

ジャケットワークを用いた意外性のある情報推薦システムの提案 Proposal of unexpected information recommendation system using jacket work

土田 滉也[†]
Hiroya Tsuchida

松本 幾也[†]
Ikuya Matumoto

和田 雄次[†]
Yuji Wada

1. はじめに

近年インターネットの普及により、オンラインショップが発達し、大量のアイテムを扱っている。多くのアイテムの中から欲しいアイテムを見つけるのは自力で見つけるのは困難である。現在では、ユーザの欲しい商品を見つける手助けとして、様々な推薦システムが利用されている。これまでの推薦システムの評価として、ユーザにどれだけ好みのアイテムを推薦できたかのような精度が重要な指標となっていたが、近年では、精度だけではなく推薦されたアイテムの意外性が着目されている。Web サービスの多くは、ユーザと似たような嗜好を持つユーザが共通して良いと評価をしたアイテムを推薦する協調フィルタリングを採用している。協調フィルタリングでは、似たような嗜好のユーザの行動履歴を用いる為、推薦リストが似通った内容になってしまい、既知であるアイテムが含まれてしまう問題がある。

本論文では推薦システムのこれらの問題を踏まえて、音楽 CD を題材とし、ユーザにとって目新しく、意外性を感じる音楽 CD の推薦システムを提案する。本システムでは、アーティストやその曲をまったく知らない状態で、ジャケットデザインのみで惹かれて商品を買うジャケット買いに着目し推薦を行う。

2. 既存研究

既存の楽曲推薦では、好みが類似する他のユーザが好む楽曲の推薦や、楽曲のソーシャルタグなどのメタデータに基づいた推薦[1]が存在するが、似たような楽曲が推薦されてしまう問題がある。

また意外性の関連研究としては、住元ら[2]は Pixiv[3]において、ユーザが興味のあるイラストタグから意外性のある候補を抽出し、イラスト推薦を行っている。向山ら[4]は楽曲の意外性ある推薦を行うために、楽曲の再生時刻と調性着目し、楽曲におけるセレンディピティ要因を分析している。

本論文では楽曲のメタデータではなく、CD アルバムのジャケットワークを用いることで、ユーザにとって意外性のある音楽 CD を推薦する手法を提案する。

3. 推薦手法

提案する手法は、ユーザが選んだ音楽 CD のジャケットワークを推薦要素とし、類似しているジャケットワークの音楽 CD を推薦する。本論文ではこれをジャケットワーク推薦と呼ぶ。音楽 CD の従来の推薦要素としてアーティスト名、ジャンルがよく使われるが、その情報を元に推薦を

行った場合、同じアーティスト、同じジャンルのアイテムが推薦されやすくなるため、既知であるアイテムが推薦されやすい。ジャケットワークには音楽 CD の雰囲気、アーティストの感性が反映されているので、類似しているジャケットワークのアイテムを推薦することで、アーティストやジャンル情報に捉われず、似たような雰囲気、感性のアイテムを推薦ができることが推察できる。

3.1 事前調査



図 1 調査用プログラム(一部)

ジャケットワークを推薦要素として推薦する際に、はたしてユーザはジャケットワークを見てアイテムを購入したい、聞いてみたいと思うかどうか調査した。図 1 のようなプログラムを作成し、アイテム 1500 個の中からランダムで 21 件アルバム画像のみを表示し、ユーザに対してこのアイテムは知っているか、気になるか(購入したい、聞いてみたい)どうかを回答してもらった。

3.1.1 調査結果

協力者 40 人に調査を行った。今回の調査ではユーザが知らないかつ気になるか評価したアイテムを意外性を感じるアイテムとする。21 個のアイテムの中から 1 人当たり平均 2.3 個の意外性を感じるアイテムがあったという結果であった。結果からユーザはジャケットワークからアイテムを購入したい、聞いてみたいと思うことが分かった。

3.2 類似計算

ジャケットワーク間の類似計算を行うためジャケットワーク画像間の類似計算を行う。類似画像間の検索の手法として画像に紐づけられているテキスト情報を元に計算を行う TBIR(Text Based image Retrieval)、画像の特徴量を基盤として計算 CBIR(Content Based image Retrieval)がある。今回はジャケットワーク画像のみで類似アイテムを計測するため CBIR の手法を使う。ジャケットワーク画像の特徴量を抽出し、抽出した特徴量をアイテム間で計算し類似アイテムを選定し、推薦を行う。

