

(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2020-201435

(P2020-201435A)

(43) 公開日 令和2年12月17日(2020.12.17)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G10L 13/00 (2006.01)</b>	G10L 13/00 100K	5K127
<b>G06F 3/16 (2006.01)</b>	G06F 3/16 690	
<b>G06F 1/16 (2006.01)</b>	G06F 3/16 630	
<b>H04M 1/00 (2006.01)</b>	G06F 3/16 610	
<b>G09B 21/00 (2006.01)</b>	G06F 1/16 313A	

審査請求 有 請求項の数 15 O L (全 20 頁) 最終頁に続く

(21) 出願番号 特願2019-109747 (P2019-109747)  
 (22) 出願日 令和1年6月12日 (2019.6.12)  
 (11) 特許番号 特許第6773844号 (P6773844)  
 (45) 特許公報発行日 令和2年10月21日 (2020.10.21)

(71) 出願人 397019380  
 株式会社ポニーキャニオン  
 東京都港区六本木一丁目5番17号  
 (74) 代理人 110000338  
 特許業務法人HARAKENZO WORLD PATENT & TRADEMARK  
 (72) 発明者 黒澤 格  
 東京都港区六本木1-5-17 株式会社ポニーキャニオン内  
 Fターム(参考) 5K127 AA31 BA03 BB18 CA31 CB33  
 GD07 JA25 KA04 KA05

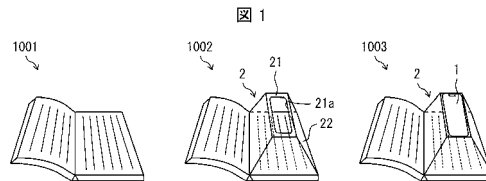
(54) 【発明の名称】 情報処理端末及び情報処理方法

(57) 【要約】

【課題】 視覚障害者に対してより好適に音読を通して複数頁で構成される文書を把握し、文章を楽しむ体験を提供することが可能な情報処理端末を実現する。

【解決手段】 情報処理端末(1)は、複数の頁で構成される文書の各頁を撮像するカメラ(11)と、テキストデータ生成処理、及び、音声信号生成処理を実行するコントローラ(12)と、前記音声信号を出力する出力部、及び、前記音声信号を音波に変換して出力するスピーカ的一方又は両方と、を備えている。

【選択図】 図1



## 【特許請求の範囲】

## 【請求項 1】

複数の頁で構成される文書の各頁を撮像するカメラと、

前記カメラにて得られた画像データを参照して、前記頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理、及び、前記テキストデータを参照して、前記テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理を実行するコントローラと、

前記音声信号を出力する出力部、及び、前記音声信号を音波に変換して出力するスピーカの一方又は両方と、を備えている、  
ことを特徴とする情報処理端末。

10

## 【請求項 2】

前記コントローラは、前記画像データを参照して、前記頁が前記カメラの撮像対象領域から外れた後に、前記テキストデータ生成処理を実行するか、又は、

前記コントローラは、前記画像データを参照して、前記頁が前記カメラの撮像対象領域から外れた後に、前記音声信号生成処理を実行する、  
ことを特徴とする請求項 1 に記載の情報処理端末。

## 【請求項 3】

前記音声信号生成処理において、前記コントローラは、前記音声信号として、ユーザにより指定された読み上げ速度で前記頁に記載されたテキストを読み上げた音声を表す音声信号を生成する、

20

ことを特徴とする請求項 1 又は 2 に記載の情報処理端末。

## 【請求項 4】

前記テキストデータ生成処理において、前記コントローラは、前記テキストデータとして、前記頁に記載された頁番号を表す頁番号データ、及び、前記頁に記載された本文を表す本文データを生成し、

前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記頁番号を読み上げた音声を表す音声信号の後に、前記本文を読み上げた音声を表す音声信号を付加した音声信号を生成する、

ことを特徴とする請求項 1 ~ 3 の何れか 1 項に記載の情報処理端末。

## 【請求項 5】

30

前記テキストデータ生成処理において、前記コントローラは、前記本文データとして、  
(1) 前頁本文の末尾が未完単語である場合、現頁本文の先頭に当該未完単語を付加すると共に、  
(2) 現頁本文の末尾が未完単語である場合、現頁本文の末尾から当該未完単語を削除したものを表す本文データを生成する、

ことを特徴とする請求項 1 ~ 4 の何れか 1 項に記載の情報処理端末。

## 【請求項 6】

前記テキストデータ生成処理において、前記コントローラは、前記テキストデータに含まれる見出しを特定し、

前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記見出しを聞き分け可能に前記テキストを読み上げた音声を表す音声信号を生成する、

40

ことを特徴とする請求項 1 ~ 5 の何れか 1 項に記載の情報処理端末。

## 【請求項 7】

前記頁が前記複数の頁で構成される文書の表紙である場合、

前記テキストデータ生成処理において、前記コントローラは、前記テキストデータとして、前記表紙に記載された題名を表す題名データ、及び、前記表紙に記載された著者名を表す著者名データを生成し、

前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記題名を読み上げた音声を表す音声信号の前又は後に、前記著者名を読み上げた音声を表す音声信号を付加した音声信号を生成する、

ことを特徴とする請求項 1 ~ 6 の何れか 1 項に記載の情報処理端末。

50

**【請求項 8】**

前記テキストデータ生成処理において、前記コントローラは、前記表紙に記載された文字列の中で、最も大きいフォントで記載された文字列を前記題名として特定すると共に、前記表紙に記載された文字列の中で、2番目に大きいフォントで記載された文字列を著者名として特定する、  
ことを特徴とする請求項 7 に記載の情報処理端末。

**【請求項 9】**

複数の頁で構成される文書の表紙を表す画像データを入力とし、当該複数の頁で構成される文書の題名及び著者名を表すテキストデータを出力とする分類器、又は、当該分類器を備えたサーバと通信するための通信インターフェースを更に備え、

前記テキストデータ生成処理において、前記コントローラは、前記カメラにて得られた画像データを入力すると共に、前記分類器から出力されるテキストデータから前記題名データ及び前記著者名データを生成する、  
ことを特徴とする請求項 7 に記載の情報処理端末。

**【請求項 10】**

複数の著者名が予め登録されたデータベース、又は、当該データベースを備えたサーバと通信するための通信インターフェースを更に備え、

前記テキストデータ生成処理において、前記コントローラは、前記表紙に記載された文字列の中で、前記データベースに著者名として登録された文字列を前記著者名データとする、  
ことを特徴とする請求項 7 に記載の情報処理端末。

**【請求項 11】**

前記コントローラは、前記カメラにて得られた画像データが表す画像が前記頁の逆立画像あるか否かを判定する逆立画像判定処理、及び、前記逆立画像判定処理にて前記画像が前記逆立画像であると判定された場合に、その旨をユーザに通知する逆立画像通知処理を更に実行する、

ことを特徴とする請求項 1 ~ 10 の何れか 1 項に記載の情報処理端末。

**【請求項 12】**

前記コントローラは、前記頁に記載されたテキストの文字数が予め定められた閾値を下回るか否かを判定する文字過少頁判定処理、及び、文字過少頁判定処理にて前記文字数が前記閾値を下回ると判定された場合、その旨をユーザに通知する文字過少頁通知処理を更に実行する、

