

閲覧者による印象のフィードバックを用いたイラスト投稿者の検索

宮嶋 清人^{1,a)} 関 洋平^{2,b)}

概要: イラスト共有 SNS において、閲覧者による印象を利用してイラスト投稿者を検索する手法を提案する。イラスト投稿者の印象は、その人が描いたイラスト群に付与されたタグから、決定木を用いて推定する。選定した 12 の印象が人によらず共通して認識できるか調査したところ、すべての印象の平均で κ 値が 0.58 であった。印象の推定には、クラウドソーシングの考えに基づいて、SNS 上で利用されている「評価タグ」を手がかりとしたところ、12 の印象のうち、8 つの印象で 0.6 以上の F 値が得られた。また、モチーフと印象によるイラスト投稿者のランキング方針について検討したところ、モチーフと印象の両者を利用してランキングを作成する必要があることが確認できた。さらに、より適切な検索結果を得るために、ユーザの情報要求に適合するイラスト投稿者の情報をフィードバックし、検索クエリの拡張を行った。その際、モチーフや印象に関するタグをクエリの拡張に利用することで、 $nDCG$ が 11% 向上した。

1. はじめに

近年、動画やイラスト等のコンテンツをユーザが投稿し、共有する SNS が増加している。こうしたサービスは、知名度の高い投稿者に対して制作物の作成を依頼するといった使い方もされている。



図 1 イラスト投稿者への仕事の依頼

イラスト製作のように、制作物にその人の個性が明確に現れ、かつ求められるテーマが多様なコンテンツの場合、目的に合った制作者を探索できる利益は大きい。しかし、既存のイラストレータ検索 ([1] など) は、人手でデータを登録するため、SNS に見られるような多数のイラスト制作者に対応することは困難である。

そこで、本研究ではイラスト共有 SNS の 1 つである

Pixiv[2] に着目して、各投稿者が描いたイラスト群に含まれるタグからその投稿者の印象を推定し、イラスト投稿者を検索する手法を提案する。提案手法では、タグからその投稿者が描くイラストのモチーフと印象を推定し、検索に利用する。また、入力したクエリだけでは、表現することが難しい情報要求を満たすために、適合する投稿者の情報をフィードバックすることでクエリ拡張を行う。

本論文の構成は次の通りである。2 節では関連研究の紹介を、3 節では、提案手法の紹介と、イラスト投稿者の検索に用いる印象の選定と印象に対する認識の一貫性についての検証を行う。4 節では、決定木による印象の自動推定について評価実験を行う。5 節では、モチーフと印象を用いたイラスト投稿者のランキング方針の検討と、フィードバックを組み込んだシステムの評価実験を行う。最後に、6 節で本論文のまとめを行う。

2. 関連研究

人物検索に関するサービスや研究として、イラストレーターズ JP ネット [1]、桑原ら [3] の研究が挙げられる。イラストレーターズ JP ネット [1] は、イラストレータの得意なモチーフや印象などを人手で登録して検索を行なっている。また、桑原ら [3] は、マイクロブログにおいてユーザの発言から、トピックとそれに対する感情を抽出し、類似ユーザの推薦を行った。本研究では、イラスト共有 SNS において、投稿イラスト群に付与されたタグから印象を自動推定し、まだ注目されていないようなイラスト制作者を

¹ 筑波大学大学院 図書館情報メディア研究科
Graduate School of Library, Information and Media Studies,
University of Tsukuba, Tsukuba, Ibaraki 305-8550, Japan

² 筑波大学図書館情報メディア系
Faculty of Library, Information and Media Science, University of Tsukuba, Tsukuba, Ibaraki 305-8550, Japan

a) s1121740@slis.tsukuba.ac.jp

b) yohei@slis.tsukuba.ac.jp

発見できる枠組みを実現する。

印象や嗜好を利用した検索に関する研究としては、中村ら [4]、住元ら [5]、椋木ら [6] の研究が挙げられる。

中村ら [4] は、動画に寄せられたコメントから印象を推定し、検索に利用した。住元ら [5] は、ブックマークしたイラスト群に付与されたタグから探索者の嗜好を表現し、推薦に利用した。本研究では、ある投稿者の投稿イラスト群に付与されたタグからその人の印象を推定する。また、イラスト共有 SNS において有用と考えられる印象について、椋木ら [6] の画像検索に関する研究を参考にし、不足している印象を補うことで本研究で用いる印象を選定した。

3. 提案手法：印象を利用したイラスト投稿者の検索

3.1 印象を用いたイラスト投稿者の検索

本研究では、ある投稿者が描いたイラスト群に付与されたタグから、その投稿者の印象を推定し、検索に利用する手法を提案する。投稿者の印象は、3.2 節で選定する 12 の印象を用いる。ユーザは、検索の際に、探したい投稿者の (1) 描くモチーフと、(2) 印象を入力する。システムは、(1) を満たし、(2) とタグから推定した印象が近い投稿者を上位に表示する。ユーザは、検索結果から、(3) 投稿者の印象を表すレーダーチャートと、(4) 投稿イラストを見て、自分の要求を満たす投稿者を選択する。

