

意味的な記述による画像の検索手法の一提案

3E-4

柳沼良知 坂内正夫
東京大学生産技術研究所

1 はじめに

膨大な量の画像データベース中から、必要な時に、必要な画像を効率的に獲得するためには、検索のためのキーワードを、人手に頼ることなく、自動的に抽出することが必要となる。このように、画像からキーワードを自動抽出し、画像を検索する手法としては、様々なものが提案されているが⁽¹⁾、本稿では、画像の配色情報に注目し、「山の絵」や、「暖かい感じの絵」といった、意味的な記述による画像の検索手法について述べる。

2 意味的な記述による画像の検索手法

画像中の物体の認識を行なう場合、もし、その物体に特徴的な配色が存在するならば、画像中からその配色の存在/非存在を抽出することにより、ある物体が存在するかしないかの大きな推定を行なうことができる。

例えば、画像全体のヒストグラムをとるならば、このヒストグラムは、大きな画像の配色を表しているに過ぎない。一方、画像を再帰的に（例えば、Quad treeのように）分割していき、その最下位レベルにおいて色のヒストグラムをとるならば、それは、各画素の色情報を表している。そして、これら、2つのレベルの中間の配色情

報を用いるならば、ある物体が存在するかしないかの大きな推定を行なうことができる。すなわち、あらかじめ、抽出すべき物体の配色情報を学習し、物体の画像中で想定される大きさ程度の領域における配色情報と比較することにより、ある物体が存在するかしないかの大きな推定を行なうことができる。同様に、「暖かい感じの絵」といった意味的な記述と、画像の階層的な配色情報の対応関係を調べることで、ユーザによる意味的な記述によって、画像の検索を行なうことができる。

試作システムにおいて、1枚1枚の画像の階層的な配色情報の管理は、色情報のベクトル表現⁽²⁾を用いることで、ある範囲に存在するデータのみを高速に検索するのに効率的な空間データ構造の1つである BD tree⁽³⁾による管理を行なっている。再帰的に分割した、各画像ごとの色のベクトル表現は、以下のように行なった。

- 画像の RGB を HVC に変換する。
- 色相 (H) のヒストグラムをつくる。
- 色相のヒストグラムの各度数をベクトルの成分と考え、その大きさを1に規格化する。

こうすることによって、色相のヒストグラムが n 段階とすると、画像（および、その部分）は、 n 次元空間中の単位球上（の第1象限）で管理することができる。

意味的な記述によって、画像を検索するには、あらかじめ、「山の絵」や、「暖かい感じの絵」といった、意味的な記述と、配色の情報を関連づけておく必要がある。このため、画像中の、ある記述に対応する部分を指定することによって、学習により、

A Proposal of an Image Retrieval Method using Semantic Description
Yoshitomo Yaginuma, Masao Sakauchi
Institute of Industrial Science, University of Tokyo

その記述に対応する配色を求めることを行なった。

本手法では、画像(および、その部分)を大きさ1の単位ベクトルとして表現しているため、ある記述に対応する複数の単位ベクトルから、それらを代表する単位ベクトル、および、その広がり(の程度)を求める必要がある(図1)。

ここでは、各ベクトルとの $\cos\theta$ の平均値を最大とするベクトルを、その記述を代表するベクトルとした。すなわち、ある記述に対応する複数のベクトルを \vec{x}_n 、その記述を代表するベクトルを \vec{X} とした場合、

$$\begin{aligned} \frac{1}{n} \sum \cos\theta_n &= \frac{1}{n} \sum \vec{X} \cdot \vec{x}_n \\ &= \vec{X} \cdot \left(\frac{1}{n} \sum \vec{x}_n \right) \quad (1) \end{aligned}$$

を最大とする \vec{X} を、その記述を代表するベクトルとして求めた。この値は、 \vec{X} の方向が、 $\frac{1}{n} \sum \vec{x}_n$ と同じ場合、最大値をとる。そのため、 \vec{X} の大きさが1であることを使えば、 \vec{X} は、 $\frac{\frac{1}{n} \sum \vec{x}_n}{|\frac{1}{n} \sum \vec{x}_n|}$ と求まる。また、この時の $\cos\theta$ の平均の値が大きければ、大きいほど、その記述に対応する複数のベクトルの広がり(の程度)は小さくなる。このため、この時の θ の値は、複数のベクトルの広がり(の程度)を表すパラメータとして用いることができる。 θ の値は、式(1)により、

$$\theta = \cos^{-1} \left| \frac{1}{n} \sum \vec{x}_n \right| \quad (2)$$

と求めることができる。

実験例を図2に示す。予め、海に特有な配色を学習させておき、「海が存在する絵」の検索を行なった例である。この例では、システムが提案した10候補のうち、5つが実際に海の絵であった。

3 おわりに

本稿では、画像の配色情報に注目し、「山の絵」や、「暖かい感じの絵」といった、意味的な記述による画像の検索手法について

述べた。今後は、より多数の画像について本手法を適用し、その性能評価、および、有効性について更に、検討を行なっていく予定である。

文献

- [1] M.Sakauchi, J.Yamane: "Realization of fully automated keyword extraction in image database systems", SPIE proceeding Vol.1771 Applications of Digital Image Processing XV, pp.67-pp.76, (1992.7)
- [2] 柳沼 良知, 坂内 正夫: "色相を用いた画像のベクトル表現による画像の検索手法の一提案", 情報処理学会第47回全国大会, 3C-5 (1993).
- [3] Y.Ohsawa, M.Sakauchi: "BD-Tree: A New N-dimensional Data Structure with Efficient Dynamic Characteristics", Proc. of 9th World Computer Congress, pp.539-544, (1983).

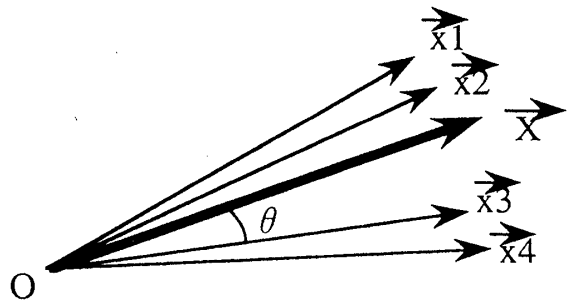


図1 配色と意味的記述の対応づけ

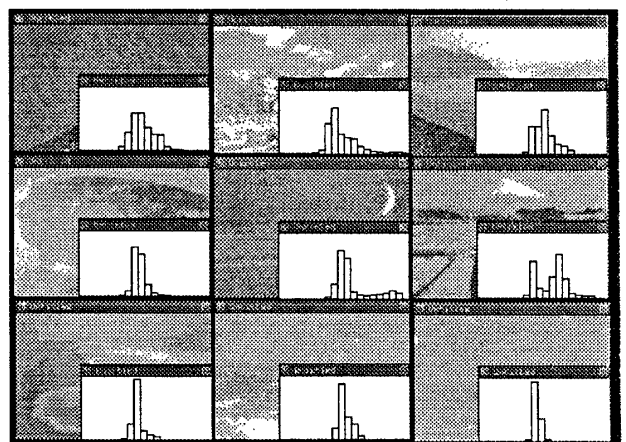


図2 「海の絵」の検索例