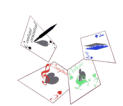


# 音情報編集のための音色を 可視化するインタフェース

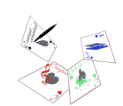
関西大学 総合情報学部 総合情報学科  
教授 山西 良典

2025年9月18日



# 音を 聴く

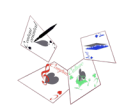




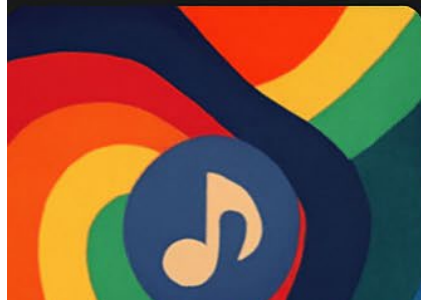
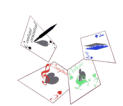
# 音を 視る







視えてるんじゃない？

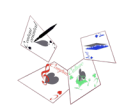


# 音は視えてない





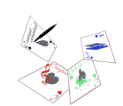




# 従来技術とその問題点

覚えなれないといけなない

同時比較が難しい



# 新技術の特徴・従来技術との比較

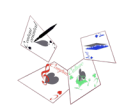
覚えなれないといけない

形象化して記録

同時比較が難しい

ネットワークで比較

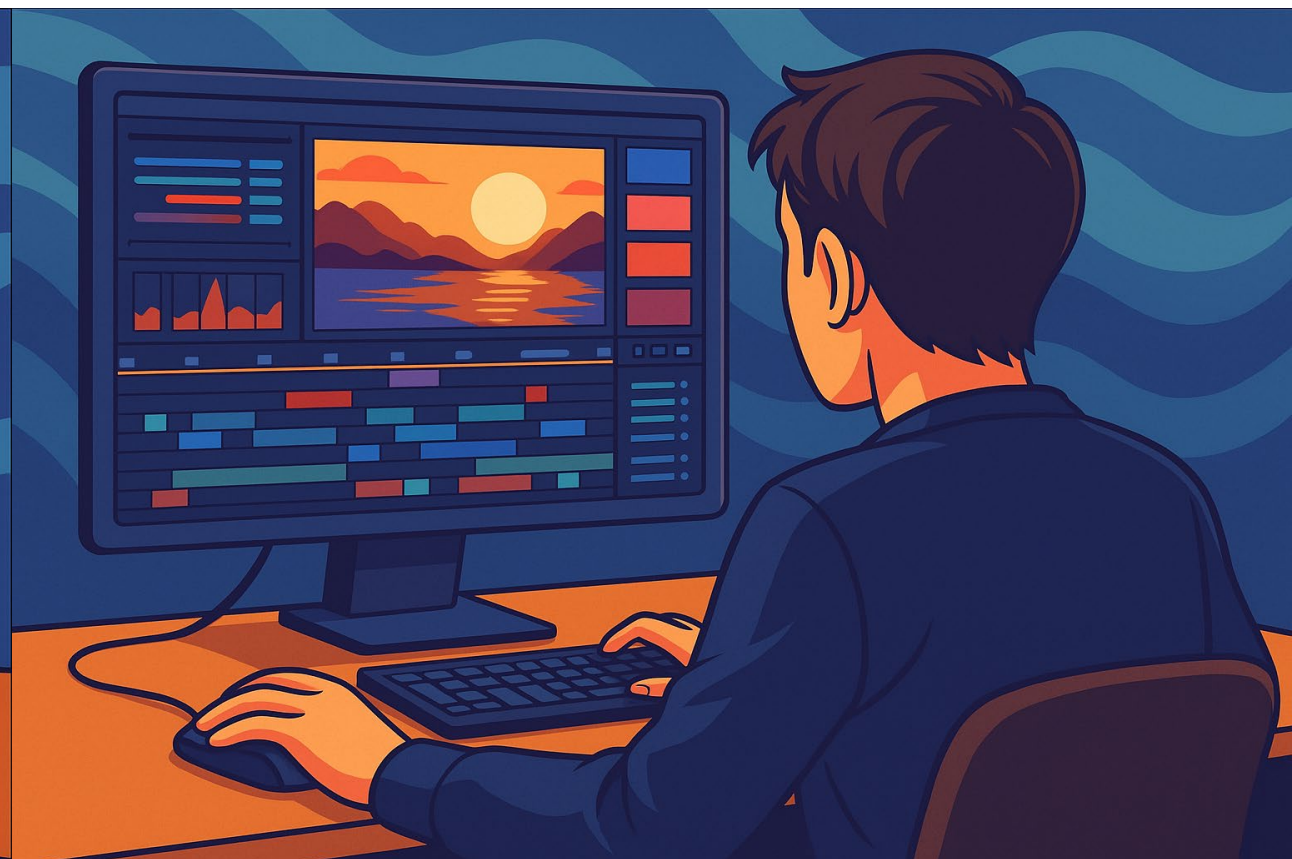


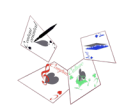


# 応用先

## 音楽制作

## 映像編集





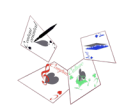
# 応用先

## 音楽制作

## 映像編集



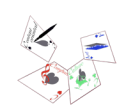
すべての音を検索する場面で



# 構成する2種類の技術

1. 音を漫画の吹き出し形状のようにして保存
2. 音の類似性をネットワークにして可視化

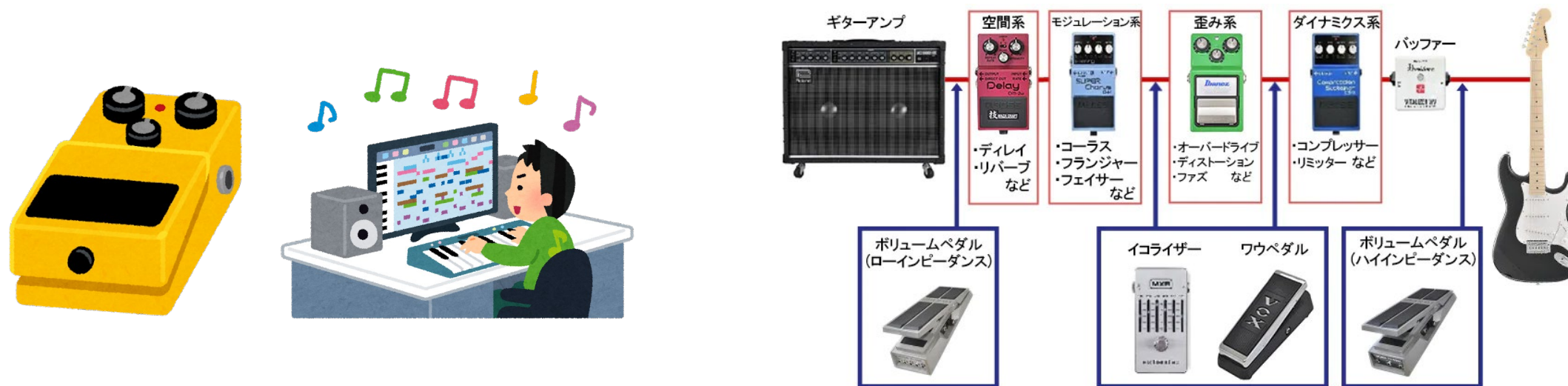




# 1. 音を漫画の吹き出し形状 のようにして保存

# 研究の背景：音響パラメータの編集（音作り）

- 複数のエフェクターの組み合わせや，様々な役割の音響特性をパラメータで調整したりすることで実現
- 音色に対する印象が明確でない初心者には困難な作業



引用元：エフェクターのつなぎ方と順番 | 空間系～補正系まで

