

# CycleGAN アーキテクチャに基づくモーションデータのスタイル変換

坂口 颯太<sup>†</sup> 白井 匡人<sup>††</sup>

<sup>†</sup> 島根大学 自然科学研究科 〒 690-8504 島根県松江市西川津町 1060

<sup>††</sup> 島根大学 学術研究院理工学系 〒 690-8504 島根県松江市西川津町 1060

E-mail: <sup>†</sup>n21m104@matsu.shimane-u.ac.jp, <sup>††</sup>shirai@cis.shimane-u.ac.jp

**あらまし** ゲーム分野などにおいて使用されるモーションデータには、そのキャラクターの個性に合わせた多種多様なスタイルが存在する。そのため、それぞれモーションキャプチャなどでデータを1つずつ作成し直すか、既存のモーションデータを手作業で修正していく必要がある。先行研究で CycleGAN アーキテクチャを用いて、既存のモーションデータを別の個性を持ったモーションデータにスタイル変換する手法があるが、生成モーションが一部激しく振動し滑らかではなくなる場合がある。本研究では、この問題を改善するために別の損失関数に置き換えて学習する手法を提案する。実験によりこの提案手法の有効性を検証し、今後の課題について検討する。

**キーワード** モーションデータ, スタイル変換, GAN, CycleGAN

## 1 前 書 き

近年、ゲーム業界では3次元コンピュータグラフィックス技術(3DCG)を用いた3Dゲームが普及しており、スマートフォン向けゲームにおいてもモーションキャプチャ技術を用いて作成されたモーションデータが採用されているなど、家庭用ゲーム機のみならず、スマートフォン向けのゲームにまで3DCGが採用される傾向にある。

これらのゲームにはそれぞれ性別や性格などの設定が異なる個性豊かなキャラクターが存在する。キャラクターの3Dモーションデータは、専門的な技術を習得した技術者らによる長時間かけた手作業もしくは、モーションキャプチャ技術を用いて演者が実際に演じるにより得られる。しかし、「歩く」「走る」などのある1つのアクションのモーションであっても、手の振る角度や足の上げる高さ、曲げる角度など複数の異なるスタイルのモーションが存在する。キャラクターごとにそれらの個性に合うモーションを作成するには、再度手作業での編集もしくは録り直しが必要になる。

そのような長期間必要とする作業の負担を軽減するための手法を検討する。Dongらによる先行研究[1]では、CycleGAN アーキテクチャ[2]を用いたモーションの大人から子供へのスタイル変換の有効性が示されている。しかし、このAdult2Child[1]の手法では、CycleGANに加えてモーションの滑らかさの向上のため「時間的コヒーレンス損失」が導入されたが、体全体にわたってフレーム間のポーズ差分を計算しそれを最小化するので、一部のボーンが激しく振動し滑らかではないモーションが生成される事象が発生する。

そこで本研究では、既存手法のAdult2Child[1]で用いられている一部の損失関数を変更し別の新たな損失関数を用いて学習を行い、体全体に渡って滑らかなモーションを生成する手法を提案する。「時間的コヒーレンス損失」の代わりに新たに「サイクル一貫性時間的差分損失」を定義する。これは、CycleGAN

のサイクル一貫性[2]の考えに基づき、フレーム間のポーズ差分を元のデータと、変換後のデータをさらに元のスタイルに戻したものとで計算し、その2つの差分を最小化することでボーン毎にポーズ差分を考慮するものである。

第2章では関連研究について、第3章では提案手法についてそれぞれ述べる。第4章で実験によって提案手法の有効性を示し、第5章にて結論をまとめる。

## 2 関 連 研 究

### 2.1 モーションリターゲティング

ある骨格構造のモーションデータを、異なる骨格構造のモーションへと再適合させるモーションリターゲティングは、これまで盛んに研究されてきている[3],[4],[5]。しかし、Dongらはモーションリターゲティングでは四肢の長さの変化に対応はできないものの、子供らしいモーションにはならないことを指摘している[1]。本研究ではボーンごとの可動域のような、キャラクターの個性などによるモーションの違いを考慮したデータの生成ができるようにする。

