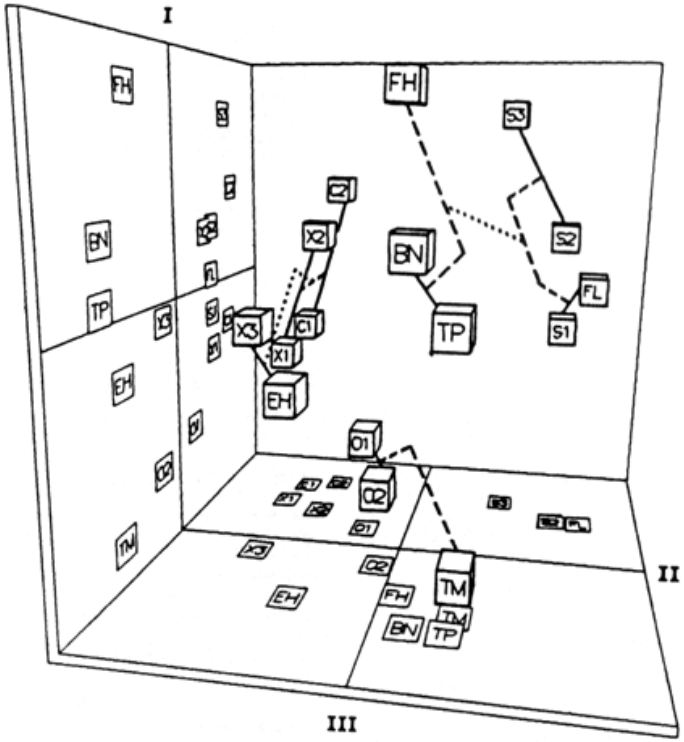


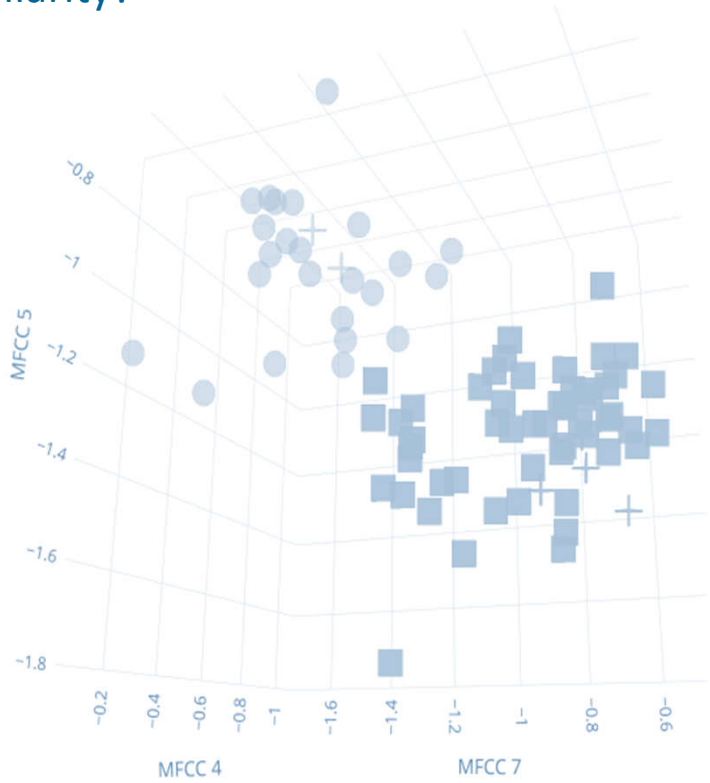
克里斯托夫·罗伊特, 伊莎贝拉-切泽迪克-艾森伯格, 萨利赫-西迪克

声学、心理学与表演艺术等：维也纳大学的系统音乐学
上海音乐学院, 上海, 27.10.-4.11.2023

A Timbre Space Odyssey

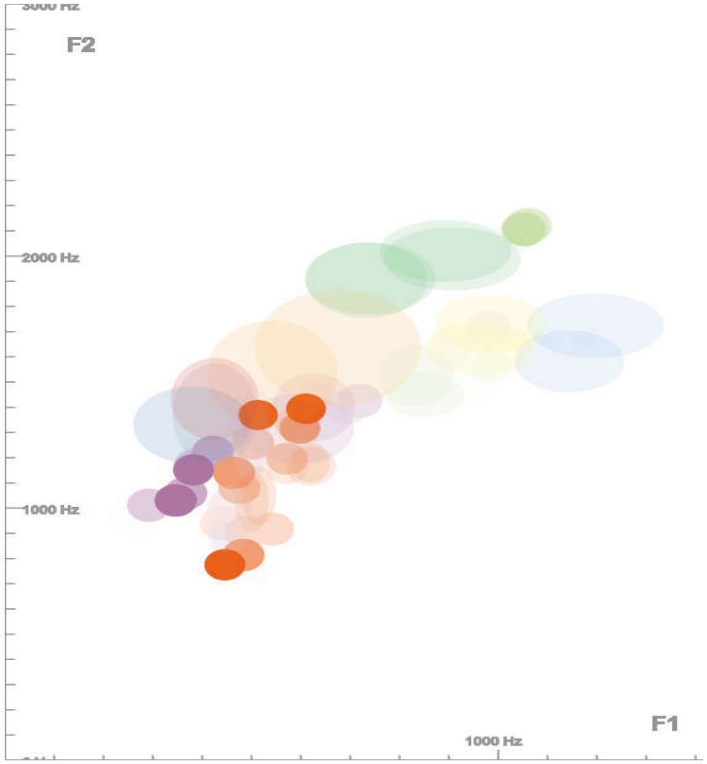
How to measure timbre similarity?





音色太空漫游

如何测量音色相似度？



A Timbre Space Odyssey

Timbre
Multidimensional Dimensions of Timbre

Space
Popular approaches of Timbre Visualisation

1 Timbre Spaces
2 MFCCs
3 Formants

Odyssey
Dynamic Timbre Maps

1 Dynamic Spectral Centroid/Flux Map
2 Dynamic MFCC Map
3 Dynamic Formant Map

Back on Earth: Take Home Message

音色太空漫游

音色
音色的多维维度

空间
流行的音色可视化方法

1 音色空间
2 梅尔频率倒谱系数（MFCCs）
3 共振峰

漫游
动态音色映射

1 动态频谱质心/通量映射
2 动态MFCC映射
3 动态共振峰映射

回到地球：带回家的信息

A Timbre Space Odyssey

音色太空漫游

Timbre
音色

A Timbre Space Odyssey

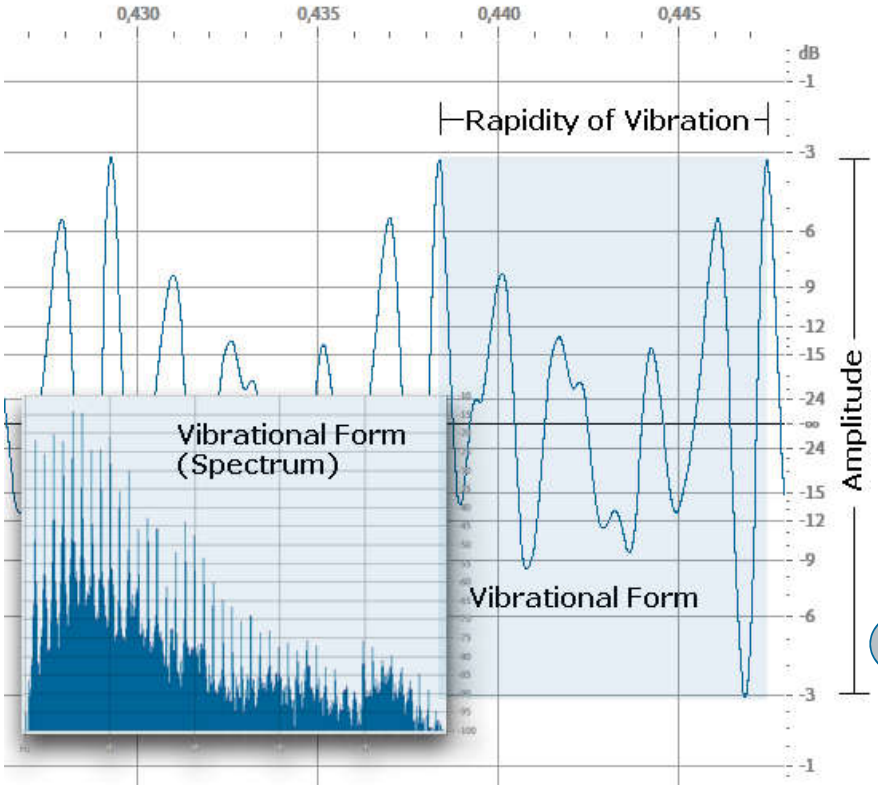
Multidimensional Dimensions of Timbre

“We have seen that [the perceived] ***force*** *depended on amplitude, and* [the perceived] ***pitch*** *on rapidity of vibration: nothing else was left to distinguish* ***quality*** [= timbre] *but vibrational form.”*

(Helmholtz 1895, p. 65)

音色太空漫游

音色的多维维度



"我们已经看到，[感知的] **力度**取决于振幅，[感知的]**音高**取决于振动的快慢：除了振动形式，没有其他东西可以区分**音质**[=音色]"。

(Helmholtz 1895 年, 第 65 页)

A Timbre Space Odyssey

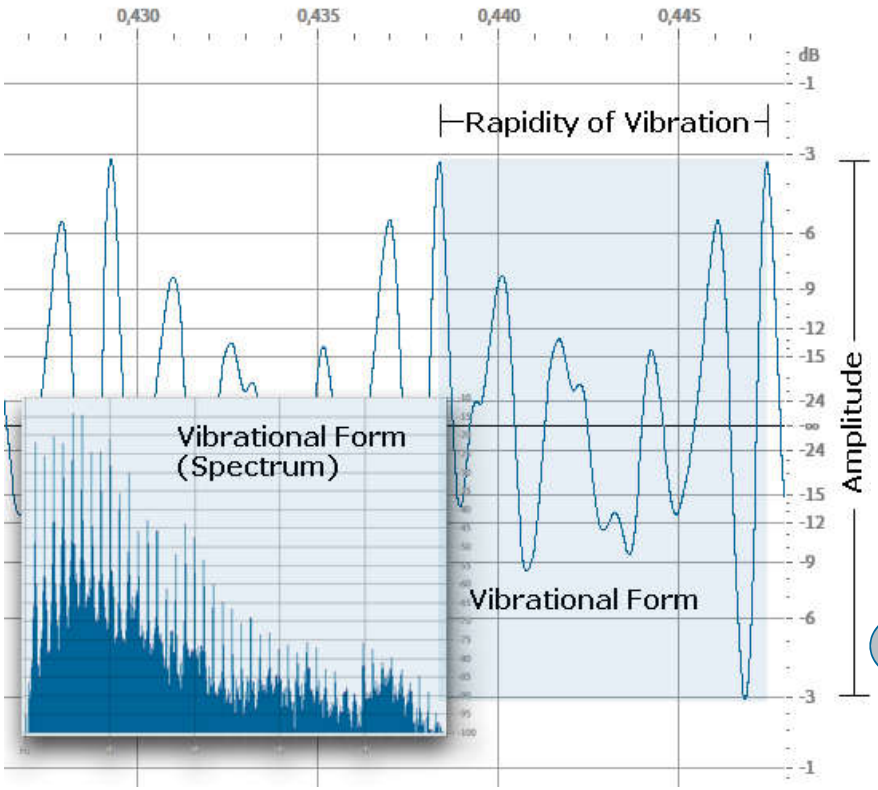
音色太空漫游

Multidimensional Dimensions of Timbre

音色的多维维度

"We have seen that [the perceived] ***force*** *depended on amplitude, and* [the perceived] ***pitch*** *on rapidity of vibration: nothing else was left to distinguish* ***quality*** [= timbre] *but vibrational form."*

(Helmholtz 1895, p. 65)



"我们已经看到，[感知的] **力度取决于振幅，[感知的]音高取决于振**动的快慢：除了振动形**式，没有其他**东西可以**区分音质**[=音色]"。

(Helmholtz 1895 年, 第 65 页)

In short:

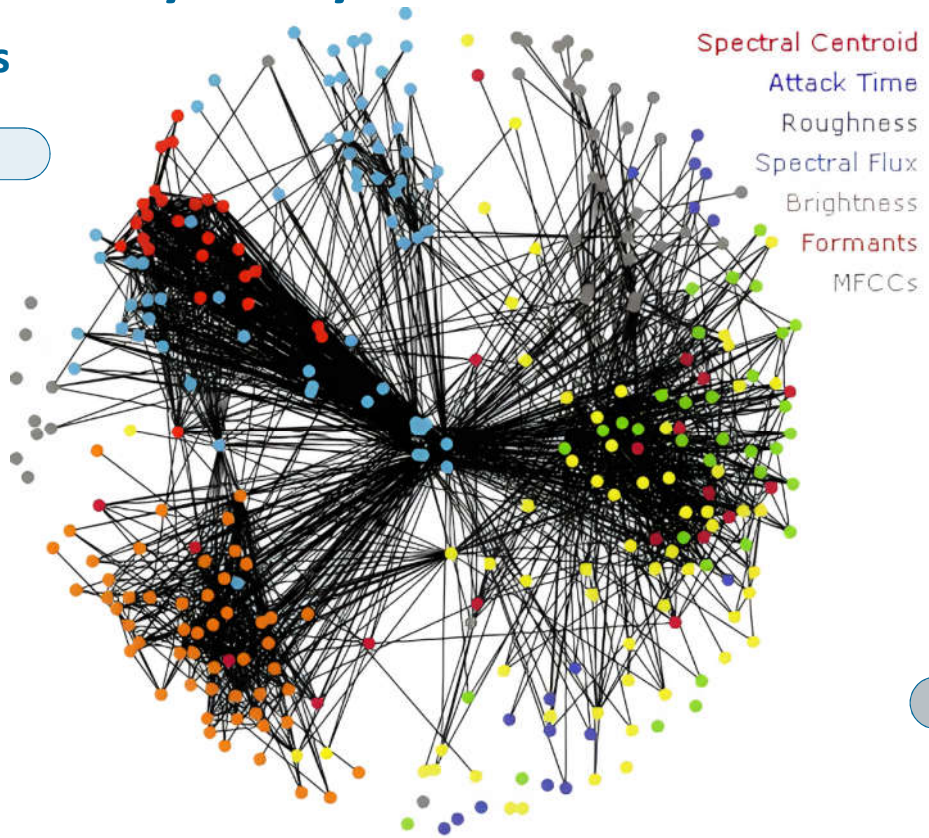
简而言之：

Amplitude	→	loudness	**振幅**	→	**响度**
Frequency	→	pitch	频率	→	**音高**
Waveform (Spectrum)	→	timbre	**波形**（频谱）	→	**音色**

A Timbre Space Odyssey 音色太空漫游

Multidimensional Dimensions

*“The manifoldness of entities [...] The features which are summarized in the word and concept of **timbre** build up such a colorful mixture that one becomes desperate to reconcile them.“*
(Stumpf 1890, p. 514)



