

## HVGH: 高次元時系列データの深層圧縮と教師なし分節化

HVGH: Segmenting High Dimensional Time Series Data Using VAE and HDP-GP-HSMM

長野匡隼 \*1  
Masatoshi Nagano中村友昭 \*1  
Tomoaki Nakamura長井隆行 \*2  
Takayuki Nagai持橋大地 \*3  
Daichi Mochihashi小林一郎 \*4  
Ichiro Kobayashi高野渉 \*2  
Wataru Takano\*1電気通信大学  
The University of Electro-Communications\*2大阪大学  
Osaka University\*3統計数理研究所  
Institute of Statistical Mathematics\*4お茶の水女子大学  
Ochanomizu University

Humans recognize perceived continuous high dimensional information by dividing it into significant segments such as words and unit motions. We believe that such unsupervised segmentation is also an important ability for robots to learn topics such as language and motions. To this end, we have been proposed Hierarchical Dirichlet processes-Gaussian Process-Hidden Semi-Markov Model (HDP-GP-HSMM). However, it has a big drawback that the model cannot divide high dimensional time-series data, and it was required to extract low dimensional features in advance. To overcome this problem, in this paper, we propose Hierarchical Dirichlet Processes-Variational Autoencoder-Gaussian Process-Hidden Semi-Markov Model (HVGH). The parameters of HVGH are estimated by mutual learning loop of VAE and HDP-GP-HSMM. Hence, HVGH can extract features from high dimensional time-series data and, simultaneously, divide it into segments in an unsupervised manner. In the experiment, we use the various motion-capture data and show that our proposed model can estimate the correct number of classes and accurate segments compared with baseline methods.

## 1. Introduction

人は知覚した時系列情報を意味を持つ単位に分節・分類することで認識している。例えば、連続している音声波形を分節することで意味を持つ単位である単語を認識し、人の連続する動作を分節することで単位動作を認識することができる。人は明示的な分節点を与えられることなく自ら学習し、時系列情報を適切に分節・分類している。ロボットにおいても、言語や動作等を柔軟に学習するために、このような教師なしで分節・分類する能力は非常に重要であるといえる。そこで、我々は隠れセミマルコフモデルにおける出力をガウス過程としたノンパラメトリックベイズモデル Hierarchical Dirichlet Processes-Gaussian Process-Hidden Semi-Markov Model (HDP-GP-HSMM) [Nagano 18] を提案し、時系列データの教師なし分節化を可能とした。

しかし、HDP-GP-HSMM では、高次元データを扱うことができず、事前に特徴量抽出を行い低次元化する必要があった。分節化はこの特徴量の設計に大きく依存し、特に高次元データの場合、有効な特徴量を設計することは容易ではない。本稿では、この問題を解決するために、HDP-GP-HSMM に Variational Autoencoder [Kingma 14] を導入した Hierarchical Dirichlet Processes-Variational Autoencoder-Gaussian Process-Hidden Semi-Markov Model (HVGH) を提案する。図1が提案手法の概要である。HVGH では、VAE のエンコーダによって圧縮された潜在変数を HDP-GP-HSMM の入力として用い、さらに HDP-GP-HSMM による学習で得られるパラメータを VAE の事前分布のパラメータとして用いる。このように交互にパラメータを最適化し、VAE により単位系列を表現しやすいような潜在空間が学習される。本稿では、複数のモーションキャプチャデータを用いた実験により、他の教師なし分節化手法よりも分節化及びクラス数推定の精度が高いことを示す。

教師なし分節化では、Hidden Markov Model (HMM) が多く利用されている [Beal 01, Taniguchi 11, Fox 11, Matsubara 14]。しかし、HMM は構造が単純なために、複雑な動作を表現することが難しいと考える。一方、我々は Gaussian processes (GP) により時系列データを表現する。GP はノンパラメトリックな手法であり、HMM よりも複雑な動作を表現するのに適していると考えている。実際に HMM を

