(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2012-109901
(P2012-109901A)
(43)公開日　平成24年6月7日(2012.6.7)

| (51)Int.Cl. | | F I | | | テーマコード（参考） |
|---|---|---|---|---|---|
| H04N 5/222 | (2006.01) | H04N | 5/222 | Z | 2H102 |
| G03B 17/18 | (2006.01) | G03B | 17/18 | Z | 2K103 |
| G03B 21/14 | (2006.01) | G03B | 21/14 | Z | 5C122 |
| G03B 21/00 | (2006.01) | G03B | 21/00 | D | |

審査請求 未請求　請求項の数 12　ＯＬ　（全 16 頁）

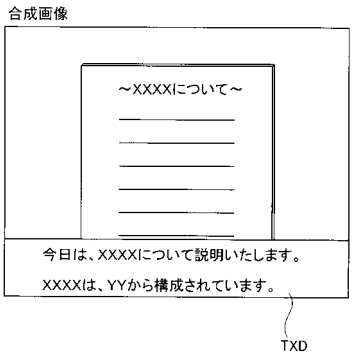
(21)出願番号　特願2010-258687 (P2010-258687)
(22)出願日　平成22年11月19日 (2010.11.19)

(71)出願人　000000424
株式会社エルモ社
愛知県名古屋市瑞穂区明前町６番１４号
(74)代理人　110000028
特許業務法人明成国際特許事務所
(72)発明者　須田　康嗣
名古屋市瑞穂区明前町６番１４号　株式会社エルモ社内

Ｆターム(参考)　2H102 BB03 BB22 BB26 BB33 CA06
2K103 AA16 AA22 AB10 BB05 BB09
CA53 CA54 CA71 CA72 CA73
5C122 DA28 EA47 FH19 FJ04 FK23
FK28 FK37 FK41 HA85 HB01
HB05

(54)【発明の名称】資料提示装置

(57)【要約】
【課題】資料提示装置に関する技術を提供する。
【解決手段】資料提示装置であって、所定領域を撮像し、該撮像した撮像画像を撮像画像データとして取得する撮像画像データ取得部と、音声を表す音声データを外部から取得する音声データ取得部と、取得した音声データに基づいて、音声に対応した所定言語の文字列を文字列データとして取得する文字列データ取得部と、撮像画像データと文字列データとに基づいて、撮像画像と文字列とを含む合成画像を合成画像データとして生成する画像合成部と、合成画像データを外部に出力する出力部とを備える。
【選択図】図５



【特許請求の範囲】
【請求項１】
資料提示装置であって、
所定領域を撮像し、該撮像した撮像画像を撮像画像データとして取得する撮像画像データ取得部と、
音声を表す音声データを外部から取得する音声データ取得部と、
前記取得した音声データに基づいて、前記音声に対応した所定言語の文字列を文字列データとして取得する文字列データ取得部と、
前記撮像画像データと前記文字列データとに基づいて、前記撮像画像と前記文字列とを含む合成画像を合成画像データとして生成する画像合成部と、
前記合成画像データを外部に出力する出力部と
を備える資料提示装置。
【請求項２】
請求項１記載の資料提示装置であって、
前記文字列データ取得部は、前記取得した音声データを認識し、前記所定言語の文字列データに変換する音声文字変換部を備える資料提示装置。
【請求項３】
請求項１記載の資料提示装置であって、
前記文字列データ取得部は、前記音声データに基づいて、回線を介して変換された前記文字列データを取得する資料提示装置。
【請求項４】
請求項１ないし請求項３のいずれか記載の資料提示装置であって、
前記変換した前記文字列データをファイルデータとして読み出し可能に記憶する文字列データ記憶部を備える資料提示装置。
【請求項５】
請求項１ないし請求項４のいずれか記載の資料提示装置であって、
前記文字列データ取得部は、前記音声データ取得部が取得した音声データに基づいて、前記音声に対応した前記所定言語とは異なる言語の文字列を文字列データとして取得する資料提示装置。
【請求項６】
請求項１ないし請求項５のいずれか記載の資料提示装置であって、
前記所定領域に含まれる被写体が変更された場合において、
前記画像合成部は、前記撮像画像データに基づいて前記被写体が変更されたことを認識可能であり、該認識後は、前記変更以前に取得した前記音声データに基づく前記文字列データは前記変更後の前記被写体を撮像した前記撮像画像データに合成しない
資料提示装置。
【請求項７】
請求項１ないし請求項５のいずれか記載の資料提示装置であって、
前記所定領域に含まれる被写体が変更された場合において、
前記画像合成部は、前記撮像画像データに基づいて前記被写体が変更されたことを認識可能であり、該認識後は、所定期間、前記変更以前に取得した前記音声データに基づく前記文字列データを、前記変更直前の前記被写体を撮像した静止画である前記撮像画像データに合成し前記合成画像データを生成する
資料提示装置。
【請求項８】
請求項１ないし請求項７のいずれか記載の資料提示装置であって、
前記画像合成部は、前記撮像画像データに基づいて前記撮像画像の余白領域を認識し、前記文字列を前記撮像画像の前記余白領域に重畳した前記合成画像に対応した合成画像データを生成する資料提示装置。
【請求項９】

