

平成 30 年 6 月 12 日現在

機関番号：32689

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00248

研究課題名(和文) ビジュアルビッグデータの高速画像検索・認識に関する研究

研究課題名(英文) Study on Fast Image Retrieval and Recognition Using Visual Big Data

研究代表者

鎌田 清一郎 (Kamata, Seiichiro)

早稲田大学・理工学術院(情報生産システム研究科・センター)・教授

研究者番号：00204602

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：ビジュアルメディアのビッグデータ(以下、ビジュアルビッグデータとする)を利活用し、画像の検索および認識に関する研究において、できるだけ計算量を削減するための情報理論における圧縮可能性について再検討し、高精度かつ高速な画像検索および認識の方法論を確立した。特に近年注目をされているディープラーニングとの融合により、顔のビジュアルビッグデータを基にしてスパースグラフニューラルネットワークという新たな研究領域を構築することができた。

研究成果の概要(英文)：Big data from visual media, which is called visual big data, was utilized in this study, and new methodologies in image retrieval and recognition were established based on a concept of compressibility which is focusing on reduction of computational complexity in information theory. Especially combining with state-of-the-art deep learning, a new research direction of sparse graph neural network was developed using face visual big data.

研究分野：知覚情報処理

キーワード：画像検索 画像認識 スパースグラフ ニューラルネットワーク ビジュアルビッグデータ ディープラーニング 圧縮可能性 基底関数分解

1. 研究開始当初の背景

安心安全な社会を実現するために、ビジュアルメディアのビッグデータ(以下、ビジュアルビッグデータとする)を利活用した画像や映像の圧縮伝送・蓄積そして検索や認識に関する研究開発は、社会的なニーズとともに益々重要となっている。本研究を開始する前には、ビデオ遠隔監視などを目的として、国際標準である画像圧縮方式(例えば MPEG4 AVC/H.264 AVC など)とは全く異なるヒルベルト曲線を利用した新たな画像圧縮・検索方式を開発してきた。本方式は、実利用テストの結果、国際標準方式に比べて低ビットレートでの圧縮効率に優れ、人物などの輪郭情報が保存されやすいことがわかった。また、本方式は、検索したい人物等の画像を、圧縮された蓄積データの中からデコードせずに、前述の国際標準方式では実現できない、高速検索できることがわかった。このような背景で、ヒルベルト走査型 Bag-Of-Feature (HBOF) 検索方式を発展させることができた。ここでヒルベルト曲線は空間充填曲線の一種であるが、階層性を有する4分木表現によって生成できるため、近傍保存性のよい空間表現ができるという利点がある。

2. 研究の目的

数十年前から大規模画像検索(Content-Based Large-scale Image Retrieval, CBLIR)は画像間の類似性を利用した方法論が数多く提案されている。本研究は、ビジュアルビッグデータを対象とし、情報理論における圧縮可能性という概念からこれまでの方法論を見直し、計算量の大幅削減と検索・認識性能向上について再検討し、高精度かつ高速な画像検索や画像認識の方法論の確立を行う。まず、ビジュアルビッグデータを活用して圧縮可能性の観点から様々な空間充填曲線や基底関数などによる画像内のオブジェクトの空間的位置情報をできるだけ保存した高速画像検索・認識法の確立を第1の目的とする。次に、近年高精度オブジェクト認識を達成しているディープラーニング(Deep Learning, DL)による機械学習を導入し、ビジュアルビッグデータに対して圧縮可能性の観点から更なる高精度化かつ高速化の実現を第2の目的とする。

3. 研究の方法

本研究を遂行するにあたっての具体的に明らかにすべき内容は次の通りである。

(1) 前述のように高速画像検索において HBOF を考案し、さらにハッシュ関数を利用した Signature ベースの高速検索方式を考案したが、本研究ではまず最適な基底曲線系、基底関数系を導入して、本検索方式の更なる高速化を実現する。例えば図1のように曲線群はいろいろと考えられるため、様々な空間充填曲線などの中でどのような曲線系が高速検索に適しているのか、などを検討し、最

