

NTTコミュニケーション科学基礎研究所
上席特別研究員

河邊隆寛 Takahiro Kawabe

視覚情報処理のメカニズムを 解明し遠隔へ質感を届ける 実験心理学者の挑戦

いかにも柔らかさそうで触りたく見える物の動き。今にも前へ走り出しそうに見える物の絵柄。人間の目を通じて、このように感じとれるのは、一体どういった知覚の仕組みが作用しているからなのでしょう。NTTコミュニケーション科学基礎研究所の河邊隆寛上席特別研究員は心理学の出身でありながら、情報処理の方法論を活用し、こうした脳内での視覚情報処理のメカニズムを解明しようとしています。また同時に、遠隔へ視覚的に質感を伝える技術の社会実装に向けても奔走されています。今回の取材で、最近の研究成果や今後の展望について伺いました。



柔らかさや触りたさを視覚で認識する

最近の視覚質感の研究について教えてください。

私は現在も視覚の研究に取り組んでいます。視覚の研究といってもさまざまですが、機械やコンピュータではなく、特に「人間」の視覚に焦点を当てて研究をしています。

前回のNTT技術ジャーナルでのインタビューでは「視覚質感の研究」についてお話をさせていただきました。私の視覚の研究は、遠隔地へ「質感」を伝えることを目的に、これを表現するパラメータを見出すことにより視覚を心理学的な手法で解明し、このパラメータを使った技術を確認し社会実装していくものです。

まず柔らかさや触りたさを視覚から認識・識別する研究について触れていきます。実験として、まずコンピュータグラフィック上の仮想空間に1枚の平面を準備します。この平面に針状の棒を押し下ろしたり、戻したりする様子を映像として記録し、どの程度深くまで押し込まれたかという「押し込み量」や押し込まれた結果、どの程度横方向に変形したかという「空間的な変形の範囲」を人間が見たとき、どう感じるのかを評価しました(図1)。

空間的な変形量8 pixelでは針の周辺にしか変形が見られませんが、32 pixelになると、より広範囲が窪んで変形したように見えます。さらに64 pixelでは一見、変形しているようには見えませんが、平面全体がたわむように形を変えています。

実験参加者にこの映像を見てもらい、「柔らかさ」や「触りたさ」の感触について質問した結果を図1のグラフに示します。2つの

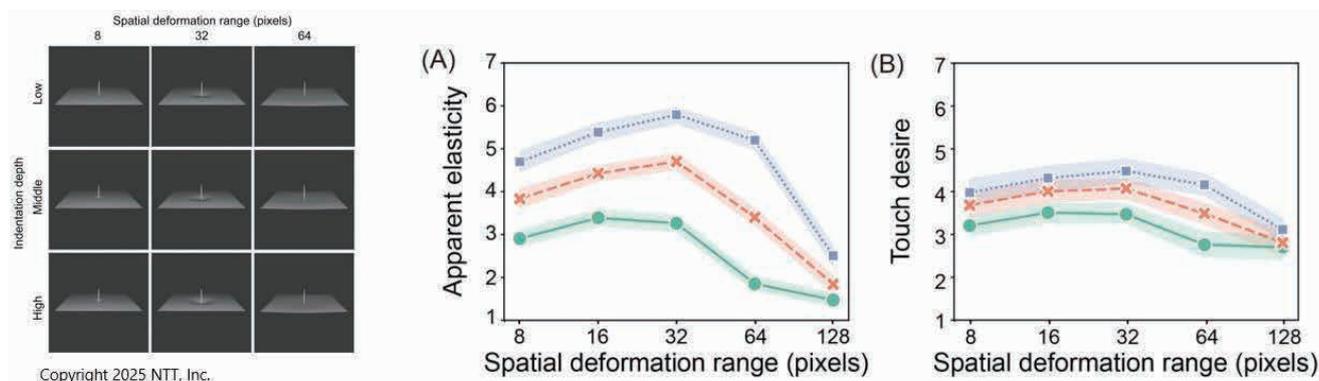


図1 物体映像への押し込み量と空間的な変形範囲から感じる柔らかさと触りたさ

グラフの横軸は共に「空間的な変形の範囲」で、グラフ(A)の縦軸が「見た目の柔らかさ」、グラフ(B)が「触りたさ」を表しています。グラフ中の各プロットでは押し込み量を変化させており、緑色< 橙色< 青色の順に大きくなっていきます。