ことを特徴とする請求項 1 ~ 11 の何れか 1 項に記載の情報処理端末。

**【請求項 13】**

前記頁に記載された頁番号を表す頁番号データを格納するメモリを更に備え、

前記コントローラは、前記メモリに格納された頁番データが表す頁番号のうち、最大の頁番号を通知する最大頁番号通知処理を更に実行する、

ことを特徴とする請求項 1 ~ 12 の何れか 1 項に記載の情報処理端末。

**【請求項 14】**

前記コントローラは、前記頁に記載されているテキストが縦書きテキストであるか横書きテキストであるかを判定する縦横判定処理を更に実行し、

前記コントローラは、前記縦横判定処理の結果を参照して前記テキストデータ生成処理を実行する、

ことを特徴とする請求項 1 ~ 13 の何れか 1 項に記載の情報処理端末。

**【請求項 15】**

前記コントローラは、朗読停止を指示する音声コマンドをユーザから取得すると、前記音声信号の出力を停止し、朗読再開を指示する音声コマンドをユーザから取得すると、前記音声信号の出力を再開する、

ことを特徴とする請求項 1 ~ 14 の何れか 1 項に記載の情報処理端末。

**【請求項 16】**

10

20

30

40

50

複数の頁で構成される文書の各頁と情報処理端末のカメラレンズとが正対するように前記情報処理端末を支持する支持体において、

前記情報処理端末が載置される天板であって、前記カメラレンズと重なる部分が開放された天板と、

上端が前記天板の外縁に連結されると共に、下端が前記頁の外縁に押し当てられる側板と、を備えている、

ことを特徴とする支持体。

【請求項 17】

前記天板及び前記側板の一方又は両方は、透光性を有している、ことを特徴とする請求項 16 に記載の支持体。

10

【請求項 18】

上記側板は、上記天板に展開可能に連結されている、ことを特徴とする請求項 16 又は 17 に記載の支持体。

【請求項 19】

前記天板に直交する方向に沿って測った、前記天板から前記側板の下端までの距離は、前記カメラレンズの被写界深度に含まれる、ことを特徴とする請求項 16 ~ 18 の何れか 1 項に記載の支持体。

【請求項 20】

前記天板及び前記側板の一方又は両方には、当該支持体の利用に適した複数の頁で構成される文書の種類を触知可能に示す構造、及び、当該支持体の利用に適した情報処理端末の種類を触知可能に示す構造の一方又は両方が設けられている、ことを特徴とする請求項 16 ~ 19 の何れか 1 項に記載の情報処理端末の支持体。

20

【請求項 21】

カメラと、コントローラと、出力部、及び、スピーカの一方又は両方と、を備えた情報処理端末を用いて複数の頁で構成される文書を音声として出力する情報処理方法であって、

前記カメラが、複数の頁で構成される文書の各頁を撮像する撮像処理と、

前記コントローラが、前記カメラにて得られた画像データを参照して、前記頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理と、

前記コントローラが、前記テキストデータを参照して、前記テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理と、

30

前記出力部が前記音声信号を出力する、及び/又は、前記スピーカが前記音声信号を音波に変換して出力する音声出力処理と、を含んでいる、

ことを特徴とする情報処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理端末に関する。また、本発明は、情報処理端末を載置するための支持体に関する。また、本発明は、情報処理方法に関する。

【背景技術】

40

【0002】

視覚障害者が利用する文書の形態として、活字図書を点字に変換した点字図書、活字図書を音声により読み上げた音訳図書、活字図書の文字や図を拡大し、見やすく整えて出版した拡大図書、及び電子書籍が知られている（特許文献 1）。

【先行技術文献】

【非特許文献】

【0003】

【非特許文献 1】 ” 視覚障害者の読書について ”、[online]、社会福祉法人日本盲人会連合、[令和 1 年 5 月 27 日検索]、インターネット < URL : <http://nichimou.org/morebooks/reading/> >

50

**【発明の概要】****【発明が解決しようとする課題】****【0004】**

しかしながら、点字図書、音訳図書、及び拡大図書は、作成に時間が必要であり、これらの図書を視覚障害者が利用可能になる時期が遅いという問題がある。さらに、点字図書及び音訳図書は、コストの観点からボランティアに頼った運用となるという問題がある。また、拡大図書及び電子書籍は全盲者が利用できないという問題がある。さらには、これらは市販された書籍を対象としたものであるが、たとえば学术论文、地方自治体などが発行する広報誌、新聞、新聞の折り込み広告、パンフレットなどの広範な印刷物が網羅的に対象とされている状況ではない。そのため、視覚障害者に向けた、より好適な、文章を音読するツールの開発が望まれている。

10

**【0005】**

本発明の一態様は、前記の課題を鑑みてなされたものであり、その目的は、視覚障害者に対してより好適に音読を通して複数頁で構成される文書を把握し、文章を楽しむ体験を提供することが可能な情報処理端末を実現することにある。

**【課題を解決するための手段】****【0006】**

前記の課題を解決するために、本発明の一態様に係る情報処理端末は、複数の頁で構成される文書の各頁を撮像するカメラと、前記カメラにて得られた画像データを参照して、前記頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理、及び、前記テキストデータを参照して、前記テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理を実行するコントローラと、前記音声信号を出力する出力部、及び、前記音声信号を音波に変換して出力するスピーカの一方又は両方と、を備えている。

20

**【0007】**

前記の課題を解決するために、本発明の一態様に係る支持体は、書籍の頁と情報処理端末のカメラレンズとが正対するように前記情報処理端末を支持する支持体において、前記情報処理端末が載置される天板であって、前記カメラレンズと重なる部分が開放された天板と、上端が前記天板の外縁に連結されると共に、下端が前記頁の外縁に押し当てられる側板と、を備えている。

30

**【0008】**

前記の課題を解決するために、本発明の一態様に係る情報処理方法は、カメラと、コントローラと、出力部、及び、スピーカの一方又は両方と、を備えた情報処理端末を用いて複数の頁で構成される文書を音声として出力する情報処理方法であって、前記カメラが、複数の頁で構成される文書の各頁を撮像する撮像処理と、前記コントローラが、前記カメラにて得られた画像データを参照して、前記頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理と、前記コントローラが、前記テキストデータを参照して、前記テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理と、前記出力部が前記音声信号を出力する、及び/又は、前記スピーカが前記音声信号を音波に変換して出力する音声出力処理と、を含んでいる。

40

**【発明の効果】****【0009】**

本発明の一態様によれば、視覚障害者に対してより好適に音読を通して複数頁で構成される文書を把握し、文章を楽しむ体験を提供することが可能な情報処理端末を実現することができる。

**【図面の簡単な説明】****【0010】**

【図1】本発明の一実施形態に係る情報処理端末の使用方法の概略を示す図である。

【図2】本発明の一実施形態に係る情報処理端末の構成を示すブロック図である。

【図3】図2に示す情報処理端末を用いて実施される情報処理方法の流れを示すフローチ

50

ャートである。

【図4】本発明の一実施形態に係る情報処理端末の支持体の構成を示す斜視図である。

【図5】本発明の一実施形態に係る情報処理端末の支持体の構成を示す展開図である。

【発明を実施するための形態】

【0011】

〔基本コンセプト〕

本実施形態の基本コンセプトについて、図1を参照して説明する。図1は、本実施形態に係る情報処理端末1及び支持体2の使用法の概略を示す斜視図である。情報処理端末1としては、例えば、カメラ付きのスマートフォンが挙げられる。

【0012】

まず、図1の1001に示すように、ユーザは、書籍をはじめとする複数の頁で構成される文書の頁を開く。複数の頁で構成される文書としては、例えば、新刊書や文庫本等の活字図書、学术论文、地方自治体などが発行する広報誌、新聞、新聞の折り込み広告、パンフレットなどが挙げられる。次いで、図1の1002に示すように、ユーザは、開いた頁の上に支持体2を載置する。これにより、開いた頁の外縁に支持体2の側板22の下端が押し当てられ、その結果、開いた頁は、状態が撮像に適した平坦な状態になる。その後、図1の1003に示すように、ユーザは、支持体2の天板21上に情報処理端末1を載置する。これにより、情報端末装置1のカメラレンズは、天板21に設けられた開口21aを介して開いた頁に正対し、情報端末装置1のカメラレンズから開いた頁までの距離は、当該カメラレンズの被写界深度に含まれるようになる。情報処理端末1には、この状態において開いた頁を撮像し、得られた画像データを参照して当該頁に記載されたテキストを読み上げる。これにより、ユーザは、開いた頁に記載されたテキストを、音声として聴取することができる。

