

令和 5 年 5 月 22 日現在

機関番号：14401

研究種目：研究活動スタート支援

研究期間：2021～2022

課題番号：21K21309

研究課題名（和文）踊りに関するヒトの情報処理の包括的解明

研究課題名（英文）Whole-brain modeling on dancing in the human brain

研究代表者

高木 優 (Yu, Takagi)

大阪大学・大学院生命機能研究科・助教

研究者番号：20828553

交付決定額（研究期間全体）：（直接経費） 2,400,000 円

研究成果の概要（和文）：ダンスは古来、全世界で親しまれてきた音と動きの総合芸術である。ダンスに関する先行研究は主に、ダンスの特定の振り付けや動作に注目し、それらと各種認知機能や脳活動との関係性、さらには熟練による変化を調べてきた。一方で、ダンス視聴中にヒトが処理している多様な情報が個々の研究では無視されてきており、包括的な理解ができていなかった。本研究では、新規データセット構築から脳活動計測実験、脳活動解析手法の開発までを行うことによってダンスに関するヒトの情報処理を包括的に明らかにすることに成功した。

研究成果の学術的意義や社会的意義

感情や技術点などの高次の情報が付与された、新たなダンスの大規模データセットを構築した。加えて、これまで難しかったダンス動画からの感覚情報及びそれらを統合した情報の自動的抽出を、大量のダンス動画を用いて訓練された深層生成モデルを用いることによって解決した。それらを用いて、ダンス視聴中のヒト脳活動で処理されている多様な情報処理の詳細を明らかにした。並行して、深層生成モデルを用いた、多様な情報を柔軟に統合して人の認知体験を再構成するための基礎技術開発も行なった。本研究は世界中60以上のメディアによって報道された。

研究成果の概要（英文）：Dance is the ancient total artwork that has been practiced by humans across cultures for thousands of years, and the cross-modal combination of movement and music can trigger diverse cognitive and emotional experiences. Previous studies on dance have mainly examined the relationship between specific movements and cognitive functions or brain activities. Therefore, the previous studies have ignored the diverse information that the human brain processes during dance viewing, and have failed to comprehensively understand the relationship between dance and various cognitive functions and brain activity. In this study, we modeled various information processing related to dance in the human brain by constructing a whole-brain modeling using large dance data set with comprehensive feature extraction with deep neural network, and developing a new brain analysis method.

研究分野：システム神経科学

キーワード：fMRI 舞踊

## 1. 研究開始当初の背景

ダンスは古来、全世界で親しまれてきた音と動きの総合芸術である。ダンスに関する先行研究は主に、ダンスの特定の振り付けや動作に注目し、それらと各種認知機能や脳活動との関係性、さらには熟練による変化を調べてきた。一方で、ダンス視聴中にヒトが処理している多様な情報が個々の研究では無視されてきており、包括的な理解ができていなかった。

ダンスが十分に探られてなかったのは、自然なダンス刺激をモデル化し全脳情報処理を探るには大きく二つの課題があったからだ。第一に十分に自然・多様・多量なデータセットがなく、短く不自然な刺激を利用せざるを得なかった。加えて、最近公開された既存のデータセットにも、高次の認知的な情報は付与されていなかった。第二に、ダンスのデータから多様な情報を効率的に抽出する手段がなかった。そのためこれまでは高コストなモーションキャプチャーデータが付与された極めて少数のデータを用いるか、そもそも情報を抽出せずにダンスの動画とダンスではない人や動物の動きを視聴中に取得された指標を比較するなどの荒い比較を行っていた。

また、上記のダンスを研究する際の特有の課題と並行して、一般に自然で複雑な刺激を用いた際に、そこから抽出された多種類・高次元の特徴量と脳活動と組み合わせるための適切な解析手法も存在しなかった。そのため、多数の脳活動・特徴量が存在する場合にそれらをうまく扱うことが難しかった。

## 2. 研究の目的

研究1：本研究ではまず、ダンスに関する高次のラベルが付与された新たな大規模データセットを構築し、その上で深層学習器を用いてダンス動画からの包括的に特徴量を抽出する。最後に多数のダンス動画を視聴中のヒト fMRI 脳活動データを取得し、全脳 Encoding/Decoding モデルを構築する。これによって一連の体験（各感覚入力、感覚統合、そこから生じる認知・情動）を詳細にモデル化し、そのモデルを用いてどんな情報が脳のどの辺に表現されているかを探索する。最後に、ダンサーと未経験者の脳情報表現を比較して、熟練により変化する情報処理の中身を詳細に明らかにすることを目的とした。

研究2：自然で複雑な刺激から抽出された多種・多次元の特徴量と、それらを視聴中の fMRI 脳活動データセットを解析するための新たな枠組みを構築する。具体的には、二つ以上の異なるモダリティの情報を統合することが可能な深層生成モデルの内部表現と、脳活動データの間を繋ぐモデルを構築することによって、マルチモーダルな情報を最大限に活かした形で被験者の体験を再構成する基礎技術を構築することに取り組んだ。

