

画像データとテキストデータに基づくマルチモーダルビデオゲーム リコメンドシステムの実現

中込佑介[†]、岡田龍太郎[†]、中西崇文[†]

武蔵野大学データサイエンス学部データサイエンス学科[†]

1. はじめに

近年、ビデオゲーム実況の普及により、幅広い世代でビデオゲームを実際にプレイするユーザーが増加している。ビデオゲーム販売サイトでは、ユーザーがプレイしたいビデオゲームを見つけやすくするために、ビデオゲームを対象とした検索機能に力を入れつつある。しかしながら現状実現されているビデオゲームを対象とした検索機能としては、新作や売上の順位によるもの、キーワード検索によるもの、RPGやアクションなどのジャンル分けによるものが主であり、ビデオゲームの内容に重きを置いたものにはなっていない。

本稿では、画像データとテキストデータに基づくマルチモーダルビデオゲームリコメンドシステムの実現方式を示す。本方式では、ビデオゲームに関する情報として、ゲーム販売サイトからテキストデータと画像データを抽出しデータセットを作成する。そして、この二つのデータのそれぞれで独立に類似度分析を行う。その上で、ユーザーが好みであるゲームと類似しているゲームを画像データおよびテキストデータの二つのデータを合算してリコメンドする。さらに、類似度計量を統合するいくつかの演算子を定義することで、画像データとテキストデータ両方が類似するゲームだけでなく、画像データは類似するがテキストデータは類似しないゲームもしくは、テキストデータは類似するが画像データは類似しないゲームのリコメンドも実現する。これにより、ゲームを簡単に見つけるだけでなく、ユーザーが簡単に見つけることができない新しいゲームのリコメンドを実現することになる。

2. 関連研究

テキストデータと画像データを組み合わせて検索を行う研究として、小邦ら[1]は、レシピサイトの調理手順を記述したテキストデータと料理画像の特徴量を組み合わせることにより、他の

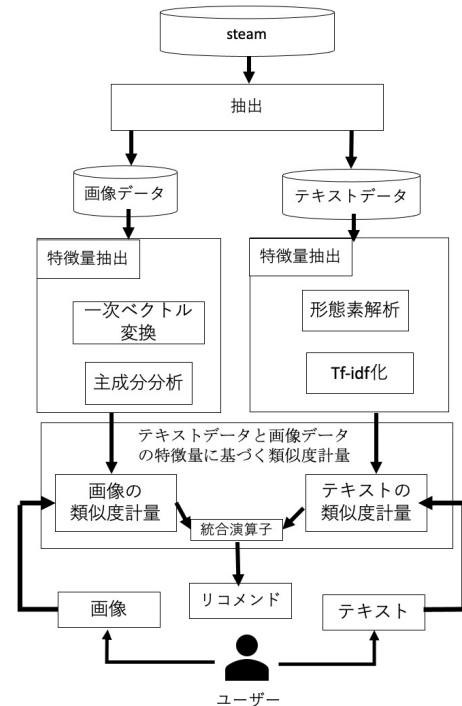


図1 マルチモーダルビデオゲームリコメンドシステムの全体像

レシピの構成方法を模倣した「重複レシピ」を検出する手法を実現している。本方式は、ゲームを画像データおよびテキストデータの二つのデータのそれぞれの類似度計量を行い、類似度計量を統合するいくつかの演算子を定義することで、ユーザーが簡単に見つけることができない新しいゲームのリコメンドを実現する

3. 提案方式

3.1 全体像

図1にマルチモーダルビデオリコメンドシステムの概要を示す。本方式では大きく分けて、テキストデータと画像データの特徴量抽出、テキストデータと画像データの特徴量に基づく類似度計量、統合演算子によって構成される。また本研究では、ゲーム配信サイトである Steam[2]から画像データおよびテキストデータを抽出する。steam は多くのゲームを販売しておりデータを抽出するサイトとして最も適していると判断した。

3.2 テキストデータからの特徴量抽出

テキストデータからの特徴量抽出は、収集機能、

Multimodal Video Game Recommendation System by Image and Text Data Similarity

Yusuke Nakagomi[†], Ryotaro Okada[†], Takafumi Nakanishi[†]

