

SNSの解析による娯楽作品についての記憶の想起・記録支援

山本愛美^{1,a)} 寺田実^{1,b)}

概要：過去に触れた娯楽作品について想起する際、いつどのような作品に触れたか、そしてその作品についてどのような感想を持っていたかを詳細に想起することには限界がある。Twitterのようなライフログを記憶想起の手助けとすることも可能であるが、膨大なツイートデータから目当ての作品についてのツイートを見つけ出すためには時間がかかる。そこで、本研究ではツイートデータからどのような作品についてよくつぶやいているかを自動的に検索し、それらの検索データからユーザーの娯楽作品年表を作成することのできるシステムを提案する。

Memory recall and record support for entertainment works by analyzing SNS

MANAMI YAMAMOTO^{1,a)} MINORU TERADA^{1,b)}

Abstract: When recalling entertainment works we have seen, there is a limit to remember in detail what you saw and what you thought about them. Although it is possible to use life logs like Twitter, it takes time to find the tweets about the target work out of enormous tweet data. In this research, we propose a system which automatically finds what kind of works are referred in the tweet data and creates the user's entertainment chronology.

1. はじめに

1.1 背景

人間はしばしば、過去に起きた出来事の記録を行う。過去の出来事を記録する意味としては、記録を付けることで自分の行動や考えを整理する、もしくは後から記録を読み返したいと考えている、などがあげられる。

その中でも「自分自身が過去に触れた作品やその感想」を記録したいと考える人々が存在する。そのような人々は「読書メーター^{*1}」などの感想、記録サイトを利用していることも多い。逐一記録を行う人々はそのような記録手法で良いものの、普段意識的に記録を付けていない人間に「○○という作品を読んだのは何年前だったか」「その時自身はどのような感想を持っていただろうか」などの疑問が生

じた際、過去自分がどのような作品に触れてきたか思い出すことが難しい場合がある。

当然ながら人間の記憶能力には限界があり、記録したいと考えている出来事が昔のものであればあるほど詳細な記憶の想起は難しくなる。そのため、人々はしばしばあらかじめ書いておいたメモ書きなどの記録を参照して記憶の想起補助を行う。この記憶の想起補助のためのメモ書きは記録する内容が昔の出来事である場合ほど効果を大きく発揮するが、当然そのように活用するにはあらかじめメモ書きを残しておく必要がある。

本研究では、この「メモ書き」にあたる記憶の想起資料としてSNSの一つであるTwitter^{*2}の投稿データを中心に活用することにより、想起のためのメモ書きが残っていないケースにおいても可能な限り記憶想起の補助を行うことが可能なシステムの提案を行う。

¹ 電気通信大学

The University of Electro-Communications

^{a)}y1631160@edu.cc.uec.ac.jp

^{b)}terada.minoru@uec.ac.jp

^{*1}<https://bookmeter.com/>

^{*2}<https://twitter.com/>

1.2 目的

本研究の目的は、ユーザーが過去に触れた娯楽作品を膨大なツイートデータから自動的に抽出、提示することによる記憶の想起支援、そして作品に触れた時期や感想を同時に記録しておくことにより、後から読み返すことで作品に触れた当時について想起することが可能な娯楽作品年表の作成、出力である。

年表の作成題材として映画、小説などの娯楽作品を選択した理由としては、後から読み返すことが可能である年表として出力することで、以後、当時の記憶の振り返りを容易に行うことが可能になるという点があげられる。Twitterの利用者の中には自分自身が触れてきた娯楽作品についての感想などを投稿する人々が存在する。映画「アバター」のように一般的に使われる「アバター」という言葉と作品名との区別が難しいなどの課題はあるものの、膨大なツイートの内そのような娯楽作品についてのツイートを自動的に判別、抽出することが可能になれば、ユーザーがこれまで触れてきた娯楽作品についてのデータや感想の整理が容易になる。

