

統計的情報処理に基づく時系列データのキーフレーム抽出・認識・再生

○ 稲邑 哲也 (東大) 谷江 博昭 (東大) 中村 仁彦 (東大/CREST)

Keyframe extraction and decompression for time series data
based on continuous Hidden Markov Models*Tetsunari Inamura, Hiroaki Tanie and Yoshihiko Nakamura
(Univ. of Tokyo, Dept. of Mechano-Informatics)**Abstract**— Memorization and generation of time series data is absolutely imperative for robot control and environment recognition. Memorization using keyframe representation is useful for motion generation and understanding of characteristic factor of time series data. In this paper, we propose an integration method of keyframe extraction and reproduction of time series data.**Key Words:** Keyframe, Hidden Markov Models, Time Series Data Recognition/Generation

1. はじめに

ロボットの行動制御や実環境認識などには時系列データの記憶と再現が必要不可欠である。この際、瞬間のデータであるキーフレーム表現を用いて記憶すると、ロボットの行動制御や環境の特徴的な要素の把握などに有効である。本稿ではキーフレームの抽出とそれからの時系列データの再現を同じ数理モデルで実現する方法について提案を行う。

ここで、ヒューマノイドのように大自由度を持つロボットの運動データや動画を扱うためには記憶容量のコストを削減するため、少ないパラメータで圧縮された記憶表現と再現手法が求められる。しかし、単なる符号化と復号化では容量コストに対処することはできても、時系列データの認識や生成に必要なコストを考慮したことにはならない。記憶表現が抽象化された表現ではなく、物理的な状態を保有したまま情報量を圧縮することがロボットにとって有利な表現形態となる。

運動の物理的な特徴を保有した圧縮表現であると考えられるのは運動データのある瞬間で切り取った「キーフレーム」と呼ばれる表現である。これはCGにおいて動画を表現するための基本的な手法で、時間と

もに変化する対象の特徴的な瞬間をキーフレームとして捉え、これを時系列方向に重ね合わせることで全体の運動を再現する手法である。近年はロボットの運動制御にも積極的に使われている¹⁾手法である。また、運動の認識にも親和性が高い表現形態である。

しかし、キーフレームの抽出基準は明確になっておらず、CGの分野ではアニメータの技術に依存している状態である。移動ロボットにおけるナビゲーションタスクを対象にして動画像からキーフレームを自動的に抽出する研究例²⁾も存在するが、前後するフレームとの識別可能性が基準となっており、動画を再現することは考慮されていない。

そこで本研究では、キーフレームを自動的に抽出する枠組みと、そのキーフレーム表現から時系列データを再現する枠組みを同時に実現するような統一的な数理モデルの提案を行う。

2. 連続分布型隠れマルコフモデルによる
キーフレーム表現とその再現

連続分布型 HMM は状態の有限集合 $Q = \{q_1, \dots, q_N\}$ 、ノード q_i から q_j へ状態が遷移する確率 a_{ij} の集合である状態遷移確率行列 $A = \{a_{ij}\}$ 、ノード q_i においてベクトル x を出力する確率分布 $b_i(x)$ の集合 B 、初期状態確率分布 $\pi = \{\pi_i\}$ 。以上の4項組 $\lambda = \{Q, A, B, \pi\}$ で表される。Fig.1のように、HMMは確率的に状態遷移を繰り返しながら、ベクトルを出力し、最終的に時系列データを出力することになる。出力確率分布 $b_i(x)$ は次のように M 個のガウス分布の線形結合で表される。

$$b_i(x) = \sum_{j=1}^M c_{ij} \mathcal{N}_{ij}(x; \mu_j, \Sigma_j) \quad (1)$$

ただし、 c_{ij} は結合係数、 μ と Σ はそれぞれガウス分布の平均ベクトルと分散行列である。ここで、平均 μ の次元は対象となる時系列データの次元と等しく、時系列データのある瞬間におけるパターンとして捉えることができる。特に、HMMの出力すべきベクトルの平均値であることから、時系列データにとって何らか

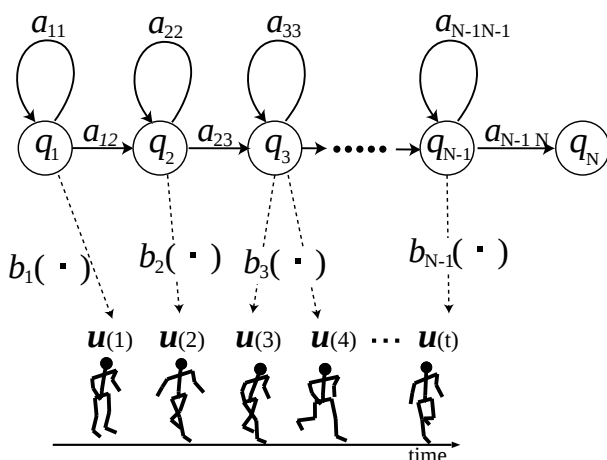


Fig.1 Motion Sequence and Hidden Markov Models

の特徴を捉えたベクトルであることが予想される．そこでキーフレーム $u = \{\mu, \Sigma\}$ と定義する．

本稿で使用する HMM は Left-to-Right モデルであるため，ノードのネットワーク構造は不変である．また初期状態は常に q_1 であるため，HMM のパラメータ $\{Q, A, B, \pi\}$ のうち $\{A, B\}$ の中に，時系列データを出力するダイナミクスが表現されていると考えられる．この二つのパラメータは EM アルゴリズム確率モデルのパラメータ推定アルゴリズムで獲得される．

