

移動ロボットの為の意識モデルと顔記憶機能に  
基づく CPU・バッテリーの有効利用と親近感の向上

筑波大学大学院  
図書館情報メディア研究科  
2019年3月  
陳 浩林

# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	研究背景 . . . . .	1
1.2	関連研究 . . . . .	1
1.2.1	CPU とバッテリーの有効利用に関する研究 . . . . .	1
1.2.2	ロボットの親近感の向上に関する研究 . . . . .	2
1.3	研究の目的と提案 . . . . .	2
1.4	論文の構成 . . . . .	5
<b>第2章</b>	<b>意識モデル</b>	<b>6</b>
2.1	Hobson の AIM モデル . . . . .	6
2.2	数理 AIM モデル . . . . .	7
2.3	数理 AIM モデルに基づくプロセス制御 . . . . .	12
<b>第3章</b>	<b>顔記憶システム</b>	<b>13</b>
3.1	システム構成 . . . . .	13
3.2	顔検出 . . . . .	15
3.3	顔認識 . . . . .	16
3.4	画像前処理と分類器訓練 . . . . .	17
3.5	音声認識と音声合成 . . . . .	18
<b>第4章</b>	<b>実験と考察</b>	<b>19</b>
4.1	実験概要 . . . . .	19
4.2	実験1 . . . . .	20
4.2.1	概要 . . . . .	20
4.2.2	システム構成 . . . . .	20
4.2.3	ソフトウェア構成 . . . . .	22
4.2.4	ハードウェア構成 . . . . .	22
4.2.5	実験結果と考察 (Part1) . . . . .	24
4.2.6	実験結果と考察 (Part2) . . . . .	25
4.3	実験2 . . . . .	26
4.3.1	概要 . . . . .	26
4.3.2	システム構成 . . . . .	26
4.3.3	ソフトウェア構成 . . . . .	27
4.3.4	ハードウェア構成 . . . . .	27
4.3.5	WoZ 法による実験 . . . . .	30
4.3.6	ババ抜きゲーム . . . . .	30
4.3.7	ロボットの発話 . . . . .	31
4.3.8	実験環境 . . . . .	31
4.3.9	評価方法 . . . . .	35

4.3.10	実験手順 . . . . .	37
4.3.11	実験結果 . . . . .	39
4.3.12	考察 . . . . .	40
4.4	実験3 . . . . .	41
4.4.1	概要 . . . . .	41
4.4.2	実験結果 . . . . .	41
4.4.3	考察 . . . . .	42
<b>第5章</b>	<b>結論</b>	<b>43</b>
	参考文献	46
付録A	ロボットの台詞	49
付録B	アンケート	52

# 目次

1.1	提案手法の概要 . . . . .	4
2.1	Hobson が提案した AIM モデル . . . . .	6
2.2	$a_{ex}(t)$ 、 $i_{ex}(t)$ 、 $m_{ex}(t)$ 、 $a_{in}(t)$ 、 $i_{in}(t)$ 、 $m_{in}(t)$ の値の時間変化の例 . . . . .	8
2.3	要素 $A$ の値の時間変化の例 . . . . .	10
2.4	要素 $I$ 、 $M$ の値の時間変化の例 . . . . .	10
2.5	数理 AIM モデルの GUI . . . . .	11
2.6	数理 AIM モデルに基づく意識状態の遷移ルート . . . . .	12
3.1	顔記憶システムの処理の流れ . . . . .	13
3.2	MTCNN による顔検出の実例 . . . . .	15
3.3	FaceNet モデルの訓練方式 . . . . .	16
3.4	画像前処理過程と分類器訓練過程 . . . . .	17
3.5	リアルタイム顔認識を行う様子 . . . . .	18
4.1	システム構成図 (実験 1) . . . . .	20
4.2	ロボットの外観 (実験 1、Part2) . . . . .	23
4.3	ハードウェア間の接続図 (実験 1、Part2) . . . . .	23
4.4	CPU の使用率の変化図 . . . . .	24
4.5	ロボットのバッテリーの電圧の変化図 . . . . .	25
4.6	システム構成図 (実験 2) . . . . .	26
4.7	ロボットの外観 (正面、実験 2) . . . . .	29
4.8	ロボットの外観 (側面、実験 2) . . . . .	29
4.9	ロボットの顔インタフェース . . . . .	29
4.10	実験室の見取り図 . . . . .	32
4.11	待機スペースの風景 . . . . .	33
4.12	実験スペースの風景 . . . . .	33
4.13	隠しカメラを収めた紙箱の正面 . . . . .	34
4.14	隠しカメラを収めた紙箱の内部 . . . . .	34
4.15	操縦者側のタブレット PC に表示される実験スペースの様子 . . . . .	34
4.16	各尺度の平均得点のグラフ . . . . .	39

# 表目次

2.1	要素 $A$ の値、要素 $M$ の値と意識状態の関係	12
4.1	使用ライブラリ (実験 1)	22
4.2	移動ロボットに搭載したノートパソコン	22
4.3	実験機器 (実験 1)	22
4.4	使用ライブラリ (実験 2)	27
4.5	操縦者側のノートパソコン	27
4.6	ロボット側のノートパソコン	28
4.7	実験機器 (実験 2)	28
4.8	Godspeed アンケートシリーズ	36
4.9	各尺度の統計データ	39
4.10	顔検出と顔認識の実行速度	41
4.11	7 人分の顔画像を用いた処理時間と分類器の正解率	41
4.12	35 人分の顔画像を用いた処理時間と分類器の正解率	42

# 第1章 序論

## 1.1 研究背景

近年では、ロボットに関する技術の継続的な発展に伴い、ロボットは徐々に人々の日常生活に浸透していき、我々人間にサービスと便利さの提供に貢献している。同時に、人間と密接に接触するこれらのロボットは、ユーザーエクスペリエンスが特に重視されており、それを改善し向上させる必要がある。現在、我々は積雪地域での地域の活性化を促進する為に、北海道最北端の稚内市で住民や観光客に地域情報を提供する移動ロボットの開発を進めており、その過程の中でいくつかの問題に直面している。

ロボットが人とインタラクションをする際は、映像や音声などのさまざまな知覚情報の並列処理が必要である。しかし、ロボットに数多くの情報処理をさせると、ロボットに搭載されている計算機のCPUを容易に使いきり、リアルタイムでの情報処理が困難となる。一方で、ロボットの周囲に人がいない場合など、知覚情報処理が必要でない場合でも同じ頻度で情報を処理し続けると、CPUとバッテリーを無駄に消費する問題が考えられる。特に移動ロボットの場合は、必要かどうかに関わらず同じ頻度で知覚情報処理を続けると、バッテリー容量の制限により稼働時間の短縮も懸念される。これは1つ目の問題点である。ロボットをより賢く柔軟に状況に応じて必要性の低い処理を省かせ、同時に別の適切な処理をさせることができると問題が改善されると考えられる。

実際に商店街などで移動ロボットの走行実験などを実施すると、人々はロボットに興味を示すが、一部の人は距離を保ってロボットに接近しないことが見受けられた。ロボットは機械的な構造と外観を持ち、かつ多くの人々はロボットと対面する経験を持っておらず、人とロボットが共生するような環境に慣れていない為、これらの人々にとってロボットは親しみやすい存在ではないものだと思われる。これは2つ目の問題点である。そこで、もしロボットが人々の注意を引くことができ、かつ親近感のある印象を人々にもたらしことができるようになると、より多くの人々がロボットに近づき、ロボットと人間のインタラクションがよりスムーズになると考えられる為、ロボットの親近感を向上させる必要がある。

## 1.2 関連研究

### 1.2.1 CPUとバッテリーの有効利用に関する研究

CPUの有効利用問題はスケジューリング関連の技術を使うことで解決する。代表的な手法としては、定時割り当、頻度に基づく優先度スケジューリング、階層スケジューリング、スケジューラといった手法が挙げられる [1]。しかし、これらの手法は事前に必要な計算資源の量を把握する必要があり、またはOSの対応上の問題などの制限が挙げられる。その他に、スケジューリング技術の実装と調整は時間と専門知識を要求する為、専門分野外の人にとっては手軽に採用できる手法だと考えにくい。

移動ロボットのバッテリーを節約するには、先行研究ではモータの制御や経路の計画を効率化する手法が研究されている。Meiらは、多項式でモータの回転速度と電力消費の関係

をモデリングし、線形変換を用いてロボットの速度と車輪の速度の関係を明確にすることにより、異なる目標地点に対する最適な路線と速度を計算する手法を提案した [2]。しかし、この手法はロボットの移動方式に着目している為、ロボットの移動に重点を置かない本研究とは方向性が異なる。

Mikawa らはロボットの知覚情報の並列処理の為に、ロボットシステムに適用できる数理 AIM(Activation-Input-Modulataion) モデルを提案した [3]。この数理 AIM モデルは、人間の意識状態を模倣することができ、これをロボットに適用することで、ロボットは外部環境の変化に応じて「覚醒」と「睡眠」といった状態を切り替えることができるようになる。ロボットの周囲に人がいる時は、ロボットは常に覚醒状態を保ち、必要に応じて知覚情報処理を行う。逆に、周囲に人がいない時に、ロボットは睡眠状態に遷移し、不要な情報処理の一部を削減することで CPU とバッテリーの使用を節約する。また、睡眠状態の時はリアルタイムでは処理しきれない重い処理を実行するができ、CPU を柔軟にかつ有効に利用することができる。CPU の有効利用に伴い、バッテリーの有効利用も期待できる。

### 1.2.2 ロボットの親近感の向上に関する研究

人のロボットへの印象を調査する為に、様々な研究が行われてきた。ロボットの外観が人に与える影響に関する研究 [4] やロボットの感情表現がどのように人を影響するのかについての研究 [5]、またはロボットの視線 [6]、ロボットの頭部の動き [7] が人にもたらす影響についての研究が挙げられる。Huang らの研究では、ロボットが人に接する際の反応時間、接近速度、個体間距離、注意深さを調整することで異なる行動パターンを作り出し、人にとって親近感の高い行動パターンを発見した [8]。Schulte らの研究では、ロボットの声と表情で人に親近感を表した [9]。これらの研究における人のロボットに対する印象評価において、親近感に関する項目で高得点を得られた方法を採用すれば、ロボットの親近感を改善することができると考えられる。しかし、これらの手法を実現するには可動部品の追加や外観の変更が必要である為、実装の難易度が高いと思われる。

Torta らが行った実験から、ロボットの音を伴う行動は速く人間の注意を引くことができ、より容易に人間に気づかれることが分かった [10]。これは人の注意を引く手法のヒントとして使える。また、神田と光永らの研究では、小学校 [11] [12]、オフィス [13]、ショッピングモール [14] などのいくつかの異なる環境で人間とロボットのインタラクション実験を行った。実験に使用されたロボットは、一人ひとりを独立した個体として認識し、人の名前を呼ぶことができる。上記の実験から、人の名前を呼ぶことは、人の注意を引きつけることができ、ロボットと人間の間でのインタラクションの円滑な開始を助け、人間により親しみやすい印象を与えることに貢献できるとの知見を得た。ロボットに人の名前を呼ぶ機能を追加するには、可動部品や外観の変更を行わずに実現できる為、本研究に適用可能な手法だと考えられる。しかし、神田と光永らの実験では、個人の特定は実験参加者に無線タグを付けることで実現した。実験参加者に事前に特定の機器を装備する必要がある為、制限が厳しい方法だと考えられる。また、これらの研究で得られたロボットが人の名前を呼ぶことは親近感の向上に役立つとの知見は、比較実験から得られたものではない為、より厳密的に比較実験を通して証明する必要がある。

## 1.3 研究の目的と提案

本研究は、移動ロボットの CPU とバッテリーの有効利用、およびロボットに対する親近感の向上を目的とする。関連研究に対する調査を踏まえ、本研究は数理 AIM モデルと顔記憶

機能を組み合わせた手法を提案する。

数理 AIM モデルは、CPU とバッテリーの有効利用の為にロボットに適用する。数理 AIM モデルはロボットの各プロセスの稼働状況を制御することができる。ロボットは周囲に人がいない時に睡眠状態に遷移し、その場合での必要性が低い処理の処理頻度を低下させることで、CPU とバッテリーの無駄な消費を省く。また、睡眠時の空き時間に有意義な処理をさせることで、CPU とバッテリーをさらに有効に利用する。

ここでの「有意義な処理」は、ロボットの親近感の向上に繋がる処理を行うものとする。本研究では、顔認識技術を用いて人を特定し、名前呼びかけることで親近感の向上を図る。顔認識は、機械学習・深層学習の手法で実現する。また、高精度の顔認識を実現するには画像の前処理および分類器の訓練が必要となる。CPU とバッテリーの有効利用の為に、これらの計算時間にかかる負荷の高い処理はロボットが休眠状態になる時に行わせる。顔記憶を可能にするシステム全体は、本論文では顔記憶システムと呼ぶ。

本提案手法の概要を図 1.1 に示す。提案手法は以下のようなステップで実行する。

まず、図 1.1 の A のように、移動ロボットがユーザと対面している時に、ユーザの許可を得てからカメラを用いて顔画像を収集する。ロボットの音声は Text-to-Speech で合成し、また音声認識で名前情報を取得する。

次に、図 1.1 の B のように、カメラを用いて周辺状況を把握し、顔検出でロボットの周辺に歩行者がいるかどうかを判断する。周辺が無人と判断する時は、数理 AIM モデルに基づきシステムを休眠状態に遷移させて、不要な処理の処理頻度を下げることで無駄な消費を省く。また、睡眠時は収集した顔画像の前処理と顔の分類器の訓練を行う。

顔画像の前処理と分類器の訓練は図 1.1 の C に示すように行われる。取得した画像を顔検出を用いて、顔の部分だけを切り取り、切り取った顔画像を、顔の持ち主の名前と組み合わせ、訓練用のデータセットを作成する。その後、データセットを用いて分類器の訓練を行う。

最後、図 1.1 の D のように、ロボットが覚醒している時に、カメラを用いてすれ違う人々の顔を検出し、それを顔認識プロセスへ渡して顔画像の認識を行う。ロボットが街で走行している間に、もし既に覚えた人の顔を検出できれば、ロボット側から自主的に人に挨拶し、名前を呼ぶことにより、人間とロボット間の親近感が高められる。



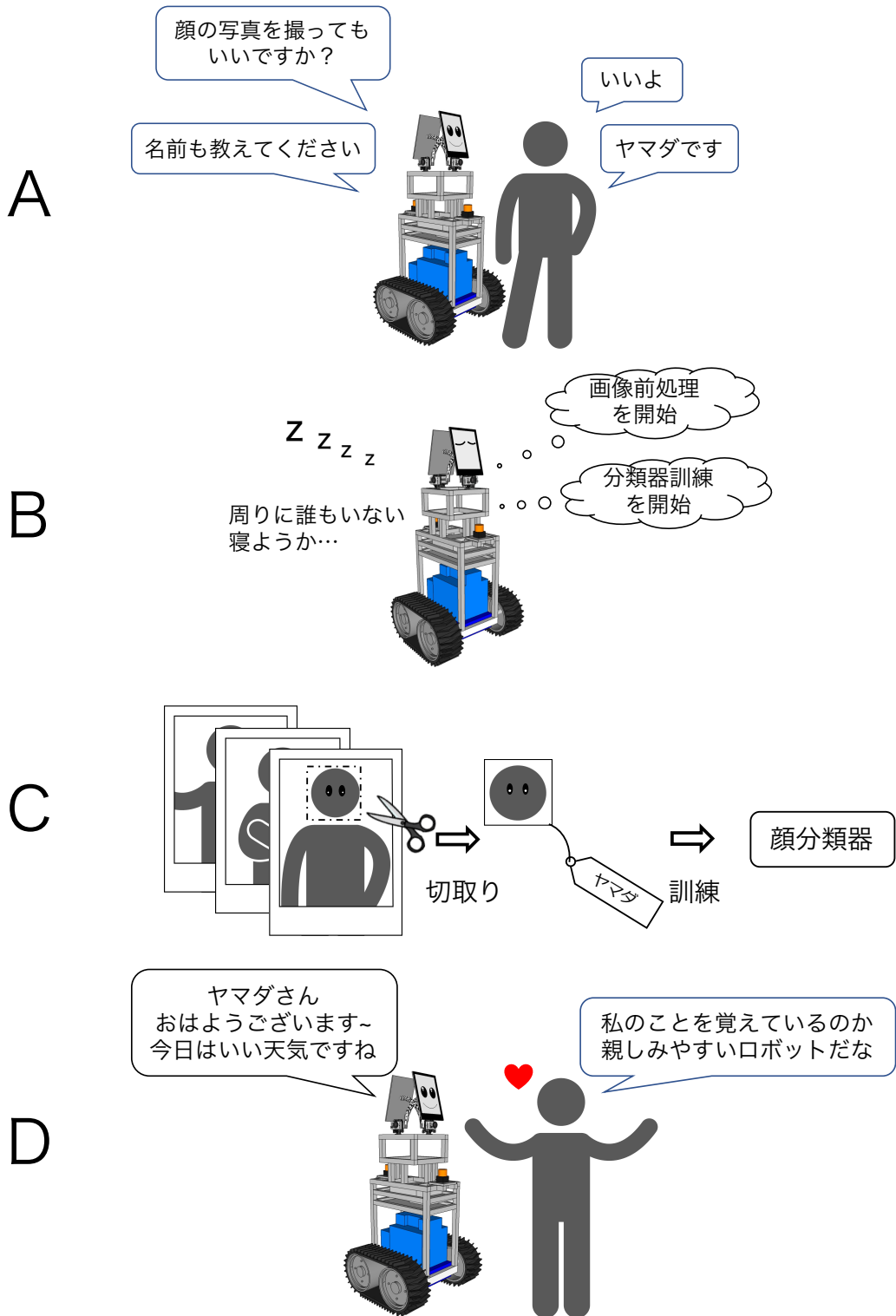


図 1.1: 提案手法の概要

## 1.4 論文の構成

本論文の第2章以降の構成は以下のようになっている。第2章では、提案手法で用いる数理 AIM モデルの元となる AIM モデルに関する説明、および数理 AIM モデルの仕組み、役割、応用方法について述べる。第3章では、数理 AIM モデルと組み合わせる為の実装した顔記憶システムの構成および関連技術についての説明をする。第4章では、数理 AIM モデルと顔記憶システムを用いた実験の詳細を説明し、その後は実験の結果と考察を述べる。第5章では、論文全体のまとめと今後の課題について述べる。

