



# 音楽情報処理技術とエンタテインメント

片寄晴弘, 森勢将雅

関西学院大学  
CREST/JST

# アウトライン

1. CrestMuseプロジェクト概要
2. 具体例（ビデオ+デモ）

# 背景

## 音楽コンテンツ制作に関するエピソード (1984年)

“Ghost Busters” / Ray Parker Jr.

“I Want a New Drug” / Huey Lewis and the News



イメージの伝達に事例が用いられた！

「ビートルズのあの編曲」

「スタンリー・キューブリック後期作品のシーン展開」…

イメージの伝達：事例 > 言語(形容詞)



# 背景

## アート・デザイン

- **創造**・自己表現, 感動, エンタテインメント
- 創造性 vs. 受容されること

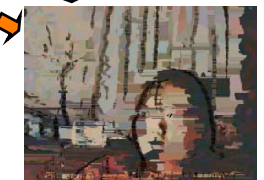
デザインの大部分は**共通理解**が前提 → **既存事例**

## メディアデザインのための工学的アプローチ 感性工学

### 本研究：

#### 既存事例の活用 (not 感性工学)

- 時系列メディアの**認知的性質**に着目・利用
- デザイン上の意図的な逸脱(= **elaboration**)の**操作, 転写方式**に焦点



工芸大  
笠尾氏提供






# 注目する音楽の性質

## 1. 記述・記録に関する歴史的積み上げ

- 音波, 楽譜, MIDI (マルチレイヤ)
  - 人間のための音楽理論の存在
- 他の芸術と比べての計算機科学との親和性：高

## 2. 時系列メディア芸術



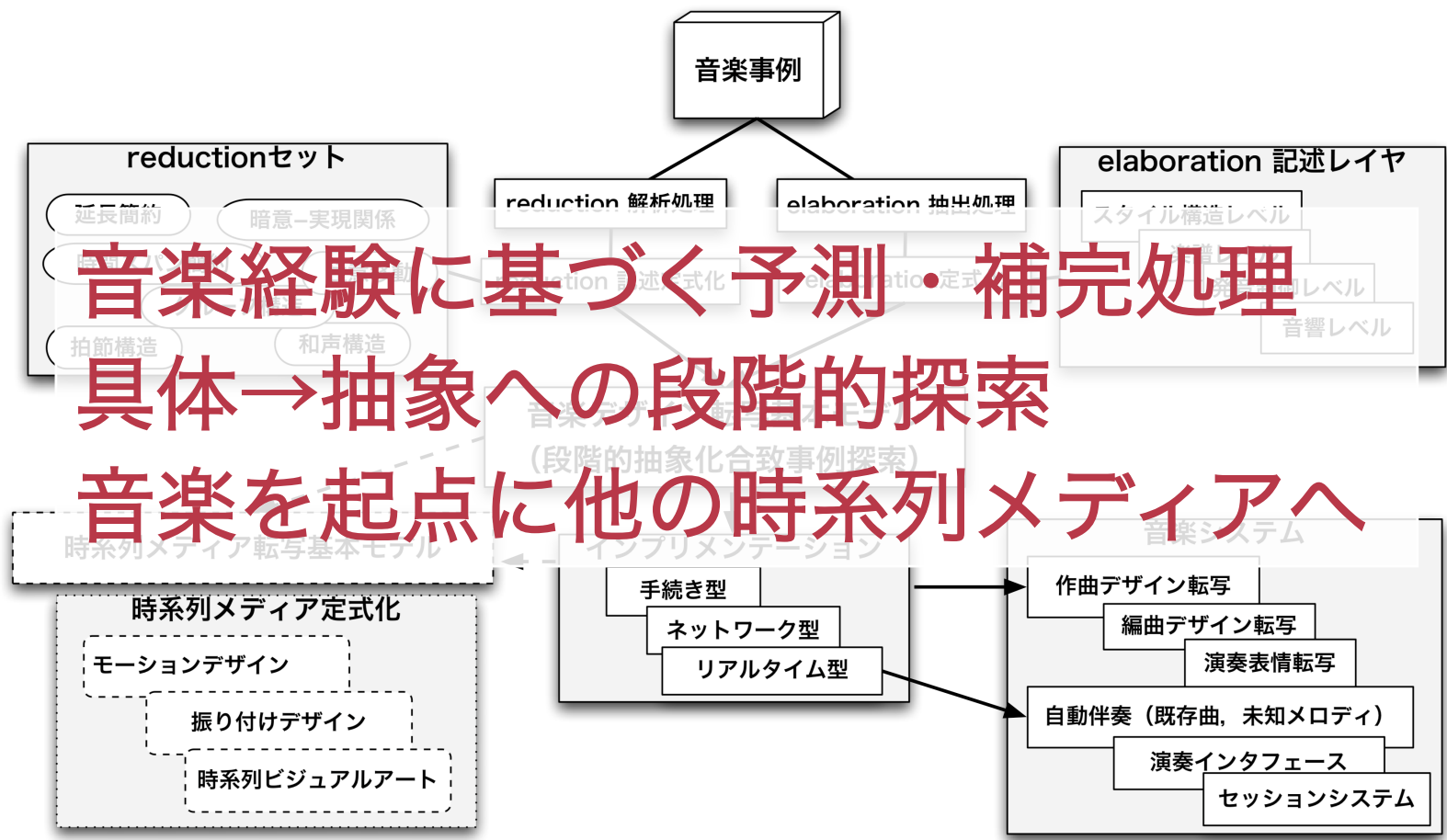
- 進行に対する予測性・因果律が基礎  
- 高次認知構造‘ (伴奏等) を伴って予測性向上 

cf. カラオケの演歌の初見歌唱しやすさ

コラールの予測モデル(Conklin, Witten 95)



