模擬育種法を用いた顔の同定支援システムに関する研究

Research of the Support System to Identify Face Using Simulated Breeding

平山 高嗣

Takatsugu Hirayama

大阪大学大学院基礎工学研究科システム人間系

Graduate School of Engineering, Osaka University.

hirayama@yachi-lab.sys.es.osaka-u.ac.jp

上田 芳弘

Yoshihiro Ueda

石川県工業試験場

Industrial Research Institute of Ishikawa.

ueda@iriii.go.jp

木村 春彦

Haruhiko Kimura

金沢大学工学部情報システム工学科

Faculty of Engineering, Kanazawa University.

kimura@ec.t.kanazawa-u.ac.jp

阿部 武彦

Takehiko Abe

金沢工業大学人間・情報・経営系

Kanazawa Institute of Technology.

abet@neptune.kanazawa-it.ac.jp

keywords: facial retrieval, subjectivity, simulated breeding, impressive adjective

Summary

In the facial retrieval, it is necessary to cope with the change of expression and consider the personal subjectivity. Generally, the matching by using linguistic keywords or image has been used in the conventional image retrieval. Therefore, this method isn't suitable for the facial image retrieval. Recently, a method treating the user's subjectivity is researched. In the method, the user evaluates facial images presented by the system, and the system retrieves new candidacies according to the evaluation. One of such method is the simulated breeding. However, this method gives heavier work to the user than the conventional method does. In this paper, we propose the facial image retrieval system united the retrieval by using linguistic keywords with the simulated breeding in order to decrease user's work. The proposed system performs the retrieval by using the impressive adjective first, and the retrieval by using the simulated breeding. We evaluate the system through the experiment using the subjects and comparing with the prior art.

1. ま え が き

犯罪を犯した容疑者の捜索に、顔写真のリストから目撃者が見た顔（ターゲット）を探し出す過程がある。この過程は、目撃者自身がターゲットの再認を行う場合や、目撃者から証言を得て、ターゲットのモンタージュを作成し、第3者が顔写真のリストからそれと似ているものを検索する場合など、その状況に応じたものとなる。しかし、いずれにしても、リストの顔写真が膨大となる場合に、目撃者や検索者の負担が大きくなり、また、多くのディストラクターを見ることにより再認成績が低下することが知られている [Laughery 72]。こうしたことから、顔画像に対して人間と同様な類似判定が行え、再認成績が向上するシステムの必要性が高まっている。

従来、画像検索手法として用いられてきたものは、言語的な検索キーを用いる方法と画像を直接検索キーとして用いる方法に分類できる。前者は一般に、画像に対し

てラベル付けを行い、検索者が入力した言語キーとのマッチングによって画像検索を行う方法 [イングベンゼン 95] であるが、正確に画像の特徴を言語で表現しきれない場合や、ラベル付けを行う設計者とシステムを利用する検索者の感性に食い違いがある場合、また、検索者が適当な言語キーを思いつかない場合がある等の問題点がある。これらの理由から、画像から受ける印象といった個人の主観性を重視する必要がある顔画像の検索に適用するには限界が生じてくる。一方、後者は主に、画像から抽出した特徴量を用いたパターンマッチングによって検索を行う方法 [坂内 88] であるが、顔の表情のように刻一刻と変化する画像に対して必ずしも有効ではない。例えば、犯行現場を撮り押さえた写真に怒りを露わにした犯人の顔が映っていた場合、顔写真リストの顔は無表情であることが多いため、パターンマッチングによる検索は適用できない。これに対して、顔の動きをアクションユニット (Action Unit: AU) と呼ばれる 44 通りの基本動作に分解し、そ

れらを組み合わせて人間の表情を合成する FACS(Facial Action Coding System) と画像工学のモーフィング(形態変化) 技術により、怒りの形相の犯人を無表情に変えることでパターンマッチングを行うことは可能であるが、現段階では正しい犯人の無表情な顔を作り出す性能が悪いため、信頼性は低い。しかし、人間はと言うと、初対面の人の顔の表情から難なくその人が持つ感情を認識することができる。つまりは、無表情な顔からいくつかの感情に対する顔を予測しているものと思われる。このように、顔の表情の変化に対する人間の認識力はすばらしく、他に勝るものはない。

そこで、システムが示す顔画像が検索者の要求に合っているか否かを検索者自身に評価させることで、検索者の主観性を重視した検索を実現させる方法が研究されている[尾田 94, 長尾 98, 杉本 98, Caldwell 91]。この方法は、従来の画像検索手法の分類から言うと、画像を直接検索キーとして用いる方法に当るが、検索者自身が検索の過程において主観的な特徴を定義することが可能であるため、顔の表情の変化にも柔軟に対応できる。例えば、[長尾 98] では、インタラクティブ GA を用いて対話的に画像検索を行うシステムが考案されている。ここで、インタラクティブ GA とは、従来の遺伝的アルゴリズムの自然選択を人為選択に代えて次世代の親を決定する操作を繰り返し、人間の好みや主観を反映した個体を発生させようとする手法である。

しかし、こういった方法は従来の手法に比べると、検索者の負担が非常に大きくなる。そのため、いかにこの負担を軽減させて、検索者の要求に適合したものを検索するかが課題であった。また、この方法を用いたこれまでの研究では、検索の対象を曖昧な顔画像、つまり、ターゲットの明確な提示を行わないものが主であった。ターゲットのイメージが明確である場合は、そうでない場合に比べ、その印象を言語で表しやすくなる。つまり、言語キーによる検索が有効な場合である。