### 2.2 モーションのスタイル変換

同じ骨格構造であっても異なるスタイルになるようにモーションを変換する手法も、モーションリターゲティングとはまた別に研究されている[6],[7],[8]。スタイルは主観的で数学的な定義が困難なため、大規模なデータベースからスタイルの特徴を推論するアプローチが研究されており、その1つであるHsuらの手法[9]をDongらが大人から子供へのスタイル変換に利用した研究[10]では、対となる学習データに依存している。これに対し、本研究では対となる学習データを必要としない。

### 2.3 敵対的生成ネットワーク

敵対的生成ネットワーク(GAN)は、より小さなデータセットでもうまく機能するよう、様々な研究が行われている[2],[11]。それらの研究のうち、CycleGAN[2]は2つの異なるスタイル

間の変換を、たとえば線画とその線画に色を塗った画像の完全な対になっているデータだけではなく、対になっていない馬とシマウマの画像データからでも行えるようにした画期的な手法である。本研究では、この CycleGAN アーキテクチャを用いることで、対になっていない少ないモーションデータでもモーションのスタイル変換を可能にする。

## 2.4 Adult2Child

Adult2Child [1] は、CycleGAN のアーキテクチャをモーションデータの生成に利用し、大人のモーションを子供らしいモーションに変換するためのアプローチを示している。Adult2Child では、モーションデータを CycleGAN に適用するため 60 フレームごとにモーションを区切り、それを 1 つの学習データとするモーションワード (motion words) を提唱している。また、大きく揺れており滑らかではないモーションが生成される問題に対する解決策を提示している。

本研究でもこの Adult2Child のフレームワークをベースとし、モーションワードを用いて学習を行ったところ、生成されたモーションデータの大きな揺れが抑えられて入るものの依然として小刻みに振動してしまっているといった、修正すべき点が見られたため、本研究ではこれの改善を試みる。

## 3 提案手法

提案手法は、Adult2Child フレームワークを基に、学習時の損失関数に変更を加え生成されるモーションデータを改善に導くものである。Adult2Child でモーションの滑らかさと安定性の向上のために導入された「時間的コヒーレンス損失」では、フレーム間のポーズ差分を全て足し合わせた値を最小化するように学習しており、一部のボーンの滑らかさが改善しない場合がある。そこで、新たに「サイクルー貫性時間的差分損失」を用いる。

「サイクルー貫性時間的差分損失」は CycleGAN のサイクルー貫性 [2] の考えを基に、ボーン毎に元のスタイルのデータと、変換後のデータを更に元のスタイルに戻したものとでそれぞれフレーム間のポーズ差分を計算し、それらを比較し同じ値に近づくように学習させるものである。これにより、ボーン全体に渡って動きを滑らかにすることができる。

### 3.1 扱うデータについて

ここでは、本研究で扱うモーションデータを定義する。Adult2Child では、モーションデータにおけるボーンの角度を四元数に変換して利用している [1] が、本研究では内部的にオイラー角を四元数に変換して処理をする Unity 上でモーションデータを扱うことを想定しているため、ボーンの角度をオイラー角のまま扱うものとする。図 1 のような状態を初期状態とし、このとき全てのボーン角度は  $0^\circ$  になる。図 1 において横軸が X 軸、縦軸が Y 軸、奥行方向の軸が Z 軸となる。例として左肘のボーンを Z 軸のみ  $30^\circ$  回転した状態を図 2 に示す。またボーン角度は  $-180^\circ$  から  $180^\circ$  の範囲を取るように成形された上で、 $-1$  から  $1$  の範囲で最大最小正規化されているものと

