

[脳情報科学が拓く AI と ICT]

## ③ 脳情報を読み解く



西本伸志 | 情報通信研究機構 脳情報通信融合研究センター (CiNet)

### 情報システムとしての脳

ヒトの脳や感覚／運動系を情報処理システムとして捉えた場合、これらはどのようなスペックや特徴を持ったシステムとして記述できるだろうか。システムへの入力については、ヒトの脳は網膜や蝸牛などの各種センサから日常的に多様な情報を受け取っている。たとえば網膜神経節細胞から視神経を通して伝達される映像信号の伝達速度は、約 10Mbps と見積もられている。一方で出力については主に発声やタイピング、ジェスチャ等の伝達手段を介して行われるが、こちらはかなり甘めに見積もっても 1Mbps もないだろう。システムのインタフェースという観点で見た場合、脳はその入出力速度に大きな非対称性を持ったシステムであるといえる。入出力の間で行われる情報処理に関する特徴の 1 つとしては、脳内では具体から抽象へ、客観から主観へと、階層的あるいは段階的な情報処理／情報表現が行われる点が挙げられる。網膜や大脳初期視覚野と呼ばれる脳領野では、輝度や色、動きといった物理的実体を反映した情報が扱われている。そこから処理が進んだ高次視覚野と呼ばれる脳領野では、より高次の意味情報（たとえば建物やヒト）や印象情報（荘厳な建物、美しい女性）が扱われる。さらにそれらが脳内に蓄積されている情報と照合されることで、「去年行ったりゾート」だとか「今年もまた行きたい」といった記憶の想起や意思、行動、予測などのより主観的な情報の生起につながっていく。

従来の脳神経科学研究においては、上記の要素は

主には個別かつ実験的に単純化された条件下において研究されてきた（たとえば、顔の画像を 20 秒見せた後に建物の画像を 20 秒見せ、誘起した脳活動の差分を取るなど）。しかし、2010 年代に入り、fMRI (functional Magnetic Resonance Imaging ; 機能的磁気共鳴画像装置) や 2 光子イメージングを代表とした脳神経活動の大規模同時計測技術の発達、および大量のデータを定量的に扱う機械学習等の多変量解析技術の高度化に伴い、映画を見ているときなどのより自然で複雑な条件下における脳の情報表現を直接的に解明する研究が盛んに行われるようになった<sup>☆1</sup> (詳しくは脚注の総説を参照)。これにより、映像を見ている際の脳活動からその人が知覚している印象内容を解読したり、それを介したマーケティング応用等が進められたりするようになった。また脳内の情報表現を定量的に解明することで、いわゆる人工知能技術の精度向上を目指す研究も進められている。さらには、脳内の想起内容等を解読することでヒト脳の入出力速度を増大させようとするなど、ほとんど SF 小説のような計画も議論されている。本稿ではこれら最近の研究や話題について紹介する。

☆1 参考資料 (総説)

- (a) Gallant, J. L., Nishimoto, S., Naselaris, T. and Wu, M. C. K. : System Identification, Encoding Models and Decoding Models : A Powerful New Approach to fMRI Research, In Kriegeskorte, N. and Kreiman, G. (eds.) : Visual Population Codes (pp.163-188), Cambridge : MIT press (2011).
- (b) 西本伸志 : エンコーディングモデルを用いた視覚情報処理研究 : 情報表現, 予測, デコーディング, 日本神経回路学会誌 19 (1), pp.39-49 (2012) .



