

調音パラメータ空間におけるベクトル量子化

1 K-2

飯島幸紀† 樺澤康夫‡ 松田郁夫‡

†日本工業大学大学院工学研究科電気工学専攻

‡日本工業大学工学部情報工学科

1. はじめに

声帯と声道の空気力学的シミュレーションに基づく声帯・声道型動的音声合成モデルは、声帯の自励振動、声道形状の時間的変化、声道内壁における熱伝導と粘性摩擦による損失、声道における音の励振、鼻腔結合、口唇からの音の放射などの各種損失を考慮しており、実際の音声生成の物理法則に則しているため、自然性の高い高品質の音声合成ができる。しかし、調音パラメータ(制御パラメータ)を精密に与えなければならない。このモデルに基づいて、音声波から調音パラメータを推定する手法として、非線形最適法によるものが提案されている¹⁾。しかし、パラメータの微小な変更のたびに音声の合成と分析を繰り返すため、計算量が多大である。本研究では、あらかじめ多数の調音パラメータにより音声を合成、分析しておき、スペクトル情報と調音パラメータを組にして蓄積して、与えられた音声波に対して、この中から最も適合する調音パラメータを選択するという方法について検討する。選択の効率化のためには、蓄積情報を制限する必要があるが、これはちょうどベクトル量子化の考え方と同じであり、実音声の確率的分布を反映した調音パラメータ(符号語)の配置を行うことが必要となる。このような符号帳の構成は最適化問題となるので、これを遺伝的アルゴリズムの応用により解くことを検討する。

2. 遺伝的アルゴリズムによる符号帳の作成

調音パラメータの推定問題をベクトル量子化問題と見ると、分析合成系における符号帳の構成問題と考えることができる。

図1に遺伝的アルゴリズムによる符号帳作成のブロック図を示す。ここでは、調音パラメータを遺伝子によって表現し、スペクトル情報として、ケプ

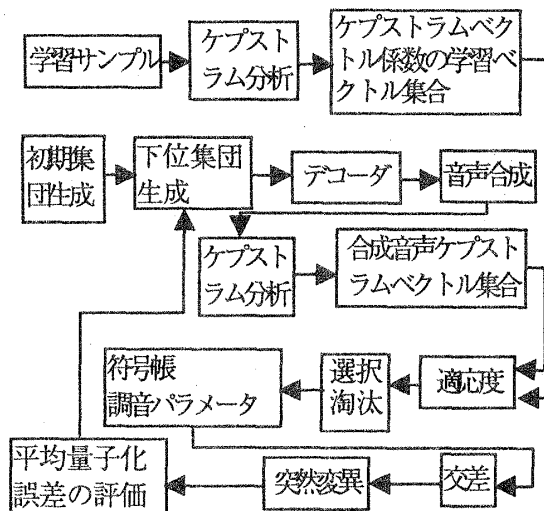


図1. 符号帳作成のブロック図

ストラム係数ベクトルを採用している。

本研究では、遺伝子がそれぞれ独立した別々の符号語になるので、遺伝子同士が影響しあつて類似した解ができないようにするため、次のようなアルゴリズムを使用している。

1. 学習サンプルのケプストラム係数ベクトル集合 S の生成 (サンプル数を N とする)

$$S = \{S_n : n = 0, \dots, N\}$$

2. 初期集団 G_0 を生成し平均量子化誤差 $err(G_0)$ を求める。(符号語数を M とする)

$$G_0 = \{g_i : i = 0, \dots, M\}$$

$$\min \leftarrow err(G_0), G_{opt} \leftarrow G_0$$

3. for ($t=0 ; t < \text{最大世代交代数} ; t++$) {

4. 下位集団の生成

符号帳 G_t の各符号語 g_i をもとに、各パラメータにランダムな振幅を持たせ類似した下位集団 G_{u_i} を作成する。(下位集団の遺伝子数を m とする)

$$G_{u_i} = \{gu_{ij} : j = 0, \dots, m\}$$

5. 適応度の計算

各下位集団の gu_{ij} の適応度 $F(gu_{ij})$ を求める。

6. 各集団毎の最良遺伝子 g_{ij} の選択

$$g_i \leftarrow \arg \max_j F(gu_{ij})$$

新符号帳 G_t を作成する。

$$G_t = \{g_i : i = 0, \dots, M\}$$

7. 符号帳 G_t の平均量子化誤差 $err(G_t)$ を求める。

8. if ($err(G_t) < \min$) {
 $\min \leftarrow err(G_t)$;
 $G_{opt} \leftarrow G_t$