適な基底曲線系、基底関数系を明らかにする。また探索対象のオブジェクトの PATCH あるいは Parts ベースの特徴記述子を再検討することにより、更に高効率化、高速化できるものと考えている。さらに、基底曲線系、基底関数系から空間情報が抽出可能となるので、検索効率のよいデータ表現の構築ができる。ここで、いかに効率よくオブジェクトの空間情報が抽出でき、どのようなデータ構造で表現できるか、などを明らかにする。また、第2の目的を達成するために DL を導入し、画像検索性能の向上を図る。最終的にビジュアルビッグデータを用い、従来手法として Spatial Pyramid Matching から最新の手法までを取り上げ、画像検索性能の比較を行い、その検索効率を明らかにする。

(2) ビジュアルビッグデータからの高精度オブジェクト認識を行うために、DL における畳込みニューラルネットワーク(Deep Convolutional Neural Network, DCNN)による認識方式を導入する。すなわち DCNN を圧縮可能性という観点から見直し、超多層ではなく、コンパクトなネットワーク構造に構築することを検討する。ビジュアルビッグデータとして顔、人物、バイオメディカル画像などを用い、従来手法として最新の手法を取り上げ、画像認識性能の比較を行い、その認識能力を明らかにする。また DCNN 法は学習に時間がかかるという問題があり、高速学習法が求められている。画像の特徴点解析において、顕著性マップや SIFT などではガウシアンフィルタが重要な役割を果たしているが、圧縮可能性の観点から離散コサイン変換(DCT)などをベースにした基底関数系を用いて解析的に最適化された定数時間ガウシアンフィルタを検討する。さらに、これはバイラテラルフィルタやガボールフィルタ、畳込みオペレータへの適用が可能である。したがって、基本となるオペレータを高速化することができ、DCNN 全体としての高速化も実現できる。最終的にビジュアルビッグデータを用いて、DCNN の従来の学習などを用いた場合と比較して、学習効率の評価により DL オブジェクト学習の時間短縮に対していかに有効であることを明らかにする。以上において開発した各方式に対して並列計算に向けた方式を検討する。またビジュアルビッグデータを利用して大規模性能評価を行い、GPU (Graphics Processing Unit) によりどこまで高速化できるかを明らかにする。

4. 研究成果

(1) 高速画像検索

まず複数の空間充填曲線の導入を検討し、相対的位置関係を考慮した階層構造による大規模画像検索アルゴリズムを開発した。すなわち、これまでのヒルベルト走査型 Bag-Of-Feature (HBOF) の問題点を再考し、空間ピラミッド表現を導入した新たな方法論を考案した。これは、局所的な特徴を連続走査に

反映させるため、4 分木生成のときの階層的 4 分割を繰り返す上で局所領域間の関係を走査パターンによって記述するものである。大規模画像データベース (Caltech256 など) を用いた画像検索実験を行った結果、従来手法である HBOF などに対して、mean Average Precision (mAP) 指標などでの精度評価を行い、12% 程度向上することがわかった。

