

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 8 日現在

機関番号：17301

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330200

研究課題名(和文) 組合せ最適化問題としてのパターン認識：理論構築と実用化に関する基礎研究

研究課題名(英文) Pattern recognition as combinatorial optimization: basic and applied research

研究代表者

酒井 智弥 (SAKAI, Tomoya)

長崎大学・工学研究科・准教授

研究者番号：30345003

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：認識対象を簡潔に説明する知識や概念の組合せを求める最適化問題としてパターン認識の理論を再構築し、既存技術では解決が困難なパターン認識の課題を合理的に解決する「組合せパターン認識」を具体化した。また、組合せ論的パターン認識の実用性・信頼性を実験的に明らかにするため、文字画像、肺聴診音、動画の解析へ応用した。その結果、典型的な線画とその組合せに基づく漢字検出、正常音・異常音の組合せによる肺音認識、自己運動・物体運動の組合せによる動画理解を可能にした。

研究成果の概要(英文)：Present research has established "combinatorial pattern recognition", which formulates the pattern recognition as a problem of finding an optimal combination of basic concepts or classes that concisely explains a given pattern. The combinatorial problem can be translated into convex optimization in which sparsity, low-rankness, co-occurrence of features, etc, are quantified as convex functions. In the applied research, the combinatorial pattern recognition has enabled, e.g., Chinese character detection based on typical combinations of lines, lung-sound recognition via signal separation into combinations of normal and adventitious sound components, and video understanding by combinations of egomotion and object motion.

研究分野：計算機科学

キーワード：パターン認識 機械学習 信号処理 圧縮センシング スパースモデリング

1. 研究開始当初の背景

(1) 識別問題に基づくパターン認識の問題点
爆発的に増大し続けるデータに適用できる検索、分類、識別等の技術は必要不可欠である。この要求に関して、既存のパターン認識技術を単に効率化するだけでは解決が見込めない下記のような重要な問題がある。

①あらゆるクラスの訓練パターンを用意し尽くせない。入力されたパターンにクラス(概念)を対応づける処理がパターン認識である。現在の主流である統計的パターン認識では訓練パターンから個々のクラスについて機械学習する。しかし、例えば画像による物体認識のために、あらゆる種類の物体について訓練パターンを用意することは現実的に非常に困難である。

②認識対象を複数のクラスに対応づける必要がある。認識対象に任意の数のクラスラベルを付与するパターン認識は多重ラベリングと呼ばれている。ひとつの画像に複数の物体が写っているときの物体認識や、文書の分類・タグ付け等、多重ラベリング問題に帰着されるパターン認識の課題が多く存在する。また、認識対象をクラスに対応づけてはならない場合がある。

(2) 原因と解決策

これらの問題は、認識対象をひとつのクラスに対応づける識別問題にパターン認識を帰着させていることに起因すると考えられる。まず、個々のクラスの学習が必要となるため①が避けられない。また、②の多重ラベリング問題に対しても、各クラスの単一ラベリングもしくは複数の2クラス識別問題に置き換える手法が典型的なアプローチとされてきた。個々の識別器は担当するクラスのパターンを識別できるが、複数のクラスの特徴が混合したパターンを適切に識別できる保証はない。クラスの組合せをひとつの新しいクラスとみなす手法も提案されているが、クラスの組合せの数だけ訓練パターンや識別器を用意する必要があり、クラス数に従って組合せ爆発を起こす。また、特定のクラス以外のあらゆる訓練パターン(負例)を用意し尽くすことも現実的ではない。これらの問題の原因を解消するためには、学習済みのクラスを組み合わせることで対象を説明する機能、言わば、知識を活用して知恵とするパターン認識が本質的に必要であると考えられる。

2. 研究の目的

(1) 「組合せパターン認識」の理論構築

認識対象を簡潔に説明する知識や概念の組合せを求める最適化問題としてパターン認識の理論を再構築し、既存のパターン認識技術では解決が困難な前述の問題点を合理的に解決する下記の機能の実現を目指す。

①組合せによって認識対象のパターンを簡潔に説明できるクラスを見つけ出す機械学習機能。複数の基本クラス(subclass)の組合せによって新しいクラスが構成されるのであれば、基本クラスについて訓練パターンから学習できればよい。また、概念が複雑でひとつのクラスとして扱うことが難しいとき、その概念を構成し得るより単純な基本クラスを学習できるのであれば、基本クラスの典型的な組合せとして識別できる可能性が生まれる。

