

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 24 日現在

機関番号：50103

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24500228

研究課題名(和文) 画像と局所特徴量の部分集合族によるマルチラベル画像分類システムの開発

研究課題名(英文) Development of Multi-Label Image Classification System by Subset Family of Images and Local Features

研究代表者

天元 宏 (TENMOTO, Hiroshi)

釧路工業高等専門学校・情報工学科・准教授

研究者番号：80321371

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：画像認識の分野において、特徴的な輝度分布(局所パターン)に基づくヒストグラムを生成し、それを特徴ベクトルとして画像分類が行われているが、1枚の画像中に複数の物体が含まれる場合、複数の物体に由来する特徴が混在するヒストグラムとなるため、従来の研究ではこれを適切に扱うことが難しかった。そこで、共通の性質を持つ画像と局所特徴量の組合せを確率的アルゴリズムで列挙し、それでデータ全体を被覆することで、マルチラベル付与による分類を試みた。実験により、ユーザーの感性と近い一致を示す分類並びに、重要キーワード画像の選択提示が可能であること等を確認できた。

研究成果の概要(英文)：In image recognition system, we often generate histograms based on local image patterns and classify the images by using the histograms as feature vectors. However, natural images usually contain multiple objects in each image, so the generated histogram are of the mixture of the histograms come from such multiple objects. Therefore, we proposed a new randomized algorithm that enumerates and covers the image set by the subset family of images and local features, which may give multi-labels to each image. Through computer simulations, we could confirm that classification results along with the users' sensibility and important visual words can be obtained by both supervised and unsupervised classification methods.

研究分野：パターン認識

キーワード：パターン認識 画像分類 機械学習 計算知能 進化計算 国際情報交換

1. 研究開始当初の背景

画像認識・理解の分野において、SIFT や HOG と呼ばれる局所特徴量を用いた物体認識に関する研究が盛んに進められており、画像から特徴的な輝度分布(局所パターン)に基づくヒストグラムを生成し、それを特徴ベクトルとして画像の分類(識別やクラスタリング)が行われている。しかし、1枚の画像中に複数の物体が含まれる場合、複数の物体に由来する局所特徴量が混在するヒストグラムとなるため、従来の研究ではこれを適切に扱うことが難しい状況であった。

一方で、報告者らは文書データに対し、1個の文書は複数の話題を含み得るとの観点から、クラスタ間の重複を積極的に認める、一般的なクラスタリングとは異なる独自の分類手法を提案しており、一定の成果を挙げている。局所特徴量を用いた画像分類においても画像がヒストグラムとして表現されることから、同様に報告者らの手法を応用することで、従来困難であった複数の物体を含む画像データの適切な分類の実現が可能なのではないかとの着想に至った。

2. 研究の目的

報告者らが提案している文書データに対する単語及びメタデータに基づく分類手法を、局所特徴量を用いた画像データの分類へと応用し、従来手法では難しい各画像の複数クラスタへの所属を積極的に認める画像分類システムの開発を目的とした。報告者らが提案している文書データに対する分類手法では、一般的な「データ全体をクラスタの集合に直和分割する」という考えではなく、「共通の性質を持つ文書と単語及びメタデータの組合せ」を確率的アルゴリズムにより検出・列挙し、その集合(文書と単語の部分集合族)でデータ全体を被覆する、という考えに基づく。本研究ではこの手法を画像データにおける局所特徴量に応用し、「共通の性質を持つ画像と局所特徴量の組合せ」でデータ全体を被覆することで、従来にない各画像へのマルチラベル付与による分類を試みた。

報告者らによる基本的なアルゴリズムは以下の通りである。1枚の画像は複数の物体を含み得るとの観点から、共通して登場すべき局所特徴量の個数等を予め条件として設定しておき、その条件を満たす範囲で限界まで拡大した部分集合を繰り返し検出し、最終的に全ての画像データを被覆する画像と局所特徴量の部分集合族を得る。これにより、従来の画像分類手法では難しかった複数の物体が同一の画像に含まれている場合の分類可能性に期待する。

また、以上のアルゴリズムで得られた各部分集合は、局所特徴量の Visual Words の列として可視化できるため、各部分集合を視覚的にラベル付けしてユーザーへ提示することも可能と思われる。ユーザーは列挙された部分集合をこのラベルで管理できる