## 第2章 意識モデル

### 2.1 Hobson の AIM モデル

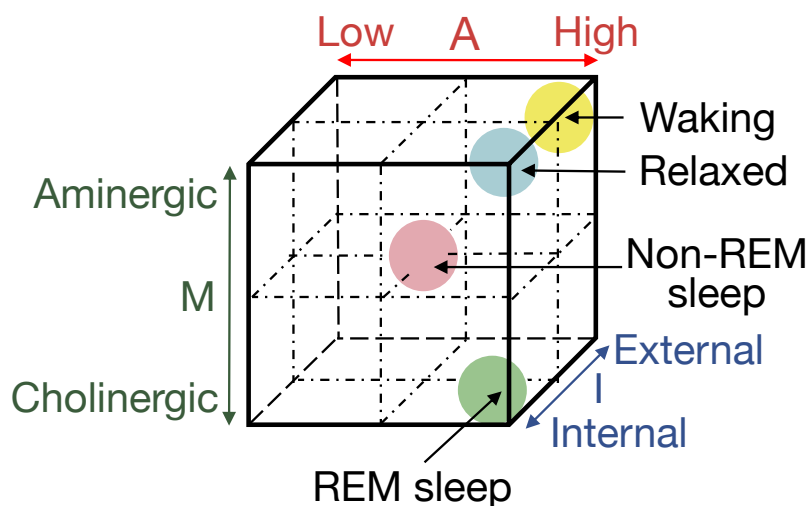


図 2.1: Hobson が提案した AIM モデル

人間の脳は状況に応じて柔軟に情報の処理を行える。Hobson は、図 2.1 のような人間の脳の意識状態を表現する為の AIM モデルを提案した [15]。AIM モデルは、*A*(Activation)、*I*(Input)、および *M*(Modulation) の 3 つの要素から構成され、3 次元の立方体空間で人間の意識状態を表現することができる。要素 *A* は、脳の全体的な活動レベルを表す。これを脳に処理されている情報の量としても理解できる。要素 *I* は処理する外部/内部情報の比率を表現する。情報は源から外部情報と内部情報に分類できる。外部情報は外部の感覚器官から得られ、内部情報は脳内の記憶などの内部情報源から生成される。要素 *M* は、外部/内部情報の処理モードを表す。実際は、*M* は脳内の化学物質のコリン作動性物質に対するアミン作動性物質の割合で決められる。アミン作動性物質の割合が高い場合は、主に外部情報が処理される。逆に、コリン作動性物質の割合が高い場合は、情報処理の方式は主に内部情報の処理するように切り替わる。

人間は主にレム睡眠の状態の時に夢を見られている [16]。レム睡眠は睡眠の状態だが、実はこの状態における脳の活動レベルは覚醒状態のように高く、夢を見るような複雑な処理が行われる。これは Hobson によって提案された AIM モデルの立方体の中で良く表現できている。それに対して、ノンレム睡眠状態の時に、脳の活動レベルが低く、情報の種類に関わらず、全体的に行われる処理が少ない。

## 2.2 数理 AIM モデル

Hobson が提案した AIM モデルに基づき、Mikawa らはロボットの知覚情報処理システムを制御する為の数理 AIM モデルを提案した [3]。AIM モデルが人の意識状態を表現するのに対し、数理 AIM モデルはロボットの意識状態を表現する。数理 AIM モデルは外部情報処理系と内部情報処理系の 2 つの処理系の動作を状態に応じて制御する。外部情報処理系では、ロボットに取り付けたセンサー類を用い、外部環境と状況の変化に対応できるようにデータを収集し、その場でリアルタイムで処理する。それに対して、リアルタイムで処理する必要がないデータ、またはリアルタイムで処理するのが困難なデータ、およびまとめて処理する必要があるデータは蓄積され、レム睡眠状態の時に内部情報処理系で処理される。

AIM モデルとは異なり、数理 AIM モデルにおける要素  $A$ 、要素  $I$ 、および要素  $M$  は異なる意味を与えられた。数理 AIM モデルにおいて、要素  $S$  は外部/内部情報から刺激情報に変換する為に使われる。要素  $A$  は刺激情報に基づき外部/内部情報の処理頻度を定める。要素  $I$  は、刺激情報の感度パラメーターを決定する。要素  $M$  は、データサンプリングの頻度を決定する。

外部情報処理系/内部情報処理系で用いる要素を使い分ける為に、要素  $A$ 、要素  $I$ 、および要素  $M$  は、さらに細かく  $a_{ex}$ 、 $a_{in}$ 、 $i_{ex}$ 、 $i_{in}$ 、 $m_{ex}$ 、 $m_{in}$  に分けられている。添字  $ex$ 、 $in$  はそれぞれ外部、内部の意味を持つ。

$a_{ex}$ 、 $a_{in}$ 、 $i_{ex}$ 、 $i_{in}$ 、 $m_{ex}$ 、 $m_{in}$  の値は経過時間  $t$  に伴って変化する。 $a_{ex}(t)$  の値は以下の式 (2.1) で決められる。

$$a_{ex}(t) = \begin{cases} L_w + b & (t < t_1) \\ \frac{L_w - L_a}{2} \left( 1 + \cos\left(\frac{2\pi(t-t_1)}{f_w}\right) \right) + L_a + b & (t_1 \leq t < t_2) \\ L_a + b & (t_2 \leq t < t_3) \\ \frac{L_a - L_n}{2} \left( 1 + \cos\left(\frac{2\pi(t-t_3)}{f_a}\right) \right) + L_n + b & (t_3 \leq t < t_4) \\ \frac{o_r}{2} \left( 1 - \cos\left(\frac{2\pi(t-t_4)}{f_r}\right) \right) + L_n + b & (t_4 \leq t) \end{cases} \quad (2.1)$$

$i_{ex}(t)$ 、 $m_{ex}(t)$  は式 (2.1) と同様の式で表せる為、ここでは説明を省略する。式 (2.1) の中の  $L_w$ 、 $L_a$ 、 $o_r$ 、 $L_n$ 、 $b$  はそれぞれ定数であり、目的に応じて自由に設定することができる。三河らは人間の睡眠パターンを参考に、以上の定数を以下の式 (2.2) のように設定した。

$$L_w = 0.75, \quad L_a = 0.50, \quad o_r = 0.25, \quad L_n = b = 0.00 \quad (2.2)$$

$a_{in}(t)$  の値は次に示す式 (2.3) で決められる。 $i_{in}(t)$ 、 $m_{in}(t)$  は同様に表せる為、それらに関する式もここでは省略する。

$$a_{in}(t) = \begin{cases} L_w + b & (t < t_1) \\ \frac{L_w - L_a}{2} \left( 1 + \cos\left(\frac{2\pi(t-t_1)}{f_w}\right) \right) + L_a + b & (t_1 \leq t < t_2) \\ L_a + b & (t_2 \leq t < t_3) \\ \frac{L_a - L_n}{2} \left( 1 + \cos\left(\frac{2\pi(t-t_3)}{f_a}\right) \right) + L_n + b & (t_3 \leq t < t_4) \\ \frac{o_r}{2} \left( 1 - \cos\left(\frac{2\pi(t-t_4)}{f_r}\right) \right) + L_n + b & (t_4 \leq t) \end{cases} \quad (2.3)$$

この場合、 $L_w$ 、 $L_a$ 、 $o_r$ 、 $L_n$ 、 $b$ の値は式(2.4)のように設定した。

$$L_w = 0.25, L_a = 0.25, o_r = 0.75, L_n = b = 0.00 \quad (2.4)$$

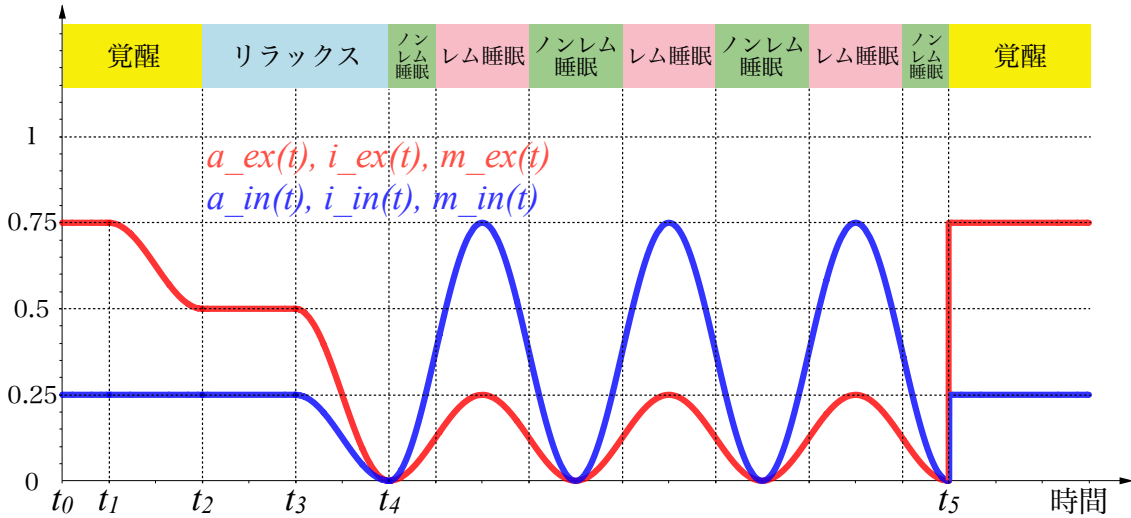


図 2.2:  $a_{ex}(t)$ 、 $i_{ex}(t)$ 、 $m_{ex}(t)$ 、 $a_{in}(t)$ 、 $i_{in}(t)$ 、 $m_{in}(t)$  の値の時間変化の例

$a_{ex}(t)$ 、 $i_{ex}(t)$ 、 $m_{ex}(t)$ 、 $a_{in}(t)$ 、 $i_{in}(t)$ 、 $m_{in}(t)$  の値の時間変化の例を図 2.2 に示す。 $t_0$  から  $t_5$  の間に、要素  $I$  によって決められる閾値  $th_s$  より大きな外部刺激が検出されていなく、各要素の値は時間の変化に伴って変化する。最初の意識状態は覚醒であったが、閾値より大きな外部刺激が検出されていない為、意識状態は睡眠の方向に遷移し、 $t_1$  時刻から各要素の値が下がる。その後、 $t_2$  から意識状態は覚醒からリラックスに変わる。リラックスの状態では、しばらく値は変化しないが、一定時間内に閾値より大きな外部刺激を検出できない場合は、 $t_3$  から各要素の値はさらに下がる。意識状態は  $t_4$  から睡眠になる。最初の睡眠状態はノンレム睡眠であり、その後はレム睡眠になる。人間と同じように、数理 AIM モデルにおいて、眠っている間に閾値より大きな外部刺激が検出されない場合、意識状態はノンレム睡眠とレム睡眠の間で交互に変化する。この図においてもそのような周期的な変化が見られる。また、ノンレム睡眠の時の各要素の値が比較的に低く、レム睡眠の時の各要素の値が比較的に高いとのことも分かる。最後に、 $t_5$  時刻で閾値  $th_s$  より大きな刺激が検出され、意識状態は覚醒状態に遷移する。

システムの外部情報処理、内部情報処理、外界センサーおよび記憶装置の個数をそれぞれ  $p_{ex}$ 、 $p_{in}$ 、 $q_{ex}$ 、 $q_{in}$  にし、要素  $A$ 、要素  $I$ 、および要素  $M$  の値は次に示す式 (2.5)~(2.7) で求められる。 $p_{ex}$ 、 $p_{in}$ 、 $q_{ex}$ 、 $q_{in}$  は 1 以上の整数であり、それぞれ外部/内部情報処理の個数、外界センサの個数、記憶装置の個数を表す。

$$A(t) = \frac{1}{p_{ex}} \sum_{i=1}^{p_{ex}} a_{ex_i}(t) + \frac{1}{p_{in}} \sum_{j=1}^{p_{in}} a_{in_j}(t) \quad (2.5)$$

$$I(t) = \frac{2}{p_{ex}} \sum_{j=1}^{p_{in}} i_{in_j}(t) - \frac{2}{p_{in}} \sum_{i=1}^{p_{ex}} i_{ex_i}(t) \quad (2.6)$$

$$M(t) = \frac{2}{q_{ex}} \sum_{i=1}^{q_{ex}} m_{ex_i}(t) - \frac{2}{q_{in}} \sum_{j=1}^{q_{in}} m_{in_j}(t) \quad (2.7)$$

$p_{ex}$ 、 $p_{in}$ 、 $q_{ex}$ 、 $q_{in}$  が全て 1 に設定される場合の要素  $A$ 、要素  $I$ 、および要素  $M$  の値の時間変化の例を図 2.3 と 2.4 に示す。これら 2 つの図は、前述の図 2.2 に対応している。図 2.3 を参照すると、要素  $A$  の値の変化範囲は  $[0,1]$  であることが分かる。要するに、現在のような設定で数理 AIM モデルを稼働させる場合、図 2.1 中の意識状態を表現するボールは立方体の中の右半分の領域のみに出現する。また、要素  $I$  の値は要素  $M$  の値の反数であることも分かる。最後に、図 2.5 に実際に稼働している数理 AIM モデルの GUI を示す。

