Greedy Kernel PCA를 이용한 화자식별*

김민석(서울시립대), 양일호(서울시립대), 유하진(서울시립대)

<차 례>

<Abstract>

Speaker Identification Using Greedy Kernel PCA

Min-Seok Kim, IL-Ho Yang, Ha-Jin Yu

In this research, we propose a speaker identification system using a kernel method which is expected to model the non-linearity of speech features well. We have been using principal component analysis (PCA) successfully, and extended to kernel PCA, which is used for many pattern recognition tasks such as face recognition. However, we cannot use kernel PCA for speaker identification directly because the storage required for the kernel matrix grows quadratically, and the computational cost grows linearly (computing eigenvector of $l \times l$ matrix) with the number of training vectors l. Therefore, we use greedy kernel PCA which can approximate kernel PCA with small representation error. In the experiments, we compare the accuracy of the greedy kernel PCA with the baseline Gaussian mixture models using MFCCs and PCA. As the results with limited enrollment data show, the greedy kernel PCA outperforms conventional methods.

* Keywords: Speaker identification, Kernel method, Kernel PCA, Greedy kernel PCA.

* 이 논문은 2007년 학술진흥재단의 지원에 의하여 연구되었음 (KRF-2007-321-A00155).

1. 서 론

화자식별은 유비쿼터스 환경에서 음성을 통하여 간편한 보안 시스템을 구현하기 위한 방법으로 그 필요성이 증대되고 있으나, 아직은 성능이 사용자의 기대에 미치지 못하여 성능향상을 위한 연구가 계속되고 있다. 최근에는 선형으로 분리 가능하지 않은 음성의 특징을 변환하여 비선형으로 모델링함으로써 성능을 향상시키는 방향으로 많은 연구가 진행되고 있다. 특히, 입력 공간(input space) 상의 특징 벡터를 고차원 특징 공간(feature space)으로 확장하여 비선형 문제를 해결하는 커널방식(kernel method)[1]이 기계학습 분야와 패턴인식 분야에서 빠르게 발전하고 있다. 커널방식은 유명한 분류 알고리즘인 support vector machine (SVM)[1]과 선형분석법인 선형판별분석 (linear discriminant analysis: LDA)[6]의 비선형 버전인 kernel fisher discriminant analysis (FDA)[1], 주성분분석 (principal component analysis: PCA)[6]의 비선형 버전인 커널(kernel) PCA[1] 등에 적용되고 있다.

본 논문에서는 커널 PCA를 통해 추출된 특징벡터가 여러 패턴인식(pattern recognition) 분야에서 좋은 성능을 나타낸 이전 연구[9]를 기반으로 이 특징벡터를 화자 식별에 적용하였다. 그러나 커널 PCA가 비선형 분류에 좋은 특징을 추출할 수 있는 장점에도 불구하고, 특징 벡터의 개수 l이 커짐에 따라 제곱배 만큼의 저장 공간이 필요한 문제를 가지고 있어, 데이터 수가 많은 문제에는 적용이 불가능하다. 또한 커널 PCA에서는 $l \times l$ 행렬의 고유벡터를 구해야하기 때문에 실행이 불가능할 수도 있다. 대부분의 화자 식별 시스템에서 학습에 쓰이는 특징 벡터 수가 10,000개 이상 (20 ms shift인 경우 200초 이상, 10 ms shift인 경우 100초 이상)이므로, 일반적인 커널 PCA를 통한 특징 추출 방법을 적용할 수 없다.

Kernel method를 이용하여 대용량 데이터를 처리하기 위한 방법으로는, kernel matrix를 구하는데 이용되는 특징 벡터를 줄이는 방법[3][4]과 특징 벡터를 조금씩 이용하여 점진적으로 문제를 해결하는 방법[5]이 있다. 본 논문에서는 특징 벡터 집합을 대표하는 부분 집합 즉, 분석에 이용되는 특징 벡터들을 대표할 수 있는 몇 개를 선택하여 kernel PCA를 수행하는 greedy kernel PCA[3] 알고리즘을 이용한 화자 식별 시스템을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 kernel PCA와 본 논문에서 제안하는 화자 식별 시스템을 소개하고, 3장에서는 greedy kernel PCA에 대해 설명한다. 4장에서는 실험에 이용하는 데이터베이스와 실험 환경에 대해 설명하고, 실험 결과와 분석 결과를 제시한다. 마지막으로 5장에서 결론을 맺는다.

2. Kernel PCA를 이용한 특징 추출 및 화자 식별

선형 PCA[1][2]는 특징 벡터 $\vec{x}_k (k=1,...,l,\ \vec{x}_k \in R^d,\ \sum_{k=1}^{l} \vec{x}_k = \vec{0})$의 공분산(covariance) 행렬의 대각 성분을 최대화하는 새로운 기저로 특징을 변환하는 방법이다. 공분산 행렬 $\vec{C} (\in R^{d \times d},\ d \times d$ 행렬)는 다음과 같이 정의된다.

$$\vec{C} = \frac{1}{l} \sum_{j=1}^{l} \vec{x}_j \vec{x}_j^T \tag{1}$$

여기서 $\vec{x}_j^T$는 $\vec{x}_j$의 전치(transpose)를 의미한다.

