Robust Speech Recognition Using KPCA-Based Noise Classification

Nattanun Thatphithakkul[1,2] , Boontee Kruatrachue[1], Chai Wutiwiwatchai[2], Sanparith Marukatat[2] and Vataya Boonpiam[2]

[1]King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand
S6060008@kmitl.ac.th and kkboontee@kmitl.ac.th,
[2]National Electronics and Computer Technology Center, Phathumthani, 12120, Thailand
chai@nectec.or.th, sanparith.marukatat@nectec.or.th and vataya.boonpiam@nectec.or.th

ABSTRACT

This paper proposes an environmental noise classification method using kernel principal component analysis (KPCA) for robust speech recognition. Once the type of noise is identified, speech recognition performance can be enhanced by selecting the identified noise specific acoustic model. The proposed model applies KPCA to a set of noise features such as normalized logarithmic spectrums (NLS), and results from KPCA are used by a support vector machines (SVM) classifier for noise classification. The proposed model is evaluated with 2 groups of environments. The first group contains a clean environment and 9 types of noisy environments that have been trained in the system. Another group contains other 6 types of noises not trained in the system. Noisy speech is prepared by adding noise signals from JEIDA and NOISEX-92 to the clean speech taken from NECTEC-ATR Thai speech corpus. The proposed model shows a promising result when evaluating on the task of phoneme based 640 Thai isolated-word recognition.

Keywords: Speech recognition, Kernel PCA, SVM

1. INTRODUCTION

It is commonly known that a speech recognition system trained by speech in a clean or nearly clean environment cannot achieve good performance when working in noisy environment. Research on robust speech recognition is then necessary. This paper focuses on the construction of robust model approach which has achieved good recognition results [1]. Generally, this model-based approach aims to create an environment-specific acoustic model or to adapt the existing model to the specific environment. Several techniques of model adaptation have been proposed e.g. linear regression adaptation and parallel model combination [2]. However, an acoustic model trained directly for specific noise is certainly superior to the adapted model, although multiple acoustic models are needed for various kinds of noise and an accurate automatic noise classification is required.

Many noise classification techniques have been studied previously. Classical technique is based on hidden markov models (HMM), linear prediction coefficients (LPC) [3] and mel-frequency cepstral coefficients (MFCC) [4], which have been proven to give better results than human listeners [4]. Another successful technique is a neural network based system with combined features of line spectral frequencies (LSF) [5], a zero-crossing (ZC) rate and energy [6]. However, implementing LSF in a real-time system is problematic. Therefore, we aim to explore a simpler feature extraction method for noise classification.

In recent years, many kernel-based classification techniques, e.g. support vector machine (SVM) [7], kernel principal component analysis (KPCA) [8-12], kernel discriminate analysis (KDA) [13], kernel fisher discriminate analysis (FDA) [14], have been proposed. These techniques have been successfully applied, not only for classification, but also for regression and feature extraction e.g. in speech recognition [8] and image recognition system [12].

This paper proposes another application of KPCA, which is noise classification. In this work, KPCA is applied to extract speech features, which are used by a pattern classifier for noise classification. An advantage of KPCA is that useful noise information can be extracted from the original feature. The computational requirement of KPCA applied to normalized logarithmic spectrums (NLS) implemented in this paper is similar to that of the MFCC or other effective features such as LSF, but with higher classification accuracy.

Our noise classification model is evaluated on 2 groups of environments. The first group contains 10 classes of environments that have been trained in the system. The second group is another set of 6 environments not trained in the system. Evaluating by the later group shows the speech recognition performance in unknown-noise environments. All noises are taken from Japan JEIDA [15] and NOISEX-92 [16]. Our Thai 640 isolated-word recognition with noise-specific acoustic models is used in the evaluation. It is noted that although the task is isolated-word recognition, phonemes are used as basic recognition units. This facilitates new word addition.

The rest of paper is organized as follows: the next section describes an overall structure of our robust speech recognition system. In Sect. 3, the KPCA algorithm is described. Sect. 4 describes our experiments, results and discussion. The last section concludes the paper and notices our future works.

2. ROBUST SPEECH RECOGNITON USING NOISE CLASSIFICATION

As described in the previous section, our robust speech recognition system uses the model-based technique, in which acoustic models are trained by speech in specific environment. An overall structure is illustrated in Fig. 1. Given a speech signal, a set of features for noise classification is extracted from a short period of silence at the beginning of signal. It is noted that this short period is assumed to be a silence where the speaker has not yet uttered. This assumption holds for our push-to-talk interface. To apply our system with other user interfaces, we need an additional module of speech/non-speech classification or other strategies to capture a non-speech portion from the input signal. Features extracted from the silence portion are then used to identify the type of environment. Once knowing the environment type, the recognizer selects a corresponding acoustic model for recognizing the rest of signal.

