本的な操作を繰り返すことによって処理するこ 40 50

とは難しい。さらに、読み取られた画像は、多くの場合、変更不能である。注釈をオリジナルテキストから分離する方法がない場合がある。これによって、注釈を使用することが困難になる。したがって、注釈を扱う改善された方法が必要である。

【0003】

【非特許文献1】 Douglas W. Clark and Lih-Jyh Weng, "Maximal and Near-Maximal Shift Register Sequences: Efficient Event Counters and Easy Discrete Logarithms," IEEE Transactions on Computers 43.5 (May 1994, pp 560-568)

【発明の開示】

【発明が解決しようとする課題】

【0004】

手書き情報を取り込む技法の1つが、書いている間に位置を判定できるペンを使用することによるものである。この機能を提供するペンの1つが、Anoto Inc. 社のAnoto penである。このペンは、カメラを使用して、事前定義のパターンを用いてエンコードされた紙の画像を取り込むことによって機能する。画像パターンの例を、図21に示す。このパターンを（Anoto社の）Anoto penによって使用し、紙の上でのペンの位置が判定される。しかし、Anoto penによって使用されるシステムによる位置の判定がどれほど効率的であるかは不確かである。取り込まれた画像の位置の効率的な判定を提供するために、取り込まれた画像の効率的なデコードを提供するシステムが必要である。

【0005】

文書に注釈を付ける時に、ユーザは、文書に関してペン先を移動することによって、文書にマークを付けることができる。ペン先の経路に、複数のストロークが含まれる場合があり、各ストロークは、一連の取り込まれた画像に対応する。したがって、文書の注釈を処理するためのペンの経路の効率的な識別が望ましい。

【0006】

透かしを入れられた迷路パターンなどの位置エンコードされた情報の一部は、テキストおよび／またはグラフィックスなどの文書の内容によって覆われる可能性がある。文書の内容が、文書の1区域の位置エンコーディング情報の比較的小さい量を覆う（または覆わない）時に、文書内の区域の位置を、位置を判定する高費用の計算を実行せずに効率的に判定することができる。しかし、文書の内容が、文書の1区域の位置エンコーディング情報の比較的多い量を覆う時には、より計算集中型の技法を使用して、文書内の区域の位置を判定する必要がある場合がある。したがって、位置エンコーディング情報が覆われる量が比較的小さい場合と、覆われる量が比較的多い場合とを区別する技法が、ユーザが位置エンコードされた情報を含む文書と相互作用する効率を高めるために望ましい。

【課題を解決するための手段】

【0007】

本発明の実施形態は、迷路パターンなどの位置エンコードされた情報を含む文書画像を分析し、取り込まれた画像にラベル付けを行うことに関する。そのような文書分析の結果を用いて、位置エンコードされた文書内のカメラによって取り込まれた画像の位置を効率的に判定することができる。

【0008】

システムに、画像取り込みモジュールおよび分析ラベル付けモジュール（analysis-and-labeling module）が含まれる。分析ラベル付けモジュールは、画像取り込みモジュールによって出力された取り込まれた画像およびオフライントレーニングデータを入力として受け取り、分析ラベル付け処理を実行し、画像ラベル情報を出力する。

【0009】

印刷された文書は、迷路パターンによって透かしを入れられる。そのような迷路パターンは、テキストなどの文書の内容によって覆われる可能性がある。そのような覆いの程度によっては、カメラによって取り込まれた画像の文書内の位置を判定するのに十分な位置

10

20

30

40

50

エンコーディングビットを、カメラによって取り込まれた画像から抽出することが不可能になる場合がある。

【0010】

透かしを入れられた文書画像を、分析し、ラベル付けを行う。分析およびラベル付けは、文書内の特定の位置での迷路パターンの可視性の程度に基づく、透かしを入れられた文書内のx-y位置を判定する能力に関連する。

【0011】

本発明の追加の特徴および長所は、後述の詳細な説明を検討する時に明かになる。

【0012】

本発明の前述の要約ならびに好ましい実施形態の次の詳細な説明は、添付する図面と共に読まれる時によりよく理解されるが、添付する図面は、請求される発明に関する制限ではなく、例として含まれる。

【発明を実施するための最良の形態】

【0013】

本発明の態様は、より大きい画像に関する取り込まれた画像の位置の判定に関する。本明細書に記載の位置判定の方法およびシステムは、マルチファンクションペンと組み合わせて使用することができる。

【0014】

下記は、読者の利益のために見出しによって区切られている。見出しには、用語、汎用コンピュータ、画像取り込みペン、配列のエンコード、デコード、エラー訂正、位置判定、ならびにカメラによって取り込まれた画像の分析およびラベル付けが含まれる。

【0015】

(I. 用語)

ペン — インクをたくわえる能力を含んでも含まなくてもよいすべての書き込み器具。いくつかの例で、インク機能を有しないスタイラスを、本発明の実施形態に従ってペンとして使用することができる。

【0016】

カメラ — 紙または他の媒体から画像を取り込むことができる画像取り込みシステム。

【0017】

(II. 汎用コンピュータ)

図1は、本発明の様々な態様を実施するために使用することができる、普通の汎用デジタルコンピュータ環境の例の機能ブロック図である。図1では、コンピュータ100に、処理ユニット110、システムメモリ120、およびシステムメモリを含む様々なシステムコンポーネントを処理ユニット110に接続するシステムバス130が含まれる。システムバス130は、メモリバスまたはメモリコントローラ、周辺バス、および様々なバスアーキテクチャのいずれかを使用するローカルバスを含む複数の種類のバス構造のいずれかとすることができる。システムメモリ120に、読み取り専用メモリ(ROM)140およびランダムアクセスメモリ(RAM)150が含まれる。

【0018】

起動中などにコンピュータ100内の要素の間で情報を転送することを助ける基本ルーチンを含む基本入出力システム160(BIOS)が、ROM 140に保管される。コンピュータ100には、ハードディスク(図示せず)から読み取り、これに書き込むハードディスクドライブ170、取外し可能な磁気ディスク190から読み取り、これに書き込む磁気ディスクドライブ180、およびCD ROMまたは他の光媒体などの取外し可能な光ディスク192から読み取るか、これに書き込む光ディスクドライブ191も含まれる。ハードディスクドライブ170、磁気ディスクドライブ180、および光ディスクドライブ191は、それぞれハードディスクドライブインターフェース192、磁気ディスクドライブインターフェース193、および光ディスクドライブインターフェース194によってシステムバス130に接続される。ドライブおよびそれに関連するコンピュー

タ読み取り可能な媒体は、パーソナルコンピュータ100のコンピュータ読み取り可能な命令、データ構造、プログラムモジュール、および他のデータの不揮発性ストレージを提供する。磁気カセット、フラッシュメモ리카ード、デジタルビデオディスク、ベルヌーイカートリッジ、ランダムアクセスメモリ(RAM)、読み取り専用メモリ(ROM)、および類似物など、コンピュータによってアクセス可能なデータを保管することができる他の種類のコンピュータ読み取り可能な媒体も、例示のオペレーティング環境で使用できることは、当業者に理解されるであろう。

#### 【0019】

オペレーティングシステム195、1つまたは複数のアプリケーションプログラム196、他のプログラムモジュール197、およびプログラムデータ198を含む複数のプログラムモジュールを、ハードディスクドライブ170、磁気ディスク190、光ディスク192、ROM 140、またはRAM 150に保管することができる。ユーザは、キーボード101およびポインティングデバイス102などの入力装置を介して、コンピュータ100にコマンドおよび情報を入力することができる。他の入力装置(図示せず)に、マイクロホン、ジョイスティック、ゲームパッド、衛星パラボラアンテナ、スキャナ、または類似物を含めることができる。これらおよび他の入力装置は、しばしば、システムバスに結合されるシリアルポートインターフェース106を介して処理ユニット110に接続されるが、パラレルポート、ゲームポート、またはUSB(universal serial bus)などの他のインターフェースによって接続することができる。さらに、これらの装置を、適当なインターフェース(図示せず)を介してシステムバス130に直接に結合することができる。モニタ107または他の種類のディスプレイ装置も、ビデオアダプタ108などのインターフェースを介してシステムバス130に接続される。モニタのほかに、パーソナルコンピュータに、通常は、スピーカおよびプリンタなどの他の周辺出力装置(図示せず)が含まれる。好ましい実施形態では、ペンデジタイザ165および付随するペンまたはスタイラス166が、手書き入力をデジタルに取り込むために設けられる。ペンデジタイザ165とシリアルポートの直接接続が図示されているが、実際には、当技術分野で既知のように、ペンデジタイザ165を、パラレルポートまたは他のインターフェースおよびシステムバス130を介して処理ユニット110に直接に結合することができる。さらに、デジタイザ165が、モニタ107と別に図示されているが、デジタイザ165の使用可能な入力領域が、モニタ107の表示領域と同一の広がりを持つことが好ましい。さらに、デジタイザ165を、モニタ107と一体化することができ、あるいは、モニタ107の上に置かれるか、または他の形で追加される別々の装置として存在することができる。

#### 【0020】

コンピュータ100は、リモートコンピュータ109などの1つまたは複数のリモートコンピュータへの論理接続を使用してネットワーク化された環境で動作することができる。リモートコンピュータ109は、サーバ、ルータ、ネットワークPC、ピア装置、または他の一般的なネットワークノードとすることができ、通常は、コンピュータ100に関して上で説明した要素の多くまたはすべてが含まれるが、図1には、メモリストレージ装置111だけを示した。図1に示された論理接続に、ローカルエリアネットワーク(LAN)112および広域ネットワーク(WAN)113が含まれる。そのようなネットワーク環境は、オフィス、会社規模のコンピュータネットワーク、イントラネット、およびインターネットで一般的なものである。

#### 【0021】

LANネットワーク環境で使用される時に、コンピュータ100は、ネットワークインターフェースまたはネットワークアダプタ114を介してローカルネットワーク112に接続される。WANネットワーク環境で使用される時に、パーソナルコンピュータ100に、通常、インターネットなどの広域ネットワーク113を介する通信を確立する、モデム115または他の手段が含まれる。モデム115は、内蔵または外付けとすることができるが、シリアルポートインターフェース106を介してシステムバス130に

接続される。ネットワーク化された環境では、パーソナルコンピュータ100に関して図示したプログラムモジュールまたはその一部を、リモートメモリストレージ装置に保管することができる。

#### 【0022】

図示のネットワーク接続が、例示的であり、コンピュータ間の通信リンクを確立する他の技法を使用できることを理解されたい。TCP/IP、イーサネット（登録商標）、FTP、HTTP、Bluetooth、IEEE 802.11x、および類似物などの様々な周知のプロトコルのいずれかが假定され、このシステムは、クライアントサーバ構成で動作し、ユーザがウェブベースサーバからウェブページを取り出すことを可能にする。様々な普通のウェブブラウザを使用して、ウェブページのデータを表示し、操作する

