

論文審査の結果の要旨および担当者

報告番号	※ 甲第	号
------	------	---

氏 名 関 翔悟

論 文 題 目

A Study on Utilization of Prior Knowledge in Underdetermined  
Source Separation and Its Application

(劣決定音源分離における事前情報の活用とその応用に関する研究)

論文審査担当者

主 査 名古屋大学教授 武田 一哉

委 員 名古屋大学教授 戸田 智基

委 員 名古屋大学准教授 工藤 博章

委 員 名古屋大学特任教授 中岩 浩巳

関翔悟君提出の学位論文「A Study on Utilization of Prior Knowledge in Underdetermined Source Separation and Its Application (劣決定音源分離における事前知識の活用とその応用に関する研究)」は、複数音響信号の重畳に関して、源信号数よりも少ない数の観測信号群から源信号を推定する問題である「劣決定音源分離問題」を扱った論文であり7章から構成される。

第1章は研究の背景を概観し、研究の焦点(源信号に関する知識の活用)を明示している。第2章は音源分離に関する従来研究、すなわち行列分解、教師あり行列分解、ディープニューラルネットワークを用いる研究を概観するとともに、研究で用いる音源分離性能の評価指標について概観している。

第3章では、複数楽器の演奏信号の重畳を対象に取り上げ、楽器それぞれが持つ調波構造の特徴を事前知識として用いることで音源分離の性能を改善する方法を議論している。すなわち、行列分解による音源分離処理の目的関数に、混合信号の再現性に加え、推定された音源信号の調波構造と事前に学習された楽器毎の調波構造との類似度を用いることを提案している。提案方法は、3から4種類の楽器の合奏から個々の楽器の信号を抽出する問題において実験的に有効性が確認されている。また、新たに加えた調波構造の類似度の計算に際しては、調波構造をメルケプストラム空間の確率分布で計算すること、確率分布には混合ガウス分布を使うこと、類似度には、KLダイバージェンスを用いることなどの工夫がなされている。

第4章では、特定の時刻・帯域が欠落した音声信号波形の回復を取り上げる。ここでは時刻・帯域が欠落した信号のパワースペクトルに行列分解を適用するが、これは数理的には第3章で扱った劣決定の信号分離と同じ構造と見なすことができる。近年の研究で提案されたTSF(Time-domain Spectrogram Factorization)法を発展させることで、この問題を解くことを提案している。TSF法を発展させることで、波形再生における振幅スペクトルと位相スペクトルの冗長性を利用することが可能になる。模擬的に作成した時間・帯域に欠落が生じた音声を用いた評価実験の結果、従来法である行列分解の結果が改善され、提案法の有効性が明らかになった。

第5章では、混合された音声の分離を取り上げ、混合過程のモデル化において、行列分解とVAE(Variational AutoEncoder)の統合(GMVAE法)を提案している。VAEによるモデル化の部分では潜在符号を活用することで、教師有り・無し of 学習を柔軟に組み替えることができるなど、行列分解に比べて有利な点が多い。また、予備的な実験によっても、行列分解に比べてVAEによってより忠実度の高い音声生成が行われていることも確認されている。さらに本章では、空間的に混合された3人の音声を2本のマイクで収録し、源音声に分離する実験を、様々な条件で行い、GMVAE法が行列分解に対して優れた性能を発揮することが確認されている。

第6章では、研究の成果を価値創造につなげるために必要なデータ循環について考察している。第7章では論文をまとめ、残された問題を整理している。

以上、関翔悟君提出の学位論文は劣決定条件下の信号分離という本質的に困難な問題に様々な数理モデルを適用するとともに、モデルのパラメータを頑健に推定するために現実的な仮定を置くことで、従来方式の性能を改善している。また、理論の実験的検証に留まることなく、価値創造への展開についても言及していることから、学術的・産業的に寄与するところが大きい。これらのことから、論文提出者の関翔悟君は博士(情報学)の学位を授与するに相応しいと判断した。