<https://www.soundhouse.co.jp/howto/guitar/effector-setting/>

# 音作りの難しさ

音色とパラメータ情報は一過性の性質を持つ



## 音色

- 時間的メディア: 音色

音色A

音色B

順番に比較

t

- 空間的メディア: イラスト

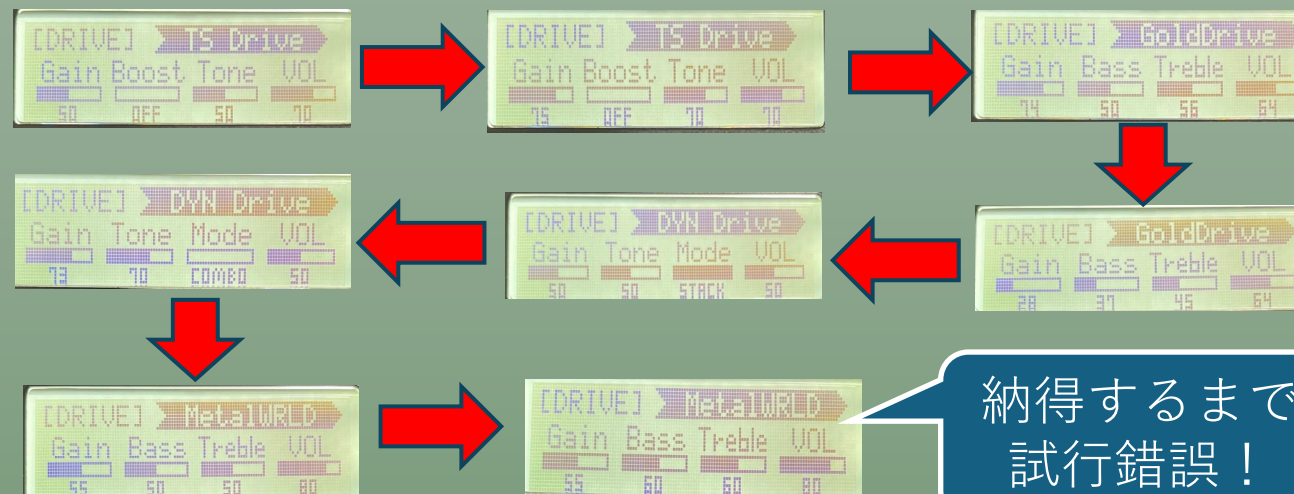
イラストA

イラストB

t

同時に比較可能

## パラメータ情報



納得するまで  
試行錯誤!



# 問題解決のためのアプローチ

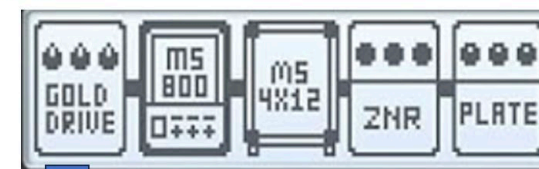
- 時間的メディアの情報が持つ意味や印象をシンボル化  
例) 五線譜, 音符, 吹き出し



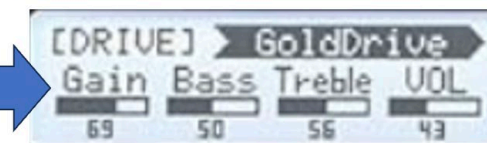
時間的メディアの情報を空間的メディアとして表現することで視覚的に解釈可能に！！

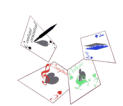
吹き出し図形

編成情報



パラメータ情報

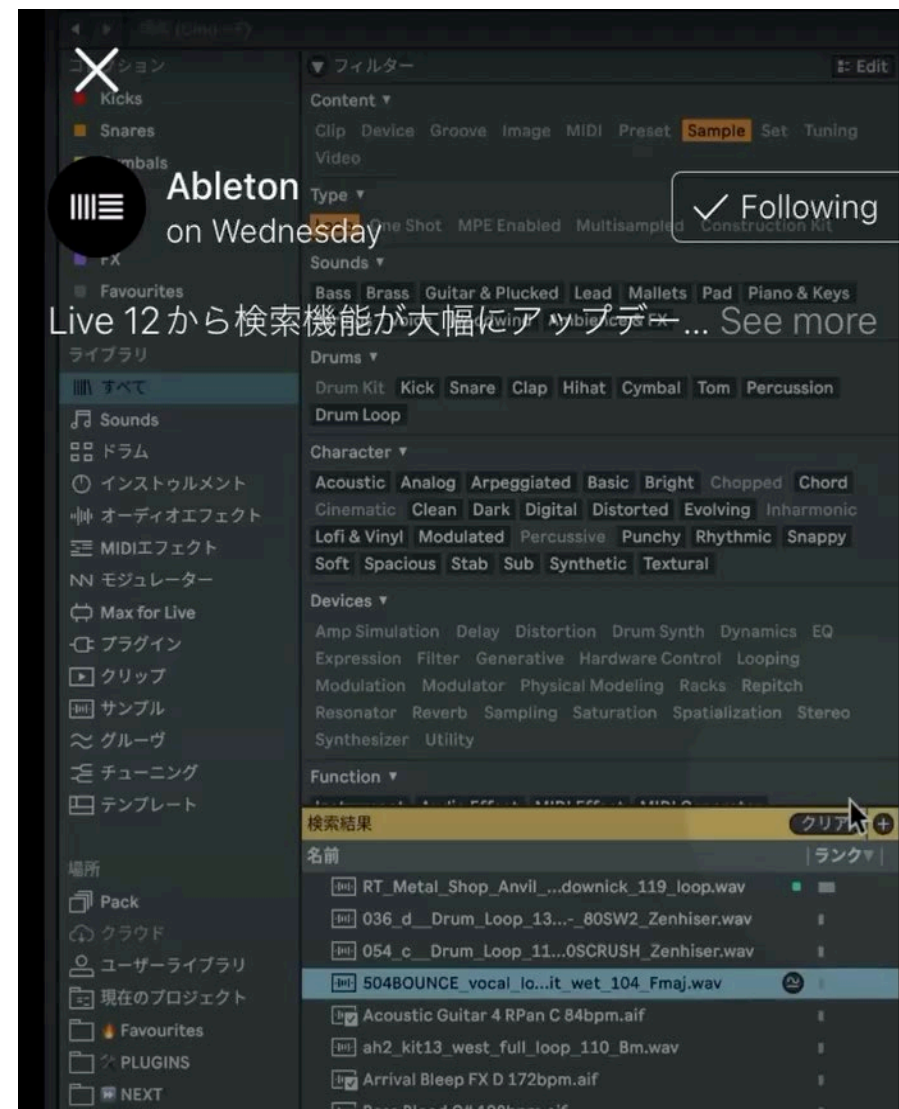


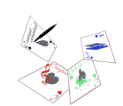


# 時間的メディアの情報に持つ印象の 解釈支援をするための問題整理

- 一過性の性質により，情報の相対的な評価が困難
  - 音色に対する印象を身に付けることは困難
    - 音楽初心者に向けた支援の仕組みが必要

どのように音色の印象の  
解釈支援をしたら良いか？





# 提案アプローチ

- 吹き出し状の図形を用いて音色を形象化
  - 音色の印象を視覚的に解釈
  - 時間的な情報の一覧化が可能に！

## 時間的メディア

一過性の性質により**解釈**が困難

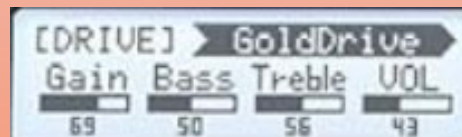
### 音色情報

- 音データ



### エフェクター

- パラメータ情報



形象化

## 空間的メディア

### 吹き出し図形

**心情や発話の印象**を  
自由に表現可能

形状から印象を**直感的**に想起



# 発想の根源

- ブーバ・キキ効果
  - 丸みの帯びた図形を“ブーバ”  
角張った図形を“キキ”と関連づける**傾向**
  - 音声の音象徴と図形の**対応問題**
- **音象徴**：音に対する認知・印象
  - 多くの人が共通理解を持つ