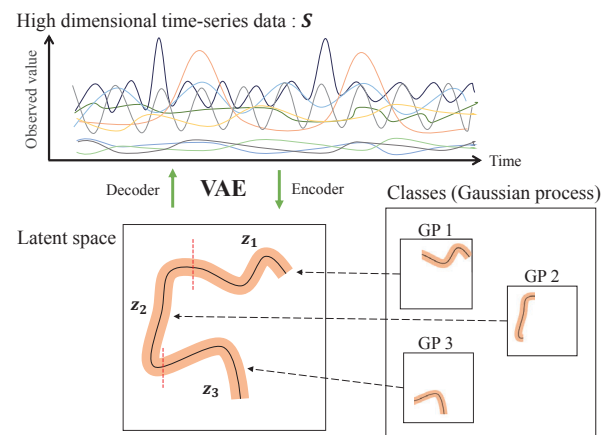


図1: Overview of the generative process of the proposed method.

ベースにした手法に比べ、GP を使った我々の手法の方が分節精度が高いことが確認できている [Nagano 18]。さらに、確率モデルに深層学習を組み合わせた研究も少ないが行われている [Johnson 17]。この研究では、VAE をクラスタリングに利用することで、複雑なデータの分類を可能としている。一方、本稿では、統計モデルである HDP-GP-HSMM と深層学習のモデルである VAE を組み合わせた高次元時系列データの分節化手法を提案する。

## 2. HVGH

図2に時系列データの生成モデルである提案手法 HVGH のグラフィカルモデルを示す。図において  $c_j (j = 1, 2, \dots, \infty)$  は単位系列のクラスであり、無限個のクラス数を仮定している [Teh 06]。 $\pi_c$  は、クラス  $c$  から各クラスへの遷移確率を表しており、 $\gamma$  をパラメータとする GEM 分布 [Pitman 02] (Stick Breaking Process [Geal 08]) によって生成された  $\beta$  と、 $\alpha$  をパラメータとする Dirichlet Process (DP) から生成される。

$$\beta \sim \text{GEM}(\gamma), \quad (1)$$

$$\pi_c \sim \text{DP}(\eta, \beta). \quad (2)$$

連絡先: 長野匡隼, 電気通信大学, 東京都調布市調布ヶ丘 1-5-1, n1832072@edu.cc.uec.ac.jp

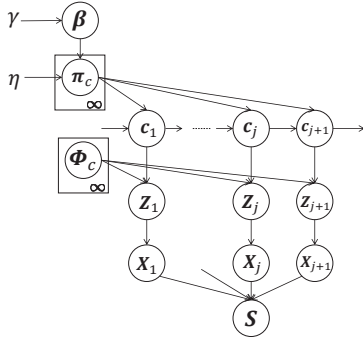


図 2: Graphical model of HVGH

$j$  番目のクラス  $c_j$  は、 $j-1$  番目のクラス  $c_{j-1}$  と遷移確率  $\pi_c$  によって決定される。ここでは、クラスと対応した  $\phi_c$  をパラメータとするガウス過程から単位潜在変数系列  $\mathbf{Z}_j$  が生成されることを仮定している。

$$c_j \sim P(c|c_{j-1}, \pi_c, \alpha), \quad (3)$$

$$\mathbf{Z}_j \sim \mathcal{GP}(\mathbf{Z}|\phi_{c_j}), \quad (4)$$

ただし、 $\phi_c$  はクラス  $c$  に分類された単位潜在変数系列の集合である。単位系列  $\mathbf{X}_j$  は、単位潜在変数系列  $\mathbf{Z}_j$  から VAE のデコーダ  $P_{dec}$  により生成される。

$$\mathbf{X}_j \sim p_{dec}(\mathbf{X}|\mathbf{Z}_j). \quad (5)$$

これらの単位系列  $\mathbf{X}_j$  を結合することで得られる系列を観測系列  $\mathbf{s} = \mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_J$  と呼び、 $\mathbf{Z}_j$  を結合することで得られる系列を潜在変数系列  $\bar{\mathbf{s}} = \mathbf{Z}_0, \mathbf{Z}_1, \dots, \mathbf{Z}_J$  と呼ぶ。また、以降、 $\mathbf{X}_j$  に含まれる  $i$  番目のデータ点を  $\mathbf{x}_{ji}$ 、 $\mathbf{Z}_j$  に含まれる  $i$  番目のデータ点を  $\mathbf{z}_{ji}$  と表記し、明らかな場合はその添字を省略する。

