

# 音場伝達系における音響信号のエンベロープ推定法の評価

正員 広林 茂樹\* 非会員 寺島 洋行\* 正員 山淵 龍夫\*

## Evaluation of Envelope Estimation Method of Acoustic Signal in Acoustic Field.

Shigeki Hirobayashi\*, Member, Hiroyuki Terashima\*, Non-member, Tatsuo Yamabuchi\*, Member

Blind convolution and deconvolution are fundamental signal processing techniques in acoustics. Recently, many blind techniques have been proposed without measuring the impulse response. The purpose of this study is to evaluate blind envelope estimations involving the carrier modeling of the source signal and room impulse response. We compare the envelope convolution method and power envelope convolution method on an experimental basis. The result of the envelope estimation obtained by the power envelope convolution method is better than that obtained by the other method. In the reverberant envelope estimation experiment, the envelope distortion can be reduced by about 6 dB.

キーワード：音場，エンベロープ，コンボリューション，デコンボリューション，信号処理

Keywords: acoustic field, envelope, convolution, deconvolution, signal processing

### 1. はじめに

近年、信号分離や残響抑圧への応用で盛んに研究されているブラインド処理では、音響信号のエンベロープ伝達系のモデル化が試みられている。これらのモデル化は、音場の順問題や逆問題で、それぞれ観測信号や音源信号などを対象信号とし、それらを簡易的に推定することを目的にしている。音場における入出力信号のエンベロープ（包絡線）の定式化においては、これまで主として J. Mourjopoulos 等が提案するエンベロープ畳み込みモデル<sup>(1),(2)</sup>や、著者等が提案するパワーエンベロープ伝達系<sup>(3),(4)</sup>などが提案されている。エンベロープ畳み込みモデルでは、音源信号と室内インパルス応答に対し、それぞれ複合正弦波を搬送波とした変調複合正弦波信号でモデル化している。パワーエンベロープ畳み込みモデルでは、それぞれ無相関な雑音信号を搬送波とした変調雑音信号でモデル化している。そこで、両方法に対し未だ実験的に評価は行われていないため、本報告では、これらを定量的に比較するとともに、各エンベロープ伝達系の特性を明らかにする。

### 2. 変調複合正弦波と変調雑音モデル

#### (2.1) エンベロープ畳み込み形モデル

J. Mourjopoulos 等<sup>(1)</sup>は、1本の正弦波信号をゆったりと変化するエンベロープで変調した信号で音源  $x(t)$  と室内インパルス応答  $h(t)$  の信号をモデル化している。さらに、複数の正弦波に拡張し、線形場における定式化を試みている。その結果、観測信号  $y(t)$  は

$$y(t) = \sum_{k=1}^N e_{xk}(t) \cos P_{xk}(t) * e_{hk}(t) \cos P_{hk}(t) \dots\dots\dots (1)$$

であらわされる。ただし  $e_{xk}(t)$ ,  $e_{hk}(t)$  はそれぞれ音源信号と室内インパルス応答のエンベロープ、 $N$  はスペクトル本数である。すなわち推定された観測信号のエンベロープ  $e_y(t)$  は次式のように導出された。

$$e_y(t) = \frac{1}{2} e_x(t) \cdot e_h(t) \dots\dots\dots (2)$$

#### (2.2) パワーエンベロープ畳み込み形モデル

また、既に著者らは、雑音を搬送波とした変調雑音モデルで各信号をモデル化し、観測信号のエンベロープを次式のように導出した<sup>(3),(4)</sup>。

$$e_y(t) = \sqrt{e_x(t)^2 \cdot e_h(t)^2} \dots\dots\dots (3)$$

以上のように、残響場においては、主に2つのエンベロープ推定法があり、これらを定量的に評価する。

### 3. 実験

両手法は広帯域処理を対象としているため、音源と室内インパルス応答の信号を複合正弦波とし、各手法で観測信号のエンベロープ  $e_y(t)$  を計算して正弦波の数に対して推定誤差を調べた。まず、音源と室内インパルス応答を次式のように定める。

$$x(t) = (1 + \cos(2\pi f_c t)) \cdot \sum_{k=1}^N a_k \cos(2\pi f_k t + \phi_k) \dots\dots\dots (4)$$

$$h(t) = e_h(t) \times \sum_{k=1}^N a_k \cdot \cos(2\pi f_k t + \phi_k) \dots\dots\dots (5)$$

なお、各正弦波の振幅  $a_k$ 、周波数  $f_k$ 、位相差  $\phi_k$  はそれぞれ独立に一様乱数で定めた。また、室内インパルス応答のエンベロープ  $e_h(t)$  は、コンサートホールにおいて 8kHz サンプ

\* 富山大学工学部  
〒930-8555 富山市五福 3190  
Faculty of Technology, Toyama University,  
3190 Gofuku, Toyama 930-8555

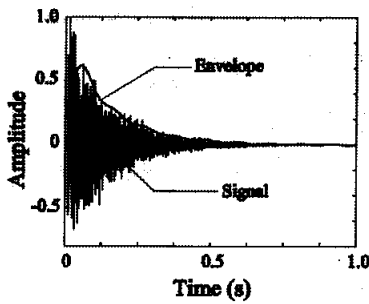


図1 室内インパルス応答とそのエンベロープ

Fig. 1. The waveform and the envelope of room impulse response.