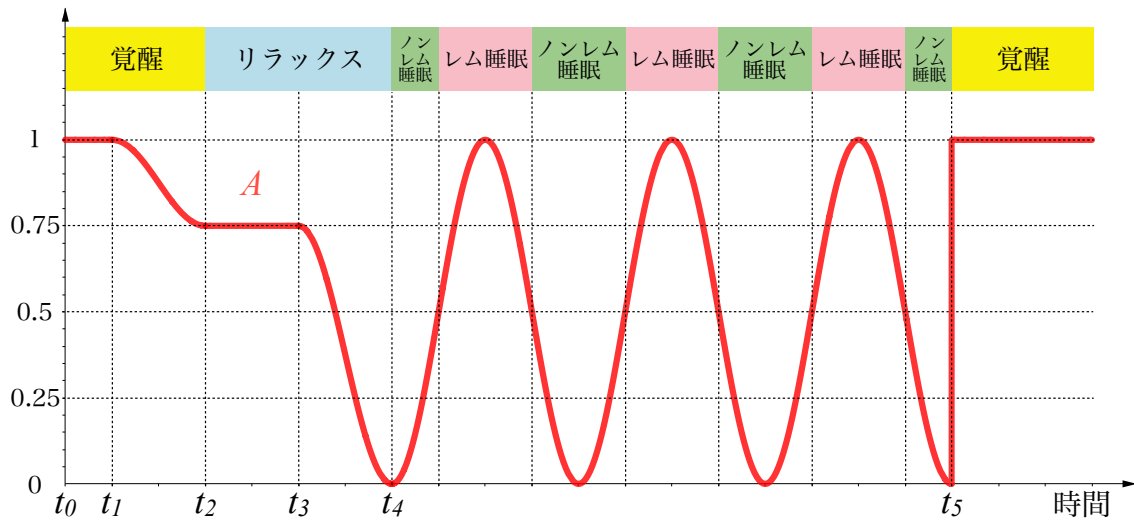


図 2.3: 要素 A の値の時間変化の例

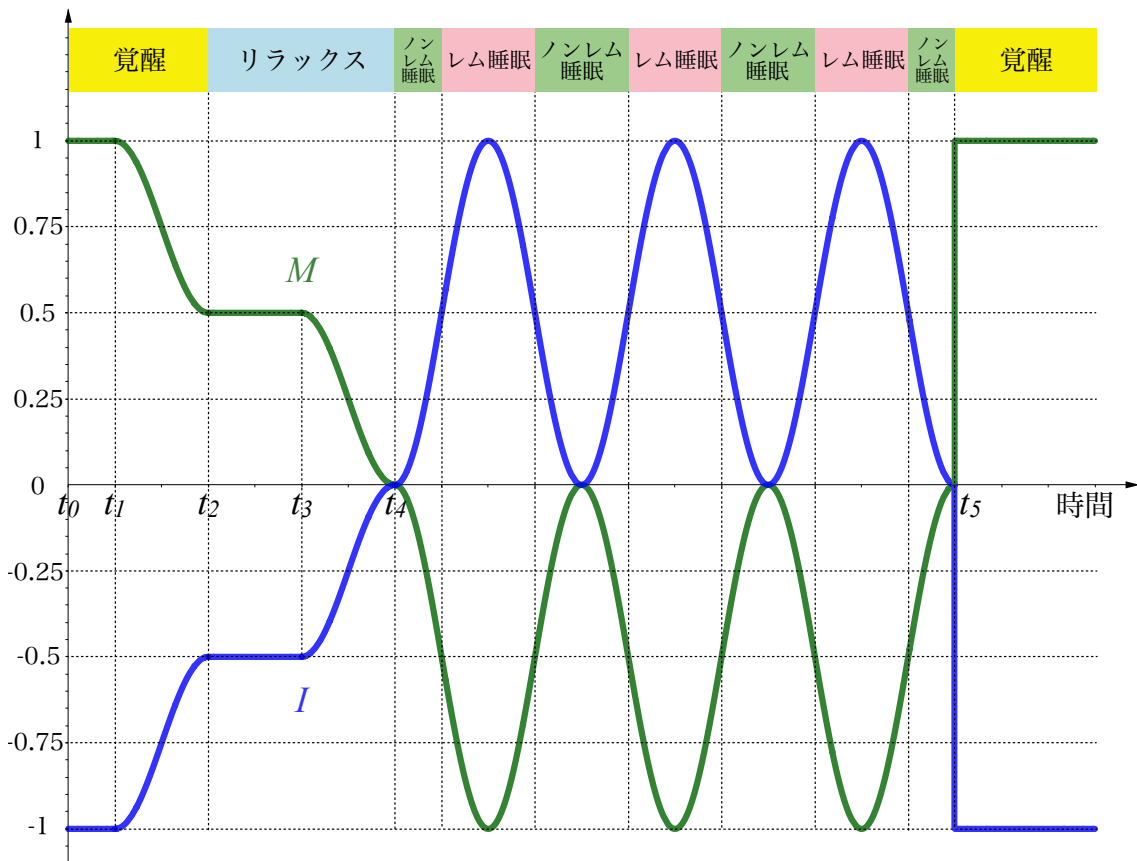


図 2.4: 要素 I、M の値の時間変化の例

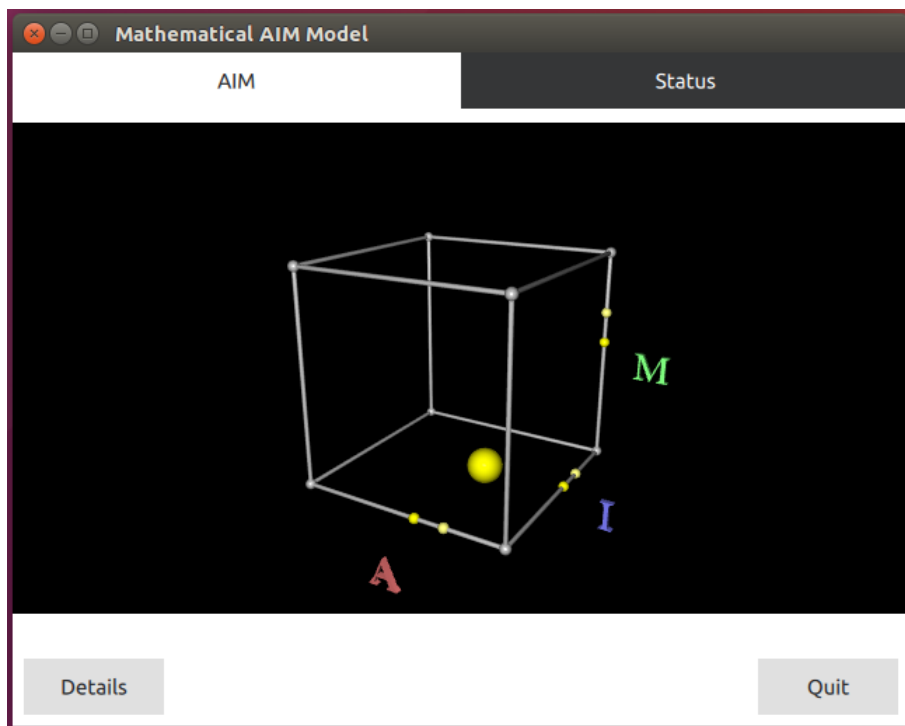


図 2.5: 数理 AIM モデルの GUI



## 2.3 数理 AIM モデルに基づくプロセス制御

本研究では、数理 AIM モデルを使用し、ロボットの意識状態を外部からの刺激情報に基づいて遷移させ、周囲に人がいない時に一部の情報処理の処理頻度を下げ、また、睡眠状態の時に、採集した顔画像の前処理と分類器の訓練をさせることで、CPU とバッテリーの有効利用を図る。前述の説明をまとめ、意識状態の遷移図を図 2.6 に示す。また、要素  $A$  と要素  $M$  の値と該当する意識状態の関係を表 2.1 に示す。

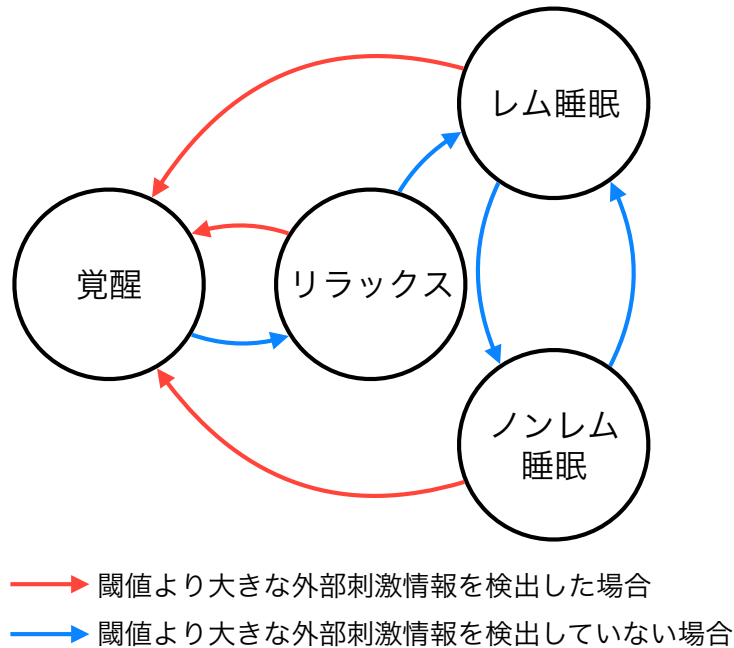


図 2.6: 数理 AIM モデルに基づく意識状態の遷移ルート

表 2.1: 要素  $A$  の値、要素  $M$  の値と意識状態の関係

要素 $A$ の値	要素 $M$ の値	該当する意識状態
$1 \geq A > 0.75$	$-0.5 > M \geq -1$	覚醒
$0.75 \geq A > 0$	$0 > M \geq -0.5$	リラックス
$0.5 > A \geq 0$	$0.5 > M \geq 0$	ノンレム睡眠
$1 \geq A \geq 0.5$	$1 \geq M \geq 0.5$	レム睡眠

数理 AIM モデルと顔記憶システムの組み合わせにおいて、数理 AIM モデル側から要素  $A$  と要素  $M$  の値を取得することで、該当する意識状態を知ることができる。その時の意識状態に応じて、顔記憶システムの中の特定のプロセスを実行することと停止させることができる。また、要素  $a_{ex}$  と  $m_{ex}$  の値によって、外界センサーに関わる特定のプロセスの実行頻度を決定する。

# 第3章 顔記憶システム

## 3.1 システム構成

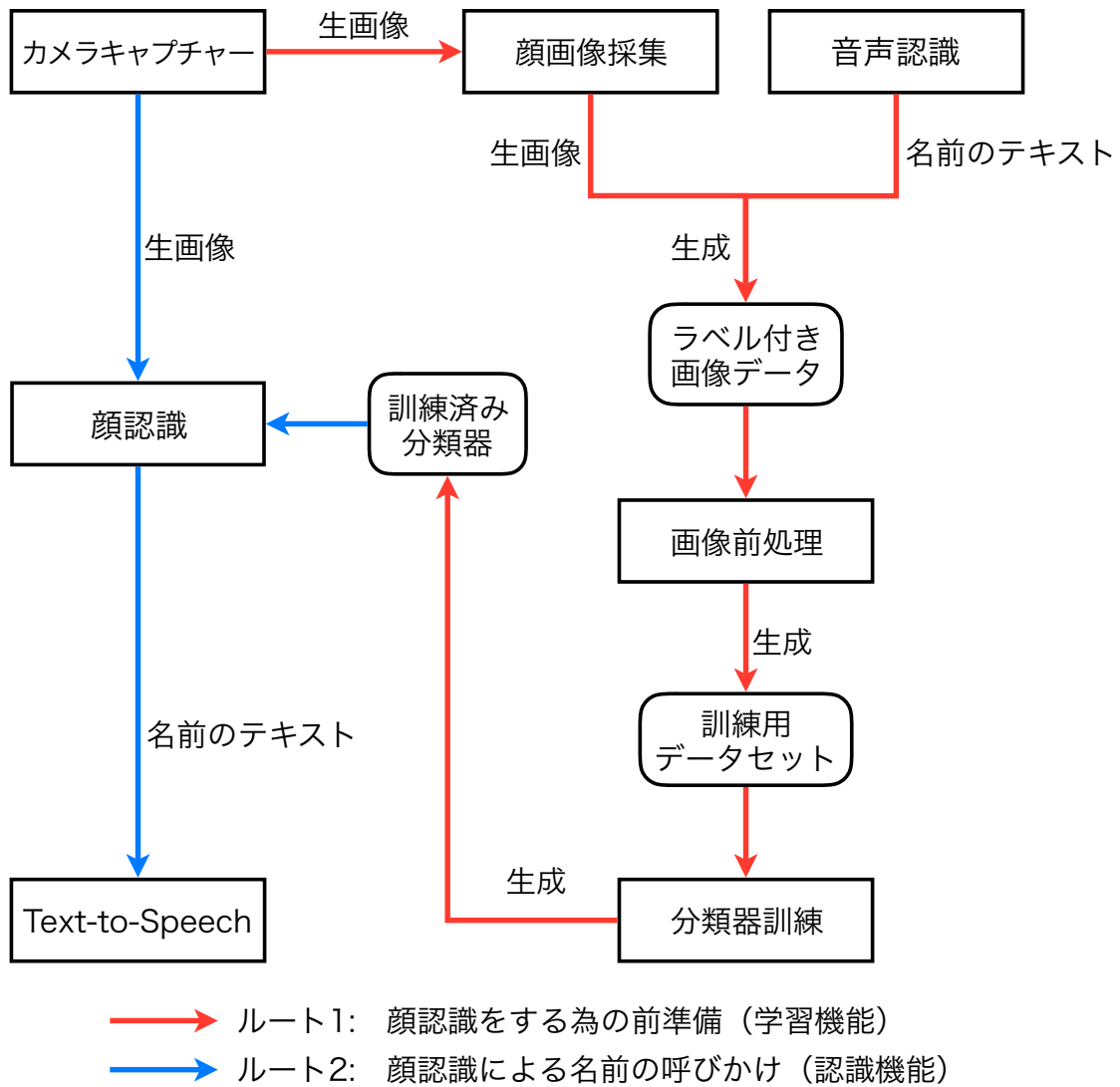


図 3.1: 顔記憶システムの処理の流れ

本研究が提案する顔記憶システムの処理の流れを図 3.1 に示す。

図 3.1 の中では、直角の四角形はプロセスを表し、角の丸い四角形は生成されたデータを表す。このシステムにおいて、2 つのルートが設定されている。なお、顔認識プロセスと画像前処理プロセスの中には、顔検出の機能が組み込まれている。

ルート 1 は、顔認識をする為の前準備となる顔画像の学習機能を実現する。まず、カメラキャプチャプロセスから、カラー画像を生画像として取得し、顔画像採集のプロセスに伝送する。それらの顔画像は、音声認識プロセスから出力される名前のテキストと組み合わせ、訓練用のラベル付き画像データとなる。その後、画像前処理を行い、訓練用データセットを生成する。最後は分類器の訓練を経て、新しく分類済みの分類器を出力する。

ルート 2 では、ルート 1 で訓練された分類器を用いて顔認識を行い、最後は認識結果を日本語テキストとして Text-to-Speech プロセスに渡し、音声を合成することで人の特定および名前呼びかけを実現する。

数理 AIM モデルをこの顔記憶システムをに組み合わせることで、以下のようなことがメリットとして挙げられる。提案手法では、ロボットが周りに人がいない状況において覚醒状態から睡眠状態に遷移すると述べた。周囲に人がいるかどうかは、カメラから取れた画像に対して顔検出をすることにより把握する。しかし、周囲にすでに人がいない時は、顔検出の処理をずっと同じ頻度で繰り返すのは無駄である。このような状況において、数理 AIM モデルは顔検出プロセスの処理頻度を下げることができ、必要のない一部の処理を省ける。また、提案した顔記憶システムの中の画像前処理プロセスと分類器訓練プロセスにおける処理は重い処理である為、覚醒時のリアルタイムでの処理には不向きである。この時間のかかる処理は数理 AIM モデルによるできた睡眠時に実行し、顔認識の精度の向上、および CPU とバッテリーの利用率の向上を図る。以下の数節では、顔記憶システムの各プロセスの仕組みと関連技術を説明する。

## 3.2 顔検出

カメラで人の画像を撮影する場合、顔の領域以外に、背景などの顔とは関係ない部分も画像の中に含まれる。そのような背景を含む画像を顔認識や分類器の訓練に使用すると、顔認識の結果が悪くなり、訓練された分類器が精度も良くなれない。したがって、以上の問題を解決する為には、カメラで取得した画像から余分な部分を取り除くことが不可欠である。ここでの顔検出は、画像から顔の領域の特定することを意味する。

顔検出を実装する方法としては、OpenCVの Haar Cascades 分類器 [17] または Dlib の顔検出器 [18] を使用することが考えられる。しかしながら、これらの方法の精度は背景を含む複雑な画像においては良くなく、実用的な用途に適用することは困難である。

そこで本研究は、MTCNN [19] と呼ばれる顔検出の手法を採用した。この手法は、複雑な環境における照明、顔の角度および表情の変化に対してよりロバストであるため、顔検出において優れた性能を示している。それだけではなく、この手法は計算量が少なく計算速度が速く、リアルタイムで画像から顔領域を検出することができる。

MTCNN のアルゴリズムは次に示す3つのステップで機能し、それぞれ1つのニューラルネットワークを使用する。まずは Proposal Network(P-Net) を使用して顔の潜在位置を予測し、それらのバウンディングボックスを出力するというのは最初のステップである。ステップ1では多数の顔が検出されるが、それらのほとんどは間違いである。2番目のステップでは、最初のステップに使われた画像と出力画像を入力として使用し、次に Refinement Network(R-Net) を使用して誤った結果のほとんどを排除する。ステップ3はステップ2とほぼ同様に機能する。ステップ2の入力画像とその出力画像は、ステップ3の入力として使用され、次に Output Network(O-Net) を使用して結果を再度絞り込み、最後にバウンディングボックスと5つの顔のランドマークの位置を出力する。

我々のシステムでは、MTCNN による顔検出を顔認識の前に実行され、また、MTCNN 法による顔検出は画像前処理プロセスの主な機能でもある。順番としては、まず、USB カメラを使って  $640 \times 480$  ピクセルの解像度を持つ生画像を取得する。次に、MTCNN 法を使用して生画像の顔領域を検出し、その領域を  $160 \times 160$  ピクセルの解像度の画像に縮小する。これらの前処理された画像は、リアルタイムの顔認識または分類器訓練時に使用するデータセットとして使用される。実際に MTCNN を使用して顔検出を行った結果を図 3.2 に示す。



図 3.2: MTCNN による顔検出の実例

### 3.3 顔認識

深層学習の手法を使った画像認識は、層の深い畳み込みニューラルネットワークを訓練することで実現できる。この方法は正確に画像を該当するクラスに分類することができるが、大量の訓練用データが必要である。さらに、ネットワークの訓練自体も相当に時間がかかる為、高性能の CPU や GPU がない状況では実現することが難しい。本研究が提案手法の場合だと、新しい人の顔が追加される為、高精度を維持する為に時間のかかる訓練を繰り返すのは非現実的だと考えられる。

そこで本研究では、顔認識を実現する為に、Schroff らによって開発された FaceNet [20] を使用する。FaceNet は LFW(Labeled Faces in the Wild) [21] データセットに対して高い精度を達成した為、近年注目を浴びている手法である。

FaceNet のモデルは、任意の顔画像に対して、最適な 128 次元の特徴ベクトルを出力できるように訓練される。この 128 次元の特徴ベクトルを用いて 128 次元のユークリッド空間における 2 つの特徴ベクトルの距離を計算する時に、同一人物の顔であればその距離はより小さく、逆に異なる人の顔である場合は、その距離はより大きいというようになる。図 3.3 に示すように、A と B は異なる人物であり、異なる顔を持つ為、FaceNet のモデルは訓練される際は、同じ A の顔であれば、その間の距離は小さく、逆に A 以外の人物の顔との距離大きくなるように最適化される。

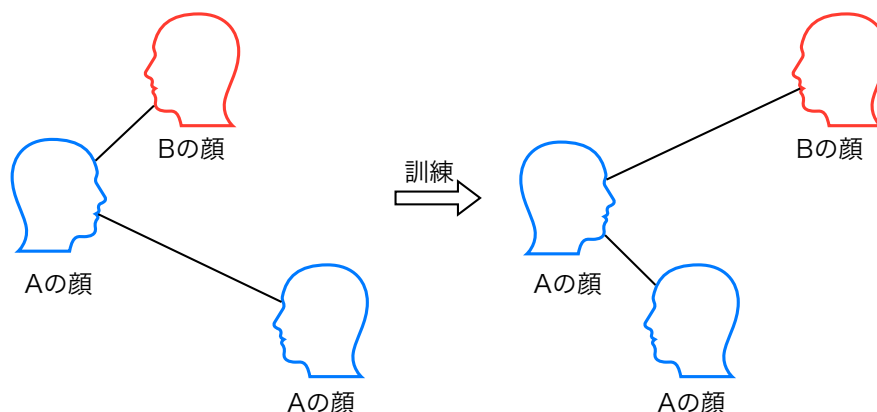


図 3.3: FaceNet モデルの訓練方式

FaceNet のモデルは、長時間をかけて数多くの人からなる多数の顔画像で構成されたデータセットで訓練される。訓練済みモデルを入手すれば、任意の顔画像からもその顔の特徴ベクトルを出力することができる。したがって、訓練済みの FaceNet のモデルを入手した後、新しいクラスの顔が追加される時もモデル自体を再度訓練する必要はない。

提案方法では、事前に訓練された FaceNet モデルを使用して顔認識を実装した。David Sandberg は GitHub レポジトリ [22] で VGGFace2 [23] データセットで訓練された FaceNet のモデルを提供しており、本研究はそれを使用した。また、David Sandberg は認識精度を向上させる為に、モデルの埋め込みサイズを FaceNet の原著論文 [20] に記載されている 128 次元から 512 次元に変更した。

