

# LPC による声門流推定のパラメータの簡易選択 GUI

平野雄大<sup>1</sup> モクタリ パーハム<sup>1</sup> 森川大輔<sup>1</sup> 平原達也<sup>1</sup>

**概要:** 本研究では、線形予測符号化 (LPC) に基づく声門逆フィルタリング (GIF) アルゴリズムを用いて、音声信号から声門流を推定する GUI を作成した。作成した GUI は、3つの GIF アルゴリズムを選択でき、GIF に必要なパラメータの値を変化させたときの結果の一覧を表示させ、その中から最も Closed phase が平坦な波形を選択することで容易にパラメータの値の設定ができる。

## Simple parameter-selection GUI for glottal flow estimation by LPC

YUDAI HIRANO<sup>†1</sup> PARHAM MOKHTARI<sup>†1</sup> DAISUKE MORIKAWA<sup>†1</sup>  
TATSUYA HIRAHARA<sup>†1</sup>

### 1. はじめに

有声音の音源である声門体積流波形 (声門流) は、直接測定することが困難だが、声門逆フィルタリング (GIF: Glottal Inverse Filtering) で推定することができる[1]。

GIF を実行する GUI ソフトウェアは、いくつか開発されており Aalto Aparat[2], DeCap[3], GloRi[4]などがある。これらのソフトウェアは、グラフィカルユーザインタフェース (GUI) を使用する。この GUI によりユーザはデジタル信号処理の知識を必要とせず、GIF に必要なパラメータを設定することによって声門流波形を得る。しかし、これらのソフトウェアは、Iterative Optimal Pre-emphasis (IOP) を用いた GIF[5]のような、新たに提案された GIF 法は実装されていない。また、これらのソフトウェアは、GIF に必要なパラメータ値の設定を各パラメータに一つ一つ行うため、最適なパラメータ値を選択することが困難であった。

そこでいくつかの GIF 法を選択でき、最適なパラメータ値の選択が容易にできる声門逆フィルタリング GUI を作成した。本報告では、その作成した GUI について述べる。

### 2. 音声生成モデルとその推定

ヒトの音声生成過程は、source-filter モデルによると、音源波の生成、声道での共鳴、口唇と鼻腔からの放射の三段階から成ると考えられる[6]。生成された音声を  $S(z)$  とすると、音源波の特性を  $G(z)$ 、声道の共鳴特性を  $V(z)$ 、放射特性を  $L(z)$  とし、 $S(z) = G(z)V(z)L(z)$  と書ける[7]。GIF は、観測した  $S(z)$  から  $G(z)$  を推定する技術である。

音声生成の各段階は、音声信号に線形予測符号化 (LPC: Linear Prediction Coding) 分析を行うことによりモデル化できる[6]。LPC は全極モデリング手法であるため、零を含む鼻音を扱ことは困難である。したがって本研究では、主に非鼻化母音を分析した。

### 3. 声門逆フィルタリングのアルゴリズム

#### 3.1 音声データの前処理

サンプリング周波数 48 kHz で AD 変換した音声信号を、カットオフ周波数 70 Hz の 2700 次の線形位相ハイパスフィルタを通した後に、任意の周波数にダウンサンプリングした。以下では、このダウンサンプリングした音声信号を 50 ms の方形窓で切り出し、音声信号を  $s(n)$  とし GIF 処理を行った。

#### 3.2 GIF アルゴリズム

よく知られている LPC に基づく GIF アルゴリズム (Standard LPC), Iterative Adaptive Inverse Filtering (IAIF)[8], IOP を用いた IAIF (IOP-IAIF) の 3 種類の GIF アルゴリズムを MATLAB で実装した。

上述した 3 種類の GIF アルゴリズムのフローチャートを図 1 に示す。まず、声門  $G$  と口唇放射  $L$  の複合特性を除去するために、 $s(n)$  に  $1 - \mu z^{-1}$  のフィルタでプリエンファシスをかけて  $v_1(n)$  を求める[7]。  $\mu$  の値は、  $s(n)$  に 1 次の LPC 分析を行うことで算出する。LPC 分析の前には毎回分析する信号にハン窓をかけた。

次に、次数  $M_v$  の LPC 分析を行い、  $V_1(z)$  を求め、  $s(n)$  に  $V_1^{-1}(z)$  をかけて  $v_{11}(n)$  を求める。さらに、  $v_{11}(n)$  に放射特性  $L(z) = 1 - bz^{-1}$  の逆フィルタをかけて得られる波形が、Standard LPC による声門流波形  $g_{\text{stdLPC}}(n)$  である。

IAIF では、声道特性  $V(z)$  の推定精度を高くするために、Standard LPC 処理において  $V(z)$  を推定する前に、  $s(n)$  から音源特性  $G(z)$  をあらかじめ除去する。すなわち、一旦 Standard LPC 分析を行い、得られた  $g_{\text{stdLPC}}(n)$  を次数  $M_g$  の LPC 分析をして、音源特性  $G_{\text{int}}(z)$  を推定する。そして、  $s(n)$  に  $G_{\text{int}}^{-1}(z)$  と  $L^{-1}(z)$  をかけて得られる  $v_2(n)$  に対して次数  $M_v$  の LPC 分析を行い、  $V_2(z)$  を求める。さらに、  $s(n)$  に  $V_2^{-1}(z)$  と  $L^{-1}(z)$  をかけて得られる波形が  $g_{\text{IAIF}}(n)$  である。

