

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6126437号
(P6126437)

(45) 発行日 平成29年5月10日 (2017.5.10)

(24) 登録日 平成29年4月14日 (2017.4.14)

(51) Int. Cl.

F I

G 0 6 T 7 / 0 0 (2017.01)

G 0 6 T 7 / 0 0 C

G 0 6 T 7 / 6 0 (2017.01)

G 0 6 T 7 / 6 0 1 5 O P

請求項の数 16 (全 22 頁)

(21) 出願番号	特願2013-74860 (P2013-74860)	(73) 特許権者	000001007
(22) 出願日	平成25年3月29日 (2013.3.29)		キヤノン株式会社
(65) 公開番号	特開2014-199584 (P2014-199584A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成26年10月23日 (2014.10.23)	(74) 代理人	100076428
審査請求日	平成28年3月18日 (2016.3.18)		弁理士 大塚 康德
		(74) 代理人	100112508
			弁理士 高柳 司郎
		(74) 代理人	100115071
			弁理士 大塚 康弘
		(74) 代理人	100116894
			弁理士 木村 秀二
		(74) 代理人	100130409
			弁理士 下山 治
		(74) 代理人	100134175
			弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 画像処理装置および画像処理方法

(57) 【特許請求の範囲】

【請求項 1】

入力画像から物体を検出する画像認識処理において参照される辞書の生成に用いられる、該物体の学習画像を生成する画像処理装置であって、

前記物体の輝度画像と、前記物体の距離画像と、を入力する第1の入力手段と、

前記輝度画像に基づいて、前記距離画像に含まれている複数の第1の領域の法線方向を取得する第1の取得手段と、

前記複数の第1の領域の法線方向と、前記輝度画像において前記複数の第1の領域に対応する複数の第2の領域の輝度値と、の間の対応関係を表す観測データ分布を取得する第2の取得手段と、

前記観測データ分布に基づいて、前記物体の表面における輝度分布を推定する推定手段と、

前記輝度分布に基づいて、前記物体の学習画像を生成する生成手段と、

を有することを特徴とする画像処理装置。

【請求項 2】

前記推定手段は、前記観測データ分布と所定の輝度分布モデルとに基づいて、前記輝度分布を推定することを特徴とする請求項 1 に記載の画像処理装置。

【請求項 3】

前記推定手段は、

複数の輝度分布モデルを取得する手段と、

前記複数の輝度分布モデルに、前記観測データ分布を割り当てる割り当て手段と、
前記複数の輝度分布モデルを、該輝度分布モデルに割り当てられている観測データ分布
に基づいて更新する更新手段と

を有することを特徴とする請求項 2 に記載の画像処理装置。

【請求項 4】

前記割り当て手段は、前記観測データ分布が示す法線方向を入力して得られる輝度値が
該観測データ分布が示す輝度値に最も近い前記輝度分布モデルに対して、該観測データ分
布を割り当てることを特徴とする請求項 3 に記載の画像処理装置。

【請求項 5】

更に、

前記更新の前後における前記観測データ分布の割当先が一致するまで、前記割り当て手
段及び前記更新手段に処理を繰り返し行わせる判断手段を有することを特徴とする請求項
3 に記載の画像処理装置。

【請求項 6】

前記生成手段は、前記物体の複数の姿勢に対応する射影変換を、前記物体のモデル情報
に対して行い、該射影変換後の前記モデル情報によって示される前記物体の表面の法線方
向に、前記輝度分布に従う輝度値を与えて、前記姿勢ごとの前記学習画像を生成すること
を特徴とする請求項 1 に記載の画像処理装置。

【請求項 7】

前記モデル情報は CAD データであり、

前記学習画像が CG 画像である

ことを特徴とする請求項 6 に記載の画像処理装置。

【請求項 8】

前記第 1 の入力手段は、検出対象となる物体が山積みされた状態を撮影して得られた、
前記輝度画像及び前記距離画像を入力することを特徴とする請求項 1 乃至 7 のいずれか 1
項に記載の画像処理装置。

【請求項 9】

前記第 1 の入力手段は、複数の姿勢における前記物体を撮影して得られた、前記輝度画
像及び前記距離画像を入力することを特徴とする請求項 1 乃至 7 のいずれか 1 項に記載の
画像処理装置。

【請求項 10】

前記生成手段は、前記物体の表面の法線方向に応じた輝度値を前記輝度分布の分散から
決定することを特徴とする請求項 1 に記載の画像処理装置。

【請求項 11】

更に、

前記輝度画像と前記物体のモデル情報から前記距離画像を生成する距離画像生成手段を
有することを特徴とする請求項 1 乃至 10 のいずれか 1 項に記載の画像処理装置。

【請求項 12】

前記距離画像は、前記物体の表面の位置を表すことを特徴とする請求項 1 乃至 11 の何
れか 1 項に記載の画像処理装置。

【請求項 13】

更に、前記生成された学習画像に基づいて、前記物体を認識するための識別器を生成す
る手段を有することを特徴とする請求項 1 乃至 12 のいずれか 1 項に記載の画像処理装置
。

【請求項 14】

更に、前記物体を含む画像を取得する手段を有し、

前記識別器に基づいて、前記物体を含む画像から前記物体を認識することを特徴とする
請求項 13 に記載の画像処理装置。

【請求項 15】

入力画像から物体を検出する画像認識処理において参照される辞書の生成に用いられる

10

20

30

40

50

、該物体の学習画像を生成する画像処理装置が行う画像処理方法であって、
前記画像処理装置の第1の入力手段が、前記物体の輝度画像と、前記物体の距離画像と、
を入力する第1の入力工程と、
前記画像処理装置の第1の取得手段が、前記輝度画像に基づいて、前記距離画像に含ま
れている複数の第1の領域の法線方向を取得する第1の取得工程と、
前記画像処理装置の第2の取得手段が、前記複数の第1の領域の法線方向と、前記輝度
画像において前記複数の第1の領域に対応する複数の第2の領域の輝度値と、の間の対応
関係を表す観測データ分布を取得する第2の取得工程と、
前記画像処理装置の推定手段が、前記観測データ分布に基づいて、前記物体の表面にお
ける輝度分布を推定する推定工程と、
前記画像処理装置の生成手段が、前記輝度分布に基づいて、前記物体の学習画像を生成
する生成工程と
を有することを特徴とする画像処理方法。

10

【請求項16】

コンピュータ装置で実行されることにより、該コンピュータ装置を請求項1乃至14の
いずれか1項に記載の画像処理装置の各手段として機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、入力画像から対象物体の画像を検出する画像認識処理に用いられる辞書生成
用の学習画像を生成する画像処理装置および画像処理方法に関する。

20

【背景技術】

【0002】

従来より、対象物体を撮像して得られた画像から該対象物体の画像を検出する画像認識
に関し、さまざまな研究開発が行われている。画像認識技術はさまざまな分野に応用され
、例えば顔認識や、工場における部品認識など、多くの実問題に利用されてきている。

【0003】

このような画像認識をパターン認識の観点で考えることができる。このパターン認識に
おいても、入力された情報をどのようにしてクラス分類するか、という分類器の研究が行
われている。このような研究として例えば、ニューラルネットワークやSVM(Support Vect
or Machine)、Randomized Treeといった様々な手法が提案されている。

30

【0004】

これらの手法においては、予め画像認識用の辞書を生成しておく必要があるが、この辞
書生成の際には学習画像が必要となる。近年の工業ロボットにおける視覚認識としては、
例えば山積みされた複数種類の部品から所望の部品を検出する部品ピックアップ処理など、
3次元的に姿勢自由度の高い対象物体を認識するニーズもある。このように3次元的な姿勢
を検出するためには、対象物体のさまざまな姿勢に対応した学習画像が必要となる。

【0005】

このようにロボットによる部品ピックアップ等を目的とした認識タスクにおいては、対象
物体の姿勢情報が極めて重要となる。学習画像に対応する姿勢は、オイラー角や四元数等
のパラメータによって表現されるが、このような姿勢を既知とするような対象物体の学習
画像を、実写画像として用意することは困難である。そのため、CADデータによる任意姿
勢のCG画像を生成し、これを学習画像とすることが一般的である。

40

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2010-243478号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

50

上記従来の、CADデータから学習画像を生成する方法としては、CADデータのポリゴン表現による結合部分をエッジとして扱い、2値化されたエッジ画像を生成することが一般的である。そして実際の検出処理時には、部品群の実写入力画像に対してエッジ抽出処理を行い、エッジ画素をベースとしたマッチングを行うことによって、対象物体の位置および姿勢を同定する。このような方法においては、実写画像に対するエッジ抽出処理の結果が、検出性能を大きく左右する。一般にエッジ抽出処理は、対象物体の材質や環境光の影響等による変動が大きく、作業による合わせ込みの手間が非常に大きい。

【0008】

これに対し、レンダリングにより、学習画像を実写に近い画像として生成する方法も利用されている。この方法では、対象物体の各面がどのような輝度値となるのかを推定する必要がある。対象物体のBRDF(双方向反射率分布関数)と環境光の状況が既知であれば、それらを用いることで、対象物体の表面に、推定される輝度値を与えてCGを生成することは可能である。しかしながら、対象物体のBRDFを正確に知るためには特殊な機材による測定が必要であり、実環境における環境光の条件を正確に数値として獲得するための作業にも相当な手間がかかる。

