

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4982583号
(P4982583)

(45) 発行日 平成24年7月25日 (2012. 7. 25)

(24) 登録日 平成24年4月27日 (2012. 4. 27)

(51) Int. Cl.	F I
G 0 6 T 1/00 (2006. 01)	G O 6 T 1/00 3 1 5
G O 1 B 11/00 (2006. 01)	G O 1 B 11/00 H
B 2 5 J 13/08 (2006. 01)	B 2 5 J 13/08 A

請求項の数 10 外国語出願 (全 18 頁)

(21) 出願番号	特願2010-90384 (P2010-90384)	(73) 特許権者	503113186 ホンダ リサーチ インスティテュート ヨーロッパ ゲーエムベーク Honda Research Inst itute Europe GmbH ドイツ連邦共和国 デー 6 3 0 7 3 オ ッフエンバッハアムメイン カール・レギ エン・シュトラーク 3 0
(22) 出願日	平成22年4月9日 (2010. 4. 9)	(74) 代理人	110000246 特許業務法人 OFH 特許事務所
(65) 公開番号	特開2011-22992 (P2011-22992A)	(72) 発明者	マティアス・フランジウス ドイツ国 6 3 0 6 9 オッフエンバッハ、セ ネフェルダークシュトラーク 4 7
(43) 公開日	平成23年2月3日 (2011. 2. 3)		
審査請求日	平成22年6月10日 (2010. 6. 10)		
(31) 優先権主張番号	09159753. 4		
(32) 優先日	平成21年5月8日 (2009. 5. 8)		
(33) 優先権主張国	欧州特許庁 (EP)		

最終頁に続く

(54) 【発明の名称】 視覚に基づく 3 次元形状認識を備えたロボット

(57) 【特許請求の範囲】

【請求項 1】

ビデオセンサからのビデオ信号に表された物体の 3 D 形状に関連する情報であって当該物体の少なくとも 1 つの物理的特性を含む情報を抽出するため、前記ビデオ信号を処理する方法であって、

物体を 3 D 形状空間に格納するためのメモリを提供するステップであって、該形状空間は、該物体の 3 D 形状及び該 3 D 形状に基づく前記少なくとも一つの物理的特性をコード化する抽象特徴空間である、ステップと、

前記形状空間に、物体の 2 D のビデオ信号の表現をマッピングするステップであって、該形状空間における該物体の座標は、該物体の 3 D 形状および前記少なくとも一つの物理的特性を示す、ステップと、

を含む方法。

【請求項 2】

2 D のビデオ信号、および、前記ビデオセンサに対する物体の位置に関する付加的な深度情報が供給される、

請求項 1 に記載の方法。

【請求項 3】

前記物体の形状に関して得られた情報に従って、物体を操作する機械的アクチュエータを制御するための、前記請求項 1 または 2 に記載の方法の使用。

【請求項 4】

請求項 1 から 3 のいずれかに記載の方法を実行するよう設計された、自律型ロボットのためのコンピューティング・ユニット。

【請求項 5】

請求項 4 に記載のコンピューティング・ユニットにビデオ信号を提供する少なくとも 1 つのビデオセンサを備えた自律型ロボットであって、前記コンピューティング・ユニットは、前記ロボットの、機械的な物体マニピュレータを制御する、

自律型ロボット。

【請求項 6】

コンピューティング・デバイス上で実行されたときに請求項 1 に記載の方法を実行する、

コンピュータ・ソフトウェア・プログラム。

【請求項 7】

物体を表すビデオ信号に基づいて、該物体の 3 D 形状および少なくとも 1 つの物理的特性に関する情報を抽出することのできる、コンピューティング・システムを設定するための教師無し学習方法であって、

それぞれ異なる姿勢の訓練物体、または、運動を発生させるための予め定められた刺激を訓練物体に与えた時の当該訓練物体の軌跡を、2 D のビデオカメラの入力視野内に提示するステップと、

前記訓練物体の異なる回転姿勢のビューの、低速に変化する特徴を抽出し、前記訓練物体の形状空間表現をパラメータ化するように、該抽出した特徴をクラスタリングするステップであって、該形状空間は、物体の 3 D 形状及び該 3 D 形状に基づく前記少なくとも一つの物理的特性をコード化する抽象特徴空間である、ステップと、

を含む方法。

【請求項 8】

教師有り学習方法によって追従される、請求項 7 に記載の方法。

【請求項 9】

ビデオセンサからのビデオ信号を処理するための方法であって、

メモリ構造内に形状特徴空間を確保するステップと、

異なる姿勢の訓練物体または訓練物体の軌跡を前記ビデオセンサに提示して、該形状特徴空間を設定するステップと、

前記訓練物体のそれぞれから、少なくとも 1 つ物理的特性を含む特徴を抽出するステップと、

同じ物体の異なるビューが、1 つのクラスタに属する類似の特徴を生成すると共に、異なる物体の類似のビューが分離されるように、前記形状特徴空間において、前記抽出された特徴から特徴クラスタを計算するステップと、

前記特徴クラスタの中心を識別するステップと、

ビデオ信号から目標物体のビューを取得するステップと、

前記目標物体のビューを、前記形状特徴空間にマッピングするステップであって、該形状特徴空間における前記目標物体の座標は、前記目標物体の形状および前記目標物体についての前記少なくとも一つの物理的特性を示す、ステップと、

前記特徴クラスタの中心に対する距離によって表される形状の類似度で、前記目標物体の形状を特徴付けるステップと、

を含む方法。

【請求項 10】

さらに、前記訓練物体の異なる回転姿勢のビューの、低速に変化する特徴を抽出し、前記訓練物体の前記形状特徴空間の表現をパラメータ化するよう、前記抽出された特徴をクラスタリングするステップであって、前記形状特徴空間は、物体の 3 D 形状及び該 3 D 形状に基づく前記少なくとも一つの物理的特性をコード化する抽象特徴空間である、ステップと、

を含む、請求項 9 に記載の方法。

10

20

30

40

50

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、概して、視覚信号、すなわち、ビデオカメラもしくはスチルカメラ(still camera)から供給される信号を処理する分野に関する。本発明は、より具体的には、カメラの入力視野に存在する対象物の3次元(3D)形状(または、たとえば摩擦係数のような、そこから導き出される物理的な特性)を認識するための画像処理の使用に関する。

【背景技術】

【0002】

3D形状認識は、典型的には、対象物を操作するアクチュエータを有するロボットや、他の自動化された対象物取り扱い機械のためのアプリケーションに適用されることができ、これによって、該ロボットや該他の自動化された機械は、自身の挙動を適応させることができ、とりわけ、対象物の認識された3D形状に基づいて、対象物を操作する操作装置(manipulator)の動作(把持力、把持方向、等)を適応させることができる。さらに、該ロボットや該他の機械は、対象物の認識された3D形状から、対象物の二次的な特徴(重さ、脆弱性、剛性(rigidity, stiffness)等)を推定することができ、それに応じて、該二次的な特徴(対象物の3D形状が、一次的特徴である)に対して自身の挙動を適応させることができる。

【0003】

対象物の表面の幾何学形状、すなわち対象物の形状は、環境と相互に作用しあうための関連する情報である。なぜならば、該形状は、対象物の物理的特性を大部分決定するからである。特に、形状は、対象物のアフォーダンス(affordance)に影響する(すなわち、それを何に使用することができるか)。対象物形状の目に対するその二次元的投影から、該形状を推定するのは困難であると共に、概して曖昧であるけれども、人間は、1つのビュー(view)からでさえ対象物の形状を良好に推定することができる。さらに、人間はまた、そのようなビューから、対象物の物理的特性およびアフォーダンスを容易に推定することもできる。

