

# 組合せ最適化に基づく未学習クラスの識別と 顔画像認識における評価

小松 祐太<sup>1</sup> 酒井 智弥<sup>1</sup>

**概要:** 本研究では、パターン認識を、可能な限り簡潔にテストデータを説明できる学習データの組合せを求める最適化問題を解くことに基づく識別問題として再解釈する。その組合せ最適化に基づく識別において、少数の学習データでテストデータを説明できるかどうかで「知っている」、「知らない」と答えることができるパターン認識を達成する。提案手法は組合せ最適化問題の解のスパース性を判定する方法である。アピアランスベースの顔画像認識で評価した結果、負例を用意することなく既学習クラスの識別率と同等の未学習クラスの識別率が得られた。また、組合せ最適化に基づくパターン認識において、学習データを構造化する意味や役割についても議論する。

**キーワード:** パターン認識, 顔認識, スパース正則化, 構造正則化

## Classification to Learned or Unlearned Class via Combinatorial Optimization and its Evaluation in Appearance-Based Face Recognition

KOMATSU YUTA<sup>1</sup> SAKAI TOMOYA<sup>1</sup>

**Abstract:** We interpret pattern classification as a problem of finding a combination of training data or classes that can explain a given test data as concisely as possible. This combinatorial optimization problem can be formulated as decomposition of test data into a sparse linear combination of training data. Our classification method assigns a label of a learned class indicated by the sparse solution if it exists, and “unlearned” otherwise. Experimental evaluation in appearance-based face recognition shows that the recognition rate for unlearned subjects is comparable to those for learned faces. We also discuss the role and significance of structuring training data from the viewpoint of pattern recognition as a combinatorial optimization problem.

**Keywords:** pattern recognition, face recognition, sparse regularization, structure regularization

### 1. はじめに

パターン認識 [1] とは、観測されたパターンをあらかじめ定められた複数の概念のうちの一つに対応させる処理のことである。先行研究では、テストデータと最も似ている学習データからクラスを決定する方法や、特徴空間におけるクラスの境界を求めて識別する方法 [2] などが提案されてきた。

このような識別問題に帰着させるパターン認識では、既知のクラスを何れかにパターンを分類するように識別器が設計されている。このような識別器で、知らないものに「知らない」と識別する機能を持たせるには、識別関数などの出力に閾値を設定する、または負例を学習する必要がある。

そもそもパターンとは、パターンから何らかの規則性や構造を見出すことができ、その規則性や構造を利用してパターンを簡潔に説明することであると解釈することができる。今、学習データによってテストデータを説明しようと

<sup>1</sup> 長崎大学  
Nagasaki University

したとき、少数の学習データでテストデータを説明することができたなら、そのテストデータは認識されたことになり、そのテストデータが属するクラスを少数の学習データから判断することができる。つまり、パターン認識とは、可能な限り簡潔にテストデータを説明できる学習データまたはクラスの組合せを求める最適化問題である。

本論文では、組合せ最適化に基づき、未学習クラスのテストデータに対して「知らない」と識別することを目的とする。テストデータを簡潔に説明する組合せとその係数を求める組合せ最適問題としてスパース正則化 [5][6][7] を用いる。そのスパース正則化の解に基づいて未学習データに対して「知らない」と識別する手法を提案する。

## 2. 代表的な識別関数法

従来手法では、学習データから決定境界や識別関数を事前に求め、テストデータに対する識別関数の値からクラスのラベルを決定するような識別を行ってきた。決定境界や識別関数を用いて識別を行う手法は総称して識別関数法と呼ばれる [1]。

### 2.1 部分空間法

部分空間法 [4] では、1つのクラスを特徴空間中のひとつの部分空間で表す。その部分空間を学習データから求め、テストデータから最も近い部分空間のクラスへ識別する。

