

Kernel PCA を用いたロバスト特徴量抽出の検討

滝口 哲也[†] 有木 康雄[†]

[†] 神戸大学 〒 657-8501 神戸市灘区六甲台町 1-1

E-mail: †{takigu,ariki}@kobe-u.ac.jp

あらまし 本稿では, Kernel PCA (Principal Component Analysis) を用いた残響にロバストな特徴量抽出法について検討する. 画像分野では, Kernel PCA はノイズ除去などに優れた性能を発揮することが知られている [1]. 例えば, 画質の粗い画像を高画質な画像へと変換を行う Super-Resolution 等が提案されている [2]. 一方, これまで Kernel PCA を音声特徴量抽出において適用した研究も行われてきている. 例えば [3] [4] では, クリーン音声に対して MFCC 計算後, 音声のスペクトル包絡成分に対応する低次ケプストラムに対して, Kernel PCA を適用した結果が報告されている. 本稿では, ケプストラムよりもロバストな特徴量抽出法として, 離散コサイン変換の代わりに Kernel PCA を用いた手法を提案し, 残響下音声認識により有効性を示す.

キーワード kernel PCA, 特徴量抽出, 残響, 音声認識

A Study on Robust Feature Extraction Using Kernel PCA

Tetsuya TAKIGUCHI[†] and Yasuo ARIKI[†]

[†] Kobe University 1-1 Rokkodai, Nada, Kobe, 657-8501 Japan

E-mail: †{takigu,ariki}@kobe-u.ac.jp

Abstract We investigate robust feature extraction using kernel PCA (Principal Component Analysis). Kernel PCA has been suggested for various image processing tasks requiring an image model such as, e.g., denoising. Image denoising is the task of constructing a noise-free image from a noisy input image. From the point of view of a kernel PCA, image denoising can be also regarded as the same problem as image super-resolution [2]. Also, an approach for feature extraction in speech recognition systems using kernel PCA has been proposed [3] [4], where kernel PCA was applied to the low-dimension cepstrums. In this paper, we propose robust feature extraction based on kernel PCA instead of DCT for MFCC. Its effectiveness is confirmed by word recognition experiments on reverberant speech.

Key words kernel PCA, robust feature extraction, reverberation, speech recognition

1. はじめに

現在, 会議などの書き起こし, ロボットとの対話など, ハンズフリーでの音声認識機能を使用するタスクに関する要求が, 多く存在する. しかしながら, 現状のシステムではユーザがマイクロフォンから離れて発話すると, 入力音声は周囲雑音及び残響の影響を受けて認識性能が劣化してしまう. またデスクトップマイクロフォンやピンマイクロフォンを用いた場合でも, ユーザが横を向くと音響伝達特性の影響により音声がひずみ, 認識性能が劣化する場合がある.

従来, 音声の伝達経路による影響に対処する方法として, ケプストラム平均減算法 (Cepstrum Mean Subtraction: CMS) が使われている. この手法は, 例えば電話回線の影響などのように, 伝達特性のインパルス応答が比較的短い場合には有効で

あるが, マイクロフォンから離れて発話した際には, 残響の影響を受けて十分な性能が得られない. 残響成分を除去する方法として, 複数のマイクロフォンを利用し, 逆フィルタを設計して観測信号から残響成分を除去する方法 [5] [6] 等が提案されているが, 音響伝達特性のインパルス応答が, 最小位相とならない場合があり逆フィルタの設計は難しい. また使用環境下においてコストや物理的な配置状況により, 複数のマイクロフォンを設置できない場合がある. シングルマイクロフォンによる残響除去手法として, 短時間分析窓と長時間分析窓を組み合わせる残響を除去する手法 [7], 調波構造に基づく逆フィルタ設計法 [8], パワートラジェクトリー残響モデルによる残響除去 [9] 等が提案されている. これらの手法の多くは, スペクトル領域において演算が行われ, 音声認識を実行する際には, 通常ケプストラム分析が用いられている. ケプストラム分析では, 対数

スペクトルに離散コサイン変換が適用される．その後、音声のスペクトル包絡成分に対応する低次のケプストラムが抽出され、音声認識の特徴量として使われる．

本稿では、離散コサイン変換よりも、ノイズロバストな特徴量抽出法として、Kernel PCA (Principal Component Analysis) を検討する．画像の分野では、Kernel PCA はノイズ除去などに優れた性能を発揮することが知られている．例えば、画質の粗い画像を高画質な画像へと変換を行う Super-Resolution 等が提案されている [2]．一方、これまで Kernel PCA を音声特徴量抽出において適用した研究も行われてきており、例えば [3] では、クリーン音声に対して MFCC 計算後、音声のスペクトル包絡成分に対応する低次ケプストラムに対して Kernel PCA を適用した結果が報告されているが、ノイズロバスト性に関しては報告がされていない．本稿では、残響などにもロバストな特徴量抽出法として、MFCC における DCT (Discrete Cosine Transformation) の代わりに、Kernel PCA を用いた手法を提案し、残響下音声認識においてその有効性を報告する．

