

視線計測を利用して音楽から想像された 色の推定

SHAN Junjie, 西澤太次郎, 西原陽子

立命館グローバル・イノベーション研究機構
立命館大学情報理工学部

2024-03-05

目次

- 背景
- 問題点
- 解決した内容
- 提案システム
 - ランダム色パネルの生成
 - 視線計測で注視色の記録
 - 記録色の可視化
- 実験
- 結果と考察
- おわりに

背景

- 色彩と音楽は、生まれつきな視覚と聴覚能力として、心理学や応用的な領域には注目が集まる
- 音楽に合わせて、異なる色彩や照明で雰囲気作り
 - 音楽はジャンルにより人間に与える印象が異なる
 - その印象から想像される色も異なる
- 音楽と色の関連性を判明するのは、マルチメディア時代の応用にとって、重要な一環である

問題点

- 「音楽理論」や「色彩構成」などの専門知識が必要
- 「絶対音感」や「信号処理」などの複雑な技術
→ 結果は直観しない、比較しにくい

K. Itoh, et al. (2017), Musical pitch classes have rainbow hues in pitch class-color synesthesia.

岩井大輔, et al. (2002), 音と色のノンバーバルマッピング-色聴保持者のマッピング抽出とその応用-

解決した内容

1. 視線計測装置でユーザが注視した色の抽出

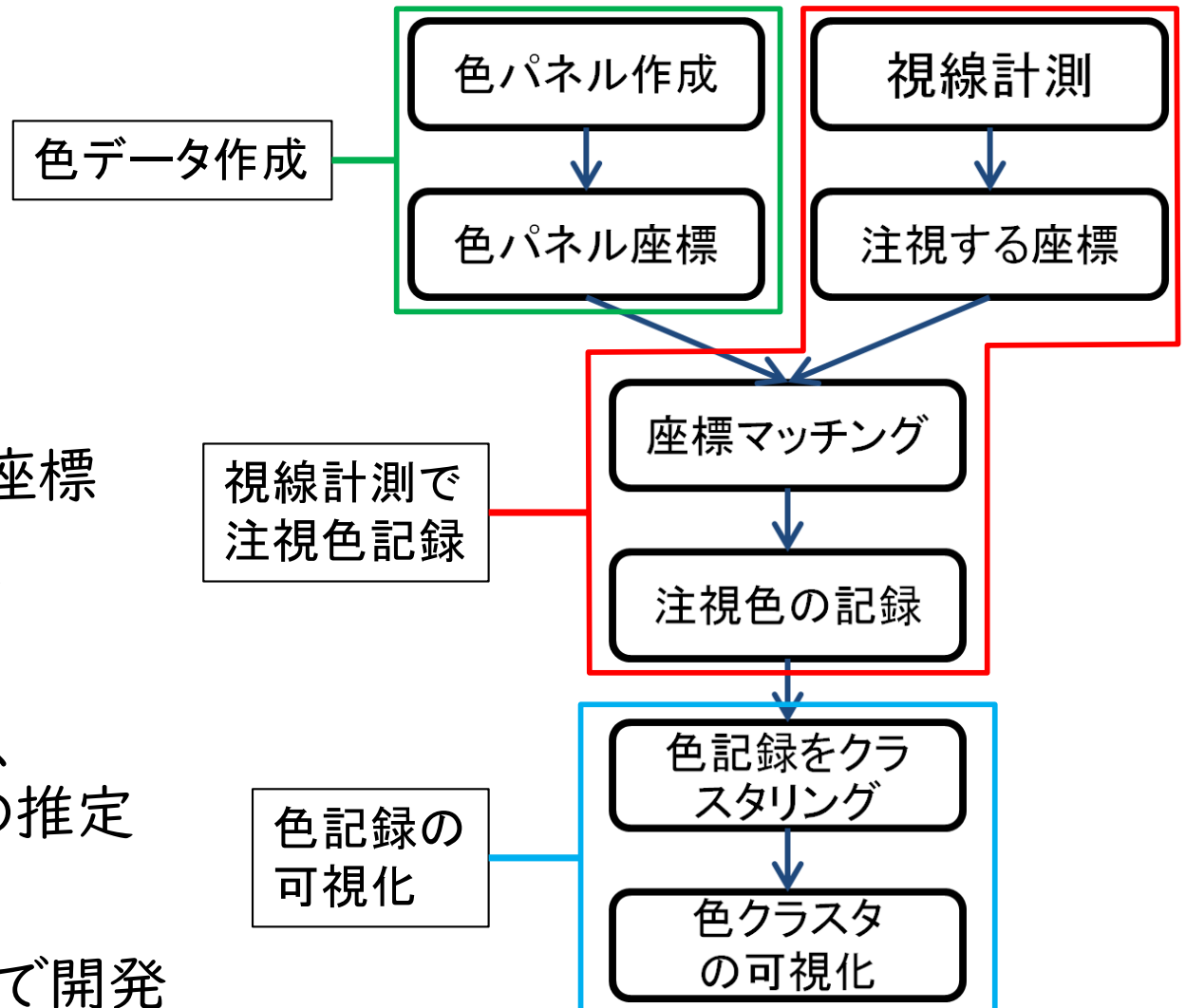
視線計測装置を用いてユーザに見られた色を記録するシステムの開発、及び記録した色をクラスタリングにより可視化するインタフェースの実装

