

(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2019-202131

(P2019-202131A)

(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) |
| A 6 1 B 3/113 (2006.01) | A 6 1 B 3/113 | 2 H 0 4 0 |
| G 0 6 F 3/01 (2006.01) | G 0 6 F 3/01 5 1 0 | 2 H 0 5 2 |
| G 0 2 B 21/00 (2006.01) | G 0 2 B 21/00 | 4 C 3 1 6 |
| G 0 2 B 23/24 (2006.01) | G 0 2 B 23/24 A | 5 E 5 5 5 |
| G 0 6 F 3/16 (2006.01) | G 0 6 F 3/16 5 2 0 | |
| 審査請求 未請求 請求項の数 20 O L (全 44 頁) | | |

(21) 出願番号 特願2019-90107 (P2019-90107)
 (22) 出願日 令和1年5月10日 (2019.5.10)
 (31) 優先権主張番号 特願2018-95520 (P2018-95520)
 (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

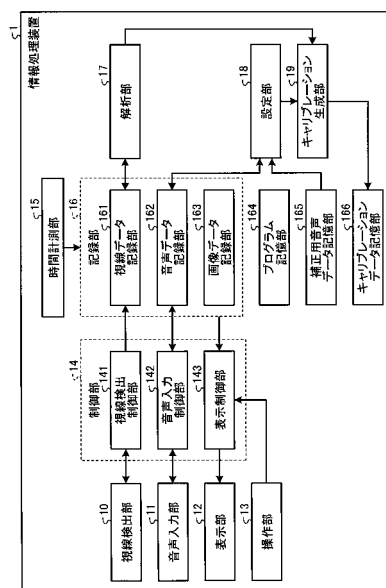
(54) 【発明の名称】 情報処理装置、情報処理方法およびプログラム

(57) 【要約】

【課題】視線の注視点と音声が発せられた時間との時間差を補正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1は、視線データに基づいて、表示部12が表示する補正用画像の複数の観察箇所の各々に対する利用者の視線の注視度が所定の値以上となる注視期間を解析する解析部17と、補正用音声データ記憶部165が記録する補正用音声データに基づいて、音声データに対して音声が発せられた時間を重要音声期間として設定する設定部18と、注視期間と重要音声期間との時間差を利用者のキャリブレーションデータとして生成して記録部16へ記録するキャリブレーション生成部19と、を備える。

【選択図】図1



【特許請求の範囲】**【請求項 1】**

複数の観察箇所の各々の座標位置が設定された補正用画像を表示する表示部と、
前記補正用画像の前記複数の観察箇所の各々で発声される音声が発設定された補正用音声データを記録する補正用音声データ記録部と、
利用者の視線を連続的に検出することによって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ることによ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を生成する音声入力部と、
前記視線検出部によって生成された視線データに基づいて、前記複数の観察箇所の各々に対する前記利用者の視線の注視度が所定の値以上となる注視期間を解析する解析部と、
前記補正用音声データ記録部が記録する前記補正用音声データに基づいて、前記音声データに対して前記音声が発声された期間を重要音声期間として設定する設定部と、
前記注視期間と前記重要音声期間との時間差に基づき前記利用者のキャリブレーションデータを生成して記録部へ記録するキャリブレーション生成部と、
を備える情報処理装置。

10

【請求項 2】

前記キャリブレーション生成部は、前記注視期間と前記重要音声期間との時間差を複数回算出し、該複数回の算出結果の統計的な特徴に基づき、前記キャリブレーションデータを生成する

請求項 1 に記載の情報処理装置。

20

【請求項 3】

前記キャリブレーションデータは、前記視線データおよび前記音声データのどちらか一方の開始時刻または終了時刻を基準にしたときの時刻差にかかわるものである

請求項 1 または 2 に記載の情報処理装置。

【請求項 4】

前記キャリブレーションデータは、前記視線データおよび前記音声データのどちらか一方を基準にしたときの期間の長さにかかわるものである

請求項 1 または 2 に記載の情報処理装置。

【請求項 5】

前記キャリブレーション生成部は、前記注視期間の開始時刻と前記重要音声期間の開始時刻との時間差または前記注視期間の終了時間と前記重要音声期間の終了時刻の時間差を前記キャリブレーションデータとして生成する

請求項 1 ～ 4 のいずれか一つに記載の情報処理装置。

30

【請求項 6】

前記設定部は、外部から入力された操作信号に応じて予め指定されたキーワードが発声された期間を前記重要音声期間としてさらに設定する

請求項 1 ～ 5 のいずれか一つに記載の情報処理装置。

【請求項 7】

前記視線検出部は、

前記キャリブレーション生成部が前記キャリブレーションデータを生成した後に、前記利用者の視線を連続的に検出することによって新たな視線データを生成し、

前記音声入力部は、

前記キャリブレーション生成部が前記キャリブレーションデータを生成した後に、前記利用者の音声の入力を受け付けることによって前記新たな視線データと同じ時間軸が対応付けられた新たな音声データを生成し、

前記解析部は、前記新たな視線データに基づいて、前記利用者の視線の新たな注視度が所定の値以上となる新たな注視期間を解析し、

前記設定部は、前記キャリブレーションデータに基づいて、前記新たな注視期間と前記新たな音声データに含まれる新たな重要音声期間の時間差を補正する

請求項 1 ～ 6 のいずれか一つに記載の情報処理装置。

40

50

【請求項 8】

前記設定部は、前記新たな音声データに対して、前記新たな注視度と前記キャリブレーションデータとを用いて前記新たな注視度に応じた重要度を割り当てて前記記録部へ記録する

請求項 7 に記載の情報処理装置。

【請求項 9】

前記設定部は、

前記新たな音声データに対して、前記新たな音声データの中で発声された発声期間を設定し、かつ、前記新たな音声データに対して、前記発声期間と前記新たな注視度と前記キャリブレーションデータとを用いて前記新たな注視度に応じた重要度を割り当てて前記記録部へ記録する

請求項 7 に記載の情報処理装置。

【請求項 10】

前記設定部は、前記新たな音声データの中で所定のキーワードが発声された期間を新たな重要発声期間として設定し、かつ、前記新たな視線データに対して、前記新たな重要発声期間と前記キャリブレーションデータとを用いて前記新たな重要発声期間に応じた重要度を割り当てて前記記録部へ記録する

請求項 7 に記載の情報処理装置。

【請求項 11】

前記設定部は、前記新たな音声データの中で所定のキーワードが発声された期間を新たな重要発声期間として設定し、かつ、前記新たな視線データに対して、前記新たな注視期間と前記新たな重要発声期間と前記キャリブレーションデータとを用いて前記重要発声期間に応じた重要度を割り当てて前記記録部へ記録する

請求項 7 に記載の情報処理装置。

【請求項 12】

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一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 8 ～ 11 のいずれか一つに記載の情報処理装置。

【請求項 13】

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新たな注視度および該新たな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 8 ～ 11 のいずれか一つに記載の情報処理装置。

【請求項 14】

前記解析部は、前記新たな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 13 に記載の情報処理装置。

【請求項 15】

前記新たな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 13 または 14 に記載の情報処理装置。

【請求項 16】

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、

10

20

30

40

50

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 13 に記載の情報処理装置。

【請求項 17】

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 13 に記載の情報処理装置。

【請求項 18】

前記新たな注視期間に該当する前記新たな注視度を有する前記新たな視線データと、前記新たな重要音声期間に該当する前記重要度を有する前記新たな音声データと、を関連付けたデータを生成する生成部をさらに備える、

請求項 8 ～ 11 のいずれか一つに記載の情報処理装置。

【請求項 19】

複数の観察箇所の各々の座標位置が設定された補正用画像を表示する表示ステップと、利用者の視線を連続的に検出することによって視線データを生成する視線検出ステップと、

前記利用者の音声の入力を受け付けることによ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を生成する音声入力ステップと、

前記視線検出ステップによって生成された視線データに基づいて、前記複数の観察箇所の各々に対する前記利用者の視線の注視度が所定の値以上となる注視期間を解析する解析ステップと、

前記補正用画像の前記複数の観察箇所の各々で発声される音声の設定された補正用音声データを記録する補正用音声データ記録部が記録する前記補正用音声データに基づいて、前記音声データに対して前記音声が発声された時間を重要音声期間として設定する設定ステップと、

前記注視期間と前記重要音声期間との時間差を前記利用者のキャリブレーションデータとして生成して記録部へ記録するキャリブレーション生成ステップと、

を含む情報処理方法。

【請求項 20】

情報処理装置に、

複数の観察箇所の各々の座標位置が設定された補正用画像を表示する表示ステップと、利用者の視線を連続的に検出することによって視線データを生成する視線検出ステップと、

前記利用者の音声の入力を受け付けることによ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を生成する音声入力ステップと、

前記視線検出ステップによって生成された視線データに基づいて、前記複数の観察箇所の各々に対する前記利用者の視線の注視度が所定の値以上となる注視期間を解析する解析ステップと、

前記補正用画像の前記複数の観察箇所の各々で発声される音声の設定された補正用音声データを記録する補正用音声データ記録部が記録する前記補正用音声データに基づいて、前記音声データに対して前記音声が発声された時間を重要音声期間として設定する設定ステップと、

前記注視期間と前記重要音声期間との時間差を前記利用者のキャリブレーションデータとして生成して記録部へ記録するキャリブレーション生成ステップと、

を実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

10

20

30

40

50

本開示は、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】

10

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】

20

【発明が解決しようとする課題】

【0005】

ところで、使用者によっては、視線の注視点と音声を発する時間とが異なる場合がある。しかしながら、上述した特許文献1、2では、視線の注視点と音声が発せられた時間との時間差については何ら考慮されていなかった。

【0006】

本発明は、上記に鑑みてなされたものであって、視線の注視点と音声が発せられた時間との時間差を使用者に合わせて補正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】

30

【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、複数の観察箇所の各々の座標位置が設定された補正用画像を表示する表示部と、前記補正用画像の前記複数の観察箇所の各々で発声される音声が発設定された補正用音声データを記録する補正用音声データ記録部と、利用者の視線を連続的に検出することによって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ることによ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を生成する音声入力部と、前記視線検出部によって生成された視線データに基づいて、前記複数の観察箇所の各々に対する前記利用者の視線の注視度が所定の値以上となる注視期間を解析する解析部と、前記補正用音声データ記録部が記録する前記補正用音声データに基づいて、前記音声データに対して前記音声が発声された期間を重要音声期間として設定する設定部と、前記注視期間と前記重要音声期間との時間差に基づき前記利用者のキャリブレーションデータを生成して記録部へ記録するキャリブレーション生成部と、を備える。

40

【0008】

また、本開示に係る情報処理装置は、上記開示において、前記キャリブレーション生成部は、前記注視期間と前記重要音声期間との時間差を複数回算出し、該複数回の算出結果の統計的な特徴に基づき、前記キャリブレーションデータを生成する。

【0009】

また、本開示に係る情報処理装置は、上記開示において、前記キャリブレーションデータは、前記視線データおよび前記音声データのどちらか一方の開始時刻または終了時刻を

50

基準にしたときの時刻差にかかわるものである。

【 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、前記キャリブレーションデータは、前記視線データおよび前記音声データのどちらか一方を基準にしたときの期間の長さにかかわるものである。

【 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、前記キャリブレーション生成部は、前記注視期間の開始時刻と前記重要音声期間の開始時刻との時間差または前記注視期間の終了時間と前記音声重要期間の終了時刻の時間差を前記キャリブレーションデータとして生成する。

10

【 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、前記設定部は、外部から入力された操作信号に応じて予め指定されたキーワードが発声された期間を前記重要音声期間としてさらに設定する。

【 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、前記視線検出部は、前記キャリブレーション生成部が前記キャリブレーションデータを生成した後に、前記利用者の視線を連続的に検出することによって得られる新たな視線データを取得し、前記音声入力部は、前記キャリブレーション生成部が前記キャリブレーションデータを生成した後に、前記利用者の音声の入力を受け付けることによって得られ、前記新たな視線データと同じ時間軸が対応付けられた新たな音声データを生成し、前記解析部は、前記新たな視線データに基づいて、前記利用者の視線の新たな注視度が所定の値以上となる新たな注視期間を解析し、前記設定部は、前記キャリブレーションデータに基づいて、前記新たな注視期間と前記新たな音声データに含まれる新たな重要音声期間の時間差を補正する。

20

【 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、前記設定部は、前記新たな音声データに対して、前記新たな注視度と前記キャリブレーションデータとを用いて前記新たな注視度に応じた重要度を割り当てて前記記録部へ記録する。

【 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、前記設定部は、前記新たな音声データに対して、前記新たな音声データの中で発声された発声期間を設定し、かつ、前記新たな音声データに対して、前記発声期間と前記新たな注視度と前記キャリブレーションデータとを用いて前記新たな注視度に応じた重要度を割り当てて前記記録部へ記録する。

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、前記設定部は、前記新たな音声データの中で所定のキーワードが発声された期間を新たな重要発声期間として設定し、かつ、前記新たな視線データに対して、前記新たな重要発声期間と前記キャリブレーションデータとを用いて前記新たな重要発声期間に応じた重要度を割り当てて前記記録部へ記録する。

40

【 0 0 1 7 】

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、前記新たな音声データの中で所定のキーワードが発声された期間を新たな重要発声期間として設定し、かつ、前記新たな視線データに対して、前記新たな注視期間と前記新たな重要発声期間と前記キャリブレーションデータとを用いて前記重要発声期間に応じた重要度を割り当てて前記記録部へ記録する。