# 時系列メディアの デザイン転写アーキテクチャ



(片寄他 人工知能学会誌 2005.3)

# ポイント（起案時）

メディア操作技術

コンテンツ(事例)

- デザイン支援（コンテンツ制作の効率化）
- 能動的音楽鑑賞（深く音楽を識る）
- 新しい音楽の愉しみ（新エンタテインメント）

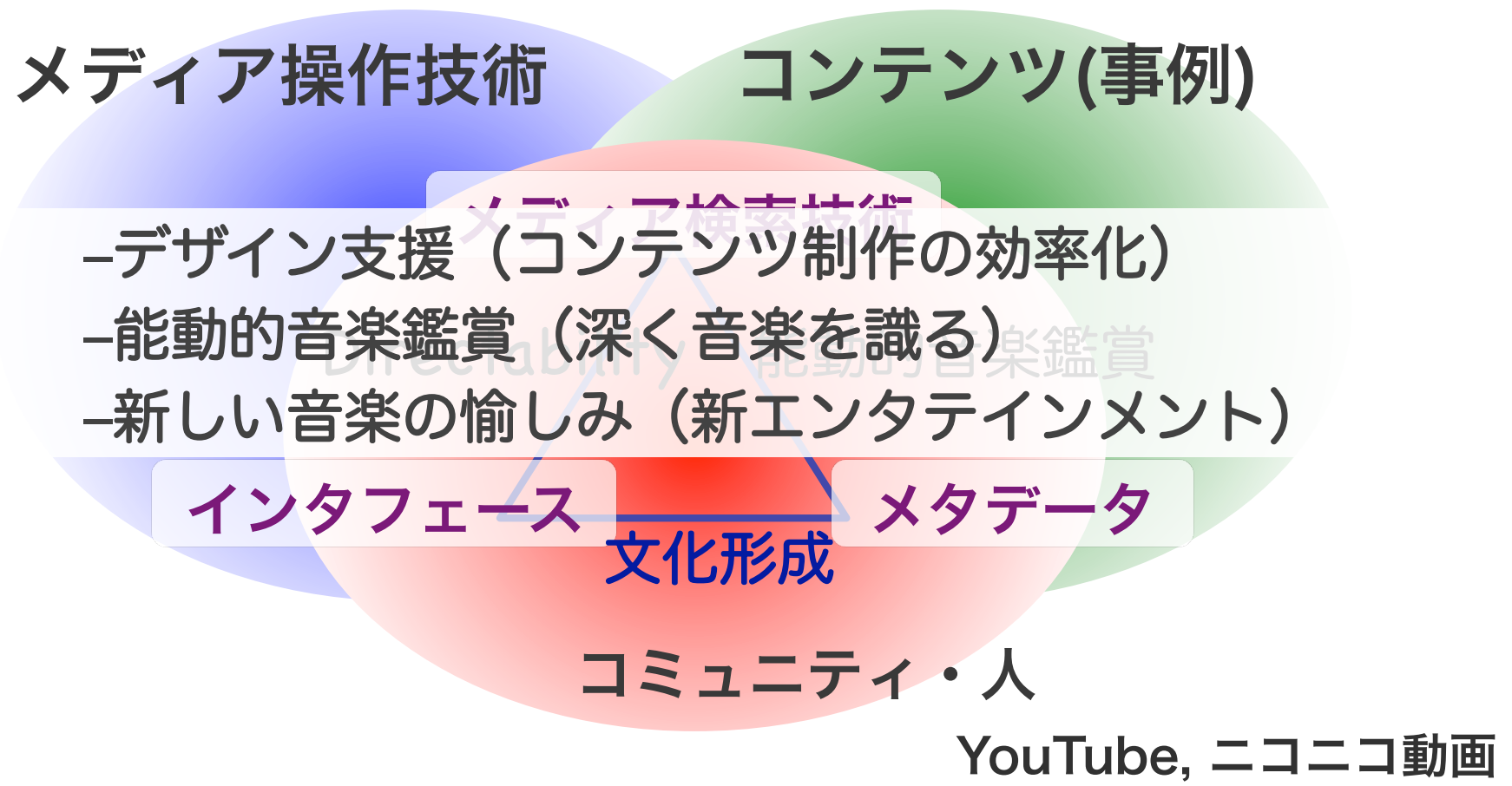
面白みに欠ける！？

**+ Directability**

人の知覚・認知に適合したより抽象度の高い操作子を用  
いての事例の制御

# とりくみの発展

## 事例参照 + Directability



# アウトライン

1. CrestMuseプロジェクト概要
2. 具体例（ビデオ+デモ）

# Video

## CrestMuse プロジェクトシンポジウム 2006

「MusicRainbow: A New User Interface to Discover Artists」

「Musiccream: 楽曲を流してくっつけて並べる音楽再生インタフェース」

「歌真似と声真似, STRAIGHTのモーフィングは, どこまで人間に迫るか?」

「STRAIGHTの原理とrealtime STRAIGHT, そして, 応用へ

「音楽音響信号をMIDIデータへ変換する」

「MIDIデータを楽譜へ変換する」

「自動伴奏システム」

「Drumix:音楽CDのドラムパート編集機能付きオーディオプレイヤー」

「Instrogramによる類似楽曲検索」

「混合音中の歌声と歌詞との時間的対応付け」

「混合音中のパートの音量を操作可能なオーディオプレイヤー」

「名演奏追体験インタフェース: iFP2」

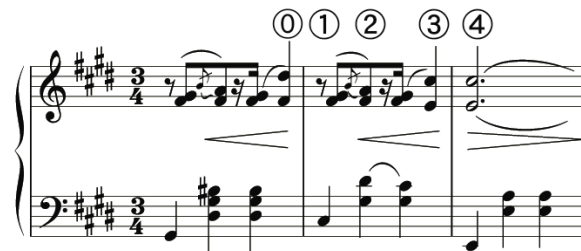
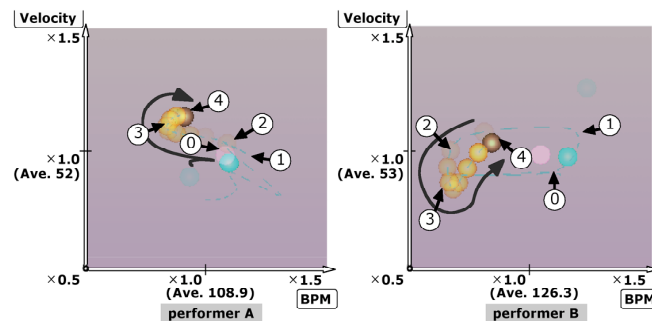
「音楽を体験している脳をはかる」

「表情付けシステム: Pop-E」