2. 音楽ジャンルから想像された色の測定

視線計測システムを利用して、音楽理論や信号処理などの専門知識を触れず、より直観的に音楽の感覚と色のマッチングを測定した

提案システム

- ランダムに生成された色パネル:
 - 色のRGB値
 - 各色ブロックの座標
- 視線計測で色記録:
 - 視線と色ブロックの座標
 - 色ブロックのRGB値
- 色記録の可視化:
 - クラスタリングにより、
注視された色種類の推定
- TkinterとMatplotlibで開発



ランダム色パネルの作成

- 4×4、合計16色のパネル
- RGB値で16色をランダムに生成
 - 各2色のRGB値のマンハッタン距離に閾値(Thre)を設置する

$$|R_i - R_j| + |G_i - G_j| + |B_i - B_j| \geq Thre;$$

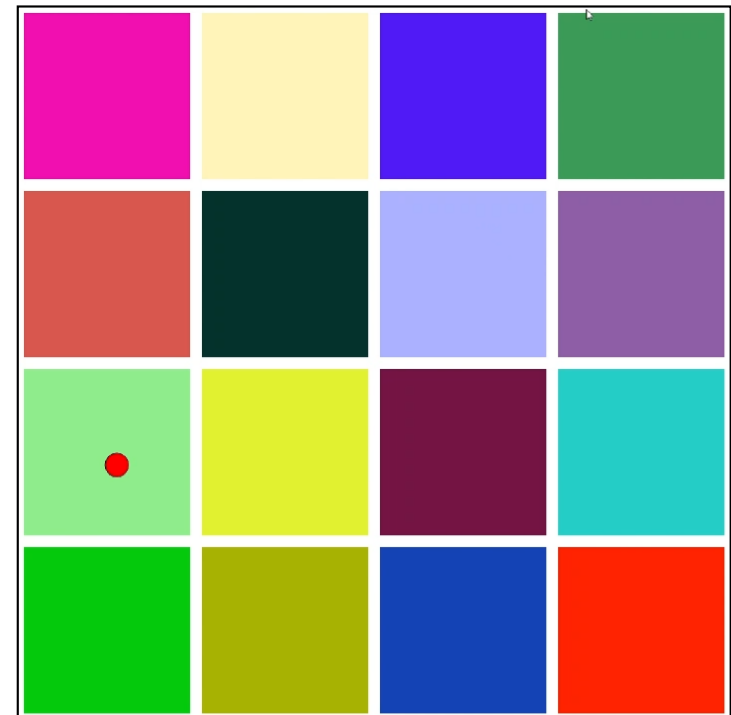
$(i, j \in \text{Generated 16 Colors})$

- 色パネルは2.5秒の間隔で更新する
 - 0.5秒の反応時間 + 2秒の計測時間



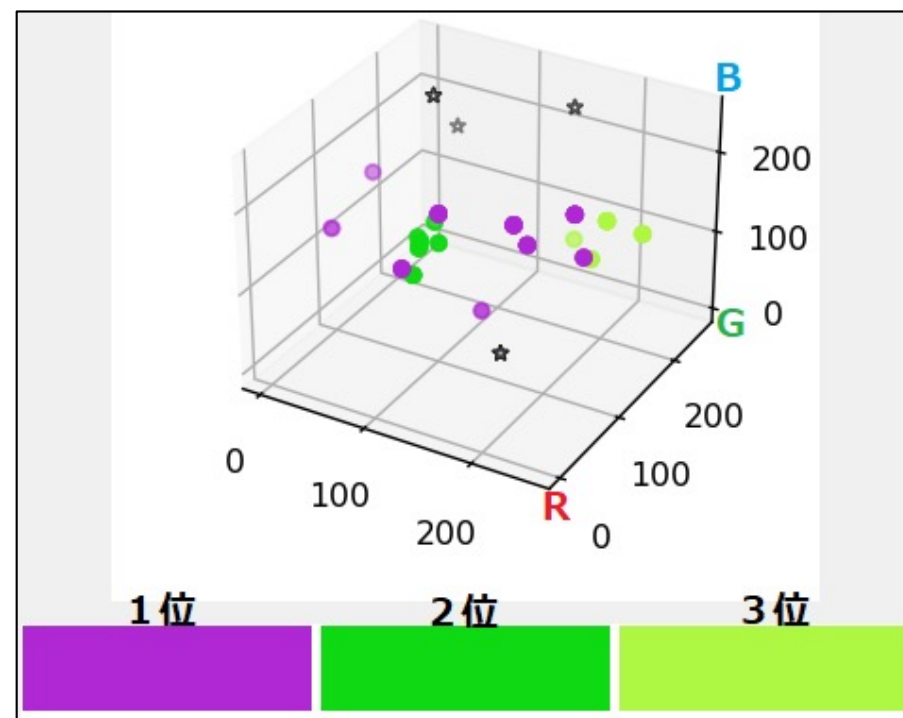
視線計測で注視色の記録

- 使われた視線計測装置は「Tobii Pro Nano」
モニタの下部に設置し、モニタに向いているユーザの左目と右目の
