

早稲田大学審査学位論文
博士（人間科学）
概要書

ヒト型エージェントに対する否定的感情表出過程を表現する定性的脳機能モデルの構築
Qualitative Brain Function Model Explaining the Mechanism of
Negative Emotion towards Humanlike Agent

2020年1月

早稲田大学大学院 人間科学研究科
田和辻 可昌
TAWATSUJI, Yoshimasa

研究指導担当教員：松居 辰則 教授

近年産業技術の進展により、様々な形でロボットやアバターなどのエージェントとインタラクションをとる機会が増えてきた。このような中で、コンピュータの振る舞いを人間にとってその意図の可読性が高くなるように実現させることは、両者間の円滑なコミュニケーションを考える上で重要な課題である。人間の擬人化作用を利用し、エージェントの外的表現を人に近づける試みは重要なアプローチであると考えられる。ところが、このような人間に対する類似性を高める試みがある一方で「不気味の谷」と呼ばれる現象が問題視されている。本研究では、このような不気味の谷がどのようなメカニズムで形成されるのかを理解することを志向して、脳を構成する各領域（脳部位）の機能的側面をモデル化し、それらを繋ぎ合わせることで脳全体の振る舞いを表現する定性的脳機能モデルの構築を試みた。本論文の第II部では、ヒト型エージェントの静的側面（静止画）に対する人間の知覚処理の特性とその神経基盤に関するモデル構築、第III部では、ヒト型エージェントの動的側面（動画）に対する人間の知覚処理の特性とその神経基盤に関するモデル構築を行った。続く第IV部では、これまでのモデルを体系的かつ統一的に記述する枠組みとしてデバイスオンロジーに着目し、各モデルを統一的な観点から位置づけるためのアプローチの有効性について論じた。

第II部では、ヒト型エージェントの「顔」の静的側面、すなわち、顔の形態的特徴が人間の印象形成に与える影響を対象とした。第四章では、ヒト型エージェントと人間の顔画像をそれぞれ刺激として用いて、顔刺激に対する情報処理を観察中の視線を計測することで抽出することを試みた。顔刺激の提示時間長の観点から顔刺激の各顔特徴領域に対する視線停留時間の長さを分析した。この結果、人間とは異なると判断されたヒト型エージェントの右目に対する停留時間は、刺激が提示される時間が長くなるにしたがって、人間の右目に対する停留時間よりも定性的に長くなる傾向があることが示唆された。ついで第五章ではこのような情報処理過程を説明するための定性的脳機能モデルの構築を試みた。このモデルでは、ヒト型エージェントに対する知覚情報処理は観察対象に対する雑多な情報処理過程と観察対象に対する精密な情報処理過程からなり、この二つの情報処理の齟齬の結果否定的な感情が形成されると考えた。モデルは、扁桃体や線条体などによる快-不快評価機構、大脳皮質による高次認知処理、海馬による記憶や評価の一貫性評価の機構からなり、第四章で見た視線遷移を説明する脳機能モデルの構築を試みた。シミュレーションの結果、形態的特徴が人間から離れるに従い、評価が振動することが示唆された。この評価の振動が不気味さの様相を表現していると考えられた。

第III部では、人間の「顔」の中でもさらに動的側面、すなわち顔の表情動作速度に着目した。第六章では、ヒト型エージェントにおける表情動作速度が、人間の典型的な表情動作速度から逸脱するにしたがって否定的な印象を形成させる、と仮説を立て、心理

学的実験を実施した。人間の自発的な表情表出を行う際の速度を典型的な動作速度と定義し、典型的な動作速度および非典型的な動作速度（典型的動作速度よりも速いあるいは遅い速度で表出される）の笑顔・怒り顔をヒト型エージェントに実装した。実験参加者にそれぞれの表情動作を観察してもらい、違和感、快-不快評価、表情強度、表情速度の適度について7段階で評価を求めた。この結果、外見で否定的な評価がなされたエージェントについては、表出速度が速い場合の非典型的な笑顔は、典型的な速度と比較して否定的な印象を与えることが示唆された。また、外見の評価あるいは表出する感情に関わらず、遅い速度での表情表出は観察者に否定的な印象を与えることが示唆された。さらに、重要な点として、表情表出速度を遅くすると人間はその背後に文脈を読み取り、何らかの理由があってその表情を表出していると感じることが示唆された。これはヒト型エージェントの表情動作であっても社会的な存在として認知させることができるということを示唆しており、学習教育文脈など非言語コミュニケーションが必要な場面においても多様な表現をヒト型エージェントが表情で伝達できることを示唆している。ついで第七章では、人間の表情認知が動作予測の情報処理と動作知覚の情報処理の二つの情報処理によって実現されていると仮説を立て、動作予測を行う神経基盤として小脳を検討した。このモデルでは、初期動作が受容されるとその情報が小脳に連絡され、小脳の内部モデルによってそのあとの動作系列が計算され予測されることがモデル化されている。シミュレーションの結果、典型的な表情動作を行うヒト型エージェントについては否定的な印象は形成されないが、非典型的な動作を与えた場合は、否定的な評価がなされることが示唆された。

第IV部では、第II部・第III部で構築された定性的脳機能モデルの記述観点が統一されていないという点を受け、それらを体系的に記述する枠組みとして拡張デバイスオントロジーの枠組みに着目した。拡張デバイスオントロジーは人工物における機能を体系的に記述するための概念を提供するが、本枠組みが生物器官としての脳における機能に対しても適用可能であることを、I-goalsとNI-goalsの観点から検討した。さらに、この記述の一環として、神経科学的知識に基づく神経細胞群間の活性伝播に関するオントロジーをロール概念に基づいて記述した。この結果、活性伝播に関するオントロジーはデバイスオントロジーに基づいた対象の捉え方ではないことが明らかとなった。

本研究の成果は、人間の否定的な感情を形成する神経情報処理メカニズムを説明する定性的脳機能モデルの構築と、人間にとってその作用機序を可読な形で表現するための記述体系の足掛かりとなる記述観点の提供である。今後はこれまでに構築した脳機能モデルを体系的に記述するために必要なデバイスオントロジー的視点について検討し、脳機能を体系的に記述可能な枠組みの詳細化を行うことが挙げられる。