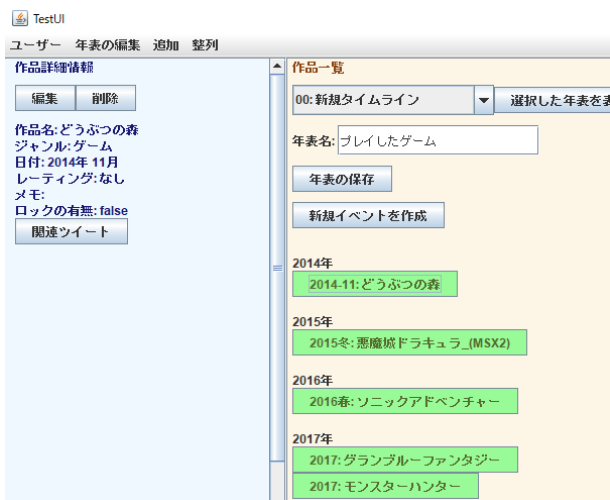


図 1 娯楽年表作成システムにて作成された年表の例

1.3 娯楽作品年表について

本研究では Twitter による投稿データをユーザーが参照することによる映画、漫画、小説などの娯楽作品についての想起効果に着目している。具体的には投稿データをシステム内にて解析し、発言データや投稿画像データからユーザーがこれまでに発言したと思われる娯楽作品名を自動的に抽出、そして抽出した作品群により構成された年表を作成することにより、ユーザーが過去触れてきた娯楽作品の年表を作成する。

図 1 は実際に本研究にて提案した娯楽年表作成システムで作成された年表の一例である。

2. 先行研究

2.1 思い出構築閲覧システム【yourStory】[1][2]

【yourStory】とは仲谷らが作成したコンピュータを用いた自分史作成支援システムである。

このシステムはコミュニティに着目しており、ユーザーの記述するエピソードはその当時所属していたコミュニティに関連付けられる形で登録される。

2.2 楽らく自分史 [3]

このソフトウェアは、自分の学歴や年齢を入力することで自分史の作成を支援する。ソフトウェア上での表示のほか、CSV、HTML などの形式で出力することも可能である。

2.3 LIFE STORY: 自分史作成支援サービス

朝日新聞社の行っている自分史の作成を支援するサービス^{*3}である。様々な社会的ニュースや、それに付随する写真を見ながら自分史を作成することができ、さらに製本が可能となっている。

2.4 スマートフォンのカメラ画像に基づくユーザープロフィール抽出 [4]

Google による画像検索サービスを用いることで、スマートフォンに保存している画像に何が映っているかについての推測結果を調べ、そこからユーザーがどのような写真を多く保存しているかについてのプロフィール作成を行う研究である。また、これらの画像が何についての画像であったかという推測結果を用いて、保存されている画像に自動タグ付けを試みている。

Google による画像検索サービスを用いて画像推測を行うという点において、本研究の先行研究であると言える。

2.5 ツイートのまとめ作成サービス

本家 Twitter におけるモーメント機能や、Togetter^{*4} など、ツイートをまとめることに特化した Web サービスは既に存在している。本研究で行うツイートのまとめとの相違点としては、これらの Web サービスは完全に自力での編集を必要としているが、本研究ではあらかじめテーマとなる娯楽作品を指定することで関係のあるツイートの推測、提示が可能となっている。

3. 提案手法

3.1 娯楽作品年表作成システム

本研究では過去にユーザーが触れた経験のある娯楽作品について整理、保存、出力を目的とした年表作成システム

^{*3}<https://lifestory.asahi.com/jibunshi>

^{*4}<https://togetter.com/>

の製作を行った。

本システムではユーザーが過去に触れた娯楽作品の作品名、触れた日時、その作品について言及したツイートなどをまとめたデータを項目とする独自形式の年表を作成することができる。また、作成した年表は Web ページとして出力することも可能である。

本システムではユーザーが娯楽作品について想起する手段を二通り実装している。

一つ目の手法では Twitter の発言データから過去にユーザー自身がどのような作品について発言していたのか自動抽出を行い、娯楽作品年表へ一括登録を行う。こちらの手法ではあらかじめ Twitter 上で娯楽作品について発言していることが条件となるものの、各ユーザーに沿った作品名提示を行うことが可能である。

もう一つの手段では、ユーザーが想起したい年をシステムへ入力することにより、「年代流行」*5 という各年どのような有名娯楽作品が発表されていたかをまとめた Web サイトの該当年の表示を行い、それを手掛かりとして記憶の想起を行う。こちらの手法ではユーザーの嗜好を考慮せずその年に大きく話題になった作品のみの提示となるものの、Twitter にて発言していない作品についても想起できる可能性がある。

4. 実装

娯楽作品年表作成システムの中でも、Twitter の発言データを用いて年表を作成する部分の実装について述べる。

なお、本研究における娯楽作品年表作成システムは Java を用いて実装を行った。

4.1 作品名データベースの作成

Wikipedia に存在する映画、小説などのカテゴリから記事名を抽出し、作品名データベースを作成する。カテゴリは映画、小説、漫画、アニメ、ドラマを作成した。

また正確な作品名だけでなく、各記事に転送設定が行われている単語（リダイレクト名称）についても対応を行い、リダイレクト名称がツイート中に出現した場合もその作品名についてつぶやいた扱いとして処理を行うようにデータベースを構成している。