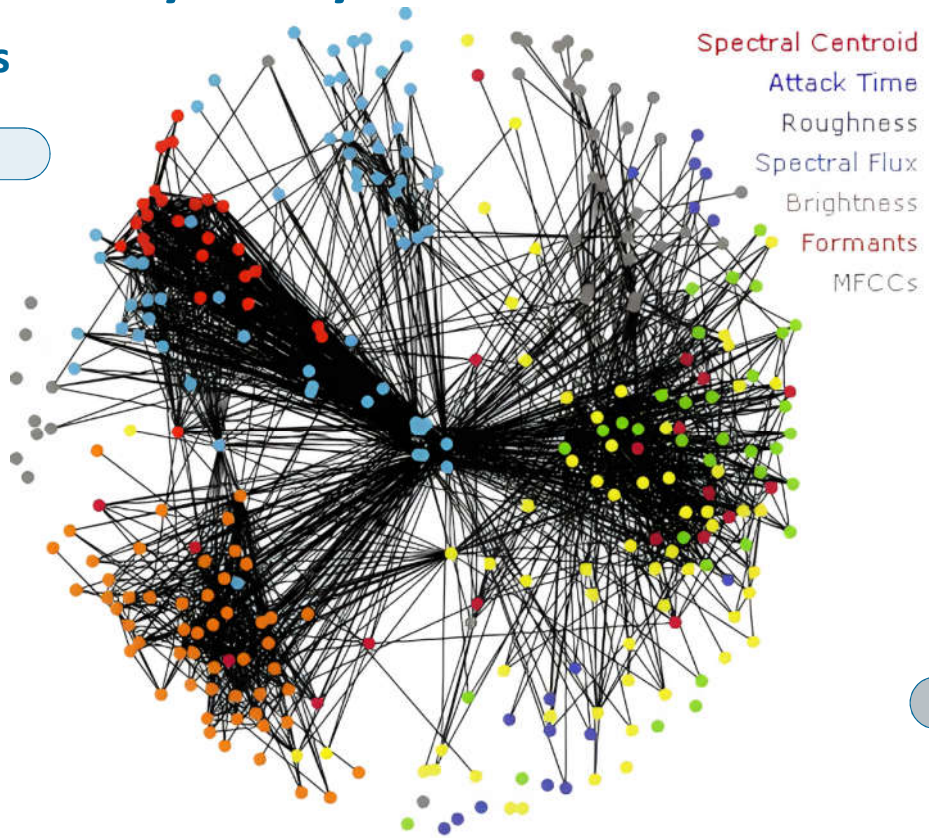
"实体的多面性[.....]**音色**这个词和概念所概括**的特征构成了五彩**缤纷**的混合物**，让人急于将**其**调和"。

(Stumpf 1890 年，第 514 页)

A Timbre Space Odyssey 音色太空漫游

Multidimensional Dimensions

*“The manifoldness of entities [...] The features which are summarized in the word and concept of **timbre** build up such a colorful mixture that one becomes desperate to reconcile them.“*
(Stumpf 1890, p. 514)



"实体的多面性[……]**音色**这个词和概念所概括**的特征构成了五彩**缤纷**的混合物**，让人急于将**其**调和"。

(Stumpf 1890 年，第 514 页)

In short:

"Until careful scientific work has been done on the subject, it can hardly be possible to say more about timbre than that it is a multidimensional dimension".

(Licklider 1951, p. 1019)

简而言之：

*"**在**对这一主题进行细致的科学研究之前，**我**们很难对音色有更多的了解，**只能**说它是一个多维的维度"。*

(Licklider 1951 年，第 1019 页)

A Timbre Space Odyssey

Multidimensional Dimensions of Timbre

*“**Timbre**. Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are **dissimilar**.”*

(most popular definition of timbre; American National Standards Institute (ANSI) 1960, p. 45; § 12.9)

音色太空漫游

音色的多维维度

*"**音色**。音色是听觉的一种属性，听者可以根据这种属性来判断两种声音的音量和音高是**否相似**"。*

(最流行的音色定义；美国国家标准协会 (ANSI) 1960, p. 45; § 12.9)

A Timbre Space Odyssey

Multidimensional Dimensions of Timbre

*"**Timbre**. Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are **dissimilar**."*

(most popular definition of timbre; American National Standards Institute (ANSI) 1960, p. 45; § 12.9)

In short: Timbre = Auditory Sensation ***minus*** Pitch ***minus*** Loudness

音色太空漫游

音色的多维维度

*"**音色**。音色是听觉的一种属性，听者可以根据这种属性来判断两种声音的音量和音高是**否相似**"。*

(最流行的音色定义；美国国家标准协会 (ANSI) 1960, p. 45; § 12.9)

简而言之：音色 = 听觉减去音高减去响度

A Timbre Space Odyssey

Multidimensional Dimensions of Timbre

*"**Timbre**. Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are **dissimilar**."*

(most popular definition of timbre; American National Standards Institute (ANSI) 1960, p. 45; § 12.9)

In short: Timbre = Auditory Sensation ***minus*** Pitch ***minus*** Loudness

1. It is a definition ex negativo ("timbre is not … ")
2. Timbre is not an independent feature, but it is something in comparison to.
3. Timbre is defined as something with a pitch. But there exist many timbres without a pitch.

音色太空漫游

音色的多维维度

*"**音色**。音色是听觉的一种属性，听者可以根据这种属性来判断两种声音的音量和音高是**否相似**"。*

(最流行的音色定义；美国国家标准协会 (ANSI) 1960, p. 45; § 12.9)

简而言之：音色 = 听觉减**去音高减去响度**

1. 这是一个否定式定义（**"音色不是……"**）。
2. **音色并不是一个独立的特征，但可以与之相比。**
3. **音色被定**义成有**音高的**东西。但也存在许多没有**音高的音色。**

A Timbre Space Odyssey

Multidimensional Dimensions of Timbre

*"**Timbre**. Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are **dissimilar**."*

(most popular definition of timbre; American National Standards Institute (ANSI) 1960, p. 45; § 12.9)

In short: Timbre = Auditory Sensation ***minus*** Pitch ***minus*** Loudness

1. It is a definition ex negativo ("timbre is not … ")
2. Timbre is not an independent feature, but it is something in comparison to.
3. Timbre is defined as something with a pitch. But there exist many timbres without a pitch.

"I think the definition of timbre by the American Standards Association should be this: 'We do not know how to define timbre, but it is not loudness and it is not pitch.'"

(Bregman 1990, p. 93)

音色太空漫游

音色的多维维度

*"**音色**。音色是听觉的一种属性，听者可以根据这种属性来判断两种声音的音量和音高是**否相似**"。*

(最流行的音色定义；美国国家标准协会 (ANSI) 1960, p. 45; § 12.9)

简而言之：音色 = 听觉减去音高减去响度

1. 这是一个否定式定义（**"音色不是……"**）。
2. **音色并不是一个独立的特征，但可以与之相比。**
3. **音色被定**义成有音高的东西。但也存在许多没有**音高的音色。**

"我认为美国标准协会对音色的定义应该是这样的： **我**们不知道如何定义**音色，但它不是响度，也不是音高。**"

(Bregman 1990年，**第** 93 页

A Timbre Space Odyssey

Multidimensional Dimensions of Timbre

Timbre has mainly to do with the waveform or spectrum of a sound

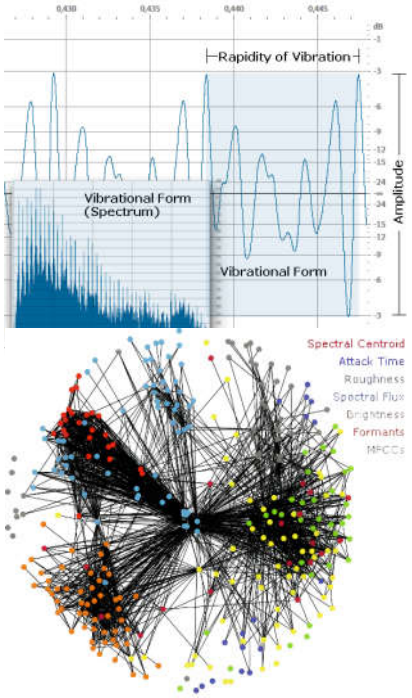
Timbre depends on acoustical features

There seems to be no proper definition for timbre.

(„The problem with timbre is that it is the name for an ill-defined wastebasket category“, Bregman 1990, p. 92)

音色太空漫游

音色的多维维度



音色主要与声音的波形或频谱有关

音色取决于声学特征

音色似乎没有一个恰当的定义。

("音色的问题在于，它是一个定义不清的垃圾桶类别的名称", Bregman 1990 年，第 92 页)

A Timbre Space Odyssey

Multidimensional Dimensions of Timbre

Timbre has mainly to do with the waveform or spectrum of a sound

Timbre depends on acoustical features

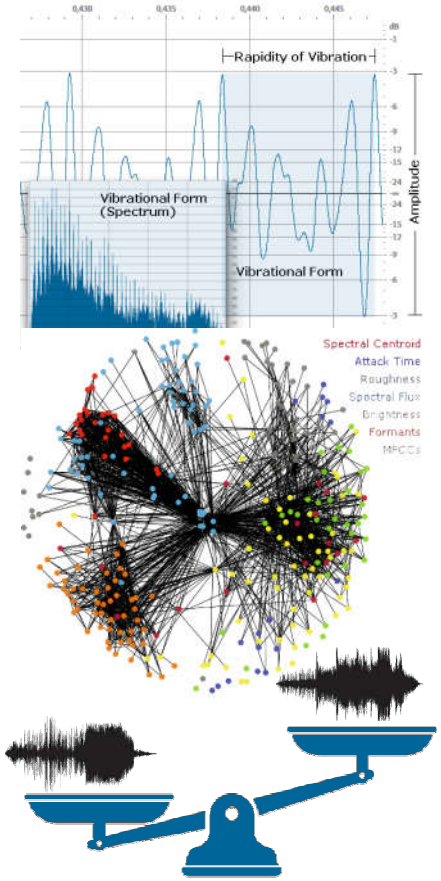
There seems to be no proper definition for timbre.

(„The problem with timbre is that it is the name for an ill-defined wastebasket category“, Bregman 1990, p. 92)

ANSI-Definition

(Timbre = Auditory Sensation *minus* Pitch *minus* Loudness)

- Most popular timbre definition in literature
- Most useless timbre definition



音色太空漫游

音色的多维维度

音色主要与声音的波形或频谱有关

音色取决于声学特征

音色似乎没有一个恰当的定义。

("音色的问题在于，它是一个定义不清的垃圾桶类别的名称", Bregman 1990 年，第 92 页)

ANSI 定义

(音色 = 听觉感觉***减去***音高再***减去***响度)

- 文学作品中最流行的音色定义
- 最无用的音色定义

A Timbre Space Odyssey

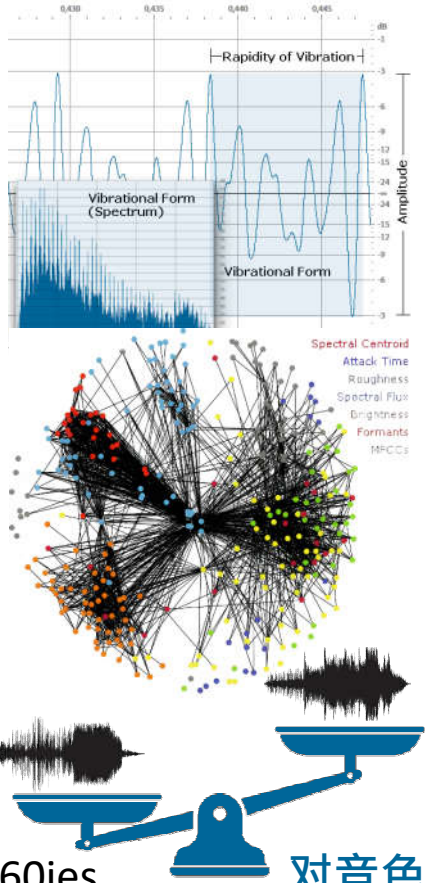
Multidimensional Dimensions of Timbre

Timbre has mainly to do with the waveform or spectrum of a sound

Timbre depends on acoustical features

There seems to be no proper definition for timbre.

(„The problem with timbre is that it is the name for an ill-defined wastebasket category“, Bregman 1990, p. 92)



ANSI-Definition

(Timbre = Auditory Sensation *minus* Pitch *minus* Loudness)

- Most popular timbre definition in literature
- Most useless timbre definition

Implications for timbre research: since 1960ies, instrument timbres have been mostly compared based on one sole single (mostly not typical) pitch.

音色太空漫游

音色的多维维度

音色主要与声音的波形或频谱有关

音色取决于声学特征

音色似乎没有一个恰当的定义。

("音色的问题在于，它是一个定义不清的垃圾桶类别的名称", Bregman 1990 年，第 92 页)

ANSI 定义

(音色 = 听觉感觉*减去*音高再*减去*响度)

- 文学作品中最流行的音色定义
- 最无用的音色定义

对音色研究的影响：自 19 世纪 60 年代以来，乐器音色的比较大多基于唯一的单一音高（大多不典型）。

A Timbre Space Odyssey 音色太空漫游

Space
空间

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

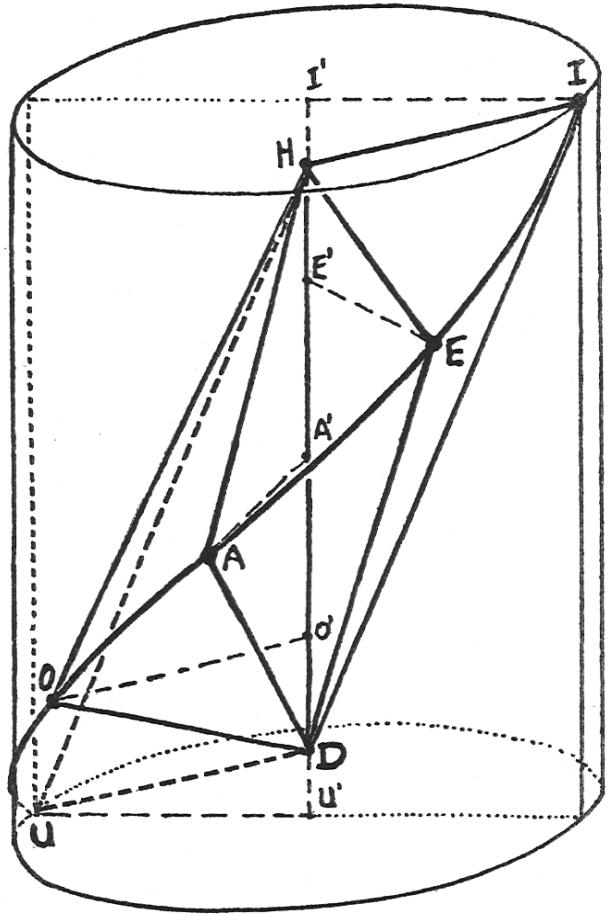
流行的音色可视化方法

1 Timbre Spaces

Albersheim 1939: First attempt at a three-dimensional representation of timbre

Acoustical cylinder of sound colors („Akustischer Farbenkörper“)

- Middle axis: brightness
- Perimeter (helical hue): type of vowel
- Radius: saturation (vowel-likeliness)



Acoustical cylinder of sound colours
声色音响圆柱体
(Albersheim 1939 年，第 353 页)

1 音色空间

阿尔伯斯海姆 1939: **首次**尝试音色的**三**维表现形式

声色音响圆柱体
(„Akustischer Farbenkörper“)

- **中**轴：**亮度**
- **周**长（螺旋角度）：**元音的**类型
- **半径**：饱和度（**元音相似性**）

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

流行的音色可视化方法

1 Timbre Spaces

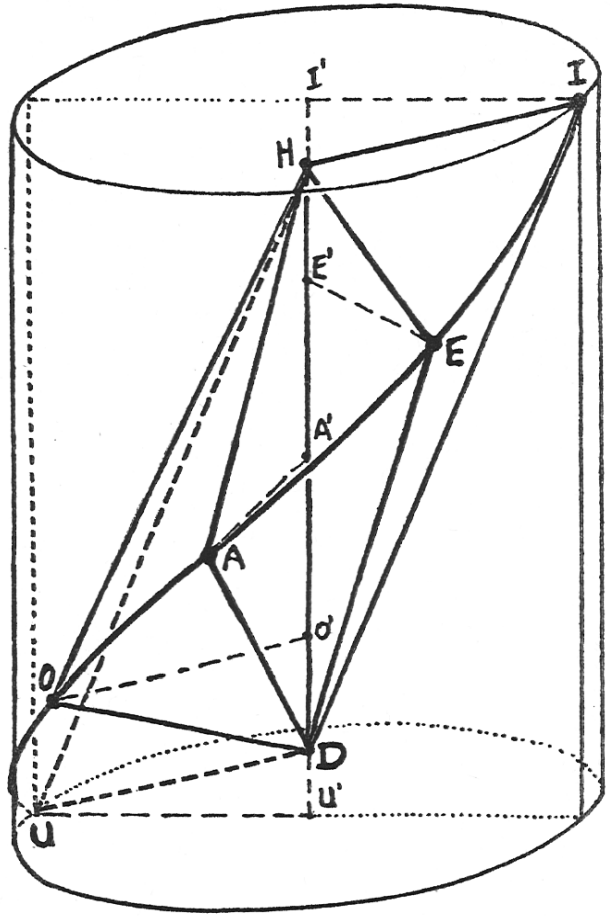
Albersheim 1939: First attempt at a three-dimensional representation of timbre

Acoustical cylinder of sound colors („Akustischer Farbenkörper“)

- Middle axis: brightness
- Perimeter (helical hue): type of vowel
- Radius: saturation (vowel-likeliness)

NICE TRY!

only imagined by Albersheim, not calculated, but what would it have sounded like?



