

論文

不気味の谷の陥落

伊藤京平*

1. はじめに

1970年、工学者の森政弘が企業広報誌『Energy』に寄稿したエッセイのタイトルは「不気味の谷」であった。当時は第一次人工知能ブームの末期とその後訪れる冬の時代の境界に位置するため、ロボット工学に対する期待と失望の入り交じった複雑な風潮の中で執筆されたものと推測される。不気味の谷は、ロボットの外観を人に近づけるにつれて、そのロボットに対する親和感は上昇していくが、類似度が一定の閾値を超えると親和感が急激に減少し、その後また上昇するという仮説を表す用語である。森によれば、不気味の谷底にあるのは死人であり、特にそれが動いているときの親和感が最も低くなる。奇しくもジョージ・A・ロメロ監督の『ナイト・オブ・ザ・リビングデッド』が1968年に公開されているが、動く死人とは今で言えばゾンビといったところであろうか。森のエッセイは、なまじ人間に近いロボットを作ろうとすると受け入れ難いものが出来上がることに對して警鐘を鳴らすと同時に、われわれの素朴な疑問——身近な存在の差異を顕著に見いだすことができるのはなぜか——を改めて提示している。

ロボットとか人形とか義手とかいうものの場合、人間に対する類似度がかなり高い状態にある時には、一歩ふみはずせば、不気味の谷に急転直下おち込んでしまうのである（森1970: 35）。

森の言説は確かに「それらしい」。しかし、なぜそのような現象が起るのか、あるいは起らないのか、ということは体系的に検討されてこなかったと言っていいだろう。本稿は不気味の谷仮説とその実証的研究の批判を通じて、不気味さとは何かという問いに議論を展開する。これにより、不気味さの概念的な揺らぎを明らかにし、その揺らぐ情動を生み出すものとの共生を論じる。

2. 不気味なもの／不気味の谷

不気味の谷に目を向ける前に、ここでは不気味なものに関する議論を手短に取りあげておきたい。ドイツの精神科医エルンスト・イエンチュは、1906年のエッセイ「不気味なものへの心理学に向けて」(*Zur Psychologie Des Unheimlichen*)の中で、不気味なものとは心理的な不確実性をもたらす対象であると論じた。イエンチュの学説が日の目を浴びたのは著名な精神分析学者・精神科医ジークムント・フロイトによるところが大きい。フロイトは1919年のエッセイ「不気味なもの」(*Das Unheimliche*)でイエンチュの学説を覆すものとしての不気味さを定義した。フロイトのいう不気味さの妙味は、それが親しみのあるものの中に見いだされるという点にある。ドイツ語の不気味さ(*unheimlich*)は家でない(*un-heim*)という意味を含んでいるが、フロイトは逆説的に、親しみの象徴たる潜在的なわが家が立ち顕れる、そのような瞬間にこそ不気味さの本質があることを指摘しているのだ。ここでフロイト自身の経験に基づく記述を引用したい。

ある夏の暑い午後、わたしはイタリアの小さな町の人気のない見知らぬ通りを気ままに散歩していて、すぐに

キーワード：不気味の谷、人のようなもの、共生

*立命館大学大学院先端総合学術研究科 2013年度入学 表象領域

も怪しいとわかる地区にうっかり入り込んでしまった。小さな家々の窓から覗いているのは厚化粧をした女の顔ばかりだった。わたしは急いで次の角を曲がり、その狭い通りから離れようとした。ところが、道を尋ねる人もないままにしばらく彷徨ったのち、気がつくと、わたしはまた同じ通りに戻っているのだった。そしてその通りにいる人々の注目を集めはじめていた。わたしは大急ぎでそこから遠ざかろうとしたが、結果はまた別の回り道をして同じところに三度立ち戻る羽目に陥った。そのときにわたしを襲ったのは、不気味としか言いようのない感情であった (Freud 1919[2016]: 237)。

フロイトの感じた不気味さは予期せぬものの回帰・反復に起因するが、その予期せぬものは決して不確かなものではなく、確かに以前も目にした狭い通りであった。フロイトの言説を抽象化すると、情動としての不気味さは「予期せぬ既知性の露出」と紐づけることができる。だが、それと同時にフロイトによれば、「不気味なもの」がぞっとさせるもの、不安や恐怖をひき起こすものであることは疑いもない」(Freud 1919[2016]: 206)。すなわち、フロイトは (1) 情動としての不気味さ、(2) 不気味なものが生み出す情動の 2 つを論じているのである。不気味さと不気味なものとの定義が独立していないならば、不気味さを不安や恐怖と近接した複合的な情動と解釈することもできるが、いずれにしてもフロイトの記述からは不気味さ、あるいは不気味なものの捉え難さが垣間見える。