する。

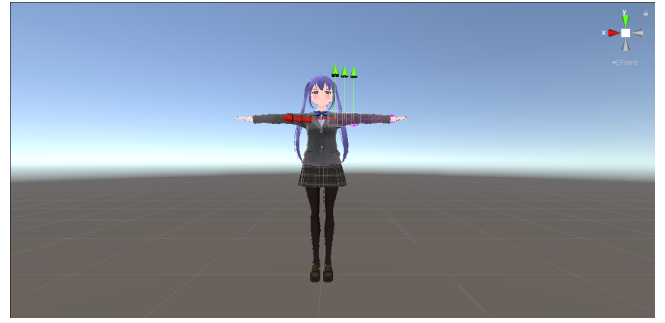


図 1 全てのボーン角度が  $0$  の状態 (初期状態)。

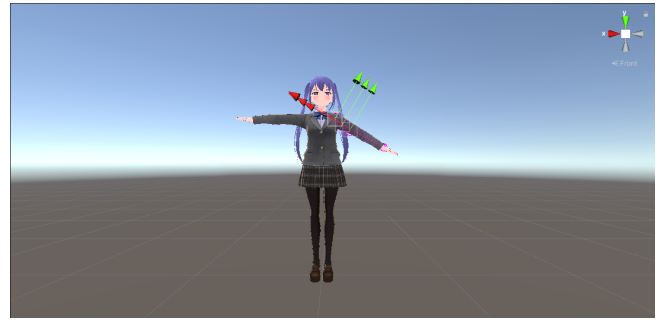


図 2 初期状態のポーズから左肘のボーンの Z 軸のみ  $30^\circ$  回転した状態。

ボーン構造は、ボーン位置のプロパティはルートボーン 1 つだけが持ち、 $j$  個のすべてのボーンが 3 つのボーン角度のプロパティを持つものとする。本研究ではボーン角度のみを利用するため、ルートボーンのボーン位置のプロパティは初期位置で固定するものとする。なお、実際の実験で用いたボーン構造などについては、第 4 章にて述べる。

本研究では毎秒 60 フレーム、すなわち 60fps の BVH 形式のモーションデータを扱い、Adult2Child [1] にて提唱されている 60 フレームごとに分割した、 $60 \times j \times 3$  のモーションワードを 1 つのデータとする。

### 3.2 フレームワーク

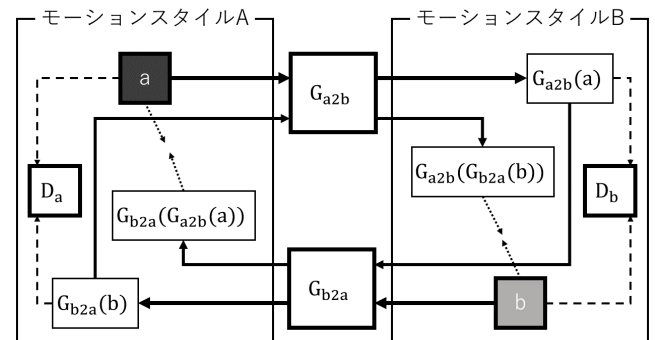


図 3 Adult2Child に基づく「motionA2B」フレームワーク。

提案する「motionA2B」フレームワークを図 3 に示す。スタ

イル A,B の 2 つのモーションデータをそれぞれ  $a, b$  とし、スタイル A,B を本物かどうか見分ける判別器をそれぞれ  $D_a, D_b$  とする。スタイルを A から B, B から A へと変換する生成器をそれぞれ  $G_{a2b}, G_{b2a}$  とし、任意のモーションワード  $x$  を入力した、これらの判別器・生成器の出力をそれぞれ  $D_a(x), D_b(x), G_{a2b}(x), G_{b2a}(x)$  とする。また、モーションワード 1 つあたりのフレーム数を  $t_{fps}$ 、ボーンの数  $j$ 、ボーン角度のプロパティ数を  $i$  とする。

### 3.3 損失関数

既存手法との違いを比較するため、Adult2Child [1] でのみ用いられる時間的コヒーレンス損失と、提案手法でのみ用いられるサイクルー貫性時間的差分損失とを合わせて述べる。

#### 3.3.1 敵対的損失

本物らしいデータを生成するために用いる GAN の敵対的損失には、平均二乗誤差 (MSE) を用いる。判別器  $D_a, D_b$ 、生成器  $G_{a2b}, G_{b2a}$  の敵対的損失  $\mathcal{L}_{D_a}, \mathcal{L}_{D_b}, \mathcal{L}_{G_{a2b}}, \mathcal{L}_{G_{b2a}}$  を、