ここで、テストデータの特徴ベクトル  $\mathbf{x} \in \mathcal{R}^m$ 、あるクラスの  $r$  次元部分空間を張る正規直交基底を  $\{\varphi_1, \dots, \varphi_r\}$  とする。テストデータと学習データの類似度はテストデータの部分空間への射影の長さで定義される。

$$S = \sum_{i=0}^{r-1} (\mathbf{x}, \varphi_i)^2 \quad (1)$$

テストデータはこの類似度が最も大きいクラスへ識別される。

この部分空間法で識別することには問題がある。まず、テストデータ  $\mathbf{q}$  が張る部分空間が図1のように与えられたとする。テストデータ  $\mathbf{q}$  は、ラベル  $c_1$  を持つ学習データの線形結合で表せる特徴ベクトルであり、識別結果としてラベル  $c_1$  が付与されるべきである。ところが、図1のように、 $\mathbf{q}$  はどのクラスの部分空間に対しても射影成分  $u_1, u_2, u_3$  を持つ。このように、複数のクラスの部分空間がすべて互いに直交していない限り、テストデータは識別されるべき正しいクラス以外の部分空間にも射影成分を持っており、他のクラスの類似度がゼロになることはない。

通常では、これでもある程度の識別を達成することができるが、図2のように、テストデータが  $c_1$  の部分空間から離れているとき、部分空間  $c_1$  への射影成分  $u_1$  が、他の部分空間への射影成分より小さくなり、類似度を手がかりに正しいクラスラベル  $c_1$  を付与することが困難になりうる。また、

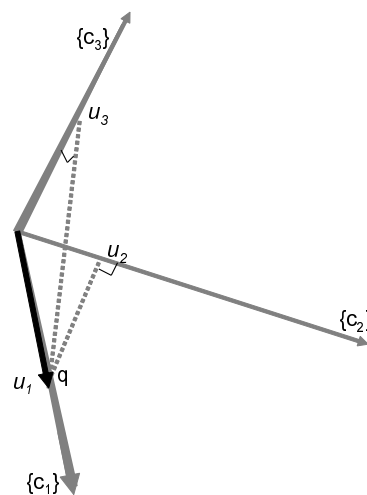


図1 各部分空間  $c_1, c_2, c_3$  へのテストデータ  $\mathbf{q}$  の射影成分  $u_1, u_2, u_3$ .

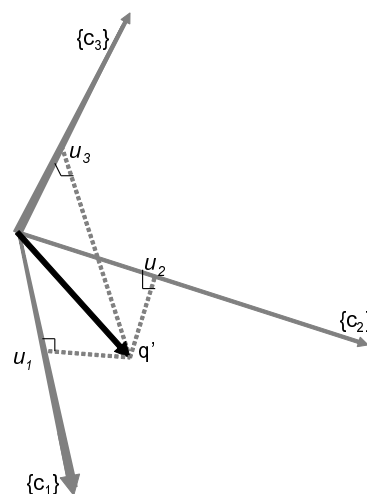


図2 部分空間  $c_1, c_2$  の和空間の成分を持つテストデータ  $\mathbf{q}'$  と  $\mathbf{q}$  の各部分空間の射影成分  $u_1, u_2, u_3$

このような場合、テストデータに複数のクラスラベルの組合せ  $c_1, c_2$  を付与するほうが望ましいと考えられる。

このように、テストデータを識別するのに関係のないクラスの類似度が非ゼロにならず、最も類似度の大きいひとつだけで識別しようとすることに起因する問題がある。

## 3. 組合せ問題に基づく識別法

### 3.1 スパース表現による識別法

類似度に基づく識別関数法ではなく、学習データによるテストデータの簡潔な表現に基づく識別を考える。クラスが線形部分空間法で表せるとき、学習データによる簡潔な表現とは、テストデータを学習データの線形結合で表したとき、多くの線形結合係数がゼロになることである。すなわち、学習データひとつひとつが列ベクトルとなっている行列  $\mathbf{D} \in \mathcal{R}^{m \times n}$ 、テストデータ  $\mathbf{q} \in \mathcal{R}^m$ 、線形結合係数を成分とするベクトルを  $\boldsymbol{\alpha} \in \mathcal{R}^n$  としたとき、

$$\mathbf{q} = \mathbf{D}\boldsymbol{\alpha} \quad (2)$$

を満たすスパースなベクトル  $\alpha$  を求める問題を考える。

このようなスパース正則化を解くアプローチは大きく 2 種類に分類できる。ひとつは、

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_0 \quad s.t. \quad D\alpha = q \quad (3)$$