10

#### 【0023】

（III. 画像取り込みペン）

本発明の態様に、エンコードされたデータストリームを、それを表す表示形式に置換することが含まれる（例えば、図4Bについて説明するように、エンコードされたデータストリームを使用して、グラフィカルパターンを作成する）。表示形式は、印刷された紙（または他の物理媒体）とするか、別の画像または画像群と共にエンコードされたデータストリームを投影するディスプレイとすることができる。例えば、エンコードされたデータストリームは、（例えば、文書のテキストを表す）紙の上の物理的なグラフィカル画像または表示される画像にオーバーレイするグラフィカル画像とするか、ディスプレイスクリーン上の物理的（変更不能）グラフィカル画像とすることができる（したがって、ペンによって取り込まれた画像部分を、ディスプレイスクリーン上で突き止めることができる）。

20

#### 【0024】

取り込まれた画像の位置のこの判定を用いて、紙、媒体、またはディスプレイスクリーンとのユーザとの相互作用の位置を判定することができる。本発明のいくつかの態様で、ペンを、紙に書き込むインクペンとすることができる。他の態様で、ペンを、ユーザがコンピュータディスプレイの表面で書くのに用いられるスタイラスとすることができる。任意の相互作用を、文書上のエンコードされた画像の知識を有するか、またはコンピュータスクリーンに表示される文書をサポートするシステムに戻すことができる。ペンまたはスタイラスが文書を横切る際にペンまたはスタイラス内のカメラを用いて画像を繰り返し取り込むことによって、システムは、ユーザが制御するスタイラスの移動を追跡することができる。表示された画像または印刷された画像を、空白の紙またはコンテンツ豊富な紙に関連する透かしとすることができ、または、スクリーンに重ねられる、もしくはスクリーンに組み込まれている、表示される画像もしくは固定コーディングに関連する透かしとすることができる。

30

#### 【0025】

図2Aおよび2Bに、カメラ203を有するペン201の例を示す。ペン201に、インクつばを含んでも含まなくてもよいペン先202が含まれる。カメラ203は、表面207からの画像204を取り込む。ペン201には、さらに、破線の箱206で表される追加のセンサおよび／またはプロセッサを含めることができる。これらのセンサおよび／またはプロセッサ206に、別のペン201および／またはパーソナルコンピュータに情報を（例えば、Bluetoothまたは他の無線プロトコルを介して）送信する能力を含めることもできる。

40

#### 【0026】

図2Bは、カメラ203が捉える画像を表す。一例では、カメラ203の視野（すなわち、カメラの画像センサの解像度）が、 $32 \times 32$ ピクセルである（ $N = 32$ である）。この実施形態では、取り込まれる画像（ $32$ ピクセル $\times 32$ ピクセル）が、カメラ203によって取り込まれる表面平面の約 $5\text{mm} \times 5\text{mm}$ の区域に対応する。したがって、図2Bには、長さ $32$ ピクセル $\times$ 幅 $32$ ピクセルの視野が示されている。Nのサイズは、調整

50

可能であり、より大きいNは、より高い画像解像度に対応する。また、カメラ203の視野が、例示のためにこの図では正方形として示されているが、当技術分野で既知のように、視野に他の形状を含むことができる。

#### 【0027】

カメラ203によって取り込まれる画像は、画像フレームの列  $\{I_i\}$  として定義することができ、この  $I_i$  は、ペン201によって、サンプリング時刻  $t_i$  に取り込まれる。サンプリングレートは、システムコンフィギュレーションおよび性能要件に応じて、これより小さくまたはこれより大きくすることができる。取り込まれる画像フレームのサイズは、システムコンフィギュレーションおよび性能要件に応じて、これより小さくまたはこれより大きくすることができる。

10

#### 【0028】

カメラ203によって取り込まれる画像は、処理システムによって直接に使用することができ、あるいは、事前フィルタリングを受けることができる。この事前フィルタリングは、ペン201内で行うことができ、あるいは、ペン201の外部（例えばパーソナルコンピュータ内）で行うことができる。

#### 【0029】

図2Bの画像サイズは、 $32 \times 32$  ピクセルである。各コーディングユニットサイズが  $3 \times 3$  ピクセルである場合に、取り込まれるエンコードされたユニットの数は、約100ユニットになる。エンコーディングユニットサイズが  $5 \times 5$  ピクセルの場合には、取り込まれるエンコードされたユニットの数は、約36になる。

20

#### 【0030】

図2Aには、位置204からのパターンの画像210が形成される像平面209も示す。物体平面207のパターンから受け取られる光が、レンズ208によって集光される。レンズ208は、単一のレンズまたは複数部分のレンズ系とすることができるが、ここでは、図を単純にするために単一レンズとして表されている。画像取り込みセンサ211が、画像210を取り込む。

#### 【0031】

画像センサ211は、画像210を取り込むのに十分な大きさとするすることができる。またはそれに替えて、画像センサ211を、位置212においてペン先202の画像を取り込むのに十分な大きさとするすることができる。参照のため、位置212での画像を、仮想ペン先と呼ぶ。画像センサ211に関する仮想ペン先位置は、ペン先、レンズ208、および画像センサ211の間の一定の関係のゆえに、固定されていることに留意されたい。

30

#### 【0032】

次の変換  $F_{S \rightarrow P}$  によって、カメラによって取り込まれた画像内の位置座標が、紙の上の実際の画像の位置座標に変換される。

#### 【0033】

$$L_{paper} = F_{S \rightarrow P} (L_{sensor})$$

書込中に、ペン先と紙は、同一平面にある。したがって、仮想ペン先から実際のペン先への変換も、 $F_{S \rightarrow P}$  になる。

#### 【0034】

$$L_{penti p} = F_{S \rightarrow P} (L_{virtual-penti p})$$

変換  $F_{S \rightarrow P}$  を、アフィン変換として推定することができる。これは、 $F_{S \rightarrow P}$  の推定値として、

#### 【0035】

40

【数 1】

$$F'_{S \rightarrow P} = \begin{Bmatrix} \frac{s_x \sin \theta_y}{\cos \theta_x \sin \theta_y - \cos \theta_y \sin \theta_x}, & -\frac{s_x \cos \theta_y}{\cos \theta_x \sin \theta_y - \cos \theta_y \sin \theta_x}, & 0 \\ \frac{s_y \sin \theta_x}{\cos \theta_x \sin \theta_y - \cos \theta_y \sin \theta_x}, & \frac{s_y \cos \theta_x}{\cos \theta_x \sin \theta_y - \cos \theta_y \sin \theta_x}, & 0 \\ 0, & 0, & 1 \end{Bmatrix}$$

【0036】

10

として単純化される。ここで、 $\theta_x$ 、 $\theta_y$ 、 $s_x$ 、および $s_y$ は、位置204で取り込まれるパターンの2つの方位の回転およびスケールである。さらに、取り込まれた画像を紙の上の対応する実際の画像と突き合わせることによって、 $F'_{S \rightarrow P}$ を洗練することができる。「洗練」とは、再帰法と称する種類の最適化アルゴリズムによって、変換 $F_{S \rightarrow P}$ のより正確な推定値を得ることを意味する。再帰法では、行列 $F'_{S \rightarrow P}$ を初期値として扱う。洗練された推定によって、SとPの間の変換がより正確に記述される。

【0037】

次に、較正によって仮想ペン先の位置を判定することができる。

【0038】

ペン先202を、紙の上の固定位置 $L_{pentip}$ に置く。次に、ペンを傾け、カメラ203が異なるペン姿勢で一連の画像を取り込めるようにする。取り込まれる画像ごとに、変換 $F_{S \rightarrow P}$ を得ることができる。この変換から、仮想ペン先の位置 $L_{virtual-pentip}$ を得ることができる。

【0039】

$$L_{virtual-pentip} = F_{P \rightarrow S} (L_{pentip})$$

ここで、 $L_{pentip}$ は、 $(0, 0)$ として初期化され、

$$F_{P \rightarrow S} = (F_{S \rightarrow P})^{-1}$$

である。

【0040】

各画像から得られた $L_{virtual-pentip}$ の平均をとることによって、仮想ペン先の位置 $L_{virtual-pentip}$ を判定することができる。 $L_{virtual-pentip}$ を用いて、 $L_{pentip}$ のより正確な推定値を得ることができる。数回の反復の後に、仮想ペン先の正確な位置 $L_{virtual-pentip}$ を判定することができる。

【0041】

ここで、仮想ペン先の位置 $L_{virtual-pentip}$ がわかっている。取り込まれた画像から、変換 $F_{S \rightarrow P}$ も得ることができる。最後に、この情報を使用して、実際のペン先の位置 $L_{pentip}$ を判定することができる。

【0042】

$$L_{pentip} = F_{S \rightarrow P} (L_{virtual-pentip})$$

(I V. 配列のエンコード)

40

1次元列を折り畳むことによって、2次元配列を構築することができる。十分に多数のビットを含む2次元配列の任意の一部を使用して、完全な2次元配列でのその位置を判定することができる。しかし、1つまたは少数の取り込まれた画像から位置を判定する必要がある場合がある。取り込まれた画像が2次元配列内の複数の位置に関連する可能性を最小にするために、非反復的な列を使用して、配列を作成することができる。作成される列の特性の1つが、列がある長さ（またはウィンドウ） $n$ にわたって繰り返されないことである。次に、1次元列の作成と、配列への列の折り畳みを説明する。

【0043】

(I V. A. 列構築)

50

エンコーディングシステムの出発点として、数列を使用することができる。例えば、列（ $m$ —列とも称する）を、体  $F_q$  の  $q$  個の要素の集合として表すことができる。ここで、 $q = p^n$  であり、 $n \geq 1$ 、 $p$  は素数である。列または  $m$ —列は、多項式除算を含むがこれに限定されない様々な異なる技法によって生成することができる。多項式除算を使用して、列を次のように定義することができる

