原著論文

画像の内容検索における類似度評価アルゴリズムの利用者フィードバックを用いた最適化†

岡 望実*1・亀山 啓輔*2・寅市 和男*3

近年、画像データベースの検索の際に、画像の特徴を用いて得られる画像間の類似度をもとに検索を行う画像内容検索（CBIR）に関する研究が行われてきている。CBIRでは、共通した客観的定義の存在しない「類似度」が用いられる。この「類似度」は様々な方法で決めることが可能であり、また、ユーザの主観によって異なるものである。この類似度の多様性ゆえに、CBIRにおいてはユーザの類似度基準に合わせた検索の実現が必要となる。

本研究においては、検索結果に対する評価基準を用いて、類似度評価アルゴリズム中のパラメータを最適化することで、画像の類似度評価をユーザの要求に適合するよう改善する枠組みを提案する。提案する枠組みでは、検索結果に対する評価基準 J を定め、Particle Swarm Optimization を用いて、 J を最小化することで最適化を行う。

本研究の提案手法の有効性を検証するシステムとして、2 値画像の輪郭を特徴とする画像の内容検索システムを用いた。そして、検索結果の類似度ランキングに対してユーザのフィードバックに基づいて計算される評価基準関数 J を定め、類似度評価アルゴリズム中のパラメータ最適化を行った。第一実験では、提案した枠組みにより最適化を行い、評価基準にあった検索結果が得られるようになることを確認した。第二実験では、2 つの異なる評価基準 J を定め、それぞれの評価基準 J について最適化を行い、要求される基準の違いにより最適化後のパラメータが異なることや、検索結果に要求に応じた違いが出ることを確認した。以上により、提案した枠組みを用いて、評価基準 J を適切に定め、それに基づいてパラメータ最適化を行うことにより、ユーザの類似度基準に合わせた検索システムを実現できることを確認した。

キーワード：画像内容検索（CBIR）、類似度、Particle Swarm Optimization、弛緩整合法、関連性フィードバック

1. 序章

1.1 画像の内容検索

画像データベースの検索においては、キーワードを問い合わせとする場合と、画像そのものを問い合わせとする場合がある。また、キーワードと画像情報の両方を用いて検索を行う場合もある。この中で、画像そのものを問い合わせとして用いる場合の画像の検索を画像の内容検索（CBIR）と呼ぶ。CBIRでの画像の検索は、画像を解析することで得られる特徴量を用いて行われる。この技術を利用すれば、キーワード付与など

大規模画像データベースの構築にかかる労力を大幅に削減でき、画像検索の基準（類似度など）をユーザの注目する画像の特徴に適するよう変更することが可能となる。

CBIRは、図1に示すように、各辞書画像と問い合わせ画像に対して、モデル化・特徴ベクトル化を行い、その後、各辞書画像と問い合わせ画像の間の類似度を評価し、検索を行うものが一般的である。それらの処理は

1. 画像の部品化
2. 各部品からの特徴抽出
3. 問い合わせ画像と辞書画像の部品間の対応関係の最適化
4. 特徴量と対応関係から両画像の類似度を評価の手順からなる。

Katoらによって行われたTRADEMARKという初期のCBIRシステム[1]においては、商標などの2値図形画像を、あらかじめ決められたグリッドを用いて領域分割し、それを部品としている。ここでは、部品の対応関係は明らかである。次に分割された画像から黒画素の数、ランレングス、局所コントラスト、局所相関

† Optimization of Similarity Evaluation Algorithm Using Relevance Feedback in Content-Based Image Retrieval

Nozomi OKA, Keisuke KAMEYAMA and Kazuo TORAICHI

*1 筑波大学 大学院システム情報工学研究科 リスク工学専攻
Dept. of Risk Engineering, Graduate School of Systems and Information Engineering, University of Tsukuba

*2 筑波大学 大学院システム情報工学研究科 コンピュータサイエンス専攻
Dept. of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba

*3 筑波大学先端学際領域研究センター
Center for Tsukuba Advanced Research Alliance, University of Tsukuba

図1 画像の内容検索の一般的枠組み。

などの特徴量が抽出され、それらは2つの画像の類似度評価に使われる。

また、山本[2]によって手書き文字認識に応用された輪郭の部品化と弛緩整合法によるマッチング手法が、Kwan[3]、Kim[4]、Kameyama[5]らによってそれぞれ商標や花押・家紋CBIRシステムとしてに応用されている。これらのシステムにおいては、部品は輪郭線分であり、問い合わせ画像と辞書画像の部品の対応関係は、弛緩整合法によって最適化され、また、その結果を用いて2つの画像の類似度が評価される。

さらに、近年報告されているCBIRシステムとしては、Squire[6][7]らによるViperなどがある。Viperでは、色やテキストチャの特徴量を用い、検索を行っている。また、これ以外にも人の主観を検索の結果に反映させるための関連性フィードバックに関する研究が盛んに行われている[8]。

画像検索においてキーワードを用いて検索を行うものとしては、Google Image Searchが有名である。また、キーワードと画像特徴の両方を用いる画像収集システムとして、柳井によるProbabilistic Image Collector[9]が発表されているが、そこでは、まず指定キーワードに関連のあるWebページを収集し、html解析により、キーワードに関連があると思われる画像を収集する。その後、色やテキストチャなどの画像特徴を用い、関連画像の絞り込みを行う。

1.2 画像の「類似度」

上記のCBIRやその他多くのCBIRシステムにおいて、画像の検索は、共通した客観的な定義の存在しない「類似度」で規定される。実際、この「類似度」は様々な方法で決めることが可能である。

例えば、赤いバラの問い合わせ画像に対して要求される類似画像の集合は、様々な赤い物体の画像の集合であるかもしれないし、様々な色のバラの画像の集合であるかもしれない。このように、検索で用いる「類似度」はユーザーの要求に合わせる必要がある。この例は、類似度に関する着目点が画像の色や形などの特徴

