

# Kinectにより観測された人の動作を説明する確率的言語生成への取り組み

小林 瑞季<sup>†,a</sup>      小林 一郎<sup>†,b</sup>      麻生 英樹<sup>†,c</sup>

† お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻 情報科学コース

‡ 産業技術総合研究所 知能システム研究部門

a) [kobayashi.mizuki@is.ocha.ac.jp](mailto:kobayashi.mizuki@is.ocha.ac.jp) b) [koba@is.ocha.ac.jp](mailto:koba@is.ocha.ac.jp) c) [h.asoh@aist.go.jp](mailto:h.asoh@aist.go.jp)

**概要** センサなどによって観測される情報の殆どは時系列データであり、ビッグデータを扱う時代においては、観測された時系列データの中から有益な情報を取得し、その内容を理解する手法の開発が重要となる。時系列データの分析方法には、トレンドの予測や複数データ間の相関関係の分析など様々な方法が存在する。一方で、時系列データの内容を理解するには可視化などの手法が用いられている。しかし、ロボットなど複数のセンサによって取得された時系列データの情報に基づき状況を認識する必要がある場合、取得された情報をより抽象度の高いレベルで観測されたデータを表現する必要がある。そのことに着目し、我々は観測された時系列データの振る舞いを言語で説明する手法の開発を目指し、その一つとして、Kinect から得られた動画像の情報を入力とした確率的なテキスト生成手法を提案する。

**キーワード** Kinect, 時系列データ, SAX, 対数線形モデル, バイグラムモデル, 動的計画法

## 1 はじめに

センサなどによって観測される情報の殆どは時系列データであり、ビッグデータを扱う時代においては、観測された時系列データの中から有益な情報を取得し、その内容を理解する手法の開発が重要となる。時系列データの分析方法には、トレンドの予測や複数データ間の相関関係の分析など様々な方法が存在する。一方で、時系列データの内容を理解するには可視化などの手法が用いられている。しかし、ロボットなど複数のセンサによって取得された時系列データの情報に基づき状況を認識する必要がある場合、取得された情報をより抽象度の高いレベルで観測されたデータを表現する必要がある。そのことに着目し、我々は観測された時系列データの振る舞いを言語で説明する手法の開発を目指し、その一つとして、Kinect から得られた動画像の情報を入力とした確率的なテキスト生成手法を提案する。

## 2 関連研究

近年、マルチモーダル情報を言語情報で表す研究が盛んに研究されてきている。2013年の自然言語処理分野での最高峰とされる国際会議 ACL における最優秀論文賞には、Yu ら [1] による動画に映る人と物との相互作用を説明するテキスト生成研究が選ばれた。彼らの研究では、物体動作の認識の方法として、物体動作を一つの時系列データとして表し、隠れマルコフモデルを用いてその動作モデルを学習・認識し、言語を表す意味ラベルを付与することにより言語化を行っている。また、人の動

作を説明するテキスト生成として、Regneri ら [2] は調理を行っている動画像から、被写体の行動を説明するテキスト生成を行っている。彼らも Yu らと同様に人の動作とそれを説明する文章のペアから動作に対して適切な表現を検証している。Ushiku ら [3] は、静止画とそれを説明するキャプションのペアを学習し、類似した画像に類似したキャプションを生成する手法を提案している。また、動画像の言語化において、Takano ら [4, 5] は、一連の人の行動を運動記号の n-gram として表現し、運動記号と呼ばれる人の動作を示した意味ラベルから単語の連想構造を表現したモデルを通じて単語の並びを作る文生成手法を提案している。

また、Liang ら [6] は、テキストと意味との関係を学習する手法を提案しており、ここでは、イベントはデータベースのレコードで表せると仮定し、レコードと自然言語で表記された説明文との関連を機械学習によって取得している。Angeli ら [7] は、Liang ら [6] が提案したモデルに基づく潜在情報と表層情報をテキスト生成する手法を提案している。また、Konstas ら [8, 9] は、入力情報固有の構造を説明する確率的な自由文脈文法を定義しており、Liang ら [6] や Angeli ら [7] と同様に、データベースのレコードと説明文を用いている。彼らは、重みを加えたグラフによって文法を表現し、与えられた入力に対してもっとも適切な導出木を見つけることでテキスト生成を行っている。