【0004】

一般に、1つの2次元(2D)画像から3D形状を再現することは、不確定なものである。なぜなら、多くの3D形状が、同じ2Dビューを生成する可能性があるからである。簡単な例として、陰影を付けられた3Dの球の写真でテクスチャされた(模様がつけられた)フラットな2Dの円は、該球自体と同じビューを生成することができる。それにもかかわらず、人間は、ほとんど日常の状況において、当該プロセスが実現可能であることを実証している。形状認識に関する非常に多くの研究が利用可能であるが、すべて、特定の副次的な問題に焦点を当てている。我々の知見によると、形状の特徴空間において、ビュー不変(視点不変)のクラスタがどのように関連しているのか、および、3D形状を特徴付けるのに該クラスタをどのように用いることができるのか、について、利用可能な出版物は存在しない。さらに、我々は、1つの(単眼の)2Dビューを用い、かつ定義されたカメラおよび照明の特性という要件を用いずに動作するアプローチに着目する。これは、潜在的に、システムを、より小さく、よりロバストに、かつより柔軟性あるものにする。

【0005】

形状の再現を解決する単刀直入な「エンジニアリング」のアプローチは、特定のアプリケーションのシナリオについての入力(すなわち、ピクセル)空間における有益な特徴を手動で識別し、その後、これらの特徴の組み合わせを識別して個々の形状を特徴付けるものである。不変の対象物認識(invariant object recognition)自体がいまだ非常に困難であるため、そのようなアプローチは、非常に制約された環境についてさえ、作用するとの保証はない。さらに、そのような特定の調整された解決手法の欠点として、各ケースについて関連する特徴を見つけるという作業が必要とされ、さらに、訓練されていない(untrained)形状に対する明白な一般化能力が無い。

10

20

30

40

50

【0006】

アプローチの一つのクラスは、スローネスの原理(slowness principle)の最適化に基づくものである(下記の非特許文献1~6を参照)。これらのアプローチの焦点は、姿勢および照明の変化を含めて、いくつかの変換(transformation)下における不変の対象物認識である。文献6は、さらに、ビュー不変の対象物認識を、対象物の姿勢推定とどう組み合わせることができるかを示している。しかしながら、これらのすべてのアプローチは、本発明によって提案しているような、一般的な形状空間を考慮していない。

【0007】

アプローチの他のクラスは、「Shape from X (SFX)」と呼ばれ、輝度画像から、表面のメッシュ表現を再現しようとするものである(下記の非特許文献7を参照)。これらのアプローチは、陰影からの形状の復元法(shape from shading (SFS))、テクスチャからの形状の復元法(shape from Texture (SFT))、および線画からの形状復元法(shape from line drawings(SFL))を含む。典型的には、これらのアプローチは、特定の光源および表面反射特性を必要とする。SFXアプローチが、その2Dビューを明示的に引き起こした3D対象物表面を再現することを目指しているのに対し、我々のアプローチは、その姿勢に関係なく、全体論的な3D形状を特徴付けることを目指している。表面および照明特性に関するSFXアプローチの前提は、明示的にコード化される必要があるが、我々のアプローチは、教師無しの学習フェーズ中に訓練(training)データの表面特性の関連する側面を抽出する。SFXアプローチに比べて、我々のアプローチにおける認識は、1つのフォワード・パスのみをもたらし認識であるので、計算上非常に効率が良い。完全な3Dのメッシュ表現と比較して、形状特徴空間における活性化パターンは、よりコンパクトであり、3D形状空間において有意な類似度計測を実現するが、このような計測を、再現された3Dメッシュについて見つけることは容易でない。

【0008】

M.Franziusらによる文献「Invariant Object Recognition with Slow Feature Analysis」は、ここで部分的に引用されている、遅い特徴(slow feature、低速に変化する特徴)の解析(SFA)の原理を述べている。

【0009】

W.Einhauserらによる文献「Learning Viewpoint Invariant Object Representations Using a Temporal Coherence Principle」は、複雑なセルおよび対象物セルの分類について、教師無しのクラスタリングを使う方法を述べている。当該述べられている方法は、特定の量の対象物の分類がどのレベルまで成功するかを示す分類パフォーマンス計測を提供することができる。分類に用いられるクラスタの数は、分類されるべき対象物の数に一致する必要がある、各クラスタは、1つの対象物に割り当てられる。クラスタリングのパフォーマンス計測は、正確に分類されたテストパターンの一部について簡単に述べている。

【0010】

H.Muraseらによる文献「Visual Learning and Recognition of 3D Objects from Appearance」は、固有空間に対する対象物の投影が、たとえば照明における変動下でどのように変化するかを述べている。固有空間における2つのポイント間の距離は、対応する輝度画像間の相関の尺度(指標)である。

【先行技術文献】

【非特許文献】

【0011】

【非特許文献1】G. Wallis and E. Rolls: Invariant face and object recognition in the visual system. Progress in Neurobiology, 51(2): 167-194, 1997.

【非特許文献2】S. Becker: Implicit learning in 3d object recognition: the importance of temporal context. Neural Computation, 11(2):347-374, 1999.

【非特許文献3】S. Stringer and E. Rolls: Invariant object recognition in the visual system with novel views of 3d objects. Neural Computation, 14:2585-2596, 2002.

【非特許文献4】W. Einhaeuser, J. Hipp, J. Eggert, E. Koerner and P. Koenig: Learning view-point invariant object representations using a temporal coherence principle. Biological Cybernetics, 93:79-90, 2005.

【非特許文献5】S. Stringer, G. Perry, E. Rolls and J. Proske: Learning invariant object recognition in the visual system with continuous transformations. Biological Cybernetics, 94:128-142, 2006.

【非特許文献6】M. Franzius, N. Wilbert, and L. Wiskott. Invariant object recognition with slow feature analysis. Neural Computation, submitted, 2009.

【非特許文献7】J.-D. Durou, M. Falcone, and M. Sagona: Numerical methods for shape from shading: A new survey with benchmarks. Computer Vision and Image Understanding, 109(1): 22-43, 2008.

【発明の概要】

【発明が解決しようとする課題】

【0012】

本発明の目的は、2次元(2D)の視覚信号から、対象物の3D(3次元)形状情報を取得するための改良された手法を提案することである。

【課題を解決するための手段】

【0013】

上記の目的は、独立請求項の特徴によって達成される。従属請求項は、さらに、本発明の主要なアイデアを発展させるものである。

【0014】

本発明は、教師無し学習方法を提案し、これは、1つのビューから、空間において自由に回転された3次元対象物を認識することができる。たとえば、ゲオン(geon: geometric icon)のような2つの入力クラス、すなわち、レンダリングされた(描画された)対象物と、カメラに提示される対象物の映像(video)、を使用することができる。教師無し学習システムは、視野角とは無関係に、類似形状の対象物をクラスタリングし、訓練対象物(training object)間での3Dモーフィング(morphing)から結果として得られる、それまでは見られなかった形状に一般化する。この一般化の実行は、学習される特徴空間が、より一般的な形状空間のパラメータ化(parametrization, パラメトリゼーション)であることを提案し、該空間では、ビューに固有の特徴に代えて、幾何学的な形状特性がコード化される。そのような形状表現は、アフォーダンス(affordance)に基づく表現について必要な基礎となるものであり、ロボットシステムにおける自動化された対象物操作について重要な潜在性を有している。

