

ガウス過程回帰に基づく個人的・集団的嗜好分布の効率的な推定

開発功太郎[†] 松原 正樹^{††} 伊藤 寛祥^{††} 森嶋 厚行^{††}

[†] 筑波大学情報学群情報メディア創成学類 〒 305-8550 茨城県つくば市春日 1-2

^{††} 筑波大学図書館情報メディア系 〒 305-8550 茨城県つくば市春日 1-2

E-mail: [†]kotaro.kaihotsu.2021b@mlab.info, ^{††}{masaki,ito,mori}@slis.tsukuba.ac.jp

あらまし マーケティングにおいて、ある個人がどのようなアイテムを好むか、あるアイテムを好む集団がどのくらい存在するかを把握することは重要である。特に音楽・美術・デザインなどのドメインにおいては、文化的背景によって評価が一致するとは限らず、集団の好みは多様となることが多い。また、アイテム数の増加により必要な回答数が増えてしまう問題もある。そこで本研究では、個人と集団の嗜好分布をガウス過程回帰によってモデル化し、多様な好みをより効率よく推定する手法を提案する。アイテムに対する評価値が何も与えられていない状況から、評価者が提示アイテムに対して逐次評価を与えることで対話的に嗜好分布を推定する問題を想定し、個人の嗜好分布を推定する際に集団の嗜好分布の推定結果を考慮したアイテムの提示を行うことで、より効果的な評価データのサンプリング戦略を実現する。主観評価タスクを対象とするシミュレーション実験の結果、ベースライン手法のUCB方策と比較して、提案手法は少ない回答数で評価者の好むアイテムを推定することが示された。

キーワード ガウス過程回帰, クラウドソーシング, 主観評価, 社会的価値観

1 はじめに

今日、個人や集団の嗜好分布を把握することはマーケティングにおいて価値がある。ある人がどのようなアイテムを好むか、あるアイテムを好む人達がどの程度存在するかが理解でき、それによりどのような製品に需要があるかやどのような人たちにどのような広告を提示することが効果的であるかを把握できるためである。

現在、個人的・集団的嗜好分布を把握するとき、アンケートやABテストのような大規模な事前調査結果を分析するデータマイニング的アプローチが取られることがあるが、提示するアイテム数の増加により必要な回答数が増えるにつれて、回答者が正しく評価できない、負担が高くなる、主観評価の信頼性や品質に問題が生じてしまう。そこで、多腕バンディット問題に形式化して解決するなど、対話的で動的なアイテム提示のサンプリング手法が用いられている。

一方、音楽・美術・デザインなどの人間の主観が介入するドメインにおいては、個人の価値観や文化的背景によって必ずしも評価が一致するとは限らず、集団の好みは多様となることがある。そのため、クラウドソーシングやマーケティングなどで主観評価の結果を統合する際には、単純な多数決ではなく、回答者がどのような好みや背景の基づいて回答しているのか等の情報を考慮しながら、集団の評価の分布を推定する必要がある [1], [2].

これらの問題に対処するため、本研究では、集団の嗜好分布は多様ではあるがいくつかのグループに類型化することができるという仮定をおき、対話的に個人の嗜好分布を推定する過程において、既に回答した他の評価者たちの情報を用いることでより効果的な手法を提案する (図 1)。提案手法により、集団の

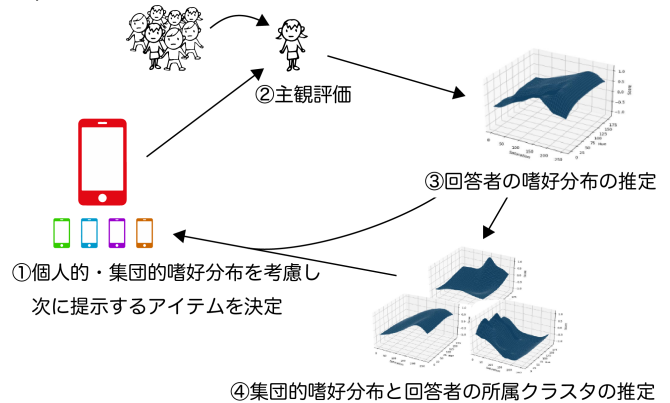


図 1: 提案手法の全体像: 個人的・集団的嗜好分布を考慮し次に提示するアイテムを決定する。回答者の主観評価をもとにガウス過程回帰により個人的嗜好分布と集団的嗜好分布を推定する。

多様な嗜好分布を推定しながら、新たな評価者に対して問い合わせを行う時にサンプル数を増やすに従い漸近的に解に近づくこと (RQ1), 評価者から高い評価をもらい続けること (RQ2) が期待される。

検証にあたり、ベースラインには動的サンプリング手法として広く用いられている、UCB 方策を採用し、シミュレーション実験と実際の評価者を用いた実験を行った。実験の結果、提案手法はベースラインよりも少ない回答数で評価者の好むアイテムを推定することが示された。

本論文の構成は以下である。2章で関連研究について説明を行い、3章で既知の評価者情報を用いて個人的・集団的嗜好を効率よく推定する手法を提案する。4章ではシミュレーション実験と実際の評価者を対象とした実験とそれぞれの結果について示す。最後に、5章において実験に対する考察を行い、6章にて本論文をまとめる。

$$\mathbf{f} = (f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_N)) \quad (1)$$

2 関連研究

2.1 多腕バンディット問題

個人的嗜好分布を予測する問題は提示するアイテムの集合をアームと捉えることで多腕バンディット問題 [3] として扱うことができ、広告配信システムなどで用いられている [4]。多腕バンディット問題とは複数のアームと呼ばれる候補の中から最も良いものを逐次的に探す問題のことである。多腕バンディット問題には累積報酬最大化や最適腕識別という 2 つの大きな枠組みがあるが、本研究では累積報酬最大化を扱う。累積報酬最大化のためのアルゴリズムでは UCB 方策やトンプソンサンプリングアルゴリズム [5] が知られている。