### 3.4 画像前処理と分類器訓練

高精度の顔認識を実現する為に、画像前処理と分類器訓練の2つの処理が必要である。

画像前処理は、覚醒時に大量に集めた顔画像に対し、MTCNNを使って顔検出を行い、処理後は160×160ピクセルの解像度の画像を出力する。その後、これらの処理された顔画像を使用して分類器の訓練を行う。

3.3節において、FaceNetでは、2つのベクトルの距離を比較するだけでも顔認識機能の実現が可能であると述べたが、実はそれはクラスの数が多い状況に限った話である。分類するクラスの数が多い場合は、入力顔画像のベクトルと以前に保存された顔画像のベクトルとの距離を比較することで顔認識を実現することが可能である。しかし、この方法ではクラス数が増えると高精度を保つのが難しくなる。この場合、訓練済みのFaceNetモデルを用いて顔検出処理された画像を512次元のベクトルに変換し、そのベクトルを使ってSVM分類器を訓練することで問題は解決できる。本研究では、scikit-learn [24] ライブラリのSVC(C-Support Vector Classification)メソッドを使用し、分類器の訓練を行うことで、認識精度を向上させる。画像前処理過程と分類器訓練過程を図3.4に示す。また、実際に画像前処理、分類器訓練を経た後のリアルタイム顔認識を行っている様子を図3.5に示す。

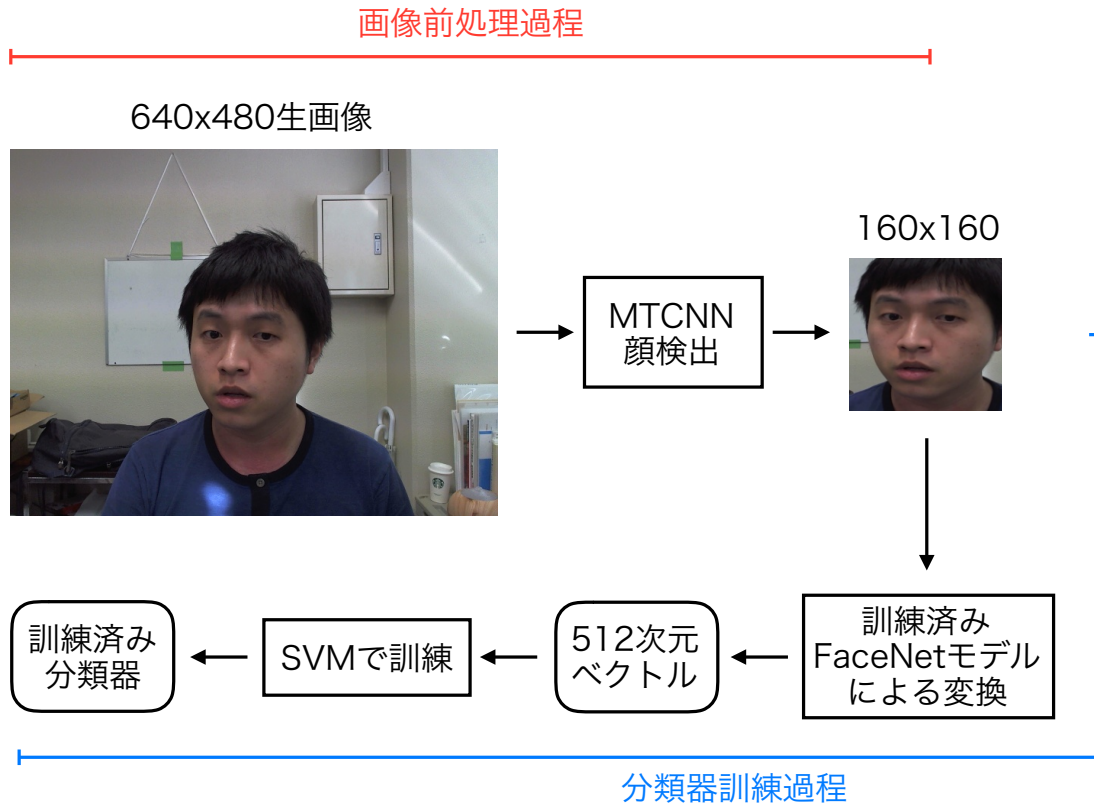


図 3.4: 画像前処理過程と分類器訓練過程

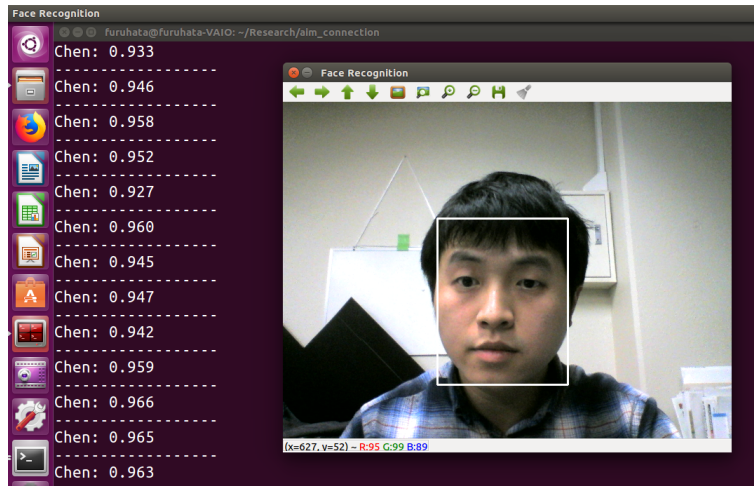


図 3.5: リアルタイム顔認識を行う様子

### 3.5 音声認識と音声合成

人の名前情報を収集する際、音声認識を使用して名前情報を取得する。その為に、オープンソースの大語彙連続音声認識エンジンである Julius [25] を用いて音声認識のプログラムを実装した。音声認識を使い、音声情報からテキストに変換し、その後は分類器訓練を行う時のクラスのラベル名として扱う。

より正しく日本人の苗字の発音を覚える為に、専用の辞書を作る必要がある。実際に研究室のメンバーの苗字の辞書を作り、少人数の場合だと高い精度が得られることが分かった。しかし、苗字の種類が増えるにつれ、精度が下がることも分かった。日本人の苗字の読み方は難しいものであり、その上に種類も多い為、多くの人の苗字を高精度で認識することは困難である。実際に応用する際、音声認識のミスによる名前の呼ぶ時の人間違いをもたらす悪影響のリスクも無視できない。以上のことを考えると、提案手法を検証する為の実験協力者による印象評価実験において、ミスを防ぐ為に音声認識を使わず、Wizard-of-OZ(WoZ)法 [26](詳細は 4.3.5 項に述べる)を使用することにした。

ロボットの発話は、Open JTalk [27] という日本語の音声合成システムを使用して日本語テキストを音声に合成することで実現した。ロボットが話す言葉を事前に考え、シナリオのように日本語のテキストとして書き、適切なタイミングで合成された音声を再生することで、人の話しかけ、名前を呼ぶことができる。実際の人との対面における発話内容および音声合成に関する設定などは 4.3.7 項に述べる。

## 第4章 実験と考察

### 4.1 実験概要

本章では次に示す2種類の実験を通して提案手法の有効性を評価する。

実験1では数理 AIM モデルと顔記憶システムを組み合わせ、数理 AIM モデルありと数理 AIM モデル無しの2つの状況におけるコンピュータの CPU の使用率およびロボットのバッテリーの電圧の変化を確認する。

実験2では実験協力者による印象評価実験を行う。名前を呼ぶと名前を呼ばないの2つの条件において、実験協力者にロボットとのインタラクションを実際に体験してもらい、その後ロボットへの印象評価を通して提案手法の中の名前を呼ぶことの効果の検証する。

実験3では、実装した顔検出機能と顔認識機能の実行速度の評価する。また、集めた顔画像を用いて画像前処理と分類器訓練を行い、処理に必要な時間および訓練済みの分類器の正解率を調査してまとめる。



## 4.2 実験 1

### 4.2.1 概要

数理 AIM モデルと顔記憶システムを組み合わせることによる計算機の CPU およびロボットのバッテリーの有効利用の効果を検証する為に、実験 1 を行った。実験 1 はさらに Part1 と Part2 に分かれる。Part1 と Part2 ではそれぞれ以下のことをした。

- Part1 : 数理 AIM モデルと顔記憶システムを組み合わせ、顔認識は常時に稼働し、画像前処理と分類器訓練を行い、この全過程の計算機の CPU の利用率の変化を見る。数理 AIM モデルあり条件と数理 AIM モデル無し条件の 2 つの条件で行う。
- Part2 : 顔認識プロセスを稼働させ、数理 AIM モデルあり条件と数理 AIM モデル無し条件の 2 条件におけるロボットのバッテリーの電圧の変化を計測する。画像前処理と分類器訓練は行わない。

Part1 の実験では、ノートパソコンと USB カメラのみを使用し、ロボットを使わなかった。Part2 の実験では、ロボットのバッテリーの残量を直接計測することが難しい為、使用するバッテリーの電圧が残量の減少に伴って下がる特性を利用し、代わりに計測が容易に行える電圧をノートパソコンで計測・記録した。ノートパソコンはコンセントを使わず、ロボットのバッテリーから電力を供給した。

### 4.2.2 システム構成

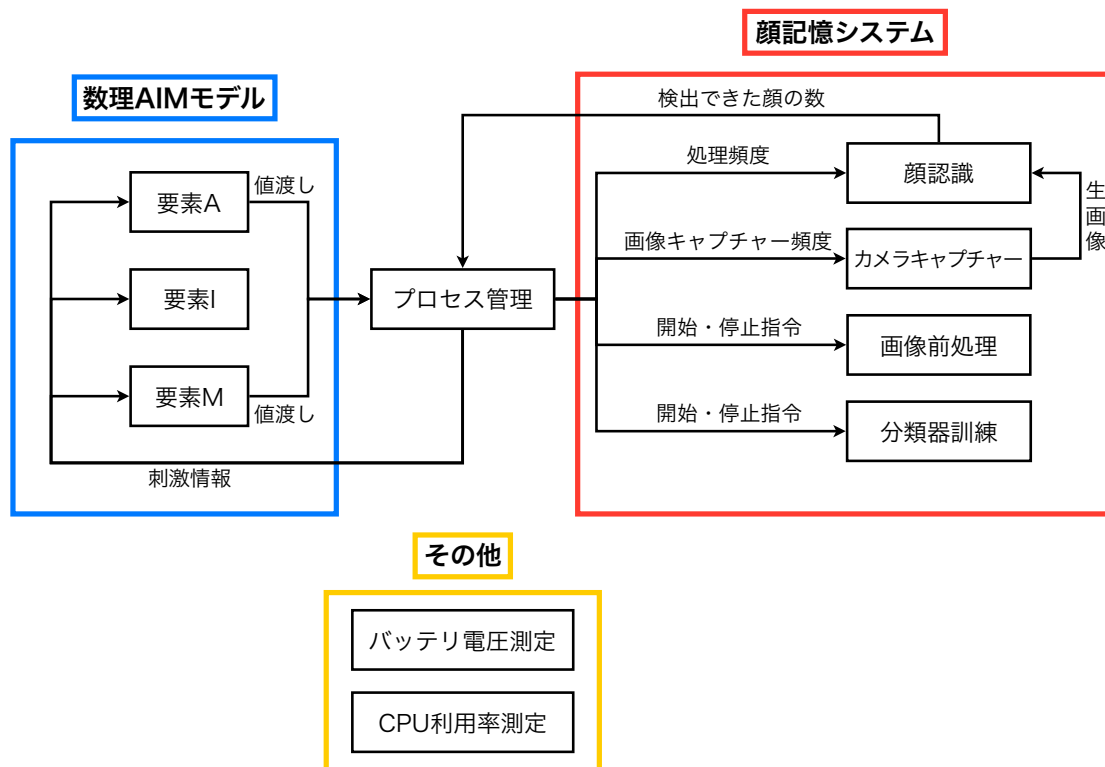


図 4.1: システム構成図 (実験 1)

本実験のシステム概要を図 4.1 に示す。この実験では、数理 AIM モデルの三要素の値を使用し、顔記憶システム側のプロセスの処理頻度と稼働状態を制御する。ここで示した顔認識プロセスと画像前処理プロセスの中に、既に顔検出の機能が組み込まれている。数理 AIM モデル側は、要素  $A$  と要素  $M$  の値をプロセス全体を管理する為の管理プロセスに渡し、意識状態を判断する。レム睡眠の時のみ、画像前処理プロセスと分類器訓練プロセスが作動する。レム睡眠とノンレム睡眠が交互で変化する為、画像前処理プロセスと分類器訓練プロセスも「作動-停止-作動-停止」のようになる。顔認識プロセスの実行頻度  $R_{fr}$  は  $a_{ex}(t)$  の値を使って式 (4.1) で決められる。同時に、カメラキャプチャーが USB カメラから画像を取り込む頻度  $R_{cp}$  は  $m_{ex}(t)$  の値を使って式 (4.2) で決められる。 $a_{ex}(t)$  と  $m_{ex}(t)$  は 0 から 0.75 まで変化する為、顔認識プロセスの実行頻度とカメラキャプチャーの画像取り込み頻度は、1 秒間 1 フレームから 1 秒間 10 フレームまでの間で変化することが分かる。

$$R_{fr} = \frac{a_{ex}(t)}{0.75} * 9 + 1 \quad (4.1)$$

$$R_{cp} = \frac{m_{ex}(t)}{0.75} * 9 + 1 \quad (4.2)$$

第 2 章における数理 AIM モデルの説明では、刺激情報を検出した場合、閾値  $th_s$  と比較し、閾値より高い場合に意識状態を覚醒に遷移させると説明したが、本実験では刺激情報だと認められるものを検出した場合、つまり USB カメラが顔を検出したら、近くに人がいると解釈し、閾値の比較なしでロボットの意識状態を覚醒に遷移させることにした。

プロセスの稼働には順番がある。顔認識プロセスとカメラキャプチャープロセスは常に稼働するが、その処理頻度は  $a_{ex}(t)$  と  $m_{ex}(t)$  の値で変化する。分類器訓練プロセスは、画像前処理プロセスの作業が全て終了してから動き始める。

本実験の画像前処理と分類器訓練を行う為に、あらかじめ 7 人から 1400 枚の顔を含める写真 (カラー画像、640x480 ピクセル) を収集した。USB カメラを用いて室内環境で 1 人あたり 200 枚を集め、ランダムで保存した。200 枚の中の 100 枚は分類器訓練に用い、残りの 100 枚は分類器の効果进行测试する為に用いた。

また、実際に画像前処理プロセスと分類器訓練プロセスを実行してみると、処理が重すぎるせいで他のプロセスの実行に支障を与えてしまうことが分かった為、画像前処理プロセスと分類器訓練プロセスの CPU 使用に制限を掛けた。本実験は 8 コア (実際は 4 コア 8 スレッド) の CPU を使用した。そこで、画像前処理プロセスは最大 2 コア使用可能、分類器訓練プロセスは最大 1 コア使用可能というように制限した。

### 4.2.3 ソフトウェア構成

本実験に使用したライブラリ一覧を表 4.1に示す。

表 4.1: 使用ライブラリ (実験 1)

ライブラリ名	用途	バージョン
ROS(Robot Operating System) [28]	プロセス間通信、プロセス処理頻度の制御	Kinetic
psutil [29]	CPU の利用率データの取得	3.4.2
Aria(Advanced Robotics Interface for Application) [30]	バッテリーの電圧データの取得	2.9.4
TensorFlow [31]	顔検出、顔認識のモデルの使用	1.12.0
scikit-learn [24]	分類器の訓練	0.20.1
OpenCV(Open Source Computer Vision Library) [17]	画像処理	3.3.1

ROS はプロセス間通信とプロセス処理頻度の制御、psutil は CPU の利用率データの取得、Aria はバッテリーの電圧データの取得、TensorFlow は顔検出および顔認識のモデルの使用、scikit-learn は顔認識の分類器の訓練、OpenCV は USB カメラからの画像の取得、加工、保存などを行う為に利用した。

### 4.2.4 ハードウェア構成

本実験で使用した機器を表 4.2、4.3に示す。

表 4.2: 移動ロボットに搭載したノートパソコン

種類	仕様
CPU	Intel Core i7-2630QM 2.00 GHz x 8
RAM	8GB
OS	Ubuntu 16.04 LTS

表 4.3: 実験機器 (実験 1)

種類	メーカー	製品名
USB カメラ	Logicool	HD Pro Webcam C920
移動ロボット	MobileRobots	Pioneer-3DX
インバータ	SANWA SUPPLY	CAR-DAV200W

実験 1 の Part2 では、移動ロボット Pioneer-3DX を使用した。他に、USB カメラは顔検出・顔認識の為に使用した。Part2 において、ロボットのバッテリーからノートパソコンの電力を供給する為に、インバータを使ってノートパソコンをロボットの電源に繋いだ。また、CPU の利用率を計測する実験では、ノートパソコンはロボットのバッテリーからの電力を使わず、直接コンセントからの電力を使った。実験機器の外観を図 4.2に示す。Part2 の実験でのハードウェア間の接続を図 4.3に示す。



図 4.2: ロボットの外観 (実験 1、Part2)