【0015】

本発明の第1の側面は、ビデオ信号において表された対象物に関して3D形状の情報(もしくは、3D形状から導出される、たとえば摩擦係数のような、少なくとも1つの特性)を抽出するために、ビデオセンサからのビデオ信号を処理するための方法に関し、該方法は、以下のステップを含む。

【0016】

ー3D形状空間において対象物が格納されるメモリを提供すること。該形状空間は、対象物の3D形状の特性をコード化する抽象特徴空間(abstract feature space)である。

【0017】

ー該形状空間において、対象物の2Dビデオ信号の表現をマッピングすること。ここで、該形状空間における対象物の座標は、該対象物の3D形状(もしくは、そこから導出される物理的特性)を示す。

【0018】

2Dビデオ信号、および、付加的には、ビデオセンサに対する対象物の位置に関する深度情報(depth information)を、供給することができる。

【0019】

ビデオセンサからの信号(すなわち、一連の画像)は、通常、たとえば640*480

10

20

30

40

50

*3というように、いくつかの画像平面からなり、幅×高さの空間レイアウト、およびいくつかの色チャネルを有している。そのような画像平面に代えて、もしくはそのような画像平面に加えて、同じ空間解像度を持つようスケーリングされる深度（奥行き）マップを用いることができる。深度マップは、典型的には、ステレオカメラの設定および深度推定に基づく相互相関によって生成される。適切なスケーリングの後、RGB入力の場合については、該深度マップを、代替の入力として（幅×高さ×1）、もしくは付加的に（幅×高さ×4）、用いることとなる。

【0020】

本発明は、また、対象物の形状についてこうして取得された情報に従って対象物を操作する機械的アクチュエータを制御するため、先行するクレームのいずれかに従う方法の使用に関する。

10

【0021】

本発明のさらなる側面は、自律型ロボットについてのコンピューティング・ユニットに関し、これは、先行するクレームのいずれかに従う方法を実行するよう設計される。

【0022】

本発明は、さらに、自律型ロボットを提案し、これには、コンピューティング・ユニットにビデオ信号を提供する少なくとも1つのビデオセンサが設けられる。該コンピューティング・ユニットは、ロボットの機械的な対象物マニピュレータ（manipulator、操作装置）を制御する。

20

【0023】

さらなる側面は、対象物を表すビデオ信号に基づいて、該対象物の3D形状に関する情報を抽出することができる、コンピューティング・システムを設定するための教師無し学習方法に関し、これは、以下のステップを含む。

— 2Dビデオカメラの入力視野において、動きを誘起する定義された刺激に対する、対象物の異なる回転位置または軌跡のそれぞれに、訓練対象物(training object)を提示すること。

— 該訓練対象物の、低速に変化する特徴(slow feature)を抽出し、該訓練対象物の形状空間表現をパラメータ化するよう、該抽出された特徴をクラスタリングすること。ここで、該形状空間は、対象物の3D形状の特性をコード化する抽象特徴空間である。

30

【0024】

本発明のさらに他の側面は、ビデオセンサからのビデオ信号を処理するための方法を提案する。該方法は、以下のステップを含む。

— メモリ構造において、形状特徴空間を提供すること。

— ビデオセンサに対し、対象物の異なる位置または軌跡において訓練対象物を提示して、該形状特徴空間を設定すること。

— 各訓練対象物から、少なくとも1つの特徴を抽出すること。

— 該形状特徴空間において、該抽出された特徴から、特徴クラスタを計算すること。これは、同じ対象物の異なるビューが、1つのクラスタに属する類似の特徴を生成しつつ、異なる対象物の類似のビューが分離されるように、行われる。

— 該特徴クラスタの中心を識別すること。

40

— ビデオ信号から対象物のビューを取得すること。

— 該形状特徴空間に対して、該対象物のビューをマッピングすること。ここで、該形状特徴空間における対象物の座標は、対象物の形状を示す。

— 該特徴クラスタの中心に対する距離によって表される形状の類似度で、該対象物の形状を特徴づけること。

【0025】

さらなる側面において、クレーム1の方法は、さらに、以下のステップを含む。

— 訓練対象物の異なる回転ビューの、低速に変化する特徴を抽出すること。

— 訓練対象物の形状特徴空間の表現をパラメータ化するよう、該抽出された特徴をクラスタリングすること。ここで、該形状特徴空間は、対象物の3D形状の特性をコード化する

50

抽象特徴空間である。

【0026】

既知の方法が、ビュー不変の対象物表現を学習するやり方を既に確立している。要約すれば、本願のこの提案の新しい革新的な側面は、以下の通りである。

【0027】

—当該分野の最新のアプローチに比して、本システムは、メッシュとして形状を明示的にモデリングすることは行わず、2Dから形状空間に、ダイレクトな抽象化(abstraction)を実現する。これは、ロボットシステムによる対象物の操作について、よりシンプルでより効率的な表現である。

【0028】

—本システムは、1つの（単眼の）2Dカメラのビューから、3D形状の特徴化および比較について、一般化された形状特徴空間を生成する。先行するアプローチでは、任意の3D対象物のビューから形状を特徴化するのに、ビュー不変のクラス間での関連性を考慮していなかった。形状空間における活性化パターンの距離は、3D形状についての類似度計測を実現する。

【0029】

—本システムは、教師無しの視覚に基づく対象物特性および教師有りの視覚に基づくまたは他の対象物特性の効率的な統合を可能にする。こうして、生成された表現は、視覚および非視覚データに基づく、有意な類似度計測を実現する。ほんの少しの教師有りデータが、学習に必要とされる。

【0030】

—本システムは、対象物の物理的な特性（たとえば、形状、摩擦）によって誘起される対象物の動きの統計からの情報を使用して、自律的に、これらの特性の表現を生成することができる。これらは、物理的な対象物と効率的に相互作用する方法を学習するのに、大いに関係している。

【0031】

上記に提示した方法とは別に、本発明は、不変の対象物認識についての方法を目指すものではない。代わりに、本発明は、対象物識別のためだけでなく、計算された特徴空間における距離を用いる。すなわち、この特徴空間における距離を用い、対象物形状の類似度のような、物理的特性を特徴づける。

【0032】

本発明によるさらなる特徴、目的および有利な点は、添付の図面と共に、以下の本発明の実施形態の詳細な説明を読むことにより当業者には自明となろう。

【図面の簡単な説明】

【0033】

【図1】プロセスの概要を示す図。Aは、物理的な3次元対象物が、カメラの前面で動かされていることを示す。Bは、カメラが、対象物のビューを取得していることを示す。Cは、時間的なビューのシーケンスおよびスローネス・オブジェクティブ(slowness objective)の最適化（スローネスを目的とした最適化）に基づいて、特徴表現が計算されることを示す。或る制約条件（たとえば、対象物の動きの軌跡（運動軌跡）のタイプ（種類））が与えられると、この特徴空間は、一般的な(generic)形状空間の特徴を持つ。すなわち、同じ形状を持つ対象物の異なるビューは、対象物の姿勢および外観（たとえば、対象物のテクスチャにより引き起こされる）に関係なく、特徴空間のコンパクトなクラス内で活性化することとなる。大きく異なる形状の異なる対象物のビューは、形状空間の異なる領域にクラスタされる。中間形状の対象物のビューは、それらの間にクラスタされる。複数の軸は、最も遅い(低速、slowest)出力 y_1 、 y_2 、 y_3 を示す。この図におけるクラスタリングの特性は、とりわけ、最も遅い特徴のこれらの軸について明らかである。スローネス最適化のSFA(slow feature analysis)実現では、これらは、SFAの最初のN個の出力である。