## 2.1 Gaussian Process (GP)

本稿では、単位系列のタイムステップ  $i$  における潜在変数  $z_i$  をガウス過程回帰で学習することで、各クラスの単位潜在変数系列を連続的に表現する。ガウス過程では、クラス  $c$  に分類された分節の集合を  $\phi_c$  とし、それと対応したタイムステップ  $i$  の組  $(i, \phi_c)$  が得られた時、タイムステップ  $i^{new}$  における  $z^{new}$  の予測分布は次式のガウス分布となる。

$$p(z^{new}|i^{new}, \phi_c, \mathbf{i}) \propto \mathcal{N}(z|\mu, \sigma^2) \quad (6)$$

$$\mu = \mathbf{k}^T \mathbf{C}^{-1} \mathbf{i} \quad (7)$$

$$\sigma^2 = c - \mathbf{k}^T \mathbf{C}^{-1} \mathbf{k} \quad (8)$$

ただし、 $k(\cdot, \cdot)$  をカーネル関数として、 $\mathbf{C}$  は、

$$C(i_p, i_q) = k(i_p, i_q) + \omega^{-1} \delta_{pq}. \quad (9)$$

を要素に持つ行列である。ここで、 $\omega$  はノイズを表現するハイパーパラメータを示す。また、 $\mathbf{k}$  は  $k(i_p, i^{new})$  を要素に持つベクトルであり、 $c$  は  $k(i^{new}, i^{new})$  である。ガウス過程では、カーネル関数を用いることで複雑に変化する時系列データを学習することが可能である。本稿では、ガウス過程回帰に一般的に用いられる以下のようなガウスカーネルを用いる。

$$k(i_p, i_q) = \theta_0 \exp\left(-\frac{1}{2}\theta_1 \|i_p - i_q\|^2 + \theta_2 + \theta_3 i_p i_q\right), \quad (10)$$

ただし、 $\theta_*$  はカーネルのパラメータである。

さらに、観測系列が多次元のベクトルの場合は、各次元が独立に生成されると仮定し、時刻  $i$  の出力値  $\mathbf{z} = (z_0, z_1, \dots)$  が

クラス  $c$  に対応するガウス過程から生成される確率  $\mathcal{GP}(\mathbf{z}|\phi_c)$  を以下のように計算する。

$$\begin{aligned} \mathcal{GP}(\mathbf{z}|\phi_c) &= p(z_0|i, \phi_{c,0}, \mathbf{i}) \\ &\quad \times p(z_1|i, \phi_{c,1}, \mathbf{i}) \\ &\quad \times p(z_2|i, \phi_{c,2}, \mathbf{i}) \cdots \\ &= \mathcal{N}(z|\mu_0, \sigma_0^2) \mathcal{N}(z|\mu_1, \sigma_1^2) \mathcal{N}(z|\mu_2, \sigma_2^2) \cdots \end{aligned} \quad (11)$$

この確率により類似する単位潜在変数系列を同じクラスへ分類することが可能となる。さらに、各次元が独立に生成されるため、 $\mathcal{GP}(\mathbf{z}_{ji}|\mathbf{Z}_c)$  の平均ベクトル  $\mu_c(i)$  と分散・共分散行列  $\Sigma_c(i)$  は、各次元の平均  $(\mu_0, \mu_1, \mu_2, \dots)$  と分散  $(\sigma_0^2, \sigma_1^2, \sigma_2^2, \dots)$  を用いて次式のようになる。

$$\mu_c(i) = (\mu_0, \mu_1, \mu_2, \dots), \quad (12)$$

$$\Sigma_c(i) = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \ddots \end{bmatrix} \quad (13)$$