# 音の形象化は音色にも有効か？

- 漫画は様々な表現技法を用いてキャラクターの心情や状況を伝達
  - 吹き出し
    - キャラクターのセリフに含む心情や印象を**形状で表現**可能
      - 発話の音響情報に含まれる**意図**や**印象**を表現するために利用

こんにちは



こんにちは



音声は形象化できるけど  
音色はできるのか？

# 実験データ：音色

Comunitaらが作成し公開するエレキギターエフェクターを使用した音源のデータセット[1]からエレキギターのモノフォニックな音源を利用

- 役割系統を基準に4種の音色を対象
- 各種エフェクターの音源内から，Zero Crossing Rateが中央値となる音源を採用

各種エフェクターの  
典型的な音を用意

| 種類        | 役割系統     | Midi番号 |
|-----------|----------|--------|
| OverDrive | 歪み       | 62(D)  |
| Chorus    | モジュレーション | 65(F)  |
| Delay     | 空間       | 71(B)  |
| NoFX      | —        | 51(D#) |

[1]Comunit`a, M., Stowell, D. and Reiss, J. D.: Guitar Effects Recognition and Parameter Estimation with Convolutional Neural Networks, The Journal of the Audio Engineering Society, Vol. 69, No. 7, pp. 594–604 (2020).

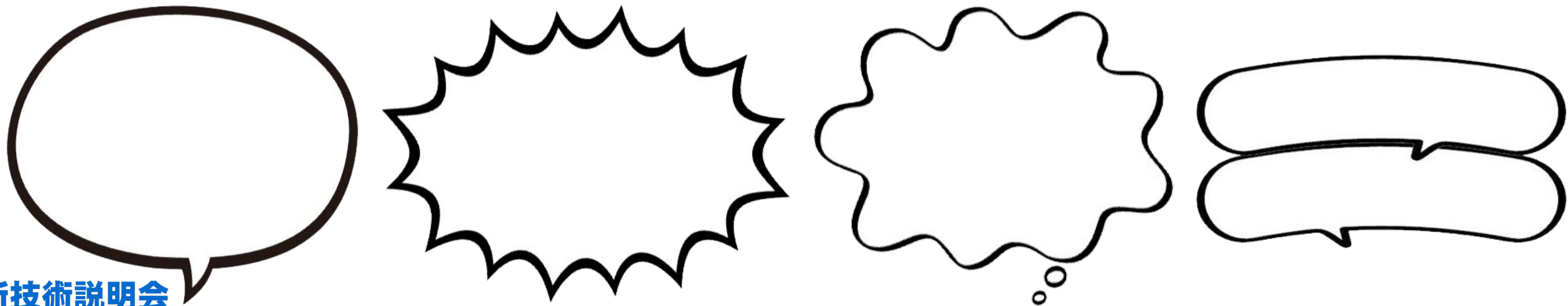


# 実験データ：吹き出し図形

WEBサイトの「フキダシデザイン」から図形を選定

各シチュエーションごとに使用される代表的な吹き出しを選択

- 日常的な会話で用いられる形状
- 叫びや怒りなどの感情の昂りを表現する形状
- キャラクタの思考を表現する形状
- 連続するセリフを表現する形状



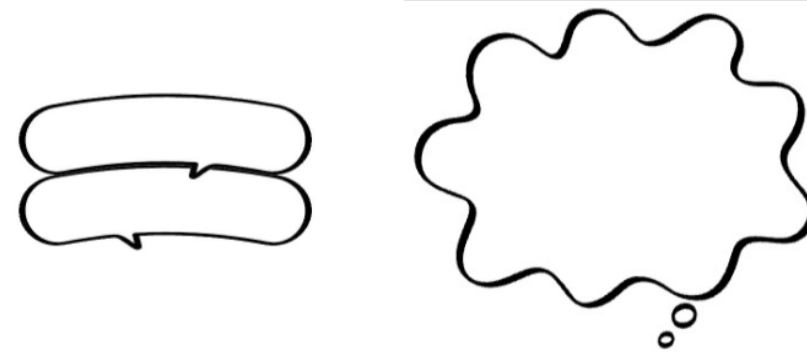
# 実験条件

- 2種類の図形と2種類のギター音源を  
呈示→**視覚**情報と**聴覚**情報を紐付け

## 実験手順

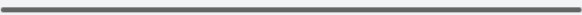


1. 音源の聴取
  2. 音源の印象と合致した図形を  
それぞれ回答
  3. アンケートへの回答
- 120名が参加し，40名ずつ3群に分割
    - 各々2問ずつ回答
      - 音楽経験あり：20名
      - 音楽経験なし：20名

Q.それぞれの吹き出しのイメージに  
合う音源を選んで回答してください

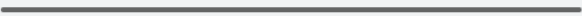



選択肢

音源A

▶ 0:00 / 0:02   

音源B

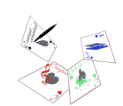
▶ 0:00 / 0:02   

Q1.左の図形と対応する音源を選択してください

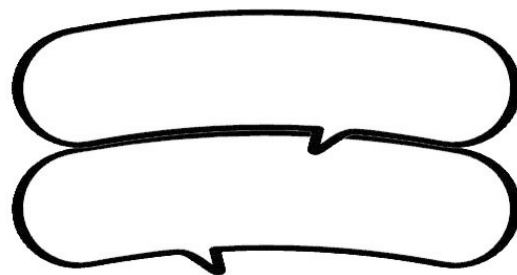
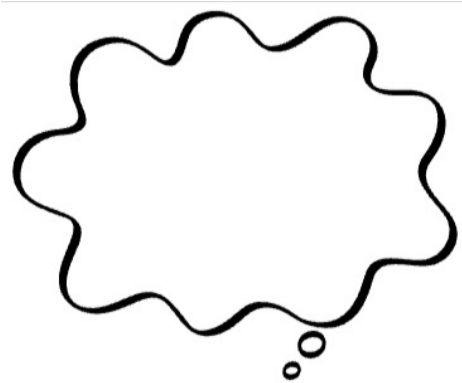
☒ A  
☐ B

Q2.右の図形と対応する音源を選択してください

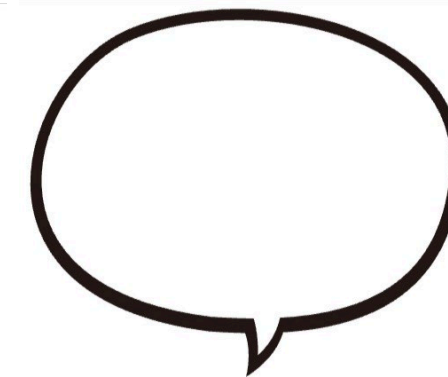
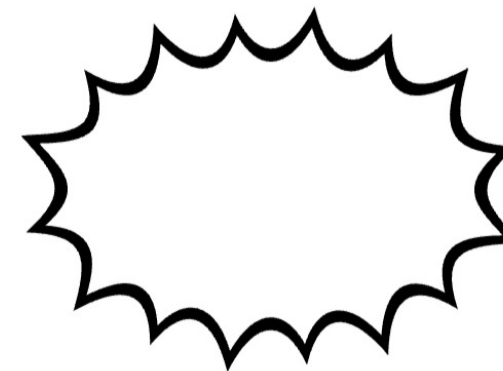
☐ A  
☒ B