【図2】モデルのアーキテクチャと刺激を示す図。入力画像が、階層ネットワークに供給

10

20

30

40

50

される。各層（レイヤ）の円は、オーバーラップする受容野(receptive field)を示し、最上層に向けて収束する。同じ一組のステップが各層に適用され、これが、右側に視覚的に示されている。

【発明を実施するための形態】

【0034】

用語

－対象物(Object)：或る視覚的な外観を備える、任意の3次元の物理的エンティティであり、これは、その形状、および表面の反射特性（たとえば、色、反射率、テクスチャ（模様））により決定される。

【0035】

－形状(Shape)、3D形状：形状は、対象物の表面の幾何学的なものである。形状は、色のような単なる光学的な表面特性とは無関係である。

【0036】

－ビュー(View)：対象物のビューは、或る距離、回転角、および照明状態で、対象物を2次元に投影したものである。カメラまたは目は、対象物のビューを生成することができる。

【0037】

－形状空間(Shape space)：ビューの特性についてはほとんど、もしくは全くコード化せず、対象物の形状の特性をコード化する、抽象的なn次元の特徴空間である。本発明に従う意味での特徴空間は、姿勢および照明の変化によって誘起される外観の変化に対しては不変である。このような空間において、（たとえばMuraseによって記述される）モデルの対象物軌跡は、コンパクトな（小さな）クラスタ内に収まる。対象物形状間の類似度を、該形状空間における2つの表現間の距離として、該形状空間内において計測することができる。Muraseによる該モデルにおける1つのポイントpは、たとえば、特定の照明状態下における特定の対象物Aの特定の姿勢をコード化する。

【0038】

－ゲオン(Geons)：後述する参考文献[Biederman1987]において最初に提案された、一組の幾何学的形状のプリミティブであり、球、円柱、立方体、円錐等を含む。

【0039】

－不変性および特定性(Invariance and specificity)：刺激(stimulus)に対して変換(transformation)が適用されるときに表現が変化しなければ、表現rは、変換t下において不変である。たとえば、ビュー不変(view invariance)は、（特徴の）表現が、異なる視覚(perspective)からの対象物のビューについて一定のままであることを意味する。自明な不変性（たとえば、すべてのビューについて一定の表現）は、表現が、最小の変動(variance)を持つ場合に回避されることができる。適用される変換tおよび表現の間にマッピングが存在すれば、該表現は、変換tに固有(specific)である。たとえば、ビュー不変の対象物認識について表現の所望の特性は、視点に対しては不変であるが、対象物のアイデンティティ(identity)については特定性を持つ。

【0040】

本発明は、2D画像および画像のシーケンスから、対象物の幾何学的な形状特性を抽出する手法を提示する。さらに、本システムは、より一般的な類似度計測を実現するため、視覚的、もしくは非視覚的な情報を使用することができる。この抽出プロセスは、2次元のビューに基づいているが、該システムは、その動作フェーズ中において、3次元の対象物の特定のビューからは大きく独立している。最初のフェーズ、すなわち訓練フェーズにおいて、該システムは、ラベル付けされていない画像シーケンスから、関連する画像の統計を抽出する。訓練フェーズは、訓練データセットにおいて、典型的には、非常に低速な、もしくははめったに変化しない特徴を見つけるため、「スローネス・オブジェクティブ(slowness objective)」に基づく確立されたアルゴリズムを使用し、これについては、たとえば、「Slow Feature Analysis（後述する参考文献のWisSej2002）」、「the trace rule (Foeldiak1991)」、「temporal stability (Kayser2001)」を参照されたい。さらなる制

10

20

30

40

50

約条件は、複数の非自明解が（定常解のように）計算されることを保証する。当該提案されるシステムは、スローネス・オブジェクティブを最適化（スローネス（遅さ）を目的とした最適化）する1つまたは複数のステップを使用する。典型的には、該複数のステップは、後述する参考文献の(Franzius2007a)にあるような、収束する階層的なレイアウトにおいて実行される。既存のシステムに比して、教師有りの情報を、スローネス・オブジェクティブの教師無し最適化によって生成された特徴を適応するのに使用することができ、対象物の物理的特性の情報を自律的に取り入れることができる。

【0041】

訓練フェーズ後の動作フェーズにおいて、該システムの出力は、2D入力ビューから、「形状特徴空間(shape feature space)」へのマッピングMである。このマッピングは、瞬時的に計算されることができるので（算術的な意味で）、非常に効率的である。すなわち、1つの2D入力ビューは、1つの形状特徴表現を生成することとなる。抽出された特徴は、形状空間において、クラスタを形成することができ、これは、同じ対象物の異なるビューが、類似の特徴を生成すると共に、異なる（別個の）3D対象物の類似のビューが分離されるように、行われる。

【0042】

このクラスタリングの特性については以前に出版されているけれども、新しい品質が、該クラスタの中心の相互関係を識別して使用することにより、ここで取り入れられる。さらに、これらのクラスタの距離を直接的に制御するための手法が、導入される。抽出された形状空間は、2Dの外観空間ではなく、3Dの形状空間において、他の類似度計測と共に、類似度計測を実現することができる。具体的には、該システムは、訓練フェーズ中に提示されていない形状の対象物のビューに一般化する（すなわち、有意の結果を生成する）。たとえば、該システムが、2つの形状の画像シーケンスで訓練され（たとえば、立方体および球）、中間の形状（たとえば、変形された”球-立方体”）のビューに適用されるならば、該抽出された特徴は、立方体と球の間の中間値を持つこととなる。こうして、任意の形状のビューの該抽出された特徴は、既知の形状クラスタの中心に対する距離で、有意に、その形状を特徴づける（たとえば、立方体に対する30パーセントの類似度、および球に対する70%の類似度）。

【0043】

本発明は、さらに、以下のように発展が可能である。

【0044】

代替の入力データ

本システムは、2Dビューを処理することに制限されず、代替的に、または付加的に、3Dの深度データ（たとえば、レーザスキャナからの）、複数の2Dビュー（たとえば、空間的に離れたカメラの）、およびこれらの組み合わせに適用されることも可能である。

【0045】

事前処理

後述する参考文献中の(Franzius2007a、Franzius2009)に記載のような以前のモデルは、人工的にレンダリングされたデータにのみ適用された。それに対し、当該提案されるシステムは、取り散らかったシーン（状況）で起こりうる、カメラによって生成された実世界のデータに適用される必要がある。したがって、典型的には、セグメント化のステップが、生の画像データの事前処理について実行される。

【0046】

非剛性の（変形可能な）対象物

選択的に、非剛性の（変形可能な）対象物を、本システムによって認識することができる。訓練フェーズ中に、非剛性変形(non-rigid deformation)を受けている対象物のビューのシーケンスが提示されたならば、本システムは、該非剛性の対象物の異なる構成(configuration)を、1つの形状に関連づけるよう学習することができる。このようなシーケンスはまた、たとえば形状Aと形状Bの間の変形(morph)として、人工的に生成されることもでき、これにより、該システムによってAおよびBの形状表現がより類似したものと