に関連するので、単純なケースであるといえるが、ユーザーの着目点が必ずしも単一の特徴と簡単には結びつくとは限らない。

前述のTRADEMARKシステムでは上記の問題に対し、特徴空間とユーザーの主観的特徴空間との間に線形写像を定義することで、ユーザーの類似度基準に適用しようとした[1]。しかしながら、類似度評価が、特徴ベクトルの距離計算だけでなく、部品の対応のさせ方にも依存する場合、部品の対応関係を解決するアルゴリズム自体の調整も必要となる。

1.3 本研究の目的

類似度の多様性という観点から、CBIRにおいては、ユーザーの類似度基準に合わせた検索の実現が課題となる。しかし、類似度評価が特徴ベクトルの距離のみでなく、様々なパラメータに影響を受ける場合、最終的に得られる類似度とアルゴリズム中のパラメータの関係が明らかではなく、勾配法などの最適化手法を用いて最適化を行うことが難しい場合がある。

本研究では、直接的な手法では最適化を行うのが難しい類似度評価を行うアルゴリズムを探索的手法を用いて最適化することで、画像の類似度評価をユーザーの要求により適合するよう改善する枠組みを提案することを目的とする。

ここで、類似度評価の変更は図2に示すように、アルゴリズム中のパラメータを調節することで行う。本研究では、アルゴリズム中のパラメータはCBIRの結果に伴う評価値を用いて、必ずしも全域で滑らかではないが、最適点近傍では滑らかであると予想される評価関数に鑑み、過去の探索結果に対するヒューリスティクスに基づく多点探索であるParticle Swarm Optimization (PSO)を用い最適化される。

以下、2章では、弛緩整合法による2値画像のマッチングと類似度評価について説明し、そのアルゴリズムにおけるパラメータの調整の必要性を述べる。3章では、Particle Swarm Optimizationを用いたパラメータの最適化の枠組みについて述べる。4章と5章では、パラメータ最適化実験の結果について述べ、6章

図2 画像の類似度評価アルゴリズム中のパラメータの最適化の枠組み。IとDはそれぞれ問い合わせ画像の集合と辞書画像の集合。

でまとめとする。

2. 弛緩整合法を用いた部品の対応付けによる画像の類似度評価

2.1 モデリング・部品の対応・類似度評価

本章では輪郭の区分的関数近似と弛緩整合法に基づいた2値画像の類似度評価法について説明する[2][3][4][5]。ただし、提案する枠組は弛緩整合法を用いたケースのみを対象とはしておらず、類似度を評価するアルゴリズム内に、類似度に影響を与えるパラメータが存在する手法一般に関して有効であると考えられる。本研究で扱う画像は図3(a)に示すような2値画像をとし、各画像は輪郭を抽出後、直線・円弧・自由曲線を用いて区分的に関数近似することで部品化される(図3(b))[10]。

2つの画像の類似度を評価する際には、各画像の輪郭線分の組の対応関係の良し悪しが評価される。

もし、多数の輪郭線分において、対応関係のよい他方の画像の輪郭線分が見えできれば、2つの画像は似ていると判断される。この手法は、山本らによって手書き漢字の認識のために開発され[2]、Kwan[3]、Kim[4]、Kameyama[5]によってそれぞれ商標や花押・家紋等CBIRシステムとしてに應用されている。

この手法では、2つの画像の輪郭線分の集合要素間の最良な対応関係が弛緩整合法[11]によって求められる。初期的な対応関係より開始し、部品の対応確率の更新を繰り返すことで、輪郭線分の対応関係が最適化されていく。また、システムのエネルギー関数が2つの画像の輪郭線分の集合の違いの指標となる。

図3 (a) 家紋画像の例 (b) 区分的関数近似による輪郭線分のモデリング

2.2 モデリング

弛緩整合法を用いたCBIRにおいて、画像の i 番目の輪郭線分に関して、図4に示す次のような特徴量が計算される。

図4 輪郭線分の特徴量。

- 始点: $s_i = (x_i^{(s)}, y_i^{(s)})$
- 終点: $e_i = (x_i^{(e)}, y_i^{(e)})$
- 線分の長さ: L_i
- 1/3点から始点への角度: $\phi_i^{(s)}$
- 2/3点から終点への角度: $\phi_i^{(e)}$
- 始点の近傍線分リスト: $N_i^{(s)}$
- 終点の近傍線分リスト: $N_i^{(e)}$

ここで、 $N_i^{(s)}$ と $N_i^{(e)}$ は最大で4近傍まで考慮するものとする。問い合わせ画像(オブジェクト集合)と辞書画像(ラベル集合)のすべての線分に関して、上記の特徴量が抽出される。

2.3 弛緩整合法による部品の対応とアルゴリズム中のパラメータ

2.3.1 弛緩整合法

弛緩整合法[11]はラベル集合の要素とオブジェクト集合の要素との対応を最適化する手法である。

比較される問い合わせ画像と辞書画像は、それぞれ N 個のオブジェクト集合 $B = \{b_1, \dots, b_N\}$ と M のラベル集合 $\Lambda = \{\lambda_1, \dots, \lambda_M\}$ を構成する。オブジェクト i へのラベルの対応は対応確率ベクトル $p_i = [p_{i(1)}, \dots, p_{i(M)}]^T \in R^M$ で表される。ここで“ T ”は転置を表す。ベクトル p_i の要素は $\sum_{k=1}^M p_{i(k)} = 1$ を満たすものとする。全 N 個の p_i の集合は画像と画像の対応の関係を表す状態ベクトル $p = [p_1, \dots, p_N]^T \in R^{MN}$ を構成する。

