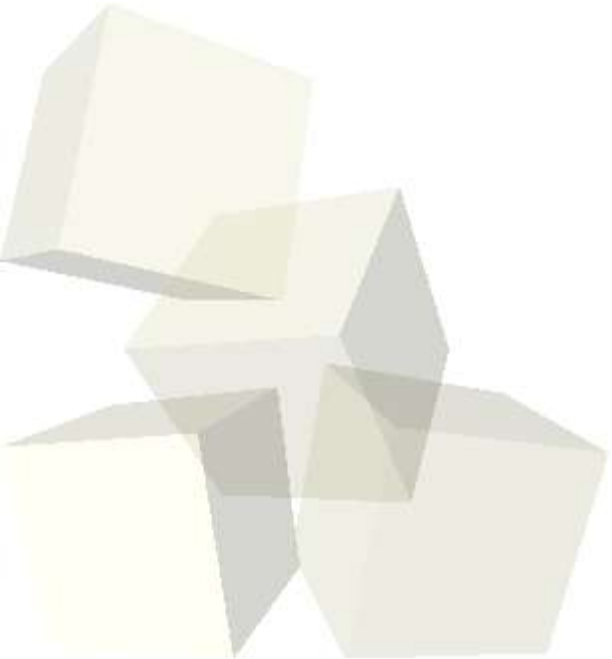
Noise Classification based on PCA

Nattanun Thatphithakkul, Boontee Kruatrachue,
Chai Wutiwiwatchai, Vataya Boonpiam



Outline

- Introduction
- Principle component analysis (PCA)
- Classification using PCA
- Experiment
- Conclusion

Introduction

- Problem
  - One way to make a robust speech recognizer is to train by noisy speech. However, with various kinds of noise, the recognizer cannot achieve an acceptable accuracy.

- Solution
  - To classify noise and select an acoustic model generated for that noise.

Introduction

- Propose Method
  - Principal Component Analysis (PCA) is used to reduce the dimension of analyzed feature.
  - A new classification structure is introduced.