なるようにすることができる。

【0047】

訓練データのシーケンスが、モデル出力の不変な特性を決定するので（後述する参考文献中のFranziusu2007aを参照）、該訓練データのシーケンスは、該モデルがビュー不変および形状固有（形状に特定）の特性を生成するよう、所定の特性を持つ必要がある。たとえば、対象物は、たとえばそれらが深度方向において回転を受けている間、長時間にわたり視覚的でなければならない。訓練データセットの動き特性が、該訓練フェーズ中に直接的に制御されることができず、かつ、該動き特性が、ビュー不変の形状特徴を抽出するのに適していなければ（たとえば、異なる対象物のビュー間での即時の切り換え）、学習が、適した動きの最中にのみ起こるように、該学習のレート（速度）を制御することが可能である（後述する参考文献中のFranziusu2007aを参照）。この原理は、自己中心性空間コード（egocentric spatial codes）の学習のコンテキストにおいて、以前に提案されている。

10

【0048】

姿勢学習

不変の形状認識に対する可能な拡張は、姿勢および位置抽出である。姿勢および位置抽出は、それ自体新しいものではないが（後述する参考文献中のFranziusu2009を参照）、ここで取り入れられる一般化された形状認識については新しい品質がある。

【0049】

学習に基づくアフォーダンス(affordance)

20

以前のアプローチは、スローネス・オブジェクティブ(slowness objective)を最適化することに基づいて、システムの不変の特性について、対象物の動きの統計のいくつかの側面を考慮していた（後述する参考文献中のFranziusu2007aを参照）。しかしながら、以前の作業では、このコンテキストにおいて、対象物の物理的特性によって誘起される動きの統計を考慮するものは無かった。一例として、刺激の後の対象物の動きを考慮すると（たとえば、1つの空間方向にインパルスを適用することによって）、これは、対象物の形状特性に依存した動きの軌跡（運動軌跡）を引き起こす。球状の対象物のような、いくつかの対象物は、比較的高速度で、長い時間にわたって動くであろうが、立方体のような形状の他の対象物は、動きを速やかに止めるであろう。これらの動きの軌跡が与えられると、スローネス原理を最適化することは、異なる軌跡を有する対象物の異なる不変性および特定性に至る。具体的には、動きの軌跡が、同じ最大振幅を持つ（たとえば、対象物の動きが、ボックスによって制約される場合）が、異なる周波数を持つとき（たとえば、球状の対象物については高く、立方体状の対象物については低い）、対象物固有のクラスタは、この周波数に依存した直径を持つこととなる。スローネス・オブジェクティブを最適化しつつ、最小の変動（variance, 分散）という制約を満たすため、対象物の典型的な動きの周波数が高くなるほど、対象物に固有のクラスタはよりコンパクトな（より小さな）ものとなる。その後、クラスタの直径が計算され、教師無しの手法で解釈されることができる。こうして、当該提案するシステムは、自律的に、かつ教師無しの手法で、対象物の物理的特性（たとえば、形状）に基づいて該対象物の表現を構築する。このような能力を備えているので、システムは、自律的に、対象物との相互作用を行う（たとえば、所望の位置において対象物を握ったり、位置付けたりする）のに必要とされる、それまでは未知であった対象物の関連する物理的特性について、自律的に学習することができる。

30

40

【0050】

教師有り情報の統合

教師無しの訓練フェーズおよび動作フェーズを、中間の教師有り学習フェーズによって拡張することが可能である。ビューのシーケンスが提示される訓練フェーズの最初の部分は、変化しないままである。訓練フェーズの最後において、新しい教師有り訓練フェーズが挿入される。スロー特徴解析(低速に変化する特徴の解析、Slow Feature Analysis(SFA))でスローネスの最適化を実現する場合には、このステップは、データの共分散行列およびデータの導関数を計算した後であって、最適にスローな解を見つけるために該行列を対

50

角化する(diagonalize)前に、挿入される。

【0051】

新しい教師有り訓練フェーズは、所望の類似度計測(similarity measure)に対してより類似するよう、計算された特徴を適応させるのに寄与する。いくつかのビューについて類似度行列 S が与えられると(たとえば、以下に示すように、アフォーダンスに基づいて)、該システムは、スローネスの学習ルールによって計算された特徴表現を、それが、 S に対してより類似したものとなるよう、適応させる。これは、ビューの対 (v_i, v_j) を、教師無し学習フェーズからのものに付加的に、該システムに対して「マイクロシーケンス(micro sequences)」として提示することによって、実現される。具体的には、 S において高い類似度の要素 s_i, j を有するビューの対 (v_i, v_j) が、システムに示される。代替的に、すべての対 (v_i, v_j) を、システムに提示することができ、該システムは、 s_i, j に比例して、その学習速度を適応させる。 SFA の場合には、そのような学習速度の適応は、データの共分散行列およびデータの導関数に対する更新を、 s_i, j に比例した因子で重み付けすることを含む。それに対し、勾配法(gradient-descent)に基づく実現は、それらの本来存在する学習レートを、 s_i, j に比例した因子を乗算することによって適応することができる。

10

【0052】

類似度行列は、該システムに対する要件として考えられる。結果として、形状特徴空間における一対の入力 a, b の距離は、以下に近づくものとなる。

【数1】

20

$$\text{norm}(S(a), S(b)) \sim 1/M(a, b)$$

【0053】

この結果は、ここで説明しているように、いくつかの刺激について、教師有りのやり方で時間的な提示順序を変更する手段によって達成される。こうして、形状空間における距離が計測されることとなり、これらの距離は、所望の、かつ所与の類似度行列 S に近づく。

【0054】

他の以前に記述されている方法に対する主な違いは、対象物の軌跡からの情報を、対象物の形状のような物理的特性に関する情報を収集するのに使用することである。以前の作業においては、異なる対象物の動きの軌跡は、可能な限り類似したものとなるよう注意深く選択されてきた。

30

【0055】

既知の方法において、異なる対象物間において、いかなる時間的なコンテキスト(ないし、いかなるシステマティックな(systematic)時間的コンテキスト)も、訓練データには存在しない。これは、一般に、カメラの前の対象物を切り換えるのに、時間がかかり、その間はビューにおいて背景のみが存在するためである。理論上、クラスタの中心の相対的な位置は、この場合には任意である。すなわち、クラスタ中心の任意の並べ替えは、スローネス最適化の等しく良好な解となる。これらの可能な並べ替えは、實際上自由であり、該システムの制御されないパラメータである。実際の実現において、特に階層的な場合において、これらの自由パラメータは、中間表現の、制御されていない、かつ理解が十分でない特性に基づいて設定される。ここで提案される教師有り訓練フェーズは、以前には決定されなかった、これらの表現の自由パラメータを、有意な方法で固定する。すなわち、所望の類似度計測を表すように固定する。この点における S の統合(integration)は、その後の教師有り分類に比べて、より効率的である。なぜなら、相対的なクラスタ距離の設定は、学習プロセスのこの点において、ほとんど付加的な情報を必要としないからである。

40

【0056】

効果として、結果として生じる特徴表現は、訓練フェーズ中の同一の対象物の、時間的

50

に近傍した異なるビューから学習されるように、ビューの不変性を直接実現すると共に、対象物間の所望の類似度計測を直接実現することができる。そのような所望の類似度計測の一例は、テクスチャ不変の、かつビュー不変の形状空間である。