【0009】

また、環境内に球体を置くことで輝度値と面方向に関する環境マップを作成し、そのマップに従って学習画像を生成する方法もある。例えば、鏡面物体の学習画像を生成するために、鏡面球体を環境内に置くことで環境マップ画像を生成する(特許文献1参照)。しかし、例えば対象物体が一般的なプラスチック等の材質による物体であった場合、たとえ素材が同じであっても成形金型や表面加工によってその反射特性は変動するため、対象物体と同じ反射特性である球状物体を用意することは非常に困難である。

【0010】

本発明は上記問題に鑑み、実環境下で対象物体を撮影した情報に基づき、環境条件を反映して対象物体の表面輝度を近似した学習画像を、容易に生成する画像処理装置および画像処理方法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】

【0011】

上記目的を達成するための一手段として、本発明の画像処理装置は以下の構成を備える。すなわち、入力画像から物体を検出する画像認識処理において参照される辞書の生成に用いられる、該物体の学習画像を生成する画像処理装置であって、

前記物体の輝度画像と、前記物体の距離画像と、をする第1の入力手段と、
前記輝度画像に基づいて、前記距離画像に含まれている複数の第1の領域の法線方向を
取得する第1の取得手段と、

前記複数の第1の領域の法線方向と、前記輝度画像において前記複数の第1の領域に対応する複数の第2の領域の輝度値と、の間の対応関係を表す観測データ分布を取得する第2の取得手段と、

前記観測データ分布に基づいて、前記物体の表面における輝度分布を推定する推定手段と、

前記輝度分布に基づいて、前記物体の学習画像を生成する生成手段と、
を有することを特徴とする。

【発明の効果】

【0012】

本発明によれば、対象物体を撮影した情報に基づき、環境条件を反映して対象物体の表面輝度を近似した学習画像を、容易に生成することができる。

【図面の簡単な説明】

【0013】

【図1】第1実施形態に係る画像認識処理を行う画像処理装置において、特に学習画像の生成を行うための概要構成を示すブロック図、

【図2】第1実施形態の画像処理装置において、ランタイム時に対象物体を検出するため

10

20

30

40

50

の概要構成を示すブロック図、

【図3】第1実施形態におけるランタイム処理および辞書生成処理を示すフローチャート

、
【図4】対象物体の表面性状を説明する図、

【図5】対象物体における輝度分布の観測例を示す図、

【図6】CG画像による学習画像生成の様子を示す図、

【図7】複数色からなる対象物体における輝度分布の観測例を示す図、

【図8】第2実施形態における輝度推定部の詳細構成を示すブロック図、

【図9】第2実施形態における輝度推定処理の詳細を示すフローチャート、

【図10】第2実施形態における輝度分布の推定方法を説明する図、

10

【図11】第2実施形態における色分け指定用のGUI例を示す図、

【図12】第2実施形態における色分けによる輝度分布関数の対応付けの例を示す図、

【図13】第3実施形態における輝度値の予測分布例を示す図、

【図14】第4実施形態において学習画像の生成を行うための概要構成を示すブロック図

、
【図15】第4実施形態における学習処理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】

【0014】

以下、本発明実施形態について、図面を参照して説明する。なお、以下の実施の形態は特許請求の範囲に関わる本発明を限定するものではなく、また、本実施の形態で説明され

20

ている特徴の組み合わせの全てが本発明の解決手段に必須のものとは限らない。

【0015】

<第1実施形態>

本発明は、入力画像から対象物体の画像を検出する画像認識処理において参照される辞書の生成に用いられる学習画像を、実環境下で対象物体を撮影した情報に基づき、環境条件を反映して対象物体の表面輝度を近似させるように生成する。

【0016】

概要構成

図1(a)に、本実施形態における画像認識処理を行う画像処理装置において、特に学習画像の生成を行うための構成の概要を示す。モデル設定部1010は、対象物体のモデルを設定し、モデル記憶部1020に記憶する。画像取得部1110は、対象物体を撮影することで事前取得画像を取得し、事前取得画像記憶部1120に記憶する。観測データ分布取得部1130は、事前取得画像記憶部1120に記憶された事前取得画像から得られる情報に基づき、輝度値の観測データ分布を取得する。輝度推定部1210では、観測データ分布取得部1130で得られたデータ分布に基づき、対象物体表面の輝度分布を推定する。画像生成部1220は、モデル記憶部1020に記憶されているモデルと、輝度推定部1210で推定された輝度分布に基づき、対象物体の様々な姿勢のCG画像を生成する。生成されたCG画像は、学習画像として学習画像記憶部2010に記憶される。本実施形態では、このように対象物体を撮影して得られた輝度分布に基づく学習画像を容易に生成することを特徴とする。

30

【0017】

以上のように生成された学習画像を用いた学習処理により、画像認識用の辞書を生成する。すなわち、学習画像設定部2020が、学習画像記憶部2010に記憶されている複数の学習画像を読み出し、学習部2100が、これらの学習画像を用いて学習処理を行い、対象物体を認識するための辞書を生成する。生成された辞書は辞書記憶部2200に記憶される。

40

【0018】

以上の構成により生成された辞書を用いて、ランタイム処理が行われる。ここでランタイムとは、本実施形態によって作成された学習画像に基づいて得られた辞書を用いて、実際の入力画像に対して対象物体の認識(検出)処理を行うことである。ランタイム時には、まず辞書設定部3010が、辞書記憶部2200に記憶された辞書を読み出し、該読み出した辞書を認識部3100に inputs する。一方、入力画像取得部3020は、対象物体を撮影した入力画像を

50

取得し、認識部3100に入力する。認識部3100は入力画像について、読み込んだ辞書に従って対象物体の位置および姿勢の推定を行う。認識部3100で推定された対象物体の位置および姿勢は、認識結果として、認識結果出力部3200にて所定の方法で提示される。

【 0 0 1 9 】

ロボット作業概要

以下、上述した構成からなる本実施形態の画像処理装置を、ロボットによる作業へ応用する場合を例として、説明を行う。まず、ランタイム時に対象物体を検出するための装置構成を図2に示す。図2において、トレイ500に検出の対象物体400が配置されている。300は画像情報および撮影位置からの距離情報を得るための撮像装置であり、図1における入力画像取得部3020に相当する。撮像装置300としては、撮影時に画像情報とともに距離情報を得ることができれば、ステレオカメラやTOFセンサ、あるいは、カメラと投光器の組み合わせによる光切断や空間コード化などの装置でもよく、本発明において限定されるものではない。また、後述するようにトラッキング技術を利用したモデルとの位置合わせを行う場合には、距離情報を得る必要はなく、撮像装置300はカメラのみによる構成であってもよい。撮像装置300は有線もしくは無線を介して計算機100に接続されている。

10

【 0 0 2 0 】

計算機100には、図1に示す認識部3100および認識結果出力部3200に相当する構成がプログラムもしくは回路として組み込まれている。計算機100の内部もしくは外部に備えられたハードディスク(不図示)が辞書記憶部2200に相当し、辞書を記憶している。なお、これら認識部3100、認識結果出力部3200、および辞書記憶部2200は、上記のような構成に限定

20

【 0 0 2 1 】

計算機100は、ロボットコントローラ210へと接続されている。ロボットコントローラ210はロボットアーム220に接続されており、ロボットアーム220はロボットコントローラ210からの命令信号を受けて動作する。ロボットアーム220には被作業物体への把持作業等、所定の作業を行うためのエンドエフェクタ230が設置されている。

【 0 0 2 2 】

ランタイム処理

30

図2に示す構成におけるランタイム処理について、図3(a)のフローチャートを用いて説明する。上述したようにランタイムとは、本実施形態で作成された学習画像に基づく辞書を用いて、実際の入力画像に対して対象物体の認識(検出)処理を行うことである。図2に示す構成例では、例えば実際の工場などで次々と搬送されてくる対象物体を撮影し、物体の位置および姿勢を認識することを指す。ただし、本発明は上記のようなランタイム処理に限定されるものではなく、例えば本発明を顔認識に応用した場合であれば、カメラを人に向けて撮影して顔認識を行う場面なども、ランタイム処理であると言える。

【 0 0 2 3 】

図3(a)において、まず辞書設定工程S3010で辞書設定部3010が、予め生成されて辞書記憶部2200に記憶されている辞書を読み出して、認識部3100に出力する。なお、辞書生成の詳細については後述する。

40

【 0 0 2 4 】

入力画像取得工程S3020では、トレイ500に配置された対象物体400がカメラ300によって撮影され、得られた画像(輝度画像)と距離情報は計算機100に送られる。次に認識工程S3100で認識部3100が、入力画像に対し、辞書設定部3010が読み込んだ辞書を用いて画像認識処理を行い、対象物体400の位置および姿勢を推定する。推定された位置および姿勢は、認識結果として認識結果出力部3200へと送られる。

【 0 0 2 5 】

ここで行われる画像認識処理とは、対象物体の位置および姿勢をクラス分類器によってクラス分けする処理であり、その際に用いられる辞書とは、このクラス分類器を定義する

50

ものである。辞書によって定義されるクラス分類器は、画像の一部に写っている対象物体がどのクラスに当てはまるかを判定することで、その位置および姿勢を認識する。例えば、周知の技術であるSVM(Support Vector Machine)やRT(Randomized Trees)によるクラス分類器を用いてもよい。また、クラス分類器への入力として、入力画像に所定の画像処理を行ったデータを入力してもよい。ここで入力画像に施す画像処理とは、入力画像をクラス分類器が扱い易い形式に変換する処理の総称であって、その処理内容は限定されない。例えば、ガウスフィルタやメジアンフィルタ等を用いたノイズ除去や、Sobelフィルタ、LoGフィルタ、ラプラシアンフィルタ、Cannyエッジ検出器によるエッジ抽出処理等が考えられる。また、拡大縮小、ガンマ補正等の前処理や、HOGやSIFTなどの特徴抽出処理も、ここで述べた画像処理として考えられる。また、画像処理は上記各処理のうち1つのみを選択的に行うものに限定されず、例えば、ガウスフィルタによるノイズ除去後にSobelフィルタによるエッジ抽出を行う、という一連の組み合わせ処理であってもよい。

