

明瞭な子音音声合成のための駆動音源信号の生成法

5G-6

太田垣浩之[†] 田頭徹也^{††} 清水忠昭[†] 菅田一博[†] 井須尚紀[†][†]鳥取大学 工学部 知能情報工学科 ^{††}ヤクモ (株)1. はじめに

LPC 分析合成方式 (以下,LPC 方式) は, 音声生成機構の調音部を全極モデルで近似し, 有声音源をインパルス列, 無声音源をホワイトノイズでモデル化している. このモデル化を行うと, 音源と調音に関する特徴パラメータにより, 少ない情報量での音声符号化が可能であるが, 合成音声の品質に限界がある. 本研究では, 予測残差信号のパワーの変化を考慮した無声音源を提案し, 子音部の合成音声の高品質化を図る.

2. LPC 分析合成方式の問題点

LPC 方式は, 簡易的な処理によって, 音声信号において重要な極の特徴を捉えることができ, ほぼ実用的な明瞭度を有する合成音声を得られる. しかし, 子音のように, 非定常的で細かく振動する信号の特徴は, 全極モデルの極と, パワーの均一なホワイトノイズでは捉え切れない. 従って,LPC 方式による子音部の合成音声は, 明瞭性に欠ける.

LPC 方式の音源として, 予測残差信号を直接利用することが考えられ, MPC,TOR,CELP 等の諸方式が提案されているが, いずれも高いコストがかかる.

本研究では, 予測残差信号のパワーの変化を捉えた駆動音源信号を生成し, ローコストで子音部の合成音声の高品質化を図る.

3. 本研究で考案した駆動音源生成法

子音は, スペクトルの時間変化によって特徴づけられているものが多い.

LPC 方式において子音の情報は, 予測残差信号に多く残っていると考えられる. 本研究では, 子音のパワーの細かな変化を考慮し, 高精度な駆動音源信号を生成する手法を提案する.

ピッチ抽出は, 変形相関法を用いる. 変形相関法により分析フレームごとに求められる有声度¹⁾(0 から 1 までの実数) が 1 未満のときは, 子音部とみなし, 以下の処理を行う.

1. 単純分割法【Method 1】

図 1 に示すように, 予測残差信号の分析フレームを等分割 (分割数 n) し, 分割区間毎の細かな音源振幅情報を抽出することで, 子音のパワーの変化を考慮した無声音源を作成する.

2. 可変分割法【Method 2】

単純分割法は, 子音の情報を高精度に捉えられるが, 情報量が増加する. 可変分割法は, 予測残差信号の振幅パワーが同程度の区間を 1 つにまとめ, 分割区間を可変にする.

4. 合成音声品質評価

合成音声は, 分析次数 12 次の LSP 分析²⁾ により作成した. サンプルングは, 10 kHz, 12bits, LSP 分析は, ハミング窓, 分析フレーム 20.0ms, 分析インターバル 12.0ms とした. 有声度が 1 未満のときは, 分割数を 30 とした.

合成音声品質を検討するために, LPC 方式の音源パラメータから作成した合成音声と本研究で考案した音

Improved method of Source Making

for Consonant Speech Synthesis

Hiroyuki Otagaki[†], Tetsuya Tagashira^{††}Tadaaki Shimizu[†], Kazuhiro Sugata[†], Naoki Isu[†][†]Dept. of Information and Knowledge Engineering,
Faculty of Engineering, Tottori University^{††}YACMO CO.,LTD.

源パラメータから作成した合成音声の対比較試験を行った。音声資料は、FM放送のニュースより採取した男声2種類、女声2種類を用いた。対比較試験は、30名の被験者を対象に行った。

図2に対比較試験のプレファレンススコアを示す。図2より、本研究で考案した駆動音源生成法は、高品質音声合成に非常に有効であるといえる。本手法は男声に対して有効であり、単純分割法、可変分割法で作成した合成音声の品質には明確な差がなかった。

本手法で生成した女声資料の駆動音源信号の波形は、図3に示すように、本来の予測残差信号の波形の特徴を捉えていたが、対比較試験で女声に対して有効でなかった。女声に有効でなかった理由として、女声のピッチ周期が短いこと、分析結果がピッチ周期の高調波成分に影響を及ぼされる偏向を受けることが考えられる。