【0044】

【数2】

$$\frac{R_1(x)}{P_n(x)}$$

10

【0045】

ここで、 $P_n(x)$  は、体  $F_q[x]$  ( $q^n$  個の要素を有する) の次数  $n$  の原始多項式である。 $R_1(x)$  は、体  $F_q[x]$  の次数  $l$  ( $l < n$ ) の非0多項式である。列は、2つのステップ、すなわち、(体  $F_q$  の要素がもたらされる) 2つの多項式を分割する第1のステップと、剰余に  $x$  をかける第2のステップとからなる反復手順を使用して作成することができる。この計算は、出力が繰返しを始める時に停止する。この処理は、線形フィードバックシフトレジスタを使用して実施することができる(非特許文献1参照)。この環境で、列の循環シフトと多項式  $R_1(x)$  の間の関係が確立され、 $R_1(x)$  を変更することによって、列が循環シフトされるだけであり、すべての循環シフトが、多項式  $R_1(x)$  に対応する。結果の列の特性の1つは、列が、 $q^n - 1$  の周期を有し、1周期内で幅(または長さ)  $n$  にわたって、すべての部分が、列の中に1回だけ存在することである。これを、「ウィンドウプロパティ(window property)」と称する。周期  $q^n - 1$  を列の長さとも称し、 $n$  を列の次数とも称する。

20

【0046】

上述した処理は、ウィンドウプロパティを有する列の作成に使用できる様々な処理の1つにすぎない。

【0047】

(IV. B. 配列構築)

(その一部をカメラによって取り込むことができる) 画像の作成に使用できる配列(または  $m$ —配列)は、1次元の列または  $m$ —列の拡張である。 $A$  が、周期  $(m_1, m_2)$  の配列である、すなわち、 $A(k + m_1, l) = A(k, l + m_2) = A(k, l)$  であるものとする。 $n_1 \times n_2$  ウィンドウが、 $A$  の周期の中でシフトされる時に、体  $F_q$  のすべての非0  $n_1 \times n_2$  行列が、1回だけ現れる。この特性を、各ウィンドウが一意であるという点で、「ウィンドウプロパティ」とも称する。ウィンドウを、周期  $(m_1, m_2)$  ( $m_1$  および  $m_2$  は、配列内に存在するビットの水平の個数および垂直の個数である) および次数  $(n_1, n_2)$  の配列として表すことができる。

30

【0048】

2進配列(または  $m$ —配列)は、この列を折り畳むことによって構築することができる。1つの手法は、列を入手し、配列の長さが  $L = m_1 \times m_2 = 2^n - 1$  である場合に、 $m_1 \times m_2$  のサイズに折り畳むことである。代替案では、所定のサイズの含めたい空間(例えば1枚の紙、30枚の紙、またはコンピュータモニタのサイズ)から始め、区域  $(m_1 \times m_2)$  を判定し、そのサイズを使用して、 $L \geq m_1 \times m_2$  (ただし  $L = 2^n - 1$ ) にすることである。

40

【0049】

様々な異なる折り畳み技法を使用することができる。例えば、図3Aから3Cに、3つの異なる列を示す。これらのそれぞれを、図3Dに示された配列に折り畳むことができる。3つの異なる折り畳み方法は、図3Dでオーバーレイとして示され、図3Eおよび3Fでラスタパスとして示されている。ここでは、図3Dに示された折り畳み方法を使用する。

【0050】

50

図 3 D に示された折り畳み方法を作成するために、長さ  $L$ 、次数  $n$  の列  $\{a_i\}$  を作成する。次に、サイズ  $m_1 \times m_2$  の配列  $\{b_{k,l}\}$  (ただし、 $\gcd(m_1, m_2) = 1$  かつ  $L = m_1 \times m_2$ ) を、配列の各ビットを

$$b_{k,l} = a_i, \text{ ただし } k = i \bmod(m_1), l = i \bmod(m_2), i = 0, \dots, L-1 \quad (1)$$

によって計算することによって、列  $\{a_i\}$  から作成する。

#### 【0051】

この折り畳み方法は、その代わりに、配列の対角線に列を置き、端に達した時に反対の端から継続するように表現されることができる。

10

#### 【0052】

図 4 A に、図 3 D の配列をエンコードするのに使用できるサンプルのエンコード技法を示す。他のエンコーディング技法を使用できることを理解されたい。例えば、代替コーディング技法を、図 11 に示す。

#### 【0053】

図 4 A を参照すると、第 1 ビット 401 (例えば「1」) が、暗いインクの列によって表される。第 2 ビット 402 (例えば「0」) が、暗いインクの行によって表される。任意の色のインクを使用して、様々なビットを表すことができることを理解されたい。選択されるインクの色における唯一の要件は、画像取り込みシステムによって区別可能になるように、媒体の背景に対して大きなコントラストをもたらすことである。図 4 A のビットは、セルの  $3 \times 3$  行列によって表される。行列のサイズは、画像取り込みシステムのサイズおよび解像度に基づく任意のサイズに変更することができる。ビット 0 および 1 の代替表現を、図 4 C から 4 E に示す。図 4 A から 4 E のサンプルコーディングの 1 または 0 の表現を、影響なしに切り替えられることを理解されたい。図 4 C に、インターリーブ配置で 2 つの行または列を占めるビット表現を示す。図 4 D に、破線の形で行または列を占めるピクセルの代替の配置を示す。最後に、図 4 E に、不規則な間隔フォーマット (例えば、2 つの暗いドットと 1 つの空白ドット) で行および列を占めるピクセル表現を示す。

20

#### 【0054】

図 4 A に戻ると、あるビットが、 $3 \times 3$  行列によって表され、画像処理システムが、 $3 \times 3$  領域内で 1 つの暗い行と 2 つの白い行を検出する場合に、0 (または 1) が検出される。1 つの暗い列と 2 つの白い列を有する画像が検出される場合に、1 (または 0) が検出される。

30

#### 【0055】

ここでは、複数のピクセルまたはドットが、1 つのビットを表すのに使用される。単一のピクセル (またはドット) を使用して 1 つのビットを表す方法は、脆弱である。ほこり、紙の折れ目、平らでない表面などによって、データ単位の単一ビット表現の読み取りが難しくなる。しかし、異なる手法を用いて、表面で配列をグラフィカルに表現できることを理解されたい。いくつかの手法を、図 4 C から 4 E に示す。他の手法も使用できることを理解されたい。空間的にシフトされたドットだけを使用する 1 つの手法を、図 11 に示す。

40

#### 【0056】

ビットストリームを使用して、図 4 B のグラフィカルパターン 403 を作成する。グラフィカルパターン 403 には、12 個の行と 18 個の列が含まれる。行および列は、ビット表現 401 および 402 を使用してグラフィカル表現に変換されたビットストリームによって形成される。図 4 B は、次のビット表現を有するものと見ることができる。

#### 【0057】

【数 3】

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

【0058】

(V. デコード)

ある人が、図2Aのペンを用いて書くか、エンコードされたパターンの近くにペンを移動する時に、カメラが、画像を取り込む。例えば、ペン201が、紙に押し付けられる時および紙の上の文書を横切る際に、圧力センサを使用することができる。その画像を処理して、エンコードされた画像の完全な表現に関して取り込まれた画像の方位を判定し、取り込まれた画像を構成するビットを抽出する。 10

【0059】

取り込まれた画像の、エンコードされた領域全体に関する方位を判定するために、図5Aから5Dに示された4つの知覚可能な角のすべてがグラフィカルパターン403に存在することができないことがわかる。実際に、正しい方位であれば、図5Aに示された型の角は、グラフィカルパターン403に存在し得ない。したがって、図5Aに示された型の角がない方位が、正しい方位である。 20

【0060】

図6に移ると、カメラ203によって取り込まれた画像を分析して、その方位を判定することにより、その取り込まれた画像は、画像601によって実際に表される位置に関して解釈可能にすることができる。まず、画像601を再検討して、ピクセルが水平および垂直に整列するように画像を回転するのに必要な角度 $\theta$ を判定する。水平および垂直以外の配置への基礎になるグリッドの回転（例えば45°）を含む代替のグリッド配置が可能であることに留意されたい。ユーザは、まず垂直パターンおよび水平パターンに気付く傾向があるので、水平および垂直以外の配置を使用することにより、視覚的にユーザの注意をそらすものを排除する可能性のある利点を提供することができる。説明を単純にするために、グリッドの方位（基礎になるグリッドの水平、垂直、および他の回転）を、まとめて予め定められたグリッド方位と称する。 30

【0061】

次に、画像601を分析して、どの角が欠けているかを判定する。画像601を、デコードの準備ができた画像603に回転するのに必要な回転量 $\alpha$ は、 $\alpha = (\theta + \{\text{どの角が欠けているかによって定められる}\} \text{回転量})$ として示される。回転量は、図7の式によって示される。図6に戻ると、角度 $\theta$ は、まず、ピクセルの水平および垂直（または他の事前定義のグリッド方位）配置に達するピクセルのレイアウトによって決定され、画像が、602に示されているように回転される。次に、分析を行って、欠けている角を判定し、画像602を画像603に回転して、デコードのために画像をセットアップする。ここで、画像は、反時計回りに90°回転され、その結果、画像603が、正しい方位を有するようになり、デコードに使用できるようになる。 40

【0062】

回転角度 $\theta$ を、欠けている角を考慮するために画像601を回転する前または後に適用できることを理解されたい。また、取り込まれた画像のノイズを考慮することによって、4つのすべての型の角が存在し得ることを理解されたい。各型の角の数を数え、最も少ない数を有する型を、欠けている角の型として選択することができる。

【0063】

最後に、画像603内のコードを読み出し、画像403の作成に使用された最初のビットストリームと相関させる。この相関は、複数の形で実行することができる。例えば、この相関は、回復されたビットストリームを、最初のビットストリーム内の他のすべてのビ 50

ットストリームの断片と比較する、再帰的手法によって実行することができる。第2に、回復されたビットストリームと最初のビットストリームとの間で、例えば2つのビットストリーム間のハミング距離を使用することによって、統計分析を実行することができる。様々な手法を用いて、元のビットストリーム内の回復されたビットストリームの位置を判定できることを理解されたい。

#### 【0064】

回復されたビットを得たならば、最初の配列（例えば、図4Bに示されたもの）で取り込まれた画像の位置を突き止める必要がある。配列全体の中でのビットのセグメントの位置を判定する処理は、複数の項目によって複雑になる。第1に、取り込まれた実際のビットが、隠される場合がある（例えば、カメラが、元のコードを隠す手書きの画像を取り込む場合がある）。第2に、ごみ、折れ目、反射などによって、取り込まれる画像のエラーが作られる場合もある。これらのエラーによって、位置突き止めプロセスが難しくなる。これに関して、画像取り込みシステムが、画像から抽出される非シーケンシャルビットを用いて機能する必要がある場合がある。次に、画像からの非シーケンシャルビットを用いて動作する方法を示す。

