

対象物スケッチによる風景画像検索における候補画像の選択法の検討

2G-3

椋木 雅之 美濃 導彦 池田 克夫

京都大学工学部

1 はじめに

画像を検索条件とする画像検索では、画像間の類似度を数値的に定義し、検索者が選択または描画した画像に類似した画像を順に提示する手法がとられている。このような類似画像検索では、自動生成したインデックスに誤りがある場合でも複数の候補を提示することにより、検索要求にこたえられる。しかし、検索システムが蓄積しているデータ内に検索要求に合う画像がない場合にも、いくつかの候補を提示してしまう⁽¹⁾。

一方、通常のキーワード検索では、検索者が指定したキーワードを付加されたデータのみを提示する。このような条件検索では、システム内に検索要求に合うデータがない場合には、候補を提示しない。大量のデータを様々な観点から何度も検索したい場合、不要な候補を提示しない方が望ましい。

そこで本研究では、画像検索において候補画像の数を限定できるような検索条件を定め、条件検索を可能とする方法を検討する。検索条件は、類似画像検索の「インデックスに含まれる誤りに強い」という特性を残したものをを用いる。また、条件検索に適したインデックスの自動生成方法についても検討する。

2 画像検索における条件検索

対象物ラベルの種類(L)毎に、以下の条件を満たすものを候補画像とする条件検索を考える。

$$\left(\begin{array}{l} \text{対象物スケッチの対} \\ \text{象物ラベル} = (L) \\ \text{かつ} \\ \text{インデックス画像対} \\ \text{象物ラベル} = (L) \\ \text{の画素の数} \end{array} \right) > \left(\begin{array}{l} \text{対象物スケッチの対} \\ \text{象物ラベル} = (L) \\ \text{かつ} \\ \text{インデックス画像対} \\ \text{象物ラベル} \neq (L) \\ \text{の画素の数} \end{array} \right)$$

スケッチによる類似検索⁽²⁾では、対象物スケッチとインデックス画像とで同じ位置にある画素の

Another candidate selection method for outdoor image retrieval system using object sketch,
MUKUNOKI Masayuki,
mukunoki@kuis.kyoto-u.ac.jp, Kyoto Univ.

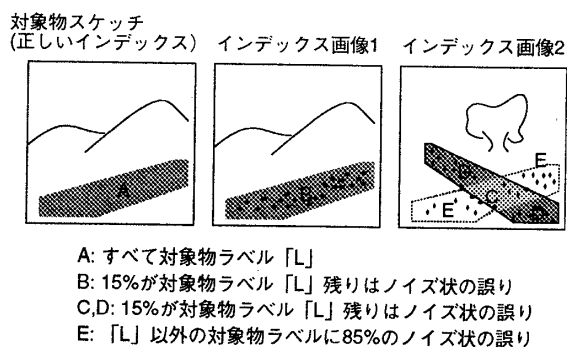


図1: スケッチによる類似検索の特性

対象物ラベルを比較して類似度を計算している。図1の対象物スケッチにランダム誤りを加えた画像をインデックス画像1とし、別の原画像に対する対象物スケッチにランダム誤りを加えた画像をインデックス画像2とする。このとき、対象物スケッチとインデックス画像1との類似度が、インデックス画像2との類似度より大きければ、正しく検索できる。

図1のA, BおよびC∪Eが同じ位置にある画素の集合とする。認識誤りをランダム誤りによりシミュレートし、認識率が15%になるように設定した。この場合、AとBの部分では、85%のランダム誤りが加わっているので、類似度は、0.15程度になる。AとCの共通部分でも類似度は、0.15程度になるが、AとEの共通部分では、Eがもとも(L)以外の対象物ラベルもっていたとし、ランダム誤りが等確率で生じるとすると、対象物ラベルの種類はすべてで10種類であるので、類似度は、 $0.85/9 = 0.094$ 程度になる。従って、Aとの共通部分で比較すると、インデックス画像1がインデックス画像2よりも類似度が大きい。対象物スケッチの他の部分についても同様なので、15%の認識率でも正しく検索できる。

