

第三章 文獻回顧

3.1 對數能量特徵值之強建式技術

當語音訊號受到環境噪音干擾時，在語音資料上的現象可由對數能量參數在時間軸上的變化明顯觀察出，尤其是在非語音訊號的段落，原本的語音對數能量參數值應該偏低，但受噪音的干擾後對數能量參數值卻因此而增加。然而參考文獻中[Bocchieri and Wilpon 1992]學者也特別研究能量特徵值對於自動語音辨識的影響。本章節所回顧的對數能量特徵值的技術，目前從作法上大致可分為兩類：第一類作法是只對訓練語料的對數能量特徵值做預先處理，方法為音框能量消去(Frame Energy Subtraction, FES)和對數能量動態範圍正規化法(Log-Energy Dynamic Range Normalization, LEDRN)。第二類作法則是同時對訓練語料和測試語料用同方法做處理，方法為靜音音框對數能量正規化法(Silence Log-Energy Normalization, SLEN)。以下小節將詳細介紹各能量特徵值強建式技術。

3.1.1 音框能量消去法(Frame Energy Subtraction, FES)

音框能量消去法[Gomez et al. 2004]主要假設噪音語音為語音訊號與噪音訊號加成的結果，與頻譜消去法(Spectral Subtraction, SS)的假設與作法相類似，希望藉由預先估測噪音的能量值，用此噪音能量值來降低噪音對語音訊號在能量上的影響，進而減少存在訓練語料和測試語料之間環境不匹配的現象。基本上必須假設：語音訊號與噪音訊號二者在統計上是不相關(Uncorrelated)，因此噪音訊號對語音訊號的影響只是加成性(Additive)的改變。因此若希望能由觀察到的雜訊語音(Noisy Speech)重建出乾淨語音(Clean Speech)訊號，只需要將含有噪音的語音能量值扣掉噪音能量值，特別是針對測試語料每一語句的任何音框來做處理。根據上述假設與處理方法，我們可以令 $E_y[i]$ 、 $E_x[i]$ 與 $E_n[i]$ 分別為音框 i 的噪音(Noise Speech)語音能量、乾淨語音(Clean Speech)能量以及噪音(Noise)本身的能

量，而三者間有式(3.1.1)的關係：

$$E_y[i] \cong E_x[i] + E_n[i] \quad (3.1.1)$$

然而， $E_n[i]$ 事先並未能知道，必需透過特定方式估測而得。我們可進一步假設每一語句的前 K 音框為噪音音框，以取得估計值 \hat{E}_n 來取代 $E_n[i]$ ，如式(3.1.2)所示：

$$\hat{E}_n = \frac{1}{K} \sum_{i=1}^K E_y[i] \quad (3.1.2)$$

在得到 $\log \hat{E}_n$ 估計值後，採用噪音遮蔽法以防止 $E_n[i] - \hat{E}_n$ 出現負值來取得近似的乾淨語音訊號能量，如下(3.1.3)：

$$E_x[i] = \begin{cases} E_y[i] - \alpha \cdot \hat{E}_n & , E_y[i] > \hat{E}_n & , 0 \leq \alpha \leq 1 \\ \beta \cdot E_y[i] & , E_y[i] \leq \hat{E}_n & , 0 \leq \beta \leq 1 \end{cases} \quad (3.1.3)$$

其中 α 為過度估測因子， β 為底限因子。

3.1.2 對數能量動態範圍正規化法(Log-Energy Dynamic Range Normalization, LEDRN)

對數能量動態範圍正規化法[Weizhong and Douglas 2005]的目標是針對於乾淨的語音訊號做預處理(Preprocessing)，讓乾淨語音訊號的對數能量可以逼近噪音語音訊號的對數能量，使得經由對數能量動態範圍正規化法(LEDRN)處理過的訓練語料所練出的聲學模型會與測試的噪音語音訊號兩者間互相匹配。

對數能量動態範圍正規化法(LEDRN)的考慮是語音訊號能量曲線波峰(Peaks)段落，不容易受到噪音影響，而在波谷(Valley)段落則會受到噪音嚴重干擾，以至於乾淨語音與噪音干擾之語音在能量特徵上有不匹配的現象。因此可以利用一個線性處理或一個非線性的處理方法，使得能量特徵波峰值維持不變，而波谷的值相對上升以達到訓練與測試的訊號能有匹配的效果。

