

遺伝的アルゴリズムの可視化技術に関する基礎研究～設計者の意図する最適解探索支援を目指して～

代表研究者	伊藤伸一	徳島大学大学院ソシオテクノサイエンス研究部
共同研究者	斎藤隆文	東京農工大学大学院生物システム応用科学府
共同研究者	満倉靖恵	東京農工大学大学院生物システム応用科学府

1 はじめに

多目的最適化問題では、パレート最適解が示すように多数の最適解候補がある。その中から最終的にたった一つの最適解を選択することは容易ではない。また、評価関数は設計者の経験的・実験的検証により決定されるか、もしくは関連研究論文などから引用されることは少なくなく、最終的な結果（評価値）によってその有効性が議論されている。もし、パレート最適解におけるパラメータと評価値が一目で容易に把握可能になれば、それらの関連性を確認することが可能になり、設計者が意図している最適解が探索されているか否かの判断が容易となる。そこで我々は、多目的最適化において一般的に利用される遺伝的アルゴリズム（genetic algorithm: GAs）を対象とし、情報可視化技術を適用する。情報可視化（Information Visualization）とは、情報を色や形状などで表現する技術である。

GAs の可視化に関する先行研究は、直接可視化型と間接可視化型に大別される。直接可視化型では、遺伝子の値、評価値・適応度、探索空間をそれぞれ可視化し、探索過程および結果を目視により確認するものである⁽¹⁾⁻⁽³⁾。しかしながら、遺伝子の値、適応度、探索空間が個別に表示されるため、直感的にその相互関係を把握することが困難となる。そのため、設計者が意図している解探索が実行されているか否かの判断も容易ではない。また、最適解候補（準最適解）を意味する個体数が増加するに従って、設計者の負担（確認作業）が増加することになる。さらに、可視化する探索空間は2または3次元で表現されるが、その空間を構成する軸の決定は容易ではなく、とくに多目的最適化においては困難となる。一方、間接可視化型では、探索空間の軸を情報圧縮技術により決定するもの⁽⁴⁾⁽⁵⁾と、染色体（遺伝子の集合）と評価値・適応度の相互関係を統計的処理により定量化することで可視化するもの⁽⁶⁾とがある。多数存在する情報圧縮技術の中において、自己組織化マップ（Self Organizing Map: SOM）を応用した可視化技術は、その有用性が実験的に検証されている⁽⁴⁾⁽⁵⁾。これには、評価関数の出力をそのまま SOM に入力し、2次元空間にマッピングする手法⁽⁵⁾と、定量化された染色体間の相互関係を SOM の入力とし、2または3次元空間にマッピングする手法が提案されている⁽⁴⁾。しかしながら、これらの間接可視化型では、遺伝子の値がどのように解探索および適応度に関与しているのかを一目で容易に把握・確認することができない。

これらの問題を解決するために、我々はこれまでに、探索空間を生成することなく、遺伝子の値、評価値・適応度を擬似カラーで表現する可視化法を提案した⁽⁷⁾⁽⁸⁾。これは、遺伝子の値と評価値・適応度の関連性を示すと共に、設計者が意図する解探索が実行されているか否かの判断を容易にするためである。また、多目的0-1ナップサック問題に適用し、実験的に提案手法の有効性を検証した。このとき、複数の評価関数を用いて比較検討している。その結果、評価関数の差異によって選択された品物の違いを一目で容易に把握することができ、設計者の意図を最も反映している評価関数の選出を支援することができた。しかしながら、遺伝子が01表現（binary-coded）であったため、局所解に陥る可能性が示唆された。また、他の多目的最適化問題への適用が困難であった。実数値表現（real-coded）へ展開することで、これらの問題点が解決可能であると考えられた。そのため、実数値遺伝的アルゴリズム（real-coded genetic algorithm: RGA）を対象とし、各最適解候補の個体の遺伝子の値と目的関数の値を可視化する技術を提案する。

提案手法は、パレートフロント（多目的最適化問題において最適解探索に利用されるグラフ）を用いることなく、疑似カラー表示法のみを用いて最適化されたパラメータ（遺伝子）とその評価値（目的関数の値）を視覚的に表示する。このとき、遺伝子と目的関数の値を正規化し、それぞれを別々に可視化する。また、可視化するに当たり、重要となるのが正規化方法である。提案手法では、3種類の正規化方法を用いて可視化し、各値を目視可能にする。さらに、遺伝子と目的関数の値の関連性を一目で理解可能な表示方法を提案する。