2.3.2 対応候補のフィルタリング(パラメータ: C_1, C_2)

まず、全てのオブジェクトとラベルの対応を考慮に入れると計算量が膨大となるため、その削減のために、

$$\begin{aligned} |L_i - L_k| &\leq C_1 \min(L_i, L_k), \\ |\phi_i^{(s)} - \phi_k^{(s)}| &\leq C_2 \\ |\phi_i^{(e)} - \phi_k^{(e)}| &\leq C_2 \end{aligned} \quad (1)$$

を満たすオブジェクトとラベルのペアのみを計算の対象とする。ここで、 C_1 と C_2 はどの程度の特徴量の違いを許容するかを決定するパラメータである。

2.3.3 初期対応(パラメータ: W_1, C_3)

候補に挙げた各ペアに関して、ラベルとオブジェクトの対応確率が以下の式に従って計算される。台形型の点広がり関数を $x \geq 0$ の範囲で

$$f_{ps}(x, W, C) = \max\{1 - W \max(x - C, 0), 0\} \quad (2)$$

と定義して、初期対応確率を以下のように計算する。

$$p'_{i(k)} = f_{ps}(|L_i - L_k|, W_1, C_3) \quad (3)$$

$$p^0_{i(k)} = \frac{p'_{i(k)}}{\sum_{k'} p'_{i(k')}} \quad (4)$$

ここで、式(3)は点広がり関数であり、長さの近い線分に対して、高い値を割り当てる。また、 $p^0_{i(k)}$ は $\mathbf{p}_i$ の第 k 要素である。パラメータ W_1 と C_3 は点広がり関数の形を決定するものである。

2.3.4 適合係数(パラメータ W_2, W_3, C_4, C_5)

状態 $\mathbf{p}$ の更新に用いる対称行列 $R = [r_{ij}(k, l)] \in R^{MN \times MN}$ の要素 $r_{ij}(k, l)$ は、 k と i 、 l と j の2つのラベル付与の対応の適切さを反映する値をとるように定義される。適合係数 $r_{ij}(k, l)$ における i, j, k, l の関係については図5を参照されたい。

始点に関する適合係数は

$$\begin{aligned} r_{ij}^{(s)}(k, l) &= \alpha + \beta \\ \alpha &= f_{ps}(|d_{ij} - d_{kl}|, W_2, C_4) \\ \beta &= f_{ps}(|d_{ik}^0 - d_{jl}^0|, W_3, C_5) \end{aligned} \quad (5)$$

のようにして、求められる。ここで、 $d_{ij} = \|\mathbf{s}_i - \mathbf{e}_j\|_E$, $d_{kl} = \|\mathbf{s}_k - \mathbf{e}_l\|_E$, $d_{ik}^0 = \phi_i^{(s)} - \phi_k^{(s)}$, $d_{jl}^0 = \phi_j^{(e)} - \phi_l^{(e)}$ である。

図5 適合係数 $r_{ij}(k, l)$: オブジェクト i とラベル k およびオブジェクト j とラベル l の2対応の適切さの指標。

終点の適合係数 $r_{ij}^{(e)}(k, l)$ も同様にして求められる。パラメータ W_2, W_3, C_4, C_5 は2組のオブジェクトとラベルの対応の違いに関する点広がり関数である。また、この適合係数のために、この手法は画像のシフトや回転に関して頑強である。

2.3.5 弛緩整合法による部品対応の最適化

状態ベクトルは以下に示すように更新される。

$$\mathbf{p}(t+1) = \mathbf{f}(\mathbf{p}(t), \mathbf{q}(t)) \quad (6)$$

$$\mathbf{f}(\mathbf{p}, \mathbf{q}) = [f_{11}, \dots, f_{NM}]^T \quad (7)$$

$$f_{ik} = q_{ik} p_{ik} / \sum_{k'=1}^M q_{ik'} p_{ik'} \quad (8)$$

$$\mathbf{q}(t) = R\mathbf{p}(t) \quad (9)$$

ここで、 t は時間を表す。対称行列 R に関して定められる。

$$-A(\mathbf{p}(t)) = -\mathbf{p}(t)^T R \mathbf{p}(t) \quad (10)$$

は、局所解 $-A(\mathbf{p})$ に収束することが知られている[12]。関数 $A(\mathbf{p})$ の値はラベリングの総合的な評価基準となる[13]。

この繰り返し演算の詳細は山本[2]を参照されたい。

式(1)を満たさず、対応するペアのなかったオブジェクトとラベルは、未対応オブジェクトおよび未対応ラベルとして扱われる。未対応のラベルとオブジェクトの集合はそれぞれ B', Λ' とあらわす。

2.4 類似度評価

最後に2つの画像の距離を以下の式で求める。

$$\Delta = \frac{1}{\sum_i L_i} \left\{ \sum_i \{(1 - R_{i(\kappa)}) L_i\} + \sum_{i' \in B'} L_{i'} + \sum_{k' \in \Lambda'} L_{k'} \right\}$$

ここで

$$\begin{aligned} R_{i(\kappa)} &= \frac{p_{i(\kappa)}^{(Z)} \{\zeta + \xi\}}{2(\sum_{j \in N_i^{(s)}} L_j + \sum_{j' \in N_i^{(e)}} L_{j'})} \\ \zeta &= \sum_{j \in N_i^{(s)}} L_j r_{ij}^{(s)}(\kappa, \sigma) p_{j(\sigma)}^{(Z)} \\ \xi &= \sum_{j' \in N_i^{(e)}} L_{j'} r_{ij'}^{(e)}(\kappa, \sigma') p_{j'(\sigma')}^{(Z)} \end{aligned} \quad (11)$$

とする. 式(11)は, 式(10)に対応するものであり, κ, σ, σ' は, それぞれオブジェクト i, j, j' に最もよく対応したラベルを指すものとする. また, $p_{i(\kappa)}^{(Z)}$ は p を Z 回更新した後の $i(\kappa)$ 成分である.

