Multiple Kernel Learning を用いた階層的画像分類

小林 大輔† 山名 早人‡§

†早稲田大学大学院基幹理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1

‡早稲田大学理工学術院 〒169-8555 東京都新宿区大久保 3-4-1

§国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

E-mail: {d_kobayashi, yamana}@yama.info.waseda.ac.jp

あらまし 画像分類は，画像一枚に対して一つのカテゴリを付与することにより，コンピュータに画像の内容を理解させる有効な手段の一つである．近年では，分類対象のカテゴリに含まれている画像から特徴量を抽出し， 特徴量の傾向を元に，カテゴリの集合を分割する操作を繰り返すことによって，階層的に分類を行う研究が進められている．しかし，既存の階層的な画像分類の手法には，複数の特徴量の重みを自動的に設定する枠組みは備わっていない．複数の特徴量を用いて分類を行う場合，分類に有効な特徴量は階層ごとに異なるため，特徴量の重みが適切でなければ分類の精度が低下してしまう可能性がある．本論文では，Multiple Kernel Learning(MKL)と呼ばれる学習器を 既存の画像分類の枠組みに統合することで，分類対象の画像に合わせて，複数の特徴量を重み付けて統合しながら階層的に分類を行う手法を提案する．提案手法を用いることで，既存の階層的な画像分類と比較して分類精度が向上することを示す.

キーワード Image Recognition, Multiple Kernel Learning, Hierarchical Categorization

1. はじめに

画像分類は，画像一枚に対してひとつのカテゴリを付与することにより，コンピュータに画像の内容を理解させる有効な手段のひとつである．Web やカメラなどの発展により，我々の身の回りの画像情報は増加しており，画像分類研究の重要性が高まっている.

画像分類には様々な手法があるが，本論文に特に関係の深い研究として，カテゴリの階層構造を構築し分類を行う手法と，複数の特徴量を統合して分類を行う手法がある.

カテゴリの階層構造を構築し分類を行う手法では,複数のカテゴリに画像を分類する際に,カテゴリ同士に階層関係を設定することによって，上位の階層から下位の階層へ，再帰的に分類を実行する.

Deng，J らの研究[1]では，与えられたカテゴリについて，人手で木構造を構成し，k 最近傍法を再帰的に適応することよって，与えられた画像データセットを分類する手法を提案している．Li，L.J らの研究[2]では Hierarchical CRP と呼ばれる確率モデルを用い，カテゴリの木構造を構成する手法を提案している．また，Gao.T らの研究[3]では，サポートベクターマシン(SVM)を再帰的に適応し,全カテゴリを 2 つのクラスに分割する操作を繰り返すことで，画像を階層的に分類する手法を提案している.

複数の特徴量を統合して分類を行う手法としては，複数の特徴量を重み付けて統合し，重みを調整することによって分類の精度を向上させる手法がある．Varma らの手法[4]では，画像から複数の特徴量を抽出し，Multiple Kernel Learning(MKL)と呼ばれる学習器を用いて,画像に有効な特徴量の重み付けを学習することで，カテゴリごとに最適な重みを学習させている．MKL によって，Caltech101[5]や Caltech256[6]などのデータセットにおいて単一の特徴量を用いた場合よりも精度が向上したと報告している．Gonen らの手法[7]や Yang らの手法[8]では,MKL を改良し，特徴量の重みをカテゴリごとに一律に決定するのではなく，予測対象の特徴ベクトルに応じて変化させる手法を提案している.

MKL やその改良手法と比較して，Gao.T らの手法は効率的に分類を実行することができる．しかし Gao.T らの手法には，複数の特徴量の重みを自動的に設定する枠組みは備わっていない．分類に有効な特徴量は階層ごとに異なるため，特徴量の重みが適切でなければ分類の精度が低下してしまう可能性がある.そこで本論文では，Gao.T らの手法を改良し，複数の特徴量を考慮しながら，階層的に画像分類を行う手法を提案する．具体的には，Gao.T らの手法において，階層ごとに解く必要のある最適化問題の一部を MKL の最適化問題に変形することによって，階層ごとに特徴量の最適な重みを学習する手法を提供する.

