

令和 元年 5 月 12 日現在

機関番号：56203

研究種目：研究活動スタート支援

研究期間：2017～2018

課題番号：17H06572

研究課題名(和文)統計的独立性と種々の数理的構造に基づく音源分離理論とその拡張

研究課題名(英文)Extended theories of audio source separation based on statistical independence and various mathematical structures

研究代表者

北村 大地 (Kitamura, Daichi)

香川高等専門学校・電気情報工学科・助教

研究者番号：40804745

交付決定額(研究期間全体)：(直接経費) 2,300,000円

研究成果の概要(和文)：本研究課題は、音源分離技術に関する数理的・実用的拡張を目的としている。音源分離とは、複数の音源が混合した観測信号のみから混合前の個々の音源信号を推定する逆問題であり、スマートスピーカー、音声認識、補聴器デバイス、音響VR・AR等、多くの応用が期待されている。本課題では特に、従来より申請者が提案している独立低ランク行列分析(ILRMA)の確率的生成モデルの一般化について、当初の計画通り取り組み、より高精度な音源分離結果が得られることを実験的に示した。また、低ランク性以外の性質や、音源にとって適切な性質を学習データと深層学習によって獲得する「独立深層学習行列分析(IDLMA)」を新たに提案した。

研究成果の学術的意義や社会的意義

音源分離技術の精度が向上すれば、補聴器等の人支援デバイスへと直接的に応用できる他、音楽の新しい楽しみ方やVR技術への援用など、これまでの芸術・文化の振興につながることを期待されている。また、近年は音声認識やスマートスピーカー等が身近な技術となったが、これらのデバイスが雑音の多い環境下でも頑健に動作するためにも、音源分離技術の応用が必須となる。このように、音源分離技術はあらゆる音響機器のフロントエンドとして必要な最も基本的な信号処理である。また、「混合信号から潜在的な因子を推定する」という観点では、音響信号のみならず、画像や電波などあらゆるメディアへの活用も期待される。

研究成果の概要(英文)：This research project aims to improve the performance of conventional audio source separation techniques by extending their theories from mathematical and practical aspects. Audio source separation is a technique for extracting specific audio sources from the observed mixture signal. This technique can be applied for many devices and systems including hearing-aid system, smart speaker, speech recognition, and so on. In this project, the generalization of probabilistic model assumed in "independent low-rank matrix analysis (ILRMA)" (state-of-the-art audio source separation method) was carried out, and its validity was confirmed by practical experiments. Also, various types of mathematical model were introduced into ILRMA to enhance its separation quality. Furthermore, data-driven approach was newly employed to ILRMA, which was named as independent deeply learned matrix analysis. The efficacy of the proposed methods was confirmed.

研究分野：音響信号処理

キーワード：音響信号処理 統計的信号処理 音源分離 深層学習

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

音源分離技術は、複数の音源が混合された状態で観測された信号の中から、混合する前の個々の音源を推定・抽出する技術である。この技術は、音声認識の精度向上、音声通信の品質向上、補聴器システム、機械システムの異常音検知、音楽信号の能動的鑑賞、音拡張現実、音ヴァーチャルリアリティ等、様々な応用が挙げられ、社会的需要の高い技術といえる。また、機械学習の側面から捉えると、観測された信号中の有意な潜在因子を推定する技術であり、近年の深層学習やパターン認識等とも密接に関連している領域である。

1990年代から今日に至るまで研究されているが、いまだ万能な音源分離技術は提案されておらず、その品質も決して十分とはいえない状況である。申請者は、本研究課題において、音源信号間の統計的独立性に加え、行列の低ランク性・スパース性・グループスパース性等の数理的構造を積極的に活用する教師無し音源分離手法を提案し、既存手法よりも高精度な音源分離理論の確立を目指した。

2. 研究の目的

本研究課題の目的は、大きく以下の2点に大別される。

- (1) 現実の音波動の混合現象に則した新しい確率的生成モデルへの一般化  
従来の教師無し音源分離手法で仮定されてきた確率的生成モデルでは、現実の音波動の混合現象に合致しない（近似的な）モデルを利用しており、音源分離性能の低下を招いている可能性があった。そこで、本研究課題では、確率的生成モデルを「 $\alpha$ 安定分布」と呼ばれる一般的な確率密度関数に修正することで、現実の物理現象に則したモデルに基づく教師無し音源分離手法を提案する。
- (2) 低ランク性以外の性質を持つ音源信号への構造変換及び最適な構造パラメータ推定  
音声等のように音色が連続的に変化する信号では、既存手法で仮定されている行列の低ランク性が成立せず、適切な音源分離がなされない問題があった。そこで、音声信号の高品質分離のため、行列のスパース構造を仮定した音源分離手法を提案する。さらに、このような音源依存の適切な「音源構造の仮定」を学習データから自動的に見つける教師あり音源分離手法として、深層学習を援用した新しい手法を提案する。

