

(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2006-155238
(P2006-155238A)

(43) 公開日 平成18年6月15日(2006.6.15)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 3/048 (2006.01)	G06F 3/00 654A	5B057
G06F 3/01 (2006.01)	G06F 3/00 680A	5E501
G06T 1/00 (2006.01)	G06T 1/00 340A	5L096
G06T 7/20 (2006.01)	G06T 7/20 300B	

審査請求 有 請求項の数 12 O L (全 27 頁)

(21) 出願番号	特願2004-344849 (P2004-344849)	(71) 出願人	504136568 国立大学法人広島大学 広島県東広島市鏡山1丁目3番2号
(22) 出願日	平成16年11月29日(2004.11.29)	(74) 代理人	110000338 特許業務法人原謙三国際特許事務所
		(74) 代理人	100080034 弁理士 原 謙三
		(74) 代理人	100113701 弁理士 木島 隆一
		(74) 代理人	100116241 弁理士 金子 一郎
		(72) 発明者	平嶋 宗 広島県東広島市鏡山1丁目4番1号 広島 大学大学院工学研究科内

最終頁に続く

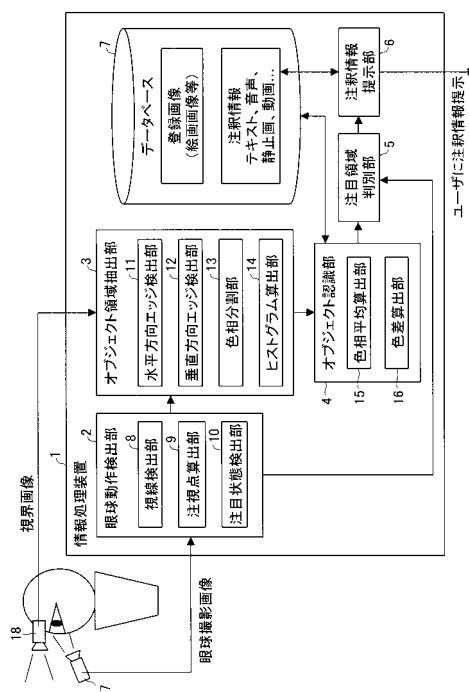
(54) 【発明の名称】 情報処理装置、携帯端末、情報処理方法、情報処理プログラム、およびコンピュータ読取可能な記録媒体

(57) 【要約】

【課題】 ユーザが注目しているオブジェクトの領域を判別し、その領域に関する注釈情報を的確にユーザに提示する。

【解決手段】 眼球撮影画像に基づき、ユーザがオブジェクトに注目している状態を検出する眼球動作検出部2と、眼球動作検出部2によりユーザが注目状態にあると検出された際、視界画像内においてユーザが注目しているオブジェクトを認識するオブジェクト認識部4と、上記オブジェクトにおいてユーザが注目している注目領域を判別する注目領域判別部5と、注目領域判別部5により判別された注目領域に関する注釈情報をユーザに提示する注釈情報提示部6とを備えている。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

ユーザの眼球運動を撮影した眼球撮影画像に基づき、ユーザがオブジェクトに注目している状態を検出する眼球動作検出手段と、

上記眼球動作検出手段によりユーザが注目状態にあると検出された際、ユーザの視界を撮影した視界画像内において、ユーザが注目しているオブジェクトを認識するオブジェクト認識手段と、

上記オブジェクトにおいてユーザが注目している注目領域を判別する注目領域判別手段と、

上記注目領域判別手段により判別された注目領域に関する注釈情報をユーザに提示する注釈情報提示手段とを備えていることを特徴とする情報処理装置。 10

【請求項 2】

上記眼球動作検出手段は、

上記眼球撮影画像から瞳孔領域を抽出する視線検出手段と、

上記瞳孔領域における中心座標から注視点の移動距離を算出する注視点算出手段と、

上記注視点の移動距離から、オブジェクトに対するユーザの視角を算出するとともに、その視角の大きさに基づき眼球の固視状態および跳躍状態を検出し、これら固視状態および跳躍状態の回数に基づきユーザがオブジェクトに注目している状態を検出する注目状態検出手段とを備えてい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 3】

上記注目状態検出手段は、ユーザが注目状態にある間の注視点の分布を内包するように複数の注視点を結ぶことにより、上記視界画像内に最小多角形領域を画定するものであるとともに、

上記注目領域判別手段は、データベース内に格納された登録画像における複数の領域のうち、上記最小多角形領域との重なりが最大となる領域を判断し、その領域に対応する上記視界画像内の領域を、上記注目領域として判別するものであることを特徴とする請求項 2 に記載の情報処理装置。

【請求項 4】

上記オブジェクト認識手段は、上記視界画像における色相と、データベース内に格納された上記視界画像の候補となり得る登録画像の色相とに基づき、ユーザが注目しているオブジェクトを認識するものであることを特徴とする請求項 1 ないし 3 のいずれか 1 項に記載の情報処理装置。 30

【請求項 5】

上記注釈情報提示手段は、さらに、上記注目領域と階層構造をなす領域に関する注釈情報をユーザに提示するものであることを特徴とする請求項 1 ないし 4 のいずれか 1 項に記載の情報処理装置。

【請求項 6】

請求項 1 ないし 5 のいずれか 1 項に記載の情報処理装置に用いる携帯端末であって、

上記眼球撮影画像を撮影する第 1 カメラを備えていることを特徴とする携帯端末。

【請求項 7】

請求項 1 ないし 5 のいずれか 1 項に記載の情報処理装置に用いる携帯端末であって、

上記視界画像を撮影する第 2 カメラを備えていることを特徴とする携帯端末。 40

【請求項 8】

請求項 1 ないし 5 のいずれか 1 項に記載の情報処理装置に用いる携帯端末であって、

上記眼球撮影画像を撮影する第 1 カメラと、上記視界画像を撮影する第 2 カメラとを備えていることを特徴とする携帯端末。

【請求項 9】

請求項 1 ないし 5 のいずれか 1 項に記載の情報処理装置に用いる携帯端末であって、

上記注釈情報提示手段から上記注釈情報を取り込み、その注釈情報の再生を行うことを特徴とする携帯端末。 50

【請求項10】

ユーザの眼球運動を撮影した眼球撮影画像に基づき、ユーザがオブジェクトに注目している状態を検出する第1ステップと、

上記第1ステップによりユーザが注目状態にあると検出された際、ユーザの視界を撮影した視界画像内において、ユーザが注目しているオブジェクトを認識する第2ステップと、

上記オブジェクトにおいてユーザが注目している注目領域を判別する第3ステップと、

上記第3ステップにて判別された注目領域に関する注釈情報をユーザに提示する第4ステップとを備えていることを特徴とする情報処理方法。

【請求項11】

請求項10に記載の情報処理方法の各ステップをコンピュータに実行させるための情報処理プログラム。

【請求項12】

請求項11に記載の情報処理プログラムを記録したことを特徴とするコンピュータ読取可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、拡張現実感に関するものであり、注目対象となるオブジェクトの注釈情報を適応的にユーザに提示し得る情報処理装置、携帯端末、情報処理方法、情報処理プログラム、およびコンピュータ読取可能な記録媒体に関するものである。

【背景技術】

【0002】

拡張現実感(Augmented Reality、以下ARと略す)とは、テキストやCG等の注釈情報を実世界のオブジェクトを重ねて、あるいは近い位置に提示することによって、実世界に情報を付加する技術であり、これによりユーザはオブジェクトの外観が持つ情報以上の情報を得ることができる。ARシステムでは、ユーザが興味を持つオブジェクトの情報を提示することを目的とするため、ユーザが注目しているオブジェクトを認識する必要がある。既存のシステムにおいて、オブジェクトの認識方法は、オブジェクト認識のためのタグ付けを行う方法と、行わない方法に大別でき、それぞれの方法により研究が盛んに行われている。

【0003】

オブジェクトにタグを付け、タグを読み取ることによりオブジェクトを認識し、関連付けられた情報を提示する方法(非特許文献1~4参照)では、2次元バーコードやRFIDタグ、赤外LEDをオブジェクトに取り付ける方法(非特許文献1~3参照)や、物理的なタグを貼り付けるのではなく、赤外線によりタグを投影する方法が提案されている(非特許文献4参照)。これらの方法では、注釈情報を提示するために、ユーザはタグリーダーやカメラなどのデバイスによりタグを読み取るという明示的な動作が必要となる。一方、タグ付けを行わずオブジェクトの認識を行い、関連付けられた情報の提示を行う方法(非特許文献5,6参照)では、事前に獲得しておいたオブジェクトの画像を基にして、ユーザに装着したカメラに映っているオブジェクトの認識を行っている。これらの方法では、オブジェクト全体が画像に映っている場合の認識精度は100%となっている。

【非特許文献1】小林元樹, 小池英樹, “電子情報の表示と操作を実現する机型実世界インターフェース「EnhancedDesk」”, インタラクティブシステムとソフトウェアV:日本ソフトウェア科学会WISS1997.

【非特許文献2】椎尾一郎, 増井俊之, 福地健太郎, “FieldMouseによる実世界インタラクティブシステム”, インタラクティブシステムとソフトウェアVII:日本ソフトウェア科学会WISS1999.

【非特許文献3】青木恒, “カメラで読み取る赤外線タグとその応用”, インタラクティブシステムとソフトウェアVIII:日本ソフトウェア科学会WISS2000.

10

20

30

40

50

【非特許文献4】白井良成，松下光範，大黒毅，“秘映プロジェクタ：不可視情報による実環境の拡張”，インタラクティブシステムとソフトウェアXI:日本ソフトウェア科学会WISS2003.

【非特許文献5】T. Kurata, T. Okuma, M. Kourogi, T. Kato, and K. Sakaue, “VizWear: Toward HumanCentered Interaction through Wearable Vision and Visualization”, The Second IEEE PacificRim Conference on Multimedia, 2001.

【非特許文献6】T. Jebara, B. Schiele, N. Oliver, A. Pentland, “DyPERS: Dynamic Personal Enhanced Reality System”, M.I.T. Media Lab. Perceptual Computing Section Technical Report, No. 468, 1998.

【非特許文献7】池田光男，“眼はなにを見ているか”，平凡社，1998.