HVGH では、 $\mu_c(i)$  と  $\Sigma_c(i)$  をパラメータとするガウス分布を VAE の事前分布とすることで、VAE と GP が相互に影響するモデルとなる。

## 2.2 Variational Autoencoder (VAE)

本稿では、Variational Autoencoder (VAE) [Kingma 14] を用いることで高次元時系列データを次元圧縮し低次元の潜在変数で表現する。VAE は、高次元データである入力  $\mathbf{x}$  と潜在変数  $\mathbf{z}$  の対応関係を学習するニューラルネットワークである。 $\mathbf{x}$  を入力としたエンコーダネットワークを通して平均  $\mu_{enc}(\mathbf{x})$  と分散  $\Sigma_{enc}(\mathbf{x})$  を推定する。潜在変数は、それらをパラメータとしたガウス分布を  $q_{enc}(\mathbf{z})$  に基づき確率的に決定される。さらにデコーダネットワーク  $p_{dec}$  を通して、データ  $\mathbf{x}'$  が生成される。

$$q_{enc}(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mu_{enc}(\mathbf{x}), \Sigma_{enc}(\mathbf{x})) \quad (14)$$

$$\mathbf{z} \sim q_{enc}(\mathbf{z}) \quad (15)$$

$$\mathbf{x}' \sim p_{dec}(\mathbf{x}|\mathbf{z}) \quad (16)$$

VAE では変分ベイズ法を用いて、尤度  $p(\mathbf{x})$  が最大となるエンコーダとデコーダのパラメータを決定する。また、通常の VAE では、事前分布として平均をゼロベクトル  $\mathbf{0}$ 、分散・共分散行列を単位行列  $\mathbf{I}$  としたガウス分布が用いられる。しかし、HVGH では、 $\mathbf{z}_{ji}$  を決定する際に、前節で計算した  $\mathbf{z}_{ji}$  が分類されたクラス  $c$  の  $\mu_c(i)$  と  $\Sigma_c(i)$  をパラメータとしたガウス分布を用いる。これにより、分節化に適した潜在空間が構成される。この VAE を用いて、観測系列  $\mathbf{s} = \mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_J$  をエンコーダによって潜在変数へ変換し、潜在変数系列  $\bar{\mathbf{s}} = \mathbf{Z}_0, \mathbf{Z}_1, \dots, \mathbf{Z}_J$  を得ることができる。

## 3. 学習

図 3 が HVGH のパラメータ推定の概要である。観測系列  $\mathbf{s}$  は VAE により潜在変数系列  $\bar{\mathbf{s}}$  へと変換される。HDP-GP-HSMM では、潜在変数系列  $\bar{\mathbf{s}}$  を  $\mathbf{Z}_0, \mathbf{Z}_1, \dots$  へと分節することで、 $\mathbf{z}$  の予測分布のパラメータ  $\mu_c(i)$  と分散・共分散行列  $\Sigma_c(i)$  を計算する。この予測分布を、VAE の事前分布として用いることで、VAE と HDP-GP-HSMM を相互に最適化することができる。

HDP-GP-HSMM のパラメータの学習では、潜在変数系列の分節とそのクラスを Blocked Gibbs Sampler [Jensen 95] によりサンプリングすることで推定する。まず、全ての観測系列を圧縮することで得られる潜在変数系列をランダムに分節・分

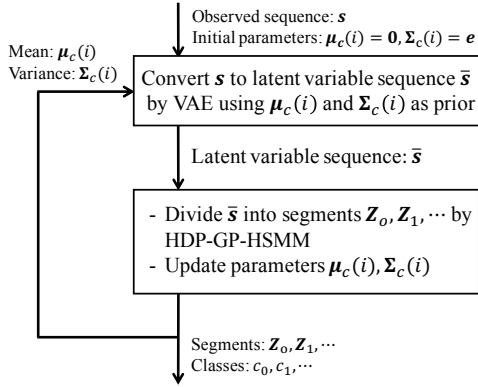


図 3: Overview of the parameter estimation of HVGH. The parameters are learned by mutual learning loop of VAE and HDP-GP-HSMM.

