

# 記憶回想時に生成される発話からのユーザモデルパラメータの推定

## Estimation of User Model Parameters from Speech Generated during Memory Recall

酒井翔伎<sup>1\*</sup> 森田純哉<sup>1</sup>  
Shoki Sakai<sup>1</sup> Junya Morita<sup>1</sup>

<sup>1</sup> 静岡大学情報学部

<sup>1</sup> Faculty of Informatics, Shizuoka University

**Abstract:** 記憶の回想を利用したメンタルヘルスケアの手法として回想法がある。本研究では回想法支援システム利用時に生成される発話からユーザの内部状態を表すモデルパラメータを推定することを目標とする。ユーザの発話の内容、韻律に注目し、状況や気分に応じた記憶の遷移をシミュレートする際に必要なパラメータを推定する方法の探索を行った。その結果、韻律情報から気分の主観評定及び写真の印象評価が推定され、モデルパラメータと対応付けられる可能性が示唆された。

## 1 はじめに

### 1.1 背景

近年、サイバーフィジカルシステムやデジタルツインなど物理世界とコンピュータ上の仮想世界を接続することで問題解決を目指す技術が注目されている。これらは Industry 4.0 の重要な技術として産業分野で活用が進んでいる。しかし、人間を対象とした文脈ではこれらの技術の活用はそれほど進んでいない。人間の特徴や機能をコンピュータ上で複製するためには、人間の身体的要素や認知機能などの内部状態について理解する必要がある。特に、人間の内部状態については非常に複雑なため、そのすべてを十分な粒度で把握することは難しい。例えば、人間とロボットのようなモデルがインタラクションを行う場合、モデルは人間の認知機能や内部状態について定量的に観測することが困難である。人間同士のコミュニケーションの場合も相手の反応から互いの意図や感情の推定を行っているが、完璧に相手の状態を把握できているわけではない。人間の内部状態をコンピュータ上で再現することが可能になれば、医学的側面やメンタルヘルスケアなどへの応用が期待される。

応用が期待されるメンタルヘルスケアの手法の一つに回想法がある。回想法は Butler によって提案された手法で、自身に関する過去の写真や音楽などから過去の

の経験や思い出を語り合う心理療法である [1]。回想法には過去を振り返り自身の人生の再評価を狙うライフレビューと、懐かしい思い出の喚起による心理的な安定を狙うレミニッセンスがある。回想法は認知症患者へのメンタルヘルスケアの一環として行われることもあるが、青年期の人々においてもポジティブな感情を促進し、ネガティブな感情を抑制する効果があるとされている [2]。こういったメンタルヘルスケアは、個人の内部状態をモデル化することで、より効果を発揮するようになるだろう。記憶の回想に関与すると考えられる内部変数としては、生体内部の覚醒度や感情価を想定できる [3, 4]。適切な回想の支援を行うためにはこれらの内部変数の状態に応じた回想の促しが必要であると考えられる。本研究では、人間と認知モデルとのインタラクションから記憶回想時の人間の内部状態を再現するインタラクションデザインを提案し、提案デザインの実現性を確かめるため、認知モデルを搭載した回想法支援システムの利用時の発話データを分析する。

### 1.2 目的

本研究では、サイバーフィジカルシステムを人間個人に適用し人間の内部状態をコンピュータ上に再現することを動機とする。これを実現するために、モデルとインタラクションする人間から得られたデータからその個人の内部状態を推定する手法を提案する。この手法の実現性を確認するため、人間から得られるデータとして発話に注目し、データから人間およびモデル

\*連絡先：静岡大学情報学部行動情報学科  
〒432-8011 静岡県浜松市中区城北3丁目5-1  
E-mail: sakai.shoki.18@shizuoka.ac.jp

の内部状態が推定可能かどうかを検討する。外化された発話から人間の内部状態を推定することで、個人のモデルをコンピュータ上に構築する可能性が示される。また、人間の反応によって、インタラクションをするモデルの内部が推定可能であることは、モデルによる働きかけの有効性を示唆することになる。分析では発話の特徴から人間およびモデルの内部状態が推定可能か検討する。

## 2 関連研究

### 2.1 サイバーフィジカルシステムとデジタルツイン

サイバーフィジカルシステムとは、多くの新しいモダリティを通じて人間と対話できる、統合された計算機能と物理機能を備えた新世代のシステムである [5]。サイバーフィジカルシステムでは、センサー等で収集した実世界のデータをサイバー空間で分析する際にデジタルツインという技術が用いられる。デジタルツインは NASA のロードマップに由来する、物理世界を仮想世界上に再現しリアルタイムに同期することで、様々な目的のシミュレーションを行う技術である [6]。これらの技術は Industry 4.0 の重要な役割として議論され [7]、製造業における製造機械などを対象に積極的に活用されている。