2. Kernel PCA による特徴抽出

2.1 提案手法

現在の音声認識システムでは、音声特徴量として MFCC (Mel Frequency Cepstral Coefficient) が広く用いられている．MFCC では、メル尺度フィルタバンクの短時間対数エネルギー出力系列に対して、離散コサイン変換 (Discrete Cosine Transformation: DCT) が適用され、ケプストラムが得られる．さらに音声のスペクトル包絡成分に対応する低次ケプストラムのみを抽出し、音声認識における特徴量として使用される．

これまでの多くのノイズ除去手法は、スペクトル領域にて演算が行われ、認識時には (対数スペクトルに対し) 離散コサイン変換が適用され、ケプストラムが求められる．そこで本稿では、よりノイズにロバストな特徴量抽出法として、離散コサイン変換の代わりに Kernel PCA を用いた手法を検討する (図 1)．スペクトル上で PCA することで、エネルギーの強い主な音声成分は低次集まり、ノイズは高次集まる．この結果、PCA によりノイズ除去が行われると期待出来る．しかしながら、PCA では線形分離可能なデータしか扱うことが出来ないため、残響や残留ノイズなどモデル表現が困難なデータに対して十分な効果が得られない場合がある．そこで線形分離不可能なデータにも対応することが可能な Kernel PCA を適用することで、ノイズ除去の効果を高める．

文献 [3] [4] では、クリーン音声に対して MFCC 計算後、低次ケプストラムに対して Kernel PCA を適用している (つまり DCT による次元圧縮後に Kernel PCA を適用している)．一方、本手法では DCT を行わずに、対数スペクトルに対して Kernel PCA を適用することにより、より有効なスペクトル情報の抽出を試みる．

2.2 Kernel PCA [10]

PCA (Principal Component Analysis) の目的は、データの本質的な構造を残しながら次元数を削減することにある．ただし、PCA はデータが非線形な構造を持つ時、有効に動作しない．

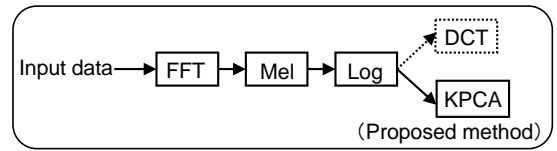


図 1 Kernel PCA による特徴抽出

Fig.1 Feature Extraction Using Kernel PCA

い．そこで Kernel PCA では、カーネル関数 Φ を用いて、元の次元よりもはるかに大きな次元へ写像を行い (データが線形表現可能な) 高次元空間において PCA が行われる．

Kernel PCA の特徴の一つは、対象に対する事前知識をカーネル関数の形で表現することにある．クリーン音声の構造を Kernel 関数で表現する事が出来れば、雑音・残響音声を特徴空間へ写像した際に、クリーン音声の構造に適した特徴量が得られ、その他のノイズ成分が除去されると期待出来る．

ここで、 d 次元観測ベクトルを \mathbf{x}_j (j はフレーム番号) とすると、共分散行列 C は、

$$C = \frac{1}{N} \sum_{j=1}^N \bar{\Phi}(\mathbf{x}_j) \bar{\Phi}(\mathbf{x}_j)^T \quad (1)$$

$$\bar{\Phi}(\mathbf{x}_j) = \Phi(\mathbf{x}_j) - \frac{1}{N} \sum_{j=1}^N \Phi(\mathbf{x}_j) \quad (2)$$

となる (N は全フレーム数)． C の固有値を λ 、固有ベクトルを \mathbf{v} と置くと、

$$\lambda \mathbf{v} = C \mathbf{v} \quad (3)$$

$$\lambda (\bar{\Phi}(\mathbf{x}_k) \cdot \mathbf{v}) = (\bar{\Phi}(\mathbf{x}_k) \cdot C \mathbf{v}), \quad k = 1, \dots, N \quad (4)$$

