

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 21 日現在

機関番号：12601

研究種目：挑戦的萌芽研究

研究期間：2014～2016

課題番号：26540078

研究課題名（和文）材質認識と物体認識を融合した画像認識の新展開

研究課題名（英文）Fusion of Object Recognition and Material Recognition for Smarter Image Recognition

研究代表者

山崎 俊彦（Yamasaki, Toshihiko）

東京大学・大学院情報理工学系研究科・准教授

研究者番号：70376599

交付決定額（研究期間全体）：（直接経費） 2,700,000円

研究成果の概要（和文）：これまで行ってきた研究内容は大きく分けて3つある。1つめは、機械学習器の確信度判定である。正しく確信度を判定できれば、確信度が低い場合に別の特徴抽出や機械学習の手法を導入ことによって判定結果を正しいものに書きできる。2つめは、進化計算を用いたCNNの構造・パラメータ自動最適化の研究である。CNNの構造・パラメータを進化計算による探索によって最適化した。3つめは文脈に基づく物体検出・認識性能の向上である。画像中に共起している他の物体の情報や、物体の出現位置や大きさを考慮して確信度を再計算する手法を提案した。

研究成果の概要（英文）： There are three major contributions in our research. The first contribution was the confidence analysis for multi-class object recognition using the intermediate values from machine learning algorithms. By using the confidence, it has been made possible to improve the recognition accuracy. The second contribution was the methods to automatically find optimal parameter settings for convolutional neural networks (CNNs) by using an evolutionary algorithm called particle swarm optimization (PSO). We have also developed two candidate pruning algorithms for efficient evolutionary process. The third contribution was taking contextual information into consideration such as the co-occurrence of objects and the location of objects in object detection. We developed candidate pruning and object rescoring methods that leverage contextual information and that can improve the state-of-the-art CNN-based object detection methods such as Fast R-CNN and Faster R-CNN.

研究分野：画像処理、パターン認識、機械学習

キーワード：物体認識 物体検出 確信度 文脈 特徴抽出

1. 研究開始当初の背景

申請時の研究の学術的背景を材質の認識、物体の認識、申請者の研究経緯の観点から述べる。

・**材質の認識**：物体の材質に関する研究は、高臨場感 CG 生成のために古くから研究がなされてきた。ビジョンの分野でも、画像による計測の観点から BRDF など表面反射関数のパラメータ推定、模様や凹凸を基にした材質分類の研究がなされている。ただし、BRDF を計測する場合、照明条件を変えて複数枚の画像撮影が必要など、制約が多い。それに対し、一般的な画像から材質認識を行う研究(図 1 参照)がなされつつあるが、技術としては未成熟である。

・**物体の認識**：物体認識は古くから研究が行われているが、近年 SIFT アルゴリズムに代表されるような局所画像特徴量の発展と Bag of Features (BoF) に代表されるような効率的な画像特徴表現形式の発展により、一般物体認識の性能は大きく改善されてきた。SIFT を用いた BoF 法は、様々な局所特徴表現の利用、Sparse coding 法などを用いた柔軟なコード割り当てなどによって、さらに発展を続けている。しかし、従来研究は画像の局所勾配に注目しているだけで、物体の材質にまで踏み込んだ認識はなされていない。

・**申請者の研究経緯**：これまで Spatial Soft Voting 法という特徴量の空間的分布を確率的に扱う手法、機械学習器が出力する確信度を処理して認識精度を向上させる手法などの物体認識技術を研究してきた。成果の一部は国際会議 ICMLA で論文賞を受賞している。これらの研究の課程で材質・質感の認識が物体認識の性能向上に重要であることに思い至り、本研究提案に至っている。

2. 研究の目的

本研究の目的は、これまで独立に行われてきた「材質認識」と「物体認識(形状認識)」を融合して処理し、互いのデータを補助的に用いることでそれぞれのタスクの認識率を向上させること、さらに材質・形状を手がかりとしたさらに細かな画像分類を可能とすることであった。しかし、深層学習など新しい手法の登場により両者の垣根がなくなり同等の問題として扱われるようになってきた。そのため、下記のように研究の力点を変え、材質認識、物体認識両方に一般的に利用可能な手法について研究を進めた。

