

科学研究費助成事業 研究成果報告書

令和 2 年 6 月 22 日現在

機関番号：27101

研究種目：若手研究(B)

研究期間：2017～2019

課題番号：17K14683

研究課題名(和文) 指向性テンソル辞書に基づく高解像度映像スパース表現手法の確立

研究課題名(英文) A study on sparse representation of high-resolution video based on directional tensor dictionary

研究代表者

京地 清介 (Kyochi, Seisuke)

北九州市立大学・国際環境工学部・准教授

研究者番号：70634616

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：本研究では指向性テンソル辞書に基づく高解像度映像スパース表現手法を確立した。大規模テンソルデータである高精細・マルチスペクトル・多視点映像を(信号圧縮・復元に向けて)スパースに表現するには、大規模なテンソル辞書が必要となる。しかしテンソル辞書学習はその多次元性により、非常に計算負荷が高いため、テンソル辞書の各次元が大規模になると学習は困難になる。本研究では、濃淡画像の各画素の方向相関(縦横斜めの方向への高い相関)が映像でも成り立つことに着目し、まず指向性(特定方向への分布形状)要素を有した非学習型大規模テンソル辞書を設計し、高解像度映像のスパース表現を行うアルゴリズムを開発した。

研究成果の学術的意義や社会的意義

高解像度映像スパース表現のための大規模テンソル辞書を非学習型辞書の多層線形結合によって生成するコンセプトは非常にユニークであり、またその効果を確認することができた。また本研究で開発したエピグラフ変形による各層のパラメータ学習手法は、近年活発に研究が進められている深層ニューラルネットワーク(DNN)のパラメータ学習に応用できると考えられる。DNNの性質はブラックボックスになっている部分が多く、これまで経験的な試行錯誤によって開発が進められてきたが、エピグラフ変形を架け橋として、理論が整備されている凸最適化学の知見を活用しながら、更に高度なDNNを構築できると考えられる。

研究成果の概要(英文)：In this study, a high-resolution video sparse representation method based on a directional tensor dictionary is established. In order to represent large scale tensor data such as high-definition, multi-spectral, multi-viewpoint video sparsely (for signal compression and restoration), a large tensor dictionary is required. However, due to the multidimensional nature of tensor, learning tensor dictionaries becomes difficult when the number of tensor dimensions becomes large. Since directional correlations between pixels of gray scale images (high correlations in the vertical, horizontal and diagonal directions) are also possible in video, we designed a fixed large-scale tensor dictionary with directional elements and developed an algorithm for sparse representation of high-resolution video.

研究分野：信号処理

キーワード：テンソル辞書 スパース表現 高解像度映像 信号圧縮 信号復元 凸最適化

様式 C-19、F-19-1、Z-19 (共通)

1. 研究開始当初の背景

高精細映像 (4K / 8K), 高フレームレート映像, 多視点映像, マルチ・ハイパースペクトル映像, 医療用画像映像 (CT・MRI) 等, 時空間・スペクトル・視点高解像度映像の普及が進んでいる。その結果, 情報量の爆発的増加が問題となっており, 映像圧縮の抜本的な高性能化が求められている。また, センサの性能不足等で高解像度映像全体を完全に取得できない状況も多く, 高精度な信号復元 (部分信号から全体信号を推定) も求められている (例: 4K 映像から 8K 映像への超解像技術)。

映像圧縮・復元の高性能化には, スパース表現を与える辞書設計が鍵となる。近年画像処理では (データ駆動) 学習型辞書が盛んに研究されており [1], 離散コサイン変換 (DCT) やウェーブレット変換 [2] 等の非学習型辞書に比べ, よりスパースな表現 (高効率スパース表現) が実現されている。