#### 【0065】

列（または $m$ —列） $I$ が、べき級数  $I(x) = 1/P_n(x)$  に対応し、 $n$ が、 $m$ —列の次数であり、取り込まれた画像に、 $I$ の $K$ 個のビット  $b = (b_0 \ b_1 \ b_2 \ \dots \ b_{K-1})^t$  が含まれ、 $K \geq n$ であり、添字 $t$ が、行列またはベクトルの転置を表すものとする。 $K$ 個のビットの位置 $s$ は、 $b_0$ が列の先頭にシフトされるようにする、 $I$ の循環シフトの回数である。そのときに、このシフトされた列 $R$ は、べき級数  $x^s/P_n(x)$ 、または  $R = T^s(I)$  に対応し、 $T$ は、循環シフト演算子である。この $s$ は、間接的に見つかる。 $P_n(x)$ を法とする多項式は、体を形成する。 $x^s \equiv r_0 + r_1 x + \dots + r_{n-1} x^{n-1} \pmod{P_n(x)}$  であることが保証される。したがって、 $(r_0, r_1, \dots, r_{n-1})$ を見つけることができ、 $s$ の解を見つけることができる。

#### 【0066】

関係  $x^s \equiv r_0 + r_1 x + \dots + r_{n-1} x^{n-1} \pmod{P_n(x)}$  は、 $R = r_0 + r_1 T(I) + \dots + r_{n-1} T^{n-1}(I)$  であることを意味する。これを2進一次方程式で書くと、

$$R = r' A \quad (2) \quad 30$$

になる。ここで、 $r = (r_0 \ r_1 \ r_2 \ \dots \ r_{n-1})^t$ であり、 $A = (I \ T(I) \ \dots \ T^{n-1}(I))^t$ であり、これは0シフトから $(n-1)$ シフトまでの $I$ の循環シフトからなる。 $r$ を解くために、疎な $K$ ビットだけが、 $R$ で使用可能である。 $R$ の $b_i$ と $b_0$ のインデックスの差が、 $k_i$  ( $i = 1, 2, \dots, k-1$ )であるものとする。と、 $R$ の1番目と $(k_i + 1)$ 番目 ( $i = 1, 2, \dots, k-1$ )の要素は、正確に $b_0, b_1, \dots, b_{k-1}$ である。 $A$ の1番目と $(k_i + 1)$ 番目 ( $i = 1, 2, \dots, k-1$ )の要素を選択することによって、次の2進一次方程式が形成される。

#### 【0067】

$$b' = r' M \quad (3) \quad 40$$

ここで、 $M$ は、 $A$ の $n \times K$ 部分行列である。

#### 【0068】

$b$ にエラーがない場合には、 $r$ の解を

#### 【0069】

#### 【数4】

$$r' = \tilde{b}' \tilde{M}^{-1} \quad (4)$$

#### 【0070】

と表すことができる。ここで

#### 【0071】

【数 5】

 $\tilde{M}$ 

【0072】

は、 $\tilde{M}$  の非縮退  $n \times n$  部分行列であり、

【0073】

【数 6】

 $\tilde{b}$ 

【0074】

は、 $\tilde{b}$  の対応する部分ベクトルである。

【0075】

$r$  がわかっているならば、Pohlig-Hellman-Silver アルゴリズム（非特許文献 1 参照）を使用して、 $x^s \equiv r_0 + r_1 x + \dots + r_{n-1} x^{n-1} \pmod{P_n(x)}$  になる  $s$  を見つけることができる。

【0076】

行列  $A$  ( $n \times L$  のサイズであり、 $L = 2^n - 1$ ) が巨大になる場合があるので、行列  $A$  全体を保管することは避けなければならない。実際に、上のプロセスで見たように、インデックス差  $k_i$  を有する抽出されたビットを与えられて、 $A$  の 1 番目および  $(k_i + 1)$  番目の列だけが、計算に関係する。そのような  $k_i$  の選択は、取り込まれた画像のサイズに対して非常に限られている。したがって、計算に使用される列だけを保存する必要がある。そのような列の総数は、 $(L = 2^n - 1$  が  $m$  列の長さである場合に)  $L$  よりはるかに小さい。

【0077】

(V I. エラー訂正)

$\tilde{b}$  にエラーが存在する場合、 $r$  の解決は複雑になる。エラー訂正を伴うデコードの従来の方法は、簡単に適用できない可能性がある。というのは、取り込まれたビットに関連する行列  $M$  が、取り込まれた画像ごとに異なる可能性があるからである。

【0078】

本発明では、統計学的手法を使用する。 $\tilde{b}$  のエラービットの数  $n_e$  が、 $K$  と比較して相対的に小さいと仮定すると、 $\tilde{b}$  および対応する、縮退していない  $M$  の部分行列

【0079】

【数 7】

 $\tilde{M}$ 

【0080】

の  $K$  個のビットから正しい  $n$  個のビットが選ばれる確率は、高い。

【0081】

選択された  $n$  個のビットがすべて正しい時に、 $\tilde{b}'$  と  $r' M$  との間のハミング距離または  $r$  に関連するエラービットの数は、最小値にならないとせず、 $r$  は、式 (4) を介して計算される。この処理を複数回繰り返すことによって、最小値のエラービットをもたらす正しい  $r$  を識別できる可能性は高い。

【0082】

最小個数のエラービットに関連する唯一の  $r$  がある場合には、それが正しい解とみなされる。そうではなく、最小個数のエラービットに関連する複数の  $r$  がある場合には、 $n_e$  が、 $M$  によって生成されるコードのエラー訂正能力を超える可能性が高く、デコードプロセスはエラーになる。システムは、次に取り込まれた画像の処理に進むことができる。もう 1 つの実施形態では、ペンの前の位置に関する情報を考慮に入れることができる。すなわち、取り込まれた画像の各々について、次にペンがあると期待できる宛先領域を識別することができる。例えば、ユーザが、カメラによって取り込まれる 2 つの画像の間でペン

10

20

30

40

50

を持ち上げなかった場合に、第2の画像の取り込みによって判定されるペンの位置は、第1の位置から遠過ぎてはならない。最小個数のエラービットに関連する各  $r$  を検査して、 $r$  から計算される位置  $s$  が、局地的な制約を満足するか、すなわち、位置が指定された宛先領域内であるか否かを調べることができる。

#### 【0083】

位置  $s$  がローカル制約を満足する場合、配列内の抽出されたビットの  $X$ 、 $Y$  位置が返される。そうでない場合には、デコード処理はエラーになる。

#### 【0084】

図8に、取り込まれた画像の列（または  $m$  列）内の位置を判定するのに使用できる処理を示す。まず、ステップ801で、取り込まれた画像に関するデータストリームを受け取る。ステップ802で、対応する列を  $A$  から抽出し、行列  $M$  を構築する。 10

#### 【0085】

ステップ803で、 $n$  個の独立の列ベクトルを、行列  $M$  からランダムに選択し、式(4)を解くことによってベクトル  $r$  を判定する。この処理を、ステップ804で  $Q$  回（例えば100回）繰り返す。ループ回数の判定は、セクション「ループ回数の計算」で説明する。

#### 【0086】

ステップ805で、 $r$  を、それに関連するエラービットの個数に従ってソートする。ソートは、当技術分野で既知の様々なソーティングアルゴリズムを用いて行うことができる。例えば、選択ソートアルゴリズムを用いることができる。選択ソートアルゴリズムは、 20  
数  $Q$  が多くない時に有益である。しかし、 $Q$  が多い場合には、多数の項目をより効率的にソートできる他のソーティングアルゴリズム（例えばマージソート）を使用することができる。

#### 【0087】

システムは、ステップ806で、複数の  $r$  が最小個数のエラービットに関連するか否かを検査することによって、エラー訂正が成功裡に実行されたか否かを判定する。そうである場合に、ステップ809でエラーを返し、デコード処理が失敗したことを示す。そうでない場合には、ステップ807で、例えば  $Pohlig-Hellman-Silver$  アルゴリズムを用いて、列（または  $m$  列）の抽出されたビットの位置  $s$  を計算する。

#### 【0088】

次に、ステップ808で、配列内の  $(X, Y)$  位置を、 $x = s \bmod m_1$  および  $y = s \bmod m_2$  として計算し、その結果を返す。 30

#### 【0089】

##### (VII. 位置判定)

図9に、ペン先の位置を判定する処理を示す。入力は、カメラによって取り込まれた画像であり、出力は、ペン先の位置座標とすることができる。また、出力に、取り込まれた画像の回転角度などの他の情報を含めることができる（含めないこともできる）。

#### 【0090】

ステップ901で、カメラから画像を受け取る。次に、（ステップ902の破線の輪郭によって示されるように）ステップ902で、任意に受け取った画像を前処理して、明るいピクセルと暗いピクセルとの間のコントラストの調節などを行うことができる。 40

#### 【0091】

次に、ステップ903で、画像を分析して、その中のビットストリームを判定する。

#### 【0092】

次に、ステップ904で、ビットストリームから  $n$  個のビットをランダムに複数回選択し、元の列（または  $m$  列）内の受け取ったビットストリームの位置を判定する。

#### 【0093】

最後に、取り込まれた画像の位置をステップ904で判定したならば、ステップ905で、ペン先の位置を判定することができる。

#### 【0094】

図10に、903および904に関するさらなる詳細を示し、取り込まれた画像内のビットストリームを抽出する手法を示す。まず、ステップ1001で、カメラから画像を受け取る。次に、(ステップ1002の破線によって示されるように)ステップ1002で、任意に画像に画像前処理を行うことができる。ステップ1003で、パターンを抽出する。ここで、様々な線のピクセルを抽出して、パターンの方位および角度 $\theta$ を見つけることができる。

#### 【0095】

次に、ステップ1004で、受け取った画像を分析して、基礎になるグリッド線を判定する。ステップ1005で、グリッド線が見つかる場合に、ステップ1006でパターンからコードを抽出する。その後、ステップ1007でコードをデコードし、ステップ1008でペン先の位置を判定する。ステップ1005でグリッド線が見つからなかった場合には、ステップ1009でエラーを返す。

10

#### 【0096】

(V I I I . 透かしを入れられた文書画像：生成、分析、およびラベル付け)

本発明の実施形態は、迷路パターンなどの位置エンコードされた情報を含む文書画像の分析と、画像のラベル付けに関する。そのような文書分析の結果を、位置エンコードされた文書内のカメラによって取り込まれた画像の位置の効率的な判定に用いることができる。

#### 【0097】

図11を参照すると、本発明の様々な実施形態によるシステム1100に、画像生成取り込みモジュール1102と分析ラベル付けモジュール1106が含まれる。後続の分析およびラベル付けのための画像の取り込みに適する技法は、上記の画像取り込みペンという題名のセクションI I I、セクションV I I I、および透かしを入れられたグレイスケール文書画像の生成という題名のセクションV I I I . Aで説明する。分析ラベル付けモジュール1106は、画像生成取り込みモジュール1102によって出力される画像データ1104およびオフライントレーニングデータ1110を入力として受け取り、下で詳細に説明する分析ラベル付け処理を実行し、画像ラベル情報1108を出力する。

