

平成 30 年 6 月 20 日現在

機関番号：15501

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00253

研究課題名(和文)人間認知理論と深層統計解析による画像表現と理解

研究課題名(英文) Study on Image Representation Learning and Understanding based on Human's Perception Principle and Deep Statistical Analysis

研究代表者

韓 先花 (Han, Xian-HUa)

山口大学・大学院創成科学研究科・准教授

研究者番号：60469195

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：本研究では、高信頼性且つ高レベルビジョン段階の画像表現学習を行い、高精度な画像認識・理解に関する研究を進めてきた。主な研究成果は以下の三点である。(1)人間視覚認知プロセスに従って、ロー画像に視覚応答強度ドメインを変換し、全ての情報を保持できる局所記述子を提案した。(2)混交ガウスモデルを用い、トレーニング記述子集合をモデリングし、テスト画像の記述子から高次元統計量を抽出することで、高信頼性な画像表現法を開発した。(3)Middle-levelな画像表現構造を積み重ねることで、深層フレームワークを発展させ、高信頼性且つ高レベルビジョンの画像表現フレームワークを構築した。

研究成果の概要(英文)：This study aimed at learning compact and inherent image representation for high-level vision problems, and developed advanced image recognition and understanding methods. Our main achievements are three-fold: 1) Based on human's perception principle, we transformed the raw-image domain into differential excitation domain and proposed to use the micro-texton as local descriptors for retaining all information, which would be distinguishable even for the subtle difference in image structures. 2) We proposed a novel middle-level image representation learning framework via exploring the deviation statistics of local descriptor set on the learned GMM model; 3) We stacked several layers of the middle level representation extraction framework, and proposed multiple-layer fisher network architecture for high-level feature learning. We applied our proposed image representation learning strategy for several image recognition applications, and proved much better performances can be achieved.

研究分野：知覚情報処理

キーワード：画像認識 機械学習 コンピュータビジョン パターン認識 画像処理

1. 研究開始当初の背景

デジタルカメラ等の情報端末やネットワークの普及に伴い、従来とは比較にならないほど膨大な画像情報が流通している。この膨大な数のデータの中から、利用者にとって最も適した情報のみを的確に提供し、活用できるような技術が求められている。特に画像や映像コンテンツの急速な普及は目立つ。このようなビジュアルコンテンツを効率よく活用化するため、画像の中に何が映っているかを機械で自動的に認識・理解することが強く求められている。

世界中で、画像理解・認識に関する研究が進められている。その基本的な枠組は画像や映像から、空間や時間および色などに関する特徴（画像表現）を抽出し、統計解析や機械学習で画像の中に何を映っているかを判別することである。従来の画像表現の研究では、画像から簡単に得られるカラー値、勾配値、方向や局所パッチなどのロー（Raw）特徴を、二値又は均一的に量子化（輝度レベルの削減）し、全体的にピラミッド空間でプーリングを行い、大域的な low-level 特徴量で画像表現を行ってきた。また、画像から局所記述子（SIFT など）を抽出し、特徴ベクトル同士の照合およびこの局所記述子をコーディングし、形成した局所記述子の統計量（例えば Bag of Feature : BOF）を Middle-level 画像表現とし、画像を検索・認識する手法が提案された。BOF 画像表現を用い、様々な異なる物体・シーンの種類（人間が日常的に扱っているデータで簡単に認識できる）の認識・理解に対して、ある程度良いパフォーマンスが得られた。しかし、今までの画像特徴はまずロー特徴の量子化を行い、情報を削減してからその統計量を抽出し画像を表現している。そこには特定分野で獲得したデータや物体の認識において異なるサンプルから微小な差異しか観測できず、その量子化によって観測された判別情報を失ってしまうという問題点がある。また、人間の認知理論を考慮せず観測データのロードメインから画像表現の抽出を行い、人間の視覚認知プロセスに乖離してしまうという点も問題である。

2. 研究の目的

本研究では、実世界で観測された画像情報から、高次統計解析によりロバスト且つコンパクトで本質的な情報を抽出し、高精度な画像認識・理解をすることを目的とする。観測された様々な情報（画像）からどのようにして本質的な（内部）表現を発見・学習し、情報理解をより高精度にできるかどうかパターン認識と人工知能等の知的情報処理分野に古くから研究課題の一つである。本研究では、ロバストな画像特徴表現を学習するため、人間認知理論に基づく視覚応答強度ドメインを変換し、マイクロテクストンで局所記述子を表現する。また、データ駆動モデル（混交ガウスモデル）を用いて画像の局所的なテクストンをコーディングし、低次だけで