【0057】

後述する参考文献中のFranzius2007,2009によって提示されているモデルの動作フェーズにおいて、教師有り分類(classification)もしくは回帰(regression)を、たとえばビュー不変の対象物分類について実行することができる。そのようなステップは、ここで提案される方法についても、なお可能である。それにもかかわらず、分類器(classifier)を、よりシンプルに(たとえば、線形に)することができ、より少ない訓練データを必要とし、これは、より効率的なものとする。なぜなら、基礎となる特徴空間は、分類タスクについてより良好に適応されるからである。こうして、ごくわずかのビューの対(v_i, v_j)のみが、結果として生じる特徴空間において所望の類似度を実現する(よって、たとえば、対象物分類を実行する)のに必要とされる。

10

【0058】

教師有り学習に何の対(v_i, v_j)も用いなければ、該システムは、後述する参考文献中のFranzius2007,2009に示されるように挙動する。何のスローネス最適化も実行されず、かつ、ビューの対(v_i, v_j)の距離の形態で該システムによってすべての情報が提供されるならば、該システムは、實際上、FDA(Fisher Discriminant Analysis)を実行する。こうして、ここで提案されるアプローチは、両方のものを統合したものを実現し、FDAの制御可能性とスローネス・オブジェクティブのパワーを組み合わせ、ビュー不変性を生成する。

20

【0059】

類似度行列Sの特定の形態

原則として、類似度行列Sは、視覚的に関連するように、もしくは関連しないように、任意の意味を持つことができる。特定の関連する類似度行列のクラスCSは、対象物または対象物クラスの物理的機能またはアフォーダンスに関連する。この情報は、通常、純粋に視覚的なデータ(すなわち、対象物の外観)からは抽出されることができない。我々は、Sによって実現される、以下のような類似度計測を提案する。

【0060】

一対象物の物理的特性。これは、平坦な表面上における摩擦および動きのタイプを含む。丸い対象物は、平坦な表面上をまっすぐに動くことができるのに対し、シャープなエッジを持つ対象物は、典型的には、より早期に停止に至る。対象物を押した後に該対象物が停止に至るまでの動きの平均時間を計測することによって、類似度行列Sを、物理的形状および摩擦特性(これらの対象物を操作するのに必要な自律型システムにとって、対象物のより重要な特徴である)を特徴付けるのに使用することができる。このような特性の計測を、手動で、もしくは自律型システム自体によって、実行することができる。前述した完全に自律型のアプローチに比べて、このアプローチは、類似度行列Sを識別するための付加的なプログラミングをいくらか必要とする。

30

【0061】

一対象物の分類(categorization): 類似度行列Sがカテゴリ・メンバシップ(category membership)に基づく場合には、別個の視覚的な外観を有しているが、類似のカテゴリ・メンバシップを備える対象物を、該システムによりクラスタリングすることができる。たとえば、類似の色の対象物のビューは、類似のサイズの対象物のビューと同様に、一対で(ペアで)、高い類似性を持ちうる。しかしながら、学習は、対象物の軌跡が、他の対象物によって影響される場合に制限され、これは、自由な回転および非常に低い摩擦の場合を除外する。

40

【0062】

関連する特徴の取得

既知の方法は、一組の問題固有の特徴(problem-specific feature)を手加工することにより、もしくは一組の適切な問題固有の特徴を、機械学習アプローチで、予め定義された

50

特徴の集まり（プール）から選択することにより、対象物認識のための特徴を選択する。これらの手法は、階層的なやり方でも適用可能である。

【0063】

本発明は、スローネス・オブジェクティブが最大化されるように、入力チャネルの重み付けされた非線形組み合わせを計算することによって、自動的に特徴を生成する既存のアプローチに基づいている。このアプローチは、多くの場合良好に動作するけれども、該特徴生成プロセスを制御する直接的な方法が無い。本発明は、該特徴生成プロセスに対して、教師有りの影響を付加する選択的な手法を提案する。固有の特性のクラスが、形状空間の生成について関連することが既知であれば（たとえば、コーナ（角）ないしエッジ（端）を示す部分的なビュー）、これらの部分的なビューを、完全な（フルの(full)）訓練ビューの代わりに（もしくは該訓練ビューに加えて）、階層的なモデルにおいて低い層を訓練するのに使用することができる。このようにして、該システムを、低い層におけるこれらの部分的なビューの各クラスについて類似の表現を計算するのに訓練することができる（たとえば、視角、照明、テクスチャとは無関係に）。

10

【0064】

提案するシステムのアプリケーションの例

自動化された形状認識は、ある程度の自律性および検知するデバイス（たとえば、カメラ）を備える任意の装置の一部となりうる、非常に基本的なツールである。

【0065】

・視覚データに基づいて、照明の状態、対象物の姿勢、および対象物の表面（テクスチャ）に対しては不変なように、形状もしくは他の物理的な特性による対象物の自動分類（ソート）。このような分類は、固定的なパターンマッチングの代わりに（たとえば、丸いポテト対細長いポテト）、形状の類似度によるソートを行うことができる。

20

【0066】

・ロボット工学のための形状認識。対象物の形状および姿勢は、ロボットのデバイスが、最も良好に、対象物を握ることのできる方法を決定し、あるいは該対象物の任意の操作を実行する方法を決定する。

【0067】

方法

スローネスな目的関数を、生の2D画像のシーケンスについて直接的に最適化することができる。しかしながら、多くの場合、入力ビューの寸法（大きさ）は、計算上、効率的な最適化には高すぎる。この場合、階層モデルが適用され、最低の層上で、入力ビューの小さなパッチを、最適化に使用されるようにする。いくつかの近傍領域の出力が、その後、次の層に供給され、該次の層は、再び、同じ目的関数を最適化する。これが、階層が所定の分解能(resolution)に収束するまで行われ、極端な場合には、1つの位置に収束するまで行われる。

30

【0068】

方法

刺激の生成

モデルは、5つの異なる凸状の対象物（球、円錐、立方体、ピラミッド、および円柱）の色づけされたビューを含む画像のシーケンスで訓練された。2つの異なる入力セット、すなわち「レンダリングされた(rendered)」ものと「映像(video)」とが使用された。レンダリングされたデータについて、視覚化ツールキット（Visualization Toolkit(VTK)）が、テクスチャを有するもしくは有しない均一な背景の前面にある、陰影付けされた対象物のビューをレンダリングするのに使用された。さらに、その形状が、5つの訓練形状のうちの任意のものの間での漸進的な補間(gradual interpolation)として設定されることのできる「変形された形(morphed figure)」から、テストデータが生成された。訓練シーケンスのための対象物の姿勢（構成;configuration）が、ランダムウォーク手法(random walk procedure)によって生成された。該シーケンスにおける該構成を生成するため、ランダムな項(random term)を、対象物の、現在の空間的、角度的、およびスケーリングの

40

50

速度に加えた。該ランダムな項は、均一の確率密度の間隔から抽出される。これらの速度は、所定のリミット値でカットオフされ、これらのリミット値を調整することによって、変換(transformation)のタイムスケールを効果的に決定することができる。位置、角度、およびスケールは、その後、これらの速度に従って更新される。対象物が、境界位置に達したならば、該対象物は戻される。この手順全体は、フラットな構成(configuration)のヒストグラムを生成し(十分な時間点が与えられると)、該速度のプロファイルは、該構成の値からは独立している。各ステップにおいて、対象物のアイデンティティは、低い確率($p = 0.02$)で変更された。刺激において、同一の姿勢にある異なる対象物をまとめてリンクするのを回避するようスイッチ(切り換え)が生じた場合には、ブランクのフレームが挿入された。これは、教師有り訓練の要素を導入することとなる。

