

モーショングラフを用いた動作生成のための強化学習

Reinforcement Learning for Motion Generation Using Motion Graphs

情報工学専攻 宮尾 健
Ken MIYAO

概要

本研究では、モーショキャプチャデータから動作生成を行う強化学習モデルを提案する。まず、モーショキャプチャデータをモーショングラフを用いて表現する。モーショングラフを利用して、指定した方向に 3D モデルが進むように学習したモデルから動作を生成する実験を行い、実験で得られた学習時のデータと生成された動作の軌跡から、提案するモデルの考察と評価を行う。

キーワード: モーショキャプチャデータ, モーショングラフ, 強化学習, 動作生成。

1 はじめに

近年、動画配信サービスの品質の向上や 3D モデリング技術の向上に伴い、配信する人が仮想空間上で自身の代わりとなる 3D モデルを用いたコンテンツやサービスが多く見られるようになった。配信者は、自身の動作をキャプチャーし 3D モデルを動かす必要がある。配信者自身の動作をキャプチャーせずに 3D モデルを動かしたいときは、3D モデルを動かすアプリケーション内で用意された動きを 3D モデルの動作に用いることで可能となる。しかし、この方法では、配信者自身の動作の特徴は反映されないため、例えば複数の配信者が同じ動きを 3D モデルの動作に用いたときに機械的な動きになってしまう。配信者の動きの特徴を反映させた 3D モデルの動作を生成するためには、配信者の動作をキャプチャーしたデータを用いて動作を生成する必要がある。

3D モデルの動作の状態によって、次にどのようなモーショキャプチャデータを利用するのか選択をする必要がある。ある状態から行動を選択を行うモデルとして、強化学習モデルがある。これは、エージェントが最適な行動し得られる報酬の和を最大にするモデルである。本研究では、モーショキャプチャデータをグラフで表現したモーショングラフを利用して、3D モデルの動作生成を行う強化学習モデルを提案する。

2 モーショングラフを用いた動作生成

2.1 モーショングラフ

モーショキャプチャデータとは、歩行や走る、跳躍などの動作している人の頭や腕、脚などの様々な関節点の情報をある時点ごとに計測したデータである。モーショキャプチャデータには動作している人の各時点の関節点の 3 次元空間座標の点群が格納されている。本稿ではモーショキャプチャデータをモーシオンデータと呼ぶ。ある時点におけるモーシオンデータをフレームという。

あるフレームから異なるフレームまでの動作をクリップという。特に、2 つのモーシオンデータによって生成されたクリップをトランジションクリップという。モーショングラフ[3]とは、モーシオンデータを有向グラフで表現したものである。モーショングラフにおいて、フレームはノード、クリップとトランジションクリップはエッジに対応する。モーショングラフ上のパス上のエッジに対応する動作をつなぎ合わせることで動作を生成することができる。

モーショングラフの関連研究として、文献 [2] では、モーショングラフを作成するときに、複数の類似した動作からより自然な動作を生成するための手法を提案している。また、文献 [1] では、モーショングラフを作成するための新たな手法を提案している。

2.2 モーショングラフの作成

距離グリッドとは、あるモーシオンデータのすべてのフレームに対し、他のモーシオンデータのすべてのフレームとの距離を、格子状に表現したものである。ある 2 つのフレームの距離は、以下の最小化問題を解くことによって計算される。

$$\min_{\theta, x_0, z_0} \sum_i w_i \|\mathbf{p}_i - \mathbf{T}_{\theta, x_0, z_0} \mathbf{p}'_i\|^2 \quad (1)$$

i は関節点の番号を表し、 $\mathbf{p}_i, \mathbf{p}'_i$ は、関節点の位置を表すそれぞれ 3 次元座標上の点である。 $\mathbf{T}_{\theta, x_0, z_0}$ は、点 \mathbf{p} を y 軸に関して θ 回転し、 x_0, z_0 だけ平行移動させる変換