本論文は以下の構成をとる．まず 2 節で提案手法の関連研究について述べ，その問題点について述べる．続いて 3 節で既存手法の問題点を解決する提案手法について述べる．

2. 関連研究

本節では，画像分類に関する手法のうち，提案手法に特に関係する手法として，画像の階層構造を構築し分類を行う手法と，複数の特徴量を統合して分類を実行する手法の二つの手法について述べる．

2.1. 画像の階層構造構築に関する手法

画像分類において，複数のカテゴリに画像を分類する際に，カテゴリ同士に階層関係を設定することによって，上位の階層から下位の階層へ，再帰的に分類を実行する手法が提案されている．

Deng ら[1]は，手動で画像の階層構造を構築する手法を提案している．具体的には，WordNet と呼ばれる，人手で定義した単語の階層構造に注目し，WordNet に含まれている単語をクエリとして画像検索を行うことによって画像を収集する．単語と合致しないノイズ画像については Amazon Mechanical Turk を用いて除去するとしている．

Li らの手法[2]では，階層的なトピックを定義する確率モデルである Hierarchical Chinese Restaurant Process (HCRP)を用いて画像に付けられた単語と画像から抽出した特徴量から階層構造を定義している．実験は自作した 4000 枚の画像データセットを 40 カテゴリに分類するタスクを行い，単語のみによって階層構造を構築した場合や，画像特徴量のみによって階層構造を構築した場合よりも精度が向上したとしている．

Gao.T らの手法[3]では，SVM を再帰的に適用することによって画像概念の階層構造を構築し，画像分類を行なっている．具体的には，まず与えられた学習用画像から特徴量を抽出する．次に，抽出した特徴量に基づいて，すべてのカテゴリを Positive, Negative, Ignore の 3 つのいずれかのクラスに分け，Positive クラスを正例，Negative クラスを負例として SVM により分類器を作る．具体的には以下に示す最適化問題**(1)**を解いて分類平面を決定すると共に，カテゴリのクラスを決定する：

$$
\begin{aligned}
\min_{\mathbf{w},b,\{\mu_k\},\{\xi_i\}} \quad & \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\left|\mu_{y_i}\right|\xi_i - A\sum_{i=1}^{N}\left|\mu_{y_i}\right| \\
& \mu_k \in \{-1,0,+1\}, \forall k \in Y \\
& \mu_{y_i}(\mathbf{w}^{\boldsymbol{T}}\boldsymbol{x}_i + b) \geq 1 - \xi_i,\ \forall i \\
& \xi_i \geq 0, \forall i \\
& -B \leq \sum_{k=1}^{K}\mu_k \leq B \\
& \sum_{k=1}^{K}\mathbf{1}\{\mu_k > 0\} \geq 1 \ \ and \\
& \sum_{k=1}^{K}\mathbf{1}\{\mu_k < 0\} \geq 1
\end{aligned} \tag{1}
$$

それぞれの変数の意味を表 1 に示す. 分類器を生成した後は, Positive クラスと Ignore クラスに属する全ての学習用画像, および Negative クラスと Ignore クラスに属するすべての学習用画像に対し，最適化問題**(1)**を再び適用することによって，画像を再帰的に分類する階層構造を構築する．

表 1　Gao.T らの手法の最適化問題におけるパラメータ

μ_k	k 番目のカテゴリのクラス (+1:Positive, 0:Ignore, -1:Negative)
$\boldsymbol{x}_i$	画像 i から抽出した特徴ベクトル
$\mathbf{w}$,b	分類平面の傾きと定数項
N	全画像枚数
Y	カテゴリの集合
y_i	i 番目のカテゴリ
K	カテゴリの総数
A,C,B	定数パラメータ
1{...}	中括弧内の条件を満たす場合に 1, 満たさない場合に 0 を返す関数