【0013】

〔情報処理端末〕

本実施形態に係る情報処理端末1の構成について、図2を参照して説明する。図2は、情報処理端末1の構成を示すブロック図である。以下、複数の頁で構成される文書が書籍である場合を一例として説明する。

【0014】

情報処理端末1には、書籍の頁に記載されたテキストを読み上げるためのアプリケーションプログラム（以下、「朗読アプリ」と記載する）がインストールされている。情報処理端末1は、朗読アプリに従って朗読処理を実行するために利用するハードウェア資源として、カメラ11と、コントローラ12と、出力端子13（出力部）と、スピーカ14と、メモリ15と、通信インターフェース16と、を備えている。情報処理端末1は、例えば、スマートフォンで実現することができる。朗読アプリは、メモリ15に保存されていてもよいし、それ以外の記録媒体に保存されていてもよい。

【0015】

カメラ11は、書籍の頁を撮像するための構成である。カメラ11としては、例えば、カメラレンズ、撮像素子、及び画像信号処理回路とにより構成されるデジタルカメラを用いられる。メモリ15としては、例えば、1又は複数の半導体RAM（random access memory）が用いられる。

【0016】

コントローラ12は、朗読アプリに従って朗読処理を実行するための構成である。コントローラ12としては、例えば、1又は複数のCPU（Central Processing Unit）が用いられる。朗読処理には、少なくともテキストデータ生成処理及び音声信号生成処理が含まれる。ここで、テキストデータ生成処理とは、カメラ11にて得られた画像データを参照して、書籍の頁に記載されたテキストを表すテキストデータを生成する処理のことを指す。また、音声信号生成処理とは、テキストデータ生成処理にて生成されたテキストデータを参照して、書籍の頁に記載されたテキストを読み上げた音声を表す音声信号を生成する処理のことを指す。朗読処理の詳細については、参照する図面を代えて後述する。

10

20

30

40

50

## 【 0 0 1 7 】

出力端子 1 3 は、音声信号生成処理にて生成された音声信号を、電気信号として外部に出力するための構成である。音声信号の出力先としては、例えば、イヤホン 3 が挙げられる。スピーカ 1 4 は、音声信号生成処理にて生成された音声信号を、音波として外部に出力するための構成である。情報処理端末 1 は、出力端子 1 3 のみを備えていてもよいし、スピーカ 1 4 のみを備えていてもよいし、出力端子 1 3 及びスピーカ 1 4 の両方を備えていてもよい。また、情報処理端末 1 は、有線インターフェースである出力端子 1 3 の代わりに近距離無線通信インターフェースを備えていてもよいし、出力端子 1 3 に加えて無線インターフェースを備えていてもよい。近距離無線通信インターフェースとしては、例えば、Bluetooth (登録商標) インターフェースが挙げられる。情報処理端末 1 が近距離無線通信インターフェースを備えている場合、出力端子 1 3 に接続して用いるイヤホン 3 の代わりに、近距離無線通信に対応したイヤホン、ヘッドホン、スピーカ等に音声を出力することができる。

10

## 【 0 0 1 8 】

通信インターフェース 1 6 は、サーバと通信するためのインターフェースである。コントローラ 1 2 は、通信インターフェース 1 6 を介してサーバとの通信を行うことによって、朗読処理の一部をサーバと連携して実行することができる。コントローラ 1 2 がサーバと連携して実行する処理の具体例については、参照する図面を代えて後述する。

## 【 0 0 1 9 】

〔朗読処理〕

20

朗読アプリに従って情報処理端末 1 が実行する朗読処理 S 1 の流れについて、図 3 を参照して説明する。図 3 は、朗読処理 S 1 (複数の頁で構成される文書を音声として出力する情報処理方法) の流れを示すフローチャートである。

## 【 0 0 2 0 】

朗読処理 S 1 は、図 3 に示すように、撮像処理 S 1 0 と、逆立画像判定処理 S 1 1 と、逆立画像通知処理 S 1 2 と、表紙判定処理 S 1 3 と、テキストデータ生成処理 S 1 4 と、音声信号生成処理 S 1 5 と、音声出力処理 S 1 6 と、縦横判定処理 S 1 7 と、テキストデータ生成処理 S 1 8 と、文字過少頁判定処理 S 1 9 と、文字過少頁通知処理 S 2 0 と、音声信号生成処理 S 2 1 と、音声出力処理 S 2 2 と、を含んでいる。

## 【 0 0 2 1 】

30

撮像処理 S 1 0 は、書籍の頁を撮像する処理である。撮像処理 S 1 0 は、コントローラ 1 2 に制御されたカメラ 1 1 によって実行される。なお、撮像処理 S 1 0 を実行するタイミングは、音声コマンドによってユーザが撮像を指示したタイミングであってもよいし、カメラ 1 1 が書籍の頁に対するフォーカスが成功したタイミングであってもよい。

## 【 0 0 2 2 】

逆立画像判定処理 S 1 1 は、撮像処理 S 1 0 にて得られた画像データの表す画像が逆立画像あるか否か(正立画像であるか)を判定する処理である。ここで、逆立画像とは、頁上方が画面下方に、且つ、頁下方が画面上方に配置された画像のことを指す。また、正立画像とは、頁上方が画面上方に、且つ、頁下方が画面下方に配置された画像のことを指す。逆立画像判定処理 S 1 1 は、コントローラ 1 2 によって実行される。

40

## 【 0 0 2 3 】

逆立画像判定処理 S 1 1 にて逆立画像であると判定された場合、逆立画像通知処理 S 1 2 が実行される。逆立画像通知処理 S 1 2 は、撮像処理 S 1 0 にて得られた画像データが逆立画像を表すことをユーザに通知する処理である。逆立画像通知処理 S 1 2 を実行するために、コントローラ 1 2 は、例えば、「本の上下が逆になっています」という音声信号を出力端子 1 3 又はスピーカ 1 4 に出力する。コントローラ 1 2 が出力端子 1 3 に当該音声信号を出力した場合、出力端子 1 3 は、当該音声信号をイヤホン 3 に出力する。イヤホン 3 は、当該音声信号を音波に変換して出力する。コントローラ 1 2 がスピーカ 1 4 に当該音声信号を出力した場合、スピーカ 1 4 は、当該音声信号を音波に変換して出力する。これにより、本を 1 8 0 ° 回転して正しい向きにする動作をユーザに促すことができる。

50

## 【 0 0 2 4 】

なお、コントローラ 1 2 は、逆立画像判定処理 S 1 1 にて画像が逆立画像であると判定された場合に、逆立画像を 1 8 0 ° 回転して正立画像として、後述する表紙判定処理 S 1 3 に進んでもよい。

## 【 0 0 2 5 】

逆立画像判定処理 S 1 1 にて正立画像であると判定された場合、表紙判定処理 S 1 3 が実行される。表紙判定処理 S 1 3 は、撮像処理 S 1 0 にて得られた画像データの表す画像に被写体として含まれる頁が表紙であるか否か（表紙以外であるか）を判定する処理である。表紙判定処理 S 1 3 は、情報処理端末 1 のコントローラ 1 2 によって実行される。

