

歌声合成の過去・現在・未来

2010年3月10日

ヤマハ株式会社 サウンドテクノロジー開発センター

剣持秀紀



過去

歌声合成の歴史



1960

1970

1980

1990

2000

2010

Physical Model

Bell Lab.
"Daisy Daisy"
(1961)



Acoustic Tube Model

P.Cock
"SPASM"
(1990)

より精緻なPhysical Model

Pseudo Physical Model

DECTalk
による歌唱合成

KTH
"Musee DIG"
(1984)

"KAE Labs"
Vocal Writer
(1999)

VirSyn
"Cantor"
(2004)

VOCALOID
(2003)

VOCALOID2
(2007~)

Spectral Model

Sinusoidal
Modeling

Macon
"Lyricos"
(1997)

NTT
"Wonder Horn"
(2004)

Time-domain
Formant合成

IRCAM
"CHANT"
(1984)

YAMAHA
"PLG-100SG"
(1997)

音声合成・信号処理
に関連する出来事・
製品など

LPC発明
(1970)

Speak & Spell
(1978)

Phase Vocoder
(1978)

PSOLA
(1989)

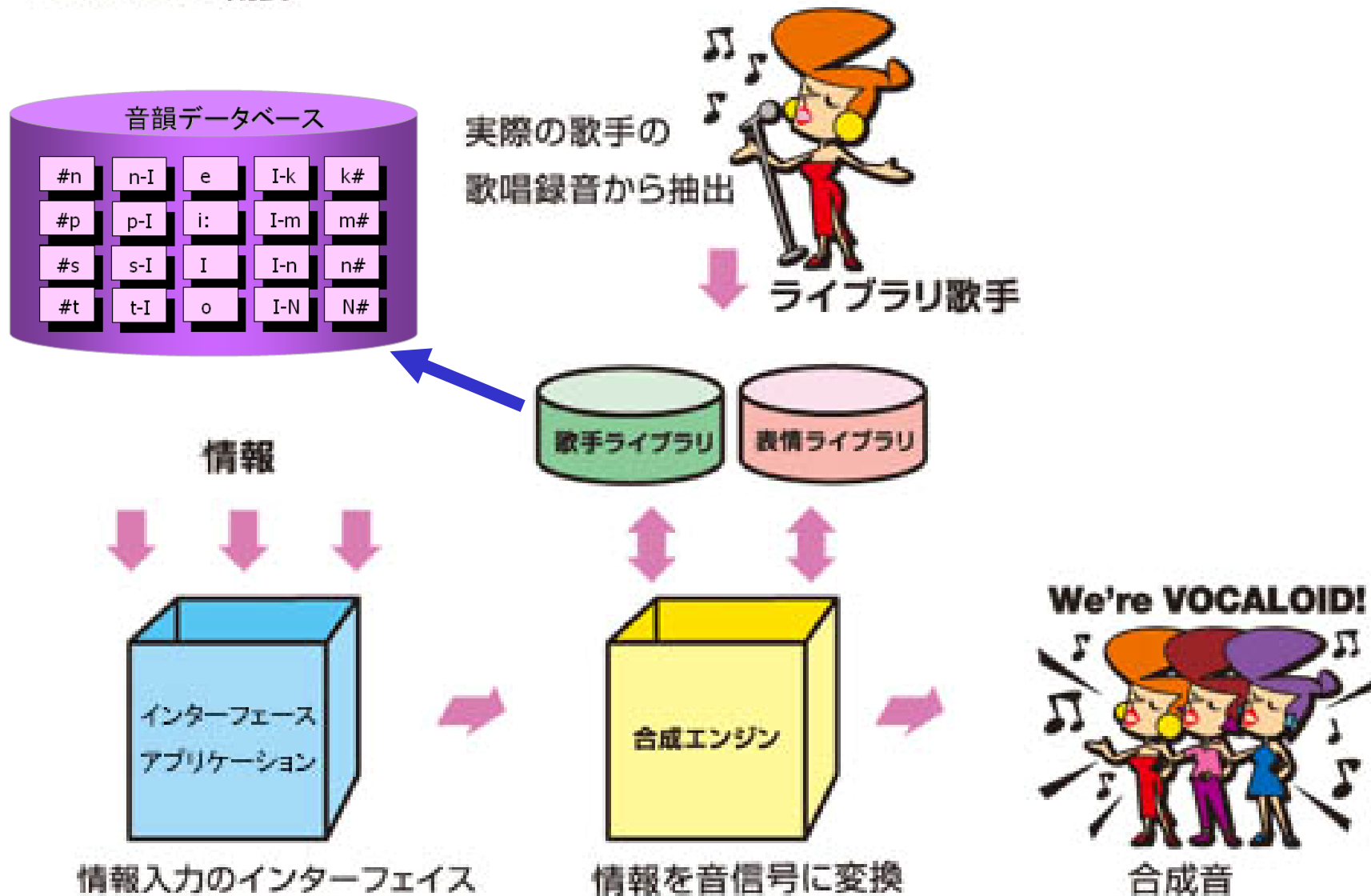
DECTalk(1983)