具體作法是先對乾淨環境訓練模式的每一則訓練語句中所有 T 個音框之對數能量中找出最大對數能量值 LE_max 以及最小對數能量值 LE_min ：

$$\begin{aligned} LE_max &= \underset{1 \leq i \leq T}{Max} LogE[i] \\ LE_min &= \underset{1 \leq i \leq T}{Min} LogE[i] \end{aligned} \quad (3.1.4)$$

當得到最大和最小對數能量值後，根據我們事前預期測試語音的噪音干擾大小決定一個動態能量的範圍比值 $D.R.$ ，在此可以依照不同的噪音環境定義出測試語音檔要調整的最小音框能量值：

$$D.R. = 10 \times \frac{LE_max}{T_min} \quad (3.1.5)$$

緊接著針對每句訓練語句做偵測，當所有音框中的最小對數能量值小於最小目標對數能量值 T_min 的時候，則對語句中每個音框的對數能量作更新，更新的方法主要分為線性(Linear)調整與非線性(Non-Linear)調整，如式(3.1.6)和式(3.1.7)：

$$\begin{aligned} & \text{if } LE_min < T_min \\ & \text{then} \end{aligned} \quad (3.1.6)$$

$$Log\hat{E}[i] = LogE[i] + \frac{T_min - LE_min}{LE_max - LE_min} \times (LE_max - LogE[i])$$

$$\begin{aligned} & \text{if } LE_min < T_min \\ & \text{then} \end{aligned}$$

$$\begin{aligned} Log\hat{E}[i] &= LogE[i] + \frac{T_min - LE_min}{\log(LE_max) - \log(LE_min)} \\ & \times (\log(LE_max) - \log(LogE[i])) \end{aligned} \quad (3.1.7)$$

其中 $Log\hat{E}[i]$ 為更新後之對數能量。由式(3.1.6)和(3.1.7)可以發現對於語音能量較大的音框在透過對數能量動態範圍正規化法處理後只被作些許地作修正，而能量較小的音框則被作大幅地作修正，如圖 3.1.1 為所顯示，圖中橫軸為原對數能量，縱軸為調整後的對數能量對應值，藍色實線為原對數能量排序後曲線，長虛線為非線性方式調整，點虛線為線性調整後的對數能量。

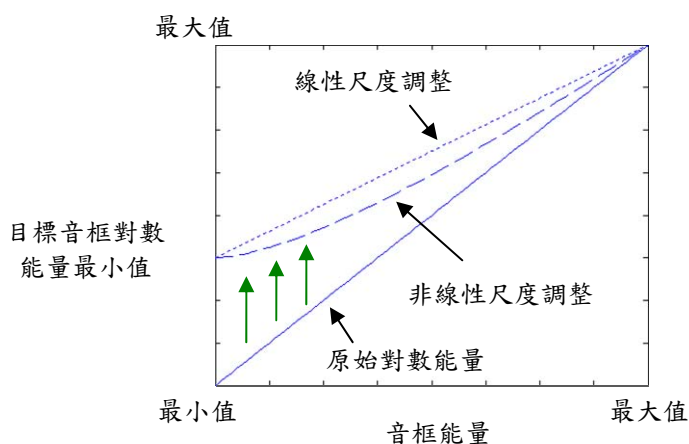


圖 3.1.1 對數能量動態範圍正規化法作用前後示意圖

3.1.3 靜音音框對數能量正規化法 I (Silence Log-Energy Normalization, SLEN I)

靜音音框對數能量正規化法 I [戴仲甫 2006] 是一個快速而有效的方法，主要利用語音端點偵測 (Voice Activity Detector, VAD) 的方法找出非語音區間作進階處理。靜音音框對數能量正規化法的原理類似能量正規法，由於噪音語音受干擾較為嚴重的部份為波谷，而波峰則比較不受影響。實驗顯示若僅保留波峰部分，經辨識後依然可以得到不錯的正確率。從實驗觀察得知，對能量特徵而言最重要的是語音能量曲線而不是非語音能量失真的部分，也就是說當有語音的段落若具有完整的能量曲線，會比音框能量值整體的降低或升高所造成的失真程度還重要。因此假設，具有完整的語音能量曲線，就可以得到好的辨識率。依照上述的假設，我們找出波型中非語音的部份，並且把它正規化為一個常數值，由於非語音的部份已經正規化為一個常數值，而語音部分又比較不受噪音的干擾，因此正規化處理過後的語料，不論在乾淨環境下語音段落的對數能量曲線或噪音環境下的語音對數能量的曲線會相似。