# Demo その1



## 演奏表情テンプレートを利用した 指揮システム／打鍵型演奏インタフェース (片寄, 奥平 2004)



### 特徴

- ・意図のマージ (ピアニストの指揮感覚)
- ・操作対象
  - ・テンポ
  - ・ダイナミクス
  - ・拍内表情
- ・モーフィング
- ・デモ
- ・没入感 cf. 音ゲーム

# Demo その2

## STRAIGHT を利用した歌唱モーフィングインタフェース (豊田, 河原, 片寄 2006~)

声マネと歌マネ

STRAIGHT のモーフィングはどこまで人間に迫るか?  
方眼上の点は STRAIGHT によるモーフィング, 黄色のパネルは人間による声真似と歌真似です。

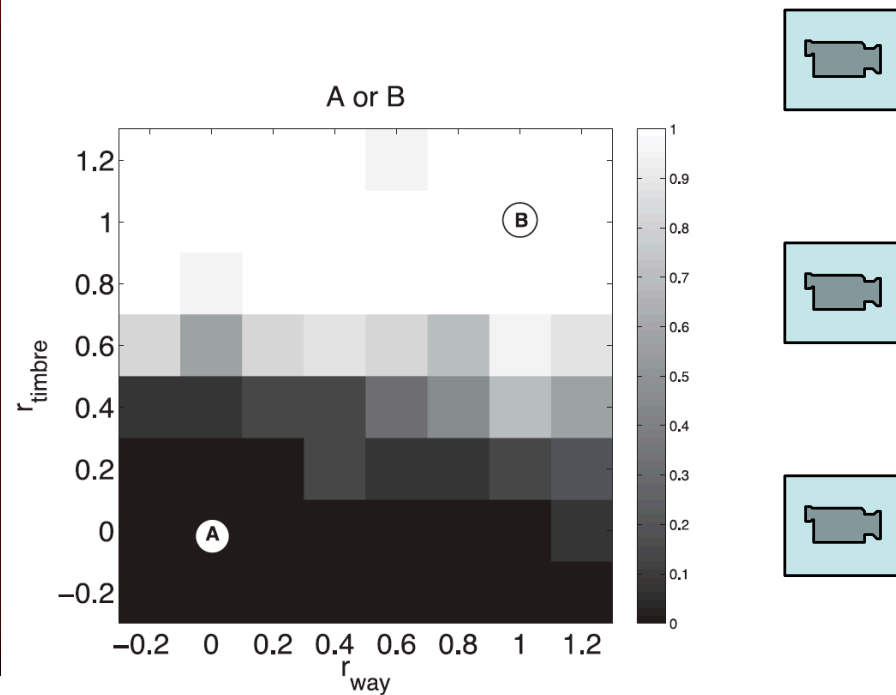
Please click on a grid point

Cursor : ( 0.00, 0.00)  
Ratio : ( , )

人による声マネ

人による歌マネ

The screenshot shows a software interface with a grid. The vertical axis is labeled '声質' (voice quality) and the horizontal axis is labeled '歌い回し' (singing style). A cursor is positioned at (0.00, 0.00). There are four speaker icons on the grid. A yellow box labeled 'Voice' is on the left and a yellow box labeled 'Style' is on the right. A speech bubble on the left says '人による声マネ' and a speech bubble on the right says '人による歌マネ'.