- Principal Component Analysis
  - Step 1 Calculate the covariance matrix

$$\mu = \frac{1}{M} \sum_{i=1}^{M} x_i \tag{1}$$

$$\Phi_j = x_j - \mu \tag{2}$$

$$C = \frac{1}{M} \sum_{j=1}^{M} \Phi_j \Phi_j^T \tag{3}$$

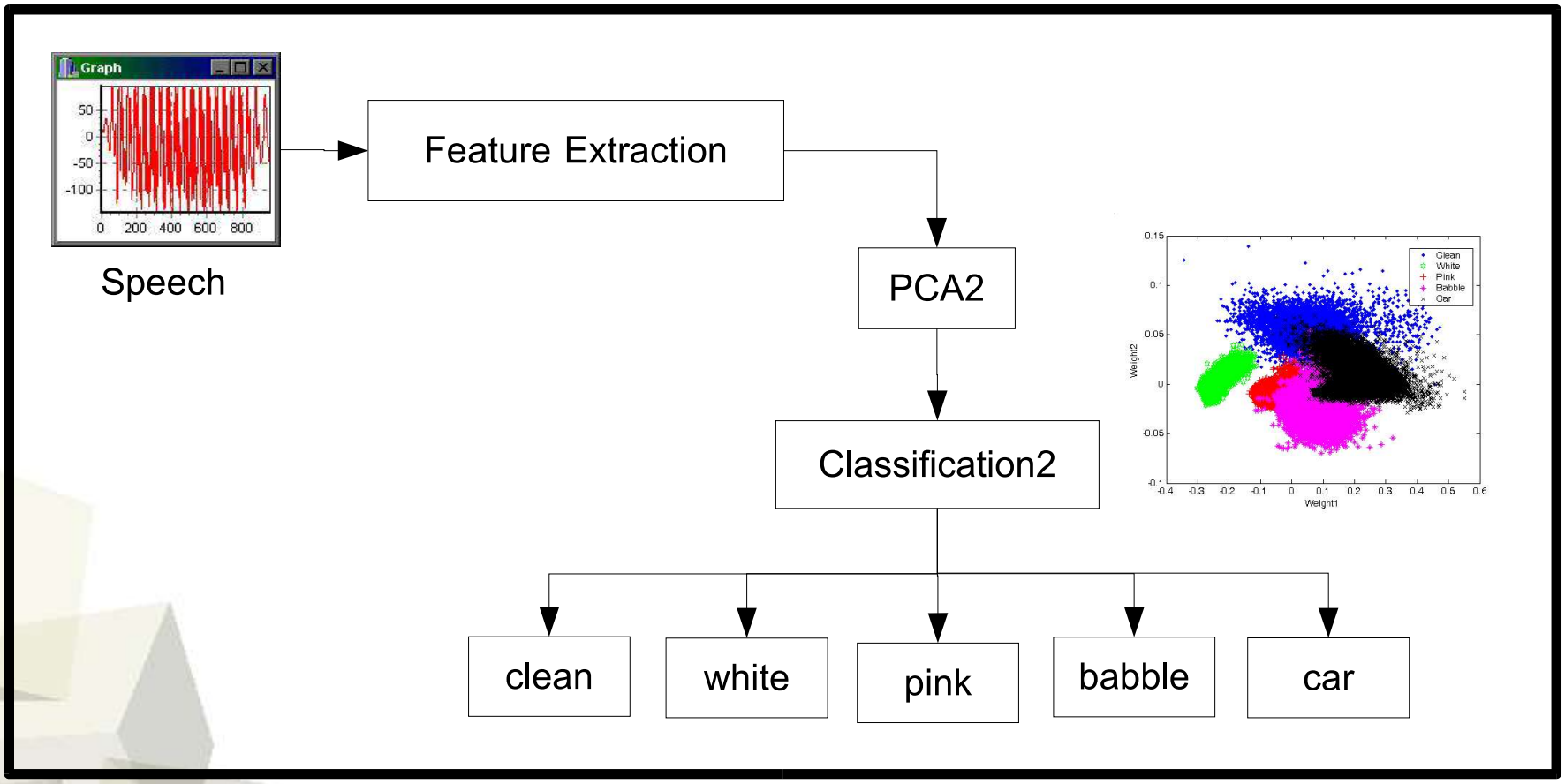
- Principal Component Analysis
  - Step 2 Calculate the eigenvectors and eigenvalues of the covariance matrix and select N eigenvectors, which correspond to the largest eigenvalue.
  - Step 3 Calculate the weight vector

