

# 猫動画群を利用したロボットの動作生成システム

佐藤 紘基<sup>†</sup> 伊嶋 郁也<sup>†</sup> 小林 稜<sup>†</sup> 池山 安杜里<sup>†</sup> 山内 翔<sup>†</sup> 鈴木 恵二<sup>†</sup>

公立ほこだて未来大学<sup>†</sup>

## 1. はじめに

aiboやLOVOTなどの人間と長い時間を共に過ごすロボットが増えてきている。このことから、人間の関心を引く、多様な魅力を持つロボットの創出が求められている。このようなロボットを創出するためには、見た目の形状の要素と行動の要素が重要となってくる。

形状の要素では、山内らによる見た目の形状を優先したロボットの設計手法が提案されており、自動設計が可能となっている[1]。

行動の要素では、人間にとって理解しやすい行動を表現するために、アニメーションからの知見を行動に取り入れた研究が報告されている[2]。しかし、このロボットの行動は、表現する行動のパターンが限られている。

これらのことから、多様な魅力を持つロボットの創出には、人間の注意・関心をより引く、多様で継続的な動作を自動で生成し続けるシステムが必要となる。そこで本稿では、このシステムにおける自然言語処理モデルを利用したロボットの動作生成が可能かを検証する。

## 2. 関連研究

継続的な動作の関連研究として、Peng らによる強化学習による提案手法が挙げられる[3]。この手法では、犬の動作を模倣することで、ロボットに犬の俊敏で継続的な動作を可能としている。しかし、犬の1つの行動パターンを継続させる動作となっているため、多様な行動はできていない。

## 3 猫動画群を利用したロボットの動作生成システム

### 3.1 システム概要

本研究におけるロボット動作の自動生成システムの概要について図1に示す[4]。このシステムは事前学習とファインチューニングの2つの学習から構成されている。事前学習の流れとしては、YouTubeから収集した猫の動画を用いて、2Dの骨格抽出と3Dへの変換を行う。その後、3Dのデータを関節とその相対角度からなるデータに変換する。このデー

Motion generation system for robots using cat video set

<sup>†</sup> Koki Sato, Future University HAKODATE

<sup>†</sup> Fumiya Ishima, Future University HAKODATE

<sup>†</sup> Ryo Kobayashi, Future University HAKODATE

<sup>†</sup> Atori Ikeyama, Future University HAKODATE

<sup>†</sup> Sho Yamauchi, Future University HAKODATE

<sup>†</sup> Keiji Suzuki, Future University HAKODATE

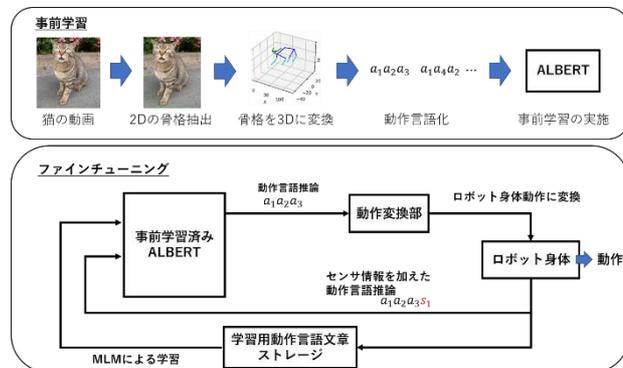


図1 ロボット動作の自動生成システム

タを動作言語と呼ぶ。動作言語化することにより、自然言語処理モデルであるALBERT(A Lite BERT)を利用し事前学習を行う[5]。また、ファインチューニングでは、事前学習済みのALBERTを用いて、ロボットの関節に合わせた学習をすることにより、ロボットの動作生成を可能とする。

### 3.2 学習データ

ロボットの動作元となるデータについては、猫の動画を採用した。猫を採用した理由は、猫が多様な行動をする点が挙げられる。また、動画データを採用した理由は、大量のデータが収集可能な点、モーションキャプチャでのデータとは異なり、動物のサイズに関係なくデータを収集可能なため、応用範囲が広い点が挙げられる。

