

6 Y - 3 機械学習を用いた感性語による画像検索

佐野健吾

雨宮秀文

宮本正邦

池田茂

中川 裕志

森 辰則

横浜国立大学 工学部

1 はじめに

近年の計算機メディアの進歩により画像データの取り扱いが増えており、その中の一つである画像検索にも様々な方法が採られている。

我々は、機械学習アルゴリズムを用いて感性語と画像の持つ物理的特徴を関連付け、その感性インデックスを基に画像検索を行うことを実現している。この方式の優れた点は一度学習により関連付け情報を取得すれば、未知の画像に対してもその関連付け情報を用いて感性インデックスを付与することができることである。また、複数の学習方式を組み合わせる多戦略学習により精度の向上を図っていることも特徴の一つである。

以下で、我々の開発した感性語による検索が出来る画像検索システムについて述べ、評価をする。

2 システム概要

システムの概要を図1に示す。本システムのコアとなる部分は二つある。一つは画像の物理的特徴を自動的に抽出する部分であり、もう一つはあらかじめ人間によって感性評価のつけられたサンプル画像と物理的特徴を用いて学習を行う学習アルゴリズムでの部分である。

2.1 感性量の取得

本システムではある画像を見た時に人間が感じる感性を表す語、すなわち感性語(優雅な、暖かい、等)を一つの感性語に対してその対立概念語(がさつ、冷たい、等)との対を用意して一つの感性軸を設けた。今回の実験では16対の感性語を採用し、その軸に対する値をアンケートにより取得した。例としては“落ち着き ↔ 賑やか”、“安心 ↔ 不安”、等が挙げられる。

Machine Learning based Image Retrieval with KANSEI Words
Kengo Sano, Hidefumi Amamiya, Masakuni Miyamoto, Shigeru Ikeda, Hiroshi Nakagawa, Tatsunori Mori
Division of Electrical and Computer Engineering,
Faculty of Engineering, Yokohama National University

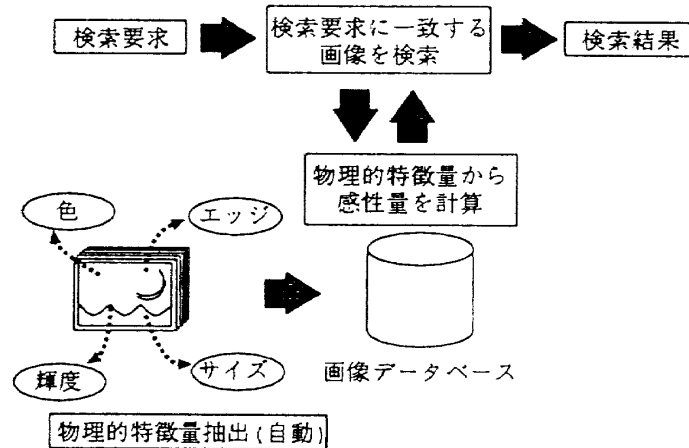


図 1: システム全体図

2.2 画像の物理的特徴

画像の物理的特徴の抽出は後述する学習の基となるものであり、どのような特徴量を抽出するかは非常に重要である。我々が使用している物理的特徴は、輝度、各色の割合(8色に減色)、画素間の微分、画面の明るさを一定方向に変化させた画像との相関、まとまった部分の色についての特徴量、R、G、Bの分布状態等である。使用した物理的特徴の数は最大108である。

2.3 学習アルゴリズム

感性語(感性量)と画像の物理的特徴量との関連付けを行うのに本システムでは以下の4つの機械学習を用いた。

- 重回帰分析 (mra) … 感性語を物理的特徴の一次式で表現
- 決定木 (d-tree) … 感性語に当てはまるか否かを物理的特徴の不等式群で判断する決定木で表現。C4.5を使用
- k近傍法 (k-nn) … 物理的特徴の張る多次元空間を感性語に対応する部分に分割する境界線を与える

- ニューラルネットワーク (neural-net) … 物理的特徴を入力すると感性語を出力するネットワークを構成

3 検索システムによる評価方法

感性量の付加された静止画像 100 枚について、80-20 と画像を無作為に分け 80 枚のデータで学習を行い、残りの 20 枚に対して予測値と真値を比較するという 5-fold cross validation を行って評価した。それらを 4 種類の学習法で学習を個別に行い、学習法の違いによる影響を検証した。

