

7ZB-01

動画メディアコンテンツを対象とした人物表情認識を用いた感情重要キーワード抽出方式

吉井 茜音[†] 岡田 龍太郎[†] 峰松 彩子[†] 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、YouTube を代表とした動画像メディアコンテンツ (以下動画) を扱う Web サイトが増えており、このことから動画は、私たちの生活に対して深く関わりを持っていると言える。

本稿では、動画像メディアコンテンツを対象とした人物表情認識を用いた感情重要キーワード抽出方式について示す。一般的に、動画の内容についてのキーワードを抽出するために、その動画内に出現する人物の発言内容を取得することが必要である。動画の中で、どの部分が重要であるかの推定と、その時の発言内容の抽出は難しいと考える。本稿では、動画内に出現する人物の表情認識を実現し、大きく表情が表出した際に発言した内容が動画の重要となる内容であるとし、表情認識の結果とともに、発言をテキストデータとして抽出する手法を提案する。本方式により、各動画の感情とその内容を同時に取得することが可能となり、検索するための感情重要キーワードを抽出することが可能となる。

2. 関連研究

浅田ら [1] は、動画に対して画像解析を用いて、シーンごとに分割したのちに、分割されたシーンの感情抽出を行い、定義した結合規則に沿ってシーンを並び替えることにより異なったストーリーを構築するシステムの開発を実現している。

本方式は表情認識によって得られた感情の数値を用いて、感情があると認識されたキーワードを動画の感情重要キーワードとして出力することを試みている。

3. 提案方式

3.1 システム全体像

本提案方式の全体図を図 1 に示す。本方式は、字幕抽出機能、感情抽出機能、感情字幕統合機能、単語重み付け統合機能、形態素解析機能、感情重要キーワード抽出機能からなる。

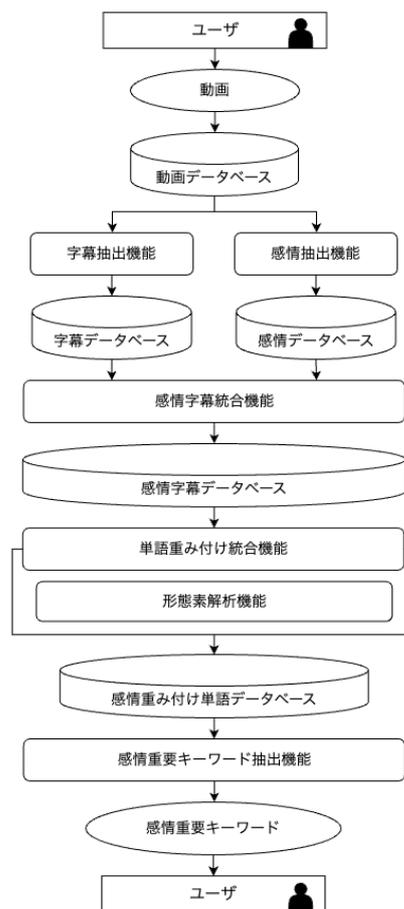


図 1: 感情重要キーワード抽出方式の全体図

3.2 字幕抽出機能

字幕抽出機能は、動画の字幕を時系列ごとに取り得る機能である。ここでは、字幕データベースとして扱う。

3.3 感情抽出機能

感情抽出機能は、FER(Face Emotion Recognizer)を用いて動画内の表情の感情分析をする機能である。本機能を用いることで、表情を angry, disgust, fear, happy, sad, surprise, neutral の 7 種類の感情にそれぞれ数値化し、取得することができる。これらの感情 7 種類の数値の合計は 1.0 である。neutral は感情が表出していない部分であるが、

An Emotional Keyword Extraction Method by Using Facial Expression Recognition for Video Media Contents
 Akane Yoshii[†], Ryotaro Okada[†], Ayako Minematsu[†], Takafumi Nakanishi[†]

[†] Department of Data Science, Musashino University

その他 6 種類は感情が表出している部分である。従って、neutral 値が低いほど、その他 6 種類の感情が表出していると考えられる。ここでは neutral の値を感情データベースとして扱う。