請求項１ないし請求項８のいずれか記載の資料提示装置であって、
前記文字列データ取得部は、ユーザーの所定の操作によって前記文字列データを取得するか否かの切替えを行う文字列データ取得切替部を備え、
前記出力部は、前記文字列データ取得部が前記文字列データ取得切替部によって前記文字列データを取得しない場合に、前記合成画像データに替えて前記撮像画像データを出力する
資料提示装置。
【請求項１０】
請求項１ないし請求項９のいずれか記載の資料提示装置であって、
前記画像合成部は、ユーザーの所定の操作によって、前記合成画像として合成する文字列の大きさ、フォントの種類、文字数、行数、文字色、背景色、表示時間のうち少なくともひとつを制御する文字表示制御部を備える資料提示装置。
【請求項１１】
請求項１ないし請求項１０のいずれか記載の資料提示装置であって、さらに、
前記文字列データ取得部が取得した前記文字列を表す前記文字列データに基づいて、前記文字列に含まれる単語に関する情報をネットワークを介して表示可能に取得する単語情報取得部を備える資料提示装置。
【請求項１２】
請求項１ないし請求項１１のいずれか記載の資料提示装置であって、
前記撮像画像データと前記文字列データとを対応付けて読み出し可能に記憶する対応付データ記憶部を備える資料提示装置。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、資料提示装置に関するものである。
【背景技術】
【０００２】
近年、資料提示装置を活用したプレゼンテーションが広く行われている。資料提示装置に関する技術としては、例えば下記特許文献１が知られている。
【０００３】
しかし、プレゼンターの話す言葉が聞き取りにくい環境でプレゼンテーションを行う場合や、視聴者の聴覚が不自由である場合には、視聴者はプレゼンターの話す言葉の内容の理解が困難な場合があった。
【先行技術文献】
【特許文献】
【０００４】
【特許文献１】特開２０１０－２４５６９０号公報
【発明の概要】
【発明が解決しようとする課題】
【０００５】
本発明は、上述した従来の課題を解決するためになされたものであり、資料提示装置を用いたプレゼンテーションにおいて、プレゼンターの話す言葉の内容を、視聴者が容易に理解することのできる技術を提供することを目的とする。
【課題を解決するための手段】
【０００６】
本発明は、上述の課題の少なくとも一部を解決するために、以下の形態または適用例を取ることが可能である。
【０００７】
［適用例１］
資料提示装置であって、所定領域を撮像し、該撮像した撮像画像を撮像画像データとし

て取得する撮像画像データ取得部と、音声を表す音声データを外部から取得する音声データ取得部と、前記取得した音声データに基づいて、前記音声に対応した所定言語の文字列を文字列データとして取得する文字列データ取得部と、前記撮像画像データと前記文字列データとに基づいて、前記撮像画像と前記文字列とを含む合成画像を合成画像データとして生成する画像合成部と、前記合成画像データを外部に出力する出力部とを備える資料提示装置。

【０００８】

この資料提示装置によると、外部から取得した音声データを文字列データに変換し、合成画像データとして外部に出力することができる。例えば、その合成画像データを画像表示装置で表示しながらプレゼンテーションを行う場合、マイク等の集音装置により外部から集音した音声（言葉）を、資料提示装置が音声データとして取得し、その音声データを文字列データに変換し、撮像した撮像画像データと合成して合成画像データを生成し、資料提示装置と接続した画像表示装置が、合成画像としてプレゼンターが話した音声に対応する文字列を表示することができる。

【０００９】

［適用例２］

適用例１記載の資料提示装置であって、前記文字列データ取得部は、前記取得した音声データを認識し、前記所定言語の文字列データに変換する音声文字変換部を備える資料提示装置。

【００１０】

この資料提示装置によると、文字列データ取得部は音声文字変換部を備えるので、外部から音声に対応する文字列データを取得する必要がない。すなわち、音声文字変換部に対応する外部機器との接続が不要であり、資料提示装置単体で音声に対応する文字列データを取得可能である。

【００１１】

［適用例３］

適用例１記載の資料提示装置であって、前記文字列データ取得部は、前記音声データに基づいて、回線を介して変換された前記文字列データを取得する資料提示装置。

【００１２】

この資料提示装置によると、文字列データを回線を介して取得するので、内部に適用例２に示したような音声文字変換部に対応する処理部を備える必要がない。

【００１３】

［適用例４］

適用例１ないし適用例３のいずれか記載の資料提示装置であって、前記変換した前記文字列データをファイルデータとして読み出し可能に記憶する文字列データ記憶部を備える資料提示装置。

【００１４】

この資料提示装置によると、文字列データを読み出し可能なファイルデータとして記憶するので、例えば、プレゼンテーションで取得したプレゼンターの話した内容を、後に、文字列データとして利用可能となる。

【００１５】

［適用例５］

適用例１ないし適用例４のいずれか記載の資料提示装置であって、前記文字列データ取得部は、前記音声データ取得部が取得した音声データに基づいて、前記音声に対応した前記所定言語とは異なる言語の文字列を文字列データとして取得する資料提示装置。

【００１６】

この資料提示装置によると、取得した音声データに基づいて、所定言語とは異なる言語の文字列データを取得する。従って、所定言語とは異なる言語の文字列データを合成画像として表示することで、視聴者が所定言語を理解できない場合でも、所定言語とは異なるその言語を理解できる場合には、プレゼンターの話した内容を理解することができる。

【００１７】
［適用例６］
適用例１ないし適用例５のいずれか記載の資料提示装置であって、前記所定領域に含まれる被写体が変更された場合において、前記画像合成部は、前記撮像画像データに基づいて前記被写体が変更されたことを認識可能であり、該認識後は、前記変更以前に取得した前記音声データに基づく前記文字列データは前記変更後の前記被写体を撮像した前記撮像画像データに合成しない資料提示装置。
【００１８】
この資料提示装置によると、例えば、プレゼンターが変更前の被写体について話していた内容を、変更後の被写体が表示されているときに文字列として表示しないので、視聴者が映像と文字列との対応関係を理解しやすい。
【００１９】
［適用例７］
適用例１ないし適用例５のいずれか記載の資料提示装置であって、前記所定領域に含まれる被写体が変更された場合において、前記画像合成部は、所定期間、前記撮像画像データに基づいて前記被写体が変更されたことを認識可能であり、該認識後は、前記変更以前に取得した前記音声データに基づく前記文字列データを、前記変更直前の前記被写体を撮像した静止画である前記撮像画像データに合成し前記合成画像データを生成する資料提示装置。
【００２０】
この資料提示装置によると、被写体の変更以前に取得した音声データに基づく文字列データを、変更直前の被写体を撮像した静止画である撮像画像データに合成し合成画像として表示する。従って、プレゼンターが変更前の被写体について話している途中に被写体を変更したとしても、変更前の被写体について話した内容の文字列を、変更前の被写体の画像とともに視聴することができる。
【００２１】
［適用例８］
適用例１ないし適用例７のいずれか記載の資料提示装置であって、前記画像合成部は、前記撮像画像データに基づいて前記撮像画像の余白領域を認識し、前記文字列を前記撮像画像の前記余白領域に重畳した前記合成画像に対応した合成画像データを生成する資料提示装置。
【００２２】
この資料提示装置によると、文字列の表示する領域を効率よく、かつ、可能な範囲で大きく確保することが可能となり、合成画像に対して文字列を大きく表示したり、より多くの文字列による情報を表示することが可能となる。
【００２３】
［適用例９］
適用例１ないし適用例８のいずれか記載の資料提示装置であって、前記文字列データ取得部は、ユーザーの所定の操作によって前記文字列データを取得するか否かの切替えを行う文字列データ取得切替部を備え、前記出力部は、前記文字列データ取得部が前記文字列データ取得切替部によって前記文字列データを取得しない場合に、前記合成画像データに替えて前記撮像画像データを出力する資料提示装置。
【００２４】
この資料提示装置によると、ユーザ（例えばプレゼンター）が所望の音声のみ、この資料提示装置に取り込み可能である。
【００２５】
［適用例１０］
適用例１ないし適用例９のいずれか記載の資料提示装置であって、前記画像合成部は、ユーザーの所定の操作によって、前記合成画像として合成する文字列の大きさ、フォントの種類、文字数、行数、文字色、背景色、表示時間のうち少なくともひとつを制御する文