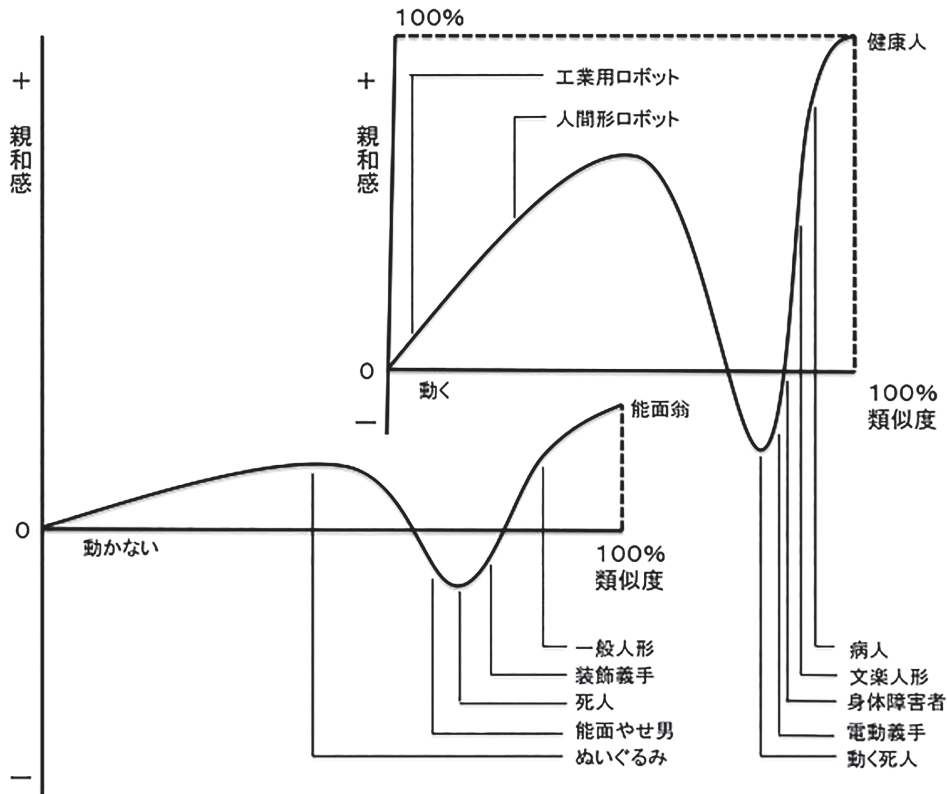


図1 不気味の谷。(森 1970: 34) を基に作成

さて、フロイトの議論から約半世紀後に提唱された不気味の谷仮説は、人に近いロボットがもたらす低親和感と不気味さを結びつけているわけだが、この不気味さに関して森は「自己防衛本能の重要な一部をなしていることだけは間違いない」(森 1970: 35) と論じている。すなわち、森によれば人には同種と近い存在の差異を鋭敏に見出すシステムが生得的に備わっているということになる。それが事実かどうかはさておき、そこには森が不気味の谷仮説を提唱することで主張しようとした内容の前提が表れている。森の主張とは、ロボットの製作における「人の類似度に捕われぬ、非人間的なデザインの推奨」であった。だが近年不気味の谷仮説が取りあげられる際に、森の主張が顧みられることはほとんどない。なおかつ仮説にすぎない不気味の谷が既成事実化している側面もあり、人型のロボットやCGに対して「不気味の谷を越えた」という表現がたびたび用いられている。したがって、森は

センセーショナルな仮説に基づいてある種の啓蒙を行おうとしたのだが、現在ではセンセーショナルな仮説のみが一人歩きしていると言っても過言ではない。

図1に示すように、不気味の谷の雛形は「動く対象」と「動かない対象」に分けられており、そこでは動きが加わることによって親和感の変動が大きくなると考えられていた。不気味の谷とはこのグラフの形状から名付けられた用語である。グラフは実験によって得られたデータを反映したものではない。なお、一般的に「不気味の谷」として紹介されるグラフは図1の「動く対象」のものに近く、明確な凹凸を持っている。

図1をさらに細かく見ると、森は「動く対象」のグラフにおいて高類似度かつ低親和感を示すものとして病人や身体障害者を挙げていることが分かるが、現在それらの対象が不気味の谷をテーマにした記述や研究の中で取りあげられることは極めて稀である。そして「動く対象」のグラフにおける類似度、親和感が共に100%の健康人とは一体何か、ということも定かでない。

デュムシェルとダミアーノは不気味の谷に関する哲学的論考の中で、われわれにとって最も不気味なものは人と区別がつかないロボットであり、類似度に基づく親和感の評価の行き着く先は超人的存在 (more human than human) による人の疎外であると指摘している (Dumouchel and Damiano 2017: 25-31)。不気味の谷はロボット研究者の妄想に過ぎないのだろうか。いずれにしてもここから、不気味の谷を越えるとはいかなることか、そもそも越えるべき谷はあるのか、という疑問が生じる。

