技術報告 43.72.Ar

日本音響学会誌 46 巻 3 号 (1990)

解析信号による音声信号零交差の考察*

荒井隆行・吉田裕一**
(上智大学理工学部)

(平成元年4月17日受付)

内容梗概 音声信号分析で用いられる特徴量の一つである零交差数について，音声信号を解析信号として扱うことによって考察する。解析信号は包絡線と位相関数によって表現されるが，その非最小位相性と零交差数との間に相関関係があると考えられ，また解析信号の非最小位相性は解析信号の零点の配置に依存することが知られている。そこで本論文では音声信号の解析信号表現よりその零点を求め，零交差数との比較を行った。そして，非最小位相性を定量化するため，新しく「非最小位相値」というものを考え，母音を用いた実験において零交差数との比較を行った。その結果，両者に強い相関があることが確認された。

1. はじめに

零交差数 (zero crossing rate) は音声信号分析で用いられる特徴量の一つで，有声部・無声部・無音部の識別に有効であるなど，その利用価値は大きい。また，ハードウェアにおいてもソフトウェアにおいても，その実現は簡単である。音声信号のような複雑な信号の零交差を完全に理論付けるのは難解であるが，零交差を用いる上で少なくとも零交差が信号におけるどの情報を表現しているかを知る必要性がある。そこで音声信号を解析信号[1] (analytic signal) で表現し，その位相情報に着目することにより，零交差を捉えてみた。解析信号を包絡線と位相に分けて考えると，包絡線は同じでも，位相が違う解析信号はいくらでも存在する。その中でも位相が最小であるとき (最小位相)，位相は包絡線から一意に決定される。解析信号の最小位相については，複素零点 (complex zeros) の概念を用いて説明されることが報告されており[2)~4)]，位相が最小位相より大きくなると (非最小位相)，次第に零交差を生じるようになることが示されている。

このことから，本論文では解析信号の非最小位相性と実際の音声信号の零点について調べてみた。更に，解析信号の最小位相性を表す特徴量として新しく“非最小位相値 (non-minimum-phase value)”という量を導入し，日本語5母音について零交差数との相関を調べた。

* Study on zero-crossings of speech signals by means of analytic signal.

** Takayuki Arai and Yuichi Yoshida (Faculty of Science and Technology, Sophia University, Tokyo, 102)

2. 最小位相

最小位相の概念は，線形システム理論の中でも扱われる。しかし，ここで扱う最小位相は解析信号の最小位相であり，線形システム理論のものとは直接的に結びつくものではない。解析信号における最小位相の概念については，Voelcker[2] によって論じられている。以下概略を述べる。

ある実信号 $s(t)$ に対して，その解析信号 $m(t)$ は，

$$m(t)=s(t)+j\hat{s}(t) \tag{1}$$

$$=|m(t)|\exp j\phi(t) \tag{2}$$

と表される。ここで，$\hat{s}(t)=H[s(t)]$ は $s(t)$ のヒルベルト変換 (Hilbert transform) 又は90度移相 (90-degree phase shift) で，$|m(t)|$ は包絡線 (envelope)，$\phi(t)$ は位相関数 (phase function) である。

解析信号のフーリエ級数 c_n は負の周波数で零になるので，帯域制限された $m(t)$ は基本角周波数を ω_0 とすると，

$$m(t)=\sum_{n=0}^{N}c_n\exp jn\omega_0 t \tag{3}$$

$$=\prod_{i=1}^{N}r_i(t) \tag{4}$$

ここで，a_i, θ_i を実数とすると，

$$r_i(t)=1-a_i\exp j(\omega_0 t-\theta_i) \tag{5}$$

時間 t を複素量 $z=\tau+j\sigma$ で置き換えたとき，$m(z)=0$ を満足する $m(z)$ の零点は $r_i(z)$ の零点 z_i の集合と考えられ，

$$z_i=\frac{1}{\omega_0}[2\pi k+\theta_i+j\ln(a_i)],\quad k\in \text{整数} \tag{6}$$

