

異常値検出を用いた単語の想起イメージの文化差検出

西村 一球[†] 村上 陽平[†] Mondheera Pituxcoosuvam[†]

立命館大学 情報理工学部 社会知能研究室[†]

1 はじめに

近年、言語や文化の違いを超えた異文化コラボレーションが求められている。こうした能力を育む地球市民教育の実現が SDGs の一つとして規定されている。このような異文化コラボレーションでは、近年、機械翻訳の品質向上により多言語コミュニケーションの言語の壁は解消されつつあるが、依然として文化差による誤解が生じている。例えば、日本でよく食べられる“ゴボウ”を機械翻訳で英訳すると“burdock”という結果が得られる。日本人の多くは“ゴボウ”から根菜を想起するが、海外では“burdock”からは赤い花をつける植物を想起することが多い。これらは実は同一の植物で“ゴボウ”は“burdock”の根の部分である。日本人は“ゴボウ”を食べる文化があるが、海外では“ゴボウ”を食べる文化がないため、想起するイメージに齟齬が生まれる。

2 関連研究

Pituxcoosuvam らによる研究では、Web 上の画像をもとに日英間で画像の特徴ベクトルの類似性を算出し、最終的に、人の主観で文化差を判定する手法を提案している[1]。柳生らは写真共有サイト上のタグ付き画像を利用して、概念の代表的な写真をタグの地域ごとに分類し、同一概念の対象物が地域ごとに文化差があるかを視覚化している[2]。Cho らは絵文字に着目し、アメリカ人と日本人に 120 種類の絵文字の印象についてアンケート調査したところ、文化の違いによって解釈のされ方が異なることを明らかにしている[3]。神田らは、アジアと欧米 8 カ国間でアバタの表情の解釈内容を比較したところ、肯定的な内容の解釈に文化差が存在することを発見している[4]。これらの既存研究では、文化差の存在を明らかにするものやその検出手法が提案されているが、最終的には人手で文化差の有無を判断する必要があるため、自動的な文化差検出手法が必要である。

3 特徴ベクトル間の類似度算出

文化差を自動的に検出するために同一概念に紐づけられた日英の単語から画像検索を行い、取得された画像の類似性に基づいて文化差検出を行う。具体的には、WordNet と日本語 WordNet という概念辞書を用いて、概念辞書上で同一概念に紐づけられた日英の単語群を取り出し、各言語の単語群で AND 検索を行う。日本語の単語群で取得し

た画像の上位 10 枚と英語の単語群で取得した画像の上位 10 枚から類似度を算出する。

ここで Keras の VGG16 を用いて、取得したそれぞれ 10 枚の画像から特徴ベクトルを抽出し、言語ごとに 10 枚分の画像ベクトルを平均して各言語の平均特徴ベクトルとする。その後、平均特徴ベクトル同士でベクトル間の類似度を Cos 類似度によって算出し文化差を判定する。

この時、10 枚の画像の中に対象概念と異なる全く関係のない画像が含まれていた場合、正確に文化差検出ができない可能性がある。そのため、異常値検出を用いて、取得画像からターゲット画像とノイズ画像に分類し、ノイズ画像の影響が小さくなるように重み付けした加重平均によって各言語の平均特徴ベクトルを生成する。今回、異常値検出の手法として、K-means, DBSCAN, Isolation Forest の 3 つの手法を用いた。

4 閾値を用いた文化差判定

算出した類似度はあくまで連続値であるため、文化差を判定の基準となる閾値を設定する。閾値を類似度が下回った場合、日英の画像同士は類似性が低いと判断して文化差ありと判定する。上回った場合、類似性が高いとして文化差なしと判定する。自動的に文化差を検出するためにこの閾値を適切に設定する必要がある。

まず、閾値を最適化するために、文化差判定の正確さ(Accuracy)を指標とする。Accuracy とは人手で文化差があると判定された概念のうち、人と同様の文化差判定をできた割合である。

$$Accuracy = \frac{\text{人と同じ文化差判定ができた数}}{\text{文化差を判定した数}} \quad (1)$$

今回は、人手によって文化差がないと判定された概念 250 個と文化差があると判定された概念 250 個の合計 500 個の概念を用いて予備評価を行い、文化差判定に最適な閾値を求める。