3. 研究の方法

本研究課題では、音源信号間の統計的独立性と行列の低ランク性を仮定した教師無し音源分離手法である「独立低ランク行列分析 (ILRMA)」という既存手法（申請者が2016年に提案した技術）を基盤とし、これを一般化・高性能化することを目指す。図1に示すように、ILRMAは「1.音源間の統計的独立性」、「2.各音源の確率的生成モデル」、「3.各音源の時間周波数構造」の3つの手掛かりを利用して音源分離を達成している。一つ目の統計的独立性は、「独立成分分析 (ICA)」や「独立ベクトル分析 (IVA)」と呼ばれる従来手法から用いられている基本原理であり、ICA理論の歴史的的成功に鑑みても、音源分離に対して非常に有効な仮定といえる。二つ目の確率的生成モデルは、音源信号の従う複素分布に関する仮定である。基本的にICA・IVA・ILRMAは尤度最大化基準で音源分離を行うため、生成モデルにどのような分布を仮定するかが重要である。特にILRMAでは、現実の音波動の混合現象（複素数の時間周波数値の和）を非負の時間周波数値の和で正当にモデリングする際の領域が生成モデルに依存する。従来のILRMAは、複素ガウス分布を生成モデルに用いることで、パワー値の和を正当化しているが、分離系の推定に最適な領域は未だ明らかにされていない。三つ目の音源構造に関する仮定は、実際に分離したい音源に強く依存する要素である。低ランク構造の有効性は申請者によって示されたが、図6に示すように、その他の構造仮定の具体的な効果についても検証する必要がある。さらに、このような音源構造の仮定は経験的な知識に基づいて決める以外にも、学習データが手に入る場合には、深層学習を導入して音源分離に効果的なパラメータとして最適化できる。

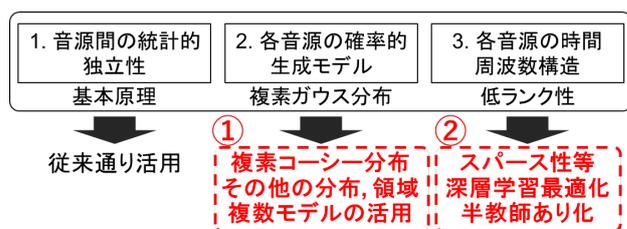


図5: ILRMAで音源分離に活用される要素とそれらの一般化及び拡張を示す。破線で囲った①及び②が具体的に実施する課題に対応している。

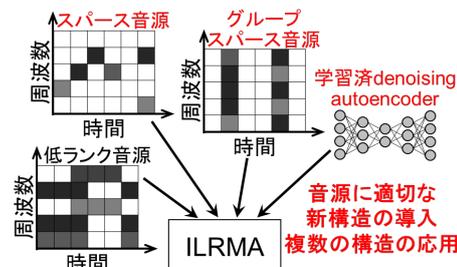


図6: 種々の制約の導入及び深層学習 (denoising autoencoder) による学習済最適制約の適用

#### 4. 研究成果

平成29年度は、従来より申請者が提案している独立低ランク行列分析 (ILRMA) の確率的生成モデルの一般化について、当初の計画通り取り組んだ。具体的には、従来の ILRMA が用いていた音源の生成モデルである複素ガウス分布を、一般化複素ガウス分布及び学生 t 分布の2種類に一般化し、それらの理論的・実用的妥当性について検証した。いずれの一般化拡張においても、従来の ILRMA の利点である高速な最適化・初期値に対する頑健性を保ったまま、より高精度な音源分離結果が得られることを示している。

