

# 体位変換指導における学習支援を目的とした動作分類手法の検討

## Study of motion classification methods for learning support in postural change guidance

小林 大将\*1, 大井 翔\*1, 森谷利香\*2, 山本裕子\*3, 野間春生\*4

Osuke KOBAYASHI\*1, Sho OOI\*1, Rika MORIYA\*2, Yuko YAMAMOTO\*3, and Haruo NOMA\*4

\*1 大阪工業大学 \*2 摂南大学 \*3 畿央大学 \*4 立命館大学

\*1Osaka Institute of Technology \*2Setsunan University

\*3Kio University \*4Ritsumeikan University

Email: osuke.kobayashi@mix-lab.net, sho.ooi@outlook.jp

**あらまし：**体位変換は介護や看護教育において重要な技能であるが、熟練者の動作を客観的に捉えることは困難である。本研究では、教育支援に向けた基礎的検討として、リザバーコンピューティングを用いた体位変換動作の分類を行った。結果として、骨格データおよび音声データを用いた結果、両者を併用することで分類精度が向上し、動作判別の可能性が示された。

**キーワード：**リザバーコンピューティング、体位変換、動作分類

## 1. はじめに

ALS（筋萎縮性側索硬化症：Amyotrophic Lateral Sclerosis）は神経難病の一つであり、神経難病は、現在の医学では根治的な治療法が確立されておらず、多くの患者が長期間にわたる療養生活を余儀なくされている。神経難病患者は様々な症状により日常生活に大きな支障をきたし、身体や心理、社会的にも苦痛を抱えることが明らかになっている。

本研究では、リザバーコンピューティング (RC) を用いて、体位変換時の動作分類を実施し、それらが看護において使用可能なのかを検証することを目的とする。これにより、看護教育における実践知の定量的理験と指導支援することを目指す。

## 2. 先行研究

看護技術教育において、熟練看護師が現場経験で培った「コツ」や「ノウハウ」といった暗黙知を新人看護師へ教育するには困難である。真嶋ら(1)は、映像メディアを活用した学習が看護スキルの修得に有効であることを示している。本研究では、骨格情報として抽出・定量化し、後述する動作分類実験における特微量の一部として活用した。

## 3. 提案手法：リザバーコンピューティング RCによるマルチモーダル学習

本研究では、音声情報と骨格情報を活用して、体位変換動作の分類を行った。音声情報は、介助者の患者への声掛けに基づいて動作のタイミングを決定し、骨格情報はその動作のデータを取得し、関節角度や姿勢の変化を反映する。

図1のように、入力時刻  $t$  における  $M$  次元の入力ベクトルを  $\mathbf{v}(t) \in R^M$ 、 $N$  次元のリザバー状態ベクトルを  $\mathbf{x}(t) \in R^N$ 、 $L$  次元の出力ベクトルを  $\mathbf{y}(t) \in R^L$  と定義する。本システムにおける入力ベクトル  $\mathbf{v}(t)$  は、体幹および四肢の関節の骨格

情報 (14 次元) と tf-idf によって抽出された 20 単語の音声特微量 (20 次元) を統合した  $M = 34$  次元の多次元時系列データである。RC の状態更新および出力は式 1・2 で定義する。

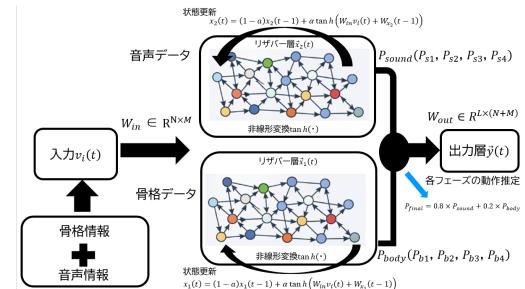


図1: リザバーコンピューティングのシステムフロー

$$\mathbf{x}(t) = (1 - \alpha)\mathbf{x}(t-1) + \alpha \tanh (\mathbf{W}_{in}\mathbf{v}(t) + \mathbf{W}\mathbf{x}(t-1)) \quad (1)$$

$$\mathbf{y}(t) = \mathbf{W}_{out} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{v}(t) \end{bmatrix} \quad (2)$$

ここで、 $\mathbf{W}_{in} \in R^{N \times M}$  は入力層からリザバー層への重み行列、 $\mathbf{W} \in R^{N \times N}$  はリザバー層内の再帰結合重み行列、 $\mathbf{W}_{out} \in R^{L \times (N+M)}$  は出力層の重み行列である。 $\alpha$  はリーケ率を表すパラメータであり、過去の状態をどの程度保持するかを決定する。RC では、 $\mathbf{W}_{in}$  および  $\mathbf{W}$  は学習過程において更新されず、 $\mathbf{W}_{out}$  のみを教師データに基づいて学習する。

RC の基本モデルを提唱した Jaeger らの研究(2)より、リザバー層における非線形変換には、双曲線正接関数  $\tanh(\cdot)$  を用いた。

$$\mathbf{W} = \rho_{target} \frac{\mathbf{W}_{init}}{\rho(\mathbf{W}_{init})} \quad (3)$$