$$w_k = v_k^T (x - \mu) \qquad k = 1, \ldots\ldots, N \tag{4}$$

$$W^T = [\, w_1 w_2 \ldots\ldots w_N \,] \tag{5}$$

Classification using PCA

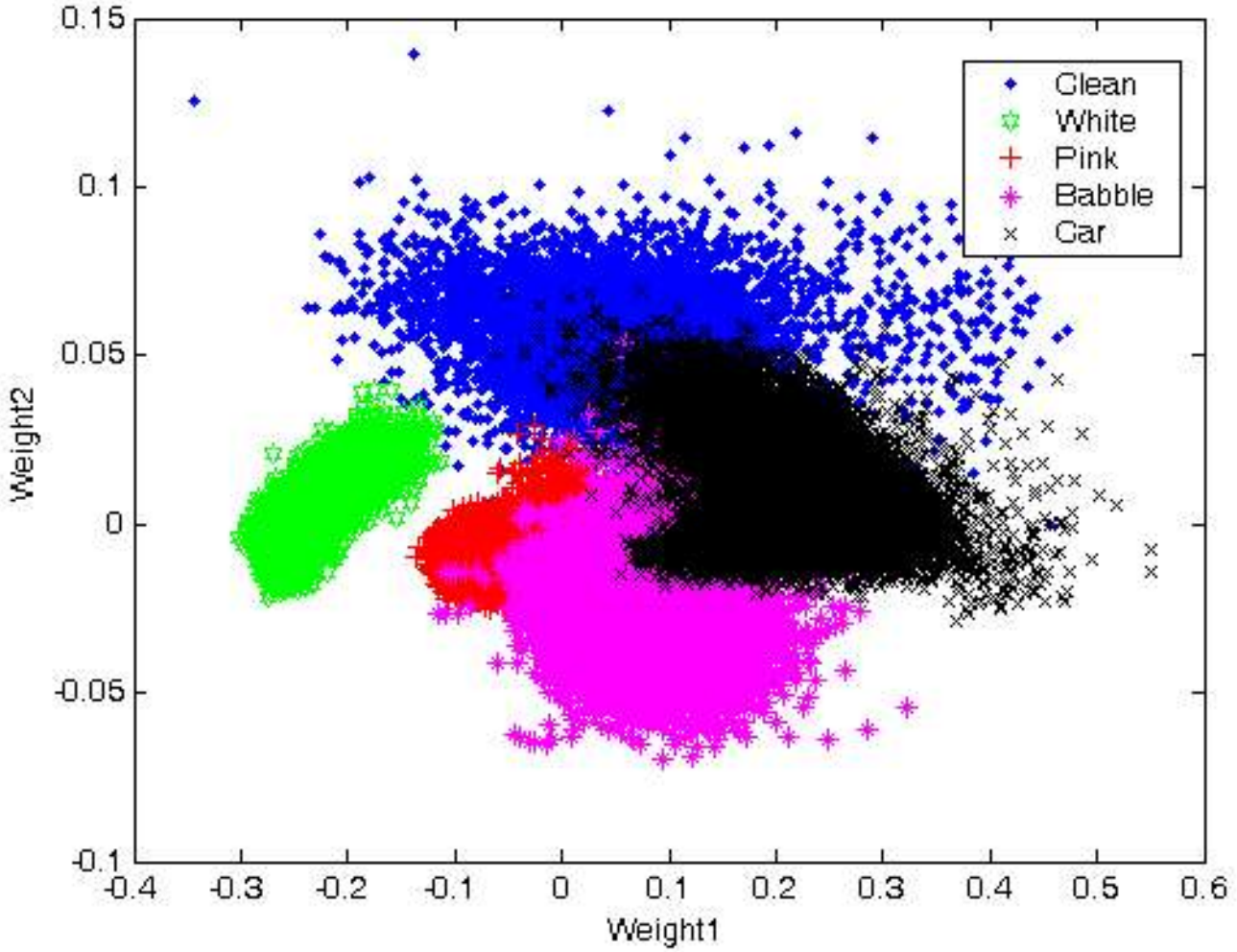
- PCA used for reducing dimension of feature vector before classification.



Structure of Classical Classification using PCA (baseline)