<sup>†1</sup> 富山県立大学 工学部  
Faculty of Engineering, Toyama Prefectural University

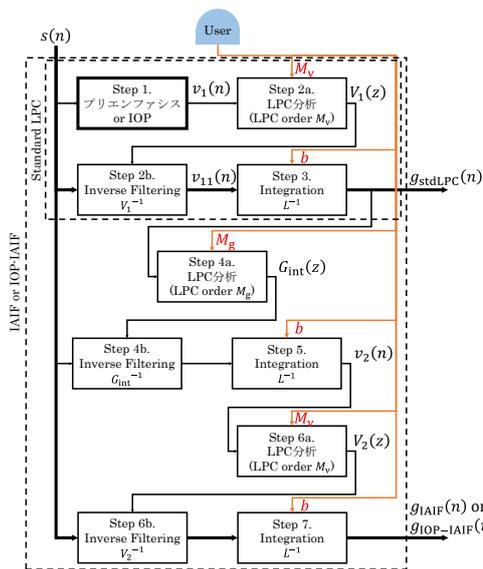


図 1: GIF アルゴリズムのフローチャート

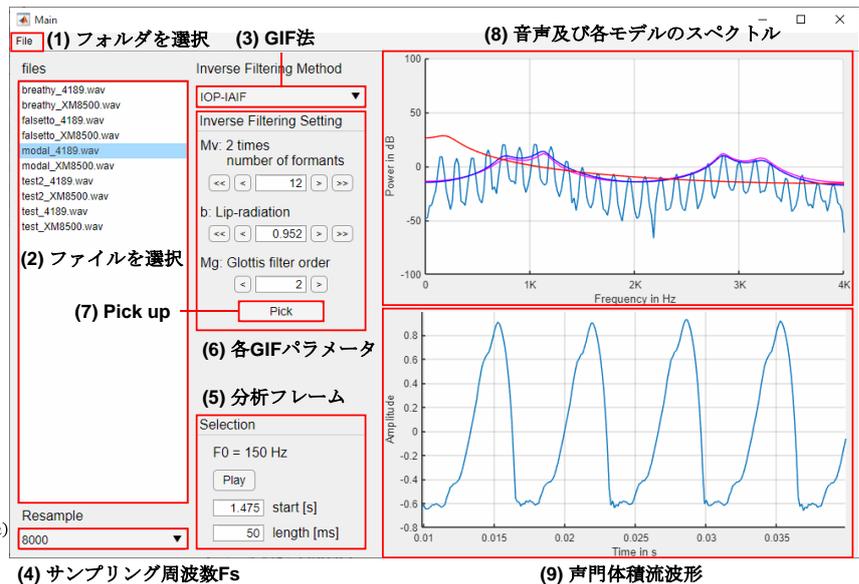


図 2: GUI の構成

IOP-IAIF[5]処理は、音源特性 $G(z)$ と放射特性 $L(z)$ の大局的な影響を除去するためのプリエンファシス・フィルタ係数 $\mu$ を最適化した IAIF 処理である。これにより、 $v_1(n)$ を source-filter モデルの声道特性により一致する、全体的にフラットなスペクトルにした。

## 4. 声門逆フィルタリング GUI

### 4.1 GUI の構成

GIF の各パラメータ ( $M_v$ ,  $b$ ,  $M_g$ ) の最適な値の設定を容易にするため声門逆フィルタリング GUI を構築した。声門逆フィルタリング GUI の構成を図 2 に示す。この GUI は、あらかじめ録音された音声信号を読み込み、GIF 法を選択して、音声信号から声門流を推定する。図 2 の例は、男性話者の地声の声質で持続性母音「a」を録音した音声进行分析した結果の一例である。

図 2(7)の「Pick」を選択すると、ウィンドウを開き、パラメータ  $M_v$  と  $b$  の値をそれぞれ 11 段階で変化させて、そのすべての  $M_v$  と  $b$  の組み合わせで推定した声門流波形を表示するようにした。図 3 に Pick で表示したウィンドウの一例を示す。ユーザは同図の波形から、リップルの量が少なく、波形が閉じている区間 (closed phase) が最も平坦な声門流波形を選ぶ。ユーザの選択を支援するために、波形の振幅の下部 10% のサンプルの標準偏差が小さいほど波形を濃く表示させた。波形の下部 10% のサンプル数が声門流の周期数の 3 倍以下の場合には、波形を淡い色で表示し、除去した。表示された波形の中で 1 つを選択すると、更に細かく  $M_v$  と  $b$  の値を変化させ、同様に推定声門流波形の一覧を表示させた。その中で、理想的な声門流波形を選択すると、メインウィンドウにパラメータが反映される。これにより一度に多くの推定結果を確認でき、最適なパラメータの値を選択することが容易になった。図 4 に GUI の