が得られる．また \mathbf{v} は以下のようにサンプル点の線形結合で表現出来る．

$$\mathbf{v} = \sum_{i=1}^N \alpha_i \bar{\Phi}(\mathbf{x}_i) \quad (5)$$

式 (1) と (5) を (4) に代入すると左辺は、

$$\begin{aligned} \lambda (\bar{\Phi}(\mathbf{x}_k) \cdot \mathbf{v}) &= \lambda \sum_i \alpha_i \bar{\Phi}(\mathbf{x}_k) \cdot \bar{\Phi}(\mathbf{x}_i) \\ &= \lambda \sum_i \alpha_i \bar{K}_{ki} \end{aligned} \quad (6)$$

となる．ここで、

$$\bar{K}_{ki} = \bar{\Phi}(\mathbf{x}_k) \cdot \bar{\Phi}(\mathbf{x}_i) \quad (7)$$

とした．また右辺は、

$$\begin{aligned} &\bar{\Phi}(\mathbf{x}_k) \cdot C \mathbf{v} \\ &= \bar{\Phi}(\mathbf{x}_k) \cdot \frac{1}{N} \sum_j \bar{\Phi}(\mathbf{x}_j) \bar{\Phi}(\mathbf{x}_j)^T \sum_i \alpha_i \bar{\Phi}(\mathbf{x}_i) \\ &= \bar{\Phi}(\mathbf{x}_k) \cdot \frac{1}{N} \sum_i \alpha_i \left\{ \sum_j \bar{\Phi}(\mathbf{x}_j) \bar{\Phi}(\mathbf{x}_j)^T \bar{\Phi}(\mathbf{x}_i) \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N} \sum_i \alpha_i \left[\bar{\Phi}(\mathbf{x}_k) \cdot \left\{ \sum_j \bar{\Phi}(\mathbf{x}_j) \bar{\Phi}(\mathbf{x}_j)^T \bar{\Phi}(\mathbf{x}_i) \right\} \right] \\
&= \frac{1}{N} \sum_i \alpha_i \sum_j \{ \bar{\Phi}(\mathbf{x}_k) \cdot \bar{\Phi}(\mathbf{x}_j) \} \{ \bar{\Phi}(\mathbf{x}_j) \cdot \bar{\Phi}(\mathbf{x}_i) \} \\
&= \frac{1}{N} \sum_i \alpha_i \sum_j \bar{K}_{kj} \bar{K}_{ji} \tag{8}
\end{aligned}$$

となる。従って、式(6)と(8)より、

$$\begin{aligned}
N\lambda\alpha &= \bar{\mathbf{K}}\alpha \\
\hat{\lambda}\alpha &= \bar{\mathbf{K}}\alpha \tag{9}
\end{aligned}$$

となり、最終的に $\bar{\mathbf{K}}$ の固有値問題に帰着する事になる。ここで $N\lambda$ を $\hat{\lambda}$ とし、また \bar{K}_{ki} を要素とする行列を $\bar{\mathbf{K}}$ とした。ただし、以下に示すように \bar{K}_{ij} は K_{ij} から計算することが可能である。

$$\begin{aligned}
\bar{K}_{ij} &= \bar{\Phi}(\mathbf{x}_i) \cdot \bar{\Phi}(\mathbf{x}_j) \\
&= (\Phi(\mathbf{x}_i) - \frac{1}{N} \sum_{m=1}^N \Phi(\mathbf{x}_m)) \\
&\quad \cdot (\Phi(\mathbf{x}_j) - \frac{1}{N} \sum_{n=1}^N \Phi(\mathbf{x}_n)) \\
&= \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) - \frac{1}{N} \sum_{m=1}^N \Phi(\mathbf{x}_m) \cdot \Phi(\mathbf{x}_j) \\
&\quad - \frac{1}{N} \sum_{n=1}^N \Phi(\mathbf{x}_n) \cdot \Phi(\mathbf{x}_i) \\
&\quad + \frac{1}{N^2} \sum_{m,n=1}^N \Phi(\mathbf{x}_m) \cdot \Phi(\mathbf{x}_n) \\
&= K_{ij} - \frac{1}{N} \sum_{m=1}^N 1_{im} K_{mj} - \frac{1}{N} \sum_{n=1}^N K_{in} 1_{nj} \\
&\quad + \frac{1}{N^2} \sum_{m,n=1}^N 1_{im} K_{mn} 1_{nj} \tag{10}
\end{aligned}$$

$$K_{ij} = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) \tag{11}$$

$$1_{ij} = 1 \quad \text{for all } i, j \tag{12}$$

よって、 \bar{K}_{ij} の行列表現は次式で与えられる。

$$\bar{\mathbf{K}} = \mathbf{K} - \mathbf{1}_N \mathbf{K} - \mathbf{K} \mathbf{1}_N + \mathbf{1}_N \mathbf{K} \mathbf{1}_N \tag{13}$$