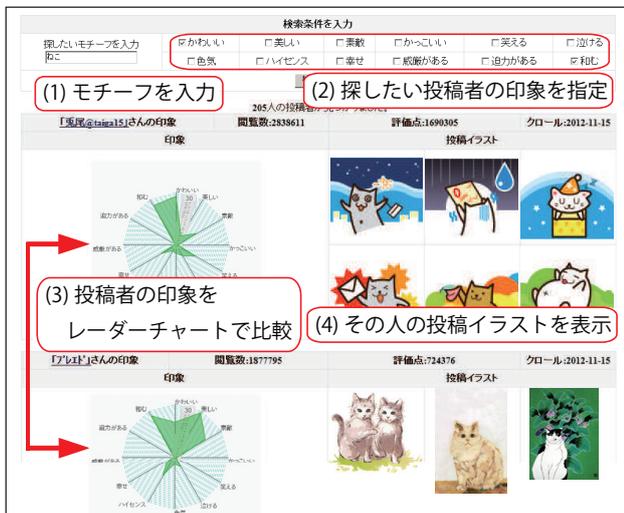


図 2 提案する投稿者検索システム

3.2 閲覧者による印象の選定

イラスト投稿者の検索に使用する印象は、椋木ら [6] の研究や、色のイメージ事典 [7] に記載されていた印象から候補を用意し、Pixiv から収集した 1,800 枚のイラストに付与されたタグと比較することで、イラスト共有 SNS における検索に有用な 12 の印象を選定した。

3.3 印象に対する認識の調査

3.3.1 調査目的

3.2 節で選定した印象に対して、人によりどの程度共通

表 1 選定した印象

| | | | |
|------|-------|-------|-------|
| かわいい | 美しい | 素敵 | カッコいい |
| 笑える | 泣ける | 色気 | ハイセンス |
| 幸せ | 威厳がある | 迫力がある | 和む |

した認識があるか調査し、イラスト投稿者の検索に利用できるかを確認する。

3.3.2 調査方法

4 名の被験者それぞれに、28 名の投稿者のイラストを 10 枚ずつ提示し、選定した 12 の印象を用いて、イラスト投稿者と、イラストに対して抱いた印象を回答させた。得られた回答から、被験者間で、投稿者と、イラストに対して抱いた印象の一致度 (κ 係数 [8]) を算出し、投稿者ごとに印象の一致度を計算した場合と、イラストごとに印象の一致度を計算した場合とを比較した。

3.3.3 調査結果

調査の結果、図 3 のような結果が得られた。

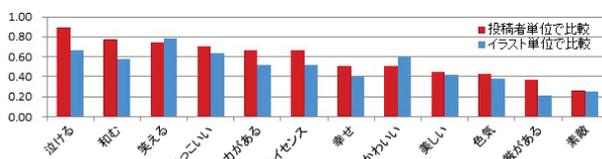


図 3 投稿者の印象の被験者間での一致度 (κ 係数)

3.3.4 考察

図 3 を見ると、投稿者単位の κ 係数は、イラスト単位の κ 係数よりも高い。両者には、 t 検定 (有意水準 5%, 両側検定) で有意差が認められた。これは、同じ投稿者のイラストを複数見ることによって、その人の特徴的な印象が明確になることによると考えられる。

次に、投稿者単位で印象を見た場合の κ 係数を見ると、ほぼ完全な一致 (almost perfect [8], $0.8 < \kappa \leq 1.0$) の 1 つの印象と、高度の一致 (substantial, $0.6 < \kappa \leq 0.8$) の 3 つの印象は、被験者間で共通した認識があるといえる。中程度の一致 (moderate, $0.4 < \kappa \leq 0.6$) の 6 つの印象も、モチーフを用いて検索候補を絞り込むことで、認識のずれを抑えることができると考えられる。軽度の一致 (fair, $0.2 < \kappa \leq 0.4$) の 2 つの印象について、「威厳がある」は、被験者の 1 人 (女性) が他の被験者 (男性) よりも「印象がある」と判断した投稿者数が極端に低かった。この被験者を除いた 3 人の間での κ 値は、0.74 となることから、印象の判断には性差がある可能性が考えられる。「素敵」は、共通して「印象がある」と判断された投稿者も複数見られたが、特定の人しか「印象がある」と判断しなかった投稿者も多い。「素敵」と感じるものが、見る人の嗜好に影響されるためと考えられる。

4. 決定木を用いた印象の自動推定

イラスト共有 SNS においては、SNS の中で生まれた独自の表現や、その時々流行に応じて様々なタグが生まれ

る。こうしたタグにも、印象推定に有効な手がかりが多数含まれている。しかし、ルールベースにより印象を判定する方法では、新たに発生したタグを手がかりとして印象を推定することは難しい。

そこで、本節では、機械学習の1つの手法である決定木 [9] を利用して印象を推定する。こうすることで、あらかじめルールに含んでおくことのできないタグも印象推定の手がかりとして利用することが可能になる。

4.1 実験目的

3 節で選定した印象の推定に用いるルールを手動で構築する場合、新たに発生したタグに対応するには、労力がかかる。そこで、クラウドソーシング [10] の考えに基づいて、手がかりとなるタグを収集することを考える。Pixiv においては、「評価タグ」 [11] と呼ばれるタグが多数存在する。評価タグとは、イラストを見て感銘を受けたユーザが付与するタグであり、多数のユーザが様々なイラストに対して付与している。評価タグには、印象の推定に有用な手がかりが含まれおり、イラスト共有 SNS においてイラスト投稿者の印象を推定する際に有効であると考えられる。評価タグは、一覧が作成されており、様々なユーザが編集を続けている。

そこで、評価タグを手がかりとして、機械学習により自動で印象を推定するルールを構築することを目指す。機械学習の1つの手法である決定木は、こうしたルールを視覚的に分かりやすい形で構築することができるため、決定木を利用する。

