

(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2019-204503

(P2019-204503A)

(43) 公開日 令和1年11月28日 (2019.11.28)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 3/16 (2006.01)	G06F 3/16 610	2H040
G02B 21/36 (2006.01)	G02B 21/36	2H052
G02B 23/24 (2006.01)	G02B 23/24 B	4C161
H04N 5/232 (2006.01)	H04N 5/232 300	4C316
H04N 5/225 (2006.01)	H04N 5/225 500	5B087
審査請求 未請求 請求項の数 18 O L (全 34 頁) 最終頁に続く		

(21) 出願番号 特願2019-88251 (P2019-88251)
 (22) 出願日 令和1年5月8日 (2019.5.8)
 (31) 優先権主張番号 特願2018-95519 (P2018-95519)
 (32) 優先日 平成30年5月17日 (2018.5.17)
 (33) 優先権主張国・地域又は機関
 日本国 (JP)

(71) 出願人 000000376
 オリンパス株式会社
 東京都八王子市石川町2951番地
 (74) 代理人 110002147
 特許業務法人酒井国際特許事務所
 (72) 発明者 堀内 一仁
 東京都八王子市石川町2951番地 オリ
 ンパス株式会社内
 (72) 発明者 渡辺 伸之
 東京都八王子市石川町2951番地 オリ
 ンパス株式会社内
 (72) 発明者 金子 善興
 東京都八王子市石川町2951番地 オリ
 ンパス株式会社内

最終頁に続く

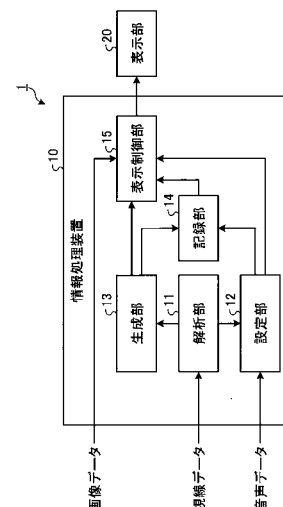
(54) 【発明の名称】 情報処理装置、情報処理方法およびプログラム

(57) 【要約】

【課題】 音声データのどの期間が重要であるか否かを把握することができる情報処理装置、情報処理方法およびプログラムを提供する。

【解決手段】 情報処理装置10は、外部から入力される利用者の視線を検出した視線データに基づいて、利用者の視線の注視度を解析する解析部11と、外部から入力される利用者の音声データであって、視線データと同じ時間軸が対応付けられた音声データに対して、解析部11が解析した注視度に応じた重要度を所定の時間間隔で割り当てて記録部14へ記録する設定部12と、を備える。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

外部から入力される利用者の視線を検出した視線データに基づいて、前記利用者の視線の注視度を解析する解析部と、

外部から入力される前記利用者の音声データであって、前記視線データと同じ時間軸が対応付けられた音声データに対して、前記解析部が解析した前記注視度に応じた重要度を割り当てて記録部へ記録する設定部と、

を備える情報処理装置。

【請求項 2】

前記設定部は、

前記解析部によって前記注視度が高いと解析された期間の時刻において前記利用者が発話した前記音声データの期間を重要期間と割り当てて前記記録部に記録する

請求項 1 に記載の情報処理装置。

10

【請求項 3】

前記設定部は、

前記注視度が高いほど前記音声データに対して前記重要度を高く割り当てて前記記録部へ記録する

請求項 1 または 2 に記載の情報処理装置。

【請求項 4】

前記設定部は、

前記解析部によって前記注視度が高いと解析された期間の時刻において前記利用者が発話した場合、該期間の時刻において前記利用者が発話した前記音声データに前記重要度を割り当てて前記記録部へ記録する

請求項 1 に記載の情報処理装置。

20

【請求項 5】

前記解析部は、

前記視線の移動速度、一定の時間内における前記視線の移動距離、一定領域内における前記視線の滞留時間のいずれか 1 つを検出することによって、前記注視度を解析する

請求項 1 ~ 4 のいずれか一つに記載の情報処理装置。

【請求項 6】

外部から入力される画像データに対応する画像上に前記解析部が解析した前記注視度および該注視度の座標情報を関連付けた視線マッピングデータを生成する生成部をさらに備える

請求項 1 ~ 5 のいずれか一つに記載の情報処理装置。

30

【請求項 7】

前記解析部は、前記視線データに基づいて、前記利用者の視線の軌跡をさらに解析し、

前記生成部は、前記解析部が解析した前記軌跡を前記画像上にさらに関連付けて前記視線マッピングデータを生成する

請求項 6 に記載の情報処理装置。

【請求項 8】

前記音声データを文字情報に変換する変換部をさらに備え、

前記生成部は、前記座標情報に前記文字情報をさらに関連付けて前記視線マッピングデータを生成する

請求項 6 または 7 に記載の情報処理装置。

40

【請求項 9】

外部から入力された操作信号に応じて指定されたキーワードを前記変換部によって変換された前記文字情報から抽出する抽出部と、

前記視線マッピングデータに対応する視線マッピング画像を表示部に表示させる表示制御部と、

をさらに備え、

50

前記表示制御部は、前記抽出部によって抽出された前記文字情報に関連付けられた前記注視度を前記表示部に強調表示させ、かつ、前記文字情報を前記表示部に表示させる請求項 8 に記載の情報処理装置。

【請求項 10】

前記利用者の視線を連続的に検出することによって前記視線データを生成する視線検出部と、

前記利用者の音声の入力を受け付けて前記音声データを生成する音声入力部と、
をさらに備える

請求項 1 ～ 9 のいずれか一つに記載の情報処理装置。

【請求項 11】

複数の利用者の各々を識別する識別情報を検出する検出部をさらに備え、
前記解析部は、前記複数の利用者の各々の視線を検出した複数の前記視線データに基づいて、前記複数の利用者の各々の前記注視度を解析し、

前記設定部は、前記検出部が検出した前記識別情報に基づいて、前記複数の利用者の各々の前記音声データに所定の時間間隔で割り当てる前記重要度を変更する

請求項 1 ～ 10 のいずれか一つに記載の情報処理装置。

【請求項 12】

標本を観察する観察倍率を変更可能であり、前記利用者が前記標本の観察像を観察可能な接眼部を有する顕微鏡と、

前記顕微鏡に接続され、前記顕微鏡が結像した前記標本の観察像を撮像することによって画像データを生成する撮像部をさらに備え、

前記視線検出部は、前記顕微鏡の接眼部に設けられ、

前記設定部は、前記観察倍率に応じて前記重要度の重み付けを行う

請求項 10 に記載の情報処理装置。

【請求項 13】

被検体に挿入可能な挿入部の先端部に設けられ、被検体内の体内を撮像することによって画像データを生成する撮像部と、視野を変更するための各種の操作の入力を受け付ける操作部と、を有する内視鏡と、

をさらに備える

請求項 1 ～ 11 のいずれか一つに記載の情報処理装置。

【請求項 14】

前記設定部は、前記操作部が受け付けた操作履歴に関する操作履歴に基づいて前記重要度の重み付けを行う

請求項 13 に記載の情報処理装置。

【請求項 15】

前記設定部は、前記解析部が解析した前記注視度が高いと判別した期間を基準にして算出した期間の前記音声データに対して、前記重要度を割り当てる

請求項 1 に記載の情報処理装置。

【請求項 16】

前記情報処理装置は、前記音声データに対して発話された発話期間を特定する発話検出部をさらに備え、

前記設定部は、前記解析部が解析した前記注視度が高いと判別した期間を基準にして、最も時間相関が高い発話期間に対して、重要度を割り当てる期間を設定することを特徴とする

請求項 1 に記載の情報処理装置。

【請求項 17】

情報処理装置が実行する情報処理方法であって、

外部から入力される利用者の視線を検出した視線データに基づいて、前記利用者の視線の注視度を解析する解析ステップと、

外部から入力される前記利用者の音声データであって、前記視線データと同じ時間軸が

10

20

30

40

50

対応付けられた音声データに対して、前記解析ステップで解析した前記注視度に応じた重要度を割り当てて記録部へ記録する設定ステップと、

を含むことを特徴とする情報処理方法。

【請求項 18】

情報処理装置に、

外部から入力される利用者の視線を検出した視線データに基づいて、前記利用者の視線の注視度を解析する解析ステップと、

外部から入力される前記利用者の音声データであって、前記視線データと同じ時間軸が対応付けられた音声データに対して、前記解析ステップで解析した前記注視度に応じた重要度を割り当てて記録部へ記録する設定ステップと、

を実行させることを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声データと視線データとを処理する情報処理装置、情報処理方法およびプログラムに関する。

【背景技術】

【0002】

近年、画像データ等の情報を処理する情報処理装置において、利用者の音声を検出したときから、所定の時間遡った期間において、表示部が表示する画像上における複数の表示領域に対して、利用者の視線が最も長く停留した画像の表示領域を注目情報として検出するとともに、この注目情報と音声とを対応付けて記録する技術が知られている（特許文献1参照）。

【0003】

また、注視注釈システムにおいて、コンピューティングデバイスの表示デバイスが表示する画像に対して、注視追跡デバイスによって検出された利用者が注視する注視点の近くに注釈アンカーを表示するとともに、この注釈アンカーに音声によって情報を入力する技術が知られている（特許文献2参照）。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特許第4282343号公報

【特許文献2】特開2016-181245号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、上述した特許文献1,2では、利用者の視線を音声入力によって行うためのキューとして利用しているにすぎないため、音声データの重要度について何ら考慮されておらず、音声データのどの期間が重要であるか否かを把握することができないという問題点があった。

【0006】

本開示は、上記に鑑みてなされたものであって、音声データの重要な期間を把握することができる情報処理装置、情報処理方法およびプログラムを提供することを目的とする。

【課題を解決するための手段】

【0007】

上述した課題を解決し、目的を達成するために、本開示に係る情報処理装置は、外部から入力される利用者の視線を検出した視線データに基づいて、前記利用者の視線の注視度を解析する解析部と、外部から入力される前記利用者の音声データであって、前記視線データと同じ時間軸が対応付けられた音声データに対して、前記解析部が解析した前記注視度に応じた重要度を割り当てて記録部へ記録する設定部と、を備える。

10

20

30

40

50

【 0 0 0 8 】

また、本開示に係る情報処理装置は、上記開示において、前記設定部は、前記解析部によって前記注視度が高いと解析された期間の時刻において前記利用者が発話した前記音声データの期間を重要期間と割り当てて前記記録部に記録する。

【 0 0 0 9 】

また、本開示に係る情報処理装置は、上記開示において、前記設定部は、前記注視度が高いほど前記音声データに対して前記重要度を高く割り当てて前記記録部へ記録する。

【 0 0 1 0 】

また、本開示に係る情報処理装置は、上記開示において、前記設定部は、前記解析部によって前記注視度が高いと解析された期間の時刻において前記利用者が発話した場合、該期間の時刻において前記利用者が発話した前記音声データに前記重要度を割り当てて前記記録部へ記録する。

10

【 0 0 1 1 】

また、本開示に係る情報処理装置は、上記開示において、前記解析部は、前記視線の移動速度、一定の時間内における前記視線の移動距離、一定領域内における前記視線の滞留時間のいずれか1つを検出することによって、前記注視度を解析する。

【 0 0 1 2 】

また、本開示に係る情報処理装置は、上記開示において、外部から入力される画像データに対応する画像上に前記解析部が解析した前記注視度および該注視度の座標情報を関連付けた視線マッピングデータを生成する生成部をさらに備える。

20

【 0 0 1 3 】

また、本開示に係る情報処理装置は、上記開示において、前記解析部は、前記視線データに基づいて、前記利用者の視線の軌跡をさらに解析し、前記生成部は、前記解析部が解析した前記軌跡を前記画像上にさらに関連付けて前記視線マッピングデータを生成する。

【 0 0 1 4 】

また、本開示に係る情報処理装置は、上記開示において、前記音声データを文字情報に変換する変換部をさらに備え、前記生成部は、前記座標情報に前記文字情報をさらに関連付けて前記視線マッピングデータを生成する。

【 0 0 1 5 】

また、本開示に係る情報処理装置は、上記開示において、外部から入力された操作信号に応じて指定されたキーワードを前記変換部によって変換された前記文字情報から抽出する抽出部と、前記視線マッピングデータに対応する視線マッピング画像を表示部に表示させる表示制御部と、をさらに備え、前記表示制御部は、前記抽出部によって抽出された前記文字情報に関連付けられた前記注視度を前記表示部に強調表示させ、かつ、前記文字情報を前記表示部に表示させる。

30

【 0 0 1 6 】

また、本開示に係る情報処理装置は、上記開示において、前記利用者の視線を連続的に検出することによって前記視線データを生成する視線検出部と、前記利用者の音声の入力を受け付けて前記音声データを生成する音声入力部と、をさらに備える。

【 0 0 1 7 】

40

また、本開示に係る情報処理装置は、上記開示において、複数の利用者の各々を識別する識別情報を検出する検出部をさらに備え、前記解析部は、前記複数の利用者の各々の視線を検出した複数の前記視線データに基づいて、前記複数の利用者の各々の前記注視度を解析し、前記設定部は、前記検出部が検出した前記識別情報に基づいて、前記複数の利用者の各々の前記音声データに所定の時間間隔で割り当てる前記重要度を変更する。

【 0 0 1 8 】

また、本開示に係る情報処理装置は、上記開示において、標本を観察する観察倍率を変更可能であり、前記利用者が前記標本の観察像を観察可能な接眼部を有する顕微鏡と、前記顕微鏡に接続され、前記顕微鏡が結像した前記標本の観察像を撮像することによって画像データを生成する撮像部とをさらに備え、前記視線検出部は、前記顕微鏡の接眼部に設

50

けられ、前記設定部は、前記観察倍率に応じて前記重要度の重み付けを行う。

【0019】

また、本開示に係る情報処理装置は、上記開示において、被検体に挿入可能な挿入部の先端部に設けられ、被検体内の体内を撮像することによって画像データを生成する撮像部と、視野を変更するための各種の操作の入力を受け付ける操作部と、を有する内視鏡と、をさらに備える。