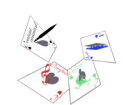
# 実験結果 1 / 3



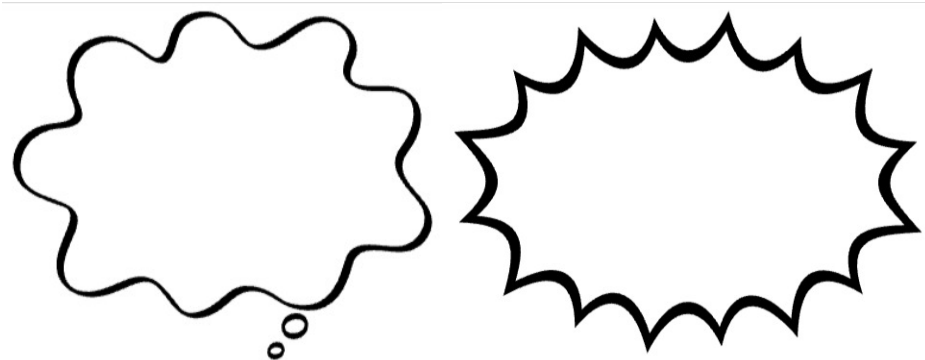
|                       | 全体    | 音楽経験者 | 非音楽経験者 |
|-----------------------|-------|-------|--------|
| Chorus:左図<br>Delay:右図 | 37.5% | 35.0% | 40.0%  |
| Chorus:右図<br>Delay:左図 | 47.5% | 45.0% | 50.0%  |
| その他                   | 15.0% | 20.0% | 0.0%   |



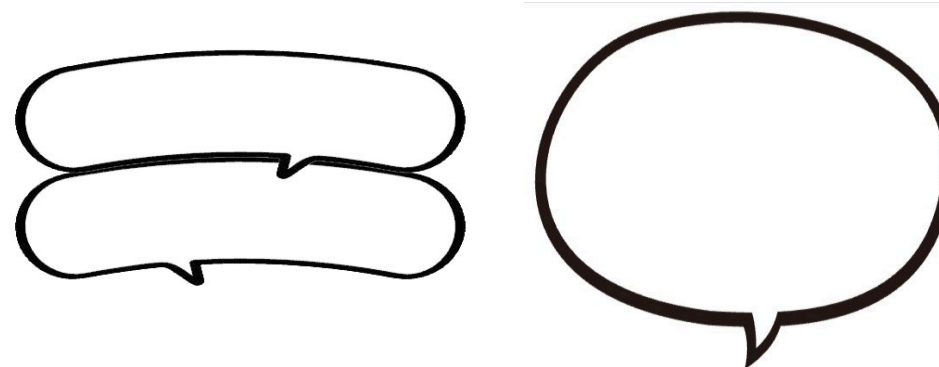
|                         | 全体    | 音楽経験者  | 非音楽経験者 |
|-------------------------|-------|--------|--------|
| OverDrive:左図<br>NoFX:右図 | 97.5% | 100.0% | 95.0%  |
| OverDrive:右図<br>NoFX:左図 | 2.5%  | 0.0%   | 5.0%   |
| その他                     | 0.0%  | 0.0%   | 0.0%   |



# 実験結果2/3

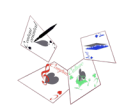


|                           | 全体    | 音楽経験者 | 非音楽経験者 |
|---------------------------|-------|-------|--------|
| Chorus:左図<br>OverDrive:右図 | 82.5% | 85.0% | 80.0%  |
| Chorus:右図<br>OverDrive:左図 | 10.0% | 15.0% | 5.0%   |
| その他                       | 7.5%  | 0.0%  | 15.0%  |

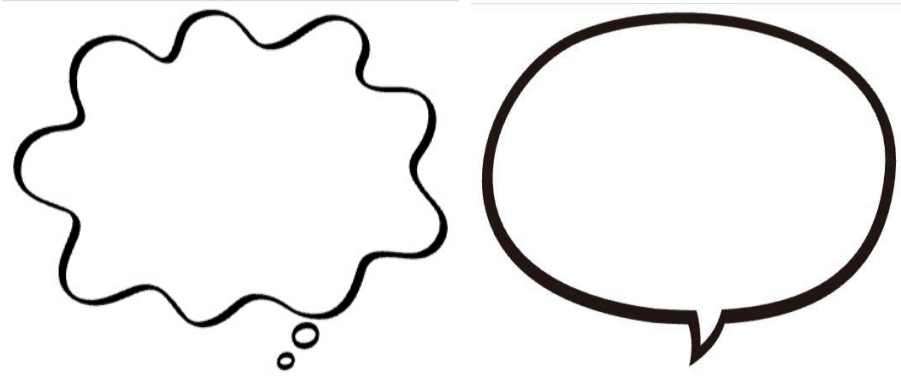


|                     | 全体    | 音楽経験者 | 非音楽経験者 |
|---------------------|-------|-------|--------|
| Delay:左図<br>NoFX:右図 | 70.0% | 65.0% | 75.0%  |
| Delay:右図<br>NoFX:左図 | 17.5% | 20.0% | 15.0%  |
| その他                 | 12.5% | 15.0% | 10.0%  |





# 実験結果3/3



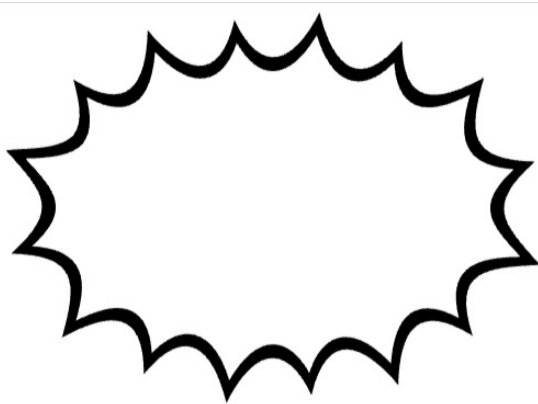
|                      | 全体    | 音楽経験者 | 非音楽経験者 |
|----------------------|-------|-------|--------|
| Chorus:左図<br>NoFX:右図 | 42.5% | 50.0% | 35.0%  |
| Chorus:右図<br>NoFX:左図 | 47.5% | 40.0% | 55.0%  |
| その他                  | 10.0% | 10.0% | 10.0%  |



|                          | 全体    | 音楽経験者 | 非音楽経験者 |
|--------------------------|-------|-------|--------|
| Delay:左図<br>OverDrive:右図 | 12.5% | 15.0% | 10.0%  |
| Delay:右図<br>OverDrive:左図 | 80.0% | 80.0% | 80.0%  |
| その他                      | 7.5%  | 5.0%  | 10.0%  |

# OverDriveはわかりやすい

- OverDriveが呈示された問題で80%を越える偏りを確認
  - NoFXとの比較で，ほぼ全ての被験者の回答が一致
  - OverDriveのような「歪み」の音色の印象と棘状の図形の関連性が高い



NoFXとの比較

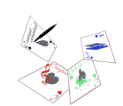
|                         | 全体    | 音楽経験者  | 非音楽経験者 |
|-------------------------|-------|--------|--------|
| OverDrive:左図<br>NoFX:右図 | 97.5% | 100.0% | 95.0%  |
| OverDrive:右図<br>NoFX:左図 | 2.5%  | 0.0%   | 5.0%   |
| その他                     | 0.0%  | 0.0%   | 0.0%   |

Chorusとの比較

|                           | 全体    | 音楽経験者 | 非音楽経験者 |
|---------------------------|-------|-------|--------|
| Chorus:左図<br>OverDrive:右図 | 82.5% | 85.0% | 80.0%  |
| Chorus:右図<br>OverDrive:左図 | 10.0% | 15.0% | 5.0%   |
| その他                       | 7.5%  | 0.0%  | 15.0%  |

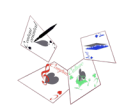
Delayとの比較

|                          | 全体    | 音楽経験者 | 非音楽経験者 |
|--------------------------|-------|-------|--------|
| Delay:左図<br>OverDrive:右図 | 12.5% | 15.0% | 10.0%  |
| Delay:右図<br>OverDrive:左図 | 80.0% | 80.0% | 80.0%  |
| その他                      | 7.5%  | 5.0%  | 10.0%  |