と考えられる。

さらに、日付情報等のメタデータも特徴量として統合することで、大規模画像データセットに対する情報粒度の可変(階層的)ブラウザーとしての可能性も広がり、応用面でも期待できる。部分集合族を得るための確率アルゴリズムに与える条件を緩めれば大きなクラスタが少数得られ、逆に条件を厳しくすれば、小さなクラスタが多数得られる。これにより、ユーザーは最初に大きなクラスタ単位で画像データセットを概観し、その後、興味深いクラスタに対して更に詳細なクラスタを掘り下げて行くことで、階層的に画像データセットをブラウズすることが可能である。

以上より、本研究では、画像と局所特徴量の部分集合族による画像データセット全体の被覆に関して以下の点を確認し、明らかにすることを目的として研究を進めた。

- (1) 得られた部分集合族とユーザーの感性(ユーザーが求める結果)との一致性
- (2) 画像データセットの規模と確率的アルゴリズムの必要計算コストの関係
- (3) Visual Words による視覚ラベルの有効性
- (4) 画像とメタデータを統合した階層的ブラウザーの可能性

3. 研究の方法

上記目的(1)(2)(4)については教師無し分類手法にて、目的(3)については教師有り分類手法を用いて遂行した。以下、まず全体に共通する処理を説明した後、教師無し及び教師有りそれぞれの手法を説明する。

画像の局所特徴量から得られる Bag of Visual Words モデルに関して、最終的に分類に用いる特徴となる Visual Words の共起性をほぼ限界まで拡大した画像の部分集合族を求める確率的アルゴリズムを計算機上に画像分類システムとして実装し、実験、評価を行った。実験データとしては、画像認識研究用に広く用いられている画像セットである Caltech256 および VOC2006 を用いた。

処理の手順としては以下の通りである。まず、個々の画像より多数の局所特徴量を抽出し、ベクトル量子化により局所特徴量の辞書を構築した。次に、個々の画像に対し、抽出した各局所特徴量と辞書を比較し、その画像に含まれる局所特徴量のヒストグラムを算出した。この処理により、個々の画像がそれぞれ一つのヒストグラムに変換される。さらに、一定の条件下で限界まで拡大した部分集合の族を求めた。部分集合の拡大を試みる順としては、まだ被覆されていない画像を重点的に優先する重み付きランダム順列を用いた。個々の画像は複数のクラスタに所属するため、マルチラベルの分類結果となる。

教師無し確率的部分クラスタ法は、非排他的な確率的クラスタリング手法であり、確率的なゆらぎを許容する代わりに、実時間に収まる範囲の計算量でクラスタリングを行うこ

とが可能である。以下に画像分類に応用した手順を示す。

- (1) 空のクラスタを生成する。
- (2) 各画像について、重み付き順列により現在所属するクラスタの数に反比例する様に重み付けをした上で、ランダムに選択した画像データ(の Bag of Visual Words 表現)を加え、代表要素とする。
- (3) 同様の重み付けをした上でランダムにもう一つ選択する。Visual Words の頻度を比較し、一定以上の共通性が認められた場合、クラスタに加える。今回は比較に Bhattacharyya 距離を用いた。
- (4) 画像を追加できなくなるまでステップ(3)を繰り返す。
- (5) 全ての画像が一つ以上のクラスタに属するようになるまでステップ(1)～ステップ(4)を繰り返す。

同じ物体が写っている画像はヒストグラムの分布が似ていると考えられ、複数の物体が写っている場合は、それぞれの物体の特徴が分布に現れるはずである。更に、データ単位で複数回の比較を行うため、複数のカテゴリに所属した分類が可能になる。なお、重み付けの初期値は全ての画像で1とし、選択確率を等しくしている。重み付けにより、割り当てられたクラスタ数が少ない画像が優先的に選択されるため、比較的短時間で処理が完了する。

次に、教師有り学習を用いて遂行した手法について述べる。

報告者らは、近年ラフ集合の分野で研究が盛んな粒度(測定精度)の考え方を学習理論へ応用し、識別対象となるデータを識別に最も適した粒度で離散化して識別を行う手法を提案している。提案手法における粒度の調整は、任意の粒度における訓練サンプルの識別状況を情報量基準により評価することで実行する。提案手法により最適に調整された各特徴の粒度はまた、その特徴の識別への貢献度評価を与える事から、提案手法は特徴選択を一般化した手法と見なすことができる。

