

## E-3 ヘリウム音声と了解性と自然性の改善

——自己相関関数を用いたヘリウム

音声復元方式 SPAC-H ——

郵政省電波研究所

鈴木 誠 史

ヘリウム音声の主要なひずみは、(1)ヘリウム空気の音速（大気の音速に対する比を $\alpha$ とする、 $\alpha=1.5\sim 2.7$ ）に比例した周波数スペクトル $F(\omega)$ の拡大、(2)スペクトルの拡大率が低い周波数ほど大、(3)無声子音レベルの低下、(4)有声音源（喉頭原音）の基本周波数 $F_0$ の上昇（上昇率を $\beta$ とする、 $\beta=1.1\sim 1.5$ ）、などである。なお、(2)、(3)は空気密度の増加に起因し、(4)は心理的要因による。これらの中で、了解性の低下に影響を与えているのは(1)で、(3)と(4)の効果は少ない。また、大陸棚程度の深度では(2)の効果は無視できる。一方、音声の自然性には $F_0$ が大きく貢献し、特に話者の同定には $F_0$ の平均値が最大の要因と考えられている。

したがって、ヘリウム音声の了解性改善のためには $\alpha$ によるひずみを修正する $F(\omega)$ の圧縮が、個人性の回復には $\beta$ を補償する $F_0$ の圧縮が必要である。ところで、先に報告した切出し伸長方式は、 $F(\omega)$ と $F_0$ を同じ比率で圧縮している。他の方式では、 $F_0$ を保存し $F(\omega)$ のみを圧縮するものが多い。また、 $F(\omega)$ の変換に際してひずみを生じたり、雑音のために品質の劣化を生じるなどの問題点があり、評価の定まった方式はない。

これらの点を考慮すると、下記の条件を満足する復元方式が必要である。

1. ひずみのない $F(\omega)$ の圧縮。
2.  $F(\omega)$ と独立した $F_0$ の圧縮。
3. SN比の低い音声でも安定に動作。
4. 装置の製作が容易。

ここでは、短時間自己相関関数 $\rho$ の性質を利用して、上記の条件を満足する復元方式を提案する。

自己相関関数の $\varphi$ には、(a)入力信号の周波数成分が保存される、(b)計算の原点が変わっても $\varphi$ の形は変わらない、(c)高調波の位相が一致するため周期性の検出が容易である、(d)ランダム雑音は $\varphi(0)$ 附近の成分になる、……などの性質がある。音声のように時間的に変動する信号に対しては、積分時間を限定した短時間自己相関関数 $\rho$ が適用され、 $\rho$ も $\varphi$ と同様の性質があるとみなされる。

波形信号を $f(t)$ とすると、その $\rho$ の1周期を次々に接続した波形が、 $f(t)$ と同じ周波数成分から成り、同じように聞こえる（性質(a)）こと、その他の $\rho$ の性質を積極的に利用する。その処

理を図1によって説明する。(図1)  $f(t)$ をヘリウム音声波形とする。その時刻  $t_1$  に関して  $\rho_1(\tau)$  を計算する。 $\rho_1$ の極大値からその周期  $T_1$  を決定する(性質(c)の利用)。なおこの  $T_1$  は一般に  $1/F_0$  に相当する。次に入力に加わっている雑音の影響を軽減するために  $\rho_1(0)$  の付近を避け、 $\rho$  の 1 周期の  $c$  倍 ( $0 < c \leq 1$ ) の波形を切り出し、これを  $kT_1$  ( $k \geq 1$ ) の時間に伸長する。次に、時刻  $t_2 = t_1 + kT_1$  として  $\rho_2(\tau)$  を求める。このとき、 $k$  によって  $\rho$  が大きく変ることはない(性質(b)の利用)。 $\rho_2$  について周期  $T_2$  の決定、 $cT_2$  の波形の切り出し伸長を行う。 $t_3 = t_2 + kT_2$ ,  $t_4 = t_3 + kT_3$  ……について同じ操作を繰り返し、出力  $g(t)$  には、 $cT_1$ ,  $cT_2$ ,  $cT_3$  …… の波形が接続したものが得られる。この  $g(t)$  は  $f(t)$  のスペクトル包絡を  $c/k$  に、 $F_0$  を  $1/k$  に変換したものである。この方式を SPAC-H (Helium Speech Processing system by use of Autocorrelation Coefficients) とよぶことにする。ヘリウム音声を  $k = \beta$ ,  $c = k/\alpha$  の SPAC-H で処理すると、正常な音声に近い  $F_0$  とスペクトル包絡の音声に変換される。

SPAC-H の系を計算機シミュレーションで構成した。50  $\mu\text{sec}$  毎に標本化したヘリウム音声 (10.6 atm) の信号を  $a_i$  で表わしたとき、 $\rho$  は次式で計算される。

$$\rho(j) = \sum_{i=1}^{400} a_i \cdot a_{i+j} \quad j = 0, 1, 2, \dots, 255$$

$\rho(j)$  の  $cT$  に相当する標本を、 $50k/c \mu\text{sec}$  の標本周期で配列することで上記の操作を行う。 $k = 1.3$  ( $\beta$ に近い),  $k/c = 2.3$  ( $\alpha$ と同じ) として処理したところ、了解性だけでなく、自然で話者の同定も容易な音声になった。このスペクトログラムを図2に示す。また、深さ 800 feet で発声された音声 (英語) を  $k = 1.2$  (但し  $\beta$  値は不明),  $k/c = 2.7$  ( $\alpha$ と同じ) で処理したところ、内容が理解できるようになり、自然音声になった。

