

科学研究費助成事業 研究成果報告書

令和 2 年 6 月 10 日現在

機関番号：14401

研究種目：研究活動スタート支援

研究期間：2018～2019

課題番号：18H06465・19K21533

研究課題名（和文）マルチモーダル品質推定に基づく機械翻訳モデルの高度化

研究課題名（英文）Machine Translation based on Multimodal Quality Estimation

研究代表者

梶原 智之（Kajiwara, Tomoyuki）

大阪大学・データビリティフロンティア機構・特任助教（常勤）

研究者番号：70824960

交付決定額（研究期間全体）：（直接経費） 2,100,000円

研究成果の概要（和文）：東京五輪に向けて、機械翻訳の精度向上が急務である。本研究では、機械翻訳の自動評価および画像を用いるマルチモーダル機械翻訳に取り組んだ。

前者について、機械翻訳の訓練においては、正解文と表層的に異なる翻訳文は意味的に正しくともペナルティを受けてしまう。そこで、文のベクトル表現を用いる自動評価手法を提案し、正解文の表現への依存を減らしつつ人手評価との高い相関を達成した。

後者について、従来のマルチモーダル機械翻訳では画像を等間隔に分割して利用しており、テキストとの対応付けが難しかった。そこで、意味的な単位に分割した画像を用いるマルチモーダル機械翻訳手法を提案し、性能の改善を確認した。

研究成果の学術的意義や社会的意義

東京五輪や大阪万博に向けて、機械翻訳の精度向上が急務である。機械翻訳モデルの効率的な改善のために、人手評価との高い相関を持つ自動評価手法の開発が重要である。本研究では、単語やフレーズの一致といった局所的な情報に頼っていた自動評価を改善し、文全体の大域的な情報をもとに人手評価との高い相関を持つ自動評価手法を構築した。機械翻訳に関する国際会議WMT-2018において開催された自動評価手法の性能を競うコンペティションにおいては、ドイツ語から英語や中国語から英語などの全7言語対において、我々の提案手法が世界最高性能を達成した。

研究成果の概要（英文）：Improving machine translation performance is an urgent task for the Tokyo Olympics. In this research, we worked on automatic evaluation of machine translation and multimodal machine translation.

In machine translation training, translated sentences that differ superficially from reference sentences are penalized even if they are semantically correct. Therefore, we proposed an automatic evaluation method using vector representation of sentences, and achieved high correlation with human evaluation while reducing the dependence on reference sentences.

In conventional multimodal machine translation, images are divided into uniform sizes, which makes matching with text difficult. Therefore, we proposed a multimodal machine translation method using images divided into semantic units and improved the translation quality.

研究分野：自然言語処理

キーワード：機械翻訳 マルチモーダル

様式 C-19、F-19-1、Z-19 (共通)

1. 研究開始当初の背景

2020年の東京五輪に向けて、機械翻訳の精度向上が急務である。日英や日中の翻訳精度は改善されてきたものの、翻訳器の訓練に必要な100万文対を超える対訳データをあらゆる言語対で用意することは困難である。テキストとテキストの対である対訳データが一部の言語対でしか大規模に利用できない一方で、画像とテキストの対はSNSの普及により多くの言語で大規模に利用できる。そこで本研究では、画像を用いるマルチモーダル機械翻訳に取り組む。

意味と表現が1対1対応しないことは自然言語処理における積年の課題である。機械翻訳においては、翻訳文と正解文の単語単位の一致率を最大化するようにニューラルネットワークを最適化するのが一般的であり、正解文と表層的に異なる翻訳文は意味的に正しくともペナルティを受けてしまう。そこで本研究では、翻訳文の品質評価において正解文の表現に対する依存を減らし、翻訳器への柔軟なフィードバックを行うことを目指す。

2. 研究の目的

- (1) 正解文の表現への依存を減らしつつ、人手評価と高い相関を持つ機械翻訳自動評価を行う
- (2) 画像を用いるマルチモーダル機械翻訳を改善する

3. 研究の方法

(1) 機械翻訳の自動評価

機械翻訳の自動評価における先行研究は、翻訳文と正解文の表層的な語句の一致率に基づく教師なし手法と編集距離や構文解析に基づく素性エンジニアリングを行う教師あり手法に大別できる。教師あり手法が高い性能を発揮するが、これらの訓練に必要な入力文・翻訳文・正解文・翻訳品質の4つ組のデータは数千件しか利用できない。