2.2. 複数の特徴量の統合に関する手法

画像分類において，階層構造の構築とは別の研究として，画像から色や形といった複数の特徴量を抽出し，それぞれの特徴量の分類への有効性に応じて，特徴量に重みづけて統合する研究が進められている．

Varma ら[4]は複数の特徴量の統合に関する手法として，Multiple Kernel Learning (MKL)を用いている．MKL はサポートベクターマシンにおけるカーネルを改良した分類器の一つである．MKL では複数の異なるカーネルを重み付けて統合する．具体的には分類平面の関数を以下のように定義する：

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i \sum_{m=1}^{M} \beta_m K_m(\boldsymbol{x}_i, \boldsymbol{x}) + b \tag{2}$$

ここでβ_mは m 番目のカーネルK_mに対する重みである．この重みは学習段階で自動的に調整される．

MKL は，複数の異なる特徴量によって分類を行う枠組を提供している．例えば一つ目のカーネルを画像から抽出した色の特徴量の類似度を測る関数，二つ目のカーネルを形の特徴量の類似度を測る関数として定義すると, MKL は，色と形のうち，どちらの特徴量が有効であるかを自動的に判定して，カーネルに重み付けを行い，分類を行う．

Gonen ら[7]は，MKL を改良した Localized Multiple Kernel Learning (LMKL)を提案している．LMKL ではカーネルに対する重みを定数ではなく，予測対象の画像の特徴ベクトルの関数として扱う．この関数のパラメータを学習することで重みを一律に与えるのではなく予測対象のベクトルに応じて変化させることができる．具体的には，分類平面の関数を以下のように定義する：

$$f(\boldsymbol{x}) = \sum_{i=1}^{N} \alpha_i y_i \sum_{m=1}^{M} \beta_m(\boldsymbol{x}_i)\beta_m(\boldsymbol{x}) K_m(\boldsymbol{x}_i, \boldsymbol{x}) + b \tag{3}$$

ここで$\beta_m(x)$は, 特徴ベクトル $\boldsymbol{x}$ に対するカーネルK_mの重みである．

Yang ら[8]は, MKL を改良した Group Sensitive Multiple Kernel Learning(GS-MKL)を提案している. GS-MKL ではあらかじめ k-means 法や P-LSA を用いて全ての特徴ベクトルをクラスタリングしておく. そしてそれぞれのクラスタに対して, MKL と同様の重みを付与する. 具体的には分類器のカーネルを以下のように定義する：

$$f(\boldsymbol{x}) = \sum_{i=1}^{N} \alpha_i y_i \sum_{m=1}^{M} \beta_m^{c(\boldsymbol{x}_i)} \beta_m^{c(\boldsymbol{x})} K_m(\boldsymbol{x}_i, \boldsymbol{x}) + b \tag{4}$$

ここで$c(x)$は特徴ベクトルxが属するクラスタ，β_m^jは，クラスタ j におけるカーネルK_mの重みである．

2.3. 既存手法に関する問題点

2.2 節で挙げた複数の特徴量を統合して分類を行う手法は，二値分類器である SVM を改良した手法であるため，複数のカテゴリを分類する場合，One vs. One 分類や One vs. Rest 分類などの複数の二値分類器を組み合わせる手法を用いる必要がある．カテゴリの個数をNとすると，必要な二値分類の回数は One vs. One 分類では$O(N^2)$，One vs. Rest 分類では$O(N)$になる．これに対し，2.1 節で挙げた Gao.T らの手法による画像分類では，二値分類を再帰的に繰り返しながら分類を実行しているため，必要な二値分類の回数は$O(\log(N))$となり，効率的に分類を行うことができる. しかし，Gao.T らの手法では，分類に複数の特徴量を用いる場合，それぞれの特徴量の重みを自動的に設定する枠組みは備わっていない．分類に有効な特徴量は階層ごとに異なるため，特徴量の重みが適切でなければ分類の精度が低下してしまう可能性がある．

3. 提案手法

本節では提案手法について説明する．3.1 項で, 提案手法の概要について述べる. 続いて 3.2 項で提案手法の最適化問題を提示し，3.3 項で最適化問題の解法について論じる.