2.2 群衆の嗜好分布が多様であることを考慮した研究

Baba ら [6] はクラウドソーシングによって集めた評価を、人がそれぞれ異なる評価軸に基づいていることを考慮し分析する手法を提示した。この研究ではフロンティアアイデアと呼ばれる少なくとも 1 つの評価観点においては最も評価されるアイデアを特定することができる。この手法はあらかじめ全探索的に一対比較法で集めたデータを分析しているため多くの問い合わせコストが発生してしまう。一方、本手法では評価者からの評価が高いと予想されるアイテムを提示することで少ない回答数で嗜好分布を把握できることが期待される。

2.3 評価者との対話により好ましいアイテムを特定する研究

小山ら [7] は多次元の嗜好分布を推定する問題において、多次元の嗜好分布の探索を人間がスライダー UI などとして認知しやすい一次元の問い合わせの連続に分解する手法を提示した。坂井ら [8] は対話的進化演算を用いて、群衆とシステムが対話を繰り返すことで好まれるデザインを生成するシステムを提案した。これらの研究は多次元の人間の嗜好分布をシステムとの対話により推定する点で共通しているが、人の嗜好がそれぞれ違うことを考慮していない。Kochanski ら [9] はベイズ最適化を用いて対話的に提示とフィードバックを繰り返すことでクッキーの味を最適化することができた。この研究では都市によって好まれるクッキーの味が異なる可能性が示唆されている。Brochu らは [10] は 3D グラフィックスの最適なテクスチャの探索のためにベイズ最適化を用いた。この研究では 3D グラフィックスの訓練を受けていない評価者でも、提示された 2 つの画像から選ぶことで密に関わる多数のパラメータを調整することができた。この研究ではランダムサンプリングに比べ半分の回答数で目的の質感を得ることができている。

3 前提知識

この章では、本手法で用いているガウス過程回帰について説明する。

ガウス過程回帰は入力変数 x から出力変数である実数値 y への関数 $y = f(x)$ を推定するモデルである [11]。 N 個のデータ $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N \in \mathcal{X}$ に対応する出力のベクトル

が平均 $\boldsymbol{\mu} = (\mu(\mathbf{x}_1), \mu(\mathbf{x}_2), \dots, \mu(\mathbf{x}_N))$, $K_{nn'} = k(\mathbf{x}_n, \mathbf{x}_{n'})$ を要素とする行列 \mathbf{K} を共分散行列とするガウス分布 $\mathcal{N}(\boldsymbol{\mu}, \mathbf{K})$ に従うとき、 f はガウス過程に従うといい、これを

$$f \sim GP(\boldsymbol{\mu}(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (2)$$

と書く [12]。

N 個の観測値 $\mathcal{D} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\}$ が与えられているとする。ここで y は平均が 0 になるように正規化されているとする。ここで、 \mathbf{x} と y の間には $y = f(\mathbf{x})$ の関係があり、この関数 f が平均 $\mathbf{0}$ のガウス過程 $f \sim GP(\mathbf{0}, k(\mathbf{x}, \mathbf{x}'))$ から生成されているとする。 $\mathbf{y} = (y_1, y_2, \dots, y_N)^T$ の時、 \mathbf{y} はガウス分布に従い、入力のすべてのペア $(\mathbf{x}_n, \mathbf{x}_{n'})$ についてカーネル関数 $k(\mathbf{x}, \mathbf{x}')$ をもちいて

$$K_{nn'} = k(\mathbf{x}_n, \mathbf{x}_{n'}) \quad (3)$$

で与えられるカーネル行列 \mathbf{K} をつかって

$$\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}) \quad (4)$$

が成り立つ。本研究では、カーネル関数として以下で定義される RBF カーネルを用いる。

$$k(\mathbf{x}, \mathbf{x}') = \theta_1 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{\theta_2}\right) \quad (5)$$

ここで、 θ_1 と θ_2 はハイパーパラメータである。

観測値 \mathcal{D} が与えられたとき、未知のデータ \mathbf{x}^* に対応する y^* の予測分布 $p(y^* | \mathbf{x}^*, \mathcal{D})$ は

$$p(y^* | \mathbf{x}^*, \mathcal{D}) = \mathcal{N}(k_* \mathbf{K}^{-1} \mathbf{y}, k_{**} - k_*^T \mathbf{K}^{-1} k_*) \quad (6)$$

ここで

$$\begin{cases} \mathbf{k}_* = (k(\mathbf{x}^*, \mathbf{x}_1), k(\mathbf{x}^*, \mathbf{x}_2), \dots, k(\mathbf{x}^*, \mathbf{x}_N))^T, \\ k_{**} = k(\mathbf{x}^*, \mathbf{x}^*) \end{cases} \quad (7)$$

である。予測分布の期待値と分散はそれぞれ

$$\mathbb{E}[y^* | \mathbf{x}^*, \mathcal{D}] = \mathbf{k}_* \mathbf{K}^{-1} \mathbf{y} \quad (8)$$

$$\mathbb{V}[y^* | \mathbf{x}^*, \mathcal{D}] = k_{**} - \mathbf{k}_*^T \mathbf{K}^{-1} \mathbf{k}_* \quad (9)$$

である。

4 問題定義

本研究の目的は、評価者集団のクラスタと各評価者の嗜好分布をより少ない評価回数で推定することである。これにより、目標とする評価者集団において嗜好が類似する部分集団の推定や、予測嗜好分布に基づく各アイテムの特徴量の最適化を、より少ない回答回数で求めるといったアプリケーションが実現できる。

入力は評価者の集合 \mathcal{W} 、アイテムの集合 \mathcal{U} 、設定するクラスタの数 n_c である。各アイテムはそれぞれ n 次元の特徴量