10

【0026】

次に、認識結果出力工程S3200では、認識結果出力部3200によって、認識結果である対象物体400の推定位置および姿勢から、ロボットが所定作業を行うための命令をエンコードし、ロボットコントローラ210へ出力する。ロボットコントローラ210は送られてきた命令をデコードし、該命令に応じてロボットアーム220およびエンドエフェクタ230を動作させ、認識された被作業物体(対象物体400)に対して所定の作業を行う。

【0027】

なお、ランタイムにおいて上記S3100の認識作業を繰り返し行う場合、S3010で設定された辞書をメモリ(不図示)に保持しておくことで、辞書設定工程S3010を繰り返す必要はない。すなわちこの場合、入力画取得工程S3020以降の処理を繰り返し実行すればよい。

20

【0028】

辞書生成処理(学習処理)

以上のようなランタイム処理を行うにあたり、対象物体400を検出するための辞書を事前に用意しておく必要がある。以下、この辞書を設定するための処理について詳細に説明する。なお上述したように、繰り返し行われるランタイム作業において辞書は再利用されるものであるから、辞書生成処理は一度だけ行えばよい。

【0029】

本実施形態における辞書生成処理は、上述した図1に示す構成により行われる。なお、図1における画像生成部1220、学習部2100、認識部3100は、いずれも図2に示す計算機100内部におけるプログラムとして実装される。また、事前取得画像記憶部1120、学習画像記憶部2010、辞書記憶部2200、ハードディスクに実装されるものとして説明する。なお、このハードディスクは計算機100の内部もしくは外部に接続されているものである。ただし、本発明の実装形態はこの例に限定されるものではなく、ランタイム時に用いるものとは別の計算機やハードディスク、あるいはカメラ内に搭載された計算機やメモリに実装されていてもよい。また、画像取得部1110は、図2に示す撮像装置300もしくは計算機100の内部に、撮像装置300を制御するプログラムとして実装される。

30

【0030】

以下、図3(b)のフローチャートに従って、本実施形態における辞書生成処理について説明する。まず、モデル設定工程S1000でモデル設定部1010が、対象物体400のモデルを設定し、モデル記憶部1020に記憶させる。ここでモデルとは、後述するように対象物体のCG画像を生成する際に必要な情報を含むモデル情報であり、具体的には対象物体400のCADデータやポリゴンモデルなどを指す。

40

【0031】

次に画像取得工程S1100では、トレイ500に対象物体400を配置して、画像取得部1110が撮像装置300で撮影を行い、輝度画像と、該輝度画像における各画素位置の距離情報(距離画像)を取得する。取得された輝度画像と距離画像をまとめて事前取得画像として、事前取得画像記憶部1050に書き込まれる。ここで事前取得画像とは、画像生成部1220での学習画像の生成の際に利用される画像であり、ランタイム時と同じ環境条件、すなわち入力画

50

像取得工程S3020と同様の環境条件下で撮影されることが望ましい。例えば、事前取得画像の撮影時の照明条件は、S3020とほぼ同じ照明条件であることが望ましい。また、事前取得画像として撮影される対象物体は、複数個がランダムに山積みされている状態であることが望ましい。また、事前取得画像としては少なくとも1枚が撮影され、本実施形態では5枚程度を撮影するものとする。このように事前取得画像を複数枚撮影する場合には、それぞれの撮影状態において対象物体400の配置が異なっていること、すなわち位置および姿勢のバリエーションが多いことが望ましい。なお、事前取得画像の撮影には、ランタイム時と同じ撮像装置を用いることが理想的ではあるが、撮像装置とトレイの位置関係や照明条件が類似しているのであれば、別の撮像装置によって事前取得画像を取得してもよい。また、事前取得画像として、単数の対象物体を様々な姿勢で撮影してもよい。その場合には、山積み状態で撮影する場合よりも多めの画像(例えば20枚程度)を撮影することが望ましい。

10

【0032】

次に、観測データ分布取得工程S1130では、上記S1100で取得された事前取得画像に基づき、輝度値の分布を示す観測データ分布を取得する。事前取得画像記憶部1120に記憶されている事前取得画像には、輝度画像における任意の画素jに対し、距離画像としてカメラ座標系位置が (X_j, Y_j, Z_j) で観測されているとする。ここでカメラ座標系とは、輝度画像の各画素に対応する3次元座標系であり、カメラ300を原点としたXYZの各軸によって示される撮影空間を示す。このとき、画素jとその近傍数点(例えば画素jと隣接8画素の計9画素)のカメラ座標系位置を集計し、それらを平面近似することで、画素jに対応するカメラ座標系位置 (X_j, Y_j, Z_j) の表面法線ベクトル N_j を算出する。事前取得画像における対象物体400の存在領域(例えばトレイ500の内部領域)内のすべての画素について、対応する表面法線ベクトルを算出することにより、輝度値と表面法線方向の対応を示す観測データ分布を得ることができる。なお、ここでは観測輝度値を各画素値として説明しているが、一般には画像上の所定位置における輝度値であることは言うまでもない。したがって、観測輝度値としては画素単体における輝度値でなくともよく、近傍数画素による局所領域としての平均値や、ノイズ除去フィルタを通した後の値等であってもよい。

20

【0033】

次に輝度推定工程S1210では、事前取得画像から得られた観測データ分布に基づき、対象物体400の表面輝度分布を推定する。対象物体400のCADデータからCGによって学習画像を生成するには、対象物体400の表面輝度分布をモデル化し、該表面輝度分布モデルにおけるパラメータ(輝度分布パラメータ)を推定する必要がある。

30

【0034】

ここで、輝度分布パラメータの具体例としては以下のようなものが考えられる。例えば、光源を単一の平行光と仮定し、対象物体400としての部品表面がランバート反射(拡散反射)するとすれば、比較的簡単な輝度分布モデルで近似することができる。この近似例を図4を用いて説明する。図4は、対象物体400に対し、光源600から光を照射し、カメラ300がその反射光を受光する様子を示している。対象物体400の表面から光源600へ向かう光源軸20のカメラ座標系における光源方向 $L = (L_x, L_y, L_z)$ と、カメラ光軸10のカメラ光軸方向 $V = (V_x, V_y, V_z)$ の中間を、反射中心軸30とする。すると反射中心軸30の方向ベクトル $H = (H_x, H_y, H_z)$ は、以下の(1)式で表わされる。

40

【0035】

$$H = (L+V) / |L+V| \quad \cdots (1)$$

そして、対象物体400表面の任意の位置における表面法線40の法線ベクトル $N = (N_x, N_y, N_z)$ と、反射中心軸30の方向ベクトル H のなす角50をとすると、50は以下の(2)式で表わされる。

【0036】

$$50 = \cos^{-1}(H \cdot N / (|H| |N|)) \quad \cdots (2)$$

このとき、上記対象物体表面の任意の位置における輝度値Jは、50の関数として以下の(3)式のようにガウス関数で近似することができる。

50

【 0 0 3 7 】

$$J(\theta_j) = C \exp(-\theta_j^2/m) \quad \dots (3)$$

なお、(3)式におけるCとmはそれぞれ、輝度分布全体の強度と輝度分布の広がりを表す輝度分布パラメータであり、これを推定することで輝度分布モデルの近似が行われる。

【 0 0 3 8 】

ここで、光源は単一であると仮定されているため、得られた観測値の中で輝度値が最大となる画素の表面法線ベクトル N_j が、光源方向 $L=(L_x, L_y, L_z)$ であると推定される。このとき、観測誤差や輝度のサチュレーション等を考慮して、近傍画素による輝度値の平均化等を行っても良いことは言うまでもない。また、光源方向 L が既知である場合には、上記光源方向の推定を行う必要がないことはもちろんである。

10

【 0 0 3 9 】

輝度分布を上記(3)式のようなガウス関数で近似する場合、まず上記のようにして得られた光源方向 L に対して、上記(1)式に従って反射中心軸の方向ベクトル H を算出する。これにより、各画素 j における反射中心軸からの角度 θ_j が上記(2)式から求まる。以下、画素 j における角度 θ_j と輝度値 J_j の組を、観測点 $p_j=(\theta_j, J_j)$ と表わす。すべての画素 j についての観測点 p_j を算出することで、図5のような観測分布を得ることができる。図5において、データ点B100は角度 θ_j と輝度値 J_j の観測点 p_j である。

【 0 0 4 0 】

図5に示す観測分布に対し、上記(3)式のモデルをB200のように最尤推定フィッティングすることで、対象物体400の表面輝度分布の推定モデルを得ることができる。具体的にはまず、誤差関数 E を推定値と観測値の差の二乗和として、以下の(4)式のように定義する。なお、(4)式において θ_j は j についての総和を示す。

20

【 0 0 4 1 】

$$E = \sum_j \{J(\theta_j) - J_j\}^2 \quad \dots (4)$$

最尤推定フィッティングをこの誤差関数 E の最小化問題として考える。すると、誤差関数 E はパラメータ C に関して下に凸の2次関数であるため、以下の(5)式を解けば、パラメータ C の更新式が(6)式のように求まる。