【0020】

また、本開示に係る情報処理装置は、上記開示において、前記設定部は、前記操作部が受け付けた操作履歴に関する操作履歴に基づいて前記重要度の重み付けを行う。

【0021】

また、本開示に係る情報処理装置は、上記開示において、前記設定部は、前記解析部が解析した前記注視度が高いと判別した期間を基準にして算出した期間の前記音声データに対して、重要度を割り当てる。

【0022】

また、本開示に係る情報処理装置は、上記開示において、前記情報処理装置は、さらに、前記音声データに対して発話された発話期間を特定する発話検出部を有し、前記設定部は、前記解析部が解析した前記注視度が高いと判別した期間を基準にして、最も時間相関が高い発話期間に対して、重要度を割り当てる期間を設定することを特徴とする。

【0023】

また、本開示に係る情報処理方法は、情報処理装置が実行する情報処理方法であって、外部から入力される利用者の視線を検出した視線データに基づいて、前記利用者の視線の注視度を解析する解析ステップと、外部から入力される前記利用者の音声データであって、前記視線データと同じ時間軸が対応付けられた音声データに対して、前記解析ステップで解析した前記注視度に応じた重要度を割り当てて記録部へ記録する設定ステップと、を含む。

【0024】

また、本開示に係るプログラムは、情報処理装置に、情報処理装置に、外部から入力される利用者の視線を検出した視線データに基づいて、前記利用者の視線の注視度を解析する解析ステップと、外部から入力される前記利用者の音声データであって、前記視線データと同じ時間軸が対応付けられた音声データに対して、前記解析ステップで解析した前記注視度に応じた重要度を割り当てて記録部へ記録する設定ステップと、を実行させる。

【発明の効果】

【0025】

本開示によれば、視線の動態に基づき、同時に記録した音声データの重要な期間を把握することができるという効果を奏する。

【図面の簡単な説明】

【0026】

【図1】図1は、実施の形態1に係る情報処理システムの機能構成を示すブロック図である。

【図2】図2は、実施の形態1に係る情報処理装置が実行する処理について説明するフローチャートである。

【図3】図3は、実施の形態1に係る解析部による視線の注視度の解析方法を模式的に説明する図である。

【図4A】図4Aは、実施の形態1に係る設定部による音声データへの重要度の割り当ての設定方法を模式的に説明する図である。

【図4B】図4Bは、実施の形態1に係る設定部による音声データへの重要度の割り当ての別の設定方法を模式的に説明する図である。

【図5】図5は、実施の形態1に係る表示部が表示する画像の一例を模式的に示す図である。

【図6】図6は、実施の形態1に係る表示部が表示する画像の別の一例を模式的に示す図

10

20

30

40

50

である。

【図 7】図 7 は、実施の形態 1 の変形例に係る情報処理システムの機能構成を示すブロック図である。

【図 8】図 8 は、実施の形態 1 の変形例に係る情報処理装置が実行する処理の概要を示すフローチャートである。

【図 9】図 9 は、実施の形態 2 に係る情報処理装置の構成を示す概略図である。

【図 10】図 10 は、実施の形態 2 に係る情報処理装置の構成を示す概略図である。

【図 11】図 11 は、実施の形態 2 に係る情報処理装置の機能構成を示すブロック図である。

【図 12】図 12 は、情報処理装置が実行する処理の概要を示すフローチャートである。

10

【図 13】図 13 は、実施の形態 2 に係る表示部が表示する視線マッピング画像の一例を示す図である。

【図 14】図 14 は、実施の形態 2 に係る表示部が表示する視線マッピング画像の別の一例を示す図である。

【図 15】図 15 は、実施の形態 3 に係る顕微鏡システムの構成を示す概略図である。

【図 16】図 16 は、実施の形態 3 に係る顕微鏡システムの機能構成を示すブロック図である。

【図 17】図 17 は、実施の形態 3 に係る顕微鏡システムが実行する処理の概要を示すフローチャートである。

【図 18】図 18 は、実施の形態 4 に係る内視鏡システムの構成を示す概略図である。

20

【図 19】図 19 は、実施の形態 4 に係る内視鏡システムの機能構成を示すブロック図である。

【図 20】図 20 は、実施の形態 4 に係る内視鏡システムが実行する処理の概要を示すフローチャートである。

【図 21】図 21 は、実施の形態 4 に係る画像データ記録部が記録する複数の画像データに対応する複数の画像の一例を模式的に示す図である。

【図 22】図 22 は、実施の形態 4 に係る画像処理部が生成する統合画像データに対応する統合画像の一例を示す図である。

【図 23】図 23 は、実施の形態 5 に係る情報処理システムの機能構成を示すブロック図である。

30

【図 24】図 24 は、実施の形態 5 に係る情報処理システムが実行する処理の概要を示すフローチャートである。

【発明を実施するための形態】

【0027】

以下、本開示を実施するための形態を図面とともに詳細に説明する。なお、以下の実施の形態により本開示が限定されるものではない。また、以下の説明において参照する各図は、本開示の内容を理解でき得る程度に形状、大きさ、および位置関係を概略的に示してあるに過ぎない。即ち、本開示は、各図で例示された形状、大きさおよび位置関係のみに限定されるものでない。

【0028】

40

(実施の形態 1)

〔情報処理システムの構成〕

図 1 は、実施の形態 1 に係る情報処理システムの機能構成を示すブロック図である。図 1 に示す情報処理システム 1 は、外部から入力される視線データ、音声データおよび画像データに対して各種の処理を行う情報処理装置 10 と、情報処理装置 10 から出力された各種データを表示する表示部 20 と、を備える。なお、情報処理装置 10 と表示部 20 は、無線または有線によって双方向に接続されている。

【0029】

〔情報処理装置の構成〕

まず、情報処理装置 10 の構成について説明する。

50

図 1 に示す情報処理装置 10 は、例えばサーバやパーソナルコンピュータ等にインストールされたプログラムを用いて実現され、ネットワークを経由して各種データが入力される、または外部の装置で取得された各種データが入力される。図 1 に示すように、情報処理装置 10 は、解析部 11 と、設定部 12 と、生成部 13 と、記録部 14 と、表示制御部 15 と、を備える。

【0030】

解析部 11 は、外部から入力される利用者の視線を検出した所定時間の視線データに基づいて、利用者の視線の注視度を解析する。ここで、視線データとは、角膜反射法に基づくものである。具体的には、視線データは、図示しない視線検出部（アイトラッキング）に設けられた LED 光源等から近赤外線が利用者の角膜赤に照射された際に、視線検出部である光学センサが角膜上の瞳孔点と反射点を撮像することによって生成されたデータである。そして、視線データは、光学センサが角膜上の瞳孔点と反射点を撮像することによって生成されたデータに対して画像処理等を行うことによって解析した解析結果に基づく利用者の瞳孔点と反射点のパターンから利用者の視線を算出したものである。

10

【0031】

また、図示していないが、視線検出部を備える装置が視線データを計測する際には、対応する画像データを使用者（利用者）に提示したうえで、視線データを計測している。この場合、図示しない視線検出部を備える装置は、使用者に表示している画像が固定している場合、すなわち表示領域の時間とともに絶対座標が変化しないとき、視線に計測領域と画像の絶対座標の相対的に位置関係を固定値として与えていれば良い。ここで、絶対座標とは、画像の所定の 1 点を基準に表記している座標を指している。

20

【0032】

利用形態が内視鏡システムや光学顕微鏡を用いる場合には、解析部 11 は、視線を検出するために提示している視野が画像データの視野となるため、画像の絶対座標に対する観察視野の相対的な位置関係が変わらない。また、利用形態が内視鏡システムや光学顕微鏡を用いる場合において、動画として記録しているとき、解析部 11 は、視野のマッピングデータを生成するために、視線検出データと、視線の検出と同時に記録された画像または提示された画像を用いる。

【0033】

一方で、利用形態が W S I（Whole Slide Imaging）では、顕微鏡のスライドサンプルの一部を視野として使用者が観察しており、時刻とともに観察視野が変化する。この場合、解析部 11 は、全体の画像データのどの部分が視野として提示されているか、すなわち全体の画像データに対する表示領域の絶対座標の切り替えの時間情報も、視線・音声の情報と同じく同期化して記録していること等が考えられる。

30

【0034】

解析部 11 は、外部から入力される利用者の視線を検出した所定時間の視線データに基づいて、視線の移動速度、一定の時間内における視線の移動距離、一定領域内における視線の滞留時間のいずれか 1 つを検出することによって、視線（注視点）の注視度を解析する。なお、図示しない視線検出部は、所定の場所に載置されることによって利用者を撮像することによって視線を検出するものであってもよいし、利用者が装着することによって利用者を撮像することによって視線を検出するものであってもよい。また、視線データは、これ以外にも、周知のパターンマッチングによって生成されたものであってもよい。解析部 11 は、例えば C P U（Central Processing Unit）、F P G A（Field Programmable Gate Array）および G P U（Graphics Processing Unit）等を用いて構成される。

40

【0035】

設定部 12 は、外部から入力される利用者の音声データであって、視線データと同じ時間軸が対応付けられた音声データに対して、所定の時間間隔毎に解析部 11 が解析した注視度に応じた重要度を音声データに割り当てて記録部 14 へ記録する。具体的には、設定部 12 は、音声データのフレーム毎に、このフレームの同じタイミングで解析部 11 が解

50

析した注視度に応じた重要度（例えば数値）を割り当てて記録部 14 へ記録する。また、外部から入力される利用者の音声データは、視線データと同じタイミングで図示しないマイク等の音声入力部によって生成されたものである。設定部 12 は、CPU、FPGA および GPU 等を用いて構成される。

【0036】

生成部 13 は、外部から入力される画像データに対応する画像上に解析部 11 が解析した注視度を関連付けた視線マッピングデータを生成し、この生成した視線マッピングデータを記録部 14 および表示制御部 15 へ出力する。具体的には、生成部 13 は、外部から入力される画像データに対応する画像上の所定領域毎に、解析部 11 が解析した注視度を画像上の座標情報に関連付けた視線マッピングデータを生成する。さらに、生成部 13 は、注視度に加えて、外部から入力される画像データに対応する画像上に解析部 11 が解析した利用者の視線の軌跡を関連付けて視線マッピングデータを生成する。生成部 13 は、CPU、FPGA および GPU 等を用いて構成される。そして、生成部 13 は、上述の WSI で使用する場合、上述の様に視線マッピングデータを画像の絶対座標として得るとき、視線を計測した際の表示と画像の絶対座標の相対的位置関係を使用することによって視線マッピングデータを生成する。また、上述の様に、生成部 13 は、観察視野が時々刻々と変化する場合には、表示領域 = 視野の絶対座標（例えば表示画像の左上が元の画像データに絶対座標でどこに位置するか）の経時変化を入力することによって視線マッピングデータを生成する。

10

【0037】

記録部 14 は、設定部 12 から入力された音声データと、所定の時間間隔毎に割り当てられた重要度と、解析部 11 が解析した注視度と、を対応付けて記録する。また、記録部 14 は、生成部 13 から入力された視線マッピングデータを記録する。また、記録部 14 は、情報処理装置 10 が実行する各種プログラムおよび処理中のデータを記録する。記録部 14 は、揮発性メモリ、不揮発性メモリおよび記録媒体等を用いて構成される。

20

【0038】

表示制御部 15 は、外部から入力される画像データに対応する画像上に、生成部 13 が生成した視線マッピングデータを重畳して外部の表示部 20 に出力することによって表示させる。表示制御部 15 は、CPU、FPGA および GPU 等を用いて構成される。なお、上述した解析部 11、設定部 12、視線マッピングおよび表示制御部 15 を CPU、FPGA および GPU のいずれか 1 つを用いて各機能が発揮できるように構成してもよいし、もちろん、CPU、FPGA および GPU を組み合わせて各機能が発揮できるように構成してもよい。

30

【0039】

〔表示部の構成〕

次に、表示部 20 の構成について説明する。

表示部 20 は、表示制御部 15 から入力された画像データに対応する画像や視線マッピングデータに対応する視線マッピング情報を表示する。表示部 20 は、例えば有機 EL (Electro Luminescence) や液晶等の表示モニタを用いて構成される。

40

【0040】

〔情報処理装置の処理〕

次に、情報処理装置 10 の処理について説明する。図 2 は、情報処理装置 10 が実行する処理について説明するフローチャートである。

【0041】

図 2 に示すように、まず、情報処理装置 10 は、外部から入力される視線データ、音声データおよび画像データを取得する（ステップ S101）。

【0042】

続いて、解析部 11 は、視線データに基づいて、利用者の視線の注視度を解析する（ステップ S102）。ステップ S102 の後、情報処理装置 10 は、後述するステップ S103 へ移行する。

50

【 0 0 4 3 】

図 3 は、解析部 1 1 による視線の注視度の解析方法を模式的に説明する図である。図 3 の (a) および図 3 の (b) において、横軸が時間を示し、図 3 の (a) の縦軸が移動速度を示し、図 3 の (b) の縦軸が注視度を示す。また、図 3 の (a) の曲線 L 1 が視線の移動速度の時間変化を示し、図 3 の (b) の曲線 L 2 が注視度の時間変化を示す。

【 0 0 4 4 】

一般には、視線の移動速度が大きいほど、利用者の注視度が低く、視線の移動速度が小さいほど、利用者の視線の注視度が高いと解析することができる。即ち、図 3 の曲線 L 1 および曲線 L 2 に示すように、解析部 1 1 は、利用者の視線の移動速度が大きいほど、利用者の視線の注視度が低いと解析し、視線の移動速度が小さいほど（視線の移動速度が小さい区間 D 1 を参照）、利用者の視線の注視度が高いと解析する。このように、解析部 1 1 は、所定時間毎（利用者が画像の観察や読影を行っている時間）の視線データに対して、利用者の視線の注視度を解析する。なお、図 3 では、解析部 1 1 は、利用者の視線の移動速度を解析することによって、利用者の視線の注視度を解析していたが、これに限定されることなく、一定の時間内における利用者の視線の移動距離および一定領域内における利用者の視線の滞留時間のいずれか 1 つを検出することによって、視線の注視度を解析してもよい。

【 0 0 4 5 】

図 2 に戻り、ステップ S 1 0 3 以降の説明を続ける。

ステップ S 1 0 3 において、設定部 1 2 は、視線データと同期化された音声データに対して、所定の時間間隔毎に解析部 1 1 が解析した注視度に応じた重要度を音声データに割り当てる設定を行って記録部 1 4 に記録する。ステップ S 1 0 3 の後、情報処理装置 1 0 は、後述するステップ S 1 0 4 へ移行する。