새로운 기저는 $\vec{C}$의 고유벡터 집합으로 구성되고, 이를 주성분(principal component)이라 부른다. 본 장에서는 특징 벡터를 고차원으로 확장하여 분석하는 커널 PCA에 대해 설명한다. 특징 벡터를 고차원 공간으로 사상하는 함수를 ϕ라 정의하면 식 (2)와 같다.

$$\phi : R^d \rightarrow F,\ \ \vec{x} \rightarrow \phi(\vec{x}) \tag{2}$$

여기서 $\phi(\vec{x})$는 $\vec{x}$가 고차원으로 확장된 특징 벡터이며, F는 임의의 고차원 공간으로 특징 공간(feature space)이라 부르고, 특징 벡터 $\vec{x}$가 속한 공간(R^d, d차원)을 입력 공간(input space)이라 부른다. 직접 ϕ 함수로 사상된 특징 공간상의 특징벡터는 지나치게 고차원이므로 처리가 어려워 2.1절에서 소개하는 바와 같이 ϕ 함수 결과의 내적을 이용한다.

2.1. Kernel Principal Component Analysis (Kernel PCA)

커널 PCA[1][2]는 input space 상의 d차원 특징 벡터를 ϕ를 통해 고차원으로 확장하여 선형 PCA를 적용하는 방법이다. Feature space 상의 특징 벡터를 $\phi(\vec{x}_1), \phi(\vec{x}_2), ..., \phi(\vec{x}_l),\ \phi(\vec{x}_k) \in F$라 하고, 이들의 평균이 $\vec{0}$[1]이라 할 때(즉, $\sum_{k=1}^{l} \phi(\vec{x}_k) = \vec{0}$), 공분산 행렬은

$$\overline{\vec{C}} = \frac{1}{l} \sum_{j=1}^{l} \phi(\vec{x}_j) \phi(\vec{x}_j)^T \tag{3}$$

1) 본 논문에서는 수식을 간단히 설명하기 위해 평균이 0이라 가정하였으나 일반적인 경우 그렇지 않다. 평균이 0이 아닌 경우에 대한 kernel PCA는 [1]에 설명되어 있다.

와 같이 정의되고, 이 공분산 행렬에서 양의 고유값 $\lambda(>0)$를 따르는 고유벡터, $\vec{v}$ $(\in F)$를 구함으로써 커널 PCA를 수행한다.

$$\lambda\vec{v}=\overline{\overline{C}}\vec{v} \tag{4}$$

여기서 $\phi(\vec{x})$와 $\vec{v}$가 고차원의 특징공간 상에 존재하기 때문에 직접 구하지 못하는 문제를 해결하기위해 커널방식(kernel method)을 이용하고, 이 때문에 커널 PCA라 불린다.

$\vec{v}$를 $\phi(\vec{x}_1),\phi(\vec{x}_2),\ldots,\phi(\vec{x}_l)$의 선형 결합이라 가정[2]하고, 계수를 $\vec{\alpha}=\{\alpha_1,\alpha_2,\ldots,\alpha_l\}$이라 하면 다음과 같이 표현할 수 있다.

$$\vec{v}=\sum_{i=1}^{l}\alpha_i\phi(\vec{x}_i). \tag{5}$$

또한 식 (4)의 양변에 $\phi(\vec{x}_j)$를 곱하면,

$$\lambda(\phi(\vec{x}_j)\bullet\vec{v})=(\phi(\vec{x}_j)\bullet\overline{\overline{C}}\vec{v}) \tag{6}$$

이고, 이 식에 식 (3)과 식 (5)를 대입하고, kernel matrix인 $l\times l$ 크기의 행렬 $\vec{K}$의 i행 j열의 각 원소, $\vec{K}_{ij}$를 $\phi(\vec{x}_i)$와 $\phi(\vec{x}_j)$의 내적으로 정의하면,

$$\vec{K}_{ij}\equiv(\phi(\vec{x}_i)\bullet\phi(\vec{x}_j)) \tag{7}$$

이고, 이로부터 다음과 같은 식을 유도할 수 있다.

$$l\lambda\overrightarrow{K\alpha}=\vec{K}^2\vec{\alpha} \tag{8}$$

여기서 양변을 $\vec{K}$로 나누면, $l\lambda\vec{\alpha}=\overrightarrow{K\alpha}$가 되고, $\vec{\alpha}$는 고유값 문제로 풀 수 있다. 식 (8)의 0이 아닌 고유값이 큰 순으로 정렬된 고유벡터 k개로 구성된 $\vec{A}^k=\{\vec{\alpha}_1,\vec{\alpha}_2,\ldots,\vec{\alpha}_k\}$가 있을 때, 식 (4)에서 같은 방법으로 선택된 k개의 고유벡터로 구성된 $\vec{V}^k=\{\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k\}$은 고유 벡터의 크기가 1이라는 정의를 만족시키기 위해 다음과 같이 정규화(normalization)를 한다.