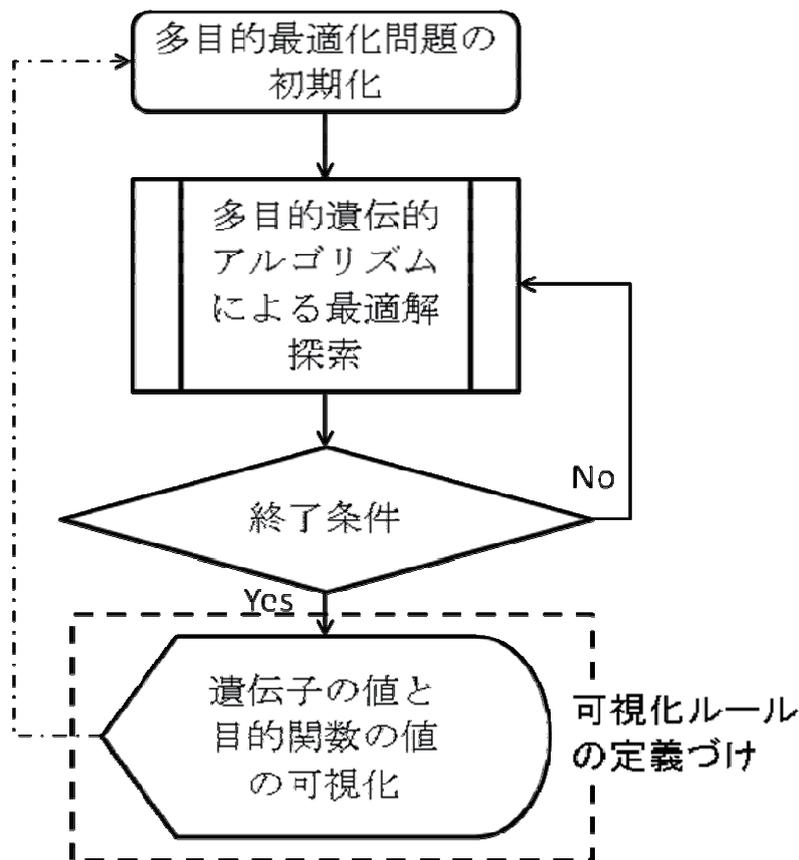


図1 多目的最適解探索アルゴリズムのフローチャート

多目的最適化問題におけるベンチマーク問題に提案手法を適用し、問題設定およびパレートフロントから読み取れる情報が、可視化結果から読み取ることが可能であることを評価の基準として、提案手法の有効性を検証する。

2 提案手法

提案手法は、多目的最適化問題における各目的関数と最適化された遺伝子の値を疑似カラー表示法により可視化する。また、多数の最適化候補間の遺伝子の値の差を確認するために、異なる3種類の正規化方法を用いて可視化する。図1に多目的最適解探索のフローチャートを示す。提案する可視化手法は、破線部で囲まれた最適化探索結果の可視化ルールの定義づけ部である。なお、最適解探索アルゴリズムは既存の手法 (Multi-objective genetic algorithm: MOGA) を用いる。

2-1 正規化方法

(1) 遺伝子の値の正規化

疑似カラー表示方法を用いた場合、正規化基準に基づいて確認可能な情報が異なる。設計者の意図する最適解の探索を支援するために、最適解の詳細を効率的に設計者に把握させる必要がある。そこで、提案手法では、異なる3種類の正規化方法を用いる。

1種類目 (Normalization technique 1: N1) は、最適解候補すべての遺伝子の値の最大値と最小値を用いて正規化する。最適解候補は、設定された複数の目的関数の要件をすべて満たす最適解である。一つの最適解を特定するためには、各最適解候補の相対的差を確認する必要がある。この正規化法は、すべての遺伝子の値の最大値・最小値を用いて正規化することで、最適化されたパラメータの相対的な違いを目視可能になる利点を持つ。

2種類目 (Normalization technique 2: N2) は、各遺伝子の値の最大値・最小値を用いて正規化する。スケールの異なるパラメータが最適化される場合、N1では遺伝子内差読み取ることが困難となる。たとえば、最適解探索に重要である要素のスケール幅が狭く、それほど重要ではない要素のスケール幅が広い場合、重