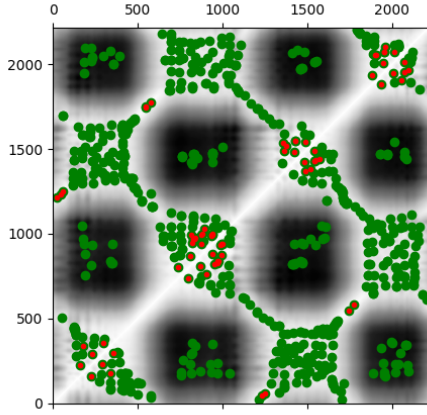


図1 距離グリッドと局所最小点

である。この最小化問題の解は以下ようになる。

$$\theta = \arctan \frac{\sum_i w_i (x_i z'_i - x'_i z_i) - \frac{1}{\sum_i w_i} (\bar{x} \bar{z}' - \bar{x}' \bar{z})}{\sum_i w_i (x_i x'_i + z_i z'_i) - \frac{1}{\sum_i w_i} (\bar{x} \bar{x}' + \bar{z} \bar{z}')}$$

$$x_0 = \frac{1}{\sum_i w_i} (\bar{x} - \bar{x}' \cos(\theta) - \bar{z}' \sin(\theta))$$

$$z_0 = \frac{1}{\sum_i w_i} (\bar{z} + \bar{x}' \sin(\theta) - \bar{z}' \cos(\theta))$$

ただし、 $\bar{x} = \sum_i w_i x_i$ である。

図1は距離グリッドを可視化したものである。縦軸と横軸はそれぞれ2つのモーションデータのフレームに対応する。2つのフレームの組の距離がグレースケールによって表現されており、黒い点は距離の小さいフレームの組、白い点は距離の大きいフレームの組である。

次に、距離グリッド上において、ある点に対して隣接する8つの点より最小となるような局所最小点を求める。図1の緑色の点は局所最小点をプロットしたものである。また、これらの緑色の点のうちいくつかの点に対応する2つのフレームは、モーショングラフのノードとなる。

次に、距離に対する閾値を定義し、すべての局所最小点のうち距離が閾値以下となる点を求める。図1の赤色の点はこれらの点をプロットしたものである。

この手順によって、モーショングラフ上のノードは、距離グリッド上の、前述した手順によって求めた点に対応する2つのフレームから定義される。

\mathbf{f}, \mathbf{f}' をそれぞれ動作を生成するのに必要な2つのフレームとすると、本研究ではトランジションクリップを以下のように作成する。

$$\mathbf{tf}(\mathbf{p}) = \alpha(\mathbf{p})\mathbf{f} + (1 - \alpha(\mathbf{p}))\mathbf{f}' \quad (2)$$

\mathbf{tf} はトランジションクリップのあるフレームの点群を表し、 p はフレームの番号である。 $\alpha(p)$ は以下の式で

ある。

$$\alpha(p) = \begin{cases} 1 & (p \leq 1) \\ 2\left(\frac{p+1}{k}\right)^3 - 3\left(\frac{p+1}{k}\right)^2 + 1 & (1 < p < k) \\ 0 & (k \leq p) \end{cases} \quad (3)$$

ここで、 k はユーザーによって定められたトランジションクリップのフレーム数である。式(2)によって、トランジションクリップは、 \mathbf{f} に対応する動作から \mathbf{f}' に対応する動作へと徐々に変化するような動作となる。

2.3 動作生成

モーショングラフを用いて動作生成を行う場合、あるエッジから次のエッジ、すなわち、ある動作から異なる動作を逐次的に結合することによって動作が生成される。2種類の動作を結合するとき、それぞれのエッジに対応する動作はそのまま結合すると、フレームの点群の位置がずれることで不自然な動作を生成してしまう。そこで、生成されている動作の位置の変換を行う必要がある。本研究では、パスの先頭のエッジに対応する動作の最後のフレームと結合するエッジの最初のフレームから、式(1)で求めた $\mathbf{T}_{\theta, \mathbf{x}_0, \mathbf{z}_0}$ を用いて変換を行う。結合する動作に含まれるすべてのフレームに対して、 $\mathbf{T}_{\theta, \mathbf{x}_0, \mathbf{z}_0}$ を用いた動作の位置の変換を行い、動作を結合することで自然な動作が生成されるようにする。

3 強化学習

3.1 強化学習とは

強化学習とは、行動する主体であるエージェントがある未知の環境において、収益と呼ばれる得られる報酬の累積が最大となるような、最適な行動の系列を求める理論的枠組みである[4]。機械学習の一分野として捉えられることが多いが、機械学習と異なる点として、アルゴリズムを繰り返し行うことでデータを収集し、最適な出力を求めることが挙げられる。

エージェントが最適な行動の系列を求めるための指標として行動価値がある。これは、ある状態においてある行動を取るときに得られる収益の期待値である。行動価値の計算は、未来に得られる報酬が未知であることから困難である。そこで、Q学習と呼ばれる、逐次的に更新を行い行動価値を推定する方法をとる。 $Q(s, a)$ を状態 s 、行動 a における行動価値とすると、Q学習では次の更新方法によって、エージェントが状態遷移するごとに行動価値の推定値を更新し最適な行動価値を求める。

$$Q(s, a) \leftarrow (1 - \mu)Q(s, a) + \mu[r(s, a, s') + \gamma \max_{a' \in A(s')} Q(s', a')]$$