また、評価タグ以外にも印象の推定に有用なタグが多数存在するため、評価タグに他の品詞を組み合わせることで、より正確に印象が推定できるかを検証する。得られた結果から、印象の推定に適切な組み合わせを検証する。

4.2 実験方法

4.2.1 決定木に用いる素性

決定木に用いる素性として、4.1 節で述べた評価タグ [11] を利用する。また、評価タグとして一覧が作成されていないタグにも、有用な手がかりとなるものが存在するため、形容詞、副詞のような印象語を含む品詞に関する素性として検証する必要がある。そこで、形態素辞書の UniDic [12] に含まれている様々な品詞に関する語を抽出し、印象の推定に利用できるか調査する。UniDic には、形態素を細かく分解できる、他の形態素辞書では扱っていない「形状詞」の辞書が定義されているといった特徴がある。タグのように短い文字列の場合、より短い単位で定義された辞書の方が有効と考えられる。また、形状詞は、形容動詞の語幹に当たるものであり、口語における出現頻度が高い。そのため、口語を多数含むタグから印象を推定する際に有効であると考えた。

各投稿者のイラストに付与されたタグと、上記の素性の部分一致を行い、投稿イラストに付与されたタグに含まれた各素性の数を利用して印象を推定する。実際にイラストに付与されているタグは、単語だけでなく、複数の語の組み合わせであるものが多いため、完全一致による判定では、うまく判定できないことを考慮し、部分一致を採用した。

以上を踏まえて、本実験で使用する属性は、表 2 の 4 種類とした。これらを用いて、イラスト投稿者の印象を推定する決定木を構築する。

表 2 印象の推定に使用する属性

| | |
|------|---------|
| 評価タグ | 427 個 |
| 形容詞 | 705 語 |
| 副詞 | 2,321 語 |
| 形状詞 | 2,100 語 |

4.2.2 訓練データとテストデータ

訓練データとして、Pixiv のイラスト投稿者 3,300 名のデータを準備した。そして、投稿者一人につき 30 枚のイラストを取得し、付与されているタグを収集した。これらの投稿者に対する印象を、人手ですべてクラス分けすることは、時間と労力の面で現実的でないため、ルールベースによりクラスを自動判定した。判定に使用したルールには、3.2 節の予備調査において収集した 1,800 枚のイラストに付与されていたタグから抽出した 520 の手がかり語を利用している。

テストデータには、3.3 節において被験者実験により印象を判定した 28 名のイラスト投稿者のデータを利用した。これらの投稿者に対する各印象の有無は、4 人の被験者のうち、誰か 1 人でも印象があると判定していれば印象があるものとした。

4.2.3 実験の手順

4.2.1 節で挙げた素性について、次の組み合わせで印象を推定し、得られた結果を比較する。評価尺度として、精度と再現率の調和平均である F 値を利用する。

- (1) 評価タグ、形容詞、副詞、形状詞のそれぞれ単体での推定
- (2) 評価タグ+他の品詞の組み合わせ
- (3) ルールベースによる判定

上記の組み合わせに対して、作成した決定木により推定した印象の F 値を計算する。まずは、評価タグや各品詞を単体で印象の推定に使用した場合に得られる F 値を計算し、評価タグや、他の品詞の有用性を検討する。次に、評価タグと、他の品詞を組み合わせた場合の F 値を計算し、評価タグ単体での F 値との比較を行う。また、ルールベースによる印象推定との比較を通して、決定木による印象推定の有効性について検討する。なお、ここで挙げたルールベースによる判定は、4.2.2 節で挙げた 1,800 枚のイラストに付与されたタグから抽出した 520 の手がかりを利用して

行なっている。

決定木の構築には、フリーのデータマイニングツールである Weka[13] を利用する。また、決定木の作成には、J48 を用いた。J48 は、C4.5[14] を利用したライブラリであり、決定木の作成とデータの更新、枝刈りなどの処理を一括して行う。

4.3 実験結果

実験により、表 3、表 4 のような結果が得られた。

表 3 評価タグや品詞を単体で使用した場合の F 値

| 印象 | 評価タグ | 形容詞 | 副詞 | 形状詞 |
|-------|--------------|-------|--------------|--------------|
| かわいい | 0.690 | 0.552 | 0.778 | 0.467 |
| 美しい | 0.783 | 0.500 | 0.435 | 0.471 |
| 素敵 | 0.600 | 0.182 | 0.483 | 0.500 |
| カッコいい | 0.737 | 0.000 | 0.706 | 0.353 |
| 笑える | 0.667 | 0.286 | 0.500 | 0.000 |
| 泣ける | 0.400 | 0.500 | 0.000 | 0.800 |
| 色気 | 0.556 | 0.400 | 0.316 | 0.462 |
| ハイセンス | 0.778 | 0.000 | 0.091 | 0.400 |
| 幸せ | 0.182 | 0.000 | 0.182 | 0.000 |
| 威厳がある | 0.500 | 0.000 | 0.000 | 0.000 |
| 迫力がある | 0.000 | 0.000 | 0.000 | 0.000 |
| 和む | 0.364 | 0.000 | 0.000 | 0.000 |

太字は、各手がかりを用いた推定結果のうち、最良の F 値を示している。

表 3 から、評価タグや他の品詞をそれぞれ単体で印象の推定に使用した場合、9 つの印象で評価タグを使用した場合の F 値が最も高くなった。「かわいい」、「泣ける」については、それぞれ副詞、形状詞で推定を行った場合の F 値が最も高い。また、「幸せ」の印象は、副詞による推定も評価タグによる F 値と同じになった。「迫力がある」の印象は、評価タグや品詞による推定では、決定木を作成することができなかった。

