

# 対話型進化計算システムにおける一対比較評価の有用性

竹之内宏<sup>1</sup> 徳丸正孝<sup>2</sup>

<sup>1</sup>福岡工業大学 <sup>2</sup>関西大学

本稿では、ユーザの感性情報を用いて対象を最適化する対話型進化計算（Interactive Evolutionary Computation：IEC）手法において、ユーザが解候補を評価する際に一対比較評価手法を用いたシステムの有用性について述べる。筆者らはこれまでに、IECにおける一対比較評価手法として、トーナメント式評価手法を用いたIECシステムを提案し、基礎実験を通して解候補の進化性能やユーザの評価負担に関する有効性を示している。本研究では、トーナメント式評価手法の改良版となる勝ち残り一対比較評価手法（Winner-based Paired Comparison method：WPC）を提案し、商品カスタマイズにおける応用を視野に入れたIECシステムを構築し、その有用性を検証した。WPCでは、ユーザは解候補の一対比較評価を繰り返し、各対戦において好みと判定された解候補がそのまま次の対戦でも提示される。また、本研究では、WPCにおいて、ユーザが一対比較評価に迷った場合に、「どちらも好き」「どちらも嫌い」の判定ができる機能を付け加えたシステムに関する検証も行った。本稿では、実ユーザを対象とした評価実験により、商品カスタマイズにおける提案システムの有用性を検証した。実験結果より、提案システムはユーザの解候補評価のしやすさの面で有用であることを確認した。

## 1. 対話型進化計算の研究背景と一対比較評価の導入

近年、ユーザがWebサイト上で、自分好みの商品デザインをカスタマイズし、購入できるサービスが普及している。たとえば、時計ブランドKnot[1]のWebサイトでは、ユーザは腕時計本体の形や色、ベルトの形、素材などを組み合わせて、オリジナルの腕時計を作成し、購入できる。また、シューズブランドNew Balance[2]のWebサイトでは、いくつかのシューズの形から好みのシューズを選択し、さらにシューズの紐やソール、ロゴの素材や色を選択し、オリジナルのシューズを作成し、購入できる。しかし、このようなサービスでは、ユーザはオリジナルのデザインを作成し購入するまでに、多数のパーツの素材や色などを選択し、組み合わせなければならない。作成過程では、一部のパーツを変更するたびにデザイン全体のイメージを確認する作業も入るため、ユーザはオリジナルのデザインを作成するまでに、多くの時間と労力を費やすことになる。

この問題を解決できると考えられる手法の1つに、対話型進化計算（Interactive Evolutionary Computation：IEC）手法がある[3]。IECでは、ユーザはコンピュータが提示したデザインを評価するだけで、ユーザの感性に合ったデザインを進化計算（Evolutionary

Computation : EC) 技術によって動的に生成する。このため、IECでは、ユーザがデザインに対する専門的な知識を有していなくても、提示されたデザインの良し悪しや好みを回答するだけでユーザの好みのデザインを生成できる。IECを用いたシステムはこれまでに数多く提案されており、既存の配色データベースから所望の配色を見つけ出す配色支援システム[4]や毛筆フォントを自動で作成するシステム[5]などがある。しかし、従来のIECシステムでは、ユーザは提示されるデザインに対し、5段階や10段階などの段階評価を求められることが多かった。このため、ユーザは、各デザインに対する印象を数値に置き換えて何世代にもわたって評価しなければならず、ユーザの評価負担が膨大になるという問題があった。

これを解決する方法の1つに、評価インタフェースの改良がある。本研究では、ユーザが解候補に対して点数付けを行う評価ではなく、解候補の一対比較評価によって、解候補の良し悪しを決定する勝ち残り一対比較評価手法 (Winner-based Paired Comparison method : WPC) を提案する。WPCでは、一対比較評価において、ユーザが好みと判定した解候補が次の対戦に進出し、他の解候補とともに提示され、一対比較評価を繰り返していく。このように、WPCでは、対戦経過における2つの提示解候補に関連性を持たせているため、ランダムに2個のデザインを提示するよりは、ユーザの解候補の評価に対するモチベーションが向上すると考えられる。

しかし、一対比較評価には、ユーザが好みのデザインを選択するときに迷った場合、ユーザはどちらかのデザインを強制的に選択しなければならないという問題がある。そこで本研究では、システムが提示する2種類のデザインに対してユーザが優劣を付け難い状況を考慮し、提示された2種類のデザインに対して「どちらも好き」「どちらも嫌い」の判定ができる機能をWPCの評価インタフェースに追加した。これにより、ユーザはデザインの選択に迷ったとき、強制的に好みのデザインを選択する状況を避けられると考えられる。

本研究では、「どちらも好き」「どちらも嫌い」の判定機能を追加しているWPCおよび追加していないWPCをそれぞれ提案システムAおよびBとして、実ユーザを対象とした解候補の進化性能やユーザの評価におけるシステムの使用感などを比較した。本研究では、IECシステムが一般に普及した際に問題となるであろうユーザの評価作業を単純にし、より評価のしやすいインタフェースの提案を目指した。

本研究では、実ユーザを対象とした評価実験により、提案システムで生成される女性用腕時計デザインの満足度やユーザの評価負担軽減に関する有効性を検証した。また、「どちらも好き」「どちらも嫌い」の判定が行われた際のタイミング、使用回数などについても考察した。

---

## 2. 関連研究

---

先行研究では、ユーザが提示された複数のデザインから好みのデザインを選択する評価インタフェースが提案されている[6],[7],[8],[9]。筆者らはこれまでに、トーナメント表にランダムに配置された解候補に対して、一対比較評価によって解候補の良し悪しを決定するトーナメント式評価手法を提案している[6]。トーナメント式評価手法では、ユーザはトーナメント表の各対戦の2つの解候補について、どちらが好みであるかの一対比較評価を行う。トーナメント式評価手法はユーザの評価作業の単純さから、評価負担は従来の段階評価手法よりも軽減できることが確認されている[6]。しかし、トーナメントの組合せや各対戦の勝敗により、全解候補に必ずしも正確な評価値が与えられず、ユーザの評価負担を軽減できても、解候補の進化性能が低下してしまうことが想定される。文献[6]では、このような懸念に対して、コンピュータ上で擬似的にユーザの