$$\begin{aligned}&(\vec{v}\bullet\vec{v})=1,\\&1=\sum_{i,j=1}^{l}\alpha_i\alpha_j(\phi(\vec{x}_i)\bullet\phi(\vec{x}_j))=(\vec{\alpha}\bullet\overrightarrow{K\alpha})=\lambda(\vec{\alpha}\bullet\vec{\alpha}).\end{aligned} \tag{9}$$

새로운 특징 벡터 $\vec{y}$를 앞에서 구한 커널 주성분 $\vec{v}$로 사상하는 방법은 다음과 같다.

$$(\vec{v} \cdot \phi(\vec{y})) = \sum_{i=1}^{l} \alpha_i (\phi(\vec{x}_i) \cdot \phi(\vec{y})). \tag{10}$$

식 (7)과 식 (10)은 $\phi(\vec{x})$의 내적 꼴로 나타나 있고, 이는 커널 함수(kernel function) k를 통해 쉽게 유도할 수 있다.

$$k(x,y) = \phi(x) \cdot \phi(y) \tag{11}$$

대표적인 커널 함수로는 다항식 함수인 $k(x,y) = (x \cdot y)^a$, radial basis function인 $k(x,y) = \exp(-\|x-y\|^2/(2\sigma^2))$ 등이 있다[1].

2.2. 화자 식별 시스템

본 연구에서는 Gaussian mixture model (GMM)[8] 기반의 화자 식별 시스템을 이용한다. GMM으로 모델링된 화자모델은 특징 벡터의 평균과 공분산으로 표현되는 가우시안 확률분포함수, $g_i(\vec{x})$,의 M(mixture 개수)개의 선형 결합으로 나타낸다.

$$p(\vec{x}|\lambda) = \sum_{i=1}^{M} w_i g_i(\vec{x}), \qquad \sum_{i=1}^{M} w_i = 1 \tag{12}$$

$$g_i(\vec{x}) = \frac{1}{(2\pi)^{D/2}|\vec{\Sigma}_i|^{1/2}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu}_i)\vec{\Sigma}_i^{-1}(\vec{x}-\vec{\mu}_i)^T\right\} \tag{13}$$

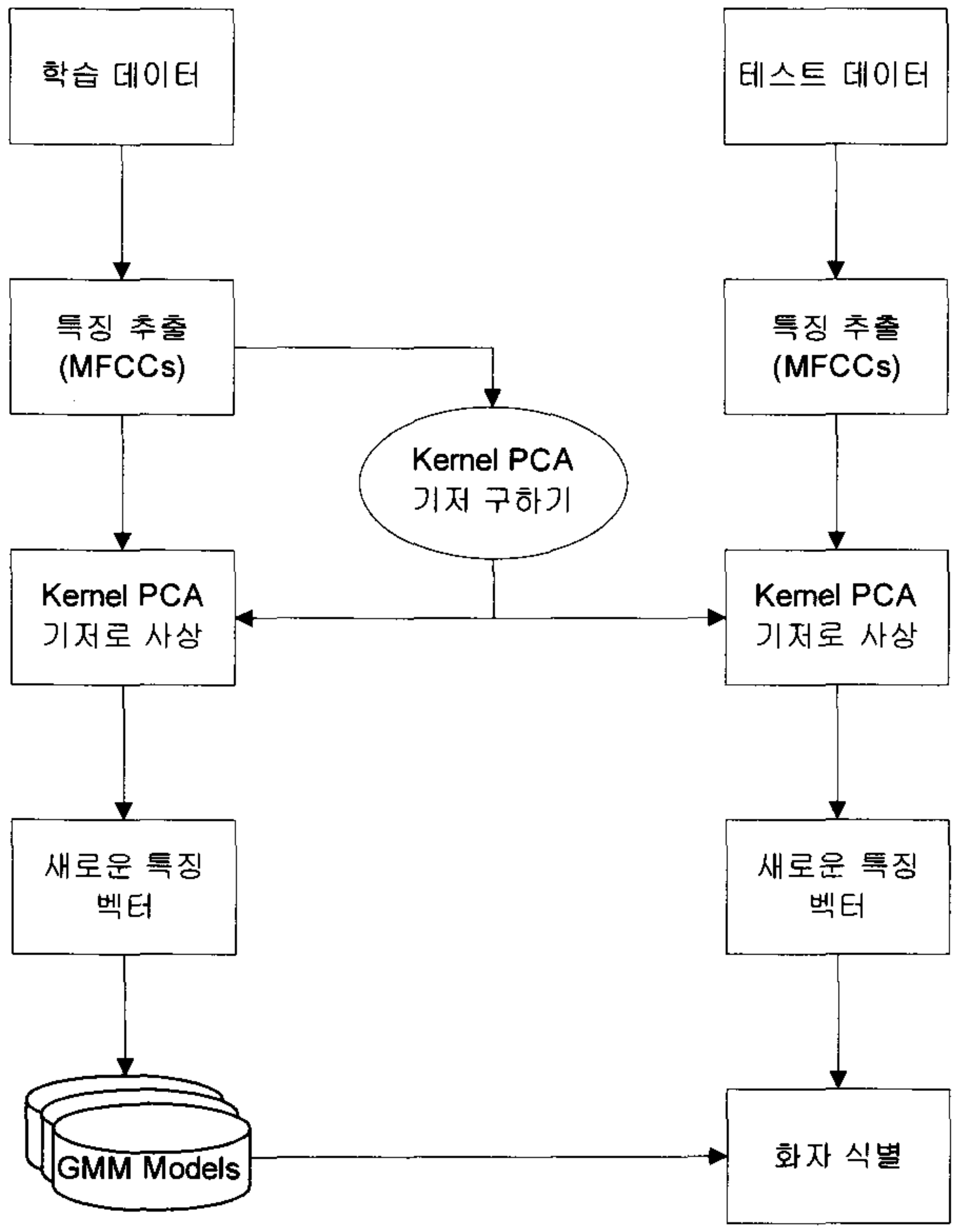
여기서 $\vec{\mu}_i$는 i번째 pdf의 평균이고, $\vec{\Sigma}_i$는 i번째 pdf의 공분산이다. λ는 화자 모델 파라미터를 나타낸다. 본 논문에서는 대각 공분산 행렬(diagonal covariance matrix)을 이용한다.