を貪欲法で求めるものである。その代表的なアルゴリズムとして OMP[5] がある。

ここで、 $\|\cdot\|_p$  は、で表される  $l_p$  ノルムのことであり、 $\|\alpha\|_0$  は、ゼロでない成分の個数である。式 (3) では、テストデータを説明する学習データの組み合わせとその係数であるスパースなベクトル  $\alpha$  の非ゼロ成分の個数が最小となる  $\alpha$  を探すことを意味する。

$\min_{\alpha} \|\alpha\|_0$  のような最小化問題のことを  $l_0$  ノルムの最小化問題と呼ぶ。これを  $l_1$  ノルムの最小化問題に緩和した

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1 \quad s.t. \quad D\alpha = q. \quad (4)$$

は、式 (3) の  $l_0$  ノルムの最小化解と高い確率で一致する一意の解を持つ。その代表的なアルゴリズムとして Lasso[6] がある。

式 (3) と式 (4) のような最小化問題を解くことにより、テストデータ  $q$  を学習データ  $D$  で簡潔に説明するための組み合わせとその係数  $\alpha$  を求める。このときの  $\alpha$  は多数のゼロ成分と、少数の非ゼロ成分で成り立つスパースなベクトルになっている。その求めた  $\alpha$  を使って第  $c$  クラスの学習データの結合係数からなるベクトル  $\delta_c(\hat{\alpha})$  を用いたときのテストデータとの残差を

$$r_c(q) = \|q - D_c \delta_c(\hat{\alpha})\|_2 \quad c = 1, \dots, k. \quad (5)$$

のようにして求め、最も残差を小さくできたクラスが識別結果として与えられる。

Wright らの先行研究 [9] では、このアルゴリズムを前面から撮影した様々な表情や明るさの顔画像の画素値を特徴量に用いて、テストデータを簡潔に説明できるような学習データの組み合わせを見つけることができ、識別性能も既存手法よりも高いことが示されている。さらに、式 (3) や式 (4) を用いることにより、ノイズを考慮した最小化問題を解くことになっており、ノイズに対する頑健性が強くなることも示されている。

### 3.2 スパース解の存在を評価する指標

スパース解の非ゼロ成分はテストデータが属するクラスを示している。一方、 $\alpha$  がスパースではない場合、テストデータを説明する簡潔な学習データの組合せが存在しないことを意味する。

スパース表現を検証するための指標として、Wright ら [9] により Sparsity Concentration Index(SCI) が定義されている。

$$SCI(\alpha) = \frac{k \cdot \max_c \|\delta_c(\alpha)\|_1 / \|\alpha\|_1 - 1}{k - 1} \in [0, 1]. \quad c = 1, \dots, k. \quad (6)$$

SCI とは、どの程度  $\alpha$  をスパースな解にすることが出来たかを表す数値であり、値が 1 に近いほど  $\alpha$  がスパースであることを表し、値がゼロに近いほど  $\alpha$  はスパースではないことを表している。

### 3.3 帰属値

スパース正則化により求めた  $\alpha$  を用いて識別結果を出す際に、SRC では、スパース正則化により求めた  $\alpha$  が、各クラス毎にどのくらいの説明能力を持っているのかを表す指標として、テストデータ  $q$  への帰属値というものを酒井 [8] が定義している。また、説明能力とは、テストデータを説明する学習データの簡潔な組合せが、どの程度テストデータの合成に貢献するのかを表す能力のことである。