本研究では、大規模に利用可能な生コーパスのみから訓練できる文の符号化器を用いて、文ベクトルに基づく回帰モデルを構築する。図1に示すように、翻訳文と正解文をそれぞれ文の符号化器によってd次元のベクトルに変換する。そして、これらの文ベクトルから素性抽出を行い、多層パーセプトロン (MLP: multi-layer perceptron) に基づく回帰モデルによって翻訳品質を推定する。素性には、翻訳文の文ベクトル・正解文の文ベクトル・文ベクトル間の要素ごとの差・文ベクトル間の要素ごとの積の4種類を使用し、これらを連結した4d次元のベクトルを用いる。

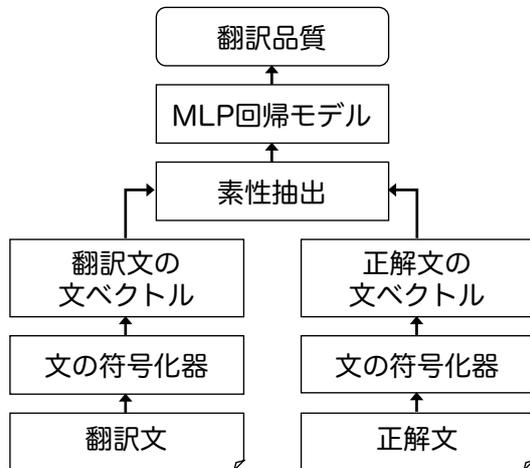


図1: 文ベクトルに基づく自動評価

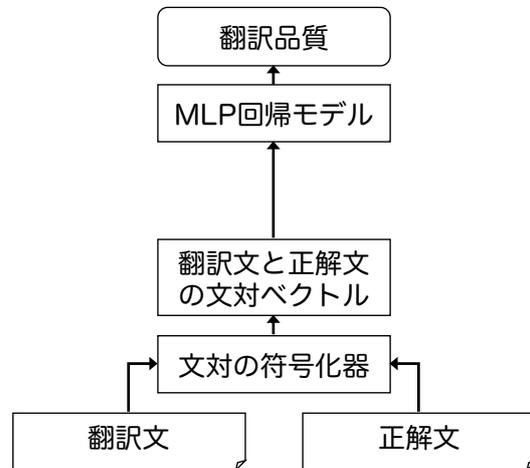


図2: 文対ベクトルに基づく自動評価

文の符号化器の代わりに文対の符号化器を用いても良い。この場合、図2に示すように、素性抽出のステップが不要になり、文対ベクトルを用いて回帰モデルを訓練する。

これらの自動評価モデルを訓練および評価するために、WMT-2017 Metrics Shared Taskのデータセットを用いる。これは、入力文・翻訳文・正解文・翻訳品質の4つ組のデータであり、本研究では入力文以外の情報を用いる。チェコ語・ドイツ語・中国語などの7言語から英語への翻訳に関する約1万文対を用いて評価実験を行う。

(2) マルチモーダル機械翻訳

画像を用いるマルチモーダル機械翻訳において、先行研究では図3のように画像を等間隔に分割して利用していた。このような物体の一部を切り取った画像はテキストとの対応関係の学習が難しいと考え、本研究では物体検出モデルを通して図4のような意味的な単位に分割した画像領域を利用する手法を提案する。画像領域を用いることで、テキスト(語句)と画像(画像領域)の対応関係が明確化でき、翻訳品質の改善に貢献できると期待できる。



