

最頻共通特徴に着目した画像検索結果の洗練化手法

A method of sophisticating outputs of image retrieval using the most frequent common image feature

尾崎 佑樹†
Yuki Ozaki

長尾 智晴†
Tomoharu Nagao

1.はじめに

我々の研究グループでは汎用画像認識エンジンの開発を行っている。そのためには、知識として物体ごとの特徴データベースを作製する必要がある。この特徴データベースの作製には、大規模データセットを用いることも可能であるが、データセットには存在しないカテゴリに対しても認識が行えるように WWW(World Wide Web)上に存在する大量の画像データを活用することを考えている。WWW上の画像に対してキーワードを基に検索可能な検索エンジンとして様々な検索エンジンが存在するが、画像が含まれるページに存在するキーワードだけを用いて検索が行われているため、キーワードとは関係ない画像が出力されることがある。そこで、画像検索エンジンが出力した画像に対して画像特徴を利用した分類を行うことで、キーワードとは関係ない画像を除去し、画像検索結果を洗練する。その結果、欲しい画像を自動で大量に収集することが可能になることが期待される。

インターネットから画像検索エンジンを使って画像を収集した後、画像特徴を利用して検索結果を洗練化する研究はいくつか存在する。HTML文書解析と画像特徴を利用した研究[1]やキーワードによる画像検索の後、検索結果の画像をユーザがキーとして指定して類似画像検索を行う研究[2]がある。HTML文書に含まれるテキスト情報が必ずしも正しい関連があるとは限らないと考えられるため、本稿では画像特徴の解析のみを用いて画像を収集する。また、あらゆるカテゴリに対し、自動で大量の画像を収集することを目的としているためインタラクティブな検索は適さないと考えられる。本稿では、特徴ベクトルの特徴空間上での分布を用いて最も多くの画像で共通している物体の特徴である最頻共通特徴を抽出する。最頻共通特徴を含む画像を抽出することで、ラベルなしデータのみから画像検索結果を洗練する方法を提案する。

2.提案手法

画像検索結果として出力される画像セットはキーワードに合致した目的の物体が写っている画像の割合が高いと考えられる。そこで、画像セットの各画像から抽出される特徴のうち、最も多くの画像で共通している特徴である最頻共通特徴がキーワードに関する物体の特徴であると考えられる。計算機がその特徴を抽出することで、人間が知識を与えなくても自動でそのカテゴリの特徴を学習することが期待される。その特徴を利用して検索後の画像を目的の画像と目的とは異なる画像に分類することで画像検索結果を洗練化することができる。

2.1.特徴抽出

一枚の画像には様々な物体が存在することが多いため、目的の物体を表す特徴を抽出するには、局所的な特徴を利用する必要がある。そのため、画像をブロック分割し、ブロック単位で特徴量を記述する。特徴量は RGB 色空間の各軸を 4 等分とした 64 次元のカラーヒストグラムを特徴量とした。次に、全ての画像のブロックを k-means 法を用いてクラスタリングを行う。また、ブロック同士の距離の測定に Bhattacharyya 係数による類似度を用いた。

2.2 画像分類

最も多くの画像に共通している物体の特徴が最頻共通特徴であり、各クラスタのブロックが所属する画像の数が最も多いクラスタが最頻共通特徴となる。そのブロックを含む画像を目的の画像、含まないものを目的とは異なる画像として分類する。

3.最頻共通特徴を用いた画像分類実験

提案手法の有効性を検証するため、Google 画像検索でキーワードを入力して出力された画像を用いて分類実験を行った。キーワードを「トラ」として入力した時に出力された 100 枚のトラ画像と同時に出力された 100 枚の非トラ画像を使用した。画像のサイズは 128×128[pixel]、ブロックのサイズは 16×16[pixel]とした。なお、ブロックの特徴量が切り出す位置に依存しないようにブロックの切り出す位置をブロックの半分だけずらしたのも使用している。

3.1.k-means 法によるブロックのクラスタリング

実験的に k=15 としてブロックをクラスタリングした。各クラスタに様々なパターンが混在しないように全ての代表ベクトルとの類似度が 0.7 未満のブロックは使用しなかった。各クラスタのブロック数と画像数は表.1 のようになった。Cluster3 のとき最も画像数が多かったため、この Cluster3 のブロックが最頻共通特徴となる。