また、表 4 から、評価タグと他の品詞を組み合わせた場合には、「かわいい」、「カッコいい」の印象で F 値が向上している。「かわいい」の場合は、形容詞を含む組み合わせにおいて、F 値が向上している。「カッコいい」については、ほとんどの組み合わせで F 値が向上しているが、特に評価タグと形状詞の組み合わせの場合、ルールベースによる推定に近い F 値まで向上した。その他の印象に関しては、評価タグ単体での推定と同じ F 値か、やや低い値となった。

4.4 考察：決定木による印象推定の有効性について

まず、評価タグや各品詞単体での印象推定の F 値に注目する。表 3 から、評価タグを利用した推定が、多くの印象で最も良い F 値を得られていることがわかる。その他の品詞による推定は、「かわいい」の印象で副詞、「泣ける」の印象で形状詞を利用した場合に、評価タグより F 値が高かった。これらの印象も、大部分は評価タグに一致する語により判定を行われていた。以上のことから、決定木によるイラスト投稿者の印象の推定には、評価タグが有効であると

いえる。

次に、表 4 において、評価タグと他の品詞を組み合わせた場合は、「かわいい」、「カッコいい」を除く 10 個の印象で評価タグのみによる推定と同じか、それ以下の F 値であった。「かわいい」の印象は、形容詞を組み合わせた場合、「カッコいい」の印象は、他の品詞のいずれを加えた場合も F 値が向上していた。この結果から、評価タグ単体でもある程度判定を行うことが可能であり、一部の印象では、他の品詞を組み合わせて推定を行うことで F 値が向上する可能性があることがわかった。

最後に、ルールベースとの比較について考察を行う。評価タグに様々な品詞を加えて推定を行うことで、「笑える」、「色気」、「ハイセンス」の 3 つの印象については、ルールベースによる推定よりも高い F 値が得られ、「威厳がある」は、ルールベースによる推定と同じ F 値であった。「かわいい」などの 4 つの印象についても、0.6 以上の F 値が得られた。特に、「笑える」、「色気」、「ハイセンス」の印象に関しては、ルールベースによる推定よりも F 値が高い。一方で、「幸せ」、「迫力がある」、「和む」の 3 つの印象については、今後改善していく必要がある。

5. イラスト投稿者検索の評価

本節では、イラスト投稿者検索の評価に関する実験として、イラスト投稿者のランキング方式の決定と、ユーザの情報要求に適合したイラスト投稿者を指定し、フィードバックすることによるクエリ拡張についての検討を行った。

5.1 実験 1：モチーフと印象を用いた

投稿者のランキング方式の比較実験

5.1.1 実験目的

投稿者をランキングする際に、検索条件として入力した、探したい投稿者が描く「モチーフ」と「印象」の両者をスコア計算式に組み込むべきかを検証する。また、スコア計算の際に、「モチーフ」と「印象」のどちらを重視するべきかについての検証も行う。

5.1.2 実験方法

検索者が指定したモチーフと印象を含むイラストの枚数をスコアとしてランキングする方法について比較実験を行った。まず、様々なモチーフや印象を描く投稿者を検索する課題を 20 個作成し、以下の 4 つの方針で検索結果をランキングした。課題の一覧は、表 5 に示す。

- (1) 指定したモチーフを含むイラストの枚数
- (2) 指定した印象を含むイラストの枚数
- (3) 指定したモチーフを含むイラストの枚数 + 指定した印象を含むイラストの枚数
- (4) 課題に応じてモチーフによるスコアと印象によるスコアの重みを調整

方針 (4) は、課題ごとに、モチーフによるスコアと、印

表 4 評価タグに品詞を組み合わせた場合の F 値

| 印象 | 評価タグ 形容詞 | 評価タグ 副詞 | 評価タグ 形状詞 | 評価タグ 副詞 形状詞 | 評価タグ 形容詞 副詞 | 評価タグ 形容詞 形状詞 | 全項目 使用 | 評価タグ | ルール ベース |
|-------|--------------|--------------|--------------|-------------------|-------------------|--------------------|--------------|-------|------------|
| かわいい | 0.741 | 0.643 | 0.690 | 0.643 | 0.714 | 0.714 | 0.741 | 0.690 | 0.857 |
| 美しい | 0.783 | 0.783 | 0.783 | 0.783 | 0.783 | 0.783 | 0.783 | 0.783 | 0.857 |
| 素敵 | 0.600 | 0.581 | 0.600 | 0.581 | 0.581 | 0.533 | 0.581 | 0.600 | 0.778 |
| かっこいい | 0.750 | 0.778 | 0.824 | 0.778 | 0.778 | 0.706 | 0.778 | 0.737 | 0.833 |
| 笑える | 0.462 | 0.400 | 0.667 | 0.400 | 0.364 | 0.600 | 0.400 | 0.667 | 0.400 |
| 泣ける | 0.400 | 0.400 | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 | 0.400 | 0.444 |
| 色気 | 0.522 | 0.500 | 0.476 | 0.545 | 0.545 | 0.455 | 0.520 | 0.556 | 0.455 |
| ハイセンス | 0.778 | 0.778 | 0.778 | 0.778 | 0.778 | 0.778 | 0.778 | 0.778 | 0.720 |
| 幸せ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.182 | 0.556 |
| 威厳がある | 0.500 | 0.500 | 0.000 | 0.500 | 0.500 | 0.400 | 0.500 | 0.500 | 0.500 |
| 迫力がある | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.600 |
| 和む | 0.364 | 0.364 | 0.364 | 0.364 | 0.364 | 0.364 | 0.364 | 0.364 | 0.615 |