## 脳内情報表現の解読

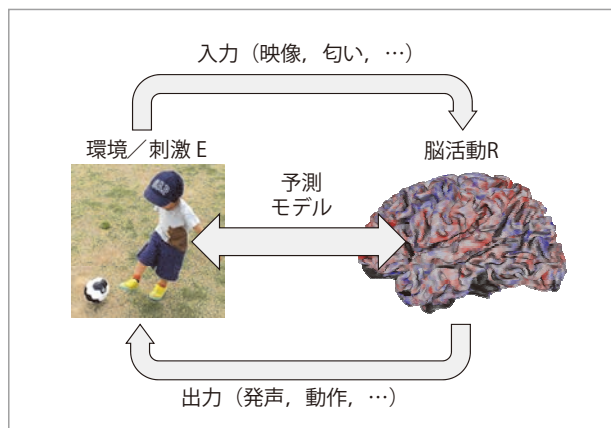
脳が感知する外界の状況や刺激を E とし、それによって生起する、もしくは内在的なダイナミクスによって変化する脳活動を R とする (図-1)。この枠組みにおいて脳の働きを理解するための定量的な方法論としては、共に多次元時空間情報である E と R の関係性を解明する (両者の関係を記述する予測モデルを構築する) ことが挙げられる。たとえば映像視聴下の脳活動について考えてみると、E は提示した映像 (輝度の時空間パターン) であり、R は誘起脳活動 (fMRI や多点電極等で計測した脳活動の時空間パターン) である。両者の関係性が分かれば、脳のどのような場所でどのような情報が符号化 (エンコード) されているかを定量的に理解することにつながるだろう。また逆に、脳神経活動を逆符号化 (デコード/解読) することで、その人が何を感じて/考えているかを推定することも可能になる。

上記のような研究の具体例として、初期視覚野と呼ばれる大脳皮質後頭葉の情報から脳内の映像情報を解読した例を紹介する<sup>1), 2)</sup>。初期視覚野においては、映像情報が視野内の局所的な時空間周波数成分 (運動エネルギー) の集合として表現されていることが知られている。このため、周波数成分で張る特徴空間を媒介とすることで、脳が知覚する映像情報と脳活動の関係を説明する予測モデルを構成する

ことができる。このようなモデルを介して、たとえば動画視聴下の脳活動から知覚する映像を推定したり<sup>1)</sup> (図-2)、映像想起下の脳活動から想起映像の内容を推定することが可能になっている<sup>2)</sup>。従来、いわゆる脳・機械間インタフェース (Brain-Machine Interface ; BMI) と呼ばれる脳活動解読を介した意思伝達手段は、主に運動意図を対象に研究が進められてきた。将来的にはこのような視覚的な情報に関する脳活動解読を行うことで、より伝達効率の良い BMI が実現する可能性がある。

上記の実用化を阻む直近の課題の1つとしては、先述の想起映像解読<sup>2)</sup>を例にとると、その実現に7T MRI 等の比較的高額 (7億円程度) の計測機器を要することが挙げられる。しかし情報技術の進展の歴史を見ると、技術革新によって3~4桁程度の価格低減が起きることは珍しくない。そのような計測技術上の革新が起きたとして、その際に利用可能な神経科学と数理解析の基盤は、整いつつある。

さらに処理が進んだ高次視覚野と呼ばれる大脳皮質側頭葉を中心とする脳領野では、映像から一步処理が進んでその意味内容 (きれいな建物や走る車等) が表現されていることが知られている。近年の研究において、これらの主に言語 (単語) で表現できるような脳内情報表現については、各種の自然言語処理技術を用いることで定量的な扱いが可能になることが知られている。例としては、視聴映像と誘