[†] Department of Data Science, Musashino University

1	peach beach splash	-0.941597
2	STEINS;GATE	-1.135649
3	Red Dead Redemption 2	-1.240788
4	RUST	-1.283219
5	Project winter	-1.328482
6	fallout	-1.39665
7	SEKIRO	-1.443532
8	Ultimate Chicken Horse	-1.508433
9	Fallguys	-1.573826
10	hitman	-1.602944

表1 両方のデータが類似している時の検索結果(クエリ: ATRI)

1	Ultimate Chicken Horse	-0.790257
2	ARK	-0.901111
3	PUBG	-0.917098
4	golf with your friends	-1.069272
5	Rocket league	-1.074095
6	GTA5	-1.118922
7	Speed runner	-1.203352
8	fallout	-1.213085
9	peach beach splash	-1.332217
10	DBD	-1.41829

表2 テキストデータが類似しており画像データが類似していない時の検索結果(クエリ: Fall Guys)

形態素解析機能, 抽出機能からなる.

3.3 画像データからの特徴量抽出

画像データからの特徴量抽出は, 画像特徴量, ベクトル変換機能, 主成分分析機能からなる.

3.4 テキストデータと画像データの特徴量に基づく類似度計量方式

特徴量抽出を行なった全てのテキストデータおよび画像データをベクトルと見なし, ベクトル同士の \cos 類似度計量することでビデオゲーム間の類似度を求める. 3.4.1 節と, 3.4.2 節には, これらの類似度の統合方式として類似度合算方式, 類似度高低組み合わせ方式について示す.

3.4.1 類似度合算方式

画像データとテキストデータが共に類似しているゲームをリコメンドするための計算方式を以下に示す.

$$sim_a = \log_{10} sim_{img} + \log_{10}(5 * sim_{txt} + 0.001)$$

なお, sim_{img} は画像データの \cos 類似度計量で求められた値, sim_{txt} はテキストデータの \cos 類似度計量で求められた値である. テキストデータの \cos 類似度計量では0が出ることがあるため0,001を足しスムージングする. また, 画像データとテキストデータの類似度の違いから, 予備実験によりテキストデータの類似度を5倍することとする.

3.4.2 類似度高低組み合わせ方式

画像データが類似しており, テキストデータが類似していないリコメンドを行うための計算方式を次に示す.

$$sim_{b1} = \log_{10} sim_{img} + 1 - \log_{10}(sim_{txt} + 0.001)$$

また, テキストデータが類似しており画像が類似していないリコメンドをするための計算方式を次に示す.

$$sim_{b2} = (1 - \log_{10} sim_{img}) + \log_{10}(sim_{txt} + 0.001)$$

これにより類似していないものと類似しているものの合算をする統合を行うことができる.

4. 実験

表1は両方のデータが類似している時の計算結果である. クエリとして, ATRI というゲームを選択した. 表1より, 最も値が大きいのは peach beach splash である. その次に大きいのは STEINS;GATE である. これらは共に特徴量の \cos 類似度を確認すると, 類似しているビデオゲームとなっている.

図3はテキストデータが類似していて画像データが類似していない時の計算結果である. クエリとして, Fall Guys を選択した. 表2より上位3つのゲームはテキストデータの特徴量類似度が高く画像データの特徴量類似度は低い. よって, 正しくリコメンドできていると考えられる.

5. おわりに

本稿では, 画像データとテキストデータを用いた類似度検索を行うことで自分の好みのゲームと類似したゲームさらに片方のデータは類似していてもう片方は類似していないゲームの検索も可能となる. 本方式を利用することで, ユーザが簡単に見つけられないゲームのリコメンドを実現することができる.

今後の課題として, 画像データ, テキストデータの増加や計算結果の向上などが挙げられる.

参考文献

- [1] 小邦将輝, Lasguido Nio, 平手勇宇, 関洋平, 調理手順テキストと料理画像の特徴量の最近傍探索に基づく重複レシピの検出手法, 電子情報通信学会技術研究報告, IEICE technical report: 信学技報, 118(278), pp.19-24, 2018.
- [2] <https://store.steampowered.com/>