そこで、本論文では、インタラクティブ GA の一種である模擬育種法と言語キーによる検索の技術を融合させた顔画像検索システムを提案する。具体的には、言語キーによって、初期集団の検索を行い、その後は検索手法を模擬育種法に切り替え、検索者が提示された顔画像に対して要求に合っているか否かを評価することで対話的に検索の絞り込みを早めるものである。模擬育種法を利用する前に言語キーによって検索することで、ある特徴を共通に持つ顔画像を初期集団として検索することができ、初期集団をランダムに検索していた従来のインタラクティブ GA や模擬育種法だけを用いる手法に比べて、検索回数の削減が期待できる。

2. 模 擬 育 種 法

本章では、本論文で提案する手法の中心概念となる模擬育種法について概説する。

2.1 概 要

模擬育種法は評価者である人間とそれを支援するコンピュータとの相互作用を通して最適化を進めるための方法論[畝見 94]であり、自明ではない主観的な評価基準を持つ問題にも適用可能な点に特徴がある。すなわち、個々の解候補(インスタンス)の良さが評価される方程式を明示できない場合や、その評価にはある種の「認識」処理が必要となる場合には、このような評価をコンピュータに自動的に実行させるという試みは、ほぼ不可能か、たとえ可能であっても膨大な計算コストが要求されることが多い。ただ、こういった困難を引き起こす問題には人間の主観的な要素が絡む場合が多いため、個々の解候補の良さを人間自身が評価すれば、人間とコンピュータとの相互作用を通じた最適化が可能となるはずである。このようなアイデアと遺伝的アルゴリズム(Genetic Algorithm:GA)といった進化的計算論の枠組とを有機的に組み合わせた方法論が模擬育種法である[北本 98]。言い換えれば、模擬育種法は、自然界における生物進化の原理に着想を得た最適化技法である GA の応用的手法となる。また、模擬育種法は、GA における自然選択を人為選択に代えることにより次世代の親を決定する操作を繰り返すことで、人間の主観に反映した個体を発生させる手法といえる。この意味で、模擬育種法は、自明でない評価基準を持つ人間の主観的要素が絡むような問題に対して極めて有効である。

2.2 問題点と対処方法

ここまで、模擬育種法についてその特徴を含めて述べてきたが、もちろん問題点も存在する[北本 98]。それは、人間とシステムのインタラクションを考える場合に重要となる『待ち時間』という要素についてであり、システムが人間を待たせる時間と、人間がシステムを待たせる時間が挙げられる。前者は、遺伝子型から表現型への変換にかかる時間を指し、待ち時間があることで、評価者が操作性を損なう可能性が考えられる。ここで、遺伝子型とは、人間とシステムのインタラクションに直接用いられる媒体をシステム内で遺伝的アルゴリズム処理を行うために数値データに変換したものであり、3章に示す染色体と同意である。また、表現型は、この媒体そのものを指し、提案システムでは顔画像に当たる。一方、後者は評価者が個体の比較及び選択にかかる時間を指し、システムの効率性の問題に関連する。

これらの待ち時間を解消する方法として「待ち行列型遺伝的アルゴリズム(queue-based genetic algorithm:QGA)」[北本 98]を活用した模擬育種法が提案されている。しか

し、本研究では、遺伝子型から表現型である顔画像を生成する処理に描画システムを適用せず、あらかじめ顔の各部位の形状をパターン化し、それらを組み合わせる処理で遺伝子型から表現型への変換を実現しているため、大きな計算コストを要することはない。また、システムの効率性よりも人間にかかる負担を軽減させることに重きを置くこととし、これらの理由から、本研究ではQGAを導入しなかった。

さて、人間の負担を軽減させるという点では、多くの画像に評価を付与するという繰り返し作業に対する人間の忍耐力に絡む問題もある。これは、まず個体の選択時の迷いに影響されるが、1つの世代でシステムが検索者に提示する個体数を小さく設定し、生成される個体に多様性を持たせるような遺伝的操作を適用することで検索者の迷いを軽減できる。また、検索者の負担は検索開始からの検索枚数の多さと時間の長さに影響される。[Laughery 72]によれば、一つ一つの顔画像に対する吟味にかかる時間よりも多くの顔画像を見ることが検索者の負担を大きくし、正しい認識を妨げるということである。これより、本研究では検索枚数を重要視することとし、初期集団の検索を言語キーを用いて行うことで準最適なものに収束させることができれば、検索枚数を減らすことが可能と考える。初期集団として提示される画像がターゲットに類似していれば、その後の検索での検索者の負担を軽減させられることは[北本 98]により主張されている。ただし、検索者の負荷を軽減することと人間の主観を多く反映させることは、ある程度トレードオフの関係にある[杉本 98]ため、システムに用いるパラメータの設定は慎重に行う必要がある。

3. 顔の同定支援システム

本システムは犯罪捜査の過程での容疑者の同定や絞り込みを支援することを目的としたものである。この過程は目撃者の証言が信頼できるものであるかどうかなど、その状況によって取り組む課題が変わるが、本研究では主に目撃者の証言から容疑者のモンタージュを作成し、顔写真リストからそれと似ているものを検索するケースを対象とする。以下に、本システムの詳細について述べる。