10

【0069】

ビデオ(映像)セットについては、紙ないし発泡スチロールの対象物の3つの変形物(variant)が使用された。1つの変形物は、白のままであり、他の変形物は、テクスチャ(模様付け)された。さらに、6個のソーダ缶、6個のゴム製のアヒル(rubber duck)、6個のボール箱(たとえば、ティーバッグ)、および6個の球状の対象物(たとえば、ボールおよびオレンジ)から成る、日常の対象物のHRI50データベースの一部が使用された。これらの対象物は、取り散らかったオフィス環境において、カメラの前面の手元にあるよう維持された。ステレオカメラの深度キュー(奥行きの手がかり, depth cue)および皮膚の色の検知に基づいて、対象物のビューは、参考文献中にあるWersingKirsteinEtAl2007に記載のさらなる処理の前に、背景からセグメント化された。このセグメント化ステップにおける変動に起因して、対象物は、常に、完全にセグメント化され、中央に置かれ、もしくはスケールされるわけではなかった。

20

【0070】

スローな特徴解析 (Slow Feature Analysis (SFA))

最適化問題: 関数空間 F および I 次元の入力信号 $x(t)$ が与えられた場合、以下の制約条件下で、出力信号 $y_j(t) := g_j(x(t))$ が、 $\Delta y_j(t) := \langle y_j'^2 \rangle_t$ を最小にするような、一組の J 個の実数値(real-valued)の入出力関数 $g_j(x) \in F$ を見つける。

【0071】

- (1) $\langle y_j \rangle_t = 0$ (zero mean (平均ゼロ)),
- (2) $\langle y_j'^2 \rangle_t = 1$ (unit variance (分散1)),
- (3) $\forall i < j: \langle y_i y_j \rangle_t = 0$ (decorrelation and order (無相関および順序付け)),

30

ここで、 $\langle \rangle$ および y' は、時間的な平均および y の導関数をそれぞれ示す。

【0072】

上記に導入された Δ 値は、信号 $y(t)$ の時間的なスローネス(すなわち、遅さ、低速さ)の尺度である。これは、該信号の時間的導関数の二乗平均により与えられ、小さな Δ 値が、低速に変化する信号を示すようにする。上記の制約条件(1)および(2)は、自明な定常解を回避し、制約条件(3)は、異なる関数 g_j が、入力異なる側面についてコード化することを保証する。制約条件(3)のために、 g_j は、また、それらのスローネス(低速さ)に従って順序付けられ、 g_1 は、最も小さい Δ を持つ。実際のアプリケーションにおいて、典型的には、最初の n 個の解のみを用い、より高速の g_i を破棄して、結果として生じるデータの次元数(dimensionality)を制御する。

40

【0073】

スローネスを目的としたものだけでも(スローネス・オブジェクティブ)、関数 g_j は、入力に対して瞬時に計算される関数であり、よって、該スローネスは、ローパスフィルタリングによっては達成できない点に注意するのが重要である。入力信号が、関数 g_j によって瞬時に抽出されることのできる、低速に変化する特徴を制約する場合にのみ、スロー(低速)出力信号を取得することができる。また、同じ理由で、一旦訓練されると、該システムは高速に(低速ではない)動作する点に注意されたい。

【0074】

50

「低速に変化する特徴(slowly varying feature)」という用語は、たとえば、「Slow Feature Analysis」についてのデルタ値のように、スローネス学習ルールの最適化によって生成される特徴を示す。こうして、「低速に変化する特徴」は、算術的に良好に定義されるコンセプトである。

【0075】

Fが有限次元(finite-dimensional)である、計算的に関連する場合においては、最適化問題に対する解を、後述する参考文献中のWisSej2002に記載のような、Slow Feature Analysis (SFA) の手段によって、およびBerkWisk2005cにおけるわずかに異なる定式化において、見つけることができる。このアルゴリズムは、固有ベクトルのアプローチに基づいており、グローバルな最適性を見つけていることを保証する。我々は、オープンソースのMDPライブラリ (Modular toolkit for Data processing (MDP)) におけるSFA実現を使用する。

【0076】

ネットワーク設定

計算上のモデルは、SFAノードの、収束する階層レイヤからなる。各SFAノードは、その入力から、SFAアルゴリズムに従って、最も低速の特徴(slowest features)を見つけ、以下のシーケンスの動作、すなわち、加法性ガウス白色雑音 (10^{-6} の分散)、次元数低下のための線形SFA、二次的拡張(quadratic expansion)、低速な特徴抽出のための他の線形SFA、および±4における極値のクリッピング、を実行する。ネットワークは、Python (パイソン) において実現され、すべての必要なエレメント (並列化を含む) は、MDPライブラリ (後述する参考文献中のMDPを参照) で利用可能である。

【0077】

参考文献

[Becker1999]: S. Becker: Implicit learning in 3d object recognition: the importance of temporal context. Neural Computation, 11(2):347-374, 1999.

[BerkWisk2005c]: P. Berkes and L. Wiskott: Slow feature analysis yields a rich repertoire of complex cell properties. Journal of Vision, 5(6): 579-602, 2005.

[Biedermann1987]: I. Biedermann: Recognition-by-Components: A theory of human image understanding, Psychological review, 94(2): 115-147, 1987.

[Durou2008]: J.-D. Durou, M. Falcone, and M. Sagona: Numerical methods for shape from shading: A new survey with benchmarks. Computer Vision and Image Understanding, 109(1): 22-43, 2008.

[Einhauser2005]: W. Einhaeuser, J. Hipp, J. Eggert, E. Koerner and P. Koenig: Learning view-point invariant object representations using a temporal coherence principle. Biological Cybernetics, 93:79-90, 2005.

[Foeldiak1991]: P. Foeldiak: Learning invariance from transformation sequences. Neural Computation, 3:194-200, 1991.

[Franzius2007a]: M. Franzius, H. Sprekeler, and L. Wiskott: Slowness and sparseness lead to place-, head-direction and spatial-view cells. PLoS Computational Biology, 3(8):e166, 2007.

[Franzius2009]: M. Franzius, N. Wilbert, and L. Wiskott. Invariant object recognition with slow feature analysis. Neural Computation, submitted, 2009.

10

20

30

40

50

[Hashimoto2003]: W. Hashimoto: Quadratic Forms in Natural Images. Network: Computation in Neural Systems, 14(4): 756-788, 2003.

[Kayser2001]: C. Kayser, W. Einhaeuser, O. Dueemmer, P. Koenig and K. Koerding: Extracting slow subspaces from natural videos leads to complex cells. Artificial neural networks – ICANN 2001 Proceedings, 1075-1080, 2001

[MDP]: P. Berkes and T. Zito: Modular Toolkit for Data Processing (MDP), version 2.0. <http://mdp-toolkit.sourceforge.net>, 2005.

10

[Sprekeler2007]: H. Sprekeler, C. Michaelis, L. Wiskott: Slowness: An objective for spike-timing dependent plasticity? PLoS Computational Biology, 3(6): e112, 2007

[Stringer2002]: S. Stringer and E. Rolls: Invariant object recognition in the visual system with novel views of 3d objects. Neural Computation, 14:2585-2596, 2002.

