

# 付加情報を用いた教師なし転移学習手法の提案

岡本 昌也<sup>1,a)</sup> 中山 英樹<sup>1,b)</sup>

## 1. 背景

物体認識の分野では大量の教師データが必要とされる中で、データに対する手動でのラベル付けは非常に人的コストが大きい。そこで、既にラベルが付いているドメイン（照明や解像度などの環境）が異なる画像群で学習した知識を用いて、対象とするドメインを認識するという、物体認識における転移学習手法が Saenko らの手法 [7] をはじめとして多数提案され、注目を集めている。転移学習とは、異なるドメイン（照明や背景など環境の条件が同じデータの集まり）から学習した知識を活用する方法のことである。異なるドメインの画像を図 1 に示す。

一方で、ウェアラブルコンピューティングの発展や、センシングデバイスの小型化などに伴い、画像の取得と同時に、位置情報や方位情報など付加的な情報を容易に取得出来るようになってきている。特に距離情報に関しては、Kinect などの機器が普及し、容易に画像に対する付加情報として取得する事が出来るようになってきた。そこで、本論文では、距離情報をはじめとする付加情報を用いた転移学習の新たな枠組みを提案する。

## 2. 関連研究

物体認識における転移学習の問題は、Saenko らの論文 [7] が提起した。この研究は、information-theoretic metric learning (ITML) の手法 [2] を応用し、異なるドメイン間のサンプルが、同じクラス間では近く、異なるクラスのサンプル間では離れる様に、マハラノビス距離尺度を求める手

法を提案している。ターゲット側にも少数の教師データを前提とする転移学習の教師学習の手法である。

目的を考えた場合、ターゲット側に教師データを一切前提としない教師なしの手法が望ましいが、非常に難しい問題である。教師なしの手法では、ソースとターゲット間の補完的な部分空間を用いるアプローチが有望であると考えられており、その最初の研究が Gopalan らのもの [5] である。この手法は、ソース側の特徴空間からターゲット側へ明示的に部分空間を作りそれらの中間的な部分空間へ射影した点を全てつなげて特徴ベクトルとする手法である。

Gopalan らの手法を改良した、解析的に解く方法として Boqing らの Geodesic Flow Kernel (GFK) の手法 [4] がある。この手法は、明示的に部分空間を作り出すのではなく、問題を固有値問題として解析的に解く方法である。

いずれの手法もターゲット側のデータとして、画像情報のみ、もしくは、画像情報と少数の教師信号を用いる手法である。本論文では、GFK の手法を改良することで、ターゲット側の画像に対応する付加的な情報を用いる、新たな転移学習手法を提案する。

## 3. 提案手法

提案手法ではまず、ソース側の特徴空間においてラベルを教師信号として PLS [8] を行う。転移学習ではソース側には、正確なラベルが付いている前提になっているので、その正確な教師信号を手がかりに PLS を適用し、ラベルが同じサンプルが近くに配置される様な部分空間を選ぶことで、PCA を適用する場合に比べて最終的な分類精度が向上することが知られている。

次に、ターゲットドメインの距離情報から、各画像についての距離特徴を求める。本論文で使用する距離情報からの特徴は、Bo らの提案した Kernel descriptors という手法 [1] を使用して、深度画像から生成した。実験では、Bo らの配布しているコード<sup>\*1</sup>を利用した。

距離情報から与えられる特徴をクラス内の特徴にある程度の相関を期待する、画像特徴を説明する説明変数と解釈して、ターゲット側の特徴空間において PLS を適用する。つまり、ソース側と同様にある種の分布の改善を予め行う



図 1 異なるドメイン間の違いの例

<sup>1</sup> 東京大学情報理工学系研究科創造情報学専攻 〒113-0033 東京都文京区本郷 7 丁目 3 番 1 号

a) okamoto@nlab.ci.i.u-tokyo.ac.jp

b) nakayama@ci.i.u-tokyo.ac.jp

<sup>\*1</sup> <http://www.cs.washington.edu/robotics/projects/kdes/>

ことで、最終的な精度の向上を目指す。

最後に、GFK の手法を適用する転移学習を行う部分は GFK と同じ手法をとる。配布されているコードをベースとして、それを改良することで実装を行った。

この後、分類手法を用いて最終的な物体認識の結果を得る。転移学習の精度を比較する為、実験では分類手法として一番単純な最近傍法を使用した。

## 4. 実験

### 4.1 使用するデータセット

実験を行う為、ソース側のデータとして ImageNet[3]、ターゲット側のデータとして Janoch らの RGB-D 物体データセット (B3DO)[6] を使用した。B3DO は、各画像に対応する距離データを表す深度画像が、画像に加えて提供されている。実際の画像の例を図 2 に示す。B3DO に含まれる画像数が多いクラスを対象とし、対象とする物体の部分のみを、画像及び深度画像を切り出して使用した。