第  $c$  クラスへの帰属値  $\nu_c$  は、テストデータ  $q$ , 復元  $\hat{q} = D\hat{\alpha}$  によって

$$\nu_c = \frac{q^T D_c \delta_c(\alpha)}{\|q\|_2 \|\hat{q}\|_2} \quad (7)$$

と記される。

## 4. 未学習クラスへの識別

本研究では、3 で述べた組合せを用いた識別方法で未学習クラスに対して「知らない」と識別する方法としてスパース性を用いてスパース解の有無を判定する方法を試みる。

### 4.1 スパース正則化の特徴を用いた未学習クラスの識別

スパース正則化の性質であるスパース性を用いることにより、未知データを識別するための学習データを用意することなく、知らないと識別することができる。スパース正則化の性質として、テストデータが既知の画像の場合、学習データのうち、テストデータと同じラベルを持つ学習データのみでテストデータのほとんどを復元することができる。また、テストデータが未知の画像の場合は、学習データのうち、複数のラベルの学習データを使ってテストデータを復元しようとする。そのような性質を用いて、未知のテストデータを識別していくことを提案する。

スパース正則化により求められた解がどの程度スパースなのかを表す指標として 3.2 で SCI というものを紹介した。この SCI は、 $\alpha$  がどのくらいスパースな解かを表しており、言い換えれば、 $\alpha$  がテストデータを復元するときにクラス毎でどのくらい使用されているのかを表していることになる。

しかし、Wright ら [9] の SCI で使用されている  $\alpha$  とは、学習データからテストデータを復元するための結合係数であり、テストデータの復元に貢献する値ではない。

そこで、テストデータの説明に用いられる割合を表す指

標として式 (7) で述べた帰属値というものがある. この帰属値を SCI の  $\alpha$  の代わりに使用することで, テストデータの説明に用いられる割合を定量化することができる. その割合を定量化すると, 各学習データに付与されたラベルを  $c$  で表し, 各クラスごとの帰属値を  $\nu_c$ , 全クラスの帰属値の和を  $\nu$  とするとき,

$$CCI(\nu_c) = \frac{k \cdot \max_c \nu_c / \nu - 1}{k - 1} \in [0, 1]. \quad (8)$$

となる. これを, CCI(Contribution Concentration Index) と定義する. CCI がゼロに近いほど複数のクラスでテストデータを復元しようとしており, CCI が 1 に近いほど 1 つのクラスでテストデータを復元していることがわかる.

つまり, SCI と同じように,

$$CCI(\nu_c) \geq \tau \quad (9)$$

のような閾値  $\tau$  を定めることにより, 閾値より小さい場合は多くのクラスを使ってテストデータを説明していることになる. これは簡潔に説明していることにならないため, 「知らない」と識別する. 閾値より大きい場合はひとつのクラスのみでテストデータの大部分を説明できたこととなり, そのクラスに識別する. このようにして未学習データを「知らない」と識別する.

## 4.2 識別方法

スパース正則化によって, スパースな解を求めることができた. この求めたスパースな解から, テストデータがどの学習データを使って説明できたのかを定量化し, クラスを決定する. 学習データ  $D \in \mathcal{R}^{m \times n}$ , テストデータ  $q \in \mathcal{R}^m$ , スパースな解  $\alpha \in \mathcal{R}^n$ , クラスの数  $k$  のとき,  $\alpha$  のひとつのクラスごとの要素と学習データを用いて,

$$\hat{q}_c = D_c \hat{\alpha}_c \quad s.t. \quad c = 1, \dots, k \quad (10)$$

のようにして, 学習データとスパースなベクトル  $\alpha$  によって求められる  $\hat{q}_1 \dots \hat{q}_k$  をテストデータの復元と呼ぶ. そして,

$$\hat{c} = \arg \min_{c=1, \dots, k} \|q - \hat{q}_c\|_2 \quad (11)$$

のようにクラス  $\hat{c}$  を決定することができる.