表 1: アイテム提示戦略の変数関数定義

変数	変数の説明
\mathcal{W}	評価者集合 $\{w_1 \dots w_{ \mathcal{W} }\}$
\mathcal{U}	アイテム集合 $\{u_1 \dots u_{ \mathcal{U} }\}$
\mathcal{C}	クラスタの集合 $\{C_1 \dots C_{n_c}\}$
$\mathcal{D} \subseteq \mathcal{W} \times \mathcal{U} \times \mathbb{R}$	評価者のアイテムに対する評価データ集合
n_c	クラスタ数
$\mathbf{x} \in \mathbb{R}^k$	各アイテムの特徴量
$y \in \mathbb{R}$	各アイテムの評価値
$\mathbf{r}_w \in \{0, 1\}^{n_c}$	各評価者の所属クラスタ
$\mu_k : \mathbb{R}^k \rightarrow \mathbb{R}$	各クラスタのセントロイド関数
$acquisit : \mathbb{R}^k \times \mathcal{W} \rightarrow \mathbb{R}$	評価者ごとのアイテムに対する獲得関数
N_R	何名の評価者が終了する度に クラスタを再計算するか
T	一人の評価者からの回答回数

を持ち, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_u} \in \mathbb{R}^n$ とかく. 例えば, 各アイテムの写真の色の RGB 値あるいは HSL 値, 音声を録音するときのマイクの座標と角度などが特徴の例として挙げられる. 評価者から与えられる評価データは, 評価者とアイテムと評価値のタプルとし, 評価データ全体の集合を \mathcal{D} とする. ここで, $(w, u, y) \in \mathcal{D} \subseteq \mathcal{W} \times \mathcal{U} \times \mathbb{R}$ である.

出力は, 評価データ集合 \mathcal{D} が与えられたときの各評価者の予測嗜好分布 $\{p(y^*|\mathbf{x}^*, \mathcal{D}, w)\}_{w \in \mathcal{W}}$, 評価者のクラスタ $\{C_k \in \mathcal{P}(\mathcal{W})\}_{k \in [n_c]}$ および各クラスタの予測嗜好分布 $\{p(y^*|\mathbf{x}^*, \mathcal{D}, C_k)\}_{k \in [n_c]}$ である. ここで, 各評価者および各クラスタの予測嗜好分布はガウス過程回帰モデルで表現されるものとする.

本研究で用いる記号の定義を表 1 に示す.

5 提案手法

提案手法の基本的なアイデアは, 個人の嗜好分布の推定と同時に評価者のクラスタの嗜好分布を推定し, 各個人がどのクラスタに属するかという情報を考慮したアイテムの提示を行うことで, より少ない回答数で推定を行うというものである.

提案手法は大きく 2 つの部分から構成される. 1 つ目は評価者にアイテムを提示する部分である. アイテムの提示では, 評価者の嗜好分布のうち未知の部分の情報を得るための探索と, 評価者が好むアイテムを提示する活用のバランスを考慮しながら行う. 2 つ目は回答者から得たフィードバックをもとに評価者の嗜好分布を予想し更新する部分である. このとき評価者の各アイテムへの評価の期待値と分散, 評価者の所属クラスタ, 各クラスタの各アイテムへの評価の期待値と分散を再計算する.

5.1 評価者のクラスタの推定とクラスタの嗜好分布の推定

本研究では評価者のクラスタを, 予測嗜好分布をクラスタリングすることで求める. ここで, 評価者のクラスタは嗜好の期待値の特徴空間全域における類似度に基づいて構成されると仮定する. これらより, クラスタリングの最適化問題は以下になる.

$$\{\mathbf{r}_w\}_{w \in \mathcal{W}}, \{\mu_k\}_{k \in [n_c]} = \arg \min_{\{\mathbf{r}_w\}, \{\mu_k\}} J \quad (10)$$

ここで, 目的関数 J は以下である.

$$J = \sum_{w \in \mathcal{W}} \sum_{k \in [n_c]} r_{wk} \int \|\mathbb{E}[y|\mathbf{x}, \mathcal{D}, w] - \mu_k(\mathbf{x})\|_2^2 d\mathbf{x} \quad (11)$$

関数 $\mu_k : \mathbb{R}^k \rightarrow \mathbb{R}$ はクラスタのセントロイドを意味する. $\mathbf{r}_w = (r_{w1}, \dots, r_{wk})^\top$ は評価者 w がどのクラスタに属するかを表現するベクトルで, 各要素 r_{wk} は以下のように表される.

$$r_{wk} = \begin{cases} 1 & k = \arg \min_k \int \|\mathbb{E}[y|\mathbf{x}, \mathcal{D}, w] - \mu_k(\mathbf{x})\|_2^2 d\mathbf{x} \text{ のとき} \\ 0 & \text{それ以外.} \end{cases} \quad (12)$$

本研究では上記のクラスタリング問題を k-means アルゴリズムにより最適化する. 各クラスタ C_k は以下のように表される.

$$C_k = \{w \in \mathcal{W} \mid r_{wk} = 1\} \quad (13)$$

クラスタの嗜好分布も個人の嗜好分布と同様にガウス過程回帰モデルで表現する. このとき, クラスタ C_k の嗜好分布 $p(y^*|\mathbf{x}^*, \mathcal{D}, C_k)$ の期待値は以下のように求める.

$$\mathbb{E}[y^*|\mathbf{x}^*, \mathcal{D}, C_k] = \frac{1}{|C_k|} \sum_{w \in C_k} \mathbb{E}[y^*|\mathbf{x}^*, \mathcal{D}, w] \quad (14)$$

5.2 アイテム提示戦略

評価者に提示するアイテムを選択するための戦略として, UCB 方策をベースとした戦略を提案する. この戦略では, 評価者の嗜好分布の期待値と分散の項からなる通常の UCB 戦略の獲得関数に加え, 評価者が属するクラスタの嗜好分布の期待値の項を加えた獲得関数により, 次に評価者に提示するアイテムを選択する. このような獲得関数を利用することで, 評価者から十分な回答数が得られていない場合でも, その評価者がどのクラスタに属するかという情報を考慮することで, より効果的なアイテムを提示することができる.