2.5 本マッチング法をCBIRに応用する上での問題点

以上のCBIR法においては, 入力画像と辞書画像の違いは, 式(11)に示す Δ によって評価される. この基準は2.3節にあげた8つのパラメータの関数となっている.

従来, これら8つのパラメータは経験的に定められたものを用いてきた. 従来の弛緩整合法を用いたCBIRを家紋画像に対して適用した検索例を表1に示す. 問い合わせ画像に対する類似度の上位にある画像で, 問い合わせ画像に似ていると思われるものもあれば, 似ていないと思われるものもあり, また, 主観的に似ていると思われる画像が下位に位置していたりもする.

表1の結果の良し悪しは, 検索結果の関数としての何らかの評価基準 J を定義しなければ評価をすることができない. また, 評価基準 J が異なれば, 検索結果の良し悪しも変わることから, ユーザごとに異なる「類似度」は, 異なる基準 J であると考えることができる.

そこで, 特定の基準 J に適合するように類似度評価アルゴリズムを調整する枠組みが必要となる. しかしながら, 弛緩整合法を用い, 部品の対応関係を最適化し類似度を評価する手法においては, パラメータと評価基準 J との関係を直接的に求めることは容易ではなく, J の勾配に基づく従来の最適化手法などを用いることはできない. 一般にこのような場合, 遺伝的アルゴリズムなどの多点探索手法が用いられることが多い.

本研究においては, パラメータ値の変化は点拡がり関数の形の変化となるので, 最適点付近においては評価関数が連続であると考えられる. このような場合, 評価関数の勾配に相当する情報が有用となるが, 拡がりを持った複数の点の評価値をもとに探索方向を決定するParticle Swarm Optimization (PSO) [14] では勾配に相当する情報を利用可能であると考えられるので, CBIRの検索の結果を改善するためのパラ

メータの最適化を行う上で, PSOを用いる.

本研究では, 図2に示すように, 検索結果に対する評価基準 J を用いることでCBIRシステムの画像の類似度評価アルゴリズム中のパラメータを調整する一般的枠組みを提案する. 本手法は, 特定の評価基準に従う類似度評価と検索結果の改善を可能とする. さらに, 評価基準の変更を可能とし, 違う種類の類似度へ適応するための再調整を可能とする.

3. PSOによる類似度評価の最適化の枠組み

3.1 パラメータの最適化に関する提案手法

ここでは, 画像の類似度評価アルゴリズム中のパラメータの最適化手法について説明する. 図2に示すように, CBIRの結果 $S(I, D; \theta_1, \theta_2, \dots, \theta_K)$ はパラメータ $(\theta_1, \theta_2, \dots, \theta_K)$ に影響を受けている. ここで, I, D, K はそれぞれ問い合わせ画像集合, 辞書画像集合, パラメータの数である. パラメータの最適化は, 評価基準 $J(S(I, D; \theta_1, \theta_2, \dots, \theta_K))$ に対して行われる.

3.2 Particle Swarm Optimization (PSO)

Particle Swarm Optimization は探索空間中の滑らかな軌跡に沿って動く多点における探索を並列的に行う最適化手法である. この手法は1995年にKennedyとEberhartにより, 群知能などをヒントに考案された[14].

粒子と呼ばれる各個体は個々の位置・速度などの情報や, グループ間で共有される情報を用い, 簡単な規則に従って探索を行う. 粒子 i は探索空間内の座標 $x_i(t)$ と速度 $v_i(t)$ の情報を持つ. $x_i(t)$ の位置での適合度は評価関数を用いて評価され, 粒子の動きに反映される. 各粒子は過去最も高い適合度を与えた位置 p_i を記憶しており, また, 群れの中で最も高い適合度を得た位置 g はすべての粒子で共有される.

各粒子は

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (12)$$

$$v_i(t+1) = v_i(t) + \lambda_1(p_i - x_i(t)) + \lambda_2(g - x_i(t)) \quad (13)$$

表1 (A)問い合わせ画像. (B)表2にあげる辞書画像の集合との検索結果. 順位は問い合わせ画像との類似度による.

		(B) Similarity ranking									
Rank		1	2	3	4	5	6	7	8	9	10
Image

(A) Query

に基づき座標と速度を更新し、探索を行う。

式(13)にあるように、速度はその粒子の現在位置から、最もよい適合度を得た位置と、グループで最もよい適合度を得た位置へ向かう方向の速度が加えられるように更新される。また、これらの影響の大きさはパラメータ λ_1 と λ_2 によって決定される。

PSOは、簡単な規則に従っているが、さまざまなベンチマークテストにおいて最適点を効率的に探し出すことが報告されている。また、株式市場や電圧制御などの現実の問題に対するPSOの応用も行われている[15][16]。

3.3 PSOによる画像の類似度評価アルゴリズム中のパラメータの最適化

PSOは以下に示す手順に従って2つの画像の類似度評価アルゴリズム中のパラメータを最適化する。ここで、探索空間は粒子が探索するパラメータ $(\theta_1, \theta_2, \dots, \theta_K)$ によって張られる K 次元ベクトル空間であり、粒子の位置ベクトル x_i が特定のパラメータベクトルに相当する。画像の類似度評価アルゴリズム、およびパラメータの組で決定される検索結果を評価する評価関数 J は、適宜決められるものとする。最適化の処理は図6に示すようになっており、以下の擬似コードであらわされる。ここで G, H はそれぞれ、PSOの最大世代数と粒子の数である。

```

1 Begin
2 for  $i = 1$  to  $H$ 
3   Initialize  $x_i$  and  $v_i$ 
4 for  $t = 0$  to  $G$  {
5   for  $i = 1$  to  $H$  {
6     Set  $x_i(t)$  in the image similarity
      evaluation algorithm and obtain
      image retrieval  $S(x_i(t))$  using
      query and dictionary image sets.
7     Evaluate retrieval as  $J(S(x_i(t)))$ .
8     Update  $p_i$  and/or  $g$  as necessary.
9   }
10 }
11 return  $g$  as the optimized parameter set.
12 End