■図-1 環境/刺激と脳のインタラクションの概念図

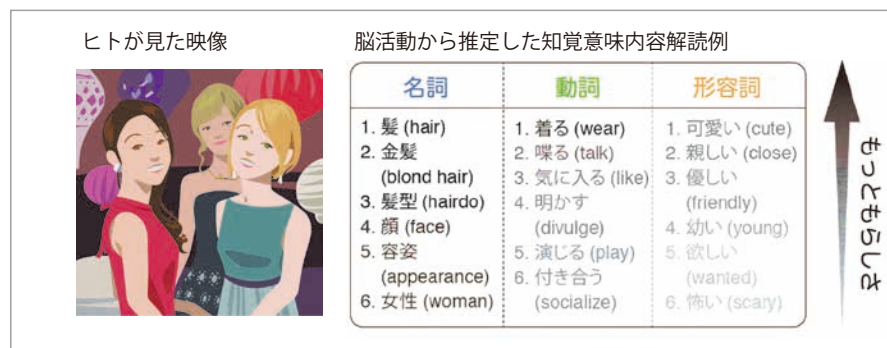


■図-2 ヒトが見た映像 (左) と脳活動から推定した知覚内容映像化例<sup>1)</sup>。運動エネルギー表現と呼ばれる脳内情報表現を介したモデルを構築することで脳活動から知覚内容を一定精度で解読。映像は著作権侵害回避のためイラスト化

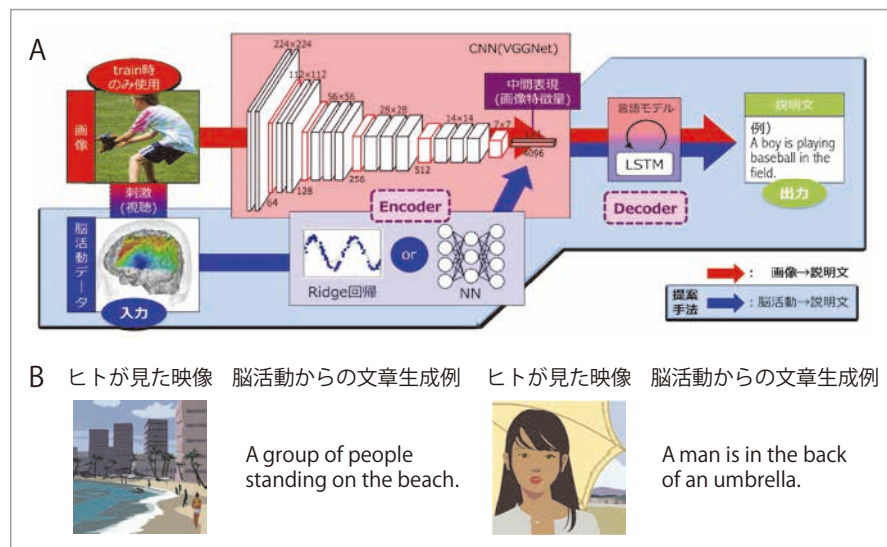
起脳活動の関係を word2vec と呼ばれる大規模コーパスから単語のベクトル表現を学習する仕組みを媒介してモデル化することで、脳活動からヒトの知覚意味内容を推定することも可能になっている<sup>3)</sup> (図-3)。この例では脳活動から上記のベクトル空間への対応関係を学習することで、名詞(物体)、動詞(動作)といった比較的客観的な内容から、形容詞(印象)で表現される比較的主観的な内容まで、約1万語の単語について知覚内容の有意な推定が可能になっている。またこのような技術の応用例として、映像素材等に関する印象定量評定を行う商用サービスの提供が始まっている。たとえば、テレビコマーシャルは視聴者に一定の印象を与える意図を持って作られるものであるが、上述のような脳活動解析技術を用いることで、意図した印象が正しく伝わっているかについて一定精度の定量的推定を行うことが可能になる(このような技術の社会実装の実際やそのより具体的な位置づけ等については、本号に萩原一平先生が寄稿された内容を参照されたい)。

より先進的な試みとして、脳活動から知覚内容を文章として解読する研究も進められている<sup>4)</sup> (図-4)。このような研究は、いわゆる人工知能技術、その中でも画像からその内容に関する説明文を生成する Neural Image Caption Generator (図-4A上) と呼ばれる技術の進展の影響が大きい。この技術では、画像からその特徴を抽出する畳み込