### 3.3 ALBERT

本研究では、動作生成のために自然言語処理モデルであるALBERTを採用する。これには2つの理由が挙げられる。1つ目は、ALBERTがマスク付き言語モデルと呼ばれるタスクが可能のためである。このタスクは、文章の一部を隠し、その部分を推論し、穴埋めをすることで、文章を自然な流れにすることを可能とするタスクとなっている。このタスクを動作間の推論に用いることで、動作間を自然な流れで補完するためにALBERTを用いる。2つ目は、BERTと比較して、軽量なモデルであることが挙げられる。BERT-baseとALBERT-baseのモデルを比較して、96Mのパラメータ削減が報告されている[5]。この軽量なモデルを利用することで、事前学習におけるデータ量を削減可能なため採用した。

## 4 動作言語化

動作言語は、各関節とその相対角度からなる形式

表 1 動作言語の概要

アルファベット(aからu)	各関節のロール, ピッチ, ヨーを表す
数字(0から9)	関節角の位置を表す
*	シーケンスの先頭を表す
#	シーケンスの末端を表す

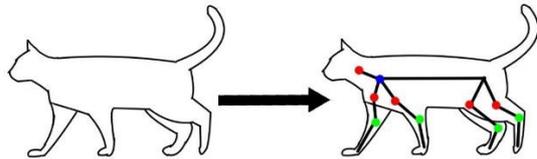


図 2 猫の関節自由度

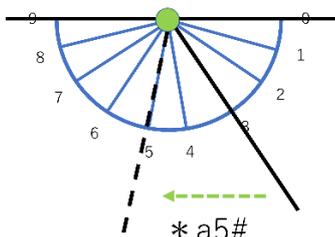


図 3 相対角度

で定義する. 表 1 に動作言語の概要を示す.

関節については, 3D の骨格データを動作言語にするため, 猫の関節に対して自由度を定義する. 図 2 に猫の関節自由度を示す. 関節は, 頭, 首, 肘, 膝となっており, 赤が 3 自由度, 青が 2 自由度, 緑が 1 自由度となっている. この自由度に対して, アルファベットの a から u までを割り当てている.

相対角度については, 各関節に対して, 180 度を 10 等分に離散化した角度で定義する. 図 3 に相対角度を示す. ここでは, 関節 a が相対角度で 3 の位置にいた時に, 5 の位置に移動する様子となっている.

### 5. ALBERT の出力結果

独自に定義した動作言語の形式を ALBERT で事前学習し, 動作言語の推論が可能か検証を行った. データは, 動作言語形式のデータをランダムに 1000000 シーケンス生成し実験を行った. ALBERT に 30 シーケンス出力させた結果を図 4 に示す.

この結果から, 動作言語の形式にあったシーケンスの生成が可能となっている. しかし, シーケンス生成の後半になるにつれて, 同じ値を交互に繰り返すような傾向が見られた. これは, 学習したデータがランダムであるために, 動作言語で表現される動作パターンがなく, ALBERT がうまく学習できなかったことが原因にあると考察する.

