

音声合成

- 音声合成とは？
- 音声合成の品質尺度
- 録音編集とテキスト音声合成
- 技術各論＞韻律の制御
- 技術各論＞素片の接合

音声合成

Copyright © by Takeshi Kawabata

音声合成とは

- コンピュータに人間の音声を、話させる技術

第一会議室を
予約して
ください



第一会議室を
予約です
か？

音声合成

Copyright © by Takeshi Kawabata



音声合成の品質尺度

- 音響的ひずみ

- 明瞭度

- 自然性

音声合成

Copyright © by Takeshi Kawabata



録音編集とテキスト音声合成

- 録音編集 (Pre-recorded voice)

- テキスト音声合成 (TTS: Text To Speech)

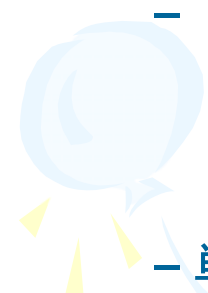
音声合成

Copyright © by Takeshi Kawabata

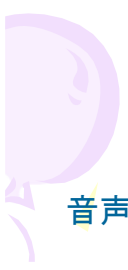


録音編集

- 録音編集 (Pre-recorded voice)

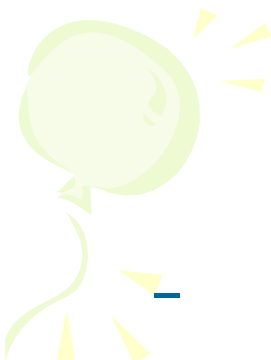


⇒
— 単語・句レベルの表現を前もって録音



⇒
⇒
音声合成

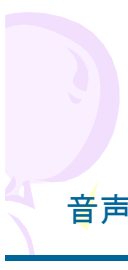
Copyright © by Takeshi Kawabata



録音編集 (2)



⇒ 公共サービスにおいては、その「声」の発声者を、
長期間にわたって確保する必要がある。



—
音声合成

Copyright © by Takeshi Kawabata

テキスト音声合成

- テキスト音声合成

(TTS: Text To Speech)

⇒

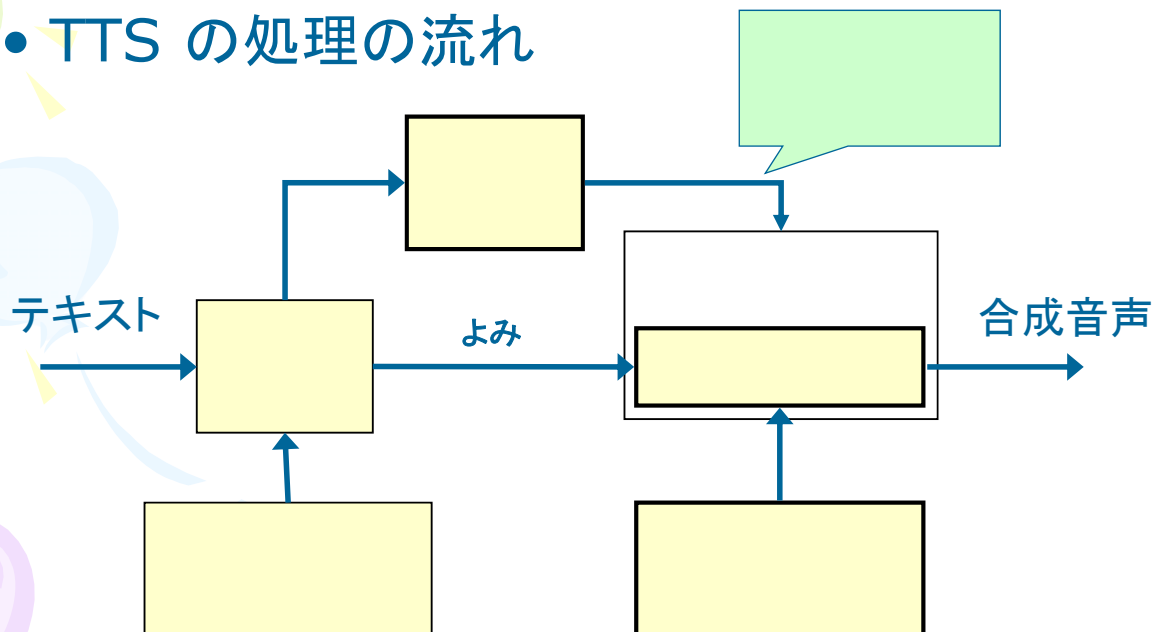
⇒

音声合成

Copyright © by Takeshi Kawabata

テキスト音声合成 (2)

- TTS の処理の流れ



音声合成

Copyright © by Takeshi Kawabata

技術各論＞韻律の生成

- アクセント

- イントネーション(話調)

