

チームの目標：実世界における音環境理解

課題：実環境の音源・残響・雑音は未知かつ多様性に富む → 遠隔では限定された種類のコマンド音声認識が限界

目標：あらゆる環境に適応しあらゆる音響信号（音声・音楽・環境音・雑音）を対象とした音環境理解技術の開発

音源分離・音源定位・認識・雑音抑圧・残響除去・マイクキャリブレーションなどを自動的に実行・同時最適化

アプローチ：深層ベイズ学習の新しい枠組みと音環境理解への応用展開

言語モデル：音声スペクトルの深層生成モデル（事前学習）

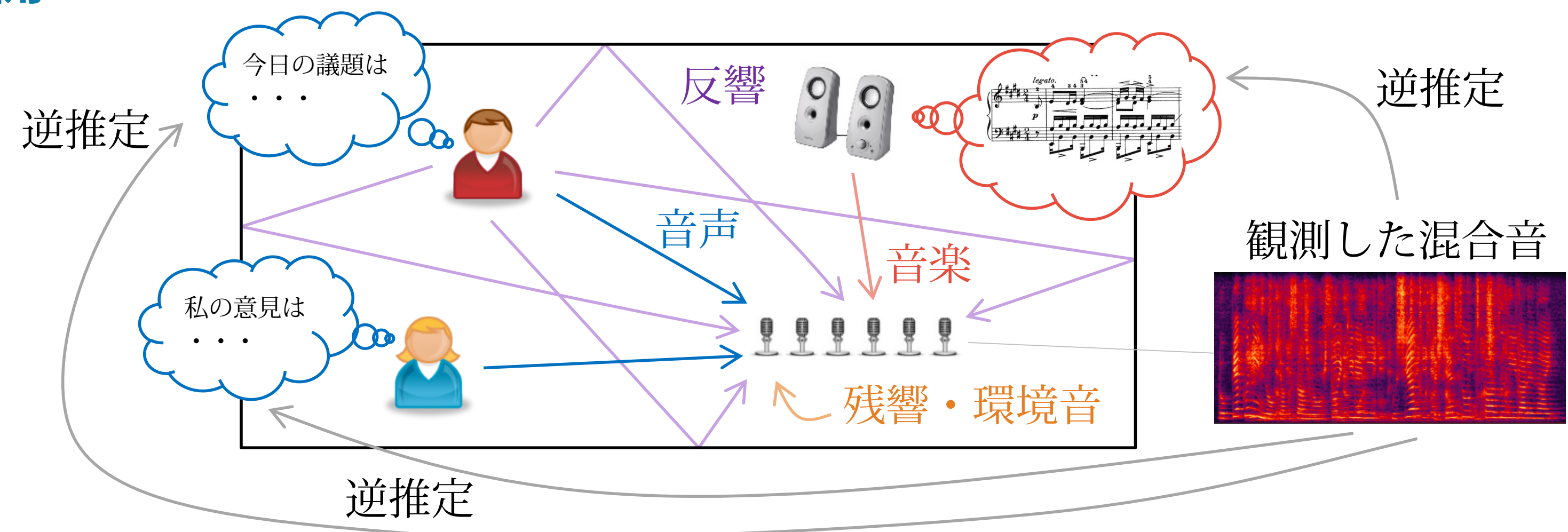
音響モデル：音響信号の重畳・伝播モデル（on-the-fly 学習）

現在の成果：音響モデル・深層ベイズ学習に関して大きな進展

NMF の究極的な拡張である CTF と ILRTA に基づく音源分離

VAE と NMF を内包する階層ベイズモデルに基づく音声強調

音声信号の複素スペクトログラムの深層生成モデル



基礎：非負値性・低ランク性・独立性に基づくブラインド音源分離の統一理論

主要な音源分離法が
特殊形として
統一的に説明可能

変換学習NMF
(TL-NMF) [Fagot+, Wendt+ 2018]

高速複数チャンネルNMF
(FastMNMF) [Ito+ 2019, Sekiguchi+ 2019]
独立低ランク行列分析
(ILRMA) [Kitamura+ 2016]

周波数
共分散

時間
共分散

非負値行列/テンソル分解
(NMF/NTF) [Févotte 2009]

同時対角化共分散モデル

複数チャンネルも
同様の枠組み

独立低ランク
テンソル分析 (ILRTA)
[Yoshii+ 2018]

独立半正定値テンソル分析
(IPSDTA) [Ikeshita 2018]
半正定値テンソル分解
(PSDTF) [Yoshii+ 2013]

制約なし共分散モデル

複数チャンネルも
同様の枠組み

相関テンソル分解 (CTF)
[Yoshii 2017]

複数チャンネルNMF
(MNMF) [Sawada+ 2013]
チャンネル
共分散

周波数
共分散

非負値行列/テンソル分解
(NMF/NTF) [Févotte 2009]

半正定値テンソル分解
(PSDTF) [Yoshii+ 2013]

軸の線形変換による要素の無相関化：独立ベクトル分析 (IVA) [Ono 2011]

応用：深層生成モデルを用いた半教師ありマルチチャンネル音声強調