字表示制御部を備える資料提示装置。
【００２６】
この資料提示装置によると、例えば、文字列の大きさ、フォントの種類、文字数、行数、文字色、背景色、表示時間などをユーザーによる所定の操作によって制御可能であり、ユーザーが所望する文字列の表示方法で合成画像に表示可能である。
【００２７】
［適用例１１］
適用例１ないし適用例１０のいずれか記載の資料提示装置であって、さらに、前記文字列データ取得部が取得した前記文字列を表す前記文字列データに基づいて、前記文字列に含まれる単語に関する情報をネットワークを介して表示可能に取得する単語情報取得部を備える資料提示装置。
【００２８】
この資料提示装置によると、例えば、単語情報取得部が取得した情報を、合成画像データに含まれる文字列の単語にハイパーリンクとしてリンク付けすることができる。このようにすると、視聴者は、プレゼンテーションの内容をより深く理解することが可能である。
【００２９】
［適用例１２］
適用例１ないし適用例１１のいずれか記載の資料提示装置であって、前記撮像画像データと前記文字列データとを対応付けて読み出し可能に記憶する対応付データ記憶部を備える資料提示装置。
【００３０】
この資料提示装置によると、撮像画像データと文字列データとを対応付けて読み出し可能に記憶するので、例えば、文字列の表示・非表示の切り替え可能な動画のフォーマット形式で記憶することによって、視聴者はその動画データを再生してプレゼンテーションを視聴する際に、文字列の表示が不要な場合に、文字列を非表示とすることができる。
【００３１】
なお、本発明は、種々の態様で実現することが可能である。例えば、資料提示方法および装置、プレゼンテーションシステム、それらの方法または装置の機能を実現するための集積回路、コンピュータプログラム、そのコンピュータプログラムを記録した記録媒体等の形態で実現することができる。
【図面の簡単な説明】
【００３２】
【図１】資料提示システム１０の構成を説明する説明図である。
【図２】資料提示装置２０の内部構成を説明するブロック図である。
【図３】文字列表示処理の流れを説明するフローチャートである。
【図４】撮像画像データに対応する撮像画像を示している。
【図５】合成画像データに対応する合成画像を示している。
【図６】合成画像（ａ）を説明する説明図である。
【図７】合成画像（ｂ）を説明する説明図である。
【図８】合成画像（ｃ）を説明する説明図である。
【図９】合成画像（ｄ）を説明する説明図である。
【図１０】合成画像（ｅ）を説明する説明図である。
【発明を実施するための形態】
【００３３】
次に、本発明の実施の形態を実施例に基づいて説明する。
Ａ．第１実施例：
（Ａ１）資料提示システムの構成：
図１は、本発明の実施例としての資料提示システム１０の構成を説明する説明図である。資料提示システム１０は、資料提示装置２０と、プロジェクタ４０とを備える。資料提

示装置２０とプロジェクタ４０とはデータ転送用のケーブルで互いに接続されている。資料提示システム１０は、資料提示装置２０の撮像領域ＲＡに載置された資料ＲＳを資料提示装置２０が撮像し、撮像画像をプロジェクタ４０がスクリーン上の投写領域ＩＡに投写表示する。投写表示された投写資料ＩＳが資料ＲＳに対応する。また資料提示装置２０にはマイク３０が接続されており、外部からの音声、ここではプレゼンター（発表者）が話した言葉（音声）をマイク３０が集音し、集音した音声を資料提示システム１０が音声認識をして、プレゼンターが話した言葉に対応する文字列をプロジェクタ４０が投写領域ＩＡの文字列表示領域ＴＸＡに投写表示する。

【００３４】

資料提示装置２０は、机などに設置される本体２２と、本体２２に設けられた操作部２３と、本体２２から上側に伸びた支柱２４と、支柱２４の先端に取り付けられたカメラヘッド２６とを備える。カメラヘッド２６にはＣＣＤを用いたビデオカメラが内蔵されており、机などに載置された資料ＲＳを単位時間当たりに所定のフレーム数で撮像する。また資料提示装置２０には、付属のリモコン２８が備えられており、赤外線によって互いに通信を行う。ユーザーは、リモコン２８を介した操作によって、マイク３０で音声を集音するか否かの切り替え、および、文字列表示領域ＴＸＡに音声に対応する文字列を表示するか否かの切り替えを行うことができる。

【００３５】

図２は、資料提示装置２０の内部構成を説明するブロック図である。資料提示装置２０は、撮像部２１０、画像処理ユニット２２０、ＣＰＵ２３０、ＲＡＭ２４０、ハードディスク（ＨＤＤ）２５０、ＲＯＭ２６０を備える。また、資料提示装置２０は、音声入力インターフェース（音声入力ＩＦ）２７２、デジタルデータ出力インターフェース（デジタルデータ出力ＩＦ）２７６、アナログデータ出力インターフェース（アナログデータ出力ＩＦ）２７８、ＵＳＢインターフェース（ＵＳＢＩＦ）２８０、操作部２３、赤外線受信部２９を備える。

【００３６】

撮像部２１０にはレンズユニット２１２と電荷結合素子（ＣＣＤ）２１４とが備えられており、ＣＣＤ２１４は、レンズユニット２１２を透過した光を受光して電気信号に変換するイメージセンサである。画像処理ユニット２２０は、ＡＧＣ（Automatic Gain Control）回路やＤＳＰ（Digital Signal Processor）から構成され、ＣＣＤ２１４から出力される電気信号を入力し撮像画像データを生成する。画像処理ユニット２２０が生成した撮像画像データは、ＲＡＭ２４０が備える撮像画像バッファ２４２に記憶される。

