Constructing Modulation Frequency Domain-Based Features for Robust Speech Recognition

Reporter:黃重翔
Professor:陳嘉平

Abstract

Abstract—Data-driven temporal filtering approaches based on a specific optimization technique have been shown to be capable of enhancing the discrimination and robustness of speech features in speech recognition. The filters in these approaches are often obtained with the statistics of the features in the temporal domain. In this paper, we derive new data-driven temporal filters that employ the statistics of the modulation spectra of the speech features. Three new temporal filtering approaches are proposed and based on constrained versions of linear discriminant analysis (LDA), principal component analysis (PCA), and minimum class distance (MCD), respectively. It is shown that these proposed temporal filters can effectively improve the speech recognition accuracy in various noise-corrupted environments. In experiments conducted on Test Set A of the Aurora-2 noisy digits database, these new temporal filters, together with cepstral mean and variance normalization (CMVN), provide average relative error reduction rates of over 40% and 27% when compared with baseline Mel frequency cepstral coefficient (MFCC) processing and CMVN alone, respectively.

Index Terms—Modulation frequency, noise-robust features, speech recognition.

Outline

- Introduction
- Temporal Filter Design In The Modulation Frequency Domain
- Constrained Optimization Problem
- Experimental Results and Discussion

Introduction

- 語音辨識系統的效能常常會由於在訓練與應用的環境中聲音條件的配置不當而下降。
- 和傳統時域濾波技術如RASTA（相對頻譜）和CMS（倒頻譜平均消去法）這些濾波器形式經過修整且對提供的語音特徵參數無關的技術比較，資料導向時域濾波器可以被調頻以便於適應語音特徵參數或是環境參數，而且常使用具體優化技術。

Introduction

- 根據資料導向時域濾波法，本論文建議時域濾波器應根據語音特徵參數的調變頻譜而產生。
- 爲了驗證這種方法，我們產生一系列根據兩種資料庫而設計的實驗，以取得時域濾波器以及它們在辨識任務上所對應的效能。

Temporal Filter Design In The Modulation Frequency Domain

- 一個M維的有序序列特徵向量{x(n), n=1, 2, 3, …, N}，n表時間索引。
- 設 $x_m(n) = x(n, m)$，n=1,2, …,N，m=1,2,…, M m表特徵參數索引。

$$
\begin{array}{ccccccccc}
\begin{bmatrix} x(1,1) \\ x(1,2) \\ \vdots \\ x(1,m) \\ \vdots \\ x(1,M) \end{bmatrix}
\begin{bmatrix} x(2,1) \\ x(2,2) \\ \vdots \\ x(2,m) \\ \vdots \\ x(2,M) \end{bmatrix}
\begin{bmatrix} x(3,1) \\ x(3,2) \\ \vdots \\ x(3,m) \\ \vdots \\ x(3,M) \end{bmatrix}
& \cdots &
\begin{bmatrix} x(n,1) \\ x(n,2) \\ \vdots \\ x(n,m) \\ \vdots \\ x(n,M) \end{bmatrix}
& \cdots &
\begin{bmatrix} x(N,1) \\ x(N,2) \\ \vdots \\ x(N,m) \\ \vdots \\ x(N,M) \end{bmatrix}
&
\begin{matrix} \rightarrow \{x_1(n)\} \\ \rightarrow \{x_2(n)\} \\ \vdots \\ \rightarrow \{x_m(n)\} \\ \vdots \\ \rightarrow \{x_M(n)\} \end{matrix}
\end{array}
$$

$$
\mathbf{x}(1) \quad \mathbf{x}(2) \quad \mathbf{x}(3) \quad \cdots \quad \mathbf{x}(n) \quad \cdots \quad \mathbf{x}(N)
$$

Fig. 1. Representation of the time trajectories of feature parameters.

