

平成30年度 情報工学コース卒業研究報告要旨

村瀬（洋）研究室	氏 名	梅 村 和 紀
卒業研究題目	画像キャプションの質的評価に向けた 文の心像性推定	

近年，画像認識技術や自然言語処理技術の発達に伴い，画像キャプション技術は目ざましい発達を遂げている．この技術により自動生成されたキャプション（図 1(A)，(B)）は，画像内の具体的事象の描写能力に長けており，画像内容に基づく検索のための索引付けや映像描写として有効である．しかし，ニュース記事中の画像に対するキャプション（図 1(C)）としてふさわしいとは限らない．そこで，本研究では用途に応じた適切なキャプションの実現を目指し，キャプションが元の画像をどれほど思い浮かべやすいものであるかに着目する．単語に対しては，Paivio ら [1] が心的イメージの喚起しやすさを表す単語属性である「心像性 (imageability)」を定義している．用途に応じたキャプションを実現するための第一歩として，語数による量的な評価だけでなく，心像性による質的な評価をすることが必要である．そこで本研究では，画像キャプションの質的評価に向けて，文の心像性を推定することを目的とする．本論文では，文中の単語の心像性と文の構造に基づいて，文の心像性を算出する手法を提案する．そして，それが人間の感覚に則した指標となっていることを検証する．

本手法では，構文解析により得られる構文木の構造に基づき，葉ノードからボトムアップに単語の心像性スコアを組み合わせることで，文全体の心像性を算出する（図 2）．その際，単語の心像性として既存の心像性データベースを利用する．クエリの単語に対し，小文字化，ステミング，レンマ化の正規化処理を順に施した上で検索する．

提案手法の妥当性を確認するため，推定した文の心像性の評価実験を行なった．比較手法として，文を構成する全ての単語の心像性の平均値による推定と，最大値による推定を行なった．実験方法は，被験者に心像性の定義を示した上で，本実験のために作成した英文対を提示し，「どちらの文の方が心像性が高いと思うか」という設問に回答させた．過半数の被験者が選択した文と各手法によるスコアが高かった文が一致した場合を正解として評価した．その結果，比較手法の正解率（平均値：60%，最大値：86%）に比べ，提案手法（93%）が上回っており，提案手法の妥当性を確認した．



心像性 [100,700]

- | | | | |
|---|-----|-----|---|
| A) a man holding a tennis racket in his hand | 490 | 低 | |
| B) it's Kei Nishikori swinging a racket at a ball and he seems 😊 | 542 | ↑ ↓ | |
| C) Japan's Kei Nishikori in action during the match against Portugal's Joao Sousa | 576 | | 高 |

図 1 キャプションごとの心像性の例

- (A) ImageCaptioning.pytorch, (B) CaptionBot, (C) ニュース記事

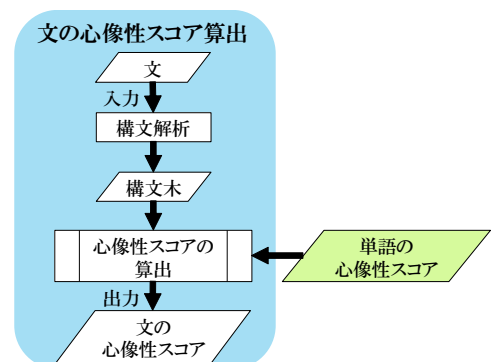


図 2 提案手法の処理手順

[1] Paivio et al., "Concreteness, imagery, and meaningfulness values for 925 nouns," Journal of Experimental Psychology, vol.76, no.1, pp.1-25, Jan. 1968.