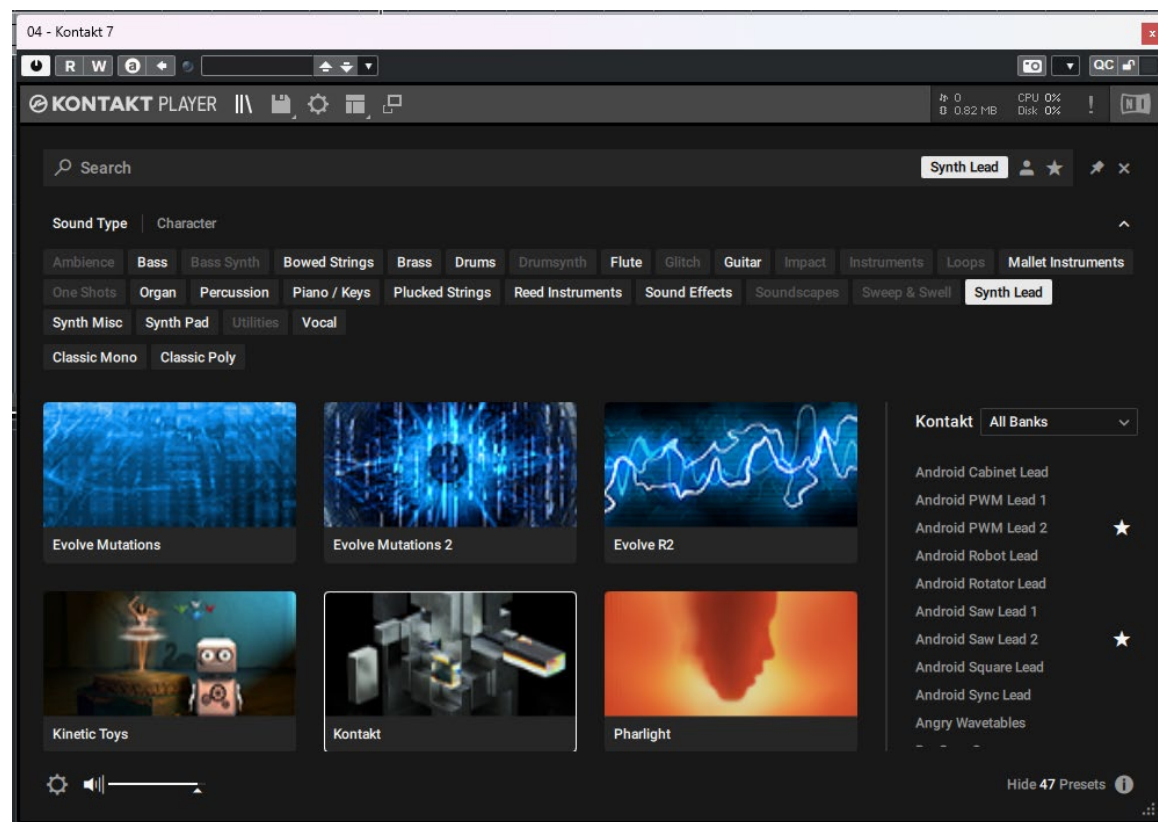
## 2. 音の類似性を ネットワークにして可視化



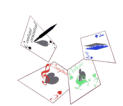


# 音色選択の問題（再確認）

一般的に、音色は**音源名のみ**で表示



- 問題①：音色同士の相対的な評価が困難・  
言語でアクセスしづらい
- 問題②：試聴した音色がどんな音色だったか  
試聴しているうちに忘れてしまう
- 問題③：全ての音色を網羅して試聴すると  
手間がかかる



# デスクトップミュージックとは

Desktop Music(DTM)

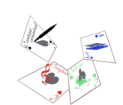


パソコンを使用する音楽制作手法の総称  
DTMに音色選択は不可欠

音色

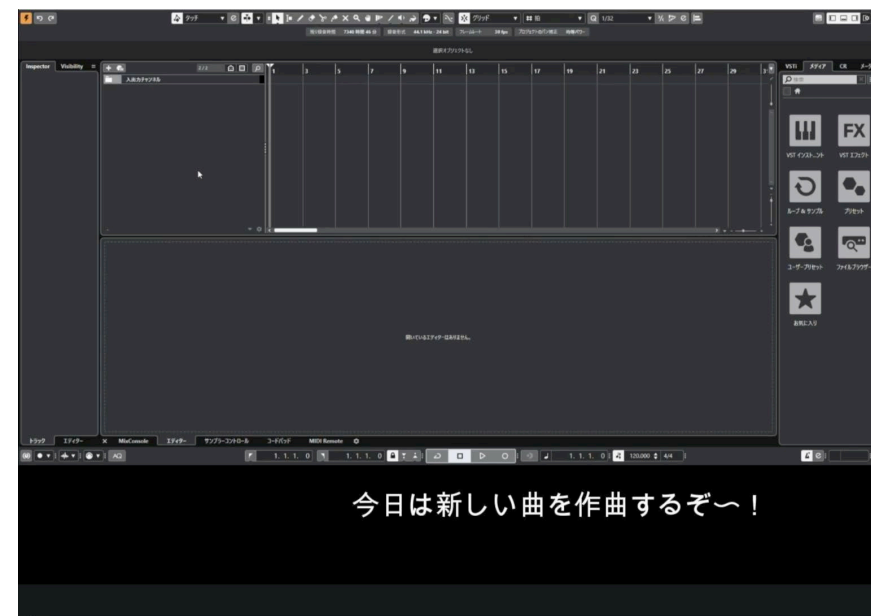


プリセットとして格納されている波形や  
エフェクトによって構築される音の特色

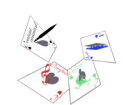


# 問題提起と解決策：Syntherium構想

- 問題①：音色同士の相対的な評価が困難・言語化しづらい
  - 解決：音色同士の関係性を可視化する
- 問題②：どんな音色だったか試聴しているうちに忘れてしまう
  - 解決：お気に入りの音色をお気に入り度を反映させてブックマークできる
- 問題③：全ての音色を網羅して試聴すると手間がかかる
  - 解決：複数の音色を一度に比較できるシステム