Acoustical cylinder of sound colours
声色音响圆柱体
(Albersheim 1939 年，第 353 页)

1 音色空间

阿尔伯斯海姆 1939: 首次尝试音色的三维表现形式

声色音响圆柱体
(„Akustischer Farbenkörper“)

- 中轴：亮度
- 周长（螺旋角度）：元音的类型
- 半径：饱和度（元音相似性）

NICE TRY!

这只是阿尔伯斯海姆的想象，而不是计算出来的，**但它听起来会是什么样子呢？**

A Timbre Space Odyssey

音色太空漫游

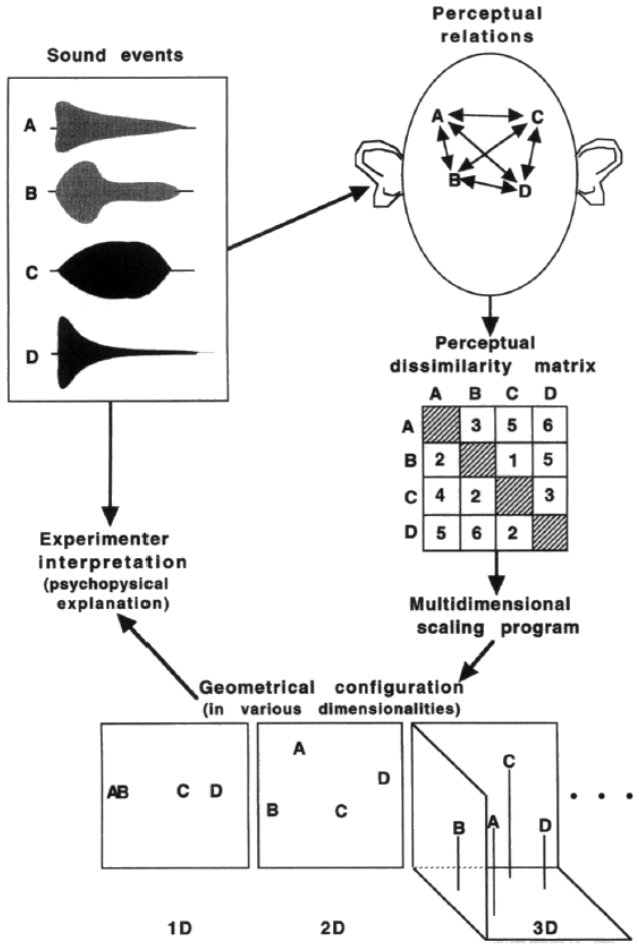
Popular Approaches of Timbre Visualisation

流行的音色可视化方法

1 Timbre Spaces

How to built up a timbre space?

1. Test subjects compare timbres (A-B comparison) and evaluate the perceived (dis)similarity on a (dis)similarity scale.



Building up a Timbre Space / 建立音色空间
(McAdams 1999 年，第 87 页)

1 音色空间

如何建立音色空间？

1. 测试对象对音色进行比较（A-B 比较），并用（不）相似度量表对感知到的（不）相似度进行评估。

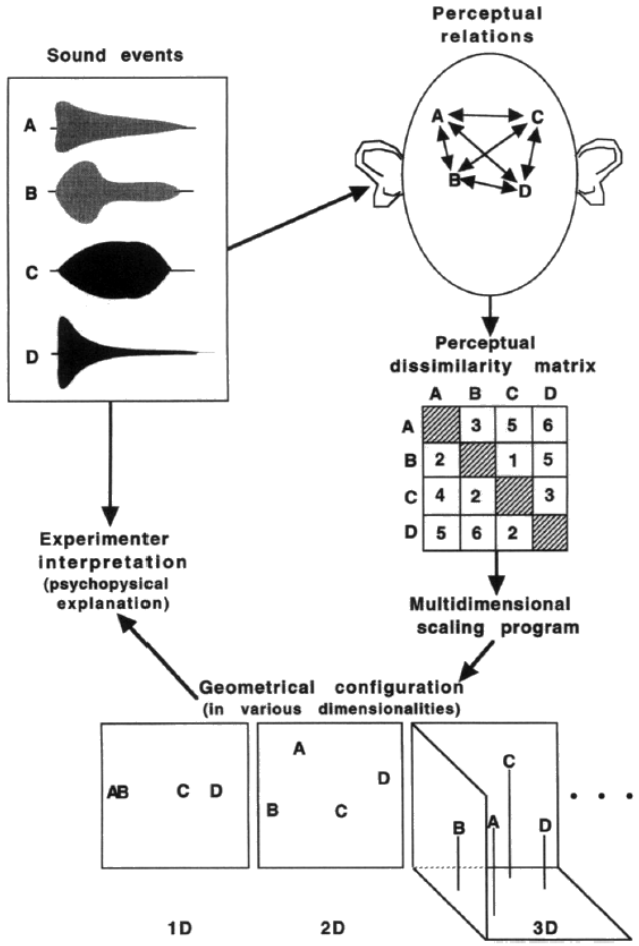
A Timbre Space Odyssey

Popular Approaches of Timbre Visualisation

1 Timbre Spaces

How to built up a timbre space?

1. Test subjects **compare timbres** (A-B comparison) and evaluate the perceived (dis)similarity on a **(dis)similarity scale**.
2. The perceived (dis)similarities of all timbres are listed as numbers in a perceived dissimilarity matrix.



Building up a Timbre Space / 建立音色空间
(McAdams 1999 年，第 87 页)

音色太空漫游

流行的音色可视化方法

1 音色空间

如何建立音色空间？

1. 测试对象对音色进行比较（A-B 比较），并用（不）相似度量表对感知到的（不）相似度进行评估。

2. 所有音色的感知（不）相似度在感知不相似度矩阵中以数字形式列出。

A Timbre Space Odyssey

Popular Approaches of Timbre Visualisation

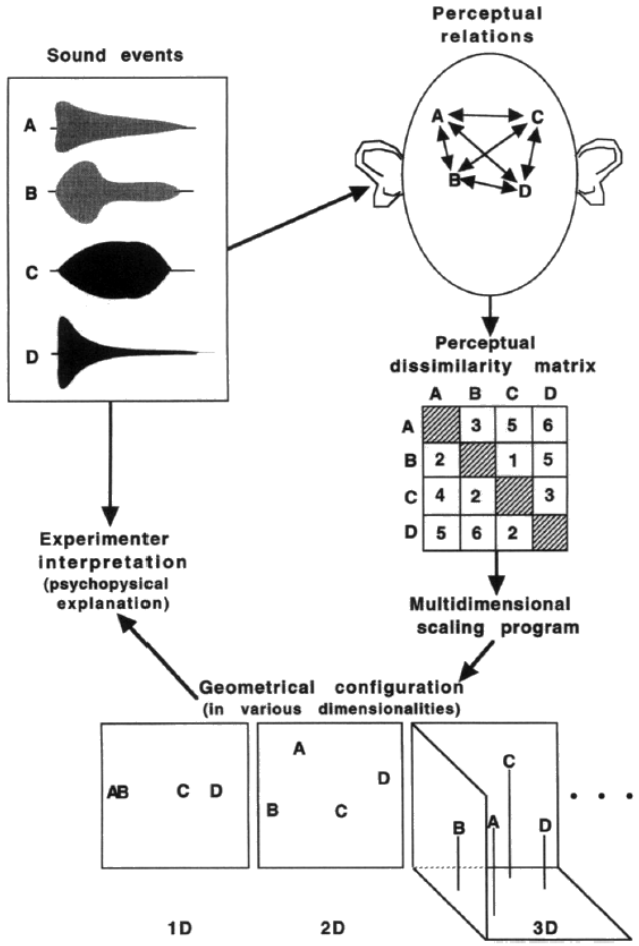
1 Timbre Spaces

How to built up a timbre space?

1. Test subjects compare timbres (A-B comparison) and evaluate the perceived (dis)similarity on a (dis)similarity scale.
2. The perceived (dis)similarities of all timbres are listed as numbers in a perceived dissimilarity matrix.
3. With the help of Multidimensional Scaling the number of perceptual dimensions are calculated.
The closer the entities on these dimensions, the more similar the timbre perception.

音色太空漫游

流行的音色可视化方法



Building up a Timbre Space / 建立音色空间
(McAdams 1999 年，第 87 页)

1 音色空间

如何建立音色空间？

1. 测试对象对音色进行比较（A-B 比较），并用（不）相似度量表对感知到的（不）相似度进行评估。

2. 所有音色的感知（不）相似度在感知不相似度矩阵中以数字形式列出。

3. 借助多维尺度计算出感知维度的数量。这些维度上的实体越接近，音色感知就越相似。

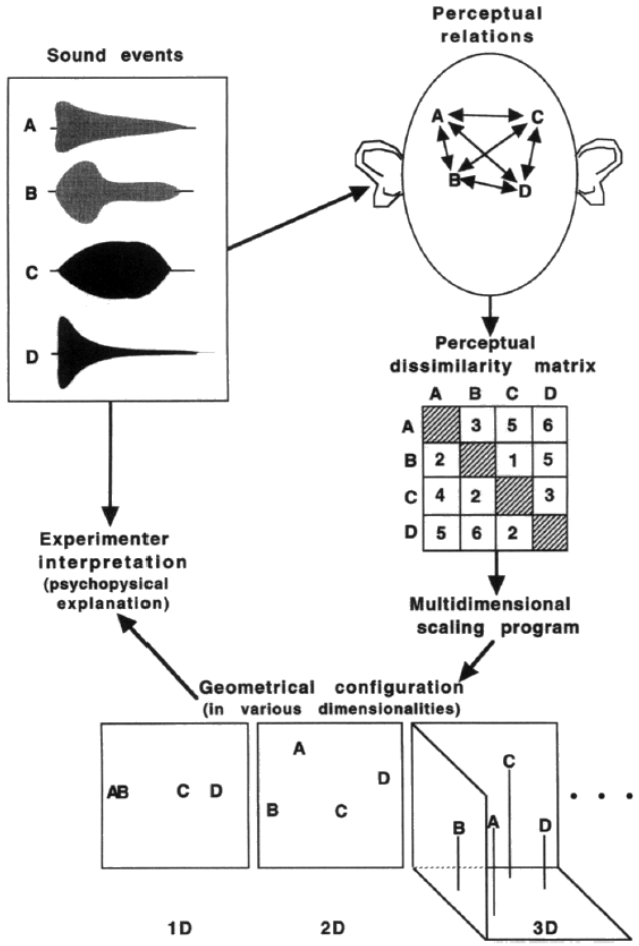
A Timbre Space Odyssey

Popular Approaches of Timbre Visualisation

1 Timbre Spaces

How to built up a timbre space?

1. Test subjects compare timbres (A-B comparison) and evaluate the perceived (dis)similarity on a (dis)similarity scale.
2. The perceived (dis)similarities of all timbres are listed as numbers in a perceived dissimilarity matrix.
3. With the help of Multidimensional Scaling the number of perceptual dimensions are calculated.
The closer the entities on these dimensions, the more similar the timbre perception.
4. The dimensions of the space are tested for correlations with timbre features.



Building up a Timbre Space / 建立音色空间
(McAdams 1999 年，第 87 页)

音色太空漫游

流行的音色可视化方法

1 音色空间

如何建立音色空间？

1. 测试对象对音色进行比较（A-B 比较），并用（不）相似度量表对感知到的（不）相似度进行评估。

2. 所有音色的感知（不）相似度在感知不相似度矩阵中以数字形式列出。

3. 借助多维尺度计算出感知维度的数量。这些维度上的实体越接近，音色感知就越相似。

4. 测试空间维度与音色特征的相关性。

A Timbre Space Odyssey

Popular Approaches of Timbre Visualisation

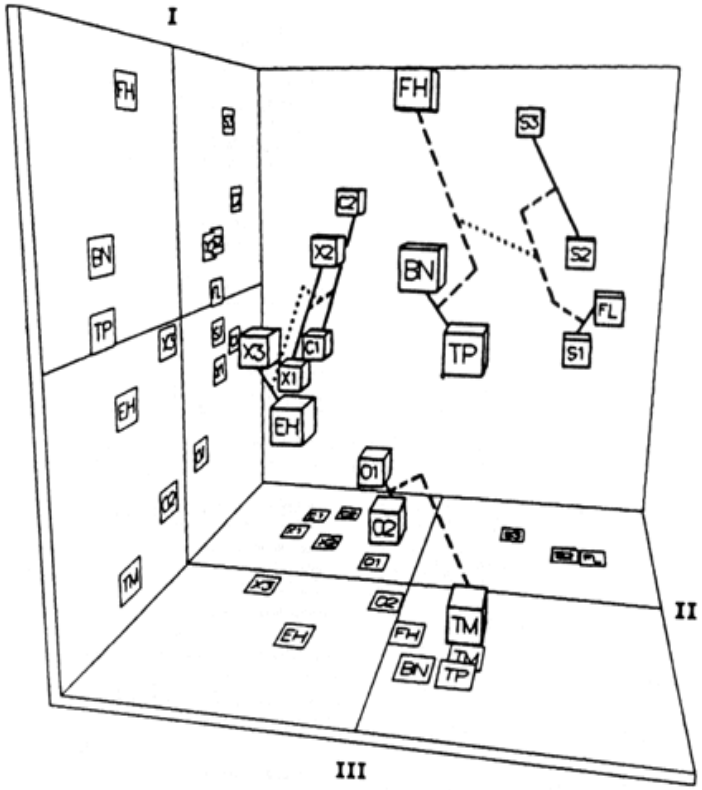
1 Timbre Spaces

Grey 1975: First known Timbre Space

Transfering subjective similarity judgments into spatial distribution using MDS.