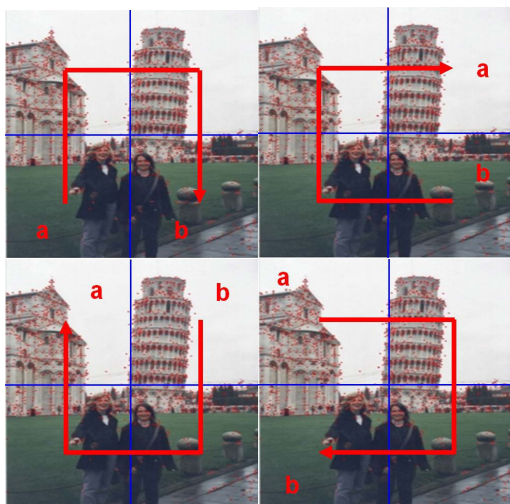


図 1 様々な曲線群による走査

近年ハッシュ関数を利用した深層学習法が高速検索として注目されている。このハッシュ法には大きく分けてデータに依存しない方法とデータ依存による方法がある。前者は、Locality-Sensitive Hashing (LSH) などに代表されるランダム写像が基本であるが、長いコードを必要とする。後者は、大規模なデータに対してコンパクトなコードを生成できるので、より効率的である。例えば Spectral Hashing (SH) や Iterative Quantization Hashing (ITQ) などが代表的である。本研究では、後者に着目し、ハミング符号を用いた教師付 Deep supervised hashing 法を確立した。これは、DCNN を用いて画像の特徴抽出を行うが、図 2 のように学習データの各特徴からカテゴリ (ラベル) 間ハミング距離最大となるターゲットハッシュ符号を生成し、この符号集合に基づき DCNN により各カテゴリ (ラベル) に対するハッシュ符号を再学習することに特徴がある。本手法と従来手法との性能を比較評価するため、公開されている MNIST および CIFAR-10 などを用いた実験結果、従来手法で有名な CNNH+ (Xia et al, AAI2014) などを超える性能を得ることができた。

次に、前述の大規模画像検索の性能向上を図るため、図 3 のように 2 段階の DCNN を利用したハッシュ関数学習法による新たな方法論を確立した。これは、予め一段目の DCNN にビジュアルビッグデータ ImageNet の特徴を学習しておき、探索に必要なラベル情報に対してハミング距離がカテゴリ内で最小、カテゴリ間で最大となるターゲットハ

ッシュ符号を生成し、この符号集合を対象として 2 段目の DCNN を学習させ、より最適なハッシュ探索を行うものである。本手法と従来手法 CNNH+ や NINH などとの性能を比較評価するため、公開されているデータセット CIFAR-10 などを用いた実験結果、従来手法を超える性能を得ることができた。また、ハッシュ関数による探索手法とは別に、距離学習法などに関して、識別ヒストグラム間距離に基づくマッチング距離学習法を検討した。

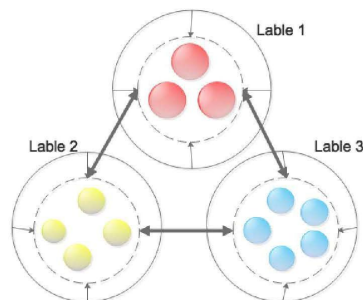


図 2 カテゴリ (ラベル) 間距離

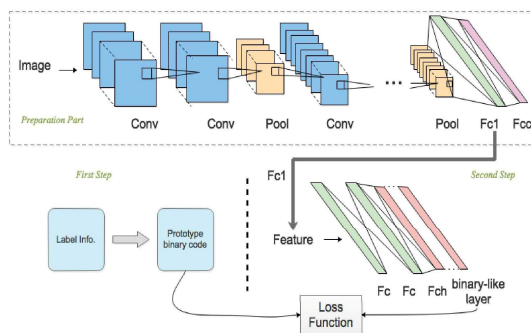


図 3 高速画像検索の概要

(2) 画像特徴記述

画像特徴記述として有名な SIFT (Scale-Invariant Feature Transform) 特徴による特徴点検出および記述に関して、圧縮可能性の観点から高速化を図った。その結果、近年注目されている従来手法 Spectral SIFT と同等のロバスト性を持ち、Spectral SIFT よりも計算コストが低い特徴点検出・特徴記述が可能であることがわかった。具体的には、スケールパラメータを変化させた際のフィルタ関数の値の変動が緩やかである点に着目し、スケール空間のフィルタ関数をスケールパラメータの M 次多項式で近似する。この近似により、1 画素あたり $O(M)$ の計算コストで特徴点検出・特徴記述ができるようになった。実験の結果、Spectral SIFT と同等のロバスト性を持ち、なおかつ約 30% の計算時間の削減を達成した。また、圧縮可能性に着目した定数時間バイラテラルフィルタ (Compressive Bilateral Filter) の設計手法について検討した。これは、図 4 のようにレンジカーネル部を空

間圧縮することにより、計算量を大幅に削減するものである。さらに、この関連研究として Discriminative and Sharable Patch (DSP)を新たに導入した画像記述子を考案し、クラスタリングを導入したオブジェクトマッチング法を提案した。公開されているデータセット MIT-67 を用いた評価実験の結果、従来手法 DeCAF などと比べて精度向上が得られた。