$\mathbf{1}_N$ は全ての要素が $1/N$ である $N \times N$ 行列である。

ここで、固有値を $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ とし、それに対応する固有ベクトルを $\alpha^{(1)}, \dots, \alpha^{(N)}$ とした際、

$$\mathbf{v}^{(l)} \cdot \mathbf{v}^{(l)} = 1, \quad \text{for all } l = p, \dots, N \tag{14}$$

を満たすように、 α を正規化する (p 番目の固有値が、正の固有値の中で一番小さい値とする。) 式(5)と(9)より、式(14)は

$$1 = \sum_{i,j} \alpha_i^{(l)} \alpha_j^{(l)} (\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j))$$

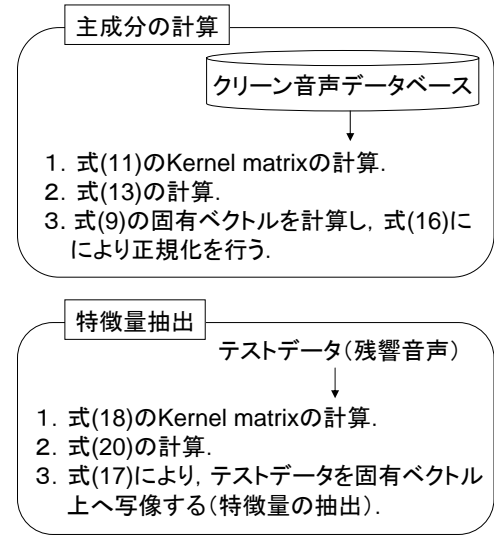


図2 Kernel PCA の計算手順

Fig. 2 Procedure of Kernel PCA

$$\begin{aligned}
&= \sum_{i,j} \alpha_i^{(l)} \alpha_j^{(l)} K_{ij} \\
&= (\alpha^{(l)} \cdot \bar{\mathbf{K}} \alpha^{(l)}) \\
&= \hat{\lambda}_l (\alpha^{(l)} \cdot \alpha^{(l)}) \tag{15}
\end{aligned}$$

となる。よって、 $\bar{\mathbf{K}}$ の固有ベクトル α に対して次式のように正規化を行う。

$$\hat{\alpha}^{(l)} = \frac{\alpha^{(l)}}{\sqrt{\lambda_l}} \tag{16}$$

次に、高次元空間において主成分を抽出するために、テストデータ \mathbf{y} の高次元における値 $\bar{\Phi}(\mathbf{y})$ を、固有ベクトル $\mathbf{v}^{(l)}$ 上に写像する。

$$\begin{aligned}
(\mathbf{v}^{(l)} \cdot \bar{\Phi}(\mathbf{y})) &= \sum_{i=1}^N \hat{\alpha}_i^{(l)} (\bar{\Phi}(\mathbf{x}_i) \cdot \bar{\Phi}(\mathbf{y})) \\
&= \sum_{i=1}^N \hat{\alpha}_i^{(l)} \bar{K}^{test}(\mathbf{x}_i, \mathbf{y}) \tag{17}
\end{aligned}$$

ここで、テストデータ \mathbf{y} と学習データとの内積 K_{ij}^{test} を計算する。

$$K_{ij}^{test} = \Phi(\mathbf{y}_i) \cdot \Phi(\mathbf{x}_j) \tag{18}$$

式(17)における \bar{K}^{test} は、 K^{test} から求めることができる。

$$\begin{aligned}
\bar{K}_{ij}^{test} &= \left(\Phi(\mathbf{y}_i) - \frac{1}{N} \sum_{m=1}^N \Phi(\mathbf{x}_m) \right) \\
&\quad \cdot \left(\Phi(\mathbf{x}_j) - \frac{1}{N} \sum_{n=1}^N \Phi(\mathbf{x}_n) \right) \tag{19}
\end{aligned}$$

$$\bar{\mathbf{K}}^{test} = \mathbf{K}^{test} - \mathbf{1}'_N \mathbf{K} - \mathbf{K} \mathbf{1}_N + \mathbf{1}'_N \mathbf{K} \mathbf{1}_N \tag{20}$$

テストデータ \mathbf{y} のフレーム数が L の場合、 $\mathbf{1}'_N$ は要素が全て $1/N$ の $L \times N$ 行列となる。図2に、Kernel PCA の計算手順の概要を示す。