$$\lambda = \left(w_i, \vec{\mu}_i, \vec{\Sigma}_i\right), \quad \text{for } i = 1, 2, \ldots, M \tag{14}$$

모델의 학습에는 expectation-maximization (EM) 알고리즘을 이용한다. 화자식별은 T개의 음성 특징 벡터를 S명의 식별 대상 각각에 대하여 유사도를 구하고, 유사도가 가장 큰 $\hat{S}$를 선택한다.

$$\hat{S} = \arg\max_{1 \le k \le S} \sum_{t=1}^{T} \log p(\vec{x}_t|\lambda_k) \tag{15}$$

커널 PCA로 추출된 특징을 이용한 화자 식별 시스템은 <그림 1>과 같다. 본 논문에서는 학습 데이터에서 mel-frequency cepstral coefficient (MFCC)로 특징 벡터를 추출하고, 커널 PCA 기저를 구한 후, 이 기저로 학습 데이터와 테스트 데이터를 사상하여 화자 식별을 수행한다.



<그림 1> Kernel PCA로 추출된 특징을 이용한 화자 식별 시스템 구성도

3. Greedy Kernel PCA

2.1절에서 설명한 커널 PCA는 다양한 패턴인식 분야에서 특징 추출법으로 이용되고 있고, 분류문제에서 좋은 성능을 나타낸다[9]. 그러나 분석하려는 특징 벡터의 수 l이 증가함에 따라 kernel matrix $\vec{K}$를 저장하기 위해서는 l^2크기의 공간이 필요하고, $l \times l$ 행렬의 고유벡터를 구하는 알고리즘은 실행 불가능 할 수 있다. 이 문제를 해결하기 위한 여러 연구가 진행되고 있다[3][4][5]. 본 논문에서는 greedy kernel PCA[3]를 이용한 화자 식별 시스템을 제안한다.

Greedy kernel PCA[3]는 분석하려는 l개의 특징 벡터 집합, $\vec{X}=\left\{\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_l\right\}$를 feature space로 확장한 $\phi(\vec{X})=\left\{\phi(\vec{x}_1),\phi(\vec{x}_2),\ldots,\phi(\vec{x}_l)\right\}$를 대표하는 m개의 특징 벡터, $\phi(\vec{S})=\left\{\phi(\vec{s}_1),\phi(\vec{s}_2),\ldots,\phi(\vec{s}_m)\right\}(\phi(\vec{S})\subseteq\phi(\vec{X}))$를 선택하여 앞에서 설명한 문제를 해결하는 알고리즘이다. $\phi(\vec{S})$를 통해 추정된 특징의 집합을 $\phi(\widetilde{X})=\left\{\phi(\tilde{x}_1),\phi(\tilde{x}_2),\ldots,\phi(\tilde{x}_l)\right\}$라 하면, $\phi(\widetilde{X})$는 $\phi(\vec{S})$의 선형 결합으로 나타낼 수 있다.

$$\phi(\tilde{x}_i)=\phi(\vec{S})\vec{\beta}_i,\quad \text{for}\, i=1,\ldots,l \tag{16}$$

여기서 $\vec{\beta}$는 $m\times 1$ 계수이며, m은 kernel matrix를 계산할 수 없을 정도로 크지 않아야한다. 즉, $m\ll l$. $\phi(\vec{x})\in\phi(\vec{X})$에 대한 mean square error (MSE)는 다음과 같고, greedy kernel PCA는 MSE를 최소화하는 $\phi(\vec{S})$를 선택한다.