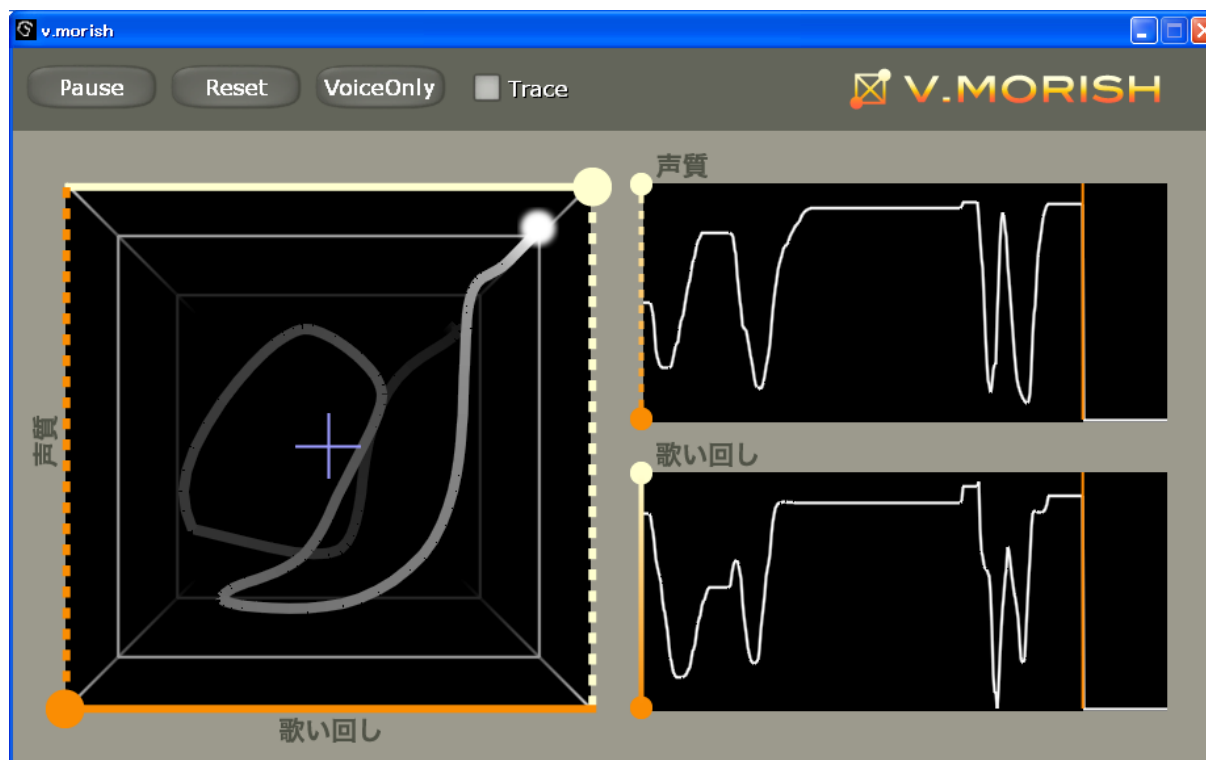






# 歌唱モーフィングインタフェース v.morishデモとその原理

# 歌声を混ぜるインタフェース v.morish



- ・ マウスを使ったリアルタイム歌唱変換
- ・ 右図を使った特徴の調整