デジタルツインを、人間が関与する領域に適用しようとする試みも行われている。例えば、NTT が掲げるデジタルツインコンピューティング構想では、ヒト、特に個人の内面のデジタル表現に挑戦することが大きな特徴として挙げられている<sup>1</sup>。また、人間の内部状態を明らかにする試みとして、脳活動からの人間の価値観の推定・再現や [8]、意思決定能力の推測に関する研究 [9] など、脳をモデル化することで人間の内部状態を明らかにする試みがされている。

上記のような、人間を対象とするデジタルツインの実現には、人間の内部状態の理解が必要である。これまで、心理学や神経科学の研究の成果により、人間の感情や認知に関する様々な原理が明らかになり、計算機を利用した人間の内的過程に関するシミュレーションの技術も多く開発されてきた。しかし、機械によるプロセスに対して、人間の認知プロセスは個人差やそのときどきの体調によって大きく変化する。そのため、特定の個人内部で生じるプロセスのシミュレーションを行うことは、機械内部のシミュレーションや平均的な人間の認知プロセスのシミュレーションを行うのとは異なる性質の困難さがあると考えられる。

### 2.2 認知アーキテクチャと記憶のモデル

先述したとおり、人間の内部状態をコンピュータ上の仮想世界で再現する手法は過去に多く研究されてきた。本研究では、人間個人の認知プロセスをモデル化するために、認知アーキテクチャを用いる。認知アーキテクチャは、人間の一般的な認知プロセスをモデル化するプラットフォームのことである。共通の基盤の上に、多様なタスクや個人に対応した人間の多様な認知機能をシミュレートすることが可能となる。様々な認知アーキテクチャがあるなかで、John. R. Anderson によって開発された認知アーキテクチャ ACT-R は多くの研究で用いられている [10]。ACT-R は心理学実験の積み重ねによって構成されている。また、視覚モジュールや運動モジュールなど様々なモジュールから構成されており、人間の各脳部位と対応付けられたモジュールが情報を伝達し合うことで人間の認知の多様な側面を捉えることができる。これらの特徴から ACT-R は人間の内部状態を再現するための有用な特徴を持つと考えられる。

ACT-R を用いた個人のモデルの構築も行われている。Sommer らは、ACT-R を用いた Cognitive Twin という概念を提唱した。彼らの研究では、個人の意思決定のモデルが作成され、それを利用することで、ACT-R によるパーソナライズされた意思決定支援の有効性が示された [11]。こういった ACT-R による認知機能の個別化は、モデルに搭載される知識（記憶）によって実現される。モデルに搭載する知識、あるいは知識の利用に関わるパラメータに応じて、ACT-R のモデルは同一の課題に対しても多様な振る舞いを生み出す。

ACT-R の知識は宣言的知識と手続き知識から構成される。宣言的知識はチャンクと呼ばれる単位によって構成され、個々のチャンクが結びつくことで、属性を介した意味ネットワークが構成される。このようなネットワークを系列的に探索するためには、手続き知識であるプロダクションルールが必要になる。プロダクションルールは、現在想起されている知識と接続する属性から一つを選択することで、次の知識に遷移する。

上記のような ACT-R における記憶の検索プロセスにバリエーションをもたせるパラメータとして、チャンクに対する活性値、プロダクションルールに対するユーティリティ値が用意されている。活性値は、記憶の利用頻度や記憶の減衰などの影響を受ける。ユーティリティ値は報酬値と学習率の影響を受ける。意味ネットワークからの宣言的記憶の呼び出しにおいては、属性の選択においてプロダクションルールのユーティリティが関与し、選択された属性のなかで複数の候補が合った際には活性値によって検索される知識が定まる。

<sup>1</sup><https://www.rd.ntt/iown/0003.html>

## 2.3 モデルベース回想法

森田らは、上記のような ACT-R の記憶の仕組みを活用した回想法支援のための写真スライドショーシステムを開発した [12]. また、現在の時間ではない時間の経験を、まるでタイムトラベルをする可能ように鮮明に思い浮かべることの意味するメンタルタイムトラベル [13] のモデル化を行い、構築したモデルによるモデルベース回想法を提案した [14]. モデルベース回想法ではユーザはディスプレイ上にスライドショー形式で動的に切り替えられる写真によって回想を行う. このときモデルはユーザの認知的状態、情動的状态を学習し、学習結果に基づく刺激提示を行うと考えられる.