$$\epsilon_{MS}=\frac{1}{l}\sum_{i=1}^{l}\|\phi(\vec{x}_i)-\phi(\tilde{x}_i)\|^2 \tag{17}$$

즉 $\phi(\vec{S})$로 추정된 특징 벡터 $\phi(\widetilde{X})$가 원래의 특징 벡터라 할 때, $\phi(\vec{X})$를 가장 잘 대표할 수 있는 $\phi(\vec{S})$를 추정하는 것이 greedy kernel PCA의 목적이다. 식 (17)을 kernel function으로 치환하면 다음과 같다(유도 과정은 [3] 참고).

$$\epsilon_{MS}=\frac{1}{l}\sum_{i=1}^{l}\left(k(\vec{x}_i,\vec{x}_i)-2\vec{K}_{tr}k_{tr}(\vec{x}_i)+\left\langle k_{tr}(\vec{x}_i),\vec{K}_{tr}k_{tr}(\vec{x}_i)\right\rangle\right) \tag{18}$$

여기서 $\vec{K}_{tr}$은 S의 kernel matrix이고, $k_{tr}(\vec{x}_i)=\left[k(\vec{s}_1,\vec{x}_i),k(\vec{s}_2,\vec{x}_i),\ldots,k(\vec{s}_m,\vec{x}_i)\right]^T$이고 $<\vec{a},\vec{b}>$는 $\vec{a}$와 $\vec{b}$의 내적을 의미한다. 식 (18)을 이용하여 모든 가능한 m개의 부분집합을 선택하는 경우의 수는 ${}_lC_m$개이다. 효과적으로 m개의 부분집합, $\phi(\vec{S})$를 선택하기 위해서 다음 식을 이용한다.

$$\epsilon_{MS}=\frac{1}{l}\sum_{i=1}^{l}\|\phi(\vec{x}_i)-\phi(\tilde{x}_i^{(t)})\|^2\leq\frac{1}{l}(l-t)\max_{\vec{x}_i\in\vec{X},\,\vec{x}_i\notin\phi(\vec{S}^{(t)})}\|\phi(\vec{x}_i)-\phi(\tilde{x}_i^{(t)})\|^2 \tag{19}$$

이 식은 식 (18)에서 오류($\|\phi(x_i)-\phi(\tilde{x}_i^{(t)})\|^2$)의 합 대신 오류의 상한값으로 치환한 것이다. 여기서 t는 1부터 m까지 점진적으로 구하고, $\phi(\vec{S}^{(t)})$는 t번 째 반복에서 선택된 특징들로 구성된 집합이고, $\phi(\tilde{x}_i^{(t)})$은 $\phi(\vec{S}^{(t)})$로 추정된 특징이다. 자세한 알고리즘은 [3]에 설명되어 있다.

4. 실험 및 결과

본 논문에서는 YOHO[7] 데이터베이스의 음성을 이용하여 제안한 방법의 성능을 평가하였다. YOHO 음성 데이터베이스는 화자인식 실험을 위해 만들어진 데이터베이스로 총 138명의 화자로 구성되어 있다. 각각의 화자는 등록용으로 4번, 인식용으로 10번의 세션을 갖고 녹음하였으며, 녹음은 고급 전화용 마이크를 사용하여 조용한 사무실 환경에서 하였다. 발성 내용은 (35-72-41)과 같이 2자리 숫자를 3개씩 연이어 발음하는 형식으로, 등록 세션에는 24개씩(한 화자 당 등록 음성 $4 \times 24 = 96$개), 그리고 인식 세션에는 4개씩(한 화자 당 인식 음성 $10 \times 4 = 40$개)의 음성이 있다.

본 논문에서는 138명의 화자 중 20명, 50명의 화자를 선택하여 실험하였다(2개의 실험 환경). 학습에는 한 개 등록 세션의 5개 음성을 이용하고, 인식에는 모든 인식 세션을 이용하였다. 특징으로는 20 ms Hamming window를 사용하여 분석한 MFCC 15차와 에너지(16차원)를 이용하였고, 식별단계에서 계산량 감소를 위해 delta와 delta- delta는 이용하지 않았다. 또한 채널 왜곡을 감소시키고 특징 벡터 공간의 불일치를 해결하기 위하여 cepstral mean subtraction (CMS) 방법을 사용하였다. 또한 에너지를 기반으로 하여 음성 앞뒤의 무음구간을 제거하였다.