20

#### 【0098】

本発明の様々な実施形態によれば、カメラ203が取り込む画像のサイズは、 $32 \times 32$ ピクセルである。組み込まれた $m$ -配列の順序(表面207で位置を位置エンコードするのに使用される)は、36である、すなわち、 $m$ -配列のサイズは、 $(2^{18} + 1) \times (2^{18} - 1)$ である。したがって、サブブロック $m$ -配列から一意の $x-y$ 位置をデコードするために、表面207での位置を判定するのに使用されるサブブロックのビット数は、少なくとも36でなければならない。

30

#### 【0099】

本発明の様々な実施形態によれば、印刷された文書が、迷路パターンによって透かしを入れられる。そのような迷路パターンは、テキストなどの文書の内容によって覆われる場合がある。そのような覆う程度に応じて、カメラによって取り込まれた画像の文書内の位置を判定するのに十分な $m$ -配列ビットを、カメラによって取り込まれた画像から抽出することが不可能になる場合がある。

40

#### 【0100】

本発明の様々な実施形態によれば、透かしを入れられた文書画像が、分析され、ラベル付けされる。分析およびラベル付けは、文書内の特定の位置の迷路パターンセルの可視性の程度に基づく、透かしを入れられた文書内の $x-y$ 位置を判定する能力に関する。

#### 【0101】

(V I I I . A . 透かしを入れられたグレイスケール文書画像の生成)

図12に、本発明の様々な実施形態に従って分析し、ラベル付けを行うことができる、透かしを入れられたグレイスケール文書画像を生成するステップを示す。1200および1204に示されているように、電子文書を印刷された文書に対応するビットマップとするか、1202および1206に示されているように紙の文書を読み取ることによるなど

50

、文書画像を入手する。1208に示されているように、ビットマップを再サンプリングして、適当な解像度にする。解像度の決定は、(1) 取り込まれた画像の位置をm-配列デコードによって判定できない場合、位置を判定するために、取り込まれた画像が透かしを入れたグレイスケール文書画像とマッチングされるので、文書画像の解像度が取り込まれた画像の解像度より低くならないという考慮事項に基づく。例えば、カメラの解像度が0.15mm/ピクセルである、すなわち、物理的世界での0.15mmが1つのカメラピクセルにマッピングされる場合に、印刷される迷路パターンセルサイズは、0.45mm×0.45mmである、すなわち、印刷される迷路パターンのセルが、カメラセンサの3×3ピクセルにマッピングされ、文書画像の解像度にも、0.15mm/ピクセルがセットされなければならない、その結果、印刷される迷路パターンセルが、文書画像の3×3ピクセル領域にマッピングされるようになる。

10

20

30

40

50

#### 【0102】

得られた画像に、透かしを入れる。図13に、文書の内容の一部1300、位置エンコードされた迷路パターンの一部1302、組み合わせられた内容および迷路パターン1304、ならびに内容および迷路パターンの一部の拡大図1306の画像を示す。拡大図1306には、(1) 重ねられたテキストによって覆われるので迷路パターンセルが可視でない位置1308と、(2) 迷路パターンセルが文書内容によって覆われず、したがって可視である位置1310が示されている。

#### 【0103】

(V I I I . B . 文書画像の分析およびラベル付け)

本発明の様々な実施形態によれば、文書画像は、その画像が、迷路パターンセルと実質的に同一のサイズを有する比較的小さいブロックに分割され；これらの小さいブロックが、文書テキストなどの文書の内容によって覆われるか否かを判定し；ピクセルごとに、ウィンドウの中心としてそのピクセルを有する隣接ウィンドウ内の完全に可視のブロックの数を数え；その数に基づいてピクセルにラベル付けを行うことによって分析される。隣接ウィンドウは、カメラ203によって取り込まれる画像のサイズと実質的に同一のサイズすなわち、32×32ピクセルを有することができる。

#### 【0104】

前述の処理が、グレイスケール文書画像に透かしを入れることを必要としない、すなわち、内容を迷路パターンと組み合わせることを必要としないことに留意されたい。文書画像が、迷路パターンセルと実質的に同一のサイズを有する比較的小さいブロックに分割される限り、文書内容による迷路パターンセルの覆いの分析は、文書内容によるブロックの覆いの分析と同等である。

#### 【0105】

図14に、そのような32×32ピクセル隣接ウィンドウ1400およびその中心ピクセル1402の例を示す。本発明の様々な実施形態によれば、文書画像のピクセルに、次の4つの型の1つとしてラベル付けを行うことができる。

#### 【0106】

型I：(検討中のピクセルを中心とする) 32×32ピクセル隣接ウィンドウに、実質的に位置エンコードされた迷路パターンセルだけが含まれる。

#### 【0107】

型II：32×32ピクセル隣接ウィンドウに、60個以上の可視迷路パターンセルが含まれる。

#### 【0108】

型III：32×32ピクセル隣接ウィンドウに、36個以上60個以下の可視迷路パターンセルが含まれる。

#### 【0109】

型IV：32×32ピクセル隣接ウィンドウに、35個以下の可視迷路パターンセルが

含まれる。

#### 【0110】

本発明の様々な実施形態によれば、取り込まれた画像の中心が、型Ⅰまたは型ⅠⅠの領域に配置される時に、文書内の画像の位置を、m-配列デコードによって一意に決定することができる。

#### 【0111】

図15に、各ピクセルの $32 \times 32$ 隣接ウィンドウ内で可視である迷路パターンセルの量に基づいて文書の諸部分がどの型（すなわち型Ⅰ、ⅠⅠ、ⅠⅠⅠ、またはⅠⅤ）にラベル付けされるかを表すために文書の領域が着色（または陰付け）されている、例の文書の文書分析結果を示す。

10

#### 【0112】

可視迷路パターンセルの数を数える際に、 $32 \times 32$ ピクセル隣接ウィンドウを $360^\circ$ 回転して、カメラがある角度で文書の画像を取り込んだ可能性があるという事実を考慮に入れなければならないことに留意されたい。しかし、実施を簡単にするために、傾いていないビューだけが分析され、傾斜角度によって引き起こされるエラーは、ピクセルの型を判定する際の閾値で考慮に入れられる。例えば、図16に、取り込まれた画像が $45^\circ$ 回転されている場合を示す。これは、可視迷路パターンセルの $17.2\%$ という最大値が失われる、ワーストケースシナリオである。その場合であっても、型ⅠⅠピクセルは、少なくとも $49$ 個（ $= 60 \times (1 - 17.2\%)$ ）の可視迷路パターンセルを有し、これによって、m-配列デコードによる一意のx-y位置判定が可能である。言い換えると、ピクセルが、型ⅠまたはⅠⅠとしてラベル付けされる場合に、カメラがどの角度で画像を取り込んだかに無関係に、m-配列デコードを用いてx-y位置を判定するのに十分な数の可視迷路パターンセルがある。その結果、取り込んだ画像をデコードできない場合には、その画像が型ⅠⅠⅠまたはⅠⅤの領域にある可能性が高い。

20

#### 【0113】

十分な迷路パターンセルが可視ではないので、ペンストロークの画像の列からx-y位置をデコードできない場合には、検索アルゴリズムを使用して、文書内の型ⅠⅠⅠおよびⅠⅤピクセル領域などの画像の位置を判定する。そのような検索アルゴリズムを、型ⅠⅠⅠおよびⅠⅤの領域だけで使用することによって、文書の実質的にすべての部分に関する検索アルゴリズムの使用に対して、画像の位置の判定の計算費用が減る。例えば、図15の例の文書では、領域の $13.7\%$ だけが、型ⅠⅠⅠ（ $11.1\%$ ）または型ⅠⅤ（ $2.6\%$ ）としてラベル付けされる。文書全体を検索することに関連する計算費用は、型ⅠⅠⅠおよびⅠⅤ領域だけの検索の費用よりはるかに高い。一実施形態で、文書全体の検索に10秒超を要するが、型ⅠⅠⅠおよびⅠⅤ領域だけの検索に1秒未満を要する場合がある。文書画像分析によって、計算費用が大幅に減る。

30

#### 【0114】

（ⅠⅩ．カメラによって取り込まれた画像のラベル付け）

カメラによって取り込まれた画像に、文書画像の分析およびラベル付けと題した上のセクションⅤⅠⅠⅠ．B．で説明した4つの型によってラベル付けを行うことができる。カメラによって取り込まれた画像に、実質的に位置エンコードされた迷路パターンだけが含まれるかどうかに関する判定を行うことができる。言い換えると、画像が型Ⅰ画像であるか否かに関する判定を行うことができる。この初期判定を、下で詳細に説明する。

40

#### 【0115】

カメラによって取り込まれた画像が、型Ⅰ画像でない場合に、カメラによって取り込まれた画像の位置エンコードされた迷路パターンを分析して、画像から抽出できる位置エンコーディングビットの数を判定することができる。迷路パターングリッド線のパラメータ（各次元に沿ったスケールおよび回転すなわちアフィン変換）が、まず計算され、その後、迷路パターンが埋め込まれた最初の方角（または四分円）が判定される。その結果、m-配列ビット情報が、グリッド線およびバーの方角に基づいて判定される。抽出されるビット数が約60を超える場合に、画像は型ⅠⅠとしてラベル付けされ、約36ビットと6

50

0ビットの間のビットが抽出される場合に、画像は型ⅠⅠⅠとしてラベル付けされ、約36ビットより少ないビットが抽出される場合に、画像は型ⅠⅤとしてラベル付けされる。カメラによって取り込まれた画像をこの形で分類するために、下の、スレッシュホルディングアルゴリズム (thresholding algorithm) と題するセクションⅠⅩ. B. で説明するスレッシュホルディングアルゴリズムを用いて、テキスト領域から迷路パターン領域を分離することができる。その後、抽出されたビットの数に基づいて、画像にラベル付けを行う。型Ⅰ、ⅠⅠ、およびⅠⅠⅠから抽出されたビットは、カメラによって取り込まれた画像の、より大きい文書内のx-y位置を判定するのに使用することができる。この型の位置判定処理は、型ⅠⅤ画像について省略することができる。というのは、型ⅠⅤ画像に、x-y位置デコードに十分なビットが含まれないからである。

10

#### 【0116】

画像に、文書内容と迷路パターンの両方ではなく、実質的に位置エンコードされた迷路パターンだけが含まれるか否かを判定するために、グラディエント画像ヒストグラムのサポート間隔 (support interval of gradient image histogram、SIGH) と称する特徴を、本発明の様々な実施形態に従って使用することができる。SIGHは、純粋な迷路パターン画像に、通常はエッジが含まれないが、文書内容が、多くの場合、迷路パターンセルまたは空白領域よりはるかに暗いので、テキストなどの文書内容を有する画像に、通常はエッジが含まれるという知識に基づいて使用される。