となり，$2\pi/\omega_0$ の周期で繰り返し出現する。

もし零点が z-平面の下側半平面 (LHP, lower half

plane) 上にのみ存在するとき，すなわち z-平面の上側半平面 (UHP, upper half plane) 上に一つも存在しないとき，最小位相 (MP, minimum phase) といい，次式を満足することが知られている[2]。

$$\phi'(t)=H[\ln'|m(t)|] \tag{7}$$

$$\ln'|m(t)|=-H[\phi'(t)] \tag{8}$$

ただし，式中の ′ (ダッシュ) は時間 t に対する微分を表す。

一般に解析信号の零点は UHP 上にも存在する。このとき非最小位相 (NMP, non-minimum-phase) といい，次式を満足することが知られている[2]。

$$\phi'(t)=H[\ln'|m(t)|]+\sum_{iu}\frac{2r_i\sigma_i}{(\tau_i-t)^2+\sigma_i^2} \tag{9}$$

$$\ln'|m(t)|=-H[\phi'(t)]+\sum_{iu}\frac{2r_i(\tau_i-t)}{(\tau_i-t)^2+\sigma_i^2} \tag{10}$$

ここで，$z_i=\tau_i+j\sigma_i$ は UHP 上の i 番目の零点，r_i はその零点の位数，シグマは UHP 上のすべての零点について総和をとることを示す。式(9)から明らかなように，非最小位相の場合には第2項に示す過剰位相が加えられる。

最小位相の場合には，式(6)からすべての $a_i(i=1,2,\cdots,N)$ について $0<a_i<1$ となり，非最小位相の場合には，少なくとも一つの a_i について $a_i>1$ となる。

ここで，実信号 $\mathrm{Re}[r_i(t)]=1-a_i\cdot\cos(\omega_0t-\theta_i)$ の零交差について考えてみると，$0<a_i<1$ のときは零交差を生じず，$a_i>1$ のとき零交差を1周期につき2回生じることが分かる。すなわち，この場合非最小位相のとき零交差を生じ，一つの UHP 上の零点に対し2回の割合で零交差が生じる (ただし，解析信号が最小位相であっても，近接する零点によって実信号が零交差を生じることがあり，これらのことについては今後検討していきたい)。

3. 実際の音声信号における最小位相と零交差

実際の音声信号を用いて，その零点を求めてみた。ある成人男性の母音 /a/, /i/ の1周期分を切り出し，$m(z)=0$ を解くことによって得られた零点を **Fig. 1**, **Fig. 2** に示す。

一般に音声信号のような複雑な信号において，その解析信号の零点と原波形の零交差を理論的に関係づけることは難解である。しかしこれらの図を見ると，/a/, /i/ とも一つの UHP 上の零点により2回の零交差を生じていることが分かる。

4. 非最小位相値と零交差数

4.1 非最小位相値の定義

信号を厳密に解析するならば，すべての零点を求めなければならないが，実際上は多くの計算時間を要する。

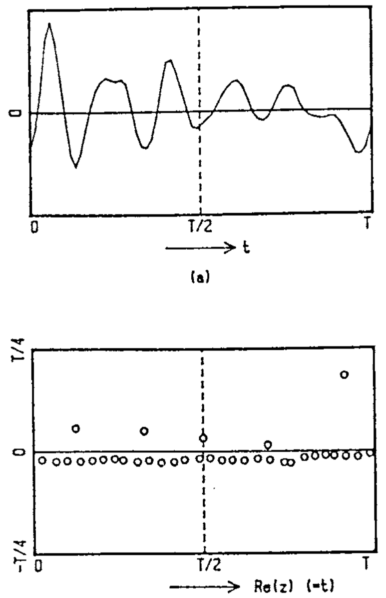

Fig. 1 Complex zeros of a Japanese vowel /a/. (a) One period of the waveform, (b) The complex zeros in the period ($0\leq\mathrm{Re}(z)(=t)<T$)