10

【発明の開示】

【発明が解決しようとする課題】

【0004】

しかしながら、従来技術では、オブジェクトに関する情報を提示する際に、ユーザがそのオブジェクトに興味を持っているか否かの判断を行っていないため、ユーザにとって不必要な情報を提示する場合がある。また、オブジェクトの認識しか行っていないため、オブジェクトのどの部分に注目したかを検出することはできない。

【0005】

本発明は、上記従来の問題点に鑑みなされたものであって、ユーザが注目しているオブジェクトに関する注釈情報を的確にユーザに提示し得る情報処理装置、携帯端末、情報処理方法、情報処理プログラム、およびコンピュータ読取可能な記録媒体を提供することを目的とする。

20

【課題を解決するための手段】

【0006】

本発明の情報処理装置は、上記課題を解決するために、ユーザの眼球運動を撮影した眼球撮影画像に基づき、ユーザがオブジェクトに注目している状態を検出する眼球動作検出手段と、上記眼球動作検出手段によりユーザが注目状態にあると検出された際、ユーザの視界を撮影した視界画像内において、ユーザが注目しているオブジェクトを認識するオブジェクト認識手段と、上記オブジェクトにおいてユーザが注目している注目領域を判別する注目領域判別手段と、上記注目領域判別手段により判別された注目領域に関する注釈情報をユーザに提示する注釈情報提示手段とを備えていることを特徴としている。

30

【0007】

上記構成によれば、眼球動作検出手段にてユーザが注目状態にあることを検出した上で、オブジェクト認識手段によりユーザが注目しているオブジェクトが認識されるので、ユーザが真に興味を示しているオブジェクトを認識することができる。

【0008】

その上、注目領域判別手段により、オブジェクトにおいてユーザが注目している注目領域を判別するので、オブジェクトにおいてどの部分にユーザが注目しているかを把握することができる。そして、注釈情報提示手段は、上記注目領域に関する情報をユーザに提示するので、ユーザが真に興味を示している領域に関する注釈情報を的確にユーザに提示することが可能であり、不必要な情報がユーザに提示されることを防止できる。

40

【0009】

このように、本発明の情報処理装置によれば、ユーザが興味を示している領域に関する注釈情報が的確にユーザに提示されるので、本発明の情報処理装置の快適な利用環境をユーザに与えることができる。

【0010】

さらに、本発明の情報処理装置は、上記課題を解決するために、上記眼球動作検出手段が、上記眼球撮影画像から瞳孔領域を抽出する視線検出手段と、上記瞳孔領域における中心座標から注視点の移動距離を算出する注視点算出手段と、上記注視点の移動距離から、オブジェクトに対するユーザの視角を算出するとともに、その視角の大きさに基づき眼球

50

の固視状態および跳躍状態を検出し、これら固視状態および跳躍状態の回数に基づきユーザがオブジェクトに注目している状態を検出する注目状態検出手段とを備えていることを特徴としている。

【0011】

人間が何かに注目している際には、固視状態と跳躍状態とを頻繁に繰り返すことが知られている。したがって、この眼球運動の特徴を利用すれば、ユーザが注目状態にあることを的確に検出することができる。

【0012】

そこで、上記構成では、眼球動作検出手段に、視線検出手段と、注視点算出手段と、注目状態検出手段とが備えられている。

10

【0013】

すなわち、固視状態および跳躍状態は、ユーザの視角を検出することにより検出可能であり、ユーザの視角は、注視点の移動距離を算出すれば求めることができる。

【0014】

上記構成の情報処理装置は、視線検出手段によりたとえば2値化処理を用いて瞳孔領域を抽出する一方で、注視点算出手段により注視点の移動距離を算出する。そして、注目状態検出手段は、注視点算出手段により求められた注視点の移動距離からユーザの視角を算出して固視状態および跳躍状態の頻度を検出するので、的確にユーザの注目状態を検出することができる。

【0015】

このように、上記構成によれば、ユーザの注目状態を的確に把握することができるので、ユーザが興味を示しているオブジェクトがオブジェクト認識手段によりの確に認識され、注釈情報提示手段はさらに的確な注釈情報をユーザに提示することができる。

20

【0016】

さらに、本発明の情報処理装置は、上記課題を解決するために、上記注目状態検出手段が、ユーザが注目状態にある間の注視点の分布を内包するように複数の注視点を結ぶことにより、上記視界画像内に最小多角形領域を画定するものとともに、上記注目領域判別手段が、データベース内に格納された登録画像における複数の領域のうち、上記最小多角形領域との重なりが最大となる領域を判断し、その領域に対応する上記視界画像内の領域を、上記注目領域として判別するものであることを特徴としている。

30

【0017】

上記構成によれば、注目状態検出手段により最小多角形領域が画定される。この最小多角形領域は、注視点の分布を内包するように複数の注視点を結んで得られるものであるから、視界画像においてユーザが注目している領域を的確に表しているといえる。

【0018】

そして、注目領域判別手段は、上記最小多角形との重なりが最大となる領域を、データベース内の登録画像について判断し、その領域に対応する視界画像内の領域を注目領域として判断するので、ユーザが注目している領域を的確に注目領域として判別することができる。

【0019】

さらに、本発明の情報処理装置は、上記課題を解決するために、上記オブジェクト認識手段が、上記視界画像における色相と、データベース内に格納された上記視界画像の候補となり得る登録画像の色相とに基づき、ユーザが注目しているオブジェクトを認識するものであることを特徴としている。

40

【0020】

視界画像がたとえば絵画画像のような色彩のある画像であれば、その色相により視界画像に含まれるオブジェクトを特徴付けることが可能である。そこで、上記構成では、視界画像の色相と、登録画像の色相とに基づき、ユーザが注目しているオブジェクトを認識するので、オブジェクトの認識精度を高めることができる。

【0021】

50

さらに、本発明の情報処理装置は、上記課題を解決するために、上記注釈情報提示手段が、さらに、上記注目領域と階層構造をなす領域に関する注釈情報をユーザに提示するものであることを特徴としている。

【0022】

上記構成によれば、注目領域と階層構造をなす領域に関する注釈情報もユーザに提示されるので、ユーザは注目領域に関してより多くの情報を得ることができ、注目領域に関する理解を深めることができる。

【0023】

また、上記眼球撮影画像を撮影する第1カメラ、および上記視界画像を撮影する第2カメラの一方または双方を携帯端末に搭載することにより、ユーザが移動しても眼球撮影画像や視界画像を撮影することができるので、本発明の情報処理装置の利便性をさらに高めることができる。

10

【0024】

さらに、携帯端末により注釈情報の再生を行うことで、移動するユーザに対しても注釈情報の提示を行うことができるので、本発明の情報処理装置の利便性をより高めることができる。

【0025】

また、本発明の情報処理方法は、上記課題を解決するために、ユーザの眼球運動を撮影した眼球撮影画像に基づき、ユーザがオブジェクトに注目している状態を検出する第1ステップと、上記第1ステップによりユーザが注目状態にあると検出された際、ユーザの視界を撮影した視界画像内において、ユーザが注目しているオブジェクトを認識する第2ステップと、上記オブジェクトにおいてユーザが注目している注目領域を判別する第3ステップと、上記第3ステップにて判別された注目領域に関する注釈情報をユーザに提示する第4ステップとを備えていることを特徴としている。

20

【0026】

上記構成の情報処理方法によれば、本発明の情報処理装置と同様の作用効果を得ることができる。

【0027】

なお、コンピュータに上記情報処理方法における各ステップを実行させる情報処理プログラムにより、コンピュータを用いて本発明の情報処理方法と同様の作用効果を得ることができる。さらに、上記情報処理プログラムをコンピュータ読取り可能な記録媒体に記憶させることにより、任意のコンピュータ上で上記情報処理プログラムを実行させることができる。

30

【発明の効果】

【0028】

本発明の情報処理装置によれば、オブジェクトに全体的に注目しているならばオブジェクト全体に関する注釈情報を提示し、オブジェクトに部分的に注目しているならばその部分に関する注釈情報を提示するなど、ユーザが注目した箇所に対応した注釈情報を適応的にユーザに提示することが可能となる。

【発明を実施するための最良の形態】

40

【0029】

〔1. 情報処理装置の構成〕

本発明の一実施形態に係る情報処理装置について、図1～図16に基づいて説明する。図1に示すように、情報処理装置1は、眼球動作検出部2と、オブジェクト領域抽出部3と、オブジェクト認識部4と、注目領域判別部5と、注釈情報提示部6と、データベース7とを備えている。

【0030】

眼球動作検出部2は、ユーザが何かに注目していることを検出するために、眼球を撮影するカメラ17の画像（以下、眼球撮影画像とする）から、ユーザの視線の動きを検出するものである。より具体的には、眼球動作検出部2は、視線検出部8と、注視点算出部9

50

と、注目状態検出部 10 とを備えている。これらの眼球動作検出部 2 に設けられた各ブロックにおける処理内容については後述する。

【0031】

オブジェクト領域抽出部 3 は、ユーザが何かに注目している際に、視界を撮影するカメラ 18 の画像（以下、視界画像とする）から、オブジェクト領域を抽出するものである。オブジェクト領域の抽出手法に関しては、種々の方法を採用することができるが、ここではその一例として、オブジェクト領域抽出部 3 に設けられる水平方向エッジ検出部 11、垂直方向エッジ検出部 12、色相分割部 13、およびヒストグラム算出部 14 によりオブジェクト領域を抽出する手順について説明する。なお、これらのオブジェクト領域抽出部 3 に設けられた各ブロックにおける処理内容の詳細については後述する。

10

【0032】

オブジェクト認識部 4 は、視界画像からオブジェクト領域抽出部 3 により抽出された領域に映っている画像と、データベース 7 内に格納された登録画像とのマッチング処理を行い、ユーザが注目している画像を認識するものである。なお、登録画像とは、予めデータベース内に格納された、視界画像の候補となる画像である。

【0033】

このマッチング処理の手法に関しても、オブジェクト領域の抽出手法と同様に種々の方法を採用することができる。ここでは、マッチング処理の一例として、オブジェクト認識部 4 に設けられた色相平均算出部 15 および色差算出部 16 を用いて、オブジェクト領域抽出部 3 により抽出された画像とデータベース 7 内の登録画像とのそれぞれにおける色相

20

の平均値・色差に着目する手法について説明する（詳細は後述する）。

【0034】

注目領域判別部 5 は、オブジェクト認識部 4 によりユーザが注目していると認識された画像において、ユーザがどの領域に注目しているかさらに詳細に判別するものである。この注目領域判別部 5 は、データベース 7 に登録された画像における領域であって、注視点座標に基づいて決定される最小凸多角形領域（詳細は後述する）との重なりが最大となる領域を、ユーザが注目していると判断する。注目領域の判別処理についての詳細は後述する。