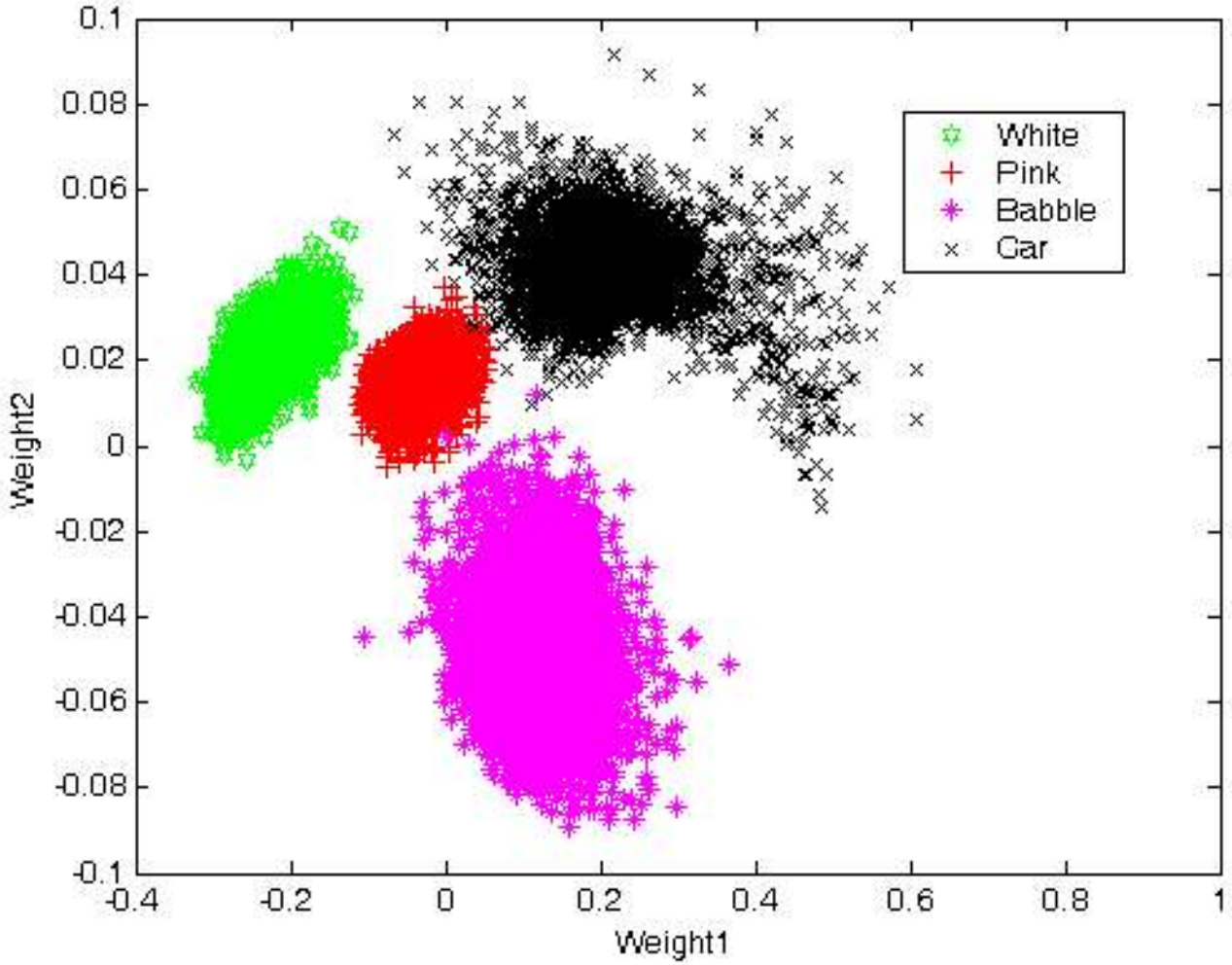
Classification using PCA

- Weight vector of principle components (Train -> 5 types)