評価を模倣した数理モデルを用いて数値シミュレーションを行い、解候補の進化性能を定量的に検証した。その結果、トーナメント式評価手法は従来の段階評価方式と比較して、遜色のない解候補の進化性能をもつことが示された[6]。このようなシミュレーション結果については、実ユーザを対象とした評価実験によっても同様の傾向が確認され、トーナメント式評価手法では、ユーザが段階評価手法と同等に満足いくデザインを生成でき、段階評価手法よりもユーザの評価負担を軽減できることが確認された[6]。

さらに、一対比較による評価では、ユーザが一度に比較する解候補数が2個となる。このため、IECの評価対象が特に音楽や動画などの時系列データの場合は、一度の評価で比較する解候補が2個で済むため、段階評価方式よりもユーザの評価負担は軽減できることが確認された[6]。この結果は、評価対象が時系列データではない静止画のときでも同様であった。また、一対比較評価を用いる際には、単純に全解候補の中から2つの解候補を取り出しユーザに評価を求めた場合、解候補数を $n$ とすると、ユーザは $(n(n-1)/2)$ 回の評価を行わなければならない。たとえば、解候補数が16個の場合、ユーザは毎世代120回もの一対比較を行うことになる。このため、ユーザの評価回数が膨大になり、解候補評価負担が大きくなってしまう。しかし、トーナメント式評価手法においては、解候補数が $n$ のとき、ユーザの評価回数は毎世代 $(n-1)$ 回で済む。たとえば、解候補数が16個の場合、ユーザは毎世代15回の評価を行うことになる。このため、トーナメント式評価手法は一対比較法と比較して、ユーザの評価回数は十分少なく、解候補評価負担を軽減できる。

しかし、トーナメント式評価手法では、ユーザが連続的に行う一対比較評価において、特に1回戦における対戦間の解候補対の関連性が薄い。このため、ユーザは連続的に関連性のない一対比較を複数回繰り返すため、評価に対するモチベーションが低下してしまうおそれがある。このような問題は、たとえば、勝者となった解候補が次の対戦で他の解候補と対戦するなど、対戦間に関係性を保持させることでより軽減できると考えられる。このような手法として、本研究では、WPCを提案し、さらにユーザの評価の迷いを軽減するため、「どちらも好き」「どちらも嫌い」の判定ができる機能を追加した際の検証を行った。また、WPCの評価回数はトーナメント式評価手法と同等で、解候補数が $n$ のとき、 $(n-1)$ 回で評価負荷が増大することはない。

---

## 3. 提案システム

---

### 3.1 概要

提案システムにおいては、たとえば時計店など実際の店舗やWebページ上での商品カスタマイズにおいて使用されることを想定している。このような場面では、ユーザが一つひとつのデザインに対して、詳細な評価を入力するよりは、ある程度評価が粗くても手軽に評価できるインタフェースの方が、評価のしやすさの観点からは有用であると考えられる。このため、本研究では、より実デザインに近い評価対象として腕時計デザインを作成し、提案システムを構築した。

図1に提案システムの概要を示す。まず、初期解候補群を生成し、その中からランダムに選択された2種類の女性用腕時計デザインをユーザに提示する。次に、ユーザに好みの腕時計デザインを選択、もしくは「どちらも好き」「どちらも嫌い」の判定による評価を行ってもらう。このとき、提案システムAでは、「どちらも好き」「どちらも嫌い」の判定ができるが、提案システムBではできない。ユーザが1世代分の評価を終えるとEC処理によって、新たな解候補が生成さ

れる。本研究では、IECのECアルゴリズムには遺伝的アルゴリズム（Genetic Algorithm：GA）を用いている。そして、生成された女性用腕時計デザインを再びユーザに提示する。これらの操作を繰り返し、ユーザの満足のいく解候補を生成する。



図1 提案システムの概要

### 3.2 勝ち残り一対比較評価手法

WPCでは、同一世代に生成されたすべての解候補に初期評価値として1点を与え、一対比較による対戦を行い、勝利解候補の評価値に敗者解候補の評価値を加算する。そして、敗者となった解候補を他の解候補と入れ換え、再び一対比較評価が行われていく。1世代分の評価が終了すれば、これらの評価値を用いてGA処理を行う。