【0035】

注釈情報提示部 6 は、注目領域判別部 5 によりユーザが注目していると判別された領域に関する注釈情報（たとえばテキストデータ、音声データ、静止画像データ、動画像データ）をデータベース 7 から読み出し、情報処理装置 1 の外部に設けられた出力装置（たとえばノートパソコン、ヘッド・マウンテッド・ディスプレイ、ヘッドホン、スピーカ）に出力するためのものである。

30

【0036】

データベース 7 は、ユーザにその注釈情報が提供される登録画像と、登録画像内のオブジェクトに関するデータとして、オブジェクトの色、オブジェクト上で注釈情報が付加されている領域の座標、その注釈情報とを格納している。

【0037】

なお、カメラ 17 は、たとえば CCD カメラであり、眼球の下部に設置される。また、カメラ 18 はたとえば CCD カメラであり、眉間に設置される。これら 2 つのカメラ 17・18 は、ユーザが正面を見ているときの眼球と、そのときの視界の中心とが、それぞれのカメラ画像において中心に映るような位置関係に配置される。

40

【0038】

上記構成により、情報処理装置 1 は、ユーザの眼球運動から、ユーザがオブジェクトに注目している状態を検出することにより、ユーザが注目しているオブジェクトの注釈情報を提示することができる。この注釈情報が提示されるまでの一連の処理を、図 2 のフローチャートを用いて説明する。

【0039】

まず、図 2 に示すように、ユーザが何かのオブジェクトに注目している状態が、眼球動

50

作検出部 2 により検出される (ステップ 1、以下ステップを単に S と記載する)。その後、オブジェクト領域抽出部 3 により、視界画像が獲得されるとともに (S 2)、その視界画像からオブジェクト領域の抽出が行われる (S 3)。さらに、オブジェクト認識部 4 によりユーザが注目している画像の認識が行われ (S 4)、ユーザが注目していると認識された画像において、ユーザがどの領域に注目しているか、注目領域判別部 5 によりさらに詳細に判別される (S 5)。そして、注釈情報提示部 6 により、ユーザが注目していると判別された領域に関する注釈情報がユーザに提示され (S 6)、一定時間経過した後に注釈情報の提示が終了する (S 7)。

【0040】

最後に、注目の検出を再開するか否かの判断が行われ (S 8)、再開する場合には S 1 の処理に戻り、再開しない場合は、一連の処理が終了する。なお、S 8 における判断は、たとえば、スイッチ、キーボード、マウス等の入力デバイスから、ユーザにより「注目の検出を再開する」旨の入力があつたか否かを判断することにより行うことができる。

10

【0041】

本実施形態における情報処理装置 1 は、上述した構成および処理フローにより、ユーザが注目している領域に関する注釈情報を、的確にユーザに提示することができる。以下、情報処理装置 1 における各ブロックにより行われる処理の詳細について説明する。

【0042】

〔2. 眼球動作検出部 2 における処理〕

眼球動作検出部 2 は、上述したように、ユーザが何かに注目していることを検出するために、眼球撮影画像から、ユーザの視線の動きを検出するものである。なお、以下の説明では、ユーザが注目状態にあることが検出されたときの、注目開始時から注目終了時までを区間を「注目区間」と定義する。

20

【0043】

また、眼球動作検出部 2 は、上述したとおり、視線検出部 8 と、注視点算出部 9 と、注目状態検出部 10 とを備えており (図 1 参照)、視線検出部 8 によりユーザの視線を検出するとともに、注視点算出部 9 により視線に対応した注視点座標を算出する。なお、注視点座標とは、ユーザが実世界において視線を合わせていた点 (以下、注視点と呼ぶ) を、視界画像における座標で示したものとする。また、注目区間における注視点の広がりをも「注視点分布」と定義する。

30

【0044】

〔2-1. 視線検出部 8 における処理〕

視線検出部 8 によるユーザの視線の検出は、瞳孔の位置を検出することにより行われる。つまり、より正確な視線の位置を求めるため、カメラ 17 から眼球に赤外線が照射されており、瞳孔と虹彩とのコントラストが強調されている。そして、視線検出部 8 は、眼球撮影画像の二値化処理、つまり瞳孔とそうでない部分とを区別することによって、瞳孔領域を抽出する。

【0045】

さらに、視線検出部 8 は、瞳孔領域の中心座標 $(x(t), y(t))$ を、時刻 t (フレーム) における瞳孔の位置として算出する。なお、眼球撮影画像の水平方向を x 軸、垂直方向を y 軸とし、原点を画像の左下として設定している。なお、以下の説明における各種設定値は、画像の大きさを水平方向 160 ピクセル、垂直方向 120 ピクセルとし、毎秒 10 フレームで獲得する場合の一例であり、獲得画像の大きさの変更やここで想定している物とは特性の異なるカメラ等の使用を前提とした場合は適宜最適な値を採用することができる。

40

【0046】

〔2-2. 注視点算出部 9 における処理〕

次に、注視点算出部 9 により、注視点座標 $(G_u(t), G_v(t))$ を、以下の式 (1) によって求める。なお、視界画像における水平方向を u 軸、垂直方向を v 軸とし、原点を画像の左下とする。

【0047】

50

【数 1】

$$\begin{cases} G_u(t) = Cv(i, j)_u - (x(t) - D(i, j)_x) \times \alpha(i, j) \\ G_v(t) = Cv(i, j)_v + (y(t) - D(i, j)_y) \times \beta(i, j) \end{cases} \dots \text{式 (1)}$$

【0048】

ここで、視界画像および眼球撮影画像の一部領域(40 x 120、40 y 80)を4×4の16ブロックに分割する。また、 $(D(i, j)_x, D(i, j)_y)$ ($i, j=1, \dots, 4$)は眼球撮影画像のブロック(i, j)の中心座標とし、 $(Cv(i, j)_u, Cv(i, j)_v)$ は視界画像のブロック(i, j)の中心座標とする。

10

【0049】

なお、ブロック(i, j)とは、画像の左下から数えて、水平方向にi番目、垂直方向にj番目のブロックのことである。また、 (i, j) 、 (i, j) は各ブロックにおける、瞳孔の中心の移動距離に対する注視点の移動距離の比を表す値である。また、瞳孔の中心の移動距離dは、以下の式(2)に基づき、注視点算出部9により単位をピクセルとして算出される。

【0050】

【数 2】

$$d = \sqrt{(x(t) - x(t-1))^2 + (y(t) - y(t-1))^2} \dots \text{式 (2)}$$

20

【0051】

〔2-3. 注目状態検出部10における処理〕

眼球運動は人間の注意や関心を表しており、人間が静止物体に注目している場合は、約300ミリ秒間の固視状態と、約30ミリ秒間に起こる跳躍運動とを頻繁に繰り返すことが知られている(非特許文献7参照)。注目状態検出部10は、この眼球運動の特徴を利用することにより、ユーザがオブジェクトに注目している状態を検出するものである。

【0052】

すなわち、注目状態検出部10は、瞳孔の位置が3フレームの間(約300ミリ秒)動かないか、または微小に動いており、4フレーム目で瞳孔の位置が跳躍する状態を検出し、この状態を固視・跳躍運動として検出する。そして、注目状態検出部10は、固視・跳躍運動が連続して3回以上検出されたとき、ユーザが注目状態にあると判定する。

30

【0053】

ただし、注目状態検出部10は、3フレーム間で、瞳孔の位置の移動が視角にして 2.1° 未満であれば固視とみなし、固視状態から瞳孔の位置の移動が視角にして 2.1° 以上であれば跳躍とみなす。また、跳躍が生じてから3秒間で次の跳躍が生じないとき、注目状態検出部10は固視・跳躍が終了したと判断する。なお、 2.1° なる値は、固視であるか跳躍であるかを区別するしきい値の一例であり、他の値であっても構わない。

【0054】

ここで、視角 θ は、以下の式(3)に基づいて算出される。

40

【0055】

【数 3】

$$\theta = \tan^{-1} \left(\frac{d \times 0.024}{2.0} \right) \dots \text{式 (3)}$$

【0056】

上記式(3)における「0.024」という数値は、眼球撮影画像における1ピクセルあたりの距離(cm)である。もちろん、画像の精細度に応じて、0.024なる値は、他の値となっても構わない。また、式(3)における「2.0」という数値は、眼球と眼球を撮影するカメラ17との距離(cm)である。もちろん、眼球の撮影条件に応じて、2.0なる値は、他の

50

値にされてもよい。

【0057】

また、注目区間において発生する注視点座標のそれぞれは、ユーザが注目状態であったときの瞳孔の位置から算出しているため、ユーザが注目しているオブジェクト上、またはその付近に現れる。そのため、注目状態検出部10は、1注目区間、すなわち注目の始まりから注目の終わりを検出するまでの間の注視点分布が内包されるように、複数の注視点座標を結ぶことにより最小凸多角形領域（最小多角形領域）を画定する。したがって、この最小凸多角形領域には、注目していたオブジェクト全体、またはその一部が含まれることになる。

【0058】

〔2-4. 眼球動作検出部2における処理フロー〕

次に、眼球動作検出部2において実行される、ユーザが何かに注目していることを検出するための処理について、図3のフローチャートを用いて説明する。

【0059】

まず、注目状態検出部10により、連続回数が0にリセットされるとともに(S10)、固視回数も0にリセットされる(S11)。なお、連続回数とは、固視・跳躍運動が連続した回数を示している。以下、これと同様の意味において「連続回数」の文言を用いる。

【0060】

その後、視線検出部8により、眼球撮影画像がカメラ17から獲得される(S12)。さらに、視線検出部8は、S12において獲得された眼球撮影画像を2値化処理することで、眼球撮影画像から瞳孔領域を抽出する(S13)。

【0061】

その後、注視点算出部9は、上記式(2)により瞳孔の中心の移動距離を算出するとともに(S14)、視線検出部8が求めた瞳孔領域の中心座標(x(t)、y(t))から、上記式(1)に基づいて注視点座標($G_u(t)$ 、 $G_v(t)$)を算出する(S15)。

【0062】

そして、注目状態検出部10は、上述の式(3)に基づき、視角 θ を算出し(S16)、その値が 2.1° 以上となるか否かを判断する(S17)。視角 θ が 2.1° 以上でない場合、注目状態検出部10はユーザが固視状態にあると判断し、固視回数を1回増やす(S18)。そして、S18の処理が終了したら、再度S12の処理に移行する。

