

複数の特徴空間における最頻共通特徴の共起に着目した 画像検索結果の洗練化

尾崎 佑樹[†] 長尾 智晴[†]

[†] 横浜国立大学 大学院環境情報学府

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。画像中の物体を認識する際、知識として物体ごとの特徴データベースを構築する必要がある。ウェブ上に存在する膨大な画像データを有効活用したいと考えている。ウェブ上の画像を検索する手段として様々な画像検索エンジンが存在するが、出力には目的とは異なる画像が混在する。そこで、画像検索から得られた画像セットをもとに画像内容に基づく洗練化を行うことで、より純度の高い画像セットを獲得する手法を提案する。これにより、ウェブ上の画像データを様々なカテゴリにデータベース化することで、画像認識における学習画像の確保に応用したいと考えている。従来の研究でも、ウェブから目的の画像を収集することを目的とした研究 [1][2][3] がある。本研究では、画像検索に Google 画像検索を用い、複数の特徴空間における最頻共通特徴の共起による画像検索結果の洗練化を行う。また、人手で教師データを与えることなく自動で洗練化を行うことを目的としており、正解が与えられていないラベルなしデータを扱うものとしている。

2 提案手法

提案手法である複数の特徴空間における最頻共通特徴の共起による画像検索結果の洗練化手法について説明する。提案手法では、カテゴリとは異なる画像が少量しか混在していないという Google 画像検索の検索上位の出力の特性を利用するという制約の下、検索上位の画像セットを学習画像として用い、カテゴリの物体の特徴を学習する。ただし、正解・不正解の信号は与えない。また、画像を目的のカテゴリの画像とそうでない画像に分類するだけでなく、目的の画像と分類された画像がどれだけそのカテゴリらしいかという尺度を表す「カテゴリらしさ」を最頻共通特徴のパターンの帰属度から算出し、カテゴリらしさ順にソートして結果を出力する。提案手法は、2つのステップから構成されている。まず、以下のステップで学習画像から最頻共通特徴を学習する。

1. 各学習画像から $w \times w$ [pixel] のブロックを、 s [pixel] 間隔で抽出し、各ブロックから特徴ベクトルを算出する。

2. 抽出したブロックを k-means 法を用いてクラスタリングする。
3. 得られたクラスタの画像数を算出し、多くの画像数から構成されているクラスタを最頻共通特徴クラスタ $C^{(M)}$ とする。

次に、以下のステップで未知画像の分類を行う。また、画像をカテゴリらしさ順にソートを行う。

1. 未知画像から $w \times w$ [pixel] のブロックを、 s [pixel] 間隔で抽出し、各ブロックから特徴ベクトルを算出する。
2. 抽出したブロックを学習で得られた代表ベクトルを用い、クラスタリングを行う。所属クラスタ決定後、以下の式で、クラスタ j に属する未知画像のブロックの特徴ベクトル $\mathbf{x}$ の帰属度 A_x を算出する。

$$A_x = \begin{cases} 1 - D(\mathbf{x}^{(j)}, \mathbf{c}^{(j)}) / D_{\max}(\mathbf{v}_i^{(j)}, \mathbf{c}^{(j)}) \\ (D(\mathbf{x}^{(j)}, \mathbf{c}^{(j)}) \leq D_{\max}(\mathbf{v}_i^{(j)}, \mathbf{c}^{(j)})) \\ 0 \quad (\text{otherwise}) \end{cases} \quad (1)$$

3. i 番目の未知画像 P_i のブロックを代表ベクトルで量子化し、以下の式で画像のカテゴリらしさ E_i を求める。

$$E_i = \begin{cases} \frac{1}{n} \sum_{x=1}^n A_x^{(i)} & (\mathbf{x} \in C^{(M)}) \\ 0 & (\text{otherwise}) \end{cases} \quad (2)$$

E_i が 0 でないとき正例（目的の画像）、0 のとき負例（目的とは異なる）の画像とする。

4. 以上の処理を全ての特徴量とブロックサイズの組み合わせで行う。
5. しきい値以上の分類器で正例と分類された画像を最終的な正例画像とする。各組み合わせで算出された E_i の総和を、最終的な未知画像 P_i のカテゴリらしさとする。

$D(\mathbf{a}, \mathbf{b})$ は、ベクトル $\mathbf{a}$ とベクトル $\mathbf{b}$ の距離を表す。距離は、ユークリッド距離を使用している。 $\mathbf{v}_i^{(j)}$ はクラスタ j 内の i 番目のベクトル、 $\mathbf{c}^{(j)}$ はクラスタ j の代表ベクトル、 $A_x^{(i)}$ は i 番目の画像の x 番目のブロックの帰属度、 $C^{(\cdot)}$ はクラスタを表し、 $C^{(M)}$ は最頻共通特徴のクラスタである。

A Method of Improving Outputs of Image Retrieval Using Co-occurrence of the Most Frequent Common Image Feature in Some Feature Space

Yuki Ozaki[†], Tomoharu Nagao[†]

[†] Graduate School of Environment and Information Sciences, Yokohama National University