CHATR
(1994)

テキスト音声合成は
大規模コーパス利用の時代へ

現在

VOCALOID概要



VOCALOID Package Products Lineup

VOCALOID
NEW SINGING SYNTHESIS TECHNOLOGY

VOCALOID₂



なぜ歌声合成を利用するか？



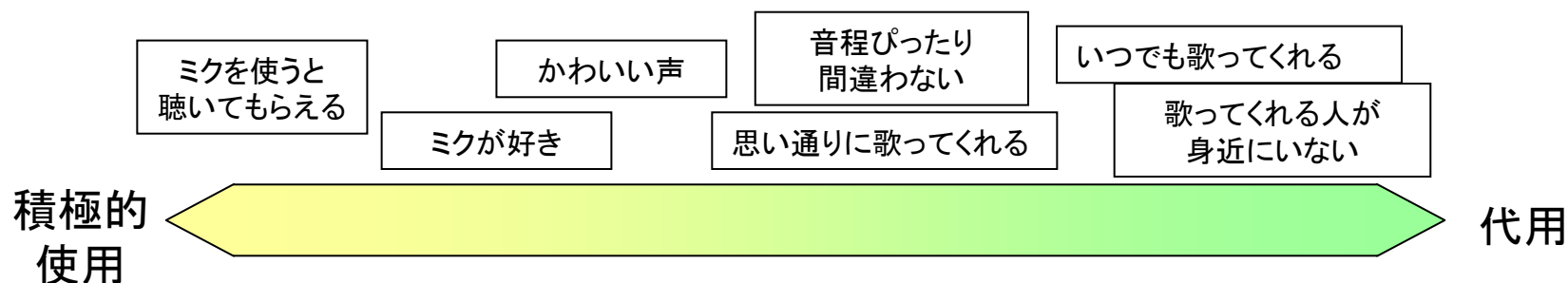
持ち運びやすい
正確
3連符や変拍子も
チューナー機能もある
etc.