【0063】

一方、S17において視角 θ が 2.1° 以上であると判断された場合、注目状態検出部10は、ユーザの視線が跳躍したと判断し(S19)、現在までの固視回数が3回以上か否かを判断する(S20)。

【0064】

S20において固視回数が3回以上ではないと判断されれば、注目状態検出部10により固視回数が0にリセットされてから(S21)、S12の処理に移行する。また、S20において固視回数が3回以上であると判断したら、注目状態検出部10は、固視回数を0にリセットした後に(S22)、ユーザが固視・跳躍状態にあると判断する(S23)

【0065】

そして、S23の後、注目状態検出部10は、最後の跳躍が生じてから3秒以内に次の跳躍が生じるか否かを判断する(S24)。そして、S24において3秒以内に次の跳躍が生じたと判断したら、注目状態検出部10は、ユーザの固視・跳躍運動が連続しているものと判断して、連続回数を1回増やす(S25)。そして、S25の処理が終了したら、再度S12の処理に移行する。

【0066】

また、S24において3秒以内に次の跳躍が生じなかったと判断したら、注目状態検出部10は、連続回数が3回以上か否かを判断する(S26)。S26において連続回数が

10

20

30

40

50

3回以上でなければ、注目状態検出部10は、連続回数を0にリセットする(S27)。S27の後、S12の処理に移行する。また、S26において連続回数が3回以上であれば、注目状態検出部10は、ユーザが注目状態にあるものと判断する(S28)。

【0067】

このようにしてS10～S28のステップを踏むことにより、眼球動作検出部2による注目状態の検出処理が終了する。

【0068】

〔3. オブジェクト領域抽出部3における処理〕

オブジェクト領域抽出部3は、上述したとおり、視界画像からオブジェクト領域を抽出するものである。本実施形態では、美術館等で展示されている絵画をオブジェクトの一例として、注目時の視界画像に映っている絵画領域を抽出する方法について説明する。

10

【0069】

オブジェクト領域抽出部3がどのようにして絵画領域を抽出するかについて、簡単に説明する。絵画が映っている視界画像において、壁と額縁との境界で水平方向・垂直方向に直線のエッジが現れる。オブジェクト領域抽出部3は、水平方向エッジを水平方向エッジ検出部11により検出するとともに、垂直方向エッジを垂直方向エッジ検出部12により検出する。

【0070】

また、HSV表色系の色相において、壁と額縁との境界で色相が変化する。オブジェクト領域抽出部3は、この色相の変化を色相分割部13およびヒストグラム算出部14により検出し、直線のエッジと色相の境界とに基づいて、視界画像から絵画領域を抽出する。

20

【0071】

次に、オブジェクト領域抽出部3における一連の処理について、図4のフローチャートを用いて説明する。図4に示すように、まず、水平方向エッジ検出部11および垂直方向エッジ検出部12のいずれか一方、または双方により、視界画像の濃淡画像から、たとえばrobertsフィルタを用いてエッジ画像が生成される(S30)。S30の後、水平方向エッジ検出部11により、水平方向に現れる直線エッジの長さが算出される(S31)。また、S31の後、垂直方向エッジ検出部12により、垂直方向に現れる直線エッジの長さが算出される(S32)。

【0072】

なお、S32は必ずしもS31の後に行われる必要は無く、S31の前にS32が実行されてもよいし、S31の処理とS32の処理とを並行して行ってもよい。これらのS31およびS32における処理の詳細については後述する。

30

【0073】

さらに、水平方向エッジ検出部11は、水平方向で隣り合うエッジでより長いエッジを検出し、額縁のエッジ候補とする。また、垂直方向エッジ検出部12は、垂直方向で隣り合うエッジでより長いエッジを検出し、額縁のエッジ候補とする。なお、図4のフローチャートでは、上述の水平方向エッジ検出部11および垂直方向エッジ検出部12によるエッジ候補の検出処理を、まとめてS33として記載している。

【0074】

その後、色相分割部13により、エッジ画像の色相が7等級に分割される(S34)。そして、ヒストグラム算出部14により、額縁のエッジ候補から色相のヒストグラムが算出される(S35)。なお、S34およびS35における処理の詳細に関しては、後述する。

40

【0075】

その後、全ての額縁のエッジ候補に対して、ヒストグラム算出部14は、エッジ候補の両側の領域に関して、色相のヒストグラムの差CS₀を、以下の式(4)により算出する(S36)。

【0076】

【数 4】

$$CS_e = \sum_{n=1}^7 |hist_{l_n} - hist_{r_n}| \quad \dots \text{式 (4)}$$

【0077】

なお、 $hist_{l_n}$ は、額縁のエッジ候補を境界にして左（または上）に1画素ずらした領域における色相のヒストグラムであり、 $hist_{r_n}$ は、額縁のエッジ候補を境界にして右（または下）に1画素ずらした領域における色相のヒストグラムである。なお、 n は1から7までの整数であり、色相の等級を示している。なお、色相の等級数を7とすれば、視界画像に映っている壁、額縁、絵画の境界で色相が大まかに分かれることは実験により確認されている。

10

【0078】

式(4)による CS_e の算出処理が終了したら、水平方向エッジ検出部11は、 CS_e の大きさが上位となるエッジを水平方向において2本検出する(S37)。また、垂直方向エッジ検出部12は、 CS_e の大きさが上位となるエッジを垂直方向において2本検出する(S37)。そして、オブジェクト領域抽出部3は、S37において検出された4本のエッジで囲まれた領域を、絵画領域として抽出する(S38)。

【0079】

上述したS30～S38の一連のステップを踏むことにより、オブジェクト領域抽出部3は、視界画像からオブジェクト領域としての絵画領域を抽出する。

20

【0080】

〔3-1. 水平方向のエッジの長さの算出処理〕

次に、上述した図4のフローにおけるS31の処理の詳細について説明する。たとえば、図5(a)に示すような視界画像が得られたとする。この場合、図6に示すように、水平方向エッジ検出部11は、視界画像において $y = a$ ($1 \leq a \leq 118$ 、 a は整数)で示される直線に関して、以下の処理を行う。

【0081】

まず、水平方向エッジ検出部11は、エッジとなる画素が直線 $y = a$ 上において連続する数(連続画素数)を、0にリセットする(S40)。そして、水平方向エッジ検出部11は、 a の値を固定した状態で、 x 座標の値を示す b の値を、1から159までの整数値として変更することにより、以下のS41～S45の処理を行う。

30

【0082】

まず、S41において、水平方向エッジ検出部11は、座標 $(b, a - 1)$ 、 (b, a) 、および $(b, a + 1)$ における画素のうち、いずれかがエッジとなる画素であるか否かを判断する。

【0083】

S41における判断がYESの場合、つまり、座標 $(b, a - 1)$ 、 (b, a) 、および $(b, a + 1)$ における画素のうちいずれかがエッジとなる画素ならば、水平方向エッジ検出部11は、座標 (b, a) の画素をエッジとみなす(S42)。その後、水平方向エッジ検出部11は、座標 $(b - 1, a)$ 、および (b, a) における画素が、ともにエッジとなる画素であるか否かを判断する(S43)。

40

【0084】

なお、このように $(b, a - 1)$ および $(b, a + 1)$ の画素がエッジとなるか否かを考慮するのは、求めたいエッジが大まかに直線となるエッジだからである。また、S41における判断がNOの場合、つまり、座標 $(b, a - 1)$ 、 (b, a) 、および $(b, a + 1)$ における画素のうちどれもエッジとなる画素ではないと判断されれば、水平方向エッジ検出部11は、S42の処理を行うことなくS43の処理を行う。

【0085】

S43において、座標 $(b - 1, a)$ および (b, a) における画素が、ともにエッジ

50

となる画素であると判断したら、水平方向エッジ検出部 11 は、連続画素数を 1 増やす (S 44)。一方、S 43 において、座標 (b - 1, a) および (b, a) における画素が、ともにエッジとなる画素ではないと判断したら、水平方向エッジ検出部 11 は、連続画素数を 0 にリセットする (S 45)。

【0086】

このようにして、S 41 ~ S 45 の処理が、b の値を 1 から 159 までの間に含まれる整数値として変化させつつ行われる。そして、各整数値について求められた連続画素数のうち、最大となる連続画素数を、水平方向エッジ検出部 11 は、 $y = a$ における直線エッジの長さとして算出する (S 46)。

【0087】

上述した S 40 ~ S 46 における処理は、a の値を 1 から 118 までの間に含まれる整数値として変化させつつ行われる。このようにして求められた水平方向のエッジの長さを、図 5 (b) の右側のグラフに示す。

【0088】

〔3 - 2 . 垂直方向のエッジの長さの算出処理〕

次に、上述した図 4 のフローにおける S 32 の処理の詳細について説明する。たとえば、図 5 (a) に示すような視界画像が得られたとする。この場合、図 7 に示すように、垂直方向エッジ検出部 12 は、視界画像において $x = b$ ($1 \leq b \leq 158$ 、b は整数) で示される直線に関して、以下の処理を行う。

【0089】

まず、垂直方向エッジ検出部 12 は、エッジとなる画素が直線 $x = b$ 上において連続する数 (連続画素数) を、0 にリセットする (S 50)。そして、垂直方向エッジ検出部 12 は、b の値を固定した状態で、y 座標の値を示す a の値を、1 から 119 までの整数値として変更することにより、以下の S 51 ~ S 55 の処理を行う。

【0090】

まず、S 51 において、垂直方向エッジ検出部 12 は、座標 (b - 1, a), (b, a), および (b + 1, a) における画素のうち、いずれかがエッジとなる画素であるか否かを判断する。

【0091】

S 51 における判断が YES の場合、つまり、座標 (b - 1, a), (b, a), および (b + 1, a) における画素のうちいずれかがエッジとなる画素ならば、垂直方向エッジ検出部 12 は、座標 (b, a) の画素をエッジとみなす (S 52)。その後、垂直方向エッジ検出部 12 は、座標 (b, a - 1), および (b, a) における画素が、ともにエッジとなる画素であるか否かを判断する (S 53)。

【0092】

また、S 51 における判断が NO の場合、つまり、座標 (b - 1, a), (b, a), および (b + 1, a) における画素のうちどれもエッジとなる画素ではないと判断されれば、垂直方向エッジ検出部 12 は、S 52 の処理を行うことなく S 53 の処理を行う。