はなく、高次（high-order）な統計量を抽出することで、より高信頼性な画像表現をすることができる。更に、データ駆動構造を積み重ねることで、深層フレームワーク（深層学習）を発展させ、よりハイレベル（high-level）な視覚表現を学習し、先端的画像理解システムの開発を目指す。

3. 研究の方法

本研究では、高信頼性且つ高レベルビジョン段階の画像表現学習を行い、高精度な画像認識・理解をすることを目的とする。主に以下の研究手法で進めてきた：

- (1) 人間視覚認知プロセスに従って、ロー画像に視覚応答強度ドメインを変換し、全ての情報を保持したマイクロテクストンを抽出することで微小な構造差異も判別できる局所記述子を提案した。
- (2) 混交ガウスモデルを用い、トレーニング記述子集合をフィッチングできるパラメータを学習し、テスト画像の記述子はそのモデルパラメータに対しての偏移統計量（低次と高次）を抽出することで、高信頼性な画像表現（中層：Middle-level 表現）法を提案した。
- (3) Middle-level な画像表現構造を積み重ねることで、深層フレームワークを発展させ、高信頼性且つ高レベルビジョンの画像表現フレームワークを構築した。これにより、理解精度を向上させ、異なる種類で微小な差異しか観測できない実用コンテンツの理解を用いて提案法の有効性を検証した。

4. 研究成果

(1) 刺激信号の強度に応じた量子化手法の開発：画像表現において局所パッチ（テクスチャ情報）の量子化によって有限な代表記述子を生成し、全体の画像に渡りその統計量（出現頻度と空間的関連構造）を抽出し大域的な画像特徴とする。従来よく用いられた手法では主に局所パッチの二値量子化（Local Binary Pattern : LBP）を行い、多大な情報の損失だけではなく、ノイズに非常に敏感という問題を生じる。本研究では、注目画素の刺激強度ベースに一定の許容度を設け、周り画素の刺激状態（ネガティブ、ポジティブ、無し：人間認知理論）に応じて三値量子化する。また、感覚に関する精神物理学の基本法則としてよく知られているウェーバーの法則によって、刺激の弁別閾（丁度可知差異：気づくことができる最小の刺激差）は、基準となる基礎刺激の強度に比例する。具体的には局所パッチの二値量子化ではなく、人間認知理論に基づく三値量子化することで新たな回転に不変なウェーバー-LTP（Local Ternary Pattern）を提案し、テクスチャ画像および HEp-2 細胞画像の Staining Pattern 認識への応用において、高精度な認識率を得ることを検証した。図 1 に提案した人間認知理論に基づく三値量子化で得られた統計特

徴と従来研究でよく用いられた手法による HEp-2 細胞画像認識の比較結果を示している。

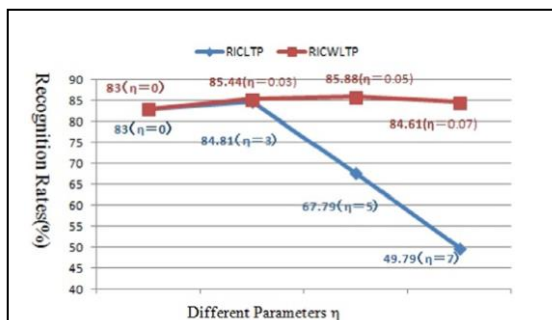


図1 提案した人間認知理論に基づく統計特徴と従来法による HEp-2 細胞画像認識の比較結果

(2) 局所パッチの量子化ではなく、人間視覚認知プロセスに従って、ロー画像に視覚応答強度ドメインを変換し、全ての情報を保持できる局所記述子の抽出法を提案した。従来の手法では、画像のロー特徴(カラー値、勾配、方向など)を二値又は均一的に量子化し、全体的に又は領域ベースプーリングすることで低次統計量を用いて画像表現する。しかし、この量子化によって判別的な情報を失ってしまうという問題点がある。そこで、本研究では視覚応答強度画像から量子化を行わず全ての情報を保持できるマイクロテクストンを直接的に局所記述子とする。提案した視覚応答強度ドメインへの変換手法は図2に示す。