【 0 0 4 2 】

$$E/C = 0 \quad \dots (5)$$

【 0 0 4 3 】

$$C = \frac{\sum_j J_j \exp\left(-\frac{\theta_j^2}{m}\right)}{\sum_j \exp\left(-\frac{2\theta_j^2}{m}\right)} \quad \dots (6)$$

30

【 0 0 4 4 】

またパラメータ m に関しては、計算を簡単にするため $m=1$ とし、 C の最適化問題として解く。ここで、誤差関数 E は C に関して凸の関数ではないため、以下の(7)式のように誤差関数 E をデータごとに分解し、それぞれに関して解く。

40

【 0 0 4 5 】

$$E_j = \{J(\theta_j) - J_j\}^2 \quad \dots (7)$$

(7)式を最急降下的に解くと、逐次更新式は以下の(8)式ようになり、これはRobbins-Monroの手続きと呼ばれる。なお、(8)式における係数 γ は正の値で定義される定数であり、観測データ数の逆数として与えるのが一般的である。

【 0 0 4 6 】

$$\begin{aligned} \gamma^{new} &= \gamma^{old} - \eta \frac{\partial E_j}{\partial \gamma} \\ &= \gamma^{old} - \eta \left[-2C\theta_j^2 \exp(-\gamma\theta_j^2) \{C\theta_j^2 \exp(-\gamma\theta_j^2) - J_j\} \right] \end{aligned} \quad \dots (8)$$

50

【 0 0 4 7 】

以上、対象物体表面が拡散反射する場合に、輝度分布パラメータ C_m を推定することによって、ガウス関数による輝度分布モデルでの近似を可能とする例を示した。一方、対象物体表面における鏡面反射成分を考慮する場合には、以下の(9)式に示すようなTorrance-Sparrowの輝度分布モデルを適用すればよい。

【 0 0 4 8 】

$$J(\theta, \phi, \psi) = K_d \cos \theta + K_s (1/\cos \theta) \exp(-\theta^2/m) \cdots (9)$$

なお、(9)式において K_d, K_s, m はこのモデルにおける輝度分布パラメータである。このモデルを図4に適用すると、 θ は(2)式の θ と同様であり、すなわち表面法線40の方向ベクトル N と反射中心軸30の方向ベクトル H のなす角50である。また、 ϕ は表面法線40の方向ベクトル N と光源軸20の方向ベクトル L のなす角70、 ψ は表面法線40の方向ベクトル N とカメラ光軸100の方向ベクトル V のなす角60であり、それぞれ以下の(10),(11)式で表わされる。

【 0 0 4 9 】

$$\theta = \cos^{-1}(L \cdot N / (|L| |N|)) \cdots (10)$$

$$\psi = \cos^{-1}(V \cdot N / (|V| |N|)) \cdots (11)$$

各観測画素 j に対応する(9)式における角度 θ_j と ψ_j は、上記(10),(11)式から得ることができ、これにより θ_j 、 ψ_j 、 ϕ_j に対応する輝度値 J_j の観測分布を得ることができる。この観測分布に対して(9)式のモデルを最尤推定によってフィッティングすることで、対象物体400の表面輝度分布の推定モデルを得ることができる。

【 0 0 5 0 】

光源が複数、もしくは環境光などによる外乱光がある場合には、表面法線方向ベクトル N を入力とし、輝度値 J を出力とするノンパラメトリックな回帰モデル $J(N)$ で輝度分布を近似してもよい。観測値における各画素 j に関する表面法線方向ベクトル N_j に対して、輝度値 J_j を教師値とし、所定のノンパラメトリックモデルを学習させることで、輝度分布推定関数を得る。ノンパラメトリックな回帰モデルとしては、SVMやSVR(Support Vector Regression)、ニューラルネットワーク等さまざまな方法を利用することができる。これらノンパラメトリックモデルを利用する場合には、フィッティングの前に予め光源方向を推定する必要はない。

【 0 0 5 1 】

また、回帰モデルの引数としてカメラ座標系位置 (X, Y, Z) を与えて $J(N, X, Y, Z)$ を近似することによって、位置による照明条件の違いを考慮した輝度分布推定関数を得ることもできる。また、輝度値が多チャンネルで得られている場合には、それぞれのチャンネルに関する輝度分布を別々に推定すればよい。多チャンネルである場合とは、例えばRGBによるカラー画像や、赤外光あるいは紫外光による非可視光画像を追加情報として含むような場合である。

【 0 0 5 2 】

以上のように輝度推定工程S1210で対象物体400の表面輝度が推定されると、次に画像生成工程S1220において、辞書を生成するために必要な複数の学習画像を設定する。学習画像は、S1000で設定された対象物体400のモデル(例えばCADデータ)に基づくCG画像として生成される。例えば、BRDFで表わされる対象物体400の表面性状と作業環境における光源情報が既知であれば、周知のレンダリング技術を用いて、モデルから様々な姿勢に関する対象物体400の見えをCG画像で再現することができる。ここで図6に、CG画像による学習画像の生成の様子を示す。図6に示すように、学習画像は対象物体400の物体中心404を中心とした測地ドーム401における各視点403と、該視点403における画像面内回転402によるバリエーションで生成され、それぞれに姿勢クラスのインデックスが与えられる。例えば、72視点、30度毎の面内回転によるバリエーションで学習画像を生成すると、辞書としては $72 \times (360/30) = 864$ クラスの分類器を学習させることになる。具体的には、モデル記憶部1020に記憶されている対象物体400のモデルに対して、上記各姿勢に対応する射影変換を行い、該射影変換後の各画素に対応するモデル上の点の法線方向(表面法線方向)を算出する。そして、輝度推定工程S1210で得られた結果に従って、その法線方向に対応する輝度値

10

20

30

40

50

を与えることで、各姿勢に応じた学習画像が生成される。

【 0 0 5 3 】

次に学習工程S2000において、画像生成工程S1220にて生成された複数姿勢の学習画像を用いて、認識部3100にて利用されるクラス分類器の形式に従って辞書を生成する。生成された辞書は辞書記憶部2200に記憶される。

【 0 0 5 4 】

以上説明したように本実施形態によれば、実環境において山積み状態、もしくは複数姿勢の単品状態での対象物体を撮影して得られた輝度情報及び距離情報に基づき、照明等の環境条件を反映して対象物体の表面輝度を近似した輝度画像を容易に生成する。この近似された輝度画像がすなわち、学習画像として辞書生成の際に利用される。

10

【 0 0 5 5 】

< 変形例 >

第1実施形態では、撮像装置300での撮影によって、対象物体を含む輝度画像および距離画像を取得する例を示したが、撮像装置300が測距機能を有さない場合にも、本実施形態は適用可能である。以下、撮像装置300にて距離画像を取得できない場合に、第1実施形態を適用する変形例について説明する。本変形例においては、周知のトラッキング技術を用いて対象物体の位置姿勢を推定することで、距離画像を生成する。図1(b)に、本変形例における学習画像生成用の構成を示すが、図1(a)と同様の構成には同一番号を付し、説明を省略する。すなわち、図1(b)における距離画像生成部1140にて、距離画像を生成する。この場合、撮像装置300で撮影した輝度画像に対し、ユーザが対象物体のモデルを画像上に重畳させて概位置および概姿勢を設定し、既存のトラッキング技術を用いて詳細な位置および姿勢に合わせ込む。合わせ込まれたモデルとの対応から、対象物体が写っている領域に関する各画素のカメラ座標系における推定位置を算出することで、推定された距離画像として扱うことができる。

20

【 0 0 5 6 】

< 第2実施形態 >

以下、本発明に係る第2実施形態について説明する。上述した第1実施形態では、対象物体が単色物体であることを前提として説明を行った。しかしながら、対象物体が複数の色から成る場合ももちろんある。例えば、複数部品によるアッセンブリとして供給される部品が対象物体である場合、該対象物体の一部が黒いプラスチック、一部が白いプラスチックで構成されるといったように、部位によって輝度特性が異なることが考えられる。第2実施形態ではこのように複数の輝度特性を有する対象物体について、学習画像を生成する例について説明する。

30

【 0 0 5 7 】

第2実施形態における画像認識処理を行うための基本構成は、第1実施形態で示した図1(a)と同様である。特に、モデル設定部1010、モデル記憶部1020、画像取得部1110、事前取得画像記憶部1120、観測データ分布取得部1130については、その動作も第1実施形態と同様であるため、説明を省略する。第2実施形態では、観測データ分布から対象物体の表面輝度を推定する輝度推定部1210における処理の詳細が、第1実施形態とは異なる。図8に、第2実施形態における輝度推定部1210の詳細構成を示し、以下に説明する。なお、画像生成部1220以降の処理、およびランタイム時の処理についても、第1実施形態と同様であるため説明を省略する。

40

【 0 0 5 8 】

第2実施形態の輝度推定部1210は、初期化部1211、データ割り当て部1212、近似部1213、収束判定部1214に細分化される。入力された観測データ分布に対し、まず初期化部1211にて、輝度分布を近似する複数の関数が初期化され、次にデータ割り当て部1212で、観測データ分布を該複数の関数のいずれかに割り当てる。そして近似部1213で、それぞれ割り当てられた観測データ分布に対して、輝度分布関数をフィッティングさせ、収束判定部1214で輝度分布推定計算が収束したか否かを判断する。

【 0 0 5 9 】

50

辞書生成処理(学習処理)

