様　式　Ｃ－１９、Ｆ－１９、Ｚ－１９　（共通）

科研費 KAKENHI

科学研究費助成事業　　研究成果報告書

平成 ２７ 年　 ６ 月　 ４ 日現在

機関番号：８２６２６
研究種目：若手研究(B)
研究期間：2012～2014
課題番号：２４７００１８４

研究課題名（和文）テンソル表現に基づくパターン識別法に関する研究

研究課題名（英文）A Study on pattern classification for feature tensor

研究代表者
小林　匠（Kobayashi, Takumi）

独立行政法人産業技術総合研究所・知能システム研究部門・主任研究員

研究者番号：３０４４３１８８

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）：行列・テンソルで表現される特徴量に対して２種の識別手法（行列識別及び特徴テンソル間類似度）を提案した。提案法では行列識別器の高速な最適化方法を示し、訓練誤差と行列のランクを同時に最小化する。得られる低ランク行列識別器は高い汎化性能を有するだけでなく、その識別重みは物理的に解釈することが容易となり更なる解析に資する。また、提案した類似度は特徴テンソル間での部分マッチングに基づき、特徴に内在する共通パターンを自動的に抽出することができる。そのため、背景ノイズなどの変動に頑健に類似度を算出することが可能となる。各種画像認識実験において、これら提案法の有効性を定量的に確認した。

研究成果の概要（英文）：In this study, we have proposed two novel methods for classifying features represented in a form of matrix or tensor; one is a matrix classifier, and the other is a similarity measure between the feature tensors which is used for exemplar-based classification. The proposed method fast optimizes the matrix classifier by minimizing classification errors as well as a matrix rank to produce a low-rank classifier of high generalization performance. Such low-rank classifier also facilitates to physically interpret the classifier weights for further analysis. The proposed similarity measure is based on partial matching of pair-wise feature tensors. It automatically extracts common patterns shared by those feature tensors and thereby produces effective similarity in disregard of noisy background patterns. In the experiments on various visual recognition tasks, the proposed methods exhibited favorable performance.

研究分野：パターン認識

キーワード：パターン識別　特徴行列　特徴テンソル　部分マッチング

１．研究開始当初の背景
(1) 情報基盤の発展に伴い大規模かつ多様な計測データを蓄積することが可能となった一方で、それらを人手で処理することは困難な状況にある。多種多様なデータを有効に活用するためには データの自動認識処理が必要不可欠である。

(2) 実世界から計測したデータには、物理座標といった物理的構造が自然に含まれ、近年ではそのような構造に基づく特徴量（構造化特徴量）を抽出することで自動認識の性能を高めている。構造化特徴量は、特徴次元×物理次元といった形の行列や高階テンソルとして表現されることが一般である。

(3) しかしながら、従来から用いられている識別方法では、そのような構造化特徴量をベクトルへと伸長し、特徴ベクトルとして扱っていた。ベクトル化により、特徴に含まれる構造は無視されてしまうため、その弁別性能が十分に活かされているとは言えない。さらに伸長されたベクトルは高次元となるため過学習も懸念される。

２．研究の目的
(1) 特徴量の構造を考慮したパターン識別法を開発する。特に、構造化特徴量は一般にテンソル（や行列）として表現されるため、特徴テンソル（行列）の識別手法を開発する。

(2) データに基づく学習を通して、特徴テンソルに対する重みを最適化する。これにより行列識別器や、事例ベース識別で用いられるサンプル間類似度を構築する。

(3) 構造に由来する性質として、(行列）重みのランクや、(テンソル）重みの滑らかさや局在性に着目し、特徴量の構造をパターン識別に有効に活用する。

(4) さらに、一般の特徴ベクトルも認識タスクによっては特徴行列として表現され、行列識別器が統一的に適用可能であることも示す。

３．研究の方法
特徴テンソルに対するパターン識別法を構築するために、〔1〕行列識別器、〔2〕特徴テンソル間類似度、の２通りのアプローチをとる。

(1) 行列識別器の学習
特徴行列を $\boldsymbol{X}$ と表すと、行列重み $\boldsymbol{W}$ を用いて行列識別器は次式で表される。

$$y=\mathrm{tr}(\boldsymbol{W}^\top\boldsymbol{X})+b$$

ここでは、行列構造の主要な性質であるランクに注目し、重み $\boldsymbol{W}$ が極力低ランクとなるように最適化（学習）する。低ランクとすることで過学習が抑えられ、識別器の汎化性能の向上が期待できる。ランクを加味した最適化問題は次式で定式化できる。

$$\min_{\boldsymbol{W},b}\mathrm{rank}(\boldsymbol{W})+C\sum_{i=1}^{n}\max[0,1-y_i\{\mathrm{tr}(\boldsymbol{W}^\top\boldsymbol{X}_i)+b\}]$$

