

令和 6 年 5 月 23 日現在

機関番号：57101

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K11878

研究課題名（和文）l1ノルム損失における畳み込み型スパース表現と分散圧縮符号化・深層学習への展開

研究課題名（英文）Convolutional sparse representation of l1 norm error criterion and its development for distributed video coding and deep learning

研究代表者

黒木 祥光（Kuroki, Yoshimitsu）

久留米工業高等専門学校・制御情報工学科・教授

研究者番号：60290847

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）：近年のAIの発展は、畳み込みニューラルネットワークが画像分類のコンテストで既存の手法を凌駕したことがきっかけである。本研究では画像を畳み込みフィルタとその係数の和で近似する畳み込みスパース表現に着目した。ここで、スパースとはフィルタ係数ができるだけ疎、つまり多くのゼロを含むことを意味し、近似精度が同じであればスパース性が高いほど特徴をよく表現していると考えられる。本研究では分散圧縮符号化と畳み込みニューラルネットワークへの応用を検討し、既存手法を上回る成果を得た。

研究成果の学術的意義や社会的意義

本研究では畳み込みスパース表現のスパース性と近似精度の双方を向上させる方法として近似精度をL1ノルムと称する誤差ベクトルの絶対値で評価する手法を提案した。また、計算負荷を低減し、大規模データで適用可能なコンセンサス方式を導出した。これらの成果は分散圧縮符号化および小規模な畳み込みニューラルネットワークにおける精度向上へとつながり、国際会議ICIBMSにおけるStudent Best Paper Awardとして評価された。

研究成果の概要（英文）：The recent advancements in AI were triggered by convolutional neural networks surpassing existing methods in the image classification contest. This study focuses on convolutional sparse representations, which approximate images using a sum of convolutional filters and their coefficients. Here, "sparse" means that the number of filter coefficients contain many zeros as possible. If the approximation accuracy is the same, a higher sparsity is considered to better representation of image features. In this study, we examined the application of distributed compression coding and convolutional neural networks and have achieved results superior to conventional methods.

研究分野：知覚情報処理

キーワード：畳み込み型スパース表現 分散圧縮符号化 深層学習 凸最適化

1. 研究開始当初の背景

与えられた画像信号 s (一般に数ベクトルで表現されるため、以下、 $s \in \mathbb{R}^N$ とする) をできるだけ少数のベクトル d_k ($k = 1, 2, \dots, K$) $\in \mathbb{R}^N$ の線形結合で表現する方法は、スパース表現と呼ばれ、圧縮センシングやパターン認識など、多くの分野で研究されている。結合係数ベクトルを $x \in \mathbb{R}^K$ とすると、この線形結合による近似は $s \approx \sum_k d_k x_k = Dx$ ($D = (d_1 d_2 \dots d_K) \in \mathbb{R}^{N \times K}$) で表現される。ここで、 D は辞書、 d_k はアトムとしばしば呼ばれる。信号のスパース性は x の非ゼロ係数の個数、つまり、 L_0 ノルムで評価されるが、 L_0 ノルムを L_1 ノルムのような凸関数 $R(x)$ に緩和することにより、スパース係数の導出は次の凸最適化問題で表現される。

$$\arg \min_x \|Dx - s\|_2^2 + \lambda R(x)$$

この式における $\lambda \in \mathbb{R}$ は第 1 項の損失項と第 2 項の正則化項のバランスをとるパラメータである。この種の問題は LASSO (Least Absolute Shrinkage and Selection Operator) や ADMM (Alternating Direction Method of Multipliers) といった、凸最適化問題のソルバーで解くことができる。しかし、以下の 2 点について疑問が残る。

