

研究テーマ：実環境下での音声認識における騒音抑制法	
研究代表者（職氏名）：経営情報学科教授・生田 顕	連絡先 (E-mail等)：ikuta@pu-hiroshima.ac.jp
共同研究者（職氏名）：教授・肖 業貴， 准教授・小川仁士， 教授・韓 虎剛， 教授・田中稔次朗	

1. 目的と概要

デジタル情報技術の進展に伴い、音声認識システムが開発され、様々な分野に応用されている。例えば、工場などにおける施設・設備巡回点検修理業務や工事現場の記録・状況報告などハンズフリー業務の支援に応用されている。このような実環境下での音声認識においては、周囲環境からの騒音対策が不可欠となる。

従来から、音声認識における騒音対策法はいくつか提案されてきたが、それらは単一マイクロホンに基づく手法とマイクロホンアレーに基づく手法に大きく分類される。騒音抑制システムを実現する際、後者の手法では複数の騒音源が存在するとき騒音源数以上のマイクロホンが必要となるため、設備が大掛かりになり、かつ騒音源の数が既知でなければならない。したがって、単一マイクロホンに基づく手法が、騒音抑制システムの実装の観点から有利である¹⁾。多くの単一マイクロホンによる多くの騒音抑制法は、ピッチ検出や騒音スペクトルの推定が不要であるなどの利点により、線形予測に基づいている。一方、この手法は線形モデルのパラメータを音声信号に基づき推定する必要があるが、実際の応用においては、音声信号に騒音が入ったデータのみが得られることから、何らかの方法を用いてパラメータを推定することが必要となる。

本研究では、工場など生産の現場における施設・設備巡回点検修理業務や工事現場の記録・状況報告などハンズフリー業務の支援に音声認識を応用することを目的とし、周囲環境からの騒音対策技術に関し新たな実用的手法を開発する。具体的には、騒音源の方向が特定できない高騒音環境下において、騒音スペクトルの事前情報やピッチ検出が不要な、単一マイクロホンのみを用いる騒音抑制法について検討を行った。このような音声認識における雑音抑制法として、カルマンフィルタを適用した手法が提案されている²⁾。しかし、カルマンフィルタは雑音のガウス性や白色性を前提としており、しかも雑音の分散情報を必要とする。多くの騒音変動は非ガウス型の複雑な形態を示し、一般には非白色で、その統計量は実際には未知である場合が多い³⁾⁴⁾。したがって、本研究では、実環境下での音声認識を可能にするため、騒音変動のガウス性や白色性などの前提を必要とせず、かつ騒音の統計情報が未知の場合でも適用できる新たな騒音抑制法を研究開発する。

具体的には、音声信号の数時点間にわたる時系列と騒音混入下での音声との間の相関情報を反映した結合確率分布に着目した。そして、各変数に関する低次はもちろん高次の各種相関情報をも展開係数に反映した級数展開表現による非線形の音声時系列モデルを導出した。さらに、この非線形モデルにおける各展開係数を未知パラメータとみなし、これらのパラメータを音声信号と同時に推定する手法を提案した。その際、確率のベイズ原理に基づき、騒音の影響を抑制した音声信号の検出のための新たなアルゴリズムの研究開発を行った。さらに、本手法の有効性は、統計量未知の騒音混入下での音声信号の推定に本理論を適用することにより実験的にも確認した。

2. 音声信号検出のための騒音抑制アルゴリズムの開発

周囲の騒音が存在する実環境において、 k 時刻における音声信号を x_k 、統計量未知の騒音が入った観測値を y_k とする。 y_k の逐次観測に基づき、 x_k を推定するための信号情報処理法を導出する。

まず、 x_k と y_k の関係を、各種相関情報を各展開係数に反映した次の条件付確率分布の展開表現で捉

える。

$$P(y_k|x_k) = P(y_k) \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} E_{rs} \theta_r^{(1)}(x_k) \theta_s^{(2)}(y_k), \quad E_{rs} = \langle \theta_r^{(1)}(x_k) \theta_s^{(2)}(y_k) \rangle \quad (1)$$