2. 研究の目的

本研究では指向性テンソル辞書に基づく高解像度映像スパース表現手法の確立を目的とした。大規模テンソルである高精細・マルチスペクトル・多視点等の高解像度映像を (信号圧縮・復元に向けて) スパースに表現するためには, 学習された大規模テンソル辞書が必要となる。しかしテンソル辞書学習はその多次元性により, 行列の場合と比べて遥かに計算負荷が高く, 特にテンソル辞書の各次元サイズが大きくなる場合学習は困難になる。本研究では, 濃淡画像の基本性質の一つである「各画素における高い指向性 (縦横斜めの方向への高い相関)」が様々な高解像度映像でも成り立つことに着目する (図 1)。本性質は様々な場面で観測でき, 例えば濃淡画像向けの辞書学習アルゴリズム KSV [1] を適用して得られた辞書のカーネルを見ると (図 2), 特定方向に振動する波形が頻出しているのが確認出来る。以上を踏まえ, まず 1) 指向性 (特定方向への分布形状) 要素を導入した非学習型大規模テンソル辞書を設計し, 2) 高解像度映像のスパース表現最適化に適用する方針で検討を進めた。



図 1: 高解像度映像の指向性

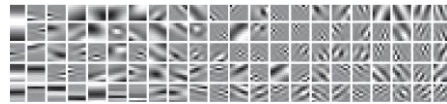


図 2: KSV 辞書に現れる指向性カーネル

3. 研究の方法

本研究課題では, 平成 29 年度, 平成 30 年度, 平成 31 年度 (令和) の 3 年計画で以下のように研究を進めた。

【平成 29 年度】

辞書に対する制約条件下 (可逆性, 直交性, テンソルのサイズ) で, 多様な多次元指向性を備えたテンソル辞書を周波数変調 (以下変調と略記) 方式によって設計するアプローチを検討した。

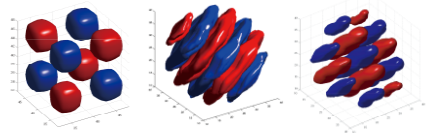
変調方式は, 他の 2 次元指向性辞書 (例: Dual-tree 複素ウェーブレット [3]) とは異なり, 可逆性, 直交性等の良好な性質を満足したまま, 大規模な辞書を容易に設計できる特徴を持つ。具体的には (低周波スペクトルを持つ) 基本ベクトルを一つ設計すれば, それの周波数領域上の平行移動 (変調) によるスペクトル配置によって全辞書要素が実現できる (図 4(a) (b) は, 図 3(a) の中・右に対応)。以上の議論を, 多次元指向性テンソル辞書に拡張するためには多次元周波数空間上でのスペクトル配置を考えればよい。実際に, 図 3(b) 中・右の形状は, それぞれ図 5(a) と (b) に示す低周波数スペクトル (中央の立方体) の変調で構成したものである。更に, 多次元空間 (3 次元, 4 次元) では図 5(c) を始めとして多様な基本スペクトル形状が考えられ, 回転や折り返し等の操作によって多様な指向性形状を実現できると考えられる。しかし回転や折り返しに基づく変調方式が存在しないため, それらの操作まで含めた一般化変調方式に基づき, 実現可能指向性形状の理論解析と設計を行った。

【平成 30 年度】

前年度の指向性テンソル辞書を高解像度映像のスパース表現に適用した。高解像度映像をスパースに表現するためには辞書が多様な指向性形状の要素を有することが望ましいが, 要素数が膨大になるほどスパース表現の計算負荷は増加する。そこで計算負荷低減のために, テスト映



(a) 二次元辞書の要素



(b) 三次元辞書の要素

図 3: 非指向性辞書と指向性辞書

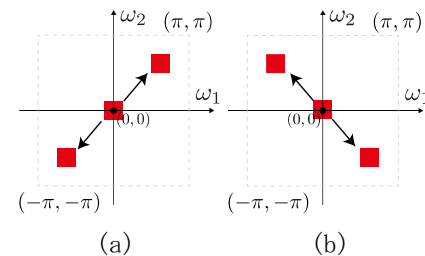


図 4: 二次元辞書のための周波数変調スペクトル配置

像群に対する事前スパース表現解析結果に基づき有効な辞書要素の事前選択 (図6 行列P の決定) を行い, 要素数を限定してから, 入力映像のスパース表現 (図6 行列B の決定) を行う (二層スパース表現). 有効な辞書要素を効率良く選択するためには, 図6 のように選択行列の非ゼロ係数がスパースに分散するだけでなく, 低ランクとなるのが望ましい. そこで低ランク性の評価も考慮するために, コスト関数に適切な核型ノルムを導入することを検討する.

