

動作の性別属性を変換する

男女双方向スタイル変換フィルタとその評価

○佐藤 志保理 山口 明彦 竹村 憲太郎 高松 淳 小笠原 司 (奈良先端科学技術大学院大学)

1. はじめに

近年、映画やゲームなどの3Dアニメーションやロボットの動作生成にモーションキャプチャが広く利用されている。しかし、個性を表現しようとする、動作の種類ごとに男性、女性、子ども、大人など、多様な属性を持った動作が必要となる。したがって、1つの動作サンプルから多様な動作を生成できれば有用であり、CGやロボティクスの分野で盛んに研究が行われている [1, 2, 3]。

例えば、Wangら [2] は、カーネル関数を用いたガウス過程回帰モデルを利用して、個性を表すパラメタ (スタイルパラメタ) を抽出する Multifactor Gaussian Process Models (MGPM) を提案した。一方、稲邑ら [3] は個々の動作を隠れマルコフモデルによって表現し、各モデル間の距離パラメタを用いて動作を内挿・外挿することにより動作を合成する手法を提案した。これらの手法は、サンプル動作間の補間や外挿によって動作を多様化する手法である。したがって、与えられた入力動作の抽象的な属性 (スタイル) を変化させたい場合、対極にあるスタイルの動作をサンプルとして保持していなければ、スタイルを変化させた動作を生成することは困難だと考えられる。

これに対して、我々は、性別や年齢の違いから生じる動作の差をモデル化することによって、動作を多様化する手法を提案した [4]。この手法は、抽象的な属性の差を直接モデル化するため、学習データと異なる動作に対してもスタイルを変換可能である。先行研究 [4] では、多様性の中でも性別による違いに着目し、任意の入力動作を女性らしく変化させるスタイル変換フィルタを提案した。このフィルタは、男性のポーズ (全身の関節角) に対して対応する女性のポーズへの差を出力する回帰モデルであり、入力動作のフレームごとにこのフィルタを適用することで、動作を女性らしく変換している。

本論文では、このフィルタを、任意の性別のポーズに対して男性・女性いずれのポーズへの差でも出力できるように改良する。これによって、与えられた任意の動作の性別属性を、任意に変更できるようになる。さらに、先行研究 [4] では行われていなかったアンケート調査を実施し、フィルタの妥当性を検証する。具体的には、男女の中間に位置するポーズから男性のポーズへの差をモデル化し、任意の動作の各ポーズに対して性別属性を変化させるための差を推定する。推定された差を入力動作に加算または減算することで、入力動作のスタイルを男女双方向に変換する。

フィルタ適用前後の動作に対して、男性/女性らしさを評価するためにアンケート調査を実施した。評価対象の動作は、モーションキャプチャによって取得した男女の歩行動作データベースから学習したフィルタ

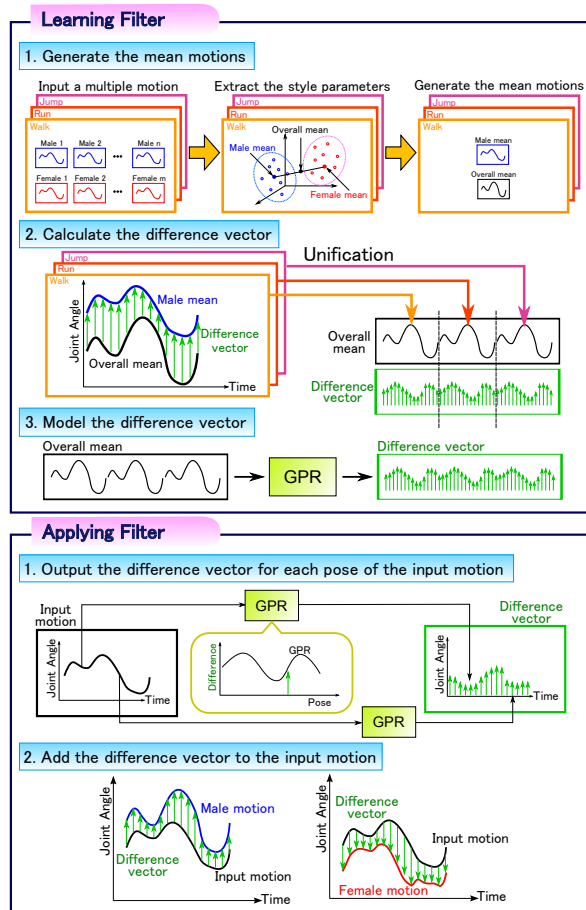


図1 男女双方向スタイル変換フィルタ

を、男女の歩行、走行、跳躍、スウィング動作に適用した動作である。アンケート調査を23人の被験者に対して実施したところ、フィルタを適用することで入力動作の性別属性を指定した方向に概ね変化させられていることが分かった。