SPAC-H のアルゴリズムは簡単で、ハードウェアもデジタル回路により容易に製作できる。また SN 比 0 dB でも安定に動作するだけでなく、雑音レベルを低減する効果もあるので実用機として役立つものと期待される。

今後は、SPAC-H の定量的評価とハードウェアの製作に努力する予定である。

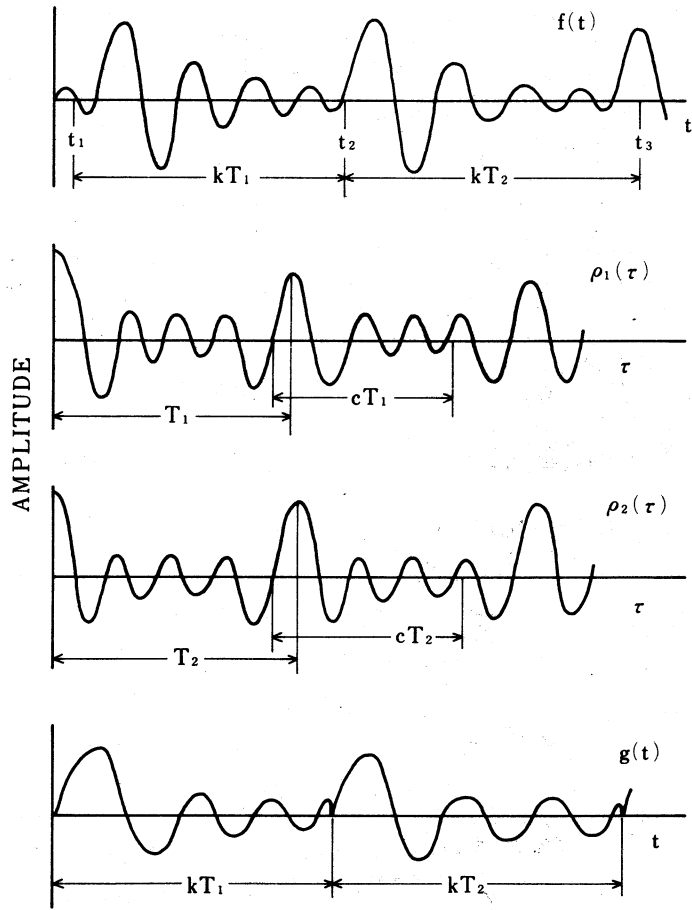


図1 SPAC-Hの原理説明図

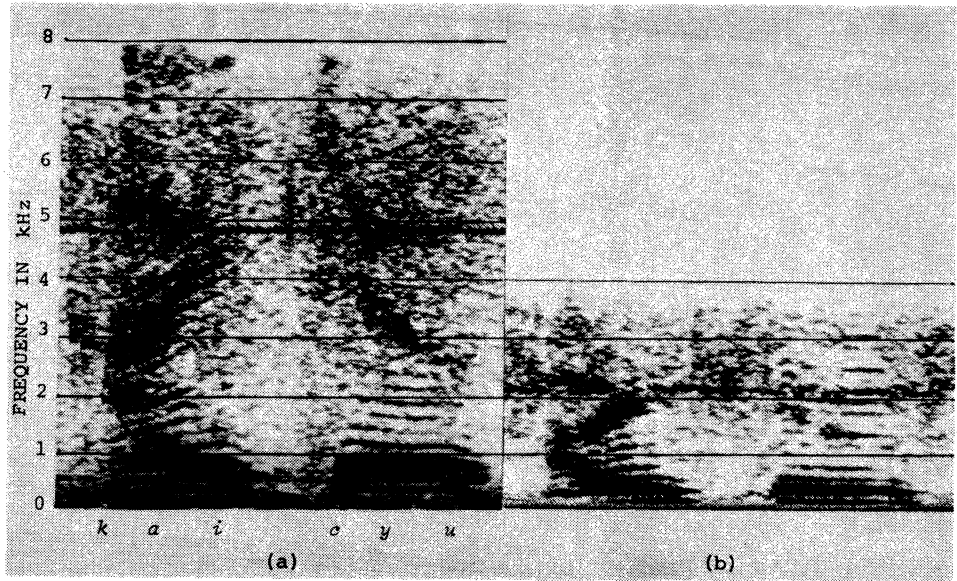


図2 SPAC-Hの実験例のスペクトログラム

(a) ヘリウム音声 (10.6 atm)

(b) (a)の $F(w)$ を $1/2.3$ に、 $F_0$ を $1/1.3$ に変換した音声

〈質問〉 埼玉医科大学 梨本一郎

本装置を、よりcompactに出来る可能性があるか。

〈答〉 郵政省電波研究所 鈴木誠史

現在のところでも、手で持てる程度の大きさにすることは可能である。又最終的には弁当箱程度になると考えている。