【 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、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一つを検出することによって、前記注視度を解析する。

50

【 0 0 1 9 】

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新たな注視度および該新たな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。

【 0 0 2 0 】

また、本開示に係る情報処理装置は、上記開示において、前記解析部は、前記新たな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 2 1 】

また、本開示に係る情報処理装置は、上記開示において、前記新たな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 2 2 】

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。

【 0 0 2 3 】

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。

【 0 0 2 4 】

また、本開示に係る情報処理装置は、上記開示において、前記新たな注視期間に該当する前記新たな注視度を有する前記新たな視線データと、前記新たな重要音声期間に該当する前記重要度を有する前記新たな音声データと、を関連付けたデータを生成する生成部をさらに備える。

【 0 0 2 5 】

また、本開示に係る情報処理方法は、複数の観察箇所の各々の座標位置が設定された補正用画像を表示する表示ステップと、利用者の視線を連続的に検出することによって視線データを生成する視線検出ステップと、前記利用者の音声の入力を受け付けることによ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を生成する音声入力ステップと、前記視線検出ステップによって生成された視線データに基づいて、前記複数の観察箇所の各々に対する前記利用者の視線の注視度が所定の値以上となる注視期間を解析する解析ステップと、前記補正用画像の前記複数の観察箇所の各々で発声される音声が発設定された補正用音声データを記録する補正用音声データ記録部が記録する前記補正用音声データに基づいて、前記音声データに対して前記音声が発声された時間を重要音声期間として設定する設定ステップと、前記注視期間と前記重要音声期間との時間差を前記利用者のキャリブレーションデータとして生成して記録部へ記録するキャリブレーション生成ステップと、を含む。

【 0 0 2 6 】

また、本開示に係るプログラムは、情報処理装置に、複数の観察箇所の各々の座標位置が設定された補正用画像を表示する表示ステップと、利用者の視線を連続的に検出することによって視線データを生成する視線検出ステップと、前記利用者の音声の入力を受け付けることによ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を生成する音声入力ステップと、前記視線検出ステップによって生成された視線データに基づいて、前記複数の観察箇所の各々に対する前記利用者の視線の注視度が所定の値以上となる注視期間を解析する解析ステップと、前記補正用画像の前記複数の観察箇所の各々で発声される

10

20

30

40

50

音声が発せられた補正用音声データを記録する補正用音声データ記録部が記録する前記補正用音声データに基づいて、前記音声データに対して前記音声が発せられた時間を重要音声期間として設定する設定ステップと、前記注視期間と前記重要音声期間との時間差を前記利用者のキャリブレーションデータとして生成して記録部へ記録するキャリブレーション生成ステップと、を実行させる。

【発明の効果】

【0027】

本開示によれば、視線の注視点と音声が発せられた時間との時間差を補正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】

10

【0028】

【図1】図1は、本開示の実施の形態1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2】図2は、本開示の実施の形態1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3】図3は、本開示の実施の形態1に係る解析部による視線の注視度が所定の値以上となる注視期間の解析方法を模式的に説明する図である。

【図4】図4は、本開示の実施の形態1に係る設定部による音声データの重要音声期間の設定方法を模式的に説明する図である。

【図5A】図5Aは、本開示の実施の形態1に係るキャリブレーション生成部によるキャリブレーションデータの生成方法を模式的に説明する図である。

20

【図5B】図5Bは、本開示の実施の形態1に係るキャリブレーション生成部によるキャリブレーションデータの別の生成方法を模式的に説明する図である。

【図6A】図6Aは、本開示の実施の形態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6B】図6Bは、本開示の実施の形態2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7A】図7Aは、本開示の実施の形態2の変形例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7B】図7Bは、本開示の実施の形態2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。

30

【図8】図8は、本開示の実施の形態3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9】図9は、本開示の実施の形態3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10】図10は、本開示の実施の形態3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11】図11は、本開示の実施の形態3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である。

【図12】図12は、本開示の実施の形態3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。

40

【図13】図13は、本開示の実施の形態3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14】図14は、本開示の実施の形態4に係る情報処理装置の構成を示す概略図である。

【図15】図15は、本開示の実施の形態4に係る情報処理装置の構成を示す概略図である。

【図16】図16は、本開示の実施の形態4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17】図17は、本開示の実施の形態4に係る情報処理装置が実行する処理の概要を

50

示すフローチャートである。

【図 1 8】図 1 8 は、本開示の実施の形態 4 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 1 9 A】図 1 9 A は、本開示の実施の形態 4 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 1 9 B】図 1 9 B は、本開示の実施の形態 4 の変形例に係る視線データと音声データの関連付けられたデータの一例を示す図である。

【図 2 0】図 2 0 は、本開示の実施の形態 5 に係る顕微鏡システムの構成を示す概略図である。

【図 2 1】図 2 1 は、本開示の実施の形態 5 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 2 2】図 2 2 は、本開示の実施の形態 5 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 2 3】図 2 3 は、本開示の実施の形態 6 に係る内視鏡システムの構成を示す概略図である。

【図 2 4】図 2 4 は、本開示の実施の形態 6 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 2 5】図 2 5 は、本開示の実施の形態 6 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 2 6】図 2 6 は、本開示の実施の形態 6 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 2 7】図 2 7 は、本開示の実施の形態 6 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 2 8】図 2 8 は、本開示の実施の形態 7 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 9】図 2 9 は、本開示の実施の形態 7 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】

【0029】

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。

【0030】

(実施の形態 1)

〔情報処理装置の構成〕

図 1 は、実施の形態 1 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 1 に示す情報処理装置 1 は、視線検出部 10 と、音声入力部 11 と、表示部 12 と、操作部 13 と、制御部 14 と、時間計測部 15 と、記録部 16 と、解析部 17 と、設定部 18 と、キャリブレーション生成部 19 と、プログラム記憶部 164 と、補正用音声データ記憶部 165 と、キャリブレーションデータ記憶部 166 と、を備える。

【0031】

視線検出部 10 は、近赤外線を照射する LED 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 CMOS、CCD 等）と、を用いて構成される。視線検出部 10 は、制御部 14 の制御のもと、表示部 12 が表示する補正用画像に対する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、この視線データを制御部 14 へ出力する。具体的には、視線検出部 10 は、制御部 14 の制御のもと、LED 光源等から近赤外線を利用者の角膜に照射し、光学センサが利用者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 10 は、制御部 14 の制御のもと、光学

10

20

30

40

50

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 1 4 1 へ出力する。なお、視線検出部 1 0 は、単に光学センサのみで利用者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 3 2 】

音声入力部 1 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 1 4 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 1 1 は、制御部 1 4 の制御のもと、利用者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 1 4 へ出力する。なお、音声入力部 1 1 は、音声の入力以外にも、音声を出力することができるスピーカ等を設け、音声出力機能を設けてもよい。

【 0 0 3 3 】

表示部 1 2 は、制御部 1 4 の制御のもと、複数の観察箇所の各々の座標位置が設定された補正用画像を表示する。表示部 1 2 は、液晶または有機 E L (Electro Luminescence) 等の表示パネル等を用いて構成される。

【 0 0 3 4 】

操作部 1 3 は、情報処理装置 1 に関する各種操作の入力を受け付ける。操作部 1 3 は、例えば、スイッチ、タッチパネル、キーボードおよびマウス等を用いて構成される。

【 0 0 3 5 】

制御部 1 4 は、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、視線検出部 1 0、音声入力部 1 1 および表示部 1 2 を制御する。制御部 1 4 は、視線検出制御部 1 4 1 と、音声入力制御部 1 4 2 と、表示制御部 1 4 3 と、を有する。

【 0 0 3 6 】

視線検出制御部 1 4 1 は、視線検出部 1 0 を制御する。具体的には、視線検出制御部 1 4 1 は、視線検出部 1 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 1 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 1 4 1 は、視線検出部 1 0 から入力された視線データに対して、各種の画像処理を行って記録部 1 6 へ出力する。

【 0 0 3 7 】

音声入力制御部 1 4 2 は、音声入力部 1 1 を制御し、音声入力部 1 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 1 6 へ出力する。

【 0 0 3 8 】

表示制御部 1 4 3 は、表示部 1 2 の表示態様を制御する。表示制御部 1 4 3 は、記録部 1 6 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 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 1 2 に表示させる。

【 0 0 3 9 】

時間計測部 1 5 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 1 0 によって生成された視線データおよび音声入力部 1 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。

【 0 0 4 0 】

記録部 1 6 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 1 6 は、視線データ記録部 1 6 1 と、音声データ記録部 1 6 2 と、画像データ記録部 1 6 3 と、を有する。

【 0 0 4 1 】

視線データ記録部 1 6 1 は、視線検出制御部 1 4 1 から入力された視線データを記録す

10

20

30

40

50

るとともに、視線データを解析部 17 へ出力する。

【0042】

音声データ記録部 162 は、音声入力制御部 142 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 135 へ出力する。

【0043】

画像データ記録部 163 は、複数の画像データを記録する。具体的には、画像データ記録部 163 は、複数の観察箇所の各々の座標位置が設定された補正用画データを記録する。

【0044】

プログラム記憶部 164 は、情報処理装置 1b が実行する各種プログラム、各種プログ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憶する。

【0045】

補正用音声データ記憶部 165 は、表示部 12 が表示する補正用画像の複数の観察箇所の各々で発声される音声が発声された補正用音声データを記憶する。

【0046】

キャリブレーションデータ記憶部 166 は、後述するキャリブレーション生成部 19 によって生成されたキャリブレーションデータを記憶する。

【0047】

解析部 17 は、視線データ記録部 161 が記録する視線データに基づいて、表示部 12 が表示する補正用画像上における複数の観察箇所の各々に対する利用者の視線の注視度が所定の値以上となる注視期間を解析する。具体的には、解析部 17 は、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）の注視度が所定の値以上となる注視期間を解析する。解析部 17 は、例えば CPU、FPGA および GPU 等を用いて構成される。

【0048】

設定部 18 は、補正用音声データ記憶部 165 が記録する補正用音声データに基づいて、音声データ記録部 162 が記録する音声データに対して補正用画像における複数の観察箇所の各々で発声される音声が発声された期間（時間）を重要音声期間として設定する。設定部 18 は、CPU、FPGA および GPU 等を用いて構成される。

【0049】

キャリブレーション生成部 19 は、解析部 17 によって解析された注視期間と設定部 18 によって設定された重要音声期間との時間差を利用者のキャリブレーションデータとして生成して記録部 16 のキャリブレーションデータ記憶部 166 へ記録する。キャリブレーションデータは、前記視線データおよび前記音声データのどちらか一方の開始時刻または終了時刻を基準にしたときの時刻差、または視線データおよび音声データのどちらか一方を基準にしたときの期間の長さにかかわるものである。なお、キャリブレーション生成部 19 は、注視期間と重要音声期間との時間差を複数回算出し、この複数回の算出結果の統計的な特徴に基づき、キャリブレーションデータを生成する。また、キャリブレーション生成部 19 は、注視期間の開始時刻と重要期間の開始時刻との時間差または注視期間の終了時間と音声重要期間の終了時刻の時間差をキャリブレーションデータとして生成してもよい。例えば、校正者や利用者は、図示していないが、補正用の画像の表示と重要単語の両方を用意する。具体的には、顕微鏡の画像の場合、校正者や利用者は、あらかじめ所定の病変が有ることが分かっている画像を用意し、重要単語としては、その病変にかかわる用語（キーワード）を用意する。キャリブレーション生成部 19 は、使用者が病変に着目してから所定の重要単語を発するまでの時間差を計測することでキャリブレーションデータとしてもよい。また、キャリブレーション生成部 19 は、キャリブレーションデータの精度を上げるために、複数回の測定を行うことが望ましい。

【0050】

〔情報処理装置の処理〕

次に、情報処理装置 1 が実行する処理について説明する。図 2 は、情報処理装置 1 が実行する処理の概要を示すフローチャートである。

【0051】

図 2 に示すように、まず、表示制御部 143 は、画像データ記録部 163 が記録する補正用画像データに対応する補正用画像を表示部 12 に表示させる（ステップ S101）。

【0052】

続いて、制御部 14 は、視線検出部 10 によって生成された視線データおよび音声入力部 11 によって生成された音声データの各々と時間計測部 133 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 161 および音声データ記録部 162 に記録する（ステップ S102）。

10

【0053】

その後、操作部 13 から表示部 12 が表示する補正用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 S103：Yes）、情報処理装置 1 は、後述するステップ S103 へ移行する。これに対して、操作部 13 から表示部 12 が表示する補正用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 S103：No）、情報処理装置 1 は、上述したステップ S102 へ戻る。

【0054】

ステップ S104 において、解析部 17 は、視線データ記録部 161 が記録する視線データに基づいて、表示部 12 が表示する補正用画像上における複数の観察箇所の各々に対する利用者の視線の注視度が所定の値以上となる注視期間を解析する。ステップ S104 の後、情報処理装置 1 は、後述するステップ S105 へ移行する。

20

【0055】

図 3 は、解析部 17 による視線の注視度が所定の値以上となる注視期間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 L1 が視線の移動速度の時間変化を示し、図 3 の（b）の曲線 L2 が注視度の時間変化を示す。

【0056】

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 L1 および曲線 L2 に示すように、解析部 17 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注視期間 D1 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 17 は、利用者が補正用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し、かつ、この注視度が所定の値以上となる注視期間（例えば注視期間 D1）として解析する。なお、図 3 では、解析部 17 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。