$$\begin{aligned} \mathcal{L}_{D_a} = & 0.5 \frac{1}{t_{fps} j i} \sum_{t_{fps}} \sum_j \sum_i \{1 - D_a(a)\}^2 \\ & + 0.5 \frac{1}{t_{fps} j i} \sum_{t_{fps}} \sum_j \sum_i \{D_a(G_{b2a}(b))\}^2 \quad (1) \end{aligned}$$

$$\begin{aligned} \mathcal{L}_{D_b} = & 0.5 \frac{1}{t_{fps} j i} \sum_{t_{fps}} \sum_j \sum_i \{1 - D_b(b)\}^2 \\ & + 0.5 \frac{1}{t_{fps} j i} \sum_{t_{fps}} \sum_j \sum_i \{D_b(G_{a2b}(a))\}^2 \quad (2) \end{aligned}$$

$$\mathcal{L}_{G_{a2b}} = \frac{1}{t_{fps} j i} \sum_{t_{fps}} \sum_j \sum_i \{1 - D_b(G_{a2b}(a))\}^2 \quad (3)$$

$$\mathcal{L}_{G_{b2a}} = \frac{1}{t_{fps} j i} \sum_{t_{fps}} \sum_j \sum_i \{1 - D_a(G_{b2a}(b))\}^2 \quad (4)$$

とする。

#### 3.3.2 サイクルー貫性損失

スタイル変換後のモーションデータの品質を保証するために CycleGAN のサイクルー貫性損失 [2] を採用し、入力モーションを 2 つの生成器  $G_{a2b}, G_{b2a}$  を通して元のスタイルに戻るよう学習させる。損失関数には平均絶対誤差 (MAE) を用いる。スタイル A,B のサイクルー貫性損失  $\mathcal{L}_{cycle,a}, \mathcal{L}_{cycle,b}$  を、

$$\mathcal{L}_{cycle,a} = \frac{1}{t_{fps} j i} \sum_{t_{fps}} \sum_j \sum_i ||| G_{b2a}(G_{a2b}(a)) - a ||| \quad (5)$$

$$\mathcal{L}_{cycle,b} = \frac{1}{t_{fps} j i} \sum_{t_{fps}} \sum_j \sum_i ||| G_{a2b}(G_{b2a}(b)) - b ||| \quad (6)$$

とする。

### 3.4 時間的コヒーレンス損失

出力モーションの滑らかさと安定性の向上のために Adult2Child において導入された時間的コヒーレンス損失 [1] について述べる。尚、提案手法ではこの損失は用いていない。モーションワードあたりのフレームごとのポーズの差分

を計算するため、スタイル A,B の時間的コヒーレンス損失  $\mathcal{L}_{coherence,a}, \mathcal{L}_{coherence,b}$  を、

$$\mathcal{L}_{coherence,a} = \sum_{t_{diff}} \sum_j \sum_i ||| G_{a2b}(a)(t) - G_{a2b}(a)(t-1) ||| \quad (7)$$

$$\mathcal{L}_{coherence,b} = \sum_{t_{diff}} \sum_j \sum_i ||| G_{b2a}(b)(t) - G_{b2a}(b)(t-1) ||| \quad (8)$$

とする。ここで、モーションワード内の任意のフレームを  $t$  ( $1 \leq t \leq 60$ ) とし、フレーム間の差分の数を  $t_{diff} = t_{fps} - 1$  とする。

### 3.5 サイクルー貫性時間的差分損失

時間的コヒーレンス損失でモーションの滑らかさを向上することはできるが、フレーム間のポーズの差分を最小化するため、全く動かないモーションデータが生成されたり一部のボーンの振動が改善されない場合がある。そこで、本研究では新たにサイクルー貫性時間的誤差損失を提唱する。

これは、現実的なフレーム間のポーズ差分を持つデータを生成するため、元のモーションデータと、2 つの生成器  $G_{a2b}, G_{b2a}$  を通して元のスタイルに変換したモーションデータとでそれぞれのフレーム間のポーズ差分を計算し、その誤差を計算するのである。サイクルー貫性時間的誤差損失  $\mathcal{L}_{diff,a}, \mathcal{L}_{diff,b}$  を

$$\mathcal{L}_{diff,a} = \sum_{t_{diff}} \sum_j \sum_i ||| diff(G_{b2a}(G_{a2b}(a))) - diff(a) ||| \quad (9)$$

$$\mathcal{L}_{diff,b} = \sum_{t_{diff}} \sum_j \sum_i ||| diff(G_{a2b}(G_{b2a}(b))) - diff(b) ||| \quad (10)$$