## 【 0 0 2 6 】

表紙判定処理 S 1 3 にて頁が表紙であると判定された場合、テキストデータ生成処理 S 1 4 が実行される。テキストデータ生成処理 S 1 4 は、撮像処理 S 1 0 にて得られた画像データを参照して、頁に記載されたテキストを表すテキストデータを生成する処理である。テキストデータ生成処理 S 1 4 は、情報処理端末 1 のコントローラ 1 2 によって実行される。特に、頁が書籍の表紙である場合に実行されるテキストデータ生成処理 S 1 4 において、コントローラ 1 2 は、テキストデータとして、表紙に記載された題名を表す題名データ、及び、表紙に記載された著者名を表す著者名データを生成する。題名データ及び著者名データの生成は、例えば、以下の方法により実現される。

## 【 0 0 2 7 】

方法 1：表紙に記載された文字列の中で、最も大きいフォントで記載された文字列を対題名として特定すると共に、表紙に記載された文字列の中で、2 番目に大きいフォントで記載された文字列を著者名として特定する。そして、特定した題名及び著者名を表すテキストデータを、題名データ及び著者名データとする。

## 【 0 0 2 8 】

方法 2：書籍の表紙を表す画像データを入力とし、その書籍の題名及び著者名を表す題名及び著者名を出力するよう、機械学習された分類器（AI）を利用する。この場合、この分類器に撮像処理 S 1 0 にて得られた画像データを入力したとき、この分類器から出力される題名及び著者名を表すテキストデータを、題名データ及び著者名データとする。なお、この分類器は、情報処理端末 1 に実装されていてもよいし、通信インターフェース 1 6 を介して情報処理端末 1 と通信可能なサーバに実装されていてもよい。後者の場合、情報処理端末 1 は、サーバに実装された分類器を利用して、表紙に記載された題名及び著者名を特定する。

## 【 0 0 2 9 】

方法 3：様々な書籍の題名が登録された題名データベース、及び、様々な書籍の著者名が登録された著者名データベースを利用する。この場合、表紙に記載された文字列の中で、題名データベースに登録された題名に合致する文字列を題名として特定すると共に、表紙に記載された文字列の中で、著者名データベースに登録された著者名に合致する文字列を著者名として特定する。そして、特定した題名及び著者名を表すテキストデータを、題名データ及び著者名データとする。なお、これらのデータベースは、情報処理端末 1 に実装されていてもよいし、通信インターフェース 1 6 を介して情報処理端末 1 と通信可能なサーバに実装されていてもよい。後者の場合、情報処理端末 1 は、サーバに実装されたデータベースを利用して、表紙に記載された題名及び著者名を特定する。

## 【 0 0 3 0 】

なお、著者名データベースには、漢字表記の著者名と共に、平仮名表記又は片仮名表記の著者名が登録されていることが好ましい。一例を挙げると、漢字表記の著者名「古沢良太」と共に、平仮名表記の著者名「こさわりょうた」が登録されていることが好ましい。この場合、著者名データを平仮名表記の著者名に基づいて生成すれば、後述する音声信号生成処理 S 1 5 において、著者名が正しく音読された音声データを生成することができる。題名データベースについても、同様のことが言える。

## 【 0 0 3 1 】

なお、コントローラ 12 は、撮像処理 S 10 にて得られた画像データを参照して、頁がカメラ 11 の撮像対象領域から外れた後に、テキストデータ生成処理 S 14 を実行してもよい。

【0032】

音声信号生成処理 S 15 は、テキストデータ生成処理 S 14 にて生成されたテキストデータを参照して、頁に記載されたテキストを読み上げた音声を表す音声信号を生成する処理である。音声信号生成処理 S 15 は、テキストデータ生成処理 S 14 を実行した後に、情報処理端末 1 のコントローラ 12 によって実行される。特に頁が表紙である場合に実行される音声信号生成処理 S 15 において、コントローラ 12 は、音声信号として、題名を読み上げた音声を表す音声信号の前又は後に、著者名を読み上げた音声を表す音声信号を付加した音声信号を生成する。

10

【0033】

なお、音声信号生成処理 S 15 において、コントローラ 12 は、撮像処理 S 10 にて得られた画像データを参照して、頁がカメラ 11 の撮像対象領域から外れた後に、音声信号生成処理を実行してもよい。また、音声信号生成処理 S 15 において、コントローラ 12 は、音声信号として、ユーザにより指定された読み上げ速度で頁に記載されたテキストを読み上げた音声を表す音声信号を生成してもよい。

【0034】

音声出力処理 S 16 は、音声信号生成処理 S 15 にて生成された音声信号を電気信号として出力端子 13 から出力する処理、及び、音声信号生成処理 S 15 にて生成された音声信号を音波としてスピーカ 14 から出力する処理の一方又は両方である。音声出力処理 S 16 は、コントローラ 12 によって実現される。

20

【0035】

表紙判定処理 S 13 にて頁が表紙以外であると判定された場合、縦横判定処理 S 17 が実行される。縦横判定処理 S 17 は、頁に記載されているテキストが縦書きテキストであるか横書きテキストであるかを判定する処理である。縦横判定処理 S 17 は、コントローラ 12 によって実行される。

【0036】

テキストデータ生成処理 S 18 は、縦横判定処理 S 17 の結果、及び、撮像処理 S 10 にて得られた画像データを参照して、頁に記載されたテキストデータを生成する処理である。テキストデータ生成処理 S 18 は、情報処理端末 1 のコントローラ 12 によって実行される。特に、頁が表紙以外である場合に実行されるテキストデータ生成処理 S 18 において、コントローラ 12 は、テキストデータとして、頁に記載された頁番号を表す頁番号データ、及び、頁に記載された本文を表す本文データを生成する。

30

【0037】

なお、テキストデータ生成処理 S 18 において、コントローラ 12 は、本文データとして、(1) 前頁本文の末尾が未完単語である場合、現頁本文の先頭に当該未完単語を付加すると共に、(2) 現頁本文の末尾が未完単語である場合、現頁本文の末尾から当該未完単語を削除したものを表す本文データを生成してもよい。例えば、前頁本文の末尾に「ラー」と記載され、現頁本文の先頭に「メン」と記載されている場合、すなわち、「ラーメン」という単語が 2 頁に亘って記載されている場合、コントローラ 12 は、現頁本文の先頭に「ラー」を付加したものを表す本文データ生成する。また、現頁本文の末尾に「チャ」と記載され、次頁本文の先頭に「ハン」と記載と記載されている場合、すなわち、「チャーハン」という単語が 2 頁に亘って記載されている場合、コントローラ 12 は、現頁本文の末尾から「チャ」を削除したものを表す本文データを生成する。なお、テキストデータ生成処理 S 18 において、本文の末尾が未完単語であるか否かを判定するアルゴリズムは、特に限定されず、公知のアルゴリズムを用いることができる。

40

【0038】

また、テキストデータ生成処理 S 18 において、コントローラ 12 は、テキストデータに含まれる見出しを特定してもよい。見出しを特定する方法としては、例えば、テキスト

50

データに含まれている文字列のうち、太字になっている文字列、又は、周囲にスペースが空いている文字列を見出しとして特定する方法が挙げられる。

【0039】

文字過少頁判定処理 S 1 9 は、頁に記載されたテキストの文字数が予め定められた閾値を下回るか否かを判定する処理である。文字過少頁判定処理 S 1 9 は、コントローラ 1 2 によって実行される。

【0040】

文字過少頁判定処理にて文字数が閾値を下回ると判定された場合、文字過少頁通知処理 S 2 0 が実行される。文字過少頁通知処理 S 2 0 は、記載されたテキストの文字数が少ない頁である旨をユーザに通知する処理である。文字過少頁通知処理 S 2 0 を実行するために、コントローラ 1 2 は、例えば、「頁全体が図になっています」との音声信号を出力端子 1 3 又はスピーカ 1 4 に出力する。コントローラ 1 2 が出力端子 1 3 に当該音声信号を出力した場合、出力端子 1 3 は、当該音声信号をイヤホン 3 に出力する。イヤホン 3 は、当該音声信号を音波に変換して出力する。コントローラ 1 2 がスピーカ 1 4 に当該音声信号を出力した場合、スピーカ 1 4 は、当該音声信号を音波に変換して出力する。これにより、ユーザが、当該頁には図が記載されていることを認識することができるため、ユーザに次頁の撮影に移行する動作を促すことができる。