本研究では、Liang ら [6] の手法を参考にして、視覚情報として取得される人と物体の振る舞いを表す時系列データと動作を表す自然言語の説明文との対応を対数線形モデルを用いて学習し、動作の意味を表す中間表現を

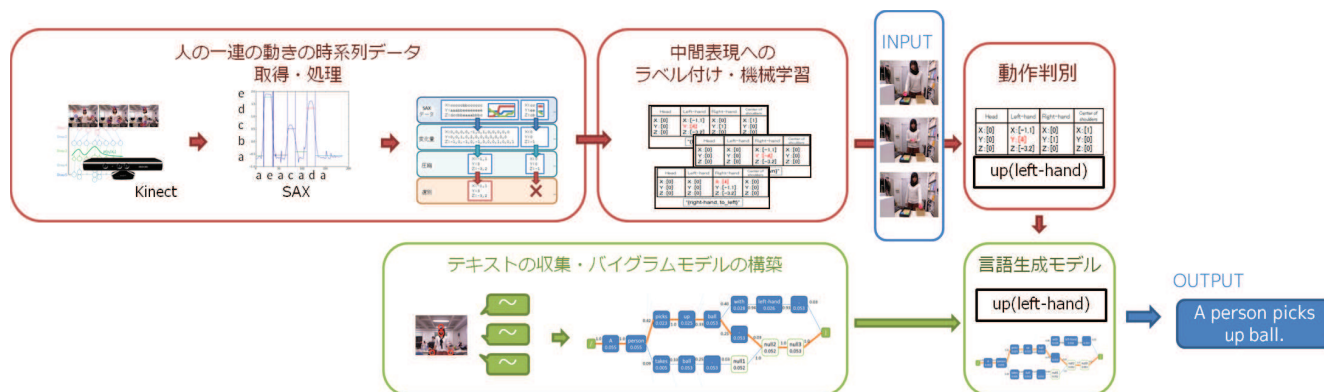


図1 動画像を入力とする確率的テキスト生成の枠組み

判別する。中間表現ごとに用意されたバイグラムネットワークを用いることにより、その動作の意味を表現する尤もらしい文を生成する。提案するテキスト生成手法には、文法を必要とするような複雑な文は生成することができないが、動画像を入力とし、尤度が高い表現で構成される文を容易に生成することが可能である。

### 3 視覚情報の言語化の枠組み

本研究の概要を図1に示す。まず、Kinect<sup>1</sup>がもつ人の骨格を追跡するライブラリとパーティクルフィルタを用いることで、人と物の動きを時系列データとして取得する。取得された時系列データはいくつかの次元圧縮作業を行い、データと自然言語の仲立ちをする中間表現とともにデータベースに格納される。その後、データベース内に蓄積された時系列データと中間表現の対応関係を機械学習することで、動作判別器を生成する。テキスト生成に用いられる言語資源は、人の動作の表現を被験者実験によって収集し、それぞれの中間表現に対してバイグラムモデルを構築する。これにより中間表現を選択すると、その中間表現に対応したバイグラムモデルが選択され、そのモデルに動的計画法を適用することで、人の動作を表現するもっともらしい語の組み合わせを選ぶことができる。

#### 3.1 時系列データの取得と処理

人間の動作の時系列データは、Kinectカメラを用いて取得する。Kinectの開発元であるMicrosoft社は、人間の骨格を推定できる標準ライブラリも提供しており、そのライブラリを用いると人の関節の3次元情報を推定することができる。