#### 【0117】

##### (ⅠⅩ. A. 特徴抽出)

本発明の様々な実施形態によれば、特徴抽出の第1ステップは、グラディエント画像 (gradient image) を得るのに使用できる、Sobelエッジ演算子 (Sobel edge operator) または別のグラディエント演算子などのグラディエント演算子である。図17を参照すると、2つの畳み込みカーネル1700および1702が、Sobelエッジ演算子を形成する。Sobelエッジ演算子を使用するために、画像の各ピクセルを、畳み込みカーネル1700および1702の両方を用いて畳み込む。一方のカーネルは、全般に垂直のエッジに最大限応答し、他方のカーネルは、水平のエッジに最大限応答する。この2つの畳み込みの「ベクトル和」値が、特定のピクセルの出力値としてとられる。その結果が、グラディエント画像である。

20

#### 【0118】

グラディエント演算子によって作られたグラディエント画像のヒストグラムを計算することができる。グラディエント画像ヒストグラム (SIGH) のサポート間隔を、グラディエント画像のヒストグラムから得ることができる。例えば、図18に、迷路パターンだけを含む取り込まれた画像のグラディエント画像のヒストグラムを示す。SIGH特徴は、17であり、これは、図18のヒストグラムの、非0の値を有する、x軸に沿った最大値である。

30

#### 【0119】

図19に、テキストおよび／または1つまたは複数の図などの文書内容を含む取り込まれた画像のグラディエント画像のヒストグラムを示す。SIGH特徴は、図19のヒストグラムでは44である。

40

#### 【0120】

##### (ⅠⅩ. B. スレッシュホルディングアルゴリズム)

本発明の様々な実施形態に従って画像が型Ⅰであるか否かを判定するスレッシュホルディングアルゴリズムは、2つのセッションすなわち (1) オフライントレーニング、および (2) オンラインラベル付けに分割することができる。

#### 【0121】

##### (ⅠⅩ. B. 1. オフライントレーニング。)

オフライントレーニングセッションでは、SIGH特徴を、既知のラベル (すなわち型Ⅰ、ⅠⅠ、ⅠⅠⅠ、またはⅠⅤ) を有する比較的多数の画像について計算することができる。純迷路パターン画像のSIGHは、多くの場合、他の型の画像のSIGHより

50

低い。最適閾値  $\eta$  は、トレーニングデータの処理の結果に基づいて選択することができる。トレーニングセット中の実質的にすべての純迷路パターン画像について、その S I G I H は  $\eta$  未満であり、実質的にすべての非純迷路パターン画像について、その S I G I H は  $\eta$  を超える。図 20 に、 $\eta$  すなわち最適閾値 2004 が、32 になるように選択されたオフライントレーニングの例の結果を示す。この図には、純迷路パターン画像に対応するヒストグラムバー 2002 と、非純迷路パターン画像（すなわち、文書内容を含む画像）に対応するヒストグラムバー 2006 も示されている。

#### 【0122】

（IX. B. 2. オンラインラベル付け）

オンラインラベル付け中に、取り込まれた画像のそれぞれの S I G I H が計算される。10  
S I G I H が  $\eta$  より小さい場合に、画像は型 I 画像としてラベル付けされる。そうではなく、S I G I H が  $\eta$  より大きい場合、画像は型 I 以外の型（すなわち、型 I I、I I I、および I V）としてラベル付けされる。

#### 【0123】

上の文書分析およびラベル付けに関する議論では、取り込まれた画像の内容が、比較的一貫性を有することが前提である。言い換えると、画像センサの増幅器および／またはカメラの照明条件が比較的一貫性を有することが前提である。この条件が、大幅に変化する場合に、特徴パラメータを更新するために新しいトレーニングを実行することができる。

#### 【0124】

（X. 最後に）

上で説明したものは、本発明の原理の適用の例にすぎない。当業者は、本発明の趣旨および範囲から逸脱せずに他の配置および方法を実施することができる。本発明の方法のいずれをも、コンピュータディスクまたは他のコンピュータ読み取り可能な媒体に保管できるソフトウェアで実施することができる。

#### 【図面の簡単な説明】

#### 【0125】

【図 1】本発明の実施形態と共に使用することができるコンピュータの全般的な説明を示す図である。

【図 2 A】本発明の実施形態による画像取り込みシステムを示す図である。

【図 2 B】本発明の実施形態による取り込まれた画像を示す図である。

【図 3 A】本発明の実施形態によるシーケンスを示す図である。

【図 3 B】本発明の実施形態によるシーケンスを示す図である。

【図 3 C】本発明の実施形態によるシーケンスを示す図である。

【図 3 D】本発明の実施形態によるシーケンスおよび折り畳み技法を示す図である。

【図 3 E】本発明の実施形態による折り畳み技法を示す図である。

【図 3 F】本発明の実施形態による折り畳み技法を示す図である。

【図 4 A】本発明の実施形態によるエンコーディングシステムを示す図である。

【図 4 B】本発明の実施形態によるエンコーディングシステムを示す図である。

【図 4 C】本発明の実施形態によるエンコーディングシステムを示す図である。

【図 4 D】本発明の実施形態によるエンコーディングシステムを示す図である。

【図 4 E】本発明の実施形態によるエンコーディングシステムを示す図である。

【図 5 A】図 4 A および 4 B によるエンコーディングシステムに関連する 4 つの可能な結果の角の 1 つを示す図である。

【図 5 B】図 4 A および 4 B によるエンコーディングシステムに関連する 4 つの可能な結果の角の 1 つを示す図である。

【図 5 C】図 4 A および 4 B によるエンコーディングシステムに関連する 4 つの可能な結果の角の 1 つを示す図である。

【図 5 D】図 4 A および 4 B によるエンコーディングシステムに関連する 4 つの可能な結果の角の 1 つを示す図である。

【図 6】本発明の実施形態による、取り込まれた画像の一部の回転を示す図である。

【図 7】図 4 A から 4 E のコーディングシステムと共に使用される回転の様々な角度を示す図である。

【図 8】本発明の実施形態による、取り込まれた配列の位置を判定する処理を示す図である。

【図 9】本発明の実施形態による、取り込まれた画像の位置を判定する処理を示す図である。

【図 10】本発明の実施形態による、取り込まれた画像の位置を判定するもう 1 つの処理を示す図である。

【図 11】本発明の様々な実施形態による、画像取り込みモジュールと分析ラベル付けモジュールを示す図である。

【図 12】本発明の様々な実施形態に従って分析し、ラベル付けを行うことができる、透かしを入れられたグレイスケール文書画像を生成するステップを示す図である。

【図 13】文書の内容の一部、位置エンコードされた迷路パターンの一部、組み合わされた内容および迷路パターン、ならびに内容および迷路パターンの一部の拡大図の画像を示す図である。

【図 14】図 13 に示された組み合わされた文書および迷路パターンの、サブウィンドウおよびサブウィンドウの中心のピクセルを示す図である。

【図 15】本発明の様々な実施形態による例の文書の文書分析結果を示す図である。

【図 16】非 0 回転角度および 45° 回転角度を有する、カメラによって取り込まれた画像を示す図である。

【図 17】Sobel エッジ演算子を形成する 2 つの畳み込みカーネルを示す図である。

【図 18】迷路パターンだけを含む取り込まれた画像のグラディエント画像のヒストグラムを示す図である。

【図 19】文書内容を含む取り込まれた画像のグラディエント画像のヒストグラムを示す図である。

【図 20】純迷路パターンの画像と文書内容を含む画像の間の区別のために選択される閾値を含む、オフライントレーニングの例の結果を示す図である。

【図 21】従来技術による文書内のエンコーディングスペースの代表を示す図である。

【符号の説明】

【0126】

- 100 コンピュータ
- 170 ハードディスクドライブ
- 180 磁気ディスクドライブ
- 191 光ディスクドライブ
- 201 ペン
- 202 ペン先
- 203 カメラ
- 204 位置
- 205 位置
- 206 追加のセンサおよび／またはプロセッサ
- 207 物体平面
- 208 レンズ
- 209 像平面
- 210 画像
- 211 画像センサ
- 212 位置
- 1100 本発明の様々な実施形態によるシステム
- 1300 本文の内容の一部
- 1302 迷路パターンの一部
- 1304 組み合わされた内容および迷路パターン

10

20

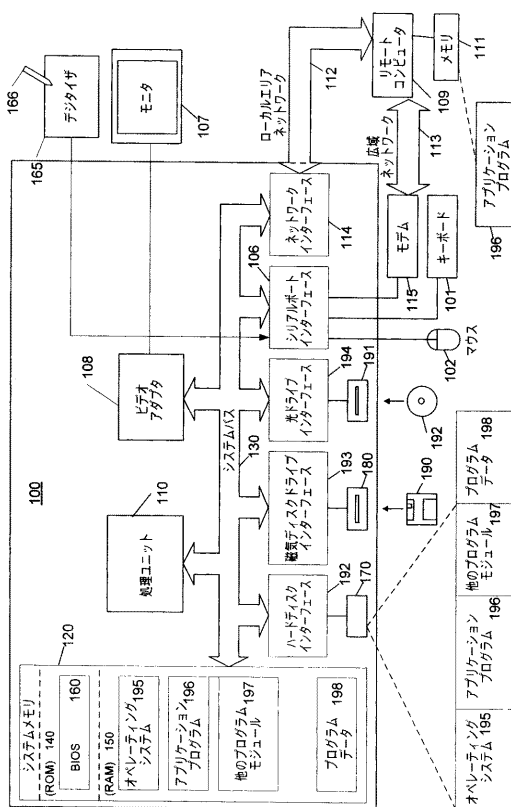
30

40

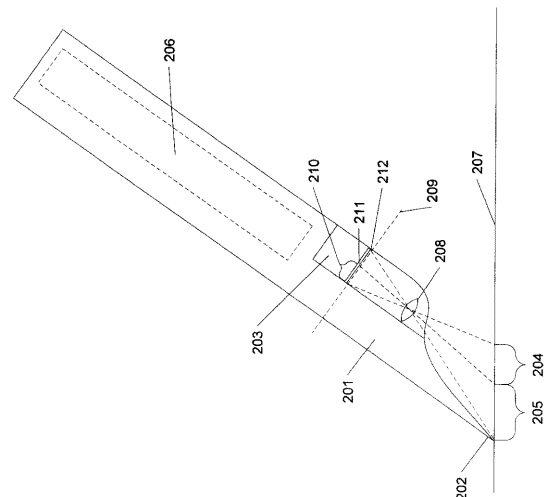
50

- 1 3 0 6 内容および迷路パターンの一部の拡大図  
 1 4 0 0 32×32ピクセル隣接ウィンドウ  
 1 7 0 0 畳み込みカーネル  
 1 7 0 2 畳み込みカーネル