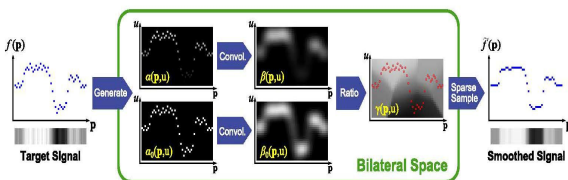


図4 圧縮可能性に着目したバイラテラルフィルタ設計

3次元ボリュームデータを対象とし、GPUに適した高速バイラテラルフィルタを設計した。3次元データに適用すると膨大な計算量が必要となるため、バイラテラルフィルタが複数の畳み込みに近似的に分解できることを利用して高速化を図った。その結果、3Dバイオメディカル画像に対して、大幅な計算量の削減が実現できた。

(3) 高速画像認識

まず顔画像認識を対象とし、新たに考案した大規模グラフ拘束スパース表現による深層学習法を開発した。当該グラフは、大量の画像に対してすべての部分画像をノードとし、ノード間のリンクを特徴記述子の類似性によって表現する大規模なものである。例えば、画像の種類にもよるが、100万枚の画像がある場合、ノード数はほぼ同数となり大規模なグラフとなる。顔画像認識シミュレーションによる比較評価実験の結果、公開されている顔画像の標準データセット LFW (Labeled Faces in the Wild) では99%を超え、YouTube Faces DBでは、膨大なYouTubeビデオのデータセットであるが、従来手法のDeepID2とほぼ同等の約92%の認識性能を得ることができた。

次に、顔認証および顔検索において、年齢による影響を軽減するために、図5のようなEigen-Aging Reference Codingという新たな年齢別顔表現法を考案した。これは、年齢ごとの固有顔を登録しておき、それらとのスパースマッチングを行うことを特徴とするものである。公開されているデータセット CACD(Cross-Age Reference Coding)を用いて、従来手法 CARC(Cross-Aging Reference Coding)と比較した結果、数%の精度向上が得られた。

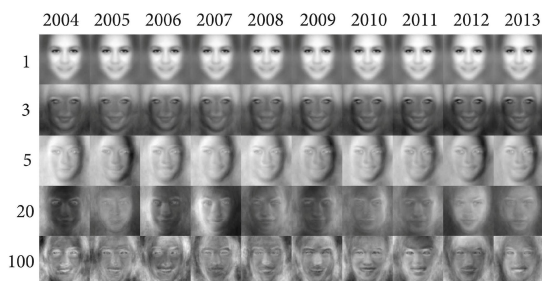


図5 Eigen-Aging Reference Codingの年齢別顔表現法

さらに、2017年9月英国マシンビジョン国際会議 (British Machine Vision Conference, BMVC)内の Irregular Domain 深層学習ワークショップにおいて、スパースグラフ表現によるグラフニューラルネットワーク (Sparse Graph Neural Network, SGNN)を用いた顔画像認識に関する研究論文を発表した。これは、図6のように画像からスパースグラフ表現により疎なグラフを構成し、近年着目されているグラフニューラルネットワーク (Graph Neural Network, GNN)との融合を図ったものである。顔画像認識の比較評価実験の結果、SGNNは、そのネットワーク構造が約10層のコンパクトなものにも拘わらず、同じような構造のDCNNに比べてより高い認識能力を有することがわかった。また標準顔画像データセット LFW や YouTube Faces などを使用した比較評価実験では、超多層DCNNのFaceNetとほぼ同等の認識精度を示すことができた。

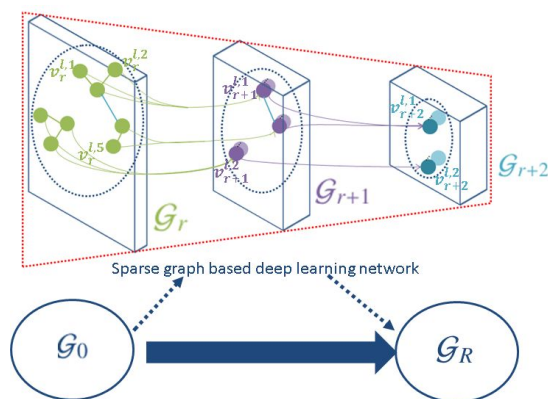


図6 SGNNの概念図

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18件)

Renju WU, Sei-ichiro KAMATA, Sparse Graph based Deep Learning Networks for Face Recognition, IEICE Transactions on Information

and Systems, vol. E101-D, no. 9, 2018, (to appear).