み神経回路(Convolutional Neural Network; CNN)と画像特徴から文書系列を生成する再帰型神経回路(Recurrent Neural Network; RNN)を組み合わせることで、画像入力からその内容を説明する文章を出力する。このうちCNNについては元々は脳視覚野の階層的情報表現を模したものであり、そこで扱われる特徴表現はヒト大脳皮質のそれと一定の対応関係がある。このため、任意の脳活動からCNNの高次表現を推定することが可能であり(図-4A下)、そこからRNNへの経路を利用することで文章を生成することが可能になる。このようにして得られた文章例を図-4Bに示す。文章の正確さなど



■図-3 ヒトが見た映像(左)と脳活動から推定した知覚意味内容解読例(右)<sup>3)</sup>。自然言語特徴空間を用いたモデルを構築することで、約1万語の候補の中から尤度の高い単語を品詞ごとに抽出。映像は著作権侵害回避のためイラスト化



■図-4 A:脳活動解読による文章生成の概念図<sup>4)</sup>。大量に利用できる画像-説明文ペアでネットワーク学習を行った後、脳活動からCNN高次層への関連を追加学習する。B:映像視聴中の脳活動から推定した知覚内容の文章解読例。映像は著作権侵害回避のためイラスト化



についてまだまだ改善の余地があるが、これはヒトの脳活動から知覚内容を文章として取り出した世界で初めての例である。このような技術は、将来的には頭の中で想起した内容をそのまま文章化するなどの高効率 BMI の数理基盤となる可能性がある。