3.1 顔 画 像

本研究ではシステム設計及び評価実験の簡単化のため、顔写真は[李 97]のような手法により、あらかじめ線画の顔画像に変換されているものとし、線画による顔画像を用いる。顔の構成は各要素の形状や位置をパターン化したもの^{*1}を組み合わせることで行う。つまり、顔画像は顔の形状、眉毛の角度、目の丸さなど、表1に示される計22個のパラメータから構成される。なお、眉毛の角度と目の丸さに関するパラメータは $\{-1, -0.5, 0, 0.5, 1\}$ 、

表1 顔画像の構成に用いたパラメータ

顔	形状	四角/丸/ホームベース
	長さ	長い/普通/短い
	幅	広い/普通/狭い
眉毛	角度	垂れ/少し垂れ/普通/少し吊り/吊り
	太さ	太い/普通/細い
	長さ	短い/普通/長い
	間隔	広い/普通/狭い
	位置	上/中/下
	位置	上/中/下
目	角度	垂れ/普通/吊り
	丸さ	丸い/少し丸い/四角/少し細い/細い
	大きさ	大きい/普通/小さい
	間隔	広い/普通/狭い
	位置	上/中/下
耳	耳たぶの形状	大きい/普通/小さい
	大きさ	大きい/普通/小さい
	位置	上/中/下
鼻	高さ	低い/普通/高い
	大きさ	大きい/普通/小さい
	位置	上/中/下
口	大きさ	大きい/普通/小さい
	唇の厚さ	厚い/普通/薄い
	位置	上/中/下

図1 本システムで用いた顔画像のサンプル

その他のパラメータは $\{-1, 0, 1\}$ のある値を取るように設定し、眉毛と目のパターン数は他に比べて多くなっている。これは、人間が顔を認識する際、下部にある要素(口、顎など)より上部にある要素(頭髪、目など)を重要視していることが知られ[乾 95]、これらの要素に対してより細かな認識を行っていると考えたためである。ただし、本研究では重要視すべき要素である頭髪については扱わない。なぜなら、実際の利用形態を考えると、髪型は日常生活で容易に変化させることができ、犯行時と顔写真リストの容疑者の髪型が一致する可能性は低く、現在の画像処理技術によれば簡単に除去できる頭髪は含めずに検索の方が妥当であると考えたためである。本システムでは設計の簡単化のため、これらのパラメータの取り得る値の全ての組合せを可能とし、約870億枚の顔画像が検索対象としてシステム上に存在すると想定できる。この数は実用段階のシステムで非現実的な値であるように思えるが、一人の人間の顔が表情や時間経過により無数に変化することを考慮すると決して非現実的なものではないとも考えられる。また、システム内部では、22個のパラメータから構成される顔画像Fを以下のようにベクトルとみなし、22次元空間の1つの点に対応づける。

$$x_f = (x_{f1}, x_{f2}, \dots, x_{fi}, \dots, x_{f22}) \quad (1)$$

*1 各パターンの描画は著者が行った。

ここで、式 (1) が染色体、右辺の各要素が遺伝子にあたる。図 1 に顔画像の例を示す。左から、遺伝子の値が全て -1 の顔画像、各遺伝子の取り得る値の中央値とした顔画像、全て 1 の顔画像である。

3.2 印象語

本システムでは初期世代の検索に用いる言語キーに、顔の全体像に対する印象を表す言葉 (印象語) と顔の要素ごとの特徴を表す具体的な言葉 (特徴語) を適用する。検索者には、まず、ターゲットに対する全体的な印象を、次に顔の要素ごとの特徴を細かに入力させることになる。これは、イメージがデザインとして表現されるまでにはいくつかのプロセスを経る [市野 99] ことと同様に、目前にある顔画像から受ける感性にも、例えば「おとなしそう」といった抽象度の高いものから「眉毛が垂れている」といった抽象度の低いものへと移るプロセスがあると考えたからである。

さて、言語キーによる検索は、入力する言語を検索者が自由に表現するかあらかじめシステム側が用意したものから選択するかのいずれかになるが、本研究では後者を用いることとした。そこで、印象語は [日本ファジィ学会 93] と [岩下 98] から対義関係にある次の 7 組 14 種

おとなしそう-活発そう
 きゃしゃ-ふくよか
 無邪気な-理知的な
 大人びた-子供っぽい
 やさしそう-冷たそう
 男っぽい-女っぽい
 強そう-弱そう

を抜粋し、特徴語は顔画像の描画に用いたパラメータと対応づけて用意した。また、言語キーによる検索は模擬育種法による検索に入る前にある程度、顔の特徴を絞り込むことにその目的があるが、選ばれた特徴語によっては検索される顔が誤った方向に収束してしまう恐れがある。これは選べる候補が増えるほど、可能性が高くなると思われるので、検索者には必要最小限の選択肢を与えることとし、検索者へ提示する特徴語は選ばれた印象語に応じて一意に設定する。つまり、各印象語に対して模範的な特徴を定める。本研究では、以下のような印象調査アンケートを実施し、その分析結果からこれを定めた。

