

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

天元 宏*¹

Development of Multi-Label Image Classification System by Subset Family of Images and Local Features

Hiroshi TENMOTO

本稿では、平成 24 年度 (2012 年度) から平成 26 年度 (2014 年度) までの期間に科学研究費の補助を受けて行った一連の研究 (基盤研究 (C), 課題番号 24500228) について、その概要を抜粋して報告する。

1. 研究開始当初の背景

画像認識・理解の分野において、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。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であり、ユーザーはこのラベルで画像を管理できる。部分集合族を得るための確率的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となる。

以上より、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

*¹ 釧路高専情報工学科

3. 研究の方法

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。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。

教師無し分類の処理手順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序を用いた。個々の画像は複数のクラスに所属し得るため、マルチラベルの分類結果となる。

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。この特徴選択手法により重要 Visual Words の選出を行い、ユーザーへの視覚ラベルの提示を試みた。

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。例えば、文房具 (ボールペン、瓶の蓋、はさみ、USB メモリ、ステープラー等) を単独で撮影した画像および複数同時に撮影した画像を混合したデータを入力に用い、提案手法による分類を行った所、いずれのクラスにおいても特定の物体と、それを含む複数物体を撮影した画像が多く分類さ

れていることが確認できた。また、特定の画像が複数のクラスに現れており、要素が重複する様にクラスタリングが行えていることも確認できた。

次に、重要 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。例えば、前述のハサミ、USB メモリ、ステープラー等を撮影した画像を用い、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、Visual Words 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% に減らせることが確認できた。これは本手法が Visual Words を重要なもの (ゴミ特徴) 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重要 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といえる。

まとめとして、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、マルチラベル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うことができ、従来研究の弱点を補完する様な分類を実現できたといえる。

5. 主な発表論文等

1. 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.
2. 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.
3. 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.