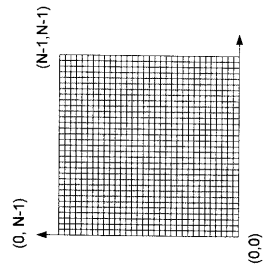
【図 1】



【図 2 A】



【図 2 B】



【図 3 A】

00000010 00 01100 0101 001 111010 0 0111001 0 010110 111 0110 01101 01 0111111

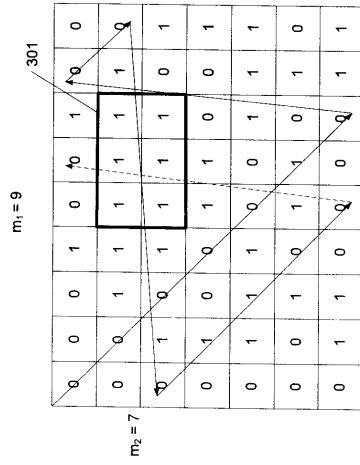
【図 3 B】

00000000 0011101 0100111 1110100 0111010 1000101 1110100 0100111 0011101

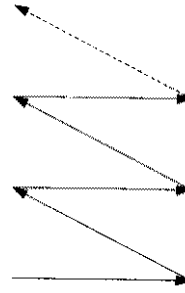
【図 3 C】

000100100 001111110 010111101 010011001 011100111 0010110101010 011000011

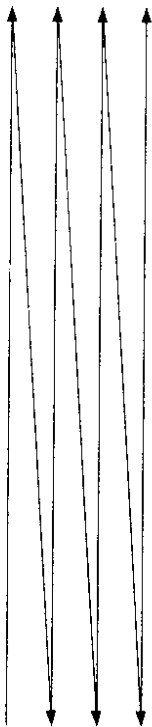
【図 3 D】



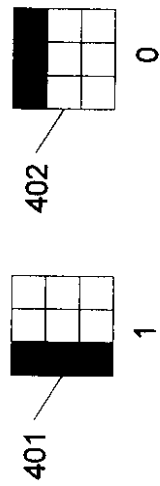
【図 3 E】



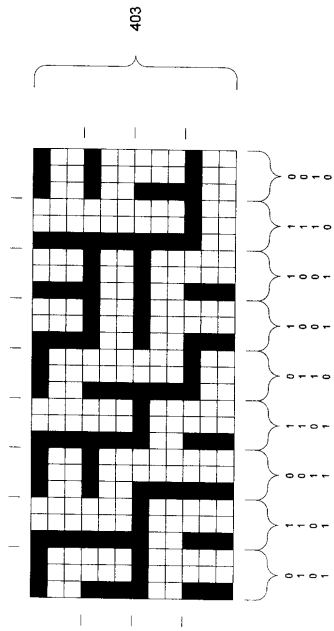
【図 3 F】



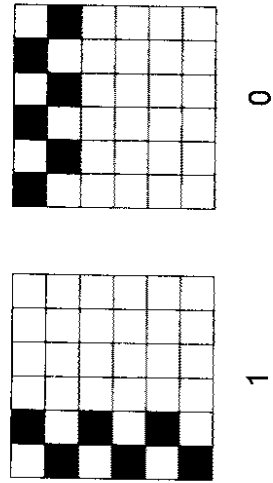
【図 4 A】



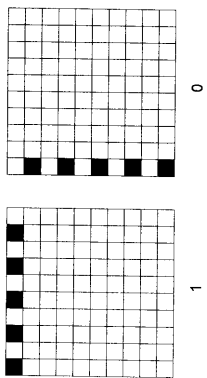
【図 4 B】



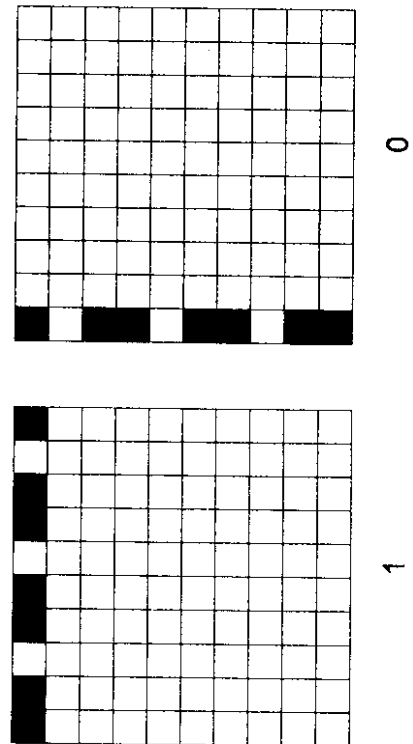
【図 4 C】



【図 4 D】



【図 4 E】



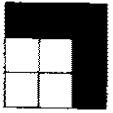
【図 5 A】



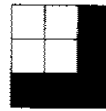
【図 5 B】



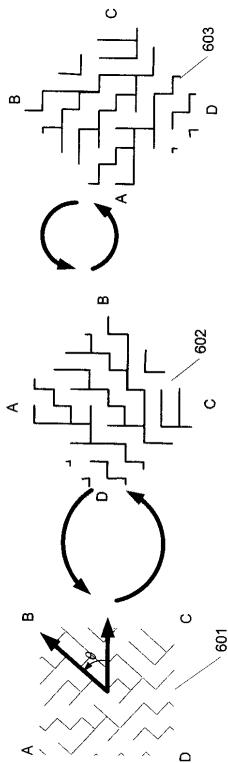
【図 5 C】



【図 5 D】



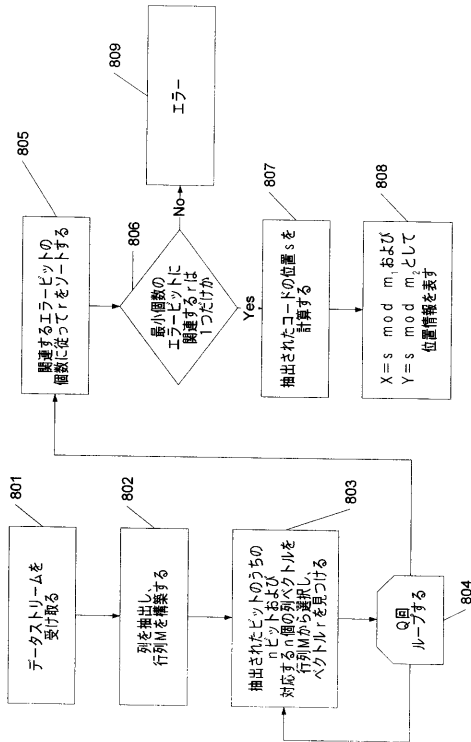
【図 6】



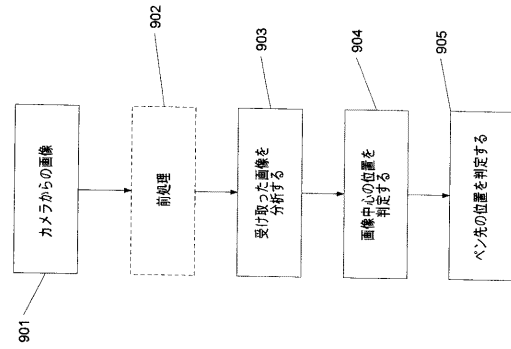
【図 7】

$$o = \theta + \begin{cases} 0 - a \text{ なし} \\ \pi - b \text{ なし} \\ \pi - c \text{ なし} \\ 3\pi - d \text{ なし} \end{cases} \frac{1}{2}$$

