

H-013

3次元特徴量を用いた構造表現による一般物体認識 Generic Object Recognition by 3D Feature-Based Structural Expression

堀 貴博†
Takahiro Hori岩橋 直人††
Naoto Iwahashi中野 幹生††
Mikio Nakano有木 康雄†††
Yasuo Ariki

1. はじめに

近年、色情報とデプス情報の両方を提供する高品質なセンシング技術が登場した。この技術は、物体認識、物体検出、動作追跡、ロボット視覚など、様々な分野での応用が期待されている。本稿では、この3次元特徴量であるデプス情報を用いた一般物体認識を提案する。

一般物体認識とは、制約のない実世界シーンを撮影した画像に含まれる物体を、計算機が一般的な名称で認識することを指し、ロボットビジョンや画像検索のために実現が強く求められている。従来研究には、2次元画像から局所の特徴を抽出し、ヒストグラム表現などに変換する手法がある。しかし、これらの手法は、画像全体を対象とするために、特徴点の位置情報や特徴点間の関係性、それに伴い対象とする物体の構造が失われるという問題がある。

そこで、この問題に対応するために、提案手法では、図1に示すように、デプス情報から抽出したD-SIFT特徴点を線でつなぎ、3次元グラフを作成する。このグラフにより、特徴点間に関係性を持たせ、位置情報を伴う物体の構造表現を実現する。ただし、グラフ表現は統計処理に適していないので、グラフをグラフ編集距離により、ベクトル表現に変換してクラス分類を行う。

色情報とデプス情報を統合したRGB-Dオブジェクトデータセットを用いて評価実験を行った。



図1 グラフの作成

2. 提案手法の概要

提案手法の概要を図2に示す。

まず、観測したデプス値を輝度値に変換し、デプス情報からデプスイメージを作成する。そのデプスイメージからSIFT特徴[1]を抽出する。こうして得られるデプス情報に基づくSIFTを本稿ではD-SIFTと呼ぶ。次に、3次元座標を持つD-SIFT特徴点をエッジで接続して3次元グラフを作成する。このようにして、学習画像からトレーニンググラフ、テスト画像からテストグラフを作成し、トレーニンググラフからプロトタイプグラフを選択する。

そして、各グラフとプロトタイプグラフとのグラフ編集距離(GED:Graph Edit Distance)[2]を用いて、グラフをベクトルに変換[3]し、クラス分類を行う。

以降で各処理について、それぞれ詳細に述べる。

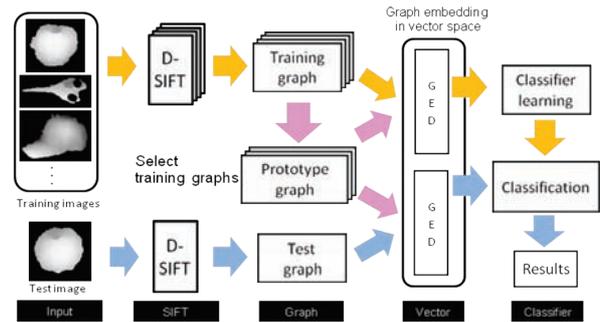


図2 提案手法の概要

3. 3次元グラフ構造表現

本研究では、[4]で提案された2次元グラフの構造表現を3次元に拡張する。各画素についてデプス値が観測できるので、抽出された特徴点に3次元座標を与え、それらをエッジで接続し、3次元グラフを作成する。また、各グラフを近接グラフに、プロトタイプグラフを擬似階層グラフに変換する。

3.1. 近接グラフ

特徴点をエッジで接続してグラフを作成するが、全ての点を相互に接続すると完全グラフとなり、通常それは計算上適切ではない。また、遠方の特徴点との関係性は薄いので、近傍の特徴点のみと接続し、近接グラフを作成する。作成された近接グラフは計算量を軽減し、検出性能を改善する。

3.2. 擬似階層グラフ

一般的に、SIFT特徴点は、スケールが大きいと高い信頼度を示す。そこで、特徴点のスケールの大きさをグラフを階層に分割し、信頼度の高い、大きいスケールの階層からグラフマッチングを開始して、徐々に階層を下げていくことにより認識性能の向上を図る。これを擬似階層グラフと定義する。また、上位の階層でマッチングしない場合には、早期に処理を終了することで、全てのノードを比較する必要がなくなり、計算時間を短縮できる。階層3に分割したグラフの例を図3に示す。

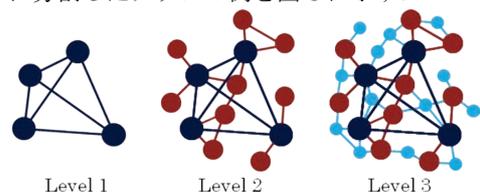


図3 擬似階層グラフ

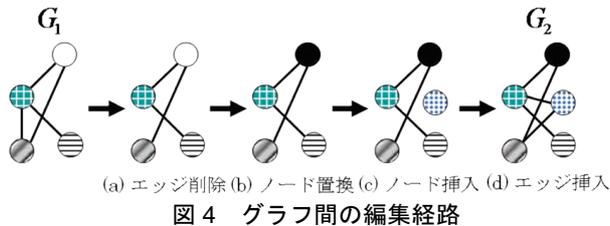
† 神戸大学大学院 システム情報学研究科, Graduate School of System Informatics, Kobe University.

