

平成 2 6 年 6 月 1 2 日現在

機関番号 : 1 3 6 0 1

研究種目 : 基盤研究(C)

研究期間 : 2011 ~ 2013

課題番号 : 2 3 5 0 0 1 7 3

研究課題名 ( 和文 ) 再利用可能認識・学習系の構築とその応用に関する研究

研究課題名 ( 英文 ) Study on learning techniques which utilize existing classifiers

研究代表者

丸山 稔 ( MARUYAMA, Minoru )

信州大学・工学部・教授

研究者番号 : 8 0 2 8 3 2 3 2

交付決定額 ( 研究期間全体 ) : ( 直接経費 ) 3,700,000 円、( 間接経費 ) 1,110,000 円

研究成果の概要 ( 和文 ) : 本研究の目的は、既存識別関数系を利用することにより、新規カテゴリ等の認識・識別を効率化・高度化するための手法を確立し、これを画像認識を中心とした問題に適用することである。本研究においては主として画像類似度検索と画像領域分割について手法の検討を行った。画像類似度検索においては、クエリ毎に適切な重み関数を学習する方式において、既存識別系の線形結合を用いる手法の有効性を検証した。実験の結果、高速性では優れているものの、次元低減に基づく手法に精度の点で劣っていることを確認した。画像領域分割においては画像全体を対象とする既存識別系によるカテゴリ制限手法を考案し、手法の有効性を実験により示した。

研究成果の概要 ( 英文 ) : We have investigated methods to utilize still existing classifiers to construct classifier for a novel category. We have studied the methods for similar image search and image segmentation problems. For image search, given a query image, appropriate weight vector is obtained via learning from examples. Although the search quality of the method is very good, it is time consuming. To reduce the computational cost, we proposed a method which approximates weight vectors from existing classifiers (weight vectors). We compared the search quality with the method based on the dimensionality reduction and found PCA-based method can outperform the other approximation techniques. For image segmentation, the task is formulated as the energy minimization. We tried to improve the energy function using the existing image classifier. We have shown the segmentation quality can be improved by limiting possible object categories.

研究分野 : 総合領域

科研費の分科・細目 : 情報学・知能情報学

キーワード : 知能情報処理 パターン認識 画像認識 機械学習 類似度計算 領域分割

## 1. 研究開始当初の背景

現在、文書、画像、動画、音響など多種多様で膨大なデータが入手可能となりつつあるが、これらのデータ群から有用な情報、知識を抽出するためには、例題からの学習に基礎を置く機械学習技術が重要である。例題からの学習、特に識別器の学習において、新規なカテゴリを学習するためには、例題を与え学習を一から繰り返す必要がある。しかしながら、もし類似の特性を有するカテゴリを既に学習してある場合にはこれらの既存識別系を再度利用して、新規カテゴリの学習を高速化・高度化できる可能性があると考えられる。これらの既存識別系利用技術を確立、進展できれば実用上の意義は非常に大きいものと考えられる。

## 2. 研究の目的

本研究の目的は、既存の識別系を利用することにより、新規カテゴリの識別を効率化・高度化するための手法を確立することである。本研究においては、主として画像認識を対象として研究を実施した。

## 3. 研究の方法

研究は、画像認識（画像類似度検索及び画像領域分割）を対象として実施した。画像類似度検索は検索クエリ画像を一時的な新規カテゴリとした識別問題と考えることができる。本研究ではクエリ毎に線形識別関数を構築する手法を高速化するために、既存識別関数の線形結合に基づく手法を検討し、他の高速化手法との比較を試みた。また画像領域分割を行う際に、画像全体を対象とする既存識別系を利用して識別精度の向上を図る手法の検討を行い、有効性を検証した。

## 4. 研究成果

## 4. 1 画像検索の高速化への応用

画像検索において、画像間の類似度を計算する典型的な手法としては、画像を特徴ベクトルで表現し、ベクトル間の距離や内積を類似度として用いることが考えられる。近年では、画像検索においてクエリのベクトルに対してベクトルの成分のユニークさを示す重みを学習し、これを用いて画像検索を行う手法が提案されている。この手法では画像検索にて良好な性能を示すことが報告されているが、クエリ画像が与えられる度に大量のサンプルを用いた SVM 学習が必要になり、計算コストがとてつもないことが欠点である。本研究では、重み成分を高速に計算するための手法として、次元削減による手法を提案し、その有用性について検証を行った。