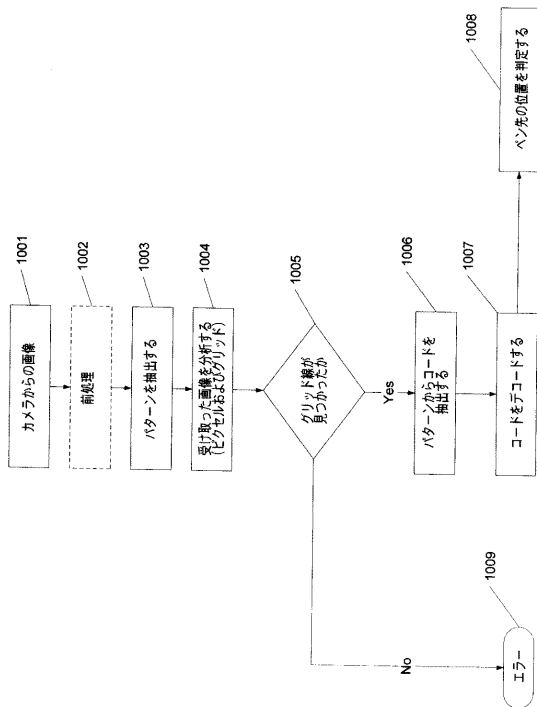
【図 8】



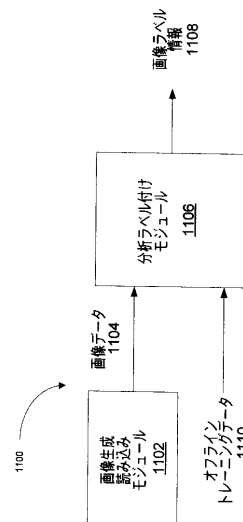
【図 9】



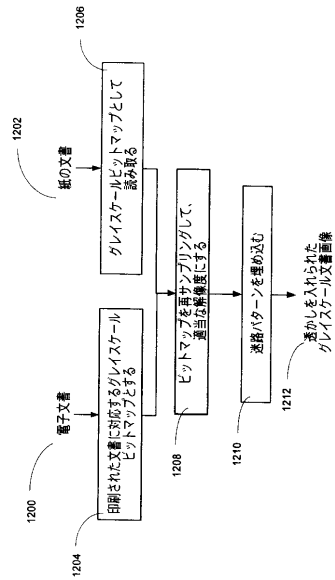
【図 10】



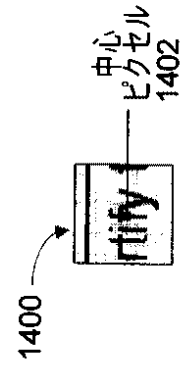
【図 11】



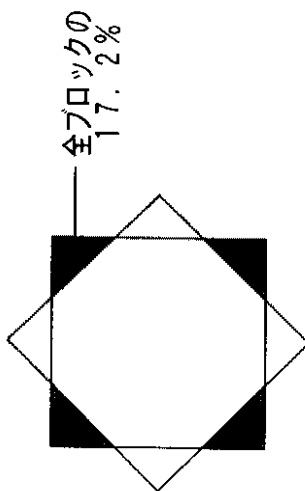
【図 1 2】



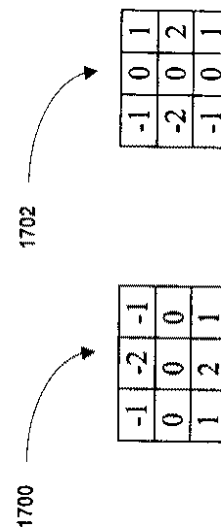
【図 1 4】



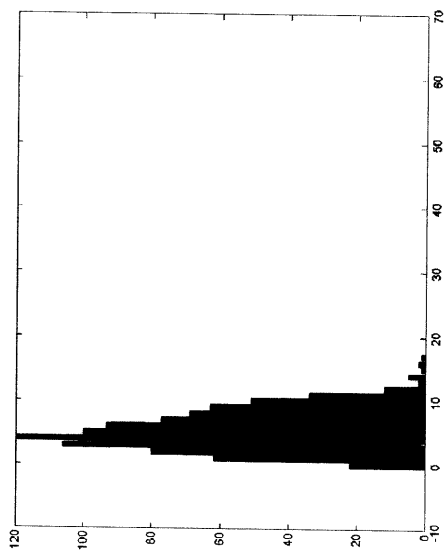
【図 1 6】



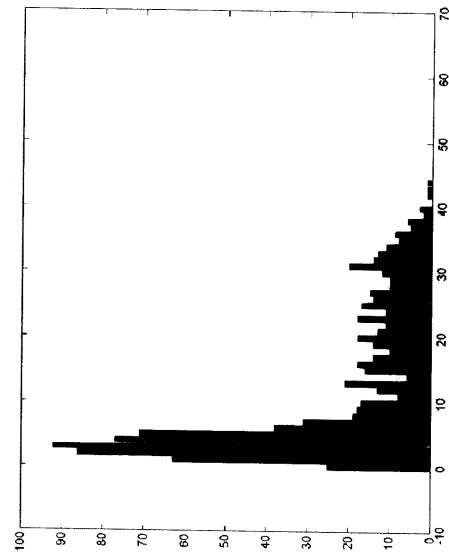
【図 1 7】



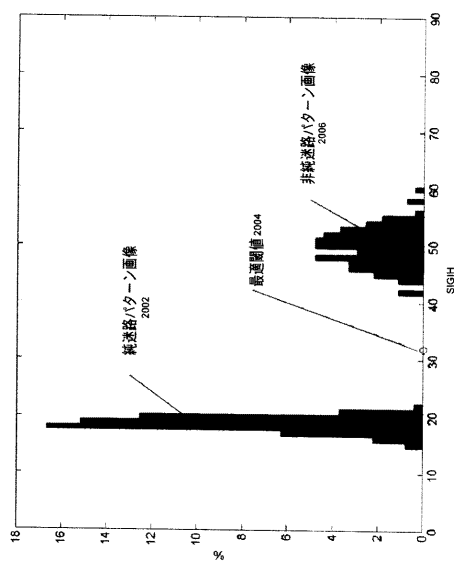
【図 18】



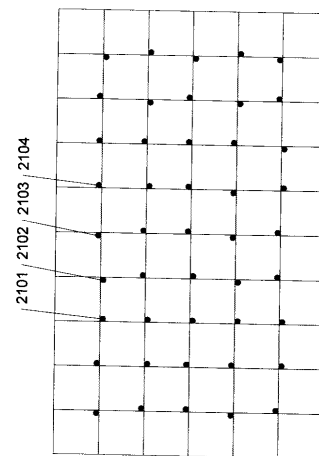
【図 19】



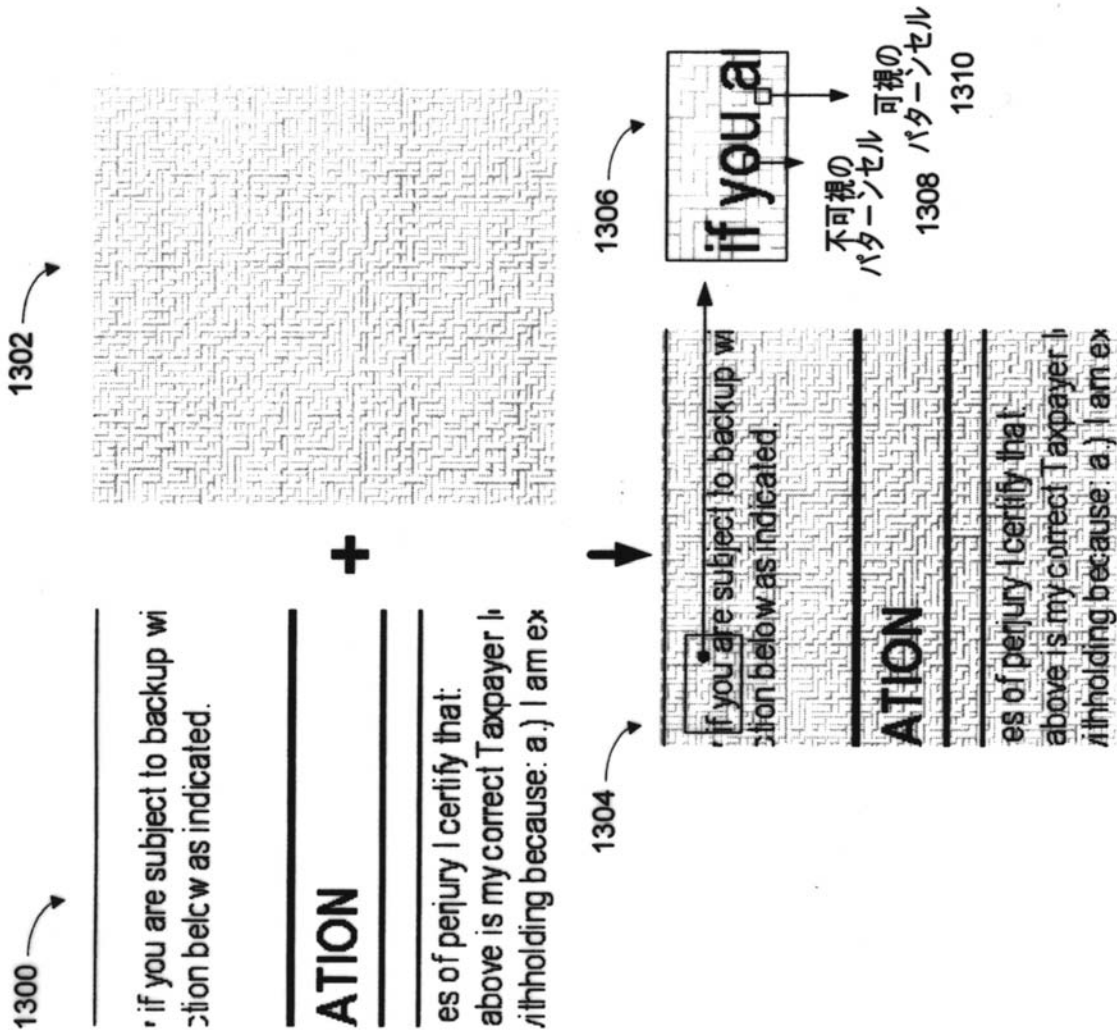
【図 20】



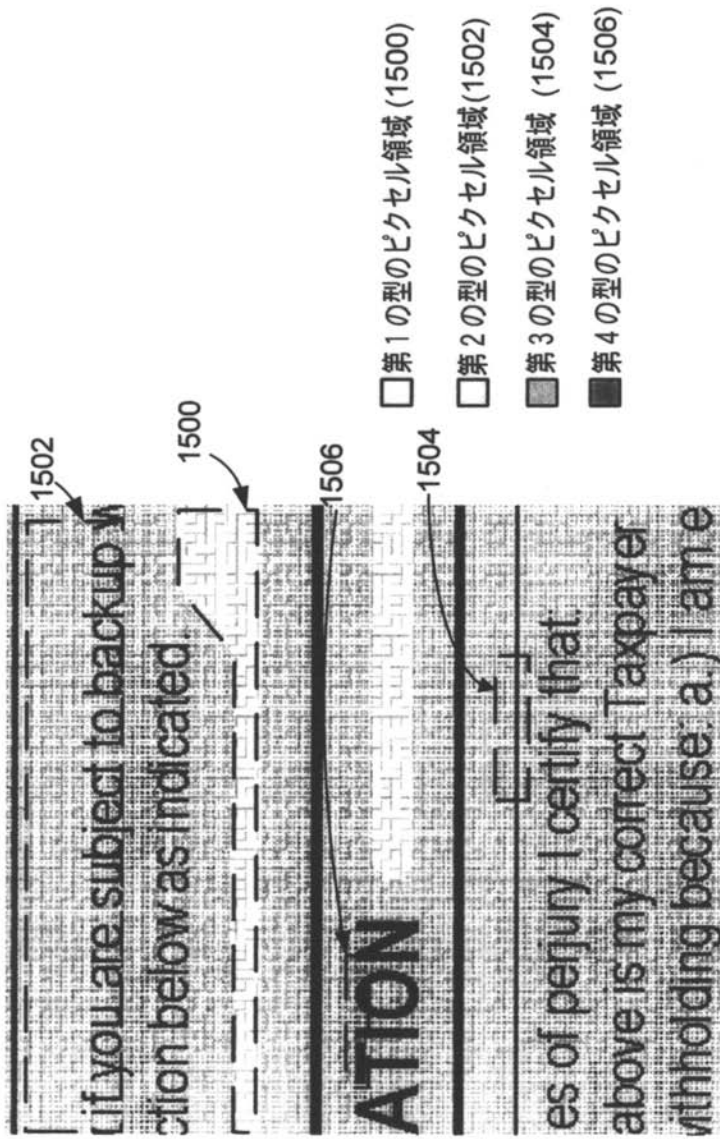
【図 21】



【図 13】



【図 15】



---

フロントページの続き

(72)発明者 リヨン チェン  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション内

(72)発明者 インノン ダン  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション内

Fターム(参考) 5B057 AA11 BA02 CA08 CA12 CA16 CB18 CC02 CG01 CH09 DA08  
DA17 DB02 DB09 DC03 DC36  
5B068 AA05 BD02 BD09 CC06  
5C076 AA14 BA01 BA06

【外国語明細書】  
2005228302000001.pdf