10

【0041】

音声信号生成処理 S 2 1 は、テキストデータ生成処理 S 1 8 にて生成されたテキストデータを参照して、頁に記載されたテキストを読み上げた音声を表す音声信号を生成する処理である。音声信号生成処理 S 2 1 は、コントローラ 1 2 によって実行される。特に、頁が表紙以外の場合に実行される音声信号生成処理 S 2 1 において、コントローラ 1 2 は、音声信号として、頁番号を読み上げた音声を表す音声信号の後に、本文を読み上げた音声を表す音声信号を付加した音声信号を生成する。

20

【0042】

なお、音声信号生成処理 S 2 1 において、コントローラ 1 2 は、音声信号として、見出しを聞き分け可能にテキストを読み上げた音声を表す音声信号を生成してもよい。例えば、見出しを聞き分け可能にテキスト「・・・である。＜改行＞（3）旅立ち＜改行＞いよいよ私は、決心した。今日こそ・・・」を読み上げた音声としては、「である（休止）みだし（休止）さん（小休止）たびだち（休止）いよいよわたしはけっしんした（小休止）きょうこそ・・・」といった例が挙げられる。別の例として、見出しとそれ以外の部分とを、異なる声（例えば、見出しは男性の声、それ以外の部分は女性の声）で読み上げた音声データを用いてもよい。

30

【0043】

音声出力処理 S 2 2 は、音声信号生成処理 S 2 1 にて生成された音声信号を電気信号として出力端子 1 3 から出力する処理、及び、音声信号生成処理 S 2 1 にて生成された音声信号を音波としてスピーカ 1 4 から出力する処理の一方又は両方である。音声出力処理 S 2 2 は、コントローラ 1 2 によって実行される。

【0044】

音声出力処理 S 2 2 において、コントローラ 1 2 は、朗読停止を指示する音声コマンドをユーザから取得すると、前記音声信号の出力を停止し、朗読再開を指示する音声コマンドをユーザから取得すると、前記音声信号の出力を再開する。例えば、コントローラ 1 2 は、「停止」という音声コマンドをユーザから取得すると、音声信号の出力を停止する。また、「再開」という音声コマンドをユーザから取得すると、音声信号の出力を再開する。あるいは、コントローラ 1 2 は、ユーザによる情報処理端末 1 の画面をタップする操作によって音声信号の出力を停止し、更に、ユーザによる情報処理端末 1 の画面をタップする操作によって音声信号の出力を再開してもよい。なお、画面をタップする操作は、画面上に設けられた特定の GUI 要素（ボタン等）をタップする操作であってもよいし、画面の任意の場所をタップする操作であってもよい。

40

【0045】

50

以上のように、本実施形態に係る朗読処理 S 1 は、カメラ 1 1 にて得られた画像データを参照して、頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理 S 1 4、S 1 8、及び、テキストデータを参照して、テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理 S 1 5、S 2 1 を含んでいる。したがって、本実施形態に係る朗読処理 S 1 によれば、視覚障害者に対してより好適に音読を通して複数頁で構成される文書を把握し、文章を楽しむ体験を提供することができる。具体的には、視覚障害の度合いに依らずに、重度の視覚障害者であっても、軽度の視覚障害者であっても、視力の衰えた高齢者であっても、健常者であっても、読書することが可能である。また、点字図書、音訳図書、及びオーディオブックと比べて低コストである。さらに、活字図書を利用して読書することができるため、幅広い種類の書籍を即座に読書することができる。以上のように、本実施形態に係る情報処理端末 1 は、視覚障害者等のユーザが徹底的に使いやすい仕様となっている。

10

#### 【0046】

なお、朗読処理 S 1 は、前述した各処理以外にも、最大頁番号通知処理を含んでもよい。最大頁番号通知処理は、メモリ 1 5 に格納された頁番データが表す頁番号のうち、最大の頁番号を通知する処理である。最大頁番号通知処理は、情報処理端末 1 のコントローラ 1 2 によって実行される。

#### 【0047】

最大頁番号通知処理を実行するために、コントローラ 1 2 は、例えば、「30 頁まで読みました」という音声信号を出力端子 1 3 又はスピーカ 1 4 に出力する。コントローラ 1 2 が出力端子 1 3 に当該音声信号を出力した場合、出力端子 1 3 は、当該音声信号をイヤホン 3 に出力する。イヤホン 3 は、当該音声信号を音波に変換して出力する。コントローラ 1 2 がスピーカ 1 4 に当該音声信号を出力した場合、スピーカ 1 4 は、当該音声信号を音波に変換して出力する。これにより、ユーザに何ページまで読んだかを認識させることができる。

20

#### 【0048】

##### 〔支持体の構成〕

本実施形態に係る支持体 2 の構成について、図 4、5 を参照して説明する。図 4 は、支持体 2 の構成を示す斜視図である。図 5 は、支持体 2 の構成を示す展開図である。なお、支持体 2 の使用方法については、上述した通りである。

30

#### 【0049】

支持体 2 は、書籍の頁と情報処理端末 1 のカメラレンズとが正対するように情報処理端末 1 を支持するための台である。支持体 2 は、1 つの天板 2 1 と、4 つの側板 2 2 と、を備えている。

#### 【0050】

天板 2 1 は、支持体 2 の上面を構成する板状部材である。本実施形態において、天板 2 1 の平面視形状は、長方形である。情報処理端末 1 は、天板 2 1 の上に載置される。天板 2 1 には、情報処理端末 1 のカメラレンズと重なる部分を開放するための開口 2 1 a が形成されている。天板 2 1 及び開口 2 1 a の形状及びサイズは、対象とする情報処理端末 1 の種類に応じて最適化されている。例えば、天板 2 1 の形状及びサイズは、情報処理端末 1 の背面の形状及びサイズに応じて設定されている。また、開口 2 1 a の形状及びサイズは、情報処理端末 1 の背面におけるカメラレンズの位置に応じて設定されている。

40

#### 【0051】

側板 2 2 は、支持体 2 の側面を構成する板状部材であり、側板 2 2 の上端は、天板 2 1 の外縁に連結されている。本実施形態において、側板 2 2 の平面視形状は、台形である。側板 2 2 の下端は、書籍の頁の外縁に押し当てられる。側板 2 2 の形状及びサイズは、対象とする情報処理端末 1 の種類、及び、対象とする書籍の種類に応じて最適化されている。例えば、天板 2 1 に直交する方向に沿って測った天板 2 1 から側板 2 2 の下端までの距離が対象とする情報処理端末 1 のカメラレンズの被写界深度に含まれるように、且つ、側板 2 2 の下端が対象とする書籍の頁の外縁に押し当てられるように設定されている。

50

## 【0052】

本実施形態において、各側板22は、天板21の外縁に展開可能に連結されている。各側板22を展開して各側板22と天板21との成す角を180°とすると、支持体2は、図5に示すように平面化される。逆に、各側板22を折り畳んで各側板22と天板21との成す角を所定の角度(180°未満)とすると、支持体2は、図4に示すように立体化される。支持体2を販売するとき、及び、支持体2を保管するときには、可搬性を向上させるために支持体2を展開して平面化することが好ましい。

## 【0053】

なお、天板21及び側板22の一方又は両方は、透光性を有する材料により構成されていることが好ましい。これにより、天板21又は側板22を透過した外光を利用して、撮像対象とする頁の照度を向上させることができる。なお、カメラレンズと同じ面に光源を備えた情報処理端末1をユーザが使用する場合、この光源を撮像対象とする頁の照度を確保するために利用してもよい。