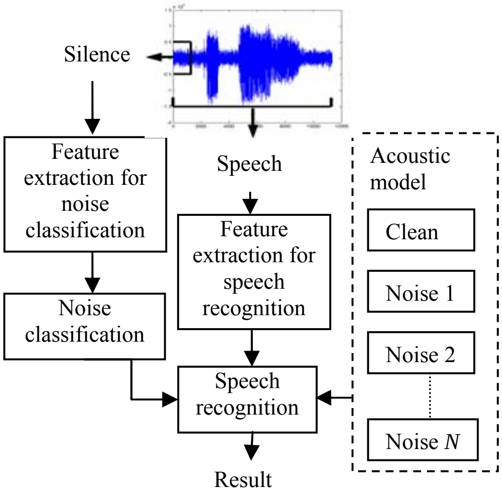

Fig. 1: *Overall structure of robust speech recognition.*

With this model, there are 3 particular difficulties:

- How to construct a robust acoustic model for a variation of signal-to-noise ratios (SNR)? In our system, a particular acoustic model is trained on noisy speech with various levels of SNR. Clean speech, whose SNR exceeds 30 dB is also combined in the training set of each noisy acoustic model.
- How to construct the environment or noise classification module? Time consuming by the noise classification module should be as low as possible, so that the overall system can achieve an acceptable processing time. The construction of such module is the main objective of this paper.
- How can the robust speech recognition model deal with unknown noises, i.e. noises not trained in the model? Normally, several major noises are trained in the system and each of other noises is expected to be classified as one of the major noises. This paper also reports the effect of our model for unknown-noise classification.

In this paper, speech features evaluated for noise classification include NLS, LSF, LPCC and MFCC. PCA and KPCA are applied to these basic features in order to extract meaningful features and enhance noise classification performance. For the noise classification algorithm, a fast and efficient technique is needed. In our experiment, a well-known SVM algorithm is evaluated. Speech recognition utilizes a state-of-the-art algorithm of HMM with MFCC as speech features.

3. KERNEL PRINCIPAL COMPONENT ANALYSIS

3.1 Kernel functions

The use of nonlinear kernel functions is a strategy to raise the capability of simple algorithms such as PCA in dealing with more complicated data. Indeed, extending these algorithms for a non-linear case may be done by replacing the involved variables by their values on a new feature space. Transformation from the original space to a new space may be done by some mapping function Φ. However, by choosing an appropriate mapping function, the dot product in the new feature space can be performed by a nonlinear function in the input space, the so-called kernel function. Hence, by replacing the dot product involving in a classical algorithm by some kernel function, we can extend this algorithm to the non-linear case. This is usually referred to as the kernel trick [10]. The commonly used kernels are shown in Table 1.

Table 1: *Some useful kernel functions.*

Kernel function	Equation
Polynomial	$k(x_i, x_j) = (x_i \cdot x_j + 1)^d$
RBF	$k(x_i, x_j) = \exp(-\left\|x_i - x_j\right\|^2 / g)$

3.2 KPCA

The idea of KPCA [8-9] is to extend the classical PCA for non-linear projection using the kernel trick. Given a set of M samples x_i, i =1,2,.....,M with $x_i \in R^n$. The classical PCA is done by computing eigenvectors and eigenvalues of the covariance matrix of these examples. Let X = [x_1; x_2; …; x_M] be the matrix of these M examples, the covariance matrix is defined by $C = M^{-1} X X^T$. The normalized eigenvectors of C form the principal subspace on which the data will be linearly projected. To extend this approach using the kernel trick, we first notice that if we dispose an eigen-couple (λ,v) of the dot product matrix $X^T X$ then we can also derive an eigen-couple (λ',v') of the covariance matrix C. Indeed, we have $\lambda.v = X^T X\, v$, so by pre-multiplying both sides of the equation by $M^{-1} X$ we get $(\lambda M^{-1})(X v) = (M^{-1} X X^T)(X v) = C\,(X v)$. This means that $\lambda' = \lambda M^{-1}$ and $v' = X v$ forms an eigen-couple of the covariance matrix C. The kernel trick is then applied by replacing the dot product in X^TX by a kernel function. It should be noted that the eigenvector v' produced by this procedure may not be properly normalized. Therefore an additional normalization step is needed. The overall KPCA algorithm is as follow:

- Compute the kernel matrix K with $K_{ij} = k(x_i, x_j)$ where k is a kernel function.
- Compute the eigen-couples of K. Let (λ_k, v_k), $k = 1, ..., M$ be these eigen-couples.
- Normalize the k^{th} principal axis by computing $v_{ki} = v_{ki}\,\lambda_k^{-1/2}$. $(\lambda_k > 0)$
- The projection of a vector $y \in R^n$ onto the k^{th} principal axis is done by computing $\sum_{i=1}^{M} v_{ki} k(x_i, y)$. For simplification, we will call the feature vector projected on the principal subspace, the "weight vector" hereafter.

For simplification, we will call the feature vector projected on the principal subspace, the "weight vector" hereafter. While a basic speech feature such as NLS is effective, an optimal order of the NLS is considerably large. With limited training set, computing the eigen decomposition from a dot matrix, or kernel matrix, can be done more accurately [11].