## 5. 学習データの構造化に基づくパターン認識

### 5.1 学習データの構造化と未学習クラスへの識別

テストデータが属すべきクラスが既に学習データにある場合, その正解クラスの学習データの枚数を少なくする制限は不要であり, 正解クラスのみを学習データを多数使用してもよい. これは, 学習データをクラス毎にグループ化し, テストデータを説明するグループ数を最小化する解に基づく識別である. また, 学習データやその集合間で関係を持つようなグラフ構造を作成し, 構造の一部でテスト

データを簡潔に説明できれば, 構造化された知識に基づく高度な識別の実現を期待できる.

例えば, 学習データを階層関係で表すことにより, 知識に深さを持たせることができるのではないだろうか. 学習データやその集合間で階層構造を構築し, テストデータを簡潔に説明できる学習データの部分構造を探すことができれば, テストデータと無関係な学習データの部分構造が除外され, 識別性能の向上が期待できる. 小数の部分構造でテストデータを簡潔に説明できる階層レベルは, 認識の深さを表す. また, 構造に基づいて簡潔に説明できなければ, 「知らない」として未学習クラスへ識別できる.

### 5.2 構造に基づくスパース正則化

学習データの任意のグループを設定し, 学習データやその集合関係をグラフ構造で表現することで, 学習データを整理する. そして, テストデータを説明する学習データのグラフ構造が小さくなるようにスパースな解を求める. このとき, 同じグループの学習データは, 類似の学習データの集合であるべきで, 類似の学習データの集合にすることにより, そのグループらしさを表現できる集合となる.

テストデータ  $q$  を説明する学習データのグラフ構造が小さくなるようにスパースな解は, 学習データ  $D \in \mathcal{R}^{m \times n}$ , テストデータ  $q \in \mathcal{R}^m$ , スパースな解  $\alpha \in \mathcal{R}^n$  としたとき,

$$\min_{\alpha} \frac{1}{2} \|q - D\alpha\|_2^2 + \lambda \Omega(\alpha) \quad (12)$$

のような最小化問題を解くことにより得られる.  $\lambda$  は解のスパース性を調整する正則化のパラメータである. また,

$$\Omega(\alpha) = \sum_{g \in \mathcal{G}} w_g \|\alpha_{|g}\| \quad (13)$$

であり,  $\alpha_{|g} \in \mathcal{R}^n$  は学習データの小グループ  $g$  の要素に対応する結合係数からなるベクトルを表す. また,  $w_g$  は, 小グループ  $g$  の重みを表す. なお, 本論文では, グラフ構造の最も単純な構造であるグループ構造を実験で扱っており, グループ構造を解くグループスパース正則化の代表的な手法としてグループ OMP[10] とグループ Lasso[11] の 2 つを用いる.

## 6. 顔画像認識による評価実験

### 6.1 実験と評価方法

スパース正則化のアルゴリズムの違いで識別率がどのように変化するのか, また, 未学習クラスの識別の判断材料として CCI を用いた場合にどの程度未知のデータを知らないで識別できたのかを測る.

本実験では, 2414 枚の顔画像データベース CroppedYale[14] を使用したアラビアンベースの顔画像認識で提案手法の未学習クラスへの識別性能を評価する. 学習データとテストデータは正面から撮った照明条件

表 1 各スパース正則化の CCI の閾値  $\tau$  の一覧

解法	CCI の閾値 $\tau$
OMP	0.45
Lasso	0.35
グループ OMP	0.35
グループ Lasso	0.25

が異なる顔画像であり、サイズを  $32 \times 32=1024$  画素に縮小し、画素値を成分とするベクトルで表す。画像には 38 人分のクラスのラベルがそれぞれ付与されている。先行研究 [9] でスパース表現に基づく識別法の性能が明らかにされている。

