

# ノンパラメトリックベイズ二重分節解析器による ドライバ意図変化点の予測手法

A Prediction Method for Changing Point of Driver's Intention  
by Nonparametric Bayesian Double Articulation Analyzer

長坂翔吾\*<sup>1</sup>  
Shogo Nagasaka

谷口忠大\*<sup>1</sup>  
Tadahiro Taniguchi

坂東 誉司\*<sup>2</sup>  
Bando Takashi

竹中 一仁\*<sup>2</sup>  
Takenaka Kazuhito

人見 謙太郎\*<sup>3</sup>  
Hitomi Kentaro

立命館大学\*<sup>1</sup>  
Ritsumeikan University

株式会社デンソー\*<sup>2</sup>  
DENSO CORPORATION

株式会社トヨタ IT 開発センター\*<sup>3</sup>  
Toyota InfoTechnology Center Co.,Ltd.

To assist a driver, it is important to estimate driver's intention. It is known that some methods based on analysis of double articulation structure can predict driver's intention with high accuracy and chunks which are estimated by the methods are similar to human's recognition. In this paper, we improve these model to predict duration time of changing point of driver's intention. To model duration of chunk, we use Hierarchical Dirichlet Process Hidden Semi-Markov Model instead of Sticky Hierarchical Dirichlet Process Hidden Markov Model. We test our proposed model using synthetic data. As a result, we show that proposed model can predict duration time more precisely than a comparative model.

## 1. はじめに

本論文では自動車運転挙動データからドライバの運転意図が一定である区間をチャンクとして抽出し、その未来の未観測データ上のチャンクの変化点を予測する手法の提案を行う。

運転挙動データはドライバのアクセル、ブレーキ操作等の運転操作より生成される。これらの運転操作はドライバの運転意図を実現するための手順であり、例えば「右折する」といった運転意図を達成するためには減速やハンドル操作等の運転操作を決まった順序で行う必要がある。このように運転意図はこれを実現するための運転操作の系列からなり、実際の運転挙動はこの運転操作から生成される。この構造は音声言語学における二重分節構造と類似していると言える。音声言語学において意味の伝達は単語により表現され、単語はそれらを構成する音素からなる。本研究では運転挙動データとの対応として音声言語学における単語をチャンク、音素をセグメントと定義する。また、チャンク、セグメント、観測データの関係を図1に示す。

運転挙動の二重分節性を示す研究として [Takenaka 12] が挙げられる。[Takenaka 12] では二重分節を仮定した解析で得られるチャンクの分節点と、被験者によりタグ付けされた自動車の進行方向、走行環境、シーンの切り替わり点のデータと比較することにより二重分節解析器により推定されるチャンクの分節点が人間の認識する運転挙動の切り替わりに近いことを示した。また [Taniguchi 12] では二重分節解析器により推定できるチャンクとその遷移確率を利用することにより、観測データから次のチャンクを高精度で予測できることを示している。

これらの先行研究より未来の二重分節点を予測することができれば、ドライバの運転意図が切り替わるタイミングを予測することができると考えられる。本研究では従来の二重分節解析器における持続時間分布の推定を改良することによりチャンクの持続時間を考慮した予測手法を提案する。