Interpretation of spatial dimensions:

- Dimension I: Spectral energy distribution
- Dimension II: Attack transients and synchronicity of the higher partials there
- Dimension III: Fluctuations and inharmonicity



Timbre Space 音色空间
(Grey 1975年，第 62 页)

音色太空漫游

流行的音色可视化方法

1 音色空间

格雷 1975: 首个已知的音色空间

利用 MDS 将主观相似性判断转化为空间分布。

空间维度的解释：

- 维度1：频谱能量分布
- 维度2：起音瞬态和较高分音的同步性
- 维度3：波动和不和谐性

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

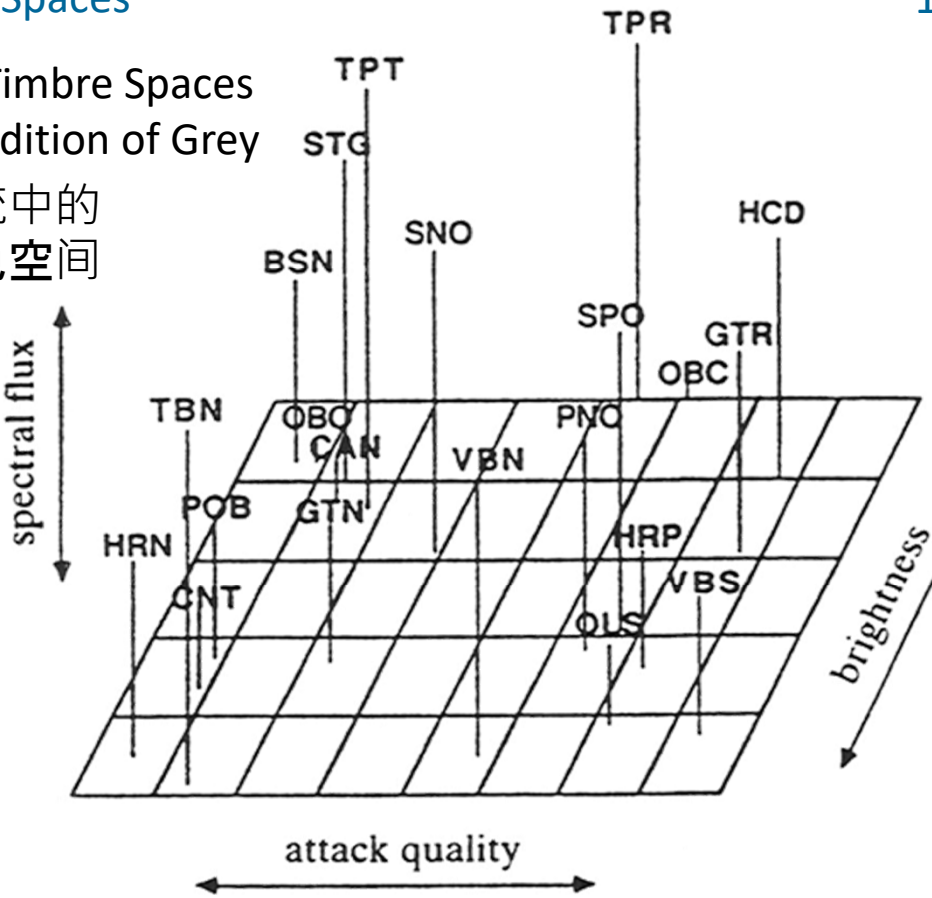
流行的音色可视化方法

1 Timbre Spaces

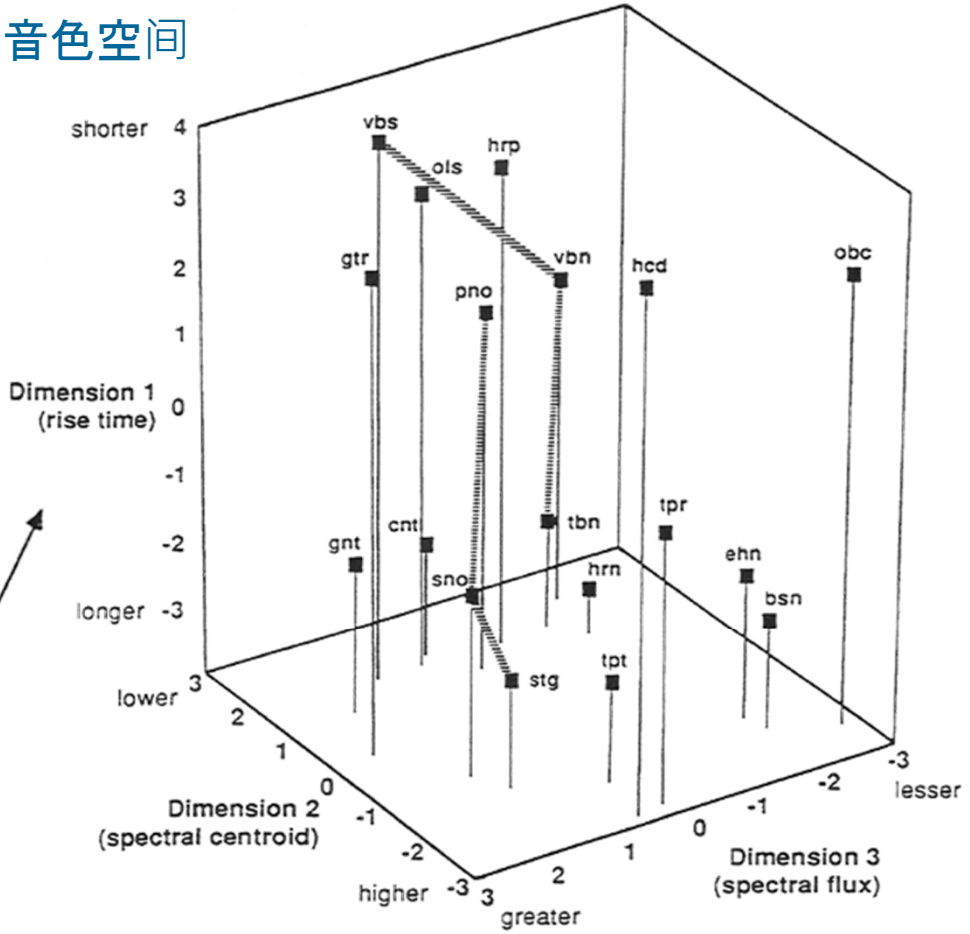
1 音色空间

Further Timbre Spaces in the tradition of Grey

格雷传统中的其他音色空间



Timbre Space based on synthetic FM sounds
基于调频合成声音的音色空间
(Krumhansl 1989, p. 47)



Timbre Space based on synthetic FM sounds
基于调频合成声音的音色空间
(McAdams et al. 1995, p. 185; McAdams 1999, p. 89)

A Timbre Space Odyssey

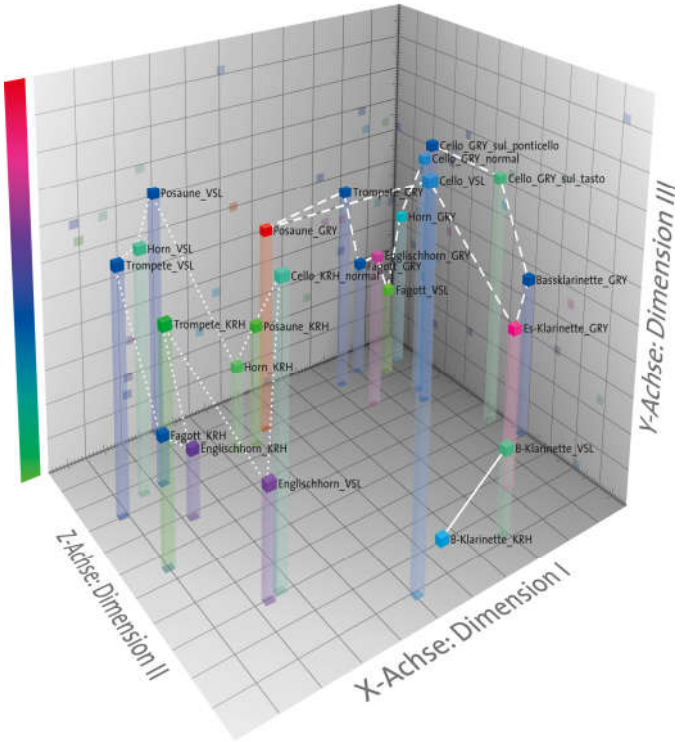
Popular Approaches of Timbre Visualisation

1 Timbre Spaces

Siddiq et al. 2015: Meta Timbre Space

Empirical comparison of three Timbre Spaces:

- Grey (10 sounds, GRY)
- Krumhansl 1989/McAdams et al. 1995
  (7 sounds, KRH)
- Vienna Symphonic Library
  (7 sounds, VSL)



Meta Timbre Space based on the sounds of Grey, Krumhansl/McAdams and Vienna Symphonic Library
基于格雷、克鲁姆汉斯尔/麦克亚当斯和维也纳交响乐声音库声音的元音色空间
(Siddiq, Reuter, Czedik-Eysenberg, Knauf 2015, 第 812 页)

音色太空漫游

流行的音色可视化方法

1 音色空间

Siddiq 等， 2015： **元音色空间**
三种音色空间的实证比较：

- Grey (10种声音, GRY)
- Krumhansl 1989/McAdams 等人, 1995年
  (7种声音, KRH)
- 维也纳交响乐声音库
  (7种声音, VSL)

A Timbre Space Odyssey

Popular Approaches of Timbre Visualisation

1 Timbre Spaces

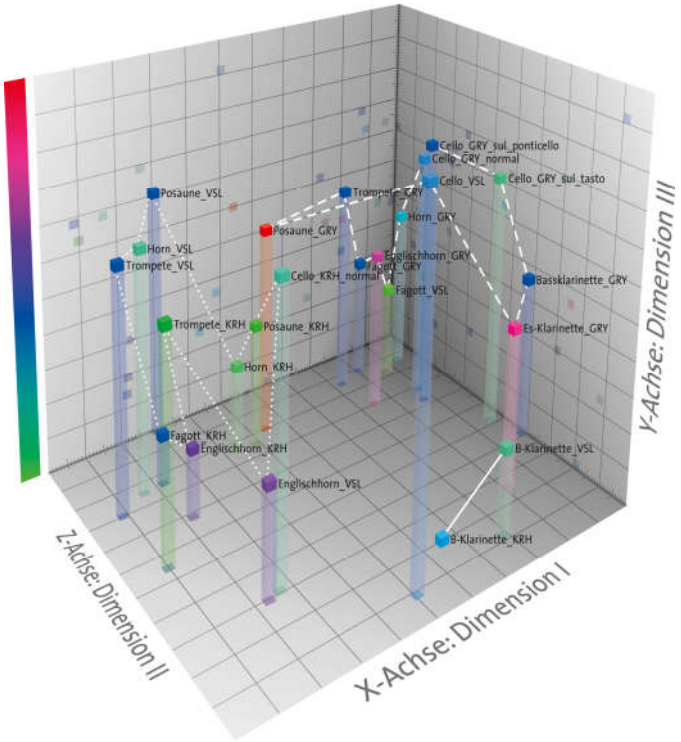
Siddiq et al. 2015: Meta Timbre Space

Empirical comparison of three Timbre Spaces:

- Grey (10 sounds, GRY)
- Krumhansl 1989/McAdams et al. 1995
  (7 sounds, KRH)
- Vienna Symphonic Library
  (7 sounds, VSL)

Results: Timbres of the same stimuli set are more similar than timbres of the same instrument.

Timbre Spaces are hardly generalizable or comparable.



Meta Timbre Space based on the sounds of Grey, Krumhansl/McAdams and Vienna Symphonic Library
基于格雷、克鲁姆汉斯尔/麦克亚当斯和维也纳交响乐声音库声音的元音色空间
(Siddiq, Reuter, Czedik-Eysenberg, Knauf 2015, 第 812 页)

音色太空漫游

流行的音色可视化方法

1 音色空间

Siddiq 等， 2015： 元音色空间
三种音色空间的实证比较：

- Grey (10种声音, GRY)
- Krumhansl 1989/McAdams 等人, 1995年
  (7种声音, KRH)
- 维也纳交响乐声音库
  (7种声音, VSL)

结果：同一刺激集的音色比同一乐器的音色更为相似。

音色空间几乎不具有普遍性或可比性。

A Timbre Space Odyssey

音色太空漫游

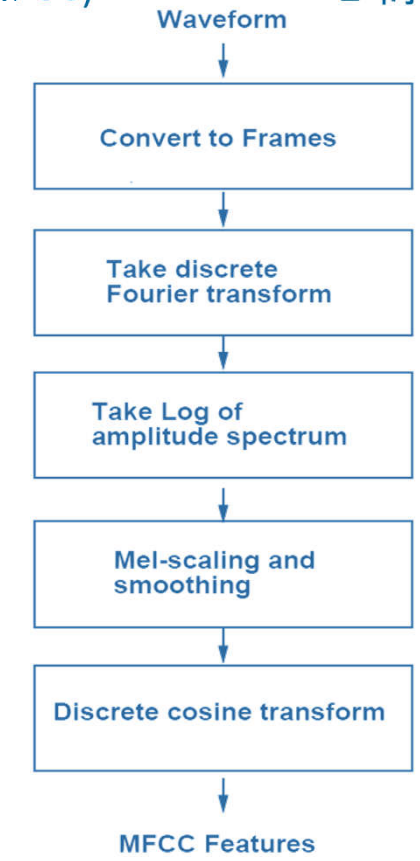
Popular Approaches of Timbre Visualisation

流行的音色可视化方法

2 Mel Frequency Cepstral Coefficients (MFCC)

2 梅尔频率倒谱系数（MFCC）

Davis & Mermelstein 1980:
Automatic calculation for speaker recognition and speech similarity rating.



MFCC calculation scheme
MFCC 计算方案
（Logan，2000 年，第 2 页）

戴维斯和默默斯坦，1980：
说话人识别和语音相似度评价的自动计算

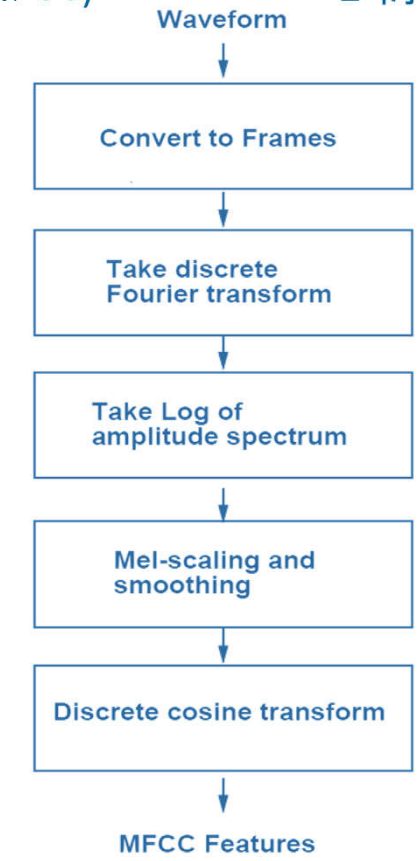
A Timbre Space Odyssey

Popular Approaches of Timbre Visualisation

2 Mel Frequency Cepstral Coefficients (MFCC)

Davis & Mermelstein 1980:
Automatic calculation for speaker recognition and speech similarity rating.

In short, the method calculates for each 20 ms frame of the waveform a Mel-scale adapted Cepstrum (spectrum of a spectrum) and compares the resulting envelope (as a 13-dimensional vector) with a set of standard envelopes (the coefficients).