１項目はランク数を、２項目は訓練誤差を意味する。しかしこのランク最小化問題は NP 困難となり、効率的に最適化することが難しい。そこで、次の３ステップを経ることで、効率的に最適化可能な問題へと帰着させる。
①ランクを線形近似するトレースノルムを導入する。

$$\min_{\boldsymbol{W},b}\|\boldsymbol{W}\|_\Sigma+C\sum_{i=1}^{n}\max[0,1-y_i\{\mathrm{tr}(\boldsymbol{W}^\top\boldsymbol{X}_i)+b\}]$$

②重み $\boldsymbol{W}$ を $\boldsymbol{W}=\boldsymbol{W}_R\boldsymbol{W}_C^T$ と分解し、$\boldsymbol{W}_R$ に関する部分最適化を、その双対問題[1]に置き換える。

$$\min_{\boldsymbol{W}_C}\frac{1}{2}\mathrm{tr}(\boldsymbol{W}_C\boldsymbol{W}_C^\top)+\sum_i\alpha_i^*(\mathbf{W}_C)-\frac{1}{2}\sum_{ij}\alpha_i^*(\mathbf{W}_C)\alpha_j^*(\mathbf{W}_C)y_iy_jK_{ij}(\mathbf{W}_C)$$

③さらに新たな変数 $\Sigma_C=\boldsymbol{W}_C\boldsymbol{W}_C^T$ を導入することで Σ_C のみの最適化問題へと帰着させる。

$$\min_{\boldsymbol{\Sigma}_C\succcurlyeq 0}\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}_C)+\sum_i\alpha_i^*(\boldsymbol{\Sigma}_C)-\frac{1}{2}\sum_{ij}\alpha_i^*(\boldsymbol{\Sigma}_C)\alpha_j^*(\boldsymbol{\Sigma}_C)y_iy_jK_{ij}(\boldsymbol{\Sigma}_C)$$

$$\text{where }\boldsymbol{\alpha}^*(\boldsymbol{\Sigma}_\mathrm{w})=\arg\max_{\boldsymbol{\alpha}\in\Omega}\sum_i\alpha_i-\frac{1}{2}\sum_{i,j}\alpha_i\alpha_jy_iy_jK_{ij}(\boldsymbol{\Sigma}_\mathrm{w}).$$

これは凸最適化問題であり、勾配法により効率的に最適化できる。さらに、ここではランク数を事前に規定することなく、自動的に低ランクへと最適化されることも利点である。また、元の識別重みは次式により復元される。

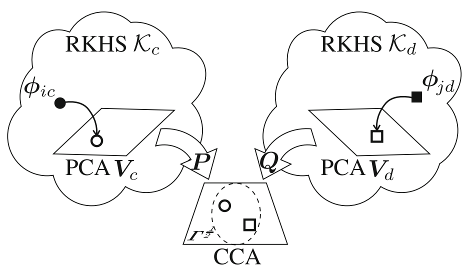
$$W^*=\sum_{i=1}^{n}\alpha_i^*(\boldsymbol{\Sigma}_C^*)y_i\boldsymbol{X}_i\boldsymbol{\Sigma}_C^*$$

(2) 行列識別器の展開
特徴行列が直接的に得られていない場合にも、①複数カーネル統合学習や②多クラスクロスモーダル学習において、特徴表現を行列へと拡張し、上記行列識別器が適用可能となる。