4. EXPERIMENTS

4.1 Data preparation

Noises used in our experiments are from the JEIDA and NOISEX-92. They are clustered to 2 groups. The first group contains 8 kinds of noise from JEIDA, including crowded street, machinery factory, railway station, large air-condition, trunk road, elevator, exhibition in a booth, and ordinary train, 1 large-size car noise from NOISEX-92, and an additional clean environment. The second group contains other 6 kinds of noise from JEIDA, including exhibition in a passage, road crossing, medium-size car, computer room, telephone booth, and press factory. The former group of environments is reserved for training the noise classification and speech recognition models, and for testing the system for "known" noises (noises recognizable by the system). The later group is used for evaluating the system for "unknown" noises (noises not trained in the system).

Noisy speech was prepared by adding the noise from JEIDA or NOISEX-92 to the clean speech of NECTEC-ATR [17] at various SNRs (0, 5, 10 and 15 dB). The pre-processed data were then clustered into several sets for noise classification and speech recognition experiments as summarized in Table 2.

4.1.1. Data set for noise classification

Three sets were prepared: a PCA and KPCA training set, a classifier training set and classifier test sets. The first set was used for computing PCA and KPCA weight vectors. The second set was used for training the noise classifier and the rest were used for evaluating the classifier.

A small frame of 1,024 samples at the beginning of the speech signal, which was expected to be silence, was used for PCA, KPCA and noise classification. As described in the Sect. 3, our speech recognizer is designed for a push-to-talk interface. With this interface, we can control the recorder to start record a silence signal before the beginning of speech. NLS and LSF used for noise classification were computed from this silence frame.

4.1.2 Data set for speech recognition

The speech recognition task in our experiment was phoneme-based 640 isolated-word recognition. 32000 speech utterances from 32 speakers were allocated for a training set. Another set of 6400 utterances from other 10 speakers are used for testing in both known and unknown-noise modes. The HMMs representing 35 Thai phones [18]. Each triphone HMM consisted of 5 states and 8 Gaussian mixtures per state. MFCC 39 dimensional vectors (12 MFCC, 1 log-energy, and their first and second derivatives) were used as recognition features.

Table 2: *Number of utterances in experimental data sets.*

TASK	DATA SET	AMOUNT
Noise classification	PCA/KPCA training	3,900
	Classifier training	24,000
	Known-noise test	256,000
Speech recognition	Recognizer training	32,000*
	Known-noise test	6,400*
	Unknown-noise test	6,400*

*Number of samples per noise per SNR

4.2 Noise classification results

Our proposed classification model using KPCA and SVM described in the Sect. 3 was compared to the classical technique using a HMM classifier [3-4], which served as a baseline system in our experiment. The noise-classification data sets are used in this section. The followings are details of noise classification experiments.

4.2.1 Classification using a HMM system

For the HMM [19] based noise classification system, we have varied the number of states as well as the number of Gaussian mixtures per state. The same set of MFCC and LPC features are used as classification features. This baseline system will be referred to as "HMM_MFCC" and "HMM_LPC". Fig. 2 and Fig. 3 present results of the evaluation of this system on the known-noise test set.

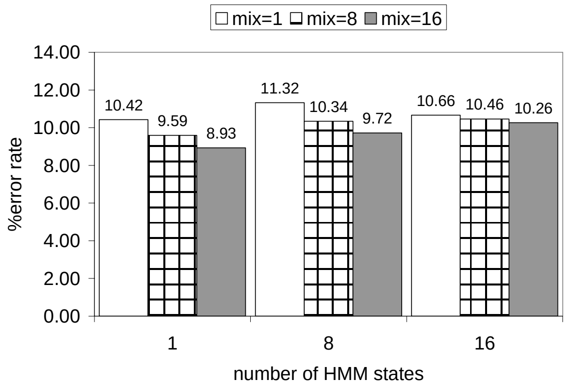

Fig. 2: *Error rate results (%) of known-noise classification based on HMM_MFCC*

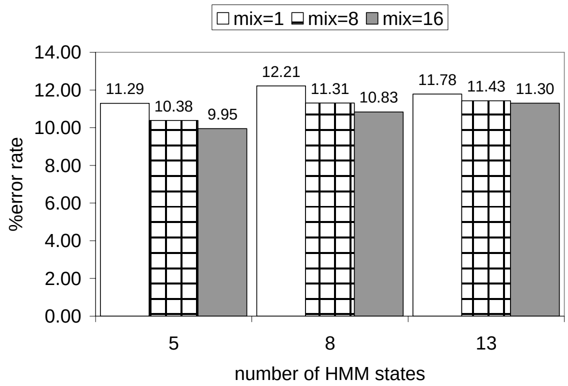

Fig. 3: *Error rate results (%) of known-noise classification based on HMM_LPC.*

4.2.2 Classification using SVM systems

A multi-class SVM [20] classifier based on one-against-one algorithm. Two kinds of kernel functions, RBF and Polynomial, are evaluated. PCA and KPCA are applied to three types of speech features including NLS (511 orders), LSF (10 orders) and MFCC (10, 12, 16 and 20 orders without energy and derivative features). The order of PCA and KPCA weight vectors is empirically tuned for each comparison. The known-noise test set is also used for evaluation in this section. Results and discussions are as follows.