図2にWPCにおける解候補の評価方法を示す。まず、A～Eの5つの解候補を生成し、それぞれの解候補に初期評価値1点を与える。

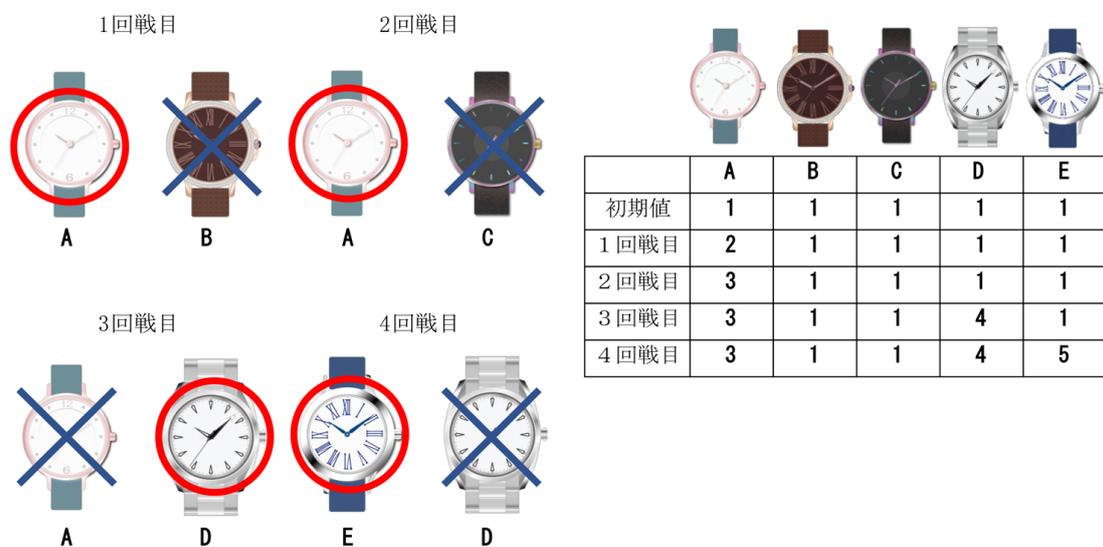


図2 WPCの解候補評価方法

1回戦目では、ユーザは解候補AとBを比較し、Aを好みの解候補として選択している。この結果に基づいて、WPCは敗者解候補Bの評価値1点を勝者解候補Aの評価値1点に加算する。これにより、解候補Aの評価値は2点になる。その後、勝者解候補Aを次の対戦に進出させ、敗者解候補Bを次の解候補に入れ換える。

2回戦目では、ユーザは解候補Aと解候補Cを比較し、解候補Aを好みの解候補として選択している。この結果に基づいて、WPCは敗者解候補Cの評価値1点を勝者解候補Aの評価値2点に加算する。これにより、解候補Aの評価値は3点になる。その後、解候補Aを次の対戦に進出させ、敗者解候補Cを次の解候補へ入れ換える。

以降、3回戦目では解候補D、4回戦目では解候補Eが好みの解候補として選択されている。これらの結果に基づいて、3回戦目では勝者解候補Dの評価値が4点、4回戦目では勝者解候補Eの評価値が5点となる。4回戦終了後の解候補A、B、C、D、Eの評価値は、それぞれ3、1、1、4、5点となる。

提案システムBでは、これらの評価値はGA処理において用いられる。

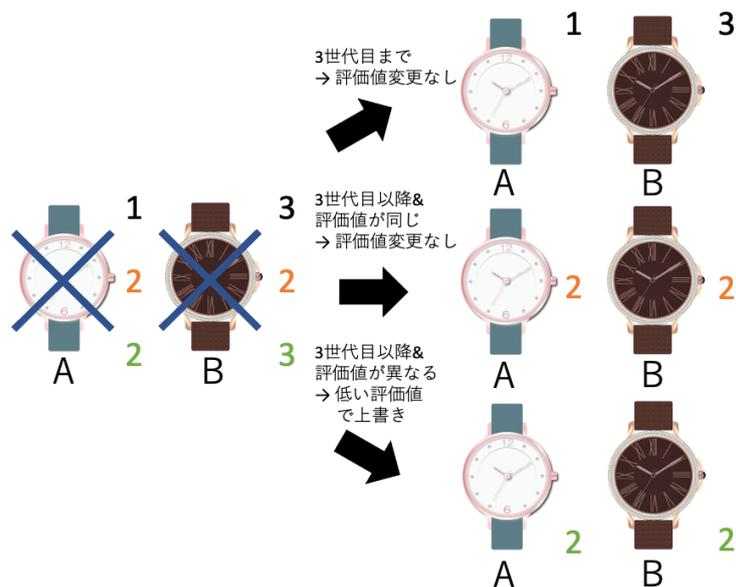
### 3.3 どちらも好き／どちらも嫌い判定の導入

3.2節で述べたWPCでは、各対戦において、ユーザが好みと判定した解候補は勝ち残り、もう一方の解候補は新たにランダムに入れ換えられる。提案システムAでは、WPCの解候補評価に、「どちらも好き」「どちらも嫌い」の判定ができる機能を追加している。これらの機能がユーザによって使用された際には、使用されたときの世代交代数および両解候補の評価値によって、解候補の評価方法が異なる。

図3に「どちらも好き」「どちらも嫌い」判定時の解候補評価方法を示す。図3のように、両デザインの評価値が同じ場合とそうでない場合、また「どちらも嫌い」判定のときは判定時の世代交代数によって評価方法が異なる。これらの評価方法は、両デザインの評価値の整合性を考慮して決定している。



(a) 「どちらも好き」評価



(b) 「どちらとも嫌い」評価

図3 WPCにおける「どちらとも好き」「どちらとも嫌い」評価

「どちらとも好き」と判定されたときは、図3 (a) のように、解候補AとBの評価値が同じ場合と異なる場合で解候補の評価方法が異なる。

2つの解候補の評価値が同じ場合は、解候補AとBそれぞれに1点を加算する。このため、解候補AとBの評価値はそれぞれ2点となる。これは、「どちらとも好き」判定によって両解候補が好みと判定されていることから、両解候補に対して評価値を加算するためである。その後、解候補AとBのどちらかを次の対戦に進出させ、もう一方の解候補を次の解候補に入れ換える。

2つの解候補の評価値が異なる場合は、評価の低い解候補の評価値を評価の高い解候補の評価値で上書きする。図3 (a) の場合では、解候補Aの評価値を評価値の高い解候補Bに合わせている。このため、解候補AとBの評価値は3点となる。すなわち、「どちらとも好き」判定によって両解候補が好みと判定されていることから、対戦前は評価が低かった解候補も現在の対戦によって、評価の高かった解候補と同等の評価値になっている。その後、評価値が同じ場合と同様に、解候補AとBのどちらかを次の対戦に進出させ、もう一方の解候補を次の解候補に入れ換える。