MFCC calculation scheme
MFCC 计算方案
（Logan，2000 年，第 2 页）

音色太空漫游

流行的音色可视化方法

2 梅尔频率倒谱系数（MFCC）

戴维斯和默默斯坦，1980：
说话人识别和语音相似度评价的自动计算

简而言之，该方法为波形的每个 20 毫秒帧计算一个梅尔尺度适应的倒频谱（频谱的频谱），并将得到的包络（作为一个 13 维向量）与一组标准包络（系数）进行比较。

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

流行的音色可视化方法

2 Mel Frequency Cepstral Coefficients (MFCC)

2 梅尔频率倒谱系数（MFCC）

Davis & Mermelstein 1980:
Automatic calculation for speaker recognition and speech similarity rating

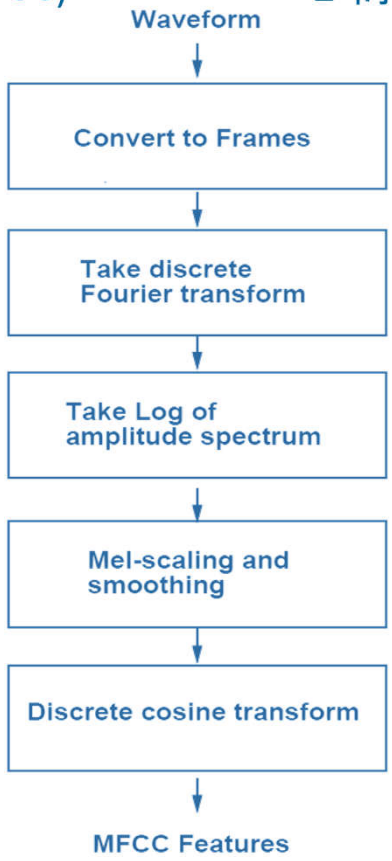
戴维斯和默默斯坦，1980：
说话人识别和语音相似度评价的自动计算

In short, the method calculates for each 20 ms frame of the waveform a Mel-scale adapted Cepstrum (spectrum of a spectrum) and compares the resulting envelope (as a 13-dimensional vector) with a set of standard envelopes (the coefficients).

简而言之，该方法为波形的每个 20 毫秒帧计算一个梅尔尺度适应的倒频谱（频谱的频谱），并将得到的包络（作为一个 13 维向量）与一组标准包络（系数）进行比较。

This method does not explain the mechanism of our audio perception, but the results are very convincing.

这种方法并不能解释我们的音频感知机制，但其结果却非常令人信服。



MFCC calculation scheme
MFCC 计算方案
（Logan，2000 年，第 2 页）

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

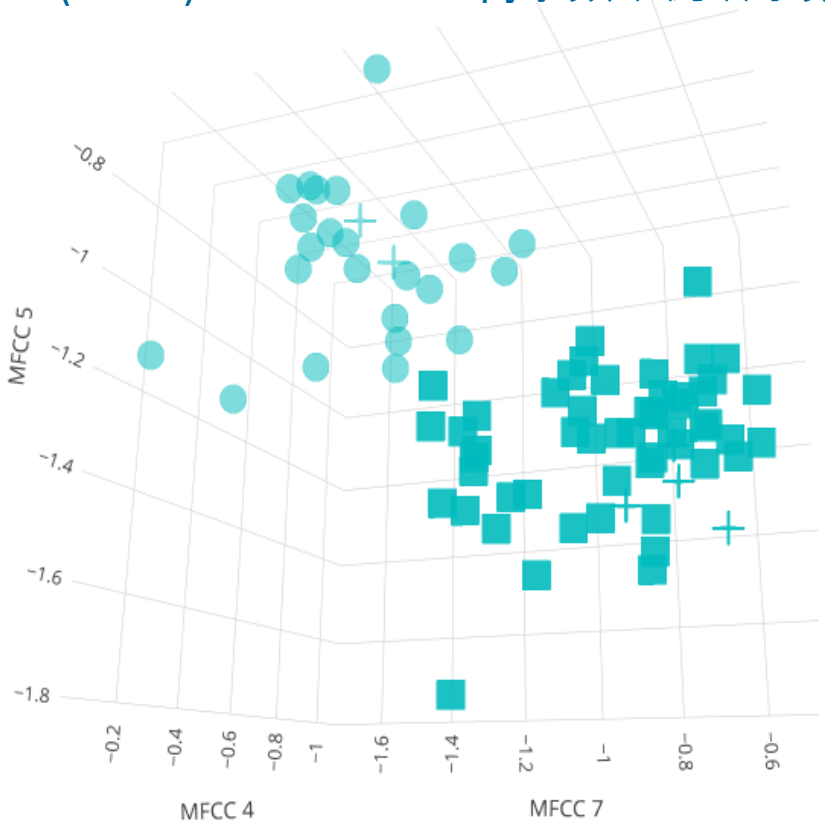
流行的音色可视化方法

2 Mel Frequency Cepstral Coefficients (MFCC)

2 梅尔频率倒谱系数（MFCC）

Pros: well suited for calculating similarities in speech / music / noises / instrumental timbres

优点：非常适合计算语音/**音**乐/**噪音**/乐器音色的相似**性**



Using MFCC4, 5 and 7 to distinguish between dove perch and courtship coos
使用 MFCC4、5 和 7 区分鸽栖和求偶鸣叫
(Reuter, Quigley 2020年)

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

流行的音色可视化方法

2 Mel Frequency Cepstral Coefficients (MFCC)

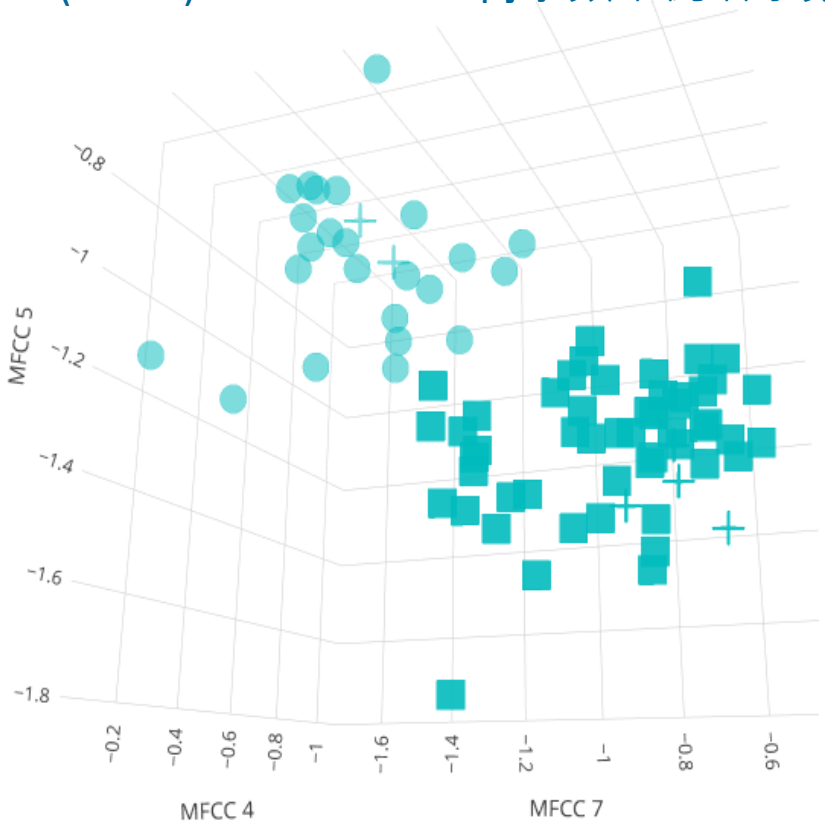
2 梅尔频率倒谱系数（MFCC）

Pros: well suited for calculating similarities in speech / music / noises / instrumental timbres

优点：非常适合计算语音/音乐/**噪音**/乐器音色的相似性

Cons: Less intuitive, numerical output is difficult to understand by humans.

缺点：不够直观，数字输出难以被人类理解。



Using MFCC4, 5 and 7 to distinguish between dove perch and courtship coos
使用 MFCC4、5 和 7 区分鸽栖和求偶鸣叫
(Reuter, Quigley 2020年)

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

流行的音色可视化方法

2 Mel Frequency Cepstral Coefficients (MFCC)

2 梅尔频率倒谱系数（MFCC）

Pros: well suited for calculating similarities in speech / music / noises / instrumental timbres

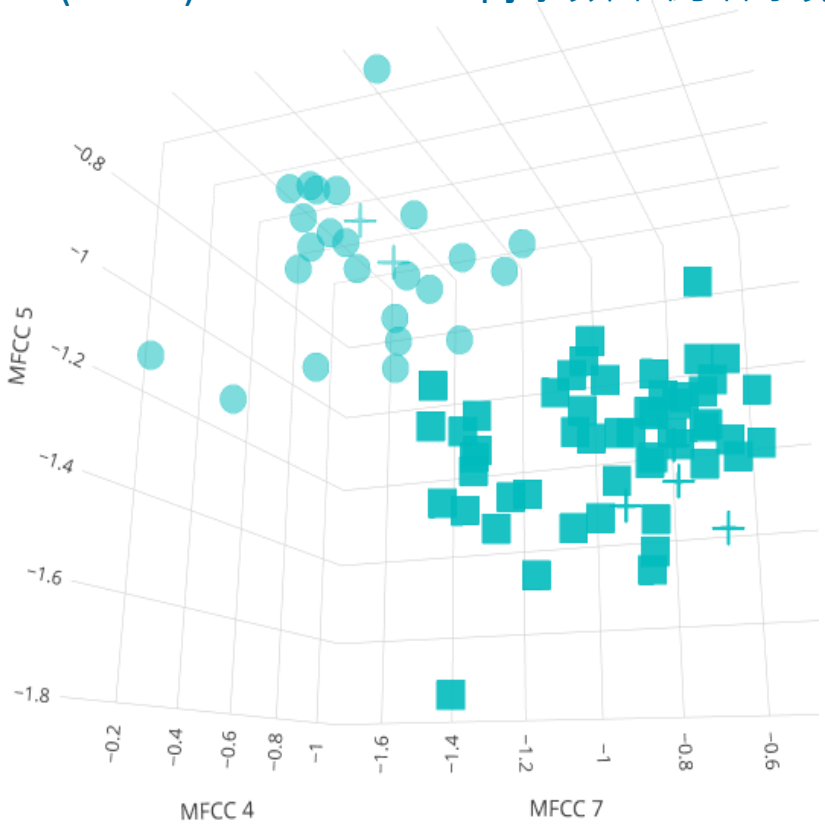
优点：非常适合计算语音/音乐/**噪音**/乐器音色的相似性

Cons: Less intuitive, numerical output is difficult to understand by humans.

缺点：不够直观，数字输出难以被人类理解。

MFCCs play a key role in modern sound similarity computation

MFCC **在**现代声音相似性计**算中**发挥着**关**键作用



Using MFCC4, 5 and 7 to distinguish between dove perch and courtship coos
使用 MFCC4、5 和 7 区分鸽栖和求偶鸣叫
(Reuter, Quigley 2020年)

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

流行的音色可视化方法

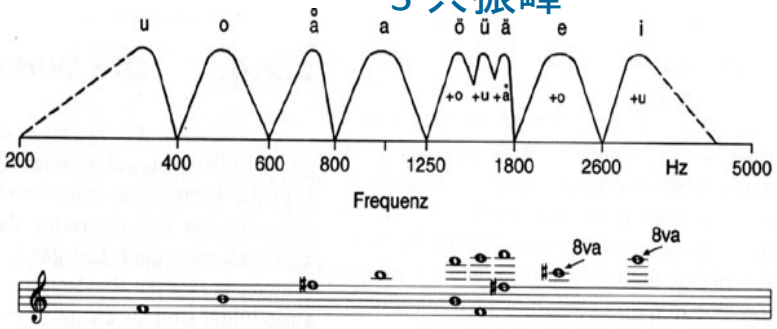
3 Formants

Schumann 1929: Typical maxima in spectra of (mostly) wind instruments.

- pitch-independent (stable)
- like vocal formants (characteristic)

In the German-speaking realm, formants are often used to describe the sound of instruments based on their vowel character.

3 共振峰



Vowel formants and their frequency ranges
元音声母及其频率范围
(Meyer，2015 年，第 33 页)

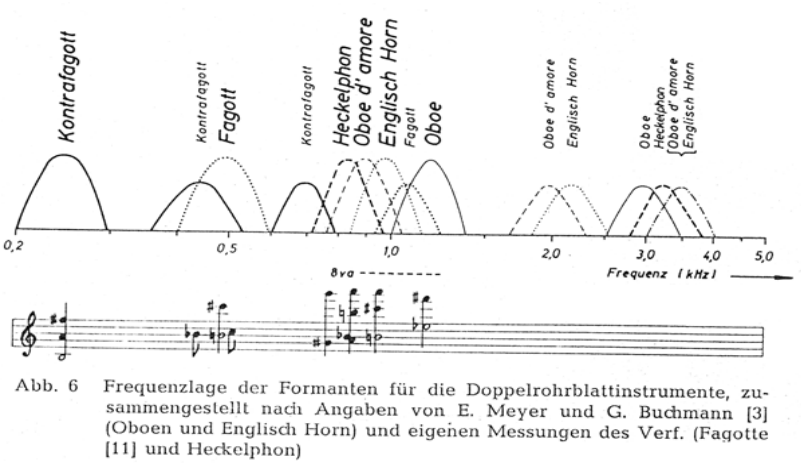


Abb. 6 Frequenzlage der Formanten für die Doppelrohrblattinstrumente, zusammengestellt nach Angaben von E. Meyer und G. Buchmann [3] (Oboen und Englisch Horn) und eigenen Messungen des Verf. (Fagotte [11] und Heckelphon)

Formants of double-reed instruments / 双簧乐器的共振峰
(Meyer 2015，第 63 页)

舒曼 1929：（**大部分**）管乐器频谱中的典型**最大值**。

- **与音高无关**（稳定）
- **像声**乐的共振峰（**特征**）

在德语领域，共振峰通常用于**描述**乐器的声音，基于它们的**元音特征**。

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

流行的音色可视化方法

3 Formants

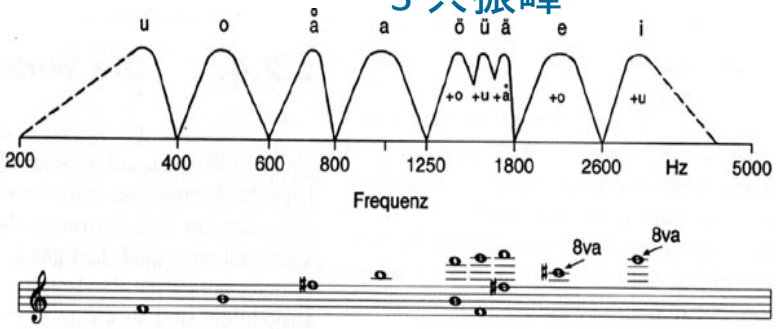
Schumann 1929: Typical maxima in spectra of (mostly) wind instruments.

- pitch-independent (stable)
- like vocal formants (characteristic)

In the German-speaking realm, formants are often used to describe the sound of instruments based on their vowel character.

If formants are removed from a timbre (or the spectral gaps between them), it no longer sounds typical for the respective instrument.