以下、2章では男女双方向のスタイル変換フィルタの学習および適用方法について述べる。3章ではアンケートによる評価結果を示し、5章でまとめる。

2. 男女双方向スタイル変換フィルタ

本章では、提案するスタイル変換フィルタの学習および適用方法について述べる。男女双方向スタイル変換フィルタの概要を図1に示す。このフィルタは、任意のポーズ入力に対して、男性のポーズへの差を出力する回帰モデルである。以下、ポーズの差を差分ベクトルと呼ぶ。差分ベクトルは、複数の動作集合から男女全体の平均動作および男性の平均動作を生成し、計算する。なお、男女全体の平均動作から男性平均動作、女性平均動作への差は、符号は逆転するが大きさは等しいと仮定する。したがって、ここでは男性平均動作のみを用いてポーズの差を計算する。そして、ガウス

過程回帰モデルを用いて差分ベクトルをモデル化することで、フィルタを構成する。フィルタ適用時には、与えられた動作に対して、各時間フレームごとに差分ベクトルを推定する。変化させたい性別属性の変化量をパラメタとして与え、それを掛けた差分ベクトルを入力ポーズに加算する。以下詳細を述べる。

2.1 フィルタの学習

複数の被験者の、複数の種類で構成される動作の集合から、フィルタを学習する方法について述べる。動作の集合を \mathcal{M} と表す。ここで、各動作 $m \in \mathcal{M}$ には、動作の種類を表すラベル $g_m \in \mathcal{G}$ 、被験者を表すラベル $s_m \in \mathcal{S}_M \cup \mathcal{S}_F$ が与えられている。 \mathcal{G} , \mathcal{S}_M , \mathcal{S}_F はそれぞれ動作の種類、男性の被験者、女性の被験者の集合である。各動作 m は、状態 \mathbf{y}_m の時系列であり、 $m = \{\mathbf{y}_{m1}, \dots, \mathbf{y}_{mT_m}\}$ と表す。ただし、 T_m は動作 m のフレーム数である。状態 $\mathbf{y}_m = (\mathbf{v}_m, \mathbf{q}_m, \boldsymbol{\omega}_m)$ は、世界座標系における体中心座標系の並進速度 \mathbf{v}_m 、関節角 \mathbf{q}_m (ポーズ)、関節角速度 $\boldsymbol{\omega}_m$ で構成される。世界座標系における体中心座標系の姿勢は \mathbf{q}_m に含まれる。

2.1.1 平均動作の生成

本節では、動作集合 \mathcal{M} から男女全体の平均動作および男性の平均動作を生成する方法について述べる。男女全体の平均動作および男性の平均動作から差分ベクトルを計算するためには、動作間で位相が同じである(ポーズの対応がとれている)必要がある。そこで、この動作の位相あわせに Multifactor Gaussian Process Model (MGPM) [2] を用いる。MGPM は複数の隠れ変数を入力し、直接観測可能な対象変数を出力するガウス過程回帰モデルである。ここでは、隠れ変数を内部状態変数 \mathbf{x} および個人を表すスタイルパラメタ $\boldsymbol{\xi}$ とし、対象変数をポーズ \mathbf{y} とする。この MGPM を用いて動作の種類 g ごとに男性の平均動作と男女全体の平均動作を生成する。

具体的には、動作データ $\mathcal{M}_g = \{m | g_m = g\} \subseteq \mathcal{M}$ を MGPM でモデル化する。ここで、時刻 t における内部状態 \mathbf{x}_t を、 $\mathbf{x}_t = [\cos \theta_t, \sin \theta_t]$, $\theta_t = \theta_0 + t\Delta\theta$ としてモデル化する。ただし、 θ_0 は初期位相、 $\Delta\theta$ は位相のステップサイズである。次に、学習した MGPM のスタイルパラメタ $\mathcal{P} = \{\boldsymbol{\xi}_m | g_m = g\}$ から男性の平均スタイルパラメタ $\boldsymbol{\xi}_{Mg}$ および男女全体の平均スタイルパラメタ $\boldsymbol{\xi}_{Ag}$ を計算する。計算したスタイルパラメタと学習されたガウス過程回帰モデルを用いて、男性の平均動作 m_{Mg} と男女全体の平均動作 m_{Ag} を生成する。このとき、動作間で初期位相 θ_0 とステップサイズ $\Delta\theta$ を共通にすることで、内部状態変数 \mathbf{x} の系列をそろえる。そうすることで、男性の平均動作と男女全体の平均動作間でポーズの対応がとれた動作が、同じフレーム数 $T_{\Delta m_g}$ の動作データとして生成される。

