

機関番号：15401

研究種目：若手研究（B）

研究期間：2008～2010

課題番号：20700163

研究課題名（和文） 学習に基づく姿勢推定における画像変化の解析と3自由度への拡張

研究課題名（英文） Analysis of appearance change in
example-based pose estimation and extension to 3DOF

研究代表者

玉木 徹 (TAMAKI TORU)

広島大学・大学院工学研究院・准教授

研究者番号：10333494

研究成果の概要（和文）：

本研究はカメラを用いた物体の姿勢推定（認識）手法の開発を目的とし、学習を行う姿勢推定において、線形の範囲内での理論的な発展と、新たな手法の開発を目指した。まず、1自由度の回転姿勢に対する姿勢パラメータと画像ベクトルとの関係を実験的に明らかにした。次に、自由度の回転姿勢を扱うための最適な姿勢パラメータと、理論的な考察を行い、球関数を用いた新しい姿勢表現と、回転行列の高周波情報を利用する手法を提案した。

研究成果の概要（英文）：

In this research project we aimed for an example-based pose estimation from appearance of objects taken by a camera, and developed a theoretical aspect for linear methods and proposed a new estimation method. First, we have investigated experimentally the relation between the 1DOF pose estimation of a rotating object and its appearance. Second, we have proposed a new pose representation of 3DOF rotation by using spherical functions, and also a new concept of the use of higher order rotation matrices.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	900,000	270,000	1,170,000
2009年度	1,600,000	480,000	2,080,000
2010年度	800,000	240,000	1,040,000
年度			
年度			
総計	3,300,000	990,000	4,290,000

研究分野：コンピュータビジョン，画像認識

科研費の分科・細目：情報学，知覚情報処理・知能ロボティクス

キーワード：姿勢推定，画像認識，球関数，学習に基づく手法

1. 研究開始当初の背景

本研究はカメラを用いた物体の姿勢推定（認識）手法の開発を目的とする。「物体の姿勢推定」とは、ある物体がカメラから見てどの位置にあり、どの方向を向いているのかを認

識することである。これは、近年の大規模工場における産業用ロボットには必須の機能である。ロボットは、アームをどのように伸ばし、どのように物体をつかめばよいのかを、瞬時に計算しなければならない。また自律移動ロボットは、障害物を検知し除去するため

に、その位置姿勢を認識することが求められる。以上のように、姿勢推定はさまざまな方面に応用されている重要な基礎技術である。本研究は、従来手法よりも優れ、理論的にも新しい姿勢推定手法を研究する。

2. 研究の目的

姿勢推定手法は、形状モデルを用いた手法と学習を行う手法に分けられる。形状モデルを用いる手法は、入力画像と3次元形状モデルを照合する方法である。一方学習を行う手法は、推定する姿勢値とそのときの画像ベクトルをセットにして学習しておく手法である。本研究は学習を行う姿勢推定において、線形の範囲内での理論的な発展と、新たな手法の開発を目的とする。

3. 研究の方法

本研究では、学習に基づく姿勢推定手法に対して、次の2つのアプローチをとる。1つ目は、1自由度の回転姿勢に対する姿勢パラメータと画像ベクトルとの関係を明らかにする。そのために、学習画像枚数と姿勢推定精度についての関連性などについて調査する。2つ目は、3自由度の回転姿勢を扱うための最適な姿勢パラメータと、理論的な考察を行う。

4. 研究成果

(1) 1 自由度の回転姿勢に対する姿勢パラメータと画像ベクトルとの関係

画像の学習による姿勢推定手法の精度は、学習サンプル数に依存している。よって、推定する際に重要となってくることは、どのぐらいの枚数を学習サンプルとするかである。学習サンプルが多いほど精度が良くなると考えられるが、様々な姿勢を推定するとなると、膨大な学習サンプルが必要である。そのため、なるべく学習サンプルは少なくしたい。しかし、どのぐらいまで学習サンプルを減らして、精度良く姿勢推定できるかはまだわかっていない。よって、学習サンプル数と推定精度の関係を明らかにし、学習サンプル数の基準を作る。

推定精度の関係を明らかにするために、今回、学習サンプル数を変化させて精度検討をした。その結果、物体により精度変化が異なり、物体によっては学習サンプル数が多い場合より、少ない場合の方が精度が良くなった。その原因を探るため学習サンプル数をより詳細に検討、また画像の相関の検討を行った。その結果、学習サンプル数が多い場合より、