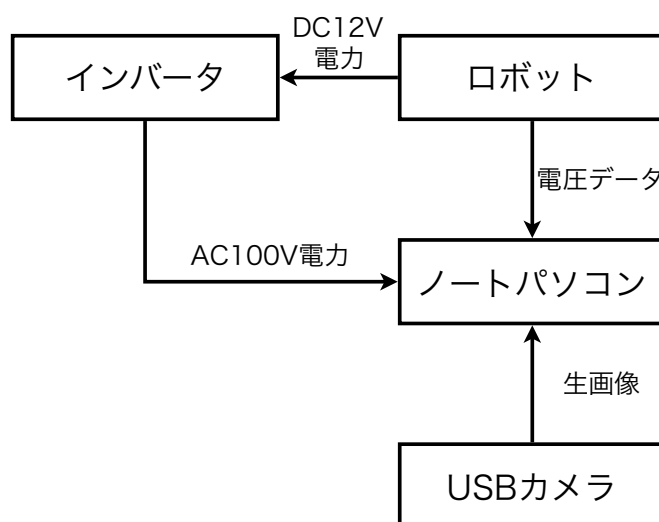


図 4.3: ハードウェア間の接続図 (実験 1、Part2)

#### 4.2.5 実験結果と考察 (Part1)

Part1 の実験を行い、CPU の利用率を取得し、そのデータをもとにグラフを描いた。Part1 の実験結果を表現するグラフを図 4.4 に示す。

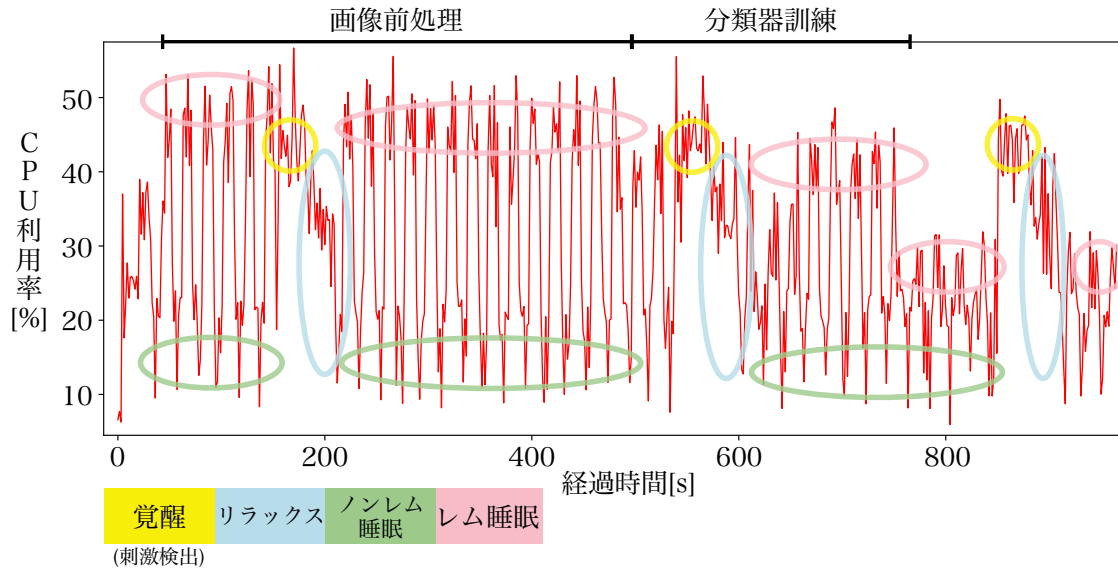


図 4.4: CPU の使用率の変化図

Part1 の実験では、実験システムに数理 AIM モデルを適用した。顔認識プロセスはこの実験において常時に稼働させた。実験の開始時刻から、USB カメラは刺激を検出できず、意識状態は覚醒-リラックス-ノンレム睡眠-レム睡眠という順番に変化していた。意識状態がレム睡眠状態になってから、画像前処理プロセスは稼働し始めた。刺激情報がない場合でも、意識状態はレム睡眠-ノンレム睡眠の間で交互に遷移する為、画像前処理プロセスは稼働と停止を繰り返していた。結果として、CPU の利用率が上下に周期的に変化することが分かった。分類器訓練の段階でも、同じような変化が見られた。画像前処理と分類器訓練がまだ完了していない段階で、それぞれ一度だけ顔をカメラの前に示し、刺激情報を与えた(約 30 秒)。その刺激を受け、意識状態はすぐに覚醒に戻り、CPU の利用率も急激に上昇した。最後に、分類器訓練が終わり、睡眠状態に入った後にもう一度刺激を与えた。刺激を受けた後、また CPU の利用率の急激な上昇が見られた。数理 AIM モデルなし状況では、意識状態は常時に覚醒状態に保つ為、CPU の利用率も常時 40% の水準になる。それに比べ、数理 AIM モデルを適用した後、必要ない時に CPU の利用率を抑えることができ、かつ睡眠時は意味のある作業ができる為、数理 AIM モデルと顔記憶システムの組み合わせることは CPU の有効利用に役立つと考えられる。

#### 4.2.6 実験結果と考察 (Part2)

Part2の実験を行い、バッテリーの電圧のデータを取得し、そのデータをもとにグラフを描いた。Part2の実験結果を表現するグラフを図4.5に示す。青い線は数理AIMモデルを適用した場合のバッテリー電圧の時間変化条件であり、赤い線は数理AIMモデルを適用しない場合での時間変化である。

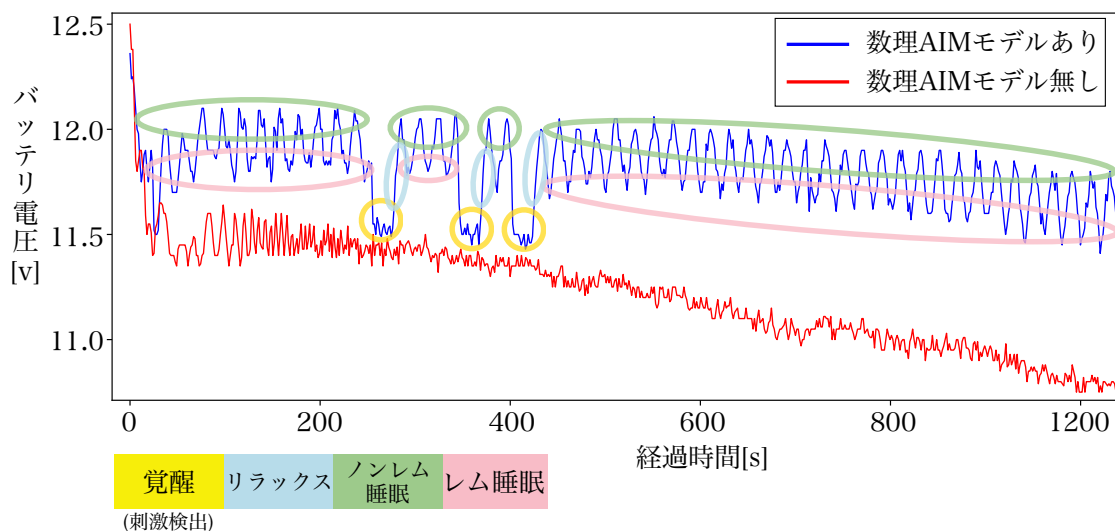


図 4.5: ロボットのバッテリーの電圧の変化図

Part2の実験では、顔認識プロセスは常時に稼働しており、画像前処理と分類器訓練は行わなかった。数理AIMモデルを適用した条件の実験において、合計3回刺激を与えた。刺激を受けた後、電圧の一時的な急速な下降が見られた。また、全体的に、青い曲線において明らかな周期的な電圧の上昇-下降の変化が見られた。数理AIMモデルを適用しない条件(赤い線)では、顔認識プロセスを常時に高い処理頻度で稼働していた為、バッテリーの消耗も青い線よりは激しかった。以上のことから、数理AIMモデルの適用は、バッテリーの有効利用に効果があると考えられる。

## 4.3 実験 2

### 4.3.1 概要

ロボットが人に名前を呼びかけることにより、人がロボットに対する親近感が良くなるかを検証する為の実験 2 を行った。この実験では、実験協力者にロボットとの三回の対面を体験し、ロボットに対する印象をアンケートで評価してもらった。実験協力者はこの三回の対面において、対話およびゲームを通してロボットとのインタラクションを実際に体験する。各対面の概要を以下のように示す。

- 一回目の対面：ロボットは実験協力者に挨拶や実験の説明をし、その後、実験協力者の苗字と顔画像を採集する。
- 二回目と三回目の対面：実験協力者はロボットとババ抜きゲームを遊ぶ。実験協力者は二枚のカードをシャッフルし、ロボットはジョーカーの位置を当てる。それぞれの対面において 6 ターンを繰り返す。

実験協力者は X 組と Y 組の 2 つの組に均等の人数に分けられる。一回目の対面は、本研究の提案手法の中のロボットによる名前情報と顔画像の採集を模擬するものである。実験協力者全員は同じ条件でロボットとの一回目の体験を体験する。しかし、実験協力者が属する組により、ロボットが二回目と三回目の対面における発話内容が異なり、条件 A と条件 B に分かれる。条件 A と条件 B を以下に示す。

- 条件 A：ロボットは該当する対面において、実験協力者の名前を呼ぶ。
- 条件 B：ロボットは該当する対面において、実験協力者の名前を呼ばない。

X 組の実験協力者は、条件 A-条件 B の順番で二回目と三回目の対面を体験する。それに対して、Y 組の実験協力者は、条件 B-条件 A の順番で二回目と三回目の対面を体験する。実験の実施順番は実験協力者に知らせない。また、二回目と三回目の対面は異なる条件で同じゲームを同じやり方で実施すると実験協力者に教示するが、条件の違いは最初は知らせず、実験終了後に知らせる。

### 4.3.2 システム構成

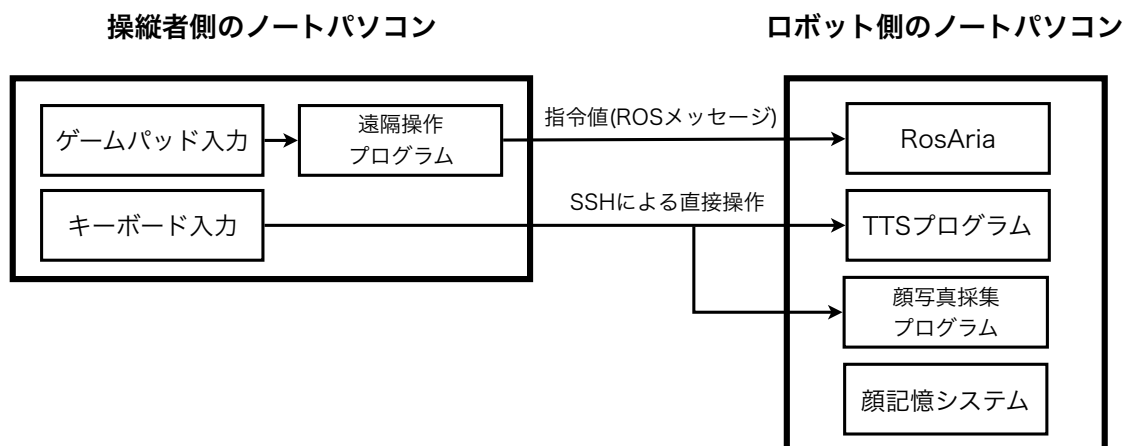


図 4.6: システム構成図 (実験 2)

本実験のシステム概要を図 4.6に示す。この実験では、4.3.5 項に後述する WoZ 法 [26] を使用する為、システムの構成は簡易化した。操縦者はノートパソコンに繋いだゲームパッドとキーボードを使用し、ロボットの発話と移動を操縦する。ゲームパッドからの指令値は、移動指令送信プログラムを通して ROS メッセージとしてロボット側ノートパソコンの ROSAria に送信され、ロボットの移動を制御する。また、SSH(Secure Shell) 接続で、操縦者は操縦者側のノートパソコンから直接ロボット側のノートパソコン上の TTS プログラムを動かし、キーボードの入力によってロボットの発話内容とタイミングを操縦できる。顔画像を採集する際も同様の方法で行う。実験協力者による印象評価実験のミスを防ぐ為に、本実験では顔記憶システムは後述の実験担当者が手動で動かす。

### 4.3.3 ソフトウェア構成

本実験に使用したライブラリ一覧を表 4.4に示す。

表 4.4: 使用ライブラリ (実験 2)

ライブラリ名	用途	バージョン
ROS(Robot Operating System) [28]	ロボットの操作、プロセス間通信	Kinetic
Aria(Advanced Robotics Interface for Application) [30]	ロボットの制御	2.9.4
Open JTalk [27]	日本語テキストからの音声合成	1.07
TensorFlow [31]	顔検出、顔認識のモデルの使用	1.12.0
scikit-learn [24]	分類器の訓練	0.20.1
OpenCV(Open Source Computer Vision Library) [17]	画像処理	3.3.1
RosAria [32]	Aria の ROS 上のインタフェース	0.9.0

ROS はロボットの関わる操作全般およびプロセス間通信、Aria はロボットの制御、OpenJTalk は日本語テキストからの音声合成、TensorFlow は顔検出および顔認識のモデルの使用、scikit-learn は顔認識の分類器の訓練、OpenCV は USB カメラからの画像の取得、加工、保存などの為に利用した。また、RosAria は ROS で Aria を使用する為のインタフェースとして用いた。

### 4.3.4 ハードウェア構成

本実験で使用した機器を表 4.5、4.6、4.7に示す。

表 4.5: 操縦者側のノートパソコン

種類	仕様
CPU	Intel Core i7-2630QM 2.00 GHz x 8
RAM	8GB
OS	Ubuntu 16.04 LTS



表 4.6: ロボット側のノートパソコン

種類	仕様
CPU	Intel Core i7-4700MQ 2.40 GHz x 8
RAM	8GB
OS	Ubuntu 16.04 LTS

表 4.7: 実験機器 (実験 2)

種類	メーカー	製品名
USB カメラ	Logicool	HD Pro Webcam C920
ゲームパッド	ELECOM	JC-U3712F
移動ロボット	Adept MobileRobots Inc.	Pioneer-3DX
大型タブレット PC	Apple	iPad Pro(12.9-inch、第 1 世代)
タブレット PC	Apple	iPad(9.7-inch、第 6 世代)
スマートフォン	Apple	iPhone 5s

顔画像の採集と顔認識の為に、USB カメラを使った。ゲームパッドはロボットの移動操作の為に用いた。実験 2 は実験 1 と同様に、移動ロボット Pioneer-3DX を使用した。しかし実験 2 では、ロボットは実験協力者の会話の相手である為、移動ロボットの本体のみを使用すると実験協力者にとってロボットとしての印象が薄く、ロボットというよりも小型の車という印象を避ける為に、移動ロボットの本体に加え、フレキシブルアームを使い、大型のタブレット PC をロボットの頭として設置した。また、3D の顔インタフェースをタブレット PC の画面に表示させ、それをロボットの顔として実験協力者に見せた。顔インタフェース上では、まばたきのアニメーションを表現できるロボットの両眼、および目の後ろに燃えているように揺れる炎が表示される。これはあくまでもロボットの外観を調整する為に使用したものであり、それを用いた情報処理などはない。ロボットの外観を図 4.7、4.8 に示し、顔インタフェースのスクリーンショットを図 4.9 に示す。

その他、本実験は 2 台のノートパソコンを使用した。また、実験室の状況を監視する為にもう 1 台のタブレット PC とスマートフォンを使用した。これらを使用する理由の詳細を 4.3.6 項に示す。



図 4.7: ロボットの外観 (正面、実験 2)



図 4.8: ロボットの外観 (側面、実験 2)

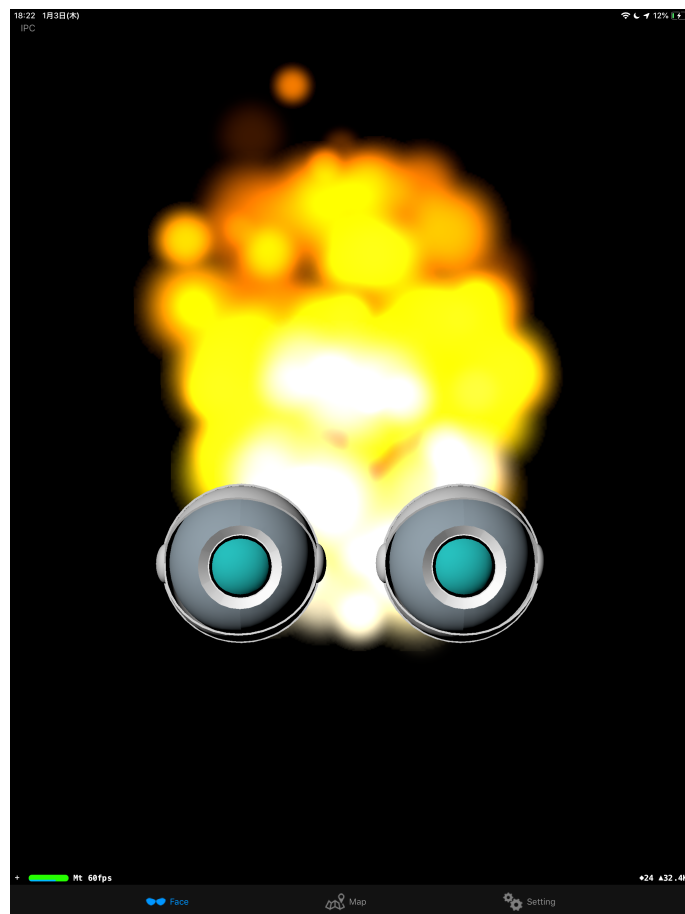


図 4.9: ロボットの顔インターフェース

#### 4.3.5 WoZ 法による実験

実験中のプログラムなどによるミスを防ぎ、およびババ抜きゲームの結果をコントロールする為に、WoZ(Wizard of Oz)法 [26] を使用した。今回の実験において、ロボットを自律的に行動させるのではなく、人が操作して作動するようにした。その為、この実験の遂行には実験協力者以外に二人が必要である。一人は実験担当者として、実験室で実験協力者の案内役を担当する。実験室での準備、実験内容の説明、およびアンケート調査の実施などは全て実験担当者が行う。もう一人は移動ロボットの遠隔操縦者として、実験室以外の別の場所でロボットの移動、発話内容、発話のタイミングを制御する。操縦者の存在は、アンケート内容の実施が全て終わるまで実験協力者に知らせない。