板橋は、森田らのシステムを改良するために、感情モデルに基づくユーザの状態の対話的なモニタリング手法を開発した [15, 16]. 板橋のシステム [15] では、対話的パラメータとして心拍変動の時系列データと、ユーザインタフェースによるフィードバックデータを設定した. ユーザは、図 1 に示すインタフェースによって現在提示されている写真に対する気分状態を 6 段階で評定する. 心拍変動とインタフェースによるフィードバックはそれぞれ、写真の活性値に付与されるノイズ値と、写真の検索ルールの報酬値と対応付けられている. 活性値に基づく記憶の検索が行われることで、覚醒度が高い場合写真の撮影時期に基づいた直近の写真へ検索の集中が起こる. このような写真の検索は心拍間隔の短い緊張状態における記憶の回想と対応する. 対して、覚醒度低い場合は直近の写真だけではなく古い写真への検索も行われる. このような写真の検索は、心拍が穏やかでリラックス状態における記憶の回想と対応する.

板橋はこのようなモデルベース回想法を評価するオンライン実験を実施した. その結果、活性値に基づく写真の検索が、有意な気分評定の変化を導くことを確かめた. しかし、板橋においてはモデルの挙動を変化させるパラメータのうち、活性値については固定され、報酬値についてもその効果を十分に検討できていなかった. 個人の回想を適切なタイミングで支援するためには、個人の反応から個人の内部状態を推測しつつ、リアルタイムに適切なパラメータに設定していく必要がある.

## 3 ユーザモデルパラメータの対話的推定

本研究では、モデルと人間のインタラクションから人間の内部状態を推定する手法として、モデルが人間に与える刺激に対する反応を基に対話的な推定を行う

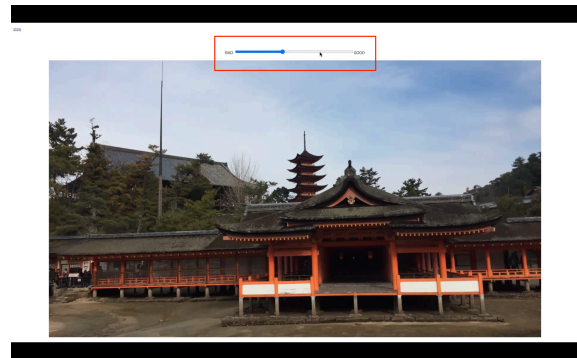


図 1: 気分状態評定インタフェース

ことを提案する. 提案するモデルと人間の関係図を図 2 に示す.

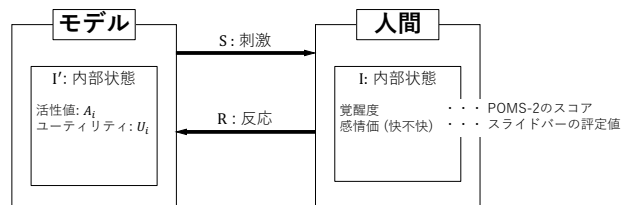


図 2: モデルと人間の関係図

モデルは人間に刺激を与え ( $S$ ), それに対する人間の反応を受け取る ( $R$ ). 人間の反応からモデルは自身の内部状態 ( $I'$ ) を更新する. 更新された  $I'$  を基に次の刺激を決定する. このやり取りを繰り返すことで  $I'$  は人間の内部状態 ( $I$ ) を反映するものになっていくと予想する.

この枠組を、前節で述べたモデルベース回想法に適用した場合、モデルの与える刺激 ( $S$ ) は提示写真、人間の反応 ( $R$ ) は発話に相当する. 人間の内部状態 ( $I$ ) は直接観測できないものの、参加者自身の気分の状態の評定が内省報告として得られる. その一方で、モデルの内部状態 ( $I'$ ) はシステム設計者にとって把握可能なパラメータとなる. モデルベース回想法における具体的なパラメータは、活性値  $A_i$  とユーティリティ  $U_i$  であり、それぞれ次のように定められる.