少ない場合の方が精度が良くなった原因は、画像の相関が高い画像同士が学習されていないためであるということがわかった。

また線形回帰のための逐次更新による回帰係数算出手法を提案して画像ベクトルの次元と同等のサンプル数まで調査した結果、線形回帰によるパラメータ推定ではサンプル画像の有効画素数まで計算が破綻しないことが確認された。しかし、物体や条件によって程度は異なるが、サンプル数増加とともにある時点で推定誤差が増加に転じる現象が確認された。

(2) 3 自由度の回転姿勢を扱うための最適な姿勢パラメータと、理論的な考察

本研究が扱う主題は、学習に基づく姿勢推定手法である。この手法は、与えられた画像と姿勢パラメータの関係を学習する、一種の回帰問題である。まず、画像に写る物体の姿勢パラメータ p_j と、その姿勢のときに撮影された画像 x_j を対にしたものを学習セットとして、その間の関係 $p_j = f(x_j)$ を学習しておく。そして学習された関係 f を用いて、新たに入力された未知の画像 x に対応する姿勢 p を推定する (つまり $p = f(x)$)。

本研究では、3 自由度の回転で表現される単一物体の姿勢を表す、新しい表現方法を考案した。推定手法 (つまり f) に関する研究は多数存在するが、これまで姿勢 p に関してはまったく注意が払われていなかった。しかし後述するように、見えに基づく姿勢推定には、他の応用にはない特有の問題が存在するために、3 自由度の回転に対しては適切な姿勢パラメータを用いなければならない。このことを考慮していない従来研究はすべて、関係 f が線形・非線形のどちらであっても、姿勢の表現方法に本質的な問題を持っている。本研究においてはこのことを指摘し、続いて必要とされる姿勢パラメータについて議論し、提案する球関数を用いた姿勢表現手法が適していることを示した。

3 自由度の姿勢表現は、連続性、一対一を満たしていなければならない。しかし、従来の3 自由度の姿勢表現の多くは(オイラー角や固定角、回転軸・回転角表現、四元数)、これらの条件を満たしていない。3×3 回転行列は唯一条件を満たす。

本研究では、3 自由度の姿勢を「 $S^3(3)$ 上で定義される連続関数 $C^\infty(S^3(3))$ によって一意に表現する」ことを提案した。これは、 S^3 上の球関数 Y のうち、偶関数で最も低周波の9つの関数であり、Legendre 陪関数、Gegenbauer 陪関数、複素単位円の積で表さ

れる。

そして、線形回帰手法を用いて実画像の姿勢推定実験を行い、提案する球関数表現が回転行列と等価であり、四元数に比べて推定誤差が小さいことを示した。しかし、球関数表現 Y には、まだいくつもの問題がある。 $aY^1 + bY^2$ という計算には物理的な意味がなく、球関数同士の補間も定義できていない。四元数や行列の積に対応する、球関数同士の演算を考える必要がある。行列 RY を直交化するための正規化もまだない。数学的に厳密な導出を必要とする一意性の証明などの議論も残っている。これらは今後の研究における課題である。

また、100 物体の 3 次元回転画像データを独自に作成し、大規模な実験を行い、一般的な 3 自由度姿勢表現 (オイラー角, 回転角・回転軸表現, 四元数) よりも高い推定精度が得られることが示された。

さらに、 3×3 回転行列 R をより高精度に推定するために、高次の回転行列 R^2, R^3, \dots, R^n を用いる手法も提案し、その有効性を示した。コンピュータビジョンの主要な問題の多くが、回転行列 R と並進ベクトル t で表わされる 3 次元姿勢の推定を必要としている。それらの手法は、回転行列 R そのものを推定している。言い換えれば、その回転を特徴づける (フーリエ解析の意味で) 最も低周波な成分を用いている、と言える。そこで、もし R そのものに加えて R^2, R^3, \dots, R^n のような高次の (つまり高周波の) 情報が得られれば、回転の推定がさらに高精度に行えることが期待できる。一様乱数を加えた回転行列を用いて数値シミュレーション実験を行い、その推定誤差を角度とフロベニウスノルムを用いて評価した。実験結果から、もし観測行列がそれぞれ独立にノイズを加えられた場合には、推定誤差は減少し、提案手法は非常に有効であることが示された。見えに基づく姿勢推定に適用した結果、推定誤差が減少することが確認され、本手法が有効であることが示された。 R がノイズを含むデータから推定される応用問題において、観測行列が持つノイズの独立性という仮定が満たされるようにするにはどのように問題を定式化すればよいかということについては、個々の応用問題において今後示す必要がある。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 2 件)

1. 木村将之, 天野敏之, 玉木徹, 星田孝

憲, 奥川裕之, 梅崎太造: 「単位円特徴空間を用いた高速でコンパクトな三次元物体認識」, 電子情報通信学会論文誌 D, Vol. J93-D, No. 11, pp. 2505-2515 (2010 11). 査読有