10

## 【0054】

また、天板21及び側板22の一方又は両方には、支持体2の利用に適した書籍の種類を触知可能に示す構造22a、及び、支持体2の利用に適した情報処理端末の種類を触知可能に示す構造22bの一方又は両方が設けられていることが好ましい。これにより、目の不自由なユーザであっても、支持体2が対象とする情報処理端末1及び/又は書籍に適したものであるか否かを触知することができる。なお、これらの構造22a、22bは、例えば、天板21又は側板22の表面に形成された凹凸によって実現することができる。

20

## 【0055】

〔ソフトウェアによる実現例〕

情報処理端末1のコントローラ12は、集積回路(ICチップ)等に形成された論理回路(ハードウェア)によって実現してもよいし、ソフトウェアによって実現してもよい。

## 【0056】

後者の場合、情報処理端末1は、各機能を実現するソフトウェアであるプログラムの命令を実行するコンピュータを備えている。このコンピュータは、例えば少なくとも1つのプロセッサ(制御装置)を備えていると共に、上記プログラムを記憶したコンピュータ読み取り可能な少なくとも1つの記録媒体を備えている。そして、上記コンピュータにおいて、上記プロセッサが上記プログラムを上記記録媒体から読み取って実行することにより、本発明の目的が達成される。上記プロセッサとしては、例えばCPU(Central Processing Unit)を用いることができる。上記記録媒体としては、「一時的でない有形の媒体」、例えば、ROM(Read Only Memory)等の他、テープ、ディスク、カード、半導体メモリ、プログラマブルな論理回路などを用いることができる。また、上記プログラムを展開するRAM(Random Access Memory)などをさらに備えていてもよい。また、上記プログラムは、該プログラムを伝送可能な任意の伝送媒体(通信ネットワークや放送波等)を介して上記コンピュータに供給されてもよい。なお、本発明の一態様は、上記プログラムが電子的な伝送によって具現化された、搬送波に埋め込まれたデータ信号の形態でも実現され得る。

30

## 【0057】

〔まとめ〕

本発明の態様1に係る情報処理端末は、複数の頁で構成される文書の各頁を撮像するカメラと、前記カメラにて得られた画像データを参照して、前記頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理、及び、前記テキストデータを参照して、前記テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理を実行するコントローラと、前記音声信号を出力する出力部、及び、前記音声信号を音波に変換して出力するスピーカの一方又は両方と、を備えている、構成である。

40

## 【0058】

本発明の態様2に係る情報処理端末は、上記の態様1において、前記コントローラは、前記画像データを参照して、前記頁が前記カメラの撮像対象領域から外れた後に、前記テ

50

キストデータ生成処理を実行するか、又は、前記コントローラは、前記画像データを参照して、前記頁が前記カメラの撮像対象領域から外れた後に、前記音声信号生成処理を実行する、構成としてもよい。

【0059】

本発明の態様3に係る情報処理端末は、上記の態様1又は2において、前記音声信号生成処理において、前記コントローラは、前記音声信号として、ユーザにより指定された読み上げ速度で前記頁に記載されたテキストを読み上げた音声を表す音声信号を生成する、構成としてもよい。

【0060】

本発明の態様4に係る情報処理端末は、上記の態様1～3の何れか1項において、前記テキストデータ生成処理において、前記コントローラは、前記テキストデータとして、前記頁に記載された頁番号を表す頁番号データ、及び、前記頁に記載された本文を表す本文データを生成し、前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記頁番号を読み上げた音声を表す音声信号の後に、前記本文を読み上げた音声を表す音声信号を付加した音声信号を生成する、構成としてもよい。

10

【0061】

本発明の態様5に係る情報処理端末は、上記の態様1～4の何れか1項において、前記テキストデータ生成処理において、前記コントローラは、前記本文データとして、(1)前頁本文の末尾が未完単語である場合、現頁本文の先頭に当該未完単語を付加すると共に、(2)現頁本文の末尾が未完単語である場合、現頁本文の末尾から当該未完単語を削除したものを表す本文データを生成する、構成としてもよい。

20

【0062】

本発明の態様6に係る情報処理端末は、上記の態様1～5の何れか1項において、前記テキストデータ生成処理において、前記コントローラは、前記テキストデータに含まれる見出しを特定し、前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記見出しを聞き分け可能に前記テキストを読み上げた音声を表す音声信号を生成する、構成としてもよい。

【0063】

本発明の態様7に係る情報処理端末は、上記の態様1～6の何れか1項において、前記頁が前記複数の頁で構成される文書の表紙である場合、前記テキストデータ生成処理において、前記コントローラは、前記テキストデータとして、前記表紙に記載された題名を表す題名データ、及び、前記表紙に記載された著者名を表す著者名データを生成し、前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記題名を読み上げた音声を表す音声信号の前又は後に、前記著者名を読み上げた音声を表す音声信号を付加した音声信号を生成する、構成としてもよい。

30

【0064】

本発明の態様8に係る情報処理端末は、上記の態様7において、前記テキストデータ生成処理において、前記コントローラは、前記表紙に記載された文字列の中で、最も大きいフォントで記載された文字列を前記題名として特定すると共に、前記表紙に記載された文字列の中で、2番目に大きいフォントで記載された文字列を著者名として特定する、構成としてもよい。

40

【0065】

本発明の態様9に係る情報処理端末は、上記の態様7において、複数の頁で構成される文書の表紙を表す画像データを入力とし、当該複数の頁で構成される文書の題名及び著者名を表すテキストデータを出力とする分類器、又は、当該分類器を備えたサーバと通信するための通信インターフェースを更に備え、前記テキストデータ生成処理において、前記コントローラは、前記カメラにて得られた画像データを入力すると共に、前記分類器から出力されるテキストデータから前記題名データ及び前記著者名データを生成する、構成としてもよい。

【0066】

50

本発明の態様10に係る情報処理端末は、上記の態様7において、複数の著者名が予め登録されたデータベース、又は、当該データベースを備えたサーバと通信するための通信インターフェースを更に備え、前記テキストデータ生成処理において、前記コントローラは、前記表紙に記載された文字列の中で、前記データベースに著者名として登録された文字列を前記著者名データとする、構成としてもよい。

【0067】

本発明の態様11に係る情報処理端末は、上記の態様1～10の何れか1項において、前記コントローラは、前記カメラにて得られた画像データが表す画像が前記頁の逆立画像あるか否かを判定する逆立画像判定処理、及び、前記逆立画像判定処理にて前記画像が前記逆立画像であると判定された場合に、その旨をユーザに通知する逆立画像通知処理を更に実行する、構成としてもよい。

10

【0068】

本発明の態様12に係る情報処理端末は、上記の態様1～11の何れか1項において、前記コントローラは、前記頁に記載されたテキストの文字数が予め定められた閾値を下回るか否かを判定する文字過少頁判定処理、及び、文字過少頁判定処理にて前記文字数が前記閾値を下回ると判定された場合、その旨をユーザに通知する文字過少頁通知処理を更に実行する、構成としてもよい。

【0069】

本発明の態様13に係る情報処理端末は、上記の態様1～12の何れか1項において、前記頁に記載された頁番号を表す頁番号データを格納するメモリを更に備え、前記コントローラは、前記メモリに格納された頁番号データが表す頁番号のうち、最大の頁番号を通知する最大頁番号通知処理を更に実行する、構成としてもよい。

20

【0070】

本発明の態様14に係る情報処理端末は、上記の態様1～13の何れか1項において、前記コントローラは、前記頁に記載されているテキストが縦書きテキストであるか横書きテキストであるかを判定する縦横判定処理を更に実行し、前記コントローラは、前記縦横判定処理の結果を参照して前記テキストデータ生成処理を実行する、構成としてもよい。