本研究では、RGB画像と深度センサー、またそれらを用いた人物の関節位置推定も使い、RGB動画像と人物の肩の右手・左手・右肘・左肘・肩の中心の5箇所のxyz

座標の時系列データを取得する(図2参照)。また、物体の動作の時系列データは、パーティクルフィルタ(3.1.1に詳説)を用いることで取得する。

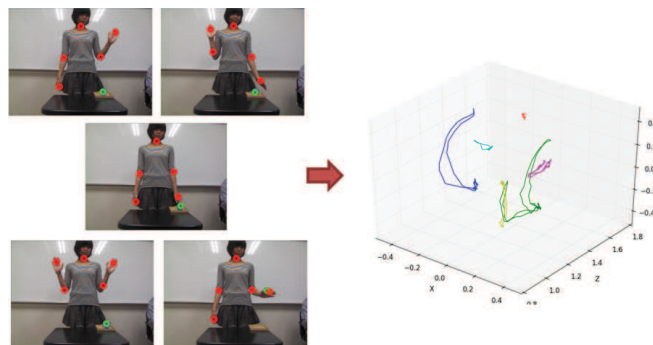


図2 Kinectとパーティクルフィルタを用いた時系列データ取得

#### 3.1.1 パーティクルフィルタ

パーティクルフィルタは状態の確率密度関数に単峰性・ガウス性といった制約はなく、非線形・非ガウス性の状態空間を推定することができる。時系列フィルタの1つである。コンピュータビジョンの分野においては、対象追跡などにパーティクルフィルタを使った手法が数多く提案されている。パーティクルフィルタは、観測できない状態ベクトル  $\mathbf{x}_t$  を観測可能な観測ベクトル  $\mathbf{y}_t$  から推定する。状態  $\mathbf{x}_t$  と観測値  $\mathbf{y}_t$  は、それぞれ以下に示すシステムモデル(1)、観測モデル(2)によって得られる。

$$P(\mathbf{x}_t | \mathbf{x}_{t-1}) \quad (1)$$

$$P(\mathbf{y}_t | \mathbf{x}_t) \quad (2)$$

$K$  個の重み付けされた粒子の集合  $X_t = \{(\mathbf{x}_t^{(k)}, \pi_t^{(k)})\}_{k=1}^K$  で状態  $\mathbf{x}_t$  の事後分布を表す。ここで  $\pi_t^{(k)}$  は個々のパーティクルの重みを示している。一般的なパーティクルフィルタのアルゴリズムを以下に示す。

<sup>1</sup><http://www.microsoft.com/en-us/kinectforwindows/>

- step1** 初期設定 : ランダムな  $K$  個の  $\mathbf{x}_{t-1}$  を生成する.  
これを初期パーティクルとする
- step2** 予測 : 式 (1) に従い  $\mathbf{x}_t^{(k)}$  を分布  $P(\mathbf{x}_t^{(k)} | \mathbf{x}_{t-1}^{(k)})$  からサンプルする
- step3** 尤度計算 :  $w_t^{(k)} = P(y_t | \mathbf{x}_t^{(k)})$   $\pi_t^{(k)} = w_t^{(k)} / \sum_k w_t^{(k)}$  のように各粒子の重み  $\pi_t^{(k)}$  を計算する
- step4** リサンプリング :  $\pi_t$  に比例した確率で  $\mathbf{x}_t$  を  $K$  個抽出する
- step5** 時間更新 :  $t \rightarrow t+1$  として step2~step4 を繰り返す

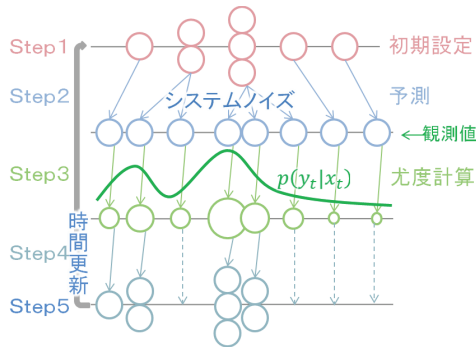


図3 パーティクルフィルタ

パーティクルフィルタは、図3に示すように、予測、尤度計算、リサンプリングを繰り返して時間更新を行う。このようにして、観測値  $y_t$  を多数の粒子  $X_t$  により追跡していく。本研究では、物体の色を尤度計算の対象とすることで、物体の追跡を行う。

### 3.1.2 時系列データの処理

人の骨格と物体の色を追跡することで得られた時系列データは、Symbolic Aggregation approXimation (SAX) [Lin 2003] を使い、文字列に変換する。

SAXとは、時系列データの近似表現方法の1つで、時系列データを文字列に変換する方法である。SAXを行う際、まずPAA(Piecewise Aggregate Approximation)というデータ圧縮作業を行う。長さ  $n$  の時系列データ  $C$  を用いて、 $w$  次元の空間ベクトル  $\bar{C} = \bar{c}_1, \dots, \bar{c}_w$  に変換すると仮定する。 $\bar{C}$  の  $i$  番目の要素は式 (3) を用いて計算される。

$$\bar{C}_i = \frac{w}{n} \sum_{j=\frac{n}{w}(i-1)+1}^{\frac{n}{w}i} C_j \quad (3)$$