図3：等間隔に分割された画像

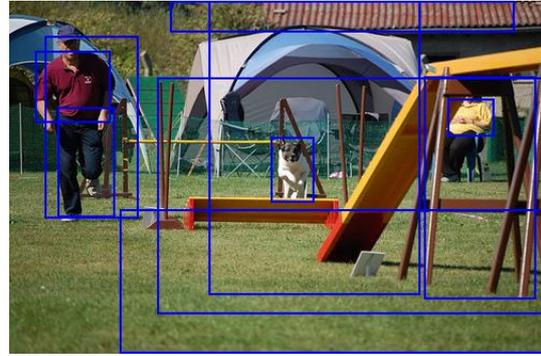


図4：意味的な単位で分割された画像

対応するテキストと画像領域の対応を捉えるために、テキストへの注意と画像領域への注意を備えた図 5 のような注意機構に基づくマルチモーダル機械翻訳のモデルを考える。テキストへの注意は、入力単語の系列を双方向 GRU で符号化し、各単語に対応するベクトルへの注意を張る。画像への注意は、Faster R-CNN および ResNet-101 を用いて入力画像を画像領域に分割し、各画像領域に対応するベクトルへの注意を張る。これらの情報を統合して、機械翻訳を行う。

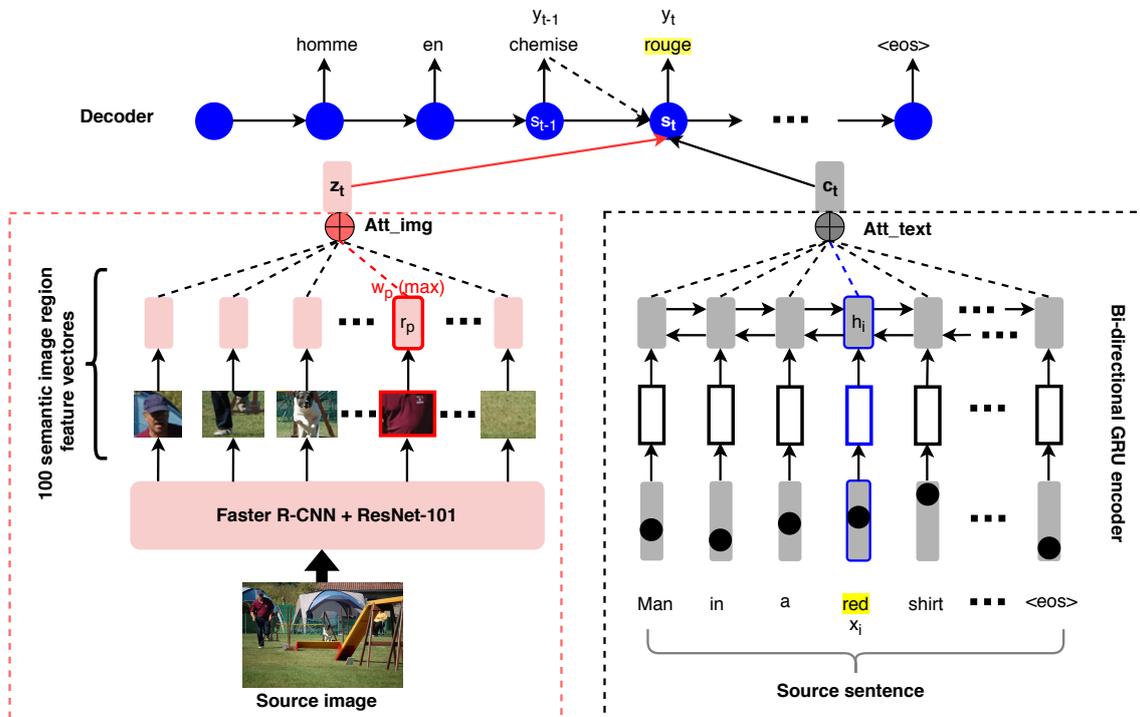


図5：画像領域に基づくマルチモーダル機械翻訳

4. 研究成果

機械翻訳の自動評価に関する実験結果を図 6 に示す。比較手法としては、語句の一致率に基づく SentBLEU および chrF++, 2017 年時点での最高性能を持つ教師あり手法 BLEND の 3 つを載せる。これらの比較手法と、文ベクトルまたは文対ベクトルに基づく提案手法を、人手評価とのピアソン相関係数によってメタ評価する。図 6 の結果は、7 言語対における平均値である。