「どちらとも嫌い」と判定されたときは、図3 (b) のように解候補AとBの評価値および評価時点での経過世代交代数によって、解候補の評価方法が異なる。

評価時点での経過世代交代数が3世代目までは、初期世代にランダムに生成された解候補群が十分に進化しておらず、ランダム探索の段階とも考えられる。このため、ユーザが提示された2つの解候補を「どちらとも嫌い」と判定することは、解候補がある程度収束するまでは多いと考えられる。これらを考慮し、経過世代交代数が3世代目までは、両解候補の評価値は変更しない。このため、解候補Aの評価値は1点、解候補Bの評価値は3点となる。その後、両解候補を次の解候補に入れ換える。

経過世代交代数が3世代目以降になると、ある程度の解候補収束が見込まれる。このため、ユーザが提示された2つの解候補を「どちらも嫌い」と判定した場合は、両解候補の評価値を維持または低い方の評価値で上書きするようにしている。3世代目以降で両解候補の評価値が同じ場合は、双方の評価値をそのままにする。図3 (b) では解候補AとBの評価値は2点になる。その後、解候補AとBのどちらかを次の対戦に進出させ、もう一方の解候補を次の解候補に入れ換える。3世代目以降で両解候補の評価値が異なる場合は、評価の高い解候補の評価値を評価の低い解候補の評価値で上書きする。図3 (b) の場合では、解候補Bの点数を評価値の低い解候補Aに合わせている。このため、解候補AとBの評価値は2点となる。その後、解候補AとBのどちらかを次の対戦に進出させ、もう一方の解候補を次の解候補に入れ換える。

提案システムAでは、3.2節で述べた評価方法に本節で述べた評価方法を合わせて得られた評価値がGA処理において用いられる。

### 3.4 評価インタフェース

図4に提案システムの評価インタフェースを示す。本研究における提案システムは、実店舗などでの利用を想定し、Androidアプリおよびタブレット端末 (Google Nexus 9) を用いて実装している。ユーザが解候補を評価する際には、図4のレイアウトでデザインが提示される。ユーザは、2種類の腕時計デザインを比較し、好みのデザインをタップ、もしくは両方の腕時計デザインが好きな場合は、画面左下部にある「どちらも好き」ボタンをタップ、両方の腕時計デザインが好みでない場合は、画面右下部にある「どちらも嫌い」ボタンをタップする。提案システムBでは、これらの「どちらも好き」「どちらも嫌い」ボタンが表示されない。ユーザが1世代分の評価を終えると、3.2節および3.3節で述べたようにして、各デザインに評価値が与えられ、次の世代の腕時計デザインがGA処理により生成される。



図4 WPCの評価インタフェース

### 3.5 評価対象

図5に女性用腕時計デザインの構成を示す。本研究における女性用腕時計デザインは、文字盤、ケース、針、リューズ、ベルトの5つのパーツで構成されている。

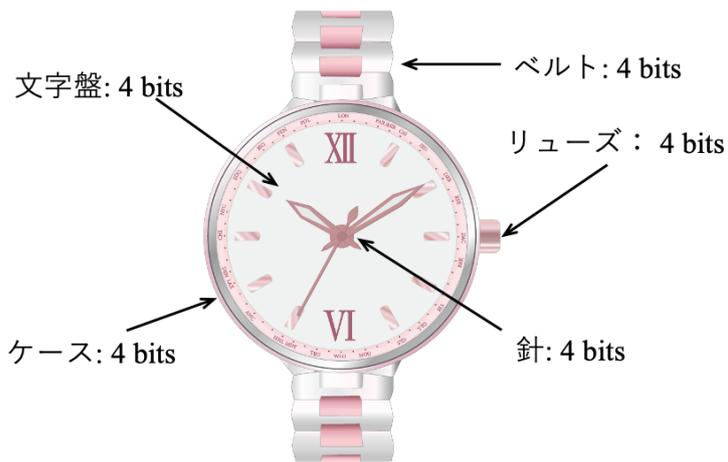


図5 女性用腕時計デザインの構成

図6に各パーツの遺伝子コーディングを示す。図6では、リューズパーツの大きさは実際のものよりも拡大している。各パーツは、それぞれ16種類のデザインを有しており、遺伝子列にはそれぞれ4bitsが割り当てられている。このため、遺伝子長は20bitsとなり、提案システムでは、合計約100万種類(=2<sup>20</sup>)の腕時計デザインを生成できる。なお、遺伝子型におけるビットパターンと表現型における腕時計デザインパーツの見た目の類似度との遺伝子コーディングは、筆者らの主観によって決定している。