총 학습에 이용되는 특징 벡터의 개수(3.1절에서 l을 의미)는 화자 수가 20명일 때, 10,394개, 50명일 때, 26,317개였다. 또한 20명 화자에 대한 식별 실험에서는 1,000개의 벡터를, 50명 화자에 대한 실험에서는 300개(3.1절에서 m을 의미)의 벡터를 선택 하여 greedy kernel PCA를 수행하였다. 커널함수(kernel function)는 radial basis function인 $k(\vec{x},\vec{y}) = \exp(-\|\vec{x}-\vec{y}\|^2/(2\sigma^2))$을 이용하였고, $\sigma = 1$로 설정하였다. 또한 커널함수를 효과적으로 적용하기 위해서 모든 특징들이 -1과 1사이에 속하도록 다음과 같이 정규화 하였다(실험 결과에서 정규화된 특징이라 부른다). 학습데이터의 각 차원의 최대값과 최소값을 각각 max, min이라 하면, max, min의 평균인 mean은 (max+min)/2이고, 구간의 크기, scale은 max-min이다. 따라서 특징 x에 대한 정규화된 특징은 (x-mean)/scale×2로 구할 수 있다.

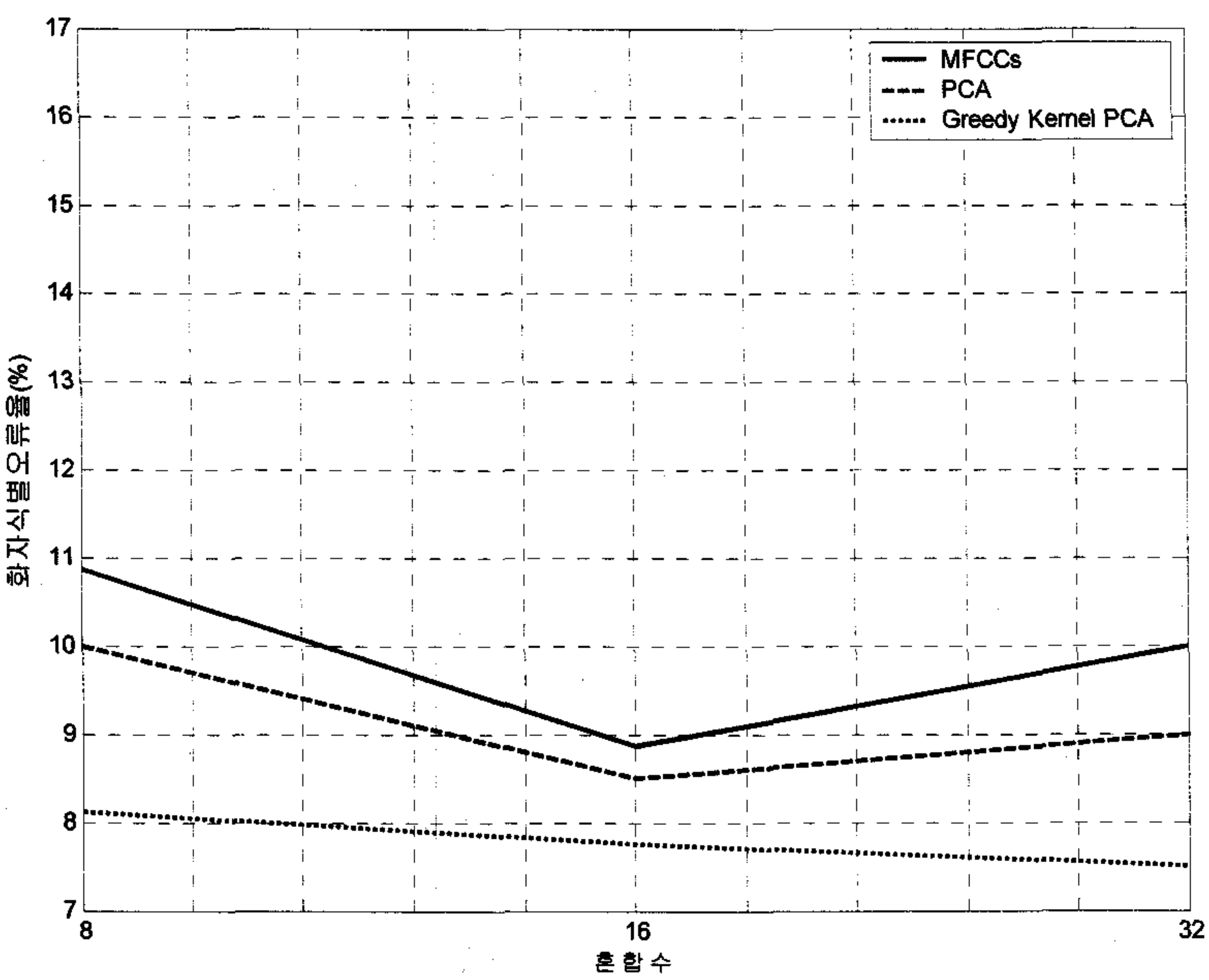
제안한 화자 식별 시스템은 MFCCs 특징을 이용한 결과와 PCA를 통해 특징을 추출한 결과와 비교하였다. 실험결과는 <표 1>, <표 2>, <그림 2>, <그림 3>과 같다. -1과 1사이로 정규화된 특징 벡터를 이용한 화자 식별 실험은 MFCCs와 완벽하게 일치한다.