以上で，隠れマルコフモデルを用いたキーフレームの抽出が可能となった．

3. 連続分布型 HMM からの時系列データ再現

逆にキーフレーム表現から元の時系列データを再現する際には，単純な補完を行うのではなく，HMM が保持しているダイナミクスを用いて正確なデータを再現する．一回の確率的データ生成試行を行うと確率的なゆらぎのため，運動の周期が異なり，関節角度にもノイズが大きく含まれる．この問題は以下のような方法で複数回の生成試行の平均をとることで解消させる．

step1 遷移確率に従って状態間を遷移させ状態遷移列 Q を得る．

step2 *step1* を n_q 回くり返し $Q_1, \dots, Q_m$ を得る．それらの平均を取って状態遷移列の平均 $\hat{Q}$ を得る．

step3 $\hat{Q}$ に従って各状態ノードから出力ベクトルを出力させ，出力時系列パターン O を得る．

step4 *step1* ~ *step3* を n_o 回くり返し $O_1, \dots, O_n$ を得る．それらの平均をとって最終的な出力時系列パターン $\hat{O}$ を得る．

ただし， n_q, n_o の値は実験的に定める．

4. 応用例

4.1 行動の模倣モデルへの応用

我々は今までにミメシス理論に基づく行動模倣のモデルを提案してきた³⁾．この枠組みでは，行動の最中の離散的な姿勢情報が行動要素として抽出され，その組み合わせとして元の行動を復元することで模倣行動を行う．

Fig.2 に関節角度を対象とした時系列データの再現結果を示す．グラフは全身の関節角のうち 1 次元分のデータを示している．グラフから分かるように，一回の生成試行 (一点鎖線) では正しい時系列データが生成されていないが，平均化を行うことで (実線) 元のデータとほぼ同等のデータを再現していることが分かる．

4.2 動画認識への応用

同様の枠組みを動画にも応用可能であることを確かめた．オフィスの廊下を進みながら 32×32 [pixel] の画像を撮影した動画から抽出された 20 個のキーフレームを Fig.3 に示す．部屋の入り口やライトのある場所など特徴的なシーンが抽出されていることが分かる．逆にこれらのキーフレームから動画を再現することも可能である．

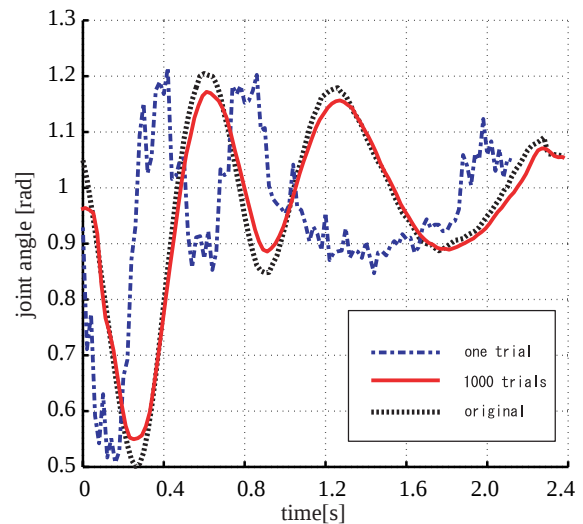


Fig.2 Original time series data and generated data

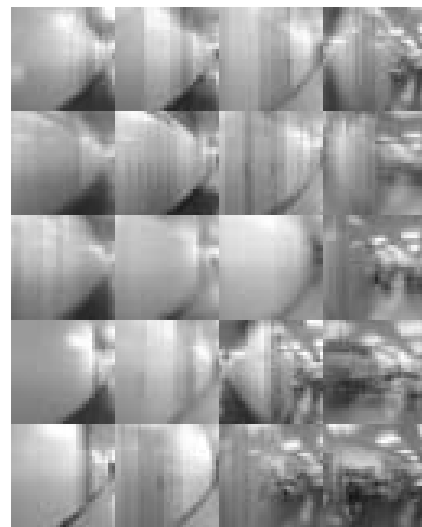


Fig.3 Extracted visual keyframes from video image

5. おわりに

本稿では，連続分布型 HMM を用いて時系列データからキーフレームを抽出し，逆に元の時系列データを再現する統一的な数理モデルを提案した．この手法は運動や動画などの時系列データの認識・再生にとどまらず，応用範囲として行動をシンボルとして抽象化し模倣を行うモデルへの応用が可能となる．この HMM をシンボルの原始的な表現形態と見立てて時系列パターンを空間的に配置し抽象化する事例を文献⁴⁾で解説する．

参考文献

- 1) 山根克，中村仁彦．ヒューマンフィギュアの全身運動生成のための協応構造化インターフェース．日本ロボット学会誌，Vol. 20，No. 3，pp. 335-343，2002．
- 2) 松本吉央，稲葉雅幸，井上博允．ビューベーストアプローチに基づく移動ロボットナビゲーション．日本ロボット学会誌，Vol. 20，No. 5，pp. 506-514，2002．
- 3) Tetsunari Inamura, Iwaki Toshima, and Yoshihiko Nakamura. Acquisition and embodiment of motion elements in closed mimesis loop. In *the Proc. of IEEE Int'l Conf. on Robotics & Automation*, pp. 1539-1544, 2002.
- 4) 谷江博昭，稲邑哲也，中村仁彦．統計的情報量を用いた原始シンボル空間の構成とそれに基づく原始シンボル系列の記号化．第 21 回日本ロボット学会学術講演会予稿集，2003．