具體作法是利用噪音偵測法來找出非語音音框，以便輔助靜音音框對數能量正規化法 I。靜音音框對數能量正規化法 I 主要利用能量來判斷語音與非語音的門檻值，並將非語音音框正規化為一個常數。如下式(3.1.8)與(3.1.9)：

$$\tau = \frac{1}{T} \sum_{i=1}^T \text{Log}E[i] \quad (3.1.8)$$

$$\text{Log}\hat{E}[i]_{i=1\dots T} = \begin{cases} \text{Log}E[i] & , \text{if } \text{Log}E[i] \geq \tau \\ \Phi & , \text{otherwise} \end{cases} \quad (3.1.9)$$

其中 T 為一段語音的音框數， $\text{Log}E[i]$ 與 $\text{Log}\hat{E}[i]$ 分別為音框對數能量以及修正後之音框對數能量， τ 則是門檻值而 Φ 為一個常數。

3.1.4 靜音音框對數能量正規化法 II (Silence Log-Energy

Normalization, SLEN II)

靜音音框對數能量正規化法 II [Tai and Hung 2006]，同靜音音框對數能量正規化法 I 的假設，但是判斷語音與非語音的門檻值方法捨棄靜音音框對數能量正規化法 I 的作法改用變動音框位移率 (Variable Frame Rate, VFR) 取代。變動音框位移選擇噪音音框方法如式 (3.1.10)：

$$y[i] = \frac{1}{2} (\text{Log}E[i+1] - y[i-1]) \quad (3.1.10)$$

其中 $y[i]$ 為 $\text{Log}E[i]$ 通過高通濾波器 (High-Pass Filter) 的輸出值； $\text{Log}E[i]$ 則為每個音框的對數能量。音框判斷法如式 (3.1.12)，若高通濾波器輸出值的 $y[i]$ 值小於門檻值 τ 則認定該音框為非語音音框，而 τ 的計算值與整個演算法如下式 (3.1.11) 與 (3.1.12)：

$$\tau = \frac{1}{T} \sum_{i=1}^T \log y[i] \quad (3.1.11)$$

$$\log \hat{E}[i]_{i=1\dots T} = \begin{cases} \log E[i] & , \text{if } \log y[i] \geq \tau \\ \Phi & , \text{otherwise} \end{cases} \quad (3.1.12)$$

最後如 3.1.3 節的假設，依照靜音音框對數能量正規化後，我們找出波型中非語音的部份，並且把它正規化為一個常數值，因此正規化處理過後，受噪音影響的語音能量的曲線會與乾淨的語音訊號能量波形將相似。

3.2 對數能量特徵強建技術實驗

根據 Aurora-2.0 的實驗設定如表 2.1.1，章節中的語音訊號特徵參數為梅爾倒頻譜參數。在此將介紹 3.1 節的語音強健技術實做於對數能量特徵值上。實驗數據採分別表示乾淨環境訓練模式與複合情境訓練模式的測試結果，最後由實驗結果顯示兩種環境下的訓練語料和測試語料於音框能量上的不匹配情況，皆可以在語音強健技術處理後能夠有效降低噪音的干擾，並在詞正確率上得到明顯的提升。實驗結果比較三組測試語料(測試組 A、測試組 B 和測試組 C)，不同的噪音設定包含有乾淨(Clean)環境結果與不同訊噪比(SNR) -5dB~20dB 干擾下的結果，表格中列平均為各組別在相同下訊噪比干擾下的平均值，其次表中最後一行的 0~20dB 平均為相同噪音於訊噪比 0~20dB 干擾下的平均值結果。

3.2.1 音框對數能量消去法(FES)

音框能量消去法假設噪音語音為語音能量與噪音能量加成的結果，式(3.1.3)中 α 過度估測因子設定為 0.95， β 底限因子為 0.05。實驗結果如表 3.2.1 與 3.2.2。