모든 혼합수에서 greedy kernel PCA가 다른 방법들보다 높은 성능을 나타내었다. 특히 혼합수가 32개일 때, 20명 화자에 대해서는 MFCCs, PCA 특징을 이용한 화자식별 실험보다 각각 2.5%, 1.5%의 성능 향상이 있었고, 50명 화자에 대해서는 각각 1.55%, 1.15%의 성능 향상이 있었다. 제안한 방법에서는 혼합수가 32개 일 때가 16일 때 보다 오류가 감소하지만 MFCCs를 사용한 기존의 방법에서는 오히

려 오류가 증가하는 것을 관찰할 수 있다. 여기서 MFCCs와 이의 선형 변환결과인 PCA는 혼합수 32개에서 과훈련(overtraining) 때문에 오류가 증가하지만, 커널 PCA는 혼합수 32개에서도 학습이 잘되는 것을 알 수 있다. [7]에 비해 절대 식별 오류율이 낮은 이유는 [7]에서는 한 화자당 96개의 학습데이터를 이용한 반면 본 연구에서는 10개만을 이용하였기 때문이다. 그러나 두 실험에서 기존의 방법에 비해 절대 오류율이 약 1~2% 감소한 것을 알 수 있다.

<표 1> 20명 화자에 대한 식별 오류율

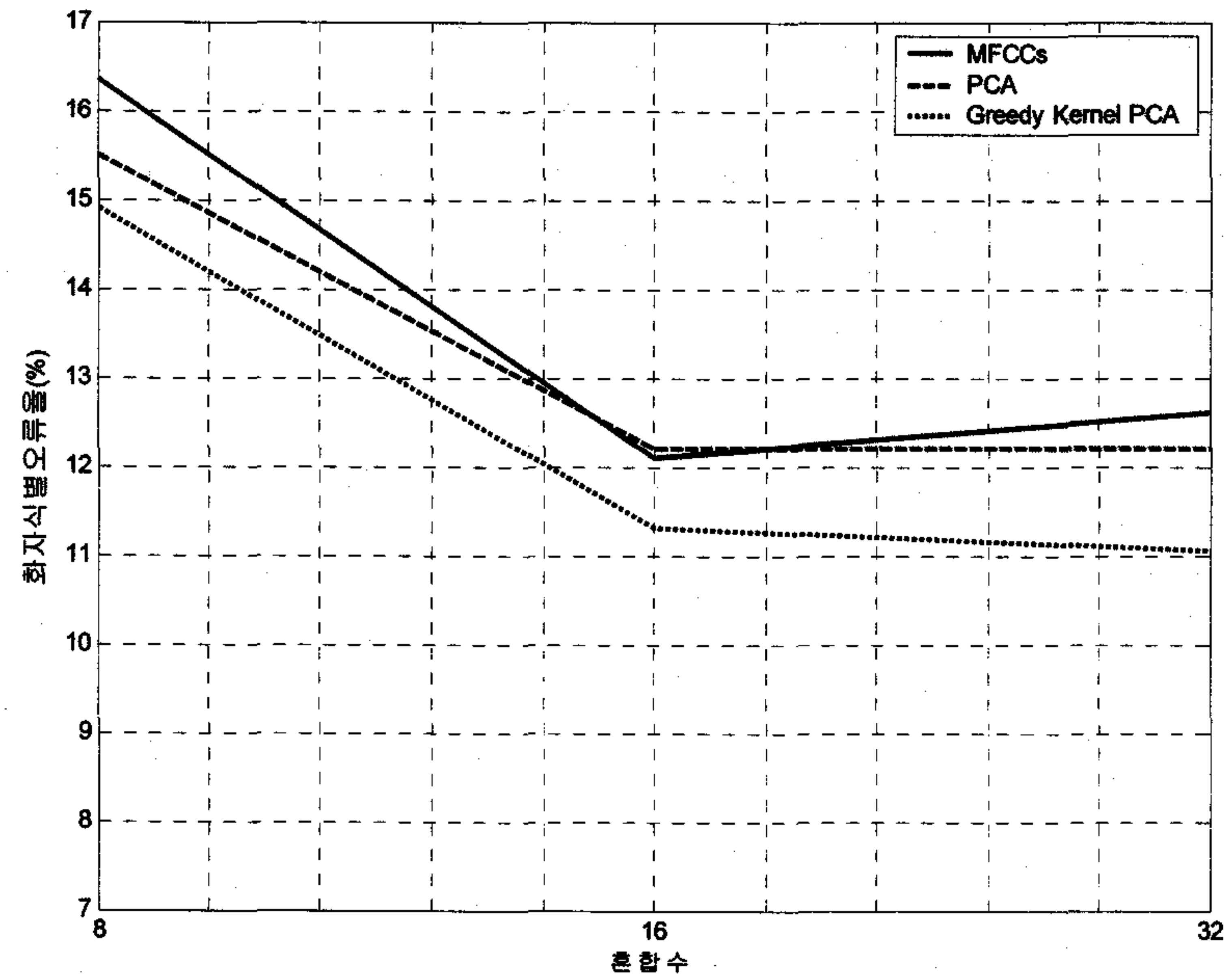
혼합수	8	16	32
MFCCs	10.87	8.87	10.00
정규화된 특징	10.87	8.87	10.00
PCA	10.00	8.50	9.00
Greedy Kernel PCA	8.12	7.75	7.50