## 3. 研究の方法

研究1：14名の被験者で5時間の自然なダンス動画を視聴する fMRI 実験を行った。14名の中には学業以外で10年以上のダンス経験があり現在も日常的にダンスを仕事またはスクールなどで行っているエキスパートを6名含んでおり、残りの8名は学業以外でのダンス経験がほとんどない素人だった。刺激として、30名のダンサーが60曲の音楽に合わせて異なる振り付けをした動画が収録されている大規模ダンス動画セット (AIST DanceDB) を用いた (引用)。また、各ダンス動画から感覚入力 (モーションと音)、モーションと音のクロスモーダルな情報、そして感情に関わる特徴量を抽出した。モーションは、AIST++が公開している深層学習技術を用いて自動的に抽出された情報を利用した。音は、メルスペクトラムなどの標準的な音声特徴量を用いた。クロスモーダルな情報は、音楽から自然なダンスを生成する AI Choreographer として開発された、過去のモーションと過去と将来の音を入力として将来のモーションを予測する深層生成モデルを用いた (引用)。感情は、アノテーションのための実験を実施することで34カテゴリーの感情ラベルを各ダンス動画について付与した。

研究2：米国ミネソタ大学のグループが公開している Natural Scenes Dataset (NSD) をヒト脳活動のデータとして利用した。NSDは各被験者が大量の自然画像を視聴中の脳データを公開している。また、刺激として用いられている自然画像には意味情報も付与されている。そこで、Stability AI が公開している潜在拡散モデルである Stable Diffusion にそれらの刺激情報を読み込ませ、Stable Diffusion の内部画像・意味表現と脳活動との対応をとるモデルを構築した。それによって、異なるモダリティ (意味・画像) の情報を最大限活用した脳活動デコーディングの枠組みの構築を試みた。

## 4. 研究成果

研究1：まず、感覚入力、複数の感覚入力の統合 (クロスモーダル)、そして感情の3種類の特

微量がそれぞれどの程度脳活動を説明する力があるかを確認した。その結果、どの特徴量も脳の広い部位を予測することがわかった (図 1)。これは古典的な機能分化の理解とは異なり、脳には幅広く様々な情報が分布していることが示された。

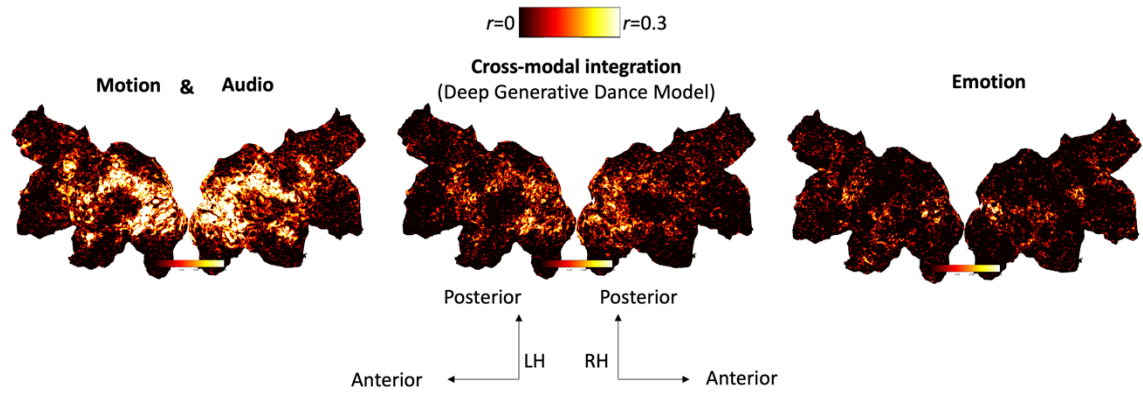


図 1

次に、本研究の特徴であるクロスモーダル特徴量が脳皮質上でどのように表現されているかを、クロスモーダル特徴量から脳活動を予測するモデルの重みに主成分分析を適用することにより可視化した。すると、脳皮質上でクロスモーダル特徴量の主成分は、異なる自然な全身協調パターンを表現していることがわかった。興味深いことに、これは過去にダンスで重要とされている基本的な全身協調パターンにかなり類似しているものであった。

更に、ダンスのモーション情報が脳皮質上でどのように処理されているかを詳細に見たところ、各関節の絶対座標 (position) は脳活動から均等に読み取れるわけではなく、下半身に比べて上半身がより高い精度で読み取れることがわかった。これは、ダンスを視聴中にヒトの脳では上半身の動きをより正確に処理していることを示唆する (図 2)。