①複数カーネル統合学習
種類 c のカーネル特徴ベクトルを ϕ_c として、特徴行列を次式により定義する。

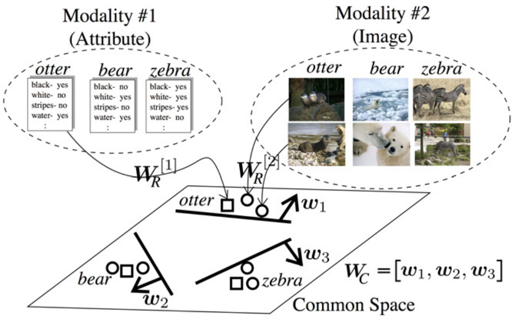
$$\boldsymbol{X}=\boldsymbol{Z}\boldsymbol{X}^\phi,\qquad \boldsymbol{X}_i^\phi=\begin{bmatrix}\boldsymbol{\phi}_{i1} & & \\ & \ddots & \\ & & \boldsymbol{\phi}_{iw}\end{bmatrix}$$

ここで $\boldsymbol{Z}$ はカーネル正準相関分析（下図参照）に基づく射影行列である。



このように、複数カーネルによる特徴表現を行列へと拡張することで、行列識別器学習[3-1]が適用可能となる。この場合には、最適化された Σ_c は各カーネル及びカーネル間に対する重みとして働き、その結果行列識別器は複数カーネルを統合した識別器として作用する。

②多クラスクロスモーダル学習
多クラスの対象に対して、多モダリティ（画像やテキスト）から特徴ベクトルが得られる場合に、上記行列識別器により、複数モダリティ間の共通表現を得つつ識別することが可能となる。



クラス c の特徴ベクトル $\boldsymbol{x}$ が得られたとき、特徴行列を

$$\boldsymbol{X} = [\boldsymbol{0}, \ldots, \boldsymbol{0}, \boldsymbol{x}, \boldsymbol{0}, \ldots, \boldsymbol{0}]$$

c 列目

と定義する。このとき、各モダリティに対する識別重みを連結した重み $\boldsymbol{W}$ は次式のように分解される。

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}^{[1]} \\ \vdots \\ \boldsymbol{W}^{[M]} \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}_R^{[1]} \\ \vdots \\ \boldsymbol{W}_R^{[M]} \end{bmatrix} \boldsymbol{W}_C^\top$$

これは、共通空間への射影 $\boldsymbol{W_R}$ とそこでの識別重み $\boldsymbol{W_C}$ が自然と内包されていることを意味する。この $\boldsymbol{W}$ の最適化に上記手法を適用することで、射影と識別重みを同時に最適化することができる。

(3) 特徴テンソル間類似度
実世界から計測した特徴量は一般に特徴次元×物理次元の形式のテンソルとして表現される。例えば、静止画像では XY 座標に基づく２次元データであり、多チャンネル時系列信号からは時間軸 T に沿って並ぶ信号列（チャンネル×時間）となる。そのような特徴テンソル内では、注目すべき対象パターンの出現位置は任意であり、さらに対象に関係のない背景パターン等も多分に含まれている。そこで、そのような背景パターンを除外しつつ位置が不定な対象パターンを抽出する部分マッチングを提案した。マッチングにより類似度が算出され、それは k-NN に代表される事例ベース識別器で用いられる。物理次元における重み $\boldsymbol{w}$ を導入すると、特徴テンソル $\boldsymbol{X},\boldsymbol{Y}$ のマッチングはその重み $\boldsymbol{w}$ の最適化問題へと帰着できる。

$$\min_{\boldsymbol{w}_X,\boldsymbol{w}_Y} \arccos \frac{\boldsymbol{x}^\top \boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}, \ s.t.\ \boldsymbol{x}=\boldsymbol{X}\boldsymbol{w}_X, \boldsymbol{y}=\boldsymbol{Y}\boldsymbol{w}_Y,$$
$$\Leftrightarrow \min_{\boldsymbol{w}_X,\boldsymbol{w}_Y} \frac{1}{2}\|\boldsymbol{X}\boldsymbol{w}_X - \boldsymbol{Y}\boldsymbol{w}_Y\|^2, \ s.t.\ \|\boldsymbol{X}\boldsymbol{w}_X\|^2 = \|\boldsymbol{Y}\boldsymbol{w}_Y\|^2 = 1.$$