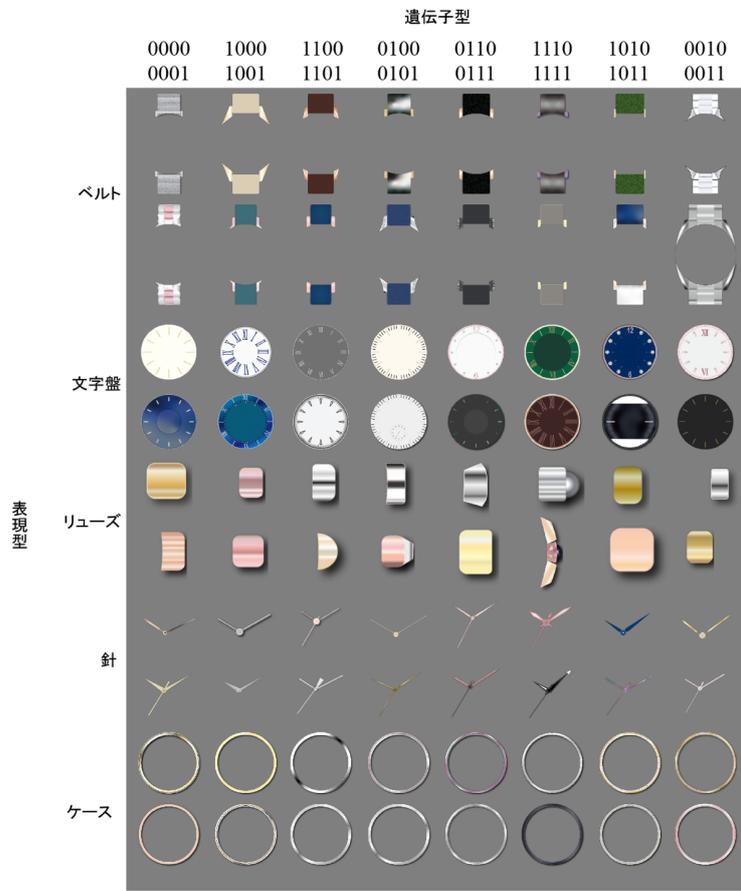


図6 女性用腕時計デザインの遺伝子コーディング

## 4. 評価実験

### 4.1 実験概要

本研究では提案システムにおいて、解候補の進化性能およびユーザの評価のしやすさの観点から、実ユーザを対象とした評価実験を行い、提案システムの有用性を検証した。本実験では、提案システムにおいて、「どちらも好き」「どちらも嫌い」判定が行われた際のタイミングや使用回数、さらに提案システムにおいて評価に要した時間、生成された女性用腕時計デザインの満足度およびユーザの評価負担について検証した。また、被験者から取得したアンケート結果についても考察した。

本実験の被験者は、10代から20代の大学生41名（男性20名、女性21名）である。被験者には、2つの提案システムAおよびBを使用してもらった。実験開始前には、被験者に提案システムを使用する上でのデザインコンセプトを教示した。デザインコンセプトは、男性被験者の場合は好きな女性にプレゼントするための腕時計を作成すること、女性被験者の場合は自分がプライベートで使用する腕時計を作成することとした。また、本実験では、順序効果による被験者の回答値のバイアスを排除するため、被験者によって2つの提案システムの使用順序はランダムに設定した。

表1に本実験のパラメータを示す。1世代における解候補数は10個であり、腕時計デザインの構成より、遺伝子長は20bitsである。本実験では、世代交代数は10とし、どの被験者においても10世代分の評価を実施してもらうこととした。これは、各被験者に一定の評価作業をしてもらい、評価負担や評価に要した時間を公平に比較するためである。選択方法、交叉方法はそれぞれGAにおける一般的な方法を用いた。突然変異率は、同様のデザインが多く提示されることによって、被験者が腕時計デザインの評価に飽きることを防ぐため、通常のGAよりは高い15%とした。これによって、解候補群にある程度のバリエーションを持たせることができると考えられる。提案システムにおける突然変異処理は、各遺伝子座に対して突然変異率によってビット反転を行う処理とした。

表1 実験パラメータ

解候補数	10
遺伝子長	20 bits
世代交代数	10
選択方法	ルーレット選択+エリート保存
交叉方法	一様交叉
突然変異率	15%

被験者は各システムを使用した後、各世代の女性用腕時計デザインのエリート個体に対する満足度を5段階評価で回答する。次に、被験者は以下に示す項目a~dのアンケートに回答する。項目aでは、1. ベルト、2. リューズ、3. ケース、4. 文字盤、5. 針の5つの選択肢から1つを選択してもらう。項目b~dでは、それぞれ5段階評価で回答してもらう。

項目a（デザインパーツの重要度）：

- ・腕時計を評価する上で、どのパーツを重視したか

項目b（解候補群の進化性能）：

- ・腕時計を評価するごとに、好みの腕時計が表示されたか

項目c（解候補群の収束性）：

- ・腕時計を評価するごとに、同様の腕時計が表示されたか

項目d（レイアウト面の評価）：

- ・腕時計は評価しやすかったか

最後に、被験者は、提案システムAとBのどちらが腕時計デザインを評価しやすかったかについて、「どちらも好き」「どちらも嫌い」判定の有無による観点より、回答する。また、一連の実験において、気づいたことなどのヒアリングを被験者に対して行った。

#### 4.2 「どちらも好き」「どちらも嫌い」判定の利用結果

本節では、被験者が提案システムAを使用する中で「どちらも好き」「どちらも嫌い」判定がどの程度行われたかについての検証結果について述べる。

まず、実験結果より、これらの判定を1回以上使用した被験者は、全被験者の88%であった。図7に世代ごとの両判定回数の平均、表2に被験者別の両判定の使用内訳を示す。図7より、いずれの世代においてもいずれかの被験者がどちらかの判定を使用していることが確認された。また、表2より、「どちらも好き」のみ利用した被験者が男女ともに最も多くなっている。さら

に、すべての被験者の全評価において、通常の一対比較評価は全評価中83%、どちらも好き評価は11%、どちらも嫌い評価は6%であった。この結果より、両判定は全評価数と比較すると少数であったことを確認した。

