

## 印象語フィルタを用いた映画推薦方式

## Movie Recommendation Method Using Filter of Impression

有本 裕亮†  
Yusuke Arimoto芋野 美紗子‡  
Misako Imono土屋 誠司‡  
Seiji Tsuchiya渡部 広一‡  
Hirokazu Watabe

## 1. はじめに

近年では、スマートフォンの普及や動画配信サービスの普及に伴い、気軽に映画を見れる環境が整ってきている。しかし映画は膨大な量が存在しているため、自分の嗜好に合った映画を選ぶことは困難である。この状況を解決するため、映画推薦システムの開発が行われている<sup>[1]</sup>。

従来の映画推薦システムでは、ジャンルや出演者、監督などの要素を用いて推薦を行うが、映画の内容分析は浅い。そこで本研究では、新たな要素を追加し、より内容分析を深く行う映画推薦システムの構築を目的とする。

新たな要素として注目したのは、映画がもつ印象である。人は、視聴する映画を決定する際、「この映画は切なそう」のように映画の印象を想像する。この印象は、映画の要素として扱えると同時に、視聴する映画を決定する際の要因になると考えた。そこで本稿では、映画の要素に印象を追加した映画推薦方式を提案する。

## 2. 印象語フィルタ

印象を映画の要素として扱うために、映画のあらすじから印象を抽出する印象語フィルタを構築した。

本稿では、映画の印象を印象語を用いて表現する。例えば、「天空の城ラピュタ」の印象を、「懐かしい」「儂い」という印象語を用いて表現する。本稿の印象語とは「映画の印象を表すのにふさわしい形容詞」であり、先の例の他には「切ない」や「残酷な」などの形容詞がある。

印象語フィルタは、映画のあらすじから印象語を抽出する。抽出の流れを図 1 に示す。

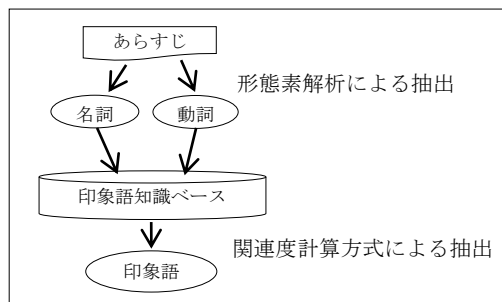


図 1 印象語抽出の流れ

まず、映画のあらすじ文を形態素解析し、名詞と動詞を抽出する。次に、抽出した名詞・動詞と意味の近い印象語を印象語知識ベースから抽出する。意味の近さを計算する際には、関連度計算方式<sup>[2]</sup>を用い、求めた関連度の平均値を印象語の重みとして取得する。

関連度計算方式は、概念ベース<sup>[3]</sup>を基に、概念間の関連の

† 同志社大学大学院理工学研究科

Graduate School of Science and Engineering, Doshisha University

‡ 同志社大学理工学部

Faculty of Science and Engineering, Doshisha University

強さを定量的に評価するものである。関連度の値は 0.0~1.0 の実数を取り、1.0 に近づくほど関連が強いことを意味する。

また、印象語知識ベースには、「映画の印象を表すのにふさわしい形容詞」が格納されている。格納されている形容詞は、TSUTAYA on-line<sup>[4]</sup>内の映画レビューで使われている形容詞のうち、大学生 12 名にアンケートを行った結果、被験者が見たことのある映画の印象を表現するために多く使用された形容詞である。

## 3. 映画知識ベース

映画の要素をシステムで扱うため、TSUTAYA on-line を参考に映画知識ベースを構築した。映画知識ベースは、映画 135 作品それぞれについて、タイトル、制作年度、監督、出演者、脚本、制作国、受賞歴、ジャンル、印象が格納されている知識ベースである。印象については、2 章で説明した印象語フィルタを用いて抽出した印象語と重みのセットを格納している。表 1 に映画知識ベースの一例を示す。

表 1 映画知識ベースの例

タイトル	製作年度	監督	...	ジャンル	印象
天空の城ラピュタ	1986年	宮崎駿	...	ファンタジーアニメ	懐かしい(0.069) 若い(0.007) :

## 4. 映画推薦方式

映画がもつ印象を考慮した推薦方式を実現するために、映画知識ベースを用いて、ユーザの視聴履歴を基に映画を推薦するシステムを構築した。本システムで対象とした 135 本の映画の中で見たことのある映画を入力すると、5 本の映画が出力されるというシステムである。5 本の映画は、視聴履歴から抽出したユーザの嗜好情報を基に、未視聴映画に点数付けを行うことで決定する。

## 4.1 嗜好情報の抽出

視聴履歴は、ユーザが見たいと思った映画の集まりであるため、映画を推薦するための嗜好情報が含まれていると考えた。そこで、ユーザが見た映画それぞれの要素を映画知識ベースから取得し、要素ごとに嗜好情報を抽出する。それぞれの要素について、嗜好情報の扱い方を説明する。

制作年度については、見た映画の平均制作年度を嗜好情報とする。受賞歴については、見た映画のうち受賞歴のある映画の割合を嗜好情報とする。受賞歴のある映画の割合が高いほど、ユーザは受賞歴のある映画を好むと考える。監督、出演者、脚本、制作国、ジャンル、印象については、映画知識ベースから獲得された回数が多いものを嗜好情報とする。また、獲得回数の多さを考慮するため、見た映画から獲得されたものすべてにおける割合も数字として獲得する。例えば、制作国の嗜好情報と