A preliminary experiment consists in comparing the three speech features namely NLS, LSF and MFCC as well as the kernel used in the SVM classifier. The Fig 4 and 5 show the results obtained from MLS and LSF features using polynomial and RBF kernel respectively. The results obtained from MFCC with various orders are shown in Fig 6 and 7 for polynomial and RBF kernel respectively.

From these 4 figures, we can see that the best result is obtained by the RBF-kernel SVM using NLS.

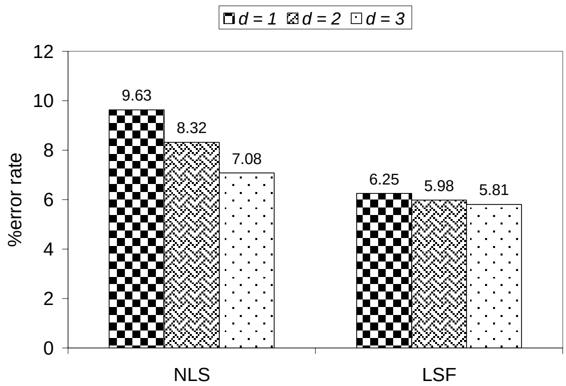

Fig. 4: *Error rate results (%) of known-noise classification based on SVM (10-order LSF and 511-order NLS, kernel functions of SVM: Polynomial).*

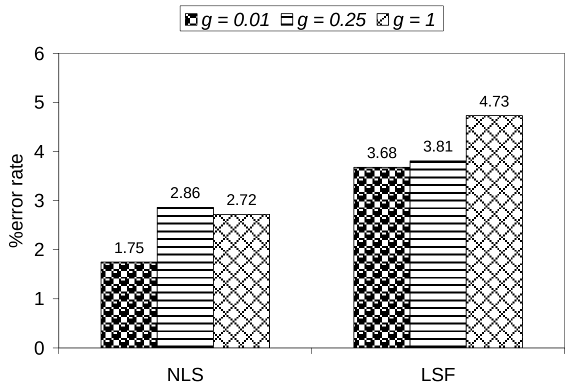

Fig. 5: *Error rate results (%) of known-noise classification based on SVM (10-order LSF and 511-order NLS, kernel functions of SVM: RBF).*

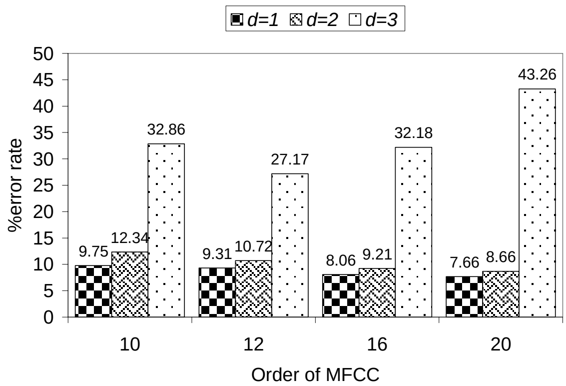

Fig. 6: *Error rate results (%) of known-noise classification based on SVM (MFCC with various orders, kernel functions of SVM: Polynomial).*

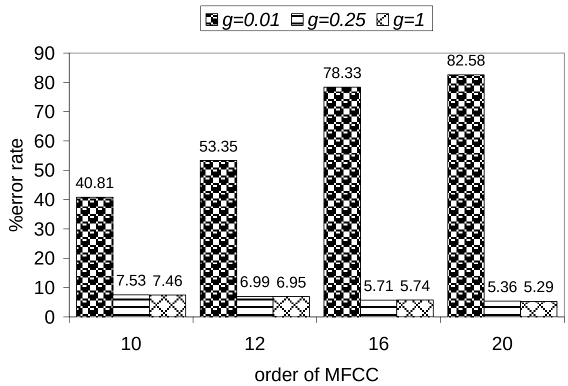

Fig. 7: *Error rate results (%) of known-noise classification based on SVM (MFCC with various orders, kernel functions of SVM: RBF).*

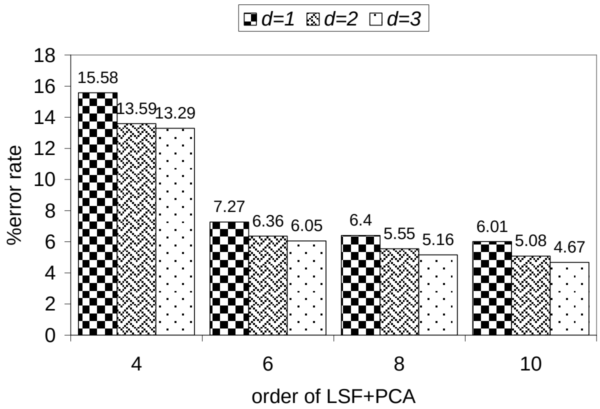

Fig. 8: *Error rate results (%) of known-noise classification based on SVM (LSF+PCA with various orders, kernel functions of SVM: Polynomial).*