要な情報が埋もれてしまうことになる。この正規化法は、遺伝子内差を表現可能である利点を持つため、特定の遺伝子に着目して最適解を探索する場合に有効になる。また、N1 と N2 の結果が類似する場合、遺伝子の持つスケール幅が類似することを意味し、N1 の結果の中に埋もれてしまっている情報が少ないことが示唆される。

3 種類目 (Normalization technique 3: N3) は、各個体内における遺伝子の値の最大値・最小値で正規化する。この正規化方法は、設定した最適する要素の中で、特異的な値を示す要素を特定することが可能な利点を持つ。N1 と N3 の結果が類似することは、N1 と N2 の結果が類似することと同様に、遺伝子のスケール差が少ないことを意味し、N1 の結果の中に埋もれてしまっている情報が少ないことが示唆される。N2 と N3 の結果が類似することは、特異的な値を示す要素が目的関数の値に影響を及ぼしている可能性が示唆されかもしくはまったく影響を与えない要素であることが示されている可能性がある。

N1, N2 および N3 の結果が類似したコントラスト持つ場合、適切なスケールで構成された最適化問題であることが示唆され、加えて、そのコントラストの差が直接的もしくは間接的に目的関数に影響を与えることを意味する。

(2) 目的関数の値の正規化

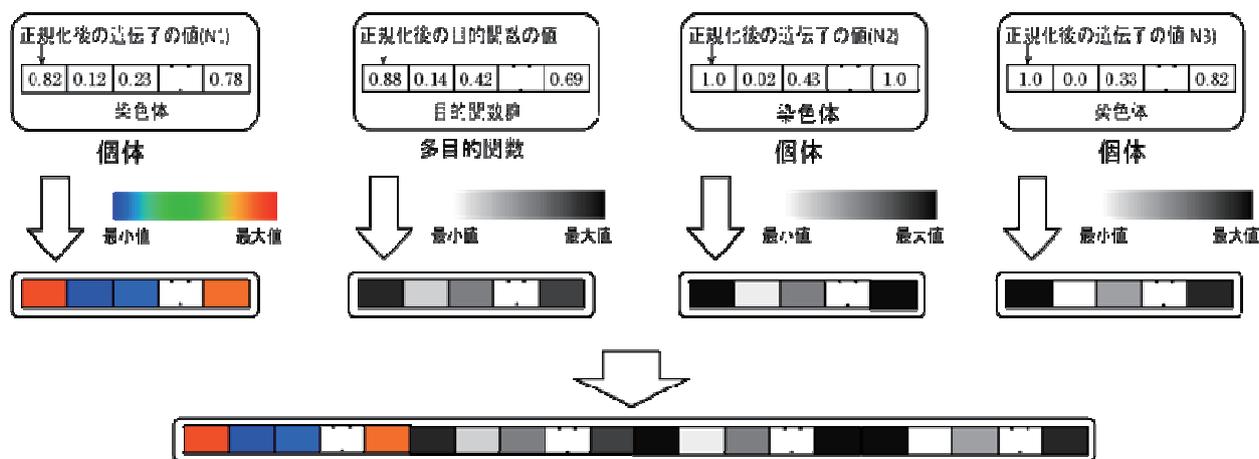
目的関数の値の正規化では、各目的関数の値の最大値・最小値を用いる。目的関数は遺伝子とは異なり、独立である場合が多く、トレードオフの関係を持つ可能性がある。くわえて、そのスケールは統合されにくい場合が多い。もし、複数の目的関数のすべての値の最大値・最小値を用いて正規化した場合、その関連性にバイアスがかけられ、誤った情報をユーザに提供することになることが容易に想定される。

2-2 可視化方法

図 2 に示すように、疑似カラー表示法により、個体の遺伝子と関連付ける画素に色を割り当てる。まず、2.1 節で記述した正規化方法により、最適候補を意味する個体の遺伝子の値と目的関数の値が正規化される。次に、最適化候補の個体の一つを取り出す。そして、遺伝子の値と目的関数の値に分ける。さらに、疑似カラー法により、各個体の遺伝子の値と各目的関数の値を可視化する。このとき、N1 により正規化された遺伝子の値はカラーバーに基づいて色づけされ、目的関数の値はグレーバーに基づいて色づけされる。加えて、N2 および N3 により正規化された遺伝子の値はグレーバーに基づいて色づけされる。

