1版

様　式　C－19、F－19－1、Z－19　(共通)

科学研究費助成事業　　研究成果報告書

科研費 KAKENHI

令和　４ 年　 ６ 月 ２６ 日現在

機関番号：６３９０５

研究種目：若手研究

研究期間：2019～2021

課題番号：１９Ｋ２０３９０

研究課題名（和文）Study on the hierarchy of cross-modal processing using functional MRI and Deep Learning

研究課題名（英文）Study on the hierarchy of cross-modal processing using functional MRI and Deep Learning

研究代表者

Ｐｈａｍ　Ｑｕａｎｇ・Ｔｒｕｎｇ（PHAM, QUANG TRUNG)

生理学研究所・システム脳科学研究領域・特任研究員

研究者番号：６０８３７７２２

交付決定額（研究期間全体）：（直接経費）　　3,200,000円

研究成果の概要（和文）：本研究では私は触覚で収集された情報の断片は視覚処理経路に活かして、物体認識を行ったと言う仮説を検討した。頭頂間溝(IPS)または外側後頭皮質(LOC)で情報が集まり、視覚腹側経路に伝達することである。20人の被験者が視覚・触覚実験に参加したもらった。3x5ドット数字のような刺激を受け、ボタンクリックで答えてもらった。同時に、MRIスキャナーを使って脳活動画像を撮れた。被験者の正答率はチャンスレベルより高ったと分かった。MRI情報解析によりIPSおよびLOCの活動が確認でき、さらに他のワーキングメモリに関する領域の活動も観察できた。IPC及びLOCにおける情報表現も似ていることも観察できた。

研究成果の学術的意義や社会的意義
The results of this study extend our current understanding of the cross-modal processing of haptic and visual information in the brain. In the future, I expected these results to become hint for building a more efficient robotics system that mimics human brain.

研究成果の概要（英文）：In this study, I proposed a hypothesis where the fragments of haptic information are integrated into visual imagery at the intraparietal sulcus (IPS) and lateral occipital cortex (LOC) before passing to the visual ventral pathway for object recognition. For examining that hypothesis, I designed an experiment where the participants try to recognize the 3x5 dot-digits (from 0 to 9) from visual and haptic stimuli. During the experiment, I captured their brain signals using an MRI scanner. Behavioral analysis revealed that most participants can recognize the dot-digit above the chance level. Their visual performance is better than haptic performance. The fMRI analysis confirmed the involvements of IPS and LOC, and the dorsolateral prefrontal cortex (dlPFC) which is related to working-memory. I performed a representational analysis and found similar representations for both haptic and visual information at the IPS and LOC.

研究分野：computational neuroscience

キーワード：cognitive science　somatosensory　haptics　perception

１．研究開始当初の背景

A conventional model (1) describes the path of the shape computations of haptic perception as follows. The haptic information is perceived by the mechanoreceptors, then arrives at thalamus, then travels to various subareas of the primary somatosensory cortex (S1), such as Area 3b, Area 1, and Area 2. In these areas, the haptic information is broken into low-level features such as touch, indentation, velocity, texture, et cetera.

Previous studies have shown that the intraparietal sulcus (IPS) and the lateral occipital complex (LOC) are essential for haptic object recognition (2, 3). The LOC is also well-known to be engaged in visual object recognition (4, 5) and preserving the high-level shape information. It is unknown whether this representation at the LOC could be shared between visual and haptic object recognition. Even though detailed evidence has not been shown, the model proposed by Lacey and Sathian, and its related works (6, 7) suggest that the IPS may manipulate the information in working memory to mediate the visual imagery for nonvisual object recognition. How IPS processes such visual imagery from the haptic object information remains unsolved.

２．研究の目的

Different from the eyes in vision, the mechanoreceptors in haptic perception can only receive the shape of an object as a sequence of haptic fragments. Thus, it is reasonable to think that the IPS may play the role of an integrator that uses such haptic fragments to construct visual imagery.