スクリーン内での座標をそれぞれ計測できる
- 計測されたユーザの左目座標の右目座標の「中点」を
ユーザの「注視点座標」として算出する
- 注視点座標がある色ブロックの
範囲にある時、そのブロックの
その時点での色のRGB値を
記録する
 - 視線の移動閾値は50ピクセル
(色探するための視線移動途中の
色ブロックを記録しないように)



色記録から注視色種類の推定と可視化

- 色記録リストに対して、クラスタリングによりユーザに注視された色の種類を推定する
 - ノイズ検出のため、DBSCAN手法を使った
 - 分けられた「クラスタの数」は、注視された「色の種類数」、各クラスタにある「記録数」は、その種類の色を注視した「頻度」と見なしている
 - 各クラスタに含まれる色記録のRGB値の平均を、その種類の代表色として推定する
 - 各クラスタの記録数の順番により、上位三つのクラスタの色記録のRGB値の3D空間分布と平均色(代表色)を可視化する



音楽から想像された色の推定実験

• 実験手順

1. ジャンル別の音楽を用意する
2. 被験者に用意した音楽を聴かせる
3. 音楽を聴かせ始めてから5秒後に色パネルを表示し、被験者の視線を計測する
4. 被験者に色パネルの中から、聴いている音楽から想像される色を目で探させる
5. 提案システムにより、被験者が注視した色を記録する
6. 聴かせる音楽を変えて、手順2～手順5を繰り返す