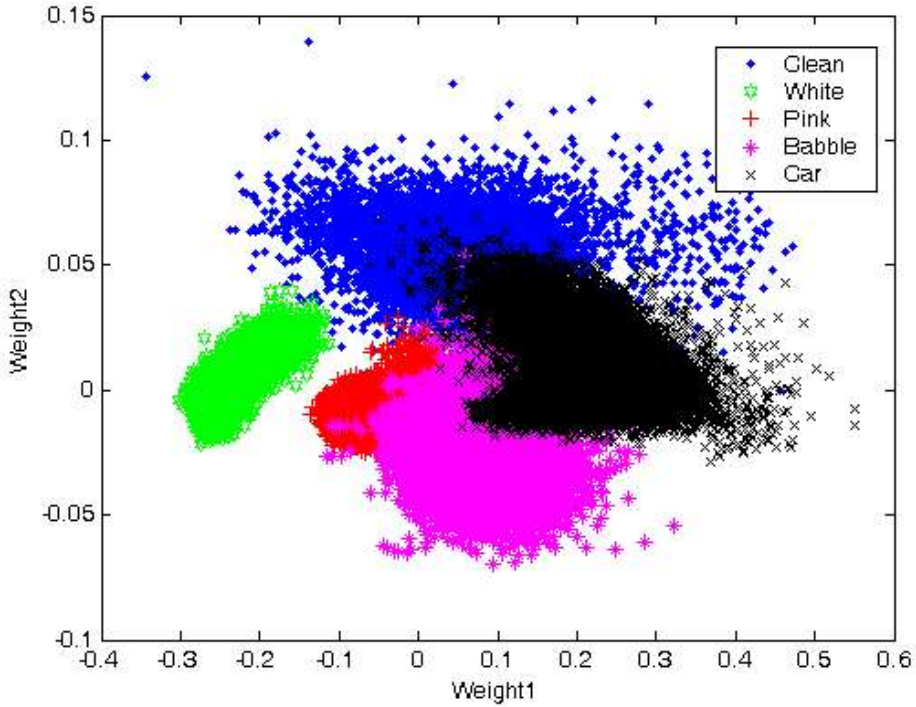
Classification using PCA

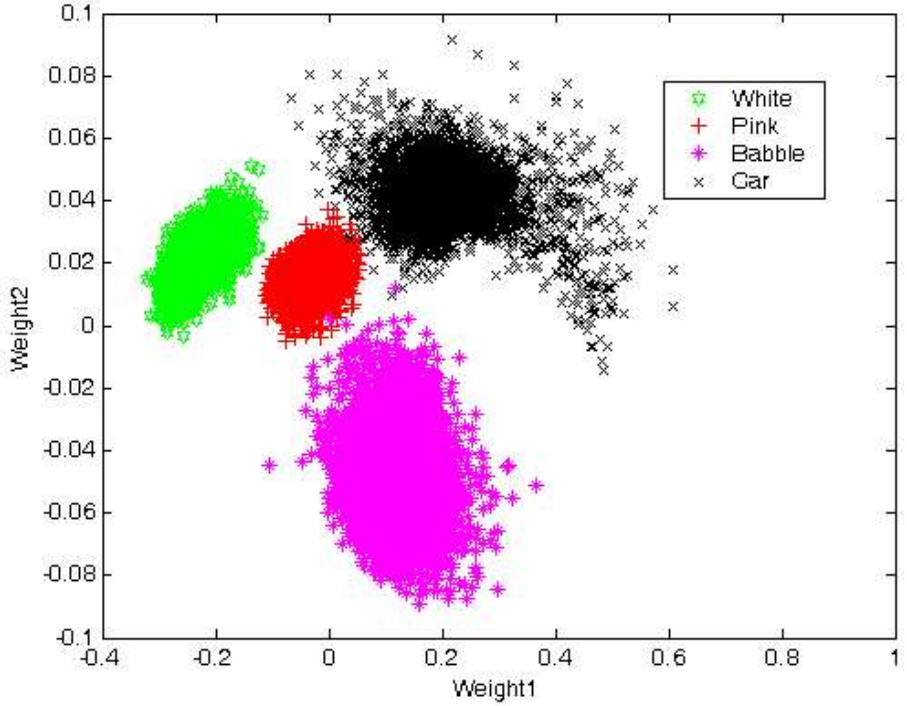
- Weight vector of principle components (Train -> 4 types)