とする。ここで、任意のモーションワード  $x$  のフレーム間のポーズ差分を計算する関数を  $diff(x)$  とする。

### 3.6 全体的な損失関数

全体的な損失関数は、次のように計算できる。

$$\begin{aligned} \mathcal{L}_{G_{a2b}, G_{b2a}, D_a, D_b} = & \mathcal{L}_{G_{a2b}} + \mathcal{L}_{G_{b2a}} + \mathcal{L}_{D_a} + \mathcal{L}_{D_b} \\ & + \lambda_{cycle} \times (\mathcal{L}_{cycle,b} + \mathcal{L}_{cycle,a}) \\ & + \lambda_{diff} \times (\mathcal{L}_{diff,b} + \mathcal{L}_{diff,a}) \quad (11) \end{aligned}$$

ここで、 $\lambda_{-}$  は各損失に対する重みである。

次のように、損失を最小するように 2 つの生成器と 2 つの判別器を学習する。

$$L = \arg \min_{G_{a2b}, G_{b2a}} \min_{D_a, D_b} \mathcal{L}_{G_{a2b}, G_{b2a}, D_a, D_b} \quad (12)$$

## 4 実験

### 4.1 モーションデータセット

実験では、HTC 社の VIVE Pro ヘッドセットとコントローラー、並びに VIVE トラッカーによって取得したモーション

データを使用する。図 4 のように頭部にヘッドセット、両手に 2 つのコントローラーを持ち、両肘・腰・両膝・両足の甲に計 7 つの VIVE トラッカーを装着し、モーションキャプチャソフトウェア「VirtualMotionCapture」を用いて得られるモーションを Unity 上で取得し、BVH 形式のモーションデータとして出力する。



図 4 モーションキャプチャ環境.

実験に使用するのはそれぞれ 200 秒、その場で手をしっかり前後に振り、陸上部のようなしっかりしたフォームのスタイル A (図 5) と、脇を締め続けることを意識し、疲労により頭を左右に振り崩れてしまったフォームのスタイル B (図 6) とで足踏みしたモーションデータを学習データとして利用し、追加でスタイル A のモーションを 170 秒収録し、検証用データとして利用する。



図 5 スタイル A.



図 6 スタイル B.

ボーン構造は、3D 人型アバター作成ソフト「Vroid Studio」を用いて作成された VRM 形式準拠のモデルを使用し、表 1 と図 7 に示す 24 個のボーンの、X,Y,Z 軸の 3 つのボーン角度を用いる。

## 4.2 比較手法

実験では、以下の 2 つの手法を比較する。

(1) Adult2Child ベース: CycleGAN の手法 [2] に Adult2Child の時間的コヒーレンス損失 [1] 加えたもの

表 1 実験に用いる 24 個のボーン

添字	ボーン名	身体部位
0	J_Bip_C_Hips	尻, ルートボーン
1	J_Bip_C_Spine	腰
2	J_Bip_C_Chest	背骨 (腹部)
3	J_Bip_C_UpperChest	背骨 (肩甲骨)
4	J_Bip_C_Neck	首
5	J_Bip_C_Head	頭
6	J_Bip_L_Shoulder	左肩
7	J_Bip_L_UpperArm	左上腕
8	J_Bip_L_LowerArm	左前腕
9	J_Bip_L_Hand	左手
10	J_Bip_R_Shoulder	右肩
11	J_Bip_R_UpperArm	右上腕
12	J_Bip_R_LowerArm	右前腕
13	J_Bip_R_Hand	右手
14	J_Bip_L_UpperLeg	左大腿
15	J_Bip_L_LowerLeg	左下腿
16	J_Bip_L_Foot	左足首
17	J_Bip_L_ToeBase	左足の甲
18	J_Bip_L_ToeBase_end	左つま先
19	J_Bip_R_UpperLeg	右大腿
20	J_Bip_R_LowerLeg	右下腿
21	J_Bip_R_Foot	右足首
22	J_Bip_R_ToeBase	右足の甲
23	J_Bip_R_ToeBase_end	右つま先

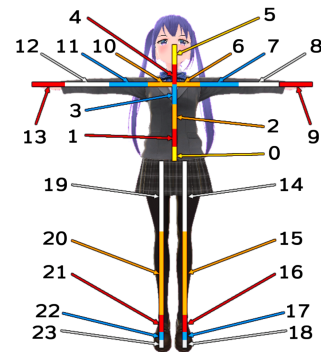


図 7 ボーン位置と添字の対応.