3. 不気味の谷の実証的研究

仮説にすぎなかった不気味の谷が実証的研究において注目されるようになった背景には、2000年代初頭にロボットの外観を「人」に近づけることに執着した研究者のグループが人とロボットの円滑なコミュニケーションを目的とする研究の中で森の理論を積極的に援用した経緯がある (Minato *et al.* 2004; MacDorman 2005; MacDorman *et al.* 2005)。これらロボット研究者の理想——「人」とロボットの類似度を限りなく高めること——は森の主張に反するが、少なくとも外観を追い求めるロボット研究者にとって不気味の谷は非常に魅力的な仮説であり、研究を進めるための足場でもあったのだろう。

ロボット研究から派生した不気味の谷の実証的研究には既に多くの問題点が指摘されている。実証的研究の多くは森の作成したグラフを基に、複数の対象における類似度と親和感を評価することによって行われている。ここではまず評価尺度に関する批判的見解を次の3つに分類する。

[1] 用語のばらつき

森のいう不気味さ、親和感といった用語に対応する訳語の統一性に関する問題が指摘されている (Kätsyri *et al.* 2015)。ここでは既存の実証的研究において確認できる不気味さに対応する訳語として *eeriness*、*creepiness*、*strangeness*、親和感に対応する訳語として *affinity*、*familiarity*、*likeability* が挙げられている。訳語は実験上重要な意味を持っており、それらが不統一であることは関連した研究の相互参照を難しくする要因となっている。

[2] 低親和感は不気味さを意味するか？

人と類似したロボットの不気味さがどのような感情なのか、そして不気味さが親和感と対をなすものなのか不明であるという指摘がある (Ho *et al.* 2008)。そもそも親和感とは何か十分に定義されておらず、親和感のない状態が不気味であることを意味するののかもよく分かっていない。ここでは人型ロボットの不気味さに対応する用語として *strange* よりも *eerie* あるいは *creepy* が適切であり、それらは嫌悪 (*disgusted*) や恐怖 (*fear*) に近い情動であることが論じられているが、「低親和感=不気味である」ことを示すものでもなく、「高親和感=不気味でない」ことを示すものでもない。特に前節で取りあげたフロイトの言説を振り返ると、慣れ親しんだ対象の異化にこそ不気味さが見出されるわけであるから、従属変数の親和感が不気味さと如何に相関するのかということは議論の余地があるだろう。

[3] 恣意的な類似度の設定

評価対象に著名なキャラクターやロボットなどが含まれていたとき、それらの親和感は類似度に関わらず高くなる傾向にあり、実験結果に大きな影響を与えるため、そういった対象の類似度の恣意的な設定が問題視されている (Laja and Manaf 2015)。例えば、低類似度側の親和感のピークを任天堂のキャラクターであるマリオ、高類似度側の親和感のピークをララ・クラフトとするような実証的研究も存在する (Tinwell *et al.* 2011)。こういった問題を省みると、実験者による類似度の設定を極力回避すること、さらに認知度の高い対象は除外するといった対策が必要になるだろう。

上述した研究成果は実験条件の設定やデータを定量的に扱うことの困難さを示すものであり、これらの課題を全て克服した研究は今のところ存在しない。しかし、それに関わらず実証的研究は行われている。本節はこれ以降、[A] 不気味の谷仮説を否定する実証的研究、[B] 不気味の谷仮説を肯定する実証的研究の中でも代表的なものをいくつか取りあげ、特に [B] に関して批判的考察を行う。

[A] 不気味の谷仮説を否定する実証的研究

不気味の谷の存在を否定的に捉える実証的研究は数少ないが、ロボット工学者のデヴィッド・ハンソン¹は不気味の谷に対して懐疑的な見解を示している (Hanson 2005)。ハンソンが行った2つの実験を振り返りたい。1つめの実験では、人の頭部を模した2体のロボットの映像を用い、それらに対する印象を調査したところ、73%の被験者がロボットを好ましいものであると回答し、不快感を示した被験者はいなかった。さらに、85%の被験者がロボットは生き生きとしており、死人のように見えなかったと回答した。2つめの実験では、ディズニー映画「アラジン」に登場するジャスミンから本物の人へ段階的に変化するモーフィング画像を作成し、6段階に分類した。被験者による親和感の採点 (-100 ~ 100) の結果、全ての画像において評価の平均値が70以上であった。これらの実験結果から、ハンソンはロボットの造形が現実的になるにつれて人々が敏感になることは認めながらも、特定のポイントで親和感が急激に減少することなどないと論じている。なお、ハンソンは発展的な研究 (Hanson 2006) においてもモーフィング画像を用いた実験を行っているが、そこでは合成画像の作成の仕方によって実験結果が変わることを指摘しており、一貫して不気味の谷仮説に懐疑的な姿勢を示している。

[B] 不気味の谷仮説を肯定する実証的研究



図2 モーフィング画像 (Seyama and Nagayama 2007: 13)