類することで初期化する。次に、 $n$  番目の潜在変数系列  $\bar{s}_n$  の分節  $Z_{n,j} (j = 1, 2, \dots, J_n)$  を、それらが分類されているクラスからすべて削除し、各クラスのガウス過程のパラメータ  $\phi_c$  と遷移確率  $P(c|c')$  を更新する。次に、次式に従い  $n$  番目の潜在変数系列の分節と、そのクラスをサンプリングする。

$$(Z_{n,1}, \dots, Z_{n,J_n}), (c_{n,1}, \dots, c_{n,J_n}) \sim P((Z_0, Z_1, \dots), (c_0, c_1, \dots) | \bar{s}_n). \quad (17)$$

これらの値を用いて各クラスのガウス過程のパラメータ  $\phi_c$  と遷移確率  $P(c|c')$  を更新する。以上の操作を全ての観測系列に対して繰り返すことで、分節とガウス過程のパラメータを交互に最適化することができる。しかし、このままでは遷移先のクラス数は無限個存在するため、各分節が各クラスに分類される確率を全て計算することは困難である。そこで、Slice Sampler [Geal 08] を用いることで、確率的にクラス数を有限個に制限し計算を可能とする。さらに、式 (17) では、あり得る分節・分類の全ての組み合わせに対して確率を与えなければならない、単純には計算できない。そこで、Forward Filtering-Backward Sampling [Uchiumi 15] を用いることでこの確率を計算し、サンプリングする。

VAE の学習では、尤度  $p(\mathbf{x})$  を最大化するエンコーダとデコーダのパラメータを決定する。一般的な VAE では、次式の変分下限を最大化するパラメータを決定する。

$$L(\mathbf{x}_{ji}, \mathbf{z}_{ji}) = \int q_{enc}(\mathbf{z}_{ji} | \mathbf{x}_{ji}) \log p_{dec}(\mathbf{x}_{ji} | \mathbf{z}_{ji}) d\mathbf{z}_{ji} - D_{KL}(q_{enc}(\mathbf{z}_{ji} | \mathbf{x}_{ji}) || p(\mathbf{z}_{ji} | \mathbf{0}, \mathbf{e})) \quad (18)$$

$\int q_{enc}(\mathbf{z}_{ji} | \mathbf{x}_{ji}) \log p_{dec}(\mathbf{x}_{ji} | \mathbf{z}_{ji}) d\mathbf{z}_{ji}$  は復元誤差を表している。また、 $p(\mathbf{z}_{ji} | \mathbf{0}, \mathbf{e})$  は  $\mathbf{z}_{ji}$  の事前分布であり、平均が  $\mathbf{0}$ 、分散が単位行列  $\mathbf{e}$  となるガウス分布である。 $D_{KL}(q_{enc}(\mathbf{z}_{ji} | \mathbf{x}_{ji}) || p(\mathbf{z}_{ji} | \mathbf{0}, \mathbf{e}))$  はカルバック・ライブラー・ダイバージェンスであり、正則化項となっている。一方 HVGH では、HDP-GP-HSMM により推定された  $\mathbf{z}_{ji}$  が分類されているクラス  $c$  の予測分布のパラメータである平均  $\mu_c(i)$  と分散  $\Sigma_c(i)$  を事前分布  $p(\mathbf{z}_{ji} | \mu_c(i), \Sigma_c(i))$  のパラメータとして用いる。

$$L(\mathbf{x}_{ji}, \mathbf{z}_{ji}) = \int q_{enc}(\mathbf{z}_{ji} | \mathbf{x}_{ji}) \log p_{dec}(\mathbf{x}_{ji} | \mathbf{z}_{ji}) d\mathbf{z}_{ji} - D_{KL}(q_{enc}(\mathbf{z}_{ji} | \mathbf{x}_{ji}) || p(\mathbf{z}_{ji} | \mu_c(i), \Sigma_c(i))) \quad (19)$$