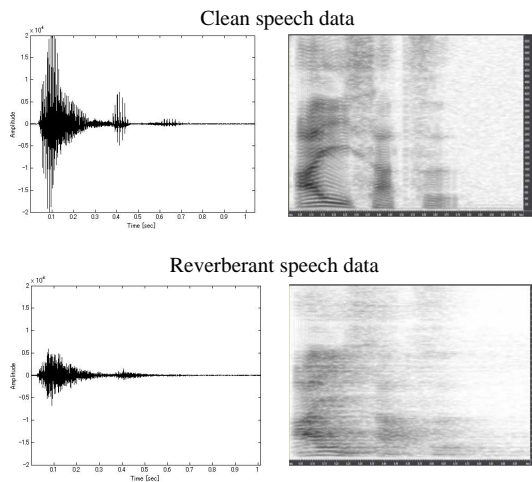


図3 クリーン音声と残響音声（残響時間 470 ms）の波形データとスペクトルグラム。 /a i sa tsu/

Fig. 3 Clean speech and reverberant speech (reverberation time = 470 ms): the speech waveform and spectrogram of the Japanese utterance /a i sa tsu/.

3. 認識実験

3.1 実験条件

評価用データとして残響音声を使用し、提案手法の有効性を検討する。残響音声の作成には、RWCP 実環境音声・音響データベース [11] より残響時間 380 ms, 470 ms, 600 ms の 3 種類のインパルス応答を使用した。マイクロホンまでの距離は約 2 m である。図 3 にクリーン音声と残響音声の波形とスペクトルグラムを示す。カーネル行列の計算に使用したクリーン音声データのフレーム数は、 $N = 2500$ とした。これは音響モデルの学習データ (2620 単語) からランダムに選択した。本実験では、カーネル行列の計算の際に次式の多項式カーネルを使用した。

$$K(x, y) = (x \cdot y + 1)^p \quad (21)$$

音声のサンプリング周波数は 12 kHz、窓幅は 32 ms、窓シフトは 8 ms とした。タスクは語彙 1000 単語として、テストデータは男性話者 3 人が対象語彙を 1 回発声したものである。音響モデルは特定話者 HMM (54 音素 HMM) を使用した。HMM は 3 状態 3 ループ、各状態が 4 混合ガウス分布とした。CMS 適用後の MFCC + Δ MFCC 32 次元での評価データに対する認識率は 63.9% (Baseline) である。提案手法では、メルフィルタバンク出力 32 次元に対し Kernel PCA を適用した。得られた値を基本係数とし、基本係数 + Δ 係数を音声認識の特徴量とした。

3.2 実験結果

図 4 に多項式カーネル次数 $p = 1$ の認識結果を示す (式 (21) において $p = 1$)。Baseline は、CMS 適用後の MFCC + Δ MFCC (32 次元) の結果である。DCT の代わりに Kernel PCA を適用することにより、主成分 16 個で 75.0% まで認識率が改善された。DCT よりも Kernel PCA の方が、ノイズに

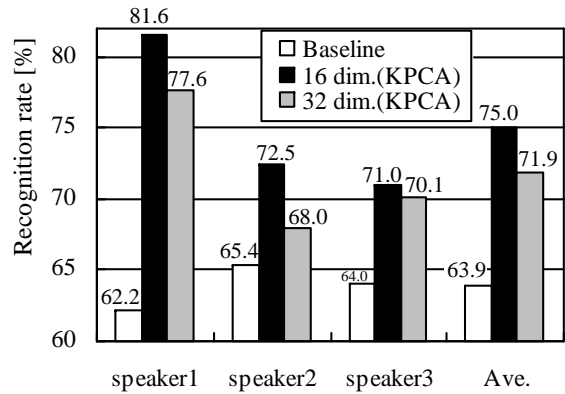


図4 提案手法による残響音声認識率 (多項式カーネル次数 $p = 1$, 残響時間 470 ms)

Fig. 4 Recognition rates for the reverberant speech by the proposed method. ($p = 1$)

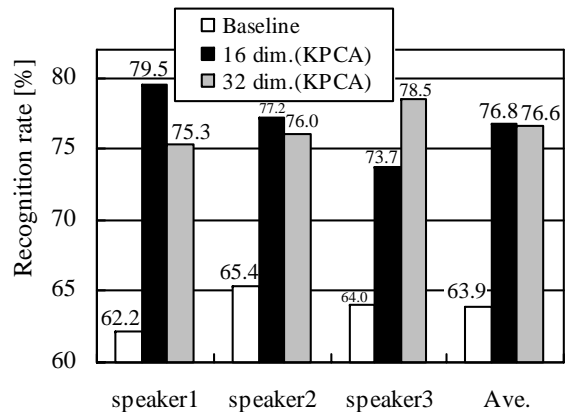


図5 提案手法による残響音声認識率 (多項式カーネル次数 $p = 2$, 残響時間 470 ms)