2.1.2 差分ベクトルの計算

男性の平均動作 m_{Mg} と男女全体の平均動作 m_{Ag} から、各ポーズの差分ベクトルの集合 Δm_g を、

$$\Delta m_g = \{\Delta \mathbf{q}_{gt} | \Delta \mathbf{q}_{gt} = \mathbf{q}_{m_{Mg}t} - \mathbf{q}_{m_{Ag}t}, t = 1, \dots, T_{\Delta m_g}\} \quad (1)$$

で計算する。

各動作の種類 g ごとに差分ベクトルを計算し、それらを差分ベクトルの集合 $\Delta m = \bigcup_{g \in \mathcal{G}} \Delta m_g$ として統合する。また、対応する男女全体の平均動作も同様に、 $m_A = \bigcup_{g \in \mathcal{G}} m_{Ag}$ とする。

2.1.3 差分ベクトルのモデル化

本節では、前節で計算した差分ベクトル集合 Δm および男女全体の平均動作の集合 m_A を用いてガウス過程回帰モデルを学習する方法について述べる。モデルの入力は任意の動作の各ポーズ \mathbf{q} 、出力は対応するポーズの差分ベクトル $\Delta \mathbf{q}$ である。このとき、入力 \mathbf{q} から出力 $\Delta \mathbf{q}$ を求めるガウス過程回帰モデルを

$$\Delta \mathbf{q} = \mathbf{f}(\mathbf{q}) + \epsilon \quad (2)$$

と表す。ここで、 ϵ はガウスノイズである。したがって、学習時には、男女全体の平均動作および差分ベクトルの2組のデータ ($m_A, \Delta m$) を与え、モデルのパラメタを学習する。本論文で使用したカーネル関数は、

$$k_d(\mathbf{q}, \mathbf{q}') = \exp\left(\lambda - \frac{\mu \|\mathbf{q} - \mathbf{q}'\|^2}{2}\right) + b \quad (3)$$

で表されるガウスカーネルである。ここで、 d は各ポーズの次元であり、カーネルのパラメタである λ , μ , b は、スケーリングを施した共役勾配法によって最適化される。なお、ガウス過程回帰モデルの実装には Matlab のライブラリである NETLAB toolbox¹ を使用した。

2.2 フィルタの適用

次に、学習したフィルタを用いて、任意の動作を男性もしくは女性らしく変化させる方法について述べる。まず、入力動作 m_{in} に対して、学習したガウス過程回帰モデル \mathbf{f} を用いて差分ベクトルの集合 Δm を推定する。推定された差分ベクトル Δm を入力動作 m_{in} に加算または減算することによって、男性もしくは女性らしい動作 m_{out} を生成する。

$$m_{out} = \{\mathbf{q}_{m_{out}t} | \mathbf{q}_{m_{out}t} = \mathbf{q}_{m_{in}t} + \alpha \mathbf{f}(\mathbf{q}_{m_{in}t}), t = 1, \dots, T_{\Delta m}\} \quad (4)$$

ここで、 α は性別属性の変化量を表すパラメタである。なお、推定される差分ベクトルは男女の中間ポーズと男性のポーズの差であるため、女性と男性のポーズの差にするためには2倍する必要がある。したがって、通常、女性のポーズを男性のポーズに変換する場合は+2、男性のポーズを女性のポーズに変換する場合は-2に、 α の値を設定する。 $T_{\Delta m}$ は入力動作のフレーム数である。

3. 男女双方向スタイル変換フィルタの評価

本章では、フィルタへの入力動作および適用結果の男性/女性らしさをアンケートによって評価する。フィルタは、男女4人ずつ計8人の歩行動作を用いて学習し、男女の歩行、走行、跳躍、スウィング動作に適用する。学習データおよびテストデータには、Xsens Technologies 社製の慣性式モーションキャプチャシステム MVN を用