【 0 0 4 6 】

図 4 A は、設定部 1 2 による音声データへの重要度の割り当ての設定方法を模式的に説明する図である。図 4 A において、横軸が時間を示し、図 4 A の (a) の縦軸が注視度を示し、図 4 A の (b) の縦軸が音声データ（発音）を示し、図 4 A の (c) の縦軸が重要度を示す。また、図 4 A の (a) の曲線 L 2 が注視度の時間変化を示し、図 4 の (b) の曲線 L 3 が音声データの時間変化を示し、図 4 A の (c) の曲線 L 4 が重要度の時間変化を示す。

【 0 0 4 7 】

図 4 A に示すように、曲線 L 2 , 曲線 L 3 , 曲線 L 4 では、音声データに変化があった場合において、利用者が注視していないとき（利用者の視線が移動（変化）しているため）、利用者が重要なことを発音している可能性が低いので、重要度が低いと推定することができる。これに対して、図 4 A に示すように、曲線 L 2 , 曲線 L 3 , 曲線 L 4 では、音声データに変化があった場合において、利用者視線の注視度が高いとき（区間 D 1 ）、利用者が重要なことを発音している可能性が高いため、重要度が高いと推定することができる。

【 0 0 4 8 】

即ち、設定部 1 2 は、音声データに対して、所定の時間間隔毎に解析部 1 1 が解析した注視度に応じた重要度を音声データに割り当てる設定を行って記録部 1 4 に記録する。具体的には、図 4 に示す場合、設定部 1 2 は、解析部 1 1 が注視度を高いと解析した区間 D 1 の音声データに対して重要度（例えば数字、視線が滞在していた時間および大中小を示す記号等）を高いと割り当てる設定を行って記録部 1 4 に記録する。このとき、設定部 1 2 は、解析部 1 1 が注視度を高いと解析した区間 D 1 と、音声データの発音区間 D 2 にずれのディレイ期間 d 1 が生じている場合、解析部 1 1 が注視度を高いと解析した区間 D 1 に対応する音声データの前後の近傍である発音区間 D 2（例えば 1 秒前後の区間）を含む発音区間に重要度が高い割り当てを設定して記録部 1 4 に記録する。

【 0 0 4 9 】

なお、実施の形態 1 では、予め利用者の注視度と発音（発声）との時間差を算出し（キャリブレーションデータ）、この算出結果に基づいて利用者の注視度と発音（発声）との

10

20

30

40

50

ずれを補正するキャリブレーション処理を行ってもよい。

【0050】

また、図4Aにおいては、視線データの注視度と発話データとの時間的なずれに着目して、区間D1と発音区間D2に遅延時間を設けるようにしているが、図4Aの変形例として、設定部12は、視線データの注視度が高い区間の前後にマージンを設けることによって、音声データの重要度が高い期間としても良い。すなわち、設定部12は、発音区間D2の開始時間が区間D1の開始時間よりも先になっており、発音区間D2の終了時間が区間D1の終了時間よりも遅くするという状態としてもよい。

【0051】

また、図4Aでは、設定部12は、注視度に応じた時間的な区間D1を設定し、この設定した時間的な区間D1とディレイ期間d1とを考慮して音声データの重要度の設定を行っている。なお、設定部12は、注視度のデータと、発話を検出する機能と、を組み合わせてもよい。具体的には、設定部12は、音声データに対して発話された発話期間を特定してもよい（発話検出部（図示せず）の機能を有してもよい）。この場合、図4Bに示すように、設定部12は、注視度が高い区間D1に基づいて、ウインドウ区間Dxを設定する。そして、設定部12は、入力された音声データに対して、利用者が発話した発話区間（発話期間）、具体的には発話区間P1、P2を検出（特定）する。ここで、設定部12は、発話区間P1、P2のうち、解析部11によって解析された注視度に基づくウインドウ区間Dxと重なりが大きい区間（例えば最も時間相関が高い発話期間）、具体的には重なりが大きい区間が発話区間P2なので、発話区間P2の区間内の音声データに高い重要度を設定する。すなわち、設定部12は、重要度を、注視度が高い期間と発話区間P2の開始終了区間の両方を考慮して設定する。より具体的には、設定部12は、解析部11によって注視度が高いと解析された区間D1（期間）内の時刻において利用者が発話した音声データの発話区間P2（期間）を重要期間と割り当てて記録部14に記録する。この場合、設定部12は、注視度が高いほど音声データに対して重要度を高く割り当てて記録部14へ記録する。さらに、設定部12は、解析部11によって注視度が高いと解析された期間の時刻において利用者が発話した場合、この期間の時刻において利用者が発話した音声データに重要度を割り当てて記録部14へ記録してもよい。

【0052】

図2に戻り、ステップS104以降の説明を続ける。

ステップS104において、生成部13は、画像データに対応する画像上に解析部11が解析した注視度を関連付けた視線マッピングデータを生成する。

【0053】

続いて、表示制御部15は、画像データに対応する画像上に、生成部13が生成した視線マッピングデータを重畳して外部の表示部20に出力する（ステップS105）。ステップS105の後、情報処理装置10は、本処理を終了する。

【0054】

図5は、表示部20が表示する画像の一例を模式的に示す図である。図5に示すように、表示制御部15は、画像データに対応する画像上に、生成部13が生成した視線マッピングデータに重畳した視線マッピング画像P1を表示部20に表示させる。図5においては、視線の注視度が高いほど、等高線の本数が多いヒートマップM1～M5の視線マッピング画像P1を表示部20に表示させる。

【0055】

図6は、表示部20が表示する画像の別の一例を模式的に示す図である。図6に示すように、表示制御部15は、画像データに対応する画像上に、生成部13が生成した視線マッピングデータに重畳した視線マッピング画像P2を表示部20に表示させる。図6においては、視線の注視度が高いほど、円の領域が大きい注視度のマークM11～M15が重畳された視線マッピング画像P2を表示部20に表示させる。さらに、表示制御部15は、利用者の視線の軌跡K1および注視度の順番を数字で表示部20に表示させる。なお、図6においては、表示制御部15は、各注視度の期間（時間）で利用者が発した音声デー

タを、周知の文字変換技術を用いて変換した文字情報を、マーク M 1 1 ~ M 1 5 の近傍または重畳して表示部 2 0 に表示させてもよい。

【 0 0 5 6 】

以上説明した実施の形態 1 によれば、設定部 1 2 が外部から入力された音声データに対して、所定の時間間隔毎に解析部 1 1 が解析した注視度に応じた重要度を音声データに割り当てる設定を行って記録部 1 4 に記録するので、音声データのどの期間が重要であるかを把握することができる。

【 0 0 5 7 】

さらにまた、実施の形態 1 によれば、生成部 1 3 が外部から入力される画像データに対応する画像上に解析部 1 1 が解析した注視度、およびこの注視度の座標情報を関連付けた視線マッピングデータを生成するので、利用者が直感的に画像上における重要な位置を把握することができる。

10

【 0 0 5 8 】

また、実施の形態 1 によれば、記録部 1 4 が設定部 1 2 によって重要度を割り当てた音声データを記録するので、ディープラーニング等の機械学習で用いる視線のマッピングに基づく画像データと音声との対応関係を学習する際の学習データを容易に取得することができる。

【 0 0 5 9 】

また、実施の形態 1 によれば、設定部 1 2 が解析部 1 1 によって注視度が高いと解析された区間（期間）内の時刻において利用者が発話した音声データの期間を重要期間と割り当てて記録部 1 4 に記録するので、他の音声データの区間と区別できるように記録部 1 4 に記録することができる。

20

【 0 0 6 0 】

また、実施の形態 1 によれば、設定部 1 2 が解析部 1 1 によって解析された注視度が高いほど音声データに対して重要度を高く割り当てて記録部 1 4 へ記録するので、音声データの区間毎の重要度を相対的に把握することができる。

【 0 0 6 1 】

また、実施の形態 1 によれば、設定部 1 2 が解析部 1 1 によって注視度が高いと解析された期間の時刻において利用者が発話した場合、この期間の時刻において利用者が発話した音声データに重要度を割り当てて記録部 1 4 へ記録するので、単に注視度が高い場合であっても、音声データがない期間に対して重要度を低く割り当てるので、ディープラーニング等の機械学習で用いる視線のマッピングに基づく画像データと音声との対応関係を学習する際に重要度に応じた適応的な学習ができる。

30

【 0 0 6 2 】

（実施の形態 1 の変形例）

次に、本開示の実施の形態 1 の変形例について説明する。上述した実施の形態 1 では、設定部 1 2 が解析部 1 1 によって解析された注視度に応じた重要度を音声データに割り当てて記録部 1 4 へ記録していたが、実施の形態 1 の変形例では、設定部 1 2 が生成部 1 3 によって生成された視線マッピングデータに基づいて、所定の時間間隔毎に重要度を音声データに割り当てて記録部 1 4 へ記録する。以下においては、実施の形態 1 の変形例に係る情報処理システムの構成を説明後、実施の形態 1 の変形例に係る情報処理装置が実行する処理について説明する。なお、上述した実施の形態 1 に係る情報処理システム 1 と同一の構成には同一の符号を付して詳細な説明は省略する。

40

【 0 0 6 3 】

〔情報処理システムの構成〕

図 7 は、実施の形態 1 の変形例に係る情報処理システムの機能構成を示すブロック図である。図 7 に示す情報処理システム 1 a は、上述した実施の形態 1 に係る情報処理装置 1 0 に代えて、情報処理装置 1 0 a を備える。情報処理装置 1 0 a は、上述した実施の形態 1 に係る設定部 1 2 に代えて、設定部 1 2 a を備える。

【 0 0 6 4 】

50

設定部 1 2 a は、生成部 1 3 によって生成された視線マッピングデータに基づいて、所定の時間間隔毎に重要度を音声データに割り当てて記録部 1 4 へ記録する。

【 0 0 6 5 】

〔 情報処理装置の処理 〕

次に、情報処理装置 1 0 a が実行する処理について説明する。図 8 は、情報処理装置 1 0 a が実行する処理の概要を示すフローチャートである。図 8 において、ステップ S 2 0 1 およびステップ S 2 0 2 は、上述した図 2 のステップ S 1 0 1 およびステップ S 1 0 2 それぞれに対応する。また、図 8 において、ステップ S 2 0 3 は、上述した図 2 のステップ S 1 0 4 に対応する。

【 0 0 6 6 】

ステップ S 2 0 4 において、設定部 1 2 は、生成部 1 3 によって生成された利用者の注視度を画像上に関連付けた視線マッピングデータに基づいて、所定の時間間隔毎に注視度に応じた重要度を音声データに割り当てる設定を行って記録部 1 4 に記録する。ステップ S 2 0 4 の後、情報処理装置 1 0 a は、ステップ S 2 0 5 へ以降する。ステップ S 2 0 5 は、上述した図 2 のステップ S 1 0 5 に対応する。

【 0 0 6 7 】

以上説明した実施の形態 1 の変形例によれば、設定部 1 2 が生成部 1 3 によって生成された利用者の注視度を画像上に関連付けた視線マッピングデータに基づいて、所定の時間間隔毎に注視度に応じた重要度を音声データに割り当てる設定を行って記録部 1 4 に記録するので、音声データのどの期間が重要であるか否かを把握することができる。

【 0 0 6 8 】

（ 実施の形態 2 ）

次に、本開示の実施の形態 2 について説明する。実施の形態 1 では、外部から視線データおよび音声データの各々が入力されていたが、実施の形態 2 では、視線データおよび音声データを生成する。以下においては、実施の形態 2 に係る情報処理装置の構成を説明後、実施の形態 2 に係る情報処理装置が実行する処理について説明する。なお、上述した実施の形態 1 に係る情報処理システム 1 と同一の構成には同一の符号を付して詳細な説明は適宜省略する。

【 0 0 6 9 】

〔 情報処理装置の構成 〕

図 9 は、実施の形態 2 に係る情報処理装置の構成を示す概略図である。図 1 0 は、実施の形態 2 に係る情報処理装置の構成を示す概略図である。図 1 1 は、実施の形態 2 に係る情報処理装置の機能構成を示すブロック図である。

【 0 0 7 0 】

図 9 ~ 図 1 1 に示す情報処理装置 1 b は、解析部 1 1 と、表示部 2 0 と、視線検出部 3 0 と、音声入力部 3 1 と、制御部 3 2 と、時間計測部 3 3 と、記録部 3 4 と、変換部 3 5 と、抽出部 3 6 と、操作部 3 7 と、設定部 3 8 と、生成部 3 9 と、を備える。

【 0 0 7 1 】

視線検出部 3 0 は、近赤外線を照射する L E D 光源と、角膜上の瞳孔点と反射点を撮像する光学センサ（例えば C M O S、C C D 等）と、を用いて構成される。視線検出部 3 0 は、利用者 U 1 が表示部 2 0 を視認可能な情報処理装置 1 b の筐体の側面に設けられる（図 9 および図 1 0 を参照）。視線検出部 3 0 は、制御部 3 2 の制御のもと、表示部 2 0 が表示する画像に対する利用者 U 1 の視線を検出した視線データを生成し、この視線データを制御部 3 2 へ出力する。具体的には、視線検出部 3 0 は、制御部 3 2 の制御のもと、L E D 光源等から近赤外線を利用者 U 1 の角膜に照射し、光学センサが利用者 U 1 の角膜上の瞳孔点と反射点を撮像することによって視線データを生成する。そして、視線検出部 3 0 は、制御部 3 2 の制御のもと、光学センサによって生成されたデータに対して画像処理等によって解析した解析結果に基づいて、利用者 U 1 の瞳孔点と反射点のパターンから利用者の視線や視線を連続的に算出することによって所定時間の視線データを生成し、この視線データを後述する視線検出制御部 3 2 1 へ出力する。なお、視線検出部 3 0 は、単に

光学センサのみで利用者 U 1 の瞳を周知のパターンマッチングを用いることによって瞳を検出することによって、利用者 U 1 の視線を検出した視線データを生成してもよいし、他のセンサや他の周知技術を用いて利用者 U 1 の視線を検出することによって視線データを生成してもよい。