```

図6 CBIRの反復試行と基準 J による検索の評価を含んだ、PSOによるパラメータの最適化。

4. 実験と考察1：提案手法の検証

4.1 実験の目的

まず、提案手法を用いることで、与えられた評価基準 J に基づいて最適化が行われることを確認するため、家紋画像のコレクションを用いて、弛緩整合法のパラメータを最適化する実験を行った。

4.2 画像セット

表2に示すようにコレクションは6つの主観的に類

表2 実験に用いられたデータベース画像。30枚の画像は主観的に似ている6枚の画像からなる5つのグループに分けられている。

	Group1	Group2	Group3	Group4	Group5
Set1					
Set2					
Set3					
Set4					
Set5					
Set6					

似していると思われる画像を1グループとした5つのグループから構成されている。

画像はそれぞれ白黒2値の400×400ピクセルの大きさの画像である。輪郭の関数化に関する詳細は[10]を参照されたい。

4.3 実験の手順と評価関数

実験では、表2にあるセット1の5枚の画像を問い合わせ画像として順次入力する。各問い合わせについて、問い合わせ画像との類似度の高い順に30枚のすべての画像がランク付けされる。問い合わせがグループ*i*の画像である場合、同じグループの画像を上位にランクした結果を高く評価するため、ランキングは同じグループの画像の順位の和 J_i によって評価される。従って、より低い値の J_i がより良い結果となる。このスコアは $J = \sum_{i=1}^5 J_i$ のようにして5つのすべてのグループについて足し合わされ、現在のパラメータの評価に使用される。ただし、ランキングは0から始まるものとし、最小値(最適値)は5つすべての問い合わせに対し、それぞれ同じグループの画像が上位6位を独占したとき、 $(0+1+2+3+4+5) \times 5 = 75$ となる。この評価値の平均値と最大値(最悪値)はそれぞれ435と795となる。

4.4 実験条件

PSOによって最適化されるパラメータは2章に示した独立な8つのパラメータである。それらの中で C_1 , C_3 , C_4 , C_5 , W_1 , W_2 , W_3 は $[0, 2]$ の範囲で調整され、 C_2 に関しては $[0, 80]$ の範囲で調整されるものとする。PSOの最大世代数は300、弛緩整合法の反復回数は5回とする。また、PSO中のパラメータ λ_1 と λ_2 は各世代で生成される $[0, 2]$ 区間の一様乱数とし、粒子の最大速度は0.05とする。

4.5 結果1：評価値の変化

図7にはPSOの各世代における J の変化を示した。また、最適化において問い合わせ画像として用いていないセット2から5までの画像を用いて問い合わせた平均評価値 J_{test} も示している。この過程で J が極小値に収束することは[17]にて報告した。11世代の後、 J は最小値に収束しているが、 J_{test} は2世代目に最小値

図7 PSOの世代における検索評価値 J の変化。問い合わせ画像として、最適化の問い合わせ画像集合(セット1)とそのほかのセットの平均に関する値をプロット。

をとっている。

本手法においても、学習アルゴリズムにおいて知られている過学習が問題となりうると考えられる。過学習を防止するためにPSOを早い世代でとめることがパラメータの値を最適化するのに貢献するかについては、更なる検証が必要である。本論文においては、表3に示す $J=75$ の場合に相当するパラメータセットを検索実験に使用する。

4.6 結果2：最適化後のアルゴリズムによる検索結果

表4と5には、最適化前のパラメータと最適化後のパラメータを用いた場合の検索結果のランキングを示している。

表4に示す問い合わせ画像は表2に示す画像集合には含まれているが最適化の際の問い合わせ画像に使われたセット1には含まれていない画像である。評価値 J_5 が43から16に減少していることから検索の結果が改善されていることは明らかである。

表5には全く新しい問い合わせ画像に対する結果を示している。表5の問い合わせ画像は、表2には含まれていないので評価値 J に関する考察は行えないが、最適化後の検索結果の方が、菱形の枠を持った画像や銀杏の葉を持った画像が上位に位置しており、より良い結果になっているものと思われる。

表3 最適化前後のパラメータの値。

Parameter	C_1	C_2	C_3	C_4	C_5	W_1	W_2	W_3
Initial value (empirical)	0.1	15	0.1	0.5	0.5	0.1	0.05	0.01
Optimized (300 generation)	1.835	3	0.815	1.691	1.964	0.004	0.227	0.696

表4 表2に含まれる問い合わせ画像(A)との類似度による検索結果。(B)に示される画像はパラメータの最適化を行う前の検索結果($J_5=43$)。(C)最適化後の検索結果($J_5=16$)。

(B) Before optimization										
Rank	1	2	3	4	5	6	7	8	9	10
Image										
Distance	0.0	1.270	1.325	1.395	1.397	1.400	1.407	1.427	1.457	1.502

(C) After optimization										
Rank	1	2	3	4	5	6	7	8	9	10
Image										
Distance	0.0	0.793	0.809	0.904	0.916	0.987	1.002	1.055	1.060	1.061

表5 表2に含まれない画像(A)を問い合わせとしたときの検索結果。(B)は最適化前の検索結果。(C)最適化後の検索結果。

(B) Before optimization										
Rank	1	2	3	4	5	6	7	8	9	10
Image										
Distance	1.633	1.727	1.745	1.774	1.868	1.870	1.908	1.914	1.967	1.977