## 2. 提案手法

二重分節性を仮定したデータの生成過程を図1に示す。まず、チャンクの遷移が言語モデルより生成される。つぎに各チャ

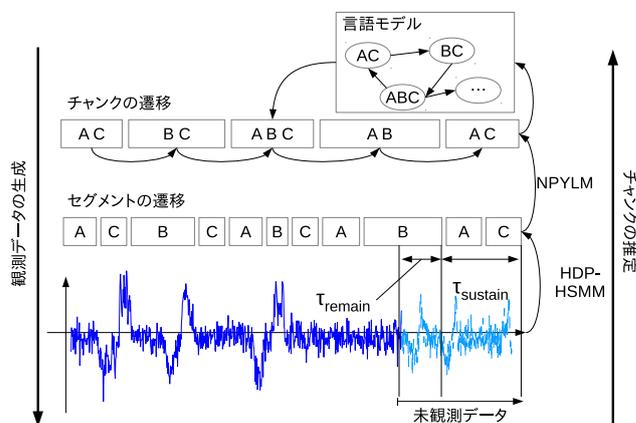


図 1: 二重分節性を持つデータの生成仮定と提案手法概要

ンクを構成するセグメントごとに自己遷移長が持続時間分布により決定され、各セグメントごとの観測値が生成される。

意図の変化点の予測を行うためには運転意図に対応するチャンクの予測とチャンクごとの持続時間分布の推定が必要となる。[Taniguchi 12] ではセグメントの推定に Sticky Hierarchical Dirichlet Process Hidden Markov Model (Sticky HDP-HMM) [Fox 08] を用いているため、セグメントの持続分布が指数分布となり変化点のモデリングに適していないと考えられる。本研究では Hierarchical Dirichlet Process Hidden Semi-Markov Model (HDP-HSMM) [Johnson 12] を用いて観測データより隠れ文字とその持続時間分布を推定した。HSMM は持続時間の分布を指数分布から一般的な持続時間分布に拡張している。本研究では持続時間分布にポアソン分布を仮定した。

チャンクの推定、予測には Nested Pitman-Yor Language Model (NPYLM) [持橋 09] を用いた。NPYLM を用いたチャンクの推定によりチャンクの遷移確率モデルを学習することができる。このモデルと HDP-HSMM により推定された各隠れ文字の持続時間分布パラメータによりチャンクの持続時間分布により、観測データから次のチャンクの切り替わり時間の予測を行うことができると考えられる。

## 2.1 二重分節解析によるチャンク変化点の時間予測

図 1 に示すように変化点までの残り時間は次の二つの時間からなることがわかる。

- $\tau_{remain}$  末尾文字が引き続き観測される時間。
- $\tau_{sustain}$  末尾文字部分を先頭を含むチャンクの持続時間。

これより次の変化点までの予測時間は

$$\tau = \tau_{remain} + \tau_{sustain} \quad (1)$$

となる。

### 2.1.1 末尾隠れ文字持続時間 $\tau_{remain}$

$\tau_{remain}$  は一部の観測が終了した後のポアソン分布により推定が可能となる。末尾文字が  $\tau_{obs}$  時間観測されている時、ポアソン分布のパラメータを  $\lambda$  とし、確率密度関数を  $\Phi$ 、累積密度関数を  $\Psi$  とすると、

$$E[\tau_{remain}] = \frac{\int_{\tau_{obs}}^{\infty} (x - \tau_{obs}) \Phi(x|\lambda) dx}{\int_{\tau_{obs}}^{\infty} \Phi(x|\lambda) dx} \quad (2)$$

$$= \lambda \frac{1 - \Psi(\tau_{obs} - 2)}{1 - \Psi(\tau_{obs} - 1)} - \tau_{obs} \quad (3)$$

により計算できる。

### 2.1.2 末尾チャンク持続時間 $\tau_{sustain}$

チャンク内での後続文字列については [Taniguchi 12] での手法と同様に Nested Pitman-Yor Language Model [持橋 09] で確率を予測する。その後、チャンクごとの予測確率の割合により持続時間分布の混合を行う。

教師なし形態素解析ではフォワードフィルタリングで前向き確率が計算される。この前向き確率より末尾チャンクの生成確率を計算し、その比により分布の混合を行う。

$$P(\tau_{sustain}|D) \propto \sum_K \sum_{w \sqsupset c_{T-K+1}^T} F_{K,w} f_D(\tau|w - c_{T-K+1}^T) \quad (4)$$

ここで文字列の引き算は、先頭文字列の削除を表し、 $w \sqsupset u$  は単語  $w$  が先頭部分に部分文字列  $u$  を含むことを示す。文字  $c$  のポアソン分布のパラメータを  $\lambda(c)$  とする。文字列  $w = [c_1, c_2, \dots, c_n]$  のポアソンパラメータはポアソン分布の再生性より

$$\lambda(w) = \sum_{i=1}^n \lambda(c_i) \quad (5)$$

となる。また  $F_{K,w}$  は未末端文字列が末尾  $K$  個チャンク化されてチャンク  $w$  となる確率であり、

$$F_{K,w} = \sum_{j=1}^{T-K} P(w|c_{T-K-j+1}^{T-K}) \alpha[T-K][j] \quad (6)$$

ただし、 $w \sqsupset c_{T-K+1}^T$  である。ここで  $\alpha[t][k]$  は  $t$  文字目に至る文字列が最後の  $k$  文字でチャンクを形成する確率を求めるための前向き確率であり、[持橋 09] の論文に同じである。

訓練データにより得たチャンク  $w$  のパラメータ  $\lambda(w)$  は式 (5) により計算できる。しかし、チャンクの切り替わり時間を正確に予測するためには未観測チャンクの持続時間分布の混合を考慮する必要がある。未知語に対する基底分布は文字  $n$ -gram によって定義されるが、その無限混合を同じ平均値を持つポアソ