【 0 0 7 2 】

音声入力部 3 1 は、音声が入力されるマイクと、マイクが入力を受け付けた音声をデジタルの音声データに変換するとともに、この音声データを増幅することによって制御部 3 2 へ出力する音声コーデックと、を用いて構成される。音声入力部 3 1 は、制御部 3 2 の制御のもと、利用者 U 1 の音声の入力を受け付けることによって音声データを生成し、この音声データを制御部 3 2 へ出力する。なお、音声入力部 3 1 は、音声の入力以外にも、

10

【 0 0 7 3 】

制御部 3 2 は、CPU、FPGA および GPU 等を用いて構成され、視線検出部 3 0、音声入力部 3 1 および表示部 2 0 を制御する。制御部 3 2 は、視線検出制御部 3 2 1 と、音声入力制御部 3 2 2 と、表示制御部 3 2 3 と、を有する。

【 0 0 7 4 】

視線検出制御部 3 2 1 は、視線検出部 3 0 を制御する。具体的には、視線検出制御部 3 2 1 は、視線検出部 3 0 を所定のタイミング毎に近赤外線を利用者 U 1 へ照射させるとともに、利用者 U 1 の瞳を視線検出部 3 0 に撮像させることによって視線データを生成させる。また、視線検出制御部 3 2 1 は、視線検出部 3 0 から入力された視線データに対して、

20

【 0 0 7 5 】

音声入力制御部 3 2 2 は、音声入力部 3 1 を制御し、音声入力部 3 1 から入力された音声データに対して各種の処理、例えばゲインアップやノイズ低減処理等を行って記録部 3 4 へ出力する。

【 0 0 7 6 】

表示制御部 3 2 3 は、表示部 2 0 の表示態様を制御する。表示制御部 3 2 3 は、記録部 3 4 に記録された画像データに対応する画像または生成部 1 3 によって生成された視線マッピングデータに対応する視線マッピング画像を表示部 2 0 に表示させる。

【 0 0 7 7 】

時間計測部 3 3 は、タイマーやクロックジェネレータ等を用いて構成され、視線検出部 3 0 によって生成された視線データおよび音声入力部 3 1 によって生成された音声データ等に対して時刻情報を付与する。

30

【 0 0 7 8 】

記録部 3 4 は、揮発性メモリ、不揮発性メモリおよび記録媒体等を用いて構成され、情報処理装置 1 b に関する各種の情報を記録する。記録部 3 4 は、視線データ記録部 3 4 1 と、音声データ記録部 3 4 2 と、画像データ記録部 3 4 3 と、プログラム記録部 3 4 4 と、を有する。

【 0 0 7 9 】

視線データ記録部 3 4 1 は、視線検出制御部 3 2 1 から入力された視線データを記録するとともに、視線データを解析部 1 1 へ出力する。

40

【 0 0 8 0 】

音声データ記録部 3 4 2 は、音声入力制御部 3 2 2 から入力された音声データを記録するとともに、音声データを変換部 3 5 へ出力する。

【 0 0 8 1 】

画像データ記録部 3 4 3 は、複数の画像データを記録する。この複数の画像データは、情報処理装置 1 b の外部から入力されたデータ、または記録媒体によって外部の撮像装置によって撮像されたデータである。

【 0 0 8 2 】

プログラム記録部 3 4 4 は、情報処理装置 1 b が実行する各種プログラム、各種プログ

50

ラムの実行中に使用するデータ（例えば辞書情報やテキスト変換辞書情報）および各種プログラムの実行中の処理データを記録する。

【0083】

変換部35は、音声データに対して周知のテキスト変換処理を行うことによって、音声データを文字情報（テキストデータ）に変換し、この文字情報を抽出部36へ出力する。

なお、音声の文字変換はこの時点で行わない構成も可能であり、その際には、音声情報のまま重要度を設定し、その後文字情報に変換するようにしても良い。

【0084】

抽出部36は、後述する操作部37から入力された指示信号に対応する文字や単語（キーワード）を、変換部35によって変換された文字情報から抽出し、この抽出結果を設定部38へ出力する。なお、抽出部36は、後述する操作部37から指示信号が入力されていない場合、変換部35から入力されたままの文字情報を設定部38へ出力する。

10

【0085】

操作部37は、マウス、キーボード、タッチパネルおよび各種スイッチ等を用いて構成され、利用者U1の操作の入力を受け付け、入力を受け付けた操作内容を制御部32へ出力する。

【0086】

設定部38は、所定の時間間隔毎に解析部11が解析した注視度と抽出部36によって抽出された文字情報とに基づいて、視線データと同じ時間軸が対応付けられた音声データに重要度および変換部35によって変換された文字情報を割り当てて記録部34へ記録する。

20

【0087】

生成部39は、表示部20が表示する画像データに対応する画像上に解析部11が解析した注視度および変換部35が変換した文字情報を関連付けた視線マッピングデータを生成し、この視線マッピングデータを画像データ記録部343または表示制御部323へ出力する。

【0088】

〔情報処理装置の処理〕

次に、情報処理装置1bが実行する処理について説明する。図12は、情報処理装置1bが実行する処理の概要を示すフローチャートである。

30

【0089】

図12に示すように、まず、表示制御部323は、画像データ記録部343が記録する画像データに対応する画像を表示部20に表示させる（ステップS301）。この場合、表示制御部323は、操作部37の操作に応じて選択された画像データに対応する画像を表示部20に表示させる。

【0090】

続いて、制御部32は、視線検出部30が生成した視線データおよび音声入力部31が生成した音声データの各々と時間計測部33によって計測された時間とを対応付けて視線データ記録部341および音声データ記録部342に記録する（ステップS302）。

【0091】

40

その後、変換部35は、音声データ記録部342が記録する音声データを文字情報に変換する（ステップS303）。なお、このステップは後述のS306の後に行っても良い。

【0092】

続いて、操作部37から表示部20が表示する画像の観察を終了する指示信号が入力された場合（ステップS304：Yes）、情報処理装置1bは、後述するステップS305へ移行する。これに対して、操作部37から表示部20が表示する画像の観察を終了する指示信号が入力されていない場合（ステップS304：No）、情報処理装置1bは、ステップS302へ戻る。

【0093】

50

ステップ S 3 0 5 は、上述した図 2 のステップ S 1 0 2 に対応する。ステップ S 3 0 5 の後、情報処理装置 1 b は、後述するステップ S 3 0 6 へ移行する。

【 0 0 9 4 】

ステップ S 3 0 6 において、設定部 3 8 は、所定の時間間隔毎に解析部 1 1 が解析した注視度と抽出部 3 6 によって抽出された文字情報とに基づいて、視線データと同じ時間軸が対応付けられた音声データに重要度および変換部 3 5 によって変換された文字情報を割り当てて記録部 3 4 へ記録する。この場合、設定部 3 8 は、抽出部 3 6 によって抽出された文字情報に対応する音声データの重要度の重み付けを行って記録部 3 4 へ記録する。例えば、設定部 3 8 は、重要度に、抽出部 3 6 によって抽出された文字情報に基づく係数を注視度に乗じた値を重要度として音声データに割り当てを行って記録部 3 4 へ記録する。

10

【 0 0 9 5 】

続いて、生成部 3 9 は、表示部 2 0 が表示する画像データに対応する画像上に解析部 1 1 が解析した注視度および変換部 3 5 が変換した文字情報を関連付けた視線マッピングデータを生成する（ステップ S 3 0 7 ）。

【 0 0 9 6 】

続いて、表示制御部 3 2 3 は、生成部 3 9 が生成した視線マッピングデータに対応する視線マッピング画像を表示部 2 0 に表示させる（ステップ S 3 0 8 ）。

【 0 0 9 7 】

図 1 3 は、表示部 2 0 が表示する視線マッピング画像の一例を示す図である。図 1 3 に示すように、表示制御部 3 2 3 は、生成部 1 3 が生成した視線マッピングデータに対応する視線マッピング画像 P 3 を表示部 2 0 に表示させる。視線マッピング画像 P 3 には、視線の注視領域に対応するマーク M 1 1 ~ M 1 5 および視線の軌跡 K 1 が重畳されるとともに、この注視度のタイミングで発せられた音声データの文字情報が関連付けられている。また、マーク M 1 1 ~ M 1 5 は、番号が利用者 U 1 の視線の順番を示し、大きさ（領域）が注視度の大きさを示す。さらに、利用者 U 1 が操作部 3 7 を操作してカーソル A 1 を所望の位置、例えばマーク M 1 4 に移動させた場合、マーク M 1 4 に関連付けられた文字情報 Q 1、例えば「ここで癌があります。」が表示される。なお、図 1 3 では、表示制御部 3 2 3 が文字情報を表示部 2 0 に表示させているが、例えば文字情報を音声に変換することによって音声データを出力してもよい。これにより、利用者 U 1 は、重要な音声内容と注視していた領域とを直感的に把握することができる。さらに、利用者 U 1 の観察時における視線の軌跡を直感的に把握することができる。

20

30

【 0 0 9 8 】

図 1 4 は、表示部 2 0 が表示する視線マッピング画像の別の一例を示す図である。図 1 3 に示すように、表示制御部 3 2 3 は、生成部 1 3 が生成した視線マッピングデータに対応する視線マッピング画像 P 4 を表示部 2 0 に表示させる。さらに、表示制御部 3 2 3 は、文字情報と、この文字情報が発声された時間とを対応付けたアイコン B 1 ~ B 5 を表示部 2 0 に表示させる。さらに、表示制御部 3 2 3 は、利用者 U 1 が操作部 3 7 を操作してマーク M 1 1 ~ M 1 5 のいずれかを選択した場合、例えばマーク M 1 4 を選択した場合、マーク M 1 4 を表示部 2 0 に強調表示するとともに、マーク M 1 4 の時間に対応する文字情報、例えばアイコン B 4 を表示部 2 0 に強調表示させる（例えば枠をハイライト表示または太線で表示）。これにより、利用者 U 1 は、重要な音声内容と注視していた領域とを直感的に把握することができるうえ、発声した際の内容を直感的に把握することができる。

40

【 0 0 9 9 】

図 1 2 に戻り、ステップ S 3 0 9 以降の説明を続ける。

ステップ S 3 0 9 において、操作部 3 7 によって複数の注視領域に対応するマークのいずれか一つが操作された場合（ステップ S 3 0 9 : Y e s）、制御部 3 2 は、操作に応じた動作処理を実行する（ステップ S 3 1 0）。具体的には、表示制御部 3 2 3 は、操作部 3 7 によって選択された注視領域に対応するマークを表示部 2 0 に強調表示させる（例えば図 1 3 を参照）。また、音声入力制御部 3 2 2 は、注視度の高い領域に関連付けられた

50

音声データを音声入力部 3 1 に再生させる。ステップ S 3 1 0 の後、情報処理装置 1 b は、後述するステップ S 3 1 1 へ移行する。

【0100】

ステップ S 3 0 9 において、操作部 3 7 によって複数の注視度領域に対応するマークのいずれか一つが操作されていない場合（ステップ S 3 0 9 : N o ）、情報処理装置 1 b は、後述するステップ S 3 1 1 へ移行する。

【0101】

ステップ S 3 1 1 において、操作部 3 7 から観察の終了を指示する指示信号が入力された場合（ステップ S 3 1 1 : Y e s ）、情報処理装置 1 b は、本処理を終了する。これに対して、操作部 3 7 から観察の終了を指示する指示信号が入力されていない場合（ステップ S 3 1 1 : N o ）、情報処理装置 1 b は、上述したステップ S 3 0 8 へ戻る。

10

【0102】

以上説明した実施の形態 2 によれば、生成部 1 3 が表示部 2 0 によって表示される画像データに対応する画像上に解析部 1 1 が解析した注視度および変換部 3 5 が変換した文字情報を関連付けた視線マッピングデータを生成するので、利用者 U 1 は、重要な音声内容と注視していた領域とを直感的に把握することができるうえ、発声した際の内容を直感的に把握することができる。

【0103】

また、実施の形態 2 によれば、表示制御部 3 2 3 は、生成部 1 3 が生成した視線マッピングデータに対応する視線マッピング画像を表示部 2 0 に表示させるので、画像に対する利用者の観察の見逃し防止の確認、利用者の読影等の技術スキルの確認、他の利用者に対する読影や観察等の教育およびカンファレンス等に用いることができる。

20

【0104】

（実施の形態 3）

次に、本開示の実施の形態 3 について説明する。上述した実施の形態 2 では、情報処理装置 1 b のみで構成されていたが、実施の形態 3 では、顕微鏡システムの一部に情報処理装置を組み込むことによって構成する。以下においては、実施の形態 3 に係る顕微鏡システムの構成を説明後、実施の形態 3 に係る顕微鏡システムが実行する処理について説明する。なお、上述した実施の形態 2 に係る情報処理装置 1 b と同一の構成には同一の符号を付して詳細な説明は適宜省略する。

30

【0105】

〔顕微鏡システムの構成〕

図 1 5 は、実施の形態 3 に係る顕微鏡システムの構成を示す概略図である。図 1 6 は、実施の形態 3 に係る顕微鏡システムの機能構成を示すブロック図である。

【0106】

図 1 5 および図 1 6 に示すように、顕微鏡システム 1 0 0 は、情報処理装置 1 c と、表示部 2 0 と、音声入力部 3 1 と、操作部 3 7 と、顕微鏡 2 0 0 と、撮像部 2 1 0 と、視線検出部 2 2 0 と、を備える。

【0107】

〔顕微鏡の構成〕

40

まず、顕微鏡 2 0 0 の構成について説明する。

顕微鏡 2 0 0 は、本体部 2 0 1 と、回転部 2 0 2 と、昇降部 2 0 3 と、レボルバ 2 0 4 と、対物レンズ 2 0 5 と、倍率検出部 2 0 6 と、鏡筒部 2 0 7 と、接続部 2 0 8 と、接眼部 2 0 9 と、を備える。