# v.morishにできること

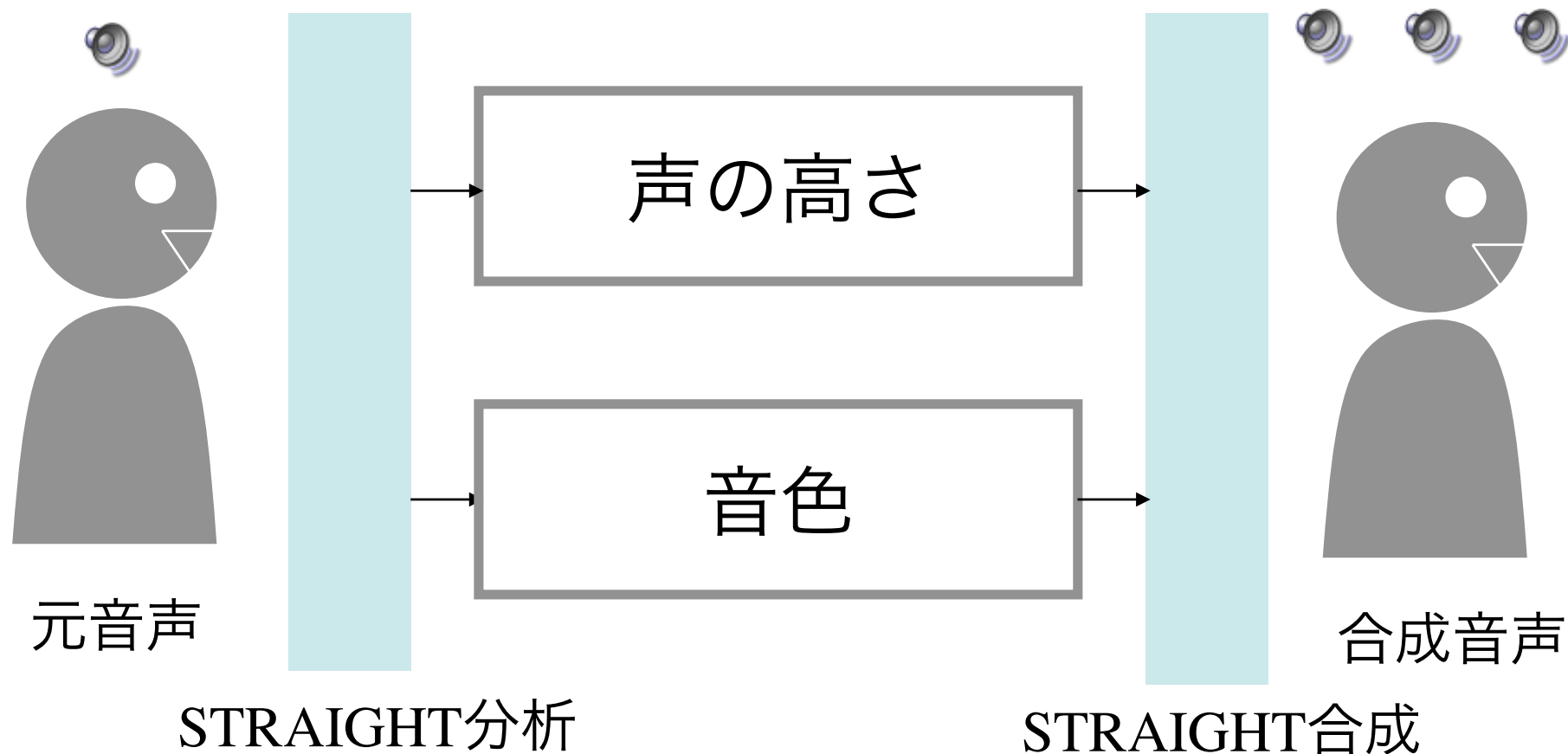
2つの歌声の特徴を混ぜ合わせる。

- 自分の歌にプロの特徴を転写
- 自分の声質に別人の声質を加える



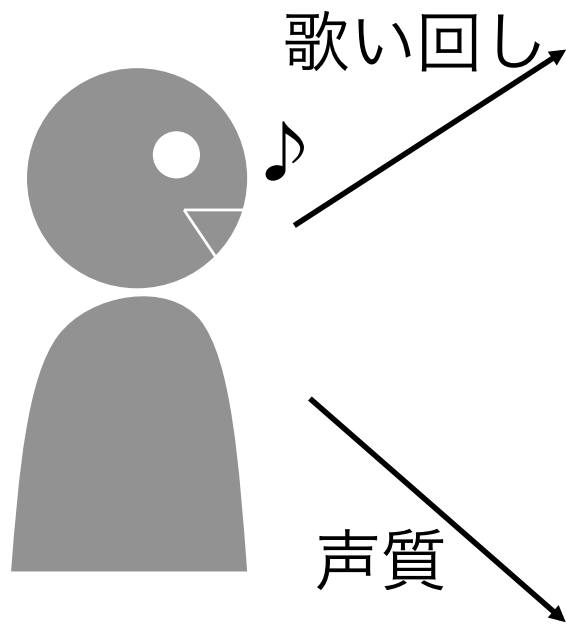
歌手の歌い方を事例として蓄積することにより、  
クリエイターが行う歌声の作り込みを支援