Temporal Filter Design In The Modulation Frequency Domain

- 當一個長度爲L的FIR濾波器 $h_m(n)$，被應用於時間軌跡 $\{x_m(n)\}$ ，則輸出採樣 $\{y_m(n)\}$ 爲

$$y_m(n) = \sum_{u=0}^{L-1} h_u(u) x_m(n-u)$$

- 用一長度爲L的移動窗口來進行 $\{x_m(n)\}$ 程序以保持總長度爲L段的參數

$$\tilde{x}_m(n) = x_m(n-L+1) \ldots \; x_m(n-1) \; x_m(n)$$

- 將 $h_m(n)$ 以及每個 $\tilde{x}_m(n)$ 填入K-L個0（當K>=2L且K爲偶數），根據Parseval's定理改寫爲

$$y_m(n) = \frac{1}{K} \sum_{k=0}^{K-1} H_m(k) X_m^*(n,k) = \frac{1}{K} \sum_{k=0}^{K-1} H_m^*(k) X_m(n,k)$$

Temporal Filter Design In The Modulation Frequency Domain

- 將$h_m(n)$以及每個$\tilde{x}_m(n)$填入K-L個0（當K>=2L且K爲偶數），根據Parseval's定理改寫爲

$$y_m(n)=\frac{1}{K}\sum_{k=0}^{K-1}H_m(k)X_m^*(n,k)=\frac{1}{K}\sum_{k=0}^{K-1}H_m^*(k)X_m(n,k)$$

此時$H_m(k)$與$X_m(n,k)$分別爲$h_m(n)$與$\tilde{x}_m(n)$的k點DFT。

- 在時間n時時域濾波器輸出的瞬時能量爲

$$|y(n)|^2=\left|\frac{1}{K}\sum_{k=0}^{K-1}H_m(k)X_m^*(n,k)\right|^2$$

可以改寫爲

$$|y(n)|^2=\frac{2(k+2)}{K^2}\sum_{k=0}^{K/2}|H_m(k)|^2|X_m(n,k)|^2=\frac{2(k+2)}{K^2}H^TX(n)$$

Temporal Filter Design In The Modulation Frequency Domain

- 證明 *Let* $\{c_i, 1 \le i \le L\}$ *be a set of real number. Then*

$$\left(\sum_{i=1}^{L} c_i\right)^2 = \sum_{i=1}^{L} c_i \sum_{j=1}^{L} c_j = \sum_{i=1}^{L}\sum_{j=1}^{L} c_i c_j$$

$$\le \frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}\left(c_i^2 + c_j^2\right) \quad \left(\because \left(c_i - c_j\right)^2 = c_i^2 + c_j^2 - 2c_i c_j \ge 0\right)$$

$$= \frac{1}{2}\sum_{i=1}^{L}\left(Lc_i^2 + \sum_{j=1}^{L} c_j^2\right) = \frac{1}{2}\left(L\sum_{i=1}^{L} c_i^2 + L\sum_{j=1}^{L} c_j^2\right) = L\sum_{i=1}^{L} c_i^2$$

$$Therefore, \left(\sum_{i=1}^{L} c_i\right)^2 \le L\sum_{i=1}^{L} c_i^2$$

both $\{H(k)\}$ and $\{X(n,k)\}$ are conjugate symmertic with respect to k=K/2. That is

$H(k) = H^*(K\text{-}k)$,

$X(n,k/2) = X^*(n,K-k)$ *for* $1 \le k \le k/2-1$

and $H(0), H(k/2)$, $X(n,0)$, and $X(n,k/2)$ are all real numbers.

Temporal Filter Design In The Modulation Frequency Domain

Therefore, the instataneous energy of the temporal filter output is

$$|y(n)|^2 = \left|\frac{1}{K}\sum_{k=0}^{K-1} H(k)X^*(n,k)\right|^2$$

$$= \frac{1}{K^2}\left|H(0)X(n,0) + H(K/2)X(n,K/2) + \sum_{k=1}^{K/2-1}\left(H(k)X^*(n,k) + H^*(k)X(n,k)\right)\right|^2$$

$$\leq \frac{1}{K^2}\left||H(0)||X(n,0)| + |H(K/2)||X(n,K/2)| + 2\sum_{k=1}^{K/2-1}|H(k)||X(n,k)|\right|^2$$

$$\leq \frac{1}{K^2}\left|2\sum_{k=1}^{K/2}|H(k)||X(n,k)|\right|^2$$

$$\leq \frac{4}{K^2}\left(\frac{K}{2}+1\right)\sum_{k=1}^{K/2}|H(k)|^2|X(n,k)|^2$$

$$= \frac{2(K+2)}{K^2}\sum_{k=1}^{K/2}|H(k)|^2|X(n,k)|^2$$

Therefore

$$|y(n)|^2 \leq \frac{2(K+2)}{K^2}\sum_{k=1}^{K/2}|H(k)|^2|X(n,k)|^2$$

Temporal Filter Design In The Modulation Frequency Domain

- 若定義濾波器輸出的瞬時調制頻譜能量爲

$$\varepsilon_Y(n)=\sum_{k=0}^{K/2}\left|H_m(k)\right|^2\left|X_m(n,k)\right|^2=H^TX(n)$$