評価尺度 システムが算出した感性量と人間がアンケートでつけた値が一致していた場合 1 ポイント、一致していなくても評価尺度が 1 段階しかずれていなかった場合 0.5 ポイント得点をつける。そして、その得点を合計したものを画像の総数で除したものを正答率と定義する。

3.1 学習による検索システムとその評価

各学習法の違いによる感性語別の正答率を図 2 に示す。

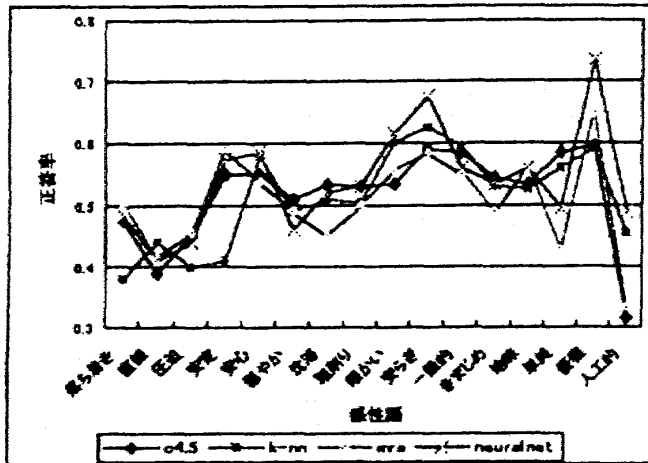


図 2: 学習法別の正答率

各学習法毎の 16 種類の感性語の正答率の平均値は、重回帰分析が 0.506、決定木が 0.517、k 近傍法が 0.512、ニューラルネットワークが 0.542 であった。平均だけ見ると正答率は似ているが感性語毎に正答率をみると学習方法によって得意、不得意な感性語があることが分かる。

よって各感性語毎に最も正答率の高い学習法を選んだ場合の正答率と感性語毎の平均の正答率を図 3 に示す。最も正答率の高い学習法を選んだ場合の正答率の平均は 0.559 である。図 3 より学習法を複合させる方法は有効であるといえる。

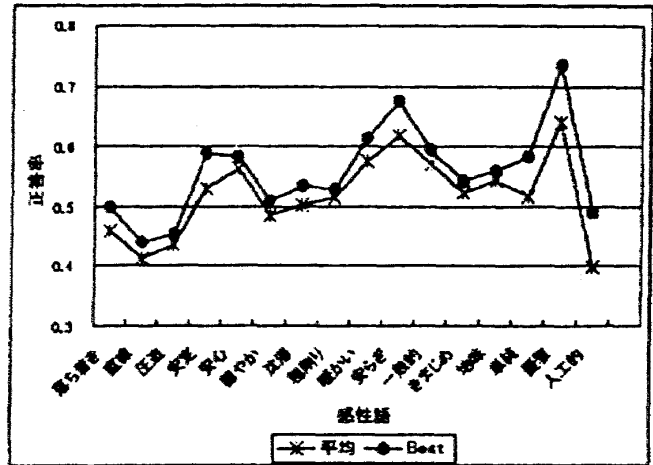


図 3: 複合方法による正答率

4 まとめ

感性語による画像検索システムを紹介した。また学習システムを複数用いることにより感性語画像検索の効率は良くなることを確認した。

今後の課題として、特定分野の画像に対して有効な物理的特徴を抽出し検索効率をさらにあげる予定である。

参考文献

- [1] 加藤俊一. サイバースペースにおけるマルチメディア情報技術. 重点領域研究「高度データベース」総括班. p67-74. 1997.
- [2] 鄭絳宇, 安部憲広. 画像の伝える感性情報の抽出と利用. 重点領域研究研究成果報告書 感性情報処理の情報学・心理学的研究. p165-170. 1995.
- [3] 八木伸行, 井上 誠喜, 林 正樹, 中須 英輔, 三谷 公二, 奥井 誠人, 鈴木 正一, 金次 保明. C 言語で学ぶ実践画像処理. オーム社. 1992.