# 音声モーフィングを実現する技術 STRAIGHT



- ・ 3つのパラメタを独立して操作可能

# 声質・歌い回し



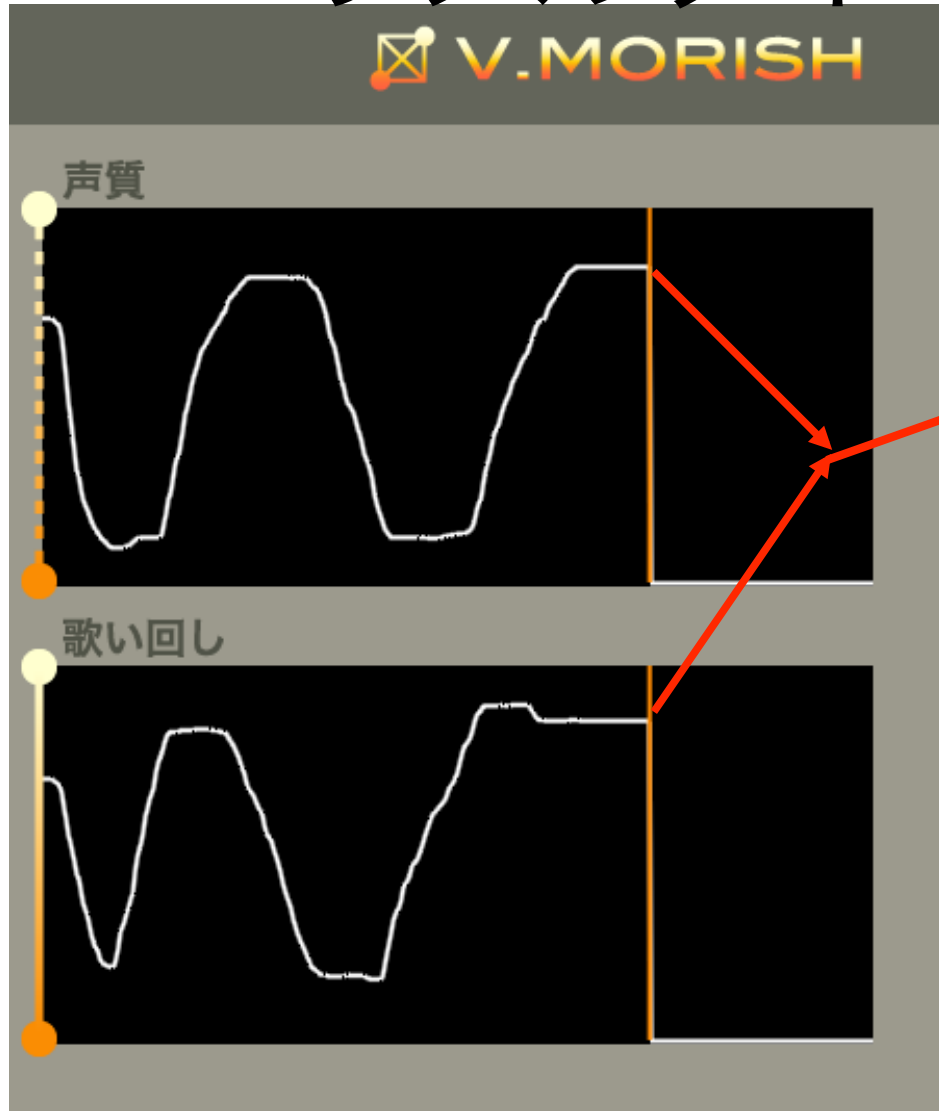
声の高さ	ビブラート等
時間情報	タイミング等

音色	歌手の個人性
----	--------

音色を変える歌唱表現もあるが、今回は考慮しない

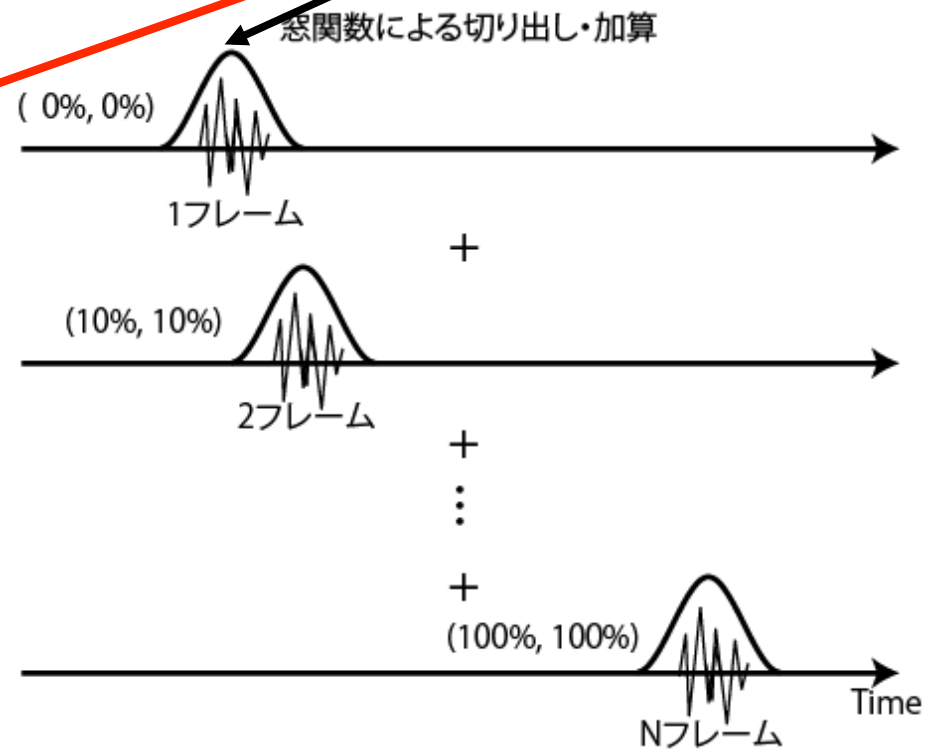