【0108】

本体部 2 0 1 は、標本 S P が載置される。本体部 2 0 1 は、略 U 字状をなし、回転部 2 0 2 を用いて昇降部 2 0 3 が接続される。

【0109】

回転部 2 0 2 は、利用者 U 2 の操作に応じて回転することによって、昇降部 2 0 3 を垂直方向へ移動させる。

50

【 0 1 1 0 】

昇降部 2 0 3 は、本体部 2 0 1 に対して垂直方向へ移動可能に設けられている。昇降部 2 0 3 は、一端側の面にレボルバが接続され、他端側の面に鏡筒部 2 0 7 が接続される。

【 0 1 1 1 】

レボルバ 2 0 4 は、互いに倍率が異なる複数の対物レンズ 2 0 5 が接続され、光軸 L 1 に対して回転可能に昇降部 2 0 3 に接続される。レボルバ 2 0 4 は、利用者 U 2 の操作に応じて、所望の対物レンズ 2 0 5 を光軸 L 1 上に配置する。なお、複数の対物レンズ 2 0 5 には、倍率を示す情報、例えば I C チップやラベルが添付されている。なお、I C チップやラベル以外にも、倍率を示す形状を対物レンズ 2 0 5 に設けてもよい。

【 0 1 1 2 】

倍率検出部 2 0 6 は、光軸 L 1 上に配置された対物レンズ 2 0 5 の倍率を検出し、この検出した検出結果を情報処理装置 1 c へ出力する。倍率検出部 2 0 6 は、例えば対物切り替えのレボルバ 2 0 4 の位置を検出する手段を用いて構成される。

【 0 1 1 3 】

鏡筒部 2 0 7 は、対物レンズ 2 0 5 によって結像された標本 S P の被写体像の一部を接続部 2 0 8 に透過するとともに、接眼部 2 0 9 へ反射する。鏡筒部 2 0 7 は、内部にプリズム、ハーフミラーおよびコリメートレンズ等を有する。

【 0 1 1 4 】

接続部 2 0 8 は、一端が鏡筒部 2 0 7 と接続され、他端が撮像部 2 1 0 と接続される。接続部 2 0 8 は、鏡筒部 2 0 7 を透過した標本 S P の被写体像を撮像部 2 1 0 へ導光する。接続部 2 0 8 は、複数のコリメートレンズおよび結像レンズ等を用いて構成される。

【 0 1 1 5 】

接眼部 2 0 9 は、鏡筒部 2 0 7 によって反射された被写体像を導光して結像する。接眼部 2 0 9 は、複数のコリメートレンズおよび結像レンズ等を用いて構成される。

【 0 1 1 6 】

〔 撮像部の構成 〕

次に、撮像部 2 1 0 の構成について説明する。

撮像部 2 1 0 は、接続部 2 0 8 が結像した標本 S P の被写体像を受光することによって画像データを生成し、この画像データを情報処理装置 1 c へ出力する。撮像部 2 1 0 は、C M O S または C C D 等のイメージセンサおよび画像データに対して各種の画像処理を施す画像処理エンジン等を用いて構成される。

【 0 1 1 7 】

〔 視線検出部の構成 〕

次に、視線検出部 2 2 0 の構成について説明する。

視線検出部 2 2 0 は、接眼部 2 0 9 の内部または外部に設けられ、利用者 U 2 の視線を検出することによって視線データを生成し、この視線データを情報処理装置 1 c へ出力する。視線検出部 2 2 0 は、接眼部 2 0 9 の内部に設けられ、近赤外線を照射する L E D 光源と、接眼部 2 0 9 の内部に設けられ、角膜上の瞳孔点と反射点を撮像する光学センサ（例えば C M O S、C C D）と、を用いて構成される。視線検出部 2 2 0 は、情報処理装置 1 c の制御のもと、L E D 光源等から近赤外線を利用者 U 2 の角膜に照射し、光学センサが利用者 U 2 の角膜上の瞳孔点と反射点を撮像することによって生成する。そして、視線検出部 2 2 2 は、情報処理装置 1 c の制御のもと、光学センサによって生成されたデータに対して画像処理等によって解析した解析結果に基づいて、利用者 U 2 の瞳孔点と反射点のパターンから利用者の視線を検出することによって視線データを生成し、この視線データを情報処理装置 1 c へ出力する。

【 0 1 1 8 】

〔 情報処理装置の構成 〕

次に、情報処理装置 1 c の構成について説明する。

情報処理装置 1 c は、上述した実施の形態 2 に係る情報処理装置 1 b の制御部 3 2、記録部 3 4 および設定部 3 8 に換えて、制御部 3 2 c、記録部 3 4 c、設定部 3 8 c と、を

10

20

30

40

50

備える。

【0119】

制御部32cは、CPU、FPGAおよびGPU等を用いて構成され、表示部20、音声入力部31、撮像部210および視線検出部220を制御する。制御部32cは、上述した実施の形態2の制御部32の視線検出制御部321、音声入力制御部322、表示制御部323に加えて、撮影制御部324および倍率算出部325をさらに備える。

【0120】

撮影制御部324は、撮像部210の動作を制御する。撮影制御部324は、撮像部210を所定のフレームレートに従って順次撮像させることによって画像データを生成させる。撮影制御部324は、撮像部210から入力された画像データに対して処理の画像処理（例えば現像処理等）を施して記録部34cへ出力する。

10

【0121】

倍率算出部325は、倍率検出部206から入力された検出結果に基づいて、現在の顕微鏡200の観察倍率を算出し、この算出結果を設定部38cへ出力する。例えば、倍率算出部325は、倍率検出部206から入力された対物レンズ205の倍率と接眼部209の倍率とに基づいて、現在の顕微鏡200の観察倍率を算出する。

【0122】

記録部34cは、揮発性メモリ、不揮発性メモリおよび記録媒体等を用いて構成される。記録部34cは、上述した実施の形態2に係る画像データ記録部343に換えて、画像データ記録部345を備える。画像データ記録部345は、撮影制御部324から入力された画像データを記録し、この画像データを生成部13へ出力する。

20

【0123】

設定部38cは、所定の時間間隔毎に解析部11が解析した注視度と倍率算出部325が算出した算出結果とに基づいて、視線データと同じ時間軸が対応付けられた音声データに重要度および変換部35によって変換された文字情報を割り当てて記録部34cへ記録する。具体的には、設定部38cは、解析部11が解析した注視度に、倍率算出部325が算出した算出結果に基づく係数を乗じた値を、音声データのフレーム毎の重要度（例えば数値）として割り当てて記録部34cへ記録する。すなわち、設定部38cは、表示倍率が大きいかほど重要度が高くなるような処理を行う。設定部38cは、CPU、FPGAおよびGPU等を用いて構成される。

30

【0124】

〔顕微鏡システムの処理〕

次に、顕微鏡システム100が実行する処理について説明する。図17は、顕微鏡システム100が実行する処理の概要を示すフローチャートである。

【0125】

図17に示すように、まず、制御部32cは、視線検出部30が生成した視線データ、音声入力部31が生成した音声データ、および倍率算出部325が算出した観察倍率の各々を時間計測部33によって計測された時間に対応付けて視線データ記録部341および音声データ記録部342に記録する（ステップS401）。ステップS401の後、顕微鏡システム100は、後述するステップS402へ移行する。

40

【0126】

ステップS402～ステップS404は、上述した図12のステップS303～ステップS305それぞれに対応する。ステップS404の後、顕微鏡システム100は、ステップS405へ移行する。

【0127】

ステップS405において、設定部38cは、所定の時間間隔毎に解析部11が解析した注視度と倍率算出部325が算出した算出結果とに基づいて、視線データと同じ時間軸が対応付けられた音声データに重要度および変換部35によって変換された文字情報を割り当てて記録部34cへ記録する。ステップS405の後、顕微鏡システム100は、ステップS406へ移行する。

50

【 0 1 2 8 】

ステップ S 4 0 6 ~ ステップ S 4 1 0 は、上述した図 1 2 のステップ S 3 0 7 ~ ステップ S 3 1 1 それぞれに対応する。

【 0 1 2 9 】

以上説明した実施の形態 3 によれば、設定部 3 8 c が解析部 1 1 によって解析された注視度と倍率算出部 3 2 5 が算出した算出結果とに基づいて、視線データと同じ時間軸が対応付けられた音声データに重要度および変換部 3 5 によって変換された文字情報を割り当てて記録部 3 4 c することによって、観察倍率および注視度に基づいた重要度が音声データに割り当てられるので、観察内容および注視度を加味した音声データの重要な期間を把握することができる。

10

【 0 1 3 0 】

なお、実施の形態 3 では、倍率算出部 3 2 5 が算出した観察倍率を記録部 1 4 に記録していたが、利用者 U 2 の操作履歴を記録し、この操作履歴をさらに加味して音声データの重要度を割り当ててもよい。

【 0 1 3 1 】

(実施の形態 4)

次に、本開示の実施の形態 4 について説明する。実施の形態 4 では、内視鏡システムの一部に情報処理装置を組み込むことによって構成する。以下においては、実施の形態 4 に係る内視鏡システムの構成を説明後、実施の形態 4 に係る内視鏡システムが実行する処理について説明する。なお、上述した実施の形態 2 に係る情報処理装置 1 b と同一の構成には同一の符号を付して詳細な説明は適宜省略する。

20

【 0 1 3 2 】

〔 内視鏡システムの構成 〕

図 1 8 は、実施の形態 4 に係る内視鏡システムの構成を示す概略図である。図 1 9 は、実施の形態 4 に係る内視鏡システムの機能構成を示すブロック図である。

【 0 1 3 3 】

図 1 8 および図 1 9 に示す内視鏡システム 3 0 0 は、表示部 2 0 と、内視鏡 4 0 0 と、ウェアラブルデバイス 5 0 0 と、入力部 6 0 0 と、情報処理装置 1 d と、を備える。

【 0 1 3 4 】

〔 内視鏡の構成 〕

まず、内視鏡 4 0 0 の構成について説明する。

内視鏡 4 0 0 は、医者や術者等の利用者 U 3 が被検体 U 4 に挿入することによって、被検体 U 4 の内部を撮像することによって画像データを生成し、この画像データを情報処理装置 1 d へ出力する。内視鏡 4 0 0 は、撮像部 4 0 1 と、操作部 4 0 2 と、を備える。

30

【 0 1 3 5 】

撮像部 4 0 1 は、内視鏡 4 0 0 の挿入部の先端部に設けられる。撮像部 4 0 1 は、情報処理装置 1 d の制御のもと、被検体 U 4 の内部を撮像することによって画像データを生成し、この画像データを情報処理装置 1 d へ出力する。撮像部 4 0 1 は、観察倍率を変更することができる光学系と、光学系が結像した被写体像を受光することによって画像データを生成する C M O S や C C D 等のイメージセンサ等を用いて構成される。

40

【 0 1 3 6 】

操作部 4 0 2 は、利用者 U 3 の各種の操作の入力を受け付け、受け付けた各種操作に応じた操作信号を情報処理装置 1 d へ出力する。

【 0 1 3 7 】

〔 ウェアラブルデバイスの構成 〕

次に、ウェアラブルデバイス 5 0 0 の構成について説明する。

ウェアラブルデバイス 5 0 0 は、利用者 U 3 に装着され、利用者 U 3 の視線を検出するとともに、利用者 U 3 の音声の入力を受け付ける。ウェアラブルデバイス 5 0 0 は、視線検出部 5 1 0 と、音声入力部 5 2 0 と、を有する。

【 0 1 3 8 】

50

視線検出部 5 1 0 は、ウェアラブルデバイス 5 0 0 に設けられ、利用者 U 3 の視線の注視度を検出することによって視線データを生成し、この視線データを情報処理装置 1 d へ出力する。視線検出部 5 1 0 は、上述した実施の形態 3 に係る視線検出部 2 2 0 と同様の構成を有するため、詳細な構成は省略する。

【 0 1 3 9 】

音声入力部 5 2 0 は、ウェアラブルデバイス 5 0 0 に設けられ、利用者 U 3 の音声の入力を受け付けることによって音声データを生成し、この音声データを情報処理装置 1 d へ出力する。音声入力部 5 2 0 は、マイク等を用いて構成される。

【 0 1 4 0 】

〔入力部の構成〕

10

入力部 6 0 0 の構成について説明する。

入力部 6 0 0 は、マウス、キーボード、タッチパネルおよび各種のスイッチを用いて構成される。入力部 6 0 0 は、利用者 U 3 の各種の操作の入力を受け付け、受け付けた各種操作に応じた操作信号を情報処理装置 1 d へ出力する。

【 0 1 4 1 】

〔情報処理装置の構成〕

次に、情報処理装置 1 d の構成について説明する。

情報処理装置 1 d は、上述した実施の形態 3 に係る情報処理装置 1 c の制御部 3 2 c、記録部 3 4 c、設定部 3 8 c、生成部 3 9 に換えて、制御部 3 2 d、記録部 3 4 d、設定部 3 8 d および生成部 3 9 d を備える。さらに、情報処理装置 1 d は、画像処理部 4 0 をさらに備える。

20

【 0 1 4 2 】

制御部 3 2 d は、CPU、FPGA および GPU 等を用いて構成され、内視鏡 4 0 0、ウェアラブルデバイス 5 0 0 および表示部 2 0 を制御する。制御部 3 2 d は、視線検出制御部 3 2 1、音声入力制御部 3 2 2、表示制御部 3 2 3、撮影制御部 3 2 4 に加えて、操作履歴検出部 3 2 6 を備える。

【 0 1 4 3 】

操作履歴検出部 3 2 6 は、内視鏡 4 0 0 の操作部 4 0 2 が入力を受け付けた操作の内容を検出し、この検出結果を記録部 3 4 d に出力する。具体的には、操作履歴検出部 3 2 6 は、内視鏡 4 0 0 の操作部 4 0 2 から拡大スイッチが操作された場合、この操作内容を検出し、この検出結果を記録部 3 4 d に出力する。なお、操作履歴検出部 3 2 6 は、内視鏡 4 0 0 を経由して被検体 U 4 の内部に挿入される処置具の操作内容を検出し、この検出結果を記録部 3 4 d に出力してもよい。

30

【 0 1 4 4 】

記録部 3 4 d は、揮発性メモリ、不揮発性メモリおよび記録媒体等を用いて構成される。記録部 3 4 d は、上述した実施の形態 3 に係る記録部 3 4 c の構成に加えて、操作履歴記録部 3 4 6 をさらに備える。

