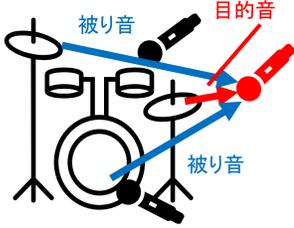


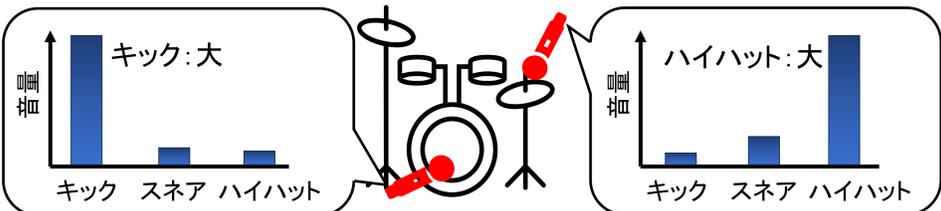
1. 研究背景

- ドラムセットを収録する際のマイキング
 - スネアやハイハットなど各ドラムパーツに**近接させて**マイクを配置
 - 近接しているドラムパーツの音(目的音)のみ録音することを狙う
 - 理想的に各ドラムパーツの音のみが録音されれば, 下記が実現可能
 - パーツ毎に最適な処理(ゲイン, EQ, コンプ等)
 - 音響信号からのドラム譜面の自動採譜
 - ドラム演奏の教育支援 [細谷ら, 2021]
- ドラムの録音における被り音の問題
 - 目的音以外に混ざり込む音を**被り音**という
 - 目的音**を残し**被り音**を抑圧することが望まれる
- 被り音抑圧問題の難しさ
 - 複数のマイクを用いた音源分離手法としてブラインド音源分離(BSS)が代表的
 - BSSはマイクアレイの使用を前提(マイク間隔が狭い録音条件)
 - 音波の到達時間差(位相差)の情報を最大限活用
 - ドラムのマイキングではマイク間隔が数m程度広がるため, **空間エイリアシング問題**により, BSSでの被り音抑圧は困難



着眼点

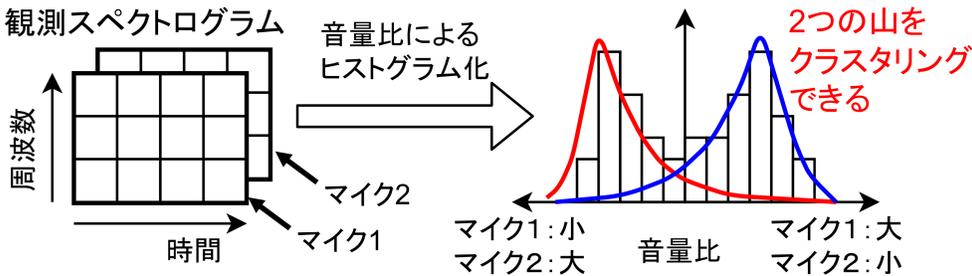
- ドラムのマイキングは各マイクをドラムパーツに近接させている
 ➡ 各マイクの音源間の音量比が大きく異なると仮定できる