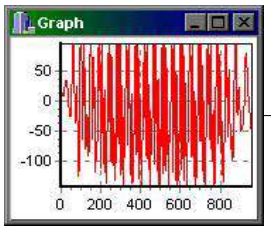
Classification using PCA

- Compare weights between train 5 types and train 4 types





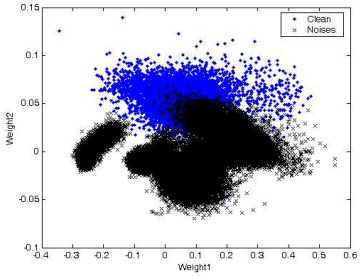
Classification using PCA



Speech

Feature Extraction

PCA1

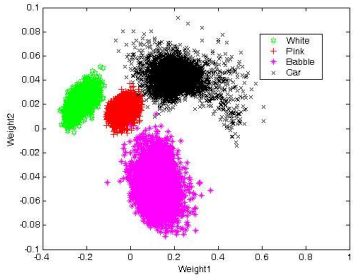


Classification1

clean

noise

PCA2



Classification2

white

pink

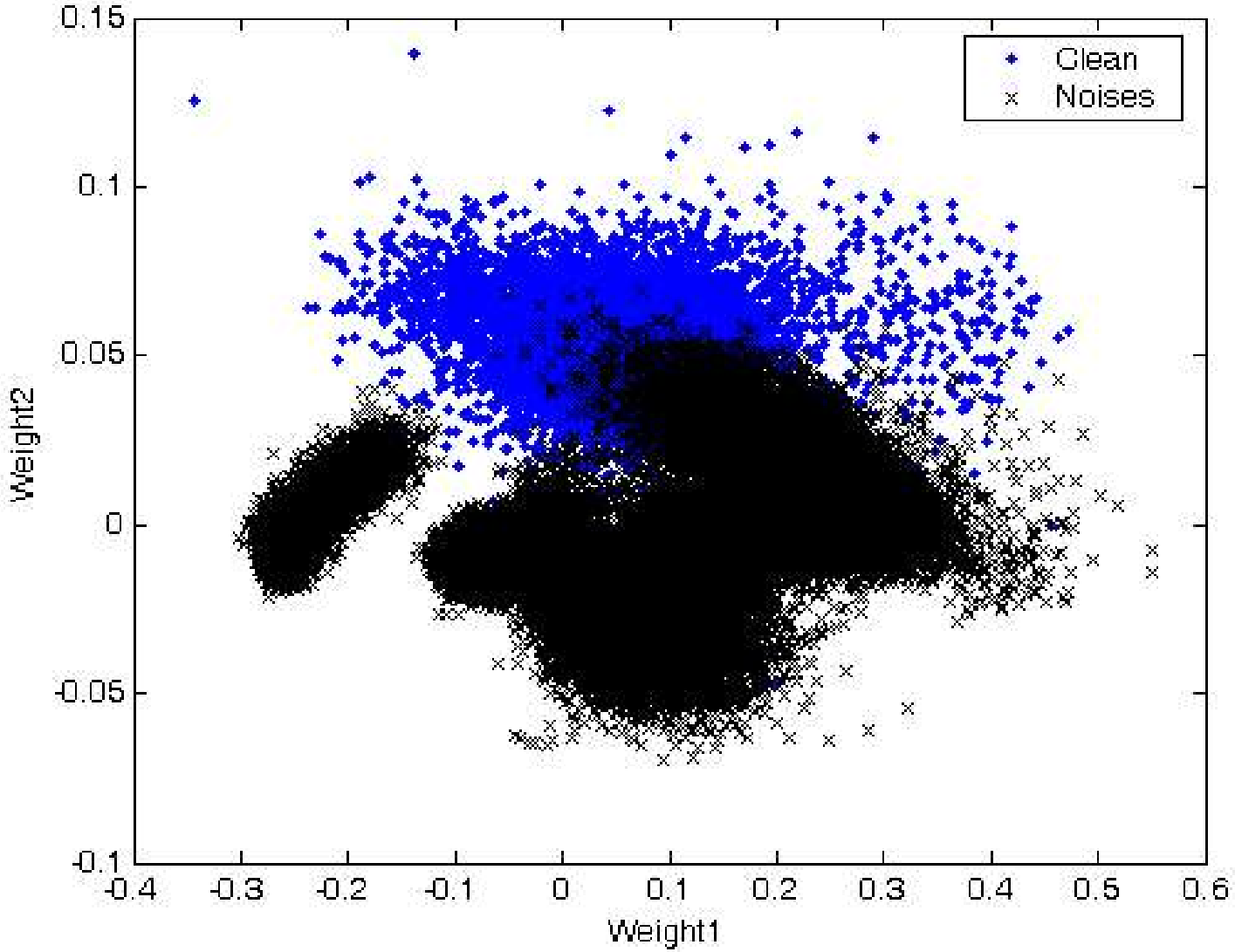
babble

car

Structure of Propose Classification using PCA

Classification using PCA

- Weight vector of principle components (Train -> 2 types)