【 0 1 4 5 】

操作履歴記録部 3 4 6 は、操作履歴検出部 3 2 6 から入力された内視鏡 4 0 0 の操作部 4 0 2 に対する操作の履歴を記録する。

40

【 0 1 4 6 】

設定部 3 8 d は、所定の時間間隔毎に解析部 1 1 が解析した注視度と操作履歴記録部 3 4 6 が記録する操作履歴とに基づいて、視線データと同じ時間軸が対応付けられた音声データに重要度および変換部 3 5 によって変換された文字情報を割り当てて記録部 3 4 d へ記録する。具体的には、設定部 3 8 d は、解析部 1 1 が解析した注視度と操作履歴記録部 3 4 6 が記録する操作履歴とに基づいて、音声データのフレーム毎に重要度（例えば数値）を割り当てて記録部 3 4 d へ記録する。すなわち、設定部 3 8 d は、操作履歴の内容に応じて設定された係数が大きいほど重要度が高くなるような処理を行う。設定部 3 8 d は、CPU、FPGA および GPU 等を用いて構成される。

【 0 1 4 7 】

50

生成部 39d は、画像処理部 40 が生成した統合画像データに対応する統合画像上に、解析部 11 が解析した注視度および文字情報を関連付けた視線マッピングデータを生成し、この生成した視線マッピングデータを記録部 34d および表示制御部 323 へ出力する。

【0148】

画像処理部 40 は、画像データ記録部 345 が記録する複数の画像データを合成することによって 3 次元画像の統合画像データを生成し、この統合画像データを生成部 39d へ出力する。

【0149】

〔内視鏡システムの処理〕

次に、内視鏡システム 300 が実行する処理について説明する。図 20 は、内視鏡システム 300 が実行する処理の概要を示すフローチャートである。

【0150】

図 20 に示すように、まず、制御部 32d は、視線検出部 30 が生成した視線データ、音声入力部 31 が生成した音声データ、および操作履歴検出部 326 が検出した操作履歴の各々を時間計測部 33 によって計測された時間と対応付けて視線データ記録部 341、音声データ記録部 342 および操作履歴記録部 346 に記録する（ステップ S501）。ステップ S501 の後、内視鏡システム 300 は、後述するステップ S502 へ移行する。

【0151】

ステップ S502 ～ステップ S504 は、上述した図 12 のステップ S303 ～ステップ S305 それぞれに対応する。ステップ S504 の後、内視鏡システム 300 は、ステップ S505 へ移行する。

【0152】

ステップ S505 において、設定部 38d は、所定の時間間隔毎に解析部 11 が解析した注視度と操作履歴記録部 346 が記録する操作履歴とに基づいて、視線データと同じ時間軸が対応付けられた音声データに重要度および変換部 35 によって変換された文字情報を割り当てて記録部 34d へ記録する。

【0153】

続いて、画像処理部 40 は、画像データ記録部 345 が記録する複数の画像データを合成することによって 3 次元画像の統合画像データを生成し、この統合画像データを生成部 39d へ出力する（ステップ S506）。図 21 は、画像データ記録部 345 が記録する複数の画像データに対応する複数の画像の一例を模式的に示す図である。図 22 は、画像処理部 40 が生成する統合画像データに対応する統合画像の一例を示す図である。図 21 および図 22 に示すように、画像処理部 40 は、時間的に連続する複数の画像データ $P_1 \sim P_N$ （ N = 整数）を合成することによって統合画像データに対応する統合画像 P_{10} を生成する。

【0154】

その後、生成部 39d は、画像処理部 40 が生成した統合画像データに対応する統合画像 P_{100} 上に、解析部 11 が解析した注視度、視線および文字情報を関連付けた視線マッピングデータを生成し、この生成した視線マッピングデータを記録部 34d および表示制御部 323 へ出力する（ステップ S507）。この場合、生成部 39d は、画像処理部 40 が生成した統合画像データに対応する統合画像 P_{100} 上に、解析部 11 が解析した注視度、視線 K_2 および文字情報に加えて、操作履歴を関連付けてもよい。ステップ S507 の後、内視鏡システム 300 は、後述するステップ S508 へ移行する。

【0155】

ステップ S508 ～ステップ S511 は、上述した図 12 のステップ S308 ～ステップ S311 それぞれに対応する。

【0156】

以上説明した実施の形態 4 によれば、設定部 38d が解析部 11 によって解析された注

10

20

30

40

50

視度と操作履歴記録部 3 4 6 が記録する操作履歴とに基づいて、視線データと同じ時間軸が対応付けられた音声データに重要度および変換部 3 5 によって変換された文字情報を割り当てて記録部 3 4 d へ記録することによって、操作履歴および注視度に基づいた重要度が音声データに割り当てられるので、操作内容および注視度を加味した音声データの重要な期間を把握することができる。

【 0 1 5 7 】

また、実施の形態 4 では、内視鏡システムであったが、例えばカプセル型の内視鏡、被検体を撮像するビデオマイクロスコープ、撮像機能を有する携帯電話および撮像機能を有するタブレット型端末であっても適用することができる。

【 0 1 5 8 】

また、実施の形態 4 では、軟性の内視鏡を備えた内視鏡システムであったが、硬性の内視鏡を備えた内視鏡システム、工業用の内視鏡を備えた内視鏡システムであっても適用することができる。

【 0 1 5 9 】

また、実施の形態 4 では、被検体に挿入される内視鏡を備えた内視鏡システムであったが、副鼻腔内視鏡および電気メスや検査プローブ等の内視鏡システムであっても適用することができる。

【 0 1 6 0 】

(実施の形態 5)

次に、本開示の実施の形態 5 について説明する。上述した実施の形態 1 ~ 4 は、利用者が一人の場合を想定していたが、実施の形態 5 では、2 人以上の利用者を想定する。さらに、実施の形態 5 では、複数の利用者で画像を閲覧する情報処理システムに情報処理装置を組み込むことによって構成する。以下においては、実施の形態 5 に係る閲覧システムの構成を説明後、実施の形態 5 に係る情報処理システムが実行する処理について説明する。なお、上述した実施の形態 2 に係る情報処理装置 1 b と同一の構成には同一の符号を付して詳細な説明は適宜省略する。

【 0 1 6 1 】

〔 情報処理システムの構成 〕

図 2 3 は、実施の形態 5 に係る情報処理システムの機能構成を示すブロック図である。図 2 3 に示す情報処理システム 7 0 0 は、表示部 2 0 と、第 1 ウェアラブルデバイス 7 1 0 と、第 2 ウェアラブルデバイス 7 2 0 と、検出部 7 3 0 と、情報処理装置 1 e と、を備える。

【 0 1 6 2 】

〔 第 1 ウェアラブルデバイスの構成 〕

まず、第 1 ウェアラブルデバイス 7 1 0 の構成について説明する。

第 1 ウェアラブルデバイス 7 1 0 は、利用者に装着され、利用者の視線を検出するとともに、利用者の音声の入力を受け付ける。第 1 ウェアラブルデバイス 7 1 0 は、第 1 視線検出部 7 1 1 と、第 1 音声入力部 7 1 2 と、を有する。第 1 視線検出部 7 1 1 および第 1 音声入力部 7 1 2 は、上述した実施の形態 4 に係る視線検出部 5 1 0 および音声入力部 5 2 0 と同様の構成を有するため、詳細な構成は省略する。

【 0 1 6 3 】

〔 第 2 ウェアラブルデバイスの構成 〕

次に、第 2 ウェアラブルデバイス 7 2 0 の構成について説明する。

第 2 ウェアラブルデバイス 7 2 0 は、上述した第 1 ウェアラブルデバイス 7 1 0 と同様の構成を有し、利用者に装着され、利用者の視線を検出するとともに、利用者の音声の入力を受け付ける。第 2 ウェアラブルデバイス 7 2 0 は、第 2 視線検出部 7 2 1 と、第 2 音声入力部 7 2 2 と、を有する。第 2 視線検出部 7 2 1 および第 2 音声入力部 7 2 2 は、上述した実施の形態 4 に係る視線検出部 5 1 0 および音声入力部 5 2 0 と同様の構成を有するため、詳細な構成は省略する。

【 0 1 6 4 】

〔検出部の構成〕

次に、検出部 730 の構成について説明する。

検出部 730 は、複数の利用者の各々を識別する識別情報を検出し、この検出結果を情報処理装置 1e へ出力する。検出部 730 は、複数の利用者の各々を識別する識別情報（例えば ID や名前等）を記録する IC カードから利用者の識別情報を検出し、この検出結果を情報処理装置 1e へ出力する。検出部 730 は、例えば、IC カードを読み取るカードリーダー等を用いて構成される。なお、検出部 730 は、複数の利用者の顔を撮像することによって生成した画像データに対応する画像に対して、予め設定された利用者の顔の特徴点および周知のパターンマッチングを用いて利用者を識別し、この識別結果を情報処理装置 1e へ出力するようにしてもよい。もちろん、検出部 730 は、操作部 37 からの操作に応じて入力された信号に基づいて、利用者を識別し、この識別結果を情報処理装置 1e へ出力するようにしてもよい。

10

【0165】

〔情報処理装置の構成〕

次に、情報処理装置 1e の構成について説明する。

情報処理装置 1e は、上述した実施の形態 4 に係る情報処理装置 1d の制御部 32d、記録部 34d および設定部 38d に換えて、制御部 32e、記録部 34e および設定部 38e を備える。

【0166】

制御部 32e は、CPU、FPGA および GPU 等を用いて構成され、第 1 ウェアラブルデバイス 710、第 2 ウェアラブルデバイス 720、検出部 730 および表示部 20 を制御する。制御部 32e は、視線検出制御部 321、音声入力制御部 322、表示制御部 323 に加えて、識別検出制御部 327 を備える。

20

【0167】

識別検出制御部 327 は、検出部 730 を制御し、検出部 730 が取得した取得結果に基づいて、複数の利用者の各々を識別し、この識別結果を記録部 34e へ出力する。

【0168】

記録部 34e は、揮発性メモリ、不揮発性メモリおよび記録媒体等を用いて構成される。記録部 34e は、上述した実施の形態 3 に係る記録部 34c の構成に加えて、識別情報記録部 347 をさらに備える。

30

【0169】

識別情報記録部 347 は、識別検出制御部 327 から入力された複数の利用者の各々の識別情報を記録する。

【0170】

設定部 38e は、解析部 11 が解析した解析結果と、抽出部 36 が抽出した文字情報と、識別情報記録部 347 が記録する識別情報と、に基づいて、所定の時間間隔毎に視線データと同じ時間軸が対応付けられた音声データに重要度および変換部 35 によって変換された文字情報を割り当てて記録部 34e に記録する。さらに、設定部 38e は、識別情報記録部 347 が記録する各利用者の識別情報に応じて重要度の重み付けを行う。すなわち、設定部 38e は、重要な利用者（例えば役職に応じて設定されたランク）ほど重要度が高くなる処理を行う。

40

【0171】

〔情報処理システムの処理〕

次に、情報処理システム 700 が実行する処理について説明する。図 24 は、情報処理システム 700 が実行する処理の概要を示すフローチャートである。

【0172】

図 24 に示すように、表示制御部 323 は、画像データ記録部 343 が記録する画像データに対応する画像を表示部 20 に表示させる（ステップ S601）。

【0173】

続いて、制御部 32e は、第 1 ウェアラブルデバイス 710 および第 2 ウェアラブルデ

50

バイス 7 2 0 の各々が生成した視線データ、音声データ、および検出部 7 3 0 が取得した識別情報の各々を時間計測部 3 3 によって計測された時間と対応付けて視線データ記録部 3 4 1、音声データ記録部 3 4 2 および識別情報記録部 3 4 7 に記録する（ステップ S 6 0 2）。ステップ S 6 0 2 の後、情報処理システム 7 0 0 は、ステップ S 6 0 3 へ移行する。

【0174】

ステップ S 6 0 3 およびステップ S 6 0 4 は、上述した図 1 2 のステップ S 3 0 3 およびステップ S 3 0 4 それぞれに対応する。ステップ S 6 0 4 の後、情報処理システム 7 0 0 は、後述するステップ S 6 0 5 へ移行する。

【0175】

ステップ S 6 0 5 において、解析部 1 1 は、第 1 ウェアラブルデバイス 7 1 0 が生成した第 1 視線データおよび第 2 ウェアラブルデバイス 7 2 0 が生成した第 2 視線データに基づいて、各利用者の視線の注視度を解析する。

【0176】

続いて、設定部 3 8 e は、所定の時間間隔毎に解析部 1 1 が解析した各注視度と、識別情報記録部 3 4 7 が記録する識別情報とに基づいて、視線データと同じ時間軸が対応付けられた第 1 音声データおよび第 2 音声データの各々に重要度および変換部 3 5 によって変換された文字情報を割り当てて記録部 3 4 e へ記録する（ステップ S 6 0 6）。

【0177】

ステップ S 6 0 7 ~ ステップ S 6 1 1 は、上述した図 1 2 のステップ S 3 0 7 ~ ステップ S 3 1 1 それぞれに対応する。

【0178】

以上説明した実施の形態 5 によれば、設定部 3 8 e が解析部 1 1 によって解析された各利用者の注視度と、識別情報記録部 3 4 7 が記録する識別情報とに基づいて、視線データと同じ時間軸が対応付けられた第 1 音声データおよび第 2 音声データの各々に重要度および変換部 3 5 によって変換された文字情報を割り当てて記録部 3 4 e へ記録することによって、識別情報および注視度に基づいた重要度が第 1 音声データまたは第 2 音声データに割り当てられるので、利用者に応じた注視度を加味した音声データの重要な期間を把握することができる。

【0179】

なお、実施の形態 5 では、設定部 3 8 e が解析部 1 1 によって解析された各利用者の注視度と、識別情報記録部 3 4 7 が記録する各利用者の識別情報とに基づいて、視線データと同じ時間軸が対応付けられた第 1 音声データおよび第 2 音声データの各々に重要度および変換部 3 5 によって変換された文字情報を割り当てて記録部 3 4 e へ記録していたが、これに限定されることなく、例えば複数の利用者の各々の位置を検出し、この検出結果と各利用者の注視度とに基づいて、第 1 音声データおよび第 2 音声データの各々に重要度および変換部 3 5 によって変換された文字情報を割り当てて記録部 3 4 e へ記録してもよい。