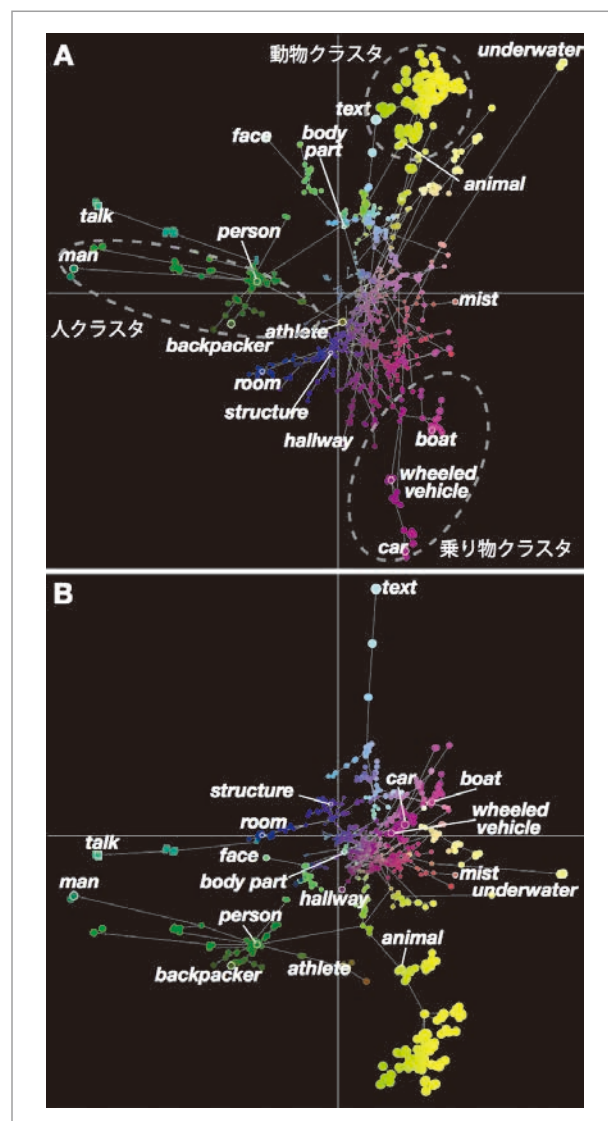
## 脳内情報表現と人工知能

多様な外界の状況 E と脳神経活動 R の対応を定量的に理解することは、脳内でどのような情報がどのように表現されているかを解明する基盤となる。このような観点から、脳内で多様な事物がどのように構造化されて表現されているかを示す脳内意味空間の定量を行った例を図-5 に示す<sup>5)</sup>。この図においては、各点が単語で表される事物（ヒト、電話、山、など）を示し、各点間の距離が脳内でそれらの事物がどの程度似たものとして表現されているかを示す。この空間においては、たとえば動物（ネコ、サルなど）、ヒト（男、バックパッカーなど）、乗り物（車、ボートなど）はそれぞれ類似概念としてクラスタ状に分布していることや（図-5A）、テキスト（文字）は特殊な事物としてほかの事物とは孤立して表現されていることなどが分かる（図-5B）。

上述のような脳内意味空間は、ヒト脳が表現する事物の相対的な関係性を定量可視化したものであり、脳を持つ世界観や常識をデータドリブンに抽出したものと考えることができる。端的な例を挙げると、図-5 の意味空間において、ヒトと動物は別々のクラスタを構成しており、これはヒトと動物は異なるものという脳の常識が表現されていると解釈できる。これは、生物学の教科書で教えらる常識（ヒトは動物の一種）とは異なる、脳が持つある種の偏見といえよう。このような空間定量の応用例としては、人工知能への援用が挙げられる。つまり、もしこのような脳内意味空間に似た情報表現を行う人工知能を構築することができれば、それはよりヒトの常識に寄り添った判断をされると考えられる。このような

観点から、深層学習等によって得られる情報表現にヒト脳活動由来の摂動を加えたところ、一部の言語課題や物体識別課題について脳活動を使わなかった場合よりも性能が向上したとする研究が報告されている<sup>6)</sup>。

上記のような神経科学と機械学習の融合的研究は興味深いものであるが、このような研究のさらなる精緻化を阻む一因としては、脳活動について学習に足る十分なサイズのデータを得ることが困難である



■ 図-5 動画視聴下脳活動から定量化した脳内意味空間<sup>5)</sup>。各点が事物（顔、部屋等）を表し、点間の距離が脳内情報表現における類似度を表す。A, B はそれぞれ別角度から主要次元を 2 次元投射したもの

点が挙げられる。一般的な脳神経科学の実験においては、特定の情報表現に関する仮説検定を行うため、高度に厳選した数サンプル～数百サンプルの条件（刺激／タスク）を用いて脳活動データ取得が行われている。このような少数サンプルから汎用的な知見をデータドリブンに得ることは困難である。筆者らが2014年に公開した自然動画刺激下ヒト脳活動データ<sup>☆2</sup>においては、個体あたりサンプルが約8,000と、従来に比べて1～2桁大きなデータを提供している。これはヒトオープンデータとしては世界最大級であり、教育・研究を用途として国内外で利用されている。しかし、画像認識においてImageNetが果たしたような大規模汎化学習を促進する役割（コンテスト用に約百万サンプルを提供）を担うには、まだデータサイズが数桁足りない。そのような大規模データの構築が、今後の脳神経科学と人工知能のさらなる融合を進めるにあたって解決すべき課題の1つであると考えられる。

## 情報通信とBMIをめぐる近況

2017年に入り、ICT分野における企業の脳活動解読分野への参入表明が相次いだ。大きなところでは、Tesla社やSpaceX社の創始者として有名な投資家Elon Musk氏が侵襲型（つまり脳外科手術による電極埋め込みを必要とする）BMI開発を目指すNeuralink社の設立を発表した。同社は4年程度で麻痺患者等を対象とした臨床応用、8～10年程度で健常者を対象としたBMIの実現を目指すとしている。現在の倫理基準では臨床的必然性のないヒト脳内電極の埋め込みは原則的に認められておらず、同社の主張はかなり挑戦的といえる。またSNS世界最大手であるFacebook社は、元米国国防高等研究計画局（DARPA）所長／Google ATAP 副所長のRegina Dugan氏を中心とした非侵襲型BMIの開発