【0071】

本発明の態様15に係る情報処理端末は、上記の態様1～14の何れか1項において、前記コントローラは、朗読停止を指示する音声コマンドをユーザから取得すると、前記音声信号の出力を停止し、朗読再開を指示する音声コマンドをユーザから取得すると、前記音声信号の出力を再開する、構成としてもよい。

30

【0072】

本発明の態様16に係る支持体は、複数の頁で構成される文書の各頁と情報処理端末のカメラレンズとが正対するように前記情報処理端末を支持する支持体において、前記情報処理端末が載置される天板であって、前記カメラレンズと重なる部分が開放された天板と、上端が前記天板の外縁に連結されると共に、下端が前記頁の外縁に押し当てられる側板と、を備えている、構成である。

【0073】

本発明の態様17に係る支持体は、上記の態様16において、前記天板及び前記側板の一方又は両方は、透光性を有している、構成としてもよい。

40

【0074】

本発明の態様18に係る支持体は、上記の態様16又は17において、上記側板は、上記天板に展開可能に連結されている、構成としてもよい。

【0075】

本発明の態様19に係る支持体は、上記の態様16～18の何れか1項において、前記天板に直交する方向に沿って測った、前記天板から前記側板の下端までの距離は、前記カメラレンズの被写界深度に含まれる、構成としてもよい。

【0076】

本発明の態様20に係る情報処理端末の支持体は、上記の態様16～19の何れか1項

50

において、前記天板及び前記側板の一方又は両方には、当該支持体の利用に適した複数の頁で構成される文書の種類を触知可能に示す構造、及び、当該支持体の利用に適した情報処理端末の種類を触知可能に示す構造の一方又は両方が設けられている、構成としてもよい。

【 0 0 7 7 】

本発明の態様 2 1 に係る情報処理方法は、カメラと、コントローラと、出力部、及び、スピーカの一方又は両方と、を備えた情報処理端末を用いて複数の頁で構成される文書を音声として出力する情報処理方法であって、前記カメラが、複数の頁で構成される文書の各頁を撮像する撮像処理と、前記コントローラが、前記カメラにて得られた画像データを参照して、前記頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理と、前記コントローラが、前記テキストデータを参照して、前記テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理と、前記出力部が前記音声信号を出力する、及び/又は、前記スピーカが前記音声信号を音波に変換して出力する音声出力処理と、を含んでいる、構成としてもよい。

10

【 0 0 7 8 】

〔付記事項〕

本発明は上述した各実施形態に限定されるものではなく、請求項に示した範囲で種々の変更が可能であり、異なる実施形態にそれぞれ開示された技術的手段を適宜組み合わせ得られる実施形態についても本発明の技術的範囲に含まれる。

20

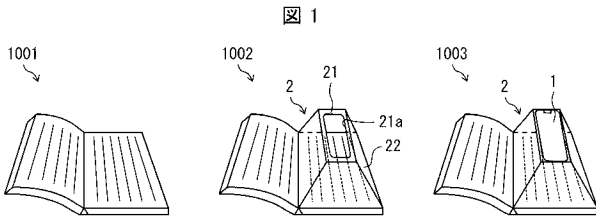
【符号の説明】

【 0 0 7 9 】

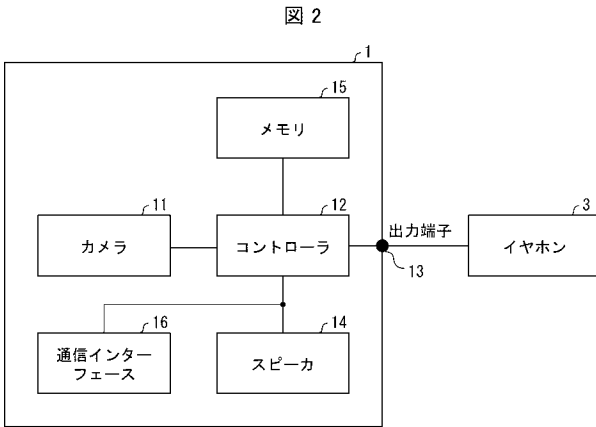
- 1 情報処理端末
- 2 支持体
- 3 イヤホン
- 1 1 カメラ
- 1 2 コントローラ
- 1 3 出力端子（出力部）
- 1 4 スピーカ
- 1 5 メモリ
- 2 1 天板
- 2 2 側板

30

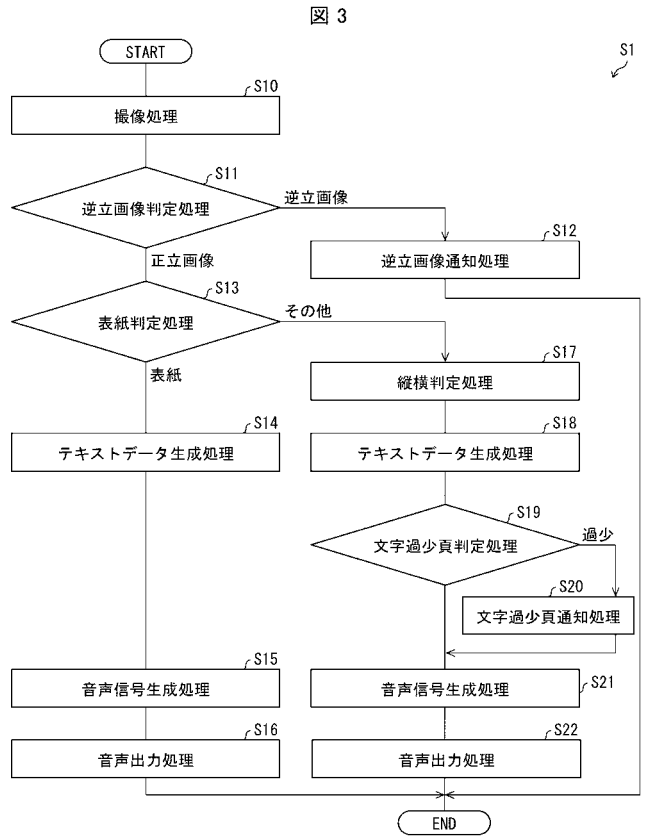
【 図 1 】



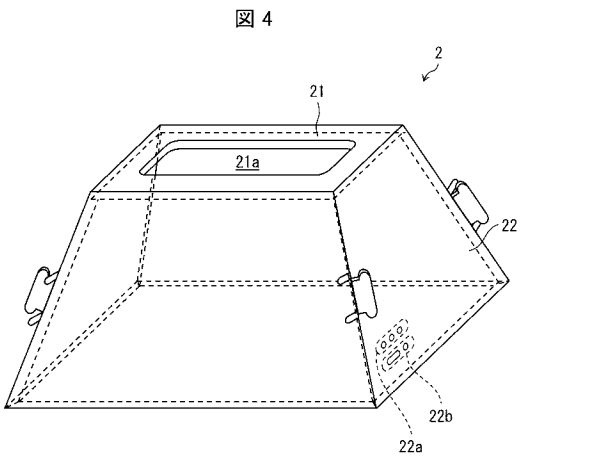
【 図 2 】



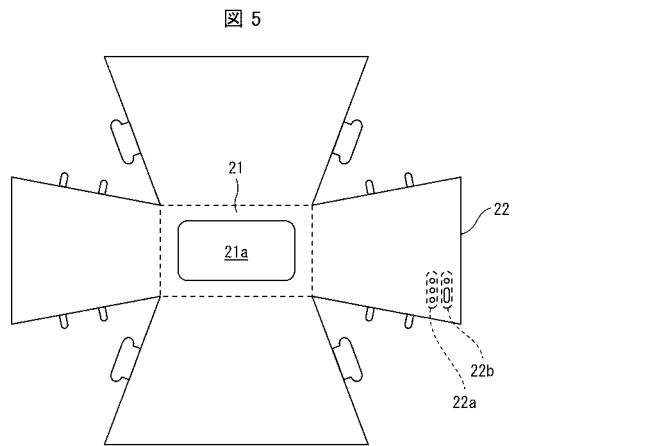
【 図 3 】



【 図 4 】



【 図 5 】



## 【手続補正書】

【提出日】令和2年7月21日(2020.7.21)