(1) 画像信号 s は画像そのものではなく、重複しないように分割された矩形ブロックが用いられる。各ブロックは一般的な符号化・復号化処理と同様に独立に処理されるが、位置ずれに対して頑健ではないため、符号化における動き補償やオプティカルフロー、深層学習におけるプーリングなどで補う必要がある。位置ずれに対して高い頑健性を有する代替法はないのであろうか。詳細は後述するが、この点を解決できれば、ノンキーフレームにサイド情報を必要としない、新たな分散圧縮符号化の枠組みや、プーリングを必要とせず、位置ずれに対して頑健な深層学習を実現できる。

(2) 辞書 D として、離散コサイン基底など既知のものを用いるだけでなく、複数の画像信号 s_j ($j = 1, 2, \dots, J$) から設計することもできる。信号 s_j は加法性の雑音を含む場合があり、それは正規分布に従うとは限らず、例えば、スパイク雑音も存在する。また、 s_j に外れ値が含まれる可能性もある。損失項の評価は L_2 ノルムが最善なのであろうか。

2. 研究の目的

本研究の目的は上に述べた 2 つの疑問を「 L_1 ノルム損失項を用いた畳み込み型スパース表現」で解決することにある。

まず、(1) は畳み込みスパース表現にて解決する。畳み込みスパース表現では、画像 s_j を画像サイズより小さな畳み込み辞書フィルタ $d_k \in \mathbb{R}^B$ ($B < N$) とその畳み込み係数 $x_{j,k} \in \mathbb{R}^N$ を用いて次式で近似する。ここで、辞書フィルタ d_k は全画像に対して共通であるため添字 j を含まない。

$$s_j = d_1 * x_{j,1} + d_2 * x_{j,2} + \dots + d_K * x_{j,K} = \sum_{k=1}^K d_k * x_{j,k}$$

画像が J 枚与えられた時に辞書フィルタおよびフィルタ係数の最適化は次式で定式化される。

$$\arg \min_{d_k, x_{j,k}} \sum_{j=1}^J \left\| \sum_{k=1}^K d_k * x_{j,k} - s_j \right\|_2^2 + \lambda \sum_{j=1}^J \sum_{k=1}^K R(x_{j,k})$$

解法の詳細は割愛するが、ここで求めた辞書フィルタおよびフィルタ係数を用いて分散圧縮符号化と畳み込みニューラルネットワーク (Convolutional Neural Networks: CNNs) に適用する。画像の位置ずれは畳み込み係数 $x_{j,k} \in \mathbb{R}^N$ のずれで表現され、辞書フィルタ d_k は変更されないことに注意されたい。その性質を活用して深層学習の畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) と分散圧縮符号化等に適用する。

次に (2) については L_1 ノルムを用いた損失項の導入によって実現したい。確率変数に対する L_2 ノルム最小化による代表値が期待値であるのに対し、 L_1 ノルムのそれは中央値で与えられるように、 L_1 ノルムの最小化は外れ値に対して頑健性が高い。この性質を損失項に適用する。特に正則化項 $R(x)$ を L_1 ノルムとした場合、損失項と正則化項が共に凸関数ではあるが不連続な L_1 ノルムで評価されるため、最適解を求めることは困難である。その解法を導きたい。

3. 研究の方法

(1 - 1) 畳み込みスパース表現における辞書フィルタの CNN への適用

例えば CNN の初期値に畳み込みスパース表現の辞書フィルタを用いる等の応用により、CNN の精度向上を試みる。

(1 - 2) 畳み込みスパース表現における辞書フィルタを用いた分散圧縮符号化

従来の分散圧縮符号化は画像を矩形ブロックに分割して個別に処理を行っていたが、畳み込みスパース表現の枠組みを活用し、より高い圧縮率で高精度な再現画像の獲得を目指す。