		乾淨環境訓練模式							
訊噪比		Clean	20dB	15dB	10dB	5dB	0dB	-5dB	0~20dB 平均
測試組 A	地下鐵	98.99	96.99	92.54	79.86	52.99	24.65	9.98	69.41
	人聲	99.03	97.04	91.32	76.54	49.12	19.50	5.74	66.70
	汽車	98.78	98.21	96.30	88.13	61.94	26.78	10.08	74.27
	展覽會館	99.01	96.76	92.75	80.19	59.09	31.38	12.59	72.03
	平均	98.95	97.25	93.23	81.18	55.79	25.58	9.60	70.60
測試組 B	餐廳	98.99	96.84	91.50	78.20	52.23	23.67	6.85	68.49
	街道	99.03	97.28	93.86	81.53	55.89	28.69	12.12	71.45
	機場	98.78	96.99	94.90	82.97	57.89	27.11	10.17	71.97
	火車站	99.01	97.75	94.66	85.99	59.80	26.26	9.60	72.89
	平均	98.95	97.22	93.73	82.17	56.45	26.43	9.69	71.20
測試組 C	地下鐵	99.23	91.93	81.36	61.59	35.52	15.57	9.43	57.19
	街道	98.94	94.44	86.76	69.89	46.58	25.39	15.15	64.61
	平均	99.09	93.19	84.06	65.74	41.05	20.48	12.29	60.90

表 3.2.1 音框能量消去法於乾淨環境訓練模式實驗結果

		複合情境訓練模式							
訊噪比		Clean	20dB	15dB	10dB	5dB	0dB	-5dB	0~20dB 平均
測試組 A	地下鐵	98.22	93.40	89.35	82.44	70.46	45.10	18.08	76.15
	人聲	98.34	96.16	92.26	86.03	75.57	50.70	18.95	80.14
	汽車	98.45	97.49	96.81	94.15	86.79	65.05	30.78	88.06
	展覽會館	98.40	96.14	94.79	90.28	80.65	54.92	23.94	83.36
	平均	98.35	95.80	93.30	88.23	78.37	53.94	22.94	81.93
測試組 B	餐廳	98.22	96.01	91.74	83.42	73.35	49.25	17.44	78.75
	街道	98.34	96.64	94.17	90.11	78.72	52.48	23.46	82.42
	機場	98.45	97.55	96.09	90.90	80.97	60.93	27.05	85.29
	火車站	98.40	96.98	94.91	91.18	79.91	59.52	28.23	84.50
	平均	98.35	96.80	94.23	88.90	78.24	55.55	24.05	82.74
測試組 C	地下鐵	98.28	94.29	90.57	82.81	64.42	33.25	13.05	73.07
	街道	98.37	96.22	93.86	86.06	68.65	41.11	19.17	77.18
	平均	98.33	95.26	92.22	84.44	66.54	37.18	16.11	75.12

表 3.2.2 音框能量消去法於複合情境訓練模式實驗結果

其中表 3.2.2 能量消去法於複合情境訓練模式需同時對訓練和測試語料作處理。

3.2.2 對數能量動態範圍正規化法(LEDNR)

對數能量動態範圍正規化針對乾淨環境訓練模式的語音對數能量做預處理，讓語料在乾淨環境下的對數能量模擬出該語料在噪音環境下的語音對數能量大小，因此在小節中的實驗以乾淨環境訓練模式為主。實驗中我們測試非線性動態範圍正規化在不同動態範圍值(D.R.)的正確率：如表 3.2.3，實驗結果發現當動態範圍值(D.R.)設定為 12dB 時，針對 Aurora-2.0 的所以測試語料會有最佳的正確率，詳細各項數據如表 3.2.4：

(D.R.)	測試組 A	測試組 B	測試組 C	總平均
10dB	68.73	69.57	57.02	66.72
11dB	73.92	76.39	62.98	72.72
12dB	74.18	76.45	62.52	72.75
13dB	71.88	74.72	61.36	70.91
14dB	70.23	73.30	61.04	69.62
15dB	68.19	71.14	60.31	67.80
16dB	68.29	71.24	60.88	67.99
17dB	66.84	69.56	60.46	66.65
18dB	66.30	68.76	60.35	66.09