(C) After optimization										
Rank	1	2	3	4	5	6	7	8	9	10
Image										
Distance	0.992	0.999	1.0	1.0005	1.001	1.0052	1.0059	1.006	1.014	1.016

2つの結果は基準値 J を用いた類似度評価の最適化の有効性を示している。

5. 実験と考察2: 評価基準 J の違いによる変化

5.1 実験の目的

次に、最適化すべき基準を変更した場合にそれが検索結果に反映されるかを検証するための実験を行った。

5.2 画像セット

表6に実験に用いた画像を示す。これらの画像は、表2のグループ5の6枚の画像を基本とし、これをセット1とする。さらに、セット2をセット1の画像を5度回転させた画像、セット3を-5度回転させた画像、セット4を95%に縮小した画像、セット5を105%に拡大した画像とする。

表6 実験に用いられたデータベース画像。30枚の画像は実験1に用いた表2のグループ5の画像を基本とし、5度回転、-5度回転、95%に縮小、105%に拡大した画像からなる。(*)を付与した画像は、最適化の際に問い合わせ画像として用いた画像。

	Group1	Group2	Group3	Group4	Group5	Group6
Set1 (基本)						
Set2 (回転+5度)						
Set3 (回転-5度)						
Set4 (縮小95%)						
Set5 (拡大105%)						

5.3 評価関数

ここでは、表6にあるグループ4セット1の画像を問合わせ画像として入力する。この時、回転を優先させて最適化する場合には、検索結果のランキングに対して、グループ4のセット1、4、5の画像の順位とグループ4のセット2、3の画像の順位を2倍にしたものの和を評価値 J とする。スケールを優先させて最適化する場合には、グループ4のセット1、2、3の画像の順位とグループ4のセット4、5の画像の順位を2倍にしたものの和を評価値 J とする。これにより、ユーザの要求する種類の画像の重み付けを変更し、それぞれの評価基準 J について最適化を行った場合に、それらの意図が結果に反映されうるかを見るものである。

ただし、ランキングは0から始まるものとし、最小も値(最適値)は同じグループの画像が上位5位を独占したとき、 $0+1+2+3+4+(1+2)=13$ となる。

5.4 実験条件

PSOによって最適化されるパラメータは2章に示される独立な8つのパラメータである。それらの中で C_1 , C_3 , C_4 , C_5 , W_1 , W_2 , W_3 は $[0, 4]$ の範囲で調整され、 C_2 に関しては $[0, 80]$ の範囲で調整されるものとする。 C_2 以外に関してパラメータの変化しうる範囲が $[0, 2]$ から $[0, 4]$ となっているのは、今回の実験が、問合せ画像とスケールや回転に差異がある画像を検索結果の上位に位置するようにパラメータの最適化を行う実験であり、回転やスケールに関連のあるパラメータが従来よりも大きく変化する可能性があるため、2倍の範囲とした。PSOの最大世代数は100とし、粒子の初期値は実験1で得た最適化後の値を用い、弛緩整合法の反復回数は5回とする。また、PSO中の λ_1 と λ_2 は各世代で生成される $[0, 2]$ 区間の一様乱数とし、粒子の最大速度は0.5とする。

5.5 結果：評価基準による差異

図8に両実験のPSOの各世代における J の変化を示す。最適化の結果、回転優先の場合には、最良値13を達成できたが、スケール優先の場合には、14までにしかなかった。

表7にスケール優先で最適化を行った場合のパラメータの値を、表8に回転優先で最適化を行った場合のパラメータの値を示す。また、表9に最適化を行ったパラメータを用いて検索をした結果を示す。

表9(C)を見ると、検索結果の4番目に105%に拡大した画像があり、3番目に5度回転した画像がある。このため、最適値13を達成することができなかった。表8を見ると、 C_2 の値が1となっている。このパラメータは、対応候補のフィルタリングに用いる角度の閾値であるが、モデリング時に角度の表現が、360度を80等分した値を1としているので、実際には4.5度となる。すると、画像を5度回転させた画像は、フィルタリングの際にはじかれることはなく、その後の弛緩法においても、特徴量が似ているため、区

図8 PSOの世代における検索評価値 J の変化。問合わせ画像として、表6のグループ5、セット1の画像を用い、スケール優先の場合の変化と回転優先の場合の変化をプロット。

表7 スケール優先で最適化を行った場合の最適化前後のパラメータの値。

Parameter	C_1	C_2	C_3	C_4	C_5	W_1	W_2	W_3
Initial value (empirical)	0.1	15	0.1	0.5	0.5	0.1	0.05	0.01
Optimized (100 generation)	2.138	1	0.275	1.102	2.046	0.0006	0.132	0.694

表8 回転優先で最適化を行った場合の最適化前後のパラメータの値。

Parameter	C_1	C_2	C_3	C_4	C_5	W_1	W_2	W_3
Initial value (empirical)	0.1	15	0.1	0.5	0.5	0.1	0.05	0.01
Optimized (100 generation)	0.590	3	2.754	1.950	2.918	0.230	0.464	0.515

別が難しいものと思われる。このため、最良値13を達成することができなかったと考えられる。

しかし、表9(B)(C)を見ると、最適化前に比べ、105%の画像が、上位に来るようになっており、改善が見られる。

また、表7、8を見ると、 C_1 の値はスケール優先

の場合に値が大きくなっており、 C_2 に関しては、逆に、回転優先の方が値が大きくなっている。これから、フィルタリングの際に、それぞれ優先度に関連のあるパラメータの値を大きくし、柔軟性を確保するようになっていることがわかった。