本研究では動作判別の精度を高めるため、一般にデータを等間隔に  $w$  個のフレームに分割するところを、各データに動的計画法を用いもっともらしい区切りを取得することで、よりデータに沿った文字列を取得する(図4参照)。

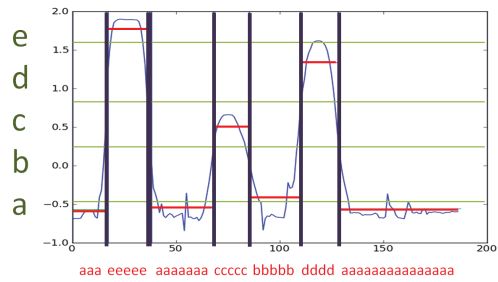


図4 動的計画法によるフレーム分割を用いたSAXの一例

SAXによって変換して得られた文字列から動作とみられる箇所を取り出す。ここでは、ある動画像データ中の全ての文字列において3つ前の文字から変化がなければ「動きがない」、変化があれば「動きがある」とみならず(図5参照)。

		動きがある	動きがない
左手	x	eeeeeeeeeeeeeeeeeeeeeeeeeeeeeeee	eeeeeeeeeeeeeeeeeeeeeeeeeeeeeeee
	y	cccccccccccccccccccccccccccccc	cccccccccccccccccccccccccccccc
	z	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
右手	x	fffffefeeeeeeeeefffffeedd	fffffefeeeeeeeeefffffeedd
	y	ccffgggggggggggggggffb	ccffgggggggggggggggffb
	z	eeeeeeeeeeeeeeeeeeeeeeeeeeeeee	eeeeeeeeeeeeeeeeeeeeeeeeeeeeee
左肘	x	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
	y	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
	z	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
右肘	x	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
	y	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
	z	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
肩の中心	x	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
	y	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
	z	dddddddddddddddddddddddddddddd	dddddddddddddddddddddddddddddd
緑のボール	x	dddddddddddddddddddddddddddddd	eeeeeeeeeeeeeeeeeeeeeeeeeeeeee
	y	eeeeeeeeeeeeeeeeeeeeeeeeeeeeee	eeeeeeeeeeeeeeeeeeeeeeeeeeeeee
	z	dddddddddddddddddddddddddddddd	eeeeeeeeeeeeeeeeeeeeeeeeeeeeee

図5 動きの抽出例

その後、「動きがある」とみなされた個所の文字列を変化量(図6中のアルファベットの下の数値)に変換し、圧縮する(図7参照)。これは同じ動作でも位置やスピードによっては文字列がある一定の間隔でずれたり文字列の長さが変化したりしてしまい、同じ動きとして学習されないためである。これにより、一定の間隔でずれてしまったものも長さが違うものでも、同じ動きとしてとらえることを可能とする。また、より特徴的な動作を抽出するために、圧縮された変化量うち最大の大きさが閾値に満たないものは取り除く(図7参照)。

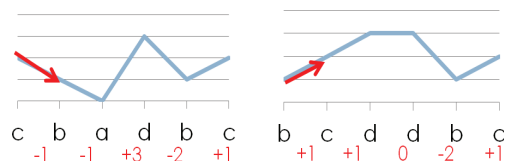


図6 文字列の変化量

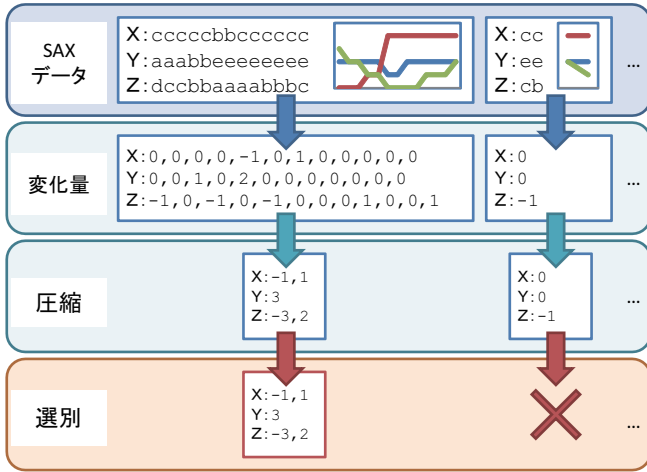


図7 データの圧縮・選別の例

### 3.2 中間表現

テキスト生成では、時系列データと自然言語文をつなぐ中間表現を用いることでテキスト生成に使う言語資源を選択する。中間表現は表1のように定義する。

表1 中間表現

action	中間表現	意味
up	“up(joint,null)”	upward movement
down	“down(joint,null)”	downward movement
pick	“up(joint,object)”	pick up movement
put	“down(joint,object)”	put movement