## 4. 1. 1. クエリ毎の類似度関数設定

クエリに基づく類似度関数では、クエリ画像の特徴ベクトルを、検索対象のデータベース中の任意画像の特徴ベクトルをとした時、類似度は以下のように定義される。

$$\text{query-based similarity}(q, x) = w_q^T x$$

ここに重み成分のベクトルである  $w_q$  は線形 SVM の学習によって与えられる。SVM 学習では、以下の目的関数の最小化を行う。

$$L(w_q) = \sum_{x_i \in P \cup q} h(w_q^T x_i) + \sum_{x_j \in N} h(-w_q^T x_j) + \lambda \|w_q\|^2$$

$$h(x) = \max(0, 1 - x)$$

上式において、P はクエリベクトルに対して小さな幾何変換を加えた小数のベクトル集合であり、ポジティブクラスの学習データとして設定される。一方、N は無作為に選択された大量(1 万枚)の画像の特徴ベクトルの集合であり、ネガティブクラスの学習データとして設定されている。関数  $h(x)$  はヒンジ損失関数である。また、本研究では線形 SVM の学習に LIBSVM を用いた。

## 4. 1. 2. 既存識別関数を用いた高速化手法の検討

画像類似度計算においては、与えられたクエリは一時的な新規カテゴリとみなすことができる。従ってクエリ毎の SVM 学習は新規カテゴリに関する学習問題と考えることができる。このとき計算時間の高速化のために、既存識別系、即ち、過去に検索対象として与えられたクエリに対して学習された重みベクトルを利用することが考えられる。本研究では予めいくつかの画像について重みベクトルを計算しておき、それらの線形結合によって近似的に重みベクトルを計算する。また、線形結合を行う際には、各重みベクトルに対して重み付けを行う。それぞれの重みの計算方法はクエリベクトルとの内積を用いた。

$$w_q = \sum_{i=1}^N \alpha_i(q) w_i \quad (\alpha_i(q) = q^T w_i)$$

既存識別関数系の利用方法以外の選択肢として、SVM の学習時間を短縮するためには、学習対象のデータを削減することも考えられる。本研究では学習時にデータベース中の要素全てを用いるのではなく、クエリに対して内積を用いた類似度の高いデータをサンプリングし、それらのみを用いて SVM を学習する手法も比較のために用いた。

## 4. 1. 3. 次元削減による高速化

SVM 学習の高速化として、データの次元削減を適用することが考えられる。しかしながら、SVM 学習の計算時間は、与えられる学習データ数とデータの次元に単純に依存しておらず、与えられた学習データの分布等によって変化することが考えられるため、次元削減は必ずしも SVM 学習計算の高速化を保障するものではないが、内部で各データを扱う際の計

算コストは下がるため、それに応じた高速化が行われることを期待し適用を行った。本研究では、次元削減手法として autoencoder と PCA(principal component analysis)を用いた。

autoencoder とは、最も単純なモデルでは 3 つの層からなる多層ニューラルネットワークであり、各層はそれぞれ入力層、隠れ層、出力層と呼ばれる。入力層と隠れ層の部分は「Encoder」と呼ばれる。「Encoder」では、入力層にベクトル が与えられる時、隠れ層のベクトル は、ネットワークのパラメータ、を用いて以下の式から求められる

$$y = \sigma(Wx + b)$$

ここに  $\sigma$  はシグモイド関数である。シグモイド関数は非線型の関数であり、線型で次元削減データを表現する PCA と異なる。また、「Encoder」では隠れ層の次元数を入力層より少なくすることで次元削減が行われる。

隠れ層と出力層の部分を「Decoder」と呼び、隠れ層のデータを再び入力層と同じ次元のベクトルへ変換する。隠れ層のベクトルから出力層のベクトル を計算するには、以下の式を用いればよい。

$$z = \sigma(W'y + b')$$

隠れ層の計算に必要なネットワークのパラメータは学習によって求められる。学習では入力層のデータが出力層で復元されるように、以下のクロスエントロピー誤差関数を最小化する形で求められる。

$$L(x, z) = -\sum_{i=1}^d \{x_i \log z_i + (1 - x_i) \log(1 - z_i)\}$$

#### 4. 1. 4. 比較実験結果

本研究で検討した高速化のための各手法の有効性を検証・比較するために、画像検索実験を行った。まず、まず、次元削減手法により学習の高速化実現出来るかを検証するため、元データと次元削減後のデータを用いた SVM 学習時間の比較を行った。次に、各手法検索性能を比較するため、画像検索実験を行った。