このとき、 μ は学習率と呼ばれるもので、 $0 \leq \mu \leq 1$ で

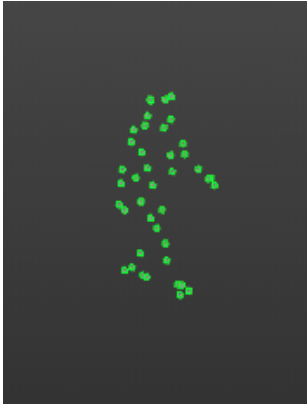


図2 使用するモーションデータ

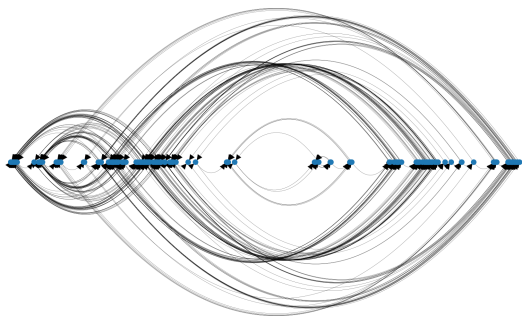


図3 モーショングラフ

ある。また、 γ は割引率と呼ばれるもので、 $0 \leq \gamma \leq 1$ である。Q 学習を用いた強化学習では、はじめにすべての状態と行動についての行動価値を 0 に初期化しておき、エージェントが状態遷移をするごとにこの更新式によって価値関数を更新する。

4 強化学習による動作生成

4.1 使用するデータ

実験では、カーネギーメロン大学から公開されているモーションキャプチャーデータベース (Graphics Lab Motion Capture Database)[5] のうち、8 の字に歩くモーションデータを利用する。モーションデータのフレーム数は 2217、関節点数は 41 である。図 2 は、モーションデータを可視化したものである。

4.2 強化学習の適用

本研究では、ユーザーが指定した方向に進むような動作を生成するために、強化学習を利用する。モーショングラフ上のノードとユーザーが指定した方向の組を状態、エッジを行動とし、強化学習によって行動列、すなわち、モーショングラフ上のパスを求めることで動作生成

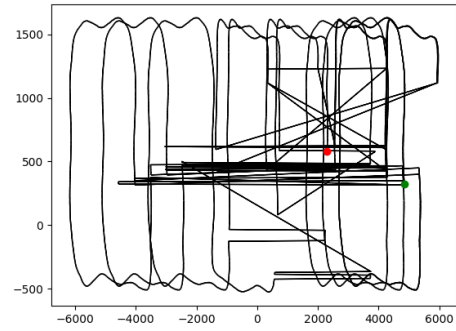


図4 学習前の軌跡 (実験1)

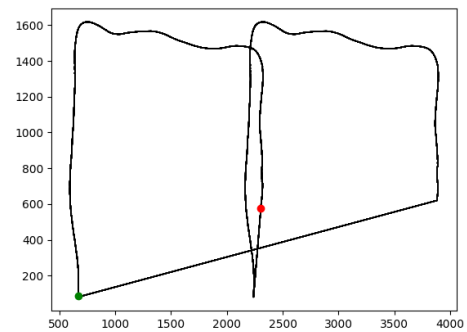


図5 学習後の軌跡 (実験1)

を行う。

実験で用いる強化学習モデルにおけるパラメータの設定は、最大エピソード数を 10000、最大ステップ数を 50 に設定した。モーショングラフに関するパラメータの設定は、遷移点の選択の割合は 20%, 25%, 50% とした。図 3 は、遷移点の選択の割合が 20% の場合のモーショングラフである。

実験で用いる報酬関数は、以下のように定義する。

$$r(s, a, s') = \frac{v_a \cdot v_o}{\|v_a\| \|v_o\|} \quad (4)$$

v_a は、行動 a によってある状態 s から次の状態 s' に遷移したときの、ある関節点の移動方向を表す 3 次元ベクトルである。 v_o は、指示した方向を表す 3 次元ベクトルである。

4.3 実験 1: 1 方向の学習による動作生成

はじめに、1 方向にのみ 3D モデルが動くよう、強化学習を用いた動作生成を行った。この実験では、式 (4) において、 v_o を x 軸に平行な単位ベクトルとした。