### 3.3 時系列データの動作判別

本研究では人の動作の判別を行うために対数線形モデルを用い、処理された時系列データと中間表現の対応を機械学習させる。3.1で述べた時系列データ処理を施したデータ  $d$  と、人の動作を表す中間表現  $r$  から構成した素性ベクトル  $\phi$  を用いて、式(4)の対数線形モデルを構成することで、データが与えられた下での各中間表現が選ばれる確率  $P(r|d)$  をモデル化した。ここで、 $Z_{d,w}$  は正規化係数である。

$$P(r|d) = \frac{1}{Z_{d,w}} \exp(\mathbf{w} \cdot \phi(d, r)) \quad (4)$$

### 3.4 バイグラムモデルによるテキスト生成

本研究では、バイグラムモデルを用いた単純なテキスト生成を行う。それぞれの動作に対しバイグラムモデルを構築するために被験者実験を行い、特定の動作に対して様々な自然言語表現を集めた。これにより、観測された時系列データに対して特定の中間表現が与えられたとき、言語資源としてバイグラムモデルを選択しテキスト生成を行う。

しかし、例えば同じ動作でも、ある人は10語で表現し、またある人は15語で表現するなど、表現の仕方が異なる。文の生成確率は、バイグラムネットワーク上での選択された単語の生起確率と単語間の遷移確率の積によって決まるため文中に含まれる語数が多くなるほど文の生成確率が下がってしまう。このことから、文の長さに依存しないテキスト生成が行えるよう、バイグラムモデルに null ラベルを導入する。

null ラベルは、文の中の単語として扱われ、他の単語と同じようにユニグラムとバイグラムの構成要素となる。このように null ラベルを扱うために、構成されたバイグラムモデルに対して動的計画法を適用する前に以下に続く前処理をそれぞれの文に対して行う。まず、全ての文で単語数の最大値  $max$ 、最小値  $min$  を得る。次に、 $max$  から  $min$  を引き、null に振る番号の最大値  $null\_max$  を求める。最後に、それぞれの文に対し、単語数が  $max$  に満たなければ、 $null\_max$  から1ずつ引いた値を、足りない数だけ文末から文頭に向け挿入していく。null ラベル導入のイメージを、図8に示す。

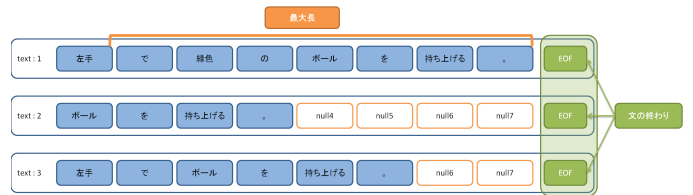


図8 null ラベル導入のイメージ

文中の各 null ラベルに違う番号をつけることによって別の単語として見なし、それぞれがバイグラムモデルの1要素として扱う。また、本研究ではバイグラムモデルを構築する際に、使用する文の取捨選択を行わないことで、多くの語と関連づけることができるため、より複雑なテキスト生成を行うことができる。

人の動作を説明するのによく用いられる文を生成するためには、このバイグラムモデルに動的計画法を適用することで尤度が最も高くなる単語の組み合わせからなる文を選ぶ。

## 4 実験

ここでは、「手を上げる」「手を下げる」「ボールを取る」「ボールを置く」という簡単な動作(図9)を言葉で表現することを目的とする。

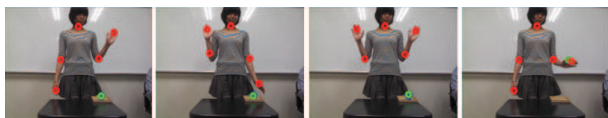


図9 言語化の対象となる動作

### 4.1 実験仕様

言語化の対象となる動作を「左手をあげる」「右手を上げる」「両手を下げる」「ボールを取る」「ボールを置く」「両手を上げる」「右手を下げる」「左手を下げる」の8つの基本動作から成ると定義する。また、それぞれの動作に対し、自然言語での説明文を生成することとする。被験者実験として、対象となる人の動作の Kinect ビデオを観賞し、それについて自然言語で説明してもらおうという実験を12人に対し行った。収集した日本語の説明文を形態素解析機 MeCab を用いて単語ごとに分け、これを言語資源としてバイグラムモデルを構築した。言語資源となった説明文の全文数、語数、語の種類数を表2に示す。