一回目の対面において、実験協力者の苗字の収集を行うが、エラーが生じると後の画像処理や分類器の訓練の妨げになる為、ロボットが実験協力者の名前を聞く時点で得られる苗字データは使わない。実験参加の連絡をする時にあらかじめ実験協力者の名前情報を入力し、それをクラスのラベルとして使い、ロボットの発話内容の中にも事前に入れる。同様の理由で、顔画像の採集も操縦者によって行う。こうすることによって、より少ないミスで円滑に実験を進められる。

本実験は2台のノートパソコンを使用した。2台のノートパソコンを同一の無線ネットワークに設置した。操縦者はその中の1台を使用し、SSHを使いもう1台のロボットに繋いでいるノートパソコンのプログラムを稼働させ、ゲームパットを使ってロボットに移動指令を送信する。

#### 4.3.6 ババ抜きゲーム

本実験で行うババ抜きゲームは、普段遊んでいるババ抜きゲームと異なるところがある。まず、今回のババ抜きゲームでは、二枚のカードのみを使用し、その中の一枚はジョーカーである。また、実世界における移動ロボットと人間の関係を考慮すると、単純な競争関係でババ抜きゲームをするのは妥当ではないと考えられる。その為に実験協力者に、「今回のゲームにおいて、あなたとロボットは競争関係ではなく、協力関係である。また、テクニックを使ってロボットにジョーカーの位置を隠すことは目標ではなく、逆にロボットにジョーカーの位置を当てさせることこそが目標です」と伝えた。

本実験は、より正確な実験データを取得する為に、実験の本当の意図を隠蔽し、実験協力者に実験の目的を「ロボットの顔記憶機能、微表情検出・分析機能の性能評価およびロボットに対する印象評価」と教えた。その他、以下の内容も実験協力者に教示した。ロボットは一回目の対面において集めた氏名情報と顔画像によって実験協力者の名前と顔を覚えてくれる。また、これらのデータの収集は二回目と三回目の対面でのババ抜きゲームの前準備である。ババ抜きゲームにおいて、ロボットは実験協力者の微表情を分析し、ジョーカーの位置を特定できる。実験協力者側としては、無理に表情を作る必要がなく、ジョーカーの位置を心の中で思えばそれが微表情としてロボットに伝わり、ロボットは実験協力者の表情を分析してジョーカーの位置を特定する。

しかし、ロボットに実装された機能の中には、顔記憶機能は存在するが、表情分析機能は存在しない。アンケート調査結果に影響を与えない為には、ゲームにおける各ターンの結果をコントロールする必要がある。そこで、本実験は、カードの内容の特定と実験時の様子を確認する為に、実験協力者の後ろの左斜め方向の所に隠しカメラを設置した。操縦者は隠しカメラ越しにジョーカーの位置を特定し、各ターンの結果をコントロールしていた。二回目と三回目のババ抜きゲームは、それぞれ6ターンを繰り返されるが、二回目と

三回目の対面の両方とも、4ターン目にロボットが必ず外すようにし、それ以外のターンでは必ず当たるようにした。

本実験でババ抜きゲームを採用したのは、以下の理由がある。まず、実験協力者をロボットとのババ抜きゲームに参加させることで、実験協力者は本実験の本当の目的に気づくことが難しくなる。また、実験協力者にはロボットの表情分析機能は事前に採集した顔画像に基づくものであると教示した為、名前を呼ばない条件で実施するゲームにおいても、一回目の対面での顔画像の採集の目的が怪しくならない。他に、もし二回目と三回目の対面において、実験協力者はただロボットと会話するだけというふうの設定すると、その会話自体が自然な感覚で行うものになるのが難しい。ゲームの中で適切なタイミングで適切な発話内容を挟むほうが自然な会話になりやすい。

#### 4.3.7 ロボットの発話

ロボットの音声発話は、Text-to-Speech(TTS) 音声合成システム OpenJTalk [27] を用いて実現された。使用する音声は Mei(Normal) [33] である。音声合成する際のパラメーターは、話速を 0.9 にし、その他は全てデフォルト値に設定した。ロボットの発話内容をあらかじめ日本語のテキストとして書き、プログラムの中に埋め込んだ。操縦者は、TTS プログラムを立ち上げる直後に、実験協力者の苗字の片仮名を入力し、その後は「名前を呼ぶ／名前を呼ばない」の中から一つを選ぶ。以上の入力によってプログラムは発話内容のテキストを構成する。実験実施時、操縦者は適宜なタイミングで各台詞に対応する番号または単語を入力することで、音声を即時に合成・再生する。

ロボットの発話内容は、実験協力者の名前と該当する実験条件によって異なる。ロボットの台詞に出現する実験協力者の名前は各実験協力者に対応する。二回目と三回目の対面において、名前を呼ぶ条件では実験協力者の名前を合計 10 回呼ぶ。それに比べ、名前を呼ばない条件では、名前が出現する 10 箇所の中の 8 箇所は空白になり、残りの 2 箇所は「あなた」に変化する。それ以外の発話内容は全て同様である。

コミュニケーション上の便利さの為に、実験で用いた移動ロボットの型番 Pioneer-3DX に因んで「パイオ」という名前をつけた。ロボットの発話内容全体は敬語を使い、丁寧で謙遜な雰囲気設定された。特に、ババ抜きゲームにおいて、協力的な関係を作る為に、ロボット側の台詞は実験協力者がロボットにジョーカーの位置を当てさせることで頑張っていることに対して感謝しているように設定した。また、実験の流れをより簡単に統一できる為に、三回の対面ともロボット側が主導権を握って会話を進めるような感じにした。実験協力者は実験内容の説明を受ける時、決められた言葉でロボットに返事するように教示された。実験中に実際に使用した台詞は付録 A に示す。

#### 4.3.8 実験環境

本実験は 2 つの部屋を使用した。その中の一つの部屋は、操縦者がロボットを操作する為の空間であり、特に配置などを考慮していない。実験自体はその隣の部屋「実験室」で行った。以下は実験室の配置についての説明をする。実験室の見取り図を図 4.10 に示す。実験室の配置に着手する前に、既に多くのものが現場に置かれている為、実験とは無関係のものは見取り図の上には示さない。

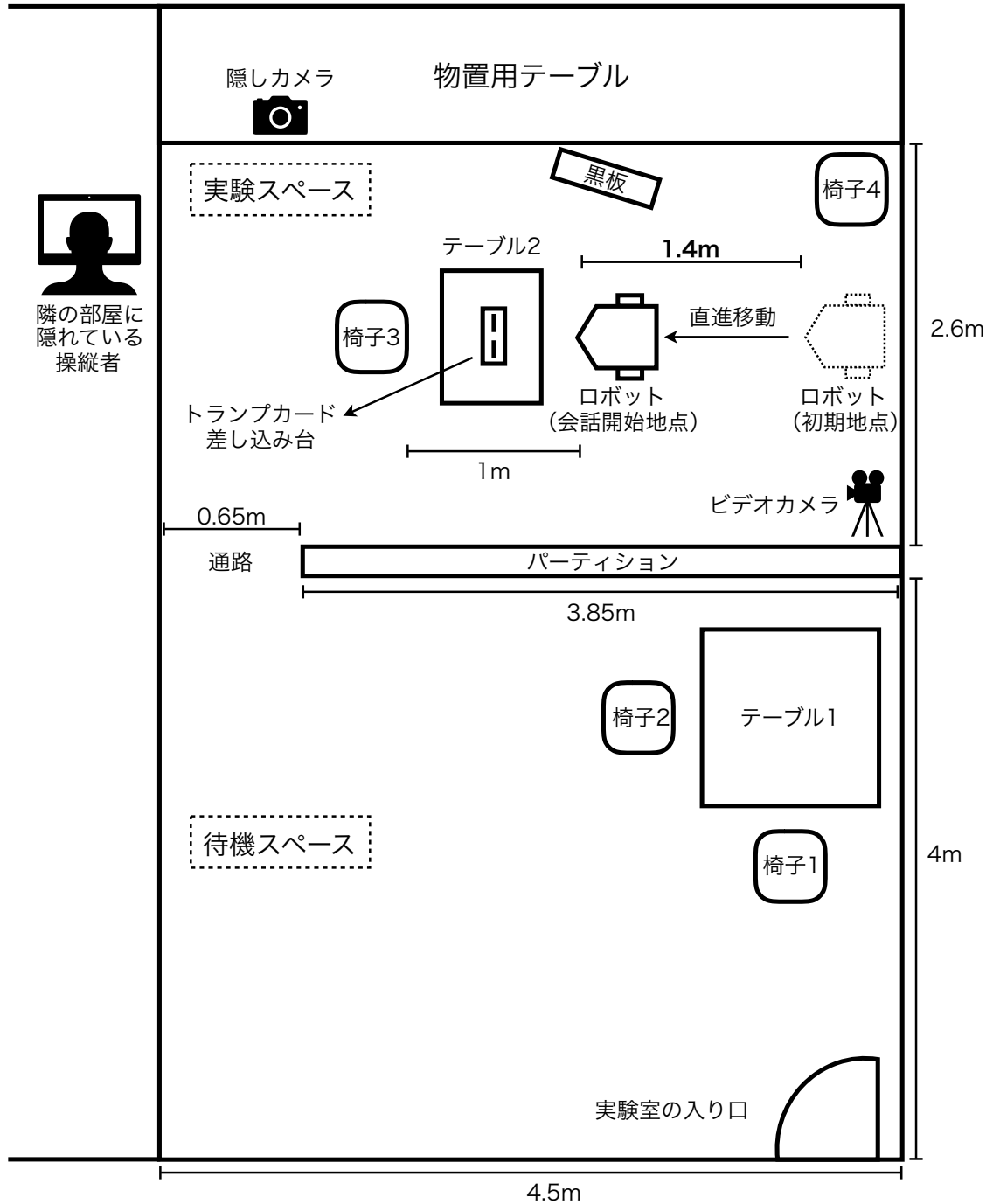


図 4.10: 実験室の見取り図

実験中は、実験担当者が操縦者と連絡を取っている姿を実験協力者に見せない必要があり、また、実験協力者が外部の影響を受けずに一人でロボットとのインタラクションを体験できるようにする為、高さが約1.82mの断熱材で大型のパーティションを作り、そのパーティションで実験室を「実験スペース」と「待機スペース」の2つの領域に分けた。両スペースの間に幅が約0.65mの通路を残した。通路付近でない限り、身長が約1.9m以下の人は立ち上がりの状態でも隣のスペースの様子を確認することができない。実験内容の説明、およびアンケート調査の実施は待機スペースのテーブル1の辺りで行われ、実験協力者とロボットの対面は実験スペースのテーブル2の辺りで行われた。実験協力者は椅子2に座って実験内容の説明などを聞き、アンケートを記入する。また、実験協力者がロボットとの対面を体験する際は、椅子3に座る。それに対し、実験担当者は椅子1に座って実験内容の説明をし、実験協力者がアンケートを記入している時は椅子4に座って記入の完了を待つ。休憩スペースの風景と実験スペースの風景を、図4.11、4.12に示す。



図 4.11: 待機スペースの風景



図 4.12: 実験スペースの風景

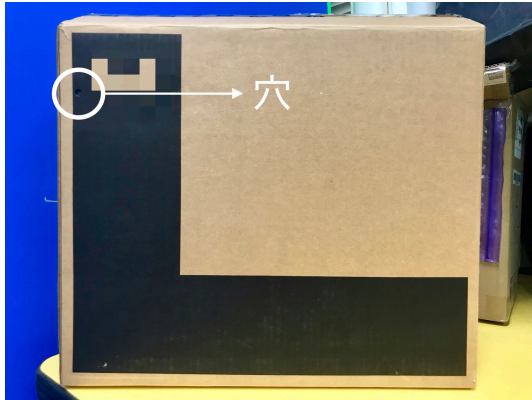


図 4.13: 隠しカメラを収めた紙箱の正面

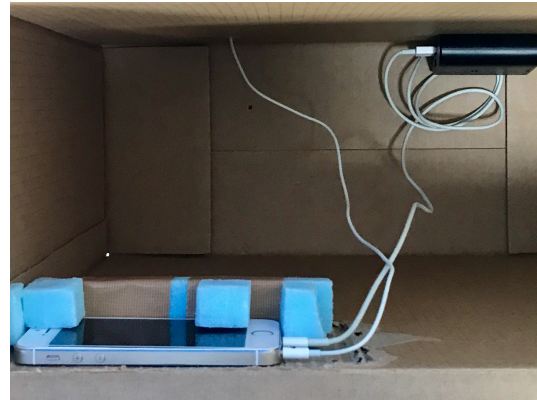


図 4.14: 隠しカメラを収めた紙箱の内部



図 4.15: 操縦者側のタブレット PC に表示される実験スペースの様子

隠しカメラは、テーブル 2 の後ろの左斜めの位置に設置した。安定性などを考慮した結果、本実験では一台のカメラ付きのスマートフォン (iPhone 5s) を隠しカメラとして使用した。スマートフォンの FaceTime を使い、操縦者側のタブレット PC にビデオ通話をかけ、実験スペースの様子を後方から観察できるようにした。スマートフォンを隠す為に、紙箱の上の黒色のエリアに小さな穴を開け、箱の中にスマートフォンを固定する為の治具を作成し、そこに隠しカメラを隠した。スマートフォンを正しい向きで箱の中の治具に入ると、カメラの位置が自動的に穴の位置に当てはまるようにした。また、スマートフォンを箱の

中に入れ、箱を閉じると、箱の外部の音声を綺麗に拾えなくなる為、箱の背面の下にさらに穴を開けた。その穴越しにマイク付きのイヤホンをスマートフォンに繋ぎ、イヤホンとマイクを箱の背後に設置し、操縦者は実験スペースの音がはっきり聞こえるように工夫した。その他、実験中のスマートフォンのバッテリー切れを防ぐ為に、モバイルバッテリーで電力を供給した。使用した紙箱の外観と内部写真を図 4.13、4.14に示す。実験時の操縦者側のタブレット PC 上に表示される隠しカメラが中継する実験スペースの様子を図 4.15に示す。

対面の開始条件として、実験協力者にロボットに挨拶するように教示した。実験協力者が椅子3に座り、ロボットに「こんにちは、パイオ」と挨拶した後、操縦者はそれを聞いてロボットを操縦し、ロボットを初期地点から 0.2m/s の速度で5秒間前進させる。前進距離は約 1.4m である。ロボットは会話開始地点に着いた後、会話を始める。実験時、会話開始地点と実験協力者の間の距離は約 1m である。テーブル2の上には、断熱材で作られたトランプカードの差し込み専用の台が置かれている。実験協力者はカードのシャッフルをした後に、カードをこの台に差し込む。この台を設置する理由としては以下の2つである。まず、実験協力者がシャッフルした後のカード位置の変更を防ぐことができ、操縦者の判断ミスを軽減できる。また、カードを決められた場所に置くことで、実験協力者の個人の好みによるカードの持ち方の違いを無くすことができ、台に差し込まれている状態であれば操縦者側は常に隠しカメラ越しにカードの内容を確認できるメリットも挙げられる。

その他、実験協力者が実験中にロボットに対する返事の言葉を忘れることを防ぐ為に、テーブル2の前の左斜めの位置に小型の黒板を置き、返事用の言葉が印刷されている紙をその上に貼った。

#### 4.3.9 評価方法

「親近感」というものは、主観的なものであり、定量的に計測するのは難しい。そこで本実験は、実験協力者がロボットと対面し、実際にロボットとのインタラクションを体験し、その後、アンケート調査を行うことで、実験協力者がロボットに対してどのような印象を持ったのかを調査・分析する。アンケートの項目は、近年 HRI(Human-Robot-Interaction) 領域で良く使われる Godspeed アンケートシリーズ [34] を用いた。

Godspeed アンケートシリーズは、Christoph Bartneck によって提案され、HRI において人がロボットに対する印象および人の心の状態を評価する為の標準化ツールである。現時点まで、このアンケートシリーズを提案した論文は既に 734 回引用され(2018年12月28日に Google Scholar による調査結果)、近年 HRI 領域でもっとも参照された先行研究の一つである。このアンケートシリーズは5つのアンケートを含めており、以下の5つの尺度で人がロボットに対する印象を評価する。

- Anthropomorphism 擬人化
- Animacy 有生性
- Likeability 好ましさ
- Perceived intelligence 知性
- Perceived safety 安全性

アンケートシリーズの中では、SD 法 [35] を使い、「Unfriendly-Friendly(親しみやすい-親しみにくい)」のような反対の意味を持つ形容詞対項目を使用し、1~5 までのスケールで測定を行う。



本研究では、親近感という言葉に注目している。Godspeed アンケートシリーズの中の Likeability(好ましさ) 尺度は、日本語に訳すると好ましさになるが、好ましいという言葉自体は「感じがよくて、心を引かれる。好感がもてる。」という風に解釈されており [36]、親近感があるという意味に近い為、今回は Godspeed アンケートシリーズの中の三番目の尺度「好ましさ」を用い、ロボットの親近感を評価する。その他の4つの尺度は本研究に直接に関わることはないが、名前を呼ぶことによってどのような影響が出るかを見る為に、このアンケートシリーズ全体を使用した。

Godspeed アンケートシリーズの評価項目の詳細を表 4.8に示す。アンケートの日本語訳は、原著論文 [34] に載っている日本語訳をそのまま使用した。原著論文の中で並べ替えて調査の意図を隠すと良いとのアドバイスがあった為、実際にアンケート用紙を作成する際も順番をランダムに並べ替えた。また、Artificial-Lifelike(人工的-生物的) という評価項目は、「擬人化」と「有生性」の両方に出現している為、混乱を招かない為に重複した項目を一つにまとめた。他に、「安全性」の最後の評価項目の「Quiescent-Surprised(平穏な-驚いた)」について、形容詞対の順番が妥当ではないという指摘がいくつか見かけた為 [37] [38]、その左右の順番を逆にした。実際に実験協力者による印象評価実験で使用した印象評価アンケートは付録 C に示す。得られたアンケートの得点を分析する際は、Godspeed アンケートシリーズの5つの尺度が対応する評価項目の平均得点を使う。各尺度が名前を呼ぶ条件と名前を呼ばない条件における平均得点に対して t 検定を行うことで、有意差の有無を確認する。

