

表紙画像とストーリーに着目したユーザの好みに類似する漫画推薦

11571073 玉木 一真 (灘本研究室)

あらまし：近年，漫画の新刊の数が増加の傾向にあり，膨大な漫画の中からユーザ好みのものを見つけるのは困難である．また漫画を推薦する際，小説と異なりストーリーが類似しているだけでなく絵の類似性も重要である．そこで本研究では，絵とストーリーが類似する漫画の推薦手法を提案する．具体的には絵の類似性を図るために表紙画像を，ストーリーの類似性を図るためにレビューを用いる．この両方がユーザ好みの漫画と類似する漫画を推薦する手法の提案を行う．

1. はじめに

近年，漫画の新刊の数が増加の傾向にあり，2016年の1年間において12,591タイトル^[1]出版されている．その膨大な漫画の中からユーザが好みに合った漫画を探すのは困難である．なぜなら，ユーザごとにそれぞれ好みの漫画が異なると共に，漫画も多種多様であるため，新たに漫画を探す場合，その好みを参考にして探さなければいけないからである．一方でAmazon¹に代表されるようにネットショッピングでは商品の推薦システムがある．このような推薦システムでは，ユーザが好みの漫画を購入した際，おすすめ機能に他の漫画が表示されるが，それは同じ作品の別の巻の漫画や，現在人気の漫画，また同時期に出版された漫画がほとんどである．従って，ユーザの好みを把握して漫画を推薦しているシステムとは言いがたい．そのため，絵とストーリーの両方が好みに合う漫画を探すのは非常に困難である．そこで我々は，ユーザの好みにその両方が類似する漫画を推薦するシステムがあると便利であると考えた．本研究では，ユーザの好みに類似する漫画を提示することを目的とし，絵とストーリーが類似する漫画の推薦手法を提案する．

具体的には以下の手順で絵とストーリーが類似する漫画の推薦を行う．

- ① ユーザの好みの漫画を入力
- ② 類似表紙画像を持つ漫画の決定
- ③ レビューから各漫画のトピックの抽出
- ④ 類似レビューを持つ漫画の決定
- ⑤ 類似表紙画像の漫画と類似レビューを持つ漫画が一致する漫画の提示

これにより，膨大な数の漫画の中からユーザの好みを把握し，絵とストーリーの類似する漫画の推薦を行う事ができ，ユーザは好みの漫画を容易に取得することが可能となる．

2. 関連研究

漫画の推薦に関する研究は多数あり，松村^[2]らはあらすじやその漫画の特徴に絞りそれに類似

する漫画の推薦を行っている．

村瀬^[3]らは，利用者より読みたい漫画の作品名のリクエストを受け，作成した類似度データを比べて，漫画を推薦している．

これらに対し本研究は，絵とストーリーがユーザの好みに類似する漫画の推薦を行う点が異なる．

3. 類似表紙画像の抽出

ユーザが好みの絵であると判断する際，表紙画像を見て，色合いや構図，絵のタッチなどを判断材料として決めている場合が多いと考えられる．そこで本研究では，色合いを用いて類似表紙画像の抽出を行う．その理由として，構図や，絵のタッチは表紙画像の一部だけであるが，色合いは全体に使われるものだからである．また連載している漫画を読む時，ユーザは1巻目から読み始めるのが一般的であると考え，その漫画の1巻の表紙画像を用いる．本研究では，漫画の表紙画像の色合いによる類似度を求めるのにカラーヒストグラムの手法を用いて，カラーヒストグラムの類似度が上位の表紙画像を類似画像とする．カラーヒストグラムとは色の三原色であるRGBを用いて，画像の中にRGBの各色が何ピクセルあるか数えて作成したグラフのことである．それを用いて類似する色合いの表紙画像を抽出する．

4. 類似レビューの抽出

漫画のストーリーすべてを取得し，漫画同士のストーリーの類似度を求めるのは困難である．そこで本研究では，レビューにストーリーの要約が書かれている場合が多い事に着目し，それを用いて漫画のストーリーの類似度を求める．

まず，漫画毎にレビュー内容を取得するために本研究では漫画レビュー.com²を用いる．このとき，レビューが極端に少ない場合は，ストーリーの真意に関わると考え，レビュー数が5件以上の漫画を対象とする．また，レビュー200件を目視した結果，評価点が極端に低いレビューはその漫画のストーリーを顕著に表していない場

¹ Amazon: <http://www.amazon.co.jp/>

² 漫画レビュー.com: <http://www.mangareview.com/>

合が多い。漫画レビュー.com はレビューの評価点が 10 点満点であるため、4 点以上のレビューを対象とする。このように 5 件以上の漫画かつ評価点が 4 点以上のレビューを取得する。