また、当初の計画では平成30年度に取り組む予定であった「低ランク性とは異なる性質を持つ音源信号への構造変換及びその最適化」について、その一部を平成29年度に取り組んだ。これは、ILRMA のような「音源の性質」と「統計的独立性」の2つの性質を用いた音源分離技術をより一般的なフレームワークへと拡張するための検討課題である。平成29年度では、低ランク性以外の性質（グループスパース性、低ランク性+スパース性等）を用いた手法に加え、音源にとって適切な性質を学習データと深層学習によって獲得する「独立深層学習行列分析 (IDLMA)」を新たに提案した。実験的な評価により、IDLMA が既存の音源分離手法と比べて驚異的な性能改善を実現できることを確認した。

平成30年度は、IDLMA のデータ拡張タスクについて取り組んだ。IDLMA において音源モデルを事前に構築するためには、学習データとして完全に分離された音源信号が大量に必要となる。現実的には、そのような学習データを用意できる状況は非常に稀であるため、より少ないデータからその特質を捉え、音響的に適切な制約の下で音源信号を新たに生成する方法について提案した。これは即ち、少数のデータから深層ニューラルネットワークを転移学習する「半教師あり IDLMA」と捉えられる。このような拡張は、より現実の問題解決に即したアルゴリズムといえる。この半教師 IDLMA について実験的に性能を確認したところ、事前学習データを大量に利用する全教師あり IDLMA と同程度の音源分離が達成できることを確認した。これらの成果については、現在論文を投稿中である。

以上より、本研究課題で目標とした事項はおおむね達成しており、学術論文3編と学会発表34件を実施するに至った。また、得られた成果のさらなる検証と発展については、以後の研究活動を通して実施し、最終的には補聴器への応用を見据えた高精度なリアルタイム音源分離システムの開発を目指す。

#### 5. 主な発表論文等

##### 〔雑誌論文〕 (計3件)

1. Hiroaki Nakajima, Daichi Kitamura, Norihiro Takamune, Hiroshi Saruwatari, and Nobutaka Ono, "Bilevel optimization using stationary point of lower-level objective function for discriminative basis learning in nonnegative matrix factorization," *IEEE Signal Processing Letters*, vol. 26, no. 6, pp. 818-822, June 2019.  
DOI: 10.1109/LSP.2019.2909079
2. Shinichi Mogami, Yoshiki Mitsui, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari, Yu Takahashi, Kazunobu Kondo, Hiroaki Nakajima, and Hirokazu Kameoka, "Independent low-rank matrix analysis based on generalized Kullback-Leibler divergence," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E102-A, no. 2, pp. 458-463, February 2019.  
DOI: 10.1587/transfun.E102.A.458
3. Daichi Kitamura, Shinichi Mogami, Yoshiki Mitsui, Norihiro Takamune, Hiroshi Saruwatari, Nobutaka Ono, Yu Takahashi, and Kazunobu Kondo, "Generalized independent low-rank matrix analysis using heavy-tailed distributions for blind source separation," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, p. 28, May 2018.  
DOI: 10.1186/s13634-018-0549-5

##### 〔学会発表〕 (計34件)

1. Naoki Makishima, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari, Yu Takahashi, and Kazunobu Kondo, "Column-wise update algorithm for independent deeply learned matrix analysis," *Proceedings of International Congress on Acoustics (ICA 2019)*, Aachen, Germany, September 2019 (accepted).
2. Naoki Makishima, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari, Yu Takahashi, Kazunobu Kondo, and Hiroaki Nakajima, "Generalized-Gaussian-distribution-based independent deeply learned matrix analysis for multichannel audio source separation," *Proceedings of International Congress and Exhibition on Noise Control Engineering (INTERNOISE 2019)*, Madrid, Spain, June 2019 (accepted).
3. Kohei Yatabe and Daichi Kitamura, "Time-frequency-masking-based determined BSS with

