

論文 / 著書情報
Article / Book Information

論題(和文)	二乗器によるピッチ抽出について
Title(English)	
著者(和文)	古井貞熙, 石井 泰
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1969年講演論文集, , , pp. 367-368
Citation(English)	, , , pp. 367-368
発行日 / Pub. date	1969, 10

古井 貞熙 石井 泰 (東大宇宙研)

音声波においてしばしば見られるような 基本波を欠いた波を二乗器に通すと、ピッチを抽出することができることは よく知られた事実である。しかし、種々の非線形変換が考えられる中で、なぜ二乗を選ぶのがよいのか、ピッチ抽出に關してどの程度の効果が期待できるのか、その際の向題点は何か、ということに關する考察は今まであまりなされていないようである。今回の報告では それらの向題に關する考察と、二乗器を用いた簡単なピッチ表示装置の試作について述べる。

1. 考察 信号波を(1)式のように表わすと、これをP乗したものは一般に(2)式のように書くことができる。

$$x(t) = \sum_{n=-N}^N C_n \exp(jn\omega t) \quad C_n: \text{complex } (C_0=0 \text{ とする}) \quad (1)$$

$$x^P(t) = \underbrace{\sum_{k=-N}^N \sum_{l=-N}^N \cdots \sum_{m=-N}^N}_{P_1} \underbrace{C_k \cdot C_l \cdots C_m}_{P_2} \exp\{(k+l+\cdots+m)j\omega t\} \quad (2)$$

(2)式で $k+l+\cdots+m$ が ± 1 となるものが 変換後に得られるピッチに対応し、 ± 1 となるものが 1 次高調波に対応する。一般に、 ± 1 となる $(k, l, \cdots, m)$ の組合わせの数が ± 2 あるいはそれ以上の値をとる組合わせの数に比べて圧倒的に大きくなることが ピッチ抽出という目的のためには好ましいことは言うまでもない。そこで、(1)式及び(2)式から基本波を除いた場合についてそれぞれ、二乗、三乗、四乗、五乗を行なった場合に得られる上記の組合わせの数を求めてみたところ、二乗が最もよい結果を示し、四乗、三乗、五乗の順に効果が悪くなり、二乗と四乗の間には期待される効果においてかなりの差のあることが確かめられた。このことから、一応、二乗あるいは二乗で近似される変換 ($|x|$, $\sqrt{|x|}$, $|x|^3$ etc. Legendre 関数によって展開できる) が ピッチ抽出のためには最も好ましいと考えられる。

(2)式で $P=2$ 即ち二乗の場合は次のようになる。

$$\begin{aligned} x^2(t) &= \sum_{n=-N}^N \sum_{l=-N}^N C_n \cdot C_l \exp\{(k+l)j\omega t\} \\ &= \sum_{n=-N}^N |C_n|^2 + \sum_{r=1}^{2N} \left\{ \left(\sum_{n=-N+r}^N C_n \cdot C_{-n+r} \right) \exp(rj\omega t) + \left(\sum_{n=-N-r}^N C_n \cdot C_{-n-r} \right) \exp(-rj\omega t) \right\} \quad (3) \end{aligned}$$

このパワースペクトル及び抽出されるピッチのパワースペクトルはそれぞれ、

$$\Phi_{x^2}(f) = \left(\sum_{n=-N}^N |C_n|^2 \right) \delta(f) + \sum_{r=1}^{2N} \left\{ \left| \sum_{n=-N+r}^N C_n \cdot C_{-n+r} \right|^2 \cdot \delta\left(f \pm \frac{r\omega}{2\pi}\right) \right\} \quad (4)$$

$$\Phi_{x^2}(f_0) = \left| \sum_{n=-N+1}^N C_n \cdot C_{-n+1} \right|^2 \quad (5)$$

これから明らかなように、二乗したときに抽出されるピッチのパワーは もとの信号波の各フーリエ成分の振幅と相互の位相関係によって左右されるが、計算したと

* Pitch Extraction by the Squarer. By Sadaoki Furui and Yasushi Ishii
(Institute of Space and Aeronautical Science, University of Tokyo)

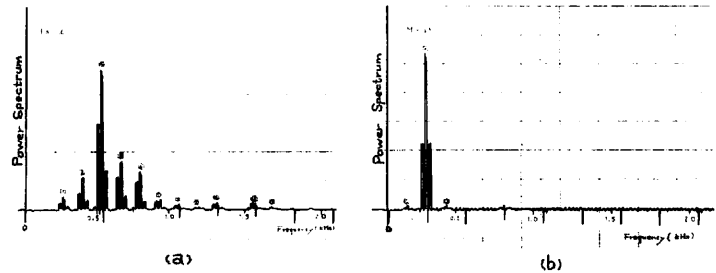
$$8(N-1)$$

ここで、この分布は自由度2の χ^2 分布を横軸方向へ約 $(2 \rightarrow \log_2(N-1))$ 倍引伸ばしたようなものになり、 N がある程度大きければ、ピッチのパワーが非常に小さくなってしまふという恐れはほとんどないと考えられる。

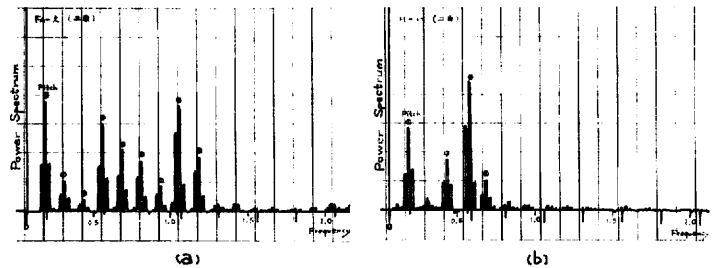
又、信号波に白色雑音に加え合われている場合についても、実験と考察を行なったところ、目で見ただけでは周期性が全くわからない程度に波形が歪んでいても十分ピッチを抽出できることが確かめられた。

2. 実験

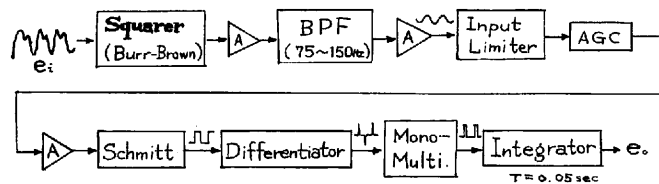
[Fig. 1.] Power spectra of human voice whose fundamental component has very small power.
(a) : "e" of Mr. A. F.
(b) : "i" of Mr. S. M.



[Fig. 2.] Power spectra of the squared wave
(a) and (b) correspond to those of Fig. 1.



[Fig. 3.] Block diagram of the pitch indicator



[Fig. 4.] Results of application of the arrangement to actual speech.