不気味の谷の存在を肯定する初期の実証的研究として瀬山と永山によるものが挙げられる (Seyama and Nagayama 2007)。ここではCGから本物の人顔に至るモーフィング画像の親和感を採点 (-2 ~ 2) することで実験が行われている。実験結果は画像の合成の度合いによって親和感に優位な差は生じないというものであった。しかし、図2に示すようにモーフィング画像の一部 (中央3つ) の目の大きさを150%に修正して実験を行ったところ、不気味の谷の存在を肯定的に捉えうる結果が得られた。同実験で不気味の谷底付近にあるのは2種類のモーフィング画像共

に目の大きさが150%の高類似度の画像（図2の右から2番目の画像）であった。

瀬山と永山による研究は目と親和感の関連を示唆した点で大きな価値を持つが、不気味の谷の実証的研究としてはおおよそ以下の3つの問題を孕んでいる。第一に、画像に対する類似度が恣意的に決定されていること、第二に、目の大きさを150%にした画像を追加する実験操作に積極的な理由が見当たらないこと、第三に、不気味の谷の閾値が存在しないことである。第一の問題はすでに本節の前半で検討したため、第二、第三の問題に目を向けたい。まず第二の問題は、たとえ目が注意を誘引することが自明であるとしても、仮説を裏付ける極めて特殊な実験結果が強調されているのではないかという疑問に繋がる。これが正しければ、当研究の実験結果を一般化することはできない。さらに、第三の問題は不気味の谷の判別基準となる特定のデータの変動があらかじめ設定されておらず、最終的な判別が実験者の主観にゆだねられていること＝客観的かつ定量的な指標の不在を意味している。

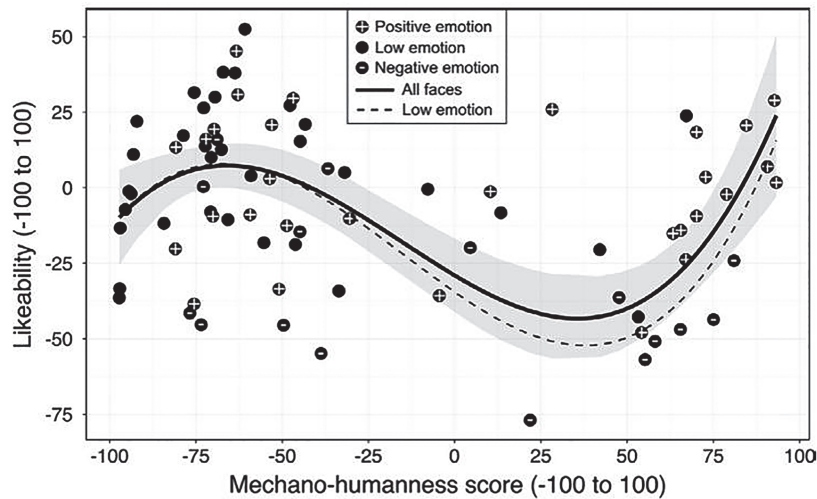


図3 ロボットに対する類似度と親和感の評価
(Mathur and Reichling 2016: 26)

上述した問題点の一部を克服し、なおかつ不気味の谷の存在を肯定する研究として、マトゥールとライヒリングによるものが挙げられる(Mathur and Reichling 2016)。ここでは80種類のロボットの画像に対する評価が行われた。これらロボットには全て頭部があり、抽象的な顔を持つロボット(e.g. マサチューセッツ工科大学の Kismet)から具体的な顔を持つロボット(e.g. ハンソン・ロボティクス社のフィリップ・K・ディック型アンドロイド)まで段階的に用意されている。被験者による「ロボットと人の類似度」および「ロボットに対して抱く情動値」の採点(共に-100~100)の結果、類似度と情動値は相関しないことが示された。その後、被験者(先の被験者とは異なるグループ)はロボットの画像に対する親和感の評価を行った。その結果を図3に示す。なお、グラフの縦軸の Likeability は親和感、横軸の Mechano-humaness score は類似度に相当し、情動値の平均値を V とすると、Positive emotion: $V > 10$ 、Low emotion: $-10 \leq V \leq 10$ 、Negative emotion: $V < 10$ を表している。

では図3を基にマトゥールとライヒリングの研究に対する考察を加えていきたい。実験者による近似曲線を脇において全体的なデータを眺めると、親和感が50をわずかに超えるロボットが1体あるのみで、過半数のロボットに対する親和感は-25~25の間にあることがわかる。すなわち、類似度を問わず親和感が高い傾向にあるとは言い難い。さらに、データが低類似度側と高類似度側に二極化しており、類似度が-25~25のロボットは7体にすぎない。したがって、データを近似する際に連続関数として扱うこと、換言すれば安易に定型的な曲線当てはめを行うことにも議論の余地がある²。不気味の谷の有無の判別にはデータの提示の仕方も大きな影響を与えるため、より慎重な吟味が必要なのではないだろうか。