3 画像洗練化実験と考察

キーワードをクエリとした Google 画像検索で得られた画像セットに対し、本提案手法の有効性を検証した。本研究では, airplane, car, desk, dog, horse, motorbike, piano, plastic bottle, rose, tiger の 10 カテゴリに対し, 実験を行った。現在, Google 画像検索では最大 1000 枚まで画像が出力され, 検索下位になるほど目的とは異なる画像の割合が高くなる。Google 画像検索の検索上位数における適合率を検証したところ, 上位 200 枚は, 多くのカテゴリで 8 割程度の目的の画像が得られることが分かったため上位 200 枚を学習画像として利用した。また, 特徴量には, 13 種の基本統計特徴量 (最大値, 最小値, 平均値, 標準偏差, 1σ 内確率, 3σ 外確率, 中央値, 第一四分位点, 第三四分位点, 最頻値, レンジ, 歪度, 尖度) と ULBP (Uniform Local Binary Pattern) 特徴量を用いた。画像サイズを 128[pixel], ブロックサイズ w は $w = 16, 32, 64, 128$ [pixel], 最頻共通特徴のクラスタを画像数が最も多いクラスタから各クラスタの画像数の $\sigma/3$ だけ少ない範囲のクラスタとした。正例と分類した分類器の数が半数以上のとき, 正例と分類した。

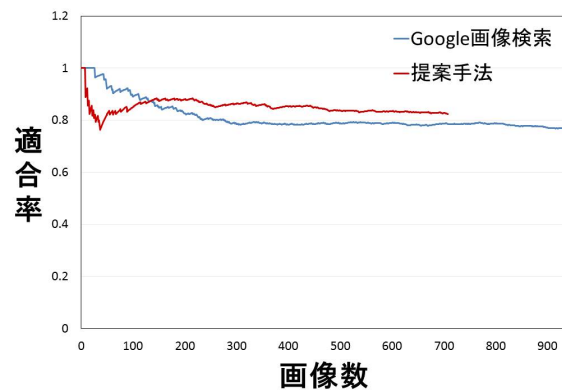
4 洗練化実験結果

Google 画像検索から得られた画像セット, 10 カテゴリに適用した洗練化結果を表 1 に示す。

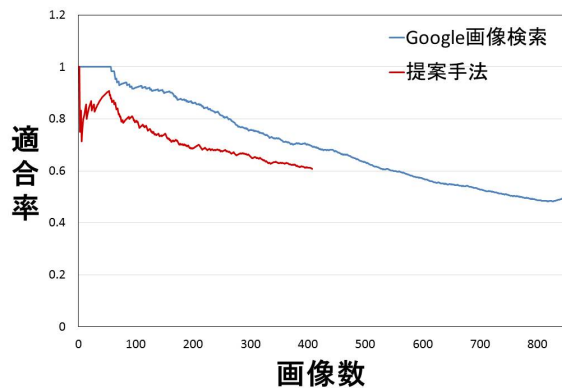
表 1: 10 カテゴリにおける提案手法と Google 画像検索の適合率

	提案手法 (%)	Google 画像検索 (%)
airplane	68.0	63.0
car	82.4	76.9
desk	81.0	79.7
dog	66.6	58.0
horse	73.7	68.5
motorbike	79.0	75.5
piano	88.1	74.4
plastic bottle	87.1	68.3
rose	60.8	49.5
tiger	58.7	42.4

そして, 抽出した画像をカテゴリらしき順にソートしたときの画像数と適合率の関係と Google 画像検索の検索上位数と適合率の関係の一例を図 4 に示す。提案手法が目的の画像と分類した画像数と同数における Google 画像検索の上位の適合率を比較すると, 10 カテゴリのうち, car, dog, motorbike, plastic bottle の 4 カテゴリで提案手法が上回ることができた。提案手法が上回った 4 カテゴリは, 物体の向きや姿勢による変化が比較的少ないカテゴリであったため, 今回設定した最頻共通特徴のクラスタ数で物体の特徴パターンを捉えることができたと考えられる。また, 色の特徴が顕著なカテゴリでは選択的にカラーヒストグラムのような特徴を利用することができれば, rose, tiger のようなカテゴリも精度を高 j 評させることができると考えられる。カテゴリらしき上位の画像でも, 目的とは異なる画像が混在しているため, 特徴量の追加やカテゴリの物体とその他の領域を分離する有効な特徴表現を検討する必要がある。



(a) カテゴリ "car" の場合



(b) カテゴリ "rose" の場合

図 1: 抽出した目的画像のカテゴリらしき上位数における適合率と Google 画像検索の検索上位数における適合率の関係例

5 まとめ

複数の特徴空間における最頻共通特徴の共起を利用した提案手法を, Google 画像検索から得られた画像セットに適用し, 有効性を検証した。全てのカテゴリにおいて, もとの画像セットの適合率よりも高い適合率を示すことができたため, 洗練化の有効性が確認できた。しかし, 得られた画像数と同数における Google 画像検索の検索上位の適合率との比較では, 精度が劣るカテゴリが多く見られたため, 実用するためには精度の向上が必要であることが分かった。今後は, 特徴量の追加や特徴の位置を利用するなど最頻共通特徴の表現方法の改善により精度を向上させていきたい。

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.SIG10 (TOD11), pp.79-91 (2001/10).
- [2] Sclaroff Stan: "Unifying textual and visual cues for content-based image retrieval on the World Wide Web", Computer Vision and Image Understanding, pp.86-98(1999).
- [3] Michael J.Swain and Frankel Chales and Vasilis Athitsos: "WebSeer: An image search engine for the world wide web", Technical Report TR-96-14, University Chicago(1996).