さらに、特徴構造として物理次元内での局在性や滑らかさに注目し、それらを重みの正則化として導入する。

$$\min_{\boldsymbol{w}} \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{A}\boldsymbol{w} + \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{L}\boldsymbol{w} + \rho_X \Omega(\boldsymbol{w}_X) + \rho_Y \Omega(\boldsymbol{w}_Y),$$
$$s.t.\ \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{B}\boldsymbol{w} = 1,$$

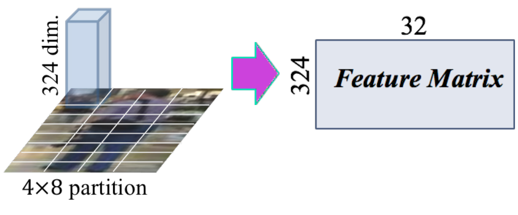
$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{X}^\top \boldsymbol{X} & -\boldsymbol{X}^\top \boldsymbol{Y} \\ -\boldsymbol{Y}^\top \boldsymbol{X} & \boldsymbol{Y}^\top \boldsymbol{Y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{K}_X & -\boldsymbol{K}_{XY} \\ -\boldsymbol{K}_{XY}^\top & \boldsymbol{K}_Y \end{bmatrix},$$
$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{X}^\top \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Y}^\top \boldsymbol{Y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{K}_X & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}_Y \end{bmatrix}, \ \boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_X \\ \boldsymbol{w}_Y \end{bmatrix}$$

ここで、$\boldsymbol{L}$ は物理座標由来のラプラシアン行列、Ω は structured sparseness[2] として定式化される局在化を誘発する正則化である。これらの正則化項は特徴テンソルの形式に基づいて適切に設定することで、任意の階数の特徴テンソルを扱うことができる。上記最適化問題は、逐次的な２次形式最適化問題へと帰着され、共役勾配法を適用することで効率的に最適化することが可能となる。これにより得られる重みは、共通対象領域のみで非零の値をとることが期待され、無関係な背景パターンを除外しつつ、本質的な対象パターンのみを自動的に抽出することができ、効果的な類似度も算出できる。

４．研究成果
様々な画像認識実験において提案法の有効性を定量的に確認した。

(1)行列識別器[3-1]を画像・動画像認識へと適用した。

①人識別
INRIA Person Dataset を用いて人 vs 非人の画像識別タスクにおける性能を評価した。局所特徴として GLAC 特徴[3]を採用すると、下図のような特徴行列が得られる。



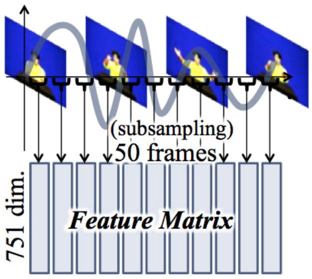
識別性能の比較結果を下表に示す。

	ランク	Error Rate
従来法 [Dalal&Triggs,2005]	−	2.25%
ベクトル識別法	32	0.55%
行列識別法	12	0.53%

従来法やベクトル識別法よりも高い性能を示し、かつ識別器も遥かに低いランクへと最適化されている。低ランクは識別器を適用す

る際の計算コストの削減にもつながる。

②動作識別
次に、RWCP 動画像データセットにおける動作識別に適用した。時系列上の各フレームから CHLAC 特徴[4]を抽出すると、下図のような特徴行列が得られる。



識別性能の比較は下表のようになる。

	ランク	Error Rate
従来法 [Kobayashi&Otsu, 2009]	–	4.14%
ベクトル識別法	50	2.11%
行列識別法	12	0.98%

先の画像認識の場合と同様に、従来法やベクトル識別法の性能を凌駕し、かつ低ランクの識別器が得られていることが分かる。
以上の結果により、特徴行列表現及びその形式での識別の有効性を示すことができた。

(2) 行列識別器の拡張手法[3-2]を各種画像認識課題へ適用した。

①複数カーネル統合
行列識別の拡張の１つとして、複数カーネル統合タスクへ適用した。Flower Dataset において、7 種のカーネルを統合して識別を行ったところ下表の性能を得た。