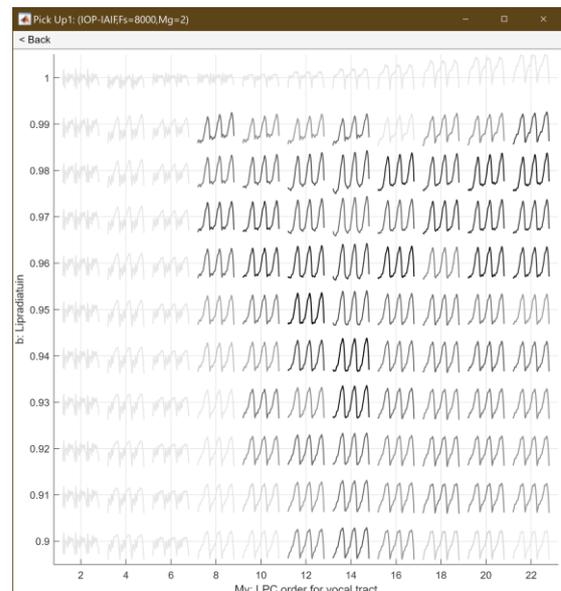


図 3: Pick ウィンドウ

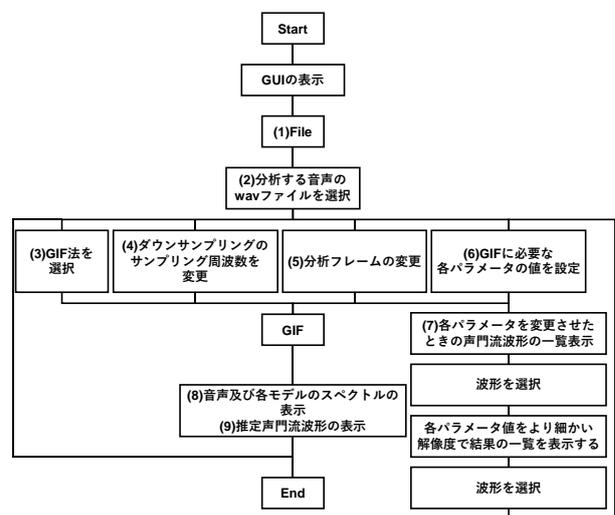


図 4: GUI のフローチャート

フローチャートを示す.

## 4.2 考察

GIF の各パラメータ ( $M_v$ ,  $b$ ,  $M_g$ ) の値を変化させたときの声門流波形の一覧表示する機能をさらに向上させ、自動的に最適なパラメータの値を選択できるアルゴリズムを開発することが今後の課題である。また本研究では、全極分析法である LPC に基づく GIF 法のみを実装しているため、鼻音の分析には適していない。表現力豊かな、または感情的な音声は鼻音を含む地声ではない声質になる可能性がある。このような地声ではない声質をもつ音声の分析を行うためには、極と零を扱う GIF 法が必要である。

## 5. まとめ

3 つの GIF アルゴリズムを選択でき、GIF に必要な各パラメータの値を変化させたときの声門流波形の一覧の中から波形の 1 つを選択することで、分析パラメータの値の細かな違いによる推定された音源波形の細かな違いを観察できる GUI を構築した。これにより、ユーザは、音声生成メカニズムの知識に基づいて、GIF に必要な最適なパラメータの値を設定しやすくなった。

## 参考文献

- [1] P. Alku, "Glottal inverse filtering analysis of human voice production – A review of estimation and parameterization methods of the glottal excitation and their applications," *Sadhana*, vol.36, 623-650, 2011.
- [2] P. Alku, *et al.*, "Aalto Aparat – A Freely Available Tool for Glottal Inverse Filtering and Voice Source Parameterization," in *Proc. Subsidia: Tools and Resources for Speech Sciences*, Malaga, Spain, 2017.
- [3] J.Sundberg *et al.*, "Formant tuning strategies in professional male opera singers," *J. Voice*, vol.27 No.3, 278-288, 2013.
- [4] J.Dalton *et al.*, "GloRi - the Glottal Research Instrument," in *Proc. Speech Prosody*, 944-948, 2014.
- [5] P. Mokhtari, H. Ando, "Iterative Optimal Preemphasis for Improved Glottal-Flow Estimation by Iterative Adaptive Inverse Filtering," in *Proc. Interspeech 2017*, 1044-1048, 2017.
- [6] 古井貞熙, "新音響・音声工学," (近代科学社, 東京, 2006), pp-99-107,126-131.
- [7] J.D.Markel, A.H.Gray, "音声の線形予測," (コロナ社, 東京, 1980), pp.2-15,265-266.
- [8] P. Alku, "Glottal wave analysis with Pitch Synchronous Iterative Adaptive Inverse Filtering," *Speech Communication*, vol.11, no.2, 109-118, 1992.