評価者 $w_i \in \mathcal{W}$ と特徴量 \mathbf{x} を持つアイテムが与えられたとき, そのアイテムが評価対象としてどれだけ適切かを返す獲得関数 $acquisit : \mathbb{R}^k \times \mathcal{W} \rightarrow \mathbb{R}$ は次式で表される.

$$acquisit(\mathbf{x}, w) = \mathbb{E}[y^*|\mathbf{x}^*, \mathcal{D}, w] + \alpha_1 \mathbb{V}[y^*|\mathbf{x}^*, \mathcal{D}, w] + \alpha_2 \sum_{k \in [n_c]} r_{wk} \mathbb{E}[y^*|\mathbf{x}^*, \mathcal{D}, C_k] \quad (15)$$

ここで, α_1, α_2 はハイパーパラメータで, この重みを調整することで, 探索と活用のバランスを調整することができる. 本研究では回答者となる評価者は与えられると仮定し, 上記の $acquisit$ 関数を最大化する特徴量をもつアイテムを評価者に提示する. すなわち, 以下のアイテム u^* を提示する.

$$u^* = \arg \min_{u \in \mathcal{U}} acquisit(\mathbf{x}_u, w) \quad (16)$$

手法の全体の流れを Algorithm1 に示す. 各評価者について, データ点の評価と予測分布の再計算を行うステップと, 評価者

Algorithm 1 Algorithm

Input: $\mathcal{W}, \mathcal{U}, n_c$ **Output:** $\{p(y^*|\mathbf{x}^*, \mathcal{D}, w)\}_{w \in \mathcal{W}}, \{p(y^*|\mathbf{x}^*, \mathcal{D}, C_k)\}_{k \in [n_c]}, \{C_k\}_{k \in [n_c]}$

```
1:  $\mathcal{D} \leftarrow \emptyset$ 
2: for  $t$  in  $[T]$  do
3:   for  $w \in \mathcal{W}$  do
4:      $u^* = \arg \max_{u \in \mathcal{U}} \text{acquisit}(\mathbf{x}_u, w)$ 
5:      $y_w^{u^*} \leftarrow$  評価者から受け取った評価
6:      $\mathcal{D} \leftarrow \mathcal{D} \cup \{(w, u^*, y_w^{u^*})\}$ 
7:      $\mathbb{E}[y^*|\mathbf{x}^*, \mathcal{D}, w], \mathbb{V}[y^*|\mathbf{x}^*, \mathcal{D}, w]$  を再計算
8:   end for
9:   if  $t \% N_R == 0$  then
10:     $\{\mathbf{r}_w\}_{w \in \mathcal{W}}, \{\boldsymbol{\mu}_k\}_{k \in [n_c]} = \arg \min_{\{\mathbf{r}_w\}, \{\boldsymbol{\mu}_k\}} J$ 
11:    for  $k$  in  $[n_c]$  do
12:       $\mathbb{E}[y^*|\mathbf{x}^*, \mathcal{D}, C_k]$  を再計算
13:    end for
14:  end if
15: end for
```

のクラスタリングを行うステップを適用する。ここで、計算量の削減のため、クラスタリングとクラスタの予測分布の再計算は一定人数が評価した際に行う。また、クラスタリングを行う際は、評価者の嗜好分布のすべての点ではなく、嗜好分布中に等間隔で設定したサンプリング点の期待値に基づいて計算する。

6 実験

本章では実データを用いたシミュレーションと、実際に評価者を用いた実験の双方により提案手法の有用性を確認する。どちらの実験でも指標は嗜好分布全体の誤差 (指標 1)、一番好む点の X 座標の誤差 (指標 2)、評価者からの評価の大きさ (指標 3) を設定した。これらの指標について詳しく説明する

6.1 指標

6.1.1 指標 1：嗜好分布全体の誤差

この指標では提案手法によって評価者の好みを全体を把握できているかを計測する。具体的には評価者の実際の嗜好分布と推定した嗜好分布の平均絶対誤差であり、以下の式で定義される。各評価者 w の t 回目時点での評価値の絶対誤差として以下を定義する。

$$AE(w, t) = \int \left| \hat{\mathbb{E}}[y^*|\mathbf{x}^*, \mathcal{D}_{all}, w] - \mathbb{E}[y^*|\mathbf{x}^*, \mathcal{D}^{(t)}, w] \right| d\mathbf{x}^* \quad (17)$$

さらに、すべての評価者の平均を以下に定義する。

$$MAE(t) = \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} AE(w, t) \quad (18)$$

ここで、 $\mathcal{D}_{all} \subseteq \mathcal{W} \times \mathcal{U} \times \mathbb{R}$ はアイテム全域について評価者から評価を受けた際の評価データの集合を意味し、 $\mathcal{D}^{(t)} \subseteq \mathcal{W} \times \mathcal{U} \times \mathbb{R}$ は、各評価者に対して t 回の評価を受けた段階での評価データ

の集合を意味する。本研究では、アイテムの特徴 \mathbf{x}^* について一定の間隔を設定し、その間隔に従ってデータをサンプリングすることで、上記の値を近似的に計算する。

6.1.2 指標 2：一番好む点の X 座標の誤差

この指標では評価者が一番好むアイテムを把握できているかを計測する。この指標はある評価者が実際に一番好むアイテムと手法が推定した評価者が一番好むアイテムのパラメータの値の間にどのくらい差があるかを表す。具体的にはアイテムのパラメータのユークリッド距離であり、以下の式で定義される。各評価者 w の t 回目時点での最も好むアイテムの評価値のパラメータのユークリッド距離として以下を定義する。

$$TAE(w, t) = \left\| \begin{aligned} & \arg \max_{\mathbf{x}^* \in \mathbb{R}^k} \hat{\mathbb{E}}[y^*|\mathbf{x}^*, \mathcal{D}_{all}, w] \\ & - \arg \max_{\mathbf{x}^* \in \mathbb{R}^k} \mathbb{E}[y^*|\mathbf{x}^*, \mathcal{D}^{(t)}, w] \end{aligned} \right\|_2 \quad (19)$$