Fig. 5 Recognition rates for the reverberant speech by the proposed method. ($p = 2$)

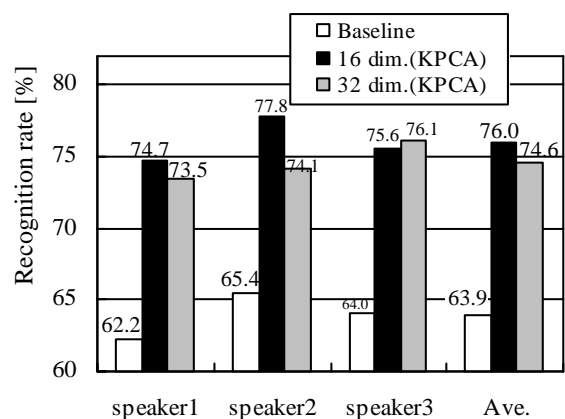


図6 提案手法による残響音声認識率 (多項式カーネル次数 $p = 3$, 残響時間 470 ms)

Fig. 6 Recognition rates for the reverberant speech by the proposed method. ($p = 3$)

ロバストな特徴量抽出法であるといえる。また、主成分 32 個で 71.9% の認識率となり、このケースでは、次元数を増やしても、認識率の改善は得られなかった。認識時のパラメータは、主成分 16 個の時は基本係数 16 次元 + Δ 係数 16 次元とな

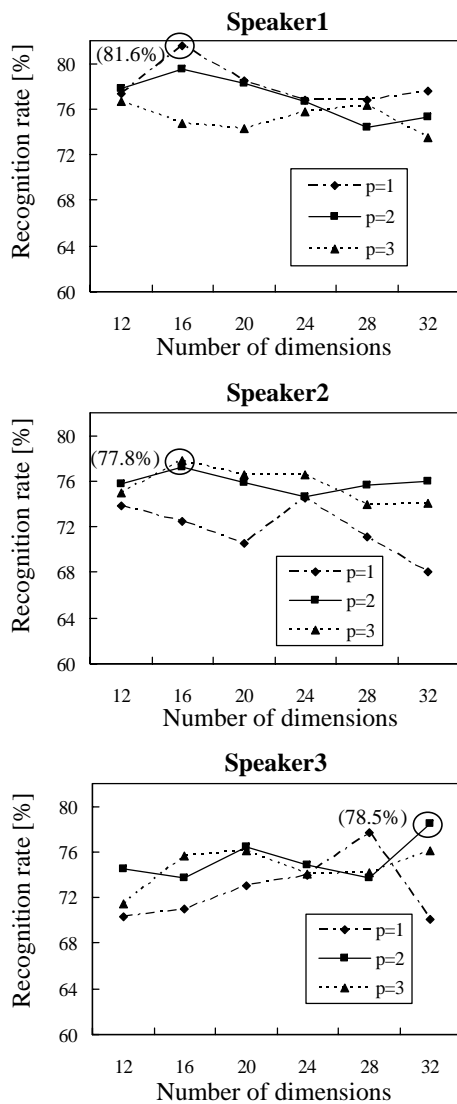


図 7 次元数（主成分数）と認識率の関係（ p は多項式カーネルの次数，残響時間 470 ms）
 Fig.7 Recognition rates of kernel PCA. In these figures, the recognition rate versus the number of dimensions is plotted.

り，全体で 32 次元の特徴量となっている．図 5 に多項式カーネル次数 $p = 2$ の認識結果を示す．平均認識率は，主成分 16 個で 76.8% となり， $p = 1$ と比べると，1.8% の認識率の改善が得られた．主成分 32 個でも 76.6% の認識率が得られており，speaker3 では，主成分 32 個の方が，16 個よりも高い認識率となった．図 6 に多項式カーネル次数 $p = 3$ の認識結果を示す． $p = 2$ の結果（図 5）と比較すると，若干平均認識率は低下している．ただし，テスト speaker1 では， $p = 1$ の場合が最も良い認識率を与え，一方，テスト speaker3 では $p = 3$ の場合が高い認識率を示している．最適なパラメータは話者毎に異なっているのが分かる．

図 7 に，次元数（主成分数）と認識率の関係を示す．テスト speaker1 は $p = 1$ ，16 次元の時，最大認識率 81.6%，テスト speaker2 は $p = 3$ ，16 次元の時，最大認識率 77.8%，テスト speaker3 は $p = 2$ ，32 次元の時，最大認識率 78.5% となってい

