

眼から読み取る心の動き ——Heart-Touching-AIのキー技術

「心を読む」技術は、人間の心と身体を理解し、サポートするHeart-Touching-AIの重要な構成要素です。「眼は心の窓」といった言葉のとおり、眼の示す瞳孔反応や眼球運動は、心の動きに関する情報をもたらします。本稿ではNTTコミュニケーション科学基礎研究所で行っている、眼から心を読み取る原理の解明と、現実問題へ適用可能な基礎技術の開発について紹介します。

ふるかわ しげと よねや まこと
古川 茂人 / 米家 惇

かしの まきお
Hsin-I LIAO / 柏野 牧夫

NTTコミュニケーション科学基礎研究所

心に響くAI技術

心と身体に充足感を与えてくれる幸福な社会を目指して、近年では、技術に求められる領域は、論理的な推論を行う知能だけでなく、人間の知性、感性、情動といった「心」にも及んでいます。機械学習などの技術発展により、表情や声の解析に基づいて人間の感情を認識するロボットも登場してきました。しかし、人間の感情や心理状態は、表情や声といった顕在的な情報に常に現れるとは限りません。また、顕在的な情報だけを用いては、人間が可能な範囲での感情認識を超えることはできません。NTTは、眼球運動や脈波などの無意識の身体反応を計測することで、「コンピュータでしか認識できない」人間の潜在的な心の動きを読み取り、「人間さえも超える」コミュニケーション能力を持ったAI（人工知能）を実現することを目指しています。

潜在的な心の動きを知るための基礎技術として、これまで脳計測に基づくBrain Computer Interface (BCI) という技術が発展してきました。しかしBCIには、拘束性の強さ、ノイズの大きさ、ロバスト性の低さなどの問題が

あるため、あくまでも医療などの限定的な利用にとどまっております。日常場面にまで普及しているとはいえない状況です。

NTTコミュニケーション科学基礎研究所では、脳計測に替わり、「眼」に着目し、心の動きを知るための基礎技術を開発しています。眼の情報は、カメラなどによって非侵襲的・非接触にとらえることができ、計測のために人間の自然な感覚や動きを妨げずにデータを取得できるという実用上のメリットもあります。

眼が表すもの

眼から得られる情報には、どのようなものがあるのでしょうか。瞳孔は虹彩に囲まれた黒く見える孔で、基本的には網膜に入射する光量を制御する役割を持ちます。明所で瞳孔径が縮小（縮瞳）し、暗所で拡大（散瞳）する対光反射はよく知られていますが、さまざまな認知プロセス（ターゲット検出、知覚、学習、記憶、意思決定など）と瞳孔径との関連を示唆する研究結果が次々と出てきています⁽¹⁾。瞳孔径は、互いに拮抗する瞳孔括約筋と瞳孔散大筋の活動のバランスによって決定されますが、それぞれの筋は、副交感神経

系および交感神経系の支配を受けています。瞳孔径は意識的（随意的）にコントロールすることはできません。眼球運動は、動眼神経に支配される外眼筋によって制御されます。眼球運動は、見たい対象に視線を向けるといった随意的なものばかりではありません。新規な刺激・顕著な刺激などに向けて無意識に視線を向けてしまうこともありますし、視線を固定している状態でも微小振幅のランダムな視線跳躍（マイクロサッカド）といったものが、不随意に生じます。この不随意的な眼球運動には、外部刺激に誘導される注意、潜在的な（対象に向けての視線変化を伴わない）注意の状態が反映されるといわれています⁽²⁾。マイクロサッカドは、その発生頻度によって特性が評価されることが多いですが、私たちはさらに、その詳細な動特性も評価対象に取り入れて成果を上げています。

瞳孔径・眼球運動の制御メカニズムにはさまざまな神経経路が複雑に関与しています。そのシステムには、認知プロセスと深くかかわりがあると考えられる神経核が含まれています。特に、脳幹に位置する神経核である青斑核（locus coeruleus）と上丘（superior colliculus）の働きが注目されていま

す。青斑核は、覚醒、不安、ストレス、注意、意思決定において何らかの重要な役割を果たしていると考えられ、その活動はノルアドレナリン経路を通してさまざまな脳部位に作用し、交感神経系の活動によって散瞳を引き起こします^{(1),(3)}。上丘は視線や空間的注意の方向制御に関与していると従来から知られていますが、近年では交感神経系、副交感神経系を介して瞳孔にも作用することが知られてきました^{(1),(2)}。上丘には、視覚・体性感覚・聴覚刺激に反応するニューロンが存在し、それらが刺激の空間的位置に対応して配列する「空間マップ」があることが知られて