  則可改寫前式爲

$$\left|y(n)\right|^2\leq\frac{2(K+2)}{K^2}\varepsilon_Y(n)$$

  大致上可以用 $\varepsilon_Y(n)$ 特徵化 $\left|y(n)\right|^2$ 的行爲。
- 現在最佳化向量H被用於最大化一具體的目標函數 $\varepsilon_Y(n)$，這和統計X有關。

Temporal Filter Design In The Modulation Frequency Domain

- 本論文提出三種最佳化技術：C-LDA，C-PCA與C-MCD。
- 當大致取得理想的H值後，並因此找出時域濾波器的振幅同調響應 $\{h[n], 0 \le n \le L-1\}$，可以用一些濾波器設計演算法取得相近的對應脈衝響應。
- 一個完美的對稱FIR濾波器爲擁有線性相位的回饋，導致通過濾波器的延遲在任何頻率下都相同，且濾波器不會產生相位與延遲失真。

Constrained Optimization Problem

- 由於每個H的元素必須限定爲非負實數，這產生一限定最佳化問題

$$H^{*} = \arg\max_{H} J(H), \text{ subject to } \mathrm{H} \geq 0$$

- 爲了處理H的元素使其非負，我們使用一中間變數向量

$$\bar{H} = \left[\bar{H}_0 \;\; \bar{H}_1 \;\; \cdots \;\; \bar{H}_{K/2} \right]$$

$$H_k = \left(\frac{\exp\left(\bar{H}_k\right)}{\sum_{m=0}^{K/2} \exp\left(\bar{H}_m\right)} \right)^{\frac{1}{p}}, \quad k = 0, 1, 2, \ldots K/2 \qquad\qquad \sum_{m=0}^{K/2} \bar{H}_k^{\,p} = 1$$

Constrained Optimization Problem

- 根據中間向量 $\bar{H}$ 找出最佳的H，使J(H)爲最大
- 使用梯度法更新 $\bar{H}$

$$\bar{H}^{(\theta+1)} = \bar{H}^{(\theta)} + \varepsilon \left.\frac{\partial J}{\partial \bar{H}}\right|_{\bar{H}=\bar{H}^{(\theta)}}$$

- ε 爲step size，且 $\frac{\partial J}{\partial \bar{H}} = \frac{\partial H}{\partial \bar{H}} \frac{\partial J}{\partial H}$ ，第i，j項爲

$$\left(\frac{\partial H}{\partial \bar{H}}\right)_{ij} = \frac{1}{p}\left(\frac{\exp\left(\bar{H}_j\right)}{\sum_{m=0}^{K/2}\exp\left(\bar{H}_m\right)}\right)^{\frac{1}{p}-1} \times \left(\frac{\exp\left(\bar{H}_j\right)\delta_{ij}\sum_{m=0}^{K/2}\exp\left(\bar{H}_m\right) - \exp\left(\bar{H}_i + \bar{H}_j\right)}{\left(\sum_{m=0}^{K/2}\exp\left(\bar{H}_m\right)\right)^2}\right), 0 \le i, j \le K/2$$

- 而 $\partial J / \partial H$ 可由所選的目標函數J(H)決定。

Constrained Optimization Problem

- LDA被廣泛應用於段落辨識，其目的爲找出資料中最有差異性的代表。
- 用來導出H的演算法稱爲限定性LDA(C-LDA)
- 每個視窗內的段落$\tilde{\mathbf{x}}(n)$的平方頻譜振幅X(n)首先被標爲J的類別之一，或是語音模型。
- 標記的程序可以經由預先經過訓練的模型其時序校準的平均來產生。

Constrained Optimization Problem

- 對屬於各個類別j進行標記後的X(n)之平均值$\boldsymbol{\mu}^{(j)}$與共變異數矩陣$\Sigma^{(j)}$可以表示爲

$$\mu^{(j)}=\frac{1}{N_j}\sum_{n=1}^{N_j}X^{(j)}(n) \ and \ \Sigma^{(j)}=\frac{1}{N_j}\sum_{n=1}^{N_j}\left(X^{(j)}(n)-\mu^{(j)}\right)\left(X^{(j)}(n)-\mu^{(j)}\right)^T$$