さらに、すべての評価者の平均を以下に定義する。

$$MTAE(t) = \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} TAE(w, t) \quad (20)$$

6.1.3 指標 3：評価者からの評価の大きさ

この指標では評価者に対して好ましいアイテムを提示し続けることができるかを計測する。この指標は t 回目の評価において評価者 w から受けた評価の大きさであり、以下の式で定義される。

$$AS(w, t) = y_w^{(t)} \quad (21)$$

さらに、すべての評価者の平均を以下に定義する。

$$MAS(t) = \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} AS(w, t) \quad (22)$$

6.2 実験 1：シミュレーション

6.2.1 設定

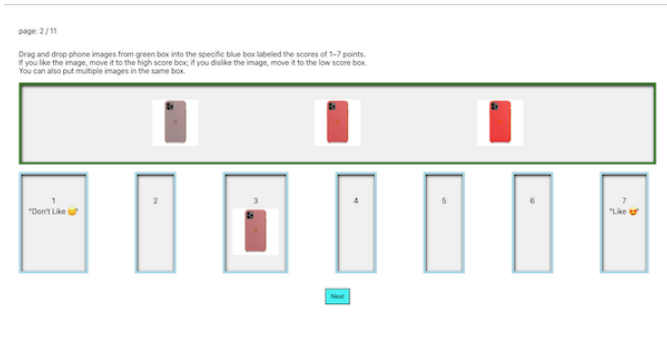
クラウドソーシングで主観評価の Ground Truth データを集めたあと、Ground Truth データの期待値分布に基づいて仮想的に回答を行う評価者 (以下、仮想回答者) によるシミュレーション実験を行った。評価対象物は色の異なるスマートフォンの画像であり、最適化を行うパラメータは携帯電話の色相と彩度である。色相は 180 段階、彩度は 255 段階に区切られており、それらをかけ合わせた 45,900 種類の画像が用意される。

Ground Truth データは以下の流れで作成した。

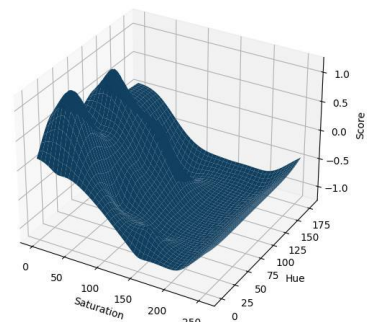
(1) 45,900 種類の画像の中から色相 8 種類、色相 4 種類を等間隔にとった 32 枚の画像を抽出した。

(2) 32 枚の画像を対象に図 2a の 1~7 点の得点を集めるタスクを作成した。このタスクでは表示された 4 枚の携帯電話の画像を得点がかかれた箱にドラッグアンドドロップする作業を 8 回繰り返す。

(3) クラウドソーシングプラットフォームである Amazon Mechanical Turk で 300 人の評価者を募った。評価者はオンラインで上記タスクにより嗜好を入力した。その後すべての画像

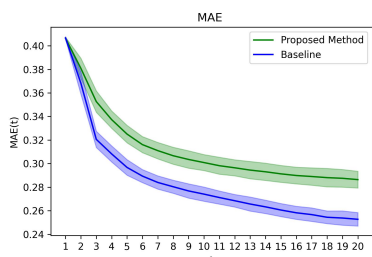


(a) 主観評価タスクのインターフェース。上部の緑色の枠内のアイテムを下部の青い枠の箱に移動させて 1 から 7 点の得点をつける。

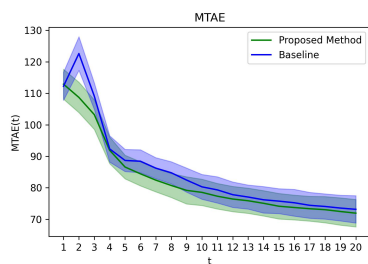


(b) 推定された嗜好の期待値分布の例。この評価者は彩度が低い、暗い色を好むことがわかる。

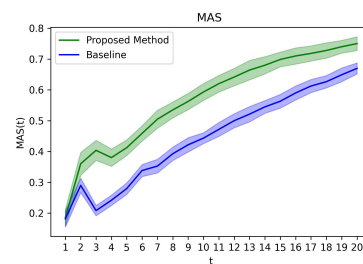
図 2: 実験 1: シミュレーションのための主観評価 Ground Truth データの収集



(a) $MAE(t)$: 嗜好分布全体の二乗誤差。ベースライン手法の方が誤差の収束が少し早い。



(b) $MTAE(t)$: 推定された一番好むアイテムと Ground Truth との距離。提案手法の方が少ない回答で好むアイテムを見つけられている。



(c) $MAS(t)$: 評価者からの評価値の回答ごとの変化。

図 3: 実験 1: シミュレーションにおける 3 つの指標での比較

に対し等しい評価をした者 3 名を省き、297 人のデータを実験のために利用した。

(4) 集めた回答の得点の 1~7 を -1~1 の範囲に変換しガウス過程回帰を用いて、評価者の嗜好分布を推定した。図 2b は推定された嗜好の期待値分布の例である。

一人の仮想回答者からの回答数 T は 20 回、クラスタリングは N_R は 1 人が回答を完了するごとに毎回実行されるように設定した。シミュレーションでは事前分布が与えられることの有効性を確認するため、はじめの 50 人はクラスタの形成にのみ用いて評価には含まなかった。獲得関数におけるクラスタの期待値に対する係数 α_1, α_2 は 1 に設定する。