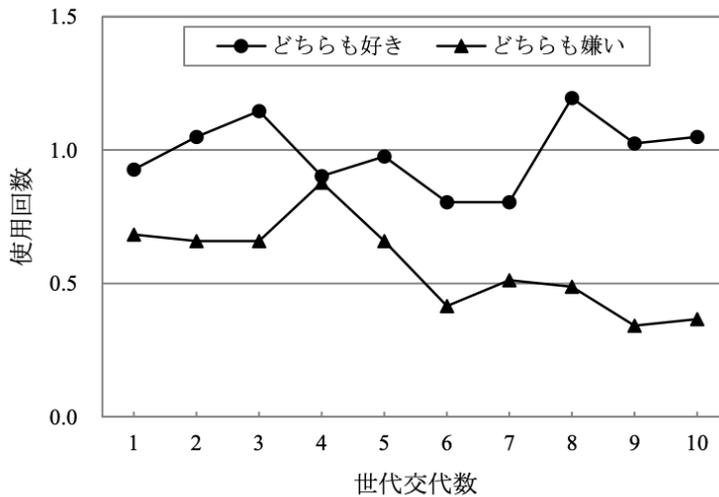


図7 「どちらも好き」「どちらも嫌い」の判定回数の結果

表2 「どちらも好き」「どちらも嫌い」判定の使用内訳

	男性被験者	女性被験者
「どちらも好き」のみ	50%	43%
「どちらも嫌い」のみ	15%	14%
両方利用	30%	19%
両方利用なし	5%	19%

図8に世代ごとに両判定を使用した被験者の割合を示す。図8より、「どちらも好き」判定は、各世代ともに全被験者の40%程度が使用していることを確認した。「どちらも嫌い」判定は、初期世代では全被験者の30%程度が使用しているが、世代交代を経るに連れて、10%程度に低下している。しかし、世代交代の後半においても「どちらも嫌い」判定が行われているため、必ずしも満足のいくデザインを生成できなかった被験者が一定数いることが示唆された。

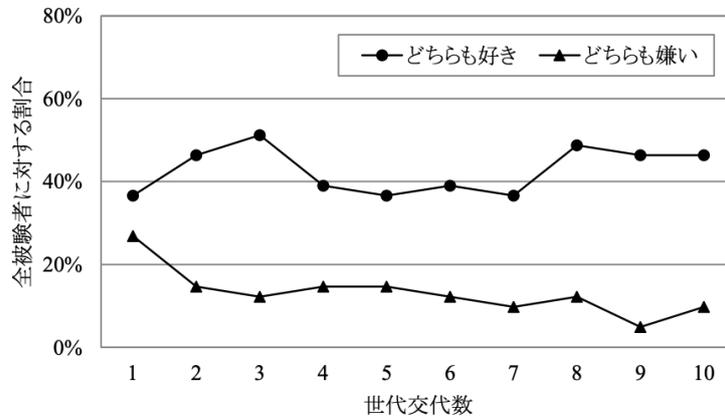


図8 「どちらも好き」「どちらも嫌い」判定時のデザイン間の関連

### 4.3 評価に要した時間の結果

本節では、被験者が両システムを利用して、好みの腕時計デザインを作成するためにどの程度の時間を要したかについて述べる。評価時間はIECシステムにおいてユーザの評価負担を軽減するためには、重要な指標である。本実験における評価時間は、被験者に最初の対戦デザインが提示されてから最終対戦の評価を終了するまでに要した時間としている。

図9に世代ごとの評価時間の結果を示す。両システムとも、初期世代の評価に要した時間が長く、その後世代交代を経るに連れて各世代の評価時間は短くなる傾向が確認された。初期世代の評価時間が長い理由には、被験者がシステムの操作に慣れていない、システム利用初期であったため提示されたデザインを慎重に評価したことが考えられる。両システムの10世代分の平均評価時間は、提案システムAが182.86[s]、提案システムBが156.96[s]であり、提案システムBの方が短かった。この結果について統計的有意差を確認したところ、有意水準5%で統計的有意差を確認した。

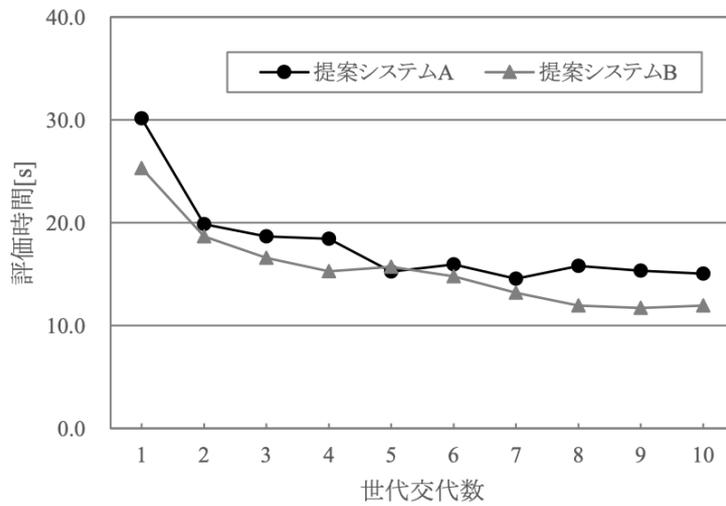


図9 世代別の評価時間の結果

次に、本実験では、提案システムにおいて、「どちらも好き」「どちらも嫌い」判定が下された際にどの程度の時間を要したかについて考察した。図10に判定方法の違いによる評価時間の結果を示す。図10において、「通常評価」および「提案システムB」は、被験者が提示された2つのデザインについて二者択一の評価を行った場合を示す。図10より、「どちらも好き」「どちらも嫌い」判定が使用されたときの評価時間は、4世代目を除いて、二者択一評価時より長くなっていることを確認した。これは、提案システムAでは、被験者にとって評価選択肢が増加した分、どのような評価を下そうか被験者が迷ったためと考えられる。このため、提案システムAでは、提案システムBよりも評価時間が長くなったと考えられる。