今日は新しい曲を作曲するぞ〜！



# Syntherium構想の要素

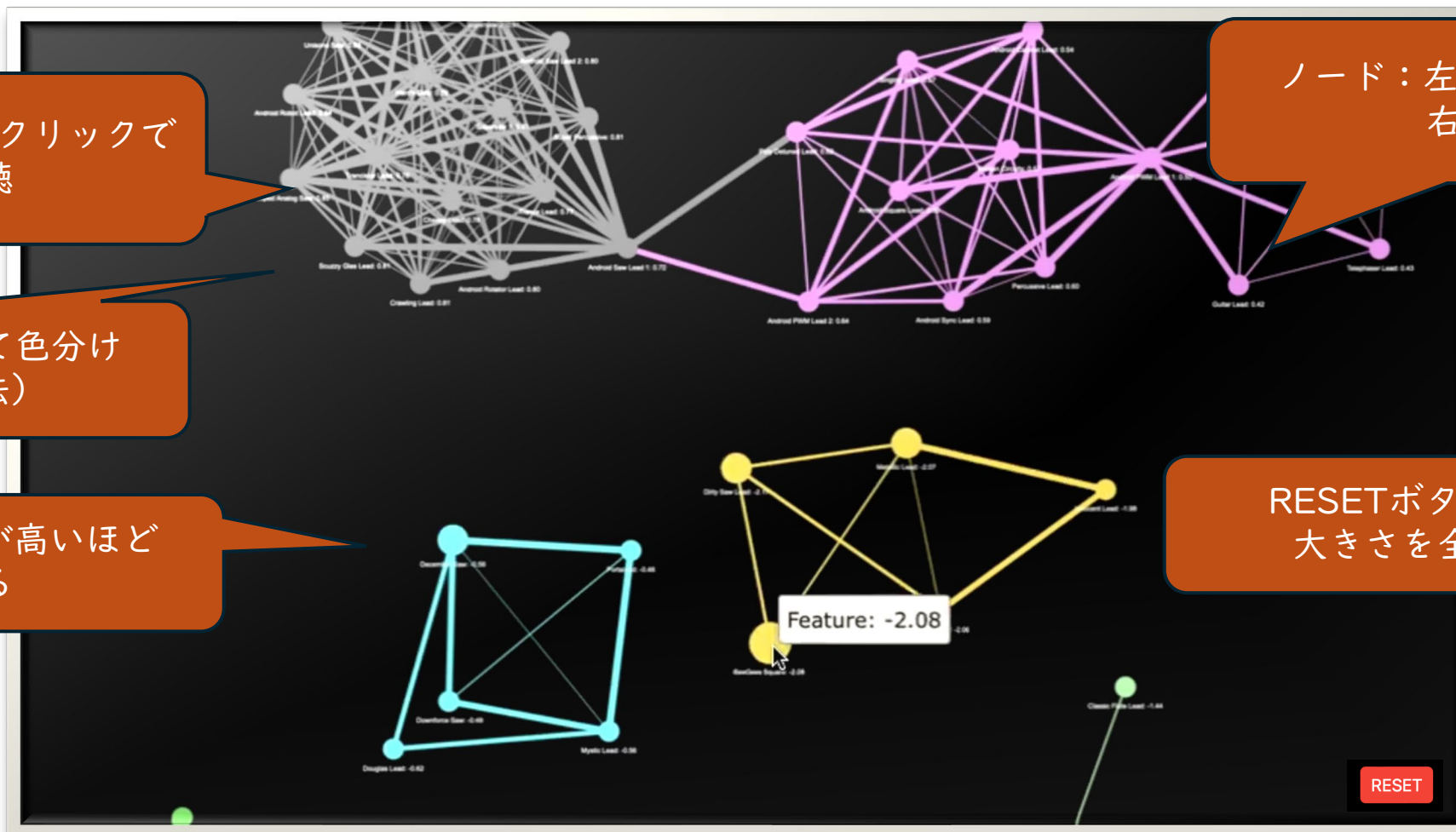
ノードを左ダブルクリックで  
音源視聴

クラスタによって色分け  
(Louvain法)

エッジ：類似度が高いほど  
太くなる

ノード：左クリックで拡大  
右クリックで縮小

RESETボタンでノードの  
大きさを全てリセット



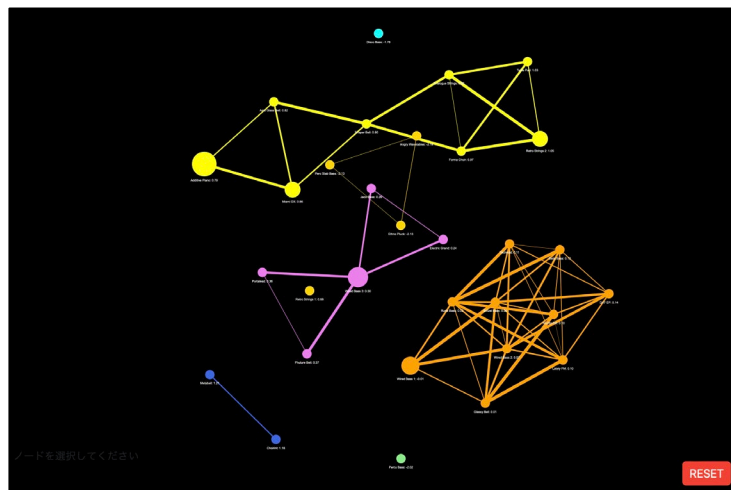


# Syntheriumの評価

## 実験群

### Syntherium使用

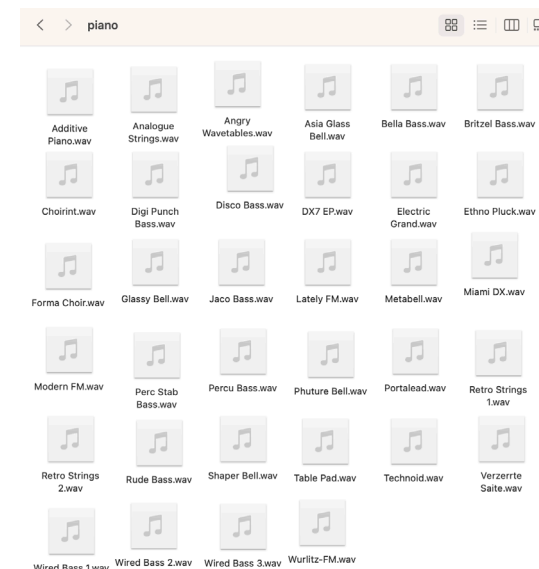
テストユーザ6人  
(音楽経験者3人・音楽未経験者3人)

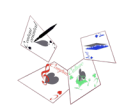


## コントロール群

### Syntherium不使用

テストユーザ6人  
(音楽経験者3人・音楽未経験者3人)

