

令和元年6月19日現在

機関番号：82401

研究種目：基盤研究(A) (一般)

研究期間：2014～2017

課題番号：26240021

研究課題名(和文) 顔のホリスティック表現の実体を顔細胞応答から探る

研究課題名(英文) Investigation of holistic representation of faces from visual features encoded by face neurons in the brain

研究代表者

谷藤 学 (Tanifuji, Manabu)

国立研究開発法人理化学研究所・脳神経科学研究センター・チームリーダー

研究者番号：60197530

交付決定額(研究期間全体)：(直接経費) 32,300,000円

研究成果の概要(和文)：視覚的に提示された顔画像は、目、鼻、口など顔のパーツに分解されているのか、あるいは、パーツの配置の情報が重要なのか？、脳の顔に反応する細胞がコードしている特徴の自然画像の断片による近似によってこの問題に取り組んだ。得られた特徴は、顔細胞の顔の向きに対する選択性をよく再現し、さらに複数の細胞の特徴を組み合わせることで、顔の向きに非依存的な個人の顔の弁別が可能になった。これらの特徴には顔の中の局所的な特徴に依るものと局所的な特徴の間の空間的な配置を捉えているものがあったが、局所的な特徴は顔パーツそのものではなく、パーツ、髪の生え際、フェースラインなどによって作られるコントラストの境界であった。

研究成果の学術的意義や社会的意義

脳において、視覚物体像は何層にもわたる多層ネットワークを経て処理されている。私たちの研究は、この多層なネットワークが低次の局所的なエッジを抽出する初期層と局所的な情報の特定の組み合わせを作る層の2層によって近似できることを示した。この結果は、脳における視覚物体処理の本質の理解に大きく貢献している。近年、AIの一分野として、視覚物体処理を行うネットワークの開発されている。しかし、それらのネットワークは多層構造を持つため、中で行われている演算はわからないという問題点がある。我々の研究で用いた手法はAIにも適用できるので、AIの解析に大きな貢献をすることが期待される。

研究成果の概要(英文)：How faces are decomposed in the brain? Faces may be decomposed into facial parts, such as eyes, nose and mouth, and/or elements representing relational information of facial parts. We addressed this question by approximating visual features encoded by face neurons as natural image fragments. Resulting features reproduced response preference to particular views, and combination of them made view invariant face identification possible. Visualizing sub-region of faces captured by these features revealed that our brain decomposed faces into local features and complex features each of which detects a spatial arrangement of the local features. However, these local features are not explicit representation of facial parts. They are edge and color contrast derived from facial parts, hairline, and face lines. We suggest that the deep layered ventral visual pathway can be approximated by a shallow network comprising layers of low level processing and of combining low level information.

研究分野：システム神経科学

キーワード：物体認識 顔認知 深層学習 AI DCNN

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

ヒトは優れた顔認識能力を持っている。表情や向きなどによって見え方が変化しても、顔から人物を特定することができる。しかし、このような優れた顔認知の能力は、顔を逆さにすると失われてしまう、或は、減弱することが心理学の研究からよく知られている。顔のイメージは脳の中でどのように分解されて処理されているのだろうか？

物体像を表現する枠組みとして、物体像を「パーツ（部分）」の表現と「パーツとパーツの空間的な位置関係」の表現の2つに分ける考え方がある。しかし、顔の表現は単純にこの枠組みでは説明できない。例えば、Tanaka と Farah (1993) は、様々な条件の下で、顔のパーツが誰のものかを問う課題を被験者に課した（図4）。被験者の正答率をみると、パーツ（例えば、鼻）だけでは誰のパーツかわかりづらいが、当然のことながら本人の顔の中にそのパーツがあればすぐわかる。しかし、すべてのパーツがあっても、そのパーツの配置がでたらめだったり、あるいは、顔を逆さにしたりすると、正答率は大きく低下するのである。すなわち、パーツが揃っているだけではだめで、パーツが正しく配置（二つの目の下に鼻が、鼻の下に口が並ぶ）されていないと、パーツの持っている情報を使って顔を認識することができない。顔の認識においては、パーツの持つ情報（例えば、個人を識別するための特徴）は、パーツの配置の情報（目鼻口の配置）と不可分なのだ。このことを指して、彼らは顔の「holistic」な表現と呼んだ。しかし、「holistic」表現は心理実験の説明として考えられた操作的な概念なので、研究者の間でもその解釈を巡って多くの混乱がある。

さて、ヒトと似た視覚系を持つサルの高次視覚野（下側頭葉視覚連合野 TE）には顔に特異的に反応する細胞がある。他の TE 野の細胞と同様に、同じ反応選択性を持つ顔細胞は集まってコラムを形成している。顔コラムや顔細胞がコードしている特徴が分かれば、顔細胞の性質から心理実験から提案する「holistic」な表現の実体がわかるのではないかと考えた。