表10に最適化後のパラメータを用いて、最適化に用

表9 最適化で用いた問い合わせ画像(A)との類似度による検索結果。(B)に示される画像はパラメータの最適化を行う前の検索結果(スケール優先 $J=25$ 、回転優先 $J=20$)、(C)最適化後の検索結果($J=14$)、(D)最適化後の検索結果($J=13$)。右下Oは元画像、R1は5度回転、R2は-5度回転、S1は95%に縮小、S2は105%に拡大させた画像。

(B) Before optimization										
Rank	1	2	3	4	5	6	7	8	9	10
Image	O	R1	S1	R2	S2	R1	S1	R2	S2	R1
Distance	0.0	0.817	0.888	0.981	0.9901	0.9906	1.071	1.073	1.075	1.081

(C) After optimization (scale)										
Rank	1	2	3	4	5	6	7	8	9	10
Image	O	S1	R1	S2	R2	S1	R1	S2	R2	S1
Distance	0.0	0.522	0.638	0.850	0.856	0.916	0.917	0.924	0.938	0.940

(D) After optimization (rotation)										
Rank	1	2	3	4	5	6	7	8	9	10
Image	O	R1	R2	S1	S2	R1	S1	R2	S2	R1
Distance	0.0	1.079	1.216	1.243	1.409	1.438	1.489	1.494	1.507	1.526

表10 最適化で用いなかった表6に含まれる問い合わせ画像(A)との類似度による検索結果。(B)に示される画像はパラメータの最適化を行う前の検索結果(スケール優先 $J=27$ 、回転優先 $J=21$)、(C)最適化後の検索結果($J=26$)、(D)最適化後の検索結果($J=25$)。右下Oは元画像、R1は5度回転、R2は-5度回転、S1は95%に縮小、S2は105%に拡大させた画像。

(B) Before optimization										
Rank	1	2	3	4	5	6	7	8	9	10
Image	O	R2	R1	S1	S2	R1	S1	R2	S2	R1
Distance	0.0	0.674	0.679	0.774	0.841	0.881	0.953	1.0177	1.0179	1.189

(C) After optimization (scale)										
Rank	1	2	3	4	5	6	7	8	9	10
Image	O	S1	R1	S2	R2	S1	R1	S2	R2	S1
Distance	0.0	0.257	0.358	0.425	0.487	0.649	0.667	0.752	0.872	0.913

(D) After optimization (rotation)										
Rank	1	2	3	4	5	6	7	8	9	10
Image	O	R2	R1	S1	S2	R1	S1	R2	S2	R1
Distance	0.0	0.870	0.969	1.012	1.237	1.327	1.344	1.440	1.546	1.617

いなかった表6にある画像を問い合わせ画像としたときの検索結果を示す。スケール優先の場合には、評価値は27から26と改善しており、拡大・縮小された画像が最適化前に比べ、上位に現れている。回転優先の場合には、拡大・縮小された画像が最適化前に比べ下位に位置しており、評価値は悪くなっているが、元の画像に似ている画像の回転している画像が現れるようになっており、回転優先検索が達成されている。

6. まとめ

本論文ではPSOを用いた画像の類似度評価アルゴリズムのパラメータ最適化の枠組みを提案した。また、輪郭線分に着目した画像の内容検索アルゴリズムを用いて実験を行い、アルゴリズムの内部のパラメータが最適化され、画像の類似度評価のランキングが改善されることを確認した。さらに、本手法において、ユーザの類似度基準に合わせたアルゴリズムの調整が可能であることを確認した。

本論文における枠組みは、画像の内容検索だけではなく、画像の収集システム[9]に関しても、応用可能であると思われる。そのような場合には、例えば各特徴量に対して重みパラメータを導入し、各キーワードに対して、それらの重みパラメータを最適化することなどが考えられる。

参 考 文 献

- [1] T. Kato, T. Kurita, and H. Shimogaki. Intelligent visual interaction with image database systems-toward the multimedia personal interface. *Journal of Information Processing*, Vol. 41, No. 2, pp. 134-143, 1991.
- [2] 山本和彦. 弛緩整合法による手書き教育漢字認識. 電子情報通信学会論文誌, Vol. J65-D, No. 9, pp. 1167-1174, 1982.
- [3] P. W. H. Kwan, K. Kameyama, and K. Toraichi. Trade-mark retrieval by relaxation matching on fluency function approximated image contours. In *Proc. of IEEE Pacific Rim Conference on Communication, Computer and Signal Processing*, pp. 255-258, 2001.
- [4] K. Kameyama, S. Kim, K. Toraichi, and T. Yamamoto. Content-based image retrieval of kaou images by region characterization and probabilistic relaxation. In *Proc. of Modeling Decisions in Artificial Intelligence 2005*, 2005. paper no. 109.
- [5] T. Tachikawa, K. Kameyama, T. Kawabe, K. Toraichi, N. Otsu, T. Kurita, and P. W. H. Kwan. Content-based retrieval of kamon images by image smoothing and relaxation. In *Proc. of Modelling, Identification and Control 2005*, pp. 422-427, 2005.
- [6] D. McG. Squire, W. Muller, H. Muller, and T. Pun.

Content-based query of image databases : inspirations from text retrieval. In *Pattern Recognition Letters*, Vol. 21, pp. 1193-1198, 2000.