【００３７】

音声入力ＩＦ２７２は、マイク３０からアナログ音声信号を受信する。アナログ／デジタル変換部（Ａ／Ｄ変換部）２７４は、音声入力ＩＦ２７２が受信したアナログ音声信号をデジタルの音声データに変換する。変換された音声データは、ＲＡＭ２４０が備える音声データバッファ２４４に記憶される。

【００３８】

ＣＰＵ２３０は、資料提示装置２０全体の動作を制御するとともに、ＲＯＭ２６０に記憶されているプログラムを読み込んで実行することにより、音声・文字変換処理部２３２、画像合成部２３４、表示設定処理部２３６として動作する。音声・文字変換処理部２３２は、音声データバッファ２４４に記憶した音声データを読み込んで認識し、日本語の文字列に対応した文字列データに変換する。変換された文字列データは、ＲＡＭ２４０が備える文字列データバッファ２４６に記憶される。音声・文字変換処理部２３２としては、例えば、アミボイス（登録商標）やビアボイス（登録商標）などの音声認識エンジンを採用することができる。本実施例では、アミボイスを採用した。また、本実施例では、音声・文字変換処理部２３２は日本語の音声データを日本語の文字列データに変換するとしたが、例えば、プレゼンターが話す言葉はフランス語である場合は、そのフランス語の音声データを認識し、フランス語の文字列に対応する文字列データに変換するとしてもよい。

【００３９】

画像合成部２３４は、撮像画像バッファ２４２に記憶されている撮像画像データと、文字列データバッファ２４６に記憶されている文字列データとを合成し、撮像画像と文字列とが含まれる合成画像データを生成する。すなわち、その合成画像がプロジェクタ４０を介してスクリーンに投写表示された際に、図１の投写領域ＩＡに表示される投写画像となるように、撮像画像データと文字列データとを合成する。画像合成部２３４によって生成された合成画像データは、ＲＡＭ２４０が備える合成画像バッファ２４８に記憶される。画像合成部２３４における処理については、後で詳しく説明する。

【００４０】

表示設定処理部２３６は、操作部２３、リモコン２８を介したユーザーからの指示に従って、投写領域ＩＡに表示する投写資料ＩＳの拡大・縮小、文字列表示領域ＴＸＡに表示される文字列の大きさ、フォントの種類、文字数、行数、文字色、背景色、表示時間の制御、および、投写領域ＩＡへの文字列表示領域ＴＸＡの表示・非表示の制御を行う。

【００４１】

デジタルデータ出力ＩＦ２７６は、合成画像バッファ２４８に記憶されている合成画像データを符号化し、デジタル信号として外部に出力する。合成画像バッファ２４８には、合成画像データを符号化する符号化処理部が含まれる。本実施例におけるデジタルデータ出力ＩＦ２７６は、外部機器との接続の規格としてＵＳＢ接続を採用する。

【００４２】

アナログデータ出力ＩＦ２７８は、合成画像バッファ２４８に記憶されている合成画像データをデジタル／アナログ変換し、ＲＧＢデータとして外部に出力する。アナログデータ出力ＩＦ２７８にはＤ／Ａコンバータ（ＤＡＣ）が含まれる。本実施例では、アナログデータ出力ＩＦ２７８にプロジェクタ４０が接続されている。

【００４３】

ＨＤＤ２５０は、大容量磁気ディスクドライブである。ＨＤＤ２５０は、音声ファイルデータ記憶部２５２、文字列ファイルデータ記憶部２５４、合成画像ファイルデータ記憶部２５６を備える。音声ファイルデータ記憶部２５２は、音声データバッファ２４４に記憶された音声データを外部に読み出し可能なファイルデータとして記憶する。文字列ファイルデータ記憶部２５４は、文字列データバッファ２４６に記憶された文字列データを外部に読み出し可能なファイルデータとして記憶する。合成画像ファイルデータ記憶部２５６は、合成画像バッファ２４８に記憶された合成画像データを外部に読み出し可能なファイルデータとして記憶する。

【００４４】

（Ａ２）文字列表示処理：

次に、資料提示システム１０が行う文字列表示処理について説明する。文字列表示処理は、マイク３０から集音した音声に対応する文字列を、撮像領域ＲＡに載置された資料ＲＳとともに、投写領域ＩＡに表示する処理である。図３は、文字列表示処理の流れを説明するフローチャートである。文字列表示処理は、操作部２３に備えられている資料提示装置２０の電源をユーザーがＯＮにすることによって開始される。文字列表示処理が開始されると、ＣＰＵ２３０は、撮像部２１０および画像処理ユニット２２０が生成した撮像画像データを取得し、撮像画像バッファ２４２に記憶する（ステップＳ１０２）。

【００４５】

次に、ＣＰＵ２３０は、マイク３０から音声入力ＩＦ２７２およびＡ／Ｄ変換部２７４を介して、プレゼンターが話した言葉を音声データとして取得し音声データバッファ２４４に記憶する（ステップＳ１０４）。ＣＰＵ２３０は、取得した音声データを読み込み、音声・文字変換処理部２３２の機能として音声認識エンジンを用いて、音声データを日本語の文字列のデータに変換し、変換後の文字列データを文字列データバッファ２４６に記憶する（ステップＡ１０６）。音声・文字変換終了後、ＣＰＵ２３０は、画像合成処理を行う（ステップＳ１０８）。具体的には、撮像画像バッファ２４２および文字列データバッファ２４６から撮像画像データおよび文字列データを読み込み、２つのデータを合成し合成画像データを生成する。

【００４６】
図４は、撮像画像データに対応する撮像画像を示している。図５は、合成画像データに対応する合成画像を示している。ＣＰＵ２３０は、画像合成処理によって、文字列表示領域ＴＸＡ（図１参照）に対応する無地の画像に文字列を重畳し、画像データ（文字列画像データＴＸＤ）を生成する。その後、図５に示すように、撮像画像データの下方に文字列画像データＴＸＤを重畳し、合成画像データを生成する。この時、操作部２３を介したユーザーからの指示に従って、表示設定処理部２３６が、文字列のフォントの種類、大きさ、色など制御を行い、表示設定処理部２３６によって処理された文字列を、画像合成部２３４が文字列表示領域ＴＸＡに対応する無地の画像に重畳し文字列画像データＴＸＤを生成し、その後、合成画像データを生成する。このような処理として、例えば、一般にＯＳＤ（on screen display)に利用される技術を利用することができる。

