

複合的類似度を用いた内容に基づく画像検索における OPF 分類器の導入

廣瀬 雄真[†] 服部 元信^{††}

[†]山梨大学大学院医学工学総合教育部 ^{††}山梨大学大学院総合研究部工学域

1 はじめに

従来、画像検索は人手により索引付けされた注釈を用いて検索が行われてきた。しかし、近年の情報技術の発展による画像データの爆発的増加により、人手による索引付けは困難になった。また、人の主観がラベル情報に影響し、検索結果に悪影響を及ぼしてしまう恐れもある。

このような問題をコンピュータに解決させることを目的とした画像検索システムの1つとして、内容に基づく画像検索 (CBIR: Content-Based Image Retrieval) が存在する。近年では、ユーザの検索要求をシステムが把握するために、適合性フィードバック (RF:Relevance Feedback) が盛んに採用されている。Wuらは、検索毎に行われる Short-term RF(SRF) と、過去のフィードバックのログ情報を蓄積し、現在の検索に役立てる Long-term RF(LRF) の2つのアプローチを融合した複合的類似度を用いて検索する HySim(Hybrid Similarity measure) 法を提案し、優れた検索性能が得られることを示している [1]。

本研究では複合的類似度を用いた検索システムを改良することを目的とする。蓄積されたログの質をより高めるために、フィードバックの学習にクラス分類において良い性能を持つ Optimum-Path Forest(OPF) 分類器を導入することにより、検索性能の向上を目指す。

2 HySim(Hybrid Similarity measure) 法

HySim 法は、SRF と LRF によって求められた類似度を統合した複合的類似度を用いることによって、よりユーザの要求に近い画像を少ない RF によって取り出すことを目的に開発された手法である。また、画像の概念が一致し、なおかつ視覚的にも類似した画像をユーザに提示することが可能となる。

2.1 意味論的類似度

画像 i の意味論的類似度 $f_{\mathbf{R}}(Z_i)$ を求めるために、まずログ情報から画像 i, j 間の意味論的相関係数 $c_{i,j}$ を求める。意味論的相関係数 $c_{i,j}$ は、画像 i, j 間の関連性を示す。 $r_{t,i}$ と $r_{t,j}$ はそれぞれ t 回目の検索において正例の場合 1、負例の場合 -1、評価されなかった場合は 0 となる。式 (1)、(2) から、画像 i, j のどちらかが正の評価を受けた場合のみ $\delta_{t,i,j} = 1$ となり、 $c_{i,j}$ の算出に寄与する。

$$c_{i,j} = \sum_{t=1}^m \delta_{t,i,j} r_{t,i} r_{t,j} \quad (1)$$

$$\delta_{t,i,j} = \begin{cases} 1 & , r_{t,i} + r_{t,j} \geq 0 \\ 0 & , r_{t,i} + r_{t,j} < 0 \end{cases} \quad (2)$$

ログ情報の中に正例が多い場合には $c_{i,j}$ の算出に寄与する項が多くなり、より精密な相関関係が表せるのに対し、正例が少ない場合にはこれが少なくなるため、おおまかな相関関係しか表せない。そのため、ログの質を高めるためには、より多くの正例を蓄積させることが必要となる。

$f_{\mathbf{R}}(Z_i)$ は、 $c_{i,j}$ を用いて以下のように求める。

$$f_{\mathbf{R}}(Z_i) = \max_{k \in P} \{c_{i,k}\} - \max_{k \in N} \{c_{i,k}\} \quad (3)$$

P は RF を得るためにユーザに提示された画像のうち正の評価を受けた画像群、 N は負の評価を受けた画像群を表している。式 (3) より、 $f_{\mathbf{R}}(Z_i)$ が正の値をとれば画像 i は正例、負の値をとれば負例と推測される。

2.2 視覚的類似度

視覚的類似度 $f_{\mathbf{X}}(Z_i)$ は画像の特徴ベクトル間のマンハッタン距離を用いて得られる類似度である。これを用いてデータを2分する境界を求め、正と負の識別を行う。