3.1. 手法の概要

提案手法では既存手法の問題点に関する解決策として， Gao. T らの手法に MKL の枠組みを導入し，複数の特徴量を用いて，階層構造を構築し画像分類を行う枠組みを考える．具体的には，まず画像から色や形といった複数の特徴量を抽出する．得られた特徴量を最適化問題に適応することによって, 各カテゴリのクラス，各特徴量の重み，分類平面のパラメータの最適化を行う. 以下, 分類されたカテゴリに対して，同じ最適化問題を適用する操作を再帰的に繰り返す．カテゴリのクラスと階層構造の関係を図 1 に示す.

3.2. 最適化問題

提案手法では，Gao.T らの手法における最適化問題を MKL の枠組みを導入できるように改良する．具体的には，Gao.T らの手法における最適化問題に複数の特徴量とその重みを導入し，Gao.T らの手法で最適化するパラメータとともに，特徴量の重みを最適化する．さらに，カーネルを導入するため，特徴ベクトルを高次元空間へ写像する関数を導入する．最適化問題を**(5)**に示す：

$$\min_{\mathbf{w}, b\{\mu_k\},\{\xi_i\},\{\beta_l\}} \frac{1}{2}\left\|\sum_{l=1}^{L} \beta_l \mathbf{w}_l\right\|^2 + C\sum_{i=1}^{N}\left|\mu_{y_i}\right|\xi_i - A\sum_{i=1}^{N}\left|\mu_{y_i}\right| \tag{5}$$

subject to

$$\begin{aligned}
&\mu_k \in \{-1,0,+1\}, \forall k \in Y \\
&\mu_{y_i}\left(\textstyle\sum_{l=1}^{L} \beta_l \mathbf{w}_l \boldsymbol{\phi}_l(\boldsymbol{x}_i) + b\right) \geq 1 - \xi_i,\ \forall i \\
&\xi_i \geq 0, \forall i \\
&-B \leq \sum_{k=1}^{K} \mu_k \leq B \\
&\sum_{k=1}^{K} \mathbf{1}\{\mu_k > 0\} \geq 1 \ \ and \\
&\sum_{k=1}^{K} \mathbf{1}\{\mu_k < 0\} \geq 1
\end{aligned}$$

ここで$\boldsymbol{\beta_l}$はl番目の特徴量に対する重みである．また，$\boldsymbol{\phi}_l(\boldsymbol{x}_i)$は特徴ベクトル$\boldsymbol{x}_i$を高次元空間へ写像させる関数であり，ベクトル$\boldsymbol{x}_i, \boldsymbol{x}_j$に対する$l$番目のカーネルを$\mathbf{k}_{\boldsymbol{l}}(\boldsymbol{x}_i, \boldsymbol{x}_j)$で表すと，$\mathbf{k}_{\boldsymbol{l}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{\phi}_l(\boldsymbol{x}_i)^T \boldsymbol{\phi}_l(\boldsymbol{x}_j)$を満たす関数として定義される．

3.3. 最適化問題の解法

最適化問題**(5)**は閉形式で解くことはできないため，変数を２つのグループに分けて，それぞれについて最適化問題を考える．具体的には，分類平面に関する変数$\mathbf{w}, b, \{\xi_i\}, \{\beta_l\}$とカテゴリのクラスを決定する変数$\{\mu_k\}$に分割する．まず，分類平面に関する変数についての最適化は，**(5)**において最小化する式の第三項を定数とみなすことができ，以下の最適化問題に帰着される：

$$\min_{\mathbf{w}, b, \{\xi_i\}, \{\beta_l\}} \ \frac{1}{2}\left\|\sum_{l=1}^{L} \beta_l \mathbf{w}_l\right\|^2 + C\sum_{i=1}^{N} |\mu_{y_i}| \xi_i \qquad \textbf{(6)}$$