音楽ジャンルの想像色を推定する実験

- 実験データ

- 「ロック」、「バラード」、「ポップス」のジャンルで
男声ボーカルと女声ボーカルの合計6つのジャンル
- 各ジャンルに3曲ずつ, 合計18曲
- 言語の影響を抑えるため、全部英語の曲

	ロック	バラード	ポップス
男声	KiLLiNG ME Remember Jagerbomb	8 Letters One Call Away Too Good At Goodbyes	Lucky Strike Runaway Baby Say Somethin'
女声	Feel A Thing Lonely Girl Despondency	Lose You To Love Me Thinking About You Rise Up	Black Magic No Excuses Into You

- 実験時間: 1曲30秒 (5秒の導入+25秒の視線計測)
- 被験者: 10名の日本人大学生

実験結果

- 25秒間で計測した色記録に対するクラスタリングの結果

被験者 ID	音楽ジャンルごとのクラスタ数					
	男声 ロック	女声 ロック	男声 バラード	女声 バラード	男声 ポップス	女声 ポップス
ID1	15	15	19	24	15	18
ID2	22	9	15	15	17	6
ID3	5	9	11	16	15	12
ID4	7	10	19	18	12	12
ID5	14	19	19	22	12	9
ID6	25	16	17	21	15	16
ID7	26	13	14	25	17	10
ID8	7	14	12	21	10	15
ID9	9	14	10	18	21	14
ID10	7	18	8	23	12	11
平均	13.7	13.7	14.4	20.3	14.6	12.3

被験者 ID	音楽ジャンルごとのノイズ数					
	男声 ロック	女声 ロック	男声 バラード	女声 バラード	男声 ポップス	女声 ポップス
ID1	60	50	44	68	48	39
ID2	41	46	63	50	64	46
ID3	28	16	27	33	37	12
ID4	18	32	37	39	17	21
ID5	39	49	51	40	35	42
ID6	18	34	39	30	27	40
ID7	81	49	57	74	77	45
ID8	13	24	42	36	21	32
ID9	58	22	25	58	52	29
ID10	31	21	28	42	29	17
平均	38.7	34.3	41.3	47.0	40.7	32.3