†† (株) ホンダ・リサーチ・インスティテュート・ジャパン, HRI

††† 神戸大学自然科学系先端融合研究環, Organization of Advanced Science and Technology, Kobe University.

4. グラフ編集距離

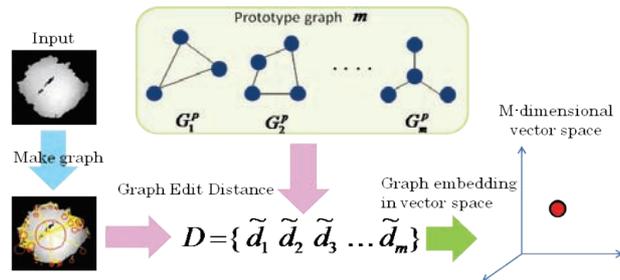
グラフ編集距離は、1つのグラフを他のグラフに変換するのに必要である最小の編集数で、2つのグラフの相違を定義する。つまり、ノードとエッジの挿入、削除、置換で構成される編集操作のそれぞれに編集コストを与え、その合計で表される。図4で2つのグラフ G_1 と G_2 の間の編集経路の例を与える。



5. グラフ-ベクトル変換

本稿で用いる変換手法は、[3]で提案された手順に従う。その概略を図5に示す。

まず、トレーニンググラフから m 個のプロトタイプグラフを選択する。次にトレーニンググラフ、テストグラフそれぞれとプロトタイプグラフとのグラフ編集距離を計算する。従って、各グラフについて m 個の相違評価値が求められ、この評価値を m 次元ベクトルとする。このように、プロトタイプグラフのセットを用いることで、全てのグラフを m 次元ベクトルに変換できる。



6. 評価実験

評価実験には、RGB-D オブジェクトデータセット[5]を用いる。このデータセットは、300物体、51カテゴリのデプス情報と色情報を持ち、デプス情報を用いて背景と対象物体の分離が行われている。撮影は、各物体をターンテーブルに乗せ、それを1メートル離れた位置から、通常のカメラとデプス情報を観測するのに用いるRGB-Dカメラにより、3つの高さから行われた。1メートル離すのは、RGB-Dカメラがデプス情報を正確に得るために、最低限必要な距離だからである。つまり、各物体につき、表面を一周撮影した3つの動画像が用意されている。このデータセットの詳細は、[6]で紹介されている。

今回実験では、RGB-D データセットの各物体につき、3つの動画像の1フレーム目のみを使用する。また、テストデータとして各カテゴリから1物体を選択し、残りを学習データとする。よって、実験に用いるデータは、51

カテゴリで画像900枚、そのうち学習画像747枚、テスト画像153枚である。

学習画像から各カテゴリ1枚を選択し、プロトタイプグラフを作成する。そのため、グラフ-ベクトル変換後のベクトルの次元は51次元となる。そうして得られたベクトルのクラス分類には、マルチクラスSVM(非線形カーネル)を用いる。表1に実験結果を示す。

表1の比較手法は、[6]において提案されている、本手法と同様にデプス情報のみを用いて、一般物体認識を行った手法の実験結果である。ただし、この手法は実験において、各物体の単一動画像から、一周分のフレームを学習データとして用いている点が本手法とは異なる。

表1 実験結果(%)

	比較手法	提案手法
認識率	66.8	26.2

7. まとめ・考察

本稿では、3次元特徴量であるデプス情報を用いて、3次元グラフ構造により物体を表現する新しい一般物体認識手法を提案した。提案手法の認識精度は現時点では比較手法よりも劣っている。これは、学習画像の枚数が比較手法に比べて著しく少ないために、認識率の低下を招いていると考えられる。今後、同数の学習画像を用いた実験を行い比較する必要がある。また、学習画像の増加に伴い、より多くの有用なプロトタイプグラフを選択し、グラフ-ベクトル変換で作成されるベクトルの次元を増やしていく予定である。次元が増えることで、より多くのグラフの特徴を表すことができ、認識率の向上が期待できる。また、3次元グラフにおいて、近接グラフと擬似階層グラフを用いることで計算時間を軽減しているが、まだ計算に膨大な時間を要している。この短縮も今後の大きな課題である。

参考文献

- [1] D. G. Lowe, "Object recognition from local scale invariant features," IEEE International Conference on Computer Vision, pp.1150-1157, 1999.
- [2] Sanfeliu A, Fu KS, "A distance measure between attributed relational graphs for pattern recognition," IEEE Transactions on Systems, Man, and Cybernetics, vol.13, no.3, pp.353-362, 1983.
- [3] Riesen K et al, "Graph Embedding in Vector Spaces by Means of Prototype Selection," in Graph-based representations in pattern recognition, ed. Escolano F et al, pp.383-393, Springer Verlag Berlin, Heidelberg, 2007.
- [4] J. Revaud et al, "Scale-Invariant Proximity Graph for Fast Probabilistic Object Recognition," CIVR, pp.412-421, 2010.
- [5] <http://www.cs.washington.edu/rgb-d-dataset/>
- [6] K. Lai, L. Bo, X. Ren, and D. Fox, "A Large-Scale Hierarchical Multi-View RGB-D Object Dataset," ICRA, 2011.