

音韻と人体部位の動きの関係に着目した オノマトペによる歩容の zero-shot 記述に向けた検討

加藤 大貴*, 平山 高嗣, 川西 康友, 井手 一郎, 出口 大輔, 村瀬 洋(名古屋大学)

A study on zero-shot gaits description by onomatopoeia based on the relationship between phoneme and body-parts movement.

Hirokata Kato, Takatsugu Hirayama, Yasutomo Kawanishi, Ichiro Ide, Daisuke Deguchi, Hiroshi Murase (Nagoya University)

Table 1 Experimental result

真値	提案手法 (CNN)		比較手法 (LSTM)	
	記述結果	距離	記述結果	距離
すたすた	てこてこ	69.7	とろとろ	41.9
のろのろ	とろとろ	33.8	つらつら	32.9
よろよろ	のろのろ	29.1	つそつそ	39.3
どっしどっし	とことこ	93.9	よろよろ	82.6
せかせか	すったすった	27.7	とろとろ	52.0
てくてく	とろとろ	76.0	すったすった	71.5
とぼとぼ	よろよろ	49.9	つきつき	54.5
ぶらぶら	のろのろ	62.2	のろのろ	84.2
平均		55.0		57.4

1. まえがき

人間の歩行動作は、その見た目に応じて多様なオノマトペで表現される⁽¹⁾。また、オノマトペには音象徴性があり、オノマトペから連想されるイメージはその音韻と強い関係があるとされる。このことから、音象徴性に基づく「音韻空間」を歩容の特徴空間と対応付け、音韻空間上で歩容を取り扱えるようにすることができれば、歩容の微妙な違いをオノマトペの音韻の違いで記述することができると考えられる。人間が歩容を見るときには、部位間の相対的な動きからオノマトペを連想する可能性が従来研究⁽²⁾で示唆されていることをふまえて、本発表では人体部位の相対的な動きに基づく映像特徴を利用し、深層学習を用いた回帰モデルにより両空間を対応付けることで、歩容に対して音象徴性に基づくオノマトペの zero-shot な記述を行なう手法を提案する。

2. オノマトペによる歩容の記述手法

はじめに、映像特徴の計算方法について述べる。まず、CPM (Convolutional Pose Machines)⁽³⁾を用いて映像から人体の部位検出を行なう。これにより14か所の部位座標を得て、そのすべての2部位の組み合わせ(91通り)の距離を求める。この処理を映像中のすべてのフレームに施し、得られた91個の距離系列を映像特徴とする。

次に、音韻空間の構成について述べる。本研究ではABAB型のオノマトペを対象を限り、これを構成する4つの音韻それぞれを小松ら⁽⁴⁾が提案する8次元属性ベクトルで表現する。そして、これらを結合して得られる32次元のベクトルによって張られる空間を音韻空間と定義する。

続いて対応付けに用いる回帰モデルについて述べる。本発表では回帰モデルとして1次元のCNNを利用する。入力として距離系列、教師信号としてオノマトペから計算された32次元のベクトルを与えることで学習を行なう。このとき、91個の距離系列それぞれをチャンネルとみなして用いる。

学習した回帰モデルにオノマトペが未知の映像を入力すると、32次元のベクトルが出力される。このベクトルを8次元ずつに分解し、それぞれにNearest Neighbor法によって音韻空間上で最も近い音韻を割り当てることにより、歩容映像をオノマトペで記述することができる。

3. データセットの作成

データセットとして、歩行動作を正面および背面から撮影

した映像を用意した。これを被験者に提示し、歩容がどのように見えるかを、すたすた、のろのろ、よろよろ、どっしどっし、せかせか、てくてく、とぼとぼ、ぶらぶらの8種類のオノマトペの中から選択させることにより、歩容に対応するオノマトペのアノテーションを行なった。ここで、用いた8種のオノマトペは日本語オノマトペ辞典⁽¹⁾から抜粋した。

4. 実験および結果

Leave-one-onomatopoeia-out 交差検定により、歩容映像をオノマトペで記述する実験を行なった。比較手法として、1次元CNNの代わりにLSTMを利用する手法を用いた。実験結果を表1に示す。表中の距離は、回帰モデルが出力した32次元ベクトルから真値までの音韻空間上での距離を示している。提案手法は全クラス平均で比較手法よりも真値に近い記述結果が得られていることから、提案手法の有効性が確認できた。

5. むすび

本発表では、音象徴性に基づいて歩容をオノマトペによって zero-shot 記述する手法を提案した。実験により、提案手法の有効性が確認できた。今後の課題として、記述結果の妥当性について被験者実験による評価を行なうことや、音韻空間や回帰モデルのアーキテクチャを改良することが挙げられる。謝辞 本研究の一部は栢森情報科学振興財団からの助成と科学研究費補助金による。

文 献

- (1) 小野, 小学館, 2007.
- (2) 杉山 他: 人工知能学会全国大会, 1C1-OS4a-4, 2011.
- (3) S. Wei, et al.: Proc. IEEE CVPR, pp.4724-4732, 2016.
- (4) 小松 他: 信学論 (A), Vol.J92-A, No.11, pp.752-763, 2009.