【平成31年度 (令和元年度)】

前年度の辞書二層スパース表現に更に辞書選択行列を乗じて多段階に拡張することによって, 辞書の多層スパース表現へと拡張できる. ただし辞書選択行列の係数学習は更に演算量が膨大となり, 高効率な学習アルゴリズムが必要となる.

この問題を解決するべく, 本研究では各層の係数行列のパラメータの高効率な学習アルゴリズムの検討を行った. 具体的には, 各層のパラメータ行列を学習する際のコスト関数は多層の合成関数となり, 最小化の際に必要な近接写像の計算が困難となるため, 合成関数の近接写像計算のための高効率アルゴリズムを検討した.

4. 研究成果

(1) 多重解像度コサイン/サイン (CS) 変調フィルタバンクの設計

変調方式に基づく指向性テンソル辞書の設計方式として, 本研究ではコサイン/サイン変調フィルタバンクを構成した. 多次元フーリエ変換を始め, 周波数変調を用いた指向性変換がいくつか提案されているが,

- フーリエ変換のように変換中に複素数演算が含まれる (その結果出力係数は複素数になる).
- 急峻な阻止域減衰を有するフィルタ特性が実現できない.

などの問題がある. 本研究で構成したCS変調フィルタバンクは, 従来の周波数変調における変調量を修正することで実現され, 実数演算かつ急峻な阻止域減衰を有する多次元変換を設計することに成功した (図7).

また本研究では, 任意のスケールに適したサイズのカーネルを生成できる多重解像度CS変調フィルタバンクを構成した (図8).

(2) マルチスケール構造テンソル正則化

二層スパース表現の係数行列は学習データの各画像の詳細成分を効率よく表現するよう決定されなければならない. そのためには, 二層スパース表現で表された画像の局所領域が鮮明画像のエッジ情報 (例: 図9の原画像の拡大領域) を忠実に再現できている必要があり, その精度を評価するための正則化関数を適切に設計する必要がある.

その評価関数として, 本研究では映像の局所領域における勾配ベクトルセットの従属性を考慮した最適化手法を導入した. 高解像度映像の局所領域では, 時空間方向の勾配ベクトル (隣接画素の差分情報) が似通った方向を示す傾向がある. 従来手法として提案された「構造テンソル全変動量正則化」では, 勾配ベクトルの類似性を低ランク性によって特徴づけた凸最適化を構成し, 詳細なテクスチャを豊富に含む映像に対しても柔軟な表現を可能にしている. 本研究では「隣接画素の差分情報のみならず, 更に離れた場所に位置する画素との差分情報も類似する」と考え, 複数の解像度の勾配ベクトルに基づく「多重解像度構造テン

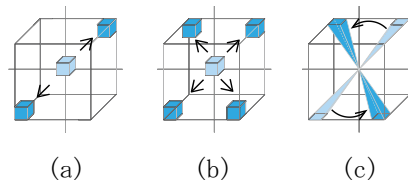


図5: 三次元辞書のための周波数変調スペクトル配置

$$\arg \min_{\mathbf{P}, \mathbf{B}} \{ \|\mathbf{Y} - \mathbf{FPB}\|_F^2 + \lambda_1 \|\mathbf{B}\|_0 + \lambda_2 \|\mathbf{P}\|_0 + \lambda_3 \|\mathbf{P}\|_* \}$$