4.2 ツイートの解析

ユーザーが過去にどのような作品に触れていたかについて解析を行うため、ユーザーの TwitterID から過去の発言、RT 内容の形態素解析を行う。

具体的な実装方法としてはツイートの本文を mecab*6 による形態素解析にかけ、どのツイートにどの単語が含まれているのか抽出を行っている。

4.3 作品名の検索

4.3.1 作品名データベースとのマッチング

ツイートの解析により取り出された単語が作品名データベースにおける作品名やリダイレクト名称と一致しているか調べ、図 2 のような形で結果をユーザーへと提示する。これにより各作品について言及したツイートが特定されるほか、言及回数やそのツイート内での「面白かった」「楽しかった」などの感想と思われる発言の有無などに着目して表 1 を基にスコア付けを行うことにより、感想を述べている可能性の高い娯楽作品を上位に提示する。

表 1 形態素解析後のツイートに対するスコア算出表

加算対象	スコア [点]
作品のフルタイトル	10
作品のリダイレクト名	3
作品名の一部	1
同一ツイート内での「アニメ」発言	3
「面白かった」「楽しかった」などの形容詞	3



図 2 娯楽作品群がユーザーへ提示される様子

4.3.2 画像解析による作品名検索

ユーザーがどの作品に触れたかについてのデータをより多く収集するため、ユーザーが投稿した画像、RT にて投稿されていた画像の解析を行う。

画像の解析には Google 検索にて用いられている「画像による検索」を用いた。この「画像による検索」機能を用いることで、Google による画像の推測結果を文字列として得ることができる。この得られた文字列も発言データと共にデータベースとのマッチングを行うことにより、娯楽作品に関係のある画像の検索を可能にした。

4.4 年表の作成

本システムでは娯楽作品を年表の項目として年表作成を行うが、この年表の項目とユーザーが該当作品についてつぶやいたツイートの紐づけを行うことにより、作品についての感想などを保存しておくことが可能である。

*5<http://nendai-ryuukou.com/>

*6<http://taku910.github.io/mecab/>

作品とツイートとの紐づけシステムにおいて、膨大なツイートの中から一つの作品についてつぶやいているツイートを見つけ出すことは手間がかかる。そのため、本研究では紐づけるツイートをユーザー自身の発言とユーザーが行ったリツイートの中から作品に関連するであろうツイートの候補群を生成し、そこから任意のツイートの選択、編集を行い紐づけを行うシステムとなっている。具体的な候補群の生成タイミングとしては、作品名とツイートのマッチング時にマッチングが成功したツイートを候補の一つとして記録を行っている。



図 3 Google による「画像を用いた検索」

5. 評価実験

5.1 実験:「作品名の抽出」機能による作品抽出数の測定

本システムのツイートからユーザーが触れた経験のある作品名を抽出する機能について、実際にツイートがユーザーの知識と一致していたか、つまりその作品の存在を知っていたかを評価するために実験を行った。

具体的には、被験者 4 名に被験者自身が普段使用している Twitter のアカウントを用い実際にシステムを用いて「作品名の抽出」機能を利用し、アカウント内で触れられているアニメ作品の抽出を行った。抽出された作品群は表 1 のスコア算出表に従ってスコア 4 以上の物が提示され、被験者自身によって提示された作品が「見たことがあるアニメ」かどうか、もしくは「見たことはないが知っているアニメ」かどうかを手動で選択させた。

なお、Twitter の発言データ収集数は直近 1000 件に設定し、発言に対する解析はそれら 1000 件すべてに行っているものの、画像解析による作品名の検索はそれらの発言の内時期が偏らないように画像を 20 件選択して行った。

また、検索対象のメディアをアニメからゲームに変更しその抽出精度を調べる実験も同時に行った。詳細な実験結果は省くが、結果は 4 のようになった。

5.2 結果

実験結果は表 2、表 3 の通りとなった。

なお、表における「見た」作品群は「知っている」作品

群にすべて含まれており、「知っている」作品群は「抽出作品数」にすべて含まれている。よって、抽出作品の合計である「抽出作品数」は「知っている」作品群と「知らない」作品群を足したものとなる。

表 2 発言データ 1000 件に対する作品抽出数 [件]

被験者	抽出作品数	見た	知っている	知らない
A	1	0	1	0
B	42	20	31	11
C	6	5	5	1
D	9	6	7	2

表 3 画像データ 20 件に対する作品抽出数 [件]

被験者	抽出作品数	見た	知っている	知らない
A	10	2	10	0
B	3	1	2	1
C	4	3	3	1
D	0	0	0	0

表 4 作品の抽出精度 [%]