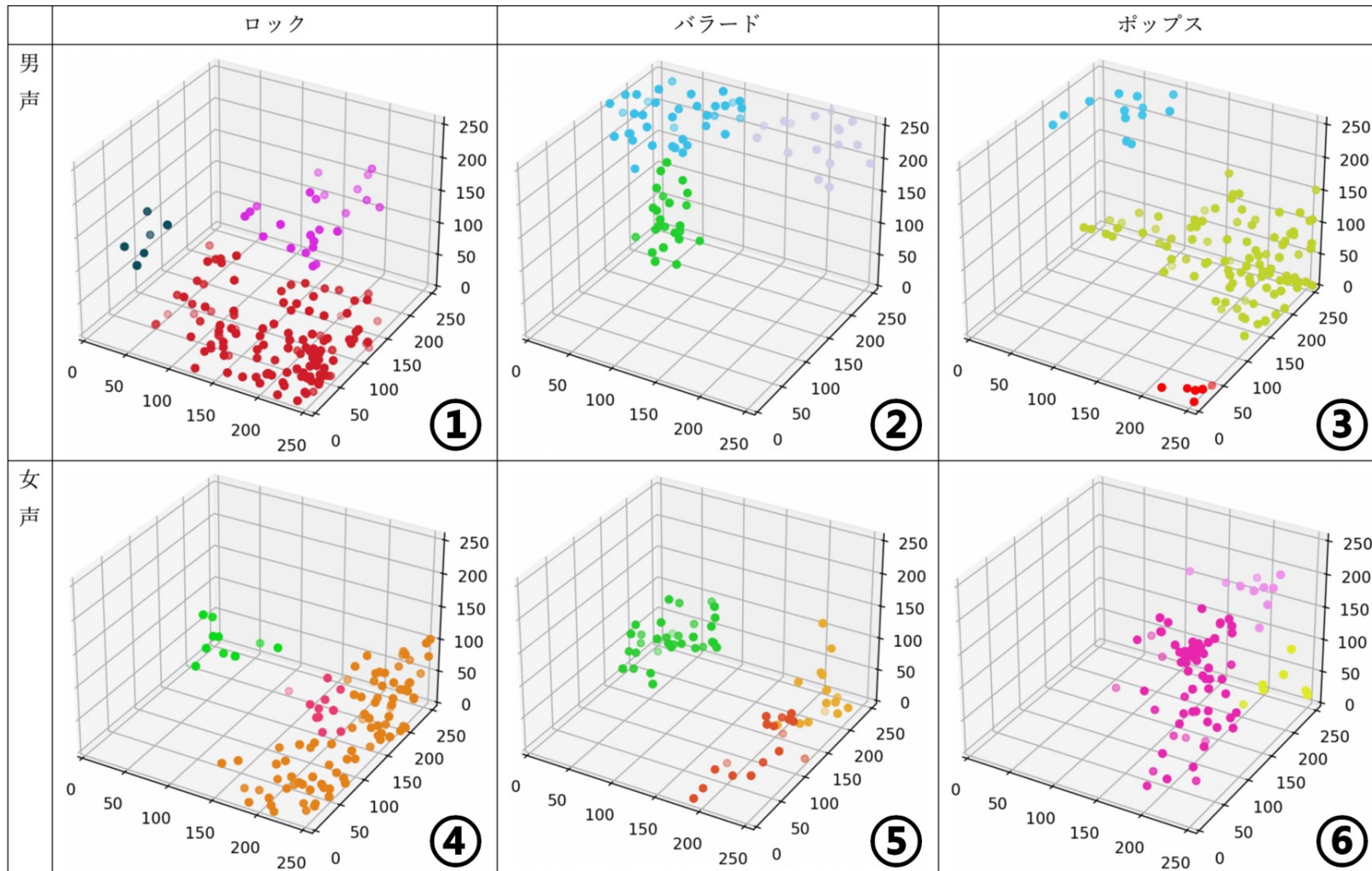
計測されたジャンル、被験者ごとに注視された色記録のクラスタ数

計測されたジャンル、被験者ごとに注視された色記録のノイズ数

- 「女声バラード」で計測された平均の色種類数(クラスタ数)と平均のノイズ数、両方も一番多かった(20.3 & 47.0) → 発散
- 「女声ポップス」で計測された平均の色種類数(クラスタ数)と平均のノイズ数、両方も一番少なかった(12.3 & 32.2) → 集中

可視化結果 (I)