【0093】

S 53 において、座標 (b, a - 1) および (b, a) における画素が、ともにエッジとなる画素であると判断したら、垂直方向エッジ検出部 12 は、連続画素数を 1 増やす (S 54)。一方、S 53 において、座標 (b, a - 1) および (b, a) における画素が、ともにエッジとなる画素ではないと判断したら、垂直方向エッジ検出部 12 は、連続画素数を 0 にリセットする (S 55)。

【0094】

このようにして、S 51 ~ S 55 の処理が、a の値を 1 から 119 までの間に含まれる整数値として変化させつつ行われる。そして、各整数値について求められた連続画素数のうち、最大となる連続画素数を、垂直方向エッジ検出部 12 は、 $x = b$ における直線エッジの長さとして算出する (S 56)。

【0095】

10

20

30

40

50

上述した S 5 0 ~ S 5 6 における処理は、b の値を 1 から 1 5 8 までの間に含まれる整数値として変化させつつ行われる。

【 0 0 9 6 】

〔 3 - 3 . 色相の分割処理およびヒストグラムの算出処理 〕

次に、上述した図 4 のフローにおける S 3 4 および S 3 5 の処理の詳細について説明する。まず、色相分割部 1 3 は、図 8 に示すように、視界画像全体から 3 6 0 等級の色相のヒストグラムを算出し、頻度が最大となる色相を検出する (S 6 0)。そして、色相分割部 1 3 は、S 6 0 において検出した色相を中心として $\pm 2 5 ^\circ$ の範囲を、一等級の範囲として設定する (S 6 1)。

【 0 0 9 7 】

その後、色相分割部 1 3 は、S 6 1 で設定した範囲以外の範囲を 6 等分することで、色相を 7 等級に分割する (S 6 2)。

【 0 0 9 8 】

この S 6 2 における処理が終了したら、ヒストグラム算出部 1 4 は、S 3 3 (図 4 参照) において設定されたエッジ候補の全てに対して、図 9 に示すように、エッジ候補の両側の領域について、範囲を 7 等級に分割した色相のヒストグラムを算出する (S 6 3)。このようにして、図 4 における S 3 4 および S 3 5 の処理が終了する。

【 0 0 9 9 】

〔 4 . オブジェクト認識部における処理 〕

次に、オブジェクト認識部 4 における処理の詳細について説明する。オブジェクト認識部 4 は、視界画像からオブジェクト領域抽出部 3 により抽出された領域に映っている画像と、データベース 7 内に格納された登録画像とのマッチング処理を行い、ユーザが注目している画像を認識するものである。ここでは、ユーザが注目している画像が絵画であり、データベース 7 に格納された登録画像としての絵画画像と、ユーザが注目している画像とのマッチングが行われる場合を例に挙げて説明する。

【 0 1 0 0 】

まず、図 1 0 に示すように、オブジェクト認識部 4 は、視界画像からオブジェクト領域抽出部 3 により抽出された絵画領域を、横・縦方向に各々 8、6 等分し、4 8 ブロックに分割する (S 7 0)。さらに、オブジェクト認識部 4 における色相平均算出部 1 5 は、各ブロックの色相の平均値を算出する (S 7 1)。

【 0 1 0 1 】

S 7 1 における色相の平均値の算出処理は、図 1 1 に示すように、各ブロックにおける画素の色相の総和を画素数で除算した値を平均値とする処理を、全てのブロックに対して行うことで実行される。

【 0 1 0 2 】

そして、図 1 0 に示すように、色差算出部 1 6 は、データベース 7 中の全ての絵画画像に対して、各ブロックでの色差から、全ブロックでの色差 CS_{p_i} を、以下の式 (5) により算出する。

【 0 1 0 3 】

【 数 5 】

$$CS_{p_i} = \sum_{j=1}^8 \sum_{k=1}^6 \left| \overline{H}_q(j,k) - \overline{H}_{d_i}(j,k) \right| \quad \dots \text{式 (5)}$$

【 0 1 0 4 】

上記式 (5) において、

【 0 1 0 5 】

10

20

30

40

【数 6】

$$\overline{H}_q(j,k)$$

【0106】

は、視界画像から抽出した絵画領域におけるブロック(j,k) (j=1,...,8), (k=1,...,6)の色相の平均値を示している。

【0107】

また、

【0108】

【数 7】

$$\overline{H}_{di}(j,k)$$

【0109】

は、データベース7中のi番目の絵画画像におけるブロック(j,k)の色相の平均値を示している。

【0110】

なお、オブジェクトや領域に2次元バーコードのようなタグを付けることにより、オブジェクト認識や注目領域の判別を行ってもよい。

【0111】

〔5. 注視点分布に基づく注釈情報の提示処理について〕

本実施形態における情報処理装置1は、注目領域判別部5によりオブジェクトにおける注目領域を判別し、注釈情報提示部6により注目領域に応じた注釈情報および階層情報の提示を行う。また、データベース7中の登録画像のそれぞれには、色情報と、注釈情報が付加される領域の凸多角形領域の頂点座標、その注釈情報、その画像を構成する領域の階層情報が定義されている。

【0112】

図12(a)は、登録画像がたとえば絵画画像である場合における、注釈情報が付加されている領域を示す図であり、図12(b)は、その絵画画像の階層情報の構成を示す図である。

【0113】

図12(a)に示す絵画を大まかな2つの領域に分けると、領域1(屋敷)と、領域2(林)とに分けることができる。また、領域1(屋敷)を更に細分化すると、領域11(屋根)と領域12(ベランダ)とに分けることができる。よって、図12(b)に示すように、階層情報は、領域1の下位に、領域11および領域12が含まれるツリー構造となっている。

【0114】

以下、注目領域判別部5による注目領域の判別処理、および注釈情報提示部6による注釈情報の提示処理について、登録画像が絵画画像である場合を例に挙げて、より具体的に説明する。

【0115】

〔5-1. 注目領域判別部5における処理〕

注目領域判別部5は、オブジェクト領域抽出部3およびオブジェクト認識部4における処理の後に、注目区間の注視点分布を含む最小凸多角形領域と、データベース中の絵画に対して注釈情報が付加されている領域の位置情報とを基にして、注目領域を判別する。

【0116】

ここで、視界画像から抽出された絵画領域と、データベース7中の絵画領域とは大きさが異なる。そのため、注目領域判別部5は、データベース7中の絵画領域に対する、抽出された絵画領域の縮小・拡大率を基に、絵画領域の大きさに関して正規化を行う。その後

10

20

30

40

50

、注目領域判別部 5 は、最小凸多角形領域に含まれる注目領域を、図 1 3 に示すフローチャートに従って判別する。

【0117】

先ず、注目領域判別部 5 は、絵画領域の左下を基準として最小凸多角形領域の各頂点までの水平方向、垂直方向の距離を求め、絵画領域上での最小凸多角形領域の位置を算出する (S 8 0)。

【0118】

また、注目領域判別部 5 は、データベース 7 内の絵画領域についても、絵画領域の左下を基準として注釈情報が付加されている凸多角形領域の各頂点までの水平方向、垂直方向の距離を算出する (S 8 1)。

10

【0119】

なお、S 8 1 の処理は、データベース 7 内における絵画画像で、予め注釈情報が定義された領域について行われる。また、S 8 1 の処理は必ずしも S 8 0 の後に実行される必要はなく、予め行われるものであっても構わない。また、注目領域判別部 5 以外の処理部により、S 8 1 が行われても構わない。

【0120】

そして、注目領域判別部 5 は、データベース 7 中の絵画領域の大きさを基準として、視界画像から抽出した絵画領域の大きさを正規化する (S 8 2)。また、S 8 0 で距離を算出した絵画領域上での最小凸多角形領域の各頂点について、正規化後の位置を算出する (S 8 3)。

20

【0121】

そして、注目領域判別部 5 は、視界画像の絵画領域における最小凸多角形領域と、データベース 7 中の注釈情報を付加された各領域との重なりがどの程度の大きさであるか判断し、その重なりが最大となる領域を、ユーザが注目した領域として判別する (S 8 4)。

【0122】

〔5 - 2 . 注釈情報提示部 6 における処理〕

注釈情報提示部 6 は、注視点分布から注目領域判別部 5 により注目領域の判別が行われた後に、注目領域に関する注釈情報を提示する。たとえば、ユーザが絵画画像の屋敷部分に注目していると注目領域判別部 5 により判別されたら、注釈情報提示部 6 は、図 1 4 に示すように、屋敷に関する情報を情報処理装置 1 の外部に設けられた出力装置 (ノート PC、ヘッド・マウンテッド・ディスプレイ等) により表示する。なお、図 1 4 に示すように、注釈情報として絵画そのものに関する情報 (作者、作者名) を表示してもよい。

30

【0123】

また、注釈情報提示部 6 は、注目領域に関する注釈情報を提示するだけでなく、注目領域の階層情報を提示してもよい。つまり、階層情報にて示される階層構造において、注目領域の親となる領域や子となる領域の注釈情報を、注釈情報提示部 6 を用いてユーザに提示することにより、注目領域に関するその他の情報を得る手がかりをユーザに提供できる。この際に表示される注釈情報が付加されている領域と、注目領域との階層が近ければ近いほど、注目領域に関する情報を的確に補足する情報を得ることができる。

【0124】

また、注釈情報が表示デバイスの画面上に提示されている間は、情報処理装置 1 にて眼球運動の解析が行われないうちにもよい。そして、ユーザが眼を一定時間閉じることを検出して、眼球運動の解析を再開するトリガとしてもよい。これにより、ユーザは目を閉じるだけで再び注釈情報の提示を受けられるので、情報処理装置 1 の利便性が向上する。

40

【0125】

また、ユーザが絵画全体に注目した場合に、絵画そのものに関する注釈情報と、絵画に対して子となる領域とを、絵画に重ね合わせて表示してもよい。また、ユーザがある領域に注目している場合、注目領域に関する注釈情報と、注目領域の親または子となる領域を、ユーザに提示してもよい。

50

【0126】

〔6. 実験と評価〕

本実施形態の情報処理装置1により判別された注目領域が、実際に注目した領域であるかを評価するために実験を行ったので以下に説明する。なお、注目領域を判別する前処理として、視界画像から絵画領域を抽出し、抽出した絵画領域の認識を行っているため、絵画領域の抽出と、絵画の認識とについての評価実験も行った。