2-3 評価方法

図 3 に可視化結果の例を示す。可視化結果の表示フレームは、左から順番にカラーバー、疑似カラー表示された遺伝子の値、グレーカラー表示された個体の目的関数の値、異なる正規化方法によりグレーカラー表示された遺伝子の値、最後にグレーバーで構成されている。情報可視化技術の評価方法は確立されていないため、評価の方法として、ベンチマーク問題および各種パラメータ設定とパレート最適解で読み取ることが可能な情報が、可視化結果から読み取ることができるか否かを確認する。このとき、結果表示フレームにおける色づけされた遺伝子および目的関数の値のコントラストに基づいて情報は読み取られる。



一つの個体に対する可視化結果の例

図 2 疑似カラー表示方法の概要

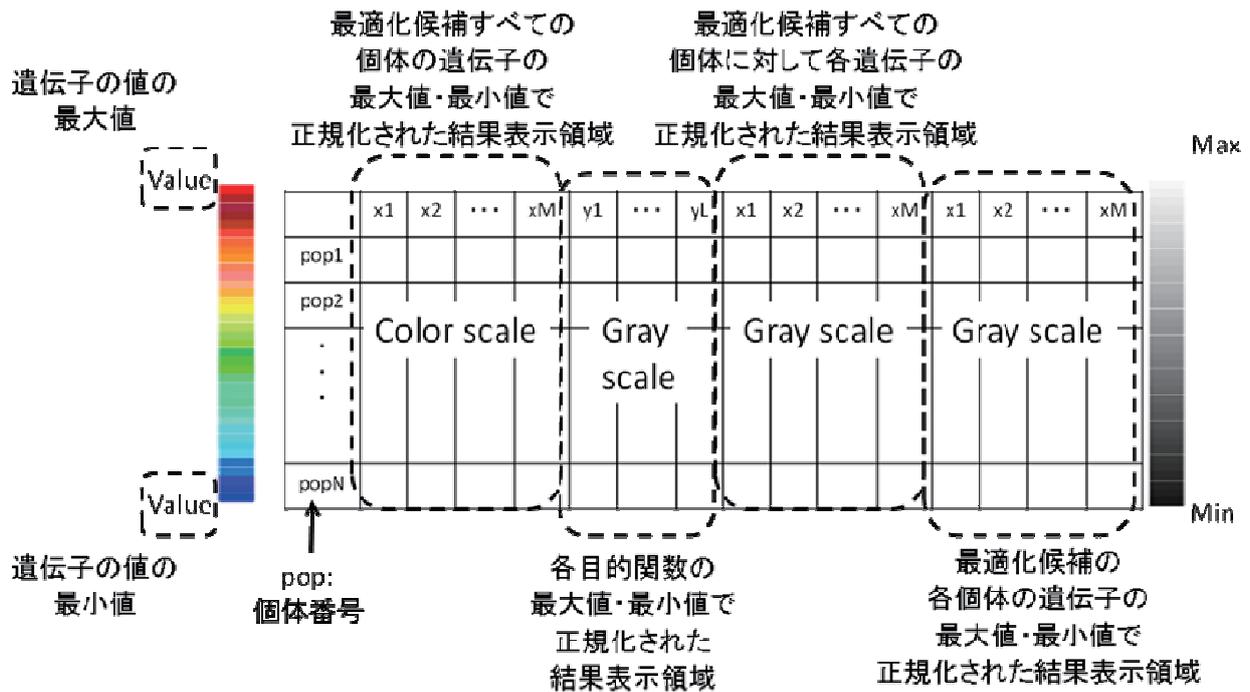


図3 可視化結果のサンプル

3 計算機シミュレーション

提案手法の有効性を示すために、多目的最適化問題のベンチマークに適用する。ベンチマーク問題には、KUR (Kursawe test function: KUR⁽⁹⁾) と ZDT6 (Zitzler, Deb, and Thiele test function: ZDT6⁽⁶⁾⁽¹⁰⁾) のテスト関数を採用した。なお、パレート最適探索には、MATLAB Global Optimization Toolbox を使用し、MOGA のアルゴリズムを採用した。最適解探索のための個体数は 50 個とした。また、選択淘汰処理にはトーナメント方式を採用し、交叉には一様交叉を採用した。さらに、交叉率および突然変異率は 0.8, 0.3 と設定した。