¹<http://www1.aston.ac.uk/eas/research/groups/ncrg/resources/netlab/>

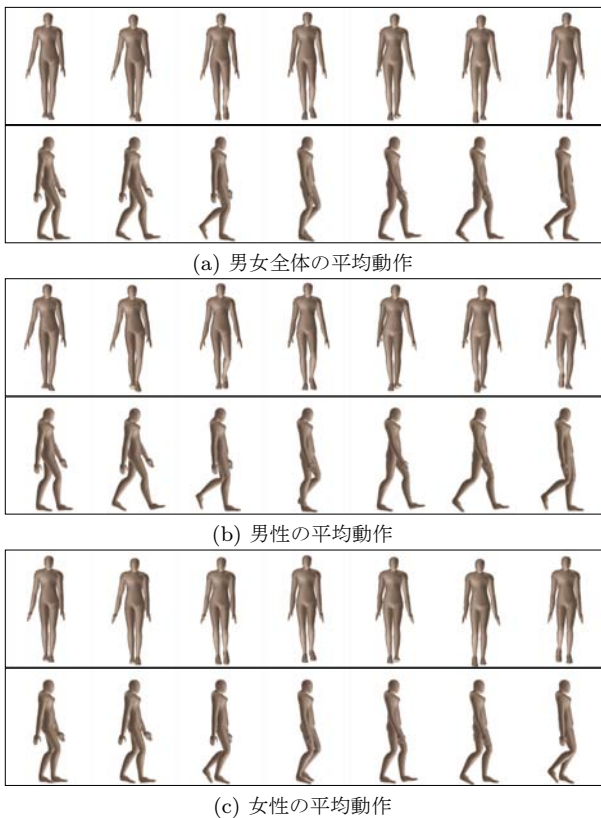


図2 歩行の平均動作（上段:正面図，下段:側面図）

いて計測したデータを 30Hz にダウンサンプリングして使用する。各動作において状態 $\mathbf{y}_m = (\mathbf{v}_m, \mathbf{q}_m, \boldsymbol{\omega}_m)$ は 129 次元であり，世界座標における体中心座標の並進速度 \mathbf{v}_m は 3 次元，関節角 \mathbf{q}_m は 63 次元，関節角速度 $\boldsymbol{\omega}_m$ は 63 次元である。

3.1 歩行動作からのフィルタの学習

男女 8 人の歩行動作を MGPM でモデル化し，男女全体の平均動作および男性平均，女性平均動作を生成した結果を図 2 に示す。MGPM のスタイルパラメタ ξ の次元は 5 に設定し，初期値は乱数で与えた。各平均動作を比較すると，特に腕，膝，足首の関節において差が大きいことが分かる。具体的には，各図の右端のフレームの正面図を見ると，男性の平均動作は膝が外側，女性の平均動作は膝が内側になっている。足首の関節においても同様の差がみられる。また，側面図より，歩幅については男性平均が一番広く，次いで男女全体の平均動作，女性の平均動作の順で狭くなっていることが分かる。本節で示した男女全体の平均動作と男性の平均動作を用いて差分ベクトルを計算し，フィルタを学習した。

3.2 フィルタの適用

前節で学習したフィルタを，歩行，走行，跳躍，テニスのスウィング動作に適用した結果をそれぞれ図 3, 4, 5, 6 に示す。差分ベクトルのスケール α は，女性の動作を男性らしく変換する場合は +2，男性の動作を女性らしく変換する場合は -2 とした。また，男性動作の男性らしさを強調する場合には +1，女性動作の女性らしさを強調する場合には -1 とした。各実験結果において，入力動作とフィルタ適用後の動作には，前

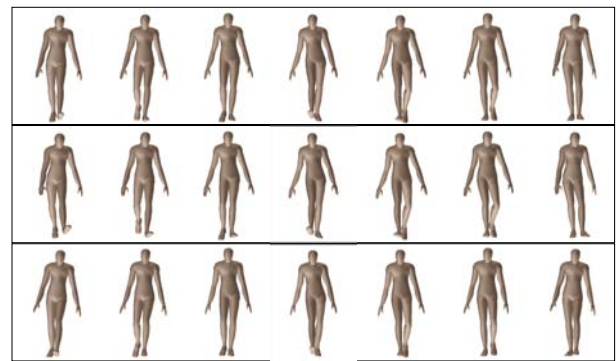
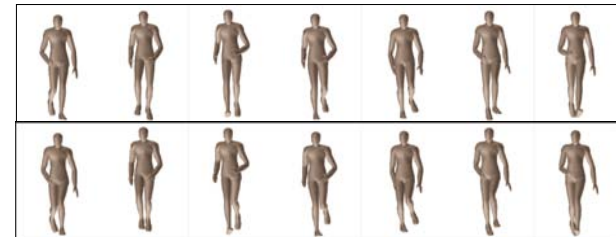