印象調査アンケートは、20 歳代男女計 30 名に対して行った。被験者は本システム上からランダムに選択された 100 枚の顔画像を一括で提示され、その中から 14 種の印象語それぞれにあてはまるものを選択した。なお、提示画像は被験者間で同じものとし、また、一つの印象語に対して複数の回答と印象語間での回答の重複を可能とした。得られたデータの分析手法は、印象語ごとに、回答から各要素のパラメータ値 (つまり遺伝子) の頻度を調べ、これに対し有意水準 1 % でカイ 2 乗検定を行った。正確には、カイ 2 乗値が高く、他と有意差があるパラメー

図 2 顔の同定支援システムのユーザインタフェース

タ値を印象語に対する顔の特徴とし、これから特徴語を生成した。なお、要素間の交互関係については無視した。この分析結果は、例えば「おとなしそう」という印象語に対して、輪郭が四角い、眉毛が垂れていて短い、眉間が狭い、目が垂れている、鼻が小さい、口が小さいと得られている。

3.3 検索方法

(1) 検索手順

検索は以下のステップで行われる。

- (a) 検索者にターゲットに対する印象を、用意した印象語の中から 1 つ選択させる。
- (b) 選ばれた印象語に対して設定した特徴語を検索者に提示し、その中からターゲットの特徴を表すものを複数個選択させる。
- (c) (b) で選ばれた特徴を重視し、印象語の特徴を考慮して顔画像を 10 個検索し、検索者に提示する。
- (d) ターゲットに似たものを複数枚選択させる。選ばれた顔画像を親とする。
- (e) 親の染色体に対して遺伝的操作を適用し、新たに 10 個の染色体を生成する。これらの子孫と呼ぶ。
- (f) それぞれの子孫の表現型 (顔画像) を検索者に提示する。
- (g) 終了条件を満たせば終了。そうでなければ、(d) へ戻り、次の世代に移る (終了条件については後述する)。

(a) ~ (c) が印象語検索、いわゆる言語キーによる検索、(d) ~ (g) が模擬育種法にあたる。この一連の作業効率を考慮して、図 2 のような GUI を用意する。

検索を進めるにあたって、場合によっては検索者の主観に適合しない顔画像のみが提示され、選択ができないことがある。この状況には、初期世代においては再度印象語による検索を行い、第二世代以降では、先代の子孫を生成し直すことで対処する。

本システムは、検索者によって選択された顔画像を管理するためのバッファ (図 3) を備え、選択された顔画像

図 3 バッファ用ユーザインタフェース

は各世代を通じて 10 枚まで蓄えられる。また、バッファに 10 枚の顔画像が蓄えられている状態で新たに顔画像が選択された場合には、検索者にそれとバッファ内の顔画像を再評価させ、バッファがよりターゲットと類似度の高い顔画像で満たされるように入れ換え作業を行わせる。これにより、終了条件を満たした時点でターゲットに似た顔画像が 10 枚バッファに蓄えられることになる。

(2) 印象語検索

印象語検索では 3・2 節で述べた印象語と特徴語を用いて初期集団となる個体を検索する。

本システムのデータベースには、印象調査アンケートで得られたデータの分析結果から、顔の各要素のパラメータ値と印象語それぞれに対して、そのパラメータ値が表す特徴の適合度が記録されている。例えば「おとなしそう」という印象語に対して、輪郭が四角いことを表すパラメータ値 (-1) は 55 の適合度を持ち、輪郭が丸いことを表すパラメータ値 (0) は 10 の適合度を持つ、というようになっている。適合度が高いほど、その印象語とパラメータ値が表す特徴が適切なものであることを意味する。

検索の方法は、まず、検索者にターゲットの印象を、用意した印象語の中から 1 つ選択させる。次に、選ばれた印象語に対して設定した特徴語を検索者に提示し (図 4)、その中からターゲットの特徴を表すものを複数個選択させる。この時、選択された特徴には、それに対応するパラメータ値を持つ適合度を 1.5 倍し、選択されなかったものには、その適合度を 0.5 倍するという処理がなされる。これにより、データベースに記録されている適合度が検索者の主観に合ったものに变换される。そして、システムはこのデータベースに記された適合度の高い特徴を持つ顔画像を順に検索して行く。

(3) 遺伝的操作

本システムでは以下の 2 種類の遺伝的操作を行うことができる。

(a) 交叉

2 個体の親の特徴を組み合わせ、2 個体の子孫を生成する。交叉には一様交叉を用いる。一様交叉とは、2 個体の親それぞれの染色体上で任意個の交叉点をとれるような交叉法であり、一般にはランダムに選ばれた 0 と 1 からなるビット列のマスクを用いて実現する。本システムではマスクの生成をランダムな値のビット列ではなく、選択された印象語を考慮して行う。つまり、各印象語に対して設定した特徴に関わるパラメータの並びを崩さずに交叉が行われるようなマスクを生成する。これは、生成される個体が検索者の主観を反映したものとなるようにするためである。

(b) 突然変異

染色体上の任意の遺伝子座の遺伝子において、その値をランダムに変更する。これは、交叉だけでは生成できない子孫を生成して、個体群の多様性を維持する働きをする。本システムでは、遺伝子一つを 1 bit として、20%/bit の確率で値を変更する。ただし、眉毛の角度と目の丸さに関するパラメータは、その取り得る値の範囲が他に比べて大きいので、40%/bit としている。