記憶を検索するプロダクションルール  $i$  は時間、場所、写っている人、写っている物の 4 つの属性を持つ. これらの属性をそれぞれ、撮影時期  $t$ , 撮影場所  $l$ , 写っている人物  $p$ , 写っている物  $o$  とするとユーティリティ  $U_i$  は  $i \in \{t, l, p, o\}$  と表すことができる.  $U_i^n$  は前回の

ユーティリティ値  $U_i^{n-1}$  と報酬値  $R_i^n$ , 学習率  $\alpha$  から,

$$U_i^n = U_i^{n-1} + \alpha[R_i^n - U_i^{n-1}] \mid i \in \{t, l, p, o\} \quad (1)$$

のように計算される.

活性値  $A_i$  はベースレベル活性値  $B_i$ , 活性化拡散  $S_i$ , ノイズ  $\epsilon_i$  の合計値として

$$A_i = B_i + S_i + \epsilon_i \quad (2)$$

のように計算される. 本研究では, ベースレベル活性値の計算に式 3 を用いる.

$$B_i = \ln\left(\sum_{j=1}^n t_j^{-d}\right) + \beta_i \quad (3)$$

$n$  はチャンク  $i$  の出現回数,  $t_j$  は  $j$  番目の出現からの経過時間,  $d$  は減衰率,  $\beta_i$  はオフセット値を示す.

式 4 の第二項である活性化拡散は現在の刺激  $S$  のコンテキスト  $C$  に対するチャンク  $i$  の連想強度  $S_i$  として計算される.

$$S_i = \sum_{j \in C} W_j S_{ji} \quad (4)$$

$C$  は, 現在の刺激  $S$  の属性値  $j$  の集合を表す.  $W_j$  は属性値  $j$  に付与される注意の重みを表し,  $S_{ji}$  は属性値  $j$  と宣言的記憶のチャンク  $i$  の連想強度を表す.

これらのモデルの内部状態  $I'$  に応じて, 実験参加者に提示される刺激  $S$  が定まり, 間接的に人間の状態  $I$  に影響が及ぶ. 本研究はそのような関係から, 人間から観測される反応  $R$  による  $I'$  の推定が, 人間の内部状態  $I$  の推定の代替となると考える.

上記の考えから, 本研究では人間の反応  $R$  から人間とモデルの 2 つの内部状態の推定を試みる. 本研究のリサーチクエスチョンは,

1. 刺激をうけた人間の反応からモデルの内部状態を推定すること ( $R \rightarrow I'$ ) は可能かどうか
2. 人間の反応から人間の (内省報告により得られた) 内部状態を推定すること ( $R \rightarrow I$ ) は可能かどうか

の 2 点であり, これを先行研究において得られた発話データから検討する.

## 4 方法

### 4.1 データセット

分析に使用するデータセットは板橋の行った実験の録画データをもとに作成した [16]. 実験ではモデルの

内部状態による人間の内部状態の変化を検討するために, 二つのパラメータを操作した, 2 (活性値算出無 vs. 活性値算出有)  $\times$  2 (報酬値無 vs. 報酬値有) の 4 条件下で参加者の行動を比較した. 実験参加者は 21 - 61 歳の 24 名で, 男性 12 名, 女性 11 名であった. これらの参加者は, クラウドソーシングサイトランサーズ<sup>2</sup>にて募集された. 実験はビデオチャットツール Zoom<sup>3</sup>で実施された. 各条件の閲覧時間は 5 分であった. 写真スライドショー閲覧前, 各条件の閲覧終了後に写真スライドショーの閲覧による気分状態の変化を評価するために, Profile of Mood States 2nd Edition (POMS-2) の日本語版 [17] による気分状態の調査を写真閲覧前, 各条件の閲覧後に実施した. また, 参加者は写真閲覧中に思考の発話をする事と, 図 1 の赤枠で囲んだ位置に表示されるユーザインタフェースを使用して提示された写真に対する現在の気分を 0, 2, 4, 6, 8, 10 の 6 段階で評定するタスクが付与された.

### 4.2 データの前処理

前節の実験の録画データを基に分析で使用するデータを作成した. 分析には発話のテキストデータと音声データを用いた. テキストデータは録画データを音声認識による自動文字起こしが可能なツール Vrew<sup>4</sup>を用いて作成した. 自動文字起こしを行った後, 誤字の修正, 実験の開始・終了及び写真の切り替わりのタイミングを示すタグの挿入を手作業によって行った. 音声データは録画データから音声部分を抽出することで作成した. 分析のため音声データは提示写真が切り替わる時間の平均値である 11 秒で分割した.