(A)、(B)の全体的なグラフ形状は類似しており、人間が柔らかいと感じることと、触りたいと感じることは明確に相関があると判断できます。また、両グラフにおいて、大きく押し込んでいる青色ではもっとも柔らかさや触りたさを感じ、小さく押し込んでいる緑色では、柔らかさや触りたさをあまり感じない結果となりました。さらに、押し込み量には依存せず、空間的な変形の範囲が中程度において、もっとも柔らかく、かつ触りたく感じることも明らかになりました。

今までの研究では、押し込み量が柔らかさや触りたさに影響することは分かっていたのですが、空間的な変形の範囲が影響していることは、今回の研究で初めて実証されました。従来は、対象の柔らかさを遠隔へ伝えるとき、押し込み量だけで表現されていたため、この感触を伝える手段の少なさが課題でした。しかし、今回、横方向への変形範囲も柔らかさや触りたさを伝えるパラメータになることを見出し、表現する引き出しの数を増やすことができました。

このような視覚的質感の研究について、ビューティテックを志向する資生堂社に説明する機会があり、例えば化粧品の質感や肌触りを視覚情報として伝えることはできないのかと2025年9月から共同研究を開始しています⁽¹⁾。例えば拡張現実なども視野に入れ、商品の映像だけでは難しい質感についても伝えられるのではないかと検討しています。現在はNTTの技術を資生堂社でテストしていただいている状況で、今後、皆さんへも発表できる機会があるのではないかと考えています。

物の生物性や意図性を視覚の切り口から研究されているそうですね。

次に視覚から感じる生物性・意図性の研究についてです。人は

柔らかさや重さといった質感とは別に、目の前の物がロボットやCGキャラクタなど無生物であっても「生きているように感じる、意図を持って動いているように感じる場面があるか」と思います。私たちはそういった対象に、生物性と意図性がどのように発現していくのかを追求しています。そしてこの研究の実用的な成果としては「生き物らしい」「人が意図を読める」ロボットやCGキャラクタのデザインにつなげていければと考えています。

具体的には、二次元平面上を自由に動くことのできるキューブ型ロボット（箱型ロボット）のどこに目をつけたら、視線をどの向きにつけ描いたら生物らしく見えるのかについて研究をしています。結論は「目は動く方向の前方」についていると生き物らしさが発現する、または自立的に動いている印象を受けるといものでした。興味深いのは、その目の位置は重要であるが、視線の方向はあまり影響を及ぼしていないという事実でした（図2）。

私たち人間は他の霊長類と比べて白目の割合が一定程度あり、強膜が明瞭に露出しているため他者の視線の方向性や何かしようとする意図を読む能力に長けています。しかし、同じ哺乳類でも犬や猫は白目の割合が小さい動物です。生物が進化するうえで、目がどちらを向いているのかではなく、実は目の位置がどこにあるのかがより重要な手掛かりであったのではと、1つの仮説として考えています。

そして、将来のロボットの目の位置をデザインしていくうえでも、こういった研究成果が創出されていると、何らかの手掛かりになるのではと感じています。

図3は三角形が自由に動きまわる映像を作成し、その映像の生き物感度を評価させた実験ですが、その際、三角形の頂点が動く方向に合っている、つまり、頂点を先頭に三角形が自由に動き回ると生き物っぽさを感じるということが分かりました。一方で、三角形の頂点が動き方向とずれている場合は、水や風に流されてしまっているように見える。つまり、生き物が動いているのではなく物体が流されているように見えるという評価になりました。そして、三角形に目をつけた場合にどうなるかと実験したところ、たとえ頂点の向きと動く方向が合っても、目がその動きの前