表 3.2.3 動態範圍值於動態範圍非線性正規化平均實驗結果

訊噪比	Clean	20dB	15dB	10dB	5dB	0dB	-5dB	0~20dB 平均
測試組 A	地下鐵	87.04	96.41	92.48	80.99	58.46	28.55	71.38
	人聲	87.15	97.31	95.50	88.54	67.59	30.65	75.92
	汽車	86.01	97.41	95.71	87.29	66.39	34.75	76.31
	展覽會館	86.15	96.48	93.24	82.57	60.69	32.52	73.10
	平均	86.59	96.90	94.23	84.85	63.28	31.62	74.18
測試組 B	餐廳	87.04	96.38	94.14	86.21	66.41	30.73	74.77
	街道	87.15	97.34	94.56	86.15	65.18	36.79	76.00
	機場	86.01	97.38	96.06	91.50	74.23	41.07	80.05
	火車站	86.15	96.30	93.00	85.90	64.95	34.62	74.95
	平均	86.59	96.85	94.44	87.44	67.69	35.80	76.45
測試組 C	地下鐵	86.67	91.40	81.58	63.13	37.52	16.58	58.04
	街道	87.18	94.44	88.78	72.43	50.24	29.05	66.99
	平均	86.93	92.92	85.18	67.78	43.88	22.82	62.52

表 3.2.4 最佳動態範圍值(12dB)於動態範圍非線性正規化實驗結果

緊接著，對數能量動態範圍正規化實驗中我們測試線性動態範圍正規化在不同動態範圍值(*D.R.*)的正確率：如表 3.2.5，實驗結果發現當動態範圍值(*D.R.*)設定為 16dB 時，針對 Aurora-2.0 的所以測試語料會有最佳的正確率，詳細各項數據如表 3.2.6：

(<i>D.R.</i>)	測試組 A	測試組 B	測試組 C	總平均
10dB	15.43	14.67	13.68	14.78
11dB	41.36	39.47	32.46	38.83
12dB	61.66	61.06	47.25	58.54
13dB	69.29	70.55	54.17	66.77
14dB	72.10	74.11	57.07	69.90
15dB	73.11	75.47	58.65	71.16
16dB	73.08	75.83	59.85	71.53
17dB	72.56	75.32	60.14	71.18
18dB	71.48	74.49	60.07	70.40

表 3.2.5 動態範圍值於動態範圍線性正規化平均實驗結果

	訊噪比	Clean	20dB	15dB	10dB	5dB	0dB	-5dB	0~20dB 平均
測試組 A	地下鐵	97.24	96.13	93.09	83.21	61.59	29.26	11.48	72.66
	人聲	97.49	96.07	93.74	86.19	64.12	32.29	9.85	74.48
	汽車	96.90	95.94	93.14	83.51	61.80	32.78	11.12	73.43
	展覽會館	96.88	96.05	92.41	80.78	58.41	31.60	11.97	71.85
	平均	97.13	96.05	93.10	83.42	61.48	31.48	11.11	73.11
測試組 B	餐廳	97.24	96.10	94.14	85.97	64.78	33.34	11.64	74.87
	街道	97.49	96.49	92.35	83.98	63.39	37.00	15.93	74.64
	機場	96.90	96.06	94.33	88.55	70.59	40.20	14.41	77.95
	火車站	96.88	95.34	93.06	84.67	63.56	35.54	12.65	74.43
	平均	97.13	96.00	93.47	85.79	65.58	36.52	13.66	75.47
測試組 C	地下鐵	96.07	91.93	81.55	59.78	31.01	13.72	9.73	55.60
	街道	95.98	93.11	84.79	65.87	41.96	22.82	14.48	61.71
	平均	96.03	92.52	83.17	62.83	36.49	18.27	12.11	58.65

表 3.2.6 最佳動態範圍值(16dB)於動態範圍線性正規化實驗結果

3.2.3 靜音音框對數能量正規化法 I (SLEN I)

靜音音框對數能量正規化法 I 主要利用能量來判斷語音與非語音的門檻值，並將非語音音框正規化為一個常數 Φ ，在此設定 Φ 常數為 1。結果如表 3.2.7 與 3.2.8。