太字は、評価タグ単体の F 値より改善しているものを示す。

象によるスコアのそれぞれに 1~3 の 3 段階の重みを掛けて作成したランキングである。この際の重みは、最も高い nDCG が得られたときの重みを採用した。

次に、得られた結果に対して、正規化減損累積利得 (Normalized Discounted Cumulative Gain)[15] を上位 5 件 (nDCG5), および上位 10 件 (nDCG10) について計算し、比較を行った。

nDCG を計算する際の、イラスト投稿者が検索課題に適合するか否かの評価は、以下の基準で行った。

- A(最も適合)：課題に適合するイラストを 5 枚以上投稿
- B(やや適合)：課題に適合するイラストを 3~4 枚投稿
- C(一部適合)：課題に適合するイラストを 1~2 枚投稿
- D(適合せず)：課題に適合するイラストを投稿していない

この際の、モチーフや印象の有無は、イラストを見て判定した印象や、該当する印象のタグ付与とされているかにより判断した。

5.1.3 実験結果

実験 1 により、表 6 の結果が得られた。

5.1.4 考察

表 6 を見ると、モチーフと印象を利用してスコアを計算するランキング方針は、どちらか一方のみによるスコアと比較して、nDCG5, nDCG10 共に高く、t 検定 (有意水準 0.1%, 両側検定) で有意差が認められた。このことから、モチーフと印象を組み合わせることにより、イラスト投稿者を効果的に検索できることが明らかとなった。

また、モチーフと印象の重みを適切に調整することで、より適切な検索結果が得られることも確認できた。この際、特に、印象とモチーフの組み合わせが一般によく見られるもの (「かわいい」印象の「女の子」など) の場合は、印象の重みを高く設定したほうが良い結果が得られ、印象と組み合わせた場合に、モチーフが少ない組み合わせ (「か

表 5 検索課題の一覧

| 課題 | 検索したい投稿者 |
|----|--|
| 1 | かわいさを感じさせる男性を描く投稿者 |
| 2 | かわいい女の子を描く投稿者 |
| 3 | かわいい猫を描く投稿者 |
| 4 | 美しい男性を描く投稿者 |
| 5 | 美しい女性を描く投稿者 |
| 6 | 美しい風景を描く投稿者 |
| 7 | 行ってみたいと思わせるような素敵な景色を描く投稿者 |
| 8 | いい雰囲気のある、素敵な人間関係を描く投稿者 |
| 9 | 格好いい男性キャラを描く投稿者 |
| 10 | 格好良く、強そうな印象を与えるロボットを描く投稿者 |
| 11 | 笑えるネタを描く投稿者 |
| 12 | 面白いパロディを描く投稿者 |
| 13 | 感動的なストーリーのマンガを描く投稿者 |
| 14 | 色気を感じさせるような男性を描く投稿者 |
| 15 | 色気を感じさせる女性を描く投稿者 |
| 16 | ボーカロイドを題材として、个性的でセンスの良いイラストを描く投稿者 |
| 17 | 幸せそうなカップルを描く投稿者 |
| 18 | 威厳やカリスマ性を感じさせるような一次創作のイラストを描く投稿者 |
| 19 | ファンタジーものを題材として、かつ威厳や迫力を感じさせるような壮大なイラストを描く投稿者 |
| 20 | 二次創作以外で、見ていて和むような人や動物を描く投稿者 |

表 6 実験 1 の各方針における nDCG

| ランキング方針 | nDCG5 | nDCG10 |
|-------------------------|--------|--------|
| 指定したモチーフを含むイラストの枚数 | 0.61 | 0.65 |
| 指定した印象を含むイラストの枚数 | 0.50 | 0.52 |
| 指定したモチーフの枚数 + 指定した印象の枚数 | 0.79** | 0.82** |
| モチーフと印象の重みを調整 | 0.81** | 0.85** |

** モチーフ、印象のどちらか一方の場合に対して
有意水準 1%, 両側検定で有意差有り

わいい」印象の「男子」) の場合は、モチーフの重みを高く設定したほうが良い結果が得られた。これらのことから、入力されたモチーフに応じて重みを調整するような仕組みを組み込むことで、検索の精度を向上させることができる可能性があると考えられる。

5.2 実験 2: 投稿者の情報のフィードバックを利用した検索クエリの拡張

5.2.1 実験目的

5.1 節の実験を通じて、イラスト投稿者をその人が描くイラストに含まれるモチーフや、印象を利用することで、検索することができることが分かった。

しかし、探したいモチーフや印象によっては、検索者の要求をうまくクエリに表現することが困難な場合や、検索結果の上位にノイズとなる投稿者が表示されることがある。

そこで、適合性フィードバックを利用することで、検索結果の再ランキングを行うことを考えた。適合性フィードバックを利用することで、クエリとして表現することが難しい要求であっても、検索結果の中から適切なイラスト投稿者を選択し、類似した投稿者を探索することが可能になる。また、検索者の要求に合わない投稿者に対して、負のフィードバックをかけることで、ノイズを減らすことも可能となる。