(2) 提案手法: CycleGAN の手法 [2] にサイクル一貫性時間的差分損失を加えたもの

尚, Adult2Child ベースでは、3.6 節の全体的な損失関数の、 $\lambda_{diff} \times (\mathcal{L}_{diff,b} + \mathcal{L}_{diff,a})$  を  $\lambda_{coherence} \times (\mathcal{L}_{coherence,a} + \mathcal{L}_{coherence,b})$  に置き換えて学習を行う。

## 4.3 実験内容と評価指標

実験の前に、それぞれの手法でいくつかの重み  $\lambda$  の値の組み合わせを変えて試し、生成器や判別器がうまく学習できる値を算出しておく。事前に算出された重みを用いて学習を行ったモデルを利用し、スタイル A である検証用データを用いてスタイル B に変換されたモーションデータの滑らかさを比較し評価する。モーションの滑らかさとは、ここではモーションが止まることなくスタイルの変換がうまくできているという前提のも

と、フレーム間のポーズ差分が小さく均一に保たれているほど滑らかであるものと定義する。

実験に用いた重み  $\lambda_-$  の値を表 2 に示す。また、共通して実験結果は 90 エポック学習させたものとする。

表 2 実験に用いた重み

手法	$\lambda_{cycle}$	$\lambda_{coherence}$	$\lambda_{diff}$
Adult2Child ベース	20	0.001	-
提案手法	20	-	0.01

生成されるモーションが滑らかであるかどうかを示す評価指標として、ボーン毎に 1 フレーム間における X 軸,Y 軸,Z 軸のボーン角度のユークリッド距離を計算し、1 フレーム進む毎のその距離の変化量を算出する。この変化量の変動が少ないとき、つまり波が小さいほどモーションが滑らかであると判断する。尚、この値が 0 であるとき、そのボーンはそのフレーム間において全く動いていないものとなる。本研究では、生成された初めの 3 秒のモーションデータにおけるボーンごとのユークリッド距離の変化量を Adult2Child ベースと提案手法とで比較する。

また、定量的な評価指標として、ボーンごとに 1 フレームあたりの差分の合計値をテストデータの全てのモーションワードで計算し 1 モーションワードあたりの平均を算出する。この数値の全ボーンの合計値と平均値で評価する。値が小さいほど滑らかであり、0 であれば最初から最後まで全く動かないものとなる。

#### 4.4 実験結果

実験により得られたユークリッド距離の変化量のボーンごとの比較結果を図 8 と図 9 に示す。横軸は時間経過、縦軸は指標の数値である。尚、縦軸の範囲はボーンごとに異なる。青色が Adult2Child [1]、橙色が提案手法のものである。

また、モーションワードのフレーム間の差分の全ボーンの合計値、1 ボーンあたりの平均値を表 3 に示す。変換前のテストデータで計算した指標も合わせて示す。

表 3 モーションワードのフレーム間の差分の比較

指標	変換前	Adult2Child	提案手法
1 ボーンあたりの平均	0.0093	0.0210	0.0123
全ボーンの合計値	0.2224	0.5031	0.2943

#### 4.5 考察

図 8 と図 9 より、ほとんどのフレーム間において提案手法は既存手法より滑らかに動いていることが確認できる。また、表 3 から滑らかさが改善していることが確認できる。これらの実験結果から、モーションワード内のフレーム間の変位を元のモーションの変位と合わせる事ができるサイクル一貫性時間的差分損失の有効性が示されているといえる。しかしながら、ボーン 7,9,11 など一部のボーンにおいて、60,120 フレーム目付近でそれぞれ変化量が著しく大きくなっているのが確認できる。この結果より、モーションワードの繋ぎ目において、モーション生成が滑らかでなくなっているといえる。