表2 収集された文の特徴

動き	文数	語数	語の種類数
左手を上げる	27	145	43
左手を下げる	28	146	47
右手を上げる	25	131	44
右手を下げる	31	174	51
両手を上げる	32	163	50
両手を下げる	30	165	53
ボールを取る	29	162	37
ボールを置く	29	170	43

テストデータには、「左手をあげる」「右手を上げる」「両手を下げる」「ボールを取る」「ボールを置く」「両手を上げる」「右手を下げる」「左手を下げる」の順で動作を行った Kinect 動画を使用した。動作判別には3.3で示した対数線形モデルを適用し、テキスト生成に使われる中間表現の判別に用いた。

### 4.2 実験結果

構築した識別機を用いてテストデータから判別された中間表現は、順に

1. “up((left\_hand),null)”
2. “up((right\_hand),null)”
3. “down((left\_hand,right\_hand),null)”
4. “up((right\_hand),green)”
5. “down((right\_hand),green)”
6. “up((left\_hand,right\_hand),null)”
7. “down((right\_hand),null)”
8. “down((left\_hand),null)”

となった。次に、選ばれた中間表現に対してあらかじめ構築されたバイグラムモデルに動的計画法を適用することで、動作を説明するもっともらしい文を生成する。

結果として、それぞれの動作に対して尤度の高かった上位3文を表3に示す。

### 4.3 考察

実験結果から、人の動作を正確に表現する文が生成出来ていることが確認できた。また、表3の生成文を見ると、いくつかの文で終端文字「EOF」が出てきていないことが分かる。これは、バイグラムモデルが集めた文に現れる語のバイグラムの組み合わせによって構成されているためである。これにより、バイグラムモデルへ null ラベルを加えた文が、集められたどの文よりも長く生成される可能性がある。また一方で、文が長くなればなるほど、その文の尤度が低くなっていく。したがって、集められた文より長い文は生成されないという仮定の下で、集めた文の最大の単語数を生成文の単語数とした。

## 5 まとめと今後の課題

本研究では、動画像中の人の動作を表現する確率的言語生成の枠組みを提案した。Kinect ビデオで抽出された人の動作およびパーティクルフィルタで取得された物体の軌跡は、時系列データとして取得され、SAXに動的計画法を用いもっともらしい区切りを導入して記号化する手法など、いくつかの次元圧縮手法を適用することで機械学習に適した形に変換される。また観測された人の動きを表現するために、被験者実験によって集められた自然言語文に基づきバイグラムモデルを構築し、動的計画法を適用することで、もっともらしい語の組み合わせを取得する。さらに、バイグラムモデルに番号を付け