います。上丘は、「低次」の感覚系からの入力を受け、その情報を大脳皮質感覚野などの「高次」の機構に伝達するだけでなく、逆に、大脳皮質感覚野、連合野といった高次機構からの入力も複数の経路を介して受けています。大脳基底核、青斑核からも投射を受けています。瞳孔径・眼球運動に基づく「心の読み取り」とは、この複雑な神経システムの出力に基づいて、背後にある認知プロセスを推し量ることといえます。

NTTの取り組み

本稿では、NTTコミュニケーション

ン科学基礎研究所の最近の取り組みを紹介します。瞳孔反応や眼球運動が、音の目立ちやすさ（顕著性）、予測からの外れ度合い（サプライズ）に対して特徴的に反応することが、複数の実験で確認されました。その成果の一部は本誌の過去の記事でも紹介しています^{(4),(5)}。音楽を考えた場合、音色やメロディなどの顕著性・サプライズは、その曲に対する親しみや好みを規定する1つの重要な要因であると考えられます。瞳孔反応や眼球運動から適切な多次元の特徴量を抽出し、機械学習技術を用いて曲への親しみや好みを推定するモデル（図1）を作成したとこ

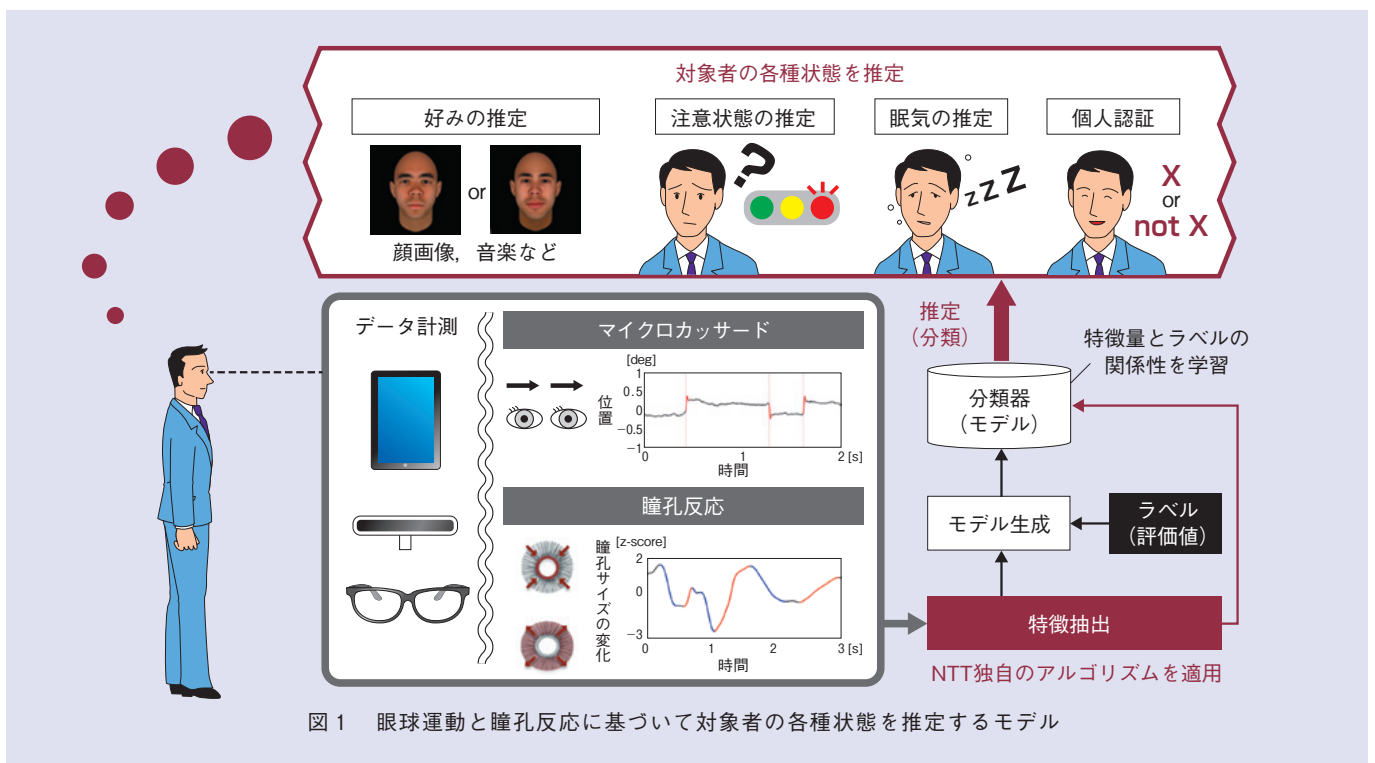


図1 眼球運動と瞳孔反応に基づいて対象者の各種状態を推定するモデル

ろ、聴取者の主観的な評価値をある一定の精度で推定できることが分かりました⁽⁴⁾。

同様なアプローチは、音楽の評価以外にも適用することができます。例えば、顔に対する好み（選好）です。ある実験では、瞳孔径を測りながら、顔の映像を観察者に示してみました。この場合、映像の輝度変化によって縮瞳が生じます。しかし、その縮瞳の大きさは、輝度変化の量を等しく設定していたにもかかわらず、呈示する顔によって異なりました（図2）。瞳孔計測実験の後に顔画像の好みについての主観評価実験を行ったところ、縮瞳の程度と主観評価とが対応することが分かりました⁽⁶⁾。前述と同様なモデルを用いて、この瞳孔反応や眼球運動の特