### 4.3 分析手法

発話から機械的に得られる特徴からモデルおよび人間の内部状態 (図 2 の  $S \rightarrow I'$ ) の推定 (図 2 の  $S \rightarrow I$ ) が可能かどうかを検討する. 音声データから特徴ベクトルを構成し, SVM による分類を行う.

#### 4.3.1 特徴量

分割された音声ファイルに openSMILE を適用することで韻律特徴を抽出した [18]. openSMILE は音声信号から特徴量を自動抽出することができるソフトウェアで, 主に自動感情認識の分野で活用されている. 使用する特徴量は openSMILE で提供されている特徴量セットのうち, eGeMAPS を用いるものである [19].

<sup>2</sup><https://www.lancers.jp>

<sup>3</sup><https://zoom.us/>

<sup>4</sup><https://vrew.voyagerx.com/ja/>

eGeMAPS では、62 種類のミニマリストティックパラメータセットと 26 種類の拡張パラメータセットの合計 88 種類の特徴量が提供される。これらに、音声に対応する発話内容の感情分析の結果の数値を加えた 89 次元の特徴ベクトルを分析に用いた。発話内容の感情分析には Google の Cloud Natural Language API<sup>5</sup>の感情分析機能を用いた。また、特徴ベクトルは事前に標準化を行った。

### 4.3.2 SVM による分類

実験条件 4 条件の分類、POMS-2 の TMD 得点が一つ前の閲覧条件から上がったか下がったかの二値分類、スライドバーのスコアが一つ前の提示写真に対するスコアから上がったか下がったか、変化しなかったかの 3 クラス分類を行う。SVM のパラメータの最適化のため、表 1 に示すカーネルとパラメータの範囲でグリッドサーチを行った。評価は 5-fold のクロスバリデーションを行った際の正解率 (Accuracy) と F 値 (F-measure) を用いた。ただし、正解率と F 値は表 2 に示す取りうる分類結果と (5) - (8) に示す式から計算されるものとする。

表 1: グリッドサーチに用いたカーネルとパラメータ

カーネル	C	ガンマ
線形	0.001, 0.01, 0.1, 1, 10, 100	
RBF	0.001, 0.01, 0.1, 1, 10, 100	0.0001, 0.001, 0.01, 0.1, 1, 10

表 2: 取りうる分類結果

		真の値	
		0	1
分類結果	0	TP (True Positive)	FP (False Positive)
	1	FN (False Negative)	TF (True Negative)

$$\text{正解率 (Accuracy)} = \frac{TP + TN}{TP + FP + FN + TN} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{F 値 (F-measure)} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

<sup>5</sup><https://cloud.google.com/natural language>

## 5 分析結果

### 5.1 実験条件の分類

はじめに、(活性値算出無 vs. 活性値算出有) × (報酬値無 vs. 報酬値有) の 4 条件の分類を行った。実験条件の分類の正解率は 0.485, F 値は 0.480 であった。正解率, F 値ともに高い値とはならなかったが、混合行列を見ると SVM による判断の結果の内、4 条件全てで真の値に分類されている割合が一番高くなっているため、分類の傾向が見られることがわかる。分類結果の混合行列のヒートマップを図 3 に示す。また、本章の残りの部分で示すヒートマップは共通して横軸が真の値、縦軸が SVM によって分類された結果の値を示すこととする。

次に、4 条件では複数の要素が関係し合うため、活性値の有無、報酬値の有無それぞれに注目して二値分類を行った。その結果、活性値の分類の正解率は 0.615, F 値は 0.667, 報酬値の分類の正解率は 0.657, F 値は 0.651 であった。分類結果の混合行列のヒートマップをそれぞれ図 4a, 4b に示す。これらの結果から、発話から写真閲覧時の活性値の高低、報酬値の有無による写真提示の違いに対する反応の違いを検出することの可能性が示唆される。

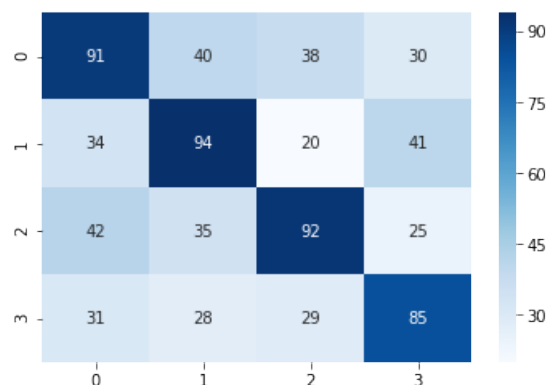


図 3: 実験条件の分類の混合行列のヒートマップ: 0 は活性値無・報酬値無条件, 1 は活性値無・報酬値有条件, 2 は活性値有・報酬値無条件, 3 は活性値有・報酬値有条件を示す。

