

**科学研究費助成事業 研究成果報告書**

平成 28 年 6 月 9 日現在

機関番号：13901  
研究種目：挑戦的萌芽研究  
研究期間：2014～2015  
課題番号：26540081  
研究課題名(和文)ウェブ資源を用いたビジュアル・イベント学習

研究課題名(英文)Visual Event Learning with Web Resources

## 研究代表者

加藤 ジェーン (KATO, Jien)

名古屋大学・情報科学研究科・准教授

研究者番号：70251882

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：本研究では、低コストで手軽に利用可能なイベント学習枠組みの開発を目標として、データ収集の省力化・効率化を達成するためのウェブデータの収集、ウェブノイズを許容できるイベント学習、およびウェブドメインで学習した識別モデルを目標タスクに適応させるドメイン適応に関して研究開発を行った。我々は、画像認識と自然言語処理を統合した画像・映像検索手法、Approximate Sparse Coding(ASC)に基づいた独自のイベント学習法、及びイベント識別器の学習に十分なラベル付きデータがある別ドメインのデータを利用するドメイン適応技術をそれぞれ開発し、当初の目標を概ね達成した。

研究成果の概要(英文)：The objective of this research is to develop a low cost event learning framework to enable easy event learning and recognition. The proposed framework has the following three components: (1) web data collecting, which helps to prepare learning data efficiently; (2) robust event learning, which guarantees the event recognition performance; and (3) domain adaption, which helps to transform the event models learned from the web domain to the target domain. In our work, we developed (1) a flexible event retrieval approach by integrating image recognition and nature language processing; (2) an approximate sparse coding based high performance event recognition approach; and (3) a feature selection based domain transform approach for adapting event model between different domains. The proposed objectives of this research have been mostly achieved.

研究分野：画像・映像の内容理解

キーワード：イベント認識 行動認識 学習データ収集の省力化 適応学習 近似スパースコーディング 画像・映像  
検索 特徴次元選択 データ選択

## 1. 研究開始当初の背景

従来のイベント認識手法では、まず、イベント認識が必要とされる映像(テストデータ)と同じドメイン(同じ題材、撮影環境等)に属し、すでにラベリングされた学習データを、事前に大量に収集・用意し、次に、これらの学習データを用いて、イベントの構成要素であるセマンティック・コンセプトを検出するための分類器を学習させる。学習したコンセプト分類器に基づき、テストデータに対して実際のイベント認識が行われる。

しかし、従来のイベント認識の枠組みには、以下の問題がある。[1]学習データの収集、正規化、ラベリングに大量の手作業を要する、[2]学習データとテストデータが同じドメインに属することを前提として、ドメイン内イベント特徴の分散を許容できる安定性の高いコンセプト分類器を学習させるためには大量の学習データが必要であるが、同一ドメインの学習データを多数用意できない場合も多い。

これらの問題は、イベント認識において最も基本的・本質的な問題であり、イベント認識技術が有用であるにもかかわらず、広く利用されない原因となっている。このため、早急な抜本的解決が必要である。

## 2. 研究の目的

本研究の目的は、映像イベント(たとえば、サッカー試合の映像におけるゴールシーン等)を認識するための知識を、ウェブ上に散在している多くの映像・画像から獲得する新しいイベント学習機構を実現することである。従来のイベント学習は、高精度な学習のためには、同一領域(ドメイン)の学習データを大量に必要とする上、多くの手作業を要し、高コストである。これがネックとなり、実用的な用途では、イベント学習技術の採用例は非常に少ない。本研究では、ウェブを学習リソース(学習データ)として用いることを可能にし、低コストで手軽に利用可能なイベント学習機構を実現する。

## 3. 研究の方法

本研究では、ウェブ資源を用いた新たな学習の枠組みを構築することを目指す。この学習の枠組みは、

- (1) データ収集の省力化・効率化を達成するためのワンストップ型データ収集、
- (2) ウェブノイズを許容できる学習、および
- (3) ウェブドメインで学習した識別モデルを目標ドメイン(テストデータのドメイン)に適応さ