3 共振峰



Vowel formants and their frequency ranges
元音声母及其频率范围
(Meyer，2015 年，第 33 页)

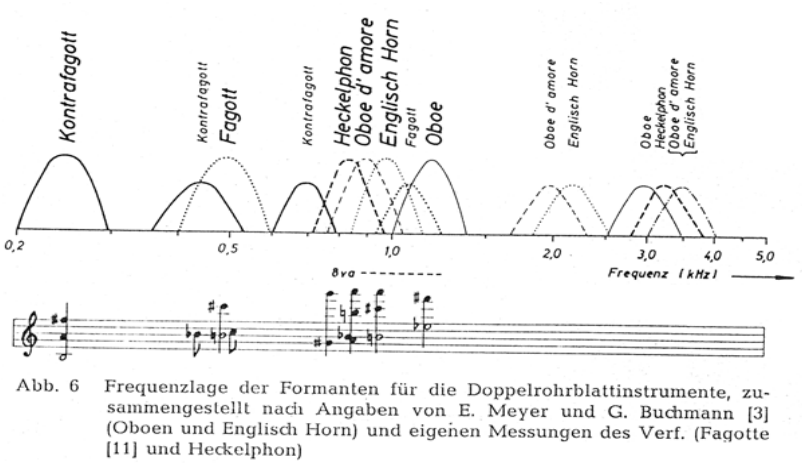


Abb. 6 Frequenzlage der Formanten für die Doppelrohrblattinstrumente, zusammengestellt nach Angaben von E. Meyer und G. Buchmann [3] (Oboen und Englisch Horn) und eigenen Messungen des Verf. (Fagotte [11] und Heckelphon)

Formants of double-reed instruments / 双簧乐器的共振峰
(Meyer 2015，第 63 页)

舒曼 1929：（**大部分**）管乐器频谱中的典型最大值。

- **与音高无关**（稳定）
- **像声**乐的共振峰（特征）

在德语领域，共振峰通常用于**描述**乐器的声音，基于它们的**元音特征**。

如果去掉音色中的共振峰（**或它**们之间的频谱间隙），音色**就不再是相**应乐器的典型音色。

A Timbre Space Odyssey

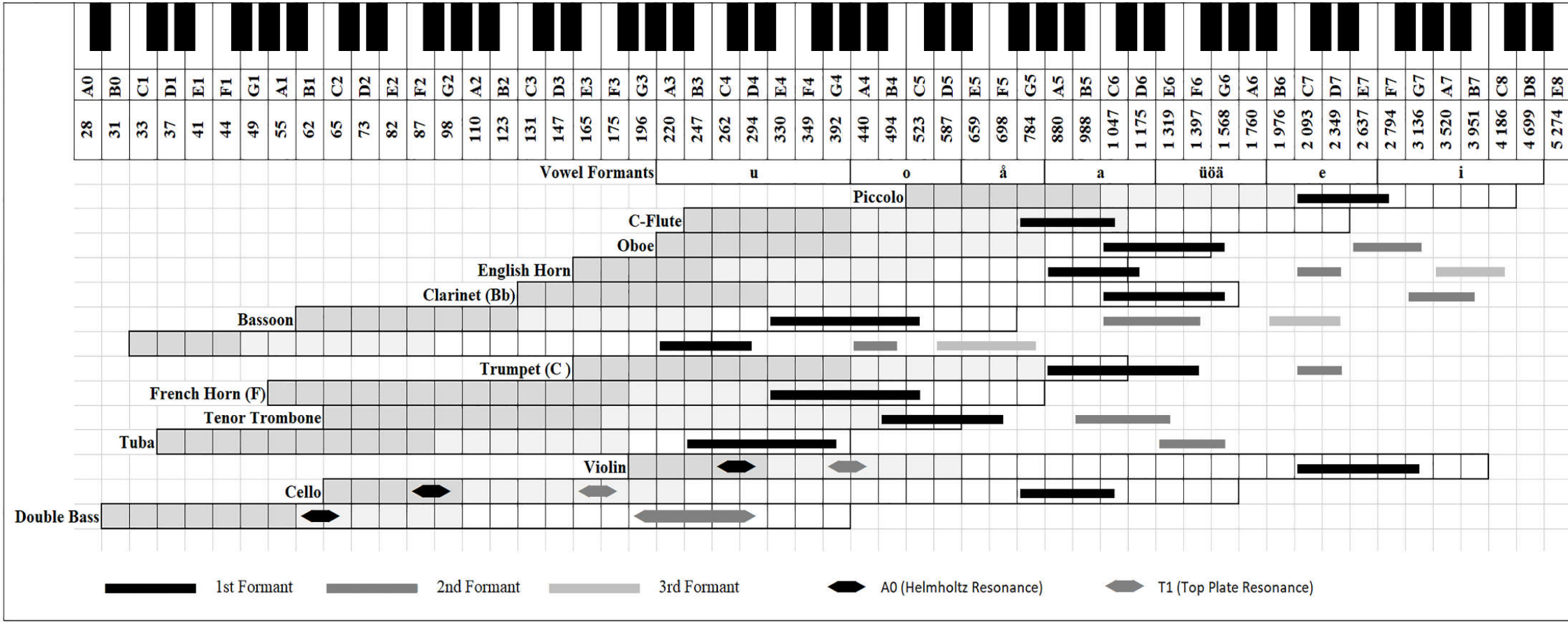
Popular Approaches of Timbre Visualisation

3 Formants

音色太空漫游

流行的音色可视化方法

3 共振峰



Formant areas and main resonances of various orchestral instruments / 各种管弦乐器的共振峰区和主要共鸣
(根据 Reuter 2014, 第 401 页; 音域边界根据 Siedenburg 等 2021 年，第 3717 页)

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

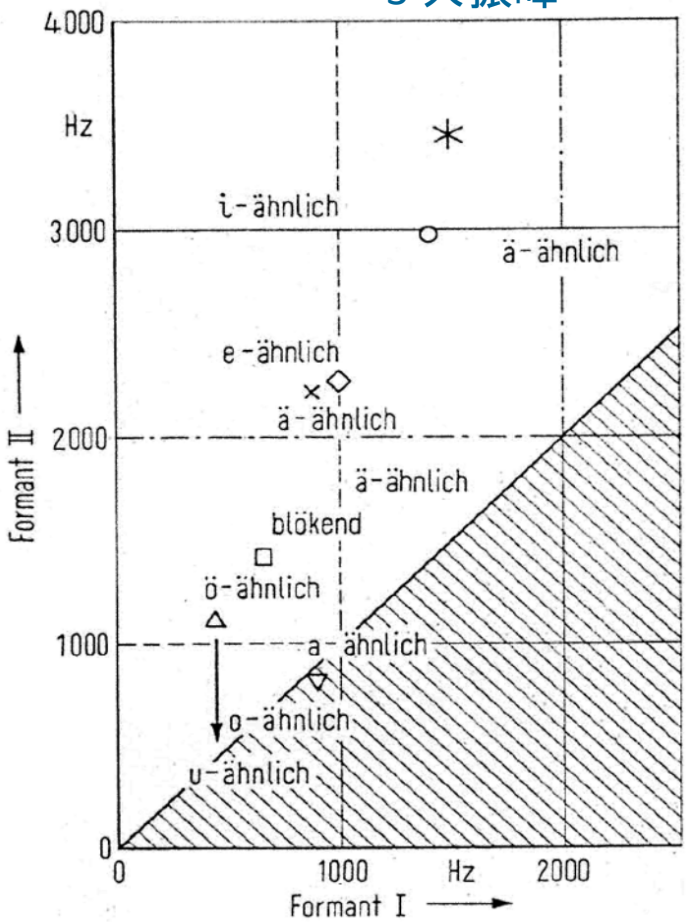
流行的音色可视化方法

3 Formants

Sirker 1974: First Formant Map: “Perceptual Space for two-formant Sounds“

Matching double reed instruments to vowels based on their first two formant areas (determined from the strongest amplitudes in their respective spectra).

3 共振峰



„Wahrnehmungsraum bei zweiformantigen Klängen“
„双共振峰声音的感知空间“ (Sirker 1974，第 52 页)

西尔克 1974：首个共振峰映射：“双共振峰声音的感知空间“

根据元音的前两个共振峰区（根据各自频谱中的最强振幅确定），将双簧乐器与元音进行匹配。

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

流行的音色可视化方法

3 Formants

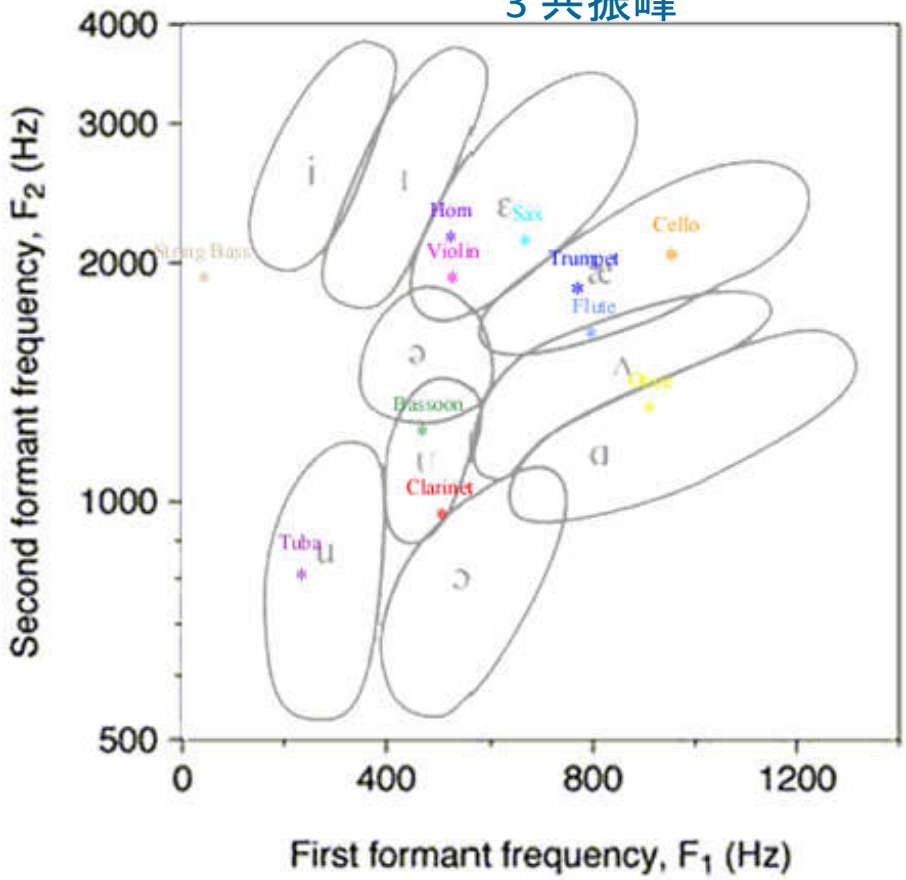
3 共振峰

McCarty 2003: Timbres of common orchestral instruments in a "vowel space"

- one pitch (different for each instrument)
- Formants estimated via Colea

"Now I can say that the tuba's timbre has a 'U' sound or (oo) as in the word who, or the trumpet has an 'ae' timbre like the work actor."

(McCarty, Stanford.edu 2003)



麦卡蒂 2003： 常见管弦乐器在"元音空间"中的音色

- 一个音高（每种乐器不同）
- 通过 Colea 预测共振峰

"现在我可以说，大号的音色像'who'一词中的'U'音或(oo)音，小号的音色像'actor'中的'ae'音"。

(麦卡蒂， Stanford.edu 2003）

Sounds of saxophone, flute, oboe, clarinet, bassoon, horn, trumpet, tuba, violin, cello, bass in the "Vowel Space".
萨克斯管、长笛、双簧管、单簧管、巴松管、圆号、小号、大号、小提琴、大提琴、贝司在 "元音空间 "中的声音
(麦卡蒂， Stanford.edu 2003)

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

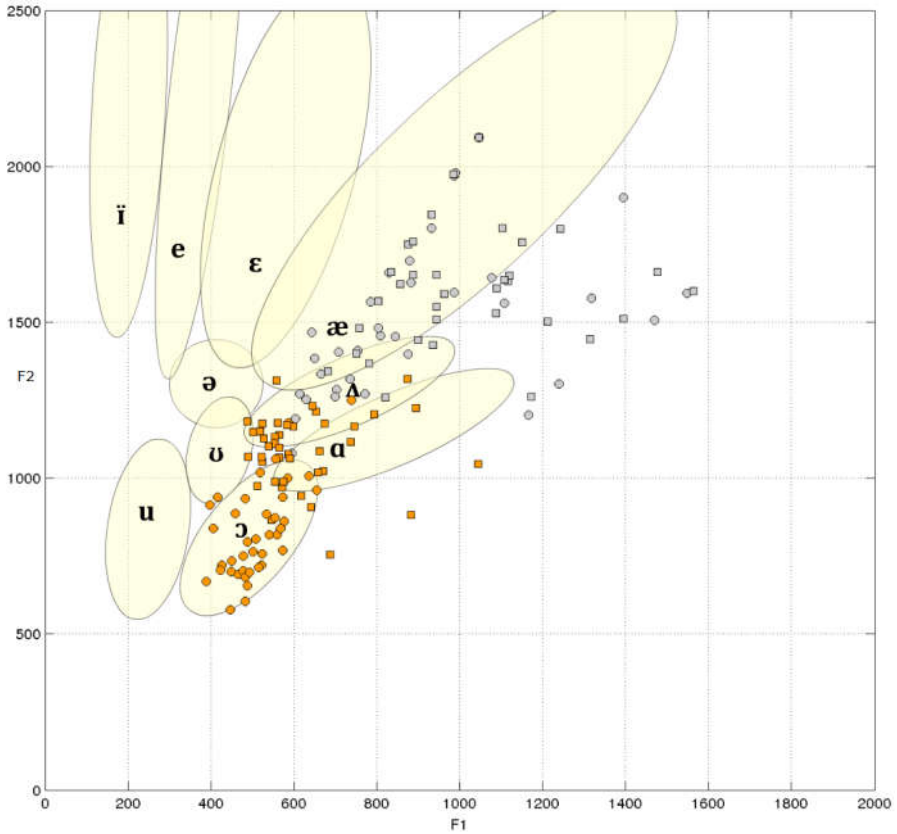
流行的音色可视化方法

3 Formants

3 共振峰

Reuter et al. 2016: Interactive Formant Map

- common woodwind and brass instruments from the VSL
- all reachable pitches
- two dynamic levels (***ff*** and ***pp***)
- 593 single sounds



罗伊特 等 2016： 交互式共振峰映射

- VSL 中常见的木管乐器和铜管乐器
- 所有可达到的音高
- 两个动态级别（***ff*** 和***pp***）
- 593 个声音

Interactive Formant Map with the sounds of bassoon (orange) and oboe (gray) in all reachable pitches in ***ff*** and ***pp***
巴松管（橙色）和双簧管（灰色）在ff和pp和所有可达到音高的交互式共振峰映射
(Reuter, Czedik-Eysenberg, Siddiq, & Oehler, 2017).