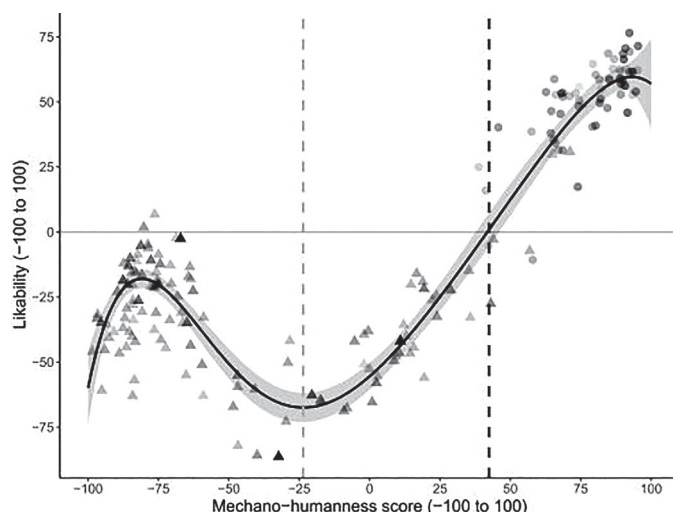


図4 人とロボットに対する類似度と親和感の評価
(Mathur and Reichling 2016: 26)

上記の実験はロボットの画像のみを評価対象とするものだったが、マトゥールとライヒリングを中心とした研究グループは本物の人の画像を評価対象に加えて同様の実験を行っている (Mathur *et al.* 2020)。ここでは、122種類のロボットと60種類の本物の人の顔画像に対する類似度と親和感の評価が行われた。その結果を図4に示す。なお、グラフ中の○は本物の人、△はロボットを表しており、データは逆分散法によって重み付けされている。

図4から、先の研究 (Mathur and Reichling 2016) よりもロボットに対する評価が相対的に低くなっていることが分かる。対称的に、人に対する親和感の評価は2体を除いて0より大きい。すなわち、対象が人であるか否かによって評価がおおよそ二分しているのである³。そして低類似度側の頂点がおおよそ高親和感を示していないことや、親和感の最小値が低類似度側にあることから、近似曲線の正当性に関する問題を譲歩したとしても、実験結果が不気味の谷を支持するとは言い難い。

本節は不気味の谷の実証的研究の批判的な検討を行った。もはやこれ以上説明するまでもなく、不気味の谷を科学的証拠に基づくものとして扱うならば、より精緻な裏付けが必要になるだろう。少なくとも現状、不気味の谷という言葉はロボットやCGと「標準的・模範的な人」の類似度に関する無根拠な指標として用いられていると言わざるを得ない。発展的な議論の素地を醸成するために不気味の谷と不気味な現象 (uncanny phenomenon) を区別して考えるべきだということは既に指摘されている (Wang *et al.* 2015)。だが、そういった問題提起を引き受けるとしても、そこにはさらなる未解決の問題が潜んでいる。つまり不気味な現象の「不気味さ」は自明なものとして扱われているわけだが、一体不気味さとは何だろうか。

4. 「人のようなもの」の不気味さ

ここで話題を不気味の谷から不気味な現象に移し、われわれに近い存在でありながらも得体の知れない「人のようなもの」と不気味さを結びつける内的システムについて情報处理的な観点から検討していきたい。

認知科学における「人のようなもの」の不気味さの研究はおおまかに分類曖昧性 (categorization ambiguity) および知覚的不一致 (perceptual mismatch) の観点に基づいて行われている (Kättsyri *et al.* 2015)。分類曖昧性を持つものは異なるカテゴリーの境界に位置する対象や、特定のカテゴリーに分類することが困難な対象である。これに対して、知覚的不一致を引き起こすものは前節でも取りあげたモーフィング画像のように、部位によってリアリズムのレベルが異なる画像や、特定の部位が誇張された画像など、予測誤差が大きい対象である。

分類曖昧性に関する研究の中で、われわれが人のようなロボットを知覚するとき、脳の情報処理系に混乱が生じているという説がある (Shimada *et al.* 2007)。嶋田らは、ジョゼフ・ルドゥーによる情動の二重経路説を援用し、視床から扁桃核へ直接至る低位経路と大脳皮質の認知システムを通過して扁桃核に至る高位経路、これら2つの経

路を通過した情報が示すカテゴリの矛盾によって不気味さが生まれると論じている。低位経路は粗雑な情報を扱うため、人のようなロボットを人として誤伝搬するが、高位経路はより厳密な情報を扱うと考えられており、それを人ではないもの=ロボットとして正確に伝えるのである。なお、ルドゥーは扁桃体が不安や恐怖の座であるという説を流布した人物として知られているが、現在自説の誤りを認めている (LeDoux 2019)。したがって、ここでも二重経路全体を不気味さと関連するシステム (あるいはその一部) であると考えた方がよいだろう⁴。