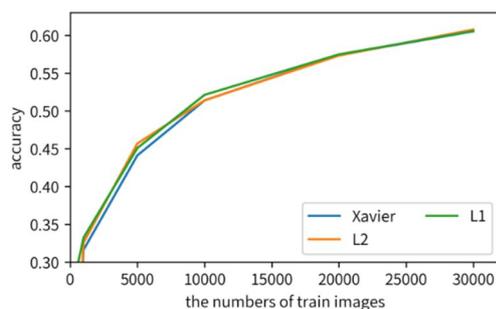
(2) L_1 ノルムを損失項とした畳み込みスパース表現の精度向上

損失項を L_1 ノルムで評価することにより、外れ値等の混入に対して頑健な辞書フィルタを求める。ここで得られた成果は上記(1-1)にも適用する。

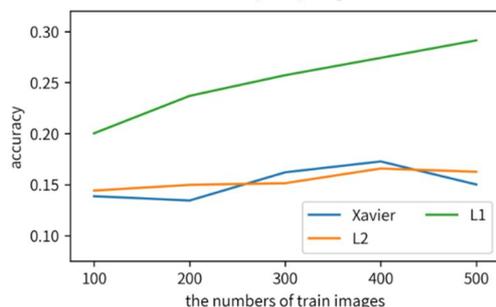
4. 研究成果

(1-1) 畳み込みスパース表現における辞書フィルタのCNNへの適用[1]-[4]

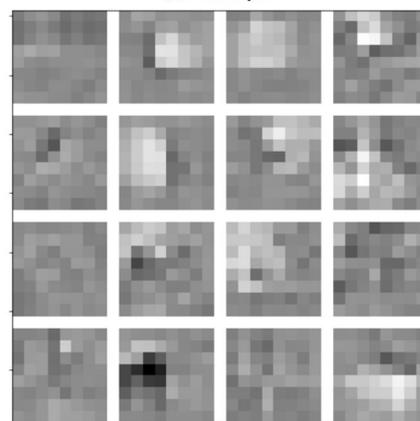
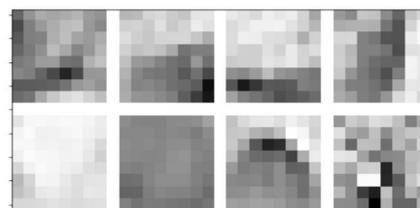
紙幅の都合上、引用文献の[1]のみについて示す。畳み込みスパース表現における辞書フィルタをCNNの初期パラメータに適用し、画像データセットCIFAR-10の分類実験を行った。辞書フィルタの設計には損失項として L_1 ノルムを用いた場合、 L_2 ノルムを用いた場合の双方について比較を行った。実験の結果を図1に示す。 L_1 ノルムを損失項として用いた提案法は L_2 ノルム損失項の場合ならびに既存手法であるXavier法と比較して、少ないトレーニング画像数で最も高い分類精度を示した。また、図2は L_1 ノルムを損失項としたときの辞書フィルタであり、画像の特徴を抽出していることが分かる。



(a) Numbers of training images up to 30,000



(b) Small numbers of training images



(b) Second layer

図1：認識率の比較

図2：生成されたフィルタ

(1-2) 畳み込みスパース表現における辞書フィルタを用いた分散圧縮符号化[5]

分散圧縮符号化では動画画像をキーフレームとノンキーフレームの2種類に分割する。キーフレームは他のフレームを参照せず、静止画像として取り扱い、ノンキーフレームより先に符号化・復号化処理を行う。ノンキーフレームは既に復号化済みのキーフレームを参照することができる。本研究では、多数のキーフレームを用いて作成した辞書フィルタを用い、ノンキーフレームを分散圧縮符号化する手法を検討した。従来のブロック単位の分散圧縮符号化と異なり、ブロックノイズ生じないことに加え、畳み込みスパース表現により画像の局所的な特徴を表す辞書と画像全体にわたるフィルタ係数を用いることで復号精度を向上させることができた。図3に