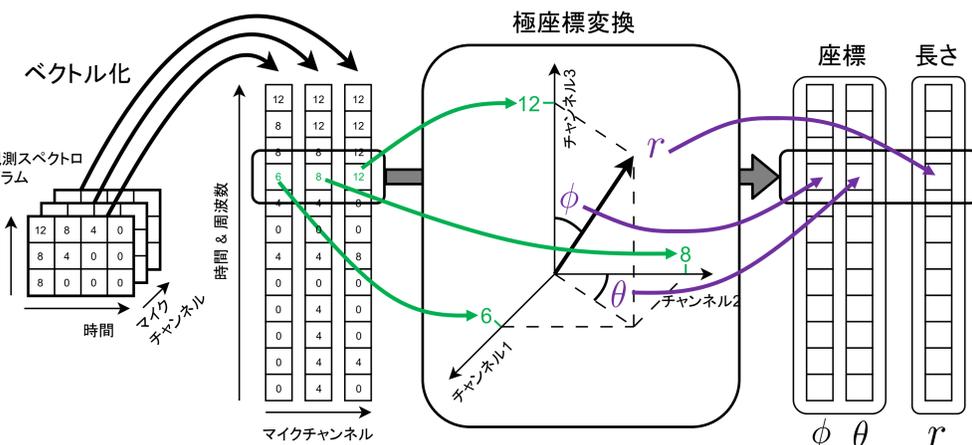
各音源の音量比の違いが被り音抑圧の手掛かりになる

2. 提案手法

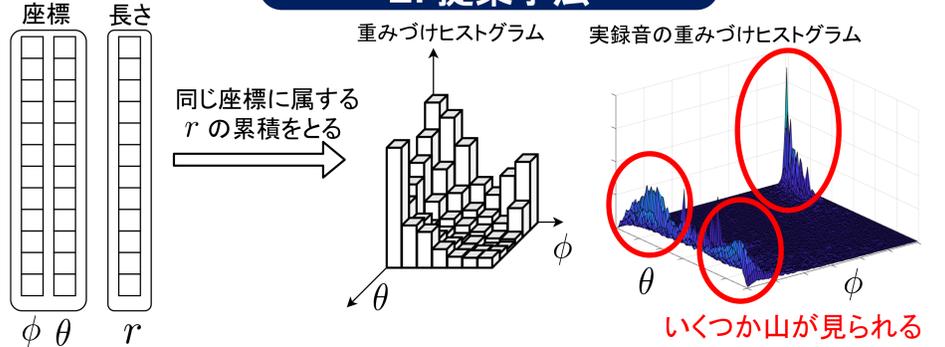
- 仮定する状況とアプローチ
 - あるドラムパーツに近接させているマイクでは, 目的音が大きな音量で観測されている
 ➡ 振幅比に基づく**教師無しクラスタリング**による音源分離が可能



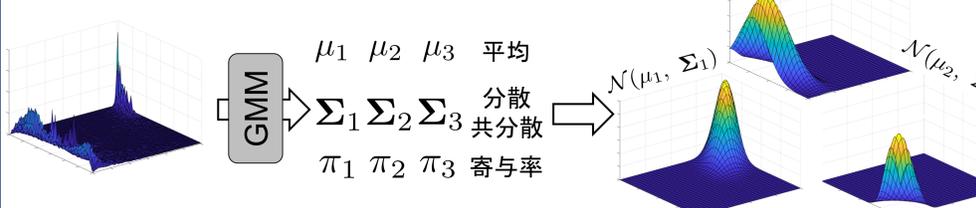
- マイク間の振幅比と絶対値による特徴量の重みづけヒストグラム
 - マイク間の時間周波数毎の観測振幅比で音量差を表現
 - 絶対的な音量が大きい時間周波数成分を重みづけ



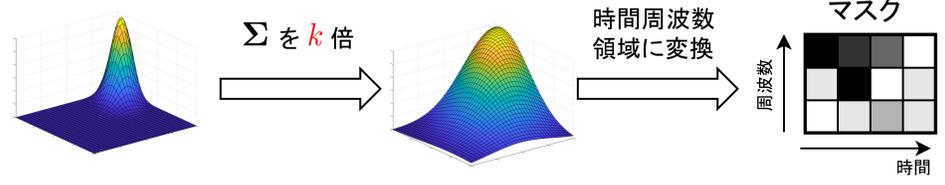
2. 提案手法



- 混合ガウス分布モデル(GMM)によるフィッティング
 - 散布図データを複数のガウス分布の組み合わせで近似 $\mathcal{N}(\mu_3, \Sigma_3)$
 - 各ガウス分布の平均, 分散・共分散, 寄与率を導出



- 時間周波数マスクの生成と被り音の抑圧
 - 観測スペクトログラムにマスクをかけることで混合音の分離を実現
 - 1. GMMで得られた各分布の母数(μ と Σ)を抽出
 - 2. 分散共分散行列 Σ を k 倍
 - 3. θ, ϕ の特徴量空間上のガウス分布 $\mathcal{N}(\mu_i, \Sigma_i)$ を時間周波数領域に変換
 - 4. 観測スペクトログラムとマスクの要素積が推定分離スペクトログラムとなる



3. 実験条件・結果

条件	設定値	ドラムパターン
サンプリング周波数	48000 Hz	
STFTの窓関数, 窓長, シフト長	Blackman, 2048, 1024	