  其中$X^{(j)}(n)$表示被屬於類別j下標記的X(n)，而N_j爲$X^{(j)}(n)$的總個數。

- X中的類別間與類別內矩陣可分別表示爲

$$S_B=\sum_{j=1}^{J}N_j\left(\mu^{(j)}-\mu\right)\left(\mu^{(j)}-\mu\right)^T \ and \ S_w=\sum_{j=1}^{J}N_j\Sigma^{(j)} \ where \ \mu=\left(1/\Sigma_{j=1}^{J}N_j\right)\Sigma_{j=1}^{J}N_j\mu^{(j)}$$

Constrained Optimization Problem

- 設σ_B^2 與σ_W^2 爲ε_Y的變異數，則C-LDA的目標函數爲

$$J_{LDA} = \frac{\sigma_B^2}{\sigma_W^2} = \frac{H^T S_B H}{H^T S_W H},\ subject\ to\ H \geq 0$$

而可得梯度$\partial J / \partial \mathbf{H}$ 爲

$$\frac{\partial J_{LDA}}{\partial H} = \frac{2\left(H^T S_W H\right) S_B H - 2\left(H^T S_B H\right) S_W H}{\left(H^T S_W H\right)^2}$$

Constrained Optimization Problem

- C-PCA與C-LDA最大不同在X(n)不需要根據不同類別進行標記，而當作單一隨機向量X的取樣。
- X的平均與共變異數可表示爲

$$\mu = \frac{1}{N}\sum_{n=1}^{N} X(n) \ and \ \Sigma = \frac{1}{N}\sum_{n=1}^{N}\left(X(n)-\mu\right)\left(X(n)-\mu\right)^{T}$$

- ε_Y 的全域變異數爲 σ^2，則目標函數爲

$$J_{PCA} = \sigma^2 = H^T \Sigma H, \ subject \ to \ H \geq 0$$

梯度爲 $\frac{\partial J_{LDA}}{\partial H} = 2\Sigma H$

Constrained Optimization Problem

- C-MCD和C-LDA相似，根據j個類別或語音模型對X(n)進行標記。
- 假設$X^{(j)}$與平均$\boldsymbol{\mu}^{(j)}$和共變異數矩陣$\Sigma^{(j)}$爲多變元高斯分布，針對第j個類別的濾波器輸出$\mathcal{E}_Y$，求得頻譜能量$\mathcal{E}_Y^{(j)}$的機率密度函數$g^{(j)}(x)$爲

$$
\begin{aligned}
g^{(j)}(x) &= \mathcal{N}\left(x;\mathbf{H}^T\boldsymbol{\mu}^{(j)},\mathbf{H}^T\boldsymbol{\Sigma}^{(j)}\mathbf{H}\right) \\
&= \frac{1}{\sqrt{(2\pi)\mathbf{H}^T\boldsymbol{\Sigma}^{(j)}\mathbf{H}}}\exp\left[-\frac{\left(x-\mathbf{H}^T\boldsymbol{\mu}^{(j)}\right)^2}{2\mathbf{H}^T\boldsymbol{\Sigma}^{(j)}\mathbf{H}}\right]
\end{aligned}
$$

Constrained Optimization Problem

- 而不同的類別i與j在 ε_Y 的間距爲

$$d_{ij} \triangleq \int_{-\infty}^{\infty} g^{(i)}(x)\log\frac{g^{(i)}(x)}{g^{(j)}(x)}dx = \log\frac{H^T\Sigma^{(j)}H}{H^T\Sigma^{(i)}H} + \frac{H^T\left(\mu^{(i)}-\mu^{(j)}\right)\left(\mu^{(i)}-\mu^{(j)}\right)H}{H^T\Sigma^{(j)}H} + \frac{H^T\Sigma^{(i)}H}{H^T\Sigma^{(j)}H} - 1$$