ン分布で近似する。簡単のため  $u = c_{T-K+1}^T$  と置く。また、 $c_t$  を  $w$  を構成する文字とする。

$$\sum_{w \sqsupset u} \prod_{t=1}^{\#(w)-\#(u)} P(c_{T+t}|c_{T+t-1}) \Phi(\tau|\lambda(w) - \lambda(u)) \quad (7)$$

この和は、ポアソン近似と文字 bi-gram の性質により計算でき、この期待値は

$$\sum_{w \sqsupset u} \prod_{t=1}^{\#(w)-\#(u)} P(c_{T+t}|c_{T+t-1}) (\lambda(w) - \lambda(u)) \quad (8)$$

となる。これを  $w$  の長さ毎の和に書き換える。ここで  $P(c'|c)$  の EOW への遷移を除いた bi-gram の遷移確率行列を  $A$  とすると長さ  $n$  までの和は

$$\lambda^T A \left( \sum_{j=0}^{n-1} A^j \right) i_{c_T} \quad (9)$$

ただし  $A^0 = I$

により計算できる\*1。ここで  $i_{c_T}$  は  $c_T$  番目の次元のみが 1 で、残りが 0 の縦ベクトル。 $\lambda$  は  $\lambda(c)$  を並べた縦ベクトルである。

本研究では NPYLM の学習時の最大チャンク長の制約より、未観測チャンクの持続時間分布を以下の式により計算した。

$$\lambda^T A \left( \sum_{j=0}^{L_{max}-K-1} A^j \right) i_{c_T} \quad (10)$$

ここで  $L_{max}$  は NPYLM の最大チャンク長である。

## 3. 評価実験

提案モデルにより推定される混合ポアソン分布の評価のため実験を行った。実験では提案モデルにより推定される混合ポアソン分布のパラメータにより真のチャンクの切り替わりが起る時間の確率の計算し、モデルの評価を行った。

### 3.1 実験設定

提案モデル検証のため入力データとして、明確な二重分節構造を持つデータセットを以下の手順により作成した。

1. 文字単位での遷移確率パラメータを式 (12) により生成する。これは先頭文字 (BOW) からの遷移パラメータと、各文字から終端文字へ遷移する確率 (EOW) を含む。

$$\beta_{letter} \sim \text{GEM}(\alpha_{letter}) \quad (11)$$

$$\pi_{letter,i} \sim \text{Dir}(\beta_{letter}) \quad (12)$$

$i = 0, 1, 2, \dots, C$

2. BOW から遷移パラメータに従って EOW に遷移するまで文字を生成する。
3. 1-2 を繰り返し、チャンクのリストを得る。

\*1  $n \rightarrow \infty$  の場合、この期待値は Neumann 級数として以下の式で計算できる。

$$E[\tau] = \lambda^T A (I + A + A^2 + \dots) i_{c_T} = \lambda^T (I - A)^{-1} i_{c_T}$$

4. チャンクの遷移パラメータを式 (13) により生成する.

$$\beta_{chunk} \sim \text{GEM}(\alpha_{chunk}) \quad (13)$$

$$\pi_{chunk,i} \sim \text{Dir}(\beta_{chunk}) \quad (14)$$

$$i = 0, 1, 2, \dots, W$$

5. チャンクの遷移確率行列より 2 と同様の方法によりチャンク列を生成する.

6. 各文字に出力パラメータと持続時間分布パラメータを設定し, 生成されたチャンク列より観測データを生成する.

本実験では  $\alpha_{letter} = 1.0$ ,  $\alpha_{chunk} = 1.0$ , 文字種類数  $L = 10$ , チャンク種類数  $W = 20$ , チャンク列の数を 10 個に設定した. 提案モデルでの評価は以下の手順で行った.

1. HDP-HSMM によりデータセットの隠れ状態とその持続時間分布を推定する.
2. データセットを leave-one-out(LOO) 法により 10 個の検証データセットに分割し, それぞれの検証データセットで, 訓練データにより提案モデルの学習を行う.
3. テストデータの未隠れ状態を 300 個分削除し, 提案モデルにより混合ポアソン分布のパラメータを推定する.
4. 推定された混合ポアソン分布により真のチャンクの切り替わり時間にチャンクが切り替わる確率を計算し, 評価値とする.