[†] 東京電機大学大学院
Tokyo Denki University

3.3 特徴量抽出

ジャケットワーク画像の特徴量の抽出をするため、畳み込みニューラルネットワーク(CNN)を利用する。CNNとは数多くの提案されている深層学習のひとつである。図2に示すように、画像の局所的な特徴抽出を担う畳み込み層と、特徴ごとに特徴をまとめあげるプーリング層を繰り返した構造となっている。直観的には、入力解像度を少しずつ落としながら異なるスケールで隣接する特徴の共起をとり、識別に有効な情報を選択的に上層へ渡していくネットワークである。CNNは多くのタスクで驚異的な性能を達成しており、特に、ImageNetに代表される物体認識データセットを用いて学習させたCNNの中間層から抽出される特徴量は非常に汎用性が高くさまざまなタスクで利用可能である。このことからジャケットワークにも利用可能であると推察できる。

既に大規模のデータセットで学習済みのモデルを特徴量抽出のみに利用する。利用するモデルはAlexnet[5]を利用する。抽出方法としては、入力画像をフィードフォワードし、適当な中間層の出力する値の抽出を行う。CNNでは、入力から近い層から識別層へ近づくにつれて、意味的な特徴に構造化することが知られる。今回は中間層の6層目の4096次元ベクトルを特徴量として抽出を行う。

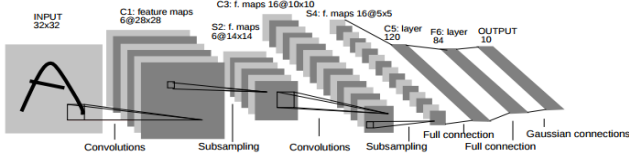


図2 畳み込みニューラルネットワーク図は[7]より引用

3.4 類似計算

特徴量の4096次元のベクトルをユークリッドノルムが1になるように正規化を行い、全画像間のユークリッド距離で計算を行い、距離が近いアイテムを推薦アイテムとする。

$$d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

d は p と q 間の距離を表す。特徴量が4096次元であるため n が4096の総和を求める。

3.5 ジャケットフュージョン

ユーザの好みのジャケットワーク画像から新たなジャケットワーク画像を生成し、生成した画像から類似アイテムを計算する手法である。ジャケットワーク推薦を行いユーザが好みと判断したアイテムを行動履歴として蓄積し、蓄積したアルバム画像を利用しユーザの好み反映された画像の生成を行う。選んだアイテムの好み特徴をとらえた新たな画像を生成しそのアイテムで類似アイテムを選ぶこ

とで、ユーザのジャケットワークの好み反映されることが推察できる。

画像の生成には、DCGAN(Deep Convolutional Generative Adversarial Networks)[6]のように、選んだ画像の共通した特徴を学習させその学習したモデルを使って、特徴が反映された画像を新たに生成する。DCGANの事例としては図3のような寝室画像の自動生成や、顔の表情変化などが挙げられる。

このDCGANを応用し、ユーザが好みと判断したアルバム画像間の特徴を学習させその特徴が反映された画像を生成し、生成した画像から全アイテムとの類似計算を行う。ジャケットワークの好みは感性によるものなので、好みの基準が曖昧である。この手法を使うことでユーザが潜在的に好みと判断する共通した特徴を掴むことができる可能性が高い。



図3 DCGANによる寝室画像生成 図は[6]から引用

4. おわりに

本論文では、推薦アイテムに目新しさ、意外性がない問題を解決するために推薦するために、ジャケットワーク推薦を提案した。ジャケットワーク画像を推薦要素に用いることによって、画像の雰囲気、感性による推薦が可能になる。感性、雰囲気は音楽ジャンルのように自然言語では表現するが困難であるため、既存の手法より意外性のあるアイテムが推薦できることが期待できる。

今後の展望として、提案したジャケットワーク推薦システムの構築を行い意外性の評価を行う。同時にジャケットワークで選んだCDの楽曲は好み合っているかの検証を行いジャケット買いの有用性の証明を示す。

参考文献

- [1] 吉井和佳, 後藤真孝, 音楽推薦システム, 情報処理 50(8), 751-755, 2009
- [2] 住元宗一郎, 中川博之, 田原康之, 大須賀昭彦, コンテンツ投稿型SNSにおける未知性と意外性を考慮した推薦エージェントの提案, 電子情報通信学会論文誌(D), Vol. J94-D, No. 11, pp. 1800-1811, 2011
- [3] Pixiv, <https://www.pixiv.net/>
- [4] 向山輝, 岸上順一, 再生時間に着目した音楽プレイリストのセレンディビティ要因の分析, 情報処理 509-510, 2016
- [5] AlexNet, <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>
- [6] Alec Radford, Luke Metz, Soumith Chintala, Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks
- [7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition. Proc. of the IEEE, pages 2278-2324, 1998.