画像検索における検索精度は、各画像を表現する特徴ベクトルの性質に依存することが考えられる。本研究では 画像認識の際によく用いられる HOG 特徴量と D-SIFT(dense sampling SIFT)特徴量を用い、まず両者による検索結果の比較を行った。

図 1 はそれぞれの特徴量を用いた場合の検索例である。図において a, b, c, d はそれぞれ、a:内積、b:クエリ毎の SVM 学習に基づく検索、c:既存識別系の線形結合モデルによる検索、d:学習データのサンプリングに基づく SVM 学習による検索、を示している。本研究ではこの事前実験に基づき、以下では特徴ベクトル表現として D-SIFT を使用する。

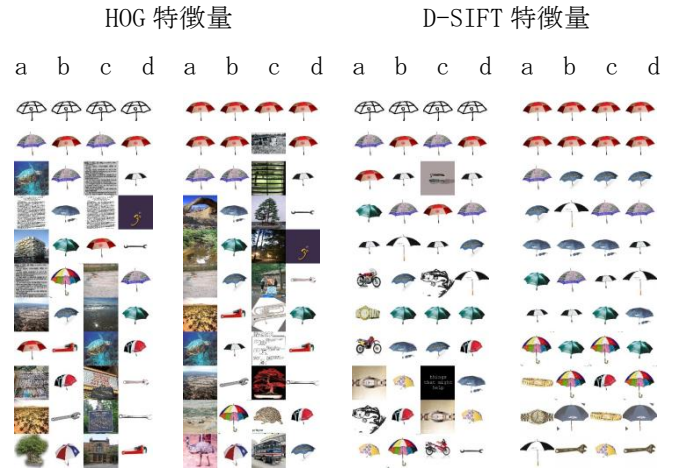


図 1. 特徴量による検索結果の比較例

#### ・ SVM 学習時間の比較

手法の性質より、SVM 学習を用いない既存識別系の線形結合手法が最も高速であることは明らかである。他の手法による SVM 学習の高速化の効果を比較するために、各手法に基づく平均計算時間 (SVM 学習時間) を表 1. に示す。学習サンプルの削減や次元削減手法の適用により、SVM 学習が高速化できていることが分かる。また、100 次元と 500 次元の場合を比較すると、より次元の大きな 500 次元が平均的には高速に学習が終了することが分かり、学習における計算時間の次元数への依存が単純な線形関係でないことも確認できる。

表 1. SVM 学習の所要時間

次元削減手法	次元	平均計算時間 (秒)
-	8192	11.63
PCA	500	0.335
PCA	100	0.925
autoencoder	500	0.335
autoencoder	100	0.925
学習データ減	8192	0.260

#### ・ 検索精度の比較実験

手法の効果を比較するためには計算時間だけではなく、検索精度を比較する必要がある。検索精度評価のために以下のような実験を行った：①カテゴリ分けされた画像データ群を用意する。②上記画像データ群を含む拡大画像 DB を作成する。③カテゴリを決定し、同カテゴリ群中の画像からクエリを選択し画像検索を実施する。④類似度が上位と判定された画像について、そのカテゴリを確認し、クエリと同一のカテゴリに所属する画像数をカウントし、以下のような比率を計算する

(同一カテゴリ所属画像数) / (上位判定画像数)

このとき、上位判定画像数は 5, 10, 15, 20 の 4 通りについて実験を行う。実験においてはカテゴリ分けされた画像集合としては Caltech101 を用いた。拡大画像 DB は大量のランダムな画像 1 万枚と上記画像集合から構成した。

まず、通常の画像を対象として実験を行った結果を表 2 に示す。実験においてはクエリ画像は Caltech101 の各カテゴリから 3 枚をランダムに選択した。表 2 より、計算時間の点では最速の既存識別系の利用に基づく手法は、識別精度の点では次元削減に基づく手法、特に PCA に基づく手法に劣ることが分かる。次元削減前の特徴次元は 8192 であるが、次元削減により 100 まで減少させても、PCA+SVM による手法は良好な識別性能を示している。