表 4.8: Godspeed アンケートシリーズ

アンケート名	項目数	構成項目(英語)	構成項目(日本語)
Anthropomorphism 擬人化	5	Fake - Natural	偽物のような - 自然な
		Machinelike - Humanlike	機械的 - 人間的
		Unconscious - Conscious	意識を持たない - 意識を持っている
		Artificial - Lifelike	人工的 - 生物的
		Moving rigidly - Moving elegant	ぎこちない動き - 洗練された動き
Animacy 有生性	6	Dead - Alive	死んでいる - 生きている
		Stagnant - Lively	活気のない - 生き生きとした
		Mechanical - Organic	機械的な - 有機的な
		Artificial - Lifelike	人工的な - 生物的な
		Inert - Interactive	不活発な - 対話的な
Apathetic - Responsive	無関心な - 反応のある		
Likeability 好ましさ	5	Dislike - Like	嫌い - 好き
		Unfriendly - Friendly	親しみにくい - 親しみやすい
		Unkind - Kind	不親切な - 親切な
		Unpleasant - Pleasant	不愉快な - 愉快的な
		Awfu - Nice	ひどい - 良い
Perceived Intelligence 知性	5	Incompetent - Competent	無能な - 有能な
		Ignorant - Knowledgeable	無知な - 物知りな
		Irresponsible - Responsible	無責任な - 責任のある
		Unintelligent - Intelligent	知的でない - 知的な
		Foolish - Sensible	愚かな - 賢明な
Perceived Safety 安全性	3	Anxious - Relaxed	不安な - 落ち着いた
		Agitated - Calm	動揺している - 冷静な
		Quiescent - Surprised	平穏な - 驚いた

#### 4.3.10 実験手順

実験2は、一人の実験協力者に対して、アンケート調査を合計2回行った。この2つのアンケートは同じものであり、前述の Godspeed アンケートシリーズを使用した。実際に使用したアンケートを付録Bに示す。

本実験の二回目と三回目の対面において、顔認識プログラムを起動し、認識結果を記録していた。それは本研究での顔記憶システムの評価する為に記録したものであり、たとえ実験中に誤認識が生じた場合でもロボットの台詞などには一切影響しない。ロボットは完全に操縦者によって操作される。

他に、本実験は全てのアンケート調査の実施が完了した後、実験実施前に示す「実験参加同意書」に加え、実験の真相が書かれている「同意書(二)」を示し、再度実験協力者の承諾を得る。

実験2の具体的な実施手順を以下のように示す。

- 実験協力者が入室する前
  1. 実験の連絡をする時点で実験協力者の名前の読み方を把握する。
  2. 隠しカメラを設置し、視界と音声を確認する。
  3. 実験用品の設置、プログラムの動作確認をする。
  4. これから来る実験協力者が属する組を確認し、実験の実施順番と実験協力者の名前の読み方を操縦者に教える。
- 実験協力者が入室した後
  1. 実験協力者に実験参加同意書の内容を確認してもらい、同意のサインをしてもらう。
  2. 実験担当者は実験協力者に一回目の対面についての説明をする。
  3. 実験担当者は実験スペースに入り、再度動作確認をする。その後、操縦者に一回目の対面がまもなく開始すると伝えた。
  4. 実験協力者に実験スペースに入ってもらい、一回目の対面が始まる。操縦者が実験協力者からの挨拶の言葉を聞いたら、ロボットを会話開始地点まで前進させ、その後、会話を始める。また、顔画像の採集の時間になったら、操縦者が顔画像採集のプログラムを起動する。
  5. 一回目の対面が終了し、実験協力者が待機スペースに戻る。
  6. 実験担当者は二回目と三回目の対面についての説明資料を実験協力者に渡す。
  7. 実験担当者は実験スペース内に入り、一回目での対面で集めた顔画像を訓練用とテスト用に分け、訓練用の顔画像の中の顔の部分を切り出し、訓練データを生成する。その後、新しく入ってきた訓練データを含め、分類器の訓練を行う。
  8. 実験担当者は待機スペースに戻り、実験協力者に口頭で二回目と三回目の対面についての説明を行う。
  9. 実験担当者は実験スペースに入り、顔認識プログラムを起動する。次に、各プログラムの動作確認をもう一度行う。その後、操縦者に二回目の対面がまもなく開始と教える。操縦者はロボットを初期地点に戻す。

10. 実験協力者に実験スペースに入ってもらい、二回目の対面が始まる。操縦者が実験協力者からの挨拶の言葉を聞いたら、ロボットを会話開始地点まで前進させ、その後、会話を始める。実験協力者がカードをシャッフルし、差し込み台にこれらのカードを設置し終わったら、操縦者はジョーカーの位置を確認し、そのターンの勝敗に応じて必ず当たるように、または外すようにする。
11. 二回目の対面が終了し、実験協力者が待機スペースに戻る。
12. 実験担当者はアンケート1を実験協力者に渡し、記入方法を説明する。実験協力者はアンケート1を記入する。
13. 実験担当者は実験スペースに入り、顔認識プログラムを終了する。その後、実験協力者がアンケートの記入を完了したのを待つ。その間は各プログラムの動作確認をもう一度行う。
14. 実験協力者は実験担当者にアンケートの記入が完了したとの意思を示す。実験担当者は顔認識プログラムを再度起動する。その後、操縦者に三回目の対面がまもなく開始と教える。操縦者はロボットを初期地点に戻す。実験担当者は待機スペースに戻り、アンケート2を回収し、実験協力者を実験スペースに入らせる。
15. 実験協力者に実験スペースに入ってもらい、三回目の対面が始まる。操縦者は実験協力者からの挨拶の言葉を聞いたら、ロボットを会話開始地点まで前進させ、その後、会話を始める。実験協力者がカードをシャッフルし、差し込み台にこれらのカードを設置し終わったら、操縦者はジョーカーの位置を確認し、そのターンの勝敗に応じて必ず当たるように、または外すようにする。
16. 実験担当者はアンケート2を実験協力者に渡す。実験協力者はアンケート2を記入する。
17. 実験担当者は実験スペースに入り、顔認識プログラムを終了する。その後、実験協力者がアンケートの記入を完了したのを待つ。
18. 実験協力者はアンケートの記入が完了したとの意思を示したら、実験担当者は待機スペースに戻り、アンケート2を回収する。
19. 実験担当者は実験協力者に同意書(二)を渡し、実験の真相を打ち明ける。実験協力者のサインをもらった後、実験協力者に謝礼を渡し、実験協力者を帰らせる。

### 4.3.11 実験結果

22名の大学生(男子12名、女子10名、平均年齢22.1歳)を集め、実験2を実施した。その内の11名はX組に属し、残り11名はY組に属する。また、実験はX組の実験協力者1名、Y組の実験協力者1名の順番で交代で行われた。全員分のアンケート1とアンケート2のデータを使い、以下の仮説に対して有意水準5%でt検定を行った。

- 仮説：ロボットが人の名前を呼ぶ条件は人の名前を呼ばない条件より親近感が高い。

アンケートデータの分析結果を図4.16、表4.9に示す。

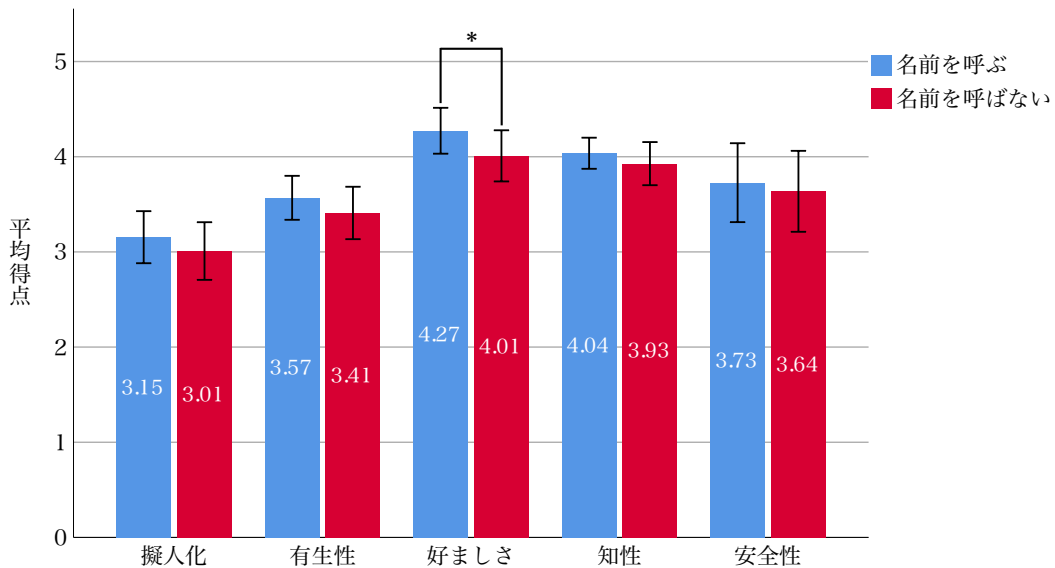


図 4.16: 各尺度の平均得点のグラフ

表 4.9: 各尺度の統計データ

尺度	統計量	名前を呼ぶ	名前を呼ばない
Anthropomorphism 擬人化	平均値	3.155	3.009
	標準偏差	0.617	0.683
	有意確率(両側)	0.195	
Animacy 有生性	平均値	3.568	3.409
	標準偏差	0.521	0.621
	有意確率(両側)	0.226	
Likeability 好ましさ	平均値	4.273	4.009
	標準偏差	0.544	0.606
	有意確率(両側)	0.039	
Perceived Intelligence 知性	平均値	4.036	3.927
	標準偏差	0.368	0.511
	有意確率(両側)	0.117	
Perceived Safety 安全性	平均値	3.727	3.636
	標準偏差	0.935	0.959
	有意確率(両側)	0.701	

#### 4.3.12 考察

データの分析を行ったところ、Godspeed アンケートシリーズの5つの尺度は全て、名前を呼ばない条件より名前を呼ぶ条件の方の平均得点が高いことが分かった。ただし、「好ましさ」の尺度においては有意差 ( $p < 0.05$ ) が見られ、他の尺度においては有意差が見られなかった為、仮説の「ロボットが人の名前を呼ぶ条件は人の名前を呼ばない条件より親近感が高い」が認めれた。この結果を見ると、名前を呼ぶ条件は名前を呼ばない条件よりロボットの親近感が高いと解釈することができる。従って、本研究の提案手法の顔記憶システムによる親近感の向上の効果は期待できる。以下に考察を示す。

5つの尺度とも、名前を呼ぶ条件のほうの平均得点が高いが、名前を呼ばない条件に比べると差が小さいことが分かる。「好ましさ」の尺度において有意差が認めれたが、その差も大きいとは言えない。これらのことに対して、以下のことが原因として考えられる。

まず、実験協力者の中で、二回目と三回目の対面の条件の違いに明確に気づく人が少なかった。実験協力者はロボットが本当に自分の表情を分析できると強く信じ、ババ抜きゲームに集中し過ぎたせいでロボットが実験協力者の名前を呼ぶ／呼ばないという違いがあることに気づかない現象が見られた。これにより、アンケートで点数を付ける際も、両条件の違いについて良く分からない為、アンケート1とアンケート2に近い点数を付けた。また、ロボットがかなり高い確率でジョーカーの位置に当たるように設定した為、実験協力者は「このロボットは凄い」と感じ、ロボットに対する印象が良く、点数が全体的に高くなり、差が付きにくい状態になった。それに加え、ロボット自身の発話内容は、丁寧で謙遜なイメージに設定され、合成されたロボットの声も可愛いく聞こえる為、ロボットに対する印象も悪くなりにくく、全体的な点数は高い方向に持ち上げられ、差が付きにくくなった。

今回の実験で用いたロボットの顔インタフェースは特別な処理などは一切してなく、実験協力者に見せるものだけという風に扱った為、アンケートの自由記述を見ると顔インタフェースに対する考慮が足りないことが分かった。顔インタフェース上の目はまばたきすることができるが、実験協力者の視線に合わせることはできない。それに対して無機質的だと記述した実験協力者は数人がいた。また、顔インタフェース上の目の後ろに炎が表示される為、炎を表示する意味が分からないという記述もあった。以上のことは、実験協力者の注意を無駄に関係のないところに向けさせた為、2条件間の平均得点の差が小さい原因の一つとも考えられる。

明確に2条件間の差に気づいた人は多くなかったが、数人は名前を呼ばれる条件において無意識的に親しみやすく感じたと述べた。また、2条件間の差に気づいたと明言した人の中は1人を除き、全員名前を呼ぶ条件のほうの「好ましさ」の評価項目に高い点数を付けた。以上のことから、名前を呼ぶことは親近感向上の役に立つと考えられる。その他、実験2の二回目と三回目の対面では、名前を呼ぶかどうかという点だけが違い、t検定の結果もちょうど「好ましさ」の尺度だけに有意差が見られた為、名前を呼ぶことは「好ましさ」以外の4つの尺度に対する影響が薄いとも考えられる。

## 4.4 実験3

### 4.4.1 概要

実験3では、実装した顔検出機能と顔認識機能の実行速度を調べた。また、これまでに集めた顔画像を用いて画像の前処理と分類器の訓練を行い、その処理時間、およびテストデータに対する訓練済みの分類器の正解率について調査した。

実験1を実施する前、既に7人分の顔画像を取得した。その後の実験2でまた28人分の顔画像を取得した（予備実験で6人分を取得）。1人の200枚の顔画像を均等の2部分に分け、訓練とテストに使用する。まずは7人分の顔画像に基づいて、数理AIMモデル適用ありと数理AIMモデル適用なしの2条件で画像の前処理と分類器の訓練を行い、処理時間を記録する。その後、最初の7人分を含む35分の顔画像でまた同じ作業を繰り返す。数理AIMモデルを適用する場合は、システムに刺激情報を与えない。本実験においても画像前処理プロセスと分類器訓練プロセスが使用するCPUコア数の最大値を制限した。画像前処理プロセスの2コアまで、分類器訓練プロセスは1コアまでと制限した。

実験3で使用したソフトウェアとハードウェアは、実験1とは同じものである為、説明を省略する。

### 4.4.2 実験結果

カメラの前に一個の顔のみを提示し、その場合の顔検出と顔認識の実行速度は表4.10に示す。ここで示す実験データは、100枚の画像を処理する所要時間から計算されたものである。

表 4.10: 顔検出と顔認識の実行速度

	顔検出のみ	顔検出+顔認識
実行速度	6.92[frame/sec]	4.77[frame/sec]

7人分の顔画像を用いた実験の結果を表4.11に示し、また、35人分の顔画像を用いた実験の結果を表4.12に示す。各過程は3回実行し、その平均時間を記録した。

表 4.11: 7人分の顔画像を用いた処理時間と分類器の正解率

		数理AIMモデル適用あり	数理AIMモデル適用無し
画像前処理		341.12[sec]	204.72[sec]
分類器訓練	FaceNetベクトル変換	158.58[sec]	70.02[sec]
	SVM分類器訓練	0.35[sec]	0.43[sec]
テストデータに対する正解率		100.0%	

表 4.12: 35 人分の顔画像を用いた処理時間と分類器の正解率

		数理AIMモデル適用あり	数理AIMモデル適用無し
画像前処理		1589.43[sec]	1040.05[sec]
分類器訓練	FaceNetベクトル変換	791.72[sec]	356.07[sec]
	SVM分類器訓練	7.88[sec]	7.23[sec]
テストデータに対する正解率		100.0%	

#### 4.4.3 考察

実験で使用したノートパソコンは古いものだが、このような条件においても顔検出と顔認識は満足できるような速度で実行できた。また、数理 AIM モデル適用無し条件において、35 人からの 3500 枚の画像に対する画像前処理は時間がかかるよう思えるが、7 人の場合と比較すると、1 人からの 100 枚の顔画像に対して処理時間は 30 秒ほどしかかからないことが分かる。既に前処理された画像は再度処理する必要がない為、システムに新しく人を追加する場合でも処理時間の心配はない。他に、分類器訓練の結果データを見ると、SVM 分類器訓練よりも顔検出された画像を FaceNet ベクトルへ変換することのほうが時間がかかることが分かった。現状では分類器訓練を行う度、FaceNet ベクトル変換が最初から実行するようになっているが、変換結果を保存し、同じ画像に対して二度目以降の変換を省くようにすれば、処理時間の短縮が期待できる。しかし、SVM 分類器の訓練時間は画像の数の増加に比例して増加するのではない為、分類するクラスの数が増えたと訓練時間が長くなる恐れがある。それを改善するのに SVM 分類器以外の分類機を検討する必要がある。また、正解率に関しては、7 人の場合と 35 人の場合とも、分類器はテストデータに対して 100.0%の正解率を達成できたが、実際に運用する際は外部環境の変化により精度が落ちることも考えうる為、トラッキングなどの方法と組み合わせて安定性を上げる必要がある。

## 第5章 結論

本研究は、移動ロボットのCPUとバッテリーの有効利用、およびロボットの親近感の向上を目的とし、数理AIMモデルと顔記憶システムを組み合わせた手法を提案した。人間の脳の意識状態を模擬する数理AIMモデルをロボットに適用することで、ロボットは取り付けられている外界センサーで外部の状況を把握し、その時の状況に応じて意識状態を遷移させることができる。これにより、ロボットは周囲に人がいない時に自動的に睡眠状態に遷移し、不要な処理の処理頻度を下げることで、CPUとバッテリーの無駄な消費を省く。また、ロボットが睡眠状態になった場合は、その時間を活用する為に、リアルタイムでは処理しきれない重い処理をさせる。人を特定する為に、顔認識機能を有する顔記憶システムを実装した。精度の良い顔認識を実現する為の画像の前処理および分類器の訓練が必要であり、これらの重い処理はロボットが睡眠の時に行わせ、睡眠による空き時間を使用することで、CPUとバッテリーをさらに有効に利用する。顔認識は人の特定と人の名前を呼ぶことに使用し、名前を呼ぶことでロボットの親近感を向上させる。