30

40

【0057】

図 2 に戻り、ステップ S105 以降の説明を続ける。

ステップ S105 において、設定部 18 は、補正用音声データ記憶部 165 が記録する補正用音声データに基づいて、音声データ記録部 162 が記録する音声データに対して補正用画像における複数の観察箇所の各々で発声される音声が発声された時間を重要音声期間として設定する。ステップ S105 の後、情報処理装置 1 は、後述するステップ S106 へ移行する。

【0058】

図 4 は、設定部 18 による音声データの重要音声期間の設定方法を模式的に説明する図

50

である。図 4 において、横軸が時間を示し、図 4 の (a) の縦軸が音声データ (発音) を示し、図 4 の (b) の縦軸が音声重要度を示す。また、図 4 の (a) の曲線 L 3 が音声データの時間変化を示し、図 4 の (b) の曲線 L 4 が音声データの音声重要度の時間変化を示す。

【 0 0 5 9 】

図 4 の曲線 L 3 および曲線 L 4 に示すように、設定部 1 8 は、補正用音声データ記憶部 1 6 5 が記録する補正用音声データに基づいて、音声データ記録部 1 6 2 が記録する音声データに対して補正用画像における複数の観察箇所の各々で発声される音声が発声された期間 (時間) を重要音声期間として設定する。例えば、設定部 1 8 は、補正用音声データ記憶部 1 6 5 が記録する補正用音声データに「癌」がある場合、利用者が「癌」を発声した時間 (音声重要区間 D 2) を重要音声期間として設定する。

10

【 0 0 6 0 】

図 2 に戻り、ステップ S 1 0 6 以降の説明を続ける。

ステップ S 1 0 6 において、キャリブレーション生成部 1 9 は、キャリブレーションデータを生成する。ステップ S 1 0 6 の後、情報処理装置 1 は、後述するステップ S 1 0 7 へ移行する。

【 0 0 6 1 】

図 5 A は、キャリブレーション生成部 1 9 によるキャリブレーションデータの生成方法を模式的に説明する図である。図 5 A において、横軸が時間を示し、図 5 A の (a) の縦軸が注視度を示し、図 5 A の (b) の縦軸が音声重要度を示す。また、図 5 A の (a) の曲線 L 2 が注視度の時間変化を示し、図 5 A の (b) の曲線 L 4 が音声重要度の時間変化を示す。

20

【 0 0 6 2 】

図 5 A の曲線 L 2 および曲線 L 4 に示すように、キャリブレーション生成部 1 9 は、解析部 1 7 によって解析された注視期間 D 1 の開始時刻 t_1 と設定部 1 8 によって設定された重要音声期間 D 2 の開始時刻 t_2 との時間差 D 3 を利用者のキャリブレーションデータとして生成する。

【 0 0 6 3 】

また、この変形例として、使用者の注視度が高い期間と重要単語を発話する期間とのタイミングが測定毎にばらつく場合は、そのばらつき期間を考慮してもよい。すなわち、図 5 B で示すように、キャリブレーション生成部 1 9 は、ばらつきの大きさに応じて、時間枠を設定する前後のマージン D D 2 , D D 3 を調整し、ばらつきが大きいほどマージンを大きくしたキャリブレーションデータとする。

30

【 0 0 6 4 】

図 2 に戻り、ステップ S 1 0 7 以降の説明を続ける。

ステップ S 1 0 7 において、キャリブレーション生成部 1 9 は、ステップ S 1 0 6 で生成したキャリブレーションデータをキャリブレーションデータ記憶部 1 6 6 に記録する。ステップ S 1 0 7 の後、情報処理装置 1 は、本処理を終了する。

【 0 0 6 5 】

以上説明した実施の形態 1 によれば、キャリブレーション生成部 1 9 が解析部 1 7 によって解析された注視期間 D 1 と設定部 1 8 によって設定された重要音声期間 D 2 との時間差 D 3 を利用者のキャリブレーションデータとして生成するので、視線の注視点と音声が発せられた時間との時間差を補正することができる。

40

【 0 0 6 6 】

また、実施の形態 1 によれば、キャリブレーション生成部 1 9 が解析部 1 7 によって解析された注視期間 D 1 の開始時刻 t_1 と設定部 1 8 によって設定された重要音声期間 D 2 の開始時刻 t_2 との時間差 D 3 を利用者のキャリブレーションデータとして生成するので、視線の注視点と音声が発せられた時間との時間差をより正確に補正することができる。

【 0 0 6 7 】

なお、実施の形態 1 では、キャリブレーション生成部 1 9 が複数の注視期間と複数の重

50

要音声期間の各々の時間差を複数回算出し、この複数回の算出結果の統計的な特徴に基づき、キャリブレーションデータを生成してもよい。これにより視線の注視点と音声が発せられた時間との時間差をより正確に補正することができる。

【0068】

(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。上述した実施の形態1では、設定部18が補正用音声データに基づいて、音声データの重要音声期間を設定していたが、実施の形態2では、操作部13から入力された操作信号に応じて予め指定されたキーワードが発声された期間を音声データの重要音声期間として設定する。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に係る情報処理装置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。

10

【0069】

〔情報処理装置の構成〕

図6A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6Aに示す情報処理装置1aは、上述した実施の形態1に係る情報処理装置1の記録部16に換えて、記録部16aを備える。さらに、情報処理装置1aは、変換部20と、抽出部21と、重要単語記憶部168と、を備える。

【0070】

記録部16aは、上述した実施の形態1に係る記録部16の構成に加えて、操作部13から過去に入力された複数のキーワードを記録するキーワード履歴記録部167を備える。ここで、過去に入力された複数のキーワードとは、例えば、操作部13を経由して過去に入力された単語の頻度を解析することで設定されたキーワードを抽出したものである。

20

【0071】

変換部20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21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。

【0072】

抽出部21は、操作部13から入力された指示信号に対応するキーワード、キーワード履歴記録部167が記録するキーワードおよび重要単語記憶部168が記録するキーワードを、変換部20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18へ出力する。

30

【0073】

重要単語記憶部168は、予め設定された重要単語である複数のキーワードを記憶する。ここで、予め設定された重要単語とは、情報処理装置1aが出荷させる前または図示しないネットワークを経由してメーカー等によって設定された単語(キーワード)である。

【0074】

〔情報処理装置の処理〕

次に、情報処理装置1aが実行する処理について説明する。図6Bは、情報処理装置1aが実行する処理の概要を示すフローチャートである。図6Bにおいて、ステップS201およびステップS202は、上述した図2のステップS101およびステップS103それぞれに対応する。

40

【0075】

ステップS203において、変換部20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S203の後、情報処理装置1aは、後述するステップS204へ移行する。ステップS204およびステップS205は、上述した図2のステップS103およびステップS104それぞれに対応する。ステップS205の後、情報処理装置1aは、ステップS206へ移行する。

【0076】

50

続いて、抽出部 2 1 は、変換部 2 0 が変換した文字情報からキーワード履歴記録部 1 6 7 が記録するキーワードまたは重要単語記憶部 1 6 8 が記録するキーワードが出現する発声時間を抽出する（ステップ S 2 0 6）。

【0077】

その後、設定部 1 8 は、抽出部 2 1 が抽出した発声時間に基づいて、音声データ記録部 1 6 2 が記録する音声データに対して補正用画像における複数の観察箇所の各々で発声される音声が発声された時間を重要音声期間として設定する（ステップ S 2 0 7）。

【0078】

ステップ S 2 0 8 およびステップ S 2 0 9 は、上述した図 2 のステップ S 1 0 6 およびステップ S 1 0 7 それぞれに対応する。ステップ S 2 0 9 の後、情報処理装置 1 a は、本処理を終了する。

10

【0079】

以上説明した実施の形態 2 によれば、キャリブレーション生成部 1 9 が解析部 1 7 によって解析された注視期間の開始時刻と設定部 1 8 によって設定された重要音声期間の開始時刻との時間差を利用者のキャリブレーションデータとして生成するので、視線の注視点と音声が発せられた時間との時間差を補正することができる。なお、実施形態 1 と同様に複数回の測定においてばらつきが生じる場合はそれを考慮したキャリブレーションデータとする。

【0080】

（実施の形態 2 の変形例）

20

次に、本開示の実施の形態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 8 が補正用音声データに基づいて、音声データの重要音声期間を設定していたが、実施の形態 2 の変形例では、音声データを文字情報に変換せずに音声が発せられた期間を重要音声期間として設定する。以下においては、実施の形態の変形例に係る情報処理装置の構成を説明後、実施の形態 2 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 に係る情報処理装置 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。

【0081】

〔情報処理装置の構成〕

図 7 A は、実施の形態 2 の変形例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 7 A に示す情報処理装置 1 a b は、上述した実施の形態 1 に係る情報処理装置 1 から、補正用音声データ記憶部 1 6 5 を削除した構成となっている。また、設定部 1 8 a b は実施の形態 1 に係る情報処理装置 1 の設定部 1 8 と異なり、音声データ記録部 1 6 2 から出力された音声データのうち、音声の内容に関わらず発話がある重要音声期間であって、例えば音量を表すレベルが所定の期間において一定値を超えている、と認識される期間を重要音声期間として設定する。

30

【0082】

〔情報処理装置の処理〕

次に、情報処理装置 1 a b が実行する処理について説明する。図 7 B は、情報処理装置 a b が実行する処理の概要を示すフローチャートである。図 7 B において、ステップ S 2 1 1、ステップ S 2 1 2、ステップ S 2 1 3、及びステップ S 2 1 4 は、上述した図 2 のステップ S 1 0 1、ステップ S 1 0 2、ステップ S 1 0 3 およびステップ S 1 0 4 にそれぞれに対応する。

40

【0083】

ステップ S 2 1 5 において、設定部 1 8 a b は、音声データ記録部 1 6 2 が記録する音声データに対して内容によらず補正用画像における複数の観察箇所の各々で発声される音声が発声された時間を重要音声期間として設定する。ステップ S 2 1 5 の後、情報処理装置 1 a b は、後述するステップ S 2 1 6 へ移行する。

【0084】

ステップ S 2 1 6、及びステップ S 2 1 7 は、上述した図 2 のステップ S 1 0 6 および

50

ステップ S 1 0 7 それぞれに対応する。ステップ S 2 1 7 の後、情報処理装置 1 a a は、本処理を終了する。

【 0 0 8 5 】

ここで、重要音声期間としては、発せられた言葉の内容によらず、注視期間の近傍時間（注視期間及びその前後の所定の期間）で声が発せられた期間で定義するようにしている。具体的には、設定部 1 8 a b は、注視期間の近傍時間内における音声期間が複数存在する場合、例えば最も音量（レベル）の大きい音声が発せられた期間を重要音声期間として、および注視期間の中央との時間差が最小となる音声が発せられた期間を重要音声期間として、いずれかを 1 つ以上を設定する。

【 0 0 8 6 】

また、キャリブレーションデータのもととなる注視期間と重要な音声期間の時間差としては、上述した図 5 A および図 5 B にある注視期間の開始時刻と重要音声期間の開始時刻、または注視期間の終了時刻と重要音声期間の終了時刻の少なくとも一方とだけに限定されるわけでない。具体的には、キャリブレーションデータのもととなる注視期間と重要な音声期間の時間差としては、注視期間の中間に相当する時刻と重要音声期間の中間に相当する時刻との差、または、注視期間の注目度および重要音声期間の音量（レベル）の大きさと時間を積分することで算出される注視期間および重要音声期間の面積の重心に相当する時刻の差である。

【 0 0 8 7 】

また、キャリブレーション生成部 1 9 は、複数の注視期間と重要音声期間から算出される時間差を統合してキャリブレーションデータを生成するようにしてもよい。キャリブレーション生成部 1 9 は、例えば複数の時間差の平均から一意のキャリブレーションデータを生成する、複数の時間差の統計情報をふまえて、平均から標準偏差の所定の倍数のばらつきを持たせた「時間範囲」としてのキャリブレーションデータを生成してもよい。さらに、キャリブレーション生成部 1 9 は、対応する注視期間と重要音声期間の面積の大きさに応じて重みを乗じて統計情報を算出してキャリブレーションデータを生成ようにしてもよい。

【 0 0 8 8 】

以上説明した実施の形態 2 の変形例によれば、キャリブレーション生成部 1 9 が解析部 1 7 によって解析された注視期間と設定部 1 8 a b によって設定された重要音声期間との時間差に基づき利用者のキャリブレーションデータとして生成するので、視線の注視点と音声が発せられた時間との時間差を補正することができる。

【 0 0 8 9 】

（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 1 , 2 では、利用者に対するキャリブレーションデータを生成していたが、実施の形態 3 では、上述した実施の形態 1 , 2 で生成されたキャリブレーションデータを用いて利用者の視線と発声との時間差を補正する。なお、上述した実施の形態 1 に係る情報処理装置 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。

【 0 0 9 0 】

〔情報処理システムの構成〕

図 8 は、実施の形態 3 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 に示す情報処理システム 1 c 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 1 0 0 と、情報処理装置 1 0 0 から出力された各種データを表示する表示部 1 2 0 と、を備える。なお、情報処理装置 1 0 0 と表示部 1 2 0 は、無線または有線によって双方向に接続されている。

【 0 0 9 1 】

〔情報処理装置の構成〕

まず、情報処理装置 1 0 0 の構成について説明する。

図 8 に示す情報処理装置 1 0 0 は、例えばサーバやパーソナルコンピュータ等にインス

10

20

30

40

50

ツ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100は、解析部111と、設定部112と、生成部113と、記録部114と、表示制御部115と、を備える。