【0180】

（その他の実施の形態）

上述した実施の形態 1 ~ 5 に開示されている複数の構成要素を適宜組み合わせることによって、種々の発明を形成することができる。例えば、上述した実施の形態 1 ~ 5 に記載した全構成要素からいくつかの構成要素を削除してもよい。さらに、上述した実施の形態 1 ~ 5 で説明した構成要素を適宜組み合わせてもよい。

【0181】

また、実施の形態 1 ~ 5 において、上述してきた「部」は、「手段」や「回路」などに読み替えることができる。例えば、制御部は、制御手段や制御回路に読み替えることができる。

【0182】

また、実施の形態 1 ~ 5 に係る情報処理装置に実行させるプログラムは、インストール

10

20

30

40

50

可能な形式または実行可能な形式のファイルデータでＣＤ－ＲＯＭ、フレキシブルディスク（ＦＤ）、ＣＤ－Ｒ、ＤＶＤ（Digital Versatile Disk）、ＵＳＢ媒体、フラッシュメモリ等のコンピュータで読み取り可能な記録媒体に記録されて提供される。

【０１８３】

また、実施の形態１～５に係る情報処理装置に実行させるプログラムは、インターネット等のネットワークに接続されたコンピュータ上に格納し、ネットワーク経由でダウンロードさせることにより提供するように構成してもよい。さらに、実施の形態１～５に係る情報処理装置に実行させるプログラムをインターネット等のネットワーク経由で提供または配布するようにしてもよい。

【０１８４】

また、実施の形態１～５では、伝送ケーブルを経由して各種機器から信号を送信していたが、例えば有線である必要はなく、無線であってもよい。この場合、所定の無線通信規格（例えばＷｉ－Ｆｉ（登録商標）やＢｌｕｅｔｏｏｔｈ（登録商標））に従って、各機器から信号を送信するようにすればよい。もちろん、他の無線通信規格に従って無線通信を行ってもよい。

【０１８５】

なお、本明細書におけるフローチャートの説明では、「まず」、「その後」、「続いて」等の表現を用いてステップ間の処理の前後関係を明示していたが、本発明を実施するために必要な処理の順序は、それらの表現によって一意的に定められるわけではない。即ち、本明細書で記載したフローチャートにおける処理の順序は、矛盾のない範囲で変更することができる。

【０１８６】

以上、本願の実施の形態のいくつかを図面に基づいて詳細に説明したが、これらは例示であり、本発明の開示の欄に記載の態様を始めとして、当業者の知識に基づいて種々の変形、改良を施した他の形態で本発明を実施することが可能である。

【符号の説明】

【０１８７】

- １ａ，７００ 情報処理システム
- １０，１０ａ，１ｂ，１ｂ，１ｃ，１ｄ，１ｅ 情報処理装置
- １１ 解析部
- １２，１２ａ，３８，３８ｃ，３８ｄ，３８ｅ 設定部
- １３，３９，３９ｄ 生成部
- １４，３４，３４ｃ，３４ｄ，３４ｅ 記録部
- １５，３２３ 表示制御部
- ２０ 表示部
- ３０，２２２，５１０ 視線検出部
- ３１，５２０ 音声入力部
- ３２，３２ｃ，３２ｄ，３２ｅ 制御部
- ３３ 時間計測部
- ３５ 変換部
- ３６ 抽出部
- ３７ 操作部
- ４０ 画像処理部
- １００ 顕微鏡システム
- ２００ 顕微鏡
- ２０１ 本体部
- ２０２ 回転部
- ２０３ 昇降部
- ２０４ レボルバ
- ２０５ 対物レンズ

10

20

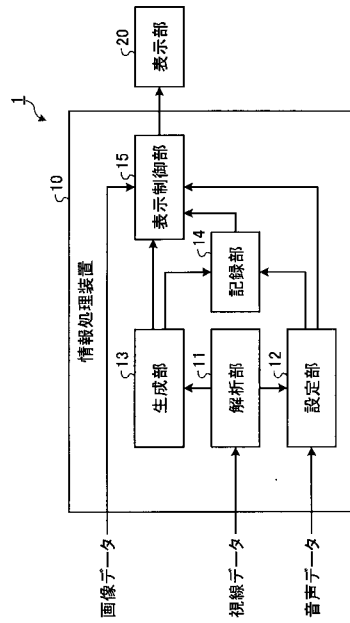
30

40

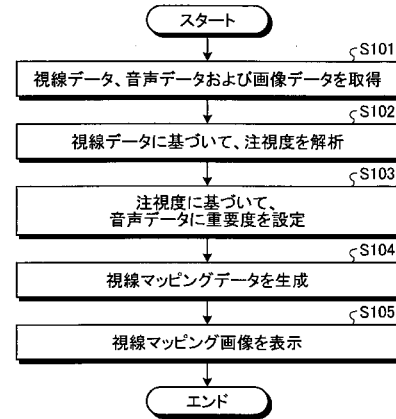
50

2 0 6	倍率検出部	
2 0 7	鏡筒部	
2 0 8	接続部	
2 0 9	接眼部	
2 1 0 , 4 0 1	撮像部	
3 0 0	内視鏡システム	
3 2 1	視線検出制御部	
3 2 2	音声入力制御部	
3 2 4	撮影制御部	
3 2 5	倍率算出部	10
3 2 6	操作履歴検出部	
3 2 7	識別検出制御部	
3 4 1	視線データ記録部	
3 4 2	音声データ記録部	
3 4 3 , 3 4 5	画像データ記録部	
3 4 4	プログラム記録部	
3 4 5	画像データ記録部	
3 4 6	操作履歴記録部	
3 4 7	識別情報記録部	
4 0 0	内視鏡	20
4 0 2	操作部	
5 0 0	ウェアラブルデバイス	
6 0 0	入力部	
7 1 0	第1ウェアラブルデバイス	
7 1 1	第1視線検出部	
7 1 2	第1音声入力部	
7 2 0	第2ウェアラブルデバイス	
7 2 1	第2視線検出部	
7 2 2	第2音声入力部	
7 3 0	検出部	30