- application to sparse IVA," Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2019), Brighton, U.K., May 2019 (accepted).
4. Shinichi Mogami, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari, Yu Takahashi, Kazunobu Kondo, Hiroaki Nakajima, and Nobutaka Ono, "Independent low-rank matrix analysis based on time-variant sub-Gaussian source model," Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018), pp. 1684-1691, November 2018.
  5. Masakazu Une, Yuki Saito, Shinnosuke Takamichi, Daichi Kitamura, Ryoichi Miyazaki, and Hiroshi Saruwatari, "Generative approach using the noise generation models for DNN-based speech synthesis trained from noisy speech," Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018), pp. 340-344, November 2018.
  6. Shinnosuke Takamichi, Yuki Saito, Norihiro Takamune, Daichi Kitamura, and Hiroshi Saruwatari, "Phase reconstruction from amplitude spectrograms based on von-Mises-distribution deep neural network," Proceedings of International Workshop on Acoustic Signal Enhancement (IWAENC 2018), pp. 286-290, Tokyo, Japan, September 2018.
  7. Shinichi Mogami, Hayato Sumino, Daichi Kitamura, Norihiro Takamune, Shinnosuke Takamichi, Hiroshi Saruwatari, and Nobutaka Ono, "Independent deeply learned matrix analysis for multichannel audio source separation," Proceedings of European Signal Processing Conference (EUSIPCO 2018), pp. 1571-1575, Roma, Italy, September 2018.
  8. Kohei Yatabe and Daichi Kitamura, "Determined blind source separation via proximal splitting algorithm," Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018), pp. 776-780, Calgary, Canada, April 2018.
  9. Yoshiki Mitsui, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari, Yu Takahashi, and Kazunobu Kondo, "Vectorwise coordinate descent algorithm for spatially regularized independent low-rank matrix analysis," Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018), pp. 746-750, Calgary, Canada, April 2018.
  10. Moe Takakusaki, Daichi Kitamura, Nobutaka Ono, Shoji Makino, Takeshi Yamada, and Hiroshi Saruwatari, "Ego-noise reduction for hose-shaped rescue robot using basis-shared semi-supervised independent low-rank matrix analysis," Proceedings of RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP 2018), pp. 351-354, Hawaii, USA, March 2018.
  11. Yoshiki Mitsui, Daichi Kitamura, Norihiro Takamune, Hiroshi Saruwatari, Yu Takahashi, and Kazunobu Kondo, "Independent low-rank matrix analysis based on parametric majorization-equalization algorithm," Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2017), pp. 98-102, Curaçao, Dutch Antilles, December 2017.
  12. 福重敢太, 高宗典玄, 北村大地, 猿渡洋, 池下林太郎, 中谷智広, "収束保証型独立半正定値テンソル分析に基づくブライント音源分離," IEICE Technical Report, EA2018-127, vol. 118, no. 496, pp. 167-172, Nagasaki, March 2019.
  13. 久保優騎, 高宗典玄, 北村大地, 猿渡洋, "ブライント音源分離における多変量複素 Student's  $t$  分布に基づくランク制約付き空間共分散モデルの推定," IEICE Technical Report, EA2018-128, vol. 118, no. 496, pp. 173-178, Nagasaki, March 2019.
  14. 牧島直輝, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, 中嶋広明, "時変複素一般化ガウス分布に基づく独立深層学習行列分析," 日本音響学会 2019 年春季研究発表会講演論文集, 1-6-7, pp. 189-192, Tokyo, March 2019.
  15. 久保優騎, 高宗典玄, 北村大地, 猿渡洋, "乗算型更新式に基づくランク制約付き空間共分散モデルの推定," 日本音響学会 2019 年春季研究発表会講演論文集, 2-6-1, pp. 245-248, Tokyo, March 2019.
  16. 最上伸一, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, 中嶋広明, "独立低ランク行列分析における majorization-equalization アルゴリズムを用いた空間パラメータの高速更新," 日本音響学会 2019 年春季研究発表会講演論文集, 2-6-2, pp. 249-252, Tokyo, March 2019.
  17. 矢田部浩平, 北村大地, "白色化の影響を考慮したスパース独立ベクトル分析," 日本音響学会 2019 年春季研究発表会講演論文集, 2-6-5, pp. 259-260, Tokyo, March 2019.
  18. 牧島直輝, 最上伸一, 高宗典玄, 高道慎之介, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, 中嶋広明, "教師あり及び半教師あり条件下における独立深層学習行列分析の実験的評価," 日本音響学会 2019 年春季研究発表会講演論文集, 2-Q-28, pp. 385-388, Tokyo, March 2019.
  19. 牧島直輝, 最上伸一, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, 中嶋広明, "ヘビータイル生成モデルに基づく独立深層学習行列分析による多チャンネル音源分離,"