まず、これら 500 個の概念を 100 個ずつのグループに 5 分割し、バリデーショndata とテストデータが 4:1 になるように設定する。このうち、バリデーショndata に対して、本手法を用いて文化差判定を行う。閾値を 0.0~1.0 までの 0.05 ずつ変動させ、Accuracy が最も高かった閾値を文化差検出に最適な閾値とする。評価実験では、テストデータに対してこの最適な閾値を設定した本手法で文化差検出を行い、Accuracy で評価

する。また、バリデーションデータとテストデータに分割する際に、データの偏りを防ぐために、5分割交差検証を用いて、各データグループがそれぞれテストデータになる組み合わせを検証し、それらの結果の平均をもとに評価を行う。

5 評価実験

表1はバリデーションによって見つけた最適な閾値を使い、テストデータを文化差判定した時のAccuracyとその時の閾値である。また、文化差データはテストデータのうち、文化差ありデータのみに対してのAccuracyと文化差なしデータのみに対してのAccuracyである。バリデーションの結果、DBSCANでは最適な閾値が0.5と0.55の2つあったため、両方での結果を載せている。

手法は、K-means, DBSCAN, Isolation Forestのそれぞれを用いて、加重平均で平均特徴ベクトルを生成する3つの手法とそれらを用いず、単純平均で平均特徴ベクトルを生成する手法を合わせた4つの手法である。これら4つの手法を比較したとき、63.0%で単純平均手法が最もAccuracyが高かった。しかし、文化差データに着目した場合、どの手法も閾値0.55の場合、文化差ありデータに対するAccuracyは68.0%と同じであった。

表1: 評価実験の結果

| 手法 | 閾値 | Accuracy (%) | | |
|------------------|------|--------------|------|-------|
| | | 文化差データ | | Total |
| | | あり | なし | |
| 単純平均 | 0.55 | 68.0 | 58.0 | 63.0 |
| K-means | 0.55 | 68.0 | 54.0 | 61.0 |
| DBSCAN | 0.5 | 54.0 | 64.0 | 59.0 |
| | 0.55 | 68.0 | 50.0 | |
| Isolation Forest | 0.55 | 68.0 | 56.0 | 62.0 |

6 考察

単純平均では文化差検出に成功したが、3つの加重平均手法では、検出に失敗した例として、(日)“薬指”, (英)“ring finger, annualry”という単語が属している概念がある。これは人手では文化差なしとされているデータであるが、単純平均では、類似度:0.582なのに対して、他の類似度は(K-means)0.509, (DBSCAN)0.508, (Isolation Forest)0.544で文化差判定に失敗している。取得された画像では、日英どちらも10枚とも人の手の写真が取得されている。このような、ノイズが入っていない画像群に各異常値検出手法で重み付けを行ったことにより、平均特徴ベクトルの類似性が低くなったと考えられる。今後は異常値検出の結果、ノイズ画像が含まれていない場合は加重平均ではなく単純平均で平均特徴ベクトルを

生成するといった改善が必要である。

7 おわりに

多言語コミュニケーションでの文化差の問題を解消するために、私たちは画像特徴ベクトルの類似度に基づく手法を提案してきた。問題点として、ノイズ画像が含まれる可能性があるため、異常値検出を用いてノイズデータに影響を受けない文化差検出手法を提案してきた。その結果、異常値検出を用いた手法では、Accuracyが6割程度であった。今後の改善として、ノイズ画像が含まれない場合には加重平均ではなく、単純平均を用いることで更なるAccuracyの向上を図る。

謝辞

本研究は、日本学術振興会科学研究費基盤研究(B)(21H03561, 令和3年度～令和6年度, 21H03556, 令和3年度～令和5年度)および日本学術振興会科学研究費若手研究(21K17794, 令和3年度～令和6年度)の支援を受けた。

参考文献

- [1] pituxcoosuvam, M., Murakami, Y., Lin, D. and Ishida, T.: Effect of Cultural Misunderstanding Warning in MT-Medical Communication, Proceeding of the International Conference on Collaboration Technologies and Social Computing, Springer, pp. 112-127 (2020)
- [2] Yanai, K., Yaegashi, K. and Qiu, B.: Detecting cultural differences using consumer-generated geotagged photos, Proceedings of the 2nd International Workshop on Location and the Web, pp. 1-4 (2009)
- [3] Cho, H., Ishida, T., Yamashita, N., Inaba, R., Mori, Y. and Koda, T.: Culturally-situated pictogram retrieval, International Collaboration, Springer, pp. 221-235 (2007)
- [4] Koda, T.: Cross-cultural comparison of interpretation of avatars' facial expressions, Proceedings of the IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06) (2006)