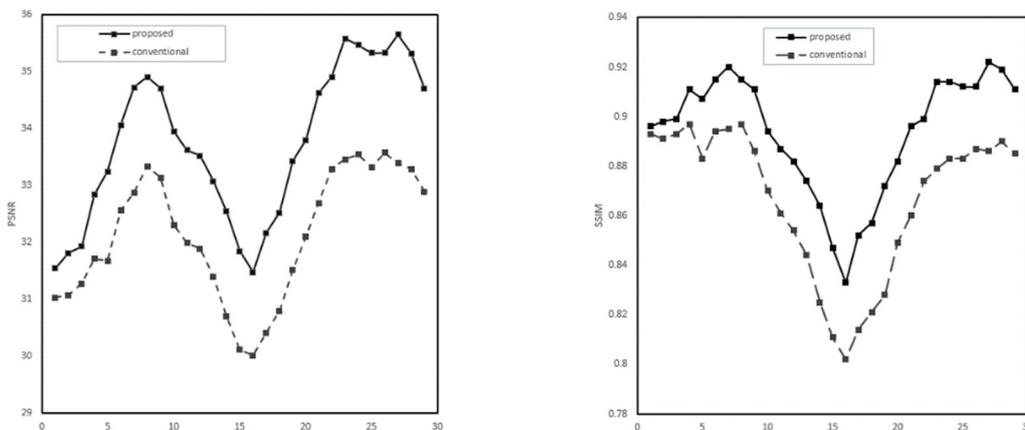


図3：フレーム番号182から120における再生画像の比較（左図はPSNR，右図はSSIM）

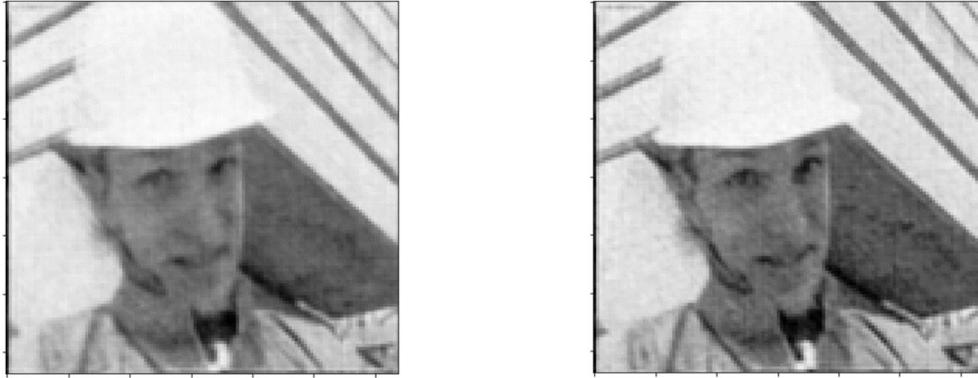


図 4：再生画像（左図は従来法，右図は提案法）

従来方法との再生画像品質の比較を示す．画像の品質は PSNR と SSIM を用いた．また，図 4 に再生画像を示す．

(2) L_1 ノルムを損失項とした畳み込みスパース表現の精度向上[6][7]

紙幅の都合上引用文献[7]のみについて示す．畳み込みスパース表現で辞書フィルタを求めるには多くのメモリを要し，計算負荷が大きい．それを解決する方法として 1 枚の画像事に辞書フィルタを求め，全画像にわたって統括するコンセンサスフレームワークが提案されている．本研究では， L_2 ノルムで定式されていた損失項を L_1 ノルムに変更し，外れ値に対して頑健なフィルタ設計を行った．辞書学習にコンセンサスフレームワークを使用しない場合，学習画像の数はメモリ容量により最大で約 100 枚程度に制限されるが，提案方法により 1,000 枚，10,000 枚の学習画像を用いた辞書を作成することができた．また， L_1 損失項の採用により，同等の非ゼロ係数で再現画像の品質を大幅な改善を達成できた．PSNR と SSIM の比較を図 5 に，実際に再現した画像を図 6 に示す．