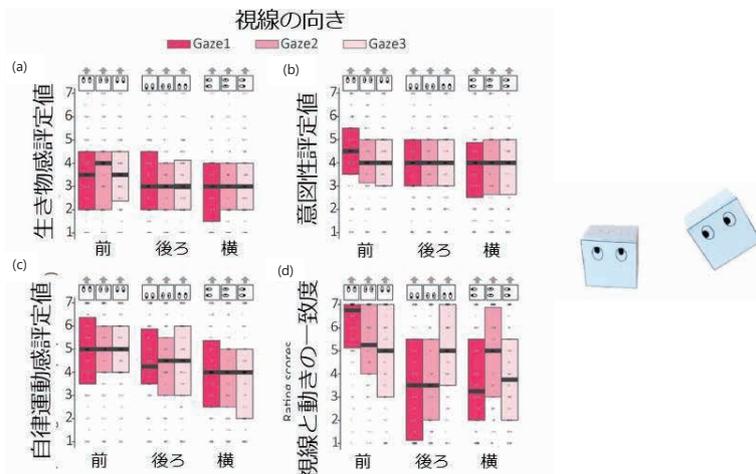


図2 キューブ型ロボットにつけた目の位置と視線が与える生き物感や意図性への影響

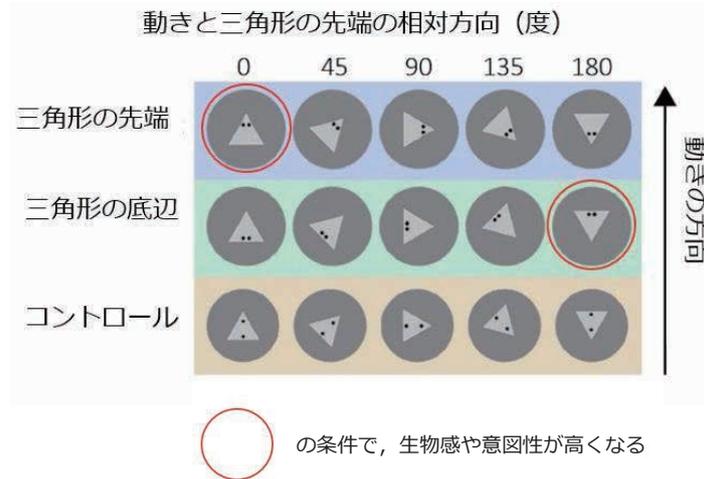


図3 目の位置と形状による生き物感や意図性

方にあるほうが生き物らしく見えているということが分かりました。

この実験でも、人間は目の位置と対象の形とを組み合わせる処理をしている中で、やはり「目の位置」が動く方向との関連性において重要であることが示されています。したがって、この実験結果は工業的なデザインを創作するうえでも、目の位置を工夫することが極めて大切であることを物語っています。

時間の流れる方向の知覚を例に視覚メカニズムを解明する基礎研究について教えてください。

最後に、視覚のメカニズムを解明する基礎的な研究についてお話しします。私たちは目に光が入ってくることで物を見ることができます。目の中ですべての情報処理が行われているように感じがちですが、実際には、そう単純ではありません。目から入った情報はまず網膜に当たり、そこで神経活動へと変換され、その信号が脳へと伝わり、脳の中で段階的な処理が行われることで、私たちはそれを「見えている」と感じているのです。

脳の中では、時間を処理する場所、空間を処理する場所、動きや移動を処理する場所、方向を処理する場所など、階層的にさまざまな処理を担う領域に分かれています。そうした処理の仕組みを、一つひとつ細かく、丁寧に調べていくことで、光のパターンから人間がどのように物を見ているのかを明らかにしていく。これが視覚のメカニズムを解明する基礎研究です。

ここでは、時間の流れる方向や液体が流れる方向の知覚メカニズムを調べた研究を紹介します。これらの研究は学術的な観点において、人間が物を見たときに、脳がどのように働いているのかを理解するうえで役立ちますし、さらに、そうした脳の仕組みを理解することで、将来の情報提示手法を新たに生み出すためにも役立つと信じています。

