

クロスドメインマッチング相関分析による画像情報を反映した単語埋め込み

福井 一輝^{1,3}, 押切 孝将^{2,3}, 下平 英寿^{1,3}

¹ 京都大学大学院情報学研究科, ² 大阪大学大学院基礎工学研究科

³ 理化学研究所革新知能統合研究センター

自然言語処理において、コーパス内に含まれる単語の共起関係から単語を数百次元の実ベクトルで表現する手法を単語埋め込みという。単語埋め込みによって得られたベクトル表現 (単語ベクトル) には単語間の類似度が反映されており、それら単語ベクトルに対して機械学習の手法を適用することで自然言語処理の様々なタスクやマルチモーダルデータ分析に応用することができる。

いくつかの代表的な単語埋め込みの手法ではニューラルネットワークが用いられているが、この他にも正準相関分析を用いた Eigenwords [1] という手法も提案されている。Eigenwords では、コーパス内の各単語とそれら単語の周辺コンテキストを、one-hot エンコーディングに基づいた単純な表現に変換し、単語とコンテキストのペアに対して正準相関分析を適用することで単語ベクトルを得る事ができる。

本研究では、Eigenwords における単語のドメインに対して単語と画像間の対応関係を与えることでマルチモーダルなベクトル表現を得る手法、Multimodal Eigenwords (MM-Eigenwords) [2] を提案する。具体的には Eigenwords のモデルにおける正準相関分析を、より一般的なクロスドメインマッチング相関分析 [3] に置き換えることによって、互いに関係のある単語と画像特徴量を共通空間に同時に埋め込むことができる。MM-Eigenwords によって得られた単語ベクトルには、単語間の類似度と単語-画像間の類似度の両方が反映される。また、従来の単語埋め込み手法において確認されている加法構成性 (“king” - “man” + “woman” ≈ “queen”) と似た性質が画像と単語間においても確認された。

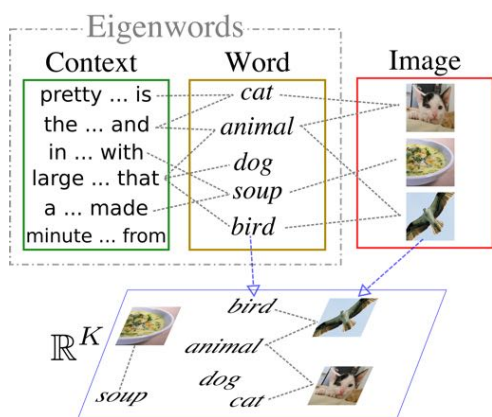


図 1: Multimodal Eigenwords の概略



図 2: 単語-画像間での加法構成性

参考文献

- [1] Paramveer S Dhillon, Dean P Foster, and Lyle H Ungar. Eigenwords: Spectral word embeddings. *Journal of Machine Learning Research*, Vol. 16, pp. 3035–3078, 2015.
- [2] Kazuki Fukui, Takamasa Oshikiri, and Hidetoshi Shimodaira. Spectral graph-based method of multimodal word embedding. In *In Proceedings of the 11th Workshop on TextGraphs in conjunction with ACL 2017*, 2017. (to appear).
- [3] Hidetoshi Shimodaira. Cross-validation of matching correlation analysis by resampling matching weights. *Neural Networks*, Vol. 75, pp. 126–140, 2016.