しかし結局のところ不気味さが如何なる情動かという問題は未だ残存している。嶋田らはルドゥーを介することで暗に不気味さと不安や恐怖を繋げている反面、それが不安あるいは恐怖のサブカテゴリなのか、複合的なものなのか、ということは明らかにしていないのである。

一方、正確には知覚的不一致とは言えないが、それと近い研究の中で、人のようなロボットが不気味なのは、それが心の帰属を促すからであると指摘されている (Gray and Wegner 2012)。ここではわれわれが心の知覚を行う際の独立した指標として、行為者性 (agency) と経験 (experience) が取りあげられている。行為者性はプランの実行や、行為の制御をする能力である。これに対して経験は感じ、感覚する能力である。グレイとウェグナーはこれら2つのうち、特に経験が心の知覚にとって重要であり、不気味さを引き起こす要因であると論じている。例えば、ロボットは本質的に経験を欠くものであると考えられるが、外観的な特徴から経験を有しているかのように知覚されるとき、われわれは不気味さを感じる。対照的に、人は本質的に経験を持つものであると期待されるが、それを欠いているように知覚されれば、不気味さを誘発する。このように、不気味さは人のようなロボットだけでなく、ロボットのような人も生じさせることが分かる。

さて、シャインとグレイによれば経験の知覚における最も重要な特徴は目である (Schein and Gray 2015)。シャインとグレイは本物の人の顔画像を「目を白抜きにした画像」と「鼻を白抜きにした画像」に加工したうえで、それらと元画像の3種類の画像に対して不気味さ (uncanny)、生気の無さ (soulless)、行為者性の欠如 (lack-agency)、経験の欠如 (lack-experience) の4つを調査した。その結果、行為者性の欠如以外の3つの指標において、「目を白抜きにした画像」に対する評価が極めて高くなったのである。実験結果はアニメーション知覚にとって顔の中でも目が最も重要な手がかりであるという研究 (Looser and Wheatley 2010) を補強するといえる。このように、目は他者の感じ、感覚する能力、さらには生物性の判別にも関与している可能性がある。

他者への心の帰属に目が重要な役割を持つことを応用して、ロボットに「心」を実装しようとする取り組みも存在する。その中でも瞬目に着目した研究の1つは、ロボットに人の生理学的データに基づいたまばたき行動を実装すると、まばたきをしない場合や5秒毎にまばたきをする場合よりも知覚的に見えると報告している (Lehmann *et al.* 2016)。したがって、「心」を行為者と被行為者の相互関係に見出されるものと捉えれば、人とロボットのコミュニケーションにおいても人同士の場合と同様に、他者と「心」を通わせること、共感するということが起こりうるのである。もしその対極にあるのが不気味さであるとするれば、それは「心」が無く共感することのできない、得体の知れない相手から回避すること、自らを防御することに繋がる。この極めて動物的な反応は人の排外的な側面を露呈するが、裏を返せば「心」と地続きにあるとも考えられる。

以上、本節は不気味な現象に関する考察を行ってきた。ここで得られた知見を本稿全体に繋げると、興味深い事実が浮かび上がってくる。それはフロイトから現在に至るまで一貫して不気味さという概念自体に揺らぎが見られるということである。本稿は不気味さを中心に据えた議論を行ってこなかったが、そもそも不気味さに対応する特定の物理状態が存在するかどうかは定かではないのだ。しかし、その揺らぎこそがわれわれを不気味なもの、ひいては不気味の谷に惹きつけてきたのではないだろうか。

5. おわりに——人のようなものへの生成変化

われわれが単純な3つの点の集合——(∴)——を顔であるかのように知覚してしまうことはシミュラクラ現象⁵と呼ばれており、人のパターン探索器としての一面を端的に示している。そして顔の原型とも言える記号(∴)は宮崎駿監督の『もののけ姫』に登場する「こだま」を想起させる。「こだま」は木の精霊であり、一種のアニミズムがそこに表されているわけである。

生物学者ヤーコプ・フォン・ユクスキュルは著作の中で、ひとの顔に似たこぶのあるカシワの木を取り上げ、木こりはそれにあまり注意を払わないが、少女は樹皮に怒った顔を見だしぎょっとする、という旨の記述を行っている (Uexküll 1934[2005]: 146)。ここでは木こりと少女それぞれの環世界の差異が示されている。木こりは世界をわかりきったものであると考えている。対照的に少女の世界は魔術的であり、未知に溢れている。ユクスキュルの記述は人の社会文化的な営みを象徴しており、これを未知なるものの既知化と捉えることもできるだろう。