## せるドメイン適応

という3つのキーステップにより実現する。

## 4. 研究成果

### (1) データ収集の省力化・効率化

この研究項目について、我々は、①ウェブから学習サンプル(画像・映像)の収集、②収集した学習サンプルの精練という2つのステップを通して、高品質の学習データを得るアプローチを採った。

ステップ①では、画像認識と自然言語処理を統合した画像・映像検索手法を開発した。具体的には、まず、ラベル付き学習サンプルを用いて、様々な一般物体や動作の識別器を学習し、これらの識別器を、ウェブから収集した画像・映像に適用することで、画像・映像から記述子を得る。次に、自然言語処理技術により、画像・映像の記述子をワードベクトルに変換させ、さらに、ワードベクトルを用いた画像・映像に対するハイレベル検索を行い、必要な学習サンプルを迅速に収集する。本研究で、我々は、一般物体認識や動作認識に selective search, saliency analysis, non-maxima suppression, deep learning 等の技術を導入し、高精度な画像・映像検索を達成した。その研究成果により、国際会議 FCST2015 で最優秀論文賞を受賞した。提案手法により検索した画像例を図1に示す。



図1 「動物園」、「動物園+子ども」、「水泳+トランクス」というキーワードでウェブから検索した画像(上から下へ)。

ステップ②では、ウェブから収集した画像・映像を精練する処理パイプラインを開発した。このパイプラインは、連続シーンと動作が含まれるショットの検出、ショットのクラスタリング、アウト라이어・ショットの削除等の処理からなる。このパイプラインを用いて、我々は半自動的に大規模暴力シーンのデータセットを作成し、パイプライン

の有効性を確認した。作成した暴力シーンデータベースを用いて映像暴力度の自動レーティング実験を行った。良好な識別結果を得られた。

### (2) ウェブノイズを許容できる学習法

この研究項目について、当初の計画では、ガウシアンプロセス学習法を開発する予定であった。しかし、実験では、ガウシアンプロセス学習法(GP)はラベルの不確実性を学習プロセスに反映させることができるという利点があるが、予測時の計算複雑度が非常に高いという問題があることがわかった。そこで、我々は、approximate sparse coding(ASC)に基づいた独自のイベント学習法を提案した(図2)。GP と比べて、ASC はより高速でロバストに学習でき、また、より高い認識精度を達成することができる。したがって、本研究では、ASC を用いた学習法を採用した。ASC に基づいた認識手法を提案した論文は、国際会議ICIP2015でベストテンパーセント優秀論文賞を受賞した。

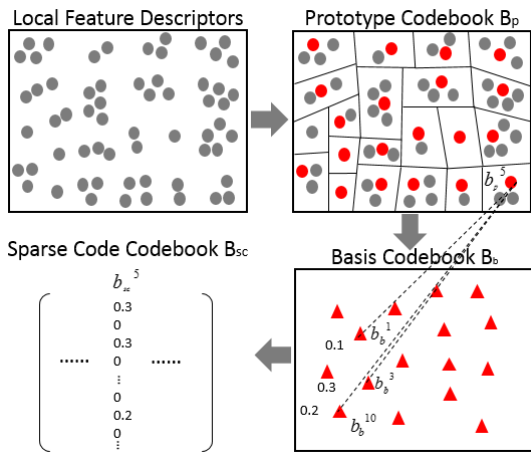


図2 Approximate Sparse Coding (ASC)手法の概要

### (3) ドメイン適応

この研究項目について、我々は、まず、目標ドメインごとに学習サンプルを用意するコストを完全に削除するために、ラベルなしデータのみからなるデータセット(目標データセット)のイベント識別器の学習に十分なラベル付きデータがある別のデータセット(元データセット)を利用するセルフトレーニングに基づいた2つのアプローチを提案し、実証実験を行った(図3)。