②クラスの簡潔な組合せによって認識対象を説明する機能。クラスの組合せをそれぞれ機械学習するのではなく、①と同様に、パターンを簡潔に説明できる最少のクラスを求める組合せ最適化問題を解く。その解は、認識対象に対応する複数のクラスを示しており、多重ラベリング問題の解法となる。また、簡潔にパターンを表現できるクラスの組合せが見つからないときはクラスに対応づけるべきではないと判断できるようになる。

(2) 実用性の検討

組合せパターン認識を達成する組合せ最適化アルゴリズムの設計し、①線画の組合せによる文字検出、②正常音・異常音の組合せによる肺聴診音のパターン認識、③動きの組合せによる動画像理解へ応用することで、組合せ論的パターン認識の実用性・信頼性を実験的に明らかにする。

3. 研究の方法

(1) 「組合せパターン認識」の理論構築

既存のパターン認識における識別問題は、クラスの選択をひとつに制限した「組合せない最適化問題」として解釈できる。この制限を解除すると「組合せ最適化問題」としてパターン認識理論を展開できる。パターン認識の問題設定を組合せ最適化問題の視点から整理し、関連研究の動向を調査しつつ理論を構築する。また、整理した問題設定毎に、基本クラスの学習や組合せ最適化の効率的なアルゴリズムの設計基盤を整備し、応用問題毎に適切に設計方針の選択とアルゴリズムの具体化を可能にする。

(2) 実用性の検討

組合せパターン認識の枠組みによって以下の3つの応用問題の解決を試み、実用性を検証する。

①線画の組合せによる文字検出

様々な文字画像から機械学習することで、線画の基本クラスや、典型的な線画の組合せを表す構造が得られることを確認する。特に漢字については、基本クラスの組合せとして偏や旁等のクラスが構成されること、それらを基本クラスを組み合わせた上位のクラスとして漢字が構成されるかどうか検証する。また、偏や旁などの基本クラスの組合せから

なる図形か否かを判定することで漢字を検出する仕組みが実現可能かどうか確認する。

②肺聴診音のパターン認識

電子聴診器によって心肺音の記録や伝送が可能であり、特に肺音の異常を発見する聴診の支援が期待されている。肺聴診音のパターン認識とは、肺聴診音を構成する正常・異常な肺音（呼吸音、連続性ラ音、断続性ラ音等）をすべて言い当てる多重ラベリング問題である。ところが、聴取される音は既にこれらの構成音の混合であるため、各構成音を基本クラスとして機械学習するための訓練データを用意できない。この問題点を組合せパターン認識で解決されるかどうか検証する。なお、肺音検査システムに関する研究業績を有する喜安千弥（長崎大学 教授）と宮原末治（元 長崎大学 教授）の両氏をそれぞれ連携研究者と研究協力者とする。

③動きの組合せによる動画像理解

動画像の見かけの動きの構成要素を基本クラスとして算出し、動きを自動的に分類する。見かけの動きは、カメラの自己運動による動画像全体の動きと、動画像中に存在する物体の動きに起因する画素値の流動（オプティカルフロー場）として定量化される。オプティカルフロー場を動きの重ね合わせと見なし、基本となるオプティカルフロー場とその組合せによって動画像を理解する組合せパターン認識のアルゴリズムを設計し、動きの分類が可能であることを実験的に確認する。

4. 研究成果

(1) 「組合せパターン認識」の理論構築

複数の基本クラス（subclass）の組合せによって認識対象を説明する組合せパターン認識の問題を整理した。組合せの要素となる基本クラスが既知か否か、クラスが階層関係等の何らかの構造や共起性を持つか否かによって、表1のように問題を(C0)～(C5)の6種類に分類した。

表1 組合せパターン認識問題の分類

基本クラス \ 構造	既知	未知
構造が無い	C0	C3
既知の構造をもつ	C1	C4
未知の構造をもつ	C2	C5

例えば、顔認識において各個人の顔画像を用意できる問題設定のように、各基本クラスについてパターンの事例を用意できる場合は(C0)～(C2)の問題として整理でき、識別問題に帰着できることが多い。