6.2.2 結果

図 3a, 3b, 3c で結果を示す。すべての図において値は 50 回繰り返し行ったシミュレーション結果の平均値である。塗られた範囲はシミュレーション毎の標準偏差である。横軸は評価者にとって何回目の回答かを表している。

指標 1: $MAE(t)$ の結果は図 3a である。20 回の回答終了時に提案手法とベースライン手法の間に大きな誤差の違いは見られない。しかし、ベースライン手法のほうがやや収束が早いと言える。指標 2: $MTAE(t)$ の結果は図 3b である。提案手法のほうが全体を通してベースライン手法よりも低い誤差を保っている。指標 3: $MAS(t)$ の結果は図 3c である。提案手法のほうがはじめは高いスコアを得ている事がわかる。一方で 6 回目の回答あたりからベースライン手法のほうがやや高い評価を得てい

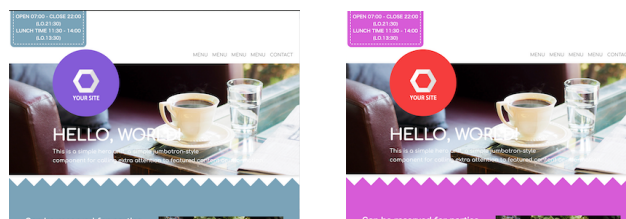


図 4: 最適化される WEB デザイン。色相と彩度のパラメータによりサイトのメインカラーとアクセントカラーが決定される。

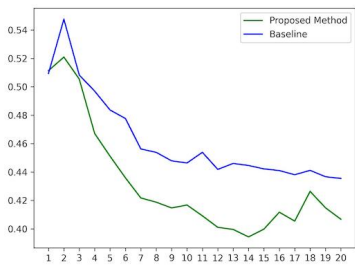
ることがわかる。

6.3 実験 2: 評価者を対象とした実験

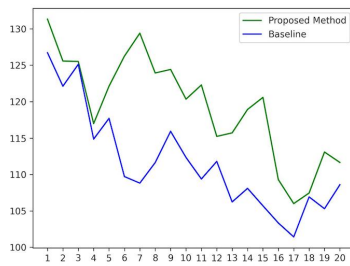
6.3.1 設定

実験 2 では実世界の評価者を対象に実験を行った。評価対象物はデザイン最適化が広く用いられている [13] [14] [15] [16], WEB サイトのデザインである¹。図 4 のようにサイトのメインカラーとアクセントカラーが変化する。最適化を行うパラメータはメインカラーの色相と彩度である。アクセントカラーの色相と彩度はメインカラーと比べ 60° 大きく 30% 鮮やかになるようになっている。シミュレーション実験のときと同じく、色相は 180 段階、彩度は 255 段階に区切られており、それらをかけ合わせた 45,900 種類のデザインが用意される。

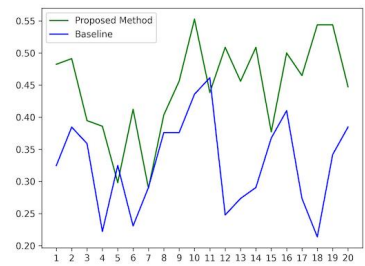
1: Web ページのテンプレートは <https://free-hp.net/> で配布されているものを利用した



(a) $MAE(t)$: 嗜好分布全体の二乗誤差. 提案手法のほうが全体的に早く収束している.

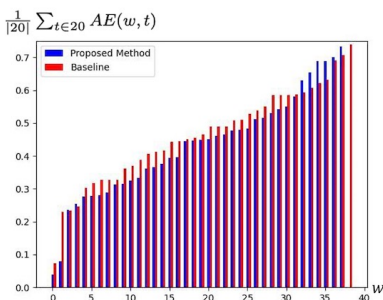


(b) $MTAE(t)$: 推定された一番好むアイテムと Ground Truth との色相と彩度のユークリッド距離.

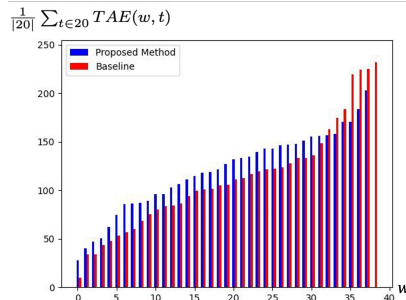


(c) $MAS(t)$: 評価値の回答ごとの変化. 提案手法の方が良い評価を受けている.

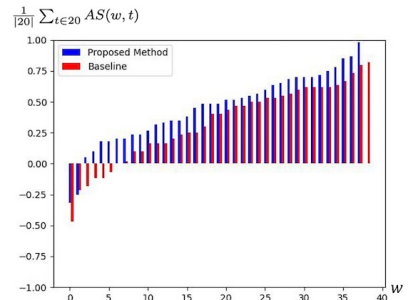
図 5: 実験 2: 実世界の評価者を対象とした実験における 3 つの指標での比較



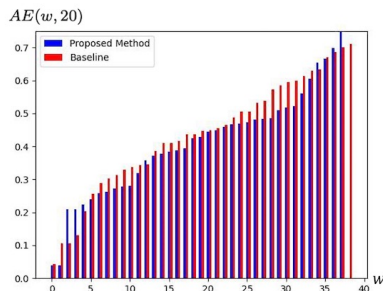
(a) 嗜好分布全体の二乗誤差 (20 回の回答の平均値)



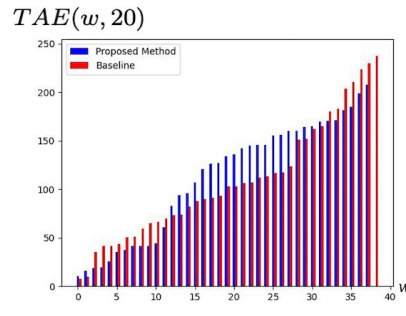
(b) 一番好むアイテムパラメータの Ground Truth との距離 (20 回の回答の平均値)



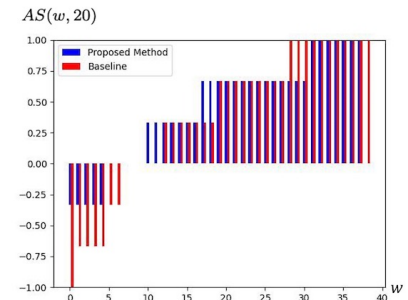
(c) 評価値 (20 回の回答の平均値). 提案手法の方がより高い評価平均値を獲得している.