$f_{\mathbf{X}}(Z_i)$ は、以下のように求める。

$$f_{\mathbf{X}}(Z_i) = \frac{\sum_{k \in P} 1/S(\mathbf{x}_{Z_i}, \mathbf{x}_k)}{|P|} - \frac{\sum_{k \in N} 1/S(\mathbf{x}_{Z_i}, \mathbf{x}_k)}{|N|} \quad (4)$$

\mathbf{x}_{Z_i} は画像 i の特徴ベクトル、 $|P|, |N|$ は正例数と負例数を表している。また、 $S(\mathbf{x}_{Z_i}, \mathbf{x}_k)$ は画像と画像間のマンハッタン距離である。式 (4) により、 $f_{\mathbf{X}}(Z_i)$ が正の値をとれば画像 i は正例、負の値をとれば負例と推測される。

2.3 複合的類似度

意味論的類似度 $f_{\mathbf{R}}(Z_i)$ と視覚的類似度 $f_{\mathbf{X}}(Z_i)$ から、複合的類似度 $f_{\mathbf{H}}(Z_i)$ を求める。

$$f_{\mathbf{H}}(Z_i) = \begin{cases} \tanh(f_{\mathbf{R}}(Z_i) + 1)|f_{\mathbf{X}}(Z_i)| & , f_{\mathbf{R}}(Z_i) > 0 \\ f_{\mathbf{X}}(Z_i) & , f_{\mathbf{R}}(Z_i) = 0 \\ \tanh(f_{\mathbf{R}}(Z_i) - 1)|f_{\mathbf{X}}(Z_i)| & , f_{\mathbf{R}}(Z_i) < 0 \end{cases} \quad (5)$$

このように、 $f_{\mathbf{R}}(Z_i)$ が正の場合には、 $f_{\mathbf{X}}(Z_i)$ の正負に関わらず $f_{\mathbf{H}}(Z_i)$ も正になり、 $f_{\mathbf{R}}(Z_i)$ が負の場合には $f_{\mathbf{H}}(Z_i)$ も負の値をとることから、意味論的類似度が重視された統合方法であることがわかる。この複合

Applying optimum-path forest classifier to CBIR using hybrid similarity

[†] Yuma HIROSE

^{††} Motonobu HATTORI

Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi ([†])

Faculty of Interdisciplinary Research, University of Yamanashi (^{††})

的類似度を用いることで、画像の概念が一致し、なおかつ視覚的特徴が類似した画像の検索が可能になる。これを用いてデータベース内の画像をランク付けし、上位の画像を検索結果としてユーザに提示し (Return set), 境界付近の画像を RF としてユーザに提示する (Label set).

3 提案手法

HySim 法では、式 (4) よりマンハッタン距離の逆数に基づいて境界を求め、SRF の学習を行う。しかし、この学習方法はその他の教師あり学習の中でも消極的な学習に分類され、ユーザからの RF が十分に活用されていない恐れがある。このような学習によって求められた境界は信頼性が低く、正負の真ん中ではなく、どちらかに偏ったものになると考えれ、全体の比率から考えると境界は負例側に偏ることになる。すると、ログには負例が多く、正例は少なく蓄積され、意味論的類似度は画像間のおおまかな相関関係しか考慮することができず、検索に悪影響を及ぼす可能性がある。

そこで、提案手法では複合的類似度を用いた検索における SRF の学習に OPF 分類器を導入することによって、よりログの質を高める方法を提案する。また、クラス分類において良い性能を持つ OPF 分類器に未分類画像を分類させることで更なる検索精度の向上を目指す。