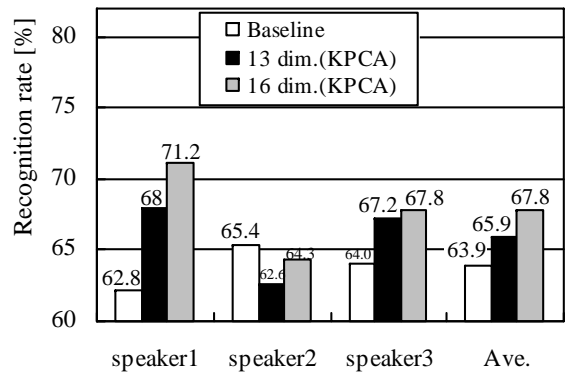


図 8 MFCC (Baseline) に Kernel PCA を適用した結果（多項式カーネル次数 $p = 1$ ，残響時間 470 ms）
 Fig.8 Recognition rates for the reverberant speech, where the kernel PCA is applied to MFCC (Baseline). ($p = 1$)

る．次元数に関しては，16 次元程度が適当であるといえるが，その際に適切なカーネル次数を設定する必要がある．

次に，Baseline の MFCC 基本係数 16 次元に対して，Kernel PCA を適用した結果を図 8 に示す [3]．(DCT 出力に対して，スペクトル包絡成分に対応する低次ケプストラムを抽出し，その後 Kernel PCA を適用する．) Baseline (63.9%) と比べて，3.9% 認識率が改善されている．一方，提案手法（図 5）と比べると 9% 近くの差がある．この理由の一つは，提案手法では 32 次元の対数スペクトル（メルフィルタバンク出力）に対して Kernel PCA を適用するため，多くのスペクトル情報から主成分が抽出可能なためであると考えられる．また，スペクトル上で Kernel PCA することで，エネルギーの強い主な音声成分は低次に集まり，ノイズは高次に集まると考えられる．この結果，Kernel PCA により一層ノイズを除去する効果が高くなっていると考えられる．

図 9 に残響時間が 380ms の場合の認識率を示す（ここではテスト speaker3 のみの結果）．Baseline の認識率は 67.5% である．提案手法を用いることにより多項式カーネル次数 3 の場合，80.6% まで認識率が改善された（主成分数 16）．また，図 10 に残響時間 600ms の場合の認識率を示す（テスト speaker3 のみの結果）．Baseline の認識率は 57.6% である．提案手法を用いることにより，70.5% まで認識率が改善された．これらの結果より，様々な残響に対して，提案手法は有効であると考えられる．

最後に，図 11 にクリーン音声に対する提案手法の認識率を示す．Baseline では 97.3%，Kernel PCA では主成分 16 個で 97.6% となり，クリーン環境下では同程度の結果が得られている．従って本手法では，ノイズ成分の有る無しに関係なく，比較的安定した認識率を得ることが出来る．

4. おわりに

残留ノイズなどにもロバストな特徴量抽出法として，MFCC における DCT の代わりに，Kernel PCA を用いた音声特徴量抽出法について検討した．残響下音声認識の結果（残響時間 470 msec），Baseline 63.9% の単語認識率に対して，提案手法

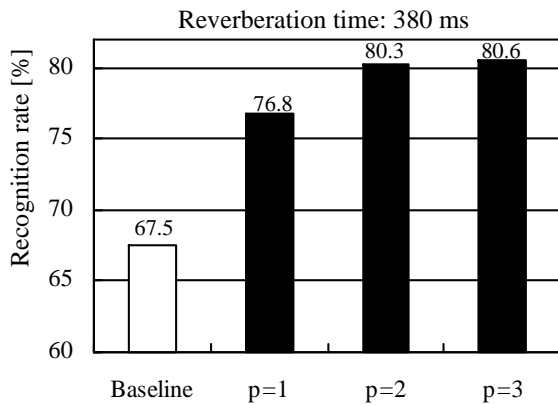


図9 提案手法による残響音声認識率 (テスト speaker3, 残響時間 380 ms, 主成分数 16 次元)

Fig. 9 Recognition rates for the reverberant speech by the proposed method (speaker3, reverberation time: 380 ms, number of dimensions: 16).