2. Toru Tamaki, Toshiyuki Amano, Kazufumi Kaneda: "Representing images of a rotating object with cyclic permutation for view-based pose estimation", Computer Vision and Image Understanding, Vol. 113, No. 12, pp. 1210-1221 (2009 12). 査読有

[学会発表] (計 22 件)

1. Toru Tamaki, Miho Abe, Bisser Raytchev, Kazufumi Kaneda: "Softassign and EM-ICP on GPU", Proc. of UPDAS2010; The 2nd Workshop on Ultra Performance and Dependable Acceleration Systems (2010 11), Nov 17, 2010, Higashihiroshima, Hiroshima, Japan.
2. Bisser Raytchev, Yuta Kikutsugi, Toru Tamaki, Kazufumi Kaneda: "Class-Specific Low-Dimensional Representation of Local Features for Viewpoint Invariant Object Recognition," Proc. of ACCV2010; The 10th Asian Conference on Computer Vision, Vol. 3, pp.1552-1563 (2010 11), Queenstown, New Zealand, November 8-12, 2010.
3. Bisser Raytchev, Tetsuya Mino, Toru Tamaki, Kazufumi Kaneda: "View-Invariant Object Recognition with Visibility Maps," Proc. of ICPR2010; The 20th International Conference on Pattern Recognition, pp. 1040-1043 (2010 08), Istanbul Convention & Exhibition Centre, Istanbul, Turkey, 2010/8/23-26.
4. Toru Tamaki, Shunsuke Tanigawa, Yuji Ueno, Bisser Raytchev, Kazufumi Kaneda: "Scale matching of 3D point clouds by finding keyscales with spin images," Proc. of ICPR2010; The 20th International Conference on Pattern Recognition, pp. 3480-3483 (2010 08), Istanbul Convention & Exhibition Centre, Istanbul, Turkey, 2010/8/23-26.
5. Toru Tamaki, Miho Abe, Bisser Raytchev, Kazufumi Kaneda, Marcos Slomp: "CUDA-based implementations of Softassign and EM-ICP," Demonstration presented at CVPR2010; IEEE Conference on Computer Vision and Pattern Recognition, June 15-17, 2010,