韻律特徴には個人差が含まれる可能性を考慮し、個人ごとの活性値と報酬値の分類を行った。個人ごとの正解率と平均の正解率を表 3 に示す。24 名の平均の正解率を見ると、全体で分類を行った場合より正解率が向上していることが分かる。このことから個人ごとに分類を行うほうが、複数人の特徴を合わせて分類を行うよりも分類性能が向上することが示唆される。個人ごとの結果を見ると、高い参加者では 8 割を超える正解率が得られた。反対に、低い参加者では 5 割近い数

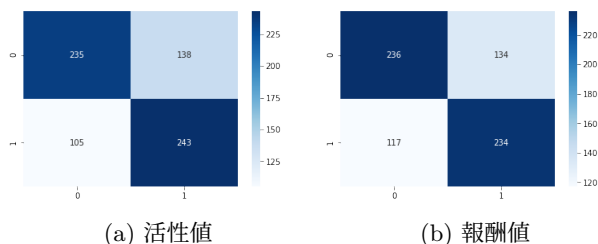


図 4: 実験条件の分類の混合行列のヒートマップ: 0 は無し条件, 1 は有り条件を示す。

値となりほぼランダムに近い分類となった。また, ID 7, 19, 23 の参加者のように活性値と報酬値で正解率に大きな差が見られた場合もあった。これらの結果は, 人間の内部状態を推定する際に, 個人差が大きく影響することを改めて示すものである。

全体での分類と個人ごとの分類の結果, 活性値の分類の可能性が見られたことから人間の韻律から覚醒度を推定できる可能性が示唆される。また, 報酬値の有無による韻律の差の分類からは人間の感情に対応した刺激を与えることで韻律に特徴が生まれる可能性が考えられる。

## 5.2 POMS-2 の結果の変化の分類

POMS-2 は標準的に利用されている気分の尺度であり, 総合的な気分状態が TMD (total mood disturbance) 得点として測定される。写真閲覧後に回答した POMS-2 の TMD 得点が閲覧前, もしくは一つ前の条件の閲覧後と比べて上がっているか下がっているかを分類した。ただし, 2 名は 4 回の閲覧で下がり続けたため除外した。TMD 得点の変化の分類の正解率は 0.711, F 値は 0.703 であった。分類結果の混合行列のヒートマップを図 5 に示す。活性値の有無, 報酬値の有無の分類の場合よりも正解率, F 値が共に向上し, 混合行列を見てもより良い分類ができていることが示されている。ここから発話から参加者の気分変動の推定が可能なが示唆される。

個人ごとの正解率と平均の正解率を表 3 に示す。個人の分類の平均を見ると 0.733 と, 全体での分類からやはり向上している。高い参加者では 0.986 とほぼ正確に分類ができている。また, 一番低い参加者でも 0.594 と活性値, 報酬値の場合よりも向上している。しかし, 高い参加者と低い参加者の差は大きいことや個人差が大きいことが示唆される。

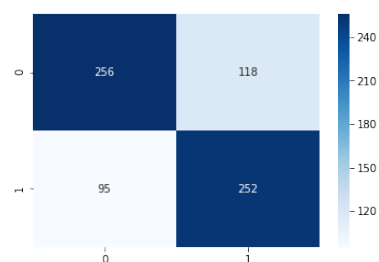


図 5: POMS-2 の混合行列のヒートマップ: 0 は上昇, 1 は下降を示す。

## 5.3 スライドバーの値の変化の分類

スライドバーの評定値が一つ前の提示写真の評定値から上がったか下がったか, 変化がなかったかを分類した。スライドバーの変化の分類の正解率は 0.480, F 値は 0.430 であった。分類結果の混合行列のヒートマップを図 6 に示す。混合行列を見ると, 変化しないことを示す 0 は比較的分類ができていることが示されているが, 上昇を示す 1 と下降を示す 2 はほぼ分類できていないことが分かる。これは, 参加者によってスライドバーをほとんど操作していない場合があったためと考えられる。

個人ごとの正解率と平均の正解率を表 3 に示す。個人ごとの平均の正解率は全体での分類から向上している。個人ごとの正解率を見ても 8 割を超える参加者も見られる。しかし, これらの数値は前述の通りスライドバーの操作をほとんどしていない参加者も含まれており信頼できる数値ではない。そのため, 板橋の実験で得られたデータについては, スライドバーによる分類から人間の内部状態を推定することは難しいといえる。