(1) テスト関数 KUR

テスト関数 KUR は、2 目的関数であり、そのパレート最適解は非連続でかつ非凸関数であり、多峰性の特性を持つ⁽¹¹⁾。図 4 にパレートフロントを示し、図 5 に提案手法による再撤回探索結果の可視化結果を示す。なお、遺伝子の個数 N は 3 とした。また、テスト関数 KUR の目的関数を下記に記す。第一目的関数 ($y1$) は、隣り合った変数 (遺伝子) の 2 乗和の総和に左右される関数である。第二目的関数 ($y2$) は、変数 1 ($x1$) を除くすべての変数の総和ならびに \sin 関数の総和に左右される関数である。

$$\text{Minimize } y1 = \sum_{i=1}^{N-1} g(x) \cdot \cdot (1)$$

$$\text{Minimize } y2 = \sum_{i=2}^{N-1} (|x_i|^{0.8} + 5 \sin(x_i^3)) \cdot \cdot (2)$$

$$g(x) = -10 \exp(-0.2 \sqrt{x_i^2 + x_{i+1}^2}) \cdot \cdot (3)$$

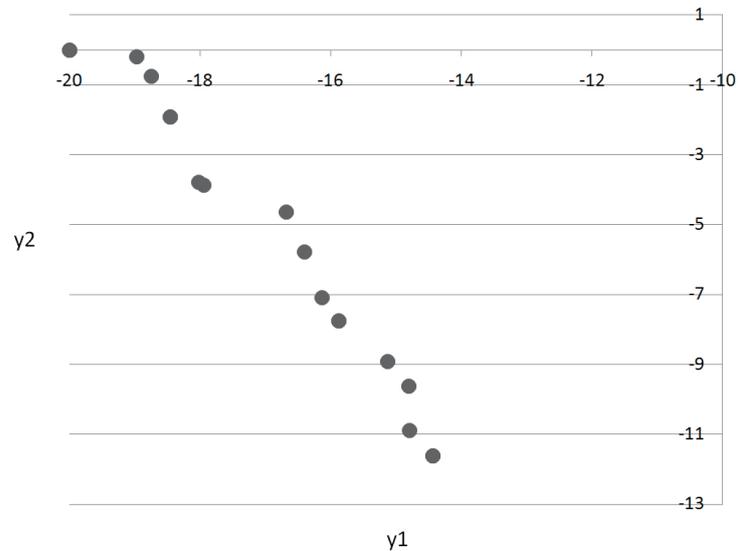


図4 テスト関数 KUR のパレートフロント

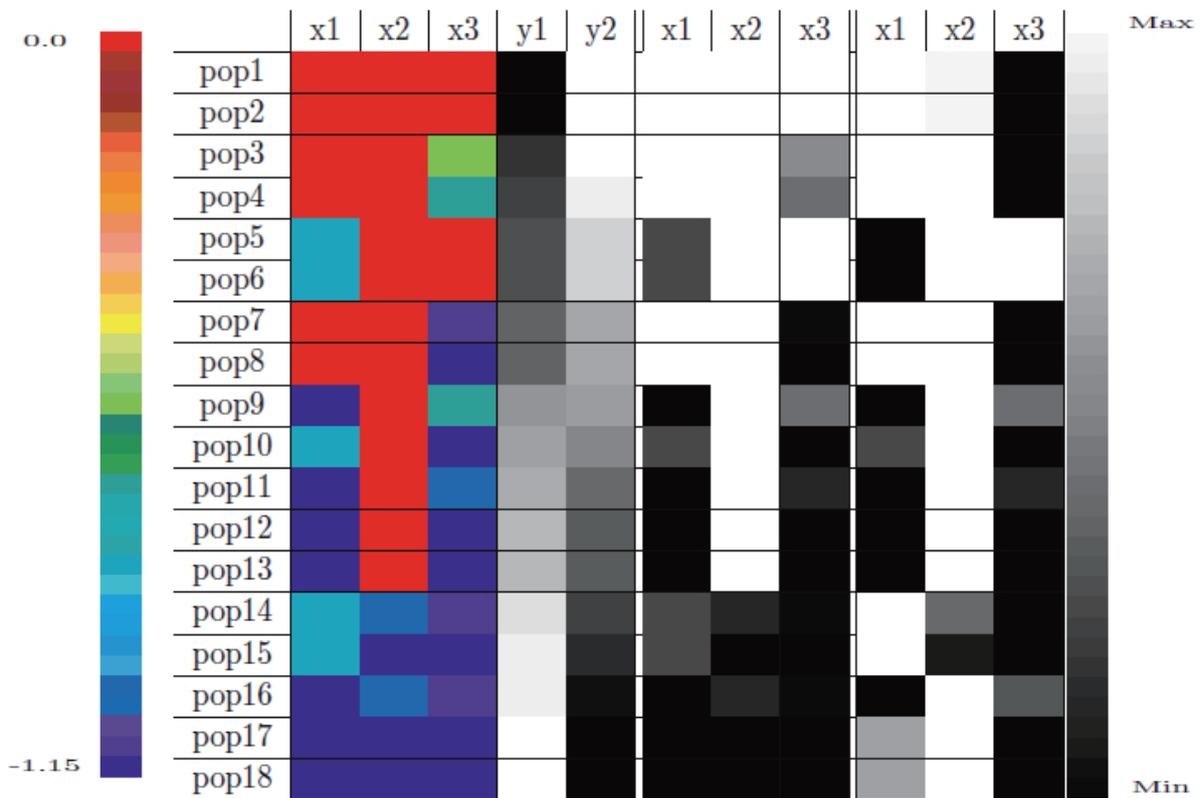


図5 テスト関数 KUR に対する最適解探索結果の可視化