(d) 嗜好分布全体の二乗誤差 (20 回目の回答の値)



(e) 一番好むアイテムパラメータの Ground Truth との距離 (20 回目の回答の値)



(f) 評価値 (20 回目の回答の値)

図 6: 実験 2 の 3 つの指標における評価者の個人差: 上段は 20 回の回答の平均値, 下段は 20 回目の回答時の評価値. 指標の値が小さい順に評価者 (X 軸) をソートしている.

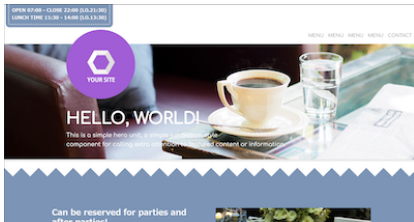
評価者は既存手法と提案手法の実験のそれぞれで 40 人を Amazon Mechanical Turk で募った. 評価者はオンラインで実験 1 と同様なインタフェースにより提示された Web デザインに対して 1~7 点で得点を入力する作業を 20 回繰り返す. 獲得関数におけるクラスタの期待値に対する係数 α_2 は, 各評価者に対する 2 回目の問い合わせまでは $\alpha_2 = 3$ に設定し, 3 回目の問い合わせ以降は $\alpha_2 = 0$ に設定する. その後, 実験 1 と同様に評価者の Ground Truth の嗜好分布を求めた. 45,900 種類のデザインの中から色相 6 種類, 色相 4 種類を等間隔に抽出した 24 枚の画像に対して 1~7 点で得点を入力し, 評価データを正規化しガウス過程回帰によって計算した.

はじめの 20 回すべての提示に同じ評価をした者をスパムと

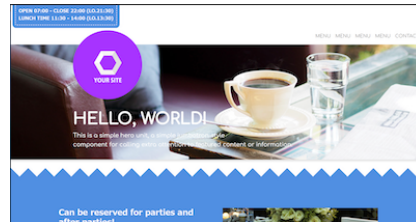
して除外し, その後の分析ではベースライン手法では 39 人, 提案手法では 38 人のデータを利用した.

6.3.2 結果

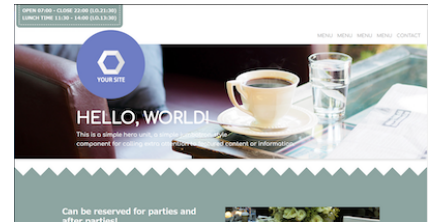
全評価者における比較: 結果は図 5a, 5b, 5c の通りである. すべての図において値は評価者全員の平均になっている. 横軸は評価者にとって何回目の回答であることを示している. ts はベースライン手法であるトンプソンサンプリングを, cts は提案手法を表している. 図 5a からは提案手法のほうが回答を通して低い誤差を保っている事がわかる. 図 5b からどちらの手法も最も好むアイテムへの誤差を評価の数が増えるのに従って減らすことができていることがわかる. また, ベースライン手法のほうが, 早い段階で制度の高い推定ができていることがわかる.



(a) 色相 107, 彩度 73. 属する評価者は 19 人.



(b) 色相 107, 彩度 183. 属する評価者は 10 人.



(c) 色相 85, 彩度 30. 属する評価者は 10 人.

図 7: 集団的嗜好推定で得られた各クラスタの代表アイテム

図 5c からは提案手法のほうが評価者から一貫して高い評価を得ることができていることがわかる。一方でどちらの手法も評価の数の増加と評価値の増加に相関は見られない。

評価者の個人差: 図 6a~6f はすべての評価者の 20 回の回答の平均, 20 回目の回答におけるそれぞれの指標の値である。評価者は指標の値が小さい順にソートしてある。図 6a, 6d からは提案手法のほうが, ほとんどの評価者の嗜好分布をより小さな誤差で予測できていることがわかる。図 6b, 6e からは指標 2 において, 提案手法のほうが多くの評価者では誤差が大きいことがわかる。一方でベースライン手法では一部の評価者の誤差が飛び抜けて大きくなる問題が見られる。図 6c, 6f からは提案手法のほうが, ほぼすべての評価者にとってより高い評価平均値を獲得していることがわかる。

推定された集団的嗜好: 図 7a~7b は実験の結果得られた 3 つのクラスタのセントロイドのアイテムである。色相は 0~179 の範囲でそれぞれ 107, 107, 85 であり, 彩度は 0~255 の範囲でそれぞれ 73, 183, 30 である。各クラスタの人数は 19, 10, 10 人であった。

7 考 察

この章では 2 つの RQ について実験結果からの考察を述べる。

7.1 新たな評価者に対して問い合わせを行う時にサンプル数を増やすに従い漸的に解に近づくこと (RQ1)

図 3a, 3b, 5a, 5b から, シミュレーション実験においても実際の評価者を用いた実験においても評価者から受け取った回答数の増加に伴い漸的に解に近づいていることがわかる。シミュレーション実験では提案手法は, 評価者が最も好むアイテムの推定では高いパフォーマンスを出すことができた。一方で評価者の嗜好分布全体を把握する指標ではベースライン手法のほうが高いパフォーマンスを出すことができた。これは提案手法は好ましくない部分への問い合わせを減らすことができていたため, 既存手法ほど探索を行っていないためだと考えられる。

