

感覚知覚世界の可視化技術の研究開発 (072107013)

Development of Visual Image Reconstruction Technology

研究代表者

宮脇陽一 株式会社国際電気通信基礎技術研究所
Yoichi Miyawaki Advanced Telecommunications Research Institute International

研究期間 平成 19 年度～平成 21 年度

概要

われわれの眼から入った視覚情報は、脳において解釈され、視知覚を構成する。本研究開発課題では、この脳内での視知覚世界を、非侵襲的に計測された脳活動信号を解析することで再構成し、画像として可視化する技術の開発を行った。その結果、プログラムの学習に用いていない任意の図形パターンをヒトの脳活動から再構成できることを示すことに成功した。この成果は、脳情報の読み取り技術を飛躍的に進歩させるものであるとともに、柔軟で多自由度なブレインマシンインタフェースの技術の開発にも貢献するものである。

Abstract

Human brain interprets neural activity conveyed from the retina, and constructs visual image perception. In this research project, we developed the 'visual image reconstruction' technique, which allows us to reconstruct the visual image perception by analyzing human brain activity patterns. We demonstrated that arbitrary patterns that are not used to train the computer program can be reconstructed from the brain activity. The results advance the neural decoding technique and also contribute to development of flexible and high-dimensional brain-machine interface.

1. まえがき

われわれの眼から入った視覚情報は、脳において解釈され、視知覚を構成する。本研究では、この脳内での視知覚世界を、非侵襲的に計測された脳活動パターンを解析することで再構成し、画像として可視化する技術を開発した。

脳活動パターンを解析することによってそのヒトの知覚や運動状態を予測する研究アプローチは、脳情報のデコーディング（復号化）と呼ばれ、近年最も盛んな分野のひとつである。しかしながら、これまでのデコーディング研究のほとんどは、計測した脳活動パターンが、あらかじめ決められた有限個のカテゴリの中のどれに最も近いかという規範で分類（classification）するだけであり、予測結果は各カテゴリに割り当てられたラベルとして出力されるのみであった。このような背景のもと、本研究では、ヒトの知覚を有限個の画像に対応したカテゴリへと分類するのではなく、脳活動からヒトが見ている画像を「画像」そのものとして再構成し、可視化することを目標とした。

画像は、ピクセルの組み合わせで表現される高次元のデータである。よって、ヒトが見ている画像を「画像」そのものとして予測しようとするれば、仮に 10×10 ピクセルからなる白黒の 2 値画像であったとしても、 2^{100} 通りものバリエーションを扱わなければならないことになる。従来のデコーディング手法では、これらの全てのバリエーションに対応する脳活動を事前に実験にて取得しプログラムを学習しなければならないが、これは現実的に不可能である。高次元のデータを一括して扱うのではなく、小規模な部分データへと分解し、その組み合わせで複雑な対象を予測する新しいアプローチが必要である。

そこで本研究では、まず画像全体を複数の小領域へと分割し、その小領域ごとに予測を行い、その予測値を組み合わせることで画像全体を予測する「モジュラ・デコーディング法」を提案した。この手法に基づいて画像再構成が可能かどうかを検証し、またその性質を詳細に解析することにより、画像再構成に必要な新しいデコーディング手法の確立を目指した。

2. 研究内容及び成果

モジュラ・デコーディング法に基づく視覚像再構成は、

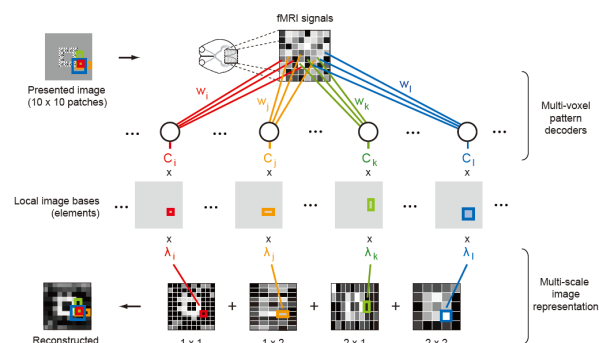


図1 fMRI 信号を用いた視覚像再構成の概念図。

画像全体が小さな画像領域の線形重ね合わせで表現可能であるとする画像再構成モデルに基づく（図1）。そして、この各画像小領域の特徴量を脳活動から並列に予測し、その予測結果を適切な重みで足し合わせることで、画像全体が再構成される。