[Stringer2006]: S. Stringer, G. Perry, E. Rolls and J. Proske: Learning invariant object recognition in the visual system with continuous transformations. Biological Cybernetics, 94:128-142, 2006.

20

[Wallis1997]: G. Wallis and E. Rolls: Invariant face and object recognition in the visual system. Progress in Neurobiology, 51(2): 167-194, 1997.

[WisSej2002]: L. Wiskott and T. Sejnowski. Slow Feature Analysis: unsupervised learning of invariances. Neural Computation, 14(4): 715-770, 2002.

【 0 0 7 8 】

さらに、以下の文献も参考にされる。

Mathias Franzius et al: "Invariant Object Recognition with Slow Feature Analysis" Artificial Neural Networks - ICANN 2008; Lecture Notes in Computer Science, Springer Berlin Heidelberg, vol. 5163. 3 September 2008 (2008-09-03), pages 961-970, ISBN: 978-3-540-87535-2

30

Wolfgang Einhauser et al: "Learning viewpoint invariant object representations using a temporal coherence principle" Biological Cybernetics ; Advances in Computational Neuroscience, Springer, Berlin, vol. 93, no. 1, 1 July 2005 (2005-07-01), pages 79-90, ISSN: 1432-0770

Murase H et al: "Visual Learning and Recognition of 3-D Objects from Appearance" International Journal of Computer Vision, Kluwer Academic Publishers, Norwell, US, vol. 14 no. 1, 1 January 1995 (1995-01-01), pages 5-24, ISSN: 0920-5691

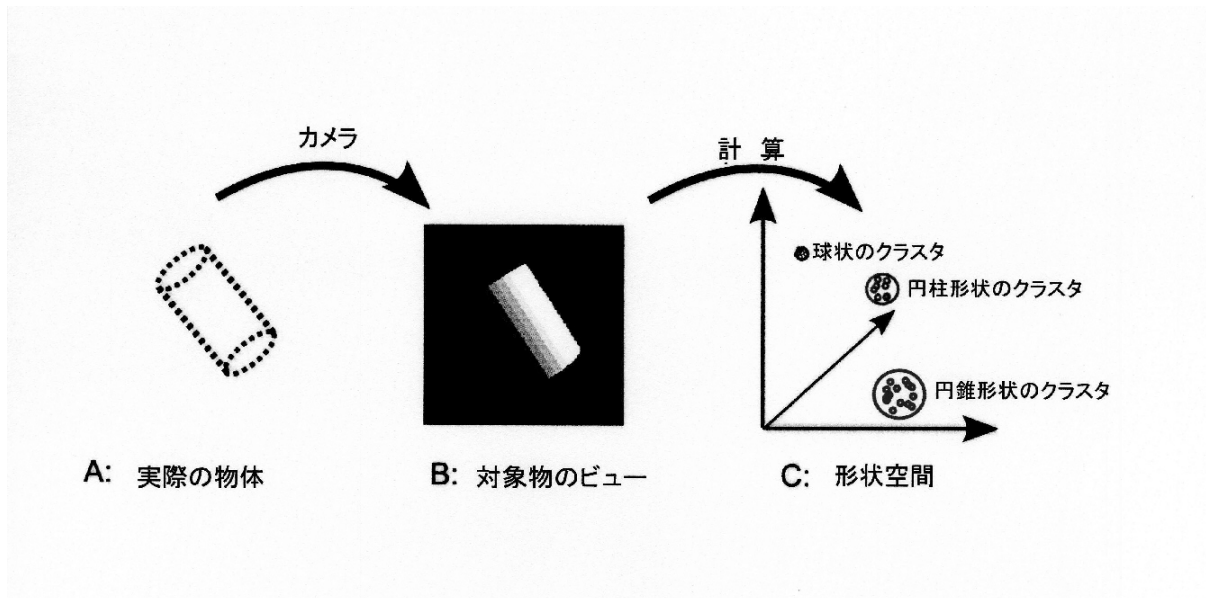
40

M. Franzius et al.: "Slowness and sparseness lead to place, head-direction, and spatial-view cells" PLoS Computational Biology, vol. 3, no. 8. August 2007 (2007-08), pages 1605-1622,

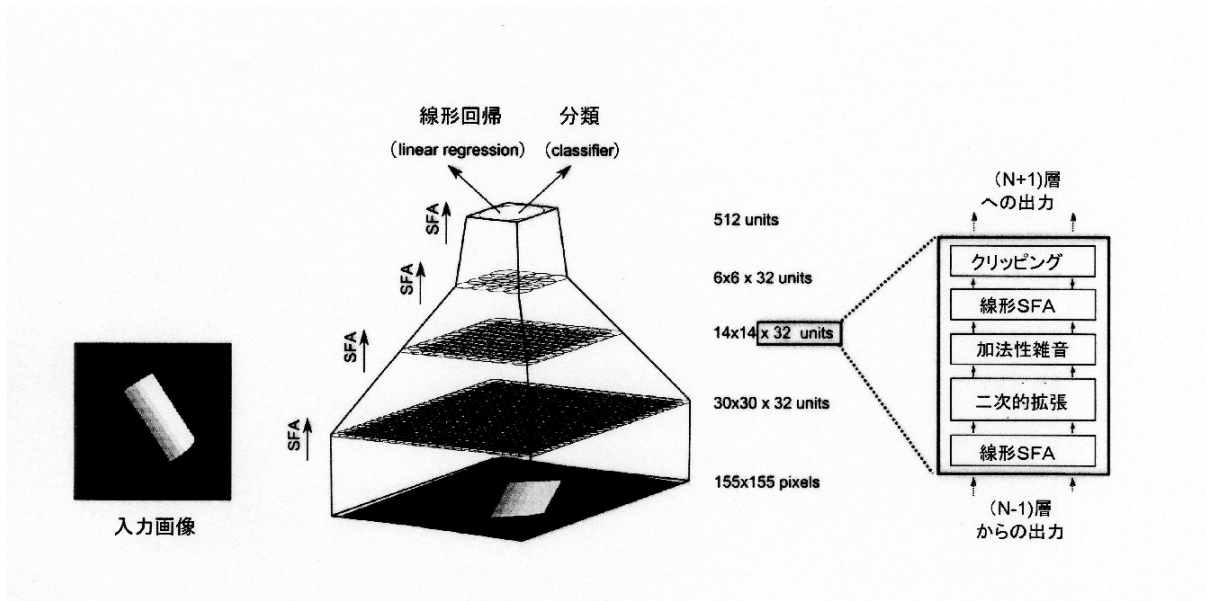
Raytchev B et al: "Unsupervised recognition of multi-view face sequences based on pairwise clustering with attraction and repulsion" 1 July 2003 (2003-07-01), Computer Vision and Image Understanding, Academic Press.

50

【図 1】



【図 2】



フロントページの続き

(72)発明者 ハイコ・ヴェージング

ドイツ国60318フランクフルト、ジセルシュトラッセ 9

審査官 鹿野 博嗣

(56)参考文献 特開2003-346152 (JP, A)

特開平08-153198 (JP, A)

特開平05-073663 (JP, A)

M.Franzius, Invariant Object Recognition with Slow Feature Analysis, ICANN '08 Proceedings of the 18th international conference on Artificial Neural Networks, 米国, 2008年, p.961-970

(58)調査した分野(Int.Cl., DB名)

G06T 1/00

B25J 13/08

G01B 11/00