また, 実際の評価者を用いた実験でも, 提案手法はより高い精度で評価者の嗜好分布全体を把握することができた。一方で評価者が最も好むアイテムの推定においてはベースライン手法のほうが早く収束した。

“そこそこ好ましい” アイテムを提示することが増え, 類似したアイテムをその後に重点的に提示してしまったことが原因として考えられる。

7.2 評価者から高い評価をもらい続けること (RQ2)

図 5c, 6c から提案手法のほうが高いスコアを得ていることがわかる。これは既知の他の評価者の情報を用いることで探索を減らし活用を増やす事ができているからだと考えられる。

また図 3c からはシミュレーション実験では, ベースライン手法と同様に評価者から受け取った回答数の増加に伴う形により高い評価を得られていることがわかる。また, 提案手法のほうが 2~3 回目の回答で高い評価を得ることができている。これはアイテムを提示する時に未知の部分が多い評価者嗜好の情報だけでなく, 他の評価者の嗜好の情報を利用しているため, 効果的な提示ができていないからだと考えられる。一方で図 5c からは実際の評価者を用いた実験では提案手法, ベースライン手法ともに評価者から受け取った回答数が増えたとしても, 受け取る評価を上げることができていないことがわかる。原因として小山ら [7] が指摘しているように主観評価を絶対評価することが難しいことが原因であることが考えられる。この問題への改善策は一対比較法などの相対評価手法を用いる事が挙げられる。

8 結 論

本論文ではこれまでに回答を得た評価者のデータを用いることで, 新たな評価者の嗜好を少ない回答数で把握する手法を提案した。具体的には評価者にアイテムを提示し, そのアイテムへの評価を受け取り嗜好分布を再計算する枠組みの中で, 新たな評価者にアイテムを提示する時に評価者の情報に既知の他の評価者の嗜好の情報を利用する戦略を取る。既知の他の評価者の嗜好の情報は, 評価者たちの嗜好分布をクラスタリングし作られたクラスタの嗜好の情報として利用される。

結果を検証したところ, 評価者が理想的な回答をする場合では提案手法は少ない回答数で評価者の嗜好分布を把握することが可能であることがわかった。また, 高い評価値を得ることが可能であることもわかった。一方で実際の評価者を対象とする場合はまだ課題が残る形となった。今後の課題として, ガウス過程回帰のハイパーパラメータ・クラスタ数を人手で調整する必要があること, 1 つのドメインでしか検証できていないこと, 空間の次元数が増えた際に機能するか検証されていないこと, 主観評価を絶対評価であることが難しいことが挙げられる。

謝 辞

本研究の一部は JST CREST (JPMJCR16E3), JSPS 科研費 (JP21H03552), JST 未来社会創造事業 JPMJMI19G8 の支援を受けたものである。ここに謝意を示す。

文 献

- [1] Jiyi Li, Yukino Baba, and Hisashi Kashima. Hyper questions: Unsupervised targeting of a few experts in crowdsourcing. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM '17*, pp. 1069–1078, 2017.
- [2] Muhammad Zahid Tunio, Haiyong Luo, Wang Cong, Zhao Fang, Abdul Rehman Gilal, Ahsanullah Abro, and Shao Wenhua. Impact of personality on task selection in crowdsourcing software development: A sorting approach. *IEEE Access*, Vol. 5, pp. 18287–18294, 2017.
- [3] 小宮山純平. 多腕バンディット問題. *人工知能*, Vol. 31, No. 5, pp. 743–747, 2016.
- [4] Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. *Proceedings of the 19th international conference on World wide web - WWW2010*, 2010.
- [5] William R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, Vol. 25, No. 3/4, pp. 285–294, 1933.
- [6] Yukino Baba, Jiyi Li, and Hisashi Kashima. Crowdea: Multi-view idea prioritization with crowds. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 8, pp. 23–32, 2020.
- [7] Yuki Koyama, Issei Sato, Daisuke Sakamoto, and Takeo Igarashi. Sequential line search for efficient visual design optimization by crowds. *ACM Transactions on Graphics (TOG)*, Vol. 36, No. 4, 2017.
- [8] 坂井将之, 竹之内宏, 徳丸正孝. デジタルサイネージを用いた複数のユーザの投票によるデザイン生成支援システム. *日本感性工学会論文誌*, Vol. 15, No. 4, pp. 503–511, 2016.
- [9] Benjamin Solnik, Daniel Golovin, Greg Kochanski, John Elliot Karro, Subhodeep Moitra, and D. Sculley. Bayesian optimization for a better dessert. In *Proceedings of the 2017 NIPS Workshop on Bayesian Optimization*, 2017.
- [10] Eric Brochu, Vlad M Cora, and Nando De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.
- [11] 赤穂昭太郎. ガウス過程回帰の基礎. *システム／制御／情報*, Vol. 62, No. 10, pp. 390–395, 2018.
- [12] 持橋大地, 大羽成征. *ガウス過程と機械学習*. 講談社サイエンスフィク, 2019.
- [13] Ron Kohavi, Alex Deng, Roger Longbotham, and Ya Xu. Seven rules of thumb for web site experimenters. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14*, pp. 1857–1866, 2014.
- [14] Katharina Reinecke and Krzysztof Z Gajos. Quantifying visual preferences around the world. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 11–20, 2014.
- [15] Dave Gehrke and Efraim Turban. Determinants of successful website design: relative importance and recommendations for effectiveness. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences*.

1999. *HICSS-32. Abstracts and CD-ROM of Full Papers*, p. 8 pages, 1999.

- [16] O. Byung Kwon, Choong-Ryuhn Kim, and Eun Jong Lee. Impact of website information design factors on consumer ratings of web-based auction sites. *Behaviour & Information Technology*, Vol. 21, No. 6, pp. 387–402, 2002.