2. 研究の目的

本研究のユニークな点は、顔を我々がどのように分解して理解しているかという問題を、脳の中で顔の視覚情報処理を担っている神経細胞の応答特性から理解しようという点である。人と同じ霊長類であるサルの脳には、人とよく似た視覚情報処理機構がある。目に移る視覚像は、大脳皮質の初期視覚野でまず局所的な方位を持つ線分と色の情報に分解される。初期視覚野の細胞の働きによって、自然画像から物体の輪郭部分など濃淡の大きく変化する部分が抽出されるが、個々の細胞は画素空間上の位置の情報を持っている。視覚情報は、初期視覚野を経たのち、下側頭葉視覚連合野（TE 野）へと運ばれる。TE 野の細胞は、初期視覚野のような物体像の局所的な線分のような単純な情報ではなく、もっと複雑な図形特徴に対して反応する。また、その図形特徴が視野のどこにあっても反応する。すなわち、視野の位置の情報を失う。大まかにいうと、目に映る物体像は、初期視覚（初期視覚野）における単純な特徴の視野空間の表現から、高次視覚（TE 野）において視野情報を持たない複雑な図形特徴の組み合わせの表現に変換される。数学的に表現すれば、物体像は、初期視覚においては空間座標で張られた高次元空間の一点であり、高次視覚においては様々な図形特徴で張られた高次元空間の一点である。ただし、TE 野の細胞が複雑な図形特徴に反応することは、様々な研究から示唆されているが、図形特徴の記述は定性的で、具体的な形として確定していない。物体像が視覚情報処理の過程でどのように分解されているのかという問いの答えは、まさに、この図形特徴を明らかにすることに他な

らない。

さて、TE 野には顔によく応じる細胞の一群がある。これらの細胞は顔だけにしか応じないわけではないが、たくさんの顔と非顔の物体像に対する応答を比較すると、有意な差があるので、顔細胞と呼ばれる。また、顔を含む様々な物体像に対する応答特性がよく似た細胞は、脳表面から垂直に伸びる円柱状の領域に集まっている。この円柱状の構造を顔コラムと呼ぶ。

我々の研究の目的は、TE 野のそれぞれの顔コラムが符号化している図形特徴を明らかにすることで、顔のイメージが脳の中でどのように分解されているかという問題に示唆を与えることである。具体的には、(1) 多くの顔コラムから視覚刺激応答を記録し、(2) 開発した特徴抽出法により、それぞれのコラムの図形特徴を決定し、(3) 図形特徴が顔のどこに対応するかを明らかにした。

3. 研究の方法

TE 野の個々の顔コラムが符号化している図形特徴を先験的な知識を導入することなく推定するために、我々は数多くの自然画像の断片を用意すれば、その中からターゲットとなる図形特徴を見つけ出すことができるという仮定を置いた。そもそも私たちの視覚系が自然画像の分析のためにあることを考えるとこのような仮定を置くことはそれほど不自然ではない。また、前述のように、顔に反応する細胞であっても非顔の物体に全く応じないわけではない。したがって、画像断片を切り出す自然画像として、必ずしも顔だけではなく、様々なカテゴリの物体を含んだものを用意した。さて、このようにして用意された図形特徴の候補の一つが神経細胞の符号化している図形特徴であるとすれば、その図形特徴から推定される物体像に対する応答と実際の神経細胞の物体像応答の間には相関があるはずだ。そこで我々は、多くの物体像に対する顔コラムの応答を記録し、得られた多くの物体像に対する神経応答を対象に、候補となるすべての画像断片から推定された物体像に対する応答と実際の神経細胞の応答の相関係数が最大になるものを探索した。

一つの画像断片について、ある物体像に対する応答の推定には、(1) 画像断片と物体像の両方を局所的な方位と色の空間に変換した後、(2) これらの間のユークリッド距離を計算し、(3) それを動径基底関数に通すことで0と1の間に規格化したものを用いた。局所的な方位と色の空間に変換したのは、初期視覚野での処理を考慮したからである。

また、視覚刺激物体は、一般に画像断片より小さい。細胞は、物体像の中でその細胞が符号化している図形特徴がある部分に対して反応していると考えられるので、刺激物体の中で、最もユークリッド距離が小さくなる画像断片の位置での距離から応答を計算する。この手法を前節で述べた顔に反応する65個の顔に反応する記録部位に適用し、611個(300個のコントロールされた向きを持つ顔は除かれている)の視覚刺激物体の応答について、神経応答との間に統計的に有意な相関を持つ図形特徴を抽出することに成功した(図1)。このようにして推定された図形特徴は、推定に用いられなかった611個の視覚刺激に対する応答についても有意な相関を示した。