図 2 データセットの画像例

### 4.2 精度比較実験

4.1 節のデータセットに対して 6 クラス及び 12 クラス分類時の比較実験を行う。比較する手法は次の 4 つである。

- (1) 提案手法 1 (ソース: PCA ターゲット: PLS)
- (2) 従来手法 1 (ソース: PCA ターゲット: PCA)
- (3) 提案手法 2 (ソース: PLS ターゲット: PLS)
- (4) 従来手法 2 (ソース: PLS ターゲット: PCA)

2,4 は GFK 法を適用する。ソース側に PCA を適用する場合である提案手法 1 と従来手法 1 を比較し、PLS を適用する場合である提案手法 2 と従来手法 2 を比較する。部分空間の次元数は、それぞれの手法について最も精度が高くなるように 10,20,30,40,50 の中から選択した。各種法は異なる部分空間を生成するので、次元数を固定する事で特定の手法に有利にならない様手動で選択している。

画像特徴として SIFT 特徴の Bag of Feature(BoF) 表現を使用する。コードブックはソース側の画像のみから作成しサイズは 1000 とした。ターゲット側の画像は 1 クラス約 200 枚、ソース側の画像数を 1 クラス 100 枚とした。

#### 4.2.1 実験結果

精度比較実験を行った結果は、次の表 1 の様になった。

表 1 各手法の分類精度

手法名	6 クラス分類精度 (%)	12 クラス分類精度
提案手法 1	<b>29.49</b>	<b>18.64</b>
従来手法 1	28.60	17.99
提案手法 2	<b>32.53</b>	<b>19.66</b>
従来手法 2	28.87	19.47

結果として、6 クラス、12 クラスどちらの場合においても、ソース側とターゲット側の両方で PLS を行った提案手法 2 が最も良いという結果になった。

## 5. まとめ

本研究では、物体認識における画像とともに取得された距離情報を使用する新たな転移学習手法を提案した。また同時に、付加情報を用いた転移学習の枠組みの提案も行った。結果として、従来手法よりも提案手法の方が、物体認識精度において、優れていることを示した。実験では、一般的な従来手法に対して精度が向上することを示した。特に、物体認識の転移学習に画像情報以外の付加情報を用いる手法は我々の知る限り初めてである。

今後は、現状ではひとつの付加情報をもとに分布の改善を行っているが、実際に使用する場合、音声や位置情報など他の多くの付加情報が同時に収集できると考えられる。そういった情報を統合的に扱い、転移学習に応用していく手法を検討して行く予定である。

## 謝辞

本研究は、JST CREST の研究課題「複雑データからのディープナレッジ発見と価値化」の支援を得て実施された。

## 参考文献

- [1] Bo, L., Ren, X. and Fox, D.: Depth Kernel Descriptors for Object Recognition, *Proc. of IROS* (2011).
- [2] Davis, J. V., Kulis, B., Jain, P., Sra, S. and Dhillon, I. S.: Information-theoretic metric learning, *ICML* (2007).
- [3] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L.: ImageNet: A Large-Scale Hierarchical Image Database, *CVPR09* (2009).
- [4] Gong, B., Shi, Y. and Sha, F.: Geodesic Flow Kernel for Unsupervised Domain Adaptation, *Proc. of IEEE conference on Computer Vision and Pattern Recognition* (2012).
- [5] Gopalan, R., Li, R. and Chellappa, R.: Domain adaptation for object recognition: An unsupervised approach, *Proc. of ICCV* (2011).
- [6] Janoch, A., Karayev, S., Jia, Y., Barron, J., Fritz, M., Saenko, K. and T., D.: A Category-Level 3-D Object Dataset: Putting the Kinect to Work, *Proc. of ICCV Workshop on Consumer Depth Cameras in Computer Vision* (2011).
- [7] Saenko, K., Kulis, B., Fritz, M. and Darrell, T.: Adapting Visual Category Models to New Domains, *Proc. of ECCV* (2010).
- [8] Wold, H., Kotz, S. and Johnson, N. L.: Partial least squares, *Encyclopedia of Statistical Sciences* (1985).