識別対象となるデータを離散化して識別を行うということは、ヒストグラムに基づく識別機を構成することを意味する。ヒストグラムに基づく識別機は、訓練サンプル数が無限大へと向かう際にベイズの意味で最適な識別機への収束が保証されている強力な識別機の一つである。

現実には小規模や中規模のデータに直面しなければならない場合が多い。その場合、良く取られるアプローチの一つは、特徴選択による次元の削減である。特徴選択段階を経ることで、現実に対処しなければならない難しい問題を、高い予測性能を備えた、言い換えると識別機のパラメータのより正確な推定が可能な識別機を構成できる様な、別の小さな問題へと変換することがきできる。

この目的のため、特徴選択手法として様々なアルゴリズムが提案されてきた。提案法で

は、これらの両方を達成することを考える。つまり、小規模や中規模なデータに対しては特徴選択を、大規模なデータに対してはベイズの意味での最適性を追求したいと考える。提案法では、これを各特徴軸上で異なる大きさの分割とすることで達成する。もし一切分割が必要なければ、その特徴軸は識別に貢献しない。その一方で、目の細かい分割が必要な特徴軸に対しては、その特徴軸は識別にとって重要である。提案法では、その様な最適な分割を MDL 基準に従って探求する。

本手法では評価関数を最小化する解を探索するために、GA(遺伝的アルゴリズム)を用いている。

4. 研究成果

まず、ユーザーの感性と近い一致を示す分類が可能であることを確認するため、教師無し分類による有効性を実験的に検証した。その結果の一例を以下に示す。

ここで入力に用いた画像データは、文房具等(ボールペン、瓶の蓋、はさみ、USBメモリ、ステープラー)の物体を撮影したものである。画像中のどこかに常に物体が存在する様な制約を設定している。また、相対的に写る物体の大きさも変更している。これにより「異なる状況の画像」を再現した。データセットの詳細は、[セット1]がボールペンと瓶の蓋、[セット2]がはさみ、USBメモリ、ステープラーである。

物体が一つだけ写っている画像は200枚、複数写っている画像は100枚ずつである。従って、[セット1]は、ボールペンと瓶の蓋が一つだけ写ったものを200枚ずつと、両方写っているものが100枚の合計500枚、[セット2]は全ての組合せはなく、はさみ、USBメモリ、ステープラーが一つだけ写ったものをそれぞれ200枚、はさみとUSBメモリ、はさみとステープラー、USBメモリとステープラー、三種類全ての組合せについて100枚ずつの合計900枚である。なお、Bhattacharyya 距離のしきい値は、実験的に検討し、データ1では0.41、データ2では0.48と設定した。

入力データセットに対し、提案手法による分類を行った結果、クラスタ中の物体の比率を見ると、いずれにおいても、特定の物体と、それが写っている複数物体を写した画像が多く分類されていることから、提案手法の有効性が確認できた。また、この一部の結果であっても、画像の合計数がそれぞれ974枚、1794枚となっており、入力画像より明らかに多い枚数となっていることがわかった。このことから、要素が重複する様なクラスタリングが行えていることがわかる。

教師無し部分クラスタ法を用いた提案手法により、一つの画像に複数の物体が写っていた場合でも、マルチカテゴリの分類が可能であることを一定程度、確認できた。本研究により、従来研究の多くが制約としている「一つの画像に一つの物体が写っている画像」と

いう条件を解除し、また、一般物体認識では、場合により膨大になる可能性の高いカテゴリの数に左右されない教師無し分類を行った。従って、従来研究の弱点を補完する様な分類を実現でき、本研究は画像分類の発展に寄与するものであるといえる。

次に、重要キーワード画像 (Visual Words) の提示を目的として、特徴選択手法との統合による教師有り分類の有効性を実験的に検証した。その結果の一例を示す。

ここでクラスはそれぞれ八サミ、USB メモリ、ステープラーとした。各クラスに 200 枚の画像を用い、その全体の半数を特徴選択および訓練に、残りの半数を誤差の推定に用いた。局所特徴量の算出には SIFT を用い、生の特徴を Bag of Visual Words モデルを用いて特徴ベクトルに変換した。また、準最適解である分割を見つけるために基本的な GA を用いた。Visual Word の個数を 10 から 1000 まで変化させ、各場合での誤差および選択された特徴の個数、つまり、分割数が 2 以上となる特徴の個数を調査した。