3. 研究の方法

関連の研究者らが当研究の成果を再現できるように、標準的なデータセット (Caltech-101, Caltech-256, Caltech-UCSD



図 1 Sharan ら [Sharan, IJCV13] が用いた材質認識用画像。



図 2 これまでの物体認識の問題点。確信度が負であっても最大のものが選ばれていた。多くの場合、これは不正解となる。

Bird200, CIFAR10, CIFAR100, PASCAL VOC 2007, MS-COCO)、ソースコードが入手できるライブラリ (libsvm, liblinear, Caffe, Fast-RNN, Faster-RNN) などを用いた。

4. 研究成果

研究成果は大きく分けて 3 つある。それぞれの詳細については発表文献などを参照していただくとして、本報告書ではその概要を述べるに留める。

(1) 機械学習器の確信度判定

正しく確信度を判定できれば、確信度が低い場合に別の特徴抽出や機械学習の手法を導入ことによって判定結果を正しいものにかんがりの精度で書き示すことを示した。この確信度判定について、解析的な手法と機械学習を用いる手法を提案した。これは、申請当時主流だった SVM による他クラス物体認識においては、分離超平面までの符号吹き距離が負であっても最大のものが出力され、その多くの場合が誤答であるという観測に基づいている(図 2)。

具体的には、一般物体認識において多クラス分類器が出力する各クラスへの帰属確率や各クラスの分離平面までの符号付き距離などの中間的な値を処理して認識結果に対する確信度を推定することで、正答の可能性が高い場合と誤答の可能性が高い場合を精度よく判別する手法について研究した。この確信度判定について、各クラスへの中間出力値を複数用いることで確信度判定の精度向上を実現した。中間出力値を処理する手法と

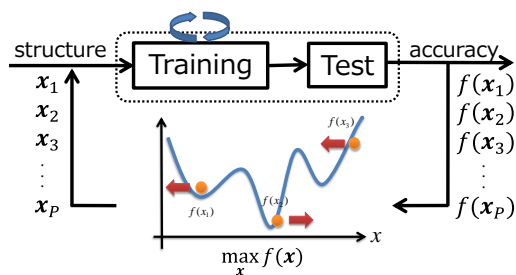


図3 進化計算を用いた CNN の構造・パラメータ自動最適化のフローチャート。

して、数値計算によるものと機械学習を用いたものという 2 つの新たな手法を提案した。また、これまで one-vs-all SVM のほか、最近注目を集めている Convolutional Neural Network (CNN) にも適用し、提案手法の一般性と有効性を確認した。

本件では 1 件の雑誌論文を発表したほか、国際会議 MMSP で Top 10% Paper Award、国内会議で IMPS2014 ベストポスター賞などを受賞した。

(2) 進化計算を用いた CNN の構造・パラメータ自動最適化

CNN に代表される深層学習は、研究者が手作業で特徴量抽出アルゴリズムを決定する必要がない画期的な手法とされているが、構造・パラメータについては研究者がやはり手作業で決定する必要がある。これらの最適化は探索空間が広く手作業による最適化は非常に困難であるため、進化計算による自動最適化について研究した。特に粒子群最適化法による最適化を行った。提案手法の概念図を図3に示す。また、単純な方法では膨大な探索時間がかかってしまうため、効率的にパラメータ最適化を行うための枝刈り手法についても併せて研究した。

本件では国内会議 MIRU にてインタラクティブ発表賞を受賞し、TOYOTA・博報堂と共同でシートベルト着用の有無を認識して交通安全意識を高めるためのスマートフォンアプリ「Seat Selfie」を開発・公開するに至った。

(3) 文脈に基づく物体検出・認識性能の向上

画像中に共起している他の物体の情報や、物体の出現位置や大きさを考慮して確信度を再計算する手法を提案した。具体的には、図3に示した通り、MRF で共起情報をモデル化し、絶対にはありえない組み合わせを除去したあと、共起する物体の種類、位置、大きさなどの文脈情報を再度 SVM にて判定し、物体検出結果のランキングする手法を提案した。これにより、深層学習を用いた最新の物体検出手法である Fast R-CNN においても VOC2007 データセットを用いた場合 mAP が 66.9%から 67.3%へ、F1 が 3.5%から 26.2%へ向上することを示した。