低ランク

指向性テンソル辞書 \mathbf{F}

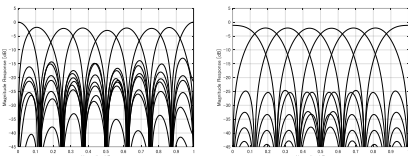
$n (\gg m)$

辞書要素選択行列 \mathbf{P}

m

×

図6: 二層スパース表現



(a) 従来法 (b) 提案法

図7: 周波数変調フィルタバンクの周波数特性

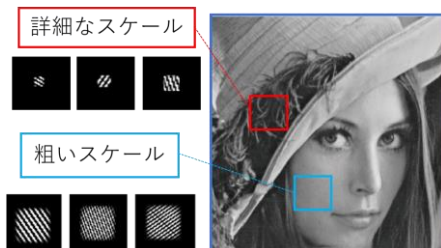


図8: 多重解像度CS変調フィルタバンク

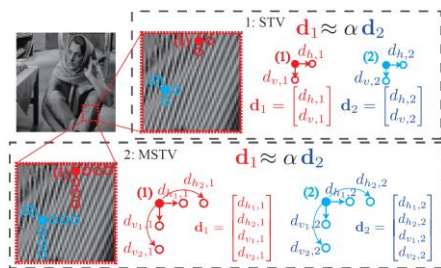


図9: STV (従来) とマルチスケールSTV (提案)

更に離れた場所に位置する画素との差分情報も類似する」と考え, 複数の解像度の勾配ベクトルに基づく「多重解像度構造テン

「ソル全変動量正則化」を導入し、頑健性の高い最適化手法を構成した (図 10)。

(3) 多チャンネル画像スパース表現のためのイントラ/インターチャンネル変換の設計

例えば RGB カラー画像や、ハイパースペクトル画像のように、多くの高解像度映像ではスペクトル情報を有しており、隣接波長に対応するスペクトル情報 (例えば図 11 に示す、カラー画像の R と G 成分, G と B 成分) には高い相関があることがよく知られている。高解像度映像をスパースに表現するためにはこの性質を活用することが重要となる。

そこで本研究ではまずチャンネル内の変換によって空間信号の相関を除去した後、チャンネル間の変換によってチャンネル間の相関を除去する方式を構成し、画像圧縮符号化にて圧縮効率の向上を実現した (図 12)。

(4) エピグラフ変形に基づく多層合成関数最小化アルゴリズム

二層スパース表現、すなわち非学習型指向性辞書に対して線形結合を作用する結合係数行列を乗じて画像を表現する方式は、更に線形結合係数を乗じることによって多層のスパース表現へと拡張できる。ただし各層のパラメータ行列は何らかのコスト関数に従って最適化しなければならないが、係数行列が多層構造を形成している観点から、コスト関数も多層の合成関数となる。ADMM に代表される近年の最適化アルゴリズムにおいてコスト関数を最小化する際には近接写像の計算が必要となる。しかし、多層の合成関数の近接写像の計算は困難となるため、間接的に近接写像を計算する代替アルゴリズムを構成する必要がある。

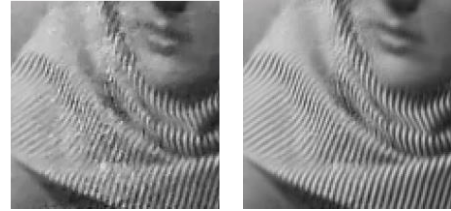
本研究ではこの問題に対してエピグラフを用いた合成関数の変形 (エピグラフ変形) を構成し、変形された問題に対する最小化アルゴリズムを開発した。図 13 のように、合成関数として定義されたコスト関数に適宜、1) 新しい変数の導入による等式制約変形、2) 等式制約から不等式制約への変形を施すことで、凸性を保ちながら効率良く近接写像が計算できることを示した。ただし本変形は不等式緩和を用いていることから、初期最適化問題から別問題に変換され、両問題の最適解が一致する保証はない。しかしながら厳密単調性という関数の性質を仮定することによって、エピグラフ変形前後の問題の最適解が一致することを理論的に証明することに成功した。

【得られた成果の国内外における位置づけとインパクト、今後の展望】

本研究では高解像度映像スパース表現のための大規模テンソル辞書の効率的な学習方法について検討を行った。非学習型辞書と多層結合行列によって様々な辞書を生成するコンセプトは非常にユニークであり、またその効果を確認することができた。