第2実施形態においても第1実施形態と同様に、後述するように生成された学習画像から、対象物体を検出するための辞書を生成する。第2実施形態における辞書生成処理の概要は第1実施形態で示した図3(b)のフローチャートと同様であるが、その詳細が異なる。なお、図3(b)においてモデル設定工程S1000、画像取得工程S1100、観測データ分布取得工程S1130については第1実施形態と同様であるため、説明を省略する。

【0060】

第2実施形態の輝度推定工程S1210においては、事前取得画像から得られた観測データ分布に基づき、対象物体の表面輝度を推定する。まず、事前取得画像記憶部1120に記憶されている画像から、輝度値と表面法線方向の対応に関する観測データ分布を得る。ここで、光源が単一の平行光であって、部品表面がランバート反射するものと仮定して、輝度分布モデルの近似例を図4を用いて説明する。

10

【0061】

図4において、対象物体400の表面から光源600へ向かう光源軸20のカメラ座標系における光源方向Lと、カメラ光軸10のカメラ光軸方向Vの間を、反射中心軸30とする。すると、反射中心軸30の方向ベクトルHは、第1実施形態と同様に(1)式で表わされる。また、表面法線40の法線ベクトルNと、反射中心軸30の方向ベクトルHのなす角 θ も、第1実施形態と同様に(2)式で表わされる。このとき、対象物体400の表面輝度特性が、T種類の特性の組み合わせによって表わされたとする。t番目($t=1, \dots, T$)の輝度分布関数 $J_t(\cdot)$ は、以下の(12)式のようにガウス関数で近似することができる。

20

【0062】

$$J_t(\cdot) = C_t \cdot \exp(-\frac{\cdot^2}{m_t}) \quad \dots (12)$$

なお、(12)式における C_t と m_t はそれぞれ、輝度分布全体の強度と輝度分布の広がりを表すパラメータである。

【0063】

対象物体400の輝度分布特性は、T個の輝度分布関数 $J_t(\cdot)$ ($t=1, \dots, T$)によって近似される。T=2のときの輝度分布関数の例を図7に示す。図7において、曲線B210, B220はそれぞれ、 $t=1, 2$ で推定された輝度分布関数による曲線である。なおTの値は、CADデータ等により材質の異なる部分として分解できる数が事前に分かっているならば、Tの値を該数(固定値)として事前に決定することができる。以下ではTが事前に設定されているものとして説明するが、Tが未知である場合の処理についても後述する。

30

【0064】

各輝度分布関数 $J_t(\cdot)$ のパラメータ C_t, m_t を推定するために、第2実施形態における輝度推定工程S1210は、図9(a)に示すように細分化される。また図10に、このパラメータ推定の概念図を示す。以下、これらの図を用いてパラメータ推定処理の手順を説明する。

【0065】

まず初期化工程S1211では、パラメータ C_t, m_t を初期化する。具体的には、初期値をランダムに選択しても良いし、T=1として最尤推定した後、該推定結果から僅かにずらした異なる値を、複数の C_t, m_t の初期値として設定する、等が考えられる。

【0066】

40

次にデータ割り当て工程S1212では、各観測点 $p_j = (\theta_j, J_j)$ を、最も近い各輝度分布関数 $J_t(\cdot)$ に対して割り当てる。詳細には例えば、複数の輝度分布関数 $J_t(\cdot)$ のうち、観測点の表面法線方向を入力としたときに得られる推定輝度値が観測点の輝度値と最も近い輝度分布関数に対して、該観測点を割り当てればよい。すなわち、輝度分布関数 $J_t(\cdot)$ を推定するためのデータ集合 S_t が、以下の(13)式のように定義される。

【0067】

$$p_j \in S_t \text{ if } \arg \min_c |J_c(\theta_j) - J_j| = t, \text{ for } \forall p_j \quad \dots (13)$$

【0068】

(13)式は、各観測点に対し、輝度分布関数のインデックスでラベリングすることに相当

50

する。この例を図10(a)に示す。同図において曲線B210-a,B220-aは、それぞれ異なるパラメータで初期化された輝度分布関数である。これに対し、各観測点が最も近い輝度分布関数に割り当てられる。すなわち、点B110-a等の白丸は曲線B210-aに割り当てられた観測点群であり、点B120-a等の黒丸は曲線B220-aに割り当てられた観測点群である。

【 0 0 6 9 】

次に、近似工程S1213では、各輝度分布関数に割り当てられた観測点群 S_t を用いて、それぞれの輝度分布関数 $J_t(\cdot)$ を最尤推定フィッティングにより更新する。この例を図10(b)に示す。同図における曲線B210-b,B220-bはそれぞれ、図10(a)に示す曲線B210-a,B220-aを自身に割り当てられた観測点群によって更新した曲線を表している。

【 0 0 7 0 】

そして全て(この場合2つ)の輝度分布関数 $J_t(\cdot)$ を更新した後、再び各観測点 p_j に対して最も近い輝度分布関数を特定する。全ての観測点 p_j に対して特定された輝度分布関数が、各観測点について既に割り当てられている輝度分布関数と同じであれば、輝度推定工程S1210は収束したものと見て、次の画像生成工程S1220へと進む。一方、特定された輝度分布関数が既に割り当てられている輝度分布関数と異なる観測点があればデータ割り当て工程S1212へと戻り、上記処理を繰り返す。

【 0 0 7 1 】

ここで、対象物体の輝度分布に鏡面反射モデルを取り入れる場合には、第1実施形態と同様にTorrance-Sparrowの輝度分布モデルを適用する。その場合、 t 番目の輝度分布関数 $J_t(\cdot, \cdot, \cdot)$ を以下の(14)式のように近似すればよい。

【 0 0 7 2 】

$$J_t(\cdot, \cdot, \cdot) = K_{dt} \cos \theta + K_{st} (1/\cos \theta) \exp(-\theta^2/mt) \cdots (14)$$

なお、(14)式において K_{dt} , K_{st} , mt はこのモデルにおけるパラメータであり、 θ は上記(10),(11)式で表わされる。これらのパラメータも第1実施形態と同様に関数フィッティングによって推定される。なお、輝度値が多チャンネルで得られている場合にも、第1実施形態と同様にそれぞれのチャンネルに関する輝度分布を別々に推定すればよい。

【 0 0 7 3 】

以上、対象物体の輝度分布特性を構成する輝度分布関数の数 T が既知である場合(上記例では $T=2$)について説明したが、以下、 T が未知である場合の推定例について説明する。 T が未知である場合、まず複数の T を設定して推定を行い、それらの中で、分布が最もよく分離しているものを選べばよい。この場合の輝度推定工程S1210における処理を、図9(b)を用いて説明する。図9(b)において、初期化工程S1211、データ割り当て工程S1212、近似工程S1213における処理内容は、図9(a)の場合と同様である。すなわち図9(b)では図9(a)と同様のS1211～S1213の処理を、複数種類の T のバリエーション $T=1, \dots, T_{\max}$ に対して行った後、分離度評価工程S1214を行うことを特徴とする。なお、 T_{\max} は対象物体を構成する色の上限数であり、例えば $T_{\max}=5$ 等を設定する。

【 0 0 7 4 】

分離度評価工程S1214では、各 T に関する分離評価値 λ_T を、以下の(15),(16),(17)式のように定義する。

【 0 0 7 5 】

$$\lambda_T = \frac{1}{T} \sum_{t=1}^T \|\zeta_t\| \cdots (15)$$

【 0 0 7 6 】

$$\zeta_t = \frac{1}{\|S_t\|} \sum_{j \in S_t} \frac{(J_j - J_t(\theta_j))^2}{\varepsilon_t^3} \cdots (16)$$

【 0 0 7 7 】

10

20

30

40

$$\varepsilon_t^2 = \frac{1}{\|S_t\|} \sum_{j \in S_t} (J_j - J_t(\theta_j))^2 \quad \cdots (17)$$

【 0 0 7 8 】

なお、(17)式における ε_t は推定値に対する二乗誤差であり、(16)式における ε_t は推定値を中心とした歪度を表す。各輝度分布関数 J_t に割り当てられているデータ集合 S_t が J_t に対して正規分布に近い形になるほど、歪度 ε_t の値は0に近づく。したがってこの場合、分離評価値 ε_t の値が最も小さくなる T の値を、輝度分布関数の数 T の推定値とする。

【 0 0 7 9 】

以上のように、対象物体の輝度分布関数の数 T が既知または未知のいずれの場合であっても、輝度推定工程 S1210 で輝度分布関数が推定されると、次に画像生成工程 S1220 で、該推定された輝度分布関数に基づく学習画像が生成される。以下、第2実施形態における学習画像の生成処理について説明する。学習画像を生成するにあたり、まず、対象物体の各部分に輝度分布関数を対応付ける必要があるが、この対応付けは以下のような方法で行うことができる。

【 0 0 8 0 】

例えば、輝度分布の拡散反射成分の大小比較によって、自動的に対応付けを行うことができる。上記(12)式に示す近似式であれば、 θ の値が大きい部分(例えば $\theta = 1 \text{ rad}$)における輝度値の大小によって、明るい材質の部分と暗い材質の部分のそれぞれに対応する輝度分布関数を判断し、対応付けることができる。また、上記(14)式に示す近似式であれば、パラメータ K_{d_i} が拡散反射成分の強さを表しているため、その大小によって輝度分布関数を対応付けてもよい。

【 0 0 8 1 】

また、対象物体のカラー画像等に対し多チャンネルの輝度分布関数が推定されている場合には、特徴のあるチャンネルの拡散反射成分を比較してもよい。例えば、赤い部位と緑色の部位を対応付ける場合、RチャンネルとGチャンネルにおける拡散反射成分の強さを比較することで、容易に対応づけることができる。