性も含めたさまざまな特徴量とを組み合わせることによって、顔の好みに関する個人ごとの判断を高い精度で予測できることも確認できています。

前述の実験で得られた瞳孔径と顔の好みとの相関は、顔画像の好みに関する認知プロセスが、瞳孔径の変化となって表れたと解釈するのが通常です。しかし、私たちは、逆の因果関係の可能性もあると考えています。つまり、個々の画像の何らかの特徴の違いによって縮瞳に違いが生じ、それが認知プロセスに作用する可能性です。それを確かめるため、顔画像ではなく、その背景画面の輝度変化を導入することによって縮瞳量を操作する実験を行いました⁽⁶⁾。その結果、実験的に導入した縮瞳量に伴って、顔画像の好みに

関する主観的評価が変わることが示されました。この現象は、技術的に重要な示唆を含んでいます。瞳孔は心を表すだけでなく、心に働きかけるための窓ともいえるのです。

今後の展望

これまでの研究では、主に音楽の好みや、顔画像の好みを推定対象としてきました。しかし、眼球の計測によって推定可能な対象は好みだけに限られたものではありません。冒頭で述べたとおり、眼球は自律神経系による支配を受けるため、快不快、眠気、緊張、恐怖など、さまざまな認知プロセスを潜在的に反映していると考えられます。私たちの技術のコアである特徴抽出、機械学習のフレームワークは、任意の推定対象に対して統一的に用いることができるため、（精度は別として）原理的には学習が可能と考えられます。ただし、学習のためには「正解データ」が「正しく」取得できることが必要条件です。例えば眠いときにはどんな眼球的動きが生じるかを学習するためには、実際に眠い状態にある人の眼球的動きを記録する必要があります。これまで、音楽の好みや顔画像の好みなどは、実験室でのアンケートにより取得可能であるため、比較的容易に正解データを集めることができました。しかし、恐怖や怒りなど、本人でも定量化が困難な対象や、倫理的な問題から実験室での制御、記録が困難な推定