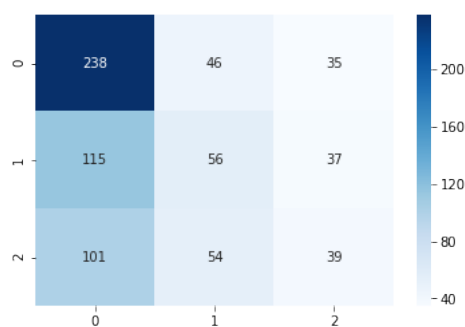


図 6: スライドバーの混合行列のヒートマップ: 0 は変化なし, 1 は上昇, 2 は下降を示す。

## 6 おわりに

本研究では、サイバーフィジカルシステムやデジタルツイン技術を人間に適用することを目標とし、人間とモデルのインタラクションから人間の内部状態を再現する手法を提案した。

提案手法の実現性を確認するため、モデルベース回想法利用時のユーザの発話を用いた内部状態の分類を行った。この分析では、モデルの内部状態（実験条件）、および人間の内部状態（POMS-2、スライドバー）の分類を試みた。前者の結果から、高い性能は得られなかったものの発話からの分類の可能性が示唆された。この結果から図2の  $R \rightarrow I'$  の推定の可能性が示された。後者のうち、POMS-2の分類の結果は、実験条件の分類よりも性能が向上した。このことから、発話の韻律特徴からユーザの気分状態の変化を推定可能であることが示唆された。この結果は、図2の  $R \rightarrow I$  の推定の可能性を示すものである。ただし、人間の内部状態を表す別の指標であるスライドバーの値の分類においては、十分な性能は得られなかった。

これらの結果から、本研究の2つのリサーチクエスチョンに対し、限定付きではあるものの可能であるとの回答が得られた。ただし、この限定性については十分に注意する必要がある。モデルの内部状態を推定することで、人間の内部状態を推測することが可能であるという前提は、その妥当性自体を改めて検討する必要がある。また、人間の内部状態を目的変数とした分類も、間接的な指標を用いたものであることに注意する必要がある。スライドバーを目的変数とした分類における困難さは、このような内部状態の間接的な測定の問題を示している。

これらを総合し、本研究では、人間とモデルのインタラクションからモデルが人間の内部状態を対話的に推定することの実現性と課題を確認したといえる。本研究で提案した手法により人間個人の内部状態を再現することが可能になれば、人間を対象としたデジタルツイン、サイバーフィジカルシステムの発展が期待される。

今後は、上記の知見を踏まえつつ、内部状態を動的に推定しつつ、ユーザを誘導するインタラクティブシステムを開発していく必要がある。その試みを行う中で、先述した限界を克服する手段が開発されていくと考える。人間の内部状態を観測する手段に関して、事後での計測は、状況に応じた内部状態の変化に対応できず、リアルタイムでの評価はユーザの負荷の点で問題が生じる。これらの限界を克服するために、より簡便かつリアルタイムにユーザの反応を計測する手段の開発が必要である。

表 3: SVM による個人ごとの分類の正解率

ID	活性値	報酬値	POMS-2	スライドバー
1	0.667	0.564	0.632	0.826
2	0.786	0.629	0.739	0.700
3	0.657	0.686	0.657	0.457
4	0.583	0.624		0.500
5	0.670	0.793	0.986	0.611
6	0.603	0.707	0.642	0.412
7	0.735	0.542	0.611	0.444
8	0.714	0.871	0.686	0.614
9	0.651	0.724	0.648	0.667
10	0.640	0.553	0.716	0.493
11	0.643	0.686	0.638	0.671
12	0.700	0.657	0.786	0.819
13	0.757	0.600	0.757	0.400
14	0.688	0.751	0.742	0.448
15	0.805	0.738	0.594	0.431
16	0.643	0.671	0.768	0.443
17	0.605	0.761	0.714	0.394
18	0.629	0.743	0.786	0.457
19	0.707	0.564	0.686	0.408
20	0.785	0.697	0.853	0.478
21	0.721	0.746	0.957	0.563
22	0.714	0.686	0.768	0.529
23	0.529	0.729		0.443
24	0.754	0.600	0.769	0.646
Mean	0.683	0.680	0.733	0.536