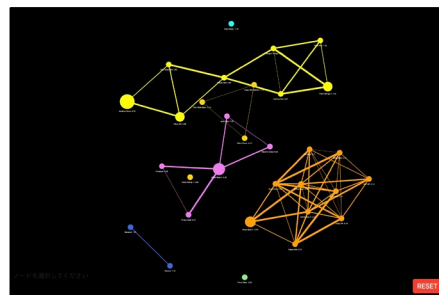




# 実験手順

## 実験群

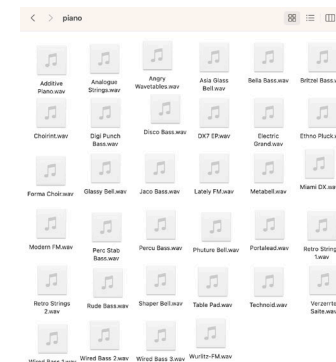
syntherium使用



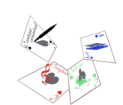
同じ音色を探す

## コントロール群

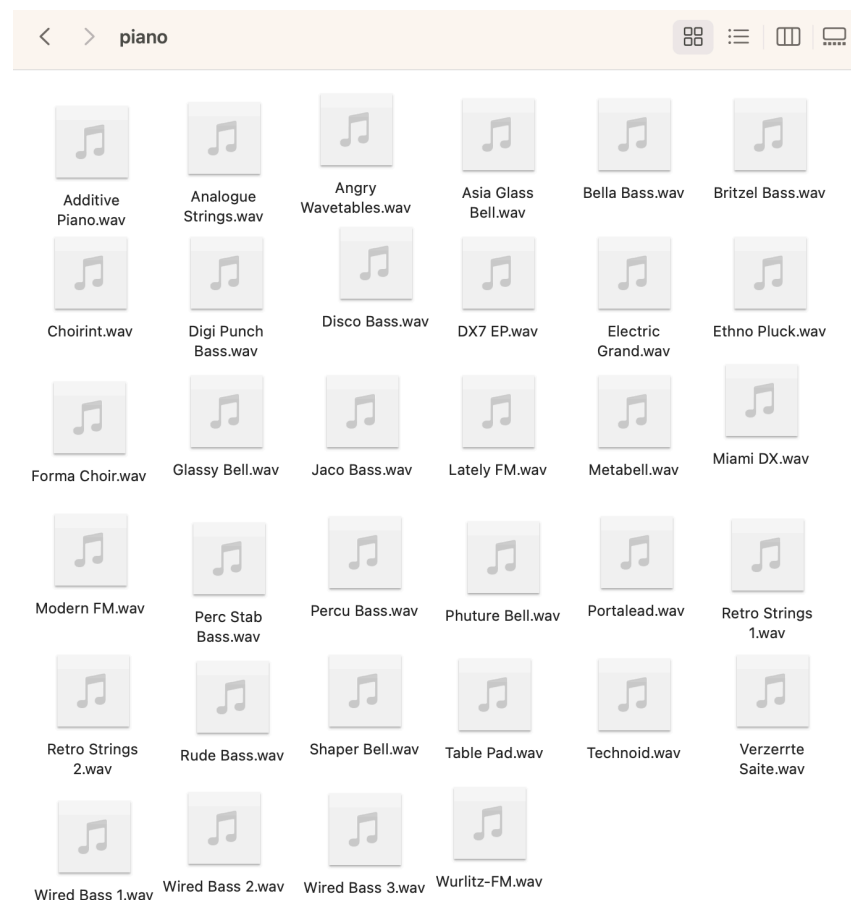
Syntherium不使用



同じ音色を探す



# 探索後のネットワーク：コントロール



## コントロール群 (Syntherium無し)の

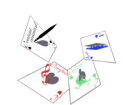
### テストユーザ

#### 結果

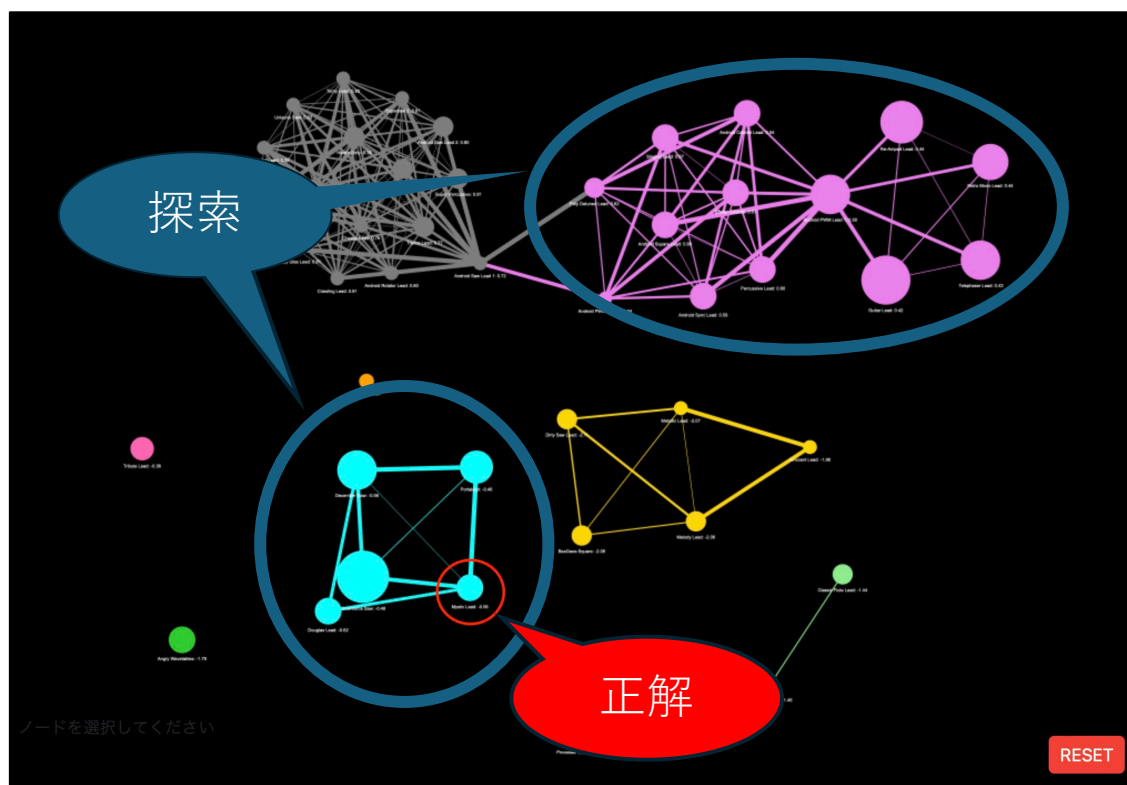
全員が上から順番（アルファベット順）に  
試聴した

#### 考察

探索において、探索順序に意味を持たせな  
かった



# 探索後のネットワーク：実験群



## 実験群(Syntherium有り)のテストユーザ

ユーザ像：楽器経験有り・音色選択経験無し

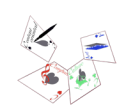
### 結果

正解の音色の周りを探索し、  
正解の音色に早い段階でたどり着いている

### 考察

システムが有効に使用されている





# 試行時間と聴取時間

## Syntherium使用の有無による試行時間と試聴回数の比較

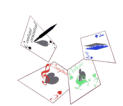
|         | Syntherium使用<br>(平均) | Syntherium非使用<br>(平均) |
|---------|----------------------|-----------------------|
| 試行時間(秒) | 117.375              | 88.292                |
| 試聴回数(回) | 27.417               | 47.500                |

### 結果

Syntheriumを使用している方が試聴回数は減少しているが、探索時間は増加した

### 考察

Syntheriumを使用することによって、探索中に思考が発生していることが推測できる



# 試行時間と聴取時間

## Syntherium使用の有無による試行時間と試聴回数の比較

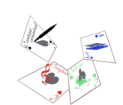
|         | Syntherium使用<br>(平均) | Syntherium非使用<br>(平均) |
|---------|----------------------|-----------------------|
| 試行時間(秒) | 117.375              | 88.292                |
| 試聴回数(回) | 27.417               | 47.500                |

### 結果

Syntheriumを使用している方が試聴回数は減少しているが、探索時間は増加した

### 考察

Syntheriumを使用することによって、探索中に思考が発生していることが推測できる

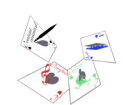


# 利用中の発言

Syntherium使用の有無と楽器経験の有無によるユーザテスト中の発話の抜粋

|        | Syntherium使用者  | Syntherium非使用者  |
|--------|--|---|
| 楽器経験者  | <ul style="list-style-type: none"><li>• (ネットワーク上で隣の音色に対して)<b>これ似てる?</b></li><li>• この(クラスタの)中にありそうだな</li></ul>                        | <ul style="list-style-type: none"><li>• どれがさっきの音色だったけ?</li></ul>                            |
| 楽器未経験者 | <ul style="list-style-type: none"><li>• <b>ぐにゃぐにゃした音</b>だな</li><li>• これは<b>にょーん</b>とした音</li><li>• (ノードが密集している箇所に) 手を出したくない</li></ul> | <ul style="list-style-type: none"><li>• さっきの音色どれだったけ?</li><li>• さっきの音色となにが違うんだろう?</li></ul> |

Syntheriumの使用によって、楽器経験問わず  
**音色の特徴や類似度を推測**する発話が見られた



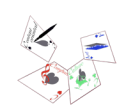
# 利用中の発言

Syntherium使用の有無と楽器経験の有無によるユーザテスト中の発話の抜粋

|        | Syntherium使用者   | Syntherium非使用者  |
|--------|---|---|
| 楽器経験者  | <ul style="list-style-type: none"><li>• (ネットワーク上で隣の音色に対して)これ似てる?</li><li>• この(クラスタの)中にありそうだな</li></ul>                | <ul style="list-style-type: none"><li>• <b>どれがさっきの音色だっけ?</b></li></ul>                            |
| 楽器未経験者 | <ul style="list-style-type: none"><li>• ぐにゃぐにゃした音だな</li><li>• これはにょーんとした音</li><li>• (ノードが密集している箇所に)手を出したくない</li></ul> | <ul style="list-style-type: none"><li>• <b>さっきの音色どれだっけ?</b></li><li>• さっきの音色となにが違うんだろう?</li></ul> |