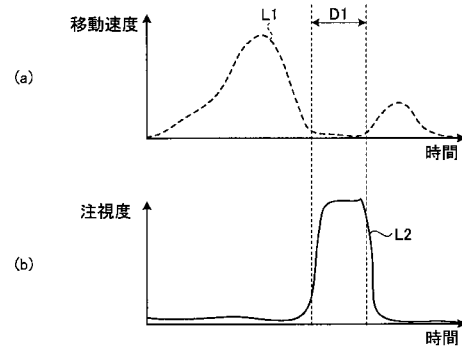
【図 1】



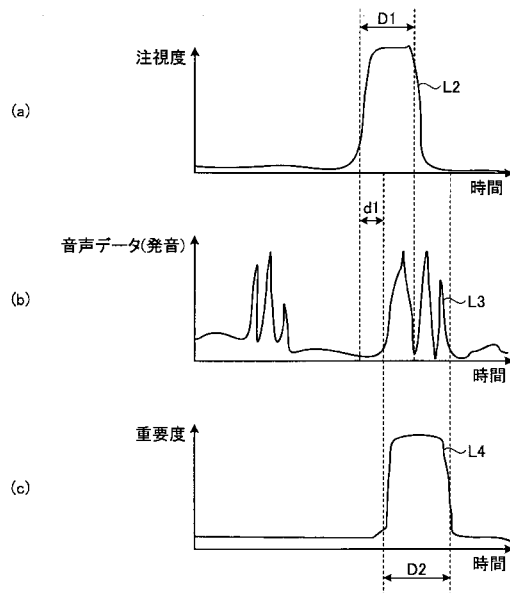
【図 2】



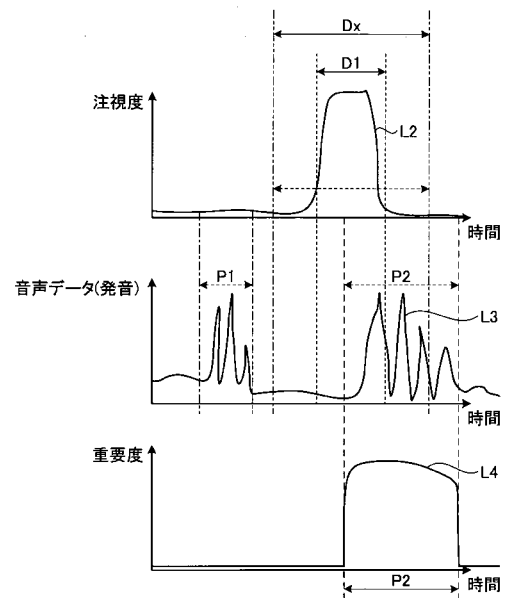
【図 3】



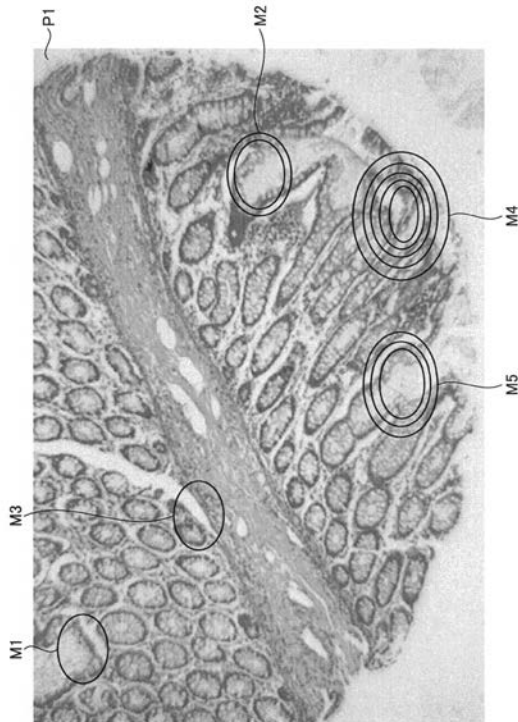
【図 4 A】



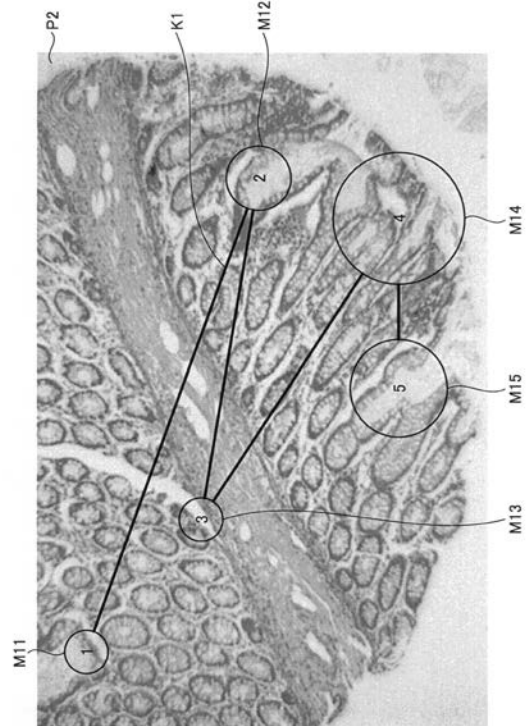
【図 4 B】



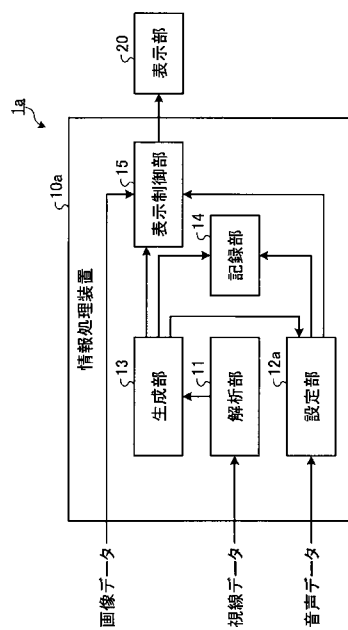
【図 5】



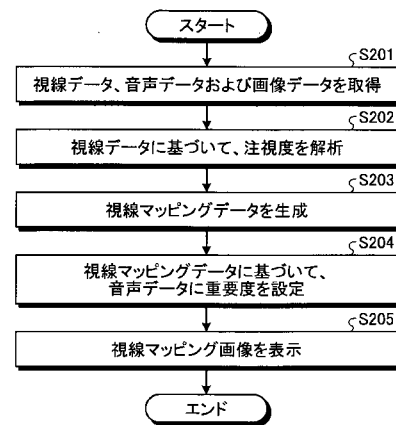
【図 6】



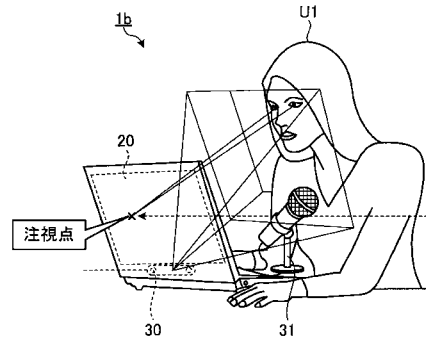
【図 7】



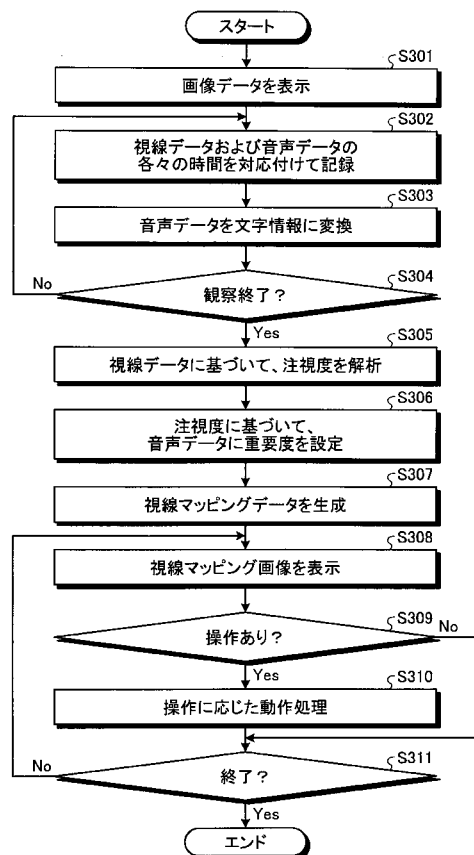
【図 8】



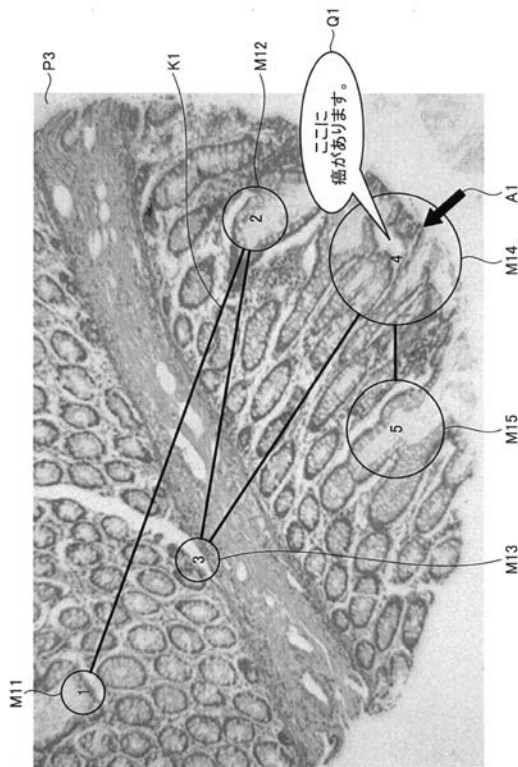
【 図 1 0 】



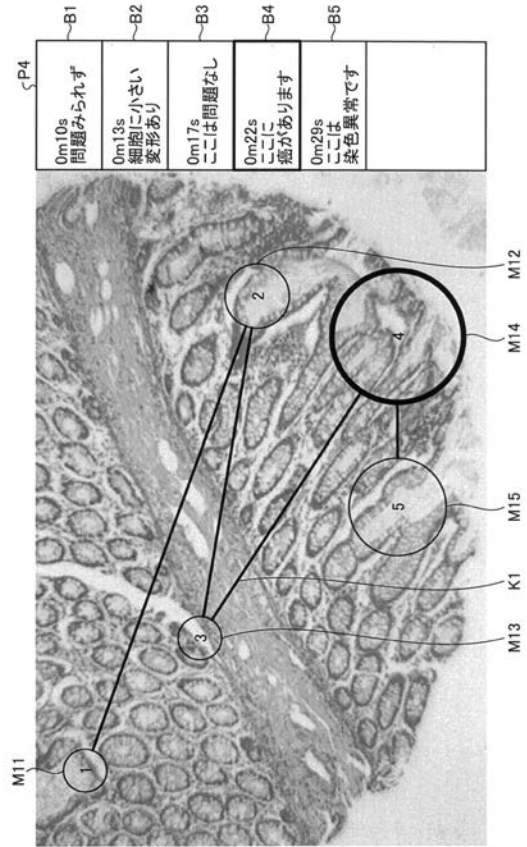
【 図 1 2 】



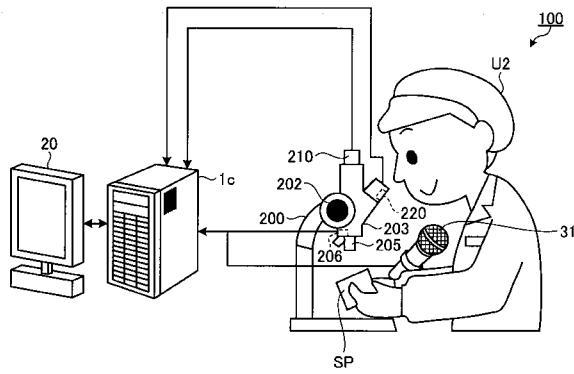
【図 13】



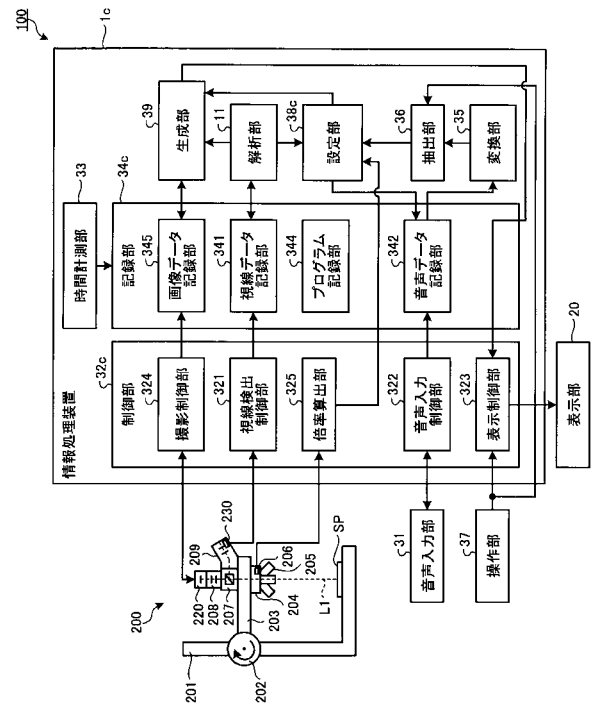
【図 14】



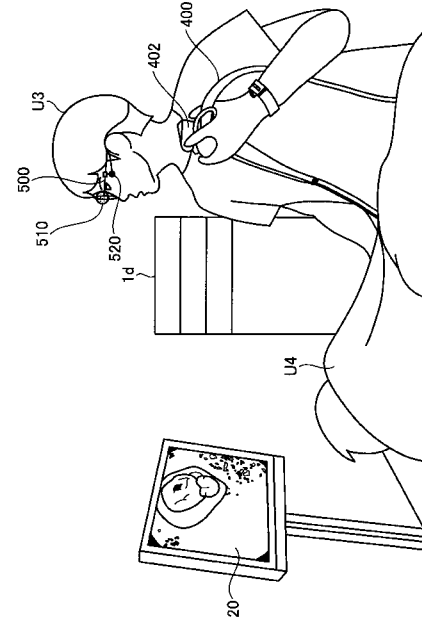
【図 15】



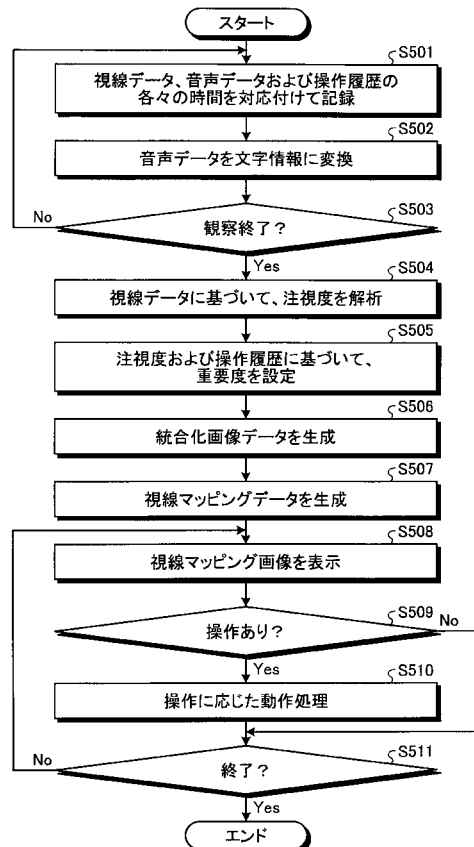
【図 16】



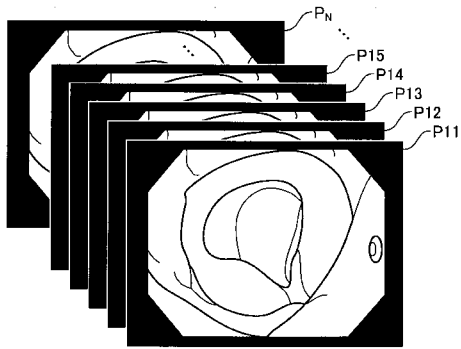
【 図 1 8 】



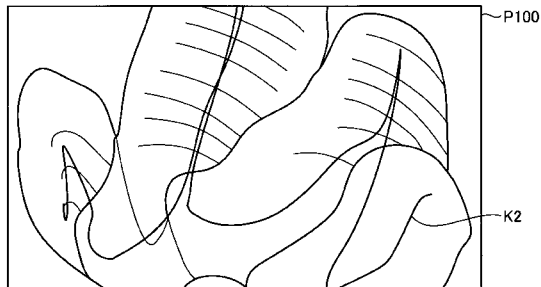
【 図 2 0 】



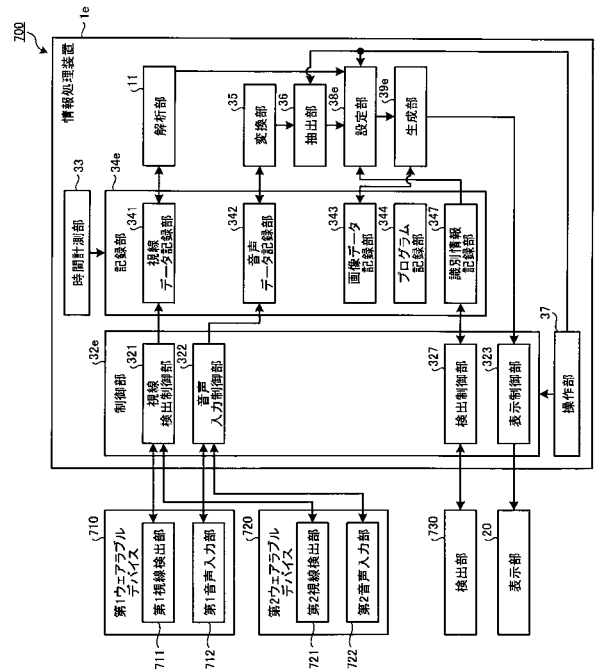
【 図 2 1 】



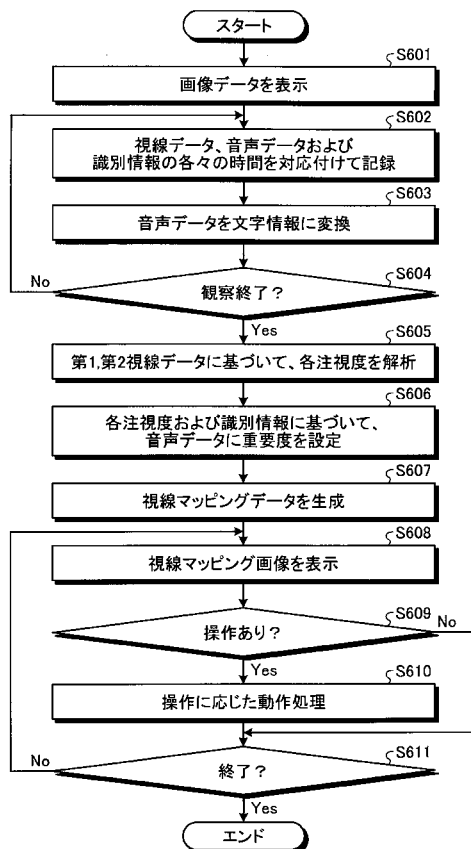
【 図 2 2 】



【 図 2 3 】



【 図 2 4 】



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	3/0346	(2013.01)	G 0 6 F	3/0346	4 2 3	5 C 1 2 2
A 6 1 B	3/113	(2006.01)	A 6 1 B	3/113		
A 6 1 B	1/00	(2006.01)	A 6 1 B	1/00	7 3 0	

(72)発明者 西村 英敏

東京都八王子市石川町 2 9 5 1 番地 オリンパス株式会社内

F ターム(参考) 2H040 DA11 DA12 DA21 GA02 GA11
2H052 AF02 AF14 AF21
4C161 CC06 DD03 FF11 LL02 NN01 PP12 RR06 WW04 YY12
4C316 AA21 AB16 FC28 FZ01 FZ03
5B087 BC12 BC32 DD03 DD17 DD18
5C122 DA26 EA42 FH09 FH11 FH18 FJ01 FJ15 FK41 GA18 GA34
HA01

专利名称(译)	信息处理设备，信息处理方法和程序		
公开(公告)号	JP2019204503A	公开(公告)日	2019-11-28
申请号	JP2019088251	申请日	2019-05-08
[标]申请(专利权)人(译)	奥林巴斯株式会社		
申请(专利权)人(译)	奥林巴斯公司		
[标]发明人	堀内一仁 渡辺伸之 金子善興 西村英敏		
发明人	堀内 一仁 渡辺 伸之 金子 善興 西村 英敏		
IPC分类号	G06F3/16 G02B21/36 G02B23/24 H04N5/232 H04N5/225 G06F3/0346 A61B3/113 A61B1/00		
FI分类号	G06F3/16.610 G02B21/36 G02B23/24.B H04N5/232.300 H04N5/225.500 G06F3/0346.423 A61B3/113 A61B1/00.730		
F-TERM分类号	2H040/DA11 2H040/DA12 2H040/DA21 2H040/GA02 2H040/GA11 2H052/AF02 2H052/AF14 2H052/AF21 4C161/CC06 4C161/DD03 4C161/FF11 4C161/LL02 4C161/NN01 4C161/PP12 4C161/RR06 4C161/WW04 4C161/YY12 4C316/AA21 4C316/AB16 4C316/FC28 4C316/FZ01 4C316/FZ03 5B087/BC12 5B087/BC32 5B087/DD03 5B087/DD17 5B087/DD18 5C122/DA26 5C122/EA42 5C122/FH09 5C122/FH11 5C122/FH18 5C122/FJ01 5C122/FJ15 5C122/FK41 5C122/GA18 5C122/GA34 5C122/HA01		
优先权	2018095519 2018-05-17 JP		
外部链接	Espacenet		

摘要(译)

为了提供一种信息处理设备，信息处理方法和程序，其能够掌握语音数据中的哪个周期是重要的。解决方案：信息处理设备10包括：分析单元11，其分析用户的固定程度。根据通过检测用户从外部输入的视线获得的视线数据获得的视线；设置单元12，以预定的时间间隔，根据由分析单元11分析的固定程度，将重要性分配给从外部输入的用户语音数据，该语音数据的时间轴与从外部输入的语音轴的时间轴相同。关联视线数据，并将其记录在记录单元14中。选定的图纸：图1