### 6. まとめ

本研究では, ロボット動作の自動生成システムに

```
*a6b6c9d1e1f1g1h1j9m9n9o2p1q6r6s6t6u6#
*a9b9c5d5e5f5g4h4i4j1k6l4m6n6o6p0q5r0#
*a1b0c0d1e0f1g9h0i3j7k118m4n4o9p1q1u0#
*a6b9c6d3e4f4g4h5i6j9k6l0m6o6p3q3r6s5#
*a1b0c1d4e9f9g9h9i5j6n0o5p6q4r3t0u6#
*u3#
*c0d3e4g4h0k415m4n5o0p0q6r6s0t6u2#
*a3b2c4d0e0f4g0h8i4k6l0m6p6q3r3s3u6#
*a0b3c1d3e3f9g4h4i0j4k5l8m4o6p3q6r0#
*a3b9c0d0e4f4g9h0i3j1k9l7m2n4o3p9q0#
*a9b3c3d1e3f3g4h4i4j0k4l5m4n5o8p4q4r6#
*a6b8c6d4e4f4g9b2i2j4k5l4m5n8o4p9q0u4#
*a9b6c4d9e0f9g4h4i4j9k9l9n4o8p8q9t9#
*a4b8c9d6e9f4g9h9i9j4k4l4m4n8o9p9q4u9#
*a9b6c6d9e4f9g6h4i6j6k6l9m2n4p8s9t6#
*a6b9c9d6e9f4g9h9i9j9k9l4m4n6o2p9s6u8#
*a9b6c6d9e6f9g6h4i6j6k6l9m2n4o9r9t9#
*a2b9c4d6e9f6g9h9i9j9k9l6m4n2o6p2q9u5#
*a6b6c9d9e6f9g6h6i6j6k6l4m2n4o9p9t0#
*a2b9c6d6e9f6g9h9i9j9k9l6m4n2o6p2q6u9#
*a6b6c9d9e6f9g6h6i6j6k6l4m2n4o9p9q#
*a9b9c6d4e9f4g9h9i9j9k9l6m4n6o6p6q6u6#
*a6b6c9d9e4f9g6h4i0j6k4l4m2n4o9p9t9#
*a9b9c6d6e9f4g9h9i9j9k6l6m4n6o2p2q9u9#
*a6b6c9d9e6f9g6h6i6j6k9l4m2n4o9p9q0#
*a4b9c6d6e9f4g9h9i9j9k6l2m4n6o4p4q9u6#
*a6b6c9d9e6f9g6h6i0j6k9l4m2n4o9p9q0#
*a4b9c6d6e9f4g9h9i9j9k6l2m4n6o4p4q9u9#
*a6b6c9d9e6f9g6h6i6j6k9l4m2n4o9p9q0#
*a4b9c6d6e9f4g9h9i9j9k6l2m4n6o4p4q9u8#
```

図 4 ALBERT の出力結果

において, ALBERT が動作言語の学習・推論が可能か検証を行った. 結果として, 動作言語形式のデータは出力可能であった. しかし, 学習データがランダムであるため, 動作的に意味のある生成はできていないため, 今後は猫動画を動作言語に変換したデータでの検証を行っていく.

### 参考文献

- [1]山内 翔, 鈴木 恵二: ” 結晶構造を参考にした内部構造と 3D プリンタによるモデル形状を優先したロボット設計”, 電気学会論文誌 C (電子・情報・システム部門誌), Vol. 139, No. 9, pp1051-1058 (2019)
- [2]A. J. N. van Breemen: ” Bringing Robots To Life: Applying Principles Of Animation To Robots”, <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.125.8136&rep=rep1&type=pdf> (参照日 2023 年 1 月 11 日)
- [3]Xue Bin Peng, Erwin Coumans, Tingnan Zhang, Tsang-Wei Lee, Jie Tan, Sergey Levine: ” Learning Agile Robotic Locomotion Skills by Imitating Animals”, <https://arxiv.org/abs/2004.00784> (参照日 2023 年 1 月 11 日)
- [4]佐藤 紘基, 伊嶋 郁也, 小林 稜, 池山 安杜里, 山内 翔, 鈴木 恵二: ” ALBERT によるキャラクター的ロボット動作の自動生成”, 令和 4 年度電気・情報関係学会北海道支部連合大会, pp201-202 (2022)
- [5]Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, Radu Soricut: ” ALBERT: A Lite BERT for Self-supervised Learning of Language Representations”, <https://arxiv.org/abs/1909.11942> (参照日 2023 年 1 月 11 日)