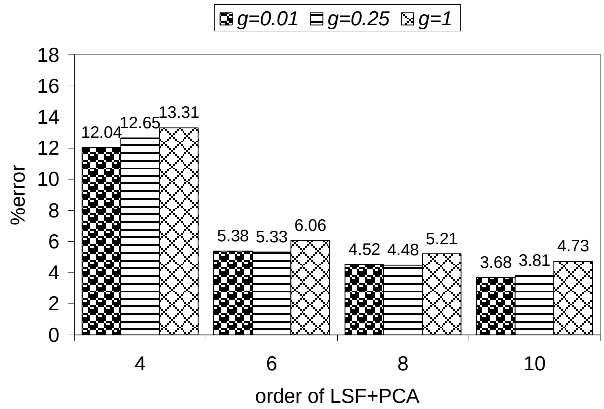

Fig. 9: *Error rate results (%) of known-noise classification based on SVM (LSF+PCA with various orders, kernel functions of SVM: RBF).*

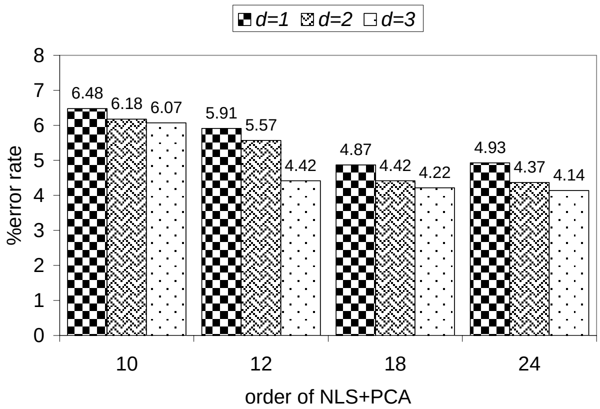

Fig. 10: *Error rate results (%) of known-noise classification based on SVM (NLS+PCA with various orders, kernel functions of SVM: Polynomial).*

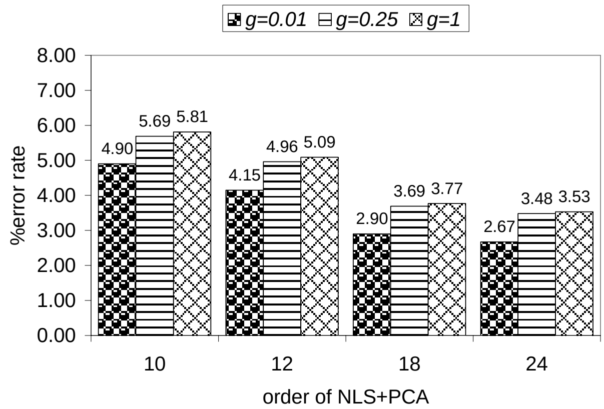

Fig. 11: *Error rate results (%) of known-noise classification based on SVM (NLS+PCA with various orders, kernel functions of SVM: RBF).*

However, a large order of NLS is needed to achieve such performance (511 orders in our case). The large number of features requires a longer time and larger storage to process. Reducing the order of NLS without a drawback of performance degradation is thus interesting.

Next, we investigate the effect of dimension reduction via PCA on the accuracy of our classifier. Applications of PCA on the 10-order LSF (denoted as LSF+PCA) and 511-order NLS (denoted as NLS+PCA) are then performed and results are shown in Fig. 8-11. The Fig 8 and 9 show the results obtained from LSF+PCA feature when using polynomial and RBF kernel respectively. The Fig. 10 and 11 show the error rate obtained with NLS+PCA. From our preliminary experiments, the classification accuracy trends to be saturated when the order of PCA exceeds 24. Hence these 2 figures (10 and 11) show only the results obtained from NLS+PCA up to the order of 24.

From these 4 figures, it is clear that using the PCA-based feature of NLS and LSF does degrade the classification accuracy, with the advantage of faster processing time. For LSF+PCA, changing from 10 orders to 6 orders, we increase about 2% error rate while the gain in processing time is not significant. For NLS+PCA, reducing from full 511 orders to 24 orders allows us gaining a significant processing time, while increasing only a slight error rate. It should be noted that, even if the order of NLS+PCA is higher than that of the LSF, computing the LSF is much more complex than the NLS+PCA. From these results, the 24 first principal components of NLS with RBF kernel is a suitable choice for the noise classification module.

The objective of the next experiment is to see whether moving from the classical linear PCA to the non-linear analysis of KPCA allows further improvement KPCA has proved to be efficient for speech recognition [4]. In this experiment, RBF kernel is used for the KPCA (RBF at g = 0.1). Results of applying KPCA to the NLS (NLS+KPCA) are shown in Fig. 12 and Fig. 13 for polynomial and RBF kernel of the SVM classifier respectively. The lowest error rate achieved is 2.35% obtained from 24-order KPCA and RBF-kernel SVM, which is also the best case comparing to all previous experiments of PCA and KPCA. This also underlines the advantage of using non-linear analysis in extracting significant features by KPCA.