ただし、 $\theta_r^{(1)}(x_k), \theta_s^{(2)}(y_k)$ はそれぞれ $P(x_k), P(y_k)$ を重みとする正規直交多項式であり、シュミットの直交化法により構成される。 x_k と y_k の関係は未知であるため、(1) における展開係数 E_{rs} を未知パラメータ α_k とみなし、音声信号 x_k と同時に推定する。具体的には、ベイズ定理 $P(x_k, \alpha_k | Y_k) = P(x_k, \alpha_k, y_k | Y_{k-1}) / P(y_k | Y_{k-1})$ の展開表現に着目することにより、音声信号の平均 \hat{x}_k 、および分散 P_{x_k} に対する推定式が次のように得られる（導出過程は省略する）。

$$\hat{x}_k = \langle x_k | Y_k \rangle = \frac{\sum_{m=0}^1 \sum_{p=0}^{\infty} A_{m0p} g_{1m} \varphi_p^{(3)}(y_k)}{\sum_{p=0}^{\infty} A_{00p} \varphi_p^{(3)}(y_k)}, \quad P_{x_k} = \langle (x_k - \hat{x}_k)^2 | Y_k \rangle = \frac{\sum_{m=0}^2 \sum_{p=0}^{\infty} A_{m0p} g_{2m} \varphi_p^{(3)}(y_k)}{\sum_{p=0}^{\infty} A_{00p} \varphi_p^{(3)}(y_k)} \quad (2)$$

$$A_{mnp} = \langle \varphi_m^{(1)}(x_k) \varphi_n^{(2)}(\alpha_k) \varphi_p^{(3)}(y_k) | Y_{k-1} \rangle$$

$$\left(\begin{array}{l} g_{1m}, g_{2m} \text{ は、それぞれ } x_k \text{ および } (x_k - \hat{x}_k)^2 \text{ を直交展開したときの展開係数、} \\ \varphi_m^{(1)}(x_k), \varphi_n^{(2)}(\alpha_k), \varphi_p^{(3)}(y_k) \text{ は、それぞれの変数に対する基準分布を重みとする直交多項式} \end{array} \right)$$

ただし、 Y_k は k 時刻までの観測値の集合 $\{y_1, y_2, \dots, y_k\}$ である。さらに、展開係数 A_{mnp} は(1)式を考慮することにより、 x_k, α_k に対する予測値の関数として、次のように表され、 x_k と α_k それぞれの逐次推定が可能となる。

$$A_{mnp} = \langle \varphi_m^{(1)}(x_k) \varphi_n^{(2)}(\alpha_k) \int \varphi_p^{(3)}(y_k) P(y_k | x_k) dy_k | Y_{k-1} \rangle = \langle \varphi_m^{(1)}(x_k) \varphi_n^{(2)}(\alpha_k) \sum_{r=0}^{\infty} \sum_{s=0}^p c_{ps} E_{rs} \theta_r^{(1)}(x_k) | Y_{k-1} \rangle \quad (3)$$

$$(\varphi_p^{(3)}(y_k) = \sum_{s=0}^p c_{ps} \theta_s^{(2)}(y_k) : c_{ps} \text{ は展開係数})$$

3. 音声信号への適用

統計量未知の騒音が混入した観測値からもとの音声信号を推定するために、本手法と拡張カルマンフィルタの2つの方法を適用した。実験結果から、拡張カルマンフィルタに比べて本手法の方が、音声信号をより精度よく推定していることが確認された。

参考文献

- 1) 川村・藤井・伊藤・副井：線形予測分析に基づく騒音抑圧法，電子情報通信学会論文誌 A, Vol.J85-A, pp.415-423 (2002).
- 2) W. Kim and H. Ko: Noise Variance Estimation for Kalman Filtering of Noisy Speech, IEICE Transactions on Information and Systems, Vol.E84-D, pp.155-160 (2001).
- 3) A. Ikuta, M. O. Tokhi and M. Ohta: A Cancellation Method of Background Noise for a Sound Environment System with Unknown Structure, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E84-A, pp.457-466 (2001).
- 4) 生田・太田・益池：音環境計測における外来雑音対策と幹線交通騒音評価への適用，電気学会論文誌 C, Vol.126, pp.63-71 (2006).