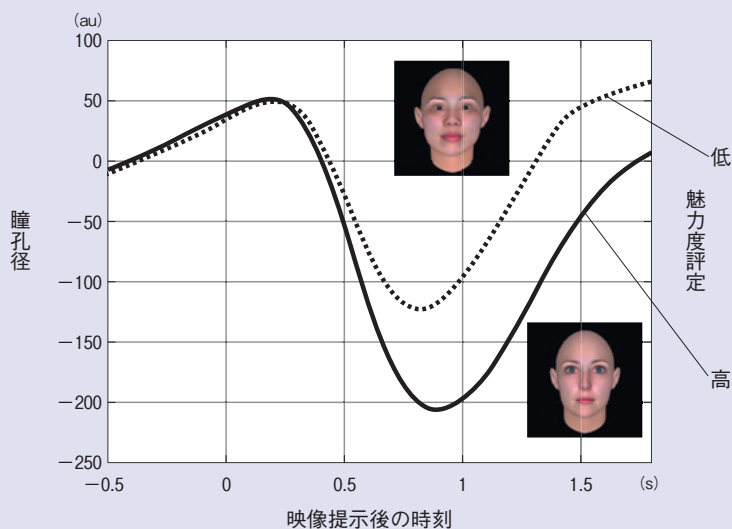


図2 顔映像に対する瞳孔反応の例

対象があるのも事実です。瞳孔反応・眼球運動が、さまざま外部的・内部的要因によって影響を受け、特定の認知プロセスや神経活動と単純な対応関係にないことも課題です。今後、無意識的な認知を的確に評価するためには、洗練された心理・行動実験や独立した生体反応計測が必要です。評価方法自体確立していないことが多い中で、これらの課題を解決していくためには、さまざまな計測・分析手法をその背後にあるメカニズムと制約に注意しながら使い、少しずつ理解の輪を絞っていくしかありません。NTTでは生理学、心理物理学から機械学習技術に至る広い分野の基礎研究とそれらの相互連携によって、課題に取り組んでいます。

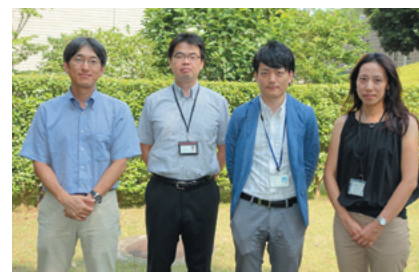
計測技術そのもののさらなる進歩も、もちろん必要です。瞳孔や眼球運動の計測は、カメラ撮像に基づく画像処理によって行われるものが主流です。現在、私たちが使っているスマートフォンやタブレット、PCなどにはWebカメラが搭載されており、眼球計測が可能となる基盤は身近に広がっているものと考えられます。ただし一般的なWebカメラのフレームレートは数10 fpsであることがほとんどです。高速度の撮像が可能になれば、例えばマイクロカッサードの動特性など、時間的に短い潜在的な応答をとらえることもできます。私たちの研究では、眼球計測専用の機器を用いて、1000 fpsでの撮像を行っていますが、この成果

を広く一般の機器で適用可能とするためには、カメラの高速化も1つのカギとなってくるでしょう。

基礎的な知見の蓄積、着実な実験に基づくデータの収集、計測技術の改良により、AIが人の眼を見て心を読む（あるいは心に働きかける）ことができるようになる日はそれほど遠くないと思われれます。

■参考文献

- (1) C. A. Wang and D. P. Munoz : "A circuit for pupil orienting responses: implications for cognitive modulation of pupil size," *Current Opinion in Neurobiology*, Vol.33, pp.134-140, August 2015.
- (2) S. Martinez-Conde, S. L. Macknik, X.G. Troncoso, and D.H. Hubel : "Microsaccades: a neurophysiological analysis," *Trends in Neuroscience*, Vol.32, No.9, pp.463-475, 2009.
- (3) G. Aston-Jones and J. D. Cohen : "An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance," *Annual Review of Neuroscience*, Vol.28, pp.403-450, 2005.
- (4) 柏野・米家・Liao・古川 : "身体から潜在的な心を読み解くマインドリーディング技術," *NTT技術ジャーナル*, Vol.26, No.9, pp.32-36, 2014.
- (5) 古川・山岸・Liao・米家・大塚・柏野 : "身体反応に現れる「聞こえ」とそのメカニズム," *NTT技術ジャーナル*, Vol.27, No.9, pp.13-16, 2015.
- (6) H. I. Liao, S. Shimojo, and M. Kashino : "Correspondence between pupillary response and facial attractiveness," Oral presentation at Asia-Pacific Conference on Vision, Singapore, July 2015.



(左から) 古川 茂人/ 柏野 牧夫/
米家 惇/ Hsin-I Liao

この分野は、基礎的な神経科学・心理物理学と機械学習技術の組み合わせによって進展しています。私たちのチームはまさにそのそれぞれの分野の専門家で構成されています。

◆問い合わせ先

NTTコミュニケーション科学基礎研究所
人間情報研究部
TEL 046-240-4795
FAX 046-240-4716
E-mail furukawa.shigeto@lab.ntt.co.jp