【0092】

解析部111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（新たな視線データ）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、上述した実施の形態1の視線検出部10（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10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

【0093】

また、図示していないが、図示しない視線検出部10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10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

【0094】

利用形態が内視鏡システムや光学顕微鏡の場合、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、画像の絶対座標にたいする観察視野の相対的な位置関係は変わらない。また、利用形態が内視鏡システムや光学顕微鏡においては、動画として記録している場合、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。

【0095】

一方で、利用形態がWSI（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、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。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する。

30

【0096】

解析部111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）の注視度が所定の値以上となる新たな注視期間を解析する。なお、図示しない視線検出部10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111は、例えばCPU、FPGAおよびGPU等を用いて構成される。

40

【0097】

設定部112は、外部から入力される利用者の音声データ（新たな音声データ）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111が解析した注視度およびキャリブレーションデータ記憶部116が記録するキャリブレーションデータに応じた重要度を音声データに割り当てて記録部114へ記録する。具体的には、設定部112は、音声データのフレーム毎に、キャリブレーションデータに基づいて、利用者の音声の発声時間と視線が注視する注視時間との時間差を補正し、この補正した音声

50

データに対して、解析部 1 1 1 が解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 1 1 4 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 1 1 2 は、CPU、FPGA および GPU 等を用いて構成される。

【0098】

生成部 1 1 3 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 1 1 1 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 1 1 4 および表示制御部 1 1 5 へ出力する。具体的には、生成部 1 1 3 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 1 1 1 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 1 1 3 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 1 1 1 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 1 1 3 は、CPU、FPGA および GPU 等を用いて構成される。上述の WSI で使用する場合は、生成部 1 1 3 は、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。また、上述の様に、生成部 1 1 3 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。

【0099】

記録部 1 1 4 は、設定部 1 1 2 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たれた重要度と、解析部 1 1 1 が解析した注視度と、を対応付けて記録する。また、記録部 1 1 4 は、生成部 1 1 3 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 1 1 4 は、情報処理装置 1 0 0 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 1 1 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る。

【0100】

表示制御部 1 1 5 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 1 1 3 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 1 2 0 に出力することによって表示させる。表示制御部 1 1 5 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 1 1 1、設定部 1 1 2、視線マッピングおよび表示制御部 1 1 5 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。

【0101】

〔表示部の構成〕

次に、表示部 1 2 0 の構成について説明する。

表示部 1 2 0 は、表示制御部 1 1 5 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 1 2 0 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。

【0102】

〔情報処理装置の処理〕

次に、情報処理装置 1 0 0 の処理について説明する。図 9 は、情報処理装置 1 0 0 が実行する処理の概要を示すフローチャートである。

【0103】

図 9 に示すように、まず、情報処理装置 1 0 0 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 S 3 0 1）。

【0104】

続いて、解析部 1 1 1 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 S 3 0 2）。具体的には、解析部 1 1 1 は、上述した実施の形態 1 の図 3 で説明した方法によって利用者の視線の注視度を解析する。

10

20

30

40

50

【 0 1 0 5 】

その後、設定部 1 1 2 は、視線データと同期化された音声データに対して、キャリブレーションデータ記憶部 1 1 6 が記録するキャリブレーションデータおよび解析部 1 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 1 4 に記録する（ステップ S 3 0 3）。具体的には、設定部 1 1 2 は、上述した実施の形態 1 の図 4 および図 5 で説明した方法によって、所定の時間間隔毎に解析部 1 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 1 4 に記録する。

【 0 1 0 6 】

続いて、生成部 1 1 3 は、画像データに対応する画像上に解析部 1 1 1 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（ステップ S 3 0 4）。

10

【 0 1 0 7 】

続いて、表示制御部 1 1 5 は、画像データに対応する画像上に、生成部 1 1 3 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 1 2 0 に出力する（ステップ S 3 0 5）。ステップ S 3 0 5 の後、情報処理装置 1 0 0 は、本処理を終了する。

【 0 1 0 8 】

図 1 0 は、表示部 1 2 0 が表示する画像の一例を模式的に示す図である。図 1 0 に示すように、表示制御部 1 1 5 は、画像データに対応する画像上に、生成部 1 1 3 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 P 1 を表示部 1 2 0 に表示させる。図 1 0 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 M 1 ~ M 5 の視線マッピング画像 P 1 を表示部 1 2 0 に表示させる。

20

【 0 1 0 9 】

図 1 1 は、表示部 1 2 0 が表示する画像の別の一例を模式的に示す図である。図 1 1 に示すように、表示制御部 1 1 5 は、画像データに対応する画像上に、生成部 1 1 3 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 P 2 を表示部 1 2 0 に表示させる。図 1 1 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 M 1 1 ~ M 1 5 が重畳された視線マッピング画像 P 2 を表示部 1 2 0 に表示させる。さらに、表示制御部 1 1 5 は、利用者の視線の軌跡 K 1 および注視度の順番を数字で表示部 1 2 0 に表示させる。なお、図 1 1 においては、表示制御部 1 1 5 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 1 2 0 に表示させてもよい。

30

【 0 1 1 0 】

以上説明した実施の形態 3 によれば、設定部 1 1 2 が外部から入力された音声データ（新たな音声データ）に対して、キャリブレーションデータ記憶部 1 1 6 が記録する利用者のキャリブレーションデータおよび解析部 1 1 1 が新たに解析した新たな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 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 1 1 1 】

さらにまた、実施の形態 3 では、生成部 1 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 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。

40

【 0 1 1 2 】

なお、実施の形態 3 では、設定部 1 1 2 が外部から入力される音声データ（新たな音声データ）に対して、この音声データの中で発声された発声期間を設定し、かつ、この音声データに対して、発声期間と解析部 1 1 1 が新たに解析した新たな注視度とキャリブレーションデータ記憶部 1 1 6 が記憶するキャリブレーションデータとを用いて解析部 1 1 1 が解析した新たな注視度に応じた重要度を割り当てて記録部 1 1 4 へ記録してもよい。これにより、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 1 1 3 】

また、実施の形態 3 では、設定部 1 1 2 が外部から入力される音声データ（新たな音声

50

データ)の中で所定のキーワードが発声された期間を新たな重要発声期間として設定し、かつ、外部から入力された視線データ(新たな視線データ)に対して、新たな重要発声期間とキャリブレーションデータ記憶部116が記憶するキャリブレーションデータとを用いて新たな重要発声期間に応じた重要度を割り当てて記録部114へ記録してもよい。視線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。

【0114】

また、実施の形態3では、設定部112が外部から入力される音声データ(新たな音声データ)の中で所定のキーワードが発声された期間を新たな重要発声期間として設定し、かつ、外部から入力される視線データ(新たな視線データ)に対して、新たな注視期間と新たな重要発声期間とキャリブレーションデータ記憶部116が記憶するキャリブレーションデータとを用いて重要発声期間に応じた重要度を割り当てて記録部114へ記録してもよい。これにより、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。

10

【0115】

また、実施の形態3では、記録部114が設定部112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。

【0116】

(実施の形態3の変形例)

20

次に、本開示の実施の形態3の変形例について説明する。上述した実施の形態3では、設定部112が解析部111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114へ記録していたが、実施の形態3の変形例では、設定部112が生成部113によって生成された視線マッピングデータに基づいて、重要度を音声データに割り当てて記録部114へ記録する。以下においては、実施の形態3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3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3に係る情報処理システム1c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。

【0117】

〔情報処理システムの構成〕

30

図12は、実施の形態3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12に示す情報処理システム1aは、上述した実施の形態1に係る情報処理装置100に代えて、情報処理装置100aを備える。情報処理装置100aは、上述した実施の形態1に係る設定部112に代えて、設定部112aを備える。

【0118】

設定部112aは、生成部113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114へ記録する。

【0119】

〔情報処理装置の処理〕

次に、情報処理装置100aが実行する処理について説明する。図13は、情報処理装置100aが実行する処理の概要を示すフローチャートである。図13において、ステップS401およびステップS402は、上述した図9のステップS301およびステップS302それぞれに対応する。また、ステップS403は、上述した図9のステップS304に対応する。

40

【0120】

ステップS404において、設定部112は、キャリブレーションデータ記憶部116が記録するキャリブレーションデータと生成部113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、キャリブレーションデータおよび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114に記録する。ステップS404の後、情報処理装置100aは、ステップS405へ移行する。ステ

50

ステップ S 4 0 5 は、上述した図 9 のステップ S 3 0 5 に対応する。

【 0 1 2 1 】

以上説明した実施の形態 3 の変形例によれば、設定部 1 1 2 が生成部 1 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 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 1 2 2 】

また、実施の形態 3 および実施の形態 3 の変形例では、注視度とキャリブレーションデータを用いて、音声・発話の重要度を付与する実施形態について説明した。これらの変形例として、所定の単語（キーワード）を発した期間である重要音声期間に対応する視線データの期間（あるいは画像にマッピングした後は画像の領域・位置）抽出するために、重要音声期間とあらかじめ取得しキャリブレーションデータを用いてもよい。

【 0 1 2 3 】

（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 3 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 4 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 4 に係る情報処理装置の構成を説明後、実施の形態 4 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 3 に係る情報処理システム 1 c 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 2 4 】

〔 情報処理装置の構成 〕

図 1 4 は、実施の形態 4 に係る情報処理装置の構成を示す概略図である。図 1 5 は、実施の形態 4 に係る情報処理装置の構成を示す概略図である。図 1 6 は、実施の形態 4 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 1 2 5 】

図 1 4 ~ 図 1 6 に示す情報処理装置 1 b は、解析部 1 1 1 と、表示部 1 2 0 と、視線検出部 1 3 0 と、音声入力部 1 3 1 と、制御部 1 3 2 と、時間計測部 1 3 3 と、記録部 3 4 と、変換部 1 3 5 と、抽出部 1 3 6 と、操作部 1 3 7 と、設定部 1 3 8 と、生成部 1 3 9 と、プログラム記憶部と、キャリブレーションデータ記憶部 3 4 5 と、を備える。

【 0 1 2 6 】

視線検出部 1 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 1 3 0 は、利用者 U 1 が表示部 1 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 1 4 および図 1 5 を参照）。視線検出部 1 3 0 は、制御部 1 3 2 の制御のもと、表示部 1 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 1 3 2 へ出力する。具体的には、視線検出部 1 3 0 は、制御部 1 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 1 3 0 は、制御部 1 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 1 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 1 2 7 】

音声入力部 1 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

132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131は、制御部132の制御のもと、利用者U1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132へ出力する。なお、音声入力部131は、音声の入力以外にも、音声を出力することができるスピーカ等を設け、音声出力機能を設けてもよい。

【0128】

制御部132は、CPU、FPGAおよびGPU等を用いて構成され、視線検出部130、音声入力部131および表示部120を制御する。制御部132は、視線検出制御部321と、音声入力制御部322と、表示制御部323と、を有する。

【0129】

視線検出制御部321は、視線検出部130を制御する。具体的には、視線検出制御部321は、視線検出部130を所定のタイミング毎に近赤外線を利用者U1へ照射させるとともに、利用者U1の瞳を視線検出部130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321は、視線検出部130から入力された視線データに対して、各種の画像処理を行って記録部34へ出力する。

【0130】

音声入力制御部322は、音声入力部131を制御し、音声入力部131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34へ出力する。

【0131】

表示制御部323は、表示部120の表示態様を制御する。表示制御部323は、記録部34に記録された画像データに対応する画像または生成部113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120に表示させる。

【0132】

時間計測部133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130によって生成された視線データおよび音声入力部131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。

【0133】

記録部34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1bに関する各種の情報を記録する。記録部34は、視線データ記録部341と、音声データ記録部342と、画像データ記録部343と、プログラム記憶部と、キャリブレーションデータ記憶部345と、を有する。

【0134】

視線データ記録部341は、視線検出制御部321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111へ出力する。

【0135】

音声データ記録部342は、音声入力制御部322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135へ出力する。

【0136】

画像データ記録部343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1b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。

【0137】

プログラム記憶部は、情報処理装置1bが実行する各種プログラム、各種プログ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。

【0138】

キャリブレーションデータ記憶部345は、利用者毎のキャリブレーションデータを記録する。

【0139】

10

20

30

40

50

変換部 135 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 136 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。

【0140】

抽出部 136 は、後述する操作部 137 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 135 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 138 へ出力する。なお、抽出部 136 は、後述する操作部 137 から指示信号が入力されていない場合、変換部 135 から入力されたままの文字情報を設定部 138 へ出力する。

10

【0141】

操作部 137 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 132 へ出力する。

【0142】

設定部 138 は、キャリブレーションデータ記憶部 345 が記録するキャリブレーションデータと、所定の時間間隔毎に解析部 111 が解析した注視度と、抽出部 136 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 135 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

【0143】

生成部 139 は、表示部 120 が表示する画像データに対応する画像上に解析部 111 が解析した注視度および変換部 135 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。

【0144】

〔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 17 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。

【0145】

図 17 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 120 に表示させる（ステップ S501）。この場合、表示制御部 323 は、操作部 137 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 120 に表示させる。

30

【0146】

続いて、制御部 132 は、視線検出部 130 が生成した視線データおよび音声入力部 131 が生成した音声データの各々と時間計測部 133 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 S502）。

【0147】

その後、変換部 135 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 S503）。なお、このステップは、後述のステップ S506 の後に行っても良い。

40

【0148】

続いて、操作部 137 から表示部 120 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 S504：Yes）、情報処理装置 1b は、後述するステップ S505 へ移行する。これに対して、操作部 137 から表示部 120 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 S504：No）、情報処理装置 1b は、ステップ S502 へ戻る。