ここでは一例として、ある映像において、時間の流れる方向の知覚メカニズムについて紹介します。私たちは普段、例えば動画が巻き戻されていると、すぐこの動きに気付くことができます。しかし、なぜそれに気付けるのかは、はっきりと分かっていませ

んでした。そこで、そのヒントとなる現象として、今回1つの発見がありました。それは、「映像を上下に反転させると巻き戻しに気付きにくくなる」という現象です。

例えば、川（液体）が流れている映像について、正立映像と上下反転させた倒立映像を作成し、視聴者に4秒間ほど見てもらいます。そして、通常どおり再生する場合と、巻き戻しをして再生する場合と2つを提示し、今見た映像はどちらの映像であったか、と質問しました。その結果、正立映像では巻き戻し動画を皆さん容易に検出することができましたが、倒立映像になると巻き戻し動画の検出率が低下したのです（図4）。

これはとても不思議な現象で、もし人間が映像や目に入ってくる情報の中で「時間情報」だけを抽出していれば、映像を上下反転させても影響はでないはずで、ところが、実際には上下反転させると検出率に大きな差が出てしまう。これはおそらく、映像を上下反転させることで、物と物との関係性、つまり空間的な構造が分かりづらくなってしまいうからだと考えられます。

上下反転した映像では、そこに映っている個々の物は認識できても、それらどうしの関係が把握しづらくなります。例えば顔のパターンも同じで、顔を逆さまにすると、誰の顔なのか急に分かりにくくなります。さらに、有名な例としてサッチャー錯視*1があります。顔全体が逆さまになっている一方で、目や口だけは正しい向きのまま配置されている画像に、私たちは特に違和感を覚えずにそれを見てしまいます。顔全体が逆さまになることで、個々の要素がどうなっているのかを把握しにくく、結果として全体を違和感なく見てしまう、という現象が起こるのです。

今回の結果も、これと同じような現象だと考えています。映像を逆さまにすると空間構造が分からなくなり、例えば川の映像では、水がどちらへ流れているのか把握しづらくなります。そのため、

*1 サッチャー錯視：上下を逆さまにした顔写真では、目や口が反転していても異常に気付きにくく、正立させると不気味に見える心理現象。1980年にピーター・トンブソンがマーガレット・サッチャーの写真を用い提示したことで知られ、人間の顔認知が「配置」に強く依存していることを示します。

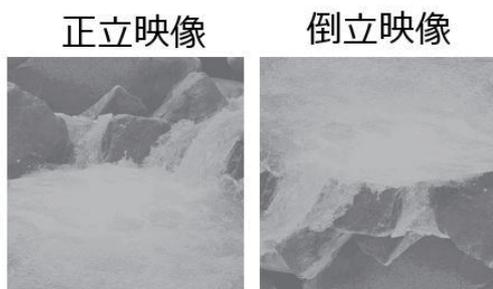
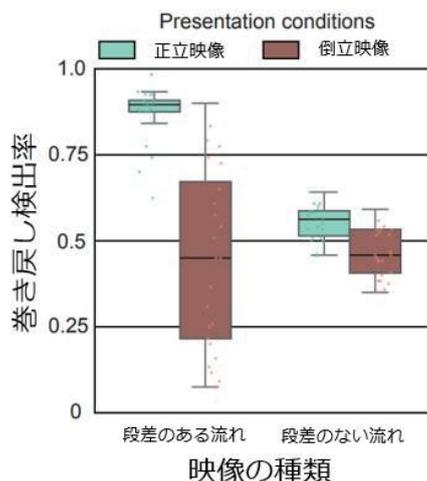