一方、組合せパターン認識でなければ克服できない課題は、複数のクラスの特徴が混合

しているパターン事例しか用意できない場合である。そのような課題は(C3)～(C5)に該当する。これらの問題が帰着する組合せ最適化問題とアルゴリズムを列挙できる：(C3)スパースコーディングのための基底学習による基本クラスの教師なし学習、非負行列分解、低ランク+スパース近似等、(C4)グループスパース性等の既知の構造を指定した基底学習（例：J. Mairal+, 10）、局所性を奨励する構造スパース主成分分析（Jenatton+, 2010）等、(C5)任意の典型的な組合せ構造を持つことを基底自体に奨励するダブルスパース正則化、スパースコーディングの多層化と見なせる深層学習等。基本クラスに関する先見情報をスパース性や冗長性、構造や共起性として凸関数で定量化する技法、および凸最適化の原理と最近の進歩を、組合せパターン認識の具体的なアルゴリズム設計に導入できる。

(2) 実用性の検討

①線画の組合せによる文字検出(C4)(C5)

基本クラスとなるべき線画と、その組合せが偏や旁等を構成する関係構造は共に未知であるとする(C5)、文字を構成する基本的な線画ではなく偏や旁等が基本クラスとして基底学習された。そこで、構造スパース主成分分析を応用して基本クラスとなるべき文字の構成要素に画素の局所性の先見情報を与えると(C4)、特徴的な線画が基本クラスとして得られ、更にその組合せとして偏や旁等の高次の特徴を上位のクラスとして獲得できた。また、得られた上位クラスの簡潔な組合せの有無に基づき文字・非文字を頑健に判別できる可能性を確認できた（図1）。

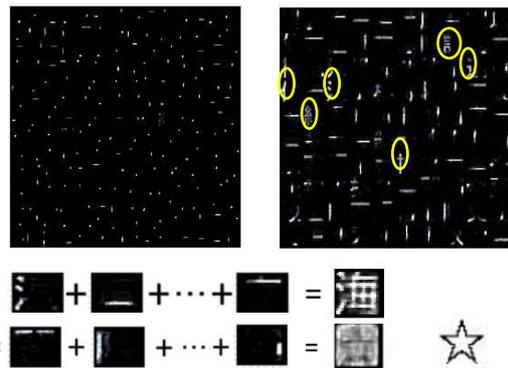


図1 線画の組合せによる文字検出。左上：常用漢字を構成する特徴的な線画の基本クラス。右上：基本クラスの典型的な組合せで構成された上位クラス。偏や旁等の高次の特徴が構成されている。下図：上位クラスの簡潔な組合せで表現できる文字「海」と、簡潔に表現できない図形「☆」。

②肺聴診音のパターン認識(C3)(C4)

肺聴診音のスペクトログラムの低ランク性とスパース性を先見情報とすると(C3)、呼吸音と断続性ラ音の構成要素が部分空間の

基底として得られ、それ以外の構成音が連続性ラ音として分離されることを確認できた。更に、次の(a)~(c)の先見情報のみから肺聴診音を構成する肺音（呼吸音、連続性ラ音、断続性ラ音）を同時に検出できることが示された：(a)呼吸音は限定された少数の周波数成分の組み合わせであること（ただし周波数帯は未知）、(b)断続性ラ音は類似したパルスが反復していること（ただし波形は未知）、(c)連続性ラ音は限定された瞬時周波数成分の組み合わせであること（ただし瞬時周波数は未知）。なお、(b)の先見情報は、断続性ラ音の構成要素が低ランク構造もつことを要求しており、(C4)の問題に対するアプローチと見なせる。聴診音から抽出された各構成音の例を図2に示す。