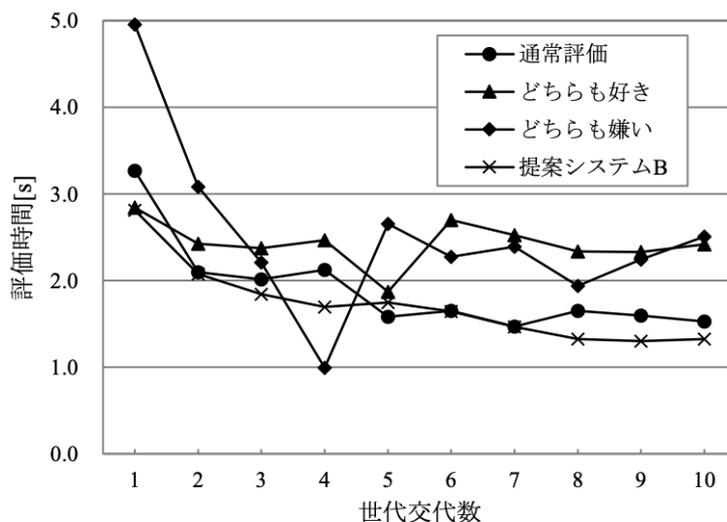


図10 対戦ごとの評価時間の結果

#### 4.4 生成されたデザインに対する満足度の結果

本節では、両システムにおける各世代のエリートデザインの満足度に関する検証結果を述べる。IECシステムにおいて、世代交代をある程度経過した後の腕時計デザインがユーザの満足いくものになっているかどうかは、IECシステムの進化性能を評価する上で重要である。

図11に生成されたデザインに対する平均満足度の結果を示す。図11より、両システムの生成されたデザインに対する満足度は、世代を経るごとに増加し、解候補が進化する傾向を確認した。しかし、生成されたデザインに対する満足度は、提案システムAより提案システムBの方が高くなっている。これは、提案システムAでは、2つのデザインに対して被験者が「どちらも好き」「どちらも嫌い」の評価ができることに起因していると考えられる。これらの評価が複数回行われた場合、デザイン間の評価値に差が付きにくく、GA処理において解候補群が進化しにくくなることが考えられる。このため、生成されたデザインに対する満足度は、提案システムAの方が低くなったと考えられる。しかし、両システムの生成されたデザインに対する満足度は、5段階評価において4点を超えている。したがって、両システムともに被験者が満足のいくデザインを生成できたといえる。

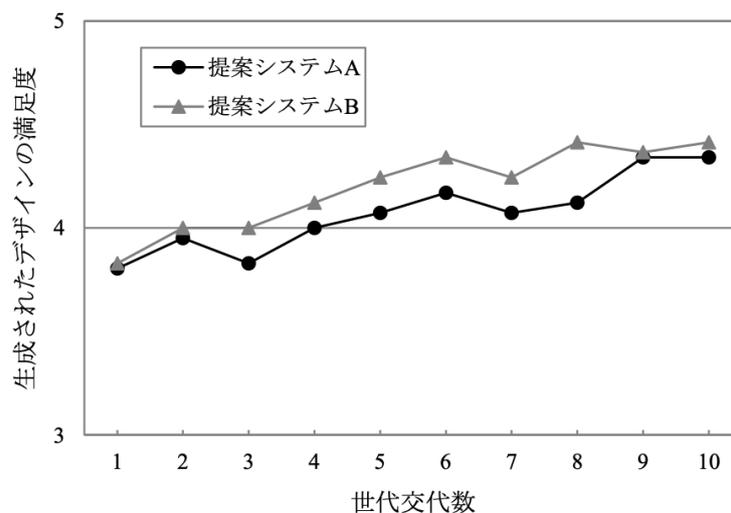


図11 生成されたデザインに対する満足度の結果

図12に、最終世代に生成された全エリートデザインを示す。図12中で各デザイン下の△および○は、それぞれ被験者が3点以下を付けたデザイン、文字盤が白地のデザインを示している。図12より、男性被験者が生成したデザインには、文字盤デザインが白地のものが全40デザイン中22デザインであった。女性被験者が生成したデザインにも文字盤デザインが白地のものが確認できるが、全42デザイン中15デザインであり、男性被験者と比較すると少ないことが確認された。女性被験者の場合は、白地の文字盤だけでなく、黒や青、茶系などの文字盤デザインも多いことも確認された。



(a) 提案システムA



(b) 提案システムB

図12 全エリートデザインの結果

## 4.5 アンケート回答に関する結果

### 4.5.1 評価のしやすさ

本実験では、被験者に2つのシステムの使用後、提案システムAとBのどちらのシステムが腕時計デザインを評価しやすかったかアンケートに答えてもらった。この結果、提案システムAが評価しやすかったと答えた被験者は全体の61%、提案システムBが評価しやすかったと答えた被験者は全体の39%であった。

提案システムAでは、被験者が評価に迷った際に「どちらも好き」「どちらも嫌い」の判定ができる反面、評価時の選択肢が増え、操作に煩わしさを感じる被験者がいるとも考えられる。このため、約4割の被験者はよりシンプルに評価できる提案システムBを支持したと考えられる。

しかし、被験者のコメントからは「迷ったときには「どちらも好き」「どちらも嫌い」判定があったため、提案システムAの方が評価しやすかった」という指摘があり、「どちらも好き」「どちらも嫌い」の判定機能はユーザの評価負担軽減に有用であったことが示唆される。

#### 4.5.2 その他のアンケート項目

本項では、被験者に対して行った提案システムの使用に関するアンケート結果について考察する。

図13に項目aからdの回答結果を示す。図13より、項目aであるシステム利用時に重要視したデザインパーツについては、両システムともに文字盤パーツの割合が最も大きいことを確認した。これは、腕時計デザインにおいて、文字盤パーツが全体の視覚的な印象を大きく左右するためであると考えられる。

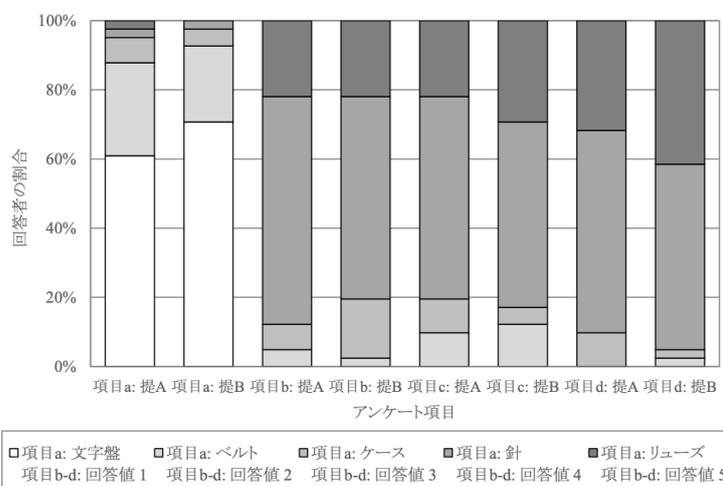


図13 アンケート結果

項目bでは、両システムとも全体の80%以上の被験者が好みの腕時計デザインが生成されたと回答している。これは、4.4節における解候補の進化が被験者の使用感においても体感されたことを示唆している。