【００４７】
画像合成処理後、ＣＰＵ２３０は、生成した合成画像データを合成画像バッファ２４８に記憶し、順次、アナログデータ出力ＩＦ２７８を介して、ＲＧＢデータに変換された合成画像データをプロジェクタ４０に出力する（ステップＳ１１０）。ＣＰＵ２３０は、これらの処理（ステップＳ１０２～ステップＳ１１０）を、ユーザーが資料提示装置２０の電源をＯＦＦにするまで繰返し行う（ステップＳ１１２）。なお、ユーザーがリモコン２８を介して、投写領域ＩＡへの文字列の非表示を指示した場合には、ＣＰＵ２３０は、合成画像データに替えて、撮像画像バッファ２４２に記憶した撮像画像データをアナログデータ出力ＩＦ２７８またはデジタルデータ出力ＩＦ２７６から出力する。

【００４８】
また、ＣＰＵ２３０は、文字列表示処理以外に、文字列表示処理の間に取得した音声データ、文字列データ、合成画像データを読み出し可能なファイルデータとして、ＨＤＤ２５０に記憶する処理を行う。具体的には、ＣＰＵ２３０は、音声ファイルデータは音声ファイルデータ記憶部２５２に、文字列ファイルデータは文字列ファイルデータ記憶部２５４に、合成画像ファイルデータは合成画像ファイルデータ記憶部２５６にそれぞれ記憶する。例えば、ＣＰＵ２３０は、音声データファイルはＷＭＡやＭＰ３やＡＡＣ等の音声ファイルのフォーマット形式で、文字列ファイルデータはＴＸＴやＤＯＣ等のテキストファイルのフォーマット形式で、合成画像ファイルデータはＭＰＧやＡＶＩやＷＭＶ等の動画や静止画のフォーマット形式でそれぞれＨＤＤ２５０に記憶する。本実施例においては、これらファイルデータは、ＵＳＢＩＦ２８０を介して接続したコンピュータやハードディスク、ＳＳＤ（Solid State Drive）等の記憶装置に読み出し可能である。

【００４９】
本実施例では、音声入力ＩＦ２７２に接続されたマイク３０から音声信号を受信するとしたが、例えば、音声入力ＩＦ２７２に接続されたＭＰ３プレーヤ、ｉＰｏｄ（登録商標）、テープレコーダ、ＭＤプレーヤ等の音声出力器から音声を受信するとしてもよい。また、本実施例では、合成画像データをプロジェクタ４０に出力してプロジェクタ４０がスクリーンに合成画像を投写表示するとしたが、デジタルデータ出力ＩＦ２７６またはアナログデータ出力ＩＦ２７８に接続したテレビや、コンピュータに接続されているディスプレイ等の画像表示装置に合成画像データを出力して、その画像表示装置によって合成画像を表示するとしてもよい。また、資料提示装置２０の音声出力インターフェースを設けて、スピーカを接続し、音声入力ＩＦ２７２を介して受信した音声信号を、そのスピーカから音声として出力するとしてもよい。

【００５０】
さらに、資料提示装置２０は、撮像領域ＲＡに載置されている被写体（資料）が変更された場合に、その変更を認識して、変更前に取得した音声データに対応する文字列データを変更後の撮像画像データに合成しないとしてもよい。具体的には、ＣＰＵ２３０が画像合成部２３４として画像合成処理を行う際に、随時処理に供する撮像画像データの輝度の変化を認識し、撮像画像データにおいて所定以上の領域（面積）で、所定以上の輝度の変化を認識した際には、撮像領域ＲＡ（図１参照）に載置されている資料ＲＳが変更された

と認識する。そして、仮に、予め音声・文字変換後の文字列を所定時間以上、文字列表示領域ＴＸＡに表示するように設定したとしても、資料ＲＳの変更を認識した際には、資料ＲＳの変更を認識する以前に取得した文字列データについては、その所定時間より短い時間であっても表示を停止し、変更後の資料ＲＳを投写領域ＩＡに表示する時点では、変更前の資料ＲＳは表示しないとしてもよい。すなわち、ＣＰＵ２３０が撮像画像データに基づいて資料ＲＳが変更されたことを認識し、認識後は、資料ＲＳの変更以前に取得した音声データに基づく文字列データは、変更後の資料ＲＳを撮像した撮像データには画像合成しない。

【００５１】

その他、ＣＰＵ２３０が資料ＲＳの変更を認識した場合には、資料ＲＳの変更以前に取得した音声データに基づく文字列データを、資料の変更直前の資料ＲＳを撮像した静止画である撮像画像データに画像合成するとしてもよい。そして、資料ＲＳの変更以前に取得した音声データに基づく文字列データを全て表示し終えるまでは、その静止画像に対応する撮像画像データを画像合成処理に用いる。

【００５２】

以上説明したように、資料提示システム１０は、プレゼンターが話した言葉を音声認識して、認識した音声を文字列として、投写領域ＩＡの文字列表示領域ＴＸＡに表示する。よって、プレゼンターの話す言葉が聞き取りにくい環境でプレゼンテーションを行う場合や、視聴者の聴覚が不自由である場合であっても、視聴者が文字列表示領域ＴＸＡに表示された文字列を視認することによって、プレゼンターの話す言葉の内容を容易に理解することができる。また、プレゼンテーションにおいて学術用語や専門用語が使われ、それらの用語を視聴者が知らない場合やなじみの無い場合に、文字列、特に漢字で表示されれば、視聴者はそれら用語の意味を理解しやすくなる。

【００５３】

また、ＣＰＵ２３０は、ＨＤＤ２５０の音声ファイルデータ記憶部２５２、文字列ファイルデータ記憶部２５４、合成画像ファイルデータ記憶部２５６に、音声ファイルデータ、文字列ファイルデータ、合成画像データの各ファイルデータを読み出し可能に記憶するので、プレゼンターのプレゼンテーションを直接視聴しなかった者であっても、各ファイルデータを閲覧または再生することによって、プレゼンテーションを視聴することができる。

【００５４】

さらに、資料提示装置２０を用いてプロジェクタ４０等の画像表示装置に画像を表示させる場合は、資料提示装置２０と画像表示装置との間に、所定の演算処理を行うコンピュータを接続するのが通常であるが、本実施例における資料提示システム１０はコンピュータを必要としない。よって、ユーザーは簡易に資料提示装置を用いてプレゼンテーションを行うことが可能である。