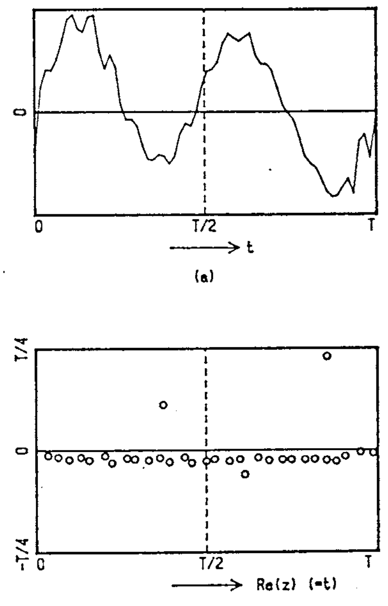

Fig. 2 Complex zeros of a Japanese vowel /i/. (a) One period of the waveform, (b) The complex zeros in the period ($0\leq\mathrm{Re}(z)(=t)<T$)

そこで，式(7)と(9)の積分値の差，すなわち，

$$\Phi=\frac{1}{T}\int_{t_0}^{t_0+T}\{\phi'(t)-H[\ln'|m(t)|]\}dt \tag{11}$$

$$=\frac{1}{T}\int_{t_0}^{t_0+T}\left\{\sum_{iu}\frac{2r_i\sigma_i}{(\tau_i-t)^2+\sigma_i^2}\right\}dt \tag{12}$$

をもって，非最小位相値 Φ(non-minimum-phase value)

Table 1 Correlation coefficient between zero-crossing rate and non-minimum-phase value for Japanese five vowels.

Vowel	Correlation coefficient
/a/	0.851
/i/	0.888
/u/	0.956
/e/	0.933
/o/	0.919
Average	0.909

を定義する。

この式は，複素時間 (z) 平面における UHP 上の $\ln'|m(z)|$ の極，(すなわち $m(z)$ の零点) に関する実軸上1周期当たりの留数の総和を表し，厳密な対応ではないが UHP 上の $m(z)$ の零点数を表すものと考えられる。

4.2 非最小位相値と零交差数

実際の音声信号を用いて非最小位相値と零交差数との比較を行った。実験に使用した音声試料は，成人男性27人による日本語5母音である。この音声試料をカットオフ周波数 3kHz の LPF (low pass filter) に通した後，サンプリング周波数 10kHz，量子化ビット数12ビットで A/D 変換を行った。そして，各音声試料の 90ms における単位サンプル当たりの平均零交差数と平均非最小位相値を求めた。両者の相関係数を **Table 1** に示す。ここで零交差は，隣接するサンプルの符号が異なるときに生じるものとした。

Table 1 に示されるとおり，どの母音も 0.9 近くの強い相関を示していることが分かる。このことから，零交差数はここで定義した非最小位相値を示すものと言える。

5. おわりに

音声信号分析で用いられる特徴量の一つである零交差数について解析信号による考察を行い，実信号を解析し，比較した。この結果，著者らの定義する非最小位相値と零交差数が強い相関を持つことが明らかとなった。今後音声認識への応用を目指し，更に解析信号手法による音声信号解析を進めていきたい。音声認識上，上記二者のいずれがより有効なパラメータであるかは，今後多数例について検証する必要があるが，少なくとも零交差の意味づけは得られたものと考えている。

文　献

1) A. V. Oppenheim and R. W. Schafer (伊達 玄訳)，ディジタル信号処理 (下) (コロナ社，東京，1978)，p. 22.
2) H. B. Voelcker, "Toward a unified theory of modulation, Part I: Phase-envelope relationships," Proc. IEEE 54, 340-353 (1966).
3) H. B. Voelcker, "Toward a unified theory of modulation, Part II: Zero manipulation," Proc. IEEE 54, 735-755 (1966).
4) S. Doshita, "Studies on the analysis and recognition of Japanese speech sounds," Ph. D. Thesis, Kyoto Univ. (1965).