本研究では、視覚刺激の最も基本的な特徴である輝度コントラストを特徴量とし、 10×10 パッチ領域からなる画像の輝度コントラスト空間パターンを再構成することを目指した。このパッチサイズを最小単位として、画像内の各パッチ位置で 1×1 , 1×2 , 2×1 , 2×2 のパッチ領域をカバーする複数の空間スケールを持つ画像小領域（画像基底）を定義し、その画像小領域内の平均コントラストを fMRI 信号から予測できるように機械学習プログラムをトレーニングした。各画像基底に対応する結合係数は、トレーニングデータ内での予測値の 2 乗誤差が最小になるように決定した。

プログラムの学習時には、オン・オフの 2 値コントラストを空間的にランダムに割り振ったランダムパターン刺激と、その際に fMRI 信号として計測された視覚野の脳活動パターンを用いた。プログラムの汎化能力を検証するため、数種類の幾何学画像（正方形や十字など）とアルファベット文字の画像を提示した際の fMRI 信号の計測実験も行った。

学習アルゴリズムにはスパース多項回帰アルゴリズム

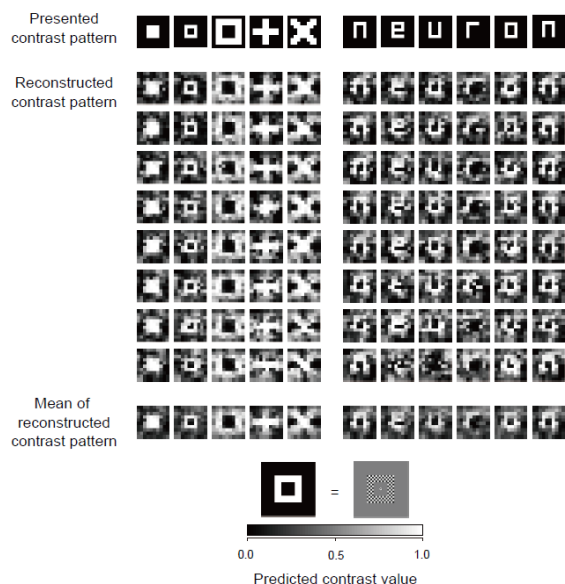


図2 視覚像再構成結果。上段は提示刺激。中段8行は、全8試行の結果を成績の高いものから順に並べたもの。下段は、それらの平均画像。アルファベットの左端と右端のデータは同じ。

(Yamashita et al, 2008)を用いた。このアルゴリズムは、入力変数の特徴次元にかかる重み係数のほとんどがゼロであるという仮定を事前分布として導入しており、その結果、予測に関連ある極少数の脳部位のみを自動的に選択しながら学習するという特徴をもつ。この特徴により、予測に関係ない無駄な特徴次元が削除されるので、オーバーフィッティングを防ぎ、汎化性能が向上すると期待される。またアルゴリズムに導入されている事前分布の仮定は、空間的に狭い位置の情報は視覚野のごく狭い脳部位で表現されているという視覚野の特徴に関する先見知識ともフィットするものである。

提案手法に基づき、ランダム画像で学習した視覚像再構成モデルを、幾何学画像とアルファベット画像を見た際の脳活動に適用して得られた再構成画像を図2に示す。これらの図形パターンは、モデルの学習には一切用いられていない新規な画像であるにも関わらず、非常に高い予測精度で図形パターンの再構成に成功しているのが見て取れる。ここでは特定の5パターンの図形と文字を用いたが、原理的には学習時に用いたランダム画像の解像度と同じ解像度の（コントラストで定義された）任意画像を再構成することが可能である。

図2の再構成画像は、同じ視覚刺激画像を12秒間見続けた際のfMRI信号を平均化したものを使って得られたものだが、この平均処理なしに、fMRI装置から送られてくる2秒ごとの単一サンプルを使っても高い精度で視覚像再構成が可能である。これらの再構成画像を連続してつなげると、被験者がfMRI装置の中でみていた画像を動画として再生できる（動画はCell Press社のホームページあるいはYouTube等で参照可能）。

