

## 二個体比較と自然言語での評価を用いた IGA による背景画像生成システム

大谷 紀子 研究室

2172056 小林 遼太

### 1. 背景と目的

遺伝的アルゴリズム (GA; Genetic Algorithm) の一種に対話型遺伝的アルゴリズム (IGA; Interactive Genetic Algorithm) がある。IGA は GA において個体の評価を人間が行うものとなっており、配色やメロディなど感性に即したものを作る手段として用いられる。幾何学模様シンボルマーク生成システム[1]をはじめとした IGA に基づく最適化システムでは、ユーザに解候補を評価させたうえで次世代の解候補を生成することで、ユーザの望む解を効率よく探索できる。しかし、数値での評価は難しいことや、ユーザの評価軸を汲み取れないことで、ユーザの好みを正確に反映しにくいことが課題である。

本研究では、IGA における評価がよりの確に次世代に反映されることを目的として、自然言語による評価と次世代生成の手法を導入した IGA システムを開発する。

### 2. 事前調査

評価文の解析に必要なデータセットの形式を決定するため、本学学生を対象に事前調査を実施した。ランダムに生成した 20 個の模様について、良し悪しなどの抱いた印象を自由記述で回答させた。得た回答には、色合いや配色などの視覚的特徴、および好ましさや力強さなどの感覚的評価に関連する語がさまざまな表現や表記で書かれていたことから、視覚的特徴と感覚的評価について定義したデータセットに加え、類似表現や表記ゆれに対応するためのデータセットも用意することにした。

### 3. システムの概要

最初に初期集団をランダムに生成する。各個体の染色体は、背景となるグラデーションの色や種類、細かさをはじめ、図形の種類、色、透明度、位置、回転のほか、図形を繰り返し描画する際の回数、範囲、変化を決定する計 48 の遺伝子から構成される。集団の多様性と模様の探索効率を考慮して、個体数は 20 とした。生成した個体の中から無作為に抽出した 2 個体を評価対象として表示する。同時に、色や柄などに関する評価をユーザが自然言語で入力できるようにする。評価画面の例を図 1 に示す。

ユーザが入力した評価文を 1 文ごとに解析する。前述のデータセットを用いて、文中の単語から「好きな特徴」、「改善してほしい特徴」などの着眼点を絞り込んだ上で、表示している模様の印象と良し悪しも推定する。絞り込んだ着眼点と推定した印象をもとに、第*i*世代*j*番目の評価文に関する個



図 1 システムの評価画面

体 $x$ の適応度 $f_{ij}(x)$ を式(1)のように定義する.

$$f_{ij}(x) = w_{ij} \sum_{k=1}^{p_{ij}} e^{-s_{ij}(r_{ijk}-c_{ijk})^2} \quad (1)$$

ここで,  $w_{ij}$ は推定した良し悪しの度合い,  $p_{ij}$ は絞り込んだ着眼点の数,  $s_{ij}$ は程度を表す語を考慮する度合い,  $r_{ijk}$ は一つの着眼点において評価文から推定した基準値,  $c_{ijk}$ は着眼点における個体の特徴を示す値である. また, 第 $n$ 世代における個体 $x$ の適応度 $g_n(x)$ を式(2)のように定義する.

$$g_n(x) = \sum_{i=1}^n a^{n-i+1} \sum_{j=1}^{m_i} f_{ij}(x) \quad (2)$$

ここで,  $m_i$ は第 $i$ 世代の評価文の数,  $a$ は1世代前の評価文による適応度を考慮する度合いである.

全個体の適応度を算出した後, 表示個体とそれ以外の個体それぞれからルーレット選択で親を選択し, 一様交叉と突然変異によって次世代の個体を生成して適応度を求める. 適応度が最も高い個体 A と, A との類似度が一定以下でかつ適応度が最高の個体 B を次世代の評価対象として表示する. ユーザが望む個体が得られるまで, 評価と次世代生成を繰り返す.

また, 初期集団を生成すると同時に, システムの使い方を説明する画面を表示する. 説明画面に入力例を示すことで, ユーザの入力をより反映されやすい文章に誘導することができる.

#### 4. 評価実験

本学学生 11 名を被験者として, 評価実験を実施した. 評価実験ではシステムを使用させた上で, 被験者の入力とシステムの出力, およびシステムの使いやすさや出力に対する満足度などの意見をアンケートフォームで収集した.

アンケートでは, 「初期候補と目標イメージの近さ」, 「入力した評価文の反映度」, 「システムの使いやすさ」などを 5 段階評価で回答させた. また, 被験者が目標としたイメージや特に反映された評価文と反映されなかった評価文, システムの改善点などに関して記述式で回答させた.

アンケートの結果, 「入力した評価文の反映度」の平均値は 3.45 となった. 特に反映された評価文の項目では, 「もっと明るくして欲しい」という趣旨の文や, 左右で良かった方の画像を選ぶような文が得られた. 一方で, 「夜景みたいにしたい」, 「線が濃すぎて嫌い」, 「特定の色に変更したい」などの評価文は反映されなかったとの声があった.

システムの使いやすさでは平均値が 3.82 となり, 自由記述では「画像の良し悪しを言える」, 「前までの文を見直せる」などの回答を得られた. 一方で, 「初期候補を表示する前に入力をしたい」, 「過去の画像を参照したい」などの意見が挙げられた.

#### 5. 考察

評価実験の結果, 使いやすさや評価の反映度に一定の評価を得られた一方で, 改善の余地も多く挙げられた. 本システムでは, 次世代の評価対象を適応度の高さと 2 個体間の類似性の低さで決定しているが, ユーザに複数の候補から選択させることで, 理想の個体をより少ない評価回数で得ることができると考える. また, 評価のポイントを汲み取る過程において, データセットに登録する単語を増やすだけでなく, 評価ルールの作成アルゴリズムを改善することも今後の課題といえる.

被験者による評価入力の回数において中央値は 9 回となった. 9 回を超えて次世代生成を行った被験者の出力記録からは, およそ 7 回目から終了時まで入力内容にかかわらずほとんど同じ模様が出力されていたことが確認できた. 入力済みの好みを引き継いでまったく別の個体を表示できるようにしたり, 多く使われた評価から即座に次世代生成へ移行できるようにしたりすることで, より多くの模様を少ない負担で探索できると考える.

#### 参考文献

- [1] 高石 麻貴奈, “個人の持つイメージに即した幾何学模様シンボルマーク生成システムの構築”, 東京都市大学メディア情報学部情報システム学科卒業論文, 2023.