このことから、認識誤りをランダム誤りにより近似できるならば、最初に挙げた条件検索でも同様に、認識率が低くても必要な画像は候補として提示され、不必要な画像は提示されにくいと考えられる。

表 1: 認識率

項目	インクリメンタル学習 (%)	従来の学習 (%)
非学習データ	56.9	70.7
学習データ	61.0	73.8

3 条件検索に適したインデックス生成法

上記の検索手法の特性は、「同じ対象物ラベルを持つべき画素の集合それぞれで、正しくラベルづけられた画素の数が、他のどのラベルに誤ってラベルづけられた画素の数より多ければ検索できる」と言い表せる。この特性を考慮した学習法として、インクリメンタルな学習法を導入する。この学習法は以下のような処理からなる。

1. 人間が原画像を見ながら対象物スケッチを描き、学習サンプルを用意する。
2. まず、1枚の画像からとった学習サンプルを使って、学習する。
3. 学習は、学習データの各対象物（カテゴリ）ごとに、そのカテゴリの学習データで、出力が正しいデータの数が、他のどのカテゴリと誤るデータより多くなるまで続ける。
4. 別の画像からとった学習サンプルを一つずつ順に学習データに加えて、3.を繰り返す。

4 実験・結果

実験ではまず、インクリメンタルな学習法をインプリメントし、学習した。30枚の画像から学習データを、他の100枚の画像から評価用データを生成した。

対象物モデルの構築には、3層の階層型ニューラルネットを用いた⁽²⁾。入力の特徴量は、色・位置・テクスチャ性からなる19種類の特徴量、識別する対象物ラベルの種類は10種類、中間層数は11ノードとした。学習済みの対象物モデルについて、認識率を表1に示した。従来の学習法⁽²⁾にくらべ、認識率では劣っている。

次に、この対象物モデルを用いてインデックス画像を生成し、上記の条件検索を行った結果を表2に示す。検索には、学習データ/評価用データを生成する際にも用いた、人間が原画像を見ながら描いたスケッチを用いた。表中の検索率は、検索件数のうち検索目標画像が提示された件数、再現率は、

表 2: 検索率, 再現率

すべての画像	インクリメンタル学習	従来の学習
検索率	49/130 (37.7%)	27/130 (20.8%)
再現率	49/198 (24.7%)	27/236 (11.4%)
学習した画像	インクリメンタル学習	従来の学習
検索率	26/30 (86.7%)	4/30 (13.3%)
再現率	26/42 (61.9%)	4/30 (13.3%)

候補として提示された画像に対する正しい検索目標画像の割合である。

類似検索は、候補画像を提示する順序を決めるものであるのに対し、条件検索では、提示すべき画像の条件が明確に定まるので、検索率・再現率が共に高い値を示すことが望ましい。この検索方法で比較すると、インクリメンタルな学習法の方が、検索率・再現率とも従来の学習法より高くなっている。特に、対象物モデルの学習に用いた画像では、高い値を示しており、検索事例から追加学習する場合特に有効であることが示されている。しかし、非学習データも含めた検索率・再現率は、認識率の高さにくらべ、必ずしも高くない。これは、

- 認識誤りにはカテゴリ毎に偏りがあり、ランダム誤りによる近似からはずれる
- インクリメンタル学習法は、カテゴリ単位で学習の終了条件を定めているため、各画像毎に検索条件を満たすインデックスが生成されとは限らないことが原因と考えられる。

5 おわりに

画像検索システムにおいて、条件検索を可能とする候補画像選択法について検討した。現状では、検索率・再現率が不十分であるため、提案したインクリメンタル学習法を改良して、この問題に対処することが今後の課題である。

参考文献

- [1] 坂内 正夫: 画像検索技術, 信学誌 vol.71 no.9 pp.911-914 (1988-9).
- [2] 椋木 雅之, 美濃 導彦, 池田 克夫: 対象物スケッチによる風景画像検索とインデックスの自動生成, 信学論 vol.79-D-II no.6 pp.1025-1033 (1996-6).