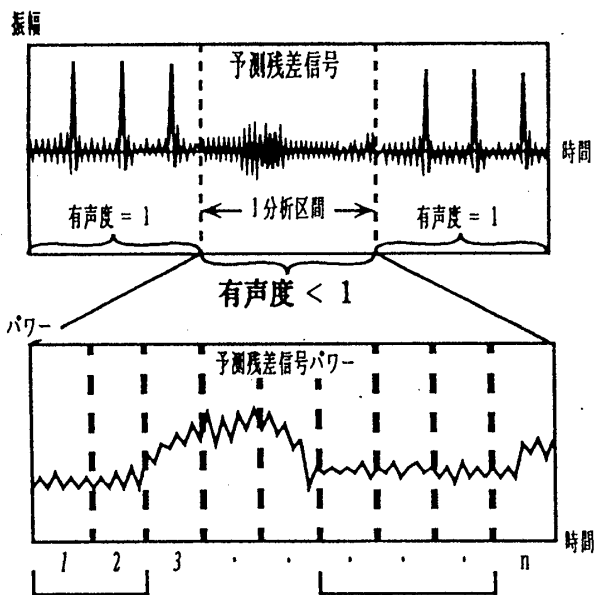


図1 単純・可変分割法

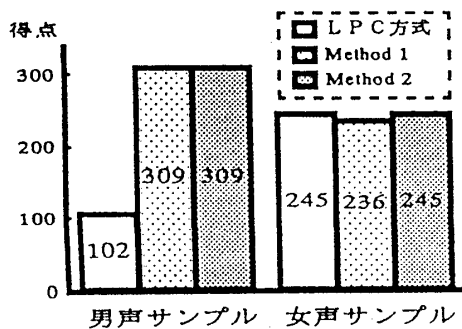


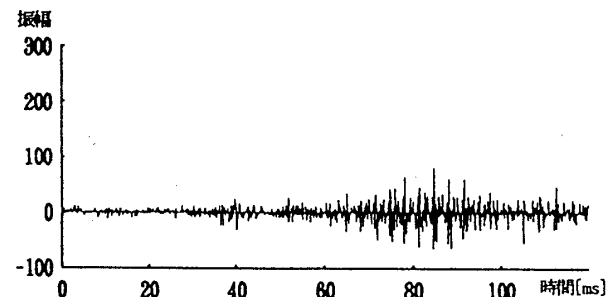
図2 合成音声のプレファレンススコア

5. おわりに

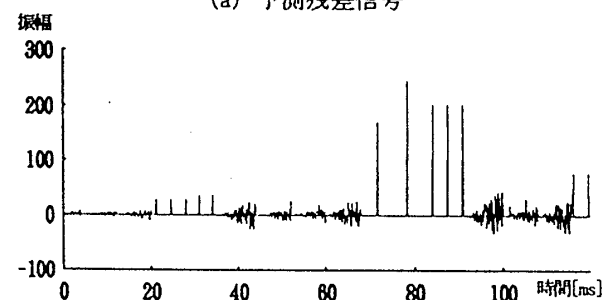
本研究では、子音部の予測残差信号の分析フレームを分割し、細かな振幅情報を求めることによって子音部の特徴を詳しく抽出した結果、LPC方式の合成音声より明瞭な合成音声を得ることができた。今後の課題としては、高精度なピッチ抽出法の開発、短いピッチ周期にも対応できる音声合成法の開発、合成音声の客観的評価法の確立が挙げられる。

参考文献

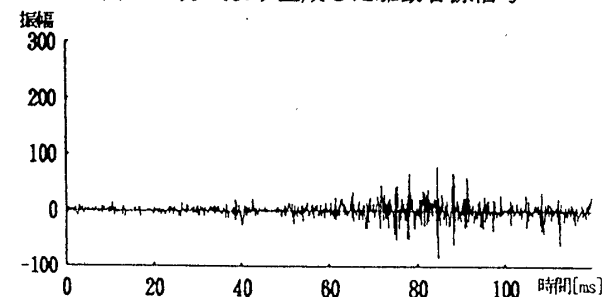
- (1) 安居院，中嶋：コンピュータ音声処理，秋葉出版，pp.65~67(1980).
- (2) 管村，板倉：“線スペクトル対 (LSP) 音声分析合成方式による音声情報圧縮”，信学論 (A), J64-A,8,pp.599~606 (1981-08).



(a) 予測残差信号



(b) LPC方式より生成した駆動音源信号



(c) 本研究の手法により生成した駆動音源信号

図3 予測残差信号と駆動音源信号の比較