抽出ソース	対象メディア	見た	知っている	知らない
発言	アニメ	0.53	0.76	0.26
画像	アニメ	0.35	0.88	0.59
発言	ゲーム	0.19	0.51	0.49
画像	ゲーム	0.06	0.53	0.47

6. 考察

実験で得られた結果より、作品の抽出精度を算出すると表 4 の通りとなる。

表 2、表 3 では被験者により抽出作品数が大きく異なっているが、これは Twitter アカウントの普段の利用用途の違いであると考えられる。実験後被験者に簡単なヒアリングを行ったが、被験者 A は普段からあまりアニメを見ないとのことだった。

発言データの抽出数と比較し画像データの抽出数が少ないが、これは発言データの解析数が 1000 件であったのに比べ画像データの解析数は投稿画像の内 20 件と低く抑えていたためだと考えられる。このように画像の検索数を少なめに設定した理由としては、評価実験の性質上一度に多くの被験者の画像を Google 画像検索によって解析を行うと検索が弾かれてしまうためであり、ユーザーが個人でシステムを使う場合においては 100 件前後まで検索が可能であると考えられる。

実験後のヒアリングにて想定していた作品名が出てきたかを尋ねたが、被験者 A によるとゲームの作品名を検索した際、一番多く発言していたであろう作品が出てこなかったとのことだった。詳細を聞いたところその作品は一般的な単語を組み合わせた作品名であったため、形態素解析に

て作品名が取得されていなかった可能性が高い。具体的には、「DARK SOULS」というゲーム作品は解析過程において「DARK」と「SOULS」に分割されてしまい、ゲーム名として作品名のマッチングが行われなかった。

固有性の高い作品名は形態素解析により正確に作品名を切り出すことが可能であり、今回の実験で抽出された作品の多くもそのようなタイトルの作品であった。具体的には、「ポケットモンスター」「ラブライブ！」などのアニメ作品名は形態素解析において一単語の固有名詞として解析されていたため、フルタイトルを発言したと判断され判定スコアにおいても上位となっていた。

発言しているにも関わらず知らない作品名が抽出されているという点については、恐らく被験者が発言した一般的な単語がシステムによって別の意味に捉えられ、作品名として解釈されているケースが多い。具体的には被験者が食べ物のつもりで発言した「ナルト」という単語から、漫画作品「ナルト」に関連する作品名が提示されてしまうケースが見られた。

7. 結論

表4の通り、ユーザーがTwitterにて発言したことのある作品の抽出はある程度できていると考えられる。しかし、本システムではユーザーが名前を知っているのみの作品と実際に触れたことのある作品を判別することができず、実際に触れた作品について記録する年表を作成するためには可能な限り多くの作品を提示する必要があると考えられる。

提示する作品数を増やす方法としては、解析する発言数や画像数を直接増やす他に、算出されたユーザーとの関連スコアが低い作品についても提示を行うことで精度は下がる可能性があるものの作品抽出数自体は増加するものと考えられる。

8. 課題

本システムの課題としては、考察で述べた通り一般的な単語が作品名であった場合、ツイッターの発言からその単語が作品名として発言されたのか否かの判別をつけることが難しいという点があげられる。想定される改善方法としては、前後のツイートの参照、ハッシュタグの意味を解析することによって作品について発言していることの証明を行う、などがあげられる。

また、ツイートと作品名との紐づけ機能について、現在は直接作品名を発言しているツイートとそれらのツイートの前後に発言されたツイートを紐づけ候補としてユーザーに見せている。しかしそれらの候補群選択手法では選択されないようなツイートの選択を目的として、任意のタグ、もしくは単語による検索を行い、それらの検索結果のツイートを紐づけするような機能の実装を予定してい

る。そして、これらのツイート紐づけ機能の有効性を証明するため、実際に被験者に紐づけ機能を使ってもらい、システム無しで作品についてのツイートを収集する場合と紐づけシステムを実際に用いて作品についてのツイートを収集した場合の効率、使いやすさについて評価実験を行う予定である。

参考文献

- [1] 橋本克哉, 仲谷善雄. 思い出構築閲覧システム【yourStory】の評価. 情報処理学会第69回全国大会講演論文集 2007(2), pp. 149-150.
- [2] 山崎和紘. 認知症者と家族の思い出共感支援システム. 情報処理学会第75回全国大会講演論文集 2013(1), pp. 197-198.
- [3] インクリメントP. “楽らく自分史”. 2000.
- [4] 篠田美香子, 酒田信親, 西田正吾. スマートフォンのカメラ画像に基づくユーザプロフィール抽出. 研究報告コンピュータビジョンとイメージメディア (CVIM) 2015-CVIM-195(35), pp. 1-5.