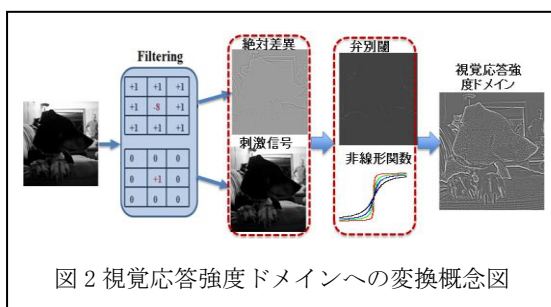


図2 視覚応答強度ドメインへの変換概念図

(3) 記述子群の高次統計量抽出法を提案した。データ駆動モデルを用いて、統計解析により、低次だけではなく高次 (High-order) な統計量を抽出することで、より高信頼性な画像表現法を開発した。その開発した手法を用いて画像から得られたテクストン記述子に対して、高信頼性な高次統計量を抽出することで、テクスチャ、食事およびバイオメディカル画像認識への応用において、高精度な結果を検証された。従来の低次統計量 (Order0) と提

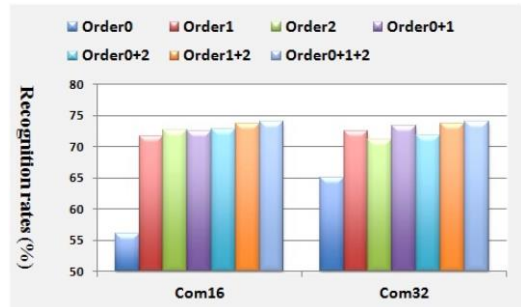


図3 低次(従来法)と高次(提案法)統計量による KTH TIPS 2a テクスチャ画像認識の比較結果

案した高次統計量 (Order1/Order2) を用いて KTH TIPS 2a テクスチャデータセットの画像認識を行い、比較結果は図3に示す。また、異なる種類でも微小な差異しか観測できない顕微鏡細胞画像に対して、熟練技師や医者の主観的な判断と同程度の認識率を得た。

「Middle-Level な画像表現の抽出法」

(4) 提案した高次統計量を学習するフレームワークを積み重ねること High-level な画像表現を抽出できる深層構造を構築した。得られた局所記述子のコードドベクトルに対し、全体画像に渡り統計量を纏めるのではなく領域ベースプーリングを行う。これにより画像領域から得られた中層 (Middle-level) の画像表現をもう一度データ駆動モデルを用いてコーディングし、グローバルやピラミッド空間でのプーリングによって高レベルな構造を学習でき、よりロバストな画像表現を得られる。本提案ではバイオメディカル画像認識に応用し、この有効性を検証した。

「High-Level な画像表現の抽出法」

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 11 件)

- 1) Yen-Wei Chen, Xinyin Huang, Dingye Chen, Xian-Hua Han, Generic and Specific Impression Estimation and Its Application to KANSEI Based Clothing Fabric Image Retrieval, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 32, No. 10 (2018) 1854024, pp. 1-18. (査読: 有)
- 2) Yingying Xu, Lanfen Lin, Hongjie Hu, Dan Wang, Yitao Liu, Jian Wang, Xian-Hua Han, Yen-Wei Chen, Texture-Specific Bag of Visual Words Model and Spatial Cone Matching Based Method for the Retrieval of Focal Liver Lesions Using Multiphase Contrast-Enhanced CT Images, International Journal of Computer Assisted Radiology and