- 最初10秒間の被験者全員の色記録に対する上位3つの色クラスタの可視化

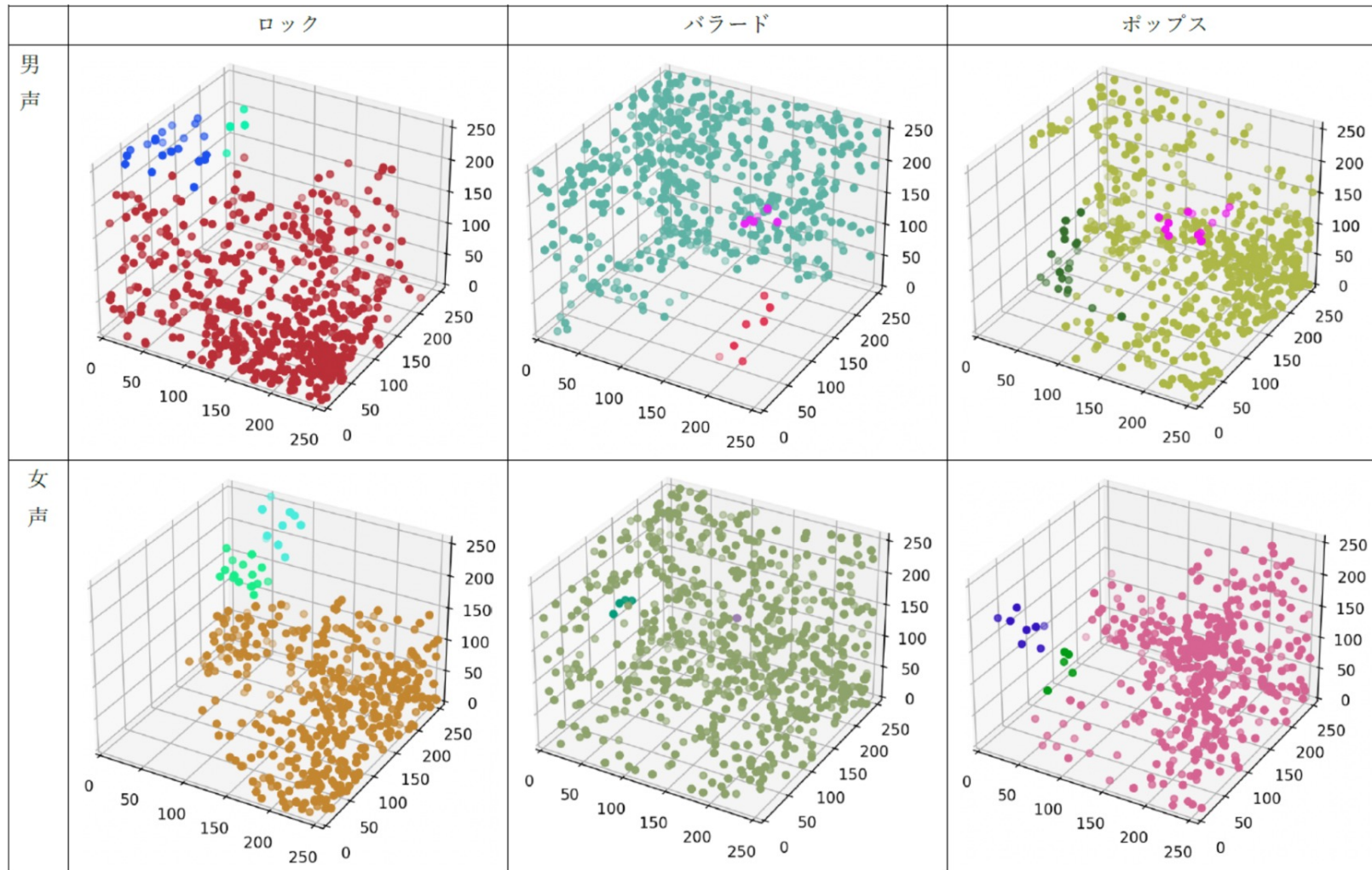


考察

- 男声ロック(①)、男声バラード(②)、男声ポップス(③)、女声ロック(④)、女声ポップス(⑥)で計測された色記録はRGB空間上で一箇所の所に集中している傾向がある
- 女声バラード(⑤)での色記録は、緑色やオレンジ色のような空間上に離れている色がそれぞれ多く記録されていた

考察 (2)