クエリ毎の SVM 学習の場合、クエリ画像と検索対象画像の 픽셀情報が異なっている場合でも良好な検索性能を示すことが期待できる。このようなクエリと DB の画像性質が異なる場合の検索能力を比較するために、スケッチ画像を用いて上の実験で用いた DB の検索を行った。



図 2. クエリスケッチ画像の例

図 2 にはクエリとして用いたスケッチ画像の例を示す。実験結果の評価は通常画像による検索実験と同様である。実験結果を表 3. に示す。この実験でも PCA による次元削減と SVM 学習を用いた手法が比較的良好な精度を示した。

#### 4. 2 画像領域分割への応用

画像認識には様々なものが考えられるが、代表的なものとして前節で取り扱ったような画像類似度認識や画像カテゴリ認識が挙げられる。これらの場合、対象となるのは画像全体であり、画像全体の特徴量を用いて、他の画像との類似度やカテゴリラベルを出力することがある。これに対して、画像に含まれる各オブジェクトを抽出し、それらのカテゴリを認識する問題は画像領域分割 (セグメンテーション) と呼ばれる。本節では、画像全体に関する認識を行う識別系が得られているとき、それを再利用して画像領域分割の精度を向上させることを考える。

表 2. 画像検索実験結果 (識別精度)

手法	次元	5	10	15	20
内積	8192	0.61	0.42	0.33	0.29
SVM	8192	0.65	0.48	0.39	0.36
線型結合	8192	0.60	0.40	0.32	0.28
学習データ減	8192	0.63	0.44	0.37	0.33
PCA	500	0.63	0.45	0.36	0.32
PCA-SVM	500	0.67	0.49	0.41	0.37
autoencoder	500	0.61	0.42	0.33	0.29
autoencoder-SVM	500	0.64	0.46	0.38	0.34
PCA	100	0.61	0.42	0.38	0.34
PCA-SVM	100	0.64	0.46	0.38	0.34
autoencoder	100	0.58	0.39	0.31	0.28
autoencoder-SVM	100	0.59	0.41	0.33	0.29

表 3. スケッチ画像をクエリとした場合の画像検索実験結果 (識別精度)

手法	次元	5	10	15	20
内積	8192	0.83	0.76	0.68	0.62
SVM	8192	0.89	0.77	0.71	0.67
線型結合	8192	0.83	0.74	0.66	0.59
学習データ減	8192	0.77	0.74	0.68	0.64
PCA	500	0.83	0.76	0.68	0.61
PCA-SVM	500	0.86	0.77	0.74	0.69
autoencoder	500	0.83	0.69	0.60	0.57
autoencoder-SVM	500	0.86	0.73	0.62	0.58
PCA	100	0.74	0.69	0.62	0.57
PCA-SVM	100	0.80	0.71	0.68	0.64
autoencoder	100	0.74	0.64	0.55	0.51
autoencoder-SVM	100	0.66	0.60	0.54	0.54

##### 4. 2. 1. 画像領域分割の定式化

領域分割は 픽셀値が類似する領域を一つの領域とするスーパー 픽셀単位で、エネルギー関数最小化に基づいて行われる。領域分割のために、まずにスーパー 픽セルを対象としたクラス識別を行う識別器を例題からの学習によって構築する。識別のための特徴量としては、SIFT, Self Similarity Descriptor, 色 (RGB) を用いる。スーパー 픽セル内の各 픽셀について抽出されたこれらの特徴量から k-means 法によってベクトル量子化 (visual words 抽出) を行い、ヒストグラム (bag of words 表現) を生成する。このヒストグラムを各スーパー 픽セルの特徴量として識別器 (RBF-SVM) を学習した。

スーパーピクセルに対してクラスラベルを割り当てるためにエネルギー関数  $E(\mathbf{x})$  を以下のように定義する.

$$E(\mathbf{x}) = \sum_{i \in V} \varphi_i(x_i) + \sum_{(i,j) \in \epsilon} \psi_{ij}(x_i, x_j) + \sum_{i \in V} g_i(x_i)$$

$\mathbf{x}$  はスーパーピクセル列に割り当てられるクラスラベル集合を表す.  $\varphi_i(x_i)$  はスーパーピクセル  $i$  にラベル  $x_i$  を与えたとき, コストが発生する unary potential である.  $\psi_{ij}(x_i, x_j)$  は隣接するスーパーピクセル  $i, j$  に異なるラベルが与えられたときのコストを表す smoothness potential である.  $g_i(x_i)$  は類似画像によるクラス制限項である. unary potential は以下で定義する.

$$\varphi_i(x_i)$$

$$= -K_i \left( \log \frac{1}{1 + \exp(-a(x_i) \text{score}(x_i; \mathbf{z}_i) - b(x_i))} \right)$$

$$+ \log \text{loc}(\hat{\mathbf{i}}, x_i)$$

ここに  $K_i$  はスーパーピクセル  $i$  内のピクセル数である.  $a(x_i), b(x_i)$  は学習により最適化されるパラメータである.  $\text{score}(x_i; \mathbf{z}_i)$  はスーパーピクセル  $i$  から生成される特徴ベクトル  $\mathbf{z}_i$  を, クラス  $x_i$  の識別器に入力したときの出力値である.  $\text{loc}(\hat{\mathbf{i}}, x_i)$  はクラス  $x_i$  がスーパーピクセルの中心座標  $\hat{\mathbf{i}}$  に出現する確率である. Smoothness potential  $\psi_{ij}(x_i, x_j)$  は以下のように定義する.

$$\psi_{ij}(x_i, x_j) = \left\{ a(x_i, x_j) K_{ij} \exp \left( \frac{-|c_i - c_j|^2}{b} \right) \right\} T[x_i \neq x_j]$$

$$a(x_i, x_j) = \begin{cases} \lambda & f[x_i][x_j] > 0 \\ \mu & f[x_i][x_j] = 0 \end{cases} \quad \frac{\mu}{2} < \lambda \leq \mu$$

ここに  $\mathbf{c}_i, \mathbf{c}_j$  はスーパーピクセル内の RGB 値の平均を表し,  $K_{ij}$  は隣接するスーパーピクセル  $i, j$  間で, 境界を共有するピクセル数を表す.  $f$  は隣接スーパーピクセル間でクラスの共起性を表すパラメータであり,  $f[x_i][x_j] > 0$  であれば, 共起性ありを示し, 共起性なしのパラメータ  $\mu$  より低い値を  $\lambda$  に設定する. これにより隣接領域間では共起関係にあるクラスが選択されやすくなることが期待出来る.

これらの項に加え, 本研究では画像類似度検索のための既存識別関数の出力に基づく類似画像によるクラス制限項を加え, 領域分割能力の向上を図った. 類似画像によるクラス制限項  $g_i(x_i)$  は以下で定義する.

$$g_i(x_i) = KT[x_i \notin X_g]$$

類似画像によるクラス制限項は画像内で出現している物体のクラスを類似画像より推測し, 推測されたクラスを認識しやすくする項である.  $X_g$  は類似画像により許容されるクラスの集合である. 許容クラス外のクラスがラベリングされるときペナルティ  $K$  が発生す

る. これにより, 許容クラス内のクラスがラベリングされやすくなることが期待される.

類似画像によるクラス制限ではすべてのスーパーピクセルに対して許容されるクラスが同一である. そのため, スーパーピクセルが許容クラス内のクラスで誤認識される問題が存在する. そこでスーパーピクセルごとに許容クラスを変化させる類似スーパーピクセルによるクラス制限も導入した. これは, 認識対象のスーパーピクセルに対して類似するスーパーピクセルを類似画像から取得し, スーパーピクセルごとに許容クラスを設定するものであり, エネルギー関数に以下の項を追加することで実現される.

$$s_i(x_i) = RT[x_i \notin X_i]$$

ここに  $X_i$  はスーパーピクセル  $i$  の類似スーパーピクセルにより許容されたクラスの集合である. 許容クラス外のクラスがラベリングされるときペナルティ  $R$  が発生する.

#### 4. 2. 2. 画像領域分割実験結果

提案手法の有効性を検証するために MSRC21 dataset を利用した検証実験を行った. 学習においては 296 枚に画像, 認識には 296 枚の画像を用いた. カテゴリ数は 21 であり, 類似画像の選出数  $k=5$ , 類似スーパーピクセルの選出数を  $t=10$  とした. 画像領域分割・認識実験結果の精度を表 4 に示す. 類似画像に関する制限項までを含むエネルギー関数を用いる手法を baseline とし, 類似スーパーピクセルによるクラス制限項を加えて拡張した関数を用いる手法を  $s$  とした. 画像類似度検索に関する既存識別系を用いることにより, 若干の精度向上が実現できることが分かる.

表 4. 画像認識実験結果 (認識率%)

カテゴリ	baseline	s
building	78.4	79.4
grass	90.7	92.2
tree	88.0	87.9
cow	62.9	63.9
sheep	68.1	69.2
sky	91.6	91.3
aeroplane	69.5	69.1
water	70.3	73.8
face	81.4	81.1
car	71.9	77.7
bicycle	76.2	79.0
flower	78.7	84.3
sign	21.7	22.5
bird	35.3	36.8



book	89.3	90.5
chair	56.8	60.6
road	80.8	80.5
cat	33.3	33.7
dog	61.2	61.5
body	54.4	53.0
boat	30.5	35.6
mean	66.2	67.8

## 5. 主な発表論文等

〔雑誌論文〕(計 1 件)

- ① Takuma Yamaguchi and Minoru Maruyama, Character type classification via probabilistic topic model, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 5, No. 2, pp. 123 - 140 (2012). (査読有)

〔学会発表〕(計 11 件)

- ① 宮下徹也, 丸山稔, 宮尾秀俊, 重みベクトルの学習に基づく類似画像検索の性能, 電子情報通信学会信越支部大会講演論文集, 5D-3, pp. 85, 10月5日, 長岡技術科学大(2013). (査読無)
- ② 染谷昌史, 宮尾秀俊, 丸山稔, クラス制限に基づく画像中のオブジェクト認識能力の向上, 電子情報通信学会信越支部大会講演論文集, 5D-4, pp. 86, 10月5日, 長岡技術科学大(2013). (査読無)
- ③ 桜井翔, 宮尾秀俊, 丸山稔, 確率的トピックモデルを用いた複数種類の論文データの解析, 電子情報通信学会信越支部大会講演論文集, 10B-1, pp. 165, 10月5日, 長岡技術科学大(2013). (査読無)
- ④ Tsubasa Yagi, Hidetoshi Miyao and Minoru Maruyama, Robust on-line handwritten object retrieval using a similarity measure based on overlap area of circles, Proc. 21st International Conference on Pattern Recognition (ICPR2012), pp. 1 - 4, November 14, Tsukuba (2012). (査読有)
- ⑤ 井本吉彦, 丸山稔, 画像カテゴリ認識のための特徴量組み合わせ手法の検討, 電子情報通信学会信越支部大会講演論文集, 5B-2, 10月13日, 新潟大 pp. 82 (2012). (査読無)
- ⑥ 小田誠也, 丸山稔, 識別関数出力を特徴量に用いる顔画像認識手法の検討, 電子情報通信学会信越支部大会講演論文集, 5C-1, pp. 85, 10月13日, 新潟大(2012). (査読無)
- ⑦ 染谷昌史, 丸山稔, オブジェクト相互関係を考慮したポテンシャル関数による画

像認識, 電子情報通信学会信越支部大会講演論文集, 5C-3, pp. 87, 10月13日, 新潟大(2012). (査読無)

- ⑧ 宮下徹也, 丸山稔, 重みベクトルの学習に基づく類似画像検索の高速化, 電子情報通信学会信越支部大会講演論文集, 6C-1, pp. 101, 10月13日, 新潟大(2012). (査読無)
- ⑨ 井本吉彦, 丸山稔, SVMを用いた画像カテゴリ認識のための特徴量の検討, 電子情報通信学会信越支部大会講演論文集, 5A-1, pp. 76, 10月8日, 新潟工科大(2011). (査読無)
- ⑩ 小田誠也, 丸山稔, kNN法を用いた画像のカテゴリ認識, 電子情報通信学会信越支部大会講演論文集, 5C-1, pp. 86, 10月8日, 新潟工科大(2011). (査読無)
- ⑪ 和田誠也, 丸山稔, 高次元空間内における近接データ検索手法の比較, 電子情報通信学会信越支部大会講演論文集, 3D-5, pp. 58, 10月8日, 新潟工科大(2011). (査読無)

〔その他〕

<http://soar-rd.shinshu-u.ac.jp/profile/ja.WCnCbpkh.html>

## 6. 研究組織

### (1) 研究代表者

丸山 稔 (MARUYAMA, Minoru)

信州大学・工学部・教授

研究者番号: 80283232

### (2) 研究分担者 なし

### (3) 連携研究者 なし