出力に関しては、独立して学習された音声モデルおよび骨格モデルの出力を統合し、最終的な動作分類を行う手法として、出力層における後段統合を採用した(式 4)。

$$\mathbf{p}_m = [p_{m1}, p_{m2}, p_{m3}, p_{m4}]^\top \quad (m \in \{\text{audio, body}\}) \quad (4)$$

ここで、 $\mathbf{p}_{\text{audio}}$  および  $\mathbf{p}_{\text{body}}$  はそれぞれ音声モデルと骨格モデルから得られる推定確率ベクトルである。これらの異なるモダリティから得られる情報を次式に示す重み付き線形結合した。

$$\mathbf{p}_{\text{final}} = w_a \mathbf{p}_{\text{audio}} + w_b \mathbf{p}_{\text{body}} \quad (5)$$

ここで、 $w_a$  および  $w_b$  は各モダリティに対する重み係数であり、 $w_a + w_b = 1$  を満たす。本研究における重み係数の設定にあたっては、10分割交差検証を実施した。検証の結果、本システムでは重み係数を  $w_a = 0.8$ ,  $w_b = 0.2$  と定義した。

最終的な動作カテゴリの判定は、統合された確率ベクトルから、一番値が大きいものを選択するものとした。

## 4. 実験

本実験では、体位変換における動作分類の検証を目的とし、骨格情報と音声情報を用いたマルチモーダルデータ収集実験を図2のように実施した。実験協力者として、看護経験のない学生16名（男性5名、女性11名）を募集し、1人あたりの総データ数は80個撮影したため、合計1280個のデータ収集を行った。データ収集は、介助動作を行う介助者役とベッド上に横たわる患者役に役割を交代させながら実施し、体位変換における4つの動作「患者の手を胸に置く」、「患者の両膝を立てる」、「回転動作」、「体位を安定させる動作」を記録した。

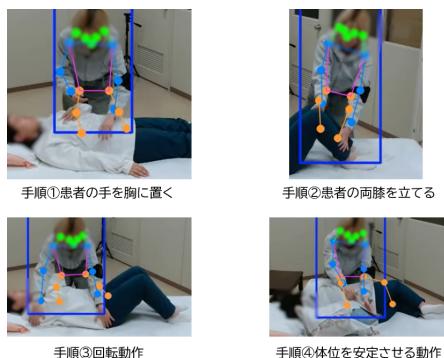


図2: データ収集の様子

## 5. 結果と考察

構築したRCモデルの基本性能および入力データの寄与を検証するため、収集したデータセットを用いて10分割交差検証（10-fold Cross-Validation）を実施した。本検証では、(1)骨格データのみ、(2)音声データのみ、および(3)骨格データと音声データを統合したマルチモーダルデータの3条件について分類精度の計測を行った。各条件で得られた結果を表1に示す。

表1: 交差検証による各条件の分類精度

条件	精度(%)
骨格データのみ	87.67
音声データのみ	98.29
マルチモーダル（骨格+音声）	98.37

単一の入力データを用いた場合、骨格データのみでは87.40%、音声データのみでは98.46%の精度という結果が得られた。これに対して、音声データと骨格データを統合して学習させたマルチモーダルモデルにおいては、精度は98.54%に達した。

これらの結果から、単一の入力データのみを用いるよりも、骨格による映像の情報と音声情報を組み合わせることで、分類精度が向上することが確認された。本研究で構築したRCモデルが、収集したデータにおける動作識別において、マルチモーダル情報の統合により識別性能が向上することが示された。

## 6. おわりに

本研究では、リザバーコンピューティングを用いて、体位変換時の動作分類を実施し、それらが看護において使用可能なのかを検証した。結果として、骨格のみ87.67%や音声のみ98.29%を単独で用いた場合に対して、マルチモーダルモデルにおいて98.37%の分類精度を達成した。これにより、介助動作に伴う音声情報を骨格検出による空間情報が補完するマルチモーダルにより看護師動作の分類ができることが確認された。

今後の展望としては、本手法にMEMS触覚センサから得られる物理的な接触圧力情報を統合する。骨格および音響情報に加えて、動作の質を決定づける情報の付加により、手順間の識別性を高める必要がある。これにより、初学者が安全かつ効率的に体位変換技術を習得できる指導支援の実現を目指す。

## 謝辞

本研究はJSPS科研費JP22H00542, JP23K10021の助成を受けた。

## 参考文献

- 真嶋由貴恵: “看護技術のスキル学習とノウハウ集約における映像活用”, 映像情報メディア学会誌, Vol.66, No.8, pp.645–649 (2012)
- Jaeger, Herbert: “The “echo state” approach to analysing and training recurrent neural networks-with an erratum note”, Bonn, Germany: German National Research Center for Information Technology GMD Technical Report, Vol.148 (2001)