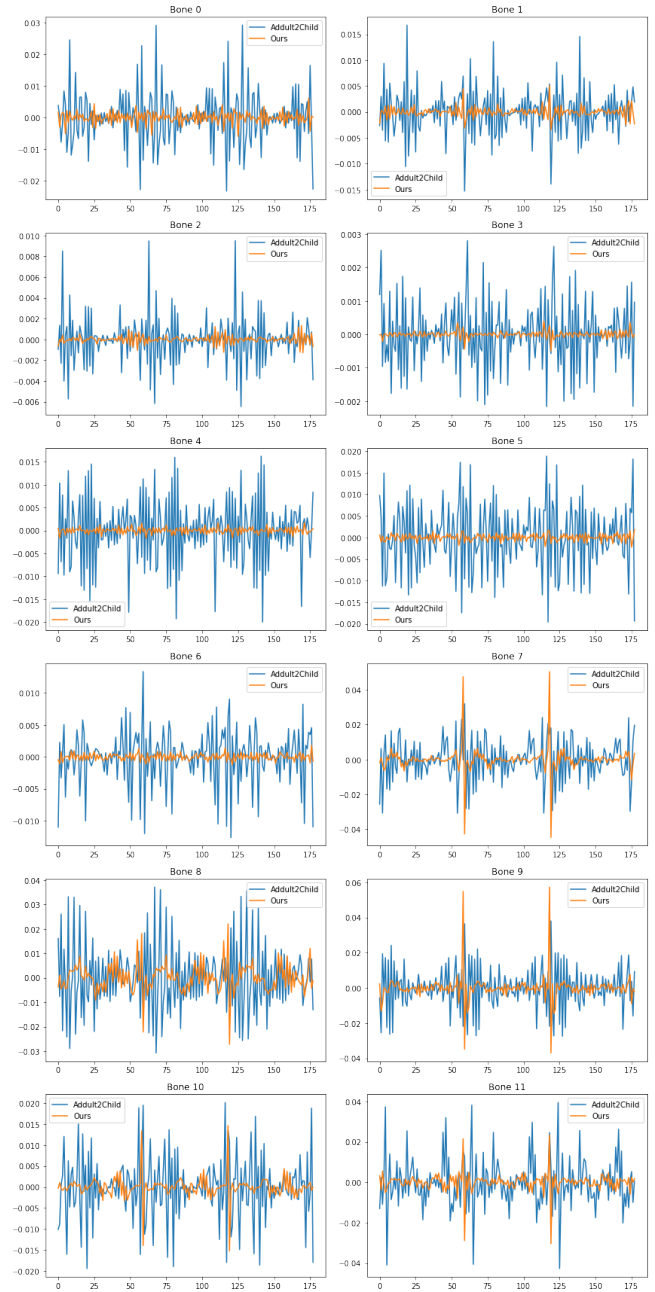


図 8 ボーン 0 から 11 における評価指標の比較。

## 5 結論

本研究では、異なるスタイル間のモーションデータの変換において、新たな損失関数を導入することにより、モーションの滑らかさが向上することを示した。しかし、モーションワード間の繋ぎ目のモーション生成がうまくできておらず、これについてはまだ課題が残されている。

## 文献

- [1] Yuzhu Dong, Andreas Aristidou, Ariel Shamir, Moshe Mahler, and Eakta Jain, "Adult2child: Motion Style Transfer using CycleGANs," MIG '20: Motion, Interaction and Games, Article No. 13, pp. 1–11, 2020.
- [2] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A.