図4 水が流れる映像における正立画像と倒立画像での巻き戻し検出率



時間的に逆転しているのか、それとも正常なのかを判断しにくくなり、結果として巻き戻しを検出しにくくなったのではと考えています。

今回見てもらった映像は、正立の場合、巻き戻しをすると下から上に水が流れます。一方、倒立の場合、巻き戻しをすると上から下に流れる動画となり、図4左の茶色のグラフのとおり検出率は低くなり、実際には巻き戻した映像のほうが自然に見えたという人も結構いる結果が得られました。この事実から、さまざまな動きのあるシーンにおいて、その空間構造が時間の流れる方向の知覚に大きく影響を与えていると考えられます。

視覚にかかわる脳内細胞の働きを計算でシミュレーションする

液体の流れる方向の知覚を例に視覚メカニズムを解明する基礎研究について教えてください。

もう1つの例として、液体の流れる方向の知覚についての研究です。例えば川が流れているとき、私たちは特に意識しなくても、その方向を自然に理解しています。では、なぜそれが分かるのか調べてみました。

通常、物は多少動いても形や色はあまり変わりませんが、液体はとても面白い物体で、動くと波が立ち、その波も同じ形のまま流れ続けるわけではなく、途中で消えたり、新しく生まれたり複雑な動きをしますし、それに伴い光の反射パターンも変化します。それにもかかわらず、私たちは液体の流れる方向を簡単に識別することができます。なぜそれが可能なのかという点が、視覚の基礎研究において非常に重要だと考えます。

私の研究では、「脳の中で動きの方向を検出する細胞の動き」に注目し、それを畳み込み処理に基づいて計算でシミュレーションしました。そして、さまざまな液体が流れる映像を、シミュレーションしている脳に見せて、「この液体はどちらに流れていますか」と問いかけます。そのとき、どのような計算を実施すれば、人間

の回答に近い結果を出せるのか、つまり、どんな計算過程が人間の視覚の仕組みに近いのかを調べています。

脳の中に実は「動きの方向」を認識する細胞が多種類存在しています。例えば、目に映った映像の上方向の動きに反応する細胞、下方向に反応する細胞、左方向に反応する細胞、さらには斜め方向に反応する細胞など、それぞれ特定の方向に反応する細胞たちです。

そこでこの研究では、どのような細胞からの情報を使って私たちは液体の流れを見ているのかを調べました。これまでの研究でよくいわれてきたのは、「たくさんある細胞の中で、もっとも強く反応している細胞の出力だけを見れば、方向は分かるのではないか」という推論です。これは「Winner-take-all」と呼ばれ、一番強く反応した細胞の意見が判断を支配する考え方です。

しかしこの方法を使うと、図5左に示すとおり、実験参加者が報告した方向との誤差が48 degと大きくなりましたので、私たちは「重み付けして平均する」方法を採用しました。すなわち、すべての細胞の反応を考慮し、それぞれの反応の強さに重みを付け、全体として平均を取る（加重平均）方法です。実際にこの計算を実施してみると、実験参加者が報告した方向との誤差はおおよそ10 deg程度まで縮小し、人間の判断に近づいてきました。

今後の展望についてお聞かせください。

私が所属しているNTT先端技術総合研究所は世界唯一の発見、世界一の技術をめざすことを目標に掲げていますので、私自身も視覚情報処理のメカニズムの分野において学術的に価値が高く、世界で初めての発見をめざして研究を進めています。

また、最近には特に「物の形」に興味を持っています。「形」とは使いやすい形など、何かしらの意味を持ってつくられた工業デザインはもちろんですが、自然に存在しているものでも、その形には何かしらの意味があるのではないかと考えています。