(6)は MKL における分類平面の最適化問題に一致するため，MKL の最適化手法を用いることで解くことができる．続いて$\{\mu_k\}$について最適化を行う．Gao.T らは$\{\mu_k\}$以外の変数$\mathbf{w}, b, \{\xi_i\}$を定数とみなすことで，最適化問題**(1)**が**(7)**に示す最適化問題に帰着されることを示している：

$$\min_{\{\mu_k\}} \ C\sum_{i=1}^{N}\left[n_k \mathbf{1}\{\mu_k = +1\} \times \left(\frac{1}{n_k}\sum_{i:y_i=k} \xi_{\mathrm{i}}^{+} - \frac{A}{\mathrm{C}}\right) + \mathbf{1}\{\mu_k = +1\}\left(\frac{1}{n_k}\sum_{i:y_i=k} \xi_{\mathrm{i}}^{+} - \frac{A}{\mathrm{C}}\right)\right]$$

subject to

$$\begin{aligned}
&\mu_k \in \{-1,0,+1\}, \forall k \in Y \\
&-B \leq \sum_{k=1}^{K} \mu_k \leq B \\
&\sum_{k=1}^{K} \mathbf{1}\{\mu_k > 0\} \geq 1 \ \ and \\
&\sum_{k=1}^{K} \mathbf{1}\{\mu_k < 0\} \geq 1
\end{aligned} \qquad \textbf{(7)}$$

ここで $\xi_{\mathrm{i}}^{+} = \max\{0, 1 - (\mathbf{w}^{\boldsymbol{T}} \boldsymbol{\phi}_{\boldsymbol{l}}(\boldsymbol{x}_i) + b)\}$， $\xi_{\mathrm{i}}^{-} = \max\{0, 1 + (\mathbf{w}^{\boldsymbol{T}} \boldsymbol{\phi}_{\boldsymbol{l}}(\boldsymbol{x}_i) + b)\}$である．

この制約付き最適化問題の解を見つけるには，まずそれぞれの$\{\mu_k\}$について，制約条件$-B \leq \sum_{k=1}^{K} \mu_k \leq B$を無視して最小化を行う．具体的には，それぞれの k について，

$$\frac{1}{n_k}\sum_{i:y_i=k} \xi_{\mathrm{i}}^{+} - \frac{A}{\mathrm{C}}, \qquad 0, \qquad \frac{1}{n_k}\sum_{i:y_i=k} \xi_{\mathrm{i}}^{-} - \frac{A}{\mathrm{C}} \qquad \textbf{(8)}$$

の３つの式のうち，最小の値をとるものを探し出す．一つ目の式が最小となるときには$\mu_k = +1$，二つ目の式が最小となるときには$\mu_k = 0$，３つ目の式が最小となるときには$\mu_k = -1$とする．

次に制約条件$-B \leq \sum_{k=1}^{K} \mu_k \leq B$を考える．制約を満たしていれば，制約条件を無視して求めたμ_kの値がそのまま解となる．もし$B < \sum_{k=1}^{K} \mu_k$ならば， すべてのμ_kの中から値が 1 であるものを取り出し，それぞれの値を 0 としたとき，最適化問題**(6)**の値の変化が小さい順にソートする．その後，制約条件が満たされるまで値の変化が小さいものから順に$\mu_k = 0$とする操作を繰り返し，最終的なμ_kの値を求める．$\sum_{k=1}^{K} \mu_k < -\mathrm{B}$についても，すべての$\mu_k$の中から値が- 1 であるものを取り出して同様の操作を行う．

謝辞

本研究の一部は，科研費（基盤（B）21300038）によるものである．

参　考　文　献

[1] Deng, J. and Dong, W. and Socher, R. and Li, L.J. and Li, K. and Fei-Fei, L: ImageNet: A Large-Scale Hierarchical Image Database, Proc.of IEEE Computer Vision and Pattern Recognition (2009).

[2] Li, L.J. and Wang, C. and Lim, Y. and Blei, D.M. and Fei-Fei, L.: Building and Using a Semantivisual Image Hierarchy, Proc. of IEEE Computer Vision and Pattern Recognition (2010).

[3] Gao, T. and Koller, D: Discriminative Learning of Relaxed Hierarchy for Large-scale Visual Recognition, Proc. of IEEE International Conference on Computer Vision (2011).

