

階層成長型自己組織化マップを用いた 内容に基づく画像検索の距離尺度の改良

Improvement of Image Distance in Content-based Image Retrieval
Using Growing Hierarchical Self-Organizing Quadtree Map

中島 聖志†

Kiyoshi Nakajima

服部 元信†

Motonobu Hattori

1. まえがき

従来、デジタル画像の検索は人間の手により作成されたラベルによって行われてきた。しかし、近年の情報技術の進歩に伴い扱う画像の数が増大しているため、従来の手法に代わるものが必要とされている。

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)システムの研究が行われている。このシステムは、画像の内容の情報の特徴ベクトルとして表現して画像間の距離を比較することにより、データベース内に蓄積された画像群の中からユーザの提示した問い合わせ画像に類似した画像を提示するシステムである。

CBIR システムに、ニューラルネットワークの手法である木構造型の自己組織化マップ(SOM: Self-Organizing Map)を組み込むことによって検索の効率化を図ったモデルとして、Wu らによる階層成長型四分木 SOM(GHSOQM: Growing Hierarchical Self-Organizing Quadtree Map)[1]がある。Wu らは、GHSOQM を CBIR システムに組み込むことにより、検索速度を向上させるとともに、良好な検索精度が得られることを示した。

しかし、Wu らが用いた画像距離は2枚の画像の類似性を単方向のみの尺度で測るものであるため、全く似ていない画像同士でも近い距離になることがあるという問題がある。本研究では、GHSOQMを用いたCBIRシステムにおける画像距離において、人間が行うような比較方法である双方向の距離尺度を用いることにより、検索精度の向上を図る。

2. GHSOQMを用いたCBIR

GHSOQMは四分木の構造を持ち、各階層にてSOMによりデータを自己組織的に分類する。図1に示すように、GHSOQMは、上の階層のニューロン(親)が 2×2 の4つのニューロンからなるSOM(子)を持つ木構造となっている。

GHSOQMをCBIRシステムに組み込むにあたり、画像の内容としての特徴ベクトルを各ニューロンに学習させる。GHSOQMのアルゴリズムは次のようになる。

Step.1 初期化

初期状態は根のみである。根(親)から1層目のSOM(子)を生成し、根に全学習データを割り当てる。Step.2へ移行する。

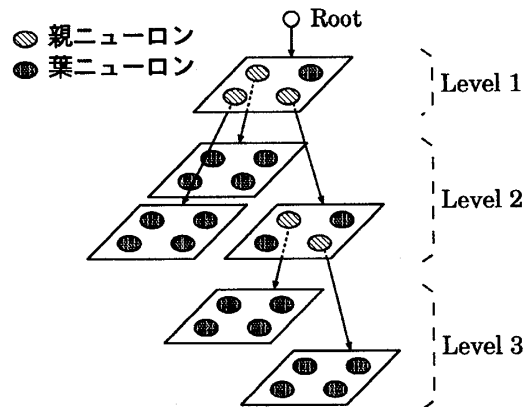


図1: GHSOQMの構造

Step.2 学習およびデータの割り当て

子のSOMに親が持つデータ群を学習させた後、親が持つデータ群の特徴ベクトルそれぞれに対して最も近い重みを持つニューロンにデータを割り振る。Step.3へ移行する。

Step.3 判定

子SOMのニューロンが持つデータ数が定数 τ を越えるとき、 2×2 のニューロンからなる子SOMを生成し親となり、Step.2へ移行する。そうでなければ、そのニューロンは葉となる。全ての子が葉となったとき、GHSOQMは学習を終える。

学習を終えたGHSOQMの葉ニューロンは、最大 τ 枚の画像へのインデックスを持つ。

このアルゴリズムでは、全データは最初の階層のSOMで大きく4つに分類される。そして親が生成した子SOMに、親が持つデータ群を学習させることにより、子SOMはそのデータ群の特徴を、上の階層よりも細かく分類していくように学習することになる。GHSOQMを組み込んだCBIRシステムでは、データベース部を木構造として構築することにより、画像の検索時間を減少させることができる。

3. 単方向による画像距離

GHSOQMを組み込んだCBIRシステム[1]における画像の内容の抽出方法、内容の比較の方法に関して述べる。

†山梨大学大学院医学工学総合教育部, 甲府市

†山梨大学大学院医学工学総合研究部, 甲府市

まず、画像全体から視覚的に類似したピクセルの集合を1つの領域とし、原画像を複数の領域に分割する。次に、分割した領域から「色×6」、「テクスチャ×6」の特徴量からなる特徴ベクトルとして R 、「画像全体の面積に対する領域の面積の割合」として f を抽出する。即ち、1枚の画像の内容は領域数分の R と f によって表現される。

Wuらは、画像 X における i 番目の領域の特徴ベクトルと面積の割合を、それぞれ R_i^X , f_i^X と表すようにしたとき、 n 個の領域を持つ画像 A から、 m 個の領域を持つ画像 B に対する距離 $Dis(A, B)$ を次の手順により計算している。