人が形を見たときに感じる印象は、どのような仕組みでそう感じるのか。形と人が抱くさまざまな印象とを結び付ける人間の知

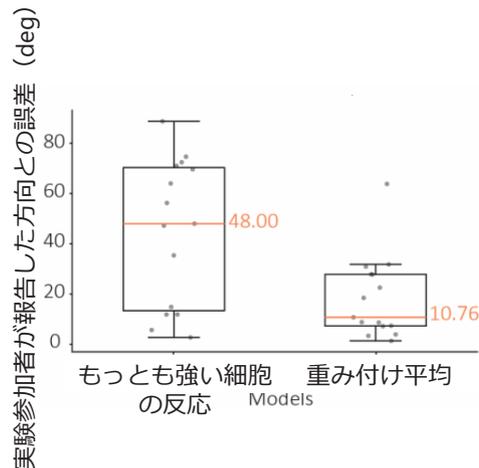


図5 各計算モデルと実験参加者が報告した結果との誤差

覚特性を解明していきたいのです。例えば「美味しそうな形」とは、どのような形だろうか。「生き生きとした形」とは何だろうといった、一見、結び付きそうもない修飾語で表現された形を認識するメカニズムについて興味を持っています。既存研究では、「甘い形」は少し丸みを帯びており、「しょっぱい形」はギザギザしていると、味覚と形の関係性について示唆した例があります。

しかし、まだ視覚の詳細な仕組みは分かっていませんので、私はその辺を明らかにしていきたいです。例えば「美味しそうな形」とはどのような形なのか、その形を人間はどのように目や脳で処理をし、そういった印象が生まれているのか、というメカニズムを解明していきます。そして、将来的にはこういった知見をとおして、食品の広告デザインやロボットデザインに活かされていくのではと考えています。

さらに、研究成果に基づいた社会実装もめざします。近々では、前述した資生堂社との共同研究を想定していますが、遠隔地への質感情報の伝達を可能にする視覚的特徴を発見し、それを世の中で活用することをめざしていきます。

やはり、実装といっても強固な基盤がないと、当面の問題を解決することにしかならず、それ以上の発展が見込めない、その場限りの問題解決になってしまいがちです。深いところまで仕組みを理解しておくことでさまざまな問題に汎化していくことができるのだと、かつて変幻灯^{*2}を開発したときにも痛感しました。基礎研究でしっかりと足元を固めつつ、縁あってNTTという企業で活動している以上、成果を世の中で活かすための道筋も考えていきたいです。

後進へのメッセージをお願いします。

日々の忙しさの中でも、「自分を見つめる時間を大切にしてい

い」と思います。自分のやりたいこと、自分がやるべきことを、しっかりと反芻し、自分なりの言葉で整理したうえで、将来のビジョンを持つことが大切です。

もう少しかみ砕いて言うと、具体的にこれからどのような仕事をやっていきたいのか、そしてそれに向かって現在、自分は何に取り組んでいるのかと繰り返し問い続け、自分の言葉にして言えるようにすること、そして、自分の仕事が将来的にどうなっているのか、どのように役立っているのか絵図を自分の中でしっかりと描き、それをいつでも誰にでも説明できる状態であることが大事であると考えます。

皆さんもこれから出会う多数の人とのかかわり合いにより、自分が予想もしていなかった仕事に挑戦することになるかもしれません。現在、私は上席特別研究員とグループリーダーを兼務しています。もともと研究者として身を立てていくことを志していた中、マネージャーも拝命することになり、今は会社の目標に向かってチームをまとめしていく立場としても期待されています。この経験は自分自身が成長できる良い機会と、ポジティブにとらえて活動はしていますが、時には、「本来自分のやりたかったこと、やるべきことは一体何だったのか」と振り返り、自分を見つめ直すことがあります。

自分のやりたいことと所属組織の活動が一致しているときはよいのですが、常にそうであるとは限りません。自分がもともと考えていたビジョンやポリシーとの間に少しズレを生じてしまうことも現実的にはあるのです。しかし、ズレてしまったことを単に仕方がないと片付けて終わるのではなく、このズレの間にどういった関係があるのかを読み解き、その関係をしっかりと位置付けておくことが大切だと思います。そうした積み重ねが、将来、自分を振り返ったときに、自分自身の選択に納得感をもたらし、支える軸になってくれるはずだと信じています。

■参考文献

(1) <https://group.ntt.jp/newsrelease/2024/09/18/240918a.html>

* 2 変幻灯：光のパターンを投影することで、止まっているはずの画像が動いて見える印象を与えることのできる、NTTが開発した全く新しい発想に基づく光投影技術。この技術では人間の錯覚を巧みに利用しています。