Xue Ting LIM, Kenjiro SUGIMOTO, Sei-ichiro KAMATA, Nuclei detection based on secant normal voting with skipping ranges in stained histopathological images, IEICE Transactions on Information and Systems, vol. E101-D, no. 2, 2018, pp. 523-530.

DOI:10.1587/transinf.2017EDP7326

Fan QIU, Sei-ichiro KAMATA, Lizhuang MA, Deep Face Recognition under Eyeglass and Scale Variation Using Extended Siamese Network, Proc. Asian Conference on Pattern Recognition, Nanjing, 2017, pp.471-476,

Xinhui WU, Sei-ichiro KAMATA and Lizhuang MA, Supervised Two-Step Hash Learning for Efficient Image Retrieval, Proc. Asian Conference on Pattern Recognition, Nanjing, 2017, pp.184-189.

Renjie WU, Sei-ichiro KAMATA, Toby BRECKON, Face Recognition via Deep Sparse Graph Neural Networks, Proceedings of Workshop on Deep Learning on Irregular Domains, in British Machine Vision Conference 2017, London, 2017, pp.1-10.

Norishige FUKUSHIMA, Kenjiro SUGIMOTO, Sei-ichiro KAMATA, Complex Coefficient Representation for IIR Bilateral Filter, Proc. IEEE international Conference on Image Processing, Beijing, 2017, pp.1-4, DOI: 10.1109/ICIP.2017.8296724.

Yongwen LAI, Sei-ichiro KAMATA, Zhizhong FU, Robust Registration of Serial Cell Microscopic Images Using 3D Hilbert Scan Search, 15th IAPR International Conference on Machine Vision Applications, Nagoya, 2017, pp. 500-503,

DOI: 10.23919/MVA.2017.7986917.

Pengyi HAO, Yang XIA, Xiaoxin LI, Sei-ichiro KAMATA, Shengyong CHEN, Discriminative Histogram Intersection Metric Learning and Its Applications, Journal of Computer Science and Technology, Springer, Vol.32, No.3, 2017, pp.507-519, DOI 10.1007/s11390-017-1740-0.

Kenjiro SUGIMOTO, Norishige FUKUSHIMA, Sei-ichiro KAMATA, Fast Bilateral Filter for Multi-channel Images via Soft- Assignment Coding, Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA2016), Jeju, 2016, pp. 1-4, DOI: 10.1109/APSIPA. 2016.7820813.

Kaihua TANG, Sei-ichiro KAMATA, Xiaonan HOU, Shouhong DING, Lizhuang MA, Eigen-Aging Reference Coding for Cross- Age Face Verification and Retrieval Proceedings of Asian Conference on Computer Vision, Taiwan, Vol.3, 2016, pp.389-403, DOI:https://doi.org/10.1007/978-3-319-54187-7_

26.

Ryo OKUTANI, Kenjiro SUGIMOTO, Sei-ichiro KAMATA, Efficient Keypoint Detection and Description Using Filter Kernel Decomposition in Scale Space, Proc. IEEE International Conference on Image Processing, Poenix, 2016, pp.31-35, DOI 10.1109/ICIP.2016.7532313.

Renjie WU and Sei-ichiro KAMATA, A Jointly Local Structured Sparse Deep Learning Network for Face Recognition, Proc. IEEE international Conference on Image Processing, Poenix, 2016, pp.3026- 3030, DOI 10.1109/ICIP. 2016.7532915.

Kenjiro SUGIMOTO, Toby BRECKON, Sei-ichiro KAMATA, Efficient Constant- Time Bilateral Filter via Eigenvalue Decomposition, Proc. IEEE international Conference on Image Processing, Poenix, 2016, pp.3319-3323, DOI: 10.1109/ICIP.2016.7532974.