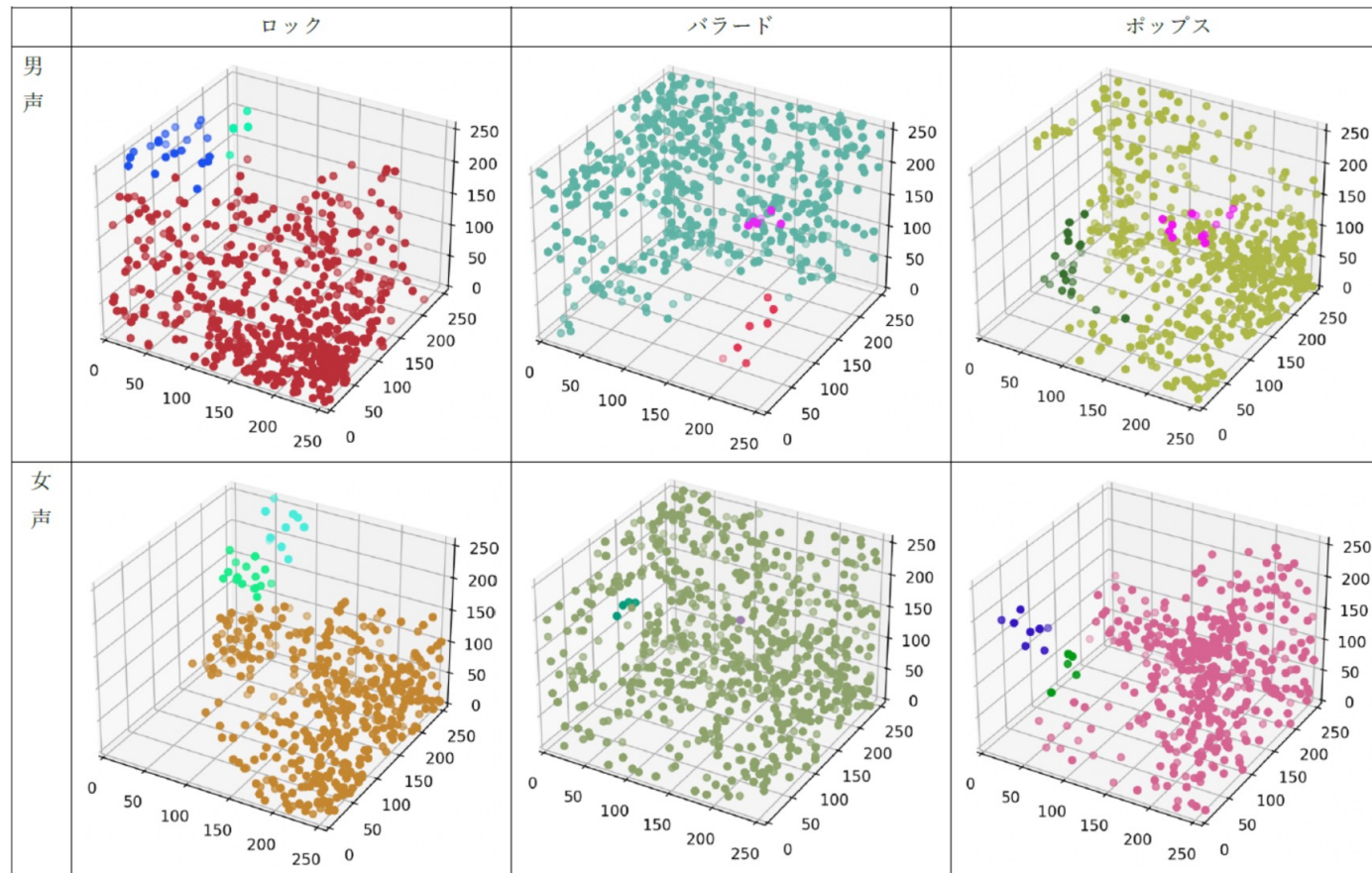
- 全25秒間の被験者全員の色記録に対する上位3つの色クラスタの可視化



– 女声バラードはRGB空間上のあちこちに点在している

考察 (3)