- Hyatt Regency San Francisco, San Francisco, USA.
6. 奥川裕之, 玉木徹, Bisser Raytchev, 金田和文 : 「アピランススペースの物体認識と姿勢推定」, MIRU2010 画像の認識・理解シンポジウム論文集, pp. 888-893, 釧路市観光国際センター, 北海道 (2010 07), 2010 年 7 月 27 日~29 日.
 7. 菊次優太, ライチェフ ビセル, 玉木徹, 金田和文 : 「クラス情報を埋め込んだ不変特徴量による 3 次元物体認識」, MIRU2010 画像の認識・理解シンポジウム論文集, pp. 1652-1657, 釧路市観光国際センター, 北海道 (2010 07), 2010 年 7 月 27 日~29 日.
 8. 箕野徹哉, ライチェフ ビセル, 玉木徹, 金田和文 : 「Visibility Map による 3 次元物体の視点不変な検出と認識」, MIRU2010 画像の認識・理解シンポジウム論文集, pp. 1658-1662, 釧路市観光国際センター, 北海道 (2010 07), 2010 年 7 月 27 日~29 日.
 9. Toru Tamaki, Bisser Raytchev, Kazufumi Kaneda, Toshiyuki Amano : "Estimating a Rotation Matrix R by using higher-order Matrices \hat{R}^n with Application to Supervised Pose Estimation," Proc. of SPPRA 2010: The Seventh IASTED International Processing, Pattern Recognition and Applications, pp. 58-64 (2010 02). Innsbruck, Austria, 2010/February/17-19.
 10. Kengo Harada, Satoko Tanaka, Toru Tamaki, Bisser Raytchev, Kazufumi Kaneda, Toshiyuki Amano : "Comparison of 3 DOF Pose Representations for Pose Estimations," Proc. of FCV2010 ; the 16th Korea-Japan Joint Workshop on Frontiers of Computer Vision, pp. 408-413 (2010 02), Aki Grand Hotel, Hiroshima, Japan, 2010/Feb/4-5.
 11. Toru Tamaki, Toshiyuki Amano, Kazufumi Kaneda : "Yet Another Representation of $SO(3)$ by Spherical Functions for Pose Estimation," Proc. of FCV2010 ; the 16th Korea-Japan Joint Workshop on Frontiers of Computer Vision, pp. 402-407 (2010 02), Aki Grand Hotel, Hiroshima, Japan, 2010/Feb/4-5.
 12. Toshiyuki Amano, Toru Tamaki : "Linear 3-D Object Pose Estimation with Dense Sample Images --Discussions about Limitation of Parameter Estimation Ability by the Linear Regressions--", Proc. of MVA2009; IAPR Conference on Machine Vision Applications, pp. 182-186 (2009 5), Keio University, Japan, 2009/5/20-22.
 13. Toshiyuki Amano, Toru Tamaki : "An Appearance Based Fast Linear Pose Estimation", Proc. of MVA2009; IAPR Conference on Machine Vision Applications, pp. 287-290 (2009 5), Keio University, Japan, 2009/5/20-22.
 14. 田中 聡子, 原田 健吾, 玉木 徹, Bisser Raytchev, 金田 和文, 天野 敏之 : 「姿勢推定における姿勢表現の比較: 100 物体を用いた実験的評価」, 電子情報通信学会技術報告パターン認識・メディア理解研究会 PRMU2009-93 Vol.109, No. 249, pp.127-132, 広島大学, 広島(2009 10), 2009 年 10 月 22 日(木)~23 日(金).
 15. 玉木 徹 : 「姿勢推定と回転行列」, 電子情報通信学会 スマートインフォメディアシステム研究会 (SIS) 信号処理研究会 (SIP) オーディオビジュアル複合情報処理研究会 (IPSJ-AVM), 電子情報通信学会技術報告 SIP2009-48, SIS2009-23, Vol.109, No. 202, pp. 59-64, 広島大学, 広島 (2009 09), 2009 年 9 月 24 日.
 16. 玉木 徹, Bisser Raytchev, 天野 敏之, 金田 和文 : 「 \hat{R}^n は R より高精度に R を推定できるか?」, MIRU2009 画像の認識・理解シンポジウム論文集, pp. 984-991, くにびきメッセ, 島根 (2009 07), 2009 年 7 月 20 日 (月) ~7 月 22 日 (水) .
 17. 原田 健吾, 玉木 徹, Bisser Raytchev, 天野 敏之, 金田 和文 : 「線形的な 3 自由度姿勢推定のための姿勢表現の性質の実験的評価」, MIRU2009 画像の認識・理解シンポジウム論文集, pp. 576-580, くにびきメッセ, 島根 (2009 07), 2009 年 7 月 20 日 (月) ~7 月 22 日 (水) .
 18. 玉木徹, 天野敏之, 金田和文 : 「球関数による姿勢表現と姿勢推定」 MIRU2008 画像の認識・理解シンポジウム論文集, pp. 1134-1141, 軽井沢プリンスホテル, 長野 (2008 07), 2008 年 7 月 29 日 (火) ~7 月 31 日 (木) .
 19. 玉木徹, 天野敏之, 金田和文 : 「球関数による 3 自由度回転物体画像系列の展開と画像補間・姿勢推定への応用」 MIRU2008 画像の認識・理解シンポジウム論文集, pp. 1169-1174, 軽井沢プリンスホテル, 長野 (2008 07), 2008 年 7 月 29 日 (火) ~7 月 31 日 (木) .
 20. 天野敏之, 玉木徹 : 「稠密なサンプル

画像を用いた3次元物体の線形姿勢推定—線形回帰によるパラメータ推定の能力限界に関する考察—」MIRU2008 画像の認識・理解シンポジウム論文集, pp. 1128-1133, 軽井沢プリンスホテル, 長野 (2008 07), 2008年7月29日(火)～7月31日(木).

21. 奥川裕之, 原田健吾, 玉木徹, 天野敏之, 金田和文: 「画像の学習による姿勢推定手法の学習サンプル数による精度変化の検討」, 第10回IEEE広島支部学生シンポジウム(HISS)論文集CD-ROM, pp. 98-71, 広島産業会館, 広島 (2008 11), 2008年11月21日(金)～23日(日).
22. 奥川 裕之, 原田 健吾, 玉木 徹, 天野敏之, 金田 和文: 「画像の学習に基づく姿勢推定手法の学習サンプル数と精度の関係性の検討」, 平成20年度電気・情報関連学会中国支部第59回連合大会講演論文集, p. 470, 鳥取大学, 鳥取 (2008 10), 2008年10月25日.

[その他]

ホームページ等

<http://home.hiroshima-u.ac.jp/tamaki/>

6. 研究組織

(1) 研究代表者

玉木 徹 (TAMAKI TORU)

広島大学・大学院工学研究院・准教授

研究者番号: 10333494

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号: