

ユーザとの相互作用を通じた画像分類システムの開発

Image Classification System based on Interaction with Human

岡田 将吾 神田 賢一 大本 義正 西田 豊明
Shogo Okada Kenichi Kanda Yoshimasa Ohmoto Toyoaki Nishida

京都大学情報学研究科知能情報学専攻

Dept. of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

京都大学工学部情報学科

School of Informatics and Mathematical Science Faculty of Engineering, Kyoto University

The goal of our research is to achieve a classification system of multimedia data based on interaction with users. To help achieve this goal, we propose a novel Semi-supervised Classification (SSC) algorithm : Graph based SSL with Metric Learning (GSSCML) which is a fundamental mechanism of the system. Proposed algorithm:GSSCML integrates the graph based SSC algorithm and Semi-supervised Adaptive Metric Learning (SSAML) algorithm. The graph is constructed by using metric which is optimized by SSAML, and the graph based SSL with SSAML gives higher classification accuracy than other graph based approach. Experimental results on benchmark datasets show that the proposed algorithm is significantly better than the graph based semi-supervised learning algorithms in terms of classification accuracy. 1 to 11 percent is attributed to the better and improvement in classification accuracy than the state of the art graph based approaches. We develop the interactive image classification system based on GSSMLML algorithm, too.

1. はじめに

デジタル技術や通信技術の進歩やデジタルカメラの普及により、大量の画像データのやりとりが可能となり、個々のPC上に蓄積される画像データは大容量となる傾向にある。蓄積された大量の画像データの分類には手間がかかる上、見通しの良い形で整理されていなければ、これらを活用することも難しくなる。従って、これらの画像データをいかにして効率的に見通しの良い形で分類するかは重要な問題の一つである。

自動的に分類するための技術の一つとして教師なし学習が挙げられる。一般的な教師なし学習手法であるクラスタリングでは予めオブジェクト間の距離を定義しておき、その距離に応じた基準に基づきオブジェクトを各カテゴリに分類する。この距離の尺度はユーザによって異なるため、予め設定した距離空間でクラスタリングした分類結果が必ずしもユーザの期待する分類結果と一致する保証はない。このためユーザは教師ラベルを与える必要があるが、全てのデータにラベルを付与するのは手間がかかるため、教師ラベルなしデータを学習に利用可能な半教師付き学習 [Chapelle 06] は有用である。

本研究ではユーザが最小限の教示 (少数のラベル付きデータ) に基づいて大量の画像データを分類可能な半教師学習手法:Graph based Semi Supervised Classification with Metric Learning (GSSCML) を提案し、この手法に基づく画像分類システムを構築する。

GSSCML はサンプル間の類似度行列を、半教師付き距離学習によって最適化した距離空間に基づき定義し、類似度グラフを学習することで、画像のような高次元の入力データに対して頑健な分類精度を実現する。距離学習には [Ye 07] で提案されている Adaptive Metric Learning (AML) を用いる。グラフ

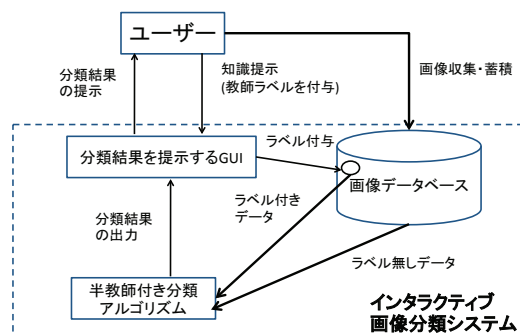


図 1: 提案するシステムの概要

に基づく半教師付き学習には Multiclass Transductive Learning with Graph (RMGT)[Liu 09] を用いる。

評価実験では、複数のデータセットを用いて、半教師付き分類タスクを行い、従来手法との比較を行い、GSSMLML の分類精度における優位性を示す。また GSSMLML と分類結果をユーザに提示する GUI を組み合わせ、相互作用に基づく画像分類システムを提案し、画像分類に関する実験を行う。

2. 提案システムの概要

提案システムの概要を図 1 に示す。システムは画像分類結果を提示する GUI、画像データベース、半教師付き学習システムの三つのモジュールで構成されている。ユーザが蓄積した画像データはデータベースに蓄積され、GUI を通じてユーザは画像の分類状況を閲覧可能である。ユーザは代表的な画像データを指示し教師ラベルを付与する。このラベル付きデータに基づき大量のラベル無しデータは分類され、画像データは分類される。分類結果は GUI に反映され、ユーザは分類結

連絡先: 岡田将吾, 京都大学 情報学研究科知能情報学専攻,
京都市左京区吉田本町 (工学部 10 号館 214 号室), 075-
753-5371, 075-753-4961, okada.s@i.kyoto-u.ac.jp

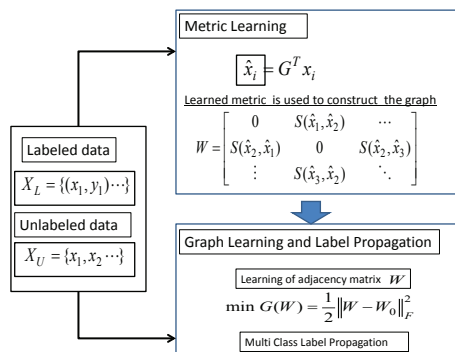


図 2: 提案する半教師付き学習手法の概要

果を閲覧する。ユーザはさらに分類したい部分に教示を与える。この操作を繰り返すことにより画像の分類が行われる。

3. 半教師付き学習アルゴリズム

提案手法は距離学習に基づく類似度グラフの構築 (3.1 節) と類似度グラフの学習 (3.2 節) の二つの処理により成り立っており、ラベル付きデータとラベルなしデータは両方の処理に利用される。提案手法の概念図を図 2 に示す。

3.1 距離学習に基づく類似度グラフの構築

3.1.1 AML による距離学習

本手法はクラスタリングにおける最適化の枠組みに距離学習が組み込まれている。

まず X を n 個のサンプルで構成されるデータセットとすると、 $\{x_j\}_{j=1}^n \in \mathbb{R}^m$ と表せる。また m 次元空間を l 次元空間に変換する線形変換行列を $G \in \mathbb{R}^{m \times l}$ と定義すると以下の式が導ける。

$$G : x_i \in \mathbb{R}^m \rightarrow \hat{x}_i = G^T x_i \in \mathbb{R}^l (l < m) \quad (1)$$

G は直行行列 ($G^T G = I_l$) とする。サンプル間の距離計算には以下の式に示すマハラノビス距離を用いる。

$$d_M(\hat{x}_i, \hat{x}_j) = \sqrt{(\hat{x}_i - \hat{x}_j)^T \hat{\Sigma}^{-1} (\hat{x}_i - \hat{x}_j)} \quad (2)$$

AML ではクラス間分散 (Sum of Squared Intra Error (SSIE)) の最大化を目標として学習が行われる。各クラスへの各サンプルの割り当てを記述する行列 $L \in \mathbb{R}^{n \times K}$ を用いて、SSIE は行列の積の形式で以下のように書ける。

$$SSIE(\{C_j\}_{j=1}^K) = \frac{1}{n} \text{tr}(L^T X^T G (\hat{\Sigma})^{-1} G^T X L). \quad (3)$$

式 (3) で $\text{tr}()$ は行列の対角成分の和を算出する演算子である。AML の目的関数 SSIE は線形変換行列 G と各サンプルの各カテゴリへの割り当て L の二つの変数を持つ。SSIE の最大化問題を解くために、EM アルゴリズムの枠組みで G と L を交互に最適化する。

G の最適化は固有値問題、 L の最適化は kernel K-means の問題に置き換えられる。ラベル付きデータ集合の内、同じラベルを持つデータは同じクラスへ、異なるラベルのデータは異なるクラスに割り当てるように kernel K-means に制約を加える [Wagstaff 01]。

3.1.2 学習された距離関数を用いたグラフの構築

サンプルセット $X = \{x_1, \dots, x_l, \dots, x_n\} \subset \mathbb{R}^m$ において、最初の l 個のサンプルをラベル付きデータ、残りの $u = n - l$ 個のサンプルをラベルなしデータとする。 X を用いて非対称の $n \times n$ 類似度グラフを下記のように定義する。

$$A_{i,j} = \begin{cases} \exp(-\frac{d_M(\hat{x}_i, \hat{x}_j)^2}{\sigma^2}), & \text{if } j \in \mathcal{N}_i \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

式 (4) において \mathcal{N}_i は x_i の k 最近傍のサンプルの集合である。 σ は $\sigma = \sum_{i=1}^n d_M(x_i, x_{ik})$ として計算される。ここで x_{ik} は x_i の k 番目の最近傍サンプルである。

式 (4) における $d_m()$ の G のサイズは $m \times l_p$ ($l_p \geq K - 1$) と定義する。適切な l_p を探索によって求め、扱うデータに応じて G のサイズを変更する。

A に基づき類似度行列を $W : W = A + A^T$ と定義する。 W は対称性を有し、 $W_{ii} = 0$ とする。

3.2 グラフに基づく半教師付き分類

提案手法では類似度グラフ W の学習に RMGT [Liu 09] を用いる。

3.2.1 類似度行列 W の学習

RMGT における類似度行列 W の学習では異なるカテゴリに属する教師データの制約 (not link) のみを用いる。 W の学習は異なるラベルを持つデータの制約を利用し、以下のように定式化出来る。

$$\begin{aligned} \min \mathcal{G}(W) &= \frac{1}{2} \|W - W^0\|_F^2 \\ \text{s.t.} \quad &\sum_{(i,j) \in \mathcal{T}} W_{ij} = 0, \\ &W \mathbf{1} = \mathbf{1}, W = W^T, W \geq 0, \end{aligned} \quad (5)$$

式 (5) において $\|\cdot\|_F$ は Frobenius ノルムであり、式 (5) は Quadratic Programming (QP) 問題の一種と見なせる。QP 問題を効率的に解くために、式 (5) は二つの凸最適化問題 $\text{s.t.} \sum_{(i,j) \in \mathcal{T}} W_{ij} = 0, W \mathbf{1} = \mathbf{1}, W = W^T$ と $\text{s.t.} W \geq 0$ に分割され、最適化される。

最後にラベル付きデータからラベル無しデータへラベルを伝播する。各要素は 0 から 1 の間の実数値を取るラベルなしデータのクラス割り当て行列を F_u とし、これを求める。多クラス分類問題のラベル伝播アルゴリズムにおける損失関数は $Q(F) = \text{tr}(F^T L F)$ と定式化出来る。グラフラプシアン L を $L = D - W$ として計算した後、部分行列形式を $L = \begin{bmatrix} L_{ll} & L_{lu} \\ L_{ul} & L_{uu} \end{bmatrix}$ と書き下し、要素が全て 1 である u 次元と l 次元のベクトルをそれぞれ $\mathbf{1}_l, \mathbf{1}_u$ とすると Q を最小化する F_u は次式で求められる。

$$F_u = -L_{uu}^{-1} L_{ul} \mathbf{Y}_l + \frac{\mathbf{1}_{uu}^{-1} \mathbf{1}_u}{\mathbf{1}_u^T L_{uu}^{-1} \mathbf{1}_u} (n \omega^T - \mathbf{1}_l^T \mathbf{Y}_l - \mathbf{1}_u^T F_u^0) \quad (6)$$

ここで F_u^0 は [Zhu 03] で提案された調和関数を多クラス分類用に拡張した関数であり、 ω は各クラスにサンプルが存在する事前確率であり、本研究では均等に $1/K$ とした。

以上をまとめ、提案する半教師付き学習アルゴリズムを表 2 に示す。

表 1: Segment データセットと YaleB データセットに対する誤分類率

Error Rate (%)	Segment		Yale Face B		USPS	
	14 labels	100 labels	20 labels	100 labels	20 labels	100 labels
AMLCC	35.64±3.60	34.14± 2.48	11.40±4.10	0.27± 0.46	28.90 ±3.57	28.60± 1.30
RMGT(W)	29.45± 5.10	12.80 ± 1.61	13.04±8.08	0.26± 0.56	22.53± 4.49	12.18 ± 1.54
GSSCML	24.90±5.79	12.80± 1.50	1.82±4.17	0.25± 0.34	21.29±4.56	10.94± 1.39

表 2: AML に基づく半教師付き学習アルゴリズム

Input: X, K, Y_l
Output: Y_u
Step 1. AML を全サンプルに対し実行して変換行列 G と共分散行列 Σ を求める。ここで与えられた l 個のラベル付きデータをクラスタリングの制約に用いる。
Step 2. 対称型類似度行列 W_0 を, AML により最適化した距離関数 $d_M(\hat{x}_i, \hat{x}_j)$ に従い計算する (3.1.2 節)。
Step 3. W を 3.2.1 節の手法に従い学習する。
Step 4. 式 (6) に基づきラベル情報を教師なしデータに伝播させる。
Step 5. Step 2. から Step 4. を繰り返し, Q を最小化する l_p^* を探索する。

4. 評価実験

4.1 半教師付き分類精度に関する評価実験

本章では提案手法:GSSCML の評価実験の詳細を述べる。実験に用いたベンチマークデータセットは UCI Machine Learning Repository [C.Blake 98] の Segment データセット (総数 2310, 7 クラス, 19 次元), Yale Face B データセット (総数 5850, 10 クラス, 1200 次元), USPS 手書き文字データセットのテストセット (総数 2007, 10 クラス, 256 次元) の 3 種類である。ここで YaleB 顔画像データセットについては原画像のサイズ 648×480 から 40×30 に画像サイズを縮小した。

教師ラベルデータは各クラスにつき最低 1 データ以上含む条件下で実験を行う。全データの内ラベルデータの集合をランダムに選択し, 計 40 回の実験を行った後に平均の誤分類率を算出する。比較手法には RMGT と制約付きクラスタリングに基づく AML (AMLCC) を用いた。ここで RMGT については 3.2 節と同様の方法で学習を行う。グラフの構築における距離 (式 (4) の d) にはユークリッド距離が用いられる。

AML と GSSCML に共通して用いられる共分散行列の平滑化パラメータを $\lambda = 100$ と設定し, RMGT と GSSCML に共通して用いられる k -近傍グラフのパラメータを $k = 20$ と設定し, 全てのデータセットに対し実験を行った。

4.1.1 実験結果

各データセットに対する分類実験結果を表 1 に示す。いずれのデータセットについても, ラベルデータが少ない場合, 多い場合の 2 種類について実験を行った。表 1 よりいずれのデータセットに対しても GSSCML の分類精度は比較手法のそれを上回っていることが示された。

ラベルデータが多い場合, GSSCML は RMGT は Segment, YaleB データセットに対し分類精度は同等 (t 検定の結果, 5% 水準で有意差なし) であった。ラベルデータが少ない場合, Segment で 4%, YaleB で 11%, USPS で 1%, GSSCML は RMGT より分類精度を改善した。[Liu 09] において RMGT は他のグラフに基づく手法よりも多クラス分類の精度が優れていると

報告されていることと, 上記の実験結果より GSSCML は少数の訓練データを利用して分類精度を改善可能であることが示された。

5. インタラクティブ画像分類実験

本節では, 提案する半教師付き学習手法を搭載したシステムにより, インタラクティブな画像分類実験を行う。画像データには CORELL データセットから鹿, 飛行機, ペンギン, 船, 鳥, 虎, 犬, 馬, 冰山, 男性, 女性のカテゴリを選択し, この中から計 120 枚の画像を選択して用いた。画像データを学習空間にランダムに配置した状態 (図 3) から, ユーザは教師ラベルの付与を行い, 画像の分類を行う。

5.1 画像の特徴量

画像の特徴量には, アピアランス特徴として SIFT, 色特徴としてカラーテクスチャモーメント, カラーヒストグラムを用いた。SIFT は Bag of Words の方法により 300 次元として求めた。カラーヒストグラムは 64 次元, カラーテクスチャモーメントは 48 次元とし, 計 412 次元の特徴量を抽出した。

5.2 画像の配置方法

[神田 10] で提案された GUI を用い, 2 次元空間上に画像を配置する。高次元データの二次元上へのマッピング方法として ISOMAP やラプシアン固有マップ法 [Yang 06] など考えられるが, ここでは半教師付き分類の結果を直接配置に反映させるため, 中心から半径 R の円周上に等角度間隔で分類後の各クラスを配置した。

5.3 画像分類例

インタラクティブな画像分類の様子を示すために, 以下のシナリオで分類を行う。

手順 1 最初に三枚の画像に対しラベルを付与し, 半教師付き分類後の結果をユーザに提示する。

手順 2 さらに分類したい画像についてラベルを付与し半教師分類を行い結果をユーザに提示する。

手順 1 後の GUI の様子を図 4 に示す。ここではペンギン, 冰山, 男性のカテゴリから 1 サンプルずつにラベルを付与した。この三サンプルを代表とするクラスが 3 つ生成された。1 番目のクラスは主にペンギンを含むクラスであるが, 鹿, 飛行機, 船, 冰山, 男性, 女性などの画像も混同している。2 クラス目は鹿以外全てのカテゴリデータを含むクラスである。3 クラス目は男性, 女性の画像が 7 割を占めるクラスである。

次に分類結果を閲覧した後, ユーザは 17 点のラベルをデータ群に教示した。各カテゴリにつき最低 1 回の教示を行った。この教示に基づき半教師付き分類を行った後の GUI の様子を図 4 に, 分類結果を表 3 に示す。表 3 より鹿, 馬のカテゴリの画像を除き, 概ね各クラスに各カテゴリのデータが分類

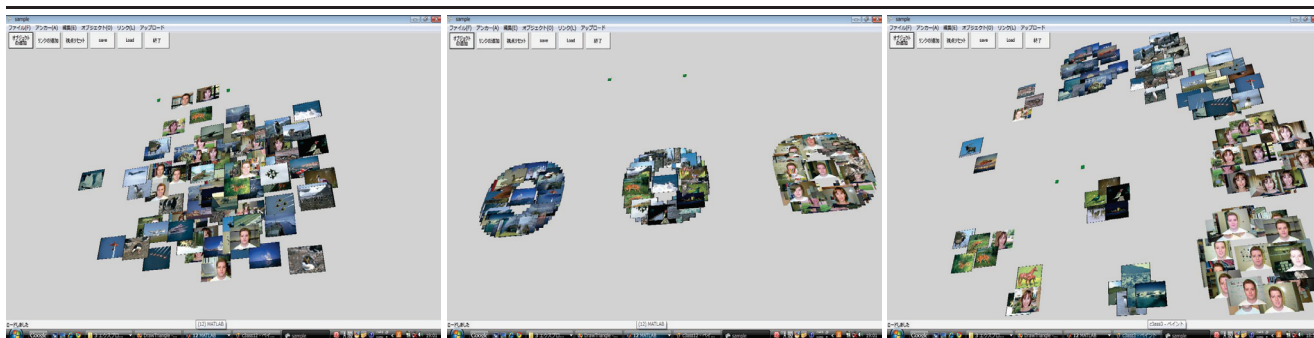


図 3: 提案システムの GUI

図 4: 一回目の教示に基づく分類の様子

図 5: 二回目の教示に基づく分類の様子

表 3: 2 回目の教示後の分類結果 (表は Confusion Matrix を示す.)

	鹿	飛行機	ペンギン	船	鳥	虎	犬	馬	氷山	男性	女性
Total	2	17	17	18	6	4	2	2	8	22	22
C1	1	1	1	0	0	0	0	0	0	0	0
C2	0	15	1	1	1	0	0	0	0	3	0
C3	0	0	13	0	0	0	0	0	0	0	0
C4	0	0	1	13	1	1	0	0	1	0	0
C5	0	1	1	1	3	0	0	0	0	0	0
C6	0	0	0	0	0	3	0	0	0	0	0
C7	0	0	0	0	0	0	1	0	0	0	0
C8	1	0	0	2	0	0	1	2	0	0	0
C9	0	0	0	1	0	0	0	0	6	0	0
C10	0	0	0	0	1	0	0	0	0	19	3
C11	0	0	0	0	0	0	0	0	1	0	18

されていることがわかる．上記のように二段階の教示を経て合計 20 個のラベルデータを与えた後に，120 個のデータは 78.9% の精度で分類された．

6. 考察

今回一人のユーザによる実験を行ったが，複数人によって分類した結果の比較や，インタラクティブ性の心理評価などを行うことでこのシステムを評価する必要がある．またインターフェイスについても知球 [久保田 06] や [Lopez-Gulliver 04] 等のようなより広範囲を俯瞰可能なものに拡張する必要がある．

7. 結論

本研究では少数のラベル付きデータと大量のラベルなしデータを利用して，距離学習の一種である AML に基づいて類似度グラフを構築し，このグラフをさらに学習することによって多クラスのパターンを高精度で分類可能な手法:GSSCML を提案した．またこの手法を搭載したインタラクティブ画像分類システムを開発した．半教師付き学習機能と GUI を通じた分類結果提示機能を持つことで，ユーザが少数のラベルデータを段階的に付与することにより分類を行うことが可能であることを示した．

参考文献

[C.Blake 98] C.Blake, and Merz, C.: UCI repository of machine learning databases (1998)

[Chapelle 06] Chapelle, O., Scholkopf, B., and Zien, A.: *Semi-supervised learning*, MIT Press (2006)

[Liu 09] Liu, W. and Chang, S.-F.: Robust multi-class transductive learning with graphs, in *Proc. IEEE Conference on CVPR*, pp. 381–388 (2009)

[Lopez-Gulliver 04] Lopez-Gulliver, R., Tochigi, H., Sato, T., Suzuki, M., and Hagita, N.: SenseWeb: collaborative image classification in a multi-user interaction environment, in *Proc of the ACM Multimedia*, pp. 456–459ACM (2004)

[Wagstaff 01] Wagstaff, K., Cardie, C., Rogers, S., and Schrodl, S.: Constrained k-means clustering with background knowledge, in *Proc. ICML*, pp. 577–584 (2001)

[Yang 06] Yang, L. and Jin, R.: Distance metric learning: A comprehensive survey, *Michigan State University* (2006)

[Ye 07] Ye, J., Zhao, Z., and Liu, H.: Adaptive distance metric learning for clustering, in *Proc. IEEE Conference on CVPR* (2007)

[Zhu 03] Zhu, X., Ghahramani, Z., and Lafferty, J.: Semi-supervised learning using Gaussian fields and harmonic functions, in *Proc. ICML*, Vol. 20, p. 912 (2003)

[久保田 06] 久保田 秀和, 西田 豊明, 角 康之: 等高線表現を用いたコンテンツアーカイブの可視化, 人工知能学会全国大会 (第 20 回) 論文集, No. 2F1-2 (2006)

[神田 10] 神田 賢一, 大本 義正, 西田 豊明: 知識空間の可視化と共有支援システムの開発とその利用, 情報処理学会全国大会論文集, No. 3ZG-8 (2010)