さて、顔細胞は、ある特定の顔の向きに対して選択的に反応するというユニークな特性を持っている。例えば、図2の実線は、2つの神経活動部位について得られた顔の向きに対する神経応答特性である。左の活動部位は、正面の顔に、右の活動部位は左右の横顔に

応答する（ただし、左向きの横顔の方が応答は強い）という顔の向きに対する特性を持っている。同定された図形特徴からこの顔の向きに対する応答を推定すると、この顔の向きに対する応答も再現できた（図2，点線）。私たちは、70%の記録部位について、図のように両者の間により相関関係があるという結果を得た。この結果は、得られた自然画像の断片の中に、顔細胞の向きに関する応答特性を説明するような性質があることを示している。

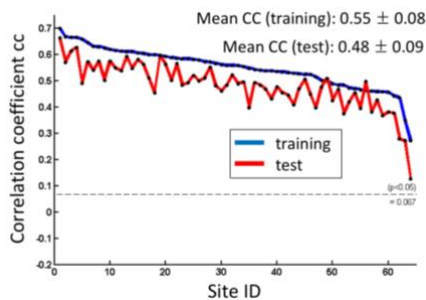


図1 推定された図形特徴と神経応答との相関。(青) 611 個の物体像 (トレーニングセット) に対する神経応答ともっとも高い相関を示す図形特徴との間の相関係数。(赤) 推定に用いられなかった 611 個の物体像 (テストセット) に対する神経応答とトレーニングで推定された図形特徴の物体像応

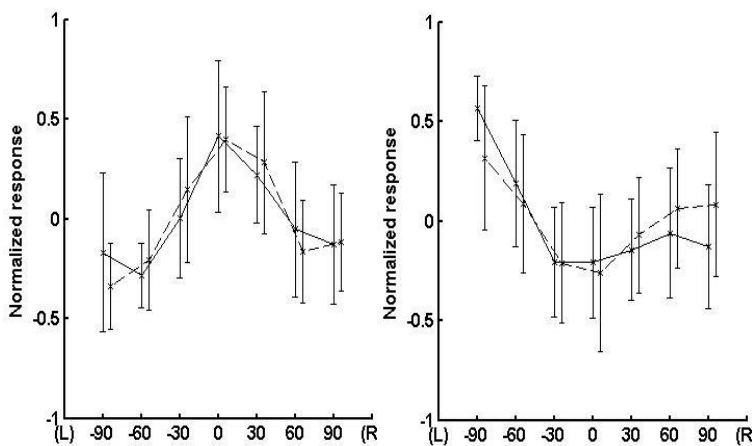


図2. 様々な向きの顔に対する顔細胞の神経応答 (実線) と推定された図形特徴から予想された向きに対する応答 (点線). 2つの記録部位の神経応答について例示してある. 横軸は、左向き (-90 度) から右向き (90 度) を、縦軸は、それぞれの顔について、各向きに対する応答と標準偏差で z 正規化した応答を示している.

4. 研究成果

開発した手法を用いて、顔カラムが符号化している多くの図形特徴を明らかにした。図3はその一例である。図3Aは、あるコラムの応答特性をもっともよく説明する画像断片である。左に元となる自然画像を示したが、上に述べたように、与えられた画像断片に対して物体像に対する応答を推定するときは、局所的な方位と色に変換してから、計算を行う。したがって、右のように方位と色の空間で表現するのがより適切である。図に示すように複数の局所的な方位を含む図形特徴が見出された。それぞれは、目、鼻、口、および、ヘアライン、顔の輪郭に由来する局所的な方位になっている。例えば、図3の例であれば、2つの水平の方位が目の部分と口の部分に対応している。しかし、これを、顔のパーツの空間的な関係の表現であるという、少し言い過ぎであるかもしれない。まず、局所的な方位は、目鼻口ばかりでなく、ヘアラインや顔の輪郭に由来するものもある。また、これらの局所的な方位はそれぞれのパーツの存在によって生じる濃淡が作る局所的な方位であるからだ。例えば、図形特徴の元となる自然画像断片にこの特徴は含まれている。明示的な顔のパーツの表現の組み合わせの表現になっていないことは、顔のパレイドリアが生じるメカニズムの一つになっているかもしれない。一方、単一の局所的な特徴を表現していると考えられる顔カラムも見出された（図4）。この例では、図形特徴は2つの異なる色の

コントラストで構成される45度の輪郭から成り (A)、人の横顔のヘアラインを検出している (B)。このように局所的な特徴を表現しているカラムにおいても、顔のパーツを明示的に表現していると考えられるものは見出されなかった。