ここで定義した各遺伝的操作の発生確率は、各世代で検索者によって選択される顔画像の数によって変化させる。例えば、被験者によって選択された顔画像が一つの場合、交叉を行うことができないため、交叉率を 0%、突然変異率を 100% とし、選択された顔画像の数が多くなるほど、検索者の主観を反映しやすくするため、交叉率を高くし、突然変異率を低くする。これは交叉率を一定にすると、親の数が増えるほど、各親が交叉に用いられる確率が下がることを懸念したためでもある。

4. 検 索 実 験

提案システムの有効性を確認するため、被験者を用いた評価実験を行った。ここでは、評価実験の内容と結果について示す。

図 4 印象語検索用インタフェース

4.1 実験方法

実験は、大学生 7 名に対して、本手法の有効性を検証するために 3 種類のシステムを用いて行った。3 種類のシステムとは、本論文で提案した顔の同定支援システム (PS)，人為選択のみによって次世代の親が決定される従来の模擬育種法を適用したシステム (SB)，顔画像データベースからランダムに検索を行い、検索者に提示するシステム (RN) である。SB の概要は本システムの模擬育種法による検索にあたる部分と同様であるが、遺伝的操作の一樣交叉に用いるマスクを生成する部分で異なる。本システムはこのマスクの生成に検索者によって選択された印象語を考慮するが、SB は 3.3 節 (1) の (d) から検索が開始されるため、印象語を選択する過程がなく、マスクはランダムなビット列によって生成される。

各被験者はまず、実験の目的とシステムの操作方法に関する説明を受け、操作の一例を観察した後、実験を開始する。実験の目的は「ターゲットによく類似した顔画像を検索していき、最終的にターゲットであると思われる顔を 10 枚検索する」というものである。それぞれのシステムを使用する順序は、実験結果にその順序による影響が出る可能性が十分に考えられるため、3 種類のシステムの組合せによる 6 通りの実験順序を各被験者に割り当てた。なお、7 名の被験者のうち 1 名は他の被験者の一人と同じ順序で行った。検索対象となるターゲット顔画像には評価実験の分析の簡単化のため、データベースに存在する図 5 を用いた。このターゲットは印象調査アンケートで被験者に提示した 100 枚の顔画像に含まれ、アンケート終了後に多くの被験者から口頭で「最も犯罪者っぽい顔」と報告を受けた顔画像である。なお、ターゲットと同じ顔は約 870 億分の 1 の割合でデータベースに含まれている。

ここで、終了条件について述べる。顔画像から受ける印象といった主観性を扱う検索においては、複数の顔画像が検索者にとってターゲットとよく似ている、もしくはターゲットであると思われる場合がある。また、そういった複数の顔画像の中から最終的な判断を行った方がより良い判断を行うことができるという報告 [長尾 98] がある。このような理由により、検索者がターゲットであると思う画像を 10 枚検索した時点で検索を終了するものとする。

表 2 に、模擬育種法の遺伝的操作に適用したパラメー

図 5 検索実験に用いたターゲット顔画像

表 2 交叉率と突然変異率

親数	交叉率	突然変異率
1	0	1
2	0.5	0.5
3	0.65	0.35
4	0.75	0.25
5 ~	0.8	0.2

タを示す。3.3 節で述べた理由から、親の数によって各遺伝的操作の発生確率を変動させている。また、これらのパラメータは提示する顔画像の多様性を維持し、誤った方向への収束を防ぐことができるように設定した。

4.2 実験結果

提案システムの有効性を検索の効率、実験目的の達成度、そして本研究で最も注目している初期世代に検索される個体の収束性という点から評価する。ここで、検索の効率は検索回数と検索時間であるが、検索回数と検索時間のどちらを重要視するかについては、2.2 節で述べたように検索回数を重要視する。また、実験目的の達成度は検索結果の精度の良さを表す。

まず、表 3、表 4、図 6 に実験結果を示す。表 3 における検索回数は検索開始から検索終了までに要した世代数、検索時間は RN と SB に関しては検索開始から検索終了までの時間、PS に関しては特徴語選択の終了時点から模擬育種法による検索の終了までの時間、表 4 におけるバッファ内ターゲット平均距離は検索終了時にバッファに残った 10 枚の各画像とターゲット顔画像間のユークリッド距離の平均を示す。また、図 6 は検索者に提示された顔画像とターゲット顔画像間のユークリッド距離の世代ごとの平均が変移する様子を示す。ここで、顔画像とターゲット画像間のユークリッド距離とは、顔画像とターゲット画像を表すそれぞれの染色体の各要素（遺伝子）から求まるユークリッド距離のことで、次式で求まる。

$$distance = \sqrt{\sum_{i=1}^{22} \{x_i - t_i\}^2} \quad (2)$$

x_i : 顔画像の遺伝子

t_i : ターゲット画像の遺伝子

顔画像間のユークリッド距離の近さがその類似性に比例する [ブルース 90] ことから、図 6 は各世代の提示顔画像とターゲットの類似性の移り変わりをみることができる。なお、表 3、表 4、図 6 の数値は全て、7 人の被験者による実験の平均である。また、PS での印象語選択と特徴語選択に幾ばくかの時間と労力がかかっているのに、それらを考慮しなかった理由は、検索者がターゲットに似た顔画像の選択を繰り返すことよりも、その労力は非常に小さく、その時間は顔に対する吟味にかかる時間と言えるため、2.2 節で述べたように重要視しないか