視認性が悪い
聞こえにくい
味がない

視認性が良い
よく聞こえる
味がある

持ち運びにくい
よたる
カツカツとチーンだけ

■ 人間の歌手の単なる代用以上のものでなければならない

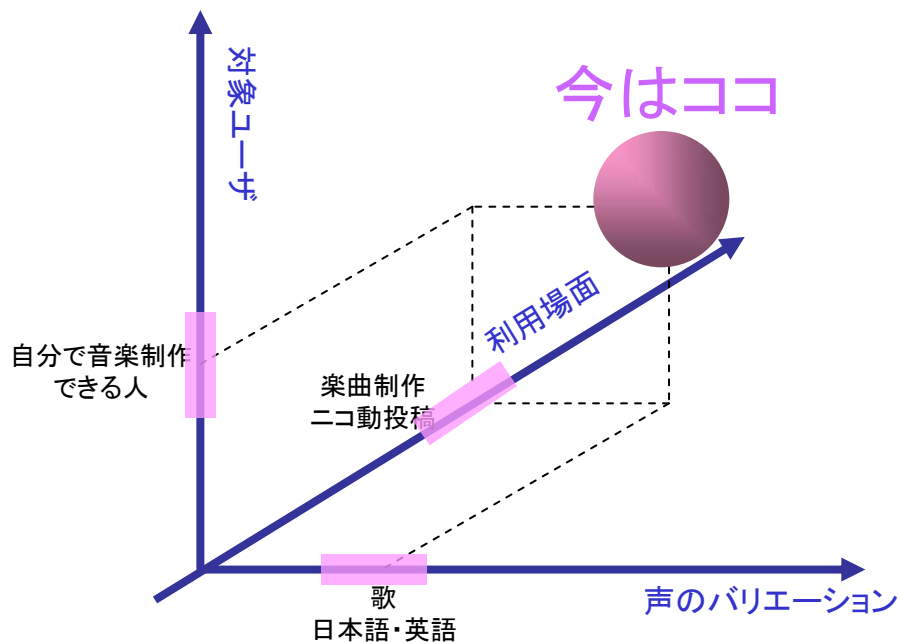


未来

- 声のバリエーション拡大

- 利用場面拡大

- ユーザ層拡大
→ 歌声合成の普遍化



- 歌声のバリエーション
 - 音楽ジャンル、歌声のスタイル
 - 言語バリエーション
- 歌声から歌声以外へ

“VOCALOID-Flex”

- 韻律を直接操作可能な合成エンジン
- 歌声とそれ以外を区別しない
- 「テキスト」から離れた発話

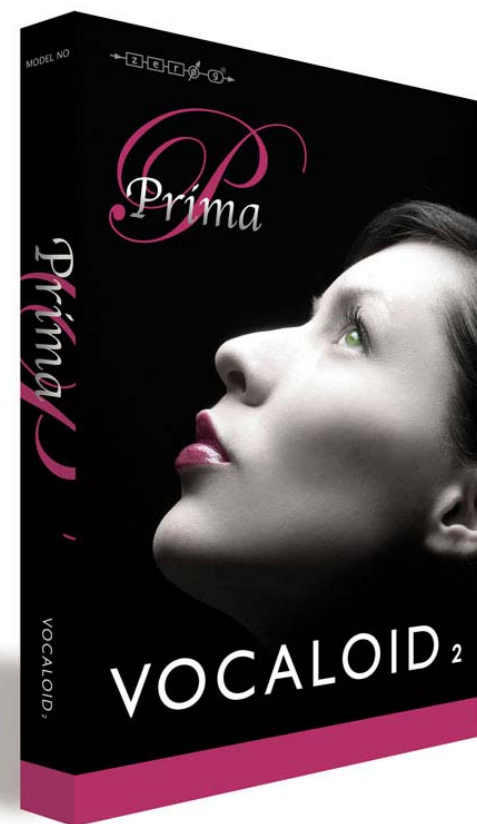
歌声

話し声
(朗読)

利用場面の拡大



- 現状：音楽制作→ニコ動に投稿
今後はそれ以外の用途にも広く使われるように
- ライブ、コンサートでの活用
- 家電、ロボットへの組み込み



サイバネティックヒューマンHRP-4C(産総研)

Center for Advanced Sound Technologies, Yamaha Corporation

ユーザ層拡大→歌声合成の普遍化



- 「私作る人」「私聴く人」???
- 「私も作ってみよう」



- 「みんなで作ろう」
 - 「歌のパートを作りたい、変えてみたい」



Soundation Studio (PowerFX)

- 15年後には...
 - 「とりあえず、まずはVOCALOIDの録音しておこうか」
 - 「あの歌手は合成じゃなくて生だって。イマドキ珍しいねえ。」