図3 歩行（上段:入力動作（男性），中段:男性らしさを強調，下段:女性らしく変換）



(a) 男性の動作を女性らしく変換



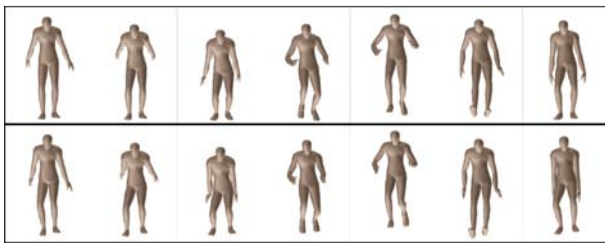
(b) 女性の動作を男性らしく変換

図4 走行（上段:入力動作，下段:フィルタ適用後）

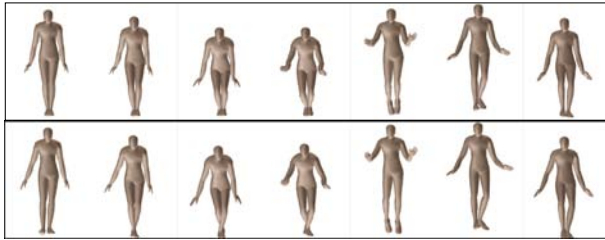
節で示した歩行動作の差と同様の違いがみられた。図 3 の左から 6 番目のフレームをみると，入力動作に対して，男性らしさを強調した動作は膝が外側になっている。それに対し，女性らしく変換した動作は膝が内側になっている。図 4, 6 の走行，スウィング動作も，歩行動作と同様に，主に下半身の関節に変化が確認できる。図 5 では，入力動作とフィルタ適用後の動作の違いが表れているものの，(a) の 1, 4, 6 番目のフレーム，(b) の 2, 3, 4, 7 番目のフレームでは差は小さいことが分かる。

3.3 アンケート評価

学習したフィルタを用いて，歩行（男女各 2 人），走行（男女各 1 人），跳躍（男女各 1 人），テニスのスウィング（男女各 1 人）の 4 種類の動作に適用し，入力動作とフィルタ適用後の動作の男性／女性らしさを評価した。被験者は 23 人であり，各動作の男性／女性らしさを -1, -0.5, 0, 0.5, 1 の 5 段階で評価するように依頼した。アンケート評価結果を図 7 に示す。各グラフで，横軸は提示した動作を表し，縦軸はアンケート評価結果の平均と ± 1 の標準偏差を表す。入力動作とフィルタ適用結果の間でアンケート評価の差について片側 t 検定を行ったところ，12 個中 9 個が 5% の有意水準で有意差がみられた。この結果より，フィルタを適用することで性別属性が指定した方向に有意に変化したと言える。有意差が認められなかった動作につ

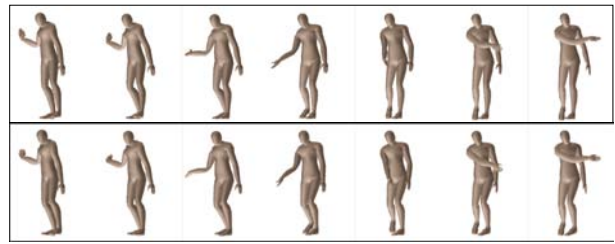


(a) 男性の動作を女性らしく変換

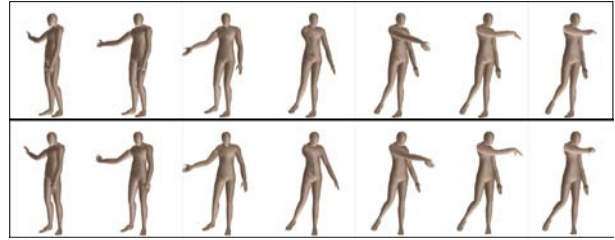


(b) 女性の動作を男性らしく変換

図5 跳躍 (上段:入力動作, 下段:フィルタ適用後)