表 3 各システムにおける検索回数と検索時間の平均

	検索回数	検索時間
PS	16.7	4 分 40 秒
SB	19.7	4 分 22 秒
RN	52.0	5 分 46 秒

表 4 各システムにおける検索終了時のバッファ内顔画像とターゲット間の平均距離

	バッファ内ターゲット平均距離
PS	4.28
SB	4.46
RN	5.20

らである．すなわち，本研究では，初期検索自体のコストよりも，それによりもたらされる効果に注目する．

表 3 から検索回数は，PS が 16.7，SB が 19.7，RN が 52.0 であり，PS が最も良い結果を示している．しかし，検索時間は PS が 4 分 40 秒，SB が 4 分 22 秒，RN が 5 分 46 秒とわずかながら PS よりも SB の方が短かった．表 3 に示す検索回数と検索時間それぞれの平均値間の差の検定を行ったところ，検索回数に関して PS と RN 間 ($t=5.51$)，SB と RN 間 ($t=5.15$) に統計的有意差が得られた (有意水準 5%， $t=2.45$) が，その他については得られなかった．

次に，表 4 から，PS は 4.28，SB は 4.46，RN は 5.20 となり，検索結果の精度の良さは PS が最も良かった．これら平均値間の差の検定を行ったところ，PS と RN 間 ($t=5.19$)，SB と RN 間 ($t=3.42$) に統計的有意差 (有意水準 5%， $t=2.45$) が得られた．PS と SB 間に有意差が得られなかったが，PS は他システムより高い達成度を示したため，SB と RN に比べ効率が良いことが実証できた．

図 6 からは，PS と SB において 15 世代目付近から提示顔画像とターゲット間の平均ユークリッド距離が 4.7 を越えないことがわかる．この 4.7 という値が本システムで設定している検索者による主観的な終了条件ではなく，客観的な終了を示すものであると考えると，SB は 11 世代目，PS は SB より 7 世代早く 4 世代目で検索を終了していることになる．つまり，表 3 に示す検索回数

図 6 提示顔画像とターゲット間の世代ごとの平均ユークリッド距離

表 5 各システムにおける選択画像数と選択画像の分散の平均

	選択画像数	選択画像の分散
PS	16.0	0.168
SB	16.9	0.172
RN	12.9	0.481

は主観的な観点から導いたものであり，客観的な観点からの検索回数は PS が 4，SB が 11 となる．また，世代ごとの PS と SB 間の差の検定を行なったところ，1 世代目 ($t=3.22$)，5 世代目 ($t=2.68$)，6 世代目 ($t=4.26$)，8 世代目 ($t=2.77$)，9 世代目 ($t=2.65$) で有意差が得られた (有意水準 5%， $t=2.45$)．これらにより PS で提示される顔画像の収束性が SB より良いことが認められた．さらに，1 世代目で有意差が得られたため，初期検索に印象語を用いることによる効果も実証された．これらのことから，検索者に蓄積する負担が比較的軽いうちに，検索者は多くの顔画像を選択し，バッファに貯めることができるといえ，これが検索効率と実験目的の達成度に良い影響を及ぼすと考えられる．

これらの結果から，SB と RN に対する提案システムの有効性が検索の効率，実験目的の達成度，そして本研究で注目している初期世代に検索される個体の収束性という点から確認された．

ここで，本研究では重要視していないが，時間短縮ができなかった理由を考察する．検索時間を検索回数で割った 1 世代あたりの所要時間を求めると，PS が約 17 秒，SB が約 13 秒，RN が約 7 秒となる．このデータから推測できることは，PS の場合には提示される顔画像の多くが検索者の主観に適合したものとなり，1 世代あたりの評価時間が長くなったということである．または，バッファに蓄積されている顔画像のほとんどが似たものとなり，さらに検索者に選択された顔画像もそれらに似ていることで，バッファの入れ換え作業に手間がかかったからとも考えられる．つまり，これらに PS の検索時間が長くなった理由があると考えられる．

そこで，1 世代あたりの所要時間が PS で長くなった原因を表 5 と図 6 から検討する．表 5 における選択画像数は検索者によって選択された顔画像の数，分散は検索者によって選択された顔画像の分散を示す．数値は全て，7 人の被験者による実験の平均値である．選択画像数は PS で 16.0 枚，SB で 16.9 枚，RN で 12.9 枚となった．また，その選択画像の分散は PS が 0.168，SB が 0.172，RN が 0.481 という値を示した．RN より PS，SB の選択画像数が多く，選択画像の分散が小さいから，提示される顔画像の多くが検索者の主観に適合したものとなった可能性が高い．図 6 からは 1，5，6，8，9 世代目で PS の平均ユークリッド距離が他より低く，有意差があるという結果が出ている．つまり，これは PS の提示画像が SB と RN のものよりターゲットに似ていることを示す．これらの結果から，PS は SB に比べターゲット

に似ている顔画像が多く提示されるにもかかわらず、選択画像数が少なく、それでいて選択画像の分散も小さいことから、検索者が顔画像の吟味に時間をかけていたことがわかる。なお、これらのデータから RN に関して、図 6 のグラフの変動が小さいわりに、選択画像の分散が大きいことから、ランダムに提示される顔画像の影響で、検索者が自己の主観を一樣に保てないということが推測できる。