## 【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

## 【特許請求の範囲】

## 【請求項1】

複数の頁で構成される文書の各頁を撮像するカメラと、

前記カメラにて得られた画像データを参照して、前記頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理、及び、前記テキストデータを参照して、前記テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理を実行するコントローラと、

前記音声信号を出力する出力部、及び、前記音声信号を音波に変換して出力するスピーカの一方又は両方と、を備え、

前記コントローラは、前記画像データを参照して、前記頁が前記カメラの撮像対象領域から外れた後に、前記テキストデータ生成処理を実行するか、又は、

前記コントローラは、前記画像データを参照して、前記頁が前記カメラの撮像対象領域から外れた後に、前記音声信号生成処理を実行する、  
ことを特徴とする情報処理端末。

## 【請求項2】

前記音声信号生成処理において、前記コントローラは、前記音声信号として、ユーザにより指定された読み上げ速度で前記頁に記載されたテキストを読み上げた音声を表す音声信号を生成する、

ことを特徴とする請求項1に記載の情報処理端末。

## 【請求項3】

前記テキストデータ生成処理において、前記コントローラは、前記テキストデータとして、前記頁に記載された頁番号を表す頁番号データ、及び、前記頁に記載された本文を表す本文データを生成し、

前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記頁番号を読み上げた音声を表す音声信号の後に、前記本文を読み上げた音声を表す音声信号を付加した音声信号を生成する、

ことを特徴とする請求項1又は2に記載の情報処理端末。

## 【請求項4】

前記テキストデータ生成処理において、前記コントローラは、前記頁に記載された本文を表す本文データとして、(1)前頁本文の末尾が未完単語である場合、現頁本文の先頭に当該未完単語を付加すると共に、(2)現頁本文の末尾が未完単語である場合、現頁本文の末尾から当該未完単語を削除したものを表す本文データを生成する、

ことを特徴とする請求項1～3の何れか1項に記載の情報処理端末。

## 【請求項5】

前記テキストデータ生成処理において、前記コントローラは、前記テキストデータに含まれる見出しを特定し、

前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記見出しを聞き分け可能に前記テキストを読み上げた音声を表す音声信号を生成する、

ことを特徴とする請求項1～4の何れか1項に記載の情報処理端末。

## 【請求項6】

前記頁が前記複数の頁で構成される文書の表紙である場合、

前記テキストデータ生成処理において、前記コントローラは、前記テキストデータとして、前記表紙に記載された題名を表す題名データ、及び、前記表紙に記載された著者名を

表す著者名データを生成し、

前記音声信号生成処理において、前記コントローラは、前記音声信号として、前記題名を読み上げた音声を表す音声信号の前又は後に、前記著者名を読み上げた音声を表す音声信号を付加した音声信号を生成する、

ことを特徴とする請求項 1 ~ 5 の何れか 1 項に記載の情報処理端末。

【請求項 7】

前記テキストデータ生成処理において、前記コントローラは、前記表紙に記載された文字列の中で、最も大きいフォントで記載された文字列を前記題名として特定すると共に、前記表紙に記載された文字列の中で、2 番目に大きいフォントで記載された文字列を著者名として特定する、

ことを特徴とする請求項 6 に記載の情報処理端末。

【請求項 8】

複数の頁で構成される文書の表紙を表す画像データを入力とし、当該複数の頁で構成される文書の題名及び著者名を表すテキストデータを出力とする分類器、又は、当該分類器を備えたサーバと通信するための通信インターフェースを更に備え、

前記テキストデータ生成処理において、前記コントローラは、前記カメラにて得られた画像データを入力すると共に、前記分類器から出力されるテキストデータから前記題名データ及び前記著者名データを生成する、

ことを特徴とする請求項 6 に記載の情報処理端末。

【請求項 9】

複数の著者名が予め登録されたデータベース、又は、当該データベースを備えたサーバと通信するための通信インターフェースを更に備え、

前記テキストデータ生成処理において、前記コントローラは、前記表紙に記載された文字列の中で、前記データベースに著者名として登録された文字列を前記著者名データとする、

ことを特徴とする請求項 6 に記載の情報処理端末。

【請求項 10】

前記コントローラは、前記カメラにて得られた画像データが表す画像が前記頁の逆立画像あるか否かを判定する逆立画像判定処理、及び、前記逆立画像判定処理にて前記画像が前記逆立画像であると判定された場合に、その旨をユーザに通知する逆立画像通知処理を更に実行する、

ことを特徴とする請求項 1 ~ 9 の何れか 1 項に記載の情報処理端末。

【請求項 11】

前記コントローラは、前記頁に記載されたテキストの文字数が予め定められた閾値を下回るか否かを判定する文字過少頁判定処理、及び、文字過少頁判定処理にて前記文字数が前記閾値を下回ると判定された場合、その旨をユーザに通知する文字過少頁通知処理を更に実行する、

ことを特徴とする請求項 1 ~ 10 の何れか 1 項に記載の情報処理端末。

【請求項 12】

前記頁に記載された頁番号を表す頁番号データを格納するメモリを更に備え、

前記コントローラは、前記メモリに格納された頁番データが表す頁番号のうち、最大の頁番号を通知する最大頁番号通知処理を更に実行する、

ことを特徴とする請求項 1 ~ 11 の何れか 1 項に記載の情報処理端末。

【請求項 13】

前記コントローラは、前記頁に記載されているテキストが縦書きテキストであるか横書きテキストであるかを判定する縦横判定処理を更に実行し、

前記コントローラは、前記縦横判定処理の結果を参照して前記テキストデータ生成処理を実行する、

ことを特徴とする請求項 1 ~ 12 の何れか 1 項に記載の情報処理端末。

【請求項 14】

前記コントローラは、朗読停止を指示する音声コマンドをユーザから取得すると、前記音声信号の出力を停止し、朗読再開を指示する音声コマンドをユーザから取得すると、前記音声信号の出力を再開する、  
ことを特徴とする請求項 1 ~ 1 3 の何れか 1 項に記載の情報処理端末。

【請求項 1 5】

カメラと、コントローラと、出力部、及び、スピーカの一方又は両方と、を備えた情報処理端末を用いて複数の頁で構成される文書を音声として出力する情報処理方法であって

、  
前記カメラが、複数の頁で構成される文書の各頁を撮像する撮像処理と、  
前記コントローラが、前記カメラにて得られた画像データを参照して、前記頁に記載されたテキストを表すテキストデータを生成するテキストデータ生成処理と、  
前記コントローラが、前記テキストデータを参照して、前記テキストを読み上げた音声を表す音声信号を生成する音声信号生成処理と、  
前記出力部が前記音声信号を出力する、及び / 又は、前記スピーカが前記音声信号を音波に変換して出力する音声出力処理と、を含み、  
前記コントローラは、前記画像データを参照して、前記頁が前記カメラの撮像対象領域から外れた後に、前記テキストデータ生成処理を実行するか、又は、  
前記コントローラは、前記画像データを参照して、前記頁が前記カメラの撮像対象領域から外れた後に、前記音声信号生成処理を実行する、  
ことを特徴とする情報処理方法。

---

フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

G 0 6 F	3/16	5 0 0
H 0 4 M	1/00	R
G 0 9 B	21/00	D