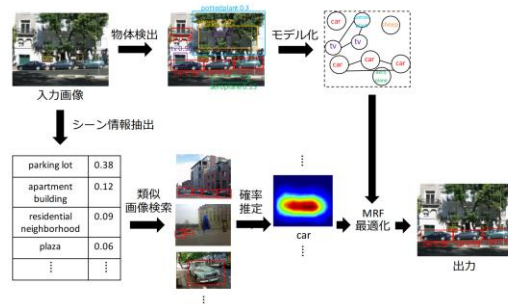


図4 文脈を考慮した画像認識・認識。

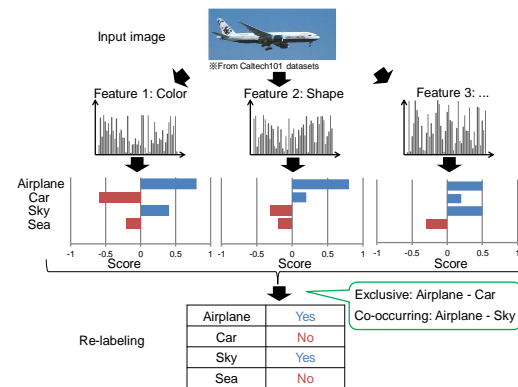


図5 マルチラベル・特徴融合同時最適化

また、マルチラベル、複数の特徴量を用いた認識にもこの文脈判別の手法を取り入れ、MRF による同時最適化を実現した。これにより、それぞれを別々に最適化する場合に比べて認識精度を向上扱せられることを確認した。

本件では、国内会議で MVE 賞、IMPS 優秀論文フロンティア賞、IE 賞など多数の賞を受賞した。

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 1 件)

[1] 山崎俊彦, 大島 辰之輔, 相澤清晴, “各クラスへの中間出力値を用いた多クラス認識のための確信度処理 (Confidence Analysis for Multi-Class Object Recognition using the Intermediate Values from Machine Learning Algorithms),” 映像情報メディア学会誌, vol. 69, no. 8, pp. J257-J260, 2015. <http://doi.org/10.3169/itej.69.J257> (査読有)

〔学会発表〕(計 29 件)

(査読付き国際会議)

[1] T. Yamasaki, T. Honma, and K. Aizawa, “Efficient Optimization of Convolutional Neural Networks using Particle Swarm Optimization,” Proc. IEEE International

Conference on Multimedia Big Data (BigMM), short paper, pp. 70-73, Laguna Hills, California, USA, Apr. 19-21, 2017.

[2] N. Inoue, R. Furuta, T. Yamasaki, and K. Aizawa, "Object Detection Refinement using Markov Random Field Based Pruning and Learning Based Rescoring," Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017), pp. 1652-1656, New Orleans, USA, March 5-9, 2017.

[3] R. Furuta, Y. Fukushima, T. Yamasaki, and K. Aizawa, "Multi-Label Classification using Class Relations Based on Higher-Order MRF Optimization," BigVision Workshop in conjunction with CVPR 2016, Las Vegas, USA, June 26 – July 1, 2016.

[4] T. Yamasaki, Y. Fukushima, R. Furuta, and K. Aizawa, "Towards Online Impression Prediction of Oral Presentations using Soft Coding," Proc. 1st workshop on Attractiveness Computing in Multimedia (ACM) in conjunction with IEEE International Conference on Multimedia Big Data (BigMM), pp. 462-465, Taipei, Taiwan, Apr. 20-22, 2016.

[5] T. Yamasaki, Y. Fukushima, R. Furuta, and K. Aizawa, "Towards Online Impression Prediction of Oral Presentations using Soft Coding," Proc. 1st workshop on Attractiveness Computing in Multimedia (ACM) in conjunction with IEEE International Conference on Multimedia Big Data (BigMM), pp. 462-465, Taipei, Taiwan, Apr. 20-22, 2016.

[6] T. Yamasaki, Y. Fukushima, R. Furuta, L. Sun, K. Aizawa, and D. Bollegala, "Prediction of User Ratings of Oral Presentations using Label Relations," ACM-MM Workshop on Affect and Sentiment in Multimedia (ACMMM ASM), pp. 33-38, Brisbane, Australia, Nov. 30th, 2015.

[7] T. Yamasaki, "Future Design and User Navigation Using Big Multimedia Data," The 1st International Conference on Advanced Imaging (ICAI), Tokyo, June 17-19, 2015. (invited)

[8] T. Maeda, T. Yamasaki, and K. Aizawa, "Multi-stage object classification featuring confidence analysis of classifier and inclined local naive Bayes nearest neighbor," Proc. International Conference of Image Processing (ICIP), pp. 5177-5181, Paris, France, Oct 27-30, 2014.