根據兩高斯統計分布的KL分歧度可得。

- 而目標函數可用所有類別間距總和取得最大值

$$J_{MCD}(H) = \sum_i \sum_{i\neq j} d_{ij} \triangleq \sum_i \sum_{i\neq j}\left[\log\frac{H^T\Sigma^{(j)}H}{H^T\Sigma^{(i)}H} + \frac{H^T\left(\mu^{(i)}-\mu^{(j)}\right)\left(\mu^{(i)}-\mu^{(j)}\right)H}{H^T\Sigma^{(j)}H} + \frac{H^T\Sigma^{(i)}H}{H^T\Sigma^{(j)}H} - 1\right],\ \textit{subject to } H \geq 0 \quad A^{(i,j)} = \left(\mu^{(i)}-\mu^{(j)}\right)\left(\mu^{(i)}-\mu^{(j)}\right)^T$$

梯度則爲

$$\frac{\partial J_{MCD}(H)}{\partial H} = \sum_i \sum_{i\neq j} d_{ij} \triangleq \sum_i \sum_{i\neq j}\left[2\left(\frac{H^T\Sigma^{(j)}H}{H^T\Sigma^{(i)}H}\right)\times\frac{\left(H^T\Sigma^{(i)}H\right)\Sigma^{(j)}H-\left(H^T\Sigma^{(j)}H\right)\Sigma^{(i)}H}{\left(H^T\Sigma^{(i)}H\right)^2} + \frac{2\left(H^T\Sigma^{(j)}H\right)A^{(i,j)}H-2\left(H^TA^{(i,j)}H\right)\Sigma^{(j)}H}{\left(H^T\Sigma^{(j)}H\right)^2} + \frac{2\left(H^T\Sigma^{(j)}H\right)\Sigma^{(i)}H-2\left(H^T\Sigma^{(i)}H\right)\Sigma^{(j)}H}{\left(H^T\Sigma^{(j)}H\right)^2}\right]$$

Experimental Results and Discussion

- 使用AURORA Ver.2.0 專案資料庫
  - 來自乾淨的訓練集的每段音聲首先被轉換爲一13維MFCC的序列
  - 濾波器長度L，DFT長度K，指數P設爲101，256與4
  - 根據C-LDA，C-PCA與C-MCD分別產生時序濾波器
  - 加入delta與delta-delta特徵，則最後會使用共39維特徵

Experimental Results and Discussion

- 技術詳細內容
  - RASTA
  - 線性相位RASTA：一對稱FIR濾波器用來近似RASTA
  - 空間對時間LDA：一超向量用串聯九個相鄰的24維(216)對數頻譜特徵參數來建構。這些超向量被投影矩陣（LDA產生）轉換來構成39維特徵參數
  - 時間LDA：直接根據時域中特徵參數的特性來產生
  - 線性相位時間LDA：近似線性相位RASTA

Test	System	clean	20dB	15dB	10dB	5dB	0dB	-5dB	average (0~20dB)	Relative WER reduction
Test Set A	plain MFCC	98.91	94.99	86.93	67.28	39.36	17.07	8.40	61.13	
	RASTA	98.70	95.91	89.50	66.96	34.63	20.06	11.94	61.41	0.72
	LP RASTA	98.94	96.95	92.47	75.81	43.84	23.18	13.00	66.45	13.69
	ST_LDA	98.52	88.65	70.59	43.50	20.77	9.84	7.15	46.67	-37.20
	ST_LDA+D+A	98.78	95.45	88.05	70.05	43.07	18.00	7.09	62.92	4.61
	T_LDA	98.63	94.49	83.67	60.43	30.60	11.11	6.56	56.06	-13.04
	LP T_LDA	98.79	94.25	86.28	71.80	45.11	22.53	12.06	63.99	7.36
	MF_C-LDA	98.80	96.67	92.94	81.34	57.17	26.24	8.92	70.87	25.06
	MF_C-PCA	98.66	95.54	89.16	73.15	46.80	22.32	9.77	65.39	10.96
	MF_C-MCD	98.18	95.89	91.59	78.97	53.74	26.01	12.81	69.24	20.86