【0149】

50

ステップ S 5 0 5 は、上述した図 9 のステップ S 3 0 2 に対応する。ステップ S 5 0 5 の後、情報処理装置 1 b は、後述するステップ S 5 0 6 へ移行する。

【 0 1 5 0 】

ステップ S 5 0 6 において、設定部 1 3 8 は、キャリブレーションデータ記憶部 3 4 5 が記録するキャリブレーションデータと、解析部 1 1 1 が解析した注視度と、抽出部 1 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 1 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 1 3 8 は、利用者 U 1 に対応するキャリブレーションを用いて視線が注視する注視期間と音声が発声された音声発声時間との時間差を補正した後に、抽出部 1 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 1 3 8 は、重要度に、抽出部 1 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乘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 1 5 1 】

続いて、生成部 1 3 9 は、表示部 1 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 1 が解析した注視度および変換部 1 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 5 0 7 ）。

【 0 1 5 2 】

続いて、表示制御部 3 2 3 は、生成部 1 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 1 2 0 に表示させる（ステップ S 5 0 8 ）。

20

【 0 1 5 3 】

図 1 8 は、表示部 1 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 8 に示すように、表示制御部 3 2 3 は、生成部 1 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 1 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 1 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 8 では、表示制御部 3 2 3 が文字情報を表示部 1 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。

30

【 0 1 5 4 】

図 1 9 A は、表示部 1 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 9 A に示すように、表示制御部 3 2 3 は、生成部 1 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 1 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 1 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 1 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 1 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 1 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 1 5 5 】

また、注視度のもととなる視線データと、文字情報のもととなる音声データは、キャリブレーションデータにより補正されることで、時間差が補正前に比べて小さくなっている。そのため、ユーザが画像のある部分を注視しながら発声している場合において、視線デ

50

ータの注視度が割り当てられている期間と音声データの重要度が割り当てられている期間は、時間的に重なる割合が補正前に比べて大きくなる。これに基づき、生成部139は、視線マッピングデータの中に視線データと音声データを時刻によって紐づけて関連付けるフォーマットを含むデータを生成する。

【0156】

図19Bは、視線データと音声データを時刻によって関連付けたデータの一例であって、データのデータフォーマットの内容を表す。図19Bの表T1では、視線データの開始時刻と終了時刻、音声データの開始時刻と終了時刻がIndex番号により関連付けられている。

【0157】

これにより、生成部139は、図19Bの表T1に示すように、例えばIndex番号により視線データを取り出して、それに該当する音声データを取り出す、ということが出来る。これに対して、生成部139は、音声が発せられた期間（重要音声期間）からそのときに画像のどこを注視していたか、という（視線データの）情報をIndex番号により取り出すこともできる。逆に、生成部139は、観察画像における視線の注視位置（領域）から、そこを見ていたに発していた音声（文字情報）をIndex番号により取り出して、注視位置に対応する文字情報を視覚化して提示することも可能である。上記以外の方法として、生成部139は、視線の特徴量（移動距離、停留時間）の経時変化と音声データを関連付けても良い。

【0158】

図17に戻り、ステップS509以降の説明を続ける。

ステップS509において、操作部137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S509：Yes）、制御部132は、操作に応じた動作処理を実行する（ステップS510）。具体的には、表示制御部323は、操作部137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120に強調表示させる（例えば図18を参照）。また、音声入力制御部322は、注視度の高い領域に関連付けられた音声データを音声入力部131に再生させる。ステップS510の後、情報処理装置1bは、後述するステップS511へ移行する。

【0159】

ステップS509において、操作部137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S509：No）、情報処理装置1bは、後述するステップS511へ移行する。

【0160】

ステップS511において、操作部137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S511：Yes）、情報処理装置1bは、本処理を終了する。これに対して、操作部137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S511：No）、情報処理装置1bは、上述したステップS508へ戻る。

【0161】

以上説明した実施の形態4によれば、生成部113が表示部120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111が解析した注視度および変換部135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U1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。

【0162】

また、実施の形態4によれば、表示制御部323は、生成部113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120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。

【0163】

（実施の形態5）

10

20

30

40

50

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 3 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 5 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 5 に係る顕微鏡システムの構成を説明後、実施の形態 5 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 4 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。

【0164】

〔顕微鏡システムの構成〕

図 20 は、実施の形態 5 に係る顕微鏡システムの構成を示す概略図である。図 21 は、実施の形態 5 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。

10

【0165】

図 20 および図 21 に示すように、顕微鏡システム 110 は、情報処理装置 110 c と、表示部 120 と、音声入力部 131 と、操作部 137 と、顕微鏡 200 と、撮像部 210 と、視線検出部 220 と、を備える。

【0166】

〔顕微鏡の構成〕

まず、顕微鏡 200 の構成について説明する。

顕微鏡 200 は、本体部 201 と、回転部 202 と、昇降部 203 と、レボルバ 204 と、対物レンズ 205 と、倍率検出部 206 と、鏡筒部 207 と、接続部 208 と、接眼部 209 と、を備える。

20

【0167】

本体部 201 は、標本 S P が載置される。本体部 201 は、略 U 字状をなし、回転部 202 を用いて昇降部 203 が接続される。

【0168】

回転部 202 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 203 を垂直方向へ移動させる。

【0169】

昇降部 203 は、本体部 201 に対して垂直方向へ移動可能に設けられている。昇降部 203 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 207 が接続される。

【0170】

レボルバ 204 は、互いに倍率が異なる複数の対物レンズ 205 が接続され、光軸 L 1 に対して回転可能に昇降部 203 に接続される。レボルバ 204 は、利用者 U 2 の操作に応じて、所望の対物レンズ 205 を光軸 L 1 上に配置する。なお、複数の対物レンズ 205 には、倍率を示す情報、例えば IC チップやラベルが添付されている。なお、IC チップやラベル以外にも、倍率を示す形状を対物レンズ 205 に設けてもよい。

30

【0171】

倍率検出部 206 は、光軸 L 1 上に配置された対物レンズ 205 の倍率を検出し、この検出した検出結果を情報処理装置 110 c へ出力する。倍率検出部 206 は、例えば対物切り替えのレボルバ 204 の位置を検出する手段を用いて構成される。

【0172】

鏡筒部 207 は、対物レンズ 205 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 208 に透過するとともに、接眼部 209 へ反射する。鏡筒部 207 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。

40

【0173】

接続部 208 は、一端が鏡筒部 207 と接続され、他端が撮像部 210 と接続される。接続部 208 は、鏡筒部 207 を透過した標本 S P の被写体像を撮像部 210 へ導光する。接続部 208 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。

【0174】

接眼部 209 は、鏡筒部 207 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 209 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。

50

【 0 1 7 5 】

〔 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、接眼部 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 1 0 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 7 6 】

〔 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 1 0 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 1 0 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 1 0 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 1 0 c へ出力する。

【 0 1 7 7 】

〔 情報処理装置の構成 〕

次に、情報処理装置 1 1 0 c の構成について説明する。

情報処理装置 1 1 0 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 1 3 2、記録部 3 4 および設定部 1 3 8 に換えて、制御部 1 3 2 c、記録部 3 4 c、設定部 1 3 8 c と、を備える。

【 0 1 7 8 】

制御部 1 3 2 c は、C P U、F P G A および G P U 等を用いて構成され、表示部 1 2 0、音声入力部 1 3 1、撮像部 2 1 0 および視線検出部 2 2 0 を制御する。制御部 1 3 2 c は、上述した実施の形態 4 の制御部 1 3 2 の視線検出制御部 3 2 1、音声入力制御部 3 2 2、表示制御部 3 2 3 に加えて、撮影制御部 3 2 4 および倍率算出部 3 2 5 をさらに備える。

【 0 1 7 9 】

撮影制御部 1 3 2 4 は、撮像部 2 1 0 の動作を制御する。撮影制御部 1 3 2 4 は、撮像部 2 1 0 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 1 3 2 4 は、撮像部 2 1 0 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 3 4 c へ出力する。

【 0 1 8 0 】

倍率算出部 3 2 5 は、倍率検出部 2 0 6 から入力された検出結果に基づいて、現在の顕微鏡 2 0 0 の観察倍率を算出し、この算出結果を設定部 1 3 8 c へ出力する。例えば、倍率算出部 3 2 5 は、倍率検出部 2 0 6 から入力された対物レンズ 2 0 5 の倍率と接眼部 2 0 9 の倍率とに基づいて、現在の顕微鏡 2 0 0 の観察倍率を算出する。

【 0 1 8 1 】

記録部 3 4 c 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 c は、上述した実施の形態 2 に係る画像データ記録部 3 4 3 に換えて、画像データ記録部 3 4 6 を備える。画像データ記録部 3 4 6 は、撮影制御部 1 3 2 4 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 1 1 3 へ出力する。

【 0 1 8 2 】

設定部 1 3 8 c は、キャリブレーションデータ記憶部 3 4 5 が記録するキャリブレーション

ョンデータと、解析部 1 1 1 が解析した注視度と、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 1 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 へ記録する。具体的には、設定部 1 3 8 c は、解析部 1 1 1 が解析した注視度に、倍率算出部 3 2 5 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 3 4 c へ記録する。すなわち、設定部 1 3 8 c は、表示倍率が高いほど重要度が高くなるような処理を行う。設定部 1 3 8 c は、CPU、FPGA および GPU 等を用いて構成される。

【0183】

〔顕微鏡システムの処理〕

次に、顕微鏡システム 1 1 0 が実行する処理について説明する。図 2 2 は、顕微鏡システム 1 1 0 が実行する処理の概要を示すフローチャートである。

【0184】

図 2 2 に示すように、まず、制御部 1 3 2 c は、視線検出部 1 3 0 が生成した視線データ、音声入力部 1 3 1 が生成した音声データ、および倍率算出部 3 2 5 が算出した観察倍率の各々を時間計測部 1 3 3 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 3 4 1 および音声データ記録部 3 4 2 に記録する（ステップ S 6 0 1）。ステップ S 6 0 1 の後、顕微鏡システム 1 1 0 は、後述するステップ S 6 0 2 へ移行する。

【0185】

ステップ S 6 0 2 ～ステップ S 6 0 4 は、上述した図 1 7 のステップ S 5 0 3 ～ステップ S 5 0 5 それぞれに対応する。ステップ S 6 0 4 の後、顕微鏡システム 1 1 0 は、ステップ S 6 0 5 へ移行する。

【0186】

ステップ S 6 0 5 において、設定部 1 3 8 c は、キャリブレーションデータ記憶部 3 4 5 が記録するキャリブレーションデータと、所定の時間間隔毎に解析部 1 1 1 が解析した注視度と、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 1 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 へ記録する。ステップ S 6 0 5 の後、顕微鏡システム 1 1 0 は、ステップ S 6 0 6 へ移行する。

【0187】

ステップ S 6 0 6 ～ステップ S 6 1 0 は、上述した図 1 7 のステップ S 5 0 7 ～ステップ S 5 1 1 それぞれに対応する。

【0188】

以上説明した実施の形態 5 によれば、設定部 1 3 8 c がキャリブレーションデータ記憶部 3 4 5 によって記録されたキャリブレーションデータと、解析部 1 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 1 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。

【0189】

なお、実施の形態 5 では、倍率算出部 3 2 5 が算出した観察倍率を記録部 1 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。

【0190】

（実施の形態 6）

次に、本開示の実施の形態 6 について説明する。実施の形態 6 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 6 に係る内視鏡システムの構成を説明後、実施の形態 6 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 3 に係る情報処理装置 1 b と同一の構成に

10

20

30

40

50

は同一の符号を付して詳細な説明は適宜省略する。

【0191】

〔内視鏡システムの構成〕

図23は、実施の形態6に係る内視鏡システムの構成を示す概略図である。図24は、実施の形態6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。

【0192】

図23および図24に示す内視鏡システム300は、表示部120と、内視鏡400と、ウェアラブルデバイス500と、入力部600と、情報処理装置1dと、を備える。

【0193】

〔内視鏡の構成〕

まず、内視鏡400の構成について説明する。

内視鏡400は、医者や術者等の利用者U3が被検体U4に挿入することによって、被検体U4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1dへ出力する。内視鏡400は、撮像部401と、操作部402と、を備える。

【0194】

撮像部401は、内視鏡400の挿入部の先端部に設けられる。撮像部401は、情報処理装置1dの制御のもと、被検体U4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1dへ出力する。撮像部401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CMOSやCCD等のイメージセンサ等を用いて構成される。

【0195】

操作部402は、利用者U3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1dへ出力する。

【0196】

〔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500の構成について説明する。

ウェアラブルデバイス500は、利用者U3に装着され、利用者U3の視線を検出するとともに、利用者U3の音声の入力を受け付ける。ウェアラブルデバイス500は、視線検出部510と、音声入力部520と、を有する。

【0197】

視線検出部510は、ウェアラブルデバイス500に設けられ、利用者U3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1dへ出力する。視線検出部510は、上述した実施の形態3に係る視線検出部220と同様の構成を有するため、詳細な構成は省略する。

【0198】

音声入力部520は、ウェアラブルデバイス500に設けられ、利用者U3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1dへ出力する。音声入力部520は、マイク等を用いて構成される。

【0199】

〔入力部の構成〕

入力部600の構成について説明する。

入力部600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600は、利用者U3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1dへ出力する。