5. 評価

前章では、提案システム (PS)、顔画像データベースからランダムに検索を行うシステム (RN)、人為選択のみによって次世代の親が決定される従来の模擬育種法を適用したシステム (SB) の 3 手法の比較を行った。しかし、これまではインタラクティブ GA や模擬育種法を応用した顔画像検索システムがいくつか提案され、その有効性が述べられているため、それらのシステムとの比較が必要となる。そこで、本章では [長尾 98] を比較の対象とし有効性の検討を行う。[長尾 98] を取り挙げた理由は、提案システムに最も類似しているからである。

まず、[長尾 98] の手法の概略を述べる。この手法は、従来のインタラクティブ GA や模擬育種法における人為選択によってのみ親が決定されるものではなく、検索履歴を考慮し、評価関数による自然選択を併用するというものである。また、自然選択も用いることで大規模な選択を行い、さらに検索履歴から過去に得られた情報も利用することによって、単純な評価により満足のいく検索結果を得ることが可能である。次に、この手法を用いたシステム (IGA) の実験についてであるが、顔画像は特徴数 10 の線画によるもので、システムには約 6 万枚の顔画像が存在する。比較に用いたシステムは本論文と同様に RN と SB である。また、被験者は大学生 6 名で、実験の終了条件は本論文と同様であった。表 6 に実験結果を示す。なお、ターゲット顔画像までの距離は検索者によって検索された各顔画像とターゲット顔画像間のユークリッド距離の総和である。

本論文の実験結果と比較するため、表 3、表 4、表 6 のそれぞれの提案システムによるデータをそれらに対応する RN のデータで割った値を求めた (表 7)。これらの値はより小さい方が提案システムの効率の良さを示す。なお、SB のデータで割った値を求めなかった理由は、[長尾 98] で用いられた SB が、

- 検索者による選択が行われない場合には、提示され

表 6 [長尾 98] のターゲット顔画像を用いた検索実験における実験結果

	検索回数	検索時間	ターゲット顔画像までの距離
IGA	36.3	6 分 9 秒	62.6
SB	63.3	8 分 49 秒	54.6
RN	77.1	8 分 9 秒	65.7

表 7 PS と IGA の効率比較

	PS/RN	IGA/RN
検索回数	0.321	0.471
検索時間	0.809	0.755
ターゲットまでの距離	0.823	0.953

る画像をランダムに決定

- 突然変異による遺伝子の変更確率が 10%/bit
- 遺伝的操作の交叉率と突然変異率は親の数によらず一定

など、本論文におけるものと若干異なるためである。

表 7 により、検索時間に対しては IGA が PS より良い結果となったが、本論文で重要視した検索回数に関しては、PS が IGA より効率が良いことがわかる。

実験目的の達成度の指標となるターゲット画像までの距離については、PS が IGA より良い成果を得ている。[長尾 98] では、IGA におけるターゲット画像までの距離が大きくなった理由に提示画像集合とターゲット間距離の収束の遅さを挙げている。一方、本論文では、4.2 節で本システムの収束の早さを確認している。つまり、この収束を早くすることで、検索者の負担を軽減でき、より検索者の満足度の高い検索が行えることが立証され、本研究の着目点の正しさも確認できたといえる。

以上より、先行研究における実験結果と比較することで、検索効率と実験目的の達成度に関する本手法の有効性が明らかになった。

6. むすび

本論文では、模擬育種法と言語キーによる検索の技術を組み合わせた顔の同定支援システムを提案し、被験者を用いた検索実験の結果から印象語による初期集団の検索が有効であることを確認した。また、[長尾 98] の手法との性能比較を行い、検索効率と実験目的の達成度という点で良い成果が見られ、インタラクティブ GA や模擬育種法において初期集団の個体群の生成をランダムに行わず、問題解決に適当となるようなところで収束させることが有効であることが明らかになった。

今後の課題としては、検索回数と検索時間に関して PS と SB で直接的に統計的有意差を出すことが挙げられる。現段階で PS で提示される顔画像の収束性が SB より良いため、あらゆる顔画像に対して多くの被験者を用いた実験を行えば、本手法の検索効率に関するさらなる有効性が確認できると思われる。本論文では [ブルース 90] から顔の類似度に各要素から求まるユークリッド距離を採用しているが、[杉本 98] の指摘にあるように、ある顔画像から同じ距離にある顔画像でも人間が感じる類似度は異なる。ゆえに、実験結果の評価に単なる距離を用いるのではなく、人間の主観をより反映した尺度を検討し、それを評価に用いる必要がある。さらにシステムにも組み込むことで人間が感じる類似度と一致した検索が可能と

なるであろう。

また、本システムは被験者の負担を軽減できるとしたが、それでも被験者は検索の終了までに平均して 167 枚 (= 検索回数 × 1 世代あたりの提示画像 10 枚) の顔画像に対して評価を下しているため、そのストレスが完全に無くなったわけではない。これについては、遺伝的操作に検索履歴を利用することや、本システムに用いた模倣育種法に変えて [長尾 98] で提案されたインタラクティブ GA を導入すること、適切なパラメータを設定することによって、さらに軽減させることは可能であると思われる。