A Timbre Space Odyssey

音色太空漫游

Popular Approaches of Timbre Visualisation

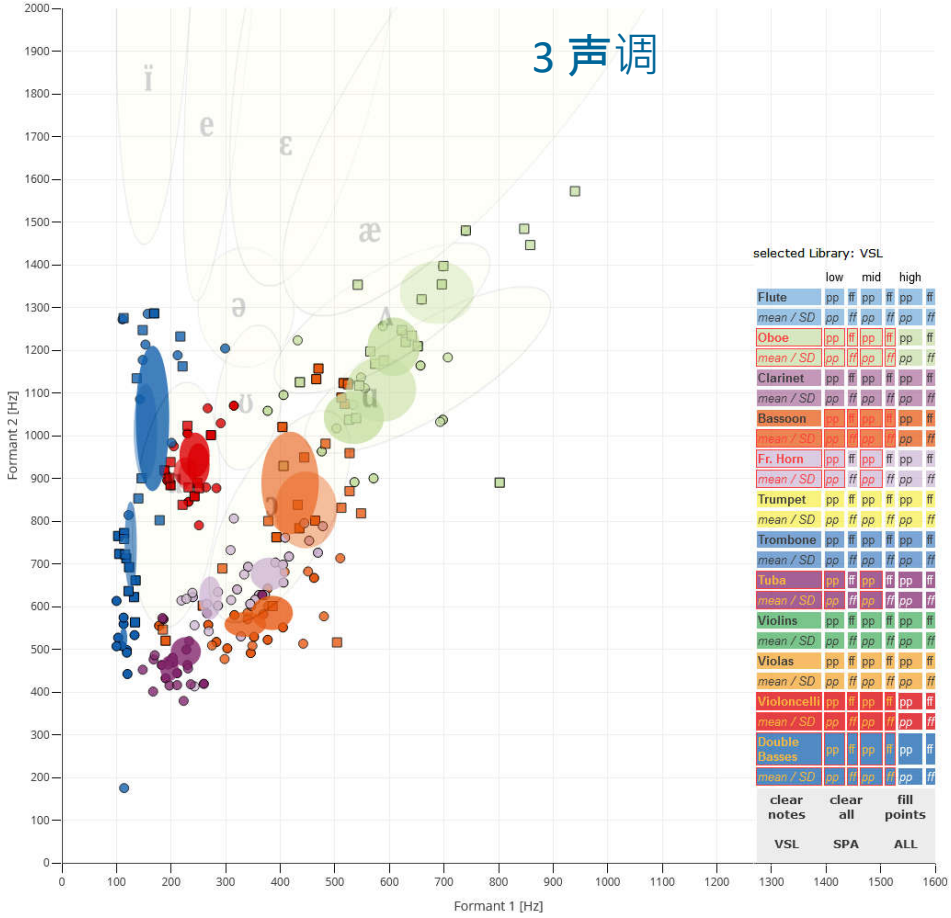
流行的音色可视化方法

3 Formants

3 声调

Reuter 2020: Update of the Interactive Formant Map:

- JavaScript (instead of Flash)
- Extension with **strings** (violin, viola, cello, bass)
- Additional Library: Spitfire Audio
- 1100 single sounds



罗伊特 2020：交互式共振峰映射的更新:

- JavaScript （取代 Flash）
- 扩展到弦乐（小提琴、中提琴、大提琴、低音提琴）
- 添加声音库：Spitfire Audio
- 1100 个声音

Formants and their mean values and standard deviations of oboe, bassoon, horn, tuba, cello and bass iin the low and middle registers.
双簧管、巴松管、圆号、大号、大提琴和低音提琴在中低音区的共振峰及其平均值和标准偏差。
(Reuter 2020)

A Timbre Space Odyssey 音色太空漫游

Odyssey
漫游

A Timbre Space Odyssey

Dynamic Timbre Maps

1 Dynamic Spectral Centroid/Spectral Flux Map

- Tracking the spectral centroid and spectral flux of all single instruments with the MIRToolbox
- Synchronization of audio and tracked features in an interactive timbre map via Plotly and P5.

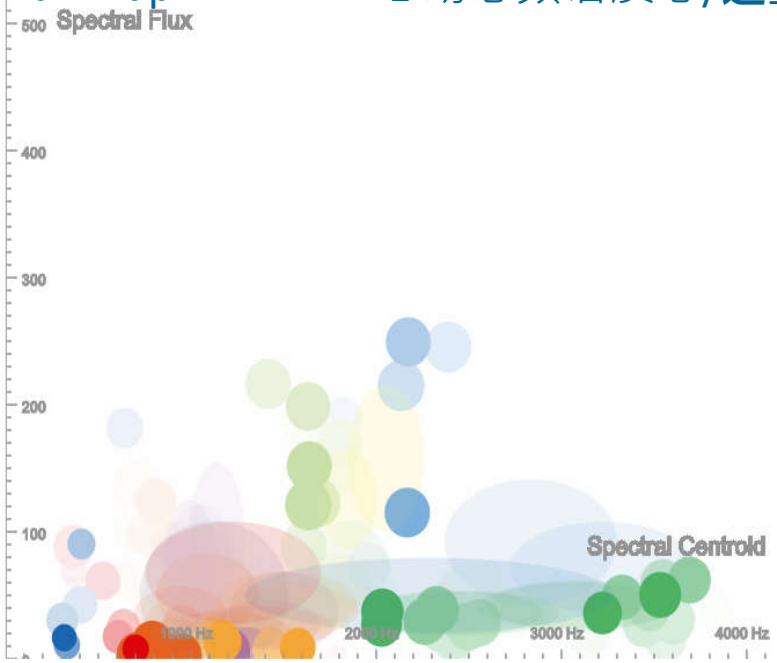
音色太空漫游

动态音色映射

Example: 1st movement of Beethoven's 7th symphony
示例：贝多芬第七交响曲第一乐章
(Recordings of the single tracks by Pätynen 等 2008 年的单轨录音)

1 动态频谱质心/通量映射

- 使用 MIRToolbox 跟踪所有单乐器的频谱质心和频谱通量
- 通过 Plotly 和 P5 在交互式音色图中同步音频和跟踪的特征。



Mean values and standard deviations of the orchestral instruments involved, here in interaction: flutes (light blue), oboes (light green), bassoons (orange), violins (green), violas (light orange), cello (red) and bass (blue)
相关管弦乐器的平均值和标准偏差（此处交互的是）：长笛（浅蓝色）、双簧管（浅绿色）、巴松管（橙色）、**小提琴**（绿色）、中提琴（浅橙色）、大提琴（红色）**和低音提琴**（蓝色）

A Timbre Space Odyssey

Dynamic Timbre Maps

1 Dynamic Spectral Centroid/Spectral Flux Map

- Tracking the spectral centroid and spectral flux of all single instruments with the MIRToolbox
- Synchronization of audio and tracked features in an interactive timbre map via Plotly and P5.

Behavior of instrumental timbre features can be visualized in "free wilderness":

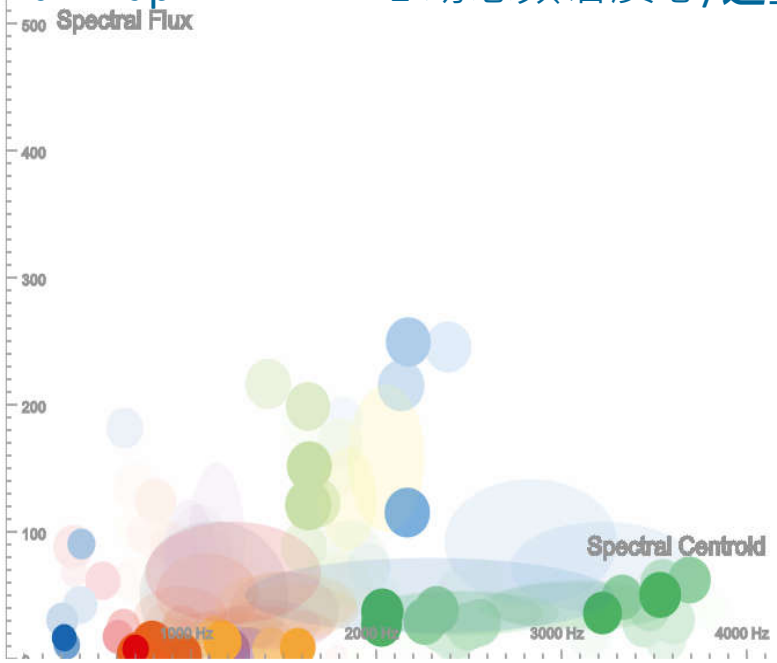
- Instrument timbres have their own territories.
- Instruments position depends on dynamic behaviour

音色太空漫游

动态音色映射

Example: 1st movement of Beethoven's 7th symphony
示例：贝多芬第七交响曲第一乐章
(Recordings of the single tracks by Pätynen 等 2008 年的单轨录音)

1 动态频谱质心/通量映射



Mean values and standard deviations of the orchestral instruments involved, here in interaction: flutes (light blue), oboes (light green), bassoons (orange), violins (green), violas (light orange), cello (red) and bass (blue)
相关管弦乐器的平均值和标准偏差（此处交互的是）：长笛（浅蓝色）、双簧管（浅绿色）、巴松管（橙色）、**小提琴**（绿色）、中提琴（浅橙色）、大提琴（红色）**和低音提琴**（蓝色）

- 使用 MIRToolbox 跟踪所有单乐器的频谱质心和频谱通量
- 通过 Plotly 和 P5 在交互式音色图中同步音频和跟踪的特征。

乐器音色特征的表现可以在“**自由原野**”**中直**观地表现出来：

- 乐器音色有**自己的**领域。
- 乐器的位置取决于动态表现

A Timbre Space Odyssey

Dynamic Timbre Maps

2 Dynamic MFCC Map

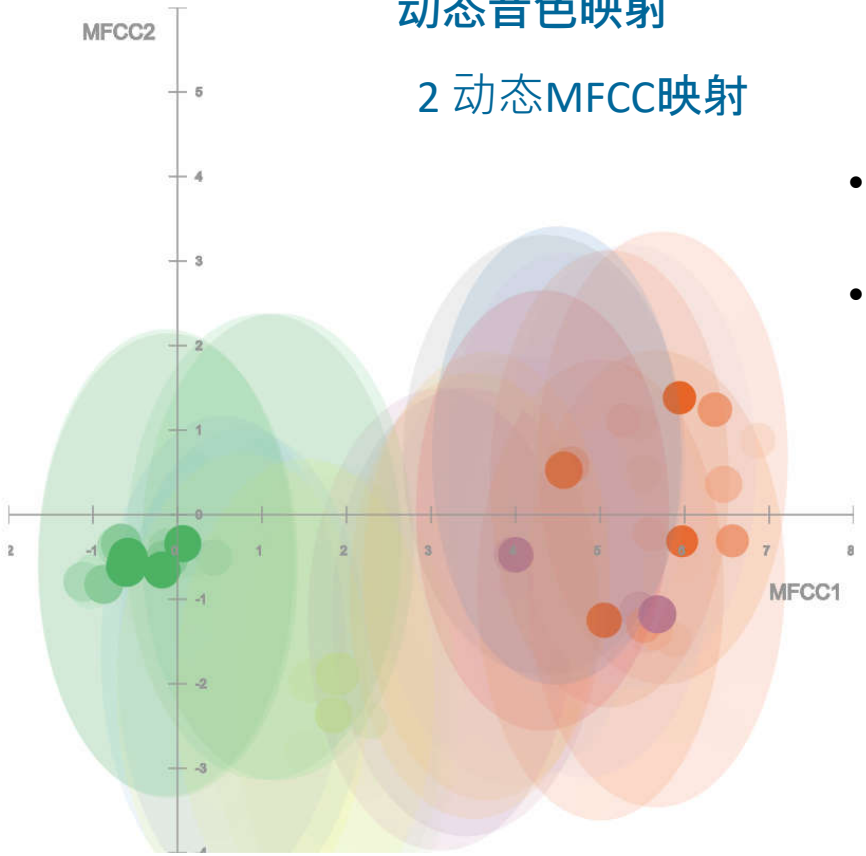
- Tracking the MFCC1 and 2 of all single instruments with the MIRToolbox
- Synchronization of audio and tracked MFCCs in an interactive timbre map via Plotly and P5.

音色太空漫游

动态音色映射

2 动态MFCC映射

Example: 1st movement of Beethoven's 7th symphony
示例：贝多芬第七交响曲第一乐章
(Recordings of the single tracks by Pätynen 等人 2008 年的单轨录音)



Mean values and standard deviations of the orchestral instruments involved, here in interaction: violins (green), horns (purple), bassoons (orange) and oboe (light green)
相关管弦乐器的平均值和标准偏差，此处交互的是：小**提琴**（绿色）、圆号（紫色）、巴松管（橙色）和双簧**管**（**浅**绿色）

- 使用 MIRToolbox 跟踪所有单一乐器的 MFCC1 和 2
- 通过 Plotly 和 P5 在交互式音色映射图中同步音频和跟踪的 MFCC。

A Timbre Space Odyssey

Dynamic Timbre Maps

2 Dynamic MFCC Map

- Tracking the MFCC1 and 2 of all single instruments with the MIRToolbox
- Synchronization of audio and tracked MFCCs in an interactive timbre map via Plotly and P5.

Behavior of instrumental MFCCs can be visualized in "free wilderness":

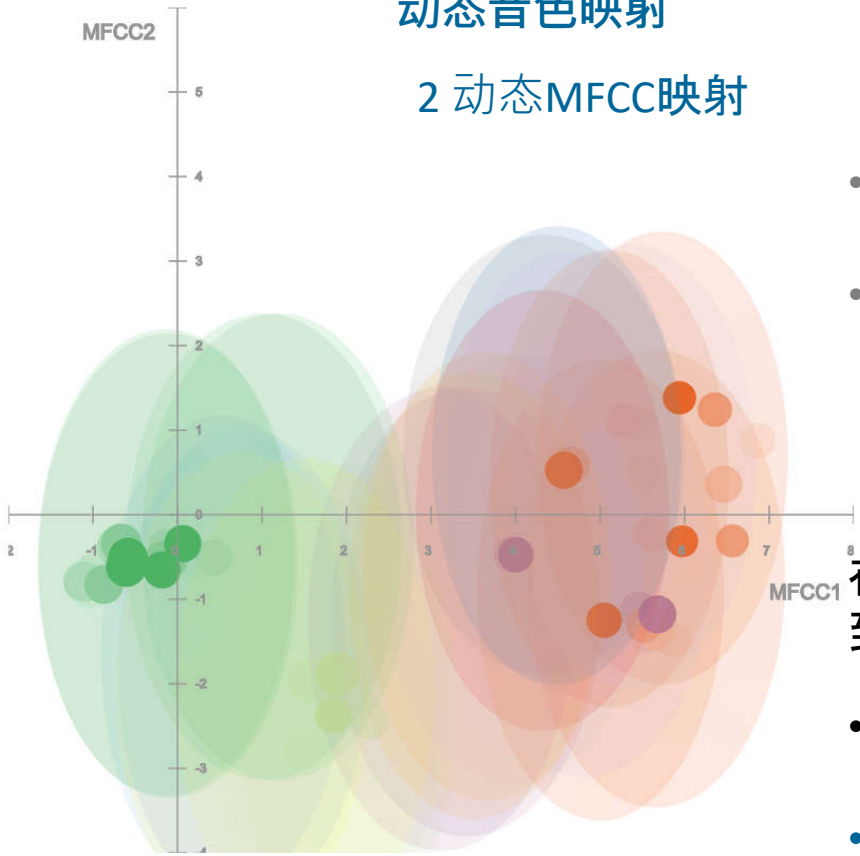
- Instrument timbres have their own territories.
- Similar timbres have overlapping MFCCs (next door territories).