本研究の画像再構成モデルでは、4種類の異なる空間スケールを持つ画像小領域を冗長に用いた多重解像度表現が大きな特徴であり、実際、この多重解像度表現によって画像再構成精度が有意に向上していることが分かった。再構成誤差を各パッチの視野位置毎に解析すると、1×1の解像度では、中心視領域で最も精度が高く、周辺視ではほぼチャンスレベルになり、逆に2×2の解像度では、中心視では誤差が大きく、周辺視においてはむしろ1×1より誤差が小さくなった。また1×2, 2×1はそれらのちょうど中間の精度を示した。ヒトの視覚

系でも中心視では受容野サイズの小さい神経細胞が多く精細な空間情報を表現するのに適しており、周辺視では受容野サイズが大きくなり精細な空間情報の表現が困難になるという知見と符合する結果である。全ての解像度を組み合わせた場合は、ちょうど単一解像度の結果の最小包絡線をたどる精度になる。これらの結果は、各視野偏位位置において最も精度の高いスケールの予測値が最適に組み合わせられ、高い再構成精度が達成されていることを示している。

3. むすび

モジュラ・デコーディング法に基づいて、ヒトが見ている画像を脳活動から高精度で再構成することに成功した。現段階ではヒトが見ているものを画像化しているだけだが、画像を頭の中で想像するだけでも実際に画像をみているときと似た活動が生じると言われているため (Kosslyn et al., 1995)、本手法を応用することで将来的には想像や夢の内容も画像化できるようになるかもしれない。

また本手法は視知覚像のみでなく、手指や四肢の運動の読み取りにも同じく応用可能である。一見複雑に見える運動であっても、少数個の運動プリミティブの組み合わせとして脳内で表現されているという可能性が指摘されている (Poggio and Bizzi, 2004)。本手法を用いて少数個の運動プリミティブ (モジュール) の状態と脳活動の関係のみを学習し、それらモジュールの組み合わせることで、任意の運動が脳活動から再構成できるかもしれない。本研究での提案手法は、我々の複雑な知覚像を解読するツールとしてのみならず、運動機能麻痺患者の柔軟なサポートに役立つブレイン・マシン・インタフェースの新たな要素技術としても、その可能性が期待される。

【誌上発表リスト】

- [1] Yusuke Fujiwara, Yoichi Miyawaki, Yukiyasu Kamitani, “Estimating image bases for visual image reconstruction from human brain activity”, *Advances in Neural Information Processing Systems* (2010 年掲載予定)
- [2] Yoichi Miyawaki, Hajime Uchida, Okito Yamashita, Masa-aki Sato, Yusuke Morito, Hiroki C Tanabe, Norihiro Sadato, Yukiyasu Kamitani, “Visual image reconstruction from human brain activity: a modular decoding approach”, *Journal of Physics: Conference Series*, Volume 197, 012021 (2009 年 9 月 15 日)
- [3] Yoichi Miyawaki, Hajime Uchida, Okito Yamashita, Masa-aki Sato, Yusuke Morito, Hiroki C Tanabe, Norihiro Sadato, Yukiyasu Kamitani, “Visual image reconstruction from human brain activity using a combination of multi-scale local image decoders”, *Neuron*, Volume 60, Issue 5 pp915-929 (2008 年 12 月 11 日)

【受賞リスト】

- [1] 宮脇陽一、2009 年度日本神経回路学会論文賞、“Visual image reconstruction from human brain activity using a combination of multiscale local image decoders”, 2009 年 9 月 25 日

【報道発表リスト】

- [1] “脳から知覚映像を読み出すーヒトの脳活動パターンから見ている画像の再構成に成功ーニューロン誌 12 月 11 日号掲載予定 (表紙デザインに採用)ー”, 2008 年 12 月 9 日

【本研究開発課題を掲載したホームページ】

- [1] http://www.atr.jp/html/topics/press_081211_j.html
- [2] http://www.cns.atr.jp/dni/?page_id=109