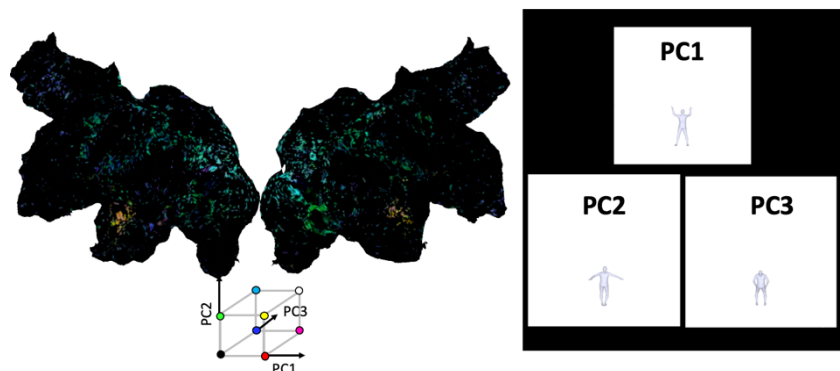


図 2

最後に、ダンスから惹起される感情の脳皮質上での処理を探った。クロスモーダル特徴量のとく同様、感情特徴量から脳活動を予測するモデルの重みに主成分分析を適用することにより可視化したところ、上位 3 つの主成分が有意に抽出された。これらの脳内での分布をみると、ダンス視聴中には全脳で異なる種類の感情表現がされていることがわかった (図 3)。

以上のモデルの情報表現を個人間で比較したところ、興味深いことに、より高次な情報処理に関わるものほどダンサーの方が強く処理していることがわかった。これは、従来の研究で調べられてきたダンサーと素人の違いの神経基盤が、高次な情報処理に関わるものであることを示唆する結果である。

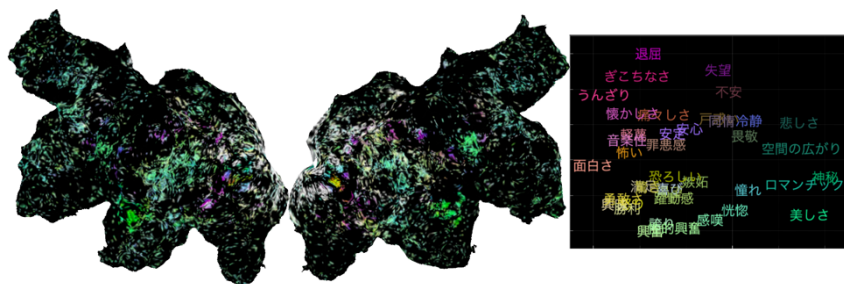


図 3

研究 2 : 図 4 に、初期視覚野の脳活動から推定される視覚情報と、高次視覚野の脳活動から推定された意味情報を、Stable Diffusion が解読可能な内部情報表現へと変換した上で、視覚再構成を試みた結果を提示する。視覚的にも意味的にも一定程度の視覚体験のデコーディングができることを示した。これによって、異なるモダリティの情報を持つ自然な刺激を脳活動から再構成する際に、マルチモーダルな深層学習器の潜在表現と紐づけることによって再構成する枠組みを提案することに成功した。

上記に加えて、認知神経科学の先行研究では主として畳み込みニューラルネットワークと再

帰的ニューラルネットワークと脳活動との対応関係が探られてきた一方で、新たに登場した(潜在)拡散モデルという生成モデルとの関係が不明であったが、潜在拡散モデルの仕組みやダイナミクスを生物学的観点から検証することにも成功した。

本研究は画像処理のトップ国際会議 CVPR2023 に採択されただけでなく (Takagi and Nishimoto, CVPR 2023) 世界中 80 以上のメディア・ニュースで報道されるなど高い注目を浴びた。

以上二つ一連の研究によって、ダンス視聴中の脳活動を解析する基盤技術開発およびダンスの脳基盤解明に成功した。

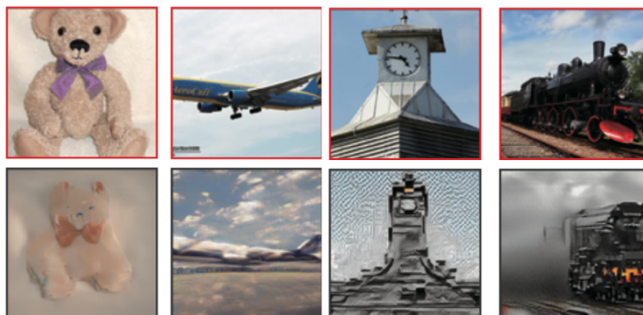


図4：提示画像（上段）と再構成画像（下段）

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 高木優、西本伸志
2. 発表標題 High-resolution image reconstruction with latent diffusion models from human brain activity
3. 学会等名 The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2023) (国際学会)
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------