	Error Rate
従来法による統合 [simpleMKL]	18.92%
行列識別法による統合	13.53%

行列識別法によって、カーネル間の関係性も考慮した統合が行われ、従来法に比べても高い識別性能を示している。

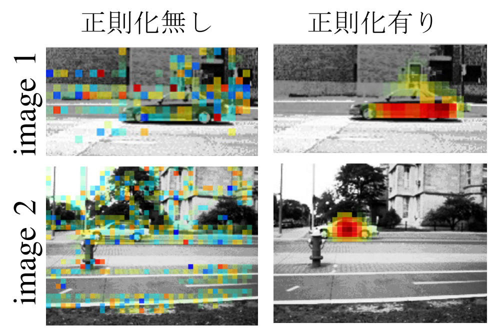
②多クラスクロスモーダル学習
２つ目の拡張として、クロスモーダル学習への適用を試みた。Animals with Attributes Dataset において、動物の属性情報と画像情報のクロスモーダルによる識別を行った。

	Accuracy
従来法 [Lampert et al.2009]	36.76%
行列識別法による学習	41.43%

行列識別器学習の枠組みにおいて、画像と属性という異なるモダリティ間での学習が効果的に行えることが示された。

(3)特徴テンソル間類似度[3-3]に基づき、画像・動画像のマッチングを行った。

①教師無し対象検出
類似度算出における重み最適化の有効性を評価するため、共通パターンとして「車」を含む画像間でのマッチングを行った（下図）。



上図では重みの大きさを疑似色で表示している。左図は正則化無し、つまり特徴の構造を無視してマッチングを行った場合、右図は提案法により正則化を導入した場合の結果である。提案法により、共通パターンである「車」領域が正しく抽出されていることが分かる。ここでは、対象に関する情報（「対象が何か」や「どこにいるか」）を一切与えていないにも関わらず、共通パターン（「車」）が自動的に抽出されている。これにより、提案した類似度算出法において、有意な重みが得られていることが示された。

②類似度に基づく動作識別
次に、動画像から得られる特徴テンソル間での類似度に基づき、事例ベース識別法を用いて動作識別を行った。Weizmann action dataset に対して(1)②と同様に CHLAC 特徴[4]をフレーム毎に抽出し、特徴行列を構成した。識別には k-NN (k=3)法を採用した。

	Accuracy
従来法 [Jhuang et al. 2007]	98.8%
従来法 [Wang&Mori, 2009]	100%
部分マッチング類似度	100%

上表に示すように従来法と同様に高い識別性能が得られた。
以上の実験により、部分マッチングにおける重み及びそこから算出される類似度の有効性を示すことができた。

＜引用文献＞
[1] V. N. Vapnik, “Statistical Learning Theory,” Wiley-Interscience, 1998.
[2] R. Jenatton, J.-Y. Audibert, and F.

Bach, “Structured variable selection with sparsity-inducing norms,” Journal of Machine Learning Research, vol. 12, pp. 2777-2824, 2011.
[3] T. Kobayashi and N. Otsu, “Image Feature Extraction Using Gradient Local Auto-correlations,” Proc. European Conference on Computer Vision (ECCV), pp. 346-358, 2008.
[4] T. Kobayashi, N. Otsu, “Three-way Auto Correlation Approach to Motion Recognition,” Pattern Recognition Letters, Vol. 30, No. 3, pp. 212-221, 2009.

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 T. Kobayashi, Low-Rank Bilinear Classification: Efficient Convex Optimization and Extensions, 査読有, International Journal of Computer Vision, Vol. 110, No. 3, 2014, pp. 308-327
DOI: 10.1007/s11263-014-0709-5

〔学会発表〕（計３件）
① T. Kobayashi, N. Otsu, Efficient Optimization For Low-Rank Integrated Bilinear Classifiers, 査読有, Proc. European Conference on Computer Vision (ECCV), 2012, pp. 474-487
② 小林匠, 大津 展之, Efficient Optimization for Low-Rank Integrated Bilinear Classifiers, 招待講演 ,第 16 回画像の認識・理解シンポジウム MIRU2013
③ T. Kobayashi, S3CCA: Smoothly Structured Sparse CCA For Partial Pattern Matching, 査読有, Proc. International Conference on Pattern Recognition (ICPR), 2014, pp. 1981-1986

〔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

６．研究組織
(1)研究代表者
小林 匠（KOBAYAHSI Takumi）
国立研究開発法人産業技術総合研究所・知能システム研究部門・主任研究員
研究者番号：30443188

(2)研究分担者

(3)連携研究者