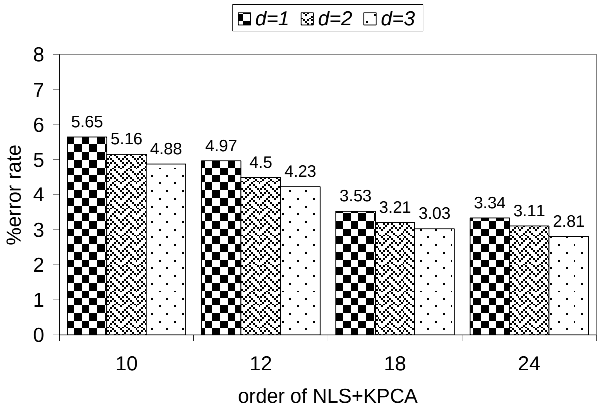

Fig. 12: *Error rate results (%) of known-noise classification based on SVM (NLS+KPCA (RBF at g = 0.1) with various orders, kernel functions of SVM: Polynomial).*

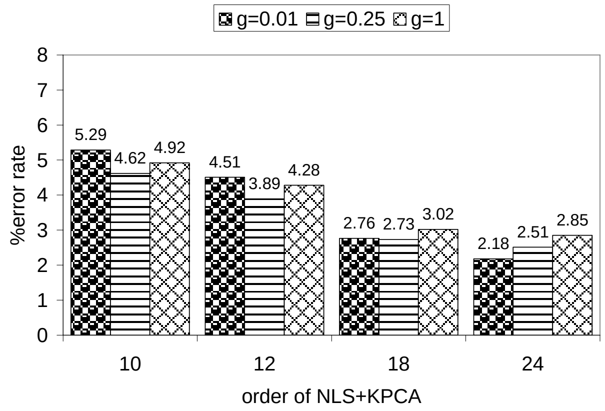

Fig. 13: *Error rate results (%) of known-noise classification based on SVM (NLS+KPCA (RBF at g = 0.1) with various orders, kernel functions of SVM: RBF).*

4.2.4 Comparison to other noise classification techniques

In this section, we evaluate the SVM classifier working on features extracted from 511 order NLS using PCA and KPCA against other approaches. The two systems are denoted as "SVM_PCA" and "SVM_KPCA" respectively. We use the order of 24 for the extracted feature from both PCA and KPCA. This order is selected empirically in previous experiments.

Fig. 14 shows the results obtained from different noise classification models using various kinds of features including our proposed KPCA-based feature. Other noise environment classifiers include the HMM with LPC and with MFCC features, the SVM with full 511-order NLS, 10-order LSF and 20-order MFCC (without energy and derivative features).

From these results, the SVM classifiers outperform the HMM classifier in all case. Moreover, the SVM with LSF and MFCC give the error rate of 3.63% and 5.29% respectively. It should be noted that, the same error rate of 3.63% were obtained when applying PCA to the 10-order LSF. According to the results, the KPCA outperforms the other, except the NLS. The NLS, however, requires the largest order (511) to achieve the underlying result. Trading off between the accuracy and running time, we found the use of SVM_KPCA optimal our noise classification module.

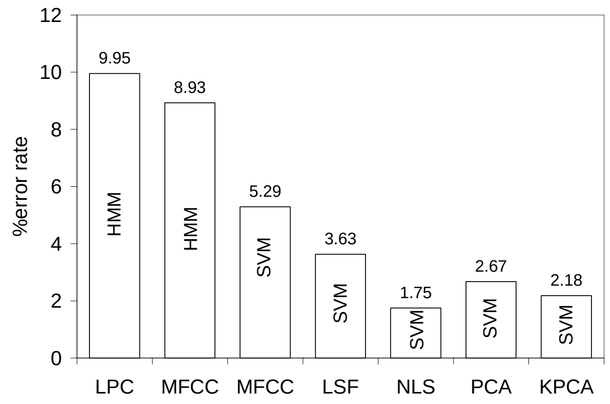

Fig. 14: *Comparative results of known-noise classification error rates using various kinds of classification system.*

4.3 Speech recognition results

In this section, several robust speech recognition techniques including our proposed model are experimentally compared. The first system (S1) was a conventional system without any implementation for robust speech recognition. The second system (S2) used zero-mean static coefficients [19], a well-known technique for noise-robust speech features. The third system (S3) was our proposed model, where input speech environment was identified and the corresponding acoustic model was chosen for recognition. In the S3 system, an acoustic model for each environment was trained by multi-SNR (5, 10, and 15 dB) data including each noise. The SVM_KPCA system (RBF at g = 0.1), which achieved the best result, was used in the S3 system. The fourth system (S4) was as similar as the S3 system except that the noise classifier was replaced by the HMM_MFCC model. The next system (S5) was an ideal system, where noise is perfectly classified, i.e. 0% noise classification error. In order to underline the importance of the classification module, we also considered the last system (S6) which is equipped with random noise classification module. These two systems, S5 and S6, indicate the upper and the lower bounds of the recognition system using noise specific HMM. In the following experiments, the speech recognition data sets are used.