Experiment

- Speech Data
  - Clean environment (from NECTEC-ATR thai corpus)
  - Noise environment (from NOISEX-92 database)
    - white
    - pink
    - babble
    - car

- Classified Techniques
  - MFCC/LPC/LPCEPSTRA+HMM
  - Norm-Log-Spectrum+PCA (1 step)+NN (nn_baseline)
  - Norm-Log-Spectrum+PCA (2 step)+NN (nn_propose)
  - Norm-Log-Spectrum+PCA (1 step)+SVM (svm_baseline)
  - Norm-Log-Spectrum+PCA (2 step)+SVM (svm_propose)

Experiment

- Experiment data

Process	Type	Num of File
Training set for PCA	Clean	1,000
	Original Noise	4,000
	Total	5,000
Training set for classification	Clean	12,306
	Original Noise	306
	Addition Noise	48,918
	Total	61,530
Testing set for classification	Clean	6,400
	Original Noise	102,400
	Total	108,800

Remark : Clean = Clean Speech

Original Noise = Noise Speech

Addition Noise = Clean + Noise Speech (SNR = 15 dB – 0 dB)

Experiment

- Noise Classification with HMM

Parameter	Error rate(%)
MFCC_E_D_A, GM=1	5.41
MFCC_E_D_A, GM=8	4.49
MFCC_E_D_A, GM=16	**4.18**
LPCEPSTRA_E_D_A, GM=1	5.80
LPCEPSTRA_E_D_A, GM=8	5.41
LPCEPSTRA_E_D_A, GM=16	5.29
LPC_E_D_A, GM=1	6.94
LPC_E_D_A, GM=8	7.01
LPC_E_D_A, GM=16	6.09

Remark : _E = Energy

_D = Delta coeffcients

_A = Acceleration coefficients

Experiment

- Noise Classification with NN

Num of hidden node	Type	Error Rate (%)
30	Baseline	4.55
	Propose	**1.20**
40	Baseline	4.37
	Propose	1.44
50	**Baseline**	**4.33**
	Propose	1.24
100	Baseline	4.60
	Propose	1.42

Remark : Baseline = Classical Classification using PCA (1 step)

Propose = Propose Classification using PCA (2 step)

Experiment

- Noise Classification with SVM (RBF kernel)

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{a^2}\right)$$

Parameter	Type	Error Rate (%)
a = 0.01	Baseline	3.99
	Propose	1.68
a = 0.1	Baseline	3.97
	Propose	1.71
a = 0.5	Baseline	3.86
	Propose	**1.57**
a = 1	**Baseline**	**3.78**
	Propose	1.64

Remark : Baseline = Classical Classification using PCA (1 step)

Propose = Propose Classification using PCA (2 step)

Experiment

- Noise Classification with SVM (Polynimial kernel)

$$k(x, y) = (x \cdot y)^p$$

Parameter	Type	Error Rate (%)
p = 1	**Baseline**	**4.01**
	Propose	**1.90**
p = 2	Baseline	19.43
	Propose	6.62
p = 3	Baseline	38.55
	Propose	27.69

Remark : Baseline = Classical Classification using PCA (1 step)

Propose = Propose Classification using PCA (2 step)

Experiment

- Noise Classification with SVM (Sigmoid kernel)