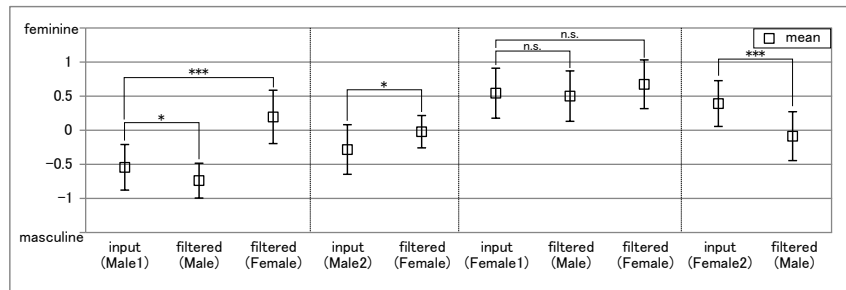


(a) 男性の動作を女性らしく変換

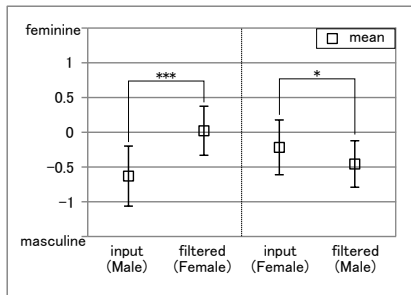


(b) 女性の動作を男性らしく変換

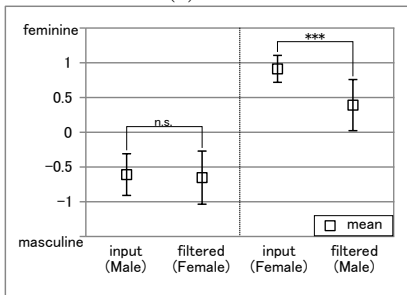
図6 テニスのスウィング (上段:入力動作, 下段:フィルタ適用後)



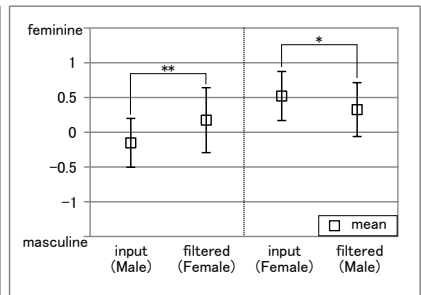
(a) 歩行



(b) 走行



(c) 跳躍



(d) テニスのスウィング

図7 アンケート評価結果 (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; n.s. 有意差なし)

いては、入力動作に対して推定された差分が、その動作を女性または男性らしくさせるには小さかったことが原因として挙げられる。また、図3 (a) の入力動作が女性1の場合では、歩行動作で学習したフィルタを用いているのにも関わらず、男性化・女性化どちらの場合にも評価に有意な差はなかった。この動作に含まれる女性の特徴は、フィルタの学習に用いた平均動作とは特徴が大きく異なるため、フィルタを適用してもその特徴が残ってしまったと考えられる。その他の動作については、動作の種類によっては、歩行動作から得られる男女の差だけでは十分ではないことも考えられる。

4. 結論

本論文では、任意の動作を男性もしくは女性らしく変化する男女双方向スタイル変換フィルタを提案し

た。実験により、提案手法を用いることで学習データと異なる入力に対しても動作を変化させることが可能であることを確認した。さらに、アンケート評価により、フィルタを適用することで入力動作の性別属性を指定した方向に変化させられていることが分かった。

参考文献

- [1] M. Brand and A. Hertzmann: "Style machines", Proceedings of SIGGRAPH 2000, pp. 183-192 (2000).
- [2] J. M. Wang, D. J. Fleet and A. Hertzmann: "Multifactor gaussian process models for style-content separation", the Twenty-Fourth International Conference on Machine Learning (ICML 2007), pp. 975-982 (2007).
- [3] 稲邑, 柴田: "動作パターンとシンボルを相互変換する原始シンボル空間における動作パターンの内挿・外挿", 日本ロボット学会誌, Vol.28, No.4, pp. 512-521 (2010).
- [4] 佐藤, 山口, 竹村, 高松, 小笠原: "任意の動作を女性らしく変化するスタイル変換フィルタ", ロボティクスメカトロニクス講演会 2011(ROBOMECH2011), pp. 1A1-O11 (2011).