この Cluster3 のブロックは図.1 のようにトラのパターンを表していた。最頻共通特徴がトラの黄色と黒の模様を表していることが分かる。しかし、図.2 のように非トラ画像のパターンも混在してしまっていた。これは色の距離が近いためトラを表すパターンと同じパターンとして分類されてしまったと考えられる。ヒストグラムのビンを小さくすると、トラを表すパターンも細かく分類されてしまい、メンバ数の多い背景などのパターンが最頻共通特徴として抽出され、物体を表すパターンを抽出できなくなってしまった。最頻共通特徴を一つの特徴だけで表現するのではなく複数の特徴を組み合わせることで表現する

†横浜国立大学大学院環境情報学部
Graduate School of Environment and Information
Sciences, Yokohama National University

ど、最頻共通特徴の表現の仕方を工夫する必要があると考えられる。

Cluster No.	メンバ数	画像数
1	2307	140
2	363	35
3	2095	164
4	2530	152
5	187	28
6	918	69
7	811	97
8	184	21
9	1174	116
10	341	31
11	3296	163
12	195	11
13	3656	140
14	2085	163
15	947	97

表1 クラスタリング後の各クラスターのブロックのメンバ数と画像数

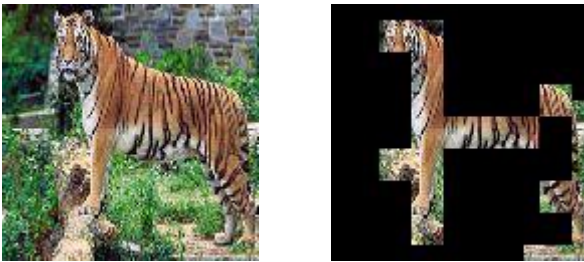


図1 トラ画像における原画像(左)と最頻共通特徴(右)のブロック例



図2 非トラ画像における原画像(左)と最頻共通特徴(右)のブロック例

3.2.分類結果

最頻共通特徴のブロックが含まれる画像をトラ画像、含まれない画像を非トラ画像とすると分類結果は表2のようになった。

トラ画像の再現率は高い数値を示しているが、非トラ画像をトラ画像と認識する誤分類率が高く、過検出傾向な結果となった。色の特徴がトラのパターンに比較的似ている地面のようなパターンを表すブロックも最頻共通特徴として含まれていたため、過検出になったと考えられる。図3のように色の特徴が大きく異なる画像に関しては非トラ画像と分類することができた。

	トラ	非トラ	再現率
トラ画像	96	4	96%
非トラ画像	66	34	34.0%
適合率	59.3%	89.5%	73.3%

表2 分類結果



図3 非トラと認識した画像

4.まとめ

キーワードをクエリとしたインターネットの画像検索エンジンを利用して得られた画像セットに対し、最頻共通特徴を抽出することでキーワードに基づく目的の画像と目的とは異なる画像に分類し、画像検索結果の洗練化を行った。

色の特徴を使って画像セットに存在する最も頻度の高い特徴を抽出し、分類を行った。しかし、最頻共通特徴を表すクラスターには関係ない物体のパターンを表す特徴も混在してしまい、過検出気味な結果となってしまった。カラーヒストグラムだけでは細かく、色を含んだ模様の特徴を抽出することは困難であるため、今後は、特徴量の改善を行っていききたい。また、最頻共通特徴の一つの特徴によって表現したが、複数の特徴を組み合わせるなど表現の仕方を工夫する必要がある。

また、本稿では同じカテゴリで色の特徴がほとんど同じであるトラを対象としたが、猫や人のような同じカテゴリでも色や模様が異なる物体に対しても分類が可能となるように、物体の形状を表現する特徴を利用して、対応できる物体の幅を広げていきたい。

参考文献

[1]柳井啓司, "キーワードと画像特徴を利用したWWWからの画像収集システム", 情報処理学会論文誌: データベース, Vol.42, No.SIG 10(TOD 11), 2001.

[2]Scaroff, S., LaCascia, M., Sethi, S. and Taycher, L.: Unifying Textual and Visual Cues for Content-Based Image Retrieval on the World Wide Web, Computer Vision and Image Understanding, Vol.75, No.1/2, pp86-98, 1999.