提案モデルとの比較のため, 正解データより各チャンクの期待値の平均をポアソン分布のパラメータとするモデルを比較モデルとして設定した. チャンク列  $[w_1, w_2, w_3, \dots, w_{N_s}]$  からなるチャンク列  $s$  の期待持続パラメータ  $\lambda(s)$  は式 (15) により計算できる.

$$\lambda(s) = \frac{1}{N_s} \sum_{i=1}^{N_s} \lambda(w_i) \quad (15)$$

ここで  $\lambda(w)$  はチャンク  $w$  の期待持続パラメータを返す関数とする.  $w$  が隠れ状態  $l$  により構成されており,  $w = [l_1, l_2, \dots, l_{N_w}]$  とすると

$$\lambda(w) = \frac{1}{N_w + 1} \sum_{i=0}^{N_w} N_w - i = \frac{1}{2} N_w \quad (16)$$

により計算できる.

### 3.2 実験結果

図 2 に各モデルごとの次のチャンクの切り替わり時間までの期待持続値の比較を示す. また図 3 にモデルごとの持続時間の平均予測誤差を示す. ただしエラーバーは最小値, 最大値を表す. 図 3 より, 提案モデルは平均 50 以下の精度で持続時間の予測できることがわかる.

次に図 4 に提案モデルと比較モデルでの対数予測確率の比較を示す. また図 5 にモデルごとの平均予測対数確率を示す. ただし, エラーバーは最小値, 最大値を表す. 検証データセット 1, 3, 7 以外では提案モデルでの尤度が比較モデルに比べ高くなっていることがわかる.

検証データセット 3 の期待持続値の計算では提案モデルの精度が高いことがわかるが, 尤度は比較モデルに比べ低くなっている. 図 6 に同データセットで推定された持続時間分布を示

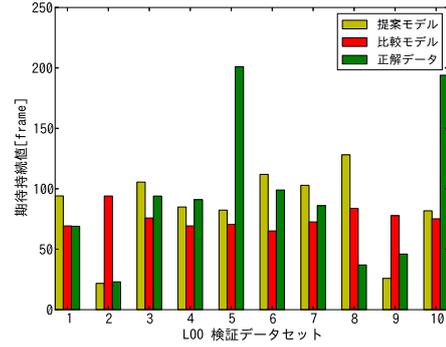


図 2: 検証データセットごとの次のチャンクの切り替わり点までの期待持続値の比較

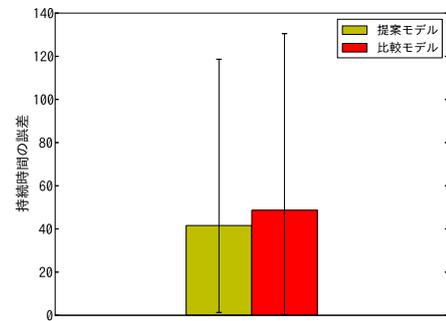


図 3: モデルごとの持続時間の平均予測誤差

す. グラフより提案モデルの持続時間分布では正解の持続値付近でのピークがないため, 尤度が低く評価されていることがわかる. また提案モデルにより推定された分布は 0 から 200 付近ではほぼ一様に分布しているため, 期待値が 100 付近となり, 期待値において比較モデルより精度が高くなったと考えられる.

次に検証データセット 8 での各モデルにより推定された持続時間分布を図 7 に示す. 提案モデルにより推定された分布は 40, 100 付近にピークを持つ混合ポアソン分布となっている. これにより期待持続値の計算では真の持続値より大きな値が計算されているが, 予測確率では正解の持続値に近いピークがあるため比較モデルに比べ高い値となっている.

## 4. おわりに

本論文では二重分節構造を持つデータよりチャンクの推定を行い, そのチャンクの切り替わり時間を予測する手法についての提案を行った.

本提案モデルでは従来の二重分節解析器での隠れ状態推定を HDP-HSMM によって行うことにより, 各文字の持続時間分布を推定することができ, NPYLM によるチャンクの分割学習により持続時間分布が混合ポアソン分布でモデル化できることを示した.