Test	System	clean	20dB	15dB	10dB	5dB	0dB	-5dB	average (0~20dB)	Relative WER reduction
Test Set B	plain MFCC	98.91	92.35	80.79	58.06	32.04	14.63	7.92	55.57	
	RASTA	98.70	96.74	91.93	74.90	44.30	23.66	13.02	66.31	24.17
	LP RASTA	98.94	97.38	93.80	81.52	53.04	27.99	14.28	70.74	34.14
	ST_LDA	98.52	83.90	64.89	37.92	16.19	6.66	5.63	41.91	-30.74
	ST_LDA+D+A	98.78	94.95	86.92	67.58	42.99	19.87	6.45	62.46	15.51
	T_LDA	98.63	93.72	87.06	66.18	36.54	17.85	9.70	60.27	10.58
	LP T_LDA	98.79	93.25	87.05	75.85	51.66	27.34	14.19	67.03	25.79
	MF_C-LDA	98.80	96.20	91.74	80.28	55.81	23.57	5.08	69.52	31.40
	MF_C-PCA	98.66	92.30	83.71	65.43	38.92	16.90	7.72	59.45	8.73
	MF_C-MCD	98.18	95.32	91.24	80.70	57.89	29.96	12.46	71.02	34.77

Experimental Results and Discussion

Test	System	clean	20dB	15dB	10dB	5dB	0dB	-5dB	average (0~20dB)	Relative WER reduction
Test Set C	plain MFCC	99.00	94.83	88.66	75.23	50.85	23.83	11.4	66.68	
	RASTA	98.69	95.40	87.70	63.31	34.94	21.12	12.72	60.49	-18.58
	LP RASTA	99.09	96.44	90.82	70.44	40.61	23.03	14.57	64.27	-7.23
	ST_LDA	98.07	86.06	72.19	52.51	31.65	17.20	9.55	51.92	-44.30
	ST_LDA+D+A	98.52	95.62	89.72	72.79	49.18	25.73	12.97	66.61	-0.21
	T_LDA	98.71	89.57	78.18	58.28	36.56	17.28	9.52	55.97	-32.14
	LP T_LDA	98.65	90.89	82.71	67.14	41.44	20.20	11.64	60.47	-18.64
	MF_C-LDA	98.87	96.00	91.70	80.66	59.34	32.28	15.52	71.99	15.94
	MF_C-PCA	98.77	94.41	88.22	74.13	52.50	26.86	12.97	67.22	1.62
	MF_C-MCD	98.36	94.85	88.88	73.48	46.13	24.81	16.25	65.63	-3.15

Experimental Results and Discussion

- 我們試著加入三個建議使用的時間濾波器與其他一些加強性的技術。
  - CMVN：

$$y_{m,CMVN}(n) = \frac{\left[x_m(n) - \mu_m\right]}{\sigma^2}$$

  - CGN：

$$y_{m,CGN}(n) = \frac{\left[x_m(n) - \mu_m\right]}{\max\left[x_m(n)\right] - \min\left[x_m(n)\right]}$$

  - AFE

WORD RECOGNITION ACCURACIES (%) AND RELATIVE WER REDUCTION (%) AS COMPARED TO THE MFCC BASELINE FOR VARIOUS APPROACHES AT DIFFERENT SNR VALUES BUT AVERAGED OVER ALL THE NOISE TYPES IN TEST SET A OF THE AURORA-2 DATABASE

Test	System	clean	20dB	15dB	10dB	5dB	0dB	-5dB	average (0~20dB)	Relative WER reduction
Test Set A	Plain MFCC	98.91	94.99	86.93	67.28	39.36	17.07	8.40	61.13	
	CMVN	98.98	95.98	91.66	80.48	57.40	26.40	10.96	70.38	23.80
	CMVN+ C-LDA	98.85	97.02	94.15	87.34	71.81	41.97	16.13	78.46	44.58
	CMVN+ C-PCA	98.51	96.48	93.42	86.47	72.87	48.76	22.68	79.60	47.52
	CMVN+ C-MCD	98.28	96.15	93.11	86.30	72.25	47.81	21.85	79.12	46.28
	CGN	98.91	96.48	93.16	85.29	69.30	40.73	15.46	76.99	40.80
	CGN+ C-LDA	98.79	96.92	94.31	88.56	76.65	52.20	22.37	81.73	53.00
	CGN+ C-PCA	98.60	96.29	93.55	87.97	76.84	54.07	23.95	81.74	53.02
	CGN+ C-MCD	98.51	96.23	93.49	87.39	74.26	48.83	20.56	80.04	48.65
	AFE	99.10	98.14	96.75	92.99	84.06	60.72	28.86	86.53	65.35
	AFE+ C-LDA	98.85	97.77	96.52	93.17	85.35	63.65	29.14	87.29	67.30
	AFE+ C-PCA	98.95	97.94	96.54	92.94	83.63	60.91	28.88	86.39	64.99
	AFE+ C-MCD	98.85	97.79	96.58	93.30	85.40	63.95	29.75	87.40	67.58