【0127】

なお、情報処理装置1として使用したPCは、周波数が500MHzのPentium(登録商標)IIIをCPUとして搭載したノート型PCである。また、カメラ17・18は、フレームサイズを160×120[pixel]、処理速度を10[fps]に設定した。なお、眼球撮影画像は256階調
10

【0128】

また、A3サイズの20種類の絵画を額縁に入れ、単一色の壁に掛けた状態で、被験者4人に対して実験を行った。20種類の絵画の内訳は、人物・動物画が10枚、風景画が10枚である。

【0129】

また、情報処理装置1のプロトタイプ使用時の様子を図15に、絵画抽出の様子を図16に示す。図16の上部5枚の画像のうち、下段右の画像は注視点を表すものである。

【0130】

〔6-1. 絵画領域の抽出精度と絵画の認識精度〕

20

〔6-1-1. 視界画像から絵画領域を抽出する精度〕

上述した方法でオブジェクト領域抽出部3により抽出された領域が、絵画領域であるかを評価するため、オブジェクト領域抽出部3により抽出された領域のうち実際に絵画領域である割合(Pr)と、視界画像に映っている絵画領域のうちオブジェクト領域抽出部3により抽出された領域の割合(Re)とに基づき評価を行った。

【0131】

そして、20種類の絵画について、ほぼ正面から見たときの絵画が映っている視界画像をシステムにより各20枚獲得し、合計400枚の視界画像から絵画領域の抽出を行った。その結果を表1に示す。表1には、各視界画像から絵画を抽出して得たPr、Reの平均値と、絵画の面積占有率の平均値とが示されている。なお、絵画の面積占有率とは、視界画像において絵画が占める面積の割合である。
30

【0132】

【表1】

	ユーザと絵画の距離			
	1.0m		1.5m	
	Pr	Re	Pr	Re
抽出精度	0.96	0.90	0.90	0.84
絵画の面積占有率	0.26		0.12	

【0133】

40

〔6-1-2. 絵画の認識精度〕

上述した手法によりオブジェクト領域抽出部3で視界画像に映っている絵画領域を抽出した後に、上述した方法によりオブジェクト認識部4で絵画を認識する。20種類の絵画について、ほぼ正面から見たときの絵画が映っている視界画像を情報処理装置1により各20枚獲得し、合計400枚の視界画像について絵画の認識を行った。そして、各視界画像において判別された絵画と、視界画像に映っている絵画と同じであれば認識の成功とした。その結果を表2に示す。

【0134】

【表 2】

	ユーザと絵画の距離	
	1.0m	1.5m
認識精度	0.91	0.86

【0135】

表 2 より、ユーザと絵画の距離が遠くなくても、ある程度の精度で絵画の認識を行うことができている。

【0136】

〔6-2. 注目領域の判別精度〕

上述した方法により注目領域判別部 5 が判別した注目領域が、実際に注目した領域であるかを評価した。なお、20種類の絵画に対して、構成要素である人や建物など5箇所を注目領域として定義しておく。

【0137】

これらの領域は、隣接している場合、離れている場合、包含している(されている)場合がある。各被験者に合計100箇所の注目領域を1回ずつ注目してもらい、注目領域判別部 5 が判別した注目領域と、実際に注目した領域が同じならば判別の成功とした。被験者4人の結果の平均を表 3 に示す。なお、ユーザと絵画との距離が1.0mの場合、絵画に対する視角は上下17.8°、左右22.6°となる。また、距離が1.5mの場合は、絵画に対する視角は上下11.4°、左右15.2°となる。

【0138】

【表 3】

	ユーザと絵画の距離	
	1.0m	1.5m
判別精度	0.88	0.81

【0139】

ユーザと絵画との距離が1.0mの場合と1.5mの場合の精度を比較すると、距離が近いほうが注目領域の判別精度はよくなっている。これは、ユーザと絵画の距離が近ければ、前処理の精度が良いことや、視界画像から抽出した絵画領域を正規化する際の、注視点の抽出上の誤差が及ぼす影響が小さいためだと考えられる。

【0140】

また、ユーザと絵画との距離が近い場合、抽出した絵画領域を正規化する際に領域の大きさの変化が小さかった。そのため、図 13 の S 8 3 で説明した正規化後の最小凸多角形の位置を計算しても、注視点の抽出上の誤差と、ユーザと絵画との距離の影響とにより、最小凸多角形が変形することはなかった。

【0141】

〔7. 補足〕

本実施形態の情報処理装置 1 の処理手順は、CPUなどの演算手段が、ROM (Read Only Memory) や RAM などの記憶手段に記憶されたプログラムを実行し、キーボードなどの入力手段、ディスプレイなどの出力手段、あるいは、インターフェース回路などの通信手段を制御することにより実現することができる。

【0142】

したがって、これらの手段を有するコンピュータが、上記プログラムを記録した記録媒体を読み取り、当該プログラムを実行するだけで、本実施形態の情報処理装置 1 の各種処理を実現することができる。また、上記プログラムをリムーバブルな記録媒体に記録することにより、任意のコンピュータ上で上記の各種機能および各種処理を実現することができる。

【0143】

この記録媒体としては、マイクロコンピュータで処理を行うために図示してはいないが

10

20

30

40

50

メモリ、例えばROMのようなものがプログラムメディアであっても良いし、また、図示していないが外部記憶装置としてプログラム読み取り装置が設けられ、そこに記録媒体を挿入することにより読み取り可能なプログラムメディアであっても良い。

【0144】

また、何れの場合でも、格納されているプログラムは、マイクロプロセッサがアクセスして実行される構成であることが好ましい。さらに、プログラムを読み出し、読み出されたプログラムは、マイクロコンピュータのプログラム記憶エリアにダウンロードされて、そのプログラムが実行される方式であることが好ましい。なお、このダウンロード用のプログラムは予め本体装置に格納されているものとする。

【0145】

また、上記プログラムメディアとしては、本体と分離可能に構成される記録媒体であり、磁気テープやカセットテープ等のテープ系、フレキシブルディスクやハードディスク等の磁気ディスクやCD/MO/MD/DVD等のディスクのディスク系、ICカード(メモリカードを含む)等のカード系、あるいはマスクROM、EPROM(Erasable Programmable Read Only Memory)、EEPROM(Electrically Erasable Programmable Read Only Memory)、フラッシュROM等による半導体メモリを含めた固定的にプログラムを担持する記録媒体等がある。

【0146】

また、インターネットを含む通信ネットワークを接続可能なシステム構成であれば、通信ネットワークからプログラムをダウンロードするように流動的にプログラムを担持する記録媒体であることが好ましい。

【0147】

さらに、このように通信ネットワークからプログラムをダウンロードする場合には、そのダウンロード用のプログラムは予め本体装置に格納しておくか、あるいは別な記録媒体からインストールされるものであることが好ましい。

【0148】

このように、本実施形態の情報処理装置1は、ユーザの眼球運動を撮影した眼球撮影画像に基づき、ユーザがオブジェクトに注目している状態を検出する眼球動作検出部2と、眼球動作検出部2によりユーザが注目状態にあると検出された際、ユーザの視界を撮影した視界画像内において、ユーザが注目しているオブジェクトを認識するオブジェクト認識部4と、上記オブジェクトにおいてユーザが注目している注目領域を判別する注目領域判別部5と、注目領域判別部5により判別された注目領域に関する注釈情報をユーザに提示する注釈情報提示部6とを備えているものである。

【0149】

上記構成によれば、眼球動作検出部2にてユーザが注目状態にあることを検出した上で、オブジェクト認識部4によりユーザが注目しているオブジェクトが認識されるので、ユーザが真に興味を示しているオブジェクトを認識することができる。

【0150】

その上、注目領域判別部5により、オブジェクトにおいてユーザが注目している注目領域を判別するので、オブジェクトにおいてどの部分にユーザが注目しているかを把握することができる。そして、注釈情報提示部6は、上記注目領域に関する情報をユーザに提示するので、ユーザが真に興味を示している領域に関する注釈情報を的確にユーザに提示することが可能であり、不必要な情報がユーザに提示されることを防止できる。

【0151】

このように、情報処理装置1によれば、ユーザが興味を示している領域に関する注釈情報が的確にユーザに提示されるので、情報処理装置1の快適な利用環境をユーザに与えることができる。

【0152】

さらに、眼球動作検出部2は、眼球撮影画像から瞳孔領域を抽出する視線検出部8と、上記瞳孔領域における中心座標から注視点の移動距離を算出する注視点算出部9と、上記

10

20

30

40

50

注視点の移動距離から、オブジェクトに対するユーザの視角を算出するとともに、その視角の大きさに基づき眼球の固視状態および跳躍状態を検出し、これら固視状態および跳躍状態の回数に基づきユーザがオブジェクトに注目している状態を検出する注目状態検出部10とを備えている。

【0153】

すなわち、固視状態および跳躍状態は、ユーザの視角を検出することにより検出可能であり、ユーザの視角は、注視点の移動距離を算出すれば求めることができる。

【0154】

上記構成の情報処理装置1は、視線検出部8により2値化処理を用いて瞳孔領域を抽出する一方で、注視点算出部9により注視点の移動距離を算出する。そして、注目状態検出部10は、注視点算出部9に求められた注視点の移動距離からユーザの視角を算出して固視状態および跳躍状態の回数を検出するので、的確にユーザの注目状態を検出することができる。

10

【0155】

このように、上記構成によれば、ユーザの注目状態を的確に把握することができるので、ユーザが興味を示しているオブジェクトがオブジェクト認識部4によりの確に認識され、注釈情報提示部6はさらに的確な注釈情報をユーザに提示することができる。

【0156】

さらに、注目状態検出部10は、ユーザが注目状態にある間の注視点の分布を内包するように複数の注視点を結ぶことにより、上記視界画像内に最小多角形領域を画定するものであるとともに、注目領域判別部5は、データベース7内に格納された登録画像における複数の領域のうち、上記最小多角形領域との重なりが最大となる領域を判断し、その領域に対応する上記視界画像内の領域を、上記注目領域として判別するものである。

20

【0157】

上記構成によれば、注目状態検出部10により最小多角形領域が画定される。この最小多角形領域は、注視点の分布を内包するように複数の注視点を結んで得られるものであるから、視界画像においてユーザが注目している領域を的確に表しているといえる。

【0158】

そして、注目領域判別部5は、上記最小多角形との重なりが最大となる領域を、データベース7内の登録画像について判断し、その領域に対応する視界画像内の領域を注目領域として判断するので、ユーザが注目している領域を的確に注目領域として判別することができる。