【0200】

〔情報処理装置の構成〕

次に、情報処理装置1dの構成について説明する。

情報処理装置1dは、上述した実施の形態5に係る情報処理装置110cの制御部132c、記録部34c、設定部138c、生成部139に換えて、制御部132d、記録部34d、設定部138dおよび生成部139dを備える。さらに、情報処理装置1dは、

10

20

30

40

50

画像処理部 1 4 0 をさらに備える。

【 0 2 0 1 】

制御部 1 3 2 d は、C P U、F P G A および G P U 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 1 2 0 を制御する。制御部 1 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 2 0 2 】

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。

10

【 0 2 0 3 】

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 5 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 7 をさらに備える。

【 0 2 0 4 】

操作履歴記録部 3 4 7 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。

20

【 0 2 0 5 】

設定部 1 3 8 d は、キャリブレーションデータ記憶部 3 4 5 が記録するキャリブレーションデータと、解析部 1 1 1 が解析した注視度と、操作履歴記録部 3 4 7 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 1 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 1 3 8 d は、キャリブレーションデータ記憶部 3 4 5 が記録するキャリブレーションデータと、解析部 1 1 1 が解析した注視度と、操作履歴記録部 3 4 7 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 1 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 1 3 8 d は、C P U、F P G A および G P U 等を用いて構成される。

30

【 0 2 0 6 】

生成部 1 3 9 d は、画像処理部 1 4 0 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 1 1 1 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 3 4 d および表示制御部 3 2 3 へ出力する。

【 0 2 0 7 】

画像処理部 1 4 0 は、画像データ記録部 3 4 6 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 1 3 9 d へ出力する。

40

【 0 2 0 8 】

〔内視鏡システムの処理〕

次に、内視鏡システム 3 0 0 が実行する処理について説明する。図 2 5 は、内視鏡システム 3 0 0 が実行する処理の概要を示すフローチャートである。

【 0 2 0 9 】

図 2 5 に示すように、まず、制御部 1 3 2 d は、視線検出部 1 3 0 が生成した視線データ、音声入力部 1 3 1 が生成した音声データ、および操作履歴検出部 3 2 6 が検出した操作履歴の各々を時間計測部 1 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 および操作履歴記録部 3 4 7 に記録する（ステップ S 7 0 1）。ステップ S 7 0 1 の後、内視鏡システム 3 0 0 は、後述するステップ S 7 0 2 へ

50

移行する。

【0210】

ステップS702～ステップS704は、上述した図17のステップS503～ステップS505それぞれに対応する。ステップS704の後、内視鏡システム300は、ステップS705へ移行する。

【0211】

ステップS705において、設定部138dは、キャリブレーションデータ記憶部345が記録するキャリブレーションデータと、解析部111が解析した注視度と、操作履歴記録部347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135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。

10

【0212】

続いて、画像処理部140は、画像データ記録部346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139dへ出力する(ステップS706)。図26は、画像データ記録部346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27は、画像処理部140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26および図27に示すように、画像処理部140は、時間的に連続する複数の画像データ $P11 \sim P_N$ ($N = \text{整数}$)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 $P100$ を生成する。

20

【0213】

その後、生成部139dは、画像処理部140が生成した統合画像データに対応する統合画像 $P100$ 上に、解析部111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S707)。この場合、生成部139dは、画像処理部140が生成した統合画像データに対応する統合画像 $P100$ 上に、解析部111が解析した注視度、視線 $K2$ 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S707の後、内視鏡システム300は、後述するステップS708へ移行する。

30

【0214】

ステップS708～ステップS711は、上述した図17のステップS508～ステップS511それぞれに対応する。

【0215】

以上説明した実施の形態6によれば、設定部138dが解析部111によって解析された注視度と操作履歴記録部347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135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。

40

【0216】

また、実施の形態6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。

【0217】

また、実施の形態6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。

【0218】

また、実施の形態6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

50

ことができる。

【0219】

(実施の形態7)

次に、本開示の実施の形態7について説明する。上述した実施の形態3～6は、利用者が一人の場合を想定していたが、実施の形態7では、2人以上の利用者を想定する。さらに、実施の形態7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7に係る閲覧システムの構成を説明後、実施の形態7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3に係る情報処理装置1b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。

10

【0220】

〔情報処理システムの構成〕

図28は、実施の形態7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28に示す情報処理システム700は、表示部120と、第1ウェアラブルデバイス710と、第2ウェアラブルデバイス720と、検出部730と、情報処理装置1eと、を備える。

【0221】

〔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710の構成について説明する。

第1ウェアラブルデバイス710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710は、第1視線検出部711と、第1音声入力部712と、を有する。第1視線検出部711および第1音声入力部712は、上述した実施の形態4に係る視線検出部510および音声入力部520と同様の構成を有するため、詳細な構成は省略する。

20

【0222】

〔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720の構成について説明する。

第2ウェアラブルデバイス720は、上述した第1ウェアラブルデバイス710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720は、第2視線検出部721と、第2音声入力部722と、を有する。第2視線検出部721および第2音声入力部722は、上述した実施の形態4に係る視線検出部510および音声入力部520と同様の構成を有するため、詳細な構成は省略する。

30

【0223】

〔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137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。

40

【0224】

〔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132d

50

、記録部 3 4 d および設定部 1 3 8 d に換えて、制御部 1 3 2 e、記録部 3 4 e および設定部 1 3 8 e を備える。

【0 2 2 5】

制御部 1 3 2 e 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 7 1 0、第 2 ウェアラブルデバイス 7 2 0、検出部 7 3 0 および表示部 1 2 0 を制御する。制御部 1 3 2 e 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 に加えて、識別検出制御部 3 2 7 を備える。

【0 2 2 6】

識別検出制御部 3 2 7 は、検出部 7 3 0 を制御し、検出部 7 3 0 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 3 4 e へ出力する。

10

【0 2 2 7】

記録部 3 4 e 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 e 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、識別情報記録部 3 4 8 をさらに備える。

【0 2 2 8】

識別情報記録部 3 4 8 は、識別検出制御部 3 2 7 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。

【0 2 2 9】

設定部 1 3 8 e は、キャリアブレーションデータ記憶部 3 4 5 が記録する利用者毎のキャリアブレーションデータと、解析部 1 1 1 が解析した各解析結果と、抽出部 1 3 6 が抽出した文字情報と、識別情報記録部 3 4 8 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 1 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 に記録する。さらに、設定部 1 3 8 e は、識別情報記録部 3 4 8 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 1 3 8 e 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。

20

【0 2 3 0】

〔情報処理システムの処理〕

次に、情報処理システム 7 0 0 が実行する処理について説明する。図 2 9 は、情報処理システム 7 0 0 が実行する処理の概要を示すフローチャートである。

30

【0 2 3 1】

図 2 9 に示すように、表示制御部 3 2 3 は、画像データ記録部 3 4 3 が記録する画像データに対応する画像を表示部 1 2 0 に表示させる（ステップ S 8 0 1）。

【0 2 3 2】

続いて、制御部 1 3 2 e は、第 1 ウェアラブルデバイス 7 1 0 および第 2 ウェアラブルデバイス 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 1 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 8 に記録する（ステップ S 8 0 2）。ステップ S 8 0 2 の後、情報処理システム 7 0 0 は、ステップ S 8 0 3 へ移行する。

40

【0 2 3 3】

ステップ S 8 0 3 およびステップ S 8 0 4 は、上述した図 1 7 のステップ S 5 0 3 およびステップ S 5 0 4 それぞれに対応する。ステップ S 8 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 8 0 5 へ移行する。

【0 2 3 4】

ステップ S 8 0 5 において、解析部 1 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。

【0 2 3 5】

続いて、設定部 1 3 8 e は、所定の時間間隔毎に解析部 1 1 1 が解析した各注視度と、

50

識別情報記録部 348 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 135 によって変換された文字情報を割り当てて記録部 34e へ記録する（ステップ S806）。

【0236】

ステップ S807～ステップ S811 は、上述した図 17 のステップ S507～ステップ S511 それぞれに対応する。

【0237】

以上説明した実施の形態 7 によれば、設定部 138e が解析部 111 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 348 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 135 によって変換された文字情報を割り当てて記録部 34e 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。

10

【0238】

なお、実施の形態 7 では、設定部 138e が解析部 111 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 348 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 135 によって変換された文字情報を割り当てて記録部 34e 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 135 によって変換された文字情報を割り当てて記録部 34e へ記録してもよい。

20

【0239】

（その他の実施の形態）

上述した実施の形態 1～7 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。

30

【0240】

また、実施の形態 1～7 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。

【0241】

また、実施の形態 1～7 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。

【0242】

また、実施の形態 1～7 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。

40

【0243】

また、実施の形態 1～7 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。

50

【 0 2 4 4 】

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。

【 0 2 4 5 】

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。

10

【 符号の説明 】

【 0 2 4 6 】

1, 1 a, 1 b, 1 d, 1 e, 1 0 0, 1 0 0 a, 1 1 0 c 情報処理装置

1 a a, 1 c, 7 0 0 情報処理システム

1 0, 1 3 0, 2 2 0, 2 2 2, 5 1 0 視線検出部

1 1, 1 3 1, 5 2 0 音声入力部

1 2, 1 2 0 表示部

1 3, 1 3 7, 4 0 2 操作部

1 4, 1 3 2, 1 3 2 c, 1 3 2 d, 1 3 2 e 制御部

1 5 時間計測部

20

1 6, 1 6 a, 3 4, 3 4 c, 3 4 d, 3 4 e, 1 1 4 記録部

1 7, 1 1 1 解析部

1 8, 1 1 2, 1 1 2 a, 1 3 8, 1 3 8 c, 1 3 8 d, 1 3 8 e 設定部

1 9 キャリブレーション生成部

2 0, 1 3 5 変換部

2 1, 1 3 6 抽出部

1 1 0 顕微鏡システム

1 1 3, 1 3 9, 1 3 9 d 生成部

1 1 5, 1 4 3, 3 2 3 表示制御部

1 1 6, 1 6 6, 3 4 5 キャリブレーションデータ記憶部

30

1 3 3 時間計測部

1 4 0 画像処理部

1 4 1, 3 2 1 視線検出制御部

1 4 2, 3 2 2 音声入力制御部

1 6 1, 3 4 1 視線データ記録部

1 6 2, 3 4 2 音声データ記録部

1 6 3, 3 4 3, 3 4 6 画像データ記録部

1 6 4, 3 4 4 プログラム記憶部

1 6 5 補正用音声データ記憶部

1 6 7 キーワード履歴記録部

40

1 6 8 重要単語記憶部

2 0 0 顕微鏡

2 0 1 本体部

2 0 2 回転部

2 0 3 昇降部

2 0 4 レボルバ

2 0 5 対物レンズ

2 0 6 倍率検出部

2 0 7 鏡筒部

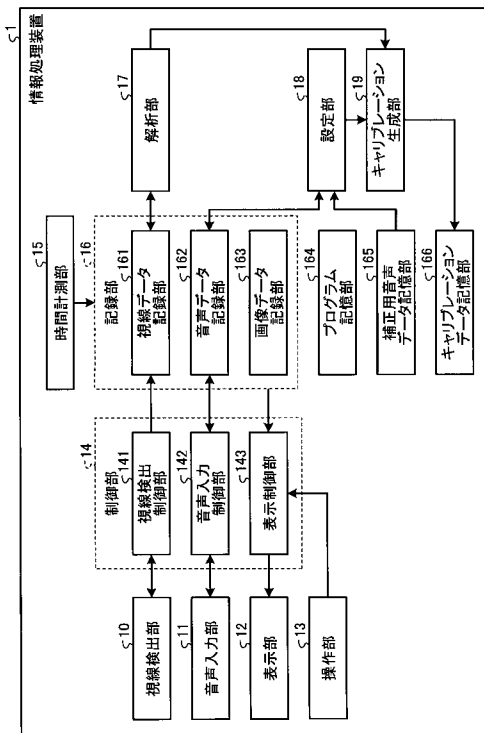
2 0 8 接続部

50

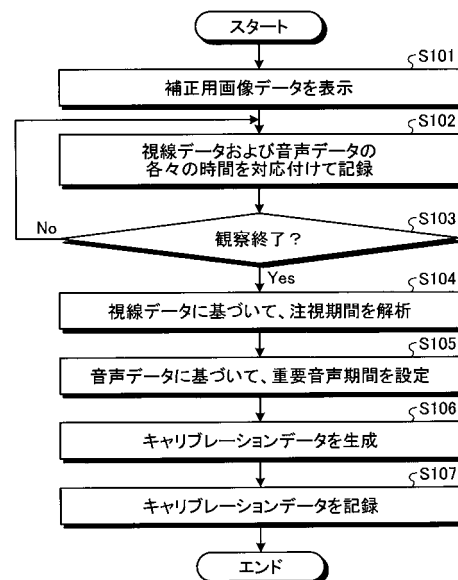
209 接眼部
 210, 401 撮像部
 300 内視鏡システム
 325 倍率算出部
 326 操作履歴検出部
 347 操作履歴記録部
 348 識別情報記録部
 400 内視鏡
 500 ウェアラブルデバイス
 600 入力部
 710 第1ウェアラブルデバイス
 711 第1視線検出部
 712 第1音声入力部
 720 第2ウェアラブルデバイス
 721 第2視線検出部
 722 第2音声入力部
 730 検出部
 324 撮影制御部
 327 識別検出制御部