以上のことから、実験を通じて適合性フィードバックの有用性の検証を行う。

5.2.2 提案: 適合した投稿者の情報を用いた再ランキング

イラストに付与されるタグは、類似した内容のイラストであっても、付与するユーザや、扱う題材などに応じて表記が揺れることが多い。このため、検索時に入力したクエリだけでは、同じモチーフを描いているイラスト投稿者であっても、うまく検索できないといった問題がある。

また、印象についても、3 節で選定した 12 の印象の他にも、あまり頻繁には使われていないが、検索に利用できれば便利な印象が存在する。以上のことから、何らかの手法によりクエリを拡張する必要があると考えた。

そこで、検索結果から、適切なイラスト投稿者を選択し、フィードバックを行うことでクエリを拡張する方法を提案する。適合性フィードバックの代表的なアルゴリズムとして、Rocchio[16] の提案した手法がある。Rocchio は、式 (1) のように、検索結果から検索者の要求に適合する文書と適合しない文書を選択し、それぞれの文書集合に含まれる文書数を用いて検索クエリの重みを調整していた。

$$q' = \alpha q + \frac{\beta}{|D_R|} \sum_{d_i \in D_R} d_i - \frac{\gamma}{|D_N|} \sum_{d_j \in D_N} d_j \quad (1)$$

本研究でも、この考えに基づいて、式 (2) を定義し、選択したイラスト投稿者の投稿イラストに付与された様々な

タグの重みを計算し、検索結果の再ランキングを行う。

$$\text{タグの重み} = \frac{\text{該当タグを含む適合ユーザ数}}{\text{全不適合ユーザ数}} - \frac{\text{該当タグを含む不適合ユーザ数}}{\text{全不適合ユーザ数}} \quad (2)$$

5.2.3 実験方法

前節で定義した、式 (2) を用いて、イラストのモチーフに関連するタグの拡張と、印象に関連するタグの拡張を試みる。そこで、以下の 4 つのフィードバック方法を比較する。

- (1) (ベースライン) 指定モチーフを含むイラストの枚数 × モチーフの重み + 指定印象を含むイラストの枚数 × 印象の重みにより、ランキングする。
- (2) (モチーフのタグ拡張) (1) のスコアに、選択した投稿者のイラストに付与されたモチーフに関連するタグについて、式 (2) で重みを計算し、そのタグを含むイラスト数にかけてスコアを再計算し、再ランキングする。
- (3) (印象のタグ拡張) (1) のスコアに、選択した投稿者のイラストに付与された印象に関連するタグについて、式 (2) で重みを計算し、そのタグを含むイラスト数にかけてスコアを再計算し、再ランキングする。
- (4) (モチーフと印象のタグ拡張) (1) のスコアに、選択した投稿者のイラストに付与されたすべてのタグについて、式 (2) で重みを計算し、そのタグを含むイラスト数にかけてスコアを再計算し、再ランキングする。

(1) は、比較のベースラインとする、指定したモチーフや印象と、その重みによるランキングである。ここで使用する重みは、5.1 節の実験 2 において、最良の結果が得られたものを用いる。(2) は、(1) の結果にモチーフに関連するタグを拡張し、再ランキングする方針である。ここで挙げたモチーフに関連するタグは、4 節において説明した評価タグ [11] 以外のタグとする。(3) は、(1) の結果に印象に関連するタグを拡張し、再ランキングする方針である。ここで挙げた印象に関連するタグは、評価タグ [11] とする。(4) は、(1) の結果にモチーフと印象に関連するタグを拡張し、再ランキングする方針である。

比較の際の評価尺度としては、nDCG[15] を利用する。検索課題としては、5.1 節の実験 1 と同様の、様々なモチーフや印象を描く投稿者を検索する 20 個の課題を用意した。各課題の検索結果の上位 5 件、および上位 10 件を対象として nDCG を計算する。この際の、イラスト投稿者の評価についても、実験 1 と同じ基準で判定している。

また、フィードバックに使用する投稿者は、この基準で「A」と判定した投稿者を適合するとし、「D」と判定した投稿者を不適合とした。「B」、「C」と判定された投稿者は、検索課題とあまり関係のないタグが拡張される恐れがあるため、フィードバックに使用しない。