【００５５】

Ｂ．変形例：

なお、この発明は上記の実施例や実施形態に限られるものではなく、その要旨を逸脱しない範囲において種々の態様において実施することが可能であり、例えば次のような変形も可能である。

【００５６】

（Ｂ１）変形例１：

上記実施例では、資料提示装置２０が音声認識エンジンとしての音声・文字変換処理部２３２（例えば、アミボイスやビアボイス）を備え、ＣＰＵ２３０が音声データを文字列データに変換する処理を行うとしたが、例えば、資料提示装置２０がネットワーク接続可能であり、ネットワーク上のサーバやコンピュータに音声データを送信し、サーバやコンピュータが備える音声認識エンジンで音声・文字変換処理を行い、変換後の文字列データをネットワークを介して資料提示装置２０が取得するとしてもよい。また、文字列データの取得方法としては、ネットワーク上のサーバやコンピュータからの取得に限らず、ＵＳ

ＢケーブルやＬＡＮケーブル等による回線によって、直接的に、音声認識エンジンを搭載したコンピュータと資料提示装置２０とを接続する構成としてもよい。そして、資料提示装置２０が取得した音声データを、そのコンピュータに送信し、そのコンピュータの音声認識エンジンが変換した文字列データを、その回線を介して取得するものとしてもよい。このようにすることで、資料提示装置２０が音声・文字変換処理部２３２（音声認識エンジン）を備える必要がない。また、ネットワーク上の音声認識エンジンを利用することで、最新の音声認識エンジンによって変換された文字列データを取得することができ、音声データから文字列データへの変換精度を向上させることができる。

【００５７】

（Ｂ２）変形例２：

上記実施例では、文字列データを所定の言語の文字列（上記実施例では日本語）に変換し、文字列表示領域ＴＸＡには日本語の文字列のみを表示するとしたが、所定の言語の文字列に加え、その音声データを翻訳した他の言語の文字列（以下、異言語文字列とも呼ぶ）を表示するとしてもよい。具体的には、資料提示装置２０が、翻訳エンジンとして、例えば、Ｇｏｏｇｌｅ翻訳（Ｇｏｏｇｌｅは登録商標）やエキサイト翻訳（エキサイトは登録商標）などに用いられている翻訳エンジンを搭載し、文字列データバッファ２４６に記憶した所定言語（例えば日本語）の文字列データに基づいて、フランス語、英語、中国語、スペイン語、ポルトガル語、ヒンディー語、ロシア語、ドイツ語、アラビア語、韓国語など、所定言語とは異なる言語に翻訳した文字列の文字列データを取得し、図６の合成画像（ａ）に示すように、異言語の文字列を所定言語と並べて、または独立して、文字列表示領域ＴＸＡに表示するとしてもよい。

【００５８】

また、資料提示装置２０がネットワーク接続可能であり、ネットワーク上のサーバやコンピュータに所定言語の文字列データを送信し、サーバやコンピュータが備える翻訳エンジンで翻訳処理を行い、翻訳後の異言語の文字列データをネットワークを介して資料提示装置２０が取得するとしてもよい。その他、ネットワーク上のサーバやコンピュータによる翻訳処理に限らず、ＵＳＢケーブルやＬＡＮケーブル等による回線によって、直接的に、翻訳エンジンを搭載したコンピュータと資料提示装置２０とを接続するとしてもよい。そして、資料提示装置２０が取得した所定言語の文字列データを、そのコンピュータに送信し、そのコンピュータの翻訳エンジンが翻訳した異言語の文字列データを、その回線を介して取得するとしてもよい。さらに、ユーザーが予め、プレゼンテーションをする分野（例えば、医学、政治・経済、工学、社会学など）を資料提示装置２０に設定可能とし、設定された分野を専門とする翻訳エンジンを資料提示装置２０内またはネットワーク上の、複数の分野の翻訳エンジンから選択して用いるとしてもよい。このようにすることで、１つのプレゼンテーションを、多数の国、地域、人種の人が視聴でき、プレゼンテーションの内容を理解することができる。また、ネットワーク上の翻訳エンジンを利用することで、常に、最新の翻訳エンジンによって翻訳された異言語文字列データを取得することができ、翻訳の精度を向上させることができる。

【００５９】

（Ｂ３）変形例３：

上記実施例では、合成画像データを、プロジェクタ４０を介して視聴者が視聴するとしたが、資料提示装置２０に回線（ネットワークを含む）を介して接続されたコンピュータや、地上デジタル放送対応のテレビを用いて視聴するとしてもよい。その際、文字列表示領域ＴＸＡに表示する文字列のうち、例えば主要な単語に対して、その単語に関する説明が掲載されたネットワーク上のホームページ、例えばWikipedia（登録商標）のホームページをハイパーリンクによってリンク付けし、視聴者がその単語に関する情報を取得可能としてもよい。その際、図７の合成画像（ｂ）に示すように、視聴者がコンピュータのディスプレイを用いてプレゼンテーションをしている際、ハイパーリンクを施した文字列の単語に下線を付し、視聴者がポインティングデバイス（例えばマウス）を用いてその下線部にカーソルを合わせた際に、その単語に関する情報をポップアップとして表示するとし

てもよい。このようにすることで、視聴者は、プレゼンテーションの内容をより深く理解することが可能となる。

【００６０】

（Ｂ４）変形例４：

上記実施例では、ＣＰＵ２３０は、画像合成処理によって、文字列を撮像画像の下方に表示するように画像合成したが（図４，図５参照）、それに限らず、図８の合成画像（ｃ）、および図９の合成画像（ｄ）に示すように、撮像画像の領域内の被写体（上記実施例では資料ＲＳ）が写っている領域以外の領域（以下、余白領域とも呼ぶ）に文字列を合成するとしてもよい。具体的には、ＣＰＵ２３０が、撮像画像データに対して、画像処理におけるラベリング処理を行うことにより、余白領域を認識可能である。撮像画像データを、各画素毎に所定の輝度を基準として二値化し、所定の輝度以上の連続する画素に同じ番号を割り振ることにより、余白領域を認識することができる。このようにすることで、文字列の表示領域を効率よく、かつ、可能な範囲で大きく確保することが可能となり、文字列を大きく表示したり、より多くの文字列による情報を表示することが可能となる。