図 4 と図 5 は、それぞれ遷移点の選択の割合を 20% として実験を行ったときの、エピソードが 1 回目と 10000 回目における 3D モデルの軌跡の図である。生成された動作は、1 回目のエピソードでは乱雑な軌跡となってい

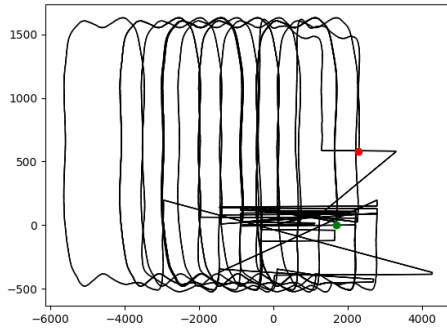


図6 学習前の軌跡 (実験 2-1)

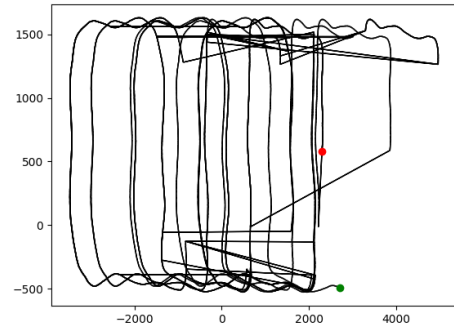


図8 学習前の軌跡 (実験 2-2)

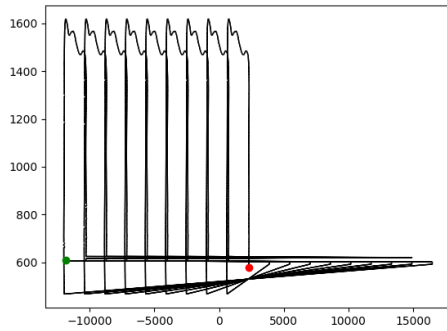


図7 学習後の軌跡 (実験 2-1)

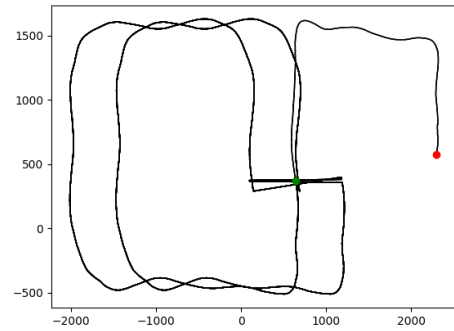


図9 学習後の軌跡 (実験 2-2)

るが、10000 回目のエピソードでは単純な軌跡となっている。これは指定した方向の動作が生成されるように学習されたためと考えられる。

4.4 実験 2：4 方向の学習による動作生成

実験 2-1 では、式 (4) において、1250 エピソードごとに $x, z, -x, -z, -x, z, x, -z$ の順に、実験 2-2 では、2500 エピソードごとに $x, z, -x, -z$ の順にこれらの軸に平行な単位ベクトルを v_o として実験を行った。

図 6 から図 9 は、遷移点の選択の割合を 20% とした場合の、エピソードが 1 回目と 10000 回目における 3D モデルの軌跡の図である。実験 2-1 のモデルでは、学習後に生成された動作は不自然な軌跡となった。また、実験 2-2 のモデルでは学習後に生成された動作は、実験 2-1 と比較して口の字を描く綺麗な軌跡となった。このことから、実験 2-2 による報酬関数の設定の方がより自然な軌跡を描く動作が生成されるように学習されていると考えられる。

5 結論

本研究では、強化学習による動作生成モデルについて、1 方向の学習による動作生成モデルと 4 方向の学習による動作生成モデルを作成し、実験と評価を行った。

今後の課題として、エピソード数を増やした実験やモーショングラフのノードとエッジの数を増やした実験を行う必要がある。より複雑な動作を生成するためには、多くのモーションデータを利用した実験や報酬関数を変えた実験を行う必要がある。

参考文献

- [1] L. Chao and A. Safonova: Achieving good connectivity in motion graphs, *Graphical Models*, vol. 71 (4), pp. 139–152, 2009.
- [2] L. Kovar and M. Gleicher: Automated extraction and parameterization of motions in large data sets, *ACM Transactions on Graphics*, vol. 23 (3), pp. 559–568, 2004.
- [3] L. Kovar and M. Gleicher: Motion Graphs, *ACM Transactions on Graphics*, vol. 21 (3), pp. 473–482, 2002.
- [4] 牧野貴樹ら編著: これからの強化学習, 森北出版, 2016.
- [5] Carnegie Mellon University: CMU Graphics Lab Motion Capture Database, available from (<http://mocap.cs.cmu.edu>), (accessed 2021-07-16).