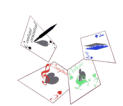
Syntherium非使用者には  
**興味を持った音色を忘れてしまう**旨の発話が多く見られた





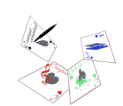
# 新技術の特徴・従来技術との比較

- 記憶が必要な音メディアを見れば分かる形で保存可能に
- 同時比較が難しかった音メディアの複数比較を可能に
- テキストでの表現しかできなかった音色を形状によって直感的に検索可能に
- 音色同士の類似性は記憶や数値のパラメータであったものを視覚的に表現



# 音が視えたら何が変わる？

音のパラメータ  $\leftrightarrow$  形状のパラメータが学習可能に

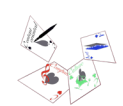


# 実用化にむけた課題

- ・ 音-図形の対応データの収集, 蓄積
- ・ 蓄積されたデータを元にした音-図形モデルの構築

# 社会実装への道筋

- ・ 現在 : 基礎理論の確立・プロトタイプ作成
- ・ ~2030 : モデル構築 → 産業応用・



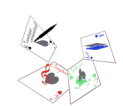
# 想定される用途と企業への貢献

## 産業的な応用の場面で

- 音色のデザインの検討段階に
- 騒音や通知音との調和のデザインに見ながら調整する音響デザイン
- 異常音を目で見てチェック

## クリエイティブな現場で

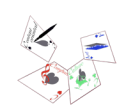
- プロフェッショナルの完成を形状として記録
- 感性の教育ツールとして
- 言語化が難しいことを形状にしてコミュニケーションを円滑に



# 本技術に関する知的財産権

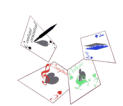
- ・ 発明の名称 : 音を含むコンテンツの制作システム、  
コンピュータ実装方法、及びコンピュータプログラム
- ・ 出願番号 : 特願2024-118433
- ・ 出願人 : 関西大学
- ・ 発明者 : 山西良典、永田大志
  
- ・ 発明の名称 : システム、方法、及び、コンピュータプログラム
- ・ 出願番号 : 特願2025-72652
- ・ 出願人 : 関西大学
- ・ 発明者 : 山西良典、米田美優





# 産学連携の経歴

- 2023年～現在 : プラットフォーム向けのゲーム関連コンテンツを提供する企業と共同研究
- 2024年～現在 : 日本のアニメ制作会社と共同研究
- 2019年 : デジタルマーケティングソリューション企業に技術指導
- 2018年～2020年 : アミューズメントマシンを開発する企業と共同研究



# お問い合わせ先

関西大学社会連携部  
産学官連携センター・知財センター

TEL 06-6368-1044

e-mail [sangakukan-mm@ml.kandai.jp](mailto:sangakukan-mm@ml.kandai.jp)

# 音情報編集のための 音色を可視化するインタフェース



**ENTERTAIN** the world

Live to **PLAY**  
**ENJOY** to live



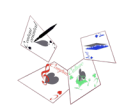
# 実験結果：アンケート結果 音色以外を図形と紐つけた意見

## 音楽経験者の感想

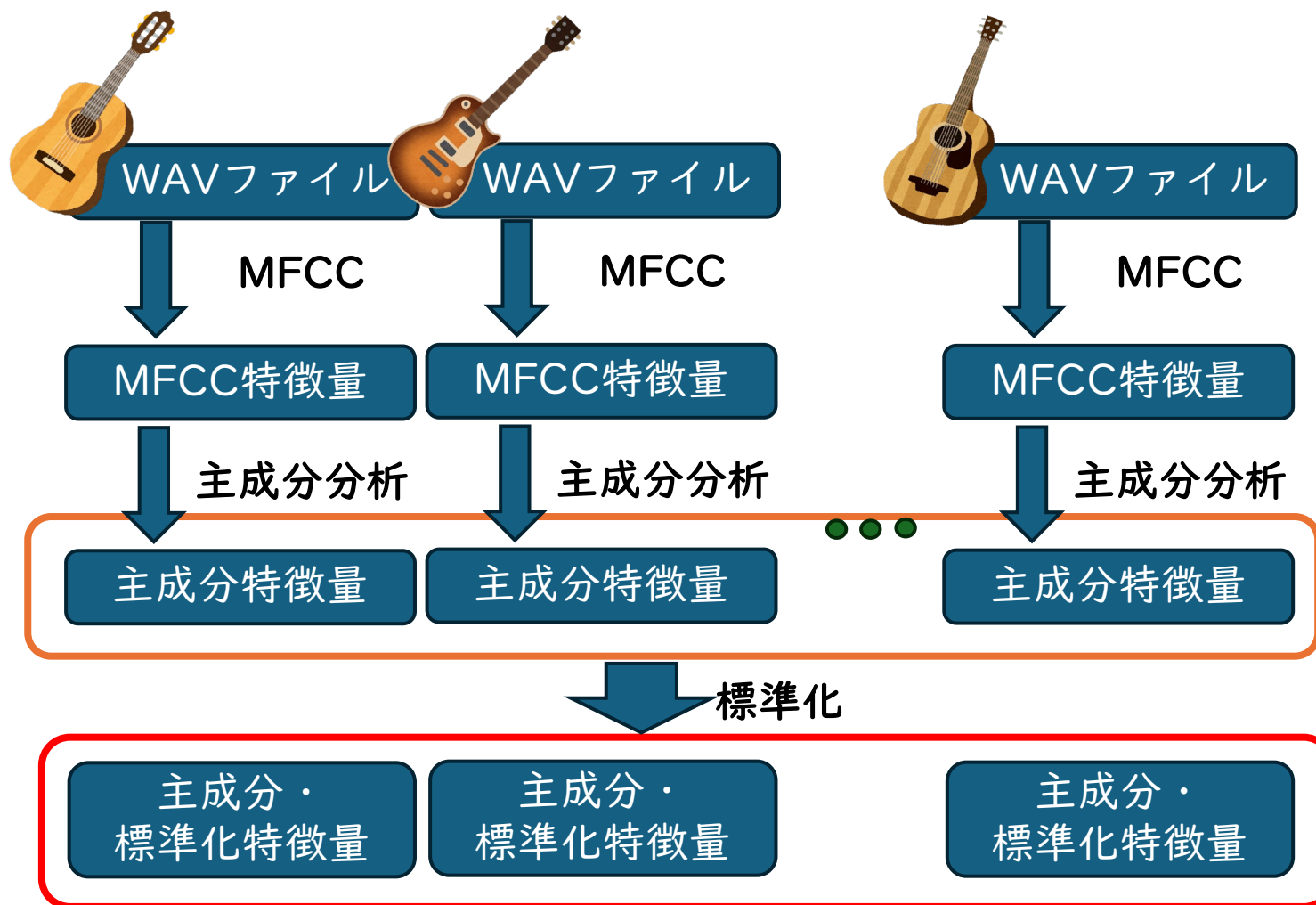
- **大きい音**や**高い音**は、吹き出しが大きい、トゲトゲしている、線が濃い等のイメージがある。また、弱い感じの音は、吹き出しがふにゃふにゃしている、丸みを帯びている、イメージがある。自分が図形を見て考えていた音のイメージではないものが多く、選択に時間がかかった。吹き出しと音源が自分の中で結びつきづらく、「強いて言えば…」といった感覚で選んだ。

## 非音楽経験者の感想

- 微妙な違いを図形に反映させる過程で、**音の高さ**や**音の強さ**を意識して選択した。**音が低い**とふにゃふにゃしたイメージがありました。
- もっと当てはまりの良い音源がありそうだなと感じました。



# 特徴量の取得



- wavファイルから128次元のMFCC（メル周波数ケプストラム係数）特徴量を算出
- 周波数の概観をまとめるために主成分分析を実施
- 得られた主成分特徴量の値域を揃えるために標準化

ここで比較！