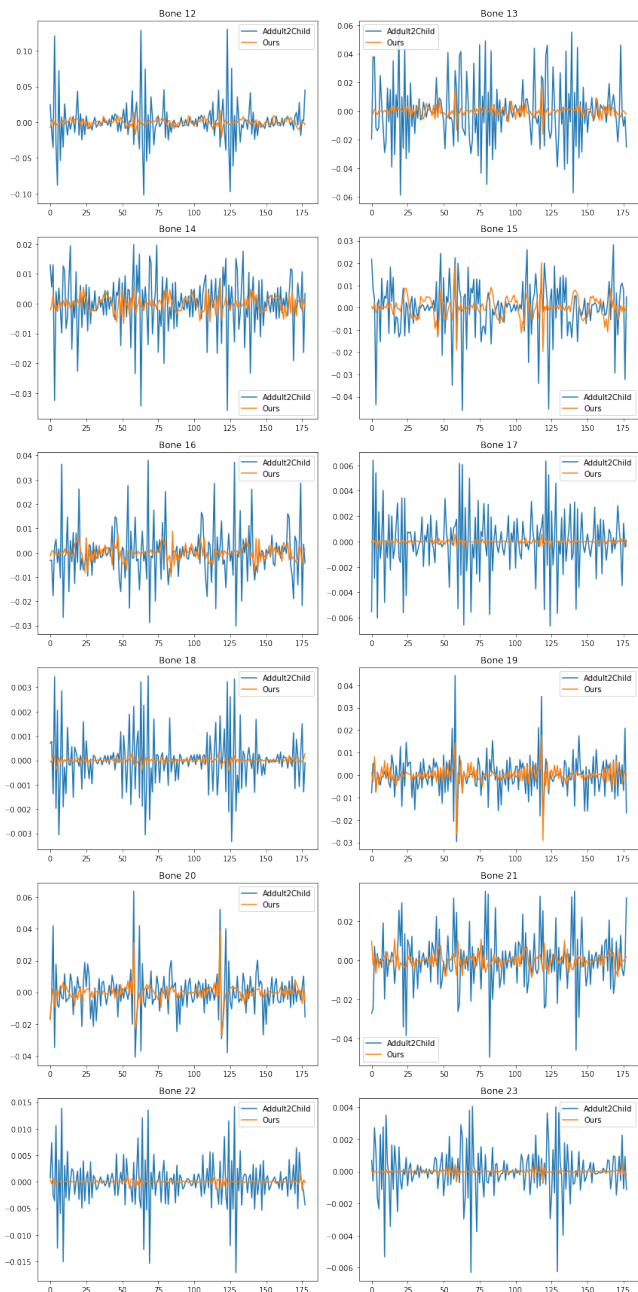


図9 ボーン 12 から 23 における評価指標の比較.

Efros, “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks,” Proceedings of the IEEE International Conference on Computer Vision (ICCV ’17), pp. 2242–2251, 2017.

- [3] Michael Gleicher, “Retargetting Motion to New Characters,” Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH ’98), pp. 33–42, 1998.
- [4] Kwang-Jin Choi and Hyeong-Seok Ko, “On-Line Motion Retargetting,” Proceedings of the 7th Pacific Conference on Computer Graphics and Applications (PG ’99), pp. 32, 1999.
- [5] Chris Hecker, Bernd Raabe, Ryan W. Enslow, John DeWeese, Jordan Maynard, and Kees van Prooijen, “Real-time motion retargeting to highly varied user-created morphologies,” SIGGRAPH ’08: ACM SIGGRAPH 2008 papers, Article No. 27, pp. 1–11, 2008.
- [6] Munetoshi Unuma, Ken Anjyo, and Ryoza Takeuchi,

“Fourier Principles for Emotion-based Human Figure Animation,” Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH ’95), pp. 91–96, 1995.

- [7] Andrew Witkin and Zoran Popović, “Motion Warping,” Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH ’95), pp. 105–108, 1995.
- [8] Kenji Amaya, Armin Bruderlin, and Tom Calvert, “Emotion from Motion,” Proceedings of the Conference on Graphics Interface ’96 (GI ’96), pp. 222–229, 1996.
- [9] Eugene Hsu, Kari Pulli, and Jovan Popović, “Style translation for human motion,” ACM Transactions on Graphics, Volume 24, Issue 3, pp. 1082–1089, 2005.
- [10] Yuzhu Dong, Aishat Aloba, Lisa Anthony, and Eakta Jain, “Style Translation to Create Child-like Motion,” Proceedings of the 39th Annual European Association for Computer Graphics Conference: Posters (EG ’18), pp. 31–32, 2018.
- [11] Tero Karras, Samuli Laine, and Timo Aila, “A Style-Based Generator Architecture for Generative Adversarial Networks,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4401–4410, 2019.