## 参考文献

- [1] Robert N Butler. The life review: An interpretation of reminiscence in the aged. *Psychiatry*, Vol. 26, No. 1, pp. 65–76, 1963.
- [2] 福島脩美, 田中勝博, 角山富雄, 張替裕子, 松田修, 森美保子, 豊嶋舞子ほか. 過去と最近の出来事の回想におけるツールとしての書記, 描画, 対話の感情効果. 目白大学心理学研究, Vol. 4, pp. 1–10, 2008.
- [3] Margaret M Bradley, Mark K Greenwald, Margaret C Petry, and Peter J Lang. Remembering pictures: pleasure and arousal in memory. *Journal of experimental psychology: Learning, Memory, and Cognition*, Vol. 18, No. 2, p. 379, 1992.
- [4] 野畑友恵, 越智啓太. 記憶に及ぼす覚醒度の効果は快・不快感情によって異なる: 覚醒度説への反証. 認知心理学研究, Vol. 3, No. 1, pp. 23–32, 2005.
- [5] Radhakisan Baheti and Helen Gill. Cyber-physical systems. *The impact of control technology*, Vol. 12, No. 1, pp. 161–166, 2011.
- [6] Elisa Negri, Luca Fumagalli, and Marco Macchi. A review of the roles of digital twin in cps-based production systems. *Value Based and Intelligent Asset Management*, pp. 291–307, 2020.
- [7] Jay Lee, Behrad Bagheri, and Hung-An Kao. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing letters*, Vol. 3, pp. 18–23, 2015.
- [8] Marcus Grueschow, Rafael Polania, Todd A Hare, and Christian C Ruff. Automatic versus choice-dependent value representations in the human brain. *Neuron*, Vol. 85, No. 4, pp. 874–885, 2015.
- [9] Tanveer Talukdar, Francisco J Román, Joachim T Operskalski, Christopher E Zwilling, and Aron K Barbey. Individual differences in decision making competence revealed by multi-variate f mri. *Human brain mapping*, Vol. 39, No. 6, pp. 2664–2672, 2018.
- [10] John R Anderson. *How can the human mind occur in the physical universe?* Oxford University Press, 2009.
- [11] Sterling Somers, Alessandro Oltramari, and Christian Lebiere. Cognitive twin: A personal assistant embedded in a cognitive architecture.
- [12] 森田純哉, 平山高嗣, 間瀬健二, 山田和範. 認知アーキテクチャを組み入れた写真スライドショーの開発: 展望と課題. 人工知能学会全国大会論文集 第 29 回全国大会 (2015), pp. 2M4NFC04b3–2M4NFC04b3. 一般社団法人 人工知能学会, 2015.
- [13] Daniel L Schacter, Donna Rose Addis, and Randy L Buckner. Remembering the past to imagine the future: the prospective brain. *Nature reviews neuroscience*, Vol. 8, No. 9, pp. 657–661, 2007.
- [14] 森田純哉, 平山高嗣, 間瀬健二, 山田和範ほか. メンタルタイムトラベルを誘導するモデルベース回想法. 研究報告ユビキタスコンピューティングシステム (UBI), Vol. 2015, No. 16, pp. 1–6, 2015.
- [15] Kazuki Itabashi, Junya Morita, Takatsugu Hirayama, Kenji Mase, and Kazunori Yamada. Interactive model-based reminiscence using a cognitive model and physiological indices. In *Proceedings of the 18th International Conference on Cognitive Modelling*, No. 2020, 2020.
- [16] 板橋和希. インタラクティブなモデルベース回想法の開発と評価. Master’s thesis, 静岡大学大学院総合科学技術研究科情報学専攻, 2021.
- [17] K Yokoyama, S Araki, N Kawakami, and T Tkakeshita. Production of the japanese edition of profile of mood states (poms): assessment of reliability and validity. [*Nihon koshu eisei zasshi*] *Japanese journal of public health*, Vol. 37, No. 11, pp. 913–918, 1990.
- [18] Florian Eyben, Martin Wöllmer, and Björn Schuller. Opensmile: the munich versatile and fast open-source audio feature extractor. In *Proceedings of the 18th ACM international conference on Multimedia*, pp. 1459–1462, 2010.
- [19] Florian Eyben, Klaus R Scherer, Björn W Schuller, Johan Sundberg, Elisabeth André, Carlos Busso, Laurence Y Devillers, Julien Epps, Petri Laukka, Shrikanth S Narayanan, et al. The geneva minimalistic acoustic parameter set (gemaps) for voice research and affective computing. *IEEE transactions on affective computing*, Vol. 7, No. 2, pp. 190–202, 2015.