本研究の今後の展望として、近年活発に研究が進められている深層ニューラルネットワーク (DNN) 学習への応用が挙げられる。DNN は入力信号に対して行列乗算と非線形関数 (活性化関数) の処理を何層にも適用する構造を有し、画像処理等で顕著な性能を示している。しかし、ネットワーク内部パラメータ学習に関する理論解析が困難なため DNN の性質はブラックボックスとなっている部分が多く、これまでトライアンドエラーによって開発が進められてきた。

本研究の多層スパース表現は DNN との親和性が非常に高く、更にエピグラフ変形による各層のパラメータ行列学習手法などは DNN の学習にも応用できるため、理論基盤が充実した従来の凸最適化工学の知見を活用しながら、更に高度な DNN を構築できると考えられる。



(a) 従来法 (b) 提案法

図 10: STV とマルチスケール STV による圧縮センシング再構成

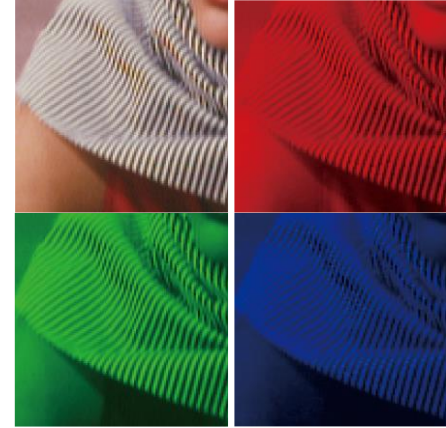


図 11: カラー画像チャンネル間相関

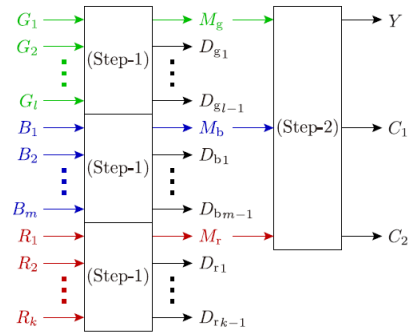


図 12: イントラ/インターチャンネル変換

$$\begin{aligned} & \underset{\mathbf{x}}{\operatorname{argmin}} f_1(f_2(\cdots f_N(\mathbf{x}) \cdots)) \\ & \underset{\mathbf{x}, \mathbf{z}_1, \dots, \mathbf{z}_N}{\operatorname{argmin}} f_1(\mathbf{z}_1) \\ & \text{s.t. } f_2(\mathbf{z}_2) \leq \mathbf{z}_1 \\ & \quad f_3(\mathbf{z}_3) \leq \mathbf{z}_2 \\ & \quad \vdots \\ & \quad f_N(\mathbf{z}_N) \leq \mathbf{z}_{N-1} \end{aligned}$$

図 13: エピグラフ変形

[参考文献]

- [1] M. Aharon, et. al., "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," IEEE Trans. Signal Process., vol 54, no. 11, pp. 4311 - 4322, 2006
- [2] G. Strang and T. Q. Nguyen, "Wavelets and filter banks," Wellesley-Cambridge Press, 1997.
- [3] R. G. B. I. W. Selesnick and N. G. Kingsbury, "The dual-tree complex wavelet transform," IEEE Signal Process. Mag., pp. 123-151, Nov. 2005.