30

【0159】

さらに、オブジェクト認識部4は、上記視界画像における色相と、データベース7内に格納された視界画像の候補となり得る登録画像の色相とに基づき、ユーザが注目しているオブジェクトを認識するものである。

【0160】

視界画像がたとえば絵画画像のような色彩のある画像であれば、その色相により視界画像に含まれるオブジェクトを特徴付けることが可能である。そこで、上記構成では、視界画像の色相と、登録画像の色相とに基づき、ユーザが注目しているオブジェクトを認識するので、オブジェクトの認識精度を高めることができる。

40

【0161】

また、注釈情報提示部6は、上記注目領域と階層構造をなす領域に関する注釈情報をユーザに提示するものである。

【0162】

上記構成によれば、注目領域と階層構造をなす領域に関する注釈情報もユーザに提示されるので、ユーザは注目領域に関してより多くの情報を得ることができ、注目領域に関する理解を深めることができる。

【0163】

また、上記眼球撮影画像を撮影するカメラ17、および上記視界画像を撮影するカメラ

50

18の一方または双方を携帯端末に搭載することにより、ユーザが移動しても眼球撮影画像や視界画像を撮影することができるので、情報処理装置1の利便性をさらに高めることができる。

【0164】

さらに、携帯端末により注釈情報の再生を行うことで、移動するユーザに対しても注釈情報の提示を行うことができるので、情報処理装置1の利便性をより高めることができる。

【0165】

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

【産業上の利用可能性】

【0166】

本発明の情報処理装置によれば、たとえば美術館のように多くの絵画を展示している環境において、作品全体に関する情報や、作品の一部に関する情報を、注釈情報としてユーザの注目領域に適応させて提示することが可能となる。これにより、ユーザは展示されている絵画に関する理解を深めることができ、また、絵画の説明をする人員を配置するための人件費も省略することができる。

【0167】

また、本発明の情報処理装置の利用環境は、絵画を展示するような環境だけに限定されるものではない。たとえば、工作機械の各部分を撮影した画像を登録画像としてデータベースに格納しておくとともに、注釈情報として工作機械の操作方法を説明するテキストデータや音声データをユーザに提示するとよい。これにより、ユーザは、工作機械の各部分に注目するだけで工作機械の操作方法を知ることができるので、初心者であっても工作機械を操作することができる。

【図面の簡単な説明】

【0168】

【図1】本発明の一実施形態に係る情報処理装置の構成を示すブロック図である。

【図2】図1の情報処理装置により注釈情報が提示されるまでの一連の処理の流れを示すフローチャートである。

【図3】図1の情報処理装置における眼球動作検出部が行う処理の流れを示すフローチャートである。

【図4】図1の情報処理装置におけるオブジェクト領域抽出部が行う処理の流れを示すフローチャートである。

【図5】図5(a)は、視界画像の一例を示す図であり、図5(b)は、図5(a)の視界画像から求められたエッジ画像ならびに水平方向のエッジの長さを示すグラフである。

【図6】水平方向のエッジに係る長さの算出処理を詳細に示すフローチャートである。

【図7】垂直方向のエッジに係る長さの算出処理を詳細に示すフローチャートである。

【図8】色相の分割処理を詳細に示すフローチャートである。

【図9】色相のヒストグラム算出処理を詳細に示すフローチャートである。

【図10】オブジェクトを分割したブロックについて色差を算出する処理を詳細に示すフローチャートである。

【図11】色相の平均値の算出処理を詳細に示すフローチャートである。

【図12】図12(a)は、登録画像がたとえば絵画画像である場合における、注釈情報が付加されている領域を示す図であり、図12(b)は、その絵画画像の階層情報の構成を示す図である。

【図13】注目領域の判別処理の詳細を示す図である。

【図14】注釈情報を提示する画面の一例を示す図である。

【図15】本発明の情報処理装置の一実施形態を使用する状態を示す図である。

【図16】本発明の情報処理装置の一実施形態により絵画画像を抽出する様子を示す図で

10

20

30

40

50

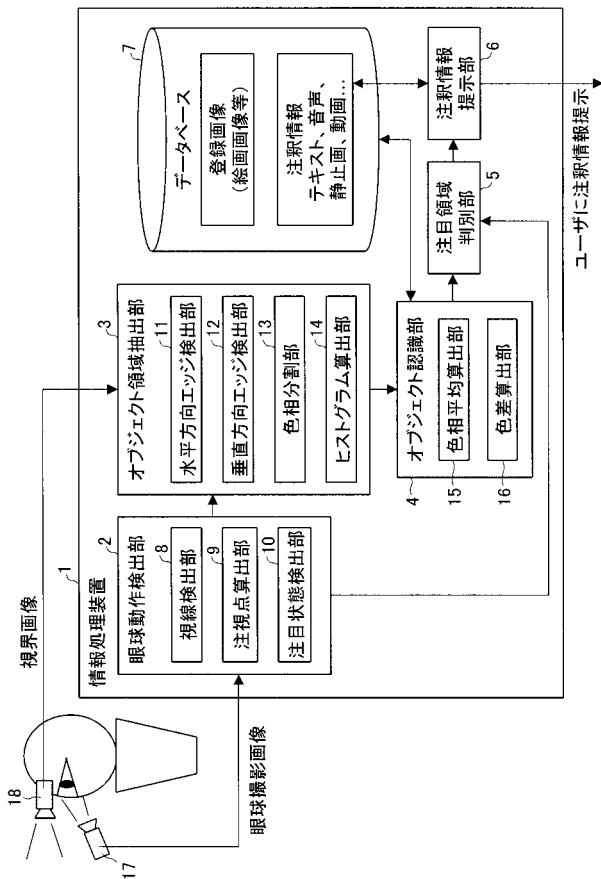
ある。

【符号の説明】

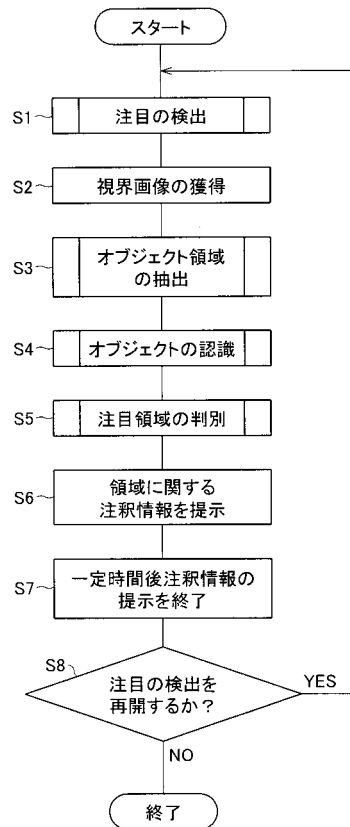
【0169】

- 1 情報処理装置
- 2 眼球動作検出部（眼球動作検出手段）
- 4 オブジェクト認識部（オブジェクト認識手段）
- 5 注目領域判別部（注目領域判別手段）
- 6 注釈情報提示部（注釈情報提示手段）
- 7 データベース
- 8 視線検出部（視線検出手段）
- 9 注視点算出部（注視点算出手段）
- 10 注目状態検出部（注目状態検出手段）
- 17 カメラ（第1カメラ）
- 18 カメラ（第2カメラ）