1つ目はフィッシャーの判断基準を用いた特徴次元選択のアプローチであり、2つ目は SVM と kNN 法を組み合わせたデータ選択のアプローチである。評価実験では、提案手法の両アプ

ローチとも安定して良い識別精度を示した。

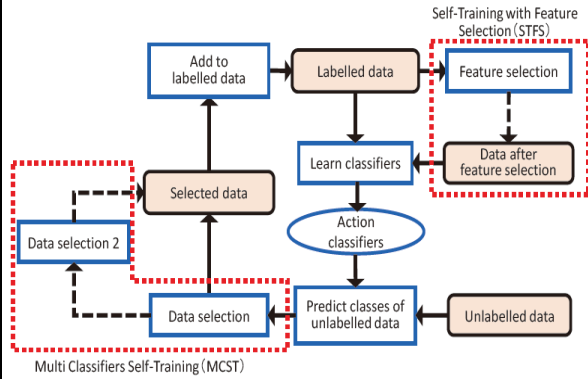


図3 提案したドメイン適応手法の枠組み

また、我々は、静止画像を対象に局所距離比較における距離メトリックを改善するための適応型メトリック学習法を開発した(図4)。即ち、既存データに含まれている物体・行動認識のための一般知識を、目標タスクのメトリック学習に利用する。適応型距離メトリックを用いると、学習サンプルの不足で起きる過学習を防ぐことができ、また、局所距離比較の性能を格段に向上させることができた。

この研究項目について当初の目標を完全に達成していないが、この研究で得た知見は、今後、ウェブドメインから目標ドメインへの適応に活用できると考える。

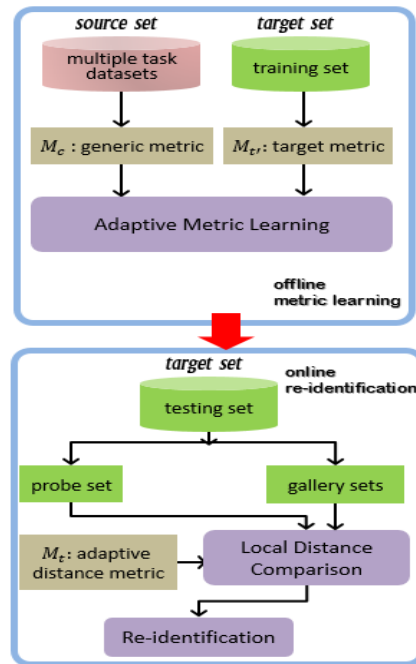


図4 適応型メトリック学習法の概要. この手法は、一般距離メトリックと目標距離メトリック間のバランスを最適化する距離メトリックを学習するものである。

## 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7件)

- ①. Yu Wang and Jien Kato, "Efficient Local Feature Encoding for Human Action Recognition with Approximate Sparse Coding", Vol.E99-D, No.4, 2016, pp.1212-1220, DOI: 10.1587/transinf.2015EDP7333 (査読有)
- ②. Guanwen Zhang, Jien Kato, Yu Wang and Kenji Mase, "Multiple-shot People Re-identification by Patch-wise Learning", IEICE Transaction on Information and Systems, Vol.E98-D, No.12, 2015, pp.2257-2270. DOI: 10.1587/transinf.2014EDP7427 (査読有)
- ③. 鈴木嵩史, ワンユ, 加藤ジェーン, 間瀬健二, 「クロスデータセットにおける行動識別器の学習アルゴリズム」, 電気学会論文誌 C, Vol.135, No.12, 2015, pp.1574-1582. DOI: 10.1587/transinf.2014EDP7427 (査読有)
- ④. Guanwen Zhang, Jien Kato, Yu Wang and Kenji Mase, "A Novel Approach for Annotation-based Image Retrieval Using Deep Architecture", Journal of Multiple-Valued Logic and Soft Computing (in press). **Best Paper Award** (査読有)
- ⑤. Guanwen Zhang, Jien Kato, Yu Wang and Kenji Mase, "Adaptive Metric Learning for People Re-identification", IEICE Transaction on Information and Systems, Vol.E97-D, No.11, 2014, pp.2888-2920. (査読有)
- ⑥. Guanwen Zhang, Jien Kato, Yu Wang and Kenji Mase, "People Re-identification with Local Distance Comparison using Learned Metric", IEICE Transaction on Information and Systems, Vol.E97-D, No.9, 2014, pp.2461-2472. (査読有)
- ⑦. Guanwen Zhang, Jien Kato, Yu Wang and Kenji Mase, "Multi-Stage Deep

