

# F0 微細変動を考慮した F0 モデルパラメータ推定手法

杉山 悠<sup>†</sup> 黒岩 眞吾<sup>†</sup> 堀内 靖雄<sup>†</sup>

<sup>†</sup> 千葉大学大学院融合科学研究科

## 1. はじめに

発声中の基本周波数(F0)の時間変化を表現するモデルとして、F0 モデルが提案されている[1]。これは、図 1 のようにアクセント指令・フレーズ指令という二つの指令に基づく F0 モデルの式で F0 の時間変化を表現可能であり、少ないパラメータで F0 データに対し、高い近似が可能であることから、有用な韻律の記述方法であると考えられているが、専門家の手作業によるパラメータ推定が必要であった。そこで、F0 モデルのパラメータを音声から自動推定する手法が報告されているが実用的な精度には至っていない[2][3]。自動推定が困難である原因の一つとして F0 微細変動がある。F0 モデルでは F0 微細変動が考慮されていないため、推定の際にパラメータの誤差を生む要因となっていた。そこで、簡易的な言語情報を元に F0 微細変動の補正を行う手法を提案する。

## 2. 提案手法

F0 微細変動の発生箇所を節境界区間、頭高型アクセントの文頭区間、無声阻害子音の前後母音区間、有声阻害子音区間に分類した。その区間に対し F0 の補正を行う事で F0 の微細変動の影響を取り除く。まず節境界区間について F0 モデルのフレーズ指令の F0 軌跡を元に F0 の外挿を行う。次に頭高型アクセント区間に関しアクセント指令の F0 軌跡を元に F0 外挿を行う。最後に無声阻害子音の前後母音、有声阻害子音区間について直線と 3 次曲線による補正の 2 種類の補正について提案を行う。予備実験により指令の大きさのパラメータ推定精度は直線補正、指令の位置推定精度は 3 次曲線による補正が優れていることが分かった。そこで F0 モデルのパラメータ推定について「指令の位置が決まれば、指令の大きさも決定する」という仮説の元、3 次曲線補正を行った F0 より得られた推定パラメータの指令位置を固定し、直線補正 F0 を用いて範囲制約付き最小二乗法によって指令大きさの再最適化を行う。

## 3. 実験データ

ATR データベース B セットの男性ナレーターによる発話音声 75 文に対し実験を行い、手作業で付与されたパラメータを正解とする。推定結果の評価指標として正解率、正解パラメータと推定パラメータの RMSE を求めた。正解率とは正解との時刻の差が 124ms (対象話者の 1 モーラの平均時間長) 以内で推定されたパ

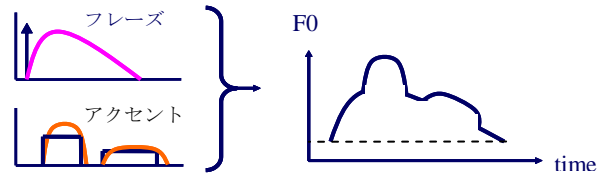


図 1. F0 モデル

表 1. 推定実験結果(正解率)

	提案手法 正解率	全自動推定[2] 正解率
フレーズ	95.1%	94.4%
アクセント	93.6%	68.1%

表 2. 推定実験結果(RMSE)

		実測 F0[3]	補正 F0
フレーズ	位置	0.068	0.069
	大きさ	0.138	0.110
アクセント	開始位置	0.048	0.046
	終了位置	0.053	0.052
	大きさ	0.121	0.091

ラメータの割合である。

## 4. 実験結果

表 1 に同じデータベースを対象とした[2]の全自動推定手法の推定精度を示す。また、表 2 に本手法と同様に言語情報を用いた先行研究[3]と提案手法における正解パラメータとの RMSE を示す。

## 5. 今後の課題

提案手法による推定結果が音声分析において実用性に足る精度を持つかを測る評価指標の検討を行う。

## 参考文献

- [1] H. Fujisaki, K. Hirose, Journal of Acoustical Society of Japan. (E)5(4), 233-242, 1984
- [2] 成澤 修一, 峯松 信明, 広瀬 啓吉, 藤崎 博也, “音声の基本周波数パターン生成過程モデルのパラメータ自動抽出法の評価”, 情処学会研報 (SLP), 2003(75(2003-SLP-047)), 1-6 (2003-07-18)
- [3] 今里 圭太, 堀内 靖雄, 篠崎 隆宏, 黒岩 眞吾, “言語情報を利用した F0 モデルパラメータの推定手法”, 日本音響学会 2013 年春季講演論文集, 3-P-36d, pp. 507-508, Mar. 2013.