テスト関数 KUR の結果において、第一目的関数 (y_1) は変数 2 (x_2) の値の増加に伴って、おおよそ減少する傾向にあることが伺える。KUR の第一目的関数の出力は隣り合った変数の 2 乗和の値に左右されるため、変数の数が 3 つの場合、変数 2 の影響を強く受けると推測される。このことは、パレートフロント (図 4) からは確認できないが、提案手法により遺伝子の値をカラー表示しているため、容易に把握することができた (図 5)。また、N1 と N2 のコントラストの傾向に差が見受けられなかった。これは、各変数 (遺伝子) に N1 と N2 で使用する最大値・最小値が含まれていたためだと推測される。さらに、N1 と N3 のコントラストの傾向において、 x_2 と x_3 にわずかな違いが見受けられた。また、N2 と N3 のコントラストの傾向に、同様の違いを確認することができた。具体的には、pop1, 2, 16, 17, 18 に違いが確認できる。N1 および N2 の結

果では、同色で表示されているが、実際には変数間の値は異なっていることをユーザは把握可能となった。しかしながら、ベンチマーク問題であるため、その差が意図する意味を把握するまでには至らなかった。

(2) テスト関数 ZDT6

テスト関数 ZDT6 は 2 目的関数であり、そのパレート最適解は単峰性の特徴を持つ⁽¹¹⁾。図 6 にパレートフロントを示し、図 7 に提案手法による再撤回探索結果の可視化結果を示す。なお、遺伝子の個数 N は 10 とした。また、テスト関数 ZDT6 の目的関数を下記に記す。第一目的関数 ($y1$) は、変数 1 ($x1$) のみで構成される関数である。第二目的関数 ($y2$) は、第一目的関数 ($y1$) と変数 1 ($x1$) を除くすべての変数の総和の除算の結果に左右される関数である。