} }

この方法は、並列 GA の離島モデルに類似した方法になっている。

3. 適応度

下位集団における適応度を適切に定めることが重要である。学習サンプルのケプストラム係数ベクトルが密集するところには、多くの符号語が置かれ、少ないところには、少なく配置されるのが良い。このような性質を反映した遺伝子の適応度の定義を行う。

第 i 遺伝子集団の遺伝子 j のケプストラムベクトルを $C(gu_{ij})$ 、第 k サンプルケプストラムベクトルを S_k とする。適応度 F を次式で定義する。

$$F(gu_{ij}) = \frac{\sum_{k=1}^N \{1 + d(C(gu_{ij}), S_k)\}^{-1}}{\sum_{j=1k=1}^M \sum_{m=1}^m \{1 + d(C(gu_{ij}), C(gu_{kj}))\}^{-1}}$$

分母は、遺伝子周辺の距離で重み付けした遺伝子数、分子は、距離で重み付けした遺伝子周辺のサンプル数と解釈できる。

4. 実験結果

学習サンプル 21300 個について、ベクトル量子化実験を行った。

実験条件

調音パラメータを図 2 に示す。各調音パラメータは表 1 に示すビット数により遺伝子にコード化した。 Lb, Lf は 8 cm に固定した。

学習サンプル N 21300 個 (母音のみ)

個体数 M 64, 下位集団の個体数 m 8

最大世代交代数 512, 交叉率 0.00

突然変異率 0.02

図 3 に、世代交代に伴う平均量子化誤差の変化を示す。

5. むすび

調音パラメータの推定にベクトル量子化の考え方

を導入し符号帳の生成に遺伝的アルゴリズムの応用を検討した。母音に対する平均量子化誤差の小さい符号帳が得られた。今後さらに、遺伝子座の配置、適応度の定義と量子化誤差の関係について検討する。

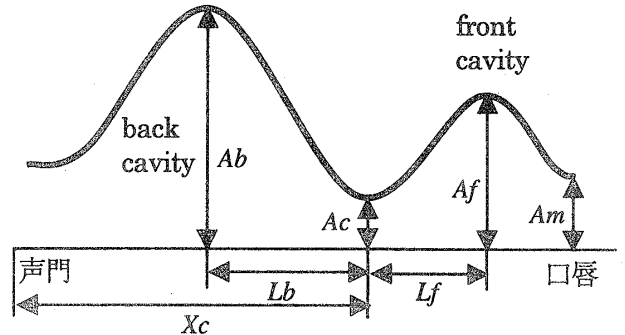


図 2. 調音パラメータ

表 1. 調音パラメータと遺伝子のビット数

調音パラメータ	X_c	A_c	A_b	A_f	A_m
ビット数	8	4	4	4	4

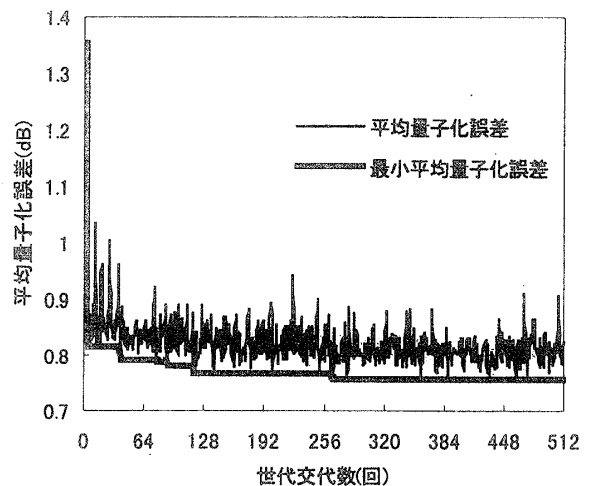


図 3. 平均量子化誤差

参考文献

[1] J.L.Flanagan et al.: "Signal models for low bit-rate coding of speech", J.A.S.A, 68(3), pp.780-791, 1980
 [2] 安居院 猛, 長尾 智晴: "ジェネティックアルゴリズム", 昭見堂, 1993
 [3] 小泉 保: "音声学入門", 大学書林, 1996
 [4] 古井 貞熙, 田崎 三郎 他: "ベクトル量子化と情報圧縮", コロナ社, 1998