取得したレビューの中には、様々なトピックについて書かれている場合がある。そこで、トピックモデルを用いてクラスタリングを行うことにより、レビューをトピック毎に分類する。トピックモデルには Repeated Bisection [4]を用いる。各種パラメータは、クラスタリングツールには Bayon, クラスタ数は Bayon の分割ポイントにより自動で決定をし、Bayon の分割ポイントは 1.5, 形態素解析には JUMAN を用いる。

漫画同士の類似度は、クラスタを 1 文書とし、その漫画に含まれるクラスタすべてを推薦対象となる漫画のクラスタ内のレビューとのコサイン類似度により求める。コサイン類似度が、上位のものをストーリーが類似している漫画とする。

5.実験

提案手法の有用性を示すために、表紙画像とレビューの 2 種類の実験を行った。

5.1 類似表紙画像の抽出実験

カラーヒストグラムを用いた類似表紙画像抽出手法が本研究に適しているかどうかの実験を行う。実験対象は「銀魂」「FAIRY TAIL」「艶姿純情 BOY」の 3 つの漫画を用いた。使用した「銀魂」の表紙画像は青を基調とした画像であり、「FAIRY TAIL」は茶を基調とし、「艶姿純情 BOY」は赤を基調としている。推薦対象の画像は各々無作為に抽出した 50 枚の表紙画像である。

実験方法

被験者の 7 名に「銀魂」の表紙画像と類似度が高かった上位 5 件の画像を見てもらい、似ているか似ていないかを直感的に判断した。

結果と考察

表 1 に「銀魂」の結果例を示す。似ている、似ていないと判断したユーザが概ね半分に分かれ、類似表紙画像を抽出するためには色合いの他に構図や絵のタッチが必要であることがわかった。

表 1：類似表紙画像上位 5 件

漫画名	類似度[R, G, B]	結果(%)
黒子のバスケ	[0.622, 0.497, 0.933]	○:43×:57
ニセコイ	[0.793, 0.727, 0.840]	○:57×:43
コウノトリ	[0.651, 0.579, 0.796]	○:43×:57
マギ	[0.573, 0.554, 0.791]	○:86×:14
七つの大罪	[0.636, 0.512, 0.726]	○:57×:43

5.2 類似レビューの抽出

実験は「銀魂」と「FAIR YTAIL」を対象データとした。推薦対象の漫画は 5.1 節と同じで

ある。対象のレビューは無作為に抽出した各々 50 件である。

実験方法

漫画毎のレビューをクラスタリングし、実験対象と推薦対象の漫画のクラスタ毎にコサイン類似度を求め、被験者 7 名に類似度が上位 5 件の漫画のストーリーが似ているか似ていないかを判断した。

結果と考察

表 2 に「銀魂」の結果を示す。「銀魂」のジャンルはギャグとバトルであり推薦された漫画はギャグとバトルとスポーツであるため結果が極端に低くなっている漫画はスポーツである。ある程度類似する漫画が抽出できたが、ジャンルの考慮が必要であることがわかった。

表 2：類似レビュー上位 5 件

漫画名	cos 類似度	結果(%)
日常	0.4277	○:71 ×:29
あひるの空	0.4173	○:0 ×:100
金色のガッシュ!!	0.4140	○:86 ×:14
アイシールド 21	0.4129	○:14 ×:86
鋼の錬金術師	0.4120	○:71 ×:29

6.まとめと今後の課題

本研究では、ユーザの好みの表紙画像とストーリーに類似する漫画の推薦手法の提案を行った。具体的には、類似表紙画像の抽出はカラーヒストグラムを用いて、類似ストーリーの抽出はレビューを用いて行った。実験により、表紙画像には色合いの他に構図や絵のタッチ、ストーリーにはジャンルの考慮が必要だとわかった。今後の課題としては、表紙画像とストーリーの両方が類似する漫画の提示と表紙画像には構図や絵のタッチの考慮、ストーリーにはジャンルの考慮をすることである。

参考文献

- [1]情報メディア白書 2018. 2018 年 2 月 22 日第 1 版発行. pp.70-71
- [2] 松村拓実, 河野一志, 延澤志保, "読者レビューでの評価情報に基づく漫画の推薦", 情報処理学会第 79 回全国大会. no.2, pp.537-538, 2017.
- [3] 村瀬尊好, 柊和佑, 安藤友晴, "マンガの概要に基づく作品推薦システム", FIT2012 第 11 回情報科学技術フォーラム, 第 4 分冊 pp.319-325, 2012.
- [4] Ying Zhao and George Karypis. Comparison of agglomerative and partitional document clustering algorithms. Technical report, Department of Computer Science, University of Minnesota, Minneapolis, MN 55455, 2002.