【 0 0 8 2 】

また、事前取得画像において複数点をユーザに指定させ、該指定点に対応する画素が寄与する輝度分布関数を検出することで、対応付けを行ってもよい。例えば図11に示すように、対象物体の山積み状態を示す事前取得画像 T500 を GUI に表示し、明輝度部位の一部である部位 T100 と、暗輝度部位の一部である部位 T200 を、カーソル T300 で指定させる。この指定により、図12に示すように輝度分布関数の対応付けがなされる。すなわち、指定した部位 T100 における観測データが、図12における C100 であり、同じく部位 T200 における観測データが C200 であったとする。図12において、観測データ C100 と C200 のそれぞれが曲線 B210, B220 で示される輝度分布関数に属することが、上記(13)式によって判定できる。

【 0 0 8 3 】

画像生成工程 S1220 では、以上のように対象物体の各部分と輝度分布関数の対応づけがなされれば、以降は第1実施形態と同様に学習画像を生成することができる。なお、その後の学習工程 S2000 で学習画像を用いて辞書を生成する処理については、第1実施形態と同様であるため説明を省略する。

【 0 0 8 4 】

以上説明したように第2実施形態によれば、対象物体表面が複数色によって構成される場合であっても、その表面輝度を近似した学習画像を生成することができる。

【 0 0 8 5 】

< 第3実施形態 >

以下、本発明に係る第3実施形態について説明する。上述した第1および第2実施形態では、対象物体の表面性状が表面法線方向に対して安定しているものとして説明を行ったが、この表面性状は必ずしも安定しているとは限らない。例えば、対象物体表面に梨地のような加工がなされていれば、同一方向を向いた対象物体の面であっても、部分によって輝

10

20

30

40

50

度値が変化する。また意図的な表面加工がなくても、成型時の型の粗さ等によって同様に輝度値の変化が発生する場合がある。このような表面正常が不安定な対象物体に対し、その輝度値の不安定さを再現するために学習画像生成時にノイズを加えることが考えられる。第3実施形態ではすなわち、学習画像生成時に加えるノイズを考慮して、対象物体の輝度分布を推定することを特徴とする。

【0086】

辞書生成処理(学習処理)

第3実施形態においても第1実施形態と同様に、後述するように生成された学習画像から、対象物体を検出するための辞書を生成するが、この辞書生成処理も図3(b)のフローチャートに従う。図3(b)においてモデル設定工程S1000、画像取得工程S1100、観測データ分布取得工程S1130については第1実施形態と同様であるため、説明を省略する。

10

【0087】

第3実施形態の輝度推定工程S1210においては、事前取得画像から得られた観測データ分布に基づき、対象物体の表面輝度を以下のように推定する。

【0088】

輝度分布を以下の(18)(19)(20)式に示すような線形ガウスカーネルモデル $y(\theta, w)$ で表わすことを考える。

【0089】

$$y(\theta, w) = w^T \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \cdots (18)$$

$$w = (w_1, \dots, w_h, \dots, w_M)^T \cdots (19)$$

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta_1), \dots, \cos(\theta_h), \dots, \cos(\theta_M) \end{pmatrix}^T \cdots (20)$$

20

ここで、 θ は図4で説明した表面法線40の方向ベクトル N と反射中心軸30の方向ベクトル H のなす角50である。また M はカーネルの数であって、ユーザによって定義される。角度は $0 \sim 90\text{deg}$ の範囲で収まるため、例えば $M=9$ と設定すれば、 10deg 程度の間隔でカーネルを配置することができる。また、パラメータ w は M 次元ベクトルであり、その要素 w_h ($h=1, \dots, M$)は正の実数値である。ベクトル $\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$ は M 次元ベクトルであり、その要素 $\cos(\theta_h)$ ($h=1, \dots, M$)は(21)式のように定義されるガウスカーネルである。

【0090】

$$\cos(\theta_h) = \exp\{-(\theta_h - \mu_h)^2 / 2S^2\} \cdots (21)$$

ここで、 μ_h はガウスカーネル $\cos(\theta_h)$ の中心位置を表す。カーネル中心 μ_h は角度 θ の定義域内に配置すればよく、例えば $M=9$ と定義した場合には、 9deg ごとに μ_h を設定してもよい。

30

【0091】

このような線形ガウスカーネルモデルで輝度分布を近似した場合、以下の(22)式のように定義される予測輝度分布を考える。

【0092】

$$p(J|R, \theta) = \int p(J|w, \theta) p(w|R, \theta) dw \cdots (22)$$

ここで、 R は観測輝度値の集合ベクトルであり、観測画素の総数が N である場合には以下の(23)式に示すような列ベクトルで表わされる。

【0093】

$$R = (J_1, \dots, J_j, \dots, J_N)^T \cdots (23)$$

ここで、 J_j は観測データの画素 j における輝度値の観測値である。

40

【0094】

上記(22)式の右辺第一項は輝度値の条件付き分布であって、以下の(24)式に示すような正規分布で表現される。

【0095】

$$p(J|w, \theta) = N(J|y(\theta, w), \sigma^2) \cdots (24)$$

ここで、 σ^2 は精度パラメータであり、以下の(25)式に示すような、第1および第2実施形態における推定輝度関数 $J(\theta_j)$ と観測輝度値 J_j との二乗誤差平均を用いる。なお、(25)式において θ_j は j についての総和を示す。

50

【 0 0 9 6 】

$$\sigma^2 = (1/N) \{J_j - J(\theta_j)\}^2 \quad \dots (25)$$

上記(24)式を重み w の尤度関数と考え、この共役事前分布を、以下の(26)式に示すような期待値 m_0 、共分散 S_0 を持つガウス分布であるとする。

【 0 0 9 7 】

$$p(w) = N(w|m_0, S_0) \quad \dots (26)$$

このとき、事後分布である(22)式の右辺第二項は、以下の(27)(28)(29)式に示すような正規分布で表わすことができる。

【 0 0 9 8 】

$$p(w|R, \theta) = N(w|m_N, S_N) \quad \dots (27)$$

$$m_N = S_N(S_0^{-1}m_0 + 1/\sigma^2 \sum_{j=1}^N R_j^T R_j) \quad \dots (28)$$

$$S_N^{-1} = S_0^{-1} + 1/\sigma^2 \sum_{j=1}^N R_j^T R_j \quad \dots (29)$$

ここで、 θ は計画行列と呼ばれ、カーネルと観測データから以下の(30)式のように決定される。

【 0 0 9 9 】

$$\Phi = \begin{bmatrix} \phi_1(\theta_1) & \dots & \phi_h(\theta_1) & \dots & \phi_M(\theta_1) \\ \vdots & & & & \vdots \\ \phi_1(\theta_j) & & \phi_h(\theta_j) & & \phi_M(\theta_j) \\ \vdots & & & & \vdots \\ \phi_1(\theta_N) & \dots & \phi_h(\theta_N) & \dots & \phi_M(\theta_N) \end{bmatrix} \quad \dots (30)$$

【 0 1 0 0 】

このとき、最小二乗法に沿って上記(18)式の線形ガウスカーネルモデルを近似すると、最終的に(22)式の予測輝度分布は以下の(31)(32)式のように表わされることが知られている。

【 0 1 0 1 】

$$p(J|R, \theta) = N(J|m_N^T \phi(\theta), \sigma_N^2(\theta)) \quad \dots (31)$$

$$\sigma_N^2(\theta) = \sigma^2 + \phi(\theta)^T S_N^{-1} \phi(\theta) \quad \dots (32)$$

なお、(32)式は予測輝度分布の分散であって、その平方根である $\sigma_N(\theta)$ は予測輝度分布の標準偏差である。

【 0 1 0 2 】

以上のように、輝度推定工程S1210で対象物体の予測輝度分布が推定されると、次に画像生成工程S1220で、該推定された予測輝度分布に基づく学習画像が生成される。第3実施形態における学習画像の生成は第1実施形態と同様に、各姿勢におけるCADモデルを射影変換したときの各画素に対応するモデル上の法線方向に対して輝度値を算出し、画素に輝度値を与えることで、学習画像を生成する。具体的には、生成する学習画像における画素 k に射影される面の法線方向と反射中心軸とのなす角 θ_k に対して、上記(31)(32)式より、予測分布 $p(J|R, \theta_k)$ を得ることができる。

【 0 1 0 3 】

ここで図13に、第3実施形態において得られる輝度値の予測分布例を示す。図13において、実線B300は変数 θ における予測分布の中心であり、破線B310および破線B320は、上記(32)式に σ_N を引数として与えて得られる標準偏差 $\sigma_N(\theta)$ で表わされる予測分布の幅を示す。また、曲線B330は θ_k における輝度方向の予測分布を示し、その幅B340は標準偏差 $\sigma_N(\theta_k)$ である。第3実施形態では、この予測輝度分布に従った乱数を発生させ、得られた値を画素 k に与えてやることで、学習画像を生成することができる。

【 0 1 0 4 】

以上説明したように第3実施形態によれば、対象物体表面における画素ごとの輝度値を、該対象物体について推定された輝度分布の分散から決定することで、対象物体の表面輝度のばらつきまで考慮した学習画像を生成することができる。