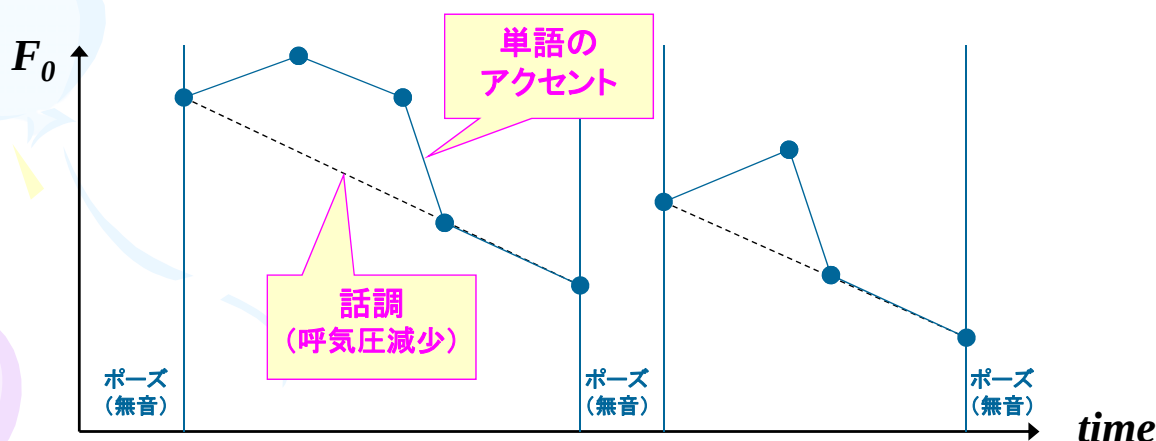
- ポーズ

音声合成

Copyright © by Takeshi Kawabata

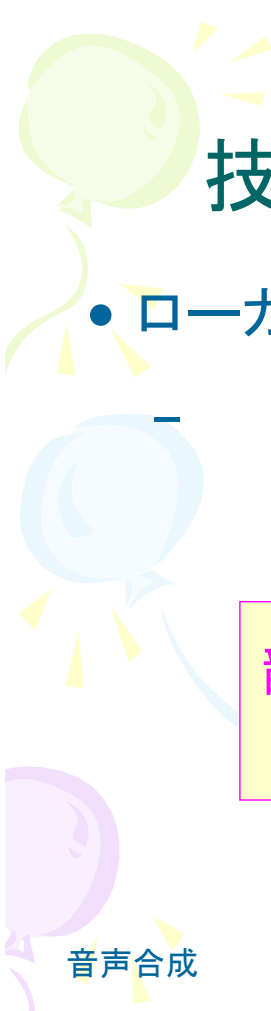
技術各論＞韻律の生成 (2)

– ハットパターンモデル[藤崎1971]



音声合成

Copyright © by Takeshi Kawabata



技術各論＞韻律の生成 (3)

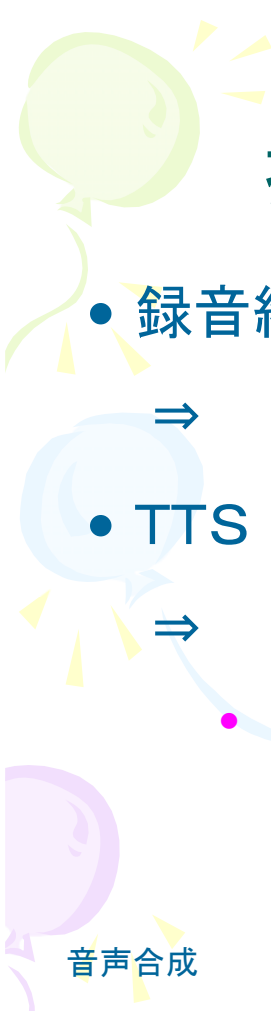
- ローカルピッチ構造／グローバルピッチ構造

韻律をうまく与えないと、合成された音声が非常に不自然になる

録音編集／TTSを問わず重要なポイント！

音声合成

Copyright © by Takeshi Kawabata



技術各論＞素片の接合

- 録音編集

⇒

- TTS

⇒

音声合成

Copyright © by Takeshi Kawabata

技術各論＞素片の接合 (2)

- PSOLA

(Pitch Synchronous Overlap Add)

—

—

音声合成

Copyright © by Takeshi Kawabata

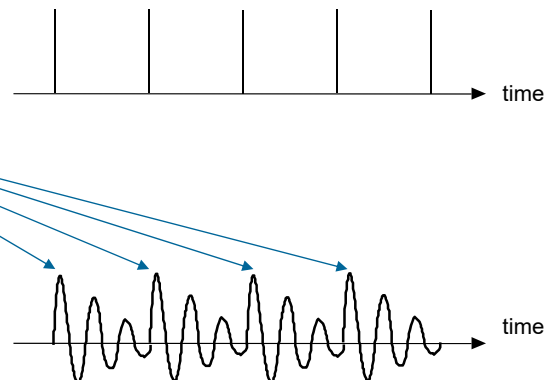
技術各論＞素片の接合 (3)

- PSOLA

— 1周期分(+ α)の波形をピッチに合わせて重ね合わせる

時間窓

原波形
(2周期分)



音声合成

Copyright © by Takeshi Kawabata



技術各論＞素片の接合 (4)

• PSOLA



Copyright © by Takeshi Kawabata

ここまでのポイント

1. 録音編集／テキスト音声合成(TTS)の違いを説明せよ
2. テキスト音声合成(TTS)技術はどのような要素から構成されているか？
3. PSOLAとはどのような技術か？