項目cでは、両システムとも全体の80%以上の被験者が腕時計デザインを選択するごとに、似たような腕時計デザインが表示されたと回答している。このため、両システムとも世代を経るごとに、ユーザの嗜好にあった腕時計デザインを複数生成できたと考えられる。

項目dでは、両システムとも全体の90%以上の被験者が腕時計デザインを選びやすかったと回答している。これは、提案システムAは提案システムBと同様に、被験者が評価しやすいインタフェースであり、被験者の評価負担が軽減されることを示唆していると考えられる。

すべての結果から、提案システムAでは、被験者が好みの腕時計デザインを評価する時間は、提案システムBに比べ約25[s]長かった。しかし、腕時計デザインを評価する際に、提案システムAの方が提案システムBよりも、評価がしやすいと答える被験者が全体の60%以上であることを確認した。また、提案システムAでは提案システムBと同様に、被験者が満足する腕時計デザインを生成できることを確認した。したがって、提案システムAでは、腕時計デザインを評価する選択肢を提案システムBの2個から4個に増やすことで、ユーザの満足のいく腕時計デザインが生成でき、ユーザの評価負担軽減にも有用であることを確認した。

---

## 5. 得られた知見と今後の展望

---

本研究では、IECの解候補評価負担軽減のため、WPCを用いた女性用腕時計デザインシステムを提案した。提案システムでは、一対比較評価に対して、ユーザが提示されたデザインのうち、「どちらも好き」「どちらも嫌い」の評価をできる機能を組み込んだ検証を行った。実験結果より、評価時間においては、「どちらも好き」「どちらも嫌い」判定がない提案システムBの方が提案システムAよりも短くなることを確認した。しかし、提案システムAは提案システムBと同等の進化性能を示し、ユーザのデザイン評価をしやすくできることも確認した。これらの結果より、一対比較評価および「どちらも好き」「どちらも嫌い」判定をIECシステムに用いることは、ユーザが手軽に商品カスタマイズをするような場面では有用であることを示唆していると考えられる。今後は、ユーザの解候補評価の過程における傾向などを調査し、提案システムの応用に関する検証を実施していく予定である。

### 参考文献

- 1) Knot公式サイト：<https://knot-designs.com/> (2020.11.3閲覧)
- 2) New Balance公式サイト：<https://shop.newbalance.jp/shop/> (2020-11-03閲覧)
- 3) Takagi, H. : Interactive Evolutionary Computation : Fusion of the Capabilities of EC Optimization and Human Evaluation, Proceedings of the IEEE, Vol.89, No.9, pp.1275–1296 (2001).
- 4) Ishibashi, K. and Miyata, K. : Statistics–Based Interactive Evolutionary Computation for Color Scheme Search, International Journal of Affective Engineering, Vol.14, No.1, pp.31–41 (2015).
- 5) 中村充志, 瀧澤 生, 星 泰成, 陳キユウ : 画像の感性を反映させたフォントの自動生成手法, 日本感性工学会論文誌, Vol.17, No.5, pp.523–529 (2018).
- 6) Takenouchi, H., Tokumaru, M. and Muranaka, N. : Tournament-style Evaluation Using Kansei Evaluation, International Journal of Affective Engineering, Vol.12, No.3, pp.395–407 (2013).
- 7) Watanabe, Y., Yoshikawa, T. and Furuhashi, T. : A Study on Application of Fitness Inference Method to PC-IGA, 2007 IEEE Congress on Evolutionary Computation (CEC 2007), pp.1450–1455 (2007).
- 8) Takagi, H. and Pallez, D. : Paired Comparisons-based Interactive Differential Evolution, World Congress on Nature and Biologically Inspired Computing (NaBIC'09), pp.475–480 (2009).
- 9) Takenouchi, H., Tokumaru, M. and Muranaka, N. : Interactive Evolutionary Computation Using a Tabu Search Algorithm, IEICE Transactions on Information and System, Vol.E96-D, No.3, pp.673–680 (2013).



竹之内宏（非会員）h-takenouchi@fit.ac.jp

2008年関西大学工学部電子工学科卒業，2010年関西大学大学院工学研究科博士課程前期課程修了，2013年同大学院理工学研究科博士課程後期課程修了，博士（工学）。同年同大学非常勤研究員，ポスドクトラルフェローを経て，2014年より福岡工業大学情報工学部システムマネジメント学科助教。現在に至る，対話型進化計算の応用システムに関する研究，ファジィ推論を用いた感性検索システムに関する研究に従事。日本知能情報ファジィ学会，日本感性工学会，電子情報通信学会，IEEE各会員。



徳丸正孝（非会員）toku@kansai-u.ac.jp

1995年関西大学工学部電子工学科卒業，1997年同大大学院工学研究科博士課程前期課程修了。同年同大工学部電子工学科助手，2016年同大システム理工学部電気電子情報工学科教授。工博，感性情報処理，対話型進化計算の応用システムの研究に従事。人工知能学会，電子情報通信学会，日本知能情報ファジィ学会，日本感性工学会，ヒューマンインタフェース学会，IEEE各会員。

採録決定：2020年12月4日

編集担当：佐藤 裕一（（株）富士通研究所）