【 0 1 0 5 】

< 第4実施形態 >

以下、本発明に係る第4実施形態について説明する。上述した第1実施形態では、対象物体表面の輝度を上記(3)式や(9)式による輝度分布モデルで近似する例を示した。第4実施形態ではさらに、輝度分布モデルについてのパラメータ(輝度分布パラメータ)として複数のパラメータ候補を用意する。そして、各パラメータ候補に基づいて対象物体の認識用の辞書を作成し、実写による入力画像に対して該辞書を適用した認識結果を評価値として、最適なパラメータ候補を選択する例を示す。ここで輝度分布パラメータとは、(3)式における C と m 、あるいは(9)式における K_d, K_s, m である。

【 0 1 0 6 】

第4実施形態における画像認識処理を行うための基本構成を図14に示す。同図において、第1実施形態で示した図1(a)と同様の構成には同一番号を付し、説明を省略する。すなわち第4実施形態では、図1(a)に示す構成から輝度推定用の構成を除き、パラメータ設定部1230と選択部2110を加えたことを特徴とする。

【 0 1 0 7 】

辞書生成処理(学習処理)

第4実施形態における辞書生成処理は、図15(a)のフローチャートに従う。なお、図15(a)においてモデル設定工程S1000、画像取得工程S1100では第1実施形態と同様に、対象物体のCADモデルと複数の事前取得画像が設定される。

【 0 1 0 8 】

画像生成パラメータ候補設定工程S1215では、学習画像を生成するための画像生成パラメータの候補を K パターン用意する。ここで画像生成パラメータとはすなわち、上述した第1実施形態で推定される輝度分布パラメータである。

【 0 1 0 9 】

すると画像生成工程S1220では、用意された K パターンの画像生成パラメータ候補のそれぞれに対応した学習画像を、第1実施形態と同様の方法によって生成する。以下、全 K パターンのうち、 k 番目の画像生成パラメータ候補によって生成された様々な姿勢の学習画像の集合を、学習画像集合 S_k とする。そして学習工程S2000では、 K 個の学習画像集合 $S_k(k=1, \dots, K)$ のそれぞれを用いて、 K 個の辞書を生成する。

【 0 1 1 0 】

そして選択工程S2110において、生成された K 個の辞書を用いて、S1100で設定された全ての事前取得画像に対する評価を行い、該評価結果に基づいて最適な辞書および画像生成パラメータ候補を選択する。

【 0 1 1 1 】

ここで、選択工程S2110の詳細を図15(b)のフローチャートに示し、説明する。

【 0 1 1 2 】

まず認識工程S2111において、第1実施形態で説明したランタイム時の認識処理(S3100)と同様に、事前取得画像に対して辞書を用いた認識処理を行い、事前取得画像中の対象物体の位置および姿勢の推定値を算出する。

【 0 1 1 3 】

次に評価工程S2112では、認識工程S2111にて得られた認識結果を以下のように評価する。まず、認識結果として得られた推定位置および姿勢に基づき、S1000で設定されたモデルから対象物体のCG画像を生成する。このとき、推定位置および姿勢に対して直接CG画像を生成してもよいが、周知のトラッキング技術を用いてより詳細にマッチングした結果に対してCG画像を生成してもよい。すなわち、認識結果として得られた対象物体の推定位置および姿勢を初期値として、既知のトラッキング技術を用いて事前取得画像に対してより詳細にマッチングさせた後の推定位置および姿勢を用いて、CG画像を生成することも可能である。そして、生成されたCG画像と事前取得画像に対してエッジ抽出処理をかけて2値化し、双方で得られたエッジ点位置を比較して距離を算出し、その和、もしくは2乗和による検出誤差を、評価値として算出する。すなわち評価工程S2112では、認識結果および

10

20

30

40

50

モデル情報から生成されるモデル画像(CG画像)と、事前取得画像における対象物体の画像との対応部位における差分に基づいて、該認識結果の評価値を算出する。

【0114】

あるいは、距離画像を利用して、距離残差による評価を行ってもよい。すなわち、認識結果として得られた推定位置および姿勢に基づき、モデルからその位置および姿勢における部品表面の距離を算出する。そして、事前取得画像に対応する距離情報と比較し、その物体表面における距離の残差の和もしくは2乗和による誤差を、評価値として算出する。

【0115】

あるいは、学習画像と入力画像の類似度を評価してもよい。その場合、認識結果における推定位置および姿勢に基づいて生成されたCG画像と、事前取得画像における対象物体の存在領域の類似度を正規化相関等で比較し、その類似度を評価値として算出する。

10

【0116】

さらに、生成されたCG画像と、入力された事前取得画像をユーザが目視で確認し、その差異を評価してもよい。例えば、位置ずれや姿勢ずれの誤差を数段階(例えば5段階程度)に定義し、主観的にユーザが入力して評価値とする。また、上記様々な評価値の組み合わせ、例えば線形和等によって一つの評価値としてもよい。

【0117】

評価工程S2112では、以上のように各事前取得画像についてK個の辞書を用いた評価値が得られたら、辞書ごとに、全ての事前取得画像で得られた評価値を累積加算した値を評価値とする。

20

【0118】

すると辞書選択工程S2113において、S2112で算出された評価値が最も良い辞書を、最適な辞書として選択する。ここで良い評価値とは、評価値がエッジ誤差や距離残差等の検出誤差であればより小さい値、評価値が類似度であればより大きい値であり、評価値の定義に依存する。

【0119】

画像生成パラメータ選択工程S2114では、辞書選択工程S2113にて選択された辞書を生成する際に用いられた画像生成パラメータ候補を、学習画像を生成するために最適な画像生成パラメータとして選択する。すなわち、画像生成工程S1220において上記選択された辞書に対応する学習画像集合Skを生成する際に用いられた画像生成パラメータ候補が選択される。

30

【0120】

以上説明したように第4実施形態によれば、実際の認識結果を評価することによって、最適な辞書を生成するための最適な学習画像を生成するパラメータ(輝度分布パラメータ)を決定することができる。したがって、該決定されたパラメータに基づいて生成された学習画像を用いて最適な辞書を作成することで、該辞書を用いた最適な認識処理を行うことができる。

【0121】

<他の実施形態>

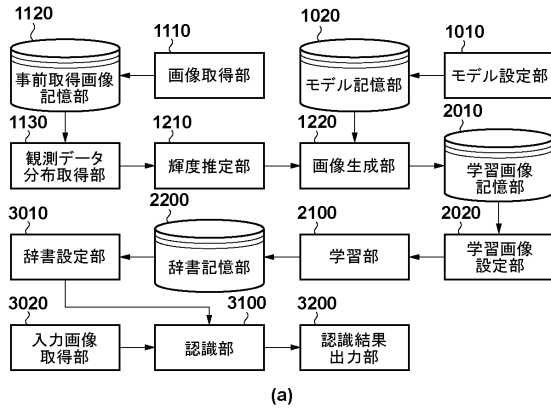
なお、上述した第1～第4実施形態においては、ランタイム時の入力画像を事前取得画像として用いてもよい。そうすることで、ランタイム時の環境変化に対して、動的に適切な学習画像を生成することが可能となる。

40

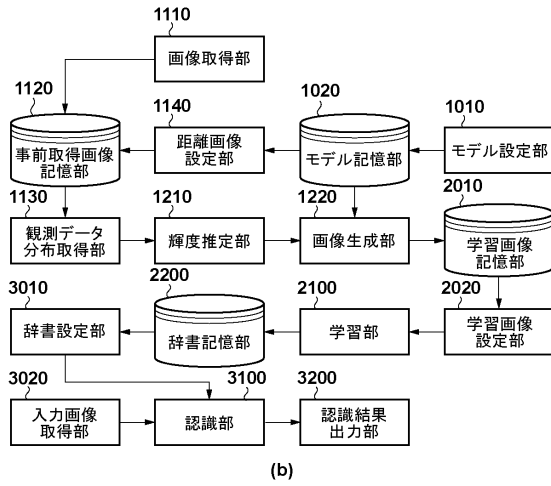
【0122】

また、本発明は、上述した実施形態の機能(例えば、上記の各部の処理を各工程に対応させたフローチャートにより示される処理)を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システム或いは装置に供給することによっても実現できる。この場合、そのシステム或いは装置のコンピュータ(又はCPUやMPU)が、コンピュータが読み取り可能に記憶媒体に格納されたプログラムコードを読み出し実行することにより、上述した実施形態の機能を実現する。

【図 1】

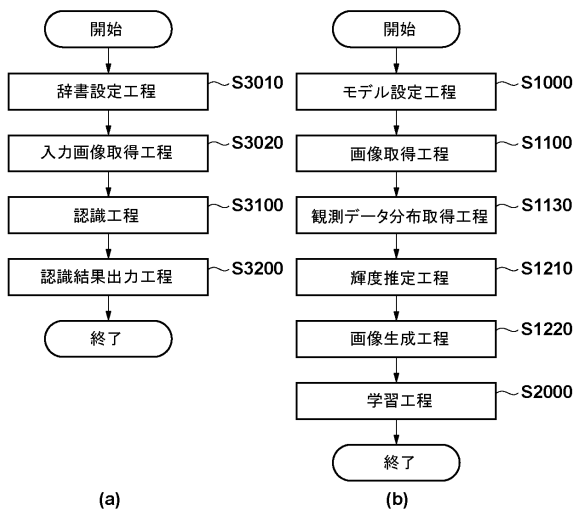


(a)



(b)

