

LSP パラメータにベクトル量子化を適用した VCV 規則音声合成法

2 J - 1

隅田庸市 清水忠昭 石坂智 井須尚紀 菅田一博

鳥取大学工学部知能情報工学科

1.はじめに

本研究では、少ない情報量で高品質な合成音声を得る事を目的とした。多量の VCV 素片音声データから得られた音声の特徴を表すパラメータに、ベクトル量子化法を適用することで大幅な情報圧縮を行った。情報圧縮した VCV 素片のパラメータ間を最適に接続することにより、任意単語を作成した。

2.研究方法

本研究では図 1 に示すように、コードブック、VCV データベース、距離データベースを用いて合成音声を作成した。合成音声作成のために用いたコードブック、VCV データベース、距離データベースの作成方法を図 2 に示す。コードブックは、(財)日本情報処理開発協会の研究用音声データベース中の男性話者 1 名の発声による音素バランス文を分析回数 12 次、フレーム長 23.22ms、更新周期 11.61ms で LSP 分析し、得られた 89162 組の LSP パラメータに LBG アルゴリズム【1】を適用することで作成した。コードブックサイズは 2048 とした。VCV データベースは、コードブック作成に用いた文章から視察により切り出した VCV 素片 217 種類 470 個を、コードブックを用いてベクトル量子化することで作成した。VCV 素片をコードの時系列で表すことにより、約 1/200 の情報圧縮が可能となった。

また、本研究ではベクトル量子化された VCV 素片を接続する上で、接続歪が小さくなるように距離データベースを用いて VCV 素片を選択した。距離データベースとは、予めコードブック内の代表ベクトル間の距離計算を行い、各代表ベクトルとのスペクトル距離が小さい代表ベクトルを求め、距離の近い順に 10 番目までの代表ベクトルの添え字番号を

Speech Synthesis Based on VCV Units Using Vector Quantization for LSP Parameters

Youichi Sumida, Tadaaki Shimizu, Satoru Ishizaka, Naoki Isu, Kazuhiro Sugata

Dept. of Information and Knowledge Engineering, Tottori

Univ.,4-101 Koyama-minami, Tottori 680, JAPAN

求めたものである。接続歪が最も小さくなる素片の選択には、膨大な計算量を必要としていたが、距離データベースを用いることで接続歪が最小となる素片の容易な選択が可能となった。

3.評価実験結果

本手法で作成した合成音声資料 8 単語を、35 人の被験者に対して、一対比較の聴覚実験を行うことにより、本手法の有効性を自然性、明瞭性の 2 点から評価した。評価実験の結果、明瞭性については高い評価を得、本研究で作成したコードブックが学習に用いた音声データを能率良く代表していることが示された。

【参考文献】

- [1] 守谷健弘・菅田雅彰：音声の適応変換ベクトル符号化，信学論(A)，J67-A.10(1984-10)

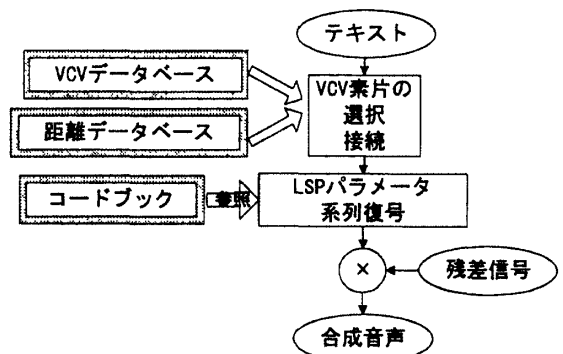


図1. 合成音声作成方法

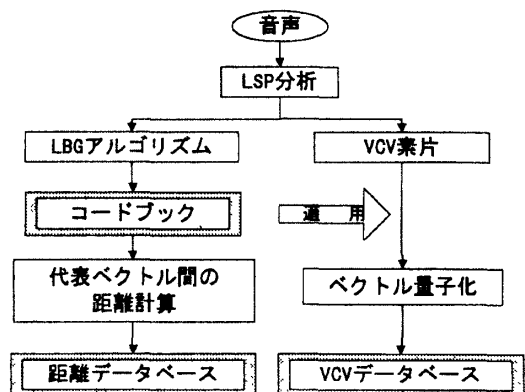


図2. コードブック、距離・VCV データベースの作成方法