1. 距離行列 $D(n \times m)$ を作成する

$$D = \begin{bmatrix} d_{11} & \cdots & d_{1m} \\ \vdots & \ddots & \vdots \\ d_{n1} & \cdots & d_{nm} \end{bmatrix}, d_{ij} = \|R_i^A - R_j^B\|^2 \quad (1)$$

2. D の各行の最小値 $D_i (i = 1, \dots, n)$ を取得する
3. 画像 A から画像 B への距離を次式により算出する

$$Dis(A, B) = \sum_{i=1}^n (D_i \times f_i^A) \quad (2)$$

即ち、画像 A の領域 i に対して最も似ている画像 B の領域 j とのユークリッド距離に、画像 A の領域の面積の情報を考慮したものが、画像間の距離となる。

以上のような、Wuらが用いた距離は単方向のみの類似性を測る尺度であり、画像 B に、明らかに画像 A が持つ内容と異なる領域が存在していても、距離が小さくなる場合がある。このようなときには、 $Dis(B, A)$ の値が大きくなることが多い。

4. 双方向の画像距離

CBIRシステムにおいては距離の定義により検索の精度が変化するため、これを人間の感覚に適合するように定義することが、内容に基づく画像の検索における精度の向上につながる。

人間は、2枚の画像の内容を比較するときに、相互に見比べることにより類似性を測る。よって、式(2)において、逆方向の距離である $Dis(B, A)$ も考慮に入れることにより、検索精度を向上できると期待される。この双方向の距離 $BiDis(A, B)$ を式(2)を用いて次式として定義する。

$$BiDis(A, B) = Dis(A, B) + Dis(B, A) \quad (3)$$

この距離は画像 A 側と画像 B 側の両方において、共通の内容を持つ領域が多ければ多い程、距離が小さくなる性質を持つものである。また、逆方向の距離である $Dis(B, A)$ においては、 $Dis(A, B)$ の計算時に作成した距離行列 D の転置行列を利用する。なお、SOMの学習時にはこの双方向の距離を考慮した更新を行う。

5. シミュレーション

従来の単方向の画像距離による GHSOQM を用いたとき(従来法)の検索精度と、提案手法である双方向の画像距離による GHSOQM を用いたとき(提案法)の検索精度を比較した。

画像群には、データクラフト社の著作権フリー画像集「フォトバイブル20000」に収録された「蝶、花、犬、街、夜空、海、料理、木、山、空」の10種類のクラスを用いた。各クラスから100枚ずつ、計1000枚を選択し、これらを用いて CBIR システムを構築した。

5.1 評価尺度

システムが提示した画像数を k 、その中に問い合わせ画像と同じクラスが存在する数を n_k としたとき、検索の精度を示す指標である適合率は $P(k) = n_k/k$ として表される。

1枚の画像を入力として与えたときの最大画像提示数 k_{max} を20とし、平均適合率 $P_{ave} = \sum_{k=1}^{k_{max}} P(k)/k_{max}$ を算出する。そして、全ての画像を1枚ずつ入力として与えたときにそれぞれ算出される、平均適合率の値の平均を評価値とした。

5.2 結果

学習回数1000における、従来法の結果と、提案法の結果を表1に示す。表1における「全画像」は全画像1000枚を入力として与えたとき、「蝶」、「花」などはそのクラスの画像群100枚を入力として与えたときの評価値を意味する。なお、GHSOQMの学習結果は毎回若干異なるため、従来法、提案法ともに10回GHSOQMを生成し、得られた評価値の平均値を表1に示した。

表1: 入力画像群に対する評価値(10回の試行の平均)

入力画像群	従来法	提案法
全画像	0.677	0.720
蝶	0.783	0.899
花	0.536	0.541
犬	0.603	0.687
街	0.590	0.484
夜空	0.761	0.924
海	0.610	0.706
料理	0.895	0.873
木	0.475	0.576
山	0.672	0.637
空	0.844	0.872

全画像、各クラスの多くにおいて提案法の値が従来法を上回る結果となった。なお、全画像を入力として与えたときの各回の評価値10個に対して、 t -検定による左片側検定を行った結果、 $t(18) = -10.45, p < .0025$ となり、有意水準0.25%で有意差が見られた。よって、提案手法である双方向の距離尺度を用いた GHSOQM の有効性が示された。

6. むすび

GHSOQMを用いたCBIRシステムにおける画像間の距離尺度を、単方向から双方向の尺度に変更し、双方向の学習を行わせた。また、従来法と提案法を比較し、検索精度が向上することを示した。

参考文献

- [1] Sitao Wu, M.K.M.Rahman, Tommy W.S.Chow, "Content-based image retrieval using growing hierarchical self-organizing quadtree map", Pattern Recognition, 38, pp.707-722, 2005.