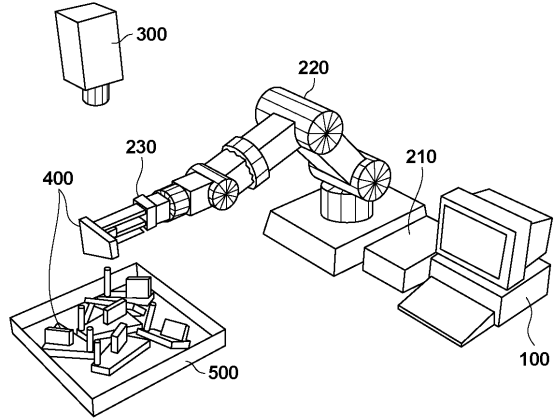
【図 3】



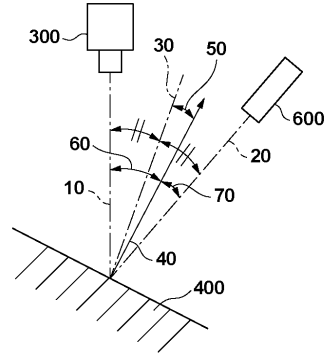
(a)

(b)

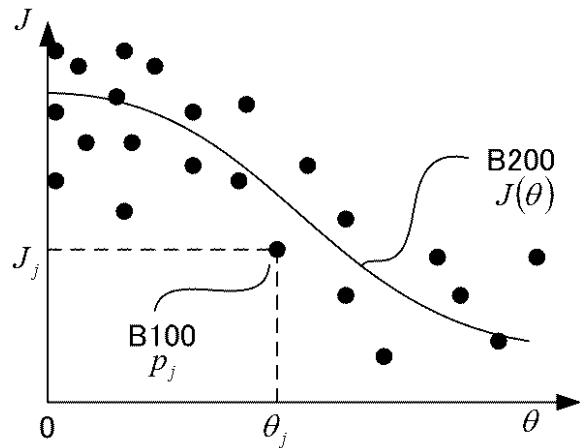
【図 2】



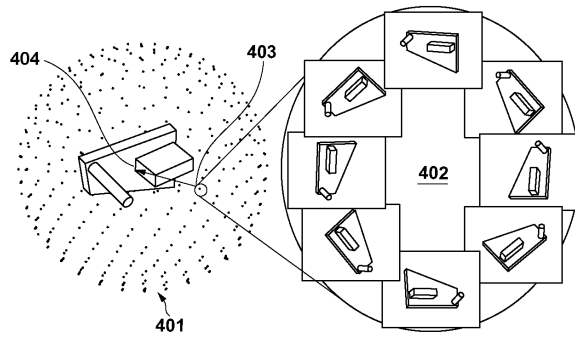
【図 4】



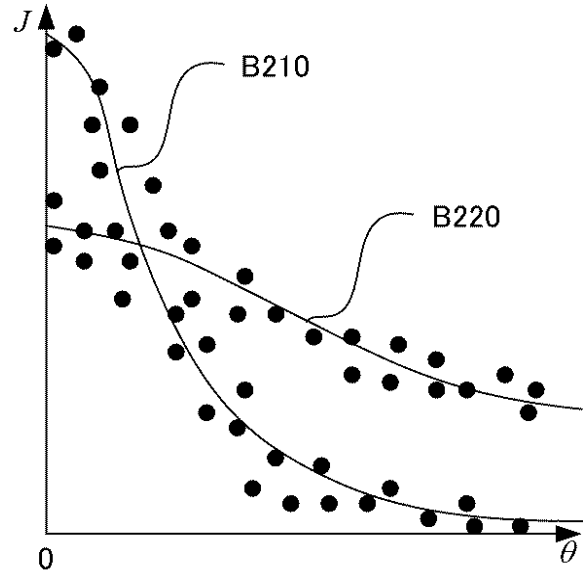
【図 5】



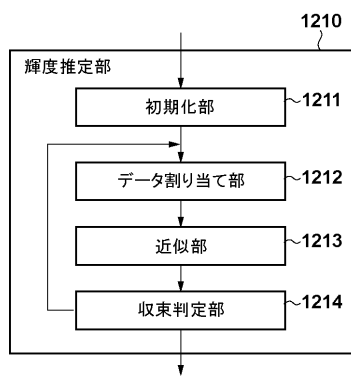
【図 6】



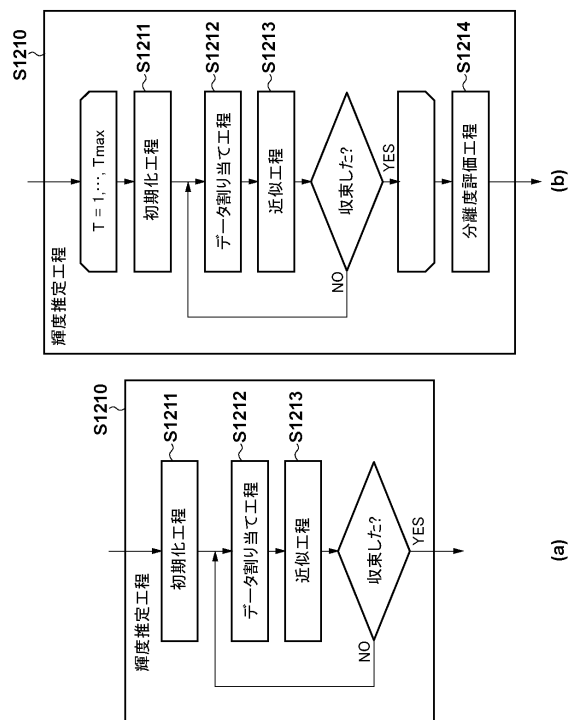
【図 7】



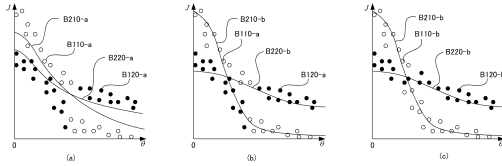
【図 8】



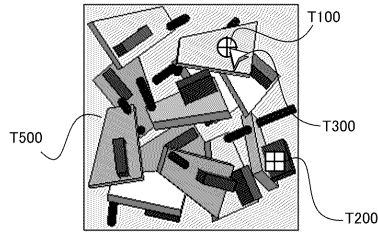
【図 9】



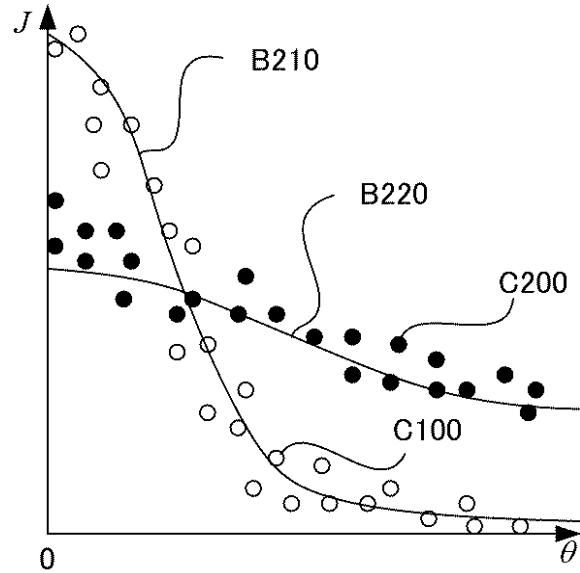
【図 10】



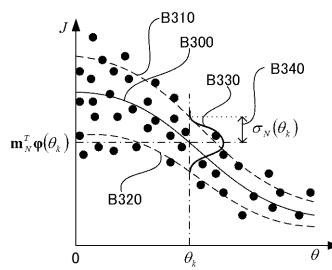
【図 11】



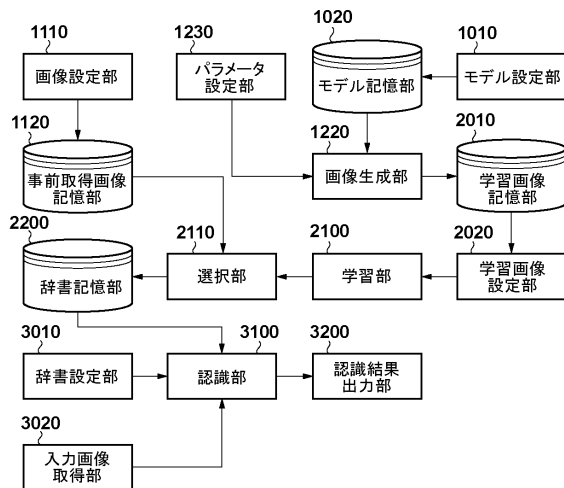
【図 12】



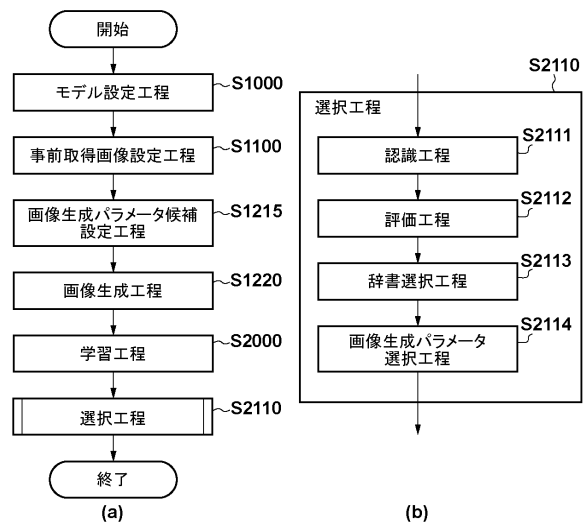
【図 13】



【図 14】



【図 15】



フロントページの続き

(72)発明者 青葉 雅人

東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 佐藤 実

(56)参考文献 特開2010-267232(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/00 - 7/90

G06T 1/00