The objective of this research is to define the hierarchical processing model that may be employed by the brain for haptic object recognition. We proposed a hypothesis of how the computation occurs hierarchically from sensory areas to IPS, then to the LOC (Fig. 1). We also address the question of whether the representation of shape obtained from haptic stimuli is identical to those obtained from visual stimuli.

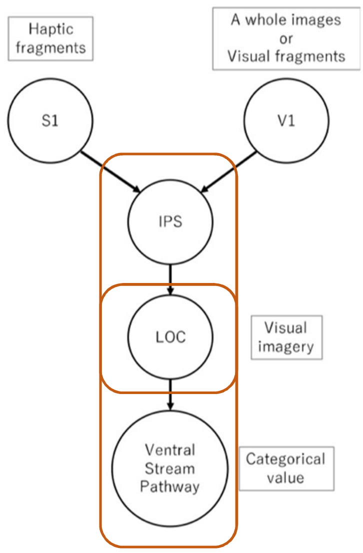

Figure 1 Hypothesis of a shared pathway for haptic and vision object recognition. The red boxes indicate the known pathways.

３．研究の方法

We recruited 20 participants for an fMRI experiment, which consists of two tasks, haptic and vision. The participants were asked to haptically recognize the 3 x 5 dot digits in the haptic task. A haptic display was developed in-house to sequentially push each row of the selected dot-digits (from top to bottom) to the participant's right palm. The dot digits were from 0 to 9 and randomly drawn for each session. In the vision task, the participants were asked to observe the similar dot digits shown on a monitor. Each stimulus last for 5 seconds, then the participants have 3 seconds to provide their answers.

During both tasks, we capture the BOLD response of participants using 3.0-Tesla Siemens Verio B MRI scanner (with a 32-element phased-array head coil) located at the National Institute for Physiological Sciences, Japan. T2*-weighted gradient echo-planar imaging (EPI) was used to obtain the function images, with the following parameters: TR=750ms, TE=31ms, flip angle=55 degrees, FOV =192 x 192 mm^2, matrix size of 98 x 98, 72 slices with isotropic voxel resolution=2 x 2 mm^2, and multiband EPI with a factor of 8. T1-weighted anatomical images were collected in each subject using the following parameters: TR=1800ms, TE=1.98ms, flip angle of 9 degrees, FOV=256 x 256 mm^2, and voxel resolution=1.0 x 1.0 x 1.0 mm^3.

All MRI data were preprocessed using an HCP Protocol as described in Glasser et al. The raw data was processed via a preprocessing pipeline including three structural pipelines (PreFreeSurfer, FreeSurfer, and Post-FreeSurfer), and two functional pipelines (fMRI Volume and fMRI surface). The FreeSurface pipeline is based on FreeSurface version 5.2. We performed the standard 1^{st} level analysis and 2^{nd} level analysis using MATLAB 9.3 and Statistic Parametric Mapping (SPM12) software.

The behavioral analysis was conducted with heavy use of Python 3.9 and MATLAB 9.3. For MRI analysis, we were thinking of using a neural network to depict our computational model. However, there were a few drawbacks, such as the lack of large data, lack of time (the COVID-19 pandemic outbreaks happened), and the complexity of the neural network. Therefore, we used a Generalized Linear Model (GLM) with specific regressors to construct the encoding model for fMRI signals. The regressors were designed to mimic the activity of working-memory related areas. Then we also performed the representational similarity analysis (RSA) to investigate the similarity between the representation of shape information during the haptic task and vision task at two regions of interest (ROI), i.e., IPS and LOC.

４．研究成果

First, most participants perform with accuracy higher than the change level. The performance in the vision task is significantly higher than those in the haptic task (paired t-test; t(14)=11.19; p<0.0001). This is an expected result because the acuity of vision is better than haptic. Among 10 dot-digits, 6 digits are easily recognizable (precision>0.5). The confusion matrices revealed that the 4 and 9 digits are highly mislabeled the most in both tasks (Fig.2). In collaboration with other colleges, we have developed a deep neural network (DNN) for providing a counterfactual explanation to a misclassification (8). Such DNN would be useful for investigating the mislabeled issue in future work.