歌唱を声質と歌い回しに分離

# リアルタイム合成の原理

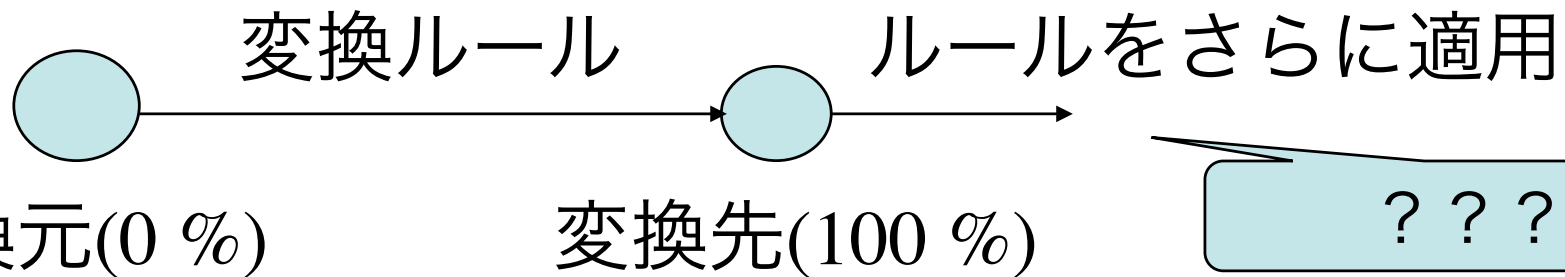
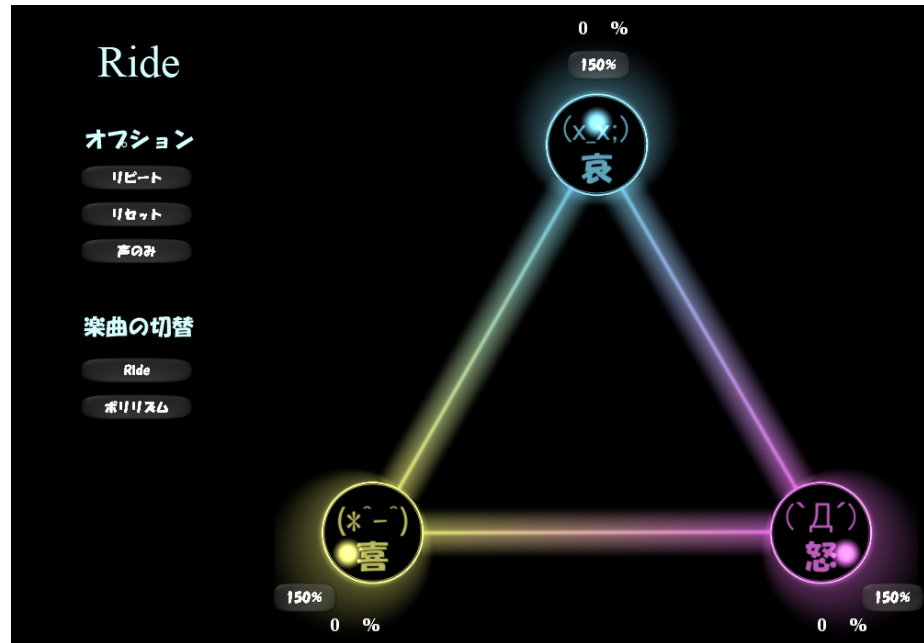


合成済歌唱

	歌い回し			
	(0%, 0%)	(0%, 10%)	...	(0%, 100%)
モーフィング率	(10%, 0%)	(10%, 10%)		
		⋮	⋮	
	(100%, 0%)			(100%, 100%)