$$\text{Minimize } y1 = 1 - \exp(-4x_1) \sin^6(6\pi x_1) \dots (4)$$

$$\text{Minimize } y2 = g(x)[1 - (y1/g(x))^2] \dots (5)$$

$$g(x) = 1 + 9[(\sum_{i=2}^N x_i)/(N-1)]^{0.25} \dots (6)$$

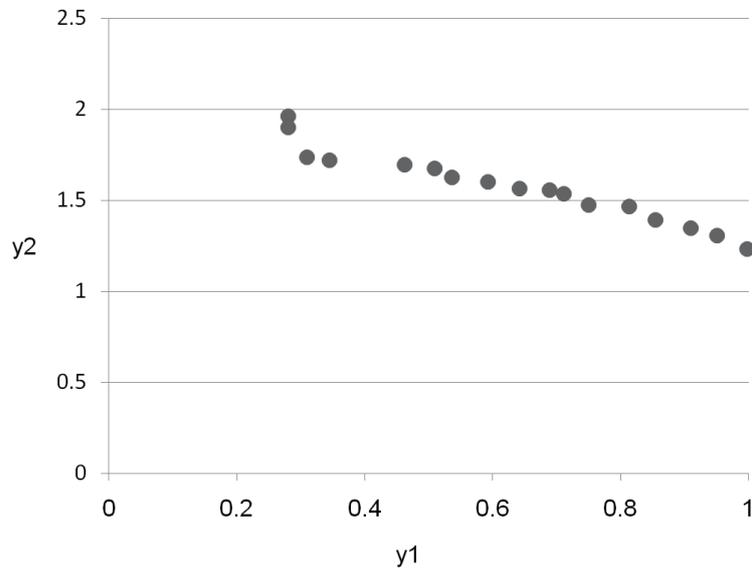


図 6 テスト関数 ZDT6 のパレートフロント

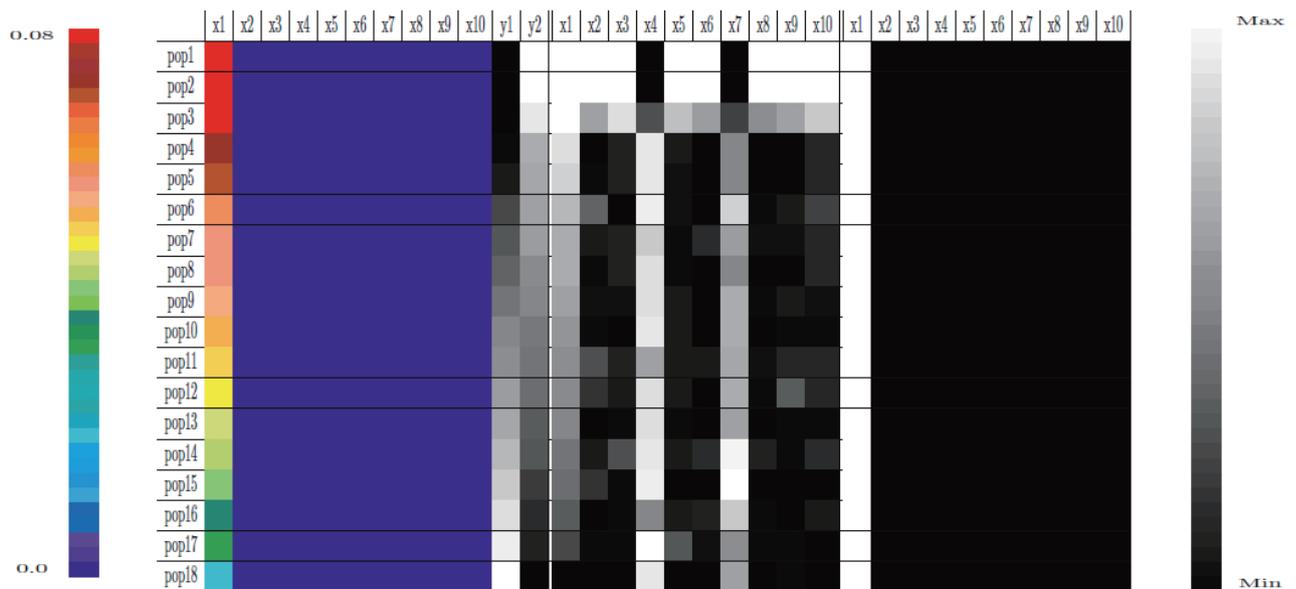


図7 テスト関数 ZDT6 に対する最適解探索結果の可視化

テスト関数 ZDT6 において、ZDT6 の第一目的関数の変数が変数 1 のみであるため、変数 1 (x_1) と第一目的関数 (y_1) との間に密接な関連性が存在する。このことは、パレート面からは確認できないが、可視化の結果 (図 7) におけるカラーバー (左サイドに位置している) に表示されている値を参照することにより容易に確認できる。しかしながら、ZDT6 は、 $x_1=[0.0, 0.2]$ における範囲でしか $y_1=[0.326, 0.7]$ を得ることができないように設計されている⁽¹¹⁾。 $x_1=[0.0, 0.2]$ は、変数 1 が極めて濃い青色で表示されることを意味するが、N1 の結果からでは濃い青色で表現された変数 1 を確認するに至らなかった。これは、Multi-Objective Genetic Algorithm (MOGA) を使用しているため、真のパレート最適解に到達していないことが原因であると推測される。また、NAG の結果では、変数 2~10 ($x_2\sim x_{10}$) の値の差を確認できなかった。これは、変数 2~10 が極めて 0.0 に近い値を持っているためだと考えられる。一方、N2 の結果では、各遺伝子の詳細な違いを確認することができた。しかしながら、その解釈および十分に考察するまでには至らなかった。これは、変数 2~10 が個別の意味を持つ変数ではないためである。さらに、N1 と N2 のコントラストの差の傾向を比較すると、正規化方法の違いによっては、把握可能な情報が異なることが明らかとなった。また、N1 と N3 のコントラストの傾向は類似した。これは、明らかに x_1 が特異的なようであることを意味し、くわえて、目的関数の値に影響を及ぼしていることを意味すると示唆される。N2 と N3 のコントラストの差の傾向は、N1 と N2 のコントラストの差の傾向と同様であった。