乾淨環境訓練模式									
	訊噪比	Clean	20dB	15dB	10dB	5dB	0dB	-5dB	0~20dB 平均
測試組 A	地下鐵	98.86	92.57	82.81	63.65	38.66	19.25	10.38	59.39
	人聲	98.61	96.77	92.90	82.01	61.25	31.47	11.76	72.88
	汽車	98.66	95.05	90.58	79.93	59.38	30.69	14.88	71.13
	展覽會館	98.95	92.93	84.20	68.44	43.47	20.46	11.69	61.90
	平均	98.77	94.33	87.62	73.51	50.69	25.47	12.18	66.32
測試組 B	餐廳	98.86	96.07	91.65	81.85	60.15	31.62	12.99	72.27
	街道	98.61	95.07	87.00	72.52	49.21	26.66	13.85	66.09
	機場	98.66	96.15	92.34	85.03	64.45	35.28	15.27	74.65
	火車站	98.95	95.71	90.81	79.88	58.99	31.29	13.21	71.34
	平均	98.77	95.75	90.45	79.82	58.20	31.21	13.83	71.09
測試組 C	地下鐵	98.83	84.77	68.28	47.31	26.71	12.99	7.77	48.01
	街道	98.64	89.12	79.96	58.65	34.95	20.47	10.73	56.63
	平均	98.74	86.95	74.12	52.98	30.83	16.73	9.25	52.32

表 3.2.7 靜音音框對數能量正規化法 I 於乾淨環境訓練模式實驗結果

		複合情境訓練模式							
訊噪比	Clean	20dB	15dB	10dB	5dB	0dB	-5dB	0~20dB 平均	
測試組 A	地下鐵	98.43	97.24	95.79	91.19	75.65	43.78	14.58	80.73
	人聲	98.52	97.70	96.86	93.68	83.25	55.20	18.56	85.34
	汽車	98.12	97.41	96.42	93.44	85.18	61.11	23.89	86.71
	展覽會館	98.52	97.25	95.74	90.93	79.33	49.71	17.40	82.59
	平均	98.40	97.40	96.20	92.31	80.85	52.45	18.61	83.84
測試組 B	餐廳	98.43	97.67	96.62	93.37	81.67	54.47	21.22	84.76
	街道	98.52	97.46	95.92	92.08	79.63	52.75	20.74	83.57
	機場	98.12	97.55	96.90	93.89	84.37	62.00	28.78	86.94
	火車站	98.52	97.78	96.14	91.51	81.46	59.36	27.21	85.25
	平均	98.40	97.62	96.40	92.71	81.78	57.15	24.49	85.13
測試組 C	地下鐵	98.65	95.73	92.54	84.19	61.01	25.24	10.04	71.74
	街道	98.19	96.19	94.62	87.70	70.31	39.63	17.93	77.69
	平均	98.42	95.96	93.58	85.95	65.66	32.44	13.99	74.72

表 3.2.8 靜音音框對數能量正規化法 I 於複合情境訓練模式實驗結果

3.2.4 靜音音框對數能量正規化法 II (SLEN II)

在靜音音框對數能量正規化法 II 中的設定與法 I 相同，將非語音音框正規化為一個常數 Φ ，並設定 Φ 常數為 1。實驗結果如表 3.2.9 與 3.2.10。

		乾淨環境訓練模式							
訊噪比	Clean	20dB	15dB	10dB	5dB	0dB	-5dB	0~20dB 平均	
測試組 A	地下鐵	98.99	94.96	89.07	71.97	44.00	19.77	10.59	63.95
	人聲	98.73	96.86	93.77	84.89	64.57	34.07	9.73	74.83
	汽車	98.90	95.35	92.10	82.82	64.84	35.73	13.81	74.17
	展覽會館	98.92	94.08	89.02	75.38	50.85	24.50	10.80	66.77
	平均	98.89	95.31	90.99	78.77	56.07	28.52	11.23	69.93
測試組 B	餐廳	98.99	96.90	93.40	85.69	64.97	36.69	12.34	75.53
	街道	98.73	96.22	91.23	78.81	57.22	31.26	12.94	70.95
	機場	98.90	96.63	93.68	87.21	68.77	41.16	15.06	77.49
	火車站	98.92	96.11	92.72	83.28	64.12	35.82	14.93	74.41
	平均	98.89	96.47	92.76	83.75	63.77	36.23	13.82	74.59
測試組 C	地下鐵	99.08	88.58	74.52	51.92	27.76	11.82	7.98	50.92
	街道	98.76	91.78	84.04	65.05	41.11	21.89	11.88	60.77
	平均	98.92	90.18	79.28	58.49	34.44	16.86	9.93	55.85