学習データにないクラスの画像がテストデータとして与えられた場合に、未学習クラスの識別性能を評価するため、38 クラスの画像のうち、30 クラスを既学習クラス、残りの 8 クラスを未学習クラスに分ける。使用する画像 2414 枚をランダムにテストデータ用と学習データ用の半分ずつに分ける。テストデータ用の画像の中からランダムに 150 枚選び、実際に使用するテストデータとする。学習データ用の画像のうち、既学習クラスの画像を学習データとして使用する。グループ正則化の場合は、学習データがクラス毎にグループになっている。未学習クラスの場合は、テストデータと識別結果で与えられるラベルとが一致しなくても互いに未学習クラスのラベルだった場合は「知らない」と出力する。

識別を  $10^1$  回反復し、既学習クラスや未学習クラス、全体の識別率の平均と偏差を求める。また、CCI の閾値を変化させることによる ROC 曲線の評価も行う。

スパースな解を求めるために最小化問題を解く必要がある。本実験では、SPAMS[15] を利用して提案手法を MATLAB で実装した。SPAMS には、グループや構造に基づく様々な最適化問題を解くアルゴリズムが実装されており、OMP や lasso などのスパースな分解問題を解くアルゴリズムの実装も提供されている。本実験では、下記の手法 1 から 4 までのスパース正則化を用いて実験をする。

手法 1 OMP

手法 2 Lasso

手法 3 グループ正則化 (グループ OMP)

手法 4 グループ正則化 (グループ Lasso)

## 6.2 実験結果

### 6.2.1 CCI を用いた各スパース正則化の性能評価

各スパース正則化の識別率をまとめたものが、表 2 である。このときの CCI の閾値  $\tau$  は、0 から 1 の中で最も性能が良くなるものに設定している。そのときの各スパース正則化の CCI の閾値  $\tau$  は表 1 である。なお、表 2 では、提案手法の識別率と比較するために、テストデータに関係なく、38 人の中からランダムに 1 人の顔を答える識別器 (ランダム) を用意した。アルゴリズム毎の性能を評価するた

表 2 未学習クラスへの識別の有無による識別率の比較 (%)

解法	未学習クラスへの識別	既知顔画像の識別率	未知顔画像の識別率	全体の識別率
ランダム	無	2.5 ± 1.4	21.5 ± 6.8	6.5 ± 1.7
OMP	無	95.9 ± 2.3	-	75.2 ± 2.6
	有	92.8 ± 2.5	92.2 ± 3.2	92.7 ± 2.7
Lasso	無	96.8 ± 1.2	-	75.2 ± 2.4
	有	94.6 ± 1.6	86.6 ± 4.6	92.8 ± 1.5
グループ OMP	無	73.9 ± 3.7	-	57.9 ± 4.0
	有	73.9 ± 3.7	65.5 ± 9.4	72.0 ± 4.0
グループ Lasso	無	92.7 ± 2.3	-	72.1 ± 3.7
	有	85.9 ± 2.8	79.4 ± 7.0	84.6 ± 2.7