3.1 OPF を用いた CBIR

OPF 分類器は画像データベースを未分類画像・訓練画像・プロトタイプに分けて分類を行う。まず、訓練画像からプロトタイプを決定し、それを根としてプロトタイプまでの距離の重みを考慮したコストを使用し、訓練画像全体にツリーを張る。次に未分類画像全体へツリーを張ることで OPF を構成し、未知の画像の正負を判断していく。OPF を導入した CBIR が Silva らにより提案されている [2]。実験により OPF 分類器は SVM 分類器よりも効率的な RF が行われることが示されている。提案手法のアルゴリズムを以下に示す。

Step1. 初期入力 ユーザは問い合わせとして正例の画像を 1 枚システムに与える。

Step2. Label Set の提示 入力された正例画像から距離的に近い、ラベル付けされていない画像を N 枚選択し、ユーザに Label Set として提示する。

Step3. Relevance Feedback ユーザは Label Set に対して正例、負例のラベル付けを行い、システムにフィードバックし、学習データに追加する。また、式 (3) で得られた意味論的類似度が正となる画像も学習データに追加する。

Step4. OPF による学習 学習データを元に複合的類似度を式 (5) より導出し、それを距離コストとして OPF 分類器の学習を行う。

Step5. Return Set の提示 OPF 分類器によって正例と判断された未分類画像のうち、負例プロトタイプ (境界) から遠い順に N 枚を検索結果として提示する。ユーザが望む結果が得られなかったとき次の手順に進む。

Step6. Label Set の提示 OPF 分類器によって正例と判断された未分類画像のうち、負例プロトタイプ (境界) から近い順に N 枚を Label Set として提示する。

Step7. Step3 に戻る

4 計算機実験

実験データとして、カリフォルニア工科大学から提供されている Caltech256 より、1000 枚 (10 クラス×100 枚) を実験データセットとした。特徴ベクトルは、Deep Convolutional Neural Net (DCNN) が獲得した 4096 次元の特徴を主成分分析により次元削減し、128 次元とした。検索では、あらかじめ画像データに付加されたラベル情報を利用した RF を行った。1 回の入力に対し 5 回の RF を行い、1 回の画像で評価される画像数は 20 枚とした。1000 枚の画像を 1 枚ずつ入力として 1000 回の検索を行い、RF の回数に対する適合率の平均を調べた。各パラメータは、各論文に準拠した値とした。結果を図 1 に示す。

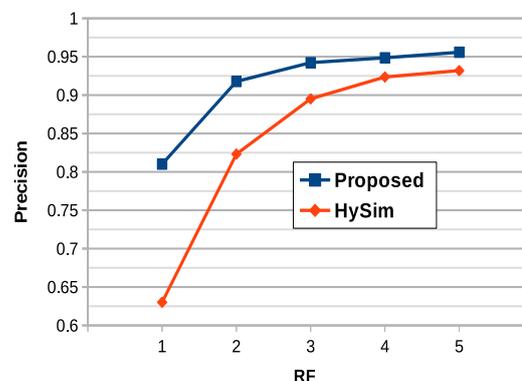


図 1: 検索回数 1~1000 における適合率の比較

この結果から、提案手法は従来法と比べて適合率が全ての RF において向上し、初期段階から上回っていることがわかる。このことから、提案手法により効率的な RF が行われ、ログの質が向上したと考えられる。

5 まとめ

本研究では、従来の複合的類似度を用いた検索システムを改良することを目的とし、SRF に OPF 分類器の学習を取り入れた複合的類似度による検索システムを提案した。計算機実験により提案手法は従来法と比べて少ない RF で高い検索精度を実現することができ、OPF 分類器導入の有効性を示すことができた。

参考文献

- [1] W. Jun, S. Hong, L. Yi-Dong, X. Zhi-Bo, L. Ming-Yu, W. Chun-Li, "A new CBIR approach based on relevance feedback and optimum-path forest classification," *Pattern Recognition*, vol.46, pp.2927-2939, 2013.
- [2] A. Silva, A. Xavier, L. Magalhaes, "Learning a hybrid similarity measure for image retrieval," *Journal of WSCG*, vol.18, no.1-3, pp.73-80, 2010.