3.4 感情字幕統合機能

感情字幕統合機能は、字幕データベースと感情データベースを用いて、時系列ごとに2つのデータベースを統合させる機能である。時系列ごとに統合させたデータベースを、感情別字幕データベースとして扱う。

3.5 単語重み付け統合機能

単語重み付け統合機能は、形態素解析機能からなる。形態素解析機能は、感情別字幕データベースの字幕から、MeCab を用いて単語を抽出する機能である。単語重み付け統合機能は、感情字幕データベースの単語ごとの感情の数値を集計する機能である。感情の数値によって重み付けされた単語のデータベースを感情重み付け単語データベースとして扱う。

3.6 感情重要キーワード抽出機能

感情重要キーワード抽出機能は、感情重み付け単語データベースから感情重要キーワードを抽出する機能である。各単語の neutral の値を、降順に並び替えることで、感情が表出している時の単語を求められることができる。その際、複数出ている単語もあるため、それらの単語は neutral の値を平均して単語の重み付けをしている。ここでは、neutral の値が低い単語を感情重要キーワードとして扱う。

4. 実験

提案方式の有効性を検証する実験を行う。動画に対して提案方式を適用し、感情重要キーワードを抽出する。また、感情の大きさにあたる neutral 値を時系列順に表示し、動画内で感情が大きくなっている箇所を可視化する。実験の流れとしては、まず、事前に動画内のどこで感情が大きくなっているかを確認しておく。その後、提案方式によって感情が大きいとされる箇所がそれに一致し、またそこに出現する単語から感情重要キーワードが抽出されていることを確認することで、本方式の有効性を検証する。

実験として入力する動画として YouTube に公開されている動画[2]を用いた。これは医療ドラマに関係する動画であり、2人の看護師が主人公である。このドラマは1人の看護師が、もう1人の看護師を一人前に育てるために、互いにぶつかり合いながらも成長していく物語である。ここでは動画内で一番感情が大きくなる場所は、2人が喧嘩する場面であると定義する。

3節で提案した方式を実装し実験を行った。感

表 1: 感情重要キーワード抽出結果

順位	キーワード	Neutral 値
1	お願い	0.02729
2	結果	0.02729
3	自業自得	0.04842
4	過ち	0.05
5	過去	0.05

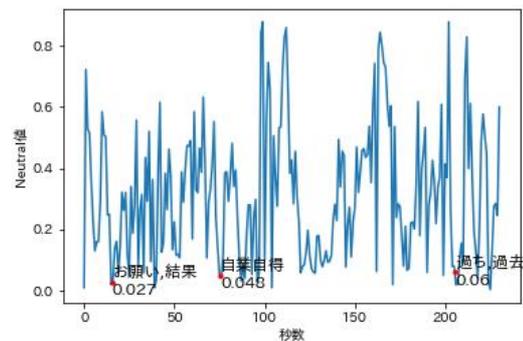


図 2: neutral 値の時系列変化

情重要キーワードの抽出結果を表 1 に示す。また、動画から抽出された感情重要キーワードを、neutral 値の時系列変化と共に図 2 に示す。

図 2 より、感情重要キーワードには「自業自得」や「過ち」といったネガティブな印象を持つ単語がある。実際、これらの単語が使われている場面は主人公の2人が喧嘩している場面であり、感情が一番大きくなったと定義した場面の単語から感情重要キーワードを抽出したことが確認された。一方、「結果」、「過去」など、あまり感情に関係なさそうな言葉も感情重要キーワードとして現れたが、そのような単語も抽出することができる。

5. 終わりに

本稿では動画を対象とした人物表情認識を用いた感情重要キーワード抽出方式を示し、新たなキーワード抽出方式を示した。

今後の展望としては、分析するデータ量を増やし、動画同士の感情重要キーワードの類似度を測ることにより、ユーザに対して、感情の表れ方の似ている動画の検索やレコメンドが可能となる。

参考文献
[1] 浅田裕也, 富士本達矢, 濱川礼, 動画からの感情抽出および感情遷移によるストーリーの自動構築, 全国大会講演論文集データベースとメディア, pp. 719-720 (2008)

[2] テレビ朝日: 【ザ・トラベルナース】今からでも間に合う! ダイジェスト, YouTube, https://youtu.be/Dh_68REfNzU (参照 2023-01-11)