Convolutional Learning for People Re-identification", International Journal of Computer Systems Science & Engineering, Vol.29, No.4, 2014, pp.243-257. (査読有)

[学会発表] (計9件)

- ①. Guanwen Zhang, Jien Kato, Yu Wang and Kenji Mase, "How to Initialize the CNN for Small Datasets: Extracting Discriminative Filters from Pre-trained Model", Proc. of The 3rd IAPR Asian Conference on Pattern Recognition (ACPR'2015), PS2-34, Kuala Lumpur, Malaysia, Nov.3-6, 2015. (査読有)
- ②. Yu Wang and Jien Kato, "Video-level Violence Rating with Rank Prediction", Proc. of The 3rd IAPR Asian Conference on Pattern Recognition (ACPR'2015), PS1-04, Kuala Lumpur, Malaysia, Nov.3-6, 2015. (査読有)
- ③. Yu Wang and Jien Kato, "Action Recognition with Approximate Sparse Coding", Proc. of IEEE International Conference on Image Processing (ICIP'2015), ARS-P1.1, Quebec City, Canada, Sept.27-30, 2015. **Best 10% Paper Award** (査読有)
- ④. 内藤貴, 王戎, 加藤ジェーン, 間瀬健二, 「時空間上暴力シーンラベリングツールの提案」, 平成 27 年度電気関係学会東海連合大会, 名古屋工業大学, M4-7, 2015.9.29. (査読無)
- ⑤. 小野瀬良祐, 王戎, 加藤ジェーン, 間瀬健二, 「チョウの種類を特定する物体認識手法の検討」, 平成 27 年度電気関係学会東海連合大会, 名古屋工業大学, M2-1, 2015.9.28. (査読無)
- ⑥. Guanwen Zhang, Jien Kato, Yu Wang and Kenji Mase, "People Re-identification Using Two-Stage Transfer Metric Learning", Proc. of IAPR International Conference on Machine Vision Applications (MVA'15), Tokyo, Japan, May 18-22, 2015. (査読有)

- ⑦. 加藤ジェーン, 王彧, 「ビデオコンテンツにおける暴力度の自動レーティング」, 電子情報通信学会技術研究報告(PRMU), 慶応大学, pp.235-238, 2015.3.19-20.  
(査読無)
- ⑧. 鈴木嵩史, 王彧, 加藤ジェーン, 間瀬健二, 「クロスデータセットに対するセルフレーニングでのデータ選択手法の提案」, 電子情報通信学会技術研究報告(PRMU), 九州大学, pp.85-89, 2014.12.11-12.  
(査読無)
- ⑨. 内藤貴, 加藤ジェーン, 王彧, 間瀬健二, 「時空間特徴を用いた YouTube 上のビデオ内の暴力シーン検出」, 第3回情報科学技術フォーラム(FIT2014), 筑波大学, pp.185-188, 2014.9.3-5.  
(査読無)

[図書](計0件)

[産業財産権]

○出願状況(計0件)

[その他]

ホームページ等

無し

## 6. 研究組織

### (1)研究代表者

加藤 ジェーン(KATO, Jien)

名古屋大学・大学院情報科学研究科・

准教授

研究者番号:70251882

### (2)研究分担者

無し

### (3)連携研究者

無し