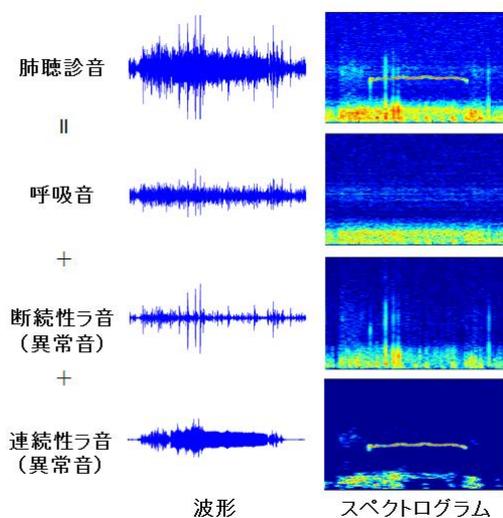


図2 肺聴診音のパターン認識。肺聴診音を簡潔に表せる構成音として、少数の周波数成分からなる呼吸音、類似したパルスの反復からなる断続性ラ音、少数の瞬時周波数成分からなる連続性ラ音が抽出されている。

③動きの組合せによる動画理解(C3)

動画像から得られる時系列2次元オプティカルフローを複素行列で表現し、低ランク+スパース近似を適用すると、図3に示すように、各カメラの自己運動と物体運動それぞれの時系列オプティカルフローが得られた。これにより、図4に示すカメラの自己運動の推定・分類と、運動する物体の検出を同時に達成できることが示された。この組合せパターン認識では、以下の(a)と(b)の先見情報のみを使用しており、各運動のモデリングは一切不要である：(a)自己運動で生じる画像の変動は少数の基本的な変動の組合せで表現できること（ただし基本的な変動は未知）、(b)運動する物体は短時間または画像の一部にしか出現しないこと（ただし物体の出現時刻や位置・形状は未知）。

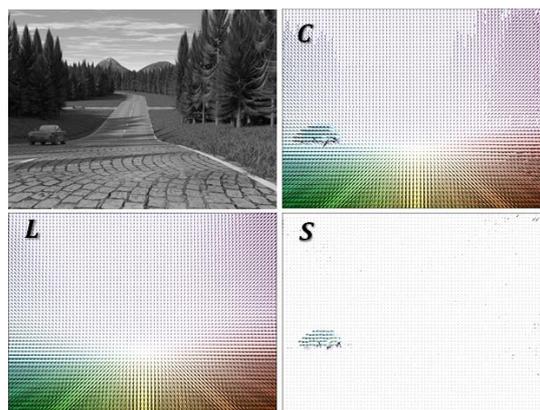


図3 時系列オプティカルフロー場の分離。左上：動画像 (EISATS SET2, IVCNZ, 08) と右上：オプティカルフロー場。左下：低ランク成分として分離された自己運動と、右下：スパース成分として分離された物体の運動によるオプティカルフロー場。対向車が検出されていることがわかる。

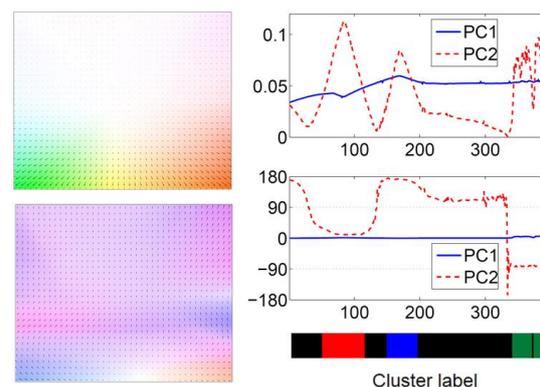


図4 自己運動の分類。低ランク成分の第1主成分（左上）と第2主成分（左下）。それぞれ、直進および回転（方向変化）に対応するオプティカルフローの基底場となっている。実際の時系列オプティカルフロー場は、これらの基底場が各時刻で大きさ（右上）と向き（右中央）を変えながら重ね合わさったものである。このような自己運動の内訳から自己運動を自動分類できる（右下、黒：直進中、赤：勾配減少中、青：勾配増加中、緑：左旋回中）。

5. 主な発表論文等

〔雑誌論文〕（計15件）

- ① 久原 裕貴, 酒井 智弥, 時系列オプティカルフロー場のスパースモデリング, 電気学会研究会資料, 査読無, no. PI-16, pp. 23-28, 2016.
- ② 酒井 智弥, 喜安 千弥, 宮原 末治, 肺聴診音のスパースモデリング, 電気学会電子・情報・システム部門大会講演論文集, pp. 534-538, 2015.
- ③ Tomoya Sakai and Hiroki Kuhara,