最後に終了条件についての検討も必要である。本論文では検索者がターゲットであると思う画像を 10 枚検索した時点で検索を終了するものとしたが、このような検索者の主観に頼る終了条件では、システムのデータベースにターゲットと同一人物の顔画像が無い場合、検索者はデータベース全体を検索して同一人物無しとするか、ある程度ターゲットに似た顔を検索して終わることになる。また、データベースにターゲットと同一人物が含まれていても、それを検索せずに途中で検索を終えてしまうこともあり得る。本システムの実用段階での利用形態が、同一人物の表情の違いにも対応した検索を行えることを必要とし、検索された顔が類似するだけでは意味がない。このように終了条件についての検討は今後の最大の課題である。

さらに、検索者がターゲットに対して適当な印象語を選択できない場合の対処法を検討しなければならない。

◇ 参 考 文 献 ◇

- [ブルース 90] ブルース, 吉川左紀子 (訳): 顔の認知と情報処理, サイエンス社 (1990).
- [Caldwell 91] Caldwell, C. and Johnston, V.S.: Tracking a criminal suspect through "face-space" with a genetic algorithm, Proc. ICGA, pp.416-421 (1991).
- [市野 99] 市野順子, 田野俊一: デザイン描画を支援するユーザインタフェース, 電子情報通信学会論文誌 (D-II), Vol. J82-D-II, No.10, pp.1693-1709 (1999).
- [イングベンゼン 95] イングベンゼン, 藤原鎮男 監訳: 情報検索研究, トッパン (1995).
- [乾 95] 乾 敏郎 (編): 認知心理学 1 知覚と運動, 東京大学出版会 (1995).
- [岩下 98] 岩下志乃, 鬼沢武久: 顔の印象と特徴を考慮した似顔絵描写とその評価, 第 14 回ヒューマンインタフェースシンポジウム講演論文集, pp.673-678 (1998).
- [北本 98] 北本朝展, 高木幹雄: 待ち行列型遺伝的アルゴリズムを用いた対話的な画像散策法, 人工知能学会誌, Vol.13, No.5, pp.728-737 (1998).
- [Laughery 72] Laughery, K.R., Alexander, J.F. and Lane, A.B.: Recognition of human faces: Effects of target exposure time, target position, pose position, and type of photograph, Journal of Applied Psychology, 55, pp.477-483 (1972).
- [李 97] 李 元中, 小畑秀文: 顔のスケッチ画像の抽出, 電子情報通信学会論文誌 (D-II), Vol. J80-D-II, No.8, pp.2161-2169 (1997).
- [長尾 98] 長尾光悦, 山本雅人, 鈴木恵二, 大内 東: インタラクティブ GA に基づく画像検索システムの評価, 人工知能学会誌, Vol.13, No.5 pp.720-727 (1998).

- [日本ファジィ学会 93] 日本ファジィ学会 (編): 講座ファジィ 10 ファジィ画像処理, 日刊工業新聞社 (1993).
- [尾田 94] 尾田政臣: 人間のイメージ形成過程の特性を利用した画像検索システム, 情報処理学会論文誌, Vol.35, No.7, pp.1449-1456 (1994).
- [坂内 88] 坂内正夫: 画像検索技術, 電子情報通信学会誌, Vol.71, No.9, pp.911-914 (1988).
- [杉本 98] 杉本富利, 西尾研一, 本多中二: IGA のための心理的尺度に基づいたファジィ推論による適合度割当て法, 人工知能学会誌, Vol.13, No.5, pp.739-745 (1998).
- [畝見 94] 畝見達夫: 遺伝的アルゴリズムとコンピュータグラフィックスアート, 人工知能学会誌, Vol.9, No.4, pp.518-523 (1994).

〔担当委員: 外山勝彦〕

2001 年 3 月 19 日 受理

著 者 紹 介

平山 高嗣

1998 年奈良工業高等専門学校情報工学科卒業, 2000 年金沢大学工学部電気・情報工学科卒業。同年, 大阪大学大学院基礎工学研究科入学。現在に至る。感性工学, 生体情報処理に興味を持つ。

上田 芳弘 (正会員)

1983 年慶応義塾大学工学部管理工学科卒。1985 年同大学院修士課程修了。同年松下電器産業入社。1989 年石川県工業試験場入所。以降, エキスパートシステム, グループウェア, データマイニングの研究開発に従事。現在, 金沢大学大学院自然科学研究科博士後期課程在籍中。電子情報通信学会, 情報処理学会, 人工知能学会各会員。

木村 春彦 (正会員)

1974 年東京電機大学工学部応用理化学科卒業。1979 年東北大学工学研究科博士 (情報工学) 課程修了。同年富士通 (株) 勤務。1980 年金沢女子短期大学講師。1984 年金沢大学経済学部助教授を経て, 現在, 同大学工学部情報システム工学科教授。工学博士。その間, 最適コード変換, プロダクションシステムの高速化の研究に従事。情報処理学会, 電子情報通信学会各会員。

阿部 武彦 (正会員)

1988 年金沢大学経済学部経済学科卒業。同年, 大和コンピュータサービス (株) (現 (株) 大和総研) 入社。1993 年石川職業能力開発短期大学校講師。1994 年金沢大学大学院自然科学研究科社会人入学。1997 年金沢工業大学講師。現在に至る。工学博士。人工知能の推論高速化の研究に従事。