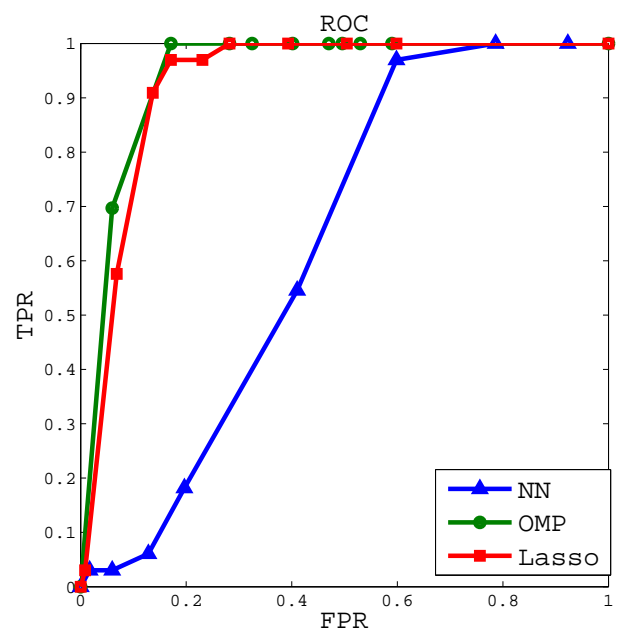


図 3 アルゴリズム別の ROC 曲線。横軸：偽陽性率、縦軸：真陽性率。

めに未学習クラスに着目したときの真陽性率 (TPR) と偽陽性率 (FPR) の値で描いた ROC 曲線を図 3, 4 に示す。図 3 はスパース正則化の解法として OMP, Lasso を用いたときの ROC 曲線を、図 4 では、スパース正則化の解法としてグループ OMP, グループ Lasso を用いたときの ROC 曲線である。ROC 曲線を描くときの TPR, FPR はクラス毎に正解、不正解の  $2 \times 2$  のクロス表で求めている。なお、図 3, 4 において、提案手法と比較するために Nearest Neighbor 法 (NN 法) の ROC 曲線を示す。これは、テストデータと学習データ 1 枚 1 枚の距離を測り、テストデータに最も近い 1 つの学習データのクラスへ識別する手法である。ここでは、距離を閾値とし、ある閾値を越えると未学習クラスに識別している。

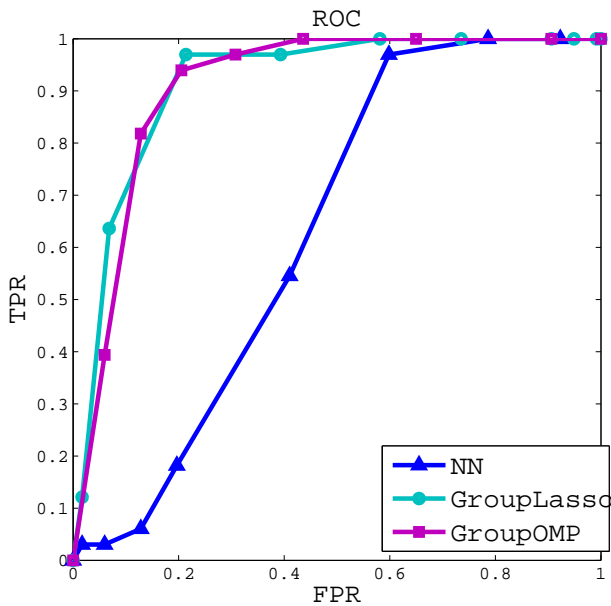


図 4 グループ正則化での ROC 曲線. 横軸: 偽陽性率, 縦軸: 真陽性率.

### 6.3 考察

#### 6.3.1 CCI を用いた各スパース正則化の性能

CCI を用いた手法は, 表 2 からわかるように, 未学習の画像は, 既学習の画像の識別率と同等か, 少し劣る程度の性能を示している. また, 未学習クラスへの識別の有無では, 既学習の画像の識別率がほぼ同等である. これは図 3 や図 4 の ROC 曲線を見るとわかるように, 既知のクラスのテストデータの CCI と未学習クラスのテストデータの CCI では, CCI の閾値  $\tau$  を決める上で適した値を決めやすい傾向があることや, 未学習クラスを識別するための学習データを用意する必要がなく未学習クラスの識別が行えることは CCI を用いる利点であるといえる.

解法に OMP や Lasso を用いたものが, グループ正則化のグループ OMP やグループ Lasso に比べて識別性能が高いことが表 2 からわかる. OMP や Lasso がスパース正則化の解  $\alpha$  をなるべく学習データの枚数を少なくしようとしているのに対して, グループ OMP やグループ Lasso はグループの個数を少なくなるようにしている. 本実験では, 学習データの枚数は 950 枚程度あるのに対し, グループの個数は 30 個と, 学習データの枚数よりも少ない. このように, スパースにする対象の数が違うため, 識別結果に違いが出たと考えられる.