- [7] H. Muller, W. Muller, D. McG. Squire, S. M. Maillet, and T. Pun. Performance evaluation in content-based image retrieval : overview and proposals. In *Pattern Recognition Letters*, Vol. 22, pp. 593-601, 2001.
- [8] Y. Rui, T. S. Huang, and S. Mehrotra. Content-based image retrieval with relevance feedback in MARS. In *Proc. of IEEE Int. Conf. on Image Processing*, pp. 815-818, 1997.
- [9] 柳井啓司. 確率的web画像収集. 人工知能学会論文誌, Vol. 22, No. 1, pp. 10-18, 2007.
- [10] F. Kawazoe, K. Toraichi, K. Nakamura, and P. W. H. Kwan. Fluency function approximation of thin-line images. *Journal of Insititute of Image Electronics Engineers of Japan*, Vol. 32, No. 4, pp. 438-445, 2003.
- [11] A. Rosenfeld, R. A. Hummel, and S. W. Zucker. Scene labeling by relaxation operations. *IEEE Trans. Sys., Man and Cybern.*, Vol. SMC-6, No. 6, pp. 420-433, 1976.
- [12] M. Pelillo. On the dynamics of relaxation labeling processes. *Proc. IEEE Int'l. Conf. on Neural Networks*, Vol. 2, pp. 1006-1011, 1994.
- [13] R. A. Hummel and S. W. Zucker. On the foundations of relaxation labeling processes. *IEEE Trans. of Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 3, pp. 267-287, 1983.
- [14] J. Kennedy and R. C. Eberhart. Particle swarm optimization. In *Proc. of IEEE Int. Conf. on Neural Networks*, pp. 1942-1948, 1995.
- [15] J. Nenortaite and R. Simutis. Adapting particle swarm optimization to stock markets. In *Proc. 5th Int. Conf. on Intelligent Systems Design and Applications*, pp. 520-525, 2005.
- [16] H. Yoshida, K. Kawata, Y. Fukuyama, and Y. Nakanishi. A particle swarm optimization for reactive power and voltage control considering voltage stability. In *Proc. IEEE Int. Conf. on Intelligent System Applications to Power Systems*, 1999.
- [17] K. Kameyama, N. Oka, and K. Toraichi. Optimal parameter selection in image similarity evaluation algorithms using particle swarm optimization. In *Proc. 2006 IEEE Congress on Evolutionary Computation*, pp. 3824-3831, 2006.

(2007年2月15日 受付)

(2007年5月5日 採録)

[問い合わせ先]

〒305-8573 茨城県つくば市天王台1-1-1

筑波大学 大学院システム情報工学研究科 リスク工学専攻
岡 望実

TEL : 029-853-5052

FAX : 029-853-5052

E-mail : oka@wslab.risk.tsukuba.ac.jp

著者紹介

おか のぞみ
岡 望実 [非会員]

2006年3月筑波大学第三学群情報学類卒，同年4月同大学大学院システム情報工学研究科リスク工学専攻博士前期課程入学，現在に至る。画像検索技術の研究に従事。

かめやま けいすけ
亀山 啓輔 [非会員]

1989年東京工業大学工学部電気電子卒，1991年同学大学院総理工精密機械システム修士了，1992年同学同大学院助手，2000年筑波大学先端学際領域研究センター講師，2007年同学大学院システム情報工学研究科准教授。適応的な信号・画像処理と検索，パターン認識，ニューラルネットワークの研究に従事。視聴覚情報研究会（AVIRG）幹事，電子情報通信学会，IEEE会員。博士（工学）。

あらい かずお
寅市 和男 [非会員]

1971年3月，北海道大学大学院博士課程修了。同年4月，電子技術総合研究所飯島特別研究室入所。1980年，筑波大学電子情報工学系助教授，1989年，同大教授，1997年より，同大先端学際領域研究センター，マルチメディア情報研究アспект教授を経て，2007年4月より筑波大学先端学際領域研究センター特任教授兼筑波大学名誉教授として現在に至る。工学博士。主としてアナログ・デジタル信号の相互変換理論を中心としたフル・エンシ情報理論の研究と，そのマルチメディアシステムへの応用による実証研究に従事。CRESTプロジェクトリーダ等。平成16年度船井情報科学振興賞，第30回井上春成賞を受賞。IEEE，電気学会，電子情報通信学会，他の会員。

Optimization of Similarity Evaluation Algorithm Using Relevance Feedback in Content - Based Image Retrieval

by

Nozomi OKA, Keisuke KAMEYAMA and Kazuo TORAICHI

Abstract :

Recently, Content-based Image Retrieval (CBIR) which evaluates the relevance of images derived by image features is being widely investigated, aiming at retrieval from large database of unlabeled images. Most commonly in CBIR, "similarity", which generally does not have common or objective definition is used. Naturally, "similarity" can be defined in various ways and it differs according to the user's subjectivity. Because of this ambiguity of similarity, realization of retrieval suited to the user's similarity criterion is needed.

This work proposes a framework for improving the similarity evaluation of images according to the user's demand, by optimizing the parameters in the relevance evaluation algorithm according to the criteria defined by the user. In this framework, we define a criteria which is a function of the retrieved result, and its parameters are optimized by using Particle Swarm Optimization (PSO).

As a system to evaluate the effectiveness of the proposed framework, we used a CBIR system of binary images which is based on the matching of image contour features. The evaluation function J based on the user's rating of the retrieval result was defined and the parameters in the relevance evaluation algorithm were optimized. In the first experiment, the parameters were optimized in the proposed framework for a given evaluation function J using PSO. There, it was found that retrieval sets more suited according to the evaluation function was obtained after the optimization. In the second experiment, two different criteria were defined. When the system was optimized for each different criterion, it was found that the final parameters were different for different requirements, and that retrieved results were also different, each adapting to the specific requirement. As a whole, it was found that CBIR systems suited to the user's subjective similarities can be obtained by using the proposed framework, defining criteria J and optimizing the parameters accordingly.

Keywords : Content-Based Image Retrieval (CBIR), Similarity, Particle Swarm Optimization (PSO), Relaxation Matching, Relevance Feedback

Contact Address : **Nozomi OKA**

*Dept. of Risk Engineering, Graduate School of Systems and Information Engineering, University of Tsukuba
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, JAPAN*

TEL : 029-853-5052

FAX : 029-853-5052

E-mail : oka@wslab.risk.tsukuba.ac.jp