BLEND は約 30 種類の素性に基づく手法であるが、文ベクトルに基づく提案手法は QuickThought と呼ばれる教師なし手法で訓練した文の符号化器から得た素性のみを用いて、より高い人手との相関を達成した。また、BERT の文対ベクトルに基づく提案手法は、出力文と正解文の単語間の関係をより適切に考慮することができ、さらに大幅に高い性能を達成した。なお、我々は WMT-2018 Metrics Shared Task のコンペティションに参加し、文ベクトルに基づく手法で全 7 言語対において世界最高性能を記録した。

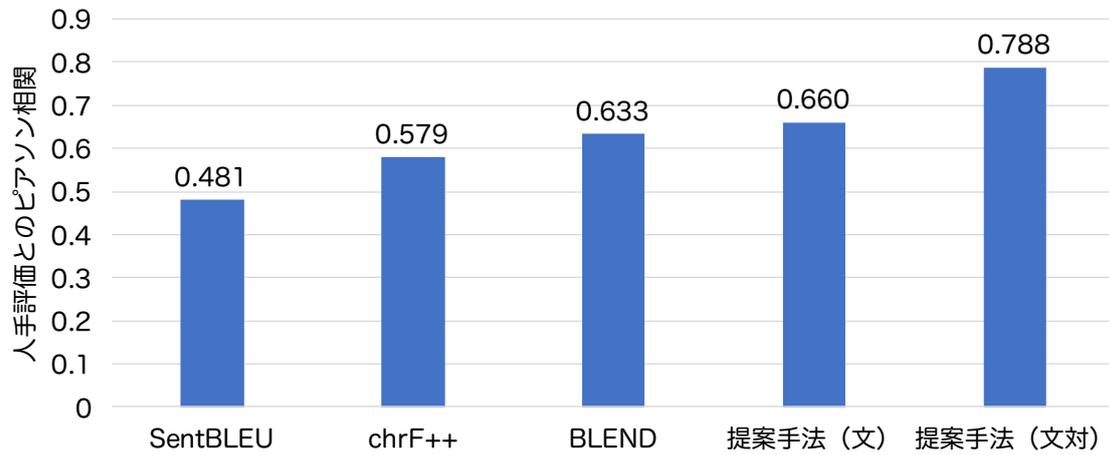


図6：機械翻訳の自動評価モデルのメタ評価

マルチモーダル機械翻訳に関する実験結果を表1に示す。Multi30k データセットにおける英語からドイツ語および英語からフランス語への機械翻訳の実験の結果、意味的な単位に分割された画像を用いる提案手法が、等間隔に分割された画像を用いるベースラインおよび画像を用いないベースラインと比べて、BLEU による自動評価において高い性能を示すことを確認できた。

表 1：マルチモーダル機械翻訳の評価結果

	英語 → ドイツ語	英語 → フランス語
テキストのみを用いる機械翻訳	34.7	56.6
等間隔に分割された画像に基づくマルチモーダル機械翻訳	36.5	57.3
意味的な単位に分割された画像に基づくマルチモーダル機械翻訳	37.0	58.2

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 嶋中宏希, 梶原智之, 小町守	4. 巻 26
2. 論文標題 事前学習された文の分散表現を用いた機械翻訳の自動評価	5. 発行年 2019年
3. 雑誌名 自然言語処理	6. 最初と最後の頁 613 ~ 634
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.5715/jnlp.26.613	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 -

〔学会発表〕 計3件（うち招待講演 0件/うち国際学会 2件）

1. 発表者名 Yuting Zhao, Mamoru Komachi, Tomoyuki Kajiwara, Chenhui Chu
2. 発表標題 Double Attention-based Multimodal Neural Machine Translation with Semantic Image Regions
3. 学会等名 Proceedings of the 22nd Annual Conference of the European Association for Machine Translation (EAMT 2020) (国際学会)
4. 発表年 2020年

1. 発表者名 嶋中宏希, 梶原智之, 小町守.
2. 発表標題 BERTを用いた機械翻訳の自動評価.
3. 学会等名 言語処理学会第25回年次大会
4. 発表年 2019年

1. 発表者名 Hiroki Shimanaka, Tomoyuki Kajiwara, Mamoru Komachi.
2. 発表標題 RUSE: Regressor Using Sentence Embeddings for Automatic Machine Translation Evaluation.
3. 学会等名 Proceedings of the Third Conference on Machine Translation (WMT 18) (国際学会)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----