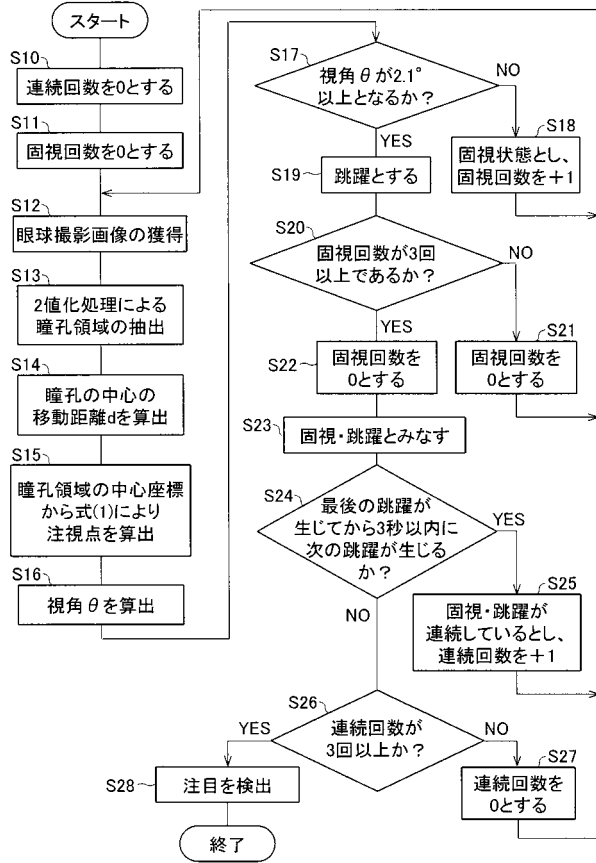
【図1】



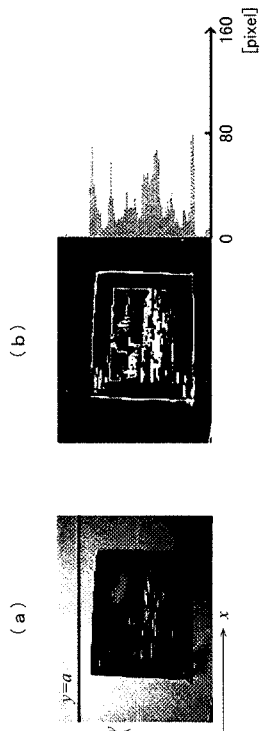
【図2】



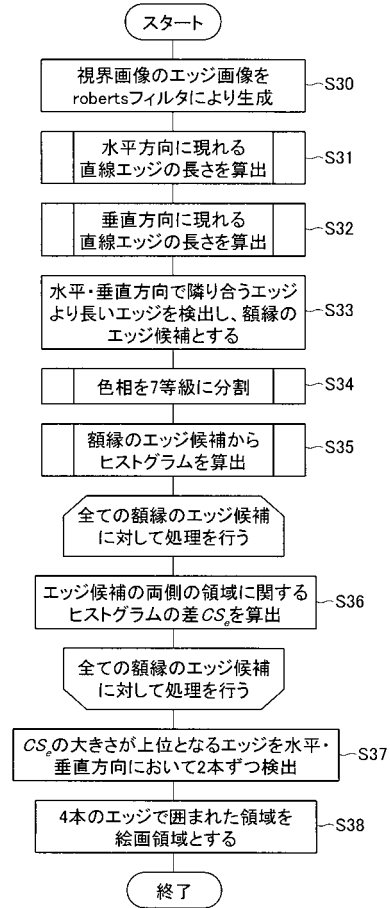
【 図 3 】



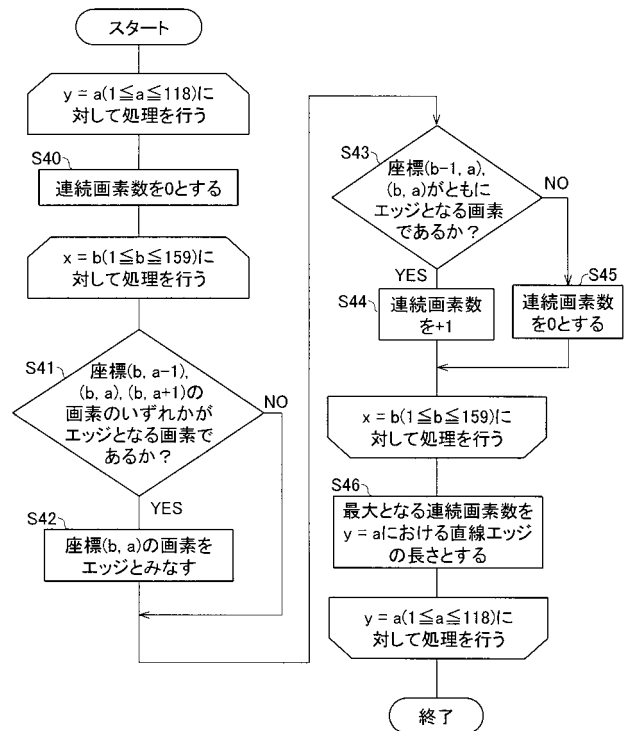
【 図 5 】



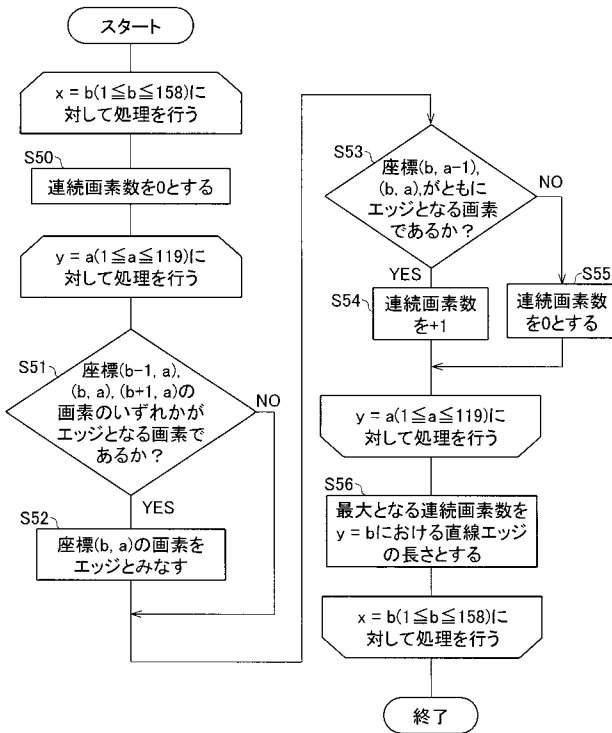
【 図 4 】



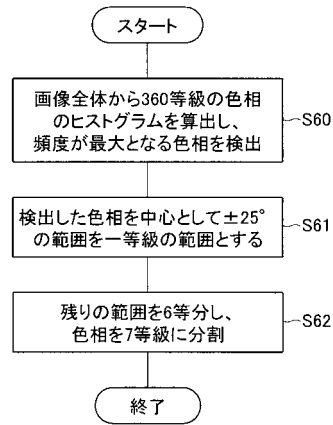
【 図 6 】



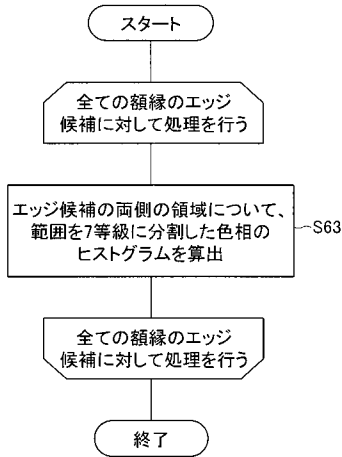
【 図 7 】



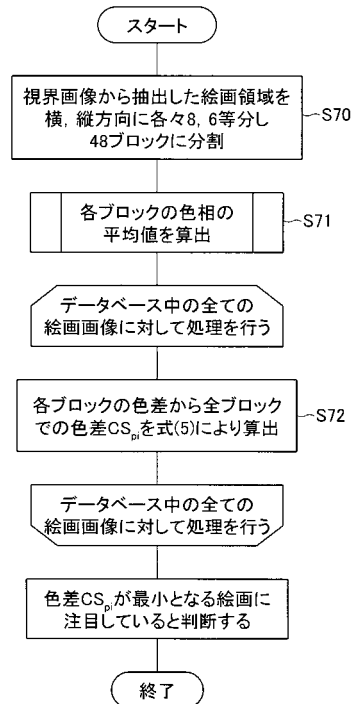
【 図 8 】



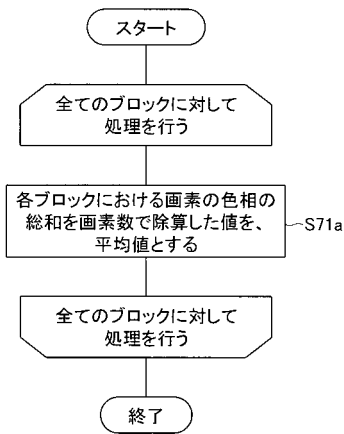
【 図 9 】



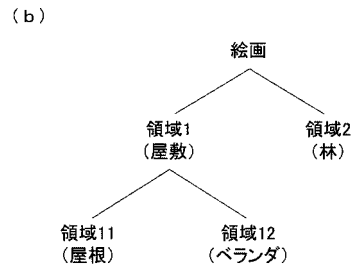
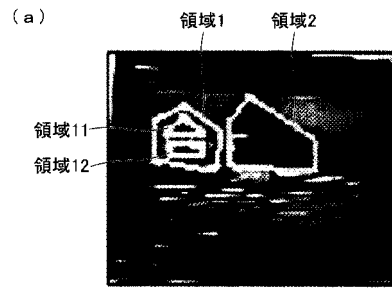
【 図 10 】



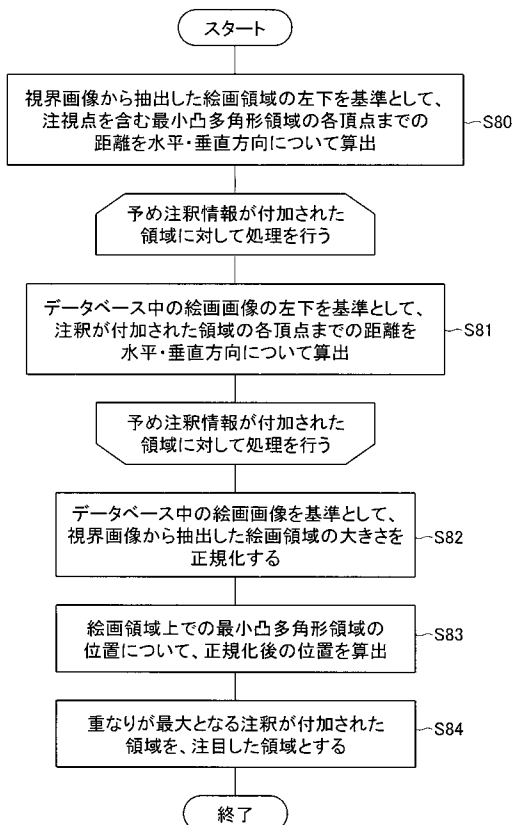
【 図 1 1 】



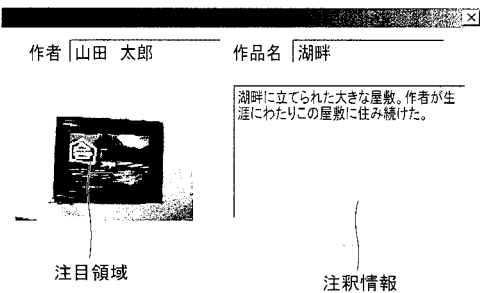
【 図 1 2 】



【 図 1 3 】



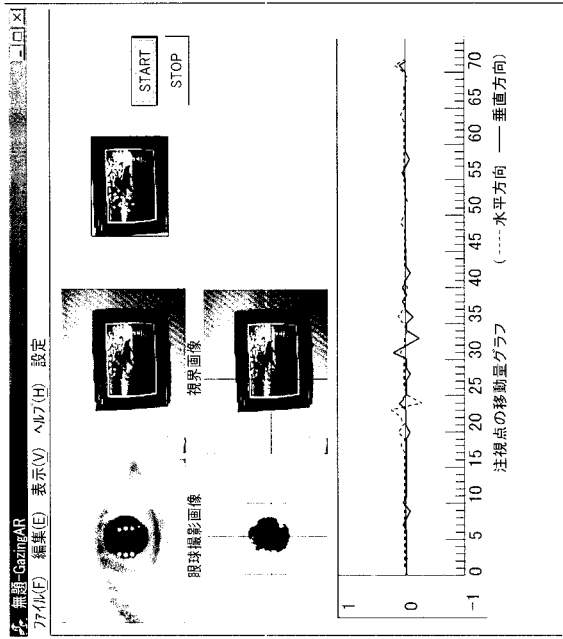
【 図 1 4 】



【 図 1 5 】



【図 16】



フロントページの続き

(72)発明者 吉高 淳夫

広島県東広島市鏡山1丁目4番1号 広島大学大学院工学研究科内

(72)発明者 竹村 知晃

広島県東広島市鏡山1丁目4番1号 広島大学大学院工学研究科内

Fターム(参考) 5B057 CA01 CA08 CA12 CA16 DA08 DA11 DA16 DC08 DC16 DC25
5E501 AA02 AA03 AA04 AA17 AA20 AA30 AB20 AC37 BA05 BA09
BA15 CA01 CB14 CC08 CC13 FA36 FA43 FB42 FB44
5L096 AA02 CA05 DA04 FA06 FA15 FA41 FA69