また, 図 3 の ROC 曲線について, OMP では, FPR20%程度で TPR100% を達成する閾値が見つかる. NN 法に比べてスパース正則化による 2 つの手法の FTR が低いうちに TPR は急激に増加している. これは, スパース正則化によりテストデータと無関係なクラスが識別先の候補から強く除外されていることを意味する.

図 4 においても, OMP や Lasso に比べるとグループ OMP やグループ Lasso が TPR100% を達成するときの FPR の値は大きい, NN 法に比べると高い性能を示している. このことから, グループスパース正則化もテストデータと無関係なクラスが識別先の候補から強く除外されていると考えられる.

## 7. おわりに

本研究では, スパース表現で組合せを求めることによる識別において, 未学習クラスを識別する機能を追加し, 知らないデータに対して「知らない」と答える識別の達成とその評価を行った. 解がスパースであるかどうかの指標 CCI を定義し, 高い性能を示すことができた.

CCI を用いることで未学習クラスを識別するために負例を用意する必要がないという利点がある. 本論文では, 「知らない」と答える性能を示し, 全く知らないのか, 該当する上位の概念 (クラス) で識別できるか, 下位の概念で個別に識別できるかなど, 認識の深さや程度を明らかにする方法の一つとして, 学習データの構造化と構造に基づく正則化によって提案手法をさらに拡張できる可能性について 5.1 で議論を行った. この可能性に関する詳しい実験評価を行うべきである.

## 参考文献

- [1] 石井健一郎, 上田修功, 前田英作, 村瀬洋, “わかりやすいパターン認識,” オーム社, 1998.
- [2] 前田英作, “痛快! サポートベクトルマシン - 古くて新しいパターン認識 -”, 情報処理 42 巻 7 号, 2007.
- [3] 高島泰斗, 香田正人, “1 クラス SVM と近傍サポートによる領域判別”, オペレーションズ・リサーチ: 経営の科学, 2006.
- [4] 黒沢由明, “解説 部分空間法の今昔” 情報処理 49 巻 5 号, pp566-572, 2008
- [5] Y. C. Pati, R. Rezaifar and P. S. krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition”, Pacific Grove. CA. USA, 1993.
- [6] B. Efron, T. Hastie, I. Jotnstone and R. Tibshirani, “Least angle regression”, Ann. Statist, 2004.
- [7] Rodolphe Jenatton, Julien Mairal, Guillaume Obozinski and Francis Bach, “Proximal methods for sparse hierarchical dictionary Learning,” INRIA 2010.
- [8] 酒井智弥, “スパース正則化に基づく多重ラベリング,” MIRU2010, 2010.
- [9] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma, “Robust face recognition via sparse representation,” MANUSCRIPT ACCEPTED BY IEEE TRANS, 2008.
- [10] T. Sakai, “Multiple pattern classification by sparse subspace decomposition,” IEEE 2009, 2009.
- [11] F. Bach, “Consistency of the group Lasso and multiple kernel learning”, HAL-00164735, 2008.
- [12] E. J. Candes, “The restricted isometry property and its implications for compressed sensing”, Comptes Rendus Mathematique, 346, 9-10, pp. 589-592, 2008.
- [13] D. Donoho and M. Elad, “Optimally sparse represen-

tation in general(non-orthogonal)dictionaries via  $l^1$  minimization”, Proc. the National Academy of Sciences of the United States of America, pp. 2197-2202, 2003.

- [14] K. C. Lee, J. Ho and D. Kriegman, “Acquiring linear subspaces for face recognition under variable lighting ”, IEEE PAMI, 2005.
- [15] J. Mairal, “SPAMS: a SParse Modeling Software, v2.1”, [http://www.di.ens.fr/willow/SPAMS/doc/doc\\_spams.pdf](http://www.di.ens.fr/willow/SPAMS/doc/doc_spams.pdf), 2011.
- [16] J. Mairal, F. Bach, J. Ponce and G. Sapiro, “Online dictionary learning for sparse coding”, ICML 2009, 2009.