4.3.1 Speech recognition in known-noise

Evaluated by the known-noise test set, comparative results are shown in Table 3. It is obvious that our proposed model (S3) achieved the best recognition

Table 3: *Comparative results of robust speech recognition in known-noise environment.*

Environments	Word accuracy (%)					
	S1	S2	S3	S4	S5	S6
Clean	93.02	92.45	93.02	92.91	93.02	86.22
Street	65.57	67.92	83.28	83.15	83.39	75.65
Factory	41.65	47.33	75.69	75.61	75.68	53.03
Station	45.12	52.79	77.62	77.44	77.67	63.69
Air condition	42.46	53.51	81.15	81.12	81.17	63.83
Road	53.90	56.99	77.30	76.35	77.39	64.68
Elevator	52.88	59.01	81.49	81.36	81.47	70.25
Exhibition	41.57	56.42	82.47	82.45	82.49	66.68
Train	25.43	45.72	79.66	79.63	79.70	49.55
Car	89.08	88.98	89.93	90.20	89.95	78.86
Average	55.07	62.11	82.16	82.02	82.19	67.24

Table 4: *Comparative results of robust speech recognition in unknown-noise environments*

Environments	Word accuracy (%)			
	S1	S2	S3	S4
Exhibition	63.17	71.69	85.25	86.35
Road	45.61	60.54	77.09	76.83
Car	78.42	82.02	86.07	86.39
Computer room	40.20	56.71	77.69	78.13
Telephone booth	37.87	45.31	70.56	71.38
Factory	31.84	49.03	72.42	72.40
Average	49.52	60.88	78.18	78.58

results in every case and the results are almost equal to the ideal case (S5).

4.3.2 Speech recognition in unknown-noise

Evaluated by the unknown-noise test set, comparative results are shown in Table 4. Although it is not significant, the S4 system outperforms the S3 system. One possible reason is that the SVM classifier might over fit to the trained classes and hence underperformed the HMM classification in handling unknown classes.

The results in Table 3 and 4 also underline the advantage of using noise classification module (S3 and S4) compared to conventional system (S2), even in unknown noise environments.

4.4 Hybrid noise classification system

Although the SVM_KPCA classifier outperformed other classifiers, an intensive analysis showed that its errors can be recovered by selecting the noise model proposed by other classifier. Hence, we have also evaluated a hybrid architecture in which the SVM_KPCA is used in conjunction with the HMM_MFCC or the SVM_MFCC. Indeed, in this hybrid system, if both classifiers agree in noise classification, the corresponding noise model is used for recognition. Otherwise, we choose among the acoustic models proposed by both classifiers, the one which maximizes the acoustic probabilities. This combined system of SMV_KPCA and HMM_MFCC gives 82.20% accuracy on known-noise test set and 78.90% on unknown-noise test set. This combined system of HMM_MFCC and SVM_MFCC gives 82.21% on known-noise test set and 78.78% on unknown-noise test set. The overall running time is increased but still being faster than the NLS.

5. CONCLUSION AND FUTURE WORKS

This paper proposed a novel technique of robust speech recognition based on model selection. The recognizer selected a specific acoustic model from a pool of acoustic models that were trained by speech data in each type of noisy environment. A noise classification module was used to identify the type of environment. KPCA applied to the NLS was proposed for the noise classification features, and SVM was used as the noise classifier Experiments showed that the proposed model gave a promising result. When combining the model to the speech recognizer, the proposed system produced almost equal recognition accuracy to the ideal system, where the type of noisy environment was given. The proposed system working with known-noise environments achieved 20.05% higher recognition accuracy over the robust system using zero-mean static coefficients, and 0.14% higher accuracy over the baseline system using the HMM and MFCC for noise classification. A hybrid system that combined our proposed model and the baseline model was also investigated. Experimental results showed a small improvement over each individual model on both known and unknown noises.

For future works, a better way to treat unknown-noises will be intensively explored. Optimization of SVM training will be performed to avoid over-training if this is the case. Other successful classifiers such as an optimal Bayes as well as applications of PCA and KPCA to other effective speech features such as MFCC will be investigated. Another interesting topic is to reduce the number of specific acoustic models by automatic clustering of noises and constructing one acoustic model for each noise cluster.

6. REFERENCES

[1] M.J.F. Gales, “Model-based techniques for noise robust speech recognition”, PhD thesis University of Cambridge, 1995.

[2] Y. Gang, “Speech recognition in noisy environments: A survey”, Speech Communication, Vol. 16, pp. 261-291, 1995.

[3] P. Gaunard, C.G. Mubikangiey, C. Couvreur, and V. Fontaine, “Automatic classification of environmental noise events by hidden markov models”, Proceedings of ICASSP1998, pp. 3609-3612, 1998.