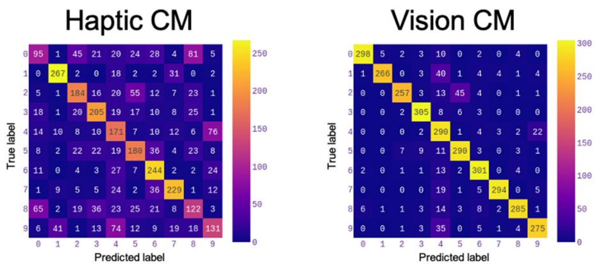

Figure 2 Confusion matrix (CM) of participant's haptic and vision task performance

The group analysis of MRI data reveals the activation of supplementary motor area (SMA) and primary motor area (M1) when the participant answered. This validates our processing method. The aIPS, vIPS, and LOC were confirmed to be involved in both haptic recognition and vision recognition (Fig.3). The dorsolateral prefrontal cortex (dlPFC) which is related to working memory, also showed high activity (one-sample t-test; t_vision(14)=4.19; t_haptic(14)=5.58). This result suggests that the top-down signal from working-memory-related high-level areas may be required for constructing the visual imagery from sensory fragments.

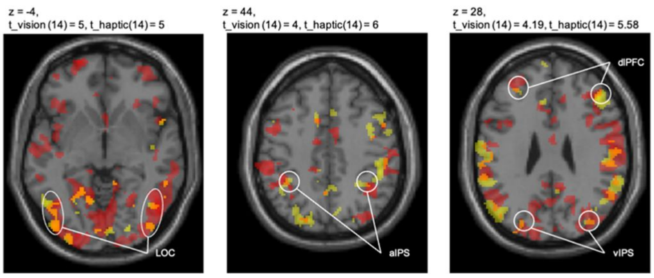

Figure 3 Slices of fMRI analysis

We performed RSA on the digits with a high correct answer rate (Fig. 4). The diagonal of the Representational Similarity Matrix (RSM) between LOC voxels found in the haptic task and those found in the vision task show a modest similarity of identical digits across both tasks. Interestingly, the RSA also revealed the similarity between visual and haptic representations at the IPS. The similarity at IPS seems to be higher than that at LOC. Taken together, it is reasonable to think that the representations at LOC may be influenced by other high-level regions (such as dlPFC). I have investigated the effects of bottom-up and top-down modulation on LOC regarding visual processing and visual imagery (9). For that purpose, I have also developed a region-to-region decoding technique capable of predicting the neural activity at one region based on the neural activity of another. Using the deep learning model and a novel region-to-region decoding technique, I found that the top-down modulation was crucial for visual feature representations during the imagery task (Fig. 5).

At the current state, our results partly confirm the proposed hypothesis, suggesting a model which consists of low-level areas (IPS, LOC), and a high-level area (dlPFC) as a shared pathway for both haptic and visual object recognition. The comparison of information representation at LOC and IPS open a novel understanding of how multi-sensory area process information from distinct modalities. We observed a modest similarity between them. This may be because the number of participants is small (the outbreak of COVID-19 made it hard for recruiting). We expect that future work with a larger dataset may clarify this observation. A neural network that mimics the behavioral response of participants would help clarify the hierarchy of the shared pathway. Furthermore, decoding the visual imagery (10) from the IPS and LOC will be an interesting approach in future work. This would help us investigate how exactly the representations of haptic object recognition and visual object recognition are similar and different inside the brain.