を発表した。同社は2年で100単語／分の脳情報伝達実現を目指すとしている。現時点において実現されているBMIの伝達速度がその1～2桁程度低いことを考えると、こちらも非常にインパクトの強い計画といえる。さらには孫正義氏の率いるSoftbank社が最近買収した大手半導体企業ARM社も、脳埋め込み型チップを開発すると表明している。

これらの著名ICT企業によるBMI開発計画が予定通り進展する可能性については、筆者も含めて懐疑的な専門家が多い。しかし、一般的な大学の研究室に比べ遥かに豊富な資本を集中投資することで、革新的な技術が生まれる可能性は考えられる。特に脳埋め込み型チップについては、小型化、ワイヤレス化、省電力化、熱設計最適化等、実現に向けた技術的な課題は多いが、ARM社はまさにそのような設計技術の専門家集団である。同社の参入は、分野としても朗報であるといえる。侵襲的な脳外科手術を伴う手法の是非は、技術的な課題であると同時に社会的な需要と受容の問題でもある。ただ、現時点でもたとえば美容整形等については臨床的必然性が必ずしもない侵襲手術が日常的に行われている。10年後、20年後の社会規範が現在の私たちと同じかどうかは、未知である。解読技術の高度化に関する進展を予測することは難しいが、たとえば前述した脳活動の文章解読<sup>4)</sup>（図-4）においては、2秒間の非侵襲脳活動計測データから1文を生成している。将来的にはこのような人工知能／自然言語処理技術の援用により、高効率な脳活動解読が実現する可能性はある。

上記のICT企業の発表に共通している点は、彼らは脳情報解読分野への参入を表明はしているものの、脳自体には（おそらく）興味がないことである。彼らの主眼はあくまで情報伝達の未来であり、主にはヒト脳からの情報出力の向上を目標としている。冒頭でも述べたとおり人間の情報出力の速度は現在の情報通信の水準に比べてあまりに遅く、たとえば今後台頭するであろう人工知能等のエージェントと

<sup>☆2</sup> <https://crcns.org/data-sets/vc/vim-2/about-vim-2>



比べて圧倒的に非力である（自分が1万語しゃべる間に1語しか応答しない隣人がいたら、どのように認識するだろうか）。あるいは、現在の携帯電話通信網は脳のごく近くまで高速無線通信を実現しているのに、脳に至る最後の、あるいは最初の数センチについては空気の振動（音声）や数本の指の動き（タイピング）などの低速伝達手段に依存している。この事実には、ある種の滑稽さを見ることもできるだろう。それが10年で克服できるかは分からないが、可能性について議論するための技術的な土台の一部は整いつつあり、興味深い時代ではある。

#### 参考文献

- 1) Nishimoto, S., et al. : Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies, *Current Biology* 21, pp.1641-1646 (2011).
- 2) Naselaris, T., et al. : A Voxel-wise Encoding Model for Early Visual Areas Decodes Mental Images of Remembered scenes, *NeuroImage* 105, pp.215-228 (2015).

- 3) Nishida, S. and Nishimoto, S. : Decoding Naturalistic Experiences from Human Brain Activity Via Distributed Representations of Words, *NeuroImage*, in press (2017).
- 4) Matsuo, E., et al. : Generating Natural Language Descriptions for Semantic Representations of Human Brain Activity. *Proc. ACL SRW 2016*, pp.22-29 (2016).
- 5) Huth, A. G., et al. : A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories Across the Human Brain, *Neuron* 76, pp.1210-1224 (2012).
- 6) Fong, R., et al. : Using Human Brain Activity to Guide Machine Learning, *arXiv 1703.05463* (2017).  
(2017年9月30日受付)

西本伸志 nishimoto@nict.go.jp

大阪大学基礎工学部生物工学コース飛び級中退。同大学院基礎工学研究科修了。博士（理学）。カリフォルニア大学バークレー校に勤務後、2013年より現職。脳情報の定量と解読に関する研究に従事。