評価実験では二重分節構造を持つデータセットを生成し, 提案モデルと各チャンクの期待持続値により単一ポアソン分布により持続長をモデリングする比較モデルで, 期待持続値と尤度

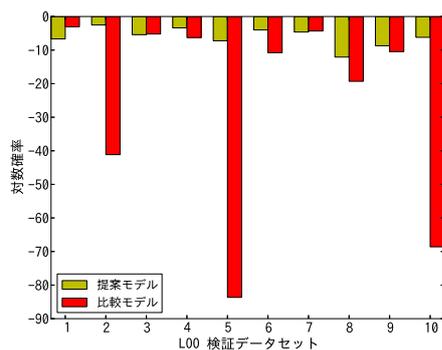


図 4: 検証データセットごとの対数予測確率の比較

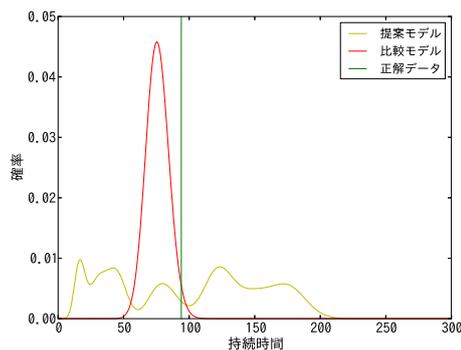


図 6: 各モデルで推定された持続時間分布 (データセット 3)

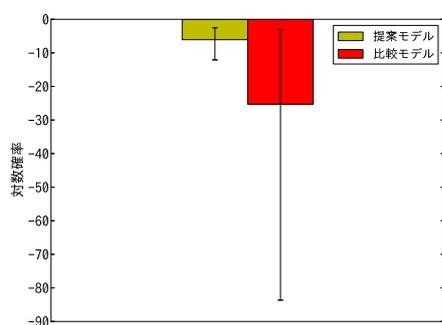


図 5: モデルごとの平均予測対数確率

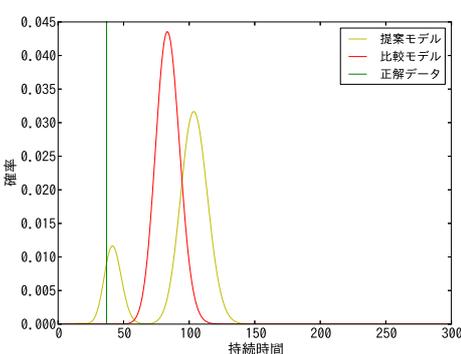


図 7: 各モデルで推定された持続時間分布 (データセット 8)

を評価することにより提案モデルの有効性を示した。また混合ポアソン分布でのモデル化を行うことにより、NPYLMにより推定された遷移チャンクの間違いの影響を抑えることができることを示した。

今後の課題として、実際の運転データを利用した解析が挙げられる。本研究ではチャンクの切り替わり時間の評価のため、二重分節構造を持つデータセットを作成したが、実際の自動車運転データでは、意図の変化点を真のチャンクの切り替わり点として、提案モデルによる解析、評価を行いたい。

また、切り替わり時間の予測の改善としてチャンクの推定精度の向上が考えられる。[Taniguchi 12]でのタスクでは、末尾に続くチャンクに含まれるセグメントのみを評価しているため、真の遷移チャンクと同じセグメント列を先頭部分に持つチャンクに高い遷移確率が割り振られた場合でも、高い精度で予測を行うことができるが、本研究のタスクではチャンクの切り替わり点が評価されるため、推定されるチャンクの長さが予測に影響する。

NPYLMでのチャンクの分割は教師なし学習であるため、頻出するチャンクの遷移がある場合、長いチャンクが推定されることや、各チャンク内に同一のセグメント列が出現する場合、短いチャンクが推定されることがある。精度の向上にはこれらのチャンクの長さを適切に推定するための手法が必要になる。

## 参考文献

- [Fox 08] Fox, E., Sudderth, E., Jordan, M., and Willsky, A.: The sticky HDP-HMM for systems with state persistence, in *Proceedings of the International Conference on Machine Learning* (2008)
- [Johnson 12] Johnson, M. J. and Willsky, A. S.: Bayesian Nonparametric Hidden Semi-Markov Models, *arXiv preprint arXiv:1203.1365* (2012)
- [Takenaka 12] Takenaka, K., Bando, T., Nagasaka, S., Taniguchi, T., and Hitomi, K.: Contextual scene segmentation of driving behavior based on double articulation analyzer, in *IROS*, pp. 4847–4852 (2012)
- [Taniguchi 12] Taniguchi, T., Nagasaka, S., Hitomi, K., Chandrasiri, N., and Bando, T.: Semiotic prediction of driving behavior using unsupervised double articulation analyzer, in *Intelligent Vehicles Symposium (IV)*, pp. 849–854 IEEE (2012)
- [持橋 09] 持橋 大地, 山田 武士, 上田 修功: ベイズ階層言語モデルによる教師なし形態素解析, *情報処理学会研究報告* (2009)