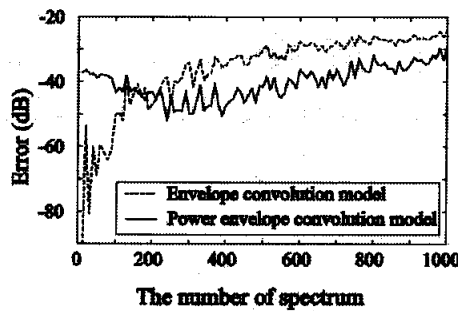


図2 複合正弦波におけるエンベロープの推定誤差

Fig. 2. The errors of the estimated envelope using sine-waves.

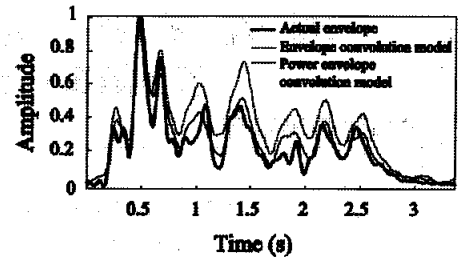


図3 実信号のエンベロープと各手法における推定エンベロープ

Fig. 3. Actual and estimated envelopes by each method.

リングで計測された室内インパルス応答  $h(t)$  を用い、ヒルベルト変換後の絶対値からエンベロープを抽出し、さらにLPFにより10Hz以下に帯域制限を行った。インパルス応答の残響時間は約1秒である。実際に抽出したエンベロープを図1に示す。次に、音源と室内インパルス応答を実際に畳み込んで、観測信号  $y(t)$  を計算し、インパルス応答のエンベロープを抽出したときと同じように観測信号のエンベロープ  $e_y(t)$  を抽出する。エンベロープ推定では式(2)を、パワーエンベロープ推定では式(3)を用いて観測信号を推定する。評価式は、実際のエンベロープ  $e_y(t)$  と各推定信号  $\hat{e}_y(t)$  の差信号の2乗平均値を用いた<sup>(3),(4)</sup>。

$$I_p = 10 \times \log_{10} \frac{1}{T} \int_0^T (e_y(t) - \hat{e}_y(t))^2 dt \quad (6)$$

ただし、 $T$ は評価時間とする。1,000本までの複合正弦波の計算結果を図2に示す(10本間隔)。点線がエンベロープ畳み込みモデル、実線がパワーエンベロープ畳み込みモデルである。150本未満の複合正弦波とした搬送波であれば、エンベロープ表記の近似は残差が小さいが、それ以上の正弦波数になるとパワーエンベロープ表記の近似が有効であることがわかる。特に300本以降では約10dBの有効性が確認できる。

また、実音声においても図3で評価実験を行った。処理例は約3秒の女声の残響音声のエンベロープを推定したものである。太線が実際のエンベロープ、点線がエンベロープ表記、実線がパワーエンベロープ表記である。実際の各音声信号には多くの正弦波が含まれているため、やはりパワーエンベロープ表記の推定法がよく近似できることがわかる。表示例では残差約6dBのパワーエンベロープの有効性が確認できた。

図4にその観測音声エンベロープとその推定エンベロープの散布図を示す。エンベロープ畳み込みモデルを点印、パワーエンベロープ畳み込みモデルを○印で図示した。パワーエンベロープ畳み込みモデルが基準線近傍に分布し、観測音声エンベロープに近いことがわかる。

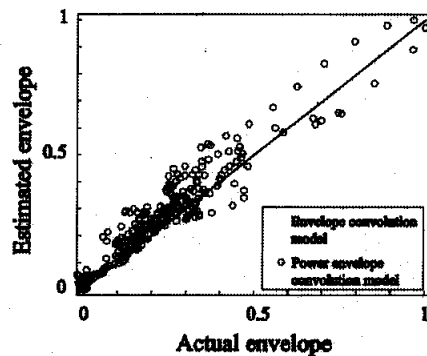


図4 実信号のエンベロープと各手法における推定エンベロープの散布図

Fig. 4. Distribution of actual and estimated envelopes.

#### 4. むすび

本報告では、音場における信号の簡易表現実現のために、エンベロープ畳み込みモデルとパワーエンベロープ畳み込みモデルの比較検討を変調複合正弦波と実音声を用いて行った。複合正弦波を搬送波とした実験においては、およそ150本以上の正弦波で構成された信号で10dB程度パワーエンベロープ畳み込みモデルの優位性が確認できた。また、実音声でも音場におけるパワーエンベロープ畳み込みモデルによる推定で6dB程度の有効性を確認した。

今後は、残響時間の変化や言語系の違いによる効果など調べる予定である。(平成15年3月17日受付)

#### 文 献

- (1) J. Mourjopoulos and J.K. Hammond: "Modelling and Enhancement of Reverberant Speech using an Envelope Convolution Method", *Proc. IEEE ICASSP 1983*, pp.1144-1147 (1983)
- (2) T. Lanthans and H.W. Strube: "speech enhancement by nonlinear multiband envelope filtering", *Proc. IEEE ICASSP*, Vol.1, pp.156-159 (1982)
- (3) 広林茂樹・野村博昭・小池恒彦・東山三樹夫: 「パワーエンベロープ伝達関数の逆フィルタ処理による残響音声の回復」, *電通学論誌 A, J81-A*, 10, pp.1323-1330 (1998-10)
- (4) 広林茂樹・山瀬龍夫: 「帯域分割を用いたパワーエンベロープ逆フィルタ処理の残響抑圧効果」, *電通学論誌 A, J83*, 8, pp.1029-1033 (2000-8)