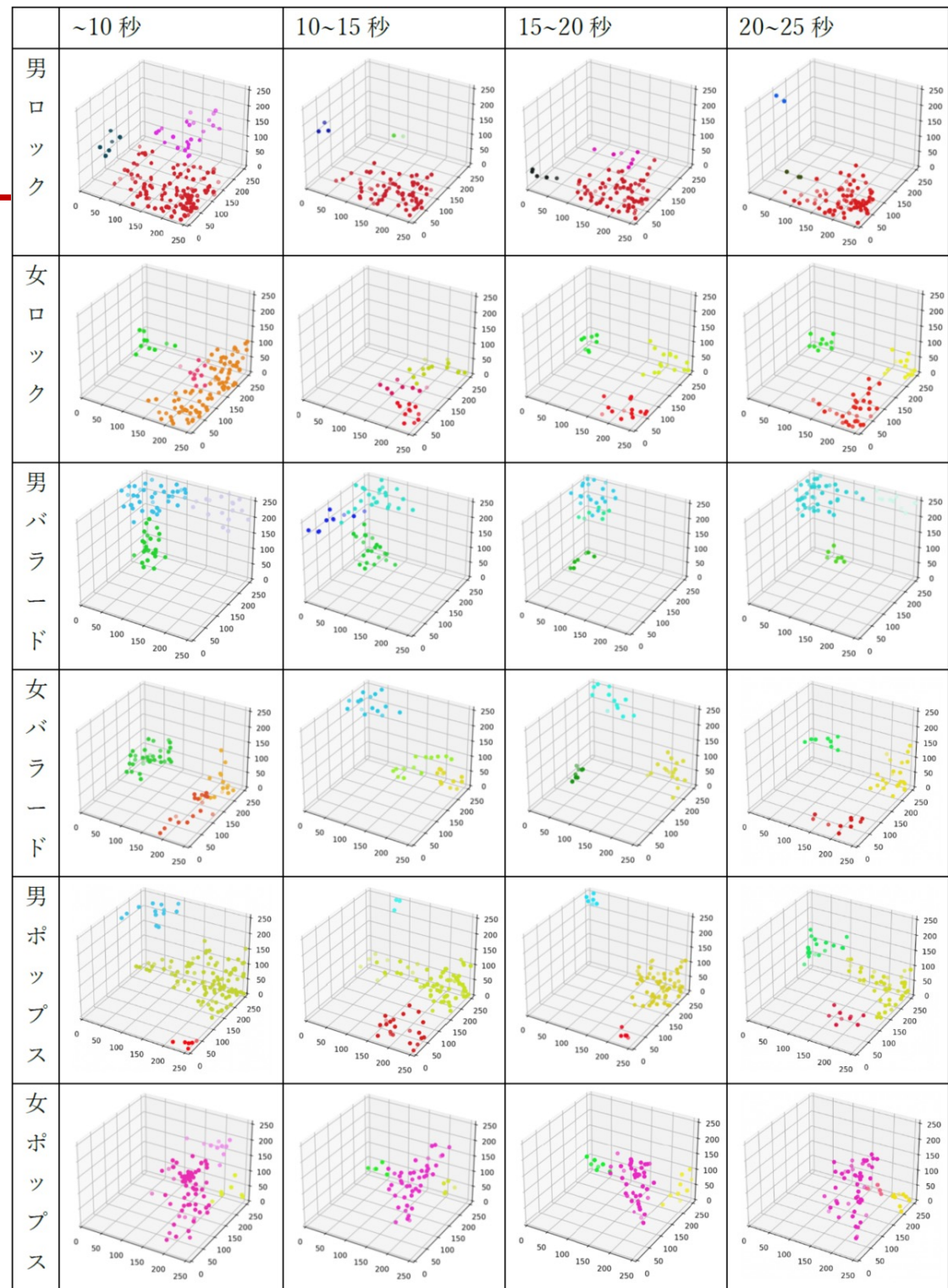
- 全25秒間の被験者全員の色記録に対する上位3つの色クラスタの可視化



- 女声バラード以外の5つのジャンルでは、2位と3位の色クラスタは、1位の色クラスタに比べて、数が少なすぎて、ほぼノイズみたいな存在になっている

考察 (3)

- 時間により被験者全員の色記録に対する上位3つの色クラスタの推移の可視化
 - 女声バラード以外の音楽ジャンルで想像された色の種類が安定している傾向が分かれる



おわりに

- 視線計測装置を用いてユーザに注視された色の記録、及び可視化するシステムを提案
 - ランダム色のパネルを作成し、色と座標の情報を記録
 - 視線計測でユーザの注視動作の判断及び注視色の記録
 - クラスタリング (DBSCAN) で色記録から注視色の推定及び可視化
- 提案システムを利用して、6つの音楽ジャンルから想像された色の計測と推定を被験者実験で調査した
 - 女声バラードでの想像色は発散、他の5つのジャンルでの想像色は集中する傾向がある
- 被験者のプロフィールを増えて、音楽の想像色が年齢層や文化背景などでの違いや関連性を調べる予定