- Surgery, Vol. 13, No. 1 (2018), pp. 151-164. (査読：有)
- 3) Chunhua Dong, Xiangyan Zeng, Lanfen Lin, Hongjie Hu, Xianhua Han, Masoud Nagedolfeizi, Dawit Aberra, and Yen-Wei Chen, An Improved Random Walker with Bayes Model for Volumetric Medical Image Segmentation, Journal of Healthcare Engineering, Volume 2017 (2017), Article ID 6506049, pp. 1-11. (査読：有)
 - 4) Xian-Hua Han, Yen-Wei Chen, Generalized Aggregation of Sparse Coded Multi-Spectra for Satellite Scene Classification, ISPRS Int. J. Geo-Inf., Vol. 6, No. 6 (2017), pp. 175-191, 2017. (査読：有)
 - 5) Xian-Hua Han, Yen-Wei Chen, Gang Xu, Co-occurrence Context of the data-driven Quantized Local Ternary Patterns for Visual Recognition, IPSJ Transactions on Computer Vision and Applications, Vol. 9, No. 4, 2017. (査読：有)
 - 6) Jian Wang, Xian-Hua Han, Yingying Xu, Lanfen Lin, Hongjie Hu, Chongwu Jin, and Yen-Wei Chen, Sparse Codebook Model of Local Structures for Retrieval of Focal Liver Lesions Using Multiphase Medical Images, International Journal of Biomedical Imaging, Volume 2017 (2017), Article ID 1413297, pp. 1-13. (査読：有)
 - 7) SiHai Yang, Xian-Hua Han, Tohsato Y, Kyoda K, Onami S, Nishikawa I, Yen-Wei Chen, Phenotype Analysis Method for Identification of Gene Functions Involved in Asymmetric Division of *Caenorhabditis elegans*, J Comput Biol., Vol. 24, No. 5, pp. 436-446, 2017 Feb 8. (査読：有)
 - 8) Xian-Hua Han, Yen-Wei Chen, HEp-2 Staining Pattern Recognition Using Stacked Fisher Network for Encoding Weber Local Descriptor, Pattern Recognition, Vol. 63, pp. 542-550, 2017. (査読：有)
 - 9) Xian-Hua Han, Yukako Tohsato, Koji Kyoda, Shuichi Onami, Ikuko Nishikawa and Yen-Wei Chen, Nuclear detection in 4D microscope images of a developing embryo using an enhanced probability map of top-ranked intensity-ordered descriptors, IPSJ Transactions on Computer Vision and Applications, Vol. 8, No. 1, pp. 1-15, 2016. (査読：有)
 - 10) Xian-Hua Han, Yen-Wei Chen, Gang Xu, Integration of Spatial and Orientation Contexts in Local Ternary Patterns for HEp-2 Cell Classification, Pattern Recognition Letter, Vol. 82, pp. 23-27, 2016. DOI: 10.1016/j.patrec.2016.02.004. (査読：有)
 - 11) Xian-Hua Han, Yen-Wei Chen and Gang Xu, High-order Statistics of Weber Local Descriptors for Image Representation, IEEE Transaction on Cybernetics, Vol. 45, Issue: 6, pp: 1180-1193, 2015. (査読：有)
- [学会発表] (計 18 件)
- 1) Xian-Hua Han, Jian Wang, Boxin Shi, Yinqiang Zheng and Yen-Wei Chen, Hyper-spectral Image Super-resolution Using Non-negative Spectral Representation with Data-guided Sparsity, The first IEEE International Workshop on Machine Learning and Computing for Visual Semantic Analysis (MLCSA2017).
 - 2) Jian Wang, Xian-Hua Han, Yingying Xu, Lanfen Lin, Hongjie Hu, Chongwu Jin, Yen-Wei Chen, Tensor sparse representation of temporal features for content-based retrieval of focal liver lesions using multi-phase medical images, The first IEEE International Workshop on Machine Learning and Computing for Visual Semantic Analysis (MLCSA2017).
 - 3) Liying Peng, Lanfen Lin, Hongjie Hu, Xiaoli