未知性と既知性をテーマにした代表的な作品にスタニスワフ・レムが1961年に発表した『ソラリス』がある。未知なる惑星ソラリスの海は、宇宙ステーションで海の観測を行う主人公の下に亡くなったはずの恋人のようなものをもたらす。この「予期せぬ既知性の露出」はフロイト的な不気味さと一致する。物語の中で海は一貫して未知性の象徴であり続けるが、最終的に幼児のようなもの=既知的な存在と解釈される。したがって、レムは未知なるものの既知化に盲進する人の性質を浮き彫りにしているのだが、その一方で不気味なものとの共生のあり方も描いている。主人公による恋人のようなものの排除—包摂を巡る葛藤はそれを如実に物語っているのではないか。

未知なるものと既知なるものの差異は個々の身体に根ざすだけでなく、それらを取り巻く社会文化的な状況によっても変わってきた。そして、われわれは絶えず未知なるものを既知的にとらえようとしている。これまで本稿で論じてきた「人のようなもの」は未知性と既知性を併せ持つ存在であり、そういった存在が生み出す情動を浮動的かつ存在すら疑わしい「不気味さ」に託すことは、われわれの機微を正確に捉えているといえる。だがそれと同時に、その根底には排除の機構が渦巻いていることも直視すべきだろう。

われわれは他者を同化しながらも、他者との差異を生み出すことによって何者かであろうとする。われわれは排外的だが、他者の権利を認めることもできる。それゆえ、われわれには「人のようなもの」を乗り越えるのではなく、ただそのまま受け入れる道もあるはずなのである。

付記

本稿は R-GIRO プログラム「次世代人工知能と記号学の国際融合拠点」によるシンポジウム「人と AI の調和が導く未来社会に向けたアプローチ」(2020年2月19日、於立命館大学大阪いばらきキャンパス)でのポスター発表をもとに作成した。

注

- 1 ハンソン・ロボティクス社の創業者・CEOである。同社のソフィアと呼ばれるロボットは2016年度のSXSWで「人類を滅ぼしたいか?」(“Do you want to destroy humans?”)とインタビューされた際、「いいとも、人類を滅ぼそう」(“OK, I will destroy humans.”)と回答した(正確には製作者によって回答させられた)ことで知られている。
- 2 ここでの主張を補足する研究として、不気味の谷仮説における類似度が直線的に変化する前提は主観的な知覚・認識を反映したのではなく、実際には対象が人カテゴリと判断されるか否かによって類似度評価に明確なカテゴリ境界が生じるという指摘 (Cheetham *et al.* 2011) が挙げられる。
- 3 被験者が本物の人と判断した対象に低評価を行わない傾向が見られることから、ある種の自己検閲が働いている可能性もあるが、本稿の主旨と外れるためここに手短かに付するにとどめる。
- 4 この問題は基本情動説と心理学的構成主義の対立とも関係しており、情動に関連するシステムの局在性自体が論争的なテーマとなっている (太田 2020)。
- 5 より広義にはパレイドリア=データの中に特定のパターンを見いだしてしまうこと。

参考文献

- Cheetham, M., Suter, P. and Jäncke, L., 2011, “The Human Likeness Dimension of the “Uncanny Valley Hypothesis”: Behavioral and Functional MRI Findings”, *Frontiers in Human Neuroscience*, 5 (126): 1-14.
- Dumouchel, P. and Damiano, L., 2017, “Living with Robots”, London: Harvard University Press.

