

# 音源分離を用いた歓声のリアルタイム認識の検討

金沢工業大学 中沢研究室 4EP2-36 丹下陽生

## 研究の背景

### 現状

近年、コロナ禍の影響が減少し、ライブやコンサート会場の声出しが解禁されつつある。そこで、観客の歓声を取り出すことで観客とマッチした演出をリアルタイムで実現できる可能性を考えた。

### 問題点

既存の技術ではライブ会場などの複雑な音響下でのデータを取り扱うことがほとんどなく、そういった環境下でも、分離ができるかどうか分からない。

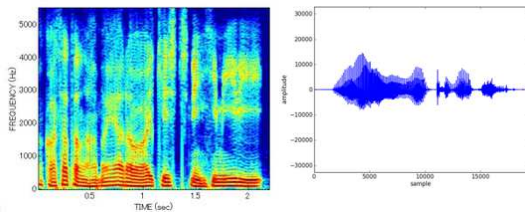
## システム概要

### 1.ビームフォーミング法による音源分離

→会場は既知であることを前提とした場合、音源位置が予測できることからこの手法に

### 2.分離音声を可視化

→スペクトログラム、音声波形を使用



### 3.画像認識処理で盛り上がり度を検出

→自分で判別したデータを用いて学習させる

### 最終目標

これらの処理を  
リアルタイムで行う

## 実験方法

### 1.ビームフォーミング法による音源分離

想定される実環境を作り、マイクロフォン、スピーカーの配置や数を変更し、音声データを取得、それらを分離する。

### 2.分離音声を可視化

取得した音声データからスペクトログラムや波形を取得し、歓声あり、歓声なし、無音（音楽なし）の3パターンのグラフから評価する。

### 3.画像認識で盛り上がり度を検出

指標を2~5段階を設定（dBの差を基準）、2から得たグラフを処理し、歓声なしとありの差を盛り上がり度を検出

dBの差	0 ~10	11 ~20	21~ 30	31~ 40	41~
指標	1	2	3	4	5

基準の1例：5段階評価とする場合（仮）

## 評価方法

- ・得られた図から正確に盛り上がり指数を判別できるかどうか。
- ・リアルタイム処理にかかる秒数の評価（目標3秒）

## 今後の展望

ライブ会場に類似した環境で音声データを取得し、その音声で盛り上がり指数を判別可能か、リアルタイムで一連の処理ができるかを確認する。