して、日本(12.143)アメリカ(10.714)と抽出された場合、ユーザは洋画よりも邦画を好むと考える。

#### 4.2 未視聴映画への点数付け

本システムは、ユーザがまだ見ていない映画の中から、ユーザの嗜好に最も近い映画を推薦する。そこで、先ほど抽出した嗜好情報を基にユーザの未視聴映画へ点数付けを行う。

映画の点数は、映画の要素それぞれについて、嗜好情報と比較して得点を求めた後、全ての要素の得点を足し合わせることで決定する。それぞれの要素について、点数の求め方を説明する。

製作年度については、嗜好情報である平均製作年度と映画の製作年度との差の絶対値を点数とする。受賞歴については、未視聴映画に受賞歴があれば、嗜好情報として求めた数字を受賞歴の点数とする。

監督、出演者、脚本、製作国、ジャンル、印象については、映画に格納されているものと嗜好情報として獲得したものが表記一致した場合に、嗜好情報として求めた数字を加算していき、点数を決定する。ただし、印象については印象語それぞれに重みを付与しているため、嗜好情報として求めた数字と印象語の重みを掛け合わせて点数を決定する。

以上のようにして求めたそれぞれの要素の得点を足し合わせて映画の得点が決定する。また、それぞれの要素の影響をそろえるために、要素の点数は0.0~1.0になるように調整した。本システムでは、得点の高い5作品を出力し、映画の推薦が行われる。

### 5. 評価実験

映画の要素に印象を追加した映画推薦方式の有効性を示すために、構築したシステムを使って評価実験を行う。点数付けに使用する要素を変えて、推薦の結果を比較する。

#### 5.1 映画評価データ

今回行った評価実験には、インターネット上で817人に対してアンケートを行い取得したデータを使用した。このデータには、本稿対象の135本の映画それぞれに表2の5段階の評価が817人分記載されている。本実験では、5段階評価の①②が付与された映画を「あるユーザが視聴した映画」、③④⑤が付与された「映画のあるユーザの未視聴映画」として扱う。

表2 選択肢の内容

①	過去(2年より前)に観た
②	最近(2年以内)観た
③	観たことはないが、今後観てみたい
④	観たことはなく、今後観てみたいか観てみたくないか分からない
⑤	観たことはないし、今後も観てみたいとは思わない

#### 5.2 実験方法

映画評価データを用いて、①②を付与した映画を入力すると、③④⑤が付与された映画から選ばれた5本の映画が出力される、という映画推薦を817回行った。

本稿では、③が付与された映画を「推薦されるべき映

画」、④が付与された映画を「推薦されてもいい映画」と考えた。そこで、推薦の精度は、システムで推薦された未視聴映画のうち、③の割合(精度A)、③④の割合(精度B)の2種類で評価する。映画評価データを用いて、点数付けを行わずランダムに推薦すると、精度Aは27.12%、精度Bは54.45%となる。

### 5.3 実験結果

結果として、点数付けの際に印象を使用することで、推薦の精度がわずかに向上した。実験の結果を表3に示す。システム1の精度は、映画の要素として制作年度、監督、出演者、脚本、制作国、受賞歴、ジャンルを使用して推薦を行った場合の精度である。システム2は、システム1で使用した要素に印象を追加した場合の精度である。

表3 実験結果

	ランダム	システム1	システム2
精度A	27.12%	28.25%	29.57%
精度B	54.45%	58.65%	59.14%

### 6. 考察

「難しい」映画を好むユーザへの映画推薦結果を見ると、システム1では「チャーリーとチョコレート工場」を推薦したが、システム2では「マトリックス・レボリューションズ」を推薦した。そこで、システム2の構築により、映画の印象を考慮した推薦方式を実現できたと考える。

また、表2の結果から、映画の要素に印象を追加したほうが精度が向上することがわかる。精度向上の原因として、映画の要素を増やすことで、映画の内容をより深く分析できたことと、人は見る映画を決める際の要素として、印象を考慮していることが考えられる。

### 7. おわりに

本稿では、従来から映画の要素としているジャンルや出演者、監督などに加え、印象も考慮した映画推薦方式を提案した。提案推薦方式の有効性を確認するため、視聴履歴を基に、映画の内容分析とユーザの嗜好情報抽出を行い、映画を推薦するシステムを構築した。実験の結果、印象を考慮することで映画推薦システムの精度がわずかに向上することが確認できた。

#### 謝辞

本研究の一部は、科学研究費補助金(若手研究(B)24700215)の補助を受けて行った。

#### 参考文献

- [1] 小野智弘, 黒川茂莉, 本村陽一, 麻生英樹, “ユーザ嗜好の個人差と状況差を考慮した映画推薦システムの実現と評価”, 情報処理学会論文誌, Vol.49, No.1, pp.130-140, 2008.
- [2] 渡部広一, 奥村紀之, 河岡司, “概念の意味属性と共起情報を用いた関連度計算方式”, 自然言語処理, Vol.13, No.1, pp.53-74, 2006.
- [3] 奥村紀之, 土屋誠司, 渡部広一, 河岡司, “概念間の関連度計算のための大規模概念ベースの構築”, 自然言語処理, Vol.14, No.5, pp.41-64, 2007.
- [4] TSUTAYA online, <http://www.tsutaya.co.jp/index.zh.html>