WORD RECOGNITION ACCURACIES (%) AND RELATIVE WER REDUCTION (%) AS COMPARED TO THE MFCC BASELINE FOR VARIOUS APPROACHES AT DIFFERENT SNR VALUES BUT AVERAGED OVER ALL THE NOISE TYPES IN TEST SET B OF THE AURORA-2 DATABASE

Test	System	clean	20dB	15dB	10dB	5dB	0dB	-5dB	average (0~20dB)	Relative WER reduction
Test Set B	plain MFCC	98.91	92.35	80.79	58.06	32.04	14.63	7.92	55.57	
	CMVN	98.98	96.41	92.15	81.78	58.69	26.47	10.98	71.10	34.95
	CMVN+ C-LDA	98.85	97.23	94.75	88.41	72.18	42.33	15.28	78.98	52.69
	CMVN+ C-PCA	98.70	96.92	94.44	88.49	74.08	48.62	20.87	80.51	56.13
	CMVN+ C-MCD	98.12	95.99	93.10	86.62	73.11	50.1	22.85	79.78	54.49
	CGN	98.91	96.86	94.09	87.01	70.20	39.95	15.09	77.62	49.63
	CGN+ C-LDA	98.79	97.06	95.07	89.79	77.39	51.36	20.33	82.13	59.78
	CGN+ C-PCA	98.60	96.41	94.17	89.10	77.77	54.87	22.94	82.46	60.52
	CGN+ C-MCD	98.51	96.59	94.30	88.53	76.08	50.84	20.33	81.27	57.84
	AFE	99.10	98.01	96.29	92.26	81.27	57.59	26.32	85.09	66.44
	AFE+ C-LDA	98.85	97.22	95.43	91.54	82.12	60.16	28.40	85.29	66.89
	AFE+ C-PCA	98.95	97.62	95.70	91.20	80.10	56.76	25.43	84.27	64.60
	AFE+ C-MCD	98.95	97.32	95.47	91.73	82.13	60.41	28.19	85.41	67.16

WORD RECOGNITION ACCURACIES (%) AND RELATIVE WER REDUCTION (%) AS COMPARED TO THE MFCC BASELINE FOR VARIOUS APPROACHES AT DIFFERENT SNR VALUES BUT AVERAGED OVER THE TWO NOISE TYPES IN TEST SET C OF THE AURORA-2 DATABASE

Test	System	clean	20dB	15dB	10dB	5dB	0dB	-5dB	average (0~20dB)	Relative WER reduction
Test Set C	plain MFCC	99.00	94.83	88.66	75.23	50.85	23.83	11.4	66.68	
	CMVN	99.12	95.51	88.71	74.21	51.25	24.30	10.49	66.80	0.36
	CMVN+C-LDA	98.83	96.80	93.30	84.63	66.66	37.82	14.90	75.84	27.49
	CMVN+C-PCA	98.66	96.10	92.43	84.67	69.79	48.04	23.04	78.20	34.57
	CMVN+C-MCD	98.25	95.62	92.37	84.29	70.02	48.02	23.81	78.06	34.15
	CGN	98.96	96.13	92.03	83.18	64.89	35.29	13.81	74.30	22.87
	CGN+C-LDA	98.91	96.94	93.95	87.80	74.11	47.11	19.55	79.98	39.92
	CGN+C-PCA	98.68	96.16	93.65	87.61	75.53	51.55	22.59	80.90	42.68
	CGN+C-MCD	98.63	96.40	93.11	86.44	72.29	46.81	19.68	79.01	37.00
	AFE	99.01	97.53	95.69	90.33	78.99	53.12	26.42	83.13	49.37
	AFE+C-LDA	98.86	97.52	95.74	91.67	81.29	56.47	25.44	84.54	53.57
	AFE+C-PCA	98.92	97.23	95.08	90.29	79.09	54.87	26.60	83.31	49.91
	AFE+C-MCD	98.87	97.53	95.83	91.81	82.07	58.10	26.63	85.07	55.19