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 3件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Ishibashi Ryoma, Suzuki Taizo, Kyochi Seisuke, Kudo Hiroyuki	4. 巻 29
2. 論文標題 Image Boundary Extension With Mean Value for Cosine?Sine Modulated Lapped/Block Transforms	5. 発行年 2019年
3. 雑誌名 IEEE Transactions on Circuits and Systems for Video Technology	6. 最初と最後の頁 1~11
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TCSVT.2017.2771461	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Matsuoka Ryo, Kyochi Seisuke, Ono Shunsuke, Okuda Masahiro	4. 巻 65
2. 論文標題 Joint Sparsity and Order Optimization Based on ADMM With Non-Uniform Group Hard Thresholding	5. 発行年 2018年
3. 雑誌名 IEEE Transactions on Circuits and Systems I: Regular Papers	6. 最初と最後の頁 1602~1613
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TCSI.2017.2763969	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ishibashi Ryoma, Suzuki Taizo, Kyochi Seisuke, Kudo Hiroyuki	4. 巻 印刷中
2. 論文標題 Image Boundary Extension with Mean Value for Cosine-Sine Modulated Lapped/Block Transforms	5. 発行年 2018年
3. 雑誌名 IEEE Transactions on Circuits and Systems for Video Technology	6. 最初と最後の頁 1~1
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TCSVT.2017.2771461	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Matsuoka Ryo, Kyochi Seisuke, Ono Shunsuke, Okuda Masahiro	4. 巻 65
2. 論文標題 Joint Sparsity and Order Optimization Based on ADMM With Non-Uniform Group Hard Thresholding	5. 発行年 2018年
3. 雑誌名 IEEE Transactions on Circuits and Systems I: Regular Papers	6. 最初と最後の頁 1602~1613
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TCSI.2017.2763969	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11件（うち招待講演 0件 / うち国際学会 5件）

1. 発表者名 Watanabe Makoto, Matsuoka Ryo, Kyochi Seisuke, Ono Shunsuke, Okuda Masahiro
2. 発表標題 Multiscale Structure Tensor Total Variation for Image Recovery
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 千々和 憂希, 京地 清介, 馬場 達也, 奥田 正浩
2. 発表標題 固有基底変換を用いた高速かつロバストな水中画像補正
3. 学会等名 電子情報通信学会信号処理研究会第33回信号処理シンポジウム
4. 発表年 2018年

1. 発表者名 Yusuke Nomura, Ryutaro Ogawa, Seisuke Kyochi, Taizo Suzuki
2. 発表標題 Multiscale directional transforms based on cosine-sine modulated filter banks for sparse directional image representation
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Kazuki Yamanaka, Seisuke Kyochi, Shunsuke Ono, and Keiichiro Shirai
2. 発表標題 Color affine subspace pursuit for color artifact removal
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Kenjiro Sugimoto, Seisuke Kyochi, and Sei-ichiro Kamata
2. 発表標題 Universal approach for DCT-based constant-time Gaussian filter with moment preservation
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Tatsuya Baba, Keishu Nakamura, Seisuke Kyochi, and Masahiro Okuda
2. 発表標題 Image enhancement method for underwater images based on discrete cosine eigenbasis transformation
3. 学会等名 IEEE International Conference on Image Processing (ICIP), 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 井本桂右, 京地清介
2. 発表標題 グラフケプストラムに基づく音響シーン分析
3. 学会等名 日本音響学会
4. 発表年 2018年

1. 発表者名 小崎祥吾, 野村優介, 京地清介, 鈴木大三
2. 発表標題 コサイン・サイン変調フィルタバンクを用いた多重解像度指向性フレームの設計
3. 学会等名 電子情報通信学会画像工学研究会
4. 発表年 2018年

1. 発表者名 友清飛海, 千々和憂希, 京地清介, 白井啓一郎
2. 発表標題 空間-スペクトル全変動正規化と二次色変換に基づく偽色抑制
3. 学会等名 電子情報通信学会画像工学研究会
4. 発表年 2018年

1. 発表者名 渡辺誠, 京地清介, 小野峻佑
2. 発表標題 マルチスケール構造テンソル全変動量に基づく画像復元
3. 学会等名 PCSJ/IMPS 2017
4. 発表年 2017年

1. 発表者名 山中和貴, 京地清介, 小野峻佑, 白井啓一郎
2. 発表標題 局所色核ノルムに基づく色アフィン部分空間追跡,"
3. 学会等名 電子情報通信学会信号処理研究会第32回信号処理シンポジウム,
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

Kyochi Laboratory -Publications- https://sites.google.com/site/s2kkyochi/publications-1?authuser=0

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