# モーフィング技術の応用(デモ)



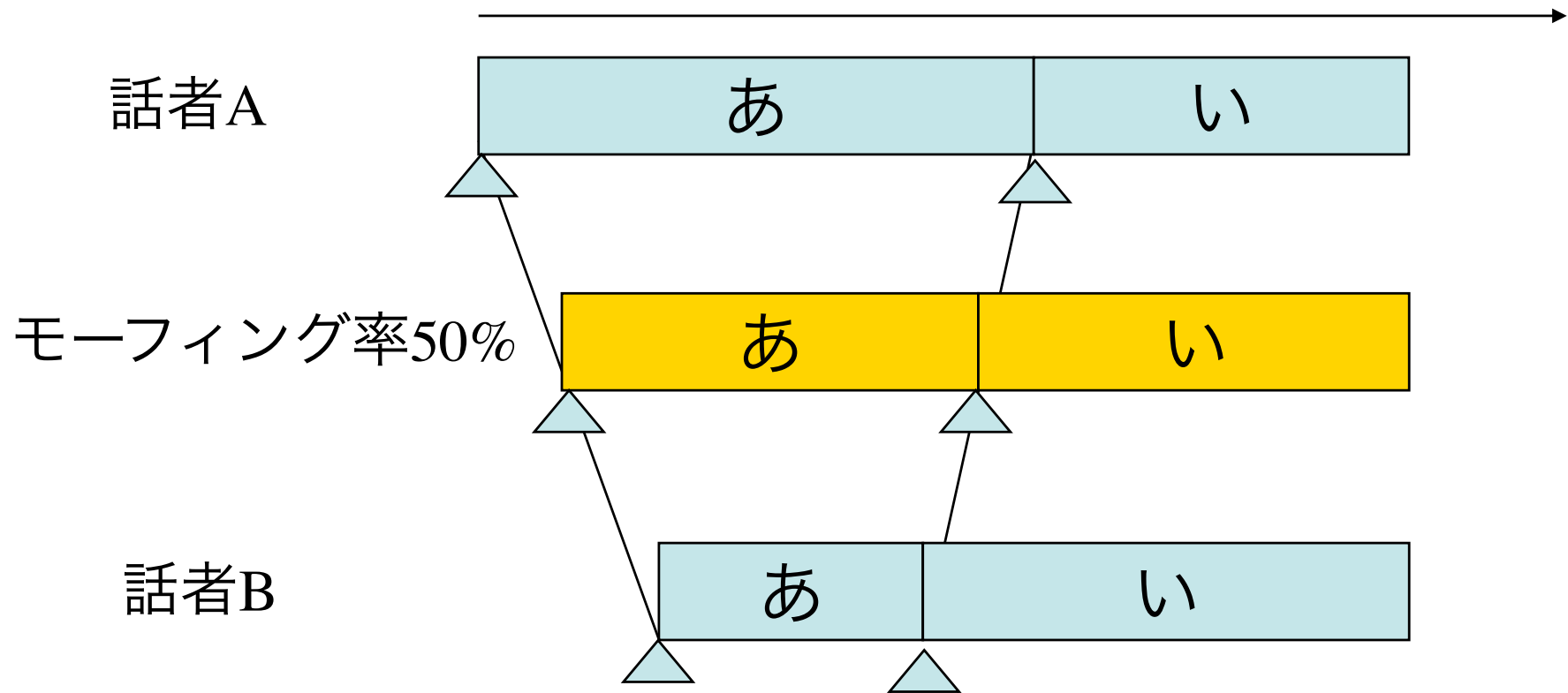
すでにある事例から，これまでに無い表現を

# リアルタイム処理への課題

- 計算時間
  - モーフィングを行うための時間
  - 音声を合成するための時間
- 音声間の対応付け
  - 2歌手の時間的な対応付けが必要

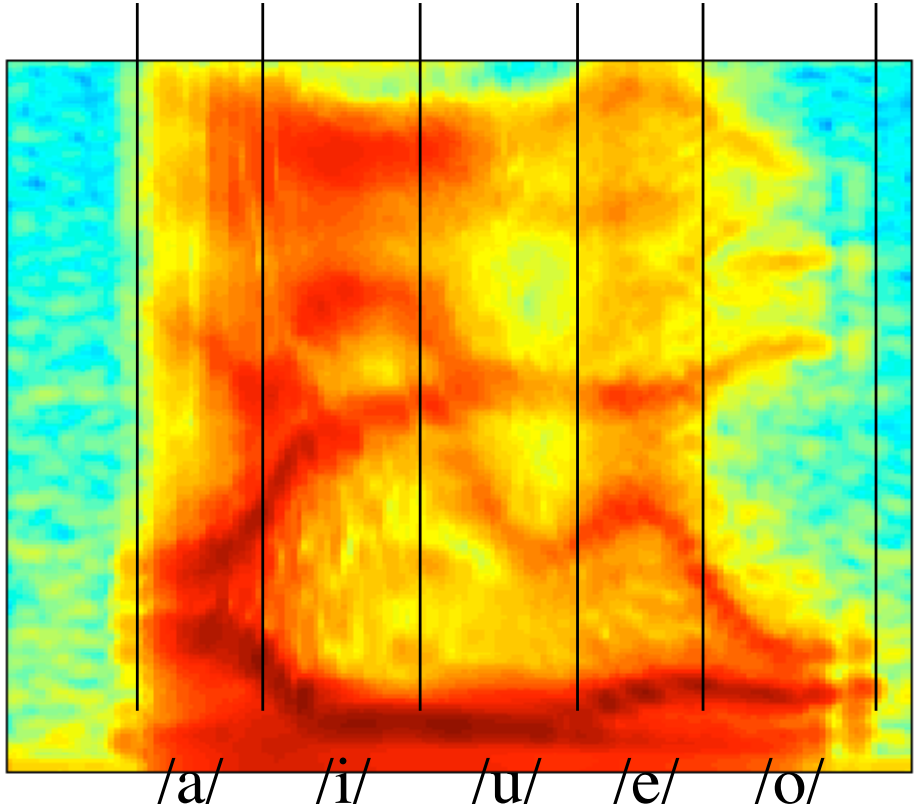


# 時間的な対応付けとは？



- ・ 音素の変わり目を正確に測ることが必要
  - 「か」の場合/k/と/a/についてそれぞれ対応付ける

# 実際の音声



- /aiueo/と発音した例
  - 音素の切れ目が無く，境界の設定が困難

# 対応付けの現状

- ・ 対応付け作業を練習した人間の経験と勘により対応付け
- ・ CrestMuseプロジェクトの別のグループでは、歌詞と歌声の対応付けを行っている

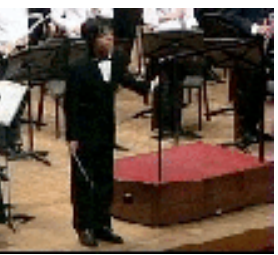
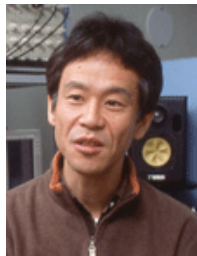


グループ間で連携することで自動化技術を推進

# その他歌唱合成の支援技術

- CrestMuseプロジェクトとしての可能性
  - 話し声を歌声に変換
  - 歌詞と歌声の対応付け
  - 歌声の特徴を抽出
- 歌声の作り込みを支援する総合的なインタフェースが構築可能

# CrestMuse Project



片寄晴弘, 森勢将雅: 音楽情報処理技術とエンタテインメント



# おわり

<http://crestmuse.jp/>

<http://ist.ksc.kwansei.ac.jp/~katayose/v.morish/>