10

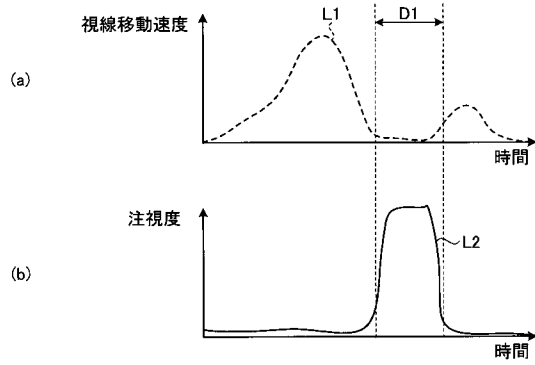
【図1】



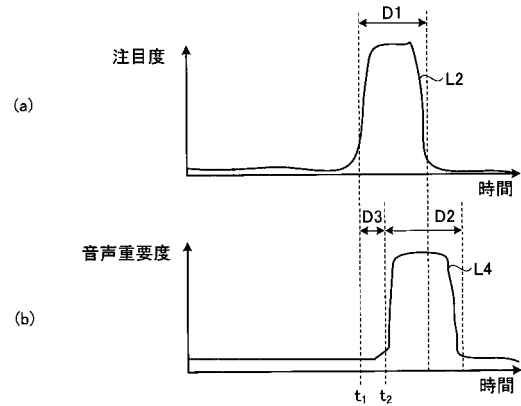
【図2】



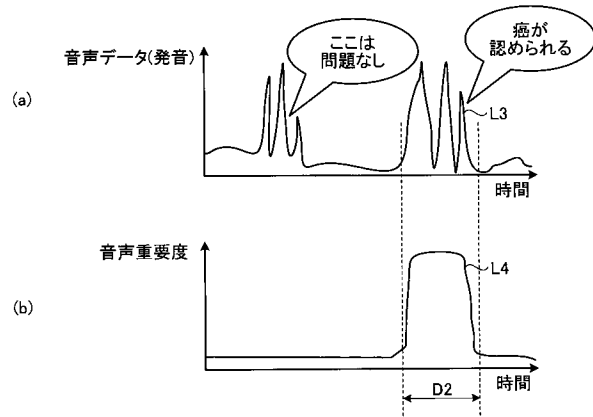
【図 3】



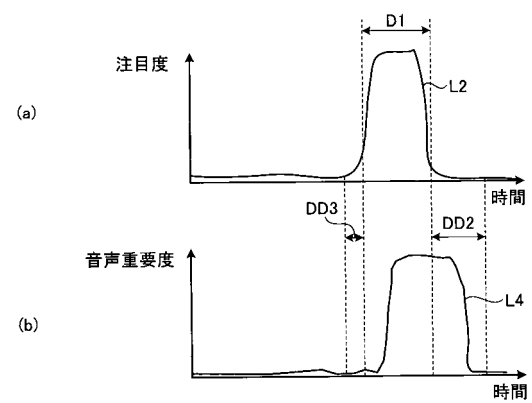
【図 5 A】



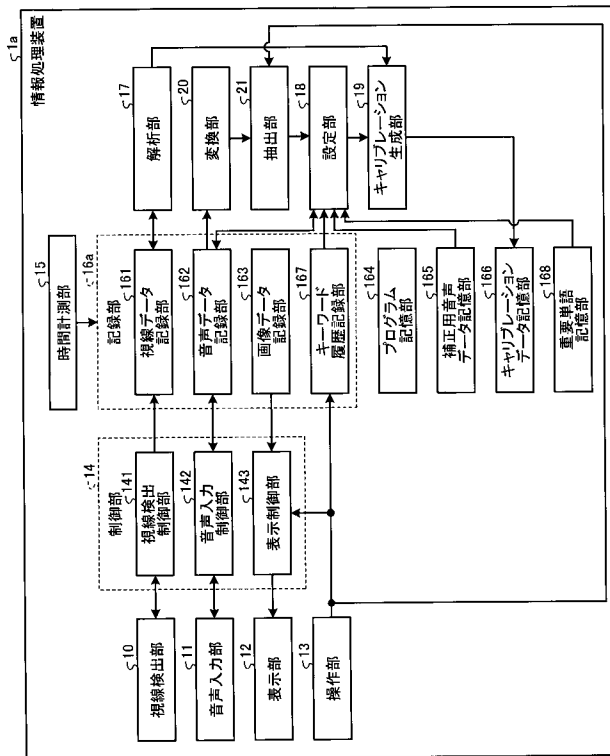
【図 4】



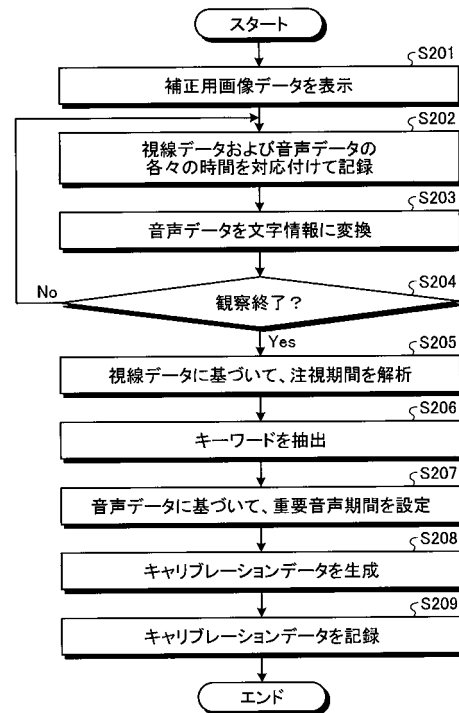
【図 5 B】



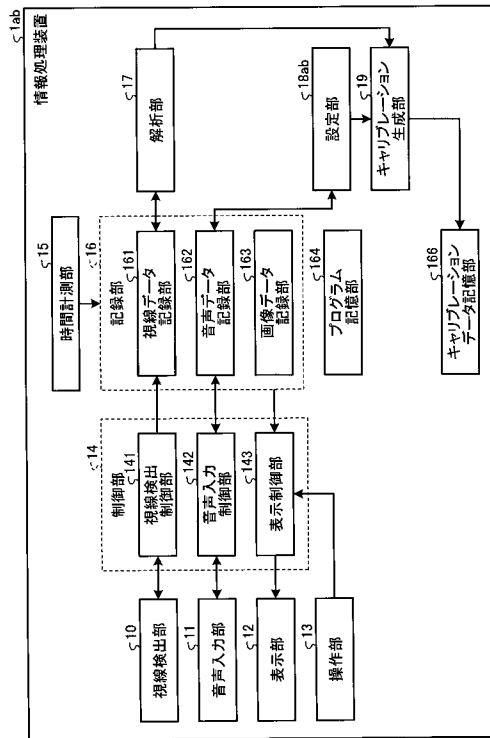
【図 6 A】



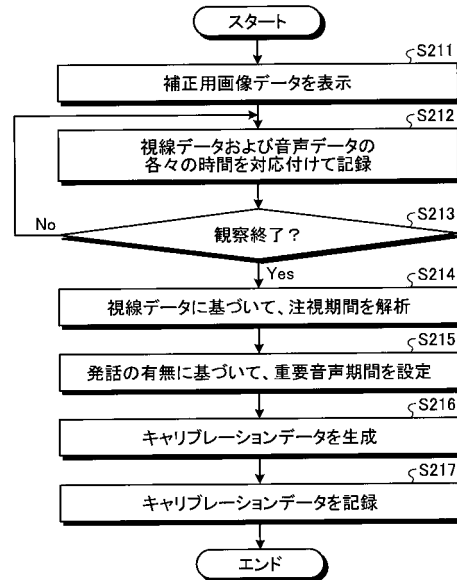
【図 6 B】



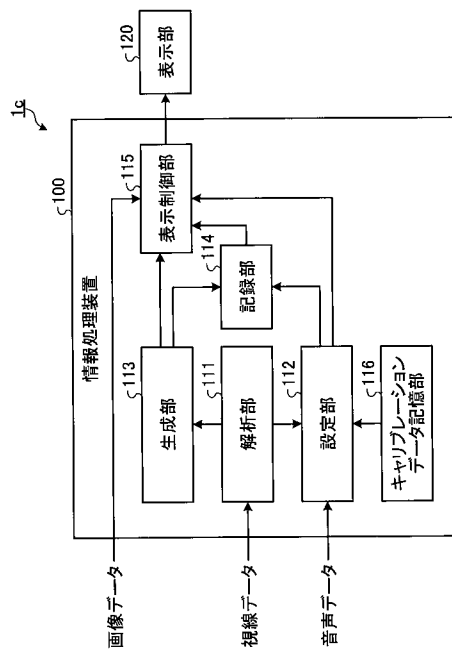
【 図 7 A 】



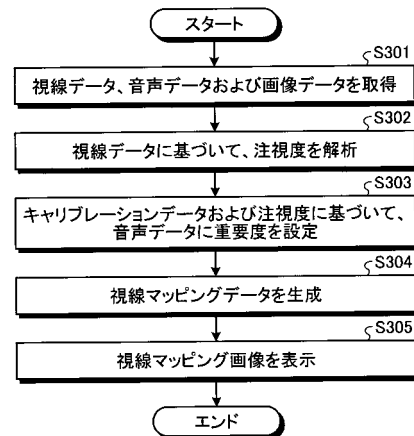
【 図 7 B 】



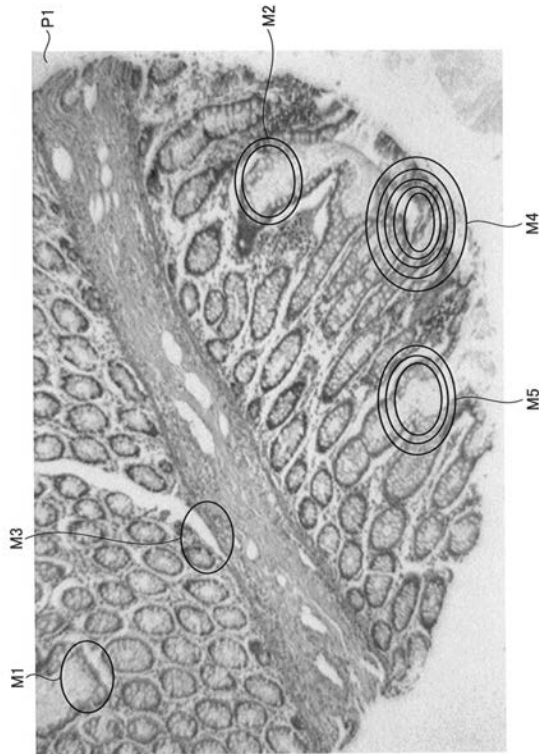
【 圖 8 】



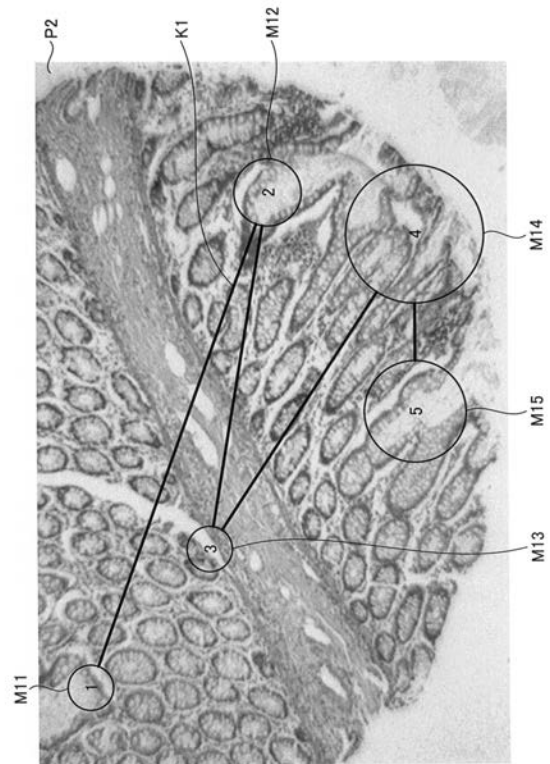
【 図 9 】



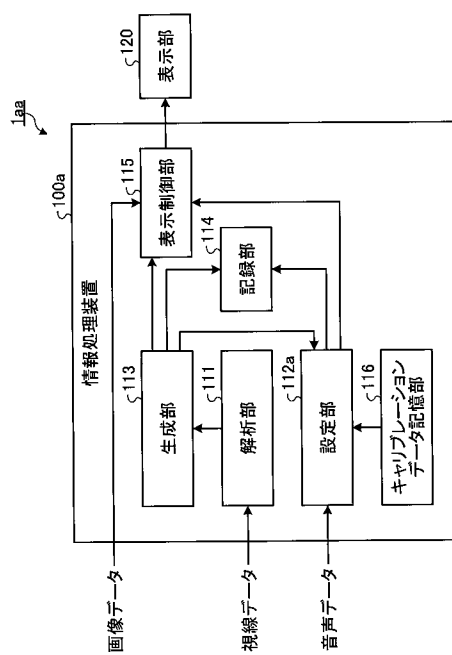
【図 10】



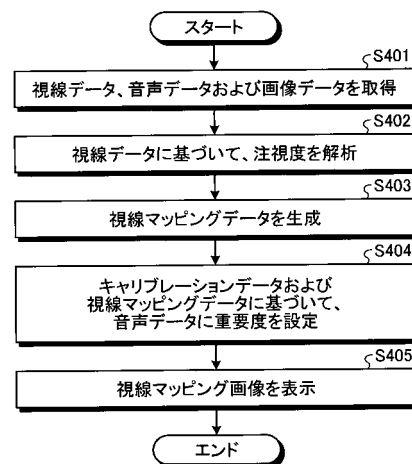
【図 11】



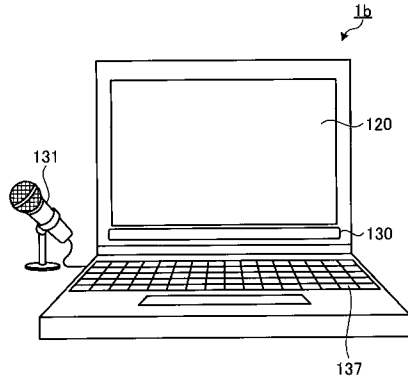
【図 12】



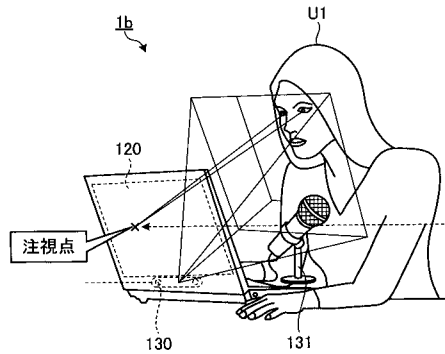
【図 13】



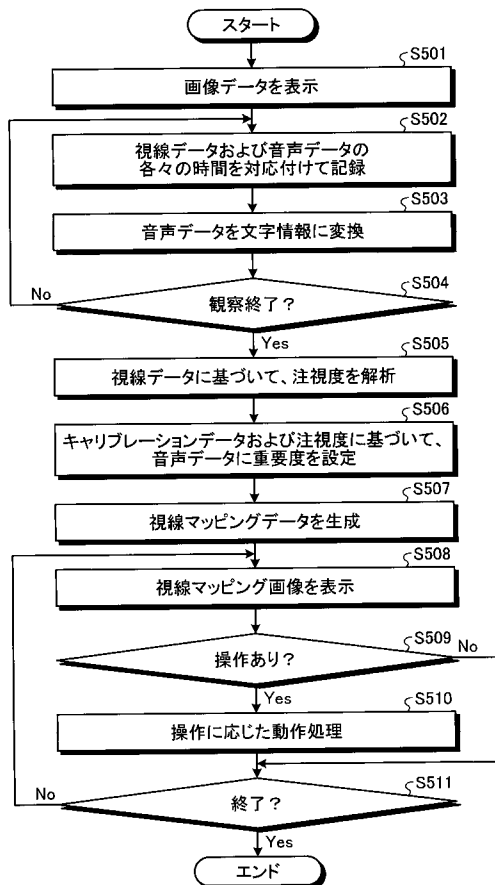
【図 14】



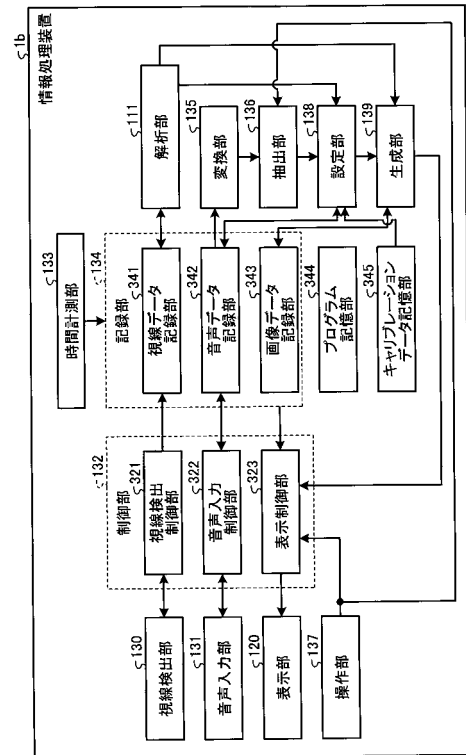
【図 15】



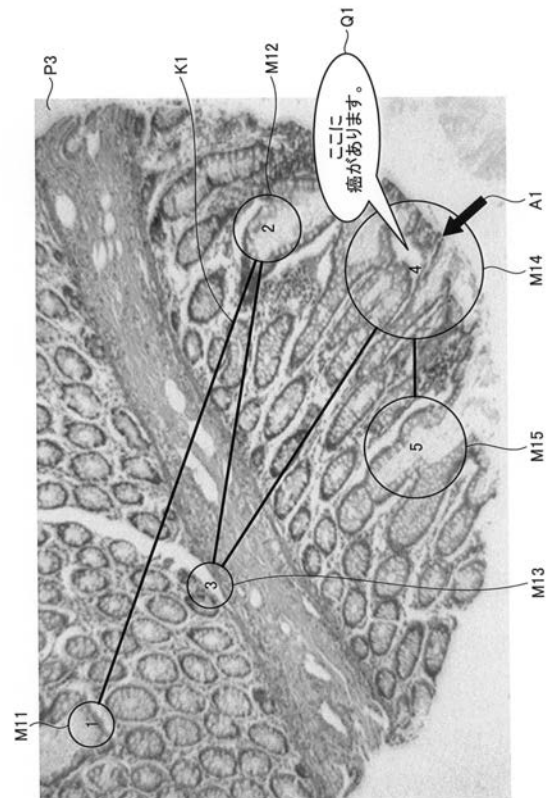
【図 17】



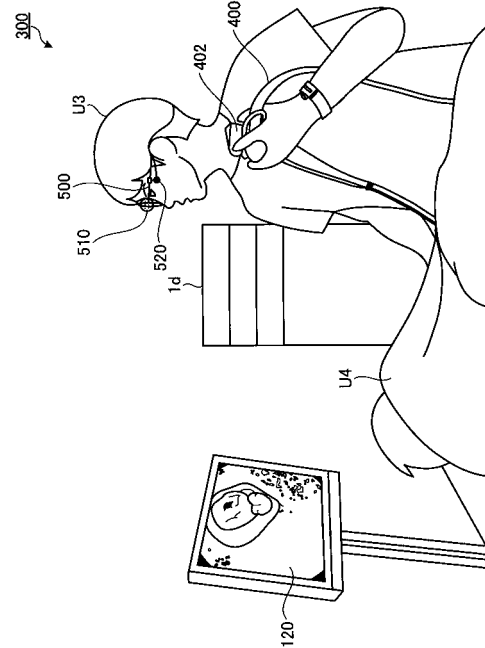
【図 16】



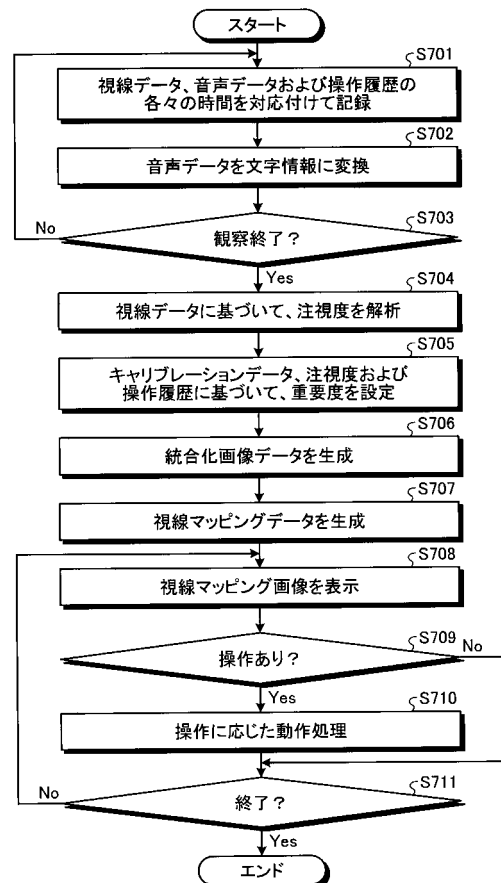
【図 18】



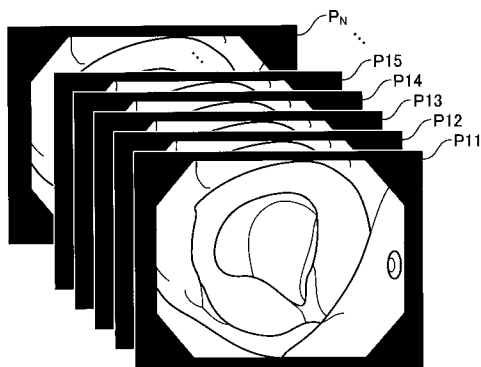
【 図 2 3 】



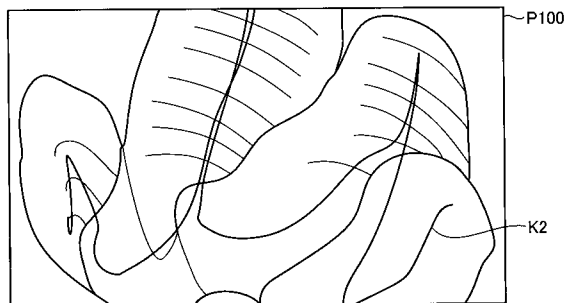
【 図 2 5 】