<그림 2> 20명 화자에 대한 식별 오류율

<표 2> 50명 화자에 대한 식별 오류율

혼합수	8	16	32
MFCCs	16.35	12.10	12.60
정규화된 특징	16.35	12.10	12.60
PCA	15.50	12.20	12.20
Greedy Kernel PCA	14.90	11.30	11.05



<그림 3> 50명 화자에 대한 식별 오류율

5. 결 론

본 논문에서는 greedy kernel principal component analysis (PCA)를 이용한 화자 식별 시스템을 제안하였다. 기존의 커널(kernel) PCA는 데이터 개수 l에 따라 저장 용량이 l^2만큼 필요하고, $l \times l$ 행렬의 고유벡터를 구해야하기 때문에 실제 화자 식별에 적용하기 어렵다. 본 논문에서는 greedy kernel PCA를 통해 이 문제를 해결하였고, 화자 식별 결과에서 MFCCs 특징을 이용한 시스템과 선형 PCA 특징 추출을 이용한 시스템에 비해 각각 2.5%와 1.5% (상대오류감소 25% 및 17%) 성능 향상 결과를 얻을 수 있었다.

커널 PCA의 특성상 많은 학습데이터를 처리하기 어려우므로, 본 연구 결과는 사용자로부터 충분한 학습데이터를 얻을 수 없는 현실적인 상황에서 효과적으로 사용될 수 있을 것이다. 또한, 화자 식별에 greedy kernel PCA를 적용할 때, 데이터의 음소가 많으면 많을수록 보다 많은 특징을 선택해야 한다. 즉 m을 보다 크게 설정하여 대표하는 값을 보다 많이 선택해야 할 것이다. 본 논문에서는 숫자음 데이터를 이용하였는데, 이 보다 많은 음소가 포함된 경우(문장 발성 등)에는 숫자음의 경우보다 많은 수의 특징을 선택할 필요가 있다.

향후 계획으로는 본 논문에서 이용한 데이터의 개수를 줄이는 방법으로 커널 PCA의 단점을 해결하는 방법뿐만 아니라, 점진적으로 커널 PCA를 구하는 incremental kernel PCA[5]를 적용하는 화자 식별 시스템에 대해 연구하는 것이다.

참 고 문 헌

[1] J. Shawe-Taylor, N. Cristianini, *Kernel Methods for Pattern Analysis*, New York: Cambridge University Press, 2004.

[2] B. Scholkopf, A. Smola, K.-R. Muller, "Kernel principal component analysis", *Proc. Int. Conf. on Artificial Neural Networks*, pp. 583-588, 1997.

[3] V. Franc. *Optimization Algorithms for Kernel Methods*, Ph.D. Thesis, Centre for Machine Perception, Czech Technical University, July 2005.

[4] B. Scholkopf, P. Knirsch, A. J. Smola, C. Burges, "Fast approximation of support vector kernel expansions, and an interpretation of clustering as approximation in feature spaces", *Proc. DAGM Symp. Mustererkennung*, pp. 124 - 132, 1998.

[5] T.-J. Chin, D. Suter, "Incremental kernel principal component analysis", *IEEE Transactions on Image Processing*, Vol. 16, No. 6, pp. 1662-1674, 2007.

[6] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*, 2nd Ed., New York: Wiley-Interscience Publication, 2000.

[7] J. P. Campbell, Jr., "Testing with the YOHO CD-ROM voice verification corpus", *Proc. ICASSP*, pp. 341-344, 1995.

[8] D. A. Reynolds, R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models", *IEEE Transactions on Speech and Audio Processing,* Vol. 3, No. 1, pp. 72-83, 1995.

[9] B. Scholkopf, A. Smola, K.-R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem", *Neural Computation*, Vol. 10, No. 5, pp 1299-1319, 1998.

접수일자: 2008년 5월 11일
게재결정: 2008년 6월 12일

▶ 김민석(Min-Seok Kim)
주소: 130-743 서울 동대문구 전농동 90
소속: 서울시립대학교 컴퓨터과학부
전화: 02) 2210-5322
E-mail: ms@uos.ac.kr

▶ 양일호(Il-Ho Yang)
주소: 130-743 서울 동대문구 전농동 90
소속: 서울시립대학교 컴퓨터과학부
전화: 02) 2210-5322
E-mail: heisco@hanmail.net

▶ 유하진(Ha-Jin Yu) : 교신저자
주소: 130-743 서울 동대문구 전농동 90
소속: 서울시립대학교 컴퓨터과학부
전화: 02) 2210-5613
E-mail: hjyu@uos.ac.kr