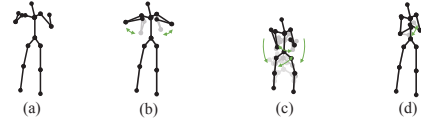


図 4: Four unit motions included in the chicken dance: (a) beaks, (b) wings, (c) tail feathers, and (d) claps.

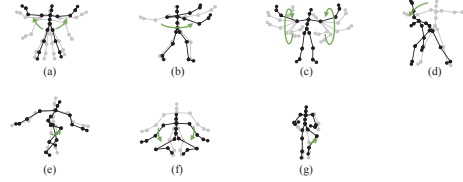


図 5: Seven unit motions included in the exercise motion1: (a) jumping jack, (b) twist, (c) arm circle, (d) bend over, (e) knee raise, (f) squatting, and (g) jogging

事前分布が  $p(\mathbf{z}_{ji} | \mathbf{0}, \mathbf{e})$  の場合、どのようなデータ点に対しても潜在変数の事前分布は  $\mathcal{N}(\mathbf{0}, \mathbf{e})$  となる。一方、 $p(\mathbf{z}_{ji} | \mu_c(i), \Sigma_c(i))$  を事前分布に用いた HVGH では、GP によって計算された平均  $\mu_c(i)$ 、分散  $\Sigma_c(i)$  をパラメータとしたガウス分布が事前分布になる。GP では時間が近いデータ点では近い値を出力するような制約が働くため、潜在変数  $\mathbf{z}_{ji}$  は、前後の潜在変数  $\mathbf{z}_{j,i-1}$  や  $\mathbf{z}_{j,i+1}$  と近い値となり、時系列データの特性を表現した潜在空間が学習される。さらに、これらのパラメータはクラス毎に計算されるため、クラス毎の潜在変数の特徴が反映される。

#### 4. 実験

提案手法 HVGH の有効性を検証するために、複数のモーションキャプチャデータに対して分節化を行ない、HDP-GP-HSMM [Nagano 18], HDP-HMM [Beal 01], HDP-HMM+NPYLM [Taniguchi 11], BP-HMM [Fox 11], Autoplaik [Matsubara 14] と比較した。評価指標は文献 [Nagano 18] と同様に、正規化ハミング距離、適合率、再現率、F 値を用いた。分節の誤差の許容範囲は観測系列の全フレーム数の  $\pm 5\%$  とした。

評価には、以下の 2 つのモーションキャプチャデータを用いた。

- **チキンダンス:** CMU Graphics Lab Motion Capture Database のチキンダンスのモーションキャプチャデータを用いた\*1。このダンスには、図 4 のように 4 つの単位動作が含まれている。
- **体操動作 1:** ダンス以外の動作として、CMU Graphics Lab Motion Capture Database の subject 13 の体操のモーションキャプチャデータを用いた。このデータは 3 つのシーケンスがあり、図 5 のように 7 個の単位動作が含まれている。

計算量を削減するため、全てのデータを毎秒 4 フレームにダウンサンプリングした。モーションキャプチャデータは 31 個の身体部位の方向が 3 次元のオイラー角で記述されており、各フレームは 93 次元のベクトルで構成されている。本実験では、この 93 次元をそのまま入力とした。また、HVGH ではハイパーパラメータが必要であり、経験的に決めた  $\lambda = 14.0, \theta_0 = 1.0, \theta_1 = 1.0, \theta_2 = 0.0, \theta_3 = 16.0$  を用いた。VAE の学習では入力データ点の 1/4 をミニバッチとして使用し、最適化には Adam [Kingma 17] を使い、更新は 150 回繰り返した。HDP-GP-HSMM の学習では、ブロックギブスサンプラーを 10 回繰り返した。さらに、VAE と HDP-GP-HSMM のループは、VAE の変分下限が収束するまで繰り返した。

\*1 <http://mocap.cs.cmu.edu/>: subject 18, trial 15



表 1: Segmentation results for the chicken dance.