[4] L. Ma, D. Smith and B. Milner, “Context awareness using environmental noise classification”, Proceedings of Eurospeech2003, pp. 2237-2240, 2003.

[5] K.E. Maleh, A. Samouelian and P. Kabal, “Frame-level noise classification in mobile environments”,

IEEE conf. Acoustics, Speech, Signal Processing, pp 237-240, 1999.
[6] C. Shao and M. Bouchard, "Efficient classification of noisy speech using neural networks", Proceedings of ISSPA2003, pp. 357-360, 2003.
[7] N. Cristianini and J.S. Taylor. "An introduction to support vector machines and other kernel-based learning methods", Cambridge: Cambridge University Press, 2000
[8] A. Lima, H. Zen, Y. Nankaku, C. Miyajima, K. Tokuda and T. Kitamura, "On the use of kernel PCA for feature extraction in speech", IEICE Tran. INF.&SYST., Vol.E87-D, pp. 2802-2811, 2004.
[9] N. Thatphithakkul, B. Kruatrachue, C. Wutiwiwatchai, S. Marukatat and V. Boonpiam, "KPCA-Based Noise classification Module for Robust Speech Recognition system", Proceeding of ECTI-CON2006, pp. 231-234, 2006.
[10] B. Scholkopf, A. Amola and K.-R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem", Neural computation, 10:1299-1319, 1998.
[11] M. Turk and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, 3(1):71-86, 1991.
[12] K.I. Kim, K. Jung and H.J. Kim, "Face recognition using kernel principal component analysis", IEEE Signal Processing. Lett. vol.9, no.2, pp.40-42, 2002.
[13] V. Roth and V. Steinhage, "Nonlinear discriminate analysis using kernel function", Advances in neural information processing systems, pp. 568–574, 2000.
[14] S. Mika, G. Ratsch, J. Weston, B. Scholkopf and K.-R. Muller, "Fisher Discriminate Analysis with Kernels", Neural Networks for Signal Processing IX, pp. 41-48, 1999.
[15] www.milab.is.tsukuba.ac.jp/corpus/noise_db.html
[16] NOISEX-92. http://www.speech.cs.cmu.edu/ comp. speech/ Section1/ Data/noisex.html
[17] S., Kasuriya, V. Sornlertlamvanich, P. Cotsomrong, T. Jitsuhiro, G. Kikui and Y. Sagisaka, "Thai speech database for speech recognition", Proceedings of Oriental COCOSDA2003, pp 105-111, 2003.
[18] S. Kasuriya, S. Kanokphara, N. Thatphithakkul, P. Cotsomrong and T. Sunpethniyom, "Context-independent acoustic models for Thai speech recognition", Proceedings of ISCIT2004, pp.991-994, 2004.
[19] The HTK book version 3.1, Cambridge University, http://htk.eng.cam.ac.uk, 2001.
[20] LIBSVM – A library for Support Vector Machines. http://csie.ntu.edu.tw/~cjlin/libsvm/

Nattanun Thatphithakkul received the B.Eng and M.Eng degree from Suranaree University, Thailand, in 2000 and in 2002, respectively. He is currently a Ph.D. student at King Mongkut's Institute of Technology Ladkrabang in Computer Engineering. His research activities are oriented toward robust speech recognition and noise model adaptation.

Boontee Kruatrachue received the BS. in Electrical Engineering from Kasetsart Univeristy, Thailand, in 1981, and M.S. and Ph.D degrees in Electrical Engineering from Oregon State Univeristy, USA., in 1984 and 1987, respectively. During 1988-1990, he was Software Engineer at Astronautics Corporation of America, Wisconsin, USA. He is now associate professor at computer engineering department, King Mongkut's Institute of Technology Ladkrabang, Thailand. His research interests include pattern recognition, data mining and machine learning.

Chai Wutiwiwatchai received B.Eng. (the first honor) and M.Eng. degrees of electrical engineering from Thammasat and Chulalongkorn University, Thailand in 1994 and 1997 respectively. He received Ph.D. from Tokyo Institute of Technology in 2004 under a scholarship of Japanese government. He is now Chief of the Speech Technology Section of the National Electronics and Computer Technology Center (NECTEC), Thailand. His research interests include speech and speaker recognition, natural language processing, and human-machine interaction.

Sanparith Marukatat received the License and Maîtrise degree from University of Franche-Compté. He has finished his DEA (a kind of French one-year Master degree) and his doctoral degree at University of Paris 6 in 2000 and 2004 respectively. He is currently a researcher in the Information R&D Division at National Electronics and Computer Technology Center (NECTEC), Thailand. His research interests include classification problem, subspace projection and sequence modelling.

Vataya Boonpiam received the B.Sc and M.Sc degree from King Mongkut's Institute of Technology North Bangkok, Thailand, in 2000 and in 2004, respectively. Her research interests include speech recognition. She is currently a researcher of Information Research and Development Division, National Electronics and Computer Technology Center (NECTEC).