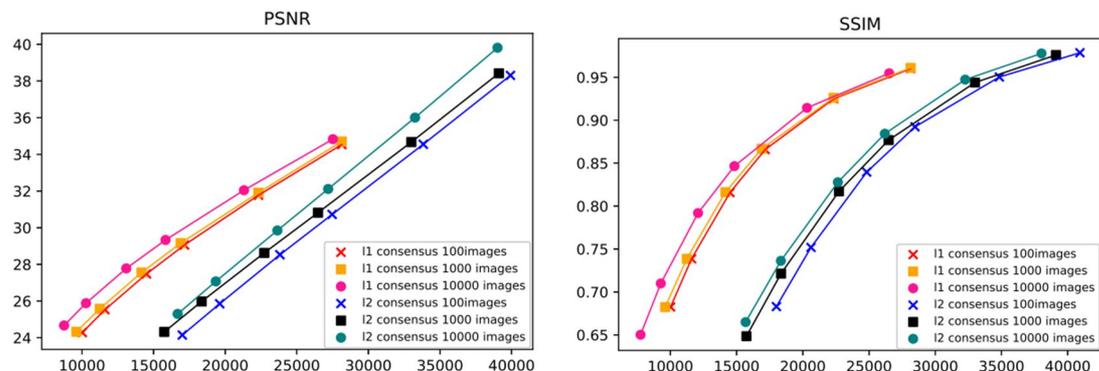


図 5：各種手法における非ゼロ係数の個数と PSNR および SSIM の比較



	L2 Consensus	L1 Consensus	original
l0 norm	17118	17088	
PSNR	24.22 dB	29.019 dB	
SSIM	0.6861	0.8647	

図 6：再生画像の比較

< 引用文献 >

[1] S. Yoda, Y. Tsukiashi, and Y. Kuroki, "Initial parameters of CNNs generated by Convolutional Sparse Representation with L1 error term", 2022 7th International Conference on Intelligent Informatics and Biomedical Science (ICIIBMS), Nara, Japan, 2022, pp. 380-381, doi: 10.1109/ICIIBMS55689.2022.9971702.

- [2] T. Yoshida, Y. Kobayashi and Y. Kuroki, "Image classification using convolutional sparse representation of l_1 error term," 2021 6th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Oita, Japan, 2021, pp. 248-249, doi: 10.1109/ICIIBMS52876.2021.9651628.
- [3] Y. Higuchi, T. Hirakawa and Y. Kuroki, "An image classification using convolutional sparse representation and cone-restricted subspace method," 2021 6th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Oita, Japan, 2021, pp. 247-248, doi: 10.1109/ICIIBMS52876.2021.9651582.
- [4] K. Kitajima and Y. Kuroki, "Image classification with multi-scale convolutional sparse representation," 2020 IEEE 7th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Kuala Lumpur, Malaysia, 2020, pp. 1-5, doi: 10.1109/ICETAS51660.2020.9484200.
- [5] Y. Higuchi and Y. Kuroki, "Distributed Compressed Video Sensing based on Convolutional Sparse Coding using a large number of keyframes," 16th International collaboration Symposium on Information, Production and Systems (ISIPS 2021).
- [6] S. Yoda, H. Kawazoe and Y. Kuroki, "Convolutional Dictionary Learning with Huber Error and l_1 Regularization Terms," 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), Hualien City, Taiwan, 2021, pp. 1-2, doi: 10.1109/ISPACS51563.2021.9651025.
- [7] M. Takanashi and Y. Kuroki, "A Consensus Framework for Convolutional Dictionary Learning based on L_1 Norm Error," 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Tokyo, Japan, 2021, pp. 1400-1404.