	Hamming distance	Precision	Recall	F-measure	# of estimated classes
HVGH	0.23	0.86	0.86	0.86	4
VAE+HDP-GP-HSMM	0.31	1.0	0.71	0.83	4
VAE+HDP-HMM	0.74	0.15	1.0	0.26	11
VAE+HDP-HMM+NPYLM	0.48	1.0	0.86	0.92	7
VAE+BP-HMM	0.34	1.0	0.86	0.92	3
VAE+Autoplaît	0.66	0.0	0.0	0.0	1

表 2: Segmentation results for the exercise motion.

	Hamming distance	Precision	Recall	F-measure	# of estimated classes
HVGH	0.16	0.66	0.93	0.75	11
VAE+HDP-GP-HSMM	0.24	0.53	0.93	0.67	12
VAE+HDP-HMM	0.75	0.05	1.0	0.09	10
VAE+HDP-HMM+NPYLM	0.61	0.30	1.0	0.45	28
VAE+BP-HMM	0.58	0.29	0.97	0.44	7
VAE+Autoplaît	0.76	0.0	0.0	0.0	2

実際に 93 次元の時系列データをそのまま入力し分節化を行った。しかし、HVGH 以外の手法では 93 次元の高次元の時系列データそのままでは適切な分節はできなかった。そこで、HVGH とパラメータを同じにした VAE を使い、93 次元のデータを圧縮した 3 次元の潜在変数を入力として分節化を行なった。表 1, 2 にそれぞれの手法により 2 つのモーションキャプチャデータを分節した結果を示す。

VAE+HDP-GP-HSMM と VAE+BP-HMM は、chicken dance のデータを比較的高い精度で分節化できている。しかし VAE+BP-HMM の体操 1 のデータの分節化では、chicken dance のデータに比べて正規化ハミング距離の値も大きく、F 値は小さくなっている。これは、chicken dance が単純な動作の繰り返しであり、含まれる各単位動作の違いが明確であるため、HMM を用いたシンプルな手法であっても分節化が容易であったと考えられる。一方、ノンパラメトリックな手法であるガウス過程を用いている HDP-GP-HSMM と HVGH は、複雑な動作が含まれている体操を分節化することができている。さらに、HVGH ではガウス過程で学習される各タイムステップの平均と分散を VAE の事前分布とすることで、各単位動作を表現する潜在変数系列が学習できたため、VAE+HDP-GP-HSMM より精度が良くなったと考えられる。

また、提案手法では chicken dance のクラス数を正しく推定できた。しかし、体操 1 においては複雑な動作が含まれており、正解より多いクラス数が推定された。体操 1 では、正解のクラス数である 7 よりも多く推定された。これは、何もしていない静止状態を 1 つの単位動作として学習し、また左右対称に同じ動きをする動作を左側に行う場合と右側に行う場合の 2 つを単位動作として学習したためである。しかし、静止動作を 1 つの単位動作とすることは必ずしも誤りとは言えず、またモーションキャプチャデータは左側に行う動作と右側に行う動作は異なる値となるため別の動作として学習することは妥当であるといえる。さらに他の手法と比較すると、HVGH で推定されたクラス数は他の手法で推定したクラス数より正解に近く、HVGH が最も正しくクラス数を推定できていることがわかる。