- Proceedings of 33th Signal Processing Symposium (SIP Symposium), B2-3, pp. 202-207, Tokyo, November 2018.
20. 高道慎之介, 齋藤佑樹, 高宗典玄, 北村大地, 猿渡洋, “方向統計 DNN に基づく振幅スペクトログラムからの位相復元,” 日本音響学会 2018 年秋季研究発表会講演論文集, 2-4-2, pp. 1127-1130, Oita, September 2018.
  21. 最上伸一, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, 中嶋広明, 小野順貴, “一般化反復射影法に基づく時変劣ガウス独立低ランク行列分析,” 日本音響学会 2018 年秋季研究発表会講演論文集, 3-1-10, pp. 215-218, Oita, September 2018.
  22. 久保優騎, 高宗典玄, 北村大地, 猿渡洋, “独立低ランク行列分析を用いたフルランク空間共分散モデルに基づくブラインド音源分離,” 日本音響学会 2018 年秋季研究発表会講演論文集, 3-1-9, pp. 211-214, Oita, September 2018.
  23. 牧島直輝, 高宗典玄, 高道慎之介, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, 中嶋広明, “半教師あり独立深層学習行列分析におけるデータ拡張に基づく音源モデルの適応,” 日本音響学会 2018 年秋季研究発表会講演論文集, 2-1-8, pp. 181-184, Oita, September 2018.
  24. 矢田部浩平, 北村大地, “一般の時間周波数マスキングに基づく独立ベクトル分析,” 日本音響学会 2018 年秋季研究発表会講演論文集, 3-1-11, pp. 219-220, Oita, September 2018.
  25. 相場亮人, 吉田実, 後藤理, 北村大地, 高道慎之介, 猿渡洋, “雑音下異常検知における前処理としての NMF 音源抽出手法の検討,” Proceedings of 119th IPSJ Special Interest Group on Music and Computer (IPSJ-SIGMUS), vol. 2018-MUS-119, no. 33, Tokyo, Jun 2018.
  26. 高道慎之介, 齋藤佑樹, 高宗典玄, 北村大地, 猿渡洋, “von Mises 分布 DNN に基づく振幅スペクトログラムからの位相復元,” Proceedings of 119th IPSJ Special Interest Group on Music and Computer (IPSJ-SIGMUS), vol. 2018-MUS-119, no. 54, Tokyo, Jun 2018.
  27. 北村大地, 角野隼斗, 高宗典玄, 高道慎之介, 猿渡洋, 猿渡洋, 小野順貴, “独立深層学習行列分析に基づく多チャンネル音源分離の実験的評価,” IEICE Technical Report, EA2017-104, vol. 117, no. 515, pp. 13-20, Okinawa, March 2018.
  28. 北村大地, 高宗典玄, 最上伸一, 三井祥幹, 猿渡洋, 高橋祐, 近藤多伸, “ヘビーテイルな分布に基づく非負値行列因子分解を用いたスパース雑音除去,” 日本音響学会 2018 年春季研究発表会講演論文集, 1-4-14, pp. 441-444, Saitama, March 2018.
  29. 角野隼斗, 北村大地, 高宗典玄, 高道慎之介, 猿渡洋, 小野順貴, “独立深層学習行列分析に基づく多チャンネル音源分離,” 日本音響学会 2018 年春季研究発表会講演論文集, 1-4-16, pp. 449-452, Saitama, March 2018.
  30. 三井祥幹, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, 中嶋広明, “空間モデル正則化を用いた独立低ランク行列分析に基づくブラインド音源分離,” 日本音響学会 2018 年春季研究発表会講演論文集, 1-4-15, pp. 445-448, Saitama, March 2018.
  31. 最上伸一, 三井祥幹, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, 中嶋広明, 亀岡弘和, “I ダイバージェンスに基づく独立低ランク行列分析の実験的評価,” 日本音響学会 2018 年春季研究発表会講演論文集, 1-4-8, pp. 423-426, Saitama, March 2018.
  32. 三井祥幹, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, “空間事前情報を用いた独立低ランク行列分析,” Proceedings of 32th Signal Processing Symposium (SIP Symposium), B8-2, pp. 360-365, Iwate, November 2017.
  33. 最上伸一, 三井祥幹, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸, “I ダイバージェンスを用いた独立低ランク行列分析,” Proceedings of 32th Signal Processing Symposium (SIP Symposium), B8-1, pp. 354-359, Iwate, November 2017.
  34. 北村大地, 小野順貴, 澤田宏, 亀岡弘和, 猿渡洋, “独立低ランク行列分析に基づくブラインド音源分離,” IEICE Technical Report, EA2017-56, vol. 117, no. 255, pp. 73-80, Toyama, October 2017.

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ等

研究成果のデモンストレーション : <http://d-kitamura.net/demo.htm>

## 6. 研究組織

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。