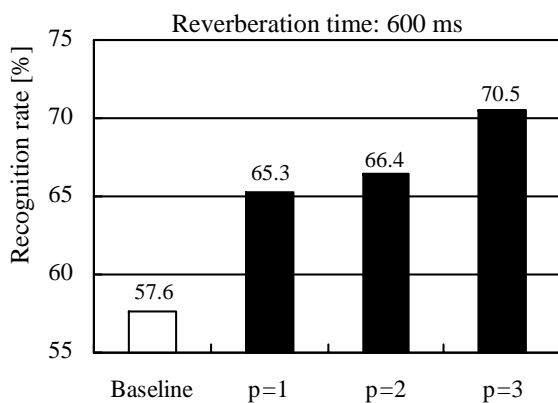


図10 提案手法による残響音声認識率 (テスト speaker3, 残響時間 600 ms, 主成分数 16 次元)

Fig. 10 Recognition rates for the reverberant speech by the proposed method (speaker3, reverberation time: 600 ms, number of dimensions: 16).

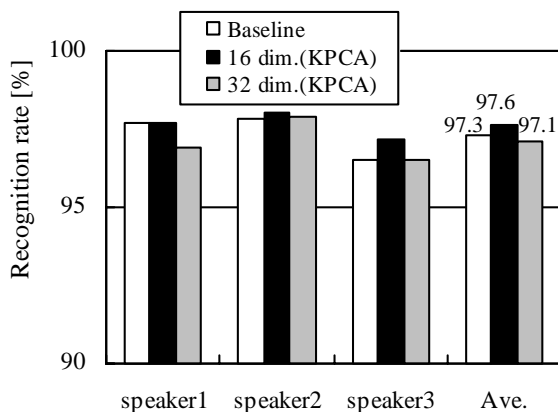


図11 クリーン音声に対する認識結果 (多項式カーネル次数 $p = 2$)

Fig. 11 Recognition rates for the clean speech by the proposed method. ($p = 2$)

により (最大) 79.3%まで認識率が改善された。今後は、様々なノイズ除去手法との統合, より適切なカーネル関数の設定方法などの検討を行い, 考察を続けていく。

文 献

[1] S. Mika, B. Scholkopf, A.J. Smola, K.-R. Muller, M. Scholz,

and G. Ratsch, "Kernel PCA and de-noising in feature spaces," In M.S. Kearns, S.A. Solla, and D.A. Cohn, editors, *Advances in Neural Information Processing Systems 11*, pp. 536–542, MIT Press (1999).

[2] K. I. Kim, M. O. Franz, and B. Schölkopf, "Kernel Hebbian Algorithm for Single-Frame Super-Resolution," *Statistical Learning in Computer Vision (SLCV 2004)*, pp. 135–149 (2004).

[3] A. Lima, H. Zen, Y. Nankaku, C. Miyajima, K. Tokuda, and T. Kitamura, "On the Use of Kernel PCA for Feature Extraction in Speech Recognition," *IEICE Trans. Inf. & Syst.*, Vol. E87-D, No. 12, pp. 2802–2811 (2004).

[4] A. Lima, H. Zen, Y. Nankaku, K. Tokuda, T. Kitamura and F. G. Resende, "Applying Sparse KPCA for Feature Extraction in Speech Recognition," *IEICE Trans. Inf. & Syst.*, Vol. E88-D, No. 3, pp. 401–409 (2005).

[5] M. Miyoshi and Y. Kaneda, "Inverse Filtering of room acoustics," *IEEE Trans. on ASSP*, Vol. 36, pp. 145–152 (1988).

[6] H. Wang and F. Itakura, "An Approach of Dereverberation using Multi-Microphone Sub-Band Envelope Estimation," *ICASSP*, pp. 953–956 (1991).

[7] C. Avendano, S. Tivrewala, and H. Hermansky, "Multiresolution channel normalization for ASR in reverberant environments," *Eurospeech*, pp. 1107–1110 (1997).

[8] 中谷智広, 三好正人, 木下慶介, "調波構造に基づくモノラル音声信号のブラインド残響除去," *電子情報通信学会論文誌, D-II*, Vol. J88-D-II, No. 3, pp. 509–520 (2005).

[9] 竹居翼, 松本弘, 山本一公, "短時間スペクトル系列残響モデルの付加雑音下での推定と音声認識による評価," *日本音響学会秋季講演論文集*, 3-7-23, pp. 147–148 (2005).

[10] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, Vol. 10, pp. 1299–1319 (1998).

[11] S. Nakamura, K. Hiyane, F. Asano, T. Nishiura, T. Yamada, "Acoustical Sound Database in Real Environments for Sound Scene Understanding and Hands-Free Speech Recognition," *Proceedings of International Conference on Language Resources and Evaluation*, Vol. 2, pp. 965–968 (2000).