- Freud, S., 1919, "Das Unheimliche". (=2016, 原章二訳, 「不気味なもの」『笑い/不気味なもの』平凡社, 205-74.)
- Gray, K. and Wegner, D., 2012, "Feeling Robots and Human Zombies: Mind Perception and the Uncanny Valley", *Cognition*, 125: 125-30.
- Hanson, D., Olney, A., Pereira, I. and Zeilke, M., 2005, "Upending the Uncanny Valley", *Conference: Proceedings, The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference*, 24-31.
- Hanson, D., 2006, "Exploring the Aesthetic Range for Humanoid Robots", *Proceedings of the ICCS/CogSci-2006 Symposium: Toward Social Mechanism of Android Science*, 39-42.
- Ho, C., Pramono, Z. and MacDorman, K., 2008, "Human Emotion and the Uncanny Valley: A GLM, MDS, and ISOMAP Analysis of Robot Video Ratings", *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, 69-76.
- Kätsyri, J., Förger, K., Mäkäräinen, M. and Takala, T., 2015, "A Review of Empirical Evidence on Different Uncanny Valley Hypotheses: Support for Perceptual Mismatch as One Road to the Valley of Eeriness", *Frontiers in Psychology*, 6: 1-16.
- Laja, L. and Manaf, A., 2015, "Overcoming the Uncanny Valley Theory in Digital Characters Based on Human Attitudes", *Pertanika J. Soc. Sci. & Hum*, 23 (5):13-22.
- Lehmann, H., Roncone, A., Pattacini, U. and Metta, G., 2016, "Physiologically Inspired Blinking Behavior for a Humanoid Robot", *International Conference of Social Robotics*, 83-93.
- LeDoux, J., 2019, "The Deep History of Ourselves: The Four Billion-Year Story of How We Got Conscious Brains", New York: Viking.
- Looser, C. and Wheatley, T., 2010, "The Tipping Point of Animacy: How, When, and Where We Perceive Life in a Face", *Psychological Science*, 21 (12):1854-62.
- MacDorman, K., 2005, "Morality Salience and the Uncanny Valley", *Proceedings of 5th IEEE-RAS International Conference on Humanoid Robots*, 399-405.
- MacDorman, K., Minato, T., Shimada, M., Itakura, S., Cowley, S. and Ishiguro, H., 2005, "Assesing Human Likeness by Eye Contact in an Android Testbed", *Proceedings of the Annual Meeting of the Cognitive Science Society*, 27:1373-8.
- Mathur, M. and Reichling, D., 2016, "Navigating a Social World with Robot Partners: A Quantitative Cartography of the Uncanny Valley", *Cognition*, 146: 22-32.
- Mathur, M. Reichling, D., Lunardini, F., Geminiani, A., Antonietti, A., Ruijiten, P., Levitan, C., Nave, G., Manfredi, D., Bassette-Symons, B., Szuts, A. and Aczel, B., 2020, "Uncanny but not confusing: Multisite Study of Perceptual Category Confusion in the Uncanny Valley", *Computers in Human Behavior*, 103: 21-30.
- Minato, M., Shimada, M., Ishiguro, H. and Itakura, S., 2004, "Development of an Android Robot for Studying Human-Robot Interaction", *Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE)*, 424-34.
- 森政弘, 1970, 「不気味の谷」, *Energy*, 7 (4): 33-5.
- 太田陽, 2020, 「基本情動説と心理構成主義」, *Contemporary and Applied Philosophy*, 11: 23-57.
- Schein, C. and Gray, K., 2015, "The Eyes Are The Window to The Uncanny Valley: Mind Perception, Autism and Missing Souls", *Interaction Studies Social Behaviour and Communication in Biological and Artificial Systems*, 16 (2): 173-179.
- Seyama, J. and Nagayama, R., 2007, "The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces", *Presence: Virtual and Augmented Reality*, 16 (4): 337-51.
- Shimada, M., Minato, T., Itakura, S. and Ishiguro, H., 2007, "Uncanny Valley of Androids and Its Lateral Inhibition Hypothesis", *16th International Conference on Robot & Human Interactive Communication*, 374-9.
- Tinwell, A., Grimshaw, M. and Williams, A., 2011, "The Uncanny Wall", *International Journal of Arts and Technology*, 4 (3): 326-41.
- Uexküll, J. von, 1934, "Streifzüge durch die Umwelten von Tieren und Menschen: Ein Bilderbuch unsichtbarer Welten", Berlin: J. Springer. (=2005, 日高敏隆・羽田節子訳, 『生物から見た世界』岩波文庫.)
- Wang, S., Lilienfeld, S. and Rochat, P., 2015, "The Uncanny Valley: Existence and Explanations", *Review of General Psychology*, 19 (4): 393-407.

The Fall of the Uncanny Valley

ITO Kyohei

Abstract:

The aim of this paper is to criticize the empirical tests for the uncanny valley hypothesis and analyze the emotion that is provoked by humanlike beings (e.g. humanoid robots). The uncanny valley hypothesis was proposed by the roboticist Masahiro Mori in 1970. According to the concept, the affinity for a robot increases with the degree of similarity between robot and human, but at a certain peak, the affinity for a robot decreases drastically, and it increases again as the degree of similarity gets close to real human level. Recently empirical studies of the uncanny valley have received much attention, but the results have been inconsistent. This paper points out the difficulties of demonstrating uncanny valley hypothesis by means of statistical methods, and clarifies the concept “uncanny” has been difficult to define from Freud’s era to the present. However, it is suggested that this elusiveness makes this hypothesis attractive. In conclusion, this paper proposes a way of coexisting with humanlike beings that arouse elusive emotion.

Keywords: Uncanny Valley, Humanlike Beings, Symbiosis

不気味の谷の陥落

伊藤京平

要旨：

本稿の目的は不気味の谷の実証的な研究成果を批判し、われわれが人型ロボットをはじめとした「人のようなもの」に抱く情動に関して考察を加えることである。不気味の谷仮説はロボット研究者の森政弘によって1970年に提唱された。この仮説によれば、ロボットを人間の外観に近づけていく過程で、類似度が大きくなるほどそのロボットに対する親和感は増加していくものの、特定のピークを迎えると親和感が急激に減少し、谷を過ぎると親和感が再び増加する。本稿は変数の設定から統計解釈に至るまで不気味の谷の実証研究に見られる複数の問題を明らかにする。さらに、われわれが不気味なものや不気味の谷に惹きつけられてきた理由として、不気味さの概念的な揺らぎを示す。最後に、その揺らぐ情動を生み出す人のようなものとの共生へと議論を展開する。