音色太空漫游

动态音色映射

2 动态MFCC映射

Example: 1st movement of Beethoven's 7th symphony
示例：贝多芬第七交响曲第一乐章
(Recordings of the single tracks by Pätynen 等人 2008 年的单轨录音)



Mean values and standard deviations of the orchestral instruments involved, here in interaction: violins (green), horns (purple), bassoons (orange) and oboe (light green)
相关管弦乐器的平均值和标准偏差，此处交互的是：小提琴（绿色）、圆号（紫色）、巴松管（橙色）和双簧管（浅绿色）

- 使用 MIRToolbox 跟踪所有单一乐器的 MFCC1 和 2
- 通过 Plotly 和 P5 在交互式音色映射图中同步音频和跟踪的 MFCC。

在“自由原野”中可以直观地看到乐器MFCC的表现：

- 乐器音色有自己的领域。
- 相似的音色有重叠的 MFCC（相邻领域）。

A Timbre Space Odyssey

Dynamic Timbre Maps

3 Dynamic Formant Map

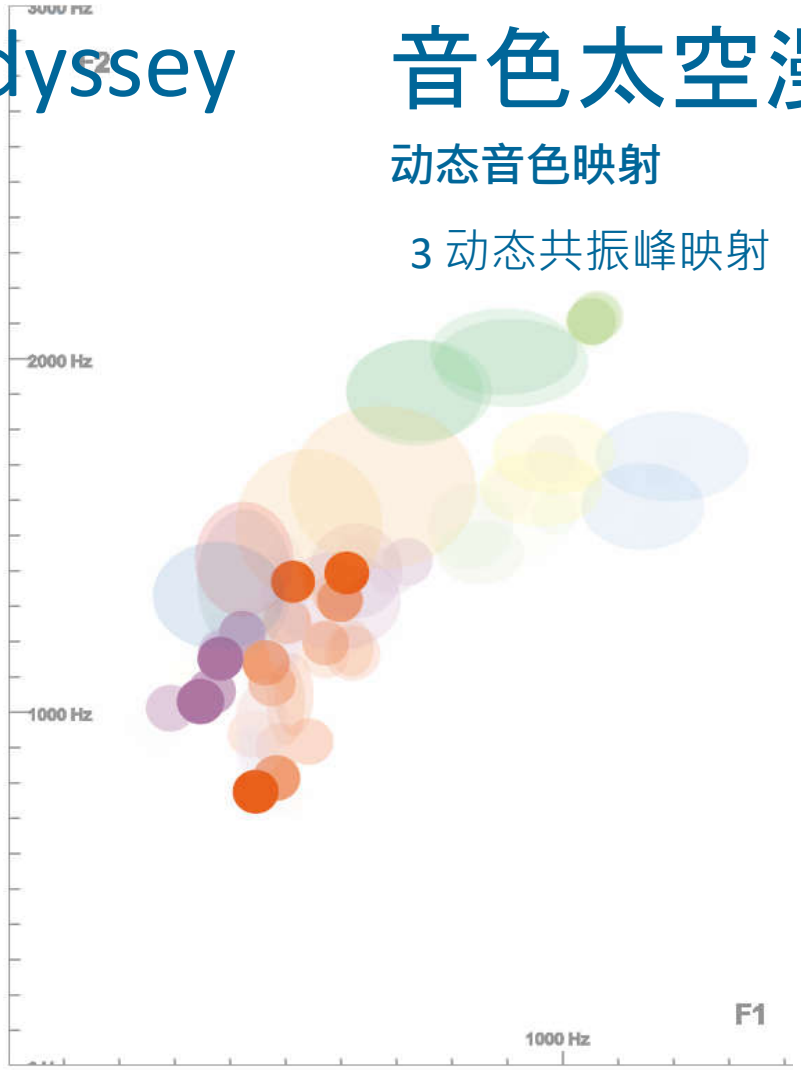
- Tracking the formants 1 and 2 of all single instruments with Parselmouth
- Synchronization of audio and tracked Formants in an interactive timbre map via Plotly and P5.

音色太空漫游

动态音色映射

3 动态共振峰映射

Example: 1st movement of Beethoven's 7th symphony
示例：贝多芬第七交响曲第一乐章
(Recordings of the single tracks by Pätynen 等人 2008 年的单轨录音)



Mean values and standard deviations of the orchestral instruments involved, here in interaction: Horns (purple), bassoons (orange) and oboe (green)
相关管弦乐器的平均值和标准偏差，此处交互的是：圆号（紫色）、巴松管（橙色）和双簧管（绿色）

- 使用 Parselmouth 追踪所有单件乐器的**第一和第二共振峰**
- 通过 Plotly 和 P5 在交互式音色映射图中**同步音**频和跟踪**的共振峰**。

A Timbre Space Odyssey

Dynamic Timbre Maps

3 Dynamic Formant Map

- Tracking the formants 1 and 2 of all single instruments with Parselmouth
- Synchronization of audio and tracked Formants in an interactive timbre map via Plotly and P5.

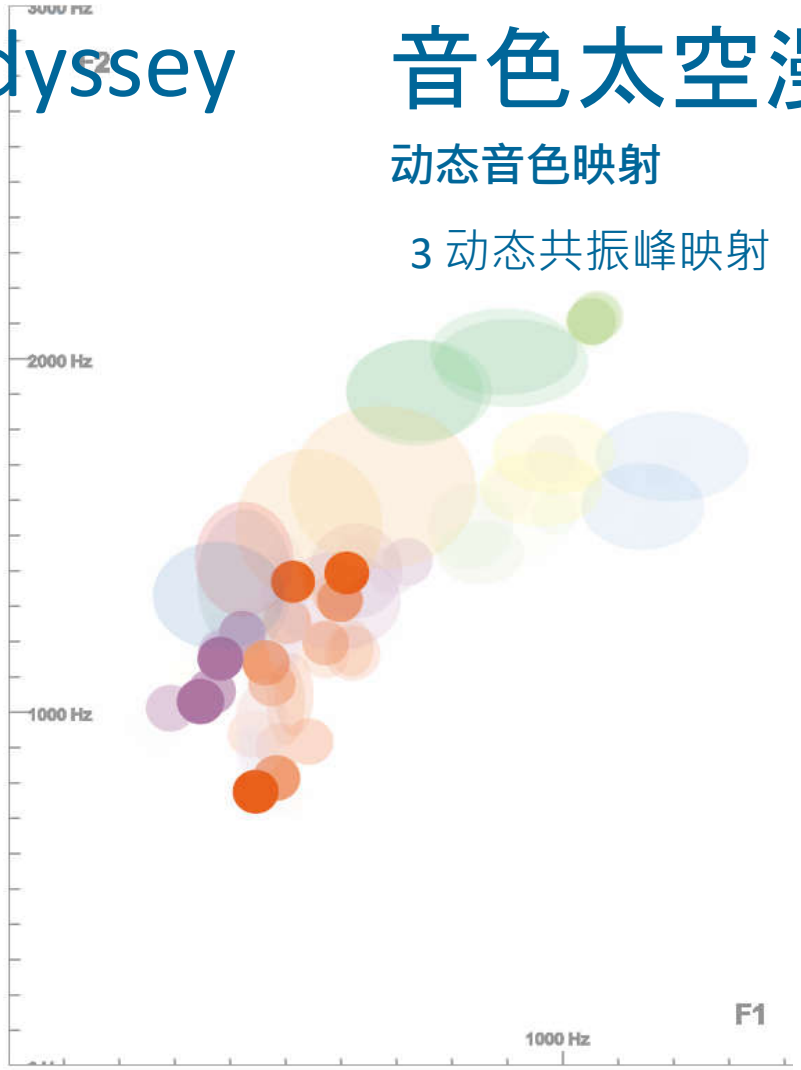
Behavior of instrumental formants can be visualized in "free wilderness":

- Instrument timbres have their own territories.
- Similar timbres have overlapping formants (next door territories).

音色太空漫游

动态音色映射

3 动态共振峰映射



Mean values and standard deviations of the orchestral instruments involved, here in interaction: Horns (purple), bassoons (orange) and oboe (green)
相关管弦乐器的平均值和标准偏差，此处交互的是：圆号（紫色）、巴松管（橙色）和双簧管（绿色）

Example: 1st movement of Beethoven's 7th symphony
示例：贝多芬第七交响曲第一乐章
(Recordings of the single tracks by Pätynen 等人 2008 年的单轨录音)

- 使用 Parselmouth 追踪所有单件乐器的第一和第二共振峰
- 通过 Plotly 和 P5 在交互式音色映射图中同步音频和跟踪的共振峰。

在“自由原野”中可以直观地看到器乐共振峰的表现：

- 乐器音色有自己的领域。
- 相似的音色有重叠的共振峰（相邻领域）。

A Timbre Space Odyssey 音色太空漫游

Back on Earth
回到地球

A Timbre Space Odyssey

Take Home Message

There seems to be no proper definition for timbre at all. Timbre can be described by its spectrum and/or by a bunch of acoustical features.

音色太空漫游

带回家的信息

对于音色，**似乎根本**没有一个恰当的定义。
音色可以用频谱和/或一系列声学特征来描述。

A Timbre Space Odyssey

Take Home Message

There seems to be no proper definition for timbre at all. Timbre can be described by its spectrum and/or by a bunch of acoustical features.

Visualisations in Timbre Spaces:
- **Pros:** intuitive, very popular
- **Cons:** not comparable, not generalizable, only one pitch

音色太空漫游

带回家的信息

对于音色，似乎根本没有一个恰当的定义。
音色可以用频谱和/或一系列声学特征来描述。

音色空间的可视化：
- 优点：直观，非常受欢迎
- 缺点：不具可比性、不具通用性、只有一个音高

A Timbre Space Odyssey

Take Home Message

There seems to be no proper definition for timbre at all. Timbre can be described by its spectrum and/or by a bunch of acoustical features.

Visualisations in Timbre Spaces:
- **Pros:** intuitive, very popular
- **Cons:** not comparable, not generalizable, only one pitch

Visualisations in MFCC Feature Spaces:
- **Pros:** very precise, very universally applicable.
- **Cons:** it is difficult to say what a single MFCC sounds like

音色太空漫游

带回家的信息

对于音色，似乎根本没有一个恰当的定义。
音色可以用频谱和/或一系列声学特征来描述。

音色空间的可视化：
- 优点：直观，非常受欢迎
- 缺点：不具可比性、不具通用性、只有一个音高

MFCC 特征空间的可视化：
- 优点：非常精确，非常普遍适用。
- 缺点：很难说清楚单个 MFCC 听起来像什么

A Timbre Space Odyssey

Take Home Message

There seems to be no proper definition for timbre at all.
Timbre can be described by its spectrum and/or by a bunch of acoustical features.

Visualisations in Timbre Spaces:
- **Pros:** intuitive, very popular
- **Cons:** not comparable, not generalizable, only one pitch

Visualisations in MFCC Feature Spaces:
- **Pros:** very precise, very universally applicable.
- **Cons:** it is difficult to say what a single MFCC sounds like

Visualisations in Formant Maps:
- **Pros:** intuitive, sound imagination possible by comparison with vowels
- **Cons:** does not work for every instrument / register

音色太空漫游

带回家的信息

对于音色，似乎根本没有一个恰当的定义。
音色可以用频谱和/或一系列声学特征来描述。

音色空间的可视化：
- 优点：直观，非常受欢迎
- 缺点：不具可比性、不具通用性、只有一个音高

MFCC 特征空间的可视化：
- 优点：非常精确，非常普遍适用。
- 缺点：很难说清楚单个 MFCC 听起来像什么

共振峰映射的可视化：
- 优点：直观，可通过与元音对比进行声音想象
- 缺点：不适用于所有乐器和音域

A Timbre Space Odyssey

Take Home Message

There seems to be no proper definition for timbre at all. Timbre can be described by its spectrum and/or by a bunch of acoustical features.

Visualisations in Timbre Spaces:
- **Pros:** intuitive, very popular
- **Cons:** not comparable, not generalizable, only one pitch

Visualisations in MFCC Feature Spaces:
- **Pros:** very precise, very universally applicable.
- **Cons:** it is difficult to say what a single MFCC sounds like

Visualisations in Formant Maps:
- **Pros:** intuitive, sound imagination possible by comparison with vowels
- **Cons:** does not work for every instrument / register

Visualisation in Dynamic Maps:
- Real-life behavior of timbre features in musical interaction becomes visible.
- Instruments show their own territories with similar sounding instrument timbres in their neighborhood

音色太空漫游

带回家的信息

对于音色，似乎根本没有一个恰当的定义。
音色可以用频谱和/或一系列声学特征来描述。

音色空间的可视化：
- 优点：直观，非常受欢迎
- 缺点：不具可比性、不具通用性、只有一个音高

MFCC 特征空间的可视化：
- 优点：非常精确，非常普遍适用。
- 缺点：很难说清楚单个 MFCC 听起来像什么

共振峰映射的可视化：
- 优点：直观，可通过与元音对比进行声音想象
- 缺点：不适用于所有乐器和音域

动态映射可视化：
- 音乐交互中音色特征的真实表现变得更清晰。
- 乐器显示出自己的领域，并在其附近显示类似的乐器音色

谢谢

A Timbre Space Odyssey　　音色太空漫游

Timbre Spaces / 音色空间

Albersheim 1939: Acoustical Cylinder of Sound colors / 阿尔伯斯海姆 1939 年 声音颜色的音响圆柱体： https://muwiserver.synology.me/p5/fft_filter_albersheim.htm
Grey 1975: Timbre Space / 格雷 1975 音色空间 : https://muwiserver.synology.me/timbrespaces/grey.htm
Krumhansl 1989, McAdams et al. 1995: Timbre Spaces / Krumhansl 1989，McAdams 等 1995：音色空间 : https://muwiserver.synology.me/formanten/more_timbre_spaces.htm
Siddiq et al. 2015 Meta Timbre Space / Siddiq 等 2015 元音色空间 : https://muwiserver.synology.me/timbrespaces/TS_mts3.htm
Reuter, Quigley 2020: Dove perch and courtship coos / 罗伊特，奎格利 2020：鸽子栖息和求偶鸣叫 : https://muwiserver.synology.me/doves/doves2_sequence.htm
Bassoon without formants / 巴松管，无共振峰 : https://muwiserver.synology.me/formanten/bassoonflattened.htm

Formant Maps / 方音图

Sirker 1974: Perceptual Space for two-formant Sounds / 西尔克 1974 双共振峰声音的感知空间 : https://muwiserver.synology.me/formanten/formantkarte_instrumente.htm
McCarty 2003: vowel space / McCarty 2003：元音空间 : https://ccrma.stanford.edu/~jmccarty/formant.htm
Reuter et al. 2016: Formant Map 1.0 / Reuter 等 2016 ： 共振峰映射 1.0 : https://muwiserver.synology.me/formanten/formantregister4_english.htm
Reuter 2020: Formant Map 2.0 / 罗伊特 2020 共振峰映射 2.0 : https://muwiserver.synology.me/formantmap/
Reuter 2022: Dynamic Formant Map / 罗伊特 2022：动态共振峰映射 : https://muwiserver.synology.me/dynamic/formants.htm
Reuter 2022: Dynamic MFCC Map / 罗伊特 2022：动态 MFCC 映射 : https://muwiserver.synology.me/dynamic/mfcc.htm
Reuter 2022: Dynamic SpectralCentroid/Flux Map /路透 2022：动态频谱质心/通量映射 : https://muwiserver.synology.me/dynamic/spectralcentroid.htm

Audio Signal Analysis Libraries/Toolboxes / 音频信号分析库/工具箱

Lartillot et al. 2007: MIRToolbox / Lartillot 等 2007： MIRToolbox : https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox
Peeters et al. 2011: Timbre Toolbox / Peeters 等 2011：Timbre Toolbox : https://github.com/VincentPerreault0/timbretoolbox
Bogdanov et al. 2013: Essentia / Bogdanov 等 2013： Essentia: https://essentia.upf.edu/
Yaafe 2013: Yaafe / Yaafe 2013： Yaafe: https://github.com/Yaafe/Yaafe
Rawlinson 2014: Meyda / 罗林森 2014：Meyda : https://meyda.js.org/
McFee et al. 2015: Librosa / McFee 等 2015： Librosa: https://librosa.org/
Jillings 2016: JS-xtract / 吉林斯 2016： JS-xtract：https://github.com/nickjillings/js-xtract
Aubio 2017: Aubio / Aubio 2017 Aubio：https://aubio.org/
Lartillot 2019: MiningSuite / Lartillot 2019：MiningSuite： http://olivierlar.github.io/miningsuite/

Formant Analysis Libraries/Toolboxes / 方音分析库/工具箱

Loizou 1998: Colea / Loizou 1998： Colea: https://ecs.utdallas.edu/loizou/speech/colea.htm
Boersma, Weenink, since 2001: Praat / Boersma、Weenink 2001 年起：Praat : https://www.fon.hum.uva.nl/praat/
Jadoul et al. 2018: Parselmouth / Jadoul et al. 2018: Parselmouth : https://github.com/YannickJadoul/Parselmouth
Kamath 2021: Formant Tracker / 卡马斯 2021 共振峰跟踪器 : https://de.mathworks.com/matlabcentral/fileexchange/8959-formant-tracker
Rabiner et al. 2021: Formant Estimation / 拉比纳等 2021：共振峰预测 : https://de.mathworks.com/matlabcentral/fileexchange/45315-formant-estimation