5.2.4 実験結果

実験の結果、表 7、表 8 のような結果が得られた。

表 8 課題ごとの nDCG の比較

| 課題 | 1. ベースライン | | 2. モチーフの拡張 | | 3. 印象の拡張 | | 4. モチーフと印象の拡張 | |
|----|-------------|-------------|-------------|-------------|-------------|-------------|---------------|-------------|
| | nDCG5 | nDCG10 | nDCG5 | nDCG10 | nDCG5 | nDCG10 | nDCG5 | nDCG10 |
| 1 | 1.00 | 0.85 | 1.00 | 0.98 | 1.00 | 0.94 | 1.00 | 0.94 |
| 2 | 1.00 | 1.00 |
| 3 | 0.68 | 0.69 | 0.68 | 0.73 | 0.55 | 0.61 | 0.68 | 0.82 |
| 4 | 1.00 | 0.83 | 1.00 | 0.98 | 1.00 | 0.92 | 1.00 | 0.96 |
| 5 | 1.00 | 0.94 | 0.81 | 0.85 | 0.96 | 0.93 | 0.81 | 0.85 |
| 6 | 0.81 | 0.84 | 0.91 | 0.89 | 0.81 | 0.87 | 0.94 | 0.92 |
| 7 | 0.95 | 0.94 | 0.91 | 0.93 | 0.96 | 0.97 | 0.91 | 0.93 |
| 8 | 0.85 | 0.93 | 1.00 | 1.00 | 0.91 | 0.97 | 1.00 | 1.00 |
| 9 | 1.00 | 0.92 | 1.00 | 0.96 | 1.00 | 0.98 | 1.00 | 0.94 |
| 10 | 0.56 | 0.65 | 0.82 | 0.86 | 0.67 | 0.71 | 1.00 | 0.97 |
| 11 | 1.00 | 0.96 | 1.00 | 0.91 | 1.00 | 0.93 | 1.00 | 0.91 |
| 12 | 0.95 | 0.88 | 0.95 | 0.84 | 0.91 | 0.79 | 1.00 | 0.89 |
| 13 | 0.45 | 0.64 | 0.92 | 0.93 | 0.79 | 0.79 | 0.92 | 0.93 |
| 14 | 0.69 | 0.77 | 0.81 | 0.76 | 0.65 | 0.69 | 0.68 | 0.74 |
| 15 | 1.00 | 0.96 | 1.00 | 0.95 | 0.92 | 0.95 | 0.92 | 0.95 |
| 16 | 0.81 | 0.87 | 0.92 | 0.90 | 0.81 | 0.83 | 0.88 | 0.90 |
| 17 | 0.94 | 0.94 | 0.94 | 0.96 | 1.00 | 1.00 | 1.00 | 0.98 |
| 18 | 0.63 | 0.57 | 0.91 | 0.87 | 0.81 | 0.76 | 0.76 | 0.78 |
| 19 | 0.50 | 0.59 | 0.83 | 0.80 | 0.57 | 0.68 | 0.83 | 0.80 |
| 20 | 0.32 | 0.37 | 0.63 | 0.64 | 0.66 | 0.70 | 0.74 | 0.75 |
| 平均 | 0.81 | 0.81 | 0.90 | 0.89 | 0.85 | 0.85 | 0.90 | 0.90 |

太字は、各フィードバック方針の中で、それぞれの課題のうちで最も高いものを示す。

表 7 各フィードバック方針における nDCG

| フィードバック方針 | nDCG5 | nDCG10 |
|-------------------|-------|--------|
| 1. ベースライン | 0.81 | 0.81 |
| 2. モチーフのタグ拡張 | 0.90* | 0.89* |
| 3. 印象のタグ拡張 | 0.85 | 0.85 |
| 4. モチーフと評価タグのタグ拡張 | 0.90* | 0.90* |

* 有意水準 5%, 両側検定でベースラインに対して有意差有り

表 7 より、各フィードバック方針の、平均の nDCG を比較すると、モチーフと印象を拡張した場合の nDCG が最も高く、次いでモチーフの拡張、印象の拡張、ベースラインという結果となった。また、モチーフと印象を拡張した場合、およびモチーフを拡張した場合の nDCG と、ベースラインの nDCG との間には、*t* 検定 (有意水準 5%, 両側検定) で有意差が認められた。

表 8 より、課題ごとの nDCG に注目すると、nDCG が大きく向上した課題は、課題 10, 13, 19, 20 であった。また、nDCG10 が、ベースラインより低下した課題は、課題 5, 11, 14, 15 であった。

5.2.5 考察

表 7 より、モチーフと印象の両者を拡張した場合の nDCG の向上率が最も高く、次いでモチーフのみを拡張した場合の向上率が高いという結果となった。このことから、モチーフと印象に関連したタグのどちらも拡張に利用することで、より適切な検索結果を得られることが分かった。特

に、モチーフによる拡張を行った場合に nDCG が向上している課題が多い。これは、適合する投稿者の情報を利用することで、表記揺れによりスコアに加算することができなかったタグや、入力したクエリに関連したモチーフを含むタグが拡張されたことが理由と考えられる。

表 8 より、nDCG が大幅に向上している課題としては、課題 10, 13, 19, 20 が挙げられる。課題 13 は、「格好良く、強そうな印象のあるロボットを描く投稿者」を探すという課題であった。この際、モチーフである「ロボット」を表現するタグには、「ロボ」、「メカ」、「機械」など、様々な表記が存在していた。そのため、フィードバックによりモチーフに関連するタグを拡張した結果、これらのタグを利用してランキングを行うことが可能となり、nDCG が向上したと考えられる。課題 19 についても、同様の理由で、入力クエリと関連性の高いモチーフに関連するタグが拡張されたことにより、nDCG が向上している。課題 13, 20 は、「感動的なストーリーのマンガを描く投稿者」、「二次創作以外で、見ていて和むような人や動物を描く投稿者」を探すといったといった漠然とした要求に基づいて検索を行う課題である。これらの課題は、適合する投稿者の情報をフィードバックすることで、検索者の要求が具体化されるため、nDCG が大幅に向上したと考えられる。

nDCG10 が、ベースラインより低下した課題のうち、課題 5, 11, 15 については、フィードバックを行う以前の段階

で、検索結果の上位のほぼすべてが適合するイラスト投稿者であった。このような場合、クエリを拡張する以前の段階で、既に検索者の要求を満たす投稿者を多数発見できているため、クエリ拡張を行う必要性が低い。フィードバックを行うことで、本来の要求から外れたモチーフや印象も拡張されてしまったことにより、nDCGが低下したと考えられる。また、課題14については、モチーフのみを拡張した場合を見ると、nDCG5が向上している。一方で、印象に関連するタグを拡張した場合は、nDCGが低下した。この課題は、「色気を感じさせる男性を描く投稿者」を探す課題であったが、印象に関連するタグを拡張したところ、「かわいい」や「笑える」など、要求と異なるタグが拡張されてしまった。こういった事例に対しては、フィードバックの戦略を検索者に選択させることにより、対応する。