提案手法の効果を検証する為に実験を行った。数理AIMモデルと顔記憶システムを組み合わせ、一定時間内のCPUの利用率とバッテリーの電圧を計測し、その変化を確認、比較することで、提案手法がCPUとバッテリーの有効利用に効果があると分かった。また、名前を呼ぶことの効果を検証する為に、実験協力者による印象評価実験を行った。実験では、実験協力者に実際にロボットとの対面を体験してもらった。名前を呼ぶ条件と名前を呼ばない条件を設け、該当する条件に応じて実験中のカードゲームにおいてロボットに実験協力者の名前を含む発話と含まない発話を行わせた。実験協力者はこの2条件でのロボットとの対面を体験し、アンケートでロボットへの印象を1~5の五段階で評価した。22名の実験協力者を集め、実験を行い、アンケートに対してt検定を行った結果、親近感に関する尺度では名前を呼ぶ条件は名前を呼ばない条件より得点が有意に高いことが分かった。これにより、ロボットが人の名前を覚え、会話の時に名前を呼ぶことは親近感の向上に役立つと分かった。

しかし、本研究はまだいくつかの課題が残っている。1つ目は、多数の日本人の苗字を高精度で認識できる音声認識を実現していないことである。音声認識で名前情報を取得する際にも誤認識が発生してしまった場合、その後の名前の呼びかけにも悪影響をもたらす為、印象を良くするどころか、逆に人に悪い印象を与えてしまうリスクも考えられる。従って今後は苗字の音声認識について他の手法を検討する必要がある。また、実装したシステムでは、評価実験環境下においては精度良く顔の検出と認識が可能だが、多数の人がロボットの周囲にいる場合や、実際に運用を目指して商店街などの複雑な環境下ではまだ他の工夫が必要である。現時点では、顔認識する際にトラッキング機能を追加することで認識結果をより安定にすることを想定している。他に、分類器の訓練時間が長いという問題も挙げられ、今後はFaceNetベクトル変換を行う際に変換結果を保存して再利用することで時間短縮を図りたい。最後に、実験協力者による印象評価実験において、親近感に関する尺度で有意差が見られたが、差が小さく効果が薄いとの問題があった。アンケートの自由記述の内容および実験協力者に対するインタビューから、実験協力者達がカードゲームに集中しすぎたせいでロボットの発話内容の変化に気づかないのが原因だと考えられる。



名前を呼ぶことの効果をより明らかに証明する為に、インタラクションの内容を他のより適切なゲームなどに変更して再度実験する必要がある。今後は、以上の問題点に着目しシステムと実験方針を改善していきたい。

## 謝辞

本研究を進めるにあたり、研究のみにとどまらず、日常生活、就職活動などにおいても数多くの貴重なご意見を下さった三河正彦准教授に心より深く感謝いたします。当初何も分からなかった私を研究生の時代から今まで、見捨てず時間をかけて優しく指導し続けて頂き、本当にありがとうございました。また、合同ゼミなどで、藤澤誠助教から適切な指摘をたくさん頂いたことにも深く感謝いたします。そして、普段の研究室生活でご協力頂いた三河研究室の先輩／後輩の皆様、特に実験協力者による印象評価実験で貴重な時間を多く割いて協力して下さいました佐藤守任さん、授業の課題や論文の執筆に付き合っ共に乗り越えて頂いた同期の呂佳芸さん、私の M1 の頃に良く細かいことの相談に乗って下さった既卒の大館雄太先輩、後は多忙な時期に実験協力者による印象評価実験に参加して下さいました参加者の皆様、本当にありがとうございました。

## 参考文献

- [1] P. DharanyaDevi, S. Poonguzhali, T. Sathiya, G. Yamini, P. Sujatha and V. Narasimhulu, "Survey on Multimedia Operating Systems," *International Journal of Computer Science & Emerging Technologies(IJCSET)*. 1(2), pages 100–108, 2010.
- [2] Yongguo Mei, Yung-Hsiang Lu, Y Charlie Hu, and CS George Lee, "Energy-efficient motion planning for mobile robots," *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, volume 5, pages 4344–4349, 2004.
- [3] M. Mikawa, M. Yoshikawa, T. Tsujimura, and Kazuyo Tanaka, "Intelligent perceptual information parallel processing system controlled by mathematical aim model," In *2007 7th IEEE-RAS International Conference on Humanoid Robots*, pages 398–403, Nov 2007.
- [4] 神田崇行, 宮下敬宏, 長田拓, 配川有二, 石黒浩, "人口ロボット相互作用における人型ロボットの外見の影響", *日本ロボット学会誌*, Vol.24, No.4, pp.497–505, 2006.
- [5] 松元崇裕, 瀬古俊一, 青木良輔, 宮田章裕, 渡部智樹, 山田智広, "人の情動体験を向上させる感情表出口ロボット", *電子情報通信学会論文誌 A*, 99(1):45–55, 2016.
- [6] 神田崇行, 石黒浩, 石田亨, "人間-ロボット間相互作用にかかわる心理学的評価", *19(3):362–371*, 2001.
- [7] Emily Wang, Constantine Lignos, Ashish Vatsal, and Brian Scassellati, "Effects of head movement on perceptions of humanoid robot behavior," In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 180–185. ACM, 2006.
- [8] Chien-Ming Huang, Takamasa Iio, Satoru Satake, and Takayuki Kanda, "Modeling and controlling friendliness for an interactive museum robot," In *Robotics: Science and Systems*, 2014.
- [9] Jamieson Schulte, Charles Rosenberg, and Sebastian Thrun, "Spontaneous, short-term interaction with mobile robots," In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, volume 1, pages 658–663. IEEE, 1999.
- [10] Elena Torta, Jim van Heumen, Raymond H Cuijpers, and James F Juola, "How can a robot attract the attention of its human partner? a comparative study over different modalities for attracting attention," In *International Conference on Social Robotics*, pages 288–297. Springer, 2012.
- [11] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro, "A practical experiment with interactive humanoid robots in a human society," In *Procs. 3rd IEEE Int. Conf. Humanoid Robots*, 2003.

- [12] Takayuki Kanda, Rumi Sato, Naoki Saiwaki, and Hiroshi Ishiguro, "A two-month field trial in an elementary school for long-term human–robot interaction," *IEEE Transactions on robotics*, 23(5):962–971, 2007.
- [13] Noriaki Mitsunaga, Takahiro Miyashita, Hiroshi Ishiguro, Kiyoshi Kogure, and Norihiro Hagita, "Robovie-iv: A communication robot interacting with people daily in an office," In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 5066–5072. IEEE, 2006.
- [14] Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita, "An affective guide robot in a shopping mall," In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, pages 173–180. ACM, 2009.
- [15] J. Allan Hobson, "The Dream Drugstore," The MIT Press, 2001.
- [16] J. Allan Hobson, "Rem sleep and dreaming: towards a theory of protoconsciousness," *Nature Reviews Neuroscience*, 10(11):803, 2009.
- [17] OpenCV team, "Opencv library," <https://opencv.org/>. [Online; Accessed 25-Aug-2018].
- [18] Davis E. King, "dlibC++Library," [dlib.net](http://dlib.net). [Online; Accessed 10-Jan-2019].
- [19] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, 23(10):1499–1503, 2016.
- [20] Florian Schroff, Dmitry Kalenichenko, and James Philbin, "Facenet: A unified embedding for face recognition and clustering," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823, 2015.
- [21] Gary B Huang, Marwan Mattar, Tamara Berg, and Eric Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," In *Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition*, 2008.
- [22] David Sandberg, "davidsandberg/facenet: Face recognition using tensorflow," <https://github.com/davidsandberg/facenet>. [Online; Accessed 25-Aug-2018].
- [23] Qiong Cao, Li Shen, Weidi Xie, Omkar M Parkhi, and Andrew Zisserman, "Vggface2: A dataset for recognising faces across pose and age," In *Automatic Face & Gesture Recognition (FG 2018)*, 2018 13th IEEE International Conference on, pages 67-74. IEEE, 2018.
- [24] scikit-learn developers, "scikit-learn: machine learning in python — scikit-learn 0.20.2 documentation," <https://scikit-learn.org/stable/>. [Online; Accessed 25-Dec-2018].
- [25] Akinobu Lee, Tatsuya Kawahara, and Kiyohiro Shikano, "Julius—an open source real-time large vocabulary recognition engine," 2001.
- [26] John F Kelley, "An empirical methodology for writing user-friendly natural language computer applications," In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 193–196. ACM, 1983.

- [27] K Oura, S Sako, and K Tokuda, "Japanese text-to-speech synthesis system: Open jtalk," In Proc. ASJ, pages 343–344, 2010.
- [28] ROS Contributors, "Ros.org | powering the world's robots," <http://www.ros.org/>. [Online; Accessed 25-Aug-2018].
- [29] Giampaolo Rodola, "psutil · pypi," <https://pypi.org/project/psutil/>. [Online; Accessed 25-Dec-2018].
- [30] Adept MobileRobots Inc, "Aria - adept mobilerobots," <http://robots.mobilerobots.com/wiki/ARIA>. [Online; Accessed 25-Agu-2018].
- [31] Google Inc, "Tensorflow - an open source software library for machine intelligence," <https://www.tensorflow.org/>. [Online; Accessed 25-Dec-2018].
- [32] Srećko Jurić-Kavelj, "Rosaria - ros wiki," <http://wiki.ros.org/ROSARIA>. [Online; Accessed 25-Dec-2018].
- [33] Akinobu Lee, Keiichiro Oura, and Keiichi Tokuda, "Mmdagent—a fully open-source toolkit for voice interaction systems," In Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, pages 8382–8385. IEEE, 2013.
- [34] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International journal of social robotics*, 1(1):71–81, 2009.
- [35] Charles E Osgood, "The nature and measurement of meaning," *Psychological bulletin*, 49(3):197, 1952.
- [36] Weblio 辞書, "好ましい (このましい) とは - 好ましいの読み方", <https://www.weblio.jp/content/%E5%A5%BD%E3%81%BE%E3%81%97%E3%81%84>. [Online; Accessed 25-Dec-2018].
- [37] Daphne Eleonora Karreman, Gilberto S epulveda Bradford, Elisabeth M.A.G. van Dijk, M. Lohse, and Vanessa Evers, "What happens when a robot favors someone?: How a tour guide robot uses gaze behavior to address multiple persons while storytelling about art," In Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction. IEEE Press, 2013.
- [38] Guido Schillaci, Saša Bodiroža, and Verena Vanessa Hafner, "Evaluating the effect of saliency detection and attention manipulation in human-robot interaction," *International Journal of Social Robotics*, 5(1):139–152, 2013.

## 付録A ロボットの台詞

### 一回目の対面

- 被験者 : こんにちは、パイオ。
- ロボット : こんにちは。やっと人が来ましたね。先ほどまで誰もいなかったので、すごく暇でした。来てくださって嬉しいです。
- では、改めまして自己紹介します。私、移動ロボットのパイオと申します。本日は実験にご参加いただき、ありがとうございます。実験担当者から既に説明を聞いたと思いますが、念の為もう一度説明します。
- 本日の実験は、私と三回の対面を体験していただきます。一回目の対面では、あなたの名前情報と顔写真を採集します。これはこの後のゲームをする為の準備です。それでは、名前情報と顔写真の採集を始めたいと思います。準備はよろしいですか？
- 被験者 : はい。
- ロボット : まず、あなたのお名前を伺います。苗字を教えてください。
- 被験者 : xx (自分の名前を答える)
- ロボット : xx さんですね。分かりました。
- 次に、顔写真を取ります。終了の合図があるまで、頭を上下左右にゆっくりと動かし続けてください。全部で十秒ほどかかります。では、始めてもよろしいですか？
- 被験者 : はい。
- ロボット : では、始めます。ゆっくりと頭を動かしてください。
- (顔写真を採集する)
- ロボット : 顔写真の採集はここで終了です。ご協力ありがとうございました。では次の対面が始まるまで、休憩スペースに戻ってお待ちください。

## 二、三回目の対面（名前を呼ぶ条件）

被験者： こんにちは、パイオ。

ロボット： こんにちは。今回は、xxさんとババ抜きゲームをします。ゲームのやり方について少し説明します。

まずは、テーブルの上にあるカードを取って、テーブルの下でシャッフルしてください。その後、カードを水色の台に差し込んでください。準備ができたなら、私に、「準備完了です」と教えてください。

では、xxさん、カードを取ってシャッフルしてください。

※（被験者はカードを取り、シャッフルした後にカードを台に差し込む）

被験者： 準備完了です。

ロボット： 私の目を見つめてください。分析を始めます。

（5秒間待つ）

ロボット： 分かりました。ジョーカーは、左/右側にあります。正解ですか？

被験者： （4ターン目以外）正解です。

（4ターン目）間違いです。

ロボット： （4ターン目以外）やはりですね。当たって嬉しいです。xxさん、ご協力ありがとうございます。

（4ターン目）残念ですね。外してごめんなさい。xxさんが頑張ってくださったのに。

ロボット： （1、2、3、5ターン目）次も頑張りましょう。では、カードを取ってシャッフルしてください。（※のところに戻り、次のターンに進む）

（4ターン目）次はミスしないように頑張ります。では、カードを取ってシャッフルしてください。（※のところに戻り、次のターンに進む）

（6ターン目）（次に進む）

ロボット： ゲームは以上です。では、結果を発表します。六回のゲームの中、合計。五回当たりました。なかなかいい成績が得られました。xxさんのおかげです。本当にありがとうございました。

今回の対面はここで終了となります。xxさん、ゲームに付き合っていただいて。すごく嬉しかったです。感謝いたします。では、休憩スペースにお戻りください。

## 二、三回目の対面（名前を呼ばない条件）

被験者： こんにちは、パイオ。

ロボット： こんにちは。今回は、あなたとババ抜きゲームをします。ゲームのやり方について少し説明します。

まずは、テーブルの上にあるカードを取って、テーブルの下でシャッフルしてください。その後、カードを水色の台に差し込んでください。準備ができたなら、私に、「準備完了です」と教えてください。

では、カードを取ってシャッフルしてください。

※（被験者はカードを取り、シャッフルした後にカードを台に差し込む）

被験者： 準備完了です。

ロボット： 私の目を見つめてください。分析を始めます。

（5 秒間待つ）

ロボット： 分かりました。ジョーカーは、左/右側にあります。正解ですか？

被験者： （4 ターン目以外）正解です。

（4 ターン目）間違いです。

ロボット： （4 ターン目以外）やはりですね。当たって嬉しいです。ご協力ありがとうございます。

（4 ターン目）残念ですね。外してごめんなさい。頑張ってくださいなのに。

ロボット： （1、2、3、5 ターン目）次も頑張りましょう。では、カードを取ってシャッフルしてください。（※のところに戻り、次のターンに進む）

（4 ターン目）次はミスしないように頑張ります。では、カードを取ってシャッフルしてください。（※のところに戻り、次のターンに進む）

（6 ターン目）（次に進む）

ロボット： ゲームは以上です。では、結果を発表します。六回のゲームの中、合計。五回当たりました。なかなかいい成績が得られました。あなたのおかげです。本当にありがとうございました。

今回の対面はここで終了となります。ゲームに付き合ってください。すごく嬉しかったです。感謝いたします。では、休憩スペースにお戻りください。



## 付録B アンケート

記入日 \_\_\_\_年\_\_月\_\_日

参加者番号		性別		年齢	
所属	学群・研究科		学類・専攻		年次

当てはまる点数を○で囲って下さい。  
 点数が低いほど、左側に賛成。高いほど、右側に賛成。

以下のスケールに基づいてこのロボットへの印象を評価して下さい

Q1	人工的な	1	2	3	4	5	生物的な
Q2	不活発な	1	2	3	4	5	対話的な
Q3	活気のない	1	2	3	4	5	生き生きとした
Q4	無責任な	1	2	3	4	5	責任のある
Q5	意識を持たない	1	2	3	4	5	意識を持っている

以下のスケールに基づいてこのロボットへの印象を評価して下さい

Q6	不愉快な	1	2	3	4	5	愉快的な
Q7	親しみにくい	1	2	3	4	5	親しみやすい
Q8	無関心な	1	2	3	4	5	反応のある
Q9	機械的な	1	2	3	4	5	有機的な
Q10	偽物のような	1	2	3	4	5	自然な

以下のスケールに基づいてこのロボットへの印象を評価してください

Q11	愚かな	1	2	3	4	5	賢明な
Q12	ひどい	1	2	3	4	5	良い
Q13	死んでいる	1	2	3	4	5	生きている
Q14	嫌い	1	2	3	4	5	好き
Q15	ぎこちない動き	1	2	3	4	5	洗練された動き

以下のスケールに基づいてこのロボットへの印象を評価してください

Q16	無知な	1	2	3	4	5	物知りな
Q17	不親切な	1	2	3	4	5	親切な
Q18	機械的	1	2	3	4	5	人間的
Q19	無能な	1	2	3	4	5	有能な
Q20	知的でない	1	2	3	4	5	知的な

以下のスケールに基づいてあなたの心の状態を評価してください

Q21	不安な	1	2	3	4	5	落ち着いた
Q22	動揺している	1	2	3	4	5	冷静な
Q23	驚いた	1	2	3	4	5	平穏な

Q24.ロボットに対する感想等を自由にご記入ください

Q25.実験で気になる所があれば自由にご記入ください

最後は、記入内容のチェックをお願いいたします。

ご協力ありがとうございます。