[9] T. Yamasaki, T. Maeda, and K. Aizawa, "SVM is not always confident: telling whether the output from SVM is true or false by analyzing its confidence values,"

Proc. IEEE International Workshop on Multimedia Signal Processing (MMSP), USB Proceedings #48, Jakarta, Indonesia, Sep 22-24, 2014.

(査読なし研究会・大会)

[1] 山崎俊彦, "Support Vector Machine を使い倒す," FIT2016 イベント企画, ここから始める情報処理 ~画像、音声、テキスト、検索、学習、一気にまとめてチュートリアル~, FIT2016 (第 16 回情報科学技術フォーラム), 富山大学五幅キャンパス, 富山県富山市, Sep. 7-9 2016.

[2] 井上直人, 古田諒佑, 山崎俊彦, 相澤清晴, "文脈・状況を考慮した物体検出の高精度化 (Object Detection Refinement Using Markov Random Field Based Pruning and Learning Based Rescoring)," 電子情報通信学会, パターン認識・メディア理解研究会 (PRMU), 名城大学 天白キャンパス, 愛知県名古屋, Mar. 20-21, 2017

[3] N. Inoue, R. Furuta, T. Yamasaki, and K. Aizawa, "Markov Random Field Based Pruning and Learning Based Rescoring for Object Detection," MIRU2016 (第 19 回画像の認識・理解シンポジウム), PS1-01, アクティシティ浜松, 静岡県浜松市, Aug. 1-4, 2016.

[4] T. Yamasaki, T. Homma, and T. Aizawa, "Efficient Optimization of Convolutional Neural Networks using Particle Swarm Optimization," MIRU2016 (第 19 回画像の認識・理解シンポジウム), PS2-11, アクティシティ浜松, 静岡県浜松市, Aug. 1-4, 2016.

[5] 井上直人, 古田諒佑, 山崎俊彦, 相澤清晴, "類似シーン画像を用いた物体検出のフィルタリング (Filtering Object Detection Results Featuring Statistical Information of Objects in Similar Scenes)," 情報処理学会第 78 回全国大会, 2N-01, 慶応大学, Tokyo. Mar 10-12, 2016.

[6] 井上直人, 古田諒佑, 山崎俊彦, 相澤清晴, "クラス間の共起情報を用いた物体検出結果のフィルタリング (Re-evaluation of object detection results by considering co-occurrence between object classes)," 映像情報メディア学会冬期大会, 14C-6, 早稲田大学, Tokyo. Dec. 15-16, 2015.

[7] 邱盟竣, 山崎俊彦, 相澤清晴, "A Fast Table-Based Approach of Bag-of-Features for Large-Scale Image Classification," 映像情報メディア学会年次大会, 24A-1, 東京理科大学葛飾キャンパス, 東京, Aug. 26-28, 2015.

[8] 本間拓人, 山崎俊彦, 相澤清晴, 川崎順平, 茶谷亮裕, 進藤俊彦, "カスケード型識別器と進化計算による自動最適化 CNN を用いたハイブリッド画像認識 (A Hybrid Image Classifier using a Cascade Classifier and a PSO-Optimized CNN)," 電子情報通信学会, 信学技報, 画像工学研究会, IE2015-117, pp.

235-240, 北海道大学, Hokkaido, Feb. 22-23, 2016.

[9] 本間拓人, 山崎俊彦, 相澤清晴, “畳み込みニューラルネットワークの最適化に関する探索の効率化 (Efficiency and Computation Cost of Searching Optimal CNN Structures),” 2015 年度映像メディア処理シンポジウム (IMPS 2015), I-4-14, ラフォーレ修善寺, Sizuoka, Shuzenji, Nov. 18-20, 2015.

[10] 古田諒佑, 福島悠介, 山崎俊彦, 相澤清晴, “高階エネルギーの MRF 最適化によるラベル共起を考慮したマルチラベル分類 (Multi-Label Classification using Label Relations Based on Higher-Order MRF Optimization),” 2015 年度映像メディア処理シンポジウム (IMPS 2015), I-4-13, ラフォーレ修善寺, Sizuoka, Shuzenji, Nov. 18-20, 2015.

