

# HMM を用いた手話の非手指信号の認識

加藤 里奈†

酒向 慎司†

北村 正†

†名古屋工業大学情報工学科

†名古屋工業大学大学院

## 1 はじめに

近年、聴覚障害者と健聴者との対話支援を目的としたコンピュータによる手話認識の研究が行われている。手話は手や指の動きにより単語の意味を表す手指信号 (Finger Signals) と、頭部動作や表情などによって文法や副詞を表す非手指信号 (Non Manual Signals, 以下 NMS) から構成される。手指信号と比較し NMS の体系化は難しく、NMS の認識に関する研究は少ない。しかし、手話文全体を理解するためには、手指信号だけでなく NMS の認識も重要である。先行研究 [1] では手話文中に出現する頭部動作 4 種類 (頷き、顎上げ、顎下げ、首振り) の認識実験を行ったが、話者の手話経験がなかったことから、現実的な評価データとは異なっていた可能性が考えられる。本研究では手話通訳士を対象とした HMM を用いた NMS の認識を試みる。

## 2 NMS の認識

NMS の体系化は難しいが、木村らの研究 [2] によると NMS には「表情」や「頭部動作」などがあり、各文法機能と対応付けることができる。本研究では、文末に出現することが多い頭部動作 3 種類 (頷き、顎下げ、顎上げ) の認識を試みる。頭部動作が 1 度出現する手話文を認識対象として想定し、HMM によってモデル化し、特徴量に頭の傾きを用いる。

### 2.1 頭の傾きの抽出

顔の各部位の 3 次元座標を取得できる Kinect v2 を用いて得られた額の座標  $(y_1, z_1)$  と頬の座標  $(y_2, z_2)$  から、次のように算出されたものを頭の傾き  $\theta$  として近似的に扱う。特徴量には  $\theta$  を平滑化したものを用いる。

$$\tan \theta = \frac{z_1 - z_2}{y_1 - y_2} \quad (1)$$

### 2.2 HMM の構成

モデルには複数の状態をもつ Left-to-Right 型 HMM を用いる。図 1 のように各状態ごとに対応する区間を切り出し、切り出された各区間に含まれる特徴パターンの平均と分散をパラメータとして設定した。本手法では、頭部動作が発生する際には頭の傾きが大きくなると仮定し、頭が傾いていない区間 (静止) と頭が大きく傾いた区間 (頷き、顎下げ、顎上げの生じる区間) の 4 種類を状態として定義する。区間の切り出しは目視によって行われた。状態数は、各頭部動作の時間的長さが異なるため、頭部動作ごとに経験的に設定した (表 1)。括弧内の数字は状態数を表す。なお、確率密度関数には 1 次元単一正規分布を用いた。

## 3 実験

実験用データは、頭部動作が文末に 1 回生起する、平叙文、肯否疑問文、命令文を各々 72 文ずつ用意し、手話

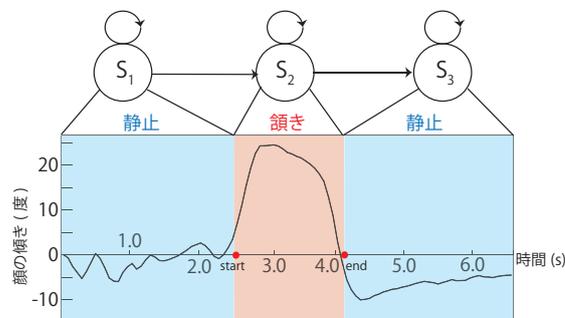


図 1 HMM の構成

文の収録は手話通訳士 1 名を対象に行った。NMS が目視で確認できないデータを除いて、学習に 66 個、認識に 125 個のデータを使用した。作成した HMM をもとに、頭部動作 3 種類の認識を行う。

### 3.1 実験結果

実験結果を表 2 に示す。3 種類の頭部動作全体の正解率は 79 % であった。誤認識の原因として、動作の類似した頷きと顎下げのモデルの作り分けが十分でなかったことや、手話に伴い NMS に由来しない頭の動きが生じたことが挙げられる。

## 4 まとめ

HMM を用いた NMS の認識手法を提案し、手話通訳士を対象とした頭部動作の認識実験を行った。その結果、頭部動作 3 種類を 79 % の正解率で認識することができた。今後の課題として、HMM の学習方法の見直しや、複数の NMS が出現する手話文への対応などが挙げられる。

### 参考文献

- [1] 不破大樹：“Kinect を用いた手話の非手指信号の認識”，電子情報通信学会，ISS-SP-163，2015。
- [2] 市田泰弘，木村晴美：“初めての手話”，日本文芸社，1995。

表 1 各状態の種類と状態数

生起する動作	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>
静止	静止 (1)	静止 (1)	静止 (1)
頷き	静止 (3)	頷き (5)	静止 (3)
顎下げ	静止 (3)	顎下げ (10)	静止 (3)
顎上げ	静止 (3)	顎上げ (2)	静止 (3)

表 2 非手指信号の認識結果

文の種類	生起する頭部動作	正解率	総合正解率
平叙文	頷き	61.1 %	79.2 %
肯否疑問文	顎下げ	95.6 %	
命令文	顎上げ	77.3 %	