【００６１】

（Ｂ５）変形例５：

上記実施例では、画像合成処理によって、文字列表示領域ＴＸＡ（図１参照）に対応する無地の画像に、文字列を重畳した画像データ（文字列画像データＴＸＤ）を生成し、その後、撮像画像データに文字列画像データを重畳し、合成画像データを生成するとしたが、それに限ることなく、図１０の合成画像（ｅ）に示すように、文字列を直接的に撮像画像に重畳し合成画像データを生成するとしてもよい。この他、文字列にシャドー効果や枠線を付与してもよい。このようにしても、上記実施例と同様の効果を得ることができる。

【００６２】

（Ｂ６）変形例６：

上記実施例では、ＣＰＵ２３０は、ファイルデータとして、音声ファイルデータ、文字列ファイルデータ、合成画像ファイルデータをＨＤＤ２５０に記憶するとしたが、それに限ることなく、動画としての撮像画像データに、文字列データを経時的に対応付けした動画ファイルデータを生成し、ＨＤＤ２５０に読み出し可能に記憶するとしてもよい。具体的には、動画再生時に文字列の表示・非表示を選択可能な動画のフォーマット形式で動画ファイルデータを生成しＨＤＤ２５０に記憶する。この動画ファイルデータを記憶するＨＤＤ２５０が、特許請求の範囲に記載の対応付データ記憶部に対応する。このような動画ファイルデータを生成することによって、視聴者は、上記実施例による効果に加え、文字列表示が不要な際には非表示とすることができる。また、この動画ファイルデータをＤＶＤやブルーレイディスク等の記録媒体に書き込み、頒布することが可能である。

【００６３】

（Ｂ７）変形例７：

上記実施例では、音声認識エンジンを用いて音声認識を行なっている。近年、高い音声認識率を実現している音声認識エンジンは、n-gram等の言語モデルを用いるが、この場合、各単語には共起性の情報が予め設定されている。そこで、ビデオカメラで撮像した資料に存在する文字列を、ＯＣＲ技術を用いて認識し、得られた文字列を形態素解析して単語群とし、この単語群を、音声認識の実施に先立って、音声認識エンジンに与える構成を採用することができる。音声認識エンジンは、与えられた単語群を認識したものとして、これらの単語群との共起が想定された単語群の認識がされやすい状況を作っておけばよい。こうすれば、プレゼンターによる話の冒頭部分の認識率の低下を防止して、全体の認識率を高めることができる。もとより、言語モデルとして、文脈自由文法が採用されている場合には、与えられた単語群により文脈を特定するものとすればよい。

【００６４】

通常のプレゼンテーションにおいて、資料に記載された文字列は、プレゼンターが説明しようとする内容と強い相関を持ち、多くの場合、話の内容が属する分野を最も端的に表わしている語群からなっている。従って、資料を更新する度に、資料に含まれる文字列を

認識して、これを形態素解析により単語群に変換してから、音声認識エンジンに出力するよう構成すれば、認識率を恒常的に高めることが可能となる。

【００６５】

音響モデルと言語モデルのいずれを用いるかにかかわらず、音声認識率を高めるために、医学分野、芸術分野など、専門性の高い分野については、専門の辞書を用意し、プレゼンターが、音声認識を実施しようとする分野を特定し、音声認識エンジンの設定（例えば、使用する辞書の変更）を行なっていることも少なくない。こうした場合に、上述したように、資料に含まれる文字列から語群を認識して音声認識エンジンに出力するのであれば、プレゼンターがいちいち、認識させようとする分野を特定して設定を手動で変更する必要がなく、優れた使い勝手を実現することができる。

【００６６】

（Ｂ８）変形例８：

上記実施例においてソフトウェアで実現されている機能の一部をハードウェアで実現してもよく、あるいは、ハードウェアで実現されている機能の一部をソフトウェアで実現してもよい。