Ryo OKUTANI, Kenjiro SUGIMOTO and Sei-ichiro KAMATA, Efficient Keypoint Detection and Description Using Via Polynomial Regression of Scale Space, Proc. IEEE International Conference on Acoustic, Signal Processing, Shanghai, 2016, pp.1357-1361, DOI: 10.1109/ICASSP.2016.7471898

Shoucheng NI, Qieshi ZHANG, Sei-ichiro KAMATA, Chongyang ZHANG, Learning Discriminative and Shareable Patches for Scene Classification, Proc. IEEE International Conference on Acoustic, Signal Processing, Shanghai, 2016, pp.1317-1321, DOI: 10.1109/ ICASSP.2016.7471890.

Qieshi ZHANG, Sei-ichiro KAMATA, A Novel Color Space Based on RGB Color Barycenter, Proc. IEEE International Conference on Acoustic, Signal Processing, Shanghai, 2016, pp.1601-1605, DOI: 10.1109/ICASSP.2016.7471947.

Li TIAN, Qi JIA, Sei-ichiro KAMATA, Integrating Multiple Global and Local Features by Product Sparse Coding for Image Retrieval, IEICE Transactions on Information and Systems, vol.E99-D, no.3, 2016, pp.731-738.

Kenjiro SUGIMOTO, Sei-ichiro KAMATA, Compressive Bilateral Filtering, IEEE Transactions on Image Processing, vol.24, no.11, 2015, pp. 3357-3369, DOI 10.1109/TIP.2015.2442916.

〔学会発表〕(計 7 件)

矢野光一, 杉本 憲治郎, 鎌田 清一郎: GPU 処理に適した三次元画像の高速近似バイラテラルフィルタ, 画像センシングシンポジウム, 2018.

杉本 憲治郎, 白井 啓一郎, 福嶋 慶繁, 鎌田 清一郎: レンジカーネルの対称性に基づく 効率的な定数時間バイラテラルフィル

タ,第 32 回信号処理シンポジウム, 2017.

杉本憲治郎, 鎌田清一郎: 特異値分解を用いた効率的な定数時間バイラテラルフィルタ, 画像符号化・映像メディア処理シンポジウム PCSJ/IMPS, 2017.

杉本 憲治郎, 京地 清介, 鎌田 清一郎: 離散コサイン変換に基づく定数時間ガウシアンフィルタの包括的性能解析, 電子情報通信学会技術研究報告(画像工学), 2017.

杉本憲治郎, 京地清介, 鎌田清一郎: 最適化によるモーメント保存型定数時間ガウシアンフィルタの設計, 画像符号化・映像メディア処理シンポジウム PCSJ/IMPS, 2016.

杉本 憲治郎, 福嶋慶繁, 鎌田 清一郎: 色クラスタに着目したカラー画像向けの効率的なバイラテラルフィルタ, 電子情報通信学会 技術研究報告(画像工学), 2016.

杉本 憲治郎, 鎌田 清一郎: "離散 KLT による定数時間バイラテラルフィルタ", 信号処理シンポジウム, 2015.

6. 研究組織

(1)研究代表者

鎌田 清一郎 (KAMATA Sei-ichiro)
早稲田大学理工学術院
研究者番号: 00204602

(2)研究協力者

杉本 憲治郎 (SUGIMOTO Kenjiro)
張 ケツ石 (ZHANG Qieshi)
柳 濟群 (RYU Jegoon)
郝 鵬翼 (HAO Pengyi)
吳 仁傑 (WU Renjie)
奥谷 遼 (OKUTANI Ryo)
笑 夕 (YE Xiaoxi)
賴 咏文 (LAI Yongwen)
林 雪コウ (LIM Xueting)
邱 帆 (QIU Fan)
湯 凱華 (TANG Kaihua)
ゴ 欣卉 (WU Xinhui)
倪 守誠 (NI Shoucheng)
矢野 光一 (YANO Koichi)
田 黎 (TIAN Li)
馬 利庄 (MA Lizhuang)
ブレッコン トビー (BRECKON Toby)