- Ling, Dan Wang, Xian-Hua Han, Yen-Wei Chen, Joint Weber-based Rotation Invariant Uniform Local Ternary Pattern for Classification of Pulmonary Emphysema in CT Images, IEEE International Conference on Image Processing (ICIP2017).
- 4) Xian-Hua Han, Jianmei Lei, Yen-Wei Chen, HEp-2 Cell Classification Using K-Support Spatial Pooling in Deep CNNs, LABELS/DLMIA@MICCAI 2016, 2016.
 - 5) Xian-Hua Han, Jian Wang, Yuu Konno, Yen-Wei Chen, Bayesian Saliency Model for Focal Liver Lesion Enhancement and Detection, ACCV Workshops (3), 2016.
 - 6) Yingying Xu, Lanfen Lin, Hongjie Hu, Dan Wang, Yitao Liu, Jian Wang, Xian-Hua Han, Yen-Wei Chen, Bag of Temporal Co-occurrence Words for Retrieval of Focal Liver Lesions Using 3D Multiphase Contrast-enhanced CT Images, 23rd International Conference on Pattern Recognition (ICPR2016), Cancun, Mexico, Dec.4-8, 2016.
 - 7) Jian Wang, Xian-Hua Han, Yingying Xu, Lanfen Lin, Hongjie Hu, Chongwu Jin, Yen-Wei Chen, Patch-based codebook model for focal liver lesions retrieval of multiphase medical volumes, Int. J. CARS, Vol.11, Suppl. S12-S13, 2016 (Heidelberg, Germany, June 21-25, 2016).
 - 8) Shota Sasano, Xian-Hua Han, Yen-Wei Chen, Food recognition by combined bags of color features and texture features, CISP-BMEI 2016, 2016.
 - 9) Yingying Xu, Lanfen Lin, Hongjie Hu, Dan Wang, Yitao Liu, Jian Wang, Xian-Hua Han, Yen-Wei Chen, A retrieval system for 3D multi-phase contrast-enhanced CT images of focal liver lesions based on combined bags of visual words and texture words, CISP-BMEI 2016, 2016.
 - 10) Xian-Hua Han, Yen-Wei Chen, Gang Xu, HEp-2 Staining Pattern Recognition Using Stacked Fisher Network for Encoding Weber Local Descriptor, Machine Learning in Medical Imaging, Volume 9352 of the series Lecture Notes in Computer Science, 2015.
 - 11) Xian-Hua Han, Yukako Tohsato, Koji Kyoda, Shuichi Onami, Ikuko Nishikawa and Yen-Wei Chen, Nuclear Detection in 4D Microscope Images of Developing Embryo Using Enhanced Probability Map of Top-ranked Intensity-ordered Descriptors, 3rd Asian Conference on Pattern Recognition (ACPR2015), Kuala Lumpur, Malaysia, Nov.3-6, 2015 (Oral Presentation).
 - 12) Xian-Hua Han, Yen-Wei Chen and Gang Xu, Co-occurrence Context of the data-driven Quantized Local Ternary Patterns for Visual Recognition, 3rd Asian Conference on Pattern Recognition (ACPR2015), Kuala Lumpur, Malaysia, Nov.3-6, 2015 (Oral Presentation).
 - 13) Misae Nakatsu, Xian-Hua Han, Ryosuke Kimura and Yen-Wei Chen, Discriminant Statistical Analysis of Local Facial Geometrical Regions, 3rd Asian Conference on Pattern Recognition (ACPR2015), Kuala Lumpur, Malaysia, Nov.3-6, 2015.
 - 14) Yingying Xu, Lanfen Lin, Hongjie Hu, Huajun Yu, Chongwu Jin, Jian Wang, Xian-Hua Han, Yen-Wei Chen, Combined Density, Texture and Shape Features of Multi-Phase Contrast-Enhanced CT Images for CBIR of Focal Liver Lesions: A Preliminary Study, Innovation in Medicine and Healthcare 2015, 2015 (Kyoto, Japan, Sep.11-12, 2015).
 - 15) Titinunt Kitrungrotsakul, Xian-Hua Han, Yen-Wei Chen, Liver Segmentation Using

Superpixel-Based Graph Cuts and Regions of Shape Constraints, IEEE International Conference on Image Processing (ICIP2015), Quebec, Canada, Sep.27-30, 2015 (Oral Presentation).

16) Qiaochu Zhao, Xian-Hua Han, Yen-Wei Chen, A Robust Registration Method using Huber ICP and Low Rank and Sparse Decomposition, Asia-Pacific Signal and Information Processing Association annual summit and conference (APSIPA ASC2015), Hongkong, Dec.15-19, 2015.

17) Titinunt Kitrungrotsakul, Yen-Wei Chen, Xian-Hua Han, Lanfen Lin, Supervoxels based Graph Cut for Medical Organ Segmentation, The 9th IFAC Symposium on Biological and Medical Systems, Berlin, Germany, Aug. 31-Sep.2, 2015. (pp.70-75).

18) Yen-Wei Chen, Dingye Chen, Xian-Hua Han, Xinyin Huang, Generic and Specific Impression Estimation of Clothing Fabric Images Based on Machine Learning, The 2015 11th International Conference on Natural Computation (ICNC'15) and the 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'15), Zhangjiajie, China, Aug.15-17, 2015.

[図書] (計 1 件)

1) Xian-Hua Han, Yen-Wei Chen, High-Order Statistics of Micro-Texton for HEp-2 Staining Pattern Classification, Artificial Intelligence in Decision Support Systems for Diagnosis in Medical Imaging (387 pages), Chapter 6, pp. 135-164, Yisong Chen and Kenji Suzuki (Eds), Springer 出版, 2018.

[産業財産権]

○出願状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況 (計 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

[その他]
ホームページ等

6. 研究組織

(1) 研究代表者

韓 先花 (Han, Xian-Hua)

山口大学・大学院創成科学研究科・准教授

研究者番号： 6 0 4 6 9 1 9 5

(2) 研究分担者

陳 延偉 (Chen, Yen-Wei)

立命館大学・情報理工学部・教授

研究者番号： 6 0 2 3 6 8 4 1