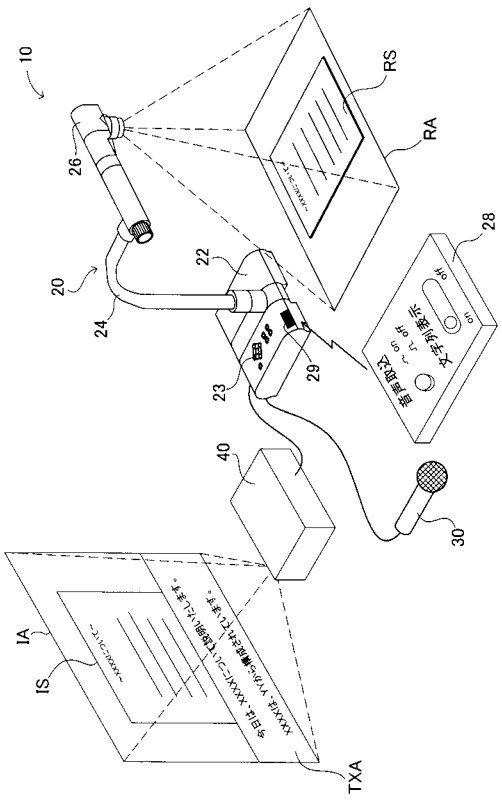
【符号の説明】

【００６７】

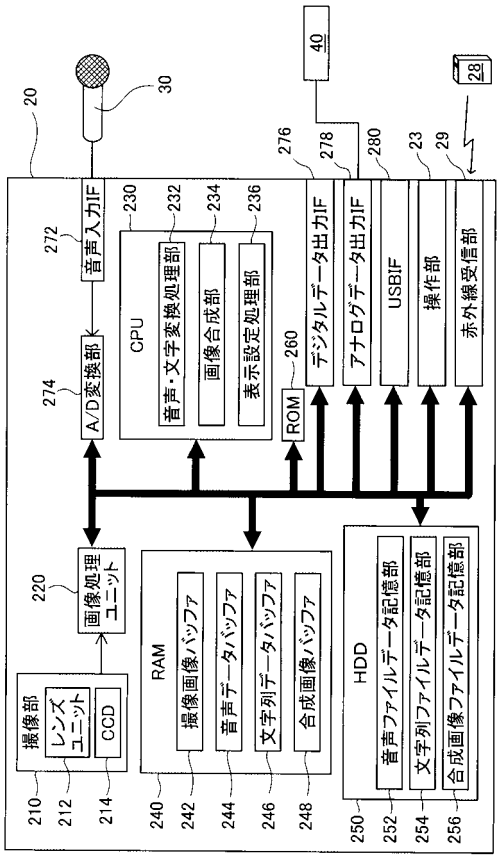
１０…資料提示システム
２０…資料提示装置
２２…本体
２３…操作部
２４…支柱
２６…カメラヘッド
２８…リモコン
２９…赤外線受信部
３０…マイク
４０…プロジェクタ
２１０…撮像部
２１２…レンズユニット
２１４…ＣＣＤ
２２０…画像処理ユニット
２３０…ＣＰＵ
２３４…画像合成部
２３６…表示設定処理部
２４０…ＲＡＭ
２４２…撮像画像バッファ
２４４…音声データバッファ
２４６…文字列データバッファ
２４８…合成画像バッファ
２５２…音声ファイルデータ記憶部
２５４…文字列ファイルデータ記憶部
２５６…合成画像ファイルデータ記憶部
２７８…アナログデータ出力部
２７２…音声入力ＩＦ
２７６…デジタルデータ出力ＩＦ
２８０…ＵＳＢＩＦ
ＲＡ…撮像領域
ＩＡ…投写領域
ＲＳ…資料
ＩＳ…投写資料

ＴＸＡ…文字列表示領域
ＴＸＤ…文字列画像データ

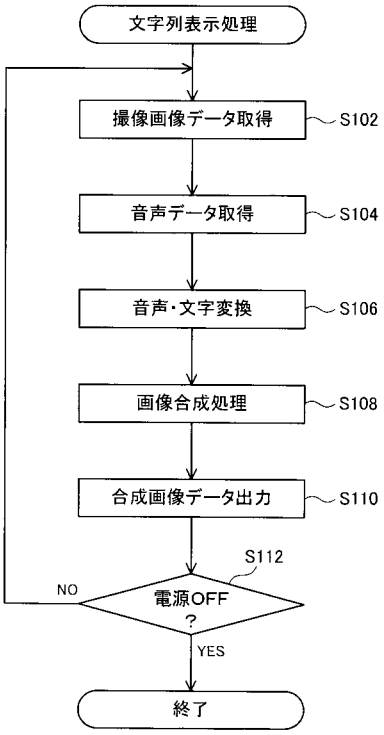
【図１】



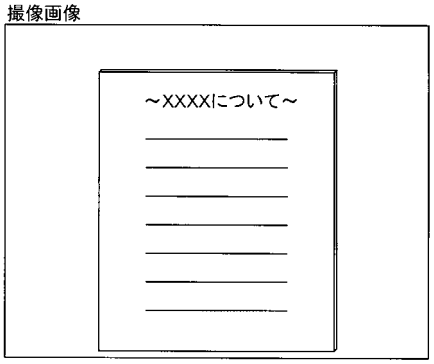

【図２】



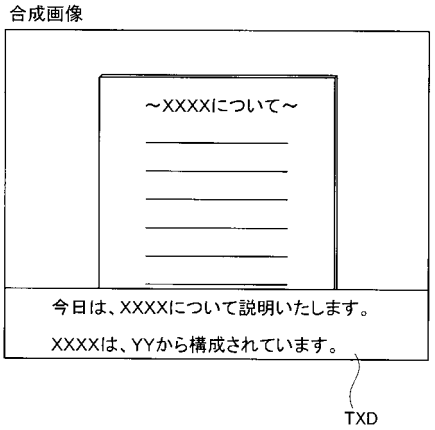
【図３】



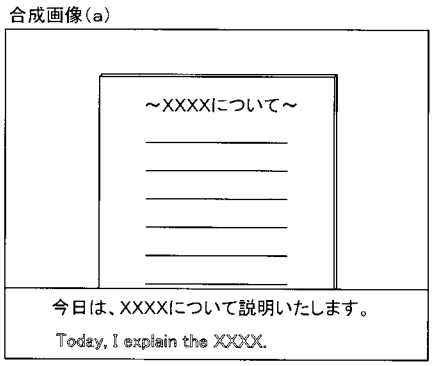

【図４】



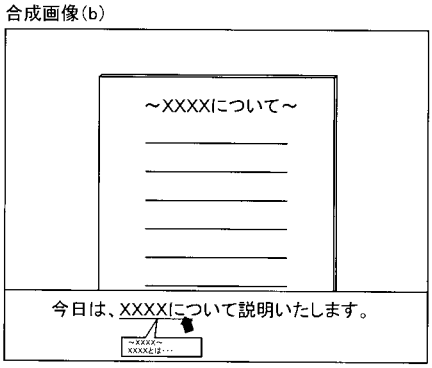

【図５】



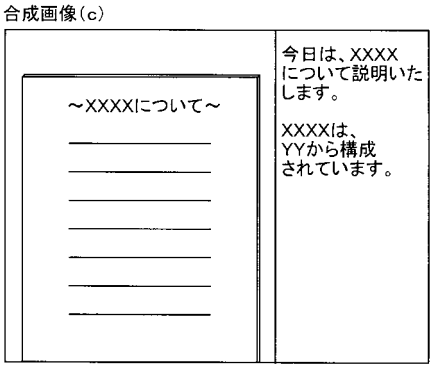

【図６】



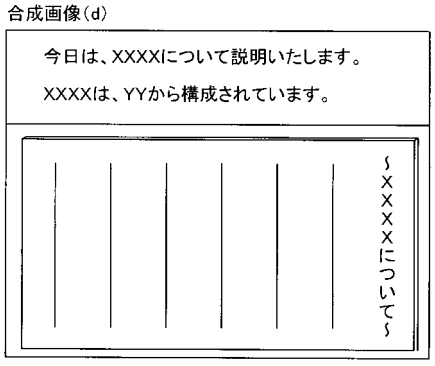
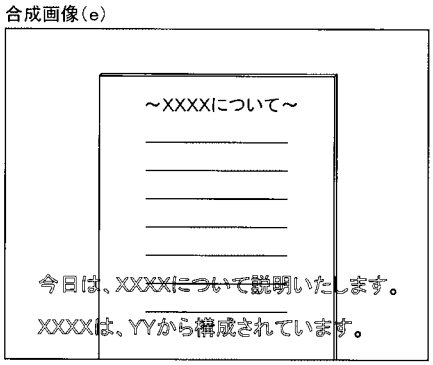

【図７】




【図８】




【図９】



【図１０】