以上の結果より、HVGH を用いることで、高次元の時系列データの分節化とクラス数の推定が可能であることが示された。

## 5. Conclusion

本稿では、VAE と HDP-GP-HSMM を相互学習することで、高次元の時系列データの分節、分類、クラス数推定を可能とした HVGH を提案した。提案手法では、VAE により単位動作の特徴を捉えた潜在変数系列が学習され、その潜在変数系列を用いることで高精度な分節化が可能となった。実際に実験により、HVGH が高次元時系列データの特徴を捉えて分節・分類が可能であることを示した。さらに、VAE と HDP-GP-HSMM をそれぞれ独立させて学習したモデルより、相互にパラメータ

を更新する HVGH の方が精度が高いことが示された。

しかし、HVGH はガウス過程を用いているため、 $N$  個のデータ点を学習させるのに  $O(N^3)$  の計算量がかかり、さらにそれを相互学習のループで繰り返すため、計算コストが非常に高くなる。そのために、HVGH では長い時系列データを分節することが困難である。そこで今後、ガウス過程の近似法を導入することによって計算コストの削減を考えている。

## 謝辞

本研究は JST CREST JPMJCR15E3, JSPS 科研費 JP18H03295 の助成を受け実施したものである。

## 参考文献

- [Beal 01] M. J. Beal, Z. Ghahramani, and C. E. Rasmussen, "The infinite hidden Markov model," In *Advances in Neural Information Processing Systems*, pp.577–584, 2001.
- [Fox 11] E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, "Joint modeling of multiple related time series via the beta process," arXiv preprint arXiv:1111.4226, 2011.
- [Geal 08] J. V. Gae, Y. Saatchi, Y. W. Teh, and Z. Ghahremani, "Beam Sampling for the Infinite Hidden Markov Model," *International Conference on Machine Learning*, pp.1088–1095, 2008.
- [Jensen 95] Jensen. Claus S, Kjærulff. Uffe and Kong. Augustine, "Blocking Gibbs sampling in very large probabilistic expert systems," *International Journal of Human-Computer Studies*, vol. 42, no.6, pp.647–666, 1995.
- [Johnson 17] Matthew James Johnson, David Duvenaud, Alexander B. Wiltschko, Sandeep R. Datta, Ryan P. Adams, "Composing graphical models with neural networks for structured representations and fast inference," arXiv preprint arXiv:1603.06277v5, 2017.
- [Kingma 14] Diederik P. Kingma, Max Welling, "Auto-Encoding Variational Bayes," arXiv preprint arXiv:1312.6114, 2014.
- [Kingma 17] Diederik P. Kingma, Jimmy Lei Ba, "ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION," arXiv:1412.6980v9, 2017.
- [Matsubara 14] Y. Matsubara, Y. Sakurai, and C. Faloutsos, "Autoplaît: Automatic mining of co-evolving time sequences," *ACM SIGMOD International Conference on Management of Data*, pp.193–204, 2014.
- [Nagano 18] M. Nagano, T. Nakamura, T. Nagai, D. Mochihashi, I. Kobayashi, and M. Kaneko, "Sequence Pattern Extraction by Segmenting Time Series Data Using GP-HSMM with Hierarchical Dirichlet Process," *International Conference on Intelligent Robots and Systems*, pp. 4067–4074, 2018.
- [Pitman 02] J. Pitman, "Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformations of an interval partition," *Combinatorics, Probability and Computing*, vol.11, pp.501–514, 2002.
- [Taniguchi 11] Taniguchi. T, and Nagasaka. S, "Double articulation analyzer for unsegmented human motion using Pitman-Yor language model and infinite hidden Markov model," In *IEEE/SICE International Symposium on System Integration*, pp.250–255, 2011.
- [Teh 06] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, "Hierarchical Dirichlet processes," *Journal of the American Statistical Association*, vol.101, no.476, pp.1566–1581, 2006.
- [Uchiumi 15] K. Uchiumi, T. Hiroshi, and D. Mochihashi, "Inducing Word and Part-of-Speech with Pitman-Yor Hidden Semi-Markov Models," *Joint Conference of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, pp.1774–1782, 2015.