$$k(x, y) = \tanh(kx \cdot y)$$

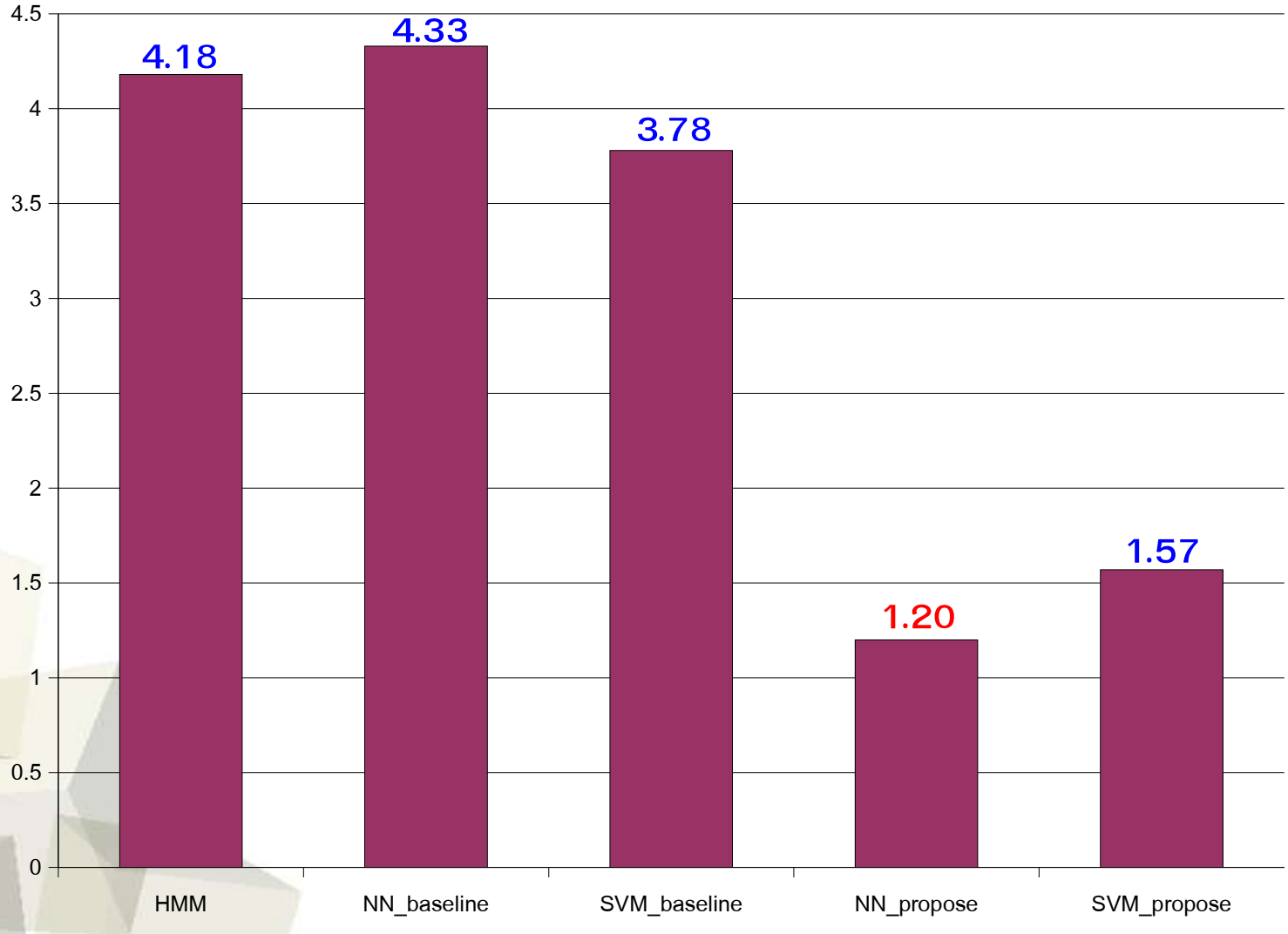
Parameter	Type	Error Rate (%)
k = 0.25	Baseline	4.02
	Propose	1.92
k = 0.5	Baseline	4.05
	Propose	1.91
k = 1	Baseline	4.35
	Propose	**1.70**
k = 2	**Baseline**	**3.98**
	Propose	2.22

Remark : Baseline = Classical Classification using PCA (1 step)

Propose = Propose Classification using PCA (2 step)

Experiment

- Summary of Classification Result



Conclusion

- This paper present a new method for noise classification used by projection to each data of the principle components.

- The new classification structure that seperates clean speech from noisy speech and then classify the type of noise gives a lower error rate comparing to the classical structure that classifies in one step

Acknowledgments

The authors would like to thank Dr.Sanparith Marukatat for his comments and suggestion.

Question & Answer