[4] M.Varma and D.Ray.: Learning The Discriminative Power-Invariance Trade-Off, Proc. of IEEE International Conference on Computer Vision (2007).

[5] L. Fei-Fei, R. Fergus, and P. Perona.: Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories. Proc. of IEEE Computer Vision and Pattern Recognition (2004).

[6] G. Griffin and A. Holub and P. Perona: Caltech256 Object Category Dataset, Technical Report 7694, California Institute of Technology (2007).

[7] M.Gonen and E. Alpaydin.: Localized Multiple

Kernel Learning, Proc. of International Conference of Machine Learning (2008).

[8] Yang, J. and Li, Y. and Tian, Y. and Duan, L. and Gao, W.: Group Sensitive Multiple Kernel Learning, Proc. of IEEE International Conference on Computer Vision (2009).

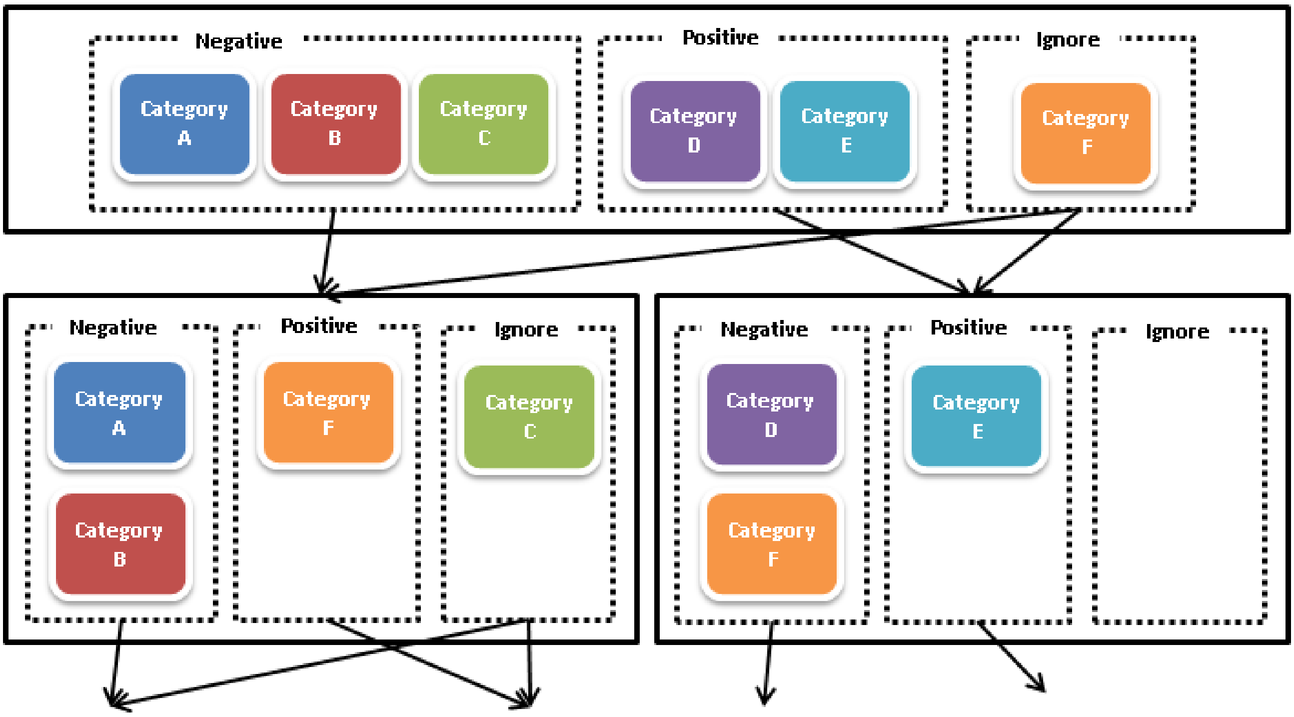

図 1 カテゴリの階層的分類