結果として、本手法では特徴選択により誤差が概ね 10% 程度上昇するが、特徴数はそれにより概ね元の特徴数の 40% に減らすことができることを確認した。これは本手法が Visual Word を重要なものとそうでないもの (ゴミ特徴) に分類できていることを意味する。

本手法の特徴選択処理により誤差がある程度上昇することは避けられないものの、それにより特徴数を大きく削減できることを確認できた。これは本手法が Visual Word を重要なものとそうでないものに分類できていることを意味する。これにより、画像セットから抽出した Visual Words が正しく識別能力を保有し、かつ、一部の Visual Words のみが分類に貢献していることが確認できた。これらをユーザーに提示すれば、当初の目的を実現できるものと考えられる。

以上、本研究では画像の局所特徴量としては既存の手法を利用するため、画像のピクセル単位での解析という観点では特色があるとは言えない。しかし、画像分類システムのバックエンドとして従来にない思想で処理を施す点において独創的であり、画像認識分野とデータマイニング分野の複合・融合として意義があると言える。

当初の目的を鑑みると、日付情報等のメタデータも特徴量として統合することで、大規模画像データセットに対する情報粒度の可変 (階層的) ブラウザーとして実現することは十分には実現できなかったが、基礎的なアルゴリズムの有効性を確認できた上で、さらに以下に挙げる今後の発展も検討することができた。

(1) 本課題の基礎理論の展開先として強化学習への応用も模索し、本課題で画像分類に用いた部分集合族を用いたアルゴリズムを試験的に環境識別機として組み込んだ強化学習の手法も試み、一定の成果を得ること

ができた。

(2) 進化計算を専門としているフィンランド・アールト大学の Seppo Ovaska 教授と本課題に関連する基礎理論に関してアイデアを交換し、本課題のアルゴリズム中で用いている最適化箇所に関して、分類精度の評価関数を分割・切替することにより改善が見込めることが分かり、現在、予備的な実験を進めている。

5. 主な発表論文等

[雑誌論文] (計 3 件)

Kizue Noro, Hiroshi Tenmoto and Akimoto Kamiya, Signal Learning with Messages by Reinforcement Learning in Multi-agent Pursuit Problem. *Procedia Computer Science*, 35(2014), 233-240. 査読有り.

DOI: 10.1016/j.procs.2014.08.103

天元 宏, MDL 基準を用いた特徴選択と Bag of Visual Words モデルへの応用の試み. 釧路工業高等専門学校紀要, 第 48 号, 2015 年(平成 27 年)1 月, 31-38. 査読無し.

<http://www.kushiro-ct.ac.jp/library/kiyo/>

天元 宏, 佐藤 一真, 局所特徴量と教師無し部分クラス法を用いたマルチラベル画像分類の試み. 釧路工業高等専門学校紀要, 第 46 号, 2012 年(平成 24 年)12 月, 43-50. 査読無し.

<http://www.kushiro-ct.ac.jp/library/kiyo/>

[学会発表] (計 2 件)

Kozue Noro, Hiroshi Tenmoto and Akimoto Kamiya, Signal Learning with Messages by Reinforcement Learning in Multi-agent Pursuit Problem. *Proceedings of the 18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*, Gdynia, Poland, 2014, September.

[http://kes2014.kesinternational.org/Hiroshi Tenmoto and Mineichi Kudo, Feature Selection on Bag of Visual Words Model by Using MDL Criterion. The ICPR Workshop on Features and Structures \(FEAST 2014\), Stockholm, Sweden, 2014, August.](http://kes2014.kesinternational.org/Hiroshi_Tenmoto_and_Mineichi_Kudo_Feature_Selection_on_Bag_of_Visual_Words_Model_by_Using_MDL_Criterion.The_ICPR_Workshop_on_Features_and_Structures_(FEAST_2014),_Stockholm,_Sweden,_2014,_August.https://sites.google.com/site/feast2014/)

<https://sites.google.com/site/feast2014/>

6. 研究組織

(1) 研究代表者

天元 宏 (TENMOTO, Hiroshi)

釧路工業高等専門学校・情報工学科・准教授

研究者番号: 80321371