5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 7件/うち国際共著 0件/うちオープンアクセス 7件）

1. 著者名 Satoshi Yoda, Yuuto Tsukiashi, and Yoshimitsu Kuroki	4. 巻
2. 論文標題 Initial parameters of CNNs generated by Convolutional Sparse Representation with L1 error term	5. 発行年 2022年
3. 雑誌名 2022 7th International Conference on Intelligent Informatics and Biomedical Science (ICIIBMS)	6. 最初と最後の頁 380-381
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICIIBMS55689.2022.9971702	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 -

1. 著者名 Yoshida Takahiro, Kobayashi Yusaku, Kuroki Yoshimitsu	4. 巻 6
2. 論文標題 Image classification using convolutional sparse representation of l1 error term	5. 発行年 2021年
3. 雑誌名 6th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)	6. 最初と最後の頁 248-249
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICIIBMS52876.2021.9651628	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 -

1. 著者名 Higuchi Yosuke, Hirakawa Tomoya, Kuroki Yoshimitsu	4. 巻 6
2. 論文標題 An image classification using convolutional sparse representation and cone-restricted subspace method	5. 発行年 2021年
3. 雑誌名 6th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)	6. 最初と最後の頁 247-248
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICIIBMS52876.2021.9651582	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 -

1. 著者名 Yoda Satoshi, Kawazoe Hironori, Kuroki Yoshimitsu	4. 巻 -
2. 論文標題 Convolutional Dictionary Learning with Huber Error and l_1 Regularization Terms	5. 発行年 2021年
3. 雑誌名 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)	6. 最初と最後の頁 1-2
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ISPACS51563.2021.9651025	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 -

1. 著者名 M. Takanashi and Y. Kuroki	4. 巻 -
2. 論文標題 A Consensus Framework for Convolutional Dictionary Learning based on L1 Norm Error	5. 発行年 2021年
3. 雑誌名 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	6. 最初と最後の頁 1400-1404
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 -

1. 著者名 Kitajima Kazuki, Kuroki Yoshimitsu	4. 巻 -
2. 論文標題 Image classification with multi-scale convolutional sparse representation	5. 発行年 2020年
3. 雑誌名 2020 IEEE 7th International Conference on Engineering Technologies and Applied Sciences (ICETAS)	6. 最初と最後の頁 1-5
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICETAS51660.2020.9484200	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 -

1. 著者名 Takanashi Mizuki, Kuroki Yoshimitsu	4. 巻 -
2. 論文標題 Image Classification Using l1-fidelity Multi-layer Convolutional Sparse Representation	5. 発行年 2020年
3. 雑誌名 2020 IEEE 7th International Conference on Engineering Technologies and Applied Sciences (ICETAS)	6. 最初と最後の頁 1-5
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICETAS51660.2020.9484323	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 -

〔学会発表〕 計4件 (うち招待講演 0件 / うち国際学会 4件)

1. 発表者名 Yosuke Higuchi and Yoshimitsu Kuroki
2. 発表標題 Distributed Compressed Video Sensing based on Convolutional Sparse Coding using a large number of keyframes
3. 学会等名 16h International collaboration Symposium on Information, Production and Systems (ISIPS 2021) (国際学会)
4. 発表年 2022年

1. 発表者名 Takuro Eguchi, Hayata Morisaki, and Yoshimitsu Kuroki
2. 発表標題 Distributed Compressed Video Sensing based on Convolutional Sparse Coding using Fourier Measurement Matrix and L1 Fidelity Term
3. 学会等名 16h International collaboration Symposium on Information, Production and Systems (ISIPS 2021) (国際学会)
4. 発表年 2022年

1. 発表者名 Satoshi Yoda, Akito Narahara, and Yoshimitsu Kuroki
2. 発表標題 On Initial CNN parameters using convolutional sparse filters
3. 学会等名 15h International collaboration Symposium on Information, Production and Systems (ISIPS 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Kaede Kumamoto and Yoshimitsu Kuroki
2. 発表標題 Multi-layer Feature Extraction with Convolutional Dictionary Learning based on L1 Norm Error with Smoothed L0 Norm Regularization and Non-negative Coefficients
3. 学会等名 14h International collaboration Symposium on Information, Production and Systems (ISIPS 2020) (国際学会)
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

Research map 黒木祥光
<https://researchmap.jp/read0047048>

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
---------	---------