[11] 山崎俊彦, 古田諒佑, 福島悠介, 相澤清晴, “文書・音声特徴によるリアルタイム・プレゼンテーション解析にむけた検討 (Towards Real-Time Impression Prediction of Oral Presentations using Document and Acoustic Features),” 2015 年度映像メディア処理シンポジウム (IMPS 2015), I-3-10, ラフォーレ修善寺, Sizuoka, Shuzenji, Nov. 18-20, 2015.

[12] Ryosuke Furuta, Yusuke Fukushima, Toshihiko Yamasaki, Kiyoharu Aizawa, “MRF-Based Multi-Label Classification using Label Relations,” 情報処理学会・第 198 回コンピュータビジョンとイメージメディア研究発表会 (CVIM), Vol. 2015-CVIM-198, No. 16, 愛媛大学, Sep. 14-15, 2015.

[13] Meng-Jiun Chiou, Toshihiko Yamasaki, Kiyoharu Aizawa, “A Fast Method of Visual Words Assignment of Bag-of-Features for Object Recognition,” The 18th Meeting on Image Recognition and Understanding (MIRU), SS4-40, 阪急エキスポパーク, Osaka, 27-30 Jul., 2015.

[14] 本間 拓人, 山崎俊彦, 相澤 清晴, “粒子群最適化を用いた深層学習のアーキテクチャ自動最適化,” The 18th Meeting on Image Recognition and Understanding (MIRU), SS2-37, 阪急エキスポパーク, Osaka, 27-30 Jul., 2015.

[15] 大島辰之輔, 山崎俊彦, 相澤清晴, “多クラス物体認識のための学習を用いた確信度処理 (Learning-Based Confidence Analysis for Multi-Class Object Recognition),” 映像情報メディア学会冬期大会, 10-8, 東京理科大学森戸記念館, Tokyo, Dec. 17-18, 2014.

[16] 山崎俊彦, 前田嵩樹, 相澤清晴, “Parametric meets inclined local NBNN: 確信度と NBNN を用いた BoF ベース画像認識の精度向上 (Parametric meets inclined local NBNN: better image recognition for

BoF-based approaches by the combination of confidence analysis and NBNN),” 映像情報メディア学会年次大会, 2-1, 大阪大学, 大阪, Aug 31st – Sep. 2, 2014.

[17] 伊藤慧, 相澤清晴, 山崎俊彦, “画像パッチ検索を利用した Bag-of-Features の高速化 (Fast Bag-of-Feature Generation Featuring Approximate Nearest Image-Patch Search),” 電子情報通信学会マルチメディア・仮想環境基礎研究会 (MVE), MVE-2014-101, pp. 125-126, 成蹊大学, 東京, Mar. 3-4, 2015.

[18] 大島辰之輔, 山崎俊彦, 相澤清晴, “多クラス物体認識のための学習を用いた確信度推定 (Learning-Based Confidence Analysis for Multi-Class Object Recognition),” 電子情報通信学会マルチメディア・仮想環境基礎研究会 (MVE), MVE-2014-100, pp. 123-124, 成蹊大学, 東京, Mar. 3-4, 2015.

[19] 山崎俊彦, “One-vs-All SVM による多クラス画像認識における認識結果に対する「確信度」解析 (Confidence Analysis for One-vs-All SVM-Based Multiclass Object Recognition),” 2014 年度映像メディア処理シンポジウム (IMPS 2014), I-3-16, ラフォーレ修善寺, Sizuoka, Shuzenji, Nov. 12-14, 2014.

[20] T. Yamasaki, T. Maeda, and K. Aizawa, “Multi-stage object classification featuring confidence analysis of classifier and inclined local naive bayes nearest neighbor,” Meeting on Image Recognition and Understanding (MIRU2014), SS1-5, 岡山コンベンションセンター, Okayama, July 28-31, 2014.

〔図書〕 (計 0 件)

〔産業財産権〕

○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕

ホームページ等

<https://www.hal.t.u-tokyo.ac.jp/~yamasaki/project/index-e.html>

6. 研究組織

(1) 研究代表者

山崎 俊彦 (Toshihiko YAMASAKI)

東京大学・大学院情報理工学系研究科・准教授
研究者番号: 70376599

(2) 研究分担者

(3) 連携研究者

(4) 研究協力者

相澤 清晴 (Kiyoharu AIZAWA)