表 3 各動作に対する生成文の上位 3 文

動作	生成文	尤度
1	● 左手, を, 上げる, 。, null_5, null_6, null_7, null_8, EOF	1.76e-12
	● 左手, を, 上げる, 。, null_4, null_5, null_6, null_7, null_8	1.57e-12
	● 左手, を, ひじ, を, 上げる, 。, null_4, null_5, null_6	1.19e-14
2	● 右手, を, 上げる, 。, null_4, null_5, null_6, null_7, null_8	8.63e-13
	● 右手, を, 上げる, 。, null_5, null_6, null_7, null_8, EOF	5.40e-13
	● 右手, を, すこし, あげる, 。, null_4, null_5, null_6, null_7	3.98e-15
3	● 両手, を, 下ろす, 。, null_4, null_5, null_6, null_7, null_8, null_9	2.91e-14
	● 両手, を, 下ろす, 。, null_5, null_6, null_7, null_8, null_9, EOF	2.68 e-14
	● 両手, を, 同時に, 下げる, 。, null_4, null_5, null_6, null_7, null_8	1.49e-16
4	● ボール, を, 持ち上げる, 。, null_5, null_6, null_7, null_8, null_9, EOF	2.64e-14
	● 左手, で, ボール, を, 持ち上げる, 。, null_6, null_7, null_8, null_9	2.03e-14
	● ボール, を, 左手, で, ボール, を, 持ち上げる, 。, null_6, null_7	2.77e-15
5	● ボール, を, 置く, 。, null_5, null_6, null_7, null_8, null_9, EOF	3.99e-15
	● ボール, を, 置く, 。, null_4, null_5, null_6, null_7, null_8, null_9	2.79 e-15
	● 左手, で, ボール, を, 置く, 。, null_5, null_6, null_7, null_8	2.57e-16
6	● 両手, を, 上げる, 。, null_5, null_6, null_7, null_8, null_9, EOF	2.53e-14
	● 両手, を, 上げる, 。, null_4, null_5, null_6, null_7, null_8, null_9	1.41e-14
	● 両手, を, 同時に, 挙げる, 。, null_4, null_5, null_6, null_7, null_8	1.32e-15
7	● 右手, を, 下ろす, 。, null_4, null_5, null_6, null_7, null_8	2.05e-12
	● 右手, を, 下ろす, 。, null_5, null_6, null_7, null_8, EOF	5.96e-13
	● 右手, を, 弧, を, 下ろす, 。, null_4, null_5, null_6	3.71e-15
8	● 左手, を, 下げる, 。, null_4, null_5, null_6, null_7, null_8, null_9, null_10	2.67e-15
	● 左手, を, 下げる, 。, null_5, null_6, null_7, null_8, null_9, null_10, EOF	8.90e-16
	● 左手, を, 伸ばし, た, 。, null_4, null_5, null_6, null_7, null_8, null_9	9.53e-18

た null ラベルを導入することにより、文生成に単語数の制限をつけずに自然言語文生成を行うことができた。また、提案手法はテンプレートによるテキスト生成ではなく、確率的なモデルによる生成であることから、例えばさらに文を収集すればそれに合わせて出力文も変化していくなど、資源となる文書によって様々な自然言語表現を得ることができる。

一方で、現段階では構文制約を取り入れてはいない。そのため今後の課題として、こうした知識を導入するとともに、より正確にイベントを説明するようなテキスト生成が行えるよう発展させていきたいと考える。また、中間表現とバイグラムモデルとの対応付けをより柔軟したり、一連の動作から自然言語文によって説明される動作を区切る問題にも取り組んでいきたい。

#### 参考文献

- [1] Haonan Yu and Jeffrey Mark Siskind, Grounded Language Learning from Video Described with Sentences, 51th Association for Computational Linguistics, Bulgaria, 2013.
- [2] Regneri, M., Rohrbach, M., Wetzels, D., Thater, S., Schiele, B., and Pinkal, M., Grounding Action Descriptions in Videos, 51th Association for Computational Linguistics, Bulgaria, 2013.
- [3] Yoshitaka Ushiku, Tatsuya Harada, and Yasuo Kuniyoshi. A Understanding Images with Natural Sentences. the 19th Annual ACM International Conference on Multimedia (ACMMM 2011), pp.679-682, 2011.
- [4] Takano, W. and Nakamura, Y.: Integrating whole body motion primitives and natural language for humanoid robots, Proc. IEEE-RAS Int. Conf. Humanoid Robots, pp.708-713, 2008.
- [5] Takano, W. and Nakamura, Y.: Incremental learning of integrated semiotics based on linguistic and behavioral symbols, Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, pp.1780-1785, 2010.
- [6] Percy Liang, Michael I. Jordan, Dan Klein 2009. Learning Semantic Correspondences with Less Supervision, ACL-IJCNLP
- [7] Angeli, Gabor and Liang, Percy and Klein, Dan, 2010. A simple domain-independent probabilistic approach to generation, Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, pp. 502-512, Cambridge, Massachusetts
- [8] Konstas, Ioannis and Lapata, Mirella, 2012. Unsupervised concept-to-text generation with hypergraphs, Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Montreal, Canada, pp.752-761
- [9] Konstas, Ioannis and Lapata, Mirella, 2012. Concept-to-text generation via discriminative reranking, Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers - Volume 1, pp. 369-378, Jeju Island, Korea
- [10] Lin, J., Keogh, E., Lonardi, S. and Chiu, B. 2003. A Symbolic Representation of Time Series, with Implications for Streaming Algorithms DMKD' 03