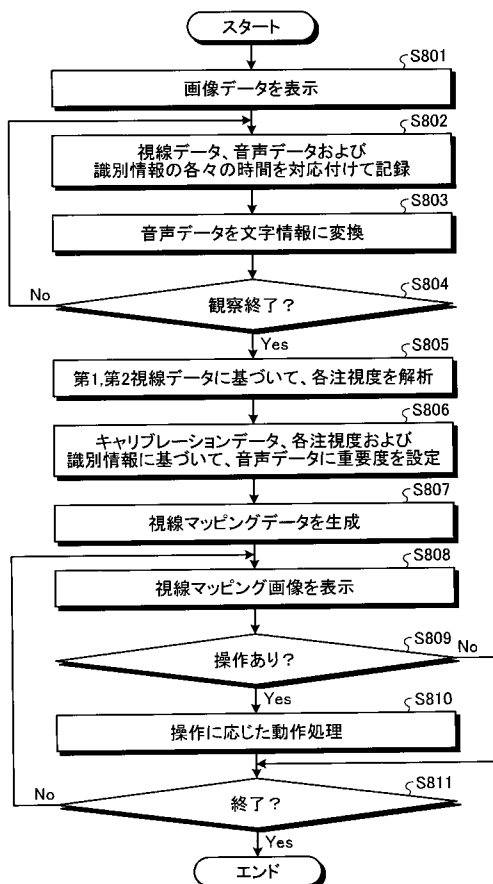
【図26】



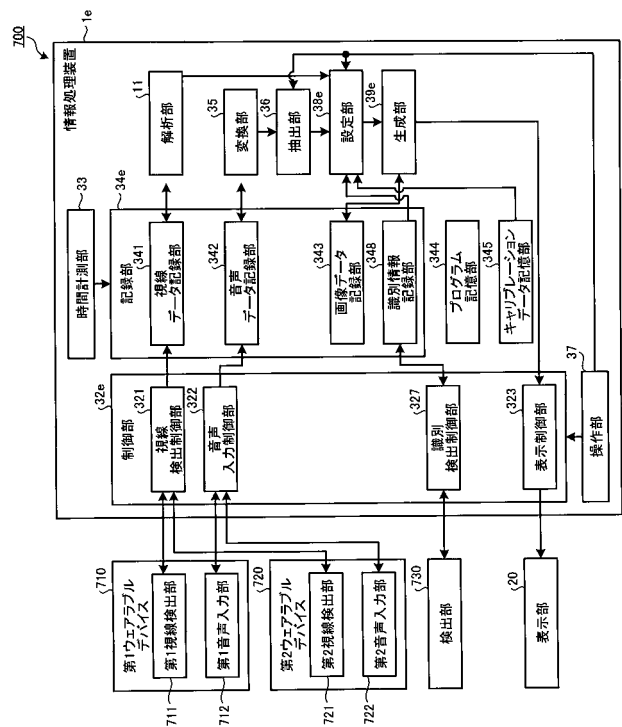
【図27】



【図29】



【図28】



フロントページの続き

(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 DA21 GA02 GA11

2H052 AD33 AF14

4C316 AA21

5E555 AA15 AA71 BA01 BB01 BC13 CA42 CA47 CB64 CB65 DA01

DB53 EA23

| | | | |
|----------------|---|---------|------------|
| 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 | JP2019202131A | 公开(公告)日 | 2019-11-28 |
| 申请号 | JP2019090107 | 申请日 | 2019-05-10 |
| [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 | A61B3/113 G06F3/01 G02B21/00 G02B23/24 G06F3/16 | | |
| FI分类号 | A61B3/113 G06F3/01.510 G02B21/00 G02B23/24.A G06F3/16.520 | | |
| F-TERM分类号 | 2H040/DA21 2H040/GA02 2H040/GA11 2H052/AD33 2H052/AF14 4C316/AA21 5E555/AA15 5E555/AA71 5E555/BA01 5E555/BB01 5E555/BC13 5E555/CA42 5E555/CA47 5E555/CB64 5E555/CB65 5E555/DA01 5E555/DB53 5E555/EA23 | | |
| 优先权 | 2018095520 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，其能够校正视线的稳定注视点与产生声音的时间之间的时间差。解决方案：信息处理设备1包括：分析单元17，用于分析凝视时段，在该凝视时段中，基于显示单元12显示的校正图像的多个观察位置中的每一个，用户的视线的凝视度等于或大于预定值 注视数据；设置单元18，基于用于校正的语音数据存储单元165记录的用于校正的语音数据，将针对语音数据说出语音的时间设置为重要语音时段；校准产生单元19产生注视周期和重要语音周期之间的时间差作为用户的校准数据，并将其记录在记录单元16中。

