

絵文字チャットによる異文化コミュニケーションの円滑化に関する研究

中村健二† 田中成典‡ 細島啓史† 吉村智史† 北野光一† ケイコウ‡

関西大学大学院総合情報学研究科† 関西大学総合情報学部‡

1. はじめに

近年、急速な国際化に伴い、異文化コミュニケーションの機会と重要性[1]が増加している。しかし、異文化環境における円滑なコミュニケーションのための外国語の完全な習得は困難である。そのため、文法の理解を必要とせず直感的に理解できる絵文字を用いた、異文化コミュニケーションを支援する研究が盛んに行われている。中でも PictNet[2]という絵文字を使用したコミュニケーションツールは、世界的なコミュニケーション支援プロジェクトに使われており、今後のグローバルコミュニケーションの流れにおいて、重要な役割を担うことが期待されている。また、絵文字を用いたコミュニケーションツールに関する既存研究[3]として、予め用意した 550 個の絵文字群から絵文字を選択し、絵文字のみで作成した文章でコミュニケーションを実施しているものがある。しかし、この既存研究では、ユーザが適切な絵文字を選択する必要があるため、ユーザに多大な負担がかかるという問題がある。そこで、本研究では、自然言語で書かれた文章から絵文字への自動変換を実現することで、ユーザの負担を軽減する手法を提案する。また、類義語辞書を自動構築[4][5]することによって、よりスムーズな会話に対応できるチャットシステムを実現する。

2. 研究の概要

本研究では、特定の言語で書かれた文章を絵文字に自動変換することにより、絵文字のみでコミュニケーションするチャットシステムを提案する。なお、対訳コーパスの更新は自動で行う。本システムの流れを図 1 に示す。本システムは、1) 絵文字チャット機能、2) 対訳コー

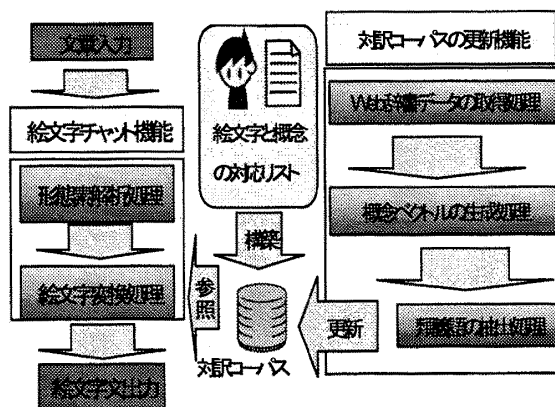


図 1 システムの流れ

パスの更新機能の 2 つの機能により構成される。また、本システムは日本語と中国語のみを対象とする。

2. 1 絵文字チャット機能

本機能では、自然言語で書かれた文章を絵文字へ自動変換する。ユーザが日本語もしくは中国語で入力した文章を形態素に分割する。助詞は絵文字で表現できないため、形態素に分割した結果から助詞以外の品詞を抽出する。なお、本研究では日本語形態素解析のツールに「茶筌」を使用し、中国語形態素解析のツールに「ICTPROP」を使用する。そして、絵文字変換処理では、形態素解析で抽出した単語を対訳コーパスによって絵文字に変換する。

2. 2 対訳コーパスの更新機能

本機能では、ベクトル空間モデル[5]を用いて絵文字と対応する概念を持つ単語の類義語を抽出する。まず、概念の定義を収集するため、Web 辞書データを取得する。Web 辞書とは、インターネット上で利用できる電子化された辞書のことである。次に、取得した Web 辞書データから概念を抽出し、その概念の説明文から概念ベクトルを生成する。概念ベクトルとは、単語同士の共起頻度に基づいてある単語が表す概念をベクトルとして表したものである。概念ベクトルを生成する処理では、概念の説明文を形態素解析し、それぞれの品詞の出現回数を取得す

Research for Supporting Cross-cultural Communication using Pictographic Chat

†Kenji Nakamura, Hirofumi Hosohata, Satoshi Yoshimura, Koichi Kitano

Graduate School of Informatics, Kansai University, 2-1-1 Ryouzenji-cho Takatsuki-shi, Osaka 569-1095, Japan