我々の研究の結果を、我々が顔をどのように分解して理解しているかという問題と結びつけてまとめると次のようになる。

(1) 顔カラムが符号化している特徴には、局所的な方位と色のコントラストを表現していると考えられるものと、このような局所的な特徴の組み合わせを表現している特徴があった。

(2) 前者について、いわゆる顔のパーツを表現していると思われる顔カラムは見出されなかった。局所的な顔カラムの応答は、これらのパーツが存在することによって生じる濃淡の境界に反応していると考えれば十分に説明できる。このような特徴は顔に多く見られるが、必ずしも顔だけに見られる特徴でない。その意味で特徴は generic である。

(3) 後者について、局所的な特徴の組み合わせは、顔のパーツ、ヘアライン、フェースラインによって生じる局所的なコントラストの組み合わせになっていた。その意味では、いわゆる「holistic」な表現の実体に近いものであるかもしれない。

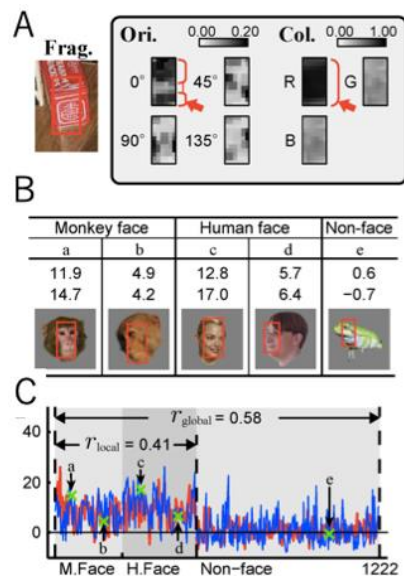


図3. ある顔カラムの図形特徴 (A) と図形特徴から推定される応答と実際のカラムの神経応答 (B、C) C に調べられた全ての物体像に対する神経応答 (赤) と図形特徴から推定された応答 (青) の比較を示した。B は、代表的な物体像について数値を具体的に示した。上段の数値が神経応答。下段の数値が推定された応答である。下の写真は視覚刺激物体像で、赤い枠の部分が、この顔カラムが検出している顔の部分である。a, b, c, s はCの矢印に対応。

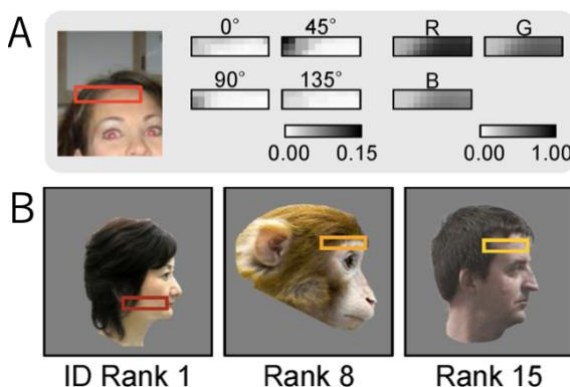


図4. 局所的な特徴を符号化していると考えられる顔カラムの一例。(A) 推定された図形特徴。(B) 顔カラムが検出している顔の部分。

5. 主な発表論文等

[雑誌論文] (計 2 件)

1. Owaki, T., Vidal-Naquet, M., Nam, Y., Uchida, G., Sato, T., Câteau, H., Ullman S., Tanifuji, M. (2018) Searching for visual features that explain response variance of face neurons in inferior temporal cortex. PLoS ONE 13, e0201192 (査読あり)
2. Obara, K., O’Hashi, K., Tanifuji, M. (2017) Mechanisms for shaping receptive field in monkey area TE. J. Neurophysiol., 118, 2448–2457 (査読あり)

[学会発表] (計 4 件)

1. Tanifuji M., “Machineries for forming center and periphery of receptive fields in monkey anterior inferior temporal cortex” IEEE International Symposium on Video and Audio Signal Processing in the Context of Neurotechnologies St. Petersburg Russia 7 2017
2. Tanifuji M., “Searching for visual features driving face neurons in the higher visual cortex, inferior temporal cortex, in macaque monkeys” IEEE International Symposium on Video and Audio Signal Processing in the Context of Neurotechnologies St. Petersburg Russia 6 2017
3. Tanifuji M., “View angle maps inside the face selective sub-region of anterior inferior temporal cortex” The 6th FAONS Congress and the 11th Biennial Conferences of CNS Wuzhen, China 9 2015
4. Tanifuji M., “Underlying Cortical Function and Structure for Object Recognition” The 5th Cinet General Conference Osaka 10 2015

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

6. 研究組織

(1)研究分担者

なし

(2)研究協力者

なし