表 3.2.9 靜音音框對數能量正規化法 II 於乾淨環境訓練模式實驗結果

		複合情境訓練模式							
訊噪比		Clean	20dB	15dB	10dB	5dB	0dB	-5dB	0~20dB 平均
測試組 A	地下鐵	97.33	96.44	94.23	88.95	75.62	47.07	16.33	80.46
	人聲	97.37	97.01	95.74	91.38	79.53	50.48	16.20	82.83
	汽車	96.99	97.05	95.82	92.99	83.95	56.58	15.96	85.28
	展覽會館	97.44	96.14	94.88	90.65	80.13	54.34	21.01	83.23
	平均	97.28	96.66	95.17	90.99	79.81	52.12	17.38	82.95
測試組 B	餐廳	97.33	96.65	94.96	90.97	78.63	54.16	19.04	83.07
	街道	97.37	96.67	95.62	90.36	78.69	53.33	21.83	82.93
	機場	96.99	96.75	95.82	92.78	84.31	63.91	28.54	86.71
	火車站	97.44	96.95	95.16	90.71	80.47	59.98	25.15	84.65
	平均	97.28	96.76	95.39	91.21	80.53	57.85	23.64	84.34
測試組 C	地下鐵	97.21	94.93	91.83	84.03	64.51	30.15	10.53	73.09
	街道	97.37	95.74	93.92	86.64	71.37	42.17	18.95	77.97
	平均	97.29	95.34	92.88	85.34	67.94	36.16	14.74	75.53

表 3.2.10 靜音音框對數能量正規化法 II 於複合情境訓練模式實驗結果

3.2.5 結論

根據以上實驗結果，將三組噪音環境整理成表 3.2.11 與表 3.2.12，表格中測試組 A、測試組 B 和測試組 C 所表示的正確率為各組實驗之 0dB 至 20dB 的算數平均數結果，平均值(Average)則為測試組 A 四份加上測試組 B 四份和測試組 C 兩份的平均值結果，如下式：

$$\frac{\text{測試組 A} \times 4 + \text{測試組 B} \times 4 + \text{測試組 C} \times 2}{10} \quad (3.2.1)$$

其中第二章之基礎實驗(Baseline)結果將納入表格中一起比較，最後一行為進步率(relative improvement, R.I.)算法如下式 3.2.2。實驗結果如表 3.2.11。

$$R.I.(%) = \frac{\text{NewScore} - \text{Baseline}}{100 - \text{Baseline}} \times 100\% \quad (3.2.2)$$

乾淨環境訓練模式					
0~20dB 平均	測試組 A	測試組 B	測試組 C	平均	進步率
Baseline	58.94	58.48	59.97	58.96	non
FES	70.60	71.20	60.90	68.90	24.23
LEDRN Linear	73.11	75.47	58.65	71.16	29.73
LEDRN Non-Linear	74.18	76.45	62.52	72.75	33.61
SLEN I	66.32	71.34	52.32	65.53	16.00
SLEN II	69.93	74.59	55.85	68.98	24.41

表 3.2.11 對數能量之強建式技術比較於乾淨環境訓練模式實驗結果

複合情境訓練模式					
0~20dB 平均	測試組 A	測試組 B	測試組 C	平均	進步率
Baseline	85.22	83.99	80.67	83.82	non
FES	81.93	82.74	75.12	80.89	-18.10
SLEN I	83.94	85.13	74.72	82.57	-7.71
SLEN II	82.95	84.34	75.53	82.02	-11.11

表 3.2.12 對數能量之強建式技術比較於複合情境訓練模式實驗結果

實驗探討：

綜合對數能量特徵值之強建式技術的實驗結果，從表 3.2.11 中可以得知以上的方法與基礎實驗比較都有明顯的進步。但是從各組別來觀察，發現結果當中的測試組 C 的效果表現有好有壞，其中則以非線性的對數能量動態範圍正規化法 (LEDRN) 進步最多。因此結果顯示上述強建式的技術只能對測試組 A 和 B 兩種情境下的噪音干擾有提高辨識率的效果。然而，在複合情境訓練模式實驗結果卻不理想，綜觀兩種訓練模式的實驗結果，其中以非線性的對數能量動態範圍正規化法 (LEDRN) 的強建式技術最佳。而測試組 C 的 MIRS 通道加成性噪音干擾，明顯是所有方法都無法有效對抗。