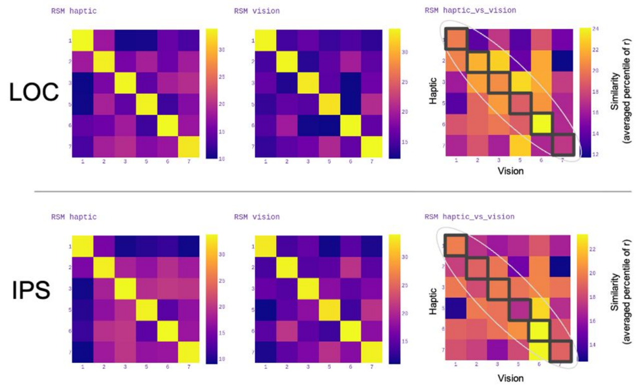

Figure 4 RSA of haptic and visual information at LOC and IPS

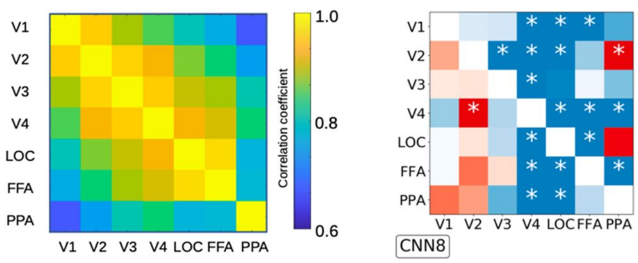

Figure 5 Region-to-Region decoding technique (left) and an example of distillation of high-level visual feature (right)

References

1. A. Bodegard et al., Hierarchical Processing of Tactile Shape in the Human Brain, Neuron, Vol. 31, 2001, pp. 317–328.
2. S. Lacey and K. Sathian, Visuo-haptic Multisensory Object Recognition, Categorization, and Representation, Front. Psychol., Vol. 5,2014.
3. K. Sathian, Analysis of Haptic Information in the Cerebral Cortex, J. Neurophysiol., Vol. 116, 2016, pp. 1795–1806.
4. R. Malach et al., Object-related Activity Revealed by Functional Magnetic Resonance Imaging in Human Occipital Cortex, PNAS, Vol. 92, 1995, pp. 8135–8139.
5. Z. Kourtzi and N. Kanwisher, Representation of Perceived Object Shape by the Human Lateral Occipital Cortex, Science, Vol. 293, 2001, pp.1506–1509.
6. G. Deshpande et al., Object Familiarity Modulates Effective Connectivity during Haptic Shape Perception, Neuroimage, Vol. 49, 2010, pp.1991–2000.
7. N. Dijkstra et al., Shared Neural Mechanisms of Visual Perception and Imagery, Trends Cogn. Sci., Vol. 23, Issue. 5, 2019, pp. 423–434.
8. T. Matsui et al., Counterfactual Explanation of Brain Activity Classifiers Using Image-To-Image Transfer by Generative Adversarial Network, Front. Neuroinform., Vol. 15, 2022, pp. 802938.
9. T. Q. Pham et al., Distillation of Regional Activity Reveals Hidden Content of Neural Information in Visual Processing, Front. Hum. Neurosci., Vol. 15, 2021, pp. 777464.

10. T. Horikawa et al., Neural Decoding of Visual Imagery During Sleep, Science, Vol. 340, Issue. 6132, 2013, pp. 639-642.

5．主な発表論文等

〔雑誌論文〕　計3件（うち査読付論文　3件／うち国際共著　3件／うちオープンアクセス　3件）

1．著者名 Pham Trung Quang、Nishiyama Shota、Sadato Norihiro、Chikazoe Junichi	4．巻 15
2．論文標題 Distillation of Regional Activity Reveals Hidden Content of Neural Information in Visual Processing	5．発行年 2021年
3．雑誌名 Frontiers in Human Neuroscience	6．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.3389/fnhum.2021.777464	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

1．著者名 Matsui Teppei、Taki Masato、Pham Trung Quang、Chikazoe Junichi、Jimura Koji	4．巻 15
2．論文標題 Counterfactual Explanation of Brain Activity Classifiers Using Image-To-Image Transfer by Generative Adversarial Network	5．発行年 2022年
3．雑誌名 Frontiers in Neuroinformatics	6．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.3389/fninf.2021.802938	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕　計2件（うち招待講演　0件／うち国際学会　2件）

1．発表者名 Pham Quang Trung
2．発表標題 Shared hierarchical cross-modal processing between visual and haptic recognition: an fMRI study
3．学会等名 NEURO2022（国際学会）
4．発表年 2022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