‡Shigenori Tanaka, Hang Xing

Faculty of Informatics, Kansai University, 2-1-1 Ryouzenji-cho Takatsuki-shi, Osaka 569-1095, Japan

る。最後に、概念ベクトル間の類似度をベクトル空間モデルによって取得し、閾値以上の概念を類義語として抽出する。概念の類似度は、ベクトル空間モデルを用いて、概念ベクトル間のコサイン値により算出する。

3. システムの実証実験と考察

本システムの実行画面を図2に示す。本システムで考案した手法の有効性を検証するために、実験実証を行った。

3.1 実証実験

本実験では、日本人学生と中国人学生それぞれ10名が、実際に提案システムを利用して会話をを行い、その後、アンケートを実施した。また、実験で利用する絵文字は、オフィス・スローライフ社が無料で提供している300組の絵文字を採用した。そして、各絵文字に対応する概念を取得するWeb辞書データとして、情報量の多い「Wikipedia」を利用した。アンケートでは、本システムに対して表1に示す6項目について、5を最良とする5段階で評価を求めた。

3.2 結果と考察

アンケート結果の平均値を表1に示す。本システムでは、絵文字文の認知率は全体の70%であった。絵文字の変換精度については、低い評価を得られた。これは、辞書に登録されている絵文字数が不足しているためだと考えられる。対訳コーパス更新機能の有効性についてのアンケート結果も低い結果となった。その原因は、類犠性の低い単語も類義語として抽出されたためである。その解決策として、ユーザによる類義語判別は必要である。「相手と通じ合えたと思いますか」と「今後まだ使いたいと思いますか」のアンケートについては、良い結果となった。なぜなら、絵文字文は自然言語の文書と違って、直感的に理解でき、ニュアンスで相手と通じ合えるためである。以上のことから、本システムが有効であることを実証できた。

4. おわりに

本研究では、ユーザが特定の言語を用いて、外国人とのコミュニケーション可能な絵文字チャットシステムを開発した。そして、実証実験からその有効性を証明した。また、「すごく楽しかったので、また使いたい」という自由記述アンケート結果から、子どもの異文化コミュニケーションへの意欲を引き立たせることによって、より大きな教育効果が得られることも期待できる。今後は、対訳コーパス更新機能の改善、固有名詞の表示や文化の違いによる絵文字の意味に対しての多様な解釈について研究する予定である。

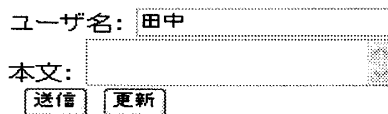


図2 チャットシステムの実行画面

表1 実験結果

項目	平均値
使いやすかったですか?	3
絵文字は正確に変換できましたか?	2
相手の言っていることが理解できましたか?	3
相手と通じ合えたと思いますか?	4
対訳コーパス更新機能は有効だと思いますか?	2
今後まだ使いたいと思いますか?	4

参考文献

- [1] 稲葉利江子, 高崎俊之, 森由美子: 絵文字コミュニケーションにおける類型の比較, 情報科学技術フォーラム, 情報処理学会, Vol.FIT 2006, pp.479-480, 2006.8.
- [2] Takasaki, T.: PictNet: Semantic Infrastructure for Pictogram Communication, The 3rd International WordNet Conference, GWC-06, pp.279-284, 2006.7.
- [3] 藤井薫和, 重信智宏, 吉野孝: 絵文字チャットによるコミュニケーションの提案と評価, 情報処理学会論文誌, 情報処理学会, Vol.47, No.7, pp.2071-2080, 2006.7.
- [4] グエン・ベト・ハー, 帆苺謙, 石川勉, 笠原要: 単語の意味の類似性判別のための大規模概念ベース, 情報処理学会論文誌, 情報処理学会, Vol.43, No.10, pp.3027-3136, 2002.10.
- [5] Salton, G., McGill, M.: Introduction to Modern Information Retrieval, McGraw-Hill, 1983.9.