Separating background and foreground optical flow fields by low-rank and sparse regularization, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2015), 査読有, pp. 1523-1527, 2015. 10.1109/ICASSP.2015.7178225

- ④ 酒井 智弥, 内田 貴大, 坂口 亜絵梨, 喜安 千弥, 宮原 末治, 岡 三喜男, 肺音認識のための信号処理 ~ 連続性ラ音・断続性ラ音の分離と識別 ~, Technical Report of IEICE, 査読無, vol. 114, no. 39, pp. 55-60, 2014.
- ⑤ 草野 周太, 酒井 智弥, 文字を構成する線画の組合せの機械学習, 第 12 回情報科学技術フォーラム講演論文集, 査読無, pp. 239-244, 2013.
- ⑥ Hayato Itoh, Tomoya Sakai, Kazuhiko Kawamoto and Atsushi Imiya, Dimension reduction methods for image pattern recognition, Lecture Notes in Computer Science, Similarity-Based Pattern Recognition (SIMBAD) Proceedings, 査読有, vol. 7953, pp. 26-42, 2013. 10.1007/978-3-642-39140-8_2
- ⑦ Hayato Itoh, Tomoya Sakai, Kazuhiko Kawamoto and Atsushi Imiya, Topology-preserving dimension-reduction methods for image pattern recognition. Lecture Notes in Computer Science, Scandinavian Conference on Image Analysis (SCIA 2013) Proceedings, 査読有, vol. 7944, pp. 195-204, 2013. 10.1007/978-3-642-38886-6_19

[学会発表] (計 15 件)

- ① 酒井 智弥, (特別講演) スパース構造を活かすセンシングと高次元データ解析, 電子情報通信学会宇宙・航行エレクトロニクス研究会, 2016 年 1 月 21 日, 長崎県美術館 (長崎県長崎市) .
- ② Toshiaki Miura, Tomoya Sakai and Misako Higashijima, Optical measurement and analysis of neck surface deformation for swallowing function evaluation. The 12th International Conference on Ubiquitous Healthcare (u-Healthcare 2015), 2015 年 11 月 30 日~12 月 2 日, 近畿大学 (大阪府東大阪市) .
- ③ 久原 裕貴, 酒井 智弥, 時系列オプティカルフロー場の低ランク+スパースモデルについて, 画像の認識・理解シンポジウム (MIRU2015), 2015 年 7 月 27 日~30 日, ホテル阪急エキスポパーク (大阪府吹田市) .
- ④ 酒井 智弥, 久原 裕貴, 志波 舜平, 時系列 2 次元オプティカルフローの複素数

表現と低ランク成分・スパース成分の抽出, 第 29 回信号処理シンポジウム, 2014 年 11 月 13 日, コミュニティ嵯峨野 (京都市) .

- ⑤ 酒井 智弥, 宮原 末治, 喜安 千弥, 朝長 正臣, 河野 哲也, 土田 朋子, 坂本 憲穂, 石松 祐二, 尾長谷 靖, 河野 茂, 呼吸音・連続性ラ音・断続性ラ音を分離する信号処理技術の紹介, 第 39 回肺音 (呼吸音) 研究会, 2014 年 10 月 26 日, 奈良県新公会堂 (奈良市) .
- ⑥ 草野 周太, 小松 祐太, 酒井 智弥, 文字画像のスパースモデリングと文字検出, 第 28 回信号処理シンポジウム, 2013 年 11 月 21 日, 海峡メッセ下関 (山口県下関市) .

[産業財産権]

特許出願状況 (計 1 件)

名称: 生体音信号処理装置、生体音信号処理方法および生体音信号処理プログラム

発明者: 酒井智弥, 喜安千弥, 宮原末治

権利者: 同上

種類: 特許

番号: 特願 2013-220240

出願年月日: 2013 年 10 月 23 日

国内外の別: 国内

6. 研究組織

(1) 研究代表者

酒井 智弥 (SAKAI, Tomoya)

長崎大学・工学研究科・准教授

研究者番号: 30345003

(2) 連携研究者

喜安 千弥 (KIYASU, Senya)

長崎大学・工学研究科・教授

研究者番号: 20234388

(3) 研究協力者

宮原 末治 (MIYAHARA, Sueharu)