6. まとめ

本研究では、イラスト共有SNSにおいて、投稿者のイラストに付与されたタグから印象を推定し、イラスト投稿者を検索する手法を提案した。

検索に使用する印象は、印象を定義している関連研究を参考に、イラスト共有SNSにおいて特に利用されているものを選定した。その中から、印象推定を実現するために、タグから推定した印象と、イラストを見て判定した印象が一致する12の印象を選定した。選定した12の印象について、被験者間での一致度(κ 係数)を計算したところ、12の印象すべてで $\kappa \geq 0.5$ となり、人によらずある程度共通して認識できることを確認した。

4節における、決定木を用いた印象の自動推定に関する実験から、評価タグ[11]や、他の品詞を属性とした決定木により、12の印象のうち、8つの印象を0.6以上のF値で推定することができた。ならびに印象の推定には評価タグ[11]が有効であることが明らかとなった。

また、印象とモチーフを、イラスト投稿者のランキングに適切に反映させるための実験を行った。まず、5.1節の実験を通じて、イラスト投稿者のモチーフと印象の両方を手がかりとして用いることにより、適切にイラスト投稿者を検索できることを明らかにした。また、5.2節では、イラスト投稿者の情報をフィードバックすることにより、投稿者を再ランキングする手法について検証した。その結果、モチーフや印象に関するタグを拡張することで、より適切な検索結果を得られることが確認できた。

最後に、本研究の貢献は以下のようにまとめられる。

- (1) イラスト共有SNSを対象とし、SNS上でソーシャルタグを利用することにより、検索者の要求に適合する作品を公開している人物を検索する手法を提案した。
- (2) イラスト投稿者検索の手がかりとして用いる印象を定義し、その自動推定の手法を提案し、実験を通して有効性を確認した。

- (3) モチーフや印象を用いて検索する上で、手がかりとなるタグの表記ゆれの問題に対応するために、情報要求に適合するイラスト投稿者の情報のフィードバックを利用することで、より適切な投稿者を検索できることを明らかにした。

謝辞

本研究の一部は、科学研究費補助金基盤研究C(課題番号24500291)ならびに筑波大学図書館情報メディア系プロジェクト研究の助成を受けて遂行された。また、調査にご協力いただいた皆様にも深く感謝します。

参考文献

- [1] イラストレーターズJPネット:イラストレーターの専門分野別検索, <http://www.illustrators-jp.net/index.shtml> (2012).
- [2] ピクシブ株式会社:イラストコミュニケーションサービス [Pixiv], <http://www.pixiv.net/> (2012).
- [3] 桑原 雄, 稲垣陽一, 奉章草野, 中島伸介, 張 建偉: マイクロブログを対象としたユーザ特性分析に基づく類似ユーザの発見および推薦方式, 情報処理学会研究報告:データベース・システム研究会, Vol. 2009-DBS-149, No. 18, pp. 1-3 (2009).
- [4] 中村聡史, 田中克己:印象に基づく動画検索, 情報処理学会研究報告:ヒューマンコンピュータインタラクション研究会, Vol. 2009-HCI-131, No. 5, pp. 77-84 (2009).
- [5] 住元宗一, 中川博之, 田原康之, 大須賀昭:未知性と意外性を考慮したイラスト推薦システムの提案:3D表示を用いたイラスト推薦インターフェース, インタラクション2011, pp. 519-522 (2011).
- [6] 椋木雅之, 田中大典, 池田克夫:対義語対からなる特徴空間を用いた感性語による画像検索システム, 情報処理学会論文誌, Vol. 42, No. 7, pp. 1914-1921 (2011).
- [7] 日本流行色協会:色のイメージ事典, 同朋舎出版 (1991).
- [8] Cohen, J.: A Coefficient of Agreement for Nominal Scales, *Educational and Psychological Measurement*, Vol. 20, No. 1, pp. 37-46 (1960).
- [9] Quinlan, J. R.: Induction of Decision Trees, *Mach. Learn.*, Vol. 1, No. 1, pp. 81-106 (1986).
- [10] Snow, R., O'Connor, B., Jurafsky, D. and Andrew, Y. N.: Cheap and Fast - But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks, *Proceedings of EMNLP 2008*, Waikiki, Honolulu, Hawaii, pp. 254-263 (2008).
- [11] ピクシブ百科事典製作委員会:Pixiv百科事典「評価タグの一覧」, <http://dic.pixiv.net/a/評価タグの一覧> (2012).
- [12] 伝 康晴, 山田 篤, 小椋 秀樹, 小磯 花絵, 小木 曾智信:UniDic version1.3.12, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h23/html/nc213120.html> (2010).
- [13] University of Waikato: Weka, <http://www.cs.waikato.ac.nz/ml/weka/> (2011).
- [14] Quinlan, J. R.: C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers, Morgan Kaufmann Publishers (1993).
- [15] Jarvelin, K. and Kekalainen, J.: Cumulated Gain-Based Evaluation of IR Techniques, *ACM Transactions on Information Systems*, Vol. 20, No. 4, pp. 422-446 (2002).
- [16] Rocchio, J. J.: Relevance Feedback in Information Retrieval, *The SMART Retrieval System - Experiments in Automatic Document Processing* (Salton, G, ed.), New Jersey, Prentice Hall, pp. 313-323 (1971).