パレート最適解からたった一つの最適解を選択する場合、パレート最適解と判断されたすべての個体の遺伝子の値を確認する必要がある。一つ一つの個体に対するその遺伝子の値の確認には、個別に対応するための操作が必要になり、最適解候補数の増加に伴って、その作業コストは高くなることが容易に推測される。提案手法を適用することで、一つ一つの個体に対するその遺伝子の値を確認する作業が容易に行なえると示唆される。また、3 種類の正規化方法を用いることで、探索を支援するための情報を効率よく可視化することができていることが確認できた。しかしながら、実用的な多目的最適化問題に適用していないため、設計者が意図する一つの最適解探索を支援するまでに至っていない。

4 おわりに

多目的最適化問題におけるパレート最適解に対して、各目的関数から導出される評価値と最適化された遺伝子の値の詳細を一目で把握可能にするために、その情報を可視化する技術を提案した。また、2 つのベンチマーク問題に適用し、提案手法の有効性を検証した。その結果、一般的に利用されるパレートフロントでは把握不可能な最適化された遺伝子の値を、提案手法を用いることで把握可能になった。くわえて、3 種類の正規化方法を採用しているため、一つの正規化方法では確認が困難であった情報を把握可能になった。今後の課題は、実用的な多目的最適化問題に提案手法を適用し、提案手法の有効性を検証すると共に、設計者の意図する最適解探索を支援することである。

【参考文献】

- [1] W. B. Shine and C. F. Eick : “Visualization the evolution of genetic algorithm search processes” , Proceedings of 1997 IEEE International Conference on Evolutionary Computation, pp.367-372, 1997.
- [2] E. Hart and P. Ross : “Gavel-A New Tool for Genetic Algorithm Visualization” , IEEE Transaction on Evolutionary Computation, Vol.5, No.4, pp.335-348, 2001.
- [3] T. D. Collins : “Applying software visualization technology to support the use of evolutionary algorithms” , Journal of Visual Languages and Computing, 14, pp.123-150, 2003.
- [4] D. Yamashiro, T. Yoshikawa and T. Furuhashi : “Understanding Effects on Genetic Operators with Visualization of Search Process” Proceedings of Computational Science Symposium, pp.33-38, 2005.
- [5] 大林茂 : “多目的最適化とデータマイニング” , 日本機械学会誌, Vol.109, No.1050, pp.383-385, 2006
- [6] 石黒英敬・吉川大弘・古橋武 : “遺伝的アルゴリズムにおける遺伝子-評価値関係の可視化と遺伝的演算へのフィードバック” , 日本知能情報ファジィ学会誌, Vol.21, No.3, pp.327-337, 2009
- [7] 伊藤伸一・宮村浩子・満倉靖恵・斎藤隆文・福見稔 : “遺伝的アルゴリズムの可視化に基づく対話的スクリーンマ決定法” , 日本シミュレーション&ゲーミング学会秋季大会, pp. 83-86, 2007年9月
- [8] S. Ito, Y. Mitsukura, H. N. Miyamura, T. Saito and M. Fukumi : “A Visualization of Genetic Algorithm Using the Pseudo-color” , Proceedings of ICONIP 2007, vol. 2, pp. 444-452, November 2007.
- [9] E. Zitzler, K. Deb and L. Thiele: “Comparison of multiobjective evolutionary algorithms: empirical results” , Evolutionary Computation 8(2):173- 195, 2000.
- [10] A. A. Sadanandan : “A Comparative Study of Diversity Preservation Techniques in Multiobjective Evolutionary Algorithms” , Phil dissertation, The University of Nottingham Malaysia Campus, August 2007.
- [11] 渡邊真也・廣安知之・三木光範 : “近傍培養型遺伝的アルゴリズムによる多目的最適化” , 情報処理学会論文誌 数理モデル化と応用, Vol.43, No.SIG10, pp.183-198, 2002.

〈発表資料〉

題 名	掲載誌・学会名等	発表年月
最適解探索支援のための遺伝的アルゴリズムの可視化技術	動的画像処理実利用ワークショップ 2011 講演論文集	2011年3月
Visualization on Pareto Solutions in Multi-objective Optimization	Proceedings of 2011 International Association of Science and Technology for Development, Artificial Intelligence and Soft Computing, IASTED ASC2011	2011年6月